Advances in Cognitive Neurodynamics

Hans Liljenström *Editor*

Advances in Cognitive Neurodynamics (IV)

Proceedings of the Fourth International Conference on Cognitive Neurodynamics - 2013

Springer

Advances in Cognitive Neurodynamics (IV)

More information about this series at http://www.springer.com/series/11163

Hans Liljenström
Editor

Advances in Cognitive Neurodynamics (IV)

Proceedings of the Fourth International Conference on Cognitive Neurodynamics – 2013

Springer

Editor
Hans Liljenström
Department of Energy and Analysis, SLU
Biometry and Systems Analysis
Uppsala, Sweden

Agora for Biosystems
Sigtuna, Sweden

ISSN 2213-3569 ISSN 2213-3577 (electronic)
Advances in Cognitive Neurodynamics (IV)
ISBN 978-94-017-9547-0 ISBN 978-94-017-9548-7 (eBook)
DOI 10.1007/978-94-017-9548-7

Library of Congress Control Number: 2014958900

Springer Dordrecht Heidelberg New York London

Printed on acid-free paper

Springer Science+Business Media B.V. Dordrecht is part of Springer Science+Business Media (www.springer.com)

Preface

Cognition in its essence is dynamic and multilayered, and the pursuit of new clues inevitably leads from one layer to the next, both bottom-up and top-down. Similarly, the nervous system can be described at different organizational levels, e.g. sub-cellular, cellular, network and the level of the entire brain, and each level is characterized by its dynamical states and processes. Knowledge and understanding of the great complexity of neural systems has increased tremendously in the last few decades. Experimental methods, such as patch clamp technique, EEG, MEG, PET, fMRI etc., have provided a huge amount of data. At the same time, statistical, mathematical, and computational analysis and modeling have contributed to an understanding of the intricate relations between structure, dynamics, and function of neural systems at different scales. Both holistic and reductionist approaches have proven essential for a more comprehensive description.

The young and rapidly growing research field of cognitive neurodynamics has evolved from the inter-play between experimental and theoretical/computational neuroscience and cognitive science. Inevitably, this field is highly interdisciplinary, where scientists from many different fields, such as neuroscience, cognitive science, psychology, psychiatry, medicine, mathematics, physics, and computer science contribute to the advancement of the field. In particular, methods to describe, analyze, and model nonlinear dynamics, including spiking, bursting, oscillating, and chaotic behavior, often in combination with stochastic processes, need to be further developed and applied. Also, advanced techniques to implement the models in artificial systems, computers, and robots are called for.

In order to promote the integration of cognitive science and neurodynamics as a whole, the International Conference on Cognitive Neurodynamics (ICCN) is held biennially since 2007 with support from the international journal *Cognitive Neurodynamics* (Springer). The first two conferences in the series were held in China (Shanghai and Hangzhou, respectively) and the third conference was held in Japan (Hokkaido).

The 4th conference, ICCN2013, on which these proceedings are based, was for the first time organized outside Asia, in Sigtuna, Sweden, on 23–27 June 2013,

right after the Swedish Midsummer. The conference was held in the inspiring and creative atmosphere of the Sigtuna Foundation, offering a stimulating forum for scientists, scholars, and engineers to review the latest progress in the field of cognitive neurodynamics, and to exchange experiences and ideas. The Sigtuna Foundation, with its unique architecture on a hill near the shore of lake Mälaren, provided an excellent setting for the talks and intense discussions, often extending late into the bright midsummer nights.

ICCN2013 attracted 153 participants from 20 different countries, who made this conference a successful and memorable event. There were four keynote talks by leading scientists in the field of cognitive neurodynamics: Prof. Walter Freeman, Prof. Riitta Hari, Prof. Fabio Babiloni, and Prof. Yoko Yamaguchi. In addition, eight plenary talks were given by Prof. Steven Bressler, Prof. Barry Richmond, Prof. Yanchao Bi, Prof. Scott Kelso, Prof. John Hertz, Prof. James Wright, Prof. Paul Rapp, and Prof. Aike Guo. In total 120 papers were presented in oral or poster sessions. The topics ranged from macro- and meso- to microscopic levels, from social and interactive neurodynamics, all the way down to neuronal processes at quantum levels.

This volume fairly well reflects the large span of research presented at ICCN2013. The papers of this volume are grouped in ten parts that are organized essentially in a top-down structure. The first parts deal with social/interactive (I) and mental (II) aspects of brain functions and their relation to perception and cognition (III). Next, more specific aspects of sensory systems (IV) and neural network dynamics of brain functions (V), including the effects of oscillations, synchronization, and synaptic plasticity (VI), are addressed, followed by papers particularly emphasizing the use of neural computation and information processing (VII). With the next two parts, the levels of cellular and intracellular processes (VIII) and finally quantum effects (IX) are reached. The last part (X), with the largest number of papers of mixed topics, is devoted to the contributions invited by the Dynamic Brain Forum (DBF), which was co-organized with ICCN2013.

We wish to express our gratitude to all those who made ICCN2013 and this Proceedings possible. In addition to all the contributing authors, we owe thanks to the special session organizers, Drs. Katarzyna Blinowska, Erik Fransén, Walter Freeman, Pavel Herman, Wlodzimierz Klonowski, Jan Lauwereyns, Sisir Roy, Emmanuelle Tognoli, Felix Tretter, Hong-bo Yu, and Tao Zhang. We gratefully acknowledge support from The Swedish Research Council, the Grant-In-Aid G4103 MEXT Japan, the Sigtuna Foundation, Agora for Biosystems, Uppsala University, the Swedish University of Agricultural Sciences, as well as from Springer. We are also grateful to all the helpful students who assisted during the conference, and to the friendly staff at the Sigtuna Foundation, including those in the kitchen, who provided us with excellent food throughout the conference.

The 5th conference in the series, ICCN2015, will be held in Sanya on the Chinese island of Hainan, 3–7 June 2015, organized by Prof. Rubin Wang and colleagues at the East China University of Science and Technology. We have no doubt that ICCN2015 will be as successful as the previous ones, bringing together

an interdisciplinary group of scientists in an inspiring atmosphere for the discussion of new exciting knowledge and surprises in the exploration of the human brain and mind. This is the guiding idea that will make each ICCN meeting a most adventurous event for many years to come.

Hans Liljenström
Hans A. Braun

Organizers

Honorary Chairs

Shun-ichi Amari, RIKEN Brain Science Institute, Japan
Jean-Pierre Changeux, Institut Pasteur, France
Walter J. Freeman, University of California, Berkeley, USA
Fanji Gu, Shanghai Society for Biophysics/Fudan University, P.R. China
Hermann Haken, University of Stuttgart, Germany
John J. Hopfield, Princeton University, USA
Barry J. Richmond, NIMH/NIH/DHHS, USA
Minoru Tsukada, Tamagawa University, Japan

General Chair

Hans Liljenström, SLU and Agora for Biosystems, Sweden

Co-Chairs

Hans A. Braun, Philipps University of Marburg, Germany
Rubin Wang, East China University of Science and Technology, P.R. China
Yoko Yamaguchi, RIKEN Brain Science Institute (BSI), Japan

Local Organizing Committee

Hans Liljenström, SLU/Agora, Uppsala/Sigtuna
Christian Finke, Carl von Ossietzky University, Germany
Alf Linderman, Sigtuna Foundation, Sigtuna
Dietrich von Rosen, SLU, Uppsala
Björn Wahlund, Karolinska Institutet, Stockholm
Peter Århem, Karolinska Institutet, Stockholm

Program Committee

Gaute Einevoll, UMB, Norway
Jeanette Hellgren, Royal Institute of Technology, Sweden
John Hertz, Nordita, Sweden
Guang Li, Zhejiang University, P.R. China
Svetlana Postnova, University of Sydney, Australia
Paul Rapp, Drexel University, College of Medicine, USA
Ichiro Tsuda, Hokkaido University, Japan
Björn Wahlund, Karolinska Institute, Sweden
Alessandro Villa, Université de Lausanne, Switzerland
Peter Århem, Karolinska Institute, Sweden

Advisory Committee

Kazuyuki Aihara (The University of Tokyo, Japan)
Fabio Babilloni (University of Rome, Italy)
Erol Basar (Istanbul Kultur University, Turkey)
Roman Borysyuk, (Russian Academy of Science/University of Plymouth, UK)
Wlodzislaw Duch (Nicolaus Coperinicus University, Poland)
Peter Erdi (Kalamazoo College, USA)
Erik Fransén (Royal Institute of Technology, Sweden)
Riitta Hari (Aalto University, Finland)
Leslie Kay (University of Chicago, USA)
Scott Kelso (Florida Atlantic University, USA)
Peter Koenig (University of Osnabrück, Germany)
Robert Kozma (University of Memphis, USA)
Anders Lansner (Royal Institute of Technology, Sweden)
Jan Lauwereyns (University of Wellington, New Zealand)
Soo-Young Lee (Korea Advanced Institute of Science and Technology, Republic of Korea)
Cees van Leuven (RIKEN Brain Science Institute, Japan)
Pei-ji Liang (Shanghai Jiao Tong University, P.R. China)
Shigetoshi Nara (Okayama University, Japan)
Alessandro Treves (International School for Advanced Studies (SISSA), Italy)
Joanna Tyrcha (Stockholm University, Sweden)
Francesco Ventriglia (Institute of Cybernetics of CNR, Italy)
Deliang Wang (Ohio State University, USA)
James Wright (Auckland University, New Zealand)

Secretariat

Maria-Pia Anderloni, Sigtunastiftelsen, Sigtuna, Sweden
Sofia af Geijerstam, Sigtunastiftelsen, Sigtuna, Sweden
Malin Lindqvist, Sigtunastiftelsen, Sigtuna, Sweden
Birgitta Lagerholm, Sigtunastiftelsen, Sigtuna, Sweden

Cover Design

Regina Clevehorn, Sigtunastiftelsen, Sigtuna, Sweden

Home Pages

Sigtunastiftelsen, Sigtuna, Sweden, www.sigtunastiftelsen.se/ICCN2013
Agora for Biosystems, www.agoraforbiosystems.se

Sponsors

Agora for Biosystems

Sigtunastiftelsen, the Sigtuna Foundation

Vetenskapsrådet, the Swedish Research Council


Grant-in-Aid G4103 MEXT Japan

Cognitive Neurodynamics, Springer

Uppsala Universitet

SLU, the Swedish University of Agricultural Sciences

Contents

Part I
Interactive and Social Neurodynamics

Brain Network Efficiency and Intelligent Scores of Children

Fang Duan, Hiroyuki Tsubomi, Yuko Yoshimura, Mitsuru Kikuchi, Yoshio Minabe, Kastumi Watanabe, and Kazuyuki Aihara

Abstract Graph theory is recently becoming a popular method to study brain functional networks and evaluate efficiency of brain function. However, only a few studies have focused on children. One of the main reasons is that children's data are usually contaminated by artifacts. We propose to construct brain graphs after applying independent component analysis (ICA) to remove artifacts in the magnetoencephalography (MEG) data of children. The clustering coefficient and the harmonic average path length are calculated from artifact-free MEG data. Analysis showed that certain values of the clustering coefficient and harmonic average path length denoted more intelligent performance, as assessed by the Kaufman Assessment Battery for Children, which may indicate better network efficiency.

Keywords Graph theory • Independent component analysis • MEG • Intelligence • Network efficiency

1 Introduction

Magnetoencephalography (MEG) is a noninvasive technique to measure the magnetic fields of the brain. The application of MEG data has succeeded in studying brain networks [1]. Graph theory has been employed to understand brain-network structure in several studies, and it has been found that brain networks hold a high-efficiency small world structure [2]. Because children are not yet fully developed,

F. Duan (✉) • K. Aihara
Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan
e-mail: duan@sat.t.u-tokyo.ac.jp

H. Tsubomi
Faculty of Humanities, University of Toyama, Toyama 930-8555, Japan

Y. Yoshimura • M. Kikuchi • Y. Minabe
Research Center for Child Mental Development, Graduate School of Medical Science, Kanazawa University, Kanazawa 920-8641, Japan

K. Watanabe
RCAST, The University of Tokyo, Tokyo 153-8904, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_1

their brains should show a more significant plasticity. Therefore, studies of the brain function of children can provide great assistance in understanding the plasticity of the brain, and its involvement in the rehabilitation of the brain after cerebropathy or many other issues. The limited self-restraint of children makes it difficult for them to fully complete the experimental tasks. This reduces the reliability of experimental data. Therefore, improving the data quality of children is also essential. In this paper, we will show an efficient way to remove the noise in the MEG data of children by ICA. Furthermore, the relationship between the network efficiency of children and intelligence scores, acquired by the K-ABC, was evaluated.

2 Methodology

The MEG recording sessions required the participants to lie as still as possible on a bed with their head inside the helmet of the MEG system. In order to help the participants stay calm, they were allowed to watch a video during the experiment. MEG was recorded with a whole-head coaxial gradiometer system (PQ 1151R; Yokogawa/KIT) for children, with 151 channel superconducting quantum interface devices. The recording session lasted for 3 min with a sampling rate of 1000 Hz. MEGs of 58 children (10 girls, 48 boys) at an age of 39–93 months were recorded. The collection of data was approved by the Ethics Committee of Kanazawa University Hospital.

The Kaufman Assessment Battery for Children (K-ABC) [3] was applied for all children to measure their intelligence. In total, three kinds of general scores, including the sequential processing, simultaneous processing, and achievement scales were measured before the experiment. To find the relationship between network efficiency and intelligence scores, we compared the clustering coefficient and harmonic average path length between high low scoring children, 25 children in each group, based on each scale.

In this paper, the MEG data were resampled to 250 Hz. We employed ICA to discriminate the noise produced by muscle movement and cardio artifact from the data. In this work, kurtosis was utilized as the index to classify the ICA components. In brain signal processing most ICA components show positive kurtosis. The artifacts, especially, always have extremely high kurtosis indexes.

We use synchronization likelihood (SL) to estimate the association matrix. SL is a non-linear measurement scheme using the method of time delay embedding [4]. For two time series $x(t)$ and $y(t)$, SL can be defined as follow:

$$SL = \frac{2}{(N-w)\,(N-w+1)\,P_{ref}} \times \sum_{i=1}^{N-w} \sum_{j=i+w}^{N} \theta\left(r_x - |\mathbf{x}(i) - \mathbf{x}(j)|\right) \theta\left(r_y - |\mathbf{y}(i) - \mathbf{y}(j)|\right), \tag{1}$$

where θ (x) is the Heaviside function. Variable w is the dead band to exclude temporal neighbors. We set $P_{ref} = 0.05$, and $w = 140$. For the time delay embedding, we set $\tau = 10$, $d = 14$. The association matrix is computed by using 5,000 sample points.

In this paper, we construct the brain graph using 43 channels covering the frontal lobe and parietal lobe, which play an important role in attention, memory, language, and other cognitive function. The value of the threshold was set under the criterion that ensures all brain graphs hold the same number of edges and can thus be easily compared between groups. In this way, the two groups of graphs hold the same number of vertexes and edges [2].

We consider brain networks as undirected and unweighted graphs. The clustering coefficient and the characteristic path length are two popular indexes for undirected unweighted graphs in small world network theory [5]. We use harmonic average path length as one kind of revised path length index in this study. The equations of clustering coefficient and harmonic average path length are given as follows:

$$C = \frac{2}{N}\sum_{i}\frac{\sum_{j=1}^{N-1}\sum_{m=j+1}^{N} a_{ij}a_{jm}a_{mi}}{k_i\,(k_i-1)}, \tag{2}$$

$$L = \frac{N\,(N-1)}{2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{1}{d_{ij}}}, \tag{3}$$

where a_{ij} is equal to 1 when vertexes i and j are connected, k_i is the number of edges incident with i, d_{ij} is the shortest path length between vertexes i and j, and d_{ij} is infinity when i and j are isolated from each other.

3 Results and Discussion

A comparison of the graph indexes between the high scoring group (HG) and low scoring group (LG) was performed. Figures 1, 2, and 3 show the means and standard deviations of the clustering coefficient and harmonic average path length as a function of the degree of graph. In the figures, the blue and red lines represent the HG and low LG, respectively. Triangles and asterisks indicate that the differences between groups were statistically significant ($p < 0.05$ and $p < 0.01$ in t-test, respectively).

Figures 1, 2, and 3 show that the differences of the harmonic average path length between HG and LG were statistically significant. The harmonic average path length of brain network in HG was lower than LG. We suggest that the brain networks of children in HG possess shorter information transfer paths than LG. Our results were compatible with the result in the reference [6].

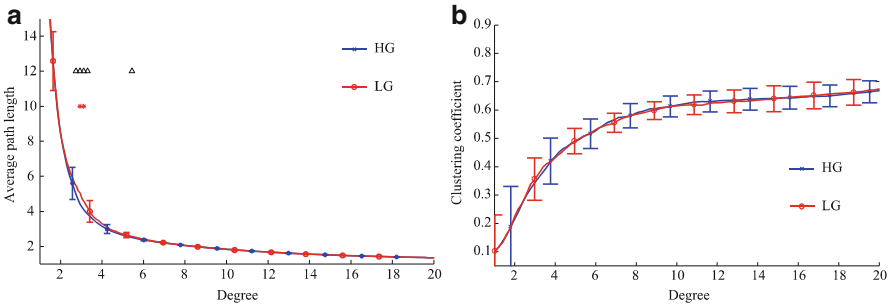


Fig. 1 The harmonic average path length (**a**) and the clustering coefficient (**b**) as a function of the degree of the networks. The groups were divided based on the sequential processing scale

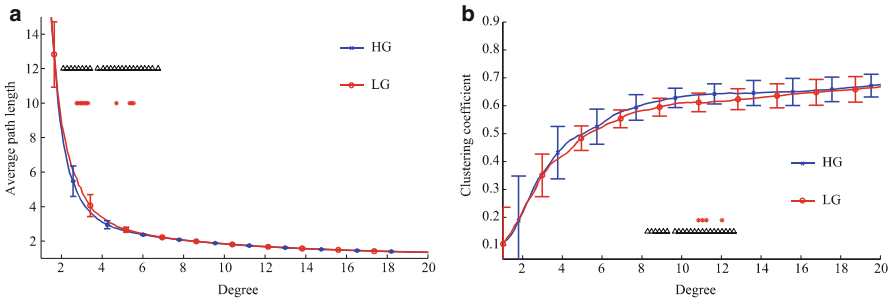


Fig. 2 The harmonic average path length (**a**) and the clustering coefficient (**b**) as a function of the degree of the networks. The groups were divided based on the simultaneous processing scale

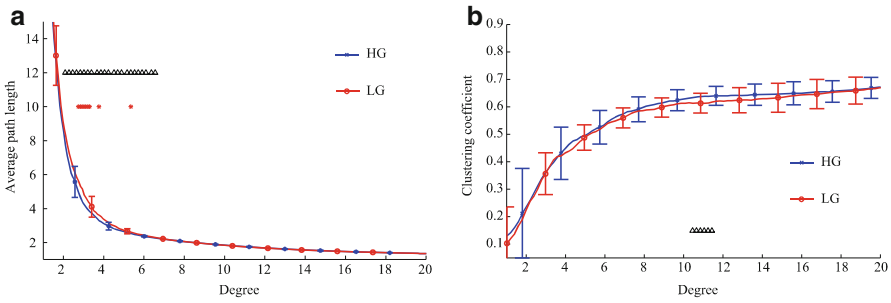


Fig. 3 The harmonic average path length (**a**) and the clustering coefficient (**b**) as a function of the degree of the networks. The groups were divided based on the achievement scale

The most significant point of the path length was shown at a degree value between 3 and 4. For a random network, the mean degree to ensure the graph is entirely connected is at least O (ln$N$) [2]. In this work, the number of nodes was 43. Therefore, the significant difference of harmonic average path length showed a broad line of degree value to make the graph connected.

Figures 2 and 3 show that the clustering coefficient was associated with the simultaneous processing and achievement scales. The clustering coefficient of HG was higher than LG. The difference, with statistical significance, could be found around a degree of 10.

In references [4, 6], the results both show that the clustering coefficient of brain networks was not associated with the participants' physiological features. In these studies, the relationship between the clustering coefficient and participants' features was assessed when the brain graphs were at a very sparse level in these previous studies. In this work, the relationship between the clustering coefficient and intelligence scores was found when the mean degree of graph were more than 10, which means that these brain networks were much denser than in the related studies.

Acknowledgment This research is partially supported the Aihara Project, the FIRST program from (MEXT Program for Fostering Regional Innovation).

References

1. Stam, C. J., de Haan, W., Daffertshofer, A., Jones, B. F., Manshanden, I., van Walsum, A. M. V., Montez, T., Verbunt, J. P. A., de Munck, J. C., van Dijk, B. W., Berendse, H. W. & Scheltens, P. 2009. Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimers disease. *Brain,* 132**,** 213–224.
2. Bullmore, E. & Sporns, O. 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience,* 10**,** 186–198.
3. Kaufman, A. S. 1983. *K-ABC: Kaufman assessment battery for children: interpretive manual*, American Guidance Service Circle Pinesˆ eMN MN.
4. Stam, C. J., Jones, B., Nolte, G., Breakspear, M. & Scheltens, P. 2007. Small-world networks and functional connectivity in Alzheimer's disease. *Cerebral Cortex,* 17**,** 92.
5. Watts, D. J. & Strogatz, S. H. 1998. Collective dynamics of 'small-world' networks. *Nature,* 393**,** 440–442.
6. van den Heuvel, M. P., Stam, C. J., Kahn, R. S. & Pol, H. E. H. 2009. Efficiency of functional brain networks and intellectual performance. *The Journal of Neuroscience,* 29**,** 7619–7624.

Computational Emotions

Mauricio Iza Miqueleiz and Jesús Ezquerro Martínez

Abstract Research on the interaction between emotion and cognition has become particularly active in the last years, and many computational models of emotion have been developed. However, none of these models is able to address satisfactorily the integration of emotion generation and its effects in the context of cognitive processes. This work tries to unify several models of computational emotions for embodied agents with the work done in cognitive architectures, based on psychological theories and applications.

Keywords Emotion processes • Computational models • Embodied agents • Communication

1 Introduction: Computational Models

Computational models take different frameworks in research and applications [1]. On the one hand, psychological models emphasize on fidelity with respect to human emotion processes. On the other hand, AI models evaluate how the modeling of emotion impacts reasoning processes or improves the fitness between agent and its environment. That is to say, the model improves and makes more effective the human-computer interaction.

Several models have been proposed and developed. However, some fundamental differences arise from their underlying emotional constructs. For instance, as we will see below, some discussions on whether emotion precedes or follows cognition disappears if one adopts a dynamic system perspective. Here, we will discuss three main approaches.

M. Iza Miqueleiz (✉)
Department Basic Psychology, University of Malaga, Malaga, Spain
e-mail: iza@uma.es

J. Ezquerro Martínez
Department of Logic and Philosophy Science, University of Basque Country, Donostia, Spain

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_2

2 Discrete Approach

This approach assumes appraisal theories of emotion [2]. In appraisal theories emotions are connected with how organisms sense events, relate them to internal needs, characterize appropriate responses and recruit cognitive, physical and social resources to act in adaptive ways.

Some models focus on appraisal as the core process to be modeled. In this sense, emotion is not completely elaborated. Mechanisms for deriving appraisal variables, via if-then rules, model specific emotion label. Emphasizing on a cognitive model of the situation, many models assume that (i) specific appraisal patterns are needed for emotion arising; and (ii) cognitive responses are determined by these appraisals.

This perspective is mainly concentrated on the cognitive structure of emotions, but doesn't account for the overall emotion process. As we will see below, the resulting computational models reflect this limitation [3]. Embodied emotion is considered as a dynamic and situated process, adjusting to the changing demands of the environment, rather than an appraisal of cognitive representations [4].

3 Dimensional Approach

Dimensional theories argue that emotions are not discrete entities. Rather, it is a continuous dimensional space [5]. These theories conceptualize emotion as a cognitive label attributed to a perceived body state, mood or core affect [6]. An agent is considered in an affective state at a given moment and the space of possible states within broad, continuous dimensions.

Although there is a relationship between both approaches, appraisal dimension is a relational construct that characterizes the relationship between some specific event (or object) and subject's emotion (belief, desire or intention). Furthermore, several appraisal variables can be active at the same time. Contrarily, the dimension of affect is a non-relational construct, indicating only the overall state of the subject.

These dimensional theories focus on the structural and temporal dynamics of core affect and often do not deal with affect's antecedent in detail. Affective states are not conceived intentional states, in the sense that affects are not about some formal object or target. Here, despite of symbolic intentional judgments, many sub-symbolic factors could contribute to a change in main affect.

Dimensional models are generally used for animated agent behavior generation. They translate emotion into a small number of dimensions that are continuously mapping features of behavior such as the spatial extent of a gesture. Similarly, these representational models can recognize human emotional behavior and are better at discriminating user affective states than the approach that only relies on discrete labels [5].

4 Communicative Approaches

These theories provide the social-communicative function of displays and can argue for dissociation between internal emotional processes and emotion displays that need not be selected on the basis of an internal emotional state. Their models often embrace this dissociation and dispense with the need for an internal emotional model, focusing on mechanisms that decide when an emotional display would effect on an agent.

For instance, Bicho et al. [7] have tried to validate a Dynamic Field model of joint action that implements neuro-cognitive mechanisms supporting human joint action. It has explained the existence of persistent inner states that lead to the emergence of high-level cognitive function. That is, cognitive processes unfold continuously in time under the influence of multiple sources of information.

Their robotics experiments show: (i) principles of DFT scale to high-level cognition in complex tasks (e.g., decision making in a social context, goal inference, error detection, anticipation, etc.); (ii) embodied view of "motor cognition" strongly contrasts with traditional AI approaches.

In this framework, taking the right actions is the result of efficient cognition: (i) action understanding and goal inference; (ii) anticipation of the user's needs (for fluency of team performance and acceptance); (iii) action monitoring, error detection and repair.

It does not look like a classical AI architecture (see [8]). It is a complex, fully integrated dynamical system, with no encapsulated modules or subsystems and an embodied view: high level cognitive functions like goal inference are based on sensorimotor representations. More, learning can be integrated in the DF framework.

5 Discussion

Several global principles can be extracted from the possibilities offered by this kind of models. Three of these are arguably the most important. Firstly, attention, as a control system to filter lower-level brain activity in order to allow few input representations to enter the higher level of thought and manipulation of neural activities (e.g., filters controlled by activity in parietal and pre-frontal cortices). Therefore, higher level processes such as thinking and reasoning work on a smaller number of input representations. Secondly, emotion, in terms of value maps learnt in orbito-frontal cortex (also coded in associated amygdala sites) in order to bias what is to be processed and to guide the choice of task goals (by their associated predicted rewards), constraining the inference chain. These value maps, joined with body activations and automatic brain-stem responses, are used to give emotional

value to decisions for action. Thirdly, long-term memory created on-line, in order to allow for incremental wisdom about the environment for use as a guide for further interactions.

A number of global principles could be added to these three possibilities, such as the use of hierarchical processing (as can be noted in vision, in order to create flexible visual codes for complex objects which can be used at a variety of scales). Another possibility involves the use of synchronization of neurons over long distances, such as 40 Hz frequency, in order to solve the binding problem of combining the different codes for objects (as happens in multi-modal hierarchical coding schemes). A further principle is that of recurrent loops of neural activity, to allow for the creation of short-term (or working memory) sites for the temporary holding of such activity for spreading around to other similar sites, in order to acting as report centers in the brain. Finally, there are the principles of integration and segregation of the neural systems of the brain that play a core role in brain processing efficiency.

Although there are a number of human psychological theories of emotion [9], it is generally agreed that emotions serve the purpose of increasing our ability to interact with our environment in a successful way. We have discussed that emotions can thus be used to increase collective behavior in a dynamic model, as it happens in the communicative approaches. We have analyzed the use of computational emotions toward increasing collaboration and collective behavior for both avatar and agent.

There exists broad experimental evidence on grounded language comprehension, such as when action related speech activates mirror system or action Sentence Compatibility Effect [10], where verbal description of spatially directional actions facilitates movements in the same direction. For instance, “Give me the wheel” activates the motor representation of a pointing/request gesture; whereas “I give you the wheel” activates the “reach-grasp-hold out” sequence.

In this sense, we have shown that Dynamic Field Theory offers a powerful theoretical language to endow autonomous robots with high-level cognitive capacities. DF-architecture for joint action represents a complex dynamical system of coupled neural populations, each with a specific functionality. Embodied and dynamic view on cognition strongly contrasts with more traditional AI approaches. It will be interesting in the future to develop a system in which each agent could evolve his communication strategies to evaluate how the agent’s dynamics and collective behavior are both affected by this additional dynamic.

Acknowledgements This work has been supported by the project (FFI2009-08999) of Spanish MICINN.

References

1. Anderson, JR (1993). *Rules of the mind*. Hillsdale NJ: LEA
2. Ellsworth PC & Scherer KR (2003). Appraisal processes in emotion. In RJ Davidson, KK Goldsmith & KR Scherer (eds.), *Handbook of the affective sciences*, pp. 572–595. New York: Oxford University Press.
3. Gratch J, Marsella S & Petta P (2009). Modeling the cognitive antecedents and consequences of emotion. *Journal of Cognitive Systems Research 10* (1), 1–5
4. Niedenthal PM (2007). Embodying emotion. *Science 316*, 1002–1005.
5. Barrett LF (2006). Emotions as natural kinds? *Perspectives on Psychological Sciences 1*, 28–58.
6. Russell JA (2003). Core affect and the psychological construction of emotion. *Psychological Review 110*, 145–172.
7. Bicho E, Louro L & Erlhagen W (2010). Integrating verbal and nonverbal communication in a dynamic neural field architecture for human-robot interaction. *Frontiers in Neurorobotics 4*, article 5.
8. Oudeyer, PY (2003). The production and recognition of emotions in speech: features and algorithms. *International Journal of Human Computer Studies 59* (1–2), 157–183.
9. Rolls E (2005). *What are emotions, why do we have emotions, and what is their computational basis in the brain*. Oxford University Press.
10. Glenbach AM & Kaschak MP (2002). Grounding language in action. *Psychon. Bull. Rev. 9*, 558–565.

Group Neurodynamics: Conceptual and Experimental Framework

Darius Plikynas, Saulius Masteika, Gytis Basinskas, Darius Kezys, and Pravin Kumar

Abstract Our multidisciplinary research challenges few essential and actual problems in the area of social neuroscience, i.e. (i) fundamental lack of understanding what is individual and collective consciousness, (ii) how to bridge group neurodynamics data to the behavioral social phenomena, and (iii) technological challenges addressing simultaneous measurements of brain EM activations and analyses in terms of spatial/temporal synchronizations and spectral coherence for the group of people. Our basic assumptions follow field theoretic consciousness interpretation approach and precede it assuming, that mutually entangled individual mind states create coherent social networks and behavior as well. The latter supposes that collective consciousness can be empirically investigated in a form of simultaneous group-wide coherent brain activations (e.g. using EEG measured delta, theta, alfa and beta oscillations). For the experimental validation of the proposed paradigm we have designed a three level experimental and simulation research framework, which starts from the modeling and benchmark estimates of correlations for the individual mind states, proceeds with estimates of the coherence and synchronicity measures for the group-wide neurodynamics and ends up with construction of oscillations based multi-agent social system simulation model.

This research project (OSIMAS) is funded by the European Social Fund under the Global Grant measure; [project No. VP1-3.1-SMM-07-K-01-137] http://osimas.vva.lt/.

D. Plikynas (✉) • G. Basinskas • D. Kezys
Research Center, Kazimieras Simonavicius University, J. Basanaviciaus 29a,
LT03109 Vilnius, Lithuania
e-mail: darius.plikynas@ksu.lt

S. Masteika
Research Center, Kazimieras Simonavicius University, J. Basanaviciaus 29a,
LT03109 Vilnius, Lithuania

Department of Informatics, Faculty of Humanities, Vilnius University, Muitines g. 8,
LT44280 Kaunas, Lithuania

P. Kumar
Meditation Research Institute, SRSG, Virbhadra Road, 249203 Rishikesh, Uttarakhand, India

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_3

Keywords Group neurodynamics • Social neuroscience • Oscillating agents • Multi-agent system

1 Introduction

Following technological advances, some new multidisciplinary research niches are opening up. One of such promising new experimental research areas is in the field of social neuroscience [1], which, enriched with the below listed conceptual field theoretic consciousness interpretation approaches, promises novel interpretation of coherent social networks and collective consciousness as well.

A review of conceptually related multidisciplinary research literature has revealed some other field theoretic cognitive and coordination interpretation approaches like the field model of consciousness (EEG coherence changes as indicators of field effects) [2], the CEMI theory of consciousness, the neurophysics of consciousness [3], the modeling of the neurodynamics complexity of teams [4], intra- and inter-cellular communication mechanisms [5], field computations in natural and artificial intelligence [6], the field-based coordination mechanisms for multi-agent systems in a robotics domain [7], amorphous or pervasive computation in contextual environments, nonlocal observations as a source of intuitive knowledge [8, 9] etc.

Based on the review of field based cognitive and coordination approaches, we made conceptual assumptions, hypothesis and postulates, which are framed by the proposed OSIMAS (oscillations based multi-agent system) simulation paradigm. It employs three conceptual models: PIF (pervasive information field), OAM (oscillating agent model) and WIM (wavelike interaction mechanism) [10]. Hence, proposed OSIMAS simulation paradigm potentially consolidates individual, group and social neurodynamics in the unified conceptual framework. Below we very briefly introduce experimental and simulation framework designed for the validation of the proposed paradigm.

2 Experimental and Simulation Framework: Individual, Group and Social Coherence Levels

For the experimental validation of the proposed paradigm we have designed a three level experimental and simulation research framework, which starts from the benchmark individual mind states, proceeds with estimates of the coherence and synchronicity measures for the group-wide neurodynamics, and ends up with the simulation of social coherence level, i.e. multi-agent system (MAS) design, see Fig. 1.

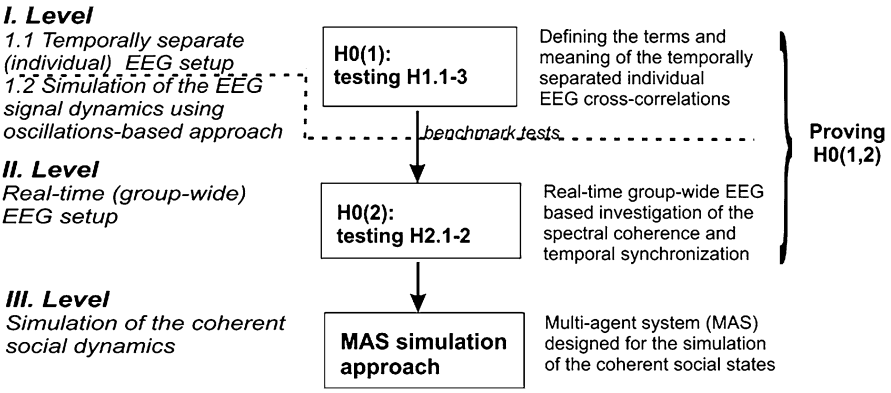


Fig. 1 Overall research framework: three major experimental and simulation research levels (individual, group-wide and MAS simulation)

2.1 Individual Coherence Level

The main purpose of the proposed individual experimental EEG framework – to find out whether brain-wave patterns, i.e. EEG-recorded mind-fields, can demonstrate mutually correlated behavior that depends solely on the states of people's minds. Hence, we tested few hypothesis (H0(1) and H0(2)), see Fig. 1. In short, our results also confirm, that temporally separated people doing the same mental activities demonstrate an increase of specific spectral coherence in their brain-wave patterns. Due to the space limitations in the current paper, plenty of obtained results can be provided in the extended version of this paper.

In short, benchmark EEG experimental results have shown partial confirmation of major OSIMAS assumptions concerning (a) oscillatory nature of the agent's mental states and (b) some common principles of power spectra redistribution in different mental states. Based on these experimental findings, we made adjustments to the conceptual oscillating agent model (OAM).

In order to find further insights about oscillatory nature of human mental states, we have designed a simulation of human brain EEG signal dynamics using a refined Kuramoto model [11], i.e. coupled oscillator energy exchange model (COEEM), which simulates human brain single channel EEG signal dynamics (see Fig. 2), using fourth-order Runge-Kuta algorithms (RK4).

We have obtained unique prognostication results, suitable for relatively accurate prognoses of the human brain EEG signal dynamics for the chosen EEG channels, see http://vlab.vva.lt.

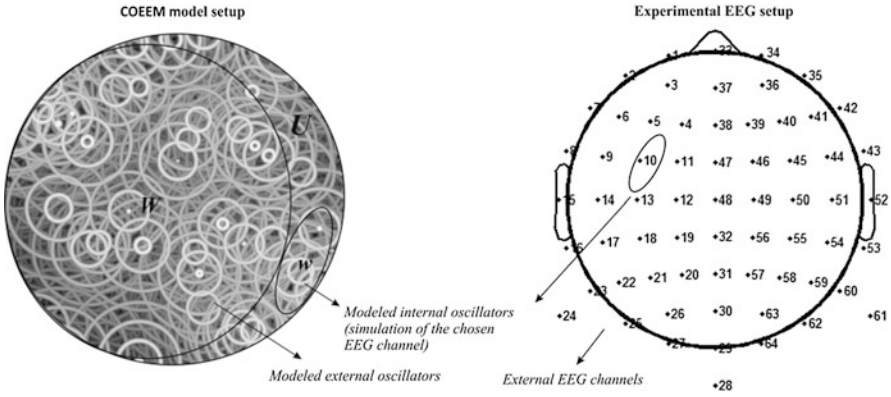


Fig. 2 The systems of the internal and external oscillators in the COEEM model

2.2 Group Coherence Level

Real-time group-wide EEG experiments (see Fig. 1) have also revealed spectral coherence and temporal synchronization patterns for the people in the same mental states (like concentration, contemplation, meditation, etc.). For analyses and visualization of group-wide neurodynamics we have designed GMIM method (URL: http://osimas-eeg.vva.lt), which shows simultaneous dynamics of the group-wide neurodynamic processes, see Fig. 3.

2.3 Social Coherence Level (MAS Approach)

From the individual and group-wide experiments we observe (see above), that mutually coherent individual mind states create coherent social networks and ordered social behavior as well. The latter supposes that collective consciousness can be empirically investigated in a form of simultaneous group-wide coherent brain activations. Technologically it is already possible. At this stage, though, more challenging is conceptual design and programmable construction of the oscillations based MAS simulation model, which is in the agenda of the OSIMAS project.

3 Conclusions

Empirical evidences of neurodynamic coherence of brain waves for the people in the similar mental states at individual and real-time group-wide levels provide conceptually novel way for the simulation of (i) social agents as oscillatory entities

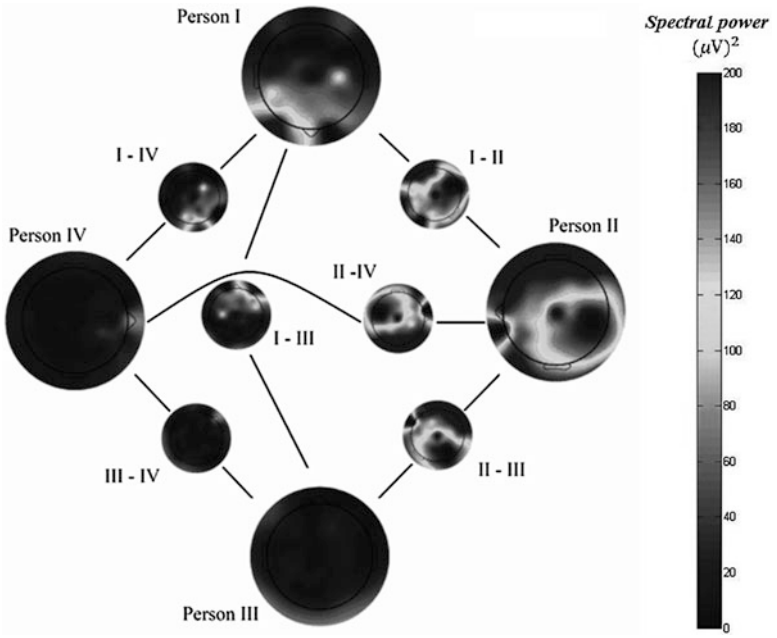


Fig. 3 Spectral power activation of brain wave dynamics for four persons in chosen mind state and spectral range (smaller diagrams and connecting lines indicate spectral power differences between respective persons)

and (ii) societies as coherent systems composed from such agents. This remarkable observation could start a new page in the social neuroscience further development.

References

1. Decety, J., Cacioppo, J.T. 2011. *Handbook of Social Neuroscience*. New York: Oxford University Press.
2. Travis, F.T., Orme-Johnson, D.W., 1989. Field model of consciousness: EEG coherence changes as indicators of field effects. International Journal of Neuroscience, 49, 203–211.
3. John, E.R., 2002. The neurophysics of consciousness. Brain Research Reviews, 39(1), 1–28.
4. Stevens, R., Galloway, T., Wang, P., Berka, C., Tan, V., Wohlgemuth, T., Lamb, J., Buckles, R., 2012. Modeling the neurodynamic complexity of submarine navigation teams. Computational and mathematical organization theory, DOI 10.1007/s10588-012-9135-9.
5. Rossi, C., Foletti, A., Magnani, A., Lamponi, S., 2011. New perspectives in cell communication: Bioelectromagnetic interactions. Seminars in Cancer Biology, 21 (3), 207–214.
6. MacLennan, B.J., 1999. Field computation in natural and artificial intelligence. Information Sciences, 119, 73–89.
7. Mamei, M., Zambonelli, F., 2006. Field-Based Coordination for Pervasive Multiagent Systems. Berlin: Springer-Verlag.

8. Radin, D., 2008. Testing Nonlocal Observation as a Source of Intuitive Knowledge. EXPLORE: The Journal of Science and Healing, 4 (1), 25–35.
9. Schwartz S A., 2008. Crossing the Threshold: Nonlocal Consciousness and the Burden of Proof. EXPLORE: The Journal of Science and Healing, 4 (1), 1–9.
10. Plikynas, D., Masteika, S., Budrionis, A., 2012. Interdisciplinary Principles of Field-Like Coordination in the Case of Self-Organized Social Systems. World Academy of Science, Engineering and Technology, 61, 754–759.
11. Pessa E. and Vitiello G. 2004. Quantum noise induced entanglement and chaos in the dissipative quantum model of brain. Int. J. Mod. Phys. B 18 (6), 841–858.

The Activity Modalities: A Priori Categories of Coordination

Lars Taxén

Abstract A conceptualization of a-priori forms of coordination as *activity modalities* is proposed. Sensations in various sensory modalities are integrated by our brain into a coherent, actionable percept, structured by the processes of *objectivation*, *contextualization*, *spatialization*, *temporalization*, *stabilization*, and *transition*. This conceptualization promises to bridge neuroscientific and applied research into the coordination problem.

Keywords Coordination • Sensations • Contextualization • Activity modalities • Integration

1 Introduction

The understanding of how actions are coordinated and integrated in the pursuit of fulfilling needs is of prime importance for brain research:

> I do not see any way to avoid the problem of coordination and still understand the physical basis of life [1, p. 176]

Coordination has been extensively investigated "internally" in various neuroscientific and cognitive studies, and "externally" in applied areas like, for example, organizational studies. Usually, neuroscientific research focuses of the inner workings of the brain, while the external environment is conceptualized in rather general, non-specific terms. As a case in point, see Fig. 1.

The brain is modeled in an elaborate way, while the environment is simply described as the "world". Thus, it is recognized that neural representations are influenced by the external world, but the character of these representations is not problematized. Also, influences in the opposite direction – from neural representations to the "world" – are not considered.

L. Taxén (✉)
Department of Computer and Information Science, Linköping University, Linköping, Sweden
e-mail: lars.taxen@gmail.com

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_4

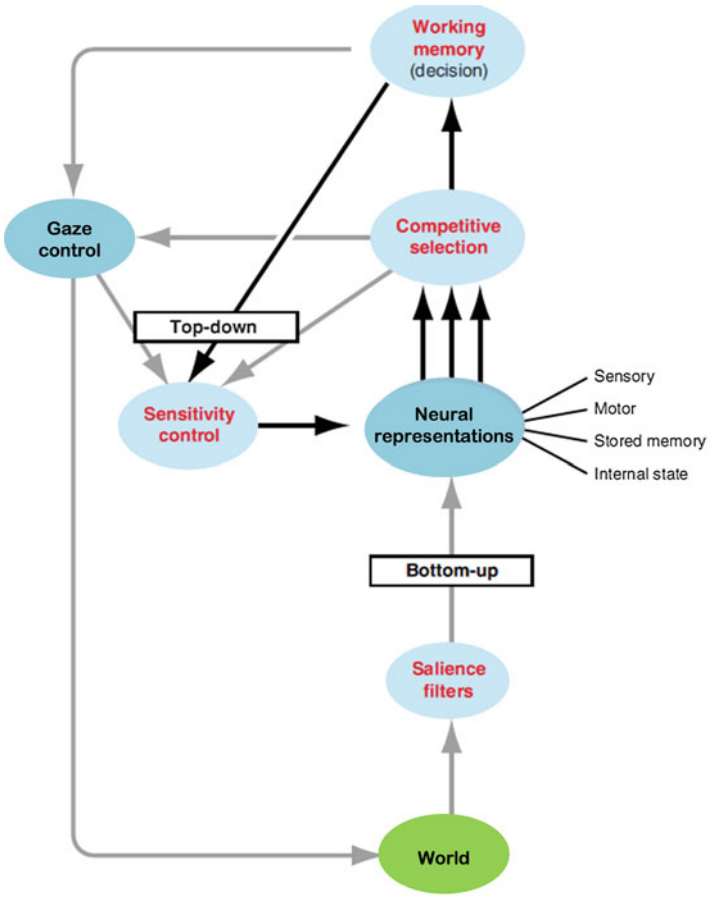


Fig. 1 An example of conceptualizing the brain – environment relation (Adapted from [2])

In a similar vein, "external" sciences such as organizational science, social sciences, information system development, system engineering, project management, etc., tend to regard the individual as a homogeneous ideal type that can be analyzed and manipulated as other, non-human elements. Thus, human abilities and limitations for acting are disregarded, which may have severe consequences for research in areas where humans are relevant. For example, models of organizations are often so complex that they are exceptionally hard to overview, understand and agree upon (see e.g. TOGAF [3]; a framework for developing so called enterprise architectures).

The separation between the internal and external research areas is understandable since research efforts have to be somehow delimited. However, this dichotomy needs to be overcome in order to advance our understanding of coordination. The brain, body and the environment should be seen in a unitary way, as succinctly stated by Llinás [4]:

> [The internal functional space that is made up of neurons must represent the properties of the external world – it must somehow be *homomorphic* with it [4, p. 65]

Concerning coordination, it is reasonable to assume that every healthy human being is born with certain phylogenetically evolved predispositions to coordinate actions in the same sense that we are born with legs for walking. Such predispositions need to be ontogenetically developed by acting in the various situations that the individual encounters during her life-span. These actions are manifested internally as a changed state of mind, and externally as various means enabling the actions. For example, the ability to orientate oneself in space is certainly alleviated by a map, once the significance of it has been understood by the individual.

The purpose of this contribution is to suggest that the homomorphy between the internal and external worlds can be conceptualized as *activity modalities*. These modalities – *objectivation*, *contextualization*, *spatialization*, *temporalization*, *stabilization*, and *transition* – should be conceived as a-priori forms, i.e. exigencies for coordinating actions in the same vein that Kant suggested time and space as a-priori forms for knowledge. Thus, the modalities are not something that can be experienced or sensed in the external world. Rather, they are evolutionary developed categories by which our brains are equipped for enabling action. The gist of this position is that sensations in various sensory modalities are integrated by our brain into a coherent, actionable percept structured by the activity modalities and their interdependencies.

The activity modality construct was conceived in my work with the coordination of extremely complex development projects in the telecommunication industry [5]. Gradually, after many years, external manifestations such as information models (spatialization), process models (temporalization), business rules (stabilization), as well as other organizational artifacts, were elaborated into the activity modality construct as putative, general categories of coordination.

The rest of the paper is organized as follows. First, I illustrate the activity modalities by the activity of ancient time mammoth hunting. Next, in order to provide a "foothold" from the external world into the intricacies of the inner world, I suggest modeling the brain as *dependencies between capabilities*. This approach is inspired from efficacious attempts to capture the essentials of complex systems in the telecom industry [6]. The paper is concluded with a discussion of the implications of the activity modality approach. In conclusion, I suggest that this line of thinking is promising enough for motivating further research efforts along this path.

2 The Activity Modalities

Imagine that you can travel some 30,000 years back in time, and you are one of the hunters in Fig. 2, motivated by the need to get food, acquire material for clothing, making arrowheads, and the like. What coordinative capabilities must you have in order to participate in this activity?[1]

A basic ability is that you can direct your attention to the *object* in focus for the activity – the mammoth. If you are unable to recognize the very target that your actions are meant to affect, you cannot coordinate your actions with the other hunters.

Second, given the object and the motive for acting, you need to conceive of a *context* – a center of gravity so to say – that enables you to recognize that which is relevant to the activity (and disregarding irrelevant things): hunters, bows, arrows, actions, shouts, gestures, and so on. For example, the river in the background is certainly relevant, since it is obstructs the mammoth from escaping in that direction. On the other hand, from participating in fishing activities you know that there are fishes in the river; but these are surely irrelevant in the mammoth hunting activity.

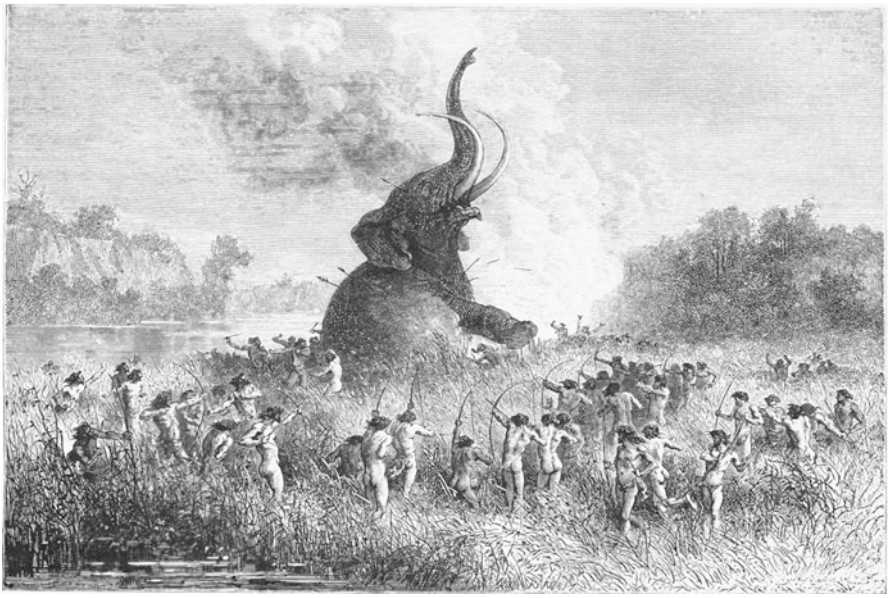


Fig. 2 Illustration of an activity (Original wood engraving by Bayard [8])

[1] "Activity" (German: Tätigkeit; Russian: deyatel 'nost') as used here refers to the rather specific sense it has in Activity Theory [7], meaning roughly "socially organized work". Thus, it is more precise than every-day English understanding of "activity". By "action" I refer to the actions of individuals participating in an activity.

The context is fundamental for making sense of individual actions. For example, it can be seen in the background of Fig. 2 that beaters scare the pray away by means of making noises and a lit fire. These actions would appear completely counterproductive if seen in isolation. It is only in the context of the entire activity that the beaters' actions become intelligible.

Third, a sense of the *spatial* structure of the activity needs to be developed in your mind. This enables you to recognize how the relevant things are related to each other, and what properties you confer upon them.

Fourth, you must acquire a sense for the *temporal* or dynamic structure of the activity. Your actions must be carried out in a certain order. For example, shooting an arrow involves the steps of grasping the arrow, placing it on the bow, stretching the bow, aiming at the target, and releasing the arrow.

Fifth, you cannot shoot your arrows in any way you like. If shooting in a wrong direction, other hunters may be hit rather than the mammoth. You need to know where to aim in order to hurt the mammoth the most. An understanding of how to perform appropriate mammoth hunting will be accrued after many successful (and, presumably, some less successful) mammoth hunts. Eventually, this lends a sense of *stability* to the activity; of the "taking for granted"; rules and norms indicating proper patterns of action that need not be questioned as long as they work.

Sixth, an activity is typically related to other activities. For example, the prey will most likely be cut into pieces and prepared to eat. This is done in a cooking activity, which in turn has its particular motive – to still hunger – and object, which happens to be the same as for the hunting activity: the mammoth. However, in this context, other aspects of the mammoth are relevant (as, for example what parts of the mammoth are edible). In order to participate in or conceive of other activities, you must be capable of refocusing your attention; to make a *transition* from one activity to another.

The six dimensions outlined above – *objectivation*, *contextualization*, *spatialization*, *temporalization*, *stabilization*, and *transition* between contexts – are denoted *activity modalities*. In order for an individual to coordinate her actions, her brain must be able to integrate multimodal sensory impressions into an actionable Gestalt in the form of the activity modalities and their interdependencies. This integration may be precluded by neurological deficiencies. For example, a brain lesion in the hippocampal area severely impairs spatial navigation, which in turn impedes orientation towards a desired target [9].

Even if coordinative capabilities are strictly individual, our social and physical environment is imperative for how these capabilities are manifested in the brain. Coordination between individuals can only occur through external means that enable individual actions to be sufficiently synchronized. Such means can be anything relevant for the activity: tools, speech, writings, gestures, symbols, and whatever else as long as they are pertinent for the coordination and integration of actions in order to achieve a common goal. This indicates that the activity should be the prime unit of analysis for coordinative inquiries. In order to pursue this line of thinking, I have suggested denoting activities characterized by the activity modalities as *activity domains* [5].

The idea behind illustrating the activity modalities by a historical activity such as a mammoth hunt is to convey that the underlying structure of activity domains is the same, regardless of time and place. The reason is simply that our phylogenetic constitution has not changed much (if at all) since we were hunting mammoths in ancient times. However, contemporary ontogenetic manifestations of the modalities are of course very different, not least because of the technological development of means.

3 An "Anatomy" of the Brain

If we hypothesize that the activity modalities are relevant for inquiring about coordination, we need to investigate how these capabilities are realized in the brain. In order to address this task, I propose to model the brain as *dependencies between capabilities*; from the most basic ones upwards.

This approach is inspired by experiences from managing the coordination of extremely complex development tasks in the telecom industry, where the system under development was conceived in precisely this way [6]. The principle behind is quite simple: the most important thing for understanding how a complex system works is to understand dependencies. Images illustrating such dependencies have, perhaps somewhat impertinent, been coined "anatomies" since it captures how the system "comes alive". So, for example, if the power on button of a computer is malfunctioning, none of the inherent capabilities of the computer can be actuated. In Fig. 3, such an "anatomy" of the brain is shown.

This anatomy should be regarded as a highly speculative and preliminary first attempt that needs further elaboration. Its character is strictly structural; there is no aspect of time or dynamics involved. Also, it is meant to illustrate the phylogenetic predispositions which are further elaborated into coordinative capabilities during ontogenesis. Thus, the ensuing capabilities of the individual are realized by the interplay between these predispositions and external influences from the environment.

How the capabilities are realized in Fig. 3 is not illustrated. This may appear strange at first glance, since the anatomy is meant to assist in finding out exactly that. However, the purpose of the anatomy is to provide a model of the whole of the brain; a model that is simple enough, yet adequate, for achieving a common understanding about it.

The anatomy should be read from the bottom up. A basic brain capability is the *motivating* one, indicating that the brain is capable of auto-activation and continual exploration of the environment. Next, a sensing capability is necessary, which is realized by the different sensory systems in the brain (visual, auditory, somatosensory, gustatory, and olfactory ones).

Sensing is a prerequisite for *attention*, which in turn depends of the capabilities *alerting* (achieving and maintaining a state of high sensitivity to incoming stimuli), *orientation* (the selection of information from sensory input), and *executive attention* (monitoring and resolving conflict among thoughts, feelings, and responses) [10].

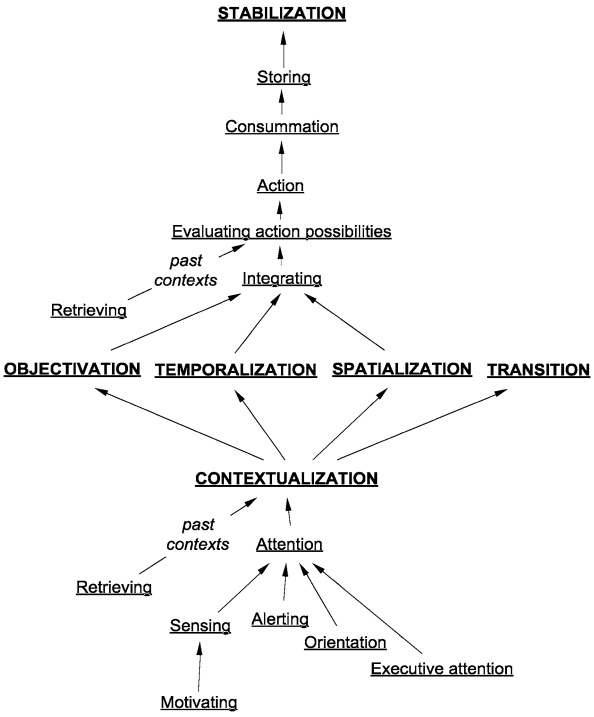


Fig. 3 Dependencies between capabilities in the brain

The *contextualization* capability is dependent on attention and the capability to resolve ambiguous percepts, which requires retrieval of akin percepts from long-term memory: "analogies are derived from elementary information that is extracted rapidly from the input, to link that input with the representations that exist in memory" [11, p. 1235].

With contextualization in place, the *objectivation*, *spatialization*, and *temporalization* capabilities can be invoked. In particular, it can be noted that objectivation depends on contextualization; indicating that the discrimination of an object in focus requires a contextual background. The *transition* modality is modeled as being dependent on contextualization, since transition is not possible if the contextualization capability is inhibited.

Next, the *integration* capability can be actuated, i.e., the formation of a pre-motor, actionable percept, which enables the *evaluation of proper action alternatives* utilizing similar percepts retrieved from long term memory. After the motor system executes the *action*, its consequences are evaluated by a "*consummation*"[2] capability, and the experience is stored in long term memory. This in turn is a prerequisite for the *stabilization* capability.

[2]Consummation is the last stage in Mead's conceptualization of the act "upon which the encountered difficulty is resolved and the continuity of organic existence re- established" [12, pp. 3–25]

4 Discussion and Conclusions

The approach presented in this paper is an attempt to address "The most important issue in brain research today [. . . ,] that of the internalization or embedding of the universals of the external world into an internal functional space" [4, p. 64].

The basic idea put forward is that our individual capabilities for action will inevitably reflect how we construct our artifacts and social institutions. Thus, the "universals of the external world" are not something that exists out there before mankind appeared on earth; rather these reflect "universals" of our brains, which evolution has brought about. This means that the activity modalities can be apprehended as "coordinative universals" by which our internal and external worlds are unified.

In fact, the capabilities that the activity modalities provide are so vital for our lives that they remain unnoticed by us, like the air we breathe. Every moment of our daily life, we effectuate these capabilities without reflection. It is only when we are struck by some lesion that the lack of a modality capability becomes apparent.

There exists a substantial amount of research that substantiates the idea of the activity modalities in the "external world" e.g. [5, 6]. However, there is a huge gap in extant research regarding how the modalities are realized in our brains. In spite of this, some interesting threads can be noticed. For example, the stabilization modality implies that every activity needs to uphold a delicate balance between the extremes of complete segregation and all-encompassing integration. At one extreme, the activity is disintegrated into non-communicating fragments, and at the other extreme, the activity is solidified and unable to change. This view has a striking resemblance to the concept of *metastability*:

> Metastability, [. . .] leads to a looser, more secure, more flexible form of function that can promote the creation of new information. [. . .] Too much autonomy of the component parts means no chance of coordinating them together. On the other hand, too much interdependence and the system gets stuck, global flexibility is lost. [13, p. 43]

Another putative connection is the formation of "global neurocognitive state" which "plays a critical role in adaptive behavior by allowing the organism to perceive and act in a manner consistent with the context of the changing situation in which it exists" [14, p. 61]. An idea worth pursuing is that the dynamical assessment of the state of an organism can be conceptualized, at least concerning coordination of actions, as activity modalities.

Concerning the grounding of the activity modality construct, relevant results from the immense knowledge base that exist in neuroscientific research today need to be addressed. In order to coordinate these results, a brain anatomy such as the one in Fig. 3, might be used. Such an image would function as a common target where pertinent research results can be related to different capabilities in the anatomy. After all, such images have been successfully used in coordinating other complex tasks!

I suggest that the activity modality approach may provide insights both in the internal and external worlds. In the external world, we may use the activity

modalities as a guiding framework for analyzing and constructing our social and physical environment. Likewise, we may study the nature of activity modality manifestations such as artifacts, norms, institutions, etc., in order to gain insight into the coordinative, functional organization of the brain. In conclusion, I claim that the conceptualization of coordination proposed in this contribution is promising enough to motivate further research efforts along this path.

References

1. Pattee, H.H. (1976). Physical theories of biological coordination. In M. Grene & E. Mendelsohn (Eds.), Topics in the philosophy of biology, 27, Boston: Reidel.
2. Knudsen, E.I. (2007). Fundamental Components of Attention. Annual Review of Neuroscience, 30, 57–78.
3. TOGAF (2013). Retrieved July 28th, 2013, from http://pubs.opengroup.org/architecture/togaf9-doc/arch/
4. Llinás, R.R. (2001). I of the vortex: from neurons to self. Cambridge, Mass.: MIT Press.
5. Taxén, L. (2009). Using Activity Domain Theory for Managing Complex Systems. Information Science Reference. Hershey PA: Information Science Reference (IGI Global).
6. Taxén, L. (Ed.) (2011). The System Anatomy – Enabling Agile Project Management. Lund: Studentlitteratur. ISBN 9789144070742.
7. Kaptelinin, V., and Nardi, B. (2006). Acting with Technology – Activity Theory and Interaction Design. Cambridge, MA: The MIT Press.
8. Bryant, W. C., Gay, S. H., (1983). A Popular History of the United States. Vol. I, Charles Scribner's Sons, New York.
9. Posner, M.I., and Petersen, S.E. (1990). The Attention System of the Human Brain. Annual Reviews of Neuroscience, 13, 25–42.
10. Posner, M.I., and Rothbart, M.K. (2007). Research on Attention Networks as a Model for the Integration of Psychological Science. Annual Review of Psychology, 2007 (58), 1–23.
11. Bar, M. (2009). The Proactive Brain: Memory for Predictions. The Philosophical Transactions of the Royal Society, (364), 1235–1243.
12. Mead, G.H. (1938). Philosophy of the act. Chicago: University of Chicago Press.
13. Kelso, J. A. S, & Tognoli, E. (2007). Toward a Complementary Neuroscience: Metastable Coordination Dynamics of the Brain. In L. I. Perlovsky, R. Kozma(Eds.) Neurodynamics of Cognition and Consciousness (pp. 39–59). Berlin Heidelberg: Springer.
14. Bressler, S.L. (2007). The Formation of Global Neurocognitive State. In L. I. Perlovsky, R. Kozma(Eds.) Neurodynamics of Cognition and Consciousness (pp. 61–72). Berlin Heidelberg: Springer.

On the Use of Cognitive Neuroscience in Industrial Applications by Using Neuroelectromagnetic Recordings

Giovanni Vecchiato, Anton Giulio Maglione, and Fabio Babiloni

Abstract Nowadays there is an emerging interest to apply the cognitive neuroscience recent scientific findings into the industrial context, to increase the efficacy of the interaction between humans and devices. The aim of the present chapter is to illustrate the possibilities of the use of electroencephalographic (EEG) signals recorded in humans to improve several industrial applications. In fact, the main message of this chapter is that the actual EEG technologies could be not only used in medical environments but also widely exploited into aeronautics and marketing contexts. It will be highlighted the use of EEG signals for application including the on-line monitoring of the mental workload of pilots, air traffic controllers and professional drivers. Physiological variables that can be used in neuromarketing analysis to assess variation of cognitive and emotional states during the perception of advertisements will be also illustrated. Finally, preliminary results of the first EEG experience of recording the brain activity during a visit of a real art gallery will be shown. Overall, such findings testify how the electrical neuroimaging seems to be now ready to pave the way to the industrial neuroscience.

Keywords Industrial neuroscience • Mental workload • Neuromarketing • Neuroaesthetic

G. Vecchiato (✉) • F. Babiloni
Department of Physiology and Pharmacology, University "Sapienza", Rome, Italy

Brainsigns s.r.l., Via Sesto Celere 7/c, 00152, Rome, Italy
e-mail: giovanni.vecchiato@uniroma1.it; fabio.babiloni@uniroma1.it

A.G. Maglione
Brainsigns s.r.l., Via Sesto Celere 7/c, 00152, Rome, Italy

Department of Anatomy, Histology, Forensic Medicine and Orthopedics, University of Rome "Sapienza", Rome, Italy
e-mail: antongiulio.maglione@uniroma1.it

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_5

1 Introduction

Brain Computer Interface (BCI) is a very active area of research in neuroscience aimed to investigate the on-line recognition and decoding of the brain activity, mainly through the use of electroencephalographic signals (EEG). BCI techniques have been developed along the last decade to improve the interaction of patients with a reduced or impaired muscular mobility with software or devices useful for them [1]. Applications have been realized in the BCI in the areas of prosthetics, or for the control of domotic applications.

2 Beyond the Use of EEG in Clinical Contexts

After focusing on neuroprosthetics, assistive technologies and domotic applications for motor impaired patients [2], Brain Computer Interfaces (BCIs) have been employed also in areas of interest for healthy subjects, such as playing videogames or controlling the mobility of helicopters or cars scale models [3, 4]. In fact, neurophysiological and electroencephalographic (EEG) measurements have been used in an attempt to characterize different mental states and to estimate the activity of the central and autonomous nervous system related to the driving tasks. This attempt is related to the generation of future devices able to detect on-line the degree of mental workload in pilots or human operators engaged in difficult surveillance tasks (reviewed by Borghini et al. [3]).

The attempt to decode brain activity of subjects related to the observation of incoming commercial stimuli is at the base of another industrial application of the EEG, usually called consumer neuroscience or neuromarketing. More precisely, such field of study concerns the application of neuroscience methods to analyse human behaviour related to marketing stimuli [5]. The reason why marketing researchers are interested in using brain imaging tools arises from the assumption that people are not usually able to fully explain their preferences when explicitly asked.

Another frontier for the industrial application of cognitive neuroscience is related to the evaluation of the brain activity during the appreciation of aesthetic stimuli, called neuroaesthetic. Such discipline was founded more than a decade ago and specifically refers to the study of the neural bases of beauty perception in art (reviewed recently in [6]). In fact, the modern EEG technology allows to record the brain activity in different environments and mobile conditions such as during the fruition of real masterpieces in fine art galleries, where they are usually experienced by visitors.

In the following sections we want to provide evidences of how the recording and analysis of brain activity through the EEG technology is now ready to be applied in several aspects of our everyday life. At this purpose, applications in BCI, on-line mental workload monitoring, neuromarketing and neuroaesthetic fields performed by our research team will be provided.

2.1 Active and Passive Brain Computer Interfaces

BCIs provide severely motor disabled people with an alternative channel to communicate and control simple devices. EEG based BCIs operate a real-time translation of predefined features into commands reflecting the users' intent. Many of the currently available BCI applications exploit different classes of EEG features as control signals, such as sensorimotor rhythms, steady-state visual evoked potentials, event-related potentials (ERPs) and the modulation of sensorimotor rhythms (SMRs; [1]). Different kind of BCI applications can be thought on the base of these kinds of potentials. In addition, such as measurements could be useful in several applications related to the monitoring of the brain states in complex tasks, such as driving cars or piloting aircrafts. The general idea of the "Passive Brain Computer Interfaces" (PBCI) is to provide to the operator a real-time feedback, by measuring the continuous EEG activity, about the mental engage required by the task. In such a case it could be possible to derive objective measurements of the internal cognitive state of the user during the task. The use of such PBCI enlarge the area of application of standard BCI toward different industrial applications, such as the on-line detection of operators' mental workload during simulated car driving. Figure 1 presents a typical experimental setup for such PBCI technology, jointly performed by the Authors and prof. Wanzeng Kong, responsible of the laboratory of the Dept. of Computer Science in the Hangzhou Dianzi University, Hangzhou, China

Fig. 1 A typical setup for a study involving the measurements of the mental workload during the drive of a simulated car. The passive BCI return real-time information about the "perceived" mental effort of the subject performing the driving task

2.2 On-Line Mental Workload Estimation and Monitoring

The aim of this area of research is to design, implement and evaluate an online, passive, EEG-based BCI system to quantify the mental workload of the subject during tasks performed at three different difficulty levels, using machine learning techniques. The system was tested by subjects performing the Multi Attribute Task Battery (MATB; [7]). We tested the system during the online run performed in each session. In particular, for each session, we trained the classifier using the data acquired during the first three runs of the session (*easy* and *hard*), and we applied the parameters to the online run in order to allow the operator to visualize directly on the visual interface the output of the classifier (mental workload index), for all the subtasks (*easy*, *medium* and *hard*) provided by the online run (Fig. 2).

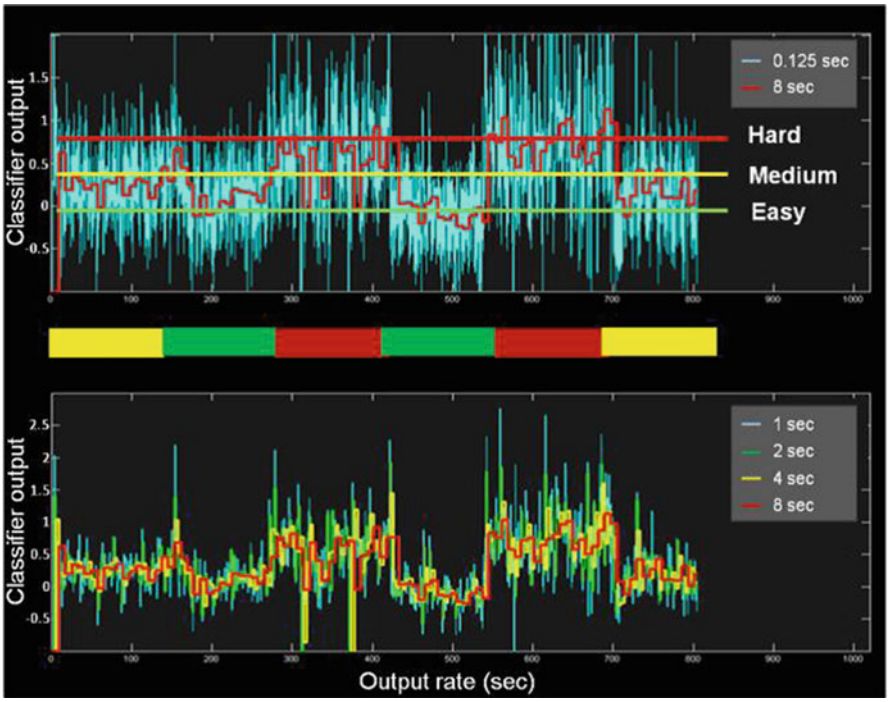


Fig. 2 Screenshot of the visual interface provided to the operator that allows to visualize the workload index (classifier output) over time. The *red line* is the average workload index, estimated in real time by the EEG recordings of the human operator. Note as during the different difficulty of tasks proposed to the subject (indicated by the different colors of the bar at the center of the figure, being *red* the most difficult task, and the *green* the easiest one) the workload index shows a different value. The higher the index, the higher the estimated workload

2.3 Neuromarketing

The final goal of the study is to link significant variations of electroencephalographic and autonomic variables with the cognitive and emotional reactions during the observation of TV advertisements. In an interview after the recording, subjects were asked to recall commercial clips they remembered and to score them according to the degree of pleasantness they perceived. Hence, the dataset has been divided into several subgroups in order to highlight differences between the cerebral activity related to the observation of the remembered and forgotten ads, and those between the liked against the disliked commercials. Both cognitive and emotional processing have been described by the Memorization (MI), Attention (AI) and Pleasantness Index (PI), which have been presented a previous work [8]. The percentage of spontaneous recall is linearly correlated with the MI values ($R^2 = 0.68$, $p < 0.01$). In particular, when both MI and AI are under their average values the percentage of spontaneous recall (18 %) is under average as well. This percentage is slightly increased (20 %) when the AI exceeds the average threshold. The highest values of spontaneous recall correspond to MI values over average. In fact, in this case the percentage reaches the value of 33 % when the AI is under average and the value of 41 % when both MI and AI are over average (Fig. 3).

As to the PI, the de-synchronization of left alpha frontal activity is positively correlated with judgments of high pleasantness.

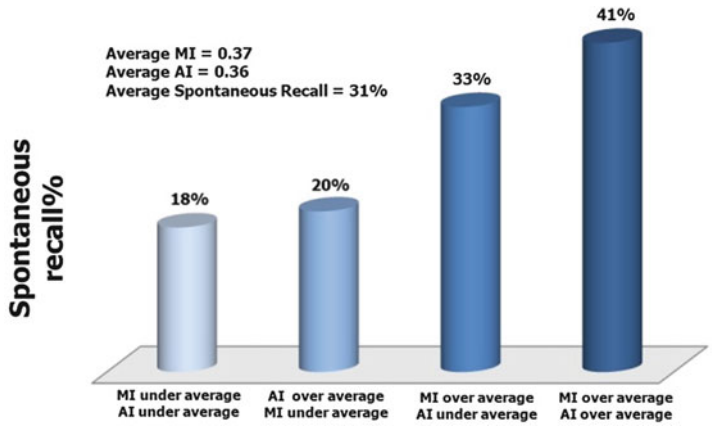


Fig. 3 Picture shows the percentage of spontaneous recall when Memorization Index (MI) and Attention Index (AI) are above or below their average value

2.4 Neuroaesthetic

The aim of this research is to investigate the neuroelectrical correlates of brain activity during the observation of real paintings in a visit of the "Scuderie del Quirinale" art gallery in Rome, Italy, during an exhibition of the XVII century Dutch's painters, including J. Vermeer (1632–1675). Interest for this study was to examine as the emotional and motivational factors, indexed by EEG asymmetry [9], could be related to the experience of viewing paintings. For this purpose, a proper Approach/Withdrawal index (AW) has been defined, as already presented in a previous study [8]. To our knowledge, this is the first study facing the collection of brain activity during the observation of real masterpieces in a national art gallery. Almost all the paintings received an appreciation substantially greater than the baseline ($p < 0.05$). It is also possible to note that the maximum value of the z-score (10.3) for the AW index is reached in occurrence of the painting number 8 of the visit, related to the J. Vermeer's "Woman with a red hat". Such evaluation was in agreement with the verbal judgments obtained in the analyzed population. The Pearson correlation between the verbal judgments provided by the sample considered and the values of the AW index was 0.505, that was indeed statistically significant ($p < 0.05$) (Fig. 4).

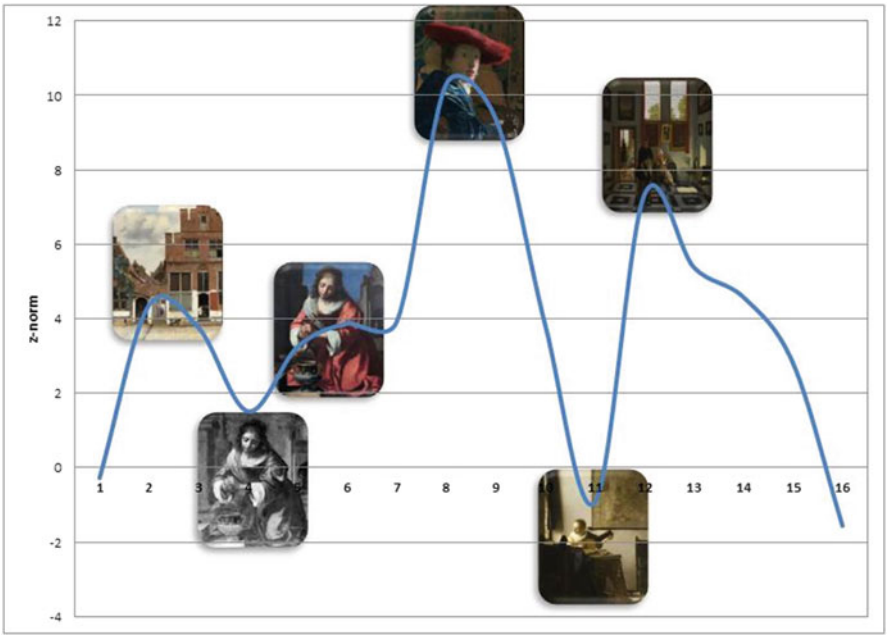


Fig. 4 Variation of the Approach/Withdrawal index along the paintings observed during the gallery visit. Number refers to the different pictures visited along the gallery in the same order by all the subjects involved in the experiment

All the above mentioned examples suggest that it is now possible to employ the neuroelectrical activity to address issues related to our perception of cognitive workload, commercial advertisements and also pictures in an art gallery. How this will open the way for a more intense use of the cognitive neuroscience in the industrial contexts it is still to be clarified in the next years. However, it is out of doubt that the process to use the knowledge from cognitive neuroscience in the industrial contexts has already started.

Acknowledgments This work has been performed with the support of the Italian Minister for Foreign Affairs "Direzione Generale per la Promozione del Sistema Paese" in the 2013.

References

1. Wolpaw J R and McFarland D J 2004 Control of a two-dimensional movement signal by a noninvasive brain–computer interface in humans Proc. Natl Acad. Sci. USA 101 17849–54
2. Babiloni F, Cincotti F, Marciani M, Salinari S, Astolfi L, Aloise F, De Vico Fallani F, Mattia D. On the use of brain-computer interfaces outside scientific laboratories toward an application in domotic environments. Int Rev Neurobiol. 2009;86:133–46.
3. Borghini G, Astolfi L, Vecchiato G, Mattia D, Babiloni F., Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. Neurosci Biobehav Rev. 2012 Oct 30. doi:pii: S0149-7634(12)00170-4. 10.1016/j.neubiorev.2012.10.003.
4. Doud AJ, Lucas JP, Pisansky MT, He B. Continuous three-dimensional control of a virtual helicopter using a motor imagery based brain-computer interface. PLoS One. 2011;6(10).
5. Lee N, Broderick A J, Chamberlain I. What is neuromarketing? A discussion and agenda for future research. Int J Psychophysiol, vol 63, no 2, pp 199–204, 2007.
6. Di Dio C, Gallese V. Neuroaesthetics: a review. Current Opinion in Neurobiology 2009, 19:682–687
7. Comstock J. R., "Matb – Multi-Attribute Task Battery For Human Operator Workload And Strategic Behavior Research," Jan. 1994.
8. Vecchiato et al., "Understanding the Impact of TV Commercials: Electrical Neuroimaging". IEEE Pulse, Vol.3(3), Page(s): 42–47. 2012.
9. Davidson RJ. (2004) What does the prefrontal cortex "do" in affect: perspectives on frontal EEG asymmetry research. Biol Psychol. 67(1–2):219−33.

A Study of c-VEP/SSVEP BCI Hybrid System

Gang Zhang, Guangyu Bin, and Xiaorong Gao

Abstract Recently, there has been an increasing interest in hybrid brain-computer interface (BCI) systems. Many hybrid BCI studies have been reported, using ERD (event-related desynchronization), SSVEP (steady-state visual evoked potentials) and other responses to construct hybrid BCI systems.

Hybrid BCI research is a novel research area. A hybrid BCI is a communication system that combines BCI with another means of sending information [3]. It requires more complex controls and processes.

c-VEP – (Coded visual evoked potential) BCI system has been approached using pseudo-random sequence encoding to evoke VEPs in BCI system. c-VEP BCI system successfully achieved 32 targets at 60 Hz display frame rate. SSVEP BCI is based on frequency modulation frequency modulation VEP. Since frequency distribution of c-VEP is wide band around nearly white noise, it can be superimposed with the SSVEP fixed frequency signal.

Keywords Hybrid BCI • SSVEP • ERD

1 Introduction

The system consists of an EEG amplifier and a PC with a LCD monitor. Figure 1 depicts the basic structure of the system. Mipower amplify (Design by Laboratory of Neural Engineering, Tsinghua University), which has a parallel port for trigger synchronization, was used in the test system. Visual stimuli were presented on a LCD monitor with a 60 Hz refresh rate. For precise timing control of stimulus

G. Zhang (✉)
Department of Biomedical Engineering, Tsinghua University, Beijing 100084, People's Republic of China

College of Biochemical Engineering, Beijing Union University, Beijing 100023, People's Republic of China
e-mail: gxr-dea@mail.tsinghua.edu.cn

G. Bin • X. Gao
Department of Biomedical Engineering, Tsinghua University, Beijing 100084, People's Republic of China

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_6

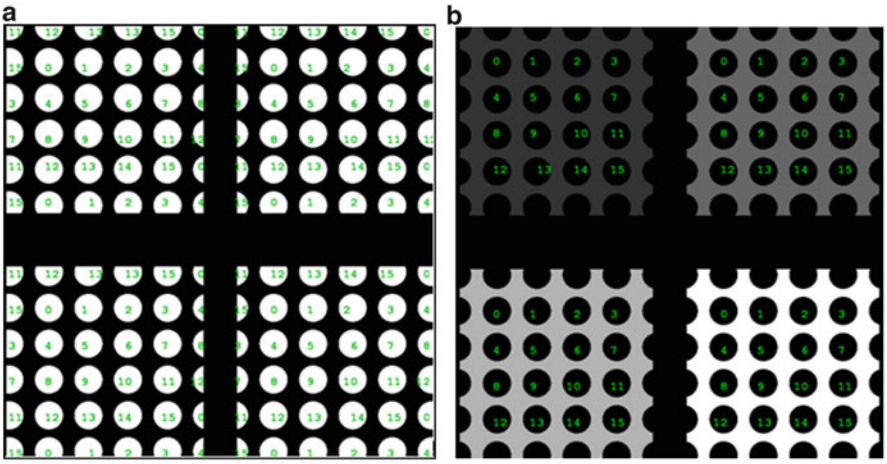


Fig. 1 (**a**) c-VEP stimulation matrix. (**b**) SSVEP stimulation blocks

presentation, DirectX technology (Microsoft Inc.) was utilized in the stimulus program, with the refresh rate of the monitor is 60 Hz. The stimulation sequence in c-VEP is: 1000011110011011111110101000101100.

As shown in Fig. 1, four Rectangle background areas represent in four different frequencies flickering inside, with 16 targets were arranged as a 4×4 matrix (in Fig. 1a). Four flicking frequency (8 Hz, 9 Hz, 11 Hz, 12 Hz) was used to evoke SSVEP (in Fig. 1b), Four flicking frequency (60/9 Hz 60/8 Hz, 60/7 Hz, 60/5 Hz) was used in former experiment.

Each target was periodically modulated by a 63-bit binary m-sequence that could evoke C-VEP. Time lag between two consecutive targets is set to four frames, and the used modulation sequences in one period were the same m-sequence for all targets. The experiment studied the feasibility of the hybrid BCI systems based on c-VEP and SSVEP. SSVEP stimulate was used as the back ground to divide the screen into four regions, each region used M pseudo-random sequence to divide into 16 targets. There were total 64 targets in this system. Canonical correlation analysis (CCA) method was performed to extract frequency information associated with the SSVEP. A template matching method was used for target identification of c-VEP [1, 2].

As shown in Fig. 2, the ellipse mask is c-VEP stimulation, the Rectangle background is SSVEP stimulation. CCA: canonical correlation analysis method.

A template matching method is generally used for target identification. To obtain the template, a training stage must be implemented. The steps of target identification are as follows:

1. In the training stage, the user is instructed to fixate on one of k targets, with the fixation target denoted by T0. During N stimulation cycles, EEG data Xn, n51, 2 . . . N is collected.

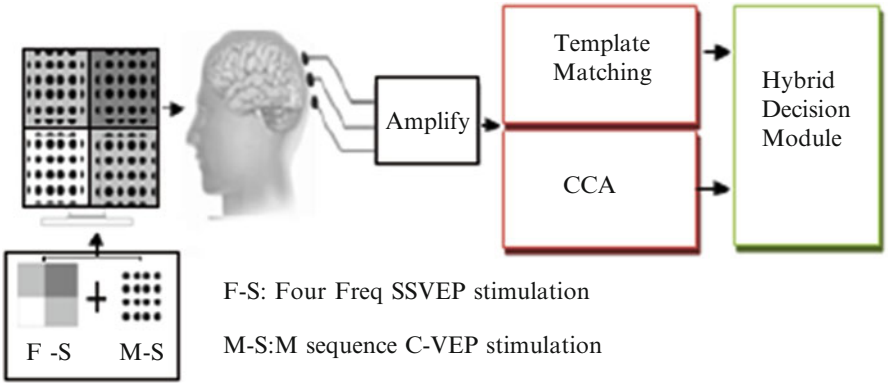


Fig. 2 A c-VEP/SSVEP BCI hybrid system structure diagraph

2. A template T is obtained by averaging over N cycles.

$$T(t) = \sum_{n=1}^{N} X_n \tag{1}$$

3. The templates of all targets are obtained by shifting T:

$$T_k(t) = T\left(\text{mod}\left(t - (\tau_k - \tau), p\right)\right. \tag{2}$$

4. For a segment of EEG data x, the correlation coefficient between x and the template T is calculated as:

$$\rho_k = \frac{T_k x^T}{\sqrt{T_k {T_k}^T x x^T}} \tag{3}$$

5. The target will be selected with the C which have the max correlation coefficient value.

$$C = \max_i (\rho_k, i = 1, 2 \ldots K) \tag{4}$$

6. The SSVEP correlations between the projections of the variables onto these basis vectors are mutually maximized. Consider the linear combinations which use canonical correlation analysis.

In Fig. 3 Circle c-VEP stimulation has almost the same Autocorrelation with Square stimulation c-VEP. In experiments, they have almost the same good perform (Fig. 4).

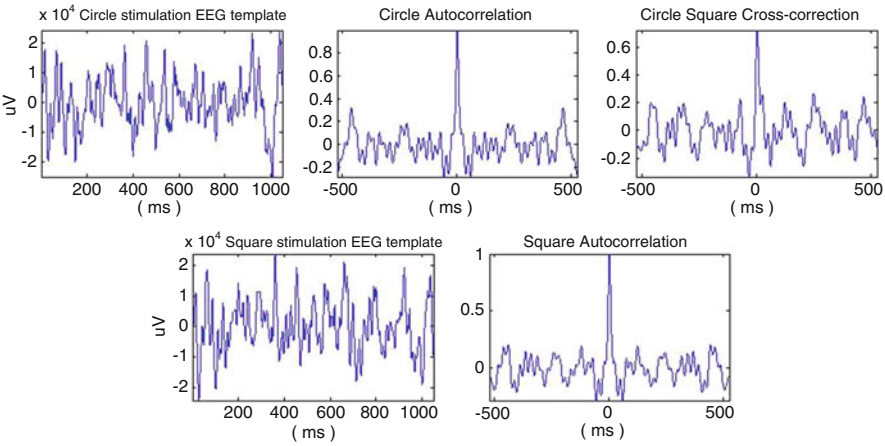


Fig. 3 Compare Circle stimulation c-VEP. With former Square stimulation c-VEP

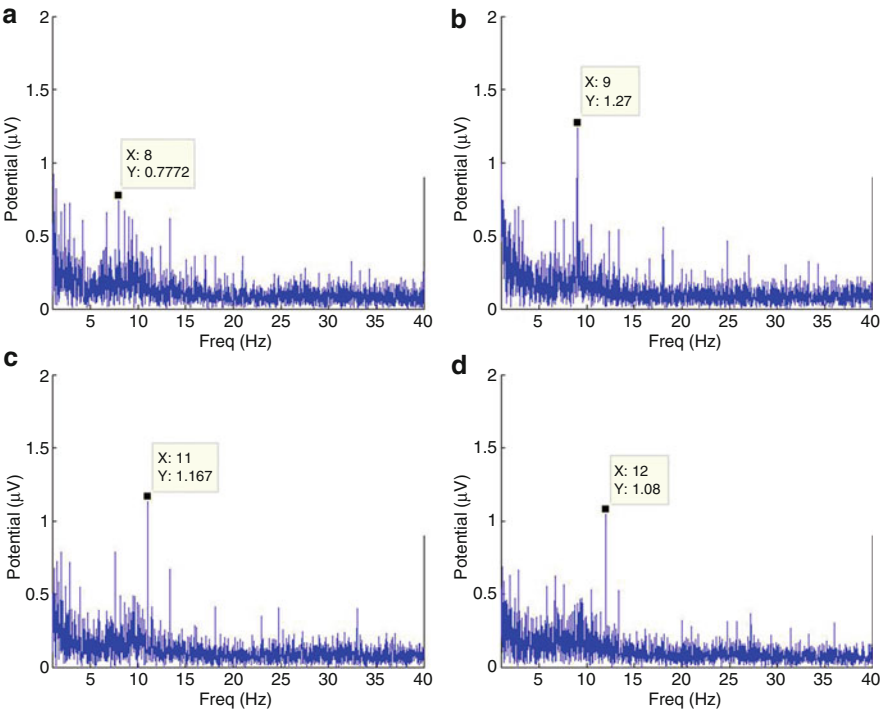


Fig. 4 SSVEP background FFT result. (**a**) Poz 8 hz ssvep background FFT. (**b**) Poz 9 Hz ssvep background FFT. (**c**) Poz 11 Hz ssvep background FFT. (**d**) Poz 12 Hz ssvep background FFT

Table 1 Results

	Accuracy (%)			ITR (bits/s)		
Subject	Hybrid	SSVEP	c-VEP	Hybrid	SSVEP	c-VEP
S1	85	100	100	80	81	55
S2	70	95	90	66	80	43
S3	75	95	80	70	82	47
S4	65	75	85	53	38	39
S5	92	98	94	129	62	87

2 Results and Conclusions

Compares the average training accuracy and ITR (information transfer rate) with c-VEP BCI system, with selection twice decision and unshielded environment. Six channel O1, O2, Oz, Po3, Po4, Poz has been used, reference was Cz channel. The results of the experiment were summarized in Table 1.

The border style and size of SSVEP stimulation would affect ITR value when using as background. The experimental results illustrated that the hybrid BCI system combines the four targets SSVEP BCI and 16 targets c-VEP BCI to get the 64 targets hybrid BCI.

Acknowledgments This work was supported by the National Basic Research Program of China (grant no 2011CB933204) and National Natural Science Foundation of China (grant no 90820304). The authors would like to thank Pujie Zheng, Zhikai Chen, Guangyu Bin for their comments and suggestions.

References

1. Bin G Y, Gao X R, et al. A highspeed bci based on code modulation vep. Journal of Neural Engineering, 2011,8(2):025015.
2. Bin G Y, Gao X R, et al. 2009 An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method J. Neural Eng. 6 046002.
3. Brendan Z Allison, et al. A hybrid ERD/SSVEP BCI for continuous simultaneous two dimensional cursor control, Journal of Neuroscience Methods 209 (2012) 299–307.

Part II
Higher Mental Functions and Dysfunctions

Memory Consolidation from Seconds to Weeks Through Autonomous Reinstatement Dynamics in a Three-Stage Neural Network Model

Florian Fiebig and Anders Lansner

Abstract Long-term memories for facts and events are not created at an instant. Rather, memories stabilize gradually over time and involve various brain regions. The shifting dependence of acquired declarative memories on different brain regions – called systems consolidation – can be tracked in time by lesion experiments and has led to the development of the Complementary Learning Systems framework, which focuses on hippocampal-cortical interaction. Observations of temporally graded retrograde amnesia following hippocampal lesions, point to a gradual transfer from hippocampus to cortical long-term memory. Spontaneous reactivations of hippocampal memories, as observed in place cell reactivations during slow-wave-sleep, are supposed to drive cortical reinstatements and facilitate this process.

We propose a functional neural network implementation of these ideas and furthermore suggest an extended three-stage framework that also includes the prefrontal cortex and bridges the temporal chasm between working memory percepts on the scale of seconds and consolidated long-term memory on the scale of weeks or months.

We show that our three-stage model can autonomously produce the necessary stochastic reactivation dynamics for successful episodic memory consolidation. The resulting learning system is shown to exhibit classical memory effects seen in experimental studies, such as retrograde and anterograde amnesia after simulated hippocampal lesioning.

Keywords Memory consolidation • Working memory • Complementary Learning Systems • Synaptic depression • Neural adaptation • Retrograde amnesia • Anterograde amnesia

F. Fiebig (✉) • A. Lansner
Computational Biology, Royal Institute of Technology KTH and Stockholm University, Stockholm, Sweden
e-mail: fiebig@kth.se

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_7

1 Introduction

Based on theoretical consideration of incremental learning in artificial neural networks, McClelland et al. [1, 2] concluded, that the existence of at least two Complementary Learning Systems (CLS) appears to be necessary, serving an adaptive function, allowing for processes of selective learning, memory strength modulation and gradual acquisition into stable long-term memory without sacrificing one-shot learning capability.

Today, multiple brain areas are thought to support declarative memory. While Working Memory is supported by the pre-frontal cortex (PFC), intermediate memory involves the medial temporal lobes and especially the hippocampus (HIP); Neocortex (CTX) provides a slow learning substrate for long-term memories. Current memory consolidation models [3, 4] assume that long-term memory consolidation of declarative memory is facilitated by the repeated reinstatement of previous activations in the neocortex.

Reactivations are known to be driven by the hippocampus and have been well studied with respect to place cells [5, 6]. Sleep and its various phases have been proposed to modulate network dynamic and plasticity, thus promoting the memory consolidation process [7–9], whereby interference between new learning (awake) and consolidation (asleep) is avoided.

Bilateral damage to the hippocampus incurs significant anterograde amnesia along with often temporally graded retrograde amnesia [10–14]. Prefrontal working memory is also aided by hippocampus, but only when demand exceeds its own very limited capacity [15].

2 Aims

We aim to build a neural network model that can autonomously consolidate memory from the scale of seconds (working memory) to weeks (long-term memory). Our dynamically interconnected model of three key brain areas is a functionally improved extension of the Complementary Learning Systems (CLS) framework [1, 16].

We focus on three major challenges:

- Replicate the temporal dynamics of memory consolidation seen in biological data, including effects of hippocampal damage.
- Implement autonomous replay as an intrinsic neural mechanism for consolidation
- Functionally bridge the full temporal range from seconds to weeks

3 Model and Methods

By implementing a Hebbian depression of synaptic connections, a biologically inspired Bayesian Confidence Propagation Neural Network (BCPNN) can autonomously reinstate previously learned attractors without need for external cues, noise bursts or other top-down controls [17]. The dynamics rest on slow synaptic depression in conjunction with neural adaptation, which allows an otherwise fast converging attractor network to shift randomly between learned attractors. Due to inherently balanced learning and forgetting in BCPNNs, the network retains its ability to learn new memories/attractors indefinitely and does not suffer from the potential pitfalls of other types of attractor networks, such as catastrophic forgetting.

Our model rests on three BCPNN-type populations of varying size and drastically different degrees of plasticity on the order of **3 min (PFC)**, **3 h (HIP)**, and **6 days (CTX)** respectively.

The simulation evolves in three phases (Figs. 1 and 2) plus one phase for recall testing afterwards, during which plasticity is turned off. Each new training pattern is

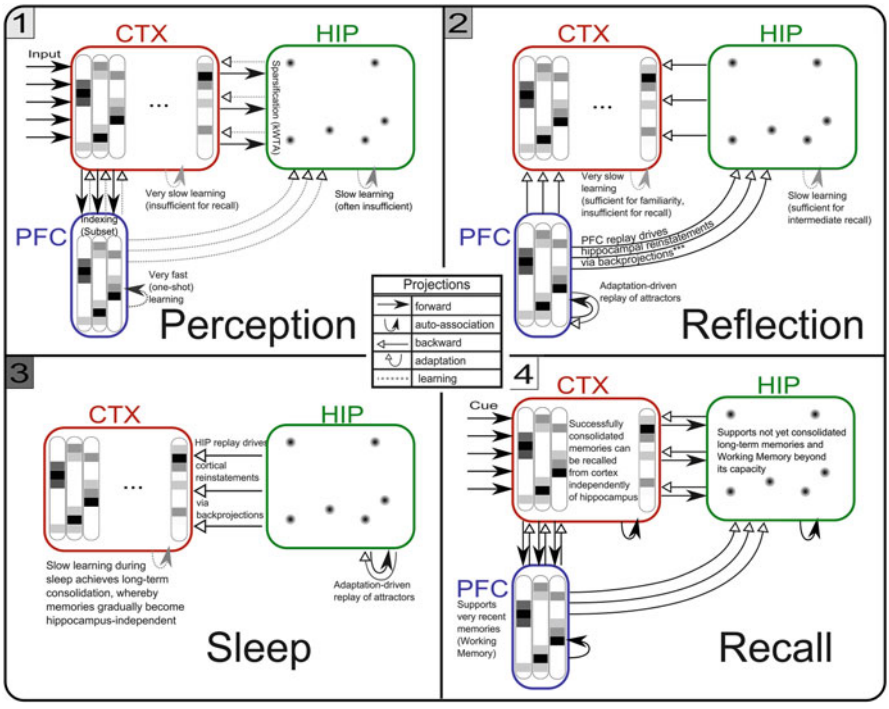


Fig. 1 From prefrontal working memory to long-term neocortical memory. The four panels show the three simulation phases (*1–3*) and their active components, as well as the configuration during cued recall (*4*). Activity in the cortical areas is organized into hypercolumns, while hippocampal activity is sparser, pattern-separated and lacks columnar organization

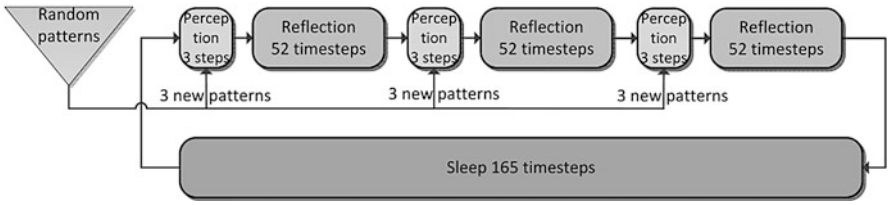


Fig. 2 The simulation cycle with its three alternating phases, named perception, reflection and sleep. Online learning occurs only during perception

shown for only one simulation timestep (brief online learning episode), forcing one-shot learning in the PFC, as the other networks learn too slowly for recall after this short exposure. Consolidation is then achieved through spontaneous reactivation of learned patterns, which will in turn cause the corresponding patterns to be projected in the next network and thus potentially learned or strengthened. Neural network dynamics are simulated in steps of 10 ms. Due to runtime considerations (simulating more than a month of time), plasticity of all networks was scaled by a factor of 26,000, so one day can be simulated in 330 steps.

4 Results

The three modeled brain regions learn and forget dynamically over time, as can be seen by plotting the probability of successful recall for each stage (Fig. 3). For added clarity, Recall was averaged (daywise) for patterns of the same age. Also note that the x-axis of Fig. 3 is logarithmic. The PFC reliably stores only the most recent patterns of the current day (age = 1d), while HIP can recall patterns for about a week. CTX memories can last for over a month, but consolidation is highly competitive as not all patterns are replayed/reinstated often enough: About a third of all patterns become consolidated into neocortical long-term memory. This dynamic can be biased, however, by introducing a plasticity modulation (e.g. relevance signal) which temporarily boosts hippocampal encoding. This, in turn, raises replay strength and frequency and thus improves consolidation odds and speed. On average, maximum systems consolidation is reached after about a week. For patterns between 2 and 5 days old, combined recall drops sharply following simulated hippocampal lesioning.

Retrograde amnesia is temporally graded, confirming experimental lesion data [1, 18, 19], see Fig. 4. Very recent pattern recall is not impacted as it is supported by working memory (PFC). Anterograde amnesia is flat, because future consolidation is permanently impaired, as well as hippocampal performance. This amnesia pattern is markedly different from simulated sleep deprivation (a consistent 50 % reduction in simulated sleep), which decreases system consolidation performance notably, but leaves hippocampal memory functional, so the anterograde effect becomes

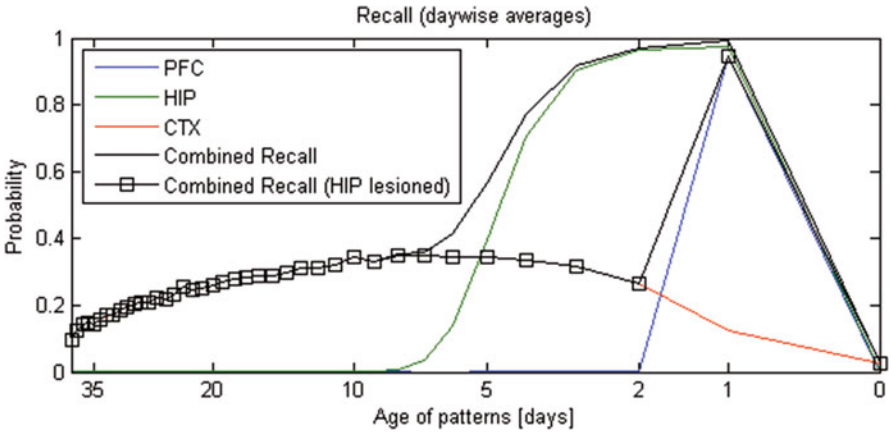


Fig. 3 The probability of successful recall plotted against age of patterns, for the three different modeled brain regions (PFC, HIP, and CTX)

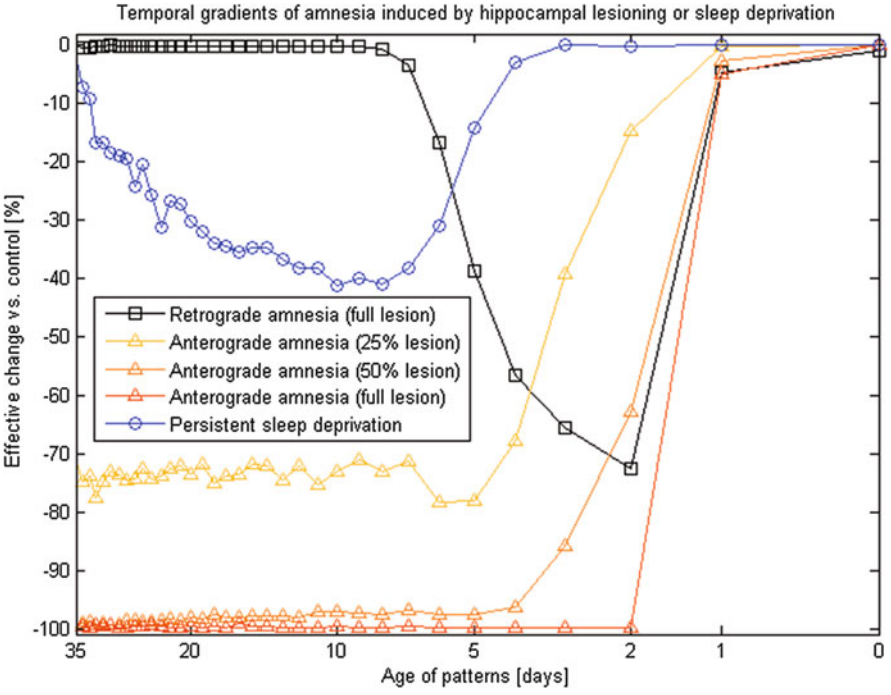


Fig. 4 Amnesia gradients, expressed as percentage wise performance changes, following partial or complete hippocampal lesions (deletion of a fraction of HIP units)

noticeable only after about 5 days. Also, because of reduced overwriting, cortical forgetting is also lower, which reduces the degree of amnesia for very old patterns.

5 Conclusion

Our model produces the necessary stochastic dynamics for episodic memory consolidation. Working Memory can hold patterns for less than a day and cortical learning is much too slow for any useful online learning from brief percepts, yet autonomous consolidation through the hippocampal system can yield stable cortical long-term memories, some of which can be recalled from cortex more than a month later.

Despite significant simplifications, the model show many of the properties and characteristics observed experimentally. These include competitive consolidation, retrograde facilitation after impaired acquisition (not specifically shown here), as well as typical amnesia effects after simulated hippocampal lesions.

With respect to the latter, we conclude that the model exhibits temporally graded retrograde amnesia (RA) qualitatively similar to pathologies seen in animal lesion experiments [12, 18, 19] and human case studies, such as Patient H.M. [10] and others [20]: Intact working memory, temporally graded RA, preserving remote cortical memories, as well as severe, flat anterograde amnesia. Our model predicts strong recall of very recent patterns, as they are supported by hippocampally-independent working memory. This is a testable prediction, given neurophysiological deactivation of hippocampal function on the timescale of working memory, such as focal cooling may allow [21]. Experimental lesion studies cannot account for the fleeting storage of new percepts in short-term memory. For example, test animals (rats, monkeys) require a resting period after the lesioning operation, before retesting and thus necessarily exclude short-term memory.

A similar network model to the one presented here, but with spiking model neurons is currently under development and future developments might include multiple trace theory, schema consolidation, hippocampal reconsolidation by letting the HIP stay plastic during replay, such that we not only consolidate neocortical traces during SWS, but replayed hippocampal attractors also reinforce, degrade or otherwise change themselves with each reinstatement event as demonstrated, for instance, by Lundqvist et al. [22].

Our results should be seen as mainly qualitative. The values of almost all our parameters – including the scaled learning rates – can be questioned on biological grounds. However, our model features a broad array of neurobiological details and to the authors knowledge, it is the first to show the viability of a three-stage consolidation chain, driven by autonomous replay that turned attractors into more useful quasi-stable attractors and thus expands the architectural options available to memory researchers looking for appropriate neural network models today.

References

1. McClelland, J.L., McNaughton, B L & O'Reilly, R C, 1995. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3), pp.419–457.
2. McClelland, J.L., 1998. Brain and Values: Is a Biological Science of Values Possible. In K. H. Pribram, ed. Mahwah, NJ: Erlbaum, pp. 535–547.
3. Norman, K. & O'Reilly, R., 2003. Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological review*.
4. Norman, K., 2010. How hippocampus and cortex contribute to recognition memory: revisiting the complementary learning systems model. *Hippocampus*.
5. Lee, A.K. & Wilson, Matthew A, 2002. Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron*, 36(6), pp.1183–1194.
6. Carr, M.F., Jadhav, S.P. & Frank, L.M., 2011. Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nature Neuroscience*, 14(2), pp.147–153.
7. Wilson, M A & McNaughton, B L, 1994. Reactivation of hippocampal ensemble memories during sleep. *Science*, 265(5172), pp.676–679.
8. Buzsaki, G., 1998. Memory consolidation during sleep: a neurophysiological perspective. *J Sleep Res*, 7(1), pp.17–23.
9. Qin, Y.L. et al., 1997. Memory reprocessing in corticocortical and hippocampocortical neuronal ensembles. *Philos Trans R Soc Lond(Biol)*, 352(1360), pp.1525–1533.
10. Scoville & Milner, 1957. Loss of recent memory after bilateral hippocampal lesions. *Neurology, Neurosurgery, Psychiatry*, 20(11), pp.11–21.
11. Squire, L R & Zola-Morgan, S., 1985. The neuropsychology of memory: New links between humans and experimental animals. *Annals of the New York Academy of Sciences*, 444, pp.137–149.
12. Zola-Morgan, S. & Squire, L R, 1990. The Primate Hippocampal Formation: Evidence for a Time-Limited Role in Memory Storage. *Science*, 250(4978), pp.288–290.
13. Squire, L R & Zola-Morgan, S., 1991. The medial temporal lobe memory system. *Science*, 253(5026), pp.1380–1386.
14. Squire, L R, Knowlton, B. & Musen, G., 1993. The structure and organization of memory: A synthesis from findings with rats, monkeys, and humans. *Annu Rev Psychol*, 44, pp.453–495.
15. Jeneson, A. & Squire, Larry R, 2012. Working memory, long-term memory, and medial temporal lobe function. *Learning & memory (Cold Spring Harbor, N.Y.)*, 19(1), pp.15–25.
16. Norman, K.A. & O'Reilly, Randall C, 2003. Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychological review*, 110(4), pp.611–46.
17. Sandberg, A. & Lansner, A, 2002. Synaptic depression as an intrinsic driver of reinstatement dynamics in an attractor network. *Neurocomputing*, 44–46, pp.615–622.
18. Winocur, G., 1990. Anterograde and retrograde amnesia in rats with dorsal hippocampal or dorsomedial thalamic lesions. *Behavioural brain research*, 38(2), pp.145–54.
19. Kim, J.J. & Fanselow, M.S., 1992. Modality-specific retrograde amnesia of fear. *Science (New York, N.Y.)*, 256(5057), pp.675–7.
20. Squire, Larry R. & Cohen, N., 1979. Memory and amnesia: resistance to disruption develops for years after learning. *Behavioral and Neural Biology*, 25(1), pp.115–125.
21. Tanaka, N. et al., 2008. Effective suppression of hippocampal seizures in rats by direct hippocampal cooling with a Peltier chip. *Journal of neurosurgery*, 108(4), pp.791–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18377260 [Accessed September 10, 2013].
22. Lundqvist, M., Herman, P. & Lansner, Anders, 2011. Theta and Gamma Power Increases and Alpha/Beta Power Decreases with Memory Load in an Attractor Network Model. *Journal of Cognitive Neuroscience*, 23(10), pp.3008–3020.

Biologically Inspired Models of Decision Making

Wlodzimierz Klonowski, Michal Pierzchalski, Pawel Stepien, and Robert A. Stepien

Abstract We present two biologically inspired models, demonstrating how people make decisions – one, called Game of Choosing (GoC), demonstrates how promotion and autopromotion influence decision making, and second, called Chaosensology Model (ChS), demonstrates how emotions may influence logical decision making. Our results confirm what has been known in psycho-sociology since William James works.

Keywords Selection • Elections • Logical thinking • Emotions • Neurodynamics

1 Introduction

We have been stressing the importance of time scale differences for a long time (cf. [1, 2] and references therein). But until very recently most economists assumed that people act rationally while undertaking economic decisions, like what to buy, when, and for how much. In 2002 Daniel Kahneman, a social psychologist, got Nobel Prize in Economy for demonstrating that process of decision making in most cases has rather emotional character than really rational one. Kahneman used notions of two Systems – 'System 2' that is responsible for conscious thinking and logical decisions, while 'System 1' is responsible for emotions and subconscious 'decisions' [3]. Psychologists suggested existence of a 'central executive structure' between our consciousness and our subconsciousness [4].

Here we present two biologically-inspired models. The first model demonstrates how small differences in initial preferences of a subject influence the decision (e.g. which mating partner to choose) and shows that such decisions are made based practically only on emotions. This model may be applied also to economic and

W. Klonowski (✉) • M. Pierzchalski • P. Stepien • R.A. Stepien
Laboratory of Biosignal Analysis Fundamentals, M. Nalecz Institute of Biocybernetics and Biomedical Engineering, Polish Academy of Sciences, 02-109 Warsaw, Poland
e-mail: wklon@ibib.waw.pl

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_8

even socio-political decision making. Our second model shows that what distinguish logical thoughts from emotions is the difference of the characteristic times of these two kinds of brain neurodynamics processes.

2 Game of Choosing Model

Our Game of Choosing (GoC) model [5] is inspired by process of choosing mating partner. That is why we have called subclasses of this model adequately 'monogamic' and 'polygamic', The subjects – system elements that make decisions – may act independently or may interact one with another. Also objects – system elements that are being chosen – may be independent or may be somehow correlated. Other subdivision is into 'natural', when environment is neutral, and 'controlled', when environment 'puts pressure' on the subjects making decisions. So we have together 16 subclasses of GoC models. The model is characterized by discrete time – subsequent moments are called steps. We have developed software to simulate such decision making processes with different parameters. The simplest class is the 'natural monogamic' model with independent subjects and uncorrelated objects – any subject after several steps makes decision to choose one given object and both the subject and the object are not taking part in further steps of the process.

Initially (step '0') each subject has some preference towards each object, so we have certain preference matrix (Fig. 1). To make decision to choose a given

	1	2	3	4	5	6	7	8	9	10	11
1	49.7592	49.4952	50.9414	49.4677	49.6746	50.3609	50.4613	50.5145	49.5433	49.6844	49.1763
2	50.2347	50.2177	50.7832	49.6456	49.3166	50.6012	49.1164	50.6732	49.1098	50.4707	50.4739
3	50.4741	50.5665	49.4662	50.1468	49.149	49.5142	50.2241	49.0547	50.9898	50.2747	50.183
4	50.1337	50.9751	50.2041	49.7141	49.3779	50.6664	49.2057	49.1954	49.7752	50.7339	50.7111
5	50.5869	49.8278	49.1371	50.1678	50.4936	49.1763	50.545	49.9146	49.0585	49.4473	50.237
6	50.8576	49.0054	50.7768	50.9223	49.6974	50.2537	50.086	50.6278	49.7109	49.5553	49.5901
7	49.2677	49.3593	49.7972	49.051	50.4597	50.4665	49.5414	50.2044	49.7312	50.9729	50.4271
8	50.3757	50.3127	49.1762	49.9739	50.9682	49.2709	50.5677	50.1342	50.1023	50.5273	49.0589
9	50.8121	49.5324	50.1652	49.5282	50.5642	50.6471	50.9906	49.7448	49.5895	49.206	50.0405
10	50.9286	50.3924	49.7686	50.5998	49.9885	49.0909	50.9552	49.0263	50.8484	49.3874	49.4674
11	50.3187	50.3424	50.9251	49.7688	50.322	50.1404	50.7283	50.6342	49.7365	49.74	49.1436
12	49.0835	50.1786	50.7604	50.4273	50.9439	50.4399	49.6544	49.8347	49.3307	49.1205	49.8711
13	50.3801	50.1769	50.1279	49.9441	49.6121	50.4414	49.9595	50.3566	50.4295	50.1329	49.5868
14	49.2345	50.2104	50.5781	49.0363	50.1797	49.9544	50.5331	49.2736	50.3538	50.5048	49.6763
15	50.9969	49.8511	50.3334	49.2958	50.7713	50.3417	50.0955	49.3126	50.3691	49.5482	50.849
16	49.4058	49.5584	50.0082	49.5548	49.5275	50.7173	50.9782	50.3626	49.9709	49.8043	49.1156
17	50.2324	50.0315	50.9206	49.925	50.661	50.8921	50.0904	49.2826	49.5848	49.1969	49.1887
18	50.3528	49.4491	49.6853	49.6364	50.1531	50.2291	49.4703	50.7644	49.9192	49.7261	50.6447
19	50.7117	50.4617	49.7374	49.5986	50.8338	49.4774	49.2752	50.5362	50.9636	49.7803	49.874
20	49.3425	49.6459	49.0637	49.0731	50.1885	50.6357	49.0813	49.3088	50.2992	49.6314	49.4254
21	49.9393	50.8181	49.6509	49.0957	50.8491	49.1632	50.6998	50.8669	50.9353	50.5577	50.7056
22	50.9441	49.3189	49.1569	50.0199	50.8569	49.544	49.8621	49.2735	49.7198	50.3959	49.6175

Fig. 1 Preference matrix; *columns* – subjects (11); *lines* – objects (22); *greenish* – rather 'yes', $p > 0.5$; *reddish* – rather 'no', $p < 0.5$

	1	2	3	4	5	6	7	8	9	10	11
1	1.00909	0.997115	0.998615	0.995672	0.998867	1.00126	0.990034	0.996345	1.00478	1.00129	1.00898
2	1.00872	0.998427	0.99688	1.00692	1.00405	0.998117	1.00543	0.995841	0.993479	1.00471	0.99932
3	0.999884	0.990174	0.992509	0.996426	0.993815	1.00367	0.992974	1.0087	0.993713	1.00591	1.00391
4	0.998563	1.00243	1.00446	0.99678	0.991882	1.00944	0.991327	1.00352	0.993444	1.00289	0.999581
5	1.00766	1.00577	0.991025	1.0032	0.993015	1.00469	1.0038	0.992591	1.00293	1.00725	0.99653
6	1.00432	0.996328	1.00739	1.00488	1.00509	1.00525	1.00391	0.992538	1.00121	0.998872	0.991003
7	1.00505	0.993262	0.998067	0.990342	1.00909	1.00465	1.00243	0.994637	1.00902	1.00985	1.00059
8	1.00055	0.990866	1.00868	0.997663	1.00242	0.993275	1.00068	1.00619	0.995939	0.995688	0.990687
9	1.00263	0.995433	0.995053	0.995923	1.00904	0.998769	1.00823	0.999644	0.99713	0.998605	1.00774
10	0.991747	1.00031	0.999197	0.998275	1.00412	1.00521	1.00256	0.994834	1.00591	1.00495	0.99378
11	1.00212	1.00908	1.00548	0.999657	0.996974	0.997144	1.00909	1.00328	0.996227	0.998274	0.996952
12	0.991699	1.00932	1.00566	0.992277	0.997461	1.00286	1.00365	1.00423	0.991567	1.00471	1.00849
13	0.995052	0.991166	1.00917	1.00271	1.00028	1.00791	1.00202	0.993868	1.00301	0.998694	1.00787
14	1.00421	0.995652	1.00331	0.994492	1.00416	0.993567	0.990823	0.994772	0.998856	0.997952	0.994643
15	0.990329	1.00426	1.00984	1.0092	0.995303	0.999817	0.999198	1.00132	1.00492	0.993235	0.990813
16	0.991226	1.0076	0.990632	1.00684	0.995517	1.00379	1.00051	0.99274	1.00134	0.993126	1.009
17	0.994282	0.996015	1.0058	1.0077	1.00331	0.992194	0.999003	1.00645	0.998101	0.993021	0.991751
18	1.00551	0.993237	1.00253	0.998009	1.00869	0.991846	1.0015	1.00395	1.00425	0.998114	1.00816
19	1.00927	0.992965	1.00659	0.999261	1.00951	0.992828	1.00163	0.999116	0.996893	1.00193	1.00251
20	0.996383	1.00262	1.00372	1.00708	1.00785	0.996502	0.993731	0.996799	0.991426	1.00777	0.999066
21	1.00962	0.999029	0.995985	1.00525	0.994381	0.99495	0.992269	1.00773	1.00055	1.00571	1.00639
22	0.990144	1.00243	0.999503	0.995824	1.00959	1.00675	1.00195	0.999578	0.991225	0.998831	1.00096

Fig. 2 Impact matrix; *columns* – subjects (11); *lines* – objects (22); *greenish* – rather 'positive', $i > 1.0$; *reddish* – rather 'negative', $i < 1.0$

object the subject's preference must exceed certain threshold. The preference matrix evolves in each step being multiplied element by element by impact matrix that represents influence of the objects on the subjects (Fig. 2). In more complicated GoC classes one also averages the impact over neighbour subjects and/or objects and takes into account external impact of the environment (Figs. 3 and 4; cf. [6]). In such cases chaos may emerge in the system [7].

For example, in Fig. 3, each of 11 subject chose one and only one object ('mating partner')'; subject 9, needed 77 repetitions (objects presentations) to make the choice while subject 10 made decision already after 46 presentations. In 'polygamic model' on Fig. 5, subject 5, chose as many as six objects (shown in dark green), while several subjects still chose only one object each; moreover subject 7 strongly rejected object 14 (shown in dark red). When interactions between subjects is taken into account objects 3, 4, 8, 14−17, and 22 are rejected by all 11 subjects, with subject 14 being strongly rejected by as many as six subjects (Fig. 6). In Fig. 7, objects belonging to class 7 are chosen by all but two subjects. In a controlled GoC with feedback between subject and environment again, like in simple monogamic case, each subject has chosen one object but rejections are scattered (Fig. 8) while when feedback between subjects is added all subjects reject objects 1, 3, 4, 8, 14, 16, 17, 22. Such situation may be observed rather during elections of MPs then in choosing a mate.

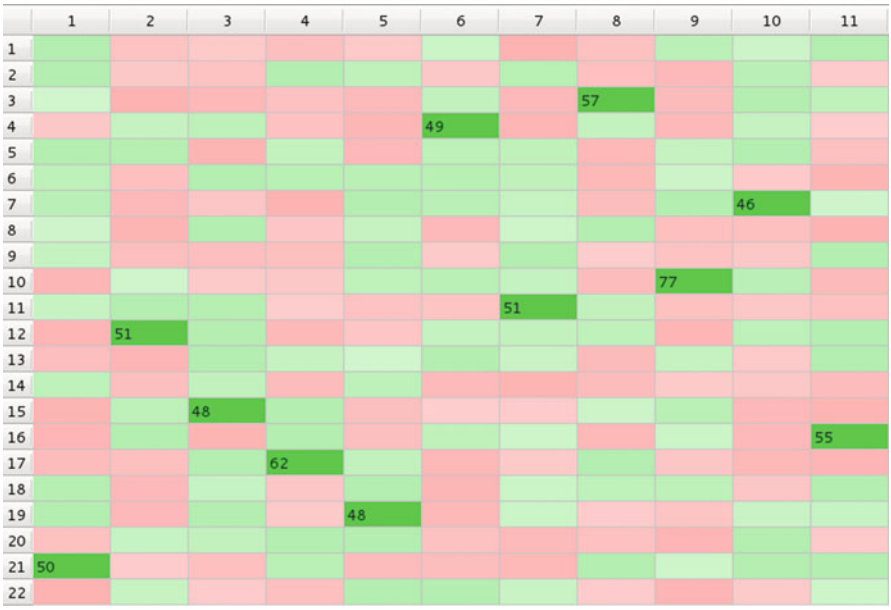


Fig. 3 Natural GoC – 'monogamic model' (additional material: mono.mpeg)

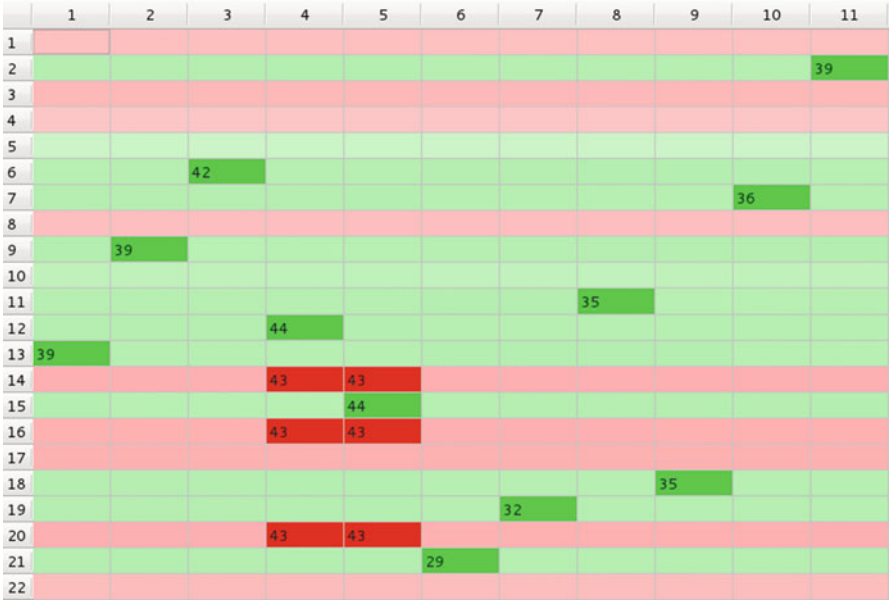


Fig. 4 Controlled GoC with feedback between subjects and the environment and with interactions between subjects (additional material: oenv_m.mpeg)

Fig. 5 Natural GoC – 'polygamic model' (additional material: poly.mpeg)

Fig. 6 Natural GoC with interaction between subjects (additional material: env_m.mpeg)

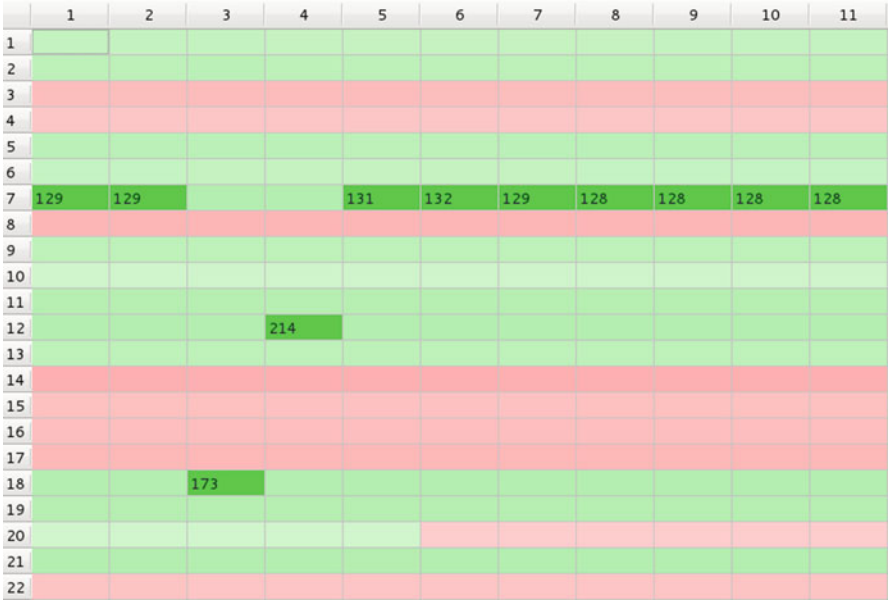


Fig. 7 Natural GoC with interaction between subjects with the number of objects in each subclass at least equal to the number of subjects (additional material: ienv_m.mpeg)

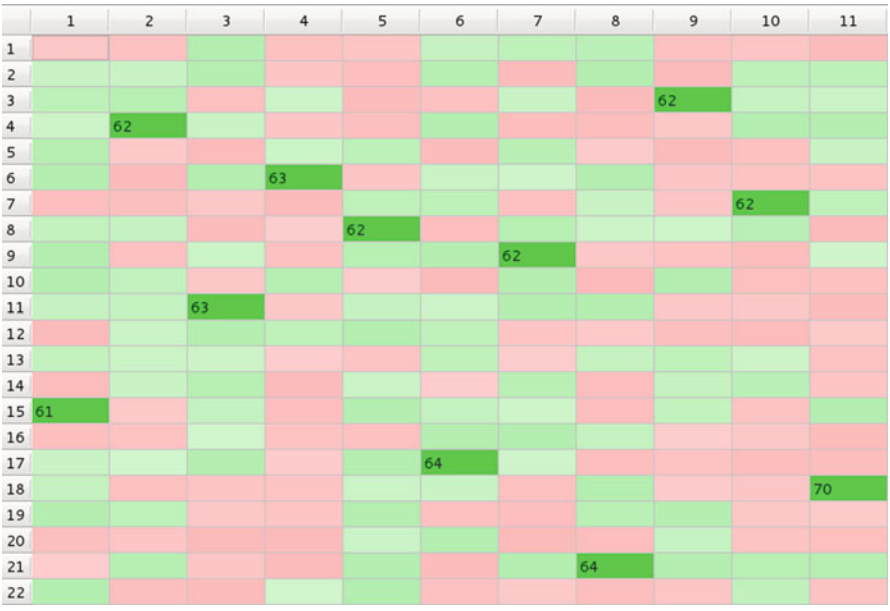


Fig. 8 Controlled GoC with feedback between subjects and the environment – without interaction between subjects (additional material: omono.mpeg)

3 Chaosensology Model

In GoC models decisions are based on emotions – a subject does not think logically, comparing 'pros' and 'cons' for the decision which subject to choose. Our Chaosensology Model (ChS) (from Latin word *sensus* – feeling, emotion) explains cognitive neurodynamics difference between logical thoughts and emotions. It is inspired by process of occupying biological 'niches' by different species – the given niche has to be vacated by one species before it may be occupied by another one. Niches are represented by lattice cells (Figs. 9 and 10). A logical thought represented by black line starting in lower-left cell that tries to move towards upper-right may be blocked because the cells in its front are occupied by randomly coiling colourful emotions (cf. [2]). We assume that τ for logical thoughts is q-times longer than for emotions, and we also change 'memory parameter' that is how long a lattice cell remains occupied (cf. Figs. 9 and 10). Similar but statical lattice models with excluded volume effect have been used e.g. for calculating entropy of mixing. Analogically, any neuron may at any given moment takes part in only one brain process and after action potential had been generated the neuron may be 'vacated' and after a certain refractory period it may be involved into another brain process.

But like any complex dynamical system human brain is characterized in any moment by momentary values of its state variables and so the brain state may be characterized by a point in a multi-dimensional phase space with appropriately defined coordinates. Then psycho-physiological processes in the brain may be represented by some trajectories in this space. Each process shows some characteristic time, τ, that defines the specific time scale for this process, i.e. time for each process should be measured as no dimensional quantity t/τ, expressed in units equal τ, rather than in 'absolute' units like seconds. Processes that are characterized by small τ are quick, these characterized by relatively large τ are slow.

Our hypothesis is [1, 2] that emotions are quick processes, while logical thoughts are relatively much slower. Quick neurodynamics processes constitute our subconsciousness while slow processes constitute our consciousness. To analyze significance of time scale difference we carry out simulations on a two-dimensional lattice (Figs. 9 and 10) and on an extremely simplified two-dimensional phase space of the brain (Fig. 11) on which logical thoughts move in horizontal direction and emotions move in vertical direction.

Optical nerve does not connect the eyes directly with visual cortex. Visual tracts pass first through limbic system that is responsible for emotions. So, emotions outside of the conscious awareness precede logical analysis of what has affected the eyes. One may say that 'we see what we feel' or that 'believing is seeing'. In our ChS model emotions always start first and logical thoughts start with delay and moves through the lattice or through the phase space on which emotions had been already moving (cf. [2]). In our model, like in reality, logical thinking may be blocked by emotions as is easily observed on Figs. 9b, c and 10. On the other

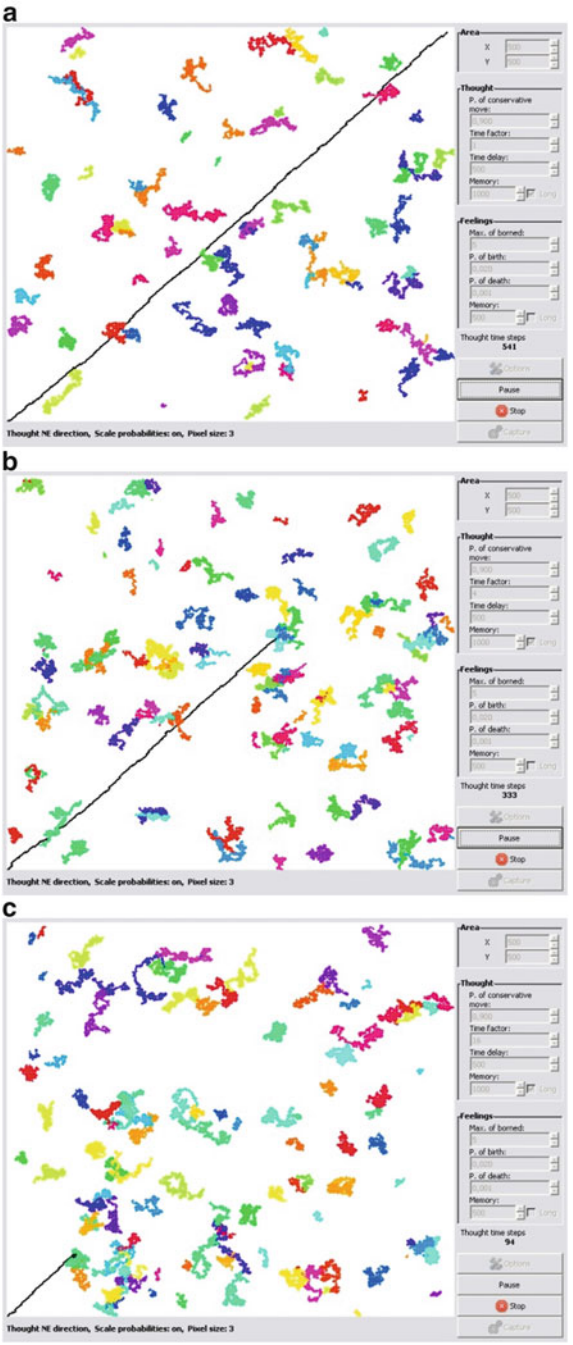


Fig. 9 Chaosensology modeling on two-dimensional lattice, short memory variant, (**a**) for q = 1; (**b**) for q = 4; (**c**) for q = 16; (additional materials: q01.avi, q04.avi, q16.avi)

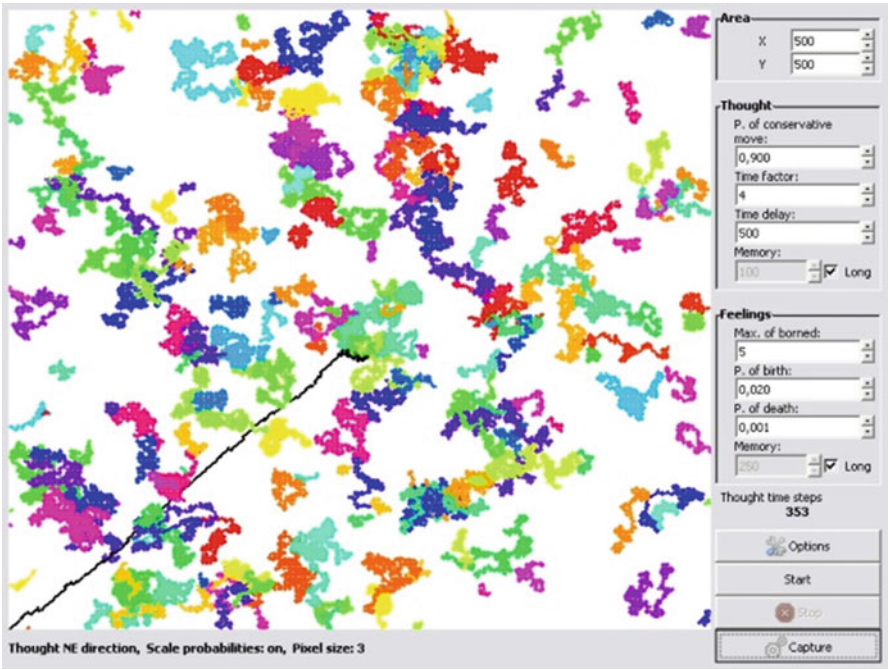


Fig. 10 Chaosensology modeling on two-dimensional lattice, q = 4, 'long memory' variant

hand, some emotions can be 'transformed into logical thoughts' and start moving in horizontal direction that is they may eventually reach conscious awareness as might be seen on Fig. 11b, c.

4 Results and Conclusions

Even the simplest GoC model with impact matrix that does not change during the whole process confirms what the father of modern psychology William James (1842–1910) said: "There's nothing so absurd that if you repeat it often enough, people will believe it". This statement was broaden and 'creatively' applied by V.I. Lenin.

Our simple simulations based on ChS model demonstrate that intensive emotions may inhibit logical thoughts and even cause a complete blockage of logical decision making [2].

Our simulations also demonstrate that there is no need for any 'central executive structure' between consciousness and subconsciousness, the existence of which were suggested by psychologists [4].

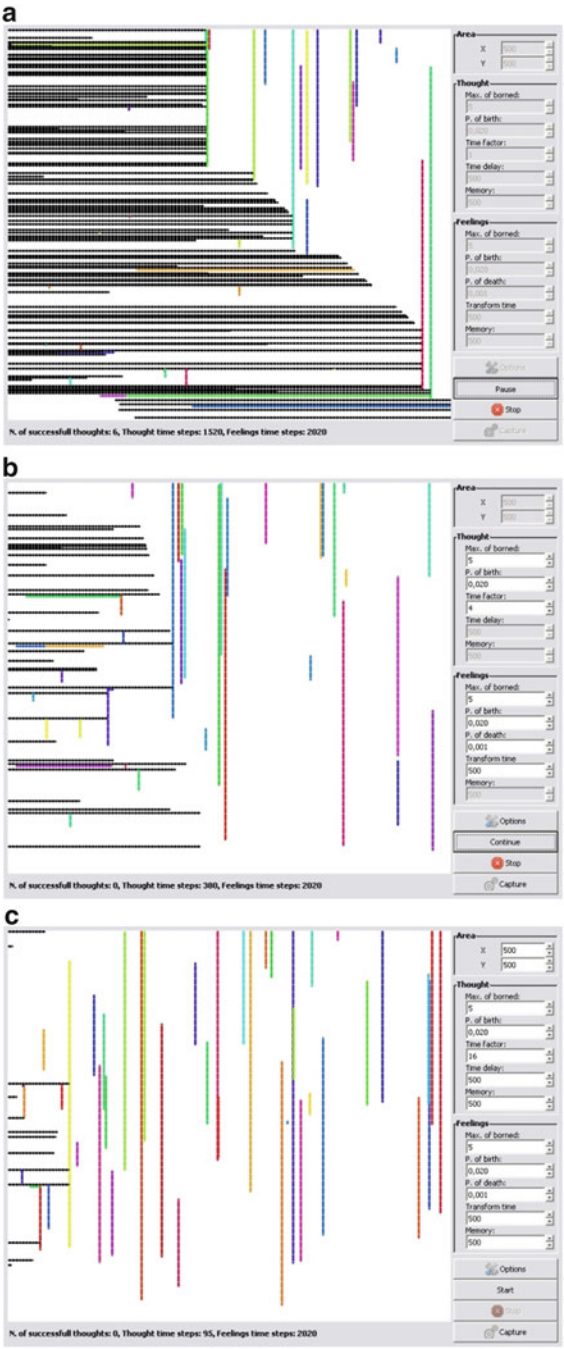


Fig. 11 Chaosensology modeling on two-dimensional phase space, (**a**) for q = 1; (**b**) for q = 4; (**c**) for q = 16; (additional materials: phase_q01.avi, phase_q04.avi, phase_q16.avi)

Acknowledgments This work was partially supported by M. Nalecz IBBE PAS through statutory project.

References

1. Klonowski W (2004) Nonlinear Dynamics in Psychophysiology – Feelings and Thoughts. In: Klonowski W (ed): Simplicity behind Complexity. Pabst Science Publishers, Lengerich, Berlin, 467–476
2. Klonowski W (2009) Significance of time scale differences in psychophysics. Cogn Process 10:S119–S126
3. Kahneman D (2011) Thinking, Fast and Slow. Farrar, Straus and Giroux, New York
4. Halligan PW, Oakley DA (2000) Greatest Myth of all. New Scientist 2265:34–39
5. Klonowski W, Pierzchalski M, Stepien P, Stepien R (2011) Econobiophysics – Choosing with Diversified Accessible Information. In: Proceedings of the Tenth International Conference on Information and Management Sciences, Gen M, Barwolff G, Li P (eds), ISSN 1539–2023, Vol.10, Series of Information and Management Sciences, Lhasa, China, August, 6–12, 2011, California Polytechnic State University, USA, pp. 328–333
6. Klonowski W, Pierzchalski M, Stepien P, Stepien R (2011) Econobiophysics – game of choosing. Model of selection or election process with diverse accessible information, Nonlinear Biomedical Physics **5:**7 http://www.nonlinearbiomedphys.com/content/5/1/7
7. Klonowski W, Pierzchalski M (2012) Influence of environment on emerging chaos in Game of Choosing models. In: Karimim H.R. (ed.): Recent Researches in Automatic Control, Systems Science and Communications, WSEAS Press, 23–31

Decision Making Mechanisms Based on Fundamental Principles of Thermodynamics

Anton P. Pakhomov

Abstract In this research, possible mechanisms of decision-making described by language of thermodynamics are considered. A three modular model of decision-making constructed from thermodynamics is offered. In this model, emotional, cognitive and mixed components in formation of the forthcoming choice are considered. As a basic function of a decision making condition entropy is considered. At the first stage (in case of emotional decision) emotion, which is comparable to energy of external system, influences closed system – emotional brain zones where fluctuations and the analysis of possible alternatives, taking into account the influence of accumulated emotional and cognitive experience, in reply to an emotion vector. At the second stage, the decision which will operate further behavior of an individual and will be transferred into environment is formed. In case of domination of a cognitive component over emotional the closed system reacts to the situation which has arisen in environment by means of the rational analysis, calculations, logics, comparing incoming information with the evolutionary accumulated experience fixed in memory. In this case emotion probably is a noise. At the following stage the created decision is transferred into the environment, then there is also an adaptation to the decision and training of the individual. In the closed system there is no fluctuations lead to chaos, bifurcation conditions are not formed and dissipative structures are not appeared as in case of emotional decision.

Keywords Decision making • Thermodynamics • Emotions • Cognition • Three modular model

A.P. Pakhomov (✉)
Decision LLC, Russian Federation, Dubininskaya str. 11/17-190, 115054 Moscow, Russia

Bando Chemical Industries Ltd. (Japan), RF, Ordzhonikidze str. 12, building 4, 119071 Moscow, Russia
e-mail: imappakhomov@gmail.com

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_9

1 Introduction

Decision making is probably one of the main types of activity, characteristic for human beings. That is why attempt to understand, explain and predict behavior of the individual making a choice became the main task of behavioral and social sciences [8]. Brain evidence complicates standard assumptions about basic preference, to include homeostasis and other kinds of state-dependence, and shows emotional activation in ambiguous choice and strategic interaction [1]. Economists generally emphasize rationality; psychologists emphasize cognitive limits and sensitivity of choices to contexts; anthropologists emphasize acculturation; and sociologists emphasize norms and social constraint. It is possible that a thermodynamics basis for behavior in neuroscience, perhaps combined with some other natural sciences standpoints, could provide some unification across the social and natural sciences [17–19]. Researchers J. Gross and R. Thompson [5] have carried out the analysis and evaluation of core features of emotion. They observed the situation – attention – appraisal – response sequence specified by the modal model of emotion. They indicate that the emotional responses generated by appraisals are thought to involve changes in experimental, behavioral and neurobiological response systems. There are several other models emphasizing the complex of choice: mechanistic models of decision making, behavioral models based on information theoretic bounds on predictive performance [2, 16]. Models include explicit decision variables that can be calculated on every trial; these quantities can then be used as proxies for the state of the subject's true internal decision variables.

2 Material and Methods

The thermodynamic investigations implemented by Nobele laureate Prigogine I. [10, 22, 23] show that any of processes occurring in the nature are irreversible. All processes are irreversible though it is possible to think up some actual processes as wished to be close to the reversible. It is impossible to return system to an initial state so that from happened event did not remain any trace. For instance, in case of fallen stone, there were following processes: its potential energy transferred at first into kinetic energy, and then into the heat dissipating in environment. The fallen stone, of course, can be lifted and put on the same place, however in the surrounded world there are indelible traces: the potential energy of the stone dissipated in the form of heat, and in our body at a rising of the stone there were various biochemical and biophysical processes.

We suppose it is apparent that proposed here model should be considered as an irreversible process as we can't say even with small probability that the system from the state II will return to the state I, without having undergone changes. Additionally, it would be reasonable to characterize system (II) considered in this research from the standpoint of thermodynamics as loop (closed) system because

the system which cannot exchange matter with environment is called closed; – the only exchange of energy is possible [10, 22, 23]. Let's consider in our case thermodynamic indicator "entropy", – which is the index of evolution. Thanks to entropy, the system in which there is initially a chaos, aspires to an equilibrium state and to the answer energetically adequate to external influence. The unit S which depends only on initial and terminating states and doesn't depend on transition from one state into another, called an entropy (from Greek – transformation), allows to describe an irreversible process as follows: $dS = dSI + dSII$; or proceeding from the record of the second law of thermodynamics offered by Clauzius in shape: $dS = \frac{\delta Q}{T} + \frac{\delta Q'}{T}$; for all irreversible changes in closed system: $dS \succ \frac{\delta Q}{T}$; entropy increment (characterizing entropy exchange of closed system with environment), in frames of our model it is possible to impress as follows: $dS = \frac{\delta Q}{T} + \delta SII$; here *dSI* is the change of system's entropy, caused by exchange of energy with environment, *dSII* is the change of entropy caused by irreversible processes in system, δQ – heat of environment, δQ′ – noncompensated heat, always positive for closed systems, T – absolute temperature. Thus, the entropy of system can be changed owing to only two reasons – either as result of transfer (transport) of entropy from environment or into environment through system borders, or as a result of entropy emergence in very system. For the closed system the stream of entropy *dS* can be both positive and negative and the entropy of system can both increase and decrease. Surely positive has to be entropy emergence *dSII* caused by changes in system, but not entropy increment *dS*. In case the system is in a condition of a thermodynamic equilibrium, the velocity of increase of entropy is equal to zero.

3 Results and Discussion

In our case emotion as the energy vector influences mentality of person, the corresponding sites of a brain being closed system. When this system starts analyzing alternatives in reply to emotion, it generally leaves equilibrium state due to fluctuations of decision searching and brings system into chaos, generating irreversible processes in closed system. After the analysis of possible alternatives spectrum is carried out and the decision is made, the entropy by an exchange with environment stabilizes closed system, returning it out of chaos, brings the zones of a brain activated by emotion into equilibrium state. As according to Kondepudi D., Prigogine I. [10] the entropy grows in irreversible processes, and in the reversible - does not change, we believe that significant increase of entropy, instead of its invariance on a phase of the alternatives choice in reply to emotion impulse, leads system to chaos and then balances it.

In the shown mechanism the possible analogy is looked through data of some biochemical investigations. In particular, the considerable heterogeneities are observed at cellular level. The potassium ion concentration in neurons is much higher, than in the extracellular environment whereas for an ion concentration of sodium the revertive situation is observed. These differences which mean the strong

disequilibrium underlie processes like innervations. Noted misbalance is supported by fissile transport of chemical components, electric and bioenergetics' reactions such as glycolysis and respiration [29].

4 Three-Modular Model of Decision Making Mechanisms Based on Thermodynamics Principles

There are several scientifically important researches of an assessment of energy efficiency in brain functioning, mechanisms of governing the evolution of cortical networks, innovative theories of informational energy coding, entropy generation [4, 13, 18, 24, 28, 30].

We will make an attempt to consider here out of principles of thermodynamics the complex theory of decision making including the following factors: emotionally and cognitively induced decisions as well as the situation when emotion and mind struggle for the choice. Let's look at the situation of emotion impact on behavior of an individual given to theoretical system, in which environmental factors (threat or usefulness levels in the context of future action / decision) are ignored. Emotion will be considered as the function of vector transfer of energy in the loop system of movement and of analysis of possible alternatives of decision-making. From the point of view of the sequence of events our model can be presented as follows: environment impacts on brain zones via emotion or/ and information (1) → analysis of alternatives or/ and cognitive estimation in brain (closed system) (2) → decision-making and adaptation to the choice (3).

At the first stage emotion which is comparable with energy of external system, influences closed system – human brain zones where fluctuations and the analysis of possible alternatives, taking into account the influence of accumulated emotional and cognitive experience, in reply to the emotion vector are going on. Though, the most probable it's seemed to be that the rational component in this situation is a noise. At the second stage the decision which will operate further behavior of an individual and will be transferred into environment is formed. Further implements adaptation to and/ or at the same time, there is to be probable further analysis of already made decision at present which will evolutionary sedimentated in reply to an affect memory in the zone of a rational choice – a dorsolateral frontal cerebral cortex (DLFCC) [14, 27]. Sometimes even at this stage made decision can be either emotionally or rationally changed (e.g. it was decided to buy a pullover which was pleasant and suitable, but on the way to the cash desk it was found another pullover which is not at all conceding to the chosen one). Than the process is coming back to one of the options of three modular model.

Extrapolating data [25] on our proposed model we receive the following integrated processes. In the environment (I) initially there is an emotional reaction or the cascade of emotional splashes, in reply to the occurred event. Emotion influences the corresponding sites of a brain and the closed system (II) in which value of a choice of this or that behavioral defined mechanism is involved in process,

the alternate options are compared, including sedimented both emotional and cognitive experience, and there is a choice of expected to be the best decision. Then process is transferred to environment (I) in which there is an adaptation to the decision and tutoring of the individual, i.e. updating of information which was stored in memory so that all subsequent actions were carried out with the greatest possible effectiveness.

In case of domination of a cognitive component over emotional the closed system reacts to the situation which has arisen in environment by means of the rational analysis, calculations, logics, comparing incoming information with the evolutionary accumulated experience fixed in memory. In this case emotion probably is a noise, a kind of background. At the following stage the created decision is transferred into environment, than also there is an adaptation to the decision and training of the individual. The cognitive analysis requiring calculations, obviously, longer while, than emotional reaction results. At the same time, in the closed system there is no fluctuations lead to chaos, bifurcation conditions are not formed and dissipative structures are not appeared. In spite of the fact that in connection with energy assumption, production of entropy *dSII* in the closed system will be still carried out, its increment *dS* will be much lower, than in case of emotionally made decision. As the decision is rigid it can hardly be reconsidered at the stage of adaptation to the made decision.

Let's consider the opportunity of decision-making in terms of competition of emotional impact and rational approach. Process of decision-making is represented generally to be similar to the way of emotional decision. The main difference is that in case of the emotional decision, the cognitive component was more noise, than in this case. Here the rational brain enters fight for leadership with emotion. Anyway ideal in this mechanism might be the situation when in a counterbalance to fight; emotion harmoniously supports the cognitive decision. From the thermodynamic point of view, as a whole, process will be proceed as the sum of two earlier described processes, and entropy increment can be expressed by the following equation:

$$dS = dSI + dSII = dSI + d\,(Se + Sr) = \frac{\delta Q}{T} + \frac{\delta Qe' + \delta Qr'}{T};$$

here *Se* - entropy produced in the closed system due to emotion, *Sr* - entropy produced in closed system due to cognitive process, $\delta Qe'$ – noncompensated heat of emotional brain, $\delta Qr'$ – noncompensated heat of rational brain. Several investigations suggest that value might be encoded by neurons in numerous brain regions, including dorsolateral prefrontal cortex [12], premotor cortex [26], frontal eye-fields [26], supplementary eye-fields [26, 27], superior colliculus [6], basal ganglia [3, 9], amygdale [20], and centromedian nucleus of the thalamus [15]. Platt M. and Padoa-Schioppa C. [21] propose that value signals expressed by different neuronal populations contribute to different mental processes. In sensory areas, value signals may contribute to perceptual attention (a process of choice between different sensory stimuli); in frontal areas, value signals may contribute to economic choice (a choice between different goods); in motor areas, value signals may

contribute to action selection (a choice between different motor acts). Miller E. and Cohen J. [14] describe that prefrontal cortex playing an important role in cognitive control, in the ability to orchestrate thought and action in accordance with internal goals. They believe that cognitive control stems from the active maintenance of patterns of activity in the prefrontal cortex that represent goals and the means to achieve them. They provide bias signals to other brain structures whose net effect is to guide the flow of activity along neural pathways that establish the proper mapping between inputs, internal states, and outputs needed to perform a given task.

In the closed system initially there is a chaos, fluctuations of possible alternatives – responses to external influence of emotion which then are ordered or/ and rational calculations in response of income information, and finally there is a choice, according to initially given entrance conditions and to the state of the closed system. The choice is carried out either spontaneously or selectively owing to sedimentation of the accumulated evolutionary experience of the individual and due to effect of emotion or cognitive approach. Other factors in our case are insignificant, in connection with the initial stipulation on ideality of system. In case of the selection choice we can assume momentary surge in entropy which as a result will balance an equilibrium state of the closed system in the form of the decision. The entropy following from the closed system, always higher than the entropy entering into the system; the difference arises because of entropy produced by irreversible processes inside the system. Our system which exchanges entropy with environment, possibly, in case of emotion or combination of emotion and mind undergoes very strong spontaneous transformations, transferring into a self-organizing mode. Order achievement through fluctuations is possible only in essentially non-linear systems (from trigger threshold processes of transfer of a nervous impulse up to evolution, irreversible development of biological systems). Such organized states are created by irreversible processes making entropy. Correspondingly, it seems that irreversible processes are that motive power which creates an order.

5 Systems of Choice Realization in Terms of Thermodynamics Knowledge

Having considered possible mechanisms of a decision making, proceeding from the principles of thermodynamics, here we would like to propose in addition to Kahneman's theory [7, 8] three main systems for discussion: System 1 which is based on an emotional priority, including moral aspect; System 2 which is characterized by rational approach to decision making and System 3 which is based on competition of affective and cognitive approaches in decision making.

The system 1 will be characterized by indeterminacy, short-time decision making which can be risky, inaccurate and unjustified. Being in emotionally excited state, the individual expecting utility of the made decision can face its real uselessness. Due to intuitively made decision the expected result hardly can provoke the true result. Probably, in evolutionary biological, financial economic and social cultural

process, decisions which analogically accepted by system 1, further become rational and will result in system 2. The system 2 will differ from system 1 with considerable determinacy, though longer-time process of decision-making as it requires the analysis and logical considering of a situation, calculations and attention, working, autobiographical, evolutionary, cognitive and some other concerned memories. As the reasonable decision is made, – it runs minimum possible risks and characterized of high degree of utility, justification of expected result. The system 3 is most complex as both emotional and rational components enter into the game. Events can develop in two ways. The first and the most optimal is when emotions elegantly accompany the cognitive decision. The second, catastrophically complex as both components enter fight for a choice. Emotions are capable to sabotage logical sense. In this system it could be reasonable to admit the assumption of LeDoux [11] concerned possibility to have a safe time due to the emotions which needs for the analysis of current situation and a probable choice of the most reasonable and adequate decision. Thermodynamic characteristics of all three systems are presented in section "Three-Modular Model of Decision-Making Based on Thermodynamic Principles".

6 Conclusions

Applying a thermodynamic approach to decision making system, author assume that emotion carries out energy function in system, thereby providing possibility of irreversible changes in system and entropy increase. Through fluctuations, being accompanied entropy increase, it is possible to observe achievement of order in systems at any level – from molecular-cellular processes to social changes at larger-scale groups. Emotions might provide one of the basic catalysts and an inductor of the energy inducing the individual to make solution of different degree of utility and polarity. In situation of the suppressed emotional hum noise the system won't transfer to other condition as there is no source of an internal energy of system. Here it is presented three-modular model of mechanisms of decision-making based on thermodynamic principles covering the following components: affective, cognitive factors and competing between them. Based on thermodynamic approach we propose in this study three systems of choice realization in support of Kahneman's theory.

References

1. Camerer C. F., Loewenstein G., Prelec D. (2004). Neuroeconomics: Why economics needs brains. Scand. J. of Economics 106 (3): 555–579.
2. Corrado G., Doya K. (2007). Understanding neural coding though the model-based analysis of decision making. J. Neurosci. 27: 8178–8180.

3. Doya K. (2008). Modulators of decision making. Nat. Neurosci. 11 (4):410–416.
4. Feldman J. (2012). The neural binding problem (s). Cogn. Neurodyn. Doi:10.1007/s11571-012-9220-2.
5. Gross J. J., Thompson R.A. (2007). Handbook of emotion regulation. London: Guilford: 3–24.
6. Ikeda T., Hikosaka O. (2003). Reward dependent gain and bias of visual responses in primate superior colliculus. Neuron 39: 693–700.
7. Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. American Psychologist, N 58: 697–720.
8. Kahneman, D., Tversky A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, N 47: 263–291.
9. Kawagoe R., Takikava Y., Hikosaka O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. Nat. Neurosci. 1: 411–416.
10. Kondepudi D., Prigogine I. (1998). Modern thermodynamics: from heat engines to dissipative structures. West Sussex: J. Willey and Sons Ltd.
11. LeDoux J.E. (1996). The Emotional Brain. NY: Simon & Schuster.
12. Lee D and X-J Wang (2008). Neural circuit mechanisms for stochastic decision making in the primate frontal cortex. In Neuroeconomics: Decision making and the brain (Glimcher PW, Camerer CF, Fehr E, and Poldrack RA, eds), Academic Press, pp 481–501.
13. Levy W.B. (1996). A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. Hippocampus 6: 579–590.
14. Miller E.K., Cohen J.D. (2001). An Integrative Theory of Prefrontal Cortex Function. Annual Reviews of Neuroscience, 24: 167–202.
15. Minamimoto T., Hori Y., Kimura M. (2005). Complimentary process to response bias in the centromedian nucleus of the thalamus. Science 308: 1798–1801.
16. O'Doherty J., Hampton A., Kim H. (2007). Model-based fMRI and its application to reward learning and decision making. Ann. NY. Acad. Sci. 1104:35–53.
17. Pakhomov A., Sudjin N. (2012). Emotional aspects of decision-making process: the thermodynamic approach. NeuroPsychoEconomics Conference Proceedings, Erasmus University, Rotterdam, p.55.
18. Pakhomov A., Sudin N. (2013). Thermodynamic view on decision making process: Emotions as a potential energy vector of realization of the choice. Cognitive Neurodynamics. DOI: 10.1007/s11571-013-9249-x.
19. Pakhomov A., Sudin N. (2013). Emotional aspects of decision making: thermodynamic approach. Experimental Psychology 3 (in press).
20. Paton J. J., Belova M. A., Morrison S. E., Salzman C. D. (2006). The primate amygdale represents the positive and negative value of visual stimuli during learning. Nature 439: 865–870.
21. Platt M., Padoa-Schioppa C. (2008). Neuronal representations of value. In Neuroeconomics: Decision making and the brain (Glimcher PW, Camerer CF, Fehr E, and Poldrack RA, eds), Academic Press, pp. 441–462.
22. Prigogine I. (1962). Introduction to thermodynamics of irreversible processes. New York, London: Wiley Interscience.
23. Prigogine I., Defay R. (1954). Chemical thermodynamics. London, New York, Toronto: Longmans Green and Co.
24. Raichle M., Gusnard D. (2002). Appraising the brain's energy budget. PNAS 99 (16): 10237–10239.
25. Rangel, A., Camerer, C., Montague, P. (2008). A framework for studying the neurobiology of value-based decision making. Nature Reviews Neuroscience, N 9: 545–556.
26. Roesch M. R., Olson C. R. (2003). Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. J. Neurophysiol. 90: 1766–1789.

27. Rosenbloom M.H., Schmahmann J.D., Price B.H. (2012). The functional neuroanatomy of decision-making. J. Neuropsychiatry Clin. Neurosci. 24: 266–277.
28. Simon B., Laughlin I., Sejnowski T.J. (2003). Communication in neural networks. Science 5641: 1870–1874.
29. Volkenstein M.V. (1988). Biophysics. Moscow: Science.
30. Wang R., Zhang Z., Chen G. (2008). Energy function and energy evolution on neural population. IEEE Trans. Neural Netw. 19 (3): 535–538.

Reward Prediction in Prefrontal Cortex and Striatum

Xiaochuan Pan, Rubin Wang, and Masamichi Sakagami

Abstract The prefrontal cortex (PFC) and striatum have mutual connections through direct and indirect pathways, and both are involved in reward prediction. But it has been suggested that the PFC and striatum may have different mechanisms in reward prediction. To understand the nature of reward process in the two areas, we recorded single-unit activity from the lateral PFC (LPFC) and striatum in monkeys performing a reward inference task. We found that prefrontal neurons could predict the reward value of a stimulus even when the monkeys had not yet learned the stimulus-reward association directly. Striatal neurons, however, could predict the reward only after directly experiencing the stimulus-reward contingency. Our results suggested dissociable functions in reward predictions: the LPFC utilized causal structure of the task or higher-order conditioning in a generative process of reward inference, whereas the striatum applied direct experiences of stimulus-reward associations in the guidance of behavior.

Keywords Prefrontal cortex • Striatum • Reward prediction • Neuron • Monkey

1 Introduction

The prefrontal cortex (PFC) and striatum are two major structures in the brain, locating at the frontal part of cerebral cortex and the forebrain, respectively. Anatomically, the PFC and striatum have mutual connections through direct and indirect pathways [1]. It is known that the PFC and striatum are involved in reward processing [2–4]. Several single-unit studies with monkeys demonstrated that PFC and striatal neurons encoded reward information related to reward type [2], amount [3] and probability [4] indicated by conditioned visual stimuli. On the other hand,

X. Pan (✉) • R. Wang
Institute for Cognitive Neurodynamics, East China University of Science and Technology, Shanghai, People's Republic of China
e-mail: pxc@ecust.edu.cn

M. Sakagami
Brain Science Institute, Tamagawa University, Tamagawagakuen 6-1-1, Machida, Tokyo 194-8610, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_10

some theories hypothesized that the PFC and striatum have different mechanisms to process reward information [5]. The striatum is thought to be involved in model-free learning to learn action values directly by trial and error, without building explicit models of the environment and task [6]. In contrast, the PFC has been shown to play an important role in model-based learning [7], such as in the encoding of abstract rules [8], in planning behaviors [9], and in dynamic decision-making [10].

To understand the nature of reward processing in the PFC and striatum, we recorded single-unit activity from the two areas in monkeys performing a reward inference task using an asymmetric reward schedule [11]. We found dissociable functions in the two areas in reward predictions. The PFC was able to infer reward value for the stimuli; but the striatum did not, instead, the striatum predicted reward based on direct experiences of stimulus-reward associations.

2 Materials and Methods

The detail description of the task can be found in [11]. Briefly, the monkeys were trained to learn two stimulus-stimulus associative sequences (Fig. 1a) in a sequential paired-association task (Fig. 1b). After that, the monkeys learned an

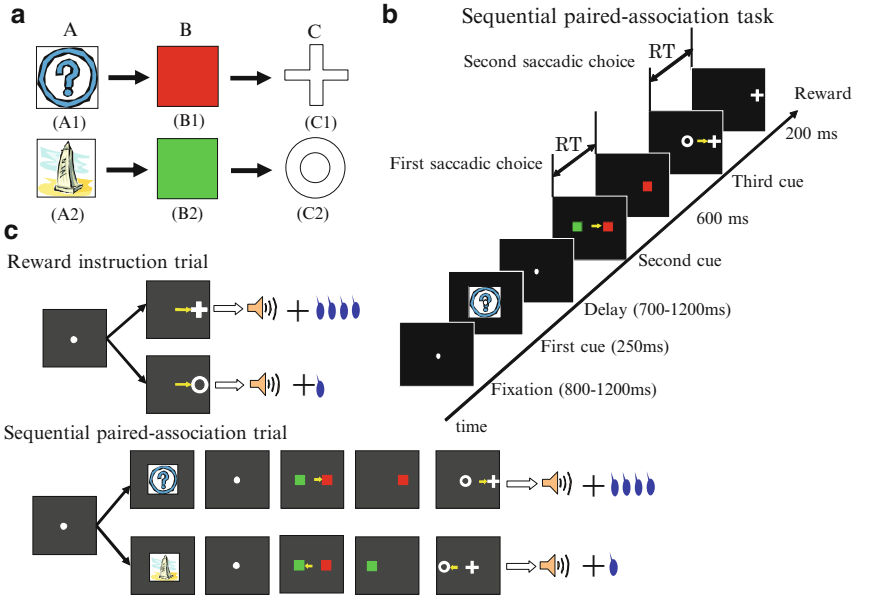


Fig. 1 The sequential paired-association task with an asymmetric reward schedule. (**a**) Two associative sequences (A1→B1→C1 and A2→B2→C2) learned by the monkey. (**b**) Schematic illustration of time events in a sequential paired-association trial (SPAT). The monkey makes a choice by a saccadic eye movement, as indicated by *small yellow arrows*. *RT* response time. (**c**) An asymmetric stimulus-reward contingency is introduced in reward instruction trials, and used in the following SPATs in one block

asymmetric reward schedule in reward instruction trials. Having fully acquired the reward instruction task and the sequential paired-association task, the monkeys were required to perform a combination of the two tasks in one block (Fig. 1c). In addition to pairing these stimuli (old stimuli) with well learned rewards, we subsequently introduced new stimuli. The monkeys learned associations between the new stimuli (e.g., N1, N2, ...) and B1 or B2 in a delayed matching-to-sample task with a symmetric reward rule.

Action potentials of single neurons were recorded extracellularly with tungsten electrodes (FHC, Bowdoinham, ME, 0.8–1.5 MΩ) from the LPFC and striatum of the monkeys. We further recorded the activity of a neuron using the new stimuli if its activity was modulated by reward using the old stimuli (A1 and A2). Once a pair of new stimuli was tested with one neuron, it could no longer be used for another neuron. We analyzed the activity of each neuron in two time periods prior to the second cues by a two-way ANOVA. The "first cue period" occurred within 100–500 ms after the first cue onset and the "early delay period" occurred within 500–900 ms after the first cue onset.

3 Results

We recorded the activity of 546 neurons from the LPFC, and the activity of 366 neurons from the striatum of the monkeys while they performed the reward-instructed SPATs with old stimuli (the first cues: A1 and A2). The activity of each neuron was analyzed using a two-way ANOVA: (stimulus (A1 or A2) vs. reward (large or small), $P < 0.01$) in the first cue and early delay periods, respectively. In this study, we focused on reward neurons that showed only a significant main effect of reward to illustrate how reward information was processed in the LPFC and striatum independently of stimulus properties.

In the LPFC we found 92 and 63 reward neurons in the first cue and early delay period, respectively. There were 113 and 44 striatal reward neurons in the first cue and early delay periods. The proportion of reward neurons in the LPFC was significantly lower than that in the striatum (28.4 % (155/546) in the LPFC; 42.9 % (157/366) in the striatum, χ^2–test, $P < 0.01$). Within these reward neurons, the activities of 106 LPFC and 100 striatal reward neurons were further recorded using the new stimuli. In the later analysis, we focused on these reward neurons whose activity was recorded using both the old and new stimuli.

We found that the majority of the 106 reward neurons in the LPFC (93/106, 87.7 %) also showed reward-type activity to the new stimuli in the first cue and/or in the early delay periods. Within those 100 striatal reward neurons, 95 neurons showed reward-modulated activity to the new stimuli in the first cue and/or in the early delay period (two-way ANOVA (stimulus vs. reward), $P < 0.05$).

We were interesting in whether the reward neurons in the LPFC and striatum could predict the reward value of the first cue stimulus (particularly for the new stimulus) presented in SPATs just after reward instruction with C1 and C2.

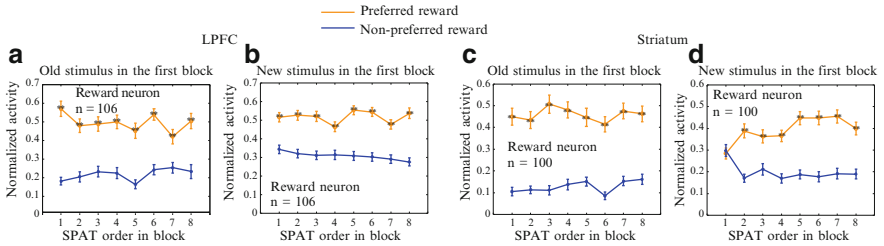


Fig. 2 Population activity of LPFC and striatal neurons as a function of SPAT order in blocks. (**a**) and (**b**) show the activity of LPFC reward neurons to old stimuli (**a**) and new stimuli (**b**). (**c**) and (**d**) show the activity of striatal reward neurons to old stimuli (**c**) and new stimuli (**d**). The normalized activity was sorted into the preferred reward condition (*orange curves*) and the non-preferred reward condition (*blue curves*). Statistical significance was determined by Mann-Whitney *U* test (** $P < 0.01$). *Error bars* indicate the s.e.m

We focused on the neural activity in the first SPAT blocks in which the new or old stimuli had been presented for the first time to each recorded neuron (Fig. 2). We found that when an old stimulus was presented as the first cue, both LPFC and striatal reward neurons discriminated the two reward conditions (large and small reward) from the first SPATs (Fig. 2a, c). However, regional differences in response activity were found when the new stimuli were presented as the first cue. LPFC reward neurons were still able to predict the reward value of the new stimuli from the first SPATs after reward instruction (Fig. 2b) despite the fact that the monkeys had never directly learned the new stimulus-reward contingency. In contrast, striatal reward neurons did not distinguish the preferred from non-preferred reward conditions in the first SPAT (Fig. 2d). After experiencing the contingency between reward and the new stimulus in the first SPAT, these neurons subsequently showed significantly differential activity in the two reward conditions from the second SPAT onwards.

4 Discussion

The current task design with new stimuli ruled out the possibility that the monkeys simply used mnemonic representations to predict reward, as they only learned associations between the new stimuli and B1 or B2 in a symmetric reward schedule. Effectively the monkeys had to integrate several independently acquired associations in order to infer the reward value for the new stimuli. The task with the new stimuli also allowed us to dissociate functions between the LPFC and striatum. LPFC and striatal neurons showed similar response patterns to those old stimuli (A1 and A2) that had been well experienced in the asymmetric task. Only when the new stimuli were presented first time in the very first SPATs, LPFC neurons still can predict reward for them, but striatal neurons cannot, indicating their different roles

in reward inference process. The LPFC might be directly involved in this reward inference process, and the striatum would be more likely engaged in rapidly learning and encoding the conditional stimulus-reward associations.

The neural activity observed in the LPFC and striatum was consistent with the predictions from the model-based and the model-free processes proposed by Daw et al. [5]. The model-based process is based on higher-order computations that allow simulations to predict outcomes using internal models, such as category being represented in the LPFC. The striatum encodes for stimulus-outcome relations through direct experiences as a result of a model-free process (e.g., temporal-difference learning). Overall, our findings supported the existence of two distinct computational strategies in the LPFC and striatum to predict reward, the former could apply inference strategy, whereas the latter utilized the direct experience of stimulus-reward contingency for guiding behavior.

Acknowledgments The work was supported by Grant-in-aid for Scientific Research on Innovative Areas, Grant-in-aid for Scientific Research (A), and Tamagawa GCOE, Japan. This work is supported by National Nature Science Foundation of China (No. 11232005), the Fundamental Research Funds for the Central Universities of China, and sponsored by Shanghai Pujiang Program (No. 13PJ1402000).

References

1. Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci 9: 357–381.
2. Watanabe M (1996) Reward expectancy in primate prefrontal neurons. *Nature* 382: 629–632.
3. Leon MI, Shadlen MN (1999) Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* 24: 415–425.
4. Samejima K, Ueda Y, Doya K, Kimura, M (2005) Representation of action-specific reward values in the striatum. *Science* 310: 1337–1340.
5. Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatum systems for behavioral control. *Nat Neurosci* 8: 1704–1711.
6. Ito M, Doya K (2011) Multiple representations and algorithms for reinforcement learning in the cortico-basal ganglia circuit. *Curr Opin Neurobiol* 21: 368–373.
7. Pan X, Sakagami M (2012) Category representation and generalization in the prefrontal cortex. *Eur J Neurosci* 35: 1083–1091.
8. White IM, Wise SP (1999) Rule-dependent neuronal activity in the prefrontal cortex. *Exp Brain Res* 126: 315–335.
9. Sakagami M, Pan X, Uttl B (2006) Behavioral inhibition and prefrontal cortex in decision-making. *Neural Netw* 19: 1255–1265.
10. Barraclough DJ, Conroy ML, Lee D (2004) Prefrontal cortex and decision making in a mixed-strategy game. *Nat Neurosci* 7: 404–410.
11. Pan X, Sawa K, Tsuda I, Tsukada M, Sakagami M (2008) Reward prediction based on stimulus categorization in primate lateral prefrontal cortex. *Nat. Neurosci. 11*, 703–712.

The Integrated Neuropsychiatric Assessment System: A Generic Platform for Cognitive Neurodynamics Research

Paul Rapp, David O. Keyser, Dominic Nathan, and Christopher J. Cellucci

Abstract The Integrated Neuropsychiatric Assessment System is a portable platform for cognitive neurodynamic research and for clinical evaluations. It provides simultaneous acquisition of heart rate variability, quantitative electroencephalography and event related potentials during a neuropsychological assessment. The analysis protocol includes measures derived from dynamical systems theory including analysis of causal networks, event-dependent interregional synchronization and small world modeling.

Keywords Neuropsychiatric diagnosis • Event related potentials • TBI • PTSD

1 Overview

This contribution reports on the continuing development of a generically applicable, physiologically based neuropsychiatric assessment system. While physiological measures constitute our present focus, an essential statement concerning the biological analysis of psychiatric illness should perhaps be made explicitly. Ultimately psychiatry is about two people facing a crisis together. Insofar as there are psychiatric cures, it is the relationship that cures. This is not an area of medical practice which will admit a dispositive quantitative measure like tumor size, blood pressure or plasma glucose concentration. As in all areas of medicine the final arbiter of psychiatric treatment is the response to the question "how are you feeling?" But while physiology, specifically psychophysiology, will never be the entire story, it

P. Rapp (✉) • D.O. Keyser
Department of Military and Emergency Medicine, Uniformed Services University, Bethesda, MD, USA
e-mail: Paul.Rapp@usuhs.edu

D. Nathan
National Intrepid Center of Excellence, Bethesda, MD, USA

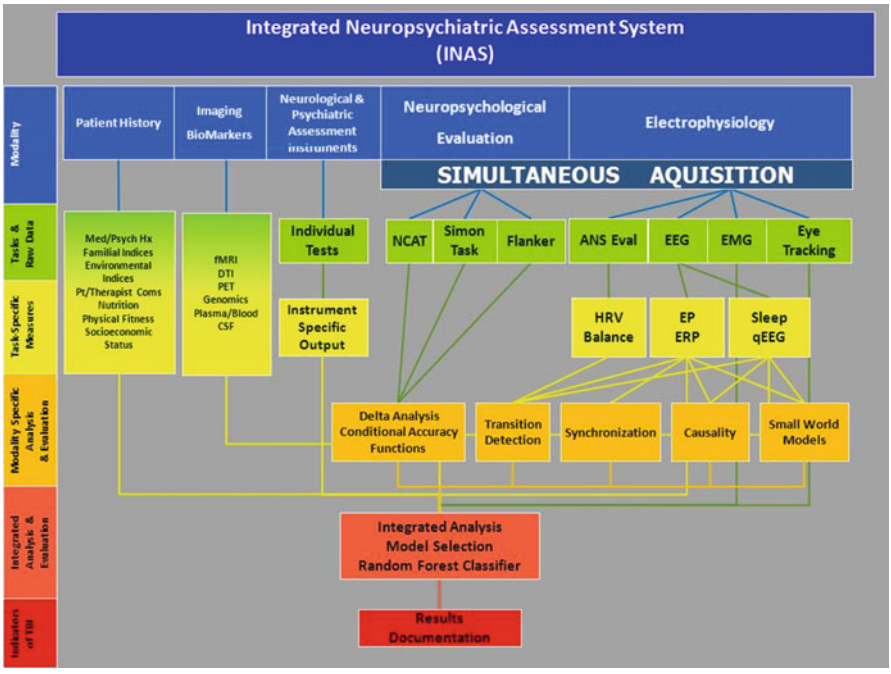
C.J. Cellucci
Aquinas LLC and the Albertus Magnus Foundation, Berwyn, PA, USA

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_11

is certainly part of the story. Our present purpose is to construct an integrated assessment technology that can address the three central questions of clinical psychophysiology.

1. Initial Assessment/Diagnosis: Can physiological measures be used to classify a patient at intake in a manner that informs treatment?
2. Longitudinal Assessment: What is the response to treatment? Do physiological variables normalize in response to treatment, where it is understood that this question is asked without prejudice as to the form of treatment be it psychopharmacological, transcranial magnetic stimulation, electroconvulsive therapy, psychotherapy or psychoanalysis?
3. Identification of Prodromes: Can psychophysiological assessments identify asymptomatic or marginally symptomatic individuals who are at risk of future presentation of a major neuropsychiatric disorder?

The system architecture is shown in the diagram. Subsystems, for example heart rate variability or neuropsychological testing, can be used in stand-alone mode. It is not necessary to utilize the entire system in every evaluation. Simultaneous signal acquisition across subsystems does, however, reduce patient burden and allows an integrated multivariate assessment. The prototype system now undergoing evaluation is portable, weighs less than 10 k and can operate on batteries. The discussion follows the system architecture diagram from left to right.



2 Patient History

The patient history is an electronic record that will be integrated with the neuropsychological and physiological data. It can be a modified version of a standard document (DoD/VA). If the system is utilized in a clinical study, the electronic history can be the FDA-negotiated Case Report Form (CRF).

3 Imaging and Biomarkers

Integrated Analysis will include biomarker data (e.g. blood/plasma components, cerebral spinal fluid components, and genomic information) and quantitative results from imaging studies when these data elements are available.

4 Neurological and Psychiatric Assessment Instruments

The Neurological and Psychiatric Assessment Instruments module contains electronic versions of standard instruments, for example Combat Exposure Scale, Mayo Classification for Traumatic Brain Injury, Rivermead Post-Concussion Symptom Scale, Short Form Health Survey, Symptom Checklist-90-Revised, Beck Depression Inventory, and PTSD Checklist [1]. There are hundreds of possibilities. Choices can be made based on the patient population and the objectives of the study. Licensing is required for proprietary instruments.

5 Neuropsychological Evaluation

The Neuropsychological Evaluation module includes standard NeuroCognitive Assessment Tests (NCAT) such as the DoD's ANAM (Automated Neuropsychological Assessment Metrics) which is a large collection of neuropsychological tests, as well as the Simon and Flanker Arrow Tasks. The INAS extends this with (1) improved timing accuracy, (2) enhanced mathematical analysis of reaction time data, (3) introduction of more demanding neuropsychological tasks, and (4) introduction of perithreshold (subliminal) stimuli. The orange analysis band of the system architecture diagram indicates the addition of conditional accuracy functions and delta analysis to the reaction time analysis section.

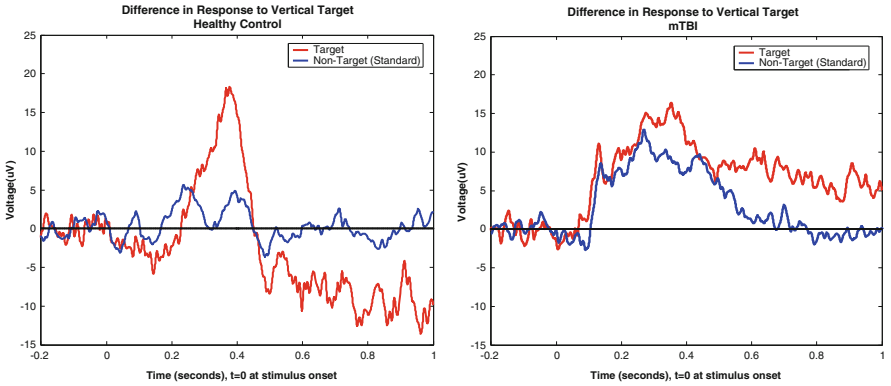
6 Electrophysiology

The electrophysiology subsystem now includes electrocardiology (heart rate variability) and electroencephalography (quantitative EEG and event related potentials). Eye tracking is scheduled for the next system iteration. The implementation of an electromyography capability is not yet in the engineering schedule.

Heart rate variability provides a noninvasive means of assessing autonomic nervous system integrity. The technology implemented here focuses on using HRV to assess psychological resilience. The HRV assessment places an emphasis on differential measures of HRV in rest and during cognitive stress Differential values (between state differences) and, most particularly, the rate of recovery from acute stress are hypothesized to be clinically disclosing when absolute values of HRV (within state values) alone are nondisclosing. In our operationalization, psychological resilience is not the absence of a response to stress. It is the ability to recover quickly from a stressful event. We argue that the derivative of a multivariate measure of heart rate variability during the recovery period will provide a physiological measure of psychological resilience generically applicable to a broad range of psychiatric presentations, and may be a prodrome of delayed onset disorders.

In the case of the event related potentials, ERPs, we're targeting the "neuropsych asymptomatic" patient who receives a clean bill of health in a conventional neuropsychological assessment but who nonetheless reports "I'm not the same." Our results suggest that an ERP study may, in some instances, make it possible to identify these individuals. Longitudinal ERP assessment may provide a quantitative mechanism for assessing response to treatment.

The results shown in the next diagrams are average ERP responses from one control participant and one TBI, traumatic brain injury, participant. While this presentation uses TBI as an example it should be noted that disease related alterations of ERPs are observed in many psychiatric populations. The TBI patient would be classed as a moderate TBI with additional orthopaedic injuries. He suffered an active duty MVA resulting in a below the knee amputation of one leg. He endorses some PTSD symptoms and a neuropsychological evaluation that we conducted showed deficits in reaction time and spatial memory. He is now in full-time, professional-level, civilian employment. Notably his current neurological examination is unremarkable. He reports, however, that neurologically "I'm not the same." The signals shown were obtained in the Brain Synchronization facility at USUHS. The test protocol has now been implemented in the INAS system and can be administered in an out-patient office visit.



Average responses to two different stimuli obtained from a control participant (left) and a clinical participant (right). Signals were recorded at scalp site Cz. The large separation between the two responses, the separation between the red and blue curves, in the case of the control participant is consistent with expectations. The absence of a separation in the case of the TBI participant is clinically significant.

The failure of response separation is indicative of CNS injury and can be observed in the absence of deficits in neuropsychological testing or in a conventional neurological examination. Additionally, these injuries are usually not observable in an imaging study. Signal separation can be quantified and used as a longitudinal measure. The EEG/ERP analysis section incorporates recently developed analysis procedures derived from dynamical systems theory. They include measures of stimulus-dependent inter-regional synchronization, analysis of CNS causal relationships and analysis of CNS networks with small world models. Alterations of these measures have been observed in psychiatric populations.

7 INAS: Integrated Analysis and Evaluation

The culmination of any combination of data element selections is represented in the red Integrated Analysis, Model Selection, Random Forest Classifier box. Here we apply non-linear analytic techniques that allow for the analysis of complex data sets. The output of this analysis is the probability that the patient's data are indistinguishable from an appropriately matched control population [2].

Acknowledgments T.R. Bashore (University of Northern Colorado), J. Kimura (Sensorium, Inc.) and B.M. Rosenberg (Thomas Jefferson University) participated in discussions concerning the design of the INAS architecture. D. Darmon (University of Maryland) participated in the development of the analysis procedures. A.M.K. Gilpin (Gilpin Biomed Research Law) provided regulatory guidance. Support has been received from the Uniformed Services University, the Marine Corps Systems Command and the Defense Medical Research and Development Program.

The opinions and assertions contained herein are the private opinions of the authors and are not to be construed as official or reflecting the views of the United States Department of Defense.

References

1. Rapp, P. E., Rosenberg, B. M., Keyser, D. O., Nathan, D., Toruno, K. M., Cellucci, C.J., Albano, A.M., Wylie, S. A., Gibson, D., Gilpin, A. M. K., and Bashore, T. R. (2013a). Patient characterization protocols for studies of traumatic brain injury and post-TBI psychiatric disorders. Frontiers in Neurology (Neurotrauma) Volume 4. Article 91, 1–34.
2. Rapp, P.E., Cellucci, C.J., Keyser, D.O., Gilpin, A.M.K. and Darmon, D.M. (2013b). Statistical issues in TBI clinical studies. Frontiers in Neurology (Neurotrauma). Volume 4. Article 177.

An Interaction Between Orbitofrontal and Rhinal Cortices Contributing to Reward Seeking Behavior

Barry J. Richmond

Abstract Monkeys given a disconnection of rhinal cortex from orbitofrontal cortex do not distinguish among different reward sizes, in a manner similar to that seen in monkeys with a bilateral rhinal cortex removal. Thus, it seems that reacting to differences among rewards requires communication between rhinal and orbitofrontal cortices. We suggest that the orbitofrontal cortex assesses value and rhinal is important for remembering the relations among the different reward sizes. This interaction provides a platform for studying information exchange across brain regions.

Keywords Reward value • Orbitofrontal cortex • Rhinal cortex • Disconnection • Monkey

The two overriding principles on which we base our interpretations of brain function are, first, that neurons and networks of neurons provide the major substrate for information processing on most time scales, and, second, that there is specialization of function in different brain areas. The consequence of these assumptions is that we expect information to be transformed as it passes from on specialized brain region to others with different, though we hope, related specializations. An historical example of how such a transformation might take place is the description of how lateral geniculate neurons with center-surround receptive fields can be aligned to give rise to the orientation specificity seen in neurons of primary visual cortex [2]. This theme of increasing specialization is followed through a sequence of visual areas into the inferior temporal cortex, where images are given their highest integrative sense in that neurons respond to complex objects such as faces and hands (Fig. 1; refs [3]). This sequence of processing in this ventral visual stream is thought to fulfill the function of giving rise to perception of complex whole objects when the signals reach the rostral end in area TE [1].

B.J. Richmond (✉)
Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Bldg 49, Rm 1B80, Bethesda, MD 20892, USA
e-mail: bjr@ln.nimh.nih.gov

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_12

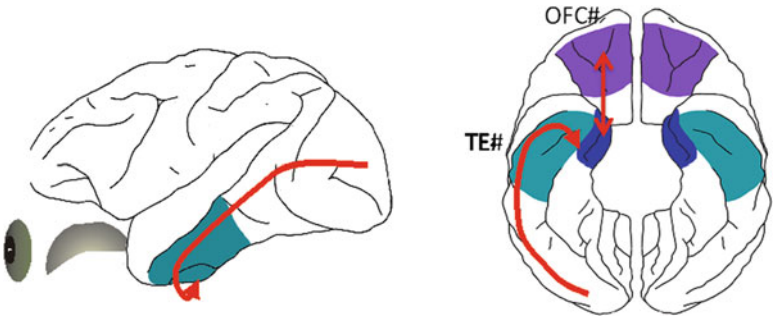


Fig. 1 *Left image*: Side view of rhesus monkey brain with eye underneath frontal lobe for orientation. The tail of the *red line* is over primary visual cortex (V1), the body overlies the visual areas in the temporal representing the cascade of the ventral visual stream, and the *arrow* head suggests that information goes underneath to medial temporal lobe structures. The turquoise region is area TE, thought to be a final step for visual pattern recognition. *Right image*: bottom view of brain, now oriented with frontal cortex at the *top* and primary visual cortex at the *bottom*. The *purple* highlights the classical orbitofrontal cortex (OFC), the *blue* rhinal cortex (Rh) which includes perirhinal and entorhinal cortices, and the turquoise the lateral inferior temporal cortex area TE. The *long red arrow* is placed to match the one on the *right image*, and the *short red arrow* indicates the existence of the reciprocal connections between OFC and Rh cortex [1]

To study how information is transformed as it passes from one brain region during reward evaluation, we have studied the contributions of the rhinal cortex in the medial temporal lobe and the orbitofrontal cortex in the monkey (Fig. 1). These regions are known to have reciprocal intrahemispheric connections, and ablations of them are followed by deficits in normal reward related behavior. Monkeys discriminate less well between rewards of different sizes after bilateral removals of orbitofrontal cortex than before the removals. After bilateral removals of rhinal cortex monkeys have a severe deficit in distinguishing among rewards also [4]. This set of results led us to ask whether the ability to distinguish among different reward sizes relied on communication between these brain regions.

We used a disconnection experiment [5]. In this study the rhinal cortex on one side and the orbitofrontal cortex on the other were removed. We then compared these monkeys to control monkeys, to monkeys with bilateral removals of rhinal cortex and to monkeys with bilateral removals orbitofrontal cortex. We used a very simple behavior in which the monkeys just touched and released a touch sensitive bar. For each set of 25 trials, a single size reward was delivered, and after each set of 25 trials, the reward side was changed randomly, with the set of rewards being 1, 2, 4, and 8 drops of juice. We measured the 'release interval', that is, the length of time from touch to release.

The control monkeys react faster for larger rewards (Fig. 2). The monkeys with bilateral OFC removals show a similar tendency, although they react faster overall, a result similar to that seen when visual cues, rather than relying on the local history (here the monkeys know what reward is forthcoming because, overall, the next reward will be the same as the preceding one). In both the case with bilateral rhinal

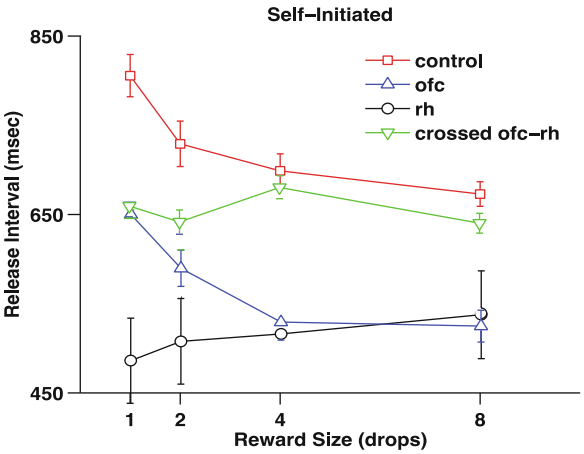


Fig. 2 Self-initiated bar pressing task. The monkey received a reward for each bar press-release sequence. The reward size was constant for 25 press-release sets, and then randomly changed to another reward size, chosen from 1, 2, 4, & 8 drop rewards. The release interval is the length of time from press to release. The results are from different sets of 4 monkeys/group. The monkeys with bilateral removal of rhinal cortex, and with the crossed rhinal-orbitofrontal removal appear to be insensitive to the difference in reward size

lesions and the rhinal-orbitofrontal disconnection, the monkeys do not change the performance with which they carry out the behavior. They act as if they unable to distinguish one reward from another.

These results seem to show that rhinal cortex must communicate with orbitofrontal for normal assessment of reward size. Our results show that the normal monkeys are sensitive to the contextual relation among the available rewards. Two processes need to occur for this to happen. First, there must be a value representation of different rewards, and second, the relative value scale needs to be remembered. Based on a considerable amount of work on OFC [6–8] and Rh we hypothesize that the OFC is critical for encoding reward value rhinal cortex must interact for the normal representation of relative reward values [9–12]. Our hope is that this OFC-Rh system provides a good substrate for studying how information is transformed as it passes between brain regions with different specialization.

Acknowledgment This work was supported by the Intralmural Research Program of the National Institute of Mental Health.

References

1. Kravitz, D.J., et al., *The ventral visual pathway: an expanded neural framework for the processing of object quality.* Trends Cogn Sci, 2013. **17**(1): p. 26–49.
2. Hubel, D.H. and T.N. Wiesel, *Receptive fields, binocular interaction and functional architecture in the cat's visual cortex.* J Physiol, 1962. **160**: p. 106–54.

3. Gross, C.G., C.E. Rocha-Miranda, and D.B. Bender, *Visual properties of neurons in inferotemporal cortex of the Macaque.* J Neurophysiol, 1972. **35**(1): p. 96–111.
4. Clark, A.M., et al., *Intersection of reward and memory in monkey rhinal cortex.* J Neurosci, 2012. **32**(20): p. 6869–77.
5. Clark, A.M., et al., *Interaction between orbital prefrontal and rhinal cortex is required for normal estimates of expected value.* J Neurosci, 2013. **33**(5): p. 1833–45.
6. Kobayashi, S., O. Pinto de Carvalho, and W. Schultz, *Adaptation of reward sensitivity in orbitofrontal neurons.* J Neurosci, 2010. **30**(2): p. 534–44.
7. Padoa-Schioppa, C., *Range-adapting representation of economic value in the orbitofrontal cortex.* J Neurosci, 2009. **29**(44): p. 14004–14.
8. Simmons, J.M. and B.J. Richmond, *Dynamic changes in representations of preceding and upcoming reward in monkey orbitofrontal cortex.* Cereb Cortex, 2008. **18**(1): p. 93–103.
9. Liu, Z., E.A. Murray, and B.J. Richmond, *Learning motivational significance of visual cues for reward schedules requires rhinal cortex.* Nat Neurosci, 2000. **3**(12): p. 1307–15.
10. Liu, Z. and B.J. Richmond, *Response differences in monkey TE and perirhinal cortex: stimulus association related to reward schedules.* J Neurophysiol, 2000. **83**(3): p. 1677–92.
11. Squire, L.R., J.T. Wixted, and R.E. Clark, *Recognition memory and the medial temporal lobe: a new perspective.* Nat Rev Neurosci, 2007. **8**(11): p. 872–83.
12. Suzuki, W.A., *Perception and the medial temporal lobe: evaluating the current evidence.* Neuron, 2009. **61**(5): p. 657–66.

Exploring Dynamic Temporal-Topological Structure of Brain Network Within ADHD

Rong Wang, Pan Lin, and Ying Wu

Abstract The **b**rain is a complex network with an anatomical and functional organization. The brain organization of attention-deficit/hyperactivity disorder (ADHD) is still not well understood. We used complex network method to investigate ADHD subject's brain network during resting state. Our results show that the node degree, clustering coefficient, local efficiency and global efficiency dynamically evolve during resting state, modularity, and the results also show the stable difference in brain network topological structure of ADHD subjects compared with normal subjects. These results would suggest the important role of the dynamic temporal-topological structure of ADHD linking to underlying dysfunctional neuronal activity.

Keywords ADHD • Resting-state functional connectivity • fMRI • Complex networks

1 Introduction

Attention-deficit/hyperactivity disorder (ADHD) is one of the most commonly diagnosed childhood neuropsychiatry and characterized by inattention, hyperactivity and impulsivity. Children with ADHD are found easily to be distracted, difficultly to focus on one task, and constantly to be in motion and impatient, which all affect their academic performance and social life.

R. Wang
State Key Laboratory for Strength and Vibration of Mechanical Structures, School of Aerospace, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China

P. Lin
Key Laboratory of Biomedical Information Engineering of Education Ministry, Institute of Biomedical Engineering, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China

Y. Wu (✉)
School of Aerospace, Xi'an Jiaotong University, Xi'an, People's Republic of China
e-mail: wying36@163.com; wying36@mail.xjtu.edu.cn

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_13

Brain is a complex dynamic system [1, 2]. Recent developments in neuroimaging have revealed the complex brain organizing of large-scale network. A large-scale brain functional network shapes brain function and facilitates cognitive processing. Numerous studies have demonstrated that functional brain complex network properties could be affected by aging and brain diseases such as schizophrenia and Alzheimer's disease. Understanding the brain network organization could facilitate better diagnosis brain disorder and guide treatment for psychiatric disorder.

More important, complex network analysis approaches have recently been shown to be sensitive to pathology in ADHD. For example, one used graph theoretical approaches based on resting state fMRI data to explore the brain networks of ADHD patients. This study suggests that the brain network topological properties of ADHD patients have been altered. Other studies also found a loss of the optimal organization pattern in ADHD [3, 4].

However, what remains unclear is how the brain network organizes dynamic configuration across resting state within ADHD. To better understand dynamic temporal topological structure of brain network of ADHD, we used complex network based on different sliding time-windows to investigate ADHD subject's brain dynamic network structure during resting state. Our results reveal that the brain networks of ADHD subjects are reorganized compared with normal subjects in brain functional networks. The results would show the important role of the dynamic temporal-topological structure of ADHD brain network linking to underlying dysfunctional neuronal activity. Our results can be a sensitive and specific biomarker of ADHD.

2 Materials and Methods

fMRI data was extracted from the open-access '1,000 Functional Connectomes Project' (http://fcon_1000.projects.nitrc.org/) in which resting-state fMRI scans have released by M.P. Milham and F.X. Castellanos at December, 2009. These data were acquired at resting state by 3 T Siemens scanner. We selected 24 ADHD subjects from group A, 19 male and 5 female, with the mean age of 34.87, and then we selected 24 normal controls from group B, 18 male and 6 female, with the mean age of 34.65. The TR is 2 s. The functional images were preprocessed using a combination of analysis of fMRI data based on AFNI (http://afni.nimh.nih.gov/afni/) and FSL's software Library (http://www.fmrib.ox.ac.uk/fsl/). EPIs were motion and slice-time corrected, and spatially smoothed using a Gaussian kernel of 6 mm FWHM. The temporal band-pass filtering (0.005 Hz $<$ f $<$0.1 Hz) was performed in order to reduce the effects of low-frequency drift and high-frequency physiological noise. After eliminating redundant information of CSF and white matter, fMRI data were further spatially normalized to the Montreal Neurological Institute (MNI) EPI template and resampled to a 3 mm cubic voxel. In the present study, we constructed brain functional networks using automated anatomical labeling (AAL) atlas to divide the brain into 108 regions of interest (ROIs). For each participant resting state fMRI data, ROIs mean time series were calculated by taking the mean of the voxel time series within each region.

The dynamic correlation coefficients were computed between each pair of brain regions for each subject based on a sliding time-window and sliding step of one TR, and then dynamic correlation matrix based on a sliding time-window for each subject was obtained. Correlation coefficient matrix represents the functional connectivity strength between the each pair of brain regions. For further statistical analysis, a Fisher's r-to-z transformation was applied to improve the normality of the correlation coefficients. Then we characterized the dynamic temporal-topological structure of whole brain network for each sliding time-window by using complex network analysis method based on BCT Matlab toolbox (http://www.brain-connectivity-toolbox.net]. To date, most brain network studies have investigated topological properties using binarized graph. The related complex network topological measures include degree, global efficiency and local efficiency, modularity, and clustering coefficient, which are defined seeing BCT Matlab toolbox reference paper [5].

3 Results and Discussion

Through the global brain functional network with 108 nodes, we investigated the differences of global brain network topological properties between ADHD and control groups for given costs at different sliding time-windows.

To assess the differences of brain network topological properties between ADHD and control groups, statistical comparisons of clustering coefficient,local efficiency, modularity and other topological properties were performed by using *t*-test based on different sliding time-windows for each value over a wide range of cost. The results are shown in Figs. 1 and 2.

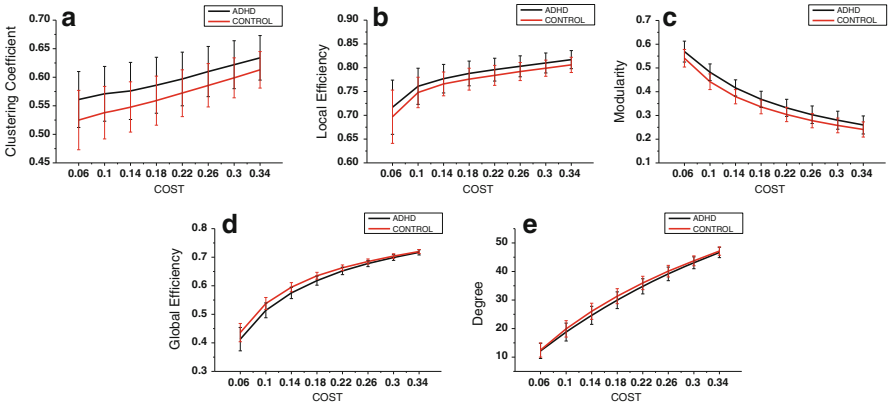


Fig. 1 Topological properties of whole brain network metrics of each group using 3 min sliding time-window. (**a**) Clustering coefficient (**b**) local efficiency (**c**) modularity (**d**) global efficiency (**e**) degree of whole brain network for ADHD and control groups as a function of cost

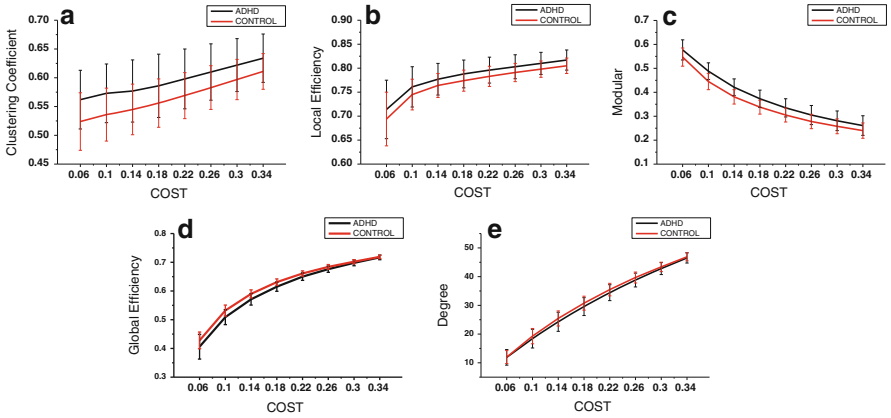


Fig. 2 Topological properties of whole brain network metrics of each group using 1 min sliding time-window. (**a**) Clustering coefficient (**b**) local efficiency (**c**) modularity (**d**) global efficiency (**e**) degree of whole brain network for ADHD and control groups as a function of cost

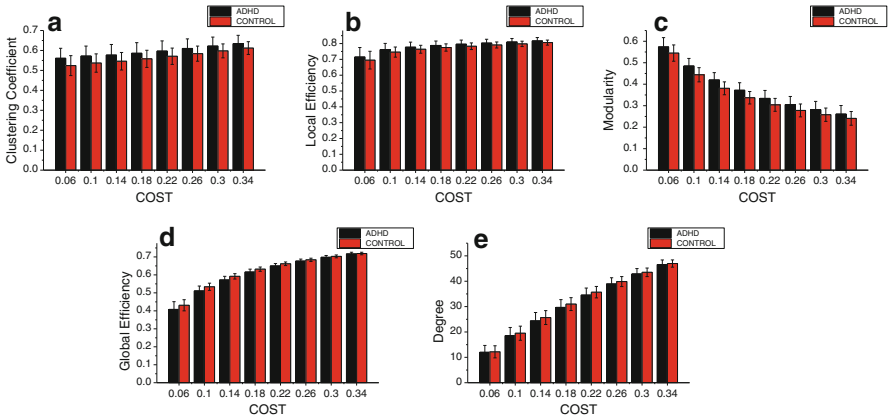


Fig. 3 Topological properties of whole brain network metrics of each group for all sliding time-windows. (**a**) Clustering coefficient (**b**) local efficiency (**c**) modularity (**d**) global efficiency (**e**) degree of whole brain network for ADHD and control groups as a function of cost

In addition, in order to study the difference of overall properties between two groups, we calculated the characteristic coefficients within the all sliding time-windows at each value over a wide range of cost. Figure 3 shows that, from the overall level of sliding time-window, the clustering coefficient, local efficiency and modularity of ADHD group are larger than that of control group for the costs, but it is converse to global efficiency and degree.

The clustering coefficient, local efficiency and modularity of ADHD group are always larger than that of control group, and global efficiency and degree of ADHD group are always smaller than that of control group at different widths of sliding time-window or all sliding time-windows for each value over a wide range of cost.

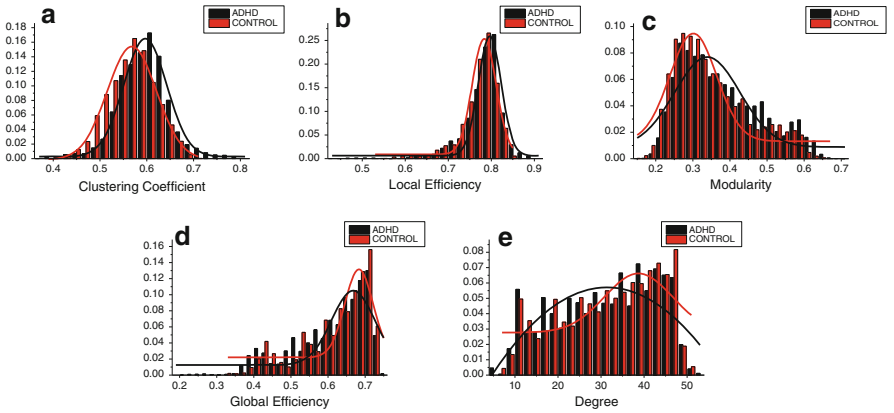


Fig. 4 The distribution of dynamic temporal-topological properties during time-windows sliding for all costs. (**a**) Clustering coefficient (**b**) local efficiency (**c**) modularity (**d**) global efficiency (**e**) degree for all time-windows and costs

Figure 4 illustrates the working scope of brain during experiment with a view point of complex network. The result show, compared with normal group, ADHD group has larger clustering coefficient, local efficiency, and modularity corresponding to the max-number which correspond to the most common working state of brain. It indicates the degree of ADHD patient's brain network are easier to gather together, the ability of information transmission at local ADHD patient's brain regions increases, and the sub-networks are more likely to be formed within ADHD patient's brain network. However, ADHD group has smaller degree and global efficiency, which means the ability of information transmission at global brain network decrease and the brain functional network shifts toward to sparse network in ADHD group compared with normal group.

Taken together, we find the stable differences in characteristic coefficients of dynamic brain networks of ADHD group compared with normal group. It is found that the clustering coefficient, local efficiency and modularity of ADHD group are always larger than that of control group, and global efficiency and degree of ADHD group are always smaller than that of control group for different sliding time-windows at given costs. These results suggest that ADHD brain functional network architecture has obviously decreased brain network integration and increased brain network segregation. Our results reveal that the important role of the dynamic temporal-topological structure of ADHD brain network linking to underlying dysfunctional neuronal activity.

Acknowledgments This work was supported by the National Natural Science Foundation of China (Nos. 11272242 60963012, 61262034,), by the Fundamental Research Funds for the Central Universities of China, by the Key Project of Chinese Ministry of Education (No. 211087), by Doctoral Fund of Ministry of Education of China (20120201120071), by the Fundamental Research Funds for the Central Universities of China, and by the New Faculty Research Foundation of XJTU.

References

1. Bullmore, E. and O. Sporns (2009). “Complex brain networks: graph theoretical analysis of structural and functional systems.” Nat Rev Neurosci 10(3): 186–198.
2. Allen, E. A. Damaraju, E. Plis, S. M.Erhardt, E. B. Eichele, T. Calhoun, V. D.“ Tracking Whole-Brain Connectivity Dynamics in the Resting State”, Cereb Cortex, in press.
3. Wang, L., C. Zhu, et al. (2009). “Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder.” Hum Brain Mapp **30**(2): 638–649.
4. Konrad, K. and S. B. Eickhoff (2010). “Is the ADHD Brain Wired Differently? A Review on Structural and Functional Connectivity in Attention Deficit Hyperactivity Disorder.” Human Brain Mapping **31**(6): 904–916.
5. Rubinov, M. and O. Sporns (2010). “Complex network measures of brain connectivity: uses and interpretations.” Neuroimage 52(3): 1059–1069.

Study on the EEG Rhythm in Meditation

Tinglin Zhang, Ruxiu Liu, Chungang Shang, Ruifen Hu, Hans Liljenström, and Guang Li

Abstract Meditation affects the brain rhythm significantly. Compared with non-meditators, the power of delta band was lower while high frequency band was higher for a meditator. Alpha band over the scalp was much more active in normal state for meditator with decreased dominant alpha frequency. Obvious transient process between normal eyes-closed rest and meditation was observed after EEG analysis. The active time and the power of beta and gamma band increased significantly in meditation. The inter-hemispheric and intra-hemispheric coherence beta and low gamma bands for meditative state were higher than normal state.

Keywords Meditation • EEG • Energy ratio • Topographic map • Coherence

1 Introduction

Meditation is a kind of mental training to turn into the "inner self" with conscious of the surroundings [1]. The individual may find inner peace and explore a particular resource that lies possibly within the subconscious to get an improved sense of well-being during the meditation. Scientific exploration confirmed the effectiveness of meditation on health promotion, anxiety reduction and stress manipulation, etc. [2, 3].

EEG is an important tool for detecting and monitoring the nervous system clinically. EEG of different meditating techniques has been researched since 1960s [4], in which Zen meditation, Transcendental Meditation (TM), Qigong and yogic meditation were the most common techniques studied [3–8]. However, it remains

T. Zhang (✉) • R. Liu • C. Shang • R. Hu • G. Li
Department of Control Science and Engineering, State Key Lab of Industrial Control Technology, Zhejiang University, 38 Zheda Road, Hangzhou 310027, People's Republic of China
e-mail: ztlyoyo@163.com

H. Liljenström
Department of Energy and Analysis, SLU, Biometry and Systems Analysis, P.O. Box 7032, SE-750 07 Uppsala, Sweden

Agora for Biosystems, P.O. Box 57, SE-193 22 Sigtuna, Sweden

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_14

an open question as different meditation techniques display diverse EEG patterns, and even for the same meditation technique the distinction between meditation and eyes-closed rest is still controversial [6, 7]. Some EEG patterns under meditation were observed in the meditators such as increased [4, 6] or blocked [5] alpha activity, or a burst of theta [6, 7] waves which differed from eyes-closed rest EEG.

The unique meditation technique in this study contains four different stages and obvious transient process between normal eyes-closed rest and meditation. The aim of this paper is to study effects of this kind of meditation on brain dynamics and rhythms as well as the difference between normal state and meditative state.

2 Methods

2.1 Experiments and EEG Recording

The study subject was a 64-year-old man with 42 years practice in meditation. The meditator could go into four different phases during meditation: first three short phases (phase1, phase2, and phase3), each sustaining less than 1 min, and the fourth phase (phase4) sustaining more than 15 min. The subject focused on different imagined images in the first three phases which could help him to search the unconscious resource and reach subconscious state finally in the fourth phase. Four males and one female, 24–26 years old, without any experience in meditation served as control group. All the subjects were right-handed.

The meditator should keep relaxed with eyes closed for 3 min before meditation. He was asked to press a button to mark the time when perceived himself entered a new phase during the meditation. EEG was also recorded for 5 min after the meditator came back to normal resting state. For the control group, the subjects were asked to keep relaxed with eyes closed for 20 min with EEG recorded. Five trials meditation experiments were performed by the meditator and five trials resting experiments were accomplished by the control group.

Neuroscan SynAmps2 Digital Amplifier with a 64-electrode cap was applied for EEG recording. Figure 1 shows the locations of 62 electrodes according to the International 10–20 System. The signal was sampled at 1,000 Hz, and referenced to the left mastoid with an analog filtering pass-band of 0.05–100 Hz.

2.2 Signal Processing

The raw EEG data were digitally referenced to the linked mastoids [9]. EEG with significant myoelectric artifacts was manually excluded. FFT, calculated in MATLAB software, was applied for time-frequency analysis on the EEG with 1,024 ms time window and sliding 512 ms. Power on each channel was computed by FFT for 1-min time window in each state. Mean energy ratio of five trials, energy of different frequency bands to total energy, for the meditator (the study subject)

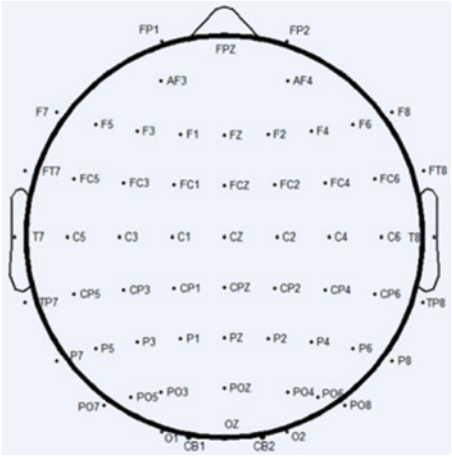


Fig. 1 Locations of the 62 electrodes

and non-meditators (control group) were compared. Topographic maps represented the power distribution of EEG over the scalp were plotted by using interpolation on a fine Cartesian grid. Mean coherence at different frequency for four pairs, F3-F4, P3-P4, F3-P3, and F4-P4 in 1 min EEG for five trials was estimated by a normalized cross spectral density function [8] implemented by function "mscohere" in MATLAB. Coherence coefficient, which varied between 0 and 1, increased with high coherence.

Meditative state is considered to be similar to a stage of sleep or drowsiness. For comparison, sleep EEG from ten male adults was recorded with the Neuroscan NuAmps Digital Amplifier on Fz channel, sampled at 250 Hz and referenced to linked mastoids followed by a band-pass filtering between 0.5 and 30 Hz. Meditation EEG on Fz channel was down-sampled to 250 Hz and band-pass filtered from 0.5 to 30 Hz before the power distribution of meditation and sleep in the low frequency band (<30 Hz) was compared.

3 Results

3.1 Power Distribution Analysis

The EEG of study subject (meditator) in normal resting state before meditation was different from control group (non-meditators). The power of delta band (<4 Hz) was lower while high frequency band (4–80 Hz) was higher for meditator than non-meditators as shown in Figs. 2 and 4.

Alpha band (8–13 Hz) over the scalp was much more active for meditator than non-meditator in normal state and dominant alpha frequency decreased in meditator which was displayed in Fig. 3 and topographic maps Fig. 5. For normal resting state, EEG of meditator was dominated with parietal occipital alpha compared with

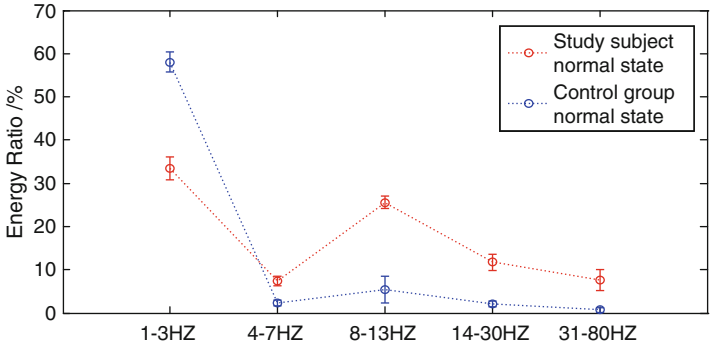


Fig. 2 Average energy ratio in different frequency bands on channel CZ for 1 min in normal state of the meditator and non-meditators

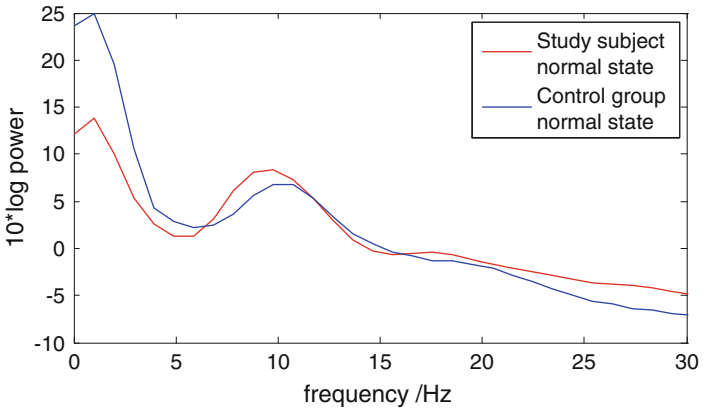


Fig. 3 Average power of EEG recorded from all channels for 1 min in normal state of the meditator compared with non-meditator

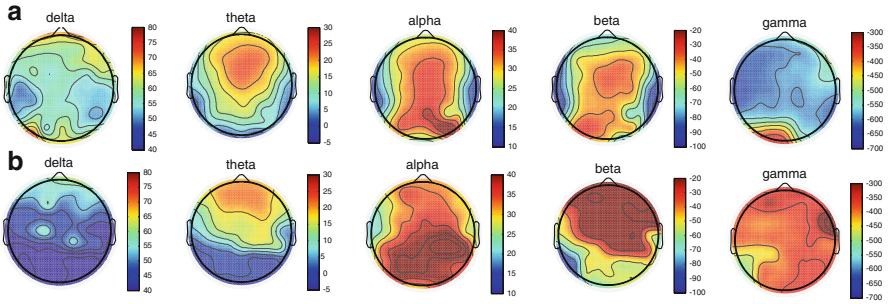


Fig. 4 Topographic map of PSD of delta (1–3 Hz), theta (4–7 Hz), alpha (8–13 Hz), beta (14–30 Hz) and low gamma band (31–80 Hz) respectively in normal state of (**a**) control group and (**b**) the study subject

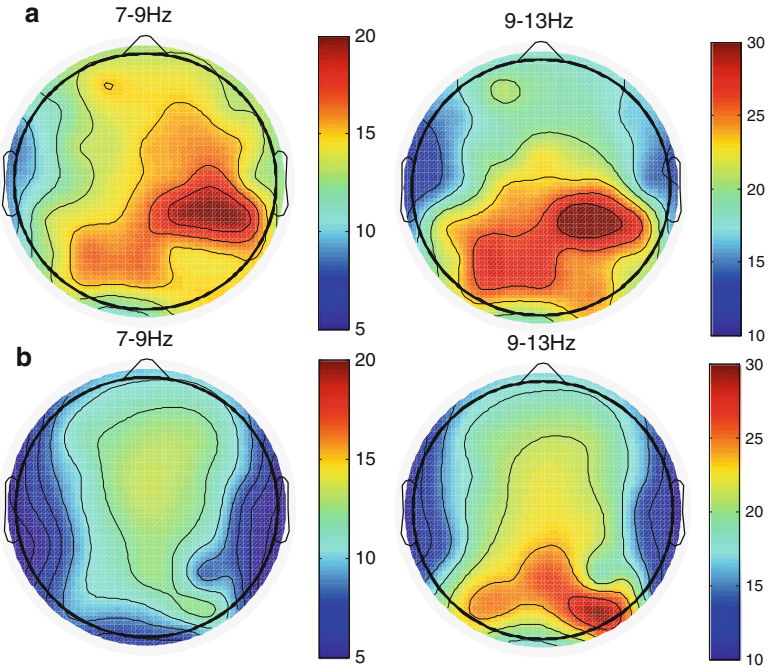


Fig. 5 Topographic comparison of mean alpha power over the scalp in 1 min normal state of (**a**) meditator and (**b**) control group for five trials

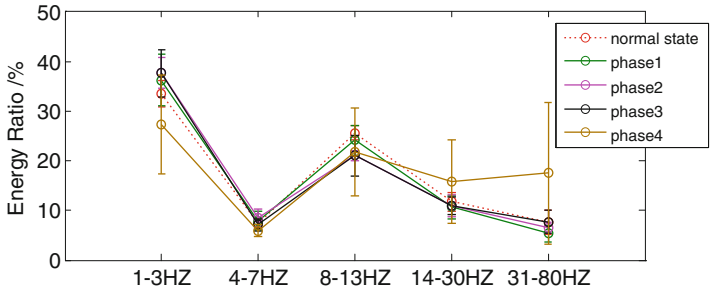


Fig. 6 Average energy ratio in different bands of normal state and meditative state of the meditator. EEG was recorded by channel CZ for 1 min

midline frontal (7–9 Hz), parietal and occipital (9–13 Hz) alpha dominance in the control group. Active alpha rhythm was correlated with feeling of calm and positive affect in early biofeedback studies [10, 11]. Therefore, the meditator was much easier to get this positive state than non-meditator.

The active time and power of the alpha band (8–13 Hz) decreased, while the power of delta-theta rhythm (<8 Hz) increased in the first three phases of meditation, followed by significantly increased gamma activity in the fourth phase according to time-frequency analysis (Fig. 6). Phase4 was apparently different from

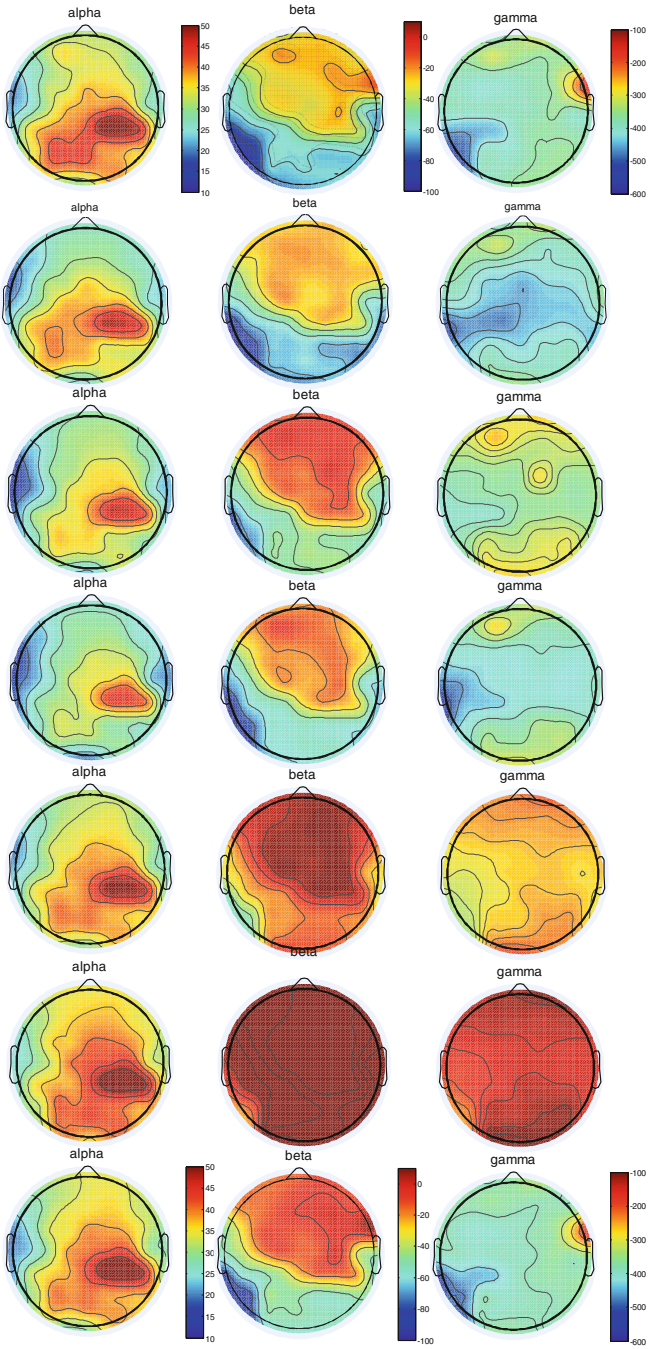


Fig. 7 Topographic display of power distribution in the first minute of (**a**) normal state before meditation, (**b**) phase1, (**c**) phase2, (**d**) phase3 and (**e**) phase4, the last minute of (**f**) phase4 and first minute of (**g**) normal state after meditation

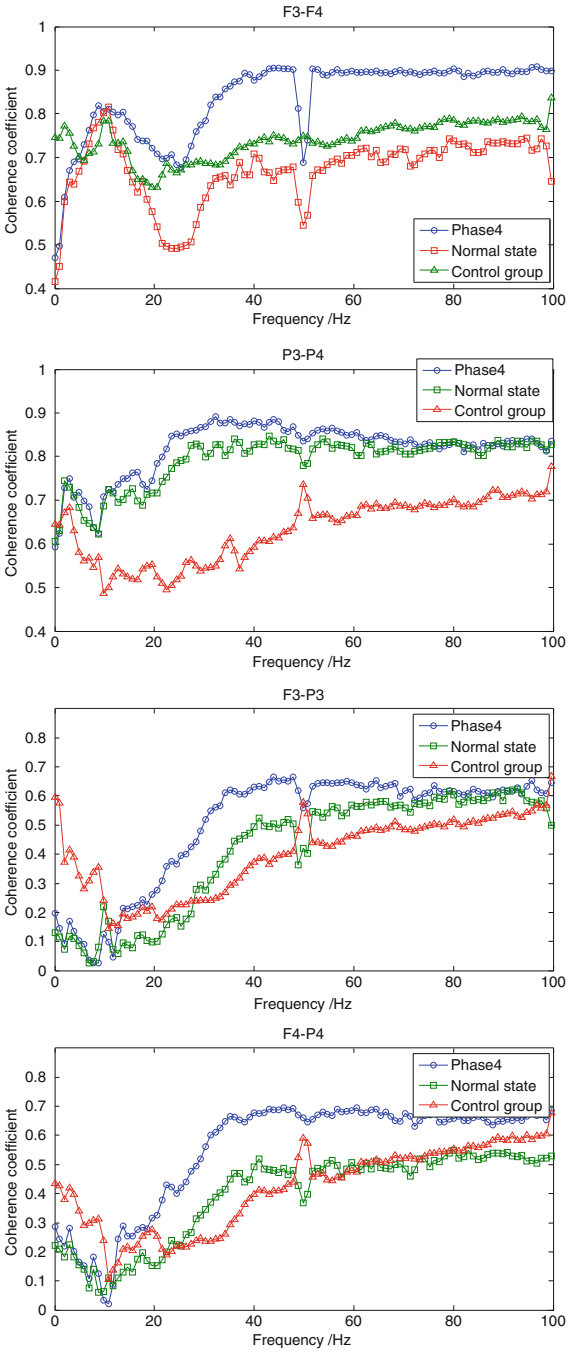


Fig. 8 Mean coherence at different frequency for four pairs, F3-F4, P3-P4, F3-P3, and F4-P4 in 1 min EEG for five trials. Coherence was higher in phase4 than normal state either for meditator or control group

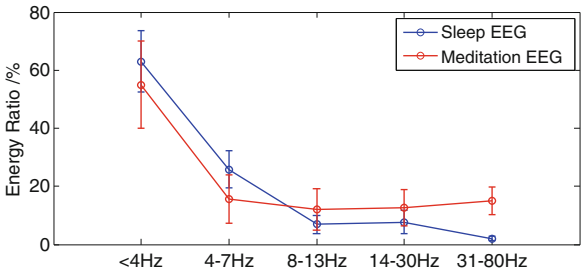


Fig. 9 Average energy ratio in different frequency bands on channel FZ. EEG was got from light sleep and meditation for 1 min

normal state and other meditative states as the first three phases were searching process and the fourth phase was final subconscious state. The increased power in gamma band might be associated with thinking activities of the meditator. EEG rhythm of the meditator could quickly return to normal state after meditation. The dynamic quantitative power variation of EEG during the whole meditation experiment could see in Fig. 7. In contrast, the EEG power maintained stable during the whole resting experiment for control group.

3.2 Coherence Analysis

The inter-hemispheric coherence was higher for frontal pairs F3-F4 as well as parietal occipital pairs P3-P4, when comparing normal with meditative state phase4 in beta and low gamma band. Large coherence coefficients of F3-P3 and F4-P4 represented the fronto-parietal intra-hemispheric coherence was also high in phase4. Coherence was higher in phase4 than normal state either for meditator or control group (Fig. 8).

Meditation was discriminated from sleep. Power distribution of meditation EEG in the low frequency band was quite different from light sleep, during which the dominant delta-theta rhythm took over. Low frequency band of sleep EEG was more active than meditative EEG, while high frequency band had less activity (Fig. 9). Furthermore, alpha coherence decreased during sleep [12], however, it increased in meditation.

4 Conclusions

Meditation has a great effect on the brain rhythms, which varies with different phases. Hence, it seems possible to study the effect of meditation on mental function. High coherence in beta and low gamma bands for meditation gives the inspiration that phase synchronization over the scalp could be tested for future study. More work is necessary on classifying normal and different meditative states or phases, which might provide a valid indicator for meditation practice.

References

1. Buzsaki, G.: Rhythms of the Brain. Oxford University Press. (2006) 214–215
2. Orme-Johnson, D.: Evidence that the transcendental meditation program prevents or decreases diseases of the nervous system and is specifically beneficial for epilepsy. Medical Hypotheses. 67 (2006) 240–246.
3. Aftanas, L.I. & Golocheikine, S.A.: Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalized attention: high-resolution EEG investigation of meditation. Neuroscience Letters. 310 (2001), 57–60.
4. Anand, B. K., Chhina, G. S., & Singh, B.: Some Aspects of Electroencephalographic Studies in Yogis. Electroencephalography and Clinical Neurophysiology. 13 (1961) 452–456.
5. Lo, P. C., Huang, M. L. & Chang K. M.: EEG alpha blocking correlated with perception of inner light during Zen meditation. The American Journal of Chinese Medicine. 31.4 (2003) 629–642
6. Travis, F.: Autonomic and EEG patterns distinguish transcending from other experiences during transcendental meditation practice. International Journal of Psychophysiology. 42 (2001), 1–9.
7. Pagano, R. R., & Warrenburg, S.: Meditation: In search of a unique effect. New York: Plenum Press. 3 (1983) 152–210.
8. Qin, Z., Jin, Y. et al: A Forty-five Year Follow-up EEG Study of Qigong Practice. International Journal of Neuroscience. 119.4 (2009) 538–52.
9. Nunez, P. L., Srinivasan, R., Westdorp, A. F., Wijesinghe, R. S., Tucker, D. M., Silberstein, R. B., & Cadusch, P. J.: EEG coherency I: Statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales. Electroencephalography and Clinical Neurophysiology. 103.5 (1997) 499–515.
10. Kamiya, J.: Operant control of the EEG alpha rhythm and some of its reported effects on consciousness. In C. T. Tart (Ed.), Altered states of consciousness (1969) 519–529. New York: Wiley.
11. Hardt, J. V., & Kamiya, J.: Anxiety change through electroencephalographic alpha feedback seen only in high anxiety subjects. Science. 201 (1978) 79–81.
12. Cantero, J. L., Atienza, M., Salas, R.M., & Gomez, C.M.: Alpha EEG coherence in different brain states: An electrophysiological index of the arousal level in human subjects. Neuroscience Letters. 271 (1999) 167–170.

Part III
Cortical Dynamics in Perception and Cognition

Set-Related Neurocognitive Networks

Steven L. Bressler

Abstract The term "set" refers to the anticipatory neurocognitive processes that prepare an individual to engage in a particular behavior. Set entails a specific configuration of anticipatory perceptual, motor, or cognitive brain processes that is initiated by task context and actively maintained for subsequent task performance. Set is made possible by prior perceptual, motor, or cognitive experience with the same or a similar task. Set-related processes are thought to establish conditions that guide and channel the fast communications between areas of the cerebral cortex that underlie perception, action, or cognition. It is proposed that set-related processes change the functional connectivity of cortical areas within and between large-scale networks of cognition, or neurocognitive networks. The operation of set has been implicated from the analysis of neuroimaging data recorded in a study of human subjects performing a demanding cued visuospatial attention task. From analysis of long-range directed (top-down) functional connectivity, a neurocognitive network of frontal and parietal cortical areas, called the Dorsal Attention Network, was inferred to modulate activity in retinotopic areas of Visual Occipital Cortex (VOC): top-down influences were larger to VOC subregions representing the attended visual hemifield than to subregions representing the unattended hemifield. This difference was maintained over seconds to minutes throughout the entire task. Bottom-up influences from the two subregions did not differ. The maintenance of task-specific top-down modulation of VOC throughout a recording session suggests that it reflects visuospatial attentional set.

Keywords Perception • Action • Cognition • Cerebral cortex • Anticipation • Neural network • Functional connectivity • Top-down • Visuospatial attention • Granger causality

S.L. Bressler (✉)
Department of Psychology, Center for Complex Systems and Brain Sciences, Cognitive Neurodynamics Laboratory, Florida Atlantic University, Boca Raton, FL, USA
e-mail: bressler@fau.edu

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_15

1 Introduction

Set is an essential component of normal human behavior. As people anticipate events in the environment and perform tasks in daily life, they prepare appropriate perceptual, motor, and cognitive processes. Perceptual set is the predisposition to perceive a specific sensory stimulus or stimuli; motor (preparatory) set is the intention to perform a specific action or actions; and cognitive set is the tendency to execute a specific cognitive function or functions. In general, set may be viewed as any configuration of perceptual, motor, and cognitive processes that is initiated by behavioral context and actively maintained for subsequent behavior [19].

Set-related effects are known in a variety of cognitive functions, including motor speech [1], saccadic eye movements [7], visual search [8], rule-based behavioral selection [13, 18], visuospatial attention [14], and visual discrimination [20]. Most studies on the neural basis of set have centered on prefrontal cortex [2, 5, 13, 19], although posterior parietal cortex [10, 15, 22] and basal ganglia [16] have also been implicated.

NeuroCognitive Networks (NCNs) are large-scale systems of distributed and interconnected neuronal populations in the central nervous system organized to perform cognitive functions [3]. NCNs involving prefrontal and posterior parietal cortical areas figure prominently in important aspects of cognition [17]. This report considers the functional configuration of NCNs as a potential mechanism for the instantiation of set in the brain.

Changing the functional interdependency relations among their component brain areas according to task-related processing demands is a potentially powerful mechanism for the set-related configuration of NCNs. It is known that a high-level frontoparietal NCN, called the Dorsal Attention Network (DAN), is responsible for controlling the selection of task-specific sensory information in humans and non-human primates [6]. The frontal and parietal regions of the DAN are consistently activated by central cues indicating where a peripheral object will subsequently appear. Set-related NCN configuration is postulated to occur in visuospatial attention as the top-down functional modulation of "low-level" visual cortical areas by "high-level" cortical areas of the DAN.

2 Methods

Is it possible to quantitatively assess the top-down functional modulation of low-level sensory areas of the brain by high-level control areas during attentional set? Wiener [25] proposed that statistical prediction of the activity in one brain region from that in another region might come from the study of "coefficients of causality running both ways" between regions. This idea would be useful for determining top-down functional modulation related to attentional set if such statistical prediction measurements could be practically derived from neural time series data recorded during set-related behavior.

Can measurements of the type proposed by Wiener actually be made? Following the lead of Wiener, Granger [11] proposed a method for measuring "coefficients of causality" from empirical data by what has come to be known as Granger Causality (GC) [12]. Granger, considering two arbitrary time series, y_t and x_t, proposed to compare two autoregressive (ar) models, called the restricted and unrestricted models, of y_t. The restricted model of y_t is an ar model that includes only past terms of y_t. The unrestricted model of y_t includes those same past terms of y_t, but also includes past terms of x_t. The improvement in predictability of y_t by inclusion of x_t in the unrestricted model, as compared to that of the restricted model, is taken as a measure of statistical "causality". In other words, if inclusion of x_t in the unrestricted model significantly improves the predictability of y_t, as compared to the restricted model, then x_t has a Granger casual influence on y_t. Determination of this improvement can be made by comparing the variances of the residual time series of each model: if the unrestricted model residual variance is significantly lower than that of the restricted model, then the unrestricted model is better at predicting y_t than the restricted model, and it is said that x_t Granger causes y_t. Not only can y_t be modeled using x_t, but x_t can also be modeled using y_t. Therefore, for any two time series, x_t and y_t, their interdependency relations may be either symmetric, i.e. $GC_{x\to y} \approx GC_{y\to x}$, or asymmetric, i.e. $GC_{x\to y} \neq GC_{y\to x}$. Multivariate autoregressive modeling may also be used in place of the bivariate models [24]. Autoregressive modeling is increasingly being used to measure directed functional connectivity in the analysis of brain networks.

To follow Wiener's proposal for measuring the statistical prediction of activity in one brain region from that in another may thus be accomplished by applying autoregressive modeling to neural time series data. To measure directed functional connectivity related to top-down attentional modulation in the brain, then, minimally requires that neural time series data be recorded during attention-demanding task performance, and that directed functional connectivity be measured from those time series. Recording modalities that currently provide potentially suitable time series data include the electroencephalogram (EEG), magnetoencephalogram (MEG), and functional Magnetic Resonance Imaging (fMRI) in normal humans; electrocorticogram (ECoG) in human patients with intracranial electrodes; and local field potential (LFP), spiking single-unit activity (SUA), and multi-unit activity (MUA) in experimental animals.

The following section discusses results obtained by measuring directed functional connectivity from fMRI Blood Oxygen-Level Dependent (BOLD) time series recorded from human subjects engaged in a demanding visuospatial attention task. The subjects performed the task in experiments conducted by the research group of Maurizio Corbetta and Gordon Shulman at Washington University School of Medicine, St Louis [23]. Whole-head fMRI BOLD data were acquired with a Siemens Allegra 3T scanner. Trial-based analysis was performed with ROIs from the DAN (FEF, IPs) and VOC (V1, V2, V3A, VP, V4) of each hemisphere [4]. GC was computed as an F-statistic from bivariate ar models, in both top-down and bottom-up directions for every DAN-VOC voxel pair. A significant F-statistic indicated that the activity in one voxel was predictable from that in another.

3 Results

The fraction of significant voxel-pair F-statistics was found to be variable across ROI pairs, indicating that ROIs were not spatially homogeneous in their influences on one another. Significantly greater GC in the top-down than bottom-up direction over all ROI pairs and subjects suggested that the Dorsal Attention Network (DAN) exerts top-down modulation of the Visual Occipital Cortex (VOC) in relation to visuospatial attention [4]. In addition, top-down GC to intermediate-tier ventral VOC regions (VP and V4) was significantly greater than that to low-tier regions (V1 and V2), suggesting that these intermediate-tier VOC regions serve as portals for top-down modulatory influence on lower-tier visual cortex.

In a subsequent analysis of the same data [24], the average net influence to voxels in the receiving ROI was computed in each direction between DAN and VOC ROI pairs. This average net influence to receiving voxels was computed from voxel-voxel influences measured as multivariate ar coefficients for voxel pairs preselected by the Least Absolute Shrinkage Selection Operator (LASSO). Advantages of this method were that: (1) the coefficients were obtained from multivariate rather than bivariate ar models, with each voxel-voxel influence conditional on all other voxels in the model; (2) the average receiving-voxel net influence was a summary statistic that represented directed functional connectivity at the inter-regional level; and (3) the LASSO procedure allowed insignificant influences to be eliminated from consideration. The directional asymmetry (top-down > bottom-up) between the DAN and VOC reported in the earlier study was confirmed using this LASSO method to compute the average receiving-voxel net influence.

4 Discussion

In unpublished results using the LASSO method, we have found that, in the visuospatial attention task, top-down influences are spatially selective, being larger to VOC subregions representing the attended visual hemifield than to subregions representing the unattended hemifield. Furthermore, spatial selectivity in the task is specific to the top-down direction, is maintained over seconds to minutes throughout the entire task, and is absent in randomized controls. These results suggest that top-down influences from the DAN modulate VOC, and that this modulation is related to task set.

Our findings indicate that set plays an important role in visuospatial attention, and furthermore, that the Wiener-Granger approach offers a practical course for the analysis of set-related behavior from fMRI BOLD time series data. The relation of the fMRI BOLD signal to underlying electrophysiological signals is still poorly understood. Some studies have suggested that ar modeling applied to fMRI BOLD data does not accurately reflect the directionality of electrophysiological interactions

[21], but others have suggested that it can be effectively employed to measure directed functional connectivity in neural systems when methodological issues are carefully considered [9].

References

1. Abbs JH, Cole KJ (1982) Consideration of bulbar and suprabulbar afferent influences upon speech motor coordination and programming. In: Speech Motor Control. Griller S, Lindblom B, Lubker J, Persson A (Eds.), pp. 156–186, Pergamon, Oxford.
2. Asaad WF, Rainer G, Miller EK (2000) Task-specific neural activity in the primate prefrontal cortex. J Neurophysiol 84:451–459.
3. Bressler SL (2008) Neurocognitive networks. Scholarpedia 3:1567.
4. Bressler SL, Tang W, Sylvester CM, Shulman GL, Corbetta M (2008) Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visuospatial attention. J Neurosci 28:10056–10061.
5. Cole MW, Etzel JA, Jacks JM, Schneider W, Braver TS (2011) Rapid transfer of abstract rules to novel contexts in human lateral prefrontal cortex. Front Hum Neurosci 5:142.
6. Corbetta M, Patel G, Shulman GL (2008) The reorienting system of the human brain: From environment to theory of mind. Neuron 58:306–324.
7. DeSouza JFX, Menon RS, Everling S (2003) Preparatory set associated with pro-saccades and anti-saccades in humans investigated with event-related FMRI. J Neurophysiol 89:1016–1023.
8. Eimer M, Kiss M, Nicholas S (2011) What top-down task sets do for us: An ERP study on the benefits of advance preparation in visual search. J Exp Psychol Hum Percept Perform 37:1758–1766.
9. Friston K, Moran R, Seth AK (2013) Analysing connectivity with Granger causality and dynamic causal modeling. Curr Opin Neurobiol 23:172–178.
10. Goodwin SJ, Blackman RK, Sakellaridi S, Chafee MV (2012) Executive control over cognition: stronger and earlier rule-based modulation of spatial category signals in prefrontal cortex relative to parietal cortex. J Neurosci 32:3499–3515.
11. Granger CWJ (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37:424–438.
12. Granger CWJ (1980) Testing for causality: A personal viewpoint. J Econ Dynam Contr 2:329–352.
13. Hoshi E, Shima K, Tanji J (2000) Neuronal activity in the primate prefrontal cortex in the process of motor selection based on two behavioral rules. J Neurophysiol 83:2355–2373.
14. t'-Jong TG, Boehler CN, Kenemans JL, Woldorff MG (2011) Differential functional roles of slow-wave and oscillatory-alpha activity in visual sensory cortex during anticipatory visual-spatial attention. Cereb Cortex 21:2204–2216.
15. Kamigaki T, Fukushima T, Miyashita Y (2009) Cognitive set reconfiguration signaled by macaque posterior parietal neurons. Neuron 61:941–951.
16. Kimura M, Rajkowski J, Evarts E (1984) Tonically discharging putamen neurons exhibit set-dependent responses. Proc Natl Acad Sci U S A 81:4998–5001.
17. Mesulam M (2009) Defining neurocognitive networks in the BOLD new world of computed connectivity. Neuron 62:1–3.
18. Reverberi C, Gorgen K, Haynes J-D (2012) Compositionality of rule representations in human prefrontal cortex. Cereb Cortex 22:1237–1246.
19. Sakai K (2008) Task set and prefrontal cortex. Annu Rev Neurosci 31:219–245.
20. Shiu LP, Pashler H (1992) Improvement in line orientation discrimination is retinally local but dependent on cognitive set. Percept Psychophys 52:582–588.

21. Smith, SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann CF, Nichols TE, Ramsey JD, Woolrich MW (2011). Network modelling methods for FMRI. Neuroimage 54, 875–891.
22. Stoet G, Snyder LH (2004) Single neurons in posterior parietal cortex of monkeys encode cognitive set. Neuron 42:1003–1012.
23. Sylvester CM, Shulman GL, Jack AI, Corbetta M (2007) Asymmetry of anticipatory activity in visual cortex predicts the locus of attention and perception. J Neurosci 27:14424–14433.
24. Tang W, Bressler SL, Sylvester CM, Shulman GL, Corbetta M (2012) Measuring Granger Causality between cortical regions from voxelwise fMRI BOLD signals with LASSO. PLoS Comput Biol 8:e1002513.
25. Wiener N (1956) The theory of prediction. In: Modern Mathematics for the Engineer. Beckenbach EF (Ed), pp. 165–190, McGraw-Hill, New York.

The Dissipative Many-Body Model and Phase Transitions in Brain Nonlinear Dynamics

Antonio Capolupo, Walter J. Freeman, and Giuseppe Vitiello

Abstract We study the energy consumption of the brain in the framework of the dissipative many-body model and the generalized Carnot cycle model. We focus our attention on the expenditure of energy to facilitate the emergence of patterns and dissipation of so-called dark energy in knowledge retrieval. The general picture of the process, by which brains construct knowledge from information and how the generalized Carnot cycle describes it, is presented in terms of Bose-Einstein condensate in the system ground state. We postulate that the extremely high density of energy sequestered briefly in cortical activity patterns can account for the vividness, richness of associations, and emotional intensity of memories recalled by stimuli.

Keywords Neural energy consumption • Many-body dynamics • Phase transitions • Cortical activity patterns • Brain dark energy

1 Introduction

Neurophysiological data [1, 2] show that an extremely intricate net of neuronal correlation forms among distant part of the brain under the action of conditioned stimuli. The correlation manifests itself to the observer in the form of coexisting, physically distinct amplitude modulated (AM) and phase modulated (PM) patterns (wave packets) that are classifiable with categories of conditioned stimuli. The onset of AM patterns into irreversible sequences is so rapid that it resemble cinematographic frames. The brain appears then as a many-body system [3, 4] and the description of the neuronal dynamics can be faced by use of the field concept

A. Capolupo (✉) • G. Vitiello
Dipartimento di Fisica, University of Salerno, 84084 Fisciano (Salerno), Italy
e-mail: capolupo@sa.infn.it; vitiello@sa.infn.it

W.J. Freeman
Division of Neurobiology, Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA
e-mail: dfreeman@berkeley.edu

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_16

proper of the many-body physics. The dissipative quantum model of brain [4, 5] is proposed along this line of thought introducing the further requirement that the brain is an open dissipative system permanently linked to the environment. The dissipative model then allows [6] the description of each AM pattern as a consequence of the spontaneous breakdown of symmetry triggered by external stimulus and associated with one of the quantum field theory unitarily inequivalent ground states [6]. The patterns rapid sequencing is associated to the non-unitary time evolution implied by dissipation.

In trained animals, use of arrays of electrodes to record the electroencephalogram (EEG) from the surfaces of the visual, auditory, somatic and olfactory cortices as the animals respond to conditioned stimuli [1,2] show that the brain is continuously moved out from its ground state activity entering a non-stationary dynamical regime, *it goes through a continuous flow of phase transitions* [4, 5]. It is observed that an abrupt decrease in the analytic power of the background activity to near zero, depicted as a null spike [1, 2, 7], initiates a perceptual phase transition. A brief state of indeterminacy, in which the significant pass band of the ECoG is near to zero and phase of ECoG is undefined, is induced by the reduction in the amplitude of the spontaneous background activity. The phase transition process to a new AM pattern can be driven by the stimulus arriving at or just before this state. The cortical dynamical regime between the null spikes is stationary and is called a frame [1, 2, 7].

Topologically non-trivial structures (phase cones and vortices) characterize the system dynamics during the non-equilibrium phase transition process [1, 2, 7]. The phase cone, a spatial phase gradient, is imposed on the carrier wave of the wave packet in a frame by the propagation velocity of the largest axons having the highest velocity in a distribution. The apex location is a random variable across frames determined by the accidents of where the null spike is lowest and the background input is highest. The null spike behaves as a vortex and has rotational energy at the geometric mean frequency of the pass band. In this paper, the dynamic formation of phase cones and vortices in brain waves are described in terms of non-homogeneous boson condensation processes in the formalism of the dissipative model [7].

2 Phase Transitions, Vortices and Phase Cones

In the process of non-instantaneous phase transition the amplitudes of the filtered ECoG [1, 2] often shows clockwise or counterclockwise rotation, giving the appearance of a vortex occupying the whole area of the phase-locked neural activity of the cortex beginning at a point in time and space [8]. The dissipative model predicts [7] the existence of singularities at the phase cone apex associated with the abrupt decrease (null spike) of the order parameter (the feature vector specifying the spatial AM pattern of the analytic amplitude) and the concomitant increase of spatial variance of the phase field (the analytic phase).

In the dissipative many-body model, the spontaneous breakdown of the rotational symmetry related to the electrical dipoles of water and other molecules [4,5] implies the existence of collective fields, called the Nambu-Goldstone (NG) boson modes [4, 5, 9, 10]. We denote them by $P(x)$ and $P^{\dagger}(x)$. The system ground state is a coherent condensate of these NG modes [7,9,10]. Let $\mathscr{P} = \rho\delta$ be the non-vanishing polarization density, where the charge density ρ and the (average) dipole length δ are real quantities. The charge density wave function $\sigma(x)$ is

$$\sigma(x) = \sqrt{\rho(x)}e^{i\theta(x)} . \tag{1}$$

The phase $\theta(x)$ is the NG boson field associated with the breakdown of phase symmetry under the (global gauge) transformation, $\sigma(x) \rightarrow e^{i\lambda}\sigma(x)$, $A_{\mu}(x) \rightarrow A_{\mu}(x)$, where $A_{\mu}(x)$ is the electromagnetic (e.m.) field and λ is space-time independent. The symmetry breakdown occurs when the charge density $\rho(x)$ has a non vanishing expectation value in the system ground state (the vacuum) $|0\rangle$: $\langle 0|\rho(x)|0\rangle = v \neq 0$.

The system is also invariant under the local gauge transformation

$$\sigma(x) \rightarrow e^{ie_0\lambda(x)}\sigma(x), \; A_{\mu}(x) \rightarrow A_{\mu}(x) + \partial_{\mu}\lambda(x), \tag{2}$$

with $\lambda(x) \rightarrow 0$ for $|x_0| \rightarrow \infty$ and/or $|\mathbf{x}| \rightarrow \infty$, $\partial_{\mu}A^{\mu}(x) = 0$, and

$$-\partial^2 A_{\mu}(x) = j_{\mu}(x) - \partial_{\mu}\theta(x) , \tag{3}$$

with $j_{\mu}(x)$ denoting the current. We remark that a shift in the $\theta(x)$ phase field, i.e.

$$\theta(x) \rightarrow \theta(x) - \alpha f(x) , \tag{4}$$

describes non-homogeneous boson condensation of the field $\theta(x)$ in the system ground state, namely the formation of coherent domains of finite size [7, 9, 10]. α is a constant, $f(x)$, called the boson condensation function, acts as a "form factor" specific for the considered domain [7,9,10] and satisfies the same equation satisfied by the $\theta(x)$ field, i.e. $\partial^2 f(x) = 0$.

In order for the condensation process to be physically detectable, $f(x)$ has to carry some topological singularity, it has to be path-dependent [7, 9, 10] (a regular $f(x)$ do not produce any observable effect). However, observables are influenced by gradients, $\partial_{\mu} f(x)$, in the boson condensate and therefore $\partial_{\mu} f(x)$ has to be single-valued. The stationary function $f(x)$ may carry a vortex singularity given by

$$f(x) = \arctan\left(\frac{x_2}{x_1}\right) , \tag{5}$$

which shows that the phase is undefined on the line $r = 0$, with $r^2 = x_1^2 + x_2^2$, consistently with the observed phase indeterminacy in the process of transition between two AM pattern frames. As a result of the single-valuedness of $\sigma(x)$, the

topological singularity is characterized by the winding number n: $\oint \nabla f(x) \cdot dl = 2\pi n, n = 0, \pm 1, \pm 2, \ldots$, when the integration path is along the closed circle $(0, 2\pi)$ (flux quantization). Phase transitions can be shown to be induced only by a singular boson transformation function $f(x)$ [7, 9, 10]. For this reason topologically non-trivial extended objects, such as vortices, are initiated during the critical regime of phase transitions. The apex of phase cones, where vortices start to be formed, is not initiated within frames, but between frames (during phase transitions). As in observations, the initial site where non-homogeneous condensation starts (the phase cone apex) is not conditioned by the incoming stimulus, but is randomly determined by the concurrence of a number of local conditions. The non-homogeneous condensation of the phase field $\theta(x)$, out of which the vortex arises, spans (almost) the whole system since it is a (quasi-)massless field. This is why in its life-time the vortex is observed to occupy the whole area of the phase-locked neural activity of the cortex.

The null spike appearing in the critical regime acts as a *shutter* that blanks the intrinsic background ECoG. When the order parameter goes to zero the microscopic activity does not decrease, but, as the model predicts, it becomes disordered, unstructured (fully symmetric). At very low analytic amplitude, the analytic phase is undefined, as it is indeed at the vortex core, and, under the incoming weak sensory input, the background activity may re-set in a new AM frame, if any, formed by reorganizing the existing activity, not by driving it by input (except for the small energy provided by the stimulus that is required to selectively excite a Hebbian nerve cell assembly needed to force the phase transition). As in the observations, in the model the reduction in activity constitutes a singularity in the dynamics at which the phase is undefined. The aperiodic shutter allows opportunities for phase transitions.

The model predicts converging (imploding) and diverging (exploding) regimes [7]. Many phase cones show little or no rotation but repetitive outward or inward pulsations with each cycle [8]. The singularity of to their rotational gradients (vortices) is associated to the one of the vortex core. The model explains all four types of these observed spatiotemporal phase gradients. The negative gradient could be explained in conventional neurodynamics (e.g. in terms of a pacemaker), but not the positive gradient. There is no explanation in the conventional framework of why both gradients, the positive and the negative one, occur, one or the other at random.

3 Brain Dark Energy

The neural mechanism of perception depends on repeated transfer of mesoscopic energy to microscopic energy and vice-versa, as the basis for the disintegration of a mesoscopic AM pattern and the formation of a new one, respectively. Provided changes in the inverse temperature β are slow, as it actually happens in mammalian brains which keep their temperature nearly constant, these energy transfers are controlled by the time derivative of the number N of the NG field condensate [6, 7]:

$$dE = \sum_k E_k \frac{dN_k}{dt} dt = \frac{1}{\beta} d\mathscr{S} \,. \tag{6}$$

E_k and $\mathscr{N}_k$ are the energy and the number of the NG excitations of momentum k. Equation (6) relates the changes in the energy $E \equiv \sum_k E_k \mathscr{N}_k$ and entropy $\mathscr{S}$ implied by the minimization of the free energy $\mathscr{F}$ at any $t, d\mathscr{F} = dE - (1/\beta) dS = 0$. Through the changes in time of the NG condensate, the entropy changes and heat dissipation, $dQ = (1/\beta) dS$, involved in the disappearance/emergence of the coherence (ordering) associated to the AM patterns, turns into energy changes. Heat dissipation is indeed a significant variable in laboratory observations. Brains require constant perfusion with arterial blood and venous removal to dispose of substantial waste heat.

Let $|0\rangle$ and $|0(\theta)\rangle_N$ denote (see Refs. [4, 6, 7] for details) the states of the brain activity corresponding to the absence of NG quanta describing the long range neuronal correlations A_k and $\tilde{A}_k$ ($|0\rangle$ is "the vacuum") and the condensate state of condensation density N_{A_k} (and $N_{\tilde{A}_k}$), respectively. The inner products of these states is given by: $\lim_{V\to\infty} \langle 0|0(\theta)\rangle_N = 0$ and $\lim_{V\to\infty} {}_N\langle 0(\theta)|0(\theta)\rangle_{N'} = 0$, for any N and any $N' \neq N$. These relations signal that in the limit of infinitely many degrees of freedom the process of condensation of the A and $\tilde{A}$ modes *is* a phase transition process: the states $|0\rangle$ and $|0(\theta)\rangle_N$, for any N, represent 'distinct' phases of the brain activity since no overlap exist among them for different values of N. These relations also express the criticality present in the brain background activity since their meaning is that no unitary transformation exists able to lead from one phase coded by N to another phase coded by N', with $N' \neq N$: they are 'unitarily inequivalent phases' and transitions from phase to phase are critical transition processes. The non-equilibrium dynamics has been studied in detail by use of the Ginzburg–Landau (GL) time dependent equation in [11]. For any $t \neq t'$ it is $\lim_{V\to\infty} {}_N\langle 0(\theta, t)|0(\theta, t')\rangle_N = 0$, which shows that brain activity is far from the equilibrium and characterized by criticality at any time t, indeed. Time evolution of the phase coded by N appears to be a far from the equilibrium critical process ($|0(\theta, t)\rangle_N$ and $|0(\theta, t')\rangle_N$ are unitarily inequivalent phases). In the non-stationary regime, the non-vanishing dF expresses the rate at which the system approaches the stationary regime at the minimum of the free energy. In [11] it is shown that the rate of change of the condensate $d\, N_A/dt$ in the non-stationary regime, named the critical GL regime, is proportional to the relaxation term $\mathscr{R}_{diss} \equiv -\Gamma_R N_A$, with the 'damping' $\Gamma_R \equiv \Gamma_1 + \Gamma_2$, where Γ_1 depends on the diffusion coefficient $D_{GL} \equiv \xi_{GL}^2/\tau_{GL}$, ($\xi_{GL}$ and τ_{GL} are the GL correlation length and the GL relaxation life-time, respectively). Γ_2 depends on non-homogeneities.

By closely following Ref. [12], we now compute the condensate energy used to feed criticality and phase transitions. Criticality means for the brain to be in a high-energy state of readiness for phase transition in order to face unpredictable events in its surrounding world. The criticality energetic needs add to the losses of energy by leakage. As a result, the brain constantly dissipates energy in self-sustained clouds of axonal action potentials and dendritic currents.

The condensate density in the state $|0(\theta,t)\rangle_N$, the time evolved at time t of $|0(\theta)\rangle_N$, is given by [4,6,12]

$$N_{A_k}(t) = {}_N\langle 0(\theta,t)|A_k^\dagger A_k|0(\theta,t)\rangle_N = \sinh^2(\Gamma_k t + \theta_k) = \frac{1}{e^{\beta(t)E_k} - 1}\,. \quad (7)$$

Γ_k denotes the life-time constant of the k-mode A_k and $\beta(t)$ is assumed to be slowly varying in time in order to ensure quasi-stationarity at any time t. The energy density of the condensate, at a given time t, is computed to be [12]

$$\rho = 4\pi^2 \int_0^\infty dk\; k^2\; \omega_k \sinh^2(\Gamma_k t + \theta_k) = 4\pi^2 \int_0^\infty dk\; k^2\; \frac{\omega_k}{e^{\beta(t)E_k} - 1}\,, \quad (8)$$

which is the energy contribution of the condensation density (7) to the energy requirements for criticality and phase transitions. It is the brain "dark energy" [13]. Setting the phase velocity $v_p = 1$ and $\hbar = c = k_B = 1$, for massless fields, such as NG fields, $\omega_k = k$, we have

$$\rho_{m=0} \;=\; 4\pi^2 \int_0^\infty dk\; \frac{k^3}{e^{\beta(t)E_k} - 1} \;=\; \frac{4}{15}\,\pi^6\, T^4(t). \quad (9)$$

Boundary effects and impurities may contribute to give a non-vanishing effective mass to the NG fields [4,9,10]. In such a case, the integral in Eq. (8) converges and its upper bound is given by Eq. (9).

Criticality is characterized by chaotic, scale-free, power laws behavior, as laboratory observations show [2,14], which are features also implied by the model since phase transition processes turn out to be described by chaotic "trajectories" from phase to phase [15], and coherent states have been shown to be characterized by scale-free, power law features [16].

The thermodynamic process of creating knowledge from information in a sensory cortex is cyclic. We have modeled it using the generalized Carnot cycle [12, 17], in which entropy is reduced by the expenditure of energy to facilitate the emergence of patterns (Fig. 1). In Fig. 1a, the gas and liquid phases coexist in varying degrees. Energy is put in by heating in $(2-3)$ and removed as waste heat in cooling $(4-1)$.

A cycle begins with a sensory cortex in a state of random background activity with low analytic power, $A^2(t)$, that is symmetric in having $1/f$ power spectral density (PSD) and no spatial or temporal pattern. The arrival of a stimulus-evoked sensory volley of pulses breaks the symmetry by initiating a narrow band oscillation that synchronizes the background pulse firings without increasing mean rates (isothermal compression). A peak on the PSD and a spatial pattern of amplitude modulation (AM) in the EEG appear. In Fig. 1b the cortex uses information from step (1) to construct, and transmit knowledge in step (3). In step 2 the AM pattern is fixed and the analytic power continues to rise, not by synchronization but by expenditure of energy in transmission of the AM pattern (adiabatic heating),

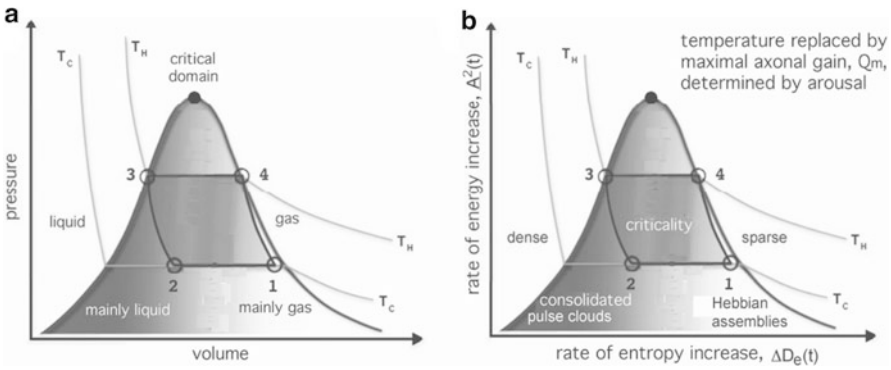


Fig. 1 (**a**) The the generalized Carnot cycle. (**b**) The cortex generalized Carnot cycle [2]

leading to maximal power and minimal entropy, with maximal classification of EEG patterns with respect to the conditioned stimuli containing information. The height of the cycle $(2-3)$ is determined by the degree of arousal as indexed by the asymptotic maximum of axonal gain, Q_m [18]. Average brain temperature is homeostatically regulated by blood circulation, but local fluctuations are widespread and closely related to analytic power. In step three the AM pattern dissolves as the firing rates diminish owing to the refractory periods of the neurons, and the strength of synaptic coupling wanes (isothermal expansion). Free energy is derived from oxidative metabolism and is dissipated as heat in all four steps. In step 4 the distribution of characteristic frequencies in the PSD spectral peak go out of phase and cancel (adiabatic cooling). The area enclosed by the loop is a measure of pragmatic information [19], i.e. the ratio of the rate of energy dissipation (power) to the rate of decrease in entropy (increase in information).

In the transition from step four to step one an existing AM pattern undergoes the rapid and reliable extinction of the commitment of energy to an attractor. This extinction is necessary in order that a new AM pattern might form in step 1–2. Our experiments suggest that a singularity [2, 20] appears during the temporal minima of analytic power. Cinematic display of $log_{10} A^2(t)$ reveals brief, sharply localized minima, the null spikes, at which the power may decrease from the mean levels by six orders of magnitude or more, coinciding with a temporal discontinuity in the analytic phase.

4 Conclusions

In the action-perception cycle, by the continual updating of the *meanings* of the flows of information information exchanged in its engagement with the environment, the brain proceeds from information to *knowledge* in its own world that we describe as its Double [5]. Our model, by resorting to the mechanism of

the spontaneous breakdown of symmetry and on dissipation, accounts for the observed dynamical formation of spatially extended domains of neuronal 'in phase' oscillations and of their rapid sequencing. It allows to study the transient non-homogeneous patterns of percepts appearing during the brain non-instantaneous phase transitions. Brains dissipate metabolic energy at rates ten-fold greater than rates in any other organ (so-called brain dark energy). The Carnot cycle and the dissipative many-body model account for the high density of energy sequestered briefly in cortical activity patterns and also account for the mobilization of the myriad microscopic details that are stored in the modified synapses into the macroscopic order that is expressed by the pulse cloud and its controlling field of dendritic currents observed in the EEG. In [12] it is suggested that ephapsis may play a role in the phase transition process (ephapsis denotes action of a neuron on others in apposition that is mediated by local chemical and/or electrical fields and not by chemical or electrical synapses).

Mesoscopic/macroscopic dynamical properties of the system are derived from the many-body dynamics: it appears that the system classical behavior cannot be explained without recourse to the underlying many-body dynamics. This brings us to the description of energy dissipation as heat in the disappearance and emergence of coherence. The neurons, the glia cells and their subcellular components are *not* quantum objects [4–6]. The quantum degrees of freedom are those associated to the dipole vibrational field and to other fields such as the phase field.

References

1. W. J. Freeman, Clin. Neurophysiol. **116** (5), 1118–1129 (2005)
2. W. J. Freeman, R. Quian Quiroga, *Imaging Brain Function with EEG*, (Springer, New York, 2013)
3. L.M. Ricciardi and H. Umezawa, Kibernetik **4**, 44–48 (1967). Reprint in *Brain and Being* ed. by G. G. Globus, K. H. Pribram, G. Vitiello, (John Benjamins, Amsterdam, 2004), pp. 255–266.
4. G. Vitiello, Int. J. Mod. Phys. B **9**, 973–989 (1995)
5. G. Vitiello, *My Double Unveiled*, (John Benjamins, Amsterdam, 2001)
6. W. J. Freeman, G. Vitiello, Phys. of Life Reviews **3**, 93–118 (2006)
7. W. J. Freeman, G. Vitiello, Int. J. Mod. Phys. B **24**, 3269–3295 (2010)
8. H. Umezawa, *Advanced field theory: micro, macro and thermal concepts*, (AIP, New York, 1993)
9. M. Blasone, P. Jizba and G. Vitiello, *Quantum Field Theory and its macroscopic manifestations*, (Imperial College Press, London 2011)
10. W. J. Freeman, Displays of movies of null spikes resembling tornados and vortices resembling hurricanes can be downloaded from: http://sulcus.berkeley.edu
"Latest Freeman Manuscripts", Section H.
11. W. J. Freeman, R. Livi, M. Obinata, G. Vitiello, Int. J. Mod. Phys. B **26**, 1250035 (2012)
12. A. Capolupo, W. J. Freeman, G. Vitiello, Phys. of Life Reviews, **10**, 85–94 (2013) W. J. Freeman, R. Kozma, G. Li, R. Quian Quiroga, G. Vitiello, T. Zhang, Advanced Models of Cortical Dynamics in Perception (this volume)
13. M.E. Raichle, Science **314**, 1249–1250 (2006)

14. W. J. Freeman, Intern J. Bifurc. Chaos **14**, 513–530 (2004)
W. J. Freeman, M. Breakspear, Scholarpedia, **2** (2) (2007) 1357. http://www.scholarpedia.org/article/Scale-free-neocortical-dynamics
15. G. Vitiello, Int. J. Mod. Phys. B **18**, 785–792 (2004) E. Pessa and G. Vitiello, Mind and Matter **1**, 59–79 (2003)
E. Pessa and G. Vitiello, Int. J. Mod. Phys. B **18**, 841–858 (2004)
16. G. Vitiello, New Math. and Nat. Comp. **5**, 245–264 (2009)
G. Vitiello, Fractals and the Fock-Bargmann representation of coherent states, in *Quantum Interaction,* ed by P. Bruza, D. Sofge, et al.. Lecture Notes in Artif. Intell., ed by R.Goebel, J. Siekmann, W.Wahlster, (Springer, Berlin 2009), pp. 6–16.
G. Vitiello, Phys. Lett. A **376**, 2527–2532 (2012)
17. W. J. Freeman, R. Kozma, G. Vitiello, Proc IEEE World Congr Comp Intell WCCI/IJCNN 2012 Brisbane QLD, Australia, (IEEE Press, 2012), pp. 3229–3236.
18. W. J. Freeman, H. Erwin, Scholarpedia **3** (2), 3238 (2008) http://www.scholarpedia.org/article/FreemanK-set
19. H. Atmanspacher, H. Scheingraber, Can. J. Phys. **68**, 728–737 (1990)
20. W. J. Freeman, Cognitive Neurodynamics **3** (1), 105–116 (2009)

Advanced Models of Cortical Dynamics in Perception

Walter J. Freeman, Robert Kozma, Guang Li, Rodrigo Quian Quiroga, Giuseppe Vitiello, and Tinglin Zhang

Abstract Here the phenomenon of interest is the flash of recognition and accompanying emotion one experiences when one receives a familiar stimulus. We explain the speed and richness of the event by postulating phase transitions in cortical neuropil: condensation from a gas-like phase to a liquid-like phase followed by evaporation. We model the process with a Carnot-like thermodynamic cycle at three successive levels of complexity: primary sensory cortices; limbic system; global neocortex. We replace the thermodynamic state variables of pressure, volume and temperature with neurodynamic variables, respectively mean beta-gamma power, pattern stability (negentropy), and neural feedback gain (mean interaction strength). We cite evidence that all sensory cortices use this cycle, necessarily so for two reasons. They all evolved from the primordial forebrain of vertebrates dominated by olfaction; they all transmit the same form of perceptual information, wave packets, so signals in all modalities armodel by linear matrix concatenation.

Keywords Carnot cycle • Criticality • Ephapsis • Hebbian assembly • Memory • Phase transition • Wave packet

W.J. Freeman (✉)
Division of Neurobiology, Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA
e-mail: dfreeman@berkeley.edu

R. Kozma
Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152, USA

G. Li • T. Zhang
Department of Control Science and Engineering, State Key Lab of Industrial Control Technology, Zhejiang University, 38 Zheda Road, Hangzhou 310027, People's Republic of China

R.Q. Quiroga
Centre for Systems Neuroscience, University of Leicester, LE1 7RH Leicester, UK

G. Vitiello
Dipartimento di Fisica, University of Salerno, 84084 Fisciano (Salerno), Italy

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_17

1 Introduction

Information is a microscopic variable that is expressed in the pulse intervals and rates of individual neurons. It brings to the cortex the kinds and intensities of all forms of signaling energy derived from the environment, and it carries from cortex all motor commands that thrust the body into the environment. Knowledge is a mesoscopic variable that is stored in modified synapses among networks of neurons and expressed in the spatiotemporal coherences of pulse firings. Each network integrates large quantities of information. Examples are the Hebbian nerve cell assembly [1], the 'cognit' of Fuster and Bressler [2], the 'bubbles' of Taylor [3] and the 'IIT' of Tononi [4]. Meaning is a macroscopic variable that can be expressed in wave packets [5]. These are observed as bursts of oscillatory field potentials in the beta and gamma ranges that recur at rates in the theta and alpha ranges (Fig. 1). Each burst has a narrow band carrier frequency that is amplitude modulated (AM) temporally and spatially. A wave packet carries perceptual information in an AM pattern. The bursts mobilize the knowledge in the cortex and use it in the action-perception cycle by which the goals of subjects are pursued. Intelligence is the ability to use knowledge to formulate and solve problems. In our view wave packets provide the neural activity that carries thoughts. Therefore to understand how brains think we need to know how cortex creates each wave packet, uses it, and ends it efficiently in the cinematic stream of thought.

In this review we focus on the process by which brains use sensory information (I) to create fragments of knowledge (K) in each sensory modality. The fragments

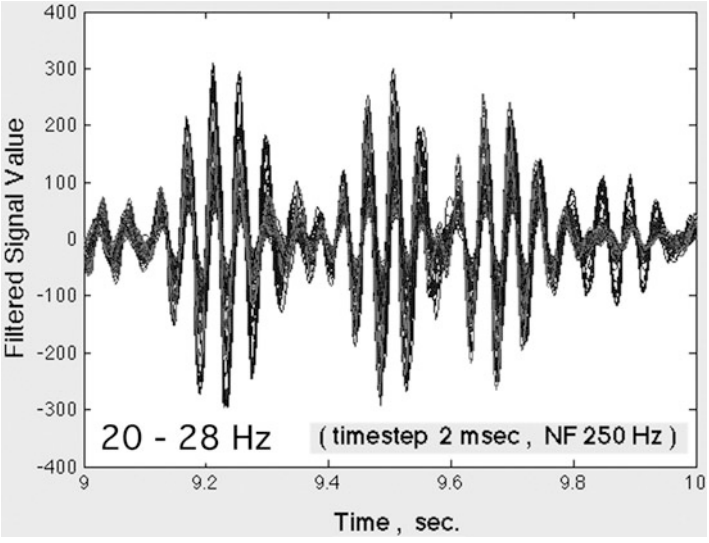


Fig. 1 64 superimposed band pass filtered ECoGs

have the same form in all modalities. They are integrated into Gestalts on transmission through the entorhinal cortex in the limbic system. We model the action-perception cycle as a thermodynamic process, because it far transcends neural networks in depending on neural mass action, and because cortices use metabolic energy at rates far exceeding those of any other tissue [6]. Sadi Carnot (1796–1832) founded thermodynamics by conceiving a cycle that incorporated the variables of pressure, volume mass, and temperature with the gas laws that interrelated them. He did this by defining the four variables, fixing mass and temperature, and graphing pressure-volume relation as an isocline, fixing temperature at another value, and calculating a new isocline. Rudolf Clausius (1822–1888) generalized the model by replacing pressure with energy and volume with entropy. Macquorn Rankine (1820–1872) extended the cycle to include gas-liquid phase transitions. We follow their lead by replacing three thermodynamic state variables with three measurable variables in neurodynamics: energy by ECoG power in frequency bands in the beta-gamma range; rate of negentropy by AM pattern stationarity (implying stability) as measured by rates of movement along trajectories in high-dimensional measurement space; and temperature by neural interaction strengths that we deduce by modeling neural activity patterns with K-sets [1, 5].

2 Methods

Square arrays of 64 electrodes were implanted on the surfaces of sensory cortices (visual, auditory, somatic, olfactory) or scalp of subjects (cats, rabbits, humans) giving windows ranging in size from 5×5 mm to the whole head. Subjects were trained in simple sensory discrimination tasks by reinforcement learning. Sets of 64 EEGs and ECoGs were recorded in blocks of 40 6-s trials, band pass filtered, and transformed to analytic amplitude and phase by the Hilbert transform. A moving window was stepped across the arrays of 64 channels by which AM patterns (given by 8×8 values of analytic amplitude at each step) were extracted from signals recorded from reinforced trials (CS+), unreinforced trials (CS-) as well as control trials (CS0). Each AM pattern was expressed as a 64×1 feature vector, $\mathbf{A}(t)$, giving a point in 64-space. Similar AM patterns projected in 2-space gave clusters of points. AM patterns were inferred to carry knowledge when the significance of correct classification with respect to behaviors gave values of $p < 0.01$ consistently in a given trial set [1].

3 Results

Bursts were extracted using band pass filters (Fig. 1). The temporal FFT was used to identify the temporal locations of beta-gamma bursts having high power, which were deemed likely to carry information yielding classifiable AM patterns. The Hilbert

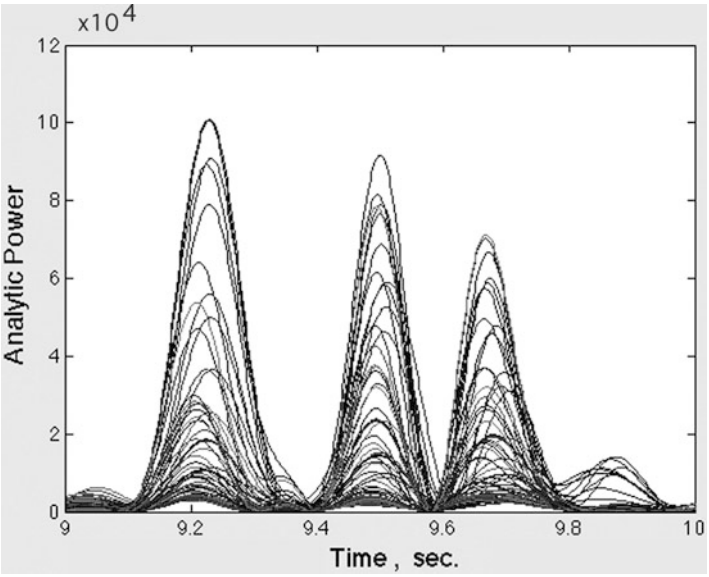


Fig. 2 Band pass filtering imposes beats in $A^2(t)$

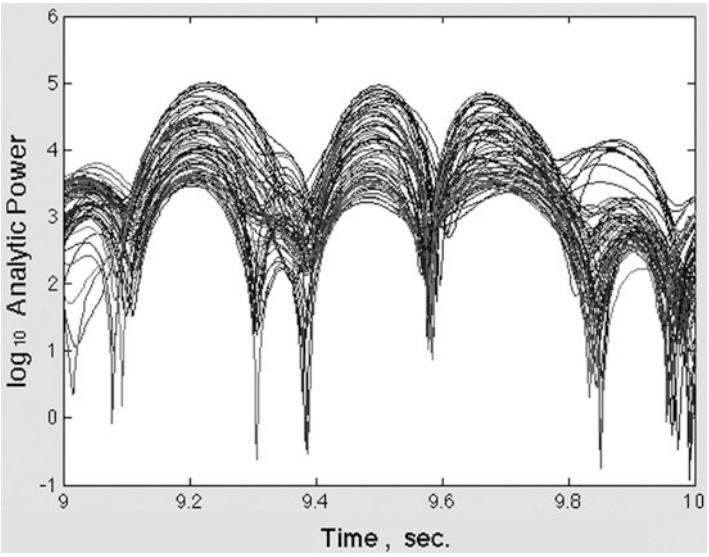


Fig. 3 Log_{10} $A^2(t)$ shows time location of beats in band pass filtered ECoG due to frequency dispersion [1]

transform was used to calculate 64 analytic amplitudes, $A_x(t)$, where $x = 1, \ldots, 64$ (Fig. 2), which were squared to get analytic power). The range between high and low power in the window often exceeded two orders of magnitude, so the measurements were displayed as log_{10} power (Fig. 3) in the beta or gamma pass

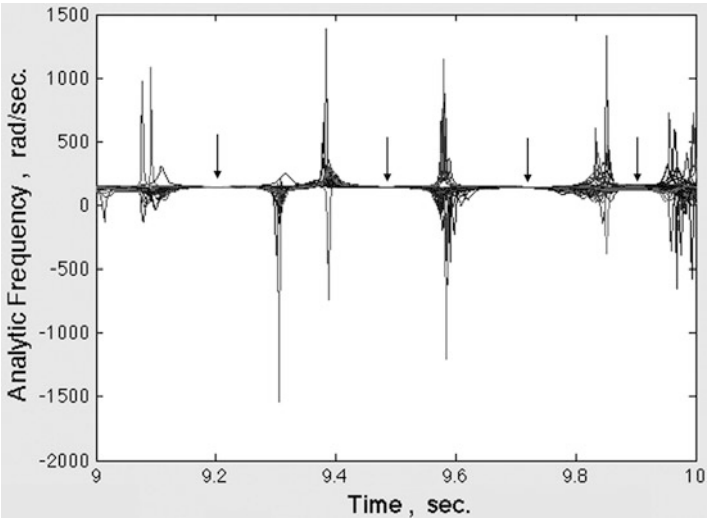


Fig. 4 The extremes show frequency indeterminacy. *Arrows* are at minima in spatial variance of frequency.

band at each step. The $\log_{10}$ mean amplitude squared estimated the ECoG power, $\underline{A}^2(t)$ in the narrow band of each wave **carrier** frequency. The analytic phase was unwrapped and converted to successive differences. Dividing each phase difference in radians by the digitizing step (2 ms) gave estimates of the analytic frequency, $\omega(t)$ in radians/s (Fig. 4).

The feature vector of *normalized* amplitudes, $\mathbf{A}(t)/\underline{A}(t)$ (spatial mean amplitude at each step) gave the AM pattern expressed as a point in 64-space. The Euclidean distance between successive points in 64-space divided by the digitizing step gave the rate of change in AM pattern, $D_e(t)$. In background chaotic ECoG the values of $D_e(t)$ varied widely and unpredictably. A burst with high values of power, $\underline{A}^2(t)$, gave an unbroken sequence of low values of $D_e(t)$ revealing stationarity of an AM pattern (Fig. 5). The persistence of a stationary AM pattern increased the certainty with each digitizing step that the AM pattern manifested an attractor that expressed a fragment of knowledge as a memory.

We verified this conjecture by collecting sets of AM patterns and classifying them with respect to CSs. The optimal measure of AM patterns for classification was the pragmatic information, $H_e(t)$ (Fig. 6).

Alternatively expressed, the product of the minimal rate of reduction in entropy, which we estimated by calculating $1/D_e(t)$, and the maximal rate of dissipated energy (power $H_{e,\,max}(t)$) in each burst gave the pragmatic information. Carnot used the area in each cycle to estimate the work done in each cycle. We used the product (corresponding to the area of the 1,234 rectangle, Fig. 7) to index the amount of knowledge in each burst [1].

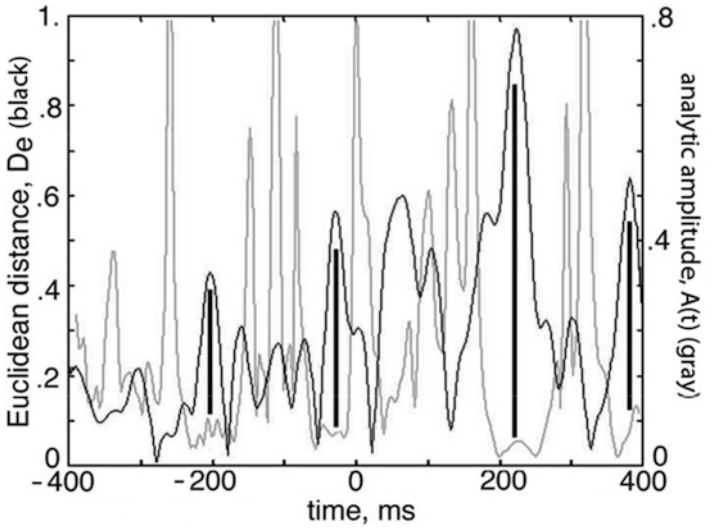


Fig. 5 $D_e(t)$ (*grey*) varied inversely with $A^2(t)$ (*black*). *Bars* show bursts >3 cycles of gamma oscillation

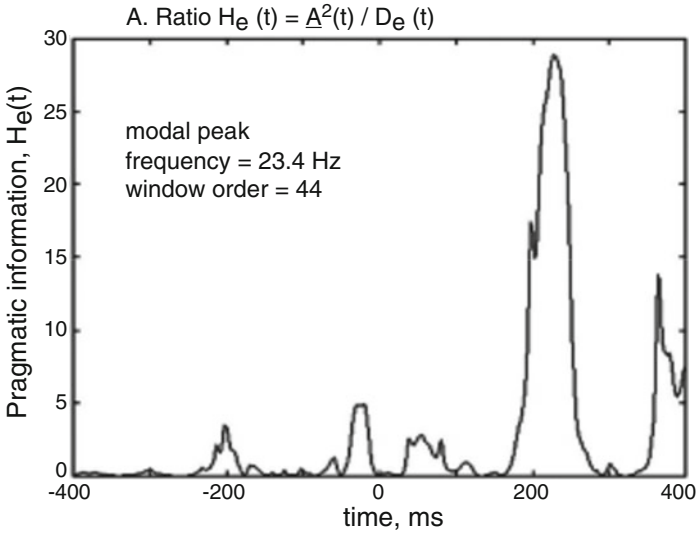


Fig. 6 Pragmatic information, $H_e(t)$, at peak values gave optimal behavioral AM pattern classification [1]

We conceive the phase transition in each burst to have four Steps (Fig. 7). In Step 1–2 (binding and coherence) a conditioned stimulus carried by a microscopic volley of action potentials from receptors ignites a mesoscopic Hebbian assembly that generalizes input to a category of equivalent receptors, abstracts by removing

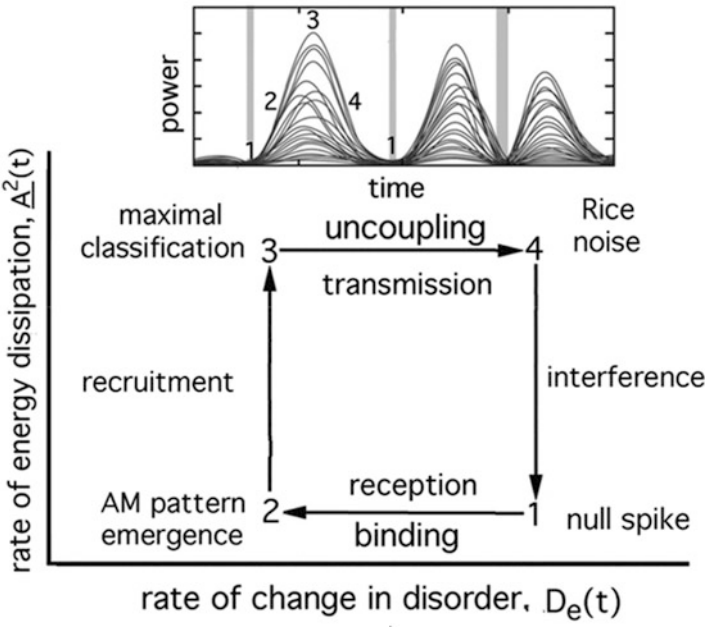


Fig. 7 The four Steps of each cycle are labeled in a graph of a set of 64 values of analytic power, $\mathbf{A}^2(t)$ [1]

irrelevant detail, amplifies the volley, and selects a basin of attraction in cortical memory. The binding in the assembly and the convergence to an attractor decrease the disorder and entropy.

In Step 2–3 (condensation and transmission) the ignition of the Hebbian assembly provides the mesoscopic transition energy needed to start a macroscopic, spatially coherent burst of oscillatory dendritic current that carries the spatial AM pattern (Fig. 7, Step 2–3). The low-density chaotic activity governed by a point attractor and synaptic transmission undergoes an irreversible phase transition by condensing to a high-density narrow-band activity that is governed by a limit cycle attractor [5], shaped by synaptic transmission, and empowered by ephaptic transmission [6, 7]. A mass of neurons is recruited into coherent subthreshold oscillation, so that all neurons in the coherent domain can contribute their bits of synaptic information to the recalled knowledge. The burst is not a representation of the physical properties of a stimulus; it manifests the rich memory of a stimulus that imparts the meaning of a stimulus.

In Step 3–4 (uncoupling, decoherence) the macroscopic pulse cloud shaping and shaped by the dendritic current field down-samples the AM pattern, using time multiplexing and pulse density modulation in cortical columns [1]. The cortex sends bursts through a fan-out-fan-in tract. The tract performs a Gabor transform that selects spatially coherent carrier waves, attenuates all else as noise due to frequency dispersion [8, 9], and delocalizes the signal content [1]. The macroscopic

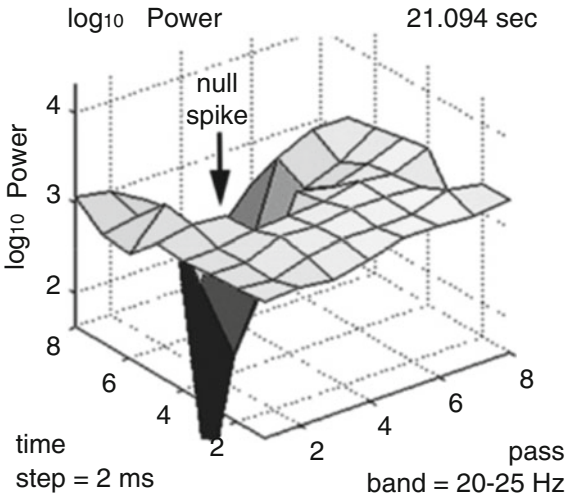


Fig. 8 The null spike is seen as a brief decrease in $\log_{10}$ analytic power in the pass band of the carrier frequency of an AM pattern $\sim 10^{-6}$ below a prevailing level. Spikes are observed by high-speed digitizing of 64 filtered ECoGs from dense 8×8 arrays [1, 8]

information density is spatially uniform. In effect the transform defines the signal as power in spatially coherent narrow band oscillation. Frequency dispersion enhances disorder by phase dispersion [interference, 8, 9]. Thereby AM patterns from all modalities and frequencies can be integrated by linear matrix concatenation [1].

In Step 4–1 (evaporation) Declining power due to refractory periods and to frequency dispersion evaporate the signal by another irreversible phase transition, after a duration that is proportional to the width of the pass band [7–9]. Extremely low values of analytic power appear at spatiotemporal points in cinematic displays of the ECoG (Fig. 8), indicating the presence of a singularity in cortical dynamics [1, 5, 10] that mediates the phase transition of neuropil from the liquid-like phase governed by a limit cycle attractor to the gas-like background phase governed by a point attractor.

We conceive that the null spike frees the cortex from the attractor by taking the order parameter of the synaptic interactions sufficiently close to zero owing to interference. Once initiated at a point on the surface, due to scale-free dynamics the transition can spread over indefinitely large areas of cortex at the conduction velocities of cortical axons as seen in conic ECoG phase gradients [1], either as an explosion or an implosion.

4 Discussion

The Carnot cycle does not model phase transitions. That is done with the Rankine cycle, which is used to model devices that use vapor instead of gas. In this form the medium is a mixture of liquid and gas in a domain of criticality [1, 11] in phase space (Fig. 9).

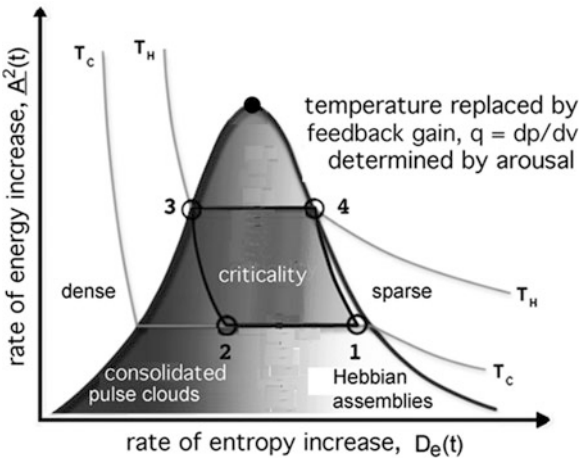


Fig. 9 The Rankine cycle is embedded in a domain of criticality, in which the both phases coexist and support gradation in density as in fog (Modified from [1])

We model the I-K process in thermodynamic terms with a Carnot-like cycle, in which metabolic energy is expended to create and use knowledge. We conjecture that our I-K cycle constitutes a modular dynamic operation at three levels of the action perception cycle in the brain [1]. First is the creation from information of a fragment of knowledge in each sensory modality. Second is the integration of the several fragments into Gestalts in the limbic system, including specification by the hippocampus of the location of the perceiver in time and space at the act of perception. The operation of Hebbian assemblies at this level is revealed by the properties of concept cells [12, 13] in the medial temporal lobe that manifest cross-modal generalization and abstraction. Third is the creation of an AM pattern over the entire forebrain [14–16], which unifies the sensory and motor cortices in the choice of the next act of search in the action-perception cycle, including the strategic plan, the next series of tactical maneuvers, and the changes in sensory input that are predicted by preafference to result from the intended actions.

The salient insight provided by the cycle is the conception of a phase transition (Step 2–3 in Fig. 6) by which the density of neurons that synchronize in subthreshold oscillation in the selected narrow frequency band rises to saturation. The mechanism may include ephaptic transmission [1, 6] based in the high packing density of terminal branches of dendrites and axons in the neuropil. This coupling by electrical interactions is well known in the dorsal spinal cord [17, 18] and has been shown to synchronize cortical neurons both in seizure [19] and in normal function [20].

This feature of the liquid-like phase will require intensive experimental and theoretical studies to determine what local forces are involved, whether there is a threshold for condensation, whether there is hysteresis in evaporation (Step 4–1), and whether the predicted local fluctuations in brain temperature can be measured [10]. We believe that the research is important, because we contend that what makes the difference between a machine that knows nothing and a brain that knows what

it is doing resides in the possibility for high density and exceedingly fine grain of ephaptic exchanges in energy and information in cortical neuropil during the condensed phase of neural activity.

References

1. Freeman WJ, Quian Quiroga R (2013) Imaging Brain Function with EEG: Advanced Temporal and Spatial imaging of Electroencephalographic Signals. New York: Springer.
2. Fuster, J.M. (2006) The cognit: a network model of cortical representation. International Journal of Psychophysiology 60:125–132.
3. Taylor JG (1998) Neural 'bubble' dynamics in two dimensions: foundations. Biol Cybern 80: 393–409.
4. Tononi G (2004) An information integration theory of consciousness. BMC Neurosci. 5: 42.
5. Freeman WJ (1975) Mass Action in the Nervous System New York: Academic. http://sulcus.berkeley.edu/MANSWWW/MANSWWW.html
6. Capolupo A, Freeman WJ, Vitiello G (2013) Dissipation of 'dark energy' by cortex in knowledge retrieval. Physics of Life Reviews, on-line. DOI: 10.1016/j.plrev.2013.01.001
7. Capolupo A, Freeman WJ, Vitiello G (this volume) The dissipative many-body model and phase transitions in brain nonlinear dynamics.
8. Freeman WJ. Deep analysis of perception through dynamic structures that emerge in cortical activity from self-regulated noise. Cognitive Neurodynamics (2009) 3(1): 105–116.
9. Rice SO. Mathematical analysis of random noise. Technical publications monographs, vol. B-1589. New York: Bell Telephone Labs (1950).
10. Freeman WJ, Livi R, Obinata M, Vitiello G (2012) Cortical phase transitions, non-equilibrium thermodynamics and the time-dependent Ginzburg-Landau equation. Int J Mod Phys B 26 (6): 1250035. DOI: 10.1142/S021797921250035Xx
11. Kozma, R, Puljic M, Freeman WJ (this volume) Thermodynamic model of criticality in the cortex based on EEG/ECoG Data. Chapter 1 in: Plenz D (ed), "Criticality in Neural Systems". NY: Wiley, (2012) 1–28
12. Quian Quiroga R. Concept cells: The building blocks of declarative memory functions. Nature Rev Neuroscience 13 (2012): 587–597.
13. Quian Quiroga R (this volume) Concept cells in the human brain
14. Ruiz Y, Pockett S, Freeman WJ, Gonzales E, Li Guang (2010) A method to study global spatial patterns related to sensory perception in scalp EEG. J Neuroscience Methods 191: 110–118. doi:10.1016/j.jneumeth.2010.05.021
15. Brockmeier AJ, Hazrati MK, Freeman WJ, Li L, Principe JC (2012) Locating spatial patterns of waveforms during sensory perception in scalp EEG. EMBC Ann Intern Conf IEEE, pp. 2531–2534, doi: 10.1109/EMBC.2012.6346479
16. Zhang T, Dai L, Wang Y, Freeman WJ, Li G (this volume) EEG spatiotemporal pattern classification of the stimuli on different fingers
17. Arvanitaki A (1942) Effects evoked in an axon by the activity of a contiguous one. J Neurophysiol 5(2):89–108. (first use of 'ephapsis')
18. Wall PD, Lidierth M (1997) Five sources of a dorsal root potential: their interactions and origins in the superficial dorsal horn. Neurophysiol 78(2): 860–871.
19. Steriade M, Amzica F (1994) Dynamic coupling among neocortical neurons during evoked and spontaneous spike-wave seizure activity. J Neurophysiol 72(5): 2051–2069.
20. Anastassiou CA, Perin R, Markram H, Koch C (2011) Ephaptic coupling of cortical neurons. Nature Neurosci 14: 217–223. doi:10.1038/nn.2727

Modeling Cortical Singularities During the Cognitive Cycle Using Random Graph Theory

Robert Kozma and Walter J. Freeman

Abstract Recent brain monitoring experiments indicate intermittent singularities in cognitive processing. We employ the biologically motivated Freeman K models to interpret these findings. In particular, we show that random graph theory provides a suitable mathematical framework to describe the experimentally-observed singularities as phase transitions across the cortical neuropil. We introduce the hypothesis that the rapid propagation of phase dispersion over the hemisphere is the manifestation of the cognitive broadcast as described in Baars' global workspace theory. In addition, our exponentially expanding brain graph model using pioneer neuron sub-plates can be used for describing recent findings on the presence of rich club structures in brain networks.

Keywords Neuropercolation • Freeman K sets • Cognitive cycle • Pioneer neurons • Conscious broadcast

1 Introduction

Criticality in the cortex emerges from the seemingly random interaction of microscopic components and produces higher cognitive functions at macroscopic scales. Recent experiments with high-resolution brain imaging techniques with electrocorticograms (ECoG), electroencephalograms (EEG), and functional magnetic resonance imaging (fMRI) provide an amazing view on the singular dynamics of cortical processes [5, 7]. In particular, singularities have been observed in analytic amplitude and analytic phase patterns over a wide range of frequencies [11].

We employ a hierarchical approach to spatio-temporal neurodynamics based on Freemans K sets [6]. The hierarchy includes microscopic, mesoscopic, and macro-

R. Kozma (✉)
Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152, USA
e-mail: rkozma@memphis.edu

W.J. Freeman
Division of Neurobiology, Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA
e-mail: dfreeman@berkeley.edu

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_18

scopic scales. The original manifestation of K models used differential equations. The present study applies neuropercolation to implement Freeman's principles of neurodynamics [6]. Neuropercolation is a family of probabilistic models based on the mathematical theory of probabilistic cellular automata over lattices and random graphs and it is motivated by structural and dynamical properties of the cortical neuropil [8]. Neuropercolation extends the concept of phase transitions to large interactive populations of nerve cells as an alternative of models based on differential equations.

We demonstrate intermittent synchronization in the neuropercolation model in accordance with experimental findings. We introduce the hypothesis that the rapidly propagating phase gradient is the neurophysiological manifestation of the conscious broadcast as described in Baars' Global Workspace Theory (GWT) [1, 12]. Finally, we show that the exponentially expanding brain graph model (EEGm) [4] can provide a feasible interpretation of recent findings about the so-called rich club graphs in large-scale cortical networks [3, 14].

2 Freeman K Sets in Neuropercolation Domain

Freeman K sets represent a hierarchical approach to spatio-temporal neurodynamics [6]. K sets are multi-scale models, describing increasing complexity of structure and dynamics, including K0, KI, KII, KIII, and KIV sets. K0 sets model the collection of non-interactive neurons forming cortical microcolumns. KI sets consist of interacting K0 sets with nonzero background activity level. KII sets consist of interacting excitatory and inhibitory KI sets with limit cycle oscillations. KIII sets are made up of multiple interacting KII sets with incommensurate frequencies, which can lead to chaotic oscillations; various sensory cortices are modeled as KIII sets. Finally, KIV is made of several KIII sets and can exhibit intermittent synchronization. KIV realizes multi-sensory fusion in the hemisphere as the basis of intentional behavior. In the present work, we limit the studies to up to KIII models.

The first generation of K sets used as mathematical tools a system of nonlinear differential equations. In order to give a more adequate characterization of singularities observed in brains during cognition, neuropercolation uses the concept of random graph theory. Random graphs provide natural means to describe criticality in the behavior of the cortex [8,9]. In the past decade, neuropercolation approach has proved to be an efficient tool to implement K sets and model cortical neurodynamics. Freeman postulated ten basic principles of neurodynamics. Neuropercolation systematically implements Freeman's principles of neurodynamics. Here we list the first five principles addressed in the present studies [9]:

- State transition of an excitatory population from a point attractor with zero activity to a non-zero point attractor with steady-state activity by positive feedback (KI).
- Emergence of oscillations through negative feedback between excitatory and inhibitory neural populations (KII).

- State transitions from a point attractor to a limit cycle attractor that regulates steady-state oscillation of a mixed excitatory-inhibitory cortical population (KII).
- Genesis of chaos as background activity by combined negative and positive feedback among three or more mixed excitatory-inhibitory populations (KIII).
- Distributed wave of chaotic activity that carries a spatial pattern of amplitude modulation made by the local heights of the wave (KIII).

Neuropercolation is a thermodynamics-based random cellular neural network model, which is closely related to cellular automata (CA). CA research has been pioneered by Von Neumann, who anticipated the significance of CA in the context of brain-like computing [13]. In this context, neuropercolation is a novel mathematical tool helping us deciphering the language of the brain. Next we introduce results of calculations with neuropercolation model system to interpret some recent experimental results on brain dynamics [7, 10].

3 Singularities Cortical Dynamics Modeled by Neuropercolation

Brain dynamics is viewed as a sequence of intermittent phase transitions in an open system with synchronization-desynchronization transitions demonstrating symmetry breaking demarcated by spatio-temporal singularities. Phase transitions represent the centerpiece of neuropercolation models as the generalization of bootstrap percolation [9]. Figure 1 gives an example of the intermittent synchronization across 256 simulated channels in an array arranged in a line. The plots in Fig. 1 show

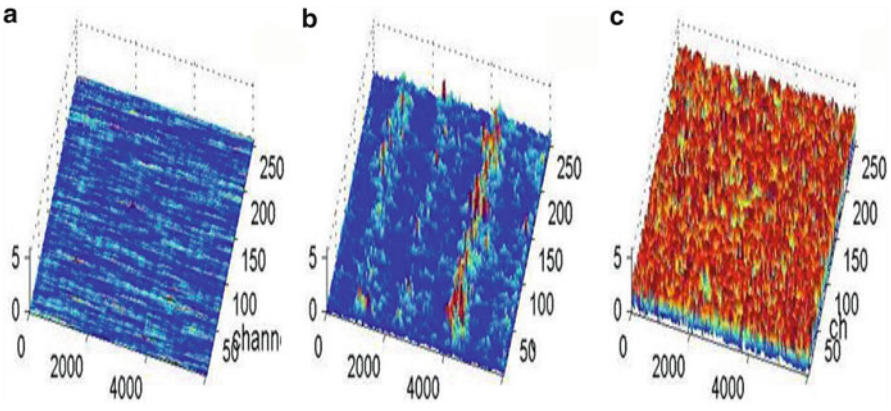


Fig. 1 Phase lag values (vertical axis) in the neuropercolation model with excitatory-inhibitory layers; x-axis: time steps; y-axis: spatial steps (channels from 1 to 256). (**a**) Noise level 13 % (subcritical): high synchrony is seen across the array; (**b**) Noise level 15 % (critical): observe intermittent desynchronization across the array; (**c**) Noise level 16 % (supercritical): the synchrony between channels is diminished [9]

a sequence of synchronization starting with high level of synchrony (subcritical state), through intermittent synchrony (critical state), up to the absence of synchrony (supercritical state).

Near critical state, we observe intermittent desynchronization over large cortical areas for a short period, see Fig. 1b. This is the period of phase transition, when large phase gradients travel rapidly across the cortex. This rapid propagation of phase synchronization-desynchronization resembles the rapid propagation of conscious broadcast as described by the Global Workspace Theory [1]. Recent ECoG experimental studies indicate that phase desynchronization and the collapse of analytic amplitudes are associated by the 'aha' effect of cognitive processing [11]. This leads us to the hypothesis that cortical singularities and phase dispersion could serve as neural correlates of the conscious broadcast.

We have developed the Exponentially Expanding Graph model (EEGm) for brains based on planar geometric principles [4]. The development of brain structures can be described as the evolution that starts from an initial set of N pioneer neurons. N is typically a small number compared to the size of the fully developed brain. Pioneer neurons are modeled as small balls densely packed in a small area of characteristic size. During the developmental stage, this tiny cortical tissue with multiplies about 10^8-fold, to have the fully developed brain with about 10^{11} neurons. The evolution of the brain is described as the inflation of the initial graph G_0 formed by the interconnected pioneer neurons, see Fig. 2. As the cortex develops, the space spanned by the neurons expands so that the distance between the original neurons increases. The space created by this expansion is filled with the newly

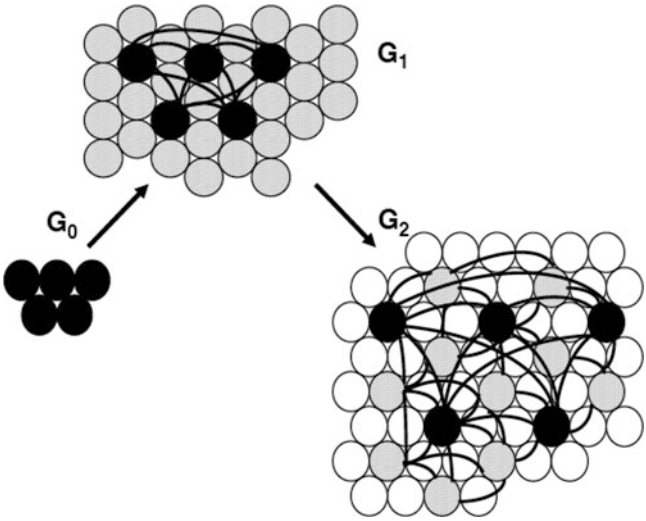


Fig. 2 Sequential evolution of the neuropil starting from a small set of pioneer neurons (*black circles*) G_0, through exponentially increasing neural populations of *grey* and *white circles*, G_1 and G_2, respectively [4]

created neurons. These new neurons grow connections to the existing ones. In the exponential model, an arriving new node makes connection to the existing neurons in its neighborhood within a disk D. A schematic view of the initial stage of development, from G_0, G_1, to G_2 graphs is shown in Fig. 2. It is shown mathematically that the in-degree distribution approximately follows exponential distribution. The edge length distribution follows a power law. The exponent is 2, which equals to the dimension of the space in which the evolving neuronal population lives.

The presence of scale-free structure of connection lengths and degree distribution is an open question in the brain network literature. Robust signal transmission can benefit from anatomical structures reflecting certain scale-free connectivity, and it is just natural to explore if brains exhibit such behavior. A key advantage of scale-free networks is their robustness to random failure of components [2]. However, scale-free structure can become a liability if a mechanism systematically targets and damages hubs. In order to provide resilience to loss of hubs, a deviation from strict scale-free connectivity may be beneficial. Some duplication between highly connected hubs is a potential solution to this problem. Interestingly, the EEGm provide a solution to this question through the key role of pioneer neurons, which remain tightly connected even if they become separated from each other spatially during the brain network exponential expansion; see the highly connected pioneer neurons (black circles) of graph G_2, Fig. 2.

Similar effects have been described as rich club features in brains [3, 14]. The rich-club is a network property that happens when the hubs of a network, the nodes with largest number of neighbors, are densely interconnected. In this sense, the EEGm model is characterized by rich club property due to the key role of pioneering neurons. In addition to rich-club feature, the exponentially exploding brain model (EEGm) reflect important experimentally observed properties of brain networks, including short processing paths, the existence of massive parallel processing paths, and the emergence of hub structures with modular architecture.

4 Conclusions

In this work, we employed neuropercolation model to describe dynamics of the neuropil. We have studied singularities in cortical space-time dynamics using neuropercolation models. Neuropercolation is a suitable tool to model several important properties of cortical neurodynamics observed in ECoG and EEG experiments. The demonstrated effects include:

- Intermittent desychronization of analytic phase across the cortex at brief intervals;
- Rapid propagation of phase gradients across the hemisphere. We posed the hypothesis that rapid desynchronization is the manifestation of the ‘aha’ effect

of cognitive processing, and thus it could be the neural correlate of the conscious broadcast in Baars' global workspace theory;
- The exponentially expanding brain graph model seems to provide a plausible interpretation of the rich club network property of brain networks.

Acknowledgements This work has been supported in part by DARPA Physical Intelligence program, through a contract with Hughes Research Laboratory (HRL), Malibu, CA (Dr. N. Srinivasa, Program Director), and by NSF CRCNS Program 13-11165.

References

1. Baars, B. J. (1997). Inside the Theater of Consciousness: The Workspace of the Mind. Oxford University Press (1997).
2. Barabasi A-L, Albert R. Emergence of scaling in random networks, Science 286: 509–512 (1999).
3. Bullmore, E., Sporns, O. The economy of brain network organization. Nature Reviews Neuroscience, 13(5), 336–349 (2012).
4. Bollobas, B., Kozma, R., Miklos, D. (eds.) Handbook of Large-Scale Random Networks, Springer (2009).
5. Buxton R.B., Introduction to Functional Magnetic Resonance Imaging: Principles and Techniques. Cambridge UK: Cambridge UP (2001).
6. Freeman, W.J., Mass Action in the Nervous System, Academic Press, NY (1975).
7. Freeman, W.J. and R. Quiroga, Imaging Brain Function with EEG, Springer (2013).
8. Kozma, R., M. Puljic, P. Balister, B. Bollobas, and W.J. Freeman, Phase transitions in the neuropercolation model of neural populations with mixed local and non-local interactions. Biol. Cyb. 92 (6): 367–379 (2005).
9. Robert Kozma (2007), Neuropercolation, Scholarpedia, 2(8):1360.
10. Kozma, R., M. Puljic, W.J Freeman, Thermodynamic Model of Criticality in the Cortex, in: Criticality in Neural Systems, D. Plenz and E. Niebur (eds.), Wiley and Sons (2014).
11. Kozma, R., J.J. Davis, W.J. Freeman, Synchronization of De-Synchronization Events Demonstrate Large-Scale Cortical Singularities As Hallmarks of Higher Cognitive Activity, J. Neurosci. and Neuro-Engng. 1(1), 13–23 (2012).
12. Madl, T., Baars, B. J., and Franklin, S. The timing of the cognitive cycle. PloS one, 6(4), e14803 (2011).
13. Von Neumann, J. The Computer and the Brain, Yale University Press, New Haven (1958).
14. Zamora-Lopez, G., Zhou, C., Kurths, J. Exploring brain function from anatomical connectivity. Frontiers in neuroscience, 5 (2011).

Concept Cells in the Human Brain

Rodrigo Quian Quiroga

Abstract In this article I illustrate and briefly describe the main characteristics of concept cells; that means, neurons in the human medial temporal lobe that respond selectively to specific concepts, like a particular person, place or object. Given the converging evidence on the behavior of concept cells and the well-known role of the medial temporal lobe in declarative memory, I argue that concept cells are the building blocks of declarative memory functions.

Keywords Perception • Memory • Neural coding • Concept cells

1 Introduction

The mere act of seeing a friend in a crowd triggers a cascade of brain processes that creates a representation leading to the recognition of our friend, the recollection of details related to him (or her) and the generation of new memories. The study of how neural populations give rise to such exquisite processes has been a subject of active research for decades, and a large number of studies have established that neurons in the ventral visual pathway – going from the primary cortical visual area (V1) to the infero-temporal cortex (IT) – are involved in visual recognition [1, 2], and that the interaction between neocortex and the medial temporal lobe (including the hippocampus, amygdala and the entorhinal, perirhinal and parahippocampal cortices) is critical for the creation of declarative memories, their consolidation and recall [3, 4]. What has remained less studied, however, are the processes and the subjacent neuronal representations that determine how external stimuli impinging our senses lead to the creation of the conceptual, internal representation we use for our thought and the formation of new memories [5].

R.Q. Quiroga (✉)
Centre for Systems Neuroscience, University of Leicester, LE1 7RH Leicester, UK
e-mail: rqqg1@le.ac.uk

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_19

2 Methods

The data comes from patients with pharmacologically intractable epilepsy. Extensive non-invasive monitoring did not yield concordant data corresponding to a single respectable epileptogenic focus and, therefore, they were implanted with chronic depth electrodes for 7–10 days to determine the seizure focus for possible surgical resection. Recordings were performed in the hippocampus, amygdala, entorhinal cortex, and parahippocampal cortex. The electrode locations were based exclusively on clinical criteria. Each electrode probe had a total of nine micro-wires at its end, eight active recording channels and one reference [6]. The differential signal from the micro-wires was amplified using a 64-channel Neuralynx system (Tucson, Arizona), filtered between 1 and 9,000 Hz and sampled at 28 kHz. Subjects sat in bed, facing a laptop computer where pictures, text or sounds were presented [5,7–9]. Each recording session lasted about 30 min. Spike detection and sorting was carried out using 'Wave_Clus' an adaptive and stochastic algorithm [10].

3 Results

Figure 1 shows a representative single unit in the entorhinal cortex, which summarizes the main findings of the recordings in the human medial temporal lobe. The neuron responded selectively not only to three different pictures of Luke Skywalker, but also to the presentation of his name written in the computer screen and pronounced by a computer synthetized voice. For the sound responses, in this case two different computer synthesized voices – one male and one female – were tried. Both voices elicited significant responses. This unit also had a significant

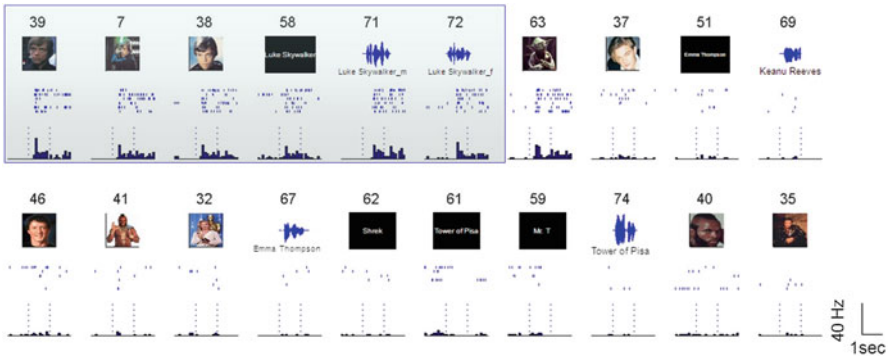


Fig. 1 Responses of a neuron in the enthorinal cortex that fired to Luke Skywalker (highlighted stimuli) and Yoda (picture nr. 63). Vertical dotted lines in the peri-stimulus time histograms denote onset and offset of the stimuli, 1 s apart. For space reasons only the largest 20 (of 76) responses are shown

response to Yoda – only one single picture of Yoda was presented – another well know character of the movie 'Star Wars', and did not respond to the other ∼70 pictures and names presented. So, we first conclude that the neuron selectively responds to the concept of 'Luke' (and Yoda) and not to particular details of a picture shown and second, that these neurons can fire to related concepts, a finding of relevance given the importance of associations for memory functions.

4 Discussion

As illustrated with the example of Fig. 1, concept cells have a sparse, abstract and explicit representation of the meaning of the stimulus. By "meaning" we refer to the subjective, internal representation created by the subject for memory functions (i.e. how something will be thought of, and potentially remembered). This conclusion is based on the finding on concept cells and the well-established role of the medial temporal lobe in declarative memory [3, 4]. It is therefore not surprising that concept cells tend to encode personally-relevant items [11] – i.e. whatever the subject may mind to store in memory. The representation by concept cells is sparse because they get activated by relatively very few of the presented stimuli [12]; it is abstract because a concept cell fires to the person or object presented and not to particular details of the picture (or other type of stimulus) shown [7, 9] and it is explicit because from the firing of a concept cell we can tell whether the particular concept is being shown [12] or even thought [13]. The finding of concept cells firing to related concepts [5, 9, 14] is in line with the key role of associations in the formation of new declarative memories. Concept cells may then be the key neural substrate for the formation of new declarative (and specifically episodic) memories.

Acknowledgments We acknowledge support from EPSRC, MRC and BBSRC.

References

1. Logothetis, N. K. and D. L. Sheinberg (1996). "Visual object recognition." Annu Rev Neurosci **19**: 577–621.
2. Tanaka, K. (1996). "Inferotemporal cortex and object vision." Annual Review on Neuroscience **19**: 109–139.
3. Squire, L. and S. Zola-Morgan (1991). "The medial temporal lobe memory system." Science **253**: 1380–1386.
4. Moscovitch, M., R. S. Rosenbaum, et al. (2005). "Functional neuroanatomy of remote episodic, semantic and spatial memory: a unified account based on multiple trace theory." Journal of Anatomy **207**: 35–66.
5. Quian Quiroga, R. (2012). "Concept cells: The building blocks of declarative memory functions." Nature Reviews Neuroscience **13**: 587–597.
6. Fried, I., C. L. Wilson, et al. (1999). "Cerebral microdialysis combined with single-neuron and electroencephalographic recording in neurosurgical patients – Technical note." Journal of Neurosurgery **91**: 697–705.

7. Quian Quiroga, R., L. Reddy, et al. (2005). "Invariant visual representation by single neurons in the human brain." Nature **435**(7045): 1102–1107.
8. Quian Quiroga, R. R., L; Koch, K; Fried, I. (2005). Decoding visual responses from single neurons in the human brain. Soc. Neurosc. Abstr. Prog. Nr.
9. Quian Quiroga, R., A. Kraskov, et al. (2009). "Explicit encoding of multimodal percepts by single neurons in the human brain." Current Biology **19**: 1308–1313.
10. Quian Quiroga, R., Z. Nadasdy, et al. (2004). "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering." Neural Comput **16**(8): 1661–1687.
11. Viskontas, I., R. Quian Quiroga, et al. (2009). "Human medial temporal lobe neurons respond preferentially to personally-relevant images." Proceedings of the National Academy of Sciences USA **106**: 21329–21334.
12. Quian Quiroga, R., L. Reddy, et al. (2007). "Decoding Visual Inputs From Multiple Neurons in the Human Temporal Lobe." J Neurophysiol **98**(4): 1997–2007.
13. Cerf, M., N. Thiruvengadam, et al. (2010). "On-line, voluntary control of human temporal lobe neurons." Nature **467**: 1104–1108.
14. Quian Quiroga, R. and G. Kreiman (2010). "Measuring sparseness in the brain: comment on Bowers (2009)." Psychological Review **117**: 291–299.

EEG Spatiotemporal Pattern Classification of the Stimuli on Different Fingers

Tinglin Zhang, Lengshi Dai, You Wang, Walter J. Freeman, and Guang Li

Abstract Synchronous activity of brain is essential for perception. To study the spatiotemporal pattern of a particular cortex area, a high definition electrode array was made to record the EEG while stimuli on different fingers. Analytic amplitude and analytic phase of high definition EEG estimated by the Hilbert Transform represented the local synchrony of brain. Spatiotemporal patterns of stimuli on different fingers can be extracted and classified by a specific algorithm. BP network combined with the KIII model, a bionic olfactory system model, was applied to achieve a better classification effect. It was proved that the patterns were nonlocal and spread over the whole array.

Keywords Amplitude modulation • Phase modulation • High definition EEG • Cone-fitting • Spatiotemporal pattern

1 Introduction

Synchrony among widely distributed neurons in large numbers is related to emergence of spatial structure in cortical activity [1]. Specific spatial patterns of cortical activity in various areas of cortex after differing perceptive stimulation will emerge by modifying cortical background activity [2]. The cortex will return to unpatterned state, waiting for other stimuli. Discrete spatial patterns, resembling

T. Zhang (✉) • Y. Wang • G. Li
Department of Control Science and Engineering, State Key Lab of Industrial Control Technology, Zhejiang University, 38 Zheda Road, Hangzhou 310027, People's Republic of China
e-mail: ztlyoyo@163.com; king_wy@zju.edu.cn; guangli@cbeis.zju.edu.cn

L. Dai
Department of Biomedical Engineering, Boston University, 44 Cummington Street, Boston, MA 02215, USA
e-mail: all-stardai@163.com

W.J. Freeman
Division of Neurobiology, Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA
e-mail: dfreeman@berkeley.edu

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_20

cinematographic frames, occur with carrier wave in the beta (13–30 Hz) and gamma (30–80 Hz) ranges. Synchrony activity in a local area of brain which may be associated with particular perception could be estimated by phase and amplitude of multiple-channel EEG records. As followed state transition in spatial patterns begin with abrupt phase resetting and dramatic increased pattern amplitude, phase modulation (PM) and amplitude modulation (AM) of shared waveforms in the high-density electrode array were constituted [1–3]. There are lots of studies on EEG synchronization [1–9], but more thorough studies should be carried on to find what the synchronous activity is in a small local area of brain by scalp EEG instead of invasive ECoG. It is generally known that the somatic sensory cortex, width is only in millimeters and length in centimeters is responsible for sensation of different parts of the body. The areas of sensory cortex related to different fingers are very small and connect one by one, and if synchronous information can be extracted from high-density EEG signals of this area when different fingers are stimulated, perceptive patterns of fingers could be distinguished. The aim of this paper was to study spatiotemporal patterns of the stimuli on different fingers by measuring the synchrony of local scalp EEG, and distinguish the patterns of the stimuli on different fingers.

2 Methods

EEGs were recorded from six right handed male adults aged 22–26 with an 8×8 electrode array made of 1.5 mm Ag/AgCl wires. The interelectrode distance was 6 mm.

The scalp of subject was cleaned and the array was bound on the left head where corresponding to somatosensory cortex that represented fingers of right hand. Two identical rubber bands were stretch to the same height to hit either index finger or little finger of right hand randomly. There was 1-min interval between two mechanical stimuli. One second before and 2 s after the stimulation EEGs were recorded. The subject was asked to close his eyes, and keep conscious and relaxing during the whole process. Forty trials hitting on each finger were accomplished. Another experiment was also designed to test if the EEG patterns could be classified with decreased distance of fingers. Index finger and ring finger of right hand were each stimulated for 40 trials with EEGs recorded.

EEG data was recorded by a Neuroscan NuAmps Digital Amplifier, referenced to the left mastoid, and sampled at 1,000 Hz with an analog filtering pass-band of 0.1–100 Hz.

Bad channels of raw EEG were replaced by adjacent signals. EEG data preprocessing included removing channel bias, normalization by dividing global standard deviation, low-pass spatial filtering and band-pass temporal filtering. Settings of spatial and temporal filters could affect the extraction of patterns. It was necessary to reduce spatial noise that was induced by high-density electrode array, therefore,

a Gussian filter [1] was applied to do spatial low pass filtering at 0.43 cycles/cm on the basis of the spatial power spectral density (PSD_x). PSD_x was calculated by 2D FFT in MATLAB software on the 64 EEG signals at each time sample in each row and column [8], and then final result was got from the mean value of all time samples in one trial. The temporal filtering, using Parks-McClellan algorithm, was implemented by function firpm in MATLAB. Different bandwidth varying from 3 to 10 Hz steps in beta band (13–30 Hz) and from 5 to 20 Hz steps in low gamma band (30–80 Hz) were explored for searching stable frames. The Hilbert transform was applied to the filtered signal to get the analytic amplitude $A_i(t)$ and analytic phase $\phi_i(t)$ of the analytic signal $v_i(t)$.

$$V_j(t) = v_j(t) + i\,u_j(t), \tag{1}$$

$$A_j(t) = \sqrt{v_j^2(t) + u_j^2(t)}, \tag{2}$$

$$\phi_j = a\tan\left(\frac{v_j(t)}{u_j(t)}\right), j = 1, \ldots, 64. \tag{3}$$

Spatial structure of the analytic phase was fitted to a cone surface due to the similarity [8], and the phase cones constitute the PM patterns. The 8×8 phase surface at each time sample was fitted to in Eq. 4.

$$\phi(t) = \phi_0(t) + \gamma(t)\sqrt{[x_i - x_0]^2 + [y_i - y_0]^2}. \tag{4}$$

Where $\phi(t)$ was the vertical offset of the cone apex, x_j, y_j was the cone apex, x_0, y_0 was the center of the electrode array, and $\gamma(t)$ was the slope of the cone.

Stable frames were obtained by applying technical criteria and physiological criteria to these parameters of phase cones and frames. There was detailed description of the criteria in [8–10]. Analytic power, which was square of analytic amplitude, specified the spatial pattern of amplitude modulation. After analytic power was subtracted by spatial ensemble average and normalized by dividing temporal standard deviation, 64×1 feature vector of each stable frame was extracted as root mean square of values over all time samples of the stable frame on each channel. Each finger was stimulated for 40 times in one experiment so that each feature matrix was 64×40.

Feature vectors were classified by a three-layer BP network with optimized parameters. There were nine neurons on the hidden layer and function traingda() in MATLAB was implemented as learning algorithm. The training set consisted of half of the samples and the other half were used as the testing set. They were then switched for crosscheck. The final result was computing the average for results of test and crosscheck. KIII model, a bionic olfactory system model, has many dynamic behaviors such as limit cycle, quasi-periodic oscillation and chaos [11,12].

It has been widely applied to pattern recognition due to its reliable pattern learning. Another classification algorithm for comparison was using KIII model to preprocess the feature vectors, afterwards, the output be classified by the same BP network applied before. Four samples of each class were used to train KIII model twice for pattern learning and others were classified by BP after preprocessed. Sixteen samples were selected as training set and 20 samples as testing set in each class for BP network.

3 Results

It was necessary to remove spatial noise as the high-density EEGs were measured locally (Fig. 1). The phase surfaces were much smoother with preprocessing of the appropriate low-pass spatial filtering than those with no spatial filtering preprocessing (Fig. 2).

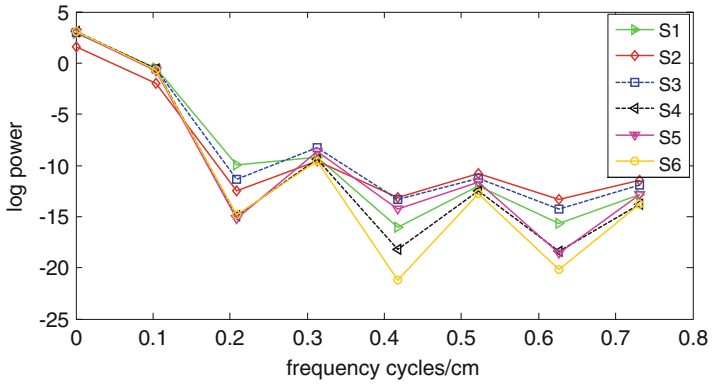


Fig. 1 Mean spatial power spectral density for each subject

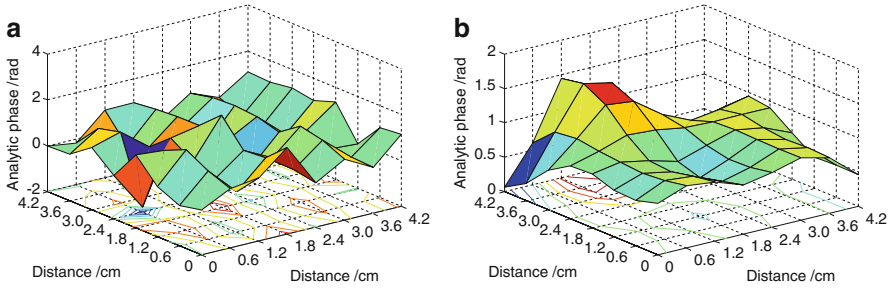


Fig. 2 Spatial phase structure obtained (**a**) before spatial filtering and (**b**) after spatial filtering

Table 1 Classification accuracy of stimuli on index finger and little finger in the first 3 stable frames

	Classifier	
Frame	BP	KIII + BP
First	71.4 ± 5.5	88.4 ± 6.2
Second	67.5 ± 4.2	73.2 ± 3.6
Third	68.2 ± 5.9	78.6 ± 5.7

Table 2 Classification accuracy of stimuli on index finger and ring finger in the first 3 stable frames

	Classifier	
Frame	BP	KIII + BP
First	72.5 ± 3.9	78.4 ± 2.3
Second	65.8 ± 3.3	69.2 ± 2.6
Third	69.8 ± 4.2	73.6 ± 5.0

Optimized preprocessing could improve classification rate. Optimal bandwidth of the temporal filter was selected repeatedly as 15–24 Hz and cutoff frequency of the spatial filter was 0.43 cycles/cm to get high classification rate.

The patterns of the stimuli on index finger and little finger of right hand were recognized only in beta range in this study as well as index finger and ring finger of right hand with shorter distance. Patterns in gamma band could not be classified.

The classifiable feature vectors were extracted from the first three stable frames appeared within 500 ms after stimulation. Each of the four subjects S1, S2, S3 and S4 attended three times the stimuli on index finger and little finger experiments. Mean of these 12 classification ratios and standard error in the first 3 stable frames emerged after stimulation was shown in Table 1. The classification rate was higher by applying the BP network preprocessed by KIII model (KIII + BP) than a single BP. Classification rate of feature vectors extracted from stable frames emerged before stimulation was approximate 50 % which represented the reliability and validity of feature extraction algorithm.

Each of the three subjects S4, S5 and S6 attended three times the stimuli on index finger and ring finger experiments. Mean of these nine classification ratios and standard error in the first three stable frames emerged after stimulation was shown in Table 2.

However, it was difficult to classify the patterns with data from different experiments, even the data was got from the same subject.

The contribution of EEGs recorded from different electrodes to spatiotemporal pattern classification was tested by deleting groups of electrodes for all subjects [13] to determine whether the information in the EEG used for classification was localized to a few channels or not. Figure 3 shows the mean classification rate by applying KIII + BP after the 40 times random removing groups of 4, 8, 16, 24, 32, 40, 48 and 56 random channels. Classification rate decreased with the increasing deleted channels which meant the spatial patterns were nonlocal and spread over the whole array of channels.

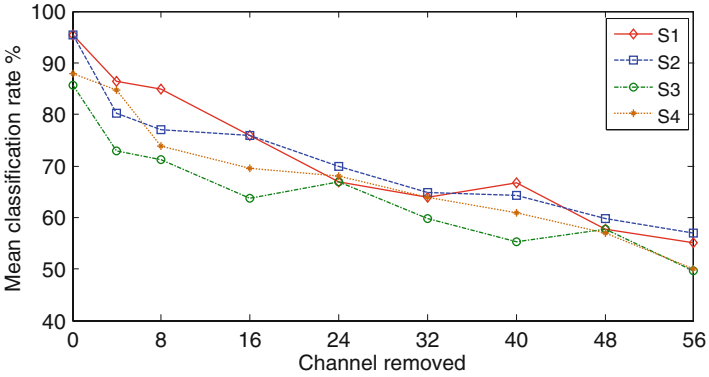


Fig. 3 Effects on classification rate of removing increasing numbers of channels for four subjects. Examples were taken from the first post-stimulus frame in stimuli on index finger and little finger experiments

4 Conclusion

Synchronous activity of brain in a special area, which could be measured by multiple-channel scalp EEG records, can be used to extract and classify the stimulus patterns of different fingers. It proved once again that perceptive pattern of human brain is modulated by phase and amplitude information. The patterns were nonlocal and spread over the whole array of channels which illustrated that they were perceptual.

References

1. Freeman, WJ: Origin, Structure, and Role of Background EEG Activity. Part 1. Analytic Amplitude. Clinical Neurophysiology. 115.9 (2004) 2077–88.
2. Freeman, WJ, Holmes, MD: Metastability, instability, and state transition in neocortex. Neural Networks. 18 (2005) 497–504.
3. Freeman, WJ: Origin, Structure, and Role of Background EEG Activity. Part 3. Neural Frame Classification. Clinical Neurophysiology 116.5 (2005) 1118–29.
4. Bhattacharya, J and Petsche, H et al: Long-range synchrony in the gamma band: role in music perception. Neuroscience. 21(2001) 6329–37.
5. Engel, AK and Kreiter AK et al: Synchronization of oscillatory neuronal responses between striate and extrastriate visual cortical areas of the cat. Proc Natl Acad Sci USA. 88 (1991b) 6048–52.
6. Freeman WJ, Rogers LJ: Fine temporal resolution of analytic phase reveals episodic synchronization by state transitions in gamma EEGs. J Neurophysiol 87 (2002) 937–45.
7. Pockett S, Bold GEJ, Freeman, WJ: EEG synchrony during a perceptual-cognitive task: Widespread phase synchrony at all frequencies. Clinical Neurophysiol. 120 (2009) 695–708.
8. Freeman, WJ: Origin, Structure, and Role of Background EEG Activity. Part 2. Analytic Phase. Clinical Neurophysiology 115.9 (2004): 2089–107.

9. Freeman, WJ and Holmes, MD et al: Fine spatiotemporal structure of phase in human intracranial EEG. Clinical Neurophysiology. 117(2006) 1228–1243.
10. Freeman, WJ and Holmes, MD et al: Dynamics of human neocortex that optimizes its stability and flexibility. International Journal of Intelligent Systems. 21(2006) 881–901.
11. Freeman, WJ: Mass Action in the Nervous System. New York: Academic Press, 1975.
12. Yao, Y and Freeman WJ: Model of biological pattern recognition with spatially chaotic dynamics. Neural Networks. 3.2(1990) 153–170.
13. Ruiz, Y and Pockett S et al: A method to study global spatial patterns related to sensory perception in scalp EEG. J Neurosci Methods. 191(2010) 110–118.

Part IV
Neurodynamics of Sensory Systems

Time Varying VEP Evaluation as a Prediction of Vision Fatigue Using Stimulated Brain-Computer Interface

Teng Cao, Chi Man Wong, Feng Wan, and Yong Hu

Abstract During the operation of the visual stimulated brain-computer interfaces (BCIs), users suffer fatigue when gazing at the stimulus. Loss of attention and decreased vigilance caused by fatigue impair the users' cognitive capabilities, which consequently degrade the EEG quality and system performance. Cognitive neurodynamics measured by EEG activity are related to the generalized performance on cognitive tasks and mental efforts. Therefore, time varying EEG spectral analysis is proposed as an objective method to evaluate fatigue during the cognitive process in the visual stimulated BCIs. The results shows that the increase in θ, α and $(\theta + \alpha)/\beta$, as well as the decrease in θ/α are associated with the increasing fatigue level during the cognitive process.

Keywords BCI • SSVEP • Fatigue • EEG • Cognitive neurodynamics

1 Introduction

Brain-computer interface (BCI) is a technology that provides a direct communication pathway between the brain and an external device which is independent of normal peripheral nervous and muscular systems [1]. Visual stimulation is the most commonly used input for practical BCI systems. However, the user of visual stimulated BCI usually claims vision fatigue. For example, steady-state visual evoked potential (SSVEP) requires the user to gaze at flashing stimuli, resulting in discomfort and tiredness, and consequently reduce the performance. Especially for stimuli with high brightness and contrast, users will easily feel extremely drowsy and have difficulty in concentration. SSVEP-based BCI, as an attention dependent system, requires significant effort to focus on the flashing stimuli to

T. Cao • C.M. Wong • F. Wan (✉)
Department of Electrical and Computer Engineering, University of Macau, Macau, People's Republic of China
e-mail: fwan@umac.mo

Y. Hu
Department of Orthopaedics and Traumatology, The University of Hong Kong, Pokfulam, Hong Kong, People's Republic of China

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_21

produce sufficiently strong SSVEP signals, and the users may easily get fatigue. The cognitive neurodynamics process of the visual system responding to an SSVEP signal is easily affected by various mental states, attention, fatigue and so on [2, 3]. Loss of attention caused by fatigue will reduce the cognitive performance, consequently reduce the SSVEP quality and system performance. There is a real demand to establish an objective and online assessment to evaluate the fatigue during cognitive process in SSVEP-based BCIs.

Previous studies proposed several questionnaires or subjects' description to indicate the fatigue degree [4–7]. However, they are quit raw, subjective, inaccurate and difficult in on-line measurement. Cognitive neurodynamics measured by EEG activity are related to generalized performance on cognitive tasks and mental efforts. More specifically, the low cognitive capacity and decreased arousal level are associated with global increase in θ and α [8–10]. Therefore, the dynamic change of EEG activity would be an indicator for fatigue. Time varying EEG spectral in theta and alpha activity is proposed as an objective method to evaluate fatigue during cognitive process in SSVEP-based BCIs.

2 Methods

An LCD monitor was used as the visual stimulator (ViewSonic 22″, refresh rate 120 Hz, 1,680 × 1,050 pixel resolution) and the stimulus was programmed under Microsoft Visual Studio 2010 using Microsoft DirectX SDK (June 2010). There was a symbol "+" in the center of the flashing stimulus to keep the subjects focusing on the target. The stimulus was flashing for 3 s for 30 trials, and 2 s pause was set between each trial. EEG was collected from Oz channel by an amplifier (g.USBamp, Guger Technologies, Graz, Austria) and were filtered by a 50 Hz notch filter and a 0.5–60 Hz band-pass filter. The amplitude of frequency bands δ, θ, α and β were calculated in each trial by using fast Fourier transforms (FFT). The average amplitude of each frequency band was selected as an index of this frequency band, and was used to calculate the EEG ratio indices θ/α and $(\theta + \alpha)/\beta$. Eleven subjects (aged from 22 to 28 years old, six males and five females) participated in the experiments and were asked to complete a validate self-reported fatigue questionnaire before and after the task work, which was called the Chalder Fatigue Scale (CFS) [11]. The CFS had high reliability and validity, and it was used in this study as a standard reference for fatigue levels. A linear regression analysis was employed to compare the association of EEG indices and fatigue states.

3 Results

The CFS score was significantly increased ($p < 0.001$) with pre-mean CFS score = 14.91, standard deviation (SD) = 1.22; post-mean CFS score = 25.73, SD = 4.98; which indicated the fatigue level was significantly increased after

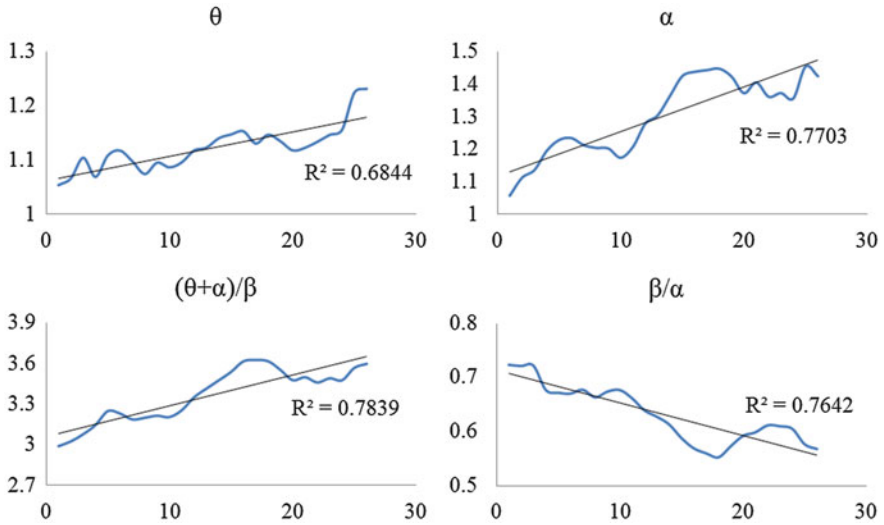


Fig. 1 Time-varying EEG features during cognitive process. *Vertical axis* amplitude (μV), *horizontal axis* number of trials. *Blue curve* time-varying EEG features, *black straight line* linear fitting line

SSVEP-based BCI experiments. For the EEG evaluation of fatigue, it would find the strong association between CFS and the EEG indices θ and α, as well as ratio indices θ/α and (θ + α)/β. Figure 1 presented the average value over all subjects for time varying EEG indices θ and α, as well as ratio indices θ/α and (θ + α)/β, which demonstrated the dynamic prediction for fatigue during cognitive process. The results indicated significant correlations between EEG indices in θ, α and the number of trials, with the R^2 values equal to 0.68 ($p < 0.001$) and 0.77 ($p < 0.001$), respectively. Both the ratio indices θ/α and (θ + α)/β were found significantly correlated with the number of trials, with the R^2 values equal to 0.60 ($p < 0.001$) and 0.78 ($p < 0.001$) respectively. Therefore, these results suggest that the increase in θ, α and (θ + α)/β, as well as the decrease in θ/α are associated with the increasing fatigue level during cognitive process.

4 Conclusion

The time varying EEG spectral analysis is proposed as an objective approach to evaluate fatigue during cognitive process in SSVEP-based BCIs. The proposed approach can provide a real-time evaluation of the fatigue with objective and quantitative measurement. The promising result suggests the potential of objective evaluation of fatigue by dynamic measurement in EEG indices θ and α, as well as ratio indices θ/α, (θ + α)/β. The time varying EEG activity would predict the fatigue

immediately, which may provide a useful online assessment of fatigue. The increase in θ, α and $(\theta + \alpha)/\beta$, as well as the decrease in θ/α are associated with the increasing fatigue level during cognitive process. In addition, this method can be used for optimal selection of visual stimuli parameters (e.g., visual stimulus frequency, duty cycle, color, etc.) to design a user-friendly BCI system, which cause less fatigue in further study.

Acknowledgments This work was supported in part by the Macau Science and Technology Development Fund (Grant FDCT/036/2009/A) and the University of Macau Research Fund (Grants MYRG2014-00174-FST, MYRG139(Y1-L2)-FST11-WF, MYRG079(Y1-L2)-FST12-VMI and MYRG069(Y1-L2)-FST13-WF).

References

1. Wolpaw J R, Birbaumer N, McFarland D J, Pfurtscheller G and Vaughan T M 2002 Brain-computer interfaces for communication and Control *Clin. Neurophysiol.* 113 767–91
2. Eriksen C W and James J D S 1986 Visual attention within and around the field of focal attention: a zoom lens model *Percept. Psychophys.* 40 225–40
3. Yamaguchi S, Tsuchiya H and Kobayashi S 1994 Electrooencephalographic activity associated with shifts of visuospatial attention *Brain* 117 553–62
4. Allison B, Lüth T, Valbunena D, Teymourian A, Volosyak I and Gräser A 2010 BCI demographics: how many (and what kind of) people can use an SSVEP BCI? *IEEE Trans. Neural Syst. Rehabil. Eng.* 18 107–16
5. Bieger J and Molia G G 2010 Light stimulation properties to influence brain activity: a brain-computer interface application *Philips Research Technical Note TN-2010-00315*
6. Volosyak I, Valbuena D, Lüth T, Malechka T and Gräser A 2011 BCI demographics II: how many (and what kind of) people can use a high-frequency SSVEP BCI? *IEEE Trans. Neural Syst. Rehabil. Eng.* 19 232–9
7. Allison B Z, Brunner C, Altstätter C, Wagner I C, Grissmann S and Neuper C 2012 A hybrid ERD/SSVEP BCI for continuous simultaneous two dimensional cursor control *J. Neurosci. Methods* 209 299–307
8. Lehmann D, Grass P and Meier B 1995 Spontaneous conscious covert cognition states and brain electric spectral states in canonical correlations *Int. J. Psychophysiol.* 19 41–52
9. Klimesch W 1999 EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis *Brain Res. Rev.* 29 169–195
10. Lafrance C and Dumont M 2000 Diurnal variations in the waking EEG: comparisons with sleep latencies and subjective alertness *J. Sleep Res.* 9 243–8
11. Chalder T, Berelowitz G, Pawlikowska T, Watts L, Wessely S, Wright D and Wallace E P 1993 Development of a fatigue scale *J. Psychosom. Res.* 37 147–53

Spike Synchronization Analysis in a Network Model of the Olfactory Bulb

Ying Du and Rubin Wang

Abstract In the olfactory system, both the temporal spike structure and spatial distribution of neuronal activity are important for processing odor information. This paper simulates the firing activity of olfactory mitral cell. By varying some key parameters of a biophysically-detailed spiking neuronal model, it is shown that the spike train of single neuron can exhibit various firing patterns. Synchronization in coupled neurons is also investigated as the coupling strength varying in the situation of two neurons and network. It is illustrated that the coupled neurons can exhibit different types of pattern when the coupling strength varies. These results may be instructive to understand information transmission in olfactory system.

Keywords Olfactory • Spike train • Maximal conductance • Synchronization • Network

1 Introduction

In the olfactory bulb, the temporal structure of neuronal activity appears to be important for processing odor information [1]. The temporal patterns produced by three simultaneously sampled projection neurons in the locust antennal lobe indicate that each odor evokes a specific temporal activity pattern [2]. Mapping of the sensory inputs has revealed that each odorant produces a reproducible spatial pattern of activation in the glomerular layer of the bulb.

Some experiments suggest that the time course of dendrodendritic inhibition is dependent on the network connectivity as well as on the intrinsic parameters on the synapses [3, 4]. In response to simulated odor stimulation, strongly activated mitral cells tend to suppress neighboring cells, the mitral cells readily synchronize their firing, and increasing the stimulus intensity increases the degree of synchronization. Preliminary experiments also suggest that slow temporal changes in the degree of synchronization are more useful in distinguishing between very similar odorants than the spatial distribution of mean firing rate. So it is necessary to explore the

Y. Du (✉) • R. Wang
Institute for Cognitive Neurodynamics, East China University of Science and Technology, Shanghai, People's Republic of China
e-mail: du.ying.buaa@gmail.com

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_22

effect of some key parameters such as maximal conductances of the olfactory model on temporal spike train pattern.

2 Method

We used a detailed, realistic model of the mitral-granule cell circuits in the olfactory bulb that introduced by Bhalla-Bower [5] to investigate the spatio-temporal processing of odor information. All parameter values are taken from the Senselab databank [6]. The effect of multi-maximal conductance to the spike train pattern of olfactory model was considered. If changing the maximal conductances of the Nap and Ks currents simultaneously, the variation tendency of spike train patterns is different to the case that only changes one parameter [7]. As shown in the Fig. 1 when two

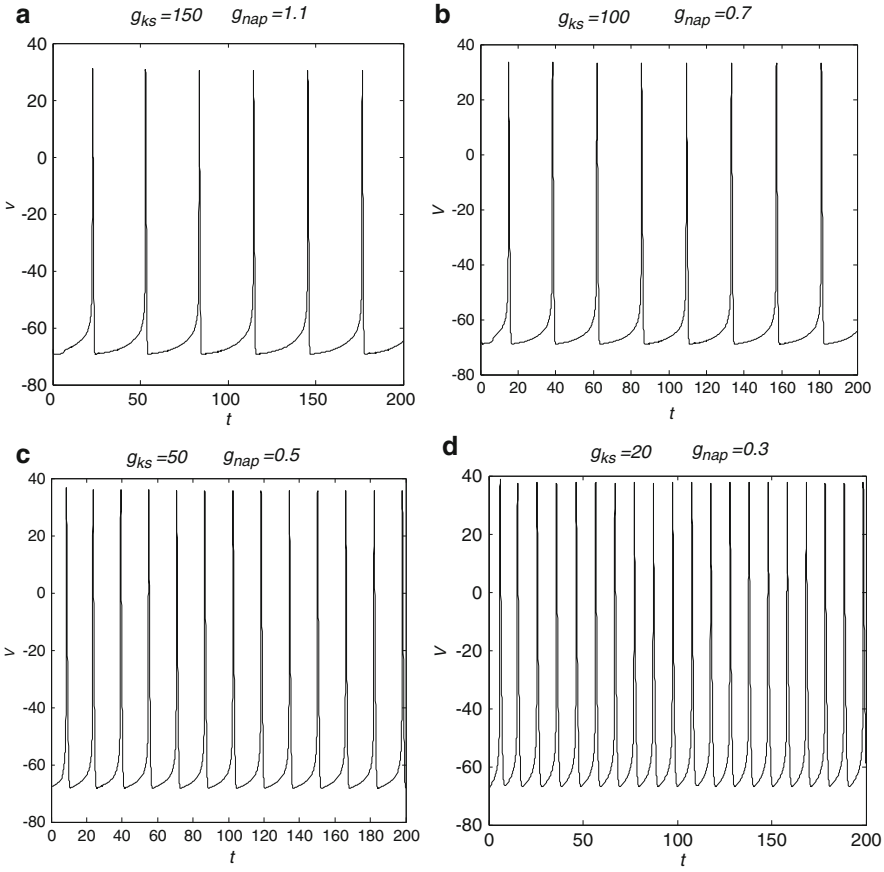


Fig. 1 Different spiking frequencies with the changing of g_{Nap} and g_{Ks}. (**a**) $g_{Nap} = 150\ \mu S/cm^2$, $g_{Ks} = 1.1\ \mu S/cm^2$, (**b**) $g_{Nap} = 100\ \mu S/cm^2$, $g_{Ks} = 0.7\ \mu S/cm^2$, (**c**) $g_{Nap} = 50\ \mu S/cm^2$, $g_{Ks} = 0.5\ \mu S/cm^2$, (**d**) $g_{Nap} = 20\ \mu S/cm^2$, $g_{Ks} = 0.3\ \mu S/cm^2$

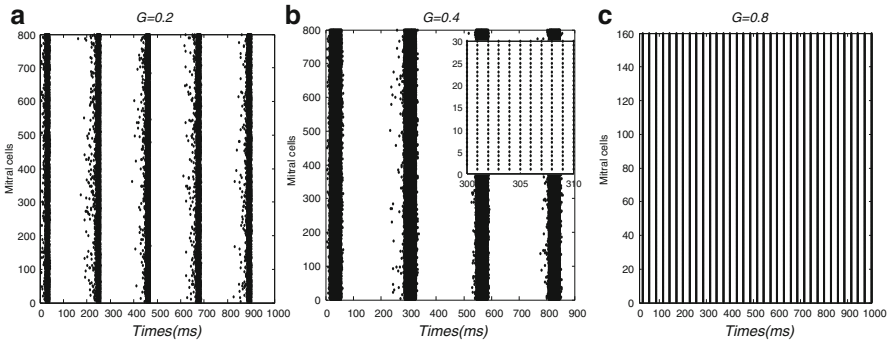


Fig. 2 Simulation of a network of 800 coupled spiking neurons. (**a**) The spike train pattern of network neurons when G is taken as 0.2. (**b**) The situation of G = 0.4. (**c**) Completely synchronization when G is taken as 0.8, right corner is enlarged situation

maximal conductance decrease simultaneously, the frequencies of olfactory spike firing are increasing, This state is similar to the case of changing single g_{Nap}, but if the value of g_{Nap} is smaller than 30 $\mu S/cm^2$, cell do not fire at all, only if the g_{Ks} decrease with the g_{Nap} simultaneously, the neuron can still generate membrane potential.

Then we use this olfactory model to simulate a sparse network of 800 spiking cortical neurons with 800×799 synaptic connections. Spatiotemporal patterns of all neurons with synaptic coupling in the complex networks are displayed in Fig. 2. We simulate a network of globally connected 800 neurons in real time, whose coupling strengths are chosen as different values. The coupled neurons were found located in different states. The synaptic connection weights between the neurons are given by the matrix $G = (g_{i,j})$, so that firing of the jth neuron instantaneously changes variable V_i by $g_{i,j}$. In the Fig. 2a–c, neurons at different synaptic coupling weights $G = (g_{i,j})$ are considered. At first, G = 0.2 is taken, that is, all $g_{i,j}$ are taken the same as 0.2. It is shown that the neurons organize into assemblies and exhibit collective rhythmic behavior. With the coupling weights increasing, we find the width of the assemblies is becoming larger (Fig. 2b), the coupled neurons could achieve local synchronization inside the assemblies as showed in the enlarged picture of Fig. 2b. When the coupling weights is added to G = 0.8, the coupled neurons eventually achieve synchronization as shown in Fig. 2c. It is evident that the synaptic coupling weight is increased and the synchronous action between neurons is strengthened.

3 Conclusion

The maximal conductances of the olfactory model have obviously effects on the spike train pattern: the frequency of spike train firing is increasing with the g_{Ks} increasing, while the frequency of spike train firing is decreasing when the g_{Ks} is

increased. With changing the maximal conductances of the *Nap* and *Ks* currents simultaneously, the variation tendency of spike train patterns is similar to the case that changing single g_{Nap}, while opposite to the case of only changing g_{Ks}. It was also shown that the coupled neurons could achieve synchronization with the variation of coupling strength in the network situation. These analyses maybe helpful to understood the integration of the many factors influencing the construction and transformation of odor representations.

Acknowledgments This work was supported by the National Natural Science Foundation of China (No.11002055 and 11232005).

References

1. Hutcheon B, Yarom Y (2000). Resonance, oscillation and the intrinsic frequency preferences of neurons. *Trends Neuroscience,* 23:216–222.
2. Jorge N.B, Leslie M.K and Nancy J.K (2009). Biophysical model for gamma rhythms in the olfactory bulb via subthreshold oscillations. *PNAS* 106(51):21954–21959
3. Bazhenov M, Stopfer M, Rabinovich M, Abarbanel HDI, Sejnovski T, Laurent G (2001) *Neuron* 30,569–581
4. Schoppa NE, Kinzie JM, Saharay, Segerson TP, and Westbrook GL (1998) Dendrodendritic inhibition in the olfactory bulb is driven by NMDA receptors. *Journal of Neuroscience*, 18:6790–6802
5. Bhalla US., Bower JM (1993) Exploring parameter space in detailed single cell models: simulations of the mitral and granule cells of the olfactory bulb. *Journal of Neurophysiology*. 69:1948–1965
6. Senselab databank, {http://senselab.med.yale.edu}
7. Du Y, Wang RB, Qu JY (2013), Spike train pattern and firing synchronization in a model of the olfactory mitral cell. *Lecture Notes in Computer Science.*7951:1

Laterality of Gamma-Oscillations in Primate Supplementary Motor Area During Performance of Visually-Guided Movements

Ryosuke Hosaka, Toshi Nakajima, Kazuyuki Aihara, Yoko Yamaguchi, and Hajime Mushiake

Abstract The neurons in the motor cortex show lateralization depending on the arm to use. To investigate if local field potential (LFP) oscillations change with contralateral and ipsilateral arm use, we analyzed the power of LFP in supplementary motor areas (SMA) and pre-SMA while animals performed a delayed-response arm use task. LFP power changed with the laterality of the arm use, but that it was frequency dependent. Specifically, power in the gamma range increased during contralateral arm use, while beta power increased with ipsilateral arm use. Our data therefore suggest that lateralized movement is executed by gamma oscillations and unit activities in the contralateral hemisphere, and is modulated by beta frequency activities in the ipsilateral hemisphere.

Keywords Motor cortex • SMA • Local field potentials • Delayed-response • Gamma oscillations

1 Introduction

Neuronal activity in the motor cortex alters in relation to the laterality of the effector [1]. Over 80 % of primary motor cortex neurons increase activity before and during movements performed using the contralateral effector. In the medial

R. Hosaka (✉)
Department of Applied Mathematics, Fukuoka University, Fukuoka, Japan
e-mail: hosaka@fukuoka-u.ac.jp

T. Nakajima • H. Mushiake
Department of Physiology, Tohoku University School of Medicine, 2-1 Seiryo-cho, Aoba-ku, Sendai 980-8575, Japan

K. Aihara
Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

Y. Yamaguchi
Neuroinformatics Japan Center, RIKEN Brain Science Institute, 2-1, Hirosawa, Wako-shi, Saitama 351-0198, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_23

motor area (such as the supplementary motor area (SMA) and pre-SMA), 31 % of neurons are contralateral neurons, and only 4 % neurons are ipsilateral neurons. In the premotor cortex, the contralateral and ipsilateral neurons constitute 13 and 9 % of total neurons, respectively. As such, movement representation is lateralized in the medial and lateral motor cortex.

Populations of neurons co-activate repeatedly for short temporal epochs, resulting in synchronized oscillation expressed by unit activity and local field potentials (LFP) [2]. Although EEG and MRI studies revealed lateralization of oscillatory activity of the lateral motor cortex, the laterality in the medial motor cortex is less well defined because the spatial resolutions of the EEG and MRI are too broad to accurately separate left and right medial motor cortices. We therefore examined the lateralization of beta and gamma oscillations in the medial motor cortex using the LFP of the SMA and pre-SMA.

2 Methods

Two Japanese monkeys (*Macaca fuscata*, M and N, 6.5 and 5.8 kg) were used, and were cared for in accordance with the Guiding Principles for the Care and Use of Laboratory Animals of the National Institutes of Health. During experimental sessions, the monkeys were seated in a primate chair facing a screen. The monkeys held manupilandums in both forelimbs, and were trained to perform a delayed-response forelimb movement task (Fig. 1). Trials started with the presentation of a white spot of light (fixation point) in the center of a screen, which the monkeys fixated on. They were then required to place the manipulandums in a neutral position. After a 500-ms holding period, one of four colored squares appeared on the display to instruct the monkey to perform supination or pronation of either forelimb. After another 500-ms waiting period, a GO signal was displayed, prompting the monkey to perform the movement as instructed. A single-site electrode, with an impedance of 0.5–2 M Ω, was inserted into the left or right side of the pre-SMA or SMA while the monkeys performed the task [3]. The epochs of the LFP were filtered (5–300 Hz) and digitally sampled (at 1 kHz), starting from the fixation point and lasting until 500 ms after the GO signal.

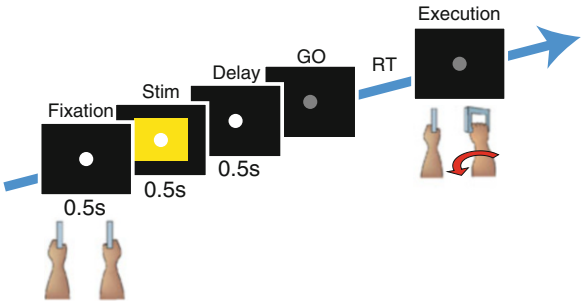


Fig. 1 The sequence of events in the delayed response task

3 Results

The monkeys performed the appropriate movements in response to the visual instructions, and LFPs were recorded in the pre-SMA and SMA during the behavioral task. A total of 57 sessions were recorded in the pre-SMA and 51 in the SMA for monkey M, and 345 were analyzed in the pre-SMA and 191 in the SMA for monkey N. Figure 2a shows an example of LFP activity recorded in the pre-SMA. The LFP fluctuated significantly, and contained several frequency oscillatory components. The beta power (10–40 Hz) of the LFP increased during the initial 500 ms and delay period, but was suppressed during presentation of the stimulus and after the GO signal. Although the power of the gamma oscillation (40–200 Hz) was increased throughout the trial, it was more pronounced during presentation of the stimulus.

Figure 2b, c show the time-frequency plots of an average of the LFP power of monkey N. The LFP power in each trial was normalized by the power during the initial 500 ms (0–500 ms from fixation). In the pre-SMA (Fig. 2b), gamma power increased phasically during the stimulus presentation and after the GO signal, as indicated by the red triangle and red arrow, respectively. During the delay period (the period between the stimulus presentation and the GO signal), gamma power remained at baseline. In addition, beta power decreased concurrently with the increase in gamma power, indicated by the black triangle and arrow. In the SMA (Fig. 2c), the increase in gamma power was not observed during stimulus presentation, and only occurred after the GO signal, shown by the red arrow. In contrast, the beta power decreased only after the GO signal, indicated by the black arrow.

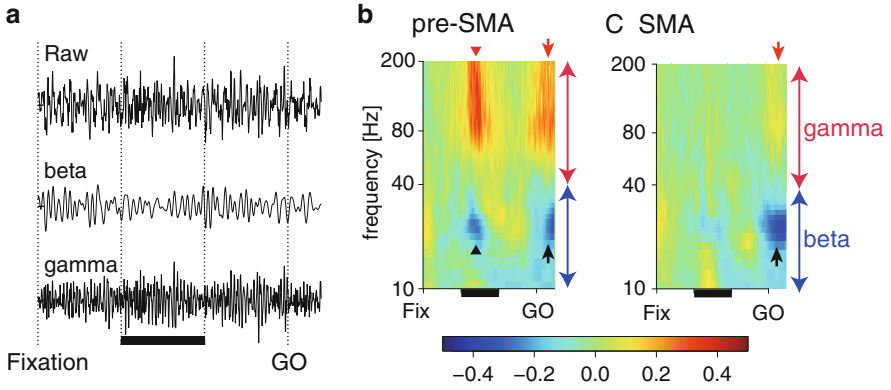


Fig. 2 Representative LFP activity. (**a**) Raw LFP data recorded from the pre-SMA of monkey M (*top*), the 10–40 Hz (beta) component of LFP (*middle*), and the 40–200 Hz (gamma) component of LFP (*bottom*). The *horizontal thick bar* indicates the period of visual stimulus presentation. (**b–c**) A time frequency plot of the average of the power of LFP in the pre-SMA (**b**) and SMA (**c**) of monkey N. The LFP activity was normalized by power during the 0–500 ms after the fixation. The colors indicate the z-score of the power: *red triangles* and *arrows* indicate the activation of gamma power, whereas *black triangles* and *arrows* indicate the suppression of beta power. An example of the LFP activity

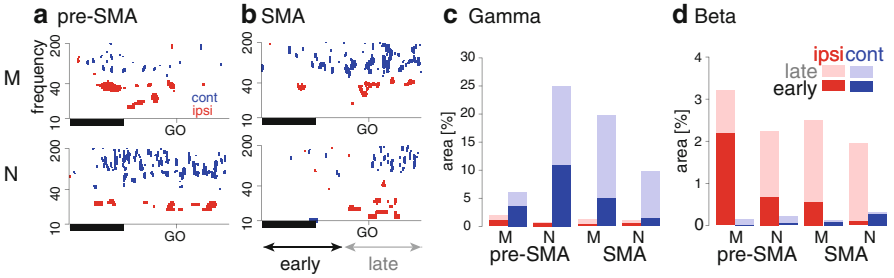


Fig. 3 (**a**–**b**) Time-frequency plots of power differences compared with arm use in pre-SMA and SMA in two monkeys. Color codes indicate the regions where the gamma oscillations were significantly stronger ($p < 0.01$) for ipsilateral (*red*) and contralateral arm use (*blue*). *Thick horizontal lines* indicate the period of the stimulus presentation. (**c**–**d**) The area of the time-frequency region showing lateralization for ipsilateral (*red*) and contralateral arm use (*blue*). The areas for the early and late periods are depicted by the *light* and *dark colors*, respectively. The area was represented by a ratio of the entire region (from stimulus onset to 500 ms after the GO signal). *M* and *N* represent monkey *M* and *N*, respectively. (**c**) The laterality of the gamma oscillation. (**d**) The laterality of the beta oscillation

To examine if the power of the LFP oscillations changed according to the arm used, we used Welch t-test to compare the LFP power with ipsilateral and contralateral arm use. The time-frequency plots of power differences with arm use are shown in Fig. 3a, b. Thick horizontal lines indicate the period of stimulus presentation, and the color codes indicate the regions where oscillations were significantly stronger for either ipsilateral (red) and contralateral arm use (blue). Beta and gamma oscillations showed different lateralization according to the arm use. Specifically, gamma power was significantly stronger when monkeys were about to perform a movement with the contralateral arm compared with the ipsilateral arm ($p < 0.01$), whereas beta power was significantly stronger with ipsilateral arm use ($p < 0.01$). The significant regions of the pre-SMA were broadly distributed, whereas the distribution in the SMA was biased to the period immediately following the GO signal. Interestingly, the distribution bias of beta power resembled that of gamma power.

The areas of significance are summarized in Fig. 3c, d. To evaluate the effects of stimulus presentation and movement execution on lateralization, the trial was separated into two periods: early (from stimulus onset to 250 ms after the termination of the stimulus), and late (from 250 ms after termination of the stimulus until 500 ms after the GO signal). In the gamma frequency range, the total area (a sum of the early and late periods) was larger for use of the contralateral arm compared with the ipsilateral. In contrast, the total area for the ipsilateral arm use was larger than the contralateral in the beta frequency range. In the SMA, the area of the late period is larger than that of the early period in both the beta and gamma frequency bands, which was not observed in the pre-SMA.

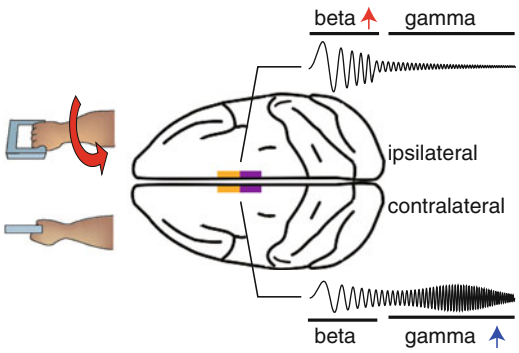


Fig. 4 Reciprocal role of gamma and beta oscillation in arm use selection

4 Discussions

We carried out LFP recordings from primate bilateral pre-SMA and SMAs, and found that beta and gamma frequency bands change the lateralization of LFP power in the medial motor cortex. The power of the gamma oscillation increased with contralateral arm use, whereas beta power was augmented during ipsilateral arm use (Fig. 4). Gamma oscillations are closely related to neuronal spikes, and as such they frequently exhibit simultaneous activation [4]. Gamma oscillation may therefore co-operate with neuronal activities to select an appropriate effector for movement. In contrast, beta power increased during ipsilateral movement [5]. The maintenance of beta oscillation during ipsilateral movement implies that beta oscillation is involved in withholding the ipsilateral arm movement.

Acknowledgements This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas "Neural Creativity for Communication (No. 4103)" (24120702) from MEXT.

References

1. Tanji, J., Okano, K. and Sato, K.C. Neuronal activity in cortical motor areas related to ipsilateral, contralateral, and bilateral digit movements of the monkey. J. Neurophysiol. 60, 325–343, 1988.
2. Buzsaki, G., Anastassiou, C. A. and Koch, C. The origin of extracellular fields and currents - EEG, ECoG, LFP and spikes. Nature reviews. Neuroscience 13, 407–420 2012.
3. Matsuzaka, Y., Aizawa, H. and Tanji, J. A motor area rostral to the supplementary motor area (presupplementary motor area) in the monkey: neuronal activity during a learned motor task. J. Neurophysiol. 68, 653–662, 1992.
4. Fries, P., Nikolic, D. and Singer, W. The gamma cycle. Trends in Neurosciences 30, 309–316, 2007.
5. Murthy, V. N. and Fetz, E. E. Coherent 25- to 35-Hz oscillations in the sensorimotor cortex of awake behaving monkeys. Proceedings of the National Academy of Sciences of the United States of America 89, 5670–5674, 1992.

Thalamocortical and Intracortical Contributions to Task-Related Cross-Frequency Coupling in Auditory Cortex

Marcus Jeschke and Frank W. Ohl

Abstract Oscillations are a hallmark of the neuronal dynamics of neocortex. Whereas earlier work investigated roles of individual frequency bands in cortical computation, newer studies demonstrated that different frequencies can be coupled and changed according to behavioral task demands. Thus, it was speculated that cross-frequency coupling could serve the dynamic coordination of local neural circuits and integrate or select larger communicating neural networks. However, it is unclear which circuit elements contribute to cross-frequency coupling. Here, we investigated the roles of thalamocortical and intracortical contributions using a method we have previously developed. Our results indicate that different cortical elements exhibit discernible cross-frequency coupling. The data further show that microsurgical interference with intracortical communication leads to specific modulation patterns of cross frequency coupling. Finally, we present a simple model to account for the observed results.

Keywords Theta-gamma coupling • Current source density analysis • Thalamocortical • Intracortical • Cross-frequency coupling • Auditory cortex

1 Introduction

Learning can elicit profound changes of the activity of cortical neuronal networks (e.g. [4]). Previously, we presented a discrimination learning and category-formation paradigm that allows dissociating mesoscopic auditory cortical activity

M. Jeschke
Department of Systems Physiology of Learning (SPL), Leibniz-Institute for Neurobiology (LIN), Magdeburg, Germany

Department of Otolaryngology, University of Göttingen, Göttingen, Germany

F.W. Ohl (✉)
Department of Systems Physiology of Learning (SPL), Leibniz-Institute for Neurobiology (LIN), Magdeburg, Germany

Institute of Biology, University of Magdeburg, Magdeburg, Germany

Center for Behavioral Brain Sciences (CBBS), Magdeburg, Germany
e-mail: frank.ohl@lin-magdeburg.de

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_24

in the gamma-band related to stimulus feature processing from activity related to semantic processing as separable elementary processes of cognition [7]. Traditionally, research has focused on individual frequency bands in neural processing (cf. [1]) while newer work on cortex demonstrated coupling of bands during various processing situations like sensory processing, attention and learning (e.g. [9, 13, 14, 16]). Therefore it was suggested that this cross-frequency coupling might serve dividing neuronal computation into frames by coordinating local neural circuits with larger communicating neural networks [3, 6]. However, the role of local cortical circuits, organized into columns and interconnected via horizontal projections of various spatial extents, has been elusive. We propose that local and global cortical circuit elements have distinct roles in organizing oscillatory cortical activity. To dissociate thalamocortical, local contributions and intracortical, global contributions to cortical activity patterns we have developed a method based on the analysis of current source density patterns [5]. This method was used here to investigate cross-frequency coupling after sensory stimulation based on the activity of presumed local and global circuit elements. We further determined discernible roles of thalamocortical and intracortical processes underlying task-relevant modulation of gamma-oscillations by theta phase using microsurgical cuts of intracortical connections.

2 Methods

Depth profiles of local field potentials (LFPs) were recorded from the auditory cortex of ten ketamine-xylazine anesthetized, adult Mongolian gerbils (*Meriones unguiculatus*). To disrupt the activity of intracortical horizontal connections microsurgical cuts were performed. Procedures were approved by an ethics committee of the state Saxony-Anhalt and conformed to the NIH *Guidelines for Animals in Research*. Stimuli were digitally synthesized and delivered via an electrostatic headphone mounted 5 cm in front of the animal. Responses were driven by 200 ms long pure tones at 64 dB SPL from 0.25 to 16 kHz in half or one octave steps, repeated 50 or 100 times every 0.6–0.8 s. Current source density (CSD)[10] based measures were calculated from LFP profiles. Average rectified CSDs (AvgRecCSD) were calculated to determine the overall activation strength. A residual analysis (RelResCSD, [5]) determined the strength of recruitment of intracortical transcolumnar circuits for cortical processing. The best frequency (BF) was defined as the pure-tone frequency that elicited the strongest AvgRecCSD. Cross-frequency coupling comodulograms (Fig. 1) of cortical activity were computed using the cross-frequency coherence (CFC) between the phase at one frequency (f_{phase}) and the amplitude at a second frequency (f_{amp}) [12]. Statistical analyses of the CFC were based on averages across all frequencies belonging to the frequency bands for f_{phase} and f_{amp} in which the maxima of the comodulograms were found (Fig. 2). When referring to classic frequency bands we used the following intervals: theta (3–8 Hz), alpha (8–13 Hz), beta (13–23 Hz), low gamma (30–50 Hz), high gamma (60–100 Hz), respectively.

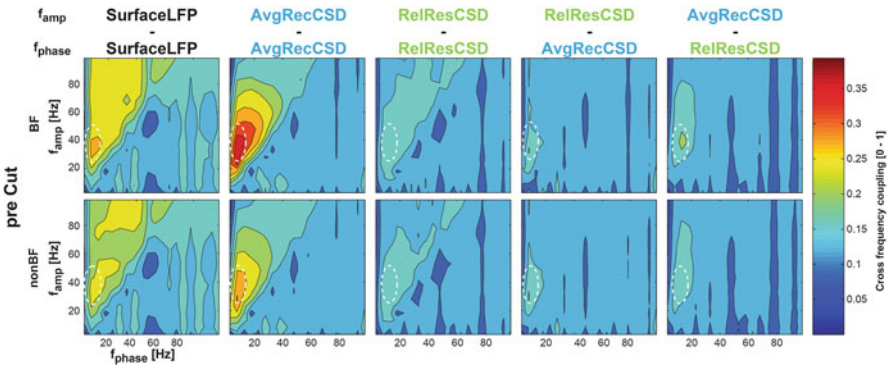


Fig. 1 Cross frequency coupling during stimulation with the best frequency (BF) and non-optimal frequencies (nonBF). In each panel the *dashed*, *white oval* indicates the CFC frequency ranges used for statistical evaluation (Fig. 2). See main text for further description

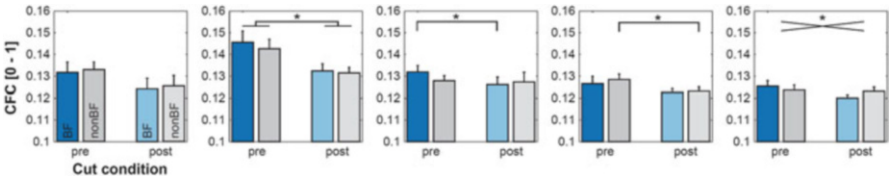


Fig. 2 Statistical evaluation of the influence of dissecting intracortical connections on CFC. The panel order follows Fig. 1. See main text for explanation

3 Results

3.1 Cross-Frequency Coupling Between Measures of Cortical Activity

Stimulation frequencies were split into a BF and a nonBF group as presenting non-optimal frequencies (nonBF) activates horizontal inputs in addition to thalamocortical inputs [5, 8]. For surface LFPs (Fig. 1 left panels) the strongest CFC was observed after 13 Hz f_{phase} and 38 Hz f_{amp} (BF stimulation) and at 8 Hz f_{phase} and 33 Hz f_{Amp} (nonBF). The modulation of low and high gamma amplitude by the theta phase appeared very similar to previous reports [3, 13, 16]. Peaks of the CFC comodulograms based on intracolumnar activity (AvgRecCSD) were found at 8 Hz f_{phase} and 33 Hz f_{amp} (BF) and at 8 Hz f_{phase} and 28 Hz f_{amp} (nonBF) (Fig. 1, column 2). Comodulograms calculated from transcolumnar activity (RelResCSD) peaked at higher frequencies (13 Hz f_{phase} and 38 Hz f_{amp} BF and nonBF]; Fig. 1, column 3). Further, the CFC was stronger for BF than for nonBF stimulation (paired *t*-test: $p < 0.032$; Fig. 1, column 3). To test how columnar activity influenced transcolumnar processing the f_{phase} was calculated from the

AvgRecCSD and f_{amp} from the RelResCSD (Fig. 1, column 4). Peaks were found at 8 Hz f_{phase} and 28 Hz f_{amp} for both stimulus groups. In contrast, the strongest CFC was found at 13 Hz f_{phase} and 38 Hz f_{amp} for BF and nonBF stimulation if f_{phase} was calculated based on RelResCSD and f_{amp} based on AvgRecCSD (Fig. 1, column 5). In other words, theta-to-high-beta/gamma coupling was observed for columnar to transcolumnar interactions and beta-to-gamma coupling was detected for transcolumnar to columnar interactions.

3.2 Contribution of Intracortical, Horizontal Connections to Cross Frequency Coupling

To investigate the role of long range processes for CFC, we performed surgical dissections of intracortical, horizontal connections. For the surface LFP theta-to-gamma coupling no significant effects were observed (Fig. 2, panel 1). In contrast, there was a significant reduction of theta-to-low gamma coupling based on the AvgRecCSD (Fig. 2, panel 2, paired *t*-test: $p = 0.041$ and $p = 0.043$ for BF and nonBF). A reduction of RelResCSD based beta-to-gamma coupling was observed after BF stimulation (Fig. 2, panel 3: paired *t*-test: $p = 0.018$). Next, we addressed whether cortical cuts modified the coupling of the phase of local activity and the amplitude of global activity (Fig. 2, panel 4). A significant decrease of theta-to-gamma coupling was observed after nonBF stimulation (paired *t*-test: $p = 0.045$). The beta-to-gamma coupling from global to local activity influences was also changed with a significant interaction between stimulation frequency and cut condition (Fig. 2, panel 5; RM-ANOVA: $p = 0.012$).

4 Discussion + Conclusions

Our data indicate that local, columnar cortical activity and global, transcolumnar activity display distinct CFC patterns. While theta-to-gamma coupling was found for local activity an alpha-to-gamma coupling was observed for global activity no matter whether global activity was coordinated in general (RelResCSD to RelResCSD CFC) or global activity tied in local activity (RelResCSD to AvgRecCSD CFC). In line with this idea, a study in human epilepsy patients reported a shift from theta-to-gamma to alpha-to-gamma coupling if larger areas of cortex needed to be coordinated for visual task performance [16].

The role of columnar and transcolumnar processes for cross-frequency coupling in cortex can be summarized in a simple model (Fig. 3). Larger neural networks have lower oscillation frequencies than smaller networks [2, 15], thus, local processes, e.g. within a column or between neighboring columns, lead to higher frequency oscillations than global processes, between distant columns. Accordingly, for local

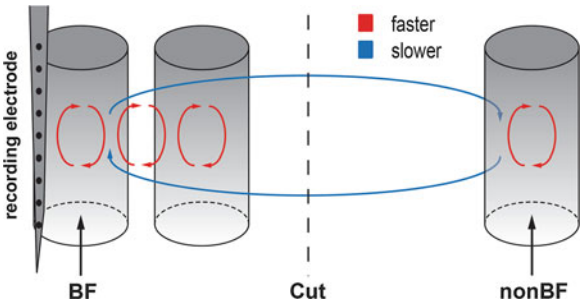


Fig. 3 Scheme illustrating the interaction of local and transcolumnar cortical processing. An electrode records activity at a given column with a certain BF. Horizontal connections provide *short* and *long-range* projections from other columns, whose contribution can be quantified by residual CSD analysis

activity (AvgRecCSD) presentation of nonBFs leads to a peak in CFC at a lower frequency than stimulation with the BF, as here columnar and transcolumnar processes contribute. For global activity (RelResCSD) no dependency on stimulation frequency was found as long range connections are recruited for all frequencies. Local-to-global interactions were observed at a lower frequency combination than for global-to-local interactions, which might be expected if local activity needs to act on slower, global activity than if global activity acts on fast, local activity. Cortical cuts of long-range connections led to reduced CFC of local activity for all stimuli. This suggests that normal local interactions need global support. CFC of columnar to transcolumnar processes was reduced for nonBFs as long-range horizontal connections which carry information about nonBFs were disrupted. Cortical cuts were more disruptive on BF based RelResCSD CFC than on nonBFs, indicating a directionality of influences such that activity originating from the observed column has a stronger impact on interactions based on horizontal connections than activity originating elsewhere. This might explain the interaction of frequency and cut condition for transcolumnar to columnar influences as BFs depend more on horizontal connections.

Although more work is needed to disentangle the complex interplay of local and global cortical circuitry, these results are fundamental for establishing a mapping of elementary processes of cognition, like stimulus processing and meaning generation, to elements of cortical neuronal circuits.

References

1. Basar E, Basar-Eroglu C, Karakas S, Schurmann M (2001) Gamma, alpha, delta, and theta oscillations govern cognitive processes. Int J Psychophysiol 39:241–248.
2. Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. Science 304:1926–1929.

3. Canolty RT, Knight RT (2010) The functional role of cross-frequency coupling. Trends Cogn Sci 14:506–515.
4. Freeman WJ, Quiroga QR (2012) Imaging Brain Function with EEG. New York: Springer.
5. Happel MFK, Jeschke M, Ohl FW (2010). Spectral integration in primary auditory cortex attributable to temporally precise convergence of thalamocortical and intracortical input. J Neurosci 30:11114–11127.
6. Jensen O, Colgin LL (2007) Cross-frequency coupling between neuronal oscillations. Trends Cogn Sci 11:267–269.
7. Jeschke M, Lenz D, Budinger E, Herrmann CS, Ohl FW (2008) Gamma oscillations in gerbil auditory cortex during a target-discrimination task reflect matches with short-term memory. Brain Res 1220:70—80.
8. Kaur S, Lazar R, Metherate R (2004) Intracortical pathways determine breadth of subthreshold frequency receptive fields in primary auditory cortex. J Neurophysiol 91:2551—2567.
9. Lakatos P, Shah AS, Huth KH, Ulbert I, Karmos G, Schroeder CE (2005) An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J Neurophysiol 94:1904–1911.
10. Mitzdorf, U (1985) Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. Physiol. Rev., 65:37—100.
11. Ohl FW, Scheich H, Freeman WJ (2001) Change in pattern of ongoing cortical activity with auditory category learning. Nature 412:733—736.
12. Onslow AC, Bogacz R, Jones MW (2011) Quantifying phase-amplitude coupling in neuronal network oscillations. Prog Biophys Mol Biol. 105: 49—57.
13. Tort ABL, Kramer MA, Thorn C, Gibson DJ, Kubota Y, Graybiel AM, Kopell NJ (2008) Dynamic cross-frequency couplings of local field potential oscillations in rat striatum and hippocampus during performance of a T-maze task. Proc Natl Acad Sci USA 105:20517—22.
14. Tort ABL, Komorowski RW, Manns JR, Kopell NJ, Eichenbaum H (2009) Theta–gamma coupling increases during the learning of item-context associations Proc Natl Acad Sci U S A 106:20942–20947.
15. Traub RD, Whittington MA, Stanford IM, Jefferys JGR (1996) A mechanism for generation of long-range synchronous fast oscillations in the cortex. Nature 383:621–624.
16. Voytek B, Canolty RT, Shestyuk A, Crone NE, Parvizi J, Knight RT (2010) Shifts in gamma phase-amplitude coupling frequency from theta to alpha over posterior cortex during visual tasks. Front Hum Neurosci 4:191.

Temporal Characteristics of the Steady State Visual Evoked Potentials

Maciej Labecki, Magdalena Zieleniewska, Karol Augustin, Jaroslaw Zygierewicz, and Piotr Suffczynski

Abstract Steady State Visual Evoked Potentials (SSVEP), observed in EEG over occipital areas of the scalp, are natural responses to flicker stimuli. This phenomenon is used in Brain-Computer Interfaces (BCI) and in psychology as an indicator of attention. Despite the fact that SSVEPs are used in these two fields of neuroscience, the knowledge of their fundamental properties is still limited.

In this study we investigated time evolution during long term flicker stimulus. The analysis was done using EEG signals recorded during a series of 60-s long stimulation epochs interleaved with 30-s rest epochs. For nine out of ten subjects we observed unequivocal decrease of SSVEP power during the first seconds of stimulation. The habituation index was $0.49\ \% \pm 0.09\ \%$ on average across subjects and it reached maximal value of $71\ \% \pm 0.10\ \%$ in a single subject.

These results may have direct implications on the research in the rapidly growing field of biomedical engineering related to brain-computer interfaces. The BCI users are often submitted to prolonged flickering light stimulation. Detailed knowledge regarding the stimulus efficiency over the course of the stimulation is therefore essential. These properties may also have an impact on the research field of information processing in the neural systems and on the scientific applications of these signals e.g. in psychological research.

Keywords SSVEP • BCI • Habituation • EEG

1 Introduction

Steady State Visual Evoked Potentials, which were discovered by [1] are the steady-state responses elicited by flicker stimulation. Frequency of oscillations of these neural responses corresponds to the stimulus frequency [2].

SSVEPs are used in two fields of neuroscience – in psychological studies as an indicator of attention [3] and in brain-computer researches [4, 5]. Despite the

M. Labecki (✉) • M. Zieleniewska • K. Augustin • J. Zygierewicz • P. Suffczynski
Faculty of Physics, University of Warsaw, Warsaw, Poland
e-mail: Maciej.Labecki@fuw.edu.pl

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_25

SSVEPs are commonly used in these research fields, the knowledge of some of their fundamental properties is still limited. Basic research mainly concerns the analysis of the relation between stimuli parameters and the magnitude of the response. There is a lack of research concerning the time evolution of these potentials and almost all of their applications are based on assumption of their stationarity. The aim of this study was to test this assumption.

2 Methods

2.1 Subjects

The study was done on group of ten healthy volunteers – five males and five females. Their ages ranged from 19 to 28 years.

2.2 Data Acquisition and Pre-processing

The EEG signal was collected using the TMSi Porti 7 amplifier and 10–20 EEG cap. The sapling frequency was 512 Hz. Collected EEG signal was filtered with 2 Hz high-pass filter (DC) and notch filter (50 Hz).

2.3 Stimuli and Stimulation Procedure

We used SSVEP stimulator which had been constructed in the Department of Biomedical Physics [6]. Subjects were concentrating on the white flickering square which was presented on the screen. We used series of 60-s long stimulation epochs interleaved with 30-s rest epochs. Amount of 50 trials for each subject was recorded. To reduce fatigue of the subjects, measurement sessions were divided to five sub-sessions separated by 15 min brakes. The stimulation frequency was 15 Hz. A signal from photodiode measuring the flickering light was recorded on a separate channel simultaneously with EEG.

2.4 Data Analysis

The SSVEP strength was estimated by the instantaneous power of the EEG signal that was bandpass filtered in a band around the stimulation frequency (14.5–15.5 Hz) using first order Butterworth filter. Then, the square of signal was computed and an envelope was estimated.

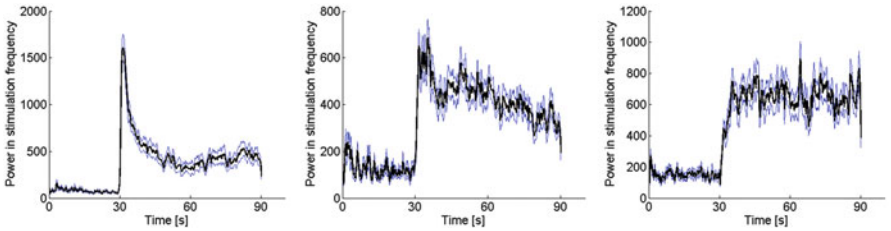


Fig. 1 Instantaneous power of SSVEP for individual subjects exhibiting: fast habituation (*left*), slow habituation (*middle*) and no habituation (*right*). The *black line* is the average across 50 trials, and the *blue lines* mark the standard error. The vertical scale of each subplot was adjusted to optimally present the power time course

3 Results

For nine out of ten subjects significant habituation of the response was observed. Different types of the response time course may be distinguished across subjects. In some subjects habituation follows a slow time course and is present throughout whole stimulation period (Fig. 1, middle panel). In other subjects the habituation is faster and is observed only during the first part of stimulation period, and afterwards the SSVEP power remains stable (Fig. 1, left panel). Still in one subject no significant habituation is observed during the 60 s of stimulation (Fig. 1, right panel).

Habituation was quantified as $100\ \% \cdot (\text{E1-E2})/\text{E1}$, where E1 is SSVEP energy between seconds 2 and 3 after the stimulation onset, E2 is SSVEP energy between seconds 3 and 2 before the stimulation offset. The habituation was $0.49\ \% + -0.09\ \%$ on average across subjects and it reached maximal value of $71\ \% + -0.10\ \%$ in a single subject.

4 Discussion

In this study we showed that for most of the subjects the strength of SSVEP decreases during long term stimulation. The habituation mechanism is not well understood. It is generally considered that attention modulates the SSVEP response [3]. Therefore, one possible explanation would be that a subject cannot maintain the same level of attention during whole stimulation period. Alternatively, the habituation may be caused by neuronal, plastic changes at the level of synapses. More experimental work is needed to provide physiological explanation for habituation phenomenon.

In most of the BCI systems the stimulation light flickers continuously during the entire period when the device is used by the subject. Habituation processes may decrease the SSVEP response affecting their proper classification. Possible solution

to the problem would be to divide the stimulation with short rest periods allowing SSVEP to recover. Influence of length of rest period for restoring SSVEP strength should be examined in further studies.

5 Conclusion

In most cases SSVEP habituates significantly. That contradicts common assumption that they correspond to stationary signals. Investigation of this phenomenon is necessary for constructing more efficient BCI devices.

References

1. Regan D, 1966. Some characteristics of average steady-state and transient responses evoked by modulated light Electroencephalogr Clin Neurophysiol. 20(3):238–48
2. Herrmann CS, 2001. Human EEG responses to 1–100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. Experimental Brain Research 137(3–4): 346–353.
3. Muller MM, Andersen S, Trujillo NJ, Valdes-Sosa P, Malinowski P, Hillyard SA, 2006. Feature-selective attention enhances color signals in early visual areas of the human brain. Proc. Natl. Acad. Sci. U.S.A. 103 (38), 14250–14254.
4. Vialatte F, Maurice M, Dauwels J, Cichocki A, 2010. Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives. Prog. Neurobiol. 90, 418–438
5. Zhu, D, Bieger J, Molina GG, Aart RM, 2010. A Survey of Stimulation Methods Used in SSVEP-Based BCIs. Computational Intelligence and Neuroscience 1, 702357
6. Durka PJ, Kus R, Zygierewicz J, Michalska M, Milanowski P, Labecki M, Spustek, T, Laszuk D, Duszyk A, Kruszynski M, 2012. User-centered design of brain-computer interfaces: OpenBCI.pl and BCI Appliances. Bull. Pol. Ac.: Tech. 60(3), 427–433

A Novel Neural Coding Mechanism Study of Interaural Time Difference Detection

Hong Zhang, Jiong Ding, and Qinye Tong

Abstract Interaural time difference (ITD) detection plays a critical role in sound localization. Some neurophysiology experiments found that the ITD discrimination threshold for single neuron in medial superior olive (MSO) is about 20 μs. In order to reveal the ITD coding mechanism essentially, an ITD detection model has been proposed in this paper based on the anatomical structure of the mammalian auditory system. Most previous research only discuss the relationship between the ITDs and the spike rate of the MSO neuron. In these results, the spike rate difference is under 1 spike/s when the ITD changes in 20 μs. It means the neuron needs more than 1 s to distinguish the ITD by spike rate coding. This explanation contradicts with the rapid and accurate sound localization response in animal psychoacoustics experiments. In our study, the neural coding in ITD detection model has been investigated via the approach of circle maps and symbolic dynamics. With these two methods, the ITDs under 20 μs can be detected in 1 s. This neural coding hypothesis agrees with the ITD detection response time in the animal behavior experiments.

Keywords Circle maps • Symbolic dynamics • Neural coding • Sound localization

1 Introduction

Sound localization plays a critical role in animal survival and predation. Based on the results from neurophysiology and psychoacoustics studies in the last century, it proved that many mammals make use of interaural time difference to perform sound localization. In the recent anatomy experiments, the neurons which are sensitive to ITDs were found in the medial superior olive. Animal behavior studies reveal that the resolution of ITD is about 20 μs [1].

The neural mechanisms of the mammalian auditory nervous system underlying high-resolution ITD detection are still not fully understood. Currently, methods in

H. Zhang (✉) • J. Ding • Q. Tong
Department of Biomedical Engineering, Zhejiang University, Hangzhou, People's Republic of China
e-mail: zhangh@mail.bme.zju.edu.cn

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_26

studying the mechanism of ITD are concentrated on animal behavior experiments. These results all show that the discharge rate of the MSO neuron is closely related to the ITDs [2]. But the rate code cannot discriminate the tiny ITDs in a short time. It conflicts with the obvious ability of sound localization in animal. For this study, two nonlinear methods, circle maps and symbolic dynamics were used to analyze the resolutions of ITDs from the neuron spike trains quantitatively.

2 Methods/Models

In our computer simulation, an ITD detection model is proposed based on the anatomical structure and physiological data from an adult gerbil auditory system. It includes a cochlea model and a neural network in the cochlear nucleus. The cochlea model is composed of an inner hair cell model and an auditory model. The ITD detection neural network consists of two globular bushy cells and an MSO neuron. The auditory nerve fiber connects the cochlea and cochlear nucleus. Finally, the ITD signal was encoded on the MSO neuron. The ITD detection circuit is shown in Fig. 1. The time difference of the sound arrival in each ear is $\triangle$T. The inner hair cell (IHC) model, auditory nerve (AN) model, neuron cell models are described in the following sections.

The inner hair cell (IHC) model corresponds to the transfer function $Vm = f_1(x)$ in Fig. 1. The basilar membrane (BM) displacement, a sine waveform, is the input signal (x), the output signal is the membrane potential (Vm) of the inner hair cell. The model is proposed by utilizing a realistic transducer conductance and membrane time constant. It also assumes that the tension-gated channels are the only apical channels [3]. The auditory nerve (AN) model corresponds to the transfer function, $FR = f_2(Vm)$ in Fig. 1. This model is based on Mountain's work in hearing model research. The final output of this model is the action potential firing rate on the auditory nerve [4]. All the neurons in the ITD detection neural network are presented by the model of Rothman (a realistic biophysical neuron model based on ionic channels) [5]. The main function of the neuron model is shown in Eq. 1.

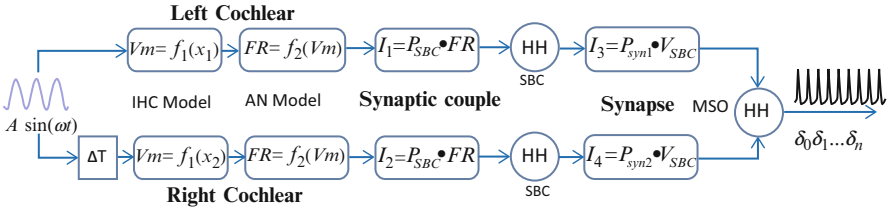


Fig. 1 The computational model of the ITD detection neural circuit. The input signal of each ear has the same frequency, and $\triangle$T is the ITD signal

The Rothman model is solved by the fourth-order Runge-Kutta method with the step is 0.02 ms.

$$C_m \frac{dV}{dt} = I_A + I_{LT} + I_{HT} + I_{Na} + I_h + I_{lk} + I_E - I_{ext} \tag{1}$$

C_m is the membrane capacitance of the neuron; V represents the membrane potential of the neuron; I_A, I_{LT}, I_{HT}, I_{Na} and I_E are ion current; I_{lk} is the leak current; I_{ext} is the external input current.

2.1 Circle Maps in the ITD Detection Model

The displacement in each cochlea BM has the same frequency. The right ear BM displacement signal has a ΔT delay compared to the left one. The output of the ITD model is the spike train of the MSO model. The ITD detection model can be considered as a dissipative dynamical system with two competing frequencies: the input BM displacement signal frequency and the oscillation frequency of MSO neuron. But the ITD detection model is a high dimensional nonlinear model, our quantitative knowledge about highly nonlinear system is meager. In order to analyze the neural coding of the ITD model, first, the complexity of this model should be reduced. Analyzing the interaction between these two frequencies is a effective method to reveal the mechanism of ITD neural coding. Circle map is a valid method to simplify a nonlinear dynamic system [6, 7].

Through the following operations, we can get the circle map of the ITD detection model.

In the ITD detection model, the variable θ_n represents the phase of the MSO spike train measured stroboscopically at periodic time point, $t_n = 2\pi n/\omega$ from the periodic input BM displacement ($x_1 = Asin(\omega t)$ in Fig. 1). A phase shift $\theta_n \to \theta_n + 1$ represents a full rotation. Due to the complexity of the ITD model, the system function $\theta_{n+1} = f(\theta_n)$ is difficult to express. But with the evolution of iterations of the circle maps, the phase train $\theta, f(\theta), f^2(\theta), \ldots$, are computed. The detailed steps are as follow. The time point train $t_1, t_2, \ldots, t_n$ is the periodic time point from the input signal $x_1 = A\sin(\omega t)$. The firing time of the spike train is recorded and denoted as $\delta_1, \delta_2, \ldots \delta_n$. δ_i is the time of ith spike of the MSO neuron. The spike phase, θ_i, is as follows:

$$\theta_i = \left(\delta_i - t_j\right) \times \omega / 2\pi \tag{2}$$

Reconstruction the phase train with the iterating operation, an one-dimensional discrete map was obtained from the ITD model, as shown in Fig. 2.

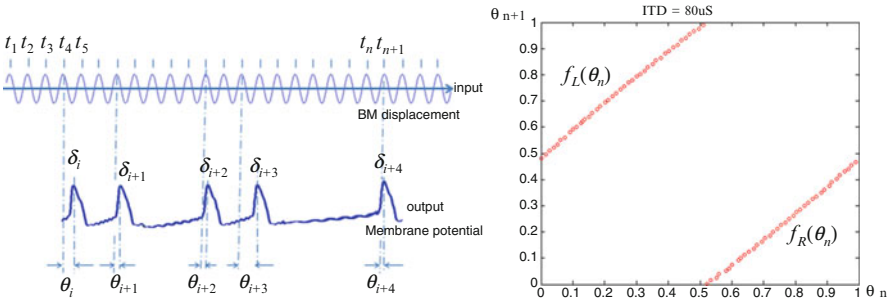


Fig. 2 *Circle* maps of the ITD detection computational model

2.2 Symbolic Dynamics of the ITD Detection Model

The basic idea of the symbolic dynamic is to compute the itineraries of orbits in terms of the regions of the phase space. According to the principle of symbolic dynamics [8, 9], the spike train symbolic rule is followed:

If the spike train is $\eta = \delta_1\delta_2\delta_3 \ldots \delta_i$, then the symbolic equation is:

$$S_i = [(\delta_i - \delta_{i-1}) \times \omega/2\pi] \tag{3}$$

The circle map of the ITD is a monotonically increasing map, therefore, the ordered rule of the symbol train is the same as S_i. If there are two sequence trains:

$$A = \{a_i\} = a_1a_2a_3 \ldots a_n, \tag{4}$$

$$B = \{b_i\} = b_1b_2b_3 \ldots b_n, \tag{5}$$

then a_i and b_i represent any one of respectively. In order to compare these two symbol trains, then the first symbols, a_1 and b_1, of the symbol sequences should be compared. When there is no less than one symbol identical to each other at the beginning of the train. The relationship is:

$$\Sigma p_0 < \Sigma p_1 < \Sigma p_2 < \cdots < \Sigma p_r \tag{6}$$

where, Σ represents the first of several successive same symbols.

3 Results

Supposed the frequency of the BM displacement is 500 Hz. The amplitude of the BM displacement is 15 nm. Setting the ΔT in Fig. 2 from 0 to 100 μs, and solving equations of the ITD model by Matlab, the symbolic sequences from the MSO

Table 1 ITD detection symbol sequence

ITD (μS)	Symbol sequence
0	$9\,8\,9\,8\,8\,9\,(8\,9)^{10}\,8\,8\,9\,(8\,9)^{12}\,8\,8\,9\,(8\,9)^{11}$
20	$9\,8\,9\,8\,8\,9\,(8\,9)^{9}\,8\,\underline{\mathbf{8}}\,9\,(8\,9)^{12}\,8\,8\,9\,(8\,9)^{12}$
40	$9\,8\,\underline{\mathbf{8}}\,9\,(8\,9)^{10}\,8\,8\,9\,(8\,9)^{11}\,8\,8\,9\,(8\,9)^{12}$
60	$9\,8\,8\,9\,(8\,9)^{9}\,8\,\underline{\mathbf{8}}\,9\,(8\,9)^{12}\,8\,8\,9\,(8\,9)^{12}$
80	$9\,8\,8\,9\,(8\,9)^{9}\,8\,8\,9\,(8\,9)^{11}\,8\,\underline{\mathbf{8}}\,9\,(8\,9)^{12}$
100	$\underline{\mathbf{8}}\,9\,(8\,9)^{9}\,8\,8\,9\,(8\,9)^{11}\,8\,8\,9\,(8\,9)^{12}$

Note: $(ab)^n = \underbrace{abab\ldots ab}_{n}$

neural spikes are listed in Table 1. The symbols which are marked as both bold and underlined letters in all the whole tables are the first different symbols between the neighboring symbolic sequences.

The ordered rule is that the symbol sequence is decreasing when the ITD value is increasing. When the length of the symbol trains reach to 49 numbers, the ITD detection resolution can be achieved at 20 μs. This means after 49 neural action potentials, the ITD signals can be discriminated by the auditory neural circuits. In the simulation, the average firing rate of the MSO neuron is 70 spikes/s, so the neural circuits are able to measure the ITD signals accurately less than 1 s.

4 Conclusions

From the simulation results, using symbolic dynamics can discriminate ITD under 20 μs in 1 s. These results are close to the limits of mammalian sound localization. Circle map is a method for reducing the dimensionality of the complex nonlinear system to a one-dimensional discrete map. The ITD detection computational models are reduced to the one-dimensional discrete map by the method of circle map. On the one-dimensional discrete map, the neural coding on the MSO neuron is decoded quantitatively by the method of symbolic dynamics. The ITD is expressed by a series of symbol sequences with the aforementioned methods. The symbol sequences make the ITD signals measurement in high-precision possible. The final result reveals that using the methods of circle maps and symbolic dynamics is a valid way to study the mechanism of sound localization.

References

1. A. N. Popper and R. R. Fay, Sound Source Localization. (New York: Springer, 2005), pp. 1–30.
2. B. Grothe, M. Pecka and D. McAlpine, “Mechanisms of sound localization in mammals,” Physiol Rev 90, 983–1012 (2010)

3. A. R. Cody and D. C. Mountain, “Low-frequency responses of inner hair cells: evidence for a mechanical origin of peak splitting”, Hearing Research 41, 89–99 (1989)
4. S. Deligeorges and D. C. Mountain, “Computational Neuroscience: Trends in Research”, (edited by Bower, New York: Plenum Press, 1997), pp. 609–616.
5. Rothman JS, Manis PB. The roles potassium currents play in regulating the electrical activity of ventral cochlear nucleus neurons. J Neurophysiol, 2003,89:3097–3113
6. M. H. Jensen, P. Bak and T. Bohr, “Transition to chaos by interaction of resonances in dissipative systems. I. Circle maps”, Physical Review A, 30, 1960–1969 (1984).
7. B. L. Hao, “Applied symbolic dynamics”, Chinese journal of physics, 36, 753–757(1998)
8. H. P. Fang and B. L. Hao, “Symbolic dynamics of the Lorenz equations”, Chaos, Solitons & Fractals, 7, 217–246 (1996)
9. P. Grassberger, H. Kantz and U. Moenig, “On the symbolic dynamics of the Henon map”, J.Phys.A: Math Gen, 22, 5217–5230 (1989).

Part V
Multi-scale Neural Network Dynamics

Representation-Implementation Trade-Off in Cortico-Limbic Ganglio-Basal Loops

Jean-Paul Banquet, Philippe Gaussier, Mathias Quoy, E. Save, F. Sargolini, and B. Poucet

Abstract Unravelling the neural substrates of behavior has made possible to dissociate a high level representation system dedicated to the build-up and storage of a world model, and an implementation system for decision, strategic choices, and sequential behavior. In most ecological situations, particularly in the animal kingdom, a tight functional association between the two blurs their boundaries. Nevertheless, some specific situations like sleep, memory consolidation, planning, or conversely habit performance tax specifically one of the two systems.

Within the paradigm of spatial-temporal learning and navigation are presented the contributions of the main structures of the representation system such as hippocampus, entorhinal, prefrontal and parietal cortices; and of the implementation system, the cortical-striatal loops in particular, monitoring the transition between goal-oriented controlled behavior and automatic habit. The electrophysiological and behavioral results of a continuous navigation task which taxes both systems, as well as goal-oriented and habit spatial-temporal strategies are presented.

Keywords Hippocampus • Entorhinal • Prefrontal • Parietal cortices; limbic • Associative-cognitive • Sensori-motor cortico-striatal loops • Goal-oriented behavior • Habit

Stimulus-response and Tolman cognitive theories which issued from the splitting of Behaviorism in early twentieth century, still provide surprisingly relevant accounts of behavior, as well as a useful framework for unravelling its neural bases. Indeed, cognitivist theory introduces the concepts of representation, and goal-oriented behavior, while S-R paradigm may include habits as a repetition-related end product of the previous mode.

J.-P. Banquet (✉) • P. Gaussier • M. Quoy
Laboratoire ETIS, UMR 8051, ENSEA, Université de Cergy-Pontoise, CNRS,
95302 Cergy Pontoise, France
e-mail: jean-paul.banquet@upmc.fr

E. Save • F. Sargolini • B. Poucet
Laboratoire des Neurosciences Cognitives, UMR 7291, Aix Marseille Université et CNRS,
Marseille, France

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_27

The outstanding progresses in identifying the neural components of behaviour allow to recognize a representation system, comprising in particular medial prefrontal (mPFC), entorhinal (EC) and posterior parietal (PP) cortices, hippocampus (HS), and amygdala, in direct hold with an implementation system made essentially of cortical-striatal-thalamic loops and cerebellum.

Early hierarchical models of information processing make a distinction between the subcortical structures responsible for automatic behavior and the cortical structures that allow representations of event-relationship and value responsible for purposive behavior. The cortical/subcortical hierarchy is supplemented by an intra-cortical hierarchy, which dissociates the posterior sensory areas and the anterior associative areas, such as the prefrontal cortex.

After fronto-striatal loops discovery, cortico-subcortical relationships are often described as segregated, parallel networks (limbic, associative-cognitive, sensori-motor) [1]. This description, very different from the aforementioned hierarchical organization of behavioral and cognitive processes, emphasizes a tight 'vertical' relationship between cortical and subcortical components within distinct, independent, functional channels.

Nevertheless, recent evidence of "spiraling" connections between components of these loops, in particular between striatum and midbrain dopamine (DA) systems, and between thalamic relay nuclei and cortex [2], suggests that the loops are not as closed as originally thought, but also support unidirectional, antero-posterior interactions and integration, supposing oriented transfer of activation and learning between devoted channels. Whereas, in the early models, the frontal pole of the brain, at the top of the hierarchy constitutes the endpoint of the long-range forward cortical connections, the frontal-limbic-striatal system, in the fronto-striatal loop model, is at the origin of a backward-oriented spiraling connectivity, and information transfer.

In this context, while cortico-striatal implementation systems stand at the forefront of neurobiological research, and in particular modeling, the limbic representation system is often under-considered. Indeed, in many situations, this limbic part of the representation system plays a role in the preparation-for-action and behavior. Nevertheless, there are conditions where active behavior is not the purpose of representation processes. Furthermore, in a phylogenetic perspective, it could be proposed that the degree of independence between the representation and the implementation systems stands as an index of 'encephalisation', in the animal kingdom.

We first draw a sketch of the relations of key components of the representation and implementation systems, as well as their functional articulation. Then will be presented the results of an experimental paradigm, the continuous navigation task, where typically spatial and temporal representations are directly bound to action, but also where learning deeply modulates the functional implication of the different systems involved in the task.

1 Representation System

As a prototypical exemplar of the representation system functions, spatio-temporal encoding within archi-, meso- and neo-cortical structures will be considered. These different levels lead to more abstract and complex representations, and eventually, to cognitive maps.

The different stages of spatial representation are relatively well known, even though their relations are not completely understood. Dorsomedial parts of the entorhinal cortex (MEC) contain cells with multiple firing fields organized in a regular grid-like structure of equilateral triangles, coextensive to the explored space (Fig. 1). Grid fields with the same triangular geometry vary across different grid cells, according to spatial frequency (field distance), orientation (tilt angle of the map), and phase (field offset relative to an external reference) [3]. Because grid cells patterns are relatively independent of the environment, and because a small number

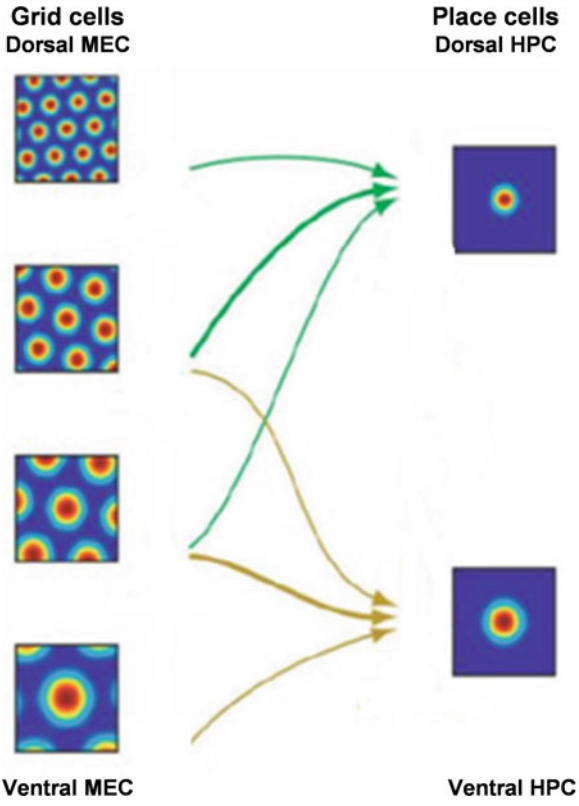


Fig. 1 A gradient of grid patterns spacing from dorsal (*small grids*) to ventral (*larger grids*) cells of the MEC poles, combined to a topographically organized anatomical connectivity between MEC and ventral-dorsal hippocampus results in dorsal (respectively ventral) hippocampus receiving predominant connections from dorsal(respectively ventral) MEC. This could explain the smaller size of the place fields of the dorsal hippocampal cells, compared to ventral ones. *Arrows*' thickness figures connection weights (Adapted from Solstad et al. [6] and Gaussier et al. [5])

of their firing patterns suffice to reconstruct animal's position during navigation, these EC patterns are thought to represent a universal metric of the environment, from which downstream hippocampal (Fig. 1) and cortical spatial representations are derived, even though some authors propose an alternative interpretation [4].

The properties of hippocampal place fields vary according to the exact location where place cells are recorded (i.e. dentate gyrus: DG, CA3, CA1). However place cells share the spatial specificity of their strongly location-related signal, and exquisite sensitivity to the environment and context, either spatial or temporal. Different models have mechanistically demonstrated how the combination of grid cells with different spatial frequencies give rise to DG place fields of different sizes [5, 6]. Further, in an entorhinal-hippocampal loop model [5, 7, 8], spatial and temporal representations combine in CA3 and CA1 to provide a dynamic representation of the animal's navigation, under the form of *transitions* from place to place, rather than pure locations, by associating allothetic and idiothetic information. Transitions form the building blocks for sequence and trajectory encoding.

Indeed, beyond the hippocampal stage, and in particular in deep MEC layers, strict spatial specificity is lost, in favor of trajectory encoding. In particular, in inverted W or in alternating T maze experiments, the firing field of EC pyramidal cells expands to an entire maze arm (Fig. 2). A supra-ordinate factor, like heading direction or path context could integrate successive firing fields in cells combining place and direction or task information.

From HS, spatio-temporal information may be routed to the posterior parietal (PP) cortex through the retrosplenium. In PP neurons, the trajectory coding undergoes further abstraction, becoming trajectory, size and direction independent (Fig. 2); mapping the order of multiple navigational epochs in a route; integrating location and self-movement information.

The head direction system which pervades all these structures seems to provide coherence between allo-centric and ego-centric representations.

Functional dissociation between representation and implementation systems can result from at least two different rationales: the process has basically nothing to do with action, like music listening, meditation in humans . . . ; the process relates to action, but the brain is in a memory or planning mode, and implementation is not relevant, like in quiescent phases or sleep, in rodents; in this case, it seems that two opposite modulations take place: first, activation of the cortical areas where memories are recorded, and simultaneously inhibition of the subcortical implementation systems.

2 Implementation Structures

The phylogenetic evolution of the distinct cortical structures (archi-, paleo-, meso-, neo-cortex) takes place through the dual process of increased complexity (e.g., number of layers) and inclusion of older structures into more recent ones. The same

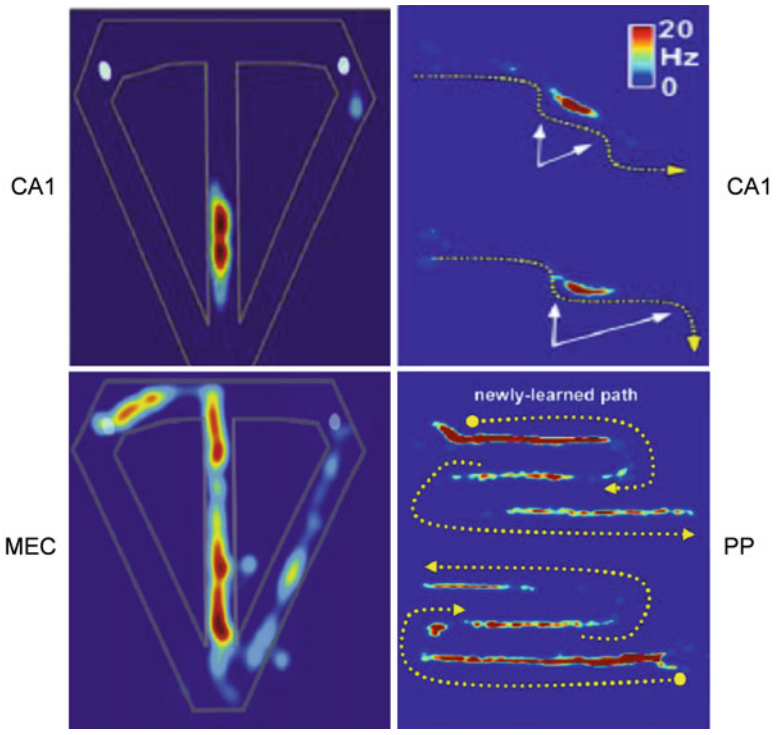


Fig. 2 Place fields in different anatomical structures: CA1, MEC, or Posterior Parietal cortex (PP) in different maze experiments. *Left*: discharge rate-maps for CA1/MEC neurons in rats performing a T-maze alternation task for right-to-left trials, to get reward at the white hole. *Right*: same rate-maps for outbound (CA1) and outbound (*upper plots*) and inbound (*lower plot*) (PP) traversal of a newly-learned path shown by *broken yellow lines*. For each of the two traversals the PP neuron discharges along each path's entire first segment and the final half of the last segment. While the spatial specificity of the place fields of CA1 place cells remains very strong whatever the geometry of the maze, this specificity is partly lost in MEC and furthermore in PP neurons, in favor of a more functional, task-related significance (Adapted from Eichenbaum et al. [9] and Nitz [10])

type of inclusion is found within ventral and dorsal (neo)striatum, through (older) patch and (recent) matrix compartments, such that authors suppose that cortical evolution may have influenced striatal evolution. This could account for the tight functional links between the two structures. Indeed, there is an anatomic, hodologic, and functional coherence between the different compartments, limbic, associative-cognitive and sensori-motor, at all levels of the cortico-ganglio-basal loops: cortex, striatum, thalamus.

Following the characterization of cortico-striato-thalamic loops [1], the functions of the different striatal compartments have been progressively refined, particularly in relation to reinforcement learning, pavlovian conditioning, and instrumental conditioning in goal oriented behavior. Thus, the nucleus accumbens (ACu) has been more particularly associated to Pavlovian conditioning and stimulus-outcome

(SO) association. The core is involved in preparatory CR and anticipatory approach; core lesions impair drug-seeking behavior triggered by drug-associated reinforcers. The shell mediates consummatory CRs and hedonic URs; its lesions preserve drug-seeking acquisition.

Similarly, dorso-medial striatum (DMS), involved in instrumental conditioning, during goal-oriented actions, plays a long-lasting role in the acquisition and expression of action-outcome (AO) learning, in clear contrast with the short-lived function of the medial prefrontal cortex (mPFC) in the same conditions. It contributes to the selection and learning of situations representing valuable parts of the task. The dorsolateral striatum (DLS) transforms goal-oriented repetitive behaviors into habits and skills, and contributes to learning behavioral sequences in general.

However, this view of dedicated functions of the striatal components of the loops can only be partial if the spiraling connections between the different loops are not taken into account [2]. Striatal neurons send direct inhibition to DA neurons from which they receive projections, and also disinhibitory connections to DA neurons projecting to the distinct next striatal area, allowing unidirectional activation-propagation from limbic to associative to motor loops (Fig. 3). A learned conditioned stimulus (CS) could simultaneously suppress a ventral tegmental area (VTA) DA learning signal, and potentiate a substantia nigra compacta (SNc) DA signal in the next cognitive loop. Thus, limbic striatal loops are in a position to control learning and processing within associative and sensori-motor cortico-striatal networks.

Pavlovian Instrumental Transfer (PIT) resorts to the same principle of learning-transfer from one structure to the next: after separate learning of a classical Pavlovian SO contingency, and of an instrumental AO contingency, with the same

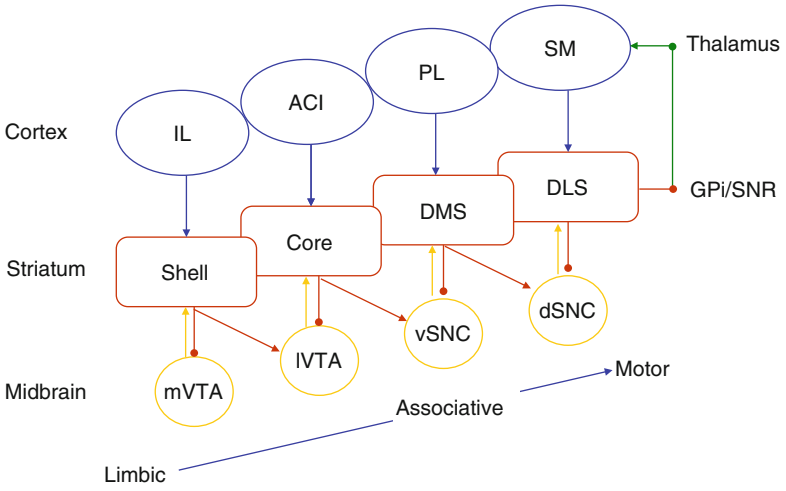


Fig. 3 Spiraling connections between striatal regions and midbrain DA system. *Abbreviations*: *ACI* Anterior cingulate, *DMS(DLS)* dorso-median (lateral) striatum, *IL(PL)* infralimbic (prelimbic), *mPFC* medial prefrontal cortex, *m(l)VTA* medial (lateral) ventral tegmental area, *v(d)SNC* ventral (dorsal) substantia nigra compacta (Adapted from Yin et al. [11])

outcome, the expression of the instrumental contingency is potentiated by the concomitant presentation of the CS, but not the reverse. The integrity of core and shell is necessary for the expression of general and specific aspects of PIT, respectively.

3 Experimental Results

In the continuous place-navigation task, the rat must reach an unmarked goal location in an open arena with a single polarizing cue card. At this goal location, it must stay immobile for a 2 s delay. A food pellet is then delivered by a food dispenser above the arena. As it bounces when hitting the ground, it can end anywhere in the arena. Therefore, the rat must explore the arena to find the food pellet.

Despite its apparent simplicity, this paradigm combines two distinct instrumental conditioning tasks: -goal-oriented navigation to the virtual goal zone, and -foraging to find the food pellet; and also two variants of Pavlovian conditioning: -classical (CS=click of food dispenser; US=food pellet); -secondary conditioning (CS=Goal zone landmark configuration; US=Click). Therefore, all cortical and subcortical areas presented in the branching tree describing learning and performance in general (Fig. 4) should be concerned, if we suppose that goal navigation becomes an habit after overtraining.

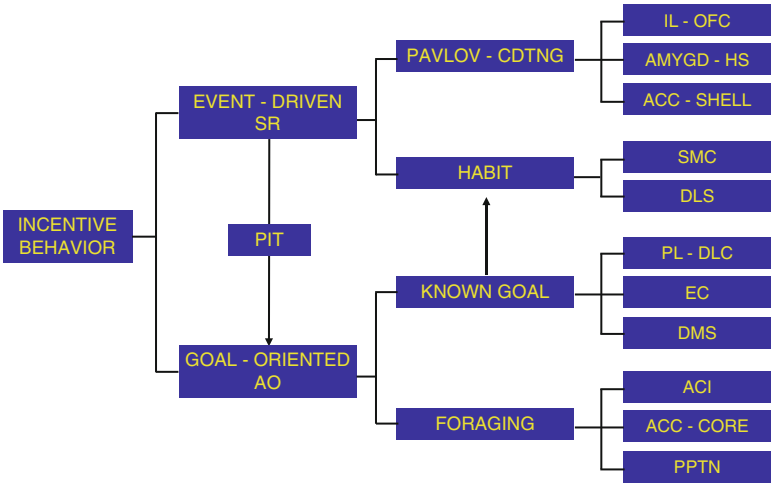


Fig. 4 The different types of behaviors and their interrelations are associated to the structures that support them. *Abbreviations: ACC* accumbens, *AMYGD* amygdala, *AO* action-outcome, *DLC* dorsolateral prefrontal cortex, *EC* entorhinal cortex, *OFC* orbitofrontal cortex, *PIT* pavlovian-instrumental transfer, *PPTN* pedonculo-pontine tegmental nucleus, *SMC*, sensori-motor cortex, *SR* stimulus-response

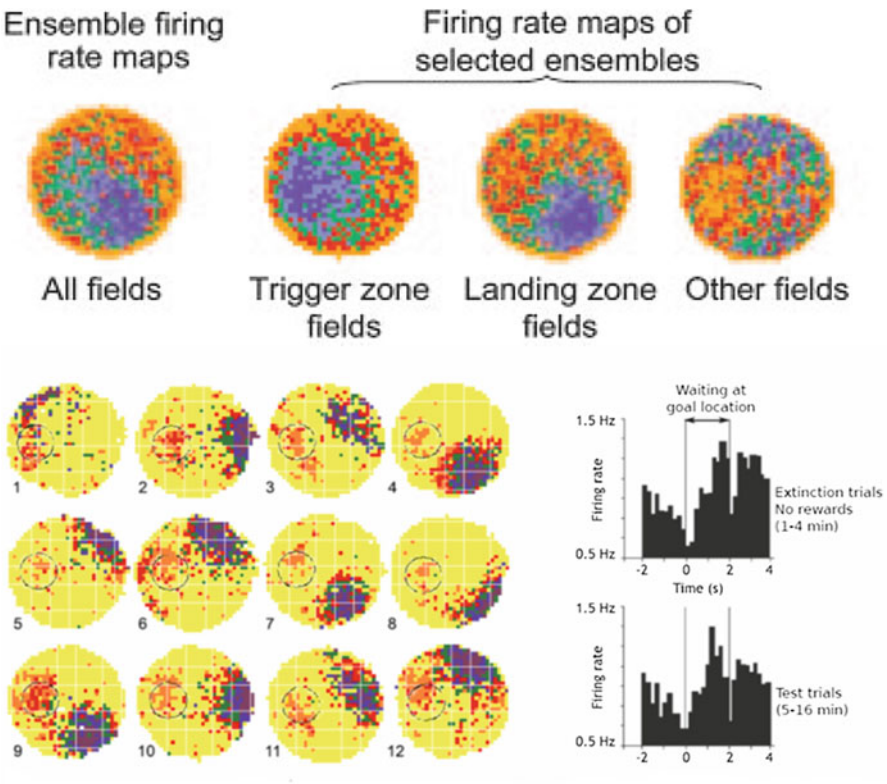


Fig. 5 *Upper*: place fields in PL-IL neurons of mPFC (Hok et al. [12]); *-Lower left*: spatial activity of hippocampal CA1 place cells (Hok et al. [12]). *Circles* mark goal location and secondary fields; *Lower right*: cumulative PETHs for all recorded CA1 place cells at the goal

Most of these results challenge classical knowledge on navigation: First, PL/IL rather than anterior cingulate neurons of mPFC have clear spatial correlates, in particular at the goal and landing zones [12, 13] (Fig. 5). This is all the more important that no spatial correlates were found in a simple foraging task [14].

In HS, place fields were not overrepresented in the goal zone, as expected. Yet, after overtraining, HS place cells presented, in addition to their location-specific main place field, a weaker secondary field at the goal location as rats were waiting for the required 2 s [13] (Fig. 5).

Third, HS secondary fields, as well as PL goal cell activity displayed a temporal profile reaching a maximum just prior the end of the 2 s waiting period (Fig. 5)

Finally, ventral HS inactivation suppressed PL place and timing activity [15], whereas mPFC inactivation did affect neither HS place cell timing activity nor secondary fields [16].

These results confirm the supposed function of mPFC in combining place and valence information to define goals; yet, confirming previous results, this function

appears to be transient and limited to early learning stages. Indeed, inactivation of mPFC after overtraining does not affect behavioral performance [16]. These findings suggest both bottom-up and top-down information transfer between mPFC and HS; top-down transfer could be expressed by the secondary fields of HS place cells. Moreover, for the first time, to the best of our knowledge, the very same HS pyramidal cells are shown to display simultaneously spatial and timing codes, which are conveyed to neocortical structures. Space and time, possibly through frequency modulation of the electrical field potentials could form a common frame for the coordination, and eventually the synchrony of distant brain structures. As a whole, these results shed new light on the role of the PF-HS circuits in goal-oriented and other types of navigation.

Acknowledgments This work was supported by the ANR project NEUROBOT (ANR-BLAN-SIMI2-LS-100617-13-01) and the AUTOEVAL Digiteo project. Many thanks to A. Gissler for assistance in the design and implementation of the illustrations.

References

1. Alexander, G.E., Delong, M.R. & Strick, P.L. (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu. Rev. Neurosci., 9, 357–381.
2. Haber, S.N. (2003) The primate basal ganglia: parallel and integrative networks J. Chem. Neuranat., 26, 317–330.
3. Fyhn, M., Molden, S., Witter, MP, Moser, El., Moser, MB. (2004) Spatial representation in the entorhinal cortex. Science 305:1258–1264.
4. Poucet, B, Sargolini, F., Song, E. Y., Hangya, B., Fox, S. E., Muller, R. (2013) Independence of landmark and self-motion guided navigation: a different role for grid cells. Submitted to Phil. Trans.R. Soc. B-Issue.
5. Gaussier, P., Banquet, JP.,Sargolini, F., Giovannangeli, C., Save, E., Poucet, B. (2007) A model of grid cells involving extrahippocampal path integration and the hippocampal loop. J of Integrative Neuroscience, 6(3), 447–476.
6. Solstad, T., Moser, E.I., and Einevoll, G.T. (2006) From grid cells to place cells: A mathematical model. Hippocampus 16:1026–1031.
7. Banquet, J. P., Gaussier, P., Quoy, M., Revel, A., Bur nod, Y. (2005). A hierarchy of associations in hippocampo-cortical systems: cognitive maps and navigation strategies. Neural Comput., 17(6):1339–1384.
8. Hirel, J., Gaussier, P., Quoy, M., Banquet, J.P., Save, E. & Poucet, B. (2013) The hippocampo-cortical loop : Spatio-temporal learning and goal-oriented planning in navigation. Neural Netw., 43:8–21.
9. Eichenbaum, H., Sauvage, M., Fortin, N., Komorowski, R., Lipton, P. (2012). Towards a functional organization of episodic memory in the medial temporal lobe. Neuroscience and Biobehavioral Reviews 36: 1597–1608.
10. Nitz, D.A. (2006) Tracking Route Progression in the Posterior Parietal Cortex. Neuron 49: 747–756.
11. Yin, H.H., Ostlund, S.B. & Balleine, B.W (2008). Re ward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. European J. of Neurosci 28:1437–1448.
12. Hok V, Save E, Lenck-Santini PP, Poucet B (2005) Coding for spatial goals in the prelimbic/infralimbic area of the rat frontal cortex. Proc Nat Acad Sci U S A 102:4602–4607.

13. Hok V, Lenck-Santini PP, Roux S, Save E, Muller RU, Poucet B (2007) Goal-related activity in hippocampal place cells. J Neurosci 27: 472–482.
14. Poucet, B. (1997). Searching for the spatial correlates of unit firing in the prelimbic area of the rat medial frontal cortex. Behav. Brain Res., 84:151–159.
15. Burton BG, Hok V, Save E, Poucet B (2009) Lesion of the ventral and intermediate hippocampus abolishes anticipatory activity in the medial prefrontal cortex of the rat. Behav Brain Res 199:222–234.
16. Hok, V., Chah, E., Save, E., & Poucet, B. (2013). Prefrontal cortex focally modulates hippocampal place cell firing patterns. J Neurosci 33:3443–3451.

Source Differences in ERP Components Between Pain and Tactile Processing

Yong Hu, Wutao Lou, Weiwei Peng, Li Hu, Zhiguo Zhang, and Jane Z. Wang

Abstract It is a promising method to use intra-epidermal electrical stimulation (IES) as a practical stimulation for nociceptive pain study. There is a question on whether IES involve other sensory receptors, e.g. tactile. This paper is to investigate whether IES could selectively activate Aδ-nociceptors during an oddball paradigm from the perspective of cortical sources by using sLORETA, and moreover, the source differences between nociceptive and non-nociceptive processing. The results provide further evidences to support that IES could effectively activate Aδ-nociceptors selectively.

Keywords Intra-epidermal electrical stimulation • Perspective of cortical sources • sLORETA • Nociceptive processing • Non-nociceptive processing

1 Introduction

Pain is an unpleasant experience that involves the conscious awareness of noxious sensations, hurting and aversive feelings associated with actual or potential tissue damage [1, 2]. The wide definition of pain recognizes that pain results not just from the physical insult but also from a combination of physical, emotional, psychological, and social abnormalities. Pain can therefore be expected to influence brain processing on many levels.

Y. Hu (✉) • W. Lou • W. Peng • L. Hu
Department of Orthopaedics and Traumatology, The University of Hong Kong, Pokfulam, Hong Kong, People's Republic of China
e-mail: yhud@hotmail.com

Z. Zhang
Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, Hong Kong

J.Z. Wang
Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_28

In the present study, we applied painful stimulation using intra-epidermal electrical stimulation (IES) to activate nociceptors selectively [3–7], because electrical pulses delivered through needle electrode inserted in the epidermis could excite selectively Aδ and C fibre (i.e. without coactivating Aβmechanoreceptors that locate deeper than the epidermis). With an increase stimulation, the Aβmechanoreceptors will be excited and Aβfiber will be inhibited, which induces tactile sensation [6].

We recorded 64-channel EEG responses from an oddball task paradigm, in which sources of different ERP components were compared to investigate the difference of somatosensory processing between pain and tactile inputs.

2 Method/Models

Eighteen right-handed healthy volunteers (nine females), aged from 19 to 29 years (21.8 ± 2.5, mean $\pm$SD), took part in the experiment.

Two rapidly-succeeding constant current square-wave pulses will be delivered in the right hand dorsum. The inter-pulse interval will be 10 ms. Each pulse will last 50 μs. Stimuli will be delivered using a stainless steel concentric bipolar needle electrode, consisting of a needle cathode (length: 0.1 mm, Ø: 0.2 mm) surrounded by a cylindrical anode (Ø: 1.4 mm). By gently pressing the device against the skin, the needle electrode will be inserted into the epidermis. In this study, two different intensities were used: (1) twice the subject's perceptive threshold, (2) twice the subject's somatosensory threshold, 5 mA. The stimulus intensity was two times of the individual perceptual threshold, which was proved to be able to selectively activate the Aδ nociceptive fibers without coactivation of the fast-conducting Aβ fibers.

Since multiple cortical regions involve in the processing of pain/somatosensory stimuli and sLORETA has been demonstrated as be a feasible tool for pain research in previous studies, sLORETA was used to estimate the source localization.

Four ERPs components, N60, P100, N120, P300, were selected for further source localization analysis, while source difference were compared between pain and somatosensory inputs.

3 Results

By using sLORETA, the source analysis of the four components were shown in Fig. 1. For early components N60 and P100, the source difference between somatosensory and pain was mainly focused on BA3, i.e., ipsilateral SI area. For N120, the difference was mainly focused on BA4, that is the ipsilateral primary motor cortex. For P300, the difference was mainly focused on BA1, that is the contralateral SI area.

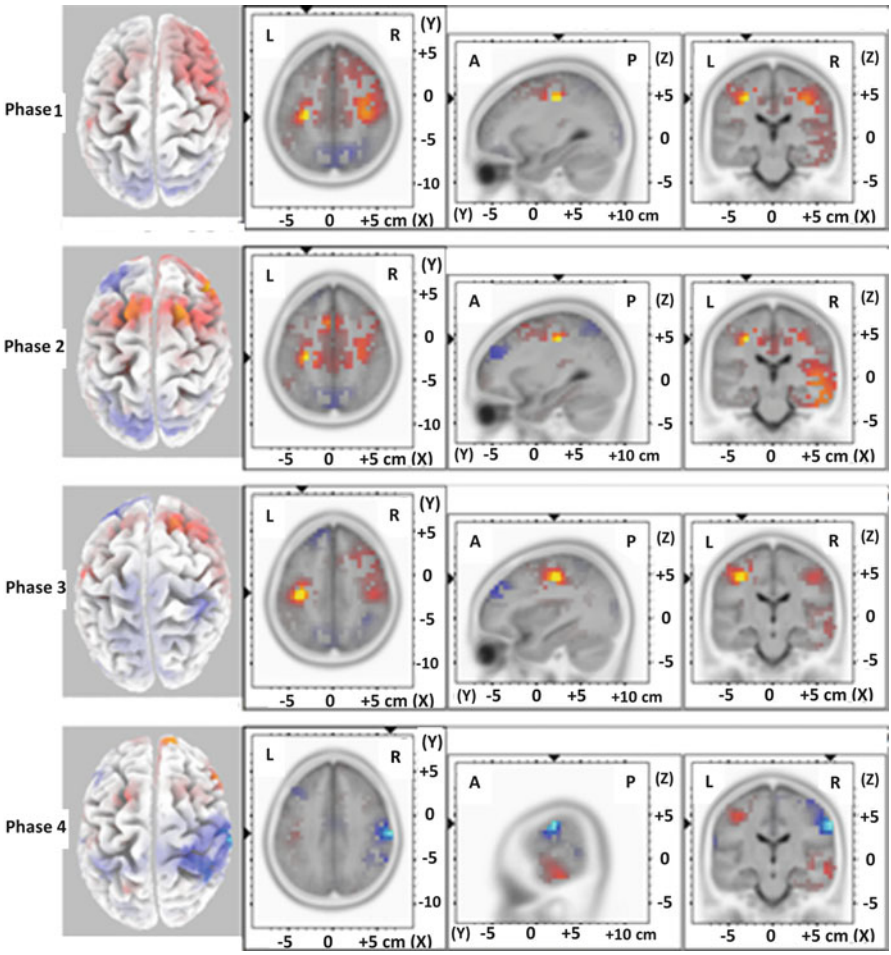


Fig. 1 Source difference between pain and tactile inputs

4 Discussion and Conclusions

During the early part of non-nociceptive/nociceptive processing, in addition to the decreased activity in SI area, the decreased contralateral frontal and left precuneus activation in the second and third phase was also found respectively. During phase 4, only significant decreased brain activity in the contralateral SI area was found for nociceptive condition compared to non-nociceptive condition. Component in this phase is consistent with the well-established ERP component of N2. The locations of N2 were consistently reported to be mainly focused on the bilateral SII areas by dipolar source analysis [8, 9]. During phase 4, only significant decreased brain activity in the contralateral SI area was found for nociceptive condition compared

to non-nociceptive condition. The role of late SI activity in pain processing is still unknown. In this phase, the nociceptive/non-nociceptive stimuli have been transferred to pain/somatosensory perception, and subjects would detect the spatial of stimulation sites so that they could make decision about whether the stimulus is target or not in next step.

The promising result suggests the different processing and pathway between pain and tactile inputs. Further study with time-varying source connectivity would provide a useful tool to identify pain and tactile sensations.

Acknowledgments This work was supported in part by National Natural Science Foundation of China (No. 81271685).

References

1. Brooks, J. C., T. J. Nurmikko, W. E. Bimson, K. D. Singh and N. Roberts (2002). "fMRI of thermal pain: effects of stimulus laterality and attention." Neuroimage **15**(2): 293–301.
2. Apkarian, A. V., M. C. Bushnell, R. D. Treede and J. K. Zubieta (2005). "Human brain mechanisms of pain perception and regulation in health and disease." Eur J Pain **9**(4): 463–484.
3. Kaube, H., Z. Katsarava, T. Kaufer, H. Diener and J. Ellrich (2000). "A new method to increase nociception specificity of the human blink reflex." Clin Neurophysiol **111**(3): 413–416.
4. Inui, K., T. D. Tran, M. Hoshiyama and R. Kakigi (2002). "Preferential stimulation of Adelta fibers by intra-epidermal needle electrode in humans." Pain **96**(3): 247–252.
5. Inui, K., T. Tsuji and R. Kakigi (2006). "Temporal analysis of cortical mechanisms for pain relief by tactile stimuli in humans." Cereb Cortex **16**(3): 355–365.
6. Mouraux, A., G. D. Iannetti and L. Plaghki (2010). "Low intensity intra-epidermal electrical stimulation can activate Adelta-nociceptors selectively." Pain **150**(1): 199–207.
7. Inui, K. and R. Kakigi (2012). "Pain perception in humans: use of intraepidermal electrical stimulation." J Neurol Neurosurg Psychiatry **83**(5): 551–556.
8. Hu, L., Z. G. Zhang and Y. Hu (2012). "A time-varying source connectivity approach to reveal human somatosensory information processing." Neuroimage **62**(1): 217–228.
9. Hu, L., W. Peng, E. Valentini, Z. Zhang and Y. Hu (2013). "Functional features of nociceptive-induced suppression of alpha band electroencephalographic oscillations." J Pain **14**(1): 89–99.

The Time-Varying Causal Coupling in Brain and Organization of Its Networks

Maciej Kaminski, Katarzyna J. Blinowska, Aneta Brzezicka, Jan Kaminski, and Rafal Kus

Abstract For investigation of time-varying brain networks an approach based on estimation of causal coupling by means of multivariate method was applied. Two cognitive experiments: Constant Attention Test and Working Memory task are considered. Time varying version of a multivariate estimator—Directed Transfer Function was used for calculating dynamically changing patterns of transmission during the tasks. Well-defined centers of activity congruent with imaging, anatomical and electrophysiological evidence were found. These centers exchanged the information only during short epochs. The strengths of coupling inside the tightly connected modules and between them was found by means of assortative mixing. The results point out to the well determined, far from randomness structure of brain networks in cognitive tasks. Very dense and disorganized structure of networks reported in literature may be explained by the presence of spurious connections produced by bi-variate measures of connectivity and further enhanced by giving all connections equal weights.

Keywords Functional connectivity • Causal coupling • Directed transfer function • Working memory • Dynamical EEG propagation • Directed networks

M. Kaminski • K.J. Blinowska (✉)
Faculty of Physics, University of Warsaw, Hoza 69 St., Warsaw, Poland
e-mail: kjbli@fuw.edu.pl

A. Brzezicka
Interdysciplinary Center for Applied and Cognitive Studies, Warsaw School of Social Sciences and Humanities, Warsaw, Poland

J. Kaminski
Department of Neurology, Nencki Institute of Experimental Biology, Warsaw, Poland

R. Kus
Department of Biomedical Physics, Faculty of Physics, University of Warsaw, Pasteura 5 St., Warsaw, Poland

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_29

1 Introduction

Herein we shall consider the organization of the brain networks in cognitive experiments with the aim to reveal networks topographical structure, strength of coupling, spectral and time-varying properties of transmission. The proposed approach is based on estimation of time-varying causal coupling between EEG signals. To this aim we applied multivariate method based on Granger causality principle—the Directed Transfer Function—DTF [1]. As a next step, for identification of specific brain modules constituting functional networks during the tasks, the assortative mixing approach [2] was used. In this paper we describe applications of DTF to two cognitive experiments: Constant Attention Test (CAT) and reasoning task involving working memory (WM).

2 Method

The Directed Transfer Function is based on calculation of the transfer function of the Multivariate Autoregressive Model (MVAR). The DTF describes causal influence of channel *j* on channel *i* in frequency domain [1]. DTF is robust in respect to noise and unlike bivariate methods does not produce spurious connections arising from the common feeding. Moreover, because DTF is based on phase differences between signals, in the absence of phase difference its value is zero. Volume conduction is a propagation of electromagnetic field, so it does not produce the phase differences on the electrodes, therefore DTF is practically not affected by the volume conduction.

Estimation of time varying version of DTF—SDTF—is based on application of a short sliding data window and ensemble averaging over realizations, which are obtained by repetitions of an experiment. The averaging concerns correlation matrices (the model is fitted independently for each short data window); the data is not averaged in the process [3].

3 Results

The CAT experiment consisted of the presentation of consecutively displayed geometrical patterns. The *target* condition was defined as any immediately repeated pattern, in this case the subject was instructed to press the button with his right thumb. The description of the experiment may be found in [4].

In both cases (*target*/*non-target*) in the initial phase mainly the transmission between prefrontal and frontal electrodes was observed, which are responsible for information storage and working memory. For *target* situation at the end of the task the burst of activity occurred from the hand motor-cortex, which can be interpreted as a signal to move the finger. In case of *non-target* in the second stage of the task the long-range transmission from the electrode overlying right Inferior Frontal Cortex

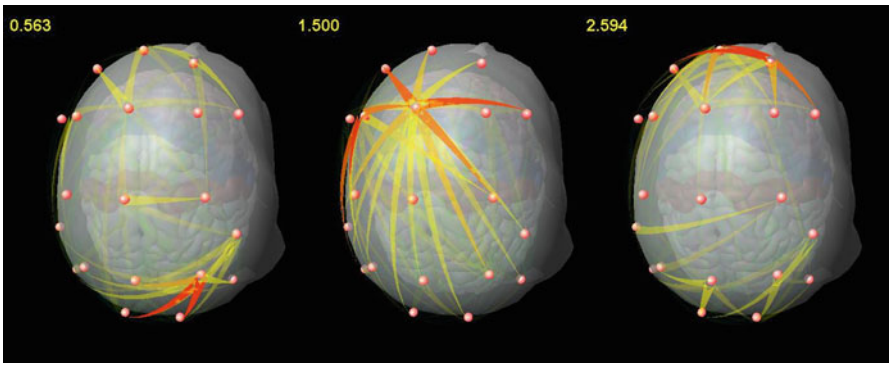


Fig. 1 Snapshots from the animation showing the time-varying pattern of propagations for the representative subject. The numbers in the *upper left* corner correspond to the time [s] after the stimulus presentation

(rIFC) or from premotor Supplementary Cortex (preSMA) to the C3 (finger motor cortex) was observed. Both structures are involved in go/no go tasks. The animations presenting the dynamical propagation of EEG activity during CAT are accessible at the address: http://brain.fuw.edu.pl/~kjbli/CAT_MOV.html.

The reasoning task involving working memory relied on memorizing and retrieval of letters and relations between them [5]. The patterns of propagations obtained by SDTF showed the existence of two main centers of activity located in the frontal and posterior regions, in agreement with the imaging experiment [6]. These two centers communicated by long range connections, which acted during some short periods. The animations showing the dynamics of the task are available at: http://brain.fuw.edu.pl/~kjbli/Cognitive_MOV.html. The snapshots from the animation are displayed in Fig. 1.

We have defined as local: connections between the neighboring electrodes in the (10–20) system and these along the diagonals of squares formed by the neighboring electrodes and as distant all other connections. Inspection of time courses of long range and short range connections revealed that at the moment of the premise presentation for all frequency bands long- and short- range connectivities were in phase and in the consecutive 3 s period they were mostly out of phase. The estimated ratio of short to long connections was in the range: 1.4 for theta to 1.5 for beta rhythm [7]. Frontal, central and two parietal modules were identified and the strength of intra-modular and inter-modular couplings were found.

4 Discussion

In CAT test the transmission between prefrontal and frontal areas observed during the initial phase complies with the involvement of frontal lobes and especially the PFC in information storage and working memory [8]. The observation of the burst

of activity in target condition in beta and gamma band from C3 (contralateral finger motor cortex) is compatible with a well known phenomena of gamma activation and beta rebound connected with hand movement [9]. In case of *non-target* the long-range transmissions occurred either from the electrode F8 overlying rIFC (six subjects) or from premSMA (three subjects) to the finger primary motor cortex. Both structures are known to be connected with "go/no go" tasks [10, 11]. The observation of transmission between distant locations for no-go tasks (e.g. between F8 and C3) is concordant with the hypothesis that inhibition of motor structures originates from long-range cortico-cortical connections [12].

In a similar experiment involving semantic priming task Schinkel et al. [13] applied as a connectivity measure joined recurrence plots followed by graph theoretical analysis. For primed stimuli almost uniform pattern of connections emerged, which was interpreted as existence of one large network component. For unprimed stimuli broadly distributed left-lateralized network components were reported. These results have no support in imaging or electrophysiological evidence.

The role of short- and long-range connections in information processing was even better visible in the working memory experiment. In our study concerning the WM task the involvement of the frontal and posterior parietal regions was observed, which is in an excellent agreement with the imaging studies [6, 14]. Sauseng et al. [15] showed that fronto-parietal coherence in theta and upper alpha reflect central executive functions of working memory.

Working memory tasks were analyzed also by means of pairwise measures of connectivity and graph theoretical analysis. Connectivity in WM tasks (namely 2Back check) was studied by Micheloyanis et al. [16] in EEG experiment and by Kitzbichler et al. [17] by means of MEG. In both studies the connectivity patterns were close to random, with some traits of "small world" structure. The patterns of connections were very dense and did not indicate the brain regions involved in the information processing.

The lack of correspondence of the above works with our findings and other evidence may be explained by the methodological flaws present in the quoted publications. Namely the bivariate measures applied in them produce a lot of spurious connections which was demonstrated in [18, 19]. In fact, because of common feeding effect, spurious connections may outnumber the true ones. If the activity of a given source is recorded at N electrodes N true and $N(N–1)/2$ false connections may be found by bivariate measures. The common practice in graph theoretical analysis, of setting the threshold very low and giving all connections equal weight, further enhances such spurious connections. As a result a very dense, disorganized and close to random architecture of connectivity emerges.

Application of multivariate measure such as DTF allows for determination of connectivity patterns and quantitative description of their structure, including the dependence on frequency and the contribution of short- and long- range interactions. By means of the proposed formalism it was possible: to determine the modular structure of the brain networks, to found the intra- and inter-modular couplings and estimate their strength. Moreover the dynamic patterns of the interactions were identified involving intra-modular persistent coupling during the task, whereas the

coupling between distant locations was less frequent. We have for the first time reported the dynamics of the interaction between the modules, showing that the exchange of information is taking place in certain specific moments.

References

1. Kamiński, M. & Blinowska, K. J. 1991 A new method of the description of the information flow in brain structures. Biol Cybern. 65, 203–210.
2. Newman, M. E. J. 2003 Mixing patterns in networks. Phys. Rev. E 67, 026126.
3. Blinowska, K. J. & Kaminski, M. (2006) Multivariate Signal Analysis by parametric Models. In Handbook of Time Series Analysis (eds. B. Schelter, M. Winterhalder & J. Timmer), pp. 387–420. Wiley-VCH Verlag.
4. K. Blinowska, R. Kus, M. Kaminski, J. Janiszewska (2010) Transmission of Brain Activity During Cognitive Task. Brain Topography 23:205–213.
5. Brzezicka, A., Kaminski, M., Kaminski, J. & Blinowska, K. J. (2010) Information transfer during transitive reasoning task. Brain Topography 24, 1–8.
6. Brzezicka A, Sędek G, Marchewka A, Gola M, Jednoróg K, Królicki L, Wróbel A (2011) A role for the right prefrontal and bilateral parietal cortex in four-term transitive reasoning: An fMRI study with abstract linear syllogism tasks. Acta Neurobiol. Exp. 71 (4), 479–495.
7. Blinowska K.J., Kaminski M., Brzezicka A., Kaminski J. (2013) Application of directed transfer function and network formalism for the assessment of functional connectivity in working memory task. Phil. Trans. Roy Soc. A. 371:20110674.
8. Romo R, Brody CD, Hernandez A, Lemus L (1999) Neuronal correlates of parametric working memory in the prefrontal cortex. Nature 399(6735):470–473.
9. Kuś, R, Ginter J. Jr., Blinowska, K. J. (2006) Propagation of EEG activity during finger movement and its imagination. Acta Neurobiol Exp. 66(3), 195–206.
10. Aron AR, Fletcher PC, Bullmore ET, Sahakian BJ, Robbins TW (2003) Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. Nat Neurosci 6:115–116.
11. Fried I, Katz A, McCarthy G, Sass KJ, Williamson P, Spencer SS, Spencer DD (1991) Functional organization of human supplementary motor cortex studied by electrical stimulation. J Neurosci 11:3656–3666.
12. Burle B, Vidal F, Tandonnet C, Hasbroucq T (2004) Physiological evidence for response inhibition in choice reaction time tasks. Brain Cogn 56:153–164.
13. Schinkel S, Zamora-López G, Dimigen O, Sommer W, Kurths J. (2011) Functional network analysis reveals differences in the semantic priming task J Neurosc. Meth. 107, 333–339.
14. Cabeza R, Nyberg L (2000) Imaging condition II. An empirical review of 275 PET and fMRI studies. J. Cogn. Neurosci. 12, 1–47.
15. Sauseng P, Klimesch W, Schabus M, Doppelmayr M (2005) Fronto-parietal coherence in theta and upper alpha reflect central executive functions of working memory. Int. J. Psychophysiol. 57(2), 97–103.
16. Micheloyannis S, Pachou E, Stam CJ, Vourkas M, Erimaki S, Tsirka V (2006) Using graph theoretical analysis of Multi channel EEG to evaluate the neural efficiency hypothesis. Neurosci. Lett. 402, 273–277.
17. Kitzbichler MG, Henson RNA, Smith ML, Nathan PJ, Bullmore ET (2011) Cognitive effort drives workspace configuration of human brain functional networks, The Journal of Neuroscience, 31(22): 8259–8270.
18. Blinowska KJ, Kuś R, Kamiński M (2004) Granger causality and information flow in multivariate processes. Phys. Rev. E 70, 050902.
19. Kuś R, Kamiński M, Blinowska KJ (2004) Determination of EEG activity propagation: pairwise versus multichannel estimate. IEEE. Trans. Biomed. Eng. 51, 1501–1510.

Enhancement of Weak Signal Detection in Parallel Arrays of Integrate-and-Fire Neurons by Negative Spatial Correlation

Yan-Mei Kang and Yong Xie

Abstract We apply the scheme of linear approximation to calculate the spectral statistics for parallel arrays of Integrate-and-fire neurons with local spatially correlated noise. Our investigation shows the curve of signal-to-noise ratio via noise intensity has more prominent peak when the internal noise correlation is negative, and thus negative correlation has advantage over positive correlation and zero correlation in weak signal detection through the integrate-and-fire arrays. Our investigation should help to understand the functional role of correlated noise.

Keywords Stochastic resonance • Negative correlation • Linear approximation

1 Introduction

The term of stochastic resonance (SR) was initially invoked in 1981 in the explanation of the alternate occurrence of the hot climate period and the cold climate period every 100,000 years [1]. Although the anti-intuitive idea cannot be reproduced in meteorology research, it has been subsequently verified in host of electrical circuit and biophysical experiments [2, 3].

Due to the prevalence of noise in neural system, the discovery that living organisms use SR to detect weak signals in a noisy environment has aroused much interest in the community of neuroscience. For instance, Longtin A checked relation between the residence time distribution and SR in FitzHugh-Nagumo (FHN) neuron model [4], Gong P L et al. disclosed the SR caused by the dynamical bistability in the same model [5], Gong Y F proposed a simple theoretical model for demonstrating the mechanism for weak sensory signal perception [6], and Liu F

Y.-M. Kang
Department of Applied Mathematics, School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China

Y. Xie (✉)
State Key Laboratory of Strength and Vibration for Mechanical Structures, School of Aerospace, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China
e-mail: yxie@mail.xjtu.edu.cn

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_30

et al. used globally-coupled FHN neuronal systems and found that the global noise spatial correlation has an inhibitive effect on signal processing [7]. Because of the complexity in the structure of neural networks and the variability in the origin of noise, the existing investigation has not completely clarified the problem of weak signal detection in neural systems.

As far as the parallel array of neural oscillators is concerned, the previous investigation was focused on independent internal noise, and as for the correlation of the internal noise, the related result is basically in blank. Motivated by the fact that negatively-correlated background noise are less noisy as a whole, we infer that the advantage of local spatial negative correlation over its statistical independent or positively-correlated counterpart should be universal in parallel systems including the neural arrays, and consequently the local spatial negative correlation should be more helpful for neural arrays to detect weak coherent signal. To confirm the above inference in this paper, let us investigate the effect of local negative spatial correlation on SR in parallel arrays of Integrate-and-Fire neurons (IF array).

2 Model and Method

The IF array under study is governed by the following Langevin equation system

$$\tau \dot{v}_i(t) = -v_i(t) + \mu + \sqrt{2D}\eta_i(t) + s(t) \tag{1}$$

for $1 \le i \le N$, where v_i denotes the membrane voltage of neuron i, μ is a dc component in the noisy synaptic input and τ is the membrane time constant for the subthreshold dynamics. We adopt the standard threshold-spike-reset condition [8]: a spike is emitted whenever $v_i(t) = V_T$, and after that the voltage is reset and held at $v_i(t^+) = V_R$ during a refractory period τ_R. Additionally, $s(t)$ is a common component in the input and the Gaussian white noise $\eta_i(t)$ $(1 \le i, j \le N)$ models the internal stochastic component for the ith neuron. For each neuron, the output spike train $y_i(t) = \sum_k s(t - t_i^k)$ with t_i^k being the kth spike time is of interest, and the average spike train $y(t) = \frac{1}{N}\sum_i y_i(t)$ is taken as the output for summing parallel array. We assume that for $1 \le i, j \le N$ there is

$$\left\langle \eta_i(t)\eta_j(t+\tau) \right\rangle = \left[\delta_{i,j} + c\delta_{i,i+1}\right]\delta(\tau) \tag{2}$$

with c being a tunable nearest-neighborhood correlation coefficient.

Since the local spatial correlation does not affect the response of each neuron, it is necessary to obtain the spectral statistics $G_{yy}(\omega) = \langle \tilde{y}(\omega)\, \tilde{y}^*(\omega) \rangle$ for the ensemble output spike train, where $\tilde{y}(\omega) = \frac{1}{\sqrt{T}} \int_0^T dt e^{i\omega t} (y(t) - r_0(D))$ is the

Fourier transform of the zero average output spike train with $r_0(D)$ being the stationary firing rate at the noise level D. For this aim, let us resort to technique of linear approximation [9–11]. According to linear approximation, each neuron can be regarded as a linear filter, and thus the frequency domain linear response for each neuron can be approximated as

$$\tilde{y}_i(\omega) = \tilde{y}_{i,0}(\omega, D) + A(\omega, D)\, s(\omega) \tag{3}$$

where $\tilde{y}_0(\omega)$ is the unperturbed part of stationary spectral density $\left\langle \tilde{y}_i^0(\omega)\, \tilde{y}_i^{0*}(\omega) \right\rangle$ and $A(\omega, D)$ is linear susceptibility on the noise level of D. The stationary spectral density [12] and the linear susceptibility [13, 14] are explicitly given as

$$A(\omega, \mu, D) = \frac{r(D) i\omega}{\sqrt{D}(i\omega - 1)} \frac{\tilde{D}_{i\omega-1}\left(\frac{\mu - v_T}{\sqrt{D}}\right) - e^{\gamma} \tilde{D}_{i\omega-1}\left(\frac{\mu - v_R}{\sqrt{D}}\right)}{\tilde{D}_{i\omega}\left(\frac{\mu - v_T}{\sqrt{D}}\right) - e^{\gamma} e^{i\omega\tau_R} \tilde{D}_{i\omega}\left(\frac{\mu - v_R}{\sqrt{D}}\right)} \tag{4}$$

$$S_0(\omega, \mu, D) \triangleq \left\langle \tilde{y}_i^0(\omega)\, \tilde{y}_i^{0*}(\omega) \right\rangle = r(D) \frac{\left|\tilde{D}_{i\omega}\left(\frac{\mu - v_T}{\sqrt{D}}\right)\right|^2 - e^{2\gamma}\left|\tilde{D}_{i\omega}\left(\frac{\mu - v_R}{\sqrt{D}}\right)\right|^2}{\left|\tilde{D}_{i\omega}\left(\frac{\mu - v_T}{\sqrt{D}}\right) - e^{\gamma} e^{i\omega t_R} \tilde{D}_{i\omega}\left(\frac{\mu - v_R}{\sqrt{D}}\right)\right|^2} \tag{5}$$

with $\gamma = \left[v_R^2 - v_T^2 + 2\overline{\mu}(v_T - v_R)\right]/4D$. And then from Eq. (3), the auto-spectral density $S_{ii}(\omega) = \left\langle \tilde{y}_i(\omega)\, \tilde{y}_i^*(\omega) \right\rangle$ for the ith neuron is obtained as

$$S_{i,i}(\omega) = |A(\omega, D)|^2 G_{ss}(\omega) \tag{6}$$

with $G_{ss}(\omega)$ standing for the spectral density of the input signal.

For calculating the cross-spectral density between the ith neuron and the (i + 1)th neuron, we decompose the corresponding Langevin equations into

$$\begin{cases} \dot{v}_i(t) = -v_i(t) + \mu + \sqrt{2D}\sqrt{1 - c_0}\xi_i(t)\Big) + \left\{\sqrt{2D}\sqrt{c_0}\eta(t) + s(t)\right\} \\ \dot{v}_{i+1}(t) = -v_{i+1}(t) + \mu + \sqrt{2D}\sqrt{1 - c_0}\xi_{i+1}(t) \\ \qquad\qquad + \sqrt{2D}\left\{\mathrm{sgn}(c)\sqrt{c_0}\eta(t) + s(t)\right\} \end{cases} \tag{7}$$

with sgn(·) being sign function and $c_0 = |c|$ being the absolute value. Now $\xi_i(t)$, $\xi_{i+1}(t)$ and $\eta(t)$ in Eq. (7) are mutually independent Gaussian white noises. Taking the terms in brackets as perturbation, then the cross-spectral density $S_{ii+1}(\omega) = \left\langle \tilde{y}_i(\omega)\, \tilde{y}_{i+1}^*(\omega) \right\rangle$ can be approximated as

$$S_{i,i+1}(\omega) = |A(\omega, D(1 - c_0))|^2 \left\{2cD + G_{ss}(\omega)\right\} \tag{8}$$

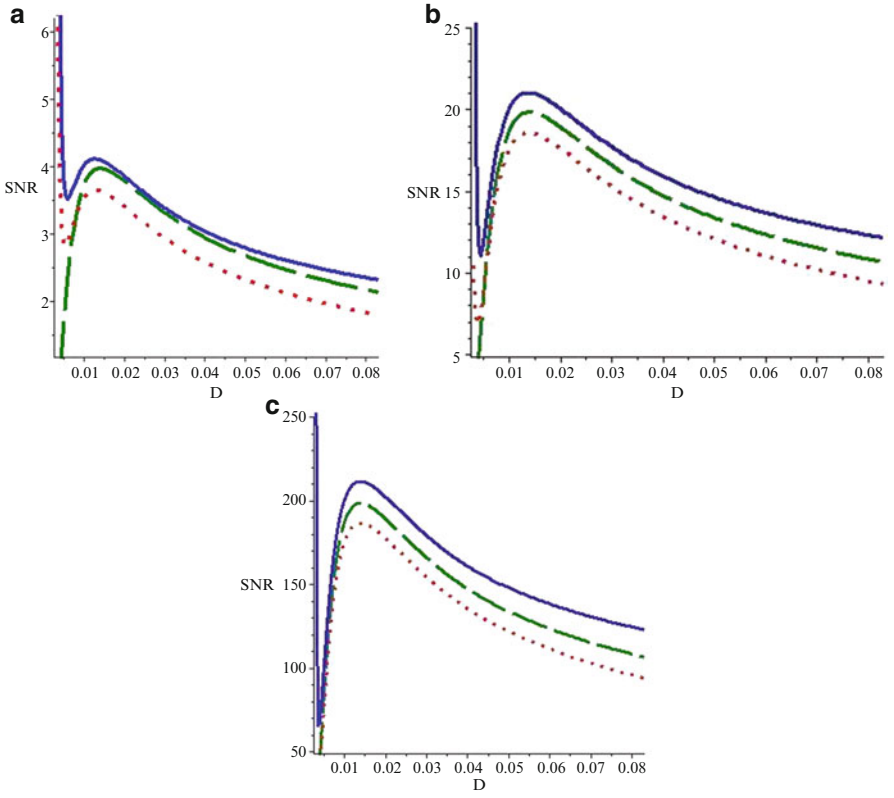


Fig. 1 Dependence of the output SNR for the IF array on noise intensity with the nearest-neighborhood correlation. The parameters $V_T = 1$, $V_R = 0$, $\tau_R = 0.1$, $\mu = 0.8$, $\Omega = 0.1$ and $\varepsilon = 0.05$ and the array size is (**a**) N = 20; (**b**) N = 100; (**c**) N = 1,000. The correlation parameter is 0.2 (*red dot*), 0 (*green dash*) and −0.2 (*blue solid*), respectively

with $G_{ss}(\omega)$ being the spectral density of the input coherent signal. Using Eqs. (6) and (8), the spectral-density for the parallel array reads

$$G_{yy}(\omega) = \frac{1}{N}\left\{S_0(\omega, D) + |A(\omega, D)|^2 G_{ss}(\omega)\right\} + \frac{N^2-3N+2}{N^2}|A(\omega, D)|^2 G_{ss}(\omega) + \frac{2(N-1)}{N^2}\left\{|A(\omega, D(1-c_0))|^2\left[2cD + G_{ss}(\omega)\right]\right\}. \tag{9}$$

Therefore by taking $s(t) = \varepsilon\cos(\Omega t)$ ($\varepsilon \ll 1$), we obtain the output signal-to-noise ratio (SNR)

$$SNR_{out} = \frac{\left\{\left(N^2 - 2N + 2\right)\left|A(\Omega, D)\right|^2 + 2(N-1)\left|A(\Omega, D(1-c_0))\right|^2\right\}\pi^2\varepsilon^2}{NS_0(\Omega, D) + 4cD(N-1)\left|A(\Omega, D(1-c_0))\right|^2}. \tag{10}$$

3 Result and Discussion

To investigate the effect of local spatial correlation on SR, Fig. 1 shows the evolutionary curve of SNR in Eq. (10) via noise intensity. From Fig. 1, it is clear that negative correlation is the best among the three correlation cases for enhancing SR using the parallel IF array in noise environment. To further check this conclusion, Fig. 2 gives the results from stochastic simulation, and again the same conclusion is observed.

We have investigated the phenomenon of SR in IF arrays with internal spatially correlated noise. Although result from the analytic approximations which could be improved has certain quantitative difference from that from stochastic simulation, our investigation shows that negative correlation is optimal for detecting weak signal among the three cases of statistically independence, positive correlation and negative correlation. Additionally, our investigation also demonstrates that local spatial negative correlation will inhibit the phenomenon of synchronization, which is completely different from the previous investigation on global spatial correlation enhancing synchronization [10], and this difference should reflect the real deviation between local spatial correlation and global spatial correlation. Noise correlation is prevalent in neural systems, and very recently negative correlation has attracted novel interest in the community of neuroscience [15]. We suggest that the conclusion in this paper might be meaningful in understanding the functional role of correlation.

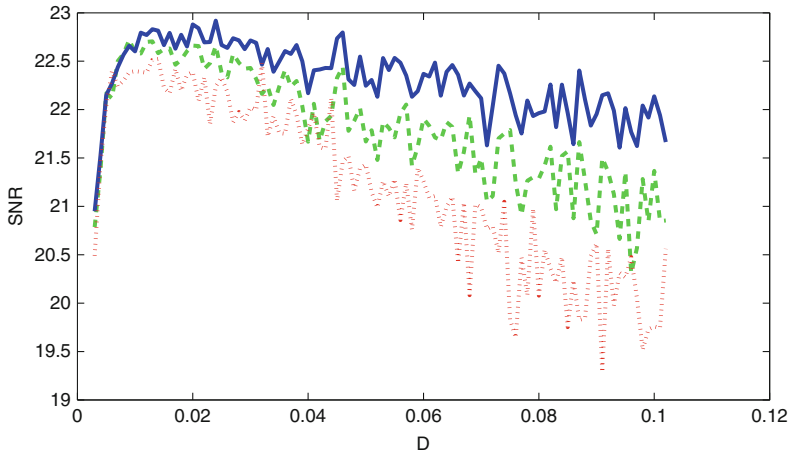


Fig. 2 Dependence of the output SNR derived from stochastic simulation for the IF array on noise intensity with the nearest-neighborhood correlation. The parameters $V_T = 1$, $V_R = 0$, $\tau_R = 0.1$, $\mu = 0.8$, $\Omega = 0.1$ and $\varepsilon = 0.1$ and the array size is $N = 20$. The correlation parameter is 0.3 (*red*), 0 (*green*) and -0.3 (*blue*), respectively

Acknowledgements The authors appreciate the support of NSFC (Grant Nos.11072182 and 11272241).

References

1. Benzi R, Sutera A, Vulpiani A. The mechanism of stochastic resonance. J. Phys. A 14, L453–L457(1981)
2. Gammaitoni L, Hanggi P, Jung P, et al. Stochastic resonance. Rev Mod. Phys. 70: 223–287(1998)
3. Douglass J K, Wilkens L, Pantazelou E, Moss F. Noise enhancement of information transfer in crayfish mechanoreceptors by stochastic resonance. Nature 365, 337–340(1993)
4. Longtin A. Stochastic resonance in neuron models. Journal of Statistical Physics 70(1/2), 309–327(1993)
5. Gong P L, Xu J X. Global dynamics and stochastic resonance of the forced FitzHugh-Nagumo neuron model. Phys Rev E 63, 031906(2001)
6. Gong Y F, Matthews N, Qian, N. A model for stochastic resonance-type behavior in sensory perception. Phys Rev E 65(3), 031904(2002)
7. Liu F, Hu B, Wang W. Effects of correlated and independent noise on signal processing in neuronal systems. Phys. Rev. E 63(3), 031907 (2001)
8. Kang Y M, Xu J X, Xie Y. Signal-to-noise ratio gain of a noisy neuron that transmits subthreshold periodic spike trains. Phys. Rev. E 72, 021902(2005)
9. Lindner B, Doiron B, Longtin A. Theory of oscillatory firing induced by spatially correlated noise and delayed feedback. Phys. Rev. E 72, 061919 (2005)
10. Shea-Brown E, Josic K, Rocha J, Doiron B. Correlation and synchrony transfer in integrate-and-fire neurons: basic properties and consequences for coding. Phys. Rev. Lett. 100, 108102(2008)
11. Trousdale J, Hu Y, Shea-Brown E, Josic K. Impact of network structure and cellular response on spike time correlations. PLoS Comput Biol 8(3):e1002408(2012)
12. Lindner B, Schimansky-Geier L, Longtin A. Maximizing spike train coherence or incoherence in the leaky integrate-and-fire neuron. Phys. Rev. E. 66, 031916(2002)
13. Brunel N, Chance F S, Fourcaud N, Abbott L F. Effects of synaptic noise and filtering on the frequency response of spiking neurons. Phys. Rev. Lett. 86(10), 2186–2189(2001)
14. Lindner B, Schimansky-Geier L. Transmission of noise coded versus additive signals through a neuronal ensemble. Phys. Rev. lett. 86(14), 2934–2937(2001)
15. Theunissen F E, Elie J E. Population code, noise correlations, and memory. Neuron 78, 209–210(2013)

A Computational Model of Hippocampal-VTA Microcircuit: Why Expectation of Reward in Rat Striatum at Choice Point Is Covert?

Yongtao Li and Ichiro Tsuda

Abstract Hippocampal-VTA microcircuit is one of the most important components of reward system. As a prominent feature in hippocampus, theta rhythm involves a strong correlation with learning, memory and decision making. However, the relation between theta rhythm and reward representation remains unclear. Based on some recent experimental discovery, a computational model of hippocampal-VTA microcircuit was proposed. By means of population activity analysis, dis-inhibition effect of dopamine neurons in VTA caused by enhanced CA3 theta rhythm can illustrate theoretically why expectation of reward in rat ventral striatum at choice point is covert.

Keywords Hippocampal-VTA microcircuit • CA3 theta rhythm • Dis-inhibition effect • Reward of expectation

1 Introduction

The VTA (ventral tegmental area) is the origin of dopaminergic neurons, which project dopamine into numerous areas in the brain, and is widely implicated in reward system, motivation, cognition, and drug addiction. As one of the most important component of reward circuitry, hippocampal-VTA microcircuit plays a crucial role in reward-related behaviors. As well known, hippocampal theta rhythm is a prominent feature in extensive studies on hippocampus, and involves a strong correlation with learning, memory. In particular, convergent evidences from various experiments in neuroscience suggest that hippocampus could have great contribution to decision making [1]. Interestingly, recent rat experiments on solving

Y. Li
Research Center for Integrated Mathematics, Hokkaido University, Kita 12 Nishi 7, Kita-ku, Sapporo 060-0812, Japan

I. Tsuda (✉)
Research Institute for Electronic Science, Hokkaido University, Kita 12 Nishi 7, Kita-ku, Sapporo 060-0812, Japan
e-mail: tsuda@math.sci.hokudai.ac.jp

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_31

multiple T-maze tasks show some possible cues to unveil the relation between hippocampal theta rhythm and reward expectation.

First, remarkably enhanced CA3 theta rhythm was observed when rats show vicarious trial and error (VTE) behaviors at choice point in the initial (unfamiliar) instead of later (familiar) stage [3]. Second, covert expectation of reward was observed at ventral striatum when rats come at choice point in the initial (unfamiliar) stage, but disappeared in the later (familiar) stage [2]. Since CA3 theta rhythm and ventral striatum firing rate occur together, two questions arise: (1) Why the representation of reward expectation in ventral striatum is covert? (2) What is the relation between enhanced oscillation in CA3 and the covert representation in ventral striatum? Despite of great efforts, the underlying mechanism of covert representation remains unclear. Here, we proposed a computational model of Hippocampal-VTA microcircuits that can answer the following questions from theoretical viewpoint.

2 Models

A hippocampal-VTA microcircuit is shown in Fig. 1. As far, at least two distinct pathways from hippocampus to VTA have been described: one originating from ventral CA1/subiculum, which is considered as the provider of novelty signal, and another one from dorsal CA3, which implements a gate operation on VTA by means of dis-inhibition effect when dorsal CA3 received a stimulus with theta rhythm [4]. In the former pathway, a novelty signal is first sent to nucleus accumbens (NAcc), where the efferent projection of medium spiny neurons (MSNs) to VTA can be

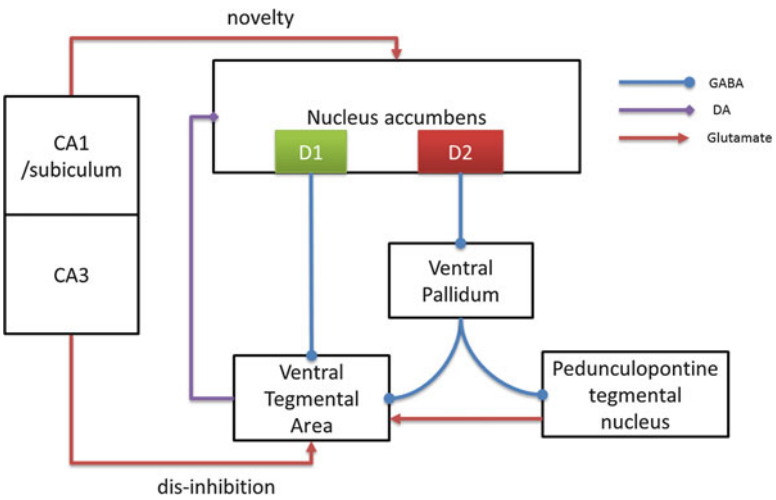


Fig. 1 A hippocampal-VTA model

divided into two distinct groups. One group of MSNs project to VTA directly, and rich type-1 dopaminergic receptors (D1Rs) are found in this direct pathway. In contrast, another group of MSNs first project to ventral pallidum (VP) whose GABAergic efferent neurons project to both VTA and pedunculopontine tegmental nucleus (PPN). VTA also received glutamatergic projection from PPN.

We model the circuit shown in Fig. 1 using mean-field population analysis. Normalized firing rate y_i of i-th population is defined by

$$y_i = f(x_i, \varepsilon_i) \tag{1}$$

$$\tau_i \dot{x}_i = -x_i + I_i \tag{2}$$

where x_i is neural activity of i-th population, and

$$I_i = \sum_{j \in G(i)} W_{i,j} y_j$$

is the total input to the population, and τ_i is time constant for activity changes in the population. A hyperbolic function is employed as the output function

$$y_i = f(x_i, \varepsilon_i) = \begin{cases} 0 & x_i < \varepsilon_i, \\ (\tanh(x_i - \varepsilon_i) + 1)/2 & x_i \geq \varepsilon_i, \end{cases} \tag{3}$$

where ε_i is the threshold of population firing rate, which can determine whether a population has a baseline activity. By virtue of the abovementioned method, we model five populations shown in Fig. 1: a population of MSNs with D1Rs that project from NAcc to VTA directly; a population of MSNs with both D1Rs and D2Rs, projecting from NAcc to VP; VP, PPN and VTA populations. The values of connection strength among five populations are chosen following the relevant anatomical features. The feedback dopamine modulation from VTA to NAcc can be described by a coefficient which depends on activity changes in VTA.

3 Results

In order to investigate why expectation of reward in ventral striatum is covert at choice point, we introduce a gate effect of hippocampal theta rhythm on VTA to the proposed population model. When enhanced hippocampal theta rhythm disinhibits VTA, a phasic increase occurs in VTA which responds to input from NAcc. Simultaneously, the feedback dopamine modulation enhanced the NAcc activity (Fig. 2a). In contrast, VTA is inhibited without hippocampal theta rhythm, VTA seldom responds to input from NAcc, so NAcc activity shows no enhanced modulation (Fig. 2b). Comparing Fig. 2a with Fig. 2b, small peak can be observed in NAcc activity when dopaminergic neurons in VTA are disinhibited. This could be

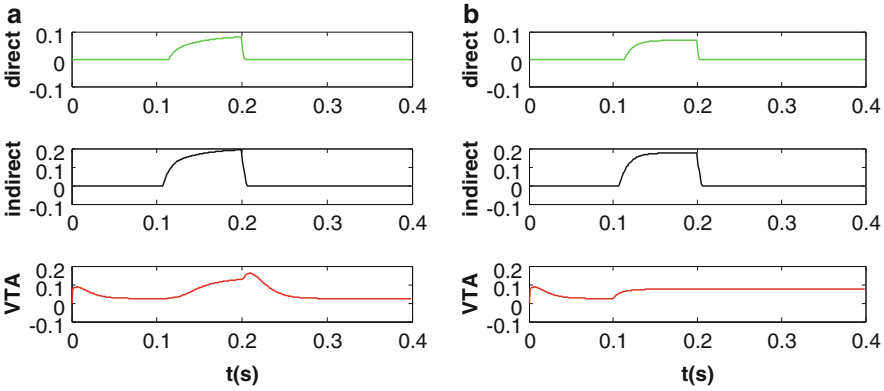


Fig. 2 Gate effects of hippocampal theta rhythm on VTA. (**a**) Dis-inhibition of VTA. (**b**) Inhibition of VTA

the reason why the representation of reward expectation in ventral striatum is covert. Importantly, it is likely that CA3 theta rhythm serves representation of expectation of reward in ventral striatum through gate effect on VTA.

4 Discussion

We proposed a computational model of hippocampal-VTA microcircuit, and the results suggest that modulation effect of dopamine from VTA to NAcc gets somewhat stronger when enhanced CA3 theta rhythm disinhibits dopaminergic neurons in VTA. Although the model is simple, the results have brought us some important implications. Particularly, it is likely to answer the questions abovementioned. Dopaminergic modulation on ventral striatum can get stronger if dopaminergic neurons in VTA are disinhibited. Despite VTA receives connections from numerous brain areas, the concurrence of enhanced CA3 theta rhythm and covert representation in ventral striatum supports the hypothesis that enhanced CA3 theta rhythm results in dis-inhibition of dopaminergic neurons in VTA. Modulator effect results in expectation of reward in NAcc is not so easy to observe, and seems covert.

5 Conclusions

Covert representation of expectation-of-reward is observed in rat ventral striatum at choice point during the earlier stage. However, the underlying neural mechanism remains obscure. Based on recent experimental reports from neuroscience, we proposed a computational model of hippocampal-VTA microcircuit, which can bring us some inspiring understanding on the problem.

- Enhanced CA3 theta rhythm could play an important role in disinhibiting the activities of dopamine neurons in VTA.
- Modulation effect of dopamine from VTA to NAcc is enhanced when hippocampal theta rhythm disinhibits VTA so that DA neurons in VTA fire in a phasic mode.
- Due to modulatory effect, the representation of expectation-of-reward in NAs seems covert.

Acknowledgments This work was partially supported by a Grant-in-Aid for Scientific Research on Innovative Areas (No. 4103) (21120002) from Ministry of Education, Culture, Sports, Science and Technology (MEXT) in Japan and partially supported by Human Frontier Science Program Organization (HFSP:RGP0039).

References

1. Wimmer, G.E., Shohamy, D. (2012). Preference by association: How memory mechanisms in the hippocampus bias decisions. *Science* 338, 270–273.
2. van der Meer, M. A., Kalensher, T., Lansink, C., Pennartz C., Berke J., Redish, D. (2010). "Integrating early results on ventral striatal gamma oscillations in the rat" *Frontiers in Neuroscience* 4 (28) 1–12.
3. Johnson, A., Redish, A. D. (2007). "Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point" *Journal of Neuroscience* 27 (45):12176–12189.
4. Luo A., Tahsili-Fahadan P., Wise R., Lupica C., Aston-Jone G. (2011). Linking context with reward: a functional circuit from hippocampal CA3 to ventral tegmental area, *Science* 333, 353–357.

A Computational Model of Cortical Pathways Formed with Electroencephalogram Synchronization

Naoyuki Sato

Abstract In this paper, a possible contribution of electroencephalogram (EEG) on spike transmission through multiple brain regions was evaluated by using computer simulations of a multi-layered network under the hypothesis that EEG and spikes are reciprocally activated. In the results, EEG activities across the multiple layers appeared to be globally synchronized, which was shown to produce a globally consistent pathway of spike transmission through the layers. This suggests that the EEG synchronization contributes to organize the activity transmission through multiple regions.

Keywords Decision making • Thermodynamics • Emotions • Cognition • Modular model

1 Introduction

Long-range electroencephalogram (EEG) synchronization is thought to act as a cue for interaction between distant regions [1]. Recent evidence [2, 3] shows a strong correlation between local field potential and individual neuronal activities, indicating that the long-range EEG synchronization may associate with spike transmission between distant brain regions. Hence, as a neural mechanism, EEG and spike activities can be hypothesized to be reciprocally enhanced. When EEG activities are synchronized in two regions, it modulates spike timing in each region, making the spike transmission more effective [4].

On the other hand, EEG can be induced by external signals, along with an increase of neuronal population activity. Here, a question arises concerning whether spike transmission in the network is still effective under the hypothesis of the EEG-spike reciprocal activation, where such conditions may result in the divergence of neuronal activity, or the locally dominated activation in a specific region.

N. Sato (✉)
Department of Complex and Intelligent Systems, School of Systems Information Science, Future University Hakodate, 116-2 Kamedanakano-cho, Hakodate-shi, Hokkaido 041-8655, Japan
e-mail: satonao@fun.ac.jp

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_32

In this paper, a possible contribution of EEG to the spike transmission through multiple regions was evaluated by using computer simulations of a multi-layered network under the hypothesis of the EEG-spike reciprocal activation. For this purpose, a neural unit consisting of a neural oscillator and a spiking neuron is proposed where the neural oscillator is associated with the neural mass model [5], and the spiking neuron is associated with a representative of the neural mass.

2 Model

A multi-layered network consisting of neural oscillators and spiking neurons was evaluated to show the relationship between EEG synchronization and spike transmission between distant regions. Figure 1 shows the basic construction of the model. Each of the N layers consists of M units. Each unit consists of an oscillatory ('O') module and a spiking ('S') module. The O-module models the population activity of neurons in the local cortical region and the S-module models a representative of neuronal activity in the population.

The S-module is described by a leaky integrate-and-fire model. It receives excitatory input from the S-modules in the neighboring layers, and from the global inhibition in the same layer. The S-module is also modulated by oscillatory activities of the O-module in the same unit. The activity of the i-th S-module in the k-th layer, v_{ik}, is given by,

$$\frac{dv_{ik}}{dt} = -C_L\left(v_{ik} - V_R\right) + C_{SE}\sum_l \sum_j w_{ikjl} D\left(t - t_{jl}^p, \tau_S\right) - C_{SI}\sum_j D\left(t - t_{jl}^p, \tau_S\right) - C_{OS}\cos\theta_{ik} + I_{app} + \varepsilon_S$$

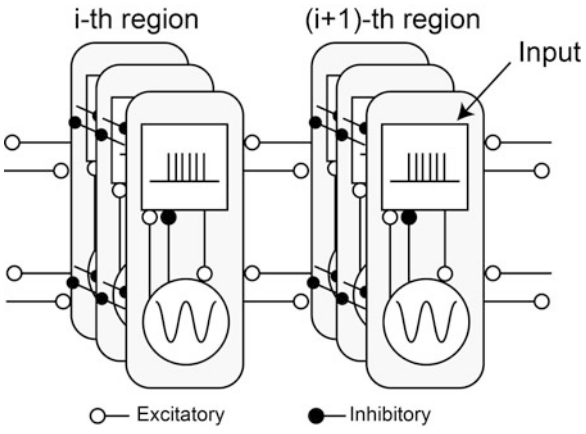


Fig. 1 Basic structure of the model. The model consists of multiple layers, of which the unit is given as spiking module (S-module) and oscillatory module (O-module). The unit receives inhibitory input from the units within the layer and excitatory input from the units in the neighbor layers

with

$$D\left(t',\tau\right)=\begin{cases}1\ (0\le t'\le\tau)\\ 0\ \ (\text{otherwise})\end{cases}$$

where $D(t,\tau)$ denotes a connection delay of a time constant τ, w_{ikjl} denotes synaptic weight from the l-th unit in the j-th layer (1 represents the existence of connection and 0 represents no connection), t^{p} denotes the time of the P-th spike, $\cos\theta_{ik}$ denotes the modulation from the O-module in the same unit, I_{app} denotes the input current, ε_s denote perturbation, and the capital letters are constants.

The O-module models neural oscillations of EEG described by a phase oscillator [6]. In contrast to the neural mass model [5] in which variables are given by the sum excitatory and inhibitory activities in a local region, the phase model does not directly associate with physiological parameters, but it is nonetheless able to sufficiently describe the temporal dynamic of neural oscillator [7]. The O-module is activated by the S-module in the same unit and modulates its spike timing. The O-module also receives excitatory input from the O-modules in the neighboring layers and from the global inhibition within the same layer. The phase of the i-th O-module in the k-th layer, θ_{ik}, is given by,

$$\frac{1}{\omega}\frac{d\theta_{ik}}{dt}=1-\left(C_{OE}\sum_{l}\sum_{k}w_{ikjl}\cos\theta_{jl}-C_{OI}\sum_{i'\neq i}\cos\theta_{i'k}+C_{SO}D\left(t_{ik}^{p},\tau_{O}\right)+\varepsilon_{O}-C_{\beta}\right)\sin\theta_{ik}$$

where ε_O denotes perturbation, and capital letters are constants.

The spike transmission through multiple layers in the network is evaluated by the uniqueness of the activity propagation pathway organized in the network. To measure the uniqueness of the pathway, a pathway formation index, P, is defined as follows,

$$P=\prod_{(i,k)\in Pathway}\frac{\text{Spike count of }(k,i)\text{ unit}}{\sum_{k'}\text{Spike count of }\left(k',i\right)\text{ unit}}.$$

where P ranges from 0 to 1 and it becomes large in the case of a continuous pathway without branches and segmentation.

Parameters of the model are given as follows; $C_L=50$, $V_R=-65$, $C_{SE}=4800$, $\tau_S=0.01$, $C_{SI}=3000$, $C_{OS}=4000$, $\omega=10\pi$, $C_{OE}=0.15$, $C_{OI}=0.2$, $C_{SO}=1$, $C_\beta=1.2$. These parameters were chosen to balance the excitatory input and the global inhibitions in each module. The frequencies of S- and O-modules' activities are given around 40 and 10 Hz, respectively.

3 Results

The Units' Activities in a Small Circuit Figure 2a shows the temporal evolution of the S- and the O-modules' activities in a single unit. The application of the input current to the S-unit resulted in the generation of spiking activity. This activity evoked neural oscillation in the O-module. Then, the spiking activity became bursting by receiving the modulation from the O-module. When the input current was turned off, both activities went back to resting states.

In a network of two units reciprocally inhibited (Fig. 2b), the oscillations in each unit quickly became anti-phase. When the third unit excitatory connected to each of the two units were added to the network (Fig. 2c), one of the two units became active and the other unit became inhibited. It is thought that the third unit stochastically chooses one of the two units and reciprocally activates the chosen unit that strongly inhibits the other unit. As illustrated in Fig. 2c, unit #2 was active, but it was found to be stochastic in repeated simulations, i.e., unit #1 was active at a percentage of around 50 %.

In the case of a network of four units forming two layers (Fig. 2d), there are four possible pathways (from #1 to #3, from #1 to #4, from #2 to #3, or from #3 to #4). In Fig. 2d, the activity propagation from units #1 to #3 appeared, while the appearance of the pathway was found to be stochastic in repeat simulations. These properties of the activity propagation are thought to also appear in the larger network.

A Network Leading to Ambiguous Pathways Figure 3 shows the results of temporal evolution of each module in a five-layer network where the units in the neighboring layers were connected all-to-all (Fig. 3a). This network was used to evaluate the relationship between activity propagation and the EEG synchronization. The input was given to two units of #1 and #15 where the former simulates a sensory input and the latter simulates a memory-dependent activity.

Temporal evolution of each module is shown in Fig. 3b. At the beginning, many units are simultaneously activated, but five units in the five layers (units #1, #5, #8, #11, and #15) remained active after a few oscillation cycles. All of them were found to be physically connected, including the initially activated units (#1 and #15). This indicates that the activity propagation pathway links the two initially activated units. In repeat simulations, the appearance of the route from unit #1 to #15 was found to be stochastic.

Cross-correlations between activities of unit #1 and #15 (Fig. 3c, d) showed that the activities of the S- and O-modules in the two units were correlated without delay. Such correlation is thought to be associated with the existence of the activity propagation pathway. In this result, the synchronization at 10 Hz was found to be dominant in both the S- and the O-module, and the synchronization in the higher frequency band was not clear.

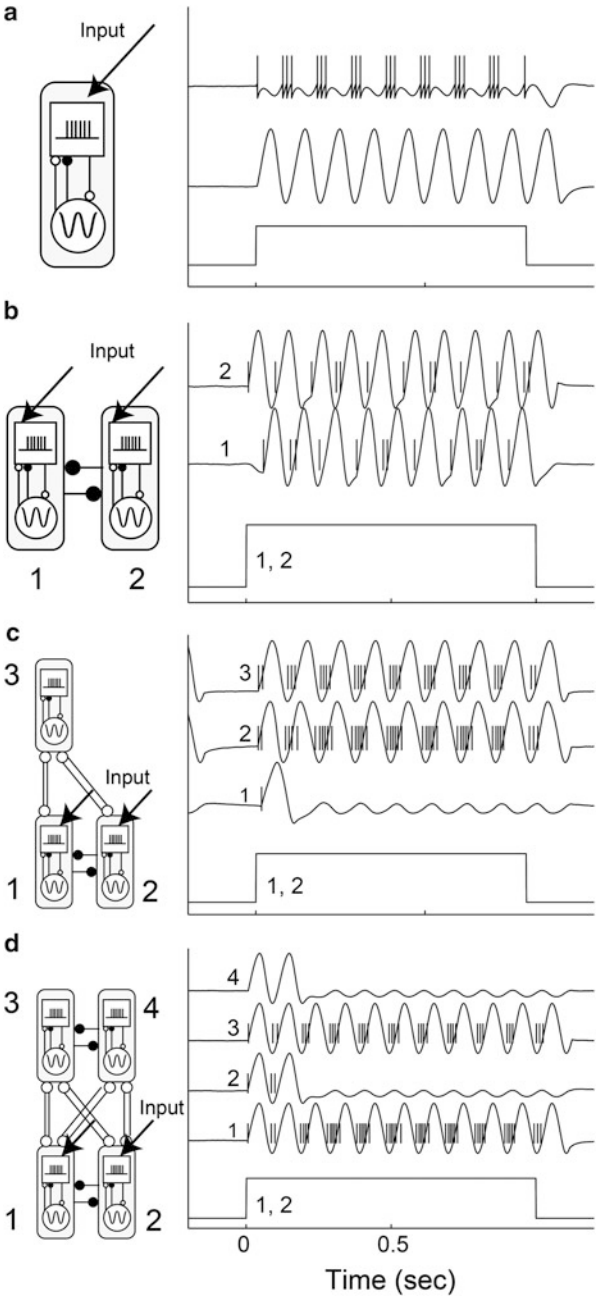


Fig. 2 Temporal evolution of each module's activity in various small circuits. (**a**) A single unit, (**b**) a network of two units reciprocally inhibited, (**c**) a network of three units in which the two units are reciprocally inhibited and the other unit reciprocally activates both of them, and (**d**) a network of four units forming two layers. In (**b**)–(**d**), activities of S- and O-modules are shown in the same raw format

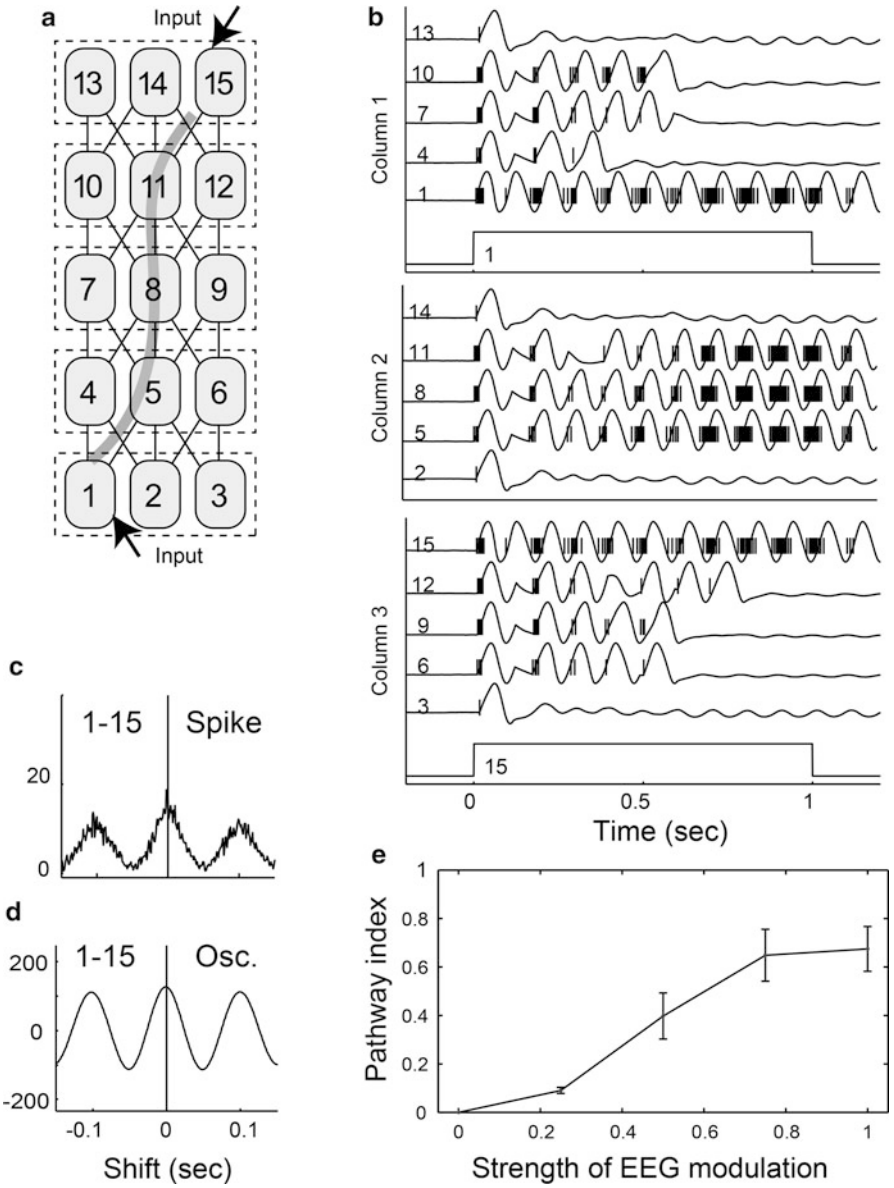


Fig. 3 Results of the temporal evolution of units' activities in the five layer network that can lead ambiguous pathway from units #1 to #15. (**a**) Network structure. A resultant pathway is shown by a *gray line*. Each numbered oval indicates units. Input was given to units #1 and #15. (**b**) Temporal evolution of each module's activity. (**c**) Cross-correlation of spikes between units #1 and #15. (**d**) Cross-correlation of oscillations. (**e**) Pathway index as a function of the strength of EEG modulation from the O-module to S-modules (N = 10)

Moreover, the stronger modulation from the O-module to the S-module was important for organizing the unique pathway (Fig. 3e). It suggests that the interaction between neural oscillation and spike activity is important for organizing activity propagation through multiple regions.

A Network Including Competing Pathways In the above result, the competition of the units within each layer may play a dominant role in the organization of the unique pathway. In this section, to evaluate the competition between pathways, a network including three competing pathways was used (Fig. 4a) where two units (#1 and #15) in different pathways were initially activated. In the results (Fig. 4b), the units in the pathway of the left column were activated, while the units in the pathway of the right column became resting. In repeated simulations, it was found that the activation of the pathways of the left and right column were stochastic.

The competition between pathways was found to depend on the modulation from the O- to the S-module. The stronger modulation was found to be important to increase the uniqueness of the pathway (Fig. 4c). Moreover, the lower frequency of the O-module's activity was found to be essential to produce the unique pathway (Fig. 4d). This result suggests that the difference of time scales between the O- and the S-module is important for organizing the global pathway.

4 Discussions

In this paper, under the hypothesis that the EEG and spike activities were reciprocally activated, a multi-layered network consisting of spiking neurons and neural oscillators was evaluated to show the possible contribution of EEG to the spike transmission. In the results, the spike-induced EEG was associated with the formation of the global activation pathway. It suggests that EEG organizes the spike transmission through multiple brain regions.

As shown in Figs. 3 and 4, all units connected to the initially activated units were initially activated, while the limited number of units forming a global pathway was only activated after a few oscillation cycles. Here, the mechanism of the reorganization of segmented local pathways is thought to be essential to produce the global pathway. A speculated mechanism is illustrated in Fig. 5 where the three units' network, identical to Fig. 2b, is used for explanation purposes. At the beginning, unit #2 activates unit #3 and unit #1 is inhibited. When an occasional spike appears in unit #1, it evokes oscillation in unit #1 and that shifts the phase of oscillation in unit #2. After that, the activity pathway can be stochastically reorganized to be #1 and #3. This scenario proposes that the activity pathway is a kind of 'meta-stable' structure that can be easily reorganized according to the spatiotemporal context of the unit. This mechanism is similar to a dynamic linking model proposed by the authors [8].

In the current study, the formation of the unique pathway is emphasized, and the model can produce multiple pathways simultaneously in different phases of EEG.

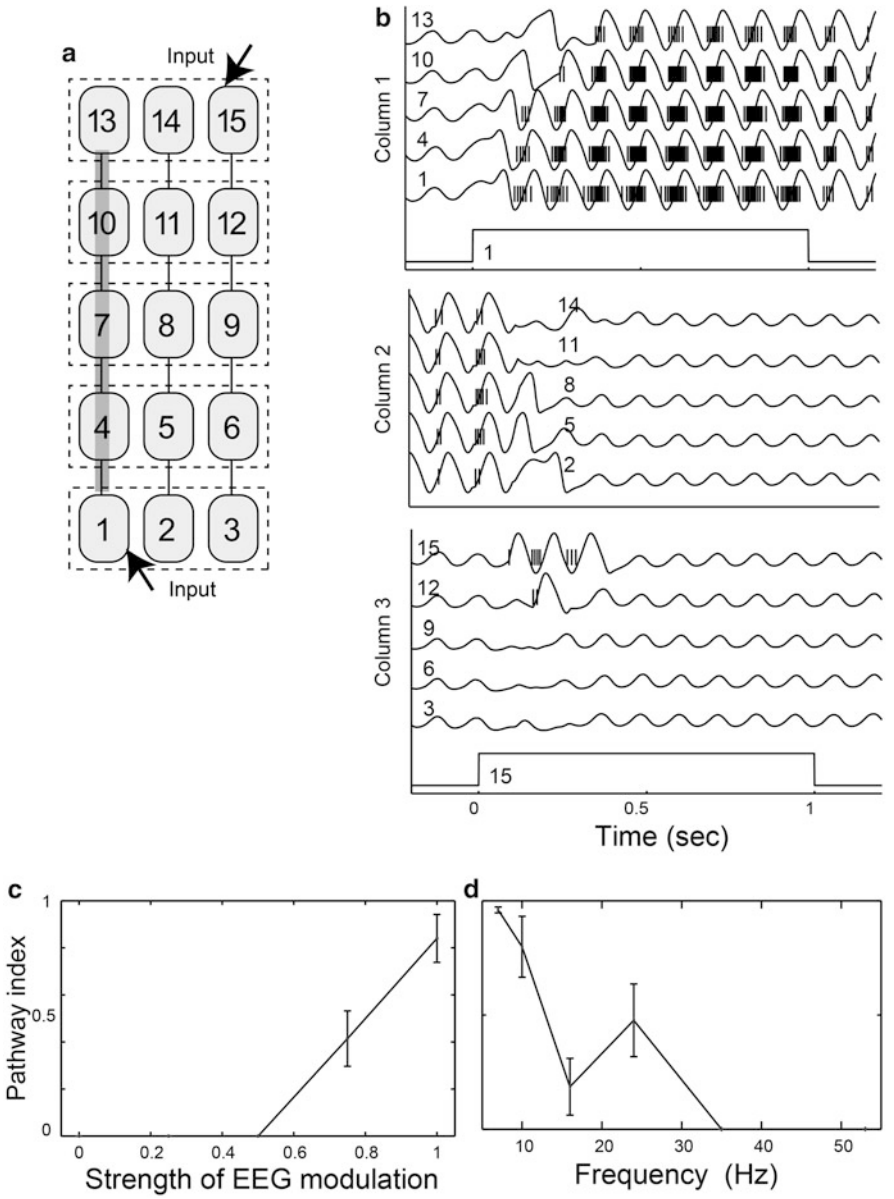


Fig. 4 Results of the temporal evolution of units' activities in the five layer network with three competing pathways. (**a**) Network structure. (**b**) Temporal evolution of each module's activity. (**c**) Pathway index as a function of the strength of EEG modulation from the O- to the S-modules (N = 10). (**d**) Pathway index as a function of the frequency of the O-module's activity (N = 10)

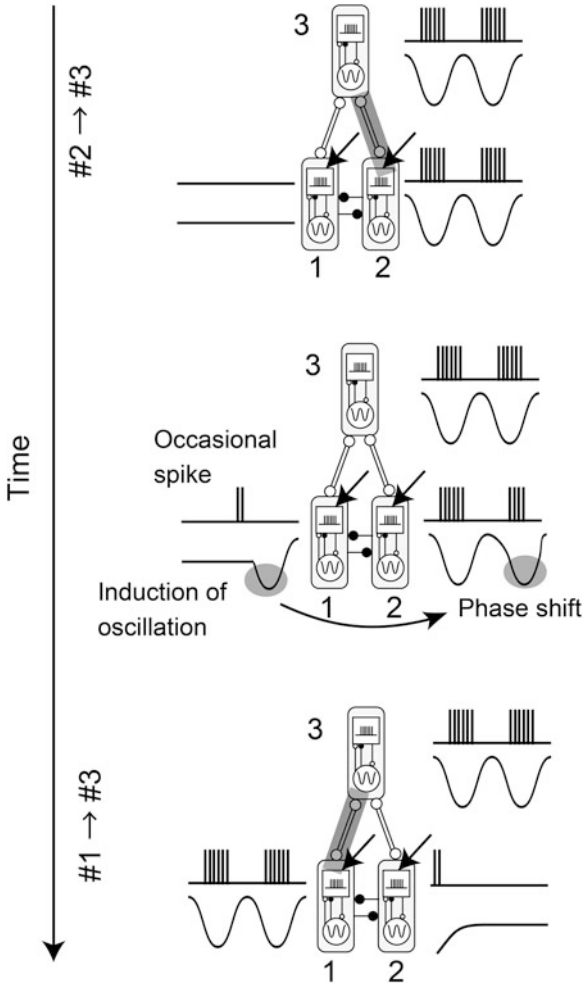


Fig. 5 A speculated mechanism of the reorganization of the activity propagation pathway

This can be implemented in the case of weaker global inhibition of the S-module (analogous to Fig. 2b). Such phase-dependent processing has been discussed in both experiment [9] and theory [10]. It may be important to evaluate the ability of the model in formation of multiple pathways in future analysis.

Oscillatory modulation was proposed to modulate spike timing of neurons, which contributes to synaptic plasticity and spike transmission [4]. In line with this effect, the current study highlights the contribution of EEG to the organization of the global pathway along multiple regions. Since the current model agrees with recent evidence of cross-frequency coupling [11], it may be important to experimentally evaluate the model in terms of EEG-spike correlation and the activity transmission pathway through distant brain regions.

Acknowledgments This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas “Neural creativity for communication (No.4103)” (24120714) of MEXT, Japan.

References

1. Varela, F., Lachaux, J.P., Rodriguez, E., Martinerie, J.: The brainweb: phase synchronization and large-scale integration. Nat Rev Neurosci 2, 229–239 (2001).
2. Nir, Y., Fisch, L., Mukamel, R., Gelbard-Sagiv, H., Arieli, A., Fried, I., Malach, R.: Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations. Curr Biol 17, 1275–1285 (2007).
3. Whittingstall, K., Logothetis, N.K.: Frequency-band coupling in surface EEG reflects spiking activity in monkey visual cortex. Neuron 64, 281–289 (2009).
4. Fries, P.: A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn Sci 9, 474–480 (2005).
5. Jansen, B.H., Rit, V.G.: Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. Biol Cybern 73, 357–366 (1995).
6. Sato, N.: Modulation of cortico-hippocampal EEG synchronization with visual flicker: A theoretical study. In: Omori, T., Yamaguchi, Y., Sakaguchi, Y., Sato, N., Tsuda, I. (eds.) Advances in Cognitive Neurodynamics (III). Springer (2013).
7. Hoppensteadt, F.C.: An introduction to the mathematics of neurons: modeling in the frequency domain. Cambridge University Press (1997).
8. Sato, N., Yano, M.: A model of binocular stereopsis including a global consistency constraint. Biol Cybern 82, 357–371 (2000).
9. Rizzuto, D.S., Madsen, J.R., Bromfield, E.B., Schulze-Bonhage, A., Kahana, M.J.: Human neocortical oscillations exhibit theta phase differences between encoding and retrieval. Neuroimage 31, 1352–1358 (2006).
10. Hasselmo, M.E., Bodelon, C., Wyble, B.P.: A proposed function for hippocampal theta rhythm: separate phases of encoding and retrieval enhance reversal of prior learning. Neural Comput 14, 793–817 (2002).
11. Canolty, R.T., Knight, R.T.: The functional role of cross-frequency coupling. Trends Cogn Sci 14, 506–515 (2010).

Neurodynamics of Up and Down Transitions in a Network Model

Xuying Xu and Rubin Wang

Abstract At the network level, populations of neurons have been observed to make upward and downward transitions synchronously. This paper focuses on the neurodynamics of these transitions in a network model which consists of 25 neurons in the network. We study the spontaneous transitions between up and down states and the ones induced by external inputs. We also compare these results with the results of a single neuron model, trying to find out the function of the cortex and the mechanism of how our brains work.

Keywords Up-down transition • Bistability • Spontaneous activity • Ion channel model

1 Introduction

Different behavioral states of an animal are characterized by distinct patterns of the global brain activity [1]. Both in vivo and in vitro recordings indicate that neuronal membrane potentials can make spontaneous transitions between distinct up and down states. Neural electrophysiology experiments show that during slow-wave sleep in the primary visual cortex of anesthetized animals [2–4] and during quiet wakefulness in the somatosensory cortex of unanesthetized animals [5, 6], the membrane potentials make spontaneous transitions between two different levels, called up and down states [7]. Transitions between up and down states can also be evoked by sensory stimulation [2, 5]. Recent findings also show that activation of a single cortical neuron can significantly modulate sensory and motor outputs [8, 9]. Furthermore, repetitive high-frequency burst spiking of a single rat cortical neuron could trigger a switch between the cortical states resembling slow-wave and rapid-eye-movement sleep [1]. This is reflected in the switching of membrane potential of the stimulated neuron from high frequency and low amplitude oscillations to low frequency and high amplitude ones, or vice versa. At the same time, cortical local

X. Xu (✉) • R. Wang
Institute for Cognitive Neurodynamics, East China University of Science and Technology, Shanghai, People's Republic of China
e-mail: xu_xuying@126.com; rbwang@163.com

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_33

field potentials (LFP) change over time. Here we use the local field potential (LFP) to describe the state of the whole cortex [10–13]. Therefore, the up and down states of a single neuron reflect distinct global cortical states, which resemble slow-wave and rapid-eye-movement sleep, respectively [14–16].

2 Models

The neural network model based on the single neuron model is described by equations as follows. This dynamic model [17] consists of the following currents: an instantaneous, inward current (sodium current), a slow h-like current, an outward current (a potassium current and a leak current) and the currents between neurons. According to the numerical simulation of this model, we observe the results that have been observed in electrophysiology experiments. The current equation for the model is:

$$C\frac{dV_i}{dt} = -\left(I_{Na}^i + I_h^i + I_K^i + I_l^i + I_{syn}^i\right)$$

The ionic currents are as follows:

$$I_{Na}^i = g_{Na} m_\infty \left(V_i - V_{Na}\right)$$

$$I_h^i = g_h h_i \left(V_i - V_h\right)$$

$$I_K^i = g_K b_i \left(V_i - V_K\right)$$

$$I_l^i = g_l \left(V_i - V_l\right)$$

The current between neurons is:

$$I_{syn}^i = \sum_{j \neq i} g_{ji} \left(V_i - V_j\right)$$

3 Results

3.1 Bistability

When we study on the single neuron model, we found that transitions between up and down states can be induced by two different kinds of stimulus. One is to

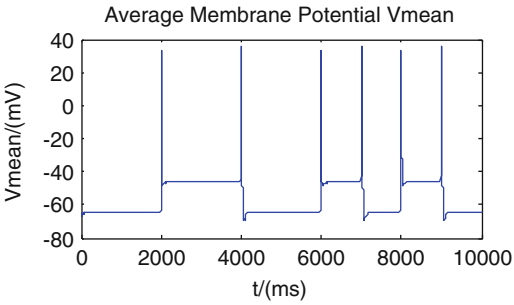


Fig. 1 In the period of 10 s, we add a pulse current which lasts 0.1 s every 2 s, with the current intensity 7.2 μA/cm^2. Average membrane potential switches between the up and down states

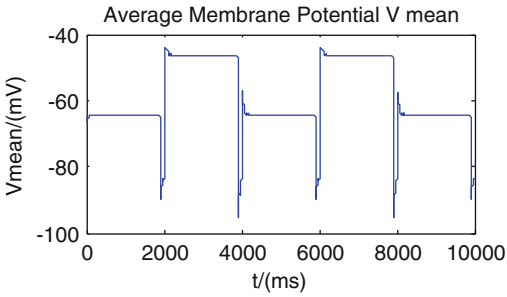


Fig. 2 In the period of 10 s, we add a stimulation which lasts 4 ms every 2 or 1 s, leading to the intensity of sodium conductance changing from 0.06 to 1.2 mS/cm^2 instantaneously. Average membrane potential switches between the up and down states

add brief outward current pulses, another is to improve the sodium conductance to a certain value instantaneously. Now, we research on the neural network in the same way to try to find out is there exist the similar phenomenon agrees with electrophysiology experiment results. Results are showed in Figs. 1 and 2.

So from the above two results, we find that the average membrane potential switches between the up state (about -45 mV) and the down state (about -65 mV). So, this dynamic model can describe the bistability of up and down transitions of neural networks. And the transitions can be modulated by external stimulations and intrinsic nature of conductance of sodium.

3.2 Directivity

We find that the changing of sodium conductance can modulate the directivity of the transitions. Figures 3, 4 and 5 describe different transition mode adjusted by different values of sodium conductance. The tops of the figures are average membrane potential V of the neural network, while the bottoms are phase plane for the mean of two kinds of dynamic variables h and V, denoted by V_{mean}, h_{mean}. The red solid line shows all the points that $\dot{h} = 0$, the blue dot line shows all the points that $\dot{V} = 0$, and the intersection of these two lines are stable point of the system. The green solid line in the figure presents the transit process from one stable point to another.

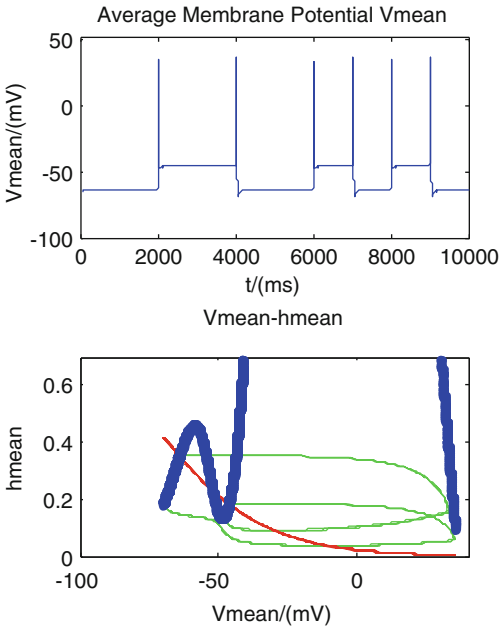


Fig. 3 Average membrane potential can transit from down to up state or from up to down state by adding a stimulation that increase sodium conductance to $g_{Na} = 1.2\,mS/cm^2$ instantaneously. The h-V phase plane further shows that the system transmits between the two stable states

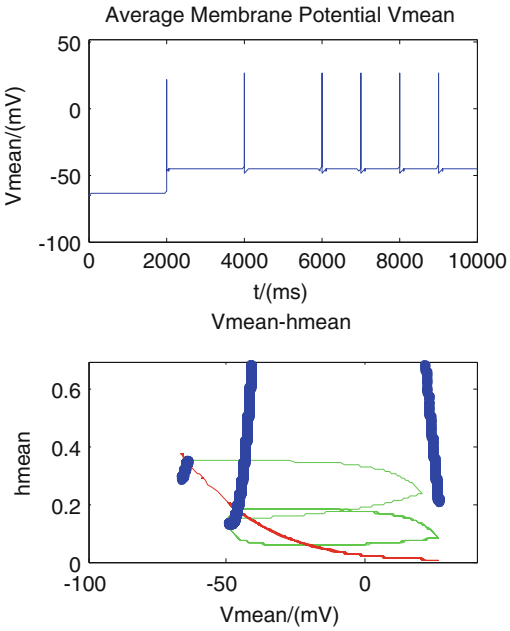


Fig. 4 If the level of sodium conductance is insufficiently activated to $g_{Na} = 0.8mS/cm^2$, the average membrane potential can only transit from the down to up state. The h-V phase plane also shows the same results

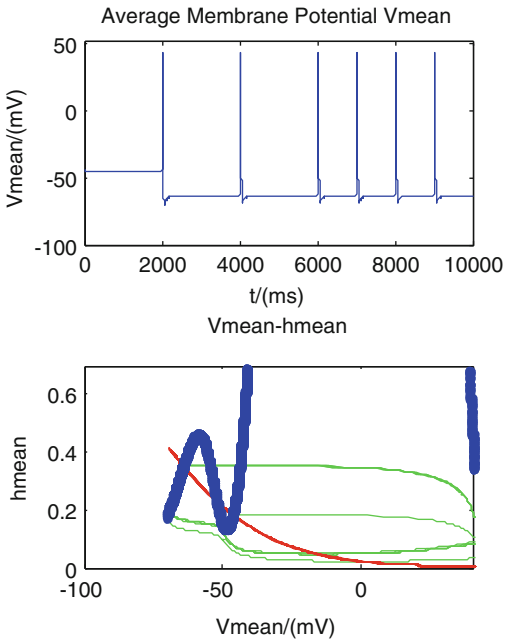


Fig. 5 If the level of sodium conductance is excessively activated to $g_{Na} = 2mS/cm^2$, the average membrane potential can only transit from the up state to the down. The h-V phase plane also presents the same results

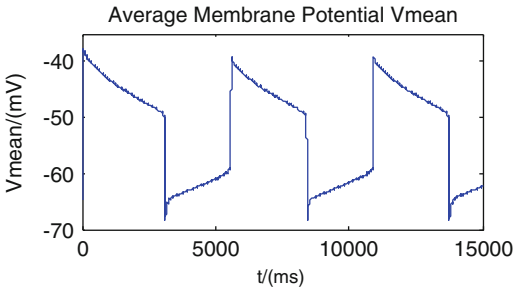


Fig. 6 Average membrane potential transit between two stable states spontaneously and periodically without external stimuli

So, the dynamic model can describe the bidirectional or unidirectional characteristic of up and down transitions of neural networks. And these transitions are controlled by intrinsic nature of conductance of sodium.

3.3 Spontaneity

In this section, we introduce the dynamic variable b, the inactivation rate of potassium conductance of each neuron, to study the spontaneous transitions of neural networks. The calculated results are shown in Figs. 6 and 7.

The dynamic model with a slowly activating potassium current can describe periodic spontaneous transitions between the up and down states in the absence of synaptic input, suggesting that they are triggered by intrinsic processes and not by external input.

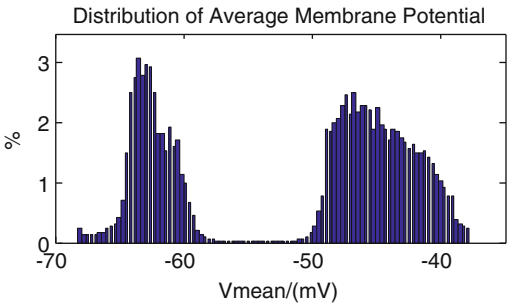


Fig. 7 The distribution of the average membrane potential, a two-peak distribution, indicating the two stable state of up and down transitions of membrane potential

4 Conclusions

This dynamic model can describe the bistability of up and down transitions of networks modulated by external stimulations and conductance of sodium and the bidirectional or unidirectional characteristic of up and down network transitions controlled by conductance of sodium, also the periodic spontaneous transitions of networks between the up and down states in the absence of synaptic input.

Acknowledgments This work is supported by the National Natural Science Foundation of China (No. 11232005).

References

1. Cheng-yu T. Li, Mu-ming Poo, Yang Dan, Burst Spiking of a Single Cortical Neuron Modifies Global Brain State. Science. 324, 643646 (2009).
2. Anderson, J., Lampl, I., Reichova, I., Carandini, M., and Ferster, D., Stimulus dependence of two-state fluctuations of membrane potential in cat visual cortex. Nat. Neurosci. 3, 617–621 (2000).
3. Lampl, I., Reichova, I., and Ferster, D., Synchronous Membrane Potential fluctuations in Neurons of the Cat Visual Cortex, Neuron. 22, 361–374(1999).
4. Steriade, M., Nunez, A., and Amzica, F., Intracellular analysis of relations between the slow (<1Hz) neocortical oscillation and other sleep rhythms of the electroencephalogram. J. Neurosci. 13, 3266–3283(1993).
5. Petersen, C. C. H., Hahn, T. T. G., Mehta, M., Grin-vald, A., and Sakmann, B., Interaction of sensory responses with spontaneous depolarization in layer 2/3 barrel cortex. PNAS 100, 13638–13643(2003).
6. Nestor Parga and Larry F. Abbott, Network model of spontaneous activity exhibiting synchronous transitions between up and down states, Neurosci. 1, 57-66 (2007).
7. Timofeev, I., Grenier, F., Steriade, M., Disfacilitation and active inhibition in the neocortex during the natural sleep-wake cycle: An intracellular study. PNAS 98, 1924–1929(2001).
8. M. Brecht, M. Schneider, B. Sakmann, T. W. Margrie, Hisker movements evoked by stimulation of single pyramidal cells in rat motor cortex, Nature 427, 704 (2004).
9. A. R. Houweling, M. Brecht, Behavioural report of single neuron stimulation in somatosensory cortex, Nature 451, 65 (2008).

10. Rubin Wang, Zhikang Zhang, Guanrong Chen, Energy coding and energy functions for local activities of brain. Neurocomputing. 73(1-3), 139-150 (2009)
11. Rubin Wang, Zhikng Zhang, Guanrong Chen, Energy function and energy evolution on neural population. IEEE Transactions on Neural Networks. Vol. 19, No. 3, 535-538 (2008)
12. Rubin Wang, Zhikang Zhang, Energy coding in biological neural network, Cognitive Neurodynamics, Vol. 1, No. 3, 203-212 (2007)
13. Yan Liu, Rubin Wang, Zhikang Zhang, Xianfa Jiao. Analysis on Stability of Neural Network in the Presence of Inhibitory Neurons. Congnitive Neurodynamics. Vol. 4, No. 1, 61-68 (2010)
14. M. Steriade, D. McCormick, T. Sejnowski, Thalamocortical Oscillations in the Sleeping and Aroused Brain, Science 262, 679 (1993).
15. A. Destexhe, D. Contreras, M. Steriade, Spatiotemporal analysis of local field potentials and unit discharges in cat cerebral cortex during natural wake and sleep states, J. Neurosci. 19, 4595(1999).
16. D. Gervasoni et al., Global Forebrain Dynamics Predict Rat Behavioral States and Their Transitions, J. Neurosci. 24, 11137 (2004).
17. Yonatan Loewenstein, Séverine Mahon, Paul Chadderton, Kazuo Kitamura, Haim Sompolinsky, Yosef Yarom & Michael Häusser, Bistability of cerebellar Purkinje cells modulated by sensory stimulation. Neuroscience. 8, 202-211 (2005).

Part VI
Oscillations, Synchronization and Synaptic Plasticity

STDP Produces Well Behaved Oscillations and Synchrony

David Bhowmik and Murray Shanahan

Abstract It has been demonstrated that, in a network of excitatory and inhibitory neurons, a synchronous response gradually emerges due to spike timing dependant plasticity acting upon an external spatio-temporal stimulus that is repeatedly applied. This paper builds on these findings by addressing two questions relating to STDP and network dynamics. Firstly, how does the choice of neuron model affect the learning of oscillation through STDP? Our experiments suggest that the earlier results hinge on the selection of a simple, biologically less realistic neuron model. Secondly, how do neural oscillators that have learned to oscillate only in response to a particular stimulus behave when connected to other such neural oscillators?

Keywords Spike timing dependent plasticity • Hebbian learning • Hodgkin-Huxley neurons • Integrate-and-fire neurons • Synchronisation

1 Introduction

Spike Timing Dependent Plasticity (STDP) is a refinement of the Hebbian learning principle for spiking neural networks, and has been reported in many experimental studies [1]. STDP has further been studied in relation to oscillations. Hosaka et al. [2] demonstrate oscillatory dynamics in a network of excitatory and inhibitory neurons that has been trained using STDP with an external spatio-temporal stimulus that was repeatedly applied. They found that a synchronous response gradually emerges, and the synchrony becomes sharp as learning proceeds. The authors state that the generation of synchrony itself does not depend on the length of the cycle of external input, however they found that synchrony emerges once per cycle of the length of the external stimulus trained upon.

D. Bhowmik (✉) • M. Shanahan
Department of Computing, Imperial College London, 180 Queen's Gate, London SW7 2AZ, UK
e-mail: d.bhowmik10@imperial.ac.uk

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_34

This paper addresses two issues relating to STDP and network dynamics. Firstly, how does the choice of neuron model affect the learning of oscillation through STDP? Secondly, how do neural oscillators that have learned to only oscillate in response to a particular stimulus behave when connected to other neural oscillators?

2 Methods

2.1 Quadratic Integrate-and-Fire Neurons

The Quadratic Integrate and Fire (QIF) model [3] displays Type I neuron dynamics [4] with a saddle node bifurcation. The time evolution of the neuron membrane potential is given by:

$$\frac{dV}{dt} = \frac{1}{\tau}(V - V_r)(V - V_t)\frac{I}{C}$$

where V is the membrane potential, with V_r and V_t being the resting and threshold values respectively. C is the capacitance of the cell membrane. τ is the membrane time constant such that $\tau = RC$ with R being the resistance. I represents a depolarizing input current to the neuron.

2.2 Izhikevich Neurons

The Izhikevich (IZ) neuron model [5] is a two variable system that can model both Type I and Type II neurons depending upon how it is parameterized. The time evolution of the model is defined as follows:

$$\frac{dV}{dt} = 0.04V^2 + 5v + 140 - U + I$$
$$\frac{dU}{dt} = a(bV - U)$$

if $V > 30$, then $\{V \leftarrow c, U \leftarrow U + d$

I is the input to the neuron. V and U are the voltage and recovery variable respectively, and a, b, c and d are dimensionless parameters. The chosen parameter values dictate that the Izhikevich neurons used in this paper are Type II neurons with a saddle node bifurcation.

2.3 Hodgkin-Huxley Neurons

The Hodgkin-Huxley (HH) model [6] is a Type II neuron with an Andronov-Hopf bifurcation. Hodgkin and Huxley found three different types of ion current: sodium (Na+), potassium (K+), and a leak current that consists mainly of chloride (Cl-) ions. From their experiments, Hodgkin and Huxley formulated the following equation defining the time evolution of the model:

$$C\frac{dV}{dt} = g_K n^4 (V - E_K) - g_{Na} m^3 h (V - E_{Na}) - g_L (u - E_L)$$

C is the capacitance and n, m and h describe the voltage dependence opening and closing dynamics of the ion channels. The standard parameterisation and rate functions for each chemical and channel are used and can be found in Hodgkin and Huxley's book [6].

2.4 Synaptic Model

A conductance synaptic model is used for experiments using the QIF and IZ models model, whereas the HH model uses synaptic reversal potentials to further scale incoming spikes. The latter model is as follows:

$$I_j(t) = \sum\nolimits_i w_{ij} t_i \left(Rev - V_j\right)$$

where $I_j(t)$ is the input to neuron j at time t, t_i is the spike from neuron i arriving at time t, and w_{ij} is the weight of the synapse connecting the two neurons. Rev is the reversal potential and V_j is the voltage of the target neuron.

2.5 Spike Timing Dependent Plasticity

The STDP update method used in this paper is an 'additive nearest neighbour' scheme. A pre-synaptic spike followed by a post-synaptic spike potentiates the synaptic weight, where as a post-synaptic spike followed by a pre-synaptic spike depresses the weight. The change in weight (Δw) is affected by the exponential of the time difference (Δt) and the learning rate constant (λ):

$$\Delta w = \lambda e \frac{-|\Delta t|}{\tau}$$

For potentiation, the learning rate value λ is 0.3, and the window τ is 20 ms. For depression, the learning rate value λ is 0.3105 and the window τ is 10 ms.

2.6 Evolution of Oscillatory Nodes

The neural architecture for generating oscillations used in this paper is pyramidal inter-neuronal gamma (PING).

Whilst the general PING architecture is well understood, the specific details required for both particular oscillatory frequencies and neuron model varies and involves a large space of parameter values within the general PING framework. In order to obtain these values we used a genetic algorithm. In the present work, all neural populations used an excitatory layer of 200 neurons and an inhibitory layer of 50 neurons. The excitatory layer drives the entire network and so is the only one to receive external input. The networks were wired up with connections between excitatory neurons, between inhibitory neurons, from excitatory to inhibitory neurons, and from inhibitory to excitatory neurons.

The parameters that were evolved were the length in milliseconds of the external stimulus presentation, the synaptic weights and delays, as well as the number of synaptic connections between source and target neurons in each pathway. The amount of time trained for was also an evolved parameter for networks that learnt. Two types of PING architecture networks were investigated. The first learnt a stimulus and then after learning would only oscillate to the learnt stimulus. The second did not use learning and so would oscillate to any input stimuli.

2.7 Synchronisation Metric

We only calculated synchrony amongst the excitatory neuron layers. The spikes of each neuron in each excitatory layer were binned over time, and then a Gaussian smoothing filter was passed over the binned data to produce a continuous time varying signal. Following this, we performed a Hilbert transform on the mean-centred filtered signal in order to identify its phase. The synchrony at time t was then calculated as follows:

$$\varphi = \frac{1}{t_{\max}} \sum_{t}^{t_{\max}} \left| \frac{1}{N} \sum_{j}^{N} \mathrm{e}^{\theta_j(t) i} \right|$$

where $\theta_j(t)$ is the phase at time t of oscillatory population j. i is the square root of -1. N is the number of oscillators, and t_{max} is the length of time of the simulation.

3 Results

3.1 Neuron Model and the Learning of Oscillation

Our first investigation explored how the neuron model affects the ability of a cluster of neurons to learn to oscillate. In order to explore this we evolved neural learning

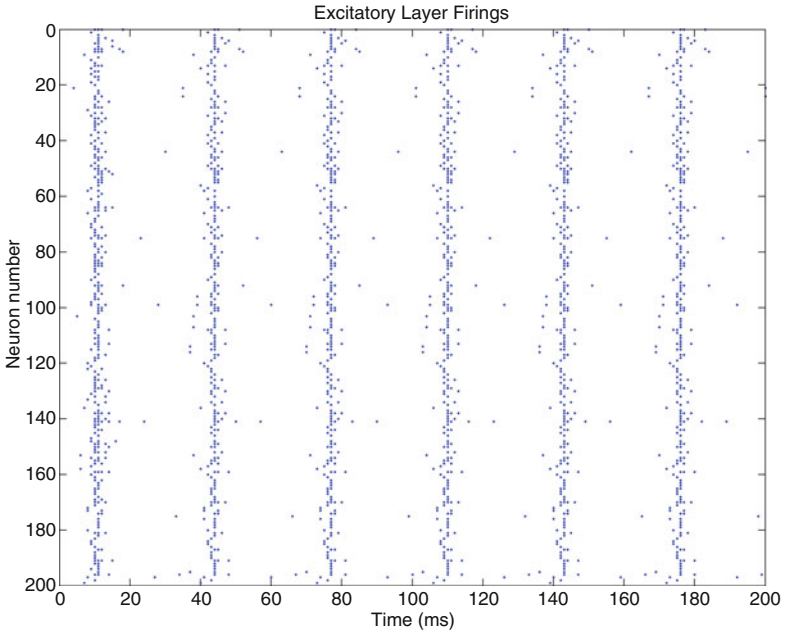


Fig. 1 Raster plot of neuron firings from the excitatory layer of a QIF PING node that has learnt to oscillate at 30 Hz

PING oscillators to oscillate at 30 Hz for QIF, IZ and HH neuron models. Figure 1 shows a raster plot of the firings of the excitatory layer from the evolved QIF solution when it has been presented with a learnt stimulus after training. In accord with the finding of Hosaka et al. [2], the network fires regularly at the stimulus presentation, and has narrow and pronounced periodic bands. These thin bands appear approximately every 33 ms giving the 30 Hz oscillation desired.

Figure 2 shows how the networks respond to between 0 and 100 % noise in the stimulus averaged over 10 runs. The aim of this study is to ascertain if the network only responds by oscillating to the learnt stimulus and no other. The QIF network performs the best, showing a gradual decline in the amplitude of the frequency response until it reaches a minimal response at 40 % noise. Less than 0.5 amplitude implies that only a few neurons are firing hence no response is really being produced, hence it is highly selective to only its learnt stimulus. The IZ model performs almost as well. The HH model performs poorest with a less pronounced frequency amplitude decline as noise rises, and also a less stable response throughout. The less stable response is due to a high variance in the amplitude over the 10 sample runs, and is indicative of the fact that the network is inherently more volatile.

Figure 3 shows the effect of changing the length in milliseconds of the stimulus trained upon (averaged over 10 runs). All learning stages for all stimulus lengths t had the same learning time. We located the frequency with the highest amplitude only. From the figure we can see that none of the models respond significantly to

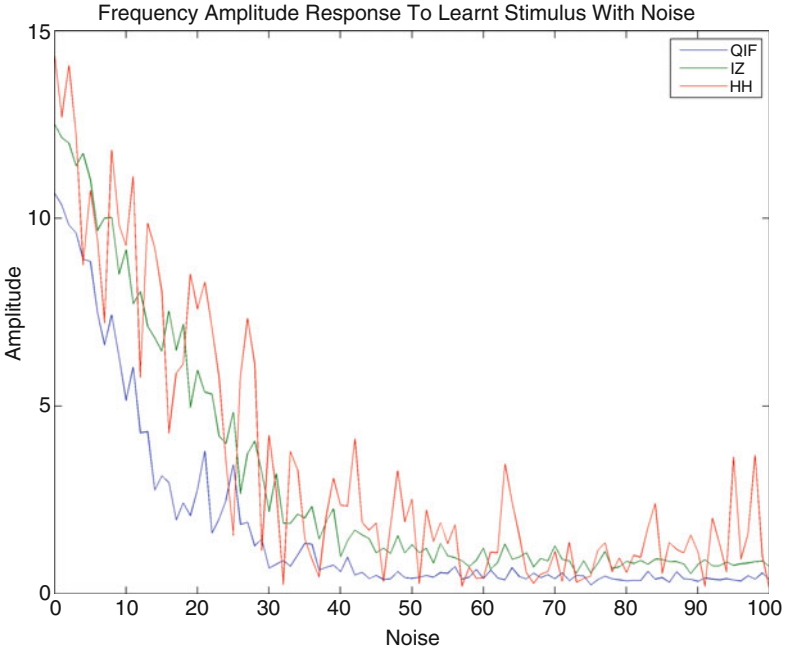


Fig. 2 Amplitude of desired 30 Hz frequency response for QIF, IZ and HH models with 0–100 % noise added to stimulus

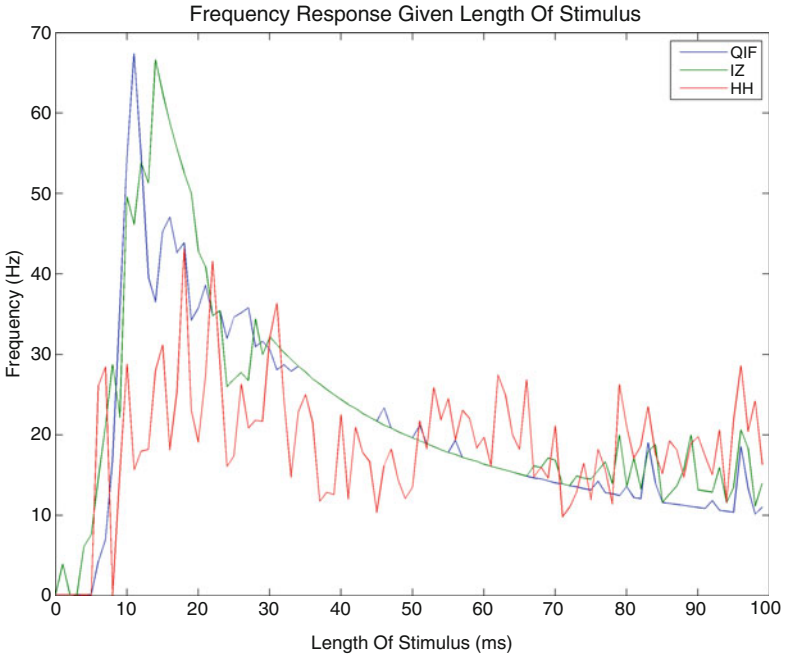


Fig. 3 Frequency response for QIF, IZ and HH models after learning with varying stimulus lengths

stimuli less than 10 ms long. Beyond this, the figure shows that for both QIF and IZ models, the length of the stimulus is roughly proportional to the frequency (f), with $f = 1{,}000/t$. This cannot be said of the HH model, which is unable to use the same network architecture to learn to oscillate at different frequencies, given only a change in the stimulus length. Having found a dependency on stimulus length, we removed the inhibitory layer from the networks and found it made no difference to the performance of QIF, IZ and HH models. We conclude that, regular repetition of a stimulus to a network that has been trained using STDP will cause oscillation at the frequency of presentation. For the HH model this further means that whilst stimulus length is important in achieving the result, the tuning of other variables is necessary to achieve the desired oscillation.

The fact that oscillatory frequency is dependent upon the length of the presentation can be elucidated by the work of Masquelier et al. [7]. They report that, '*Each time the neuron discharges in the pattern*, *it* [*STDP*] *reinforces the connections with the presynaptic neurons that fired slightly before in the pattern. As a result next time the pattern is presented the neuron is not only more likely to discharge to it*, *but it will also tend to discharge earlier*' [7]. The fact that neurons learn to always respond to a particular stimulus implies that the regular repetition of a stimulus to a recurrent network would cause the network to fire regularly at the stimulus presentation, and that this firing would become earlier and sharper, in the sense of producing narrower and more pronounced periodic bands, as learning proceeds. Hence, the resulting synchrony.

It follows from this that after an appropriate period of learning the frequency of the oscillation can be adjusted by simply altering the length of the stimulus, as it is only the beginning of the stimulus that is required to induce firing. To test this hypothesis we generated a stimulus of 100 ms, trained the network on it repeatedly until a satisfactory amplitude response was attained. We then tested the network only using the first t milliseconds repeatedly. We did this for every value of t between 13 and 100 ms. As can be seen by the results for the IZ model shown in Fig. 4, the hypothesis is correct. Hosaka et al. [2] state that in a network of excitatory and inhibitory neurons, STDP transforms a spatiotemporal pattern to temporal information. However, from the evidence above we conclude that the resultant temporality is not due to the network dynamics that result from the PING architecture, but is an artifact of repeated *periodic presentation* of a learnt stimulus. The network will respond "synchronously" whenever the stimulus is presented.

3.2 Critical Coupling Experiment

Our next investigation explored the critical coupling experiment [8] in which synchrony increases smoothly but rapidly as connection strength increases in a uniformally connected network of oscillators. We compare the results to neural oscillators that respond to any stimulus to with those that only respond to a learnt stimulus. We generated every frequency of oscillation between 10 and 50 Hz for both types of architecture using QIF neurons.

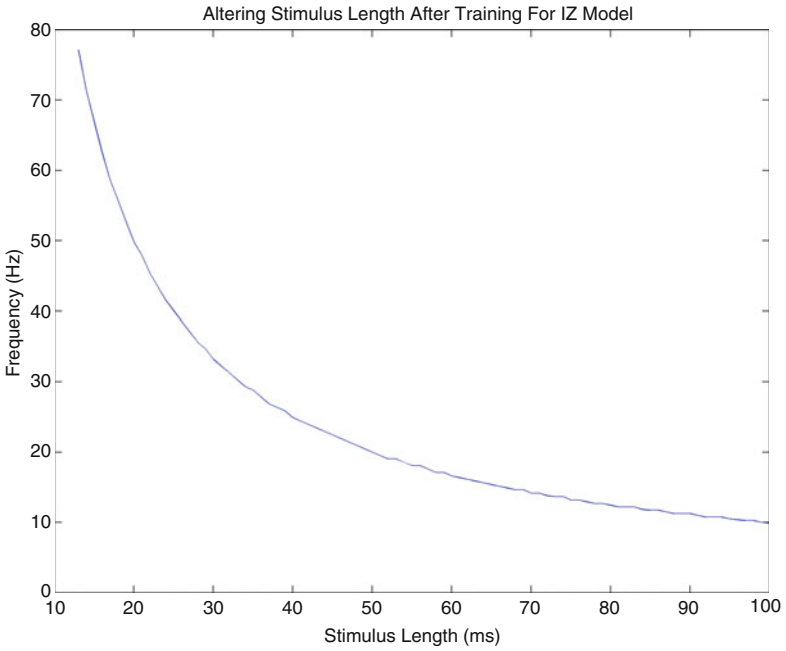


Fig. 4 IZ model after learning with stimulus. Frequency response for varying stimulus lengths presented after learning

In all our experiments we used 64 neural oscillator nodes to form a network, with frequencies selected using a Gaussian distribution with a mean of 30 HZ and a variance of 10 Hz. The phase of each oscillator was determined by the time at which external input to the oscillator was started, which varied from 0 to 100 ms. The slowest oscillator was 10 Hz and therefore a random start point ranging from 0 to 100 ms allowed for 10 Hz oscillators (as well as all oscillators of higher frequency) to be completely out of phase with each other. The neurons in the excitatory layers of each node were synaptically connected to the neurons in the excitatory layers of each other node with a connection ratio of 0.2. The experiments involved a sweep of 200 synaptic weights for all inter-node connections. Weights were set to the same value within each iteration in the parameter sweep, but with each different iteration having a different synaptic weight. On each sweep the overall synchrony of the network was measured. The networks were simulated for 2,000 ms for each iteration of the sweep. Each network comprised 16,000 neurons and 36,256,000 synapses.

Figure 5 shows the synchrony results for the evolved PING architectures that **do not use learning**. At 0 connection strength there is a synchrony of around 0.2, which indicates no synchrony at all except for coincidental alignments in phase. Synchrony rises with connection strength but so too does the spread of the dots, indicating some variation in behaviour with these systems. The synchrony levels off at 0.07 connection strength and remains the same until there is a major discontinuity at 0.17 connection strength.

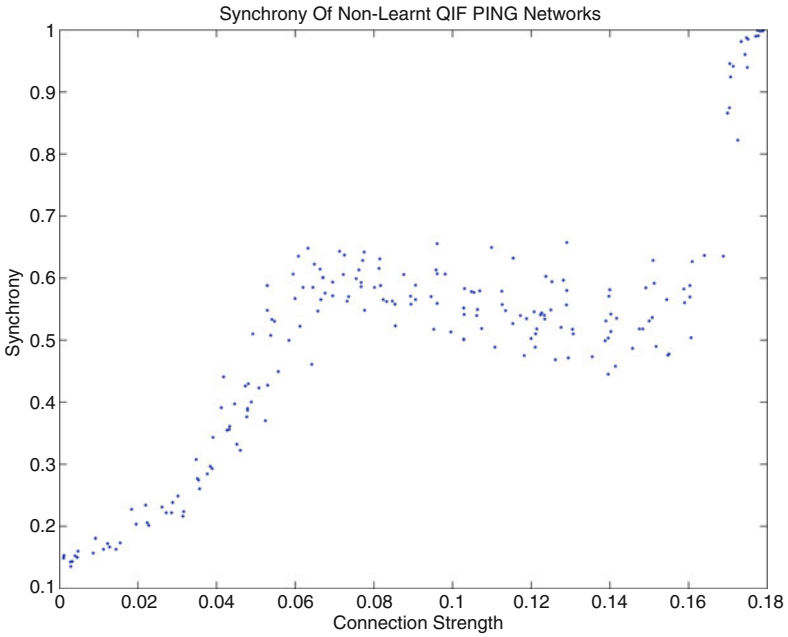


Fig. 5 Synchrony of QIF models that have not used STDP to train to respond to a particular stimulus and therefore responds to any stimuli

By contrast Fig. 6 show the synchrony for the neural oscillators **that had learnt to oscillate**. Within a critical region of connection strengths, synchrony can be seen to increase smoothly but rapidly as connection strength increases, in accord with Kuramoto's findings. The connection strength is effective at different levels from the non-learning PING model due to different sensitivities in the different architectures, Poisson process parameters, and scaling factors. However the behaviour is the key difference to note. There is a very tight sinusoidal increase, indicating little variation in behaviour with these learnt systems, unlike those in Fig. 5. There are also no discontinuities.

On reaching 100 % synchrony both types of architecture exhibited saturation, by which we mean all neurons were firing all the time. Figure 7 shows a 200 ms snippet of the pairwise synchrony between oscillator nodes at their respective maximal synchronies before saturation. The non-learning oscillator networks show deviations from full synchrony in which the network separates into sub-groups, which although they diverge, show similar phase movements indicating mutual influence between the groups. The learning architectures show little group separation behaviour but instead single oscillators seem to separate into their own phases away from full synchrony. The same behaviours are manifest at lower global synchrony levels albeit that the deviations are greater. The behaviour may be intuitively explained by the fact that in the non-learning architectures the individual intrinsic oscillatory frequencies of the nodes as well as the interaction between them are generated by

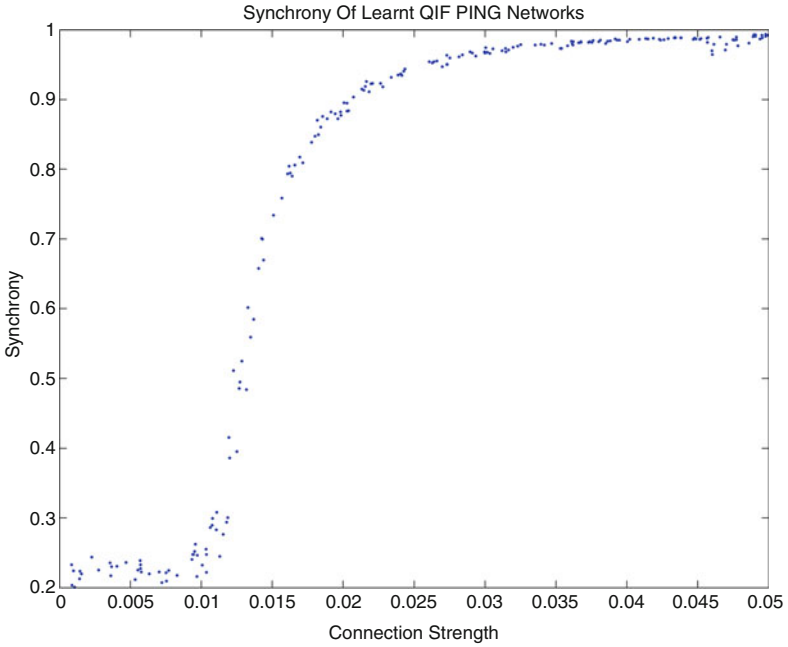


Fig. 6 Synchrony of QIF models that have used STDP to train to respond only to a particular

the network architecture which forms a complete system, whereas in the learning architectures the individual intrinsic oscillatory frequencies are created by an external stimulus that is separate from the network system and as such is unable to receive dynamic feedback and therefore facilitates more individual rather than group behaviour.

4 Discussion

It has been shown that STDP generates robust synchronous responses. After learning, the networks are highly selective for their learnt stimulus, responding at the beginning of each repeated stimulus presentation, and do not respond to other stimuli. We can conclude that repeated post-learning presentation of the stimulus overrides or interferes with the oscillations that would otherwise be caused by the delays in the PING architecture. A fast EI/IE loop will feed back and subside before the next learnt stimulus response. In this case oscillations from the periodic stimulus will take precedence over PING oscillations. Using neurons of either Type I or Type II classification produces equivalent results with STDP. However, the HH model does not perform in the same manner. The difference in the HH model is the Andronov-Hopf bifurcation and the neuron's synaptic reversal potential. The result is a less robust network that is also unable to use the same architecture to

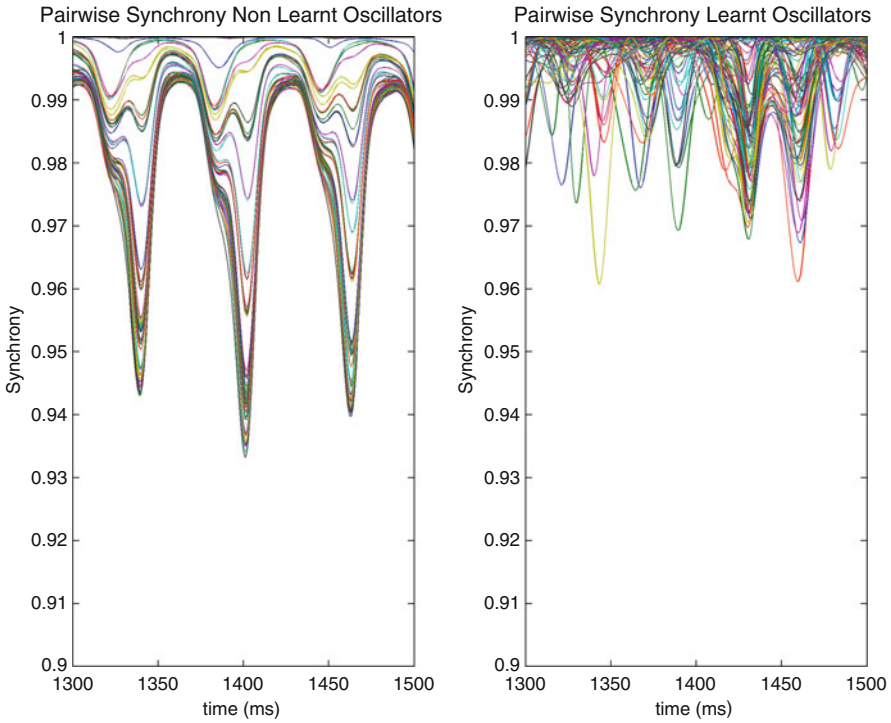


Fig. 7 Pairwise synchrony of QIF models at maximal synchrony before saturation. 200 ms shown only

learn to respond to stimuli that have a variety of presentation times. Further to this, the critical coupling experiment demonstrates that the collective behaviour of oscillatory architectures that have been pre-trained using STDP is well defined and precise, in contrast to those that have not been trained. However, the internal dynamical behaviour differs between the two architectures.

References

1. Abbott, L. F., & Nelson, S. B. (2000). Synaptic plasticity: taming the beast. Nature neuroscience, 3 Suppl(November), 1178–83.
2. Hosaka, R., Ikegnuci, T., Nakamura, H., Akaki, O. (2004). Information Transformation from a Spatiotemporal Pattern to Synchrony through STDP Network. Procs IJCNN 2004, 1475–1480.
3. Latham, P.E., Richmond, B.J., Nelson, P.G., Nirenberg, S. (2000). "Intrinsic dynamics in neuronal networks. i. theory." Journal of Neurophysiology 83(2), 808–827.
4. Ermentrout, B. (1996). Type I membranes, phase resetting curves, and synchrony. Neural Computation 8(5), 979–1001.
5. Izhikevich, E. M. (2003). Simple model of spiking neurons. IEEE transactions on neural networks. 14(6), 1569-72.

6. Hodgkin, A.L., Huxley, A.F. (1952). A quantitative description of ion currents and its applications to conduction and excitation in nerve membranes. J. Physiol. (Lond.), 117:500–544.
7. Masquelier, T., Guyonneau, R., & Thorpe, S. J. (2008). Spike timing dependent plasticity finds the start of repeating patterns in continuous spike trains. PloS one, 3(1), e1377.
8. Kuramoto, Y. (1984). Chemical Oscillations, Waves and Turbulence. Springer-Verlag, Berlin, 1984.

Robust Synchronization in Excitatory Networks with Medium Synaptic Delay

Hong Fan, Zhijie Wang, and Fang Han

Abstract The robust synchronization in a network coupled by excitatory synapses with medium synaptic delay is explored. It is found that the robust synchronization is sensitive to the reversal potential of the synapse. Furthermore, it is found that low reversal potential, strong synapse strength, and weak external drive contribute to the robust synchronization of an excitatory network. The important role of excitatory neurons in the robust synchronization is ascribed to the mechanism that the neurons firing earlier in previous cycle of the oscillation receive stronger inhibitory synaptic currents (or weaker excitatory synaptic currents).

Keywords Robust synchronization • Excitatory-inhibitory network • Synaptic currents

1 Introduction

Inhibitory neural networks are widely believed to confer advantages to the genesis of robust synchronized oscillation of high frequency [1]. As synchronized firing usually appears in neural circuits consisting of both inhibitory neurons and excitatory neurons, it is worthwhile to study the role of the excitatory neurons in the synchronization. The synchronization of excitatory neural networks may be robust when there is no synaptic delay or synaptic delay is very long (for example, longer than 10 ms). However, in locally interconnected neural circuits where synchronized firing occurs, synaptic delay is usually medium (for example, between 0.5 and 10 ms) [2]. Therefore, whether the synchronization of excitatory neural networks with medium synaptic delay is robust is an important problem. This problem has not been studied sufficiently yet and we will investigate it in this paper.

H. Fan (✉)
Glorious Sun School of Business and Management, Donghua University, 1882 West Yanan Road, 200051 Shanghai, People's Republic of China
e-mail: Hongfan@dhu.edu.cn

Z. Wang • F. Han
College of Information Science and Technology, Donghua University, 2999 North Renmin Road, 201620 Shanghai, People's Republic of China

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_35

2 Models

The membrane potential of neuron i is described by:

$$C_m \frac{dVi}{dt} = I_i^{Na} - I_i^K - I_i^L - I_i^{syn} + I_i^{app}. \quad (1)$$

The first two terms in the right side of Eq. 1 are the ionic currents responsible for the generation of the action potential; the third term is the leakage current; the fourth term is the synaptic current; the fifth term stands for the external currents, which are different neuron by neuron. The detailed description of the first three terms can be found in [3]. The external drive for each neuron is drawn from a normal distribution with mean I_{mean} and standard deviation σ. We assume that the synaptic current is the linear summation of each post-synaptic current component resulting from a single action potential. When neuron i receives a spike at time t_i, the post-synaptic current component is described by

$$I_i^{syn} = \sum_{j=1}^{N} \rho_{i,j}(t) \left[v_i(t) - E_{syn} \right] \quad (2)$$

$$\rho_{i,j}(t) = \begin{cases} 0 & t < tj + d \\ g_{syn} \frac{1}{\tau d - \tau r} \left(e^{-\frac{t - tj - d}{\tau^d}} - e^{-\frac{t - tj - d}{\tau^r}} \right) & t \geq tj + d' \end{cases} \quad (3)$$

where E_{syn} is the reversal potential of the synapse; t_j is the recent firing time of neuron j; other four parameters in Eq. 3 are the synapse strength, rise time constant, decay time constant, and synaptic delay, respectively. If E_{syn} is high (for example 0 mV), the synapse is excitatory. The purpose of this paper is to search the parameter space to find the important parameters which support the robust synchronization. Synaptic delay is medium in this paper. If not stated otherwise, the parameter values in the network are set as follows:

$$C_m = 1\ \mu F/cm^2, \tau^d = 5\ ms, \tau^r = 0.1\ ms, d = 5\ ms, N = 200.$$

3 Robust Synchronization in Excitatory Networks

As the neurons connect with each other though synapses in networks, the robust network synchronization depends on the properties of the synaptic currents. We can see intuitively from Eq. 2 that the synaptic current difference between two neurons is sensitive to the difference of the membrane potential $v_i(t)$ between two neurons, if reversal potential E_{syn} is low (Note that E_{syn} cannot be too low, as the synapse may be inhibitory if E_{syn} is too low) or synapse is strong. Consequently, the robust

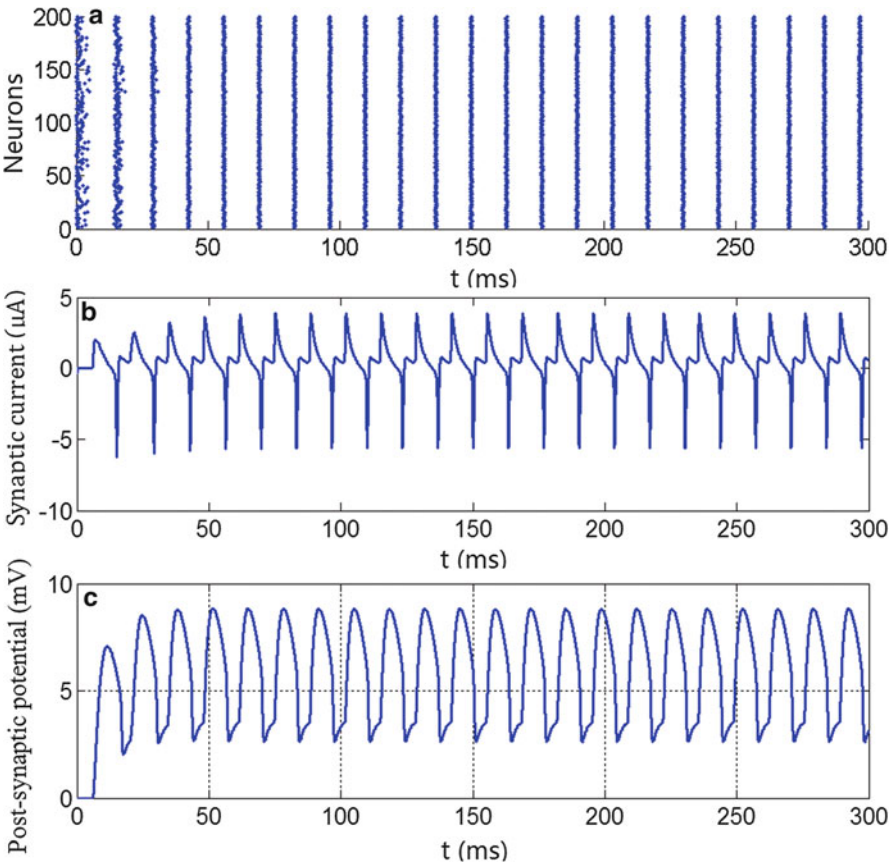


Fig. 1 Robust synchronized oscillations in excitatory networks. $E_{syn} = -52$mV, $g_{syn} = 2/N$, $I_{mean} = 0.1\mu A$, $\sigma = 0.01\mu A$. (**a**) Raster plot the spikes of neurons; (**b**) Time course of the synaptic current of a typical neuron; (**c**) Time course of the post-synaptic potential of a typical neuron

synchronization is sensitive to E_{syn} and synapse strength. Figure 1 shows that when reversal potential is low and synapse is strong, the neural firing in the network can be synchronized under a relatively high level of heterogeneity (the level of heterogeneity is characterized by the ratio of the standard deviation to the mean of the external drives). Figure 2 shows that the firing patterns in a network with medium reversal potential and medium synapse strength, where synchronization is less accurate than that in Fig. 1. When reversal potential is high and synapse strength is weak, the firing of neurons cannot be synchronized (see Fig. 3).

Figure 1b shows the time course of the synaptic currents of a typical neuron, which reveals that the synaptic current is excitatory (the negative peak of the synaptic current is caused by the action potential, as the peak value of the action potential is larger than reversal potential). Figure 1c reveals that the firing events of neurons occur at the moment when the post-synaptic potentials (PSP)

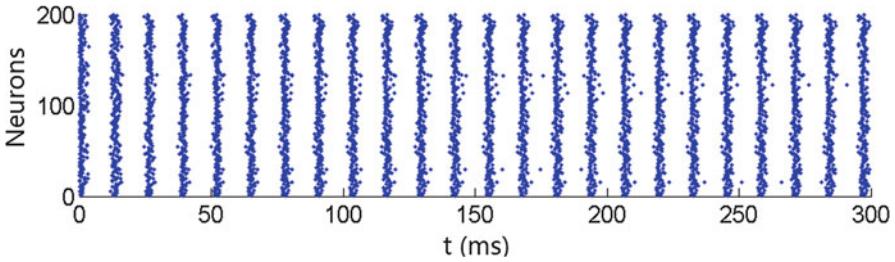


Fig. 2 Weak synchronized oscillations in excitatory networks. $E_{syn} = -45mV$, $g_{syn} = 0.7/N$, $I_{mean} = 0.5\ \mu A$, $\sigma = 0.05\ \mu A$

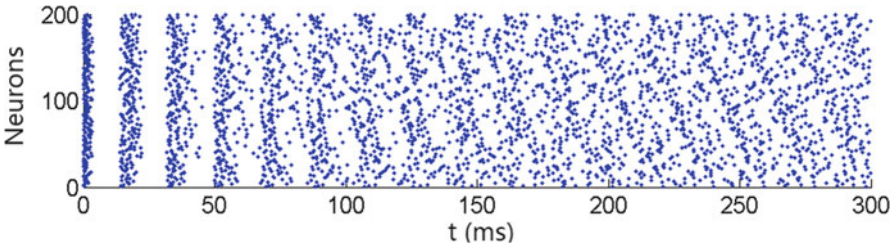


Fig. 3 Non-synchronized firing in excitatory networks. $E_{syn} = 0\,mV$, $g_{syn} = 0.1/N$, $I_{mean} = 0.5\ \mu A$, $\sigma = 0.05\ \mu A$

are decreasing. This is interesting because the stability of synchronization of a network with current-based synapses [4] requires that the firing events occur at the moment when the PSP is rising. But in a network with conductance-based synapses, there are two kinds of synaptic components contributing to the robust network synchronization [5, 6]. Stated in other words, the heterogeneity of the neurons can be overcome by the two effects of PSP differences among these neurons. One is the shape effect of PSP, which contributes to the network synchronization if PSP is rising at the moment when firing event occurs. The other is the amplitude effect, which contributes to the network synchronization due to the fact that neurons fire earlier in the previous cycle of the oscillation will receive stronger inhibitory current (or weaker excitatory synaptic current). The amplitude effect dominates in most cases [6]. Ascribing to the amplitude effect, the network synchronization is robust even the PSP is declining at the moment when firing event occurs. This explains the robust network synchronization in an excitatory network with medium synaptic delay, which is shown in Fig. 1a.

4 Conclusions and Discussion

Simulation results show that synchronization in excitatory networks with medium synaptic delay is robust, when reversal potential is low, synapses are strong, and external drives are weak. The robust synchronization in excitatory networks may

also play an important role in the genesis of neural rhythm in E-I networks. In E-I networks, each neuron receives both excitatory and inhibitory synaptic currents. As well as inhibitory synaptic currents, excitatory synaptic currents contribute to overcome the heterogeneity of the neurons. In E-I networks, synapses can be very strong while the network frequency keeps roughly unchanged, as excitatory and inhibitory synaptic currents cancel each other. However, the effects that excitatory and inhibitory synaptic currents contribute to overcome the heterogeneity of the neurons do not cancel each other. Therefore, it is conceivable that synchronization in E-I network can be rather robust when synapses are strong.

Acknowledgments This research is supported by the NSF of China (Grants No. 61075105, No. 70971021 and No. 11102038).

References

1. M. Bartos, I. Vida, and P. Jonas, Nat. Rev. Neurosci. 8 (2007):45–56.
2. I. Vida, M. Bartos, and P. Jonas, Neuron 49 (2006):107–117.
3. X.J. Wang and G. Buzsáki, The Journal of Neuroscience 16(1996): 6402–6413.
4. W. Maass and C.M. Bishop, (Eds.), Pulsed Neural Networks, MIT Press, Cambridge, MA, (1998).
5. Z. Wang, H. Fan, and K. Aihara, Phys. Rev. E 83(2011): 051905.
6. Z. Wang, W.K. Wong, Neural networks 43(2013): 55–62.

Synchronization in Neuronal Population with Phase Response

Xianfa Jiao, Danfeng Zhu, and Rubin Wang

Abstract In the present study, we have formulated a phase description of a neuronal oscillator with non-instantaneous synaptic inputs, by using the phase sensitivity function. By numerical simulation, we found that the synaptic time constant is an important factor for global network synchronization. If the synaptic time constant is smaller, perfectly synchronized behavior quickly occurs. As the synaptic time constant is increased, periodic synchronization emerges. However, synchronized activity is lost for larger synaptic time constant. The external periodic stimulation can change the synchronized patterns in the neuronal population. With a stronger stimulation or high-frequency stimulation, synchronized bursting occurred in the neuronal population.

Keywords Neuronal population • Phase response • Synaptic input • External stimulus

1 Introduction

Oscillations are ubiquitous in the nervous system. Many experiments have shown that the complex interaction between neurons can induce various rhythmic activity in the nervous system [1]. Both the normal physiological function or abnormal physiological disorders (such as Parkinson's disease, epilepsy and so on) are related to the synchronized neural activities with various frequency [2, 3]. Synchronized activity among neurons and the formation of neuronal clusters are considered as a fundamental mechanism for cognitive function and consciousness [4, 5]. Despite the ubiquity and importance of synchronized activity, the underlying mechanism and the key system parameters are not yet known, and little attention has been paid to investigating the dynamic response of an oscillator network to external stimuli. The phase

X. Jiao (✉) • D. Zhu
School of Mathematics, Hefei University of Technology, Hefei, People's Republic of China
e-mail: xfjiao@126.com

R. Wang
Institute for Cognitive Neurodynamics, East China University of Science and Technology, Shanghai, People's Republic of China

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_36

response curve (PRC) represents how an external stimulus affects the timing of spikes immediately after the stimulus in repetitively firing neurons [6]. PRC describe the phase shift of the perturbed neuronal oscillator when a neural oscillator receive external input or synaptic input [7–9]. In order to explore the dynamic mechanism of synchronous activity in the nervous system, the phase response curve is an important and effective method [10–14]. In the nervous systems, cortical neurons undergo massive synaptic bombardment and ever-found perception information stimulation. To understand the response properties of neurons operating in this regime, we investigate a model neuron as a neuronal oscillator with non-instantaneous synaptic inputs represented by $\alpha -$ function, and external periodic stimulus.

2 Model

We consider a neural population composed of N neural oscillators, where neuronal oscillators are identical and globally coupled with each other, subject to a common external periodic force. The phase of *jth* oscillator θ_j obeys the evolution equation:

$$\frac{d\theta_j}{dt} = \omega + \frac{\varepsilon}{NZ(\theta_j)} \sum_{k=1}^{N} \sum_{n} \alpha\left(t - t_k^n\right) + c \sin(\omega t)\, Z\left(\theta_j\right) \tag{1}$$

where $\frac{\varepsilon}{N} \sum_{n} \alpha\left(t - t_k^n\right)$ is the input to the *jth* neuronal oscillator from the *kth* neuronal oscillator, ε is weak coupling constant, N is the total number of neuronal oscillators, t_n^k is the *nth* firing time of the *kth* neuronal oscillator, $\alpha(t)$ is a causal coupling function. θ_j is the phase of *jth* neuronal oscillator, ω is the natural frequency of a neuronal oscillator, $Z(\theta)$ is a phase response curve of a neuronal oscillator, $c \sin(\omega_0 t)$ is an external periodic force with a strength c and frequency ω_0.

We assume that the mutual interaction shift the frequency of the mean phase of these oscillators by $\varepsilon\Omega$ from the natural frequency ω, and define the relative phase $\psi_j = \theta_j - (\omega + \varepsilon\Omega)t$. The relative phase ψ_j changes slowly compared with and will hardly change during the oscillation period θ_j. Therefore, we substitute ψ_j into Eq.(1), and average Eq. (1) over one period keeping ψ_j constant, so the relative phase ψ_j obeys the following equation:

$$\begin{aligned} \frac{d\psi_j}{dt} = & -\varepsilon\Omega + \frac{\varepsilon}{NT} \int_0^T Z\left(\psi_j + (\omega + \varepsilon\Omega)\, t\right) \sum_{k=1}^{N} \sum_{n} \alpha\left(t - t_k^n\right) \\ & + c \sin(\omega t)\, Z\left(\psi_j + (\omega + \varepsilon\Omega)\, t\right) \end{aligned} \tag{2}$$

In order to investigate the dynamic response of neural population, we introduce complex order parameters describing synchronized phenomenon in the neuronal population

$$\mathrm{Re}^{i\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\psi_j} \tag{3}$$

where R is the amplitude of the order parameters describing the degree of synchronization of neuronal oscillators, $0 \leq R \leq 1$, the bigger R show that synchronous activity is stronger, ψ is the average phase of the neuronal population.

3 Result

We investigate the response property of neuronal population to external periodic stimulus. The phase sensitivity function $Z(\theta)$ is considered as a sinusoidal sensitivity function $\sin(\theta)$ as in Ref. [15]. As synaptic time constant τ is smaller, the neuronal population quickly synchronized in-phase (Fig. 1a); but with τ increased, periodic synchronization occurred, and as the synaptic time constant is larger, the synchronization become weaker (Fig. 1b, c); even more, synchronized activity can be lost (Fig. 1d). This shows that synaptic time constant is an important condition under which the global neural network synchronized.

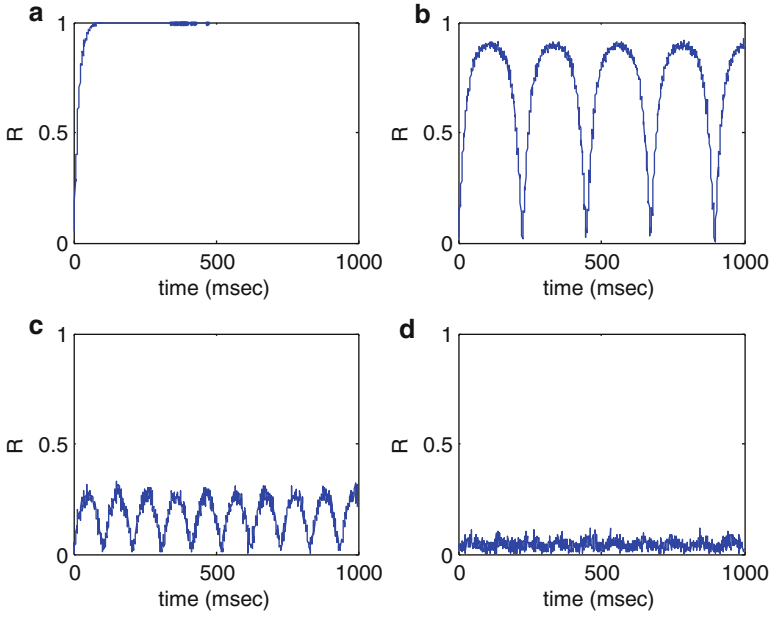


Fig. 1 Time evolution of the order parameter R by varying τ in the absence of stimulation. Parameters: $\varepsilon = 0.01$, $\omega = 3$, $\Omega = 2\pi$, $T = 2.05$, (**a**) $\tau = 0.4$, (**b**) $\tau = 0.5$, (**c**) $\tau = 0.7$ (**d**) $\tau = 1.4$

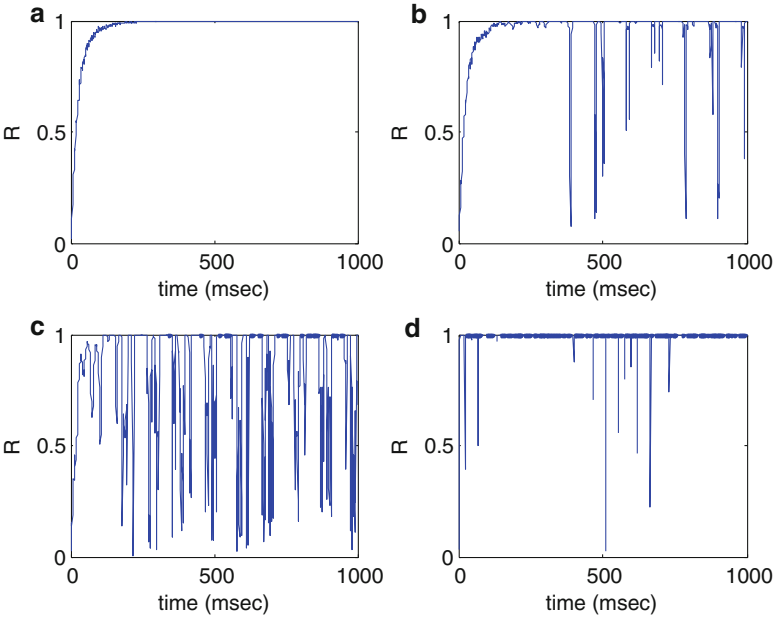


Fig. 2 Time evolution of the order parameter R in the presence of stimulation. Parameters: $\varepsilon = 0.01$, $\omega = \omega_0 = 3$, $\Omega = 2\pi$, $\tau = 0.485$, (**a**) $c = 0$, (**b**) $c = 0.003$, (**c**) $c = 0.03$ (**d**) $c = 1.5$

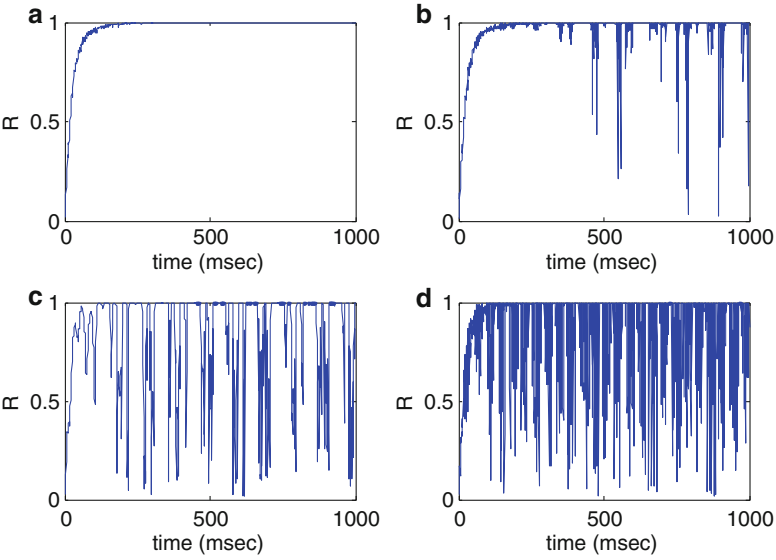


Fig. 3 Time evolution of the order parameter R in the presence of stimulation. Parameters: $\varepsilon = 0.01$, $\omega = \omega_0 = 3$, $\Omega = 2\pi$, $\tau = 0.485$, (**a**) $\omega_0 = 0.03$, (**b**) $\omega_0 = 0.03$, (**c**) $\omega_0 = 3$ (**d**) $\omega_0 = 30$

Without stimulation, the neuronal population quickly tend to perfect synchronization (Fig. 2a); With a weak stimulation, the neuronal activity transition from the perfect synchronization to synchronized bursting occurred (Fig. 2b); As the stimulus intensity was increased, the synchronized bursting duration is prolonged (Fig. 2c); However, the synchronized bursting duration is shorten in the presence of stronger stimulus (Fig. 2d). As the stimulus frequency was increased, the neuronal activity transition from the perfect synchronization to synchronized bursting occurred (Fig. 3b); As the stimulus frequency was further increased, the synchronized bursting becomes stronger (Fig. 3c, d).

4 Conclusions

In the present study, we have formulated the phase description of the neuronal oscillator with non-instantaneous synaptic inputs represented by $\alpha -$ function, and external periodic stimulus by using the phase sensitivity function. The synaptic time constant is an important parameter for perfect synchronization, periodic synchronization, synchronized bursting in the global neural network synchronized. The influence of external periodic stimulation on the neuronal population depends on stimulus intensity and frequency. With a stronger stimulation or a high frequency stimulation, synchronized bursting occurred.

Acknowledgments This work was supported by the National Natural Science Foundation of China under Grant No. 11232005 and No. 11172086.

References

1. Buzsaki G. (2006), Rhythms of the brain, Oxford University Press.
2. Hammod C. et al. (2007) Trends Neuroscience. 30, 357–364.
3. Fries, P. (2002), J. Neurosci. 22, 3739–3754.
4. Gray C.M. (1994), J Comput Neurosci. 1, 11–38.
5. Maldonado, P.E. et al. (2000), Cereb. Corter. 10, 1117–1131.
6. Dutkin, B. S. et al. (2005), J. Neurophysiol. 94, 1623–1635.
7. A.T.Winfree. (1980), The Geometry of Biological Time. Springer, New York.
8. Y. Kuramoto. (1984) Chemical Oscillations, Waves, and Turbulence. Springer, New York.
9. E. M. Izhikevich. (2007), Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting. MIT Press, Cambridge.
10. Roy M. Smeal et al. (2010), Phil.R.Soc.B. 365, 407–2422.
11. Roberto F. Galan et al. (2005), Phy. Rev. Lett. 94, 158101–4.
12. Yasuhiro Tsubo et al. (2007), Euro. J. Neurosci. 25, 3429–3441.
13. Yauhiro Tsubo et al. (2007), Phy. Rev. Lett. 99, 228101.
14. G. B. Ermentrout et al. (2008), Trends Neurosci. 31, 428.
15. A. Pikovsky et al. (2001), Cambridge University Press, Cambridge.

Geometry of Dynamic Movement Primitives in Neural Space: A FORCE-Learning Approach

Hiromichi Suetani

Abstract Dynamic movement primitives are one of key concepts for understanding dexterous and flexible movements of biological bodies. In the field of robotics engineering, simple types of nonlinear differential equations are used to generate movement primitives from demonstrations, but it remains unclear how nonlinear dynamics in the real brain can also generate movement primitives in biologically natural ways. The aim of this study is to investigate a possible role of nonlinear dynamics in random recurrent neural networks (RNNs) for skillful motor learning. We show that one-shot temporal patterns such arm reaching movements can be trained by a type of RNN-learning so-called FORCE-learning recently proposed by Sussillo and Abbott and a number of patterns are summarized as a manifold embedded in a space of synaptic weights of readout neurons. We also discuss how generalization of learning against untrained motor patterns can be achieved by identifying nonlinear coordinates (meta-parameters) on this manifold in a higher level of the central nervous system.

Keywords Dynamic movement primitives • Nonlinear dynamics • Robotics • Recurrent neural networks • Motor learning

1 Introduction

Generation of dexterous body movements is an indispensable condition of biological systems [1]. As one of strategies for acquiring skilled movements, it has been considered that human beings first acquire a number of standard movements referred to as "movement primitives" through repetitive experiences, then integrate them into rich varieties of sequential behavior.

From the viewpoint of nonlinear dynamical systems, movement primitives can be regarded as objects of nonlinear dynamical systems, e.g., such as fixed points,

H. Suetani (✉)
Department of Physics and Astronomy, Kagoshima University, Kagoshima 890-0065, Japan

Decoding and Controlling Brain Information, PRESTO, JST, Kawaguchi 332-0012, Japan
e-mail: sustain@sci.kagoshima-u.ac.jp

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_37

limit cycle attractors and heteroclinic orbits from one of attractors to another one. Actually, a number of movement primitives for humanoid robots with high degrees of freedom has been successfully obtained from nonlinear differential equations whose parameters are estimated from demonstrations of movement data [2].

On the other hand, the brain exhibits rich variety of nonlinear dynamics from the single neural cell level to the cortex level. Especially, chaotic behaviors attracted much attention in the context of biological functions of the brain [3]. From the viewpoint of neuroscience, it becomes interests how movement primitives can be interpreted and generated as nonlinear dynamics of neural networks.

In real biological brains, asymmetrical connectivity among neurons including random recurrent connection is more natural than symmetrical ones like the Hopfield networks yielding only static solution, and it actually shows more complicated dynamics including limit cycles and chaos [4]. In this case, however, such chaos does not play a functional role whereas patterns can be memorized as point attractors in the Hopfield networks.

After several years have passed, random recurrent neural networks (RNNs) attract considerable attentions again by the recent progress of the so-called reservoir computing (RC) paradigm [5]. In this study, we employ a FORCE-learning machine proposed by Sussillo and Abbott [6] along the paradigm of RC in order to explore how chaotic activity in RNNs can be a useful function to organize dynamic movement primitives.

2 Methods

A RNN for FORCE-learning is defined by the following equations

$$\begin{aligned}\tau \dot{x}_n(t) = &-x_n(t) + g \sum\nolimits_{n'=1}^{N} w^{\mathrm{rec}}_{nn'} \tanh(x_{n'}(t)) \\ &+ \sum\nolimits_{m=1}^{M} w^{\mathrm{in}}_{nm} u_m(t) + \sum\nolimits_{l=1}^{L} w^{\mathrm{fb}}_{nl} z_l(t). \end{aligned} \tag{1}$$

Here, $x_n, n = 1, \ldots, N$ is the internal state of the n-th neuron, $u_m, m = 1, \ldots, M$ is the m-th input current, and $z_l, l = 1, \ldots, L$ is the state of l-th readout neuron determined as

$$z_l(t) = \sum_{n=1}^{N} w^{\mathrm{out}}_{ln} r_n(t), \; r_n(t) = \tanh(x_n(t)). \tag{2}$$

From the spirit of RC, in the FORCE-learning process, $w^{\mathrm{rec}} = (w^{\mathrm{rec}}_{nn'})$, $w^{\mathrm{in}} = (w^{\mathrm{in}}_{nm})$ and $w^{\mathrm{fb}} = (w^{\mathrm{fb}}_{nl})$ are fixed and sparsely random matrices, and only the matrix $w^{\mathrm{out}} = (w^{\mathrm{out}}_{ln})$ is incrementally modified according to the recursive least square (RLS) or the gradient methods where the difference between the readout states and teacher signals is used as error signals [6].

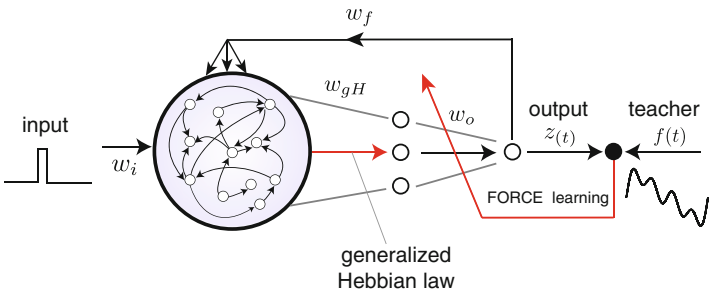


Fig. 1 Incorporation of generalized Hebbian-learning with FORCE-learning

In addition to the basic structure of the FORCE-learning, we also incorporate an intermediate layer with K neurons as shown in Fig. 1, in order to (i) reduce the number of essential dimensionality of the dynamics inside the RNN and (ii) compare RNNs with different topologies (different realizations of random numbers for preparing the recurrent connectivity matrix w^{rec}). Here, the connections between neurons in RNN and ones in the intermediate layer is incrementally updated through the "generalized Hebbian-learning (GHA)" [7] as

$$\Delta w_{kn}^{gH} = \eta(r_n(t) - \sum_{k'=1}^{k} w_{k'n}^{gH} y_{k'}(t)) y_k(t) \tag{3}$$

The k-th neurons y_k gives the k-th principal component of the dynamics of RNN. It is expected that the dynamics of two RNNs are statistically equivalent, the readout matrices for these two RNNs are almost same against the same teacher signal. We call neural networks with the structure shown in Fig. 1 the FORCE-GHA networks in the following.

3 Numerical Experiments and Results

As an illustrative example of our approach, we consider generation of reaching movement patterns executed by a two-link arm (Fig. 2a) by the training of the FORCE-GHA networks. Here, the purpose of neural networks is to generate the torques u_1 and u_2 that rotate the links of an arm. In actual motor learning, the data of the torques is not directly given as teacher signals, neural networks have to acquire them in an unsupervised way such as reinforcement learning. But, this is not a main topic in this study. So, to skip this problem, we first generate arm movement trajectories using the minimum jerk principle [8] in the joint space, then obtain the torque data using the inverse model of a two-link arm used in [9]. We employ these torque data as teacher signals for the training of the FORCE-GHA networks.

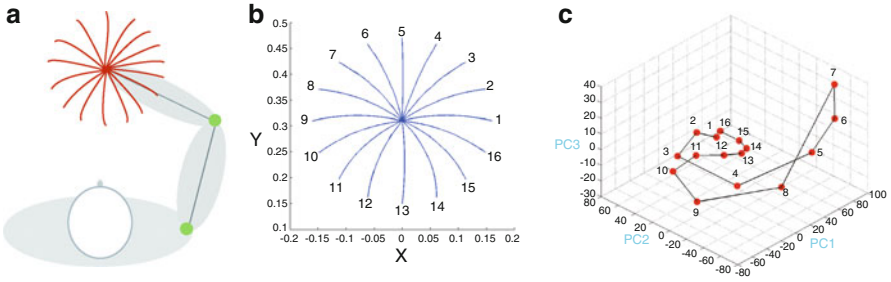


Fig. 2 (**a**) Schematic plot of reaching movements by a two-link arm. (**b**) The 16 reaching movements in the task space. (**c**) Synaptic weights of readout neurons corresponding to (**a**) in PCA-space. It forms a one-dimensional closed curve $\mathscr{S}$. $g = 1.2$

We focus on how arm movements are embedded in the neural space (here we refer the neural space as the space of the readout matrices w^{out}). Imagine that a number of the FORCE-GHA networks forms a structure such as micro columns in the cerebellum and each network is trained to generate a pair of torque patterns u_1 and u_2. For each arm movement shown in Fig. 2b, there exists a FORCE-GHA network and a fixed readout matrix w^{out} after training. Figure 2c shows a projections of 16 readout matrices w^{out} into PCA-space in the case of $g = 1.2$, which corresponds to 16 arm movements shown in Fig. 2b (here, each $2 \times K$ readout matrix w^{out} is re-aligned to a column vector with $2K$ dimensionality). We can see that these points are lying on a nonlinear closed curve (one-dimensional manifold) $\mathscr{S}$ in the PCA-space which keeps the topological continuity in arm movements in the task space. This result implies that an untrained pattern of the torques for a new reaching movement can be generated by interpolating between these trained vectors. Such an interpolation is achieved as a coordinate Θ, i.e., "phase variable", on $\mathscr{S}$, which is considered as a representation of movement in a higher level of the central nervous system(CNS). Each point obtained from the training of finite experiences is considered as a "primitive" in order to construct a manifold like $\mathscr{S}$ in the CNS.

Furthermore, we investigate how the one-dimensional manifold $\mathscr{S}$ is changed according to affine transformations of the starting point of reaching movements in the task space, and results are shown in Fig. 3. We can see that $\mathscr{S}$ is changing with preservation of its topology according to affine transformations in the task space. Therefore, in addition to the coordinate Θ, there exist other coordinates to characterize the total manifold constructed by integration of $\mathscr{S}$. CNS can acquire new, untrained patterns immediately using such higher level parameters given from finite experiences.

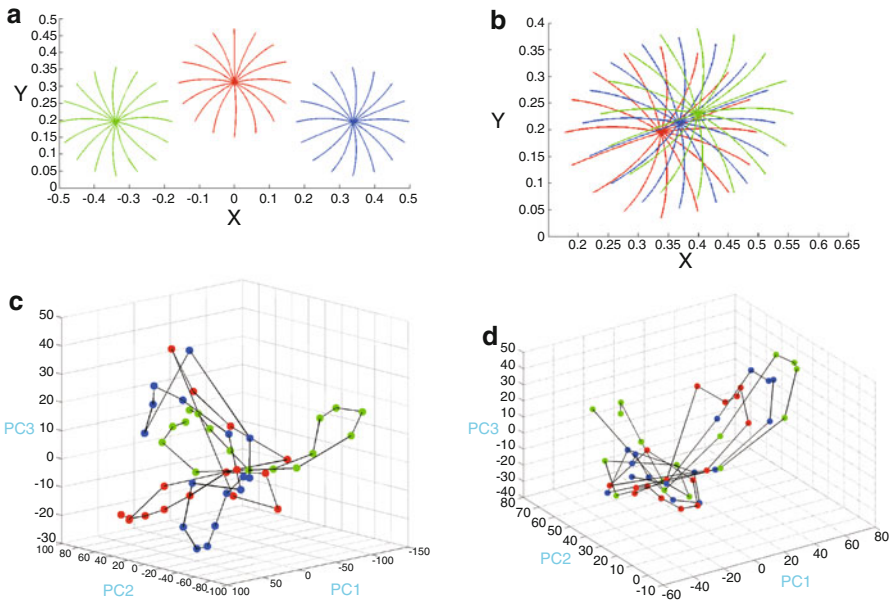


Fig. 3 (**a**, **b**) Affine transformations of 16 reaching movements in Fig. 2b. (**c**, **d**) Change of the manifold $\mathscr{S}$ against the corresponding affine transformations $g = 1.2$

4 Conclusions

In summary, we have investigated motor pattern generations by learning of RNNs, especially using a combination of FORCE-learning(supervised learning) and generalized Hebbian learning(dimensionality reduction). We showed that temporally changing motor patterns can be encoded appropriately into fixed readout matrices (synaptic weights) which keep topological feature of the task space for movements. We also discussed a possibility of acquisition of generality for motor patterns by learning of nonlinear coordinates that spans a manifold in the readout matrix space. We will develop a unified approach to hierarchical learning for generalizations of motor patterns in future study.

Acknowledgements We would like to thank I. Tsuda, S. Akaho and Y. Sakaguchi for fruitful discussions. We also thank D. Rodriguez for preparing arm reaching movements data. This study is partially supported by Grant-in-Aid for Scientific Research (No. 24120713), MEXT, Japan.

References

1. N. Bernstein, *The Coordination and regulation of movements*, Pergamon, 1967.
2. A.J. Ijspeert, J. Nakanishi and S. Schaal, Learning Attractor Landscapes for Learning Motor Primitives, In: Advances in neural information processing systems, 1523 (2002).
3. I. Tsuda, Toward an Interpretation of Dynamic Neural Activity in terms of Chaotic Dynamical Systems, Behavioral and Brain Sciences **24**, 793 (2001).
4. H. Sompolinsky, A. Crisanti and H.J. Sommers, Chaos in Random Neural Networks, Physical Review Letters **61**, 259 (1988).
5. H. Jaeger, W. Maass and J. Principe, Special Issue on Echo State Networks and Liquid State Machines, Neural Networks **20** 287 (2007).
6. D. Sussillo and L.F. Abbott, Generating Coherent Patterns of Activity from Chaotic Neural Networks, Neuron **63**, 544 (2009).
7. T.D. Sanger, Optimal Unsupervised Learning in a Single-layer Linear Feedforward Neural Network, Neural networks **2**, 459 (1989).
8. T. Flash and N. Hogan, The Coordination of Arm Movements: An Experimentally Confirmed Mathematical Model, The Journal of Neuroscience **5**, 1688 (1985).
9. E. Todorov and W. Li, A Generalized Iterative LQG Method for Locally-optimal Feedback Control of Constrained Nonlinear Stochastic Systems, In: American Control Conference, Proceedings of the 2005 IEEE, 300 (2005).

Contribution of Endogeus Acetylcholine to STDP Induction

Eriko Sugisaki, Yasuhiro Fukushima, Minoru Tsukada, and Takeshi Aihara

Abstract The synaptic plasticity is thought to play a crucial role in learning and memory. It has been reported that the relative timing of pre- and postsynaptic activity determines the direction and extent of hippocampal synaptic changes known as spike-timing dependent plasticity (STDP). Pre-post timing where presynaptic spike precedes a post synaptic spike induces long-term potentiation (LTP). Meanwhile, cholinergic inputs are considered to be integrated with sensory inputs and to play an important role in learning and memory. Cholinergic neurons in the medial septum are projecting their terminals to hippocampal CA1 to release acetylcholine (ACh). In order to investigate the influence of endogenous ACh on STDP, pre-post timing of STDP protocol was applied under interneuron-activated CA1 network. As the results, STDP was enhanced in the presence of eserine, and this enhancement was induced by the muscarinic and nicotinic activation not only on pyramidal neuron but also on interneurons. These findings suggest that ACh plays a critical role as a modulator for synaptic plasticity in hippocampal CA1 network.

Keywords Acetylcholine • Hippocampus • CA1 • Plasticity • STDP

1 Introduction

The synaptic plasticity is thought to serve as the cellular substrate for various forms of learning and memory. It has been reported that the relative timing between pre- and post-synaptic spiking determines the direction and extent of synaptic changes in a critical temporal window, process known as spike timing-dependent plasticity (STDP) [1, 2]. Pre-post timing of STDP protocol where excitatory postsynaptic potential (EPSP) induced in Schaffer collaterals precedes a back-propagating action potential (BPAP) from the postsynaptic neuron induces long-term potentiation

E. Sugisaki (✉) • M. Tsukada • T. Aihara
Brain Science Institute, Tamagawa University, 6-1-1 Tamagawagakuen, Machida, Tokyo 194-8610, Japan
e-mail: e-sugisaki@eng.tamagawa.ac.jp

Y. Fukushima
Kawasaki University of Medical Welfare, Karashiki, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_38

(LTP), while the opposite timing induces long-term depression (LTD). Meanwhile, cholinergic system is playing an important role in learning and memory [3]. Cholinergic neurons projecting to CA1 in hippocampus via fimbria [4] are distributed mainly in the medial septum. Shinoe et al. [5] reported that slow EPSP [6] induced by the repetitive stimulation of cholinergic axons enhanced LTP due to muscarinic acetylcholine (ACh) receptor (mAChR) activation. It is also reported that the synaptic plasticity was enhanced by the activation of nicotinic ACh receptor (nAChR) by the application of chronic nicotine [7].

In this study, cholinergically induced STDP in pre-post timing protocol was investigated, and resulted in STDP was enhanced in the presence of eserine. Furthermore, this regulation was influenced by the activation of mAChRs and nAChRs not only on pyramidal neurons but also on interneurons.

2 Method

All procedures were approved by the Tamagawa University Animal Care and Used Committee. Hippocampal slices (400 μm in thickness) were prepared from Wistar rats (20–25 days old). Whole cell patch clamp recordings were made from the soma of CA1 pyramidal neurons, and the membrane was kept at -64 ± 1 mV. In IPSC experiments, IPSCs were isolated by holding the membrane potential at 0 mV in voltage clamp. Depending on the experiments, eserine, atropine and mecamylamine were added to ACSF 5 min. before the application of STDP-induction protocol until it was finished. Picrotoxin was added throughout the experiments when necessary. Stimulation in Schaffer collaterals was injected to induce EPSP of 2–4 mV amplitude as a baseline. STDP protocol pattern was shown in Fig. 1. The magnitude of plasticity was defined as (averaged EPSP slopes obtained from 20 to 30 min after STDP-inducing stimulus)/(averaged baseline EPSP slopes).

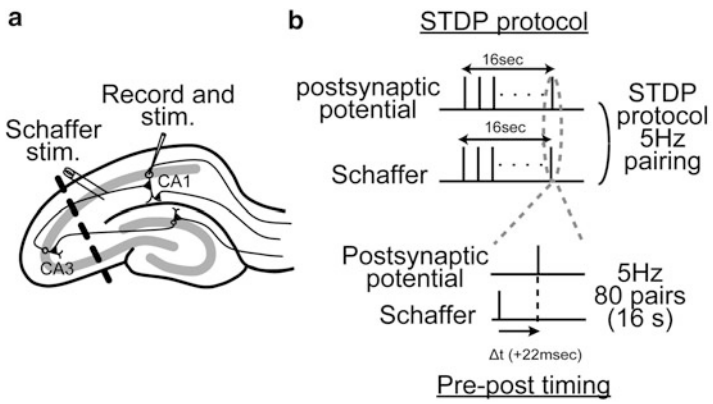


Fig. 1 Stimulation pattern

3 Result

3.1 STDP Was Enhanced by the Application of Eserine

First, in order to investigate the influence of endogenous ACh on the induction of STDP, pairing stimulation was applied in the presence of different concentrations (2 and 10 μM) of eserine. When an STDP-inducing stimulus in pre-post timing was applied, LTD in the control condition (70.1 ± 7.1 %, $n = 5$, $P < 0.05$) was changed to LTPs in the presence of 2 μM eserine (111.9 ± 2.1 %, $n = 5$, $P < 0.01$, $P < 0.01$ vs. control) and 10 μM eserine (120.5 ± 4.8 %, $n = 6$, $P < 0.01$, $P < 0.01$ vs. control) respectively as shown in Fig. 2. These results show that STDP was shifted toward potentiation by the activation of AChRs. Next, in order to clarify the effect of $GABA_A$ receptor ($GABA_AR$) activation on cholinergically induced STDP, picrotoxin (25 μM) was added to ACSF in the presence of eserine 2 μM. When the STDP-inducing protocol in pre-post timing was applied only in the presence of 2 μM eserine, LTP was observed as in Fig. 2, and the LTP was significantly enhanced in the additional application of picrotoxin (179.3 ± 9.1 %, $n = 5$, $P < 0.01$, $P < 0.01$ vs. eserine 2 μM only; Fig. 3a). As these results indicate that $GABA_AR$ contributed to the cholinergically-induced STDP, next the dependence of eserine on IPSCs were confirmed. The magnitude of IPSCs in control condition was significantly enhanced in the presence of 2 μM eserine (1.28 ± 0.06 %, $n = 4$, $P < 0.01$, $P < 0.01$ vs. control; Fig. 3b, c). These results show that interneurons were strongly activated by the application of eserine, therefore the cholinergically induced STDP was regulated by interneurons.

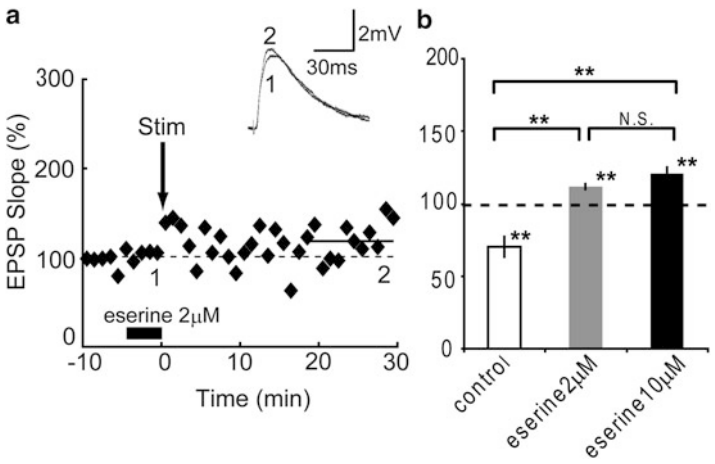


Fig. 2 Eserine effect on STDP

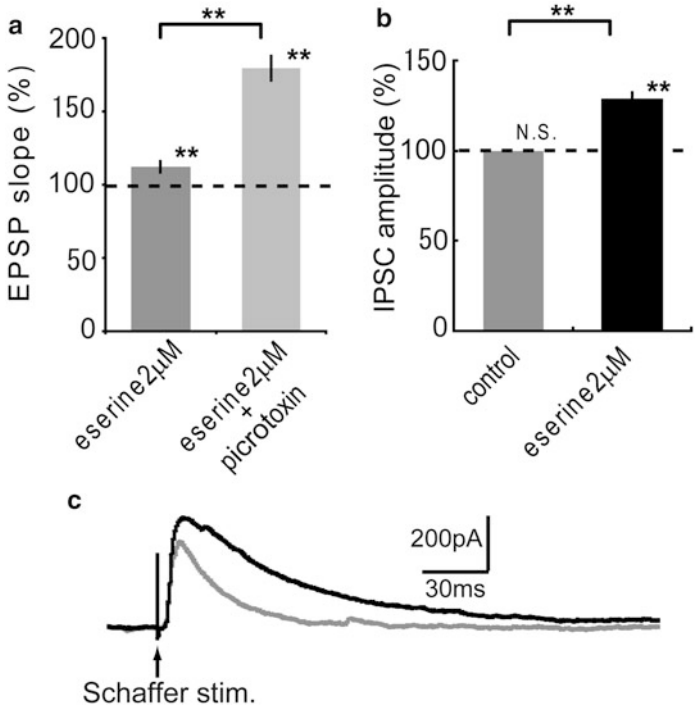


Fig. 3 Eserine effect on IPSC

3.2 Cholinergically-Induced STDP Was Contributed by mAChR and nAChR

As STDP was enhanced by the application of eserine, in second, the dependence of each AChR was clarified. When an STDP-inducing protocol in pre-post timing was applied in the presence of eserine 2 μM and picrotoxin 25 μM, LTP was observed as in Fig. 4a. When atropine at 1 μM was added to the eserine and picrotoxin treated ACSF, LTP was abolished (109.2 ± 4.0 %, $n = 6$, N.S., $P < 0.01$ vs. eserine 2 μM only; Fig. 4a), while LTP was significantly reduced in the presence of 3 μM mecamylamine instead of atropine (155.6 ± 4.3 %, $n = 5$, $P < 0.01$, $P < 0.05$ vs. eserine 2 μM only, $P < 0.01$ vs. eserine 2 μM + atropine 1 μM). According to the results, cholinergically-induced STDP was regulated by the activation of mAChRs more effectively than nAChRs on pyramidal neuron. Furthermore, in order to investigate the effect of AChRs on interneurons, similar experiments were performed in the absence of picrotoxin. LTP induced in the presence of eserine 2 μM (Fig. 2) was significantly enhanced with additional application of atropine 1 μM (131.4 ± 5.7 %, $n = 5$, $P < 0.01$, $P < 0.01$ vs. eserine 2 μM only; Fig. 4b), while STDP was changed to LTD in the presence of 3 μM mecamylamine instead

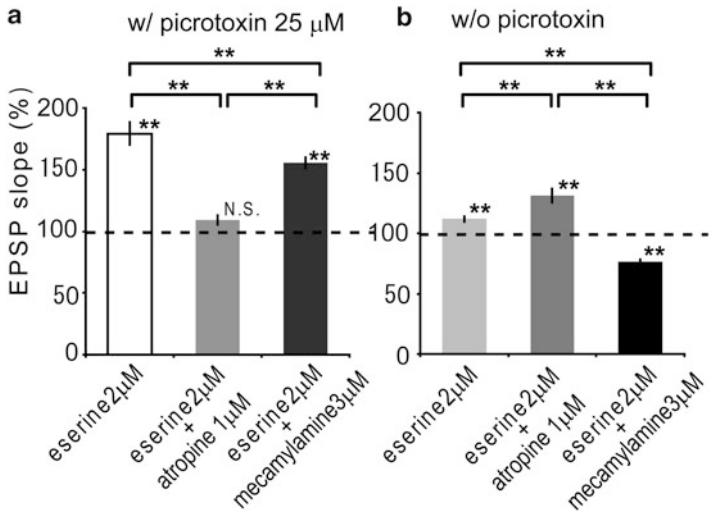


Fig. 4 Eserine effect on AChRs

of atropine (76.5 ± 2.0 %, n = 5, $P < 0.01$, $P < 0.01$ vs. eserine 2 μM only, $P < 0.01$ vs. eserine 2 μM + atropine 1 μM). Therefore cholinergically-induced STDP was regulated by AChRs on interneurons. Taken together, these findings show that the activation of mAChRs contributed more effectively than nAChRs not only on pyramidal neurons but also on interneurons.

4 Conclusion

In this study, to investigate the influence of endogenous ACh on synaptic plasticity in hippocampal CA1 neurons, STDP-inducing stimulus consisting of the evoked firing activities of neurons was injected. In order to influence ACh effect, cholinesterase was blocked by eserine application to prevent ACh from breaking down. It is generally known that LTP induction evoked by STDP protocols depends on the large Ca^{2+} influx through NMDARs in the hippocampus [8], while LTD is induced by a moderate and sustained increase in Ca^{2+} levels [9]. In addition, the responses to NMDAR are selectively potentiated by ACh in rat hippocampal CA1 neurons [10], and also our previous results showed NMDAR response was enhanced by eserine application. These mechanisms, the increase in Ca^{2+} influx through NMDAR by the activation of AChRs, support our results that the cholinergically-induced STDP was enhanced in the presence of eserine under interneuron activated CA1 network (Fig. 2). Furthermore, it is considered that the sign (LTP or LTD) and the amplitude of the cholinergically induced STDP in this study was decided depending on the amount of Ca^{2+} influx as the extended BCM rule that Nishiyama et al. have

reported [11, 2]. We demonstrated in the previous study that STDP was enhanced by the activation of mAChRs on pyramidal neurons in CA1 area. Furthermore, nAChRs on pyramidal neurons facilitated LTP [7]. On the other hand, the majority of interneurons of all layers of CA1 showed depolarization when exposed to muscarinic agonist [12]. These reports support our results that the cholinergically induced STDP was regulated by mAChRs and nAChRs not only on pyramidal neurons but also on interneurons (Fig. 4). We conclude that ACh plays a critical role as a modulator for spatial-temporal information processing in the hippocampus, and thus attention modulates learning and memory at cell level.

Acknowledgments This work was supported by the Grants-in-Aid for Scientific Research 20500278 and 21120006 from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

1. Bi GQ, Poo MM (1998) J Neurosci 18: 10464-10472.
2. Nishiyama M, Hong K, Mikoshiba K, Poo MM, Kato k (2000) Nature 408:584-488.
3. Blokland A (1995) Brain Res Brain Res Rev 21:285-300.
4. Nicoll RA (1985) Trends Neurosci 8:533-536.
5. Shinoe T, Matsui M, Taketo MM, Manabe T (2005) J Neurosci 25:11194-11200.
6. Cole AE, Nicoll RA (1984) J Physiol 352:173-188.
7. Fujii S, Ji Z, Morita N, Sumikawa K (1999) Brain Res 30:137-43.
8. Magee JC, Johnston D (1997) Science 275:209-13
9. Shouval HZ, Kalantzis G (2004) J. Neurophysiol. 93:1069-73
10. Markram H, Segal M (1990) Neurosci Lett 113:62-65.
11. Bienenstock EL, Cooper LN, Munro PW (1982) J Neurosci 2:32-48.
12. McQuiston AR, Madison DV (1999) J Neurosci. 19:5703-10.

Bidirectional Wave Propagations Can Improve Loop Finding Time

Kei-Ichi Ueda, Yasumasa Nishiura, Yoko Yamaguchi, and Keiichi Kitajo

Abstract Previous experimental studies have provided evidence that transient large-scale synchronization of neuronal oscillations plays an important role in switching brain states associated with human brain functions. In our previous study, we investigated the behavior of switching between synchronized and desynchronized states induced by inhibitory interactions between groups of neurons, and proposed a continuous model for a loop-finding system, where a closed loop is defined as a phase synchronization of a group of oscillators belonging to corresponding nodes (K-I Ueda, Phys Rev E 87:052920, 2013). In this study, we modify the previous models to improve the loop-finding time. It is demonstrated that bidirectional flows of phase synchronized waves and interaction between them can improve the loop-finding time.

Keywords Large-scale synchronization • Switching brain states • Loop-finding • EEG • Wave propagation

1 Introduction

A growing body of evidence indicates that large-scale synchronization of neural oscillations plays an important role in linking relevant brain regions for human brain functions. Previous experimental electroencephalography (EEG) studies

K.-I. Ueda (✉)
Graduate School of Science and Engineering, University of Toyama, Toyama 930-8555, Japan
e-mail: kueda@sci.u-toyama.ac.jp

Y. Nishiura
WPI-AIMR, Tohoku University, Miyagi 980-8577, Japan

Y. Yamaguchi
Neuroinformatics Japan Center, RIKEN Brain Science Institute, 2-1, Hirosawa, Wako-shi, Saitama 351-0198, Japan

K. Kitajo
Rhythm-based Brain Information Processing Unit, RIKEN BSI-TOYOTA Collaboration Center, 2-1, Hirosawa, Wako-shi, Saitama 351-0198, Japan

PRESTO, Japan Science and Technology Agency (JST), 4-1-8, Honcho, Kawaguchi, Saitama 332-0012, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_39

have demonstrated that large-scale phase synchronizations mediate human brain functions such as visual awareness [1,2], perceptual switching [3], working memory [4], and attention [5]. These results indicate that the transition between different synchrony states is important in dynamically switching brain states. In this article we study coupled oscillatory systems in a practical problem, the loop finding problem, to develop a mathematical theory to elucidate the mechanism of transient behavior of synchronized oscillatory patterns.

In recent years, inspired by adaptive behavior observed in biological systems, continuous models for the pathfinding problem have been proposed. In laboratory experiments, pathfinding algorithms based on self-organization processes have been studied [6–8]. Continuous models for [8] have also been proposed [9].

In a previous study [10], we proposed a system which is capable of finding a loop path in the network. The system can also show self-recovery properties; that is, the system finds other possible paths when an existing path is broken due to the removal of nodes and paths. The principal remaining shortcoming of the model is the increasing rate of the finding time as the network size increases. In [11], it has been found that the pathfinding time can be improved by the interaction of bidirectional propagating waves of synchronized states. In this article, we apply the method to the loop-finding system to improve the finding time.

2 Model

The graph structure we consider in this study has unidirectional edges between vertices (Fig. 1a). We propose a model that is capable of finding a loop in the graph. The model is as follows:

$$
\begin{aligned}
\dot{u}^j_{k+} &= f(u^j_{k+}, v^j_{k+})/\tau_j + \mu_1 F(s_{k+}; \bar{s}) + \mu_2 \sum_{l^+=1^+}^{K^+} a_{l^+,k^+} F(s_{l^+}; \bar{s}) + w(t), \\
\dot{v}^j_{k+} &= g(u^j_{k+}, v^j_{k+})/\tau_j \\
&\quad + \mu_3 \sum_{m^+=1^+}^{K^+} \sum_{l^+=1}^{K^+} a_{m^+,l^+} a_{m^+,k^+} F(s_{m^+}; \bar{s}) F(s_{l^+} + \sigma^- s_{l^-}; \gamma \bar{s}),
\end{aligned} \tag{1}
$$

$$
\begin{aligned}
\dot{u}^j_{k-} &= f(u^j_{k-}, v^j_{k-})/\tau_j + \mu_1 F(s_{k-}; \bar{s}) + \mu_2 \sum_{l^-=1^-}^{K^-} a_{l^-,k^-} F(s_{l^-}; \bar{s}) + w(t), \\
\dot{v}^j_{k-} &= g(u^j_{k-}, v^j_{k-})/\tau_j \\
&\quad + \mu_3 \sum_{m^-=1^-}^{K^-} \sum_{l^-=1^-}^{K^-} a_{m^-,l^-} a_{m^-,k^-} F(s_{m^-}; \bar{s}) F(\sigma^+ s_{l^+} + s_{l^-}; \gamma \bar{s}),
\end{aligned} \tag{2}
$$

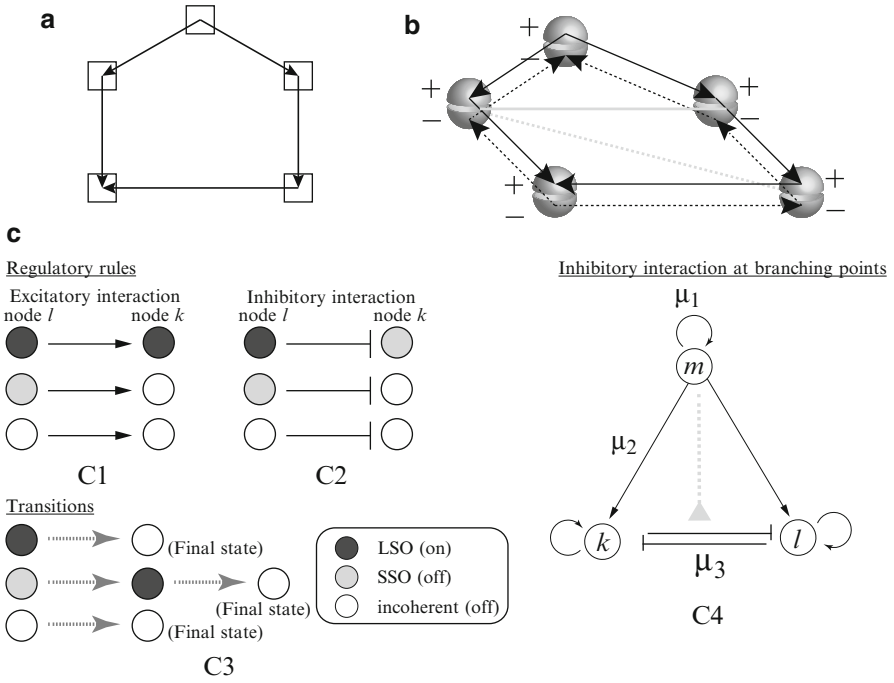


Fig. 1 (**a**) A graph whose vertices (*squares*) are connected by unidirectional edges (*arrows*). (**b**) Network structure corresponding to graph (**a**). The *black solid and dashed arrows* indicate excitatory links corresponding to systems (1) and (2), respectively. The *gray solid and dashed lines* indicate inhibitory links, which are added due to (C4). (**c**) (C1)(C2) Regulatory rules for excitatory and inhibitory interactions. (C3) Transient behavior when a node receives no input. The solution converges to an incoherent state for any initial condition. (C4) Bidirectional inhibitory interactions take place at the branching point. The activity of the inhibitory interactions is also controlled by s_m (*dashed line*)

$s_{k+} = \frac{1}{J}\sum_{j=1}^{J} u_{k+}^{j}$, $s_{k-} = \frac{1}{J}\sum_{j=1}^{J} u_{k-}^{j}$, $f(u,v) = u(1-u)(u-p) - v$, $g(u,v) = \varepsilon(u - qv + r)$, $k^{\pm} = 1^{\pm}, 2^{\pm}, \ldots, K^{\pm}$, $j = 1, 2, \ldots, J$. Here, t is dimensionless time, a dot above a variable indicates a derivative of the variable with respect to t, and μ_i $(i = 1, 2, 3)$ are positive constants. The node number $k^{\pm}$ is responsible for the k-th vertex in the graph. A node consists of a group of oscillators. J and K^{+} (or K^{-}) correspond to the number of elements belonging to each node and the number of nodes, respectively. We set $J = 15$. The state of each node $k^{\pm}$ is determined by the average of the amplitudes of the oscillators $u_j^{k^{\pm}}$ in the corresponding node, for example, $s_{k^{\pm}}$. We assume that the interaction function F has a threshold for activation. Regulation of on–off switching of the connecting nodes depends on s_k, and is simply defined by the Heaviside function with a threshold $\bar{s}$, where $F(s;\bar{s}) = 1$ for $s > \bar{s}$ and $F(s;\bar{s}) = 0$ for $s \leq \bar{s}$. The interactions affect all oscillators uniformly; that is, they are independent of j.

Excitatory interaction directed from node l toward node k is expressed as $a_{l,k}$, where $a_{l,k} = 1$ and $a_{l,k} = 0$ indicate the presence and absence of such interaction, respectively. We take $a_{k^\pm,k^\pm} = 0$ for all $k^\pm$.

The systems (1) and (2) are decoupled by taking $\sigma^+ = \sigma^- = 0$. The former and latter systems have forward and opposite directional excitatory links. That is, if a link from $i^+ \to j^+$ exists for (1) ($a_{i^+,j^+} = 1$), then a link from $j^- \to i^-$ exists for (2) ($a_{j^-,i^-} = 1$). System (1) with $\sigma^+ = 0$ is the same as the system proposed in [10], and system (2) with $\sigma^- = 0$ is the same as the previous system, except for the direction of the links. We refer to system (1) with $\sigma^- = 0$ and system (2) with $\sigma^+ = 0$ as system A and system B, respectively. We refer to the coupled system with $\sigma^\pm = 1$ as system C.

The parameters p, q, r, and ε are independent of $k^\pm$, and set $p = 0.02$, $q = 1.0$, $r = -0.04$, and $\varepsilon = 0.01$. Furthermore, $w(t)$ is a small amount of random noise in the interval $[0, 0.05]$. The time constants τ_j take random values from the interval between τ^- and τ^+, where the values of τ^- and τ^+ are set to $(\tau^-, \tau^+) = (6.0, 6.5)$. The distribution of τ_j is the same for all nodes; that is, it is independent of k.

Each node can have one of three possible states: synchronized oscillation with a large amplitude, synchronized oscillation with a small amplitude, or incoherent oscillation. We respectively refer to these states as large-amplitude synchronized oscillation (LSO), small-amplitude synchronized oscillation (SSO), and incoherent oscillation. SSO is observed due to inhibitory interactions, which are uniformly applied to all oscillators. To realize SSO, we employ the FitzHugh–Nagumo equation, which is a typical equation exhibiting a small-amplitude relaxation oscillation.

The parameter values μ_i are set so that the regulatory rules and the conditions for C1, C2, and C3 in Fig. 1 are satisfied. Parameter μ_2 is taken to be sufficiently large to ensure that phase resetting propagates along the loops, which allows the connected node to enter LSO (C1). Parameter μ_3 is taken to be sufficiently large so that SSO is observed due to the inhibitory interaction (C2). Parameter μ_1 is taken to be sufficiently close to but not at the limit of LSO, which enables the node to undergo a transition SSO $\to$ LSO $\to$ incoherent oscillation (C3). The threshold $\bar{s}$ is taken to be sufficiently large to ensure that the values of s_k are smaller than $\bar{s}$, and the state of the node converges to incoherent oscillation when no input is received (C3). The parameter values are set to $(\mu_1, \mu_2, \mu_3) = (1.8 \times 10^{-3}, 0.06, 0.07)$ and $\bar{s} = 0.825$. Inhibitory interaction takes place at branching points (C4), implemented in $a_{m^\pm,l^\pm} \times a_{m^\pm,k^\pm}$ in the $v_{k^\pm}$-equation.

3 Results

We examine numerical experiments using network L and network R, shown in Fig. 2a. Each network has three possible loops. Here, a loop is defined as a phase synchronization (or LSO) belonging to the corresponding nodes. As discussed in [10], it is observed that LSO waves propagate across nodes, and that systems A and B can autonomously find one of the possible loops. The systems also find a

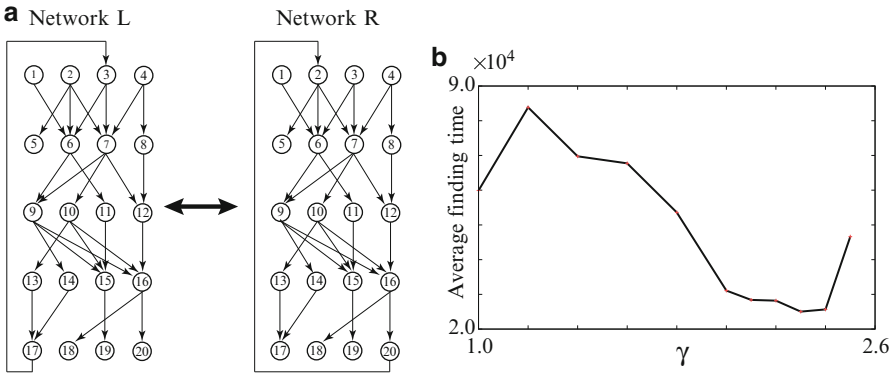


Fig. 2 (**a**) The network alternates between network L and network R, changing when the system has succeeded in finding one of the possible loops. *Arrows* indicate the forward direction of links. (**b**) The average finding time when γ is taken as a control parameter

new loop when the network structures are changed between networks L and R. It is confirmed that system C, too, can have the properties mentioned above. Due to the mutual interactions between (1) and (2), it is found that both node k^+ and k^- are in LSO when system C has succeeded in finding a loop, and the nodes $k^\pm$ correspond to the loop.

To compare the finding time between systems A, B, and C, we measured average finding times $\bar{T}_A$, $\bar{T}_B$, and $\bar{T}_C$; $\bar{T}_A := (\bar{T}_{A,L} + \bar{T}_{A.R})/2$, where $\bar{T}_{A,L}$ ($\bar{T}_{A,R}$) indicates the averaged time to find one of the possible loops in network L (network R) in Fig. 2a. During computations, the network structures are alternatively switched between the network L and R when the corresponding system has succeeded in finding a loop. $\bar{T}_{B,L}$, $\bar{T}_{B,R}$, $\bar{T}_{C,L}$, and $\bar{T}_{C,R}$ are defined similarly. The total switching number is 50 for each system, and $\gamma = 1$ is fixed for systems A and B.

It is found that $\bar{T}_A \approx 3.54 \times 10^4$ and $\bar{T}_B \approx 5.62 \times 10^4$, where the difference in the averaged times comes from the difference in position of inhibitory links. The averaged time $\bar{T}_C$ can be improved by taking γ as a control parameter. In fact, the finding time is improved by increasing γ from 1.2, and has a minimum value around $\gamma = 2.3$ ($\bar{T}_C \approx 2.50 \times 10^4$) (Fig. 2b). No loop can be found for $\gamma \geq 2.5$. Note that the finding time is smaller than that of systems A and B; $\bar{T}_C \leq \min\{\bar{T}_A, \bar{T}_B\}$. This means that the bidirectional wave propagations improves the loop finding time.

4 Discussion

The finding process of the present system fundamentally owes to that proposed in the previous study [10]. In this study, we have considered the effect of interactions of bidirectional flows of wave propagation of the LSO state. We confirmed that the finding time can be improved by taking the threshold of the inhibitory interaction functions in an appropriate regime.

Sharp transition behavior from LSO to an incoherent state, which is needed for (C3), can be generated due to the existence of the limiting point of LSO. To generate the limiting point, we employed the step function as an interaction function. In fact, when we employ $F(x;\bar{s}) = 1/(1+\exp(-\gamma(x-\bar{s})))$ as a interaction function, we observe a sharp transition from LSO to incoherent oscillation when γ is sufficiently large. Therefore, in this study, we take $\gamma = \infty$, the Heaviside function.

Acknowledgements This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas "Neural creativity for communication" (No. 4103) (21120003 and 21120005) of MEXT, Japan and a Grant-in-Aid for Scientific Research (C) (25400199) of JSPS, Japan.

References

1. Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B., Varela, F. J.: Perception's shadow: long-distance synchronization of human brain activity. Nature **397**, 430–433 (1999).
2. Kitajo, K., Doesburg, S. M., Yamanaka, K., Nozaki, D., Ward, L. M., and Yamamoto, Y.: Noise-induced large-scale phase synchronization of human brain activity associated with behavioural stochastic resonance. Europhys. Lett. **80**, 40009 (2007).
3. Shimaoka, D., Kitajo, K., Kaneko, K., Yamaguchi, Y.: Transient process of cortical activity during Necker cube perception: from local clusters to global synchrony. Nonlinear Biomedical Physics, **4** (Suppl. 1), S7 (2010).
4. Kawasaki, M., Kitajo, K., and Yamaguchi, Y.: Dynamic links between theta executive functions and alpha storage buffers in auditory and visual working memory. Eur. J. Neurol. **31**, 1683–1689 (2010).
5. Doesburg, S. M., Roggeveen, A. B., Kitajo, K., Ward, L. M.: Large-scale gamma-band phase synchronization and selective attention. Cerebral Cortex, **18**, 386–396 (2008).
6. Oliver, S., Tóth, Á, Showalter, K.: Navigating complex labyrinths: optimal paths from chemical waves, Science, 868–868 (1995).
7. Marco, D., Di Caro, G., Gambardella, L. M.: Ant algorithms for discrete optimization. Artificial life **5** 137–172 (1999).
8. Nakagaki, T., Yamada, H., Tóth, Á.: Intelligence: Maze-solving by an amoeboid organism, Nature **407** (6803), 470 (2000).
9. Tero, A., Kobayashi, R., Nakagaki T.: j A mathematical model for adaptive transport network in path finding by true slime mold, J. theo. biol. **244**(4) 553–564 (2007).
10. Ueda, K.-I.:Three-state network design for robust loop-searching systems, Phys. Rev. E **87**, 052920 (2013).
11. Ueda, K.-I., Nishiura, Y., Yamaguchi, Y., Kitajo, K., arXiv:1309.2845.

Phase Coupling Between Hippocampal CA1 and Prefrontal Cortex in a Depression-Model Rats Indicating Impaired Synaptic Plasticity

Chenguang Zheng, Zhuo Yang, and Tao Zhang

Abstract Chronic stress induces profound learning and memory deficits with a potential mechanism, impaired synaptic plasticity in hippocampus (HPC) and medial prefrontal cortex (mPFC). We propose a hypothesis that the theta-rhythm coupling in HPC-mPFC circuit is associated with the synaptic plasticity as an underlying mechanism of the cognitive dysfunction in depression.

Keywords Hippocampus • Prefrontal cortex • Local field potentials • Depression • Phase coupling

1 Introduction

Chronic stress induces profound learning and memory deficits in behavior of humans and rodents, paralleled by impaired synaptic plasticity in hippocampus (HPC) and medial prefrontal cortex (mPFC) as a potential mechanism [1, 2]. Furthermore, theta-frequency synchronization has been measured between the hippocampus and downstream targets to demonstrate their cooperation during a variety of behaviors and in several psychiatric disorders. The synchrony of theta rhythm has been shown between the HPC and mPFC during working memory [3, 4] and its impairment during anxiety and schizophrenia [5, 6]. In our previous studies, theta coupling was decreased in depression partly associated with the impaired long-term potentiation in thalamocortical pathway [7, 8]. Therefore, we propose a hypothesis here that whether the theta-rhythm coupling in HPC-mPFC circuit indicates the synaptic plasticity as an underlying mechanism of the cognitive dysfunction in depression.

C. Zheng • T. Zhang (✉)
College of Life Sciences, Nankai University, Tianjin 300071, People's Republic of China
e-mail: zhangtao@nankai.edu.cn

Z. Yang
College of Medicine Science, Nankai University, Tianjin 300071, People's Republic of China

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_40

2 Method/Models

2.1 Animals and Treatments

All the experiments were carried out according to the guidelines of the Beijing Laboratory Animal Center, and approved by the Ethical Commission at Nankai University.

Twelve Adult male Wistar rats (180–220 g body weight at the beginning of the experiment) were reared under standard laboratory conditions (lights on from 7:00 A.M. to 7:00 P.M.; room temperature 24 °C ±2 °C). Food and water were supplied during all phases of the experiments. A group of six rats were handled daily and served as controls (Con group). Another group of six rats (Dep group) were submitted to 3 weeks of depression animal model establishment, the chronic unpredictable stress (CUS) model [9].

2.2 Electrophysiological Experiments

The rats of both Con and Dep groups were placed in a stereotaxic frame under 30 % urethane anesthesia (3.5 ml/kg, i.p., Sigma-Aldrich, St. Louis, MO, U.S.A.). The recording electrode was placed in the prelimbic area (PrL) of PFC (AP 3–3.3 mm, ML 0.8–1 mm, H 3.0–3.8 mm) and a bipolar stainless-steel stimulating electrode was positioned into the CA1 area of the ventral hippocampus (AP -6.3–6.5 mm, ML 5.5 mm, H 4.9–6.0 mm). The electrode positions were determined finally via inducing a characteristic monosynaptic negative going field excitatory postsynaptic potential (fEPSP) in the PrL by the stimulation of CA1 region.

Spontaneous local field potential (LFP) were collected first for 20 min from the CA1 region and PrL respectively. The LFP signals were fed into a multi-channel differential amplifier and acquired simultaneously at 1,000 Hz (bandpass filtered from 0.3 to 200 Hz). After that, the LTP was induced in CA1-PrL circuit. After a 30 min baseline, high-frequency stimulation to induce LTP consisted of two series of ten trains (250 Hz, 200 ms) at 0.1 Hz, 6 min apart, delivered at test intensity. Following the LTP induction for 1 h, the LFPs were recorded for another 20 min at both CA1 and PrL regions.

2.3 Power Spectra and Phase Coupling Analysis

Power spectra of LFPs were computed using the multitaper method with the Chronux codes described elsewhere (http://www.chronux.org) [10]. An optimal family of orthogonal tapers were used, parameterized with time-bandwidth product $TW = 5$ and $k = 9$ tapers.

The EMA algorithm was developed as a measurement for direction of information flow [11, 12]. Its theory and algorithm were described in detail in

the previous studies [8]. Coefficients c could be employed to measure unidirectional information transfer. For the phase coupling analysis, the original LFPs signals were band-pass filtered into 1–8 Hz (delta + theta band), 3–8 Hz (theta band) and 30–100 Hz (gamma band), with a two-way least squares FIR filter (eegfilt.m from the EEGLAB toolbox, bandwidth = 5 Hz). Hilbert transform was used to obtain the instantaneous phase of neural signals approximately from the signals filtered in the frequency bands.

2.4 Data and Statistical Analysis

All the data were expressed as mean ± SEM. For the slopes of the evoked fEPSPs along with time, a two-way repeated measure ANOVA was used. In order to compare the power spectra and coupling indices of LFPs between pre- and post- LTP (in pair), Wilcoxon signed-rank test was applied in both groups. Student's t-test was used to compare the power and coupling indices between Con and Dep groups. All the statistical analyses were performed using SPSS 18.0 software and the significant level was set at 0.05.

3 Results

3.1 LTP Induction

In Con group, a lasting increase in the slope of the fEPSPs in the PrL induced by HFS in CA1. However, the LTP was robustly impaired in the PFC when the rats were exposed to CUS treatment. Two-way repeated measures ANOVA confirmed that there was significant difference on the time course of normalized fEPSPs between Con and Dep groups ($F_{(1,10)} = 21.283$, $p = 0.001$), whereas there was no statistical differences of time or time × groups interaction ($p > 0.05$) (Fig. 1a). Exposure to CUS significantly impaired LTP in the PrL for the last 20 min of recording after HFS, tested by Student's t-test (Fig. 1b). ($p = 0.005$).

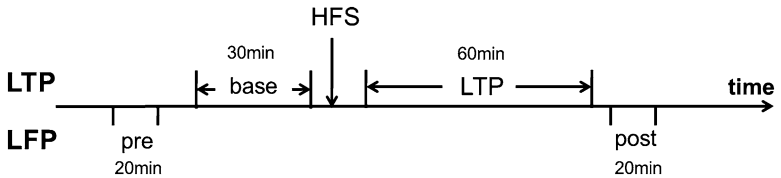


Fig. 1 The effects of CUS treatment on fEPSP slopes of LTP. (**a**) Time course of fEPSP slopes evoked in the prelimbic area (PrL) in medial prefrontal cortex (mPFC) by stimulation of the CA1 region. Each point represents mean ± SEM of averaged evoked responses for 2 min epochs. High frequency stimulation is indicated by *arrows*. (**b**) Mean fEPSP slopes for the last 20 min of recording after HFS (**p < 0.01)

3.2 *Power Spectra Analysis*

Wilcoxon signed-rank test showed that there was a significant increment of the mean power after LTP at both CA1 and PrL regions (in pair test for CA1 of Con group: $p = 0.046$; CA1 of Dep group: $p = 0.116$; PrL of Con group: $p = 0.028$; PrL of Dep group: $p = 0.046$, Fig. 2). On the other hand, Student's *t*-test was used to determine the influence of CUS treatment on the LFPs for pre-LTP. The data showed that only the mean power of LFPs at PrL increased in Dep group ($p = 0.046$), however the mean power of LFPs at CA1 didn't change ($p = 0.716$).

The power spectra in delta (1–3 Hz), theta (3–8 Hz) and gamma (30–100 Hz) were extracted, and the relative power percent of these three frequency bands were obtained by $\frac{\text{rhythm power}}{\text{sum of power}} \times 100\%$ (Fig. 3). Wilcoxon test showed that the relative power

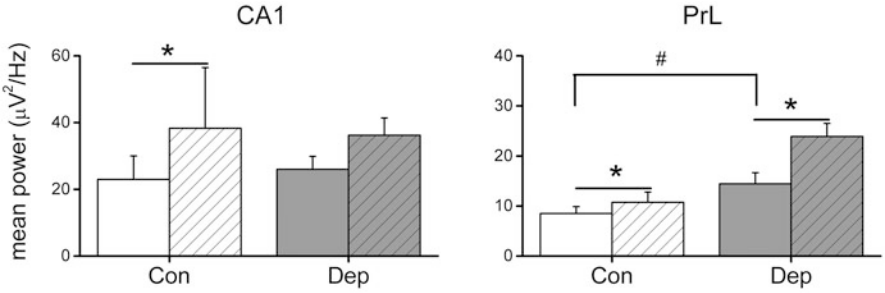


Fig. 2 Mean power spectra in 1–100 Hz of LFPs from CA1 and PrL between pre- and post LTP induction in both two groups $*p < 0.05$ between pre- and post LTP by Wilcoxon test, and $\#p < 0.05$, $\#\#p < 0.01$, $\#\#\#p < 0.001$ between two groups by Student's *t*-test

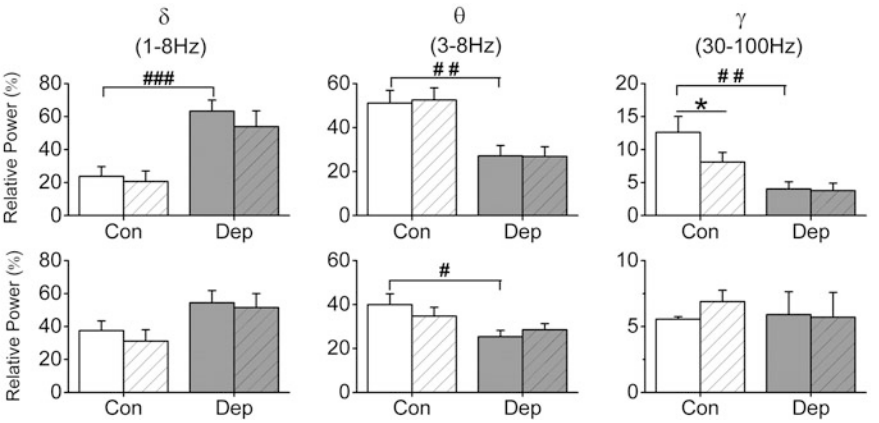


Fig. 3 Relative power percent (%) on delta, theta and gamma rhythms between pre- and post-LTP at CA1 (*upper panel*) and PrL (*lower panel*) regions in both groups. $*p < 0.05$ between pre- and post LTP by Wilcoxon test, and $\#p < 0.05$, $\#\#p < 0.01$ between two groups by *t*-test

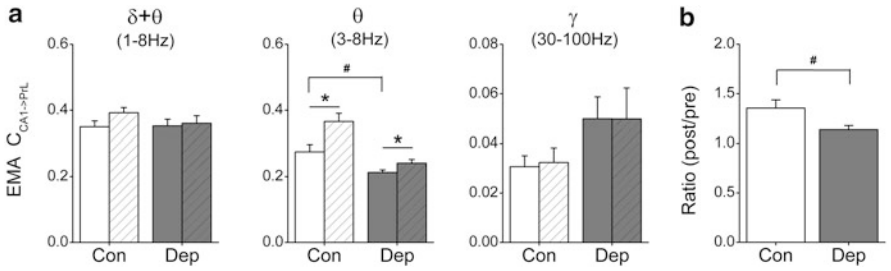


Fig. 4 Phase coupling on CA1-PrL circuit. (**a**) Unidirectional coupling index $c_{\mathrm{CA1}\rightarrow\mathrm{PrL}}$ on delta, theta and gamma rhythms between pre- and post-LTP in both groups. (**b**) EMA ratio of post/pre LTP on theta rhythm in both groups. $^{*}p < 0.05$ between pre- and post LTP by Wilcoxon test, and $\#p < 0.05$ between two groups by Student's *t*-test

percent didn't increase after LTP in all the frequency bands at both CA1 and PrL regions, especially for the gamma rhythm in CA1 which was reduced significantly ($p = 0.028$). As for the comparison between groups, it can be seen that there is an increased tendency only on delta rhythm (CA1 delta: $p = 0.001$) and a decreased tendency on theta and gamma rhythms of rats in depression (CA1 theta: $p = 0.009$; PrL theta: $p = 0.028$; CA1 gamma: $p = 0.009$).

3.3 Unidirectional Phase Coupling

The results implied that on theta rhythm the unidirectional index $c_{\mathrm{CA1}\rightarrow\mathrm{PrL}}$ of pre-LTP reduced significantly in Dep group ($p = 0.024$). However, the $c_{\mathrm{CA1}\rightarrow\mathrm{PrL}}$ index on theta rhythm was increased significantly after LTP induction in both Con ($p = 0.028$) and Dep ($p = 0.046$) groups, measured by Wilcoxon test. Also, the relative increment (post/pre ratio) of $c_{\mathrm{CA1}\rightarrow\mathrm{PrL}}$ index in Con groups was higher than that in Dep group ($p = 0.043$) (Fig. 4b).

4 Discussion

In this study, we found that the integral power spectra were increased in depression rats, whereas the relative percentage of theta and gamma reduced, partly consistent with our previous data [13, 14]. The phase coupling in theta rhythm on CA1-PrL circuit was significant lower in depressed rats, which indicated the impaired synchronization at baseline, i.e. pre-LTP, by the exposure to CUS. This result was in line with impaired synchrony of neural activity between HPC and mPFC in other psychiatric disorders such as schizophrenia and anxiety [5, 6]. Importantly, after LTP induction, the theta-frequency coupling was enhanced for both normal and

depressed rats. However, the power ratio didn't change between pre- and post-LTP. It suggested that, the phase coupling in theta rhythm are implicated in synaptic strength and could be maintained for long time course. Furthermore, the relative increment of CA1-PrL coupling strength after LTP induction was higher in Con group than in Dep group, which was strongly consistent with the LTP data. All these results supported our hypothesis that the theta-frequency phase coupling was implicated in HPC-mPFC synaptic plasticity and could further indicate the cognitive level in depression disorder.

Acknowledgments This work was supported by grants from the National Natural Science Foundation of China (31171053, 11232005) and Tianjin research program of application foundation and advanced technology (12JCZDJC22300).

References

1. J.J. Cerqueira, F. Mailliet, O.F. Almeida, T.M. Jay, N. Sousa, J Neurosci 27 (2007) 2781–2787.
2. J.J. Kim, D.M. Diamond, Nat Rev Neurosci 3 (2002) 453–462.
3. S. Fujisawa, G. Buzsaki, Neuron 72 (2011) 153–165.
4. J.A. Gordon, Curr Opin Neurobiol 21 (2011) 486–491.
5. T. Sigurdsson, K.L. Stark, M. Karayiorgou, J.A. Gogos, J.A. Gordon, Nature 464 (2010) 763–767.
6. A. Adhikari, M.A. Topiwala, J.A. Gordon, Neuron 65 (2010) 257–269.
7. C. Zheng, M. Quan, Z. Yang, T. Zhang, Neurosci Lett 490 (2011) 52–56.
8. C. Zheng, M. Quan, T. Zhang, J Comput Neurosci 33 (2012) 547–558.
9. P. Willner, Psychopharmacology (Berl) 134 (1997) 319–329.
10. P. Mitra, H. Bokil, Observed brain dynamics, Oxford University Press, USA, 2008.
11. M.G. Rosenblum, A.S. Pikovsky, Phys Rev E 64 (2001) 045202.
12. M.G. Rosenblum, L. Cimponeriu, A. Bezerianos, A. Patzak, R. Mrowka, Phys Rev E Stat Nonlin Soft Matter Phys 65 (2002) 041909.
13. C. Zheng, T. Zhang, Cogn Neurodyn 7(2) (2013) 167-172.
14. X. Xu, C. Zheng, T. Zhang, Frontiers Comput Neurosci (2013) Volume 7: Article 27.

Part VII
Neural Computation and Information Processing

Mapping of Cortical Avalanches to the Striatum

Jovana J. Belić, Andreas Klaus, Dietmar Plenz, and Jeanette Hellgren Kotaleski

Abstract Neuronal avalanches are found in the resting state activity of the mammalian cortex. Here we studied whether and how cortical avalanches are mapped onto the striatal circuitry, the first stage of the basal ganglia. We first demonstrate using organotypic cortex-striatum-substantia nigra cultures from rat that indeed striatal neurons respond to cortical avalanches originating in superficial layers. We simultaneously recorded spontaneous local field potentials (LFPs) in the cortical and striatal tissue using high-density microelectrode arrays. In the cortex, spontaneous neuronal avalanches were characterized by intermittent spatiotemporal activity clusters with a cluster size distribution that followed a power law with exponent −1.5. In the striatum, intermittent spatiotemporal activity was found to correlate with cortical avalanches. However, striatal negative LFP peaks (nLFPs) did not show avalanche signatures, but formed a cluster size distribution that had a much steeper drop-off, i.e., lacked large spatial clusters that are commonly expected for avalanche dynamics. The underlying de-correlation of striatal activity could have its origin in the striatum through local inhibition and/or could result from a particular mapping in the corticostriatal pathway. Here we show, using modeling, that highly convergent corticostriatal projections can map spatially extended cortical activity into spatially restricted striatal regimes.

Keywords Neuronal avalanches • Striatum • Cortico-striatal network • Cortex • Basal ganglia

J.J. Belić (✉)
School of Computer Science and Communication, KTH Royal Institute of Technology, 10044 Stockholm, Sweden

Bernstein Center Freiburg, University of Freiburg, Freiburg 79104, Germany
e-mail: belic@kth.se

A. Klaus • D. Plenz
Section of Critical Brain Dynamics, National Institute of Mental Health, Bethesda, MD, USA

J. Hellgren Kotaleski
School of Computer Science and Communication, KTH Royal Institute of Technology, 10044 Stockholm, Sweden

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_41

1 Introduction

Scaling laws are ubiquitous in nature. Earthquakes, landslides, and forest fires are examples for systems in which local events can propagate over long distances forming extended cascades. Cascade sizes are described by power laws that express one variable as a function of another raised to a power, $f(x) \sim x^{\alpha} (\alpha \neq 0)$ [1]. An interesting feature of power laws is that they show no characteristic scale, and when plotted in log-log coordinates they produce a straight line with a characteristic slope (α) for that system.

Power law distributions of neuronal activity at the network level were found in vitro by recording local field potentials in cortical neural networks using slices of rat cortex as well as cultured networks [2]. It was observed that the number of electrodes activated over certain threshold during each single burst of spontaneous neuronal activity was distributed according to a power law with a characteristic exponent of -1.5 and those events were called neuronal avalanches. Since then, neuronal avalanches were reported in vivo [3] and were shown to display long-term stability, diversity and fast propagation of local synchrony [4]. Neuronal avalanches also characterize networks that have a maximum dynamic range [5], maximal variability of phase synchrony [6], and avalanches might play a central role in information transmission and storage [7].

Here we study neuronal avalanches in an open-loop system of the cortex and striatum. The input region of the basal ganglia (striatum) plays an important role in reward based learning and control of actions [8]. The striatum is a major recipient of massive glutamatergic inputs from the cerebral cortex and thalamus. Understanding how the striatum responds to cortical inputs has crucial importance for clarifying the overall functions of the basal ganglia. Medium spiny neurons (MSNs) dominate in the striatum (up to 95 % in rodents). They are inhibitory (GABAergic) and have membrane properties that give them a high threshold for activation [9].

In this paper, we investigate how the striatum responds to cortical neuronal avalanches. We developed an abstract cortical model that reproduces statistics observed in experimental data [4, 10, 14]. The model is giving values for entropy and average mutual information that are in best agreement with experimental data compared to the alternative models [5, 7] with all-to-all or random connectivity, and where connectivity values were chosen from a uniform distribution. After that we extended the model in order to determine which kind of connectivity between cortex and striatum could account for the experimental observation of a steeper slope of the clustered activity in the striatum compared to the cortex [13, 14].

2 Method

Organotypic Cultures on Microelectrode Arrays All animal procedures were in accordance with National Institutes of Health guidelines. Coronal slices from rat cortex (350 μm thick), striatum (500 μm thick), and midbrain (substantia nigra pars

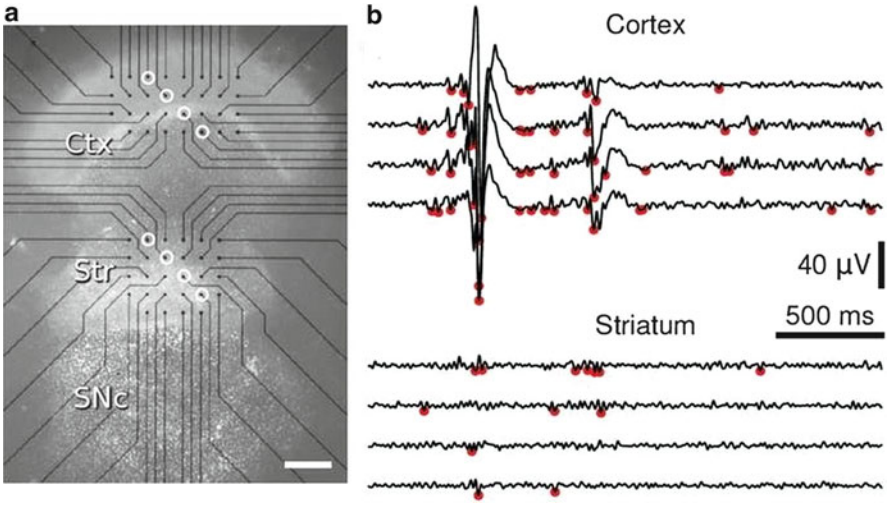


Fig. 1 (**a**) Organotypic cortex-striatum-substantia nigra pars compacta culture. The *white circles* indicate the electrodes for which representative LFP traces are shown in (**b**). (**b**) Example of LFP activity for four cortical and striatal electrodes. *Red dots* indicate significant negative LFP peaks

compacta; 500 µm thick) were cut on a vibratome in sterile Gey's balanced salt solution and cultured on planar, 60-channel microelectrode arrays (MEA) for the recording of local field activity. For all MEA recordings a custom layout with two sub-arrays for cortex (31 electrodes) and striatum (28 electrodes) was used (Fig. 1a).

Local Field Potential Analysis Local field potential (LFP) activity was recorded at (or down-sampled to) 1 kHz, and subsequently band-pass filtered at 1–50 Hz. Negative LFP (nLFP) deflections were detected by finding the minimum value of the LFP signal that crossed a threshold of $z = -4.5$ standard deviations (SDs) (Fig. 1b). Spatiotemporal clusters were detected as in previous work [2].

Computational Model Our abstract model of the cortical network was designed in order to reproduce some of the main statistics observed in experimental data [10]. First, each of the N nodes was assigned a number that determined the out-degree of each of them. The average number of connectivity was chosen to be 10 [10], and afterwards we applied the preferential attachment rule. Each node was attached to other nodes in proportion to the out-degrees of those nodes in order to get the node degree linearly related to the average node strength for both in- and out-degrees (Fig. 2a). Transmission probabilities p_{ij} (from node j to node i) were picked from an exponential distribution (Fig. 2b). There were no self-connections and no more than one connection between attached pairs of nodes. The weights were then scaled (p'_{ij}) such that the branching parameter for the entire network was equal to 1. The probability that node i fired at time $t + 1$ was:

$$p_{ij} = \left(1 - \prod\nolimits_{j \in j(\tau)} \left(1 - p'_{ij}\right)\right), \tag{1}$$

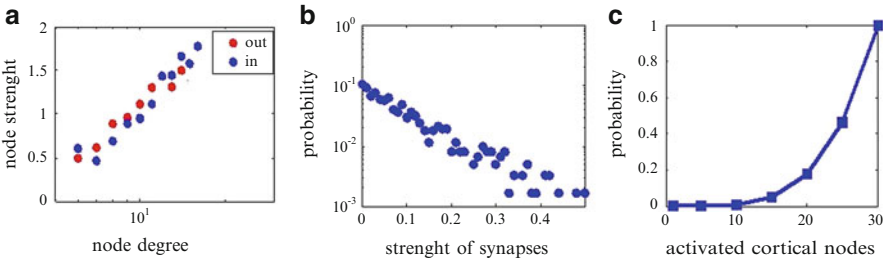


Fig. 2 (**a**) Node degree is linearly related to the average node strength for both in and out degrees. (**b**) Plot of link weight probability. (**c**) Probability of activation for striatal nodes depending on the cortical activity

where $J(t)$ was a set of nodes that fired at time t. Node i actually fired at time $t + 1$ only if $p_{iJ} > \epsilon$, where ϵ was a random number from a uniform distribution on [0, 1]. Then, each of the $N = 30$ nodes in the striatum was randomly connected with a certain number ($N_k = 4$ here) of nodes in the cortex. A node in striatum always fired if the full pattern in cortex assigned to that node was present; otherwise it fired with a very low probability (Fig. 2c). In a first approach, we focus on the contribution of the corticostriatal pathway to explain the experimental results while ignoring intrastriatal inhibition. Our future work will explore the role of the intrastriatal GABAergic network.

3 Results

We recorded spontaneous LFP activity in organotypic cortex-substantia nigra cultures simultaneously from 31 electrodes in the cortex to 28 electrodes in the striatum [12, 13]. Negative LFPs were detected by applying a negative threshold to each electrode. Analyses of the cluster size distributions in the striatum pointed to a more negative exponent than the one obtained for the cortex [2, 13] (Fig. 3). The steeper slope of the striatal cluster size distributions indicates reduced spatiotemporal correlations.

In order to investigate how the observed striatal dynamics could be explained by the corticostriatal connectivity, we developed an abstract computational model (see section "Method"). We stimulated population events by randomly choosing and triggering a single node (this procedure was repeated 30,000 times for each trial) in the cortex, while simultaneously monitoring the resulting activity in the striatum. When we gave to each striatal node the probability of activation presented in Fig. 2c, we got a steeper slope of the striatal cluster size distribution, similar to experimental results (Fig. 4).

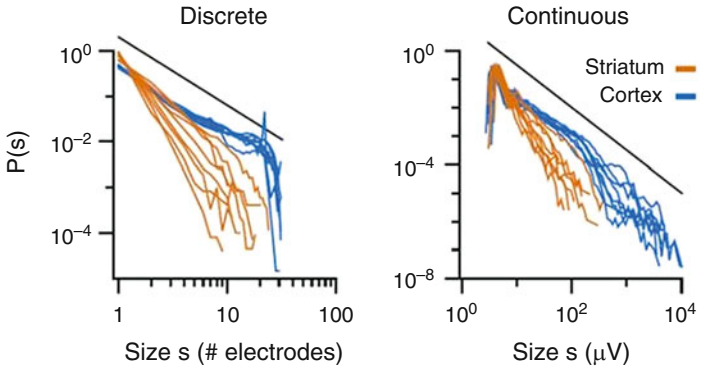


Fig. 3 Experimentally measured discrete (*left panel*) and continuous (*right panel*) cluster size distributions for cortex and striatum. *Black line* indicates a power law with $\alpha = -1.5$ for comparison

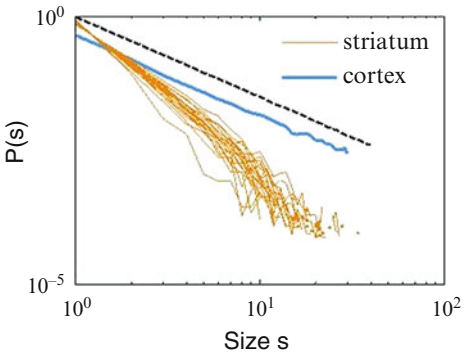


Fig. 4 Cluster size distributions for cortex and striatum in the model. *Black line* indicates a power law with $\alpha = -1.5$ for comparison

Table 1 Comparison of experimental results with results produced by different models (mean ± SD). Entropy and mutual information were calculated for N = 60 in the same way as published previously [7]

	Entropy H (bits)	Site-to-site MI (bits)
Experimental data [7]	5.7 ± 1.6	0.2 ± 0.2
Our model	5.8 ± 1.5	0.24 ± 0.06
All to all connectivity	6.5 ± 0.3	0.15 ± 0.01
Random connectivity	6.2 ± 1.3	0.14 ± 0.03

Our cortical model has shown the best agreement with experimental data regarding values for entropy and site-to-site mutual information [7] compared to models where all-to-all or random connectivity was assumed (average connectivity was set to 10) and strengths of connections were picked from a uniform distribution (Table 1). We also found bigger deviations from experimental data in those cases when the connection strengths in our model were chosen from uniform distribution

instead of exponential, and when the branching parameter for each node (sum of connection strengths of a particular node to other nodes) was set to one. By increasing the connectivity we observed a decrease of site-to-site mutual information and an increase of entropy. Parts of these results were previously published in the form of abstracts [11, 12].

4 Conclusion

In the present study, we focused on how the corticostriatal pathway might contribute to striatal processing of cortical avalanches. Experimentally, we demonstrated that striatal cluster size distributions were characterized by a more negative power law exponent as compared to that for cortical avalanches. We developed a corticostriatal model in which cortical activity was mapped on striatal neurons through convergent projections. This mapping resulted in striatal responses to cortical avalanches in line with our experimental findings. Specifically, under the assumption of a particular high activation threshold of striatal nodes we can reproduce power law like distributions with a coefficient similar to the one found experimentally. Future work is exploring the role of intrastriatal inhibition in shaping striatal responses to cortical avalanches.

References

1. Bak, P.: How nature works: The science of self-organized criticality. New York, NY: Copernicus Press (1996).
2. Beggs, M., Plenz, D.: Neuronal avalanches in neocortical circuits. J. Neurosci. 23 (2003) 11167–11177.
3. Gireesh, E., Plenz, D.: Neuronal avalanches organize as nested theta and beta/gamma-oscillations during development of cortical layer 2/3. Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 7576–7581.
4. Beggs, J., Plenz, D.: Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures. J. Neurosci. 24 (2004) 5216–5229.
5. Shew, W., Yang, H., Petermann, T., Roy, R., Plenz, D.: Neuronal avalanches imply maximum dynamic range in cortical networks at criticality. J. Neurosci. 29 (2009) 15595–15600.
6. Yang, H., Shew, W., Roy, R., Plenz, D.: Maximal variability of phase synchrony in cortical networks with neuronal avalanches. J. Neurosci. 32 (2012) 1061–1072.
7. Shew W., Yang, H., Yu, S., Roy, R., Plenz, D.: Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches. J. Neurosci. 31 (2011) 55–63.
8. Grillner S., Hellgren Kotaleski, J., Menard, A., Saitoh, K., Wikstrom, M.: Mechanisms for selection of basic motor programs-roles for the striatum and pallidum. Trends in Neurosciences. 28 (2005) 364–370.
9. Tepper, J., Koos, T., Wilson, C.: GABAergic microcircuits in the neostriatum. Trends in Neurosciences. 27 (2004) 662–669.
10. Pajevic, S., Plenz, D.: Efficient network reconstruction from dynamical cascades identifies small-world topology of neuronal avalanches. Plos Computational Biology. 5 (2009) e1000271.

11. Belić, J., Klaus, A., Plenz, D., Hellgren Kotaleski, J.: Neuronal avalanches and the cortico/striatal network. BMC Neuroscience 13 (2012) P1.
12. Klaus, A., Plenz, D.: Striatal processing of cortical neuronal avalanches. Neuroscience 2012, New Orleans, LA: Society for Neuroscience.
13. Klaus, A. The role of striatal inhibition in the processing of cortical neuronal avalanches. http://hdl.handle.net/10616/41351, Doctoral dissertation (2013).
14. Klaus, A., Yu, S., & Plenz, D. (2011). Statistical analyses support power law distributions found in neuronal avalanches. PLoS ONE, 6(5), e19779.

Gauss-Markov Processes for Neuronal Models Including Reversal Potentials

Aniello Buonocore, Luigia Caputo, Amelia G. Nobile, and Enrica Pirozzi

Abstract Gauss-Markov processes, restricted from below by a reflecting boundary, are here used to construct inhomogeneous leaky integrate-and-fire (LIF) stochastic models for single neuron's activity in the presence of a reversal hyperpolarization potential and different input signals. Under suitable assumptions, we are able to obtain the transition probability density function with a view to determine numeric, simulated and asymptotic solutions for the firing densities when the input signal is constant, decays exponentially or is a periodic function. The our results suggest the importance of the position of the lower boundary as well as that of the firing threshold when one studies the statistical properties of LIF neuron models.

Keywords LIF • Reversal hyperpolarization potential • Transition probability density function • Firing density • Statistical properties

1 LIF Models with Reversal Hyperpolarization Potential

We consider the Gauss-Markov process $\{Y(t), t \geq 0\}$ characterized by mean function

$$m(t) = \varrho\left(1 - e^{-t/\vartheta}\right) + \int_0^t \mu(\xi)\, e^{-(t-\xi)/\vartheta}\, d\xi \quad (t \geq 0) \tag{1}$$

and covariance function $c(s,t) = h_1(s)\, h_2(t)$ $(0 \leq s \leq t)$ such that

A. Buonocore (✉) • L. Caputo • E. Pirozzi
Dipartimento di Matematica e Applicazioni, Università di Napoli Federico II, Monte S. Angelo, 80126, Napoli, Italy
e-mail: anibuono@unina.it; luigia.caputo@unina.it; epirozzi@unina.it

A.G. Nobile
Dipartimento di Studi e Ricerche Aziendali (Management & Information Technology), Università di Salerno, Via Giovanni Paolo II, n. 132, 84084 Fisciano (SA), Italy
e-mail: nobile@unisa.it

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_42

$$h_1(t) = \frac{\sigma^2 \vartheta}{2}\left(e^{t/\vartheta} - e^{-t/\vartheta}\right), \quad h_2(t) = e^{-t/\vartheta} \quad (t \geq 0), \tag{2}$$

where $\vartheta > 0$, $\sigma > 0$, $\varrho \in \mathbb{R}$ and $\mu(t) \in C^1[0, +\infty)$. The transition density $f_Y(x, t|y, \tau)$ is a normal density with mean and variance (cf., for instance, [1, 3]):

$$M(t|y,\tau) = y\, e^{-(t-\tau)/\vartheta} + \varrho\left(1 - e^{-(t-\tau)/\vartheta}\right) + \int_\tau^t \mu(\xi)\, e^{-(t-\xi)/\vartheta}\, d\xi,$$

$$V(t|\tau) = \frac{\sigma^2 \vartheta}{2}\left(1 - e^{-2(t-\tau)/\vartheta}\right), (0 \leq \tau \leq t). \tag{3}$$

The infinitesimal moments of $Y(t)$ are:

$$A_1(x,t) = -\frac{x-\varrho}{\vartheta} + \mu(t), \quad A_2(t) = \sigma^2, \quad (x \in \mathbb{R}, \vartheta > 0, \sigma > 0, \varrho \in \mathbb{R}), \tag{4}$$

that identify the drift and the infinitesimal variance of a time non-homogeneous Ornstein-Uhlenbeck process defined in the interval $(-\infty, +\infty)$.

In the context of neuronal models, (4) characterize an inhomogeneous LIF diffusion process $Y(t)$, describing the evolution of the membrane potential (see, for instance, [1, 8, 9]). The time constant ϑ governs the spontaneous decay of the membrane potential to the resting level ϱ, the function $\mu(t)$ represents external signal inputs to the neuron, whereas the infinitesimal variance σ^2 gives the amplitude of the noise.

In the neuronal model (4) the state space for the underlying stochastic process is the entire real axis, implying that arbitrarily large hyperpolarization values for the membrane potential are possible. Some authors (see, for instance, [4, 6, 7]) have considered alternative models by assuming the existence of a lower boundary for the membrane potential. For this reason, in the sequel, we shall focus on the Ornstein-Uhlenbeck process confined by a time-dependent reflecting boundary $\nu(t)$ that can be looked at as the neuronal reversal hyperpolarization potential.

We consider the stochastic process $\{X(t), t \geq 0\}$, defined in the interval $[\nu(t), +\infty)$, obtained by considering the Ornstein-Uhlenbeck process (4) in presence of a reflecting boundary

$$\nu(t) = \varrho\,(1 - e^{-t/\vartheta}) + \int_0^t \mu(\xi)\, e^{-(t-\xi)/\vartheta} d\xi + B\, e^{-t/\vartheta}, \tag{5}$$

with $B \in \mathbb{R}$. The transition probability density function (pdf) of $X(t)$ is (see, [4])

$$f_X(x,t|y,\tau) = f_Y(x,t|y,\tau) + f_Y[2\,\nu(t) - x, t|y,\tau], \quad [x \geq \nu(t), y \geq \nu(\tau)], \tag{6}$$

with $f_Y(x, t|y, \tau)$ normal density with mean and variance given in (3).

We now analyze the asymptotic behavior of $X(t)$ when the input signal $\mu(t)$ is a bounded function, asymptotically constant, i.e. $\lim_{t\to+\infty}\mu(t)=\hat{\mu}$. Under such assumption, from (3) and (5) one has:

$$E(Y) := \lim_{t\to+\infty} M(t|y,\tau) = \lim_{t\to+\infty} \nu(t) = \varrho + \hat{\mu}\,\vartheta,$$

$$\mathrm{Var}(Y) := \lim_{t\to+\infty} V(t|\tau) = \frac{\sigma^2\,\vartheta}{2},$$

so that from (6) one obtains the steady-state pdf of the restricted Ornstein-Uhlenbeck process $X(t)$:

$$W(x) := \lim_{t\to+\infty} f_X(x,t|y,\tau) = \frac{2}{\sigma\sqrt{\pi\vartheta}}\exp\Big\{-\frac{(x-\rho-\hat{\mu}\,\vartheta)^2}{\sigma^2\vartheta}\Big\} \quad (x\geq\varrho+\hat{\mu}\,\vartheta)$$

and the related asymptotic mean and variance:

$$E(X) = \varrho + \hat{\mu}\,\vartheta + \sigma\sqrt{\frac{\vartheta}{\pi}}, \quad \mathrm{Var}(X) = \sigma^2\vartheta\left(\frac{1}{2}-\frac{1}{\pi}\right).$$

We note that $E(Y)=\lim_{t\to+\infty}\nu(t)$ for the free Ornstein-Uhlenbeck process $Y(t)$, whereas $E(X) > \lim_{t\to+\infty}\nu(t)$ for the restricted Ornstein-Uhlenbeck process $X(t)$. Furthermore, by comparing the asymptotic means and variances of $X(t)$ and $Y(t)$, one has $E(X) > E(Y)$ and $\mathrm{Var}(X) < \mathrm{Var}(Y)$.

We now consider the first passage time (FPT) problem to a threshold $S(t) > \nu(t)$, with $S(t)\in C^1[0,+\infty)$, for the restricted Ornstein-Uhlenbeck process, with reflecting boundary $\nu(t)$ given in (5). We denote by

$$\mathscr{T}_y = \inf_{t\geq\tau}\big\{t : X(t) > S(t)\big\}, \qquad X(\tau) = y < S(\tau),$$

the random variable FPT of $X(t)$ from $X(\tau)=y\geq\nu(\tau)$ to the threshold $S(t)$ and by

$$g_X[S(t),t|y,\tau] = \frac{\partial}{\partial t}P\big(\mathscr{T}_y\leq t\big) \qquad [\nu(\tau)\leq y < S(\tau)]$$

its FPT pdf. In the neuronal modeling context, $g_X[S(t),t|y,\tau]$ identifies the firing density of $X(t)$, i.e. the FPT pdf of $X(t)$ through the firing threshold $S(t)$ starting from y at time τ.

Figure 1 shows two simulated sample paths for the free process $Y(t)$ (green) and for the restricted process $X(t)$ (blue), both obtained via an appropriate algorithm based on the simulation of the Brownian motion in the presence of a constant reflecting boundary and on suitable spatio-temporal transformations (the details of the algorithm are object of our paper in preparation). The arrow in Fig. 1 indicates

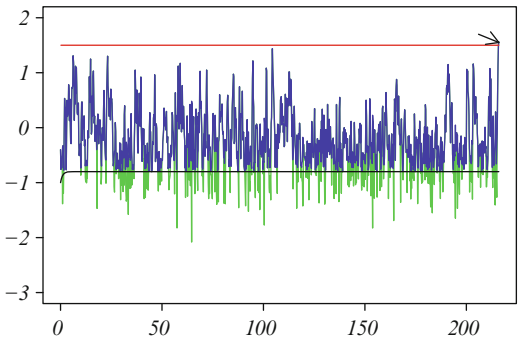


Fig. 1 Sample paths of the free process $Y(t)$ (*green*) and of the restricted process $X(t)$ (*blue*) in the presence of a reflecting boundary $\nu(t)$ (*black*) for $S = 1.5$ (*red*) and constant input signal $\mu(t) = 0.1$

the instant of the first passage time. The firing pdf $g_X[S(t), t|y, \tau]$ is solution of the nonsingular Volterra integral equation (see, [4])

$$g_X[S(t), t|y, \tau] = -2\Psi_X[S(t), t|y, \tau] + 2\int_\tau^t g_X[S(u), u|y, \tau]\Psi_X[S(t), t|S(u), u]\, du \tag{7}$$

with $\nu(\tau) \le y < S(\tau)$ and where

$$\Psi_X[S(t), t|y, \tau] = f_X[S(t), t|y, \tau] \times \left\{ \frac{S'(t) - m'(t)}{2} - \frac{S(t) - m(t)}{2\,\vartheta} \frac{1 + e^{-2(t-\tau)/\vartheta}}{1 - e^{-2(t-\tau)/\vartheta}} + \frac{[y - m(\tau)]\, e^{-(t-\tau)/\vartheta}}{\vartheta\left[1 - e^{-2(t-\tau)/\vartheta}\right]} \right\} \tag{8}$$

with $\nu(t)$ and $f_X(x, t|y, \tau)$ given in (5) and (6), respectively. If $\mu(t)$ is a bounded function such that $\lim_{t\to+\infty} \mu(t) = \hat{\mu}$, then the firing pdf $g_X[S(t), t|y, \tau]$ through the bounded threshold $S(t)$, such that $\lim_{t\to\infty} S(t) = \hat{S}$, admits the following exponential asymptotic behavior:

$$g_X[S(t), t|y, \tau] \simeq R(\hat{S})\ \exp\{-R(\hat{S})(t - \tau)\}, \tag{9}$$

for $\nu(\tau) \le y < S(\tau)$, where

$$R(\hat{S}) = \frac{2\,[\hat{S} - (\varrho + \hat{\mu}\,\vartheta)]}{\sigma\,\vartheta\,\sqrt{\pi\vartheta}}\ \exp\left\{-\frac{[\hat{S} - (\varrho + \hat{\mu}\,\vartheta)]^2}{\sigma^2\,\vartheta}\right\}.$$

The goodness of the exponential approximation (9) increases as the threshold is progressively moved away from the starting point and from the asymptotic mean of the related processes.

2 Neuronal Models with Different Input Signals

We now specialize our results to different types of functions $\mu(t)$, by considering the following cases: constant, exponential and periodic signals. In these three cases, the asymptotic behavior of the firing pdf will be explicitly determined.

2.1 Constant and Exponential Input Signals

For $t \geq 0$, we assume that the input signal is

$$\mu(t) = \mu + \lambda\, e^{-t/\vartheta} \quad (\vartheta > 0; \lambda, \mu \in \mathbb{R}). \tag{10}$$

We consider the process $X(t)$ restricted by the reflecting boundary

$$\nu(t) = (\varrho + \mu\,\vartheta)\,(1 - e^{-t/\vartheta}) + \lambda\, t\, e^{-t/\vartheta} + B\, e^{-t/\vartheta}, \tag{11}$$

with $B \in \mathbb{R}$. The transition pdf of $X(t)$ is given in (6) and the firing pdf can be computated via (7) and (8) by using the numerical algorithm proposed in [3]. For the case of exponential input signal (10), in the left part of Fig. 2 we plot the firing density $g_Y(S,t|y,0)$ of the free process $Y(t)$ (red curve) and the firing density $g_X(S,t|y,0)$ of $X(t)$ (blue curve) through the constant threshold $S = 1.5$, starting from the initial state $y = -0.4$ at time $\tau = 0$. The reflecting boundary for $X(t)$ is $\nu(t) = -0.8 - (0.2 + 0.1\,t)\,e^{-t}$. The dashed curves indicate the asymptotic behaviors of the firing densities for $X(t)$ and $Y(t)$, respectively. The integration step in the numerical algorithm is set at $\vartheta/10 = 0.1$. Instead, on the right of Fig. 2 the simulated firing pdf for the process $X(t)$ is shown.

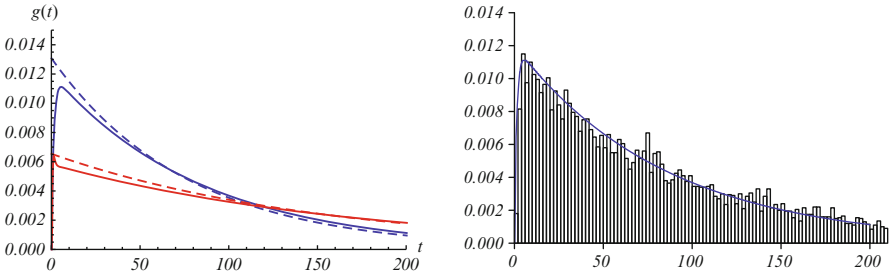


Fig. 2 On the *left*: for the exponential signal (10), $g_Y(S,t|y,0)$ (*red curve*), $g_X(S,t|y,0)$ (*blue curve*) and their asymptotic exponential behaviors (*dashed curves*) are plotted with $\vartheta = 1$, $\mu = 0.1$, $\varrho = -0.9$, $\lambda = -0.1$, $\sigma^2 = 1$, $B = -1$, $y = -0.4$ and $S = 1.5$. On the *right*: the histogram of a sample of 10,000 simulated firing times is compared with the firing density $g_X(S,t|y,0)$

2.2 Periodic Input Signal

We suppose that the input signal is the periodic function

$$\mu(t) = \mu + \lambda \cos(\omega\, t + \phi) \quad (\mu \in \mathbb{R}, \lambda \neq 0, \phi \in \mathbb{R}, \omega > 0) \tag{12}$$

for $t \geq 0$, whose period is $Q = 2\pi/\omega$. LIF neuronal models with periodic input signal are considered, for instance, in [2, 5, 9, 10] and references therein. For the process $X(t)$ with the reflecting boundary

$$\nu(t) = (\varrho + \mu\, \vartheta)\,(1 - e^{-t/\vartheta}) + \frac{\lambda\, \vartheta}{1 + \omega^2\, \vartheta^2}$$
$$\times \Big\{ \cos(\omega\, t + \phi) + \omega\, \vartheta\, \sin(\omega\, t + \phi) - \big[\cos\phi + \omega\, \vartheta\, \sin\phi\big]\, e^{-t/\vartheta} \Big\} + B\, e^{-t/\vartheta} \tag{13}$$

with $B \in \mathbb{R}$, the transition pdf is given in (6). For the case of the periodic input signal (12), in the left part of Fig. 3 we plot the firing density $g_Y(S, t\,|\,y, 0)$ of the free process $Y(t)$ (red curve) and the firing density $g_X(S, t\,|\,y, 0)$ of $X(t)$ (blue curve) through the constant threshold $S = 1.5$, starting from the initial state $y = -0.4$ at time $\tau = 0$. As shown in Fig. 3, the firing densities $g_Y(S, t\,|\,y, 0)$ and $g_X(S, t\,|\,y, 0)$ exhibit damped oscillations having the same period Q of the periodic input signal. For the periodic input signal (12), in the right part of Fig. 3 the simulated and the asymptotic firing density of the process $X(t)$ are compared with the numeric firing pdf of the left part of Fig. 3.

Detailed proofs of the results shown so far as well as some their extensions will be provided in future works.

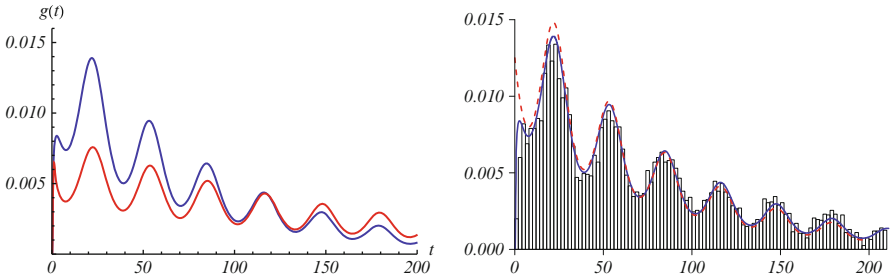


Fig. 3 On the *left*: for the periodic signal (12), $g_Y(S, t\,|\,y, 0)$ (*red curve*), $g_X(S, t\,|\,y, 0)$ (*blue curve*) are plotted for $\phi = 5$, $\vartheta = 1$, $\mu = 0.1$, $\varrho = -0.9$, $\lambda = -0.1$, $\omega = 0.2$, $\sigma^2 = 1$, $B = -1$, $y = -0.4$ and $S = 1.5$. On the *right*: the histogram of a sample of 10,000 simulated firing times is compared with the firing density $g_X(S, t\,|\,y, 0)$ (*blue curve*) and the asymptotic behavior (*dashed red curve*)

References

1. Buonocore A., Caputo L., Pirozzi E., Ricciardi L.M.: The first passage time problem for Gauss-diffusion processes: algorithmic approaches and applications to LIF neuronal model. Methodol. Comput. Appl. Probab. **13**, 29–57 (2011)
2. Burkitt A. N.: A review of the integrate-and-fire neuron model. II. Inhomogeneous synaptic input and network properties. Biol. Cybernet. **95**(2), 97–112 (2006)
3. Di Nardo E., Nobile A.G., Pirozzi E., Ricciardi L.M.: A computational approach to first-passage-time problems for Gauss-Markov processes. Adv. Appl. Prob. **33**, 453–482 (2001)
4. Giorno V., Nobile A.G., Pirozzi E., Ricciardi L.M.: Towards some computational problems arising in biological modeling. In: Moreno Diaz R., Pichler F. (eds.) Lecture Notes in Computer Science, vol 2809, pp. 360–371. Springer-Verlag, Berlin Heidelberg (2003)
5. Giraudo M.T., Sacerdote L.: Effect of periodic stimulus on a neuronal diffusion model with signal-dependent noise. BioSystems **79**, 73–81 (2005)
6. Inoue J., Doi S.: Sensitive dependence of the coefficient of variation of interspike intervals on the lower boundary of membrane potential for leaky integrate-and.fire neuron model. BioSystems **87**, 49–57 (2007)
7. Jahn P., Berg R.W., Hounsgaard J., Ditlevsen S.: Motoneuron membrane potentials follow a time inhomogeneous jump diffusion process. J. Comput Neurosci. **31**, 563–579 (2011)
8. Lansky P., Ditlevsen S.: A review of the methods for signal estimation in stochastic diffusion leaky integrate-and-fire neuronal models. Biol. Cybern. **99**, 253–262 (2008)
9. Kobayashi R., Shinomoto S., Lansky P.: Estimation of time-dependent input from neuronal membrane potential. Neural Computation **23**, 3070–3093 (2011)
10. Schindler M., Talkner P., Hänggi P.: Escape rates in periodically driven Markov processes. Physica A **351**, 40–50 (2005)

On Super-Turing Neural Computation

Jérémie Cabessa and Alessandro E.P. Villa

Abstract In this paper, we provide a historical survey of the most significant results concerning the computational power of neural models. We distinguish three important periods: first, the early works from McCulloch and Pitts, Kleene, and Minky, where the computational equivalence between Boolean recurrent neural networks and finite state automata is established. Secondly, the two breakthroughs by Siegelmann and Sontag showing the Turing universality of rational-weighted neural networks, and the super-Turing capabilities of analog recurrent neural networks. Thirdly, the recent results by Cabessa, Siegelmann and Villa revealing the super-Turing computational potentialities of interactive and evolving recurrent neural networks.

Keywords Neural computation • Recurrent neural networks • Finite automata • Turing machines • Turing machines with advice • super-Turing

1 The Early Works

In theoretical neuroscience, understanding the computational and dynamical capabilities of biological neural networks is an issue of central importance. In this context, much interest has been focused on comparing the computational powers of diverse theoretical neural models with those of abstract computing devices.

This comparative approach was initiated by McCulloch and Pitts who proposed a modelisation of the nervous system as a finite interconnection of threshold logic units [19]. For the first time, neural networks were considered as discrete abstract machines, and the issue of their computational capabilities investigated from the automata-theoretic perspective. In this context, Kleene and Minsky proved that recurrent neural networks made up of threshold activation units were computationally equivalent to classical finite state automata [13, 20].

J. Cabessa (✉) • A.E.P. Villa
Neuroheuristic Research Group, University of Lausanne, Quartier Dorigny, CH-1015 Lausanne, Switzerland
e-mail: jcabessa@nhrg.org; avilla@nhrg.org

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_43

Besides, in a seminal report entitled "Intelligent Machinery" [31], Turing brilliantly introduced many concepts which have later become central in the field of neural computation. For instance, Turing foresaw the possibility of surpassing the capabilities of finite state machines and reaching Turing universality via neural networks called "B-type unorganised machines". The networks consisted of a general interconnection of NAND neurons, and the consideration of infinitely many such cells could simulate the behaviour of a Turing machine. Moreover, Turing also introduced the key idea of "training" neural networks by considering the possibility of modifying the synaptic connections between the cells by means of what he called "connection-modifiers". Later, the Turing universality of infinite or heterogeneous neural networks has further been investigated in many directions, see for instance [8,9,11,23]. These seminal works opened up the way to the theoretical computer scientist approach to neural computation. However, the purely discrete and mechanical approach under consideration quickly appeared too restrictive, far from the biological reality.

According to these considerations, von Neumann proposed another relevant approach to the issue of information processing in the brain from the hybrid perspective of digital and analog computation [22]. He considered that the non-linear character of the operations of the brain emerges from a combination of discrete and continuous mechanisms, and therefore envisioned neural computation as something strictly more powerful than abstract machines. Almost in the same time, Rosenblatt proposed the so-called "perceptron" as a more general computational neural model than the McCulloch-Pitts units [24]. The essential innovation consisted in the introduction of numerical synaptic weights and as well as a special interconnection pattern. This neural model gave rise to an algorithmic conception of "learning" achieved by adjusting the synaptic weights of the networks according to some specific task to be completed. This study is nowadays considered as foundational for the field of machine learning. The computational capabilities of the perceptron were further studied by Minsky and Papert [21].

2 Two Significant Breakthroughs

Later, Siegelmann and Sontag made two significant steps forward concerning the precise issue of the computational power of recurrent neural networks. Firstly, they focused their attention on the consideration of more realistic activation functions for the neurons and showed that by extending the activation functions of the cells from boolean to linear-sigmoid, the computational power of the neural networks would drastically increase from finite state automata up to Turing capabilities [28]. The Turing universality of neural networks was then generalised to a broader class of sigmoidal activation functions [12]. The computational equivalence between the so-called *rational recurrent neural networks* and the Turing machines has nowadays become standard result in the field.

Secondly and most importantly, following von Neumann considerations, they assumed that the variables appearing in the underlying chemical and physical phenomena could be modelled by continuous rather than discrete numbers, and therefore proposed a precise study of the computational power of recurrent neural networks from the perspective of analog computation [27]. They introduced the concept of an *analog recurrent neural network* as a classical linear-sigmoid neural net equipped with real- instead of rational-weighted synaptic connections. This analog information processing model turns out to be capable of capturing the non-linear dynamical properties that are most relevant to brain dynamics, such as rich chaotic behaviours [7, 25, 26, 29, 32]. In this context, they proved that analog recurrent neural networks are computationally equivalent to Turing machine with advice, hence capable of super-Turing computational capabilities from polynomial time of computation already. They further formulated the so-called Thesis of Analog Computation – an analogous to the Church-Turing thesis, but in the realm of analog computation – stating that no reasonable abstract analog device can be more powerful than first-order analog recurrent neural networks [26, 27].

3 Present and Future

But until the mid 1990s, the neural models involved in the study of the computational capabilities of recurrent neural networks have always been oversimplified, lacking many biological features which turn out to be essentially involved in the processing of information in the brain. In particular, the effects that various kinds of noise might have on the computational power of recurrent neural networks had not been considered. Moreover, the ability of neural networks to evolve over time has also been neglected in the models under consideration. Biological mechanisms like synaptic plasticity, cell birth and death, changes in connectivity, etc., – which are widely assumed to be of primary importance in the processing and encoding of information –, have yet not been taken into consideration in the study of the computational capabilities of neural networks.

Concerning noise, Maass and Orponen showed that general analog computational systems subjected to arbitrarily small amount of analog noise have their computational power reduced to that of finite automata or even less [17]. In particular, the presence of arbitrarily small amount of analog noise seriously reduces the capabilities of both rational- and real-weighted recurrent neural networks to those of finite automata, namely to the recognition of regular languages. Maass and Sontag then extended this result by showing that, in the presence of gaussian or other common analog noise distribution with sufficiently large support, recurrent neural networks have their computational reduced to even less than finite automata, namely to the recognition of definite languages [18]. These two results were further generalised to the broader classes of quasi-compact and weakly ergodic Markov computational systems, respectively [1].

Concerning the evolvability of neural networks, Cabessa and Siegelmann considered a more biologically oriented model where the synaptic weights, the connectivity pattern, and the number of neurons can evolve rather than stay static [3]. The so-called *evolving recurrent neural networks* were proven to be computationally equivalent to the analog neural networks, and hence capable of super-Turing computational power, regardless of whether their synaptic weights are rational or real. These results are important, showing that the power of evolution brings up additional potentialities to first-order recurrent neural networks and provides an alternative and equivalent way to the incorporation of the power of the continuum towards the achievement of super-Turing computational capabilities of neural networks. This feature is particularly interesting since certain analog assumptions in neural models have sometimes been argued to be too strong.

However, in this global line of thinking, the issue of the computational capabilities of neural networks has always been considered from the strict perspective of Turing-like classical computation [30]: a network is viewed as an abstract machine that receives a finite input stream from its environment, processes this input, and then provides a corresponding finite output stream as answer, without any consideration to the internal or external changes that might happen during the computation. But this classical computational approach is inherently restrictive, and has nowadays been argued to "no longer fully corresponds to the current notion of computing in modern systems" [16], especially when it refers to bio-inspired complex information processing systems [14, 16]. Indeed, in the brain (or in organic life in general), information is rather processed in an interactive way, where previous experience must affect the perception of future inputs, and where older memories may themselves change with response to new inputs. Hence, neural networks should rather be conceived as performing sequential interactions or communications with their environments, and be provided with memory that remains active throughout the whole computational process, rather than proceeding in a closed-box amnesic classical fashion. Accordingly, the computational power of recurrent neural networks should rather be conceived from the perspective of interactive computation [10].

Along these lines, Cabessa and Siegelmann studied the computational power of recurrent neural networks involved in a basic interactive computational paradigm [4]. They proved that the so-called *interactive recurrent neural networks* with rational and real synaptic weights are computationally equivalent to interactive Turing machines and interactive Turing machines with advice, respectively. These achievements provide a generalisation to the bio-inspired interactive computational context of the previous classical results by Siegelmann and Sontag [27, 28]. Besides, Cabessa and Villa also provided a study of the super-Turing computational capabilities of analog neural networks involved in another kind of reactive and memory active computational framework [5].

The last advances concerning the study of the computational power of recurrent neural networks were provided by Cabessa and Villa [2, 6]. They studied the computational potentialities of a recurrent neural model combining the two relevant features of evolvability and interactivity introduced in [3, 4], and showed that the

so-called *interactive evolving recurrent neural networks* are capable of super-Turing computational potentialities, equivalent to interactive Turing machine with advice, irrespective of whether their synaptic weights are rational or real.

These results show that the consideration of evolving capabilities in a first-order interactive neural model provides the potentiality to break the Turing barrier, irrespective of whether the synaptic weights are rational or real. They support the extension of the Church-Turing Thesis to the context of interactive computation: "Any (non-uniform interactive) computation can be described in terms of interactive Turing machines with advice" [15]. As for the classical computational framework, the super-Turing computational capabilities can be achieved without the need of a framework based on the power of the continuum – in the case of interactive evolving recurrent neural networks with rational weights. This feature is particularly meaningful, since while the power of the continuum is a pure conceptualisation of the mind, the evolving capabilities of the networks are, by contrast, really observable in nature.

From a general perspective, we believe that such theoretical studies about the computational power of bio-inspired neural models might ultimately bring further insight to the understanding of the intrinsic natures of both biological as well as artificial intelligences. We also think that foundational approaches to alternative models of computation might in the long term not only lead to relevant theoretical considerations, but also to important practical applications. Similarly to the theoretical work from Turing which played a crucial role in the practical realisation of modern computers, further foundational considerations of alternative models of computation will certainly contribute to the emergence of novel computational technologies and computers, and step by step, open the way to the next computational era.

References

1. Ben-Hur A, Roitershtein A, Siegelmann HT (2004) On probabilistic analog automata. Theor Comput Sci 320(2–3):449–464
2. Cabessa J (2012) Interactive evolving recurrent neural networks are super-turing. In: Filipe J, Fred ALN (eds) ICAART (1), SciTePress, pp 328–333
3. Cabessa J, Siegelmann HT (2011) Evolving recurrent neural networks are super-turing. In: IJCNN, IEEE, pp 3200–3206
4. Cabessa J, Siegelmann HT (2012) The computational power of interactive recurrent neural networks. Neural Computation 24(4):996–1019
5. Cabessa J, Villa AEP (2012) The expressive power of analog recurrent neural networks on infinite input streams. Theor Comput Sci 436:23–34
6. Cabessa J, Villa AEP (2013) The super-turing computational power of interactive evolving recurrent neural networks. LNCS, vol 8131, pp 58–65
7. Celletti A, Villa AE (1996) Determination of chaotic attractors in the rat brain. J Stat Phys 84(5):1379–1385
8. Franklin S, Garzon M (1989) Neural computability. In: Omidvar O (ed) Progress in Neural Networks, Ablex, Norwood, NJ, USA, pp 128–144

9. Garzon M, Franklin S (1989) Neural computability II. In: Omidvar O (ed) Proceedings of the Third International Joint Conference on Neural Networks, IEEE, pp 631–637
10. Goldin D, Smolka SA, Wegner P (2006) Interactive Computation: The New Paradigm. Springer-Verlag New York, Inc., Syracuse, NJ, USA
11. Hartley R, Szu H (1987) A comparison of the computational power of neural network models. In: Butler C (ed) Proceedings of the IEEE First International Conference on Neural Networks, IEEE, pp 17–22
12. Kilian J, Siegelmann HT (1996) The dynamic universality of sigmoidal neural networks. Inf Comput 128(1):48–56
13. Kleene SC (1956) Representation of events in nerve nets and finite automata. In: Shannon C, McCarthy J (eds) Automata Studies, Princeton University Press, Princeton, NJ, pp 3–41
14. van Leeuwen J, Wiedermann J (2001) Beyond the Turing limit: Evolving interactive systems. LNCS, vol 2234, pp 90–109
15. van Leeuwen J, Wiedermann J (2000) The Turing machine paradigm in contemporary computing. In: Enquist B, Schmidt W (eds), Mathematics Unlimited - 2001 and Beyond, Springer-Verlag, Berlin / Heidelberg, pp 1139–1155
16. van Leeuwen J, Wiedermann J (2008) How we think of computing today. LNCS, vol 5028, pp 579–593
17. Maass W, Orponen P (1998) On the effect of analog noise in discrete-time analog computations. Neural Comput 10(5):1071–1095
18. Maass W, Sontag ED (1999) Analog neural nets with gaussian or other common noise distributions cannot recognize arbitary regular languages. Neural Comput 11(3):771–782
19. McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysic 5:115–133
20. Minsky ML (1967) Computation: finite and infinite machines. Prentice-Hall, Inc., Englewood Cliffs, N. J.
21. Minsky ML, Papert S (1969) Perceptrons: An Introduction to Computational Geometry. MIT Press, Cambridge, MA, USA
22. Neumann J von (1958) The computer and the brain. Yale University Press, New Haven, CT, USA
23. Pollack JB (1987) On connectionist models of natural language processing. PhD thesis, Computing Research Laboratory, New Mexico State University, Las Cruces, NM
24. Rosenblatt F (1957) The perceptron: A perceiving and recognizing automaton. Tech. Rep. 85-460-1, Cornell Aeronautical Laboratory, Ithaca, New York
25. Segundo JP (2003) Nonlinear dynamics of point process systems and data. Int J Bif Chaos 13(08):2035–2116
26. Siegelmann HT (1995) Computation beyond the Turing limit. Science 268(5210):545–548
27. Siegelmann HT, Sontag ED (1994) Analog computation via neural networks. Theor Comput Sci 131(2):331–360
28. Siegelmann HT, Sontag ED (1995) On the computational power of neural nets. J Comput Syst Sci 50(1):132–150
29. Tsuda I (2001) Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems. Behav Brain Sci 24(5):793–847
30. Turing AM (1936) On computable numbers, with an application to the Entscheidungsproblem. Proc London Math Soc 2(42):230–265
31. Turing AM (1948) Intelligent machinery. Technical report, National Physical Laboratory, Teddington, UK
32. Villa A, Tetko I (1995) Spatio-temporal patterns of activity controlled by system parameters in a simulated thalamo-cortical neural network. In: Herrmann H, Wolf D, Poppel E (eds) Supercomputing in Brain Research: from Tomography to Neural Networks, World Scientific, pp 379–388

Chasing Cognitive Neurodynamics by Single-Trial Analysis of Electroencephalogram (EEG)

Yong Hu, Li Hu, Hongtao Liu, Zhiguo Zhang, Guangju Zhang, and Hongyan Cui

Abstract Single trial electroencephalogram (EEG) and evoked potential (EP) is a very important tool to investigate cognitive neurodynamics. This paper introduced a newly developed toolbox for single trial extraction of EP/ERPs (STEP v1.0). Its application in laser evoked potential (LEP) and somatosensory evoked potential (SEP) were presented to demonstrate the use of single trial analysis. Trial-to-trial variability provide plentiful biological information, which helped us understanding the function of the nervous system.

Keywords Nociceptive processing • Non-nociceptive processing

1 Introduction

Sensory, motor or cognitive events can elicit sudden and short-lasting changes in Electroencephalogram (EEG), embedded in the ongoing spontaneous activities. In addition to triggering evoked potentials (EP) or event related potentials (ERPs), various events may also trigger transient modulations (ERS and ERD) of the ongoing oscillatory brain activity. The amount of these event-related changes is much smaller than that of the background EEG activity. Thus, changes triggered by single trials are hardly detected, and the most widely used approach to identify them is to average a large number of trials, which relies on the assumption that the shape and the latency of EP/ERPs are equal from trial to trial. However, it has been clearly

Y. Hu (✉) • L. Hu • H. Liu
Department of Orthopaedics and Traumatology, The University of Hong Kong, Pokfulam, Hong Kong, People's Republic of China
e-mail: yhud@hotmail.com

Z. Zhang
Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, Hong Kong

G. Zhang • H. Cui
Institute of Biomedical Engineering, Chinese Academy of Medical Sciences, 300192 Tianjin, People's Republic of China

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_44

shown that there is a large dynamic variability in the EEG response from trial to trial. However, this lost information is very important for EEG studies for the following reason: (1) because it contains information about how the brain varies its response to the same stimulus; (2) because it allows researchers to look for correlations between this brain response and a behavioural performance for each single trial; (3) because it allows researchers to compare this measure of the electrical brain activity with other more measures of brain function. Therefore, methods that explore ERP dynamics at the level of single trials can provide new insights into the functional significance of the different brain processes underlying these brain responses [1, 2]. In this paper, a series of various methods on single trial EP/ERPs analysis has been developed [1–4] and a single trial extraction toolbox for EP/ERPs (STEP v1.0) was proposed.

2 Method

We have developed a single trail extraction toolbox for EP/ERPs (STEP v1.0). This interactive Matlab toolbox includes: (1) Continuous Wavelet filtering (CWF), (2) Multiple linear regression (MLR), (3) Multiple linear regression with Dispersion Term (MLRd), (4) Constrained Second Order Blind Identification (CSOBI), (5) Discrete Wavelet filtering (DWT), (6) Independent Component Analysis (ICA) and Probabilistic Independent Component Analysis (PICA). In this paper, the usefulness of this new toolbox was demonstrated by its applications in laser evoked potentials (LEP) and somatosensory evoked potentials (SEP).

2.1 Time Domain Morphology of Single-Trial LEP

The signal to noise ratio (SNR) of LEP was improved by wavelet filtering. Then, the latency, amplitude and morphology of single-trial N2-P2 complex in LEPs was extracted by multiple linear regression (MLR) (Fig. 1). In MLR, variability matrices that capture the variations of latency and morphology of each LEP peak are generated by shifting and compressing the average LEP waveform (1). These variability matrices are fed to a principal component analysis (PCA) (2). The resulting three main principal components (PCs) are used to define three regressors for each peak (3). These regressors are then applied against each single trial and used to model each single-trial ERP peak (4).

2.2 Single-Trial Time-Frequency Domain LEP

Figure 2 shows the procedure when generating regressors for LEP in the TF-sMLR approach using the nonparametric approach based on PCA, while the same number

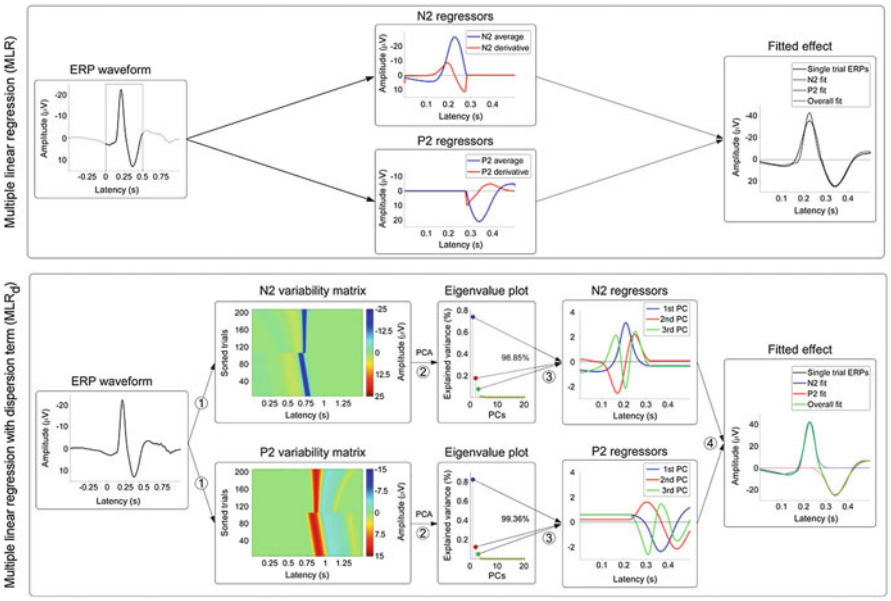


Fig. 1 Flowcharts describing the procedures of multiple linear regression (MLR) and multiple linear regression with dispersion term (MLR_d) to estimate the single-trial latency and amplitude of ERP peaks

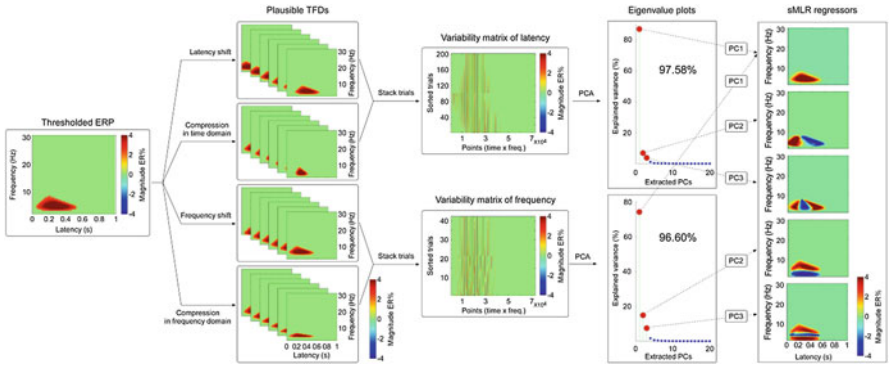


Fig. 2 Flowchart describing the procedure developed to generate the regressors for the TFD of LEP in TF-sMLR using the nonparametric approach based on PCA

of regressors has also been obtained for ERS and ERD. The first three principal component PCs obtained from the latency variability matrix explained 97.58 %, 95.30 %, and 95.36 % of the total variance for LEP, ERD, and ERS respectively, and the first three PCs obtained from the frequency variability matrix explained 96.60 %, 93.26 %, and 97.28 % of the total variance for LEP, ERD, and ERS respectively.

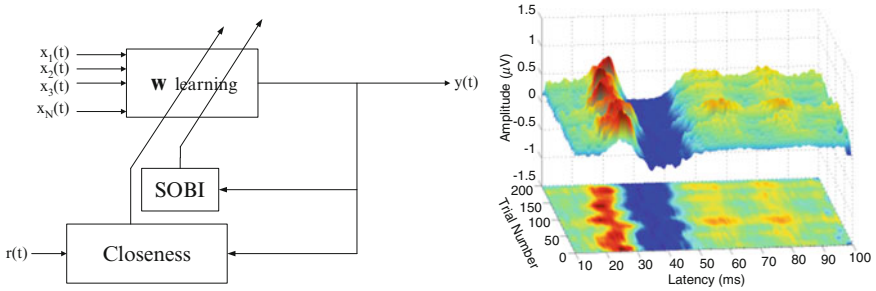


Fig. 3 Single trial SEP extracted by SOBI-R (*left panel*: Block diagram of cSOBI algorithm; *Right panel*: single trial SEP)

2.3 Single Trial SEP Extraction

SEP is the most popular technique in spinal cord monitoring to avoid possible neurological complication because of its advantages of low cost, easy to use, non-invasive and higher successful rate. However, SEP monitoring is somewhat more technically challenging than other evoked potential methods because of its weaker signal and extremely low SNR, variable waveform and time-consuming acquisition. Ensemble averaging (EA) is the most widely used method, which unavoidably requires a longer measurement time. Significant irreversible neurological injury might have already occurred during the lengthy data collection time. In addition, there are a lot of features in single trial SEP which were ignored by averaging a relatively large number of consecutive trials. We developed a new method to extract SEP with rare channel by using one-unit second-order blind identification with reference (SOBI-R) algorithm (Fig. 3).

3 Conclusions

While developing these methods of single trial EP/ERPs analysis and the toolbox STEP v1.0, we have applied them to explore the single-trial dynamics between these measures, behavioural variables and also measurements of brain activity. Trial-to-trial variability and the explored relationships provide plentiful biological information, which helped us understanding the function of the nervous system.

Acknowledgments This work was supported in part by the Hong Kong Research Grants Council (GRF HKU713006E/767511E), National Natural Science Foundation of China (No. 81271685) and the Tianjin Key Project Grant (No.11ZCKFSY01600).

References

1. Hu, L., Mouraux, A., Hu, Y., Iannetti, G.D. (2011) A novel approach for enhancing the signal-to-noise ratio and detecting automatically event-related potentials (ERPs) in single trials. Neuroimage 50, 99–111.
2. Hu L, Zhang ZG, Hung YS, Luk KD, Iannetti GD, Hu Y. (2011) Single-trial detection of somatosensory evoked potentials by probabilistic independent component analysis and wavelet filtering. Clin Neurophysiol. 2011 Jul;122(7):1429–39.
3. Liu H, Chang CQ, Luk KD, Hu Y. (2011) Comparison of blind source separation methods in fast somatosensory-evoked potential detection. J Clin Neurophysiol. 2011 Apr;28(2):170–7.
4. Liu HT, Xie XB, Xu SP, Wan F, Hu Y. (2013) One-unit second-order blind identification with reference for short transient signals. Information Sciences. 2013 April, 227: 90–101.

Brain: Biological Noise-Based Logic

Laszlo B. Kish, Claes G. Granqvist, Sergey M. Bezrukov, and Tamas Horvath

Abstract Neural spikes in the brain form stochastic sequences, i.e., belong to the class of pulse noises. This stochasticity is a counterintuitive feature because extracting information—such as the commonly supposed neural information of mean spike frequency—requires long times for reasonably low error probability. The mystery could be solved by noise-based logic, wherein randomness has an important function and allows large speed enhancements for special-purpose tasks, and the same mechanism is at work for the brain logic version of this concept.

Keywords Neural logic • Deterministic logic • Logic variable • Neural spikes • Stochastic signal

1 Noise-Based Logic and Brain versus Computer

Noise-based logic (NBL) [1], which has been inspired by the fact that the neural signals in the brain are stochastic, utilizes independent stochastic processes as well as their superposition to carry the logic signal. A basic *brain logic* version [2] of NBL has been motivated by the following observations:

L.B. Kish (✉)
Department of Electrical Engineering, Texas A&M University, College Station, TX 77843-3128, USA
e-mail: Laszlokish@tamu.edu

C.G. Granqvist
Department of Engineering Sciences, The Ångström Laboratory, Uppsala University, P.O. Box 534, SE-75121 Uppsala, Sweden

S.M. Bezrukov
Laboratory of Physical and Structural Biology, Program in Physical Biology, NICHD, National Institutes of Health, Bethesda, MD 20892, USA

T. Horvath
Division for Knowledge Discovery, Fraunhofer IAIS, Schloss Birlinghoven, 53754 Sankt Augustin, Germany

Department of Computer Science, University of Bonn, Bonn, Germany

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_45

(i) The number of neurons in the brain is similar to the number of switching elements (MOS transistors) in a modern flash drive and about 10 % of those of a Macbook Air laptop,
(ii) The maximum frequency of neural spikes is about 20 million times less than the clock frequency, and
(iii) Neural spike sequences are stochastic, which suggests that their information channel capacity is further limited.

The above facts indicate that the classical suggestions—that *neural information is statistical* and is carried by the *mean frequency of spikes* or *their cross-correlation* (between spike trains)—are likely to be false. Instead we propose that

(a) Single neural spikes potentially carry orders of magnitude more information than a single bit, and
(b) The brain uses a number of special-purpose operations that allow it to achieve *reasonably accurate but not perfect* results in a short time and with relatively small "brain-hardware" and time complexity.

The fact that the brain operates in a different way than a computer can be easily demonstrated. Figure 1 provides an illustrative example: the lines contain strings that are identical with the exception of one line where there is a small difference. The brain detects this different line immediately without reading each character in every line. A computer, however, would scan the image character-by-character and then make a comparison accordingly. If we try carry out the analysis in the computer's way—i.e., reading and comparing each element in the tasks described above—we would perform extremely slowly in a large system. Clearly, the brain is using different schemes than computers and employs various special-purpose, noise-based operations, and it must do so because its "brute force" abilities are weaker than those of a laptop.

Fig. 1 Simple demonstration of the difference between the ways of operation for the brain and a computer

Two more quick demonstrations of the difference between the brain and a computer are these: We can try to multiply the first 100 integer numbers and check how long it takes for a laptop computer, it takes less than a millisecond. And we can try to memorize the string "*saenstnsjrmenHixerLöeailarenecltcsrhel*" or its rearranged version "*Hans Liljenström is an excellent researcher*". The second version is much easier for the brain, while the rearrangement does not matter for the computer. More precisely, the first version is easier for the computer because of the lack of blank characters.

2 The Essential Feature of Noise-Based Brain Logic

Due to the limited space, we only illustrate the most essential feature of brain logic [2], which is that the superposition of orthogonal stochastic spike trains carries the information, and the identification of orthogonal components in the superposition is done by coincidence detection (since the neuron is essentially a coincidence detector). As soon as a spike belonging to a component is detected in the superposition by comparing it to the reference signal of the component, its existence is detected, as apparent from Fig. 2. Thus, though the signals are stochastic, no time averaging is needed for interpretation of the signal, and the error probability decays exponentially with increasing waiting time.

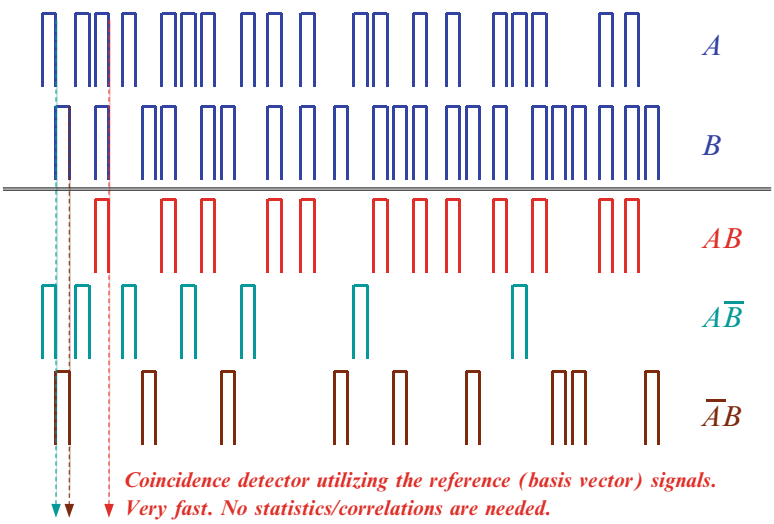


Fig. 2 Demonstration why the brain is so fast notwithstanding that neural spikes are stochastic and occur with low frequency. The existence or non-existence of any of the three orthogonal spike trains in superposition sequences A and B can be quickly observed by coincidence detection, i.e., by neurons

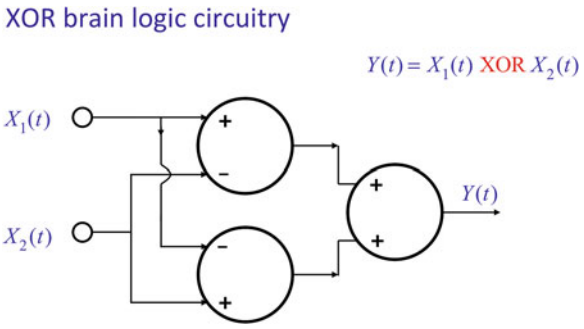


Fig. 3 Neural circuitry realizing the XOR logic function for neural spikes. Circles signify neurons; the "+" and "−" inputs are the excitatory and inhibitory inputs, respectively

3 Example: String Verification by Brain Logic

Below we show a method on how neurons can solve a string verification problem very rapidly and with non-zero error probability that decreases exponentially versus time of operation. The non-brain version of this computational scheme, based on bipolar random telegraph waves, was described in earlier work [3]. In the present paper we propose a brain version and provide the neural circuitry for that, as introduced next.

Suppose that two communicating parts of the brain, called A (Alice) and B (Bob), must verify pairs of N-long bit strings via a slow communication channel within the brain. We represent the possible bits in the strings by $2N$ partially overlapping random neural spike sequences (neuro-bits). Via the brain wiring, Alice and Bob have the ability to access these neuro-bits and use them as a reference signal. Then a *hyperspace neural signal* is generated by making the pairwise XOR function of the N neuro-bit values of the strings at each clock step. For example, comparing only 83 time steps of the hyperspace signals at Alice's and Bob's side provide an error probability of less than 10^{-25}, i.e., a value of the order of the error rate of regular computer bits [3]. Therefore it is enough to create and communicate a small number of signal bits through the information channel. It is important to note that this error probability is independent of the length N of the bit string. Figure 3 shows the neural circuitry to carry out this protocol.

Generalizing this method for the brain may show how intelligence makes reasonable decisions based on a very limited amount of information. Furthermore, our results provide a conceptual explanation why spike transfer via neurons is usually statistical with less than 100 % success rate.

References

1. Kish, L.B., Phys. Lett. A 373 (2009) 911–918.
2. Bezrukov, S.M., Kish, L.B., Phys. Lett. A 373 (2009) 2338–2342.
3. Kish, L.B., Khatri, S., Horvath, T., Eur. J. Phys. B 79 (2011) 85–90.

Time-Dependent Approximate and Sample Entropy Measures for Brain Death Diagnosis

Li Ni, Jianting Cao, and Rubin Wang

Abstract To give a more definite criterion using Electroencephalograph (EEG) approach on brain death determination is vital for both reducing the risks and preventing medical misdiagnosis. This paper presents several novel adaptive computable entropy methods based on approximate entropy (ApEn) and sample entropy (SampEn) to monitor the varying symptoms of patients, and to determine the brain death. The proposed method is a dynamic extension of the standard ApEn and SampEn by introducing a shifted time window. The main advantages of the developed dynamic approximate entropy (DApEn) and dynamic sample entropy (DSampEn) are for real-time computation and practical use. Results from the analysis of 35 patients (63 recordings) show that the proposed methods can illustrate effectiveness and well performance in evaluating the brain consciousness states.

Keywords EEG • Brain death • Entropy measures • Real-time computation • Consciousness

1 Introduction

Brain death is defined as the complete, irreversible and permanent loss of all brain and brainstem functions [1–4]. Under the definition, however, it's hard to conduct brain death judgement precisely for some clinical reasons. Traditional clinical tests are expensive, time consuming, and even dangerous in some cases (e.g., apnea test etc.). To avoid above disadvantages, we have proposed a EEG preliminary examination procedure before the test of spontaneous respiration, which makes the test easier, more effective and brings less risks [5]. To determine quasi brain death

L. Ni (✉) • R. Wang
Institute for Cognitive Neurodynamics, East China University of Science and Technology, Shanghai, People's Republic of China
e-mail: 8nl@163.com

J. Cao
Saitama Institute of Technology, 1690 Fusaiji, Fukaya-shi, Saitama 369-0293, Japan

Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_46

(QBD, where quasi- means that it's a preliminary decision), EEG which is known to us an important clinical tool for the observing brain signals, has been widely available in many countries to evaluate the absence of cerebral cortex function [5–7]. Our research aim is to provide several signal processing tools to determine brain death based on EEG analysis, and help clinicians conduct the diagnosis in the practical operation.

The complexity of nonlinear physiologic signals has been wildly used in evaluating the differences between health and disease states [8]. The information of complexity contained by a physiologic time series directly reflect the state of such physiologic system [9]. The concept of entropy has been extensively available for complexity measures [10,11]. Approximate entropy (ApEn) and sample entropy (SampEn) are effective approaches used in the complexity analysis, and help us have a better understanding of biological system. Pincus first introduced ApEn [11], a set of measures of system complexity closely related to entropy, which has well performed to analyze clinical cardiovascular and other time series. One defect of ApEn, however, is that its statistics lead to inconsistency. Therefore, Richman and Moorman has developed SampEn [12] as an improvement, due to ApEn leads to bias where SampEn does not, which is caused by self matches, so that SampEn agrees with theory much more closely than ApEn over a broad range of conditions. In our studies, we will further illustrate the improved accurracy of SampEn statistics for brain death diagnosis.

This paper presents the time-dependent extensions of ApEn and SampEn, since the static methods can only deal with a limited length of time series whereas the analysis of the data of long recording length is common in a biological system. The analysis on a small segment of the original data may probably cause a larger error and even a fault (for example, the segment is seriously contaminated by noise), causing misleadingness. So that the time-dependent method enables us to gain a more comprehensive and global view into a complex system. On the other hand, our time-dependent method can decrease the amount of calculation in a simulation process and improve the efficiency for an analysis on a full data. As a result, the analysis on the successively changing information contained by a total time series is available for us.

2 Methods/Models

2.1 Approximate Entropy (ApEn)

For a limited time series of N points, $U = [u_1, u_2, \ldots, u_N]$ is formed by the m-dimension vectors $X_i = [u_i, u_{i+1}, \ldots, u_{i+m-1}]$ and $X_j = [u_j, u_{j+1}, \ldots, u_{j+m-1}]$, where $i, j \leq N - m + 1$. The max distance between X_i and X_j can be calculated by

$$d[X_i, X_j] = max_{k=1,2,\ldots,m}[|u_{i+k-1} - u_{j+k-1}|]. \quad (1)$$

Given a threshold r and each $i \leq N - m + 1$, let B_i^m be the number of vectors X_j within r of X_i, and we define

$$C_i^m(r) = \frac{B_i^m}{N - m + 1}, \; where \; i \leq N - m + 1, \tag{2}$$

and $\phi^m(r)$ as mean of $C_i^m(r)$

$$\phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln C_i^m(r). \tag{3}$$

Equation (2) is mainly defined to calculate the possibility that for each Xi and Xj, the two vectors are similar within the threshold r, while Eq. (3) is used to calculate the average.

By finding $\phi^{m+1}(r)$, $ApEn(r, m, N)$ takes the form as

$$ApEn(m, r, N) = \phi^m(r) - \phi^{m+1}(r). \tag{4}$$

This is how ApEn is defined to measure the self-similarity of the time series [11].

2.2 Sample Entropy (SampEn)

SampEn deals with same m-dimension vectors X_i and X_j as defined in ApEn. The distance between two vectors is calculated by Eq. (1). In SampEn, let A_i^m denotes the number of vectors X_j within r of X_i times $(N - m)^{-1}$, for j ranges from 1 to $N - m + 1$ and $j \neq i$, excluding self-matches. We then define A^m as mean of A_i^m, for all $1 \leq i \leq N - m + 1$, and takes the form as

$$A^m = \sum_{i=1}^{N-m+1} \frac{A_i^m}{N - m + 1}. \tag{5}$$

By increasing the space dimension to $m + 1$, and also repeat the steps in Eqs. (1) and (5), we can obtain A^{m+1}. Then SampEn can be obtained by

$$SampEn(m, r, N) = -\ln\left(\frac{A^{m+1}(r)}{A^m(r)}\right). \tag{6}$$

This is how SampEn is defined to measure the self-similarity of the time series [11, 12].

2.3 Time-Dependent Extensions of ApEn and SampEn

The time-dependent ApEn (DApEn) and time-dependent SampEn (DSampEn) have been proposed for the analyzing physiologic time series. The values of ApEn or SampEn are calculated in a set of consecutive time windows marked with their starting instant t with a length t' (selected manually for a sufficient data length and acceptable consuming time) of the whole data with a length T, respectively. Here, the expressions of DApEn and DSampEn are then obtained as $ApEn(m, r, N)_t$ and $SampEn(m, r, N)_t$, where the subscript t represents the time windows for ApEn and SampEn computation. As a result, if the denoted variable t ranges, for example, from t_1 to t_2 with a step length λ (set $\lambda = t'$), the values of ApEn or SampEn are obtained respectively in several non-overlapping windows. Then DApEn is defined by

$$ApEn(m, r, N)_t = \phi_t^m(r) - \phi_t^{m+1}(r), \tag{7}$$

while DSampEn is defined by

$$SampEn(m, r, N)_t = -\ln\left(\frac{A_t^{m+1}(r)}{A_t^m(r)}\right). \tag{8}$$

3 Experiment

In our present study, the EEG experimental protocols were executed in the ICUs of a hospital. The EEG recording instrument was a portable NEUROSCAN ESI-32 amplifier associated with a note computer. During EEG recording, a total of nine electrodes were placed on the forehead of the patients. Six channels were placed at corresponding electrodes (Fp1, Fp2, F3, F4, F7, F8), two electrodes (A1, A2) were placed on the ears as reference and an additional channel, GND, served as the ground. Experimental data were obtained from 35 patients (19 male, 16 female) of ages ranging from 17 to 85 years old; 19 of them were diagnosed to be comatose and the left were brain deaths. The average length of the EEG recordings from these patients was about 5 min.

3.1 Results for DApEn and DSampEn

For the real-time application such as monitoring a state of the patient, it is necessary to introduce time-dependent-based analysis to explore the brain wave activity changes of the patients over time. As shown in Fig. 1, over the time-coordinate (0–800 s) of EEG signals, ApEn and SampEn in each second is calculated in a 5 s

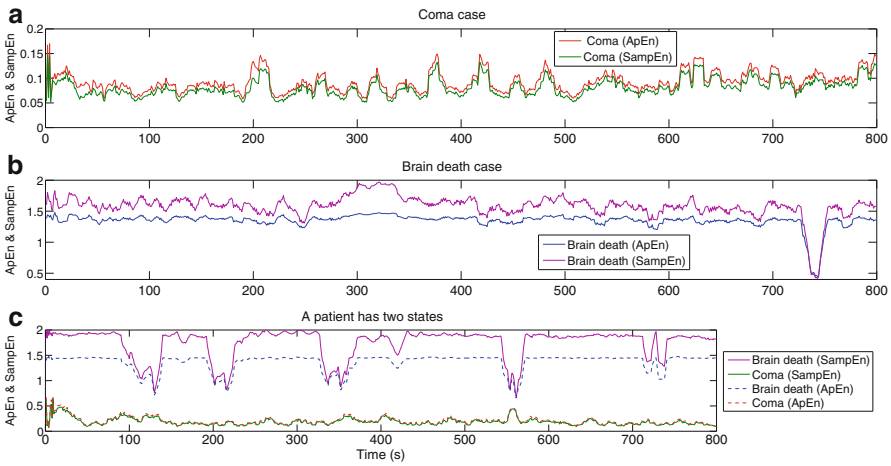


Fig. 1 time-dependent complexity measure for two different cases (**a**) vs. (**b**) and a patient has two states (**c**)

non-overlapping moving time window. Results of DApEn and DSampEn for a coma case (a), a quasi-brain-death case (b) and a case that the patient behaved from coma to quasi-brain-death (c).

Under the same experimental environment, the results of DApEn and DSampEn are obtained by Eqs. 7 and 8. To obtain a more smooth curve, the moving average method is applied to decrease the high frequency part by calculating the average of every ten points. For the same coma case, values of SampEn (green) remain low over time, while values of ApEn (red) remain slightly higher than that of SampEn. For the same quasi-brain-death case, SampEn assign a higher value (purple) than ApEn (blue). This indicates a more powerful capability for DSampEn to classify the two brain consciousness states than for ApEn. Around 750 s, the huge fluctuation is caused by the serious contamination of noise. Moreover, the results of DApEn and DSampEn of the certain patient whose coma state and brain death state are both recorded are plotted in Fig. 1c. In this case, this patient's two states are well discernible because of a huge difference of the values of both DApEn and DSampEn, however, the results of coma state are slightly different and several data segments are under the influence of interfering noise.

From all the obtained results, we firmly believe that complexity of the coma and the brain deaths can be well used for brain consciousness determination. The plotted time-dependent ApEn or SampEn indicates the state of a patient and such time-dependent methods also helps monitor the trend of a patient's state, with which clinic can carry out emergency medical care before danger. But in this paper the methods applied to predict are beyond our scope. So, with the help of our time-dependent algorithm, the on-line EEG preliminary brain death determination system comes into reality.

4 Conclusions/Discussions

The obtained results show that the states of coma and brain death are discernible by calculating the values of ApEn and SampEn, and the proposed time-dependent method illustrates a feasible solution for practical use of brain death diagnosis, because our simulation is identical to the clinical results. These methods may be applied to analysis of other physiologic time series as a reference.

Acknowledgements This work was supported by KAKENHI (21360179, 22560425) (JAPAN), also supported by the Key Project of National Science Foundation of China (11232005) (CHINA).

References

1. Chen, Z., Cao, J., Cao, Y., Zhang, Y., Gu, F., Zhu, G., Hong, Z., Wang, B., Cichocki, A.: An empirical EEG analysis in brain death diagnosis for adults. Cognitive Neurodynamics, 2 (2008), pp. 257–271.
2. Cao, J.: Analysis of the Quasi-Brain-Death EEG Data Based on a Robust ICA Approach. Cognitive Neurodynamics, 4253 (2006), pp. 1240–1247.
3. Shi, Q., Zhou, W., Cao, J., Tanaka, T., Wang, R.: Brain-Computer Interface System Using Approximate Entropy and EMD Techniques. Cognitive Neurodynamics, 6146 (2010), pp. 204–212.
4. Yang, K., Cao, J., Wang, R. and Zhu, H.: Analyzing EEG of quasi-brain-death based on dynamic approximate entropy measures. Chinese Journal of Biomedical Engineering, 30 (1) (2011), pp. 27–33.
5. Cao, J., Chen, Z.: Advanced EEG signal processing in brain death diagnosis. In: Signal processing techniques for knowledge extraction and information fusion. Springer. (2008), pp. 275–297.
6. Li, L., Saito, Y., Looney, D., Cao, J., Tanaka, T., Mandic, D. P.: Data fusion via fission for the analysis of brain death. In: Evoling intelligent systems: methodology and applications. Springer. (2010), pp. 279–320.
7. Eelco, F., Wijdicks, M.: Brain death worldwide. Neurology, 58 (2002), pp. 20–26.
8. Zhang, T., Xu, K.: Multifractal analysis of intracranial EEG in epilepticus rats. Lecture Notes in Computer Science, 7062 (2011), pp. 345–362.
9. Taylor, R. M.: Reexamining the definition and criteria of death. Seminars in Neurology, 17 (1997), pp. 265–271.
10. Hu, M., Liang, H.: Adaptive Multiscale Entropy Analysis of Multivariate Neural Data. IEEE Transactions on Biomedical Engineering, 59 (2012), pp. 12–16.
11. Pincus, S. M.: Approximate entropy as a measure of system complexity. Proceedings of the National Academy of Sciences of the United States of America, 88 (1991), pp. 2297–2302.
12. Richman, J. S., Moorman, J. R.: Physiological time-series analysis using approximate entropy and sample entropy. AJP-Heart and Circulatory Physiology, 2278 (2000), pp. 2039–2050.

Preliminary Study on EEG-Based Analysis of Discomfort Caused by Watching 3D Images

Jie Yin, Jingna Jin, Zhipeng Liu, and Tao Yin

Abstract The objective here is to investigate the effects of 3D images on viewers, study the relationship between power spectrum features of EEG and discomfort of the viewers, and try to figure out an objective index of discomfort caused by watching 3D images. Ten volunteers were divided into two groups, and EEG were recorded in both groups before and during the experiment of watching 3D and 2D images. Power spectral density (PSD) and variance for repeated measures were extracted to analyze the objective index of discomfort. There was an increase in EEG power spectral density in the frontal and occipital areas during the experiment, and especially the phenomena in 3D were more obvious than those in 2D. Meanwhile, statistical methods showed a significant difference between the two experimental groups. We conclude that the EEG power spectral density is strongly correlated with visual discomfort. PSD increases while the visual discomfort aggravates. The results of the EEG PSD are also consistent with the statistical results. Further, the EEG power spectral density is expected to serve as the objective index for analyzing the discomfort caused by watching 3D images.

Keywords 3D images • Discomfort • EEG analysis • Power spectral density

1 Introduction

The emergence of three-dimensional television (3D TV) brings the viewer a whole new experience. This technology not only improves the image quality, but also changes the figures in images [1]. 3D TV uses the angel difference while people

J. Yin • J. Jin • Z. Liu
Institute of Biomedical Engineering, Peking Union Medical College, Chinese Academy of Medical Sciences, Tianjin, People's Republic of China

T. Yin (✉)
Institute of Biomedical Engineering, Peking Union Medical College, Chinese Academy of Medical Sciences, Tianjin, People's Republic of China

IEEE Member, Beijing, People's Republic of China
e-mail: bme500@163.com

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_47

observe objects with eyes to identify the distance of objects. Stereopsis is the perception of depth that is constructed based on the difference between the left and right retinal images. The brain fuses the two images and, from retinal disparity, i.e., the distance between corresponding points in these images, the brain extracts relative depth information. Hence, in comparison to 2D technologies, 3D TV provides the user with more exact perception of depth and thus the operation can be performed with higher accuracy and promptness.

Stereoscopic display technology has the strong advantages of stereo and immersion, but if you watch the three-dimensional images for a long time (more than 30 min), some viewers will suffer the uncomfortable symptoms such as visual fatigue, dizziness, which, to a certain extent, affects the popularity and development of three-dimensional stereoscopic display technology. Health and safety are the problems which need to be solved urgently in 3D stereoscopic display technology. The viewers have to move their eyes and quickly adjust to the focus changes, thus, this could cause viewers suffering from visual fatigue and other discomfort.

At present, methods to measure the discomfort caused by 3D images could be clustered as explorative studies, psychophysical scaling, questionnaires, and electrical signal measurement. Presently, most brain activity research in depth perception concentrates on fundamental issues, such as identifying the specific pathways for binocular vision [2, 3].

Explorative studies can be used in the context of stereoscopic displays to make a fundamental evaluation [4]. Psychophysical scaling enables engineers to enhance and optimize their systems based on quantified perceptual attributes such as image quality and visual discomfort. Questionnaires have been widely applied as a specific means to determine the degree of visual discomfort [5, 6]. In clinical research, questionnaires are efficient to evaluate the degree of asthenopia due to visual deficits. Li et al. [7] used background EEG (Electroencephalogram) and ERP (event related potentials) to measure visual fatigue. The frequency spectrum of the background EEG signals is known to analyze the state of stress; i.e., high frequencies starting at ± 12 Hz denote stressful situations. Though the P300 latency of the event related potentials is delayed in stressful situations, they found that the delay was much stronger for the P700 latency. Results concluded that the power of the spectrum of the background EEG as well as the delay in the P700 latency depended on binocular parallax and viewing time, which was confirmed by subjective assessments. Hence, delays in the transmission of visual information measured with EEG seem to be an appropriate method for visual fatigue. The measurements on discomfort such as visual fatigue caused by 3D images are still preliminary.

In this study, we tried to compare the differences of the EEG in 2D and 3D presenting situations in order to provide a viable method to measure the visual discomfort. We collected the subjects' EEG signals, analyzed power spectrum density of the signals and try to discuss the specialty of EEG in 3D images on the view of power.

2 Procedure

2.1 Subjects and Facilities

The subjects participated in the research are ten graduates aged 20–30 years, with normal vision and health. In order to make comparison, all subjects viewed the images in 2D and 3D condition respectively.

In the experiment, the facilities are comprised of a 55-in. SAMSUNG 3D TV 55C8000XF, Blu-ray DVD player, 3D active glasses and two 3D cartoons. The equipment to collect the EEG is Neuroscan 64 channels platform, SynAmps amplifier system. We set the sampling rate at 1,000 Hz. Electrode positions were a subset of the international 10–20 system sites.

2.2 Design and Procedure of the Experiment

In the real experimental scene, we set the viewing distance as 3 m, as shown in Fig. 1. The procedure of the experiment is following that: First, the experimenters made introduction to subjects the specific steps and measurements. Second, we positioned the electrodes on the subjects and collected the EEG signals before

Fig. 1 Real experimental scene of watching 3D images

presentation, with the eye closed and opened 1 min respectively. Third, subjects viewed the cartoons without unnecessary movements in 80 min. After the presentation, subjects could have a short rest for 15 min. Finally, we made a questionnaire study for the subjects.

2.3 Math

We chose the Classical Power Spectral Density Estimation to analyze the signals collected. On the MATLAB 2010b platform, we used the periodogram function to calculate the mean power spectrum. Then, we calculated the percentages of each band in the total power spectrum and fitted straight lines to search the trends totally. Finally, both the power spectrum data and percentages are analyzed using ANOVA (a repeated measures analysis of variance) on SPSS 18.0 platform.

3 Result

We set the EEG in terms of four frequency bands including delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz). Each band's power spectrum was analyzed. From the analysis, the subjects' discomfort was obvious during the 50–60 min, as shown in Figs. 2, 3, and 4.

In Fig. 2, the PSD ratio of delta band in 3D condition (red line) ascends more obviously than 2D condition (blue line). Since delta band denotes the excessive fatigue and lethargy [8], the result concluded that the 3D condition could cause more obvious discomfort in frontal region.

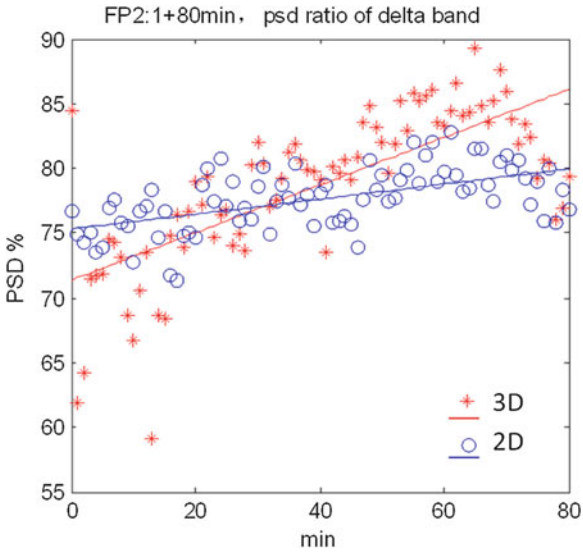


Fig. 2 Average of PSD ratio of delta band in FP2 lead

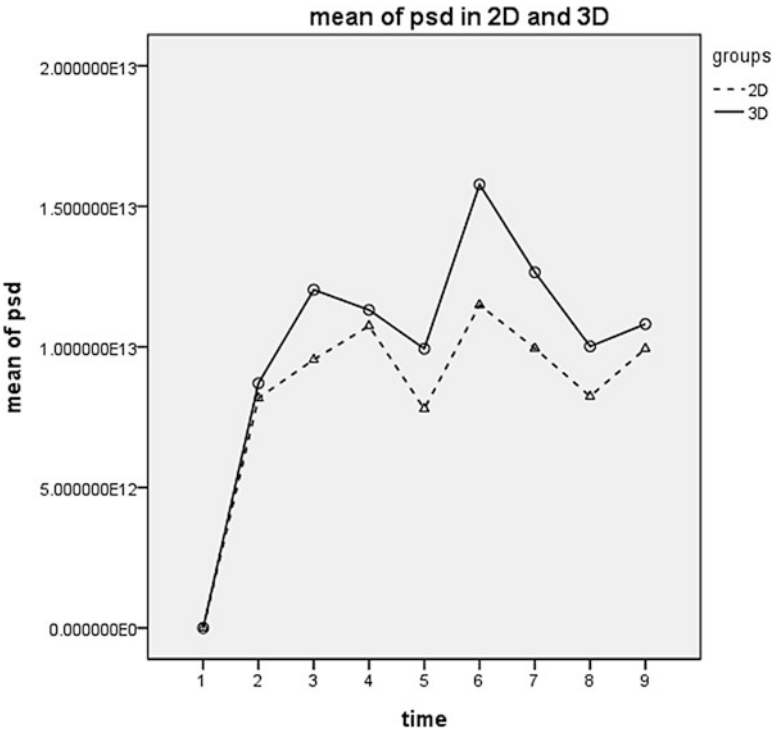


Fig. 3 Average power spectra of delta band in Oz lead, EEG signals were recorded during watching 2D image and 3D image respectively

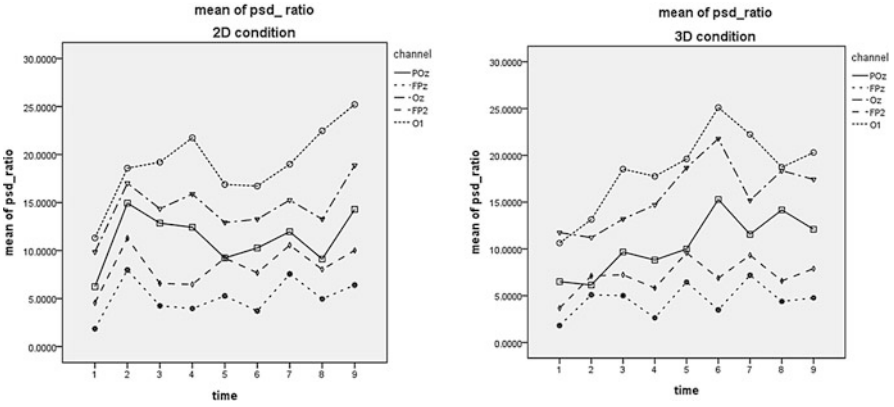


Fig. 4 Power spectra ratio of gamma band, EEG signals were recorded from two groups watching 2D image and 3D image respectively

In Fig. 3, we segmented the viewing EEG for every 10 min and calculated the mean power spectrum. Then, the analysis of variance of repeated measure was used on the SPSS 18.0 platform. The points in horizontal axis (time) 1, 2, 3, 4, 5, 6, 7, 8, 9 represent 0, 10, 20, 30, 40, 50, 60, 70, 80 min respectively. In comparison, the accelerating rate in 3D condition was more than 2D condition with the alpha band data in Oz channel. Also, there is significant difference between the two conditions ($p < 0.05$).

As shown in Fig. 4, the trends of the channels FPz and FP2 in frontal area are consistent in 2D and 3D condition. Also, the channels POz, Oz, and O1 in occipital area are consistent in two conditions, and the changes in 3D condition was more obvious than 2D condition, both of which maximized after 50 min.

There is study showing that high frequency band such as beta frequencies (i.e. frequencies between 12 Hz and 30 Hz) are observed in a stressful situation while low band frequencies are observed when people rest or when they are attentive [9]. The analysis in gamma band in Fig. 4 revealed that subjects felt physiological tension and stress after viewing 3D images for a long time.

4 Conclusion

The results above revealed that subjects were likely to suffer fatigue when viewing 3D images for a long time. The difference in EEG between 2D and 3D condition was more obvious in frontal and occipital areas. This paper concentrated on the discussion about the ascending trends of alpha band in the FP2 channel (representative of the frontal area). Besides, this study discussed the power spectral density in Oz channel. Comparison between the two conditions showed that both the PSD accelerated and the trend in 3D condition was more obvious than 2D condition.

Mental fatigue refers to people's functional decline in the daily work or learning process for a long time mental work [10, 11]. Several years study in mental fatigue revealed that EEG become the most widespread measure to analyze the changes in the central nervous system, hailed as the 'gold standard' to monitor fatigue.

Using EEG to measure the subjects' discomfort caused by viewing 3D images is a new approach, which needs more discussion and research. In this study, the experimental condition needed to be improved and standardized. Also, the design of the experiment needed to be further optimized and exclude the individual differences. Besides, the data could be deeply analyze and process in order to obtain an objective method to measure the discomfort caused by viewing 3D images. The goal of the study was to provide some useful reference for measurement of the EEG-based Analysis of Discomfort Caused by watching 3D images.

References

1. W. A. IJsselsteijn, Presence in Depth, Ph.D. dissertation, Dept. Human Technology Interaction, University of Technology, Eindhoven, The Netherlands (2004).
2. R. Blake and N. K. Logothetsis, "Visual competition", Nat. Rev. Neurosci.3, 13–21(2002).
3. J. Parker, "Binocular depth perception and the cerebral cortex", Nat. Rev. Neurosci. 8, 379-391(2007).
4. L. M. J. Meesters, W. A. IJsselsteijn, and P. J. H. Seuntiens, "A survey of perceptual evaluations and requirements of three-dimensional TV", IEEE Trans. Circuits Syst. Video Technol. 14, 381-391 (2004).
5. J. E. Sheedy, J. Hayes, and J. Engle, "Is all asthenopia the same?", Optom. Vision Sci. 80, 732-739 (2003).
6. M. Emoto, T. Niida, and F. Okana, "Repeated Vergence Adaptation Causes the Decline of Visual Functions in Watching Stereoscopic Television", J. Disp. Technol. 1, 328-340 (2005).
7. H. Li, J. Seo, K. Kham, and S. Lee, "Measurement of 3D visual fatigue using event-related potential (ERP): 3D oddball paradigm", unpublished results presented at the 3D-TVCON'08, Istanbul, Turkey, May28-30 (2008).
8. Lal S K L, Craig A. A Critical Review of the Psychophysiology of Driver Fatigue [J]. Biological Psychology, 55(3): 173-194 (2001).
9. H.K. Gomarus, M. Althaus, A.A. Wijers, and R. B. Minderaa, "The effects of memory load and stimulus relevance on the EEG during a visual selective memory search task: An ERP and ERD/ERS study." Elsevier, pp. 871-884 (2006).
10. Linder D, Frese M, Meijmam TF. Mental, fatigue and the control of mental process: effects on perseveration and planning [J]. Acta Psychologica, 113 (1): 45-65 (2003).
11. Leonard JT, Rebekah K, Karla K, et al. EEG-based estimation of mental fatigue [C]. In Caldwell JA, Wesensten NJ, eds. Proceedings of SPIE Vol. 5797: Biomonitoring for Physiological and Mental Performance During Military Operations. Orlando: SPIE: 105-115 (2005).

EEG-EMG Analysis on Corticomuscular Coherence of Middle Finger Flexion Tasks

Tianchen Zhai, Cheng Qi, Rui Xu, Anshuang Fu, Lixin Zhang, Xin Zhao, Peng Zhou, Hongzhi Qi, Baikun Wan, Xiaoman Cheng, Weijie Wang, and Dong Ming

Abstract Over the last few years much research has been devoted to investigate the synchronization between cortical motor and neuromuscular activity as electroencephalogram (EEG) and surface EMG (sEMG), which could elicit a new research idea in the field of sports medicine and rehabilitation engineering. Corticomuscular coherence (CMC) is a method combining the brain with muscles, which indicates the coordination of motor control. In order to examine modulation of the CMC in diverse motion modes, ten healthy young right-handed adults performed voluntary finger flexion, simulated finger flexion and motor imagination tasks at short intervals of 2 s. EEG and sEMG were recorded simultaneously from the primary motor cortex and the musculusflexorperforatus, respectively. The extreme points of improved cross-correlation coefficient, and partial directional coherent (PDC) algorithm were used to compare and analyze the coherences under different motion tasks. We found that the strongest coherence showed up during stimulation task. The results suggest clinical reference to build rehabilitative treatment protocol in the future.

Keywords Corticomuscular coherence • Neuromuscular activity • Median nerve • Partial directed coherence

T. Zhai • C. Qi • R. Xu • A. Fu • L. Zhang • X. Zhao • P. Zhou • H. Qi • B. Wan • D. Ming (✉)
Department of Biomedical Engineering, Laboratory of Neural Engineering and Rehabilitation, Tianjin University, Tianjin, People's Republic of China
e-mail: richardming@tju.edu.cn

X. Cheng
School of Science, Tianjin University of Technology, Tianjin, People's Republic of China

W. Wang
Department of Orthopaedics and Traumatology, Ninewells Hospital, University of Dundee, Tianjin, People's Republic of China

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_48

1 Introduction

A groundbreaking research on monkeys found a contralateral synchronous phenomenon of the cortical motor system [1]. The corticomuscular coherence (CMC) in humans was first systematically studied in 1995. They interpreted that The synchronized cortical activity is coupled with electromyographic (EEG) in beta rage (15–30 Hz) for the implication of motor unit synchronization [2]. Since then, the CMC analysis has been used not only in the function of cortical control of movement in general population [3], but also in pathological analysis in corticomuscular diseases [4, 5]. Magnetoencephalographic (MEG) and EEG are often adopted in corticomuscular studies for their high time resolution. Through a power spectra analysis, an increasing synchronization occurred in beta-range during the motor imagery task [6]. Several corticomuscular coherence studies have focused on task-dependent research. As revealed by Hashimoto, EEG and electromyography (EMG) correlates strongly during the motor execution (ME) and motor imagery (MI) tasks in a frequency band of 14–30 Hz [7]. Perez found that the corticomuscular coherence can increase with a training session in beta frequency band [8]. Johnson and Shinohara demonstrated that the corticomuscular coherence shows different intensity in alpha- and beta- frequency rage for different age groups [9]. Another research shows that concurrent tasks can cause a decreasing of beta band coherence, which is most likely because of the distractions [10]. Some of the mechanisms between central nervous system and motor system are as yet unknown.

Neuromuscular activities are always accompanied with the release of corresponding electrophysiological signal. Accordingly, the method of analyzing electrophysiological signals will be quite effective. In the pre-sent study, we try to reveal the difference of corticomus-cular interaction during different states of people by coherence analysis. To this aim, EEG data (C3, C4) and EMG data was sampled during the subjects' motor imagery, voluntary movement and stimulated movement under short intervals.

2 Method

2.1 Subjects and Procedure

Subjects were ten right handed healthy volunteers (age range 23–27 years, mean 24.2 ± 0.8), all of whom were undergraduate and graduate students. All subjects were free of neurological or psychiatric disorders or medications known to adversely affect EEG and EMG recording. The purpose of the study and the procedure of experiment were explained to each subject before preparation for recording. After this, subjects were seated in a shielded room in a comfortable chair with a straight

back. Each subject was asked to relax and attempt to achieve three tasks which included the imagination (I), voluntary (V) and stimulation (S) of middle finger flexion with a period of 2 s.

2.2 Data Collection

EEG data were recorded from bipolar electrodes placed on or near sites C3, C4 using the International 10–20 system of electrode placement. A ground electrode was placed on site A1 on the left mastoid and reference electrode on A2 on the right mastoid. The sampling rate was 16384 Hz. Two FES electrodes which attached to the subject's arm outputted pulses in the median nerve under short-interval stimulated motion mode.

The EEG and sEMG signals were band-pass filtered (EEG: 5–45 Hz, sEMG: 5–45 Hz). And data were down-sampled into 512 Hz. Additionally, stimulus pulses generated during stimulation task were removed.

2.3 Preprocessing

Keeping a far distance between EMG collection electrodes and stimulating electrodes can prevent EMG and stimulus signals from alias. Since the FES electrodes were placed quite close to the data-collection electrodes, the amplitude of EMG can be less than that of FES pulse with several orders of magnitude. A preprocessing which can remove the large pulse noise from FES is essential. The two-stage peak detection algorithm was adopted in the preprocessing of EMG generated by stimulated movement to eliminate the strong stimulus artifact in the EMG signal [11, 12].

The two-stage peak detection algorithm aims at wiping off the stimulus artifact as well as keeps EMG signals intact. The positive and negative peaks of stimulus signals were detected high level and low level of peaks which were set based on absolute value of EMG data. Stimulating pulse is a huge wave with voltage spike; the amplitude and pulse wide are under control of some factors which include output current, amplifier settings, electrode placement, and stimulus patterns. As the FES pulse will be removed when detected, the removal can be effective to both single pulse and waves which meet a given phase relation (see Fig. 1).

2.4 Partial Directed Coherence

To represent the Granger causality on a frequency domain method, partial directed coherence is proposed [13].

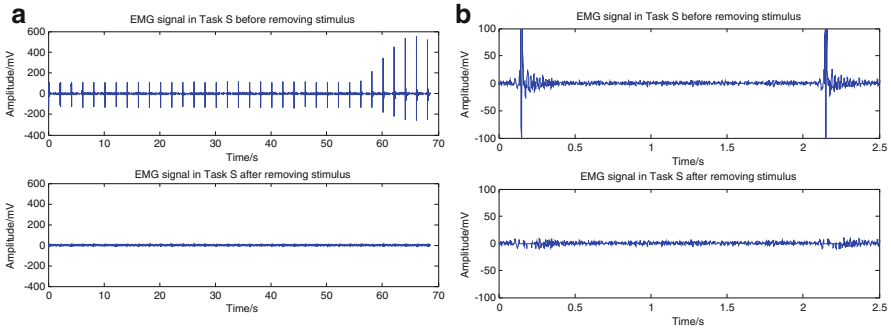


Fig. 1 An example of the stimulus artifacts removal in a stimulation task. (**a**) The first figure shows the communicated sEMG, and the other one is the filtered result. (**b**) In order to see this result clearly, this is got from the part surrounded by the *red rectangle* maximized in (**a**)

Let $x = (x(t))_{t \in z}$, when $x(t) = (x_1(t), \cdots, x_n(t))'$ is a stable n-dimensional time series that averages zero. Then a short p order VAR model as follows:

$$x(t) = \sum_{r=1}^{p} a(r)x\,(t-r) + \varepsilon(t) \tag{1}$$

Let the *pth* order vector autoregressive model describes the set of simultaneously observed second-order stationary time series:

$$\begin{bmatrix} x_1(t) \\ \vdots \\ x_Q(t) \end{bmatrix} = \sum\nolimits_{k=1}^{p} A_k \begin{bmatrix} x_1\,(t-k) \\ \vdots \\ x_Q\,(t-k) \end{bmatrix} + \begin{bmatrix} w_1(t) \\ \vdots \\ w_Q(t) \end{bmatrix} \tag{2}$$

The linear relationship between x_j and x_i are described at the by the coefficients matrix A_k *kth* past lag, and the $\varepsilon(t)$ is a multivariate Gaussian white noise process.

From the joint spectral density estimate $x = (x(t))_{t \in z}$ through:

$$\det\left(I - a(1)Z - \ldots - a(p)Z^p\right) \neq 0 \tag{3}$$

$$A_k = \begin{bmatrix} a_{1,1}(k) & \cdots & a_{Q,1}(k) \\ \vdots & \ddots & \vdots \\ a_{Q,1}(k) & \cdots & a_{Q,Q}(k) \end{bmatrix} \tag{4}$$

Do not like the original granger causality that is double variable time series, partial directed coherence is based on the treatment of multivariable autoregressive time series. To describe Granger causality relations, the concept of partial directed coherence is introduced. So for the autoregressive process of order p partial directed coherence are defined as follows:

$$\left|\pi_{i\leftarrow j}\left(\omega\right)\right| = \frac{\left|\mathbf{A}_{ij}\left(\omega\right)\right|}{\sqrt{\sum_{k}\left|\mathbf{A}_{kj}\left(\omega\right)\right|^{2}}} \tag{5}$$

Condition (2) ensures that the denominator is always positive, so the equation can well define directional coherent. As can be seen from the definition, if and only if all the coefficients $a_{ij}(r)$ are equal to zero, $|\pi_{i\leftarrow j}(\omega)|$ for all frequencies ω do not exist, thus there is no Granger causal connection between x_j and x_i. This means that the partial directional coherent $|\pi_{i\leftarrow j}(\omega)|$ delivered a measurement for the directed linear influence in the frequency domain. Also, because Eq. 5 is a normalized form, therefore partial directional coherent takes value in [0, 1]. It compared the influence of the past for and the impact of the past with other variables, therefore, for a given signal source, partial directional correlation analysis can be arranged according to the intensity of the impact of variables.

3 Result

The result of improved coherence of one subject is showed in Fig. 2. It is difficult to tell the difference among these three tasks. Then, we did some statistical analysis of the significant coherence. The frequency bands of EEG signals are divided into three typical bands: alpha bank (8–14 Hz), beta band (15–30 Hz), gamma (30–40 Hz), and correlation coefficients in the top fifth percentile were defined as extreme points of improved cross-correlation coefficient. Statistical analysis was performed to assess the coherence between EEG and sEMG in different tasks. The number of extreme points was calculated in each trial and then added across all trials, which could indicate the coherency level. Figure 3a–c shows the statistical coherence in three tasks from ten subjects. The histogram from Fig. 3d shows the average across subjects in different frequency bands and tasks. The EEG-EMG correlation can be measured through PDC. The EEG signal of C3 and C4 channels are the most relevant to the hand motion.

EEG from Channel C4 to EMG (EEGC4-EMG) can be presented as the PDC value, which helps us analyze the coherence between EEG and sEMG.

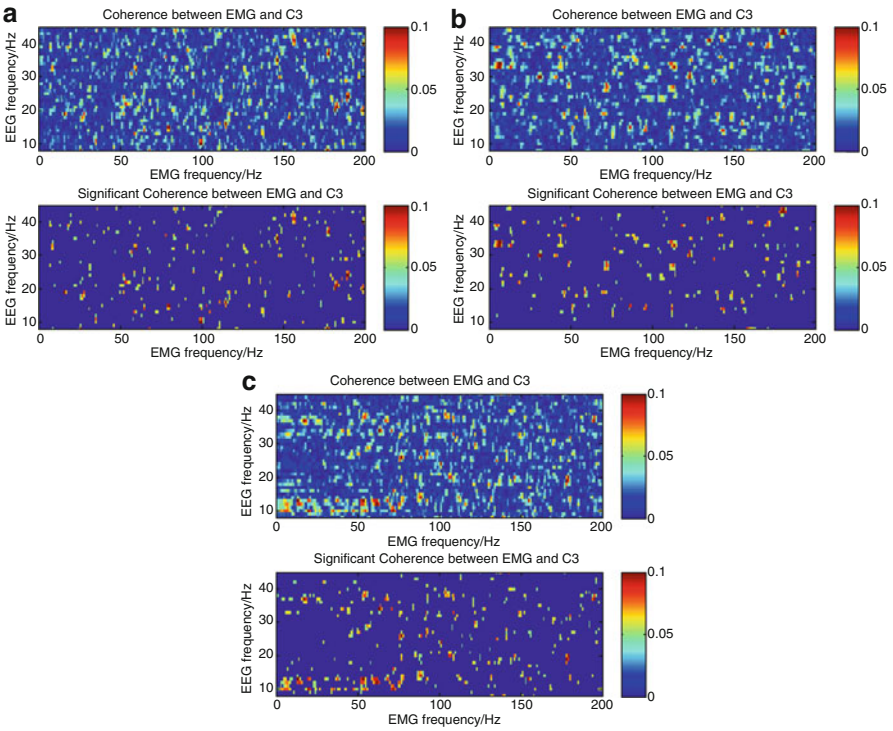


Fig. 2 An example of improved coherence of one subject in the three tasks

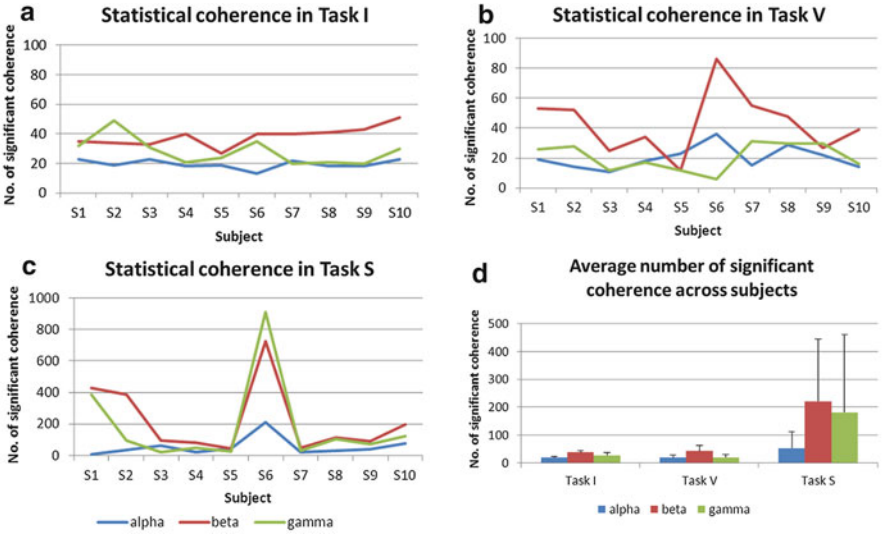


Fig. 3 (**a**)–(**c**) are the mean of point number with significant coherence among ten subjects during three tasks. (**d**) The mean of point number with significant coherence among ten subjects during three tasks

4 Conclusions

The stimulus removal result from Fig. 1 shows that the stimulus disturbance is large relative to sEMG, which could damage the coherence result if not removed. The EMG data without the stimulus disturbance could be seen clearly in maximized part. In Fig. 2, there is no strong coherence found in Task imagery. The reason may be that there is no actual movement during this task. So the EMG recorded was resting EMG without any motor information. It can make sense that EMG was not correlated with EEG during task I. The conclusion was similar with that found in prior studies. Second, not every subject showed strong EEG-EMG coherence in task V and S. This is also the bottleneck of the coherence problem. The selection of subjects and proper training will be taken into consideration in the future to enlarge and steady the coherence result.

From the result of Fig. 3d, it indicates the coherence is stronger in Task S than that in Task V. This may because of the synchronization between EEG and EMG under the electrical stimulation. But we can't exclude the influence caused by the stimulus disturbance that has not been removed overall. The remained stimulus in EEG and EMG may lead to high EEG-EMG coherence for the sources of the signals which are the same. Large standard deviation may come from the different degrees of this stimulus removal. In addition, beta band is the main frequency band for EEG-EMG coherence during this finger flexion movement because the number of significant coherence dots is larger than that of alpha and gamma band in every task.

In sum, the experiment with three tasks was designed. Improved coherence function and PDC was employed. At last, we found the strongest coherence showed up during stimulation task. This proposed evidence for the therapy of paralysis with functional electrical stimulation.

Acknowledgments This research was supported by National Natural Science Foundation of China (No. 81222021, 31271062, 61172008, 81171423, 51007063), National Key Technology R&D Program of the Ministry of Science and Technology of China (No. 2012BAI34B02) and Program for New Century Excellent Talents in University of the Ministry of Education of China (No. NCET-10-0618).

References

1. Venkatesh N. Murthy and Eberhard E. Fetz: Coherent 25-to 35- Hz oscillations in the sensorimotor cortex of awake behaving monkeys. Proceedings of the National Academy of Sciences. (1992) 5670–5674.
2. B. A. Conway, D. M. Halliday, S. F. Farmer, U. Shahani, P. Maas, A. I. Weir and J. R. Rosenberg: Synchronization between motor cortex and spinal motoneuronal pool during the performance of a maintained motor task in man. Journal of Physiology (1995) 917–924.
3. James M. Kilner, Stuart N. Baker, Stephan Salenius, Riitta Hari and Roger N. Lemon: Human cortical muscle coherence is directly related to specific motor parameters. Journal of Neuroscience (2000) 8838–8845.

4. Jan Raethjen, Michael Lindemann, Matthias Dümpelmann, Roland Wenzelburger, Henning Stolze, Gerd Pfister, Christian E. Elger, Jens Timmer, Günther Deuschl: Corticomuscular coherence in the 6–15 Hz band: is the cortex involved in the generation of physiologic tremor? Experimental Brain Research (2002) 142:32–40.
5. Luis Patino, Wolfgang Omlor, Vihren Chakarov, Marie-Claude Hepp-Reymond and Rumyana Kristeva: Absence of gamma-range corticomuscular coherence during dynamic force in a deafferented patient. Journal of Neurophysiology (2008) 1906–1916.
6. Neuper C, Scherer R, Reiner M, Pfurtscheller G: Imagery of motor actions: Differential effects of kinesthetic and visual-motor mode of imagery in single-trial EEG. Cognitive Brain Research (2005) 668-677.
7. Yasunari Hashimoto, Junichi Ushiba, Akio Kimura, Meigen Liu, and Yutaka Tomita: Correlation between EEG-EMG coherence during isometric contraction and its imaginary execution. Acta Neurobiol Experimentalis (2010) 70: 76-85.
8. Monica A. Perez, Jesper Lundbye-Jensen and Jens B. Nielsen: Changes in corticospinal drive to spinal motoneurones following visuo-motor skill learning in humans. The Journal of Physiology (2006) 843–855.
9. Ashley N. Johnson and Minoru Shinohara: Corticomuscular coherence with and without additional task in the elderly. Journal of Applied Physiology (2012) 970–981.
10. Johnson AN, Wheaton LA and Shinohara M.: Attenuation of corticomuscular coherence with additional motor or non-motor task. Clinical Neurophysiology (2011) 356–363.
11. O'Keeffe DT, Lyons GM and Donnelly AE: Stimulus artifact removal using a software-based two-stage peak detection algorithm. Journal of Neuroscience Methods (2001) 109: 137–145.
12. Baccala AL and Sameshima K: Stimulus artifact removal using a software-based two-stage peak detection algorithm. Biological Cybernetics (2001) 85: 463–474.
13. Schelter B, Winterhalder M and Eichler M: Testing for directed influences among neural signals using partial directed coherence. Journal of Neuroscience Methods (2005) 152: 210–219.

Part VIII
Signal Transmission at Cellular and Network Levels

Bursting in Two-Coupled Cells Network in the Pre-Bötzinger Complex

Lixia Duan, Dandan Yuan, and Xi Chen

Abstract The persistent sodium play an important role in respiratory rhythm generation in neurons of the pre-Bötzinger complex. We study the bursting patterns and transition mechanisms in a two-cell network model by both fast/slow decomposition and two-parameter bifurcation analysis. The results show the transitions of different types of bursting are not synchronized. This work provide a future understanding of that how NaP current works in respiratory rhythm generation.

Keywords Respiratory rhythm generation • Pre-Bötzinger complex • Bursting patterns • Two-cell network • Bifurcation analysis

1 Introduction

A network of oscillatory bursting neurons in the pre-Bötzinger complex (pre-BötC) of the mammalian brain stem is closely related to inspiratory phase of the respiratory rhythm generation [11]. Most of these neurons are endowed with a persistent sodium (NaP) current and a calcium activated nonspecific cationic (CAN) current, which have been shown potentially contribute to the generation of rhythmicity within the pre-BötC [4, 9].

The anti-phase and in-phase oscillations are common seen when the coupling is introduced. Some researches have shown that these multi-phase oscillations are important rhythm in the pre-BötC [1], but how multi-phasic oscillations depend on the persistent sodium currents in the pre-BötC is less investigated.

We have studied the pattern-dependent mechanisms with g_{Na} and g_K changing in the two-cell network model in the pre-BötC [5]. In this paper, we investigate the effects of g_{NaP} on bursting patterns and transition mechanisms. This paper is divided in four sections. The model is described in Sect. 2. In Sect. 3, the dynamic range of the firing patterns in two-parameter (g_{NaP}, g_{CAN})-space is explored. The parameter space can be divided into four regions according to multi-phasic oscillations.

L. Duan (✉) • D. Yuan • X. Chen
College of Science, North China University of Technology, Beijing, People's Republic of China
e-mail: duanlx@ncut.edu.cn

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_49

The bursting patterns and transition mechanisms between these regions with g_{NaP} varying are investigated. The conclusions are given in the last section. All the bifurcation analyses in this paper are performed using XPPAUT [6].

2 Model Description

We use a two-cell network model which is a combination of the single-compartment model developed by Park and Rubin [10], and Butera [2,3]. The model is described as follows:

$$\dot{v}_i = (-I_{NaP} - I_{CAN} - I_{Na} - I_K - I_L - I_{syn-e})/C, \tag{1}$$

$$\dot{h}_i = \varepsilon(h_\infty(v_i) - h_i)/\tau_h(v_i), \tag{2}$$

$$\dot{n}_i = (n_\infty(v_i) - n_i)/\tau_n(v_i), \tag{3}$$

$$\dot{s}_i = \alpha_s(1 - s_i)s_\infty(v_i) - s_i/\tau_s, \tag{4}$$

where $i, j = 1, 2$ and $i \neq j$. C represents the whole cell capacitance. I_{Na},I_K and I_L are a fast Na^+, delayed-rectifier K^+, passive leakage current, respectively. I_{syn-e} represents a excitatory synaptic input from other cell in the network. Specifically, $I_{NaP} = g_{NaP}m_{p,\infty}(v_i)h_i(v_i - E_{Na})$, $I_{Na} = g_{Na}m_\infty^3(v_i)(1 - n_i)(v_i - E_{Na})$, $I_K = g_K n_i^4(v_i - E_K)$, $I_L = g_L(v_i - E_L)$, $I_{tonic-e} = g_{tonic-e}(v_i - E_{syn-e})$, $I_{syn-e} = \sum_{i \neq j} g_{syn-e}s_j(v_i - E_{syn-e})$ and $I_{CAN} = g_{CAN}f([Ca]_i)(v_i - E_{Na})$.

The activation of the CAN current by the calcium concentration is given as $f([Ca]_i) = (1 + (K_{CAN}/[Ca]_i)^{n_{CAN}})^{-1}$. The calcium dynamics is described as $d[Ca]_i/dt = f_m(J_{ER_{IN}} - J_{ER_{OUT}})$, $dl_i/dt = AK_d(1 - l_i) - A[Ca]_i l_i$, in which $J_{ER_{IN}} = (L_{IP_3} + P_{IP_3}[\frac{IP_3[Ca]_i l_i}{(IP_3+K_l)([Ca]_i+K_a)}]^3)([Ca]_{ER} - [Ca]_i)$, $J_{ER_{OUT}} = V_{SERCA}\frac{[Ca]_i^2}{K_{SERCA}^2+[Ca]_i^2}$ and $[Ca]_{ER} = \frac{[Ca]_{Tot}-[Ca]_i}{\sigma}$. The meaning and values of other parameters are same as that in [10].

3 Bursting and Pattern Transition Mechanisms

The activity patterns depending on parameters g_{NaP} and g_{CAN} are illustrated, as shown in Fig. 1a, The horizontal axis represents g_{NaP} and the vertical g_{CAN}. Two-coupled cells in the pre-BötC can generate two types of oscillations: the in-phase and anti-phase oscillations [1], so the two-parameter space can be divided into four regions: region I (silence), region II (in-phase bursting and anti-phase bursting), region III (in-phase bursting and anti-phase spiking) and region IV (in-phase spiking and anti-phase spiking). We chose $g_{CAN} = 0.7$ nS as a representative and explore

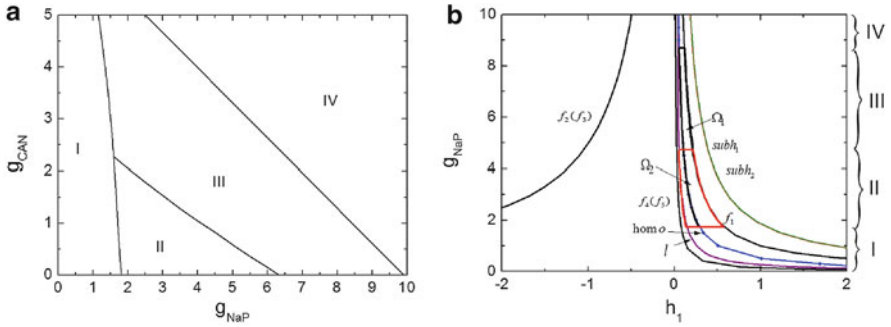


Fig. 1 (**a**) The firing patterns depending on two parameters g_{NaP} and g_{CAN}. The parameter space is divided into four parts, that is, region I, region II, region III and region IV. (**b**) Two-parameter bifurcation analysis of the fast subsystem $(1)_i$, $(3)_i$ and $(4)_i$ (i = 1, 2) with h_1 as a single slow variable as $g_{CAN} = 0.7$ nS. The *black curves* labeled with f_i ($i = 1, 2, 3, 4, 5$) denote the fold bifurcations and the *purple curve* labeled with l denotes saddle-node bifurcation on the limit cycle. The *solid green curve* ($subh_1$) and the *red curve* ($subh_2$) represent the subcritical Andronov-Hopf bifurcations respectively. The *blue line-dot curve* ($homo$) is constituted of the homoclinic bifurcation points HC of limit cycle. The region Ω_1 (in *solid black line*) and Ω_2 (in *solid red line*) represent the slow variable regions of in-phase bursting and anti-phase bursting, respectively

the bursting transition mechanisms between these regions with g_{NaP} changing. The systems yield a steady calcium concentration $[Ca]_1 = [Ca]_2 = 0.02104057$ mV with $IP_3 = 0.8\,\mu$M. h_1 and h_2 are almost equal when $g_{syn-e} = 9$ nS [1], we can consider h_1 and h_2 as one single slow variable, h_1. The two-parameter bifurcation analysis is shown in Fig. 1b. As g_{NaP} increases, the systems undergo regions I, II, III and IV. The maximum and minimum values of limit cycles in the full system (1)–(4), we named "the slow variable regions" Ω_1 and Ω_2 [8], are appended in Fig. 1b. The bifurcation curves in regions Ω_1 and Ω_2 are different which play an important role in determining which types of bursting or spiking can occur.

The fast/slow decompositions of fast systems $(1)_i$, $(3)_i$ and $(4)_i$ (i = 1,2) in the two-cell network model with $g_{CAN} = 0.7$ nS are shown in Fig. 2a–d. As the cross point of h-nullcline (the thick solid blue line) and "S-shaped" curves is under the bifurcation point F_1, the systems exhibit silence in region I, as shown in Fig. 2a. When $g_{NaP} = 4$ nS, the cross point exceeds the bifurcation point F_1 and the systems exhibit bursting or spiking. The fast/slow decomposition of fast systems with respect to the slow variable h_1 is shown in Fig. 2b. The systems exhibit in-phase bursting and anti-phase bursting in region II, as shown in Fig. 2e, f, respectively. The in-phase bursting is the "fold/homoclinic" type due to the active state disappears via homoclinic bifurcation; and the anti-phase bursting is the "fold/fold cycle" type due to the active state disappears via the fold limit cycle bifurcation (LPC) [7]. When g_{NaP} increases to 5 nS, the anti-phase bursting transit into anti-phase spiking but the in-phase bursting does not, which form the firing patterns in region III. When g_{NaP} increases to 9 nS, both in-phase and anti-phase bursting transit into spiking, which are shown in Fig. 2g, h respectively.

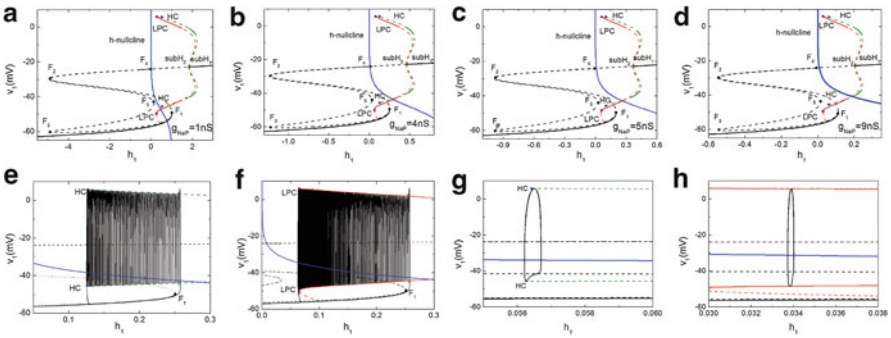


Fig. 2 The fast/slow decompositions of the two-cell network model with different values of g_{NaP}, here $g_{CAN} = 0.7$ nS. (**a**) $g_{NaP} = 1$ nS, (**b**) $g_{NaP} = 4$ nS, (**c**) $g_{NaP} = 5$ nS, (**d**) $g_{NaP} = 9$ nS. F_i ($i = 1, 2, 3, 4, 5$) represent the fold bifurcations, $subH_1$ and $subH_2$ two subcritical Andronov-Hopf bifurcations, LPC the saddle-node bifurcation on the limit cycle, HC homoclinic bifurcation of the limit cycle. The *blue curve* represents the h-nullcline. (**e**) "Fold/homoclinic" (in-phase) bursting with $g_{NaP} = 4$ nS. (**f**) "Fold/fold cycle" (anti-phase) bursting with $g_{NaP} = 4$ nS. (**g**) In-phase spiking with $g_{NaP} = 9$ nS. (**h**) Antiphase spiking with $g_{NaP} = 9$ nS

4 Conclusions

In this paper, the bursting and pattern transition mechanisms with g_{NaP} varying are concerned in a two-cell network model of the pre-BötC. The parameter space (g_{NaP}, g_{CAN}) is divide into four regions according to the multi-phasic oscillations. The "fold/homoclinic" type and "fold/fold cycle" type bursting emerge in two-coupled cells and the transition mechanisms of these bursting with g_{NaP} changing are investigated. The results show that the transitions of these two types of bursting are not synchronized with g_{NaP} increasing. The asynchronism of firing patterns in the two-cell network of pre-BötC indicaes that the selection of network, which is important for the generation of respiratory rhythm. The study can also provide a further understanding of how NaP current working in respiratory network.

References

1. J. Best, A. Borisyuk, J. Rubin, D. Terman, M. Wechselberger. *SIAM J. Appl Dyn Syst.* **4**, 1107–1139 (2005).
2. R.J. Butera, J. Rinzel, J.C. Smith. *J. Neurophysiol.* **82**, 382–397 (1999a).
3. R.J. Butera, J. Rinzel, J.C. Smith. *J. Neurophysiol.* **82**, 398–415 (1999b).
4. C.A. Del Negro, C. Morgado-Valle, J.L. Feldman. *Neuron.* **34(5)**, 821–830 (2002b).
5. L.X. Duan, D.H. Zhai, X.H. Tang. *Int. J. Bifurcat Chaos.* **22(05)** (2012).
6. G.B. Ermentrout. 1nd edition, (Philadelphia, SIAM) (2002).
7. E.M. Izhikevich. *Int. J. Bifurcat Chaos.* **10**, 1171–1266 (2000).

8. Q.S. Lu, Z.Q. Yang, L.X. Duan, H.G. Gu, R. Wei. *Chaos. Solitons. fract.* **40** 577 (2009).
9. R.W. Pace, D.D. Mackay, J.L. Feldman, C.A. Del Negro. *J. Physiol.* **582 (1)**, 113–125 (2007a).
10. C. Park, J.E. Rubin. *J. Comput. Neurosci.*, 1–22 (2013).
11. J.C. Smith, H.H. Ellenberger, K. Ballanyi, D.W. Richter, J.L. Feldman. *Science.* **254**, 726–729 (1991).

An Electrodiffusive Formalism for Ion Concentration Dynamics in Excitable Cells and the Extracellular Space Surrounding Them

Geir Halnes, Ivar Østby, Klas H. Pettersen, Stig W. Omholt, and Gaute T. Einevoll

Abstract In processes where ionic concentrations vary significantly, the standard cable equation fails to accurately predict the transmembrane potential. Such processes call for a mathematical description able to account for the spatiotemporal variations in ion concentrations as well as the subsequent effects of these variations on the membrane potential. We here derive a general electrodiffusive formalism for consistently modeling the dynamics of ion concentration and the transmembrane potential in a one-dimensional geometry, including both the intra- and extracellular domains. Unlike standard cable theory, the electrodiffusive formalism accounts for diffusive currents and concentration-dependent variation of the longitudinal resistivities.

Keywords Ion concentrations • Spatiotemporal variations • Membrane potential • Electrodiffusive formalism • Two-domain model

1 Introduction

In standard cable theory, the effect of ionic diffusion on the net electrical currents is neglected. Longitudinal resistivities, which in reality depend on ion concentrations, are assumed to be constant [6,9,12]. These are typically good approximations when modelling short-term electrical neural activity, as ion concentration typically vary little at the relevant time scale (<100 ms). However, in small intracellular volumes, such as dendritic spines, the local ion concentration can change quite dramatically

G. Halnes (✉) • I. Østby • K.H. Pettersen • G.T. Einevoll
Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, Ås, Norway
e-mail: geir.halnes@umb.no; ivar.ostby@medisin.uio.no; klas.pettersen@umb.no; gaute.einevoll@umb.no

S.W. Omholt
Department of Animal and Aqucultural Sciences, Centre for Integrative Genetics, Norwegian University of Life Sciences, Ås, Norway
e-mail: stig.omholt@umb.no

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_50

within a few milliseconds [11], and in slower, macroscopic transport processes, concentration gradients may build up over time [2, 7]. For processes involving significant ion concentration gradients, the cable model will fail to give accurate predictions.

Qian and Sejnowski [11] have previously developed a consistent, electrodiffusive scheme for modelling the dynamics on v_M and ion concentrations [11]. Like the standard cable model, the electrodiffusive model assumes that transport phenomena are essentially one-dimensional. Unlike the standard cable model, the electrodiffusive model derived v_M from the ion concentration dynamics, accounting for all ionic movements (membrane fluxes, longitudinal diffusion, and longitudinal electrical migration), as well as for the concentration-dependent variation of the intracellular resistivities. An important limitation with this model [11] is that it only includes the dynamics in the intracellular space (ICS), whereas the extracellular space (ECS) was assumed to be isopotential and with constant ion concentrations.

Here expand the electrodiffusive formalim [11] to explicitly include both the ICS and the ECS. The result is a general mathematical framework for consistently modelling the dynamics of the membrane potential and ion concentrations in the intra- and extracellular domain. We believe that this framework will be of general value for the field of neuroscience. In the discussion, we give a few examples of processes that the formalism can be applied to.

2 Electrodiffusive Formalism

We seek a general mathematical framework for consistently modelling the dynamics of the membrane potential (v_M) and the concentrations in the ICS ($[k]_I$) and ECS ($[k]_E$) of a set (k) of ionic species in a geometry as that depicted in Fig. 1.

2.1 Particle Conservation

We consider the continuity equations for an ion species k with valence z_k in domains I and E:

$$\frac{\partial j_{kI}(x,t)}{\partial x} + \frac{O_M}{a_I} j_{kM}(x,t) + \frac{O_M}{a_I} j_{kI}^{in}(x,t) + \frac{\partial [k]_I(x,t)}{\partial t} = 0 \tag{1}$$

$$\frac{\partial j_{kE}(x,t)}{\partial x} - \frac{O_M}{a_E} j_{kM}(x,t) + \frac{O_M}{a_E} j_{kE}^{in}(x,t) + \frac{\partial [k]_E(x,t)}{\partial t} = 0, \tag{2}$$

with the sealed-end boundary conditions ($n = I$ or E):

$$j_{kn}(0,t) = j_{kn}(l,t) = 0. \tag{3}$$

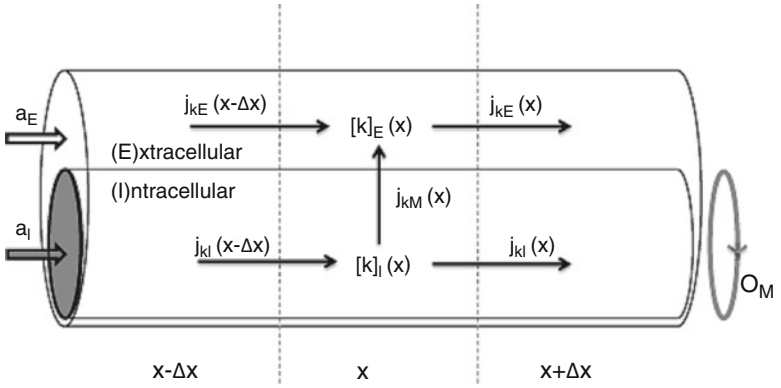


Fig. 1 A two domain-model for ion concentration dynamics in the intra- and extracellular space. The ICS is represented as a cylindrical cable (I), coated by ECS (E). The geometry is specified by three parameters, where a_I is the cross section area of the cable, a_E is the cross section area of the ECS and O_M is the circumference of the cable. The concentration of ion species k is denoted $[k]_n$ where n represents domain I or E. Ionic movement is described by the transmembrane flux density (j_{kM}) and the longitudinal flux densities due to electrical migration (j^f_{kn}) and diffusion (j^d_{kn})

Here a_I and a_E are the cross sections of the ICS and ECS, respectively, and O_M is the circumference of the membrane. The longitudinal flux densities are given by the generalized Nernst-Planck equation (to keep notation short, we skip the functional arguments (x, t) from here on):

$$j_{kn} = -\frac{D_k}{\lambda_n^2}\frac{\partial [k]_n}{\partial x} - \frac{D_k z_k}{\lambda_n^2 \psi}[k]_n \frac{\partial v_n}{\partial x}, \tag{4}$$

where the first term on the right represents the diffusive flux density (j^d_{kn}) and the last term is the flux density due to ionic migration in the electrical field (j^f_{kn}). The effective diffusion constant $D^*_k = D_k/\lambda_n^2$ is composed of the diffusion constant D_k in dilute solutions and the tortuosity factor λ_n, which summarizes the hindrance imposed by the cellular structures [2, 10]. We use $\psi = RT/F$, where R is the gas constant, T the absolute temperature, and F is Faraday's constant. The formalism we derive is general to the transmembrane flux density (j_{kM}), as long as j_{kM} is a local function of v_M, ionic concentrations in I and E, and possibly some additional local state variables. The formalism can be combined with any external input (j^{in}_{kn}) which fulfills the constraint:

$$\sum_k z_k j^{in}_{kE} = -\sum_k z_k j^{in}_{kI}, \tag{5}$$

as we shall explain later.

With N ion species, Eqs. 1 and 2 (with j_{kn}, j_{kM} and j_{kn}^{in} as described above) represent a system of $2N+3$ variables which are functions of x and t. These are the $2N$ concentration variables ($[k]_n$ for $k = 1, 2, \ldots N$ and $n = E, I$), and the three additional variables (v_M, $\partial v_I/\partial x$ and $\partial v_E/\partial x$) occurring in the expressions for the flux densities. We now seek to express v_M, $\partial v_I/\partial x$ and $\partial v_E/\partial x$ as functions of ionic concentrations, so that Eqs. 1 and 2 constitute a fully specified (and numerically solvable) system of equations.

2.2 Voltage Expressions

To reduce the number of independent variables to the $2N$ state variables ($[k]_n$) we use three additional constraints:

$$v_M = \frac{a_I}{C_M O_M}\rho_I \tag{6}$$

$$\frac{a_I}{C_M O_M}\rho_I = -\frac{a_E}{C_M O_M}\rho_E. \tag{7}$$

$$v_M = v_I - v_E \Rightarrow \frac{\partial v_M}{\partial x} = \frac{\partial v_I}{\partial x} - \frac{\partial v_E}{\partial x} \tag{8}$$

Equation 6 is the assumption that the membrane is a parallel plate capacitor. Then v_M is determined by the density of charge on the inside of the membrane, which in turn is determined by the ionic concentrations:

$$\rho_n = F\sum_k z_k[k]_n + \rho_{sn}. \tag{9}$$

For practical purposes, we have included a density of static charges (ρ_{sn}) in Eqs. 6 and 7, representing contributions from ions that are not considered in the conservation equations. If the set $[k]_n$ include all present species of ions, then $\rho_{sn} = 0$.

As a capacitor separates a charge Q from a charge $-Q$ (equal in magnitude, opposite in sign), the ECS charge density must be also consistent with v_M. This is the second constraint (Eq. 7). To our knowledge, we are the first to make use of it in an electrodiffusive model. When using Eqs. 6 and 7, we implicitly assume that all local net charge in the system is stuck on the capacitive membrane (in the Debye-layers), and that the bulk solutions in the ICS and ECS are electroneutral [4, 8]. In order not to violate Eq. 7, the external input to the system must also be locally electroneutral (cf. Eq. 5).

We shall now use Eq. 7 together with the general definition of the transmembrane potential (Eq. 8) to derive our expressions for $\partial v_I/\partial x$ and $\partial v_E/\partial x$. We start by summing the particle conservation laws (Eqs. 1 and 2) to obtain:

$$\frac{\partial j_{kI}}{\partial x} + \frac{\partial j_{kE}}{\partial x} + \frac{O_M}{a_I} j_{kI}^{in} + \frac{O_M}{a_E} j_{kE}^{in} + \frac{\partial [k]_I}{\partial t} + \frac{\partial [k]_E}{\partial t} = 0 \tag{10}$$

If we multiply this by Fz_k and take the sum over all ion species, k, the terms involving j_{kn}^{in} disappear due to Eq. 5 and the terms involving $[k]_n$ disappear due to Eqs. 7 and 9. We are left with:

$$a_I F \sum_k z_k \frac{\partial j_{kI}}{\partial x} = -a_E F \sum_k z_k \frac{\partial j_{kE}}{\partial x} \rightarrow a_I F \sum_k z_k j_{kI} = -a_E F \sum_k z_k j_{kE}. \tag{11}$$

The last implication follows from the sealed-end condition (Eq. 3). If the charge symmetry condition is satisfied at a given time $t = 0$, Eq. 11 is the condition that it remains satisfied at all times t.

The flux densities j_{kn} are defined by Eq. 4. We note that Eq. 11 contains the sum of j_{kn} over all ionic species. These sums can be converted to current densities. For convenience, we distinguish between the current densities due to diffusion (i_n^d) and migration in the electrical field (i_n^f), defined as [6]:

$$i_n^d = -\sum_k \frac{FD_k z_k}{\lambda_n^2 \psi} \frac{\partial [k]_n}{\partial x}, \tag{12}$$

and

$$i_n^f = -\sum_k \frac{FD_k z_k^2}{\lambda_n^2 \psi} [k]_n \frac{\partial v_n}{\partial x} = -\frac{1}{r_n} \frac{\partial v_n}{\partial x}, \tag{13}$$

where r_n denotes the resistivity [6]. With Eqs. 4, 12 and 13, Eq. 11 can be rewritten as:

$$a_I \left(i_I^d + \frac{1}{r_I} \frac{\partial v_I}{\partial x} \right) = -a_E \left(i_E^d + \frac{1}{r_E} \frac{\partial v_E}{\partial x} \right), \tag{14}$$

By combining Eq. 8 with Eq. 14 we may finally derive our expressions for the voltage gradients:

$$\frac{\partial v_I}{\partial x} = \left(\frac{\partial v_M}{\partial x} + \frac{r_E a_I}{a_E} i_I^d + r_E i_E^d \right) \left(1 + \frac{r_E a_I}{r_I a_E} \right)^{-1} \tag{15}$$

$$\frac{\partial v_E}{\partial x} = \left(-\frac{\partial v_M}{\partial x} + r_I i_I^d + \frac{r_I a_E}{a_I} i_E^d \right) \left(1 + \frac{r_I a_E}{r_E a_I} \right)^{-1}. \tag{16}$$

Here, r_n is given by Eq. 13, i_n^d by Eq. 12, and v_M by Eqs. 6 and 9. All voltage terms are thereby expressed in terms of ionic concentrations, and the conservation equations (Eqs. 1 and 2) represent a fully specified system.

3 Discussion

We presented a one-dimensional, electrodiffusive framework for modeling the dynamics of the membrane potential (v_M) and the ion concentrations ($[k]_n$) of all included ion species (k) in an intra- and extracellular domain.

3.1 Implementation

A step-wise procedure of how to use this formalism is a follows:

1. Specify initial conditions for the membrane potential (v_{M0}) and k ion concentrations ($[k]_{n0}$) that is to be simulated. The main charge carriers are believed to be Na^+, K^+ and Cl^-, but other species can be included.
2. Specify the static charge density (ρ_{sn}) so that $[k]_{n0}$ and v_{M0} are consistent according to Eqs. 6 and 7.
3. Specify the membrane mechanisms (functions for j_{kM}) representing ion channels, ion pumps and other membrane mechanisms relevant for the cell type that is to be modelled.
4. Specify an external input function that fulfills Eq. 5.
5. Solve the $2N$ conservation equations (Eqs. 1 and 2), with the boundary condition in Eq. 3. This can be done numerically by using a spatial discretization method (e.g., the Matlab-solver *pdepe*). For each time step, v_M, $\partial V_I/\partial x$ and $\partial V_E/\partial x$ are defined algebraically by Eqs. 8, 15 and 16.

3.2 Applications

In the most direct interpretation, the ICS in Fig. 1 represents a single neurite coated with a thin sheath of ECS. For example, the ICS could represent an axon that has been removed from the ionic solution and placed in air or oil, so that only a thin layer of the ionic solution surrounds its membrane [12]. Alternatively, the ICS could represent an individual axon in e.g., the optical nerve, where axons are densely packed in bundles, and separated by constrained gaps of extracellular space [1]. However, cases where it is biologically meaningful to consider the ECS as a relatively thin coating faithfully following a single cell may be limited. For most single cell processes, the assumption that $a_I \ll a_E$, and that conditions in

the ECS are constant, may be more reasonable. In this limit the electrodiffusive formalism reduces to the one-domain (ICS) model presented previously by Qian and Sejnowski [11].

A geometrical simplification as that in Fig. 1 has also been justified for certain macroscopic transport processes through a chunk of neural tissue [2,3,5]. Typically, the ECS comprises about 20 % of the total neural tissue volume, while the remaining 80 % is the ICS of various cells [2]. Assuming that a large number of cells (e.g., all cells belonging to a specific species) participate in simultaneous ion exchange with the ECS, the impact on the ion concentrations in the ICS and ECS may be of the same order of magnitude. This calls for a two-domain formalism such as ours. When addressing a macroscopic transport process, the ICS in Fig. 1 does not refer to a single cell, but the total amount, within the chunk of tissue, of the participatory cell type. Similarly, ECS refers to the total extracellular volume in the chunk. For the geometrical parameters a_E, a_I and O_M, one could then use, respectively, the ECS volume per total tissue volume, the ICS volume per total tissue volume, and the membrane surface area per total tissue volume. Such a macroscopic interpretation of Fig. 1 allows for a broader range of applications.

We have previously presented the electrodiffusive formalism with the specific application to a model of spatial potassium buffering by astrocytes [5]. We refer to the previous work for a specific, illustrative implementation.

Acknowledgements The project was supported by the Research Council of Norway (eVITA program; project numbers 178892 and 178901), and EU Grant 269921 (BrainScaleS).

References

1. Bokil, H., Laaris, N., Blinder, K., Ennis, M., Keller, A.: Ephaptic Interactions in the Mammalian Olfactory System. Journal of Neuroscience **21**: RC173 (2001).
2. Chen KC., Nicholson C.: Spatial buffering of potassium ions in brain extracellular space. Biophysical journal **78**: 2776–97 (2000)
3. Gardner-Medwin A.: Analysis of potassium dynamics in mammalian brain tissue. The Journal of physiology **335**: 393–426 (1983)
4. Feldberg, S.: On the dilemma of the use of the electroneutrality constraint in electrochemical calculations. Electrochemistry Communications **2**: 453–456 (2000).
5. Halnes,G., Østby, I., Pettersen, K.H., Omholt, S.W., Einevoll, G.T.: Electrodiffusive model for astrocytic and neuronal ion concentration dynamics. PLoS Comp. Biol. 9(12): e1003386 (2013).
6. Koch, C.: Biophysics of computation: information processing in single neurons. Oxford University Press: New York, 1st edition (1999).
7. Kofuji P., Newman EA.: Potassium buffering in the central nervous system. Neuroscience **129**: 1045–56: (2004)
8. Leonetti, M., Dubois-Violette, E. Theory of Electrodynamic Instabilities in Biological Cells. Physical Review Letters **81**: 1977–1980 (1998).
9. Mori, Y.: From Three-Dimensional Electrophysiology to the Cable Model: an Asymptotic Study. arXiv:09013914 [q-bioNC] (2009).

10. Nicholson, C., Syková, E. Extracellular space structure revealed by diffusion analysis. Trends in neurosciences **21**: 207–15 (1998).
11. Qian, N., Sejnowski, T.: An electro-diffusion model for computing membrane potentials and ionic concentrations in branching dendrites, spines and axons. Biological Cybernetics **15**: 1–15 (1989).
12. Rall, W.: Core conductor theory and cable properties of neurons: In: Kandel, E.R. (editor), Handbook of Physiology: The nervous system. American Physiological Society (1977)

Dispersion in Excitatory Synaptic Response

Francesco Ventriglia

Abstract Due to its extreme importance for brain activity, the function of the excitatory synapse is the object of a huge amount of researches investigating the specific contribution to the synaptic response of fairly all its structural elements. Here, we utilized a model of hippocampal synapse to describe the random dispersion of the response evaluated from the dispersion of the amplitude peak of the miniature Excitatory Post-Synaptic Current (mEPSC). The model is based on time discretized Langevin Equations which describe the Brownian motion of Glutamate molecules released by a neurotransmitter vesicle within the synaptic cleft, their collisions with the structural elements, their binding to post-synaptic receptors and their final spillover. The value of the amplitude peak of the computed mEPSCs was put in relationship with different binding probabilities and different number of AMPA receptors. The dispersion has been used to compute an appropriate value of the binding probability of Glutamate molecules to post-synaptic receptors.

Keywords Glutamate synapse response • Binding probability • EPSC peak value • Synapse computer simulation

1 Introduction

Grasping what determines the time-course of the excitatory synapse response, its variations and its random structure is of basic importance for the understanding of the brain activity. Experimental data and modeling/computational investigations made this field more and more clear, but the synaptic function still remains difficult to understand [2, 3, 5]. In a series of articles (see references in [6]), we studied a mathematical model of an (hippocampal) excitatory synapse. It is based on the description of Brownian motion of Glutamate molecules (GLUTs) within the synaptic cleft by a discrete-time Langevin equations, ruling their space position and velocity, and of their interaction with the structural elements of the synapse. The

F. Ventriglia (✉)
Istituto di Cibernetica "E.Caianiello", CNR, Via Campi Flegrei 34, 80078 Pozzuoli, NA, Italy
e-mail: franco@ulisse.cib.na.cnr.it

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_51

model is simulated on a parallel computer by using an ultra-fast time scale. In fact, an extremely short time step was chosen, $\Delta t = 40 \cdot 10^{-15}$ s. This value permitted an extremely accurate description of the collisions of GLUTs with synaptic elements. Here, we will present, discuss and utilize the *strange* phenomenon of the peak amplitude dispersion of the synaptic response, that has been brought to light by the late simulations. This phenomenon occurs when, in a series of simulations in which all the parameters of the synaptic model assume the same values and only the seed for the initialization of the Random Number Generator (RNG) changes, the miniature Excitatory Post-Synaptic Currents (mEPSCs) present amplitude peaks having different values. We will show that the dispersion of the synaptic response is reduced when the number of AMPARs is higher.

2 Model of Excitatory Synapse

The geometry of the synaptic cleft model was based on two (concentric) cylinders with a common height of 20 nm. The entire synaptic cleft was represented by the larger one, whereas the active synaptic space was simulated by the smaller cylinder, having bases on the Active Zone (AZ) and on the Post-Synaptic Density (PSD). Attacked on the top of AZ, a small sphere simulated a releasing neurotransmitter vesicle. AMPA and NMDA receptors were modeled as small cylinders protruding in the synaptic cleft from the PSD zone. Two small circles, having the diameter of the cross-section of a GLUT and located randomly on the exposed face of the receptor, simulated the binding sites for GLUTs. The annular space, external at the AZ/PSD synaptic volume till to the boundary of the cleft, was filled with Filaments [7]. We assumed that at the arrival of the Action Potential (AP), an expanding fusion pore opened between the vesicle and the synaptic cleft. It was simulated as a cylinder with a gradually increasing diameter and a fixed height of 12 nm. At the starting computation time, $t = 0$, GLUTs contained in the vesicle were distributed in space according an uniform distribution, and in velocity according to a Maxwell distribution, and the diameter of the fusion pore was equal to the diameter of the cross section of a GLUT.

The Brownian motion was described by the Langevin equations:

$$\mathbf{r}_i\,(t + \Delta) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\Delta \tag{1}$$

$$\mathbf{v}_i\,(t + \Delta) = \mathbf{v}_i(t) - \gamma \frac{\mathbf{v}_i(t)}{m}\Delta + \frac{\sqrt{2\gamma\varepsilon\Delta}}{m}\Omega_i \tag{2}$$

where i is the ith molecule ($i = 1 \ \ldots \ N$; N being the total number of Glutamate molecules), Δ (Δt) is the time step and Ω_i is a random vector with three components, each having a Gaussian distribution with mean value $\mu = 0$ and standard deviation $\sigma = 1$. The other parameters are:

m, the molecular mass; γ, a friction term, which depends on the absolute temperature: $\gamma = k_B \cdot T/D$, where k_B is the Boltzmann constant, T is the absolute temperature in Kelvin degrees, and D is the diffusion coefficient of Glutamate; $\varepsilon = k_B \cdot T$.

Concerning the GLUT/AMPAR binding we note that the results in experimental literature were obtained by imposing unrealistic conditions: stationarity of the Glutamate concentration within the synaptic cleft and a long period, of the order of milliseconds, for the exposition of receptors to Glutamate [2]. Conversely, when the neurotransmitters are released by a docked vesicle, the flow of neurotransmitters within the cleft is much faster and far from stationarity. In the lack of correct experimental data for the binding rates, we attempted a computation of the binding by using probabilistic arguments. We searched the unknown value of the binding probability on the base of a geometrical reasoning and through comparisons of the peak amplitudes between the computer simulated and the experimentally recorded mEPSCs. The main hypothesis is based on the experimental observation that the binding takes place when the GLUT presents a particular arrangement with respect to the binding site: the γ-carboxyl group of GLUT is in front [1]. To compute the binding probability we assumed an elongated shape for GLUT and we made the hypothesis that GLUT binds to the site only when the directions of the long axis belong to a restricted range of values. By considering a GLUT as a spindle set at center of a unit sphere, we hypothesized that only the directions contained within a (unknown) spherical cone, could produce the binding. Thus, the binding probability was computed as the ratio between the volume of the spherical cone and that of the unit sphere and, by varying the angle of the cone, we observed the effect of different binding probabilities on the synaptic response. In all our simulations only the AMPARs contributions have been considered for the computation of the mEPSC, since under normal conditions the ionic channels of NMDARs cannot convey currents. Thus, NMDARs were considered only as competitors of AMPARs for GLUTs binding.

The discretized Langevin equations were implemented in a parallel FORTRAN program by using MPI (message passing interface) routines. The paths of all the GLUTs were computed up to the occurrence of one of the following events: a re-uptake, a receptor binding or the spillover. The program was run on a computer based on a cluster of workstations. Single and double binding times between GLUTs and AMPARs/NMDARs were recorded in two matrices and used for the subsequent computation of the mEPSC by a program which produced the random transitions to/from open and close states and to the single binding states.

3 Simulation and Results

The new computer simulations investigated more thoroughly the phenomenon of the dispersion of the synaptic response. We analyzed the results obtained with three values for the binding probability, B_P: 0.00496, 0.00874, 0.01360 (related to

Table 1 The first column, Case- with Low and High, is related to the number of AMPARs

Case	Mean value	Standard deviation	CV
Low – 18°	−10.26	2.01	19.58
High – 18°	−18.09	3.68	20.35
Low – 24°	−17.92	1.79	10.01
High – 24°	−34.85	2.31	6.62
Low – 30°	−23.56	1.42	6.03
High – 30°	−46.47	1.07	2.31

Second column -Mean value- is related to the amplitude peaks of the computed mEPSCs. Standard deviation and CV constitute the next two columns

Spherical Segment Angle: 18°, 24°, 30°, respectively), and, for each probability value, two values for the number of AMPARs, one normal and the other very high: (A) 55 and (B) 154. (The respective values for NMDARs were 13 and 18.) For each couple (B_P, AMPARs Number), six simulations with different GNR seeds were accomplished, for a total of 36 computer simulations. In all these simulations the vesicle was located at $X_0 = 0$ and released 775 GLUTs. The total height of AMPARs (and NMDARs) was 17 nm, while the height of the portion of receptor protruding in the synaptic cleft was 6 nm. The relevant statistical values of the amplitude peaks of the computed mEPSCs are reported in Table 1.

4 Discussion

The phenomenon of the peak amplitude dispersion of the synaptic response shown by some of our previous computer simulations has been analyzed in more detail in this article to ascertain its origins. The main aim was to obtain information about the binding probability value (of GLUTs to AMPA receptors) that better approximate the biological reality. Several series of computer simulations have been carried out. In each series all the parameters of the synaptic model remain fixed and only the seed for the initialization of the RNG varied. From the results in Table 1 we could note that the dispersion is more marked in series in which the model parameters lead to amplitude peaks with a lower mean (absolute) value. The selection of a low value for the binding probability seems to be the main cause of this phenomenon. In some respects, also the presence of a lower number of AMPA receptors in the Post Synaptic Density contributes to extend the dispersion. Hence, the large variation of the mEPSC peak amplitude shown in these particular conditions is a symptom of the instability of the synaptic response due to too low binding probability values. We concluded that the best value for an estimate of the binding probability is that related to the spherical cone angle of 30°, i.e. $B_P = 0.01360$, since it produced the lower CV (dispersion: Standard deviation/Mean).

The relationship between the mEPSC peak amplitude and the synaptic dispersion presents other more interesting aspects, inviting to theoretical speculations. In fact,

when the binding probability value is high and the number of AMPARs is larger, then the dispersion is reduced until to disappear. This fact induces to think to the effects produced on the mEPSCs by the increasing of the AMPARs number that occurs in consequence of the Long-Term Potentiation process [4]. If the reduction of the dispersion in consequence of the increasing of the number of the AMPA receptors is a true, physiological aspect of the synaptic response, then learning and memory, which seem to be based on the LTP phenomenon, should be considered in a different prospective.

References

1. Armstrong N, Gouaux E (2000) Mechanisms for Activation and Antagonism of an AMPA-Sensitive Glutamate Receptor: Crystal Structures of the GluR2 Ligand Binding Core. Neuron 28: 165–181
2. Clements JD, Lester RA, Tong J, Jahr CE, Westbrook GL (1992) The time course of Glutamate in the synaptic cleft. Science 258: 11498–11501
3. Liu G (2003) Presynaptic control of quantal size: kinetic mechanisms and implications for synaptic transmission and plasticity. Curr Opin Neurobiol 13: 324 –331
4. Malinow R, Malenka RC (2002) AMPA receptor trafficking and synaptic plasticity. Annu Rev Neurosci. 25: 103–126
5. Uteshev VV, Pennefather PS (1996) A mathematical description of miniature post-synaptic current generation at central nervous system synapses. Biophys J 71: 1256–1266
6. Ventriglia, F., DiMaio, V. (2013) Glutamate-AMPAR interaction in a model of synaptic transmission. Brain Research (2013) http://dx.doi.org/10.1016/j.brainres.2013.04.051
7. Zuber B, Nikonenko I, Klauser P, Muller D, Dubochet J (2005) The mammalian central nervous synaptic cleft contains a high density of periodically organized complexes. Proc Natl Acad Sci U.S.A. 102: 19192–19197

Research on the Neural Energy Coding

Rubin Wang and Ziyin Wang

Abstract In this paper, we investigated the energy distribution caused by neural activity of the biological neural network and the neural energy coding expressed by the network energy flow. The numerical results show that the energy coding can reflect the continuous change of cerebral blood flow in the cerebral cortex, and the positive and negative energy flow reflect the energy supply in cerebral blood flow and the consumption of energy in the neuronal population. Since the neuronal network is based on electrophysiological experiment, nervous energy coding proposed is a novel neuronal coding in neural information processing, which coding pattern cannot be reproduced by the existing neural model and neural coding theory. There is a unique corresponding relation between synchronous oscillation and energy flow in the neural network. Therefore, neuronal energy coding proposed can provide the basis for the establishment and analysis of global neural coding in the brain.

Keywords Neural energy • Biological neural networks • Cerebral blood flow • Neural information processing • Synchronous oscillations

1 Instruction

Neural coding and decoding is the most challenging and important field in neuroscience [1, 2]. However, there is no effective theory of neural coding and decoding to investigate the global behavior of brain activity, many problems in Neural coding and decoding are difficult to resolve [3, 4, 5, 6, 15]. The neural activity and operations of the brain subordinated to the principle of depleting minimized energy

R. Wang (✉)
Institute for Cognitive Neurodynamics, East China University of Science and Technology, Shanghai, People's Republic of China
e-mail: rbwang@163.com

Z. Wang
School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, People's Republic of China
e-mail: wangziyinx@yeah.net

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_52

and maximizing the efficiency of signal transmission [6]. Therefore, in neural information processing, how nervous energy is transformed to control the operation of the brain and information coding ? This is our motivation to explore nervous energy.

However, up to now, the research on energy consumption in the nervous system is mainly limited to the experimental data [7–12]. Hence to explore regular pattern hided in the phenomenon, the quantitative relationship between the nerve signal transmission and energy consumption is important.

In this paper, based on the above theoretical model and results, we investigate the energy flow in the network composed of the neuronal populations subjected to various stimulations, and the energy coding when the energy changes with respect to time.

2 Neuronal Networks with Mutual Coupling

In order to simulate the energy neuron network coding, we adopt a novel biophysical model of neuron proposed and researched in [13]. Thus a fully connected biological neural network can be researched, shown in Fig. 1, where a neuron couple with all other neurons. Therefore, The following neural network structure is strictly defined in the basis of neurobiology [13, 14].

The global connected neural network is composed of 15 neurons in Fig. 1. To understand the energy coding mode in the neural network subject to stimulation, the global coupled neural network is only composed of 15 neurons in Fig. 3. The neurons are coupled with bidirectional asymmetry coupling strengths. For example, the 1*th* neuron is coupled to the 2*th* neuron with coupling strength 0.15, and the 2*th*

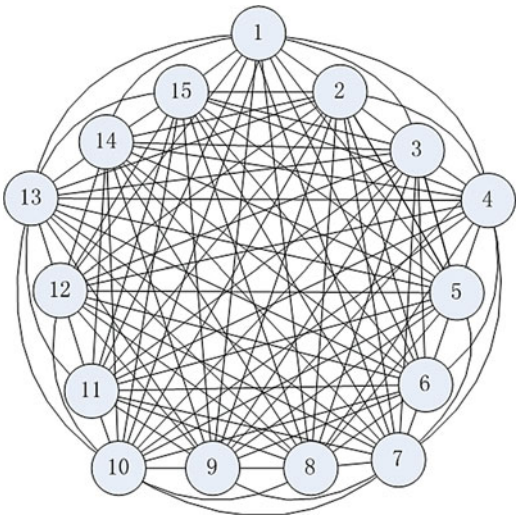


Fig. 1 Connection structure of neural networks

neuron is coupled to $1th$ neuron with coupling strength 0.22. A synaptic coupling strength between neurons is uniformly distributed in the statistical sense [13], we assume that a synaptic coupling strength is uniformly distributed in [0.1 0.3].

Let coupling strength matrix: $w_{i,j}$s coupling strength when the *ith* is coupled to *jth*.

The network operate as follows:

$$S_j(t) = W \times Q(t-\tau)^T \tag{1}$$

$$I_{m,j}(t) = \begin{cases} i_{m1} + \sum_{j=1}^{n} \left[i_{0m} (j-1) \sin\left(\omega_m (j-1) \left(t_j - t_{j-1}\right)\right)\right] \\ \quad + i_{0m}(n) \sin\left(\omega_m(n)(t-t_n)\right) \quad if \quad S_j(t) > th \\ i_{m1} + i_{0m}(n) \sin\left(\omega_m(n)(t-t_n)\right) \quad if \quad S_j(t) < th \end{cases} \tag{2}$$

we treat $I_{m,j}(t)$ as the stimulation of the neuron, then we can obtain action potential and power consumed during action potential from the model proposed by [18]

where $S_j(t)$ is the sum of the stimulations for the *jth* neuron at time t; $Q(t-\tau) = [Q_1(t-\tau), Q_2(t-\tau), \ldots, Q_j(t-\tau), \ldots Q_n(t-\tau)]$ indicates firing states of the neurons, which take value 0 at resting and 1 at firing.

3 Energy Coding Under Different Stimulus

3.1 Instantaneous Stimulus

We assume that Topological relationships between neurons are identical for a network fully connected, stimulation the neurons receiving stimulation may be an arbitrary position in the network. We assume that the $1th$, $2th$, $3th$ neurons stimulated at the same time, set it the first neuronal group. Stimulating time is 0.1 ms, stimulating voltage is 40 mV. Then, the stimulus is immediately disconnected. The coupling strength between neurons is uniformly distributed in [0.1 0.3], the transform delay is 2 ms, refractory period is set to 0.5 ms.

With the models and parameters mentioned above, the numerical simulation result can be achieve. Figure 2 shows the power consumed by the overall neural networked.

Figure 2 display that the transient response of energy network is 6 ms when the first neuronal group subject to the initial stimulus, and then the total energy of the neural network periodically change. This reflect the activity patterns of the neuronal population. In other words, the potential variation of the neural network can be estimated by using the total energy consumed.

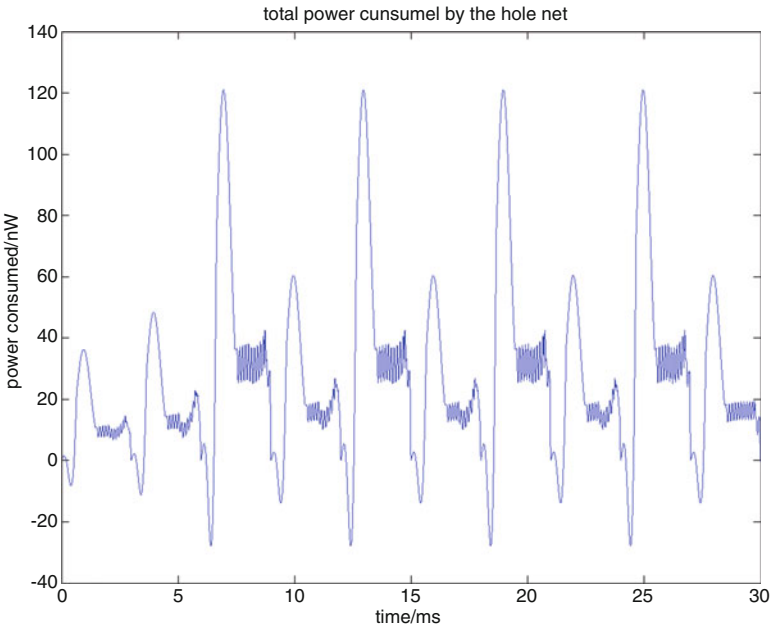


Fig. 2 The total power during the network operation

3.2 Intermittent Stimulation

Instantaneous stimulation is actually a special intermittent stimulation when intermittent time is infinite length. Therefore, intermittent stimulation and transient stimulation have the same regular pattern in some cases. Figure 3a display that the total energy of the neural network subject to intermittent stimulation. The neuronal network first operate at time 0, the 1th and 2th neurons are intermittently stimulated every 30 s, the stimulation time is 0.1 ms, the stimulation intensity is 40 mV, the intermittent time is 30 ms.

The flow of energy bring into a stable periodic stage after 10 ms (Fig. 3). An instantaneous stimulus is applied to the 1th and 2th neurons in the network at the time of 30 ms again, where the stimulation intensity and time are the same as the first time. The energy flow changes pattern from a stable phase to another stable periodic phase after the instantaneous phase.

According to Hebbian plasticity principle [16, 17], neurons will reduce sensitivity to stimuli when the cortical neural network is repetitively subject to the same stimulus, which result in the reduced demand of the cerebral blood flow. Compared to the energy encoding 30 ms later, the energy encoding contain learning information under the initial instantaneous stimulus, and the energy coding is used to maintain and store the learning information after 30 ms. In other words, the information of the initial transient stimulation is preserved by intermittent stimulation, so it is a more stable learning process.

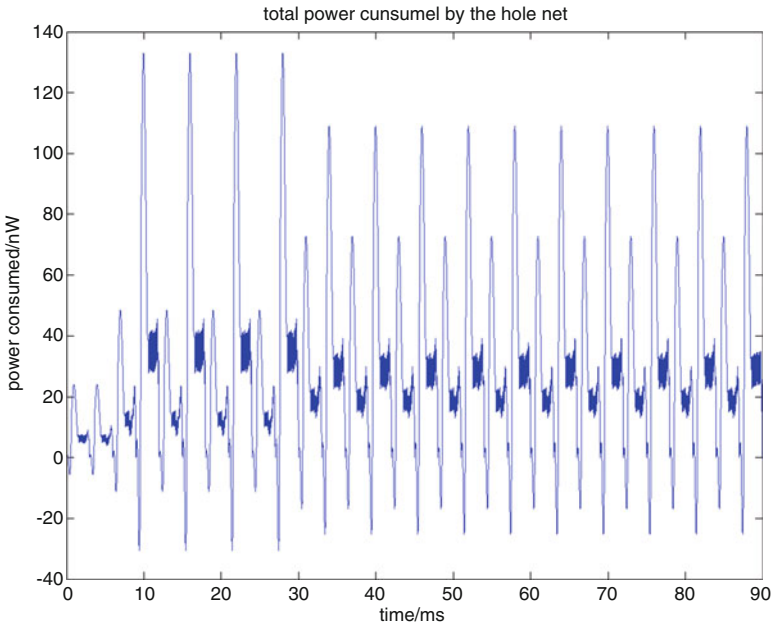


Fig. 3 The total power of the network in the intermittent stimulation

4 Conclusions

We can conclude that energy neural coding possess five properties as follows:

1. Nervous energy coding based on a neuronal network has only correspondence between the membrane potential with nervous energy. Therefore it is very possible that the nervous energy can encode the neural information.
2. The nerve energy encoding can automatically assign the nervous energy corresponding to each frequency of synchronous oscillations in accordance with the allocation ratio in the total energy of the network, and separate each neuronal group from the network.
3. The neural energy can be superimposed. This facilitates modeling and computational analysis for high-dimensional nonlinear complex neural network composed of the large-scale neurons.
4. The nervous energy can be a carrier relating neural information with cerebral blood flow.
5. The neural energy coding theory can ignore the membrane potential of single neuron, and determine the working status of the brain by the energy in the local network.

Acknowledgments This work was supported by the National Natural Science Foundation of China under Grant No. 11232005 and No. 11172086.

References

1. Shun∼ichi Amari, Hiroyuki N. Difficulty of Singularity in population coding. *Neural Computation*, (2005), 17(4): 839∼858
2. Michael S. Gazzaniga, Richard B. Ivry, George R. Mangun, *Cognitive Neuroscience. The Biology of the Mind* (3rd Edition). Norton & Company, Inc. (2009)
3. David W. McLaughlin, Ruling out and ruling in neural codes. *Proc. Natl. Acad. Sci. PNAS*. 5936–5941, Vol. 106, No. 14 (2009)
4. Gopathy Purushothanman & David C Bradley, Neural population code for fine perceptual decisions in area MT. *Nature Neuroscience*. Vol.8, 99–106 (2005)
5. Rama Natarajan, Quentin J.M. Huys, Peter Dayan, and Richard S. Zemel, Encoding and decoding spikes for dynamic stimuli. *Neural Computation*. 20, 2325–2360 (2008)
6. Simon B. Laughlin1, Terrence J. Sejnowski., Communication in neural networks. *Science.* 301, 1870 (2003)
7. Fahmeed Hyder, Douglas L. Rothman & Robert G. Shulman., Total neuroenergetics support localized brain activity: Implications for the interpretation of fMRI. *Proc. Natl. Acad. Sci. PNAS.* Vol.99, No.16, 10771–10776 (2002)
8. Arien J. Smith, Hal Blumenfeld, Kevin L. Behar, Douglas L. Rothman & Robert G. Shulman., Cerebral energetics and spiking frequency: The neurophysiological basis of fMRI. *Proc. Natl. Acad. Sci. PNAS.* Vol.99, No.16, 10765–10770 (2002)
9. Marcus E. Raichle & Debra A. Gusnard., Appraising the brain's energy budget. *Proc. Natl. Acad. Sci. PNAS.* Vol.99, No.16, 10237–10239 (2002)
10. Simon B. Laughlin., Energy as a constraint on the coding and processing of sensory information. *Current Opinion in Neurobiology*. 11:475–480 (2001)
11. Levy WB, Baxter RA. Energy efficient neural codes. *Neural Comp.* 8:531(1996)
12. Michael D. Fox, Marcus E. Raichle, Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature.* Vol.8, September, 700–711 (2007)
13. Mikail Rubinov, Olaf Sporns, Jean-Philippe Thivierge, Michael Breakspear, Neurobiologically Realistic Determinants of Self-Organized Criticality in Networks of Spiking Neurons. PLoS Computational Biology. 7(6): e1002038 (2011)
14. B. Huang, Neural networks and advanced function of brain. *Science Press*. (2000)
15. Theo Vosse, Garard Kempen, The unification space implemented as a localist neural net: predictions and error-tolerance in a constraint-based parser. *Cognitive Neurodynamic*. 3:331–346 (2009)
16. Joerg F. Hipp, Andreas K. Engel, Markus Siegel, Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron*. 69, 387–396. (2011)
17. Thomas Wennekers, Gunther Plam, Syntactic sequencing in Hebbian cel assemblies. *Cognitive Neurodynamics*. 3:429–441 (2009)
18. Rubin Wang, Zhikng Zhang, Guanrong Chen, Energy function and energy evolution on neural population. *IEEE Transactions on Neural Networks*. Vol.19, Issue 3, 535–538 (2008)

Neural Field Dynamics and the Development of the Cerebral Cortex

J.J. Wright and P.D. Bourke

Abstract As neuron precursors divide and generate action potentials, they concurrently undergo apoptosis. We propose that the ensemble of neurons competitively selected is that which generates the maximum synchrony of action potentials. Consequently, local intracortical neurons and patchy connections emerge in "ultra-small" world configurations, forming clearly defined macrocolumns and patch connections in hexagonal array where patch connections have relatively long axons, and less defined structures elsewhere in the cortex. Extension of the competitive principle to local synaptic level explains ante-natal organisation of response properties in primary visual cortex, including effects of stimulus orientation, angle relative to motion, length, and speed, on apparent orientation preference. Post-natal Hebbian consolidation of connections leads to the mature configuration. By implication, superimposed spatio-temporal images, rather than categorical feature responses, form the substrate of cortical information processing.

Keywords Macrocolumns • Cortical patch connections • Cortical embryogenesis • Cortical apoptosis • Synchronous oscillation • Neural small worlds • V1 organization • Cortical visual responses • Cortical feature responses • Spatio-temporal neural images

1 Introduction

During embryogenesis cells that become the neurons of the cerebral cortex divide and migrate to their mature positions while undergoing apoptosis –the cell death of a substantial fraction of their number – ultimately forming minicolumns in their radial disposition, while in their surface disposition they are said to form

J.J. Wright (✉)
Department of Psychological Medicine, School of Medicine, University of Auckland, Auckland, New Zealand
e-mail: jj.w@xtra.co.nz

P.D. Bourke
iVEC@UWA, University of Western Australia, Perth, Australia

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_53

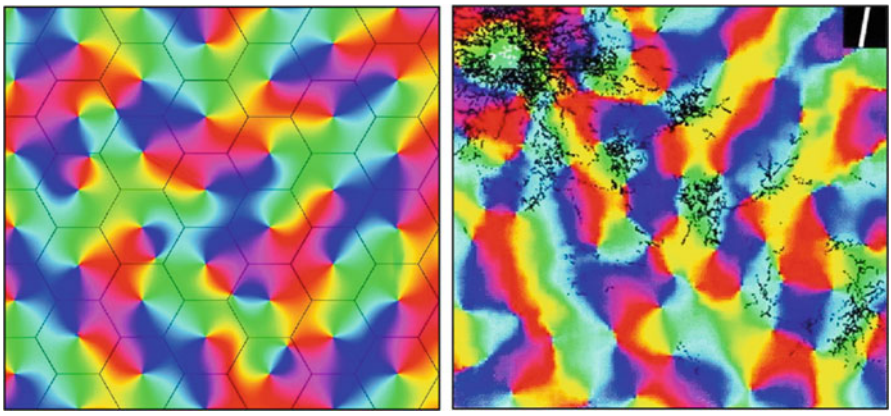


Fig. 1 *Left*: idealized diagram of organization in V1 of OP, 0–180°, about singularities, and with marginal continuity, within an hexagonal framework composed of superficial patch connections. *Right*: experimental data [18]. The *black* areas are patchy synaptic connections

macrocolumns – periodic structures that are most apparent in the primary visual (V1) and some parts of somatosensory (S1) cortices, are each about 300 μm across, and are roughly delineated by superficial patchy connections on the perimeter of each column. Within each column, individual cells in V1 respond with an orientation preference (OP) to visual lines of differing orientation [1]. The surface organization of OP exhibits significant hexagonal rotational periodicity, in which each roughly delineated macro- columnar unit exhibits all values of OP arrayed around a pinwheel [2, 3]. Varying chirality and orientation of the pinwheels achieves continuity of OP at the columnar margins, and patchy connections link areas of similar OP together, "like to like" (Fig. 1).

Hubel [4], in his Nobel address, hailed Mountcastle's original proposal that columns formed fundamental building blocks of cortex as "Surely the most important contribution to understanding of cerebral cortex since Ramon y Cajal". Enthusiasm for the explanatory power of the concept has since waned. Horton and Adams [5] described the cortical column as "a structure without a function", and terminology describing them has become confused [6].

Difficulties arise partly because columnar structure is not clearly apparent outside V1 and S1, and because there is marked interspecies variation in definition of columns even in V1, to the point of apparent absence in small animals. Attempts to model the emergence of columnar organization of OP have also struck considerable difficulty. In some species there is clear emergence of structure ante-natally, rather than post-natally, yet models of the macrocolumn are generally dependent on response to visual features [7]. Which "features" are regarded as fundamental is also controversial, and how this relates to signal processing is problematic.

We have proposed a theory of emergence of cortical columns and their functional significance [8], which differs considerably from all other explanations. We base our explanation on two findings: (1) in vitro, embryonic neurons fire synchronously and self-organize into "small worlds" [9] and (2) synchronous firing of neurons prevents their apoptosis [10].

We assume synchrony and cell survival are causally linked – perhaps because some collective pumping action allows a synchronously coupled assembly of cells to increase their uptake of one or more vital metabolic substances. Therefore the emergent cell network would be that selection of cell types, and their arrangement, that maximizes the amplitude of synchrony for a given limit of total metabolic supply. The consequences of these assumptions are as follows.

2 Selection for "Small-World" Connectivity

Our arguments are based upon properties demonstrated in simulations of cortical gamma synchrony, and travelling waves [11, 12]. Closely situated cells are able to exchange synchronous pulses with smaller phase difference of afferent and efferent pulses. Therefore minimization of the total axonal lengths of their interconnections maximises synchrony magnitude (and uptake) while minimizing axonal metabolic cost.

In the dilute network of neuronal connections, the metric distance of soma separation is proportional to "degree of separation" in the topological sense. Therefore maximization of synchrony, by minimizing axonal lengths, selects a neural network with "ultra-small world" connectivity. This requires, in turn, that the average density of synaptic connectivity decline with distance as a power function [13].

A power function is the sum of exponential functions, and pre-synaptic densities of cortical neurons decline roughly exponentially with distance from the cell body [14]. Therefore small-world connectivity can be approximated from populations of neurons with differing characteristic ranges.

3 Local Variability of Axonal Ranges in the Selected Population

Equal approximations to a power function can be achieved by combining different relative densities of a variety of cell types, each type characterised by axonal length. Simplifying to only two types, Figures 2 and 3 show where long/short axon length is large ($\lambda_\beta \gg \lambda_\alpha$) approximation of a power curve requires the ratio of local neurons to patch neurons be large ($N_\beta \gg N_\alpha$).

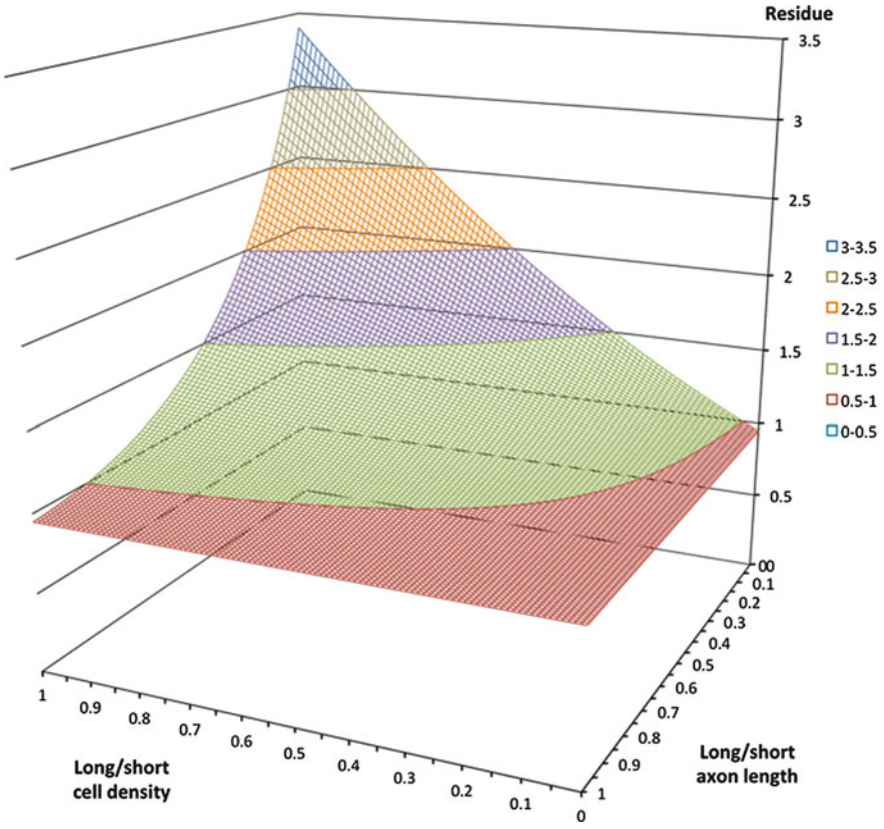


Fig. 2 Equal approximations to a power function can be achieved with a variety of combinations of cell densities and axonal ranges. *Colour margins* show fit-residual isocontours

4 Resolution into Macrocolumns

It can be shown that, where J is the magnitude of synchronous oscillation, and $\mathbf{q}$,$\mathbf{r}$ are positions of excitatory neurons in the cortex,

$$J \propto \int_{\mathbf{q}} \int_{\mathbf{r}} \left(N_\alpha \lambda_\alpha e^{-\lambda_\alpha |\mathbf{q}-\mathbf{r}|} + N_\beta \lambda_\beta e^{-\lambda_\beta |\mathbf{q}-\mathbf{r}|} \right) d\mathbf{q} d\mathbf{r}$$

Therefore synchrony is maximized by selection of that ensemble of cells in which the cells with relatively short but dense axons are closely situated to each other. Such packing forces the cells with long-range axons to form connections at longer range, enforcing a "patchy" connection system. Arrangement in an hexagonal patchwork optimizes this synchrony-facilitating orderliness, but a clearly demarcated arrangement of this type is only possible where

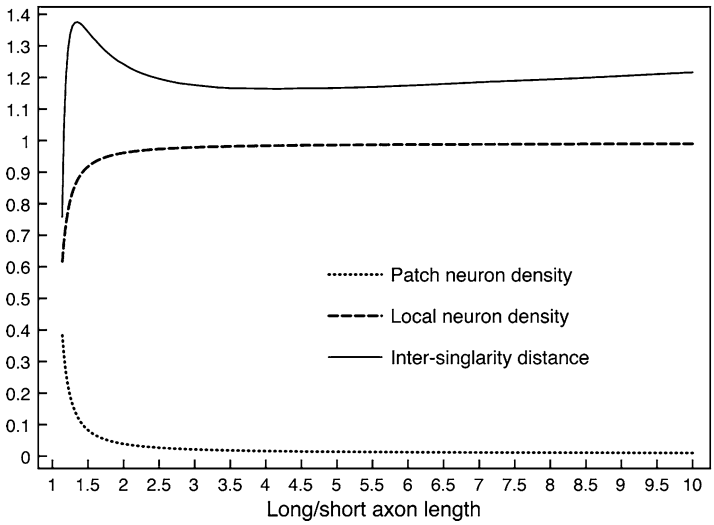


Fig. 3 Cell densities and axonal ranges of local and patch neurons vary approximately inversely for equivalent fits to a power function obtained along the lowest-residual iso-contour in Fig. 2. Intersingularity distance (cp Fig. 4) is comparatively invariant over this range

$$\frac{local\ \ cells}{local\ \ cells\ \ +\ \ patch\ \ cells} = \frac{N_\beta}{N_\alpha + N_\beta} \geq \frac{\pi}{2\sqrt{3}}$$

This follows simply from the ratio of area of a circle to a hexagon, when local cells are enclosed within an hexagonal patch-connection frame. Therefore the absence of a clearly columnar arrangement does not imply a loss of the small world organization, nor does it deny that both short-range local connections, and longer-range functional connections are present – the distinct types are merely more entangled with each other (See Fig. 4).

Thus, variation of the clarity of demarcation of columns in differing cortical areas, and between species, need not reflect major differences in function.

5 A Mobius Map Within Macrocolumns

Restated in physical terms, the maximization of J requires the populations of cells of differing axonal range be geometrically arranged so as to permit maximum resonance throughout the system. Since the amplitude of synchronous oscillation declines with distance of separation of cell bodies, then the system of patch connections and the local neurons within each macrocolumn must achieve a 1:1 connection system, promoting resonance between cells in each macrocolumn, and the surrounding patch system, and thus forming an input map of the cortical

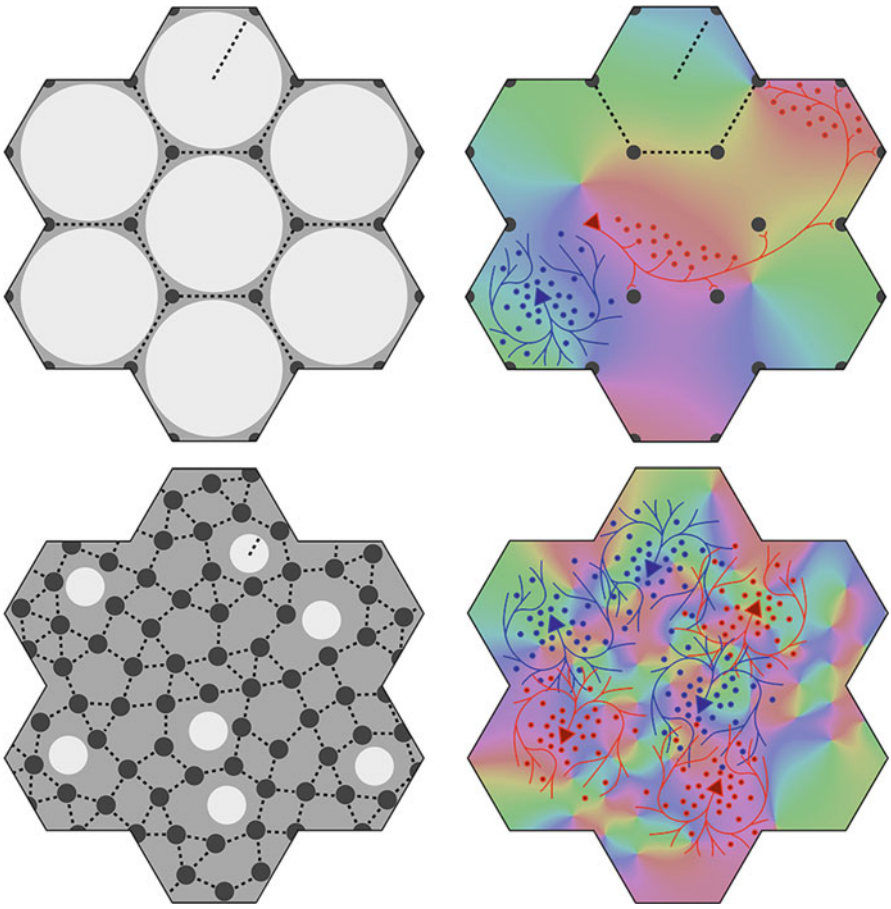


Fig. 4 Variation of the structure of macrocolumns at extremes of the axonal lengths and cell numbers in Fig. 3. *Top*: With large long/short axon length ratio, clearly resolved hexagonal organization emerges, with long (*red*) patch connections linking "like to like", and highly clustered short intracortical axons (*blue*). *Bottom*: near-complete loss of resolution when long/short axon ratio approaches 1

surround, projected onto each macrocolumn. If it is additionally assumed that the competition for crucial resources is not simply between individual neurons, but also between closely situated pre-synapses arising from the same cell, then "winner take all" competition between closely situated synaptic connections would develop, and at equilibrium each cell would then require high firing correlation with some of its neighbors, and low firing correlation with other neighbors – and be correspondingly strongly linked to some neighbors by "saturated" synapses, and weakly to others by "sensitive" synapses. This intra-cellular constraint, along with the requirement to form a 1:1 map of connections between each macrocolumn and its

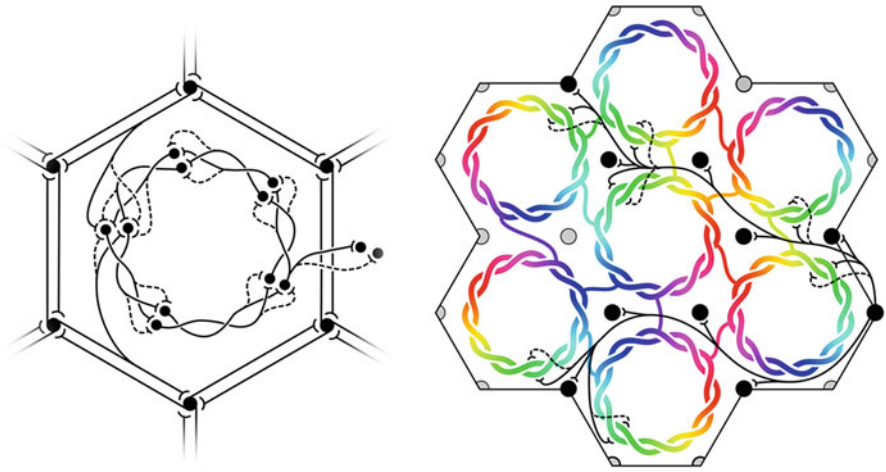


Fig. 5 Maximization of synchrony with local synaptic competition leads to Mobius ordering, within macrocolumns. *Left*: Equilibrium disposition of saturated (*solid*) and sensitive (*dashed*) synapses in the developing neocortex. *Right*: "Like to like" saturated patchy connections map the same part of the surrounding cortical field onto homologous cell positions on the Möbius configuration

patchy-connection surround, can be met if the connections within the macrocolumn form a closed system analogous to a Mobius strip. Figure 5 (left) shows how a dynamic equilibrium of synaptic connections can thus be struck. The mapping of the patch system onto the macrocolumn can be expressed as

$$\mathbf{P}\left(\left|\mathbf{R}-\mathbf{C}_j\right|,\vartheta\right)\rightarrow\mathbf{p}^{[2]}\left(\left|\mathbf{r}-\mathbf{C}_j\right|,\pm\vartheta+\varphi\right)$$

where **P** is the plane of the patchy connections, and **R** are cortical positions with reference to these, while **p** and **r** are corresponding plane and positions within a macrocolumn. The square bracketed superscript [2] indicates the map's resemblance, if viewed from a third dimension, to a 2:1 map formed by squaring a complex vector. $\mathbf{C}_j$ is the origin of both **P** and $\mathbf{p}^{[2]}$ for the $j-th$ local map, and corresponds to the position of the OP singularity in that macrocolumn. ϑ is the polar angle of **R,** chirality of the local map is indicated by $\pm\vartheta$, and φ is the orientation of the local map relative to the global map.

Figure 5 (right) shows further requirements for synchrony maximization. On the input map, radial lines on the surrounding cortex must map about a centre, analogous to an OP singularity. The Mobius strip-like folding of connections means that "OP" from 0 to 180° is mapped 0 to 360° about the singularity – concealing a superposition of diametrically opposite lines projected from the cortex to the macrocolumn. To further increase resonance, patch connections must link "like to like" OP in forming multiple 1:1 maps, and adjacent macrocolumns must also be so arranged as to increase resonance by linking "like to like" map positions on adjacent

macrocolumns, as closely as possible within a roughly hexagonal framework. Thus, the properties of V1 sketched in Fig. 1 are reproduced.

These considerations apply to the development of the cortex prior to the beginning of vision at birth.

6 Consequences of Eye-Opening and Development of Responses to Stimuli

The dynamic equilibrium of synaptic activity described above presumably gives rise to some persistence of the structure on Hebbian principles, but subsequent to birth, inputs from the direct visual pathway must produce strong perturbations from equilibrium, and overwriting of the Mobius structure by later learning.

To fire rapidly in the mature brain, individual neurons in V1 require direct visual input from their receptive fields, in summation with "contextual" signals transferred laterally by the patch system – and on firing, they give rise to further, laterally spreading, contextual signals, forming travelling waves.

The transfer of these waves from the wider cortex to each macrocolumn, considered as transfer to a Mobius-like map, is then a mapping with time lags of an image, O, given by

$$O\left(\mathbf{P}, t\right) \rightarrow O\left(\mathbf{p}^{[2]}, t + \frac{|\mathbf{R} - \mathbf{r}|}{v}\right)$$

where v is the wave speed.

This permits the simulation of V1 neuron responses to visual moving lines, by calculating the corollary inputs reaching a macrocolumn at the time the direct visual input reaches the same macrocolumn, as shown in Fig. 6.

The results match the experimental data of Basole et al. [15] – data considered incompatible with earlier notions of V1 neuron OP specificity.

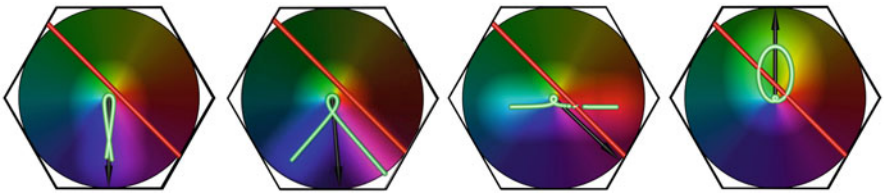


Fig. 6 In V1, lateral transmission via patchy connections, plus input from direct visual pathway, summates above threshold for action potentials. Results for a line moving from left to right, and oriented at 45° to the line of passage, with stimulus speed/wave speed, left to right, 0.1, 0.5, 1.0, 1.5

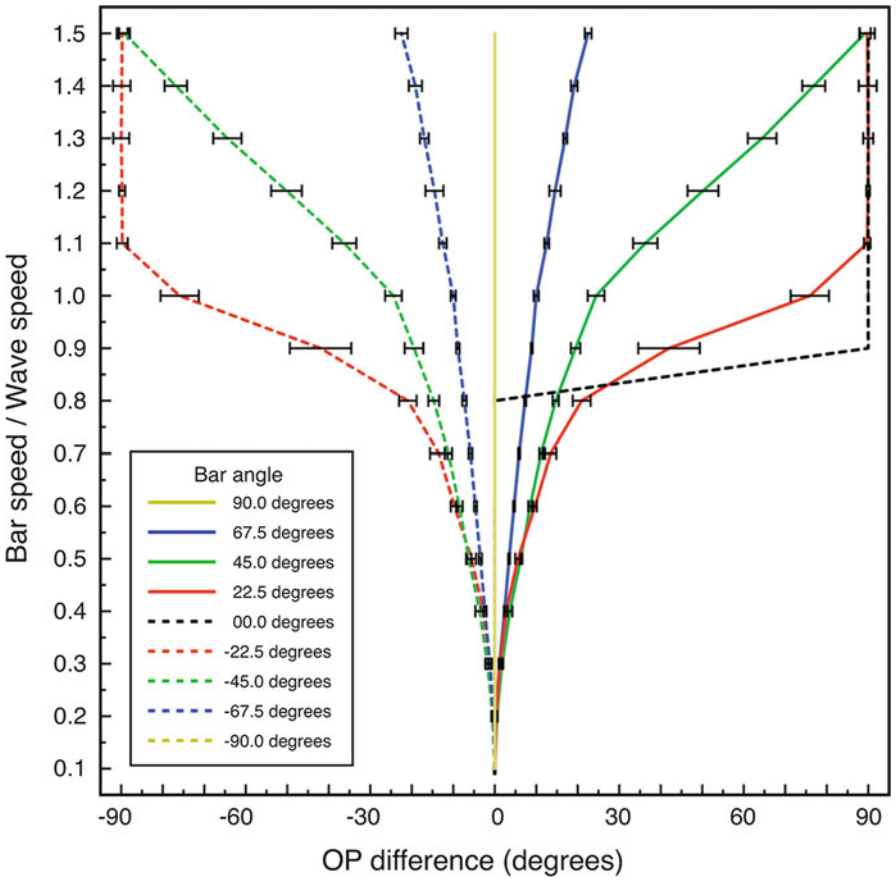


Fig. 7 Simulation results: change in apparent OP, and standard error of the estimate, as a function of bar speed to wave speed, for lines at different orientations to their directions of motion. Bar length 6 macrocolumn diameters

In the case of a visual line of given length, the selective neuron responses vary not only with line orientation, but also with its inclination to the direction of travel, and speed, as shown in Fig. 7. Notably, the "classic" property of elementary OP is seen only for low stimulus speeds.

The results of Basole et al. have been otherwise explained by assuming V1 neurons show specific tuning to combinations of object orientation, spatial frequency and temporal frequency [16]. We have shown that Issa et al.'s description is equivalent to the effect of Hebbian learning upon the properties demonstrated in our simulation. Overwriting of the pre-natal Mobius maps by post-natal Hebbian learning also explains the consolidation of "like to like" patchy connections, and the continuity and completeness required by dimension-reduction models [7] of response maps.

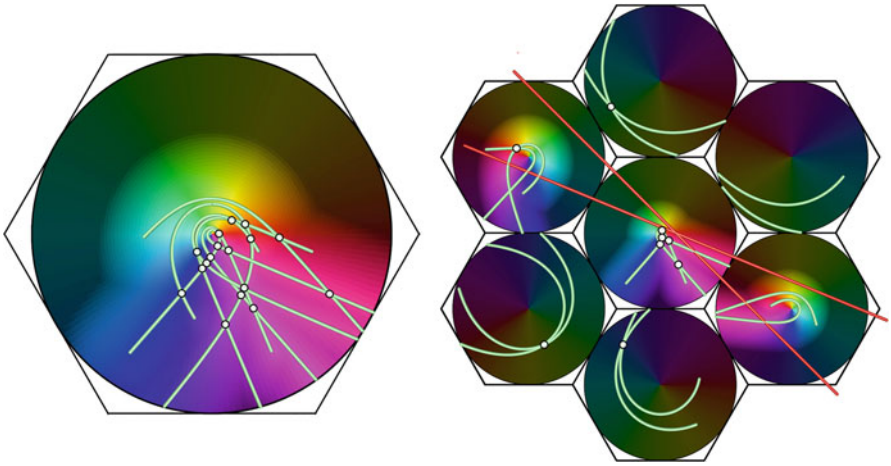


Fig. 8 Inter-areal interactions. *Right*: a complex of macrocolumns in V1, three of which are concurrently activated by a pair of moving visual lines. *Left*: A column in a higher visual area, which is co-resonant at equilibrium with the lower-centre complex. Transient signals from the lower to the higher system, result in superposition of signals representing different positions on, and times of passage of, the moving lines

7 Interactions Between Cortical Areas

The same principles of self-organization should apply widely in the cortex. If so, then cortical areas self-organized into patchy connections and macrocolumns could also interact with other cortical areas via cortico-cortical fibers. Favorov and Kursun [17] have demonstrated the potential of neocortical layer 4 to permit near-linear superposition of impulses relayed via cortico-cortical fibers. Co-resonance among sets of macrocolumns at multiple scales would thus be possible. With such an ante-natal organization, after birth, signals from the environment could then produce complex contextual superpositions of waves relayed between groups of Mobius maps.

One such instance is modeled in Fig. 8, showing how neuron responses to compound aspects of moving visual stimuli could arise. Similarly, return transmission from higher to lower cortical areas might mediate some aspects of attention.

8 Conclusions

Our account emphasizes the importance of cooperative and competitive processes in embryonic development, in addition to genetically programmed developmental cascades. It explains diverse aspects of neural architecture and function in a unified way, including the ante-natal emergence of functional structures in V1,

the origin of macrocolumns and superficial patch connections, their tendency to hexagonal periodicity, their interareal and interspecies variation, and the response properties of V1 neurons, including the post-natal abnormalities produced by visual deprivation. The model can also be combined with models utilizing Turing pattern formation, to account for the origin of OD columns. The principles may (with appropriate adjustment for local cell forms and the organization of input pathways) be applicable to other sensory modes, and even motor cortex. At the time of writing preliminary evidence has been obtained of the existence of Mobius-like organization in macrocolumns of the sensorimotor cortex.

As well as the capacity to explain empirical data, there are interesting implications for theories of neural information processing. The ultra-small-world configuration implies that the organization is near a maximum for speed and energy-efficiency of processing. Synaptic storage capacity can reach theoretical maximum entropy, under the assumption that available metabolic resources are sufficient to sustain only 50 % of synapses at maximum saturation. The modular organization offers a potential for the rapid expansion of the cerebral cortex seen in its evolution.

Perhaps most importantly, the model indicates that brain function may be built upon a primary, ab initio, spatial organization that can act as a reference framework for sensory inputs from the environment as well as for internal dynamics of the Freeman type. Not abstract features, but superimposed spatio-temporal images, may form the kernel of cerebral information processing.

Dedicated to Adrienne Edith Wright.

References

1. Hubel, D. H., and Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* 148, 574–591.
2. Paik, S.-B., and Ringach, D. L. (2011). Retinal origin of orientation maps in visual cortex. *Nat. Neurosci.* 14, 919–925.
3. Muir, D. R., Da Costa, N. M. A., Girardin, C. C., Naaman, S., Omer, D. B., Ruesch, E., et al. (2011). Embedding of cortical representations by the superficial patch system. *Cereb. Cortex* 21, 2244–2260.
4. Hubel, D. H. (1981). Acceptance Address. The Nobel Prize in Medicine.
5. Horton, C. H., and Adams, D. L. (2005). The cortical column: a structure without a function. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 837–862.
6. Rakic, P. (2008). Confusing cortical columns. *PNAS* 105, 12099-12100.
7. Swindale, N. V. (1996). The development of topography in the visual cortex: a review of models. *Network* 7, 161–247.
8. Wright, J.J., and Bourke, P.D. (2013). On the dynamics of cortical development: synchrony and synaptic self-organization. *Frontiers in Computational Neuroscience*. DOI: 10.3389/fncom.2013.00004.
9. Downes, J. H., Hammond, M. W., Xydas, D., Spencer, M., Becerra, V. M., Warwick, K., et al. (2012). Emergence of a small-world functional network in cultured neurons. *PLoS Comput. Biol.* 8:e1002522. doi: 10.1371/journal.pcbi.1002522.

10. Heck, N., Golbs, A., Riedemann, T., Sun, J.-J., Lessmann, V., and Luhmann, H. J. (2008). Activity dependent regulation of neuronal apoptosis in neonatal mouse cerebral cortex. *Cereb. Cortex* 18, 1335–1349.
11. Wright, J. J. (2009). Generation and control of cortical gamma: findings from simulation at two scales. *Neural Netw.* 22, 373–384.
12. Wright, J. J. (2010). Attractor dynamics and thermodynamic analogies in the cerebral cortex: synchronous oscillation, the back- ground EEG, and the regulation of attention. *Bull. Math. Biol.* 73, 436–457.
13. Cohen, R., and Havlin, S. (2003). Scale- free networks are ultra-small. *Phys. Rev. Lett.* 90:058701. doi: 10.1103/ PhysRevLett.90.058701
14. Braitenberg, V., and Schüz, A. (1991). *Anatomy of the Cortex: Statistics and Geometry.* Berlin, New York: Springer.
15. Basole, A., White, L. E., and Fitzpatrick, D. (2003). Mapping of multiple features in the population response of visual cortex. *Nature* 423, 986–990.
16. Issa, P., Rosenberg, A., and Husson, T. R. (2008). Models and measurements of functional maps in V1. *J. Neurophysiol.* 99, 2745–2754.
17. Favorov, O.V., and Kursun, O. (2011). Neocortical layer 4 as a pluripotent function linearizer. *J. Neurophysiol*. 105, 1342-1360. doi: 10.1152/jn.00708.2010.
18. Bosking, W. H., Zhang, Y., Schofield, B., and Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J. Neurosci.* 17, 2112–2127.

Signal Processing by Ephaptic Coupling of Hodgkin-Huxley Axons

Masashi Yamada, Hideaki Konno, and Nobuyuki Takahashi

Abstract We examined a signal processing on Hodgkin-Huxley model repeats under ephaptic coupling resembling the early stage of axon bundles in auditory nerve. We show impulse conduction synchronizations in axons partially myelinated by an oligodendrocyte with propagation time difference has a potential to select functionally maturating axons. We discuss the possible feedback process for maintaining parallel-aligned axons corresponding to the signal flow.

Keywords Signal processing • Ephaptic coupling • Axon bundle • Hodgkin-Huxley model • Impulse conduction • Myelinated axon • Functional maturation • Auditory nerve

1 Introduction: Parallel Axons Myelination in Maturation Process

Ephaptic coupling of parallel axons can underlie functional activity in the brain through the synchronization of action potentials [1] with adjustment of signal conduction [2]. Little evidence has been found for the ephaptic synchronization in mammalian axons. However, the local field potential (LFP) provides shifts in the timing of action potential generation [3]. We expect inevitable interaction between the axons partially encapsulated in an oligodendrocyte, so that parallel alignment of axons in auditory and optic nerve, for example, is maintained in the maturation

M. Yamada
Department of Information Science, Hokkaido University of Education, Hakodate, 1-2 Hachimancho, Hakodate, Hokkaido 040-8567, Japan

Uryu Elementary School, Dai-7-chonai, Uryu-cho, Uryu-gun, Hokkaido 078-2637, Japan

H. Konno • N. Takahashi (✉)
Department of Information Science, Hokkaido University of Education, Hakodate, 1-2 Hachimancho, Hakodate, Hokkaido 040-8567, Japan
e-mail: takahashi.nobuyuki@h.hokkyodai.ac.jp

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_54

process with precise relationships between soma positions and axon terminal positions holding the sensory locational information. The process is believed to progress depending on the signal flowing in the axon.

We examined a special signal processing on Hodgkin-Huxley (HH) model repeats under ephaptic coupling corresponding to the early stage of axon bundles. We show impulse synchronization similar to those found previously [1] in axons partially myelinated by an oligodendrocyte with propagation time difference has a potential to select functionally maturating axons. We discuss the possible feedback process for maintaining parallel-aligned axons corresponding to the signal flow in the maturation processes.

2 Target Material: Auditory Cochlear Axons

Auditory cochlear axons show loosely packed parallel alignment [4]. The auditory nerve starting from the spiral ganglion changes its myelin from schwann cell to oligodendrocyte at the schwann-glial border. The central axon bifurcates in the cochlear nucleus to form an ascending branch to anteroventral cochlear nucleus and a descending branch to dorsal cochlear nucleus [5]. We focus in the axon path from the main branch point to the terminal in the anteroventral cochlear nucleus. The length is approximately 2 mm.

Our discussion is limited to the prenatal period, in which the developments in auditory systems are progressing but not completed [6]. We assume that the myelin is partially constructed and the partial myelin is served by an oligodendrocyte on the adjacent parallel axons.

3 Method/Models: Simulation of Electrostatic Coupled HH Model Repeats

A ladder-like circuit consisting of tandem repeats of HH active nodes and intracellular resistances was analyzed computational. The HH nodes represent the node of Ranvier and or the partially myelinated part having HH activity. The intracellular resistances are models of the myelinated axons between the active nodes. The intracellular resistances were calculated as homogeneous cylindrical conductor of the axon diameter with the internal resistivity 100 Ω cm. Extra-cellular resistances were set for the direct connecting points between HH nodes of adjacent axons A and B interacting extracellular because of the partial myelination as partial encapsulation for the both HH nodes of adjacent axons. The values of extracellular resistances are chosen by trial. HH equations are same as the original [7] changed for the sign of the membrane potential. HH parameters for giant nerve fiber [7] are scaled with a factor

12 [8] for representing the short absolute refractory period ∼1 ms for mammalian auditory nerve fibers. The number of HH nodes representing the partial myelin is 5 with node length 100 μm and internode length 50 μm, total 700 μm for each axon (diameters da = 0.8 μm and db = 2.0 μm).

4 Results: Impulse Propagation Synchronization in Adjacent Axons

Synchronizations of action potentials in the interacting two axons were simulated depending on the extracellular resistances through extracellular potential differences between HH active nodes as same as the previous works [1, 2].

Figure 1 shows the synchronization of impulses between the parallel axons of thin (0.8 μm) and thick (2.0 μm) diameters. The action potential peaks show synchronization for HH nodes grounded with (a) 3 M ohm extracellular resistances for the 1–3 nodes and grounded 0 and 4 nodes comparing with (b) zero extracellular resistances for parallel axons of 0.8 μm and 2.0 μm diameters. This type synchronization takes place only for the case action potentials in both axons coincide to fire within the time window related with the width of the impulse.

Figure 2 shows extracellular resistance dependence of the time (ta4 and tb4) of arrival of action potential peaks to the HH node 4 for the axon A of 0.8 μm and the axon B of 2.0 μm diameters, respectively. The effect on the synchronization takes place mainly for the conduction delay for the thick axon gradually depending on the extracellular resistances.

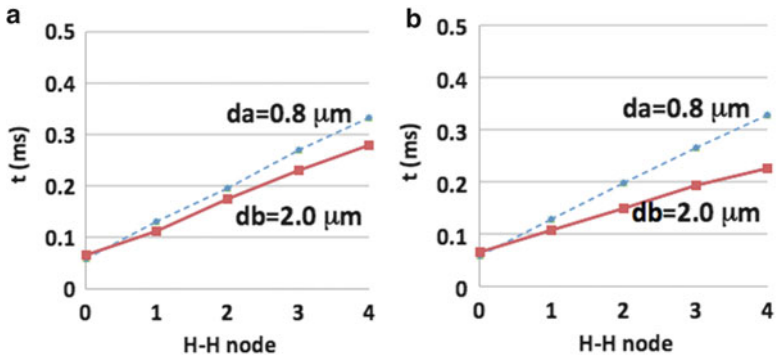


Fig. 1 The action potential peaks show synchronization for HH nodes grounded with (**a**) 3 M ohm extracellular resistances at nodes 1–3 compared with (**b**) zero extracellular resistances, for parallel axons (diameters da = 0.8 μm and db = 2.0 μm)

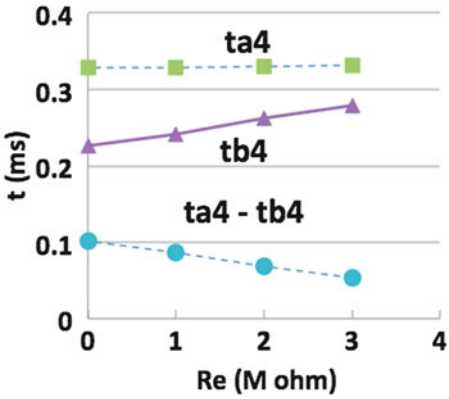


Fig. 2 The action potential peaks reach at node 4 at time ta4 and tb4 with synchronization depending on the extracellular resistances Re for parallel axons of 0.8 μm and 2.0 μm diameters, respectively

5 Discussions and Conclusions

Figure 1 indicates that ephaptic current between the axons causes decrease of the velocity of thick axon resulting the synchronized arrival of the action potentials to the end node. Thin axon, which needs longer time to reach the end node comparing with the thick axon without the ephaptic interaction, causes time-out for synchronization with other neurons. The ephaptic interaction causes sending signals within the time window of the coincident processing on next step. It is effective to maintain functional roles of the thin axon and to provide growth factors of myelination for thin axons in the maturation stages. The extracellular resistance 1 M ohm causes extracellular potential around −25 mV at peak. This high value of extracellular potential is provided by the special structure like that assumed here for the partial encapsulation of adjacent nodes of parallel axons.

The gradual change of synchronization depending on the extracellular resistivity in Fig. 2 provides a possibility of signal conduction regulation by extracellular resistance of glia contribution.

Here we showed that the ephaptic coupling causes (a) signal alignment with adjustment of sending speed. The model also shows several signal processing, i.e., (b) new signal generation and (c) simultaneous signal sending back, and (d) the shifts in the timing of the action potential generation. These processing transfer signals between parallel axons. We did not examine here the three processing, (b), (c) and (d), which will all be reported near future.

In summary, an ephaptic synchronization of impulse conduction is analyzed computational for a ladder-like circuit consisting of tandem repeats of Hodgkin-Huxley active nodes resembling prenatal auditory nerve. Possibility of signal conduction regulation controlled by local field signal transfer between parallel axons is indicated in the model. Adaptive switching of synchronized circuits based on the "memory" in axons and glia without adaptation in synaptic connectivity is targeted in future study.

Acknowledgments Prof. Dr. Wolfgang Paul, Univ. Halle-Wittenberg, is greatly acknowledged for staying of N.T. in his group. Dr. R. D. Troub, IBM Research, and Prof. K. Aihara, Tokyo University, are acknowledged for their encouragement to this study. Prof. Dr. F. W. Ohl, Leibniz Institute for Neurobiology, Magdeburg and Prof. Erik Fransen, KTH Royal Institute of Technology, Stockholm are also acknowledged for the helpful discussions.

References

1. Traub, R. D., et al. Simulation of hippocampal afterdischarges synchronized by electrical interactions. Neuroscience **14** (1985) 1033–1038.
2. Reutskiy, S., Rossoni, E., and Tirozzi, B., Conduction in bundles of demyelinated nerve fibers: computer simulation. Biol. Cybern. **89** (2003) 439–448.
3. Anastassiou, C. A., Perin, R., Markram, H., & Koch, C., Ephaptic coupling of cortical neurons. Nature neurosci. **14** (2011) 217–223.
4. Perge, Ja'nos A., et al. Why do axons differ in caliber?. The Journal of Neuroscience **32** (2012) 626–638.
5. Ryugo, D. K., and E. M. Rouiller. Central projections of intracellularly labeled auditory nerve fibers in cats: Morphometric correlations with physiological properties. Journal of Comparative Neurology **271** (1988) 130–142.
6. Moore, Jean K., Perazzo, Lucy M., and Braun, Alex, Time course of axonal myelination in the human brainstem auditory pathway. Hearing Research **87** (1995) 21–31.
7. Hodgkin, A. L., & Huxley, A. F., A quantitative description of membrane current and its application to conduction and excitation in nerve, The Journal of physiology, **117** (1952) 500.
8. Motz, H., and Rattay. F., A study of the application of the Hodgkin-Huxley and the Frankenhaeuser-Huxley model for electrostimulation of the acoustic nerve, Neuroscience **18** (1986) 699-712.

Part IX
Quantum Cognition

Modeling Concept Combinations in a Quantum-Theoretic Framework

Diederik Aerts and Sandro Sozzo

Abstract We present a modeling for conceptual combinations which uses the mathematical formalism of quantum theory. Our model faithfully describes a large amount of experimental data collected by different scholars on concept conjunctions and disjunctions. Furthermore, our approach sheds a new light on long standing drawbacks connected with vagueness, or fuzzyness, of concepts, and puts forward a completely novel possible solution to the 'combination problem' in concept theory. Additionally, we introduce an explanation for the occurrence of quantum structures in the mechanisms and dynamics of concepts and, more generally, in cognitive and decision processes, according to which human thought is a well structured superposition of a 'logical thought' and a 'conceptual thought', and the latter usually prevails over the former, at variance with some widespread beliefs.

Keywords Quantum theory • Combination problem • Decision processes • Conceptual thought • Hilbert space

1 Conceptual Vagueness and the Combination Problem

According to the 'classical view', going back to Aristotle, a concept is considered practically as a 'container of its instantiations'. This view was already criticized by Wittgenstein but definitely put at stake by Rosch's work on color, which showed that subjects rate concept membership of an instance as graded (or fuzzy, or vague). Hence mathematical notions to model such conceptual fuzziness were put forward. But, Osherson and Smith's examples of concept conjunctions revealed a fundamental difficulty of classical (fuzzy) set-theoretic approaches to model

D. Aerts (✉)
Center Leo Apostel (Clea), Brussels Free University (VUB), Krijgskundestraat 33, 1160 Brussels, Belgium
e-mail: diraerts@vub.ac.be

S. Sozzo
School of Management and IQSCS, University of Leicester, University Road LE1 7RH, Leicester, United Kingdom
e-mail: ss831@le.ac.uk

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_55

such conjunctions. These authors considered the concepts *Pet* and *Fish* and their conjunction *Pet-Fish*, and observed that, while an exemplar such as *Guppy* is a very typical example of *Pet-Fish*, it is neither a very typical example of *Pet* nor of *Fish*. Hence, the typicality of a specific exemplar with respect to the conjunction of concepts shows an unexpected behavior from the point of view of classical set and probability theory. That the 'Pet-Fish problem' (also known as 'Guppy effect') indeed revealed a fundamental conflict with classical structures, was confirmed in a crucial way by Hampton's studies [1, 2], which measured the deviation from classical set-theoretic membership weights of exemplars with respect to pairs of concepts and their conjunction or disjunction. Hampton showed that people estimate membership in such a way that the membership weight of an exemplar of a conjunction (disjunction) of concepts is higher (lower) than the membership weights of this exemplar for one or both of the constituent concepts. This phenomenon is referred to as 'overextension' ('underextension'). Several experiments have since been performed and many elements have been taken into consideration with respect to this 'combination problem' to provide a satisfactory mathematical model of concept combinations. Notwithstanding this, a model that represents the combination of two or more concepts from the models that represent the individual concepts still does not exist.

Meanwhile, it has been shown that quantum structures are systematically present in domains of the social sciences, e.g., in the modeling of cognitive and decision processes [3–10]. As such, we have developed a specific quantum-theoretic approach to model and represent concepts [3–5, 11–15]. This 'quantum cognition approach' was inspired by our research on the foundations of quantum theory, the origins of quantum probability and the identification of genuine quantum aspects, such as contextuality, emergence, entanglement, interference, superposition, in macroscopic domains. A 'SCoP formalism' was worked out which relies on the interpretation of a concept as an 'entity in a specific state changing under the influence of a context' rather than as a 'container of instantiations'. This representation of a concept was new with respect to traditional approaches and allowed us to elaborate a quantum representation of the guppy effect explaining at the same time its occurrence in terms of contextual influence. Successively, the mathematical formalism of quantum theory was employed to model the overextension and underextension of membership weights measured by Hampton [1, 2]. More specifically, the overextension for conjunctions of concepts measured by Hampton [1] was described as an effect of quantum emergence, interference and superposition, which also play a fundamental role in the description of both overextension and underextension for disjunctions of concepts [2]. Furthermore, a specific conceptual combination experimentally revealed the presence of another genuine quantum effect, namely, entanglement [11–13, 15]. In this paper, we present an elaborate and unified quantum-mechanical representation of concept combinations in Fock space which faithfully agrees with different sets of data collected on concept combinations. Our modeling suggests an explanatory hypothesis according to which human thought is a quantum superposition of an 'emergent thought' and a 'logical thought', and that the quantum-theoretic approach in Fock space enables this approach to general human thought, consisting of a superposition of these two modes, to be modeled.

2 Quantum Modeling in Fock Space

Our quantum modeling approach for the combination of two concepts is set in a Fock space $\mathscr{F}$ which consists of two sectors: 'sector 1' is a Hilbert space $\mathscr{H}$, while 'sector 2' is a tensor product Hilbert space $\mathscr{H} \otimes \mathscr{H}$.

Let us now consider the membership weights of exemplars of concepts and their conjunctions/disjunctions measured by Hampton [1, 2]. He identified systematic deviations from classical set (fuzzy set) conjunctions/disjunctions, an effect known as 'overextension' or 'underextension'.

Let us start from conjunctions. It can be shown that a large part of Hampton's data cannot be modeled in a classical probability space satisfying the axioms of Kolmogorov [5]. Indeed, the membership weights $\mu_x(A)$, $\mu_x(B)$ and $\mu_x(A \text{ and } B)$ of an exemplar x for the concepts A, B and 'A and B' can be represented in a classical probability model if and only if the following two conditions are satisfied (see [5] for a proof)

$$\Delta_x^c = \mu_x(A \text{ and } B) - \min(\mu_x(A), \mu_x(B)) \leq 0 \tag{1}$$

$$0 \leq k_x^c = 1 - \mu_x(A) - \mu_x(B) + \mu_x(A \text{ and } B) \tag{2}$$

Let us consider a specific example. Hampton estimated the membership weight of *Mint* with respect to the concepts *Food*, *Plant* and their conjunction *Food and Plant* finding $\mu_{Mint}(Food) = 0.87$, $\mu_{Mint}(Plant) = 0.81$, $\mu_{Mint}(Food \text{ and } Plant) = 0.9$. Thus, the exemplar *Mint* presents overextension with respect to the conjunction *Food and Plant* of the concepts *Food* and *Plant*. We have in this case $\Delta_x^c = 0.09 \not\leq 0$, hence no classical probability model exists for these data.

Let us now come to disjunctions. Also in this case, a large part of Hampton's data [2] cannot be modeled in a classical Kolmogorovian probability space, due to the following theorem. The membership weights $\mu_x(A)$, $\mu_x(B)$ and $\mu_x(A \text{ or } B)$ of an exemplar x for the concepts A, B and 'A or B' can be represented in a classical probability model if and only if the following two conditions are satisfied (see [5] for a proof)

$$\Delta_x^d = \max(\mu_x(A), \mu(_x B)) - \mu_x(A \text{ or } B) \leq 0 \tag{3}$$

$$0 \leq k_x^d = \mu_x(A) + \mu_x(B) - \mu_x(A \text{ or } B) \tag{4}$$

Let us again consider a specific example. Hampton estimated the membership weight of *Donkey* with respect to the concepts *Pet*, *Farmyard Animal* and their disjunction *Pet or Farmyard Animal* finding $\mu_{Donkey}(Pet) = 0.5$, $\mu_{Donkey}(Farmyard\ Animal) = 0.9$, $\mu_{Donkey}(Pet \text{ or } Farmyard\ Animal) = 0.7$. Thus, the exemplar *Donkey* presents underextension with respect to the disjunction *Pet or Farmyard Animal* of the concepts *Pet* and *Farmyard Animal*. We have in this case $\Delta_x^d = 0.2 \not\leq 0$, hence no classical probability model exists for these data.

It can be proved that a quantum probability model in Fock space exists for Hampton's data, as follows [5, 12, 13].

Let us start from the conjunction of two concepts. Let x be an exemplar and let $\mu_x(A)$, $\mu_x(B)$, $\mu_x(A \text{ and } B)$ and $\mu_x(A \text{ or } B)$ be the membership weights of x with respect to the concepts A, B, 'A and B' and 'A or B', respectively. Let $\mathscr{F} = \mathscr{H} \oplus (\mathscr{H} \otimes \mathscr{H})$ be the Fock space where we represent the conceptual entities. The concepts A, B and 'A and B' are represented by the unit vectors $|A_c(x)\rangle$, $|B_c(x)\rangle$ and $|(A \text{ and } B)_c(x)\rangle$, respectively, where

$$|(A \text{ and } B)_c(x)\rangle = m_c(x)e^{i\lambda_c(x)}|A_c(x)\rangle \otimes |B_c(x)\rangle + n_c(x)e^{i\nu_c(x)}\frac{1}{\sqrt{2}}(|A_c(x)\rangle + |B_c(x)\rangle) \quad (5)$$

The numbers $m_c(x)$ and $n_c(x)$ are such that $m_c(x), n_c(x) \geq 0$ and $m_d^2(x) + n_c^2(x) = 1$. The decision measurement of a subject who estimates the membership of the exemplar x with respect to the concept 'A and B' is represented by the orthogonal projection operator $M_c \oplus (M_c \otimes M_c)$ on $\mathscr{F}$, where M_c is an orthogonal projection operator on $\mathscr{H}$. Hence, the membership weight of x with respect to 'A and B' is given by

$$\mu_x(A \text{ and } B) = \langle (A \text{ and } B)_c(x)|M_c \oplus (M_c \otimes M_c)|(A \text{ and } B)_c(x)\rangle = m_c^2(x)\mu_x(A)\mu_x(B) + n_c^2(x)\left(\frac{\mu_x(A) + \mu_x(B)}{2} + \Re\langle A_c(x)|M_c|B_c(x)\rangle\right) \quad (6)$$

The term $\Re\langle A_c(x)|M_c|B_c(x)\rangle$ is called 'interference term' in quantum theory, since it is responsible of the deviations from classicality in the quantum double-slit experiment. In [5, 13] we have proved that a solution is obtained in the Fock space $\mathbb{C}^3 \oplus (\mathbb{C}^3 \otimes \mathbb{C}^3)$ with the interference term given by

$$\Re\langle A_c(x)|M_c|B_c(x)\rangle = \sqrt{1 - \mu_x(A)}\sqrt{1 - \mu_x(B)}\cos\theta_c(x) \quad (7)$$

with $\theta_c(x)$ being the 'interference angle for the conjunction', and $M_c = |100\rangle\langle 100| + |010\rangle\langle 010|$, where $\{|100\rangle, |010\rangle, |001\rangle\}$ is the canonical basis of $\mathbb{C}^3$. For example, in the case of *Mint* with respect to *Food*, *Plant* and *Food and Plant*, we have $m_c^2(x) = 0.3$, $n_c^2(x) = 0.7$ and $\theta_c(x) = 50.21°$.

Let us come again to the disjunction of two concepts. The concepts A, B and 'A or B' are represented in the Fock space $\mathscr{F}$ by the unit vectors $|A_d(x)\rangle$, $|B_d(x)\rangle$ and $|(A \text{ or } B)_d(x)\rangle$, respectively, where

$$|(A \text{ or } B)_d(x)\rangle = m_d(x)e^{i\lambda_d(x)}|A_d(x)\rangle \otimes |B_d(x)\rangle + n_d(x)e^{i\nu_d(x)}\frac{1}{\sqrt{2}}(|A_d(x)\rangle + |B_d(x)\rangle) \quad (8)$$

The numbers $m_d(x)$ and $n_d(x)$ are such that $m_d(x), n_d(x) \geq 0$ and $m_d^2(x) + n_d^2(x) = 1$. The decision measurement of a subject who estimates the membership of the exemplar x with respect to the concept 'A or B' is represented by the orthogonal projection operator $M_d \oplus (M_d \otimes \mathbb{1} + \mathbb{1} \otimes M_d + M_d \otimes M_d)$ on $\mathscr{F}$, where M_d is an orthogonal projection operator on $\mathscr{H}$. Hence, the membership weight of x with respect to 'A or B' is given by

$$\begin{aligned}\mu_x(A \text{ or } B) &= \langle (A \text{ or } B)_d(x) | M_d \oplus (M_d \otimes \mathbb{1} + \mathbb{1} \otimes M_d + M_d \otimes M_d) | \\ &\quad \times (A \text{ or } B)_d(x) \rangle \\ &= m_d^2(x)(\mu_x(A) + \mu_x(B) - \mu_x(A)\mu_x(B)) \\ &\quad + n_d^2(x)\Big(\frac{\mu_x(A) + \mu_x(B)}{2} + \Re\langle A_d(x)|M_d|B_d(x)\rangle\Big) \end{aligned} \tag{9}$$

The term $\Re\langle A_d(x)|M_d|B_d(x)\rangle$ is the 'interference term for the disjunction'. In [5, 13] we have proved that a solution is obtained in the Fock space $\mathbb{C}^3 \oplus (\mathbb{C}^3 \otimes \mathbb{C}^3)$ with the interference term given by

$$\Re\langle A_d(x)|M_d|B_d(x)\rangle = \sqrt{1 - \mu_x(A)}\sqrt{1 - \mu_x(B)}\cos\theta_d(x) \tag{10}$$

with $\theta_d(x)$ being the 'interference angle for the disjunction'. Concerning the *Donkey* case, we have $m_d^2(x) = 0.26$, $n_d^2(x) = 0.74$ and $\theta_d(x) = 77.34°$.

By comparing Eqs. (6) and (9), we can see that the interference terms are generally different. Indeed, the representation of the unit vectors $|A_c(x)\rangle$, $|A_d(x)\rangle$, $|B_c(x)\rangle$ and $|B_d(x)\rangle$ generally depend on the exemplar x, on the membership weights $\mu_x(A)$ and $\mu_x(B)$, and also on whether $\mu_x(A \text{ or } B)$ or $\mu_x(A \text{ and } B)$ is measured, which results in different interferences angles $\theta_c(x)$ and $\theta_d(x)$.

3 Conclusions

The probabilistic expressions in the previous section allow the modeling of almost all of Hampton's data [1, 2], describing the deviations from classical logic and probability theory in terms of genuine quantum aspects. Moreover, one of us has recently shown [14] that our quantum approach successfully models the data collected by Alxatib and Pelletier [16] on the so-called 'borderline contradictions', and it can be further tested to model data coming from future cognitive tests. One can then inquire into the existence of underlying mechanisms determining these deviations from classicality and, conversely, the effectiveness of a quantum-theoretic modeling. Our explanations is the following.

Whenever two concepts A and B are combined in human thought to form the conjunction 'A and B', or the disjunction 'A or B', a new genuine effect comes into play, namely emergence. More specifically, if a subject is asked to estimate whether

a given exemplar x belongs to the vague concepts A, B, 'A and B' ('A or B'), two mechanisms act simultaneously and in superposition in the subject's thought. A 'quantum logical thought', which is a probabilistic version of the classical logical reasoning, where the subject considers two copies of exemplar x and estimates whether the first copy belongs to A and (or) the second copy of x belongs to B. But also a 'quantum conceptual thought' acts, where the subject estimates whether the exemplar x belongs to the newly emergent concept 'A and B' ('A or B'). The place where these superposed processes can be suitably structured is the Fock space. Sector 1 of Fock space hosts the latter process, while sector 2 hosts the former, while the weights $m_c^2(x)$ ($m_d^2(x)$) and $n_c^2(x)$ ($n_d^2(x)$) measure the amount of 'participation' of sectors 2 and 1 for the conjunction (disjunction), respectively. But, what happens in human thought during a cognitive test is a quantum superposition of both processes. The abundance of over- and under- extension effects is a significant clue that the dominant dynamics in human thought is governed by emergence, and that logical reasoning is only secondary, at variance with old established beliefs.

It is interesting to observe that similar deviations from logic and classical probability theory are observed in other areas of cognitive science, e.g., decision making ('prisoner's dilemma', 'disjunction effect', 'conjunction fallcy') and behavioral economics ('Allais, Ellsberg, Machina paradoxes'). In the above perspective, our explanation for the appearence of these phenomena is that what has been identified a fallacy, an effect or a deviation, is a consequence of the dominant dynamics in human thought which is emergent in nature, while what has been typically considered as a default to deviate from, namely logical reasoning, is a consequence of a secondary dynamics within human thought, which is quantum logical in nature.

References

1. Hampton, J. A. (1988a). Overextension of conjunctive concepts: Evidence for a unitary model for concept typicality and class inclusion. J. Exp. Psychol. Learn. Mem. Cogn. 14, 12–32.
2. Hampton, J. A. (1988b). Disjunction of natural concepts. Mem. Cogn. 16, 579–591.
3. Aerts, D., Gabora, L. (2005). A theory of concepts and their combinations I: The structure of the sets of contexts and properties. Kybernetes 34, 167–191.
4. Aerts, D., Gabora, L. (2005). A theory of concepts and their combinations II: A Hilbert space representation. Kybernetes 34, 192–221.
5. Aerts, D. (2009). Quantum structure in cognition. J. Math. Psychol. 53, 314–348.
6. Pothos, E. M., Busemeyer, J. R. (2009). A quantum probability model explanation for violations of 'rational' decision theory. Proc. Roy. Soc. B 276, 2171–2178.
7. Khrennikov, A. Y. (2010). Ubiquitous Quantum Structure. Springer, Berlin.
8. Busemeyer, J. R., Pothos, E., Franco, R., Trueblood, J. S. (2011). A quantum theoretical explanation for probability judgment 'errors'. Psychol. Rev. 118, 193–218.
9. Busemeyer, J. R., Bruza, P. D. (2012). Quantum Models of Cognition and Decision. Cambridge University Press, Cambridge.
10. Haven, E., Khrennikov, A. Y. (2013). Quantum Social Science. Cambridge University Press, Cambridge.
11. Aerts, D., Sozzo, S. (2011). Quantum structure in cognition. Why and how concepts are entangled. LNCS 7052, 116–127.

12. Aerts, D., Broekaert, J., Gabora, L., Sozzo, S. (2013). Quantum structure and human thought. Behav. Brain Sci. 36, 274–276.
13. Aerts, D., Gabora, L., and Sozzo, S. (2013). Concepts and their dynamics: A quantum-theoretic modeling of human thought. Top. Cogn. Sci. 5, 737–772.
14. Sozzo, S. (2014). A quantum probability model in Fock space for borderline contradictions. J. Math. Psychol. 58, 1–12.
15. Aerts, D. and Sozzo, S. (2014). Quantum entanglement in concept combinations. Int. J. Theor. Phys. 53, 3587–3603.
16. Alxatib, S., Pelletier, J. (2011). On the psychology of truth gaps. In Nouwen, R., van Rooij, R., Sauerland, U., & Schmitz, H.-C. (Eds.), *Vagueness in Communication* (pp. 13–36). Berlin, Heidelberg: Springer-Verlag.

Beyond the Quantum Formalism: Consequences of a Neural-Oscillator Model to Quantum Cognition

J. Acacio de Barros

Abstract In this paper we present a neural oscillator model of stimulus response theory that exhibits quantum-like behavior. We then show that without adding any additional assumptions, a quantum model constructed to fit observable pairwise correlations has no predictive power over the unknown triple moment, obtainable through the activation of multiple oscillators. We compare this with the results obtained in reference (de Barros, J.A.: Decision making for inconsistent expert judgments using negative probabilities. In: Quantum Interaction, Lecture Notes in Computer Science. Springer Verlag, Berlin (2014)), where a criteria of rationality gives optimal ranges for the triple moment.

Keywords Neural oscillator model • Stimulus response theory • Quantum cognition • Reinforcement learning • Hilbert space

1 Introduction

Recently, much attention has been paid to quantum-mechanical formalisms applied to human cognition (see [6, 8, 9], and references therein). This comes from an increasing set of empirical data better described by quantum models than classical probabilistic ones (for an new effective classical approach, however, see [7], to appear in this proceedings).

The underlying origins of such quantum-like features are not well understood, but few researchers believe that actual quantum mechanical processes are responsible (see [6] but also [10] for a different view). Instead, as argued in [4], what is behind such features is a contextual influence. Interference-like effects in neuronal firings in the brain lead to outcomes that are context dependent, similar to the two-slit experiment in quantum mechanics, thus providing a possible explanation. In fact, in [3] we showed how a simplified neural model with interference emerging from the collective dynamics of coupled neurons gives origin to quantum-like effects.

J.A. de Barros (✉)
Liberal Studies Program, San Francisco State University, San Francisco, CA, USA
e-mail: barros@sfsu.edu

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_56

Such model was designed to be consistent with currently known neurophysiology and to reproduce the behavioral stimulus-response theory [11]. Here, we discuss the implications of such neural model to quantum cognition, and in particular to the predictability power of the quantum-mechanical apparatus, as opposed to its descriptive power.

2 Model and Main Results

Here we briefly present the main model shown in [3, 11], and the readers are referred to them for details. For the simple case of a continuum of responses, we start with representations of stimulus and response in terms of phase oscillators. Such oscillators, made out of collections of neurons, are synaptically coupled, and, depending on the coupling strength, may synchronize. Let $s(t)$ be the neural oscillator representing the activation of a stimulus, and $r_1(t)$ and $r_2(t)$ the oscillators for the two extremes in a continuum or responses. We focus on their phases, φ_s, φ_{r_1}, and φ_{r_1}, whose dynamics are given by

$$\dot{\varphi}_i = \omega_i + \sum_{j \neq i} k^E_{i,j} \sin(\varphi_i - \varphi_j) + \sum_{j \neq i} k^I_{i,j} \cos(\varphi_i - \varphi_j), \tag{1}$$

where $k^E_{i,j}$ and $k^I_{i,j}$ are the overall excitatory and inhibitory couplings between the neural oscillators. During reinforcement, the coupling strengths $k^E_{i,j}$ and $k^I_{i,j}$ are changed in a Hebb-like fashion. This model can easily be extended to include multiple stimulus and response oscillators. For instance, in [3] it was used with two stimulus oscillators to obtain quantum-like effects. Such effects were the consequence of couplings between the oscillators that were reinforced to respond to two different stimuli corresponding to incompatible contexts. When both stimuli were simultaneously activated, an interference effect was obtained.

Quantum-like models lead to contextual responses, in the sense that there exists no joint probability distribution for the associated random variables. Let us look at the particular example presented in reference [2] and expanded in another context in [5]. Let $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$ be ± 1-valued random variables, and consider the neural oscillator system represented in Fig. 1. For this system, the activation of one of the three stimulus oscillators, C_1, C_2, or C_3, leads to the corresponding responses computed via phase differences. For example, if C_1 is sampled, the oscillators' dynamics, dictated by the specific values of inhibitory and excitatory couplings, converge to a fixed point that may favor $\mathbf{X} = 1$ (oscillator X) instead of $\mathbf{X} = -1$ (oscillator $\sim X$), while at the same time favoring $\mathbf{Y} = -1$, thus corresponding to a negative correlation. With such oscillator system, it is possible in principle to choose couplings such that the correlations between $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$ are too strong for a joint probability distribution to exist. As a consequence, and because of the pairwise commutativity of the set of quantum-mechanical observables $\hat{X}$, $\hat{Y}$, and

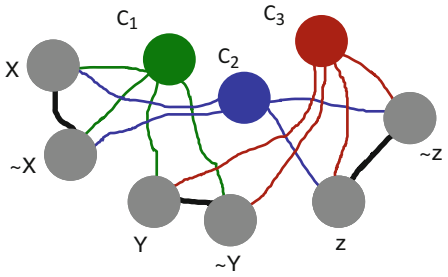


Fig. 1 Layout of a neural-oscillator system exhibiting pairwise correlations between $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$. In this oscillator system, $(\mathbf{X} = 1)$ and $(\mathbf{Y} = -1)$ corresponds to the synchronization with oscillator C_1 closer in phase to X and not to $\sim X$, while at the same time being closer to Y than to $\sim Y$

$\hat{Z}$ corresponding to the random variables $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$, it follows that there exists no state $|\psi\rangle$ in the Hilbert space $\mathcal{H}$ where such observables are defined and such that the neural correlations hold [5]. However, even in such situations a quantum model can be constructed [1], and in order to describe the correlations set by the neural-oscillator model, we are forced to expand the Hilbert space to $\mathcal{H}' \otimes \mathcal{H}$ [5]. For instance, we can write a state vector

$$|\psi\rangle = c_{xy}|A\rangle|\psi_{xy}\rangle + c_{xz}|B\rangle|\psi_{xz}\rangle + c_{yz}|C\rangle|\psi_{yz}\rangle, \tag{2}$$

where $|A\rangle$, $|B\rangle$, and $|C\rangle$ are orthonormal vectors in $\mathcal{H}'$, $\langle\psi_{xy}|\hat{X}\hat{Y}|\psi_{xy}\rangle = -2/3$, $\langle\psi_{xz}|\hat{X}\hat{Z}|\psi_{xz}\rangle = -1/2$, $\langle\psi_{yz}|\hat{Y}\hat{Z}|\psi_{yz}\rangle = 0$, and c_{xy}, c_{xz}, and c_{yz} are such that $|c_{xy}|^2 + |c_{xz}|^2 + |c_{yz}|^2 = 1$. Because each of the states $|\psi_{xy}\rangle$, $|\psi_{xz}\rangle$, and $|\psi_{yz}\rangle$ can have arbitrary triple moments (they do not fix enough of the distribution) between -1 and 1, it follows that (2) can describe the correlations but has no predictive power with respect to the neural oscillator model or human decision making.

However, the couplings encoding different responses in the Kuramoto equations do determine, within a certain range, values for the triple moment. The triple moment would be the equivalent, following [3], of a simultaneous activation of all stimulus oscillators. Thus, the neural model would provide a definite prediction, in contrast with the quantum one.

3 Final Remarks

Quantum formalisms applied to human cognition have shown a great potential for certain applications in the social sciences. However, one must ask how this is so, and also how predictive they are. For instance, as showed above, it is possible to devise an neural system whose quantum description has no predictive power. Thus, we could in principle design an experiment to test this neural system, but not its corresponding quantum description.

Could there be some principle to be added to the quantum description that could provide predictions for outcomes of the experiment proposed? For example, in [5] we proposed a minimization principle as a normative decision for quantum-like inconsistencies, which allowed signed probabilities to move from a descriptive to a normative theory. Perhaps a principle of this type added to the quantum formalism could be not only normative but predictive as well. However, if we think that the underlying dynamics for quantum cognition is actually from the complex and contextual interaction of neurons, perhaps some similar principle from it should be added to the quantum description.

Finally, we would like to emphasize that the quantum approach has suggested interesting experiments in psychology. As such, it is a promising field not only because of its ability to describe experiments, but also for the intuitions it provides for thinking about context-rich situations. Therefore, understanding its limitations and perhaps extending it would be desirable.

Acknowledgements We thank Sandro Sozzo for pointing out reference [1], and him as well as Gary Oas, Ehtibar Dzhafarov, Paavo Pilkkanen, and Sisir Roy for useful comments and discussions.

References

1. Aerts, D., Sozzo, S.: Quantum entanglement in concept combinations. arXiv e-print 1302.3831 (2013). URL http://arxiv.org/abs/1302.3831
2. de Barros, J.A.: Joint probabilities and quantum cognition. In: A. Khrennikov, A.L. Migdall, S. Polyakov, H. Atmanspacher (eds.) AIP Conference Proceedings, vol. 1508, pp. 98–107. American Institute of Physics, Växjö, Sweden (2012). DOI doi:10.1063/1.4773121. URL http://proceedings.aip.org/resource/2/apcpcs/1508/1/98_1
3. de Barros, J.A.: Quantum-like model of behavioral response computation using neural oscillators. Biosystems **110**(3), 171–182 (2012). DOI 10.1016/j.biosystems.2012.10.002. URL http://www.sciencedirect.com/science/article/pii/S0303264712001736
4. de Barros, J.A., Suppes, P.: Quantum mechanics, interference, and the brain. Journal of Mathematical Psychology **53**(5), 306–313 (2009)
5. de Barros, J.A.: Decision making for inconsistent expert judgments using negative probabilities. In: Quantum Interaction, Lecture Notes in Computer Science. Springer Verlag, Berlin (2014).
6. Busemeyer, J.R., Bruza, P.D.: Quantum models of cognition and decision. Cambridge University Press, Cambridge, Great Britain (2012)
7. Dzhafarov, E.N., Kujala, J.V.: Random variables recorded under mutually exclusive conditions: contextuality-by-default. In: H. Liljenström (ed.) Advances in Cognitive Neurodynamics (IV): Proceedings of the Fourth International Conference on Cognitive Neurodynamics – 2013. Advances in Cognitive Neurodynamics. Springer, Dordrecht (2015)
8. Haven, E., Khrennikov, A.: Quantum Social Science. Cambridge University Press, Cambridge, UK (2013)
9. Khrennikov, A.: Ubiquitous Quantum Structure. Springer Verlag, Heidelberg (2010)
10. Pylkkänen, P.: Weak vs. strong quantum cognition. In: Proceedings of the 4th International Conference on Cognitive Neurodynamics, ICCN2013 (2014). Forthcoming.
11. Suppes, P., de Barros, J.A., Oas, G.: Phase-oscillator computations as neural models of stimulus–response conditioning and response selection. Journal of Mathematical Psychology **56**(2), 95–117 (2012)

Random Variables Recorded Under Mutually Exclusive Conditions: Contextuality-by-Default

Ehtibar N. Dzhafarov and Janne V. Kujala

Abstract We present general principles underlying analysis of the dependence of random variables (outputs) on deterministic conditions (inputs). Random outputs recorded under mutually exclusive input values are labeled by these values and considered stochastically unrelated, possessing no joint distribution. An input that does not directly influence an output creates a context for the latter. Any constraint imposed on the dependence of random outputs on inputs can be characterized by considering all possible couplings (joint distributions) imposed on stochastically unrelated outputs. The target application of these principles is a quantum mechanical system of entangled particles, with directions of spin measurements chosen for each particle being inputs and the spins recorded outputs. The sphere of applicability, however, spans systems across physical, biological, and behavioral sciences.

Keywords Contextuality • Couplings • Joint distribution • Random outputs

1 Introduction

This paper pertains to any system, physical, biological, or behavioral, with *random outputs* recorded under varying conditions (*inputs*). A target example for us is a quantum mechanical system of two entangled particles, "Alice's" and "Bob's." Alice measures the spin of her particle in one of two directions, α_1 or α_2, and Bob measures the spin of his particle in one of two directions, β_1 or β_2. Here, α and β are inputs, and each trial is characterized by one of four possible *input values* (α_i, β_j). The spins recorded in each trial are realizations of random variables A and B, which, in the simplest case, can attain two values each: a_1 or a_2 for A and b_1 or b_2 for B. One can think of many examples in other domains with similar formal structure, e.g., a psychophysical experiment with an observer responding

E.N. Dzhafarov (✉)
Purdue University, West Lafayette, IN, USA
e-mail: ehtibar@purdue.edu

J.V. Kujala
University of Jyväskylä, Jyväskylä, Finland
e-mail: jvk@iki.fi

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_57

to stimuli with varying characteristics α (say, intensity) and β (say, shape). These characteristics then constitute inputs, while some characteristics of the responses, such as response time A (with a continuum of values) and response correctness B (with two possible values), are random outputs.

Accounts of the approach presented in this paper can be found in [5–7], but this paper is the first one focusing entirely on its basic principles. The approach amounts to philosophical rethinking (or at least conceptual tweaking) of the foundations of probability, specifically, of random variables and their joint distributions. Here, it is presented without technical details (that can be reconstructed from [3–6]).

2 Basic Principles

Let all or some of the random outputs of a system form a random variable X,[1] and the totality of all inputs be a variable χ. In our target example, $\chi = (\alpha, \beta)$ with input values $\chi_1 = (\alpha_1, \beta_1), \ldots, \chi_4 = (\alpha_2, \beta_2)$, whereas X can be (A, B) with values $x_1 = (a_1, b_1), \ldots, x_4 = (a_2, b_2)$, or A with values a_1, a_2, or B with values b_1, b_2. If χ itself is a random variable, so that $\chi_1, \chi_2, \ldots$ occur with some probabilities, we ignore these probabilities and simply condition the recorded outputs X on values of χ. In other words, we have a distribution of X given that $\chi = \chi_1$, a distribution of X given that $\chi = \chi_2$, etc., irrespective of whether we can control and predict the values of χ, or they occur randomly. Now, this conditioning upon input values means that *X is indexed by different values of χ*. We obtain thus, "automatically," a set of different random variables in place of what we previously called a random variable X. We have X_{χ_1} (or X_1, if no confusion is likely) which is X when $\chi = \chi_1$, X_{χ_2} (or X_2) which is X when $\chi = \chi_2$, etc. Let us formulate this simple observation as a formal principle.

Principle 1 *Outputs recorded under different (hence mutually exclusive) input values are labeled by these input values and considered different random variables. These random variables are stochastically unrelated, i.e., they possess no joint distribution.*

Thus, in our target example, we have four random variables A_{ij}, four random variables B_{ij}, and four random variables $(A, B)_{ij} = \left(A_{ij}, B_{ij}\right)$ corresponding to the four input values $\chi_k = \left(\alpha_i, \beta_j\right)$. The principle holds irrespective of how the distribution of X_k depends on χ_k. Thus, the variables A_{i1} and A_{i2} remain different even if their distributions are identical (as they should be if Bob's choice cannot influence Alice's measurements). One must not assume that they are one and the same random variable, $A_i = A_{i1} = A_{i2}$. The latter would mean that A_{i1} and A_{i2} have a joint distribution, because of which the probabilities $\Pr[A_{i1} = A_{i2}]$ are well

[1] Random variables are understood in the broadest sense, so that a vector of random variables (or any set thereof, or a random process) is a random variable too.

defined, and that these probabilities equal 1. But A_{i1} and A_{i2} do not have a joint distribution. Indeed, two random variables X and Y have a joint distribution only if their values can be thought of as observed "in pairs," i.e., if there is a scheme of establishing correspondence $x_{(i)} \leftrightarrow y_{(i)}$ between observations $x_{(1)}, x_{(2)}, \ldots$ of X and $y_{(1)}, y_{(2)}, \ldots$ of Y. In our example, the correspondence is defined by the two measurements being simultaneously performed on a given pair of entangled particles. Each such a pair of measurements corresponds to a certain input value, e.g., A_{21} and B_{21} correspond to $\chi = (\alpha_2, \beta_1)$. Therefore, no measurement outputs corresponding to different input values, such as A_{i1} and A_{i2}, or A_{i1} and B_{i2} co-occur in the same sense in which, say, A_{i1} co-occurs with B_{i1}.

However, given any two random variables X and Y, one can impose on them a joint distribution, and create thereby a random variable $Z = (X, Y)$, referred to as a *coupling* for X and Y. By definition, the distribution of a coupling Z agrees with the distributions of X and Y as its marginals.

Principle 2 *Stochastically unrelated outputs recorded under mutually exclusive input values can be coupled (imposed a joint distribution upon) arbitrarily. There are no privileged couplings.*

Thus, in our target example, the famous Bell-type theorems [1, 3, 8] implicitly impose on (A_{11}, B_{11}), …, (A_{22}, B_{22}) a coupling with $A_{i1} = A_{i2}$ and $B_{1j} = B_{2j}$. This amounts to considering a random variable $\left(A'_1, A'_2, B'_1, B'_2\right)$ such that $\left(A'_i, B'_j\right)$ is distributed as $\left(A_{ij}, B_{ij}\right)$. The Bell-type theorems show that such a coupling exists if and only if the distributions of the coupled pairs (A_{11}, B_{11}), …, (A_{22}, B_{22}) satisfy certain constraints (Bell-type inequalities, known to be violated in quantum mechanics). In our approach, however, except possibly for simplicity considerations, this coupling has no privileged status among all possible coupling for (A_{11}, B_{11}), …, (A_{22}, B_{22}). Thus, any distribution of spins satisfying Bell-type inequalities is also compatible with the coupling in which (A_{11}, B_{11}), …, (A_{22}, B_{22}) are stochastically independent pairs of random variables, as well as with an infinity of other couplings in which $\Pr[A_{i1} = A_{i2}]$ and $\Pr\left[B_{1j} = B_{2j}\right]$ may be different from 1.

If the distributions of A_{i1} and A_{i2} are not the same for $i = 1$ or $i = 2$, the situation is simple: the output A is influenced by both inputs α and β (and analogously for B_{1j} and B_{2j}). If, however, the distributions of A_{i1} and A_{i2} are always the same, and if, moreover, substantive considerations (e.g., laws of special relativity) prevent the possibility of interpreting β as "directly" influencing A, then we can say that β forms a *context* for the dependence of A on α (and analogously for α creating a context for the dependence of B on β). Principle 1 ensures that this contextuality is introduced "automatically," by labeling all outputs by all conditions under which they are recorded. The degree and form of contextuality in a given system (e.g., those with constraints more relaxed than the Bell-type inequalities [2, 9]) can be characterized by considering all possible probabilities $\Pr[A_{i1} = A_{i2}]$ and $\Pr\left[B_{1j} = B_{2j}\right]$, called *connection probabilities* in [5–7]. This approach allows

one to embark on a deeper investigation of the relationship between the classical probability theory and quantum mechanics than in the Bell-type theorems.

3 Apparent Problems with the Approach

Two objections can be raised against our approach. One is that it requires to label random variables by circumstances that cannot possibly be relevant. If reaction time X to a given stimulus is recorded in conjunction with measurements of the temperature on Mars with the values $\chi_1 = low$ and $\chi_2 = high$, would it be meaningful to "automatically" split X into stochastically unrelated X_{low} and X_{high}? The answer is: it is meaningful. If the temperature on Mars affects the distribution of X, then considering X_{low} and X_{high} as different random variables is clearly useful for understanding of X. If, as we suspect, the temperature on Mars does not affect the distribution of X, then one can impose on $\left(X_{low}, X_{high}\right)$ an arbitrary coupling, including one with $X_{low} = X_{high} = X$. The latter choice amounts to ignoring the temperature on Mars altogether.

The other objection is that if we apply Principle 1 systematically, we have to consider different realizations of a random variable X as stochastically unrelated random variables. X occurring in trial 1 as $x_{(1)}$ is labeled X_1 and considered stochastically unrelated to X_2 that occurs in trial 2 as $x_{(2)}$, and so on. But this is perfectly reasonable, and moreover, it is a standard issue in the probabilistic theory of couplings [10]. Once a coupling (e.g., the commonly used iid one) is imposed on $X_1, X_2, \ldots$, it creates a new random variable $Y = (X_1, X_2, \ldots)$, of which we have a single realization $y = \left(x_{(1)}, x_{(2)}, \ldots\right)$. One can then investigate whether this y is statistically plausible in view of the distribution of Y using standard statistical reasoning.

References

1. Bell, J.: On the Einstein-Podolsky-Rosen paradox. Physics 1, 195–200 (1964).
2. Cirel'son, B.S.: Quantum generalizations of Bell's inequality. Letters in Mathematical Physics 4, 93–100 (1980)
3. Dzhafarov, E.N. & Kujala, J.V.: Selectivity in probabilistic causality: Where psychology runs into quantum physics. *Journal of Mathematical Psychology*, 56, 54–63 (2012). (available as *arXiv*: 1110.2388.)
4. Dzhafarov, E.N., & Kujala, J.V.: Quantum entanglement and the issue of selective influences in psychology: An overview. *Lecture Notes in Computer Science*, 7620, 184–195 (2012). (available as *arXiv*:1209.0041.)
5. Dzhafarov, E.N., & Kujala, J.V.: All-possible-couplings approach to measuring probabilistic context. *PLoS ONE* 8(5): e61712. doi:10.1371/journal.pone.0061712 (2013). (available as arXiv:1209.3430.)

6. Dzhafarov, E.N., & Kujala, J.V.: A qualified Kolmogorovian account of probabilistic contextuality. *Lecture Notes in Computer Science*, 8369, 201–212 (2014). (available as arXiv:1304.4546.)
7. Dzhafarov, E.N., & Kujala, J.V.: No-forcing and no-matching theorems for classical probability applied to quantum mechanics. *Foundations of Physics*, 44, 248–265 (2014). (available as *arXiv*:1305.3649.)
8. Fine, A.: Hidden variables, joint probability, and the Bell inequalities. Physical Review Letters 48, 291–295 (1982).
9. Landau, L. J.: On the violation of Bell's inequality in quantum theory. Physical Letters A 120, 54–56 (1987).
10. Thorisson, H.: Coupling, Stationarity, and Regeneration. New York: Springer (2000).

Weak vs. Strong Quantum Cognition

Paavo Pylkkänen

Abstract In recent decades some cognitive scientists have adopted a program of *quantum cognition*. For example, Pothos and Busemeyer (PB) argue that there are empirical results concerning human decision-making and judgment that can be elegantly accounted for by quantum probability (QP) theory, while classical (Bayesian) probability theory fails. They suggest that the reason why QP works better is because some cognitive phenomena are *analogous* to quantum phenomena. This naturally gives rise to a further question about *why* they are analogous. Is this a pure coincidence, or is there a deeper reason? For example, could the neural processes underlying cognition involve subtle quantum effects, thus *explaining* why cognition obeys QP? PB are agnostic about this controversial issue, and thus their kind of program could be labeled as "weak quantum cognition" (analogously to the program of weak artificial intelligence as characterized by Searle). However, there is a long tradition of speculating about the role of subtle quantum effects in the neural correlates of cognition, constituting a program of "strong quantum cognition" (SQC) or "quantum cognitive neuroscience". This paper considers the prospects of SQC, by briefly reviewing and commenting on some of the key proposals. In particular, Bohm and Hiley's active information program will be discussed.

Keywords Quantum cognition • Quantum probability • Analogy • Active information • Implicate order • Mental causation • Representational content • David Bohm • Basil Hiley

P. Pylkkänen (✉)
Department of Cognitive Neuroscience and Philosophy, University of Skövde, P.O. Box 408, SE-541 28 Skövde, Sweden
e-mail: paavo.pylkkanen@his.se

The Finnish Center of Excellence in the Philosophy of the Social Sciences (TINT), Department of Philosophy, History, Culture and Art Studies, University of Helsinki, P.O. Box 24, FI-00014 Helsinki, Finland
e-mail: paavo.pylkkanen@his.se

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_58

1 Introduction

In their recent article "Can quantum probability provide a new direction for cognitive modeling?" Pothos and Busemeyer (PB) (2013) make a convincing case that there are empirical results concerning human decision making and judgment that can be elegantly accounted for by quantum probability (QP) theory, while classical (Bayesian) probability theory fails [15]. In particular, they point out that human judgment and preference often display order and context effects, violations of the law of total probability and failures of compositionality, and that in such cases QP – with features such as superposition and entanglement – provides a natural explanation of cognitive process. More generally, they suggest that QP is potentially relevant in any behavioral situation that involves uncertainty.

Such success in modeling raises the question of how can it be that QP which was developed to account for quantum physical phenomena could possibly be able to account for cognitive phenomena. PB do not discuss this issue at great length, but suggest that the reason is because some cognitive phenomena are *analogous* to quantum phenomena. But this gives rise to a further question: *why* are these phenomena analogous to each other? Is it a mere coincidence or is there some deeper explanation? For example, might the neural processes underlying cognition be quantum-like in some way? PB remain agnostic about this controversial issue, and thus we might call their program an instance of "weak quantum cognition" (somewhat analogously to the program of weak AI in artificial intelligence research; cf. also the program of "weak quantum theory", where one applies some, but not all formal features of quantum theory to explain cognitive phenomena, see Atmaspacher et al. 2002 [3]). However, there is a long tradition of speculating about the role of subtle quantum effects in the neural correlates of cognition, constituting a program of "strong quantum cognition" or "quantum cognitive neuroscience". While it may be a good research strategy in cognitive science to pursue weak quantum cognition without worrying about the underlying reasons for why QP works for cognition, it would clearly be a major scientific breakthrough if strong quantum cognition would turn out to be correct. It is thus worth giving attention to the current state-of-the-art in strong quantum cognition. The aim of this paper is to briefly review and comment some major developments. In particular, I will consider the prospects of Bohm and Hiley's research program [8, 18].

2 Strong Quantum Cognition: Subtle Quantum Effects in the Neural Correlates of Cognition?

There are various ways in which the neural processes underlying cognition could be quantum-like. The strongest possibility is that they literally involve subtle quantum effects. For example, following Niels Bohr, David Bohm speculated about this possibility already in 1951 in his textbook *Quantum theory* [6]. Anticipating

the current research on quantum cognition [1, 15], he drew attention to what he considered to be *remarkable point-by-point analogies* between quantum processes and thought. He added that it would provide a natural explanation of these analogies if it turned out that some key neural processes (e.g. in synapses) were subject to quantum-theoretical limitations (for a discussion of Bohm's analogies see Pylkkänen 2014 [17]).

Harald Atmanspacher (2011) has provided a useful overview of various programs of what I have above call "strong quantum cognition" [2]. First of all, there are approaches that stay within the usual interpretation of the quantum theory. There is the von Neumann-Wigner line of thought that assumes that consciousness plays a role in quantum state reductions; in Stapp's later development of this approach the neural correlates of conscious intentional acts are assumed to involve quantum state reductions. There is the Ricciardi-Umezawa-Vitiello approach that sees mental states, particularly memory states, as vacuum states of quantum fields (this approach has been given an imaginative philosophical interpretation by Globus 2003 [10]). Finally, there is the Beck-Eccles approach, where it is assumed that due to quantum mechanical processes the frequency of exocytosis at a synaptic cleft can be controlled by mental intentions, without violating the conservation of energy (for a discussion of this last approach see also Hiley and Pylkkänen 2005 [13]).

Atmanspacher also draws attention to programs of strong quantum cognition that involve further extensions or generalizations of present-day quantum theory. Most notably, there is Penrose's proposal that human (say mathematical) insight is non-computable and that the physiological correlates of such insight thus need to involve non-computable physical processes. He thinks that such process might well be related to quantum state reduction. However, Penrose is not satisfied with quantum state reduction as this is characterized in the usual interpretation of quantum theory. Instead, he proposes that gravity brings about the reduction under certain circumstances, which allows the possibility of an orchestrated objective reduction (Orch-OR) – the idea being that the reduction can take place without the activity of a human conscious observer, and in an orchestrated way. This involves an extension of current quantum theory, in which latter the state reductions obey the usual laws of quantum probability. Together with Hameroff, Penrose proposed that neural microtubules might provide a site where Orch-ORs could take place. Their assumption is that Orch-ORs in neural microtubules, when suitably integrated, constitute conscious moments. (So it is not that consciousness collapses the wave function but rather that the collapses constitute consciousness.) The idea is similar to Stapp's later ideas, but one difference is that while Stapp stays within the usual interpretation of quantum theory, Penrose's approach involves going beyond it (in that the reductions can be objective and orchestrated, and need not obey the usual quantum laws). Hameroff and Penrose (2014) have recently published an extensive review of their approach, with new features and, in the same journal, a reply to various criticisms [12].

Those who advocate strong quantum cognition typically encounter the criticism that quantum effects are washed out in the "warm, wet and noisy" conditions of the macroscopic world and brains in particular (the "decoherence problem"). It is

thus concluded that quantum theory is only relevant to physical processes in the (sub)atomic domain and should be ignored in other physical domains. However, there are many recent research developments suggesting that biological organisms at ordinary temperatures exploit subtle quantum effects, and biological evolution would thus have been able to solve the decoherence problem at least in some biological contexts (e.g. the studies on energy-harvesting in photosynthesis and avian magnetoreception; for a short review, see Ball, P. (2011) [4]). As Atmanspacher points out, it is however still a controversial issue whether subtle quantum processes play a significant role in the neural correlates of cognition and consciousness.

Note that those researchers who accept that cognition is quantum-like and seek to explain this in neural terms need not necessarily adopt the program of strong quantum cognition. For there is also the possibility that the neural processes underlying cognition involve no subtle quantum effects, but can nevertheless give rise to quantum-like neural activity. Something like this is implied by Barros and Suppes (2009) when they suggest that *classical* interference in the brain may lead to contextual processes [5]. They refer to experimental work according to which cortical oscillations may propagate in the cortex as if they were waves; and to simulations of the mammalian brain which show the presence of interference in the cortex.

3 Bohm's Active Information Program of Quantum Cognition

We already mentioned briefly above that the physicist David Bohm speculated early about strong quantum cognition in his 1951 textbook [6]. At that time he was thinking within the usual interpretation of quantum theory, and the analogies he drew attention to then reflect this. As is well known, Bohm's key long-term aim was to understand quantum theory better, and this led him to develop a number of different alternative schemes. Given his early intuition that quantum processes and thought are analogous, it is not surprising that he applied the new ideas arising from his various quantum schemes to describing the mind.

In 1952 Bohm published two articles in *Physical Review* where he proposed (similarly to deBroglie's earlier ideas) that an electron is a particle guided by a field. For example, according to this hypothesis, in the famous 2-slit experiment the particle goes through one of the slits, while the field goes through both slits, and guides the particle. (For a discussion of this approach, as well as its various criticisms, see Goldstein 2009 [11].)

In later work with Basil Hiley, Bohm emphasized that this field does not push and pull the particle mechanically but rather in-forms its energy (Bohm and Hiley 1993 [8]; see also Smith 2003 [21]; Pylkkänen 2007 [18]). It is the form of the field (rather than its intensity) that determines the quantum potential acting upon the particle. The form of the field, in turn, reflects the form of the entire experimental arrangement – and thus the behavior of the particle depends on the whole situation. Thus, Bohm and Hiley argued, the key new ontological feature of quantum theory

is the existence of objective and active information at the quantum level (see also Maroney 2002) [14]. Bohm (1990) extended this model to include higher levels of information, so that cognitive informational processes could be connected to quantum information, which in turn could control neurophysiological processes in, say, synapses [7]. This line of research has been developed by e.g. Hiley and Pylkkänen 2005 [13]; for an interesting recent related and critical discussion, see Seager 2013 [19]. They propose that the approach enables new ways of understanding such key philosophical problems as mental causation, intentionality and even consciousness.

In Bohm's active information scheme it is thus assumed that quantum theory needs to be extended into a hierarchy of levels of active information, where the human mind, for example, involves not only the lower levels, but also the more subtle levels (Bohm claimed that such an extension of quantum theory is not arbitrary, but "natural" from the physical and mathematical point of view). Some of these levels are at the manifest, classical level while others are more "subtle", quantum-like. In perception, information encoded in manifest levels (e.g. in the form of printed words) is carried toward the more subtle levels in the nervous system, where the meaning of the information is apprehended. Such apprehension of meaning is an activity, which crucially involves the organization of the lower levels of information.

Consider, for example, a case where someone encounters shadows in a dark night, while knowing there might be a dangerous assailant around. If a shadow is interpreted as "the assailant" this means "danger" and typically results in a powerful psycho-somatic (Bohm would say signa-somatic) response. Meaning, typically assumed to be a "mental" quality, organizes the physiological state of the person. With this example Bohm wants to illustrate that there is no strict ultimate division between mind and matter, as information acts as a bridge at all levels.

There is here a two-way traffic between manifest and subtle levels (perception carries information from the manifest toward subtle levels, while the perception of meaning at subtle levels results in a physiological response). It is in this way that we can understand how mind (understood as involving very subtle physical levels) can influence the more manifest aspects of the physical domain (e.g. bodily movements).

One key idea thus is that the more subtle, physical ("mental") levels are influenced by and can also influence the lower, manifest levels. Mind is not floating free from the quantum and classical levels, but can influence these latter, thus providing a new way of understanding mental causation. The level of quantum information is especially important in providing the missing link between the traditional categories of physical and mental. Of course, it is a major unsolved problem in the Bohm scheme what exactly is meant by the "subtle levels". It seems that Bohm meant both complex neurophysiological processes (already described in cognitive neuroscience), and some subtle quantum and "super-quantum" effects taking place in the brain but not yet discovered. Thus this is currently a heavily speculative scheme that needs much further critical examination and development. Yet, in my view, it has some advantages over the other schemes of strong quantum cognition (see Hiley and Pylkkänen 2005) [13]. For one thing, it can be argued that

the Bohm-Hiley ontological interpretation provides currently the best ontological scheme for quantum theory, and we need a clear quantum ontology to tackle in a quantum-theoretical way the mind-matter problem, often also characterized as the ontological problem (e.g. Churchland 2013) [9].

The possibility that something like information plays such a fundamental role in physical processes might throw new light upon some of the other perennial problems in philosophy of mind and language. For example, in current consciousness studies it is fairly popular to think consciousness in terms of representational content. To borrow an example from Fred Dretske, suppose that you are dreaming about a blue dog. Many would say that there is nothing blue nor dog-like in your brain when you are so dreaming. But what, then, is the blue dog you apparently see? One suggestion is that it is the *representational content* carried by your neural processes. Your neural processes consist in various physiological activities, which constitute the *vehicle* for your representational content.

This type of approach may sound reasonable, but we have (in my view) currently no satisfactory theory about how representational content arises from, say, neurophysiological vehicles. If I build (with currently available technology) a robot that can receive visual (and other) information, store, process and make use of it, most of us would assume that the robot experiences no representational content. For example, it can react to wavelengths in an appropriate way, but presumably it experiences no color. Now, the information in such a robot is typically stored in the "classical" level (from the physics point of view). One speculative possibility opened up by Bohm's approach is that the kind of information we meet at the quantum level can (at least in some situations) support or give rise to genuine representational content. In the case of simple systems (e.g. electrons) this is presumably very rudimentary "proto-content". But in more complex settings the quantum-like information might have the sort of representational content we encounter in experience. Thus the reason why current robots lack genuine representational content would be because they lack the quantum-like Bohmian active information humans (and other suitably organized biological organisms) have (cf. Pylkkänen 1995 [16], 2007 [18]). Such talk of "proto-content" connects with the long tradition of panpsychism in Western history of philosophy. Panpsychism has in recent years made a comeback in analytical philosophy and consciousness studies (see Seager and Allen-Hermanson 2013, and the references therein) [20].

Of course, one challenge is then to show how such Bohmian quantum information (and its higher-level analogues) would survive and operate in neurophysiological processes. The work of Hameroff and Penrose (2014) contains many suggestions about how the quantum wave function (and thus, in Bohmian terms, the quantum active information) could survive in the warm, wet and noisy conditions of the brain, albeit the issue is controversial [12]. Also, the application of Bohmian quantum field theory (for the electromagnetic field) to dendritic fields might be an especially fruitful area where to develop a more concrete model of quantum active information in neural processes (Hiley, private communication).

The above is admittedly very speculative, but it is important to realize how potentially radical Bohm's suggestion about quantum theoretical active information is.

Insofar as such information has mental (or proto-mental) aspects, Bohm's suggestion challenges many fundamental assumptions about the nature of the physical in contemporary philosophy of mind.

4 A More General Scheme: The Implicate Order

In the early 1960s Bohm began to seek a more general scheme in which one could bring quantum theory and relativity together. This framework became known as the implicate order. The implicate order refers to holistic phenomena, where, for example, information about the whole is enfolded in each region (as in the movement of lights waves, which can be recorded in a hologram). Applied to the universe, this suggests that the universe is a movement in which a holistic order, the implicate order prevails – thus the universe is *holomovement*. At each moment a three-dimensional explicate order unfolds from the holomovement, only to enfold back in the next moment. This process of unfoldment and enfoldment takes place so rapidly that we do not see it but instead typically perceive an enduring three-dimensional reality of macroscopic objects.

Bohm proposed that the implicate order is fundamental and general, and also prevails in biological and psychological phenomena. For example, the conscious experience of listening to music can be understood in terms of the implicate order. A symphony involves a movement in which a total order builds up and grows. At each moment we are most explicitly aware of certain tones, while the previously explicate tones are experienced as enfolded, actively transforming structures; our experience also involves an anticipation the future tones. Bohm thus provided a new way of thinking and modeling a central issue in phenomenology, namely time consciousness. This has been discussed in some detail by Pylkkänen (2007, ch 5) [18]. Such an application of the implicate order to describe phenomenal experience can be seen as an instance of "weak quantum cognition", as one is using a theoretical scheme inspired by quantum theory while not claiming that phenomenal experience is literally a quantum phenomenon.

5 Concluding Remarks

We have above briefly reviewed some programs of quantum cognition, focusing upon Bohm's active information program. Given the key unsolved problems in philosophy of mind (mental causation, intentionality, consciousness), it seems reasonable to explore quantum theoretical models that radically change our common notion of the physical. The possibility that something like active information might play a key role at the quantum level is particularly important, as it may help to understand issues such as mental causation and even the origin of genuine representational content.

References

1. Aerts, D. (2009) Quantum structure in cognition, *Journal of Mathematical Psychology* 53 (2009) 314–348.
2. Atmanspacher, H. (2011) Quantum Approaches to Consciousness, *The Stanford Encyclopedia of Philosophy*, E. N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2011/entries/qt-consciousness/>
3. Atmanspacher, H., Römer, H., and Walach, H. (2002). "Weak quantum theory: Complementarity and entanglement in physics and beyond". *Foundations of Physics* 32, 379–406.
4. Ball, P. (2011) The dawn of quantum biology, *Nature* 474, 272–274.
5. Barros, J.A. & Suppes, P. (2009) Quantum mechanics, interference and the brain, *Journal of Mathematical Psychology* 53, 306–313.
6. Bohm, D. (1951) *Quantum theory*. New York: Prentice Hall. Republished by Dover, 1989
7. Bohm, D. (1990): A new theory of the relationship of mind and matter, *Philosophical Psychology* 3, pp. 271–286.
8. Bohm, D. & Hiley, B.J. (1993): *The Undivided Universe. An Ontological Interpretation of Quantum Theory.* Routledge, London.
9. Churchland, P. (2013) *Matter and Consciousness. 3rd edition.* Cambridge, Mass.: MIT Press.
10. Globus, G. (2003) *Quantum Closures and Disclosures: Thinking-together postphenomenology and quantum brain dynamics*. Amsterdam: John Benjamins.
11. Goldstein, S. (2009): Bohmian Mechanics, The Stanford Encyclopedia of Philosophy (Spring 2009 Edition), E. N. Zalta Ed., URL = <http://plato.stanford.edu/archives/spr2009/entries/qm-bohm/>.
12. Hameroff, S. & Penrose, R. (2014) Consciousness in the universe: A review of the Orch OR theory. *Physics of life reviews,* forthcoming.
13. Hiley, B.J. & Pylkkänen, P. (2005): Can mind affect matter via active information?, Mind and Matter 3 (*2*), 7–26. URL = <http://www.mindmatter.de/resources/pdf/hileywww.pdf>
14. Maroney, O. 2002. *Information and Entropy in Quantum Theory*. PhD Thesis, Department of Physics, Birkbeck College, University of London. URL = <http://www.bbk.ac.uk/tpru/OwenMaroney/thesis/thesis.html>
15. Pothos, M & Busemeyer J. R. (2013) Can quantum probability provide a new direction for cognitive modeling? *Behavioral and Brain Sciences* 36, 255–327. doi:10.1017/S0140525X12001525
16. Pylkkänen, P. (1995) On Baking a Conscious Cake with Quantum Yeast and Flour,in B. Borstner and J. Shawe-Taylor (eds). *Consciousness at a Crossroads of Cognitive Science and Phenomenology*. Thorverton: Imprint Academic.
17. Pylkkänen, P. (2014) Can quantum analogies help us to understand the process of thought? *Mind and Matter* 12/1, 61–92. A modified version of a paper originally published in G. Globus, K. Pribram and G. Vitiello (eds.) *Being and Brain. At the Boundary between Science, Philosophy, Language and Arts*, 167–197. Amsterdam: John Benjamins.
18. Pylkkänen, P. (2007) *Mind, Matter and the Implicate Order.* Heidelberg and New York: Springer Frontiers Collection. <http://www.springer.com/philosophy/metaphysics+&+epistemology/book/978-3-540-23891-1>
19. Seager, W. (2013) Classical Levels, Russellian Monism and the Implicate Order. *Foundations of Physics* 43: 548–567.
20. Seager, William and Allen-Hermanson, Sean, (2013) Panpsychism, *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2013/entries/panpsychism/>.
21. Smith, Q. (2003): Why cognitive scientists cannot ignore quantum mechanics? In Smith, Q. & A. Jokic (eds): Consciousness: New Philosophical Perspectives. Oxford University Press, Oxford.

Quantum Ontology: A New Direction for Cognitive Modeling

Sisir Roy

Abstract Human cognition is still a puzzling issue in research and its appropriate modeling. It depends on how the brain behaves at that particular instance and identifies and responds to a signal among myriads of noises that are present in the surroundings (called external noise) as well as in the neurons themselves (called internal noise). Thus it is not surprising to assume that the functionality consists of various uncertainties, possibly a mixture of aleatory and epistemic uncertainties. It is also possible that a complicated pathway consisting of both types of uncertainties in continuum play a major role in human cognition. The ability to predict the outcome of future events is, arguably, the most universal and significant of all global brain functions. The ability to anticipate the outcome of a given action depends on sensory stimuli from the outside world and previously learned experience or inherited instincts. So, one needs to formulate a theory of inference using prior knowledge for decision-making and judgment. Typically, Bayesian models of inference are used to solve such problems involving probabilistic frameworks. However, recent empirical findings in human judgment suggest that a reformulation of Hierarchical Bayesian theory of inference under this set-up or a more general probabilistic framework based approach like quantum probability would be more plausible than a Bayesian model or the standard probability theory. However, as the framework of quantum probability is an abstract one needs to study the context dependence so as understand the new empirical evidences in cognitive domain.

Keywords Bayesian model • Quantum probability • Context dependence • Internal noise • Brain function • Decision making

1 Introduction

For more than 200 years mathematicians and philosophers have been using probability theory to describe human cognition. Recently through several experiments with human subjects [1, 2], violation of traditional probability theory is

S. Roy (✉)
Indian Statistical Institute, Kolkata, India
e-mail: sisir.sisirroy@gmail.com

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_59

clearly revealed in plenty of cases. Literature survey clearly suggests that classical probability theory fails to model human cognition beyond a certain limit. While Bayesian approach may seem to be a promising candidate to this problem, the complete success story of Bayesian methodology is yet to be written. The major problem seems to be the presence of epistemic uncertainty and its effect on cognition at any time point. Moreover the stochasticity in the model arises due to the unknown path or trajectory (definite state on mind at each time point) a person is following. To this end a generalized version of probability theory borrowing idea from quantum mechanics may be a plausible approach. Quantum theory allows a person to be in an indefinite state (superposition state) at each moment of time. A person may be in an indefinite state that allows all of these states to have potential (probability amplitude) for being expressed at each moment [3]. Thus a superposition state seems to provide a better representation of the conflict, ambiguity or uncertainty that a person experiences at each moment [2]. Conte et al. [4] demonstrated that mental states follow quantum mechanics during perception and cognition of ambiguous figures.

These empirical evidences indicate the applicability of quantum probability framework to the decision making in cognitive domain. However, the framework of quantum probability is an abstract framework devoid of material content like concept of elementary particle, the various fundamental constants like Planck constant, speed of light and Gravitational constant in modern physics. So this framework can be applied to any branch of science dealing with decision making such as in Biology, Social science etc. The central issue is how to make this framework context dependence so as to apply to a specific field. In this paper we make an attempt to analyze the whole situation in a critical manner.

2 New Empirical Evidences and Inadequacy of Classical Probability Theory

Various group of scientists [2] made attempts to explain some phenomena related to cognitive modeling using the concept of quantum probability (which has a non-Boolean structure). This cognition spectrum of human mind is usually classified as in six categories as follows:

1. *Disjunction effect*
2. *Categorization – decision interaction*
3. *Perception of ambiguous figures*
4. *Conjunction and Disjunction fallacies*
5. *Overextension of Category membership*
6. *Memory recognition over-distribution effect. Fallacies over-distribution Effect*

The data collected from various experiments related to these six categories clearly indicate the inadequacy of classical probability theory. For example, let us consider the first one i.e. Disjunction effect. Tversky and Shafir [5] discovered a phenomenon called the disjunction effect in the process of testing a rational axiom of decision theory called the sure thing principle [6]. According to the sure thing principle, if under state of the world X you prefer action A over B, and if under the complementary state of the world X^C you also prefer action A over B, then you should prefer action A over B even when you do not know the state of the world. Tversky and Shafir experimentally tested this principle by presenting 98 students with a two stage gamble that is a gamble which can be played twice. Classical probability theory says $P(A) > P(B)$ always. The data from 98 students say that instead of definitely being in the win or loss state, the student enters a superposition state that prevents finding a reason for choosing the gamble. In all these experiments it is claimed that the simple additivity fails for probability law i.e. $P(A + B)$ is not equal to $P(A) + P(B)$. However, they can be explained with the probability rule of quantum framework i.e. $P(A + B) = P(A) + P(B) +$ Interference term. This is a clear indication of violation of classical probability theory.

3 Quantum Probability Theory and Modeling in Cognitive Domain

In quantum theory, the famous double slit experiment clearly indicates the existence of interference term which proves that the probability of union of two mutually exclusive (complementary aspects) events like detection of particle and wave aspect is not equal to the sum of probabilities of event A and probability of the complementary events A^C. This is a clear indication of violation of Boolean structure. In 1932 John von Neumann published his pioneering book [7] where he made a rigorous formulation of the above issue in the following manner. He showed that projection operators in Hilbert space may be considered as elementary "yes"–"no" propositions about measurable properties and he constructed the important logical connectives "and", "or", and "not" in terms of projection operators. Subsequently Birkoff and von Neumann [8] showed that the projection operators and the corresponding propositions constitute an orthocomplemented "lattice" with some additional properties. In contrast to the orthocomplemented and distributive (Boolean) lattice of classical logic, the lattice of the "logic of quantum mechanics" turned out to be much weaker than the Boolean lattice and to be neither distributive nor modular. Initially the community of physicists and philosophers did not accept the "logic of quantum mechanics" as a genuine logic in the strict sense in the tradition of Aristotle, Thomas Aquinas and George Boole, which governs the rules of our rational thinking and arguing.

The main goal of quantum logic is the reconstruction of Hilbert lattices and quantum mechanics in Hilbert space. Within the quantum logic approach quantum mechanics in Hilbert space appears as an abstract and empty theory and presumably universally valid. In this approach there is no classical world and hence no borders between the two worlds from which we could read off Planck's constant. Hence, there is no hope to find the constant h. Here objects or particles can be comprehended if such abstract framework is endowed with notion like "***localizability***" ***and*** "***homogeneity***" [9]. It raise an important issue how to make this abstract framework of quantum probability can be made context dependence i.e. how Planck constant, speed of light etc. can be embedded into the structure of a physical theory. Similar arguments can be made how one applies this type of framework in the domain of cognitive science. If we are to interpret quantum probability (QP) theory as a mechanistic cognitive theory, there must be some method whereby the operations postulated by QP are implemented within the human brain. We emphasize no special quantum physics effects are needed in brain function. If we can show how neurons can compute these operations, then we can interpret QP as making strong claims about *how brains* reason, rather than merely acting as a novel behavioral description of the results of cognitive processing.

For example, if there is one vector for HAPPY and another for UNHAPPY, the current mental state representation might be 0.86HAPPY + 0.5UNHAPPY, representing a state more similar to HAPPY than to UNHAPPY. This is similar to the superposition principle in QP framework. In the above type of description, the issue is what is meant by the unity of complementary aspects at the cellular level. Can we think of any natural constant which gives rise to this kind of unit at the cellular level? Essentially it is needed to construct a framework describing cellular basis of cognition and its relevance to the shift of paradigm.

This type of issue has been debated in eastern philosophy specially in Buddhist philosophy many centuries before. For example, the concept of "Neutral Mind" and "Equanimity" have been described in Buddhist philosophy. In both the states the state of mind is superposition of two complementary states like happiness and unhappiness. However, there is subtle difference between neutral mind and equanimity. In the state of neutral mind there is a scope of decision making where as in the state of equanimity no such scope of decision making. Moreover, there exists various states of equanimity which should be analyzed in the light of modern neuroscience.

References

1. Aerts D, Aerts S, Gabora L (2009): Experimental evidence from quantum structure in cognition. arXiv:0810.5290v2 [math-ph] 22 Jan 2009.
2. Busemeyer JR and Trueblood JS (2011): Theoretical and empirical reasons for considering the application of quantum probability theory to human cognition. Quantum Physics meets TARK. University of Groningen. (Internet access)
3. Heisenberg W (1958): Physics and philosophy. Harper and Row.

4. Conte E, Khrennikov AY, Todarello O, Federici A, Mendolocchio L and Zbilut JP (2009): Mental states follow quantum mechanics during perception and cognition of ambiguous figures. Open Systems & information Dynamics, 16(1).
5. Tversky A and Shafir E (1992): The disjunction effect in choice under uncertainty, Psychological Science 3, 305–309.
6. Savage LJ (1954): The Foundations of Statistics. NY: Wiley.
7. von Neumann J (1932): Mathematical Foundation of Quantum Theory, Springer, Berlin.
8. Birkoff G. and von Neumann J (1936) The Logic of Quantum Mechanics, Annals of Mathematics vol. 37, p.823–843.
9. Mittelstaed P (2011): Rational Reconstructions of Modern Physics. Springer

Part X
Dynamic Brain Forum

Cue-Dependent Modulation of Synchrony in Primates' Medial Motor Areas

Haruka Arisawa, Ryosuke Hosaka, Keisetsu Shima, Hajime Mushiake, and Toshi Nakajima

Abstract Although β oscillations are the representative brain activity in sensorimotor areas and the basal ganglia, how they coordinate activities of multiple structures in the brain is poorly understood. To examine the coordination of the activities of the pre-supplementary motor area (pre-SMA) and the SMA through β oscillations, we recorded local field potentials simultaneously in these areas while monkeys performed a motor task. Examination of inter-area phase difference revealed that the pre-SMA became phase-advanced in β oscillations relative to the SMA when a visual cue signaled initiation of a trial. The strength of phase synchrony decreased markedly while the monkeys were visually instructed about the movement to be performed and was strengthened when the monkeys repeated the movements they had performed in the previous trial. These results suggest that visual input initializes the dynamic state of the pre-SMA and SMA when a trial starts. Additionally, sensory signals seem to be acquired and motor plans formed via modulation of the strength of inter-area synchrony of β oscillations.

Keywords Monkey • Premotor cortex • Field potential • Beta oscillation • Synchrony

1 Introduction

β oscillations are the representative brain activity observed in sensorimotor areas and the basal ganglia [1]. Although it has been established that both the pre-supplementary motor area (pre-SMA) and the SMA are crucially involved in the

H. Arisawa • K. Shima
Department of Physiology, Tohoku University School of Medicine, Sendai, Japan

R. Hosaka
Department of Applied Mathematics, Fukuoka University, Fukuoka, Japan

H. Mushiake • T. Nakajima (✉)
Department of Physiology, Tohoku University School of Medicine, 2-1 Seiryo-cho, Aoba-ku, Sendai 980-8575, Japan
e-mail: b0mb1004@st.med.tohoku.ac.jp

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_60

performance of memory-guided movements [2], few studies have addressed the issue of the coordination of the activities of the two areas from the perspective of synchronized β oscillations. To investigate this issue, we made simultaneous recordings of local field potentials (LFPs) in the pre-SMA and the SMA while monkeys memorized and performed specific movements.

2 Methods

The experimental subjects were two Japanese monkeys (*Macaca fuscata*; one male and one female; weight, 5.8–6.5 kg) cared for according to National Institutes of Health guidelines. Our previous report [3] described in detail the experimental setup and the methods used for animal surgery. During the experiment, each animal sat in a primate chair holding a handle in each hand (Fig. 1a). For each trial of this task, the animals were required to perform a motor sequence consisting of two motor elements. The temporal order of the required motor sequence was selected pseudo-randomly and altered in blocks of nine trials. The first three trials in a block were performed under the guidance of two visually presented instructional cues (visually guided trials; Fig. 1b). The color of the instructional cue indicated left forearm supination (red) or pronation (blue) and right forearm pronation (yellow) or supination (green). In the remaining six trials in the block, the animal performed the movements specified in the previous visually guided trials with no visual instructions (memory-guided trials). At the beginning of a trial, a central fixation point (FP) was presented on the screen in front of the subject to serve as a cue to initiate the trial. The monkey was required to maintain the handle placement and

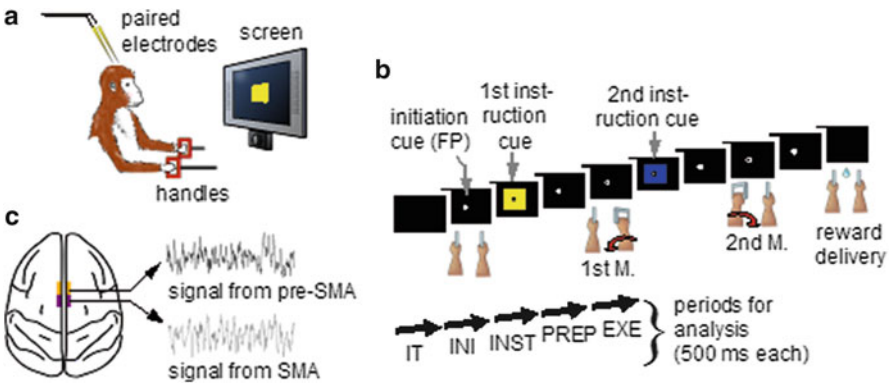


Fig. 1 (**a**) Experimental setup (**b**) Task sequence (**c**) Schematic drawing of medial motor areas (*top view*)

eye fixation for 1.5 s, during which the instructional cue for the first movement was presented for 0.5 s in the visually guided trials. Subsequently, the FP dimmed to serve as the first-movement trigger signal. The animal was required to perform the first movement and return the handle to the neutral position within the reaction-time limit (1 s). After a delay period of 1 s, during which the instructional cue for the second movement was presented for 0.5 s (in visually guided trials), the animal was given the trigger signal for the second movement. A series of correct movements without a fixation break was rewarded with the delivery of juice, followed by a 1.5-s intertrial interval (ITI).

After each animal became proficient in performing this behavioral task, an acrylic recording chamber and head-fixation bolts were implanted on the animal's skull. After complete recovery from the surgery, we used glass-insulated Elgiloy microelectrodes to make simultaneous recordings of LFPs for the SMA and pre-SMA in one hemisphere while the subjects performed the task (Fig. 1c). Online data collection was performed using a multichannel acquisition processor. Eye position was monitored using an infrared corneal-reflection-monitoring system at 1 kHz (Millennium G200, Matrox).

As a spectral measure of the correlation of the two signals across frequencies, we calculated the coherence from the cross-spectral density between the two LFPs and normalized it by the power spectral density of each [4] using the LFPs recorded in visually guided trials. Coherence values range from 0 to 1. A value of 0 indicates that the two signals were completely uncorrelated, whereas a value of 1 indicates that the signals were completely correlated at frequency f. We defined the frequency that yielded the greatest coherence in the β band (15–40 Hz) as the frequency of interest (f_0) for further analysis.

Our main interest was the cue-dependent modulation of inter-area β oscillations during visually guided trials. We thus defined five consecutive analytic periods in a visually guided trial, as shown at the bottom of Fig. 1b: intertrial period (IT), trial-initiation period (INI), instruction-cue period (INST), motor-preparation period (PREP), and motor-execution period (EXE). To extract the instantaneous phase for the frequency f_0, we applied a wavelet-based approach. The signal of the LFP was convoluted by a complex Gabor's wavelet $w(f_0, t)$:

$$w(f_0,\ t) = g(t)\left\{\exp(i2\pi f_0 t) - \exp\left(-\sigma^2(2\pi f_0)^2\right)\right\},$$

where

$$g(t) = \frac{\exp\left(-\frac{t^2}{4\sigma^2}\right)}{2\sigma\sqrt{\pi}},\ \sigma = \frac{5}{2\pi f_0}.$$

The instantaneous phase difference between a pair of electrodes $\Delta\varphi(f_0, t)$ was expressed as the relative phase of the signal recorded from the pre-SMA to that from the SMA:

$$\Delta\varphi\,(f_0,\ t) = \varphi_{pre}\,(f_0,\ t) - \varphi_{sma}\,(f_0,\ t)\,.$$

To determine the strength of inter-hemispheric phase synchronization, we calculated a phase-synchronization index (PSI) for a given 500-ms analytic period in a trial as the mean vector length of the angular dispersions of the phase differences in the analytic period [5]. If the phase difference varies little in the analytic period, the PSI is close to 1; otherwise, it is close to 0. To examine how the presentations of the initiation cue and the motor instructional cue influenced the inter-area phase relationship, we computed the mean phase difference $\overline{\Delta\varphi}$ for the IT, INI, and INST periods in each recording session. $\overline{\Delta\varphi}$ was calculated as the angle of the circular mean of the phase differences observed in the analytic period across trials.

3 Results

We performed a total of 30 paired recordings (15 sessions in each hemisphere) for monkey M and 61 (29 and 32 sessions in the left and right hemisphere, respectively) for monkey N. Coherence analysis revealed that $f_0 = 35$ Hz for monkey M, and $f_0 = 22$ Hz for monkey N. The dynamics of the inter-area phase difference and the strength of the phase synchronization between areas based on the analysis for the frequency f_0 are discussed in the next section.

3.1 Relative Phase Advance of Pre-SMA to SMA on Initiation of a Typical Trial

Figure 2a shows an example of a pair of 35 Hz LFPs simultaneously recorded from the pre-SMA (black trace) and SMA (gray) in monkey M during a 1,500-ms interval including the IT, INI, and INST periods. The superimposed red trace represents the time course of the phase difference (relative phase of the pre-SMA to that of the SMA). To quantify the mean phase difference for each period, we computed the circular mean of instantaneous phase differences (Fig. 2b; for display purposes, the instantaneous phase difference is illustrated for every 50 ms) within each period. The direction of the circular mean was defined as the mean phase difference in a period (Fig. 2c, directions of red arrows). During the IT period in the trial, the SMA was slightly phase advanced relative to the pre-SMA. In the INI period, the pre-SMA often led the SMA in the β-cycle. In the INST period, the pre-SMA became more phase advanced relative to the SMA in terms of the mean phase difference. Because

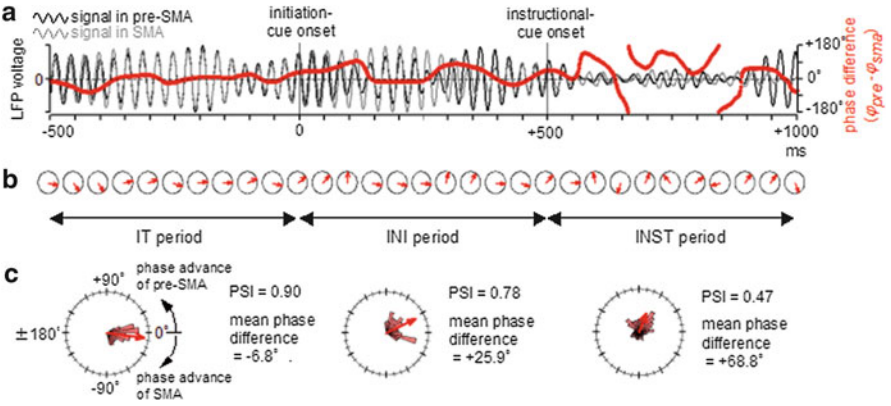


Fig. 2 Dynamics of inter-areal β-synchrony and distribution of phase differences in an example trial

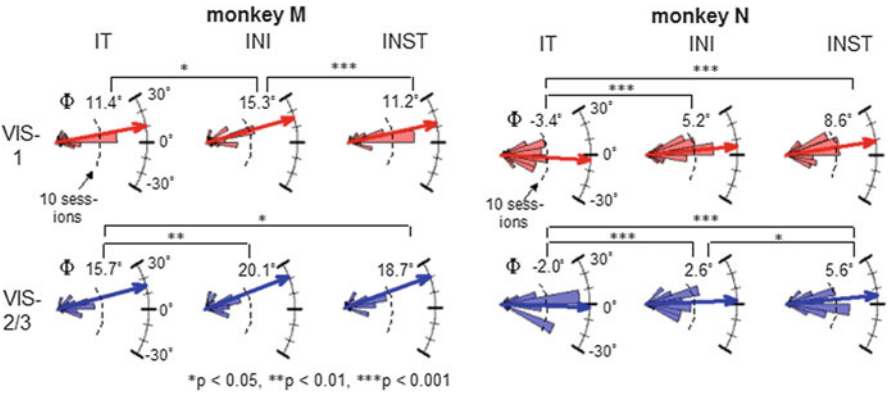


Fig. 3 Distributions of mean phase differences $\overline{\Delta\varphi}$ and their grand means (Φ)

the inter-area phase synchrony decayed toward motor onset (see next section), we applied the analysis of the phase difference to only the IT, INI, and INST periods.

To characterize the dynamics of the inter-area phase difference across sessions, we first confirmed that the mean phase difference $\left(\overline{\Delta\varphi}\right)$ computed for each of the three periods was significant in all the sessions (Rayleigh test, $p \ll 0.001$). We then performed a population analysis across sessions. Figure 3 shows distributions of $\overline{\Delta\varphi}$ in circular histograms with arrows indicating grand mean directions (Φ) across sessions. We found that the relative phase of the pre-SMA in the INI period (middle column) was significantly more advanced than it was in the IT period (left column) in each trial for both monkeys ($p < 0.05$, Moore test for paired data on a circular scale of measurement). The relative phase of the pre-SMA continued to be more advanced in the INST period than in the IT period (left vs. right column), except in the VIS-1 trial for monkey M.

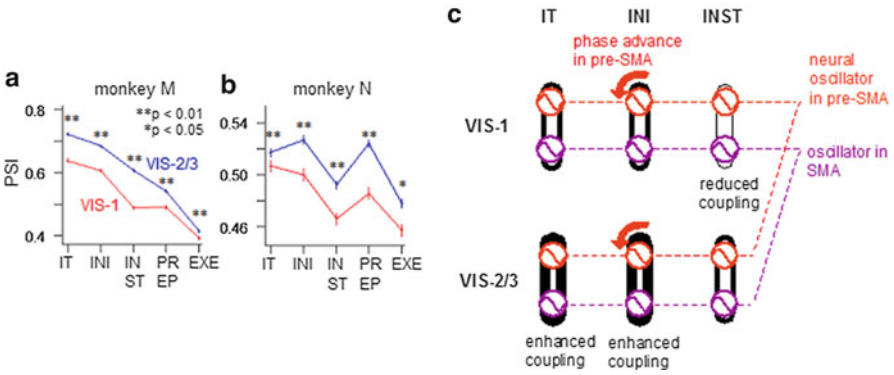


Fig. 4 (**a**, **b**) Time course of strength of phase synchrony for VIS-1 (*red*) and VIS-2/3 (*blue*) trials (**c**) Model of coupling oscillators consisting of pre-SMA (*orange*) and SMA (*purple*)

3.2 Instructional-Cue Dependent Suppression and Enhancement of Beta Synchrony During Implementation of Motor Plan

In the example trial shown in Fig. 2, the PSI was near 1 in the IT period and decreased in the INST period. According to our task design, the repetitive presentation of identical sets of instructional cues in a series of three visually guided trials prompted implementation of the motor plan. To examine the effects of task events and repetitive cue presentation on the inter-area phase synchrony, we computed the median PSIs across all recording sessions for each monkey (Fig. 4a, b). In the first visually guided trials (VIS-1; red trace), the PSI was higher during the IT and INI periods and was considerably suppressed in the INST period. The PSI rebounded in the PREP period and was lowest in the EXE period. Although the PSI exhibited similar temporal dynamics in the remaining visually guided trials (VIS-2/3; blue trace) to those in the VIS-1 trials, it was significantly greater in those trials than in VIS-1 trials across all analytic periods ($p < 0.05$, Mann–Whitney U-test).

4 Conclusions

In the present study, we examined the cue-dependent modulation of inter-area synchrony in β-band LFPs recorded in the pre-SMA and SMA. We determined the relative phase advance of the pre-SMA to the SMA on initiation of a trial. We also found that phase synchrony was generally enhanced in parallel with the repetitive performance of the identical motor sequence in a block. Moreover, we found that the presentation of a visual instructional cue for the forthcoming movement attenuated the degree of the inter-area phase synchrony, especially in VIS-1 trials.

To interpret our results in terms of interacting oscillators, we constructed a conceptual model (Fig. 4c). In this model, the pre-SMA and SMA are regarded as a pair of functionally coupled oscillators. The relative phase advance of the pre-SMA on trial initiation (orange arrows in the middle column) may be triggered by presentation of the initiation cue. This view is supported by the finding that neuronal activity in the pre-SMA often responds to visual stimuli [6] and that the pre-SMA has access to visual input via projections from the inferior parietal lobule [7]. This phase advance may, in turn, initialize the causal relationship between the activities of the pre-SMA and those of the SMA, enabling the pre-SMA to exert top-down control over the SMA. This view is supported by a simulation study [8]. Additionally, the strength of phase synchrony is represented as the thickness of a black connecting background in Fig. 4c. The remarkable reduction in phase synchrony in response to the initial presentation of the motor instructional cue (top row, right column) suggests that use of visual instructions leads to a reduction in functional coupling between the two oscillators. The enhancement of phase synchronization that accompanies repetition of the identical motor sequence indicates that the functional coupling of the two oscillators is strengthened in parallel with the implementation of a motor plan (bottom row, left and middle column). Taken together, our results indicate that the dynamics of the inter-area synchrony of β oscillations may mediate the functional coupling between the pre-SMA and the SMA, enabling the initialization of the dynamic state of the two areas, the acquisition of sensory signals, and the implementation of a motor plan.

Acknowledgments This work was supported by grants from the CREST, JST and a Grant-in-Aid for Scientific Research on Innovative Areas "Neural creativity for communication (No.4103)" (24120702) of MEXT.

References

1. Engel, AK. & Fries, P. Curr. Opin. Neurobiol. **20** (2010) 156–165.
2. Tanji, J. Ann. Rev. Neurosci. **24** (2001) 631–651.
3. Nakajima, T. et al., J. Neurophysiol. **101** (2009) 1883–1889.
4. Jerbi, K. et al., Proc. Natl. Acad. Sci. USA. **104** (2007) 7676–7681.
5. Kuramoto, Y. Chemical oscillations, waves, and turbulence. Springer, Berlin (1984).
6. Matsuzaka, Y., Aizawa, H. & Tanji, J. J. Neurophysiol. **68** (1992) 652–662.
7. Luppino, G. J. et al. J. Comp. Neurol. **338** (1993) 114–140.
8. Battaglia, D. et al. PLoS Comput Biol. **8** (2012) e002238.

Multisynaptic State Functions Characterizing the Acquisition of New Motor and Cognitive Skills

José M. Delgado-García, Raudel Sánchez-Campusano, Alejandro Carretero-Guillén, Iván Fernández-Lamo, and Agnès Gruart

Abstract Learning and decisions are precise functional states of brain cortical circuits that can only be approached by the use of multidisciplinary and complementary tools in behaving animals. The availability of genetically manipulated mice and rats, of mathematical and computational neuroscience methods, and of advanced electrophysiological techniques—susceptible of being applied in behaving animals during the acquisition of different learning paradigms—has largely facilitated this approach. Here, we have recorded activity-dependent changes in synaptic strength in different synapses of hippocampal and prefrontal circuits during the acquisition and storage of classical and instrumental conditioning paradigms. Furthermore, we have developed a dynamic approach of multisynaptic state functions to characterize the acquisition of new motor and/or cognitive skills. In our opinion, a synaptic state function is analogous to a precise picture of synaptic weights while the behaving animal learns the selected task. Therefore, the different state functions of large cortical circuits during the very moment at which learning is taking place could be specifically defined by 3D-arrays of synaptic sites, learning stages, and behaviors. Couplings between the different synaptic state functions are determined by means of weight functions that characterize the changes in synaptic strengths, the type (linear or nonlinear) of interdependences among state functions, as well as the timing and correlation relationships among them. The detailed analysis of the collected data indicates that many synaptic sites within cortical circuits modulate their synaptic strength across the successive stages of acquisition of associative learning tasks. The expected main output of this type of experimental approach would be that learning is the result of the activity of wide cortical and subcortical circuits activating particular functional properties of involved synaptic nodes, and that we can quantify these activation patterns by means of state and weight functions. In this regard, we expect that a map of state functions relating the acquisition of new motor and cognitive abilities and the underlying synaptic plastic changes will be offered in the near future for different types of learning tasks and situations. This same optimized approach

J.M. Delgado-García • R. Sánchez-Campusano • A. Carretero-Guillén • I. Fernández-Lamo
A. Gruart (✉)
Division of Neurosciences, Pablo de Olavide University, Seville, Spain
e-mail: agrumas@upo.es

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_61

could be applied to the selective stimulation of synaptic nodes across the involved circuits, in genetically modified animals, or in animals receiving selective injections of si-RNA, and other molecular-disturbing procedures.

Keywords Synaptic strength • Hippocampus • Prefrontal cortex • Conditioning • Associative learning • Cognition

1 Introduction

It is traditionally accepted that the study of the different neural sites and mechanisms underlying learning and memory processes has to be approached with the help of molecular, histological, and in vitro electrophysiological procedures. Although these experimental approaches have rendered important insides with regard to brain structural and functional properties, they have two important limitations: (i) they use to generate profound alterations of the studied nervous system; and, (ii) they do not report any information on events taking place during the very moment of the acquisition process. In this regard, learning and related cognitive and motor processes should be studied, at the end, at live [1, 2].

Taking into account that neurons are the basic functional elements characterizing the nervous tissue, we would need to know the specific functional properties that the different neural types present in cortical circuits and, very importantly, its functional contribution, moment to moment, to the global process of learning, memory storage, and recall. Thus, each neuronal type in a selected cortical circuit plays a specific role that can only be determined by the use of experimental models allowing its study in the best physiological conditions and during the acquisition of new motor or cognitive abilities. We have developed the basic technology for the study of activity-dependent changes in synaptic strength at a given relay site [1–4]. In this paper, we have extended this information to many different synapses located in hippocampal and prefrontal circuits (see Figs. 1 and 2) during the acquisition and storage of two types of associative (Pavlovian and instrumental) learning tasks in behaving rabbits and rats. We are also introducing here the mathematical tools that will be used in order to obtain relevant data on neuronal network processing during associative learning tasks.

2 Methods

In a first series of experiments, animals (rabbits) were prepared for the chronic recording of the electromyographic activity of orbicularis oculi muscle and of synaptic activity (field excitatory post-synaptic potentials, fEPSPs) evoked at the intrinsic hippocampal circuit or by the stimulation of its main input, i.e., the perforant pathway (Fig. 1). Rabbits were trained with a Pavlovian conditioning

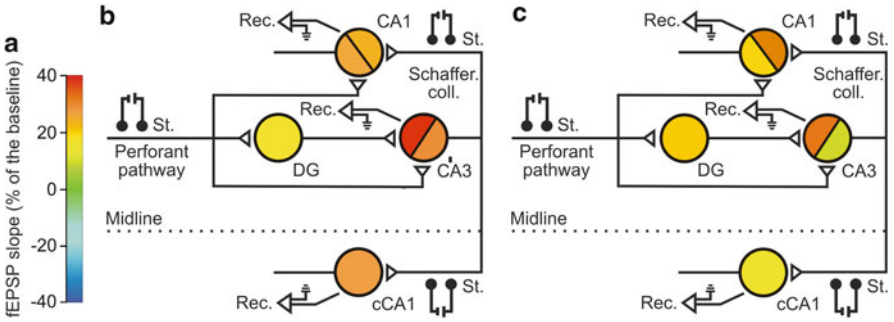


Fig. 1 A diagrammatic representation of synaptic weights present in the hippocampal circuit of behaving rabbits during classical conditioning (**b**, session 1; **c**, session 8) of eyelid responses using a delay conditioning paradigm. The color code is illustrated in (**a**). In brief, animals were implanted with stimulating (St.) electrodes in the perforant pathway, the Schaffer collateral/commissural pathway, or in the contralateral CA3 area, and with recording electrodes (Rec.) in the dentate gyrus (DG), and the hippocampal CA3 and CA1 areas. Synaptic activation took place during the CS (tone)—US (air puff) interval

protocol (i.e., a classical eyeblink conditioning). Field EPSPs were evoked in the different hippocampal synapses at the interval between conditioned (a tone) and unconditioned (an air puff presented to the cornea) stimulus presentations. The simultaneous recording of synaptic activities at different neural sites offered a still unknown picture of the specific functional states taking place at hippocampal during the actual acquisition process.

In a second series of experiments, Wistar rats were implanted with stimulating and recording electrodes in selected sites of the intrinsic hippocampal circuit and/or in the perforant pathway. Rats were trained in a Skinner box to press a lever in order to obtain a small piece of food. Animals were stimulated at different hippocampal and prefrontal synapses during their performance in the Skinner box task (see Fig. 2).

For analysis, we used here mathematical tools designed in our laboratory to obtain the state functions characterizing the acquisition of new motor and/or cognitive skills. The programs/scripts used here were developed by one of us (R.S.-C.) with the help of MATLAB (The MathWorks, Natick, MA, USA) routines [3].

3 Results

A few years ago, we showed that the hippocampal CA3→CA1 synapse presents a significant change in strength during the acquisition of a type of associative learning task in alert behaving mice: i.e., the classical conditioning of eyelid responses [1, 2]. It was also shown in this study that this learning-dependent change in synaptic strength was linearly related with the rate of acquisition of the

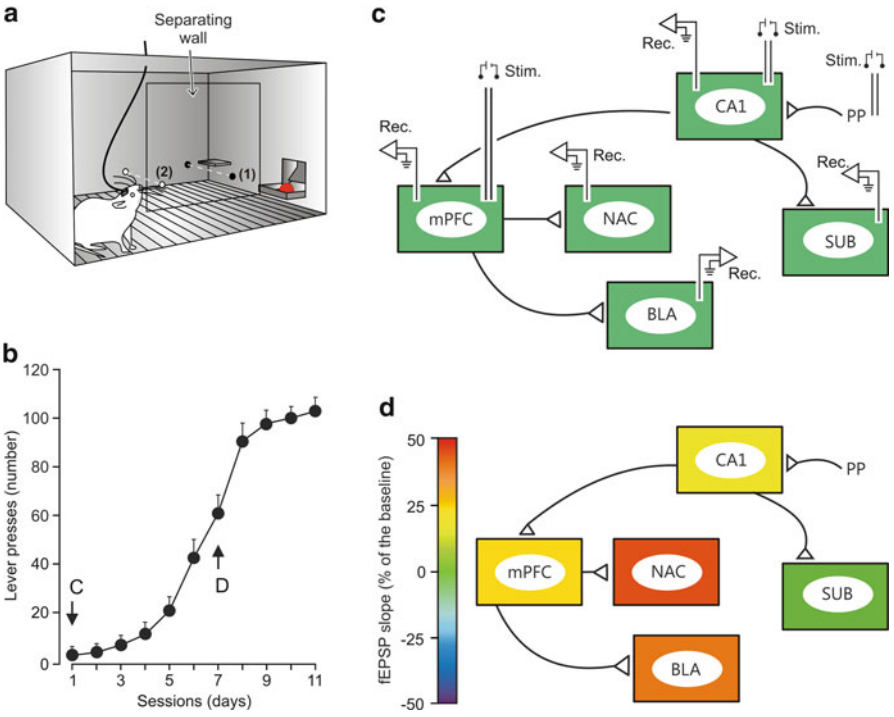


Fig. 2 An example of the quantitative analysis for changes in synaptic strength taking place in different hippocampal and prefrontal synapses during the acquisition of an instrumental learning task. (**a**) Experimental design. The selected synapses were activated when the animal approached to the lever to press it, i.e., when crossing the indicated photoelectric cells. (**b**) Acquisition curve. (**c**, **d**) Synaptic weights (fEPSP slopes) recorded during the first and the seventh training sessions. Abbreviations: *BLA* basolateral amygdala, *mPFC* medial prefrontal cortex, *NAC* nucleus accumbens septi, *Sub* subiculum

conditioned eyeblinks, suggesting a more-or-less direct relationship between the acquisition process and the underlying synaptic plastic changes. At that moment, it was assumed that other synapses present in the intrinsic hippocampal circuit and the many other related to its main inputs (perforant pathway) and outputs (other cortical structures) should present similar changes in synaptic strength [2]. Indeed, we have addressed this question in a recently published study [4] and with different ongoing experiments being carried out in our laboratory. As illustrated in Fig. 1, the different hippocampal synapses present a complex evolution of their synaptic strength across the successive conditioning sessions. Thus, it was clear that each synapse in the hippocampal network contributes in a different way to the acquisition process. In addition, we have carried out a similar classical eyeblink conditioning in behaving mice, including the analysis of fEPSPs changes taking place in nine different hippocampal synapses across conditioning (not illustrated). Here again, changes in synaptic strength across conditioning indicated the presence of a timed and specific plasticity pattern characterizing this type of associative learning.

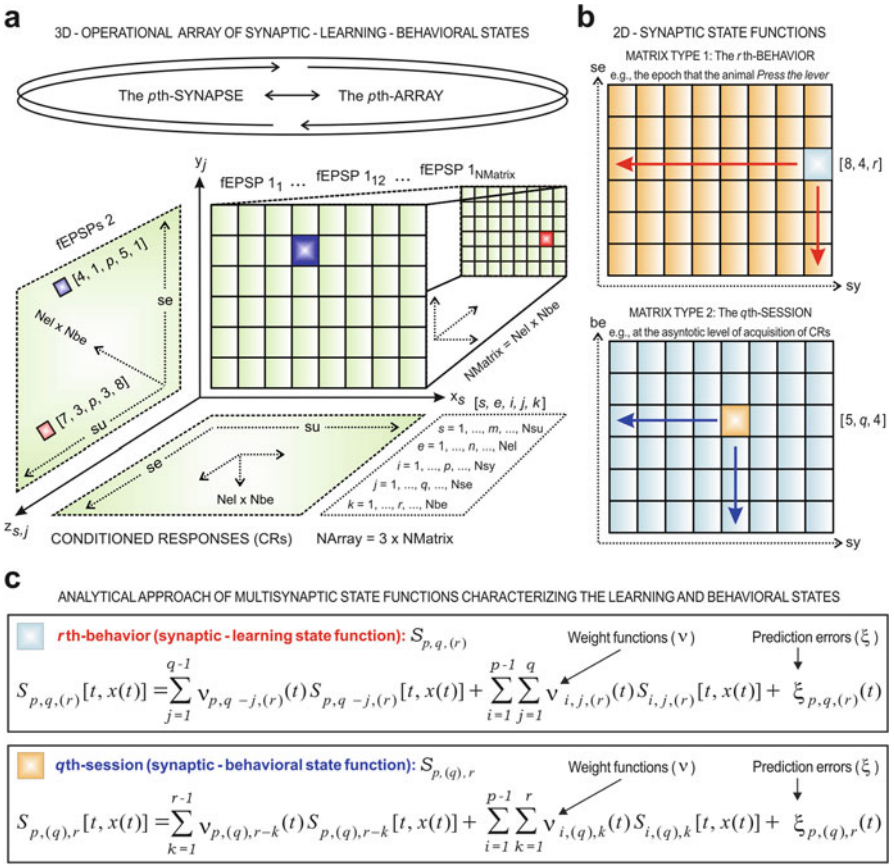


Fig. 3 Operational 3D-array of state functions. (**a**) Here, a synaptic state function is analogous to a precise picture of a synaptic pathway while the behaving animal learns the task. Therefore, the different sate functions of large cortical synaptic circuits during the very moment at which learning is taken place could be specifically defined by 3D-arrays of synaptic sites, learning stages, and behaviors. (**b**) Diagram of the matrices for the synaptic-learning (*r*th-behavior) and synaptic-behavioral (*q*th-session) state functions analyses. Two examples of the state functions are presented in the inset (e.g., [8, 4, *r*] for the *r*th-behavior; and [5, *q*, 4] for the *q*th-session). (**c**) Mathematical formulation of the multisynaptic state functions for the *r*th-behavior or the *q*th-session

In Fig. 2 is illustrated a similar experimental approach for the study of synaptic plasticity in behaving animals during another type of associative learning. In this case, we studied changes in fEPSPs evoked at specific synapses connecting the hippocampus, prefrontal cortex, accumbens septi, and the basolateral amygdala (Fig. 2c, d) during the acquisition of an instrumental learning task, using a fixed ratio (1:1) paradigm, i.e., the experimental animal received a small pellet of food every time its pressed a lever (Fig. 2b). Electrical stimulation of the selected synapse was carried out at the moment the rat approached to the lever (Fig. 2a). As illustrated in

Fig. 2c, d, there are important changes in synaptic strength in the selected cortical network from the initial conditioning session to the session when the designed task is acquired by the animals.

Finally, in Fig. 3 we present the proposal of a 3D-array of synaptic-learning-behavioral states (Fig. 3a) and the diagrammatic representation of the analytical approach of multisynaptic state functions (Fig. 3b, c). According to this analytical design, two types of matrix can be formed: (1) a matrix for the relationships between different synaptic-learning states during the *r*th-behavior; and, (2) a matrix for the relationships between different synaptic-behavioral states during the *q*th-session of conditioning (Fig. 3b). The weight functions (Fig. 3c) determine the strength and type of interdependences among states as well as the timing-causality relationships between them. The prediction error (Fig. 3c) estimates the uncertainties associated with the model and depends on the past values of all the synaptic-learning states. Preliminary results obtained with the analytical procedure suggest the presence of specific spatial-temporal patterns of synaptic weights characterizing each particular learning situation.

4 Discussion

We hope that the present experimental approach will help to offer, in the near future and for the very first time, a complete and quantifiable picture of synaptic events taking place in cortical circuits directly involved in the acquisition, storage, and retrieval of different types of associative learning tasks. Indeed, our experimental approach is susceptible of being used in different types of associative learning as classical and instrumental conditioning, as well in other types of non-associative learning tasks as object recognition and spatial orientation [1–4].

References

1. Gruart, A., Muñoz, M.D., and Delgado-García, J.M. Involvement of the CA3-CA1 synapse in the acquisition of associative learning in behaving mice. J. Neurosci. (26) (2006) 1077–1087.
2. Delgado-García, J.M., and Gruart, A. Building new motor responses: eyelid conditioning revisited. Trends Neurosci. (29) (2006) 330–338.
3. Sánchez-Campusano, R., Gruart, A., and Delgado-García, J.M. Dynamic changes in the cerebellar-interpositus/red-nucleus-motoneuron pathway during motor learning. Cerebellum (10) (2011) 702–710. doi:10.1007/s12311-010-0242.
4. Jurado-Parras, M.T., Gruart, A., Delgado-García, J.M. Observational learning in mice can be prevented by medial prefrontal cortex stimulation and enhanced by nucleus accumbens stimulation. Learn. Mem. (19) (2012) 99–106.

Visual Hallucinations in Dementia with Lewy Bodies (I): A Hodological View

Hiroshi Fujii, Hiromichi Tsukada, Ichiro Tsuda, and Kazuyuki Aihara

Abstract Patients with dementia with Lewy bodies (DLB) frequently experience the phenomenon of visual hallucination (VH), which Collerton (2005) aptly described as "people see things that are not there." The possible involvement of cholinergic deficiency in the VH seen in DLB have long been proposed, but the precise neural mechanism to account for the particular phenomenology of VH is not known. The aim of this work is to delineate the core mechanisms of VH based on considerations of symptomatology and recent data on pathophysiology, viewed from the structure of the brain's cognitive system. The key may be in the prefrontal top-down facilitations on the category of the "seen" object at the center of attention. The prefrontal cortex (PFC: VLPFC/OFC) creates such a categorical "index" as a bias for the inferior temporal (IT) cortex to reactivate a detailed image of the object. In doing so, the PFC quickly receives information on a low spatial frequency image (LFI) of the object via a cortical *short-cut* (the magnocellular pathway), together with internal signals such as expectancy and emotion, and information on the *context* and *setting* in which they appear. The VH may appear when a part of this system breaks down. The PFC then creates a bias (index) for IT cortex that is hallucinatory. This may happen when some *conduction disturbances* occur along the short-cut pathway due to either cortical degenerations (e.g., loss of cortical pyramidal neurons/pre-synaptic axons), or loss of the α7 subtype of nicotinic receptor (α7-nAChR).

H. Fujii (✉)
Department of Intelligent Systems, Kyoto Sangyo University, Kyoto 603-8555, Japan
e-mail: fujii@cse.kyoto-su.ac.jp

H. Tsukada • I. Tsuda
Research Institute for Electronic Science, Hokkaido University, Kita 12 Nishi 7, Kita-ku, Sapporo 060-0812, Japan

K. Aihara
Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_62

Keywords Dementia with Lewy bodies • Hallucinations • PFC • Top-down facilitation • Cortical degeneration • nAChR loss

1 Introduction

Patients with dementia with Lewy bodies (DLB) frequently experience the phenomenon of visual hallucination (VH), which Collerton (2005) [1] aptly described as "people see things that are not there."

The possible involvement of cholinergic deficiency in the VH seen in DLB has long been proposed, but its precise neural mechanism is not fully understood [1, 2].

2 Phenomenology

The phenomenology of VH symptoms associated with DLB shows a clear contrast with psychedelic (serotonin-related) visual hallucinations in LSD use, or schizophrenic VH. This phenomenology raises a number of important questions related to VH in DLB: Why is hallucinatory images contain mostly single entities, e.g., a vivid and colored human or animal appearing at the (foveal) center of attention? Why is a hallucinatory figure *pasted* on a background scene that is perfectly normal? Why does a figure appear irrespective of whether the eyes are closed or open? Why does the imagery continue for a few minutes (not seconds, nor hours). Moreover, why are the images generally *consistent* with the *context* and *setting* in which they appear? This particular symptomatology may reflect the internal mechanisms of VH associated with DLB.

The aim here is to delineate the root mechanisms of VH through considerations of such phenomenology and the structure of the visual cognitive system together with recent data on pathophysiology.

3 Role for Orbitofrontal Cortex in Cognitive System

Top-Down Facilitation The content and character of VH in DLB primarily reflects the nature of visual processing [1], and VH is a consequence of a dysfunction of the normal cognitive system.

The orbitofrontal cortex (OFC) and probably the ventrolateral prefrontal cortex (VLPFC) are together involved in the categorical identification of the object at the center of attention [3]. This information ("index") is then sent back to the inferior temporal (IT) cortex as a bias to form a detailed representation (Fig. 1).

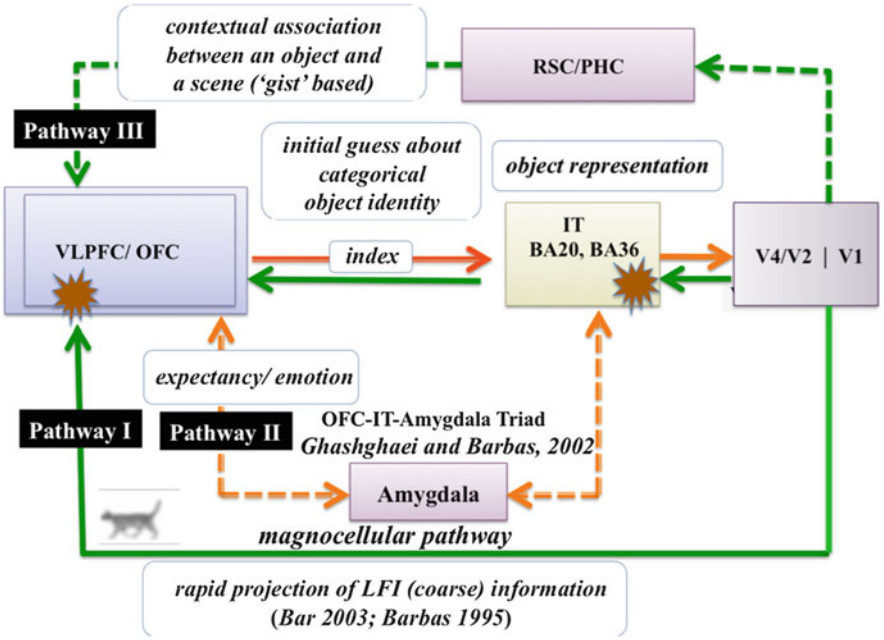


Fig. 1 Three pathways in normal scene and object perception. The stars (*red*) indicate possible sites of conduction disturbance in VH. When disturbances occur, the PFC correctly receives contextual data and emotion/expectancy as in normal cognitive processes, but the quickly arriving short-cut data of an external object via Pathway I is missing. Bottom-up signals of external objects via V4/V2 are also missing for IT, caused by the second disturbances

OFC receives, among others, at least three streams of projections important in the present context: one is the low spatial frequency image (LFI) via a "cortical shortcut", i.e., the magnocellular pathway from secondary visual systems (such as V2/V4) [3, 4]. The second and the third streams bring, respectively, "emotion and expectancy" from the OFC-ITC-amygdala triad [5] and contextual information as the "gist" of the scene, from the retrosplenial/parahippocampal cortices (RSC/PHC), which contribute to OFC processing by providing context-based top-down facilitation [4].

Hodological Dysfunction We hypothesize that a *conduction disturbance* temporarily occurs somewhere along Pathway I. We postulate that this might be caused primarily by the recently reported cortical degeneration of gray matter in the PFC and IT in hallucinating DLB patients [7], or loss of the α7 subtype of nicotinic receptors [6]. Because of these disturbances the PFC's decision on *categorical identity* is made essentially on the basis of context and expectancy alone. Hence the PFC creates and sends "hallucinating" indices to IT.

There is another site of pathology, i.e., the IT cortex [7, 8]. The possible disconnection of IT from the secondary visual cortices V4/V2 due to conduction

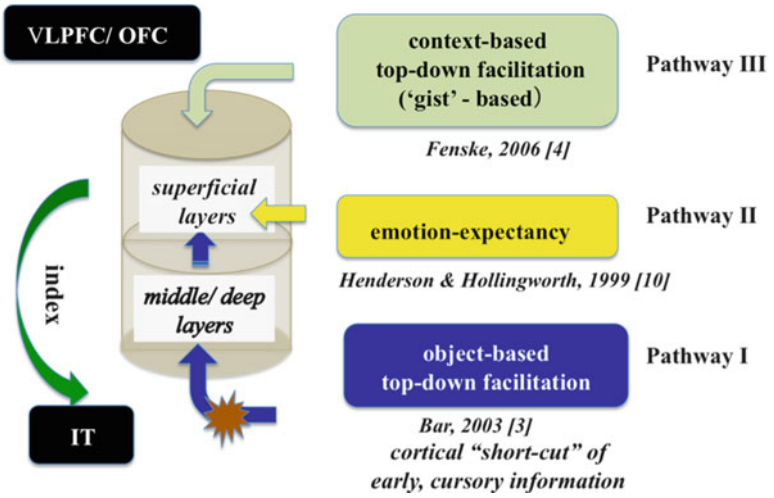


Fig. 2 PFC decision on top-down bias is based on three signals

disturbances as above could be the second causative factor for VH (Fig. 1). This may block IT's ability to compensate, based on detailed bottom-up data, potentially hallucinatory images induced by PFC (Fig. 2).

The disconnections at the two sites could be regarded as an instance of dynamically fluctuating, and unidirectional *hodological*[1] deficiency [9].

4 Proposals: Root Cause for Visual Hallucinations Associated with DLB

1. The root cause and the responsible areas: malfunction of PFC (VLPFC/OFC) and IT cortex. PFC sends a "hallucinatory" index of categorical identity on the basis of *emotion-expectancy* and the *top-down* context alone, ignoring external but cursory information via Pathway I.
2. Possible pathophysiology may be *conduction disturbances* along Pathway I conveying external LSI signals, and at the same time on the pathway from the secondary visual system to IT. This is due to either [a] cortical degenerations as loss of cortical pyramidal neurons/pre-synaptic axons [7], or [b] loss of the α7 subtype of nicotinic receptor (α7-nAChR) [8].

[1]By "hodological" dysfunction we mean dysfunction related to connections between brain regions, as compared with dysfunction localized to specific brain regions (topological) (Ffytche) [9].

5 Discussions

In the companion paper [11], we discuss such problems as: (1) The structure of the attractor space for an index in PFC and its dynamics. (2) How can the IT cortex activate detailed images based on a PFC biasing signal? As illustrative examples we show some of the results from our computational model that reproduced such a phenomena with hallucinations.

Acknowledgements We appreciate valuable discussions with Dr. Daniel Collerton, University of New Castle upon Tyne, United Kingdom, Prof. Guy Sandner, Université Louis Pasteur, France, and Dr. Yuichi Katori, The University of Tokyo, Japan. This work was partially supported by JSPS KAKENHI Grant Number 20246026, 22500281 and, also partially supported by a Grant-in-Aid for Scientific Research on Innovative Areas (No. 4103) (21120002) from Ministry of Education, Culture, Sports, Science and Technology (MEXT) in Japan.

References

1. E.K. Perry and R.H. Perry, Acetylcholine and Hallucinations: Disease Related Compared to Drug-Induced Alterations in Human Consciousness, Brain and Cognition 28 (1995), 240–258.
2. D. Collerton et al., Why people see things that are not there: A novel Perception and Attention Deficit model for recurrent complex visual hallucinations, Behavioral and Brain Sciences 28 (2005), 737–794.
3. M. Bar, A Cortical Mechanism for Triggering Top-Down Facilitation Visual Object Recognition, Journal of Cognitive Neuroscience 15 (2003), 600–609.
4. M.J. Fenske et al., Top-down facilitation of visual object recognition: object-based and context-based contributions, Progress in Brain Research 155(2006), 3–21.
5. H.T. Ghashghaei and H. Barbas, Pathways for Emotion: Interaction of Prefrontal and anterior Temporal Pathways in the Amygdala of the Rhesus Monkey, Neuroscience 115 (2002), 1261–1279.
6. R.T. Reid et al., Nicotinic receptor losses in dementia with Lewy bodies: comparisons with Alzheimer's disease, Neurobiology of Aging 21 (2000) 741–746.
7. C. Sanchez-Castaneda et al., Frontal and Associative Visual Areas Related to Visual Hallucinations in Dementia with Lewy Bodies and Parkinson's Disease with Dementia, Movement Disorders 25 (2010), 615–622.
8. J.A. Court et al., Visual hallucinations are associated with lower α bungarotoxin binding in dementia with Lewy bodies, Pharmacology, Biochemistry and Behavior 70 (2001) 571–579.
9. D.H. Ffytche, The hodology of hallucinations, Cortex 44 (2008) 1067–1083.
10. J. M. Henderson and A. Hollingworth, High-level scene perception, Annual Review of Psychology 50 (1999), 243–271.
11. H. Tsukada, H. Fujii, I. Tsuda and K. Aihara, Recurrent Complex Visual Hallucinations in Dementia with Lewy Bodies (II): Computational Aspects (*this issue*).

Dividing Roles and Ordering Information Flow in the Formation of Communication Systems: The Influence of Role Reversal Imitation

Takashi Hashimoto, Takeshi Konno, and Junya Morita

Abstract The process and mechanism behind the formation of symbolic communication systems is studied in this paper by using human cognitive experiments and the computer simulation of cognitive architecture. In the presented experiment, pairs of participants carry out a coordination task repeatedly in which a symbolic message is passed between them. Two-thirds of participant pairs ($n = 14$) formed communication systems that could stably solve the coordination task by establishing role division by utilizing the order of sending messages, namely turn-taking. The behavioral strategy behind this role division is then analyzed by using transfer entropy and simulating ACT-R to model participant behavior. The analysis of the transitions of transfer entropy shows that successful pairs behave consistently from the beginning of the experiment and begin to include their partners' behavior when deciding on role division. The comparison between human experiment and ACT-R simulation data suggests that role reversal imitation, whereby participants store partners' behaviors and utilize them to decide on their own behavior, is effective at establishing a communication system.

Keywords Communication systems • Role division • Imitation • Transfer entropy • ACT-R

1 Introduction

Recent studies have used experimental semiotics in order to examine how symbolic communication systems are formed [1, 2]. The quantitative analysis of the formation process of symbolic communication systems shows that the formation process comprises three scaffolding stages: establishing common ground, sharing the

T. Hashimoto (✉) • T. Konno
Department of Knowledge Science, School of Knowledge Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan
e-mail: hash@jaist.ac.jp

J. Morita
Institute of Innovation for Future Society, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_63

symbol system, and dividing (communicative) roles by utilizing turn-taking, namely the sending order of the symbolic messages between the first and second senders [3]. In this paper, we examine the causes of this role division in the formation process by analyzing transfer entropy and using cognitive architecture modeling.

2 Experimental Method

A total of 21 pairs of human participants repeated a coordination task that comprised several rounds by using computer terminals from different sites (Fig. 1). In each round, a proxy marker for each participant was randomly allocated to a position in a 2×2 grid, with each partner unaware of the other's position. Each participant composed a symbolic message by combining two meaningless figures and sent it to his/her partner. This message was then immediately displayed on the partner's terminal. On this basis, participants moved one step horizontally or vertically, or stayed in the same position, in order to attempt to reach the same position as their partners. After the movement phase, the initial and destination positions of both participants were displayed. The pair gained two points each if they reached the same position and lost one point each otherwise. The score started from zero points and remained non-negative throughout the experiment. This process was repeated until the score reached 50 points ("success") or 1 h elapsed ("failure"). Fourteen pairs of participants were successful, and seven pairs failed.

3 Results of the Transfer Entropy Analysis

Transfer entropy $T_{I \to J}$, a measure of information flow from a stochastic process I to another process J, is defined by

$$T_{I \to J} = \sum_{y_{n+1}, x_n, y_n} p\left(y_{n+1}, x_n, y_n\right) \log \left(\frac{p\left(y_{n+1} \middle| x_n, y_n\right)}{P\left(y_{n+1} \middle| y_n\right)} \right),$$

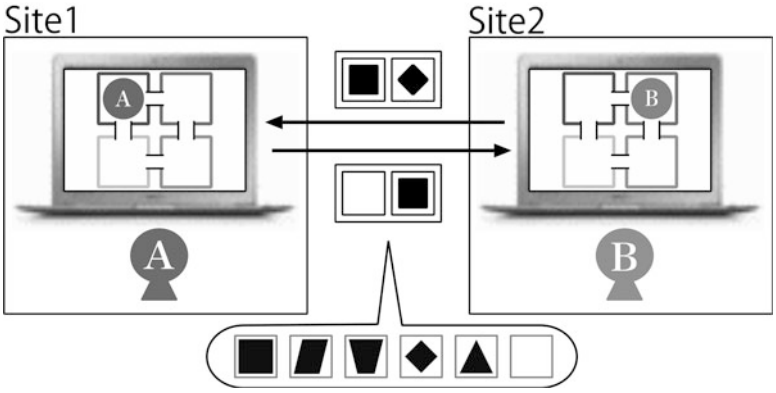


Fig. 1 Schematic view of the experimental setup

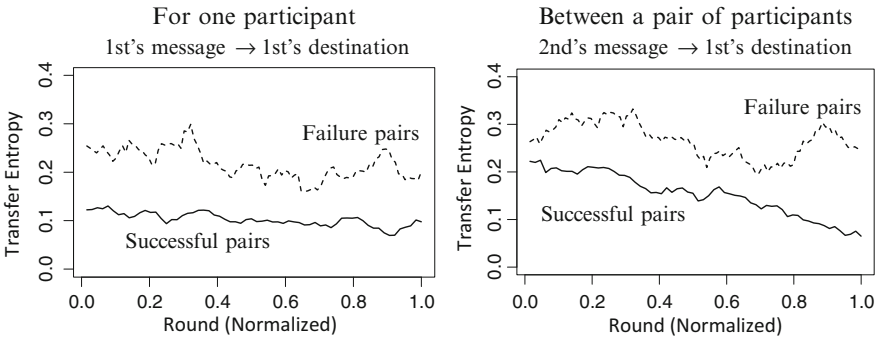


Fig. 2 Transitions of transfer entropy for one participant, from the first sender's message to the first sender's destination (*left*), and between participants, from the second sender's message to the first sender's destination (*right*). The *solid lines* represent successful pairs and the *dashed lines* failure pairs. The X-axis is the normalized round, since the number of rounds differ by pair

where x_n and y_n represent the stochastic sequence of processes I and J, respectively. Here, we analyzed this value based on turn-taking, where the first sender attempts to inform the initial position and the second sender attempts to indicate a destination position that both can reach.

The transitions of transfer entropy are shown in Fig. 2. The left-hand graph depicts transfer entropy for one participant, where I and J represent the sequences of the first sender's messages and his/her destinations, respectively. This value for successful pairs (solid line in Fig. 2 left) stayed at a low level from the beginning of the experiment, while that for the failure pairs (dashed line in Fig. 2 left) fluctuated in the mid range. By contrast, transfer entropy between a pair of participants (solid line in Fig. 2 right), where I is the sequence of the second sender's messages and J is that of the first sender's destinations, began moderately and declined over time for successful pairs, but did not decrease from this average level for the failure pairs (dashed line in Fig. 2 right). Further, transfer entropy for the successful pairs negatively correlated with the performance of the task.

These results imply that the correspondence between the messages and movements of one participant for successful pairs was consistent from the beginning of the experiment. Moreover, successful pairs could also make the information flow between a pair of participants ordered (i.e., certain and predictable), which suggests that they were able to incorporate their partners' behavior adequately.

4 Simulation Analysis Using ACT-R

In the next step, the experiment was simulated by using the cognitive architecture ACT-R (Adaptive Control of Thought-Rational) [4]. ACT-R makes a decision according to IF-THEN-type production rules [5]. It thus integrates symbolic learning and sub-symbolic learning by reinforcing the usage of these rules. This feature

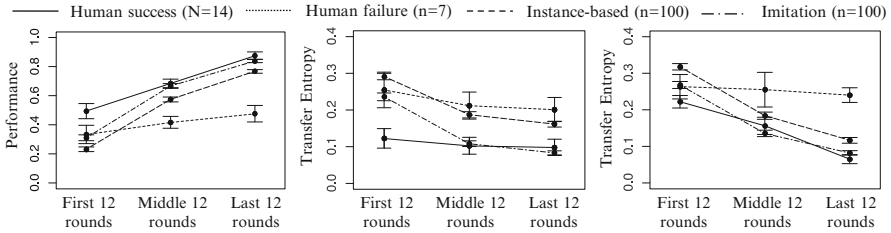


Fig. 3 Comparison between the ACT-R simulation and human experiment data for performance (*left*), transfer entropy for one participant (*middle*), and that for a pair of participants (*right*). The rounds are divided into three parts (first, middle, and last 12 rounds) in order to make the number of rounds uniform

is suitable to model our experiment, since both types of learning approaches are used to establish common ground and symbol systems in communication formation processes.

The participant model used herein stored successful decisions about messaging and movements. We then compared two strategies that differed in terms of influence over the decision-making process: Instance-Based (i.e., the use of one's own decisions) and Role Reversal Imitation (i.e., the use of the partner's decisions by reversing roles in addition to the first strategy). We introduced the imitation strategy because successful pairs in the experiment tended to accept partner's symbol use in their making process of symbol systems. In acceptance of partner's system, participants recalled the partner's behavior in the same situation and thereby took the same decision.

Figure 3 compares the ACT-R simulation and human experiment data. We found that the imitation strategy replicated the results of the human experiment in terms of performance (i.e., based on the agreement rate of destinations in 12 rounds) as well as transfer entropy for one participant and between a pair of participants. All three indices (performance and transfer entropies) show that the imitation strategy is more similar to the successful pairs than the instance-based strategy, which rather resembles the failure pairs.

5 Discussion and Conclusion

Both the analyses presented in this paper suggest that a successful participant behaves consistently from the initial stage by storing and utilizing his/her own successful decisions as well as using his/her partner's decisions (imitation) based on adequately stored experiences. We showed that incorporating the partner's behavior through role reversal imitation facilitated the establishment of role division and ordered the information flow between partners.

The higher performance and lower transfer entropy of human data compared with the imitation strategy of ACT-R in the first stage suggest that human participants

have an initial disposition (bias) in their behavior that contributes to establishing common ground. The ACT-R simulation started by demonstrating random behavior with no bias in the use of symbols and movement. However, the reinforcement of successful rule selection in the imitation strategy led to the formation of shared symbol systems.

Acknowledgments The authors thank Ichiro Tsuda and Yutaka Yamaguchi for their valuable discussions. This work was supported by a MEXT KAKENHI on Innovative Areas (No. 21120011 in No. 4103).

References

1. Galantucci, B. An experimental study of the emergence of human communication systems. *Cognitive Science*, **29** (2005) 737–767.
2. Konno, T., Morita, J. & Hashimoto, T. Symbol communication systems integrate implicit information in coordination tasks. *Advances in Cognitive Neurodynamics (III)*, Y. Yamaguchi (ed.) Springer (2013) pp. 453–460.
3. Konno, T., Morita, J. & Hashimoto, T. Process and factor of establishing role division for communicating connotations (in preparation)
4. Morita, J., Konno, T. & Hashimoto, T. The role of imitation in generating a shared communication system. *Proc. of the 34th Annual Meeting of the Cognitive Science Society*, N. Miyake et al. (eds.) (2012) pp. 779–784.
5. Anderson, J.R. *How can the Human Mind Occur in the Physical Universe?* Oxford University Press (2007).

Optical Imaging of Plastic Changes Induced by Fear Conditioning in Auditory, Visual, and Somatosensory Cortices

Yoshinori Ide, Muneyoshi Takahashi, Johan Lauwereyns, Minoru Tsukada, and Takeshi Aihara

Abstract Second-order conditioning using light, tone and foot-shock was carried out, and conditioned responses to tone and light could be observed by monitoring the cardiac pulse. Three sensory cortical areas including auditory, somatosensory, and visual cortices in the same animal were simultaneously recorded by using optical imaging with voltage sensitive dye, RH795. Cortical activities to each stimulus including foot-shock, light, and tone were compared in naïve versus conditioned animals. As a result, while only primary and secondary somatosensory cortices were activated to foot-shock alone in a naïve animal, after the activation of primary and secondary somatosensory cortices, neural activities were propagated to auditory and visual cortices in conditioned animals. Our findings illustrate the integration of three different modalities through second-order conditioning. Through this integration, sensory stimulation in one modality can lead to the retrieval of associated, but presently absent sensory information in two other modalities.

Keywords Auditory cortex • Visual cortex • Somatosensory cortex • Plasticity • Second-order conditioning • Optical imaging • Voltage sensitive dye

Y. Ide (✉) • M. Tsukada
Brain Science Institute, Tamagawa University, 6-1-1 Tamagawagakuen, Machida, Tokyo 194-8610, Japan
e-mail: yide@lab.tamagawa.ac.jp

M. Takahashi • J. Lauwereyns
Brain Science Institute, Tamagawa University, 6-1-1 Tamagawagakuen, Machida, Tokyo 194-8610, Japan

Graduate School of Systems Life Sciences, Kyushu University, Fukuoka 819-0395, Japan

T. Aihara
Brain Science Institute, Tamagawa University, 6-1-1 Tamagawagakuen, Machida, Tokyo 194-8610, Japan

Graduated College of Engineering, Tamagawa University, 6-1-1 Tamagawagakuen, Machida, Tokyo 194-8610, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_64

1 Introduction

Sensory cortices are defined by responses to physical stimuli in specific modalities. Recently, however, human neuroimaging studies have shown auditory cortex activation without sound [1], and it has been reported that auditory cortical neurons can be activated and/or modified by non-auditory stimuli [2, 3]. Little is known about how the brain produces such activity. Here we study this topic with an entirely different paradigm, based on fear conditioning, using optical imaging [3]. Our data shows that foot shocks alone can activate the auditory and visual cortices of a guinea pig through second-order conditioning.

2 Methods

The experiments were performed in accordance with the guidelines of the Animal Experiments Committee of Tamagawa University. Three to four weeks old female guinea pigs were used as experimental subjects. Two groups (conditioning and naïve) were prepared. Second-order conditioning with light, tone, and foot-shock was carried out while the animal was awake in the test cage, through a grid shock floor. The conditioning protocols are shown in Fig. 1.

On day 1, pairing of tone (8 kHz, 5 s) and foot-shock (0.7 mA, 0.5 s) was repeated five times. On day 2, pre-test for light was carried out; then, pairing of light (white LED, 10 s) and tone (8 kHz, 5 s) was repeated five times. The conditioned response to light was tested on day 3. In order to measure the conditioned response, the cardiac pulse (blood flow) was recorded by using an ear sensor. Next, the cortical activities in naïve and conditioned animals were recorded by using optical imaging with a voltage sensitive dye, RH795. Responses from the left cortical areas, including auditory, visual, and somatosensory cortices, were recorded in

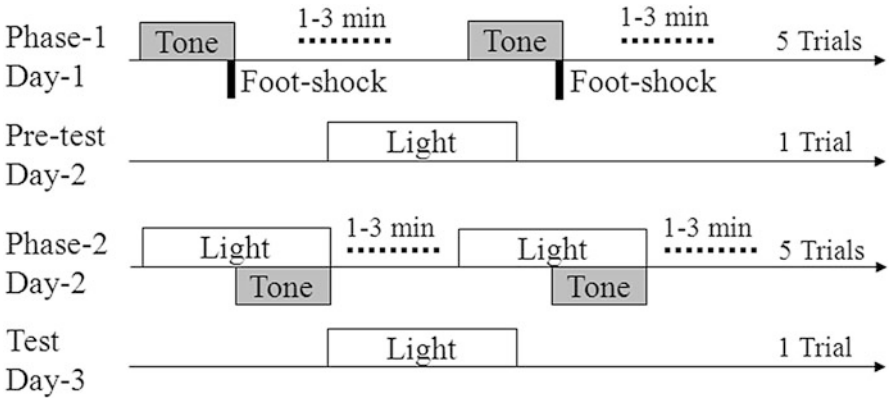


Fig. 1 Protocol of second-order conditioning using light, tone and foot-shock

three conditions: (1) presentation of a pure-tone (duration: 50 ms, frequency: 4–16 kHz, acoustic pressure: 65 dB SPL) to the right ear; (2) presentation of a light stimulus (duration 5 ms, white LED) to the right eye; and (3) application of an electric shock (duration: 0.2 s, current intensity: 0.5–1.5 mA) to the hind legs.

3 Results

Second-order conditioning using light, tone and foot-shock was carried out. On day 1, blood flow was monitored during the tone-shock conditioning; as a result, the conditioned response was observed in the third to fifth trial. On day 2, pairing of light and tone was repeated five times. Blood flow change to the light stimulus was monitored in the pre-test (day-2) and test (day-3); as a result, the conditioned response to light was observed after the light-tone conditioning, proving second-order conditioning to be successful. Next, neural activities in the auditory, visual, and somatosensory cortices were recorded by optical imaging. In order to identify the location of these three cortical areas, cortical activities in the recording sites to light, foot-shock, and tone stimuli were investigated. Figure 2a shows a schematic diagram of cortical areas including visual, somatosensory and auditory cortices as

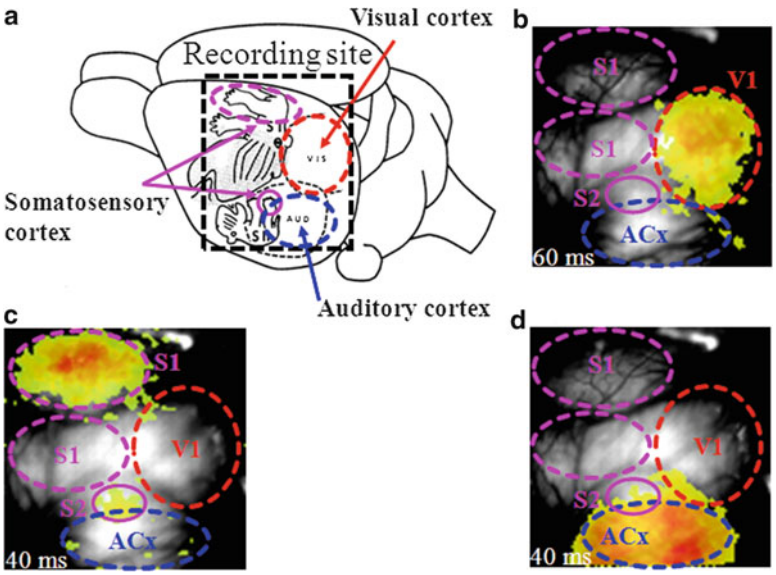


Fig. 2 (**a**) A schematic diagram of cortical areas including visual, somatosensory and auditory cortices. The areas surrounded by *red* and *blue dashed lines* are visual (V1) and auditory (ACx) cortices, respectively. The areas surrounded by *pink dashed* and *solid lines* are primary (S1) and secondary (S2) somatosensory cortices corresponding to the hind legs, respectively. (**b**) Neuronal activity in V1 to light stimulus. (**c**) Neuronal activity in S1 and S2 to foot-shock to hind legs. (**d**) Neuronal activity in ACx to tone stimulus

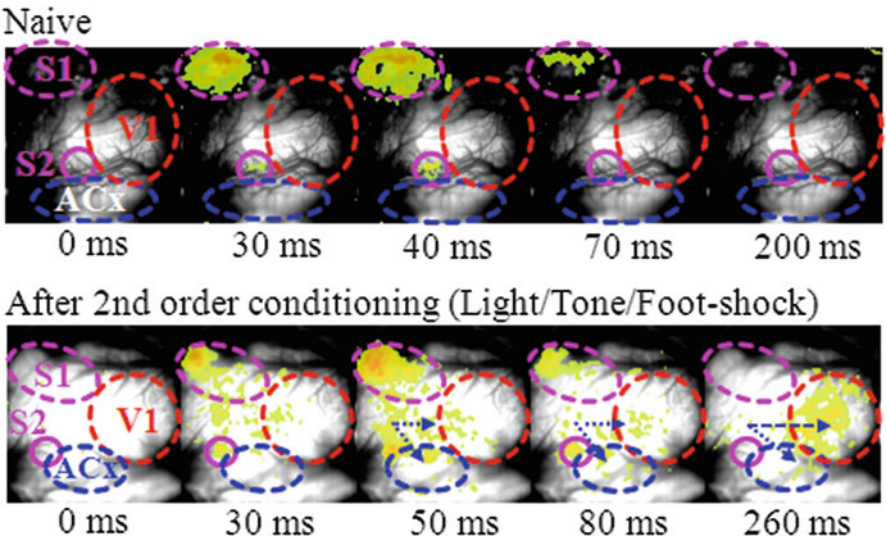


Fig. 3 Cortical activities in response to foot-shock alone in naïve and conditioned animals, respectively. *Foot-shock*: electric shock to hind legs. *S1*, *S2*: primary and secondary somatosensory cortex, *V1*: primary visual cortex, *ACx*: auditory cortex

reported by Woolsey and Van der Loos [4]. Figure 2b shows cortical activity to a light stimulus, suggesting that the activated area is primary visual cortex.

Figure 2c shows cortical activity to a foot-shock to the hind legs, suggesting that the activated areas are primary and secondary somatosensory cortices corresponding to the hind legs. Figure 2d shows cortical activity to a tone stimulus, suggesting that the activated area is primary auditory cortex. We confirmed that the configuration of visual, somatosensory, and auditory cortices shown in Fig. 2b–d corresponds to that in Fig. 2a. Finally, cortical activities to a foot-shock were compared in naïve versus conditioned animals. Figure 3 shows cortical activities in response to a foot-shock alone in naïve and conditioned animals, respectively. As a result, only S1 and S2 were activated to a foot-shock in naïve animals, without any propagation of neural activity. However, in conditioned animals, after the activation of S1 and S2, neural activities were propagated to the auditory (ACx) and visual (V1) cortices.

Acknowledgments This work was supported by Grant-in-Aid for Scientific Research on Innovative Areas 21120006 from MEXT in Japan.

References

1. Kraemer DJ., Macrae CN., Green AE., Kelley WM. Musical imagery: sound of silence activates auditory cortex. Nature 434 (2005) 158.
2. Brosch M., Selezneva E., Scheich H. Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkey. J Neurosci 25(29) (2005) 6797–6806.

3. Ide Y., Takahashi M., Lauwereyns J., Sandner G., Tsukada M., Aihara T. Fear conditioning induces guinea pig auditory cortex activation by foot shock alone. Cogn. Neurodyn. 7(2013) 67–77.
4. Woolsey TA. and Van der Loos H. The structural organization of layer IV in the somatosensory region (SI) of mouse cerebral cortex. The description of a cortical field composed of discrete cytoarchitectonic units. Brain Res. 17(2) (1970) 205–42.

Reward-Modulated Motor Information in Dorsolateral Striatum Neurons

Yoshikazu Isomura

Abstract It is well known that dorsolateral striatum neurons participate in the direct pathway (with dopamine D1 receptors) or the indirect pathway (without D1 receptors) to control voluntary movements. Little is, however, known about functional representation of motor information and its reward modulation in individual striatal neurons. In our recent study, we analyzed spike activity of single dorsolateral striatal neurons in head-fixed rats performing voluntary forelimb movement in a reward-expectable manner. Some of them were identified morphologically by a juxtacellular visualization combined with in situ hybridization for D1 mRNA. The striatal neurons showed distinct functional activations before and during the forelimb movement, regardless of the D1 mRNA expression. They were often modulated only positively by their reward expectation for the correct motor response. Our results suggest that the direct and indirect pathway neurons may play a cooperative rather than antagonistic role in spatiotemporal control of voluntary movements in the skeletomotor loop of the basal ganglia.

Keywords Striatum • Direct pathway • Dopamine receptor • Reward • Voluntary movement

1 Introduction

The basal ganglia are composed of the dorsal and ventral striatum, globus pallidus (external and internal segments), subthalamic nucleus, and substantia nigra. They connect with the cerebral cortex and thalamus to organize parallel cortico-basal ganglia thalamo-cortical loop circuits [1]. In rodents, the dorsolateral part of striatum receiving excitatory inputs from the primary motor cortex participates in the skeletomotor loop that contributes voluntary movements.

More than 90 % of striatal neurons are medium spiny neurons, which are the only cell type of striatal projection neurons, receiving glutamatergic inputs

Y. Isomura (✉)
Brain Science Institute, Tamagawa University, 6-1-1 Tamagawa-gakuen, Machida, Tokyo 194-8610, Japan
e-mail: isomura@lab.tamagawa.ac.jp

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_65

from the cortex and sending GABAergic projections to other parts of the basal ganglia [2]. The projection neurons in dorsolateral striatum participate in either the direct pathway (expressing dopamine D1 receptors) or the indirect pathway (D2 receptors), controlling voluntary movements in an antagonistic manner [2]. The D1- and D2-expressing neurons are excited and inhibited, respectively, by dopamine release from the substantia nigra neurons encoding a reward prediction error. People believe that an activation of the direct pathway neurons results in an enhancement of voluntary movements (like as a car accelerator), while an activation of the indirect pathway neurons results in a depression of them (a brake). However, it remains to be known how individual striatal neurons for the two pathways functionally represent motor information and how their motor information is modulated by reward expectation.

In our recent study [3], we juxtacellularly recorded the spike activity of single neurons in the dorsolateral striatum while rats were performing voluntary forelimb movements in a reward-expectable condition. We also recorded multi-neuronal activity in the motor cortex during the voluntary movements for comparison.

2 Methods

All experiments were carried out in accordance with the Animal Experiment Plan approved by the Animal Experiment Committee (RIKEN, Japan). Adult Long-Evans rats (150–250 g, male) were first handled to adapt to the experimental environments. A head-attachment was surgically attached to the skull under isoflurane gas anesthesia. After recovery from the surgery, they were deprived of drinking water in their home cages. Then, the rats were trained for 2 weeks, by an operant learning, to perform an alternate-reward forelimb movement task under a head-fixed condition, which was the same as used in the previous study [4, 5] except for an alternation of reward and no reward delivery. Briefly, the head-fixed rats had to start each trial by pushing a lever and hold it with the right forelimb for at least 1 s. After the lever hold period, if they pulled the lever spontaneously (with no cue signal), a high-tone sound was presented with a 0.2–0.8 s delay to notice that the response was correct in the trial. A small amount (10 ul) of saccharin water as a reward was or was not delivered alternately during the presentation of high-tone sound. Thus, they learned to anticipate reward or no reward in each trial. Once the rats completed the operant task learning, they were transferred to the recording set-up for electrophysiology.

We obtained juxtacellular recordings through a glass electrode from single neurons in the left dorsolateral striatum of individual rats performing the alternate-reward forelimb movement task. In some of the recorded neurons, we tried to electroporate Neurobiotin into them with positive current pulses typically for 15–20 min. The recorded neurons were visualized later for morphological identification in combination with in situ hybridization for D1 mRNA and immunohistochemistry for mu opioid receptor. We simultaneously recorded multi-neuronal

activity in the forelimb area of the left motor cortex through a 16ch silicon probe. The spike activity in striatum and motor cortex neurons was isolated and analyzed offline by our spike-sorting soft-ware EToS [6, 7].

3 Results

In our experiments, most of the rats (N = 63 of 67 rats) learned to perform the alternate-reward forelimb movement task well [3]. The onset of pull movements was apparently earlier in the rewarded trials than in the non-rewarded trials. The velocity of pull movements was not largely different in the rewarded and non-rewarded trials. Thus, the behavioral observations suggest that the rats successfully understood the alternate reward rule and anticipated the alternate delivery of reward and no reward for their correct task response.

We obtained juxtacellular recordings from 84 neurons in the dorsolateral striatum, while the rats were performing the alternate-reward forelimb movement task [3]. Regarding task-related discharge activity, the striatal neurons were classified into two major functional groups: Hold-type and Movement-type neurons showing functional activations before and during the forelimb movement, respectively. The Hold-type neurons (n = 21) showed lower spike rates during the push movement than during the hold period, whereas the Movement-type neurons (n = 46) increased spiking rates during both the push and the pull movements. We found the Hold- and Movement-type activities were often augmented by the reward-anticipatory condition (positive reward modulation). It is unlikely that the reward modulation was simply due to the behavioral difference in forelimb movement.

After the juxtacellular recording, we visualized a part of the recorded striatal neurons, and examined whether they expressed the mRNA for the dopamine D1 receptor by using a fluorescent in situ hybridization (Fig. 1). We attempted to visualize 57 neurons morphologically out of the 84 juxtacellularly (electrically)

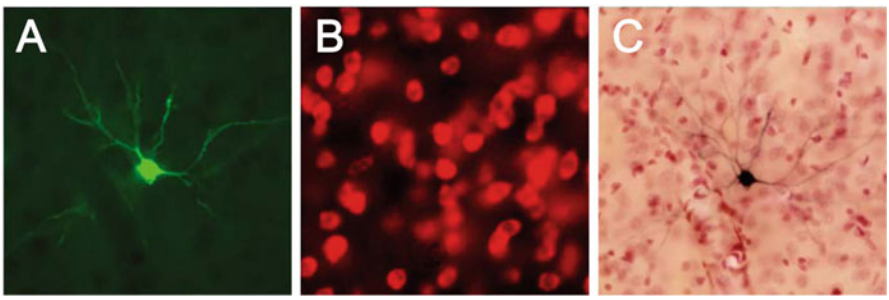


Fig. 1 A juxtacellularly recorded striatum neuron. (**a**) Fluorescent visualization for Neurobiotin-loaded neuron. (**b**) In situ hybridization for dopamine D1 receptor mRNA. (**c**) Its re-stained morphology

recorded neurons, and we obtained 27 successfully visualized neurons in the dorsolateral striatum where the FL area of the motor cortex densely innervates. We obtained 10 D1-positive neurons (Hold-type n = 5, Movement-type n = 5) and 17 D1-negative neurons (Hold-type n = 3, Movement-type n = 7, and others). Thus, the Hold- and Movement-type neurons were not specific to D1-negative or -positive subpopulation, respectively. Furthermore, the reward modulation of Hold-type and Movement-type neurons seemed independent of the expression of D1 mRNA.

Next, we examined whether the functional activity of motor cortex neurons was affected by the reward alternation, because the striatal reward information may be conveyed from the motor cortex rather than the substantia nigra. Using the multi-neuronal recording, we obtained a total of 216 neurons (164 regular-spiking (RS) and 52 fast-spiking (FS) neurons) in the FL area of the motor cortex during the task performance (N = 24 rats) [3]. We found 46 Hold-type and 106 Movement-type neurons in the motor cortex. Consistent with our previous study [4], many of the FS neurons displayed the Movement-type activity (n = 37) rather than the Hold-type activity (n = 1). Importantly, most of the motor cortex neurons were not clearly modulated by the reward alternation. The cortico-striatal pyramidal cells are located in the layers 5A and 5B of the rat motor cortex. We found no largely biased distribution in the reward modulation of Movement-type activity in the superficial (putatively layer 2/3), middle (layer 5), and deep (layer 6) layer subpopulations. It suggests that most of motor cortex neurons appeared be specialized to process motor information with no reward information, unlike the striatal neurons.

4 Discussion

As described above, we showed how motor information and reward information are represented by individual neurons in the dorsolateral striatum and the motor cortex of the skeletomotor loop in the rat. According to the classical basal ganglia model, the excitation of striatal neurons for the direct and indirect pathways, respectively, enhances and depresses the activity of motor cortical neurons antagonistically [1]. Recent studies using gene modification techniques certainly support this model [8, 9]. Therefore, one may expect that the direct pathway neurons increase the spike activity during voluntary movements, whereas the indirect pathway neurons do so during no movements. However, no large difference was found between D1-positive and -negative neurons in the dorsolateral striatum. Interestingly, Cui et al. (2013) reported that these two pathways are concurrently and similarly activated during voluntary movements [10]. Our results at a single-cell resolution level favor their observations.

Thus, the antagonistic pathway system may be accomplished by a more complex involvement of the two groups of striatal neurons. For example, for Hold-type activity, D1-positive (direct pathway) neurons may prepare for an intended movement or just maintain overall muscle tone, while D1-negative (indirect pathway) neurons

may suppress temporally incorrect expression of the movement. For Movement-type activity, D1-negative neurons may suppress concurrent expression of antagonistic muscular movements, while D1-positive neurons execute the intended movement. Taken together, this simple movement may actually be composed of several spatiotemporal motor components, and individual striatal neurons for both pathways may antagonistically code subprograms for the motor components to complete the whole movement coordinately.

The dorsal as well as ventral striatum is known to process reward-related or motivational information [11, 12]. The dopamine neurons of substantia nigra send reward prediction error signals to the striatal projection neurons. Dopamine enhances the activation of D1- expressing (direct pathway) neurons and depresses D2-expressing (indirect pathway) neurons. It is thus possible that the dopamine signals carrying reward information affect motor information represented by these projection neurons in the dorsal striatum. But our results showed a strong populational shift to the positive direction in the reward modulation of dorsal striatum neurons. Accordingly, the dorsal striatum neurons may possibly receive synaptic inputs carrying reward information from cortical (other than the motor cortex) or subcortical (non-dopaminergic) neurons, as well as the substantia nigra dopamine neurons, in the skeletomotor loop.

The two-pathway system of the basal ganglia is likely to work, not just antagonistically, but rather coordinately by various functional activations of individual striatal neurons that contribute to an integration of motor and reward information.

5 Conclusions

We have shown the functional spike activity of identified dorsolateral striatum neurons during voluntary movements and reward expectation.

1. Hold- and Movement-type neurons were not specific to D1-negative and -positive subpopulation, respectively. Therefore, It is not a simple set of an accelerator and a brake.
2. Reward modulation of Hold- and Movement-type activity seemed independent of D1 receptor expression. It may be mediated through dopaminergic as well as non-dopaminergic synaptic inputs.

Our findings will shed light on functional coordination of individual neurons in basal ganglia-related networks to accomplish goal-oriented behaviors.

Acknowledgments This work was supported by Grants-in-Aid for Young Scientists (A) (21680036); for Scientific Research on Priority Areas (20019035) and for Scientific Research on Innovative Areas (22120520, 24120715); and JST CREST, from the Ministry of Education, Culture, Sports, Science, and Technology of Japan. This paper is a short review on our recent publication [3] that was also presented in the Dynamic Brain Forum 2013.

References

1. Alexander, G. E. & Crutcher, M. D. Trends Neurosci 13 (1990) 266–271.
2. Gerfen, C. R. In The rat nervous system, 3rd ed., Elsevier (2004) 455–508.
3. Isomura, Y. et al. J Neurosci 33 (2013) 10209–10220.
4. Isomura, Y. et al. Nat Neurosci 12 (2009) 1586–1593.
5. Kimura, R. et al. J Neurophysiol 108 (2012) 1781–1792.
6. Takekawa, T. et al. Eur J Neurosci 31 (2010) 263–272.
7. Takekawa, T. et al. Front Neuroinform 6 (2012) 5
8. Kravitz, A. V. et al. Nature 466 (2010) 622–626.
9. Hikida, T. et al. Neuron 66 (2010) 896–907.
10. Cui, G. et al. Nature 494 (2013) 238–242.
11. Goldstein, B. L. et al. J Neurosci 32 (2012) 2027–2036.
12. Stalnaker, T. A. et al. J Neurosci 32 (2012) 10296–10305.

Anxiolytic Drugs and Altered Hippocampal Theta Rhythms: The Quantitative Systems Pharmacological Approach

Tibin John, Tamás Kiss, Colin Lever, and Péter Érdi

Abstract The spirit of systems pharmacology was adopted to study the possible mechanisms of anxiolytic drugs on hippocampal electric patterns. The purpose of this study was to investigate the mechanisms by which anxiolytics characteristically reduce the slope and/or intercept of the stimulus-frequency relationship of hippocampal theta. A network of neuron populations that generates septo-hippocampal theta rhythm was modeled using a compartmental modeling technique. The effects of cellular and synaptic parameters were studied. Results show that halving the rate of the rise and fall of pyramidal hyperpolarization-activated (I_h) conductance lowers intercept of nPO elicited theta frequency relationship, that increasing the decay time constant of inhibitory post-synaptic current can reduce the frequency of low nPO stimulation elicited theta rhythm, and that maximal synaptic conductance of GABA-mediated synapses has little effect on frequency within the septo-hippocampal CA1 network.

Keywords Systems pharmacology • Anxiolytic drugs • Hippocampus • Theta rhythms • Population neural network

T. John (✉)
Center for Complex Systems Studies, Kalamazoo College, Kalamazoo, MI, USA
e-mail: Tibin.John11@kzoo.edu

T. Kiss
Neuroscience Research Unit, Pfizer Global Research and Development, Pfizer Inc, 700 Main Street, Cambridge, MA 02139, USA

Wigner Research Centre for Physics, Institute for Particle and Nuclear Physics, Hungarian Academy of Sciences, Budapest, Hungary

C. Lever
Department of Psychology, Durham University, Durham, UK

P. Érdi
Center for Complex Systems Studies, Kalamazoo College, Kalamazoo, MI, USA

Wigner Research Centre for Physics, Institute for Particle and Nuclear Physics, Hungarian Academy of Sciences, Budapest, Hungary

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_66

1 Introduction

Our main aim here is driven by the repeated observation that anxiolytic drugs disrupt hippocampal theta. It is now clear that a neurochemically very-wide range of anxiolytic drugs reduce the frequency of reticular-elicited hippocampal theta in the anaesthetised rat [1, 2]. Such theta-frequency reduction remains the best preclinical (i.e. animal-based) test for predicting a clinically-efficacious anxiolytic drug. Conversely, drugs which are antipsychotic or sedative but not anxiolytic do not reduce reticular elicited theta frequency [1]. The utility of the theta-frequency reduction assay in predicting clinical efficacy of anxiolytic drugs suggests to us the importance of creating a quasi-realistic biophysical network modeling framework within which to model changes in hippocampal theta frequency.

To model changes in hippocampal theta frequency resulting from changes in conductance and decay time constants of specific currents, we extended our previous computational model [3], which was constructed to to simulate the generation and pharmacological control of septohippocampal theta rhythm. Our goals here were (i) to simulate the effect of the stimulating current delivered to the nPO on the frequency of hippocampal theta oscillation, (ii) to extend our previous model with a more biologically realistic septal innervation where we account for tonic and rhythmic cholinergic and glutamatergic, and rhythmic GABAergic inputs, and (iii) to study how anxiolytic drugs of diverse chemical structure modify the synaptic parameters in order to reduce theta frequency. The current paper describes our initial results in meeting these aims.

2 Methods

A biophysically detailed compartmental modeling technique was adopted to build a network of pyramidal cell and certain inhibitory cell populations.

Model specifications including parameter settings and full equation specifications are given at http://geza.kzoo.edu/theta/theta.html and are used without change unless otherwise specified. In order to model the recently verified role of the glutamatergic population of the medial septum in contributing to theta activity [4], a rhythmic spike-based input modulated at theta frequency was implemented to represent the effect of this population within the septo-hippocampal system (Fig. 1). To simulate an assay that reliably predicts the clinical efficacy of anxiolytics, electrical stimulation to the nPO region of the brainstem was modeled as depolarizing current to pyramidal, basket, and medial septal GABAergic (MS-GABA) neurons, as well as by the frequency of the rate modulation of the medial septal glutamatergic (MS-Glu) spiking object.

Septo-hippocampal theta rhythm was recorded as a population activity measured as the extracellular field potential around pyramidal cells, the most abundant neurons in the relevant neural tissue. The mean theta-range frequency of the LFP

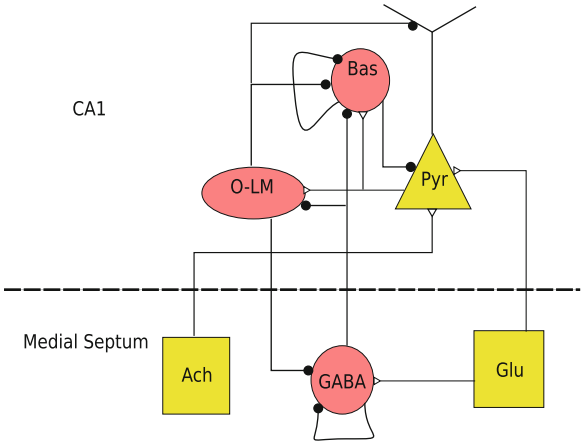


Fig. 1 Structure of septo-hippocampal network model. Three hippocampal neuron populations and one medial septal population are modeled explicitly. Two medial septal populations are modeled indirectly by their rhythmic or tonic effects on the system (*squares*). *Red symbols* and *black circles* indicate inhibitory populations and synapses, respectively; *yellow symbols* and *open triangles* indicate excitatory populations and synapses

was determined as the frequency corresponding to the peak of the power spectrum in the range 2–12 Hz. The time-dependence of synaptic conductances was modeled as following first-order kinetics according to the equation

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \alpha F\left(V_{\mathrm{pre}}\right)(1-s) - \beta s$$

with s representing the fraction of open synaptic channels at the post-synaptic membrane, weighted by a maximal conductance g_{syn} to evaluate post-synaptic current. The multiplicative inverse of the β term in this synaptic model is referred to as the decay time constant of the post-synaptic current, τ_{syn} [5].

See this webpage for the R, Octave, and GENESIS code used to generate the simulation experiments and data analysis: http://geza.kzoo.edu/theta/thetaFreq.html. ANCOVA followed by Tukey contrasts was used to compare the intercepts of nPO stimulation elicited theta frequency-stimulus relationships before and after potential anxiolytic effects were applied.

3 Results

The LFP generated by pyramidal neurons exhibited increasing frequency and amplitude of oscillation in the 3–10 Hz theta range (Fig. 2a). The frequency of synchronous firing amongst pyramidal cells also increased with increased nPO

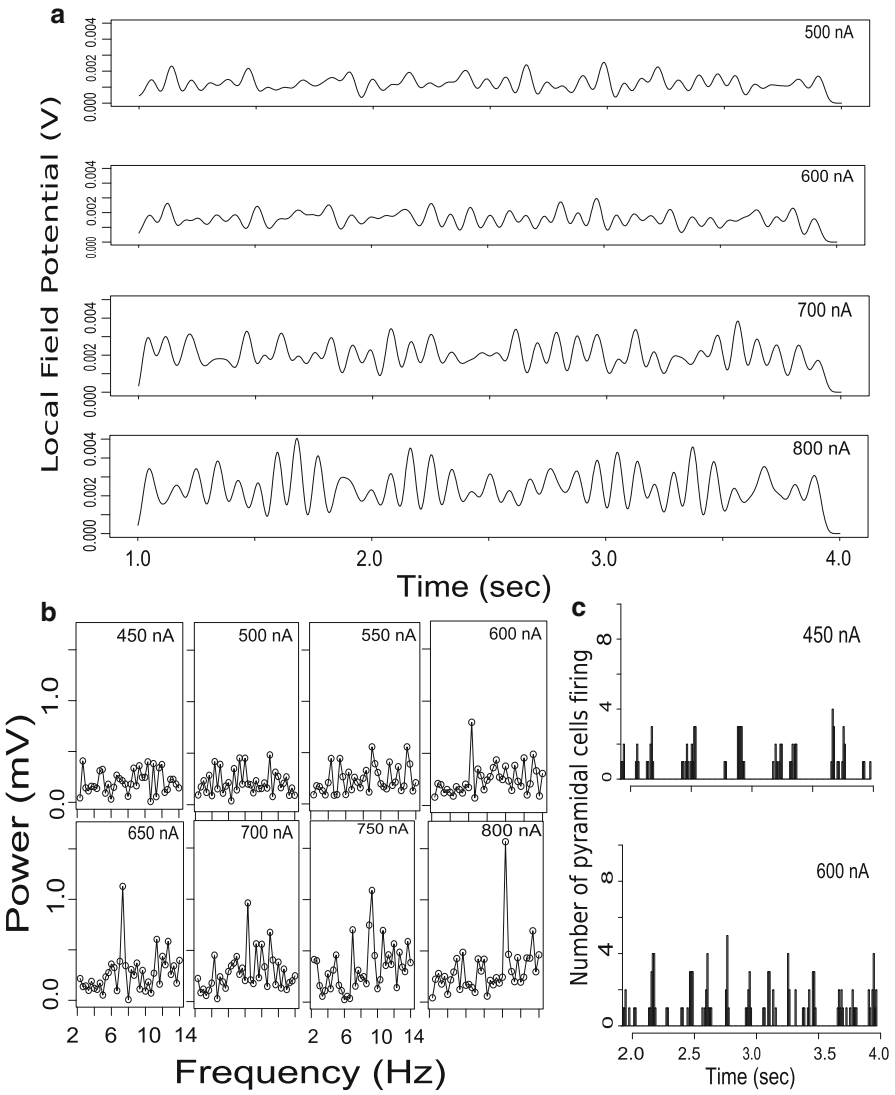


Fig. 2 (**a**) Synthetic LFP recordings from pyramidal cells for increasing nPO stimulation in model. (**b**) Representative power spectra of LFP in the theta band for increasing nPO stimulation. (**c**) Representative firing histograms showing increasing frequency of synchronous firing in pyramidal somata

stimulation (Fig. 2c). An approximately linear increase in mean theta frequency with nPO stimulus level was observed (Fig. 3; black).

The decay time constant of all GABA-receptor mediated synapses was doubled to test the effect on elicited theta frequency. A significant interaction effect between nPO stimulus level and decay time constant of GABA synapses was observed

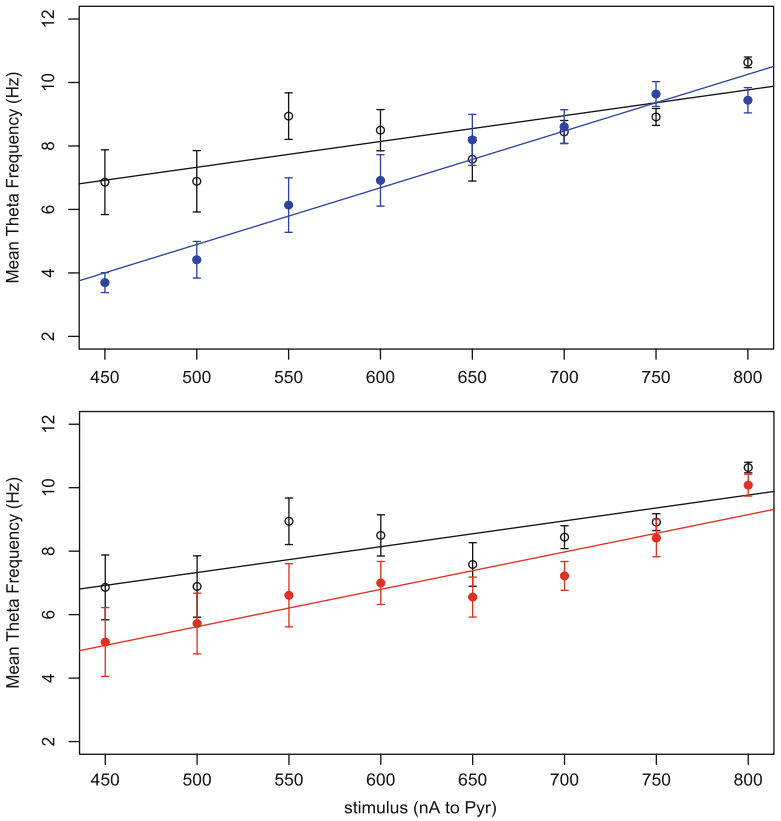


Fig. 3 Modulation of synaptic dynamics affects frequency of nPO stimulation elicited theta rhythm. Mean frequency and standard error are shown with default parameter settings (*black*) and with GABA decay time constant doubled (*blue*). Slowing down pyramidal hyperpolarization-activated (I_h) current dynamics slightly lowers intercept of nPO elicited theta frequency relationship. Mean theta frequency and standard error are shown for a range of nPO stimulation levels with default parameter settings (*black*) and with I_h conductance rise and fall rates cut in half (*red*). Both assays were conducted for N = 12 runs of the simulation per point

($F = 11.80, df = 1, 188, p < 0.001$, Fig. 3) suggesting an increased slope of the stimulus-frequency relationship when the time constant was doubled. The intercept of the regression line was also significantly reduced ($t = -4.053, p < 0.0001$) when the decay time constant was doubled.

No significant difference in mean network frequency was detected between any level of positive allosteric modulator action ($F = 0.833, df = 7, 80, p = 0.563$, Fig. 4), suggesting that these drugs may operate outside the septo-hippocampal system when modulating the frequency of septo-hippocampal theta rhythm.

When the rates at which I_h conductance rises and falls in the somata and dendritic regions of pyramidal cells in the model were each halved, reduction in the intercept of the stimulus-frequency relationship was observed, although this trend

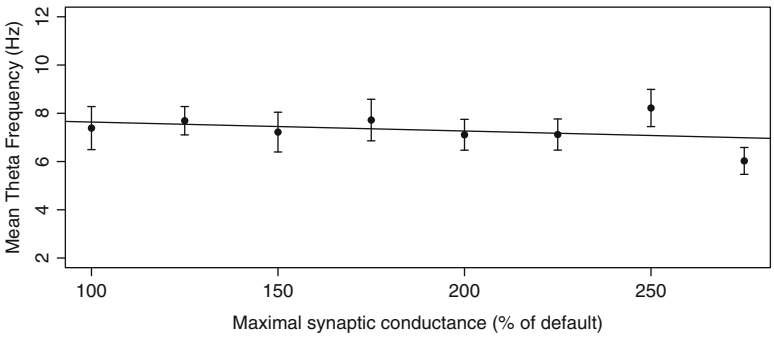


Fig. 4 Modulation of maximal synaptic conductance associated with GABA receptors within septo-hippcampal system has negligible effect on theta frequency. *Error bars* indicate standard error of the mean of N = 12 runs of the simulation Depolarizing currents of 600 nA, 1.4, and 2.2 μA to pyramidal somata, basket cells, and MS-GABA cells was used, respectively, with a MS-Glu population firing rate modulation of 6 Hz

did not quite reach significance ($t = -1.772, p = 0.078$, Fig. 3). No significant interaction between the categories and the level of nPO stimulation was observed ($F = 1.351, df = 1, 188, p = 0.247$), suggesting no difference in slope between the relationships.

4 Discussion

It has been demonstrated that the frequency of hippocampal theta rhythm can be reduced by modulating synaptic and intrinsic cellular parameters in a biophysically realistic model of the septo-hippocampal system. Research goals obtained include modeling nPO stimulation assay with increasing depolarizing currents to evaluate potential effects of anxiolytic drugs, expanding septo-hippocampal model to begin to account for the role of medial septal glutamatergic population in hippocampal theta production, and identifying potential effects of selective anxiolytics including slowing the decay and onset of hyperpolarization-activated current in CA1 pyramidal neurons, producing intercept reduction. It is interesting to note that a recent study showing that the blockade of the I_h channel resulted in slowing of nPO stimulated hippocampal theta oscillation and also that this drug has anxiolytic effects [6]. Similarly, our modeling results also suggest that decreasing the influence of the I_h channel on the evolution of the membrane potential would result in lower theta intercept. Increasing synaptic decay time constant of inhibition relative to that of excitation lowers theta frequency at some stimulation levels and might be combined with other mechanisms of frequency reduction to have more reliable anxiolytic effects. This result is comparable to the effects of zolpidem, which increases the synaptic time constant of $GABA_A$ receptors [7] and is known to have

weak anxiolytic effects. Our model also predicts that the frequency-modulating effects of positive allosteric modulators of $GABA_A$ receptors on hippocampal theta is mediated by synapses outside of the septo-hippocampal system. This prediction is consistent with evidence suggesting that the reduction of reticular-elicited hippocampal theta by systemic benzodiazepine injection is largely or wholly mediated by medial hypothalamic sites. For instance, chlordiazepoxide infusions to the medial supramammillary nucleus mimic those of systemic injections of chlordiazepoxide [8].

Acknowledgements PÉ is benefited to be a Fellow of the Institute of Advanced Studies at Durham University. He also thanks to the Henry Luce Foundation for general support. TK is a full time employee and shareholder of Pfizer Inc. Some research reported in this paper was supported by a BBSRC New Investigator grant (BB/G01342X/1 and X/2) to Colin Lever.

References

1. J. Gray and N. McNaughton, *The neuropsychology of anxiety*. Oxford university press, 2000.
2. N. McNaughton, B. Kocsis, and M. Hajós, "Elicited Hippocampal Theta Rhythm: a Screen for Anxiolytic and Procognitive Drugs through Changes in Hippocampal Function?," *Behavioural Pharmacology*, no. 5–6, pp. 329–346, 2007.
3. M. Hajós, W. E. Hoffmann, G. Orbán, T. Kiss, and P. Erdi, "Modulation of septo-hippocampal Theta activity by GABAA receptors: an experimental and computational approach," *Neuroscience*, vol. 126, no. 3, 2004.
4. C. Y. L. Huh, R. Goutagny, and S. Williams, "Glutamatergic neurons of the mouse medial septum and diagonal band of Broca synaptically drive hippocampal pyramidal cells: relevance for hippocampal theta rhythm," *The Journal of neuroscience: the official journal of the Society for Neuroscience*, vol. 30, no. 47, 2010.
5. X. J. Wang and G. Buzáski, "Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model," *Ó The Journal of Neuroscience*, vol. 15, no. 20, 1996.
6. M. Yeung, C. T. Dickson, and D. Treit, "Intrahippocampal infusion of the ih blocker ZD7288 slows evoked theta rhythm and produces anxiolytic-like effects in the elevated plus maze," *Hippocampus*, vol. 23, no. 4, 2013.
7. N. Hájos, T. F. F. Z Nusser, E A Rancz, and I. Mody, "Cell type- and synapse-specific variability in synaptic GABAA receptor occupancy," *The European Journal of Neuroscience*, vol. 12, no. 3, 2000.
8. M.-A. Woodnorth and N. McNaughton, "Similar effects of medial supramammillary or systemic injection of chlordiazepoxide on both theta frequency and fixed-interval responding," *Cognitive, Affective & Behavioral Neuroscience*, vol. 2, no. 1, 2002.

Carbachol-Induced Neuronal Oscillation in Rat Hippocampal Slices and Temperature

Itsuki Kageyama and Kiyohisa Natsume

Abstract Θ and β oscillations are functional and can be involved in the memory and learning. Temperature affects the processing. It was reported that the specific temperature range is required for induction of theta oscillations. However, affects of temperature on the other neural oscillations have not been clarified so far. In the present study, we examined the impact of temperature on the carbachol-induced oscillations. Carbachol-induced oscillations were observed at the temperature 27–35 °C. At 27 °C, the frequency covered from θ to β range. From 29 to 35 °C, the frequency was in the β range. The frequency was significantly increased with the temperature. It was correlated with the temperature. Above 35 °C, the oscillations were not observed. Instead, the epileptic discharges were induced. Next, the spatio-temporal pattern of theta oscillation and epileptic discharges in the slices was measured. The current-source density (CSD) analysis clarified that theta oscillation and epileptic discharges had the different CSD patterns.

Keywords Neuronal oscillations • Memory and learning • Hippocampus • Temperature dependence • Epileptic discharges

1 Introduction

The hippocampus of a rat in vivo has various oscillations, θ, β oscillation. These oscillations are functional and can be involved in the memory procressing [1]. Theta (<12 Hz), β (12–30 Hz) oscillations similar to the in vivo oscillations can be induced in hippocampal slices with the application of a cholinergic agent carbachol

I. Kageyama (✉)
Graduate School of Life Science and Systems Engineering, Department of Brain Science and Engineering, Kitakyushu, Japan
e-mail: kageyama-itsuki@edu.brain.kyutech.ac.jp

K. Natsume
Graduate School of Life Science and Systems Engineering, Department of Brain Science and Engineering, Kitakyushu, Japan

RCBT, Kyushu Institute of Technology, Kitakyushu, Japan
e-mail: natume@brain.kyutech.ac.jp

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_67

[2]. Fluctuations in brain temperature affect on memory process [3]. It is reported that the specific temperature range is required for the induction of θ oscillation. However, effects of temperature on the other neural oscillations have not been clarified so far. In the present study, we examined the impact of temperature on the carbachol-induced oscillations. In the recorded temperature range, carbachol-induced θ oscillation (CITO) was induced. We also recorded the spatio-temporal pattern of the oscillation, CITO and picrotoxin-induced epileptic discharges (PIED), using multi-electrode array about.

2 Matherials and Methods

2.1 Slice Preparation

Hippocampal slices were prepared from male Wistar rats (45–160 g) aged 3–6 weeks. The rats were anesthetized with isoflurane and then decapitated. Their brains were removed quickly and cooled in artificial cerebrospinal fluid (ACSF) solution at 0 °C. Transverse hippocampal slices, which were 400–450 μm thick, were prepared using a tissue slicer. The slices were incubated in a preserving chamber at 33 ± 1 °C for an hour applying ACSF solution bubbled with 95 % O_2 and 5 % CO_2.

2.2 Recording Protocol

Neural oscillation was recorded from the *stratum pyramidale* to the *stratum radiatum* of the CA3 region using a glass microelectrode filled with 2 M NaCl (1–2 MΩ). To induce neural oscillations, acetylcholine receptor agonist carbachol was applied. After an hour, a hippocampal slice was transferred to the recording chamber, which was set to a temperature of either 27 ± 1, 29 ± 1, 31 ± 1, 33 ± 1, 35 ± 1, 37 ± 1, or 39 ± 1 °C. In order to make sure that the hippocampal slice temperature remained constant, we measured the temperature of the slices continuously using the needle-thermocouple probe and thermometer. Multi-electrode array system (MED64) of Alpha MED Scientific Inc., Ltd. was used to record the spatio-temporal pattern of field potential of CITO and PIED, which were induced by a acetylcholine receptor agonist carbachol and a $GABA_A$ receptor antagonist picrotoxin.

3 Result

3.1 Temperature Dependence of Carbachol-Induced Neuronal Oscillations

The administration of carbachol to hippocampal slices, neuronal oscillations as shown in Fig. 1 were induced in the recording at the temperature of 27–35 °C. In this ranges, the burst-like neuronal oscillations were induced. The frequency of the oscillation in a burst ranged from 9.8 to 13.7 Hz at 27 °C. It ranged from θ to β frequency range. On the other hand, the frequencies in a burst were from 12.7 to 17.6 Hz at temperature of 29–35 °C. The frequency was in β frequency range. The frequency was significantly increased with temperature (***$p < 0.001$; ANOVA with Krusal-Wallis test). In addition, the duration and inter-burst interval (IBI) were significantly decreased with temperature (***$p < 0.001$; ANOVA with Krusal-Wallis test). The significant changes of the amplitude with temperature was not observed.

3.2 Different Spatio-Temporal Patterns of CITO and PIED

CSD analysis was performed using the field potentials recorded from 64 electrodes of MED64. During the generation of CITO, a cluster of a pair consisted of a source and a sink made a line, and propagated from dentate gyrus to CA1 along the *stratum*

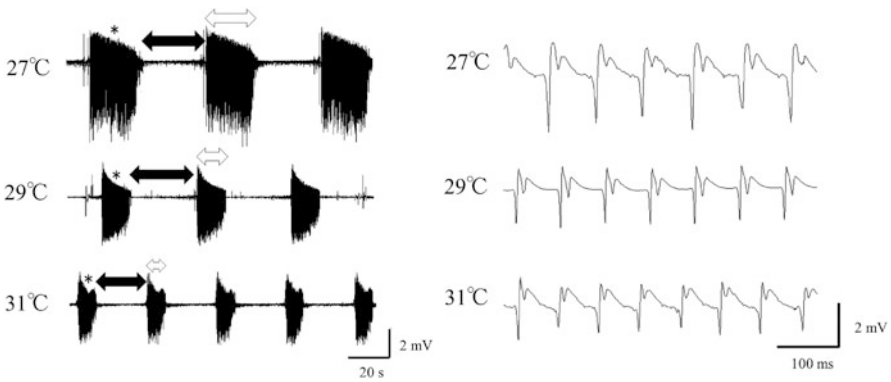


Fig. 1 Representatives of carbachol-induced neuronal oscillations at temperature of 27–35 °C. Perfusion of 30 μM carbachol induced the bursts of the oscillation. The *left* trace reveals successive bursts of neuronal oscillation with the regular burst duration (*white arrow*) and IBI (*black arrow*). The *right* trace expands the asterisk portion of the *left* trace

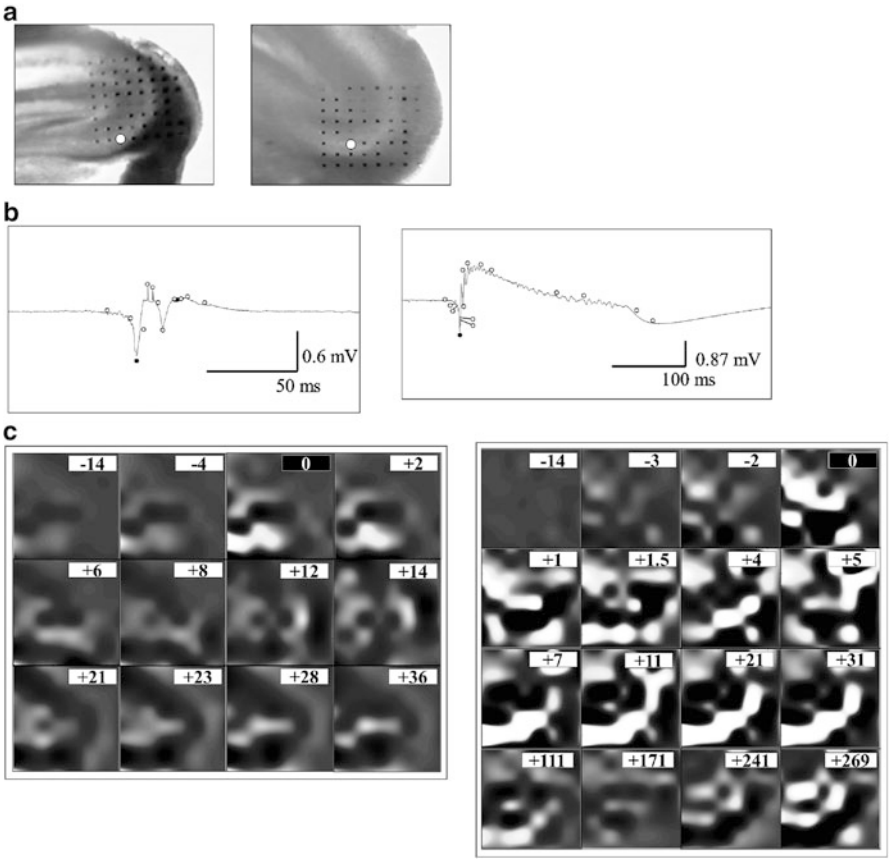


Fig. 2 The result of CSD analysis for CITO (*left column*) and PIED (*right column*). (**a**) The pictures of the hippocampal slices on the MED probe. *Black small points* show electrodes. (**b**) Typical field potential of CITO and PIED from an electrode of MED. CITO and PIED were measured at *white circles* shown in (**a**). (**c**) A time course change of the spatio-temporal patterns of sinks and sources in one cycle of a wave shown in (**b**). The time when the most negative potential was induced is defined as 0 msec. Spatio-temporal patterns of sinks and sources are shown at the time indicated at the tics shown in (**b**)

pyramidale, a pyramidal cell layer in CA3 (Fig. 2). The cluster diminished after the propagation. Then the source covered the entire cell layer. On the other hand, in the PIED, a larger cluster of a sink and source pair was induced suddenly, it propagated faster than CITO, and it lasted for a longer time (Fig. 2).

4 Discussion

There was difference in the spread of a cluster of sinks and sources between CITO and PIED. In the case of CITO, that of sinks and sources started from the neighborhood of dentate gyrus of hippocampal CA3, and it propagated along a pyramidal cell layer. Afterwards, the source covered the whole pyramidal cell layer. On the other hand, in the case of PIED, the small cluster of sinks and sources were induced, it suddenly merged to propagate along the whole CA3 pyramidal cell layer. Afterwards, the source covered the whole pyramidal cell layer for a long time. After the source disappeared, the sink covered the same area for a long time. CITO changed with temperature. Therefore, in the future, the spatio-temporal patterns of sinks and sources will be recorded with different temperature.

Acknowledgements This work was supported by JSPS KAKENHI Grant Numbers 24120712, and by a Grant-in-Aid for Scientific Research on Innovative Areas (Comprehensive Brain Science Network) from the Ministry of Education, Science, Sports and Culture of Japan..

References

1. Varela F, Lachaux JP, Rodriguez E, Martinerie J.,"The brainweb: phase synchronization and large-scale integration.", Nat Rev Neurosci., Vol. 2(4), pp.229–39 (2001).
2. Lubenov EV, Siapas AG.,"Hippocampal theta oscillations are travelling waves", Nature, Vol. 459(7246), pp. 534–9 (2009).
3. Shimono K, Brucher F, Granger R, Lynch G, Taketani M.,"Origins and distribution of cholinergically induced beta rhythms in hippocampal slices.", J Neurosci., Vol. 20(22), pp. 8462–73 (2000).

Associative Memory Network with Dynamic Synapses

Yuichi Katori, Yosuke Otsubo, Masato Okada, and Kazuyuki Aihara

Abstract We investigate the dynamical properties of an associative memory network with dynamic synapses whose connection strength changes with short-term plasticity (STP) mechanism. We use a network model composed of stochastic neurons and dynamic synapses and a corresponding mean field model. In the present study, we focuses on influences of the facilitation synapses whose connection strength is strengthened by the STP. We report that the facilitation synapses contribute to the stabilization of memory association and that the bifurcation analysis on the mean field model clarifies the stability in its macroscopic dynamics. Furthermore, the synaptic facilitation causes several types of transitive dynamics among multiple memory patterns.

Keywords Associative memory • Dynamic synapses • Short-term plasticity • Stochastic neurons • Mean field model

1 Introduction

In an associative memory network composed of mutually connected neurons, memory patterns are stored in their synaptic connections, and the network retrieves a stored memory pattern according to its network dynamics [1]. In conventional

Y. Katori (✉)
FIRST, Aihara Project, JST, 4-6-1 Komaba, Meguro-ku, Tokyo, Japan

Institute of Industrial Science, The University of Tokyo, Tokyo 153-8505, Japan
e-mail: katori@sat.t.u-tokyo.ac.jp

Y. Otsubo
Graduate School of Frontier Science, The University of Tokyo, Kashiwa, Chiba 277-8561, Japan

M. Okada
Graduate School of Frontier Science, The University of Tokyo, Kashiwa, Chiba 277-8561, Japan

RIKEN Brain Science Institute, Hirosawa, Wako, Saitama 351-0198, Japan

K. Aihara
Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_68

associative memory model, the strength of the connections are assumed to be static. However, recent physiological studies revealed that the strength of the synaptic connections changes in a short period of time with short-term plasticity mechanism; these synapses are called dynamic synapses [6]. The strength of synaptic connection decreases (depression synapse) or increases (facilitation synapse). The synaptic facilitation contributes flexible information representation in the prefrontal cortex [2]. Dynamical properties of networks with stochastic neurons and dynamic synapses have been intensively investigated [3–5]. The associative memory network with the dynamic synapses shows variety of dynamical states including the memory retrieved state and the transitive state between stored memory patterns [4]. In the present study, we explore further detailed dynamical properties of the associative memory network.

2 Model

The associative memory network we used here [4] is composed of N binary neurons, and the neurons are connected via the dynamic synapses. The state of the ith binary neuron is denoted by the variable $s_i(t)$ and takes an active state $[s_i(t) = 1]$ or a resting state $[s_i(t) = 0]$. Changes in the synaptic transmission strength is determined by the fraction of releasable neurotransmitters $x_i(t)$ and the utilization parameter $u_i(t)$ [6]. The state of the neuron and the dynamic synapses changes according to the following equations [4]:

$$\mathrm{Prob}[s_i(t+1) = 1] = \frac{1}{2}\left(1 + \tanh[\beta h_i(t)]\right), \tag{1}$$

$$x_i(t+1) = x_i(t) + \frac{1 - x_i(t)}{\tau_R} - s_i(t)x_i(t)u_i(t), \tag{2}$$

$$u_i(t+1) = u_i(t) + \frac{U_{se} - u_i(t)}{\tau_F} + U_{se}(1 - u_i(t))s_i(t), \tag{3}$$

where $h_i(t) = \sum_{j\neq i}^{N} J_{ij}[2s_j(t)x_j(t)u_j(t)/U_{se} - 1]$ represents the total input to the ith neuron and $1/\beta = T$ represents the noise intensity. The quantity J_{ij} represents the absolute strength of the connection from the jth to ith neuron; the memory patterns are stored in this connections. U_{se} represents the steady state value of the variable $u_i(t)$. The strength of synaptic transmission is given by the product of $x_j(t)$ and $u_j(t)$; the strength decreases (depression) or increases (facilitation) depending on the parameters τ_R, τ_F, and U_{se}.

The associative memory network is implemented with following absolute strength of synaptic connection $J_{ij} = \frac{1}{N}\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu$ where $J_{ii} = 0$. The p memory patterns $\xi^\mu = (\xi_1^\mu, \cdots, \xi_N^\mu), \mu = 1, \cdots, p, \xi_1^\mu \in \{-1, 1\}$ are given by the following correlated patterns. Suppose that a parent memory pattern ξ,

which satisfy Prob$[\xi_i = \pm 1] = 1/2$. The memory patterns ξ^μ are given by Prob$[\xi_i^\mu = \pm 1] = (1 \pm b\xi_i)/2$, where b is the correlation level among memory patterns and takes values in the interval $[0, 1]$.

To analyze the macroscopic properties of the associative memory network, we consider the following macroscopic mean field model that captures overall dynamical properties of the network [4].

$$m_{\boldsymbol{\eta}}(t+1) = g_\beta \left(\sum_{\boldsymbol{\eta}'} p_{\boldsymbol{\eta}'} \boldsymbol{\eta} \cdot \boldsymbol{\eta}' \left(2m_{\boldsymbol{\eta}'}(t) X_{\boldsymbol{\eta}'}(t) U_{\boldsymbol{\eta}'}(t)/U_{se} - 1 \right) \right), \tag{4}$$

$$X_{\boldsymbol{\eta}}(t+1) = X_{\boldsymbol{\eta}}(t) + \frac{1 - X_{\boldsymbol{\eta}}(t)}{\tau_R} - m_{\boldsymbol{\eta}}(t) X_{\boldsymbol{\eta}}(t) U_{\boldsymbol{\eta}}(t), \tag{5}$$

$$U_{\boldsymbol{\eta}}(t+1) = U_{\boldsymbol{\eta}}(t) + \frac{U_{se} - U_{\boldsymbol{\eta}}(t)}{\tau_F} + U_{se}(1 - U_{\boldsymbol{\eta}}(t)) m_{\boldsymbol{\eta}}(t), \tag{6}$$

where $\boldsymbol{\eta}$ indicates p-dimensional pattern vectors $\boldsymbol{\eta} = (\eta^1, \ldots, \eta^p)^T \in \{-1, 1\}^p$. A set of neurons $\{1, \ldots, N\}$ is divided into 2^p groups on the basis of these pattern vectors. Suppose that $\bar{\boldsymbol{\xi}}_i = (\xi_i^1, \ldots, \xi_i^p)^T \in \{-1, 1\}^p$, a sublattice is defined as a set of neurons belonging to a given pattern vector. The sublattice belonging to the pattern vector $\boldsymbol{\eta}$ is defined as $\mathcal{I}_{\boldsymbol{\eta}} = \{i | \bar{\boldsymbol{\xi}}_{\mathbf{i}} = \boldsymbol{\eta}\}$. $p_{\boldsymbol{\eta}} = |\mathcal{I}_{\boldsymbol{\eta}}|/N$ denotes the relative sublattice size.

3 Results

We analyze the dynamical properties of the associative memory network by numerical simulations and the mean field model. In the present study, we commonly use following parameter values: $b = 0.2$, $U_{se} = 0.1$, $N = 10^4$, and $p = 3$.

Process of the memory retrieval can be characterized as a convergence of the state of the network to an attractor. In the pseudo-constant range in which the effect of STP quickly disappears ($\tau_R = 4$), the state of the network tends to converges to a steady state. As the effects of facilitation increases (as τ_F increases), the pseudo-memory states appear, and then the memory retrieved states are stabilized. The bifurcation analysis on the mean field model clarifies the qualitative changes in this stability as shown in Fig. 1a.

In the depression-dominant range ($\tau_R = 7$), the network tends to be unstable. Increases in effects of facilitation (increases in τ_F) causes variety of oscillatory states (Fig. 1b–d) including the transitive state among the memory patterns. In the mean field model, the transitive states appear as the pseudo-periodic orbit. The orbit distributes on the several types of manifolds: circle, torus, and strange attractor like structure.

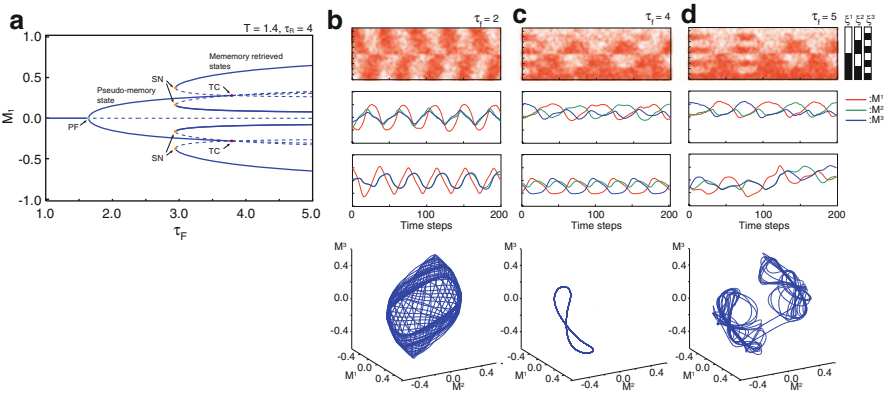


Fig. 1 (**a**) Bifurcation diagram in the pseudo-constant range. (**b–d**) Typical time courses of the network response in the depression-dominant range. In the *top row*, *dots* indicate the active state of the neuron. Remaining rows indicate the overlaps, which quantify the similarity between the state of the network and the stored patterns in the stochastic model (*second row*) and in the mean field model (*third row* and *bottom*)

4 Conclusions

The associative memory network with dynamic synapses exhibits a variety of dynamical states: memory and pseudo-memory states, as well as transitive dynamics among memory patterns. The facilitation synapses have significant influences on the network dynamics. These results imply that the dynamic synapses potentially contribute various brain functions, e.g. synaptic working memory and memory retrieval including sequential memory association. These functional aspects of the network dynamics should be investigated in the future.

Acknowledgements This research was supported by the Aihara Project, the FIRST program from PSPS and JSPS Grant-in-Aid for Young Scientist (B) Grant Number 25730175.

References

1. Hopfield, J.J., : Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. USA, **79**, 2554–2558, (1982)
2. Katori, Y., Sakamoto, K., Saito, N., Tanji, J., Mushiake, H., Aihara, K.: Representational Switching by Dynamical Reorganization of Attractor Structure in a Network Model of the Prefrontal Cortex, PLoS Comp. Biol, **7**,11, e1002266, (2011)
3. Katori, Y., Igarashi, Y., Okada, M., Aihara, K.: Stability Analysis of Stochastic Neural Network with Depression and Facilitation Synapses, J. Phys. Soc. Jpn., **81**, 114007, (2012)
4. Katori, Y., Otsubo, Y., Okada, M., Aihara, K.: Stability analysis of associative memory network composed of stochastic neurons and dynamic synapses, Front. Comp. Neurosci, **7**,6, 1–12, (2013)

5. Mejias, J.F., Torres, J.J.: Maximum memory capacity on neural networks with short-term synaptic depression and facilitation, Neural Comp., **21**, 3, 851–871, (2009)
6. Tsodyks, M., Pawelzik, K., Markram, H.: Neural networks with dynamic synapses, Neural Comp., **10**, 4, 821–835, (1998)

Neural Dynamics for a Sudden Change in Other's Behavioral Rhythm

Masahiro Kawasaki, Keiichi Kitajo, Kenjiro Fukao, Toshiya Murai, Yoko Yamaguchi, and Yasuko Funabiki

Abstract Our communication includes behavioral synchronization with sudden changes of other's rhythms. However, it is not clear about its neural mechanism. Here, we compared the behaviors and brain activities between the normal and autism spectrum disorder (ASD) subjects in an alternating tapping task which required subjects to synchronize the tapping intervals with a virtual person (PC program) who tapped at a constant interval or a variable interval. Behavioral results showed no significant difference between the normal and ASD subjects under the constant-interval condition. In contrast, the rates of synchronization of the normal subjects were higher than those of the ASD subjects under the variable-interval condition. EEG results showed alpha modulations (10–12 Hz) in the motor cortex were larger for the normal subjects than the ASD subjects under the variable condition, whereas there was no difference under the constant condition. Our results suggested that the synchronization with other's sudden variable rhythms would be associated with the motor alpha modulations.

M. Kawasaki (✉)
Rhythm-based Brain Information Processing Unit, RIKEN BSI-TOYOTA Collaboration Center, 2-1, Hirosawa, Wako-shi, Saitama 351-0198, Japan

Department of Intelligent Interaction Technology, Graduate School of Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi, Ibaraki 305-8573, Japan
e-mail: kawasaki@iit.tsukuba.ac.jp

K. Kitajo
Rhythm-based Brain Information Processing Unit, RIKEN BSI-TOYOTA Collaboration Center, 2-1, Hirosawa, Wako-shi, Saitama 351-0198, Japan

PRESTO, Japan Science and Technology Agency (JST), 4-1-8, Honcho, Kawaguchi, Saitama 332-0012, Japan

K. Fukao • T. Murai • Y. Funabiki
Department of Psychiatry, Graduate School of Medicine, Kyoto University, 54 Kawaharacho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan

Y. Yamaguchi
Neuroinformatics Japan Center, RIKEN Brain Science Institute, 2-1, Hirosawa, Wako-shi, Saitama 351-0198, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_69

Keywords EEG • Synchronization • Rhythm • Motor • Autism

1 Introduction

Our communication needs synchronization with other's behavioral rhythms [1]. For example, for listening to music, hand clapping is spontaneously synchronized among audiences. In another example, our speaking speeds often synchronized with partners. Thus the behavioral rhythms of different individuals are well known to be spontaneously synchronized through social interactions [2]. Moreover, recent studies have identified the neural mechanisms (e.g. mirror neuron system [3], mu-rhythm [4], and inter-brain synchronization [5, 6]) for the behavioral synchronization with others in verbal and non-verbal communication.

Although such other's behavioral rhythms sometimes include a sudden change, its neural mechanisms remains unclear. To address the issue, it is useful to compare the behaviors and brain activities between the normal and autism spectrum disorder (ASD) subjects in coordinating with other's behavioral rhythms. The latter group of ASD is typically characterized by communication difficulties and stereotyped behaviors [7]. Here, we conducted the alternating tapping task which required subjects to synchronize the tapping intervals with a virtual person (PC program) who tapped at a constant interval or a variable interval, and compared the behavioral and electroencephalograph (EEG) rhythms between normal and ASD subjects.

2 Method/Models

Twenty-four normal subjects (12 females; 25.5 ± 6.7 years) and 24 ASD subjects (ten females; 29.2 ± 7.2 years) took part in the tasks with their eyes closed. The ASD was evaluated by the MSPA (Multi-dimensional scale for PDD and ADHD) and ADOS (Autism Diagnostic Observation Schedule). The intelligence quotient (IQ) which was evaluated by WAIS-III (Wechsler Adult Intelligence Scale), was not different between control (111.00 ± 12.19) and ASD subjects (111.33 ± 11.21).

In the task, when one subject tapped key with right index finger, one sound ("do" or "mi") was presented as the subject's tapping through both right and left earphones (Fig. 1). After that, the other sound ("mi" or "do") was presented as a virtual person's (PC's) tapping. If a difference between time intervals of previous PC's tapping (from subject to PC) and current subject's tapping (from PC to subject) was small (threshold: 50 ms), one-octave high sound was presented. Subjects were asked to tap the key with equal time interval of previous PC's tapping whereas the tapping rhythms were not instructed.

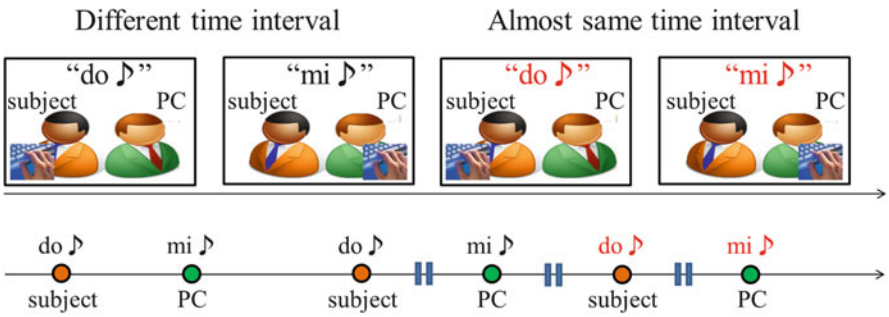


Fig. 1 Schematic illustration of the alternate tapping task. *Red-colored* words mean one-octave high sounds

Each subject completed 500 trials in two conditions; the constant condition (the intervals from the subject's tapping to the PC tapping were 600 ms) and the variable condition (the intervals were from 400 to 800 ms in five steps every 100 trials; e.g. 600 ms-400 ms-600 ms-800 ms-600 ms).

The EEG was recorded with 27 active scalp electrodes and the BrainAmp ExG MR equipment. The sampling rate was 1,000 Hz. Reference electrodes were placed on the right and left ear lobes. The independent components analyses components which were significantly correlated with the EOG were eliminated from the data as the artifact involved in eye blinks and movements [8]. And further, to reduce the effect of the spreading cortical currents due to volume conductance producing voltage maps covering large scalp areas, current source density conversion was performed. To identify the time-frequency amplitudes during the tasks, we applied wavelet transforms assuming that Morlet's wavelet function of 7-cycles length, with frequency ranging from 1 to 40 Hz (1-Hz steps).

3 Results

According to behavioral results, both the normal and ASD subjects showed high rates of synchronization with PC's tapping intervals (i.e. number of presentation of one octave high sound) under the constant condition. On the other hand, under the variable condition, the rates of synchronization of the normal subjects were higher than those of the ASD subjects (Fig. 2).

The EEG results showed the enhancements of alpha amplitudes (10–12 Hz) in the left motor cortex at the subject's tapping and the alpha decrements at the PC's tapping. Interestingly, the alpha modulations in the motor cortex were larger for the normal subjects than the ASD subjects under the variable condition, whereas there was no difference under the constant condition.

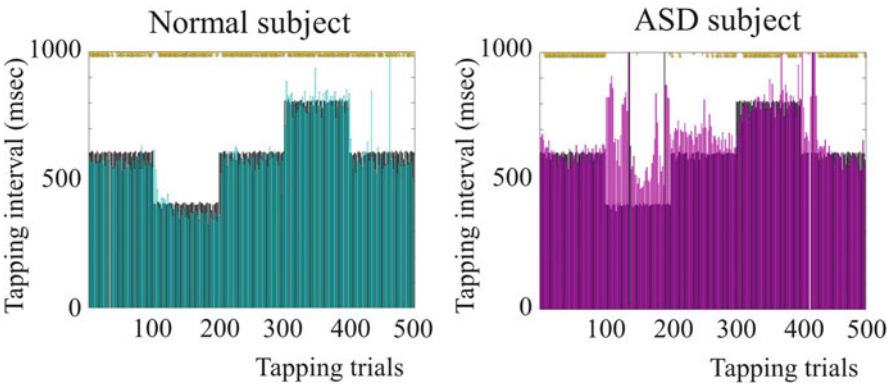


Fig. 2 Example tapping intervals under the variable condition (*cyan* healthy subject, *magenta* ASD subject, *black* PC). *Yellow dots* indicate trials which different intervals are within thresholds (50 ms)

4 Discussion

Along with the behavioral results, the brain activities were different between the normal and ASD subjects when the PC's rhythms were suddenly changed but not when the rhythms were fixed. Our results suggested that the synchronization with other's sudden variable rhythms would be associated with the motor alpha modulations. The motor alpha activities are known to be associated with the mirror neuron system [4]. Future study should address the causal relationships between the behavioral rhythms and the EEG rhythms in coordinating with sudden changes of other's rhythms.

Acknowledgments The research was supported by a Grant-in-Aid for Scientific Research on Innovative areas (21120005 and 24120706 and 25119512), a Grant-in-Aid for Young Scientists (B) (23700328) to MK. The authors would like to thank Tadao Mizutani and Miho Yoshizumi for their support in data acquisition and Yohei Yamada for their support in data analyses.

References

1. Kelso, J.A.S. Dynamic Patterns: The Self-Organization of Brain and Behavior. Cambridge: MIT Press (1995).
2. Hasson, U. et al. Intersubject synchronization of cortical activity during natural vision. Science**303**, 1634 (2004).
3. Iacoboni, M., Woods, RP., Brass, M., Bekkering, H., Mazziotta, JC., Rizzolatti, G., Cortical mechanisms of human imitation. Science 286, 2526–2528 (1999).
4. Tognoli, E., Lagarde, J., de Guzman, G. C. & Kelso, J. A. The phi complex as a neuromarker of human social coordination. Proc. Natl. Acad. Sci. USA **104**, 8190–8195 (2007).

5. Yun, K., Watanabe, K., Shimojo, S. Interpersonal body and neural synchronization as a marker of implicit social interaction. Sci. Rep. 2:959 (2013).
6. Kawasaki, M. Yamada, Y., Ushiku, Y., Miyauchi, E., Yamaguchi, Y. Inter-brain synchronization during coordination of speech rhythm in human-to-human social interaction. Sci. Rep. 2:959 (2013).
7. Funabiki Y, Kawagishi H, Uwatoko T, Yoshimura S, Murai T. Development of a multi-dimensional scale for PDD and ADHD. Res Dev Disabil. 32(3):995–1003 (2011).
8. Kawasaki, M., Kitajo, K. & Yamaguchi, Y. Dynamic links between theta executive functions and alpha storage buffers in auditory and visual working memory. Eur. J. Neurosci. **31**, 1683–1689 (2010).

Active Behaviors in Odor Sampling Constrain the Task for Cortical Processing

Leslie M. Kay, Daniel Rojas-Líbano, and Donald Frederick

Abstract Sensory perception is accomplished by means of active behaviors that help to extract information from the external environment. These behaviors become a part of the percept and also constrain the perceptual task. In olfactory perception, sniffing is the means by which individuals acquire olfactory stimuli from the environment. Rats sniff at 8–10 Hz during odor sampling, but each sniff has a different and stereotyped pattern that serves to find needed information. Higher flow sniffs are used to identify high sorption odors within mixtures, while lower flow sniffs are used to find lower sorption odors. Extended sniffing bouts (mean of 300–600 ms) are also stereotyped and tuned by the context in which rats identify odors. These sniffing bouts may be determined by the cognitive demands of the task or by particularities associated with training.

Keywords Olfactory perception • Sniffing • Sorptiveness • Odor sampling • Odor discrimination

L.M. Kay (✉)
Department of Psychology, The University of Chicago, Chicago, IL, USA

Committee on Neurobiology, The University of Chicago, Chicago, IL, USA

Institute for Mind and Biology, The University of Chicago, Chicago, IL, USA
e-mail: LKay@uchicago.edu

D. Rojas-Líbano
Committee on Neurobiology, The University of Chicago, Chicago, IL, USA

Institute for Mind and Biology, The University of Chicago, Chicago, IL, USA

Instituto de Ciencias Biomédicas, Universidad de Chile, Santiago, Chile

D. Frederick
Department of Psychology, The University of Chicago, Chicago, IL, USA

Institute for Mind and Biology, The University of Chicago, Chicago, IL, USA

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_70

1 Introduction

Cortical sensory systems have been shown to employ various strategies to represent and process sensory information. In many systems precise temporal structure in neuronal firing patterns produce oscillations at the population level that occur in the process of discriminating or recognizing stimuli. Neural oscillations have been ascribed many functions by neuroscientists, from binding of diverse sensory features to attention to pattern separation. In the rat olfactory system, neurons in the olfactory bulb respond with a collective fast oscillation in the gamma frequency range (40–100 Hz) when an individual sniffs an odor in many contexts [1]. We assume that all the precision is accounted for by the neural activity in these systems, but we often fail to acknowledge that animals in any waking behavioral state engage in volitional and goal-directed behavior while actively seeking stimuli. These behaviors adjust the incoming sensory information according to the animal's behavioral or cognitive goal, and they may be used to simplify or constrain the task for neurons in the brain. Behavioral effects on stimulus acquisition are well established in the visual arena, where investigators employ fixation strategies to normalize the sensory input. Here we summarize recent studies that address sniffing strategies in rodents; instead of normalizing out the sniff we describe emerging behavioral motifs for olfactory discrimination behavior. We address dynamics in behavior at the level of the single sniff and the sniffing bout. Single sniffs in the 7–10 Hz range have a period of 100–140 ms, while sniffing bouts composed of two to five sniffs generally range from 200 to 700 ms. Characterizing these highly dynamic behaviors is essential for understanding how odors are perceived.

2 Sorption and Sniffing Behavior

2.1 The Chromatographic Hypothesis

When odorant molecules pass over two media, such as air and mucus (as happens in the nose), they distribute between them. The sorptiveness of an odorant, quantified by the partition coefficient, describes this distribution. Besides this coefficient, the airflow at which the odorant passes over the mucus also determines how much odorant is absorbed. Faster flows cause fewer molecules to be absorbed compared to slower flows. In addition, areas of the sensory nasal epithelium that normally sustain lower airflows express a broad class of receptors that may bind low sorption odorants [2, 3]. Work by Mozell and colleagues showed that the olfactory epithelium functions as a kind of chromatograph, in which odorants strike preferentially at different zones of the epithelium depending on their sorptiveness [4, 5]. Given the abovementioned role of airflow in the odorant distribution, it has been suggested that animals may manipulate airflow through changes in sniffing to direct odorants to the epithelium zone where they are best detected [6].

2.2 Testing the Hypothesis (Methods)

We tested the chromatographic hypothesis in an experiment designed to challenge rats to identify an odorant within a binary mixture [7]. Because rats are extremely good at detecting odors, we wanted to test two rat groups, where both had to sniff the very same odor mixtures but each had to search within each mixture for specific odorants (that we called targets) of different sorptiveness for each group. We produced two sets of four monomolecular odorants, and in each set there were two high sorption and two low sorption odorants. Two groups of eight rats were trained to discriminate two novel training odors in a go/no-go task (sniff in a central odor port, then perform a nose poke in a second port for one odor, withhold responding for the other odor). After training, one rat group was assigned to find a high sorption target odor in the six binary mixtures formed from the odor set, and the other rat group was assigned to search for a low sorption odor in the same mixtures. They learned the target odor at the beginning of the session, and then they responded to randomly presented mixtures or the target odor. We repeated this for the second set of four odors, but the rat groups had their targets reversed (low sorption for the previously high sorption group, and vice versa). In addition to behavior, we also recorded EMG activity from the diaphragm muscle for each rat. This allowed us to estimate the relative airflow used for different sniffs, odors and days.

2.3 Testing the Hypothesis (Results)

We found that rats that sought the high sorption odor in either odor set learned faster than those seeking the low sorption odor. The former group learned to identify the mixtures containing that odor within the first or second session, while those tracking the low sorption odor took three sessions to learn. In the third session, when rat groups showed equal performance on the task, the group seeking the high sorption target odor sniffed with higher flow rates than those seeking the mixtures containing the low sorption target odor. The high and low flow rates were higher and lower than the flow rates that the two groups used on the first day. These data support the hypothesis that rats manipulate their sniff strategies to take advantage of the chromatographic properties of the sensory organ [7]. Another laboratory tested the hypothesis around the same time and produced opposite results [8]. The major difference between the two studies was their use of monomolecular odorants and our use of mixtures.

3 Sampling Bouts

While rats can identify an odor with information from a single sniff, they usually use more than one sniff to perform an odor discrimination. Taking additional sniffs may lead to better performance, but this result is subject to some controversy

[7, 9–12]. We have found that rats do increase performance with more sniffs, but there are bounds on this improvement. Two sniffs of the odor are better than one, but beyond that, it depends on the odor target. For low sorption odors in binary mixtures, there was an increase in performance (percent correct) even for the third sniff [7]. This was not true for the high sorption odors in mixtures. However, rats in different studies spent different amounts of time sniffing the odors, and there is near consensus that, while rats (and mice) can sometimes do better by sampling longer, they generally do not [7, 11–13].

The range of mean sampling times for odor discrimination is 250–600 ms, and it is unclear why the means are so different for different studies and laboratories. Within a laboratory, sampling times are relatively stable around a mean, but across laboratories the sampling times vary widely. This leads us to conclude that factors in training or the structure of the task encourage shorter or longer sampling times. Sampling time measures include the time an odor takes to arrive at the animal's nose, which varies depending on the distance from the odor source to the sampling port and the airflow in the olfactometer (usually 50–200 ms). Assessments of true sampling times subtract this delay and estimate the number of sniffs that contain odor. Most studies find that with one sniff of an odor rats and mice can perform above chance with all but the most difficult odor discrimination problems, and laboratories agree that with one more sniff performance can increase somewhat. With sniffs averaging 7–10 Hz, and the first sniff not assumed to contain any odor, then the minimum sampling time is approximately 200 ms, including the odor delay. In our laboratory, sampling times are 300–700 ms, (means of 400–450 ms). We do not restrict the amount of time that rats can sample odors, and we find that some of them sample for many seconds on some trials. Other laboratories restrict the maximum time of odor presentation to 1 or 2 s, and this factor might restrict the amount of time that rats choose to spend sampling an odor [9–11].

With an odor sampling bout lasting a few 100 ms, the remaining question is whether all sniffs are alike. We found that they are not alike. In the study in which we examined the role of sorptiveness, we found that the first two sniffs were similar to each other and much faster than later sniffs [7]. The first two sniffs could be as fast as 10 Hz, while later sniffs could be as slow as 7 Hz. We also found that inhalation times varied across sniffs, and exhalation times did not. So, rats played with their inhalation durations across sniffs in a stereotyped fashion even as they adjusted the sniff across sessions while learning to discern an odor within the mixtures.

4 Conclusion

Behavioral strategies constrain the type of information that the brain receives from a sensory organ. In olfaction, stereotyped sampling strategies define a sniffing bout for rats and mice. Fast sniffs predominate during odor sampling, but even within bouts the shape and speed of sniffs vary. Sampling strategies are adjusted as subjects

learn to detect odorants within a background. Thus, it is important to know the behavioral parameters under which a stimulus is processed when hypothesizing neural processing modes that may be used by an animal.

References

1. Rojas-Líbano, D., and Kay, L. M. (2008). Olfactory system gamma oscillations: the physiological dissection of a cognitive neural system. Cogn. Neurodyn. *2*, 179–194.
2. Bozza, T., Vassalli, A., Fuss, S., Zhang, J.-J., Weiland, B., Pacifico, R., Feinstein, P., and Mombaerts, P. (2009). Mapping of class I and class II odorant receptors to glomerular domains by two distinct types of olfactory sensory neurons in the mouse. Neuron *61*, 220–233.
3. Johnson, B. A., and Leon, M. (2007). Chemotopic odorant coding in a mammalian olfactory system. J. Comp. Neurol. *503*, 1–34.
4. Mozell, M. M., and Jagodowicz, M. (1973). Chromatographic separation of odorants by the nose: retention times measured across in vivo olfactory mucosa. Science (80). *181*, 1247–1249.
5. Mozell, M. M., Kent, P. F., and Murphy, S. J. (1991). The effect of flow rate upon the magnitude of the olfactory response differs for different odorants. Chem Senses *16*, 631–649.
6. Schoenfeld, T. A., and Cleland, T. A. (2006). Anatomical contributions to odorant sampling and representation in rodents: zoning in on sniffing behavior. Chem Senses *31*, 131–144.
7. Rojas-Líbano, D., and Kay, L. M. (2012). Interplay between sniffing and odorant sorptive properties in the rat. J. Neurosci. *32*, 15577–15589.
8. Cenier, T., McGann, J. P., Tsuno, Y., Verhagen, J. V, and Wachowiak, M. (2013). Testing the sorption hypothesis in olfaction: a limited role for sniff strength in shaping primary odor representations during behavior. J. Neurosci. *33*, 79–92.
9. Uchida, N., and Mainen, Z. F. (2003). Speed and accuracy of olfactory discrimination in the rat. Nat. Neurosci. *6*, 1224–9.
10. Abraham, N. M., Spors, H., Carleton, A., Margrie, T. W., Kuner, T., and Schaefer, A. T. (2004). Maintaining accuracy at the expense of speed: Stimulus similarity defines odor discrimination time in mice. Neuron *44*, 865–876.
11. Rinberg, D., Koulakov, A., and Gelperin, A. (2006). Speed-accuracy tradeoff in olfaction. Neuron *51*, 351–358.
12. Frederick, D. E. E., Rojas-Libano, D., Scott, M., and Kay, L. M. (2011). Rat behavior in go/no-go and two-alternative choice odor discrimination: differences and similarities. Behav. Neurosci. *125*, 588–603.
13. Kay, L. M., Beshel, J., and Martin, C. (2006). When good enough is best. Neuron *51*, 277–278.

Autonomous Situation Understanding and Self-Referential Learning of Situation Representations in a Brain-Inspired Architecture

Edgar Koerner, Andreas Knoblauch, and Ursula Koerner

Abstract Making sense of a scene has been considered a problem of sensory analysis traditionally. Prediction is used to cope with combinatorial explosion of possible alternative interpretations of the sensory signals. However, high complexity and variability of natural scenes limit the use of sensory appearance-based prediction dramatically. Brains of living beings seem to use a different strategy. Evolution discovered the power of storing an episode of successful behavior and re-using this memorized experience in similar situations. Such episodes consisting of intended behavior, its outcome, the spatial context, and relevant objects constitute situation models which control the selective inspection of sensory input required for its smooth execution. We argue that this behavior-based approach enables dynamically composed situation models that make the world more regular than it is indeed, and that can be learned autonomously.

Keywords Behavior-based situation models • Self-referential learning • Dynamic composition of internal models • Rhythmic control of composition process

1 The Challenge of Autonomous Situation Understanding

Situation understanding is a prerequisite for autonomous systems to behave properly in real-world environments. For establishing that capability a sufficiently rich body of knowledge on typical scenes and its constraints for behavior is required. However, acquired knowledge is only useful to the extent that it allows using experience of successful behavior in the past to select proper behavior in a similar situation, and to predict what may happen in the future [1]. But future situations are not exact replicas of past situations because of the high complexity and high variability of real-world scenes. Moreover, the intended task changes the situation with respect to what

E. Koerner (✉) • A. Knoblauch • U. Koerner
Honda Research Institute Europe GmbH, Offenbach, Germany
e-mail: edgar.koerner@honda-ri.de

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_71

are relevant and irrelevant aspects of the scene. This excludes appearance-based templates of scenes from being useful for an effective representation. Therefore, the required internal model for understanding situations must be dynamically composed according to constructive principles integrating elements and gist from stored experience that are relevant for the intended task.

How to control the composition of such a dynamic situation model? Which kind of knowledge representation could serve as the framework for guiding the process of composing a meaningful situation model for performing a certain task? For different tasks, even in case of the same sensory description of the outside scene, the situation may be different since it is the task which defines which objects and object configurations are relevant for its successful execution. Thus, the task specifies the demand on sensory analysis, providing a hypothesis of how the situation should look like to start its execution. Scene analysis is reduced in this way to a hypothesis guided sensory inspection of expected objects within expected configurations in the scene. In this way, the task demand serves to dramatically reduce the complexity of the scene, since it defines relevant objects and aspects while anything else is considered irrelevant for the time being.

2 Brain-Inspired Approach to Situation Understanding

Brains are basically control systems for organizing behavior. Evolution discovered the benefit of memorizing actions and its outcome and reusing this experience in similar situations. A proper frame for such a memorized action is an episode which consists of both the intended action itself, the spatial and behavioral context in which the action was experienced, and object configurations which have been relevant for successful execution of the intended task. The essential difference of this task-based situation model to the sensory data only based computer vision approach of scene analysis is in its capability to assign a value to an experienced situation. Action provides semantics to sensory data which enables the system to autonomously construct a consistent architecture of its knowledge representation based on self-reference. Finally, the difference in expected and actually experienced outcome provides the system with the capability to improve the situation models by adaptation and learning. In this framework, sensory analysis is not for creating an internal replica of outside world, but it is a directed search for those remembered aspects which are required to control the action to verify and adapt the internal model to the actual sensory situation.

For task-based episodes being an efficient situation model, similarity matching between memorized and actual situations at different levels of granularity must be possible, as well as autonomous learning to generalize and discriminate further within the acquired knowledge. Based on considerations of the phylogenetic development of cortex we hypothesize the following representational and control architectures may have evolved to support this function.

2.1 *The Hub-Network as Top-Level Organization of Different Representational Hierarchies*

The storehouse of the brain, the cortex, developed in vertebrates starting from a basic configuration comprising the hippocampus and a few old areas which served to represent actions, the respective spatial context, relevant objects, and the outcome of the action for later reuse in similar situations [2]. These phylogenetically oldest areas of the cortex constitute the basic system for organization of behavior and learning. At birth these heavily interconnected "old" subsystems are available and basic long-range connections are predefined, also in humans. During phylogenetic development, with increasing behavioral demand on fine-grained navigation and manipulation, additional subsystems for more specific and detailed representation emerged, which form hierarchic extensions of the concept representation beneath the "old" structures outlined above. These "old" structures are still heavily interconnected to each other and to its respective representation hierarchy so that they are referred to as "hubs" [3]. The hub network depicted in Fig. 1 may constitute the basic representational and control architecture for dynamic composition of situation models according to constructive principles.

The medial PFC selects basic actions (escape, fight, relax, consume, etc.) based on task-demand driven situation understanding provided by inputs about relevant current objects, their value, and the current spatial and behavioral context. The outputs of the mPFC to subcortical autonomous centers enable the adaptation of internal state of the body to the intended behavior. The hippocampus (HC) supports fast learning between any content including object-context-action-outcome. By activating such stored behavioral models we can construct a new scene with objects and persons, imagine an action and outcome [4]. This basic hub network is activated

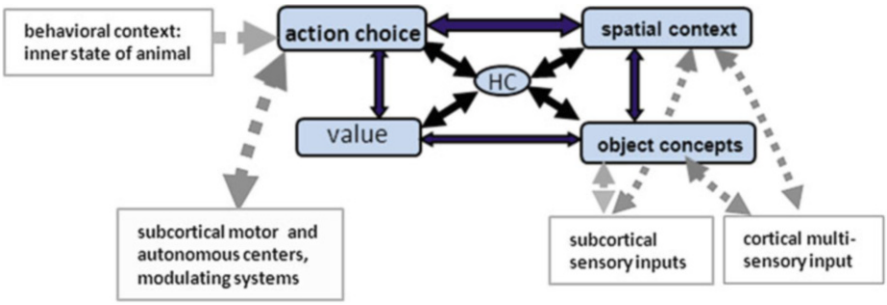


Fig. 1 Hub-network: Intensively communicating top-level systems of separate representation hierarchies, and their proposed neural substrates in the brain. *HC* hippocampus; value – orbitofrontal cortex codes value of objects and action outcomes; context – retrosplenial cortex and precuneus are structures that represent concepts of spatial context for own behavior; object concepts – perirhinal cortex represents object concepts; action choice – medial pre – frontal cortex (mPFC) selects basic actions

for any retrieval of self-related memory, as well as for planning and envisioning the future [5, 6]. Separate representation hierarchies dynamically linked at its top levels seem to be the key for that capability.

2.2 Rhythmic Control for Repetitive Composition of Dynamic Situation Model

For keeping track with the rapidly changing real world environment, the situation model must be updated with a sufficiently high frequency. Hence, to enable such a recurrent iteration of the construction process of the situation model, the length of the construction process must be limited. In the cortex, rhythmic modulation of excitability is ubiquitous and provides a natural control for iteration processes. Hippocampal nested theta and gamma rhythms may serve to link distributed subsystems of the different representational hierarchies (as outlined above) repetitively into a dynamic situation model according to the changing interaction situation. Any temporary stable composition can be learned with HC.

During slow-wave sleep so-called sharp-wave ripples (SW-R) transfer stored knowledge from HC to the hub network and the neocortex for consolidation [7]. However, the synchronizing SW-Rs during awake state seem to be involved in planning and decision making [8]. In the awake state a characteristic ripple-related sequence of events is observed starting with restriction of ongoing sensory inputs (extensive inhibition of subcortical structures), via activation of mPFC (action choice) and finally generation of SW-Rs in the hippocampus that synchronize respective sensory and motor areas for action realization [9].

It seems that this sequence of SW-R related activation reflects the generation of a dynamic situation representation according to constructive principles, including the dynamic allocation of the processing architecture for a planned action.

3 Conclusion

Self-referential learning is bound to a basic architecture that builds a “frame” of a behavioral model consisting of context, item, value, action and outcome. Specific representation hierarchies serve to extend the basic architecture on behavioral demand, and provide the means for dynamic composition of situation models even for scenes never experienced before. Task-demand based situation descriptions are highly robust situation models that can be anchored to sensory signal description of current physical environment by model guided sensory inspection and self-referential learning.

References

1. Schacter, D.L. and Addis, D.R. (2007) Nature 445; 27
2. Kaas, J.H. (2011) Brain, Behavior and Evolution 78:7–21
3. Sporns, O. (2013) Curr. Opin. Neurobiol. 23(2)162–71
4. Mullaly, S.L. and Maguire, E. (2013) Neuroscientist (in press)
5. Buckner, R.L., Andrews-Hanna, J.R., and Schacter, D.L. (2008) Ann NY Acad Sci, 1124, 1–38
6. Schacter, D.L., Addis, D.R., Hassabis, D., Martin, V.C., Spreng, R.N., & Szpunar, K.K. (2012) Neuron, 76, 677–94.
7. Buszaki, G. (1996) Cerebral Cortex 6:81–92
8. Jadhav, S.P., Kemere, C., German, P.W., and Frank, L.M. (2012) Science, 336:1454–8
9. Logothetis, N.K., Eschenko, O., Murayama, Y., Augath, M., Steudel, T., Evrard, H.C., Besserve, M., Oeltermann, A. (2012) Nature 491, 547–53

Metastable Activity Patterns in Cortical Dynamics and the Illusion of Localized Representations

Robert Kozma

Abstract This essay studies the relationship between recent advances in brain monitoring and modeling techniques, with the promise of identifying neural correlates of higher cognitive functions and consciousness. In this work we address the issue of the interpretation of the corresponding results, with special emphasis on the hypothesis of localized brain representations. We propose a dynamical systems theory framework to resolve the apparent contradiction between localized representations and large-scale, distributed (global) brain processing principles. Our dynamic model is a manifestation of Dreyfus' situated intelligence, following the traditions of Merleau-Ponty. This approach finds its applications in embodied intelligent systems and intentionality in animals and in man-made devices.

Keywords Transient dynamics • Phase transition • Metastability • Representationalism • Embodied intelligence.

1 Introduction

Advanced brain imaging techniques provide unprecedented insight into the operation of brains, which leads to approaches interpreting brain operation in the framework of neural representations. Brain activity exists simultaneously at multiple levels of a hierarchy. Based on the available experimental tools, the temporal and spatial observation windows can be granulated into the following scales: (i) microscopic level, which uses microelectrodes to measure synaptic and action potentials at the micron scale; (ii) mesoscopic level, corresponding to mm in space and tens of ms in time, measured by electrocorticograms (ECoG); and (iii) macroscopic measurements in cm and s from scalp electroencephalograms (EEG), magnetoencephalograms (MEG), and functional magnetic resonance imaging (fMRI) reflecting levels of cerebral metabolism [2, 6].

R. Kozma (✉)
Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152, USA
e-mail: rkozma@memphis.edu

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_72

High spatial resolution images by fMRI [3] can be linked with high temporal resolution EEG and MEG signatures by combining hemodynamic imaging with EEG imaging [1]. Single neurons in human and animal neocortex have been shown to generate spike trains during performance of precise cognitive tasks such, for example, as face recognition [14]. These cells are called colloquially as grandmother cells [11]. There is an apparent contradiction between the experimentally observed localized cognitive responses, and the general view of brains as massive distributed organs processing information and knowledge.

In this work, we analyze the contradiction between localized representations and embedded cognition principles [4]. Contrary to representationalist approaches by Minsky and colleagues [12, 13], who view intelligent behavior as symbolic manipulation of mental representations, Dreyfus assumes no symbolic representations in brains [4]. Following the embedded cognition approach by Dreyfus, we describe the operation of brains through a sequence of metastable states [6, 10] as brains interact with their environment. These metastable states may be viewed as intermittent 'symbols,' however, they are transient and disintegrate soon after they emerge [10]. Based on the concept of transient, metastable states, we outline the principles of intentional neurodynamics and indicate their potential benefits in practical implementations of intentional systems.

2 The Illusion of Cortical Representations

Representationalist models of brain dynamics and cognition have their roots in symbolic approaches to intelligence in biological and artificial systems. Symbolic models and knowledge-based systems proved to be powerful tools dominating the field of artificial intelligence for about three decades (from 1960s through the 1980s) [5, 12, 13]. In the representationalist point of view, external events and perceptions are transformed into inner symbols to represent the state of the world.

Due to the success of symbolic approaches of artificial intelligence, it is natural to see that neuroscientists try to identify "symbolic representations" in brains. Representationalism seems to receive experimental support from intracranial recordings in human and animal neocortex. Namely, single neuron recordings indicate the presence of dedicated neurons responding to specific input stimuli during cognitive tasks such, e.g., face recognition [14]. These and further similar results in the literature have been used as supporting arguments for localized representations and the concept of grandmother cells [11].

There is, however, a critical caveat in the representationalist interpretation of these single neuron experiments. If grandmother neurons were randomly distributed in the neocortex, the likelihood of finding any of them would be vanishingly small. It would be practically impossible to identify the visual image,which would evoke

a measurable response in a given probe by flashing even hundreds of visual images in front of the patient. The fact that such responsive cells can be found at all implies that they are concentrated in neural assemblies distributed in the neocortex. Potential candidates are Hebbian cell assemblies that are formed by learning to enact a designated task [6]. Once any component of a Hebbian cell assembly is stimulated, the whole assembly responds, which may explain the experimentally observed effect.

There are theoretical reasons for alternatives to symbolic approaches of intelligence. The concept of situated intelligence by Dreyfus is a prominent example of philosophical alternatives to symbolism. Dreyfus ascertains that intelligence is defined in the context of the environment; therefore, a preset and fixed symbol system cannot grasp the essence of intelligence [4]. Connectionism provides a useful model of brains and cognition, which is in many ways complementary to symbolic approaches. Pragmatic implementations of situated intelligence find their successful applications in embodied intelligence and robotics [9]. In the next section, we outline a neural network-based approach to interpret experimental finding of cortical spatio-temporal dynamics.

3 Phase Transition Models of Metastable Cortical Patterns

Detailed studies of EEG and ECoG data reveal that the observed brain activity patterns are intermittent both in time and space [5, 10]. Metastable patterns exist for relatively long time periods, while the transition from one state to the other is fast. There are various models to describe the experimental findings, including chaotic itinerancy [16], and the existence of transient heteroclinic channels [15]. Behavioral manifestations of metastablity and the complementarity principle is described in [7]. Here we employ the cinematic theory of brain dynamics. According to the cinematic theory [6], cognition is not a smooth, continuous process in time; rather it is a sequence of metastable patterns (cinematic frames), interrupted from time to time by discontinuities (acting as the shutter).

Intentional neurodynamics is a biologically motivated connectionist approach. We employ neuropercolation model system to implement principles of intentional neurodynamics using phase transitions [2,5]. The neuropercolation model describes various dynamical properties of metastable activity patterns in the cortical neuropil. The duration of the metastable patterns is much longer than the very rapid transition from one pattern to the other. This is illustrated in Fig. 1, where the ensemble average activity level is shown over two-dimensional neuroperolation lattices. Rewiring is a critical control parameter, which is used for tuning the characteristic length of the metastable states; for details, see [2, 8].

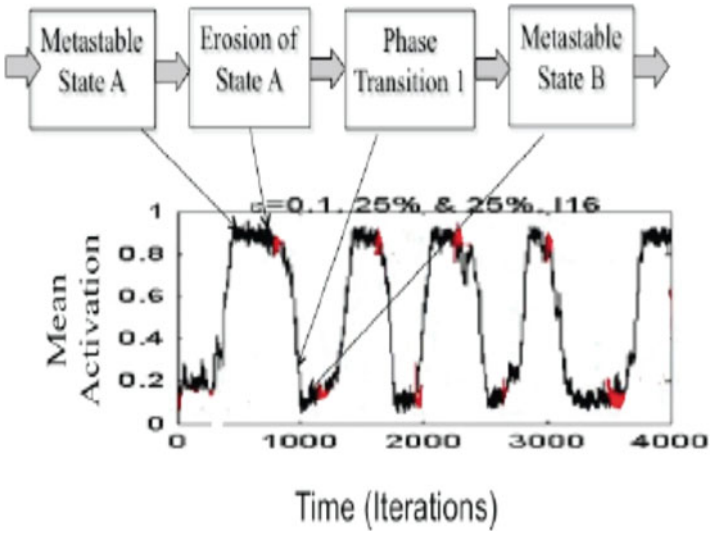


Fig. 1 Illustration of the dynamics of the metastable states modeled by neuropercolation; *red marks* show the start of the erosion of metastable states

4 Conclusion

In this work we discuss the question of localized representations in brain recordings. We argue that metastable patterns may be perceived as localized representations with specific meaning in the context of the state of the cortex and inputs. However, such percepts are transitory illusions, and any specific pattern rapidly dissolves (within a fraction of second) and new patterns emerge. Even if the internal and external states of the system remain unchanged during such transitions, the new patterns are different form the previous one due to the changed dynamic trajectory of the brain as a dynamic system.

The proposed analysis may help to reconcile the contradictions inherent in perceived mental representations and to establish a harmony between the everlasting human desire to create tangible chunks of knowledge reflecting our limited grasp of reality and the infinite complexity of the world where our body and brain reside.

Acknowledgements This work has been presented at the 2013 Dynamic Brain Forum (DBF), as part of ICCN2013, with the kind support of the DBF Foundation. via the Center for Integrative Mathematics, Hokkaido University, Japan. This research has been supported in part by DARPA Physical Intelligence program, through a contract with Hughes Research Laboratory (HRL), Malibu, CA (Dr. N. Srinivasa, Program Director), and by NSF CRCNS Program 13-11165.

References

1. Barlow JS. The Electroencephalogram: Its Patterns and Origins. Cambridge MA: MIT Press (1993).
2. Bollobas, B., Kozma, R., Miklos, D. (eds.) Handbook of Large-Scale Random Networks, Springer (2009).
3. Buxton R.B., Introduction to Functional Magnetic Resonance Imaging: Principles and Techniques. Cambridge UK: Cambridge UP (2001).
4. Dreyfus, H.L. How Representational Cognitivism failed and is being replaced by Body/World Coupling. After Cognitivism: A reassessment of Cognitive Science and Philosophy, Springer: 39–73 (2009).
5. Freeman, W.J., How Brains Make up their Mind, Columbia University Press (2001).
6. Freeman, W.J. and R. Quiroga, Imaging Brain Function with EEG, Springer (2013).
7. Kelso, J.A.S., D.A. Engstrom, The complementary nature. Cambridge, MA, MIT Press (2006).
8. Kozma, R. Intentional Systems: Review of Neurodynamics, Modeling, and Robotics Implementation, Physics of Life Reviews, 5(1), 1–21, (2007).
9. Kozma, R., T. Huntsberger, H. Aghazarian, E. Tunstel, R. Ilin, WJ Freeman, Intentional Control for Planetary Rover SRR2K, Advanced Robotics, 21(8), 1109–1127 (2008).
10. Kozma, R., Freeman, W.J., The KIV Model of Intentional Dynamics and Decision Making, Neur. Netw. 22(3): 277–285 (2009).
11. Lettvin JY, Lettvin on grandmother cells. pp 434–435 in The Cognitive Neurosciences, Gazzaniga MS (ed.). Cambridge MA: MIT Press (1995).
12. Minsky, M. A framework for representing knowledge. MIT, AI Lab, Cambridge, MA (1974).
13. Newell, A., Simon, H. A. Human problem solving. Englewood Cliffs, NJ: Prentice-Hall (1972).
14. Quiroga QR, Reddy L, Kreiman G, Koch C, Fried I, Invariant visual representation by single-neurons in the human brain. Nature 435: 1102–1107 (2005).
15. Rabinovich, M. I., Varona, P., Selverston, A. I., Abarbanel, H. D. Dynamical principles in neuroscience. Reviews of modern physics, 78(4), 1213 (2006).
16. Tsuda I. Toward an interpretation of dynamics neural activity in terms of chaotic dynamical systems. Beh. and Brain Sci., 24(5): 793–810 (2001).

Memories as Bifurcations Shaped Through Sequential Learning

Tomoki Kurikawa and Kunihiko Kaneko

Abstract Representation of memory in the neural system is one of central problems in the neuroscience. We have proposed a novel memory viewpoint "memories-as-bifurcations", in contrast to the classical viewpoint "memoiries-as-attractors." Here, we analyze generalization of memory, based on this novel viewpoint and show that such generalization is formed through a simple sequential learning process.

Keywords Memory • Bifurcations • Sequential learning • Rate-coding neurons • Generalization

1 Introduction

How memory is represented in neural system is an important problem for understanding cognitive function in our brain. Often, a memory is considered as an attractor of neural circuit [1], where an initial state of the network is set by an input and then the neural state converges to an attractor depending on the initial state. There, spontaneous activity without external stimuli is not taken into account seriously, while recent experimental studies have demonstrated that the spontaneous activity plays a key role in the response to a stimulus [2, 3]. To better understand this role, we have proposed a novel viewpoint [4], "memories-as-bifurcations", that differs from the traditional "memories-as-attractors" viewpoint. According to this memories-as-bifurcations viewpoint, memory is recalled when the spontaneous neural activity is changed to an appropriate output activity upon application of an input, known as a bifurcation in dynamical systems theory, wherein the input modifies the flow structure of the neural dynamics. Although we have exhibited that memories are formed through a simple learning rule [5], generalization process

T. Kurikawa (✉)
Brain Science Institute, RIKEN, 2-1 Hirosawa Wako-shi, Saitama 351-0198, Japan
e-mail: tomoki.kurikawa@riken.jp

K. Kaneko
Graduate School of Arts and Sciences, University of Tokyo, Komaba 3-8-1, Meguro-ku, Tokyo, Japan
e-mail: kaneko@complex.c.u-tokkyo.ac.jp

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_73

in which a system retrieves an identical pattern by applying some inputs has not been well studied yet. In this paper, we analyzed the generalization process formed through the sequential learning process used in the previous study [5].

2 Model

We consider N rate-coding neurons that take continuous activity values x_i, which evolve as

$$\dot{x_i} = \tanh(\beta(\sum_{j\neq i}^{N} J_{ij}x_j + \gamma\eta_i^{\mu})) - x_i, \tag{1}$$

where J_{ij} denotes a connection from the j-th to i-th neuron, $\gamma\boldsymbol{\eta}^{\mu}$ is an input pattern $\boldsymbol{\eta}^{\mu}$ of input strength γ, and μ is index of learned mappings. For a learned input pattern $\boldsymbol{\eta}$, we set a pattern $\boldsymbol{\xi}$ as a target (each pattern is a binary random pattern). The synaptic connection J_{ij} evolves according to

$$\dot{J_{ij}} = \alpha(\xi_i^{\mu} - x_i)x_j, \tag{2}$$

where $\alpha > 0$ is a learning parameter. We give a set of M random correlated input and output patterns, whose correlation satisfies $[\boldsymbol{\eta}^{\mu}\boldsymbol{\eta}^{\mu+1}]/N = [\boldsymbol{\xi}^{\mu}\boldsymbol{\xi}^{\mu+1}]/N = C$. Here, $[\cdots]$ means the average over random patterns of input and target. Every mapping is learned in reverse numerical order from $\mu = M-1$ to $\mu = 0$. Further, a system learns the set iteratively in the same order.

3 Results

Through the learning process, the memories of mappings are embedded in the system. First, in order to evaluate the response of the system to the learned input, we measured the average overlap $[< \overline{\boldsymbol{x}\boldsymbol{\xi}^{\mu}}/N >]$ upon η^{μ} as a function of μ for $C = 0.9$ and 0.1, shown in Fig. 1a, where $\overline{\cdots}$, $< \cdots >$ and $[\cdots]$ mean the average over time, initial states of one network, and networks, respectively. Note that the response is defined here as an activity in the presence of an input, not as an evoked activity by a transient input used only for the initial condition as in the Hopfield model. For $C = 0.9$ and 0.1, the average overlap with the latest learned target ($\mu = 0$) takes nearly unity and this target can be recalled perfectly. The average overlap with the earlier learned target decreases rapidly and then, saturates at around 0.8 for $C = 0.9$, whereas the overlap keeps nearly unity for $C = 0.1$. Interestingly, memory performance of the system that learns a set with lower correlation is greater than that with a higher correlation.

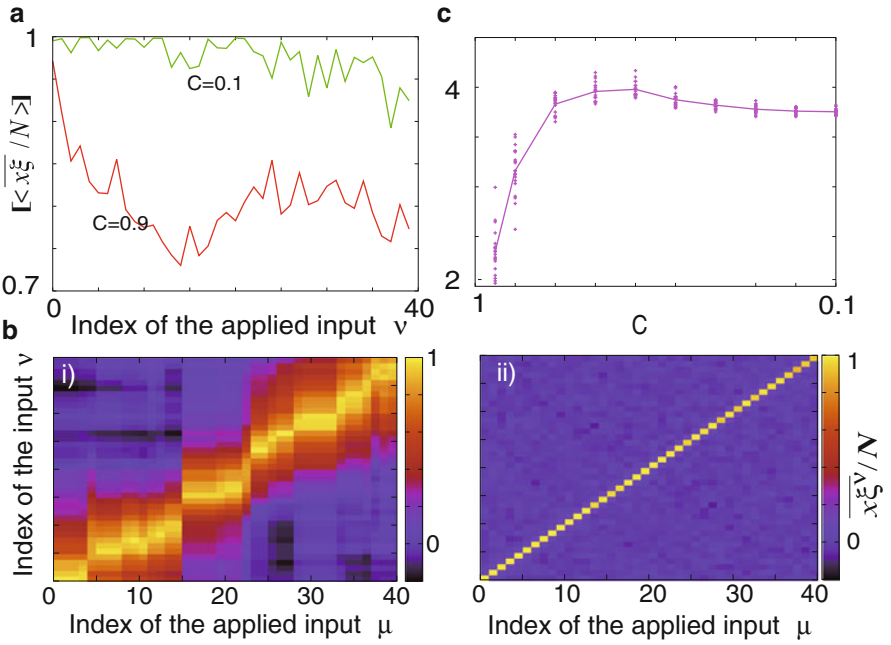


Fig. 1 (**a**) The average overlap $[< \overline{x\xi^{\nu}}/N >]$ in the presence of the input ν is plotted for $C = 0.1$ and 0.9, where larger ν represents earlier learned input. (**b**) The average overlap $< \overline{x\xi^{\mu}}/N >$ with the target μ in the presence of the input ν is shown for (i) $C = 0.9$ and (ii) $C = 0.1$. (**c**) The basin entropy $\Sigma v_i log v_i$ is plotted as a function of C

To analyze the behavior in detail, we computed the temporal average of the overlap with ξ^{μ} upon η^{μ}. We calculated the average overlap $< \overline{x\xi^{\nu}}/N >$ upon η^{μ} as a function of μ and ν, plotted in Fig. 1b.[1] For $C = 0.9$ (Fig. 1b(i)), different inputs that are learned evoke an identical pattern, which is one of the targets corresponding to the inputs. Different inputs are associated with the same target, and other targets are not recalled precisely. This lowers the memory performance, but inputs are categorized into a cluster. and some clusters are formed. For $C = 0.1$ (Fig. 1b(ii)), in contrast, every evoked pattern is distinct depending on each input. Each target is recalled correctly and higher memory capacity is achieved.

Depending on C, the behavior of evoked patterns is significantly changed. We analyzed this change by measuring basin entropy S for different C, shown in Fig. 1c, where $S \equiv -\Sigma_i \mathrm{vol}_i log \mathrm{vol}_i$, $\mathrm{vol}_i = \Sigma_{\mu=0}^{M-1} \mathrm{vol}_i^{\mu}/M$. vol_i^{μ} is basin volume of i-th attractor upon η^{μ}. For large C, the entropy takes very small value, meaning that a few large clusters are formed. With decreasing C, the entropy increases rapidly and the number of the clusters also increases (data not shown). At around $C = 0.7$,

[1] In the following analysis, we focus on behavior of one network, but the qualitative behavior is independent of networks.

the entropy and the number of attractors take the maxima, where several attractors coexist upon each of inputs. For much smaller C, the entropy value is closer to 3.689, which is value when only one different attractor exists corresponding to each input.

So far we demonstrated that the similar learned inputs can evoke the identical pattern. Now, we analyze the behavior when a novel input, which is similar to one of the learned ones, is presented. For this purpose, we measured the average overlap in the presence of the mixed input $\gamma^1\eta^1 + \gamma^2\eta^2$ with changing γ^1 and γ^2 for $C = 0.1$. As shown in Fig. 2a, in a regime with higher γ^1 and lower γ^2, $< \overline{\boldsymbol{x}\boldsymbol{\xi}^{\boldsymbol{1}}}/N >$ takes nearly one. Upon an input in this regime of (γ^1, γ^2), $\boldsymbol{\xi}^{\boldsymbol{1}}$ is perfectly recalled. Such input is novel for the present neural system, because the system has learned inputs only with $(\gamma^1, \gamma^2) = (1, 0)$ and $(0, 1)$. Thus, recognition of the novel inputs as η^1 implies the generalization. With decreasing γ^1 and increasing γ^2, there occurs bifurcation from the stable fixed-point corresponding to $\boldsymbol{\xi}^{\boldsymbol{1}}$ to the chaotic behavior, as shown in Fig. 2b. With further decreasing γ^1 and increasing γ^2, $\boldsymbol{\xi}^{\boldsymbol{2}}$ is amplified and stabilized through another bifurcation. For the lower γ^1 and the higher γ^2, there is a phase in which $< \overline{\boldsymbol{x}\boldsymbol{\xi}^{\boldsymbol{2}}}/N >$ takes nearly one, and generalization to regard the input as η^2 follows. These results demonstrate the parameter space (γ^1, γ^2) is divided to three phases; the phases in which $\boldsymbol{\xi}^{\boldsymbol{1}}$ is recalled, $\boldsymbol{\xi}^{\boldsymbol{2}}$ is recalled, and the chaotic dynamics wandering between $\boldsymbol{\xi}^{\boldsymbol{1}}$ and $\boldsymbol{\xi}^{\boldsymbol{1}}$, i.e., wondering if the input is regarded as η^1 or η^2.

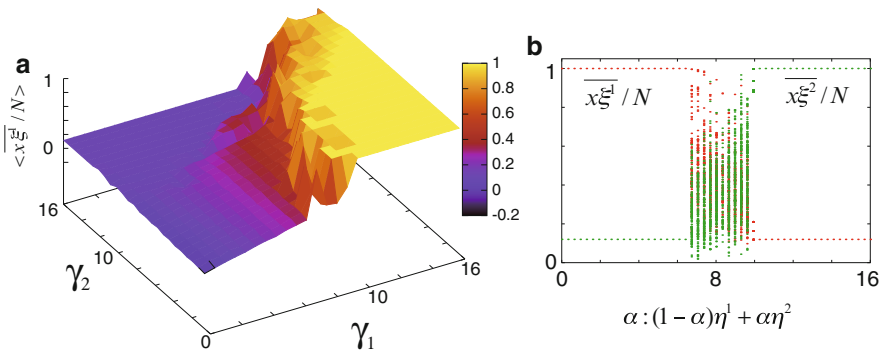


Fig. 2 Bifurcation diagrams through changing two input strengths. (**a**) The average overlap $< \overline{\boldsymbol{x}\boldsymbol{\xi}^{\boldsymbol{1}}}/N >$ upon the input $\eta = \gamma^1\eta^1 + \gamma^2\eta^2$ is plotted as a function of γ^1 and γ^2. (**b**) A bifurcation diagram with changing $\eta = \alpha\eta^2 + (1-\alpha)\eta^1$ is shown

4 Conclusions and Discussions

In this study, we analyzed the generalization of memory in memories-as-bifurcations viewpoint, which has been proposed in [4]. We demonstrated that after the simple learning, the system can retrieve a target pattern by applying not only the learned input corresponding to this target but also different patterns whose input strengths shape the distinct region in the parameter space, meaning that the generalization of memory is achieved. The present study may provide a novel viewpoint on generalization, as input-evoked bifurcation in neural activity dynamics shaped by learning. Indeed, some experimental studies suggest such generalization occurs in sensory system [6]. Although how the generalization is related to the spontaneous activity needs to be clarified in future, the bifurcation-based viewpoint will shed light on better understanding of the neural processing of memories.

Acknowledgements This work was partially supported by a Grant-in-Aid for Scientific Research on Innovative Areas "Neural creativity for communication (No.4103)" (No. 21120004) and No. 233744 from MEXT, Japan.

References

1. J. Hopfield, Proc., Natl., Acad., Sci., USA, **81**, 1984
2. T. Kenet, D. Bibitchkov, M. Tsodyks, A. Grinvald, and A. Arieli, Nature, **425**, 2003
3. P. Berkes, G. Orbn, M. Lengyel, and J. Fiser, Science, **331**, 2011
4. T. Kurikawa and K. Kaneko, EuroPhys Lett., **98**, 2012
5. T. Kurikawa and K. Kaneko, PLoS Comput. Biol., **9**, 2013
6. J. Niessing and R. Friedrich, Nature, **465**, 2010

Behavioral Interactions of Two Individual Arm Robots Using Independent Chaos in Recurrent Neural Networks

S. Kuwada, T. Aota, K. Uehara, S. Hiraga, Y. Takamura, and Shigetoshi Nara

Abstract Based on a heuristic idea and by computer experiment, we show that chaos introduced into a recurrent neural network model can enable "complex control with simple rule(s)" under ill-posed situations. Furthermore, we show behavioral interactions of two individual arm robots driven by independent chaos implemented into each arm control system using recurrent neural networks.

Keywords Neuro-dynamics • Constrained chaos • Neural networks • Functional experiments • Behavioral interactions • Inter-brain communications • Ill-posed controls

1 Introduction

Since a few decades, brain science has been greatly developed, however, the mechanisms of advanced functions of brain have been still beyond our understanding. In these situations, there are people who think that recently discovered chaos in brain or biological systems could play an important role in their advanced functions [1–3]. Nara and Davis proposed that chaotic dynamics can occur in a recurrent neural network model by changing a system parameter, and they have studied that it can be applied to solving ill-posed problems, for example, memory search or synthesis, to solve maze (labyrinth) with use of chaotic roving robot, and so on [4–6]. In their opinion, chaotic dynamics with certain dynamical structures plays an important role in complex functions. In this paper, based on the same idea with them, and by computer experiment, we propose that chaos introduced into a recurrent neural network model can enable "complex arm control with simple rule(s)" under ill-posed situations [7, 8], and as an actual example, we show that an arm robot without having advanced visual processing function (see Figs. 1 and 2) can take an target object to a set position under ill-posed situations.

S. Kuwada • T. Aota • K. Uehara • S. Hiraga • Y. Takamura • S. Nara (✉)
Electrical and Electronic Engineering Department, Graduate School of Natural Science and Technology, Okayama University, Tsushima-naka 3-1-1, Kita-ku, Okayama 700-8530, Japan
e-mail: nara@ec.okayama-u.ac.jp

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_74

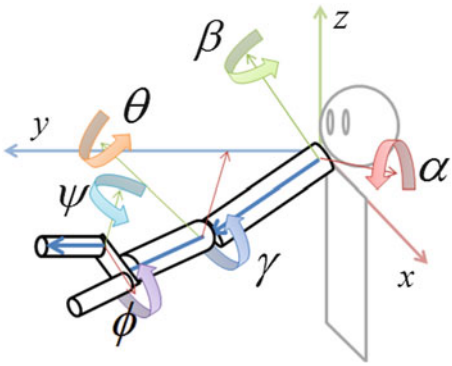


Fig. 1 Our arm model in Euler angle scheme

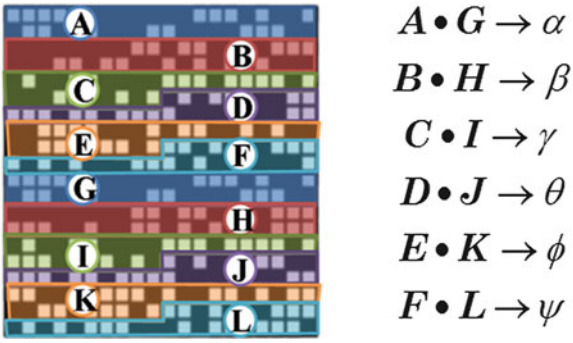


Fig. 2 Firing pattern of 400 neurons & sub-vectors' codings which correspond to incremental motions via Euler angles

Furthermore, we show behavioral interactions of two individual arm robots driven by independent chaos implemented into each arm control system using recurrent neural networks.

2 Method and Model

Our study works with an interconnected recurrent neural network model (abbreviated as RNNM hereafter) consisting of N binary neurons, and the updating rule is defined by

$$x_i\,(t+1) = \mathrm{sgn}\left\{ \sum_{j \in G_i(r)}^{N} w_{ij} x_j(t) \right\}$$

$$\mathrm{sgn}(u) = \begin{cases} +1 & u \geq 0 \\ -1 & u < 0 \end{cases}$$

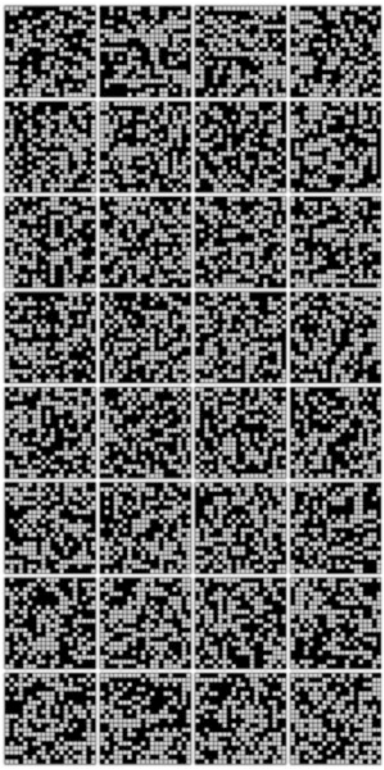


Fig. 3 The cycle attractor patterns

where $x_i(t) = \pm 1$ ($i = 1 \sim N$) is the firing state of a neuron specified by space site index i at time t, and w_{ij} is connection weight from the neuron x_j to the neuron x_i. w_{ii} is taken to be 0. r ($0 < r < N$) is fan-in number for neuron x_i, named *connectivity* that is the most important system parameter in our work. $G_i(r)$ is a spatial configuration set of connectivity r for neuron x_i, the number of which are ${}_{N-1}\mathrm{C}_r$. Therefore, with full connectivity $r = N - 1$, determination of w_{ij} by means of a kind of orthogonalized learning method enables us to embed a group of N dimensional state patterns (vectors) as cycle memory attractors in N dimensional state space. Let us employ our arm model as shown in Fig. 1, and introduce coding of sub-vectors of neuron firing pattern (vector) as increment of the Euler angles in arm motion. In our neural model system, attractor patterns consists of (K patterns per cycle) $\times$ L cycles, and each patterns has N neurons. In this work, we take $K = 4$, $L = 8$, and $N = 400$, where the firing states of $N = 20 \times 20 = 400$ neurons are represented by black pixel or white pixel (see Fig. 2). Long time updating makes an initial pattern converge to one of the embedded cycle attractors.

Now, when we reduce connectivity r by blocking signal transfer from the other neurons, then attractors gradually become unstable, and the network state changes from attractor dynamics to chaotic dynamics, where we discard the detailed description of the destabilizing processes [5]. Let us describe actual arm motions

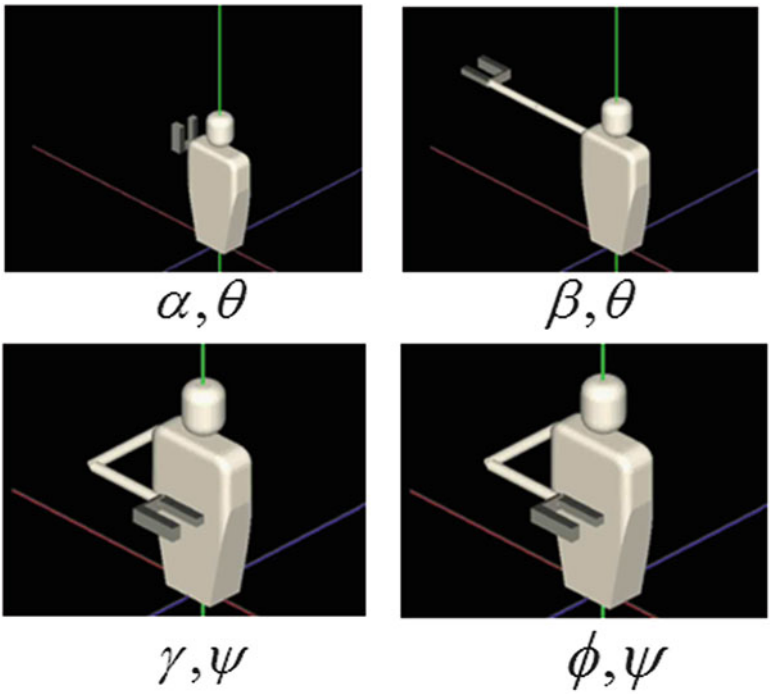


Fig. 4 The embedded four definite motions corresponding to the Fig. 3

and neural firing patterns (vectors). In the full connection state, when one of the embedded cycle attractors appears, then, at each time step, decoded quantities following the corresponding relations shown in Fig. 2 stationary give incremental Euler angles, so the generated motion by them is one of the definite motions shown in Fig. 4, whereas they are snap shots of the definite periodic motions of arm. It should be noted that, to display the results of computer experiment, we used the software "Insilico IDE" which is opened to the public on the web site "Physiome Platform". Once connectivity is reduced to one order of magnitude smaller than full connection number, then dynamics of firing pattern becomes chaotic. Correspondingly, decoded motions indicate chaotic behaviors, in which fragmental motion of the embedded definite motions are coming out, vanishing, coming out vanishing, and repeating them chaotically.

Now, we apply these two types of motions, definite motions and chaotic motions to realizing (a) catching a target object, (b) catching a target and taking it back to the set position, (c) catching a target and taking it back to the set position under the existence of unknown obstacles, (d) competitive catching a target object between two independent arm robots, where, in all cases, we assume that robot has no advanced ability of visual information processing, and only adaptive switching between attractor regime (full connection) and chaotic regime (small connection)

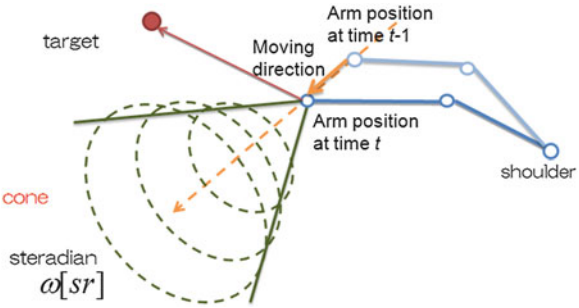


Fig. 5 A schematic description of the condition about connectivity switching. If the object is *inside a certain cone* the axis of which is an elongation of direction from the position of arm edge at time t-1 to the position of arm edge at time t, *then* $r = N - 1$, *otherwise* $r =$ small.

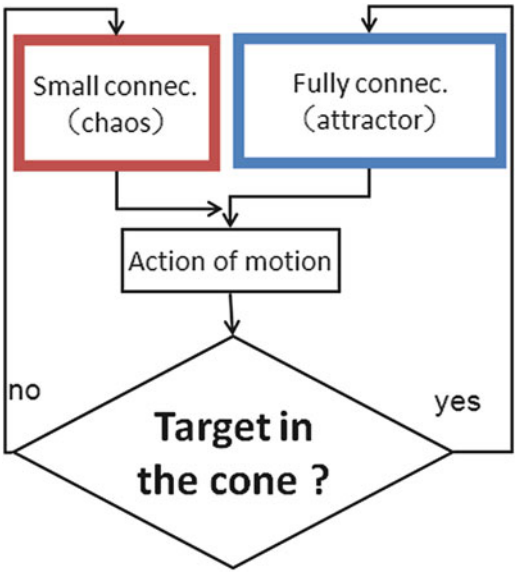


Fig. 6 Control algorithm of connectivity switching at each time step, corresponding to Fig. 5

depending on situations with including uncertainty (ill-posed situations) is used by means of simple rule. Figure 5 shows a schematic description of switching condition of connectivity and Fig. 6, the rough algorithm of this control system, where Figs. 7 and 8 are actual two examples of set situations (c) and (d) given above. All the computer experiments are successfully done and only about the case of (d), the result to evaluate the success rate with respect to connectivity is shown in Fig. 9.

One can recognize that chaos generated by rather small connectivity give better results, which means that chaos with having strong ruin of embedded attractors prevents the robots from generating various and/or adaptive motions in given environments. So, chaos having certain weak dynamical structures could be useful, however, optimization of dynamical structures is quite difficult problem and it would be big issues including learning of chaos in this scheme.

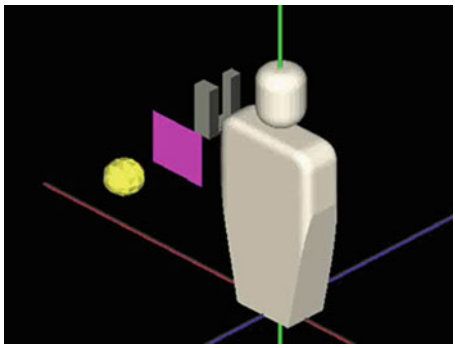


Fig. 7 Experiment of the situation (c) given in the text

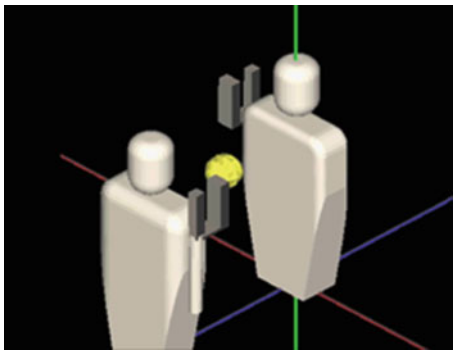


Fig. 8 Experiment of the situation (d) given in the text

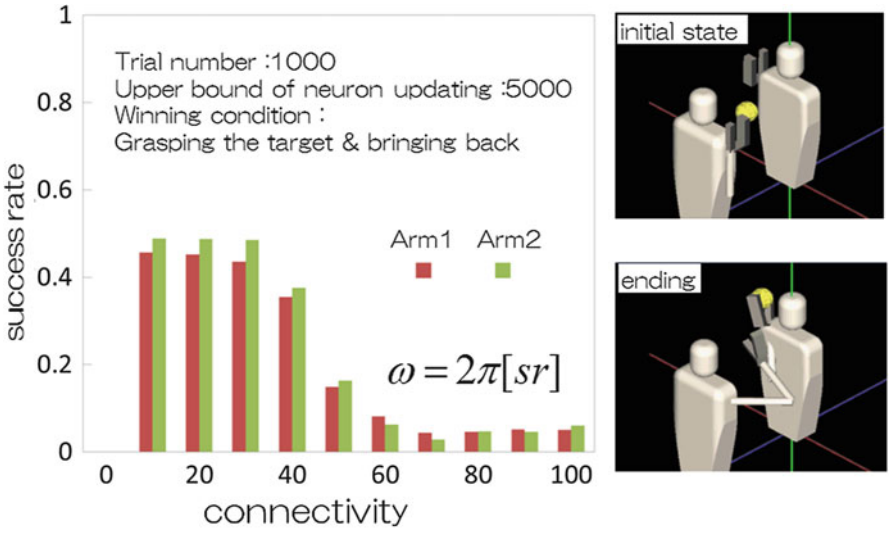


Fig. 9 Computer experiments of competitive taking of a target object between two arm robots, where evaluated success rate depending on connectivity is shown.

3 Concluding Remarks

1. Using Euler angle model of an arm robot having 6° (angles) of freedom, we made computer experiments to solve ill-posed problems using chaos in RNNM installed into the control systems of the robot, where the robot can obtain only information about target direction with uncertainty, and without any pre-knowledge about configurations of obstacles.
2. By the computer experiments, we were able to show that chaos realize autonomous and adaptive functions with use of simple rule(s)
3. The computer experiments in which the two robots competitively behave as taking off and back a target, show that functional performance strongly depends on dynamic structures of chaos generated by reducing connectivities in RNNM.

Acknowledgments This work is supported partly by Grant-in-Aid in the Ministry of Education, Science, Sports & Culture of Japanese Government, and partly by Co-operative Research Program of "Network of Joint Center for Materials & Devices"

References

1. C. A. Skarda, W. J. Freeman, Behav. Brain Sci. vol. 10, p. 161 (1987)
2. I. Tsuda, Behav. Brain Sci., vol. 24, p. 793 (2001)
3. S. Nara, P. Davis, Prog. Theor. Phys. vol. 88, p. 845 (1992)
4. S. Nara, P. Davis, M. Kawachi and H. Totuji, Int. J. Bif. & Chaos, vol. 5, p. 1205 (1995)
5. S. Nara, CHAOS, vol. 13, p. 1110 (2003)
6. Y. Suemitsu and S. Nara, Neural Computation, vol. 16 p. 1943 (2004)
7. Y. Li, S. Kurata, S. Morita, S. Shimizu, D. Munetaka and S. Nara, Bio. Cyb., vol. 99, p. 185 (2008)
8. R. Yoshinaka, M. Kawashima, Y. Takamura, H. Yamaguchi, N. Miyahara, K. Nabeta, Y. Lee & S. Nara, Studies in Computational Intelligence, Vol. 399, p.287-p.305, K. Madani, A. D. Correia, A. Rosa & J. Filipe (Eds.), Springer Verlag, 2012

Free Will and Spatiotemporal Neurodynamics

Hans Liljenström

Abstract It is widely assumed that neuroscience has shown that conscious will is an illusion. Indeed, a number of experimental results seem to indicate that conscious will is not causally related to the willed action, i.e., epiphenomenal. There are, however, alternative ways of interpreting the results that these conclusions are based on, and this paper examines the scientific arguments and experiments regarding conscious will. We argue that there is, as yet, no empirical support for epiphenomenal conscious will, and that the alternative hypothesis, that conscious will is causative, is also consistent with experimental data. We also use computational models and simulations to demonstrate that local neural impulses may trigger global oscillatory activity after a substantial delay period, supporting the idea that intentional impulses may be part of a conscious will experience.

Keywords Free will • Intention • Consciousness • Neuronal causation • EEG • fMRI • Brain stimulation • Cortical networks • Computer simulations

1 Introduction

The subjective experience of agency is so immediate that we consider it self-evident that our actions are controlled by our free will. Yet, there is an increasing amount of experiments that seem to indicate that free will is an illusion without causal effects on the nervous system. In fact, it appears that our actions are governed by neural events and processes prior to our experience of our free will. Experimental results by e.g. Soon et al. [1] or Desmurget et al. [2] are taken as support for the dominating paradigm of materialism, which excludes any mental-neural (downward)

H. Liljenström (✉)
Department of Energy and Analysis, SLU, Biometry and Systems Analysis, P.O. Box 7032, SE-750 07 Uppsala, Sweden

Agora for Biosystems, P.O. Box 57, SE-193 22 Sigtuna, Sweden
e-mail: hans.liljenstrom@slu.se

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_75

causation [2]. The issue of free will is of profound importance for understanding human nature, and a careful examination of the experimental results and their interpretations is therefore essential. Here, we consider *free will* as the conscious intention to act, and we will refer to this as *conscious will* in the following, to stress the close relation between free will and consciousness. While consciousness is not yet understood, we assume there is a neural correlate of conscious will, and that conscious will can be probed experimentally through the reports of human subjects. In the following, we will briefly summarize the experiments that have been interpreted as evidence for a non-causal, *epiphenomenal* conscious will.

2 Experimental Results as Evidence

A series of famous EEG experiments [3–7] (see [8] for a review) are often quoted as evidence for an illusory free will. The EEG *readiness potential* (RP), apparent only when averaged over a large number of trials, seems to precede the conscious will for *spontaneous* voluntary movements. The RP precedes the conscious will by about 850 ms with admitted pre-planning, and by 350 ms without admitted pre-planning. This demonstrates that the timing of the RP depends critically on pre-planning. (The actual action occurs about 200 ms after the perceived impulse to act).

While EEG (and MEG) may elucidate temporal relationships in willed actions, at time scales of seconds, or below, brain imaging techniques such as PET and fMRI may provide better spatial information of which brain regions are involved, at time scales of several seconds to minutes. An early study of regional cortical blood flow (rCBF) during willed imagined and real motor acts was performed by Ingvar and Philipson [9], where clear differences in the two cases were found. These experiments, as well as subsequent rCBF studies by Frith et al. [10] point at prefrontal cortex (PFC) as central in planning and choice of willed actions, but also the supplementary motor area (SMA), parietal cortex and the basal ganglia seem to be involved (see e.g., [11–13]).

Soon et al. [1] investigated timing correlations between subjective decisions and brain events measured by fMRI. Subjects were asked to freely select between pressing either of two buttons operated by the left and the right hand, respectively, and to indicate the time when the decision was made. For some critical brain regions, it was possible to predict the handedness (left or right) of the action with at most about 60 % accuracy, where 50 % accuracy would indicate uninformed plain guessing. It was found that fMRI maps from the frontopolar cortex acquired 10 s before the action carried information leading to an, on average, 60 % accurate prediction. The frontopolar cortex is assumed [14] to be involved in reasoning or evaluation of internally generated information. The timing of the action also seems predictable with 60 % accuracy from fMRI data obtained from the SMA. The precise role of SMA is not known, but it is typically assumed that it is involved in planning learned sequences of movements.

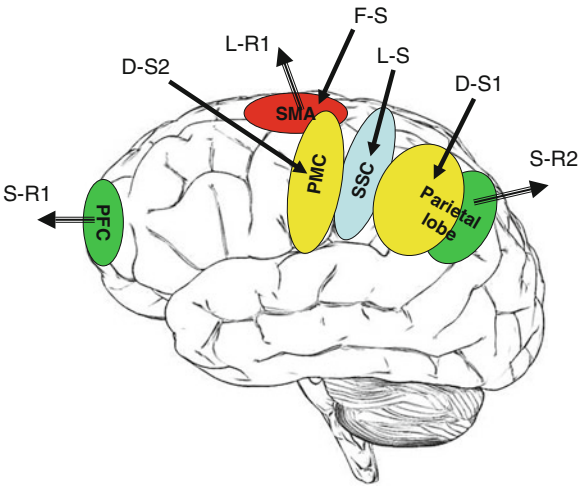


Fig. 1 Schematic description of some of the experimental loci discussed in the text. *SMA* Supplementary motor area, *PFC* Prefrontal cortex, *PMC* Premotor cortex, *SSC* Somatosensory cortex. L-S corresponds to Libet's stimulation of the somatosensory brain area. L-R1 is the recording of RP from SMA of Libet (and others). S-R1 and S-R2 correspond roughly to the brain areas, where Soon et al. detected nerve signals 10s before the awareness of a willful act. D-S1 and D-S2 correspond roughly to the areas stimulated by Desmurget et al. F-S is the stimulation to the SMA by Fried et al.

Finally, brute manipulation e.g., by open brain electrical stimulation [2, 15, 16] or transcranial magnetic stimulation [17, 18] can induce a wide range of experiences, including conscious will [19]. Such artificially enforced conscious will may occasionally be followed by appropriate motor actions or hallucinations of imaginary actions. Electrical stimulation to the SMA can e.g., induce conscious will connected to real or imagined motor actions [16] (Fig. 1).

3 Analysis and Computer Simulations

Experimental results, such as those briefly summarized above have been taken as evidence that free will is an illusion, since a neural signal apparently associated with the conscious will precedes it by a substantial time period. However, in a complex system such as the human brain, constantly interacting with its environment and with extensive feedback loops, it is difficult, if not impossible to determine any causal chains. We are in a continuous perception-action cycle, where cause and effect cannot easily be discerned. The situation is complicated by the different levels of organization of our nervous systems, where there are not only loops between different parts of the brain, but also between different levels (micro, meso, macro). We may not be able to say with certainty whether neural events precede mental events, or vice versa, or whether they are simultaneous.

Careful analysis of the experimental results above actually shows that several alternative interpretations may be equally probable, including (unconscious/unreported) pre-planning. Libet's experiments demonstrate a strong dependency on pre-planning, as discussed above. This means that the results in all experiments of this type, including those by Soon et al. [1], may be compromised by the presence of an unknown amount of pre-planning. In addition, when Soon et al. find a 60 % correlation between the next decision and the state of the PFC 10 s before the action, could that reflect a correlation between the memory of previous choices and the upcoming decision. It is conceivable that the next choice in the series could have a 60 % correlation with the accumulated history of the series, since humans are not very good at random number generation, which may be sought (unconsciously) by the subjects in these experiments.

Libet concluded there was a subjective backward referral, explaining the (false) sensation of a causative conscious will. Libet used, however, a threshold stimulus to the cortex. The electrical current was set so that it was just above the limit where it became noticeable. However, Pollen [20] showed that the 500 ms delay is an artifact of using threshold stimuli. Repeating Libet's threshold stimuli experiments on the visual cortex of anesthetized cats, Pollen elucidated the mechanism for time-delay of such inputs and found that neural inhibitions delays the expression of threshold stimuli. Normal sensory stimuli are presumably well over threshold and enters awareness much faster. The bulk of Libet's observations is thus neatly explained by Pollen, without revoking to the hypothesis of subjective backward referral.

Further, applying Bayesian network analysis to the Soon et al. [1] experiments, where multivariate probabilities are discussed, demonstrate that the alternative interpretation, that conscious will is causative, is also consistent with all experimental and anatomical facts.

In order to investigate the spatiotemporal relations between local and global events and processes in a cortical network, we have used computer simulations with cortical network models [21–24]. In particular, a spontaneous pulse in one part of the network may result in oscillatory (about 40 Hz) spatiotemporal patterns of activity in the entire network half a second, or more, later (see Fig. 2). This could be interpreted

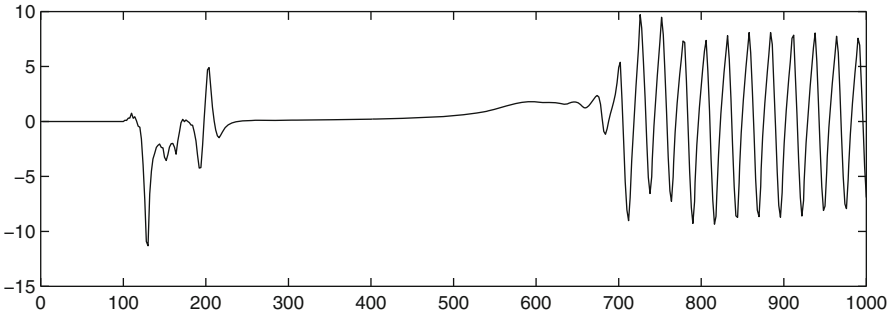


Fig. 2 Simulation results of a pulse in a cortical network node resulting in a global oscillatory dynamics of the entire network some 500 ms later. (The y-axis corresponds to EEG amplitude at an arbitrary scale, the x-axis denotes time in milliseconds)

as an intentional, spontaneous, impulse in one part of the brain that gives rise to an extensive spatiotemporal activity pattern that would correspond to a conscious experience of the intentional impulse. In this particular case, the approximately 500 ms delay between the impulse and the oscillatory pattern could correspond to the time for a subjective experience to build up. Longer time delays could be modelled with larger and more complex network models (which will be published elsewhere).

4 Discussion and Conclusions

In this paper we have focused on a series of neurophysical experiments, which has been used in the debate as evidence for an *epiphenomenal conscious will*. Many of these experiments are largely based on the findings that some "signal", e.g., the so-called readiness potential (RP), precedes a conscious will to perform a movement.

The general conclusion (e.g., [1, 25]), is that neuroscience finally has abolished free will, by showing that real decisions are made subconsciously up to 10 s before the illusory conscious act of will. A conservative interpretation of the same data is, however, that there is a weak correlation between the brain state long before the action and the experienced decision to act. This correlation could alternatively be understood as an effect of occasional unreported pre-planning, a correlation between the memory of the previous action and the next action, or a weak correlation between unconscious precursor processes and a causally connected conscious will.

With computer simulations, we have also demonstrated that a spontaneously initiated pulse in a single network node can result in a coherent global network activity much later. In reality, such an (intentional) impulse is not arising in a vacuum, but is embedded in a continuous flow of neural activity in a complex network of neurons and cortical subsystems, where the causal relationships are difficult, if not impossible to follow. Nevertheless, it is conceivable that the intentional impulse is part of a conscious experience, where the sense of a "self" acting needs time to emerge. Intention may indeed precede attention, in the perpetual action-perception cycle of consciousness exploring its environment [26].

In fact, when carefully examining the experimental procedures and results we can find no convincing evidence that conscious will is epiphenomenal. Similar conclusions have been made by others (e.g., [27–30]), including arguments from animal behavior and clinical studies of mental malfunction. Batthyany [27] points at the bias in interpretation of the experiments above, where the dominating philosophical preference suppresses alternative interpretations. The alternative hypothesis of a causative conscious will cannot be falsified or confirmed by the evidence either. In fact, it is not easy to design and perform experiments that could reveal the true nature of willful acts, especially not in an artificial environment with non-ecological tasks. One could argue that neither Libet nor Soon et al. test for free will at all, since the subjects are only asked to perform an "action" when there is an urge to move, and these movements can be said to be actions only in a very limited sense.

Indeed, the way we pose our questions, set up our experiments, and instruct our subjects, is guided by our preconceived beliefs, and hence it seems almost

impossible to get results that would contradict dominating paradigms. There is a great need for refined experiments and an unbiased analysis of the empirical evidence, which better can address the problem of conscious will in natural complex situations. The field also urgently needs a more precise and consistent terminology that avoids ambiguity and minimizes confusion. Such a terminology would facilitate interpretation and communication of hypotheses and experimental results.

Acknowledgements I am grateful for fruitful collaboration and discussions with my co-workers, notably Peter Århem, Soumalee Basu, Yuqiao Gu, and Magnus Jändel.

References

1. Soon, C. S., Brass, M., Heinze, H-J., & Haynes, J-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience* **11**, 543.
2. Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., & Sirigu, A. (2009). Movement intention after parietal stimulation in humans. *Science* **324**, 811–813.
3. Kornhuber, H. H. & Deecke, L. (1965). Hirnpotentialandeerungen bei Wilkurbewegungen und passiv Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale *Pflugers Archiv fur Gesamte Psychologie* **284**, 1–17.
4. Libet, B., Wright, E. W., & Gleason, C. A. (1982). Readiness potentials preceding unrestricted "spontaneous" vs. pre-planned voluntary acts. *Electroencephalography and Clinical Neurophysiology* **54**, 322–335.
5. Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of a freely voluntary act. *Brain* **106**, 623–642.
6. Keller, I. & Heckhausen, H. (1990). Readiness potentials preceding spontaneous motor acts: voluntary vs. involuntary control. *Electroencephalography and Clinical Neurophysiology* **76**, 351–361.
7. Haggard, P. & Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Exp. Brain Res*. **126**, 128–133.
8. Libet, B. (2004). *Mind Time*. London: Harvard University Press.
9. Ingvar, D. H. & Philipson, L. (1977). Distribution of cerebral blood flow in the dominant hemisphere during motor ideation and motor performance. Ann. Neurol. **2**, 230–237
10. Frith, C. D., Friston, K., Liddle, P.E., & Frackowiak, R.S.J. (1991). Willed action and the prefrontal cortex in man. A study with PET. *Proc. R. Soc. Lond.* (B). **244**, 241–246.
11. Ingvar, D. H. (1994). The will of the brain: Cerebral correlations of wilful acts. *J. Theor. Biol*. **171**, 7–12.
12. Spence, S. A. & Frith, C. D. (1999). Towards a functional anatomy of volition. *J. Consc. Stud.* **6**, 11–29.
13. Schultz, W. (1999). The primate basal ganglia and the voluntary control of behaviour. *J. Consc. Stud.* **6**, 31–45.
14. Christoff, K. & Gabrieli, J.D.E. (2000) The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology* **28**, 168–186.
15. Delgrado, J. M. R. (1969) *Physical control of the mind: Toward a psychocivilized society*. New York: Harper and Row.
16. Fried, I., Katz, A., McCarthy, G., Sass, K. J., Williamson, P., Spencer, S.S., & Spencer, D.D. (1991). Functional organization of the human supplementary motor cortex studied by electrical stimulation. *J. Neurosci.* **11**, 3656–3666.

17. Ammon, K. & Gandevia, S. C. (1990). Transcranial magnetic stimulation can influence the selection of motor programmes. *J. Neurol. Neurosurg. Psychiatr.* **55**, 705–707.
18. Brasil-Neto, J. P., Pascual-Leone, A., Valls-Solé, J., Cohen, L.G ., & Hallet, M (1992). Focal transcranial magnetic stimulation and response bias in a forced-choice task. *Neurol. Neurosurg. Psychiatr.* **55**, 964–966.
19. Haggard, P. (2009). The sources of human volition. *Science* **324**, 721–733.
20. Pollen, D.A. (2004) Brain stimulation and conscious experience. *Consciousness and Cognition* **13**, 626–45.
21. Liljenström H., Modeling the Dynamics of Olfactory Cortex Using Simplified Network Units and Realistic Architecture, *Int. J. Neur Syst*, **2**:1–15, 1991.
22. Liljenström, H. (2012) Mesoscopic Brain Dynamics. *Scholarpedia* 7(9):4601
23. Basu, S. and Liljenström, H. (2001) Spontaneously Active Cells Induce State Transitions in a Model of Olfactory Cortex. *BioSystems* **63**, 57–69.
24. Gu, Y. & Liljenstrom, H. (2007) A neural network model of attention-modulated neurodynamics. *Cognitive Neurodynamics* **1**:275–285.
25. Wegner, D.M. (2003). The mind's best trick: how we experience conscious will. *Trends in Cognitive Sciences* **7**, 65–69.
26. Liljenström, H. (2011). Intention and Attention in Consciousness Dynamics and Evolution. J. Cosmology **14**:4848–4858.
27. Batthyany, A. (2009). Mental causation and free will after Libet and Soon: Reclaiming conscious agency. In: A. Batthyany & A. Elitzur (eds.), *Irreducibly Conscious. Selected Papers on Consciousness*. Winter.
28. Heisenberg, M. (2009) Is free will an illusion? *Nature* **459**,164–165.
29. Sternberg, E. J. (2010). *My brain made me do it – The rise of neuroscience and the threat to moral responsibility.* New York : Prometheus Books.
30. Auletta G., Colagè I., and Jeannerod M. Eds. (2013) *Brains Top Down – Is Top-Down Causation Challenging Neuroscience?* Singapore: World Scientific.

High Frequency Oscillations for Behavioral Stabilization During Spatial Alternation

Hiroshi Nishida, Muneyoshi Takahashi, A. David Redish, and Johan Lauwereyns

Abstract It has been suggested that sharp-wave ripples (SWR) in the hippocampus contribute to memory consolidation processes. Here, we investigated behavioral performance and SWRs and gamma oscillations as rats performed a delayed spatial alternation task. We observed that the rats' behavior changed in the later trials compared to the earlier trials; the number of premature fixation breaks (trials in which the rat failed to fixate for 1 s) increased, and the duration from fixation onset to nose-poking in the choice hole reduced. SWR and gamma oscillations occurred during eating after correct choice trials. We found that the number of SWR events and the power at wide frequency range during SWR events decreased in the later trials as compared to the earlier trials. In addition, the correlation between SWR and gamma oscillations just before SWR events was higher in the earlier trials than in the later trials. Our findings support the notion that SWR serves to facilitate and stabilize the task behavior and that the inputs from CA3 and entorhinal cortex play a critical role for memory consolidation.

Keywords Hippocampus • Spatial alternation • Gamma oscillations • Sharp wave ripple

1 Introduction

High frequency oscillation in the hippocampus, accompanied with large-amplitude irregular activity, sharp-wave ripples (SWR), is thought to be involved in memory consolidation processes [1, 2]. A recent study has reported that gamma oscillations

H. Nishida
Graduate School of Systems Life Sciences, Kyushu University, Fukuoka 819-0395, Japan

M. Takahashi • J. Lauwereyns (✉)
Brain Science Institute, Tamagawa University, 6-1-1 Tamagawagakuen, Machida, Tokyo 194-8610, Japan

Graduate School of Systems Life Sciences, Kyushu University, Fukuoka 819-0395, Japan
e-mail: jan@sls.kyushu-u.ac.jp

A.D. Redish
Department of Neuroscience, University of Minnesota, Minneapolis, MN 55455, USA

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_76

co-occurred with SWR and might support the dynamic formation of coordinated CA3 and CA1 cell assemblies [3]. However, it remains unclear how the current information is reconstructed on the basis of memory, and how behavioral performance becomes stabilized in a short period.

To investigate the underlying mechanisms, we trained rats to perform a memory-guided spatial alternation task that included a 1-s fixation period and a 1.5-s delay period [4], and analyzed the local field potentials (LFP) from the hippocampal CA1 when the animal was eating after successful trials. Although the rats were well trained for the spatial choice on the alternation task, we expected that the rats' behavior would change during the fixation and the delay period in a behavioral session. In addition, if SWR and gamma oscillations contribute to memory consolidation or behavioral stabilization, these activities should show some transient changes within sessions, relating to the behavioral transition.

2 Methods

2.1 Experimental Set-Up

Four male Wistar/ST rats (weighing 280–420 g; 16–24 weeks old at the beginning of training; Japan SLC Inc., Hamamatsu, Japan) were used as subjects. The rats were trained on a delayed spatial memory-guided alternation task that included a 1-s fixation period and a 1.5-s delay period [4]. All procedures were in accordance with the U. S. National Institutes of Health guidelines for animal care and approved by the Tamagawa University Animal Care and Use Committee.

2.2 Data Analysis

As behavioral measure, we used correct choice rate, duration from fixation onset to choice response, and number of premature fixation breaks (trials in which the rat failed to fixate for 1 s).

We analyzed the LFP for the 10-s ITI period after the rat chose the correct hole, and while it remained in the half side of the box that included the reward point. SWR events were detected when the 10-ms Gaussian filtered envelope of the filtered LFP (150–250 Hz) exceeded the mean + 3SD for at least 15 ms. For the purpose of analyzing SWR associated with eating, we eliminated the data points for which the running speed was over 4 cm/s. We calculated SWR-triggered spectrograms using MATLAB (MathWorks, Natick, MA) and Chronux toolbox 2.00. The window width was set at 100-ms and moved with 10-ms steps.

The power spectrogram was normalized at each frequency band by the mean and standard deviation (i.e., Z-score) in each behavioral session. The frequency band of low gamma, high gamma, and SWR were defined as 30–45 Hz, 60–90 Hz, and 150–250 Hz, respectively.

3 Results and Discussion

We collected data from 18 behavioral sessions (112.4 trials per session with a standard deviation, SD, of 34.9). For the purpose of investigating within-session dynamics of the rats' behavior, we divided the data into two groups (the first ten trials and the last ten trials in a session) and compared behavioral performance between the earlier and the later trials. The rats' behavior changed significantly in a session even if the rats were well trained for spatial alternation. The correct choice rate in later trials was higher than in earlier trials ($p < 0.001$) and the reaching time to the choice hole reduced in later trials ($p < 0.001$). In addition, the number of premature fixation breaks increased in later trials ($p < 0.001$). Thus the rats seemed to perform the behavioral task more rapidly and to terminate the fixation nose-poking depending on internal timing after adjusting during the session.

We next examined the neural activity obtained from 31 tetrodes located in hippocampal CA1 area. First, to investigate high and low gamma oscillations during SWR events, we extracted SWR events and then calculated the SWR-triggered power spectrogram following the protocol by Carr et al. ([3]; Fig. 1a, b). Consistent with their study, we observed that the power of the gamma oscillations increased toward SWR events (Fig. 1b–d). Furthermore, we also found that low gamma oscillation positively correlated with SWR, again similar to the previous study (Fig. 1e). Subsequently, we investigated whether the neural oscillations during SWR adapted to the behavioral transitions. As shown in Fig. 2a, SWR events occurred more often in the earlier trials than in the later trials ($p < 0.001$). After computing the power spectrogram in the earlier and later trials, we calculated the power difference by subtracting the power in the later trials from that in the earlier trials. Figure 2b shows that a large fraction of the power in the earlier trials after detecting SWR events was significantly higher than that in the later trials, especially in SWR frequency band.

SWR is thought to result from a synchronized burst in hippocampal CA3 region [5]. However, some components of SWR might be caused by direct input from entorhinal cortex (EC) or by intrinsic CA1 circuitry [6]. We observed decreasing trends of power at wide frequency bands during SWR. Interpreting this result, the input strength from EC and CA3 and the number of activated neurons in CA1 might decrease. Consequently, the number of SWR to exceed threshold might reduce as well. Furthermore, our results showed that the correlation between high gamma

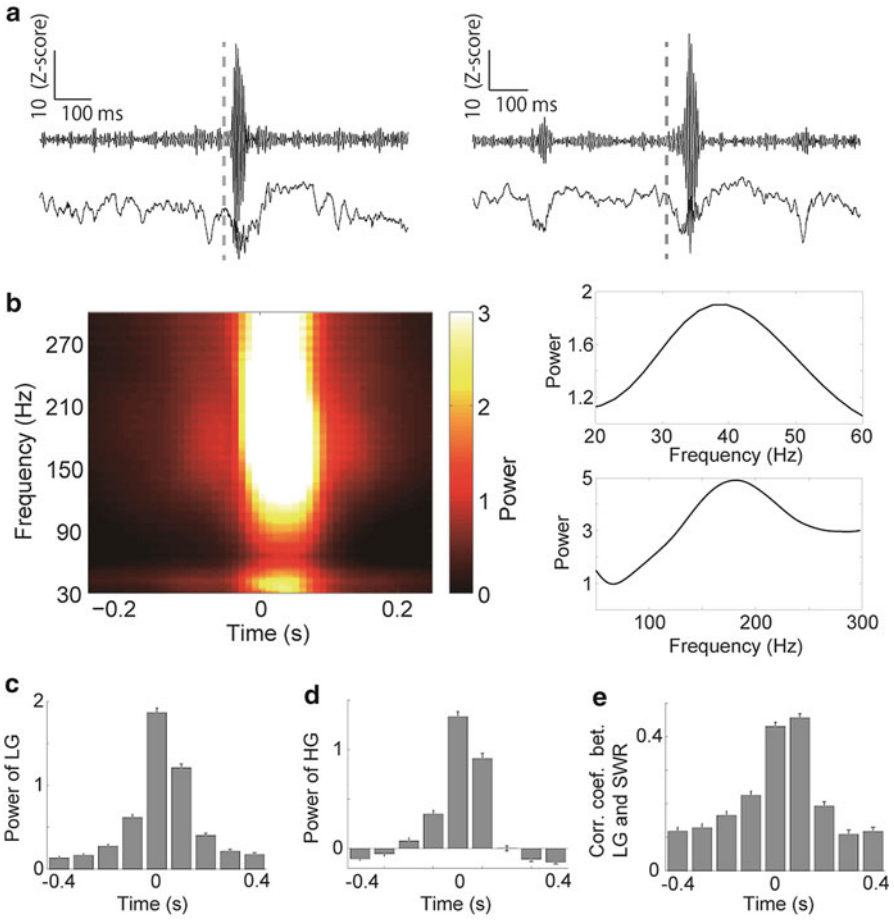


Fig. 1 (**a**) Examples of SWR; the upper signal was filtered between 150 and 250 Hz, the lower signal was non-filtered (1–475 Hz). A vertical dashed line indicates SWR detection point. (**b**) SWR-triggered power spectrogram. The right figures show the mean power of high gamma to SWR range (*upper*) and low gamma range (*lower*) over 100 ms after SWR detection. (**c**, **d**) Mean power of low gamma (**c**) and high gamma (**d**) in each 100-ms bin. (**e**) Correlation coefficient between SWR and low gamma

(which reflects input from EC) and HFO (which might be lower frequency SWR) changed in the earlier trials as compared to the later trials. Taken together, our findings support the notion that SWR serves to facilitate and stabilize the task behavior and that the inputs from CA3 and entorhinal cortex play a critical role in memory consolidation.

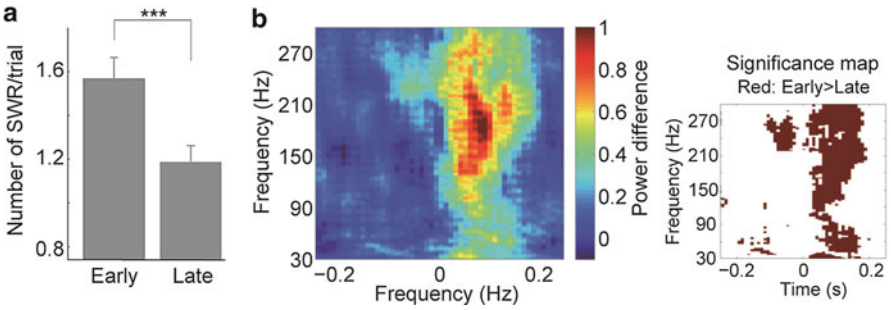


Fig. 2 (**a**) Number of SWR events per trial in earlier trial and later trials. (*** indicates $p < 0.001$ on Mann-Whitney *U* test). (**b**) The *left* figure shows the power difference between earlier and later trials. A *red color* indicates that the power is higher in earlier trials than in later trials. The *right* figure shows statistically significant results ($p < 0.05$) between earlier and later trials by Mann-Whitney *U* test. A *red color* indicates that the power is significantly higher in earlier trials than in later trials

Acknowledgments This work was supported by Human Frontier Science Program award RGP0039/2010, KAKENHI (24120710; Neural creativity for communication), Tamagawa GCOE Program, Narishige Neuroscience Research Foundation, and Grant-in-Aid for JSPS Fellows.

References

1. Buzsáki, G.: Hippocampal sharp waves: Their origin and significance. Brain Res. 398 (1983) 242–252
2. Jadhav, S. P., Kemere, C., German, P. W., Frank, L. M.: Awake hippocampal sharp-wave ripples support spatial memory. Science 336 (2012) 1454–1458.
3. Carr, M. F., Karlsson, M. P., Frank, L. M.: Transient slow gamma synchrony underlies hippocampal memory replay. Neuron 75 (2012) 700–713.
4. Takahashi, M., Lauwereyns, J., Sakurai, Y., Tsukada, M.: A code for spatial alternation in rat hippocampal CA1 neurons. J. Neurophysiol. 102 (2009) 556–567.
5. Buzsáki, G., Leung, L. W., Vanderwolf, C. H.: Cellular bases of hippocampal EEG in the behaving rat. Brain Res. 287 (1983) 139–171.
6. Remondes, M., Schuman, E. M.: Role for a cortical input to hippocampal area CA1 in the consolidation of a long-term memory. Nature 431 (2004) 699–703.

Planning Based on One's Own Past and Other's Past During a Communication Task

Jiro Okuda, Maki Suzuki, Takeshi Konno, Junya Morita, and Takashi Hashimoto

Abstract Recent behavioural and neuroimaging studies have suggested that constructing ideas about the future (prospection) is based on neural networks responsible for remembering past experiences (episodic memory). On the other hand, episodic memory involves variety of information that includes not only one's own past behaviours but also those by others. These two types of past information may influence prospection and planning of our future behaviour, especially in a social context involving communication with others. It has never been investigated, however, how the information about one's own past behaviour and those by others contribute to planning of our future behavior. In the present study, we explored evidence for differential planning processes based on one's own past behavior and that on other's past behavior. For this purpose, we used a symbolic communication task involving two subjects who sent and received arbitrary symbol messages to plan for their behaviours in the task. Our results showed that subjects' planning in the task was influenced adaptively by one's own past behaviour and those by the other subject (i.e., a partner of the task). Particularly, the subjects' behaviour favoured their own past behaviour when they sent the same message to the partner as in the previous trial, whereas their behaviour favoured the partner's past behaviour when they received the same message from the partner as in the previous trial. We speculate that such an adaptive use of differential planning strategies might contribute to successful development of an effective communication system across self and others.

J. Okuda (✉) • M. Suzuki
Faculty of Computer Science and Engineering, Khoyama Centre for Neuroscience, Kyoto Sangyo University, Kamigamo Motoyama, Kita-ku, Kyoto 603-8555, Japan
e-mail: jokuda@cc.kyoto-su.ac.jp

T. Konno • T. Hashimoto
Department of Knowledge Science, School of Knowledge Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan

J. Morita
Institute of Innovation for Future Society, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_77

Keywords Planning • Memory • Prospection • Past • Future • Communication • Language • Symbol • Co-creation • Self • Other

1 Introduction

Recent studies have clarified that imagining and planning future behaviours are closely related to recalling past episodes [1, 2]. These studies have suggested neuro-cognitive mechanisms for construction of prospective ideas by recombining retrieved one's own past experiences. On the other hand, our daily life involves not only ourselves but also others surrounding us, whose past behaviours may greatly influence our future planning. However, previous studies regarding the future planning have never tried to investigate distinction between planning processes based on one's own past and those on other's past behaviours.

In the present study, we developed a symbolic communication task involving two persons who sent and received arbitrary symbol messages with each other to decide their own trial response (spatial movement in a 2 by 2 grid). By using this task, we tried to differentiate planning based on one's own past behaviour (movement in previous trials) and that on other's past behaviour. Specifically, we analysed how subjects' planning of the movement was influenced by their own and other's movements in previous trials, in relation to messages they sent and received. Through these analyses, we explored evidence for involvement of differential planning processes during development of a symbolic communication system across two persons.

2 Materials and Method

Twenty-four young healthy subjects (mean age 22.4 years old) were anonymously paired into 12 pairs. Each pair of subjects simultaneously participated in one same task, but in different rooms. They cooperated to perform a coordination task via a game screen displayed on a computer monitor in front of each subject [3] (Fig. 1).

On the screen, each subject's initial position in the 2 by 2 grid was displayed for each subject (red or green circle). The subjects' task was to move their own position to the same position as the other subject (hereafter, 'partner') after the movement. Since they did not know the partner's initial position with each other, they were allowed to create and send a message to the partner. Each subject was able to create his/her own message as he/she liked, by choosing two symbol marks out of five possibilities (●, ■, ♦, +, and blank, see Fig. 1 for examples of the message). The subjects tried to encode information about their initial position as well as intended destination position by their own rule. They also tried to decode information about the partner's intention from the message. After the message exchange, the subjects were asked to make their own movement, followed by a screen that displayed

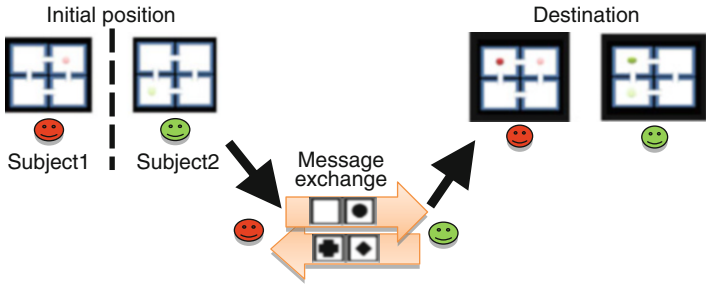


Fig. 1 Schematic illustration of a sequence of one task trial

both subjects' initial positions and destinations. Successful movement to the same position was rewarded with 1 game point to both subjects. They repeated 24 trials of this movement task. More detailed information about structure and rationale of this task was reported elsewhere [3].

By using this task, we tried to examine how the subjects' movement planning in each trial was influenced by their own and partner's past behaviours in a course of trials during which a symbolic communication system across the two subjects was gradually developed (see Konno et al. [3] for basic profiles of the development of across-subject communication protocols during the task). For this purpose, we analysed whether the subjects' decision of their destination position in each trial was the same as that by their own or the partner's destination in the previous trial, as a function of numbers of repetition of the same messages they sent or received.

Specifically, for each trial from the second to the 24th trial, we compared the subjects' destination position of that trial with their destination in the previous trial where they created and sent exactly the same message to the partner as in the current trial. We calculated percentages of trials in which they moved to the same destination as their previous destination, as a function of repetition of trials where they created and sent the same message to the partner. These data were used to index planning based on information about one's own past behaviour (planning of a movement based on information of a past message and movement the subject made). For a comparison measure to this index, we also calculated percentages of trials in which the subjects moved to the same destination as the previous partner's destination when they sent the same message to the partner.

Conversely, to evaluate planning based on *other's* past behaviour, we compared the subjects' destination position of each trial with *the partner's destination* in the previous trial where the subjects *received* exactly the same message *from the partner* as in the current trial. We calculated percentages of trials where they moved to the same destination as *the partner's past destination*, as a function of repetition of trials where the same message was *received from the partner*. This data indexed planning of a movement based on information of a past message and movement that *the partner* made. Again for a comparison measure to this index, we also calculated percentages of trials in which the subjects moved to the same destination as the

previous self destination when they received the same message from the partner. By these analyses, we tried to quantify the subjects' tendencies of planning based on the past task behaviour by themselves and that by others.

3 Results

In all the indices described above, the percentages of the same destination as the previous (one's own or partner's) destination gradually increased as repetition of the trials with the same message increased. In the trials where the subjects sent the same message to the partner, a mean percentage (across all subjects) of the same destination as previous own destination reached approximately 65 % at the third and fourth repetitions of the same messages. In contrast, a mean percentage of the same destination as previous partner's destination was at the most 40 % at the third repetition of the same messages. Analysis of variance (ANOVA) with factors of self/other (movement to previous own or partner's destination) and number of repetition of the same message the subjects sent (2nd to 6th) revealed significant main effects of the factors of self/other and repetition number ($p < 0.001$ for both effects), with no significant interaction across the two factors.

On the other hand, in the trials where the subjects received the same message from the partner, the percentages of the same destination as previous partner's destination were significantly higher than those of previous own destination. A mean percentage of the same destination as previous partner's destination reached approximately 65 % at the fourth repetition of the same messages, whereas a mean percentage of the same destination as previous own destination was at the most 50 % at the fifth repetition. Again, an ANOVA with factors of self/other and number of repetition of the same message the subjects received revealed significant main effects of the factors of self/other and repetition number ($p < 0.05$ for both effects), with no significant interaction across the two factors.

In summary, the data indicated that the subjects' planning of the movement favoured their own past movement when they sent the same message to the partner, whereas their planning favoured the partner's past movement when they received the same message from the partner.

4 Discussion

The present results showed that the subjects' planning during the task was influenced by information about one's own past behaviour and that of other's past behaviour. Specifically, the subjects gradually developed strategies to choose the same destination as the partner's previous destination as well as their own previous destination. Thus the data provide behavioural evidence for involvement of both information about one's own and other's past behaviours in planning. Moreover,

the present data suggested differential contributions of the planning based on one's own and other's past behaviours, depending on messages the subjects sent and received. That is, the planning based on one's own past was predominant when the subjects sent the same message to the partner, whereas the planning based on the partner's past movement was predominant when they received the same message from the partner. These results suggest that the subjects spontaneously imitated the partner's past movement when necessary. We speculate that such an adaptive use of the imitative planning strategy may play a pivotal role in establishment of an effective communication system across humans [4].

Acknowledgments This study was supported by Grants-in-Aid for Scientific Research on Innovative Areas (#21120007 to J.O. and #21120011 to T.H.), MEXT, Japan.

References

1. Okuda, J. et al. NeuroImage **19** (2003) 1369–1380.
2. Addis, D. R. & Schacter, D. L. Front. Hum. Neurosci. **5** (2011) 1–15.
3. Konno, T., Morita, J. & Hashimoto, T. Advances in Cognitive Neurodynamics (III) (2013), Springer, pp. 453–460.
4. Morita, J., Konno, T. & Hashimoto, T. Proceedings of the 34th Annual Meeting of the Cognitive Science Society (2012) 779–784.

Development of the Multimodal Integration in the Superior Colliculus and Its Link to Neonates Facial Preference

Alexandre Pitti, Yasuo Kuniyoshi, Mathias Quoy, and Philippe Gaussier

Abstract The question whether newborns possess inborn social skills is a long debate in developmental psychology. Fetal behavioral and anatomical observations show evidences for the control of eye movements and facial behaviors during the third trimester of pregnancy whereas specific sub-cortical areas, like the superior colliculus (SC) and the striatum appear to be functionally mature to support these behaviors. These observations suggest that the newborn is potentially mature for developing minimal social skills. In this talk, we propose that the mechanism of sensory alignment observed in SC is particularly important for enabling the social skills observed at birth such as facial preference and facial mimicry.

In a computational simulation of the maturing superior colliculus connected to a simulated facial tissue that replicate some attributes of the bio-mechanical properties of the fetus' face, we model how the incoming tactile information is used to direct visual attention toward faces. We suggest that the unisensory superficial visual layer (eye-centered) in SC and the deep somatopic layer (face-centered) in SC are combined into an intermediate layer for vision-tactile integration and that multi-modal alignment in this third layer allows newborns to detect faces and to mimic them. After we complete the learning stage within each map through Hebbian reinforcement learning, we show that the intermediate layer develops vision-tactile sensory alignment which respects the topology of the visuotopic map and of the facial map. We observe emergent properties of the global network such as sensitivity to the orientation of face-like patterns and detection of facial expressions.

Although neonate imitation is only a marker that disappears after 2–3 months in human and lets place to whole body imitation starting in human 9–12 months, we propose that the superior colliculus plays a key role in the perinatal period to bootstrap the immature cortex to develop its social abilities.

Keywords Sensory Alignment • Multimodal integration • Superior colliculus • Face modeling • Face detection • Social development • Topographic maps • Spiking networks

A. Pitti (✉) • Y. Kuniyoshi • M. Quoy • P. Gaussier
Laboratoire ETIS, UMR 8051, ENSEA, Université de Cergy-Pontoise, CNRS,
95302 Cergy-Pontoise, France
e-mail: alexandre.pitti@u-cergy.fr

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_78

1 Introduction

Neonatal imitation is perhaps the phenomenon that crystallizes the most the nature versus nurture debate. It questions us whether or not newborns possess inborn social skills, which is a radical leap out against Pagietian development that considers social cognition as the latest stage of cognitive development. Nonetheless, several evidences taken from pre-natal observations permit to infer an intermediate scenario in which sensorimotor learning at the fetal stage may give the background for a minimal social brain [1]. For instance, fetal behavioral and anatomical observations show evidences for the control of eye movements and of facial behaviors during the third trimester of pregnancy whereas specific sub-cortical areas, like the superior colliculus (SC) and the striatum appear to be functionally mature to support these behaviors [2]. These observations suggest that the newborn is potentially mature for developing minimal social skills. In this short paper, we propose that the mechanism of sensory alignment observed in SC is particularly important for enabling the social skills observed at birth such as facial preference and facial mimicry [3].

The superior colliculus has a particular neural architecture that may ease multimodal integration for simple social skills. For instance, each modality is constructed into super-imposed topographical layers that converge unidirectionally to an intermediate multimodal layer; that is there is no recurrent connections within and between the maps [4]. First, the visual map is constructed into a retinotopic layer whereas the somatotopic map is constructed into a head-centered reference frame. Second, synaptic nerves/connections from each layer, see Fig. 1.

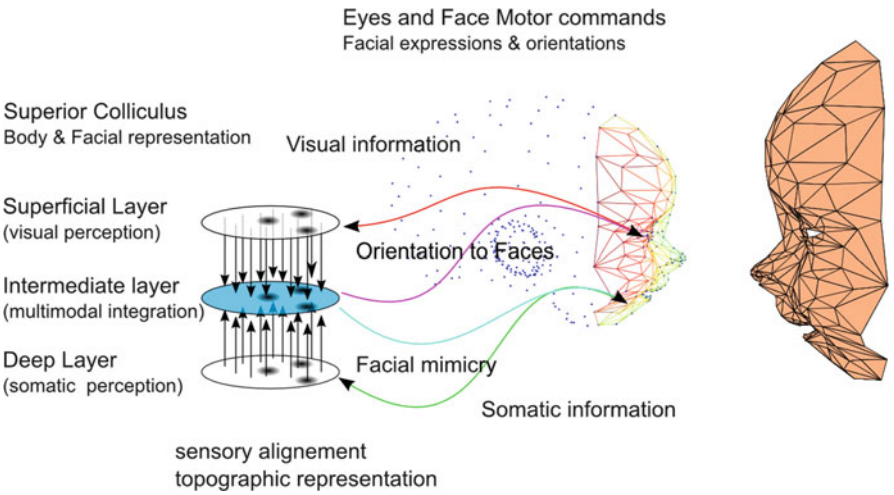


Fig. 1 Topological sensory alignment in the colliculus and its role for social development. We hypothesize that the visual and the somatotopic topographical maps bind together

2 Methods and Models

We make a computational simulation of the maturing superior colliculus connected to a simulated facial tissue that replicates some attributes of the bio-mechanical properties of the fetus' face. We model how the incoming tactile information is used to direct visual attention toward faces. We suggest that the unisensory superficial visual layer (eye-centered) in SC and the deep somatopic layer (face-centered) in SC are combined into an intermediate layer for visuo-tactile integration and that multimodal alignment in this third layer allows newborns to detect faces and to mimic them (see Fig. 1).

Neural populations are defined with integrate-and-fire neurons that capture the spatio-temporal dynamics from the two sensory modalities. The detection of structured patterns is an important attribute for preserving the topology of each modality in each map. The neural populations works similarly to a Kohonen learning systems except that we model the maturing period of SC. We add an activity-dependent mechanism based on novelty detection in order to construct the topology of the neural map by preserving at the same time the existing neurons' topology and by adding new neurons that refine it.

3 Results

After we complete the learning stage within each map through Hebbian reinforcement learning, we show that each topology respects the retinotopic topography of the eye and the somatotopic topography of the face, as seen in the SC. Then, we merge the two unisensory layers into a common intermediate layer. The multimodal layer develops synaptic links that align the visuo-tactile sensory information from each other, into a mixed spatial representation based the eye-centered reference frame and the face-centered reference frame.

As a result, when rotating a three-dots face-like pattern in front of the eye-field, we observe sensitivity of the network for certain orientations only. That is, when the three dots align well with the caricatural eyes and mouth configurational topology (i.e., facial identification), see Fig. 2a. Second, the neural activity taken from the intermediate visuo-tactile map during observation of certain facial expression like surprise and stance triggers the neurons to the characteristic visual configurational patterns of the face during rapid changes.

This situation occurs because of sensory alignment and of the high correlation with the tactile distribution of its own face, see Fig. 2b. We can imagine then that if the intermediate cells feed-forward this activity to the corresponding facial motor activity, then imitation will occur.

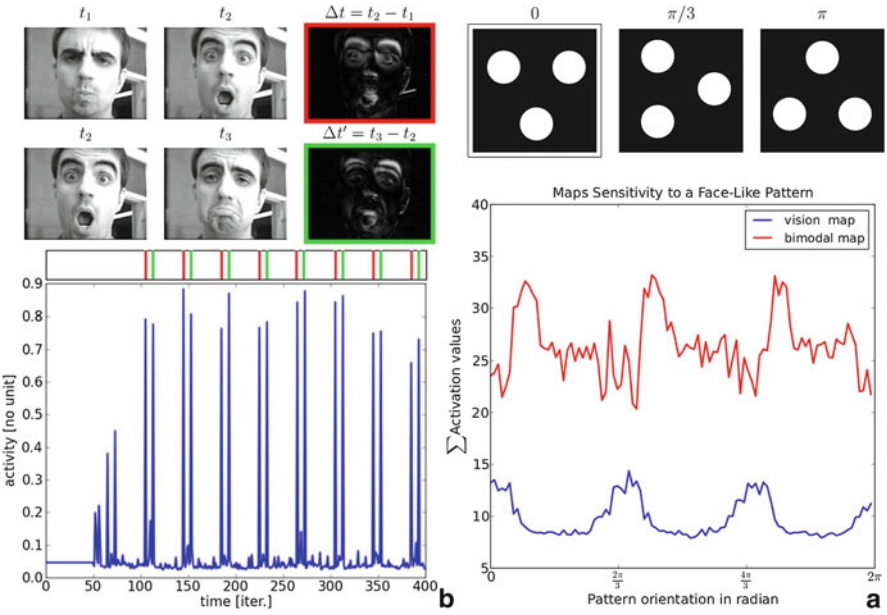


Fig. 2 Sensitivity of the multimodal topographic map to face-like visual inputs. (**a**) three dots patterns aligned. (**b**) facial expressions

4 Conclusion

We have introduced a developmental model of the superior colliculus starting from the fetal stage in the context of primitive social behaviors. In comparison to normal stimuli, we propose that faces are particular information as the visual and somatic maps in the superior colliculus are perfectly aligned topologically. We suggest that multimodal alignment may influence neonates for social skills, to recognize faces and to generate mimicry.

References

1. Johnson M., "Subcortical face processing," Nat Rev Neur, 6, 766–774, 2005
2. Stanojevic M. and Kurjak A., "Continuity between fetal and neonatal neurobehavior," J of Ultrasound in Obstetrics and Gyn (2) 3, 64–75, 2008
3. Pitti A, Kuniyoshi Y, Quoy M, Gaussier P (2013) Modeling the Minimal Newborn's Intersubjective Mind: The Visuotopic-Somatotopic Alignment Hypothesis in the Superior Colliculus. PloS ONE 8(7): e69474.
4. Stein B., Perrault Jr T.., Stanford T., and Rowland B., "Postnatal experiences influence how the brain integrates information from different senses," Front. in Int. Neuro., (30) 14, 4904–4913, 2010.

Excitation-Inhibition Balance of Prefrontal Neurons During the Execution of a Path-Planning Task

Kazuhiro Sakamoto, Naohiro Saito, Shun Yoshida, and Hajime Mushiake

Abstract The balance between excitation and inhibition in the cortex is thought to play a crucial role in the processing of information by neurons in the network. We hypothesize that the excitation-inhibition balance contributes to flexible behavioral planning in the lateral prefrontal cortex (lPFC). In this study, we analyzed the temporal development of the relative activities of inhibitory interneurons (INs) against excitatory pyramidal neurons (PNs) showing significant directional selectivity of cursor-movement at each step of the execution period in the lPFC of monkeys while they were engaged in a path-planning task. During the preparatory period, INs transiently made a greater relative contribution before enhanced cursor-movement selectivity. In contrast, no relative increase in IN activity was observed before the corresponding enhanced selectivity during the execution period. These observations suggest that the excitation-inhibition balance may contribute to planning and execution in different ways in the lPFC network.

Keywords Monkey • Prefrontal cortex • Single unit • Excitatory neuron • Inhibitory neuron

1 Introduction

Recently, the balance between excitation and inhibition in the cortex has become an important research focus, since this balance is thought to play a crucial role in the information processing of neuronal networks [1, 2]. However, the functional

K. Sakamoto (✉)
Research Institute of Electrical Communication, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan
e-mail: sakamoto@riec.tohoku.ac.jp

N. Saito
Department of Clinical Neuroscience, Yamagata University Graduate School of Medical Science, 2-2-2 Iida-Nishi, Yamagata 990-9585, Japan

S. Yoshida • H. Mushiake
Department of Physiology, Tohoku University School of Medicine, 2-1 Seiryo-cho, Aoba-ku, Sendai 980-8575, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_79

significance of this balance in flexible behavior is unclear. The lateral prefrontal cortex (lPFC) is known to be involved in flexible behavioral planning [3, 4]. We hypothesize that the excitation-inhibition balance contributes to flexible behavioral planning in the lPFC. Here, we compared the relative activities of excitatory pyramidal neurons (PNs) and inhibitory interneurons (INs) in the lPFC of monkeys while they were performing a path-planning task that required them to plan multiple cursor movements to attain a final goal in a maze on a screen (Fig. 1). In particular, we focused on the neuron types that showed selective modulation of the firing rate (FR) for cursor directions at each step of the execution period [5].

2 Methods

The details of the experimental procedures have been previously described [5–7]. Two Japanese monkeys were trained to perform a path-planning task that required step-by-step cursor movements, controlled with the manipulanda, to reach a goal in a checkerboard-like maze. Supination and pronation of each forearm was assigned to move the cursor in four directions. We recorded the single-unit activities in the lPFC of these monkeys.

Putative PNs and INs were classified based on their action potential waveform [8–10]. Spike-width was defined as:

$$\sqrt{(peak - trough)^2 + (zero\_cross \downarrow 1 - zero\_cross \downarrow 2)^2}.$$

An obvious distribution with two peaks was observed: putative PNs were defined as neurons with a spike-width of >0.8 ms, while INs had a spike-width of <0.8 ms. The index of the inhibition-excitation balance at each time was calculated using the following equation:

$$\frac{\text{average FR of INs - average FR of PNs}}{\text{average FR of INs} + \text{average FR of PNs}}$$

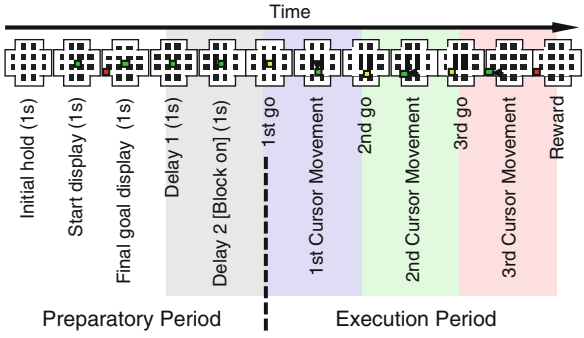


Fig. 1 The event sequence in the path-planning task. *Green*, *red* and *yellow squares* denote the cursor, final-goal, and GO signals, respectively

In addition, we used the inhibition/excitation ratio (average *FR* of INs / average *FR* of PNs) to evaluate the IN-PN balance. Statistical significance was assessed using the bootstrapping method (10,000 times).

3 Results

Of 299 neurons, 54 were classified as INs, and 245 as PNs. For the first cursor movements, 26 INs and 89 PNs showed significant selectivity during the execution period, 28 INs and 87 PNs for the second, and 25 INs and 84 PNs showed selectivity for the third.

Figure 2 shows examples of simultaneously recorded third cursor movement-selective IN (Fig. 2a) and PN (Fig. 2b). Both neurons had enhanced firing rates during the preparatory and execution periods in the trials when the monkey made leftward cursor movements at the third step. It should be noted, however, that these neurons also showed gradually elevated activity before block-onset, with slightly higher elevation detected in IN.

The relationship between the excitation-inhibition balance and third cursor-movement selectivity of the neurons (from Fig. 2) is shown in Fig. 3a–f. The mean selectivity was enhanced after block-onset in the preparatory period (Fig. 3a). Interestingly, the relative contribution of interneuronal activity is transiently

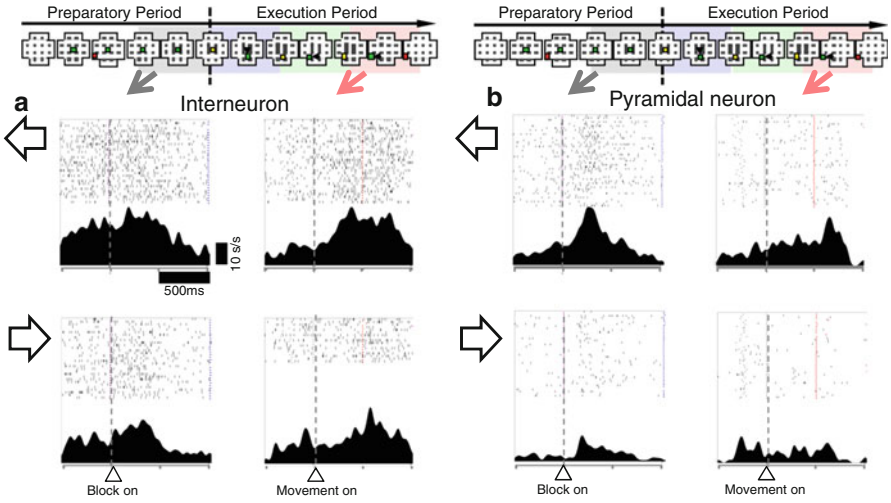


Fig. 2 An example of simultaneously recorded third cursor movement-selectivity in the neuronal pair. Activities during trials in which the monkey made leftward (*upper row* of **a** and **b**) and *rightward* (*lower row*) cursor movements at the third step are shown. *Left* and *right* columns represent preparatory and execution activities, respectively. (**a**) A putative inhibitory interneuron. (**b**) An excitatory pyramidal neuron

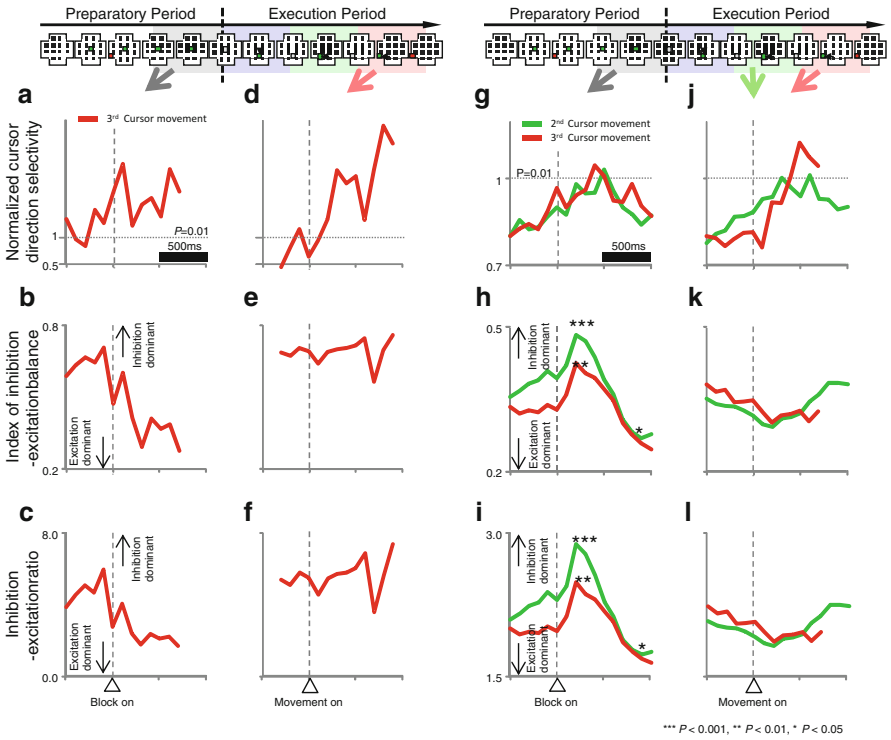


Fig. 3 Relationship between cursor-movement selectivity and excitation-inhibition balance (**a**–**f**). The neuronal pair, based on the data shown in Fig. 2. (**a**) Enhancement of selectivity after block-onset in the preparatory period. (**b**, **c**) Enhanced contributions of interneurons prior to cursor-movement selectivity. (**b**) The index of inhibition-excitation balance. (**c**) The inhibition-excitation ratio. **d** Enhanced selectivity after the onset of the third movement in the execution period. (**e**, **f**) Absence of enhanced contribution of interneurons. (**g**–**l**) The neuronal populations. *Green* and *red* indicate the second and third cursor-movement selective populations, respectively. (**g**) Enhanced selectivity after block-onset in the preparatory period. (**h**, **i**) Increased contributions of interneurons prior to the enhancement of cursor-movement selectivity. (**h**) The index of the inhibition-excitation balance. (**i**) The inhibition-excitation ratio. (**j**) Increased selectivity after the onset of movement in the execution period. (**k**, **l**) No increase in the contributions of interneurons

increased prior to enhanced selectivity (Fig. 3b, c). In contrast, the relative increase in interneuron activity was not observed at the onset of third cursor-movement (Fig. 3e, f), while third cursor-movement selectivity was enhanced (Fig. 3d).

We also carried out population analysis during the time course of the excitation-inhibition balance, and directional selectivity. Although directional selectivity was enhanced during both the preparatory and execution periods (Fig. 3g, j), the excitation-inhibition balance was transiently shifted to an inhibition-dominant state, but only in the preparatory period (Fig. 3h–l). During the execution period, the relationship between excitation and inhibition was balanced, with moderate

variability. This dynamic change of excitation- inhibition balance was consistent between the first (data not shown), second (green in Fig. 3) and third (red) cursor-movement selective populations.

4 Discussion

We analyzed the temporal development of the balance between inhibition and excitation in neuronal populations, and revealed significant cursor movement selectivity for each step of the execution period. In each population, INs made a greater relative contribution before enhanced cursor- movement selectivity in the preparatory period. In contrast, no relative increase in IN contribution was seen before the corresponding enhanced selectivity during the execution period.

lPFC is thought to play a crucial role in planning goal-oriented behavior and the monitoring of consequences [3–5]. Our findings suggest that, during the planning phase, enhanced inhibitory activity in the lPFC may be involved in shaping the top-down signals for downstream motor-related cortical areas. In contrast, a moderate excitation-inhibition balance may contribute to the maintenance of planned actions while monitoring their consequences during the execution period.

Acknowledgement This research was supported by grants #22500283, #22120504, and #24120703 from MEXT and CREST from JST.

References

1. Shadlen, M.N., Newsome, W.T.: The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J. Neurosci. 18 (1998) 3870–3896.
2. Salinas, E., Sejnowski, T.J.: Impact of correlated synaptic input on output firing Rate and variability in simple neuronal models. J. Neurosci. 20 (2000) 6193–6209.
3. Miller, E.: The Prefrontal cortex: complex neural properties for complex behavior. Neuron. 22 (1999) 15–17.
4. Tanji, J., Shima, K., Mushiake, H.: Concept-based behavioral planning and the lateral prefrontal cortex. Trends Cogn. Sci. 11 (2007) 528–534.
5. Mushiake, H., Saito, N., Sakamoto, K., Itoyama, Y., Tanji, J.: Activity in the lateral prefrontal cortex reflects multiple steps of future events in action plans. Neuron. 50 (2006) 631–641.
6. Mushiake, H., Saito, N., Sakamoto, K., Sato, Y., Tanji, J.: Visually based path-planning task by Japanese monkeys. Cogn. Brain Res. 11 (2001) 165-169.
7. Sakamoto, K., Mushiake, H., Saito, N., Aihara, K., Yano, M., Tanji, J.: Discharge synchrony during the transition of behavioral goal representations encoded by discharge rates of prefrontal neurons. Cerebral Cortex. 18 (2008) 2036–2045.
8. Wilson, F.A., O'Scalaidhe, S.P., Goldman-Rakic, P.S.: Functional synergism between putative gamma-aminobutyrate-containing neurons and pyramidal neurons in prefrontal cortex. Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 4009–4013.

9. Rao, S.G., Williams, G.V., Goldman-Rakic, P.S.: Isodirectional tuning of adjacent interneurons and pyramidal cells during working memory: evidence for microcolumnar organization in PFC. J. Neurophysiol. 81 (1999) 1903–1916.
10. Constantinidis, C., Goldman-Rakic, P.S.: Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex. J. Neurophysiol. 88 (2002) 3487–3497.

Simultaneous Multichannel Communication Using Chaos in a Recurrent Neural Network

Ken-ichiro Soma, Ryota Mori, and Shigetoshi Nara

Abstract Based on a heuristic idea and by computer experiment, we show that chaos introduced into a recurrent neural network model can enable simultaneous multichannel signal transfer as a metaphor of intra-brain communications which are experimentally observed such as certain distant fields in brain show multi synchronizations or simultaneous multi activation of neurons associated with functioning, but in intermediate regions between them, the physical channels of signal or information transfer are not observed, instead, only chaotic activities *do exist*.

Keywords Neuro-dynamics • Constrained chaos • Neural networks • Functional experiments • Multichannel communications • Intra-brain communications • Neural synchronization

1 Introduction

Nowadays, chaos have been observed in various fields. Their deterministic but unstable dynamics have been showing us interesting phenomena and attracting great interest of a large number of scientists. Since chaotic phenomena were recently discovered in living systems, in particular discovered in brain [1–3], an idea that chaos would play important roles in complex information processing in brain, and in biological systems as well, was proposed [4, 5]. In this paper, based on the same idea, a heuristic model of intra-brain communications is proposed which are experimentally observed such as certain distant fields in brain show multi synchronizations or simultaneous multi activations of neurons associated with functioning [6, 7], whereas, in intermediate regions between them, the physical channels of signal transfer or information transfer are not observed, instead, only chaotic activities *do exist*. In our model, chaotic dynamics introduced into a recurrent neural network model is used as signal transfer medium to enable us to

K. Soma • R. Mori • S. Nara (✉)
Electrical and Electronic Engineering Department, Graduate School of Natural Science and Technology, Okayama University, Tsushima-naka 3-1-1, Kita-ku, Okayama 700-8530, Japan
e-mail: nara@ec.okayama-u.ac.jp

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_80

realize simultaneous multi-channel communications, such as occurring in brain. As prototype of such communication between distant neuron groups resulting from certain learning in neural network, two or three cycle attractors are embedded in our recurrent neural network model, each of which consists of $20 \times 20 = 400$ neurons' firing pattern sequences that have specified periods. Each firing pattern of 400 neurons includes two $3 \times 3 = 9$ (always) firing neurons and the others of random firing, where the two (always firing) 9-neuron groups correspond to a sender neuron group and a receiver neuron group, respectively. It means that they are regarded as two synchronized local neuron groups in brain which are communicating each other in some on-going function. The other 382 neurons are regarded as inter-neurons to support synchronization with the use of giant redundancy. Once chaotic dynamics is introduced by reducing connectivity to one order of magnitude smaller than full connection state, then the correlation between any two neuron's pair is almost lost. However, when we apply a rather strong signal to one of the two nine neurons group which belongs to the synchronized neuron group in the embedded attractors, then without collapsing global chaos, the other (receiver) nine neurons synchronize with the sender neurons, whereas the correlations between the other neurons except the receiver and the sender, indicate still no-correlation. These results are kept same, even two or three independent inputs are applied simultaneously to two or three neuron groups belonging to the different two or three cycle attractors, respectively.

2 Method and Model

Our study works with an interconnected recurrent neural network model (abbreviated as RNNM hereafter) consisting of N binary neurons, and the updating rule is defined by

$$x_i\,(t+1) = \mathrm{sgn}\left\{ \sum_{j \epsilon G_i(r)}^{N} w_{ij} x_j(t) \right\}$$

$$\mathrm{sgn}(u) = \begin{cases} +1 & u \geq 0 \\ -1 & u < 0 \end{cases}$$

where $x_i(t) = \pm 1 (i = 1 \sim N)$ is the firing state of a neuron specified by index i at time t, w_{ij} is connection weight from the neuron x_j to the neuron x_i. w_{ii} is taken to be 0. r $(0 < r < N)$ is fan-in number for neuron x_i, named connectivity that is the most important system parameter in our work. $G_i(r)$ is a spatial configuration set of connectivity r for neuron x_i, the number of which are ${}_{N\text{-}1}\mathrm{C}_r$. Therefore, with full connectivity $r = N - 1$, determination of w_{ij} by means of a kind of orthogonalized learning method enables us to embed a group of N dimensional state patterns as

cyclic memory attractors. In our works, attractor patterns consists of (K patterns per cycle) $\times L$ cycles, and each patterns has N neurons. In this work, we take $K = 15$, $L = 2$ (see Fig. 1), or $K = 10$, $L = 3$ (see Fig. 4), and $N = 400$, where the firing states of $N = 20 \times 20 = 400$ neurons are represented by black pixel or white pixel as shown in Fig. 1. Long time updating makes an initial pattern converge into one of embedded cycle attractors. Now, when we reduce connectivity r by blocking signal transfer from other neurons, then attractors gradually become unstable, and the network state changes from the embedded patterns and chaotically wanders in the state space [8, 9]. In our computer experiments, we take the connectivity to be $r = 6$, and, in chaotic state, apply two external inputs as (Fig. 2) $[\sum w_{ij}x_j(t) + \alpha_i \cos\ (2\pi t/S_{(A,B)})_{i \in C}]$to the two $3 \times 3 = 9$ firing neurons set on the two corners (A, B), where the two periods $S_{(A,B)}$ are 137 and 181, respectively. The correlation between a sending neuron and a target (receiving) neuron are calculated as shown in Fig. 3. The correlation between the sending neuron and the other neurons are calculated and shown in Fig. 3 as well.

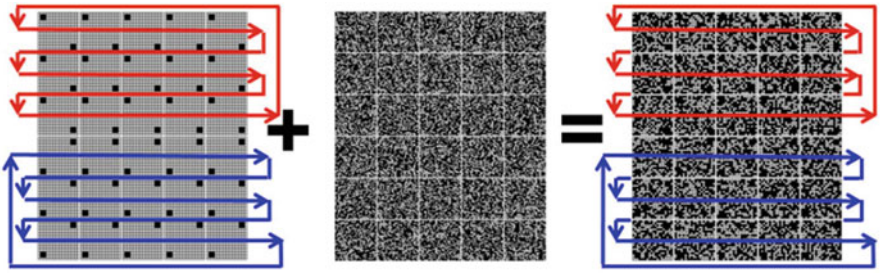


Fig. 1 The embedded two cycle attractors

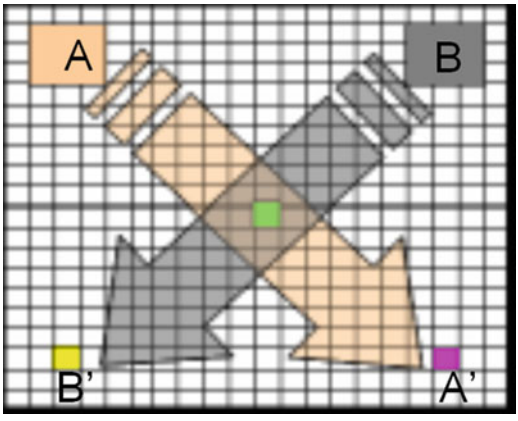


Fig. 2 Signal transfer through chaos

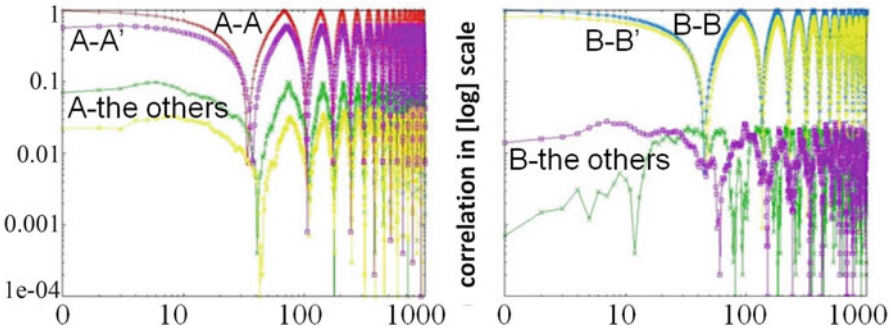


Fig. 3 The correlation function in log-log plotting

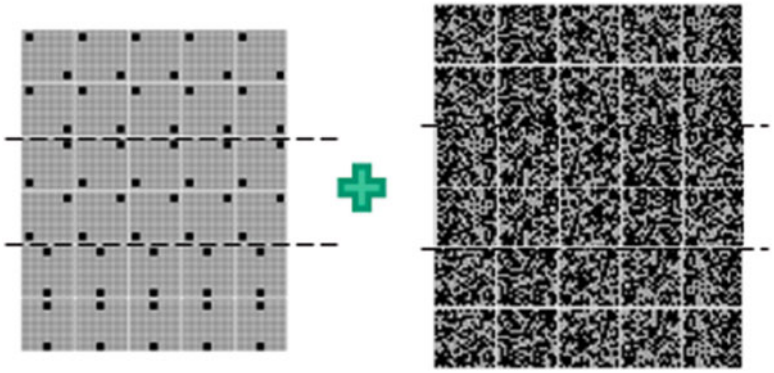


Fig. 4 The attractor patterns of three channel signal transfer

3 Results and Discussion

Let us consider more in detail about Fig. 3, the values of correlation functions between the sending neuron A(B) and the target A'(B') and the intermediate neurons as well. One can observe that the crosstalk is negligible, when we consider [log] scale in the figs. to represent correlation function f_{ij} vs. τ, where the figs. are shown using the absolute value of,

$$f_{ij}(\tau) = \left\langle x_i(t) x_j(t+\tau) \right\rangle_t \qquad \left(i, j = A^{(\prime)},\ B^{(\prime)} \text{ or the others}\right)$$

They suggest that simultaneous multi-channel communications are realized in our heuristic model systems inspired by the phenomena observed in brain dynamics. Once a fragment of learned stimulation (memorized stimuration) would be applied to the sensory neuron group, it could make the chaos constrain in the region of state space including the corresponding learned firing state (one of the learned attractors). The response of chaos to the fragmental input of the learned stimulation is quite

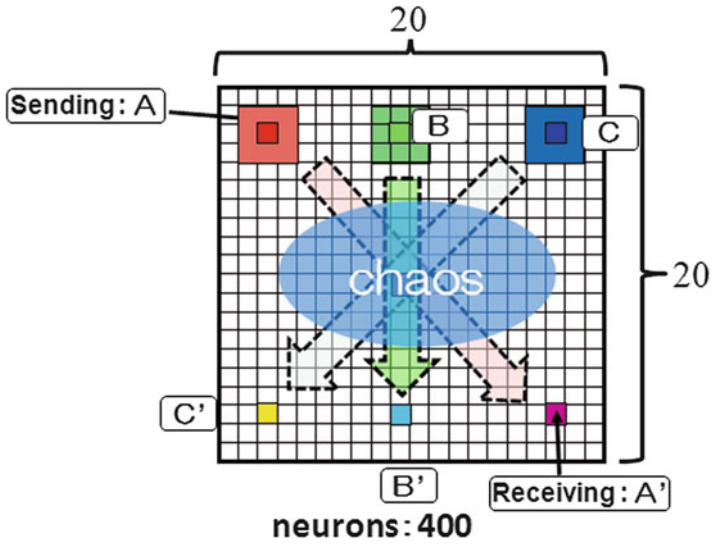


Fig. 5 A trial of simultaneous three channel signal transfer

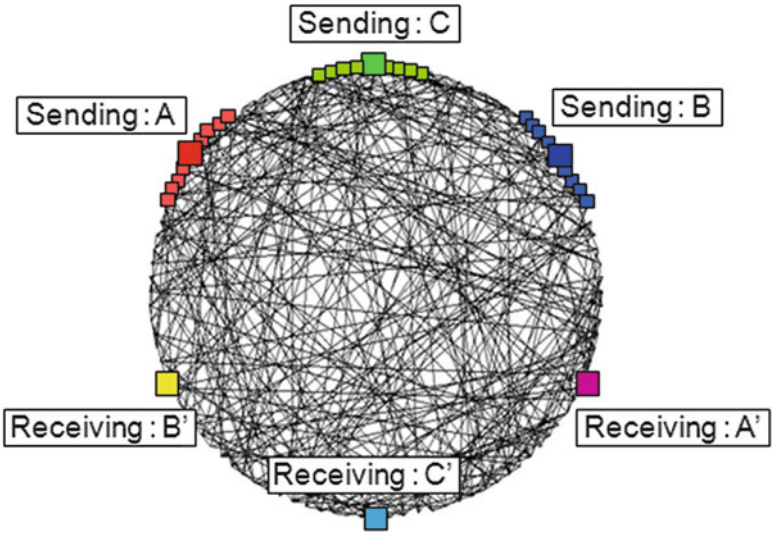


Fig. 6 A display of neuron connectivity configuration at $r = 6$

quick (sensitive), and enough to take similar or adaptively responding behaviors corresponding to learned stimulations (Figs. 5, 6 and 7).

Figure 8 shows the results of calculating the correlation functions with respect to the three pairs of the neurons corresponding to sender and receiver. One can recognize that good signal transfer without cross talk is realized in simultaneous three channel signal transfer.

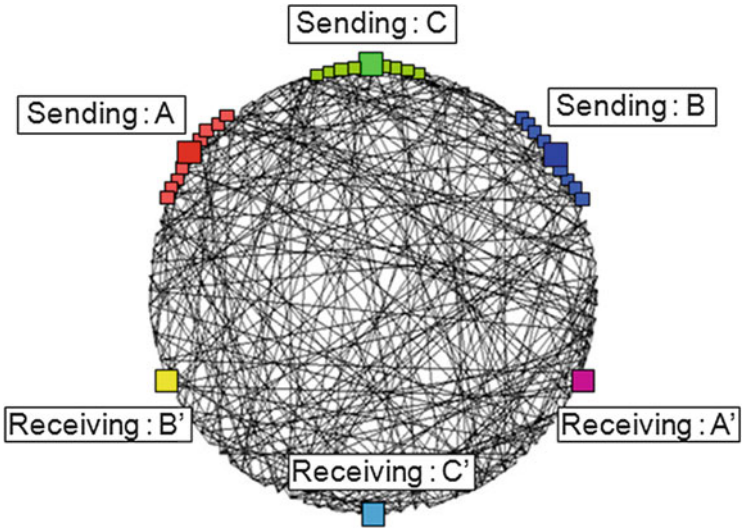


Fig. 7 The same with Fig. 6 to show the correspondences in Fig. 8. It is possible to obtain the same results in the case of *simultaneous three channel signal transfer*, which are shown in Fig. 4, 5 and 6

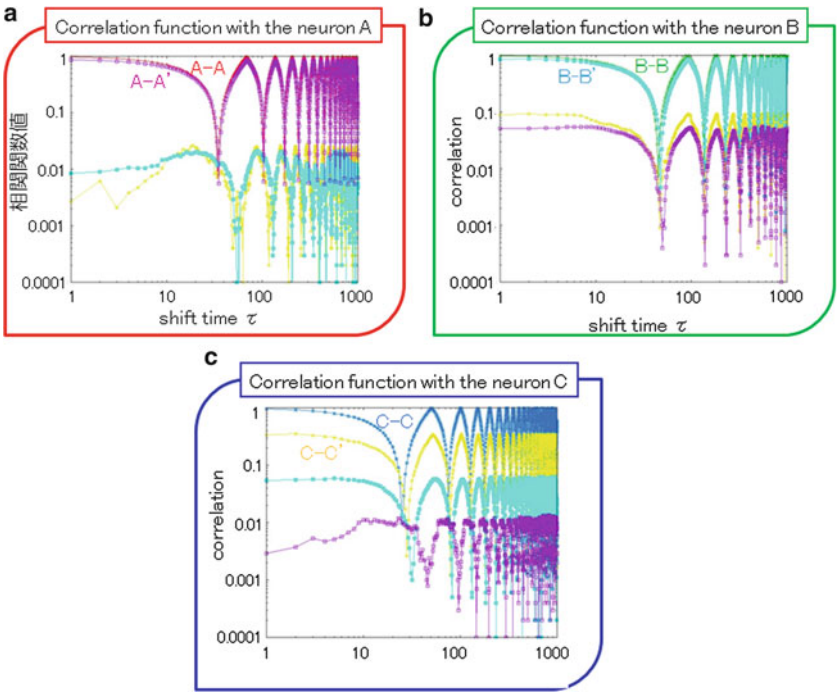


Fig. 8 (**a**) A-A is the auto-correlation. A-A′ corresponds to the signal transfer from A to A′ (**b**) B-B is the auto-correlation. B-B′ corresponds to the signal transfer from B to B′ (**c**) **C**-C is the auto-correlation. C-C′ corresponds to the signal transfer from C to C′

Acknowledgments This work is supported partly by Grant-in-Aid in the Ministry of Education, Science, Sports and Culture of Japanese Government, and partly by Co-operative Research Program of "Network of Joint Center for Materials and Devices"

References

1. H. Hayashi, S. Ishizuka, M. Ohta, K. Hirakawa - Physics Letters A, vol. 88(1982), p. 435
2. C. A. Skarda, W. J. Freeman, Behav. Brain Sci. vol. 10, p. 161 (1987)
3. K. Aihara, T. Takabe, M. Toyoda, Phys Lett. A, vol. 114, p. 333 (1990)
4. S. Nara, P. Davis, Prog. Theor. Phys. vol. 88, p. 845 (1992)
5. I. Tsuda, Behav. Brain Sci., vol. 24, p. 793 (2001)
6. C. M. Gray, P. Koenig, A. K. Engel, W. Singer, Nature, vol. 338, p. 334 (1989)
7. H. Mizuhara, Li-Qun Wang, K. Kobayashi, Y. Yamaguchi, Neuro Report, vol. 15, p. 1233 (2004)
8. S. Nara, P. Davis, M. Kawachi and H. Totuji, Int. J. Bif. & Chaos, vol. 5, p. 1205 (1995)
9. S. Nara, CHAOS, vol. 13, p. 1110 (2003)

Self-Reflective Mind in a Machine: Neural Correlates for Multidimensional Mind Perception

Hideyuki Takahashi and Takashi Omori

Abstract From three different studies, we could find the entropy in Matching Pennies Game (MPG) increasing, owing to the mind perception of the game opponent. Further, we found two components of mind perception: "mind reader" and "mind holder". We also found neural correlates related to each component in MPG. The series of our studies shed a light on the neural function of mind perception in social interaction.

Keywords Perception • Matching pennies game • Neural correlates • Social interaction • Robots

1 Introduction

We sometimes attribute a mental state to a non-animate object, such as a communication robot. Neuroimaging studies have explored neural correlates for the mind perception and these studies found that specific brain regions, such as medial prefrontal cortex, tempora-parietal junction and superior temporal sulcus, function in the mind perception [1]. Further, these brain regions are known to be activated even in the mind perception for a non-animate object, such as a human-like robot [2].

Recently, the mind perception is thought to follow multi-dimensional scales. For example, Gray and his collages proposed two dimensional scales for the mind perception, "agency" and "experience", on a basis of large-scale questionnaire surveys [3]. However, the mind perception has been dealt as a one-dimensional psychological scale in most of previous neuroimaging studies. Further there are few studies that revealed how each scale of the mind perception affects our cognition and behavior.

H. Takahashi (✉)
Graduate School of Engineering, Osaka University, Suita, Japan
e-mail: hideyuki@ams.eng.osaka-u.ac.jp

T. Omori
Brain Research Institute, Tamagawa University, Machida, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_81

In this paper, we introduce our three studies using matching pennies game (MPG) for investigating how multiple aspects of mind perception affect behavior and brain activities during the game playing.

2 About Matching Pennies Game (MPG)

MPG is a simple, zero-sum, competitive game played by two players. In this game, each player selects one decision from two options, "L" or "R," in each trial; winning/losing outcomes for each player are determined by a combination of the decisions of the two players. If both players select the same choice, one player is the winner and the other is the loser. If not, the identities of the winner and the loser are reversed. The winner receives a fixed reward, and the loser loses the same amount, in each trial. Participants played this game across multiple trials and were asked to increase their accumulated rewards as much as possible. MPG is a symmetrical, zero-sum game, and the required strategy is completely equal between the two players. Participants were required to predict the opponent's next behavior and to avoid the risk of having their behavior predicted by the opponent. In our experiment, the opponent was always a computer program, regardless of the instructions given to participants, and the program always selected each option with equal probability. Hence, the expected wining ratio was always 0.5, regardless of the participant's decision.

3 Behavioral Analysis in MPG

In MPG, generating random decisions is efficient in order for game players to avoid prediction of their next decision by their opponent: randomness of decision making is an important behavioral index in MPG. We quantified the randomness of decision making during each session of a matching pennies game as entropy H, which was calculated from the conditional frequency $p(d|c)$ of decision d (L or R) selected in current game context c (the recent choices for participants and opponents). Entropy H indicates how decision d is generated independently from the current game context, and the value of H positively correlates with the degree of randomness of decision making in a matching pennies game (see detail in [4]).

4 Study 1: Matching Pennies Game with a Human Opponent

In this experiment, we investigated how the participant's belief that his/her opponent was a human (i.e. mind holder) affects the behavior in MPG [4]. There were two conditions: the human opponent condition (HO condition), in which participants

were instructed that the opponent was a human player, and the computer opponent condition (CO condition), in which participants were instructed that the opponent was a computer program. In the HO condition, participants were instructed that their opponent was a human player without being given any extra information (e.g., how the opponent played the game). In the CO condition, participants were instructed that the opponent was a computer program.

We compared both acquired reward and behavioral entropy between the HO and CO conditions. There was no significant difference in the acquired reward between the HO and CO conditions. However, the mean of entropy in the HO condition was significantly higher than that in the CO condition. The instruction that the game opponent was a human player thus improves the randomness of participants' decision making in MPG regardless of actual game performance, and our results suggested that entropy is an efficient behavioral index for investigating the affection of mind perception in MPG.

5 Study 2: Matching Pennies Game with a Robot Opponent

In this experiment, we tried to investigate a precise factor of mind perception that affected in MPG [5]. We used a humanoid robot named "PoCoBot" as the opponent in MPG. Participants sat at a table facing the front of the robot (the distance between the participants and the robot was about 1 m). Participants were given two cards; one had "Left" and the other had "Right" printed on one side of each card. In each trial, participants selected one of these two cards and placed the selected card face down on the table. Then, the robot indicated one direction, either left or right, with gestures. Participants played MPG for 20 trials with the robot in each session. Before the session, participants had a short conversation with the robot (i.e., greeting), and the robot suddenly turned its head during the conversation. We checked whether participants followed the robot's gaze direction with their own gaze just after the robot turned its head. Participants were also asked to rate the human-likeness of the robot on a seven-point scale in each session.

We investigated whether the participant's behavior of following the gaze of the robot related to the subjective human-likeness of the robot or to the entropy of the participant's responses. We compared the means of the subjective human-likeness of the robot between sessions when the participant exhibited gaze-following behavior and sessions when the participant did not follow the robot's gaze. We did not find a significant difference between trials with and without gaze-following behavior. Next, we compared the means of the entropy of participant decision sequences between sessions with gaze-following behavior and sessions without it. We found that the mean of entropy in sessions with gaze following was significantly higher than that in sessions without gaze following. These results suggested that the entropy of participants' decision sequences tended to be high when they followed the robot's gaze regardless of the subjective human-likeness of the robot.

6 Study 3: Neural Correlates of Multiple Mind Perception in MPG

In this study, we explored the neural correlates for the multidimensional mind perception in a human-robot (or human-human) competitive game (Takahashi et al. in press). We prepared five kinds of opponents (Human, Android, Infanoid, Keepon, computer) for this study. Participants had a short conversation with each of opponents before the game and rated the impression of each opponent by a questionnaire after the conversation. After rating, participants played the matching pennies game (simple zero-sum game) with each opponent in MRI scanner. The opponent of participants was always the common computer program that selected its decision randomly regardless of the action of the participants.

From questionnaire's results, we found two psychological components related to mind perception, "mind reader" and "mind holder", by principle component analysis. We explored brain regions correlated with each of the components during the game play and found brain activities of temporal pole and the inferior part of temporal-parietal junction correlated with the degree of "mind reader" and those of medial prefrontal cortex, posterior cingulate and the superior part of temporal-parietal junction correlated with the degree of "mind reader". Further we also found that the perception of "mind reader" enhanced the entropy in MPG.

7 Conclusions

From above mentioned three studies, we could find the entropy in MPG increased owing to mind perception to the game opponent. Further, we found two components of mind perception "mind reader" and "mind holder" from questionnaire's results and the perception of the former affected the entropy in MPG. We also found neural correlates related to each component in MPG. The series of our studies shed a light on the neural function of mind perception in social interaction.

References

1. Rilling JK, Sanfey AG, Aronson J a, Nystrom LE, and Cohen JD. The neural correlates of theory of mind within interpersonal interactions. NeuroImage, 22 (4): 1694–703, 2004.
2. Krach S, Hegel F, Wrede B, Sagerer G, Binkofski F, and Kircher T. Can machines think? Interaction and perspective taking with robots investigated via fMRI. PloS one, 3 (7): e2597, 2008.
3. Gray HM, Gray K, and Wegner DM. Dimensions of mind perception. Science (New York, N.Y.), 315 (5812): 619, 2007.

4. Takahashi H, Saito C, Okada H, and Omori T. An investigation of social factors related to online mentalizing in a human-robot competitive game. Japanese Psychological Research, 55 (2): 144–153, 2013.
5. Takahashi H, Terada K, Morita T, Suzuki S, Haji T, Kojima H, Yoshikawa M, Matsumoto Y, Omori T, Asada M, and Naito E. Different impressions of other agents obtained through social interaction uniquely modulate dorsal and ventral pathway activities in the social human brain. Cortex, Vol.58, pp.289–300, 2014.

Formation of Pyramidal Cell Clusters Under Neocortical Balanced Competitive Learning

Amir Tal and Hava Siegelmann

Abstract Staggeringly complex, cortical inner-circuitry remains largely a mystery to this day. Recently, a clustered pattern has been discovered in layer 5 pyramidal cell connectivity. Distinct clusters appear at highly predictable traits across neonatal animals, implying fundamental role and perhaps a DNA-prescribed design. In this research we demonstrate natural formation of such patterns under inclusion of the two major forms of inhibition existent in neocortical layer 5, those of large basket cells and Martinotti cells. We therefore offer a simple developmental account of this seemingly innate complex structure, along with a useful biologically prevalent micro-circuit capable of complex development and learning.

Keywords Competitive learning • Brain development • Brain connectivity • Cell assemblies

1 Introduction

Standard contemporary scientific thought attributes neuronal network connectivity a pivotal role in who we are and how we function. The underlying architecture of this network is, however, still largely at debate. While connection specificity is evident on a large scale – ascribing different regions and layers of the brain unique functionality, it is still rather unclear whether single cell connectivity is predetermined or not [1, 2].

Recent studies have revealed repeating predictable sub-circuit patterns of connectivity in neocortical layer 5 network [3, 4]. Groups of pyramidal cells (PC) appear to cluster together at consistently high connectivity ratios and strengths compared to the network surrounding them, conforming to simple common

A. Tal (✉)
The Leslie and Susan Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat-Gan 52900, Israel
e-mail: amir.tal@biu.ac.il

H. Siegelmann
Computer-Science Department, Biologically Inspired Neural and Dynamical Systems (BINDS) Lab, University of Massachusetts, 140 Governors drive, 01003-9264 Amherst, MA, USA

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_82

organizational guidelines. Recurrence of such structures cross-animal at a very young age led researchers to postulate an innate tendency for it, that these structures constitute a basic building block of the neuronal network regardless of unique individual experience. This description begs the question of these clusters' origin – are they predetermined, perhaps DNA prescribed?

Apart from PCs, 20–30 % of the mammalian neocortex is made of interneurons of numerous types. In the fifth layer, two types constitute the vast majority of these cells: large basket cells (LBC) and Martinotti cells (MC) [5, 6].

LBCs are known to be the most common lateral inhibition neurons of the neocortex, extending expansive lateral axonal arborizations to neighboring and distant columns within their layer [7]. For this reason, they are frequently attributed imposing Mexican hat shaped inhibition in the neocortex [8], dynamics which are prominent in neural network literature [9].

MCs, in contrast, are characterized by their long ascending axon to layer 1, which primarily targets thick tufted PC dendrites [5]. Electrophysiological study revealed MC mediate a unique inhibitory pathway of sub-columnar dimensions. MCs provide prolonged feedback to excitatory PC neighborhoods which stirred it to action, within a distinct delay from activity onset [10–12].

Putting together PC, LBC and MC neurons, a comprehensive simulation of neocortical layer 5 circuitry is possible. Under synaptic spike-timing-dependent plasticity rules (STDP) [13] we have found this circuitry inevitably develops clustered connectivity patterns in an excitatory PC network, bearing similar traits to those empirically found. MC inhibition with its unique spatial and temporal traits constitutes a substantial addition to classic local-excitation lateral-inhibition connectivity, forming a useful competitive circuit computationally analogous to self-organizing maps [14].

2 Results

2.1 Cluster Characteristics

A neural network simulation was developed to model layer 5 PC, LBC and MC connectivity. Leaky integrate-and-fire model was chosen as the neuronal model and connection probabilities and strengths were initialized based on electrophysiological data [1, 15]. On initialization, network is connected at 12.22 ± 0.31 % based on somatic distance-dependant connection probabilities. Simulation causes a drop of this rate to 7.77 ± 0.24 % on average, however leaving PC-PC connectivity probability distance-dependent. When applying Affinity Propagation algorithm [16] to this network at the same configuration used in Perin et al. [3], 40.4 ± 1.71 clusters are detected, at an average size of 24.8 ± 1.06 members per cluster. For comparison, a rerun of the original simulation by Perin et al. for 1,000 neurons results in clusters of 31.58 ± 2.29 members.

Clusters yielded an average connectivity ratio of 30.06 ± 0.86 %, almost four times higher than the overall network connectivity ratio. This is significantly higher than the simulation outcome of Perin et al. [3], where clusters of a few dozen neurons were connected at an average 16 % alone. However, according to electrophysiological data reported in the paper, connectivity ratio seems to rise with group size, already peaking at 12.5 % (7 out of 56) in groups of eight neurons. This may imply higher connectivity in larger clusters such as ours, but is unfortunately unverifiable as groups were restricted to 12 neurons in the original study.

In terms of synaptic strength, we compared mean synaptic strength of synapses bridging different clusters with those connecting neurons of the same cluster. While the amount of the first kind reaches even ten times the amount of inner-cluster synapses, synaptic strength tends to distribute rather similarly in the two groups, with tremendous variance in both (Fig. 1). This bimodal distribution of synaptic weights is common in naïve forms of STDP learning, and may be balanced using additional parameters to plasticity [17], a complexity we thought unnecessary for our model. Despite this large variability, strong synapses constitute a significantly higher percentage of inner-synapses than outer-synapses. Synaptic strength inside clusters is on average 3.43 ± 0.31 mV (in EPSP), while the average synapse outside clusters is of 2.29 ± 0.18 mV EPSP strength.

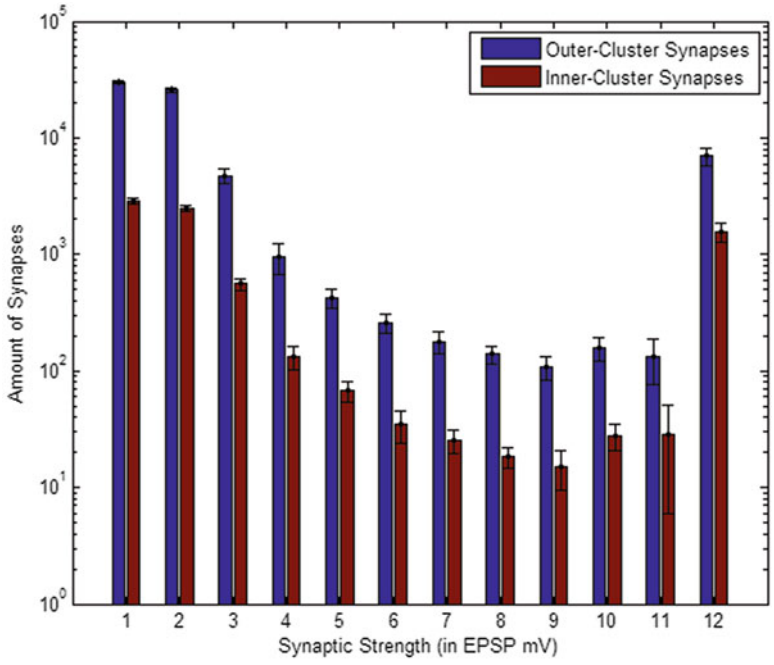


Fig. 1 Synaptic strength distribution. Histogram of outer-cluster synapses (shown in *blue*) compared with inner-cluster (*red*) on a log scale displays the bimodal distribution all synapses demonstrate, as well as the difference in volume between both types. Strong synapses constitute a larger fraction of inner-cluster synapses than outer-cluster synapses. STD bars shown in *black*

2.2 Common Neighbor Rule

Number of common neighbors (NCN), which is the principle guideline for cluster identification in this paradigm, therefore proves to be a good predictor of synaptic strength and connection probability. In consistence with what has been reported in Perin et al.[3], the more common neighbors two neurons share, the more likely they are to be connected, and the stronger the connection between them tends to be. Compliance to NCN is especially noteworthy as our networks stem from random distance dependent connectivity and Hebbian principles alone. NCN is used only as a retrospect analysis tool, in contrast to the work done by Perin et al., where it was also used as the guideline for network wiring.

2.3 Network Topology

Networks which underwent simulation are left with a degree distribution which is binomial looking with a positive skew. Median falls very close to the mean value, but right tail is evidently longer than left one (Fig. 2). Further classification of the network based on clustering coefficient and average shortest path as described in Watts and Strogatz [18], reveals that traits resemble those of a small-world

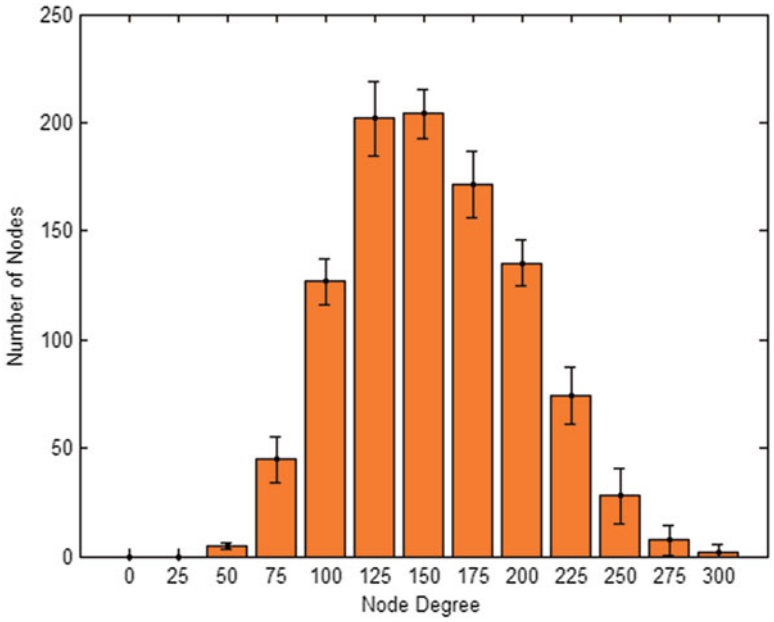


Fig. 2 Avg. network degree distribution. – median = 151.7 ± 4.8 – mean = 155.3 ± 4.7 – skewness = 0.24 ± 0.1

Table 1 Comparison of network clustering coefficient and average shortest path with two networks bearing same number of edges – a completely regular network on one hand and a completely random one on the other. Our network falls in the characteristics of a small-world network due to a high clustering coefficient relative to its short average path

	Clustering coefficient	Avg. shortest path
Regular network	0.74	7.08
Random network	0.07	1.92
Our network	0.13	1.97

network (Table 1). While the average shortest path of our network is almost the same as in a completely random network, clustering coefficient is almost twice that of a random network. Small-world network topology is found in abundance in nature. It indicates the network has a small degree of separation within it, but still connectivity is not random. Nodes connect to each other based on some logic, and the resulting clustering coefficient is higher than would be expected by chance.

2.4 Effect of Martinotti Inhibition

Isolating MC inhibition from simulation, it seems the chief influence of this pathway is the regulation of neuron activity in one area over the other. While BC inhibition is the core cause of lateral competition and local clustering, simulations lacking MC inhibition result in a highly uneven network. Without this type of inhibition, selected areas of the network grow constantly stronger with less and less competition. Lacking MC inhibition, an average of three clusters (7 %) account for over half of the spikes produced by the network, one of them over 25 % by itself (Fig. 3). This cluster reigns supreme over an obvious hierarchy shaped within the network, in terms of firing volume and average synaptic strength alike.

3 Discussion

3.1 LBC-MC Circuitry

The dual inhibition described in this research accounts for clustering of Pyramidal Cell neurons, due to dynamics of competitive activity which it promotes. Regions rival each other over different input via LBC lateral inhibition, earning input selectivity in a manner reminiscent of Self Organizing Maps. MC inhibition steps into this competitive scheme to restrict "winning" regions for a certain amount of time after they have been active. This self inhibition is crucial for an equal spread of activation in the network. Without it, most network connectivity will degenerate and the network will be deprived of input selectivity. MC behavior seems to be solving

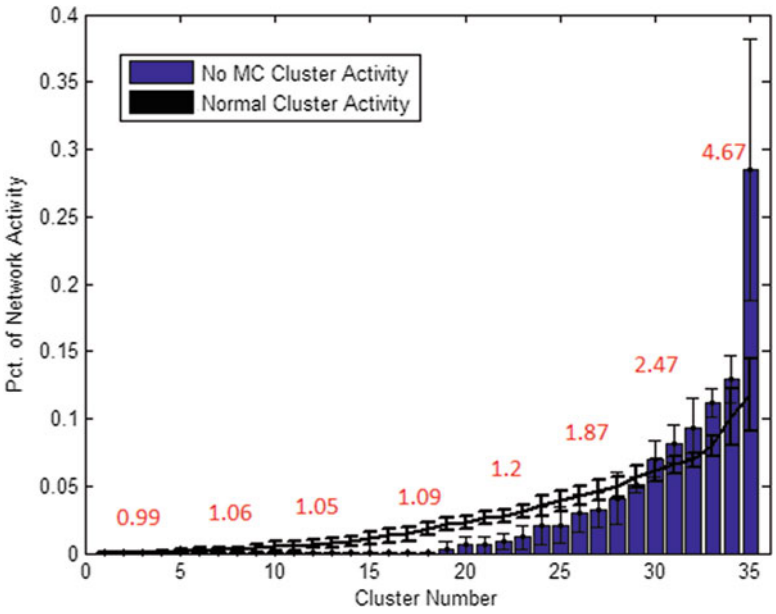


Fig. 3 Effect of MC pathway. Sorting clusters by spike volume reveals that lack of MC (*blue bars*) results in a much less even distribution of activity between clusters. One cluster produced over 25 % of network activity alone, compared to the avg. 10 % maximum clusters reach in normal simulations (*black line*. STD bars drawn vertically). *Red* numbers indicate average cluster synaptic strength (in mV EPSP), which rises dramatically in correlation with firing volume

the AI learning problem of "dead units". Dead units are a problem suffered by hard competitive learning algorithms in which computational units may be neglected and not used in the final outcome of learning. MC seems to be analogous to a learning "conscience", restricting overly-successful regions to make competitive learning "fair" [19].

In this light, PC-BC-MC wiring may be a connectivity pattern selected throughout evolution to give neuronal networks an intrinsic inclination to learn and accommodate themselves to incoming dynamic patterns. Each neuron can potentially be a cluster center to begin with. It is the unique stimulus history which will form certain clusters and not others in different networks following different experiences, but in agreement with the insight also given in Perin et al. [3], a clustered layout seems to be inevitable under these dynamics.

3.2 Clustered Network Topology

Simulated networks show traits most similar to small-world networks, maintaining an average short path between any two nodes under a relatively clustered layout.

This structure seems to afford quite an optimal balance between network transmission speed and local connectivity.

As the speed in which neuronal commands may be transmitted is bounded first and foremost by chemical synaptic transmission speed, short paths are a significant advantage in a living neural network. Achieving such connectivity with a binomial degree distribution means the network does so without the usage of global hubs, which is critical both for parallel processing, but moreover for robustness – as the network is not dependent upon any single neuron.

When considering high clustering, the most immediate advantage seems to be, like in the hub-free rational, network robustness. If we regard clusters as basic unified units, the system becomes resilient to single neuronal failures thanks to redundant connectivity. However, another attractive possibility might be drawn from the social network metaphor – an additional type of transmission speed. According to a study of epidemic processes in clustered networks, increase in network clustering coefficient leads to a decrease of epidemic threshold [20]. Clusters make it easier for information to percolate in the network, reaching new nodes significantly quicker and easier. In neural networks this would mean requiring weaker stimuli to recruit neurons to fire, so being more sensitive to input.

References

1. Fino, E., & Yuste, R. (2011). Dense inhibitory connectivity in neocortex. *Neuron*, *69*(6), 1188–203. doi:10.1016/j.neuron.2011.02.025
2. Hill, S. L., Wang, Y., Riachi, I., Schürmann, F., & Markram, H. (2012). Statistical connectivity provides a sufficient foundation for specific functional connectivity in neocortical neural microcircuits. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(42), E2885–94. doi:10.1073/pnas.1202128109
3. Perin, R., Berger, T. K., & Markram, H. (2011). A synaptic organizing principle for cortical neuronal groups. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(13), 5419–24. doi:10.1073/pnas.1016051108
4. Song, S., Sjöström, P. J., Reigl, M., Nelson, S., & Chklovskii, D. B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS biology*, *3*(3), e68. doi:10.1371/journal.pbio.0030068
5. Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., & Wu, C. (2004). Interneurons of the neocortical inhibitory system. *Nature reviews. Neuroscience*, *5*(10), 793–807. doi:10.1038/nrn1519
6. Rudy, B., Fishell, G., Lee, S., & Hjerling-Leffler, J. (2011). Three groups of interneurons account for nearly 100 % of neocortical GABAergic neurons. *Developmental neurobiology*, *71*(1), 45–61. doi:10.1002/dneu.20853
7. Wang, Y., Gupta, A., Toledo-Rodriguez, M., Wu, C. Z., & Markram, H. (2002). Anatomical, physiological, molecular and circuit properties of nest basket cells in the developing somatosensory cortex. *Cerebral cortex (New York, N.Y.: 1991)*, *12*(4), 395–410.
8. Casanova, M. F., Buxhoeveden, D., & Gomez, J. (2003). Disruption in the inhibitory architecture of the cell minicolumn: implications for autism. *The Neuroscientist: a review journal bringing neurobiology, neurology and psychiatry*, *9*(6), 496–507. doi:10.1177/1073858403253552

9. Kang, K., Shelley, M., & Sompolinsky, H. (2003). Mexican hats and pinwheels in visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(5), 2848–53. doi:10.1073/pnas.0138051100
10. Berger, T. K., Silberberg, G., Perin, R., & Markram, H. (2010). Brief bursts self-inhibit and correlate the pyramidal network. *PLoS biology*, *8*(9). doi:10.1371/journal.pbio.1000473
11. Kapfer, C., Glickfeld, L. L., Atallah, B. V, & Scanziani, M. (2007). Supralinear increase of recurrent inhibition during sparse activity in the somatosensory cortex. *Nature neuroscience*, *10*(6), 743–53. doi:10.1038/nn1909
12. Silberberg, G., & Markram, H. (2007). Disynaptic inhibition between neocortical pyramidal cells mediated by Martinotti cells. *Neuron*, *53*(5), 735–46. doi:10.1016/j.neuron.2007.02.012
13. Feldman, D. E. (2012). The spike-timing dependence of plasticity. *Neuron*, *75*(4), 556–71. doi:10.1016/j.neuron.2012.08.001
14. Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*(1), 59–69.
15. Packer, A. M., & Yuste, R. (2011). Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? *The Journal of neuroscience: the official journal of the Society for Neuroscience*, *31*(37), 13260–71. doi:10.1523/JNEUROSCI.3131-11.2011
16. Frey, B. J., & Dueck, D. (2007). Clustering by passing messages between data points. *Science (New York, N.Y.)*, *315*(5814), 972–6. doi:10.1126/science.1136800
17. Van Rossum, M. C., Bi, G. Q., & Turrigiano, G. G. (2000). Stable Hebbian learning from spike timing-dependent plasticity. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, *20*(23), 8812–21.
18. Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*(6684), 440–2. doi:10.1038/30918
19. DeSieno, D. (1988). Adding a conscience to competitive learning. *IEEE International Conference on Neural Networks*, (c), 117–124 vol. 1. doi:10.1109/ICNN.1988.23839
20. Newman, M. E. J., & Web, W. (2003). Properties of highly clustered networks. *Physical Review E*, *68*(2), 1–7.

Systems Biopsychiatry: The Cognitive Neurodynamics of Schizophrenia in an Integrated Systemic View

Felix Tretter

Abstract The atomistic research program in biological psychiatry has to admit increasingly that only a systemic re-construction of the brain might help to understand mental functions and dysfunctions ("Computational Systems Biopsychiatry"). New technologies that record multiple units by complex data analyses implicate multiple network concepts of brain structures and functions (e.g. graph theoretical models). However, these descriptive quantitative models are not sufficient to "explain" complex mental disorders such as schizophrenia. For this reason computer-based process models are necessary. A sketch of this multi-level modelling task is given by referring to "dysconnectivity hypothesis" of schizophrenia.

Keywords Biological psychiatry • Schizophrenia • Dysconnectivity hypothesis • Multi-level modelling • Brain-mind problem

1 Introduction

According to WHO, mental disorders (e.g. depression and schizophrenia) are the leading diseases that have a high socioeconomic impact on society [1]: For this reason, progress in psychiatric research resulting in effective medications with high efficacy, a low rate of side effects and low costs is very important. These psychopharmaceuticals should be based on comprehensive knowledge of brain function and dysfunction. Still, we are far away from such a conceptual framework of understanding mental disorders in the context of biological psychiatry. Therefore, it could be useful to develop a systemic multi-level view of the brain that integrates different segments of knowledge and that is based on applications of mathematical tools in data analysis and computerized modelling.

F. Tretter (✉)
Bawarian Academy of Addiction, Landwehrstr. 60-62, 80336 Munich/FRG
e-mail: felix.tretter@bas-muenchen.de

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_83

2 Methods

In psychiatric research today, two main methodological approaches are dominating: functional and structural imaging studies of the macro-anatomical circuitry and molecular biological research of the brain with emphasis on "omics" studies (genomics, proteomics, metabolomics etc.; Fig. 1). These high-throughput technologies generate extremely large data sets. For this reason, new mathematical tools that complement multivariate statistical methods must be applied and developed. As complexity of the brain mechanisms and their dynamics challenge human reasoning, the computational systems modelling approach seems to be unavoidable [2].

Additionally, neuropsychiatry also has to tackle the basic problem to relate mental phenomena to brain mechanisms (brain-mind problem). Also behaviouristic psychopathology still has to admit that subjective experience is the essential basis of mental disorders: In various stages of mental illness, the patients report strange experiences that in many cases cannot be proved by objective tests (e.g. stupor in depression and schizophrenia). Regarding this, mental disorders roughly can be characterized as an imbalance or dissociation of conscious experience regarding states and processes of cognition and emotion and their subfunctions.

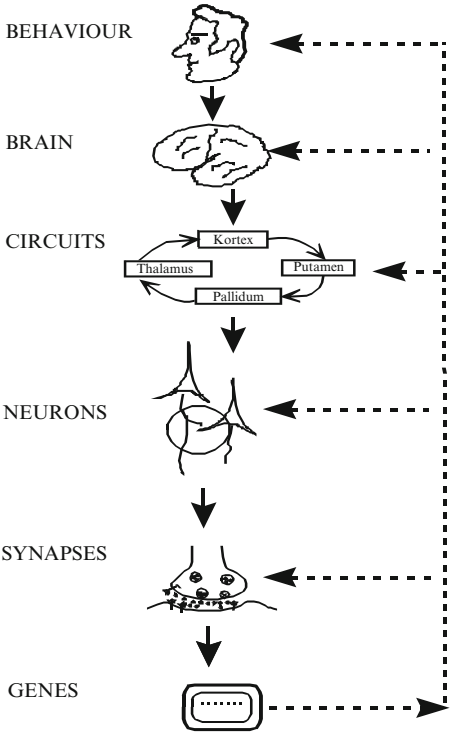


Fig. 1 The multi-level approach in systemic neuropsychiatry – top-down deconstructing analysis by biotechnologies and bottom-up reconstructing synthesis by computer-based modelling and simulations

2.1 Schizophrenia as a Systemic Mental Disorder

Schizophrenia is characterized by a complex symptomatology with a high variation of the clinical manifestation in each individual patient [3]. For instance, symptoms like dysfunctions of perception (hallucinations), thought disorders and delusions are very significant. Additionally, ambivalence, autism and affective disorders are observed. All these symptoms are dissociated, which is a basic feature of schizophrenic information processing [4].

Regarding this complexity of symptoms, a systemic approach of modelling psychopathology should focus on a few symptoms, and should not claim to describe the whole disease [5]. In line with this, multi-level modelling of neurobiological data on deficiencies of working memory functions is a fruitful example [6, 7].

For our aim to understand psychiatric illnesses by a system theoretic framework, it is important to transform the symptoms into a quantitative time-related language of functions and dysfunctions so that it is possible to match this type of characterization of the symptomatology of schizophrenia to the activity patterns that are obtained by neurobiological methods: For instance, a *hallucination* can be defined as a stimulus-free perception that is uncoupled from a sensory system. In this case, a hyperreactivity to sensory stimuli can be observed objectively (reduced prepulse inhibition;[8]). In neurobiological terms this function corresponds to a spontaneous activity or a hyperreactivity of a local perceptual network (e.g. temporal cortex) that exhibits a too strong self-activation compared to the strength of the self-inhibition. Also *thought disorders* can be understood as a "too many" or "too few" branches of the sequence of the thoughts. Also a too weak or too strong and too short or too long persistence of information in the working memory buffer can be the basis of a dysfunctional working memory function [6, 7; comp. 9].

At present, there is no complete transformation of psychopathological terms into some sort of a functionalistic systemic language as it was demonstrated here. However, this reconceptualization of basic categories of psychopathology and also the construction of a systemic concept of mental functions could be one way for theoretical psychiatry to provide a better match with neurobiological data (isomorphic relation), at least on a micro-anatomical level. Mental functions correspond to local and total brain functions and not only to functions of single neurons or to gene functions.

2.2 Schizophrenia as a Neural Network Disorder

Here we start with a vague concept of the global brain as a neural network, assuming that nearly every macroanatomically defined brain area is directly connected with each other. In this view, we focus on 11 components that are frequently mentioned in publications about relevant schizophrenic symptoms. Here we represent some of their relevant connections (Fig. 2). Although the Human Connectome Project

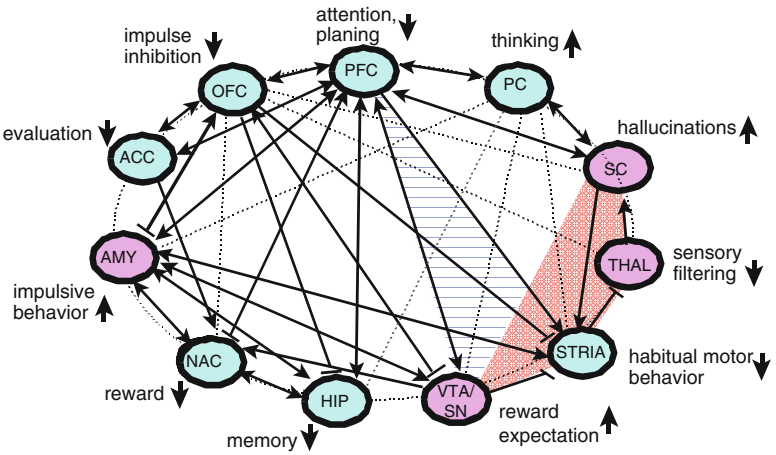


Fig. 2 Hypothetical brain networks of schizophrenia: squared red subnetwork VTA/SN-STRIA-THAL-SC-STRIA proposed by Carlsson [11, 12], lined blue subnetwork VTA/SN-STRIA-PFC-VTA/SN proposed by Weinberger and Winterer [13]. Abbreviations: *OFC* orbitofrontal cortex, *PFC* präfrontal cortex, *PC* parietal cortex, *SC* sensory cortex, *THAL* thalamus, *STRIA* striatum, *VTA* ventral tegmental area, *HIP* hippocampus, *NAC* nucleus accumbens, *AMY* amygdala, *ACC* anterior cingulate cortex *Red* higher activation, *blue* lower activation comp. to healthy subjects

is recording the connectivity of the whole brain [10], we reduce this complexity in order to demonstrate some conceptual and methodological issues: In this global framework some essential models of the schizophrenic brain circuitry are integrated and can be seen as models of subsystems of the whole network of supraregional circuits (Fig. 2). For instance, in the history of neurobiological explanation of schizophrenia one of the first successful hypothetical models was proposed by Arvid Carlsson [11, 12]. In this model, Carlsson proposed a circular dopamine-driven dysfunction: a hyperactivity of the dopamine system that starts in the brain stem (substanzia nigra, SN and ventral tegmental area, VTA) and very strongly inhibits the "striatal complex" (here: STRIA) via inhibiting D2 receptor activation. From here, GABA-based inhibitory projections start to the thalamus (THAL) that provides input for the sensory (visual and auditory) cortices (SC). These (and other) cortical areas exhibit activating projections back to the striatal complexes. In case of schizophrenia, the striatal inhibitory output is weak so that thalamus is hyperactively driving the sensory cortices (Fig. 2a). Several years later a "corticocentric" conception of the neural sources of schizophrenic symptoms was proposed by Daniel Weinberger and his co-workers [13]: Functionally weak prefrontal cortical projections down to the brain stem would not be able to inhibit the up-stream dopaminergic projections to the striatum and also to prefrontal cortex so that these projections exhibit a hyperactivation of the D2 receptors. In prefrontal cortex a pathological functional dominance of D2 receptors in relation to D1 receptors could evoke a weak working memory function (too much distraction). Additionally, imaging studies using path analysis found that a weak connectivity between frontal

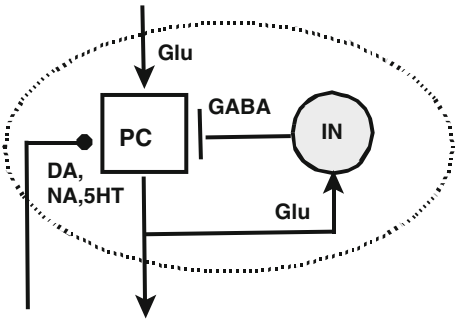


Fig. 3 Core circuit as can be seen in models proposed by Grace and coworkers ([19] for Hippocampus and by Lewis et al. [15] for Cortex. Abbreviations: *Glu* glutamate, *GABA* gamma-amino-acid, *DA* dopamine, *NA* norepinephrine, *5HT* serotonin

brain areas and temporal brain areas are typical for schizophrenia and could explain other cognitive impairments (not depicted in Fig. 2; [14]). Reviewing such articles in detail, a realistic picture of all peculiarities of the "dysfunctionally connected" schizophrenic brain network could be constructed.

On a micro-level, the corticocentric concept of schizophrenia was affirmed by the latest finding that in the prefrontal cortex a insufficient local circuitry consisting of pyramidal cells and inhibitory neurons could be the cause of a weak activity of this cortical area [15]. The core circuit of this micro-model is based on a two-component model of reciprocal connectivity consisting of an activating unit and an inhibiting unit (Fig. 3). The activity of such modules easily can oscillate [16]. Therefore this concept is in line with the finding of reduced gamma-activity in cognitive tasks in schizophrenic patients [17, 18].

Looking one level deeper, considering the genetic level, pathological properties were found for synaptic mechanisms such as degrading enzymes, e.g. a weak COMT by gene 22q11 with a Val and Met polymorphism in case of the Met allele implicates a weak COMT enzyme and in consequence a high dopamine concentration (comp.[20]). Also genes like neuregulin, dysbindin, DISC1, RGS4, GRM3 and G72 are susceptibility genes for schizophrenia.

From a molecular biological point of view the intracellular molecular signaling network in schizophrenia is still not completely understood. Simply speaking we would expect "schizophrenic neurons" but only in the sense of a certain intrinsic dynamic imbalance of activating and inhibiting mechanisms that generate pathological hypo- or hyperactivity of the neurons.

Summing up, a bottom-up view arises that many genes act on different levels of the morphological formation of connections between brain cells and areas leading to an insufficient macro-anatomic network that processes information in a highly pathological way ("dysconnectivity" hypothesis; [21]). Still a lot of gaps must be filled by data and by concept formation. This would be a task of a (Computational) "Systems Neuropsychiatry" because computer models and simulations seem to be necessary for deeper understanding of that issues of neurodynamics [2].

Already integrating current biochemical hypotheses for schizophrenia cannot be understood intuitively: A hyperfunction of dopamine transmission, a hyperfunction of serotonergic transmission (and also a hyperfunction of norepinephrine) related

to a hypofunction of glutamate and GABA transmission (and also of acetylcholine transmission) can be integrated in the concept of a "neurochemical mobile" that could be the basis of a computer-based exploration of neurodynamics [22].

3 Conclusions

Progress in understanding and explaining mental disorders on a neurobiological basis needs systemic conceptions and methodologies to analyze data and to construct computational models. This new research program could be called "Computational Systems Neuropsychiatry".

Acknowledgements This paper is produced without financial support and there is no conflict of interests.

References

1. WHO, World Health Organization (2008). The Global Burden of Disease: 2004 Update. WHO, Geneva
2. Tretter F, Gebicke-Haerter PJ, Mendoza ER et al. eds. (2010): Systems biology in psychiatric research: From high-throughput data to mathematical modelling. Weinheim: Wiley-VCH; 2010
3. Andreasen NC. Positive and negative symptoms: historical and conceptual aspects. Mod Probl Pharmacopsychiatry 1990; 24: 1–42
4. Bleuler E 1911: Dementia praecox oder Gruppe der Schizophrenien. F. Deuticke, Leipzig/Wien 1911
5. Schiepek, G, Schoppek, W & Tretter, F. (1992). Synergetics in psychiatry - simulation of evolutionary patterns of schizophrenia on the basis of nonlinear difference equations. In W. Tschacher, G. Schiepek & E. J. Brunner (Eds.), Self-organization and clinical psychology. Berlin: Springer. (pp. 163–194).
6. Durstewitz D, Seamans JK, Sejnowski TJ. (2000): Neurocomputational models of working memory. Nat Neurosci. 2000 Nov;3 Suppl:1184-91. Review.
7. Wang, XJ 2006: Toward a microcircuit model for cognitive deficits in schizophrenia. Pharmacopsychiatry S1: 80–87
8. Hazlett, EA; Romero, MJ; Haznedar, MM; New, AS; Goldstein, KE; Newmark, RE; Siever, LJ; Buchsbaum, MS (2007). "Deficient attentional modulation of startle eyeblink is associated with symptom severity in the schizophrenia spectrum". Schizophrenia research 93 (1–3): 288–95.
9. Tretter, F. Müller, W. (Eds)(2007): Computational Neuropsychiatry Vol 2: The Functional Architecture of Working Memory Networks in Schizophrenia – Data and Models. Pharmacopsychiatry Supplement 1
10. Jirsa V, Sporns O, Breakspear M, Deco G, McIntosh AR (2010). Towards the virtual brain: Network modeling of the intact and the damaged brain. Archives Italiennes de Biologie 148: 189–205.
11. Carlsson A. The current status of the dopamine hypothesis of schizophrenia. Neuropsychopharmacology 1988; 1: 179–186

12. Carlsson A. The neurochemical circuitry of schizophrenia. Pharmacopsychiatry 2006; 39 Suppl 1: S10–14
13. Winterer, G., Weinberger DR (2004): Genes, dopamine and cortical signal-to-noise-ratio in schizophrenia. Trends in Neurosciences V27(11) 74:1–58.
14. Bányai, M. Ujfalussy, B. Diwadkar V. and Érdi P. (2011): Impairments in the prefronto-hippocampal interactions explain associative learning deficit in schizophrenia. BMC Neuroscience 12(Suppl 1):93, 2011. doi:10.1186/1471-2202-12-S1-P93
15. Lewis, D.L., Hashimoto T, Volk, DW (2005) Cortical inhibitory neurons and schizophrenia Nature Rev Neurosci 6:312–324
16. Liljenström, H. (2010): Network Effects of Synaptic Modifications. Pharmacopsychiatry 2010; 43: S67-S81 DOI: 10.1055/s-0030-1252058
17. Uhlhaas, P.J. Singer, W. (2010): Abnormal neural oscillations and synchrony in schizophrenia. Nat Rev Neurosci. 2010 Feb;11(2):100-13. doi: 10.1038/nrn2774.
18. Noh K, Shin, K.S, Shin D., Hwang J.Y. Kim J.S., Jang J.H., Chung C.K, Kwon J.S., Cho1K-H *(2013): Impaired coupling of local and global functional feedbacks underlies abnormal synchronization and negative symptoms of schizophrenia. BMC Systems Biology 2013 7:30.
19. Goto Y, Grace AA. The dopamine system and the pathophysiology of schizophrenia: a basic science perspective. Int Rev Neurobiol 2007; 78: 41–68
20. Harrison PJ, Weinberger DR. 2005: Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence. Mol Psychiatry. 2005 Jan;10(1):40–68
21. Stephan, K. E., Friston, K. J. and Frith C. D. (2009). Dysconnection in Schizophrenia: From Abnormal Synaptic Plasticity to Failures of Self-monitoring. Schizophr Bull. 2009 May;35(3):509-27.
22. Tretter, F., an der Heiden, U., Rujescu, D., Pogarell, O. (2012): Computational modelling of schizophrenic symptoms – basic issues Pharmacopsychiatry 45, S1, 36–41

Modeling the Genesis of Components in the Networks of Interacting Units

Ichiro Tsuda, Yutaka Yamaguti, and Hiroshi Watanabe

Abstract From the viewpoint of system development, we investigate how components emerge in a network system consisting of interacting units. We propose two mathematical models with 'variational' principles: one treats the emergence of neuron-like components from interacting maps, and the other one treats the emergence of hierarchical module-like components from interacting neuron-like units. In both models, maximum transmission of information was used as a 'variational' principle. This type of mathematical model provides a basis for consideration of the mechanism of cell differentiation in embryos and stem cells, and of functional differentiation in the brain.

Keywords Functional differentiation • Chaos • Variational principle • Components

1 Introduction

In conventional mathematical models of biological systems, coupled dynamical systems such as coupled-map and coupled-oscillator systems have been adopted widely and investigated in detail. A coupled system can be recognized as providing a good model for the emergence of varieties of spatiotemporal patterns over various hierarchies from the microscopic to the macroscopic level. In this type of model, the presence of the element itself and its function are presupposed. This is true, for example, for chemical reactions, where starting chemical materials are given, they

I. Tsuda (✉) • Y. Yamaguti
Research Institute for Electronic Science, Hokkaido University, Kita 12 Nishi 7, Kita-ku, Sapporo 060-0812, Japan
e-mail: tsuda@math.sci.hokudai.ac.jp

H. Watanabe
Research Center for Integrative Mathematics, Hokkaido University, Kita-12, Nishi-7, Sapporo, Hokkaido 060-0812, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_84

begin to interact with each other, and the reactions develop, forming products of the system. Different processes operate in differentiation, such as cell differentiation in embryos, and functional differentiation in the brain because the interacting components cannot be determined before the system development starts. Rather, the components are produced by a certain constraint, which acts on the whole system [1, 2]. We model this process, adopting a certain 'variational' principle.

2 Method/Models

Using a certain genetic algorithm, we calculated the development of the interactions and the states of the dynamical systems. In the genetic algorithm, we applied a constraint of maximum transmission of information as a 'variational' principle to drive the system development: mutual information in the study of a genesis of neuron-like components and transfer information in the study of a genesis of module-like components. Because transfer entropy is equivalent to conditional mutual information, we describe a 'variational' principle in terms of a maximization of mutual information as a typical example. Mutual information is defined in the following way.

$$I(f) = \int p\,(x_t) \ln p(x_t)^{-1} dx_t - \int p\,(x_t)\; p\,(x_{t\prime}/x_t) \ln p(x_{t\prime}/x_t)^{-1} dx_t \qquad (1)$$

One may view the transition probability as a function of a state variable x and its time derivative x', that is,

$$p\,(x_{t\prime}/x_t) \rightarrow p\,(x\prime_t, x_t) \qquad (2)$$

The dynamical systems law is written by

$$\frac{dx}{dt} = f(x) \qquad (3)$$

Then, one has the following formula.

$$I(x) = \int_{t_1}^{t_2} F\left(x, x'\right) dt \qquad (4)$$

$$F\left(x, x'\right) = f(x) p(x) \ln p(x)^{-1} - f(x) p(x) p\left(x', x\right) \ln p(x\prime, x)^{-1} \qquad (5)$$

The first and the second terms on the right hand side of Eq. 5 may correspond to the potential energy-like and the kinetic energy-like terms, respectively.

Here, $-F(x, x')$ is a Lagrangean-like function, which implies local information flow in the present case. Therefore, the 'variational' principle is provided by the following formula.

$$\delta I = 0 \Rightarrow \frac{d}{dt}\frac{\partial F}{\partial x'} - \frac{\partial F}{\partial x} = 0 \tag{6}$$

3 Results

We report the results of the simulations for the genesis of heterogeneously interacting modules from homogeneously interacting modules and the results for the genesis of a neuronal unit. For the former model [3], we found the development of heterogeneous interactions between presumed modules that consist of homogeneous interactions among units (Fig. 1). The development of the dynamics was obtained under the constraint of maximum transmission of information measured by the product of transfer entropies from module I to II and module II to I. We used the networks of the Weyl transformations with noise, which are coupled randomly via sinusoidal transformations of those variables. Using the genetic algorithm, the networks developed from homogeneous couplings to heterogeneous couplings. We observed the following spontaneous symmetry breaking: In one module, say module I, in-phase couplings were dominant, but other types of couplings also survived. In the other module (module II), only in-phase couplings survived. The couplings from module I to II developed to become in-phase couplings only, while those from module II to I developed to become anti-phase couplings only.

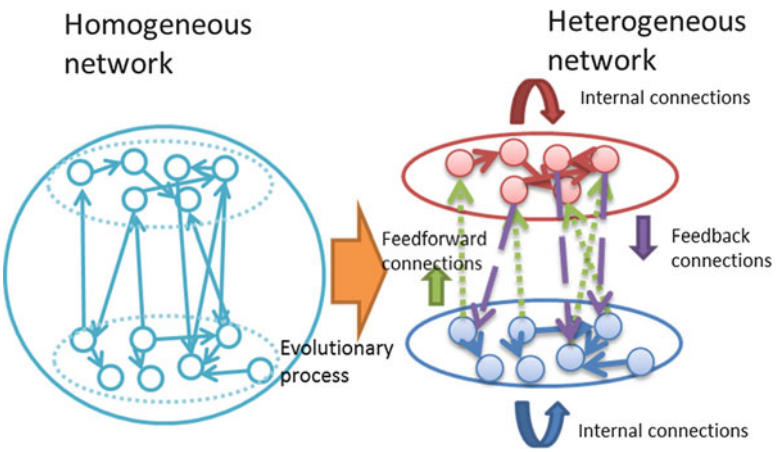


Fig. 1 Genesis of a heterogeneous network from a homogeneous network under the constraint

For the latter model [4], we found the development of neuron-like units from a set of various dynamical systems, whose units exhibit an excitable or an oscillatory behavior. This development was obtained under the constraint of maximum transmission of mutual information in the network of dynamical systems. We adopted a set of one-dimensional maps described by the following equation.

$$z(t+1) = \tanh(\gamma_1(z(t) - \alpha_1)) - \omega \tanh(\gamma_2(z(t) - \alpha_2)) + J \tag{7}$$

By changing the parameters in the right hand side of Eq. 7, we obtained constant functions, monotonic functions, and unimodal and bimodal functions. The network consisting of the same maps with a given strength of couplings was developed by a certain genetic algorithm, where the global constraint was a maximum transmission of information from external inputs measured by the mutual information between the input time series and the time series of each elementary individual map. In the case of strong couplings, a constant function survived, which means that any external input is transmitted on the network without any deformation. In the case of intermediate strengths of couplings, excitable maps survived (see Fig. 2). Three fixed points appear in the map. The leftmost one may represent equilibrium and a resting potential, and the middle one indicates a threshold. When the initial state is set below the threshold, the dynamical orbit monotonously relaxes to the equilibrium state. Whereas, when the initial state is set beyond the threshold, the dynamical orbit experiences a large excursion surrounding the third fixed point, which is unstable, overshoots the equilibrium state indicating hyper-polarization, and finally relaxes to the equilibrium state. The nature of these behaviors is similar to that of the excitability of neurons. Finally, in the case of weak couplings, oscillatory maps survived, and these may indicate the appearance of oscillatory neurons.

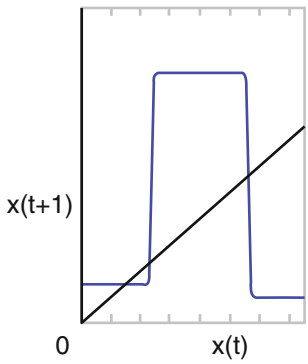


Fig. 2 Genesis of excitable map. Abscissa denotes x(t) and ordinate x(t + 1)

4 Conclusion

We found curious behaviors in the development of dynamical systems under constraints, where the constraints provided the influential factor for a whole system through a 'variational' principle. Recently, several features of dynamic behaviors in neural systems with respect to functional differentiation have been observed in various network systems. The characteristic dynamics in the system development we found here may be related to those findings.

Acknowledgments This work was partially supported by a Grant-in-Aid for Scientific Research on Innovative Areas "The study on the neural dynamics for understanding communication in terms of complex hetero systems (No. 4103)" (21120002) of The Ministry of Education, Culture, Sports, Science, and Technology, Japan. This work was also supported partially by HFSP (RGP0039/2010) "Deliberative Decision Making in Rats".

References

1. I. Tsuda, A hermeneutic process of the brain. *Prog. Theor. Phys*., suppl. 79 (1984), 241–259.
2. R. Rosen, *Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life*. Columbia University Press, New York, 1991.
3. Y. Yamaguti, in this Proceedings.
4. H. Watanabe, T. Ito, and I. Tsuda, Maximal transmission of mutual information can produce neuron-like functions. To be submitted, 2015.

Visual Hallucinations in Dementia with Lewy Bodies (II): Computational Aspects

Hiromichi Tsukada, Hiroshi Fujii, Ichiro Tsuda, and Kazuyuki Aihara

Abstract The aim of this study is, together with a companion paper, to present a hypothetical description of the core mechanisms of *recurrent complex visual hallucinations* (RCVH) associated with dementia with Lewy bodies (DLB). This paper focusses on giving some illustrative numerical examples on how the brain hallucinates, and to see whether mathematical models can describe our proposals of the root cause of these hallucinations.

Keywords DLB • RCVH • Dementia • Attention • Visual hallucinations • Attractor network

1 Introduction

Dementia with Lewy bodies (DLB) is the second most prevalent dementia, and patients with DLB frequently experience the psychic symptom of recurrent complex visual hallucinations (RCVH). The possible involvement of cholinergic deficiency in DLB has long been argued, but the precise neuropathological mechanisms are not fully understood [1, 2].

H. Tsukada (✉) • I. Tsuda
Research Institute for Electronic Science, Hokkaido University, Kita 12 Nishi 7, Kita-ku, Sapporo 060-0812, Japan
e-mail: tsukada@es.hokudai.ac.jp

H. Fujii
Department of Intelligent Systems, Kyoto Sangyo University, Kyoto 603-8555, Japan

K. Aihara
Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_85

In the companion paper [3], we presented the following as the potential root cause of RCVH in DLB (hereafter referred to as VH).

1. Malfunctions in the prefrontal cortex (PFC: VLPFC/OFC) and the inferior temporal (IT) cortex are responsible for VH.
2. In normal situations, the PFC quickly creates, to identify an external object at the center of attention, bias signals ("index") on its categorical identity to facilitate image reactivation in the IT. This PFC function is executed based on information arriving from three pathways: contextual data (as "gist") via Pathway III [4], emotion-expectancy (Pathway II), and external but cursory (low frequency) images (LFI) via Pathway I – a cortical *shortcut* [5].
3. In a hallucinating brain, the possible pathophysiology may involve conduction disturbances somewhere along Pathway I. Hence, the PFC decision on a top-down "index" is made on the basis of only emotion-expectancy and the contextual information, ignoring the external LFI via Pathway I. At the same time, a similar disturbance may occur in the pathway linking visual systems to the IT (V-IT Pathway).

These disturbances may arise from either (a) cortical degenerations as loss of cortical pyramidal neurons/ pre-synaptic axons [6], or (b) loss of the α7 subtype of nicotinic receptor (α7-nAChR) [7, 8].

In the following discussion, we will focus on two points:

1. The structure of the attractor space for an index in the PFC, and its dynamics.
2. How can the IT cortex activate detailed images based on the PFC biasing signal ("index")?

2 Models

2.1 Attractor Structure of the PFC Index Space

Let us consider first a situation in which the signals via Pathway I are blocked because of conduction disturbances. What would be the PFC's decision on the "should exist" object? A key may lie in the structure of attractor space for indices ("index space"). We assume that the index space is the product of three "property spaces", i.e., LFI images (Pathway I), emotion and expectancy (Pathway II), and contextual association between objects and scenes (Pathway III) [4] (see Fig. 1).

In the following, the two streams from Pathways II and III are considered as a single context-expectancy variable (the "top-down" stream), and the external image (the "bottom-up" stream) for simplicity. Thus the representation in the index space would be expressed by a concatenation [EXm, Cn], (m = 1,..,M; n = 1,..,N), where EXm and Cn, are respectively the m-th external image and the n-th context-

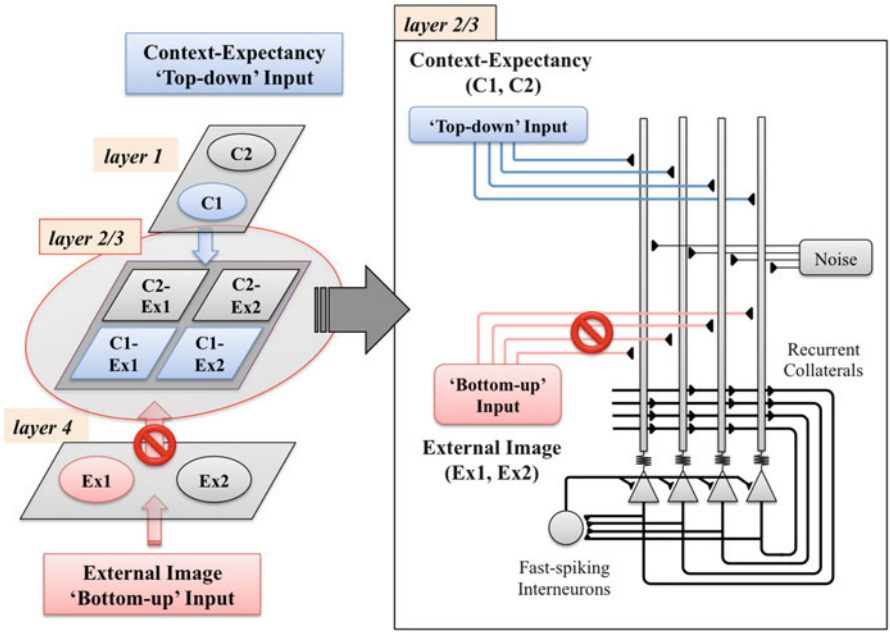


Fig. 1 A schematic diagram of PFC network architecture

expectancy expressed by a simple pattern of 0 and 1 binary sequences. (M = 2, N = 2 for the simplest case). The model is designed so that each concatenated vector [EXm, Cn] defines an attractor in the index space (see Fig. 1). This index network is tuned so that it exhibits a dynamics spontaneously *itinerant* among those ("quasi") attractors [9] in default states, i.e., with no inputs.

The question is, if only a top-down "variable" (e.g., C1) is assigned (with no information on EX's), then how would PFC behave? Possible states would be [C1, EX1] or [C1, EX2], both being C1 internal context-related. In fact, under such a situation, the index dynamics exhibits a transitory behavior among these C1-related states. That is, the network organizes a *context-dependent itinerancy* under such conditions. The staying time in a particular attractor would depend on the situation (see Fig. 2).

2.2 IT Model

The IT network receives a "categorical index" from the PFC, and this biasing signal could activate the object representation even without inputs from the visual cortices [10, 11]. This is the situation that we postulate as the one that occur in DLB hallucinators. We emphasize that IT neurons are activated by top-down signals

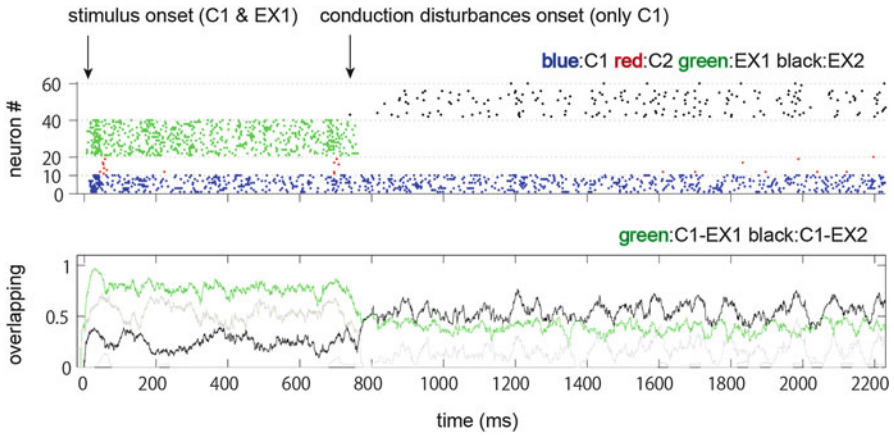


Fig. 2 Effect of conduction disturbances. (*Top*) Spike rasterplot of pyramidal neurons in the PFC network. *Different colors* indicate a different cell assembly (*blue*: C1; *red*: C2; *green*: EX1; *black*: EX2). (*Bottom*) Overlapping between spike patterns of neurons in the PFC network and stored attractor patterns

without bottom-up sensory input [10]. Moreover, attentional biasing signals can also be generated in the absence of any visual stimulation whatsoever [11]. The basic question is how the IT can activate a relevant object representation solely from a top-down index from the PFC. We propose the following two scenarios.

Scenario I. Attentional Phasic Cholinergic Projections onto the IT Phasic (transient) acetylcholine (ACh) projections from the nucleus basalis of Meynert [12] help the IT network state to make a transition from a transitory (quasi-attractor) state to an attractor state. Top-down attention may help the IT dynamics to transit to the attractor regime while the "categorical index" from the PFC contributes to jump in to the specified attractor [9, 13]. However, whether phasic cholinergic projections from the NBM takes place [12] in the case of visual hallucinator, is not known, since attentional deficiency is known as one of the principal symptoms of DLB.

Scenario II. Cognitive "Perseveration" in PFC A long-lasting signal from the PFC with a proper firing rate can sustain an image representation in the IT even without cholinergic projections.

3 Results

3.1 Hallucinatory Index (Bias)

We examined the effect of conduction disturbances in the PFC network. During the first 750 ms, the PFC network functioned correctly with external object EX1 and internal context C1. Then EX1 was turned off (but the internal context C1

remained) at t = 750 ms. Subsequently, the activation pattern of EX1 suddenly disappeared, and another alternative subjective guess EX2 was activated (see Fig. 2). This guessed object has nothing in common with the correct object EX1 except that they are both imaginable from, and consistent with, the internal context C1.

3.2 Attentional Phasic Cholinergic Projections onto the IT (Scenario I) and Cognitive "Perseveration" in the PFC (Scenario II)

Our simulation tested the above two IT scenarios. In Scenario I, under the condition of continuing ACh projections, attractor 1 remained activated while ACh was effective even when the PFC input was turned off (see Fig. 3a). However, when the ACh projections were turned off, the activation of attractor 1 could not be remained. In Scenario II, A long-lasting signal from the PFC could sustain image representation in the IT even with no cholinergic projections (see Fig. 3b). However, attractor 1 could not maintain activation without Glu spike volleys from the PFC.

4 Discussions

In our simulation, we observed that some characteristics of VH symptoms could be reproduced in our conceptual model. The PFC model provided a subjectively correct, i.e., consistent with context, but objectively wrong, index to the IT on the nature of the "seen" object. In the IT model, the activation of IT representation continued under the conditions of our two scenarios.

However, there are some additional considerations, such as the configuration of the attractor space and interactions between the PFC and IT. In addition, there are many remaining questions on the mechanisms of "perseveration" that continues for a few minutes with the occurrence of hallucinations. Experimental studies of the pathophysiology underlying our hypothesis are required.

Acknowledgments We thank Daniel Collerton, University of Newcastle upon Tyne, United Kingdom for valuable discussions. This work was partially supported by a Grant-in-Aid for Scientific Research on Innovative Areas (No.4103) (21120002) from Ministry of Education, Culture, Sports, Science and Technology (MEXT) in Japan and partially supported by Human Frontier Science Program Organization (HFSP:RGP0039).

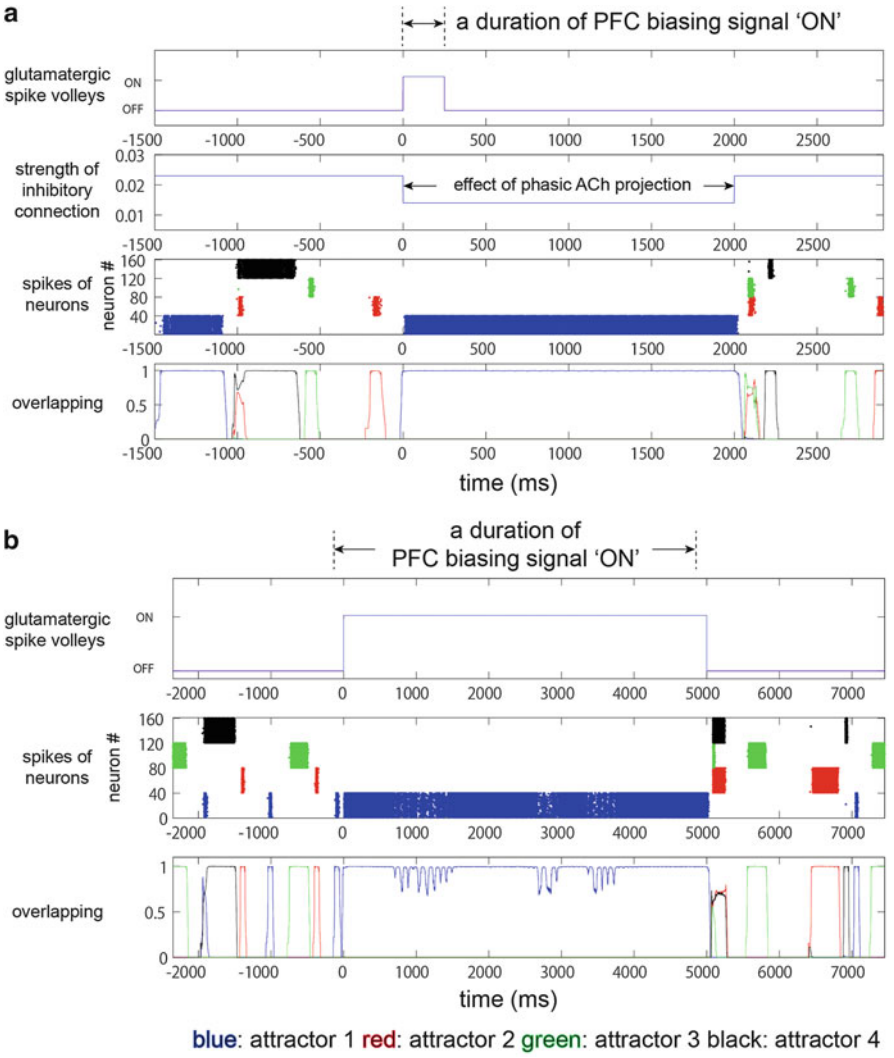


Fig. 3 (**a**) Attentional phasic cholinergic projections onto the IT (Scenario I). (**b**) Cognitive perseveration in the PFC (Scenario II)

References

1. E.K. Perry and R.H. Perry, Acetylcholine and Hallucinatiaons: Disease Related Compared to Drug-Induced Alterations in Human Consciousness, Brain and Cognition 28 (1995), 240–258.
2. D. Collerton et al., Why people see things that are not there: A novel Perception and Attention Deficit model for recurrent complex visual hallucinations, Behavioral and Brain Sciences 28 (2005), 737–794.

3. H. Fujii, H. Tsukada, I. Tsuda and K. Aihara, Visual Hallucinations in Dementia with Lewy Bodies (I): A Hodological View, this issue.
4. M.J. Fenske et al., Top-down facilitation of visual object recognition: object-based and context-based contributions, Progress in Brain Research 155 (2006), 3–21.
5. M. Bar, A Cortical Mechanism for Triggering Top-Down Facilitation in Visual Object Recognition, Journal of Cognitive Neuroscience 15 (2003), 600–609.
6. C. Sanchez-Castaneda et al., Frontal and Associative Visual Areas Related to Visual Hallucinations in Dementia with Lewy Bodies and Parkinson's Disease with Dementia, Movement Disorders 25 (2010), 615–622.
7. R.T. Reid et al., Nicotinic receptor losses in dementia with Lewy bodies: comparisons with Alzheimer's disease, Neurobiology of Aging 21 (2000) 741–746.
8. J.A. Court et al., Visual hallucinations are associated with lower αbungaro toxin binding in dementia with Lewy bodies, Pharmacology, Biochemistry and Behavior 70 (2001) 571–579.
9. H. Tsukada, Y. Yamaguti and I. Tsuda, Transitory memory retrieval in a biologically plausible neural network model, Cognitive Neurodynamics 7 (2013) 409–416.
10. H. Tomita et al., Top-down signal from prefrontal cortex in executive control of memory retrieval, Nature 401 (1999) 699–703.
11. S. Kastner and L. G. Ungerleider, The neural basis of biased competition in human visual cortex, Neuropsychologia 39 (2001) 1263–1276.
12. V. Parikh, R. Kozak, V. Martinez and M. Sarter, Prefrontal acetylcholine release controls cue detection on multiple timescales, Neuron 56 (2007) 141–154.
13. T. Kanamaru, H. Fujii and K. Aihara, Deformation of Attractor Landscape via Cholinergic Presynaptic Modulations: A Computational Study Using a Phase Neuron Model, PLoS ONE 8 (2013), e53854.

Neural Dynamics Associated to Preferred Firing Sequences

Alessandro E.P. Villa, Alessandra Lintas, and Jérémie Cabessa

Abstract In a distributed recurrent neural network equivalent changes at one synapse might correspond to different patterns of activity and changes in strength at particular links between two cells may become meaningless.The information is not necessarily resident in the links among the units, but is likely to be provided by the activity organized in a highly precise temporal mode. Precise spatio-temporal firing sequences and attractor dynamics may be strongly associated, such that the detection of spatio-temporal firing patterns may reveal the existence of underlying modes of activity controlled by few parameters in deterministic chaotic dynamics.

Keywords Recurrent neural networks • Activity patterns • Spatio-temporal firing • Attractor dynamics • Chaos

1 Spatio–Temporal Firing Patterns

The majority of neural circuits of the forebrain, i.e. the basal ganglia thalamocortical circuit, are formed by highly interconnected networks of neurons in which the activity of each cell is necessarily related to the combined activity in the neurons that are afferent to it. Reentrant activity through chains of neurons is likely to occur due to the presence of recurrent connections at various levels of the circuits. Developmental and/or learning processes determine the strengthening and weakening of synaptic interactions between the neurons of selected pathways. In cell assemblies interconnected in this way, some ordered, and precise (in the order of few ms) interspike interval relationships referred to as *spatio–temporal firing patterns*, may recur within spike trains of individual neurons, and across spike trains recorded from different neurons. For this to be true, temporal firing patterns must occur to a significant level above chance (Fig. 1). Then, whenever the same information is presented in the network, the same cell assemblies will be activated

A.E.P. Villa (✉) • A. Lintas • J. Cabessa
Neuroheuristic Research Group, University of Lausanne, Quartier Dorigny, CH-1015 Lausanne, Switzerland
e-mail: avilla@nhrg.org; alintas@nhrg.org; jcabessa@nhrg.org

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_86

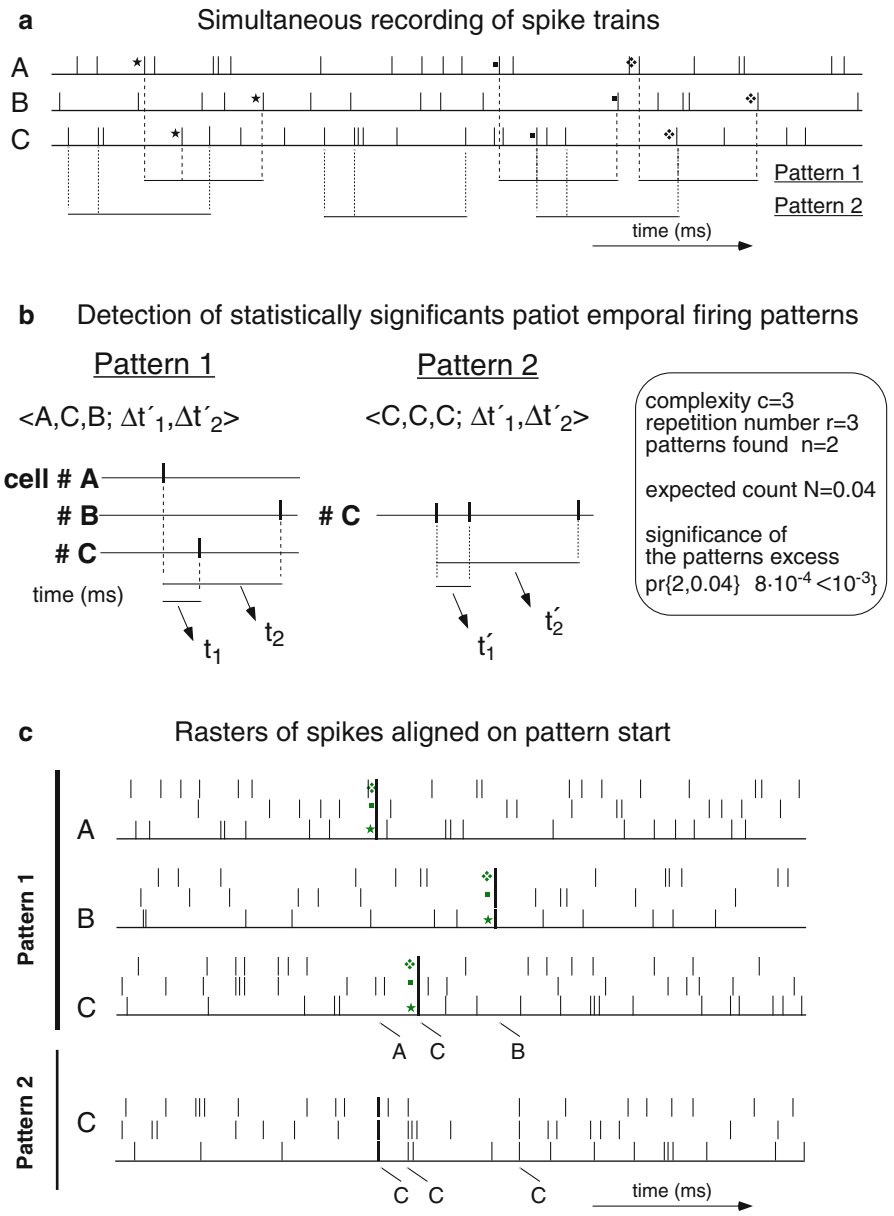


Fig. 1 Outline of the general procedure for pattern detection. (**a**) Three cells, labeled A, B, and C, participate to a patterned activity. Three occurrences of two precise patterns are detected. Each occurrence of the first pattern has been labeled by a specific marker in order to help the reader to identify the corresponding spikes. (**b**) Estimation of the statistical significance of the detected patterns. Two patterns, n = 2, <A,C,B> and <C,C,C> were found. Each pattern was formed by three neurons, c = 3, and was repeated three times, r = 3, in the analyzed record. The expected number of patterns of this complexity and repetition number was N = 0.04. The probability to observe 2 or more patterns when 0.04 patterns are expected is noted as pr{0.02, 4}. (**c**) Display of the pattern occurrences as a raster plot aligned on the patterns' start (Adapted from [19])

and the same temporal pattern of firing will be observed. A remarkable invariance in the firing times of the tested neurons, indicating a high degree of reliability of their response and not a stochastic nature of the discharges, was experimentally observed after complex patterns of stimulation [5,14]. Experimental evidence exists that correlated firing between single neurons recorded simultaneously in the primate frontal cortex may evolve within tens of milliseconds in systematic relation to behavioral events without modulation of the firing rates [21, 26]. Precise firing sequences have been described in relation to particular temporal relationships to stimuli [23], or movement [1], or differentially during the delay period of a delayed response task [16, 24, 27]. When a specific input pattern activates a cell assembly, the neurons are activated following a certain mode. Then, a *mode of activity* defines how an information is processed within a neural network and how it is associated to the output pattern of activity that is generated. In this framework the *state* of the neural network is defined by a set of parameters characterizing the neural network at a certain time. Then, the state of the network at any given time is represented by the values of these parameters and a network state were fully determined if all parameters were known for each neuron.

2 Dynamical System Analysis

The brain is characterized by biochemical reactions whose energy requirement is derived almost entirely from glucose consumption coupled with processes intended to transmit and integrate the information carried by the spikes across the neural networks. For sake of simplicity it is rationale to describe the activity of the neural network with the spike trains of all its elements. Spike trains are statistically expressed by point-like processes with the meaning that point process system are systems whose input and output are point processes. In a dynamical system the subsequent state of the system is determined by its present state. The irreversible dissipative processes associated to brain metabolism introduce an essential metastability of brain dynamics. A dynamical system in a whole is said to be *deterministic* if it is possible to predict precisely the evolution of the system in time if one knows exactly the initial conditions and the subsequent perturbations. However, a slight change or incorrect measurement in these values results in a seemingly unpredictable evolution of the system. A passage in time of a state defines a *process*. Whenever a process is completely deterministic at each step of its temporal evolution but unpredictable over the long term it is called a *chaotic process* or simply *chaos*.

An equivalent definition of a process is a path over time, or *trajectory*, in the *space of states*. The points approached by the trajectory as the time increases to infinity are called *fixed points* and the set of these points forms an *attractor*. If the evolution in time of the system is described by a trajectory forming a closed loop also referred to as a periodic orbit then the system is said to have a *limit cycle*. It is unlikely that the irreversible dissipative processes associated with brain dynamics

produces always the same repeating sequence of states. However, this aperiodic behavior is different from randomness, or stochastic process, because an iterated value of the point process (all spike trains in the network) can only occur once in the series, otherwise due to the deterministic dynamics of the system the next value should also be a repetition and so on for all subsequent values. The perturbations applied to any combination of the governing set of parameters move a dynamical system characterized by fixed points away from the periodic orbits but with passing of time the trajectory collapses asymptotically to the same attractor. If the system is deterministic, yet sensitive to small perturbations, the trajectory defining its dynamics is an aperiodic orbit, then the system is said to have a *chaotic attractor*, often referred to as a *strange attractor*. Then, the set of all possible perturbations define the inset of the attractor or its *basin of attraction*.

By extending this approach to the spike trains recorded from all elements of the neural network it is theoretically possible to develop an acceptable model for the identification of the system. Notice that the goodness of fit of a certain kernel estimate as plausible is evaluated by means of a function f describing its mode of activity–the mode of activity being defined by how an information is processed within a neural network and how it is associated to the output pattern of activity that is generated. In formal terms f is a probability function that describes how a state x is mapped into the space of states. If the function is set by a control parameter μ we can write $f_\mu(x) = f(\mu, x)$. A dynamical system x' is a subset of the space of states and can be obtained by taking the gradient of the probability function with respect to the state variable, that is $x' = \nabla f_\mu(x)$. Mathematically speaking, the space of states is a finite dimensional smooth manifold assuming that f is continuously differentiable and the system has a finite number of degrees of freedom [18].

If the activity is generated by chaotic attractors, whose trajectories are not represented by a limit set either before or after the perturbations, the attracting set may be viewed through the geometry of the topological manifold in which the trajectories mix. It is likely that several attractors may appear, moving in space and time across different areas of the network, in the dynamics of large neural networks. Such complex spatio-temporal activity may be viewed more generally as an *attracting state*, instead of simply an attractor [3]. In particular, simulation studies demonstrated that a neural circuit activated by the same initial pattern tends to stabilize into a timely organized mode or in a asynchronous mode if the excitability of the circuit elements is adjusted to the first order kinetics of the postsynaptic potentials [10, 22].

3 The Brain Catastrophe

Let us assume that the dynamical system is structurally stable. In terms of topology structural stability means that for a dynamical system x' it exists a neighborhood $\mathcal{N}(x')$ in the space of states with the property that every $Y \in \mathcal{N}(x')$ is

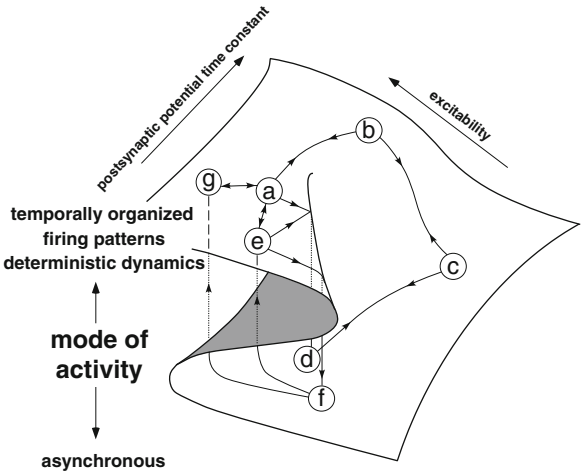


Fig. 2 Topological interpretation of neural dynamics as a function of two control parameters, the cell excitability and the kinetics of the postsynaptic potentials. The equilibrium surface is represented by a cusp catastrophe where transitions can occur either suddenly or continuously between temporally organized firing patterns and asynchronous activity

topologically equivalent to x'. This assumption is extremely important because a structurally stable dynamical system cannot degenerate. As a consequence, there is no need to know the exact equations of the dynamical system because qualitative, approximate equations, i.e. in the neighborhood, show the same qualitative behavior [2]. In the case of two control parameters, $x \in \mathbb{R}$, $\mu \in \mathbb{R}^2$, the probability function f is defined as the points μ of $\mathbb{R}^2$ with a structurally stable dynamics of $x' = \nabla f_\mu(x)$ [15]. That means the qualitative dynamics x' is defined in a neighborhood of a pair (x_0, μ_0) at which f is in equilibrium (e.g. minima, maxima, saddle point). With these assumptions, the equilibrium surface is geometrically equivalent to the Riemann-Hugoniot or cusp catastrophe [20]. The cusp catastrophe is the universal unfolding of the singularity $f(x) = x^4$ and the equilibrium surface is described by the equation $V(x, u, v) = x^4 + ax^2 + bx$, where a and b are the control parameters. We suggest that metastable modes of neural activity could lie in the equilibrium surface with postsynaptic potential kinetics and membrane excitability as control parameters (Fig. 2).

We assume that the same neural network may subserve several modes of activity through modulation of its connectivity, e.g. according to learning or pathological processes, or by modulation of its excitability, e.g. by modulation of the resting potential or of the synaptic time constants. The state of a neural network being defined by a set of characteristic control parameters at a certain time then, at any given time, the state of the network is represented by the values of control parameters and a network state is fully determined if all parameters were known for each neuron. It is not possible to know all variables determining brain dynamics, yet the analysis of experimental spike trains has confirmed the existence of deterministic chaotic dynamics in neural networks [4, 8, 11].

The paths drawn on the cusp illustrate several types of transitions between network states. In this framework at Point (a) in Fig. 2 the network state is such that an input pattern will evoke precisely time structured activity detectable by

preferred firing sequences. This point corresponds to a high level of excitability and a relatively long decay time of the postsynaptic potentials, e.g. 12 ms. This may be associated to the tonic mode of firing described in the thalamo-cortical circuit, where bistability of firing activity has been well established. Different firing patterns might be evoked by the same input if the synaptic dynamics is changed within a certain range of cellular excitability, as suggested for neuromodulatory mediators. Also, different input patterns of activity may produce similar modes of activity, somewhat like attractors. The transitions between these states are represented by paths $(a-b-a)$, $(a-e-a)$ and $(a-g-a)$ in Fig. 2. Several types of neurons tend to switch towards a rhythmic or bursty type of firing if the excitability is decreased due to a hyperpolarization of the cell membrane or by modifying the spike threshold level [9, 17]. In the former case a smooth passage between timely structured activity and asynchronous firing is likely to occur, as suggested by path $(b-c-b)$, especially if the synaptic decay is long. On the opposite, a sudden switch from temporal patterns of firing to desynchronized activity will occur, as indicated by paths $(a-d)$ and $(e-f)$, in the case of a fast synaptic decay and a modulatory input modifying the threshold potential.

Complex spatio-temporal firing patterns may also occur with low levels of excitability (point (e) in Fig. 2), as suggested by cholinergic switching [25] and control of synchronous activation within the basal ganglia thalamocortical circuit [12, 13]. Point (e) on the equilibrium surface can be particularly unstable because a further decrease in excitability, path $(e-f)$, but also an increase in synaptic decay, path $(e-d)$, may provoke a sudden change in the mode of activity, as observed in simulation studies [10]. During long lasting hyperpolarization the excitability is low and the kinetics of the postsynaptic potential is often irrelevant with regarding the input pattern such that the output activity would always tend to be organized in rhythmic bursts. Conversely, an increase in excitability from point (f) with a fast time constant of the synaptic decay, say 4–5 ms, the input patterns could turn on either stable, path $(f-g)$, or unstable temporally organized modes of activity only through sudden transitions, path $(f-e)$ [17].

4 Discussion

The detection of precise spatio-temporal firing patterns or attractors necessarily requires the stability of the generating processes over a relatively long period of time. Thus, precise spike patterns in single or across multiple neurons may be mostly involved in long-term processes (e.g., memory traces, learned motor programs), whereas the ensemble coding based on systematic firing rate modulations may be related to short-term operational processes (e.g. motor action, attentional or feature-binding). The research presented here *is not* discussing some questions that most neurophysiologists usually ask: which is the most adequate stimulus for a given neuron, how is the external world mapped in the cortex, what are precisely the receptive fields of single units, etc . . . What is being discussed here *is* the association

of neural activity in distributed brain information processing with deterministic chaotic dynamics. In the nervous system the problem of learning is crucial and can hardly be approached without taking into account synaptic modification. However, changes in strength at particular links between two cells may become meaningless because in a distributed system equivalent changes at one synapse might correspond to different patterns of activity. Therefore, the information is not necessarily resident in the links among the units, but is likely to be provided by the activity organized in a highly precise temporal mode (precise spatio-temporal firing patterns and attractors) that is considered *meaningful* if it is associated to an outcome that is validated by the re-entrant neural activity, or *spurious* otherwise [6, 7].

References

1. Abeles M, Vaadia E, Bergman H, Prut Y, Haalman I, Slovin H (1993) Dynamics of neuronal interactions in the frontal cortex of behaving monkeys. Conc Neurosci 4:131–158
2. Andronov AA, Pontryagin L (1937) Systemes grossiers. Dokl Akad Nauk SSSR 14(5):247–250
3. Asai Y, Villa AEP (2008) Reconstruction of underlying nonlinear deterministic dynamics embedded in noisy spike trains. J Biol Phys 34(3–4):325–340
4. Asai Y, Guha A, Villa AEP (2008) Deterministic neural dynamics transmitted through neural networks. Neural Netw 21(6):799–809
5. Bryant HL, Segundo JP (1976) Spike initiation by transmembrane current: a white noise analysis. J Physiol 260:279–314
6. Cabessa J, Villa AEP (2010) A hierarchical classification of first-order recurrent neural networks. Chin J Physiol 53(6):407–416
7. Cabessa J, Villa AEP (2012) The expressive power of analog recurrent neural networks on infinite input streams. Theoretical Computer Science 436:23–34
8. Celletti A, Villa AE (1996) Determination of chaotic attractors in the rat brain. J Stat Phys 84(5):1379–1385
9. Foote SL, Morrison JH (1987) Extrathalamic modulation of cortical function. Annual Review of Neuroscience 10:67–95
10. Hill S, Villa AE (1997) Dynamic transitions in global network activity influenced by the balance of excitation and inhibtion. Network: Comp Neural Netw 8:165–184
11. Iglesias J, Chibirova O, Villa A (2007) Nonlinear dynamics emerging in large scale neural networks with ontogenetic and epigenetic processes. Lect Notes Comp Sci 4668:579–588
12. Lintas A, Silkis IG, Albéri L, Villa AE (2012) Dopamine deficiency increases synchronized activity in the rat subthalamic nucleus. Brain Res 1434:142–151, DOI 10.1016/j.brainres.2011.09.005
13. Lintas A, Schwaller B, Villa AE (2013) Visual thalamocortical circuits in parvalbumin-deficient mice. Brain Res DOI 10.1016/j.brainres.2013.04.048
14. Mainen ZF, Senjowski TJ (1995) Reliability of spike timig in neocortical neurons. Science 268:1503–1506
15. Peixoto MM (1962) Structural stability on two-dimensional manifolds. Topology 1(2):101–120
16. Prut Y, Vaadia E, Bergman H, Slovin H, Abeles M (1998) Spatiotemporal structure of cortical activity: Properties and behavioral relevance. J Neurophysiol 79:2857–2874
17. Segundo JP (2003) Nonlinear dynamics of point process systems and data. Int J Bif Chaos 13(08):2035–2116
18. Smale S (1967) Differentiable dynamical systems. Bull Amer Math Soc 73:747–817

19. Tetko IV, Villa AEP (2001) A pattern grouping algorithm for analysis of spatiotemporal patterns in neuronal spike trains. 1. Detection of repeated patterns. J Neurosci Meth 105:1–14
20. Thom R (1975) Structural stability and morphogenesis: an outline of a general theory of models. Fowler DH, trans. Benjamin, Reading, PA (USA)
21. Vaadia E, Haalman I, Abeles M, Bergman H, Prut Y, Slovin H, Aertsen A (1995) Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. Nature 373:515–518
22. Villa A, Tetko I (1995) Spatio-temporal patterns of activity controlled by system parameters in a simulated thalamo-cortical neural network. In: Herrmann H, Wolf D, Poppel E (eds) Supercomputing in Brain Research : from Tomography to Neural Networks, World Scientific, pp 379–388
23. Villa AEP, Abeles M (1990) Evidence for spatiotemporal firing patterns within the auditory thalamus of the cat. Brain Res 509:325–327
24. Villa AEP, Fuster JM (1992) Temporal correlates of information processing during visual short-term memory. NeuroReport 3:113–116
25. Villa AEP, Bajo Lorenzana VM, Vantini G (1996) Nerve growth factor modulates information processing in the auditory thalamus. Brain Res Bull 39(3):139–147
26. Villa AEP, Hyland B, Tetko IV, Najem A (1998) Dynamical cell assemblies in the rat auditory cortex in a reaction-time task. BioSystems 48:269–278
27. Villa AEP, Tetko IV, Hyland B, Najem A (1999) Spatiotemporal activity patterns of rat cortical neurons predict responses in a conditioned task. Proc Natl Acad Sci USA 96(3):1106–1111

Evolution of Heterogeneous Network Modules via Maximization of Bidirectional Information Transmission

Yutaka Yamaguti

Abstract In theoretical studies, it has been proposed that maximizing information transmission between subsystems can be a principle for the development and evolution of complex brain networks. In this article, we study how heterogeneous modules develop in coupled-map networks through evolutionary processes, where selection pressure is to maximize bidirectional information transmission. Emergence of heterogeneous structure is demonstrated, showing that maximization of bidirectional information transmission between interacting modules can act as selection pressure to enhance differentiation between interacting modules.

Keywords Bidirectional information transmission • Heterogenous network modules • Evolution • Selection pressure • Coupled oscillators

1 Introduction

Modular architecture is an important concept of neural organization. Understanding the generation process of functionally differentiated modules in the brain is of great interest in the field of neuroscience. In theoretical studies, it has been proposed that maximizing information transmission between subsystems can be a principle for the development and evolution of complex brain networks. In previous studies [1–3], information transmission through successive layers of feed-forward networks, or information preservation through time in recurrent networks, has been considered a principle for designing functional neural networks. In this paper, we investigate a mathematical mechanism of functional differentiation of network modules induced by selection pressure for maximizing bidirectional information transmission. For this purpose, we try to extract the essence of the evolutionary dynamics by investigating a simple coupled-oscillator network model.

Y. Yamaguti (✉)
Research Institute for Electronic Science, Hokkaido University, Kita 12 Nishi 7, Kita-ku, Sapporo 060-0811, Japan
e-mail: yyama@es.hokudai.ac.jp

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_87

2 Models and Methods

We consider heterogeneous coupled-map networks consisting of two sub-network modules. Each module consists of N phase oscillators, which are discrete-time versions of the Kuramoto model. The structure of the network is statistically regulated by a set of parameters, p^c, q^c, and r^c. The probability of existence of a connection between two oscillators within a module (an intra-modular connection) is given by $2p^c(1-q^c)$, while the probability of existence of a connection from an oscillator in module 1 to an oscillator in module 2 or for a connection from module 2 to module 1 (an inter-modular connection), is given by $4p^cq^cr^c$ or $4p^cq^c(1-r^c)$, respectively. Here, p^c regulates the total number of connections in the network, q^c controls the fraction of inter-modular connections among all connections in the network, and r^c controls the fraction of inter-modular connections from module 1 to 2 among all inter-modular connections. Four types of oscillator-to-oscillator coupling are introduced to the network. These types of couplings lead to phase locking with 0 (in-phase), $\pi/2$, π (anti-phase), or $3\pi/2$ phase lag between the two oscillators in a two-oscillator system.

The dynamics of the k-th oscillator in the i-th module ($i = 1, 2, k = 1, \ldots, N$) is described by

$$\theta_{t+1}^{(i,k)} = \theta_t^{(i,k)} + \omega^{(i,k)} + \frac{\alpha}{Np^c} \sum_{(j,l)\in G^{(i,k)}} \sin\left(\theta_t^{(j,l)} - \theta_t^{(i,k)} - \psi_{kl}^{ij}\right) + \beta_t^{(i,k)},$$

where $\omega^{(i,k)}$ is a natural frequency, α is a coupling strength, and $\beta_t^{(i,k)}$ represents additive noise that affects each oscillator independently. $G^{(i,k)}$ represents a set of labels for those oscillators that connect to the oscillator (i,k). Each ψ_{kl}^{ij} is randomly assigned to one of four possible values $(m-1)\pi/2\, m = 1 \ldots 4$, according to probabilities p_m^{ij}. Note that $p_m^{ij} \geq 0$ and $\sum_{m=1}^4 p_m^{ij} = 1$. In the case of a two-oscillator system, these couplings lead to phase locking with 0 (in-phase), $\pi/2$, π (anti-phase), or $3\pi/2$ phase lag between the two oscillators. To characterize the macroscopic states of the modules, we define a phase coherence $R^{(i)}(t)$ and a mean phase $\Theta^{(i)}(t)$ for each module as the absolute value and the argument of $(1/N)\sum_{k=1}^N \exp(\sqrt{-1}\theta^{(i,k)}(t))$:

$$R^{(i)}(t)\exp\left(\sqrt{-1}\Theta^{(i)}(t)\right) = \frac{1}{N}\sum_{k=1}^{N}\exp(\sqrt{-1}\theta^{(i,k)}(t)). \tag{1}$$

The phase difference between mean phases is denoted as $\Phi(t) = \Theta^{(2)}(t) - \Theta^{(1)}(t)$.

Transfer entropy (TE) [4], which measures directed information transfer between two mean phases, was calculated in both directions for each network [5], and the product of the TEs for the two directions was regarded as the fitness function

of the network. TE can be represented by using conditional mutual information: $T_{1\to 2}(\tau) = I\left(\Theta^{(2)}(t+\tau); \Theta^{(1)}(t) | \Theta^{(2)}(t)\right)$, where τ is a time step for discretization and $I(X;Y|Z)$ denotes the mutual information between X and Y conditioned by Z.

A standard genetic algorithm was used to develop the networks by gradually improving the set of parameters. During the evolution, p^c was fixed while a set of other parameters, namely q^c, r^c, and p_m^{ij} ($m = 1 \ldots 4, i, j = 1$ or 2), was taken as a genetic representation to be modified to find the maximum of the fitness function. The gene population consists of 48 genes. In each generation, networks were generated from the genes and numerical simulations were performed. Then, fitness was calculated for each network. The 10 networks with the best fitness were selected for reproduction of the next generation. Exact copies of the selected genes and their mutations were made for the next generation. These processes were repeated for 2000 generations.

3 Results

The mean values of fitness of the selected networks were nearly saturated by the end of the simulations. The connection parameters of evolved networks are summarized in Fig. 1. In module 1, in-phase couplings are dominant in intra-modular couplings, but other couplings remain in the module. On the other hand, in module 2, more couplings became the in-phase type. In inter-modular couplings, almost all couplings from module 1 to module 2 became the in-phase type, while those in the opposite direction became the anti-phase type. Further, the number of couplings was larger in the 1–2 direction than in the opposite direction. These results indicate that symmetry between the two modules was broken and differentiations were developed through the evolutionary process.

The dynamics of the coherence and the inter-modular phase difference in an evolved network are depicted in Fig. 2. The phase difference as well as the intra-modular coherence $R^i(t)$ exhibited slow, chaotic oscillations. The value of the phase difference oscillated slowly between in-phase and anti-phase states. The slow

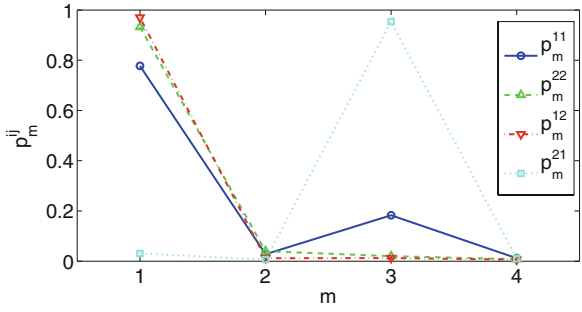


Fig. 1 Means of coupling parameters of evolved networks

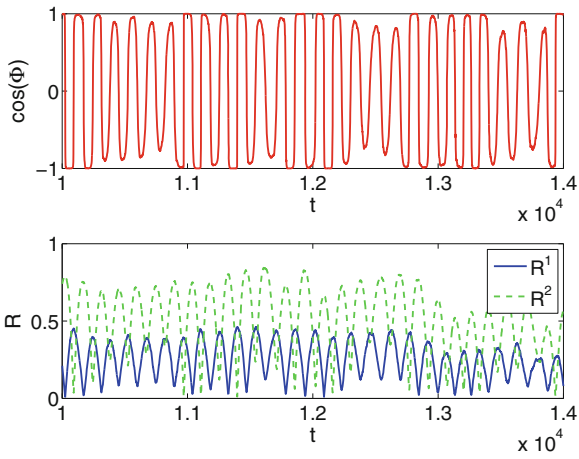


Fig. 2 Dynamics of the phase difference Φ (*upper*) and intra-modular coherence R^i (*lower*)

oscillations of the coherences of two modules had a mutually anti-phase relation. Reflecting the differentiated parameters of the intra-modular connections, their mean amplitude values were also differentiated.

4 Conclusions and Discussion

Developments of two interacting network modules through an evolutionary process that maximizes bidirectional information transmission were studied. Emergence of heterogeneous structure was demonstrated. It was shown that maximization of bidirectional information transmission between interacting modules can act as selection pressure to enhance differentiation between interacting modules. Our results support the hypothesis that maximization of bidirectional information exchange can serve as a principle for development of heterogeneous structures in complex systems such as the brain.

Acknowledgements This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas "The study on the neural dynamics for understanding communication in terms of complex hetero systems (No.4103)" (21120002) of The Ministry of Education, Culture, Sports, Science, and Technology, Japan.

References

1. Linsker, R.: Computer, 21, 105–117, (1988).
2. Tanaka T., Kaneko T., Aoyagi T., Neural computation, 21, 1038–67, (2008).
3. Watanabe H., Ito T., Tsuda I. , in preparation.
4. Schreiber, T., Phys. Rev. Let., 85, 461–4. (2000).
5. Yamaguti Y., Tsuda I., Takahashi Y., Cogn. Neurodyn., 8, 17–26, (2014).

Gamma-Band Shift in the Activity of Rat Hippocampal CA1: A Comparison of Memory-Guided and Visually-Cued Spatial Choice

Noha Mohsen Zommara, Hiroshi Nishida, Muneyoshi Takahashi, and Johan Lauwereyns

Abstract Recent studies have highlighted the contribution of theta-gamma oscillations in hippocampus in memory and cognition. Here we adapted a nose-poke paradigm in order to investigate the role of hippocampal theta-gamma oscillations in a memory-guided behavior. We compared the local field potentials (LFPs) from the hippocampal CA1 cell layer of the same rats in two tasks: a memory-guided spatial alternation task and a visual-spatial discrimination task with random cues. We noted theta-gamma band coupling during a critical fixation period, while the rat was immobile but alert; high and low gamma components were highly phase-locked to the ongoing theta oscillation in CA1 during the fixation period. These modulations from high gamma (55–90) at the onset of the fixation to low gamma (22–45) at the offset of the fixation occurred in both tasks, but were stronger in the memory-guided alternation task. Thus, these data provide further evidence for the role of hippocampus CA1 in mnemonic coding relating to spatial alternation.

Keywords Hippocampus • Spatial alternation • Visual discrimination • Theta-gamma oscillations

1 Introduction

Hippocampus has a critical role in encoding the spatial representations in sequence-dependent learning [1]. Previous studies have reported how gamma and theta oscillations supported various forms of mnemonic coding [2]. Here, we follow

N.M. Zommara • H. Nishida
Graduate School of Systems Life Sciences, Kyushu University, Fukuoka 819-0395, Japan

M. Takahashi • J. Lauwereyns (✉)
Brain Science Institute, Tamagawa University, 6-1-1 Tamagawagakuen, Machida, Tokyo 194-8610, Japan

Graduate School of Systems Life Sciences, Kyushu University, Fukuoka 819-0395, Japan
e-mail: jan@sls.kyushu-u.ac.jp

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7_88

up on this research by examining the dynamics of theta and gamma oscillations in a comparison between an alternation task and a visually-cued task, particularly during a carefully controlled delay period.

We concentrated on the role of hippocampal CA1 while the rats perform two different tasks: a memory-guided task with spatial alternation and a visual discrimination task with random cues. We used a nose-poke paradigm in which rats had to remain for 1 s with their nose in the central hole. This fixation period is crucial for recording the brain activity during an immobile but fully alert anticipation period. We designed a delayed memory-guided task in which rats should memorize the sequence of the trials to make the optimal choice. In the choice phase, both implicated holes were illuminated simultaneously so that the rats could not depend on any sensory cues to discriminate the alternation sequence. On the other hand, in the visual discrimination task, rats were guided by the visual cues; here, each cue was illuminated separately and randomly on a trial-by-trial basis.

2 Methods

Four male adult Wistar rats were used. All training and recording sessions, as well as surgical procedures were approved by the Tamagawa University Animal Care and Use Committee.

During a 1-day session we applied two different tasks in different blocks; a memory-guided spatial alternation task and a visually-guided discrimination task. In both tasks the trial started when the center hole was illuminated; rats were required to make a nose-poke response and sustain it for 1 s. When the center light went off, a delay period was introduced; following that, the left and right holes were illuminated on the frontal wall. The nose poke was recorded through an infrared photo-beam detector in each hole.

In the memory-guided alternation task both holes were illuminated simultaneously and rats were required to alternate between the right and left holes. On the other hand, in the visual discrimination task, the holes were illuminated one-by-one and not in a fixed sequence. Each correct response allowed the rats to obtain one pellet as a reward. Following an erroneous response the lights were extinguished, and the same trial was repeated. The inter-trial interval was 10 s, during which time the rats had to wait until the start of the next trial.

The neural recording was conducted after the completion of behavioral training. We recorded multi-unit activity and local field potentials (LFPs) in stratum pyramidale of the right dorsal CA1 (AP -3.6, ML 2.2 mm) using a 14-tetrode hyperdrive assembly during task performance.

3 Results and Discussion

We found that the memory-guided task had a higher correct performance rate than the visual discrimination task; the mean of correct trials were 90.1 % with SD 4.76 %, and 85.9 % with SD 9.58 %, respectively.

During the immobile period, while rats were waiting for the next event, different trends of gamma waves appeared. The population analysis of the local field potentials (LFPs) showed that gamma-band activity in CA1 shifted from high frequency (55–90 Hz) at the beginning of the fixation to low frequency (22–45 Hz) at the end of the fixation in both tasks during the 1 s fixation period (Fig. 1).

Our analysis showed that the shifts from high gamma at the onset of the fixation to low gamma at the offset of the fixation occurred in both tasks; however, these trends were more pronounced in the memory-guided task. The differential intensity in the modes of information processing in CA1 likely reflects the necessary usage of memory in the memory-guided alternation task; this information needed to be suppressed in the visually guided random cue task.

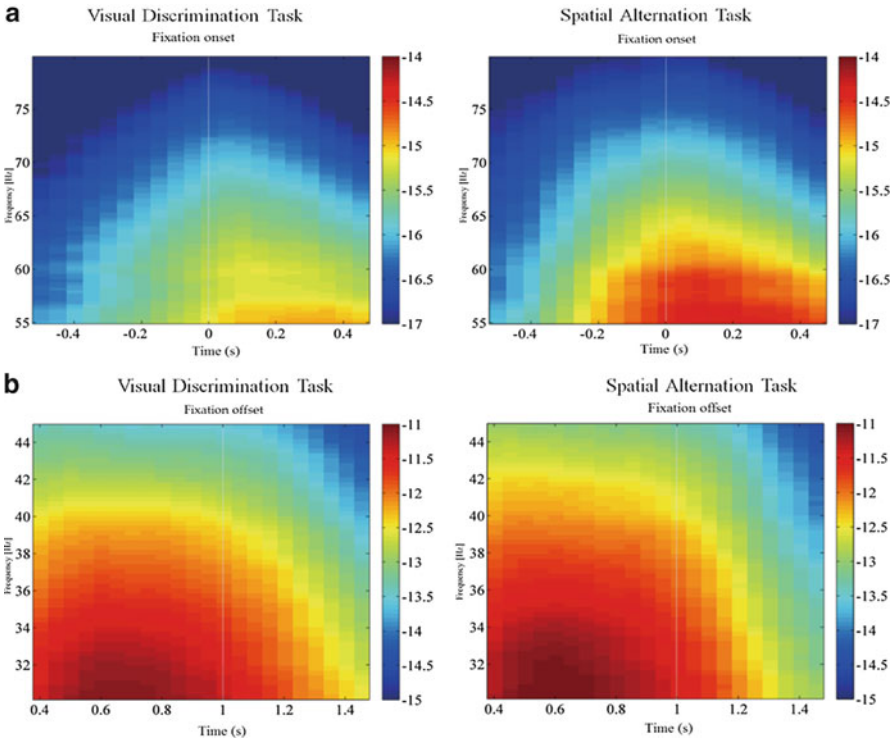


Fig. 1 CA1 oscillation activity during the fixation period in both tasks. (**a**) High gamma modulation around the onset of the fixation. (**b**) Low gamma modulation around the offset of the fixation

Gamma and theta oscillations had been observed previously in several regions in the brain. In a previous study, theta activity had been seen in free moving rats [3]. Here, we show gamma and theta association during a quiet immobile period in CA1. Hippocampal areas CA1 and CA3 contribute to the retrieval of events in memory-dependent tasks [4]. Learning and cognition, on the other hand, appear to depend on the connectivity between CA1 and entorhinal cortex [5]. These different roles in the hippocampus reflect the different manipulation of received information. A recent study has reported that different waves of gamma were correlated with input from different areas in the hippocampus, with fast gamma originating from entorhinal cortex and slow gamma from CA3; these gamma oscillations occurred during different phases of the theta rhythm in CA1 [6].

Here we find that theta and gamma coupling occurs during the fixation period while rats are completely immobile and waiting for the next event to make their choice. We noted that the activation of the low gamma band during the memory-guided alternation task in CA1 was more pronounced than in the visual discrimination task, providing further empirical support for the role of hippocampus in sequence learning during memory-guided tasks.

Entorhinal cortex receives input from different reward-related areas in the brain like amygdala, and conveys the information to hippocampus. These different manipulations enable the extraction of contextual information and the translation of this information to codes necessary for decision-making. Previous studies have shown memory-guided sequence-dependent activity in hippocampus [7] and reported that the theta phase in CA3 was modulated by the low gamma band when rats learned to make associations in the spatial context [8]. In addition, type-two theta tended to increase during the immobile period [9].

Taken together with our results, this research puts emphasis on the important role of specific theta and gamma oscillations, and their coupling, during memory processing.

Acknowledgments This study was supported by Human Frontier Science program award RGP0039/2010; Grant-in-Aid from MEXT, 24120710 (Neural creativity for communication); the Tamagawa GCOE Program; and the Narishige Neuroscience Research Foundation.

References

1. Eichenbaum, H.: The hippocampus and declarative memory: cognitive mechanisms and neural codes. Behavioral Brain Research 127 (2001) 199–207.
2. Colgin, L.L. & Moser, E.I.: Gamma oscillation in the hippocampus. Physiology 25 (2010) 319–329.
3. O'Keefe, J. & Recce, M.L.: Phase relationship between hippocampal place units and the EEG theta rhythm. Hippocampus 3 (1993) 317–330.
4. Hasselmo, M.E.: The role of hippocampal regions CA3 and CA1 in matching entorhinal input with retrieval of associations between objects and context: Theoretical Comment on Lee et al. (2005) Behav. Neurosci. 1 (2005) 342–345.

5. Brun, V.H., Otnass, M.K., Molden, S., Steffenach, H.A., Witter, M.P., Moser, M.B. & Moser, E.I.: Place cells and place recognition maintained by direct entorhinal-hippocampus circuitry. Science 296 (2002) 2243–2246.
6. Colgin, L.L, Denninger, T., Fyhn, M., Hafting, T., Bonnevie, T., Jensen, O., Moser, M.B. & Moser E.I.: Frequency of gamma oscillations routes flow of information in the hippocampus. Nature 462 (2009) 353–357.
7. Ferbinteanu, J. & Shapiro, M.L.: Prospective and retrospective memory guided coding in the hippocampus. Neuron 40 (2003) 1227–1239.
8. Tort, A.B., Komorowski, R.W., Manns, J.R., Kopell, N.J. & Eichenbaum, H.: Theta-gamma coupling increases during learning of item-context associations. Proc. Nat. Acad. Sci. USA. 106 (2009) 20942–20947.
9. Takahashi, M., Lauwereyns, J. Sakurai, Y. & Tsukada, M.: A code for spatial alternation during fixation in rat hippocampal CA1 neurons. J. Neurophysiol. 102 (2009) 556–567.

Index

H. Liljenström (ed.), *Advances in Cognitive Neurodynamics (IV)*,
Advances in Cognitive Neurodynamics, DOI 10.1007/978-94-017-9548-7