Extended Anderson Criticality in Heavy-Tailed Neural Networks

Asem Wardak and Pulin Gong*
School of Physics, University of Sydney, New South Wales 2006, Australia

(Received 15 December 2021; revised 8 May 2022; accepted 5 July 2022; published 22 July 2022)

We investigate the emergence of complex dynamics in networks with heavy-tailed connectivity by developing a non-Hermitian random matrix theory. We uncover the existence of an extended critical regime of spatially multifractal fluctuations between the quiescent and active phases. This multifractal critical phase combines features of localization and delocalization and differs from the edge of chaos in classical networks by the appearance of universal hallmarks of Anderson criticality over an extended region in phase space. We show that the rich nonlinear response properties of the extended critical regime can account for a variety of neural dynamics such as the diversity of timescales, providing a computational advantage for persistent classification in a reservoir setting.

DOI: 10.1103/PhysRevLett.129.048103

Introduction.—In a diverse range of physical, biological, financial, and ecological systems, complex dynamics fluctuating across multiple scales emerge from a large number of interacting, nonlinear units with heterogeneous properties. Understanding the organizing principles and behavior of such complex dynamics is a longstanding topic of interest across these diverse fields [1]. In neuroscience and machine learning, neural networks with many interacting neurons likewise exhibit complex dynamics with large fluctuations that are critical for their information processing abilities on real-world inputs [2]. However, the network mechanisms and fundamental computational capabilities of complex neural dynamics remain elusive.

The classical formulation of complex dynamics in systems with many interacting elements is based on neural networks with homogeneous connectivity [3]. Mathematical approaches such as mean-field theory and random matrix theory have robustly predicted a phase of chaotic activity with global, homogeneous (i.e., delocalized) fluctuations existing adjacent to an ordered, silent regime, enabling the analysis of a wide range of systems with a characteristic spatial scale [4]. The edge of the ordered and chaotic phases gives rise to critical phenomena that is thought to be necessary for these systems to perform useful computations [5]. However, growing evidence has shown that coupling heterogeneity is widespread in complex systems such as biological [6] and artificial neural networks [7,8], underscoring the need to understand the fundamental dynamical and computational mechanisms of such heterogeneity.

Here, we study the dynamics of random neural networks with heterogeneous, heavy-tailed connectivity. After describing the fixed points of the system using a Lévy mean-field approach, we develop a novel non-Hermitian random matrix theory for column-structured heavy-tailed matrices to analyze the statistical fluctuations of random neural networks around the fixed point. This theory reveals a new regime with correlated multifractal modes that are neither localized nor delocalized, but have aspects of both (see Fig. 1). Multifractality is characterized by the appearance of differing, nontrivial structures appearing simultaneously over a wide variety of scales [9], and is a hallmark of Anderson transitions (criticality) [10]. Anderson transitions were first described in the context of disordered electronic systems with localized and metallic (i.e., delocalized) phases [11], and have since been analyzed in a broad sense in a wide range of systems, including treelike

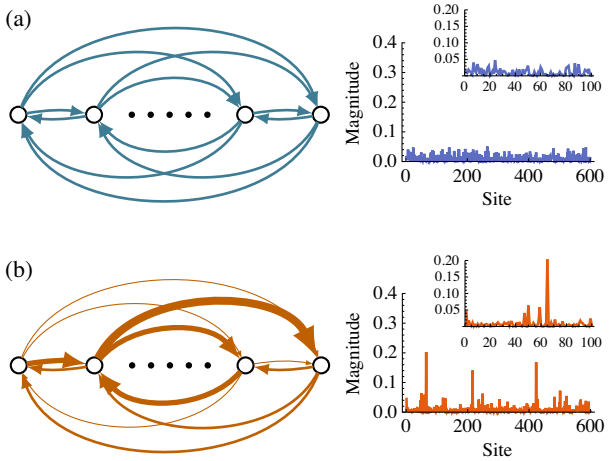

FIG. 1. Schematic of homogeneous and heavy-tailed networks (left) and their temporal activity fluctuations over neural sites (right). (a) Homogeneous neural network (blue left). Activity fluctuations are delocalized and spread evenly over the spatial extent of the system (right). (b) Heavy-tailed neural network (red left) with heterogeneous weights (line thickness). Large activity fluctuations are multifractal with a mixture of localization and delocalization over system sites (right) visible over multiple scales (right inset).

 © 2022 American Physical Society

Bethe lattices [12] and those exhibiting conventional second-order phase transitions [10].

We illustrate that the heavy-tailed heterogeneity in connectivity enables Anderson criticality to emerge in a broad parameter regime. The correlated multifractal modes characteristic of this extended critical regime are able to explain a range of realistic neural dynamics, including correlated fluctuations with low-dimensional features [13], long-range correlations [14], and a diversity of timescales [15]. Importantly, these correlated multifractal modes provide a profound computational advantage in the setting of real-time reservoir computing by allowing for a persistent form of dimensionality expansion, which is not possible in classical homogeneous systems.

Network model and fixed points.—We begin by extending the seminal random neural network model with interacting nonlinear units analysed by Sompolinsky *et al.* and others [3,4], which has the dynamics

$$h_i'(t) = -h_i(t) + g\sum_{j=1}^{N} J_{ij}\phi[h_j(t)], \tag{1}$$

where $h_i(t)$ is the input of the ith neuron at time t, J_{ij} is the strength of the connection from neuron j to neuron i, and ϕ is a scalar nonlinearity that determines the neural firing rate given the input. The theoretical results in this Letter apply for all differentiable, sublinear ϕ, and the figures use $\phi = \tanh$ for comparability with previous models [3]. The coefficient g is the gain parameter of the synaptic input. Our aim is to investigate networks with heavy-tailed heterogeneity that do not fall within the purview of the Gaussian large-size limit. Finite-size heterogeneous networks can be investigated to some degree using perturbative finite-order corrections to the homogeneous Gaussian limit [16]. We instead take an approach that is exact in the large network limit by regarding each J_{ij} as an independent random variable whose second moment is not finite, so that its probability density has a power-law asymptotic tail,

$$p_{J_{ij}}(x) \overset{|x|\to\infty}{\sim} \frac{C_\alpha}{2N|x|^{1+\alpha}}, \tag{2}$$

where $C_\alpha \coloneqq \Gamma(1+\alpha)\sin(\pi\alpha/2)/\pi$ is a normalization factor and $1 < \alpha < 2$. Such heavy-tailed connectivity has been observed in the Drosophila central brain [6], in successfully trained artificial neural networks [8], and in spin-glass systems with strong disorder [17]; theoretically elucidating the dynamical impact of such heterogeneity has drawn increasing attention [18]. As we discover in our Jacobian analysis below, the zero fixed point is unstable so that network activity is nonzero and has a macroscopic number of nonzero fixed points at any gain g for $\alpha < 2$ (see Supplemental Material [19] for Lévy mean-field theory), in contrast to homogeneous networks ($\alpha = 2$) for which the network exhibits a single stable fixed point when $g < 1$.

Network stability and heavy-tailed random matrix theory.—To determine the local stability of the network around the fixed points, we analyze the Jacobian matrix $-I + gJ\mathrm{diag}_j[\phi'(h_j)]$ obtained from Eq. (1) where $\mathrm{diag}_j\chi_j$ denotes the diagonal matrix with entries χ_j. Shifting this Jacobian matrix yields a stability matrix obtained by scaling the columns of gJ by $\phi'(h_j)$, which has the form of a column-structured non-Hermitian random matrix [24]. Since the heavy-tailed matrix J (and thus $J\mathrm{diag}_j\chi_j$) has a locally treelike structure [21], we develop a new cavity approach for column-structured non-Hermitian heavy-tailed random matrices to obtain the spectral density and eigenvector localization properties of the Jacobian [19]. In recent years, cavity approaches have been used in asymmetrically disordered contexts involving non-Hermitian random ensembles by mapping the problem back to a symmetric, Hermitian system of twice the dimensionality [22]. Our cavity approach to the column-structured matrix $J\mathrm{diag}_j\chi_j$ for any χ_j with $\langle|\chi_j|^\alpha\rangle_j < \infty$ yields the spectral density

$$\rho(z) = \frac{y_*^2 - 2|z|^2 y_* \partial_{|z|^2} y_*}{\pi}\left\langle \frac{|\chi_i|^2 SS'}{(|z|^2 + |\chi_i|^2 y_*^2 SS')^2}\right\rangle_i \tag{3}$$

(see Supplemental Material [19] for mathematical derivations) as a function of eigenvalue modulus $|z|$. The random variables $S, S' \sim L(\alpha/2, 1, 0, C_\alpha/4C_{\alpha/2})$ are independent, skewed $\alpha/2$-stable samples, $\langle..\rangle_i$ denotes averaging over i, S, and S', and the fixed-point variable y_* is found by solving the equation

$$1 = \left\langle \left(\frac{|\chi_i|^2 S}{|z|^2 + y_*^2|\chi_i|^2 SS'}\right)^{\alpha/2}\right\rangle_i. \tag{4}$$

Our random matrix theory thus unifies both classical results in block-structured non-Hermitian random matrix theory [4,23] for which $\alpha = 2$ and the random variables S, S' reduce to the constant 1, and unstructured heavy-tailed non-Hermitian random matrices [21] for which $\chi = 1$.

Figure 2(a) shows the eigenvalue density $\rho(z)$ of the Jacobian and its numerical validation as a function of eigenvalue modulus $|z|$. A key characteristic of heavy-tailed neural networks is the infinite spectral radius r_0, given by the point at which $y_* = 0$, so that the zero fixed point only occurs for zero gain. However, since the eigenvalue density of the heavy-tailed Jacobian is exponentially suppressed at large radius [19,21], a characteristic spectral radius r_p can be defined by the eigenvalue modulus $|z|$ at which the parameter y_* drops to a fraction p of its value at $z = 0$. Meanwhile, the spectral radius r_0 of the Jacobian for classical networks ($\alpha = 2$) is finite and is obtained by a spectral analysis of finite-variance block-structured random matrices [4,23]. In this sense, y_* behaves as an order parameter indicating the transition between microscopic

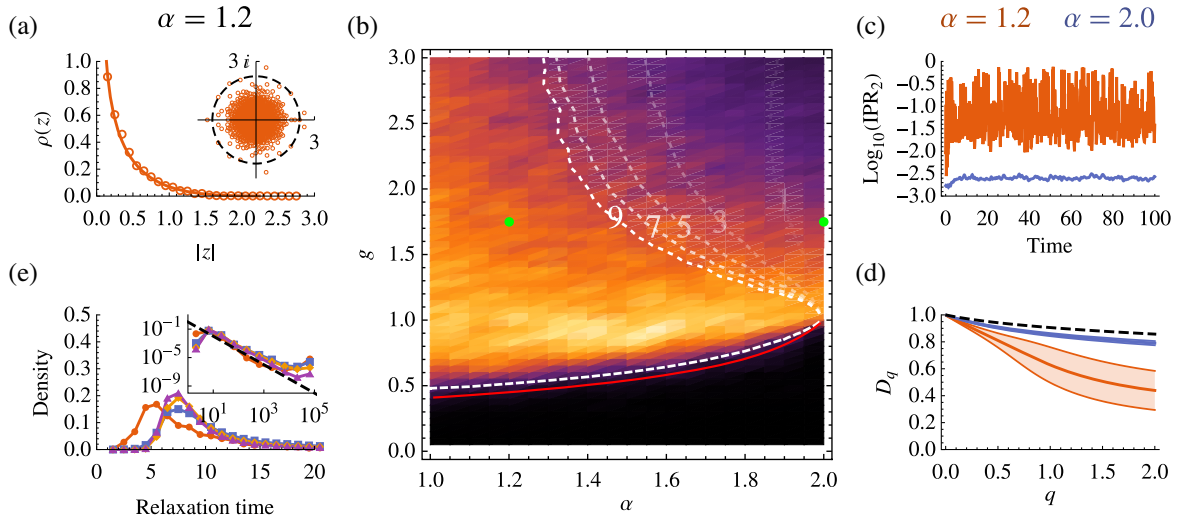

FIG. 2. Extended Anderson critical regime in heavy-tailed random neural networks ($N = 1000$ in simulations). (a) Jacobian spectra $J\mathrm{diag}_j[\tanh'(h_j)]$ in the stationary state (lines) with characteristic spectral radius $r_{0.01}$ (dashed circle) and their numerical validation (dots). (b) Phase diagram over heavy-tailed index α and gain g. Extended Anderson critical regime predicted by Eq. (5) (lines labeled by n) corresponds with high relaxation time variance (colored). Green dots at $g = 1.75$. (c) Log-IPR of the heavy-tailed network activity fluctuation vector (red). (d) Mean multifractal dimension D_q shaded to one standard deviation over time. D_q estimate for a random delocalized vector on the N-sphere (dashed) for comparison. (e) Distribution of the half-width at quarter-maximum of the local-field autocorrelation for $g = 0.75$, 1.5, 2.25, 3 (red, blue, yellow, purple). Log-log plot (inset) with comparison power-law line of exponent -2.0.

and macroscopic numbers of eigenstates relative to system size N at a given eigenvalue modulus $|z|$. Thus, heavy-tailed networks exhibit a quasiordered regime with a fixed point near zero whose magnitude is suppressed due to the microscopic number of eigenstates above the Jacobian stability line $\mathrm{Re}z = 1$. A continuous transition parametrized by $p \ll 1$ is then defined by the value of the gain parameter g at which the characteristic spectral radius $r_p = 1$, distinguishing a quiescent phase from a chaotic regime (Fig. 2(b), lower dashed line). This transition is consistent with the point at which neural activity predicted by mean-field theory [19] deviates significantly from zero (Fig. 2(b), red line).

An extended critical phase with correlated multifractal modes.—The localization of dominant Jacobian eigenmodes v determines the spatial profile of network fluctuations and is described by the inverse participation ratio $\mathrm{IPR}_q(v) = \sum_i |v_i|^{2q}$, where $q > 0$ is a continuous index. The scaling of this quantity for large system size N is asymptotically described by $\mathrm{IPR}_q(v) \sim N^{(1-q)D_q}$, where $D_q = 0$ ($D_q = 1$) corresponds to localized (delocalized) states. Multifractal states are characterized by D_q being a nontrivial function of q, which is a hallmark of Anderson transitions [10]; the quantity D_q is consequently known as the multifractal dimension. Standard results in random matrix theory have established that the eigenvectors of random matrices with independent subexponential entries are delocalized [25], which includes the Jacobian matrices of classical homogeneous random neural networks ($\alpha = 2$). The chaotic fluctuations of classical random neural networks are thus delocalized, with the IPR (Fig. 2(c), blue) staying close to the value N^{1-q} attained for a constant, maximally delocalized vector $(1/\sqrt{N}, \ldots, 1/\sqrt{N})$, and the D_q estimate via $D_q \sim (\log_N \mathrm{IPR}_q)/(1-q)$ staying close to that of a random N-dimensional spherical vector, which is delocalized with overwhelming probability [26] (Fig. 2(d), blue and black); these two D_q estimates remain slightly curved due to finite-size effects.

Using our cavity approach, we find that all of the right eigenvectors of the heavy-tailed network Jacobian around the stationary state are multifractal for sigmoidal ϕ, a hallmark of Anderson transitions (see Supplemental Material [19] for derivation). The activity of heavy-tailed neural networks is thus dominated by multifractal chaotic fluctuations in contrast to the spatially delocalized chaos appearing in classical models. This theoretical prediction on heavy-tailed network dynamics is confirmed by simulations of temporal fluctuations of homogeneous and heavy-tailed neural activity (Figs. 2(c) and 2(d), red) which are both chaotic at $g = 1.75$. In Fig. 2(c) the IPR varies widely between the asymptotic large-N delocalized ($-\log_{10} N \approx -3$) and localized (0) values, while the classical network's activity fluctuations remain delocalized (blue). This result is remarkable from a physical standpoint as the activity itself is delocalized due to the bounded activation function $\phi = \tanh$, and does not visibly differ significantly from classical networks with Gaussian dynamics (see Supplemental Material [19]).

To investigate the behavior of the system's multifractal modes over long timescales, we quantify the extent to which Jacobian eigenvalues λ_i corresponding to a given eigenvector change their modulus relative to unity when new samples are chosen from the stationary distribution of neural activity. We thus consider the Jacobian average

$$\langle(|\lambda_i| - 1)^n\rangle_i = \int_{\mathbb{C}} (|z| - 1)^n \rho(z) dz \tag{5}$$

penalizing small and rewarding large eigenvalues to a degree determined by n which we call the annealing strength, and we find a region for the gain g adjacent to and above the ordered transition line, which is characterized by a greater proportion of eigenvalues away from zero compared to the ordered transition. This region for $n = 1, 3, \ldots, 9$ is found between the lower ordered transition line and the upper dashed line bearing the label n in Fig. 2(b). Quantifying the Jacobian spectral density via the Jacobian average allows us to distinguish between an active chaotic region in which unstable fluctuations tend to be quickly suppressed in favor of new fluctuations, and a region of temporally correlated chaotic fluctuations. The continuous transition between these active and correlated chaotic regimes is parametrized by the annealing strength n in Eq. (5). The theoretically predicted correlated region closes into the well-known critical point at the ordered-chaotic phase transition ($g = 1$) for classical rate-based networks ($\alpha = 2$), thus supporting the notion that the extended region of correlated multifractal modes is a critical regime. This extended critical phase is characterized by a significantly nonzero stationary state and a macroscopic proportion of unstable eigenstates relative to system dimensionality N, which is fundamentally different from the classical edge of chaos occurring around the zero fixed point when a microscopic proportion of eigenstates crosses the stability line. Consequently, this extended critical phase remains chaotic rather than existing solely at the edge of a chaotic phase as in classical networks [4].

In summary, the extended critical phase of temporally correlated, spatially multifractal fluctuations provides a demonstration of how various aspects of realistic neural dynamics may be exhibited simultaneously, such as long-range correlations [14] and low spatial dimensionality relative to system size [13]. This latter property arises from the localization of spatially multifractal fluctuations onto a small number of sites relative to system size (Fig. 1). Such behavior, along with the non-self-averaging properties characteristic of Anderson criticality [27] that we derived for local-field autocorrelation, suggests that the timescales across neurons in the heavy-tailed neural network are diverse. To validate this theoretical prediction on the extended critical region, we compute the relaxation timescales of neural autocorrelations over many realizations of random networks across heavy-tailed index α and network gain g. The extended Anderson critical regime emerging between the lower and upper dashed lines in Fig. 2(b) predicted by Eq. (5) (labeled by n) corresponds with high relaxation time variance (colored). We find that the extended critical phase is characterized by diverse timescales [Fig. 2(e)]: the relaxation timescale distribution is power law with index -2.0, consistent with that seen in cortical memory traces [15]. This behavior only exists in the critical regime (see Supplemental Material [19] for data analysis and quantitative fitting).

Persistent reservoir computing.—To explore the computational implications of the extended critical regime of correlated multifractal chaos, we consider a reservoir computing task described in [28] exploiting the chaotic dimensionality expansion of neural representations $\Delta H_{12} = H_{12}^{(s)} - H_{12}^{(n)}$ between signal ($H_{12}^{(s)}$) and noise ($H_{12}^{(n)}$) distances (see Supplemental Material [19] for setup details). Heavy-tailed networks make use of the temporal correlations of multifractal Jacobian eigenvectors above the stability line to enact a persistent form of real-time computation in a reservoir computing context. Because the extended critical regime is chaotic in heavy-tailed networks with a significantly nonzero fixed point, this regime is able to perform dimensionality expansion on its input in contrast to the classical edge of chaos which resides around the zero fixed point. At the same time, the correlated multifractal Jacobian eigenvectors work to hold off the onset of mixing (Fig. 3, yellow) to perform persistent chaotic dimensionality expansion on its input, allowing the computed result to remain in the system and thus the classification performance to stay above the baseline

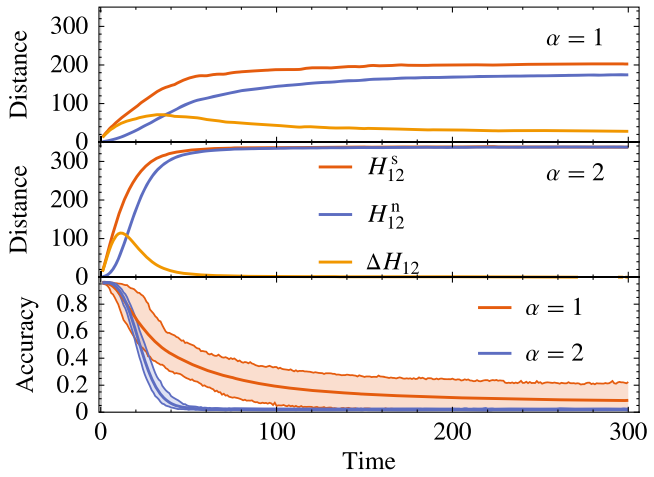

FIG. 3. Persistent dimensionality expansion in heavy-tailed random neural networks. Top, middle panels: Evolution of average signal and noise distances in a rate-based model with $N = 250$ and $g = 3$, as measured by the Hamming distance $H_{12} \coloneqq \|\phi(h_i^{(1)}) - \phi(h_i^{(2)})\|_i^2$ between states $\phi(h_i^{(1)})$ and $\phi(h_i^{(2)})$. Bottom panel: 5th–95th percentile classification accuracy of heavy-tailed (red) and homogeneous (blue) rate-based neural networks.

(0.01–0.03) for a longer period of time (Fig. 3, bottom). For example, in the parameters used in Fig. 3, the Gaussian network reaches chance level after around 70 network time constants while the heavy-tailed network remains significantly above chance level even after 3000 time constants (see Supplemental Material [19]). The extended critical regime of correlated multifractal chaos is thus able to produce efficient neural representations balancing the dimensional compression of stimuli (Fig. 3 top, red and blue), which is useful for generalization [29], and the separation of stimuli, in order to enact a form of persistent real-time computation.

Discussion.—Our theory rigorously demonstrates that heterogeneous, heavy-tailed connectivity can endow neural circuits with Anderson criticality over an extended parameter region, thus eliminating the fine-tuning needed in homogeneously connected neural networks. The Anderson criticality is characterized by correlated, low-dimensional fluctuations [13,14] and a diverse reservoir of timescales [15] as observed in biological neural systems; a reservoir of timescales can be produced by networks with spatially dependent coupling [30], plasticity mechanisms [31], or approximations to heterogeneous coupling using block-structured connectivity [4] and networks of neural clusters [32], but the ability for such networks to perform complex computations remains unexplored. Moreover, the extended Anderson criticality provides a unique mechanism for combining robust real-time computation with long-term memory of the computed output. Both homogeneous and heavy-tailed networks use chaos to enhance the separation of inputs for linear classification, but after a transient period of high separability, mixing dominates and erases the computed output from the homogeneous system. The recent observations on the ubiquity of heavy-tailed coupling in pretrained deep neural networks [8] suggest that our theory would be powerful for revealing the shared dynamical principles for persistent computation in both biological and artificial neural networks.

By making a novel link between Anderson criticality and the highly fluctuating complex dynamics of neural networks, our results suggest that complex systems operating over multiple scales should display a degree of multifractality at some level of fluctuations of system activity, even when the activity itself is bounded due to physical constraints. Multifractal phenomena have indeed been seen in a wide variety of natural systems such as turbulence [9]. In statistical and condensed matter physics, the Anderson transition appears with multifractality around the boundary of two phases characterized by localized and delocalized eigenvectors [10,33]. Our heavy-tailed neural network may thus be viewed as an inversion of these canonical models: a structurally extended Anderson regime $0 < \alpha < 2$ is bounded by delocalized Jacobian modes appearing in classical homogeneous networks with $\alpha = 2$, and localized eigenstates appearing in the limit $\alpha \to 0$ corresponding to sparse matrices and directed random graphs with a small average degree [34]. Hence, our random matrix theory could be applied to understand how complex dynamics emerge in physical systems.

This work was supported by the Australian Research Council (DP160104316).

*Corresponding author.
pulin.gong@sydney.edu.au

[1] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Critical phenomena in complex networks, Rev. Mod. Phys. **80**, 1275 (2008); H. E. Stanley and P. Meakin, Multifractal phenomena in physics and chemistry, Nature (London) **335**, 405 (1988); S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, and Y. Dodge, Turbulent cascades in foreign exchange markets, Nature (London) **381**, 767 (1996).

[2] D. R. Chialvo, Emergent complex neural dynamics, Nat. Phys. **6**, 744 (2010); 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*, Vol. 29, edited by D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett (Curran Associates, Inc., Red Hook, NY, 2016); J.-N. Teramae and T. Fukai, Local cortical circuit model inferred from power-law distributed neuronal avalanches, J. Comput. Neurosci. **22**, 301 (2007); Y. Gu, Y. Qi, and P. Gong, Rich-club connectivity, diverse population coupling, and dynamical activity patterns emerging from local cortical circuits, PLoS Comput. Biol. **15**, e1006902 (2019).

[3] H. Sompolinsky, A. Crisanti, and H. J. Sommers, Chaos in Random Neural Networks, Phys. Rev. Lett. **61**, 259 (1988); J. Kadmon and H. Sompolinsky, Transition to Chaos in Random Neuronal Networks, Phys. Rev. X **5**, 041030 (2015).

[4] J. Aljadeff, M. Stern, and T. Sharpee, Transition to Chaos in Random Networks with Cell-Type-Specific Connectivity, Phys. Rev. Lett. **114**, 088101 (2015); M. Stern, H. Sompolinsky, and L. F. Abbott, Dynamics of random neural networks with bistable units, Phys. Rev. E **90**, 062710 (2014); K. Rajan, L. F. Abbott, and H. Sompolinsky, Stimulus-dependent suppression of chaos in recurrent neural networks, Phys. Rev. E **82**, 011903 (2010); G. Wainrib and J. Touboul, Topological and Dynamical Complexity of Random Neural Networks, Phys. Rev. Lett. **110**, 118101 (2013).

[5] N. Bertschinger and T. Natschläger, Real-time computation at the edge of chaos in recurrent neural networks, Neural Comput. **16**, 1413 (2004).

[6] C.-T. Shih, Y.-J. Lin, C.-T. Wang, T.-Y. Wang, C.-C. Chen, T.-S. Su, C.-C. Lo, and A.-S. Chiang, Diverse community structures in the neuronal-level connectome of the drosophila brain, Neuroinformatics **18**, 267 (2020); L. K. Scheffer, C. S. Xu, M. Januszewski, Z. Lu, S.-y. Takemura, K. J. Hayworth, G. B. Huang, K. Shinomiya, J. Maitlin-Shepard, S. Berg *et al.*, A connectome and analysis of the adult Drosophila central brain, eLife **9**, e57443 (2020); G. Buzsáki and K. Mizuseki, The log-dynamic brain: How

skewed distributions affect network operations, Nat. Rev. Neurosci. **15**, 264 (2014).

[7] N. Perez-Nieves, V. C. H. Leung, P. L. Dragotti, and D. F. M. Goodman, Neural heterogeneity promotes robust learning, Nat. Commun. **12**, 5791 (2021).

[8] C. H. Martin, T. S. Peng, and M. W. Mahoney, Predicting trends in the quality of state-of-the-art neural networks without access to training or testing data, Nat. Commun. **12**, 4122 (2021); C. K. Qu, A. Wardak, and P. Gong, Extended critical regimes of deep neural networks, arXiv:2203.12967.

[9] R. Benzi, G. Paladin, G. Parisi, and A. Vulpiani, On the multifractal nature of fully developed turbulence and chaotic systems, J. Phys. A **17**, 3521 (1984).

[10] F. Evers and A. D. Mirlin, Anderson transitions, Rev. Mod. Phys. **80**, 1355 (2008).

[11] P. W. Anderson, Absence of diffusion in certain random lattices, Phys. Rev. **109**, 1492 (1958).

[12] R. Abou-Chacra, D. J. Thouless, and P. W. Anderson, A selfconsistent theory of localization, J. Phys. C **6**, 1734 (1973); G. Parisi, S. Pascazio, F. Pietracaprina, V. Ros, and A. Scardicchio, Anderson transition on the Bethe lattice: an approach with real energies, J. Phys. A **53**, 014003 (2020).

[13] F. Mastrogiuseppe and S. Ostojic, Linking connectivity, dynamics, and computations in low-rank recurrent neural networks, Neuron **99**, 609 (2018); C. Huang, D. A. Ruff, R. Pyle, R. Rosenbaum, M. R. Cohen, and B. Doiron, Circuit models of low-dimensional shared variability in cortical networks, Neuron **101**, 337 (2019).

[14] B. J. He, Scale-free brain activity: Past, present, and future, Trends Cognit. Sci. **18**, 480 (2014).

[15] A. Bernacchia, H. Seo, D. Lee, and X.-J. Wang, A reservoir of time constants for memory traces in cortical neurons, Nat. Neurosci. **14**, 366 (2011).

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

[17] P. Cizeau and J. P. Bouchaud, Mean field theory of dilute spin-glasses with power-law interactions, J. Phys. A **26**, L187 (1993).

[18] L. 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); A. Wardak and P. Gong, Fractional diffusion theory of balanced heterogeneous neural networks, Phys. Rev. Research **3**, 013083 (2021).

[19] See Supplemental Material, which includes Ref. [20], at http://link.aps.org/supplemental/10.1103/PhysRevLett.129.048103 for detailed mean-field and random matrix derivations, along with additional data.

[20] C. Bordenave and D. Chafaï, Around the circular law, Probab. Surv. **9**, 1 (2012); F. L. Metz, I. Neri, and D. Bollé, Localization transition in symmetric random matrices, Phys. Rev. E **82**, 031135 (2010); T. Rogers and I. P. Castillo, Cavity approach to the spectral density of non-hermitian sparse matrices, Phys. Rev. E **79**, 012101 (2009); I. Neri and F. L. Metz, Spectra of Sparse Non-Hermitian Random Matrices: An Analytical Solution, Phys. Rev. Lett. **109**, 030602 (2012); Y. Ahmadian, F. Fumarola, and K. D. Miller, Properties of networks with partially structured and partially random connectivity, Phys. Rev. E **91**, 012820 (2015); D. Schertzer, M. Larchevêque, J. Duan, V. V. Yanovsky, and S. Lovejoy, Fractional Fokker-Planck equation for nonlinear stochastic differential equations driven by non-Gaussian Lévy stable noises, J. Math. Phys. (N.Y.) **42**, 200 (2001); A. Wardak, First passage Leapovers of Lévy flights and the proper formulation of absorbing boundary conditions, J. Phys. A **53**, 375001 (2020).

[21] C. Bordenave, P. Caputo, and D. Chafaï, Spectrum of non-hermitian heavy tailed random matrices, Commun. Math. Phys. **307**, 513 (2011).

[22] F. L. Metz, I. Neri, and T. Rogers, Spectral theory of sparse non-hermitian random matrices, J. Phys. A **52**, 434003 (2019).

[23] K. Rajan and L. F. Abbott, Eigenvalue Spectra of Random Matrices for Neural Networks, Phys. Rev. Lett. **97**, 188104 (2006); Y. Wei, Eigenvalue spectra of asymmetric random matrices for multicomponent neural networks, Phys. Rev. E **85**, 066116 (2012).

[24] For clarity, we refer to $gJ\mathrm{diag}_j[\phi'(h_j)]$ as the Jacobian matrix in the remainder of the Letter as its eigenvectors remain unchanged and its eigenvalues are shifted 1 unit to the right.

[25] M. Rudelson and R. Vershynin, Delocalization of eigenvectors of random matrices with independent entries, Duke Math. J. **164**, 2507 (2015).

[26] S. O'Rourke, V. Vu, and K. Wang, Eigenvectors of random matrices: A survey, J. Comb. Theory Ser. A **144**, 361 (2016).

[27] A. Solórzano, L. F. Santos, and E. J. Torres-Herrera, Multifractality and self-averaging at the many-body localization transition, Phys. Rev. Research **3**, L032030 (2021).

[28] C. Keup, T. Kühn, D. Dahmen, and M. Helias, Transient Chaotic Dimensionality Expansion by Recurrent Networks, Phys. Rev. X **11**, 021064 (2021).

[29] C. Stringer, M. Pachitariu, N. Steinmetz, M. Carandini, and K. D. Harris, High-dimensional geometry of population responses in visual cortex, Nature (London) **571**, 361 (2019).

[30] A. Miri, K. Daie, A. B. Arrenberg, H. Baier, E. Aksay, and D. W. Tank, Spatial gradients and multidimensional dynamics in a neural integrator circuit, Nat. Neurosci. **14**, 1150 (2011); M. Joshua, J. F. Medina, and S. G. Lisberger, Diversity of neural responses in the brainstem during smooth pursuit eye movements constrains the circuit mechanisms of neural integration, J. Neurosci. **33**, 6633 (2013); R. Chaudhuri, A. Bernacchia, and X.-J. Wang, A diversity of localized timescales in network activity, eLife **3**, e01239 (2014).

[31] M. O. Magnasco, O. Piro, and G. A. Cecchi, Self-Tuned Critical Anti-Hebbian Networks, Phys. Rev. Lett. **102**, 258102 (2009).

[32] M. Stern, N. Istrate, and L. Mazzucato, A reservoir of timescales in random neural networks, arXiv:2110.09165.

[33] P. Cizeau and J. P. Bouchaud, Theory of Lévy matrices, Phys. Rev. E **50**, 1810 (1994); E. Tarquini, G. Biroli, and M. Tarzia, Level Statistics and Localization Transitions of Lévy Matrices, Phys. Rev. Lett. **116**, 010601 (2016).

[34] F. L. Metz and I. Neri, Localization and Universality of Eigenvectors in Directed Random Graphs, Phys. Rev. Lett. **126**, 040604 (2021).