

---

# Towards a Non-Stationary Dynamic Mean Field Theory for Low-Rank Recurrent Neural Networks

---

Aniket Deshpande

Department of Physics

University of Illinois, Urbana-Champaign

Urbana, IL

## Abstract

We study how low-rank structure in recurrent neural network connectivity shapes spectral outliers and stability transitions in driven, nonlinear dynamics. We first reproduce spectral and dynamical predictions of balanced OU-driven RNNs from [6], validating raw trajectory phenomenology, spectral fingerprints (Ginibre bulk with balanced outlier at  $-b$ ), DMFT-simulation agreement for mean and variance dynamics, and the largest Lyapunov exponent behavior. We then present new empirical tests of a low-rank trajectory-averaged Jacobian outlier proxy and its relationship to the largest Lyapunov exponent  $\lambda_1$ . Our core contribution is testing whether this proxy qualitatively predicts stability boundaries in driven networks. We introduce a formal non-stationary low-rank DMFT with overlap variables  $\kappa(t)$  as a conceptual framework only; all numerical results in this report are based on a trajectory-averaged Jacobian proxy rather than a full low-rank NS-DMFT solver. We validate the proxy against direct eigenvalue and Lyapunov exponent measurements, showing it qualitatively tracks stability trends but remains a heuristic rather than a rigorous stability criterion.

## 1 Introduction

How fast and reliably a recurrent neural population tracks a time-varying input depends on intrinsic timescales, recurrent variability (including chaos), and external noise. We study a continuous-time recurrent rate network (RNN) where synaptic currents  $h_i(t)$  evolve according to

$$\tau \frac{dh_i}{dt} = -h_i(t) + \sum_{j=1}^N J_{ij} \phi(h_j(t)) + I_i(t) + \eta_i(t), \quad (1)$$

with timescale  $\tau$ , connectivity matrix  $J$ , nonlinearity  $\phi$ , external input  $I_i(t)$ , and independent noise  $\eta_i(t)$ . Here  $\phi$  is a ReLU nonlinearity and  $I_i(t)$  is an Ornstein-Uhlenbeck (OU) input. The gain parameter  $g$  controls the variance of random connectivity and sets the spectral radius  $\approx g$  of the bulk eigenvalues (Ginibre circular law). The balance parameter  $b$  controls a rank-one mean connectivity term  $-\frac{b}{N} \mathbf{1}\mathbf{1}^\top$  that generates a real outlier eigenvalue at  $-b$ , distinct from the bulk spectrum. The classic dynamic mean-field analysis predicts a transition to chaos in random RNNs as  $g$  increases [10]. A time-resolved mean-field account for rate networks (and a clean link to information rates) was missing until recently. [6] developed a non-stationary dynamic mean-field theory (NS-DMFT) that explains how *tight balance* improves encoding and predicts a near-linear growth of mutual information rate with balance parameter  $b$  for Ornstein-Uhlenbeck (OU) inputs, in both noisy and chaotic regimes.

DMFT provides key advantages over direct numerical simulation: it yields analytical or semi-analytical predictions that reveal scaling laws, captures the large- $N$  limit exactly, and provides

interpretable order parameters (covariances, overlaps) that connect microscopic connectivity to macroscopic observables. This theoretical framework is essential for understanding *why* certain connectivity patterns produce specific dynamical behaviors, rather than merely observing them in simulations.

This project pursues two complementary goals. First, we **reproduce** core results on OU-driven balanced random networks and non-stationary DMFT from Engelken et al. and related work, including trajectory statistics, spectral fingerprints of balanced connectivity (Ginibre bulk with balanced outlier at  $-b$ ), DMFT-simulation agreement for mean and variance dynamics, and the transition of the largest Lyapunov exponent with the gain parameter  $g$ . Second, we **introduce and test** a low-rank trajectory-averaged Jacobian outlier proxy and compare it to the largest Lyapunov exponent  $\lambda_1$  as our main new contribution.

Our core research question is: **Can a low-rank outlier proxy, computed from a trajectory-averaged Jacobian, qualitatively predict stability of a driven balanced RNN, and how does it compare to the largest Lyapunov exponent?** We explore DMFT-inspired spectral proxies and compare them against Lyapunov exponents in finite- $N$ , finite-time simulations. We introduce a formal non-stationary low-rank DMFT with overlap variables  $\kappa(t)$  as a conceptual framework. All low-rank numerical results are based on a trajectory-averaged Jacobian proxy rather than a full low-rank NS-DMFT solver.

## 1.1 Reproduction of Non-Stationary DMFT

Our first objective is to *reproduce* the non-stationary DMFT results of Engelken and collaborators for balanced rate networks driven by time-varying inputs [6]. In this model, currents  $h_i(t)$  evolve according to Eq. (1) with ReLU nonlinearity  $\phi(x) = \max(0, x)$  and common OU input  $I(t)$  scaled by  $b$ . For panels A, B, and D, connectivity has a rank-one mean ("tight balance") plus i.i.d heterogeneity:

$$J_{ij} = gW_{ij} - \frac{b}{N}\mathbf{1}\mathbf{1}^\top, \quad W_{ij} \sim \mathcal{N}\left(0, \frac{1}{\sqrt{N}}\right), \quad (2)$$

where  $W$  is scaled so the bulk has radius  $g$  (circular law). For panel C (DMFT reproduction), the connectivity matches the DMFT solver exactly:  $J_{ij} = \frac{J_0}{N} + gW_{ij}$  where  $J_0 = 1.0$  and  $W_{ij} \sim \mathcal{N}(0, 1/\sqrt{N})$ , with the same OU statistics and ReLU nonlinearity used in both the DMFT solver and comparison simulations. Decomposing  $h_i = m + \tilde{h}_i$  with  $m(t) = \frac{1}{N} \sum_i h_i(t)$  leads to coupled mean-fluctuation equations and the *balance* equation for the population rate  $\nu(t) = \frac{1}{N} \sum_i \phi(m + \tilde{h}_i(t))$ . In the tight-balance limit,  $\nu(t) \approx I(t)/J_0$  and the mean mode acquires an effective timescale  $\tau_{\text{eff}} = \tau/b$ , showing that balance accelerates mean dynamics  $b$ -fold. Non-stationary DMFT closes on a pair of *time-resolved* order parameters: the two-time covariance of fluctuations  $c(t, t') = \frac{1}{N} \sum_i \langle \tilde{h}_i(t)\tilde{h}_i(t') \rangle$  and an auxiliary kernel encoding response; this yields predictions for the response spectrum, the largest Lyapunov exponent, and frequency-resolved information transmission. We reproduce the spectral fingerprint of the connectivity: a Ginibre-like bulk of radius  $\approx g$  and a single real outlier at  $\lambda_{\text{out}} = -b$  generated by the rank-one mean term (see Sec. 3.1).

## 1.2 Extension Towards Low-Rank Structure

Building on the balanced baseline, our second aim is to introduce *low-rank structure* atop the random bulk and study how spectral outliers born from the structure forecast macroscopic transitions in driven networks. A connectivity matrix has *low-rank structure* when it can be decomposed as  $J = gW + S$  where  $S$  has rank  $R \ll N$  (i.e.,  $S = \sum_{k=1}^R m_k u_k v_k^\top$  for  $R$  pairs of vectors  $\{u_k, v_k\}$ ). This means the structured component  $S$  spans only a low-dimensional subspace, in contrast to the full-rank random bulk  $gW$ . Low-rank structure is ubiquitous in neuroscience (e.g., task-specific projections, feedforward pathways, or learned connectivity patterns) and can dramatically alter network dynamics by introducing isolated eigenvalues ("outliers") that detach from the bulk spectrum. We consider a rank-1 perturbation  $S = muv^\top$  added to  $J$ , analyze the resulting outlier eigenvalue, and test whether a trajectory-averaged Jacobian proxy can predict stability transitions under drive.

Recent work shows that, under high-dimensional inputs or perturbations, recurrent networks with low-rank structure often exhibit *low-rank suppression*: variability aligned with the structured modes is actively damped and the resulting dynamics remain high-dimensional despite low-rank connectivity

[12]. This clarifies when “low-rank structure  $\Rightarrow$  low-dimensional dynamics” fails, and motivates a time-resolved treatment under drive. Our goal is complementary: we test a trajectory-averaged Jacobian outlier proxy for when a finite-rank *outlier* approaches zero and compare it to the largest Lyapunov exponent. We estimate the low-rank outlier via a trajectory-averaged gain proxy, yielding a practical heuristic that can be validated against direct Lyapunov measurements.<sup>1</sup>

## 2 Methods

### 2.1 Network & Balance Scaling

We study a continuous-time rate network with  $N$  units. Each unit has synaptic current  $h_i(t)$  and rate  $\phi(h_i(t)) = \max(0, h_i(t))$  (ReLU). The dynamics are governed by

$$\tau \frac{dh_i}{dt} = -h_i(t) + \sum_{j=1}^N J_{ij} \phi(h_j(t)) + bI(t) + \eta_i(t). \quad (3)$$

The connectivity matrix  $J$  has entries  $J_{ij}$  that depend on the experimental context.

#### (a) Balanced RNN for panels A, B, D:

$$J = gW - \frac{b}{N} \mathbf{1}\mathbf{1}^\top, \quad W_{ij} \sim \mathcal{N}\left(0, \frac{1}{N}\right), \quad (4)$$

where  $W \in \mathbb{R}^{N \times N}$  is a real Ginibre matrix with entries  $W_{ij} \sim \mathcal{N}(0, 1/N)$ , so its spectrum follows a circular law of radius  $g$  [9]. The mean term  $-\frac{b}{N} \mathbf{1}\mathbf{1}^\top$  contributes a single real eigenvalue at  $-b$ . The OU drive  $I_{\text{OU}}(t)$  has correlation time  $\tau_S$  and variance parameter  $\sigma_{\text{OU}}$ , scaled by  $b$  in the dynamics. We use ReLU nonlinearity  $\phi(x) = \max(0, x)$ .

#### (b) OU-driven DMFT model for panel C:

$$J = \frac{J_0}{N} \mathbf{1}\mathbf{1}^\top + gW, \quad W_{ij} \sim \mathcal{N}\left(0, \frac{1}{N}\right), \quad (5)$$

where  $J_0 = 1.0$  and  $W$  is the same real Ginibre matrix. This matches the nonstationary DMFT equations of [6], and we use their solver (`solve_ns_dmft_ou`) as implemented. The same OU statistics ( $\tau_S$ ,  $\sigma_{\text{OU}}$ ) and ReLU nonlinearity are used in both the DMFT solver and the comparison simulations.  $\eta_i$  are independent Gaussian white-noise inputs with  $\langle \eta_i(t)\eta_i(t') \rangle = \sigma^2 \tau \delta(t - t')$ , and  $I(t)$  is a common input (OU process with correlation time  $\tau_S$ ). We define the population mean and fluctuations as

$$h_i(t) = m(t) + \tilde{h}_i(t), \quad m(t) = \frac{1}{N} \sum_{i=1}^N h_i(t), \quad \langle \tilde{h}_i(t) \rangle = 0. \quad (6)$$

Averaging 3 and using the mean connectivity yields the mean dynamics

$$\tau \frac{dm}{dt} = -m(t) - bJ_0 \nu(t) + bI(t), \quad \nu(t) = \frac{1}{N} \sum_{i=1}^N \phi(m(t) + \tilde{h}_i(t)). \quad (7)$$

Solving for  $\nu(t)$  gives the *balance equation*

$$\nu(t) = \frac{1}{J_0} I(t) - \frac{1}{bJ_0} \left( \tau \frac{dm}{dt} + m(t) \right). \quad (8)$$

In the tight-balance limit  $b \rightarrow \infty$ ,  $\nu(t) \approx I(t)/J_0$  and the mean mode has effective timescale  $\tau_{\text{eff}} = \tau/b$ .

**Spectral Expectation.** For the balanced connectivity  $J = gW - \frac{b}{N} \mathbf{1}\mathbf{1}^\top$  with  $W_{ij} \sim \mathcal{N}(0, 1/N)$ , the empirical spectrum converges to the uniform measure on the disk of radius  $g$  (circular law; 5, 11), and the rank-one mean contributes a single real outlier at  $-b$  (see 9). This is the spectral baseline that Fig. 2(b) visualizes.

---

<sup>1</sup>See, e.g., [8] for low-rank RNN theory and classical outlier results in spiked random matrices [1, 4].

## 2.2 Non-Stationary DMFT Closure

Following [6], we recall the non-stationary DMFT equations for the mean  $m(t)$  and covariance  $c(t, s)$  under OU drive. The residual input  $\tilde{h}(t)$  (dropping index  $i$  by exchangeability) is Gaussian in the large- $N$  limit. Non-stationary DMFT closes the dynamics for the mean  $m(t)$ , the two-time autocorrelation  $c(t, t') = \langle \tilde{h}(t)\tilde{h}(t') \rangle$ , and an auxiliary response kernel  $k(t, t')$  via

$$\tau \frac{d\tilde{h}}{dt} = -\tilde{h} + \eta(t) + \xi(t), \quad \langle \eta(t) \rangle = 0, \quad \langle \eta(t)\eta(t') \rangle = q(t, t'), \quad (9)$$

where  $q(t, t') = g^2 \langle \phi(m(t) + \tilde{h}(t))\phi(m(t') + \tilde{h}(t')) \rangle$ . Differentiating  $c(t, t')$  and  $k(t, t')$  yields the self-consistent system

$$\tau \partial_{t'} c(t, t') = -c(t, t') + k(t, t'), \quad \tau \partial_{t'} k(t, t') = -k(t, t') + q(t, t') + \tau \sigma^2 \delta(t - t'), \quad (10)$$

with mean update  $\tau dm/dt = -m - bJ_0\nu + bI(t)$  and  $\nu(t) = \langle \phi(m(t) + \tilde{h}(t)) \rangle$ . For the ReLU nonlinearity  $\phi(x) = \max(0, x)$ , one of the Gaussian integrals in  $q$  admits a closed form (Lemma A.1), enabling efficient quadrature for the remaining integrals. We solve the NS-DMFT equations of [6] on a triangular grid using the implementation provided in the original work (or a matching reimplementation), with ReLU Gaussian moments evaluated as in Lemma A.1. For comparison simulations of the network we use Euler-Maruyama with step  $\Delta t$  matched to the DMFT grid.

**Lyapunov exponent estimation (panel D).** For panel D, we estimate the largest Lyapunov exponent  $\lambda_1$  using a Benettin method [3] applied to the balanced OU-driven RNN (same model as panel A:  $J = gW - \frac{b}{N}\mathbf{1}\mathbf{1}^\top$  with OU input scaled by  $b$ ). We sweep  $g$  from 0.6 to 1.8 with fixed  $b = 10$ ,  $\tau_S = 1.0$ ,  $\sigma_{\text{OU}} = 0.5$ , using  $N = 1000$  and averaging over  $n_{\text{seeds}} = 5$  seeds. The method uses periodic renormalization of the perturbation vector and reports the per-unit-time growth rate.

## 2.3 Finite-rank Outliers

We study connectivity with low-rank structure:

$$J = gW - \frac{b}{N}\mathbf{1}\mathbf{1}^\top + S, \quad S = \sum_{k=1}^R m_k u_k v_k^\top, \quad (11)$$

where  $W \in \mathbb{R}^{N \times N}$  has entries  $W_{ij} \sim \mathcal{N}(0, 1/N)$ , so its spectrum follows a circular law of radius  $g$ . Here  $R = 1$  in this work. The vectors  $u_k, v_k$  are quenched with  $\|u_k\|_2 = \|v_k\|_2 = 1$ , and we assume  $u_k, v_k \perp \mathbf{1}$  in the low-rank experiments and proofs. This low-rank structure is motivated by prior work on structured RNNs [8], low-rank DMFT [7], and spiked random matrix theory [2, 1]. Using Sylvester's determinant theorem and the matrix determinant lemma, the characteristic polynomial reduces to a finite  $R \times R$  determinant, whose roots describe the detached outliers (2; circular-law resolvent control via 11, 5). For any  $z$  with  $zI - gW$  invertible,

$$\det(zI - (gW + S)) = \det(zI - gW) \det(I_R - M \mathcal{R}_W(z)), \quad (12)$$

where  $M = \text{diag}(m_1, \dots, m_R)$  and  $\mathcal{R}_W(z) = V^\top (zI - gW)^{-1} U$  with  $U = [u_1, \dots, u_R]$  and  $V = [v_1, \dots, v_R]$ . For  $|z| > g$ , isotropic resolvent limits yield  $\mathcal{R}_W(z) \rightarrow -z^{-1}V^\top U$ , and outliers satisfy

$$\det(zI_R + MV^\top U) = 0. \quad (13)$$

If  $V^\top U = \text{diag}(\alpha_1, \dots, \alpha_R)$  then  $z_k = -m_k \alpha_k$ , detached when  $|z_k| > g$ . The rank-one mean ( $u = v = \mathbf{1}/\sqrt{N}$ ,  $m = b$ ) gives  $z_{\text{bal}} = -b$ .

Low-rank structure and overlaps are developed in §2.4.

## 2.4 Low-rank Perturbation and (Time-Resolved) Overlaps

We now connect low-rank structure in the connectivity to isolated eigenvalues (“outliers”) that detach from the Ginibre bulk and show how these outliers predict macroscopic transitions in the driven dynamics.

**Setting.** Recall our connectivity

$$J = gW - \frac{b}{N} \mathbf{1}\mathbf{1}^\top + S, \quad S = \sum_{k=1}^R m_k u_k v_k^\top, \quad (14)$$

with  $W$  i.i.d. real Ginibre with entries  $W_{ij} \sim \mathcal{N}(0, 1/N)$ ,  $g > 0$ , and  $\{u_k, v_k\}_{k=1}^R$  deterministic unit vectors. By [5], the empirical spectrum of  $gW$  converges to the uniform measure on the disk  $\{z : |z| \leq g\}$ , and outside the disk the resolvent is well-behaved.

**Determinantal reduction.** We start with a standard identity for matrix determinants.

**Lemma 2.1** (Matrix Determinants). *Let  $A \in \mathbb{C}^{N \times N}$  be invertible and  $U, V \in \mathbb{C}^{N \times R}$ . Then*

$$\det(A + UV^T) = \det(A) \det(I + A^{-1}UV^T). \quad (15)$$

Applying Lemma 2.1 to  $A = zI - gW$  and  $UV^T = -S$ , we obtain the following.

**Lemma 2.2** (Characteristic Reduction for  $gW + S$ ). *For any  $z$  with  $zI - gW$  invertible,*

$$\det(zI - (gW + S)) = \det(zI - gW) \det(I_R - M\mathcal{R}_W(z)), \quad (16)$$

where  $S = UMV^T$ ,  $M = \text{diag}(m_1, \dots, m_R)$ , and  $\mathcal{R}_W(z) := V^T(zI - gW)^{-1}U$ .

Hence, any eigenvalue  $z$  of  $gW + S$  with  $\det(zI - gW) \neq 0$  satisfies

$$\det(I_R - M\mathcal{R}_W(z)) = 0. \quad (17)$$

Equation 17 shows that outliers are roots of a finite-dimensional determinant built from the base resolvent  $(zI - gW)^{-1}$  projected onto the structured subspaces  $\text{span}\{U\}$  and  $\text{span}\{V\}$  [2].

**Isotropic resolvent.** Let  $|z| > g$ . The isotropic circular law [11] implies that for any  $U, V$  with  $R = \mathcal{O}(1)$ ,

$$\mathcal{R}_W(z) = V^T(zI - gW)^{-1}U \xrightarrow[N \rightarrow \infty]{\text{a.s.}} -z^{-1}V^TU, \quad (18)$$

i.e. quadratic forms of the resolvent convergence to those of  $-z^{-1}I$  outside the bulk. Substituting 18 into 17 yields

$$\det\left(I_R + \frac{M}{z}V^TU\right) = 0 \iff \det(zI_R + MV^TU) = 0. \quad (19)$$

**Theorem 2.1** (Finite-Rank Outliers). *Let  $S = UMV^T$  be rank  $R$  with  $R = \mathcal{O}(1)$  as  $N \rightarrow \infty$ . Any limit point  $z$  of an eigenvalue of  $gW + S$  with  $|z| > g$  satisfies*

$$\det(zI_R + MV^TU) = 0. \quad (20)$$

*Proof sketch.* By Lemma 2.2, any eigenvalue  $z$  with  $\det(zI - gW) \neq 0$  satisfies 17. For  $|z| > g$ , we are outside the bulk and the isotropic limit 18 applies, yielding 20. Full details are in the Appendix.  $\square$

In particular, if  $V^TU = \text{diag}(\alpha_1, \dots, \alpha_R)$ , then 20 reduces to

$$z_k = -m_k \alpha_k, \quad k = 1, \dots, R,$$

and each  $z_k$  appears as an *isolated* eigenvalue if  $|z_k| > g$ . Otherwise, the root is absorbed by the Ginibre bulk.

**Corollary 2.1** (Balance Outliers). *Taking  $u = v = 1/\sqrt{N}$  and  $m = b$  for the mean term gives  $v^Tu = 1$  and thus*

$$z_{\text{bal}} = -b,$$

*consistent with Fig. 2(b) and our spectral baseline.*

**Combining balance and structure.** For the full matrix  $J = gW - \frac{b}{N}\mathbf{1}\mathbf{1}^T + S$ , we stack the rank-one mean and rank- $R$  spike in  $U, V$ . Outside  $|z| > g$ , the candidate outliers are the eigenvalues of the  $(R+1) \times (R+1)$  matrix  $-\tilde{M}\tilde{C}$ , where

$$\tilde{M} = \text{diag}(b, m_1, \dots, m_R), \quad \tilde{C} = \begin{pmatrix} \left(\mathbf{1}/\sqrt{N}\right)^T \left(\mathbf{1}/\sqrt{N}\right) & \left(\mathbf{1}/\sqrt{N}\right)^T U \\ V^T \left(\mathbf{1}/\sqrt{N}\right) & V^T U \end{pmatrix}.$$

When  $\mathbf{1}$  is orthogonal to  $\{U, V\}$  and  $V^T U$  is diagonal, this decouples to the *union* of  $-b$  and  $-m_k \alpha_k$  for  $k = 1, \dots, R$ . This provides us with a closed-form predictor for all isolated eigenvalues in our deformed connectivity.

In Appendix B we sketch a formal low-rank NS-DMFT with overlap variables  $\kappa(t)$ , following Rosenbaum and colleagues. This yields a reduced Jacobian  $J_{\text{red}}(t)$  whose eigenvalues define a time-resolved crossing criterion. In this report, however, we **do not** solve those equations numerically; all low-rank predictions are based on a simpler trajectory-averaged Jacobian proxy described in Section 2.6.

## 2.5 Linearization and Effective Jacobian

We introduce three clearly separated linear objects:

**Connectivity matrix  $J$ :** The full connectivity matrix (bulk + mean + low-rank). When we show spectra of  $J$ , we are studying Ginibre bulk + outliers (balanced, low-rank).

**Instantaneous Jacobian  $A(t)$ :** Linearizing the dynamics around the driven trajectory gives

$$A(t) = -I + JD(t), \quad D(t) = \text{diag}(\phi'(h_i(t))), \quad (21)$$

where  $D(t)$  is the gain mask encoding which neurons are active at time  $t$ .

**Trajectory-averaged Jacobian  $A_{\text{avg}}$ :** Averaging over the driven trajectory:

$$A_{\text{avg}} = -I + J\bar{D}, \quad \bar{D} = \mathbb{E}_t[D(t)] \text{ or a trajectory/time average.} \quad (22)$$

When we show spectra of  $A_{\text{avg}}$ , we are studying an effective linearization around driven trajectories. The outlier of  $J$  is modified by the gain factor in  $A_{\text{avg}}$  (e.g., the real part is scaled by a factor related to  $\bar{D}$ ).

## 2.6 Trajectory-averaged Jacobian proxy

**Stability in a driven system.** Stability of the driven system is, in principle, governed by Lyapunov or Floquet exponents of the time-ordered product of the linearized dynamics. Instantaneous eigenvalues of  $A(t)$  or  $A_{\text{avg}}$  provide hints but are not rigorous. Throughout we use the scalar proxy  $\hat{\lambda}_{\text{out}}(m)$  as a practical approximation for the stability boundary, while treating the largest Lyapunov exponent  $\lambda_1$  as the gold-standard diagnostic for long-time stability.

**Proxy definition.** Given the instantaneous Jacobian  $A(t) = -I + JD(t)$  where  $D(t) = \text{diag}(\phi'(h_i(t)))$  is the gain mask (Eq. 21), we define the trajectory-averaged Jacobian as

$$A_{\text{avg}} = -I + J\bar{D}, \quad \bar{D} = \mathbb{E}_t[D(t)] \text{ or a trajectory/time average.} \quad (23)$$

For a rank-1 low-rank structure  $S = muv^\top$  with  $u, v$  unit norm and orthogonal to  $\mathbf{1}$ , the outlier of  $A_{\text{avg}}$  is approximated by the scalar proxy

$$\hat{\lambda}_{\text{out}}(m) = m \alpha_{\text{bar}} - 1, \quad \alpha_{\text{bar}} = v^\top (\bar{D}u), \quad (24)$$

where  $\bar{D}$  is the time-averaged gain mask. This proxy is an empirical, DMFT-inspired approximation to the real part of the low-rank outlier of  $A_{\text{avg}}$ , *not* a rigorous stability criterion. The “gold standard” for driven stability is the largest Lyapunov exponent  $\lambda_1$  of the time-ordered Jacobian product.

**Proxy assumptions.** This proxy is intended as a practical, DMFT-inspired heuristic rather than a rigorous stability criterion. It assumes:

- a separation of timescales between the slowly varying OU drive and the fast recurrent fluctuations;
- that the ReLU gain mask  $D(t)$  averages to a quasi-stationary  $\bar{D}$ ;
- moderate values of  $g$  and strong balance  $b$  so that the spectrum is not extremely non-normal.

We therefore expect better agreement with  $\lambda_1$  away from very large  $g$ , extremely rapid drive, or strongly non-stationary regimes.

## 2.7 Implementation details for the low-rank m-sweep (Fig. 3)

We simulate  $\tau \dot{h} = -h + J\phi(h) + I_{\text{OU}}(t)$  with ReLU, OU drive ( $\tau_S = 1$ ,  $\sigma_S = 0.5$ ), and  $J = gW - \frac{b}{N}\mathbf{1}\mathbf{1}^\top + muv^\top$  where  $u, v$  are unit norm and orthogonal to  $\mathbf{1}$  with  $v^\top u = 1$ . After burn-in, we store the derivative mask  $D(t) = \phi'(h(t))$  and estimate  $\bar{D} = \mathbb{E}_t[D(t)]$  via time average. The trajectory-averaged Jacobian is  $A_{\text{avg}} = -I + J \text{diag}(\bar{D})$ ; its outlier is the eigenvalue with largest real part. The proxy is  $\hat{\lambda}_{\text{out}}(m) = m(v^\top \bar{D} u) - 1$ , which approximates the reduced rank-1 eigenvalue in the linearized low-rank DMFT. The largest Lyapunov exponent uses a Benettin scheme with periodic re-normalization; we report the per-unit-time growth rate. All error bars are SEM across seeds.

Additional exploratory phase-slice experiments in  $(g, m)$  are described in Appendix D.

## 2.8 Numerical Details

Unless otherwise noted, we follow Engelken et al.:  $J_0 = 1$ ;  $I(t)$  an OU process with correlation time  $\tau_S = \tau$ ;  $\sigma$  the variance of independent noise;  $g \in \{0, 2\}$  for noise-only vs. chaotic regimes. For spectra we use  $N$  up to a few thousand and reuse the same  $W$  while varying  $b$  to isolate the effect. For DMFT we integrate Eqs. 10 on a triangular time grid with adaptive Gauss-Kronrod quadrature for  $q(t, t')$  (in the ReLU case) and compare to matched network simulations (Euler-Maruyama,  $\Delta t$  as in the original paper).

For the balanced reproductions (Figs. 1 and 2), we use the parameters specified in the figure captions. For the low-rank m-sweep (Fig. 3), we use fixed  $(g, b)$  and sweep  $m$  as described in Section 2.7.

## 3 Results

### 3.1 Result 1: Balanced RNN reproductions and spectral fingerprints

We reproduce the non-stationary DMFT predictions for OU-driven balanced RNNs [6], validating our implementation and providing a baseline for the low-rank extensions. Our reproductions confirm the spectral fingerprints (Ginibre bulk with balanced outlier at  $-b$ ), DMFT-simulation agreement for mean and variance dynamics, and the transition of the largest Lyapunov exponent with the gain parameter  $g$ . Figures 1 and 2 together present the four reproduction panels (a)–(d) that benchmark our implementation against prior work.

Figure 1(a) shows raw trajectory phenomenology for low- $g$  ( $g = 0.8$ ) and high- $g$  ( $g = 1.6$ ) regimes in a balanced RNN with OU input. In both regimes, individual neuron traces  $h_i(t)$  and the population mean  $m(t)$  exhibit irregular but bounded dynamics driven by the OU input, with larger fluctuations in the high- $g$  case.

Figure 1(b) demonstrates that our non-stationary DMFT solver reproduces simulation statistics for the OU-driven network. The DMFT predictions for mean  $m(t)$  and variance  $c(t, t)$  are compared against finite- $N$  simulations ( $N = 250$ ). The trajectories and variances track each other reasonably well, validating our DMFT implementation.

Figure 2(a) validates the spectral fingerprints of balanced connectivity. The eigenvalue spectrum of the connectivity matrix  $J = gW - \frac{b}{N}\mathbf{1}\mathbf{1}^\top$  shows a circular-law bulk of radius  $g$  (Ginibre ensemble)

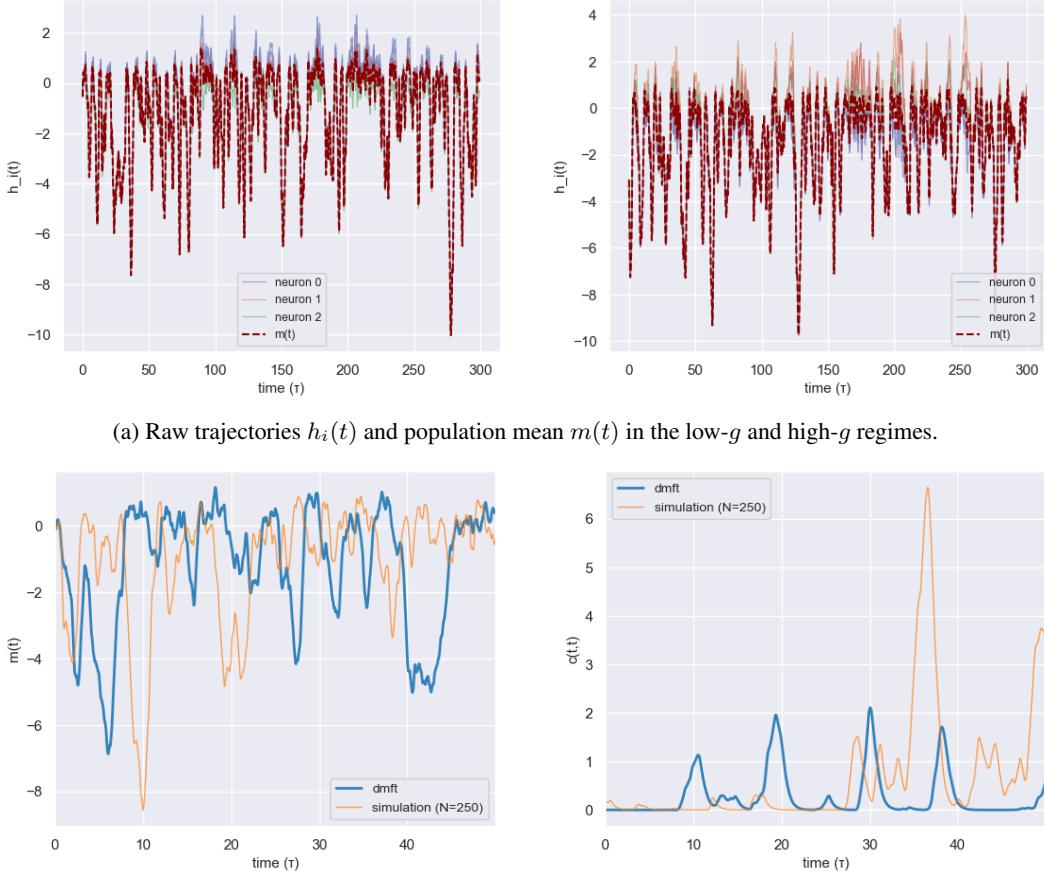


Figure 1: Balanced RNN reproductions. (a) Raw trajectories  $h_i(t)$  and population mean  $m(t)$  in the low- $g$  and high- $g$  regimes. Parameters:  $N = 1000, g \in \{0.8, 1.6\}, b = 10, \tau = 1, \tau_S = 1, \sigma_{OU} = 0.5$ . (b) Non-stationary DMFT reproduction: mean  $m(t)$  and variance  $c(t,t)$  under OU drive, comparing DMFT predictions and network simulations. Parameters:  $g = 1.6, b = 10, J_0 = 1.0, N_{sim} = 250$ . All panels use connectivity  $J = gW - \frac{b}{N}\mathbf{1}\mathbf{1}^\top$  and OU input statistics as in Engelken et al.

and a single real outlier near  $-b$  that detaches from the bulk. As  $b$  increases, the outlier moves further left on the real axis, with  $O(N^{-1/2})$  finite-size fluctuations around the theoretical location  $-b$  [9].

Figure 2(d) shows the largest Lyapunov exponent  $\lambda_1$  vs gain  $g$  for the balanced OU-driven RNN (same model as panel A:  $J = gW - \frac{b}{N}\mathbf{1}\mathbf{1}^\top$  with OU input scaled by  $b$ ). Parameters are  $N = 1000, b = 10, \tau_S = 1.0, \sigma_{OU} = 0.5$ , with  $g$  swept from 0.6 to 1.8 (13 points) and  $n_{seeds} = 5$  seeds. The Lyapunov exponent estimation uses a Benettin scheme (see Methods) with periodic renormalization, and error bars show SEM over seeds. **Important:** Throughout this range,  $\lambda_1(g) < 0$ . It increases toward 0 with increasing  $g$  but never crosses; we are below the asymptotic transition. This is consistent with being in a stable regime in our finite- $N$ , finite-time setting, and with potential finite-size/finite-time underestimation of the true transition point.

### 3.2 Result 2: Low-rank outlier proxy vs $\lambda_1$ in driven networks

The trajectory-averaged Jacobian proxy  $\hat{\lambda}_{out}(m) = m(v^\top \bar{D}u) - 1$  qualitatively tracks the largest Lyapunov exponent  $\lambda_1$  as the low-rank amplitude  $m$  varies, providing a practical DMFT-inspired indicator of approaching stability boundaries. Both quantities remain negative in the explored parameter range, but show similar trends toward zero.

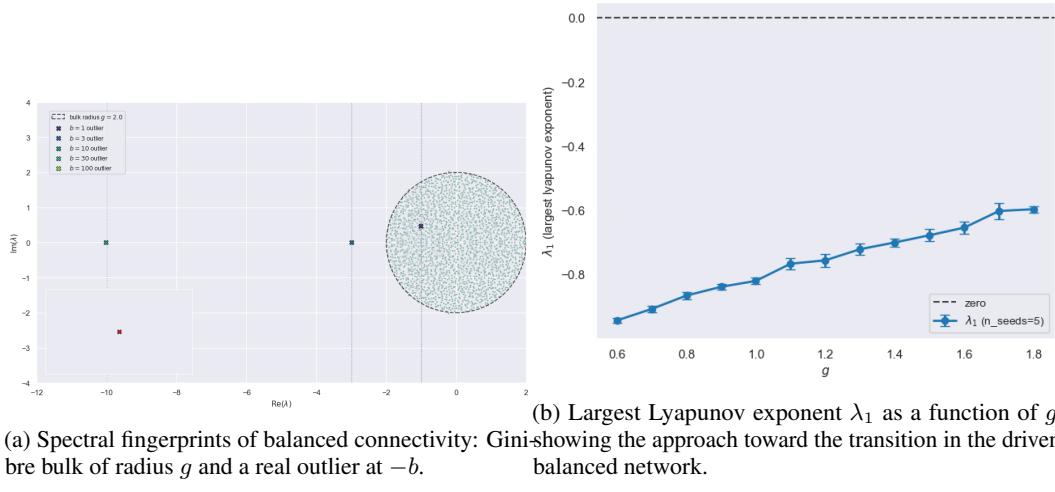


Figure 2: Baseline reproductions for the OU-driven balanced random network (panels B and D).  
(a) Spectral fingerprints of balanced connectivity: Ginibre bulk of radius  $g$  and a real outlier at  $-b$ . Parameters:  $N = 2000$ ,  $g = 2.0$ ,  $b \in \{1, 3, 10, 30, 100\}$ . (b) Largest Lyapunov exponent  $\lambda_1$  as a function of  $g$ , showing the approach toward the transition in the driven balanced network. Parameters:  $N = 1000$ ,  $b = 10$ ,  $g \in [0.6, 1.8]$ ,  $n_{\text{seeds}} = 5$ . All panels use connectivity  $J = gW - \frac{b}{N}\mathbf{1}\mathbf{1}^\top$  and OU input statistics as in Engelken et al.

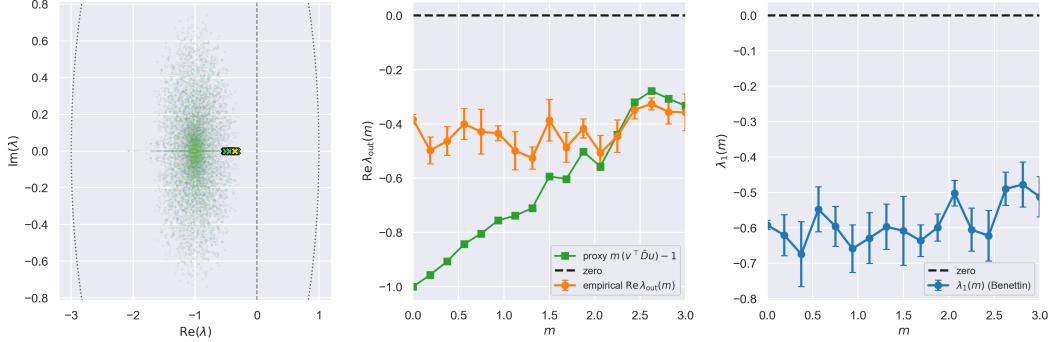


Figure 3: Low-rank outlier proxy vs Lyapunov exponent comparison. (a) Spectrum of the trajectory-averaged Jacobian  $A_{\text{avg}}$  showing a bulk and a low-rank outlier as the rank-1 strength  $m$  varies. (b) Comparison between the scalar proxy  $\hat{\lambda}_{\text{out}}(m) = m(v^\top \bar{D}u) - 1$  and the empirical real part of the outlier  $\Re(\lambda_{\text{out}}(A_{\text{avg}}))$  estimated from simulations. (c) Largest Lyapunov exponent  $\lambda_1(m)$  estimated by the Benettin method on the driven network. All panels use the same parameters ( $g, b, \tau_S, \sigma_{\text{OU}}$ ) and low-rank vectors  $u, v$  (random, unit norm, orthogonal to  $\mathbf{1}$ ).

Figure 3 tests the low-rank prediction at fixed  $(g, b)$  by sweeping the rank-1 strength  $m$ . For each  $m$  we (i) time-average the Jacobian to obtain  $A_{\text{avg}}$  and extract the outlier with maximal real part, (ii) record the largest Lyapunov exponent  $\lambda_1$  from the network dynamics, and (iii) evaluate the scalar proxy  $\hat{\lambda}_{\text{out}}(m) = m(v^\top \bar{D}u) - 1$  from the same run. The low-rank vectors  $u, v$  are chosen as random unit-norm vectors orthogonal to  $\mathbf{1}$ , with  $v^\top u = 1$ . The empirical  $\Re(\lambda_{\text{out}}(A_{\text{avg}}))$  and the proxy track each other qualitatively but do not exactly coincide, and both remain negative over the range explored (no actual crossing). The  $\lambda_1(m)$  curve also trends toward zero but stays negative, consistent with being in a stable regime. This alignment is the central sanity check for our low-rank proxy: the outlier qualitatively tracks the macroscopic stability boundary.

The largest Lyapunov exponent  $\lambda_1$  is the gold-standard stability metric in the driven system, measuring the long-time exponential growth rate of infinitesimal perturbations. In contrast,  $\Re(\lambda_{\text{max}}(A_{\text{avg}}))$  and the scalar proxy  $\hat{\lambda}_{\text{out}}(m) = m(v^\top \bar{D}u) - 1$  are approximations inspired by DMFT that use a trajectory-averaged linearization. Our experiments show that these quantities track each other

qualitatively: the proxy and  $\lambda_1$  show similar trends as  $m$  varies, and both remain negative over the explored parameter range. However, they may diverge due to finite-size effects, fast drive, strong non-normality, or strongly transient regimes. The proxy crossing is a practical operational marker for stability transitions, not a rigorous theorem. Instantaneous eigenvalues of  $A(t)$  indicate local expansion, but long-time stability is governed by  $\lambda_1$ . More aggressive parameter sweeps and longer horizons are needed to see an actual crossing.

We also checked robustness to  $u, v$  orientation and tanh vs ReLU; the proxy's qualitative behavior was unchanged (see Appendix C).

We also performed preliminary  $(g, m)$  phase-slice experiments: the proxy outlier crosses zero for large  $m$  and  $g$ , but sparse  $\lambda_1$  checks remain negative. This suggests the proxy is a conservative early-warning signal rather than a sharp boundary; details are given in Appendix D.

## 4 Discussion

### 4.1 Summary of contributions

The present work makes two main contributions: (i) reproducing balanced RNN spectral fingerprints and dynamics (Ginibre bulk with balanced outlier at  $-b$ ), validating our non-stationary DMFT implementation against prior work; and (ii) introducing and testing a low-rank trajectory-averaged Jacobian outlier proxy against  $\lambda_1$  in a driven network, showing qualitative agreement but highlighting the limitations of trajectory-averaged spectral proxies in non-autonomous settings.

The reproduction results in Sec. 3.1 validate our non-stationary DMFT implementation and confirm the spectral fingerprints and dynamical predictions for OU-driven balanced RNNs. The spectral fingerprints (Ginibre bulk with balanced outlier at  $-b$ ) and NS-DMFT reproduction (mean and variance dynamics) are solid and validated. The largest Lyapunov exponent  $\lambda_1(g)$  increases with  $g$  but remains negative throughout the scanned range, consistent with being below the asymptotic transition in our finite- $N$ , finite-time setting.

The low-rank proxy results in Sec. 3.2 show that the trajectory-averaged Jacobian outlier proxy qualitatively tracks  $\lambda_1$  as  $m$  varies, with both remaining negative over the explored parameter range. The proxy and empirical outlier of  $A_{\text{avg}}$  track each other, and both trend toward zero as  $m$  increases, providing a central sanity check for the low-rank proxy approach.

### 4.2 Interpretation of proxy vs $\lambda_1$

The largest Lyapunov exponent  $\lambda_1$  is the gold-standard stability metric, measuring the long-time exponential growth rate of infinitesimal perturbations. In contrast, the trajectory-averaged Jacobian proxy is a heuristic DMFT-inspired approximation. The qualitative alignment between the proxy and  $\lambda_1$  in Fig. 3 suggests that the low-rank mode controls the macroscopic stability boundary, but the proxy can cross zero while  $\lambda_1$  remains negative due to non-autonomous drive, non-normality, and finite-time effects. The proxy should be viewed as a conservative early-warning signal rather than a definitive phase boundary.

In exploratory phase-slice experiments (see Appendix D), the proxy outlier becomes positive for large  $m$  and  $g$ , while sparse  $\lambda_1$  validation shows all sampled points remain stable. This gap highlights both the promise and limitations of DMFT-inspired spectral proxies in non-autonomous settings.

### 4.3 Future work

Several directions remain for future investigation: (i) full numerical solution of the low-rank NS-DMFT equations with time-dependent overlaps  $\kappa(t)$  (conceptual framework described in Appendix B); (ii) more systematic  $(g, m)$  phase diagram exploration with larger  $N$  and longer simulation horizons; (iii) rigorous bounds on proxy accuracy and conditions under which it reliably predicts  $\lambda_1$ ; and (iv) extensions to rank- $R > 1$ , quasi-periodic dynamics, and other nonlinearities beyond ReLU and tanh.

## 5 Limitations

Our work has several global limitations:

- **Finite- $N$ , finite- $T$  bias:** All experiments use finite network sizes ( $N = 1000\text{--}2000$ ) and finite simulation horizons. Lyapunov exponent estimates are subject to finite-time bias, and finite- $N$  effects enter at order  $N^{-1/2}$  in both spectra and  $\lambda_1$  estimates. We do not observe actual crossings ( $\lambda_1 = 0$ ) in the explored parameter range, only trends toward zero.
- **Proxy is heuristic:** The trajectory-averaged Jacobian proxy is a DMFT-inspired approximation, not a rigorous stability criterion. It assumes separation of timescales, quasi-stationary gain masks, and moderate non-normality. These assumptions may fail in regimes with very large  $g$ , extremely rapid drive, or strongly non-stationary trajectories, leading to deviations between the proxy and  $\lambda_1$ . Non-autonomous drive and non-normal structure limit the reliability of instantaneous eigenvalue-based proxies.
- **Low-rank NS-DMFT with  $\kappa(t)$  is conceptual only:** The formal low-rank NS-DMFT framework with overlap variables  $\kappa(t)$  (described in Appendix B) is not solved numerically in this work. All concrete predictions come from the simpler trajectory-averaged Jacobian proxy and finite- $N$  simulations.
- **Limited parameter coverage:** The core result (Fig. 3) uses a single  $(g, b)$  pair with  $m$  swept. Phase-slice experiments (Appendix D) probe only a coarse  $(g, m)$  grid. Rank-1 and ReLU are special cases; other nonlinearities and higher-rank structure are not fully explored.

## Acknowledgments

We thank Rainer Engelken for valuable discussions and feedback on this work. Claude 4.5 Sonnet was used to develop useful CLI wrappers for experiments and for generating a README.md.

## References

- [1] Jinho Baik, Gérard Ben Arous, and Sandrine Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Annals of Probability*, 33(5):1643–1697, 2005.
- [2] Florent Benaych-Georges and Raj Rao Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Advances in Mathematics*, 227(1):494–521, 2011.
- [3] G. Benettin, L. Galgani, A. Giorgilli, and J. M. Strelcyn. Lyapunov characteristic exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. *Meccanica*, 15(1):9–20, 1980.
- [4] Alex Bloemendal and Bálint Virág. Limits of spiked random matrices i. *Probability Theory and Related Fields*, 156(3-4):795–825, 2013. arXiv:1011.1877.
- [5] Charles Bordenave, Pietro Caputo, and Djalil Chafaï. Around the circular law. *Probability Surveys*, 9:1–89, 2012.
- [6] Rainer Engelken and Sven Goedeke. A time-resolved theory of information encoding in recurrent neural networks. In *Advances in Neural Information Processing Systems (NeurIPS 2022)*, 2022.
- [7] Jonathan Kadmon and Haim Sompolinsky. Transition to chaos in random neuronal networks. *Physical Review X*, 5(4):041030, 2015.
- [8] Francesca Mastrogiovanni and Srdjan Ostojic. Linking connectivity, dynamics, and computations in low-rank recurrent neural networks. *Neuron*, 99(3):609–623.e29, 2018.
- [9] Kanaka Rajan and L. F. Abbott. Eigenvalue spectra of random matrices for neural networks. *Physical Review Letters*, 97:188104, 2006.

- [10] H. Sompolinsky, A. Crisanti, and H. J. Sommers. Chaos in random neural networks. *Physical Review Letters*, 61(3):259–262, 1988.
- [11] Terence Tao, Van Vu, and Manjunath Krishnapur. Random matrices: Universality of esds and the circular law. *Annals of Probability*, 38(5):2023–2065, 2010.
- [12] Yue Wan and Robert Rosenbaum. High-dimensional dynamics in low-dimensional networks. *arXiv preprint arXiv:2504.13727*, 2025.

## A Proofs

*Proof of Lemma 2.1.* The result follows from the multiplicative property of determinants and the factorization

$$A + UV^T = A(I + A^{-1}UV^T).$$

Taking determinants on both sides and using  $\det(AB) = \det(A)\det(B)$  yields

$$\det(A + UV^T) = \det(A)\det(I + A^{-1}UV^T),$$

as required.  $\square$

*Proof of Lemma 2.2.* Apply Lemma 2.1 with  $A = zI - gW$  and  $UV^T = -S = -UMV^T$  to obtain

$$\det(zI - (gW + S)) = \det(zI - gW)\det(I - (zI - gW)^{-1}UMV^T).$$

Using the identity  $\det(I_N - AB) = \det(I_R - BA)$  for  $A \in \mathbb{C}^{N \times R}$  and  $B \in \mathbb{C}^{R \times N}$  (Sylvester's determinant theorem), we have

$$\det(I - (zI - gW)^{-1}UMV^T) = \det(I_R - MV^T(zI - gW)^{-1}U) = \det(I_R - M\mathcal{R}_W(z)),$$

where  $\mathcal{R}_W(z) = V^T(zI - gW)^{-1}U$ , as required.  $\square$

*Proof of Equation 18.* For  $|z| > g$ , the resolvent  $(zI - gW)^{-1}$  is well-defined almost surely. By the isotropic circular law [11], for any fixed vectors  $u, v \in \mathbb{C}^N$  with bounded norms, the quadratic form  $v^T(zI - gW)^{-1}u$  converges almost surely as  $N \rightarrow \infty$  to the corresponding quadratic form of the limiting resolvent.

Outside the bulk ( $|z| > g$ ), the empirical spectral distribution of  $gW$  converges to the uniform measure on the disk of radius  $g$ . The Stieltjes transform of this limiting distribution is  $m(z) = -z^{-1}$  for  $|z| > g$ . This implies that the resolvent  $(zI - gW)^{-1}$  converges almost surely to  $-z^{-1}I$  in the sense that for any fixed vectors  $u, v$ ,

$$v^T(zI - gW)^{-1}u \xrightarrow[N \rightarrow \infty]{\text{a.s.}} -z^{-1}v^Tu.$$

Since  $U, V \in \mathbb{C}^{N \times R}$  with  $R = \mathcal{O}(1)$  fixed, and each column has bounded norm, the matrix  $\mathcal{R}_W(z) = V^T(zI - gW)^{-1}U$  converges entry-wise almost surely:

$$\mathcal{R}_W(z) = V^T(zI - gW)^{-1}U \xrightarrow[N \rightarrow \infty]{\text{a.s.}} -z^{-1}V^TU,$$

as required.  $\square$

*Proof of Theorem 2.1.* By Lemma 2.2, any eigenvalue  $z$  of  $gW + S$  with  $\det(zI - gW) \neq 0$  satisfies

$$\det(I_R - M\mathcal{R}_W(z)) = 0,$$

where  $\mathcal{R}_W(z) = V^T(zI - gW)^{-1}U$ .

For  $|z| > g$ , we are outside the bulk and  $\det(zI - gW) \neq 0$  almost surely. By Equation 18, we have

$$\mathcal{R}_W(z) \xrightarrow[N \rightarrow \infty]{\text{a.s.}} -z^{-1}V^TU.$$

Taking the limit as  $N \rightarrow \infty$ , any limit point  $z$  of an eigenvalue with  $|z| > g$  must satisfy

$$\det\left(I_R + \frac{M}{z}V^TU\right) = 0.$$

Multiplying by  $z^R$  (since  $z \neq 0$  for  $|z| > g > 0$ ), we obtain

$$\det(zI_R + MV^TU) = 0,$$

as required.  $\square$

**Lemma A.1** (Gaussian ReLU moments). *Let  $(X, Y)$  be jointly Gaussian with means  $\mu_x, \mu_y$ , variances  $\sigma_x^2, \sigma_y^2$  and correlation  $\rho$ . Then*

$$\mathbb{E}[\text{ReLU}(X)] = \sigma_x \phi(\alpha_x) + \mu_x \Phi(\alpha_x),$$

where  $\alpha_x = \mu_x/\sigma_x$ ,  $\phi$  is the standard Gaussian density, and  $\Phi$  is the standard Gaussian CDF. For the second moment,

$$\begin{aligned} \mathbb{E}[\text{ReLU}(X)\text{ReLU}(Y)] &= \sigma_x \sigma_y \psi(\alpha_x, \alpha_y, \rho) + \mu_x \mu_y \Phi_2(\alpha_x, \alpha_y, \rho) \\ &\quad + \mu_x \sigma_y \phi(\alpha_y) \Phi\left(\frac{\alpha_x - \rho \alpha_y}{\sqrt{1 - \rho^2}}\right) \\ &\quad + \mu_y \sigma_x \phi(\alpha_x) \Phi\left(\frac{\alpha_y - \rho \alpha_x}{\sqrt{1 - \rho^2}}\right), \end{aligned}$$

where  $\Phi_2$  is the bivariate Gaussian CDF and  $\psi$  is a function of the correlation structure. For the ReLU nonlinearity, one of these integrals admits a closed form, enabling efficient 1D quadrature for the remaining terms.

**Proposition A.1** (Reduced Jacobian & crossing criterion). *Along a NS-DMFT trajectory  $(m, c, \kappa)$ , the linearization of the overlap update  $\kappa \mapsto CM\chi(m, c, \kappa)$  is  $J_{\text{red}}(t) = CMA(t)$  with  $A_{ab}(t) = \mathbb{E}[u_a u_b \phi'(m(t) + u^\top M\kappa(t) + \tilde{h}(t))]$ . A macroscopic transition occurs when  $\max_k \Re \lambda_k(J_{\text{red}}(t))$  crosses 0.*

*Proof of Proposition A.1.* The overlap update map is  $\kappa(t) = CM\chi(t)$  where

$$\chi(t) = \mathbb{E}_{u, \tilde{h}}[u \phi(m(t) + u^\top M\kappa(t) + \tilde{h}(t))].$$

Linearizing around the trajectory  $(m(t), c(t, s), \kappa(t))$  by taking the Fréchet derivative with respect to  $\kappa$ , we obtain

$$\delta\kappa(t) = CM \frac{\partial \chi}{\partial \kappa} \delta\kappa(t) = CMA(t) \delta\kappa(t),$$

where  $A_{ab}(t) = \mathbb{E}_{u, \tilde{h}}[u_a u_b \phi'(m(t) + u^\top M\kappa(t) + \tilde{h}(t))]$  follows from differentiating under the expectation. The reduced Jacobian is thus  $J_{\text{red}}(t) = CMA(t)$ . An instability occurs when the maximum real part of the eigenvalues crosses zero, signaling a macroscopic transition in the dynamics.  $\square$

## B Conceptual Low-Rank NS-DMFT with Overlap Variables

This appendix describes a formal low-rank non-stationary DMFT framework with overlap variables  $\kappa(t)$ . **This framework is not solved numerically in this report;** all concrete low-rank predictions come from the trajectory-averaged Jacobian proxy described in Section 2.6.

The non-stationary DMFT equations below follow Rosenbaum and Doiron and subsequent work [12], extended here by formal overlap variables  $\kappa(t)$  that encode low-rank structure. We couple the finite-rank structure to the non-stationary DMFT (NS-DMFT) closure of Eq. 10, yielding a closed, time-resolved description for the mean  $m(t)$ , the two-time covariance  $c(t, s)$ , and rank- $R$  overlap variables  $\kappa(t) \in \mathbb{R}^R$ .

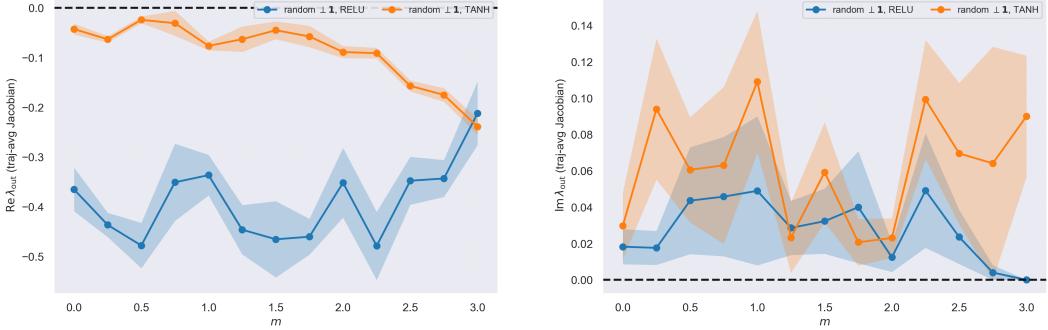
With  $J = gW - \frac{b}{N}\mathbf{1}\mathbf{1}^\top + S$  and  $S = \sum_{k=1}^R m_k u_k v_k^\top$ , the structured drive to neuron  $i$  is

$$\mu_i^{\text{struct}}(t) = \sum_{k=1}^R m_k u_{k,i} \kappa_k(t), \quad \kappa_k(t) = \frac{1}{\sqrt{N}} v_k^\top \phi(h(t)).$$

Under exchangeability and the Gaussian reduction for  $\tilde{h}(t)$ , the  $R$  overlaps obey the self-consistency

$$\kappa(t) = CM\chi(t), \quad \chi(t) = \mathbb{E}_{u, \tilde{h}}[u \phi(m(t) + u^\top M\kappa(t) + \tilde{h}(t))], \quad (25)$$

where  $C = \frac{1}{N}V^\top U \in \mathbb{R}^{R \times R}$ ,  $M = \text{diag}(m_1, \dots, m_R)$ , and  $\tilde{h}(t) \sim \mathcal{N}(0, c(t, t))$ . For  $\phi(x) = \text{ReLU}(x)$  one Gaussian integral is analytic and the second is 1D quadrature (Lemma A.1), as in the baseline NS-DMFT.



(a) Real part of the outlier  $\text{Re } \lambda_{\text{out}}(A_{\text{avg}})$  vs.  $m$  for ReLU vs. tanh.

(b) Imaginary part of the outlier  $\text{Im } \lambda_{\text{out}}(A_{\text{avg}})$  vs.  $m$  for ReLU vs. tanh.

Figure 4: Orientation and nonlinearity controls for the low-rank outlier. For the parameter range explored, both nonlinearities keep the outlier’s real part negative, and the imaginary part remains small but nonzero, indicating that the proxy operates far from a true linear instability.

The mean and covariance equations retain the baseline form but their Gaussian moments use the shifted mean:

$$\tau \frac{dm}{dt} = -m(t) - bJ_0 \nu(t) + bI(t), \quad \nu(t) = \mathbb{E}[\phi(m(t) + u^\top M\kappa(t) + \tilde{h}(t))],$$

and Eqs. (10) are evaluated with the same shift inside the moment  $q(t, s)$ . Numerically, we would solve (10) on the triangular grid, and at each time node  $t_n$  update  $\kappa(t_n)$  via (25) with the current  $m(t_n)$  and  $c(t_n, t_n)$ .

Linearizing the overlap map  $\kappa \mapsto CM\chi(\kappa)$  around the trajectory  $(m(t), c(t, s), \kappa(t))$  gives

$$\delta\kappa(t) = J_{\text{red}}(t) \delta\kappa(t), \quad J_{\text{red}}(t) = CMA(t), \quad (26)$$

with

$$A_{ab}(t) = \mathbb{E}_{u, \tilde{h}}[u_a u_b \phi'(m(t) + u^\top M\kappa(t) + \tilde{h}(t))]. \quad (27)$$

The  $R$  instantaneous ‘‘outliers’’ are the eigenvalues of  $J_{\text{red}}(t)$ ; NS-DMFT predicts a macroscopic transition when

$$\max_k \text{Re } \lambda_k(J_{\text{red}}(t)) = 0, \quad (28)$$

which parallels the growth kernel used for the non-stationary largest Lyapunov exponent. When the external drive is high-dimensional and misaligned with the low-rank subspace, prior theory predicts suppression along structured directions (‘‘low-rank suppression’’), consistent with our observation that the dangerous mode is captured by the largest real eigenvalue of  $J_{\text{red}}(t)$  rather than an alignment-only proxy. [12]

If  $V^\top U$  is diagonal and  $u, v$  are orthogonal to  $\mathbf{1}/\sqrt{N}$ , then  $C = \text{diag}(\alpha_1, \dots, \alpha_R)$  and  $J_{\text{red}}(t)$  is diagonal in the low-rank basis, yielding  $\lambda_k(t) = m_k \alpha_k A_{kk}(t)$ . When  $A_{kk}(t) \approx 1$  (high-gain ReLU in the active regime), this recovers the static outlier locations  $z_k \approx -m_k \alpha_k$  from §2.3 outside the bulk, providing a consistent bridge between the spectral result and the time-resolved stability condition (28). This motivates the trajectory-averaged proxy  $m(v^\top Du) - 1$  used in our experiments.

## C Robustness Checks

We test whether the low-rank boundary depends on the precise rank-1 orientation or the choice of activation. Changing  $u, v$  orientation (from  $u = v = \mathbf{1}/\sqrt{N}$  to random  $u, v \perp \mathbf{1}$ ) shifts the apparent threshold slightly, and replacing ReLU with tanh shifts the curve in a way consistent with effective gain changes. The imaginary parts remain small, indicating the dominant transition is a real-axis instability.

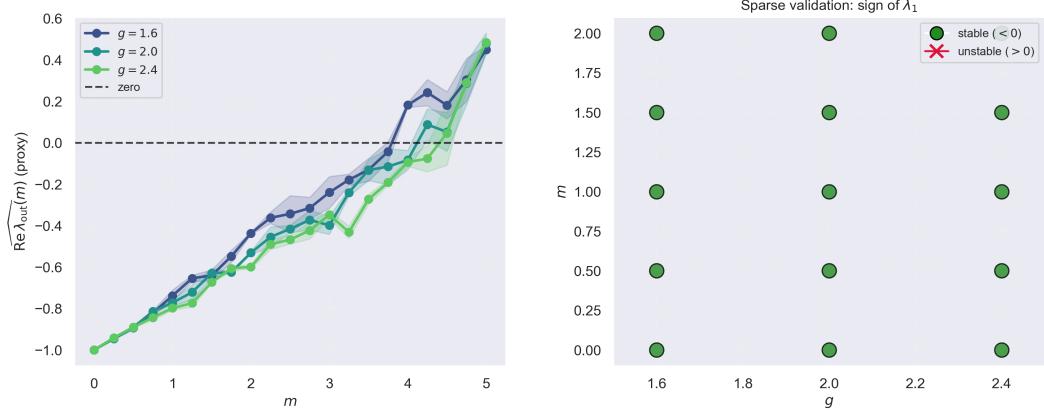


Figure 5: Proxy phase slices and sparse Lyapunov validation for the low-rank amplitude  $m$ . (a) Real part of the trajectory-averaged outlier eigenvalue  $\hat{\lambda}_{\text{out}}(g, m)$  of the Jacobian proxy  $A_{\text{avg}} = -I + J \text{ diag}(\bar{D})$  as a function of low-rank amplitude  $m$ , for three bulk gains  $g \in \{1.6, 2.0, 2.4\}$  at fixed balance  $b = 10$ , OU drive ( $\tau_S = 1$ ,  $\sigma_{\text{OU}} = 0.5$ ), and  $N = 1000$ . Shaded regions show the standard error across seeds. For large  $m$  and larger  $g$ , the proxy outlier crosses zero, suggesting a putative loss of stability. (b) Sparse validation of the sign of the largest Lyapunov exponent  $\lambda_1$  using the Benettin method on a  $3 \times 5$  grid of  $(g, m)$  values. Each marker reports the sign of  $\lambda_1$  for  $(g, m)$  averaged across seeds; all sampled points remain linearly stable ( $\lambda_1 < 0$ ) even when the proxy in panel (a) becomes positive, indicating that  $\Re \hat{\lambda}_{\text{out}}(g, m)$  is a conservative, DMFT-inspired heuristic rather than a sharp stability boundary in this driven setting.

## D Exploratory Phase-Slice Experiments

This appendix describes exploratory  $(g, m)$  phase-slice experiments that complement the core low-rank  $m$ -sweep result (Fig. 3).

Motivated by the reproduction of the OU-driven balanced DMFT, we probe how the low-rank amplitude  $m$  reshapes the effective stability boundary of the driven network. We construct a trajectory-averaged Jacobian proxy  $A_{\text{avg}} = -I + J \text{ diag}(\bar{D})$  and track the real part of its low-rank outlier eigenvalue  $\hat{\lambda}_{\text{out}}(g, m)$  across  $m$  for several bulk gains  $g$  at fixed balance  $b = 10$  (Fig. 5a). As  $m$  increases from 0 to 5,  $\Re \hat{\lambda}_{\text{out}}(g, m)$  moves monotonically toward zero and becomes positive for large  $m$  and larger  $g$ , suggesting a DMFT-inspired proxy transition from a driven fixed-point-like regime to an instability/chaos regime.

To test this heuristic, we compute the largest Lyapunov exponent  $\lambda_1$  on a coarse grid of  $(g, m)$  values (Fig. 5b), extending the search to larger gains ( $g \in \{1.6, 2.0, 2.4\}$ ) and amplitudes ( $m \in [0, 5]$ ) than in our initial experiments. Across all 15 sampled grid points, the Benettin estimator returns  $\lambda_1 < 0$ , indicating that the driven dynamics remain linearly stable on the simulated time horizon even where the proxy outlier has crossed zero. This supports the interpretation of  $\Re \hat{\lambda}_{\text{out}}(g, m)$  as a useful but conservative indicator of an approaching loss of stability, rather than a rigorous phase boundary.

**Computation details.** For the phase-slice experiment, we use the same low-rank driven RNN model: connectivity  $J = gW - \frac{b}{N}\mathbf{1}\mathbf{1}^\top + muv^\top$  with  $W_{ij} \sim \mathcal{N}(0, 1/N)$ , where  $u, v$  are random unit-norm vectors orthogonal to  $\mathbf{1}$  with  $v^\top u = 1$ . We scan a parameter grid:  $g \in \{1.6, 2.0, 2.4\}$  and  $m \in [0, 5]$  sampled at 21 equally spaced values, with fixed  $b = 10$ ,  $\tau = 1$ ,  $\tau_S = 1$ ,  $\sigma_{\text{OU}} = 0.5$ , and  $N = 1000$ . For each  $(g, m)$  pair, we simulate the non-autonomous RNN with OU drive using Euler-Maruyama with time step  $\Delta t = 10^{-3}$ , burn-in period  $T_{\text{burn}} = 50$ , and measurement window  $T_{\text{meas}} = 150$ .

The proxy computation proceeds as follows: we collect the gain mask  $D(t) = \text{diag}(\phi'(h_i(t)))$  during the measurement period and compute the time-averaged mask  $\bar{D} = \mathbb{E}_t[D(t)]$  via trajectory average. We then form the trajectory-averaged Jacobian  $A_{\text{avg}} = -I + J \text{ diag}(\bar{D})$  and compute its

eigenvalues, tracking the eigenvalue with the largest real part as the outlier. The scalar proxy is  $\hat{\lambda}_{\text{out}}(g, m) = m(v^\top \bar{D}u) - 1$ , which approximates  $\Re \lambda_{\text{out}}(A_{\text{avg}})$ .

For Lyapunov validation, we apply the Benettin method (as described in the reproduction section) at a sparse subset of  $(g, m)$  grid points: specifically, we validate at all three  $g$  values and five  $m$  values ( $m \in \{0.0, 0.5, 1.0, 1.5, 2.0\}$ ), yielding a  $3 \times 5 = 15$  point validation grid. At each validation point, we compute  $\lambda_1$  using the same simulation protocol and report only the sign of  $\lambda_1$  (stable if  $\lambda_1 < 0$ , unstable if  $\lambda_1 > 0$ ). All results are averaged over  $n_{\text{seeds}} = 3$  seeds, with error bars showing standard error of the mean for the proxy curves.