



## Topological and Dynamical Complexity of Random Neural Networks

Gilles Wainrib<sup>1,\*</sup> and Jonathan Touboul<sup>2,3,†</sup>

<sup>1</sup>LAGA, Université Paris 13, Sorbonne Paris Cité, LAGA, CNRS (UMR 7539), 99 avenue J.B. Clément, F-93430 Villetaneuse, France

<sup>2</sup>The Mathematical Neuroscience Laboratory, CIRB/Collège de France (CNRS UMR 7241, INSERM U1050, UPMC ED 158, MEMOLIFE PSL\*), 11, place Marcelin Berthelot, 75005 Paris, France

<sup>3</sup>BANG Laboratory, INRIA Paris-Rocquencourt, Domaine de Voluceau, 78153 Le Chesnay, France

(Received 6 November 2012; published 11 March 2013)

Random neural networks are dynamical descriptions of randomly interconnected neural units. These show a phase transition to chaos as a disorder parameter is increased. The microscopic mechanisms underlying this phase transition are unknown and, similar to spin glasses, shall be fundamentally related to the behavior of the system. In this Letter, we investigate the explosion of complexity arising near that phase transition. We show that the mean number of equilibria undergoes a sharp transition from one equilibrium to a very large number scaling exponentially with the dimension on the system. Near criticality, we compute the exponential rate of divergence, called topological complexity. Strikingly, we show that it behaves exactly as the maximal Lyapunov exponent, a classical measure of dynamical complexity. This relationship unravels a microscopic mechanism leading to chaos which we further demonstrate on a simpler disordered system, suggesting a deep and underexplored link between topological and dynamical complexity.

DOI: [10.1103/PhysRevLett.110.118101](https://doi.org/10.1103/PhysRevLett.110.118101)

PACS numbers: 87.18.Tt, 05.10.-a, 87.18.Nq, 87.19.lj

Heterogeneity of interconnections is a crucial property to understand the behavior of realistic networks that arise in the modeling of physical, biological, or social complex systems. Among these, a paramount example is given by neuronal networks of the brain. In these systems, synaptic connections display characteristic disorder properties [1,2] resulting from development and learning. Taking into account this heterogeneity seems now essential, as experimental studies of neuronal tissues have shown that the degree of disorder in the connections significantly impacts the input-output function, rhythmicity, and synchrony, effects that can be related to transitions between physiological and pathological behaviors [3–5]. As an example, Aradi and Soltesz [3] have shown that rats subject to febrile seizure present the same average synaptic properties but increased variance.

These properties are reminiscent of disordered physical systems such as spin glasses. The relationship between disorder and qualitative behaviors has been thoroughly studied within theoretical frameworks such as the Sherrington-Kirkpatrick model [6] describing the behavior of binary variables interacting through a random connectivity matrix. In these models, estimating the number of metastable states, deemed to be deeply related to the transition to chaos, remains an important endeavor [7]. In the context of nervous system modeling, random neural networks [8,9] constitute a prominent class of models, which describe the evolution of the activity of a neuron  $i$  in an  $n$ -neurons network through the randomly coupled system of ordinary differential equations:

$$\dot{x}_i = -x_i + \sum_{j=1}^n J_{ij} S(x_j), \quad (1)$$

where  $J_{ij}$  are independent centered Gaussian random variables of variance  $\sigma^2/n$  representing the synaptic connectivity coefficients between neuron  $i$  and  $j$  [10] and  $S$  is an odd sigmoid function with maximal unit slope at the origin [ $S'(x) \leq S'(0) = 1$ ] representing the synaptic nonlinearity.

The behavior of system (1) has been analyzed in the asymptotic regime of infinite population size  $n \rightarrow \infty$  [8] and displays a phase transition for a critical value of the disorder at  $\sigma = 1$ : For  $\sigma < 1$  all the trajectories are attracted to the trivial equilibrium  $\mathbf{x} = 0$ , and for  $\sigma > 1$  the trajectories have chaotic dynamics. This generic phase transition has been numerically observed in a number of situations in more realistic models involving multiple populations and excitable dynamics [11]. That phase transition and the chaotic regime beyond the edge of chaos appear particularly relevant to understanding the computational capabilities of neuronal networks. In particular, information processing capacity was characterized as optimal at the edge of criticality [12,13], and such random neural networks are in particular used in recent machine learning algorithms [14]. Moreover, the question of criticality has been widely debated in the theoretical neuroscience community and beyond [15–17], and it seems that, through a number of different mechanisms among which is plasticity of synapses, networks may tend to be naturally poised near criticality. In the random neural network, the microscopic mechanisms underpinning this phase transition remain unclear. It is hence of great interest to dissect precisely the behavior of such systems at the edge of chaos.

Characterizing the topological modifications of the phase space arising at the edge of chaos is precisely the question we shall address in this Letter. More precisely,

we will estimate the averaged number of equilibria in the random neural network (1). To this end, we will develop upon the theories of random matrices [18], random geometry, and Gaussian fields [19].

We denote by  $A_n(\sigma)$  the random number of equilibria [depending on the realization of the matrix  $\mathbf{J} = (J_{ij})$ ]. These are the solutions of the system:

$$x_i = \sum_{j=1}^n J_{ij} S(x_j). \quad (2)$$

For  $\sigma < 1$ , consistent with the mean-field analysis, we first show that  $\mathbb{E}[A_n(\sigma)] \rightarrow 1$  when  $n \rightarrow \infty$ . The proof proceeds by showing that the system is contracting on the whole space  $\mathbb{R}^n$ . To this purpose, one needs to characterize the eigenvalues of the random Jacobian matrix  $-\mathbf{I} + \mathbf{J} \cdot \Delta(S'(\mathbf{x}))$ , where  $\mathbf{I}$  is the identity matrix and  $\Delta(S'(\mathbf{x}))$  is the diagonal matrix with elements  $S'(x_i)$ . The matrix  $\mathbf{J} \cdot \Delta(S'(\mathbf{x}))$  is a centered Gaussian matrix with independent components, and each column has a distinct standard deviation given by  $\sigma^2 S'(x_i)^2/n$ . Rajan and Abbott [20] provide the system of equations satisfied by the squared modulus of the eigenvalues of such random matrices and solve these when considering only two different variances. Their methodology readily generalizes to our problem, and elementary algebraic manipulations (see Refs. [21,22]) show that the spectral density has support in the disk of radius  $\frac{\sigma^2}{n} \sum_{i=1}^n S'(x_i)^2$  for large  $n$ . In our case,  $S'(x) \leq 1$ , and hence all eigenvalues of the matrix  $-\mathbf{I} + \mathbf{J} \cdot \Delta(S'(\mathbf{x}))$  have a negative real part in the limit  $n \rightarrow \infty$ . This implies global contraction of the dynamics, ensuring the fact that the trivial equilibrium  $\mathbf{x} = 0$  is the unique equilibrium for  $\sigma < 1$  (Banach fixed point theorem) and its global stability.

For  $\sigma > 1$  the situation is more complex. Similar to the phase transition in spin glasses, the behavior of  $\mathbb{E}[A_n(\sigma)]$  is likely to scale as  $\exp[nC(\sigma)]$ , where  $C(\sigma)$  is the topological complexity. Our starting point is to observe that the fixed point equations (2) can be viewed as zero crossings of a Gaussian field indexed by  $\mathbf{x} \in \mathbb{R}^n$ . Moreover, regularity and measurability of that Gaussian field ensure that the Kac-Rice [23] formula can be applied in order to estimate the number of solutions:

$$\mathbb{E}[A_n(\sigma)] = \int_{\mathbb{R}^n} d\mathbf{x} \mathbb{E}[|\det[-\mathbf{I} + \mathbf{J} \cdot \Delta(S'(\mathbf{x}))]| \times \delta_0(-\mathbf{x} + \mathbf{J} \cdot S(\mathbf{x}))].$$

Recent studies [26–28] have used that formula to estimate the mean number of critical points in random energy landscapes arising from Hamiltonian systems with strong symmetry properties (spin glass with spherical symmetry in Ref. [28] and translation invariant potential in Refs. [26,27]). In our case, the situation is substantially different: (i) There is no underlying energy landscape since the system (1) is not Hamiltonian, and (ii) symmetry properties of the vector field do not enable the same kind of reduction method developed in Refs. [26–28]. A major

technical difficulty is the fact that, because of the lack of symmetry in our system, one needs to deal with the determinant of random matrices with columns of nonidentical variances, for which the spectral density is unknown.

However, near criticality, for  $\sigma = 1 + \varepsilon$  with  $0 < \varepsilon \ll 1$ , we can obtain a first-order estimate of the number of equilibria taking advantage of the fact that all equilibria remain close to  $\mathbf{x} = 0$ . More precisely, we first show that with an arbitrarily high probability  $1 - \xi(n, \varepsilon)$  with  $\xi(n, \varepsilon) \rightarrow 0$  as  $n \rightarrow \infty$ , all equilibria belong to a ball  $\mathcal{B}_{\rho(\varepsilon)}$  centered at  $\mathbf{x} = 0$  of radius  $\rho(\varepsilon)$  which tends to 0 as  $\varepsilon \rightarrow 0$ . This is a consequence of the spectral analysis of the random matrix  $-\mathbf{I} + \mathbf{J} \cdot \Delta(S'(\mathbf{x}))$ , whose eigenvalues all have negative real parts outside of  $\mathcal{B}_{\rho(\varepsilon)}$  for large  $n$ . The property that fixed points remain in a small ball around zero is nontrivial. The proof proceeds by defining  $\rho(\varepsilon)$  the unique positive solution of the scalar equation  $x/\sigma = S(x)$ , which is clearly arbitrarily small when  $\varepsilon \rightarrow 0$ , and the smooth modified sigmoid function  $S_{\varepsilon, \eta}$ :

$$S_{\varepsilon, \eta}(x) = \begin{cases} x/\sigma & |x| < \rho(\varepsilon) \\ S(x) & |x| > \rho(\varepsilon) + \eta \end{cases}$$

[the small interval  $[\rho(\varepsilon), \rho(\varepsilon) + \eta]$  allows us to define a smooth continuation]. Because of the properties of the sigmoidal function, and in particular the fact that the differential is decreasing for  $x > 0$  (and increasing for  $x < 0$ ), the same argument as used in the case  $\sigma < 1$  applied to system (1) defined with the sigmoid  $S_{\varepsilon, \eta}$  (termed *modified system*) ensures that all eigenvalues of  $-\mathbf{I} + \mathbf{J} \cdot \Delta(S'_{\varepsilon, \eta}(\mathbf{x}))$  have a negative real part and hence that the unique fixed point of the modified system is 0. In particular, there is no fixed point outside the ball of radius  $\rho(\varepsilon)$ . In that region the original and modified systems are identical, implying that the only possible fixed points of the original system are contained in the ball of radius  $\rho(\varepsilon)$ .

Therefore, one can split the expectation according to whether  $|\mathbf{x}| < \rho(\varepsilon)$  or not, yielding

$$\mathbb{E}[A_n(\sigma)] = \int_{\mathcal{B}_{\rho(\varepsilon)}} d\mathbf{x} \mathbb{E}[|\det[-\mathbf{I} + \mathbf{J} \cdot \Delta(S'(\mathbf{x}))]| \times \delta_0(-\mathbf{x} + \mathbf{J} \cdot S(\mathbf{x}))] + O(\xi(n, \varepsilon)).$$

Moreover, thanks to the differentiability of the determinant operator, we know that within the ball  $\mathcal{B}_{\rho(\varepsilon)}$ , the integrand is equal to  $|\det(-\mathbf{I} + \mathbf{J})| + O[\rho(\varepsilon)]$ , eventually yielding

$$\mathbb{E}[A_n(\sigma)] = \mathbb{E}[|\det(-\mathbf{I} + \mathbf{J})|] + O[\rho(\varepsilon) + \xi(n, \varepsilon)].$$

To evaluate this formula, we first compute the logarithm of the determinant:

$$\frac{1}{n} \log |\det(-\mathbf{I} + \mathbf{J})| = \frac{1}{n} \sum_{\lambda \in \text{sp}(\mathbf{J})} \log |\lambda - 1|,$$

where  $\text{sp}(\mathbf{J})$  denotes the spectrum of  $\mathbf{J}$ , which in the large  $n$  limit is uniformly distributed in the disk of radius  $\sigma$  [29]. Using this property one obtains in the large  $n$  limit

$$\frac{1}{n} \log |\det(-\mathbf{I} + \mathbf{J})| = c(\sigma) + R(n)$$

and  $c(\sigma) := \int_{\mathbb{C}} \log |z - 1| \mu_{\sigma}(dz)$  with  $\mu_{\sigma}(dz)$  the uniform distribution on the disk of radius  $\sigma = 1 + \varepsilon$  and  $R(n)$  the finite-size error associated to the convergence to the circular law. We have

$$\frac{1}{n} \log \mathbb{E}[|\det(-\mathbf{I} + \mathbf{J})|] = c(\sigma) + \frac{1}{n} \log \mathbb{E}[e^{nR(n)}]. \quad (3)$$

This ensures that  $\frac{1}{n} \log \mathbb{E}[A_n(\sigma)]$  is arbitrarily close to  $c(\sigma)$  when  $\varepsilon \rightarrow 0$  and  $n \rightarrow \infty$ .

We are hence left computing  $c(\sigma)$ . Since  $\log |z|$  is harmonic, we can show that for  $a, b > 0$

$$\int_0^{2\pi} \log |a - b e^{i\theta}| d\theta = 2\pi \log[\min(a, b)],$$

ensuring that  $c(\sigma) = 0$  for  $\sigma < 1$ , which is consistent with our previous analysis of this case, and

$$c(\sigma) = \frac{1}{\pi\sigma^2} \int_1^{\sigma} 2\pi r \log(r) dr = \log(\sigma) + \frac{1}{2} \left( \frac{1}{\sigma^2} - 1 \right) \quad (4)$$

for  $\sigma > 1$ . For  $\sigma$  close to  $1^+$ , we conclude that  $c(\sigma) \sim (\sigma - 1)^2$ . Therefore we have shown that

$$\mathbb{E}[A_n(\sigma)] \sim e^{n(\sigma-1)^2} \quad (5)$$

up to multiplicative polynomial factors.

This combinatorial explosion of the number of equilibria is the hallmark of the accumulation of bifurcations [30] in the neighborhood of the critical parameter value  $\sigma = 1$ . Moreover, this result provides a possible topological explanation for the emergence of chaos. For  $\sigma > 1$ , the phase space is heavily mined with equilibrium points, most of which are saddles due to the spectral properties of the Jacobian matrices as discussed above. Typical trajectories evolving in this landscape will hence wander from the vicinity of the different unstable equilibria and attractors appearing, inducing a very high sensitivity to perturbations, a distinctive feature of chaos.

The classical characterization of chaos relies on the evaluation of the maximal Lyapunov exponent of the trajectories quantifying the dynamical complexity. This quantity is defined as follows. Applying an infinitesimal external perturbation  $x_i \rightarrow x_i + \delta x_i^0$  on neuron  $i$  at time  $t_0$  induces a change on all neurons at subsequent times  $x_j(t) \rightarrow x_j(t) + \delta x_{ji}(t)$  defining a susceptibility matrix  $\Psi_{ij}(t_0 + \tau, t_0) = \delta x_{ji}(t_0 + \tau) / \delta x_i^0$ . From the trace of this matrix we shall define an averaged susceptibility  $\Psi^2(\tau) = \lim_{t_0 \rightarrow \infty} \frac{1}{n} \mathbb{E}[\sum_{i,j} \Psi_{ij}^2(t_0 + \tau, t_0)]$  and eventually obtain the maximal Lyapunov exponent:

$$\lambda = \lim_{t \rightarrow \infty} \frac{\log[\Psi^2(t)]}{t}.$$

This quantity was analyzed in Ref. [8] by using a spectral decomposition based on the one-dimensional Schrödinger equation. Near criticality, the decomposition dramatically simplifies and yields for  $\sigma \sim 1^+$

$$\lambda \sim (\sigma - 1)^2.$$

We hence conclude that the topological and dynamical complexities have the same behavior at the edge of chaos. More surprising is that, using the spectral decomposition of the Schrödinger equation in the limit  $\sigma \gg 1$ , Sompolinsky and collaborators show that  $\lambda$  diverges as  $\log(\sigma)$ , precisely as the complexity of the system given in formula (4), although our analysis rigorously only applies for  $\sigma$  close to 1. This further strengthens our microscopic interpretation of the emergence of chaos in relationship with the number of saddles.

If this interpretation is the actual phenomenon arising in random neural networks at the edge of chaos, then the same phenomenon may hold in simpler, lower-dimensional dynamical systems with a large number of unstable fixed points, and these shall display a similar relationship between the number of unstable equilibria and the Lyapunov exponent. Probably the simplest low-dimensional system with an easily controllable number of fixed points is the *fakir bed* dynamics, corresponding to the movement of a particle in a two-dimensional complex landscape with  $k$  unstable fixed points. More precisely, we now consider a particle confined in a compact subset of  $\mathbb{R}^2$ , with close to the origin a fixed number  $k$  of Gaussian hills (corresponding to the presence of  $k$  unstable fixed points) randomly chosen in space. Trajectories of particles soon get chaotic as  $k$  increases [see Fig. 1(b)], and we numerically compute the maximal Lyapunov exponent of the trajectories. Since the landscape is probabilistic due to the choice of the location of hills, we compute the averaged maximal Lyapunov exponent across 100 independent realizations of the process and plot it against the logarithm of the number of equilibria. The corresponding curve [Fig. 1(a)] indeed

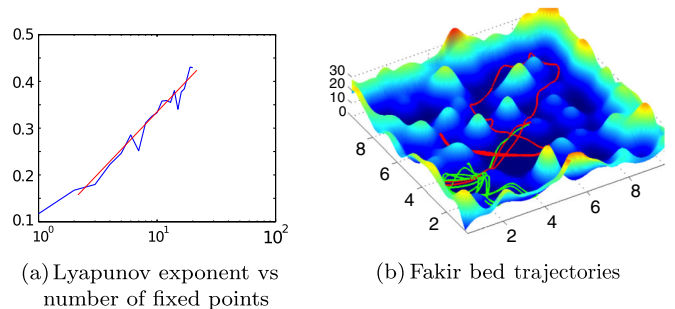


FIG. 1 (color online). The fakir bed dynamics. (a) Lyapunov exponent averaged over 100 realizations of the fakir dynamics vs number of fixed points (in semilogarithmic axes) and linear regression (slope 0.36). (b) Random landscape and two trajectories with very close initial conditions.

displays an increasing profile well approximated by a linear curve: A clear relationship again emerges between number of unstable fixed points and Lyapunov exponents, supporting our interpretation related to the random neural network.

Thus, motivated by the analysis of fine microscopic phenomena arising at the edge of chaos in random neural networks, we have shown that the average number of equilibria scales exponentially with the system size, with an exponential coefficient proportional to the Lyapunov exponent. This property, relying in part on the spectral theory of random matrices, readily inherits universality properties of the circular law [31] and therefore is valid for a large class of independent couplings beyond Gaussian. Matrices satisfying Dale's principle usually do not fall in this universality class and do not necessarily present the type of phase transition under consideration. Indeed, Dale's principle typically requires one to consider correlated or noncentered synaptic weights, generally modifying the spectrum of the connectivity matrix and the network dynamics. Combining recent results on the spectrum of such matrices [20,32] with the Kac-Rice formula should provide new insight into the complexity of these networks.

Moreover, our result shows that the complexity smoothly increases with the disorder parameter with a critical exponent 2, larger than 1, which corresponds to second-order phase transitions. This property ensures a form of structural robustness in the neighborhood of the phase transition in the sense that the complexity-related properties of the critical state would hold beyond the edge of chaos, which may have several implications in information processing capabilities [12]. Moreover, the result obtained (5) reveals a particular scaling  $\sigma = 1 + O(n^{-1/2})$  characterizing the typical thickness of the edge of chaos for large finite-size networks. This study is the first application to random neural networks of recent methods used for counting the number of metastable equilibria in spin glasses. From the theoretical viewpoint, we extended that approach to out of equilibrium, non-Hamiltonian systems at zero temperature (singular points of a vector field). This identity incidentally found between topological and dynamical complexity highlights what we conjecture to be a deep correspondence in large complex systems. Numerous questions and perspectives have emerged from this study, among which is the estimation of the distribution of equilibria and their number beyond the edge of criticality, requiring significant advances in the analysis of random matrix determinants or pursuing the exploration, in line with Refs. [33–35], of the relationship between topological and dynamical complexity with other measures such as the fractal dimensions of chaotic attractors.

\*wainrib@math.univ-paris13.fr

†jonathan.touboul@college-de-france.fr

- [1] D. Parker, *J. Neurosci.* **23**, 3154 (2003).
- [2] E. Marder and J. Goaillard, *Nat. Rev. Neurosci.* **7**, 563 (2006).
- [3] I. Aradi and I. Soltesz, *J. Physiol.* **538**, 227 (2002).
- [4] V. Santhakumar and I. Soltesz, *Trends Neurosci.* **27**, 504 (2004).
- [5] I. Soltesz, *Diversity in the Neuronal Machine: Order and Variability in Interneuronal Microcircuits* (Oxford University, New York, 2005).
- [6] D. Sherrington and S. Kirkpatrick, *Phys. Rev. Lett.* **35**, 1792 (1975).
- [7] G. Parisi, in *Mathematical statistical physics, École d'été de physique des houches session LXXXIII*, edited by A. Bovier, F. Dunlop, A. van Enter, F. den Hollander, and J. Dalibard, (ESF Summer school École thématique du CNRS, Les Houches, 2006), p. 4.
- [8] H. Sompolinsky, A. Crisanti, and H. Sommers, *Phys. Rev. Lett.* **61**, 259 (1988).
- [9] S. Amari, *IEEE Trans. Syst. Man Cybern.* **2**, 643 (1972).
- [10] This classical model does not satisfy Dale's principle: A neuron can here be both excitatory and inhibitory.
- [11] G. Hermann and J. Touboul, *Phys. Rev. Lett.* **109**, 018702 (2012).
- [12] T. Toyozumi and L. Abbott, *Phys. Rev. E* **84**, 051908 (2011).
- [13] D. Sussillo and L. Abbott, *Neuron* **63**, 544 (2009).
- [14] H. Jaeger and H. Haas, *Science* **304**, 78 (2004).
- [15] P. Bak, *How Nature Works: The Science of Self-Organized Criticality* (Copernicus, New York, 1996), Vol. 212.
- [16] J. Beggs, *Phil. Trans. R. Soc. A* **366**, 329 (2008).
- [17] M. Kitzbichler, M. Smith, S. Christensen, and E. Bullmore, *PLoS Comput. Biol.* **5**, e1000314 (2009).
- [18] T. Tao, *Topics in Random Matrix Theory* (American Mathematical Society, Providence, 2012), Vol. 132.
- [19] R. Adler and J. Taylor, *Random Fields and Geometry* (Springer-Verlag, Berlin, 2007), Vol. 115.
- [20] K. Rajan and L. Abbott, *Phys. Rev. Lett.* **97**, 188104 (2006).
- [21] L. Garcia del Molino, Master's thesis, École Polytechnique, 2012.
- [22] Y. Wei, *Phys. Rev. E* **85**, 066116 (2012).
- [23] This formula was initially introduced by Kac [24] to evaluate the average number of real roots of a random polynomial, and a modern account can be found in Refs. [19,25].
- [24] M. Kac, *Bull. Am. Math. Soc.* **49**, 314 (1943).
- [25] J. Azaïs and M. Wschebor, *Level Sets and Extrema of Random Processes and Fields* (Wiley, New York, 2009).
- [26] Y. Fyodorov, *Phys. Rev. Lett.* **92**, 240601 (2004).
- [27] Y. Fyodorov and I. Williams, *J. Stat. Phys.* **129**, 1081 (2007).
- [28] A. Auffinger and G. Arous, [arXiv:1110.5872](https://arxiv.org/abs/1110.5872).
- [29] V. Girko, *Theory Probab. Appl.* **29**, 694 (1985).
- [30] B. Cessac, *J. Phys. I (France)* **5**, 409 (1995).
- [31] T. Tao and V. Vu, *Commun. Contemp. Math.* **10**, 261 (2008).
- [32] T. Tao, *Probab. Theory Relat. Fields* **155**, 231 (2013).
- [33] F. Ledrappier, *Ann. Math.* **122**, 540 (1985).
- [34] P. Frederickson, J. Kaplan, E. Yorke, and J. Yorke, *J. Diff. Equ.* **49**, 185 (1983).
- [35] S. Chatterjee, [arXiv:0810.4221](https://arxiv.org/abs/0810.4221).