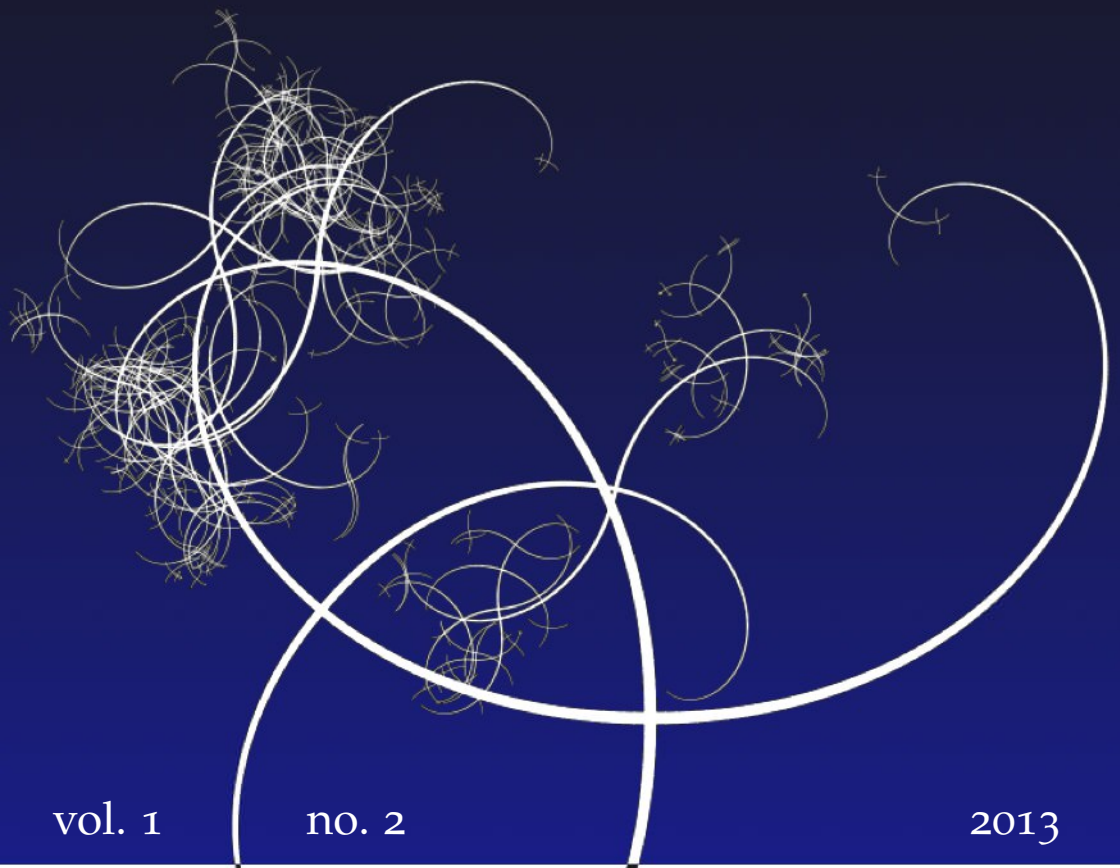
NISSUNA UMANA INVESTIGAZIONE SI PUO DIMANDARE VERA SCIENZIA
S'ESSA NON PASSA PER LE MATEMATICHE DIMOSTRAZIONI
LEONARDO DA VINCI



vol. 1 no. 2 2013

Mathematics and Mechanics *of* Complex Systems

Mathias Linkerhand and Claudius Gros

SELF-ORGANIZED STOCHASTIC TIPPING IN SLOW-FAST DYNAMICAL SYSTEMS

msp

MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS
Vol. 1, No. 2, 2013
dx.doi.org/10.2140/memocs.2013.1.129

SELF-ORGANIZED STOCHASTIC TIPPING IN SLOW-FAST DYNAMICAL SYSTEMS

MATHIAS LINKERHAND AND CLAUDIUS GROS

Polyhomeostatic adaption occurs when evolving systems try to achieve a target distribution function for certain dynamical parameters, a generalization of the notion of homeostasis. Here we consider a single rate-encoding leaky integrator neuron model driven by white noise, adapting slowly its internal parameters, threshold and gain, in order to achieve a given target distribution for its time-averaged firing rate. For the case of sparse encoding, when the target firing-rate distribution is bimodal, we observe the occurrence of spontaneous quasiperiodic adaptive oscillations resulting from fast transition between two quasistationary attractors. We interpret this behavior as self-organized stochastic tipping, with noise driving the escape from the quasistationary attractors.

1. Introduction

Self-regulation plays an important role in biological and technical systems. Homeostatically regulated steady states are a precondition to life, examples being the concentration of blood glucose controlled by insulin [Plum et al. 2006] and glucagon, the pH value of blood [Schaefer 1961; Tresguerres et al. 2010], and body temperature [Charkoudian 2003], which are all autoregulated in order to maintain stable conditions. Further examples are the concentration of ions, proteins, and transmitters in the brain; their respective levels are all self regulated [Marder and Goaillard 2006]. Furthermore, homeostasis is implemented and can be found in technical systems, for example in microrobotic swarms [Kernbach and Kernbach 2011]. Adaption typically introduces a slow time scale into the dynamical system [Gros 2010b], a process also called metalearning, a central notion in the context of neuromodulation [Doya 2002] and emotional control [Gros 2010a]. The resulting dynamical system then has both fast and slow variables and critical transitions in the form of tipping processes may occur [Kuehn 2011].

Classical homeostasis involves the regulation of a scalar quantity such as body temperature. More complex forms of homeostasis are also important in the realm

PACS2010: 05.10.Gg, 05.40.Ca, 05.45.-a, 05.45.Tp, 05.65.+b.
Keywords: stochastic tipping, complex systems, chaos, intrinsic adaption, slow-fast, metalearning.

of life. For example, an animal may want to achieve a certain time-averaged distribution of behaviors, like foraging, resting, and engaging socially, over the course of several days. This kind of adaptive behavior has been termed polyhomeostasis [Marković and Gros 2010; 2012]. It occurs when a dynamical system tries to achieve, via the continuous adaption of slow variables, a given target distribution for the time-averaged activity of a subset of fast variables. Polyhomeostatically adapting systems are typically slow-fast dynamical systems and their dynamical behavior can tip spontaneously from one state into another. Polyhomeostasis may therefore result in nontrivial dynamical phenomena. Tipping transitions from laminar flow to intermittent chaotic bursts of activities have been observed for networks of rate-encoding and polyhomeostatic adapting neurons [Marković and Gros 2010; 2012].

Tipping transitions can occur both in adaptive and in driven systems. Potential tipping scenarios are currently discussed intensively in the context of climate research [Lenton et al. 2008; Ashwin et al. 2012]. They may be related to a slow driving of external parameters [Baer et al. 1989], to noisy input inducing a stochastic escape from a local attractor [Gammaitoni et al. 1998; McDonnell and Abbott 2009], or through a dynamical effect when the rate of change of a control parameter reaches a certain threshold [Ashwin et al. 2012].

Here we study the phenomenon of self-organized tipping for a polyhomeostatic adapting system driven by a steady-state stochastic input. We examine a previously proposed model [Stemmler and Koch 1999; Triesch 2005] for regulating the firing-rate distribution of individual neurons based on information-theoretical principles. This type of model has been studied previously for the case of discrete time systems and unimodal target firing-rate distributions [Marković and Gros 2010; 2012]. Here we examine the case of continuous time and bimodal target distribution functions, corresponding to sparse coding. For bimodal firing-rate distributions the neural activity tends to switch in between states close to minimal and maximal activity. Similar bimodal activity states are also observed in many other systems, for example, dynamical gene regulation networks [Davidson and Erwin 2006]. We find that bimodal target distributions may lead to self-organized bistability within a certain range of parameters.

We consider a single leaky integrator neuron with noisy input and a sigmoidal transfer function having two degrees of freedom. To achieve a special behavior — here the temporal output distribution of the firing rate — we use polyhomeostasis to change the intrinsic parameters which are directly influencing the mapping of the membrane potential to the firing rate in order to obtain a specific output distribution. We derive these parameter-changing rules using stochastic adaption and show that two degrees of freedom already result in a good behavior approximation, for most of the parameters studied. For bimodal adaption target distributions we observe

self-organized and quasiperiodic stochastic tipping in between two quasistationary attractors resulting from competing adaption gradients.

2. Model

Biological neurons integrate incoming signals and emit an axon potential, a spike, whenever the membrane potential has reached a certain threshold. The membrane potential then returns, after a short refractory period, rapidly to its resting value. This behavior can be captured using spiking integrate-and-fire neural models [Burkitt 2006]. In many circumstances the firing rate, the number of spikes per unit time, is important and rate-encoding neural models can be used [Borst and Theunissen 1999]. Here we consider a single rate-encoding leaky integrator driven by white noise $\xi(t)$,

$$\dot{x}(t) = -\Gamma x(t) + \xi(t), \qquad \langle \xi(t)\xi(t') \rangle = Q\delta(t-t'), \tag{1}$$

where $x > 0$ is the membrane potential and $\Gamma > 0$ the relaxation rate. The firing rate $y(t) \in [0, 1]$ is a nonlinear function of the membrane potential $x(t)$, which we have selected as

$$y(t) = g(x(t)), \qquad g(x) = \frac{1}{1 + e^{-a(x-b)}}, \tag{2}$$

where $a > 0$ is the gain and b is the threshold. The polynomial transfer function (2) maps the membrane potential $x \in [-\infty, \infty]$ to the normalized firing rate $y \in [0, 1]$ which approaches zero and unity for small and large membrane potentials, respectively, compare Figure 1. The slope of $g(x)$ is $a/4$ at the threshold b.

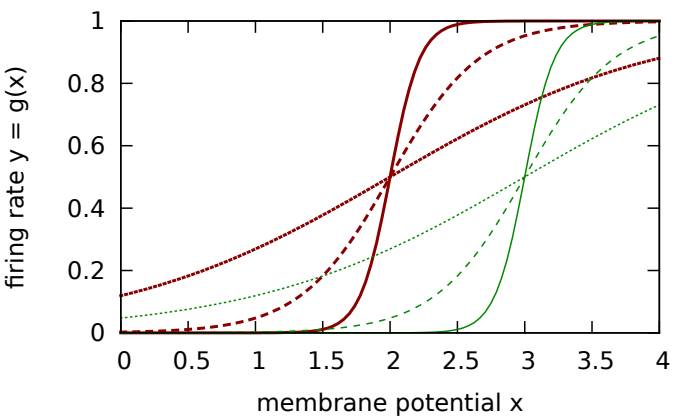


Figure 1. The transfer function $g(x)$ of (2) for thresholds $b = 2$ (red lines) and $b = 3$ (green lines) and various gains a: 1 (dotted), 3 (dashed), and 9 (solid).

Usually the intrinsic parameters of the transfer function (2), a and b, are taken as given by some a priori considerations. Here we will consider them to be slow variables, $a = a(t)$ and $b = b(t)$, adapting slowly such that a target dynamical behavior is approached on the average for the firing rate $y(t)$. The stochastic driving $\xi(t) \in [\Xi_1, \Xi_2]$ in (1) is simulated through white noise plateaus: The values are generated according to a uniform probability distribution (white noise), but they remain constant for short time intervals on the order of unity. The membrane potential averages the input driving noise, due to the leak rate Γ in (1), its distribution function $\rho(x)$ having a mean $\mu_\rho \approx (\Xi_1 + \Xi_2)/(2\Gamma)$ and variance $\sigma_\rho^2 \approx (\Xi_2 - \Xi_1)/(2\Gamma)$.

2.1. ***Polyhomeostatic adaption.*** The firing-rate statistic is given by

$$p(z) = \frac{1}{T}\int_{t_0}^{t_0+T} \delta(z - y(t))\,\mathrm{d}t, \qquad \int_0^1 p(z)\,\mathrm{d}z = 1, \tag{3}$$

where the length T of the sliding observation window is substantially larger than the relaxation rate $1/\Gamma$. The firing-rate distribution $p(z)$ is an important quantity characterizing the information processing capability of biological and artificial neurons. No information is encoded for a constant firing rate, the next value is always exactly the same as before, so no new information is transferred. One may assume that a certain distribution $q(y)$ of firing rates may constitute an optimal working regime. Possible functional dependencies for $q(y)$ can be derived by information-theoretical considerations, for example, maximizing information entropy, as discussed further below.

Considering a given target firing-rate distribution $q(y)$, the closeness of the actual firing-rate distribution $p(y)$ is measured by the Kullback–Leibler divergence D_{KL}, their relative entropy [Gros 2010b]:

$$D_{\mathrm{KL}}(p, q) = \int \mathrm{d}y\, p(y) \ln \frac{p(y)}{q(y)}, \qquad D_{\mathrm{KL}}(p, q) \geq 0. \tag{4}$$

The Kullback–Leibler divergence is positive definite and vanishes only when the two distributions coincide. The KL-divergence is generically not symmetric but becomes symmetric in the limiting case of similar distributions p and q, becoming equivalent in this limit to the χ^2 test [Gros 2010b]. Our aim is now to rewrite (4) as an integral over the membrane potential x, using

$$p(y)\,\mathrm{d}y = \rho(x)\,\mathrm{d}x, \qquad p(y) = \frac{\rho(x)}{\mathrm{d}y/\mathrm{d}x}, \tag{5}$$

where $\rho(x)$ is the membrane potential distribution. Using $y = g(x)$ and (4) and (5), we obtain

$$\frac{\partial D_{\mathrm{KL}}}{\partial \theta} = \int \mathrm{d}x\, \rho(x)\Big[-\frac{1}{g'}\frac{\partial g'}{\partial \theta} - \frac{q'}{q}\frac{\partial g}{\partial \theta}\Big] \equiv \int \mathrm{d}x\, \rho(x)\frac{\partial d}{\partial \theta} \tag{6}$$

for the derivative of the Kullback–Leibler divergence with respect to the intrinsic parameters $\theta = a, b$ of the transfer function $g(x)$; see (2).

We consider now the case in which the system does not dispose of prior information about the distribution of input stimuli and the thereby resulting distribution of membrane potential $\rho(x)$. The best strategy to minimize the Kullback–Leibler divergence is then to minimize the individual terms of the integral (6) through the stochastic adaption rules [Triesch 2005; Marković and Gros 2010]

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = -\epsilon_\theta \frac{\partial d}{\partial \theta}, \qquad \theta = a, b, \tag{7}$$

for the intrinsic parameters of the transfer function $g(x)$, where the ϵ_θ are appropriate small adaption rates.

2.2. ***Target firing-rate distribution.*** In order to evaluate (7), respectively (6), we need to specify the target firing-rate distribution $q(y)$. For this purpose we use information-theoretical considerations.

Given a continuous probability distribution function q its Shannon entropy $H(q)$ can be defined as

$$H(q) = -\int \mathrm{d}y\, q(y) \ln q(y). \tag{8}$$

Among all the real-valued distributions with specified mean μ and standard deviation σ the Gaussian distribution [Gros 2010b]

$$q(y) \propto \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) \propto \exp(\lambda_1 y + \lambda_2 y^2) \tag{9}$$

has maximal information entropy, with $\mu = -\lambda_1/(2\lambda_2)$ and $2\sigma^2 = -1/\lambda_2$, which is easily obtained using variational calculus:

$$0 = \delta\left[H(q) + \lambda_1 \int \mathrm{d}y\, y q(y) + \lambda_2 \int \mathrm{d}y\, y^2 q(y)\right],$$

where $-\lambda_1$ and $-\lambda_2$ are the respective Lagrange parameters. In Figure 2 examples for $q(y)$ are illustrated for several values of λ_1 and λ_2. The support of the target firing rates is compact, $y \in [0, 1]$, and both negative and positive λ_1 and λ_2 can be considered. The normalization factor $\int_0^1 \mathrm{d}y\, q(y)$ cancels out in (6), since only ratios are involved.

For positive $\lambda_2 > 0$ and $\lambda_1 \approx -\lambda_2$ one obtains bimodal target distributions. This is an interesting case, since sparse coding, which is realized when only a minority of neurons in a given network is active, and a majority is inactive [Olshausen and Field 2004], is characterized by a skewed bimodal distribution.

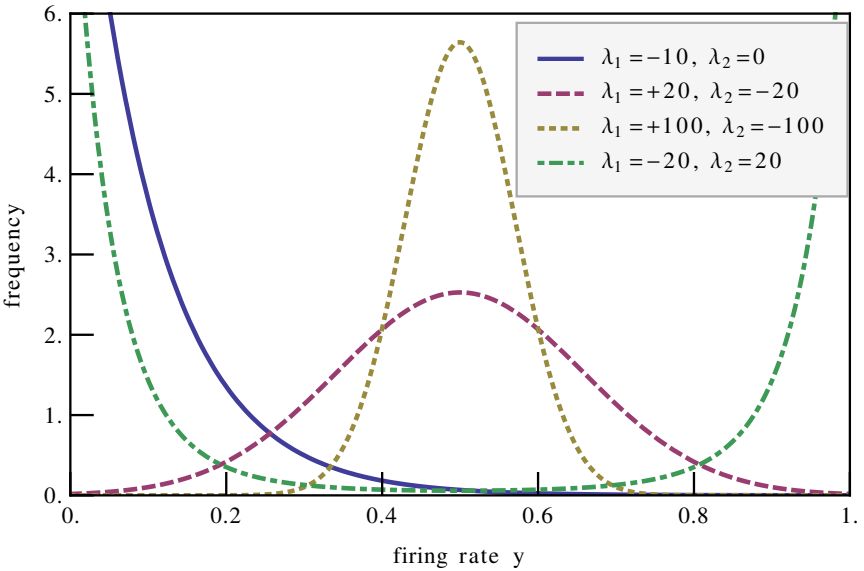


Figure 2. Target distribution $q(y)$, see (9), with some selected parameters λ_1 and λ_2. The target firing-rate distributions are bimodal for $\lambda_2 > 0$.

2.3. ***Stochastic adaption rules.*** From (9) and (2) we find the relations

$$\frac{q'(y)}{q(y)} = \lambda_1 + 2\lambda_2 y, \qquad \frac{\partial g}{\partial x} = ag(1-g),$$

and

$$\frac{\partial g}{\partial a} = (x-b)g(1-g), \qquad \frac{\partial g}{\partial b} = -ag(1-g),$$

which we can use to evaluate the stochastic adaption rules (7) as

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \epsilon_a\Big[\frac{1}{a} + (x-b)[1-2y+(\lambda_1+2\lambda_2 y)(1-y)y]\Big] \tag{10}$$

and

$$\frac{\mathrm{d}b}{\mathrm{d}t} = \epsilon_b\big[-a\big(1-2y+(\lambda_1+2\lambda_2 y)(1-y)y\big)\big]. \tag{11}$$

These two adaption rules will lead to an adaption of the time-averaged firing-rate distribution $p(y)$ towards the target distribution $q(x)$ whenever the adaption time-scales $1/\epsilon_\theta$ are substantially larger than the time constants of the neural dynamics, which in turn are determined by the time scale of the incoming stimuli and by the leak-rate Γ in (1).

The transfer function $g(x)$ contains only two free parameters, the gain a and the threshold b. Perfect adaption $p(y) \equiv q(y)$, for all $y \in [0, 1]$, can hence not

be expected. The system tries to minimize the Kullback–Leibler divergence by adapting the available degrees of freedom, which are just two in our case.

2.4. ***Numerical method.*** Equations (1), (10), and (11) form a set of first-order differential equations with respect to time. We solve them numerically using the Euler method with one evaluation per time step. We confirmed that this method is sufficiently accurate using the fourth-order Runge–Kutta method [Press et al. 2007]. The random white noise $\xi(t)$ is generated through a pseudorandom number generator with a uniform distribution. The values for the leak Γ, the time step Δt, and the learning rates ϵ_a and ϵ_b are shown in the corresponding figures.

3. Results

We performed a series of simulations with the aim of studying two issues. Poly-homeostatic adaption has been studied previously for the case of discrete time systems [Triesch 2005; Marković and Gros 2010]; here we examine the case of continuous time. The case of a bimodal target distribution is, in addition, highly interesting, as it confronts the system with a dilemma. The transfer function $g(x)$, compare Figure 1, is strictly monotonic. The distribution of the membrane potential $\rho(x)$ is hence unimodal. There is no easy way for the adapting neuron to achieve, as a steady-state time-average, a bimodal output firing-rate distribution $p(y)$. The question then is whether the system will find a way out of this dilemma through spontaneous behavioral changes.

3.1. ***Target distribution approximation.*** For most simulations we used, if not stated otherwise, $\Gamma = 1$ for the leak rate and $\Delta t = 10^{-1}$ for the integration time step. A typical time series is given in Figure 3. Note that the adaption of the intrinsic parameters a and b takes place on a slower time scale than that of the primary dynamic variables, x and y, as typical for a slow-fast dynamical system.

Applying moderate to small learning rates $\epsilon_a = \epsilon_b \lesssim 0.01$ the neuron's firing rate y approximates various types of target distributions q quite well. In Figure 4 the achieved and the respective target firing-rate distributions are compared. The respective relative entropies are well minimized and presented in Table 1. Strictly speaking the stochastic adaption rules (10) and (11) are equivalent to approximating the firing-rate statistic (3), which is a time-averaged quantity, towards the target distribution function $q(y)$ only in the limit of very small adaption rates, ϵ_a and ϵ_b. Small but finite values for the adaption rates, as used in our simulations, correspond to a trailing averaging procedure over a limited time interval, and the value of Kullback–Leibler divergence achieved hence depends weakly on the actual values used for the learning rates.

For very high learning rates, $\epsilon_b \gg 0.1$, the threshold b follows the membrane potential x nearly instantaneously, and both variables become highly correlated.

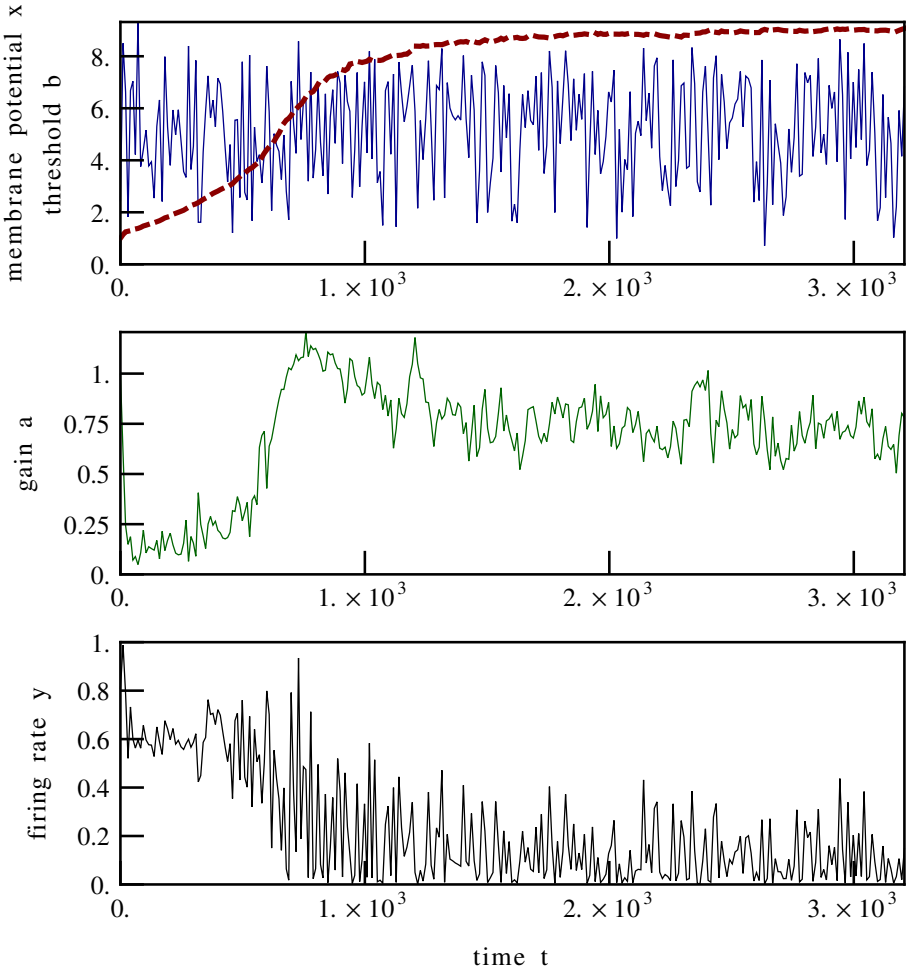


Figure 3. Typical time series for a unimodal target distribution $q(y)$ with $\lambda_1 = -10$, $\lambda_2 = 0$, compare Figure 2. Plotted are the membrane potential x (solid blue line, upper panel), the threshold b (dashed red line, upper panel), the gain a (solid green line, middle panel) and the firing rate y (solid black line, lower panel). $\Delta t = 10^{-1}$, $\epsilon_a = \epsilon_b = 10^{-2}$, and $\Gamma = 1$.

Therefore the firing-rate distribution p cannot approximate the target distribution q anymore. In fact, the resulting Kullback–Leibler divergence is then very high. The tipping in dynamic behavior as a function of adaption rate amplitude is typical for a rate-induced tipping transition [Ashwin et al. 2012].

3.2. ***Gain-threshold phase diagram.*** Due to the sigmoidal shape of the transfer function, several target distributions lead to specific fingerprints in the gain-threshold

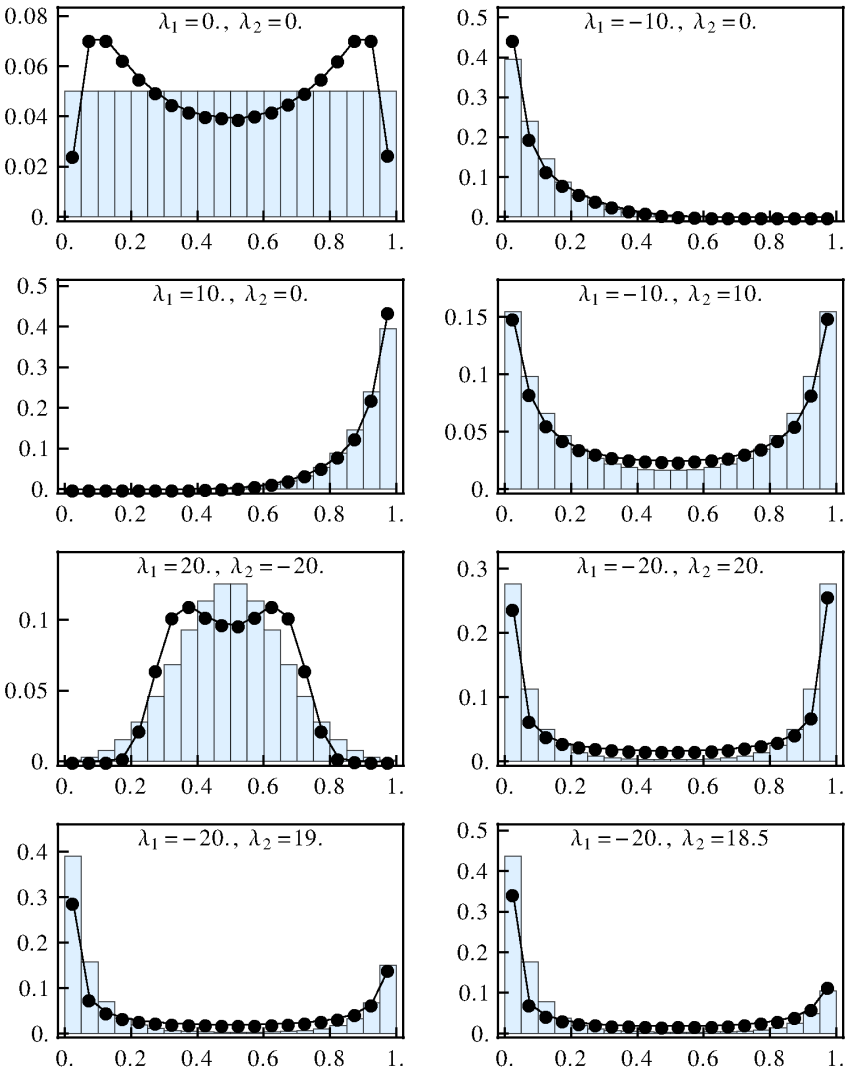


Figure 4. Target distribution q (bars) versus achieved distribution p (points) for different distributions. In each diagram λ_1 and λ_2 are given, while $\Delta t = 10^{-1}$, $t_{\max} = 10^8$, $\epsilon_a = \epsilon_b = 10^{-2}$, $\Gamma = 1$, and $\Xi = [0, 10]$.

phase diagram which we present in Figure 5. The threshold, for example, for a left (right)-dominant target distribution is high (low) and is therefore sensitive to the mean $\mu = -\lambda_1/(2\lambda_2)$ of $q(y)$. Small gains a result in quite flat transfer functions $g(x)$, compare Figure 1, mapping the membrane potentials to similar firing rates y. High gains in a discriminate, relative to the threshold b, on the other side

λ_1	λ_2	Shape	D_{KL}
0	0	Uniform	0.043
−10	0	Left-dominant	0.034
+10	0	Right-dominant	0.028
−10	+10	Left/right-dominant	0.018
+20	−20	Hill	0.076
−20	+20	Left/right, symmetric	0.175
−20	+19	Left/right, left-skewed	0.244
−20	+18.5	Left/right, left-skewed	0.283

Table 1. The relative entropies D_{KL} (4) of various target distributions (see Figure 2) compared to the corresponding achieved distribution, compare Figure 4.

between high and low membrane potentials. The gain is therefore smaller for hill-shaped and flat target distributions, as compared to the left and right-dominant target distributions (for example, $\lambda_1 = -20$, $\lambda_2 = +20$) for which intermediate values are suppressed.

Left (right)-dominant target distributions (compare Table 1) correspond directly to high (low) transfer function thresholds. Uniform, hill, and other not unilateral dominant target distributions lead to intermediate transfer function thresholds with a wide variety of the transfer function gains. For symmetrical target distributions from hill-shaped to diametrical-shaped there is a transition from low to high gains.

3.3. ***Self-organized stochastic escape.*** While the left- and right-dominant unimodal target distributions are easily approximated due to the sigmoidal shape of the transfer function, the bimodal left and right-dominant target distributions put the system in a dilemma: Since intermediate values are to be suppressed the transfer function gain a cannot be too small. Because of this there exist at least two quasistationary fixed points, one for the left part and one for the right part of the distribution.

For zero or small learning rates $\epsilon_a = \epsilon_b \approx 0$ the system is trapped in a single local fixed point. Only the left or right part of the target distribution is then approximated, and the Kullback–Leibler divergence is not well minimized.

Increasing the learning rates $\epsilon_a = \epsilon_b$ allows the system to escape stochastically from the respective local fixed points: the transfer function threshold b conquers the local gradient and moves to the other fixed point and back (compare Figure 6). In the long-term observation the system therefore approximates both the left and the right part of the target distribution and hence minimizes the relative entropy, compare Table 2. These tipping transitions between the two quasistationary fixed points are illustrated in Figure 7, which shows a typical time series for a skewed

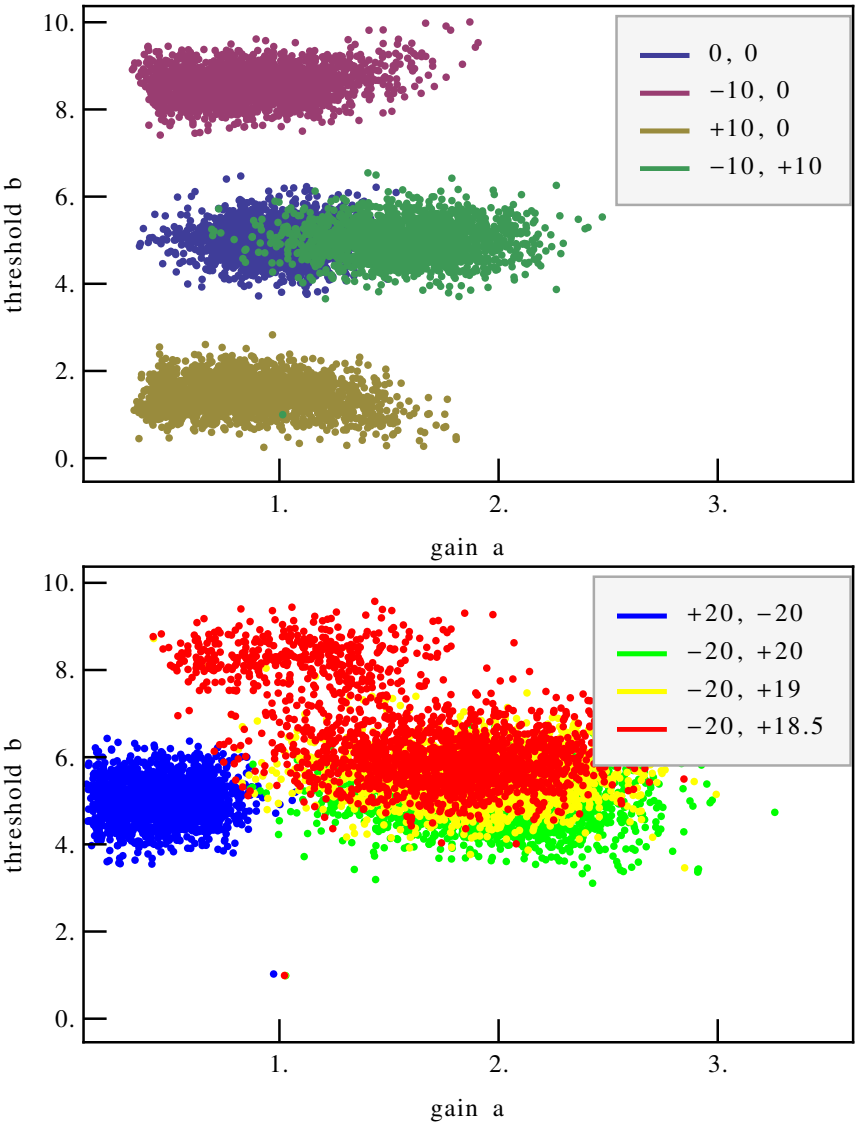


Figure 5. Phase diagram: Plotted are the gain $a(t)$ and the threshold $b(t)$ of the transfer function for various target distributions (λ_1 and λ_2 given in the legend). The respective target and achieved firing-rate distributions are given in Figure 4.

target distribution. Note that there are two fixed points for the gain and threshold and a direct correspondence to the periods of high and low firing rates $y(t)$.

Very low learning rates ϵ_a and ϵ_b lead to deep and large basins of attraction for the respective fixed points, while on the other hand high learning rates result in the

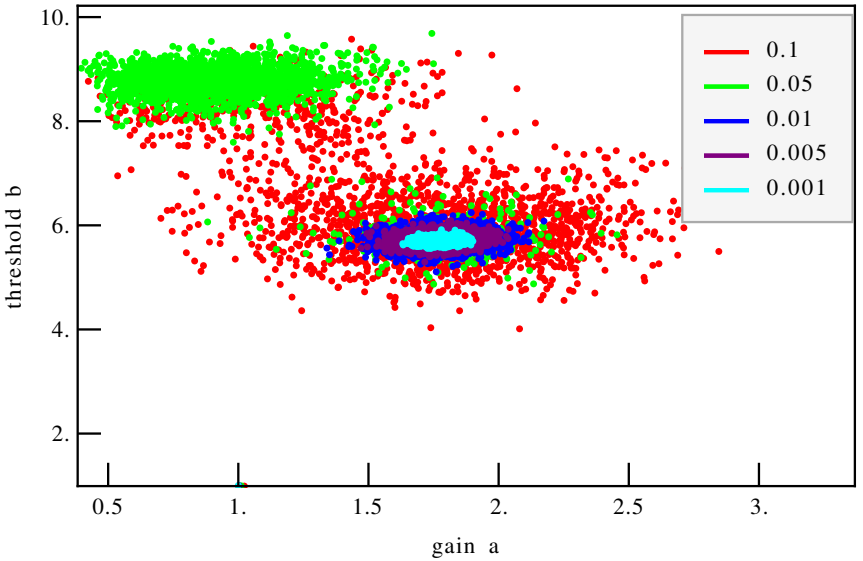


Figure 6. Stochastic escape: Phase diagram of the transfer function gain versus transfer function threshold for a convex left-skewed target distribution with various learning rates ($\epsilon_a = \epsilon_b$ given in the legend), with $\Delta t = 10^{-1}$, $\Gamma = 1$, $\lambda_1 = -20$, and $\lambda_2 = 18.5$.

close following of the threshold to the membrane potential which prohibits reaching the target distribution. This mechanism is reminiscent of the case of Langevin dynamics in a double-well potential [Hanggi 1986], where a stochastically driven particle may switch forth and back between two local minima [Gros 2010b]. The switching time is controlled for the double-well problem by the Kramers escape rate, which depends exponentially on the potential barrier height. It is difficult to formulate a quantitative mapping to the double-well problem. The local attractors visible in Figures 6 and 7, and the effective barriers in between them, are self-organized structures. Note that the strength Q of the noise term (1) is constant and influences the transition rate only weakly, due to the continuous adaption of the transfer function, via (10) and (11), to the average strength of the stochastic driving.

$\epsilon_a = \epsilon_b$	10^{-5}	10^{-4}	10^{-3}	$5 \cdot 10^{-3}$	10^{-2}	$5 \cdot 10^{-2}$	10^{-1}
D_{KL}	0.306	0.295	0.293	0.289	0.283	0.154	0.109

Table 2. Relative entropies D_{KL} (4) for the left-skewed target distribution ($\lambda_1 = -20$, $\lambda_2 = 18.5$) relative to the achieved distribution for various learning rates ϵ_a and ϵ_b, compare Figure 6.

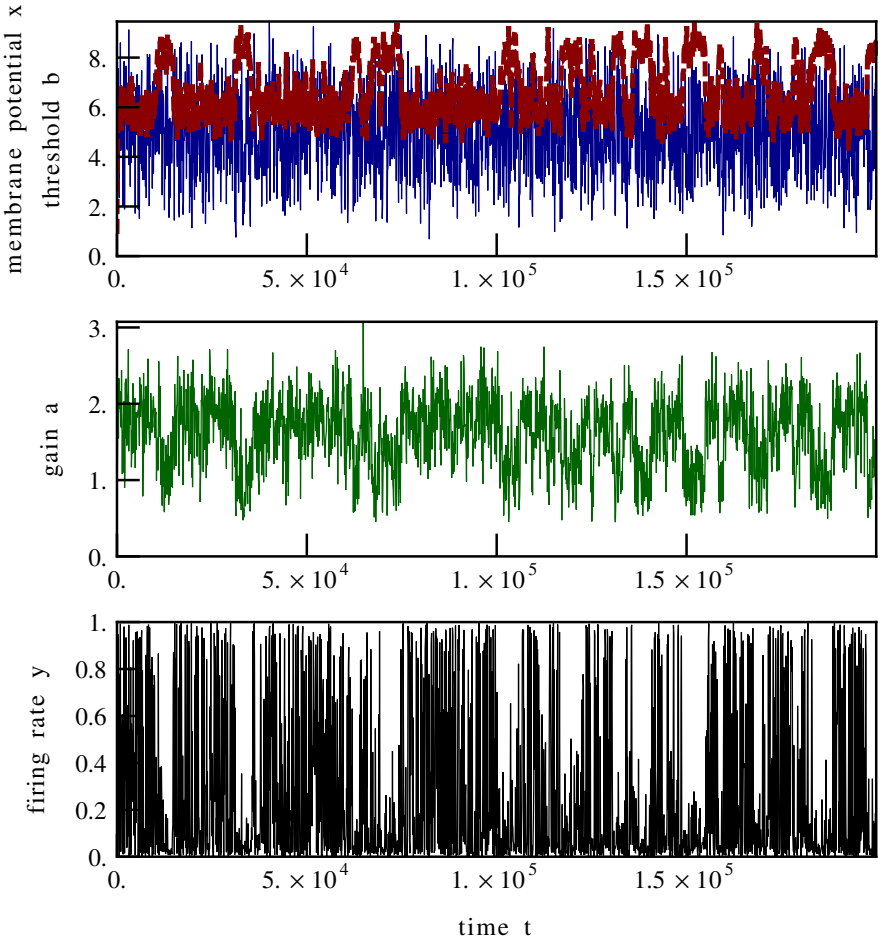


Figure 7. Time series of the membrane potential x (top, blue line), threshold b (top, red line), gain a (middle, green line), and firing rate y (bottom, black line), with $\Delta t = 10^{-1}$, $\epsilon_a = \epsilon_b = 10^{-1}$, $\Gamma = 1$, $\lambda_1 = -20$, and $\lambda_2 = 18.5$.

4. Discussion

We showed that polyhomeostatic adaption of a continuous-time leaky integrator leads to the desired firing-rate distributions. We also run further simulations using white noise and Gaussian noise input and replace the transfer function by other qualitatively different (but still sigmoidal) functions, see the Appendix. It turns out that the polyhomeostatic adaption as well as the self-organized stochastic escape are quite robust principles. However, the quality of the approximation (as seen by visual overlapping) and the value of the Kullback–Leibler divergence depend on the learning rates, and also on the input distribution and the input's strength.

The stochastic tipping as a function of adaption rates has a close relation to the phenomenon of stochastic escape. The strength of the driving input noise is constant, but its influence is averaged out for very low adaption rates. Stochastic escape from one local attractor to another is not possible. The stochasticity of the input becomes relevant for intermediate values of adaption rates and stochastic transitions between the two quasistationary attractors are most frequent. Finally, for very large adaption rates, the system tips into another dynamical state, tracking the stochastic input signal nearly instantaneously. This sequence of behaviors is self organized and can be reached from any initial state.

Appendix: Polynomial transfer function

The polyhomeostatic adaption of the system does not change qualitatively by replacing the transfer function g. Instead it turns out that the system is robust against changing the transfer function as long as it remains sigmoidal. We also applied a transfer function

$$g(x) = \frac{(x/b)^{ab}}{(x/b)^{ab} + 1}, \tag{A.1}$$

with a polynomial decay to $g(0) = 0$, which limits the membrane potential $x \geq 0$ to be nonnegative. It turns out that the shape of the target distribution q is also well approximated using this transfer function. Stochastic escape from one fixed point to another and back can be observed in addition, since two fixed points are necessary for some target distributions.

The transfer function has an inflection point for exponents $ab > 0$; it is absent for $ab < 1$, compare Figure 8. The transfer function g behaves as

$$g(x) \approx \begin{cases} (x/b)^{ab}, & x \ll b, \\ \frac{1}{2} + \frac{1}{4}a(x-b), & x \approx b, \\ 1 - (b/x)^{ab}, & x \gg b. \end{cases} \tag{A.2}$$

The slope is $a/4$ which approaches zero and unity for small and large membrane potentials respectively.

From (A.1) we find the relations

$$\frac{\partial g}{\partial x} = (1-g)g\frac{ab}{x}, \tag{A.3}$$

$$\frac{\partial g}{\partial a} = (1-g)gb\ln\left(\frac{x}{b}\right), \qquad \frac{\partial g}{\partial b} = (1-g)ga\left[\ln\left(\frac{x}{b}\right) - 1\right], \tag{A.4}$$

which we can use to evaluate the stochastic adaption rules (7) as

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \epsilon_a\left[\frac{1}{a} - b\ln(x/b)[1 - 2y + (\lambda_1 + 2\lambda_2 y)(1-y)y]\right] \tag{A.5}$$

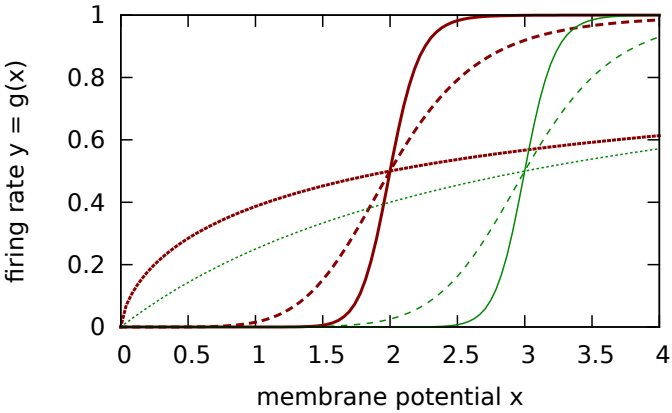


Figure 8. The transfer function $g(x)$, see (A.1), for thresholds $b = 2$ (red lines) and $b = 3$ (green lines) and various gains a: 1/3 (dotted), 3 (dashed), and 9 (solid). No inflection point is present for exponents $ab < 1$.

and

$$\frac{\mathrm{d}b}{\mathrm{d}t} = \epsilon_b\Big[\frac{1}{b} - a[\ln(x/b) - 1][1 - 2y + (\lambda_1 + 2\lambda_2 y)(1 - y)y]\Big]. \tag{A.6}$$

Applying this transfer function g it turns out that the target distribution is well approximated also in this case, even though the membrane potential is restricted to nonnegative numbers. Table 3 lists the well-minimized Kullback–Leibler divergences for several target distributions.

λ_1	λ_2	Shape	D_{KL}
0	0	Uniform	0.060131
−10	0	Left-dominant	0.069351
+10	0	Right-dominant	0.114578
−10	+10	Left/right-dominant	0.051811
+20	−20	Hill	0.148098
−20	+20	Left/right, symmetric	0.189217
−20	+19	Left/right, left-skewed	0.063934
−20	+18.5	Left/right, left-skewed	0.261215

Table 3. Relative entropies of various target distributions compared to the corresponding achieved distribution ($\epsilon_a = \epsilon_b = 10^{-2}$, bins = 100).

We conclude that the stochastic adaption rules are therefore generic and qualitatively independent on the concrete realization of the transfer function. However, quantitatively the resulting relative entropies depend on the choice of the transfer function which also influences the optimal adaption rates ϵ_a and ϵ_b.

A.1 ***Self-organized stochastic escape.*** For the nonsymmetric convex target distribution ($\lambda_1 = -20$, $\lambda_2 = 19$) there are two fixed points. Since the target distribution cannot be well approximated by only one fixed point the system escapes stochastically from one to the other and back within a certain period, compare with Figures 9 and 10. For small learning rates $\epsilon_a = \epsilon_b \lessapprox 0.01$ the system is trapped in only one fixed point. The relative entropy therefore is not well minimized.

For intermediate learning rates $0.01 \lessapprox \epsilon_a = \epsilon_b \lessapprox 0.04$ the perturbation is high enough to stochastically escape from that fixed point and approach another one. Figures 9 and 10 show a typical time series for this tipping. This has also an effect on the relative entropy which is therefore even smaller than without stochastic escape (see Table 4).

For high learning rates $\epsilon_a = \epsilon_b \gtrapprox 0.05$ the system's behavior changes: the transfer function is close to a Heaviside step function and the threshold follows the membrane potential quickly. In that state the achieved distribution is not close

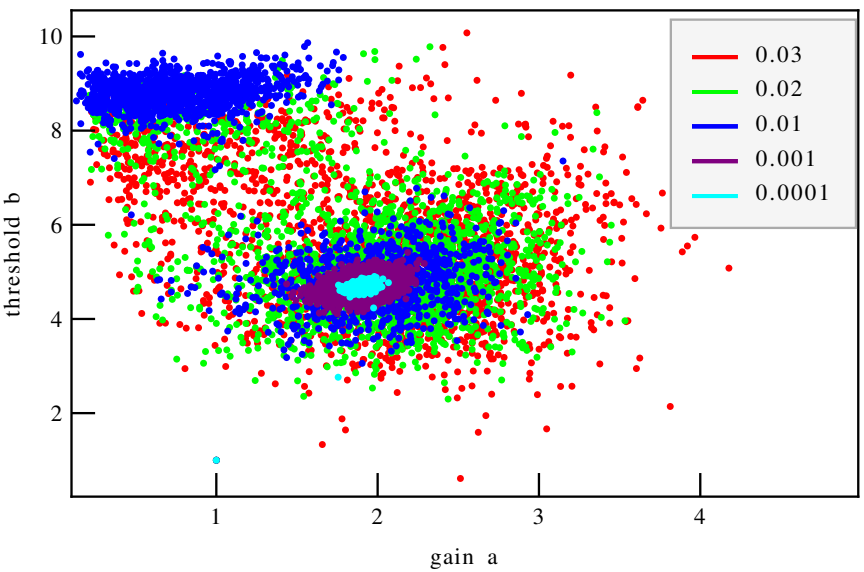


Figure 9. Stochastic escape: Phase diagram of the transfer function gain versus transfer function threshold for a convex left-skewed target distribution with various learning rates (ϵ_a and ϵ_b given in the legend), with $\Delta t = 10^{-2}$, $\Gamma = 0.1$, $\lambda_1 = -20$, and $\lambda_2 = 19$.

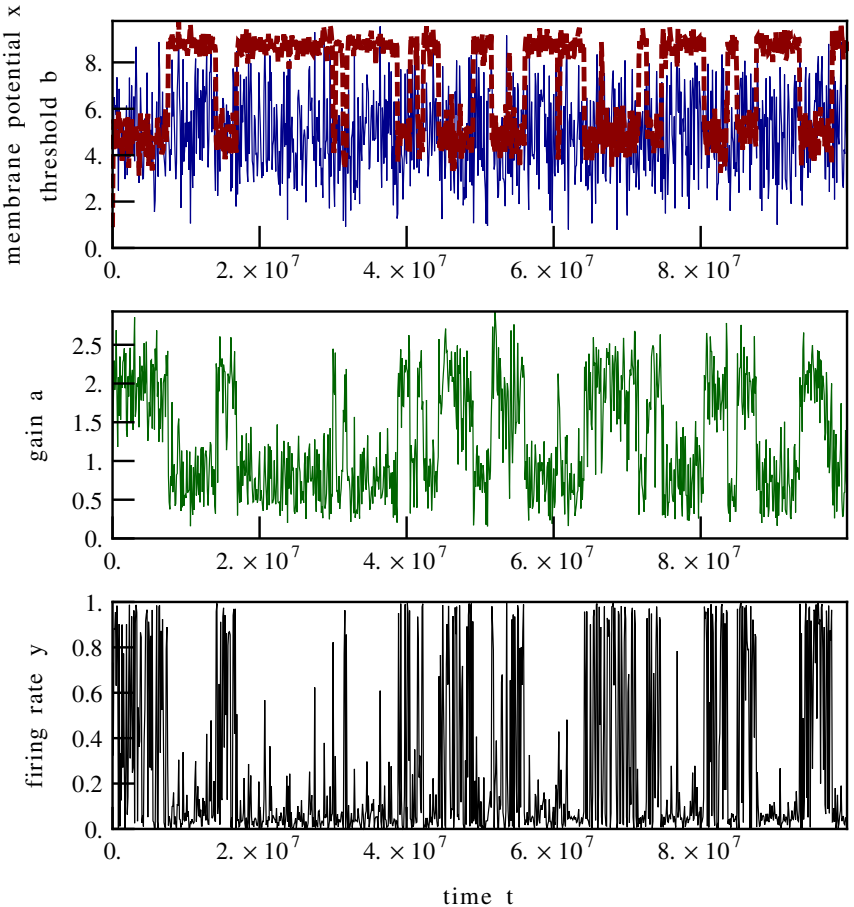


Figure 10. Time series: membrane potential x, transfer function threshold b (dashed), transfer function gain a, and firing rate y, with $\Delta t = 10^{-1}$, $\epsilon_a = \epsilon_b = 10^{-2}$, $\Gamma = 0.1$, $\lambda_1 = -20$, and $\lambda_2 = 19$.

to the target distribution, therefore the relative entropy is not minimized anymore (see Table 4).

$\epsilon_a = \epsilon_b$	10^{-4}	10^{-3}	0.01	0.03	0.04	0.05	0.06
D_{KL}	0.376	0.368	0.064	0.043	0.017	1.892	1.591

Table 4. Relative entropies of the left-skewed target distribution ($\lambda_1 = -20$, $\lambda_2 = 19$) compared to the achieved distribution for various learning rates ϵ_a and ϵ_b. Note that the Kullback–Leibler divergence is not minimized for $\epsilon_b \gtrapprox 0.05$ due to the fast correlation of the membrane potential and the transfer function threshold.

References

[Ashwin et al. 2012] P. Ashwin, S. Wieczorek, R. Vitolo, and P. Cox, "Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system", *Phil. Trans. R. Soc. A* **370**:1962 (2012), 1166–1184.

[Baer et al. 1989] S. M. Baer, T. Erneux, and J. Rinzel, "The slow passage through a Hopf bifurcation: delay, memory effects, and resonance", *SIAM J. Appl. Math.* **49**:1 (1989), 55–71.

[Borst and Theunissen 1999] A. Borst and F. E. Theunissen, "Information theory and neural coding", *Nat. Neurosci.* **2**:11 (1999), 947–957.

[Burkitt 2006] A. N. Burkitt, "A review of the integrate-and-fire neuron model, I: Homogeneous synaptic input", *Biol. Cybern.* **95**:1 (2006), 1–19.

[Charkoudian 2003] N. Charkoudian, "Skin blood flow in adult human thermoregulation: how it works, when it does not, and why", *Mayo Clin. Proc.* **78**:5 (2003), 603–612.

[Davidson and Erwin 2006] E. H. Davidson and D. H. Erwin, "Gene regulatory networks and the evolution of animal body plans", *Science* **311**:5762 (2006), 796–800.

[Doya 2002] K. Doya, "Metalearning and neuromodulation", *Neural Netw.* **15**:4–6 (2002), 495–506.

[Gammaitoni et al. 1998] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, "Stochastic resonance", *Rev. Mod. Phys.* **70**:1 (1998), 223–288.

[Gros 2010a] C. Gros, "Cognition and emotion: perspectives of a closing gap", *Cogn. Comput.* **2**:2 (2010), 78–85. arXiv 1002.3035

[Gros 2010b] C. Gros, *Complex and adaptive dynamical systems: a primer*, 2nd ed., Springer, Berlin, 2010.

[Hanggi 1986] P. Hanggi, "Escape from a metastable state", *J. Stat. Phys.* **42**:1-2 (1986), 105–148.

[Kernbach and Kernbach 2011] S. Kernbach and O. Kernbach, "Collective energy homeostasis in a large-scale microrobotic swarm", *Robot. Auton. Sys.* **59**:12 (2011), 1090–1101.

[Kuehn 2011] C. Kuehn, "A mathematical framework for critical transitions: bifurcations, fast-slow systems and stochastic dynamics", *Physica D* **240**:12 (2011), 1020–1035. arXiv 1101.2899

[Lenton et al. 2008] T. M. Lenton, H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber, "Tipping elements in the earth's climate system", *Proc. Nat. Acad. Sci. USA* **105**:6 (2008), 1786–1793.

[Marder and Goaillard 2006] E. Marder and J. M. Goaillard, "Variability, compensation and homeostasis in neuron and network function", *Nat. Rev. Neurosci.* **7** (2006), 563–574.

[Marković and Gros 2010] D. Marković and C. Gros, "Self-organized chaos through polyhomeostatic optimization", *Phys. Rev. Lett.* **105**:6 (2010), Article ID #068702. arXiv 1001.0663

[Marković and Gros 2012] D. Marković and C. Gros, "Intrinsic adaptation in autonomous recurrent neural networks", *Neural Comput.* **24**:2 (2012), 523–540.

[McDonnell and Abbott 2009] M. D. McDonnell and D. Abbott, "What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology", *PLoS Comput. Biol.* **5**:5 (2009), Article ID #e1000348.

[Olshausen and Field 2004] B. A. Olshausen and D. J. Field, "Sparse coding of sensory inputs", *Curr. Opin. Neurobiol.* **14**:4 (2004), 481–487.

[Plum et al. 2006] L. Plum, B. F. Belgardt, and J. C. BrÃ¼ning, "Central insulin action in energy and glucose homeostasis", *J. Clin. Investig.* **116**:7 (2006), 1761–1766.

[Press et al. 2007] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes: the art of scientific computing*, 3rd ed., Cambridge University Press, Cambridge, 2007.

[Schaefer 1961] E. K. Schaefer, "Blood pH and pCO_2 homeostatis in chronic respiratory acidosis related to the use of amine and other buffers", *Ann. NY Acad. Sci.* **92** (1961), 401–413.

[Stemmler and Koch 1999] M. Stemmler and C. Koch, "How voltage-dependent conductances can adapt to maximize the information encoded by neuronal firing rate", *Nat. Neurosci.* **2**:6 (1999), 521–527.

[Tresguerres et al. 2010] M. Tresguerres, S. K. Parks, E. Salazar, L. R. Levin, G. G. Goss, and J. Buck, "Bicarbonate-sensing soluble adenylyl cyclase is an essential sensor for acid/base homeostasis", *Proc. Nat. Acad. Sci. USA* **107**:1 (2010), 442–447.

[Triesch 2005] J. Triesch, "A gradient rule for the plasticity of a neuron's intrinsic excitability", pp. 65–70 in *Artificial neural networks: biological inspirations* (*ICANN 2005*), *I* (Warsaw, 2005), edited by W. Duch et al., Lecture Notes in Computer Science **3696**, Springer, Berlin, 2005.

Received 4 Apr 2012. Revised 13 Jul 2012. Accepted 3 Nov 2012.

MATHIAS LINKERHAND: linkerhand@th.physik.uni-frankfurt.de
Institute for Theoretical Physics, Goethe University, Max-von-Laue-Straße 1, D-60438 Frankfurt am Main, Germany

CLAUDIUS GROS: gros07@itp.uni-frankfurt.de
Institute for Theoretical Physics, Goethe University, Max-von-Laue-Straße 1, D-60438 Frankfurt am Main, Germany


MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS

msp.org/memocs

MEMOCS (ISSN 2325-3444 electronic, 2326-7186 printed) is a journal of the International Research Center for the Mathematics and Mechanics of Complex Systems at the Università dell'Aquila, Italy.

Cover image: "Tangle" by © John Horigan; produced using the *Context Free* program (contextfreeart.org).

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Mathematics and Mechanics of Complex Systems
vol. 1 no. 2 2013

Self-organized stochastic tipping in slow-fast dynamical systems 129
Mathias Linkerhand and Claudius Gros

Well-posedness for dislocation-based gradient viscoplasticity, II: general nonassociative monotone plastic flows 149
Sergiy Nesenenko and Patrizio Neff

Symmetry classes for even-order tensors 177
Marc Olive and Nicolas Auffray

A direct approach to nonlinear shells with application to surface-substrate interactions 211
Miroslav Šilhavý

A sufficient condition for a discrete spectrum of the Kirchhoff plate with an infinite peak 233
Fedor L. Bakharev, Sergey A. Nazarov and Guido H. Sweers

MEMOCS is a journal of the International Research Center for the Mathematics and Mechanics of Complex Systems at the Università dell'Aquila, Italy.