Slow Transition to Low-Dimensional Chaos in Heavy-Tailed Recurrent Neural Networks

Eva Yi Xie EVAYIXIE@PRINCETON.EDU
Princeton Neuroscience Institute; Allen Institute

Stefan Mihalas STEFANM@ALLENINSTITUTE.ORG
Łukasz Kuśmierz LUKASZ.KUSMIERZ@ALLENINSTITUTE.ORG
Allen Institute

Abstract

Growing evidence suggests that synaptic weights in the brain follow heavy-tailed distributions, yet most theoretical analyses of recurrent neural networks (RNNs) assume Gaussian connectivity. We systematically study the activity of RNNs with random weights drawn from biologically plausible Lévy alpha-stable distributions. While mean-field theory for the infinite system predicts that the quiescent state is always unstable—implying ubiquitous chaos—our finite-size analysis reveals a sharp transition between quiescent and chaotic dynamics. We theoretically predict the gain at which the finite system transitions from quiescent to chaotic dynamics, and validate it through simulations. Compared to Gaussian networks, finite heavy-tailed RNNs exhibit a broader gain regime near the edge of chaos, namely, a slow transition to chaos. However, this robustness comes with a tradeoff: heavier tails reduce the Lyapunov dimension of the attractor, indicating lower effective dimensionality. Our results reveal a biologically aligned tradeoff between the robustness of dynamics near the edge of chaos and the richness of high-dimensional neural activity. By analytically characterizing the transition point in finite-size networks—where mean-field theory breaks down—we provide a tractable framework for understanding dynamics in realistically sized, heavy-tailed neural circuits.*

Keywords: theoretical neuroscience; recurrent neural networks; heavy-tailed connectivity; finite-size effects; Lyapunov exponents

1. Introduction

Advances in connectomics are revealing increasingly detailed wiring diagrams across species and brain regions (Dorkenwald et al., 2024; The MICrONS Consortium, 2025), prompting questions about the structural principles that govern neural computation. A striking and ubiquitous feature is the heavy-tailed distribution of synaptic weights, observed in mammalian cortex, mammalian hippocampus, and even *Drosophila* (Song et al., 2005; Lefort et al., 2009; Ikegaya et al., 2013; Scheffer et al., 2020), yet absent from most theoretical neuroscience models (e.g., Kadmon and Sompolinsky, 2015), which typically assume Gaussian connectivity, and from common machine learning initialization schemes (Glorot and Bengio, 2010; He et al., 2015). Such heavy tails can be modeled using Lévy α-stable distributions (Feller, 1971; Borak et al., 2005), parameterized by a stability index α where smaller α produces more extreme outliers, potentially altering network dynamics. In recurrent neural networks (RNNs), a central phenomenon is the transition between quiescent

*The codebase is publicly available at https://github.com/AllenInstitute/HeavyRNN_public.

and chaotic regimes—the "edge of chaos"—long hypothesized to optimize information flow (Bertschinger et al., 2004; Schuecker et al., 2018). While Gaussian random matrices exhibit a well-defined transition determined by eigenvalue spectra (Rajan and Abbott, 2006; Aljadeff et al., 2015), heavy-tailed matrices feature unbounded eigenvalue densities (Bordenave et al., 2011), which in the infinite-size limit can abolish the transition entirely (Kuśmierz et al., 2020). Understanding how such connectivity shapes dynamics is timely, as RNNs are increasingly used to model neural population activity (Sussillo and Abbott, 2009; Pandarinath et al., 2018), simulate cognitive tasks (Yang et al., 2019; Driscoll et al., 2024), and generate hypotheses for experiments (Pinto et al., 2019; Pagan et al., 2025), yet theoretical insight into RNNs with biologically realistic heavy-tailed weights remains scarce.

Related Work Prior theoretical work has shown that infinite-width feedforward networks with Gaussian weights converge to Gaussian processes (Neal, 1996), whereas Lévy α-stable weights yield richer α-stable process priors (Favaro et al., 2020; Jung et al., 2021; Bordino et al., 2023; Favaro et al., 2023). While these studies focus on feedforward architectures, our annealed analysis reveals a distinct transition in heavy-tailed feedforward networks and extends the investigation to recurrent networks. Heavy-tailed connectivity has been linked to extended critical regimes in both RNNs and feedforward nets (Wardak and Gong, 2022; Qu et al., 2022); our work builds on this by explaining the phenomenon via maximal Lyapunov exponents, showing size-dependent transition locations, and identifying a tradeoff between regime size and neural manifold dimensionality. For the `tanh` activation function studied here, Cauchy RNNs ($\alpha = 1$) were found by Kuśmierz et al. (2020) to exhibit perpetual chaos in the infinite-size limit, but our results demonstrate that finite networks can still exhibit well-defined transitions, underscoring the importance of finite-size effects.

2. Methods

We study discrete-time RNNs with tanh activation, $x_i(t+1) = \tanh\big(\sum_j W_{ij}x_j(t) + I_i(t)\big)$, where $W_{ij} \overset{\text{i.i.d.}}{\sim} L_\alpha(\sigma)$ is drawn from a symmetric Lévy α-stable distribution with scale $\sigma = g/N^{1/\alpha}$, gain g, and $\alpha \in (0, 2]$ ($\alpha = 2$ is Gaussian). We consider autonomous ($I = 0$), stimulus-driven, and *annealed* settings, the latter resampling W each step (equivalent to a depth-T feedforward net).

Dynamical stability is quantified by the maximum Lyapunov exponent (MLE) $\lambda_{\max}$ via a QR-based algorithm (Vogt et al., 2022), with $\lambda_{\max} > 0$ indicating chaos, < 0 stability, and ≈ 0 the edge of chaos. Attractor dimensionality is measured by the Lyapunov (Kaplan–Yorke) dimension $D_{\mathrm{KY}} = k + \frac{\sum_{i=1}^{k}\lambda_i}{|\lambda_{k+1}|}$, where $\{\lambda_i\}$ are ordered exponents and k is the largest index with $\sum_{i=1}^{k}\lambda_i \geq 0$, and by the participation ratio $\mathrm{PR} = (\sum_i \tilde{\lambda}_i)^2/\sum_i \tilde{\lambda}_i^2$ from the eigenvalues $\{\lambda_i\}$ of the steady-state covariance. PR reflects linear dimensionality, D_{KY} nonlinear complexity; their comparison shows how heavy-tailed connectivity shapes stability and manifold geometry.

3. Results

3.1. Finite heavy-tailed networks exhibit a predictable quiescent-to-chaotic transition

We asked whether finite-size recurrent networks with Lévy α-stable weights retain a well-defined quiescent–chaotic transition, absent in the infinite-width limit for $\alpha < 2$ (Kuśmierz et al., 2020). An annealed analysis of the linearized dynamics (Appendix A) yields $g^* = \exp(-\langle \Xi_{N,\alpha} \rangle)$, $\Xi_{N,\alpha} = \frac{1}{\alpha} \ln\left(\frac{1}{N} \sum_{j=1}^{N} |z_j|^\alpha\right)$, $z_j \sim L_\alpha(1)$. The theory predicts g^* decreases with N for all α, approaching the mean-field limit quickly for $\alpha = 2$ but only logarithmically for $\alpha < 2$ (Fig. 1A). Simulations of annealed and quenched linear networks match these predictions: in the quenched case, g^* varies across realizations but centers near the annealed value, with larger fluctuations at smaller N (Fig. 1B; Appendix B).

3.2. Heavier-tailed RNNs exhibit a slower, more robust transition to chaos

We next examine how the network transition depends on the tail index α. Simulations of autonomous RNNs (Fig. 1C; noisy-input results in Appendix) show that for $\alpha \geq 1$, $\lambda_{\max}$ increases from negative to positive values as g grows, verifying a transition to chaos numerically. The shape and position of this transition vary systematically with α: in the Gaussian case ($\alpha = 2$), $\lambda_{\max}$ rises steeply, while for heavier-tailed networks (smaller α) the growth is more gradual, keeping the network close to the edge of chaos over a broader range of g. This extended critical-like regime may offer robustness benefits to parameter drift, enabling high-capacity dynamics without precise tuning (Bertschinger et al., 2004; Legenstein and Maass, 2007; Toyoizumi and Abbott, 2011). Increasing N shifts the transition to lower g (Appendix C), with heavier-tailed networks showing a stronger shift, consistent with our finite-size scaling theory (Fig. 1A).

3.3. Heavy-tailed RNNs compress the chaotic attractor into a lower-dimensional slow manifold

The robustness of the transition to chaos in heavy-tailed RNNs raises a natural question: does the attractor geometry also vary systematically with the tail index α? At the average critical gain $\langle g^* \rangle$, Lyapunov spectra show that Gaussian networks have many exponents near zero, indicating a broad slow manifold, whereas heavier-tailed networks have fewer (Fig. 2A). Accordingly, the Lyapunov dimension D_{KY} drops sharply with smaller α (Fig. 2B), and the participation ratio, a variance-based measure of activity dimensionality, also declines (Fig. 2C), though less steeply. This tradeoff persists under stimulus-driven dynamics (Appendix D): heavier-tailed networks gain robustness near the edge of chaos but compress activity into a lower-dimensional manifold.

4. Discussion

Our results reveal a finite-size phenomenon absent from mean-field theory: recurrent networks with Lévy-distributed synaptic weights exhibit a clear quiescent-to-chaotic transition whose location and sharpness depend systematically on both the tail index α and network

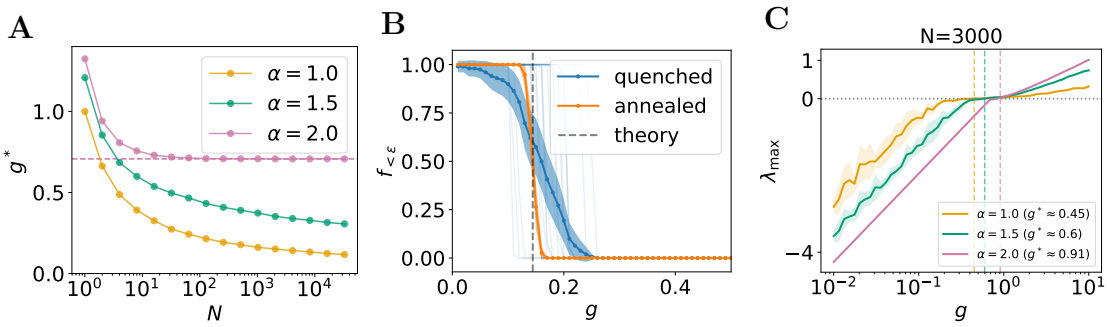

Figure 1: **Finite heavy-tailed networks exhibit a predictable and tail-dependent *slow* transition to chaos.** (**A**) Theoretical prediction of the critical gain g^* as a function of network size N for different tail indices α. In the Gaussian case ($\alpha = 2$), g^* rapidly converges to the mean-field limit (dashed line), whereas heavy-tailed networks ($\alpha < 2$) show a much slower, logarithmic decay toward zero. (**B**) Fraction of near-zero state components $f_{<\epsilon}$ ($\epsilon = 0.1$) in linearized networks with $\alpha = 1$ and $N = 3000$ evolved for $T = 100$ steps. Annealed networks (orange) display a sharp transition at the predicted g^*, while quenched networks (blue) show realization-dependent shifts, which average to a smoother curve but remain centered near the theory. (**C**) Maximum Lyapunov exponent $\lambda_{\max}$ for $N = 3000$ as a function of g and α, showing that heavier-tailed networks transition more gradually and remain near the edge of chaos over a broader range of g.

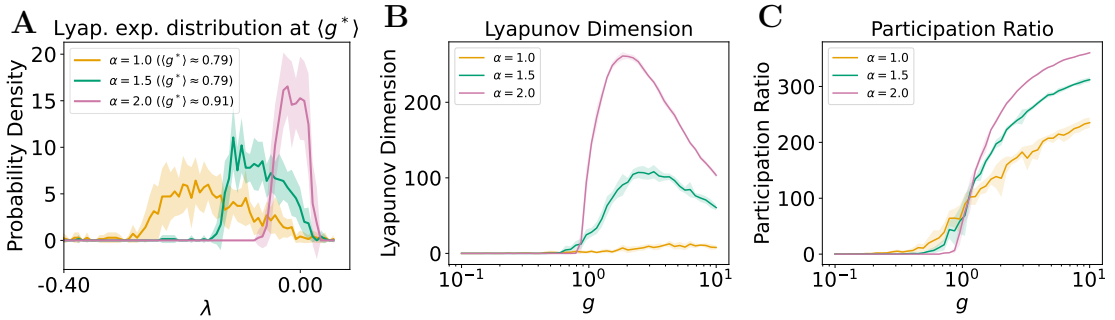

Figure 2: **Heavy-tailed RNNs operate on lower-dimensional attractors near the edge of chaos.** All panels: $N = 1000$, mean ± 1 SD across 10 trials; see Appendices for noisy-input and ablation results. (**A**) Lyapunov spectra at the average critical gain $\langle g^* \rangle$ reveal that heavier-tailed networks have fewer exponents near zero, indicating a compressed slow manifold. (**B**) Lyapunov dimension is lower for smaller α, reflecting fewer locally expanding or marginally stable directions. (**C**) Participation ratio also decreases with smaller α, indicating reduced variance spread across activity modes.

size N. Heavier-tailed connectivity widens the parameter range over which networks remain near the edge of chaos, conferring robustness to gain variations—a property that may be advantageous for biological circuits operating across multiple states or in variable environments. However, this robustness comes at a cost: both the Lyapunov dimension and participation ratio decrease with heavier tails, indicating that dynamics are compressed into a lower-dimensional slow manifold. This tradeoff implies that, for tasks requiring high-dimensional activity, heavier-tailed networks may need to be larger than Gaussian ones, whereas tasks relying on low-dimensional dynamics may benefit from the extended critical-like regime. These findings highlight the importance of finite-size corrections for interpreting neural activity and suggest that anatomical connectivity statistics can directly inform predictions about circuit-level computational capacity. Future work should investigate how these dynamical properties translate to information-processing performance, and whether learning, task structure, and biological constraints modulate or exploit the robustness-dimensionality tradeoff identified here.

References

Johnatan Aljadeff, Merav Stern, and Tatyana Sharpee. Transition to chaos in random networks with cell-type-specific connectivity. *Physical Review Letters*, 114(8):088101, 2015.

Nils Bertschinger, Thomas Natschläger, and Robert Legenstein. At the edge of chaos: Real-time computations and self-organized criticality in recurrent neural networks. In L. Saul, Y. Weiss, and L. Bottou, editors, *Advances in Neural Information Processing Systems*, volume 17. MIT Press, 2004.

Szymon Borak, Wolfgang Härdle, and Rafal Weron. *Stable distributions*. Springer, 2005.

Charles Bordenave, Pietro Caputo, and Djalil Chafaï. Spectrum of non-hermitian heavy tailed random matrices. *Communications in mathematical physics*, 307:513–560, 2011.

Alberto Bordino, Stefano Favaro, and Sandra Fortini. Infinitely wide limits for deep stable neural networks: sub-linear, linear and super-linear activation functions. *arXiv preprint arXiv:2304.04008*, 2023.

Sven Dorkenwald, Arie Matsliah, Amy R Sterling, Philipp Schlegel, Szi-Chieh Yu, Claire E McKellar, Albert Lin, Marta Costa, Katharina Eichler, Yijie Yin, et al. Neuronal wiring diagram of an adult brain. *Nature*, 634(8032):124–138, 2024.

Laura N Driscoll, Krishna Shenoy, and David Sussillo. Flexible multitask computation in recurrent networks utilizes shared dynamical motifs. *Nature Neuroscience*, 27(7):1349–1363, 2024.

Stefano Favaro, Sandra Fortini, and Stefano Peluchetti. Stable behaviour of infinitely wide deep neural networks. In *International Conference on Artificial Intelligence and Statistics*, pages 1137–1146. PMLR, 2020.

Stefano Favaro, Sandra Fortini, and Stefano Peluchetti. Deep stable neural networks: large-width asymptotics and convergence rates. *Bernoulli*, 29(3):2574–2597, 2023.

William Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971.

Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256. JMLR Workshop and Conference Proceedings, 2010.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pages 1026–1034, 2015.

Yuji Ikegaya, Takuya Sasaki, Daisuke Ishikawa, Naoko Honma, Kentaro Tao, Naoya Takahashi, Genki Minamisawa, Sakiko Ujita, and Norio Matsuki. Interpyramid spike transmission stabilizes the sparseness of recurrent network activity. *Cerebral Cortex*, 23(2): 293–304, 2013.

Paul Jung, Hoil Lee, Jiho Lee, and Hongseok Yang. *alpha*-stable convergence of heavy-tailed infinitely-wide neural networks. *arXiv preprint arXiv:2106.11064*, 2021.

Jonathan Kadmon and Haim Sompolinsky. Transition to chaos in random neuronal networks. *Physical Review X*, 5(4):041030, 2015.

Harry Kesten. Random difference equations and renewal theory for products of random matrices. 1973.

Łukasz Kuśmierz, Shun Ogawa, and Taro Toyoizumi. Edge of chaos and avalanches in neural networks with heavy-tailed synaptic weight distribution. *Physical Review Letters*, 125(2):028101, 2020.

Sandrine Lefort, Christian Tomm, J-C Floyd Sarria, and Carl CH Petersen. The excitatory neuronal network of the c2 barrel column in mouse primary somatosensory cortex. *Neuron*, 61(2):301–316, 2009.

Robert Legenstein and Wolfgang Maass. Edge of chaos and prediction of computational performance for neural circuit models. *Neural networks*, 20(3):323–334, 2007.

Lutz Molgedey, J Schuchhardt, and Heinz G Schuster. Suppressing chaos in neural networks by noise. *Physical review letters*, 69(26):3717, 1992.

Radford M. Neal. *Bayesian Learning for Neural Networks*. Springer-Verlag, Berlin, Heidelberg, 1996. ISBN 0387947248.

Marino Pagan, Vincent D Tang, Mikio C Aoi, Jonathan W Pillow, Valerio Mante, David Sussillo, and Carlos D Brody. Individual variability of neural computations underlying flexible decisions. *Nature*, 639(8054):421–429, 2025.

Chethan Pandarinath, Daniel J O'Shea, Jasmine Collins, Rafal Jozefowicz, Sergey D Stavisky, Jonathan C Kao, Eric M Trautmann, Matthew T Kaufman, Stephen I Ryu, Leigh R Hochberg, et al. Inferring single-trial neural population dynamics using sequential auto-encoders. *Nature methods*, 15(10):805–815, 2018.

Lucas Pinto, Kanaka Rajan, Brian DePasquale, Stephan Y Thiberge, David W Tank, and Carlos D Brody. Task-dependent changes in the large-scale dynamics and necessity of cortical regions. *Neuron*, 104(4):810–824, 2019.

Cheng Kevin Qu, Asem Wardak, and Pulin Gong. Extended critical regimes of deep neural networks. *arXiv preprint arXiv:2203.12967*, 2022.

Kanaka Rajan and Larry F Abbott. Eigenvalue spectra of random matrices for neural networks. *Physical review letters*, 97(18):188104, 2006.

Louis K Scheffer, C Shan Xu, Michal Januszewski, Zhiyuan Lu, Shin-ya Takemura, Kenneth J Hayworth, Gary B Huang, Kazunori Shinomiya, Jeremy Maitlin-Shepard, Stuart Berg, et al. A connectome and analysis of the adult drosophila central brain. *elife*, 9: e57443, 2020.

Jannis Schuecker, Sven Goedeke, and Moritz Helias. Optimal sequence memory in driven random networks. *Physical Review X*, 8(4):041029, 2018.

Sen Song, Per Jesper Sjöström, Markus Reigl, Sacha Nelson, and Dmitri B Chklovskii. Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS biology*, 3(3):e68, 2005.

Adiel Statman, Maya Kaufman, Amir Minerbi, Noam E Ziv, and Naama Brenner. Synaptic size dynamics as an effectively stochastic process. *PLoS computational biology*, 10(10): e1003846, 2014.

David Sussillo and Larry F Abbott. Generating coherent patterns of activity from chaotic neural networks. *Neuron*, 63(4):544–557, 2009.

The MICrONS Consortium. Functional connectomics spanning multiple areas of mouse visual cortex. *Nature*, 640(8058):435–447, 2025.

Taro Toyoizumi and Larry F Abbott. Beyond the edge of chaos: Amplification and temporal integration by recurrent networks in the chaotic regime. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 84(5):051908, 2011.

Ryan Vogt, Maximilian Puelma Touzel, Eli Shlizerman, and Guillaume Lajoie. On lyapunov exponents for rnns: Understanding information propagation using dynamical systems tools. *Frontiers in Applied Mathematics and Statistics*, 8:818799, 2022.

Asem Wardak and Pulin Gong. Extended anderson criticality in heavy-tailed neural networks. *Phys. Rev. Lett.*, 129:048103, Jul 2022.

Guangyu Robert Yang, Madhura R Joglekar, H Francis Song, William T Newsome, and Xiao-Jing Wang. Task representations in neural networks trained to perform many cognitive tasks. *Nature neuroscience*, 22(2):297–306, 2019.

Appendix A. Mathematical analysis of the transition in annealed networks

Our goal is to show that networks with α-stable weight distributions exhibit a transition between two regimes, and to find the location of this transition, which we denote as g^*. As in Gaussian networks, the quiescent state is stable and any small perturbation around it shrinks if weights are generated from a narrow enough distribution (i.e., $g < g^*$). Similarly, the quiescent state is unstable if the underlying distribution is wide enough ($g > g^*$). In contrast to Gaussian networks, however, this effect can only be observed through the analysis of finite-size effects.

As described in the main text, we study the linear stability of the quiescent fixed point by analyzing the dynamics of small perturbations around it. Since weights are randomly redrawn at each step, the evolution $\varepsilon^{(t)}$ is a stochastic process. To quantify its behavior we focus our attention on the conditional distribution $\varepsilon^{(t+1)}$ given $\varepsilon^{(t)}$. Components of this vector are independent due to the assumed independence of rows of the weight matrix. The conditional distribution of a single component can be characterized in the Fourier space as

$$\left\langle \exp\left(ik\varepsilon_i^{(t+1)}|\varepsilon^{(t)}\right)\right\rangle_W = \left\langle \exp\left(ik\sum_{j=1}^{N_t} W_{ij}^{(t)}\varepsilon_j^{(t)}\right)\right\rangle_{W_{ij}^{(t)}} = \exp\left(-|k|^\alpha g^\alpha \frac{1}{N_l}\sum_{j=1}^{N_t}\left|\varepsilon_j^{(t)}\right|^\alpha\right) \tag{1}$$

where we used $W_{ij}^{(t)} \sim L_\alpha\left(g/N_t^{1/\alpha}\right)$. Thus, for $t > 1$, the perturbation, when conditioned on the previous step, is an α-stable random variable. More specifically, it can be written as $\varepsilon_i^{(t+1)}|\varepsilon^{(t)} \sim L_\alpha\left(\gamma^{(t+1)}\right)$, where the conditional scale at step $t+1$

$$\gamma^{(t+1)} = g\left(\frac{1}{N_t}\sum_{j=1}^{N_t}\left|\varepsilon_j^{(t)}\right|^\alpha\right)^{1/\alpha} \tag{2}$$

is a deterministic function of state at time t, which itself is a random variable. We can unpack this relation one step backwards by conditioning on $\varepsilon^{(t-1)}$ instead, with $\varepsilon_i^{(t)}|\varepsilon^{(t-1)} \sim L_\alpha\left(\gamma^{(t)}\right)$. We utilize the fact that this can also be expressed as

$$\varepsilon_i^{(t)}|\varepsilon^{(t-1)} = \gamma^{(t)} z_i^{(t)} \tag{3}$$

where $\gamma^{(t)}$ depends on the perturbation at time $t-1$, and $z_i^{(t)}$ are i.i.d. α-stable variables. This leads to the recursive formula for scalar $\gamma^{(t)}$

$$\gamma^{(t+1)} = \gamma^{(t)}\xi^{(t)} \tag{4}$$

where $\left(\xi^{(t)}\right)_{t=1}^{\infty}$ is a sequence of independent random variables distributed as

$$\xi^{(t)} = g\left(\frac{1}{N_t}\sum_{j=1}^{N_t}\left|z_j^{(t)}\right|^\alpha\right)^{1/\alpha} \tag{5}$$

with i.i.d. $z_j^{(l)} \sim L_\alpha(1)$. If layers have the same width $N_t = N$, $\xi^{(t)}$ are i.i.d. and (4) is a scalar multiplicative process with i.i.d. entries. Thus, we have reduced our problem to a

simpler special case of purely multiplicative scalar Kesten process. We can easily solve this recursion and rewrite the solution as a sum

$$\ln \gamma^{(t+1)} = \ln \gamma^{(t)} + \sum_{i=1}^{t} \ln \xi^{(i)} \tag{6}$$

where $\gamma^{(1)}$ is deterministically specified by the input perturbation $\varepsilon^{(0)}$. It is known (Kesten, 1973; Statman et al., 2014) that this sum diverges to $-\infty$ almost surely if $\langle \ln \xi \rangle < 0$ and diverges to ∞ almost surely if $\langle \ln \xi \rangle > 0$. Accordingly, the sequence $\left(\gamma^{(t)}\right)_{t=1}^{\infty}$ either converges to 0 or diverges. Therefore, the critical width of the synaptic weight distribution is given by

$$g^* = \exp\left(-\langle \Xi_{N,\alpha} \rangle\right) \tag{7}$$

where

$$\Xi_{N,\alpha} = \frac{1}{\alpha} \ln \left(\frac{1}{N} \sum_{j=1}^{N} |z_j|^{\alpha} \right) \tag{8}$$

with $z_j \sim L_\alpha(1)$.

Appendix B. Derivation of the logarithmic decay of $g^*(N)$

Here, we estimate the expected value of

$$\Xi_{N,\alpha} = \frac{1}{\alpha} \ln \left(\frac{1}{N} \sum_{j=1}^{N} |z_j|^{\alpha} \right), \tag{9}$$

where $z_j \sim L_\alpha(1)$, for large N. We define $Y_{N,\alpha} = \frac{1}{N} \sum_{j=1}^{N} |z_j|^{\alpha}$ and note that the Laplace transform of $Y_{N,\alpha}$ can be calculated as

$$F_{N,\alpha}(s) = \left\langle e^{-sY_{N,\alpha}} \right\rangle = \left(F_{1,\alpha} \left(\frac{s}{N} \right) \right)^N \tag{10}$$

where

$$F_{1,\alpha}(s) = \left\langle e^{-s|z|^{\alpha}} \right\rangle_{z \sim L_\alpha(1)} \tag{11}$$

According to (10), the large N asymptotic of $\Xi_{N,\alpha}$ is dominated by the behavior of $F_{1,\alpha}(s)$ around $s = 0$. This behavior should be similar for all symmetric distributions with the same stability index. For example, take $\rho_z(x) = \frac{\alpha}{2}|x|^{-1-\alpha}$ for $|x| > 1$ and $\rho_z(x) = 0$ otherwise. The resulting expansion can be found as

$$\left\langle e^{-s|x|^{\alpha}} \right\rangle_{z \sim \rho_z} = s \int_s^{\infty} du u^{-2} e^{-u} = s\Gamma(-1, s) \approx 1 - s\left(1 - \gamma - \ln s\right) + O(s^2) \tag{12}$$

where $\Gamma(a, s)$ is the upper incomplete gamma function and γ is the Euler-Mascheroni constant. Thus, the asymptotic expansion of $F_{1,\alpha}(s)$ must take the form

$$F_{1,\alpha}(s) = 1 - A_\alpha s \left(B_\alpha - \ln s\right) + O(s^2) \tag{13}$$

for some irrelevant constants A_α, B_α. We plug (13) into (10) and arrive at

$$\ln F_{N,\alpha}(s) = -A_\alpha s \left(B_\alpha - \ln s + \ln N\right) + O\left(N^{-1}\right) \tag{14}$$

For $N \gg 1$, (14) corresponds to a random variable X_N that can be constructed as

$$X_N = X_1 + A_\alpha \ln N, \tag{15}$$

where

$$\langle \exp(-sX_1) \rangle = \exp(-A_\alpha s \left(B_\alpha - \ln s\right)) \tag{16}$$

We can rewrite the desired expected value as

$$\langle \Xi_{N,\alpha} \rangle \approx \frac{1}{\alpha} \left\langle \ln \left(X_1 + A_\alpha \ln N \right) \right\rangle_{X_1} \tag{17}$$

The distribution of X_1 is fixed and does not change with N. Thus, for large N the second term dominates, and we arrive at

$$g^* = \exp(-\langle \Xi_{N,\alpha} \rangle) \asymp \frac{1}{(\ln N)^{1/\alpha}} \tag{18}$$

Appendix C. Finite-size dependence of autonomous RNN dynamics

We next examine how network size N modulates the quiescent-to-chaotic transition in autonomous Lévy RNNs. Figure 3 shows that increasing N systematically shifts the critical gain g^* leftward, with the effect most pronounced for heavier-tailed connectivity ($\alpha < 2$). This finite-size dependence, absent in mean-field predictions, reinforces our theoretical result that g^* varies with both N and α.

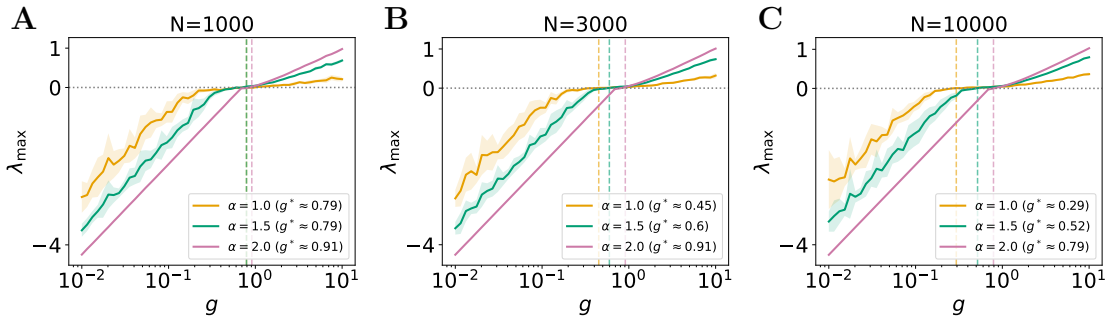

Figure 3: **Maximum Lyapunov exponent ($\lambda_{\max}$) as a function of gain g for autonomous recurrent networks with different tail indices α, shown for: (A)** $N = 1000$, **(B)** $N = 3000$, and **(C)** $N = 10000$. Curves show mean across 10 trials; shaded regions denote ± 1 SD. We let the networks evolve for $T = 3000$ steps, among which the Lyapunov exponents are accumulated over the last $K = 100$ steps. See results under noisy stimulus in Appendices D. Heavier-tailed networks (lower α) exhibit a slower, more gradual increase in $\lambda_{\max}$ near the transition (where $\lambda_{\max} = 0$), resulting in a broader edge-of-chaos regime with respect to g. Dashed lines and legend mark the average critical gain g^* at which $\lambda_{\max}$ first crosses zero. As N increases, this transition shifts leftward, especially for lower α, in line with our theoretical predictions on finite-size effects.

Appendix D. Additional results under small noisy input

We replicate our main results under i.i.d. Gaussian noise (variance = 0.01) added at each time step, testing robustness in more biologically realistic, stimulus-driven settings. As expected (Molgedey et al., 1992), noise quenches chaos but preserves the autonomous-case trends.

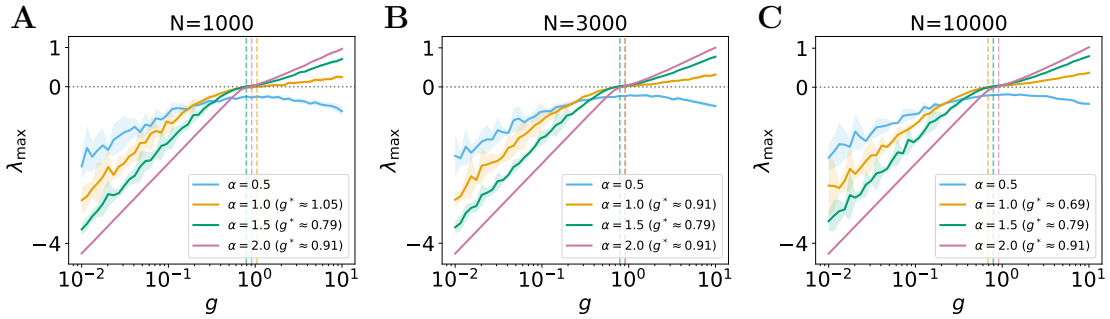

Figure 4: **Effect of network size under small i.i.d. noisy input.** Maximum Lyapunov exponent ($\lambda_{\max}$) as a function of gain g in noisy stimulus-driven recurrent networks with Lévy α-stable weight distributions. Curves show mean across 10 trials; shaded regions denote ± 1 SD. Each panel corresponds to a different network size: (A) $N = 1000$, (B) $N = 3000$, and (C) $N = 10000$. Curves show mean across 3 trials; shaded regions denote ± 1 SD. As in the autonomous case, if a transition exists, then heavier-tailed networks exhibit a slower transition and wider critical regime near $\lambda_{\max} = 0$. The critical gain g^* (dashed line) shifts leftward with increasing N, consistent with finite-size theory.

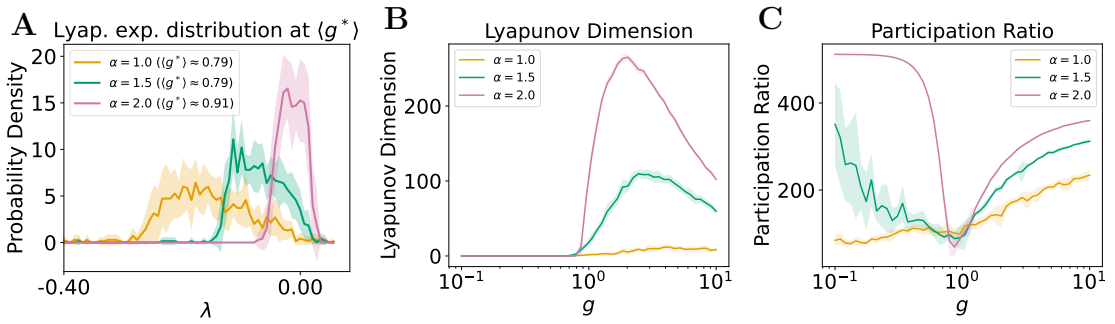

Figure 5: **Attractor geometry under noisy input (N = 1000).** Mean across 10 trials; shaded regions denote ± 1 SD. (A) Lyapunov exponent distributions at g^*: heavier-tailed networks have fewer near-zero exponents, indicating a compressed slow manifold (x-axis truncated for clarity). (B) Lyapunov dimension declines with heavier tails, as in the autonomous case. (C) Participation ratio shows a dip near transition before rising, unlike the monotonic autonomous profile, but remains lower overall for heavier tails.