Regular graphs maximize the variability of random neural networks

Gilles Wainrib  
Mathieu Galtier

February 18, 2014

Abstract

Considering a model of dynamical units interconnected through a random weighted directed graph, we investigate how the behavior of the system is influenced by intricate relationships between the in-degree distribution, the variance of the weights and the non-linearity of the model. To this end, we develop an original theoretical approach based on a combination of a classical mean-field theory originally developed in the context of dynamical spin-glass models and the heterogeneous mean-field theory developed to study epidemic propagation on graphs. Our main result is that, surprisingly, increasing the variance of the in-degree distribution does not result in a more variable dynamical behavior, but on the contrary that the most variable behaviors are obtained in the regular graph setting. We further study how the dynamical complexity of the attractors is influenced by the statistical properties of the in-degree distribution.

The modelling of systems involving many dynamical units connected by a complex network of interactions is a topic of increasing interest in many scientific domains, such as neuroscience [15], genetics [9,8], epidemiology [13], artificial intelligence [8] or social sciences [6] (see also [2] for a general overview). In general, the understanding of the emerging properties of such systems relies on the combination of three main factors: (i) the structure of the connectivity network, (ii) the way units interact with each other and (iii) the dynamical properties of single units. One of the main theoretical challenge in this line of research is to unravel the subtle links between these three factors. In particular, in this Letter we are interested in the relationship between the degree distribution of a weighted graph and the properties of a system of interacting non-linear dynamical units on such a graph, motivated by the study of neural networks, with applications both toward theoretical neuroscience and machine learning. Such questions have been also asked recently in the context of synchronization phenomenon, showing the importance of the degree distribution for characterizing the transition to the synchronous state [14,15,17].

We consider a directed graph \( G = (v,E) \) with \( n \) nodes, and an in-degree distribution \( P_n(k) \). We denote \( \alpha = k/n \in [0,1] \) the rescaled in-degree and assumes that in the large \( n \) limit, \( P_n(\alpha n) \to p(\alpha) \), called the rescaled in-degree distribution (the typical in-degree of a node is assumed to be of order \( n \)). We further make the assumption of no correlation between in- and out-degree. To our knowledge the question of assortative mixing properties of anatomical neuronal networks has not been fully elucidated so far [4], and remains a question of current research. On each edge \( e = j \to i \), one assigns a weight \( J_{ij} \), which are independent centered random variables with finite variances, which may depend on the in-degree \( \alpha \) of node \( i \): \( \text{Var}[J_{ij}] = \sigma_\alpha^2/n \). We consider the following dynamical system on this weighted di-graph:

\[
 x_i(t+1) = S \left( \sum_{j \to i} J_{ij} x_j(t) \right) 
\]

where \( S(\cdot) \) is an odd sigmoid function with \( S(0) = 0 \), \( S'(0) = 1, S'(x) > 0 \) and \( x S''(x) \leq 0 \) (for instance \( S(x) = \tanh(x) \)). This model can be seen as a dynamical version of a zero-temperature spin-glass model [10] and has been studied in the context of neural networks in [15,7,21,22] in the case where the graph is fully connected, i.e. \( p = \delta_1 \). We will study this system deriving a self-consistent equation characterizing the variance of the nodes variables \( x_i(t) \), showing how a combination of the degree distribution \( p \) and the variance profile \( \sigma^2_\alpha \) controls the transition from a stable null equilibrium to a disordered state. Based on this analysis, we will show that regular graphs are surprisingly enhancing the variability of the nodes variables, while making the system more homogeneous in terms of network topology.

In the case of a general weighted graph, the overall strategy we introduce in this Letter is to combine the classical mean-field approach [13,15,7] with the idea of partitioning the nodes according to their rescaled degree \( \alpha \), at the heart of the heterogeneous mean-field (HMF) theory developed in the field of epidemiology [13,20]. Suppose \( i \) is a node with rescaled degree \( \alpha \) and denote \( \gamma_i^\alpha(t) \) the variance of \( x_i(t) \) and \( a_i = \sum_{j \to i} J_{ij} x_j(t) \). This sum contains only \( \alpha a \) terms, which come from nodes with various degree, and not only from nodes of degree \( \alpha \). The main idea of HMF is that these terms effectively sample the overall behavior of the system. Then the key step in the classical mean-field approach is to assume the independence between the \( x_i \)'s and \( J \) (mean-field assumption). This has been rigorously justified in [1,12,5] using large deviation techniques when adding an arbitrary small white-noise term on [1], but remains an open prob-
lem in the zero-noise case. Under the mean-field assumption, one deduces that \( a_i \) behaves as a centered Gaussian variable with variance \( \alpha \sigma^2 \gamma^2(t) \) where

\[
\gamma^2(t) = \int_0^1 p(\alpha)^2(t) d\alpha \tag{2}
\]

From the iteration equation \( x_i(t + 1) = S(a_i(t)) \), we further deduce:

\[
\gamma^2(t + 1) = F(\alpha \sigma^2 \gamma^2(t)) \tag{3}
\]

with

\[
F(z^2) = (2\pi)^{-1/2} \int_{\mathbb{R}} S^2(x) e^{-x^2/2} dx \tag{4}
\]

Combining the above equations, we obtain:

\[
\gamma^2(t + 1) = \int_0^1 p(\alpha) F(\alpha \sigma^2 \gamma^2(t)) d\alpha := \bar{F}(\gamma^2(t)) \tag{5}
\]

Therefore, to understand the order-disorder phase transition in this system, the first step is to study the dynamical system \( \gamma^2(t + 1) = \bar{F}(\gamma^2(t)) \). Due to the properties of the sigmoid function \( S(\cdot) \), the function \( \bar{F} \) is increasing, concave and satisfies \( F(0) = 0 \) and \( F'(0) = 1 \). We then deduce that \( \bar{F} \) is also increasing, concave and satisfies \( \bar{F}(0) = 0 \). Therefore, beyond the trivial equilibrium \( \gamma^2 = 0 \), the existence of another non-trivial equilibrium for (5) will depend on the value of the slope at zero:

\[
\bar{F}'(0) = \int_0^1 p(\alpha) \alpha \sigma^2 d\alpha := \mu \tag{6}
\]

Therefore, \( \gamma^2(t) \) converges to \( \gamma^2_{\infty} = 0 \) if \( \mu < 1 \), and \( \gamma^2(t) \) converges to a limit value \( \gamma^2_{\infty} > 0 \) if \( \mu > 1 \). For instance, in the classical case of a complete graph with homogeneous variances, \( p(\alpha) = \delta_1 \) and \( \sigma^2_{\infty} = \sigma^2 \), so the critical condition becomes \( \sigma = 1 \) as expected. In the case of a regular graph with rescaled degree \( \alpha_0 \) (each node has \( \alpha_0 n \) incoming edges) with homogeneous variances, \( p(\alpha) = \delta_0 \) and \( \sigma^2_{\infty} = \sigma^2 \), so the critical variance parameter becomes \( \sigma = \alpha_0^{-1/2} \). In the case of a general degree distribution with homogeneous variances \( \sigma^2_{\alpha} = \sigma^2 \) for all \( \alpha \in [0, 1] \), one obtains that the critical value of \( \sigma \) is given by \( \langle \alpha \rangle^{-1/2} \) where \( \langle \alpha \rangle \) is the mean rescaled degree.

Finally, among all the possible choices of \( \sigma_{\alpha} \), the case \( \sigma^2_{\alpha} = \sigma^2 / \alpha \) is particularly interesting since the critical value is always \( \sigma = 1 \), whatever the choice of \( p(\alpha) \), as one would have expected. This choice might correspond to the concept of synaptic scaling [19], ensuring that the overall input coming to a given unit has a typical strength independent of its in-degree.

The critical parameter \( \mu \) can be seen as a weighted average of the rescaled degrees, and in the case of a homogeneous variance profile is precisely proportional to the averaged rescaled degree \( \sigma^2(\alpha) \). In this case, it is natural to investigate the impact of the rescaled degree higher moments, such as its variance \( Var(\alpha) \), on the value of the fixed point \( \gamma^2_{\infty} \) characterizing the disordered state. It is not possible in general to obtain a closed form expression of \( \gamma^2_{\infty} \), however close to the transition, \( \bar{F} \) can be approximated by \( \bar{F}(x) = x \bar{F}'(0) + \frac{x^2}{2} \bar{F}''(0) + O(x^3) \) with \( \bar{F}''(0) \) given above, and \( \bar{F}''(0) = \bar{F}''(0) \int_0^1 p(\alpha) \alpha^2 d\alpha < 0 \) since \( \bar{F}''(0) < 0 \). Restricting the analysis to the homogeneous variances case, that is \( \sigma^2_{\alpha} = \sigma^2 \), we introduce a small parameter \( \epsilon = \sigma^2(\alpha) - 1 \) and obtain:

\[
x = x \sigma^2(\alpha) + \frac{\bar{F}''(0) \sigma^4}{2} (Var(\alpha) + \langle \alpha \rangle^2) x^2 + O(x^3) \tag{7}
\]

Therefore, for a given mean-degree \( \langle \alpha \rangle \), this formula implies that increasing the variance \( Var(\alpha) \) will decrease the variability \( \gamma^2_{\infty} \) of the neural activity variables, implying that regular graphs maximize \( \gamma^2_{\infty} \) when the mean-degree is kept fixed. We have tested this theoretical prediction by numerical simulations, constructing random weighted graphs with a prescribed degree distribution \( p_i = \frac{1}{2}(\delta_c + \delta_1 - \epsilon) \), by selecting half of the nodes to have an in-degree \( cn \alpha \) and the other half with an in-degree \( (1-c)n \). In this setting, the average rescaled degree \( \langle \alpha \rangle \) is kept constant at 1/2, while the variance \( Var(\alpha) \) depends on \( c \) as \( Var(\alpha) = (c - 1/2)^2 \). To measure the variability of the node variables, we compute the temporal average of the instantaneous variances of \( x_i(t) \):

\[
\gamma^2 = \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (x_i(t) - \bar{x}(t))^2 \tag{9}
\]

where \( \bar{x}(t) = \frac{1}{n} \sum_{i=1}^n x_i(t) \). Results are summarized in Figure 1, where the estimated variance \( \gamma^2 \) is displayed against the variance of the rescaled degree \( Var(\alpha) \), showing that the variability of the node variables is maximal when the graph is regular. In fact, it is also possible to compute the \( O(\epsilon^2) \) term in the expansion of \( \gamma^2_{\infty} = a_1 \epsilon + a_2 \epsilon^2 + O(\epsilon^3) \) and one finds:

\[
a_2 = \frac{2 \bar{F}''''(0)}{3 \sigma^6 \bar{F}''(0)^3} \langle \alpha^3 \rangle^2 \tag{10}
\]

To further characterize the non-trivial attractors in the disordered phase, it is of interest to estimate the maximal Lyapunov exponent \( \lambda \), as in [11], which is below 1 in the case of a steady state, equal to 1 in the case of a limit cycle and larger than 1 in the case of a chaotic attractor. To this end, we consider two solutions of (1) starting from two different initial conditions such that \( a_i(0) = a_i(0) \sim N(0, \delta_0 \sigma^2) \). Denoting \( \Delta_i(t) = a_i(t) - a_i(t) \), we have:

\[
\Delta_i(t + 1) = \sum_{j > i} J_{ij} (S(a_j(t)) - S(a_j(t)))
\]

\[
= \sum_{j > i} J_{ij} S'((a_j(t) + a_j(t))/2) \Delta_j(t) + O(||\Delta(t)||^2)
\]
Therefore, if node $i$ has a rescaled in-degree $\alpha$, one obtains the following relationship on the variances:
\begin{equation}
\delta^2_\alpha(t + 1) = \alpha \sigma^2 \Phi(\alpha \sigma^2 \gamma^2(t)) \delta^2(t) + O(\delta^3(t)) \tag{11}
\end{equation}
with
\begin{equation}
\Phi(z^2) := (2\pi)^{-1/2} \int S^2(zx)e^{-x^2/2}dx \tag{12}
\end{equation}
Integrating with respect to the degree distribution, we obtain:
\begin{equation}
\delta^2(t + 1) = \Phi(\gamma^2(t)) \delta^2(t) + O(\delta^3(t)) \tag{13}
\end{equation}
with
\begin{equation}
\Phi(\gamma^2(t)) := \int_{\sigma^2}^{1} p(\alpha) \sigma^2 \Phi(\alpha \sigma^2 \gamma^2(t)) d\alpha \tag{14}
\end{equation}
On the attractor, $\gamma^2(t) = \gamma^2_\infty$ so we deduce that
\begin{equation}
\lambda := \lim_{t \to \infty, \delta(0) \to 0} \left( \frac{\delta^2(t)}{\delta^2(0)} \right)^{1/t} = \Phi(\gamma^2_\infty) \tag{15}
\end{equation}
From this formula, one deduce, as expected, that in the subcritical regime $\lambda = \mu < 1$ because $\gamma^2_\infty = 0$ and $\Phi(0) = 1$. In the supercritical regime, it is natural to wonder whether $\lambda$ will also be a decreasing function of $\Var(\alpha)$. The answer is not straightforward because the distribution $p$ appears in $\lambda$ both explicitly in formula \[14\] and implicitly through the dependence of $\gamma^2_\infty$. Denoting $\beta = 2/(\alpha^4 \sigma^4 \langle \alpha^2 \rangle)$ and $\epsilon = \sigma^2(\alpha) - 1$, we know that $\gamma^2_\infty \sim \beta \epsilon$ when $\epsilon \ll 1$. Therefore, from a Taylor expansion of $\Phi$ around zero, we obtain:
\begin{align*}
\Phi(\gamma^2_\infty) &= 1 + \epsilon(1 + \langle \alpha^2 \rangle \sigma^2 \beta \Phi'(0)) \\
&\quad + \epsilon^2(\langle \alpha^3 \rangle \sigma^4 \beta^2/2) + O(\epsilon^3)
\end{align*}
When substituting $\beta$ with its expression, one discovers that $\langle \alpha^2 \rangle$ disappears in the $O(\epsilon)$ term, and appears next in the $O(\epsilon^2)$ term:
\begin{align*}
\Phi(\gamma^2_\infty) &= 1 + \epsilon(1 - \frac{\Phi'(0)}{\frac{\Phi''(0)}{\sigma^2}}) \\
&\quad + \epsilon^2 \frac{\langle \alpha^3 \rangle}{\langle \alpha^2 \rangle^2} \frac{1}{\sigma^4} \left( \frac{2\Phi''(0)}{\frac{\Phi''(0)}{\sigma^2}} + \frac{2\Phi'''(0)}{3\frac{\Phi''(0)}{\sigma^2}} \right) + O(\epsilon^3)
\end{align*}
The first consequence of this formula is that, at first order close to the transition, the dynamical complexity of the trajectories as measured by the Lyapunov exponent, does not depend on the higher order moments of the degree distribution. This means that $\Var(\alpha)$ controls the variability of the variables $x_i$ without modifying complexity of the chaotic attractor. This is true only at first order in $\epsilon$, and the dependence of the Lyapunov exponent is expressed through a more complicated parameter $\nu := \langle \alpha^3 \rangle / \langle \alpha^2 \rangle^2$, which accounts for the higher moments of the rescaled degree distribution, beyond the variance.

In this paper, we have shown that neural networks defined over regular graphs lead to a higher activity variability than those defined over irregular graphs. This somewhat counter-intuitive result illustrates that more homogeneity in the definition of a dynamical system does not necessarily lead to more order in the resulting dynamics. We have also shown that this effect of increasing the variance of the network while increasing regularity does not interact strongly with the stability of the attractor whose largest Lyapunov exponent remains almost constant, hence providing a way to control the variability of a neural representation without influencing much its dynamical complexity.

References

[1] G Ben Arous and A Guionnet. Large deviations for langevin spin glass dynamics. *Probability Theory and Related Fields*, 102(4):455–509, 1995.

[2] Alain Barrat, Marc Barthelemy, and Alessandro Vespignani. *Dynamical processes on complex networks*, volume 574. Cambridge University Press Cambridge, 2008.

[3] James M Bower and Hamid Bolouri. *Computational modeling of genetic and biochemical networks*. The MIT Press, 2001.

[4] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198, 2009.

[5] Tanguy Cabana and Jonathan Touboul. Large deviations, dynamics and phase transitions in large stochastic heterogeneous neural networks. *arXiv preprint arXiv:1302.6951*, 2013.
[6] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. Reviews of modern physics, 81(2):591, 2009.

[7] Bruno Cessac, Bernard Doyon, Mathias Quoy, and Manuel Samuelides. Mean-field equations, bifurcation map and route to chaos in discrete time neural networks. Physica D: Nonlinear Phenomena, 74(1):24–44, 1994.

[8] Herbert Jaeger and Harald Haas. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. Science, 304(5667):78–80, 2004.

[9] Stuart A Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. Journal of theoretical biology, 22(3):437–467, 1969.

[10] Marc Mézard, Giorgio Parisi, and Miguel Angel Virasoro. Spin glass theory and beyond, volume 9. World scientific Singapore, 1987.

[11] L. Molgedey, J. Schuchhardt, and H. G. Schuster. Suppressing chaos in neural networks by noise. Phys. Rev. Lett., 69:3717–3719, Dec 1992.

[12] Olivier Moynot and Manuel Samuelides. Large deviations and mean-field theory for asymmetric random recurrent neural networks. Probability Theory and Related Fields, 123(1):41–75, 2002.

[13] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. Physical review letters, 86(14):3200, 2001.

[14] Alex Roxin. The role of degree distribution in shaping the dynamics in networks of sparsely connected spiking neurons. Frontiers in computational neuroscience, 5, 2011.

[15] H Sompolinsky, A Crisanti, and HJ Sommers. Chaos in random neural networks. Physical Review Letters, 61(3):259–262, 1988.

[16] H Sompolinsky and Annette Zippelius. Dynamic theory of the spin-glass phase. Physical Review Letters, 47(5):359, 1981.

[17] Bernard Sonnenschein and Lutz Schimansky-Geier. Onset of synchronization in complex networks of noisy oscillators. Physical Review E, 85(5):051116, 2012.

[18] Bernard Sonnenschein, Michael A Zaks, Alexander B Neiman, and Lutz Schimansky-Geier. Excitable elements controlled by noise and network structure. The European Physical Journal Special Topics, 222(10):2517–2529, 2013.

[19] Gina G Turrigiano and Sacha B Nelson. Hebb and homeostasis in neuronal plasticity. Current opinion in neurobiology, 10(3):358–364, 2000.

[20] Alessandro Vespignani. Modelling dynamical processes in complex socio-technical systems. Nature Physics, 8(1):32–39, 2011.

[21] Gilles Wainrib and Luis Carlos García del Molino. Optimal system size for complex dynamics in random neural networks near criticality. Chaos: An Interdisciplinary Journal of Nonlinear Science, 23(4):–, 2013.

[22] Gilles Wainrib and Jonathan Touboul. Topological and dynamical complexity of random neural networks. Physical Review Letters, 110(11):118101, 2013.