

Research article

Open Access

Measuring information integration

Giulio Tononi*¹ and Olaf Sporns²

Address: ¹Department of Psychiatry, University of Wisconsin, Madison, USA and ²Department of Psychology, Indiana University, Bloomington, USA

Email: Giulio Tononi* - gtononi@wisc.edu; Olaf Sporns - osporns@indiana.edu

* Corresponding author

Published: 02 December 2003

Received: 26 August 2003

BMC Neuroscience 2003, 4:31

Accepted: 02 December 2003

This article is available from: <http://www.biomedcentral.com/1471-2202/4/31>

© 2003 Tononi and Sporns; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: To understand the functioning of distributed networks such as the brain, it is important to characterize their ability to integrate information. The paper considers a measure based on effective information, a quantity capturing all causal interactions that can occur between two parts of a system.

Results: The capacity to integrate information, or Φ , is given by the minimum amount of effective information that can be exchanged between two complementary parts of a subset. It is shown that this measure can be used to identify the subsets of a system that can integrate information, or complexes. The analysis is applied to idealized neural systems that differ in the organization of their connections. The results indicate that Φ is maximized by having each element develop a different connection pattern with the rest of the complex (functional specialization) while ensuring that a large amount of information can be exchanged across any bipartition of the network (functional integration).

Conclusion: Based on this analysis, the connectational organization of certain neural architectures, such as the thalamocortical system, are well suited to information integration, while that of others, such as the cerebellum, are not, with significant functional consequences. The proposed analysis of information integration should be applicable to other systems and networks.

Background

A standard concern of communication theory is assessing information transmission between a sender and a receiver through a channel [1,2]. Such an approach has been successfully employed in many areas, including neuroscience [3]. For example, by estimating the probability distribution of sensory inputs, one can show that peripheral sensory pathways are well suited to transmitting information to the central nervous system [4]. When considering the central nervous system itself, however, we face the issue of information *integration* [5]. In such a distributed network, any combination of neural elements can be viewed as senders or receivers. Moreover, the goal is not just to trans-

mit information, but rather to combine many sources of information within the network to obtain a unified picture of the environment and control behavior in a coherent manner [6,7]. Thus, while a neural system composed of a set of parallel, independent channels or reflex arcs may be extremely efficient for information transmission between separate inputs and outputs, it would be unsuitable for controlling behaviors requiring global access, context-sensitivity and flexibility. A system equipped with forward, backward and lateral connections among specialized elements can perform much better by supporting bottom-up, top-down, and cross-modal interactions and forming associations across different domains [8].

The requirements for information integration are perhaps best exemplified by the organization of the thalamocortical system of vertebrates. On the "information" side, there is overwhelming evidence for functional specialization, whereby different neural elements are activated in different circumstances, at multiple spatial scales [9]. Thus, the cerebral cortex is subdivided into systems dealing with different functions, such as vision, audition, motor control, planning, and many others. Each system in turn is subdivided into specialized areas, for example different visual areas are activated by shape, color, and motion. Within an area, different groups of neurons are further specialized, e.g. by responding to different directions of motion. On the "integration" side, the information conveyed by the activity of specialized groups of neurons must be combined, also at multiple spatial and temporal scales, to generate a multimodal model of the environment. For example, individual visual elements such as edges are grouped together to yield shapes according to Gestalt laws. Different attributes (shape, color, location, size) must be bound together to form objects, and multiple objects coexist within a single visual image. Images themselves are integrated with auditory, somatosensory, and proprioceptive inputs to yield a coherent, unified conscious scene [10,11]. Ample evidence indicates that neural integration is mediated by an extended network of intra- and inter-areal connections, resulting in the rapid synchronization of neuronal activity within and between areas [12-17].

While it is clear that the brain must be extremely good at integrating information, what exactly is meant by such a capacity, and how can it be assessed? Surprisingly, although tools for measuring the capacity to transmit, encode, or store information are well developed [2], ways of defining and assessing the capacity of a system to integrate information have rarely been considered. Several approaches could be followed to make the notion of information integration operational [5]. In this paper, we consider a measure based on effective information, a quantity capturing all *causal* interactions that *can* occur between two parts of a system [18]. Specifically, the capacity to integrate information is called Φ , and is given by the minimum amount of effective information that can be exchanged across a bipartition of a subset. We show that this measure can be used to identify the subsets of a system that can integrate information, which are called *complexes*. We then apply this analysis to idealized neural systems that differ in the organization of their connections. These examples provide an initial appreciation of how information integration can be measured and how it varies with different neural architectures. The proposed analysis of information integration may also be applicable to other non-neural systems and networks.

Theory

As in previous work, we consider an isolated system X with n elements whose activity is described by a Gaussian stationary multidimensional stochastic process [19]. Here we focus on the intrinsic properties of a neural system and hence do not consider extrinsic inputs from the environment. The elements could be thought of as groups of neurons, the activity variables as firing rates of such groups over hundreds of milliseconds, and the interactions among them as mediated through an anatomical connectivity $CON(X)$. The joint probability density function describing such a multivariate process can be characterized [2,20] in terms of entropy (H) and mutual information (MI).

Effective information

Consider a subset S of elements taken from the system. We want to measure the information generated when S enters a particular state out of its repertoire, but only to the extent that such information can be integrated, i.e. it can result from causal interactions within the system. To do so, we partition S into A and its complement B ($B = S - A$). The partition of S into two disjoint sets A and B whose union is S is indicated as $[A:B]_S$. We then give maximum entropy to the outputs from A , i.e. substitute its elements with independent noise sources of constrained maximum variance. Finally, we determine the entropy of the resulting responses of B (Fig. 1). In this way we define the *effective information* from A to B as:

$$EI(A \rightarrow B) = MI(A^{H^{\max}}:B) \quad (1)$$

where $MI(A:B) = H(A) + H(B) - H(AB)$ stands for mutual information, the standard measure of the entropy or information shared between a source (A) and a target (B). Since A is substituted by independent noise sources, the entropy that B shares with A is due to causal effects of A on B . In neural terms, we try out all possible combinations of firing patterns as outputs from A , and establish how differentiated is the repertoire of firing patterns they produce in B . Thus, if the connections between A and B are strong and specialized, different outputs from A will produce different firing patterns in B , and $EI(A \rightarrow B)$ will be high. On the other hand, if the connections between A and B are such that different outputs from A produce scant effects, or if the effect is always the same, then $EI(A \rightarrow B)$ will be low or zero. Note that, unlike measures of statistical dependence, effective information measures causal interactions and requires perturbing the outputs from A . Moreover, by enforcing independent noise sources in A , effective information measures all possible effects of A on B , not just those that are observed if the system were left to itself. Also, $EI(A \rightarrow B)$ and $EI(B \rightarrow A)$ will generally differ, i.e. effective information is not symmetric.

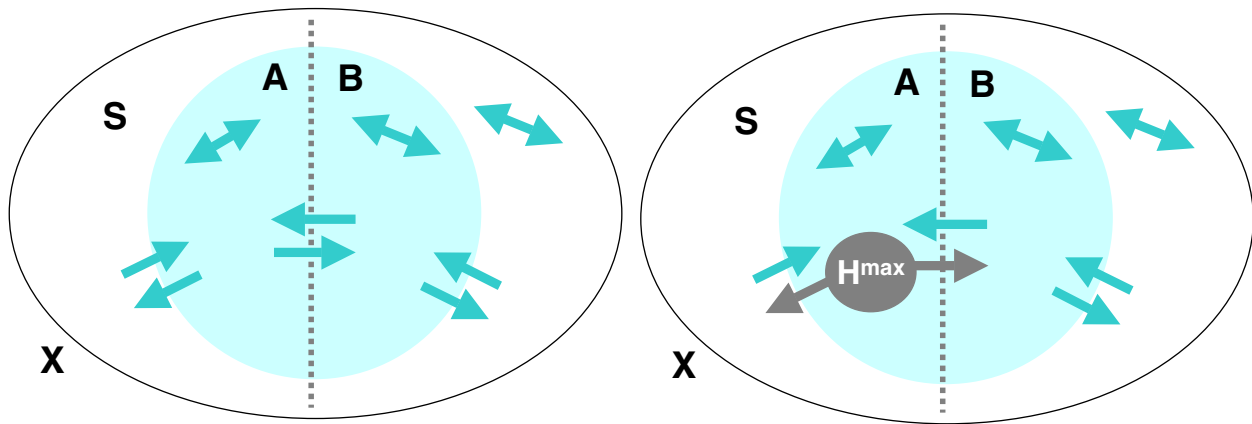


Figure 1

Schematics of effective information. Shown is a single subset S (grey ellipse) forming part of a larger system X. This subset is bisected into A and B by a bipartition (dotted line). Arrows indicate anatomical connections linking A to B and B to A across the bipartition, as well as linking both A and B to the rest of the system X. Bi-directional arrows indicate intrinsic connections within each subset and within the rest of the system. (Left) All connections are present. (Right) To measure $EI(A \rightarrow B)$, maximal entropy H^{\max} is injected into the outgoing connections from A (see Eq. 1). The resulting entropy of the states of B is then measured. Note that A can affect B directly through connections linking the two subsets, as well as indirectly via X.

For a given bipartition $[A:B]_S$ of subset S, the effective information is the sum of the effective information for both directions:

$$EI(A \rightleftharpoons B) = EI(A \rightarrow B) + EI(B \rightarrow A) \quad (2)$$

Information integration

Based on the notion of effective information for a bipartition, we can assess how much information can be integrated within a system of elements. To this end, we note that a subset S of elements cannot integrate any information if there is a way to partition S in two complementary parts A and B such that $EI(A \rightleftharpoons B) = 0$. In such a case we would be dealing with two (or more) causally independ-

ent subsets, rather than with a single, integrated subset. More generally, to measure how much information can be integrated within a subset S, we search for the bipartition(s) $[A:B]_S$ for which $EI(A \rightleftharpoons B)$ reaches a minimum. $EI(A \rightleftharpoons B)$ is necessarily bounded by the maximum entropy available to A or B, whichever is less. Thus, to be comparable over bipartitions, $\min\{EI(A \rightleftharpoons B)\}$ is normalized by $\min\{H^{\max}(A), H^{\max}(B)\}$. Thus, the *minimum information bipartition* of subset S, or $MIB(S)$, is its bipartition for which the normalized effective information reaches a minimum, corresponding to:

$MIB(S) = [A:B]_S$ for which $EI(A \rightleftharpoons B) / (\min\{H^{\max}(A), H^{\max}(B)\}) = \min$ for all A (3)

The capacity for *information integration* of subset S, or $\Phi(S)$, is simply the value of $EI(A \rightleftharpoons B)$ for the minimum information bipartition:

$$\Phi(S) = EI(MIB(S)) \quad (4)$$

The Greek letter Φ is meant to indicate the information (the "I" in Φ) that can be integrated within a single entity (the "O" in Φ). This quantity is also called *MIB complexity*, for minimum information bipartition complexity [18,21].

Complexes

If $\Phi(S)$ is calculated for every possible subset S of a system, one can establish which subsets are actually capable of integrating information, and how much of it. Consider thus every possible subset S of m elements out of the n elements of a system, starting with subsets of two elements ($k = 2$) and ending with a subset corresponding to the entire system ($k=n$). A subset S with $\Phi > 0$ is called a *complex* if it is not included within a subset having higher Φ [18,21]. For a given system, the complex with the maximum value of $\Phi(S)$ is called the *main complex*, where the maximum is taken over all combinations of $k > 1$ out of n elements of the system.

main complex (X) = $S \subseteq X$ for which $\Phi(S) = \max$ for all S (5)

In summary, a system can be analyzed to identify its complexes – those subsets of elements that can integrate information among themselves, and each complex will have an associated value of Φ – the amount of information it can integrate.

Results

In order to identify complexes and their $\Phi(S)$ for systems with many different connection patterns, we implemented each system X as a stationary multidimensional Gaussian process such that values for effective information could be obtained analytically (see Methods). In order to compare different connection patterns, the connection matrix $CON(X)$ was normalized so that the absolute value of the sum of the afferent synaptic weights per element corresponded to a constant value $w < 1$ (here, unless specified otherwise, $w = 0.5$). When considering a bipartition $[A:B]_S$, the independent Gaussian noise applied to A was multiplied by c_p , the perturbation coefficient, while the independent Gaussian noise applied to the rest of the system was given by c_i , the intrinsic noise coefficient. By varying c_p and c_i one can vary the signal-to-noise ratio (SNR) of the effects of A on the rest of the sys-

tem. In the following examples, except if specified otherwise, we set $c_p = 1$ and $c_i = 0.00001$ in order to emphasize the role of the connectivity and minimize that of noise.

An illustrative example

We first consider a system for which the results of the search for complexes should be easy to predict based on the anatomical connectivity. The system comprises 8 elements with a connection matrix corresponding to two fully interconnected modules, consisting of elements 1–4 and 5–7, respectively. These modules are completely disconnected from each other except for receiving common input from element 8 (Fig. 2A,2G). We derived the system's covariance matrix (Fig. 2B) based on the linear model and performed an exhaustive search for complexes as described in the Methods section. The search for minimum information bipartitions proceeded over subsets of sizes $k = 2, \dots, 8$, with a total of 247 individual subsets being examined. The main output is a list of all the subsets S of system X ranked by how much information can be integrated by each of them, i.e. their values of Φ . In Fig. 2C, the top 25 Φ values are ranked in descending order, regardless of the size of the corresponding subset. The composition of each subset corresponding to each Φ value is schematically represented in Fig. 2D, with different grey tones identifying the minimum information bipartition. The most relevant information is contained at the top of the matrix, which identifies the subsets with highest Φ . The subset at the top of Fig. 2C,2D consists of elements $\{1,2,3,4\}$ and has $\Phi = 20.7954$. Its minimum information bipartition (MIB) splits it between elements $\{1,2\}$ and $\{3,4\}$. The next subsets consist of various combinations of elements. However, many of them are included within a subset having higher Φ (higher up in the list) and are stripped from the list of complexes, which is shown in Fig. 2F. Fig. 2E shows the ranked values of Φ for each complex. In this illustrative example, our analysis revealed the existence of three complexes in the system: the one of highest Φ , corresponding to elements $\{1,2,3,4\}$, the second-highest, corresponding to elements $\{5,6,7\}$ with $\Phi = 20.1023$, and a complex corresponding to the entire system (elements $\{1,2,3,4,5,6,7,8\}$) with a much lower value of $\Phi = 7.4021$. As is evident from this example, the same elements can be part of different complexes.

Optimal networks for information integration

The above example illustrates how, in a simple case, the analysis of complexes correctly identifies subsets of a system that can integrate information. In general, however, establishing the relationship between the anatomical connectivity of a network and its ability to integrate information is more difficult. For example, it is not obvious which anatomical connection patterns are optimal for

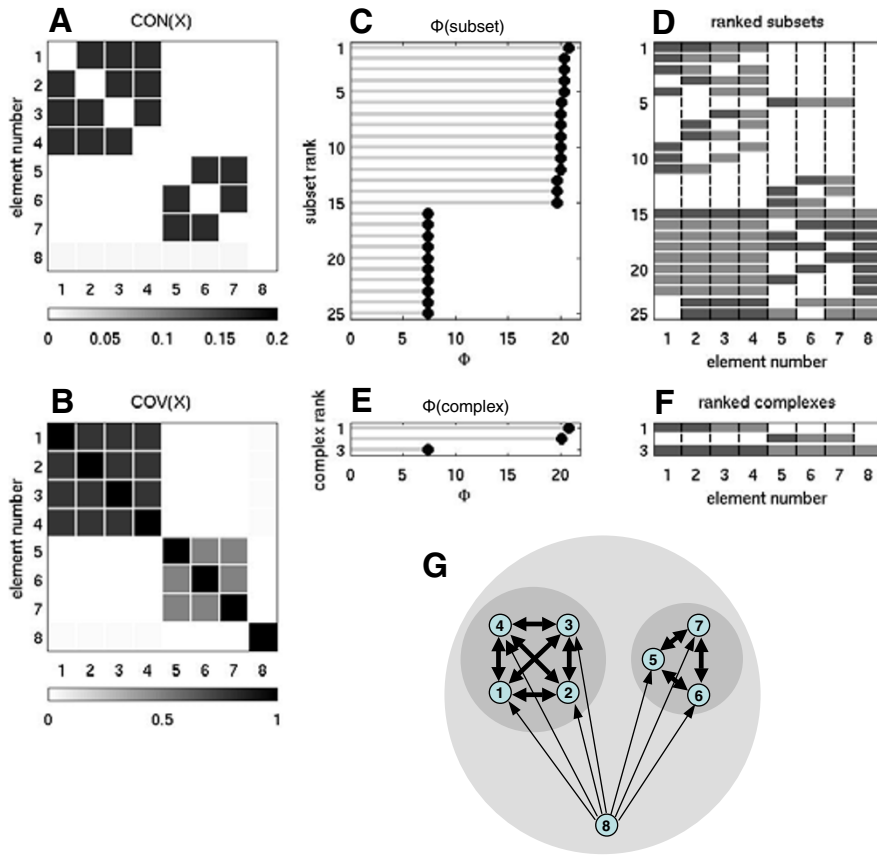


Figure 2

Measuring information integration: An illustrative example.

To measure information integration, we performed an exhaustive search of all subsets and bipartitions for a system of $n = 8$ elements. Noise levels were $c_i = 0.00001$, $c_p = 1$. (A) Connection matrix $CON(X)$. Connections linking elements 1 to 8 are plotted as a matrix of connection strengths (column elements = targets, row elements = sources). Connection strength is proportional to grey level (dark = strong connection, light = weak or absent connection). (B) Covariance matrix $COV(X)$. Covariance is indicated for elements 1 to 8 (corresponding to A). (C) Ranking of the top 25 values for Φ . (D) Element composition of subsets for the top 25 values of Φ (corresponding to panel C). Elements forming the subset S are indicated in grey, with two shades of grey indicating the bipartition into A and B across which the minimal value for EI was obtained. (E) Ranking of the Φ values for all complexes, i.e. subsets not included within subsets of higher Φ . (F) Element composition for the complexes ranked in panel E. (G) Digraph representation of the connections of system X (compare to panel A). Elements are numbered 1 to 8, arrows indicate directed edges, arrow weight indicates connection strength. Grey overlays indicate complexes with grey level proportional to their value of Φ . Figs. 2 to 7 use the same layout to represent computational results.

information integration, given certain constraints on number of elements, total amount of synaptic input per element, and SNR.

Fig. 3 shows connection patterns and complexes obtained from a representative example of nonlinear constrained optimization of the connection matrix $CON(X)$, starting from random connection matrices ($n = 8$, all columns

normalized to $w = 0.5$, no self-connections, and high SNR, corresponding to $c_i = 0.00001$, $c_p = 1$). A large number of optimizations for Φ were carried out and produced networks having high values of information integration ($\Phi = 73.6039 \pm 0.5352$ for 343 runs) that shared several structural features. In all cases examined, complexes with highest (optimized) values of Φ were of maximal size n . After optimization, connection weights

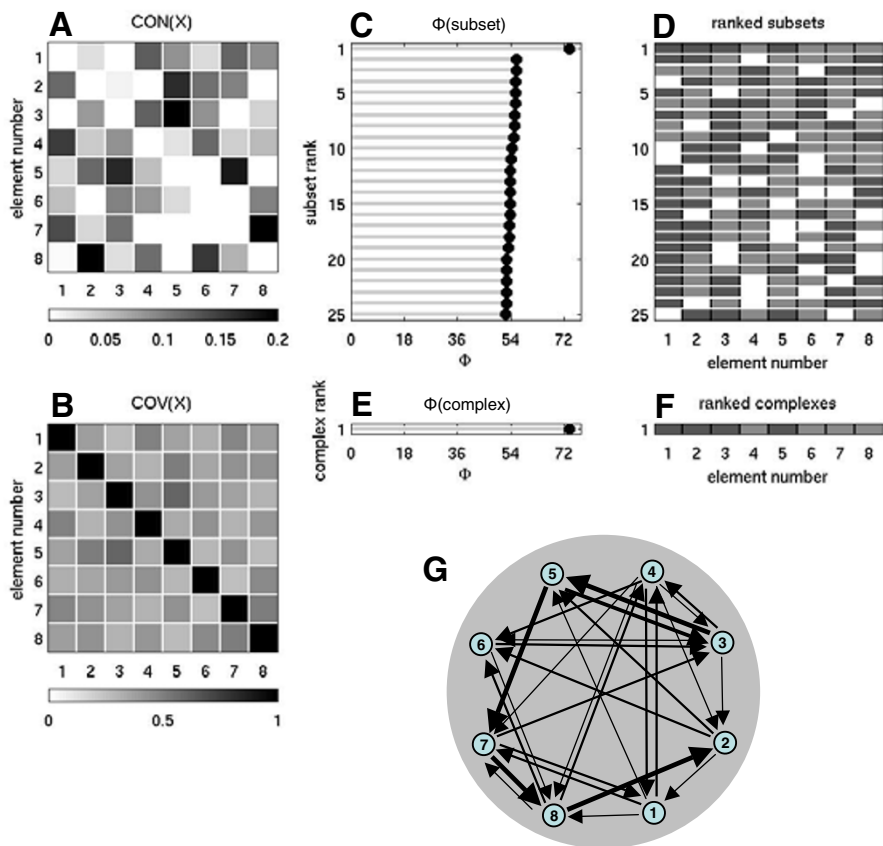


Figure 3
Information integration and complexes for an optimized network at high SNR. Shown is a representative example for $n = 8$, $w = 0.5$, $c_i = 0.00001$, $c_p = 1$. Note the heterogeneous arrangement of the incoming and outgoing connections for each element.

showed a characteristic distribution with a large proportion of weak connections and a gradual fall-off towards connections with higher strength ("fat tail"). The column vectors of optimized connection matrices were highly decorrelated, as indicated by an average angle between vectors of $63.6139 \pm 2.5261^\circ$ (343 optimized networks). Thus, the pattern of synaptic inputs were different for different elements, ensuring heterogeneity and specialization. On the other hand, all optimized connection matrices were strongly connected (all elements could be reached from all other elements of the network), ensuring information integration across all elements of the system.

Under conditions of low SNR ($c_i = 0.1$, $c_p = 1$), optimized values of Φ (albeit much lower) were again obtained for complexes of size n (Fig. 4, $\Phi = 5.7454 \pm 0.1189$ for 21

runs). However, the optimized connection patterns had different weight distributions and topologies compared to those obtained with high SNR. Connection weights showed a highly bimodal distribution, with most connections having decayed effectively to strengths of zero and a few connections remaining at high strength, corresponding to a sparse connection pattern. Again, decorrelation of synaptic input vectors was high ($78.9453 \pm 1.3946^\circ$ for 21 optimized networks), while at the same time networks remained strongly connected.

In order to characterize optimal networks for information integration in graph-theoretical terms, we examined sparse networks with fixed numbers of connections of equal strength. Starting from networks of size $n = 8$ having 16 connections arranged at random, we used an

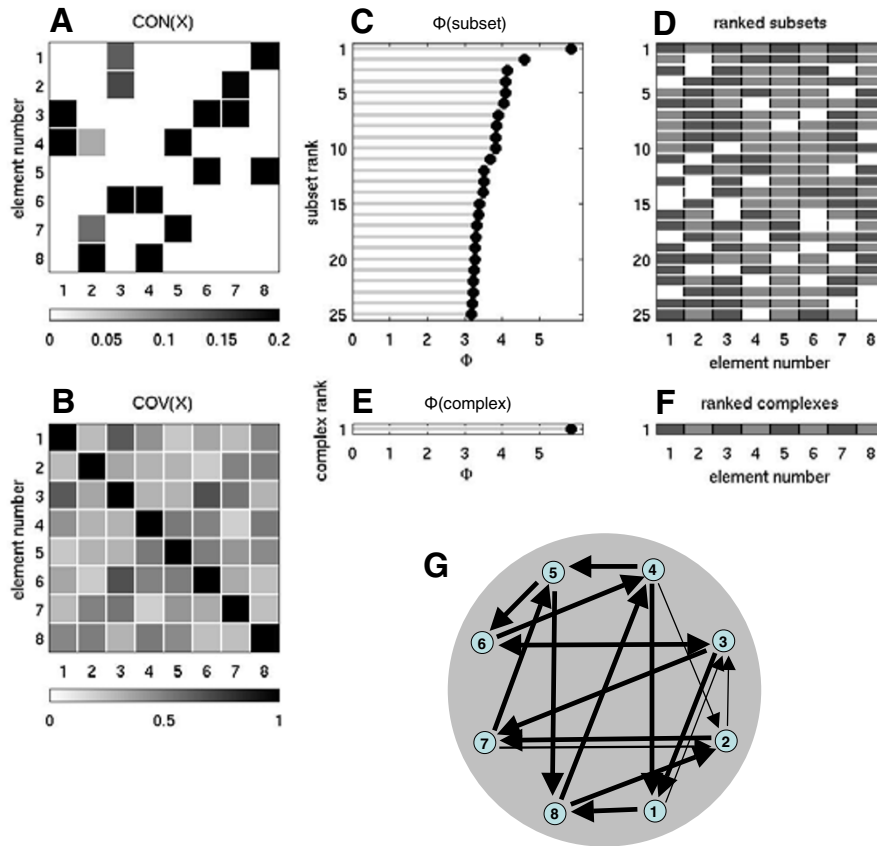


Figure 4
Information integration and complexes for an optimized network at low SNR. Shown is a representative example for $n = 8$, $w = 0.5$, $c_i = 0.1$, $c_p = 1$. Note the sparse structure of the optimized connection pattern.

evolutionary algorithm [22] to rearrange the connections for optimal information integration. Also in this case, optimal networks yielded complexes corresponding to the entire system, both for high ($c_i = 0.00001$, $\Phi = 60.7598$, Fig. 5) and low levels of SNR ($c_i = 0.1$, $\Phi = 5.943$, not shown). Optimized sparse networks shared several structural characteristics that emerged in the absence of any structural constraints imposed on the evolutionary algorithm. As with non-sparse networks, heterogeneity of the connection pattern was high, with no two elements sharing sets of either inputs or outputs. To quantify the similarity of input patterns, we calculated the matching index [23,24], which reflects the correlation (scaled between 0 and 1) between discrete vectors of input connections to a pair of vertices, excluding self- and cross-connections. After averaging across all pairs of elements of each optimized network, the average matching index was 0.1429,

indicating extremely low overlap between input patterns. Also, despite the sparse connectivity, all optimized networks were strongly connected. In-degree and out-degree distributions, which record the number of afferent and efferent connections per element, tended to be balanced, with most if not all elements receiving and emitting two connections each (see example in Fig. 5). Consistent with results from nonlinear optimization runs, networks showed few (if any) direct reciprocal connections but had an over-abundance of short cycles. For example, an optimal pattern having maximal symmetry was made up by nested cycles of length 3 (3-cycles, or indirect reciprocal, as shown in Fig. 10). Finally, structural characteristics of optimized sparse networks differed significantly from those of random sparse networks. The latter had a much lower value of Φ ($\Phi = 35.6622 \pm 5.0382$, 100 exemplars) and complexes of maximal size $n = 8$ were found only in

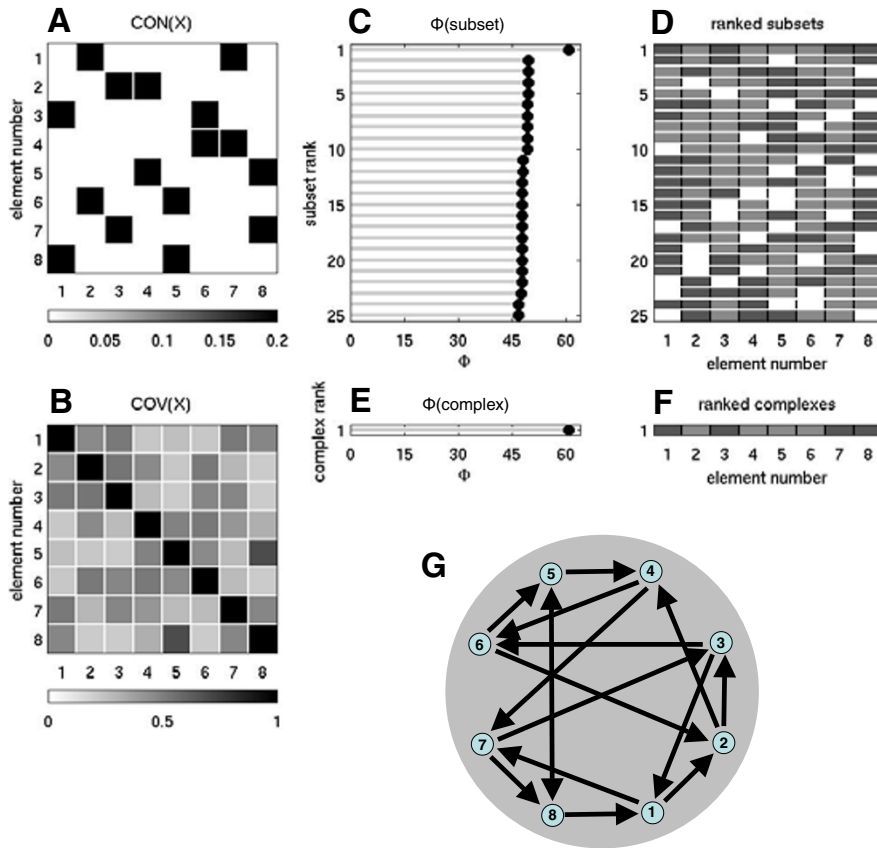


Figure 5
Information integration and complexes for an optimized sparse network having fixed number of connections of equal strength. The network was obtained through an evolutionary rewiring algorithm ($n = 8$, 16 connections of weight 0.25 each, $c_i = 0.00001$, $c_p = 1$). Note the heterogeneous arrangement of the incoming and outgoing connections for each element, the balanced degree distribution (two afferent and two efferent connections per element) and the low number of direct reciprocal connections.

a fraction (46/100) of cases. Calculation of the matching index revealed a higher degree of similarity between input vectors (average of 0.2180), despite random generation. Moreover, only a small number of random networks were strongly connected (13/100) and the fraction of direct reciprocal connections was higher.

In summary, all optimized networks, irrespective of size, sparse or full connectivity, and levels of noise, satisfied two opposing requirements – that for specialization and that for integration. The former corresponded to the strong tendency of connection patterns of different elements to be as heterogeneous as possible, yielding elements with different functional roles. The latter was revealed by the fact that optimized networks, in all cases

examined in this study and for any combination of parameters or constraints, formed a single, highly connected complex of maximal size.

Homogeneous and modular networks

To examine the influence of homogeneity and modularity of connection patterns on the integration of information, it is useful to study how values of Φ respond to changes in these parameters. Specifically, if heterogeneity (specialization) is essential, Φ should decrease considerably for connection matrices that lack such heterogeneity, i.e. are fully homogeneous. Fig. 6 shows a fully connected network where all connection weights had the same value $CON_{ij} = 0.072$, i.e. a fully *homogenous* network (matching index = 1). Its main complex contained all elements of the system,

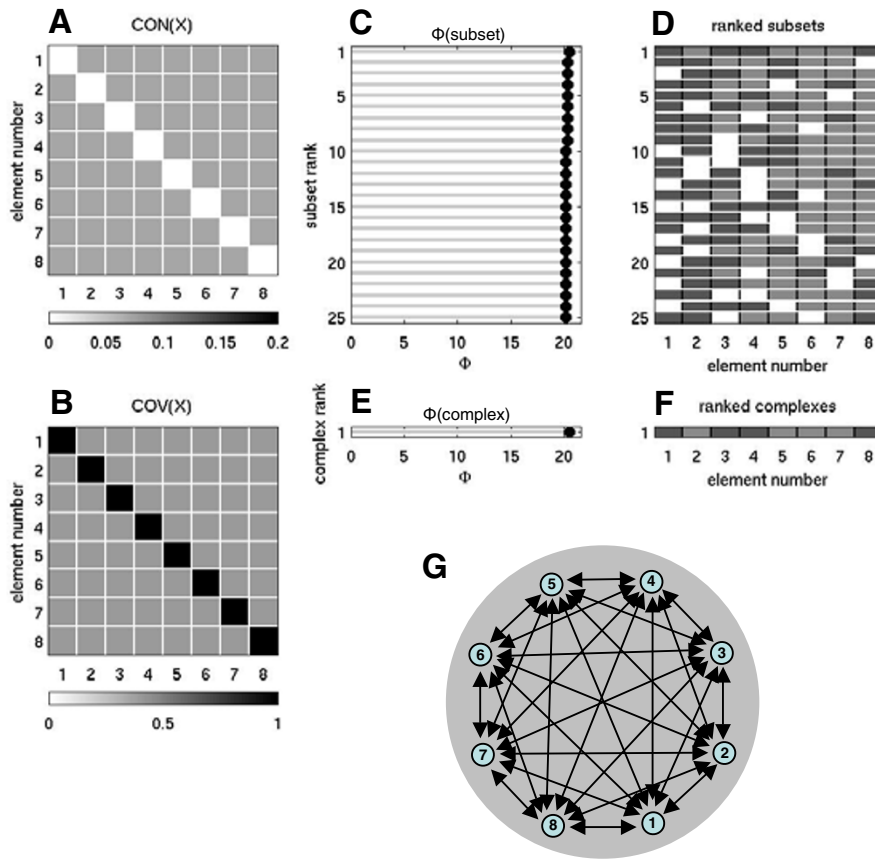


Figure 6
Information integration and complexes for a homogeneous network. Connectivity is full and all connections weights are the same ($n = 8$, $w = 0.5$, $c_i = 0.00001$, $c_p = 1$).

but its ability to integrate information was much less than that of optimized networks ($\Phi = 20.5203$ against 73.6039 , for $n = 8$ and $c_i = 0.00001$). This result was obtained irrespective of the size of the system or SNR.

On the other hand, if forming a single large complex (integration) is essential to reach high values of Φ , then a system that is composed of several smaller complexes (e.g. "modules") should integrate less information and have lower values of Φ . Fig. 7 shows a strongly modular network, consisting of four modules of two strongly interconnected elements ($CON_{ij} = 0.25$) that were weakly linked by inter-module connections ($CON_{ij} = 0.0417$). This network yielded $\Phi = 20.3611$ for each of its four modules, which formed the system's four main complexes. A complex consisting of the entire network was also obtained,

but its Φ value (19.4423) was lower than that of the individual modules.

Finally, to evaluate how Φ changes by varying parametrically along the continuum between full modularity to full homogeneity, we plotted the value of Φ for a series of Toeplitz connection matrices of Gaussian form for increasing standard deviation σ (Fig. 8A,8B; Toeplitz matrices have constant coefficients along all subdiagonals). As σ was varied from 10^{-3} (complete modularity, self-connections only) to 10^1 (complete homogeneity, all connections of equal weight), Φ reached a maximum for intermediate values of σ , when the coefficients in the matrix spanned the entire range between 0 and 1.

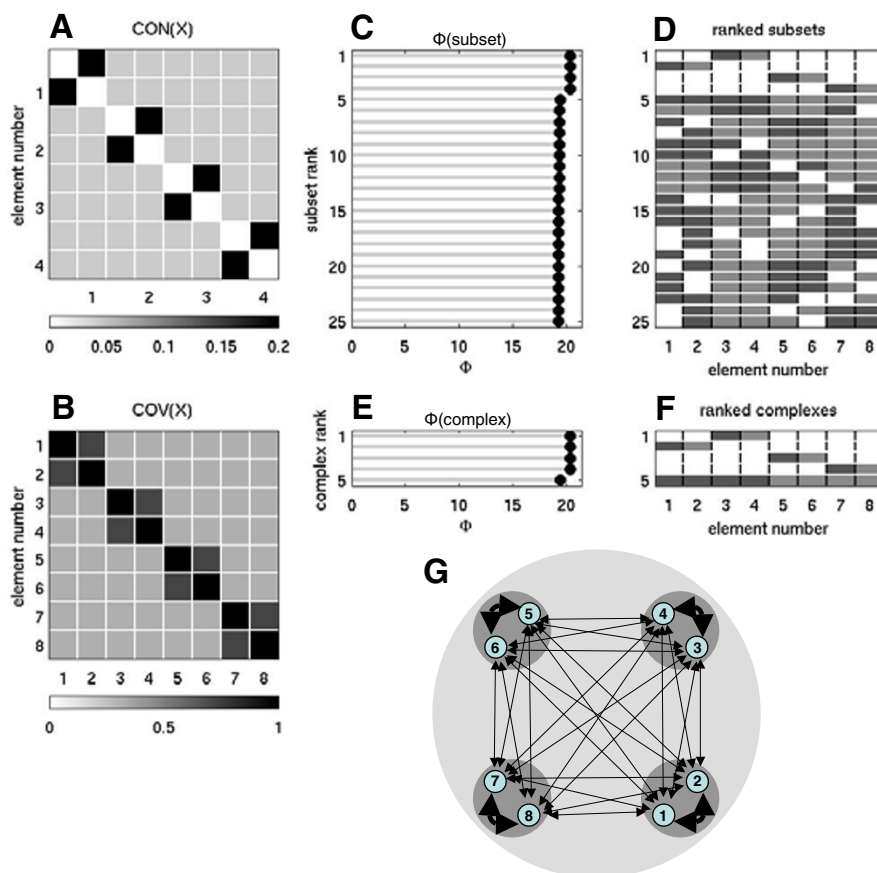


Figure 7
Information integration and complexes for a strong modular network. Weights of inter-modular connections are 0.0417 ($n = 8, w = 0.5, c_i = 0.00001, c_p = 1$).

Basic digraphs and their Φ

In this section, we examine the capacity to integrate information of a few basic digraphs (directed graphs) that constitute frequent motifs in biological systems, especially in neural connection patterns (Fig. 9). A *directed path* (Fig. 9A) consists of a series of directed connections (edges) linking elements (vertices) such that no element is repeated. A directed path composed of connections of equal weight linking n elements formed a complex of size n with a low value of $\Phi = 10.1266$ ($n = 8, \text{connection strength} = 0.25, c_i = 0.00001, c_p = 1$). The minimum information bipartition (MIB) always separated two connected elements. The value for Φ was fixed irrespective of the total length of the path. Addition of a connection linking the last and first elements results in a closed path or *cycle* (Fig. 9B). The complex integrating maximal information corresponded to the entire system and its MIB corre-

sponded to a cut into two contiguous halves. The value of $\Phi = 20.2533$ was exactly twice that for the directed path (Fig. 9A) and remained constant irrespective of the length of the cycle. A *two-way cycle* with reciprocal connections linking pairs of elements as nearest neighbors (Fig. 9C) had a main complex equal to the entire system, with a MIB that separated the system into two halves. The value for Φ was 40.5065, almost twice that of the one-way cycle, and was constant irrespective of length ($n > 4$). If the system is a *fan-out* (divergent) digraph (Fig. 9D) with a single element (master) driving or modulating several other elements (slaves), the entire system formed the main complex, the MIB corresponded to a cut into two contiguous halves, and the value of Φ was low ($\Phi = 10.8198$). However, this architecture became more efficient, compared to others, under low SNR. Predictably, the fan-out architecture was rather insensitive to size under high SNR.

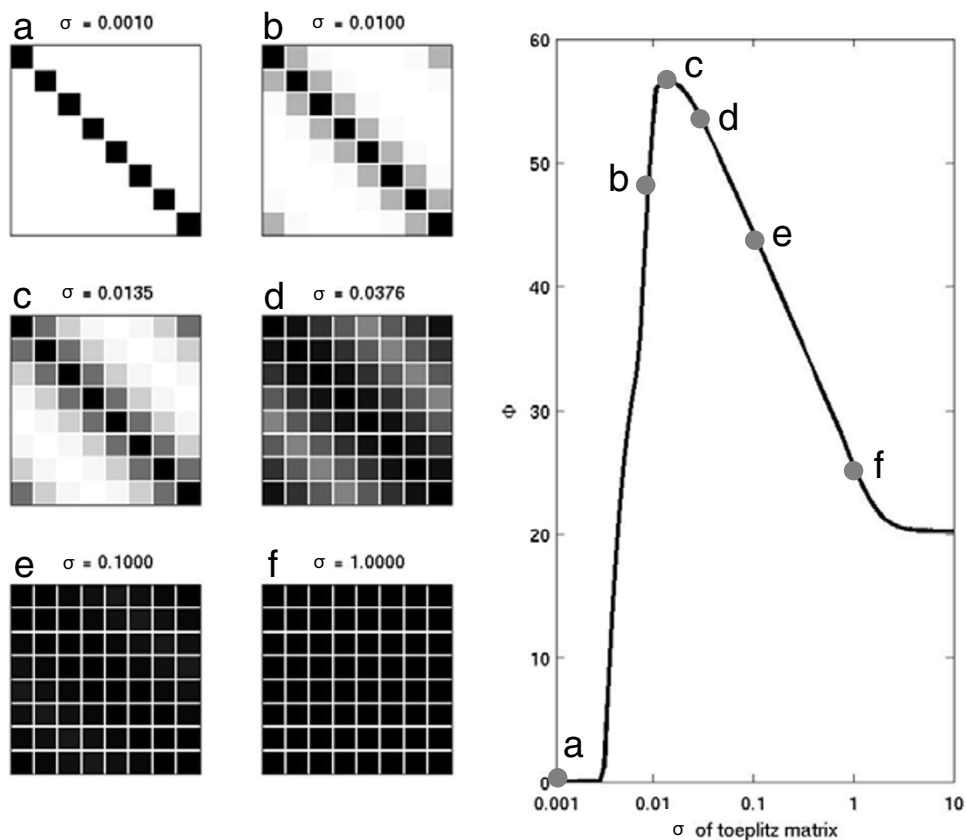


Figure 8

Information integration as a function of modularity – homogeneity. Values of Φ were obtained from Gaussian Toeplitz connection matrices ($n = 8$) with a fixed amount of total synaptic weight (self-connections allowed). The plot shows Φ as a function of the standard deviation σ of the Toeplitz connection profile. Connection matrices (cases a to f) are shown at the left, for different values of σ . For very low values of σ ($\sigma = 0.001$, case a), Φ is zero: the system is made up of 8 causally independent modules. For intermediate values of σ ($\sigma \approx 0.01$ to 0.1, cases b, c, d, e), Φ increases and reaches a maximum; the elements are interacting in a heterogeneous way. For high values of σ ($\sigma = 1$, case f) Φ is low; the interactions among the elements are completely homogeneous.

If the system is a *fan-in* (convergent) digraph (Fig. 9E), where a single element is driven by several other elements, the entire system formed the main complex and MIB corresponded to a cut into two contiguous halves with low Φ values ($\Phi = 10.8198$).

It is also interesting to consider what happens if paths, cycles, or a fan-out architecture are added onto a complex of high Φ (Fig. 10). Specifically, adding a path did not change the composition of the main complex, and did not change its Φ if the path was out-going from the complex (Fig. 10A; the source element of the path in the complex is called a *port out*). If the path was in-going (Fig. 10B; the

target element of the path in the complex is called a *port in*), the Φ value of the main complex could be slightly altered, but its composition did not change. The addition of an in-going or an out-going path did form a new complex including part or all the main complex plus the path, but its Φ value was much lower than that of the original complex, and corresponded to that of the isolated path, with the MIB cutting between the path and the port in/out. Adding cycles to a complex (Fig. 10C) also generally did not change the main complex, although Φ values could be altered. It should be noted that, even if the elements of a cycle do not become part of the main complex (except for ports in and ports out) they can provide an

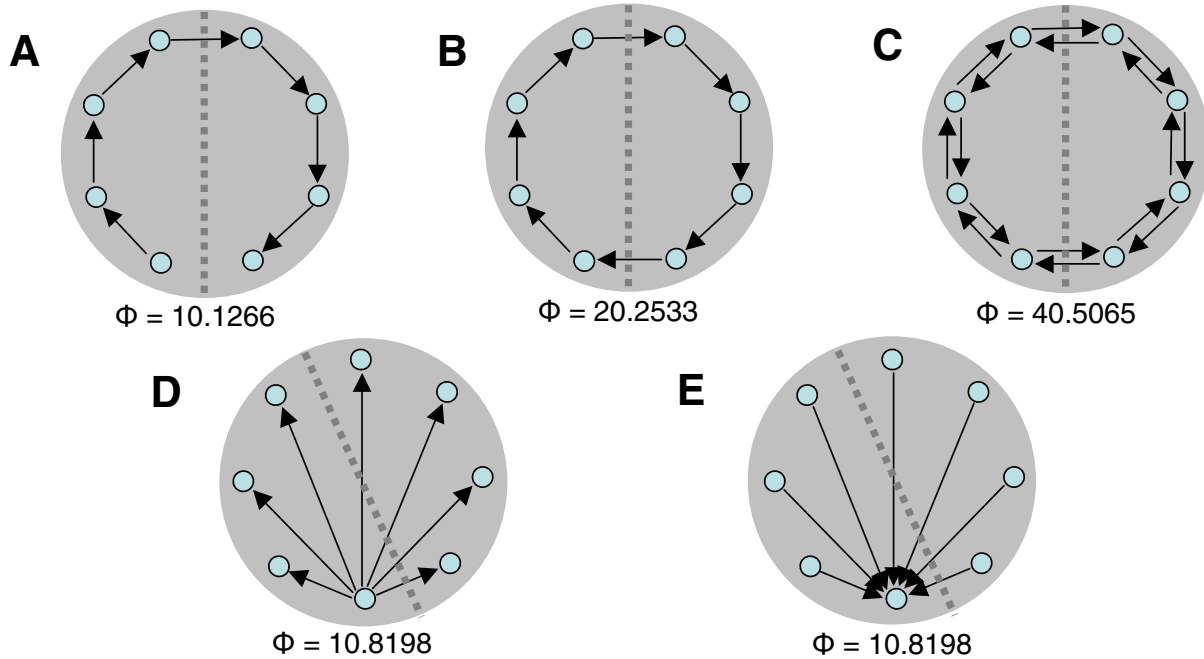


Figure 9
Information integration for basic digraphs. (A) Directed path. (B) One-way cycle. (C) Two-way cycle. (D) Fan-out digraph. (E) Fan-in digraph. Complexes are shaded and values for Φ are provided in each of the panels.

indirect connection between two elements of the main complex. Thus, information integration within a complex can be affected by connections and elements outside the complex.

Scaling and joining complexes

Because of the computational requirements, an exhaustive analysis of information integration is feasible only for small networks. While heuristic search algorithms may be employed to aid in the analysis of larger networks, here we report a few observations that can be made using networks with $n \leq 12$. Keeping all parameters except size the same, maximum values of Φ achieved using nonlinear constrained optimization grew nearly linearly at both high and low SNR, e.g. increasing from $\Phi = 39.95$ at size 4 to $\Phi = 103.96$ at size 12 ($w = 0.5, c_i = 0.00001, c_p = 1$).

For larger networks, and especially for networks of biological significance, a relevant question is whether high levels of information integration can be obtained by joining smaller components. To investigate this question, we first optimized two small component networks with $n = 8$ and then joined them through 8 pairs of reciprocal connections (Fig. 11A,11B). We refer to the connections inside each component as "intra-modular", and to those between the two components as "inter-modular". Before being joined the two optimized component networks each had $\Phi = 60.7598$. After being joined by inter-modular connections, whose strength was equal to the total strength of intra-modular connections, the value of Φ for the system of size $n = 16$ was $\Phi = 109.5520$. This represents a significant gain in information integration without further optimization, considerably exceeding the average

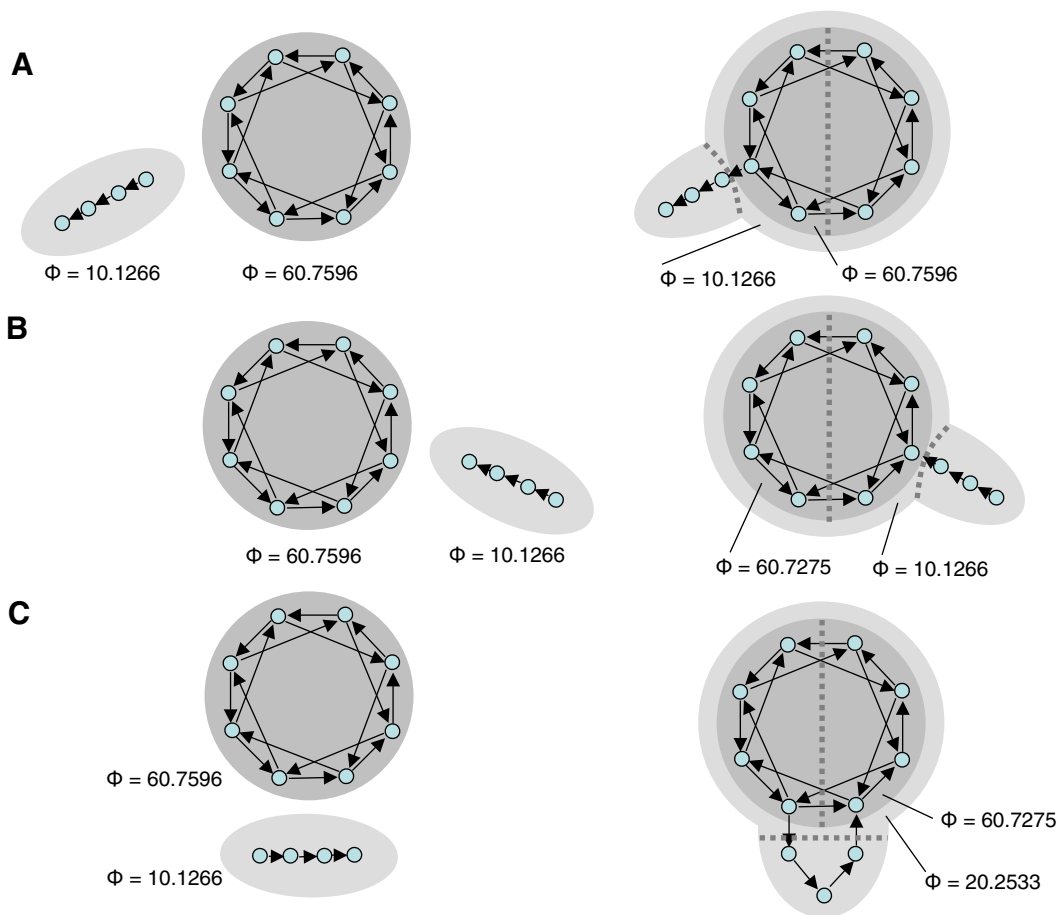


Figure 10
Adding paths and cycles to a main complex. (A) Out-going path added to a complex. Complex and directed path are shown separately on the left and joined on the right. (B) In-coming path added to a complex. (C) Cycle added to a complex. Complexes are shaded and values for Φ are provided in each of the panels.

Φ for random networks of size $n = 16$ ($\Phi = 51.5930 \pm 5.2275$, 10 exemplars).

The stability of the maximal complex and its capacity to integrate information depends on the relative synaptic strengths of intra-modular and inter-modular connections. In Fig. 11C, we plot the values for Φ and the size of the corresponding complex for different inter-modular coupling strengths, using two $n = 8$ sparse optimized components joined by 8 pairs of reciprocal connections. At very low levels of inter-modular coupling, the two component networks formed separate complexes (see modular networks above) and their Φ was close to the average Φ attained by optimizing networks of size $n = 8$ ($\Phi \approx 60.7598$). As coupling increases, at a coupling strength of approximately 0.001, the two components

suddenly joined as one main complex. This complex reached a maximal $\Phi \approx 109.6334$ at a coupling ratio of 2:3 for intra- versus inter-modular connections. Stronger coupling resulted in weaker information integration.

Our results demonstrate that the value for Φ attained by simply coupling optimized smaller components substantially exceeds that of random networks of equivalent size, although it typically does not reach the Φ of optimized networks. This gain in integrated information achieved by joining complexes suggests that larger networks capable of high information integration may be constructed efficiently starting from small optimized components.

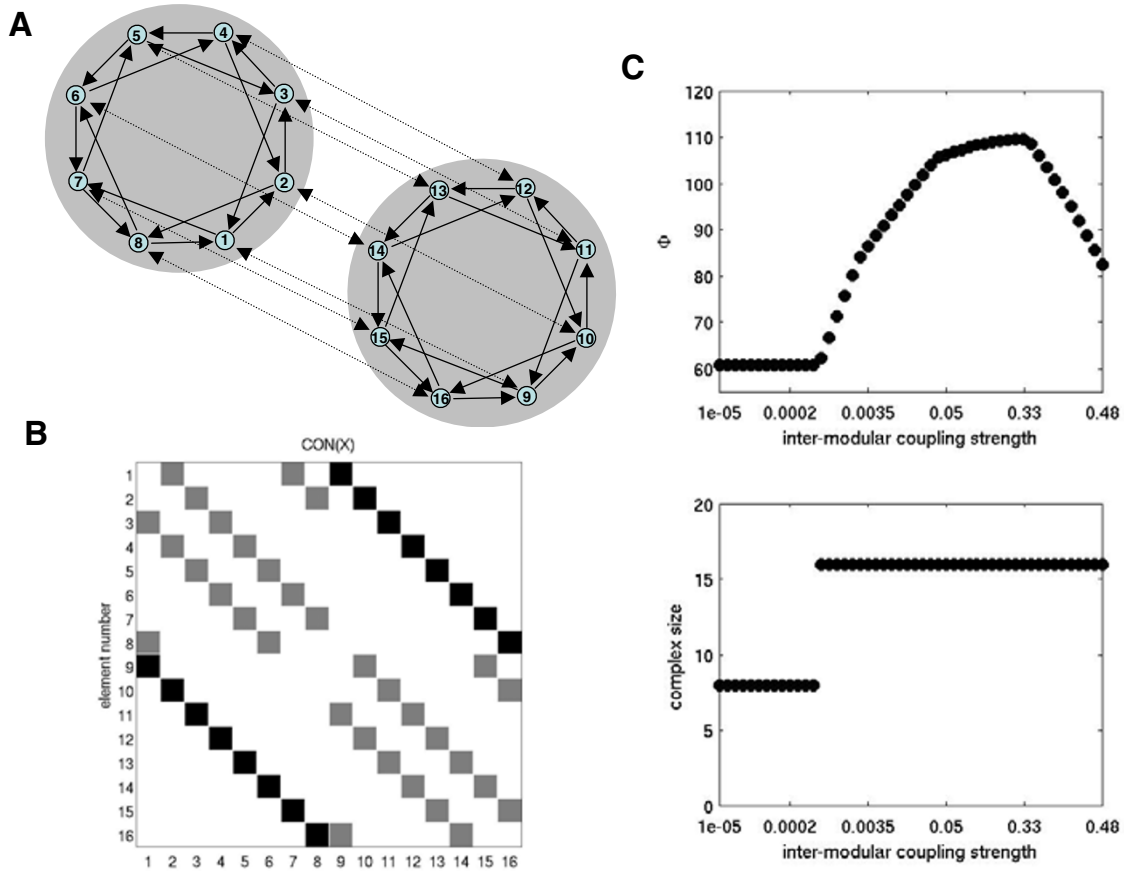


Figure 11
Joining complexes. (A) Graph representation of two optimized complexes ($n = 8$ each), joined by reciprocal "inter-modular" connections (stippled arrows). For simplicity, these connections are arranged as n pairs of bi-directional connections, representing a simple topographic mapping. (B) Connection matrix of a network of size $n = 16$ constructed from two smaller optimized components ($n = 8$). (C) Information integration as a function of inter-modular coupling strength. Connection strength is given as inter-modular CON_{ij} values; all intra-modular connections per element add up to $0.5 - CON_{ij}$. Plot at the bottom shows the size of the complex with highest Φ as a function of inter-modular coupling strength.

Discussion

In this paper, we have examined how the capacity to integrate information can be measured and how it varies depending upon the organization of network connections. The measure of information integration – Φ – corresponds to the minimum amount of effective information that can be exchanged across a bipartition of a subset. A subset of elements capable of integrating information that is not part of a subset having higher Φ constitutes a complex. As was shown here with linear systems having different architectures, these measures can be used to identify subsets of elements that can integrate information and to measure its amount.

Measuring information integration

Despite the significant development of techniques for quantifying information transmission, encoding and storage, measures for information integration have been comparatively neglected. One reason for such neglect may be that information theory was originally developed in the context of communication engineering [1]. In such application, the sender, channel, and receiver are typically given, and the goal is to maximize information transmission. In distributed networks such as the central nervous system, however, most nodes can be both senders and receivers, multiple pathways can serve as channels, and a key goal is to integrate information for behavioral control [5].

In previous work, we introduced a measure called neural complexity (C_N), which corresponds to the average mutual information for all bipartitions of a system [19]. We showed that C_N is low if the elements of the system are connected too sparsely (no integration), or if they are connected too homogeneously (no specialization), but it is high if the elements are connected densely *and* in a specific manner (both integration *and* specialization). C_N captures an important aspect of the *average* integration of information within a distributed system. However, it is insensitive to whether the system constitutes a single, integrated entity, or is actually made up of independent subsets. For example, based on C_N alone it is not possible to know whether a set of elements is integrated or is arranged into parallel, independent channels. To identify which elements of a system constitute an integrated subset, we introduced an index of functional clustering (CI), which reflects the ratio between the overall statistical dependence within a subset over the statistical dependence between that subset and the rest of the system [25]. Because it is based on a ratio, CI requires statistical comparison to a null hypothesis to rank subsets of different size. Like C_N , CI cannot detect whether elements are merely correlated, e.g. because of common input, or are causally interacting.

The procedure discussed here represents a direct attempt at measuring integration information, and it differs in several important respects from the above-mentioned two-step process of finding functional clusters and measuring the average mutual information between their bipartitions. The analysis of complexes identifies subsets in a single step based on their ability to integrate information, i.e. their Φ value. Moreover, Φ measures the actual amount of information that can be integrated, rather than an average value, and eliminates the need for statistical comparisons to a null hypothesis. Φ provides therefore a meaningful metric to cluster subsets of neural elements depending on the amount of information they can exchange. At a more fundamental level, information can only be integrated if there are actual interactions within a system. Accordingly, because it is based on effective information rather than on mutual information, Φ measures causal interactions, not just statistical dependencies. In neural terms, this means that Φ characterizes not just the functional connectivity of a network, but the entire range of its effective connectivity (in neuroimaging studies, effective connectivity is generally characterized as a change in the strength of interactions between brain areas associated with time, attentional set, etc. [26,27]). Finally, similar to the capacity for information transmission of a channel [1], the capacity for information integration of a system is a fundamental property that should only depend on the system's parameters and not be affected by observation time or nonstationarities. Accordingly, Φ

captures all interactions that are possible in a system, not just those that happen to be observed over a period of time.

Information integration and basic neuroanatomy

A characterization of the structural factors that affect information integration is especially useful when considering intricately connected networks such as the nervous system. Using linear systems of interconnected elements, we have examined how, given certain constraints, the organization of connection patterns influences information integration within complexes of elements. The results suggest that only certain architectures, biological or otherwise, are capable of jointly satisfying the requirements posed by specialization and integration and yield high values of Φ .

The present analysis indicates that, at the most general level, networks yielding high values of Φ must satisfy two key requirements: i) the connection patterns of different elements must be highly heterogeneous, corresponding to high specialization; ii) networks must be highly connected, corresponding to high integration (Figs. 3,4,5). If one or the other of these two requirements is not satisfied, the capacity to integrate information is greatly reduced. We have shown that homogeneously connected networks are not suited to integrating large amounts of information because specialization is lost (Fig. 6), while strongly modular networks fail because of a lack of integration (Fig. 7). Moreover, we have shown that randomly connected networks do not perform well given constraints on the amount of connections or in the presence of noise, since optimal arrangements of connection patterns are highly specific. We have also examined simple digraphs that are widely represented in biology, such as paths, cycles, and pure convergence and divergence (Fig. 8). Such simple arrangements do not yield by themselves high values of Φ , but they can be helpful to convey the inputs and outputs of a complex, to mediate indirect interactions between elements of a complex or of different complexes, and for global signalling. Finally, we have seen that, while optimizing connection patterns for Φ is unrealistic for large networks, high values of Φ can be achieved by joining smaller complexes by means of reciprocal connections (Fig. 11). These observations are of some interest when considering basic aspects of neuroanatomical organization.

Thalamocortical system

The above analysis indicates that Φ is maximized by having each element develop a different connection pattern with the rest of the complex (functional specialization) while ensuring that, irrespective of how the system is divided in two, a large amount of information can be exchanged between the two parts (functional integration).

The thalamocortical system appears to be an excellent candidate for the rapid integration of information. It comprises a large number of widely distributed elements that are functionally specialized [8,9]. Anatomical analysis has shown that specialized areas of cortex maintain heterogeneous sets of afferent and efferent pathways, presumably crucial to generate their individual functional properties [28]. At the same time, these elements are intricately interconnected, both within and between areas, and these interconnections show an overabundance of short cycles [12,29]. These connections mediate effective interactions within and between areas, as revealed by the occurrence of short- and long-range synchronization, a hallmark of functional integration, both spontaneously and during cognitive tasks [13-17]. Altogether, such an organization is reminiscent of that of complexes of near-optimal Φ . The observation that it is possible to construct a larger complex of high Φ by joining smaller, near-optimal complexes in a fairly straightforward manner, may also be relevant in this context. For example, while biological learning algorithms may achieve near-optimal information integration on a local scale, it would be very convenient if large complexes having high values of Φ could be obtained by laying down reciprocal connections between smaller building blocks.

From a functional point of view, an architecture yielding a large complex with high values of Φ is ideally suited to support cognitive functions that call for rapid, bidirectional interactions among specialized brain areas. Consistent with these theoretical considerations, perceptual, attentional, and executive functions carried out by the thalamocortical system rely on reentrant interactions [30], show context-dependency [30-32], call for global access [10,11], require that top down predictions are refined against incoming signals [27], and often occur in a controlled processing mode [33]. Considerable evidence also indicates that the thalamocortical system constitutes the necessary and sufficient substrate for the generation of conscious experience [11,34-36]. We have suggested that, at the phenomenological level, two key properties of consciousness are that it is both highly informative and integrated [11,18]. Informative, because the occurrence of each conscious experience represents one outcome out of a very large repertoire of possible states; integrated, because the repertoire of possible conscious states belongs to a single, unified entity – the experiencing subject. These two properties reflect precisely, at the phenomenological level, the ability to integrate information [21]. To the extent that consciousness has to do with the ability to integrate information, its generation would require a system having high values of Φ . The hypothesis that the organization of the thalamocortical system is well suited to giving rise to a large complex of high Φ would then provide some rationale as to why this part of the brain

appears to constitute the neural substrate of consciousness, while other portions of the brain, similarly equipped in terms of number of neurons, connections, and neurotransmitters, do not contribute much to it [21].

Cerebellum

This brain region contains probably more neurons and as many connections as the cerebral cortex, receives mapped inputs from the environment and controls several outputs. However, the organization of synaptic connections within the cerebellum is radically different from that of the thalamocortical system, and it is rather reminiscent of the strongly modular systems analyzed here. Specifically, the organization of the connections is such that individual patches of cerebellar cortex tend to be activated independently of one another, with little interaction between distant patches [37,38]. This suggests that cerebellar connections may not be organized so as to generate a large complex of high Φ , but rather very many small complexes each with a low value of Φ . Such an organization seems to be highly suited for the rapid, effortless execution and refinement of informationally encapsulated routines, which is thought to be the substrate of automatic processing. Moreover, consistent with the notion that conscious experience requires information integration, even extensive cerebellar lesions or cerebellar ablations have little effect on consciousness.

Cortical input and output systems, cortico-subcortical loops

According to the present analysis, circuits providing input to a complex, while obviously transmitting information to it, do not add to its ability to integrate information if their effects are entirely accounted for by the elements of the complex to which they connect. This situation resembles the organization of early sensory pathways and of motor pathways. Similar conclusions apply to cycles attached to a main complex at both ends, a situation exemplified by parallel cortico-subcortico-cortical loops. Consider for example the basal ganglia – large neural structures that contain many circuits arranged in parallel, some implicated in motor and oculomotor control, others, such as the dorsolateral prefrontal circuit, in cognitive functions, and others, such as the lateral orbitofrontal and anterior cingulate circuits, in social behavior, motivation, and emotion [39]. Each basal ganglia circuit originates in layer V of the cortex, and through a last step in the thalamus, returns to the cortex, not far from where the circuit started [40]. Similarly arranged cortico-cerebello-thalamo-cortical loops also exist. Our analysis suggests that such subcortical loops, even if connected at both ends to the thalamocortical system, would not change the composition of the main thalamocortical complex, and would only slightly modify its Φ value. Instead, the elements of the main complex and of the connected cycle form a joint complex that can only integrate the limited amount of

information exchanged within the loop. Thus, the present analysis suggests that, while subcortical cycles or loops could implement specialized subroutines capable of influencing the states of the main thalamocortical complex, the neural interactions within the loop would remain informationally insulated from those occurring in the main complex. To the extent that consciousness has to do with the ability to integrate information [18], such informationally insulated cortico-subcortical loops could constitute substrates for unconscious processes that can influence conscious experience [21,41]. It is likely that informationally encapsulated short loops also occur within the cerebral cortex and that they play a similar role.

Limitations and extensions

Most of the results presented in this paper were obtained on the basis of systems composed of a small number of elements, and further work will be required to determine whether they apply to larger systems. For example, it is not clear whether the highest values of Φ can always be obtained by optimizing connection patterns among all elements, or whether factors such as noise, available connection strength, and dynamic range (maximum available entropy) of individual elements would eventually force the largest complexes to break down. Similarly, it is not clear whether arbitrarily high values of information integration could be reached by generating connection patterns according to some probability distribution and merely increasing network size. Results obtained with small networks ($n \leq 12$) indicate that the performance of connection patterns generated according to a uniform or normal random distribution diverges strongly from that of optimal networks as soon as realistic constraints are introduced. For example, if we assume that each network can only maintain a relatively small number of connections, thus enforcing sparse connectivity, optimized networks develop specific wirings that yield high Φ ($\Phi = 60.76$, for $n = 8$ and $c_i = 0.00001$), while sparse random networks generated according to a uniform distribution have low Φ and tend to break down into small complexes ($\Phi = 33.90$). Similarly, if we assume that a certain amount of noise is unavoidable in biological systems ($c_i = 0.1$, $c_p = 1$), architectures yielding optimal Φ turn out to be sparse and are clearly superior to random patterns with connections drawn from uniform distributions ($\Phi = 5.71$ versus $\Phi = 2.92$).

The results presented in this initial analysis were obtained from linear systems having no external connections. Real systems such as brains, of course, are highly non-linear, and they are constantly interacting with the environment. Moreover, the ability of real systems to integrate information is heavily constrained by factors above and beyond their connective organization. In neural systems, for instance, factors affecting maximum firing rates, firing

duration, synaptic efficacy, and neural excitability, such as behavioral state, can radically alter information integration even if the anatomical connectivity is unchanged [21]. Such factors and their influence on Φ value of a complex need to be investigated in more realistic models [10,42,43]. Values of effective information and thus of Φ are also dependent on both temporal and spatial scales that determine the repertoire of states available to a system. The equilibrium linear systems considered here had predefined elementary units and no temporal evolution. In real systems, information integration can occur at multiple temporal and spatial scales. An interesting possibility is that the "grain size" in both time and space at which Φ/t reaches a maximum might define the optimal functioning range of the system [21]. With respect to time, Φ values in the brain are likely to show a maximum between tens and hundreds of milliseconds. It is clear, for example, that if one were to perturb one half of the brain and examine what effects this produces on the other half, no perturbation would produce any effect whatsoever after just a tenth of a millisecond ($EI = 0$). After say 100 milliseconds, however, there is enough time for differential effects to be manifested. Similarly, biological considerations suggest that synchronous firing of heavily interconnected groups of neurons sharing inputs and outputs, e.g. cortical minicolumns, may produce significant effects in the rest of the brain, while asynchronous firing of various combinations of individual neurons may not. Thus, Φ values may be higher when considering as elements cortical minicolumns rather than individual neurons, even if their number is lower.

A practical limitation of the present approach has to do with the combinatorial problem of exhaustively measuring Φ in large systems, as well as with the requirement to perturb a system in all possible ways, rather than just observe it for some time. In this respect, large-scale, realistic neural models offer an opportunity for examining the effects of a large number of perturbations in an efficient and economical manner [10,42]. In real systems, additional knowledge can often provide an informed guess as to which bipartitions may yield low values of EI . Moreover, as is the case in cluster analysis and other combinatorial problems, advanced optimization algorithms may be employed to circumvent an exhaustive search [44]. In this context, it should be noted that the bipartitions for which the normalized value of EI will be at a minimum will be most often those that cut the system in two halves, i.e. midpartitions [18]. Similarly, a representative rather than exhaustive number of perturbations may be sufficient to estimate the repertoire of different responses that are available to neural systems. Work along these lines has already progressed in the analysis of sensory systems in the context of information transmission [3,45]. It will be essential to extend this work and evaluate

the effects of perturbations applied directly to the central nervous system, as could be done for instance by combining transcranial magnetic stimulation with functional neuroimaging [46,47], or microstimulation with multi-electrode recordings. Finally, where direct perturbations are impractical or impossible, a partial characterization of the ability of a system to integrate information can still be obtained by exploiting endogenous variance or noise and substituting mutual information for effective information in order to identify complexes and estimate Φ (see examples at <http://www.indiana.edu/~cortex/complexity.html>).

Conclusions

Despite these limitations and caveats, the approach discussed here demonstrates how it is possible to assess which elements of a system are capable of integrating information, and how much of it. Such an assessment may be particularly relevant for neural systems, given that integrating information is arguably what brains are especially good at, and may have evolved for. More generally, the coexistence of functional specialization and integration is a hallmark of most complex systems, and for such systems the issue of information integration is paramount. Thus, a way of localizing and measuring the capacity to integrate information could be useful in characterizing many kinds of complex systems, including gene regulatory networks, ecological and social networks, computer architectures, and communication networks.

Methods

Model Implementation

In order to identify complexes and their $\Phi(S)$ for systems with many different connection patterns, we implemented numerous model systems X composed of n neural elements with connections CON_{ij} specified by a connection matrix $CON(X)$; self-connections CON_{ii} are generally excluded. In order to compare different architectures, $CON(X)$ was normalized so that the absolute value of the sum of the afferent synaptic weights per element corresponded to a constant value $w < 1$. If the system's dynamics corresponds to a multivariate Gaussian random process, its covariance matrix $COV(X)$ can be derived analytically. As in previous work [19], we consider the vector X of random variables that represents the activity of the elements of X , subject to independent Gaussian noise R of magnitude c . We have that, when the elements settle under stationary conditions, $X = X * CON(X) + cR$. By defining $Q = (1 - CON(X))^{-1}$ and averaging over the states produced by successive values of R , we obtain the covariance matrix $COV(X) = \langle X * X \rangle = \langle Q^t * R^t * R * Q \rangle = Q^t * Q$, where the superscript t refers to the transpose.

Mutual Information

Under Gaussian assumptions, all deviations from independence among the two complementary parts A and B of

a subset S of X are expressed by the covariances among the respective elements. Given these covariances, values for the individual entropies $H(A)$ and $H(B)$, as well as for the joint entropy of the subset $H(S) = H(AB)$ can be obtained. For example, $H(A) = (1/2) \ln [(2\pi e)^n |COV(A)|]$, where $|\bullet|$ denotes the determinant. The mutual information between A and B is then given by $MI(A:B) = H(A) + H(B) - H(AB)$ [2,20]. Note that $MI(A:B)$ is symmetric and positive.

Effective Information

To obtain the effective information between A and B within our model systems, we enforced independent noise sources in A by setting to zero strength the connections within A and afferent to A . Then the covariance matrix for A is equal to the identity matrix (given independent Gaussian noise), and any statistical dependence between A and B must be due to the *causal* effects of A on B , mediated by the efferent connections of A . Moreover, all possible outputs from A that *could* affect B are evaluated. Under these conditions, $EI(A \rightarrow B) = MI(A^{H_{max}}:B)$, as per equation (1). The independent Gaussian noise R applied to A is multiplied by c_p , the perturbation coefficient, while the independent Gaussian noise applied to the rest of the system is given by c_i , the intrinsic noise coefficient. By varying c_p and c_i one can vary the signal-to-noise ratio (SNR) of the effects of A on the rest of the system. In what follows, high SNR is obtained by setting $c_p = 1$ and $c_i = 0.00001$ and low SNR by setting $c_p = 1$ and $c_i = 0.1$. Note that, if A were to be stimulated, rather than substituted, by independent noise sources (by not setting to zero the connections within and to A), one would obtain a modified effective information (and derived measures) that would reflect the probability distribution of A 's outputs as filtered by its connectivity.

Complexes

To identify complexes and obtain their capacity for information integration, we consider every subset $S \subseteq X$, composed of k elements, with $k = 2, \dots, n$. For each subset S , we consider all bipartitions and calculate $EI(A \rightleftharpoons B)$ for each of them according to equation (2). We find the *minimum information bipartition* $MIB(S)$, the bipartition for which the normalized effective information reaches a minimum according to equation (3), and the corresponding value of $\Phi(S)$, as per equation (4). We then find the *complexes* of X as those subsets S with $\Phi > 0$ that are not included within a subset having higher Φ and rank them based on their $\Phi(S)$ value. As per equation (5), the complex with the maximum value of $\Phi(S)$ is the *main complex*.

Algorithms: Measuring Φ and finding complexes

Practically, the procedure for finding complexes can only be applied exhaustively to systems composed of up to two dozen elements because the number of subsets and bipar-

titions to be considered increases factorially. When dealing with larger systems, a random sample of subsets at each level (e.g. 10,000 samples per level) can provide an initial, non-exhaustive matrix of complexes. A search for complexes of maximum Φ can then be performed by appropriately permuting the subsets with the highest Φ (cf. [44] for optimization procedures in cluster analysis). In most cases, this procedure rapidly identifies the subsets with the highest Φ among millions of possibilities. MATLAB functions used for calculating effective information and exhaustive search for complexes are available for download at <http://www.indiana.edu/~cortex/complexity.html> or <http://tononi.psychiatry.wisc.edu/informationintegration/toolbox.html>.

Algorithms: Optimization

To search for connectivity patterns that are associated with high or optimal values of Φ , we employed two optimization strategies. When searching for connection patterns characterized by full CON(X) matrices with real-valued and continuous connection strengths CON_{ij} , we used a nonlinear constrained optimization algorithm with Φ as the cost function. When searching for sparse connection patterns characterized by binary connection matrices, we used an evolutionary search algorithm similar to an algorithm called "graph selection" described in earlier work [22]. Briefly, individual generations of varying connection matrices are generated and evaluated with Φ as the fitness function. Networks with the highest values for Φ are copied to the next generation and their connection matrices are randomly rewired to generate new variants. The procedure is continued until a maximal value for Φ is reached.

Authors' contributions

Both authors were involved in performing the simulations and analyzing the results. GT drafted the initial version of the manuscript, which was then revised and expanded by both authors. Both authors read and approved the final manuscript.

Acknowledgements

The authors were supported by US government contract NMA201-01-C-0034. The views, opinions, and findings contained in this report are those of the authors and should not be construed as an opinion, policy, or decision of NIMA or the US government.

References

- Shannon CE, Weaver W: *The mathematical theory of communication* Urbana: University of Illinois Press; 1949.
- Cover TM, Thomas JA: *Elements of information theory* New York: Wiley; 1991.
- Rieke F: *Spikes: exploring the neural code* Cambridge, Mass.: MIT Press; 1997.
- Dan Y, Atick JJ, Reid RC: **Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory.** *J Neurosci* 1996, **16**:3351-3362.
- Tononi G, Edelman GM, Sporns O: **Complexity and the integration of information in the brain.** *Trends Cognit Sci* 1998, **2**:44-52.
- James W: *The principles of psychology* New York: H. Holt; 1890.
- Sherrington CS: *The integrative action of the nervous system* New York: Yale Univ. Press; 1947.
- Mountcastle VB: *Perceptual neuroscience: the cerebral cortex* Cambridge, Mass.: Harvard University Press; 1998.
- Zeki S: *A vision of the brain* Oxford ; Boston: Blackwell Scientific Publications; 1993.
- Tononi G, Sporns O, Edelman GM: **Reentry and the problem of integrating multiple cortical areas: simulation of dynamic integration in the visual system.** *Cerebral Cortex* 1992, **2**:310-335.
- Tononi G, Edelman GM: **Consciousness and complexity.** *Science* 1998, **282**:1846-1851.
- Van Essen DC, Anderson CH, Felleman DJ: **Information processing in the primate visual system: an integrated systems perspective.** *Science* 1992, **255**:419-423.
- Munk MH, Nowak LG, Nelson JI, Bullier J: **Structural basis of cortical synchronization. II. Effects of cortical lesions.** *J Neurophysiol* 1995, **74**:2401-2414.
- Nowak LG, Munk MH, Nelson JI, James AC, Bullier J: **Structural basis of cortical synchronization. I. Three types of interhemispheric coupling.** *J Neurophysiol* 1995, **74**:2379-2400.
- Llinas R, Ribary U, Joliot M, Wang X-J: **Content and context in temporal thalamocortical binding.** In: *Temporal Coding in the Brain* Edited by: Buzsaki G, Llinas RR, Singer W. Berlin: Springer Verlag; 1994.
- Engel AK, Fries P, Singer W: **Dynamic predictions: oscillations and synchrony in top-down processing.** *Nat Rev Neurosci* 2001, **2**:704-716.
- Varela F, Lachaux JP, Rodriguez E, Martinerie J: **The brainweb: phase synchronization and large-scale integration.** *Nat Rev Neurosci* 2001, **2**:229-239.
- Tononi G: **Information measures for conscious experience.** *Arch Ital Biol* 2001, **139**:367-371.
- Tononi G, Sporns O, Edelman GM: **A measure for brain complexity: relating functional segregation and integration in the nervous system.** *Proceedings of the National Academy of Sciences of the United States of America* 1994, **91**:5033-5037.
- Jones DS: *Elementary information theory* New York: Oxford University Press; 1979.
- Tononi G: **Consciousness: Theoretical Aspects.** In: *Encyclopedia of Neuroscience* 3rd edition. Edited by: Adelman G, Smith BH. New York: Elsevier; 2003.
- Sporns O, Tononi G, Edelman GM: **Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices.** *Cereb Cortex* 2000, **10**:127-141.
- Hilgetag CC, Burns GA, O'Neill MA, Scannell JW, Young MP: **Anatomical connectivity defines the organization of clusters of cortical areas in the macaque monkey and the cat.** *Philos Trans R Soc Lond B Biol Sci* 2000, **355**:91-110.
- Sporns O, Tononi G: **Classes of network connectivity and dynamics.** *Complexity* 2002, **7**:28-38.
- Tononi G, McIntosh AR, Russell DP, Edelman GM: **Functional clustering: Identifying strongly interactive brain regions in neuroimaging data.** *Neuroimage* 1998, **7**:133-149.
- Horwitz B: **The elusive concept of brain connectivity.** *Neuroimage* 2003, **19**:466-470.
- Friston K: **Beyond phrenology: what can neuroimaging tell us about distributed circuitry?** *Annu Rev Neurosci* 2002, **25**:221-250.
- Passingham RE, Stephan KE, Kotter R: **The anatomical basis of functional localization in the cortex.** *Nat Rev Neurosci* 2002, **3**:606-616.
- Sporns O, Tononi G, Edelman GM: **Theoretical neuroanatomy and the connectivity of the cerebral cortex.** *Behav Brain Res* 2002, **135**:69-74.
- Edelman GM: **Neural Darwinism: selection and reentrant signaling in higher brain function.** *Neuron* 1993, **10**:115-125.
- McIntosh AR: **Towards a network theory of cognition.** *Neural Netw* 2000, **13**:861-870.
- McIntosh AR, Rajah MN, Lobaugh NJ: **Functional connectivity of the medial temporal lobe relates to learning and awareness.** *J Neurosci* 2003, **23**:6520-6528.
- Shiffrin RM, Schneider W: **Automatic and controlled processing revisited.** *Psychological Review* 1984, **Vol 91(92)**:269-276.
- Crick F, Koch C: **A framework for consciousness.** *Nat Neurosci* 2003, **6**:119-126.

35. Rees G, Kreiman G, Koch C: **Neural correlates of consciousness in humans.** *Nat Rev Neurosci* 2002, **3**:261-270.
36. Zeman A: **Consciousness.** *Brain* 2001, **124**:1263-1289.
37. Cohen D, Yarom Y: **Patches of synchronized activity in the cerebellar cortex evoked by mossy-fiber stimulation: questioning the role of parallel fibers.** *Proc Natl Acad Sci U S A* 1998, **95**:15032-15036.
38. Bower JM: **The organization of cerebellar cortical circuitry revisited: implications for function.** *Ann N Y Acad Sci* 2002, **978**:135-155.
39. Alexander GE, Crutcher MD, DeLong MR: **Basal ganglia-thalamo-cortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions.** *Prog Brain Res* 1990, **85**:119-146.
40. Middleton FA, Strick PL: **Basal ganglia and cerebellar loops: motor and cognitive circuits.** *Brain Res Brain Res Rev* 2000, **31**:236-250.
41. Baars BJ: *A cognitive theory of consciousness* New York, NY, US: Cambridge University Press; 1988.
42. Lumer ED, Edelman GM, Tononi G: **Neural dynamics in a model of the thalamocortical system .I. Layers, loops and the emergence of fast synchronous rhythms.** *Cerebral Cortex* 1997, **7**:207-227.
43. Tagamets MA, Horwitz B: **Integrating electrophysiological and anatomical experimental data to create a large-scale model that simulates a delayed match-to-sample human brain imaging study.** *Cereb Cortex* 1998, **8**:310-320.
44. Everitt B: *Cluster analysis* London, New York: E. Arnold, Halsted Press; 1993.
45. Buracas GT, Albright TD: **Gauging sensory representations in the brain.** *Trends Neurosci* 1999, **22**:303-309.
46. Pascual-Leone A, Walsh V, Rothwell J: **Transcranial magnetic stimulation in cognitive neuroscience – virtual lesion, chronometry, and functional connectivity.** *Curr Opin Neurobiol* 2000, **10**:232-237.
47. Paus T: **Imaging the brain before, during, and after transcranial magnetic stimulation.** *Neuropsychologia* 1999, **37**:219-224.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

