Input-Dependent Wave Attenuation in a Critically-Balanced Model of Cortex

Xiao-Hu Yan, Marcelo O. Magnasco*

Laboratory of Mathematical Physics, Rockefeller University, New York, New York, United States of America

Abstract

A number of studies have suggested that many properties of brain activity can be understood in terms of critical systems. However it is still not known how the long-range susceptibilities characteristic of criticality arise in the living brain from its local connectivity structures. Here we prove that a dynamically critically-poised model of cortex acquires an infinitely-long ranged susceptibility in the absence of input. When an input is presented, the susceptibility attenuates exponentially as a function of distance, with an increasing spatial attenuation constant (i.e., decreasing range) the larger the input. This is in direct agreement with recent results that show that waves of local field potential activity evoked by single spikes in primary visual cortex of cat and macaque attenuate with a characteristic length that also increases with decreasing contrast of the visual stimulus. A susceptibility that changes spatial range with input strength can be thought to implement an input-dependent spatial integration: when the input is large, no additional evidence is needed in addition to the local input; when the input is weak, evidence needs to be integrated over a larger spatial domain to achieve a decision. Such input-strength-dependent strategies have been demonstrated in visual processing. Our results suggest that input-strength dependent spatial integration may be a natural feature of a critically-balanced cortical network.

Introduction

The brain analyses sensory input under a wide range of input regimes, and employs a wide variety of strategies to extract information in diverse and challenging environments. A number of strategies employed in the primary sensors, such as gain control [1] or light-level-dependent receptive field shapes [2], have been intensively studied and shown to be optimal computational strategies to make the best use of the signal-to-noise available [3]. Less is understood at the level of the primary sensory cortices, such as strategies that change the scale of spatial integration according to input strength [4–6]. A recent study has shown [7] that effective lateral interactions in the primary visual cortex of cats and monkey appear to be modulated by the contrast level in the input. At high stimulus contrasts, small amount of lateral interactions are measured. As the contrast is lowered, waves propagate further away from their place of origin, until at zero contrast spontaneous waves of activity appear to decay extremely slowly with distance to their epicenter. This behavior has been interpreted as gating of the functional connectivity by the input.

We offer a simple model which quantitatively reproduces this phenomenology with minimal assumptions. The central hypothesis is that interactions between neurons are adjusted in such a way that excitation and inhibition are precisely balanced, not just globally, but in local patches. By construction our model is poised at a massive, high-dimensional Hopf bifurcation. A recent model of self-tuned critical neural nets has shown how anti-Hebbian interactions lead to such dynamical states and maintain them, and explained their

relationship to self-organized criticality [8]; in our construction, we shall forgo the dynamical poising mechanisms and assume the system is already poised at such a state, to analyze the dynamical consequences of such poising. The second postulate is that interactions are local within the two-dimensional surface of our abstract cortex, and that nonlinearities are local, in fact confined to individual neurons. From these postulates we shall show it follows that the system possesses infinitely-long-ranged susceptibilities at zero input, and that the range of the susceptibility, measured by its exponential decay length, decreases with increasing input.

The ease with which the brain changes tasks and activities, adapts to new sensory regimes or environments, and with which different areas communicate has long suggested to researchers that some features of brain function, and in fact of many biological systems, can be fruitfully studied in the context of critical systems [9]. Such systems have hitherto come in two varieties. *Statistical criticality* refers to systems related to statistical-mechanics models of phase transitions at a critical point [10,11]; the system may require tuning of an external parameter [12] (a temperature or magnetic field in the original models), or may be *self-organized*, meaning the system spontaneously operates in a regime giving rise to wide (usually power-law) distributions, such as avalanche behavior [13–15] and long-range susceptibilities [16,17].

A related but different concept, based on the notion of a dynamical bifurcation, is *dynamical criticality*, broadly inspired in the classic works of Gold [18] and Ashby [19]. These systems are poised at the onset of dynamical instabilities such as a Hopf bifurcation [20–23], or create an otherwise marginal configuration

such as a line attractor [24–26]. It again comes in two varieties, one in which a parameter has to be adjusted to achieve poising, and *self-tuned criticality* [21], in which the system spontaneously poises itself at the dynamical transitions [27] Dynamical criticality shows a number of the hallmarks of statistical criticality, including avalanche-like bursts of activity [8]. We show the presence of another hallmark: an infinitely long-ranged susceptibility. But this susceptibility does not take an arbitrary form; as we shall show, this susceptibility emerges in the absence of an input to the system, while the presence of an input dampens the responses to finite range. This specific signature relates now dynamical criticality also to input-dependent spatial integration.

Analysis

For a typical differential equation, the characteristic physical scales of its solutions are set by constant parameters that can be read-off directly from the differential equation proper; for example, the decay lengthscale in a production-diffusion-degradation model is the square root of the diffusion coefficient divided by the degradation rate, and is therefore a constant. Even if there was a nonlinear diffusion or nonlinear degradation, the lengthscale would still be the quotient of the diffusion coefficient evaluated at low concentrations over the degradation rate evaluated at low concentrations. However, when an equation is poised at the onset of a dynamical transition, the linearized equation around zero does not fully determine the solution, and higher-order terms need to be employed. As a result, the parameters in question are set by the local values of the equation's terms evaluated at the solution. For example, consider a system at the onset of a Hopf bifurcation, a model widely employed to describe cochlear dynamics [20,22,28]:

$$\dot{z} = i\omega_0 z - |z|^2 z + I(t) \tag{1}$$

If a periodic input $I(t) = Fe^{i\omega_0 t}$ at the resonant frequency ω_0 is presented, the solution $z(t)$ undergoes a transient until it settles into the steady-state solution $z = F^{1/3} e^{i\omega_0 t}$.

We now show that the timescale for relaxation of this transient is not set by the coefficients in the equation, but rather by the magnitude of the input. Consider the solution $z(t) = (\delta + F^{1/3})e^{i\omega_0 t}$, where δ is a distance to the steady-state solution, i.e. the transient proper [29]. Substituting this ansatz into the equation above and canceling terms shows δ satisfies the equation

$$\dot{\delta} = -3F^{2/3}\delta \tag{2}$$

and hence the relaxation timescale is a function of the input amplitude F rather than depending only on the internal parameters of the equation. In general, whenever physical scales such as timescales or lengthscales are sensitive functions of the input to the system, and vary with varying input, we may suspect the system to be poised near a dynamical instability conferring on the solution a great deal of sensitivity to the input's characteristics.

In our abstract model neurons interact with one another through a matrix A_{ij} that is assumed to have purely imaginary eigenvalues, and all other interactions are mediated by higher-order terms in the normal form. The equation is

$$\dot{x}_i = \sum_j A_{ij} x_j - x_i^3 + I_i(t) \tag{3}$$

where $I(t)$ is an input. A fuller treatment would include also an internal noise term $\xi(t)$ but we shall not use it here.

The simplest such matrix A_{ij} is shown in Figure 1: neurons are arranged on the sites of a chessboard and interact with their 4 nearest neighbors; neurons are declared excitatory if they are on the white squares of the chessboard, inhibitory in the black squares, and the strength of all excitation and inhibition is identical. The interaction matrix A_{ij} therefore is antisymmetric and has purely imaginary eigenvalues.

It is manifest that equation (3) is a high-dimensional version of equation (1). However care should be exercised in noting that eq (3), unlike (1) is not a normal form for the Hopf bifurcation, because the structure of the nonlinearities in this equation is not generic for a high-dimensional space: the nonlinear terms are local to each of the neurons and do not couple the internal states of different neurons. While we consider this assumption to be reasonable in our context (we do not expect faraway neurons to interact strongly by means other than through their synaptic connections), it is mathematically quite strong.

Results

We shall now show that equation (3), when equipped with the matrix described in Fig 1, has solutions whose behavior is similar to that experimentally shown in [7] for spike-triggered LFPs in cat and monkey visual cortex.

In the absence of input, eq (3) supports propagating waves which attenuate algebraically, as shown in Figure 2 by direct numerical integration of Eq. (3). This algebraic decay is central to our argument, and important in tying together the dynamically critical models with the more widely studied self-organized critical models, so we shall discuss its origin in a bit more detail. Preserving only the linear part of (3) and iterating it twice to obtain $\dot{x} = A^2 x$, we note the matrix A^2 decouples the excitatory and inhibitory sublattices. Concentrating on the excitatory sites, the matrix couples each location in space to the nearby excitatory site in a pattern

$$\begin{pmatrix} 0 & * & -1 & * & 0 \\ & -2 & * & -2 & * \\ -1 & * & -4 & * & -1 \\ & -2 & * & -2 & * \\ 0 & * & -1 & * & 0 \end{pmatrix}$$

where the -4 is the self-coupling to the center site. These connections can be seen as a discrete implementation of $-\nabla^2 - 16$, where ∇^2 is the Laplacian operator, and suggests we should take a passage to the continuum limit. We shall need to rename our original activities x as ϕ, to reuse (in an admitted abuse of notation) the x as our spatial variables now, since we need to make space explicit to be able to use ∇^2. Thus we map $x_i(t) \rightarrow \phi(x,t)$ Taking the standard passage to the limit entails taking smaller and smaller lattices, of lattice width δ, while scaling the original synaptic strength s so that the quotient $c = s\delta$ remains finite, and leads to the equation

$$\partial_t^2 \phi + 4c^2 \nabla^2 \phi + 16\phi = I(t) \tag{4}$$

where c has units of space over time, i.e. a speed. This equation, a Helmholtz equation in euclidean space-time [30], has as a Green function
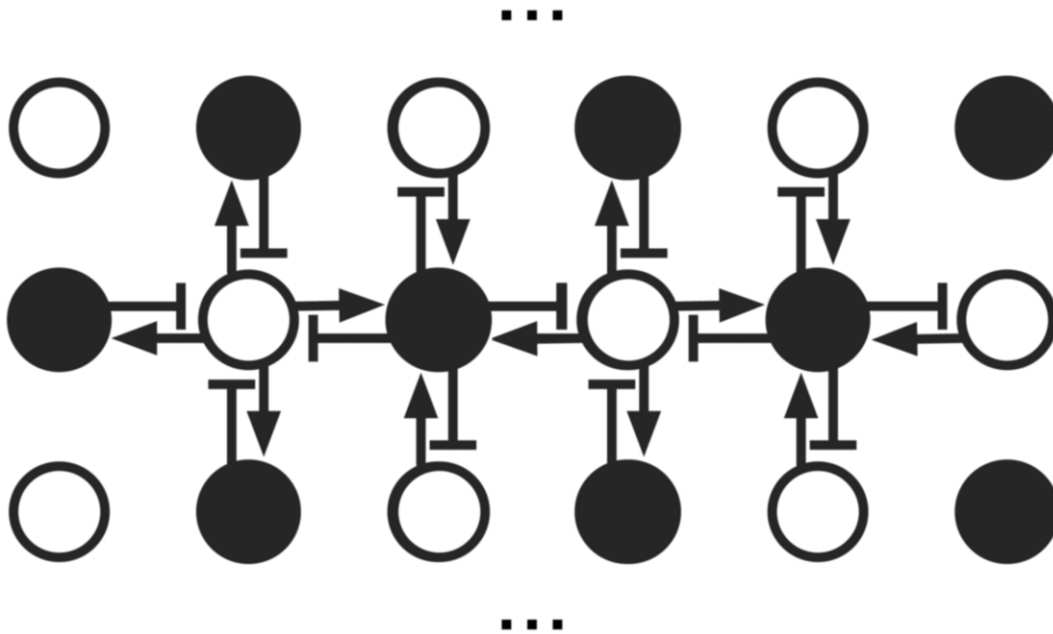
Figure 1. The simplest connectivity of a critically-balanced model of cortex. Black units are inhibitory, white units are excitatory, and all synaptic strengths are identical. Considering a lattice of edge E, and the site at column c and row c gets index $i = c + Er$, then the matrix is $A_{ij} = (-)^{p(i)} s(\delta_{ii\pm1} + \delta_{ii\pm E})$ where $p(i)$ is the parity of site i; assuming i is even, $p(i) = (i \bmod E - \mathrm{int}(i/E))$.
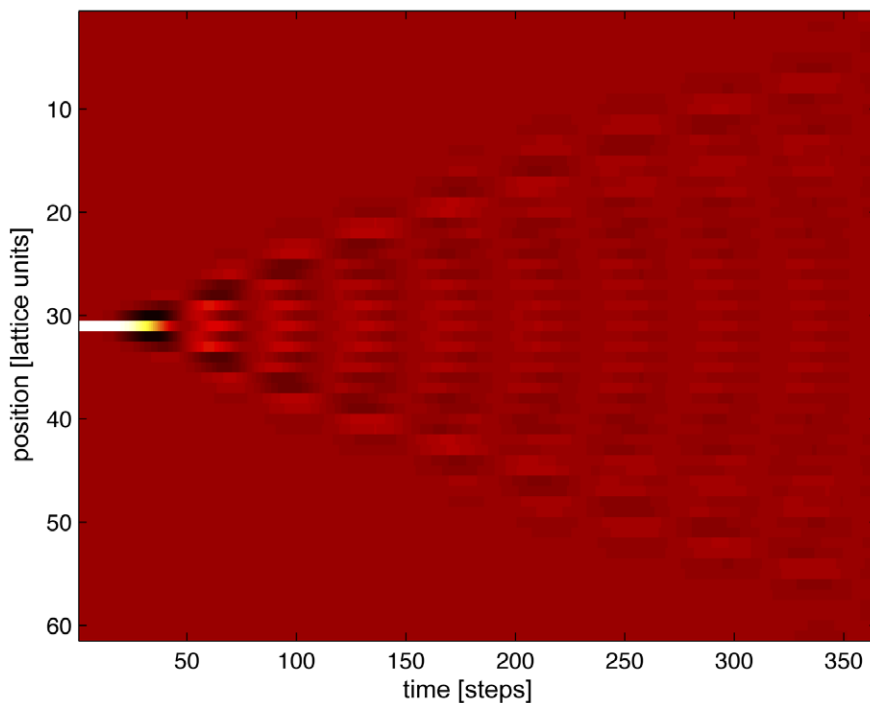doi:10.1371/journal.pone.0041419.g001

Figure 2. Numerical evolution of equation (1). The activity in a 1D slice of the 2D array (the central row) is displayed. Horizontal axis, time; vertical axis, horizontal coordinate in the array. An initial perturbation at time 0 expands outwards at asymptotically constant speed, with characteristic ripples looking like a ''wave'' which attenuates algebraically as it goes out.
doi:10.1371/journal.pone.0041419.g002

$$G(y) = \frac{e^{ik|y|}}{|y|} \qquad y = \left(\frac{x}{2c}, t\right)$$

which indeed looks like the propagating waves shown in Figure 2. Care should be taken in referring to these Green functions as waves, because *they are not the solution of a wave equation* — the equation above is elliptic, not hyperbolic. Pictorially, the equation describes the eigenmodes of a 3D membrane whose coordinates are space (2D) and time (1D). If we look at the solution for one given time, one observes a cut through this shape, which displays peaks and valleys. As one moves forward in time, the valleys and peaks appear to be moving, in what seems to be wavelike fashion. To some extent it is profitable to describe these motions in terms of waves, even though from a strict perspective they are not, because they are not the solution to a wave equation. From an alternative viewpoint we can see the matrix in equation (3), (or alternative the differential operator in equation (5)) has eigenmodes whose spatial frequency k is related to the imaginary component of the corresponding eigenvalue (its temporal frequency ω) through an equation of the form $c^2 k^2 + \omega^2 = 1$, called the *dispersion relation* of the system. Thus ω and k lie on a circle and are inversely related to one another, rather than lying on a line (directly related to one another) which is the dispersion relation of a wave equation. While the system does show a dispersion relation, and to that extent we can describe its behavior as waves, its dispersion relation goes against the normal one for wave equations, with increasing spatial frequency for decreasing temporal frequency.

Returning to the case in which there is a nonzero input, we need to distinguish between the ongoing activity $x(t)$ and the extra activity $x + \delta x$ that would be observed elsewhere due to a perturbation at a given site. In the experiment described in [7] the spike-triggered averaging of the LFP is done so as to extract the extra activity that follows a given neural spike at a particular neuron. Since our model is continuous and not formulated in terms of spikes, we shall use a perturbative formulation as the analysis most close in spirit to that experimental measurement. A small perturbation $\delta x(t)$ riding on top of a pre-existing activity level $x(t)$ (i.e. a Floquet-Lyapounov analysis [29]) satisfies

$$\frac{\partial(x + \delta x)_i}{\partial t} = \sum_j A_{ij}(x + \delta x)_j - (x + \delta x)_i^3 + I(t)$$

which to lowest order in δx will see the expectation value of x_i^2 as a damping coefficient (in time):

$$\delta \dot{x} \approx \sum_j A_{ij} \delta x_j - 3 \langle x_i^2 \rangle \delta x_i \qquad (5)$$

where the $\langle \cdots \rangle$ denote an average or expectation value. As the amplitude of stimulation is increased we observe that the waves attenuate exponentially with distance, rather than algebraically, because the variance of the ongoing activity $\langle x_i^2 \rangle$ enters the equation for the perturbation as an effective damping coefficient.

Exactly how the variance $\langle x_i^2 \rangle$ of the response changes with the input's amplitude will be dependent on the spatiotemporal structure of the input $I_i(t)$; in what follows we shall assume the input to be (or have properties essentially equivalent to) spatiotemporal Gaussian white noise of amplitude $\in$:

$$I_i(t) = \in \eta_i(t) \qquad \langle \eta_i(t)\eta_j(s) \rangle = \delta_{ij}\delta(t - s)$$

and we will assume that the perturbation $\in$ is confined, at time zero, to a single site i_0: $\delta x(t=0) = \in_0 \delta_{ii_0}$. Henceforth we shall describe the wavelike properties of δx, the perturbation riding on top of the ongoing activity $x_i(t)$ caused by $I_i(t)$

At short times δx initially spreads just like in Figure 2. As the "wave" propagates outwards at constant speed, the temporal damping of δx in equation (4) manifests as a spatial damping, so that at large distances the "wave" is exponentially attenuated with distance because it takes a time $\mathcal{O}(distance)$ to reach there. This effect is shown in Figure 3, where numerical solutions are shown of the perturbation δx riding on top of a solution $x(t)$ in response to input of specified amplitude. More specifically, Figure 3a shows the maximum value $m(d)$ that δx acquires at any given time for times $t > 0$, for a given specified distance d from the epicenter of the perturbation as a function of distance d:

$$m(d) = \max_{t > 0} \max_{|i - i_0| = d} |\delta x_i(t)|$$

and Figure 3b shows the range λ, the reciprocal of the decay constant β of $m(d)$, as a function of the input's strength $\in$:

$$m(d) \approx \exp(-\beta d) \equiv \exp\left(-\frac{d}{\lambda}\right) \qquad d \gg 1$$

Discussion

Recent work [7] addressed an ancient dispute in visual neurophysiology: whether visual cortex is primarily driven by its inputs, or whether it is primarily driven by lateral recurrent interactions within cortex, with a relatively small nudge from the inputs. In support of the first view, substantial work from many labs had shown that the responses to visual stimulation in cortex do not change substantially if cortex is inactivated—if it is inhibited from firing through pharmacological or electromagnetic means. Supporting the second view, similarly substantial work from many labs had shown that the strength of intracortical inhibition and excitation is much larger than the overall input from thalamus—it is the nearly perfect balance between excitation and inhibition which permits thalamus to sway activity one way or the other; in fact, in the absence of stimulation cortex activates spontaneously in patterns which agree with those that would have been evoked by pieces of imagery—it literally hallucinates little lines [31]. The results in [7] show that both views are correct at different strengths of the driving input, in this case the visual contrast: at high contrast there is little lateral interaction, while at low or null contrast lateral interactions dominate.

We have demonstrated an extremely compact model reproducing the results of [7]. Our model has but one strong assumption: precise balancing of excitation and inhibition. It offers over the horizontal connection gating model three advantages. First, there is no gating of connectivity at all in our model, as the model has constant parameters without nonlinear interactions; all nonlinearities are strictly local. Gating could be tricky to implement because it cannot be a function of the output of cortex, but rather only of its input, requiring thus that thalamic input to cortex be duplicated: one copy to drive the cortical neurons, the other copy to gate the lateral connections between neurons. Second, the attenuation length is naturally a graded function of the magnitude of the input. In a gated connectivity model it would be extremely
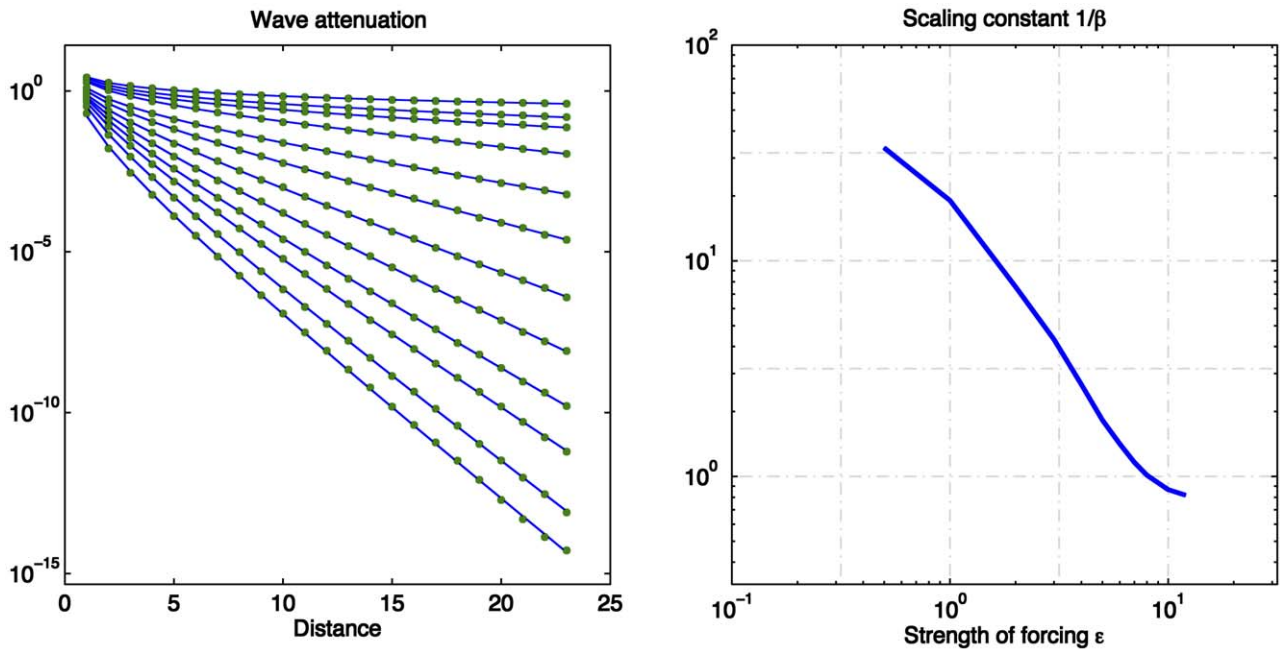
Figure 3. wave attenuation depends directly on the intensity of the input. Left panel, the spatial profile of the wave, for different values of the input intensity; the steeper curves at the bottom correspond to larger intensity, while the shallow and curved solution at the top corresponds to zero input. Right, a fit to the slope of the rightmost points of the left panel, plotted as 1/slope (i.e. attenuation length) as a function of the forcing strength.
doi:10.1371/journal.pone.0041419.g003

hard to arrange the gating to be such a graded function over a wide range of input contrasts. Third, the attenuation length, in the absence of input, is naturally infinite because the balanced cortex produces algebraically-decaying waves. In a functional gating scenario, signal propagation in the absence of input still could either grow or shrink exponentially; to have an algebraic propagation already presupposes balancing internal parameters precisely, which is mostly our only assumption.

Future work on this model should address whether the predicted propagation speed of the outward wave matches the physiologically observed $0.3m/s$, even when taking into account physiologically-relevant synaptic delays. Understanding this will require further elucidation of whether such waves "take shortcuts" through long-range connections, and what their effective speed is when propagating through a tissue with a large spread in connectivity distances. If the mechanism we propose indeed

underlies changing integration scale with input intensity, then further work should also attempt to elucidate functional dependences of the local activity level (and thus the attenuation length) with the input, and relate these predictions, measured in distances on the cortex, to the appropriate sensoritopic scales, such as visuotopic units.

Author Contributions

Conceived and designed the experiments: XHY MOM. Performed the experiments: XHY MOM. Analyzed the data: XHY MOM. Wrote the paper: XHY MOM.

References

1. Schwartz O, Simoncelli EP (2001), Natural signal statistics and sensory gain control, Nature Neuroscience 4, 819–825 doi:10.1038/90526.
2. Atick J, Redlich AN (1990), Towards a Theory of Early Visual Processing, Neural Computation 2.3 308–320, doi:10.1162/neco.1990.2.3.308.
3. Rieke F, Warland DA, de Ruyter van Steveninck R, Bialek W (1999) Spikes. Bradford, ISBN 0262681080.
4. Levitt JB, Lund JS (1997) Contrast dependence of contextual effects in primate visual cortex. Nature 387, 73–76.
5. Sceniak MP, Ringach DL, Hawken MJ, Shapley R (1999) Contrast's effect on spatial summation by macaque V1 neurons. Nature Neuroscience 2, 733–739 doi:10.1038/11197.
6. Kapadia MK, Westheimer G, Gilbert CD (1999). Dynamics of spatial summation in primary visual cortex of alert monkeys. Proc. Natl. Acad. Sci. USA, 96.21, 12073–12078.
7. Nauhaus I, Busse L, Carandini M, Ringach DL (2009) Stimulus contrast modulates functional connectivity in visual cortex. Nat. Neurosci 12.1 70–76.
8. Magnasco MO, Piro O, Cecchi GA, (2009). Self-tuned critical anti-Hebbian networks, Phys. Rev. Lett. 102 258102.
9. Mora T, Bialek W (2011). Are Biological Systems Poised at Criticality? J. Stat. Phys. 144.2 268–302.

10. da Silva L, Papa ARR, de Souza AMC (1998), Criticality in a simple model for brain functioning. Phys. Lett. A 242.6, 343–348.
11. Fraiman D, Balenzuela P, Foss J, Chialvo DR (2009) Ising-like dynamics in large-scale functional brain networks, Phys. Rev. E 79, 061922.
12. Schneidman E, Berry II MJ, Segev R, Bialek W (2006), Weak pairwise correlations imply strongly correlated network states Nature 440, 1007.
13. Beggs JM, Plenz D (2003) Neuronal avalanches in neocortical circuits. J Neurosci 23: 11167–11177.
14. Levina A, Herrmann JM, Geisel T (2007), Dynamical synapses causing self-organized criticality in neural networks. Nature Physics 3, 857.
15. Gireesh ED, Plenz D (2008), Neuronal avalanches organize as nested theta- and beta/gamma-oscillations during development of cortical layer 2/3 (2008) Proc. Natl. Acad. Sci. USA 105, 7576–7581.
16. Eguíluz VM, Chialvo DR, Cecchi GA, Baliki M, Apkarian AV (2005). Scale-Free Brain Functional Networks. Phys. Rev. Lett. 94, 018102.
17. Kitzbichler MG, Smith ML, Christensen SR, Bullmore E (2009) Broadband Criticality of Human Brain Network Synchronization. PLoS Comput Biol 5(3): e1000314. doi:10.1371/journal.pcbi.1000314.
18. Gold T (1948) Hearing. II. The Physical Basis of the Action of the Cochlea. Proc. R. Soc. Lond. B 135 no. 881 492–498.

19. Ashby WR (1948). Design for a Brain. Electronic Engineering, 20: 379–83.
20. Choe Y, Magnasco MO, Hudspeth AJ (1998) A model for amplification of hair-bundle motion by cyclical binding of Ca2+ to mechanoelectrical-transduction channels, Proc. Natl. Acad. Sci. USA 95.26 15321–15326.
21. Camalet S, Duke T, Jlicher F, Prost J (2000). Auditory sensitivity provided by self-tuned critical oscillations of hair cells. Proc Natl Acad Sci USA 97: 3183–3188.
22. Eguiluz VM, Ospeck M, Choe Y, Hudspeth AJ, et al.(2000) Essential nonlinearities in hearing, Phys. Rev. Lett. 84.22 5232–5236.
23. Moreau L, Sontag E (2003). Balancing at the border of instability. Physical review 68: 020901.
24. Seung HS (1998). Continuous attractors and oculomotor control. Neural Networks 11, 1253–1258 (1998).

25. Seung HS, Lee DD, Reis BY, Tank DW (2000) Stability of the memory of eye position in a recurrent network of conductance-based model neurons. Neuron, 26, 259–271.
26. Machens CK, Romo R, Brody C (2005) Flexible Control of Mutual Inhibition: A Neural Model of Two-Interval Discrimination. Science 307, 1121–1124.
27. Bienenstock E, Lehmann D (1999) Regulated Criticality. Advances in Complex Systems 1.4 361–384.
28. Duke T, Jlicher F (2003) Active Traveling Wave in the Cochlea, Phys. Rev. Lett. 90, 158101.
29. Guckenheimer J, Holmes P (1997) Nonlinear oscillations, dynamical systems, and bifurcations of vector fields. Springer Verlag ISBN 0387908196.
30. Pikulin VP, Pohozaev SI, Iacob A (2012) Equations in Mathematical Physics: A Practical Course. Springer Verlag ISBN 3034802676.
31. Kenet T, Bibitchkov D, Tsodyks M, Grinvald A, Arieli A (2003) Spontaneously emerging cortical representations of visual attributes. Nature 425, 954–956.