Phase Space Analysis of Chaotic Neural Networks

Jakob Stubenrauch,[1, 2, *] Christian Keup,[1, 2] Anno C. Kurth,[1, 2] Moritz Helias,[1, 2] and Alexander van Meegen[1, 3]

[1]*Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and*
JARA-Institute Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany
[2]*Department of Physics, Faculty 1, RWTH Aachen University, Aachen, Germany*
[3]*Institute of Zoology, University of Cologne, 50674 Cologne, Germany*
(Dated: October 17, 2022)

We analytically determine the distribution of fixed points in a canonical model of a chaotic neural network. This distribution reveals that fixed points and dynamics are confined to separate shells in phase space. Furthermore, the distribution enables us to determine the eigenvalue spectra of the Jacobian at the fixed points. Perhaps counter-intuitively, the velocity of the dynamics is strongly correlated with the direction imposed by the nearest fixed point despite the spatial separation. We propose that this influence of the fixed points is mediated by tangentially fixed lines.

Introduction.– Chaotic dynamics are well understood in low dimensional systems but are notoriously challenging in high dimensions [1]. In low dimensions, the first step in the analysis of a dynamical system is to determine its fixed points in phase space, for example the two instable fixed points at the centers of the Lorenz attractor [2]. For high-dimensional nonlinear systems, merely finding all fixed points rapidly becomes prohibitive [3].

Neural networks are high-dimensional nonlinear systems which can exhibit chaotic dynamics [4, 5]. Furthermore, they operate outside equilibrium due to the asymmetric coupling [6, 7]. We here consider a canonical model of a chaotic neural network [4]: N nonlinearly connected units $x_i(t)$ receiving a constant external input η_i and obeying the dynamics

$$\dot{x}_i = -x_i + \sum_{j=1}^{N} J_{ij}\phi(x_j) + \eta_i \tag{1}$$

with nonlinear transfer function $\phi(x) = \tanh(x)$, independent and identically distributed (i.i.d.) coupling weights $J_{ij} \sim \mathcal{N}(0, g^2/N)$, and i.i.d. external inputs $\eta_i \sim \mathcal{N}(0, D)$. Due to the directional nature of synapses, the coupling weights are asymmetric.

The analytical approach of dynamical mean-field theory [4, 8] led to a deep understanding of the dynamics of the recurrent network model (1) at large N. Both without [4, 9] and with [10] external input, the statistics of the activity is well captured by a zero-mean Gaussian process with self-consistent autocorrelation function and the system is chaotic above a critical value of $g = g_c$ (without external input $g_c = 1$). Dynamical mean-field theory has also been successfully applied to various extensions of the model [11–20]. Furthermore, cross-correlations [21] and the full Lyapunov spectrum have been investigated recently [22].

In contrast to the dynamics, the phase space and the fixed point structure received considerably less attention. The pioneering work [23] showed that for $g = 1 + \varepsilon$, $0 < \varepsilon \ll 1$ and in the absence of external input, the phase space contains a number of fixed points that grows exponentially with the size of the system N. Their finding

that the maximum Lyapunov exponent and the rate controlling the exponential increase of fixed points to leading order in ε coincide led the authors to hypothesize a deep link between the abundance of fixed points and the chaoticity of the dynamics. To investigate the relation between the fixed points and the dynamics, a mandatory first step is to establish the location of the fixed points.

Therefore, in this Letter, we first determine the spatial distribution of the fixed points. On the technical level, this requires to compute the expected zero-crossings of a Gaussian process with location dependent, i.e., non-homogeneous statistics. Afterwards, we compare the geometry of the fixed points with the geometry of the dynamics and show that both are confined to separate shells in the phase space. Next, we leverage the distribution of fixed points to investigate the stability of the local dynamics at the fixed points. Finally, we propose a coherent framework that exposes the topological structure connecting the two shells, thus explaining the influence of the fixed points on the dynamics despite their spatial separation.

Throughout the Letter, we assume that the network is in the chaotic regime $g > g_c$ and that the number of units N is sufficiently large to allow us focus on the leading order behavior, which we express by the abbreviated notation $a \doteq e^{Nb}$ to denote $\lim_{N\to\infty} \frac{1}{N} \ln a = b$.

Spatial distribution of fixed points.– From here on, we use vector notation to write Eq. (1) as $\dot{\boldsymbol{x}} = \boldsymbol{y}(\boldsymbol{x})$ where the velocity is given by $\boldsymbol{y}(\boldsymbol{x}) = -\boldsymbol{x} + \mathbf{J}\phi(\boldsymbol{x}) + \boldsymbol{\eta}$. Since the parameters $\mathbf{J}$ and $\boldsymbol{\eta}$ of the model are normally distributed, the velocity $\boldsymbol{y}(\boldsymbol{x})$ and the Jacobian $\boldsymbol{y}'(\boldsymbol{x}) = -\mathbb{1} + \mathbf{J}\,\mathrm{diag}[\phi'(\boldsymbol{x})]$ are Gaussian processes; note that both $\boldsymbol{y}(\boldsymbol{x})$ and $\boldsymbol{y}'(\boldsymbol{x})$ are non-homogeneous processes. Due to the randomness of $\boldsymbol{y}(\boldsymbol{x})$, the location of the fixed points $\boldsymbol{y}(\boldsymbol{x}) = 0$ is described by a distribution $\rho(\boldsymbol{x})$. This distribution counts how many fixed points are on average within an infinitesimal volume in phase space. We determine $\rho(\boldsymbol{x})$ from the Kac-Rice formula [24–26]

$$\rho(\boldsymbol{x}) = \langle \delta[\boldsymbol{y}(\boldsymbol{x})] \, |\det \boldsymbol{y}'(\boldsymbol{x})| \rangle_{\mathbf{J},\boldsymbol{\eta}}. \tag{2}$$

The expected number of fixed points N_{fp} follows from the normalization constant $N_{\mathrm{fp}} = \int_{\mathbb{R}^N} d\boldsymbol{x}\,\rho(\boldsymbol{x})$. The Jacobian determinant $|\det \boldsymbol{y}'(\boldsymbol{x})|$ in Eq. (2) ensures that every fixed point contributes to the distribution with equal weight. Eq. (2) is equivalent to a random matrix problem: Using Bayes' law to condition on $\boldsymbol{y}(\boldsymbol{x}) = 0$, $\rho(\boldsymbol{x})$ can be rewritten into [27, A.1]

$$\rho(\boldsymbol{x}) = p_{\mathrm{L}}(\boldsymbol{x}) \left\langle |\det\left[\mathbf{M}(\boldsymbol{x}) + \mathbf{X}\,\boldsymbol{\Sigma}(\boldsymbol{x})\right]| \right\rangle_{X_{ij}\sim\mathcal{N}(0,N^{-1})} \quad (3)$$

where the first factor $p_{\mathrm{L}}(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}\,|\,0, \kappa(\boldsymbol{x}) + D)$ with $\kappa(\boldsymbol{x}) = \frac{g^2}{N}\sum_{i=1}^{N}\phi(x_i)^2$ is the probability of the velocity to be zero and the second factor is the expected determinant of a random matrix with mean $\mathbf{M}(\boldsymbol{x})$ and covariance $\boldsymbol{\Sigma}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}(\boldsymbol{x})$ controlling for the fluctuations of the velocity process. Here, $\kappa(\boldsymbol{x}) + D$ is the variance of the Gaussian process $\boldsymbol{y}(\boldsymbol{x})$ and $\mathbf{M}(\boldsymbol{x})$, $\boldsymbol{\Sigma}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}(\boldsymbol{x})$ are the mean and variance, respectively, of the Gaussian process $\boldsymbol{y}'(\boldsymbol{x})$ conditioned on $\boldsymbol{y}(\boldsymbol{x}) = 0$. Extending the technique introduced in [28], the determinant can be calculated; assuming self-averaging and excluding singularities, it is given by $\left\langle |\det\left[\mathbf{M}(\boldsymbol{x}) + \mathbf{X}\,\boldsymbol{\Sigma}(\boldsymbol{x})\right]| \right\rangle \doteq \exp\left(N\zeta(\boldsymbol{x})\right)$ with [27, A.2]

$$\zeta(\boldsymbol{x}) = -\frac{1}{2}z_* + \frac{1}{2N}\sum_{i=1}^{N}\ln\left[1 + z_* g^2 \phi'(x_i)^2\right] \quad (4)$$

where z_* is the solution of

$$1 = \frac{1}{N}\sum_{i=1}^{N}\frac{g^2\phi'(x_i)^2}{1 + z_* g^2\phi'(x_i)^2}. \quad (5)$$

To summarize, the N-dimensional distribution of the fixed points is $\rho(\boldsymbol{x}) \doteq \exp\left(-NS(\boldsymbol{x})\right)$ with

$$S(\boldsymbol{x}) = \frac{q(\boldsymbol{x})}{2[\kappa(\boldsymbol{x}) + D]} + \ln\{2\pi[\kappa(\boldsymbol{x}) + D]\} - \zeta(\boldsymbol{x}) \quad (6)$$

where $q(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^{N}x_i^2$ and $\zeta(\boldsymbol{x})$ is determined by Eq. (4) and Eq. (5).

The fixed point distribution (6) has the symmetry of a hypercube. Furthermore, the distribution depends on the location in phase space $\boldsymbol{x}$ only through the scalar fields $q(\boldsymbol{x})$, $\kappa(\boldsymbol{x})$, and $\zeta(\boldsymbol{x})$. All three scalar fields are collective variables which depend on $\boldsymbol{x}$ through network-averages. This property enables us to express the fixed point distribution as a functional $\rho(\boldsymbol{x}) = \rho[\mu_{\boldsymbol{x}}]$ of the *empirical measure*

$$\mu_{\boldsymbol{x}}(y) = \frac{1}{N}\sum_{i=1}^{N}\delta(y - x_i), \quad (7)$$

i.e., the empirical distribution of vector components of $\boldsymbol{x}$. From the expected empirical measure at the fixed points $\mu_*(y) = \langle\mu_{\boldsymbol{x}}(y)\rangle_{\boldsymbol{x}\sim\rho(\boldsymbol{x})}$ all network-averaged expectation values $\left\langle\frac{1}{N}\sum_{i=1}^{N}f(x_i)\right\rangle_{\boldsymbol{x}\sim\rho(\boldsymbol{x})} = \int dy\,\mu_*(y)f(y)$ can be computed. The expected empirical measure is given,
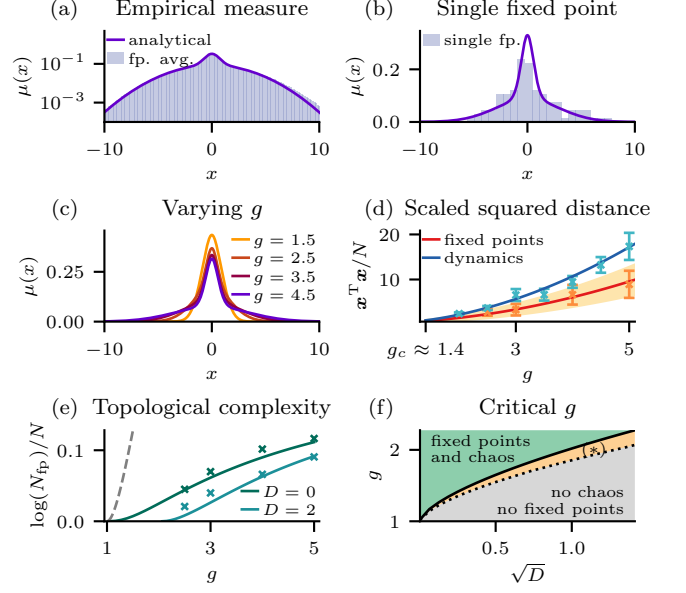

Figure 1. Quantitative characterization of fixed points. (a) Distribution of vector components of fixed points (empirical measure): theoretical result (solid line) based on Eq. (8) and histogram (bars) averaged across 4×10^4 fixed points for a single realization of the coupling weights. (b) Same as (a) for a randomly chosen single fixed point. (c) Empirical measure for different values of g. (d) Scaled squared distance to the origin of fixed points (theory: red solid line; theoretical finite-size standard deviation: yellow shading; numerical results: orange error bars) and dynamics (theory: blue solid line; numerical integration: turquoise error bars). (e) Number of fixed points (topological complexity; solid lines: theory based on Eq. (6); crosses: numerical results, see [27, E]; gray dashed line: theory for $g \to 1^+$, $D = 0$ by [23]). (f) Transition to positive topological complexity (black dotted line) and transition to chaos (black solid line) based on [10]. In the regime $(*)$ the phase space exhibits an exponential number of fixed points but the dynamics are not chaotic. Parameters: $D = 0.1$ for (a)–(d), $g = 4$ for (a) and (b), $N = 100$ for all numerical results and the shading in (d).

for large N, by the saddle point that maximizes $\rho[\mu_{\boldsymbol{x}}]$ in function space and admits the form [27, B.1]

$$\mu_*(y) \propto \sqrt{1 + \alpha\phi'(y)^2}\,e^{-\frac{y^2}{2\beta} + \gamma\phi(y)^2}, \quad (8)$$

for which the parameters α, β, and γ are determined by $1 = g^2\left\langle\left(\phi'(y)^{-2} + \alpha\right)^{-1}\right\rangle_{\mu_*}$, $\beta = \left\langle\phi(y)^2\right\rangle_{\mu_*} + D$, and $\gamma = \frac{g^2}{2\beta}\left(\beta^{-1}\left\langle y^2\right\rangle_{\mu_*} - 1\right)$ where expectation values have to be taken self-consistently with respect to the expected empirical measure.

We compare the empirical measure Eq. (8) to the distribution of vector components of numerically determined fixed points. For the numerical results, we fix the realization of the random parameters and employ a Levenberg-Marquart rootfinder starting from independent normally distributed initial conditions until satura-

tion, i.e., until almost no new fixed points are found (see [27, E]). We see in Fig. 1(a) that the theory Eq. (8) is in excellent agreement with the empirical measure averaged over all fixed points found numerically (see [27, Fig. 2(a)] for further examples). Moreover, as shown in Fig. 1(b), even single fixed points closely resemble the expected empirical measure.

To quantify the latter observation, we show in [27, B.2] that the probability distribution functional of the empirical measures takes the form $P[\mu] \doteq \exp(-NH[\mu])$ with an analytically determined rate functional $H[\mu]$. Put differently, μ obeys a large deviation principle [29, 30] with rate functional $H[\mu]$; the minimum of the rate functional is attained at the expected empirical measure μ_*. Since $P[\mu]$ quantifies both the variability within a realization of the parameters as well as across realizations [27, B.2], akin to the law of total variance, deviations of μ from μ_* are rare for large N even at the level of individual fixed points. Mismatches between $\mu_{\boldsymbol{x}}$ and μ_* for a fixed point $\boldsymbol{x}$ are thus finite size effects (see [27, Fig. 2(b)] for further examples).

Geometry of fixed point distribution.– The expected empirical measure μ_* provides insight into the geometry of the fixed point distribution in phase space. First, the sharp peak of μ_* in combination with its broad base implies that the fixed points are posed in the vicinity of the span of a subset of axes in phase space.

Next, we want to inspect the typical distance of fixed points to the origin. The expected value of the *scaled squared distance* $u_{\boldsymbol{x}} = \frac{1}{N}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$ is given by

$$u_* = \int dy\, y^2 \mu_*(y). \tag{9}$$

The distribution of the distance $P(u)$ inherits the exponential form of $P[\mu]$ because u is determined by the empirical measure; formally, this is a consequence of the contraction principle [29]. Thus, $P(u) \doteq e^{-NI(u)}$ where the rate function is

$$I(u) = \inf_{\mu:\, \int y^2 \mu(y)\, dy = u} H[\mu]. \tag{10}$$

The rate function is again $\mathcal{O}(1)$; hence, in the limit $N \gg 1$, the fluctuations of the distance vanish and we conclude that the fixed points are distributed on a thin spherical shell with radius $\sqrt{Nu_*}$. In Fig. 1(d), we show the average distance and fluctuations based on Eq. (10) (see [27, C.1]) for $N = 100$.

To put the fixed point's distance to the origin into context with the dynamics, we leverage the result from dynamic mean-field theory that the network-averaged variance $q[\boldsymbol{x}(t)] = \frac{1}{N}\sum_{i=1}^{N} x_i(t)^2$ is self-averaging for stationary statistics with fluctuations vanishing in the large N limit both across time and across network realizations [8, 20]. Hence, also the trajectory is embedded in a thin shell around the origin, which is of radius $\sqrt{Nq}$.
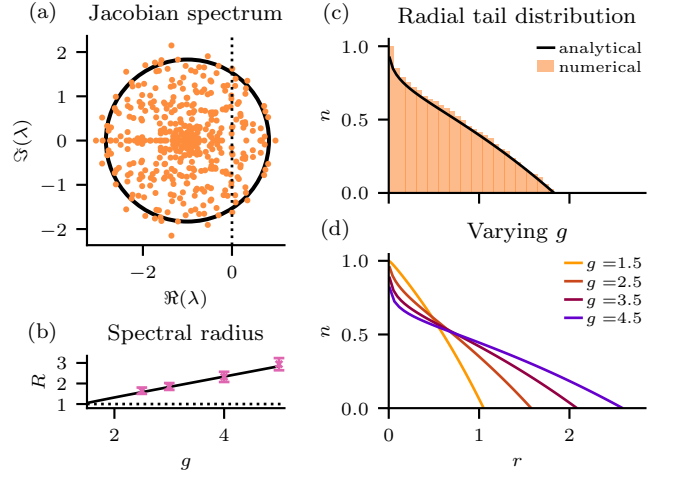

Figure 2. Jacobian spectrum at fixed points. (a) Eigenvalues (orange dots) of the Jacobian at five randomly chosen fixed points. Every eigenvalue with positive real part (black dotted line) corresponds to an unstable mode of the fixed point. The support of the spectrum (black circle) is determined by the spectral radius, Eq. (11). (b) Spectral radius; theory (black line) and numerical results (pink errorbars). (c) Radial tail distribution of eigenvalues; theory (black line) based on Eq. (12) and histogram (orange bars) based on the eigenvalues shown in (a). (d) Theory from (c) for varying g. Parameters: $N = 100$, $g = 3$ in (a) and (c), $D = 0.1$.

In Fig. 1(d), we compare the radii of both shells. We note that, for all $g > g_c$, the fixed points shell is inside of the trajectories shell. Furthermore, for $N \to \infty$, the overlap between the shells vanishes and thus the trajectory is clearly separated from the fixed points in phase space.

Number of fixed points.– A core result of [23] is that without noise, $D = 0$, the system has a transition from a single stable fixed point to an exponential number of unstable fixed points $N_{\mathrm{fp}} \doteq e^{cN}$ at $g_c = 1$. The respective rate c, called the *topological complexity* in [23], is derived in [27, C.2] and shown in Fig. 1(e). We see that the critical gain parameter g_c grows with $D > 0$; the corresponding transition line is shown in Fig. 1(f). For $D \ll 1$, the transition to an exponential number of fixed points coincides with the transition to chaos. For larger noise strengths D, however, a regime exists where the system has an exponential number of fixed points yet the dynamics are not chaotic.

Stability of fixed points.– After studying the number and location of fixed points in phase space, we now investigate the local dynamics in the vicinity of a fixed point $\boldsymbol{x}^*$. Linearizing the dynamics, $\dot{\boldsymbol{x}} = \boldsymbol{y}'(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*) + \mathcal{O}[(\boldsymbol{x} - \boldsymbol{x}^*)^2]$, shows that the eigenvalues (see Fig. 2(a)) of the Jacobian at the fixed point $\boldsymbol{y}'(\boldsymbol{x}^*)$ determine the local stability: Each eigenvalue with positive (negative) real part corresponds to an unstable (stable) eigendirection of the fixed point. The Jacobian can be written as $\boldsymbol{y}'(\boldsymbol{x}) = -\mathbb{1} + \mathbf{X}\,\mathrm{diag}[g\phi'(\boldsymbol{x})]$ where again
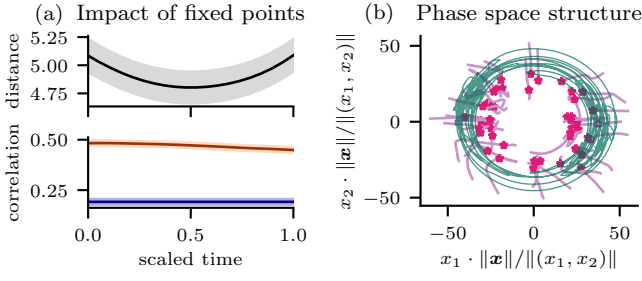
Figure 3. Impact of fixed points. (a) Distance to the nearest fixed point (upper panel) and Pearson correlation between $\boldsymbol{y}(\boldsymbol{x})$ and its linear predictor $\boldsymbol{y}'(\boldsymbol{x}^*)[\boldsymbol{x}(t) - \boldsymbol{x}^*]$ (lower panel, brown line) while the trajectory $\boldsymbol{x}(t)$ goes from a fixed point $\boldsymbol{x}^*$ being the nearest one (scaled time = 0) until another one is closer (scaled time = 1). Corresponding result using a randomly chosen fixed point (lower panel, blue line). Shading denotes standard deviation of the mean. (b) Projection of the trajectory (green line), the fixed points (pink asterisks) and the tangentially fixed lines (light pink lines) onto the first two neurons with a subsequent rescaling at every time-step / point such that the original norm is preserved in the projection. Parameters: $g = 5$, $D = 0$, $N = 100$.

$X_{ij} \sim \mathcal{N}(0, N^{-1})$. The eigenvalue spectrum of such random matrices can be computed with the method developed in [31] because $\mathrm{diag}[g\phi'(\boldsymbol{x})]$ is invertible. For large N, the eigenvalue distribution of $\boldsymbol{y}'(\boldsymbol{x})$ is centered around $-1 + 0i$ and confined within a circle of radius $R(\boldsymbol{x}) = g\sqrt{N^{-1}\phi'(\boldsymbol{x})^{\mathrm{T}}\phi'(\boldsymbol{x})}$. At a fixed point, the contraction principle attests a large deviation principle for the spectral radius, with the expected value given by

$$R_* = g\sqrt{\int dy\, \phi'(y)^2 \mu_*(y)}. \tag{11}$$

In Fig. 2, we see that this radius is always > 1 in the chaotic phase, indicating that for large N all fixed points are unstable.

Within the support, the distribution of eigenvalues is isotropic around the center. We can hence express the distribution by the fraction of eigenvalues further than r from the center $n_{\boldsymbol{x}}(r)$, i.e., the radial tail distribution. It obeys, again, a large deviation principle dominated by the solution $n_*(r)$ of

$$1 = \int dy\, \mu_*(y) \frac{g^2\phi'(y)^2}{r^2 + n_*(r)\, g^2\phi'(y)^2}. \tag{12}$$

We present the solution in Fig. 2(c). Here, we observe that, especially deep in the chaotic regime, the unstable modes of fixed points are underrepresented as opposed to a uniform spectrum.

Impact of fixed points.– In [23], it is conjectured that the dynamics can be understood as meandering around the different fixed points, first following their stable directions and then being repelled along their unstable directions. The separation of fixed points and dynamics

onto different shells in phase space seemingly contradicts this hypothesis.

To investigate the relation between fixed points and dynamics further, we compute the correlation of the actual velocity of the numerically integrated model with the velocity predicted by the dynamics linearized around the nearest fixed point. In Fig. 3(a), we observe that the velocities are indeed strongly correlated, especially while the trajectory moves towards the nearest fixed point. How can such a strong correlation arise despite the spatial separation of fixed points and dynamics? To understand this, we consider *tangentially* fixed points, i.e., points $\boldsymbol{x}$ where the velocity is purely radial $\boldsymbol{y}(\boldsymbol{x}) \propto \boldsymbol{x}$, or, put differently,

$$0 = \boldsymbol{y}(\boldsymbol{x}) - \frac{\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}}{\|\boldsymbol{x}\|^2}\boldsymbol{y}(\boldsymbol{x}). \tag{13}$$

These points also exist on the dynamics shell. Here, they play the same role as usual fixed points since the dynamics' radial component vanishes for $N \gg 1$ due to the self-averaging of the network-averaged variance q.

The true fixed points are also tangentially fixed with vanishing radial velocity. Do the true fixed points correspond to tangentially fixed points on the dynamics shell? We start from a true fixed point and move outwards (inwards) into the unique direction where only the radial velocity decreases (increases); all points along this path thus have vanishing tangential velocity, so we term the paths *tangentially fixed lines*, see Fig. 3(b). These lines can be computed using a Levenberg-Marquart rootfinder to solve Eq. (13) while imposing the distance to the origin $\|\boldsymbol{x}\| = r$. In Fig. 3(b), we observe that most of these lines intercept with the dynamical shell. Hence, for most true fixed points we find a tangentially fixed point on the dynamics shell which potentially mediates the effect of the fixed point and likely explains the correlation observed in Fig. 3(a).

Discussion.– In this Letter we characterize the phase space structure of a chaotic neural network using the distribution of fixed points. The finding that the fixed points are in a region separate from the dynamical attractor led to the question whether the fixed points even impact the dynamics. Numerically, the strong correlation between the velocity and its linear predictor based on the nearest fixed point shows that there is an impact. We propose that the influence of the fixed points is mediated by tangentially fixed lines, i.e., lines in phase space along which the velocity is purely radial. From a dynamical perspective the lines have an effect similar to fixed points: their unstable directions scatter incoming trajectories which results in chaotic dynamics.

In high-dimensional linear dynamical systems, May's pioneering stability analysis [32] enabled considerable insights into the dynamics of ecosystems [33]. In the nonlinear case, the number of fixed points can be determined if the velocity is generated by a homogeneous Gaussian

potential [34]; in this case, it is even possible to determine the number of minima of the potential [35–37] with applications in deep learning [38, 39]. The non-potential case has been addressed in [23] and [40]; in the former work the analysis is restricted to $g = 1 + \varepsilon$ while the latter work investigates a velocity field based on a homogeneous Gaussian field (for which it is possible to extend the analysis to the fraction of stable directions of fixed points [41]). A particular non-homogeneous case with dynamics constrained to a sphere has been solved in [42]. For a recent review on stationary points of random fields see [43]. Here, we go beyond the previous results by finding the complete distribution of fixed points, which includes their number, of the non-homogeneous random neural network (1) for arbitrary $g > g_c$. To this end, we extend methods from random matrix theory [28, 31] to compute the determinant of non-hermitian random matrices with a correlation structure including low-rank terms. Our numerical results show that the self-averaging assumption is justified for the model (1), i.e., quenched and annealed results coincide [44].

The results presented here pave the way towards a detailed mechanistic understanding of the velocity field underlying this fundamental model of a high dimensional chaotic network. There are several directions for further research: First, it would be interesting to extend the analysis to more structured networks, e.g., in terms of low rank perturbations [18], levels of symmetry [28, 45, 46], or population structure [13, 14, 20]. Second, we here focused on the regions in phase space containing either fixed points or ongoing dynamics. Going radially outwards from the unstable fixed point at the origin, is there a transition in the characteristics of the velocity field before the fixed-point shell? In particular, where and how do the tangentially fixed lines start? Third, the frustration created by the quenched rotation between the axes system, singled out by the element-wise application of the nonlinearity, and the eigensystem of the connectivity creates the complexity of the phase space. What is the geometric relation between the axes system and the dynamics on the chaotic attractor? Last, deep insights into trained neural networks are possible with an analysis of their phase space [47, 48]. Here, we analyzed the phase space of a random reservoir which already allows universal computation if the readout is optimized [49]—more generally, learning with chaotic networks is an exciting direction of ongoing research [19, 50, 51] that might be able to leverage the exponential number of fixed points and the associated capability for sequence processing.

* Corresponding author. jakob.stubenrauch@rwth-aachen.de

[1] S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, 2nd ed. (Westview Press, Philadelphia, PA, 2014).

[2] E. N. Lorenz, Deterministic nonperiodic flow, Journal of the Atmospheric Sciences **20**, 130 (1963).

[3] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. (Cambridge University Press, 2007).

[4] H. Sompolinsky, A. Crisanti, and H. J. Sommers, Chaos in random neural networks, Phys. Rev. Lett. **61**, 259 (1988).

[5] C. van Vreeswijk and H. Sompolinsky, Chaos in neuronal networks with balanced excitatory and inhibitory activity, Science **274**, 1724 (1996).

[6] M. I. Rabinovich, P. Varona, A. I. Selverston, and H. D. Abarbanel, Dynamical principles in neuroscience, Rev. Mod. Phys. **78**, 1213 (2006).

[7] H. Sompolinsky, Statistical mechanics of neuronal networks, Phys. Today **41**, 70 (1988).

[8] M. Helias and D. Dahmen, *Statistical Field Theory for Neural Networks* (Springer International Publishing, 2020) p. 203.

[9] A. Crisanti and H. Sompolinsky, Path integral approach to random neural networks, Phys. Rev. E **98**, 062120 (2018).

[10] J. Schuecker, S. Goedeke, and M. Helias, Optimal sequence memory in driven random networks, Phys. Rev. X **8**, 041029 (2018).

[11] L. Molgedey, J. Schuchhardt, and H. Schuster, Suppressing chaos in neural networks by noise, Phys. Rev. Lett. **69**, 3717 (1992).

[12] M. Stern, H. Sompolinsky, and L. F. Abbott, Dynamics of random neural networks with bistable units, Phys. Rev. E **90**, 062710 (2014).

[13] J. Kadmon and H. Sompolinsky, Transition to chaos in random neuronal networks, Phys. Rev. X **5**, 041030 (2015).

[14] J. Aljadeff, M. Stern, and T. Sharpee, Transition to chaos in random networks with cell-type-specific connectivity, Phys. Rev. Lett. **114**, 088101 (2015).

[15] F. Mastrogiuseppe and S. Ostojic, Intrinsically-generated fluctuating activity in excitatory-inhibitory networks, PLOS Comput. Biol. **13**, e1005498 (2017).

[16] A. van Meegen and B. Lindner, Self-consistent correlations of randomly coupled rotators in the asynchronous state, Phys. Rev. Lett. **121**, 258302 (2018).

[17] I. D. Landau and H. Sompolinsky, Coherent chaos in a

recurrent neural network with structured connectivity, PLOS Comput. Biol. **14**, e1006309 (2018).

[18] F. Mastrogiuseppe and S. Ostojic, Linking connectivity, dynamics, and computations in low-rank recurrent neural networks, Neuron **99**, 609 (2018).

[19] C. Keup, T. Kühn, D. Dahmen, and M. Helias, Transient chaotic dimensionality expansion by recurrent networks, Phys. Rev. X **11**, 10.1103/physrevx.11.021064 (2021).

[20] A. van Meegen, T. Kühn, and M. Helias, Large-deviation approach to random recurrent neuronal networks: Parameter inference and fluctuation-induced transitions, Phys. Rev. Lett. **127**, 158302 (2021).

[21] D. G. Clark, L. F. Abbott, and A. Litwin-Kumar, Dimension of activity in random neural networks (2022).

[22] R. Engelken, F. Wolf, and L. F. Abbott, Lyapunov spectra of chaotic recurrent neural networks (2020), arXiv:2006.02427.

[23] G. Wainrib and J. Touboul, Topological and dynamical complexity of random neural networks, Phys. Rev. Lett. **110**, 118101 (2013).

[24] M. Kac, On the average number of real roots of a random algebraic equation, Bulletin of the American Mathematical Society **49**, 314 (1943).

[25] S. O. Rice, Mathematical analysis of random noise, Bell Syst. Tech. J. **24**, 46 (1945), reprinted in [52].

[26] J.-M. Azaïs and M. Wschebor, *Level Sets and Extrema of Random Processes and Fields* (John Wiley & Sons, 2009).

[27] See supplemental material at XYZ for detailed derivations and further information, which includes Refs. [53–56].

[28] H. Sommers, A. Crisanti, H. Sompolinsky, and Y. Stein, Spectrum of large random asymmetric matrices, Phys. Rev. Lett. **60**, 1895 (1988).

[29] H. Touchette, The large deviation approach to statistical mechanics, Phys. Rep. **478**, 1 (2009).

[30] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications* (Springer Berlin Heidelberg, 2010).

[31] Y. Ahmadian, F. Fumarola, and K. D. Miller, Properties of networks with partially structured and partially random connectivity, Phys. Rev. E **91**, 012820 (2015).

[32] R. M. May, Will a large complex system be stable?, Nature **238**, 413 (1972).

[33] S. Allesina and S. Tang, The stability–complexity relationship at age 40: a random matrix perspective, Popul. Ecol. **57**, 63 (2015).

[34] Y. V. Fyodorov, Complexity of random energy landscapes, glass transition, and absolute value of the spectral determinant of random matrices, Phys. Rev. Lett. **92**, 240601 (2004).

[35] A. J. Bray and D. S. Dean, Statistics of critical points of gaussian fields on large-dimensional spaces, Phys. Rev. Lett. **98**, 150201 (2007).

[36] Y. V. Fyodorov and I. Williams, Replica symmetry breaking condition exposed by random matrix calculation of landscape complexity, J. Stat. Phys. **129**, 1081 (2007).

[37] Y. V. Fyodorov and C. Nadal, Critical behavior of the number of minima of a random landscape at the glass transition point and the tracy-widom distribution, Phys. Rev. Lett. **109**, 167203 (2012).

[38] Y. N. Dauphin, R. Pascanu, C. Gulcehre, K. Cho, S. Ganguli, and Y. Bengio, Identifying and attacking the saddle point problem in high-dimensional non-convex optimization, in *Adv. Neural Inf. Process. Syst.*, Vol. 27, edited by Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger (Curran Associates, Inc., 2014).

[39] A. Choromanska, M. Henaff, M. Mathieu, G. Ben Arous, and Y. LeCun, The Loss Surfaces of Multilayer Networks, in *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research, Vol. 38, edited by G. Lebanon and S. V. N. Vishwanathan (PMLR, San Diego, California, USA, 2015) pp. 192–204.

[40] Y. V. Fyodorov and B. A. Khoruzhenko, Nonlinear analogue of the may-wigner instability transition, Proc. Natl. Acad. Sci. USA **113**, 6827 (2016).

[41] G. Ben Arous, Y. V. Fyodorov, and B. A. Khoruzhenko, Counting equilibria of large complex systems by instability index, Proc. Natl. Acad. Sci. USA **118**, e2023719118 (2021).

[42] Y. V. Fyodorov, Topology trivialization transition in random non-gradient autonomous ODEs on a sphere, Journal of Statistical Mechanics: Theory and Experiment **2016**, 124003 (2016).

[43] V. Ros and Y. V. Fyodorov, The high-d landscapes paradigm: spin-glasses, and beyond (2022), arXiv:2209.07975.

[44] M. Mézard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond (World Scientific Lecture Notes in Physics, Vol 9)* (World Scientific Publishing Company, 1987).

[45] D. Martí, N. Brunel, and S. Ostojic, Correlations between synapses in pairs of neurons slow down dynamics in randomly connected neural networks, Phys. Rev. E **97**, 062314 (2018).

[46] K. Berlemont and G. Mongillo, Glassy phase in dynamically-balanced neuronal networks, BioRxiv , 484348 (2022).

[47] D. Sussillo and O. Barak, Opening the black box: Low-dimensional dynamics in high-dimensional recurrent neural networks, Neural Comput. **25**, 626 (2013).

[48] S. Vyas, M. D. Golub, D. Sussillo, and K. V. Shenoy, Computation Through Neural Population Dynamics, Annu. Rev. Neurosci. **43**, 249 (2020).

[49] W. Maass, T. Natschläger, and H. Markram, Real-time computing without stable states: a new framework for neural computation based on perturbations, Neural Comput. **14**, 2531 (2002).

[50] B. Poole, S. Lahiri, M. Raghu, J. Sohl-Dickstein, and S. Ganguli, Exponential expressivity in deep neural networks through transient chaos, in *Advances in Neural Information Processing Systems 29* (2016).

[51] M. Farrell, S. Recanatesi, T. Moore, G. Lajoie, and E. Shea-Brown, Gradient-based learning drives robust representations in recurrent neural networks by balancing compression and expansion, Nature Machine Intelligence **4**, 564 (2022).

[52] N. Wax, ed., *Selected Papers on Noise and Stochastic Processes* (Dover Publications, New York, 1954).

[53] R. L. Stratonovich, *Topics in the Theory of Random Noise* (Gordon and Breach, New York, 1967).

[54] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*, Adaptive Computation and Machine Learning (MIT Press, Cambridge, MA, USA, 2006) p. 248.

[55] T. Tao, V. Vu, and M. Krishnapur, Random matrices: Universality of esds and the circular law, Ann. Probab. **38**, 2023 (2010).

[56] R. S. Ellis, An overview of the theory of large deviations and applications to statistical mechanics, Scand. Actuar. J. **1995**, 97 (1995), https://doi.org/10.1080/03461238.1995.10413952.

Phase Space Analysis of Chaotic Neural Networks
(Supplemental Material)

Jakob Stubenrauch,[1,2] Christian Keup,[1,2] Anno C. Kurth,[1,2] Moritz Helias,[1,2] and Alexander van Meegen[1,3]

[1]*Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and*
JARA-Institute Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany
[2]*Department of Physics, Faculty 1, RWTH Aachen University, Aachen, Germany*
[3]*Institute of Zoology, University of Cologne, 50674 Cologne, Germany*
(Dated: October 17, 2022)

CONTENTS

A. Distribution of fixed points

1. Kac-Rice formula

Counting roots To find the distribution of fixed points we first need to establish how to count roots of a field $\boldsymbol{y}(\boldsymbol{x})$. The number of roots in a volume V is given by the Kac-Rice formula [1–3]

$$N_{\mathrm{fp}}(V) = \int_V d\boldsymbol{x}\, \delta\left[\boldsymbol{y}(\boldsymbol{x})\right] |\det \boldsymbol{y}'(\boldsymbol{x})| \tag{1}$$

where $\left[\boldsymbol{y}'(\boldsymbol{x})\right]_{ij} = \partial_{x_j} y_i(\boldsymbol{x})$ is the Jacobian of the field $\boldsymbol{y}(\boldsymbol{x})$. The Dirac delta $\delta\left[\boldsymbol{y}(\boldsymbol{x})\right]$ hits at every root of $\boldsymbol{y}(\boldsymbol{x})$ and the Jacobian determinant assures that every root contributes 1 to the total number of roots in N_0. The latter is necessary due to the scaling property of the Dirac delta

$$\int_V d\boldsymbol{x}\, \delta\left[\boldsymbol{y}(\boldsymbol{x})\right] h(\boldsymbol{x}) = \sum_{\boldsymbol{x}^* \in V \text{ s.t. } \boldsymbol{y}(\boldsymbol{x}^*)=0} \frac{h(\boldsymbol{x}^*)}{|\det \boldsymbol{y}'(\boldsymbol{x}^*)|}. \tag{2}$$

Fixed point distribution Since the counting formula Eq. (1) holds for every Borel set $V \subset \mathbb{R}^N$ [3], the integrand in Eq. (1) can be interpreted as a distribution of roots. On average over realizations $\mathbf{J}$ and $\boldsymbol{\eta}$, this distribution is

$$\rho(\boldsymbol{x}) = \langle \delta\left[\boldsymbol{y}(\boldsymbol{x})\right] |\det \boldsymbol{y}'(\boldsymbol{x})| \rangle_{\mathbf{J},\boldsymbol{\eta}}. \tag{3}$$

Its norm $N_{\mathrm{fp}} \equiv \left\langle N_{\mathrm{fp}}(\mathbb{R}^N) \right\rangle = \int_{\mathbb{R}^N} d\boldsymbol{x}\, \rho(\boldsymbol{x})$ is the expected total number of fixed points.

Joint distribution of velocity and Jacobian The expectation value in Eq. (3) can be simplified by transforming from the realization parameters $\mathbf{J}$ and $\boldsymbol{\eta}$ to the velocity

$$\boldsymbol{y}(\boldsymbol{x}) \equiv \dot{\boldsymbol{x}} = -\boldsymbol{x} + \mathbf{J}\phi(\boldsymbol{x}) + \boldsymbol{\eta} \tag{4}$$

and Jacobian

$$\boldsymbol{y}'(\boldsymbol{x}) = -\mathbb{1} + \mathbf{J}\operatorname{diag}\left[\phi'(\boldsymbol{x})\right] \tag{5}$$

as random fields. Here, $J_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, g^2/N)$ and $\eta_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, D)$ where we use the notation $\chi_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(a, b)$ to denote that the random variables χ_i are independently and identically distributed (i.i.d.) by the Gaussian distribution with mean a and variance b. The joint probability distribution of $\boldsymbol{y}(\boldsymbol{x})$ and $\boldsymbol{y}'(\boldsymbol{x})$ follows by computing their first moments

$$\langle y_i(\boldsymbol{x})\rangle = -x_i \equiv \mu_i(\boldsymbol{x}), \qquad \langle [\boldsymbol{y}'(\boldsymbol{x})]_{ik}\rangle = -\delta_{ik} \equiv [\boldsymbol{\mu}_i(\boldsymbol{x})]_k, \tag{6}$$

and second cumulants

$$\langle\langle y_i(\boldsymbol{x})\,[\boldsymbol{y}'(\boldsymbol{x})]_{jk}\rangle\rangle = \frac{g^2}{N}\delta_{ij}\phi(x_k)\phi'(x_k) \equiv \delta_{ij}[\boldsymbol{k}(\boldsymbol{x})]_k, \quad \langle\langle [\boldsymbol{y}'(\boldsymbol{x})]_{ik}\,[\boldsymbol{y}'(\boldsymbol{x})]_{jl}\rangle\rangle = \delta_{ij}\delta_{kl}\frac{g^2}{N}\phi'(x_k)^2 \equiv \delta_{ij}[\mathbf{K}(\boldsymbol{x})]_{kl}. \tag{7}$$

Since all higher cumulants vanish, the joint distribution of $\boldsymbol{y}$ and $\boldsymbol{y}'$ is Gaussian. Furthermore, it factorizes into velocity components y_i and respective gradients (rows of the Jacobian) $\nabla y_i = \boldsymbol{y}'_i$. Hence, we can write the probability distribution of $\boldsymbol{y}$ and $\boldsymbol{y}'$ at each point $\boldsymbol{x}$ as $p_{\boldsymbol{x}}(\boldsymbol{y}, \boldsymbol{y}') = \prod_{i=1}^N p_{\boldsymbol{x}}(y_i, \boldsymbol{y}'_i)$ with

$$\begin{pmatrix} y_i \\ \boldsymbol{y}'_i \end{pmatrix} \overset{\text{i.i.d.}}{\sim} \mathcal{N}\left[\begin{pmatrix} \mu_i(\boldsymbol{x}) \\ \boldsymbol{\mu}_i(\boldsymbol{x}) \end{pmatrix}, \begin{pmatrix} [\kappa(\boldsymbol{x}) + D] & \boldsymbol{k}(\boldsymbol{x})^{\mathrm{T}} \\ \boldsymbol{k}(\boldsymbol{x}) & \mathbf{K}(\boldsymbol{x}) \end{pmatrix}\right]. \tag{8}$$

A different formulation of the Kac-Rice formula [3, 4] gives the distribution of crossings of a Gaussian random process $(\boldsymbol{y}, \boldsymbol{y}')$ through $\boldsymbol{y} = 0$ in terms of the joint probability as

$$\rho(\boldsymbol{x}) = \int d\boldsymbol{y}'\, p_{\boldsymbol{x}}(\boldsymbol{y} = 0, \boldsymbol{y}')\, |\det \boldsymbol{y}'|. \tag{9}$$

Since the formulation in terms of the Dirac delta, Eq. (3), is self-explaining and Eq. (9) is harder to motivate from scratch, we show their equivalence by explicit substitution from $\mathbf{J}$ and $\boldsymbol{\eta}$ to $\boldsymbol{y}$ and $\boldsymbol{y}'$. Solving Eqs. (4) and (5) for $\mathbf{J}$ and $\boldsymbol{\eta}$ yields

$$\begin{bmatrix} \mathbf{J} \\ \boldsymbol{\eta} \end{bmatrix}(\boldsymbol{y}, \boldsymbol{y}') = \begin{bmatrix} (\mathbb{1} + \boldsymbol{y}')\operatorname{diag}\left[1/\phi(\boldsymbol{x})'\right] \\ \boldsymbol{y} + \boldsymbol{x} - (\mathbb{1} + \boldsymbol{y}')\operatorname{diag}\left[1/\phi(\boldsymbol{x})'\right]\phi(\boldsymbol{x}) \end{bmatrix}. \tag{10}$$

The Jacobian of the substitution is

$$D_{\begin{bmatrix} \boldsymbol{y}' \\ \boldsymbol{y} \end{bmatrix}}\begin{bmatrix} \mathbf{J} \\ \boldsymbol{\eta} \end{bmatrix} = \begin{bmatrix} D_{\boldsymbol{y}'}\mathbf{J} & D_{\boldsymbol{y}}\mathbf{J} \\ D_{\boldsymbol{y}'}\boldsymbol{\eta} & D_{\boldsymbol{y}}\boldsymbol{\eta} \end{bmatrix} = \begin{bmatrix} \operatorname{diag}\left[1/\phi(\boldsymbol{x})'\right] & & & \\ & \ddots & & \mathbf{0} \\ & & \operatorname{diag}\left[1/\phi(\boldsymbol{x})'\right] & \\ & D_{\boldsymbol{y}'}\boldsymbol{\eta} & & \mathbb{1} \end{bmatrix} \tag{11}$$

where the columns of $\mathbf{J}$ are stacked to a vector to get the usual rule for substitution. Due to the upper right block being zero, the lower left block does not contribute to the determinant. The integration in Eq. (3) can hence be written

$$\rho(\boldsymbol{x}) = \langle \delta\left[\boldsymbol{y}(\boldsymbol{x})\right] |\det \boldsymbol{y}'(\boldsymbol{x})|\rangle_{\mathbf{J}, \boldsymbol{\eta}} \tag{12}$$

$$= \int d\mathbf{J}\, d\boldsymbol{\eta}\, p_g(\mathbf{J})\, p_D(\boldsymbol{\eta})\, \delta\left[\boldsymbol{y}(\boldsymbol{x})\right] |\det \boldsymbol{y}'(\boldsymbol{x})| \tag{13}$$

$$= \int d\boldsymbol{y}'\, d\boldsymbol{y}\, \left|\prod_{i=1}^N \frac{1}{\phi'(x_i)}\right|^N p_g\left[\mathbf{J}(\boldsymbol{y}')\right] p_D\left[\boldsymbol{\eta}(\boldsymbol{y}, \boldsymbol{y}')\right] \delta\left[\boldsymbol{y}(\boldsymbol{x})\right] |\det \boldsymbol{y}'| \tag{14}$$

where $p_g(\mathbf{X}) = \prod_{ij} \mathcal{N}(J_{ij}|0, g^2/N)$ and $p_D(\boldsymbol{x}) = \prod_i \mathcal{N}(x_i|0, D)$ and $\mathcal{N}(a|b, c)$ denotes the Gaussian distribution function with random variable a, mean value b and variance c.

By plugging Eq. (10) into the first part of the integrand in Eq. (14), one can see that

$$\left| \prod_{i=1}^{N} \frac{1}{\phi'(x_i)} \right|^N p_g \left[\mathbf{J}(\boldsymbol{y}') \right] = \prod_i \mathcal{N} \left[\boldsymbol{y}'_i \middle| \boldsymbol{\mu}_i(\boldsymbol{x}), \mathbf{K}(\boldsymbol{x}) \right] \tag{15}$$

with $\boldsymbol{\mu}$, $\mathbf{K}$ as in Eqs. (6) and (7). Equation (15) is the marginal distribution of the Jacobian $\boldsymbol{y}'$, since marginalizing a Gaussian distribution does not alter diagonal cumulants. A similar procedure, and using the formula for the conditional probability of a Gaussian [5, Appendix A2]

$$p_{\boldsymbol{x}}(y_i|\boldsymbol{y}'_i) = \mathcal{N} \left[y_i \middle| \mu_i + \boldsymbol{k}^{\mathrm{T}} \mathbf{K}^{-1} \left(\boldsymbol{y}'_i - \boldsymbol{\mu}_i \right), \kappa + D - \boldsymbol{k}^{\mathrm{T}} \mathbf{K}^{-1} \boldsymbol{k} \right], \tag{16}$$

allows to identify

$$\int d\boldsymbol{y} \, p_D \left[\boldsymbol{\eta}(\boldsymbol{y}, \boldsymbol{y}') \right] \delta \left[\boldsymbol{y}(\boldsymbol{x}) \right] = p_{\boldsymbol{x}}(y_i = 0|\boldsymbol{y}'_i) \tag{17}$$

Putting all together, we get

$$\rho(\boldsymbol{x}) = \int d\boldsymbol{y}' p_{\boldsymbol{x}}(\boldsymbol{y} = 0|\boldsymbol{y}') p_{\boldsymbol{x}}(\boldsymbol{y}') \left| \det \boldsymbol{y}' \right|, \tag{18}$$

which, using Bayes' law

$$p_{\boldsymbol{x}}(\boldsymbol{y}, \boldsymbol{y}') = p_{\boldsymbol{x}}(\boldsymbol{y}') \, p_{\boldsymbol{x}}(\boldsymbol{y}|\boldsymbol{y}') = p_{\boldsymbol{x}}(\boldsymbol{y}) \, p_{\boldsymbol{x}}(\boldsymbol{y}'|\boldsymbol{y}), \tag{19}$$

gives Eq. (9).

Condition on the level Using the rightmost dissection in Eq. (19), i.e. condition on the level $\boldsymbol{y}$, is advantageous since the *level probability*

$$p_{\mathrm{L}}(\boldsymbol{x}) \equiv p_{\boldsymbol{x}}(\boldsymbol{y} = 0) = \frac{1}{\sqrt{2\pi \left[\kappa(\boldsymbol{x}) + D \right]}^N} e^{-\frac{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}{2[\kappa(\boldsymbol{x})+D]}}, \tag{20}$$

where $\kappa(\boldsymbol{x}) = (g^2/N)\phi(\boldsymbol{x})^{\mathrm{T}}\phi(\boldsymbol{x})$, can be taken out of the integral, i.e.

$$\rho(\boldsymbol{x}) = p_{\mathrm{L}}(\boldsymbol{x}) \left\langle \left| \det \boldsymbol{y}' \right| \right\rangle_{\boldsymbol{y}' \sim p_{\boldsymbol{x}}(\boldsymbol{y}'|\boldsymbol{y}=0)}. \tag{21}$$

The level probability denotes the marginal probability of the field $\boldsymbol{y}(\boldsymbol{x})$ to be zero at a point $\boldsymbol{x}$. This alone is not the fixed point distribution since the distribution of the Jacobian (think of it as the slope) controls how often the field $\boldsymbol{y}(\boldsymbol{x})$ can cross the zero level when it is close to zero.

Jacobian distribution To compute the expected determinant in Eq. (21), we have to characterize the conditional distribution $p_{\boldsymbol{x}}(\boldsymbol{y}'|\boldsymbol{y} = 0)$ first. The conditional distribution of a Gaussian is again a Gaussian. Its moments are given by [5, Appendix A2]

$$p_{\boldsymbol{x}}(\boldsymbol{y}'|\boldsymbol{y} = 0) = \prod_i \mathcal{N} \left[\boldsymbol{y}'_i | \mathbf{M}_i(\boldsymbol{x}), \mathbf{C}(\boldsymbol{x}) \right] \tag{22}$$

where the mean conditioned Jacobian is

$$M_{ij}(\boldsymbol{x}) = -\delta_{ij} - \frac{g^2}{N} \frac{x_i \phi(x_j) \phi'(x_j)}{\kappa(\boldsymbol{x}) + D} \tag{23}$$

and each row $\boldsymbol{y}'_i$ has the same covariance matrix

$$C_{nm} = \delta_{nm} \frac{g^2}{N} \phi'(x_n)^2 - \frac{g^4}{N^2} \frac{\phi(x_n)\phi'(x_n)\phi(x_m)\phi'(x_m)}{\kappa(\boldsymbol{x}) + D} \tag{24}$$

It will prove useful to write the conditioned Jacobian in the compact form

$$\boldsymbol{y}'(\boldsymbol{y} = 0) = \mathbf{M} + \mathbf{X}\boldsymbol{\Sigma} \tag{25}$$

with $X_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1/N)$ and $\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma} \equiv N\mathbf{C}$. To see this, lets verify that it has the correct moments (here we suppress the dependency $(\boldsymbol{y} = 0)$):

$$\langle y'_{ij} \rangle = M_{ij}$$

$$\langle y'_{ij} y'_{kl} \rangle - \langle y'_{ij} \rangle \langle y'_{kl} \rangle = \left\langle \sum_{n,m=1}^{N} X_{in} \Sigma_{nj} X_{km} \Sigma_{ml} \right\rangle$$

$$= \sum_{n,m=1}^{N} \frac{1}{N} \delta_{ik} \delta_{nm} \Sigma_{nj} \Sigma_{ml} \qquad (26)$$

$$= \delta_{ik} \frac{1}{N} \sum_{n} \Sigma_{jn}^{\mathrm{T}} \Sigma_{nl} = \delta_{ik} C_{jl}$$

Thus, as needed, the different rows are uncorrelated (δ_{ik}) and within a row the correlation structure is given by $\mathbf{C}$. The scaled square root $\mathbf{\Sigma}$ of the covariance matrix

$$\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma} = N\boldsymbol{C} = \boldsymbol{\Lambda}(\mathbb{1} - \boldsymbol{v}\boldsymbol{v}^{\mathrm{T}})\boldsymbol{\Lambda},$$
$$\boldsymbol{\Lambda} = g \operatorname{diag}[\phi'(\boldsymbol{x})], \qquad (27)$$
$$v_i = \frac{g}{\sqrt{N(\kappa + D)}} \phi(x_i)$$

can be computed by using

$$\sqrt{\mathbb{1} - \boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}} = \mathbb{1} - \frac{\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}}{1 + \sqrt{1 - \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}}}, \qquad (28)$$

which can be checked for general vectors $\boldsymbol{v}$. We find

$$\mathbf{\Sigma}(\boldsymbol{x}) = \left[\mathbb{1} - \frac{g^2}{N} \frac{\phi(\boldsymbol{x})\phi(\boldsymbol{x})^{\mathrm{T}}}{\kappa(\boldsymbol{x}) + D\left[1 + \sqrt{1 + \kappa(\boldsymbol{x})/D}\right]} \right] \boldsymbol{\Lambda}(\boldsymbol{x}) \quad . \qquad (29)$$

Concluding, the fixed point distribution can be expressed as

$$\rho(\boldsymbol{x}) = p_{\mathrm{L}}(\boldsymbol{x}) \left\langle |\det[\mathbf{M}(\boldsymbol{x}) + \mathbf{X}\mathbf{\Sigma}(\boldsymbol{x})]| \right\rangle_{X_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1/N)}. \qquad (30)$$

Note that both $\mathbf{M}(\boldsymbol{x})$ and $\mathbf{\Sigma}(\boldsymbol{x})$ consist of a sum of a diagonal and a rank one matrix.

2. Determinant of a non-hermitian random matrix

We want to compute the expectation of the modulus determinant in Eq. (30), i.e. of the matrix

$$\boldsymbol{y}'(\boldsymbol{x}) = \mathbf{M}(\boldsymbol{x}) + \mathbf{X}\mathbf{\Sigma}(\boldsymbol{x}), \quad X_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1/N). \qquad (31)$$

The determinant of a matrix is the product of its eigenvalues. For matrices of the form Eq. (31), but having complex entries, it is known that the eigenvalue spectrum is self-averaging [6, Theorem 1.14].. I.e., assuming self-averaging also in our case where the entries are real, we can use

$$\zeta(\boldsymbol{x}) = \frac{1}{N} \ln \langle |\det \boldsymbol{y}'(\boldsymbol{x})| \rangle = -\frac{1}{N} \ln \left\langle |\det \boldsymbol{y}'(\boldsymbol{x})|^{-1} \right\rangle. \qquad (32)$$

Determinant as Gaussian integral To compute ζ, we build on the approach by [7]. First, we use $\det \mathbf{A}^{\mathrm{T}} = \det \mathbf{A}$ and $\det(\mathbf{A}\mathbf{B}) = \det \mathbf{A} \det \mathbf{B}$ to rewrite Eq. (32) as

$$\zeta = -\frac{1}{N} \ln \left\langle \frac{1}{\sqrt{\det(\boldsymbol{y}'^{\mathrm{T}}\boldsymbol{y}' + \varepsilon\mathbb{1})}} \right\rangle \qquad (33)$$

where we also added a diagonal matrix with (infinitesimal) strength ε to avoid divergences. We note that $\boldsymbol{y'}^{\mathrm{T}}\boldsymbol{y'}$ is positive semi-definite, and thus $\boldsymbol{y'}^{\mathrm{T}}\boldsymbol{y'} + \varepsilon\mathbb{1}$ is positive definite, so we can use the Gaussian integral $\int d\boldsymbol{z}\, e^{-\frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\mathbf{A}\boldsymbol{z}} = \sqrt{\frac{(2\pi)^N}{\det\mathbf{A}}}$ to get

$$\zeta = -\frac{1}{N}\ln\left\langle\int d\boldsymbol{z}\, e^{-\frac{1}{2}\boldsymbol{z}^{\mathrm{T}}(\boldsymbol{y'}^{\mathrm{T}}\boldsymbol{y'}+\varepsilon\mathbb{1})\boldsymbol{z}}\right\rangle + \frac{1}{2}\ln(2\pi). \tag{34}$$

The exponent in the expectation value in Eq. (34) is at most quadratic in the random variables X_{ij}. Hence, it can be computed as a Gaussian integral and we arrive at

$$\zeta = -\frac{1}{N}\ln\int d\boldsymbol{z}\, e^{-\frac{1}{2}\varepsilon\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z} - \frac{\boldsymbol{z}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\mathbf{M}\boldsymbol{z}}{2(1+\frac{1}{N}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{z})} - \frac{N}{2}\ln(1+\frac{1}{N}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{z})} + \frac{1}{2}\ln(2\pi). \tag{35}$$

Collective variables   We evaluate the integral in saddle point approximation. To this end, we introduce the collective variables

$$m = \frac{1}{N}\boldsymbol{z}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\mathbf{M}\boldsymbol{z}, \qquad \sigma = \frac{1}{N}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{z} \tag{36}$$

Denoting all collective variables as $\boldsymbol{\theta}$, we arrive at

$$\zeta = -\frac{1}{N}\ln\int d\boldsymbol{\theta}\, e^{-N[\frac{m}{2(1+\sigma)} + \frac{1}{2}\ln(1+\sigma)]}\int d\boldsymbol{z}\, e^{-\frac{1}{2}\varepsilon\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}}\delta(\boldsymbol{\theta} - \boldsymbol{\theta}(\boldsymbol{z})) + \frac{1}{2}\ln(2\pi).$$

We can interpret

$$p(\boldsymbol{\theta}) = \left(\frac{\varepsilon}{2\pi}\right)^{N/2}\int d\boldsymbol{z}\, e^{-\frac{1}{2}\varepsilon\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}}\delta(\boldsymbol{\theta} - \boldsymbol{\theta}(\boldsymbol{z}))$$

as the density of the transformed random variable $\boldsymbol{\theta} = \boldsymbol{\theta}(\boldsymbol{z})$ where $x_i \overset{i.i.d.}{\sim} \mathcal{N}(0, 1/\varepsilon)$. The corresponding characteristic function is

$$\Phi(\tilde{\boldsymbol{\theta}}) = \left(\frac{\varepsilon}{2\pi}\right)^{N/2}\int d\boldsymbol{z}\, e^{-\frac{1}{2}\varepsilon\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z} + i\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\theta}(\boldsymbol{z})} = \sqrt{\frac{\varepsilon^N}{\det(\varepsilon\mathbb{1} - 2i\tilde{m}\frac{1}{N}\mathbf{M}^{\mathrm{T}}\mathbf{M} - 2i\tilde{\sigma}\frac{1}{N}\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{\Sigma})}}$$

such that we can write

$$p(\boldsymbol{\theta}) = \frac{N^2}{(2\pi)^2}\int d\tilde{\boldsymbol{\theta}}\, e^{-N[i\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\theta} - \Omega(\tilde{\boldsymbol{\theta}})]}$$

with a scaled cumulant generating function $\Omega(\tilde{\boldsymbol{\theta}}) = \frac{1}{N}\ln\Phi(N\tilde{\boldsymbol{\theta}})$ which is given by

$$\Omega(\tilde{\boldsymbol{\theta}}) = -\frac{1}{2N}\ln\det(\varepsilon\mathbb{1} - 2i\tilde{m}\mathbf{M}^{\mathrm{T}}\mathbf{M} - 2i\tilde{\sigma}\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{\Sigma}) + \frac{1}{2}\ln\varepsilon.$$

Inserting $p(\boldsymbol{\theta})$, we arrive at

$$\zeta = -\frac{1}{N}\ln\int d\boldsymbol{\theta}\int d\tilde{\boldsymbol{\theta}}\, e^{-NH(\boldsymbol{\theta},\tilde{\boldsymbol{\theta}})} + O(N^{-1}\ln N), \tag{37}$$

$$H(\boldsymbol{\theta},\tilde{\boldsymbol{\theta}}) = \frac{m}{2(1+\sigma)} + \frac{1}{2}\ln(1+\sigma) + i\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\theta} + \frac{1}{2N}\ln\det(\varepsilon\mathbb{1} - 2i\tilde{m}\mathbf{M}^{\mathrm{T}}\mathbf{M} - 2i\tilde{\sigma}\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{\Sigma}). \tag{38}$$

Note that the terms containing ε or 2π cancel. Now, we would like to perform a saddle-point approximation to get $\zeta = H(\boldsymbol{\theta}_*, \tilde{\boldsymbol{\theta}}_*) + \mathcal{O}(N^{-1}\ln N)$ where $\boldsymbol{\theta}_*$ and $\tilde{\boldsymbol{\theta}}_*$ minimize $H(\boldsymbol{\theta},\tilde{\boldsymbol{\theta}})$.

Saddle-point approximation   The corresponding saddle-point equations are

$$-2i\tilde{m} = \frac{1}{1+\sigma}, \qquad -2i\tilde{\sigma} = -\frac{m}{(1+\sigma)^2} + \frac{1}{1+\sigma} = \frac{1+\sigma-m}{(1+\sigma)^2}, \tag{39}$$

$$-2im = \frac{1}{N}\partial_{\tilde{m}}\ln\det(\varepsilon\mathbb{1} - 2i\tilde{m}\mathbf{M}^{\mathrm{T}}\mathbf{M} - 2i\tilde{\sigma}\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}), \qquad -2i\sigma = \frac{1}{N}\partial_{\tilde{\sigma}}\ln\det(\varepsilon\mathbb{1} - 2i\tilde{m}\mathbf{M}^{\mathrm{T}}\mathbf{M} - 2i\tilde{\sigma}\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}).$$

To evaluate the derivatives, we use $\partial_x \ln\det(x\mathbf{X} + y\mathbf{Y}) = \mathrm{tr}((x\mathbf{X} + y\mathbf{Y})^{-1}\mathbf{X})$ to obtain

$$m = \frac{1}{N}\mathrm{tr}\Big[(\varepsilon\mathbb{1} - 2i\tilde{m}\mathbf{M}^{\mathrm{T}}\mathbf{M} - 2i\tilde{\sigma}\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma})^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{M}\Big], \qquad \sigma = \frac{1}{N}\mathrm{tr}\Big[(\varepsilon\mathbb{1} - 2i\tilde{m}\mathbf{M}^{\mathrm{T}}\mathbf{M} - 2i\tilde{\sigma}\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma})^{-1}\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\Big]. \tag{40}$$

We can combine Eqs. (39) and (40) to

$$m = \frac{1}{N}\mathrm{tr}\Big[(\varepsilon\mathbb{1} + (1+\sigma)^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{M} + (1+\sigma-m)(1+\sigma)^{-2}\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma})^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{M}\Big], \tag{41}$$

$$\sigma = \frac{1}{N}\mathrm{tr}\Big[(\varepsilon\mathbb{1} + (1+\sigma)^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{M} + (1+\sigma-m)(1+\sigma)^{-2}\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma})^{-1}\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\Big]. \tag{42}$$

Putting it all together, we arrive at

$$\zeta = -\frac{\sigma_*(1+\sigma_*-m_*)}{2(1+\sigma_*)^2} + \frac{1}{2N}\ln\det[\varepsilon(1+\sigma_*)\mathbb{1} + \mathbf{M}^{\mathrm{T}}\mathbf{M} + (1+\sigma_*-m_*)(1+\sigma_*)^{-1}\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}] + \mathcal{O}(N^{-1}\ln N) \tag{43}$$

where we used $i\tilde{m}_*m_* = -\frac{m_*}{2(1+\sigma_*)}$ and $i\tilde{\sigma}_*\sigma_* = -\frac{\sigma_*(1+\sigma_*-m_*)}{2(1+\sigma_*)^2}$ due to Eq. (39).

Diagonal case   We first consider the case where $\mathbf{M} = -\mathbb{1}$ and $\mathbf{\Sigma} = \mathbf{\Lambda} \equiv g\,\mathrm{diag}[\phi'(\boldsymbol{x})]$, i.e., we neglect the rank 1 parts. As we will see later, this already yields the leading order contribution to the determinant.

The resulting saddle-point equations are

$$m = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\varepsilon + \frac{1}{1+\sigma} + \frac{1+\sigma-m}{(1+\sigma)^2}\lambda_i^2}, \qquad \sigma = \frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_i^2}{\varepsilon + \frac{1}{1+\sigma} + \frac{1+\sigma-m}{(1+\sigma)^2}\lambda_i^2},$$

and the determinant is

$$\zeta = -\frac{\sigma_*(1+\sigma_*-m_*)}{2(1+\sigma_*)^2} + \frac{1}{2N}\sum_{i=1}^{N}\ln[\varepsilon(1+\sigma_*) + 1 + (1+\sigma_*-m_*)(1+\sigma_*)^{-1}\lambda_i^2] + \mathcal{O}(N^{-1}\ln N).$$

For $\varepsilon \to 0$, $m, \sigma \to \infty$ is a possible solution to the saddle-point equations.

Assuming the diverging solution is the relevant one, the determinant simplifies to

$$\zeta = -\frac{1}{2}\Big(1 - \frac{m_*}{\sigma_*}\Big) + \frac{1}{2N}\sum_{i=1}^{N}\ln\Big[\varepsilon\sigma_* + 1 + \Big(1 - \frac{m_*}{\sigma_*}\Big)\lambda_i^2\Big] + \mathcal{O}(N^{-1}\ln N).$$

The saddle-point equations determine $\frac{m}{\sigma}$ and $\varepsilon\sigma$ via

$$\frac{m}{\sigma} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\varepsilon\sigma + 1 + (1 - \frac{m}{\sigma})\lambda_i^2}, \qquad 1 = \frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_i^2}{\varepsilon\sigma + 1 + (1 - \frac{m}{\sigma})\lambda_i^2}$$

which can be combined to $(\varepsilon\sigma + 1)\frac{m}{\sigma} + (1 - \frac{m}{\sigma}) = 1$. The latter equation leads to $\varepsilon\sigma = 0$, thus the saddle-point equations reduce further to

$$\frac{m}{\sigma} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{1 + (1 - \frac{m}{\sigma})\lambda_i^2}, \qquad 1 = \frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_i^2}{1 + (1 - \frac{m}{\sigma})\lambda_i^2}.$$

These two equations are equivalent; for convenience we choose the second and introduce $z = 1 - \frac{m}{\sigma}$ which is determined by

$$1 = \frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_i^2}{1 + z\lambda_i^2}. \tag{44}$$

This equation needs to be solved numerically. In terms of the solution z_*, the determinant follows from

$$\zeta = -\frac{1}{2}z_* + \frac{1}{2N}\sum_{i=1}^{N}\ln(1 + z_*\lambda_i^2) + \mathcal{O}(N^{-1}\ln N). \tag{45}$$

The result can numerically be shown to be equivalent to results based on the spectral density from Ahmadian et al. [8].

Exact determinant Here, we compute the determinant for general matrices $\mathbf{M}$, $\mathbf{\Sigma}$ to cover the case beyond the diagonal approximation, including the rank 1 terms.

If we assume that $m, \sigma \to \infty$ is also in the general case the relevant solution, the determinant simplifies to

$$\zeta = -\frac{1}{2}\left(1 - \frac{m_*}{\sigma_*}\right) + \frac{1}{2N}\ln\det\left[\varepsilon\sigma_*\mathbb{1} + \mathbf{M}^{\mathrm{T}}\mathbf{M} + \left(1 - \frac{m_*}{\sigma_*}\right)\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\right] + \mathcal{O}(N^{-1}\ln N).$$

The saddle-point equations for the relevant quantities are

$$\frac{m}{\sigma} = \frac{1}{N}\mathrm{tr}\left[\left(\varepsilon\sigma\mathbb{1} + \mathbf{M}^{\mathrm{T}}\mathbf{M} + \left(1 - \frac{m}{\sigma}\right)\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\right)^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{M}\right],$$

$$1 = \frac{1}{N}\mathrm{tr}\left[\left(\varepsilon\sigma\mathbb{1} + \mathbf{M}^{\mathrm{T}}\mathbf{M} + \left(1 - \frac{m}{\sigma}\right)\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\right)^{-1}\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\right].$$

We can combine both equations to

$$\frac{m}{\sigma} + \left(1 - \frac{m}{\sigma}\right) = \frac{1}{N}\mathrm{tr}\left[\left(\varepsilon\sigma\mathbb{1} + \mathbf{M}^{\mathrm{T}}\mathbf{M} + \left(1 - \frac{m}{\sigma}\right)\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\right)^{-1}\left(\mathbf{M}^{\mathrm{T}}\mathbf{M} + \left(1 - \frac{m}{\sigma}\right)\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\right)\right]$$

which is fulfilled for $\varepsilon\sigma = 0$. For $\varepsilon\sigma = 0$, the remaining saddle-point equations are equivalent; hence, we opt again for the second one and introduce $z = 1 - \frac{m}{\sigma}$ which obeys

$$1 = \frac{1}{N}\mathrm{tr}\left[\left(\mathbf{M}^{\mathrm{T}}\mathbf{M} + z\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\right)^{-1}\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\right]. \tag{46}$$

In terms of the solution z_*, the determinant is given by

$$\zeta = -\frac{1}{2}z_* + \frac{1}{2N}\ln\det(\mathbf{M}^{\mathrm{T}}\mathbf{M} + z_*\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}) + \mathcal{O}(N^{-1}\ln N). \tag{47}$$

Leading order determinant Equations (46) and (47) are exact. However, they require computing the trace and the determinant of a matrix. To capture the $N \gg 1$ limit, we express the equations in terms of network sums $\sum_{i=1}^{N} f(x_i)$. Here, we will also see that to leading order they correspond to the diagonal case above.

To this end, we first need an expression in terms of network sums of

$$\ln\det\left(\mathbf{M}^{\mathrm{T}}\mathbf{M} + z_*\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}\right). \tag{48}$$

We recall Eq. (23)

$$M_{ij}(\boldsymbol{x}) = -\delta_{ij} - \frac{g^2}{N}\frac{x_i\phi(x_j)\phi'(x_j)}{\kappa(\boldsymbol{x}) + D} \tag{49}$$

$$= -\left(\mathbb{1} + \frac{1}{\kappa + D}\boldsymbol{x}\boldsymbol{k}^{\mathrm{T}}\right) \tag{50}$$

where $[\boldsymbol{k}(\boldsymbol{x})]_k = \frac{g^2}{N}\phi(x_k)\phi'(x_k)$, and

$$\mathbf{\Sigma}(\boldsymbol{x}) = \left[\mathbb{1} - \frac{g^2}{N}\frac{\phi(\boldsymbol{x})\phi(\boldsymbol{x})^{\mathrm{T}}}{\kappa(\boldsymbol{x}) + D\left(1 + \sqrt{1 + \kappa(\boldsymbol{x})/D}\right)}\right]\mathbf{\Lambda}(\boldsymbol{x}) \tag{51}$$

$$\mathbf{\Lambda}(\boldsymbol{x}) = g\,\mathrm{diag}[\phi'(\boldsymbol{x})].$$

Hence, the squares are

$$\mathbf{M}^{\mathrm{T}}\mathbf{M} = \mathbb{1} + \frac{1}{\kappa + D}\left(\boldsymbol{k}\boldsymbol{x}^{\mathrm{T}} + \boldsymbol{x}\boldsymbol{k}^{\mathrm{T}} + \frac{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}{\kappa + D}\boldsymbol{k}\boldsymbol{k}^{\mathrm{T}}\right), \tag{52}$$

and

$$\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma} = N\left[\mathbf{K} - \frac{1}{\kappa + D}\boldsymbol{k}\boldsymbol{k}^{\mathrm{T}}\right] \tag{53}$$

where $[\mathbf{K}(\boldsymbol{x})]_{kl} = \delta_{kl}\frac{g^2}{N}\phi'(x_k)^2$.

Thus, we need the determinant of a matrix of the form

$$\mathbf{A} = \mathbf{D} + \boldsymbol{a}\boldsymbol{x}^{\mathrm{T}} + \boldsymbol{x}\boldsymbol{a}^{\mathrm{T}} + \boldsymbol{b}\boldsymbol{b}^{\mathrm{T}} \tag{54}$$

where $\mathbf{D} = \mathbb{1} + z_* N\mathbf{K}$ is diagonal, $\boldsymbol{a} = \frac{1}{\kappa+D}\boldsymbol{k}$, $\boldsymbol{b} = \sqrt{\frac{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}{(\kappa+D)^2} - \frac{Nz_*}{\kappa+D}}\boldsymbol{k} = \sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} - (\kappa+D)\,Nz_*}\,\boldsymbol{a} \equiv \sqrt{\vartheta(\boldsymbol{x})}\boldsymbol{a}$. Threefold application of the matrix determinant lemma yields

$$\det\mathbf{A} = \left(1 + \boldsymbol{b}^{\mathrm{T}}\left(\mathbf{D} + \boldsymbol{a}\boldsymbol{x}^{\mathrm{T}} + \boldsymbol{x}\boldsymbol{a}^{\mathrm{T}}\right)^{-1}\boldsymbol{b}\right)\times \tag{55}$$

$$\times\left(1 + \boldsymbol{x}^{\mathrm{T}}\left(\mathbf{D} + \boldsymbol{x}\boldsymbol{a}^{\mathrm{T}}\right)^{-1}\boldsymbol{a}\right)\left(1 + \boldsymbol{a}^{\mathrm{T}}\mathbf{D}^{-1}\boldsymbol{x}\right)\det\mathbf{D} \tag{56}$$

The inverse matrices can be computed with the Sherman Morrison formula

$$\left(\mathcal{A} + \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}\right)^{-1} = \mathcal{A}^{-1} - \frac{\mathcal{A}^{-1}\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}\mathcal{A}^{-1}}{1 + \boldsymbol{v}^{\mathrm{T}}\mathcal{A}^{-1}\boldsymbol{u}} \tag{57}$$

where $\mathcal{A}$ is an invertible square matrix and $\boldsymbol{u}$ and $\boldsymbol{v}$ are column vectors. Up to threefold application of the Sherman Morrison formula and some simplifications lead to

$$\zeta = -\frac{1}{2}z_* + \frac{1}{N}\sum_{i=1}^{N}\ln\left[1 + z_*g^2\phi'(x_i)^2\right] + \frac{1}{N}\ln\left[(1+\beta)^2 + \alpha(\vartheta - \gamma)\right] \tag{58}$$

where

$$\alpha(\boldsymbol{x}) = \frac{g^4}{[\kappa(\boldsymbol{x}) + D]^2}\frac{1}{N^2}\sum_{i=1}^{N}\frac{\phi(x_i)^2\phi'(x_i)^2}{1 + z_*g^2\phi'(x_i)^2} \tag{59}$$

$$\beta(\boldsymbol{x}) = \frac{g^2}{\kappa(\boldsymbol{x}) + D}\frac{1}{N}\sum_{i=1}^{N}\frac{x_i\phi(x_i)\phi'(x_i)}{1 + z_*g^2\phi'(x_i)^2} \tag{60}$$

$$\gamma(\boldsymbol{x}) = \sum_{i=1}^{N}\frac{x_i^2}{1 + z_*g^2\phi'(x_i)^2} \tag{61}$$

Note that $\alpha = \mathcal{O}(N^{-1})$, $\beta = \mathcal{O}(1)$, and $\gamma, \vartheta = \mathcal{O}(N)$. Hence, to leading order, Eq. (58) equals the diagonal approximation Eq. (45). This correspondence, however, only holds if the same is true for the defining Equation of z_*, which we show next.

The exact definition of z_* is the solution of Eq. (46). Using the Sherman Morrison formula, we find

$$\left(\mathbf{M}^{\mathrm{T}}\mathbf{M} + z\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{\Sigma}\right)^{-1} = \mathbf{B}^{-1} - \frac{\mathbf{B}^{-1}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\mathbf{B}^{-1}}{\vartheta^{-1} + \boldsymbol{a}^{\mathrm{T}}\mathbf{B}^{-1}\boldsymbol{a}} \tag{62}$$

where $\mathbf{B} = \mathbf{D} + \boldsymbol{a}\boldsymbol{x}^{\mathrm{T}} + \boldsymbol{x}\boldsymbol{a}^{\mathrm{T}}$. Hence, the right hand side of Eq. (46) is

$$\frac{1}{N}\mathrm{tr}\left[\left(\mathbf{B}^{-1} - \frac{\mathbf{B}^{-1}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\mathbf{B}^{-1}}{\vartheta^{-1} + \boldsymbol{a}^{\mathrm{T}}\mathbf{B}^{-1}\boldsymbol{a}}\right)\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{\Sigma}\right] \tag{63}$$

$$= \frac{1}{N}\mathrm{tr}\left[\mathbf{B}^{-1}N\left(\mathbf{K} - \frac{1}{\kappa+D}\boldsymbol{k}\boldsymbol{k}^{\mathrm{T}}\right)\right] - \frac{1}{N}\frac{1}{\vartheta^{-1} + \boldsymbol{a}^{\mathrm{T}}\mathbf{B}^{-1}\boldsymbol{a}}\left(\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{B}^{-\mathrm{T}}\boldsymbol{a}\right)^{\mathrm{T}}\mathbf{B}^{-1}\boldsymbol{a} \tag{64}$$

where for the second line, we used the properties of the trace. Applying twice the Sherman-Morrison formula and identifying the fields from above, the inverse of $\mathbf{B}$ is

$$\mathbf{B}^{-1} = \mathbf{D}^{-1} - \frac{\mathbf{D}^{-1}\boldsymbol{x}\boldsymbol{a}^{\mathrm{T}}\mathbf{D}^{-1}}{1 + \beta} - \frac{\mathbf{D}^{-1}\left(\boldsymbol{a}\boldsymbol{x}^{\mathrm{T}} - \frac{\gamma}{1+\beta}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}} - \frac{\alpha}{1+\beta}\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}} + \frac{\alpha\gamma}{(1+\beta)^2}\boldsymbol{x}\boldsymbol{a}^{\mathrm{T}}\right)\mathbf{D}^{-1}}{1 + \beta - \frac{\gamma\alpha}{1+\beta}}. \tag{65}$$

We have a lot of contributions. The only full rank part however is

$$\frac{1}{N}\mathrm{tr}\left(\mathbf{D}^{-1}N\mathbf{K}\right) = \frac{1}{N}\sum_{i=1}^{N}\frac{g^2\phi'(x_i)^2}{1 + zg^2\phi'(x_i)^2} = \mathcal{O}(1). \tag{66}$$

The other parts (rank 1 parts in the trace and scalar products to the right of Eq. (64)) yield scalar products that are all $\mathcal{O}(N^{-1})$. To see this, we count each inner product that arises as a factor N, track the explicit appearances of N and use the orders known for α, β, and γ. Hence, the equation defining z_*, Eq. (46), is to leading order equivalent to Eq. (44).

Lastly, we have to check that a $\mathcal{O}(N^{-1})$ correction to Eq. (44) only yields a $\mathcal{O}(N^{-1})$ correction to z_*, and furthermore to ζ in Eq. (45). To this end, assume that z_*^0 solves the approximate Eq. (44). Adding a correction term of $\mathcal{O}(N^{-1})$ to the approximate defining Eq. (44) leads to a shift $z_*^0 \to z_*^0 + \delta$, where δ is the solution of

$$1 = \frac{1}{N} \sum_{i=1}^{N} \frac{\lambda_i^2}{1 + (z_*^0 + \delta)\,\lambda_i^2} + \mathcal{O}(N^{-1}) \tag{67}$$

$$\Leftrightarrow \frac{1}{N} \sum_{i=1}^{N} \frac{\lambda_i^4}{\left(1 + z_*^0 \lambda_i^2\right)^2} \delta = \mathcal{O}(N^{-1}) + \mathcal{O}(\delta^2) \tag{68}$$

where we Taylor expanded around $\delta = 0$. Since the prefactor in front of δ is $\mathcal{O}(1)$, we find that $\delta = \mathcal{O}(N^{-1})$ is self-consistent. The change to ζ is then

$$\zeta = -\frac{1}{2}\left(z_*^0 + \delta\right) + \frac{1}{2N} \sum_{i=1}^{N} \ln\left[1 + \left(z_*^0 + \delta\right)\lambda_i^2\right] \tag{69}$$

$$= -\frac{1}{2} z_*^0 + \frac{1}{2N} \sum_{i=1}^{N} \ln(1 + z_* \lambda_i^2) + \mathcal{O}(\delta). \tag{70}$$

The leading order contribution to the scaled log determinant is hence indeed Eq. (45) with z_* being the solution of Eq. (44).

B. Empirical measure

So far we derived the distribution of fixed points $\rho(\boldsymbol{x})$. In this Section we discuss the *empirical measure*

$$\mu_{\boldsymbol{x}}(y) = \frac{1}{N} \sum_{i=1}^{N} \delta(y - x_i). \tag{71}$$

It is a function of y that is parameterized by the position $\boldsymbol{x}$. It contains all vector elements of $\boldsymbol{x}$ but forgets about their order. We want to understand the distribution of empirical measures when $\boldsymbol{x} \sim \rho(\boldsymbol{x})$. We will see that this distribution (of distributions) is strongly peaked at the expected empirical measure

$$\mu_*(y) = \langle \mu_{\boldsymbol{x}}(y) \rangle_{\boldsymbol{x} \sim \rho(\boldsymbol{x})}. \tag{72}$$

From the distribution of the empirical measure, the distribution of certain network-averaged quantities can be nicely characterized. Particularly, the expectation of a network average is given by the expected empirical measure

$$\left\langle \frac{1}{N} \sum_{i=1}^{N} f(x_i) \right\rangle_{\boldsymbol{x} \sim \rho(\boldsymbol{x})} = \int dy\, \mu_*(y) f(y), \tag{73}$$

which can be seen by plugging in the definition. Similarly, we can also express $\kappa(\boldsymbol{x})$ as a functional of $\mu_{\boldsymbol{x}}$

$$\kappa(\boldsymbol{x}) = \frac{g^2}{N} \phi(\boldsymbol{x})^{\mathrm{T}} \phi(\boldsymbol{x}) = g^2 \int dy\, \mu_{\boldsymbol{x}}(y) \phi(y)^2 = \kappa[\mu_{\boldsymbol{x}}]. \tag{74}$$

Furthermore, we can express the log determinant $\zeta(\boldsymbol{x})$ by $\mu_{\boldsymbol{x}}$

$$\zeta(\boldsymbol{x}) = -\frac{1}{2} z_*(\boldsymbol{x}) + \frac{1}{2N} \sum_{i=1}^{N} \ln(1 + z_*(\boldsymbol{x}) g^2 \phi'(x_i)^2) \tag{75}$$

$$= -\frac{1}{2} z_*[\mu_{\boldsymbol{x}}] + \frac{1}{2} \int dy\, \mu_{\boldsymbol{x}}(y) \ln(1 + z_*[\mu_{\boldsymbol{x}}] g^2 \phi'(y)^2) \tag{76}$$

$$= \zeta[\mu_{\boldsymbol{x}}] \tag{77}$$

where $z_*[\mu_{\boldsymbol{x}}]$ is the solution of

$$1 = \frac{1}{N} \sum_{i=1}^{N} \frac{g^2 \phi'(x_i)^2}{1 + z_* \, g^2 \phi'(x_i)^2} = \int dy \, \mu_{\boldsymbol{x}}(y) \frac{g^2 \phi'(y)^2}{1 + z_* \, g^2 \phi'(y)^2}. \tag{78}$$

Summing up, the complete fixed point distribution can be expressed in terms of the empirical measure $\mu_{\boldsymbol{x}}$

$$\rho(\boldsymbol{x}) = \rho[\mu_{\boldsymbol{x}}]. \tag{79}$$

1. Expected empirical measure

In this Section, we compute the expected empirical measure $\mu_*(y) = \langle \mu_{\boldsymbol{x}}(y) \rangle_{\boldsymbol{x} \sim \rho(\boldsymbol{x})}$. To this end, we follow the method presented in [9] and [10]. We consider the characteristic functional

$$Z[j] = \left\langle e^{ij^{\mathrm{T}} \mu_{\boldsymbol{x}}} \right\rangle_{\boldsymbol{x} \sim \rho(\boldsymbol{x})} = \left\langle e^{i \frac{1}{N} \sum_{i=1}^{N} j(x_i)} \right\rangle_{\boldsymbol{x} \sim \rho(\boldsymbol{x})} \tag{80}$$

where $j(y)$ is an auxiliary external source field and $j^{\mathrm{T}} \mu_{\boldsymbol{x}} = \int dy \, j(y) \mu_{\boldsymbol{x}}(y)$ denotes a functional scalar product (this notation will be implicit in the following). By $\langle \circ \rangle_{\boldsymbol{x} \sim \rho(\boldsymbol{x})}$ we mean average with respect to $\rho(\boldsymbol{x})/(\int d\boldsymbol{z} \, \rho(\boldsymbol{z}))$. We define the scaled cumulant generating functional as $W_N[j] = \frac{1}{N} \ln Z[Nj]$. The expected empirical measure is the first Taylor coefficient of W_N

$$\mu_*(y) = \left. \frac{\delta}{\delta j(y)} W_N[j] \right|_{j(y)=0}. \tag{81}$$

Plugging in the explicit result for the fixed point distribution, we have

$$W_N[j] = \frac{1}{N} \ln \int d\boldsymbol{x} \frac{1}{\left[2\pi \left(\kappa[\mu_{\boldsymbol{x}}] + D \right) \right]^{N/2}} e^{S[j; \boldsymbol{x}]} - c, \tag{82}$$

$$S[j; \boldsymbol{x}] = -\frac{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}{2 \left(\kappa[\mu_{\boldsymbol{x}}] + D \right)} + N\zeta[\mu_{\boldsymbol{x}}] + i \sum_i j(x_i), \tag{83}$$

$$c = \frac{1}{N} \ln \int d\boldsymbol{x} \rho(\boldsymbol{x}) = \frac{1}{N} \ln N_{\mathrm{fp}}, \tag{84}$$

where the expected number of fixed points N_{fp} is the norm of the fixed point distribution $\rho(\boldsymbol{x})$. The rate c is known as the *topological complexity* [11].

We want to evaluate the integral in Eq. (82). To this end, we introduce an auxiliary field $\mu(y)$ that we use to replace the $\boldsymbol{x}$-dependent field $\mu_{\boldsymbol{x}}(y)$. To ensure that this is still correct, we multiply the integrand by the functional Dirac constraint $\delta[\mu - \mu_{\boldsymbol{x}}] \equiv \lim_{M \to \infty} \prod_{i=1}^{M} \delta[\mu(y_i) - \mu_{\boldsymbol{x}}(y_i)]$ where $\{y_1, ..., y_M\} \overset{M \to \infty}{\to} \mathbb{R}$ is a discretization of the real line. Then, we have to integrate over $\mu(y)$ in a functional sense $\int \mathcal{D}\mu \equiv \lim_{M \to \infty} \int_{-\infty}^{\infty} \prod_{i=1}^{M} d\mu(y_i)$ for the replacement to be correct at every point in the $\boldsymbol{x}$ integration. Lastly, we replace the functional Dirac constraint by its Fourier integral representation

$$\delta[\mu - \mu_{\boldsymbol{x}}] = \int \mathcal{D}\tilde{\mu} \, e^{-iN\tilde{\mu}^{\mathrm{T}}(\mu - \mu_{\boldsymbol{x}})}$$
$$= \int \mathcal{D}\tilde{\mu} \, e^{-iN\tilde{\mu}^{\mathrm{T}} \mu + i \sum_{i=1}^{N} \tilde{\mu}(x_i)}, \tag{85}$$

where

$$\int \mathcal{D}\tilde{\mu} \equiv N \lim_{M \to \infty} \int_{-\infty}^{\infty} \prod_{i=1}^{M} \frac{d\tilde{\mu}(y_i)}{2\pi}. \tag{86}$$

With the auxiliary fields μ and $\tilde{\mu}$, the $\boldsymbol{x}$-integral in Eq. (82) formally factorizes

$$W_N[j] = \frac{1}{N} \ln \int \mathcal{D}\mu \mathcal{D}\tilde{\mu} \, e^{-iN\tilde{\mu}^{\mathrm{T}} \mu + N \ln \Omega[\mu, \tilde{\mu}, j]} - c \tag{87}$$

$$\Omega[\mu, \tilde{\mu}, j] = \int \frac{dx}{\sqrt{2\pi \left(\kappa[\mu] + D \right)}} e^{-\frac{x^2}{2(\kappa[\mu]+D)} + \zeta[\mu] + ij(x) + i\tilde{\mu}(x)}. \tag{88}$$

Note that this factorization into identical integrals $\int dx$ is only formal: The integrals are still coupled through their common dependence on the fields μ and $\tilde{\mu}$.

Saddle point approximation The exponent of the integrand in Eq. (87) is proportional to N. We are interested in large networks where $N \gg 1$. In this regime we can perform a saddle point approximation which yields to leading order in N^{-1}

$$W_N[j] = -i\tilde{\mu}_*[j]^{\mathrm{T}}\mu_*[j] + \ln\Omega\{\mu_*[j], \tilde{\mu}_*[j], j\} - c \tag{89}$$

where

$$\mu_*[j](y) = \left.\frac{\delta\ln\Omega}{\delta\,i\tilde{\mu}(y)}\right|_{\mu_*[j], \tilde{\mu}_*[j]}, \quad i\tilde{\mu}_*[j](y) = \left.\frac{\delta\ln\Omega}{\delta\,\mu(y)}\right|_{\mu_*[j], \tilde{\mu}_*[j]} \tag{90}$$

are the maxima of the exponent of the integrand in Eq. (87). With this, we can compute the expectation value of $\mu_{\boldsymbol{x}}$ in saddle point approximation: Evaluating Eq. (81) gives

$$\mu_*(y) = \left.\frac{\delta\ln\Omega\{\mu_*[0], \tilde{\mu}_*[0], j\}}{\delta\,j(y)}\right|_{j=0} \tag{91}$$

$$= \mu_*[0](y)$$

where we used Eqs. (90) in the first step to eliminate the chain-rule derivatives and in the second step to identify the result with $\mu_*[0]$. Therefore, the derivative in Eq. (91) only acts on the explicit dependency of Ω on j.

Saddle point equations Next, to get the expected empirical measure, we discuss the solution of the saddle point Equations (90) for $j = 0$.

The first saddle point Equation follows from straight forward differentiation

$$\mu_*(y) = \frac{e^{-\frac{y^2}{2(\kappa[\mu_*]+D)} + \zeta[\mu_*] + i\tilde{\mu}_*(y)}}{\sqrt{2\pi(\kappa[\mu_*]+D)}\Omega[\mu_*, \tilde{\mu}_*]}. \tag{92}$$

The second saddle point Equation

$$i\tilde{\mu}_*(y) = \frac{1}{\Omega}\left.\left(\frac{\partial\Omega}{\partial\kappa}\frac{\delta\kappa[\mu]}{\delta\mu(y)} + \frac{\partial\Omega}{\partial\zeta}\frac{\delta\zeta[\mu]}{\delta\mu(y)}\right)\right|_{\mu_*, \tilde{\mu}_*} \tag{93}$$

involves some chain rule derivatives. We need the derivative of κ,

$$\frac{\delta\kappa[\mu]}{\delta\mu(y)} = g^2\phi(y)^2, \tag{94}$$

and we need the derivative of $\zeta[\mu]$ as in Eq. (76). We find

$$\frac{\delta\zeta[\mu]}{\delta\mu(y)} = -\frac{1}{2}\frac{\delta z_*[\mu]}{\delta\mu(y)}\left(1 - \int dx\,\mu(x)\frac{g^2\phi'(x)^2}{1 + z_*[\mu]g^2\phi'(x)^2}\right) + \frac{1}{2}\ln(1 + z_*[\mu]g^2\phi'(y)^2) \tag{95}$$

where the first part vanishes due to the definition of z_*, see Eq. (78).

Concluding, the saddle point equation for $\tilde{\mu}$ is

$$i\tilde{\mu}_*(y) = \frac{g^2\phi(y)^2}{2(\kappa[\mu_*]+D)}\left(\frac{\langle x^2\rangle_{\mu_*}}{\kappa[\mu_*]+D} - 1\right) + \frac{1}{2}\ln(1 + z_*[\mu_*]g^2\phi'(y)^2). \tag{96}$$

The expected empirical measure is determined by the simultaneous solution of Eqs. (92) and (96). Combining them, we find

$$\mu_*(y) = \mathcal{Z}^{-1}\sqrt{1 + \alpha\phi'(y)^2}\,e^{-\frac{y^2}{2\beta} + \gamma\phi(y)^2} \tag{97}$$

where

$$\alpha = z_*[\mu_*]g^2, \quad \beta = \kappa[\mu_*] + D, \quad \gamma = \frac{g^2}{2\beta}\left(\frac{q[\mu_*]}{\beta} - 1\right) \tag{98}$$
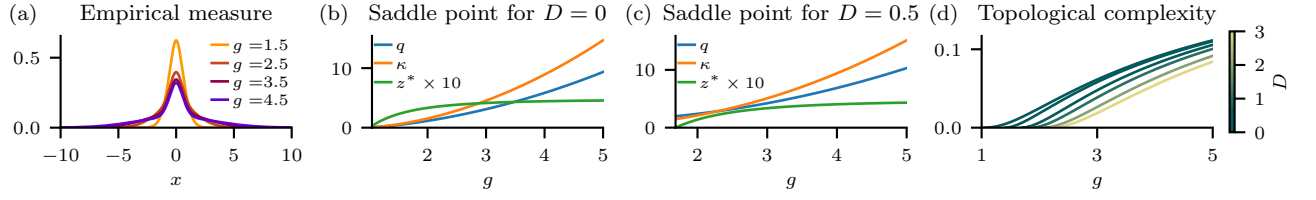
Figure 1. Solution of the saddle point approximation. (a) Expected empirical measure Eq. (97) for $D = 10^{-3}$. (b) Saddle point fields Eq. (100) for $D = 0$. (c) Same as (a), but $D = 0.5$. (d) Topological complexity Eq. (125) for several values of g and D. The transition to a positive complexity shifts for increasing noise strength.

and $\mathcal{Z}$ normalizes the empirical measure. Here, $q[\mu_*] \equiv \left\langle y^2 \right\rangle_{y \sim \mu_*}$. The parametrized form of the empirical measure in Eq. (97) allows to formulate the saddle point approximation in terms of the scalars z_*, κ and q. The respective set of Equations can be formulated in terms of standard Gauss integrals

$$q = \sqrt{2\pi\beta(\kappa)}\mathcal{Z}^{-1}\left\langle \beta x^2 \sqrt{1 + \alpha(z_*)\phi'(\sqrt{\beta(\kappa)}x)^2} e^{\gamma(\kappa,q)\phi(\sqrt{\beta(\kappa)}x)^2} \right\rangle_{x \sim \mathcal{N}(0,1)} \tag{99}$$

$$\kappa = g^2 \sqrt{2\pi\beta(\kappa)}\mathcal{Z}^{-1}\left\langle \phi(\sqrt{\beta(\kappa)}x)^2 \sqrt{1 + \alpha(z_*)\phi'(\sqrt{\beta(\kappa)}x)^2} e^{\gamma(\kappa,q)\phi(\sqrt{\beta(\kappa)}x)^2} \right\rangle_{x \sim \mathcal{N}(0,1)} \tag{100}$$

$$1 = g^2 \sqrt{2\pi\beta(\kappa)}\mathcal{Z}^{-1}\left\langle \frac{\phi'(\sqrt{\beta(\kappa)}x)^2}{\sqrt{1 + \alpha(z_*)\phi'(\sqrt{\beta(\kappa)}x)^2}} e^{\gamma(\kappa,q)\phi(\sqrt{\beta(\kappa)}x)^2} \right\rangle_{x \sim \mathcal{N}(0,1)} \tag{101}$$

where the norm $\mathcal{Z}$ in Eq. (97) can be written as

$$\mathcal{Z} = \int dy \sqrt{1 + \alpha\phi'(y)^2} e^{-\frac{y^2}{2\beta} + \gamma\phi(y)^2} \tag{102}$$

$$= \sqrt{2\pi\beta(\kappa)}\left\langle \sqrt{1 + \alpha(z_*)\phi'(\sqrt{\beta(\kappa)}x)^2} e^{\gamma(\kappa,q)\phi(\sqrt{\beta(\kappa)}x)^2} \right\rangle_{x \sim \mathcal{N}(0,1)}. \tag{103}$$

Equations (99), (100), and (101) can be solved efficiently by a damped iteration and using Gauss-Hermite quadrature for the integrals due to the formulation in terms of standard Gaussians. They are equivalent to the compact equations in the main text. The resulting saddle point fields and the expected empirical measure are shown in Fig. 1.

Furthermore, the expected empirical measure Eq. (97) is compared to numerical results (see Sec. E) in Fig. 2.

2. Fluctuations of the empirical measure

Here, we discuss some properties of the expected empirical measure.

Expected empirical measure versus marginal distribution   In the context of self-averaging variables one often wants to know if certain distributions can be swapped when computing the expected value of the respective variable. In this spirit, we here show that the expected empirical measure is not only the expected distribution of vector components at a fixed point, but also the expected marginal distribution of one vector component across all fixed points. To see this, consider the definition of the expected empirical measure

$$\mu_*(y) = \frac{1}{\int d\boldsymbol{x}'\rho(\boldsymbol{x}')}\int d\boldsymbol{x}\,\rho(\boldsymbol{x})\,\mu_{\boldsymbol{x}}(y) = \frac{1}{\int d\boldsymbol{x}'\rho(\boldsymbol{x}')}\frac{1}{N}\sum_{i=1}^{N}\int d\boldsymbol{x}\,\rho(\boldsymbol{x})\,\delta(x_i - y). \tag{104}$$

Recall that due to the realization average, the fixed point distribution is symmetric under permutations of the neurons. Hence, every part of the sum in Eq. (104) is equal, and we can write

$$\mu_*(y) = \frac{1}{\int d\boldsymbol{x}'\rho(\boldsymbol{x}')}\int d\boldsymbol{x}\,\rho(\boldsymbol{x})\,\delta(x_1 - y) \tag{105}$$

where instead of x_1 any other neuron could be chosen, as well. Carrying out the integration over x_1, we find the usual expression for marginal distributions

$$\mu_*(y) = \frac{1}{\int d\boldsymbol{x}'\rho(\boldsymbol{x}')}\int \left(\prod_{i=2}^{N} dx_i\right)\rho(y, x_2, ..., x_N). \tag{106}$$
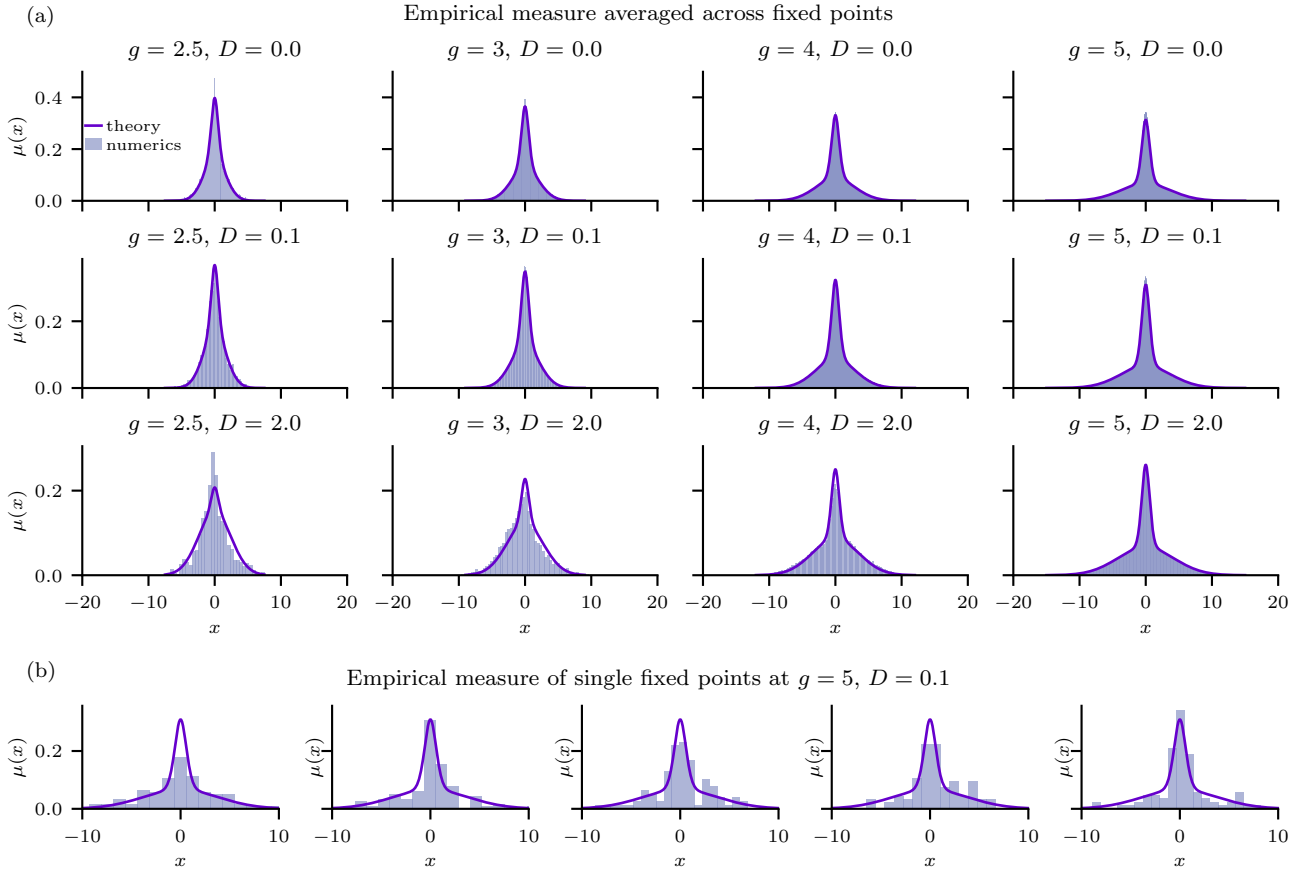
Figure 2. Empirical measure for different parameters. (a) Numerical verification of Eq. (97) by averaging across all fixed point found at saturation for a fixed realization (see Sec. E). (b) Same as (a), but using single fixed points instead of averaging across all. We note that despite the low resolution due to numerical limitations, even single fixed points resemble the expected empirical measure. This is discussed in Sec. B 2.

Origin of the variability in $\rho(\boldsymbol{x})$ So far we computed the expected empirical measure. In the next Section, we estimate fluctuations of the empirical measure around the expected empirical measure. In order to interpret them right, we here study the origin of these fluctuations. When considering the empirical measure of single fixed points in a fixed realization, the fluctuations could be across fixed points *within* the realization. When considering the average empirical measure over all fixed points in one realization, the fluctuations of this average could be *across* realizations.

We find that the fixed point density accounts for both, within- and across realization fluctuations. To see this we derive the *law of total variance*. The variance of the empirical measure is

$$T(y) = \frac{\delta^2}{\delta j(y)^2} \ln \left\langle e^{ij^\mathrm{T}\mu_{\boldsymbol{x}}} \right\rangle_{\boldsymbol{x} \sim \rho(\boldsymbol{x})} \Bigg|_{j=0} = \left\langle \left(\mu_{\boldsymbol{x}}(y) - \langle \mu_{\boldsymbol{z}}(y) \rangle_{\rho(\boldsymbol{z})} \right)^2 \right\rangle_{\rho(\boldsymbol{x})} . \tag{107}$$

Recall that the fixed point distribution is the realization average $\rho(\boldsymbol{x}) = \langle \rho_{\mathbf{J},\boldsymbol{\eta}}(\boldsymbol{x}) \rangle_{\mathbf{J},\boldsymbol{\eta}}$. In that spirit, we can dissect the total variance T (by adding a zero term) into

$$T(y) = A(y) + \langle W_{\mathbf{J},\boldsymbol{\eta}}(y) \rangle_{\mathbf{J},\boldsymbol{\eta}} \tag{108}$$

where

$$A(y) = \left\langle \left(\langle \mu_{\boldsymbol{x}}(y) \rangle_{\rho_{\mathbf{J},\boldsymbol{\eta}}(\boldsymbol{x})} - \left\langle \langle \mu_{\boldsymbol{z}}(y) \rangle_{\rho_{\mathbf{J}',\boldsymbol{\eta}'}(\boldsymbol{z})} \right\rangle_{\mathbf{J}',\boldsymbol{\eta}'} \right)^2 \right\rangle_{\mathbf{J},\boldsymbol{\eta}} \tag{109}$$

is the *across realizations variance* of the expected empirical measure and

$$W_{\mathbf{J},\boldsymbol{\eta}}(y) = \left\langle \left(\mu_{\boldsymbol{x}}(y) - \langle \mu_{\boldsymbol{z}}(y) \rangle_{\rho_{\mathbf{J},\boldsymbol{\eta}}(\boldsymbol{z})} \right)^2 \right\rangle_{\rho_{\mathbf{J},\boldsymbol{\eta}}(\boldsymbol{x})} \tag{110}$$

is the *within realizations variance* of the empirical measure. Equation (108) reflects the law of total variance.

We conclude that the variability we get from the fixed point distribution accounts for both – across- and within realizations – variances. This means that the expected empirical measure is also the expected distribution of vector components of *every* fixed point. We expect that for our case the within realizations part $\langle W_{\mathbf{J},\boldsymbol{\eta}} \rangle_{\mathbf{J},\boldsymbol{\eta}}$ strongly dominates for both the empirical measure and the norm due to a self-averaging property in the sense that $\langle G(\boldsymbol{x}) \rangle_{\rho_{\mathbf{J},\boldsymbol{\eta}}(\boldsymbol{x})} \approx \langle G(\boldsymbol{x}) \rangle_{\rho(\boldsymbol{x})}$; proving this analytically would require to compute the joint density $\langle \rho_{\mathbf{J},\boldsymbol{\eta}}(\boldsymbol{x}) \rho_{\mathbf{J},\boldsymbol{\eta}}(\boldsymbol{z}) \rangle_{\mathbf{J},\boldsymbol{\eta}}$ akin to the computation of $\langle \rho_{\mathbf{J},\boldsymbol{\eta}}(\boldsymbol{x}) \rangle_{\mathbf{J},\boldsymbol{\eta}}$ in Sec. A 1 and show that it is very close to $\rho(\boldsymbol{x})\rho(\boldsymbol{z})$.

Computing across-realization fluctuations numerically is resource-intensive because one has to find a representative set of fixed points for several realizations to track the variability of the within-realization mean. Now even finding a representative set for one realization is difficult for resource reasons, see Sec. E.

However, for the case of norms we can compare the rate function we compute in Sec. C 1 with the distribution of numerical fixed point norms in one realization. Here we find that most of the variability seems to be already within the realization, see Fig. 3, underlying our claim that the across-variances part A is subdominant.

Rate function for the empirical measure   In Fig. 2 we observe that the expected empirical measure is in excellent agreement with the empirical measure averaged across fixed point of a fixed realization. We further observe that even single fixed point's empirical measures resemble the expected empirical measure. To understand this, we here investigate the likeliness of deviations of the empirical measure from the expected empirical measure. In deriving the saddle point Equations (92) and (96) we showed the differentiability of the scaled cumulant generating functional Eq. (89). Due to the existence and the differentiability of the scaled cumulant generating functional, the Gärtner-Ellis theorem holds [10], [12, Sec. 5], stating that the empirical measure fulfills a *large deviation principle*: The family of measures μ converges to the expected measure μ_* as $N \to \infty$ in the sense of distributions. One can also call this a self-averaging property.

According to the Gärtner-Ellis theorem, the probability distribution functional of all empirical measures at fixed points is of the form $P[\mu] \doteq \exp(-NH[\mu])$, and the *rate functional* $H[\mu]$ is the Legendre transform of the scaled cumulant generating functional. It quantifies the probability of deviations of μ from the minimum μ_* of the rate functional. These are rare for large N since they are exponentially suppressed in probability.

In this Section, we compute the rate functional. Since the scaled cumulant generating functional is differentiable, the Legendre transform is

$$H[\mu] = i\mu^{\mathrm{T}} j_*^{\mu} - W_N\left[j_*^{\mu}\right] \tag{111}$$

where j_*^{μ} is determined by the stationarity condition

$$\mu(y) = \left.\frac{\delta}{\delta\, ij(y)} W_N[j]\right|_{j_*^{\mu}}. \tag{112}$$

Analogously to Eq. (91), we get

$$\mu(y) = \frac{e^{-\frac{y^2}{2\left(\kappa\left\{\mu_*[j_*^{\mu}]\right\}+D\right)}+\zeta\left\{\mu_*[j_*^{\mu}]\right\}+ij_*^{\mu}(y)+i\tilde{\mu}_*[j_*^{\mu}](y)}}{\Omega_*\sqrt{2\pi\left(\kappa\left[\mu_*[j_*^{\mu}]\right]+D\right)}} \tag{113}$$

where we introduced the short hand notation $\Omega_* = \Omega\left\{\mu_*[j_*^{\mu}], \tilde{\mu}_*[j_*^{\mu}], j_*^{\mu}\right\}$. Comparing with Eq. (91), we also see the identity

$$\mu(y) = \mu_*[j_*^{\mu}](y) \tag{114}$$

reflecting that μ is canonically conjugate to j. Formally solving Eq. (113) for the appearance of j_*^{μ} in the exponent, and plugging into Eq. (111), yields

$$H[\mu] = D_{\mathrm{KL}}(\mu\|\nu) - W_N\left[j_*^{\mu}\right] \tag{115}$$

where $D_{\mathrm{KL}}(\rho_1\|\rho_2) = \langle \ln\left[\rho_1(x)/\rho_2(x)\right] \rangle_{\rho_1(x)}$ is the Kullback-Leibler divergence between the probability distribution functions ρ_1 and ρ_2. As a functional of ρ_1 it is convex and has a single minimum at $\rho_1 = \rho_2$. The reference function here is

$$\nu(y) = \frac{e^{-\frac{y^2}{2(\kappa[\mu]+D)}+\zeta[\mu]+i\tilde{\mu}_*[j_*^{\mu}](y)}}{\Omega_*\sqrt{2\pi\left(\kappa[\mu]+D\right)}}. \tag{116}$$

A necessary condition for the expected measure μ_* is that it minimizes the rate functional

$$\frac{\delta}{\delta\mu}H[\mu]\Big|_{\mu=\mu_*} = ij_*^{\mu_*} \overset{!}{=} 0. \tag{117}$$

Indeed, since $j = 0$ maximizes $W_N[j]$, namely $W_N[0] = 0$, the expected empirical measure determined through the self consistent requirement of the Kullback-Leibler divergence is

$$\mu_*(y) = \frac{e^{-\frac{y^2}{2(\kappa[\mu_*]+D)}+\zeta[\mu]+i\tilde{\mu}_*[0](y)}}{\Omega_*\sqrt{2\pi\left(\kappa[\mu_*]+D\right)}}, \tag{118}$$

in line with the saddle point Equation (92).

Equation (115) expresses that points $\mu \neq \mu_*$ are exponentially suppressed for two reasons: First, simply because of the difference in terms of the Kullback-Leibler divergence. Second, because of a mismatch of the extremum of the scaled cumulant generating functional.

C. Further observables

1. Distance distribution, separation of shells

Here, we ask about the distance $d_{\boldsymbol{x}} = \sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}$ of fixed points to the origin. The expected distance is

$$d_* = \sqrt{N \int dy\,\mu_*(y)y^2}. \tag{119}$$

According to the *contraction principle* [10], the distance inherits the large deviation principle from the empirical measure and thus, Eq. (119) is dominant for large N. Hence, the fixed points live on a thin shell of an N-dimensional sphere of radius d_*. The thickness of the spherical shell decays exponentially with N.

We want to compute the finite size fluctuations of the distance. To this end, we compute the rate function of the squared scaled distance $u = d^2/N = \int dy\,\mu(y)y^2$. The contraction principle states that the rate function for u can be derived from the rate functional for μ

$$I(u) = \inf_{\mu:\,v^{\mathrm{T}}\mu=u} H[\mu] \tag{120}$$

where $v(y) = y^2$. The minimization is carried out using a Lagrange multiplier, hence we have to optimize

$$\mathcal{L}[\mu;\lambda] = H[\mu] - \lambda\left(u - v^{\mathrm{T}}\mu\right) \tag{121}$$

for both, λ and μ. Recalling the Legendre transform for $H[\mu] = \inf_j\left(i\mu^{\mathrm{T}}j - W[j]\right)$ in Eq. (111) and swapping the optimizations, we have

$$\frac{\delta}{\delta\mu(y)}\left[i\mu^{\mathrm{T}}j - \lambda\left(u - v^{\mathrm{T}}\mu\right)\right] \overset{!}{=} 0 \tag{122}$$

which constraints the source term $ij \overset{!}{=} \lambda v$. Since $v(y) = y^2$ is a fixed function, $\inf_j \to \inf_\lambda$. We have

$$I(u) = \inf_\lambda\left(\lambda\mu^T v - W[-i\lambda v]\right)\Big|_{v^{\mathrm{T}}\mu\overset{!}{=}u} = \inf_\lambda\left(\lambda u - W_N[-i\lambda v]\right) \tag{123}$$

where we were able to explicitly plug in the condition $v^{\mathrm{T}}\mu = u$. Thus, we have to solve

$$u = \frac{\partial}{\partial\lambda}W_N[-i\lambda v] = v^{\mathrm{T}}\mu_*[-i\lambda v] \quad . \tag{124}$$

In practice, we solve this by computing $\mu_*[-i\lambda v]$ in the same way as we computed $\mu_*[0]$ in Sec. B 1 for a sequence of values λ. Then, we check which u they correspond to using Eq. (124). Lastly, we plug into Eq. (123). In the main text, we show the respective variances for $D = 0.1$ and several values of g. Here, in Fig. 3 we also show the whole distribution for several parameters and compare it to the numerical results.

Finite size distribution of the fixed point's scaled squared distance $u = N^{-1}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$
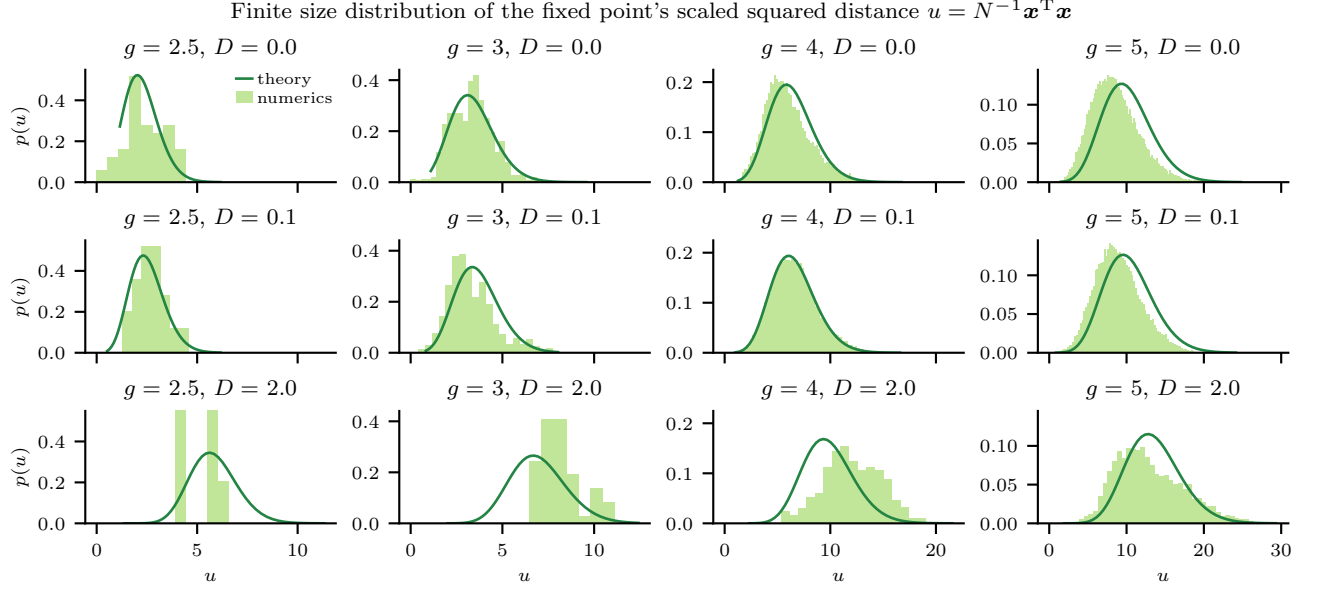

Figure 3. Distance distribution. Using Eq. (123) and Eq. (124), we compute the distribution of the scaled squared distance for several parameters (solid line) and compare it to the numerical findings (histograms) for fixed realizations of $\mathbf{J}$ and $\boldsymbol{\eta}$.

Deviations from numerical results   Here, we discuss the deviations of the distance distribution derived as above from the numerical findings using brute force fixed point finding, which can be observed in Fig. 3.

For low values of g, only few fixed points can be found due to only $N = 100$ and $c(g \approx 1) \approx 0$ and $N_{\mathrm{fp}} \doteq e^{cN}$. Hence, the distance distribution relies on a low number of samples. In contrast, for the empirical measure, we have N times more samples, since every vector component is a sample.

For high values of g, the system has many fixed points. So many that only a tiny subset of the complete number can be found within reasonable time and we therefore have to stop the fixed point finding routine before saturation. Hence, the resulting numerical distribution is strongly dependent on possible unknown biases of the Levenberg-Marquart fixed point finder (possibly due to choice of initial values) to reach fixed points with certain properties more often than others.

For intermediate values of g we should expect the best agreement, but also both effects from above might kick in.

2.   Topological complexity

In this Section, we discuss the *topological complexity*. It was introduced by Wainrib and Touboul [11] to describe the rate by which the expected number of fixed points grows with the number of neurons $N_{\mathrm{fp}} \doteq e^{c(g,D)\,N}$. Wainrib and Touboul proposed the topological complexity as a structural indicator of chaoticity, complementary to the maximum Lyapunov exponent which they refer to as *dynamical complexity*.

By construction in Eq. (3), the expected number of fixed points is the norm of the fixed point distribution. We therefore encountered the topological complexity in the scaled cumulant generating functional W_N in Eq. (84). We get the topological complexity c in saddle point approximation by the normalization of the scaled cumulant generating functional Eq. (89) $W_N[0] \overset{!}{=} 0$ as

$$c = -i\tilde{\mu}_*^{\mathrm{T}}\mu_* + \ln \Omega\left[\mu_*, \tilde{\mu}_*\right]. \tag{125}$$

Based on the results of Sec. B 1, the topological complexity is shown in Fig. 1(d). This result substantially deviates from the result in [11] for $g > 1$, hence our methods extends range of validity, $g = 1 + \epsilon$, $0 < \epsilon \ll 1$ required in [11].

Critical g.–   In the main text we discuss the transition from a system without fixed points to a positive topological complexity $c > 0$. The transition point is computed as the root of Eq. (125).

Next, in the main text, we show the transition line to chaos. The result presented there is based on the dynamic mean-field computation in [13]. The case of *quenched noise*, as [13] puts it, which also we are interested in, slightly differs from the white noise case. This is for two reasons: First, at infinity, the autocorrelation function does not

drop to zero but remains positive. Second, the stability of the quenched-noise system is completely determined by the Jacobian of the dynamics. Hence, the criterion for the transition to chaos is that of the spectral radius of the Jacobian crossing unity

$$1 = g^2 \left\langle \phi'(\sqrt{c_0}z)^2 \right\rangle \tag{126}$$

where z is standard normal. Here, c_0 is the equal-time autocorrelation, which follows from "energy conservation" in the particle-in-a-potential analogy [13] as

$$-\frac{1}{2}c_\infty^2 + g^2 f_\Phi(c_\infty, c_0) + c_\infty D = -\frac{1}{2}c_0^2 + g^2 f_\Phi(c_0, c_0) + c_0 D \tag{127}$$

where $f_u(x,y) = \iint Dz_1 Dz_2 u\left(\sqrt{y - \frac{x^2}{y}}z_1 + \frac{x}{\sqrt{y}}z_2\right)u\left(\sqrt{y}z_2\right)$ and $Dz = dz\frac{1}{\sqrt{2\pi}}e^{-z^2/2}$. Lastly, demanding a constant autocorrelation c_∞ at infinite timelag $\tau = \infty$ gives

$$c_\infty = g^2 f_\phi(c_\infty, c_0) + D. \tag{128}$$

We solve Eqs. (126), (127), and (128) with a damped fixed point iteration. This yields the critical coupling strength g_c at which the transition to chaos occurs, and which is shown in the main text.

3. Jacobian spectrum at fixed points

Here, we want to characterize the spectrum of the Jacobian Eq. (5) at fixed points. The Jacobian reads

$$\mathbf{y}'(\mathbf{x}) = -\mathbb{1} + \mathbf{J}\operatorname{diag}\left[\phi'(\mathbf{x})\right] \tag{129}$$

where $J_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, g^2/N)$. For random matrices of this type, it is known that the spectrum of eigenvalues is self-averaging [6, Theorem 1.14]. Ahmadian et $al.$ [8] provide a general framework for computing the spectrum of arbitrary random matrices of the form $\mathbf{\mathcal{M}} + \mathbf{\mathcal{L}}\mathbf{X}\mathbf{\mathcal{R}}$ where X_{ij} are i.i.d. with zero mean and variance $1/N$. Since the identity in Eq. (129) only contributes a shift of -1 of the eigenvalues of $\mathbf{y}'(\mathbf{x})$, we can use the method by Ahmadian et $al.$ for the simpler case $\mathbf{\mathcal{M}} = 0$, $\mathbf{\mathcal{L}} = \mathbb{1}$, and $\mathbf{\mathcal{R}} = g\operatorname{diag}[\phi'(\mathbf{x})]$, where we pulled the factor g out of $\mathbf{J}$ explicitly to match the notation by Ahmadian et $al.$ The special case of random matrices with zero mean is described by Ahmadian et $al.$ starting from [8, Eq. (2.13)] for the support $R(\mathbf{x}) = \sqrt{\frac{1}{N}\sum_{i=1}^N \sigma_i^2}$ where σ_i are the singular values of $\mathbf{\mathcal{L}}\mathbf{\mathcal{R}}$, thus the eigenvalues of $\mathbf{\Lambda}$ since $\Lambda_{ij} = \delta_{ij}g\phi'(x_i) > 0$. In terms of the distribution of $\mathbf{X}\mathbf{\Lambda}(\mathbf{x})$, $\sigma_i = g\phi'(x_i)$ is the standard deviation of the matrix elements in the ith column. Thus, we have the support

$$R(\mathbf{x}) = g\sqrt{\frac{1}{N}\phi'(\mathbf{x})^\mathrm{T}\phi'(\mathbf{x})}. \tag{130}$$

Outside of this support, the spectral density vanishes. Within the support, the spectral density is isotropic around the center at $-1 + 0i$, which comes from the identity $-\mathbb{1}$ that we pulled out. The spectral density

$$\nu(r) = -\frac{1}{2\pi r}\partial_r n_{\mathbf{x}}(r) \tag{131}$$

can hence be derived from the $radial$ $tail$ $distribution$, i.e. the proportion of eigenvalues further than r from the center, $n_{\mathbf{x}}(r)$ [8, Eq. (2.14)]. This is the fundamental theorem of calculus in polar coordinates. The proportion $n_{\mathbf{x}}(r)$ is given by [8, Eq. (2.15)]

$$1 = \frac{1}{N}\sum_{i=1}^N \frac{1}{\frac{r^2}{[g\phi'(x_i)]^2} + n_{\mathbf{x}}(r)} \qquad \text{for } r < R \tag{132}$$
$$\text{and } n_{\mathbf{x}}(r) = 0 \qquad \text{for } r \geq R$$

Note that for the case of identically distributed columns ($x_i = x$ for all i), Eq. (132) can be solved $n(r) = 1 - r^2/[g\phi'(x)]^2$, hence $\nu(r) = 1/[\pi R(\mathbf{x})]^2 = \text{const.}$ w.r.t. r, and reflects the circular law which states that the spectrum of i.i.d. random matrices is uniform [6].

We are, however, interested in the case where x_i are different, namely sampled from the expected empirical measure to give the expected spectral distribution of the Jacobian at fixed points. Since both the spectral radius and the defining equation of $n_{\boldsymbol{x}}$ depend on $\boldsymbol{x}$ through network sums $\sum_{i=1}^{N} f(x_i)$ only, both can be expressed in terms of the empirical measure

$$R[\mu_{\boldsymbol{x}}] = g\sqrt{\int dy\,\mu_{\boldsymbol{x}}(y)\phi'(y)^2}, \qquad 1 = \int dy\,\mu_{\boldsymbol{x}}(y)\frac{1}{\frac{r^2}{[g\phi'(y)]^2} + n_{\mu_{\boldsymbol{x}}}(r)}. \tag{133}$$

Consequently, their expected values are given by $R[\mu_*]$ and $n_{\mu_*}(r)$ respectively. The latter is computed using a bisection algorithm. The results are shown in the main text Fig. 2.

D. Correlation despite separation

In the main text we discuss the apparent contradiction that the dynamics' velocity is strongly correlated with the prediction by the nearest fixed point despite the separation of shells. Here, we give some more technical details on this discussion.

1. Impact of fixed points

To measure the impact of fixed points we fix a realization of the connectivity $\mathbf{J}$ and the static noise $\boldsymbol{\eta}$. Then, we use a high performance cluster, as outlined in Sec. E, to get a large fraction of the fixed points for that realization. Next, we integrate the model for the same realization and ask at every time point along the trajectory $\boldsymbol{x}(t)$ which one of the many fixed points is the closest. The distance $\|\boldsymbol{x}(t) - \boldsymbol{x}^*\|$, where $\boldsymbol{x}^*$ is closest at time t, is shown in the main text (Fig. 3). At the same time, we compute the Pearson correlation

$$c(t) = \frac{\boldsymbol{y}[\boldsymbol{x}(t)]^{\mathrm{T}}\,[\boldsymbol{y}'(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*)]}{\|\boldsymbol{y}[\boldsymbol{x}(t)]\|\,\|\boldsymbol{y}'(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*)\|} \tag{134}$$

of the actual velocity and the one predicted by the nearest fixed point. To compare all these 'drive-bys', we normalize the time T during which a certain fixed point is the closest to 1.

2. Tangentially fixed points and lines

To bridge the gap between the inner fixed point shell and the outer dynamics shell, we consider tangentially fixed points. Tangentially fixed points are fixed points of an equivalent model, wherein the distance is constrained $\|\boldsymbol{x}(t)\| \overset{!}{=} r$. The dynamics of the constraint model follows by subtracting the radial velocity $\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}(\boldsymbol{x})/\|\boldsymbol{x}\|^2$ from the original model, yielding

$$\dot{\boldsymbol{x}} = -\frac{\boldsymbol{x}^{\mathrm{T}}\mathbf{J}\phi(\boldsymbol{x})}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}\boldsymbol{x} + \mathbf{J}\phi(\boldsymbol{x}). \tag{135}$$

Thus, tangentially fixed points are solutions of

$$\mathbf{J}\phi(\boldsymbol{x}) = \frac{\boldsymbol{x}^{\mathrm{T}}\mathbf{J}\phi(\boldsymbol{x})}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}\boldsymbol{x} \quad \Leftrightarrow \quad \boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}\mathbf{J}\phi(\boldsymbol{x}) = r^2\mathbf{J}\phi(\boldsymbol{x}). \tag{136}$$

Since Eq. (136) is the eigenvalue equation of a rank 1 matrix, the solution is unique up to a scalar factor, $\mathbf{J}\phi(\boldsymbol{x}) = (1 + s)\boldsymbol{x}$.

For the tangentially fixed lines shown in the main text Fig. 3, we start with fixed points $\boldsymbol{x}$ and their exact distance to the origin $r = \|\boldsymbol{x}\|$ (here, $s = 0$). Then we incrementally change this distance $r \to r + \Delta r$ and compute the solution of Eq. (136) using the Levenberg-Marquart rootfinder initialized at the last known tangentially fixed point. Moving to larger (smaller) radii r along the fixed line corresponds to building up a radial velocity towards (away from) the origin. Towards the inside, all tangentially fixed lines stop; numerically this is detected by noting that no tangentially fixed points can be found anymore in the vicinity and under the distance-constraint. Note that in the linear regime of the phase space, tangentially fixed points cannot exist in exponential abundance anymore, hence tangentially fixed lines have to end. Towards the outside, most fixed lines did not end within the interval of integration, some lines however did in the same manner as towards the inside.
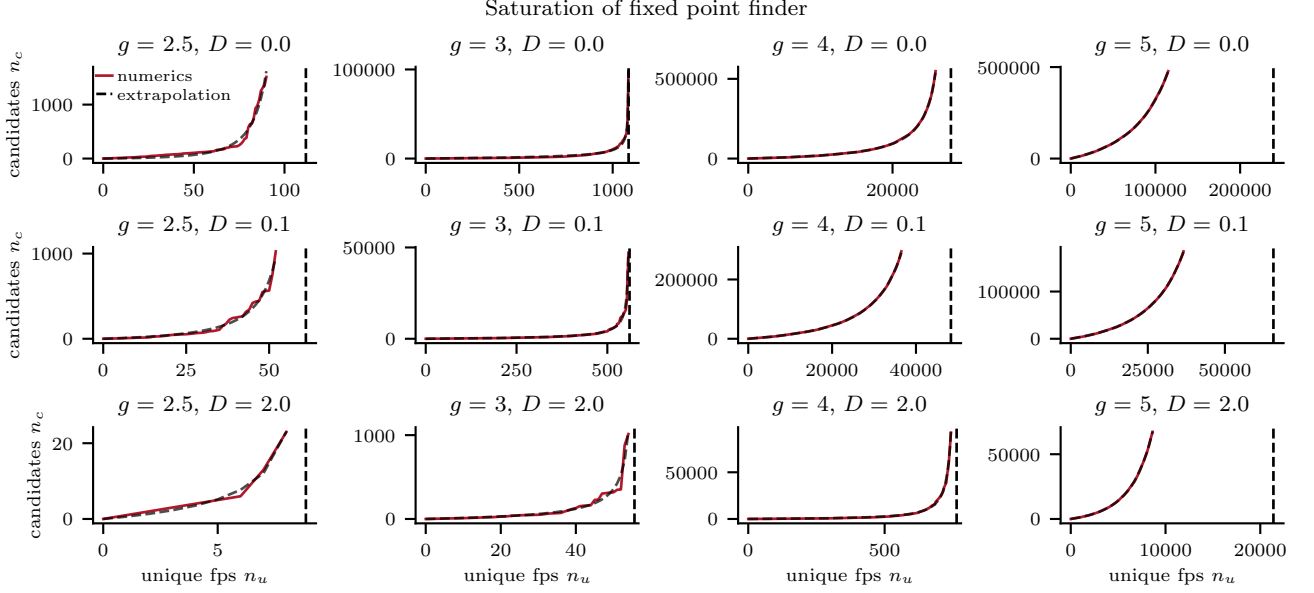
Figure 4. Saturation of the fixed point finder. For each number of unique fixed points (x-axis) we show the number of candidates needed (y-axis, red solid line). The black dashed curve shows the fit to the function $n_c(n_u) = p_1\left((n_u - p_2)^{-p_3} - p_2^{-p_3}\right)$. The dashed vertical line denotes the pole position p_2. It estimates the expected number of unique fixed points. Here, $N = 100$.

E. Numerical fixed point finding

To test the analytical results about the statistics of the fixed points, we compare them to statistics of large numbers of fixed points obtained for specific realizations of the random connectivity $\mathbf{J}$ and noise $\boldsymbol{\eta}$ by brute force fixed point finding.

To this end, we first sample $\mathbf{J}$ and $\boldsymbol{\eta}$ randomly using the variances g^2/N and D respectively. The number of neurons is $N = 100$ for all fixed point finding procedures.

Next, we employ a Levenberg-Marquart minimizer on the velocity field $\boldsymbol{y}(\boldsymbol{x}) = -\boldsymbol{x} + \mathbf{J}\phi(\boldsymbol{x}) + \boldsymbol{\eta}$ starting from 1.5×10^9 to 10×10^9 initial conditions sampled from a normal distribution of large variance. The exact number of initial conditions is chosen such that the procedure, if possible, saturates, see below.

Since Levenberg-Marquart is a minimizer, we next sort out slow points, i.e., points at which the norm of the velocity is locally minimal but not zero; for a point to be counted as fixed point, the norm of the velocity is required to be $< 10^{-6}$. We call the resulting list of fixed points *candidates*.

1. Saturation

Due to the large number of initial conditions, some fixed points are found repeatedly in the list of candidates. Furthermore, fixed points with certain properties (e.g. a small norm) seem to be found more likely than others. To remove this bias, resulting from the choice of the ensemble of initial conditions and from the mechanics of Levenberg-Marquart, we use a high number of initial conditions and sort out duplicates.

To test whether the number of initial conditions is high enough, we monitor the saturation: Iterating through the list of candidates, for each new unique fixed point found (say, number n_u), we count the number n_c of candidates needed to find n_u unique fixed points. At the beginning of the iteration, $n_c \approx n_u$, because almost every candidate has been unknown before. Along the list of candidates, n_c grows supra-linearly, as more and more candidates are already contained in the list of unique fixed points and hence more candidates are needed to find a new unique fixed point. When $n_c \gg n_d$, the procedure is saturated in the sense that a large fraction of the fixed points accessible to Levenberg-Marquart seem to be found. The flow of the saturation n_c is shown in Fig. 4.

Assuming a finite number of fixed points, we expect a pole in the function $n_c(n_u)$ at the position of the (unknown) complete number of unique fixed points $n_u \equiv p_2$. A generic function for such a pole is $(n_u - p_2)^{-p_3}$ where $p_3 > 0$. Setting the y-axis intercept to 0 requires subtracting $p_2^{-p_3}$. To allow for an arbitrary overall scale, we further multiply
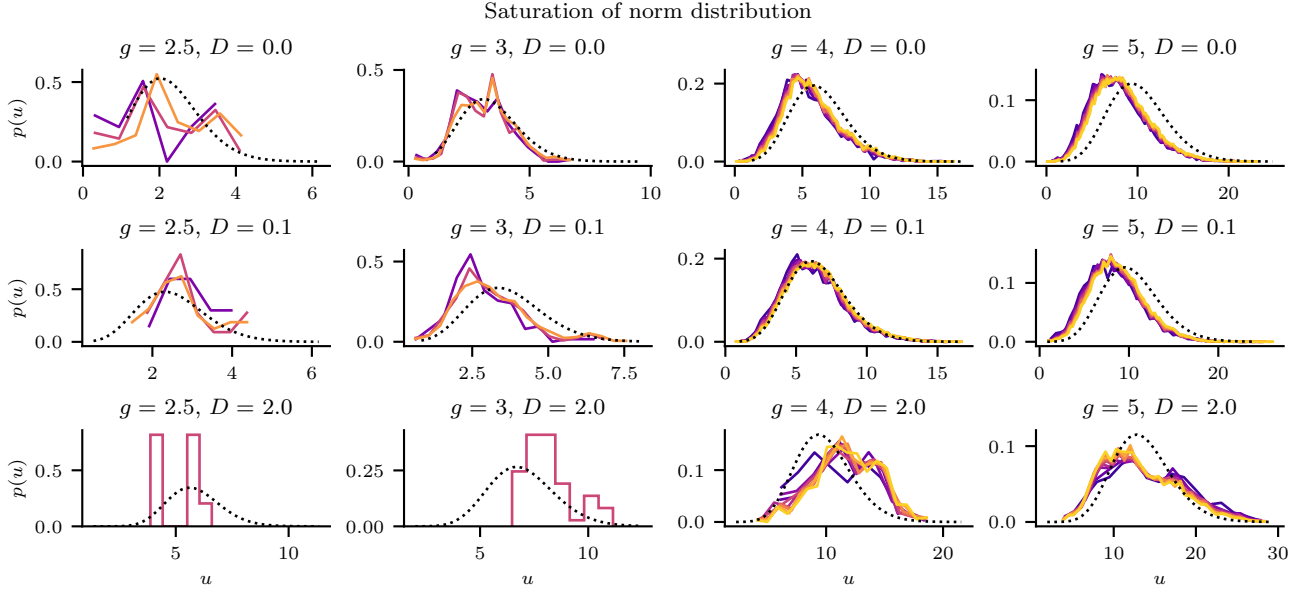
Saturation of norm distribution

Figure 5. Norm saturation. Distribution of the scaled squared norm $u = \frac{1}{N}\boldsymbol{x}^{*\mathrm{T}}\boldsymbol{x}^*$ of fixed points. The color denotes the number of fixed points taken into account: The purple line denotes the distribution of a first batch of fixed points. The yellow line uses all fixed points found. The dotted lines show the theory based on Eqs. (123) and (124).

the function by p_1. Hence, we fit the saturation flow to the function $n_c(n_u) = p_1\left[(n_u - p_2)^{-p_3} - p_2^{-p_3}\right]$, see Fig. 4. With this, p_2 is the extrapolated estimate of the number of unique fixed points. This number is shown in the main text in Fig. 1(e).

To understand the bias of the fixed point finder, we further show in Fig. 5 how the norm distribution of the fixed points changes after acquiring more and more fixed points. Here, we observe that especially for poorly saturated searches (e.g. $g = 5$, $D = 0$, see upper right panel in Fig. 4), the procedure finds fixed points with a small norm more likely. Only by sorting out duplicates, the norm distribution shifts towards the expected distribution, however not reaching it in time.

[1] M. Kac, On the average number of real roots of a random algebraic equation, Bulletin of the American Mathematical Society **49**, 314 (1943).
[2] S. O. Rice, Mathematical analysis of random noise, Bell Syst. Tech. J. **24**, 46 (1945), reprinted in [**?**].
[3] J.-M. Azaïs and M. Wschebor, *Level Sets and Extrema of Random Processes and Fields* (John Wiley & Sons, 2009).
[4] R. L. Stratonovich, *Topics in the Theory of Random Noise* (Gordon and Breach, New York, 1967).
[5] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*, Adaptive Computation and Machine Learning (MIT Press, Cambridge, MA, USA, 2006) p. 248.
[6] T. Tao, V. Vu, and M. Krishnapur, Random matrices: Universality of esds and the circular law, Ann. Probab. **38**, 2023 (2010).
[7] H. Sommers, A. Crisanti, H. Sompolinsky, and Y. Stein, Spectrum of large random asymmetric matrices, Phys. Rev. Lett. **60**, 1895 (1988).
[8] Y. Ahmadian, F. Fumarola, and K. D. Miller, Properties of networks with partially structured and partially random connectivity, Phys. Rev. E **91**, 012820 (2015).
[9] A. van Meegen, T. Kühn, and M. Helias, Large-deviation approach to random recurrent neuronal networks: Parameter inference and fluctuation-induced transitions, Phys. Rev. Lett. **127**, 158302 (2021).
[10] H. Touchette, The large deviation approach to statistical mechanics, Phys. Rep. **478**, 1 (2009).
[11] G. Wainrib and J. Touboul, Topological and dynamical complexity of random neural networks, Phys. Rev. Lett. **110**, 118101 (2013).
[12] R. S. Ellis, An overview of the theory of large deviations and applications to statistical mechanics, Scand. Actuar. J. **1995**, 97 (1995), https://doi.org/10.1080/03461238.1995.10413952.
[13] J. Schuecker, S. Goedeke, and M. Helias, Optimal sequence memory in driven random networks, Phys. Rev. X **8**, 041029 (2018).