Glassy phase in dynamically-balanced neuronal networks

Kevin Berlemont
Center for Neural Science, New York University, New York, NY 10003, USA

Gianluigi Mongillo
Sorbonne Université, INSERM, CNRS, Institut de la Vision, F-75012 Paris, France and Centre National de la Recherche Scientifique (CNRS), Paris, France

(Dated: March 14, 2022)

We present a novel mean-field theory for balanced neuronal networks with arbitrary levels of symmetry in the synaptic connectivity. The theory determines the fixed point of the network dynamics and the conditions for its stability. The fixed point becomes unstable by increasing the synaptic gain beyond a critical value that depends on the level of symmetry. Beyond this critical gain, for positive levels of symmetry, we find a previously unreported phase. In this phase, the dynamical landscape is dominated by a large number of marginally-stable fixed points. As a result, the network dynamics exhibit non-exponential relaxation and ergodicity is broken. We discuss the relevance of such a glassy phase for understanding dynamical and computational aspects of cortical operation.

Balanced networks are a key theoretical construct to describe the dynamical regime in which the cortex operates and to understand its computational implications [1]. In these networks, neuronal activity dynamically adjusts so that the average input to a neuron is of the same order as the spatial and/or temporal fluctuations [2–5]. This results in a dynamical state—the balanced regime—characterized by low and heterogeneous single-neuron firing rates, with temporally-irregular spiking [6–9] and vanishingly small pairwise correlations in the activity of different neurons [5, 10], consistent with experimental observations in the cortex.

The balanced regime is understood in random networks of binary neurons [3, 5, 10] and of rate-based neurons [11, 12], random meaning that there are no correlations in the connectivity matrix. However, local connectivity (i.e., on spatial scales of the order of 100 μm) in the cortex is not random. Reciprocal synaptic connections are both over-represented, as compared to random networks, and stronger than average [13–15]. The effects of such partial symmetry on the dynamics are not well understood [16–18], because the analytical description is difficult. In fact, fluctuations in the synaptic inputs are not Gaussian, due to the correlations in the connectivity. Currently, there is no mean-field description for partially-symmetric balanced networks.

In this letter, we show that the non-Gaussianity of the inputs can be conveniently dealt with by a cavity approximation [19], which leads to an exact mean-field theory when correlations are restricted to reciprocal connections. The theory determines the distribution of neuronal activity in the fixed point and its stability. When the mean-field solution is unstable, we study numerically the network dynamics. We find that the dynamics are not ergodic, provided the correlation in the connectivity is positive.

We consider a fully connected network of N inhibitory rate-based neurons. The state of neuron i is described by its synaptic input, $h_i$, which evolves in time (in units of the time constant) according to

$$\dot{h}_i + h_i = \sqrt{N} - \frac{1}{\sqrt{N}} \sum_{j=1}^{N} w_{ij} \nu_j,$$

where the first term on the r.h.s. describes an excitatory external input, $w_{ij} > 0$ is the efficacy of the synapse connecting neuron j to neuron i, and $\nu_j$, the rate of neuron j, is a function of its synaptic input, i.e., $\nu_j = \phi(h_j)$. We choose $\phi(h_j) \equiv h_j \Theta(h_j)$, where $\Theta(\cdot)$ is the Heaviside function. The $w_{ij}$ are given by

$$w_{ij} = e^{-\frac{1}{2} \frac{\sigma^2_{ij} + \sigma^2_{II}}{\rho_{w}}},$$

where $t_{ij}$ is normally distributed with mean zero and unit variance. Thus, the synaptic efficacies are log-normally distributed with mean $\langle w \rangle = 1$ and variance (synaptic gain) $\sigma^2_{w} = e^{\sigma^2_{t}} - 1$, which is tuned by adjusting $\sigma^2_{t}$. The correlation of reciprocal connections (level of symmetry), $\rho_{w}$, is tuned by adjusting the correlation $\rho_{w} = \langle t_{ij} t_{ji} \rangle$.

The model includes elements of biological realism lacking in previous studies [16, 17]: (i) the synaptic efficacies are sign-constrained, to comply with the so-called Dale’s law [20]; (ii) the single-neuron activity is described by a non-negative variable. The network, then, operates in the balanced regime for $N \to \infty$. The other modeling choices, such as the distribution of the $w_{ij}$ and the specific $\phi(\cdot)$, have been motivated by considerations of analytical/numerical tractability of the resulting equations. They are immaterial for the theory developed below.

We are interested in a mean-field description of the fixed point(s) of our model network (see [21] for a detailed derivation). For this, we need to determine the distribution of the synaptic inputs in a fixed point. We cannot use the central limit theorem to evaluate the sum in...
Eq. (1) because all $ν_j$ for $j \neq i$ depend on $ν_i$ and, hence, they are not independent. Moreover, when $ρ_w \neq 0$, $w_{ij}$ and $ν_j$ are also correlated because $ν_j$ depends on $w_{ji}, ν_i$, and $w_{ij}$ and $w_{ji}$ are correlated. To make progress, we assume that the $ν_j$ are correlated only through $ν_i$, and rewrite $h_i$ as

$$
h_i = \sqrt{N} - \frac{1}{\sqrt{N}} \sum_{j=1}^{N} w_{ij} ν_j - \frac{1}{\sqrt{N}} \sum_{j=1}^{N} w_{ij} δν_j \quad (3)
$$

$$
≡ u_i - \frac{1}{\sqrt{N}} \sum_{j=1}^{N} w_{ij} δν_j,
$$

where $ν_j$ is the rate of neuron $j$ when the outgoing connections from neuron $i$ are removed (i.e., $w_{ji} = 0$ for all $j$) and, hence, $δν_j$ is the change in the rate of neuron $j$ due to neuron $i$. By construction the $ν_j$ are now independent and the distribution of the cavity input, $u_i$, across realizations of the $w_{ij}$ becomes Gaussian in the limit $N \to \infty$. The corresponding mean, $μ$, and variance, $σ^2$, can be expressed in terms of the means and variances of the $w_{ij}$ and the $ν_j$. This is because the $w_{ij}$ and the $ν_j$ are independent by construction.

The second sum on the r.h.s. of Eq. (3) — the reaction term — can be computed with linear response theory, because the presence of neuron $i$ induces only a small (i.e., $O(1/\sqrt{N})$) perturbation in the input to neuron $j$. In the limit $N \to \infty$, we find

$$
-\frac{1}{\sqrt{N}} \sum_{j=1}^{N} w_{ij} δν_j = ρ_w σ_w^2 χ_j \equiv Rν_j \quad (4)
$$

with $χ \equiv N^{-1} \sum_j χ_{jj}$, where $χ_{jj}$ is the static local susceptibility of neuron $j$. Finally, we can determine self-consistently $h_i$ using Eqs. (3) and (4), i.e.,

$$
h_i = u_i + Rϕ(h_i) \quad \Rightarrow \quad h_i = ψ(u_i) \equiv \frac{u_i}{1 - RΘ(u_i)}. \quad (5)
$$

As anticipated, the distribution of the synaptic inputs in the fixed point is non-Gaussian when $ρ_w \neq 0$. However, Eq. (5) solves the problem. We can express $ν_i$ as a function of the cavity input $u_i$, i.e., $ν_i = ϕ(u_i) = φ(ψ(u_i))$. Using the effective transduction function, $ϕ(·)$, we then determine the mean and the variance of the distribution of the cavity inputs, as in a random network, by requiring self-consistency. Additionally, however, we impose self-consistency on $χ$, which determines in turn the effective transduction function self-consistently.

The mean-field theory is valid only if the fixed point is asymptotically stable, otherwise we cannot use linear response theory to compute the reaction term. The stability of the fixed point can be assessed by computing the leading order corrections to the mean-field theory, i.e., the $δν_j$ in Eq. (3). If these adjustments are non-physical, the mean-field solution, and hence the associated fixed point, is unstable [19]. Using $ϕ(·)$ we can write

$$
δν_j = ϕ(u_j) δu_j = χ_{jj} δu_j \quad (to \ leading \ order),
$$

where $δu_j$ is the change in the cavity input to neuron $j$ due to neuron $i$. The key simplification is that the $δu_j$ are Gaussian (by construction). Determining self-consistently the $δν_j$ reduces to a standard problem [21]. We find

$$
N⟨δν⟩ = -ν_i; \quad N⟨νδν⟩ = 0 \quad (6)
$$

$$
N⟨δν^2⟩ = \frac{σ_w^2 χ_{nl}}{1 - σ_w^2 χ_{nl}} ν_i^2 \equiv R_{nl} ν_i^2, \quad (7)
$$

where we have defined $χ_{nl} ≡ N^{-1} \sum_j χ_{jj}$. Clearly, $⟨δν^2⟩$ must be non-negative and finite, which requires $σ_w^2 χ_{nl} < 1$. This latter condition is satisfied whenever

$$
σ_w (1 + ρ_w) < \sqrt{2}. \quad (8)
$$

As long as the mean-field solution is stable, there is a unique, stable fixed point of Eq. (1) [21].

Next, we compare the theoretical predictions with the results of numerical simulations. We find a very good agreement as illustrated in Fig. 1 for the mean-field theory, and in Fig. 2 for the leading order corrections to the mean-field theory, i.e., Eqs. (6)-(7). As can be seen in Fig. 1d, the fraction of active neurons (i.e., with $ν_i > 0$) in the fixed point, $f$, decreases with increasing $ρ_w$. This can be understood as follows. The stability of the fixed point is controlled by the spectrum of the Jacobian, $J_{ij} = -δ_{ij} - Θ(h_j) w_{ij}/\sqrt{N}$, where $δ_{ij}$ is the Kronecker symbol. The average spectrum (i.e., over realizations of the synaptic matrix) is easily evaluated in the limit $N \to \infty$. There is one real, negative eigenvalue...
Figure 2. Comparison between the theoretical corrections to the mean-field theory (full lines) and the results of numerical simulations (symbols) at varying $\rho_w$, for $\sigma_w^2 = 1/2$. Error bars represent 95% confidence intervals. The inset in (c) shows the divergence of the non-linear response for $\delta\nu > 0$.

The results are illustrated in Fig. (3). In Fig. 3a, we plot the number of fixed points found as a function of the rescaled synaptic gain, $g_w = -1 + \sigma_w^2 (1 + \rho_w)^2 / 2$, for three values of $\rho_w > 0$. As predicted by the theory, different ICs lead to the same fixed point for $g_w < 0$, i.e., when the mean-field solution is stable (see Eq. (8)). Instead, when $g_w > 0$ (i.e., the mean-field solution is unstable), the network is multi-stable for $\rho_w \geq 0.5$. For $\rho_w < 0.5$ (and $g_w > 0$), the dynamics rarely converged to a fixed point within the allotted time (data not shown). Obviously, this does not locate the transition. There could be an exponential number of fixed points for $N \to \infty$ and yet none at a given finite $N$, depending on how fast the logarithm of the typical number of fixed points grows.

In Fig. 3b, we plot $f$ averaged over fixed points and connectivity matrices, $\tilde{f}$, as a function of $g_w$. Consistently with the stability bound Eq. (10), $\tilde{f}$ decreases with $g_w$. However, it is systematically above $f^*$ for $g_w \geq 0$, which is indicative of strong finite-size effects. In fact, $\tilde{f}$ decreases with $N$ at parity of $g_w$. Thus, we performed a finite-size scaling analysis. Interestingly, $\tilde{f}$ extrapolated at $N \to \infty$ is very close to $f^*$, suggesting that the (vast) majority of the fixed points found starting from random ICs are marginally stable for $N \to \infty$. This is confirmed by the finite-size scaling analysis of the spectral gap – the distance of the spectrum of the Jacobian from the imaginary axis – illustrated in Fig. 3c. There, we plot the spectral gap averaged over fixed points as a function $N$, for $\rho_w = 0.75$ and two values of $g_w$. For $g_w = -0.5$, the average spectral gap rapidly converges to a positive value.

By contrast, for $g_w = 2$ the average spectral gap keeps approaching 0. This decay is well fitted by $\alpha N^{-2/3}$ ($\alpha$ is a free parameter), as expected if the spectrum touches the imaginary axis for $N \to \infty$ [27]. The inset illustrates the variability of the spectral gap across fixed points, confirming that, indeed, most fixed points are at the edge of stability. For purpose of illustration, we plot in Fig. 3d the spectrum of the Jacobian for the fixed point found at $g_w = -0.5$ (top) and for the fixed points found at $g_w = 2$ (bottom). Note that $g_w = 2$ is far from the critical line.

These results are explained naturally if, for $\rho_w > 0$, the number of stable fixed points of Eq. (1) grows exponentially with $N$. At a given $\rho_w$, the marginally-stable fixed points are exponentially more than the asymptotically-stable ones. This explains why one always finds close-to-marginal fixed points when starting from random ICs. For the same $\rho_w$, however, there must be also an exponential number of asymptotically-stable fixed points.

This is needed to explain why the marginality of the dynamics is structurally stable against decreasing $\rho_w$. The number of stable fixed points decreases with decreasing $\rho_w$ until only unstable fixed points are left at $\rho_w = 0$ (and $\sigma_w^2 > 2$). This leads to the chaotic phase [26].

If the above scenario is correct, we expect broken ergodicity for any $\rho_w > 0$ via the same mechanisms.

The rationale for this choice should be clear shortly. The
Note, however, that one can no longer define an energy function to describe the dynamics when $\rho_w < 1$. Thus, there is no guarantee that the fixed points and their stability alone determine the dynamics of the network. Nevertheless, it seems plausible (to us) that if other attractors existed (e.g., limit cycles), they would have vanishing basins of attraction for $N \to \infty$ because, in the same limit, an exponential number of stable fixed points has to be accommodated in the phase space. The results we present now support this conjecture.

By definition, the network dynamics are ergodic if the average of the $h_i$ over a suitably long time interval are independent of the initial conditions. We define

$$D^{ab}(T) = \frac{1}{N} \sum_i \left[ \bar{h}^a_i(T) - \bar{h}^b_i(T) \right]^2,$$

where $\bar{h}^a_i(T)$ is the time-average of $h_i$ over a time interval $T$ when the dynamics Eq. (1) start from the IC $a$ (specifying $h^a_j(0)$ for all $j$). Thus, the dynamics are ergodic if $\lim_{T \to \infty} D^{ab}(T) = 0$ for all pairs $(ab)$ of different ICs. We numerically estimated the behavior of $D^{ab}(T)$. The results are shown in Fig. 4. When $g_w < 0$, $D^{ab}(T)$ rapidly goes to 0, regardless of $\rho_w$. When $g_w > 0$, $D^{ab}(T)$ goes to 0 for $\rho_w \leq 0$ (Fig. 4b), and to a finite value for $\rho_w = 1$ (Fig. 4c). Interestingly, for $0 < \rho_w < 1$, the decay of $D^{ab}(T)$ is so slow that $D^{ab}(2000) > 0$ for both $\rho_w = 0.26$ (Fig. 4d) and $\rho_w = 0.45$ (Fig. 4e). We have fitted $D^{ab}(T)$, averaged over $(ab)$, to a stretched exponential with an offset. As can be seen, the stretched exponential provides a good fit to the observed decay (black, dashed lines in Fig. 3d and e). In all cases, the offset was positive suggesting that the dynamics are indeed not ergodic.

In summary, we have shown that balanced networks of rate-based neurons exhibit three distinct phases. For low synaptic gain, there is a unique, stable fixed point regardless of the level of symmetry. The dynamics are chaotic and ergodic for random or anti-symmetric networks, while they are glassy for symmetric networks. In the glassy phase, settling in a fixed point becomes a very slow process and, accordingly, single neurons’ activity keeps fluctuating in time. From this point of view, the glassy phase resembles the chaotic one. The resemblance, however, is only superficial. Unlike in the chaotic phase, the memory of the initial conditions never fades away in the glassy phase.

Is there any evidence that the cortex operates in a glassy phase? Neurons in the cortex feature significant autocorrelation in spike counts separated by lags of seconds (see, e.g., [30]). This is a generic property of the glassy phase, which only requires positive levels of symmetry and large synaptic gains, as observed in the cortex. Alternative accounts [11, 12, 31–34] require significant fine-tuning of some parameter(s) (see [35] for a detailed discussion of this point). More concretely, the glassiness of the cortical dynamics could be directly assessed by looking at multi-point correlation functions [28, 29]. This appears feasible, given the availability of long, large-scale recordings of neuronal activity at the single-cell resolution.

Is there any computational advantage in operating in the glassy phase? It seems to us that there is no obvious disadvantage. Whatever benefit there is in operating in the chaotic phase, it can be only amplified by operating in the glassy phase, if only because of the broad spectrum of time scales and its robustness to changes in the synaptic gain and/or the level of symmetry. More interestingly, recent studies have shown that glassy materials can reliably perform useful computations, e.g., storing and retrieving memories, when driven cyclically [36]. These results, we believe, offer a fresh perspective for looking at old data, e.g., oscillations and short-term memory, possibly leading to new theoretical and experimental endeavors.

To conclude, we have shown that glassy dynamics can robustly occur in model systems whose dynamics are not described by an energy function. The underlying mechanism appears to be the same as the one leading to glassy dynamics in Hamiltonian systems, that is, the sudden...
Figure 4. (a) Phase diagram showing the different dynamical regimes as a function of the synaptic gain $\sigma_w^2$ and the level of symmetry $\rho_w$. Colored circles: representative points in the $\sigma_w^2 - \rho_w$ plane further illustrated in panels (b)-(e). The region marked unfeasible corresponds to pairs of $\sigma_w^2$ and $\rho_w$ that cannot be realized due to the positivity constraint for the synaptic efficacies, i.e., $w_{ij} > 0$. (b)-(e) $D^\alpha$, averaged over all pairs of initial conditions, as a function of time for different values of $\sigma_w^2$ and $\rho_w$. Shaded areas represent $\pm 1$ standard deviation from the mean. Colors identify the corresponding points in (a). Dashed black lines in (d) and (e) are least-squares fits to a stretched exponential (see main text for details).

We note that our results support earlier proposals that, to have structurally-robust glassy behavior, it is sufficient for the system to have an exponential number of fixed points with a broad distribution of spectral gaps [37]. This suggests that glassy behavior could, indeed, be quite common in biology. If so, the theoretical and experimental tools developed in physics to understand glasses could provide a suitable paradigm, not just a useful metaphor, to understand quantitatively complex behavior in biological systems.

We are grateful to B. Cessac, V. Hakim, G. La Camera, Y. Loewenstein, J.-P. Nadal, M. Tsodyks and C. van Vreeswijk for many stimulating discussions, and to V. Hakim and G. La Camera for a critical reading of the manuscript. G.M. acknowledges the hospitality and the financial support of the Simons Center for Systems Biology, at the School of Natural Sciences, Institute for Advanced Study, Princeton, where this work has been finalized. This work was supported by grants ANR-19-CE16-0024-01 and ANR-20-CE16-0011-02 from the French National Research Agency and by a grant from the Simons Foundation (891851, G.M.). K.B. acknowledges a fellowship from the ENS Paris-Saclay and support from the Office of Naval Research N00014-17-1-2041 (to X.-J. Wang).

[1] Y. Ahmadian and K. D. Miller, Neuron 109, 3373 (2021).

[2] M. V. Tsodyks and T. Sejnowski, Network: Computation in Neural Systems 6, 111 (1995).

[3] C. van Vreeswijk and H. Sompolinsky, Science 274, 1724 (1996).

[4] D. J. Amit and N. Brunel, Cerebral Cortex 7, 237 (1997).

[5] A. Renart, J. de la Rocha, P. Bartho, L. Hollender, N. Parga, A. Reyes, and K. D. Harris, Science 327, 587 (2010).

[6] A. Lerchner, C. Urrutia, J. Hertz, M. Ahmadi, P. Ruffiot, and S. Enemark, Neural Comp. 18, 634 (2006).

[7] T. Sejnowski, Network: Computation in Neural Systems 6, 111 (1995).

[8] G. Mongillo, D. Hansel, and C. van Vreeswijk, Phys. Rev. Lett. 108, 158101 (2012).

[9] M. Helias, T. Tetzlaff, and M. Diesmann, PLoS Computational Biology 10 (2014).

[10] O. Harish and D. Hansel, PLoS Computational Biology 11, e1004266 (2015).

[11] A. Crisanti and H. Sompolinsky, Physical Review X 5, 041030 (2015).

[12] S. Song, P. J. Sjöström, M. Reigl, N. Nelson, and D. B. Chklovskii, PLoS Biol. 3, e68 (2005).

[13] A. Crisanti and H. Sompolinsky, Physical Review A 37, 486 (1988).

[14] J. Kadmon and H. Sompolinsky, Physical Review X 5, 041030 (2015).

[15] N. Brunel, Nature Neuroscience 19, 749 (2016).

[16] D. Martí, N. Brunel, and S. Ostojic, Physical Review E 97, 062314 (2018).

[17] S. Rao, D. Hansel, and C. van Vreeswijk, Scientific Reports 9, 1 (2019).

[18] M. Mézard, G. Parisi, and M. Virasoro, Spin glass theory
and beyond: An introduction to the replica method and its applications (World Scientific Publishing Company, 1987).

[20] According to Dale’s law, a given neuron produces only one neurotransmitter. Hence, all its synapses have the same effect on the post-synaptic targets, either excitatory or inhibitory depending on the neurotransmitter. There are rare exceptions to this law.

[21] See Supplemental Material at XYZ for a detailed derivation of the mean-field theory and of the conditions for the stability of the mean-field solutions.

[22] H. J. Sommers, A. Crisanti, H. Sompolinsky, and Y. Stein, Physical review letters 60, 1895 (1988).

[23] A. Altieri, F. Roy, C. Cammarota, and G. Biroli, Physical Review Letters 126, 258301 (2021).

[24] B. Cessac, Journal de Physique I 5, 409 (1995).

[25] C. van Vreeswijk and H. Somplinsky, Neural Comp. 10, 1321 (1998).

[26] G. Wainrib and J. Touboul, Physical review letters 110, 118101 (2013).

[27] A. J. Bray and M. A. Moore, Journal of Physics C: Solid State Physics 12, L441 (1979).

[28] L. Berthier and G. Biroli, Reviews of Modern Physics 83, 587 (2011).

[29] T. Kirkpatrick and D. Thirumalai, Reviews of Modern Physics 87, 183 (2015).

[30] J. D. Murray, A. Bernacchia, D. J. Freedman, R. Romo, J. D. Wallis, X. Cai, C. Padoa-Schioppa, T. Pasternak, H. Seo, D. Lee, et al., Nature Neuroscience 17, 1661 (2014).

[31] G. Deco and E. Hugues, PLoS Computational Biology 8, e1002395 (2012).

[32] A. Litwin-Kumar and B. Doiron, Nature Neuroscience 15, 1498 (2012).

[33] L. Mazzucato, A. Fontanini, and G. La Camera, Journal of Neuroscience 35, 8214 (2015).

[34] D. Dahmen, S. Grün, M. Diesmann, and M. Helias, Proceedings of the National Academy of Sciences 116, 13051 (2019).

[35] C. Huang and B. Doiron, Curr Opin Neurobiol 46, 31 (2017).

[36] N. C. Keim, J. D. Paulsen, Z. Zeravcic, S. Sastry, and S. R. Nagel, Reviews of Modern Physics 91, 035002 (2019).

[37] L. F. Cugliandolo, J. Kurchan, P. Le Doussal, and L. Peliti, Physical review letters 78, 350 (1997).