

Chapter 3

Generating Functionals for Guided Self-Organization

Claudius Gros

3.1 Controlling Complex Systems

One may take it as a running joke, that complex systems are complex since they are complex. It is however important to realize, this being said, that complex systems come in a large varieties, and in many complexity classes, ranging from relatively simple to extraordinary complex. One may distinguish in this context between *classical* and *modern* complex system theory. In the classical approach one would typically study a standardized model, like the Lorentz model or the logistic map, being described usually by maximally a handful of variables and parameters (Gros 2008). Many real-world systems are however characterized by a very large number of variables and control parameters, especially when it comes to biological and cognitive systems. It has been noted, in this context, that scientific progress may generically be dealing with complexity barriers of various severities, in far reaching areas like medicine and meteorology (Gros 2012b), when researching real-world natural or biological complex systems.

Generically, a complex system may be described by a set of first-order differential equations (or maps), like

$$\dot{x}_i = f_i(x_1, x_2, \dots | \gamma_1, \gamma_2, \dots), \quad (3.1)$$

where the $\{x_i\}$ are the primary dynamical variables and the $\{\gamma_j\}$ the set of control parameters. Modern complex system theory has often to deal with the situation where the phase space of dynamical variables and parameters are both high dimensional. Everything in the macroscopic world, f.i. the brain, can be described by an appropriate set of equations of motion, like (3.1), and we are hence confronted with two types of control problems:

Claudius Gros

Institute for Theoretical Physics, Goethe University Frankfurt, Germany

e-mail: gros@itp.uni-frankfurt.de

- How do we derive governing equations of type (3.1)?
- Given a set of equations of motion, like (3.1), how do we investigate its properties and understand the resulting behavior as a function of the control parameters?

At its core, we are interested here in how to generically control, in general terms, a complex and self-organizing system. A range of complementary approaches are commonly used in order to alleviate the control problem, we discuss here some of the most prominent (non mutually exclusive) approaches.

- *Delegation to Evolution*

One is often interested, especially in biology and in the neurosciences, in biologically realistic models and simulations (Markram 2006). In this case both the functional form and the parametric dependences are taken from experiment. One may then expect, thanks to Darwinian selection, that the such constructed dynamical equation should exhibit meaningful behavior, replicating observations.

- *Exploring Phase Space*

A complete understanding would correspond, within dynamical system theory, to a full control of both the qualitative behavior of the flow in phase space and of its dependency on the control parameters. Achieving this kind of complete control is clearly very desirable, but often extremely hard to achieve when dealing with large numbers of dynamical variables and control parameters, the typical situation in modern complex system theory. The exploration of phase space, typically through a combination of analytical and numerical investigations, is in any case an indispensable tool, even when only a small fraction of the overall phase space volume can be probed.

- *Classical Control Theory*

Classical control theory deals with the objective to control a real-world system, like a rocket, such that a desired behavior is optimally achieved, in the wake of noise both in the sensor readings and in the action effectiveness (Leigh 2004). Classical control theory is of widespread use in engineering and for robot control (De Wit et al. 1996). Our present discussion deals however with the general control of working regimens of a self-organizing complex system; if we knew what the system is supposed to do, we would be done.

- *Diffusive Control*

Neuromodulators (Marder 2012), like dopamine, serotonin, choline, norepinephrine, neuropeptides and neurohormons, act in the brain as messengers of a diffusive control system (Gros 2010, 2012a), controlling intrinsic and synaptic properties like neural gain and threshold, or synaptic plasticity. Diffusive control is needed to stabilize a desired working regime, a process also denoted as metalearning (Doya 2002), and to switch between different working regimes in order to achieve behavioral flexibility (Arnsten et al. 2012). Diffusive control is a very general strategy for controlling a complex system.

- *Generating Functionals*

This is the subject we will develop here. One can achieve an improved understanding when considering classes of dynamical systems derived from superordinated functionals. In this case the equations of motion are not given a priori, but

derived from a generating principle. Here we will detail out how this approach leads to an alleviation of the control problem.

One needs to recall, coming back to the introductory statement, that there is no one-size-fits-all method for controlling complex systems (Frei and Serugendo 2011), as there are many kinds and varieties of complex systems. Here we will consider primarily systems made up of a potentially large number of similar functional units, as typical for neural networks. A related aspect of the generic control problem discussed above regards, in this context, the stability of a default working regime with regard to external influences and statistical fluctuations (Clarke 2007). This is particularly important in functional complex systems, such as an ecosystem (Holling 1973; May 2001), or cognitive architectures, the subject of our interest here.

3.2 Guiding Self-Organization

There is no strict scientific definition of what self-organization means or implies. It is however generally accepted to consider processes as self-organizing when a rich and structured dynamics results from a set of relatively simple evolution rules. The term self-organization is of widespread use (Haken 2006), ranging from classical non-equilibrium physical (Nicolis 1989) and biological (Camazine 2003) systems to the assembly of complex macromolecules (Lehn 2002); it is quite generally accepted that the foundations of life are based on self-organizing principles (Kauffman 1993). The brain in particular, possibly the most complex object presently known to humanity, is expected to result from a plethora of intertwined self-organizing processes (Kelso 1995), ranging from self-organized cognitive functions (Kohonen 1988) to self-organized critical states (Bak 1999; Chialvo 2010).

Self-organization is, per se, content free, having no semantic relevancy. The stars in a rotating galaxy, to give an example, may spontaneously organize into a set of distinct density waves, known as the arms of a spiral galaxy. Even though pretty to the eye, the spiral arms of the Milky Way do not serve any purpose; self-organization is in this case just a byproduct of Newton's law. The situation is however generically distinct for biological settings, or for man-made systems, where functionality is the key objective.

The design of functionality is of course a standard objective for the vast majority of man-made systems, and contrasts with the absence of functionality of natural phenomena. Here we are interested in self-organizing processes which are neither fully designed nor without any objective. There is a middleway, which has been denoted "*targeted self-organization*" (Gros 2008) or, alternatively, "*guided self-organization*" (Prokopenko 2009; Martius and Herrmann 2010).



For a designed system the functionality is specified in detail in order to achieve optimal performance for a given task. The target for a self-organizing process is however presumed to be a generic principle, often based on information theoretical

considerations, with the actual functionality arising indirectly through self-organizing processes. Targeted and guided self-organization are essentially identical terms, with guided self-organization having a somewhat broader breath. One could guide, for example, a dynamical system by restricting its flow to a certain region in phase space, allowing for an otherwise unrestricted development within this bounded area of phase space. Here we will neglect the differences in connotation between targeted and guided self-organization and use both terms interchangeably.

Let us come back at this point to the general formulation of a complex dynamical system through a set of parameterized first-order differential equations, as given by (3.1). The distinction between a parameter γ_j and a primary dynamical variable $x_k(t)$ is a question of time scales.

$$\boxed{\begin{array}{l} \dot{x}_k : \text{fast} \\ \dot{\gamma}_k : \text{slow} \end{array} \left. \vphantom{\begin{array}{l} \dot{x}_k \\ \dot{\gamma}_k \end{array}} \right\} \text{time evolution}}$$

The flow $(x_1(t), x_2(t), \dots)$ of the primary dynamics is taking place in the slowly changing environment of parameter space, defining the adiabatic background. The slow adaption of parameters is what controls in the end the working regime of a dynamical system, and is also denoted sometimes as metalearning (Vilalta and Drissi 2002). Not all parameters can be involved in metalearning, a small but finite set of core parameters $\{\gamma_j\} \in \{\gamma_k\}$ must be constant and immutable,

$$\dot{\gamma}_j = 0.$$

This set of core parameters is what defines in the end the system. One has achieved a dimensional reduction of the control problem if the number $|\{\gamma_j\}|$ of core parameters is small. This is the aim of guided self-organization, that a concise set of core parameters controls the development and the dynamical properties of a system, with quantitative tuning of the values of the control parameters inducing modifications of the system's characteristics, both on a quantitative and a qualitative level.

3.3 Generating Functionals

There are two principle venues on how to express guiding principles for dynamical systems, implicitly or explicitly. In analogy, one can implement conservation laws in physics by writing down directly appropriate equations of motion, demonstrating that, e.g., energy is conserved. In this case energy conservation is implicitly present in the formulation of the dynamical system. Alternatively one may consider directly a time independent Lagrange function, a condition which explicitly guarantees energy conservation for the respective Lagrange equations of motion. Here we will concentrate on the second approach, the explicit derivation of equations of motion for targeted self-organization through appropriate generating functionals.

The term generating functional has a wide range of connotations in the sciences. The action functional in classical mechanics and quantum field theory is a prominent example from physics, the generating functional $\sum_k p_k x^k$ for a distribution

function p_k (with $p_k \geq 0$ and $\sum_k p_k = 1$) another from information theory. In the neurosciences it is custom to speak of objective functions (Intrator and Cooper 1992; Goodhill and Sejnowski 1997) instead of generating functionals.

As a first example we consider a simple energy functional

$$E(\{x_k\}) = \frac{\Gamma}{2} \sum_k x_k^2 - \frac{1}{2} \sum_{kl} y(x_k) w_{kl} y(x_l), \quad (3.2)$$

which is suitable for a network of neurons with membrane potential x_k and firing rate $y(x_k)$. Here $y(x)$ is the sigmoidal transfer function

$$y(x) = \frac{1}{1 + e^{a(b-x)}}, \quad (3.3)$$

parameterized by the gain a and the threshold b . The w_{ij} in (3.2) will turn into the synaptic weights, as we will show later on, and Γ into a relaxation rate. Concerning the terminology, one could consider $E(\{x_k\})$ also to be an energy function (instead of a functional), being a function of the individual x_k . Here we use the term energy functional, for the functional dependence on the vector $\mathbf{x} = (x_1, x_2, \dots)$ of membrane potentials.

For our second example we consider a general functional based on the principle of polyhomeostasis (Marković and Gros 2010). One speaks of a homeostatic feedback loop when a target value for a single scalar quantity is to be achieved. Life per se is based on homeostasis, the concentrations of a plethora of biological relevant substances, minerals and hormones need to be regulated, together with a vast number of physical properties, like the body temperature or the heart beating frequency. Polyhomeostasis is, in contrast, typically necessary for time allocation problems.

The problem of allocating time for various tasks constitutes the foundation of behavior. Every living being needs to decide how much time to spend, relatively, on vitally important behaviors, like foraging, resting, exploring or socializing. Maximizing only a single of the possible behavioral patterns would be counterproductive, only a suitable mix of behaviors, as an average over time, is optimal. Mathematically this goal is equivalent to optimizing a distribution function, hence the term polyhomeostasis, in contrast to the case of homeostasis, corresponding to the optimization of a single scalar quantity.

All a neuron can do, at any given moment, is to fire or not to fire, a typical time allocation problem. The generic functional

$$F[p] = \int p(y) f(p(y)) dy \quad (3.4)$$

of the firing rate distribution

$$p(y) = \frac{1}{T} \int_0^T \delta(y - y(t - \tau)) d\tau \quad (3.5)$$

is an example of the polyhomeostatic principle. Minimizing $F[p]$ corresponds to optimizing a given function $f(p)$ of the neural activity distribution $p(y)$. The resulting adaption rates will then influence the timeline $y(t)$ of the neural activity. This is an example of guided self-organization, since the target functional is expressed in terms of general statistical properties of the dynamical flow, independently of an eventual semantic content. The explicit form and derivation of the adaption rates will be discussed further below, both for the polyhomeostatic functional (3.4) and for the energy functional (3.2).

3.4 Equations of Motion

There are several venues for deriving equations of motions from a given target functional. One uses variational calculus, within classical mechanics, when deriving the Lagrange equations of motion. In classical mechanics the target functional, the action, needs to be stationary with respect to an arbitrary variation of the trajectory. Here we will consider instead generic objective functions which are to be minimized.

Minimizing an objective function is a very generic task for which a wide range of methods and algorithms have been developed (Papadimitriou and Steiglitz 1998; Goldberg 1989; Kennedy and Eberhart 1995). Here we are however interested in a different aspect. Our aim is not to actually find the global minimum of a given objective function, or any stationary point, which is not of interest per se. Objective functions serve as a guiding principle and equations of motion induced by minimizing a given objective function will tend to minimize it. Other driving influence will however in general compete with this goal and it is this very competition which may result in complex and novel dynamical states.

For an objective function which is an explicit function of the dynamical variable, like the energy functional (3.2), the equations of motion just correspond to the downhill flow within the energy landscape,

$$\dot{x}_j = -\frac{1}{T_e} \frac{\partial}{\partial x_j} E(\{x_k\}), \quad (3.6)$$

where the timescale T_e of the flow is normally set to unity, $T_e = 1$. In our case we obtain

$$\dot{x}_k = -\Gamma x_k + a_k y_k (1 - y_k) \sum_j w_{kj} y_j, \quad (3.7)$$

where we have used (3.3) and

$$y'(x) = \frac{\partial y}{\partial x} = ay(1 - y). \quad (3.8)$$

The dynamical system (3.7) just corresponds to a network of leaky integrators (Hopfield 1982, 1984), with the x_k and y_k corresponding to the membrane potential and the mean neural firing rate respectively. The neurons are coupled through the weight

matrix w_{kj} , the synaptic weights. The term $a_k y_k (1 - y_k)$ in front of the inter-neural coupling is present only when deriving (3.7) from the energy functional (Linkerhand and Gros 2012a), and not when formulating equivalent neural updating rules directly from neurobiological considerations (Olshausen et al. 1993).

The polyhomeostatic functional (3.4) is used to derive adaption rules for the intrinsic parameters a_i and b_i of the transfer function (3.3). The lack of an explicit dependence on either a_i or b_i rules out adaption rules like $\dot{a}_i \propto -\partial F[p]/\partial a_i$, which would be the equivalent to (3.6). It is however possible to derive implicit adaption rules, for which the minimization of the objective functions $F[p]$ is performed stochastically in the sense that the time-averaged firing rate $p(y)$ is sampled along the flow during the time evolution. For this purpose we change variables and rewrite the generating functional

$$F[p] = \int p(x) f(p(y)/y') dx, \quad p(y) dy = p(x) dx \quad (3.9)$$

as an expectation value over the distribution $p(x)$ of the membrane potential x , the input. The transfer function (3.3) maps the input of a neuron to its output and adaption rules for the intrinsic parameters should hence not depend explicitly on the actual distribution $p(x)$ of the input, they should be universal in the sense that the intrinsic adaption rules should abstract from the actual semantic content of the information being processed. Noting that $p(x)$ does not depend explicitly on the gain a and the threshold b , we have

$$\frac{\partial}{\partial \theta} F[p] = \int dx p(x) \frac{\partial}{\partial \theta} f(p(y)/y'), \quad \theta = a, b. \quad (3.10)$$

For the overall global minimum of $F[p]$ the weighting with respect to the input distribution $p(x)$ would be needed to be taken into account. As we are however interested only in adaption rules abstracting from the actual form of the input distribution, and noting that $p(x) \geq 0$ is positive definite, we demand that the adaption process should lead to a uniform minimization of the kernel of (3.10),

$$\dot{\theta} = -\varepsilon_\theta \frac{\partial}{\partial \theta} f(p(y)/y'), \quad \theta = a, b, \quad (3.11)$$

where ε_θ are the respective adaption rates. The adaption process should generally be slow, as typical for metalearning, and the adaption rates ε_a and ε_b small. In this case the updating rule (3.11) will statistically sample the input distribution $p(x)$, as an average over time, and become equivalent with (3.10).

The adaption rates (3.11) are generic and need to be concretized for a specific polyhomeostatic function $f(p)$. A straightforward target functional for the problem of allocating time is to consider a target distribution function $q(y)$ for the neural firing rate. In this case the functional

$$F[p] = \int p(y) f(p(y)) dy, \quad f(p) = \ln(p/q) \quad (3.12)$$

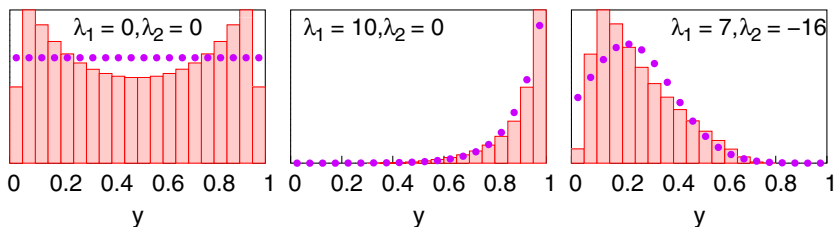


Fig. 3.1 The results of the intrinsic adaption rules (3.14) and (3.15) for the time averaged firing rate distribution (boxes, see Eq. (3.5)) of a single neuron driven by a white-noise input and for several information maximizing target distributions (points, see Eq. (3.13))

corresponds to the Kullback-Leibler divergence (Gros 2008), which is a positive definite measure for the similarity of two distribution functions p and q . The Kullback-Leibler divergence is minimal whenever $p(y)$ and $q(y)$ are as similar as possible, within the configuration of all dynamically realizable firing rate distributions $p(y)$.

The target firing rate distribution $q(y)$ could be any positive and normalized distribution function. Here we demand that $q(y)$ should maximize Shannon's information entropy $-q \ln(q)$, which can be achieved using variational calculus:

$$0 = -\delta \int q [\ln(q) - \lambda_1 y - \lambda_2 y^2] dy, \quad q(y) \propto e^{\lambda_1 y + \lambda_2 y^2}. \quad (3.13)$$

Here λ_1/λ_2 are suitable Lagrange parameters enforcing a given mean/variance. The flat distribution $\lambda_1 = \lambda_2 = 0$ maximizes information entropy in the absence of any constraint. Using (3.11) and $y' = ay(1-y)$, see Eq. (3.8), we obtain then the adaption rules (Triesch 2005, 2007; Marković and Gros 2010; Linkerhand and Gros 2012b)

$$\dot{a} = \varepsilon_a \left(\frac{1}{a} + (x-b) \Delta \tilde{\theta} \right) \quad (3.14)$$

$$\dot{b} = \varepsilon_b (-a) \Delta \tilde{\theta}, \quad \Delta \tilde{\theta} = (1-2y) + y(1-y) [\lambda_1 + 2\lambda_2 y]. \quad (3.15)$$

In Fig. 3.1 we present the results for a single polyhomeostatically adapting neuron, driven by white noise, for various target distributions $q(y)$. Note that there are only two intrinsic parameters, the threshold b and the gain a , to be optimized and that the transfer function (3.3) can hence not change, during the adaption process, its functional form arbitrarily. The firing rate distribution $p(y)$ approximates, considering this limitation, the target distribution $q(y)$ remarkably well, an exemplification of the principle of targeted self-organization.

3.5 Adaptive Phase Space

It is illuminating to investigate somewhat in detail the behavior of the adaption process in the phase space (a, b) of the intrinsic adaption parameters, and to study individual trajectories $(a(t), b(t))$. In Fig. 3.2 we present a selection of trajectories

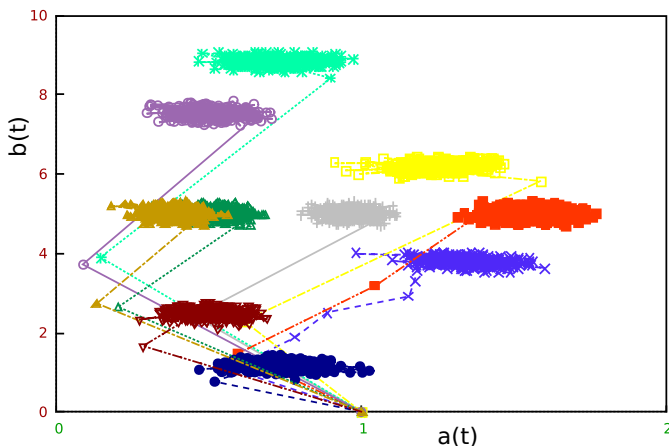


Fig. 3.2 Sample trajectories $(a(t), b(t))$ resulting from the intrinsic adaption rules (3.14) and (3.15), color coded for various parameters λ_1 and λ_2 of the target distributions $q(t)$, compare Eq. (3.13). All trajectories start at $(a(0), b(0)) = (1, 0)$ and then settle into distinct regions of phase space, where they perform a confined stochastic walk, due to the white-noise input.

for distinct realizations of the target distribution $q(y)$, as given by Eq. (3.13). The neuron is driven by a white noise input, the starting gain and threshold are $a = 1$ and $b = 0$, for all trajectories. After a relatively fast initial transient the intrinsic parameters settle to distinct respective regions in the phase space, where they perform a stochastic motion, reflecting the white-noise character of the driving input. Three of the resulting firing rate distributions $p(y)$ are shown in Fig. 3.1.

The target distribution $q(y)$, see Eq. (3.13), can be selected to be bimodal, which is generally the case for inverse Gaussians having $\lambda_1 < 0$ and $\lambda_2 > 0$. In Fig. 3.3 we present the adaptive walk through phase space $(a(t), b(t))$ for a bimodal target distribution having $\lambda_1 = -20$ and $\lambda_2 = 18.5$ and for various adaption rates $\varepsilon_a = \varepsilon_b$. When the adaption process is very slow, viz for small ε_a and ε_b the system average over extended periods of the stochastic input and the dynamics becomes smooth (Linkerhand and Gros 2012b), fluctuating with a reduced amplitude around a certain target region in phase space, just as illustrated in Fig. 3.2.

For a bimodal target distribution $q(y)$ there may however be two local minima in adaptive space, since the transfer function (3.3) is always monotonic. For any given pair of intrinsic parameters the system can hence approximate well only one of the two peaks of a bimodal transfer function. For small adaption it remains trapped in one of the local minima, but larger adaption rates ε_a and ε_b will lead to an enhanced sensibility with respect to the stochastic driving, inducing stochastic tipping transitions between the two local minima. This is a striking realization of the principle of guided self-organization.

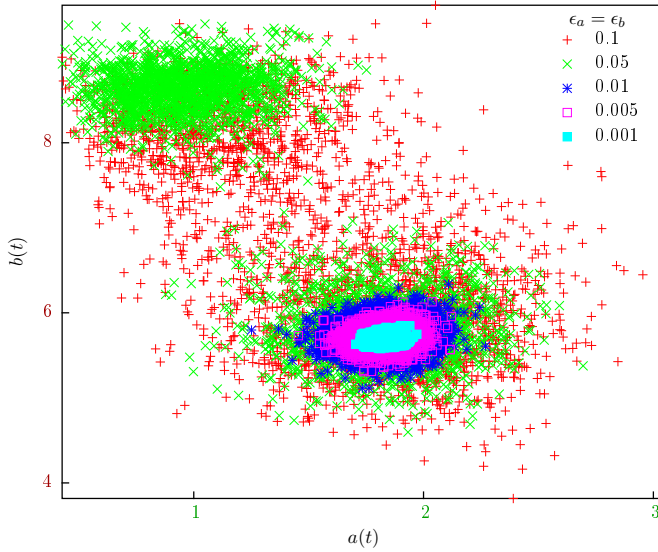


Fig. 3.3 Sample trajectories $(a(t), b(t))$ resulting from the intrinsic adaption rules (3.14) and (3.15), color coded for various adaption parameters $\epsilon_a = \epsilon_b$, as given in the legend. The single neuron is driven by white noise and the target distribution, see Eq. (3.13) is bimodal, parameterized by $\lambda_1 = -20$ and $\lambda_2 = 18.5$. For moderate large adaption rates the system is able to make stochastically driven tipping transitions between two local minima (Linkerhand and Gros 2012b).

3.6 Self-Organized Dynamical States

As a second example for the functioning of polyhomeostatic optimization we consider a network of N randomly interconnected neurons,

$$x_k = \sum_{j \neq k} w_{kj} y_j$$

which corresponds to (3.7) in the anti-adiabatic limit $\Gamma \rightarrow \infty$ (and without the factor $y(1-y)$). For the synaptic weights we select

$$w_{ij} = \begin{cases} +1/\sqrt{K} & \text{with probability } \rho_{exc} \\ -1/\sqrt{K} & \text{with probability } 1 - \rho_{exc} \end{cases}, \quad (3.16)$$

where K is the in-degree. The system is balanced for $\rho_{exc} = 1/2$. As a second control parameter, besides the fraction ρ_{exc} of excitatory links, we consider the average target activity μ ,

$$\mu = \int y q(y) dy, \quad \int q(y) dy = 1, \quad (3.17)$$

which is taken to be uniform, viz identical for all sites.

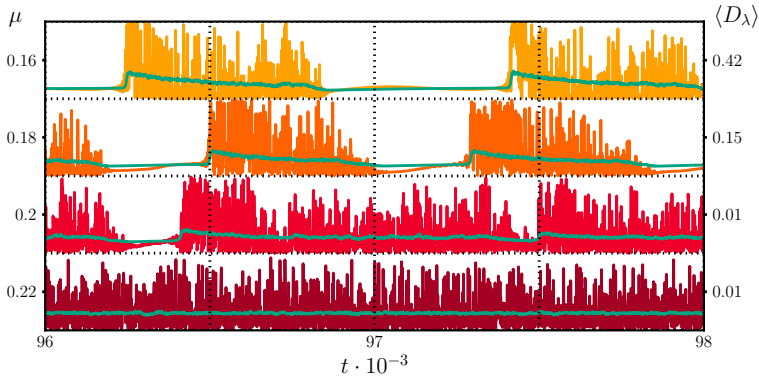


Fig. 3.4 For a network of $N = 1000$ adapting neurons, according to Eqs. (3.14) and (3.15), the activity of a randomly selected neuron and the average neural activity (green line). The network is balanced, with as many excitatory and inhibitory links, randomly selected according to Eq. (3.16). Shown are results for various target mean activities μ , see Eq. (3.17). The right-hand axis is not a scale, the numbers are the values of the network-averaged Kullback-Leibler divergence $\langle D_\lambda \rangle$, as defined by Eq. (3.12). One observes that the mean target activity μ entering the polyhomeostatic generating functional acts as a parameter controlling the resulting self-organized dynamical state (Marković and Gros 2012).

In Fig. 3.4 we present the results for a balanced network with $N = 1000$ adapting neurons, and an in-degree of $K = 100$. Shown are both the activity of a single, randomly selected site and the average activity, averaged over all sites. We notice that the network enters distinct dynamical states, as a function of the mean target activity μ (Marković and Gros 2010, 2012). For intermediate target activity levels the dynamics is chaotic, for smaller mean activities μ a regime with intermittent bursts is observed. One has hence the possibility to tune the self-organized dynamical state through the target set by the polyhomeostatic generating functional, an example of targeted self-organization. Interestingly the overall value of the network-averaged Kullback-Leibler divergence is minimal in the chaotic state.

In Fig. 3.5 we present the results for the same network of $N = 1000$ sites as in Fig. 3.4, but this time the network is not balanced, $\rho_{exc} > 1/2$. The mean target firing-rate activity is kept constant at $\mu = 0.3$. For larger values of ρ_{exc} the network synchronizes, not surprisingly, as a result of the predominance of positive feedback loops. For values of ρ_{exc} close to the balanced state, the system is chaotic, with a large variability around a partly synchronized state in between. One can regard ρ_{exc} as a controlling parameter of the energy functional (3.2), which hence allows to guide the self-organization of the resulting dynamical state. The value of the Kullback-Leibler divergence is, again, lowest in the chaotic state, which explores phase space best.

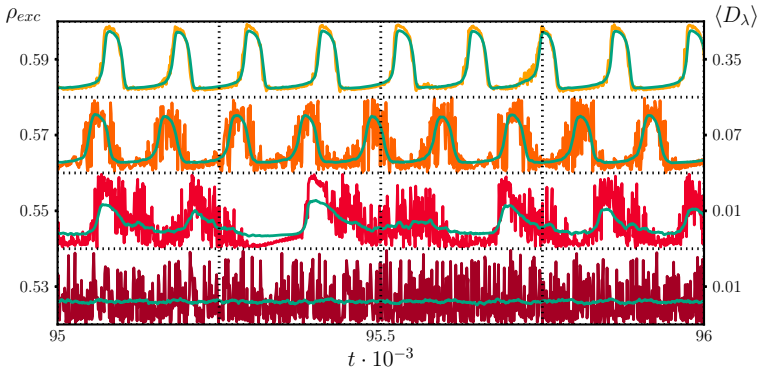


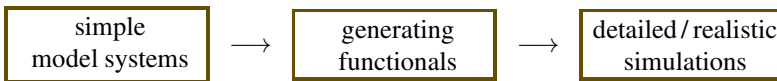
Fig. 3.5 For networks containing $N = 1000$ adapting neurons with an in-degree $K = 100$ and a target mean activity $\mu = 0.3$, see Eq. (3.17), the activity of a randomly selected neuron and the average neural activity (green line). The networks are not balanced, having a slight excess ρ_{exc} of randomly selected excitatory links see Eq. (3.16). Also given (on the right) are the respective values of the network-averaged Kullback-Leibler divergence $\langle D_\lambda \rangle$, as defined by Eq. (3.12). The network shows a transition between chaos and synchronization, as a function of ρ_{exc} (Marković and Gros 2012).

3.7 Discussion

A self-organizing process may be guided by presenting to the system one or more targets. If these targets are very concrete they may destroy the self-organizing character of the process, resulting in a driving force. One possibility to achieve a gentle way of controlling a self-organizing process is to formulate the targets in terms of statistical properties of the desired dynamical state, with a basic example being the time-average distribution function of activities. Optimizing the distribution of activities is an example of a time-allocation problem, which is intrinsically of polyhomeostatic nature.

A given set of goals may be achieved by a range of different tools, for example using evolutionary algorithms. In this treatise we have discussed the perspective, together with concrete examples, of explicitly deriving equations of motions from generating functionals incorporating polyhomeostatic and other targets. We believe that this approach offers several advantages. Having explicit time evolution equation at hand is, in our view, mandatory for time-efficient simulations and applications. Generating functionals can furthermore be seen as a route for solving the control problem, as they offer a substantial dimensional reduction in the number of free parameters. This is a particularly attractive feature, in view of the raising appreciation that the neuromodulator control system in the brain tunes the relative stability of a wide range of possible dynamical operative states of the affected downstream circuits.

From an alternative perspective one may view generating functionals also as a midway between the study of simplified model systems and biological realistic simulations.



Model systems may constitute important reference models, for understanding and developing key concepts and methods. Detailed simulations are, at the other extreme, often indispensable for obtaining a realistic comparison with experimental data, having however the drawback that an in-depth understanding is in general not achievable. We propose generating functionals as a venue for building increasingly complex dynamical systems and cognitive architectures, a venue which allows for the control of the operating modi of the system by tuning a limited number of high-level control parameters incorporating the targets of the respective generating functionals.

References

- Arnsten, A., Wang, M., Paspalas, C.: Neuromodulation of thought: flexibilities and vulnerabilities in prefrontal cortical network synapses. *Neuron* 76, 223–239 (2012)
- Bak, P.: *How nature works: the science of self-organized criticality*. Springer (1999)
- Camazine, S., Deneubourg, J., Franks, N., Sneyd, J., Theraula, G., Bonabeau, E.: *Self-organization in biological systems*. Princeton University Press (2003)
- Chialvo, D.: Emergent complex neural dynamics. *Nature Physics* 6(10), 744–750 (2010)
- Clarke, B.: *Stability of complex reaction networks*. Wiley (2007)
- De Wit, C., Bastin, G., Siciliano, B.: *Theory of robot control*. Springer, New York (1996)
- Doya, K.: Metalearning and neuromodulation. *Neural Networks* 15, 495–506 (2002)
- Frei, R., Serugendo, G.D.M.: Advances in complexity engineering. *International Journal of Bio-Inspired Computation* 3(4), 199–212 (2011)
- Goldberg, D.: *Genetic algorithms in search, optimization, and machine learning* (1989)
- Goodhill, G., Sejnowski, T.: A unifying objective function for topographic mappings. *Neural Computation* 9(6), 1291–1303 (1997)
- Gros, C.: *Complex and adaptive dynamical systems: A primer*. Springer (2008)
- Gros, C.: Cognition and emotion: perspectives of a closing gap. *Cognitive Computation* 2, 78–85 (2010)
- Gros, C.: Emotional control—conditio sine qua non for advanced artificial intelligences? In: Müller, V. (ed.) *Philosophy and Theory of Artificial Intelligence*. Springer (2012a)
- Gros, C.: Pushing the complexity barrier: diminishing returns in the sciences. *Complex Systems* 21, 183 (2012b)
- Haken, H.: *Information and self-organization: A macroscopic approach to complex systems*, vol. 40. Springer (2006)
- Holling, C.: Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 1–23 (1973)
- Hopfield, J.: Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences* 79(8), 2554–2558 (1982)
- Hopfield, J.: Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences* 81(10), 3088–3092 (1984)
- Intrator, N., Cooper, L.: Objective function formulation of the bcm theory of visual cortical plasticity: Statistical connections, stability conditions. *Neural Networks* 5(1), 3–17 (1992)

- Kauffman, S.: The origins of order: Self-organization and selection in evolution. Oxford University Press (1993)
- Kelso, J.: Dynamic patterns: The self-organization of brain and behavior. MIT press (1995)
- Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Proceedings of the IEEE International Conference on Neural Networks, vol. 4, pp. 1942–1948. IEEE (1995)
- Kohonen, T.: Self-organization and associative memory. Springer (1988)
- Lehn, J.: Toward self-organization and complex matter. *Science* 295(5564), 2400–2403 (2002)
- Leigh, J.: Control theory. Peter Peregrinus Limited (2004)
- Linkerhand, M., Gros, C.: Generating functionals for autonomous latching dynamics in attractor relict networks. arXiv preprint arXiv:1212.5054 (2012a)
- Linkerhand, M., Gros, C.: Self-organized stochastic tipping in slow-fast dynamical systems. arXiv preprint arXiv:1207.2928 (2012b)
- Marder, E.: Neuromodulation of neuronal circuits: Back to the future. *Neuron* 76, 1–11 (2012)
- Marković, D., Gros, C.: Self-organized chaos through polyhomeostatic optimization. *Physical Review Letters* 105, 68702 (2010)
- Marković, D., Gros, C.: Intrinsic adaptation in autonomous recurrent neural networks. *Neural Computation* 24(2), 523–540 (2012)
- Markram, H.: The blue brain project. *Nature Reviews Neuroscience* 7, 153–160 (2006)
- Martius, G., Herrmann, J.M.: Taming the beast: Guided self-organization of behavior in autonomous robots. In: Doncieux, S., Girard, B., Guillot, A., Hallam, J., Meyer, J.-A., Mouret, J.-B. (eds.) SAB 2010. LNCS, vol. 6226, pp. 50–61. Springer, Heidelberg (2010)
- May, R.: Stability and complexity in model ecosystems. Princeton University Press (2001)
- Nicolis, G.: Physics of far-from-equilibrium systems and self-organisation. *The New Physics* 11, 316–347 (1989)
- Olshausen, B., Anderson, C., Van Essen, D.: A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *The Journal of Neuroscience* 13(11), 4700–4719 (1993)
- Papadimitriou, C., Steiglitz, K.: Combinatorial optimization: algorithms and complexity. Dover Publications (1998)
- Prokopenko, M.: Guided self-organization. *HFSP Journal* 3, 287–289 (2009)
- Triesch, J.: A gradient rule for the plasticity of a neuron's intrinsic excitability. In: Duch, W., Kacprzyk, J., Oja, E., Zadrożny, S. (eds.) ICANN 2005. LNCS, vol. 3696, pp. 65–70. Springer, Heidelberg (2005)
- Triesch, J.: Synergies between intrinsic and synaptic plasticity mechanisms. *Neural Computation* 19(4), 885–909 (2007)
- Vilalta, R., Drissi, Y.: A perspective view and survey of meta-learning. *Artificial Intelligence Review* 18(2), 77–95 (2002)