Multiscale Model of an Inhibitory Network Shows Optimal Properties near Bifurcation

Christopher L. Buckley* and Thomas Nowotny†

CCNR, University of Sussex, Falmer, Brighton BN1 9QJ, United Kingdom

(Received 14 October 2010; revised manuscript received 24 November 2010; published 10 June 2011)

We present a systematic multiscale reduction of a biologically plausible model of the inhibitory neuronal network of the pheromone system of the moth. Starting from a Hodgkin-Huxley conductance based model we adiabatically eliminate fast variables and quantitatively reduce the model to mean field equations. We then prove analytically that the network's ability to operate on signal amplitudes across several orders of magnitude is optimal when a disinhibitory mode is close to losing stability and the network dynamics are close to bifurcation. This has the potential to extend the idea that optimal dynamic range in the brain arises as a critical phenomenon of phase transitions in excitable media to brain regions that are dominated by inhibition or have slow dynamics.

DOI: 10.1103/PhysRevLett.106.238109 PACS numbers: 87.19.lj, 87.19.ll

The extraordinary ability of brains to operate on sensory signal amplitudes across several orders of magnitude has been well established by psychophysicists [1] and behavioral biologists [2] alike, but the underlying mechanisms are still not fully understood. One key model system to understand the origin of this dynamic range (DR) is the pheromone system in the antennal lobe (AL) of male moths that allows them to find females from more than a mile away [3]. Given the limitations of individual neurons within the AL [4] it is likely that a significant contribution to the DR arises from their collective properties and growing evidence indicates that biological systems like the AL may exploit self-organized criticality [5,6]. It has been demonstrated in a series of works, e.g., [7], that dynamics close to a phase transition can underlie large DR in excitable media and Kinouchi *et al.* [8] subsequently suggested that this result extends to the gap junction coupled neurons in the olfactory bulb (OB), the vertebrate equivalent of the AL. However, gap junctions have not been found in the AL, the dynamics are dominated by slow inhibitory ($GABA_B$) synapses and, perhaps most importantly, recordings from the AL are not consistent with excitable dynamics. Here we use the olfactory system as a test case and develop a framework for describing a critical point in a nonexcitable network inspired by the ideas of [9]. We analytically reduce a biologically plausible conductance based model of the macroglomerular complex (MGC), the pheromone subsystem of the AL, by adiabatically eliminating fast variables and then construct mean field (MF) equations. We demonstrate how a disinhibitory mode in the recurrent network of inhibitory local neurons (LNs) can mediate signal transmission from the olfactory receptor neurons (ORNs) to the intrinsically active projection neurons (PNs) and show analytically how sensitivity and DR are maximized when this disinhibitory mode is close to losing stability and the network is close to bifurcation.

Conductance based model.—We model the MGC as a network of N Hodgkin-Huxley (HH) neurons interacting through a random inhibitory coupling matrix $-\hat{G} \equiv -\hat{g}_{ij}$ where $\hat{g}_{ij} \in \{1, 0\}$ with probability P and $1 - P$, respectively. The membrane potential of neurons is given by

$$C\dot{V}_i = -I_{\mathrm{Na}} - I_K - I_L - I_M - I_{i,\mathrm{DC}} - I_{i,\mathrm{syn}} - I_i, \quad (1)$$

where I_i is an external current and $I_{i,\mathrm{DC}}$ a constant bias. The leak current is $I_L = g_L(V_i - E_L)$ and the ionic currents I_{Na} and I_K are described by standard HH equations [10], as in [11]. I_M is a spike-frequency adaptation current, characterized by the slow gating variable z_i with $\dot{z}_i = [H(V_i) - z_i]/\tau_z$, and $H(V_i) = 0.01/\{1 + \exp[-(V_i + 20)/5]\}$. The remaining parameter values are $C = 0.143$ nF, $g_L = 0.02672$ μS, $E_L = -63.563$ mV, $g_{\mathrm{Na}} = 7.15$ μS, $E_{\mathrm{Na}} = 50$ mV, $g_K = 1.43$ μS, $E_K = -95$ mV, and $\tau_z = 50$ ms. Neurons are connected by first order synapses [12] with a fast rise ($\alpha = 1$ kHz) and slow decay ($\beta = 0.01$ kHz) of neurotransmitter s_i in response to spiking events, see Fig. 1(b),

$$\dot{s}_i = -\beta s_i + \alpha\Theta(t - t_s)\Theta(t_s + t_r - t), \quad (2)$$

where t_s is the last time the membrane potential exceeded the threshold V_{th}, Θ is the Heaviside function, and t_r is the duration of transmitter release. The synaptic current onto neuron i is then given by

$$I_{i,\mathrm{syn}} = \sum_j - \hat{g}_{ij} s_j (V_i - V_{\mathrm{rev}}), \quad (3)$$

where V_i denotes the potential of the ith neuron, $V_{\mathrm{rev}} = -90$ mV, $t_r = 1$ ms, and $V_{\mathrm{th}} = 20$ mV.

Reduction to a rate model.—Given that the synapses are relatively slow compared to the membrane dynamics (i.e., β is sufficiently small) we can adiabatically eliminate the latter from the network description. To do so, we start by constructing a mapping between a static input current and the steady state firing frequency F of the neuron (F-I curve). Without spike-frequency adaptation ($g_m = 0$) the neurons exhibit type 1 excitability (i.e., spiking onsets

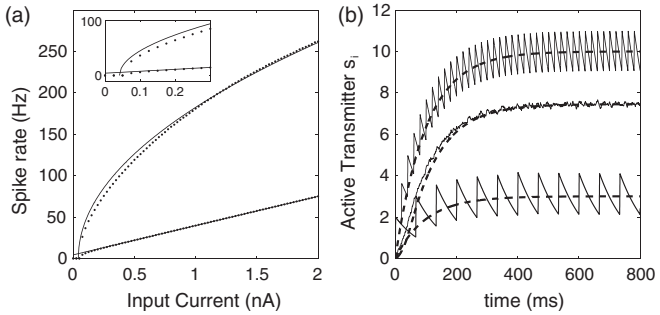

FIG. 1. (a) *F-I* curve for a HH neuron with (lower points) and without (upper points) spike-frequency adaptation and the corresponding functional fits (lines). Inset: details of the curve at onset. (b) Growth of s_i in response to spike input frequencies of 50 Hz (top) and 15 Hz (bottom). The middle traces show a sum of 30 synapses with input frequencies $\in [15, 40]$ Hz. Dashed lines are smooth approximations according to (4).

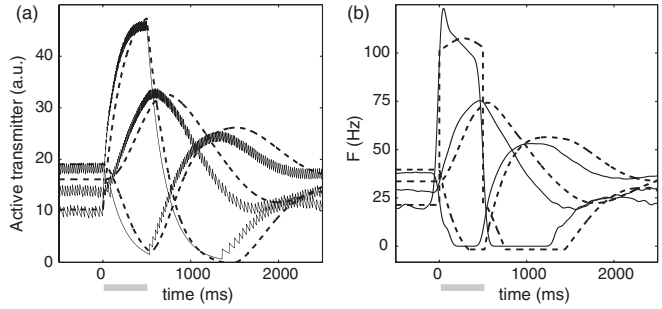

FIG. 2. The dynamics of three neurons from a network with $N_+ = 5$, $N_- = 15$, $P = 0.5$, and $p_\lambda = 0.9$ in response to a current pulse of $I = 1$ nA (gray bar). (a) active transmitter concentration s_i in the full model (solid) and the approximation (dashed), (b) spike density function (solid) and the rate approximation (dashed).

through a saddle-node bifurcation on a limit cycle) and the *F-I* curve is well fit by a square root function which has an infinite derivative at spiking onset, see Fig. 1(a). Spike-frequency adaptation ($g_m > 0$) can effectively linearize the *F-I* curve [13], $F(I) = [mI + C]_+$, see Fig. 1(a), where $[x]_+ \equiv \max\{x, 0\}$ and m and C are determined numerically by a linear fit to the *F-I* curve. We can view the response of the synapse to a constant input frequency as the leaky integration of a square wave of transmitter activation and can construct a smooth approximation to (2) as

$$\dot{s}_i = -\beta s_i + \alpha t_r F. \tag{4}$$

The approximation works best with high input frequencies or low β [Fig. 1(b)]. Summing across many synapses, assuming randomly distributed phases, further improves the approximation [Fig. 1(b), middle line]. Combining (3) and (4) we obtain

$$\dot{\mathbf{s}} = -\beta \mathbf{s} + \gamma[-\tilde{G}\mathbf{s} + \boldsymbol{\theta} + \mathbf{I}], \tag{5}$$

where $\gamma(x) = \gamma_c[x]_+$, $\gamma_c = \alpha m t_r$ and $\theta \equiv \theta_i \equiv I_{i,\mathrm{DC}} + C$ which has absorbed the constant offset, C, from the *F-I* curve. Except during a spike the membrane potential of each neuron is close to its resting potential $V^*_{i,\mathrm{rest}}$ which we calculated numerically in the baseline state of the network. We then approximate the synaptic input to each neuron as $\tilde{G} \equiv \tilde{g}_{ij} = \hat{g}_{ij}(V^*_{i,\mathrm{rest}} - V_{\mathrm{rev}})$. The rate of neuron i will then be given by $F_i = m[-\tilde{G}\mathbf{s} + \boldsymbol{\theta} + \mathbf{I}]_+$.

Globally stable dynamics.—In the absence of stimulation the baseline frequencies of the LNs are distributed between 15 and 40 Hz. Pheromone stimulation induces transient rate oscillations with periods on the order of 10^2 ms [4]. To model these phenomena we construct a system with globally stable rate dynamics. The local stability around a fixed point (FP) $\mathbf{s}^* = s^*_1, \ldots, s^*_N$ is determined by the eigenvalues of the Jacobian $J = -\gamma_c \tilde{G} - \beta 1$. The network dynamics are locally stable if [14]

$$\lambda_{\max}(-\gamma_c G) < \beta. \tag{6}$$

We therefore can control the stability of the system by scaling the weights as $G = \kappa \tilde{G}$ where $\kappa = \frac{p_\lambda \beta}{\lambda_{\max}(\gamma_c \tilde{G})}$. $p_\lambda \geq 0$ is a stability parameter, i.e., the FP is stable if $p_\lambda < 1$. In order to obtain appropriate baseline frequencies $\in [F_{\min}, F_{\max}]$ we choose $s^*_i \in [s^*_{\min}, s^*_{\max}]$ with $s^*_{\min} = \frac{\beta F_{\min}}{\alpha t_r}$, $s^*_{\min} = \frac{\beta F_{\max}}{\alpha t_r}$, see (4), and, setting $\frac{ds_i}{dt} = 0$, $\forall_i$ and $\mathbf{I} \equiv \mathbf{0}$ in (5), we get the necessary biases

$$\boldsymbol{\theta} = \gamma^{-1}(\beta \mathbf{s}^*) + G\mathbf{s}^*. \tag{7}$$

The local stability determined by p_λ is only necessary but not sufficient for global stability. However, like [15] we find that locally stable networks are almost always globally stable and use (6) as a practically sufficient condition for global stability. Figure 2 shows example traces from a stable network ($p_\lambda < 1$) with baseline frequencies $\in [15, 40]$ Hz. When the FP is unstable ($p_\lambda > 1$) the system exhibits oscillations (perhaps chaos) or saturating dynamics depending on the connectivity matrix.

A disinhibitory pathway.—Odor stimulation of ORNs increases the firing rate of PNs. While evidence of direct excitatory connections from ORNs to PNs exists [16], other evidence suggests that a significant part of the PN excitation arises from LNs through a disinhibitory pathway [2]. Here we describe how such a pathway, in which the suppression of some LNs releases intrinsically active PNs from inhibition, could arise in a random recurrent network of inhibitory neurons. We describe the MGC in terms of a set of N_+ stimulated (LNs that receive input from ORNs) s_i, $1 \leq i \leq N_+$ and N_- unstimulated LNs s_i, $N_+ < i \leq N$, i.e., $I_i = I$ for $i = 1, \ldots, N_+$ and $I_i = 0$ otherwise. We assume that stimulated LNs do not synapse on PNs and hence quantify the disinhibition of the PNs in terms of the suppression of activity of unstimulated LNs, see Fig. 4(b) (inset). This scenario is described by (5) where $\tilde{g}_{ij}$ is a random matrix with entries $\tilde{g}_{ij} \in \{0, 1\}$ if $i, j \leq N_+$ or $i, j > N_+$ (intraconnections) and $\tilde{g}_{ij} \in \{0, \rho\}$ otherwise (interconnections). The parameter ρ scales the ratio between the intra- and interpopulation connectivities. $\tilde{G}$

is scaled with κ for some p_λ and the biases θ_i are set with (7) such that the frequencies are $\in [15, 40]$ Hz. When the network coupling is moderate, stimulation causes a net decrease in the activity of the unstimulated LNs and hence a net disinhibition of PNs. We can see this by linearizing (5) around the FP to find, using (7),

$$\boldsymbol{\sigma}^* = \frac{\gamma_c}{\beta}\left[\mathbb{1} + \frac{\gamma_c}{\beta}G\right]^{-1}\mathbf{I}, \tag{8}$$

where we have set $\sigma_i^* \equiv s_i^* - s_i^{*(0)}$, the displacement of the FP s_i^* with input current $\mathbf{I}$, from the FP $s_i^{*(0)}$ without input ($\mathbf{I} \equiv 0$). For a weakly coupled system, $\beta > \gamma_c\|G\|$ (where $\|\cdot\|$ is an appropriate matrix norm), we can expand the right-hand side to obtain $\boldsymbol{\sigma}^* \approx \frac{\gamma_c}{\beta}[\mathbb{1} - \frac{\gamma_c}{2\beta}(G + G^T) + \mathcal{O}(\|G\|^2)]\mathbf{I}$ and hence find the FP of the LN populations as

$$\begin{aligned}\sigma_i^{+*} &\approx I\frac{\gamma_c}{\beta} - I\frac{\gamma_c^2}{2\beta^2}\left(\sum_{j=1}^{N_+}[g_{ij} + g_{ji}]\right)\\ \sigma_i^{-*} &\approx -I\frac{\gamma_c^2}{2\beta^2}\left(\sum_{j=1}^{N_+}[g_{ij} + g_{ji}]\right).\end{aligned} \tag{9}$$

Hence the stimulated LNs will increase in activation and the unstimulated LNs will decrease.

In a strongly coupled system we can analyze the network response by eigendecomposition of the bracket in (8), which is the Jacobian J of the system. Decomposing it as $J = Q\Lambda Q^{-1}$ we find $\boldsymbol{\sigma}^* = -\gamma_c[Q\Lambda^{-1}Q^{-1}]\mathbf{I}$ where Q is the matrix of eigenvectors and Λ is the diagonal matrix of eigenvalues of J. As $p_\lambda \to 1$, the real part of one eigenvalue, the critical eigenvalue λ_c, will approach 0. Therefore, the displacement of the FP with input is dominated by the corresponding mode which is not necessarily in the direction of the disinhibitory pathway. However, this critical mode becomes aligned with the disinhibitory pathway if the ratio between the intra- and interconnection weights is $\rho > 1$, see Fig. 4(b), which is not an unreasonable assumption as the connections in the MGC are thought to be genetically specified.

Mean field approximation.—We consider (5) for the stimulated s_i^+ and unstimulated populations s_i^-. For sufficiently large N_+, N_- one can approximate $\sum g_{ij}s_j^+ \approx P\kappa N_+\langle s_i^+\rangle$ (and similarly for all other terms) and average over all neurons i to obtain

$$\begin{aligned}\dot{\sigma}^+ &= -\beta\sigma^+ + \Phi[-\kappa P(\{N_+ - 1\}\sigma^+ + \rho N_-\sigma^-) + I]\\ \dot{\sigma}^- &= -\beta\sigma^- + \Phi[-\kappa P(\rho N_+\sigma^+ + \{N_- - 1\}\sigma^-)],\end{aligned} \tag{10}$$

where $\sigma^+ = \langle s_i^+ - s_i^{+*(0)}\rangle$, and similarly for σ^-, are the average displacements of the variables from the FP value $s_i^{+*(0)}$ in the absence of input. We have also replaced θ_i using (7). The function $\Phi(u)$ describes how the average response of the network changes as more unstimulated nodes are driven into silence; using [17] we find $\Phi[u] = \gamma_c u$ for $u \geq a$, $\Phi[u] = \frac{\gamma_c}{(b-a)}(bu + \frac{u^2+b^2}{2}) - \beta s^{+*(0)}$ for $a > u \geq b$ and $\Phi[u] = -\beta s^{+*(0)}$ otherwise. Here, $a = -\frac{\beta s_{\min}^{+*}}{\gamma}$ and $b = -\frac{\beta s_{\max}^{+*}}{\gamma}$ depend on the distribution of s_i at the FP. Phase plane analysis reveals that this system undergoes a supercritical pitchfork bifurcation as the FP becomes unstable. Note that the bifurcation results from the competition between inhibitory populations rather than from a transition to self-sustained activity described in excitable media, e.g., [7] and that, therefore, the total activity decreases when approaching the critical point from below.

Dynamic range and sensitivity.—To quantify the response of this network we define the DR as $\Delta = 10\log(\frac{I_{\max}}{I_{\min}})$ (measured in dB), where $I_{\min}$ is the input leading to a response of $0.05\sigma_\infty$, $I_{\max}$ to $0.95\sigma_\infty$ and σ_∞ is the deviation from the FP for infinite input, Fig. 4(a) (inset). The response of the unstimulated nodes exhibits maximal DR as the network approaches bifurcation. In a sample of 20 conductance based model networks both the moderately coupled system ($p_\lambda = 0.5$) and a purely feedforward system had mean DR 10 ± 6 dB. A system close to bifurcation ($p_\lambda = 0.995$) had DR of about 22 ± 6 dB (Fig. 3). When $p_\lambda > 1$ the system is displaced from the now unstable FP, even in absence of input, and the magnitude of the response and the DR are smaller.

While the rate approximation quantitatively captures the dynamics for $p_\lambda < 0.9$ (Figs. 2 and 3) deviations of the approximation from the full model become magnified as the system approaches bifurcation ($p_\lambda \to 1$). Consequently, we cannot accurately calculate the critical weight strengths using the approximation and the prediction of the location of the FP diverges as $p_\lambda \to 1$, potentially before we can exploit the dynamical divergence

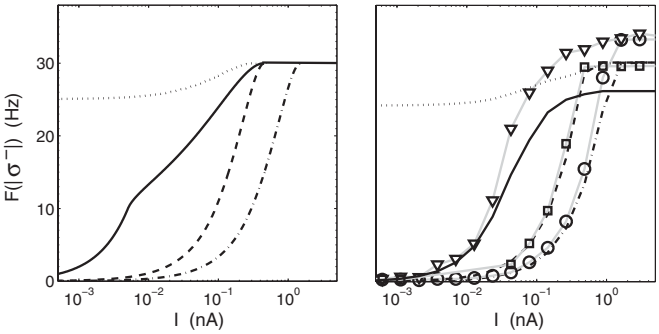

FIG. 3. Average reduction of frequency of the unstimulated population (LN^-) in response to 2s of noisy input (mean I, std $\sigma = 0.01$ nA). Left: asymptotic equilibrium of the mean field approximation for $p_\lambda = 1.1$ (supercritical) (dotted), 0.995 (solid), 0.5 (dashed), and without feedback (dot-dashed). Right: Rate model for $p_\lambda = 1.1$ (supercritical) (dotted), $p_\lambda = 0.995$ (solid) and 0.5 (dashed) and without feedback (dot-dashed) and data from the full conductance model for $p_\lambda = 0.5$ (squares), without feedback (circles) and close to critical (triangles). Note that the asymptotic and the numerically observed dynamic range differ because the system does not reach the fixed point in finite time.

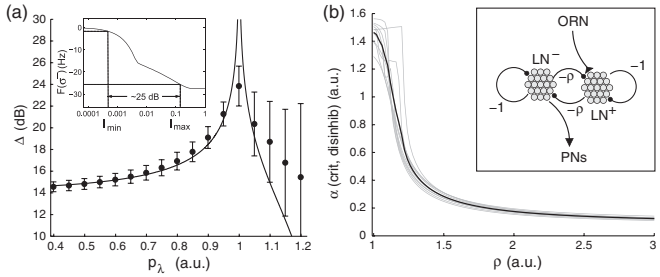

FIG. 4. (a) Dynamic range versus p_λ calculated from the nullclines of (10) (line) and for the full rate network (symbols; mean and std from 20 different connectivity matrices). Inset: average displacement of the FP by input used to obtain the dynamic range in the mean field model. (b) Angle between the critical mode and the disinhibitory mode as a function of the ratio between the inter- and intraconnection weights ρ for 10 individual connectivities (gray). The black line shows the average. Inset: A network schematic.

we are interested in. We can overcome these limitations by numerically approaching the critical point in the conductance based model while concurrently up-regulating the firing frequency of neurons that fall below 75% of their target firing frequencies (triangles in Fig. 3).

To understand the sensitivity of the system further we can analytically calculate the asymptotic gain of the network in the MF description. Using the nullclines of (10) we find the gain $\frac{d\sigma^{-*}}{dI} = \frac{\gamma_c^2 \kappa P \rho N_+}{\lambda_1 \lambda_2}$, where λ_i are the eigenvalues of the Jacobian of (10) at the FP. As we scale the system to the bifurcation, λ_1 approaches zero and the gain of the system diverges. In contrast, the gain without feedback, $\frac{d\sigma^{-*}}{dI} = -\frac{\gamma_c^2 \kappa P \rho N_+}{\beta^2}$, does not diverge. It is straightforward to construct an analytical expression for the DR from the nullclines of (10). The DR peaks at the critical point [Fig. 4(a)]. The DR estimation is less reliable for $p_\lambda > 1$ because the system often is oscillatory and the response to input is quite weak and irregular.

Discussion.—We have shown how the DR, mediated by a single disinhibitory mode, is maximized when the system is close to bifurcation. This result is largely independent of the model details and could straightforwardly be extended to dominant modes in networks containing both excitatory and inhibitory neurons. This suggests an idea of criticality in neural systems that differs from the more common statistical description and, at the network level, is more related to the work of [18], or, at the cellular level, to work on the vertebrate cochlea [19,20]. The bifurcation in an inhibitory network we investigated here is qualitatively different to the phase transitions to self-sustained activity examined in excitable media and our model makes the prediction that neural substrates could be taken into a supercritical state ($p_\lambda > 1$) by increasing inhibition rather than increasing excitation [21].

We are grateful to L. Barnett, A. Barrett, and R. Huerta for helpful discussions. This work was funded by the Biotechnology and Biological Sciences Research Council (Grant No. BB/F005113/1).

*C.L.Buckley@sussex.ac.uk
†T.Nowotny@sussex.ac.uk

[1] S. S. Stevens, *Psychophysics: Introduction to Its Perceptual, Neural and Social Prospects* (Wiley, New York, 1975).
[2] P. Kloppenburg and A. R. Mercer, Annu. Rev. Entomol. **53**, 179 (2008).
[3] D. Schneider, in *The Neurosciences Second Study Program* (Rockefeller Univ. Press, New York, 1970), pp. 511–518.
[4] D. Jarriault, C. Gadenne, P. Lucas, J. P. Rospars, and S. Anton, Chem. Senses **35**, 705 (2010).
[5] A. Levina, J. M. Herrmann, and T. Geisel, Phys. Rev. Lett. **102**, 118110 (2009).
[6] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. A **38**, 364 (1988).
[7] M. Copelli, A. C. Roque, R. F. Oliveira, and O. Kinouchi, Phys. Rev. E **65**, 060901 (2002).
[8] O. Kinouchi and M. Copelli, Nature Phys. **2**, 348 (2006).
[9] H. Haken, *Synergetics* (Springer, Berlin, 1983).
[10] R. D. Traub and R. Miles, *Neural Networks of the Hippocampus* (Cambridge University Press, New York, 1991).
[11] T. Nowotny and M. I. Rabinovich, Phys. Rev. Lett. **98**, 128106 (2007).
[12] A. Destexhe, Z. F. Mainen, and T. J. Sejnowski, J. Comput. Neurosci. **1**, 195 (1994).
[13] B. Ermentrout, Neural Comput. **10**, 1721 (1998).
[14] If the eigenvalues of a matrix A are λ then the eigenvalues of A-βI are λ_i-β and we have defined $\lambda_{\max}(A)$ to denote the maximum real eigenvalue of the matrix A.
[15] H. Jaeger, M. Lukosevicius, D. Popovici, and U. Siewart, Neural Networks **20**, 335 (2007).
[16] C. M. Root, J. L. Semmelhack, A. M. Wong, J. Flores, and J. W. Wang, Proc. Natl. Acad. Sci. U.S.A. **104**, 11 826 (2007).
[17] Consider a system of N independent variables $y_i = [-I + C_i]_+$, where $C_i \in [a, b]$. Then, given I, the probability $P_I(\frac{dy}{dI} = -1) = P_I(I < C_i) = \frac{b-I}{b-a} \equiv P_d$, and setting $Y = \langle y_i \rangle$ we have $\mathbb{E}_I(\frac{dY}{dI}) = -P_d$. In the limit $N \to \infty$ the variance becomes small and to good approximation $\frac{dY}{dI} \approx \frac{-(b-I)}{(b-a)}$. Integrating both sides with condition $Y(b) = 0$ we obtain $Y \approx -\frac{1}{b-a}(bI - \frac{I^2}{2} - \frac{b^2}{2})$.
[18] N. Bertschinger and T. Natschläger, Neural Comput. **16**, 1413 (2004).
[19] S. Camalet, T. Duke, F. Jülicher, and J. Prost, Proc. Natl. Acad. Sci. U.S.A. **97**, 3183 (2000).
[20] A. J. Hudspeth, Neuron **59**, 530 (2008).
[21] W. L. Shew, H. Yang, T. Petermann, R. Roy, and D. Plenz, J. Neurosci. **29**, 15595 (2009).