Fixed point geometry in chaotic neural networks

Jakob Stubenrauch [1,2,3,4,*] Christian Keup [1,2,5] Anno C. Kurth [1,2] Moritz Helias [1,2] and Alexander van Meegen [1,6,7]

[1] *Institute for Advanced Simulation (IAS-6), Jülich Research Centre, Jülich, Germany*

[2] *Department of Physics, Faculty 1, RWTH Aachen University, Aachen, Germany*

[3] *Bernstein Center for Computational Neuroscience Berlin, Philippstrasse 13, Haus 2, 10115 Berlin, Germany*

[4] *Physics Department of Humboldt University Berlin, Newtonstrasse 15, 12489 Berlin, Germany*

[5] *Statistical Physics of Computation Laboratory, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland*

[6] *Institute of Zoology, University of Cologne, 50674 Cologne, Germany*

[7] *Center for Brain Science, Harvard University, Cambridge, Massachusetts 02138, USA*

(Received 24 October 2022; revised 6 December 2023; accepted 18 March 2025; published 29 May 2025)

Understanding the high-dimensional chaotic dynamics occurring in complex biological systems such as recurrent neural networks or ecosystems remains a conceptual challenge. For low-dimensional dynamics, fixed points provide the geometric scaffold of the dynamics. However, in high-dimensional systems, even the location of fixed points is unknown. Here, we analytically determine the number and distribution of fixed points for a canonical model of a recurrent neural network that exhibits high-dimensional chaos. This distribution reveals that fixed points and dynamics are confined to separate shells in state space. Furthermore, the distribution enables us to determine the eigenvalue spectra of the Jacobian at the fixed points, showing that each fixed point has a low-dimensional unstable manifold. Despite the radial separation of fixed points and dynamics, we find that the principal components of fixed points and dynamics align and that nearby fixed points act as partially attracting landmarks for the dynamics. Our results provide a detailed characterization of the fixed point geometry and its interplay with the dynamics, thereby paving the way towards a geometric understanding of high-dimensional chaos through their skeleton of unstable fixed points.

DOI: 10.1103/PhysRevResearch.7.023203

I. INTRODUCTION

Complex systems with heterogeneous interactions are ubiquitous in biology, from large-scale ecosystems [1] to neural networks [2,3]. These systems are characterized by nonlinear and asymmetric interactions that generate rich collective dynamics, including stable fixed points, limit cycles, and chaos [4]. Chaotic behavior has been observed in both ecosystems [5,6] and neural networks [7–9]. Indeed, in high-dimensional systems with asymmetric, random interactions, chaos appears to be the rule rather than the exception [10].

While chaotic dynamics are well understood in low-dimensional systems, they present significant challenges in high dimensions [4], particularly for "extensive chaos" where the chaotic dynamics themselves are high dimensional. In low-dimensional systems, fixed points provide geometric insight into chaotic dynamics—exemplified by the two unstable fixed points at the centers of the Lorenz attractor [11]. However, the state space of high-dimensional nonlinear systems remains largely unexplored. Although it is known that the number of fixed points can grow exponentially with dimension [12–14], and that these fixed points must be unstable in chaotic systems [12], their spatial distribution and relationship to the dynamics remain unknown.

We address these challenges in the context of the canonical chaotic neural network model introduced by Sompolinsky et al. [7]. This model is particularly suitable because dynamical mean-field theory [7,15] (reviewed pedagogically in [16]) has provided detailed insights into its dynamics for large networks. Specifically, the theory predicts a transition to chaos above a critical connection strength [7,17]. Recent investigations of cross correlations [18] and the complete Lyapunov spectrum [19] have confirmed the extensive nature of chaos in this system. The approach based on dynamic mean-field theory has proven to be robust, revealing the dynamical properties of various extensions of the model [20–31].

In this work, we determine the spatial distribution of fixed points of the chaotic neural network model. Our approach relies on computing the expected zero crossings of a Gaussian process with location-dependent statistics. By comparing the geometries of the fixed points and dynamics, we demonstrate their confinement to separate shells in state space. Analysis of higher-order statistics reveals alignment between the principal components of the fixed points and the dynamics. After characterizing the local dynamics at the fixed points, we demonstrate that the fixed points serve as transiently attractive landmarks that guide the chaotic flow.

*Contact author: jakob.stubenrauch@rwth-aachen.de

II. MODEL

We consider a generalized version of the model from [7], including neuronal heterogeneity and external input [17]. In this generalized version, the model consists of N nonlinearly connected neurons $x_i(t)$ obeying the system of ordinary differential equations

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

with neuronal nonlinearity $\phi(x) = \tanh(x)$, independent and identically distributed (i.i.d.) coupling weights $J_{ij} \sim \mathcal{N}(0, g^2/N)$, i.i.d. set points $\eta_i \sim \mathcal{N}(0, D)$, and i.i.d. Gaussian noise ξ_i with $\langle \xi_i(t)\rangle = 0$ and $\langle \xi_i(t)\xi_i(t')\rangle = \sigma^2\delta(t-t')$. The strength of the recurrent connections is controlled by g; D and σ control the neuronal heterogeneity and the external noise input, respectively.

Throughout the manuscript, we assume that the network is in the chaotic regime $g > g_c$ [7,17] and that the number of units, N, is sufficiently large to allow us to focus on the leading, exponential scale of the behavior, which we express by the abbreviated notation $a \doteq e^{Nb}$ to denote $\lim_{N\to\infty} \frac{1}{N} \ln a = b$.

III. FIXED POINT DISTRIBUTION

We use vector notation to write Eq. (1) as $\dot{\boldsymbol{x}} = \boldsymbol{y}(\boldsymbol{x}) + \boldsymbol{\xi}(t)$ with velocity $\boldsymbol{y}(\boldsymbol{x}) = -\boldsymbol{x} + \mathbf{J}\boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{\eta}$. Since $\mathbf{J}$ and $\boldsymbol{\eta}$ are Gaussian, 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 in state space [note that both $\boldsymbol{y}(\boldsymbol{x})$ and $\boldsymbol{y}'(\boldsymbol{x})$ are nonhomogeneous, i.e., their statistics depend on absolute position $\boldsymbol{x}$]. 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 state space. We determine $\rho(\boldsymbol{x})$ from the Kac-Rice formula [32–34],

$$p_{\mathrm{fp}}(\boldsymbol{x}) = \left\langle N_{\mathrm{fp}}^{-1}\delta[\boldsymbol{y}(\boldsymbol{x})]\,|\det \boldsymbol{y}'(\boldsymbol{x})|\right\rangle_{\mathbf{J},\boldsymbol{\eta}}, \tag{2}$$

where the normalization is given by the number of fixed points, N_{fp}, the Dirac δ enforces $\boldsymbol{y}(\boldsymbol{x}) = 0$, and the Jacobian determinant $|\det \boldsymbol{y}'(\boldsymbol{x})|$ ensures that every fixed point contributes to the distribution with equal weight. We assume that the number of fixed points is strongly self-averaging such that we can separate the averages,

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

with norm $\langle N_{\mathrm{fp}}\rangle_{\mathbf{J},\boldsymbol{\eta}} = \int d\boldsymbol{x}\,\rho(\boldsymbol{x})$. Equation (3) is equivalent to a random matrix problem (see Appendix A): Using Bayes' law to condition on $\boldsymbol{y}(\boldsymbol{x}) = 0$, $\rho(\boldsymbol{x})$ can be rewritten into (see Supplemental Material [[35], A1])

$$\rho(\boldsymbol{x}) = p_{\mathrm{L}}(\boldsymbol{x})\langle|\det[\mathbf{M}(\boldsymbol{x}) + \mathbf{X}\,\boldsymbol{\Sigma}(\boldsymbol{x})]|\rangle_{X_{ij}\sim\mathcal{N}(0,N^{-1})}, \tag{4}$$

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 independent Gaussian processes $[\boldsymbol{y}(\boldsymbol{x})]_i$, $\mathbf{M}(\boldsymbol{x})$ is the mean of the Gaussian process $\boldsymbol{y}'(\boldsymbol{x})$ conditioned on $\boldsymbol{y}(\boldsymbol{x}) = 0$, and $N^{-1}\boldsymbol{\Sigma}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}(\boldsymbol{x})$ is the covariance matrix of the independent rows of $\boldsymbol{y}'(\boldsymbol{x})$ conditioned on $\boldsymbol{y}(\boldsymbol{x}) = 0$. Extending the technique introduced in [36] and excluding singularities, the determinant is given by

$$\langle|\det[\mathbf{M}(\boldsymbol{x}) + \mathbf{X}\,\boldsymbol{\Sigma}(\boldsymbol{x})]|\rangle \doteq \exp[N\zeta(\boldsymbol{x})], \tag{5}$$

with (see Supplemental Material [[35], A2])

$$\zeta(\boldsymbol{x}) = -\frac{1}{2}z_* + \frac{1}{2N}\sum_{i=1}^{N}\ln[1 + z_* g^2\phi'(x_i)^2], \tag{6}$$

where z_* is the solution of

$$1 = \frac{1}{N}\sum_{i=1}^{N}\frac{g^2\phi'(x_i)^2}{1 + z_*\,g^2\phi'(x_i)^2}. \tag{7}$$

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

$$S(\boldsymbol{x}) = \frac{u(\boldsymbol{x})}{2[\kappa(\boldsymbol{x}) + D]} + \frac{1}{2}\ln\{2\pi[\kappa(\boldsymbol{x}) + D]\} - \zeta(\boldsymbol{x}), \tag{8}$$

where $u(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^{N}x_i^2$ and $\zeta(\boldsymbol{x})$ is determined by Eqs. (6) and (7).

Since incoming and outgoing weights J_{ij} of all units are statistically equivalent, the averaged fixed point distribution $\rho(\boldsymbol{x})$ is permutation symmetric, which implies an approximate independence of coordinates $\int \rho(\boldsymbol{x})\,dx_{k+1}\ldots dx_N \approx \prod_{i=1}^{k}\mu(x_i)$ for $k \ll N$ (see Supplemental Material [[35], B3]). The permutation symmetry is reflected by the property that the fixed point distribution, given by Eq. (8), depends on $\boldsymbol{x}$ only through network averages. Consequently, we can express it as a functional $\rho(\boldsymbol{x}) = \rho[\mu_{\boldsymbol{x}}]$ of the *empirical measure*,

$$\mu_{\boldsymbol{x}}(y) = \frac{1}{N}\sum_{i=1}^{N}\delta(y - x_i), \tag{9}$$

which is 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 $\langle\frac{1}{N}\sum_{i=1}^{N}f(x_i)\rangle_{\boldsymbol{x}\sim\rho(\boldsymbol{x})} = \int dy\,\mu_*(y)f(y)$ can be computed. The expected empirical measure is given, for large N, by the saddle point that maximizes $\rho[\mu_{\boldsymbol{x}}]$ in function space and admits the form (see Appendix B)

$$\mu_*(y) \propto \sqrt{1 + \alpha\phi'(y)^2}\,e^{-\frac{\gamma^2 + \gamma\phi(y)^2}{2\beta}}, \tag{10}$$

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

We compare the empirical measure given by Eq. (10) 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 [37,38] starting from independent normally distributed initial conditions until saturation, until almost no new fixed points are found (see Supplemental Material [[35], E]). Figure 1(a) shows that the theory given by Eq. (10) is in excellent agreement with the empirical measure averaged

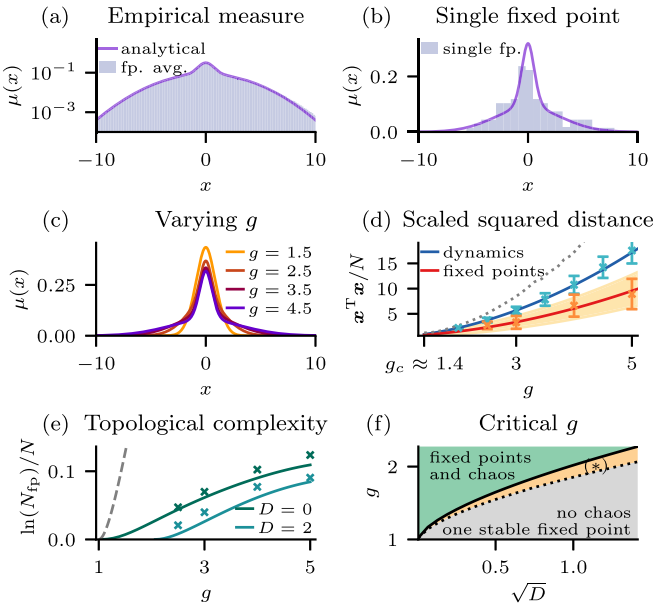

FIG. 1. Characterization of fixed points. (a) Distribution of vector components of fixed points (empirical measure): theoretical result (solid line) based on Eq. (10) and histogram (bars) averaged across 4×10^4 fixed points for a single realization of the coupling weights with $D = 0.1$ and $g = 4$. (b) Same as (a) for a randomly chosen single fixed point. (c) Empirical measure for different values of g. (d) Scaled squared norm of fixed points (theory: red solid line; theoretical finite-size standard deviation: yellow shading; numerical results: orange error bars), dynamics (theory: blue solid line; numerical integration averaged over 15 realizations per g: turquoise error bars), and maximum scaled squared norm above which the number of fixed points is no longer exponential (gray dotted line). (e) Number of fixed points (topological complexity; solid lines: theory, Eq. (13); crosses: numerical results, Supplemental Material [[35], E]; gray dashed line: theory for $g \to 1^+$, $D = 0$ by [12]). (f) Transition to positive topological complexity (black dotted line) and transition to chaos (black solid line) based on [17]. In the regime (∗), the state space exhibits an exponential number of unstable fixed points, but the dynamics are not chaotic. Parameters: $D = 0.1$ and g as indicated for (a)–(d), $N = 100$ for numerical results and the shading in (d).

over all fixed points found numerically in a single realization of $\mathbf{J}$ (see Supplemental Material [[35], Fig. 2(a)] for further examples). Moreover, as shown in Fig. 1(b), even single fixed points closely resemble the expected empirical measure.

The agreement between the expected empirical measure and distribution for a single fixed point is not a coincidence. The probability distribution functional of the empirical measures takes the form $P[\mu] \doteq \exp(-NH[\mu])$ with an analytically determined rate functional $H[\mu]$ (see Supplemental Material [[35], B2]), which means that it obeys a large deviation principle [39,40]; see Appendix C. The minimum of $H[\mu]$ 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 (see Supplemental Material [[35], B2]), 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 Supplemental Material [[35], Fig. 2(b)] for further examples).

IV. SEPARATE SHELLS OF FIXED POINTS AND DYNAMICS

First, we investigate the geometry of the fixed point distribution using the expected empirical measure. A prominent feature of the empirical measure is the sharp peak and the Gaussian base [Figs. 1(a)–1(c)]. This shape reflects the compromise between the two contributions in the fixed point density, given by Eq. (4): high probability of a vanishing velocity, captured by $p_{\mathrm{L}}(\boldsymbol{x})$, and a steep expected slope to increase the density of zero crossings, captured by the determinant. The former leads to the broad Gaussian base and the latter to the sharp peak. Geometrically, the sharp peak at zero implies that the fixed points are in the vicinity of spans of subsets of axes in state space.

Next, we consider the squared norm $u(\boldsymbol{x}) = \frac{1}{N}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$, which quantifies the distance to the origin and which is equal to the second moment of the empirical measure $u(\boldsymbol{x}) = \int dy\, y^2 \mu_{\boldsymbol{x}}(y)$, with expectation value

$$u_* = \int dy\, y^2 \mu_*(y) = \beta(1 - \gamma/g^2). \tag{11}$$

The distribution of the distance $P(u)$ inherits the exponential form of $P[\mu]$; formally, this is a consequence of the contraction principle [39]. Thus, $P(u) \doteq \exp[-NI(u)]$, where the rate function is

$$I(u) = \inf_{\mu:\, \int y^2 \mu(y)\, dy = u} H[\mu]. \tag{12}$$

Hence, for $N \gg 1$, the fluctuations of the distance vanish and the fixed points are distributed on a thin spherical shell with radius $\sqrt{Nu_*}$, where u_* [given by Eq. (11)] minimizes $I(u)$. In Fig. 1(d), we show the average distance and fluctuations based on Eq. (12) (see Appendix D 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 [16,30]. Hence, also the trajectory is embedded in a thin shell around the origin, which is of radius $\sqrt{Nq}$, and q is derived in [17].

The confinement to a thin spherical shell is a generic feature of high-dimensional, weakly correlated random variables [41], but the radius depends on the underlying high-dimensional distribution. Thus, we compare the radii of the two shells in Fig. 1(d). 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 state space. Note, however, that the vanishing overlap is the result of an exponentially large fraction of the fixed points. In absolute terms, the number of fixed points at the dynamics shell is still exponential [see Supplemental Material [[35], C2] and the gray dotted line in Fig. 1(d)].

V. NUMBER OF FIXED POINTS

A core result of [12] is that without neuronal heterogeneity, $D = 0$, the system has a transition from a single stable fixed point to an exponential number of unstable fixed points

$\langle N_{\rm fp}\rangle \doteq \exp[cN]$ at $g_c = 1$. The respective rate c, the *topological complexity*, can be computed from the empirical measure (see Appendix E), leading to

$$c = \frac{\gamma - \alpha}{2g^2} - D\frac{\gamma}{2\beta g^2} + \ln \mathcal{Z} - \frac{1}{2}\ln(2\pi\beta), \qquad (13)$$

where $\mathcal{Z}$ is the normalization of Eq. (10). Asymptotically, at $D = 0$ and for any sigmoidal nonlinearity with $\phi'(0) = 1$, $c \to -\frac{1}{\pi} + \ln\left[1 + \mathrm{erf}(\sqrt{1/\pi})e^{1/\pi}\right] \approx 0.264$ for $g \to \infty$ and $c = (\frac{2}{3})^4\epsilon^3 + O(\epsilon^4)$ for $g - 1 = \epsilon \to 0$ (see Supplemental Material [[35], C4]). Both the cubic growth at the chaos transition and the saturation at large g are consistent with the behavior of the effective dimension of the dynamics [18].

In Fig. 1(e), we see that the critical gain parameter g_c grows with D; 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 heterogeneity D, however, a regime exists where the system has an exponential number of fixed points, yet the dynamics are not chaotic (see [42] for a similar observation). Our theory and numerical results are both in agreement with the critical point $g_c = 1$ for $D = 0$ found by [12], but the quantitative value of the topological complexity clearly differs from the result by [12] and is well captured by our theory [Fig. 1(e)].

VI. ALIGNMENT OF FIXED POINTS AND DYNAMICS

The analysis thus far shows that there are exponentially many fixed points which are, however, radially separated from the dynamics. This suggests a drastic decoupling of fixed points and dynamics. We investigate this further using higher-order statistics of fixed points and dynamics.

We consider the dimensionality of dynamics and fixed points based on the cross-covariance matrix $\mathbf{C} = \langle(\boldsymbol{x} - \langle\boldsymbol{x}\rangle)(\boldsymbol{x} - \langle\boldsymbol{x}\rangle)^{\rm T}\rangle$, where the average is either over the fixed points (denoted $\mathbf{C}_{\rm fps}$) or the trajectory of the stationary dynamics (denoted $\mathbf{C}_{\rm dyn}$) for a fixed realization of $\mathbf{J}$. The variance of the fixed points or the dynamics projected on the ith eigenvector (termed principle component) of $\mathbf{C}$ is given by the respective eigenvalue λ_i, and the eigenvalue participation ratio $(N^{-1}\sum_{i=1}^N \lambda_i)^2/(N^{-1}\sum_{i=1}^N \lambda_i^2)$ is thus a linear estimate of the dimension of the set of fixed points or the dynamical attractor, respectively. For the dynamics, Refs. [18,19] showed that the dimensionality is extensive but fractional. The same is true for the fixed points [Fig. 2(a)]. Furthermore, the dimensionality grows monotonically with g for dynamics and fixed points. For the dynamics, the dimensionality saturates at approximately 6% [18]; for the fixed points, the dimensionality is approximately twice as large. Thus, both dynamics and fixed points occupy a lower-dimensional subspace of the state space.

The next question is if these two subspaces are aligned. To this end, we consider the relative cumulative variance of either fixed points or dynamics along the first n principal components of the fixed points,

$$V_n^{\circ} = \frac{\sum_{i=1}^n \boldsymbol{v}_i^{\rm T}\mathbf{C}_{\circ}\boldsymbol{v}_i}{\sum_{i=1}^N \boldsymbol{v}_i^{\rm T}\mathbf{C}_{\circ}\boldsymbol{v}_i}, \qquad (14)$$

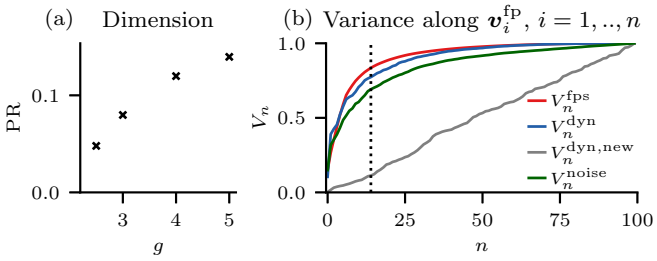

FIG. 2. Alignment of fixed points and dynamics. (a) Dimensionality of the cloud of fixed points for varying g and $D = 0$. (b) Relative cumulative variance given by Eq. (14) of the fixed points (red), the dynamics (blue), the dynamics including strong external noise $\sigma = 2$ (green), and the dynamics with resampled connectivity (gray) along the n leading principal components of the fixed points for $g = 5$, $D = 0$.

where $\circ \in \{\rm fps, dyn\}$ specifies the covariance matrix and $\boldsymbol{v}_i$ is the ith principal component of $\mathbf{C}_{\rm fps}$. The saturation of $V_n^{\rm fps}$ at $n \approx 0.1N$ (black dotted line) corresponds to the dimensionality of ≈10%. For the dynamics, $V_n^{\rm dyn}$ saturates almost as quickly, showing that the embedding manifolds of fixed points and dynamics align nearly perfectly. Importantly, this is only true for a given realization of $\mathbf{J}$: the relative variance of dynamics $V_n^{\rm dyn,new}$ with resampled $\mathbf{J}_{\rm new}$ captured by the principal components of the fixed points from the original $\mathbf{J}$ grows linearly with the number of components [Fig. 2(b), gray line]; thus, across realizations, the dynamics and fixed points do not align. Last, adding strong external noise ($\sigma = 2$) to the dynamics does not foil the alignment. This indicates that fixed points of the model for $\sigma = 0$ also provide insight into the noisy case $\sigma > 0$.

In summary, considering the higher-order statistics shows that the dynamics are contained near the manifold occupied by the fixed points. For the dynamics, the alignment can be traced back to the most repulsive eigenvectors of the Jacobian at the origin $\boldsymbol{y}'(0) = -\mathbb{1} + \mathbf{J}$ [18]. Due to the alignment of fixed points and dynamics, this holds for the fixed points as well. At the origin, the alignment of Jacobian and fixed points implies that the dynamics are repelled towards the other fixed points within the lower-dimensional subspace.

VII. LOCAL DYNAMICS

We now consider the dynamics in the vicinity of fixed points $\boldsymbol{x}^*$ outside the origin. Local dynamics are determined by the eigenvalues of the Jacobian $\boldsymbol{y}'(\boldsymbol{x}^*)$ at these fixed points. The Jacobian can be written as $\boldsymbol{y}'(\boldsymbol{x}) = -\mathbb{1} + \mathbf{X}\,\mathrm{diag}[g\phi'(\boldsymbol{x})]$ with $X_{ij} \sim \mathcal{N}(0, N^{-1})$; the corresponding eigenvalue spectrum can be computed with the method developed in [43] 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 $R(\boldsymbol{x}) = g\sqrt{N^{-1}\phi'(\boldsymbol{x})^{\rm T}\phi'(\boldsymbol{x})}$ [Figs. 3(a) and 3(b)]. At a fixed point, the contraction principle attests a large deviation principle for the spectral radius, with the expected value given by

$$R_* = g\sqrt{\int dy\, \phi'(y)^2\mu_*(y)}. \qquad (15)$$

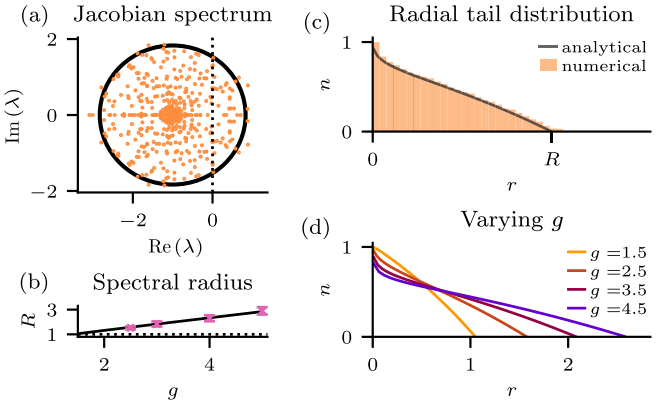

FIG. 3. Jacobian spectrum at fixed points. (a) Eigenvalues (orange dots) of the Jacobian at five randomly chosen fixed points of one realization of $\mathbf{J}$ and $\boldsymbol{\eta}$. Every eigenvalue with positive real part corresponds to an unstable mode of the fixed point. The support of the spectrum (black circle) is determined by the spectral radius, given by Eq. (15). (b) Spectral radius: theory (black line) and numerical results (pink error bars). (c) Radial tail distribution of eigenvalues: theory (black line) based on Eq. (16) 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), $D = 0.1$.

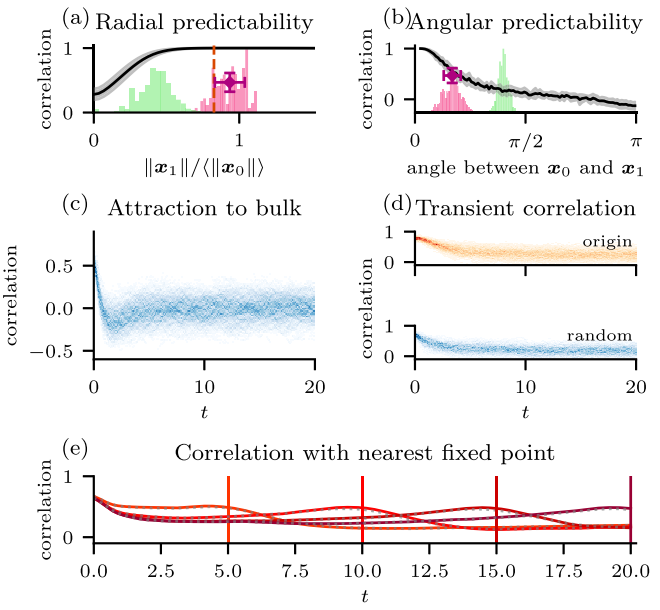

FIG. 4. Impact of fixed points. (a) Pearson correlation between $\boldsymbol{y}(\boldsymbol{x}_0)$, where $\boldsymbol{x}_0$ is on the attractor, and its linear predictor $\boldsymbol{y}(\boldsymbol{x}_1) + \boldsymbol{y}'(\boldsymbol{x}_1)[\boldsymbol{x}_0 - \boldsymbol{x}_1]$ for radially perturbed points $\boldsymbol{x}_1$ (black line) and nearest fixed points (purple error bar). Histograms show the density of the norm of the fixed points that are closest to the dynamics (pink) and to random points on the dynamics shell (green); the dashed line shows the mean norm of all fixed points. Parameters: $g = 5$, $D = 0$, and $N = 300$ (perturbations) or $N = 100$ (fixed points); shadings denote standard deviation. (b) Like (a), but for angular perturbed points $\boldsymbol{x}_1 = a\boldsymbol{x}_0 + \sqrt{1-a^2}\|\boldsymbol{x}_0\|\boldsymbol{z}$, where $z_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1/N)$, and the parameter $a \in [-1, 1]$ is varied to interpolate between $\boldsymbol{x}_0$ and $\boldsymbol{z}$ such that different angles between $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ are realized. (c) Density of Pearson correlation of the velocity $\boldsymbol{y}(\boldsymbol{x}(t))$ with the pointer $\boldsymbol{x}^* - \boldsymbol{x}$ to a randomly chosen fixed point for 400 trajectories starting from random initial points with norm $\frac{1}{N}\|\boldsymbol{x}\|^2 = c_0$, where c_0 is the asymptotic squared norm for $g = 4$ and $D = 0$. (d) Density of Pearson correlation between the velocity in (c), with the velocity predicted linearly from the fixed point at the origin (orange histogram, top), and for a randomly chosen fixed point for each trajectory (blue histogram, bottom). (e) Mean Pearson correlation between $\boldsymbol{y}(\boldsymbol{x}(t))$ and its linear predictor as in (a), where $\boldsymbol{x}_0$ is the nearest (solid lines) or second-nearest (dotted lines) fixed point at $t \in \{5, 10, 15, 20\}$.

This radius is always >1 in the chaotic phase (Fig. 3), indicating that for large N, all fixed points are unstable.

Within the support, the distribution of eigenvalues is isotropic around the center [Fig. 3(a)]. We express the distribution by the fraction of eigenvalues further than r from the center $n_{\boldsymbol{x}}(r)$, termed the radial tail distribution. It obeys, again, a large deviation principle dominated by the solution $n_*(r)$ of

$$1 = \int dy\, \mu_*(y) \frac{g^2 \phi'(y)^2}{r^2 + n_*(r)\, g^2 \phi'(y)^2}. \tag{16}$$

We present the solution in Fig. 3(c), which shows that the unstable modes of fixed points are strongly underrepresented relative to a uniform spectrum. The overrepresentation of eigenvalues on the real line [Fig. 3(a)] and the smearing of the spectral radius [Fig. 3(c)] are known finite-size effects [44,45].

These results show that most directions of the fixed points are stable. Thus, within the vicinity of a fixed point, the dynamics are attracted towards it from most directions. However, eventually the unstable directions dominate and the trajectories are expelled from the fixed point along these directions.

VIII. IMPACT OF FIXED POINTS

Wainrib and Touboul [12] proposed an intriguing "topological explanation for the emergence of chaos": the dynamics meanders around the unstable fixed points and the interplay between attraction along their stable directions and repulsion along their unstable directions leads to the high sensitivity to perturbations which is characteristic for chaos. The radial separation of fixed points and dynamics uncovered above seemingly contradicts this explanation, while the alignment of dynamics and fixed points seems to support it. Using the empirical sample of fixed points, we now investigate their relation to the dynamics in more detail.

The topological explanation implicitly assumes that the local dynamics at the nearest fixed point $\boldsymbol{x}^*$ captures the essential properties of the full dynamics. We test this assumption by checking if the linearization provides a satisfactory prediction of the actual velocity $\boldsymbol{y}[\boldsymbol{x}(t)]$: given the linear predictor $\boldsymbol{y}_1(t) \equiv \boldsymbol{y}(\boldsymbol{x}^*) + \boldsymbol{y}'(\boldsymbol{x}^*)[\boldsymbol{x}(t) - \boldsymbol{x}^*]$, we quantify the accuracy of the prediction by the average Pearson correlation between $\boldsymbol{y}[\boldsymbol{x}(t)]$ and $\boldsymbol{y}_1(t)$, where $\boldsymbol{x}^*$ is the nearest fixed point at each time t for six trajectories of length 50. On average, the correlation with the nearest fixed points' linear predictor is approximately 0.5 [Figs. 4(a) and 4(b), purple error bar]. While this is a high correlation given the high dimensionality N, there is still a gap to a perfect correlation.

The gap to perfect correlation could be caused by the radial separation of fixed points and dynamics. Accordingly, we also linearize the velocity around points that are radially shrunk

or stretched from $\boldsymbol{x}(t)$. For these points, the linearized velocity still predicts the dynamics well at radial distances equal to the separation of fixed points and dynamics [Fig. 4(a)]. Thus, the gap is not caused by the radial separation. However, linearizing around points that are rotated by a fixed angle into a random tangential direction quickly degrades the predictive power at angular distances equal to the separation of fixed points and dynamics [Fig. 4(b)]. Thus, the remaining gap to perfect predictability must be attributed to the angular separation.

Interestingly, the angular separation between dynamics and fixed points is small compared to the angular separation between fixed points and random control points that are statistically equivalent to the dynamics [Fig. 4(b), green histogram]. The small angular distance at which the dynamics passes the fixed points can be shown to result from a majority of attractive directions (see Supplemental Material [[35], D1]): Almost every sample from a sphere around a fixed point moves towards it due to the predominance of stable eigenvalues. In contrast, the expected constant velocity of equivalent points $\boldsymbol{p}$ with elements drawn i.i.d. from the empirical measure, $p_i \overset{\text{i.i.d.}}{\sim} \mu_*$, is nonzero, which renders $\boldsymbol{p}$ highly repulsive.

In this sense, fixed points can be seen as landmarks of the stationary dynamics [46]: Starting from a random initial condition, the system moves towards the bulk of fixed points; see the decreasing correlation of $\boldsymbol{y}(\boldsymbol{x}(t))$ and $\boldsymbol{x}^* - \boldsymbol{x}$, where $\boldsymbol{x}^*$ is a randomly chosen fixed point, in Fig. 4(c). This is reflected by a strong correlation between the initial velocity and its linear predictor at randomly chosen fixed points [Fig. 4(d), bottom] and, most importantly, the origin [Fig. 4(d), top]. The correlation with both the origin and random fixed points quickly declines, hinting towards the system having arrived at the fixed point's manifold, although at higher distance to the origin. From here on, the system is mostly drawn and slightly less pushed by the ever-changing current subset of nearby fixed points, as indicated by the increasing and decreasing correlation between $\boldsymbol{y}(\boldsymbol{x}(t))$ and $\boldsymbol{y}_1$ based on a nearby fixed point [Fig. 4(e)].

IX. DISCUSSION

In this manuscript, we characterize the state space structure of a chaotic neural network using the distribution of fixed points. We identify a decoupling of the chaos transition and the emergence of unstable fixed points. We furthermore show a spatial separation between fixed points and dynamics. Interestingly, the principal components of fixed points and dynamics are aligned, despite their radial separation. Last, we establish the dynamic role of the fixed points as attractive landmarks for the trajectory.

In high-dimensional linear dynamical systems, May's pioneering stability analysis [1] enabled considerable insights into the dynamics of ecosystems [47]. In the nonlinear case, the number of fixed points can be determined if the velocity is generated by a homogeneous Gaussian potential [48]; in this case, it is even possible to determine the number of minima of the potential [49–51] with applications in deep learning [52,53]. The nonpotential case has been addressed in [12] for the random network (1) at $g = 1 + \varepsilon$ with $\varepsilon \ll 1$, in [13] for 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 [54]), and in [14,55] for a Lotka-Volterra model. Other nonhomogeneous cases have been studied in [56] with dynamics constrained to a sphere and in [57] with a metastable model where the distance of the fixed points to the origin determines which initial conditions decay or escape. For a recent review on stationary points of random fields, see [58].

Here, we go beyond the previous results and determine the distribution of fixed points, which includes their number, of the random neural network (1) for arbitrary $g > g_c$. To this end, we extend methods from random matrix theory [36,43] to compute the determinant of non-Hermitian random matrices with a correlation structure including low-rank terms. The analysis is restricted to the average number of fixed points, which provides an upper bound to the typical number of fixed points [14,55]. Our empirical results suggest that the bound is tight; for the density of fixed points, we also expect no difference between quenched and annealed behavior due to the excellent match of empirical and theoretical density.

The results presented here aim to pave the way towards a geometric understanding of the state space underlying high-dimensional chaotic systems. There are several directions for further research: First, it would be interesting to extend the analysis to more structured networks, for example in terms of low-rank perturbations [27], levels of symmetry [36,59,60], or population structure [22,23,30]. Second, the frustration created by the quenched rotation between the axes system, singled out by the elementwise application of the nonlinearity, and the eigensystem of the connectivity creates the complexity of the state 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 by analyzing their state space [61,62]. Here, we analyzed the phase space of a random reservoir which already allows universal computation if the readout is optimized [63]—more generally, learning with chaotic networks [29,64–66] is a direction of research that might be able to leverage the exponential number of fixed points and the associated capability for sequence processing.

ACKNOWLEDGMENTS

We are grateful to Günther Palm and Alexandre René for discussions about chaotic dynamics, and to Larry Abbott and David Clark for the suggestion to consider the subspace alignment. C.K. would like to acknowledge a helpful discussion with Vittorio Erba about entropy. This work was partly supported by the European Union Horizon 2020 Grant No. 945539 (Human Brain Project SGA3), funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Grant No. 368482240/GRK2416, the Helmholtz Association Initiative and Networking Fund under Project No. SO-092 (Advanced Computing Architectures, ACA), the German Federal Ministry for Education and Research (BMBF Grant No. 01IS19077A), and the Excellence Initiative of the German federal and state governments (Grant No. ERS PF-JARA-SDS005). Open access publication funded by the

Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Grant No. 491111487.

DATA AVAILABILITY

The data that support the findings of this article are openly available [67].

APPENDIX A: FROM KAC-RICE TO RANDOM MATRIX THEORY

Here, we show the equivalence of Eqs. (3) and (4). The $N(N+1)$ random variables $\mathbf{J}, \boldsymbol{\eta}$ are isomorphic, at almost every fixed $\boldsymbol{x}$, to the $N(N+1)$ variables $\boldsymbol{y}(\boldsymbol{x}), \boldsymbol{y}'(\boldsymbol{x})$. Thus, the integral over $\mathbf{J}$ and $\boldsymbol{\eta}$ implied by the expectation value in Eq. (3) can be substituted into an integral over $\boldsymbol{y}(\boldsymbol{x})$ and $\boldsymbol{y}'(\boldsymbol{x})$. Carrying out the substitution (detailed in the Supplemental Material [[35], A1]), one can use the Dirac $\delta(\boldsymbol{y})$ to solve the $\int d\boldsymbol{y}$ integral, which leads to

$$\rho(\boldsymbol{x}) = \int d\boldsymbol{y}'\, p_{\boldsymbol{x}}(\boldsymbol{y}=0, \boldsymbol{y}')\, |\det \boldsymbol{y}'|, \tag{A1}$$

where $p_{\boldsymbol{x}}(\boldsymbol{y}, \boldsymbol{y}') = \prod_{i=1}^{N} p_{\boldsymbol{x}}(y_i, \boldsymbol{y}_i')$ is the joint probability distribution (at $\boldsymbol{x}$) of $\boldsymbol{y}$ and $\boldsymbol{y}'$, and here $\boldsymbol{y}_i'$ is the ith row of $\boldsymbol{y}'$. Specifically, $p_{\boldsymbol{x}}(\boldsymbol{y}, \boldsymbol{y}')$ is a multivariate Gaussian,

$$\begin{pmatrix} y_i \\ \boldsymbol{y}_i' \end{pmatrix} \overset{\text{i.i.d.}}{\sim} \mathcal{N}\left[\begin{pmatrix} \mu_i(\boldsymbol{x}) \\ \boldsymbol{\mu}_i(\boldsymbol{x}) \end{pmatrix}, \begin{pmatrix} [\kappa(\boldsymbol{x}) + D] & \boldsymbol{k}(\boldsymbol{x})^{\mathrm{T}} \\ \boldsymbol{k}(\boldsymbol{x}) & \mathbf{K}(\boldsymbol{x}) \end{pmatrix} \right], \tag{A2}$$

with

$$\kappa(\boldsymbol{x}) = \frac{g^2}{N} \sum_k \phi(x_k)^2, \tag{A3}$$

$$[\boldsymbol{k}(\boldsymbol{x})]_k = \frac{g^2}{N} \phi(x_k) \phi'(x_k), \tag{A4}$$

$$[\mathbf{K}(\boldsymbol{x})]_{kl} = \delta_{kl} \frac{g^2}{N} \phi'(x_k)^2. \tag{A5}$$

Equation (A1) is an alternative formulation of the Kac-Rice formula [34,68]. One may rewrite $p_{\boldsymbol{x}}(\boldsymbol{y}, \boldsymbol{y}') = p_{\boldsymbol{x}}(\boldsymbol{y})\, p_{\boldsymbol{x}}(\boldsymbol{y}'|\boldsymbol{y})$ into the marginal distribution of the velocity $\boldsymbol{y}$ and the conditional distribution of the Jacobian $\boldsymbol{y}'$. Following [[69], Appendix A2], we find the probability to be at zero velocity, termed level probability $p_{\mathrm{L}}(\boldsymbol{x}) \equiv p_{\boldsymbol{x}}(\boldsymbol{y}=0)$,

$$p_{\mathrm{L}}(\boldsymbol{x}) = \frac{1}{\sqrt{2\pi[\kappa(\boldsymbol{x}) + D]}^N} e^{-\frac{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}{2[\kappa(\boldsymbol{x})+D]}}. \tag{A6}$$

Furthermore, the conditional distribution of the Jacobian is [[69], Appendix A2]

$$p_{\boldsymbol{x}}(\boldsymbol{y}'|\boldsymbol{y}=0) = \prod_i \mathcal{N}[\boldsymbol{y}_i'|\mathbf{M}_i(\boldsymbol{x}), \mathbf{C}(\boldsymbol{x})], \tag{A7}$$

where the mean conditioned Jacobian is

$$M_{ij}(\boldsymbol{x}) = -\delta_{ij} - \frac{g^2}{N} \frac{x_i \phi(x_j) \phi'(x_j)}{\kappa(\boldsymbol{x}) + D}, \tag{A8}$$

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

$$C_{nm} = \delta_{nm} \frac{g^2}{N} \phi'(x_n)^2 - \frac{g^4}{N^2} \frac{\phi(x_n)\phi'(x_n)\phi(x_m)\phi'(x_m)}{\kappa(\boldsymbol{x}) + D}. \tag{A9}$$

The random matrix characterized in Eq. (A7) can be expressed as

$$\boldsymbol{y}' = \mathbf{M}(\boldsymbol{x}) + \mathbf{X}\, \boldsymbol{\Sigma}(\boldsymbol{x}), \tag{A10}$$

where $X_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1/N)$ and

$$\boldsymbol{\Sigma}(\boldsymbol{x}) = \left[\mathbb{1} - \frac{g^2}{N} \frac{\phi(\boldsymbol{x})\phi(\boldsymbol{x})^{\mathrm{T}}}{\kappa(\boldsymbol{x}) + D[1 + \sqrt{1 + \kappa(\boldsymbol{x})/D}]} \right] \times g\, \mathrm{diag}[\phi'(\boldsymbol{x})], \tag{A11}$$

which can be checked by comparing moments.

Concluding, substituting Eq. (3) to get Eq. (A1), conditioning on $\boldsymbol{y} = 0$, and plugging in the expressions Eqs. (A6) and (A10) leads to Eq. (4). The remaining technical difficulty is to solve the random matrix theory problem of computing the expected modulus determinant of Eq. (A10). As detailed in the Supplemental Material [[35], A2], the determinant is given by Eq. (5).

APPENDIX B: EXPECTED EMPIRICAL MEASURE

In this Appendix, 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 [30] and [39]. We consider the characteristic functional

$$Z[j] = \left\langle e^{i j^{\mathrm{T}} \mu_{\boldsymbol{x}}} \right\rangle_{\boldsymbol{x} \sim \rho(\boldsymbol{x})} = \left\langle e^{i \frac{1}{N} \sum_{i=1}^{N} j(x_i)} \right\rangle_{\boldsymbol{x} \sim \rho(\boldsymbol{x})}, \tag{B1}$$

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 the average with respect to $\rho(\boldsymbol{x}) / [\int d\boldsymbol{z}\, \rho(\boldsymbol{z})]$. We define the scaled cumulant generating functional as $W_N[j] = \frac{1}{N} \ln Z[N j]$. The expected empirical measure is the first Taylor coefficient of W_N,

$$\mu_*(y) = \frac{\delta}{\delta j(y)} W_N[j]|_{j(y)=0}. \tag{B2}$$

Plugging in the fixed point distribution given by Eq. (8), we have

$$W_N[j] = \frac{1}{N} \ln \int d\boldsymbol{x}\, e^{-NS(\boldsymbol{x}) + i \sum_i j(x_i)} - c, \tag{B3}$$

$$c = \frac{1}{N} \ln \int d\boldsymbol{x} \rho(\boldsymbol{x}) = \frac{1}{N} \ln \langle N_{\mathrm{fp}} \rangle. \tag{B4}$$

The rate c is the *topological complexity* [12].

Next, we evaluate the integral in Eq. (B3). To this end, we first express $S(\boldsymbol{x}) = S[\mu_{\boldsymbol{x}}]$ by the empirical measure by expressing the scalars u, κ, ζ, and z_* through $\mu_{\boldsymbol{x}}$ instead of $\boldsymbol{x}$, specifically,

$$u[\mu_{\boldsymbol{x}}] = \int dy\, y^2 \mu_{\boldsymbol{x}}(y), \tag{B5}$$

$$\kappa[\mu_{\boldsymbol{x}}] = g^2 \int dy\, \phi(y)^2 \mu_{\boldsymbol{x}}(y), \tag{B6}$$

$$\zeta[\mu_{\boldsymbol{x}}] = -\frac{1}{2} z_*[\mu_{\boldsymbol{x}}] + \frac{1}{2} \int dy\, \ln[1 + z_*[\mu_{\boldsymbol{x}}] g^2 \phi'(y)^2], \tag{B7}$$

$$1 = \int dy\, \frac{g^2 \phi'(y)^2}{1 + z_*[\mu_{\boldsymbol{x}}] g^2 \phi'(y)^2}. \tag{B8}$$

Next, we introduce an auxiliary field $\mu(y)$ that we use to replace the $\boldsymbol{x}$-dependent field $\mu_{\boldsymbol{x}}(y)$. For this to be correct, we multiply the integrand by the functional Dirac constraint $\delta[\mu - \mu_{\boldsymbol{x}}] \equiv \lim_{M\to\infty} \prod_{m=1}^{M} \delta[\mu(y_m) - \mu_{\boldsymbol{x}}(y_m)]$, where $\{y_1, \ldots, y_M\} \overset{M\to\infty}{\to} \mathbb{R}$ is a discretization of the real line. Then, we have to integrate over $\mu(y)$ in a functional sense, $\int \mathcal{D}\mu \equiv \lim_{M\to\infty} \int_{-\infty}^{\infty} \prod_{m=1}^{M} d\mu_m$, $\mu_m = \mu(y_m)$ discretizes μ, for the replacement to occur at every point in the $\boldsymbol{x}$ integration. Lastly, we replace the functional Dirac constraint by its Fourier integral representation,

$$\begin{aligned}\delta[\mu - \mu_{\boldsymbol{x}}] &= \int \mathcal{D}\tilde{\mu}\, e^{-iN\tilde{\mu}^{\mathrm{T}}(\mu-\mu_{\boldsymbol{x}})} \\ &= \int \mathcal{D}\tilde{\mu}\, e^{-iN\tilde{\mu}^{\mathrm{T}}\mu + i\sum_{i=1}^{N}\tilde{\mu}(x_i)},\end{aligned} \tag{B9}$$

where

$$\int \mathcal{D}\tilde{\mu} \equiv \lim_{M\to\infty} \int_{-\infty}^{\infty} \prod_{m=1}^{M} N\frac{d\tilde{\mu}_m}{2\pi}, \tag{B10}$$

and $\tilde{\mu}_m = \tilde{\mu}(y_m)$ discretizes $\tilde{\mu}$. With the auxiliary fields μ and $\tilde{\mu}$, the $\boldsymbol{x}$ integral in Eq. (B3) formally factorizes,

$$W_N[j] = \frac{1}{N} \ln \int \mathcal{D}\mu\mathcal{D}\tilde{\mu}\, e^{-iN\tilde{\mu}^{\mathrm{T}}\mu + N\ln\Omega[\mu,\tilde{\mu},j]} - c, \tag{B11}$$

$$\Omega[\mu,\tilde{\mu},j] = \int \frac{dx}{\sqrt{2\pi(\kappa[\mu]+D)}} e^{-\frac{x^2}{2(\kappa[\mu]+D)} + \zeta[\mu] + ij(x) + i\tilde{\mu}(x)}. \tag{B12}$$

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. (B11) is proportional to N, and thus the leading order is given by the saddle-point approximation,

$$W_N[j] = -i\tilde{\mu}_*[j]^{\mathrm{T}}\mu_*[j] + \ln\Omega\{\mu_*[j], \tilde{\mu}_*[j], j\} - c, \tag{B13}$$

where

$$\begin{aligned}\mu_*[j](y) &= \left.\frac{\delta \ln\Omega}{\delta\, i\tilde{\mu}(y)}\right|_{\mu_*[j],\tilde{\mu}_*[j]}, \\ i\tilde{\mu}_*[j](y) &= \left.\frac{\delta \ln\Omega}{\delta\, \mu(y)}\right|_{\mu_*[j],\tilde{\mu}_*[j]}\end{aligned} \tag{B14}$$

are the maxima of the exponent of the integrand in Eq. (B11). With this, we can compute the expectation value of $\mu_{\boldsymbol{x}}$ in saddle-point approximation: Evaluating Eq. (B2) gives

$$\begin{aligned}\mu_*(y) &= \left.\frac{\delta \ln\Omega\{\mu_*[0], \tilde{\mu}_*[0], j\}}{\delta\, j(y)}\right|_{j=0} \\ &= \mu_*[0](y),\end{aligned} \tag{B15}$$

where we used Eqs. (B14) 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. (B15) only acts on the explicit dependency of Ω on j.

Next, we solve the saddle-point equations (B14) for $j = 0$. The first saddle-point equation (by straightforward functional differentiation) is

$$\mu_*(y) = \frac{e^{-\frac{y^2}{2(\kappa[\mu_*]+D)} + \zeta[\mu_*] + i\tilde{\mu}_*(y)}}{\sqrt{2\pi(\kappa[\mu_*]+D)}\Omega[\mu_*,\tilde{\mu}_*]}. \tag{B16}$$

The second saddle-point equation

$$i\tilde{\mu}_*(y) = \frac{1}{\Omega}\left(\frac{\partial\Omega}{\partial\kappa}\frac{\delta\kappa[\mu]}{\delta\mu(y)} + \frac{\partial\Omega}{\partial\zeta}\frac{\delta\zeta[\mu]}{\delta\mu(y)}\right)\Bigg|_{\mu_*,\tilde{\mu}_*} \tag{B17}$$

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

$$\frac{\delta\kappa[\mu]}{\delta\mu(y)} = g^2\phi(y)^2, \tag{B18}$$

and we need the derivative of $\zeta[\mu]$ as in Eq. (B7). We find

$$\begin{aligned}\frac{\delta\zeta[\mu]}{\delta\mu(y)} = -\frac{1}{2}\frac{\delta z_*[\mu]}{\delta\mu(y)}\left[1 - \int dx\,\mu(x)\frac{g^2\phi'(x)^2}{1 + z_*[\mu]g^2\phi'(x)^2}\right] \\ + \frac{1}{2}\ln\{1 + z_*[\mu]g^2\phi'(y)^2\},\end{aligned} \tag{B19}$$

where the first part vanishes due to the definition of z_*; see Eq. (B8). Concluding, the saddle-point equation for $\tilde{\mu}$ is

$$\begin{aligned}i\tilde{\mu}_*(y) = \frac{g^2\phi(y)^2}{2(\kappa[\mu_*]+D)}\left(\frac{\langle x^2\rangle_{\mu_*}}{\kappa[\mu_*]+D} - 1\right) \\ + \frac{1}{2}\ln\{1 + z_*[\mu_*]g^2\phi'(y)^2\}.\end{aligned} \tag{B20}$$

The expected empirical measure is determined by the simultaneous solution of Eqs. (B16) and (B20). Combining them leads to Eq. (10). At the saddle point,

$$\alpha = z_*[\mu_*]g^2, \tag{B21}$$

$$\beta = \kappa[\mu_*] + D, \tag{B22}$$

$$\gamma = g^2\left(1 - \frac{u[\mu_*]}{\beta}\right). \tag{B23}$$

APPENDIX C: LARGE DEVIATIONS OF THE EMPIRICAL MEASURE

Here, we discuss how the empirical measure $\mu_{\boldsymbol{x}}(y)$ is distributed in function space due to the fixed point distribution $\boldsymbol{x} \sim \rho(\boldsymbol{x})/\langle N_{\mathrm{fp}}\rangle$ in Eq. (8). Thereby, we go beyond the expected empirical measure derived in Appendix B. The starting point is the scaled cumulant generating functional in the saddle-point approximation given by Eq. (B13). Due to its existence and differentiability, the Gärtner-Ellis theorem holds [39], [[70], 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. 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.

Carrying out the Legendre transform (see the Supplemental Material [[35], B2]), we find

$$H[\mu] = D_{\mathrm{KL}}(\mu\|\nu) - W_N[j^{\mu}_*], \tag{C1}$$

where j_*^μ is the solution of $\mu(y) = \mu_*[j](y)$ and $D_{\mathrm{KL}}(\rho_1 \| \rho_2) = \langle \ln[\rho_1(x)/\rho_2(x)] \rangle_{\rho_1(x)}$ denotes the Kullback-Leibler divergence between the probability distribution functions ρ_1 and ρ_2. The reference function here is

$$\nu(y) = \frac{e^{-\frac{y^2}{2(\kappa[\mu]+D)} + \zeta[\mu] + i\tilde{\mu}_*[j_*^\mu](y)}}{\Omega_* \sqrt{2\pi(\kappa[\mu]+D)}}. \tag{C2}$$

The rate functional given by Eq. (C1) is convex and its minimum is the expected empirical measure given by Eq. (10); deviations thereof are suppressed exponentially.

APPENDIX D: LARGE DEVIATIONS OF CONTRACTIONS

Quantities that can be derived from the empirical measure, such as the scaled squared norm of fixed points $u[\mu] = \int dy\, y^2 \mu(y)$ and the spectral radius of the the Jacobian at fixed points $R^2[\mu] = g^2 \int dy\, \mu(y) \phi'(y)^2$ are termed contractions in the context of large deviations theory. According to the *contraction principle* [39], contractions inherit the large deviations principle of the empirical measure.

1. Scaled squared norm

We start with the scaled squared norm. Specifically, the contraction principle states that $p(u) \doteq \exp[-NI(u)]$ and that the rate function $I(u)$ can be derived from the rate functional for μ,

$$I(u) = \inf_{\mu:\, \int y^2 \mu(y)\, dy = u} H[\mu], \tag{D1}$$

repeating Eq. (12). The minimization can be carried out using a Lagrange multiplier, and hence we have to optimize,

$$\mathcal{L}[\mu; \lambda] = H[\mu] - \lambda \left[u - \int dy\, y^2 \mu(y) \right], \tag{D2}$$

for both λ and μ. Writing H as the Legendre transform of W, $H[\mu] = \inf_j (i\mu^{\mathrm{T}} j - W_N[j])$, one can carry out the optimization over μ first,

$$\frac{\delta}{\delta \mu(y)} [i\mu^{\mathrm{T}} j - \lambda(u - v^{\mathrm{T}} \mu)] \stackrel{!}{=} 0, \tag{D3}$$

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

$$\begin{aligned} I(u) &= \inf_\lambda (\lambda \mu^T v - W[-i\lambda v])|_{v^{\mathrm{T}} \mu \stackrel{!}{=} u} \\ &= \inf_\lambda (\lambda u - W_N[-i\lambda v]), \end{aligned} \tag{D4}$$

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

$$u = \frac{\partial}{\partial \lambda} W_N[-i\lambda v] = v^{\mathrm{T}} \mu_*[-i\lambda v]. \tag{D5}$$

In practice, we solve this by computing $\mu_*[-i\lambda v]$ in the same way as we computed $\mu_*[0]$ in Appendix A for a sequence of values λ. Then, we check which u they correspond to using Eq. (D5). Lastly, we plug into Eq. (D4). $I(u)$ is minimized at the expected squared norm given by Eq. (11) and deviations thereof are suppressed exponentially. The standard deviation computed from Eq. (D4) is shown by the shaded area in Fig. 1(d). The full (asymmetric) distribution is shown in the Supplemental Material [[35], Fig. 3] for several parameters.

2. Spectral radius

The contraction principle and the method in Appendix D 1 can also be used to derive the distribution $p(\mathfrak{r}) \doteq e^{-NJ(\mathfrak{r})}$ of the squared spectral radius $\mathfrak{r} \equiv R^2$ with

$$J(\mathfrak{r}) = \inf_{\mu:\, \int dy\, g^2 \phi'(y)^2 \mu(y) = \mathfrak{r}} H[\mu]. \tag{D6}$$

Specifically, one needs to compute $\mu_*[-i\lambda w]$ for a sequence of values λ, then solve $\mathfrak{r} = w^{\mathrm{T}} \mu_*[-i\lambda w]$ for $\lambda(\mathfrak{r})$ and get the rate function as $J(\mathfrak{r}) = \lambda(\mathfrak{r})\mathfrak{r} - W_N[-i\lambda(\mathfrak{r}) w]$. Equation (D6) is minimized at the expected squared spectral radius, i.e., the square of Eq. (15), and deviations thereof are suppressed exponentially. As detailed in the Supplemental Material [[35], E2], Eq. (D6) captures the dominant finite-size smearing of the spectral radius observable in Fig. 3(a).

APPENDIX E: TOPOLOGICAL COMPLEXITY

The topological complexity c is the rate by which the expected number of fixed points grows with the number of units, $\langle N_{\mathrm{fp}} \rangle \doteq e^{cN}$. Reference [12] proposed the topological complexity as a structural indicator of chaoticity, complementary to the maximum Lyapunov exponent which they refer to as dynamical complexity.

To compute c, we start from the scaled cumulant generating functional given by Eq. (B13), where c appears as the normalizing constant asserting $W_N[0] \stackrel{!}{=} 0$. Thus,

$$c = -i\tilde{\mu}_*^{\mathrm{T}} \mu_* + \ln \Omega[\mu_*, \tilde{\mu}_*]. \tag{E1}$$

Plugging Eqs. (B20), (10), and (B12) into Eq. (E1), we get Eq. (13).

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

[2] S. B. Laughlin and T. J. Sejnowski, Communication in neuronal networks, Science **301**, 1870 (2003).

[3] S. Loomba, J. Straehle, V. Gangadharan, N. Heike, A. Khalifa, A. Motta, N. Ju, M. Sievers, J. Gempt, H. S. Meyer *et al.*, Connectomic comparison of mouse and human cortex, Science **377**, eabo0924 (2022).

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

[5] E. Beninca, J. Huisman, R. Heerkloss, K. D. Jöhnk, P. Branco, E. H. Van Nes, M. Scheffer, and S. P. Ellner, Chaos in a long-term experiment with a plankton community, Nature (London) **451**, 822 (2008).

[6] T. L. Rogers, B. J. Johnson, and S. B. Munch, Chaos is not rare in natural ecosystems, Nat. Ecol. Evol. **6**, 1105 (2022).

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

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

[9] M. Monteforte and F. Wolf, Dynamic flux tubes form reservoirs of stability in neuronal circuits, Phys. Rev. X **2**, 041007 (2012).

[10] I. Ispolatov, V. Madhok, S. Allende, and M. Doebeli, Chaos in high-dimensional dissipative dynamical systems, Sci. Rep. **5**, 12506 (2015).

[11] E. N. Lorenz, Deterministic nonperiodic flow, J. Atmos. Sci. **20**, 130 (1963).

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

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

[14] V. Ros, F. Roy, G. Biroli, G. Bunin, and A. M. Turner, Generalized Lotka-Volterra equations with random, nonreciprocal interactions: The typical number of equilibria, Phys. Rev. Lett. **130**, 257401 (2023).

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

[16] M. Helias and D. Dahmen, *Statistical Field Theory for Neural Networks* (Springer International Publishing, Cham, 2020), Vol. 970.

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

[18] D. G. Clark, L. F. Abbott, and A. Litwin-Kumar, Dimension of activity in random neural networks, Phys. Rev. Lett. **131**, 118401 (2023).

[19] R. Engelken, F. Wolf, and L. F. Abbott, Lyapunov spectra of chaotic recurrent neural networks, Phys. Rev. Res. **5**, 043044 (2023).

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

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

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

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

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

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

[26] I. D. Landau and H. Sompolinsky, Coherent chaos in a recurrent neural network with structured connectivity, PLoS Comput. Biol. **14**, e1006309 (2018).

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

[28] Ł. Kuśmierz, S. Ogawa, and T. Toyoizumi, Edge of chaos and avalanches in neural networks with heavy-tailed synaptic weight distribution, Phys. Rev. Lett. **125**, 028101 (2020).

[29] C. Keup, T. Kühn, D. Dahmen, and M. Helias, Transient chaotic dimensionality expansion by recurrent networks, Phys. Rev. X **11**, 021064 (2021).

[30] 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).

[31] A. Wardak and P. Gong, Extended Anderson criticality in heavy-tailed neural networks, Phys. Rev. Lett. **129**, 048103 (2022).

[32] M. Kac, On the average number of real roots of a random algebraic equation, Bull. Am. Math. Soc. **49**, 314 (1943).

[33] S. O. Rice, Mathematical analysis of random noise, Bell Syst. Tech. J. **24**, 46 (1945), reprinted in [71].

[34] J.-M. Azaïs and M. Wschebor, *Level Sets and Extrema of Random Processes and Fields* (Wiley, New York, 2009).

[35] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.7.023203 for detailed derivations and further information, which includes Ref. [72].

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

[37] K. Levenberg, A method for the solution of certain nonlinear problems in least squares, Quantum Appl. Math. **2**, 164 (1944).

[38] D. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, J. Soc. Ind. Appl. Math. **11**, 431 (1963).

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

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

[41] R. Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science*, Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press, Cambridge, 2018).

[42] K. Krishnamurthy, T. Can, and D. J. Schwab, Theory of gating in recurrent neural networks, Phys. Rev. X **12**, 011011 (2022).

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

[44] A. Edelman, E. Kostlan, and M. Shub, How many eigenvalues of a random matrix are real? J. Am. Math. Soc. **7**, 247 (1994).

[45] B. Rider and C. D. Sinclair, Extremal laws for the real Ginibre ensemble, Ann. Appl. Probab. **24**, 1621 (2014).

[46] M. Rabinovich, R. Huerta, and G. Laurent, Transient dynamics for neural processing, Science **321**, 48 (2008).

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

[48] 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).

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

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

[51] 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).

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

[53] 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*, Proc. Mach. Learn. Res. Vol. 38, edited by G. Lebanon and S. V. N. Vishwanathan (PMLR, San Diego, CA, 2015), pp. 192–204.

[54] 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).

[55] V. Ros, F. Roy, G. Biroli, and G. Bunin, Quenched complexity of equilibria for asymmetric generalized Lotka-Volterra equations, J. Phys. A: Math. Theor. **56**, 305003 (2023).

[56] Y. V. Fyodorov, Topology trivialization transition in random nongradient autonomous ODEs on a sphere, J. Stat. Mech. (2016) 124003.

[57] S. B. Fedeli, Y. V. Fyodorov, and J. R. Ipsen, Nonlinearity-generated resilience in large complex systems, Phys. Rev. E **103**, 022201 (2021).

[58] V. Ros and Y. V. Fyodorov, The high-dimensional landscape paradigm: Spin glasses and beyond, in *Spin Glass Theory and Far Beyond* (World scientific publishing, Singapore, 2023), Chap. 6, pp. 95–114.

[59] 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).

[60] K. Berlemont and G. Mongillo, Glassy phase in dynamically-balanced neuronal networks, BioRxiv: 484348.

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

[62] S. Vyas, M. D. Golub, D. Sussillo, and K. V. Shenoy, Computation through neural population dynamics, Annu. Rev. Neurosci. **43**, 249 (2020).

[63] 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).

[64] T. Toyoizumi and L. F. Abbott, Beyond the edge of chaos: Amplification and temporal integration by recurrent networks in the chaotic regime, Phys. Rev. E **84**, 051908 (2011).

[65] 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*, edited by D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett (Curran Associates, Inc., Montreal, Canada, 2016).

[66] 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, Nat. Mach. Intell. **4**, 564 (2022).

[67] J. Stubenrauch, C. Keup, A. C. Kurth, M. Helias, and A. van Meegen, Fixed point geometry in chaotic neural networks, Zenodo (2025), doi:10.5281/zenodo.15211341.

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

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

[70] R. S. Ellis, An overview of the theory of large deviations and applications to statistical mechanics, Scand. Actuarial J. **1995**, 97 (1995).

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

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