

Renormalization group for spectral collapse in random matrices with power-law variance profiles

Philipp Fleig ^{*}*Department of Cellular Biophysics, Max Planck Institute for Medical Research, Jahnstraße 29, 69120 Heidelberg, Germany*

(Received 22 December 2025; revised 1 February 2026; accepted 9 March 2026; published 2 April 2026)

We propose an index-space renormalization group (RG) approach to compare and collapse eigenvalue densities across system sizes in structured random matrix ensembles with rank-ordered power-law variance profiles. The approach is to fix a natural spectral scale by letting the model normalization run with size, turning raw spectra into comparable, collapsed density curves. We illustrate this construction on generalizations of two classic random matrix ensembles—Wigner and Wishart—equipped with rank-ordered power-law variance profiles. We use random matrix theory methods to derive self-consistent fixed-point equations for the resolvent to compute their eigenvalue densities. We define an RG scheme based on index-space decimation and compute the Beta function controlling the RG flow as a function of the variance profile power-law exponent. The running normalization leads to spectral collapse which we confirm in simulations and solutions of the fixed-point equations. We expect that the scale-fixing principle underlying our RG construction extends to other structured ensembles, provided an appropriate renormalization map can be defined.

DOI: [10.1103/zhz2-t8yz](https://doi.org/10.1103/zhz2-t8yz)

I. INTRODUCTION

Many large datasets of complex systems with N interacting units are naturally summarized by matrices: correlations between neurons [1] or genes [2], couplings in materials [3], network connectivities [4], or interaction kernels in collective behavior [5]. A compact, system-level descriptor is the spectral distribution—the distribution of matrix eigenvalues—which encodes properties such as variability, stability, and characteristic scales of the system [6–9]. In practice, across different studies or different runs of an experiment, data matrices may be recorded or simulated at different system sizes N . However, one would like to extract properties of the system in a size-independent way. This raises a basic technical question: How should one compare spectra across matrix sizes, and how can they be *collapsed* onto a common reference curve?

Random matrix theory (RMT) provides a useful language to address this question. It offers a foundation for high-dimensional statistics with macroscopic laws that are universal, i.e., insensitive to microscopic details [6,10–13]. Predictions include principled null models, signal-noise separation, detection, and stability thresholds, and phase transitions in high-dimensional inference [7,8,14–17]. Random matrix theory also supplies standard normalizations and finite- N edge fluctuations—e.g., Wigner’s variance scaling, the aspect ratio in Marčenko-Pastur, and soft Tracy-Widom

edges [6,7,12,18]. However, these asymptotic normalizations and edge scalings standardize particular regimes (e.g., bulk and microscopic edge regimes), but they do not provide a general scale-setting rule for collapsing entire spectral curves across different matrix sizes.

We address this gap by adopting a *renormalization group* (RG) viewpoint [19–21]. The idea is simple and operational: Choose a natural spectral scale (e.g., mean level and bandwidth) to hold fixed as the matrix size changes—analogueous to a renormalization condition—and let a single normalization factor run with matrix size under an index-space decimation map. The map acts on rank-ordered matrix indices rather than integrating out eigenmodes in the Wilsonian sense. We show that with such a running normalization, spectra at different sizes collapse and can be compared on equal footing, while the residual size dependence is funneled into a single running coupling governed by a Beta function [22]. We note that an RG-flavored perspective of RMT in high-dimensional regression has been explored in Ref. [23].

We illustrate the recipe on two generalizations of classic random matrix (RM) ensembles—Wigner [6] and Wishart [7]—with *power-law variance profiles*. Heterogeneity in interaction strengths, sampling variances, or participation weights is widely observed in, e.g., neural population activity and synaptic coupling, heterogeneous networks, and high-dimensional data analysis where sampling is intrinsically nonuniform (see, e.g., Refs. [1,24–28]). Such empirical contexts motivate studying how heavy-tailed or power-law heterogeneity shapes spectral observables across system sizes. We emphasize that we use idealized, analytically tractable ensembles, and we do not claim that these systems realize the specific deterministic variance profiles considered here.

Using RMT methods, we derive self-consistent fixed-point equations for the finite- N RM resolvent, solve them numerically to obtain the eigenvalue density, and find excellent

^{*}Contact author: philipp.fleig@mr.mpg.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Open access publication funded by Max Planck Society.

agreement with simulations. We then derive the running normalization that fixes a chosen spectral scale, demonstrate spectral collapse, define an index-space renormalization scheme (implemented via matrix decimation) and compute the Beta function to delineate relevant, marginal, and irrelevant regimes of the RG flow. At the marginal exponent, the Wigner-type ensemble has a nontrivial RG fixed point, whereas the Wishart-type ensemble exhibits marginal flow with logarithmic corrections and approaches the trivial limit only logarithmically slowly. Finally, we discuss the large- N limit where the coupling flow is governed by a Callan-Symanzik equation [29,30] and derive an integral equation for the large- N RM resolvent. We will provide full details for the Wigner-type ensemble and present the key results for the Wishart-type ensemble.

II. WIGNER ENSEMBLE WITH POWER-LAW VARIANCE PROFILE

A. Ensemble definition

Given system size $N \in \mathbb{N}$, we construct the symmetric $N \times N$ random matrix $\mathbf{X}_N = (X_{jk})$ with independent centered entries:

$$\mathbb{E}[X_{jk}] = 0, \quad X_{jk} = X_{kj}, \quad 1 \leq j \leq k \leq N, \quad (1)$$

$$\text{Var}(X_{jk}) = \frac{\sigma^2}{\gamma} j^{-\alpha} k^{-\alpha}, \quad j < k, \quad (2)$$

$$\text{Var}(X_{jj}) = \frac{2\sigma^2}{\gamma} j^{-2\alpha}, \quad (3)$$

where $\alpha \in \mathbb{R}_{\geq 0}$ is the power-law exponent, $\sigma^2 > 0$ a constant variance scaling factor, and $\gamma > 0$ the normalization factor whose form we are going to determine through our RG analysis. Thus, the RM ensemble consists of matrices with independent, centered entries and variances as defined in Eqs. (2) and (3). The variances follow an “index-ordered” power-law profile that decays as we move diagonally across the matrix, starting from the upper left corner. The index ordering is a modeling convention: We view the ordered power-law variance as a tractable idealization for *rank-ordered heterogeneity*, i.e., after sorting degrees of freedom by an empirical strength measure (variance, degree, participation) the rank profile is modeled as a power law. For $\alpha = 0$ the variance profile becomes constant across matrix entries and by setting the normalization to scale linearly with matrix size, $\gamma \equiv N$, the definition of the standard Wigner ensemble is recovered [6,12].

B. RMT definitions

We begin by fixing notation for the empirical resolvent, its deterministic equivalent, and the associated (smoothed) eigenvalue density. For a random matrix $\mathbf{X}_N$ of size $N \times N$, we define the *empirical resolvent* (or Green’s function) as

$$\mathbf{G}_N^{\text{emp}}(z) := (z\mathbf{I}_N - \mathbf{X}_N)^{-1}, \quad z = x + i\eta, \quad \eta > 0. \quad (4)$$

Taking the normalized trace yields at finite N :

$$G_N^{\text{emp}}(z) := \frac{1}{N} \text{Tr} \mathbf{G}_N^{\text{emp}}(z). \quad (5)$$

For a self-averaging normalized trace, $G_N^{\text{emp}}(z)$ concentrates around its *deterministic equivalent*, $G_N(z)$,

$$G_N^{\text{emp}}(z) \simeq G_N(z), \quad (6)$$

up to fluctuations of order $(N\eta)^{-1/2}$. The deterministic equivalent is defined in Appendix A. The finite- N eigenvalue density is obtained as

$$\rho_{N,\eta}(x) = -\frac{1}{\pi} \text{Im} G_N(x + i\eta), \quad (7)$$

and analogously from G_N^{emp} , up to errors. If $G_N(z)$ converges in the large- N limit, then we denote the limit by $\bar{G}(z)$ and obtain the eigenvalue density from $\rho(x) = -\frac{1}{\pi} \text{Im} \bar{G}(x + i0^+)$. In what follows we work with the deterministic fixed-point objects and only write out the label “emp” explicitly, when needed to avoid ambiguity. Here “fixed point” refers to the fixed-point solution of a self-consistent equation for the resolvent.

In the derivation of deterministic fixed-point equations for the resolvent, we will for simplicity sometimes assume that the matrix entries X_{ik} are Gaussian. However, we note that the fixed-point structure depends only on the variance profile rather than exact Gaussianity. The same set of equations can be derived without assuming Gaussianity (assuming standard moment and tail control [31–33] instead). In simulations we use Gaussian entries. Additionally, we provide numerical evidence hinting at insensitivity of the spectral density to the entry law within a finite-variance class (including a heavy-tailed example). Throughout the article, we adopt physics-style reasoning, instead of full mathematical rigor.

C. Fixed-point equation

We now derive a closed system of *self-consistent* equations for the diagonal resolvent entries, whose solution yields the eigenvalue density. We take the *matrix Dyson equation* (MDE) [32,34] for the deterministic equivalent

$$\mathbf{G}_N(z)^{-1} = z\mathbf{I}_N - \mathbf{\Sigma}(z), \quad (8)$$

where $\mathbf{\Sigma}$ is the *self-energy*, as our starting point. The MDE is a functional equation for the resolvent. In the following we use a heuristic approach to write the MDE explicitly for the Wigner ensemble with power-law variance profile. In Appendix B we give an alternative full-length derivation of the self-consistent equation, starting from the resolvent identity in Eq. (4).

For centered matrices with independent entries and variance profile $s_{jk} := \text{Var}(X_{jk})$, the self-energy $\mathbf{\Sigma}$ is diagonal and is determined by the diagonal of $\mathbf{G}_N$ via

$$\Sigma_{jj}(z) = \sum_{k=1}^N s_{jk} (\mathbf{G}_N)_{kk}(z), \quad \Sigma_{ij}(z) = 0 \quad \text{for } i \neq j. \quad (9)$$

To treat the Wigner ensemble with power-law variance profile it is convenient to introduce the following notation:

$$w_j := j^{-\alpha}, \quad \mathbf{w} := (w_1, \dots, w_N)^T, \quad (10)$$

$$\mathbf{D} := \text{diag}(w_1^2, \dots, w_N^2), \quad (11)$$

where we refer to the w_j s as variance weights. In our setting, the variance profile is encoded in the variance matrix

$\mathbf{S} = (s_{jk})$, which, by Eqs. (2) and (3), is given by

$$s_{jk} := \begin{cases} \frac{\sigma^2}{\gamma} w_j w_k, & j \neq k \\ \frac{2\sigma^2}{\gamma} w_j^2, & j = k \end{cases} \Leftrightarrow \mathbf{S} := \frac{\sigma^2}{\gamma} [w w^T + \mathbf{D}]. \quad (12)$$

This structure preserves diagonality, so the self-consistent equations close on the diagonal

$$(\mathbf{G}_N)_{jj}(z) = \frac{1}{z - \sum_{k=1}^N s_{jk}(\mathbf{G}_N)_{kk}(z)}. \quad (13)$$

With s_{ik} from Eq. (12) this becomes

$$\frac{1}{G_j} = z - \frac{\sigma^2}{\gamma} \left[w_j \sum_{k=1}^N w_k G_k + w_j^2 G_j \right], \quad (14)$$

where we have defined, $G_j := (\mathbf{G}_N)_{jj}$. From this equation, we define two parameters

$$c := \frac{\sigma^2}{\gamma}, \quad \text{and} \quad u(z) := \sum_{k=1}^N w_k G_k(z), \quad (15)$$

which we refer to as the *coupling* and the *variance-weighted mean field*, respectively. Here c is a variance prefactor that multiplies the variance kernel and controls the strength of the Dyson self-energy. We refer to it as a coupling to follow standard RG terminology (cf. Sec. III D). In Eq. (14), site j experiences the variance-weighted mean field, $w_j u(z)$, while the diagonal contribution, $w_j^2 G_j$, is a local term. Overall, for fixed $u(z)$, each G_j solves a scalar quadratic equation

$$\frac{1}{G_j} = z - c(w_j u + w_j^2 G_j), \quad (16)$$

with the physical branch given by $\text{Im} G_j(z) < 0$. Equations (15)–(16) reduce the problem to a one-dimensional, self-consistent fixed-point equation for $u(z)$, from which $G_j(z)$ and hence also the normalized trace, $G_N(z)$, follow. The fixed point serves as a deterministic predictor that approximates $G_N^{\text{emp}}(z)$ up to local-law errors, cf. Appendix A.

As a sanity check, we verify the reduction to the Wigner semicircle law. For $\alpha = 0$, $w_j = 1$ for all j . By symmetry, $G_j(z) = G_N(z)$, and $u = N G_N$. Summing the self-consistent equation in Eq. (16) over sites, we obtain a scalar Dyson equation of the form

$$\frac{1}{G_N} = z - c(u + G_N) = z - c(N + 1)G_N, \quad (17)$$

such that

$$G_N(z) = \frac{z \pm \sqrt{z^2 - R^2}}{2c(N+1)}, \quad R = 2\sqrt{\frac{\sigma^2}{\gamma}(N+1)}. \quad (18)$$

Thus, the limiting eigenvalue density is the semicircle with radius R . Keeping the normalization γ general makes the scaling explicit. The standard Wigner scaling $\gamma = N$ yields $R \rightarrow 2\sigma \sqrt{1 + 1/N} \rightarrow 2\sigma$ as $N \rightarrow \infty$, i.e., the classical Wigner semicircle law with variance parameter σ^2 . An algorithm for solving the fixed-point equations (15)–(16) by fixed-point iteration is described in Algorithm 1 of Appendix C. In Figs. 1(a) and 1(b) we compare the eigenvalue density $\rho_{N,\eta}$ computed

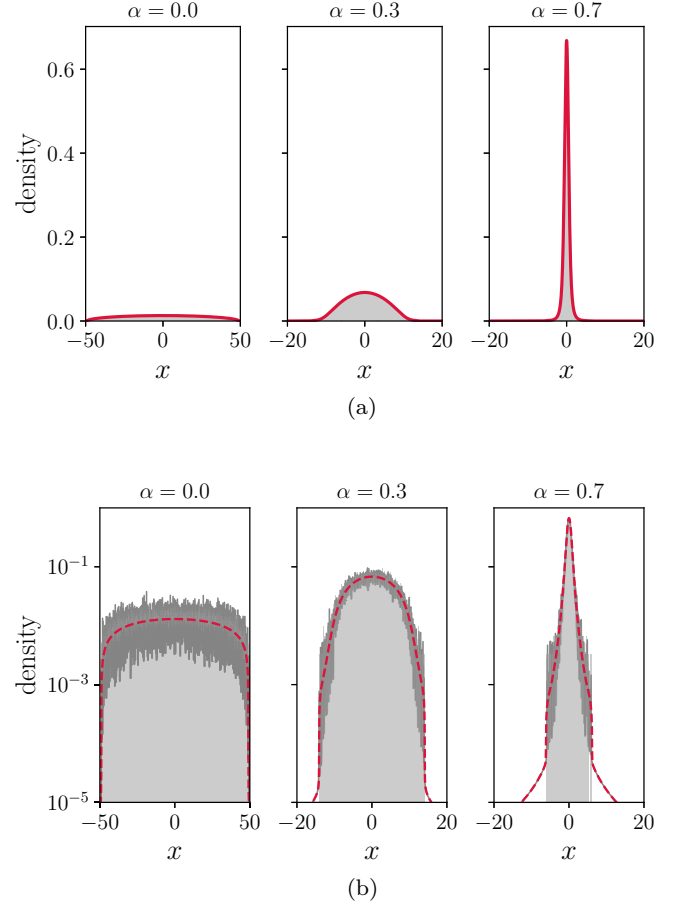


FIG. 1. Eigenvalue densities of Wigner ensemble with power-law variance profile. (a) Linear-linear plot of the eigenvalue density for different values of α from simulations (gray histogram) and the deterministic fixed-point solution of the self-consistent finite- N equations. (b) Log-linear plot of the same densities. In (b) we additionally overlay the empirically Cauchy-smoothed density $\rho_{N,\eta}^{\text{emp}}$ (gray curve) and the deterministic fixed-point prediction $\rho_{N,\eta}$ (red dashed curve) at $\eta = 5 \times 10^{-3}$. In the tails, the gray curve is visually indistinguishable from the red dashed curve. Across all plots we use $\gamma = 1$, $N = 600$, $\sigma^2 = 1$, 20 independent simulation trials, and α values as indicated on the subplots.

from the solution of the fixed-point equations to eigenvalue histograms obtained from simulations with the RM $\mathbf{X}_N$ simulated across multiple independent trials. We find a near-perfect match between fixed-point solutions (red curves) and simulations (gray) for exponent values $\alpha = 0.0, 0.3$, and 0.7 . On the log-scale panel we additionally overlay the empirical Cauchy-smoothed density $\rho_{N,\eta}^{\text{emp}}$ at the same η used in the fixed-point evaluation. Notably, in the tails the smooth empirical curve and the fixed-point solution coincide, making it explicit that the visible tails outside of the unsmoothed support are a standard artifact of Cauchy smoothing (Appendix D). Figures 1(a) and 1(b) show the *raw* eigenvalue densities. Here and below, “raw” refers to densities plotted with the trivial normalization $\gamma = 1$. We note that the support of these raw densities changes significantly across the different values of α .

III. RENORMALIZATION GROUP ANALYSIS

We now turn to an RG analysis of the power-law-structured Wigner ensemble and start by deriving a running normalization $\gamma(N)$ that collapses eigenvalue densities onto each other.

A. RG condition and spectral collapse

The second moment of $\mathbf{X}_N$ is given by

$$m_2 = \frac{1}{N} \mathbb{E}[\text{Tr } \mathbf{X}_N^2], \quad (19)$$

which equals the expected average of squared eigenvalues [the first moment vanishes by Eq. (1)]. We choose the second moment (the typical bandwidth) as the spectral scale that we want to keep constant under renormalization. Fixing the second moment to a constant value m_2^* :

$$m_2 \stackrel{!}{=} m_2^*, \quad (20)$$

is analogous to fixing an energy density per degree of freedom in statistical mechanics or a variance per site in disordered systems [20,35]. In the next section we introduce an RG scheme and we are going to interpret Eq. (20) as a *renormalization condition*. To satisfy Eq. (20) we introduce the normalization $\gamma(N)$ which *runs* with system size and equally the running coupling, $c(N) = \sigma^2/\gamma(N)$, cf. Eq. (15). It is natural to impose m_2^* to be of order unity and without loss of generalization, we set $m_2^* \equiv 1$ in all calculations and figures.

Using the ensemble symmetry $X_{jk} = X_{kj}$, from Eq. (19) we compute the following expression for the second moment:

$$m_2 = c r_N(\alpha), \quad (21)$$

where

$$\begin{aligned} r_N(\alpha) &:= \frac{1}{N} \left[\left(\sum_{j=1}^N w_j \right)^2 + \sum_{j=1}^N w_j^2 \right] \\ &= \frac{H_N(\alpha)^2 + H_N(2\alpha)}{N}, \end{aligned} \quad (22)$$

and we have used the definition of the *generalized harmonic sum*,

$$H_N(\theta) := \sum_{j=1}^N j^{-\theta}. \quad (23)$$

With the definition of the coupling in Eq. (15), and the renormalization condition Eq. (20), Eq. (21) implies

$$\gamma(N) = \frac{\sigma^2}{m_2^*} r_N(\alpha), \quad (24)$$

such that the normalization is N dependent. For the RG analysis we are interested in the dominant scaling behavior with N . We note the following scaling regimes of the generalized harmonic sum: $H_N(\theta) \sim \frac{N^{1-\theta}}{1-\theta}$ for $\theta < 1$, $H_N(\theta) \sim \log N$ for $\theta = 1$, and $H_N(\theta) \rightarrow \zeta(\theta)$ for $\theta > 1$, cf. Appendix E. We find

the dominant large- N behavior of $r_N(\alpha)$ is

$$r_N(\alpha) \sim \begin{cases} \frac{1}{(1-\alpha)^2} N^{1-2\alpha}, & 0 \leq \alpha < 1, \alpha \neq 1/2, \\ \text{const.}, & \alpha = 1/2, \\ \frac{\log^2 N}{N}, & \alpha = 1, \\ N^{-1}, & \alpha > 1, \end{cases} \quad (25)$$

which we confirm numerically in Appendix Fig. 5. Thus, we find α -dependent power-law scaling of the normalization with the system size N for $\alpha \geq 0$, except at $\alpha = 1/2$ where $\gamma(N) = O(1)$, and a logarithmic correction at $\alpha = 1$. We note that by Eq. (15) the coupling scales inverse to the normalization.

With the running normalization $\gamma(N)$ given by Eq. (24), the finite- N spectral densities obtained from simulations of the ensemble at different values of N and fixed exponent α collapse onto a common curve. In Figs. 2(a) and 2(b) we show a comparison of raw and collapsed eigenvalue densities. For $N = 200$, we also show the one standard deviation band due to simulation sampling noise. While the raw curves (left subplots with $\gamma = 1$ normalization) generally do not lie within the sampling noise band, the collapsed curves [right subplots with $\gamma(N)$ normalization] do and largely collapse onto each other. From the collapsed curves we observe some small effects of nonuniform edge convergence. Overall, this demonstrates approximate spectral invariance along the RG trajectory in finite- N simulations: After fixing the second moment via the running normalization $\gamma(N)$, the eigenvalue density is independent of system size up to finite- N corrections. We note that the collapse discussed above concerns the bulk density. For $\alpha > 1$ the bulk density degenerates to a delta mass at zero, and only a finite number of $O(1)$ —outliers associated with small index values remain. These outliers are not self-averaging.

B. Decimation procedure

Next we introduce an index-space decimation map via matrix decimation and interpret the running normalization and coupling in this context. This decimation map plays a role analogous to a coarse-graining step in RG, but it acts on the imposed index ordering rather than integrating out eigenmodes. We define that one decimation step consists of removing a fraction of matrix indices corresponding to the smallest weights w_j (largest index values). We fix a scale factor $b > 1$ and retain indices $j \leq \lfloor N/b \rfloor$. The retained block, $\mathbf{X}^{(\text{dec})}$, inherits the same variance profile with weights $w_j^{(\text{dec})} = j^{-\alpha}$ with $j = 1, \dots, \lfloor N/b \rfloor$, up to rescaling. From Eq. (21) we obtain the postdecimation second moment,

$$m'_2 = \frac{1}{N/b} \mathbb{E}[\text{Tr } \mathbf{X}^{(\text{dec})2}] = c' \frac{H_{N/b}(\alpha)^2 + H_{N/b}(2\alpha)}{N/b}, \quad (26)$$

where we implicitly apply the floor function to noninteger values of N/b . To enforce the renormalization condition in Eq. (20) under one decimation step $N \mapsto N/b$, we choose the renormalized coupling c' so that $m'_2(N/b; c') = m_2(N; c)$ at finite N . This yields the large- N discrete RG update for the running coupling and normalization,

$$\frac{c'}{c} = \frac{\gamma}{\gamma'} \sim \frac{1}{b} \frac{H_N(\alpha)^2 + H_N(2\alpha)}{H_{N/b}(\alpha)^2 + H_{N/b}(2\alpha)} \xrightarrow{N \rightarrow \infty} b^{1-2\alpha}. \quad (27)$$

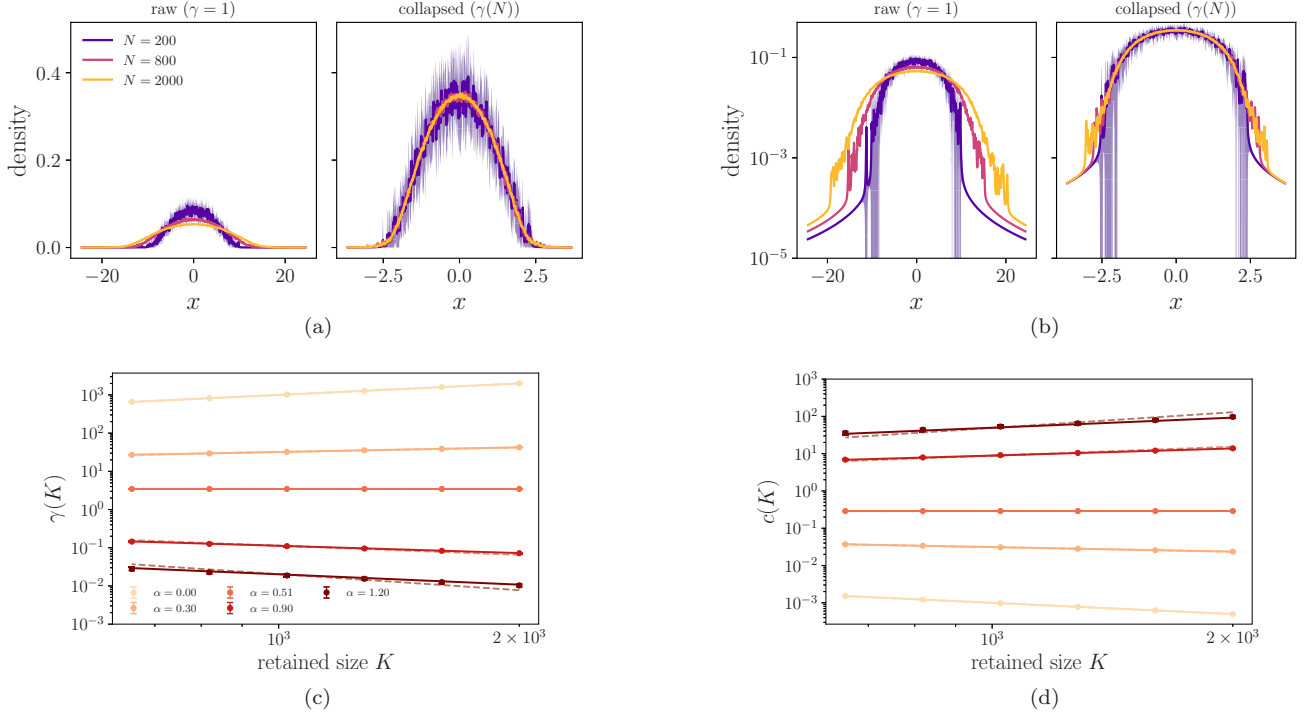


FIG. 2. Eigenvalue density collapse and running normalization. [(a) and (b)] Raw and collapsed eigenvalue densities obtained from simulations of the ensemble. The raw densities (left subplot) use $\gamma = 1$ and the collapsed densities are obtained using the running normalization $\gamma(N)$. We use $\alpha = 0.3$ and $N = 200, 800$, and 2000 and 10 independent ensemble trials. For $N = 200$ we also show the one standard deviation band of sampling noise. Linear-linear scale (a) and log-linear scale (b). (c) Decimation flow of normalization: $\gamma(K)$ versus retained size K after removing a fixed fraction of largest indices at each step, while the second moment is kept fixed at unity. Solid lines: Exact prediction. Markers with error bars: Simulation with mean and standard deviation from six independent trials. Dashed lines serve as guides indicating slope $1 - 2\alpha$. Dashed lines are mostly covered by the exact prediction. (d) Decimation flow of the coupling c .

If we take a total of n decimation steps with fixed scale factor $b > 1$, then the retained system size K_n is

$$K_n := \left\lfloor \frac{N}{b^n} \right\rfloor \Leftrightarrow n = \frac{\log(N/K_n)}{\log b}, \quad (28)$$

with $K_0 = N$. Iterating the update in Eq. (27) yields the n -step flow for the coupling

$$\frac{c_n}{c_0} \sim \prod_{\bar{n}=0}^{n-1} \frac{1}{b} \frac{H_{N/b^{\bar{n}}}(\alpha)^2 + H_{N/b^{\bar{n}}}(2\alpha)}{H_{N/b^{\bar{n}+1}}(\alpha)^2 + H_{N/b^{\bar{n}+1}}(2\alpha)} \xrightarrow{N \rightarrow \infty} b^{n(1-2\alpha)}, \quad (29)$$

such that asymptotically, $c_n \propto b^{n(1-2\alpha)}$, or, using Eq. (28):

$$c_n \propto \left(\frac{N}{K_n} \right)^{1-2\alpha} \Rightarrow \gamma(K_n; N) = \frac{\sigma^2}{c_n} \propto K_n^{1-2\alpha}, \quad (30)$$

where the proportionality is up to a constant that does not depend on K_n (for fixed initial $K_0 = N$).

In Figs. 2(c) and 2(d), we show the running of the normalization $\gamma(K)$ and coupling $c(K)$ with retained system size K , respectively. Dashed lines serve as reference to indicate the theoretical scaling exponent, $1 - 2\alpha$, for the RG flow under decimation. Observed deviations from the reference lines for $\alpha > 1/2$ arise from finite-size effects, competition of scaling, and slow log-corrections.

For the present index-ordered power-law profile, which is form-invariant under decimation (the retained block inherits

the same functional form of the variance weights, up to rescaling), the decimation flow is consistent with the scaling of $\gamma(N)$ obtained from asymptotic analysis, because the same scale-fixing condition is reimposed at the retained size. For more general profiles that are not form invariant under decimation, different renormalization maps can lead to different flows and it is not obvious if the same scaling is obtained.

C. Beta function and RG regimes

We define the running log-coupling, $\kappa := \log c$, with $c(K_n) = \sigma^2/\gamma(K_n)$. Under one decimation step with per-step scale factor b , the discrete update is

$$\kappa_{n+1} - \kappa_n = (1 - 2\alpha) \log b, \quad (31)$$

which integrates to

$$\kappa_n = \kappa_0 + n(1 - 2\alpha) \log b. \quad (32)$$

Writing the total scale factor as, $B := b^n$, and defining the RG scale, $s := \log B$, the continuous infinitesimal flow along the RG trajectory with m_2 fixed becomes

$$\beta_{\text{RG}}^b(c) := \frac{d\kappa}{ds} = \frac{d \log c}{d \log B} = 1 - 2\alpha. \quad (33)$$

This is the *Beta function* [22] for c of the RM ensemble. It reveals three, physically distinct, regimes under decimation:

Relevant ($\alpha < 1/2$; $\beta_{\text{RG}}^b > 0$). Under decimation to a smaller retained size K , $\gamma(K) \propto K^{1-2\alpha}$ decreases as K

decreases, while $c(K)$ increases under decimation. This is the RG sense in which the coupling is relevant.

Marginal ($\alpha = 1/2$; $\beta_{\text{RG}}^b = 0$). The coupling sits at an RG fixed point (a fixed line) and the Beta function changes sign at marginality. Since the fixed point remains interacting (non-vanishing self-energy and hence not the free resolvent, $1/z$), the fixed point is nontrivial. The normalization $\gamma(K)$ tends to a constant and the coupling $c(K)$ is scale stationary. Any residual drift under decimation is a finite-size effect.

Irrelevant ($\alpha > 1/2$; $\beta_{\text{RG}}^b < 0$). Under decimation to smaller K , $\gamma(K)$ increases and $c(K)$ decreases, i.e., the coupling weakens under decimation. As already noted above, for $\alpha > 1$, the bulk density degenerates and non-self-averaging outliers can appear.

D. RG flow and the Callan-Symanzik equation

We now investigate the RG flow in the large- N limit. For this, we set the RG scale, $s := \log N$, and let the coupling run as $c(s) = \sigma^2/\gamma(s)$ so that the spectral scale (second moment) remains fixed along the flow. The running of the coupling is characterized by the Beta function

$$\beta_{\text{RG}} := \frac{d \log c}{d \log N} = -(1 - 2\alpha), \quad (34)$$

with a logarithmic correction at $\alpha = 1$. In this alternative definition, the Beta function differs by a sign from the Beta function in Eq. (33). Both are valid forms of the Beta function for the RM ensemble. Along this RG trajectory, spectral observables should be invariant to changes in the scale s when the induced change in the coupling $c(s)$ is taken into account. One such observable is the mean resolvent, $G_N(z) = \frac{1}{N} \sum_{i=1}^N G_i(z)$. In the large- N limit, we denote the mean resolvent by $\bar{G}$. In Appendix F, we derive an integral equation for $\bar{G}$ and solve it numerically. This continuum equation provides a closed description of the large- N resolvent along the RG trajectory and reproduces the collapsed bulk density.

Invariance along the flow is stated as, $d\bar{G}/ds = 0$. Separating explicit and implicit scale dependence, through $c = c(s)$, yields the equation

$$(\partial_s + \beta_{\text{RG}} c \partial_c) \bar{G}(z; c) = 0, \quad (35)$$

where we have used the Beta function in Eq. (34). This equation is known as the Callan-Symanzik equation [29,30] and holds in the large- N limit along the RG trajectory. The equation captures how a shift in the scale s is compensated by an induced change of the coupling c , leaving the mean resolvent $\bar{G}$ invariant.

At finite N , the quantity $d G_N/ds$, is generally nonzero due to finite-size effects, and the Callan-Symanzik equation is not satisfied exactly. However, as we previously showed for simulated data in Figs. 2(a) and 2(b), finite-size effects are controlled, implying that approximate invariance, and therefore collapse, of the eigenvalue density across different sizes N , is still observed for a running coupling. As an additional step, we demonstrate the approximate invariance on the level of the resolvent. For fixed α , we solve the self-consistent Eqs. (15) and (16), for different N and the running normalization of Eq. (21) and extract the density $\rho_{N,\eta}(x)$. In Fig. 3, we show a comparison of raw and collapsed eigenvalue densities

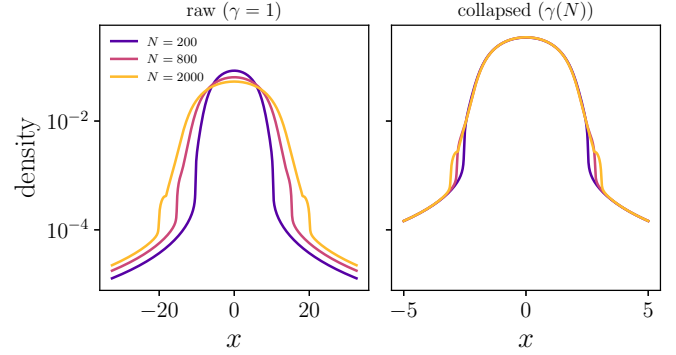


FIG. 3. Spectral collapse for the finite- N self-consistent equation. Spectral invariance test for $\alpha = 0.30$ across $N = 200, 800$, and 2000 . Left: Raw eigenvalue density $\rho_{N,\eta}(x)$ with $\gamma = 1$ computed from the finite- N self-consistent equation for different N . Right: Collapsed eigenvalue density with running normalization $\gamma(N)$. We use $\eta_{\text{raw}} = \sqrt{\gamma(N)}\eta_{\text{coll}}$, with $\eta_{\text{coll}} = 10^{-2}$ for all N .

for three values of N . Up to nonuniform edge convergence, the curves collapse onto each other, confirming spectral invariance also for finite values of N to good approximation. In Appendix D, we discuss our choice of η values for the raw and collapsed curves.

It is worth emphasizing that while the leading scaling of the running normalization can be obtained from asymptotic analysis of the scale-fixing condition, the decimation construction specifies an explicit discrete renormalization map for the running coupling under $N \mapsto N/b$. This provides a step-by-step prediction for how the coupling changes when the renormalization condition is reimposed, and it enables finite- N invariance tests along the induced trajectory (at the level of the density and the resolvent), while also classifying relevant, marginal, and irrelevant regimes under this specified map.

The predictive power of the RG comes from *universality*, which means that the RG flow is not an artifact of distributional choices of the model. For our RM ensemble this could mean that its universality class is determined by the power-law exponent and the multiplicative variance weights but that spectral observables are largely insensitive to the particular choice of matrix entry law (subject to some conditions). In Appendix Fig. 6 we show a test of several entry-law distributions under the RG normalization that fixes the second moment and observe collapse across sizes in agreement with the solution of the self-consistent equation: Gaussian, Rademacher (± 1), and Student's t with finite variance (e.g., $\nu = 4$) and the matrix entries are scaled by the variance weights w_j to obtain the power-law variance profile. Within this ensemble class, the bulk of the eigenvalue density and edges appear insensitive to the entry-law after scale fixing.

IV. WISHART ENSEMBLE WITH POWER-LAW VARIANCE PROFILE

We study a Wishart ensemble with power-law heterogeneity placed in the *rows* (e.g., observation index). Let $\mathbf{Z} \in \mathbb{R}^{T \times N}$ have independent and identically distributed centered

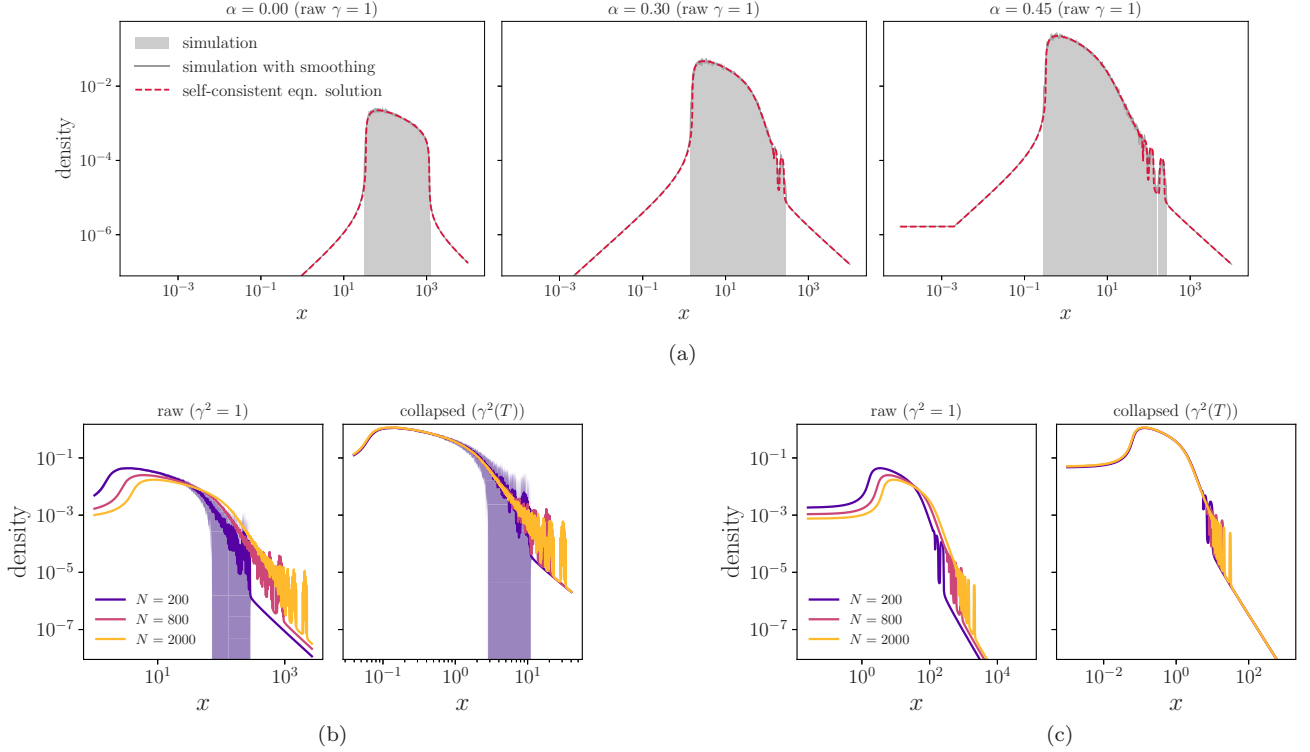


FIG. 4. Wishart ensemble with power-law variance profile. (a) Comparison of spectral distribution simulations (gray histogram), the fixed-point solution of the self-consistent equations (red dashed), and the empirically Cauchy-smoothed density (gray curve) computed with the same smoothing as the fixed-point curve. In the tails, the gray curve is visually indistinguishable from the red dashed curve. Across all subplots we use $\gamma = 1$, $N = 200$, $T = 400$ (i.e., $q = 0.5$), and $\sigma^2 = 1$, with α values as indicated above the plots. (b) Raw and collapsed eigenvalue densities obtained from simulations of the ensemble. We use $\alpha = 0.3$ and $N = 200, 800$, and 2000 , and $q = 0.5$ and simulate 100 independent trials of the ensemble. (c) Raw and collapsed eigenvalue density from the fixed-point solution for the same parameters as (b). Left: Raw densities with $\eta_{\text{raw}} = \gamma^2(T) \eta_{\text{coll}}$. Right: Collapsed densities with $\eta_{\text{coll}} = 10^{-2}$ for all N .

unit-variance entries; let $\mathbf{D} := \text{diag}(v_1, \dots, v_T)$, where $v_t = t^{-2\alpha}$ with $t = 1, 2, \dots, T$; and define

$$\mathbf{Y} := \mathbf{D}^{1/2} \mathbf{Z}, \quad \mathbf{C} := \frac{1}{\gamma^2} \mathbf{Y}^T \mathbf{Y} = \frac{1}{\gamma^2} \mathbf{Z}^T \mathbf{D} \mathbf{Z} \in \mathbb{R}^{N \times N}. \quad (36)$$

The ordered power-law row-variance weights v_t are a modeling convention, meant as a tractable idealization of rank-ordered heterogeneity across observations after sorting rows by an empirical strength measure such as variance or participation. We also define the aspect ratio of matrix dimensions, $q := N/T$ and think of $\mathbf{Y}$ as a $T \times N$ data matrix with variance decaying as a power law across its rows, and $\mathbf{C}$ is the associated sample covariance matrix scaled by γ^2 .

In Appendix G we give a cavity method derivation of the following finite- N self-consistent Dyson equation for the average resolvent of $\mathbf{C}$:

$$G_N(z) = \frac{1}{z - \Sigma_N(z)}, \quad (37)$$

$$\Sigma_N(z) = \frac{1}{\gamma^2} \sum_{t=1}^T \frac{v_t}{1 - (N/\gamma^2) v_t G_N(z)}. \quad (38)$$

For background on the derivation see Refs. [10,32,33,36,37]. As a sanity check we note that for $\alpha = 0$ the variance profile of $\mathbf{Y}$ becomes homogeneous and setting $\gamma^2 \equiv T$, the self-consistent equation reduces to a form that yields the

well-known Marčenko-Pastur distribution with aspect ratio q [7,12]. An algorithm to solve the self-consistent equations (37)–(38) is described in Algorithm 2 of Appendix G. In Fig. 4(a), we show a comparison of the simulated eigenvalue density and the fixed-point solution obtained from the self-consistent equations and find an exact match. In Appendix I we discuss sampling of large- x eigenvalue bands.

The first moment (mean eigenvalue level) is given by

$$m_1 = \frac{1}{N} \mathbb{E} [\text{Tr } \mathbf{C}] = \frac{1}{\gamma^2} \sum_{t=1}^T t^{-2\alpha}. \quad (39)$$

As a renormalization condition, we fix the first moment to a constant value m_1^* which we set to $m_1^* \equiv 1$, without loss of generality. To implement the renormalization scheme for the Wishart ensemble with power-law variance profile, we fix the mean level across sizes by choosing a running normalization $\gamma^2(T)$ that compensates the growth of the row weights. Operationally, one coarse-graining step acts on the data matrix. We remove a fraction of the smallest-weight rows of $\mathbf{Y}$ (equivalently, truncate $\mathbf{D}$ to its top-left block), retaining $T' := \lfloor T/b \rfloor$ rows with weights $v'_t = t^{-2\alpha}$, $t = 1, \dots, T'$. If only rows are decimated, then N is unchanged and the aspect ratio changes as $q' = N/T' = bq$. If a fixed aspect ratio is desired, then we decimate columns accordingly, $N' := \lfloor N/b \rfloor$, so that $q' \approx q$. Here we adopt the latter approach. The normalization

γ^2 is then updated to γ'^2 so that $m'_1 = m_1^*$, where

$$m'_1 = \frac{1}{N'} \mathbb{E}[\text{Tr} \mathbf{C}^{(\text{dec})}]. \quad (40)$$

The induced finite- T flow of γ^2 under the above decimation is shown in Appendix Fig. 8, and its large- T asymptotics is summarized by

$$\gamma^2(T) \propto \sum_{t=1}^T t^{-2\alpha} \sim \begin{cases} \frac{T^{1-2\alpha}}{1-2\alpha}, & 0 \leq \alpha < 1/2, \\ \log T, & \alpha = 1/2, \\ \zeta(2\alpha), & \alpha > 1/2 \text{ (as } T \rightarrow \infty), \end{cases} \quad (41)$$

where ζ is the Riemann zeta function. For the RG scale, $s := \log T$, it is natural to track the running *coupling*

$$g(T) := \gamma^{-2}(T) \quad (42)$$

and define its Beta function by

$$\beta_{\text{RG}}(g) := \frac{d \log g}{d \log T}. \quad (43)$$

The expression is well defined for all α and quantifies how the strength of the covariance fluctuations changes with scale. For $\alpha < 1/2$ one has $\gamma^2(T) \sim T^{1-2\alpha}$, hence $g(T) \sim T^{-(1-2\alpha)}$ and $\beta_{\text{RG}} \rightarrow -(1-2\alpha) < 0$. Therefore, the coupling strengthens when the scale is reduced, which is the RG sense of relevance. At $\alpha = 1/2$ one has $\gamma^2(T) \sim \log T$, hence $g(T) \sim 1/\log T$ and $\beta_{\text{RG}} \rightarrow 0$ with logarithmic corrections, i.e., marginal flow. For $\alpha > 1/2$ the normalization saturates, $\gamma^2(T) \rightarrow \zeta(2\alpha)$ as $T \rightarrow \infty$, so $g(T) \rightarrow \zeta(2\alpha)^{-1}$ and $\beta_{\text{RG}} \rightarrow 0$, making the heterogeneity irrelevant in the sense that the coupling approaches a size-independent limit. In Appendix H we discuss a linearized Beta function that quantifies the approach to the fixed point for $\alpha > 1/2$ and recovers a nonzero scaling exponent.

In Figs. 4(b) and 4(c), we show a comparison of the raw and collapsed eigenvalue densities, for simulations and solutions of the self-consistent Eqs. (37) and (38), respectively. Analogously to the Wigner-type ensemble, we visualize the spectra via Cauchy-kernel smoothing, choose a single η_{coll} for all sizes in the collapsed panel, and set $\eta_{\text{raw}} = \gamma^2(T) \eta_{\text{coll}}$. We find that eigenvalue densities collapse onto each other, across the different sizes. Differences in the number of density peaks at the right end of the densities are due to better sampling of the power-law profile for larger values of N .

With the running normalization $\gamma^2(T)$ given by Eq. (41), we can define an effective aspect ratio,

$$q_{\text{eff}}(T) := \frac{N}{\gamma^2(T)}. \quad (44)$$

This ratio plays a role analogous to the Marčenko-Pastur parameter q by controlling how well the dominant, large-variance rows are sampled. For $\alpha < 1/2$ the growth of $\gamma^2(T)$ yields a nontrivial bulk. At $\alpha = 1/2$ the growth is only logarithmic and, in the strict limit, almost all eigenvalues accumulate near zero, while only a vanishing fraction do not approach zero. For $\alpha > 1/2$ the normalization saturates and the bulk collapses toward zero. In this case, the renormalization fixes m_1 by a finite number of outliers of size $O(N)$,

while the $O(1)$ bulk collapses to zero. We note that at finite N , the deterministic fixed-point equation for the resolvent remains quantitatively predictive also for $\alpha > 1/2$, provided one compares Cauchy-smoothed densities at fixed smoothing width η in the plotted units (and, in simulations, averages over realizations to reduce sampling noise).

Finally, in Appendix I, we provide a derivation of an integral equation for the large- N continuum resolvent and check its correctness through a numerical comparison with the simulated eigenvalue density in Appendix Fig. 9.

V. DISCUSSION

We used an RG perspective to compare eigenvalue spectra across matrix sizes for two structured ensembles—Wigner and Wishart with rank-ordered power-law variance profiles. In this work, we treat these as tractable idealizations of rank-ordered heterogeneity.

The collapsed spectral distribution yields size-aware dynamical thresholds across different systems. For linear dynamics built from an interaction matrix, stability changes when the right spectral edge crosses the control baseline, and the slowest relaxation time is set by the distance to that edge [40]. For consensus and diffusion on networks, the Laplacian spectral gap controls mixing times and convergence rates [38,39]. Finally, in neural network dynamics, the spectral radius marks the onset of unstable or chaotic activity [1,8]. Because the RG normalization collapses spectra across sizes, the edge and gap thresholds can be set once at a reference size (in collapsed units) and then applied unchanged across N , yielding consistent stability and relaxation criteria. Applying the scale-fixing spectral-collapse recipe to concrete systems is a natural next step.

To make the recipe operational on data, one should quantify uncertainty—i.e., statistical variability across realizations at fixed N (sampling noise)—and separate it from deterministic finite-size drift ($1/N$ bias along the RG trajectory), for example by reporting confidence bands for smoothed densities and edge locations (at fixed resolution). Furthermore, sharper finite- N controls near marginality and at spectral edges would bound these drifts and strengthen guarantees behind collapse and Callan-Symanzik invariance.

We have shown that the RG behavior is nontrivial yet analytically tractable. While we do not claim full universality, our evidence provides first hints that after scale fixing, the bulk and edges are largely insensitive to the entry law. A systematic program—varying entry distributions and adding mild structural perturbations—would delineate the empirical domain of attraction of the collapsed spectral distribution and its edge.

Finally, for the two ensembles we have studied, the power-law structure suggested a simple index-space decimation consistent with the imposed hierarchy, implemented via matrix decimation. Extending our RG-like construction beyond these cases will require model-specific renormalization maps. Several complementary approaches have been explored on real and synthetic data [41–46], which develop alternative coarse-graining transformations. A key direction is to combine such transformations with scale-fixing renormalization

conditions to construct renormalization maps for broader ensemble classes.

ACKNOWLEDGMENTS

We thank Ilya Nemenman for discussions on RMT in which the idea of using RG for anisotropic matrices via decimating the smallest variance variables emerged and we also thank him for helpful comments on this work. We also thank Francesca Mastrogriuseppe for discussions.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX A: (ANISOTROPIC) LOCAL LAWS

In RMT we distinguish spectral resolutions by the scaling of the imaginary part η of $z = x + i\eta$: macroscopic resolution keeps η fixed as $N \rightarrow \infty$; mesoscopic resolution lets η shrink with N but still large enough to cover many levels; microscopic resolution probes the mean spacing scale. Concretely:

$$\eta = O(1) \text{ (macroscopic),} \quad (\text{A1})$$

$$N^{-1+\varepsilon} \leq \eta \ll 1 \text{ (mesoscopic),} \quad (\text{A2})$$

$$\eta \lesssim N^{-1} \text{ (microscopic),} \quad (\text{A3})$$

for some fixed $\varepsilon > 0$. On mesoscopic scales, *local laws* [31,32] quantitatively control and imply the self-averaging of resolvents. Recall that $\mathbf{G}_N^{\text{emp}}$ and $\mathbf{G}_N$ denote the empirical resolvent and its deterministic equivalent, respectively. Then averaged statistics concentrate uniformly off the real axis:

$$\frac{1}{N} \text{Tr} \mathbf{G}_N^{\text{emp}}(z) \simeq \frac{1}{N} \text{Tr} \mathbf{G}_N(z), \quad (\text{A4})$$

$$\left| \frac{1}{N} \text{Tr}(\mathbf{G}_N^{\text{emp}}(z)) - \frac{1}{N} \text{Tr}(\mathbf{G}_N(z)) \right| \ll 1, \quad (\text{A5})$$

with typical error scale $(N\eta)^{-1/2}$. In the equations “ $\simeq$ ” denotes equality up to controlled, vanishing errors. Or, equivalently, $\mathbb{E}[G_N^{\text{emp}}(z)] = G_N(z) + O((N\eta)^{-1/2})$ on mesoscopic domains. So either route, expectation or deterministic equivalent, yields the same self-consistent fixed point at the level of accuracy considered here.

The *anisotropic local laws* [32,33] strengthen this to uniform control of quadratic forms:

$$u^T \mathbf{G}_N^{\text{emp}}(z) v \simeq u^T \mathbf{G}_N(z) v, \quad (\text{A6})$$

$$\sup_{\|u\|=\|v\|=1} |u^T (\mathbf{G}_N^{\text{emp}}(z) - \mathbf{G}_N(z)) v| \ll 1, \quad (\text{A7})$$

where $u, v \in \mathbb{R}^N$ for real ensembles are deterministic unit vectors, $\|u\| = \|v\| = 1$, independent of the matrix randomness. The anisotropic local law provides uniform control of the quadratic form $u^T \mathbf{G}_N(z) v$ over all such u and v . The law holds on mesoscopic domains with possible model-dependent modifications near spectral edges or singularities. Physically, η sets a mesoscopic spectral coarse-graining scale: A window of

width η averages over $\sim N \eta \rho(x)$ levels (eigenvalues), making resolvent observables effectively deterministic on that scale. In this sense the imaginary part η acts as a spectral coarse-graining parameter: Increasing η implies a coarser scale and thus averaging of finer-scale fluctuations, while decreasing η moves one toward microscopic resolution. As N grows, self-averaging random observables converge to their deterministic equivalents, with vanishing fluctuations of order $(N\eta)^{-1/2}$.

APPENDIX B: DERIVATION OF THE SELF-CONSISTENT EQUATION FOR WIGNER-TYPE ENSEMBLE

The goal is to derive a site-level Dyson form for the diagonal resolvent. We start from the resolvent identity

$$(z\mathbf{I}_N - \mathbf{X}_N) \mathbf{G}_N^{\text{emp}}(z) = \mathbf{I}_N, \quad z = x + i\eta, \quad \eta > 0. \quad (\text{B1})$$

The diagonal entries satisfy

$$z G_{ii}^{\text{emp}} - X_{ii} G_{ii}^{\text{emp}} - \sum_{k \neq i} X_{ik} G_{ki}^{\text{emp}} = 1, \quad (\text{B2})$$

where for ease of notation we have dropped the label N . In the following subsection we also omit the label “emp” from the empirical resolvent to avoid clutter of notation.

1. Integration by parts

We derive expressions for the expectation values of the different terms in Eq. (B2). We assume Gaussian entries and use integration by parts (IBP), which gives the standard relation, $\mathbb{E}[X_{ab} F(\mathbf{X})] = \text{Var}(X_{ab}) \mathbb{E}[\partial F / \partial X_{ab}]$. The derivative of the resolvent is

$$\frac{\partial \mathbf{G}}{\partial X_{ab}} = \mathbf{G} \frac{\partial \mathbf{X}}{\partial X_{ab}} \mathbf{G}, \quad (\text{B3})$$

and for symmetric $\mathbf{X}$:

$$\frac{\partial \mathbf{X}}{\partial X_{ik}} = \mathbf{E}_{ik} + \mathbf{E}_{ki}, \quad \frac{\partial \mathbf{X}}{\partial X_{ii}} = \mathbf{E}_{ii}, \quad (\text{B4})$$

with $i \neq k$ and $(\mathbf{E}_{ab})_{uv} = \delta_{au} \delta_{bv}$. In particular,

$$\frac{\partial G_{ii}}{\partial X_{ii}} = (\mathbf{G} \mathbf{E}_{ii} \mathbf{G})_{ii} = G_{ii}^2. \quad (\text{B5})$$

Let s_{ik} denote the entrywise variance profile defined in Eq. (12). Taking the expectation value of Eq. (B2) and applying Gaussian IBP yields for the third term on the left-hand side:

$$\mathbb{E}[X_{ik} G_{ki}] = s_{ik} \mathbb{E} \left[\frac{\partial G_{ki}}{\partial X_{ik}} \right] = s_{ik} \mathbb{E}[G_{kk} G_{ii} + G_{ki}^2], \quad (\text{B6})$$

with $i \neq k$. Likewise, for the second term:

$$\mathbb{E}[X_{ii} G_{ii}] = s_{ii} \mathbb{E} \left[\frac{\partial G_{ii}}{\partial X_{ii}} \right] = s_{ii} \mathbb{E}[G_{ii}^2]. \quad (\text{B7})$$

2. Mean-field closure

We now pass from the empirical identity to a deterministic site equation using self-averaging, cf. Appendix A. Taking expectations in Eq. (B2) gives

$$z \mathbb{E}[G_{ii}^{\text{emp}}(z)] - \sum_{k=1}^N \mathbb{E}[X_{ik} G_{ki}^{\text{emp}}(z)] = 1. \quad (\text{B8})$$

On mesoscopic scales ($\eta \geq N^{-1+\varepsilon}$) off the real axis, the anisotropic local law implies self-averaging of diagonal resolvent entries and suppression of off-diagonals, uniformly in i and k :

$$G_{ii}^{\text{emp}}(z) = G_{ii}(z) + O((N\eta)^{-1/2}), \quad (\text{B9})$$

$$G_{ki}^{\text{emp}}(z) = O((N\eta)^{-1/2}), \quad (\text{B10})$$

with $i \neq k$. Hence in Eqs. (B6) and (B7) mixed products factorize up to local-law errors and off-diagonal squares are negligible:

$$\mathbb{E}[G_{kk}^{\text{emp}} G_{ii}^{\text{emp}}] = G_{kk} G_{ii} + O((N\eta)^{-1/2}), \quad (\text{B11})$$

$$\mathbb{E}[(G_{ki}^{\text{emp}})^2] = O((N\eta)^{-1}). \quad (\text{B12})$$

Therefore, for Eqs. (B6) and (B7) we have the approximations:

$$\mathbb{E}[X_{ik} G_{ki}^{\text{emp}}] \simeq s_{ik} G_{kk} G_{ii}, \quad \mathbb{E}[X_{ii} G_{ii}^{\text{emp}}] \simeq s_{ii} G_{ii}^2. \quad (\text{B13})$$

Substitution into the site equation (B8), and dividing by G_{ii} , yields the closed equation for the deterministic equivalent

$$\frac{1}{G_{ii}(z)} = z - \sum_{k=1}^N s_{ik} G_{kk}(z). \quad (\text{B14})$$

Thus the self-energy closes as

$$\Sigma_{ii}(z) := \sum_{k=1}^N s_{ik} G_{kk}(z), \quad \Sigma(z) = \mathbf{S} \mathbf{G}_N(z), \quad (\text{B15})$$

and, with s_{ik} from Eq. (12) we obtain the form

$$\frac{1}{G_i} = z - \frac{\sigma^2}{\gamma} \left(w_i \sum_{k=1}^N w_k G_k + w_i^2 G_i \right), \quad (\text{B16})$$

where we have defined $G_i := G_{ii}$. Using the definition of the coupling and variance-weighted resolvent in Eq. (15), this becomes the on-site Dyson form

$$\frac{1}{G_i} = z - c(w_i u(z) + w_i^2 G_i), \quad (\text{B17})$$

with the physical branch fixed by $\text{Im } G_i(z) < 0$.

APPENDIX C: FIXED-POINT ALGORITHM FOR WIGNER-TYPE ENSEMBLE

We solve for the variance-weighted mean field $u(z)$ by fixed-point iteration (indexed by superscript n) at $z = x + i\eta$ with small $\eta > 0$ [47] (Algorithm 1).

ALGORITHM 1. Fixed-point iteration for $u(z)$ and $G_N(z)$ (Wigner-type ensemble).

Input: $z = x + i\eta$ with $\eta > 0$, tolerance $\varepsilon > 0$, damping $\theta \in (0, 1]$

Init: $w_j \leftarrow j^{-\alpha}$, $c \leftarrow \sigma^2/\gamma$, $W_j \leftarrow c w_j^2$;
 $u^{(0)} \leftarrow -i c_0 \sum_{j=1}^N w_j$ with $c_0 \in (0, 1)$; $n \leftarrow 0$

(Warm start:)

for $j \leftarrow 1$ **to** N **do**

Compute $G_j^{(0)}$ from the quadratic
 $W_j(G_j^{(0)})^2 - (z - c w_j u^{(0)})G_j^{(0)} + 1 = 0$,
choosing the branch with $\text{Im } G_j^{(0)} < 0$;

$u^{(1)} \leftarrow \sum_j w_j G_j^{(0)}$; $n \leftarrow 1$;

repeat

for $j \leftarrow 1$ **to** N **do**

Compute $G_j^{(n)}$ from
 $W_j(G_j^{(n)})^2 - (z - c w_j u^{(n)})G_j^{(n)} + 1 = 0$,
choose $\text{Im } G_j^{(n)} < 0$;

$u^{(n+1)} \leftarrow (1 - \theta) u^{(n)} + \theta \sum_j w_j G_j^{(n)}$;
 $n \leftarrow n + 1$;

until $|u^{(n)} - u^{(n-1)}| < \varepsilon$;

$G_N(z) \leftarrow \frac{1}{N} \sum_j G_j^{(n)}$;
 $\rho_{N,\eta}(x) \leftarrow -\frac{1}{\pi} \text{Im } G_N(x + i\eta)$;

APPENDIX D: VISUALIZING THE EIGENVALUE DENSITIES

We visualize spectra via Cauchy kernel smoothing [12]. For the empirical resolvent

$$\begin{aligned} \rho_{N,\eta}^{\text{emp}}(x) &:= -\frac{1}{\pi} \text{Im } G_N^{\text{emp}}(x + i\eta) \\ &= \frac{1}{\pi} \left[\frac{1}{N} \sum_{j=1}^N \frac{\eta}{(x - \lambda_j)^2 + \eta^2} \right], \end{aligned} \quad (\text{D1})$$

where $\{\lambda_j\}$ are the empirical eigenvalues. For the deterministic equivalent G_N we define the smoothed density

$$\rho_{N,\eta}(x) := -\frac{1}{\pi} \text{Im } G_N(x + i\eta), \quad (\text{D2})$$

which is the same Cauchy smoothing but without an eigenvalue sum. In both cases η is the half-width (resolution) of the Cauchy kernel and fixes the resolution scale: Features narrower than η are averaged out, while broader structures are preserved. For an $N \times N$ matrix, the microscopic mean level spacing is $\Delta(x) \sim 1/(N \rho_{N,\eta}(x))$. Choosing η with $\Delta(x) \ll \eta \ll O(1)$ defines a mesoscopic resolution: The Cauchy kernel averages over $\sim N \eta \rho_{N,\eta}(x)$ eigenvalues, suppressing level-to-level noise while preserving $O(1)$ variations of the density. On this scale, the empirical resolvent concentrates around its deterministic equivalent and the Cauchy-smoothed empirical density tracks deterministic eigenvalue density.

In collapsed curves, the eigenvalues are rescaled by the running normalization $\gamma(N)$ that fixes a spectral moment

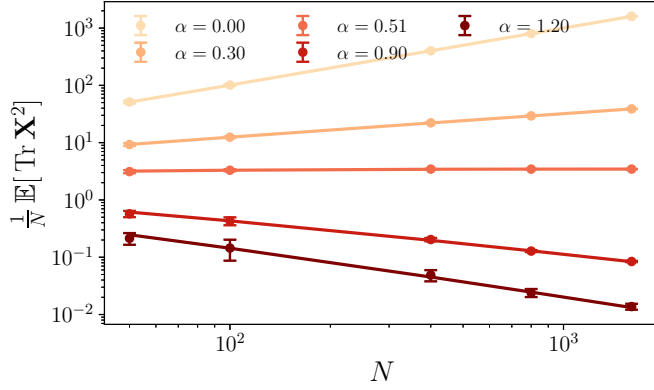


FIG. 5. Finite-size scaling of the second moment for representative exponents α . Solid lines: Exact formula in Eq. (21). Markers: Simulation estimates from eight independent trials. Dashed guides: Predicted slopes.

[see, e.g., Eq. (24)], according to $x = \sqrt{\gamma(N)}y$ [the spectral bandwidth scales like $\sqrt{\gamma(N)}$ in raw units]. To compare raw and collapsed curves with the same spectral resolution, the smoothing width must transform covariantly under this change of units: since $dx = \sqrt{\gamma(N)}dy$, keeping the kernel width fixed in y units implies

$$\eta_{\text{raw}} = \sqrt{\gamma(N)} \eta_{\text{coll}}. \quad (\text{D3})$$

Thus, we choose the same η_{coll} for all values N , and set η_{raw} via the above relation. Using a constant value of η for all raw curves would lead to under- or oversmoothing at large or small N and break comparability across sizes.

On mesoscopic scales the smoothed density $\rho_{N,\eta}$ is strictly positive outside the true support due to the Cauchy kernel. This “tail” is a visualization artifact of the width set by η and decays as $\sim \eta/(x - \text{edge})^2$. We therefore use a single mesoscopic η across sizes and, for fair comparison, apply the same kernel to empirical spectra and to the MDE solution.

APPENDIX E: LARGE- N BEHAVIOR OF THE GENERALIZED HARMONIC SUM

We define the generalized harmonic sum

$$H_N(\theta) := \sum_{j=1}^N j^{-\theta}. \quad (\text{E1})$$

The large- N behavior of the generalized harmonic sum is classical (see, e.g., Chap. 3 in Ref. [48]) and has three regimes:

$$H_N(\theta) = \frac{N^{1-\theta}}{1-\theta} + O(1), \quad 0 \leq \theta < 1, \quad (\text{E2})$$

$$H_N(1) = \log N + \gamma_E + O(N^{-1}), \quad \theta = 1, \quad (\text{E3})$$

$$H_N(\theta) = \zeta(\theta) + O(N^{-(\theta-1)}), \quad \theta > 1, \quad (\text{E4})$$

where $\gamma_E = 0.57721\dots$ is the Euler-Mascheroni constant and ζ is the Riemann zeta function

$$\zeta(\theta) := \sum_{j=1}^{\infty} j^{-\theta}, \quad \theta > 1. \quad (\text{E5})$$

Thus, the dominant scaling of the generalized harmonic number is

$$H_N(\theta) \sim \begin{cases} N^{1-\theta}, & \theta < 1, \\ \log N, & \theta = 1, \\ 1, & \theta > 1. \end{cases} \quad (\text{E6})$$

APPENDIX F: LARGE- N LIMIT AND THE CONTINUUM FIXED POINT

We start from the finite- N self-consistent Eq. (16) and pass to a large- N continuum limit [11]. This is useful because it transforms the discrete fixed-point equation into a closed integral equation with a rank-one self-energy, providing a universal description and analytic control of the spectrum. As before, we choose the running coupling $c(N) = \sigma^2/\gamma(N)$ such that the second moment in Eq. (19) is scale-fixed by the RG condition in Eq. (20).

We reparameterize indices and weights and assume a smooth resolvent profile:

$$y_j := \frac{j}{N}, \quad w_j = j^{-\alpha} = N^{-\alpha} y_j^{-\alpha}, \quad G_j(z) \simeq G(y_j; z). \quad (\text{F1})$$

We note that $y_j = j/N \in (0, 1]$ such that the coordinate y_j becomes continuous in the large- N limit. The profile $G(y; z)$ is the leading-order dependence of the diagonal resolvent on the continuous position y induced by the slowly varying weights w_j in mean-field self-energy form.

First, we treat the term capturing the interaction with the variance-weighted mean field in Eq. (16) which, using the above relations in Eq. (F1), we write as

$$c w_j \sum_{k=1}^N w_k G_k = c(N) N^{1-2\alpha} \left[\frac{1}{N} \sum_{k=1}^N y_j^{-\alpha} y_k^{-\alpha} G_k \right]. \quad (\text{F2})$$

We define the continuum amplitude

$$h_\alpha := \lim_{N \rightarrow \infty} [c(N) N^{1-2\alpha}] = \begin{cases} (1-\alpha)^2, & 0 < \alpha < 1, \\ 0, & \alpha \geq 1. \end{cases} \quad (\text{F3})$$

At $\alpha = 1$, the amplitude h_α vanishes due to log-suppression. Hence the Riemann sum in Eq. (F2) becomes an integral and, overall,

$$c w_j u \xrightarrow{N \rightarrow \infty} h_\alpha \int_0^1 y^{-\alpha} y'^{-\alpha} G(y'; z) dy', \quad (\text{F4})$$

and the rank-one field vanishes in the continuum limit for $\alpha \geq 1$.

Next, we treat the local term in Eq. (16) in an analogous manner and find that it vanishes in the large- N limit. Specifically,

$$c w_j^2 G_j = c(N) N^{1-2\alpha} \frac{y_j^{-2\alpha}}{N} G_j, \quad (\text{F5})$$

and in the large- N limit,

$$c w_j^2 G_j \xrightarrow{N \rightarrow \infty} \frac{h_\alpha}{N} y^{-2\alpha} G(y; z). \quad (\text{F6})$$

Thus, $c w_j^2 G_j = O(N^{-1})$ uniformly in x for $z = x + i\eta$ with fixed $\eta > 0$ (log-corrected at $\alpha = 1$). Since $G(y; z) = O(1)$ off

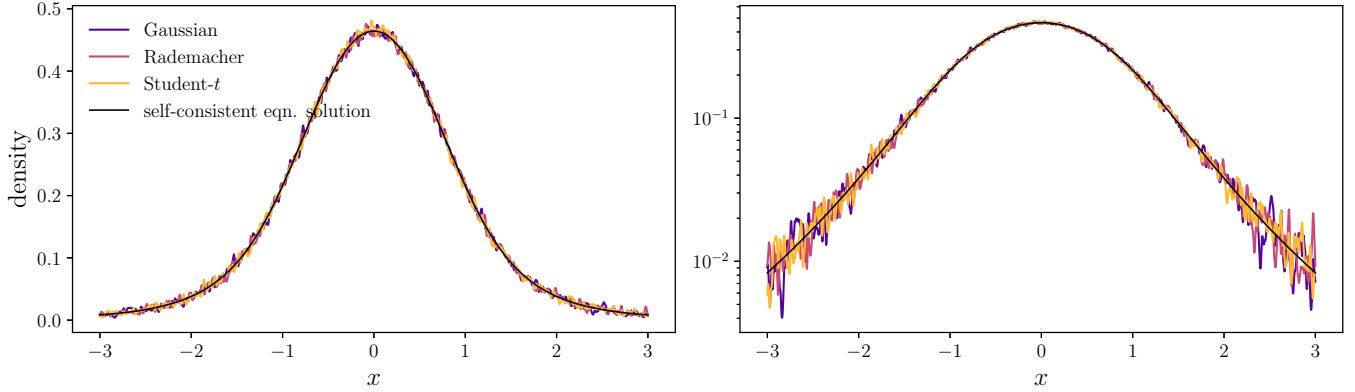


FIG. 6. Entry-law universality. Eigenvalue densities from simulations with Gaussian (purple), Rademacher (magenta), and Student's t with $\nu = 4$ (yellow) matrix entries collapse onto the solution of the self-consistent equation (black) under the RG normalization $\gamma(N)$. Linear-linear scale (left) and log-linear scale (right). Parameters are $N = 600$, $\alpha = 0.60$, and $\eta_{\text{coll}} = 10^{-2}$ and ensembles are simulated with 20 independent trials.

the real axis, the local term contribution vanishes in the large- N limit.

Thus, as the continuum limit form of Eq. (16), we find the Dyson integral equation

$$G(y; z) = \frac{1}{z - \Sigma(y; z)} \quad (\text{F7})$$

with self-energy

$$\Sigma(y; z) := \int_0^1 K(y, y') G(y'; z) dy' \quad (\text{F8})$$

and the kernel:

$$K(y, y') := h_\alpha y^{-\alpha} y'^{-\alpha}. \quad (\text{F9})$$

This form of the Dyson equation is consistent with the abstract Dyson equation point of view of Ref. [34]. Since the kernel factorizes, we can define

$$\Sigma(y; z) = y^{-\alpha} S(z), \quad S(z) := h_\alpha U(z),$$

$$U(z) := \int_0^1 y^{-\alpha} G(y; z) dy. \quad (\text{F10})$$

Then

$$G(y; z) = \frac{1}{z - \Sigma(y; z)} = \frac{1}{z(1 - \phi(z)y^{-\alpha})}, \quad (\text{F11})$$

where in the last equality we have defined the normalized self-energy amplitude

$$\phi(z) := \frac{S(z)}{z} = \frac{h_\alpha}{z} U(z). \quad (\text{F12})$$

Using the definition of $U(z)$, we obtain the scalar fixed point,

$$\phi(z) - \frac{h_\alpha}{z^2} Q_\alpha(\phi(z)) = 0, \quad Q_\alpha(\phi) := \int_0^1 \frac{dy}{y^\alpha - \phi}. \quad (\text{F13})$$

The mean resolvent follows as

$$\bar{G}(z) = \int_0^1 G(y; z) dy = \frac{1}{z} (1 + \phi Q_\alpha(\phi)). \quad (\text{F14})$$

At the marginal value $\alpha = 1/2$ the RG fixed point is non-trivial: The continuum kernel has $h_{1/2} = 1/4$, so the Dyson

self-energy does not vanish and the resolvent is not the free resolvent, $1/z$. On the other hand, for $\alpha > 1$

$$\Sigma = 0, \quad \text{and} \quad \bar{G}(z) = \frac{1}{z}, \quad (\text{F15})$$

i.e., the continuum self-energy vanishes and the bulk of the eigenvalues collapses toward zero, while a finite number of $O(1)$ -outliers (associated with small indices) remain. These outliers are not self-averaging. In Fig. 7, we show a comparison of the continuum limit form and the eigenvalue density obtained from simulated ensembles and find a near-perfect match.

Within the class of centered, independent entries with finite variance and standard moment bounds, the large- N deterministic limit supplied by the matrix Dyson equation depends on the model only through its variance operator; see, e.g., Refs. [31–33]. Under our RG normalization that fixes m_2 and for the power-law weights, the variance kernel factorizes, so the continuum self-energy is rank one and its amplitude is h_α . Consequently, the bulk limiting density depends on α only through h_α and is insensitive to higher cumulants at leading order on mesoscopic scales.

We note that the above statement is restricted to mean-field regimes with finite variance. Subleading corrections, edge behavior, and finite- N fluctuations can be model dependent, and heavy tailed (infinite variance), sparse, or banded ensembles also lie outside this scope. More generally, (nonfactorizable) variance profiles would lead to higher-rank kernels and different fixed points.

APPENDIX G: DERIVATION OF THE SELF-CONSISTENT EQUATION FOR WISHART-TYPE ENSEMBLE

We consider $\mathbf{Z} \in \mathbb{R}^{T \times N}$ with independent and identically distributed centered unit-variance entries, $\mathbf{D} = \text{diag}(v_1, \dots, v_T)$ with $v_t = t^{-2\alpha}$, and define

$$\mathbf{Y} = \mathbf{D}^{1/2} \mathbf{Z}, \quad \mathbf{C} = \frac{1}{\gamma^2} \mathbf{Y}^T \mathbf{Y} = \frac{1}{\gamma^2} \mathbf{Z}^T \mathbf{D} \mathbf{Z}. \quad (\text{G1})$$

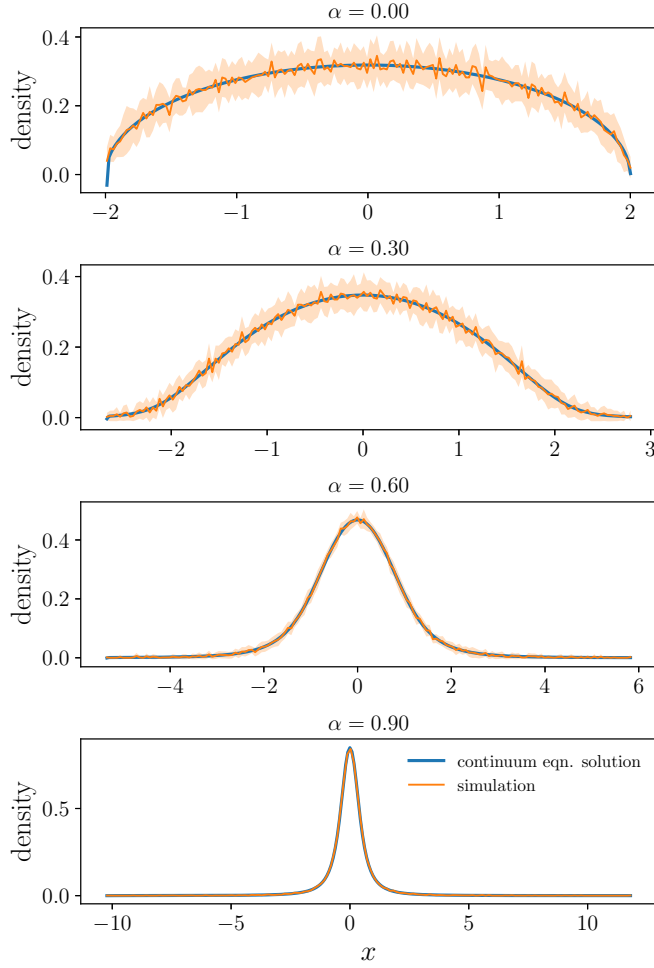


FIG. 7. Comparison of continuum equation solution and simulation. Plots show the continuum limit solution (blue) and the simulation with mean and one standard deviation band (orange) for $\alpha = 0, 0.3, 0.6$, and 0.9 . Simulation curves are obtained with $N = 600$ and 20 independent trials.

1. Row decomposition and rank-one update

We write the t th row of $\mathbf{Z}$ as $z_t \in \mathbb{R}^N$ and set

$$b_t := \frac{\sqrt{v_t}}{\gamma} z_t \sim \mathcal{N}(0, (v_t/\gamma^2) \mathbf{I}_N), \quad (\text{G2})$$

so that the covariance matrix is a sum of rank-one outer products:

$$\mathbf{C} = \sum_{t=1}^T b_t b_t^T. \quad (\text{G3})$$

Starting from the resolvent identity

$$(z \mathbf{I}_N - \mathbf{C}) \mathbf{G}_N(z) = \mathbf{I}_N, \quad (\text{G4})$$

and taking the normalized trace, we obtain

$$z G_N(z) - \frac{1}{N} \sum_{t=1}^T b_t^T \mathbf{G}_N(z) b_t = 1, \quad (\text{G5})$$

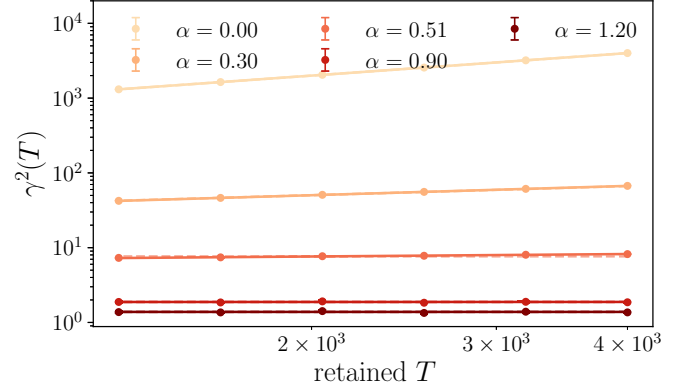


FIG. 8. Normalization flow with decimation. Decimation flow of $\gamma^2(T)$ with retained size T after removing a fixed fraction of largest indices at each step, while the first moment is kept at unity. Solid lines: Exact prediction. Markers with error bars show the simulation mean and standard deviation computed from six independent trials.

with $G_N(z) := \frac{1}{N} \text{Tr} \mathbf{G}_N(z)$. We remove the t th row from $\mathbf{Z}$ and define the *leave-one-out* quantities

$$\mathbf{C}^{(t)} := \sum_{s \neq t} b_s b_s^T, \quad \mathbf{G}_N^{(t)}(z) := (z \mathbf{I}_N - \mathbf{C}^{(t)})^{-1}. \quad (\text{G6})$$

By the Sherman-Morrison formula [12,49]:

$$(\mathbf{A} + uv^T)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1} uv^T \mathbf{A}^{-1}}{1 + v^T \mathbf{A}^{-1} u} \quad (\text{G7})$$

with $\mathbf{A}^{-1} = \mathbf{G}_N^{(t)}$ and $uv^T = \mathbf{G}_N^{-1} - \mathbf{G}_N^{(t)-1} = -b_t b_t^T$, we obtain the rank-one update of the resolvent and its trace:

$$\mathbf{G}_N(z) = \mathbf{G}_N^{(t)}(z) + \frac{\mathbf{G}_N^{(t)}(z) b_t b_t^T \mathbf{G}_N^{(t)}(z)}{1 - b_t^T \mathbf{G}_N^{(t)}(z) b_t} \quad (\text{G8})$$

$$\Leftrightarrow b_t^T \mathbf{G}_N(z) b_t = b_t^T \mathbf{G}_N^{(t)}(z) b_t + \frac{[b_t^T \mathbf{G}_N^{(t)}(z) b_t]^2}{1 - b_t^T \mathbf{G}_N^{(t)}(z) b_t}, \quad (\text{G9})$$

where we have used Eq. (G3). Hence, we obtain

$$b_t^T \mathbf{G}_N(z) b_t = \frac{b_t^T \mathbf{G}_N^{(t)}(z) b_t}{1 - b_t^T \mathbf{G}_N^{(t)}(z) b_t}. \quad (\text{G10})$$

2. Conditional Gaussian averaging

The random objects we average are the quadratic forms in the numerator and denominator on the right-hand side of Eq. (G10). We condition on the *leave-one-out* (minor) resolvent $\mathbf{G}_N^{(t)}$ [i.e., treat $\mathbf{C}^{(t)}$ as fixed], and note that b_t is independent of $\mathbf{C}^{(t)}$. Since $b_t \sim \mathcal{N}(0, (v_t/\gamma^2) \mathbf{I})$ and $\mathbf{G}_N^{(t)}$ is fixed under the conditional expectation, the standard Gaussian

contraction gives

$$\mathbb{E}_{b_t} [b_t^T \mathbf{G}_N^{(t)} b_t \mid \mathbf{G}_N^{(t)}] = \text{Tr}(\text{Cov}(b_t) \mathbf{G}_N^{(t)}) = \frac{v_t}{\gamma^2} \text{Tr} \mathbf{G}_N^{(t)}(z). \quad (\text{G11})$$

Thus, under the conditional expectation, the quadratic form reduces to a trace against the leave-one-out resolvent, and we have

$$\mathbb{E}_{b_t} [b_t^T \mathbf{G}_N(z) b_t \mid \mathbf{G}_N^{(t)}] = \frac{\frac{v_t}{\gamma^2} \text{Tr} \mathbf{G}_N^{(t)}(z)}{1 - \frac{v_t}{\gamma^2} \text{Tr} \mathbf{G}_N^{(t)}(z)}. \quad (\text{G12})$$

3. Mean-field replacement and equation closure

The anisotropic local law (cf. Appendix A and Refs. [32,33]), asserts that, uniformly for $z = x + i\eta$ with $\eta > 0$, leave-one-out resolvents $\mathbf{G}^{(t)}$ and full resolvents $\mathbf{G}_N$ differ by subleading corrections in quadratic forms and traces:

$$\frac{1}{N} \text{Tr} \mathbf{G}_N^{(t)}(z) = G_N(z) + O((N\eta)^{-1/2}). \quad (\text{G13})$$

Substituting into the normalized trace identity yields, up to vanishing errors

$$z G_N(z) - \frac{1}{N} \sum_{t=1}^T \frac{\left(\frac{v_t}{\gamma^2}\right) N G_N(z)}{1 - \left(\frac{v_t}{\gamma^2}\right) N G_N(z)} = 1, \quad (\text{G14})$$

or, equivalently, the closed fixed point,

$$G_N(z) = \frac{1}{z - \Sigma_N(z)}, \quad (\text{G15})$$

$$\Sigma_N(z) = \frac{1}{\gamma^2} \sum_{t=1}^T \frac{v_t}{1 - (N/\gamma^2) v_t G_N(z)}. \quad (\text{G16})$$

These are the finite- N fixed-point equations for the deterministic resolvent. Empirical resolvents concentrate around it with local-law accuracy. The correction term is of order $(N\eta)^{-1/2}$. The same fixed-point equations can also be derived from the Silverstein-Dozier framework for sample covariances [10,36,37].

4. Fixed-point algorithm for the Wishart-type ensemble

We solve the self-consistent equations for the Wishart-type ensemble Eqs. (37) and (38), with the following fixed-point algorithm [47] (Algorithm 2).

ALGORITHM 2. Fixed-point iteration for $G_N(z)$ (Wishart-type ensemble).

Input: $z = x + i\eta$ with $\eta > 0$; α, N, T, γ^2 ;
 tolerance $\varepsilon > 0$; damping $\theta \in (0, 1]$;
 stabilization $\varepsilon_{\text{stab}} > 0$
Init: $v_t \leftarrow t^{-2\alpha}$ for $t = 1, \dots, T$; $G^{(0)} \leftarrow 1/z$;
 $n \leftarrow 0$
 (Warm start:) Compute
 $\Sigma^{(0)} \leftarrow \frac{1}{\gamma^2} \sum_{t=1}^T \frac{v_t}{1 - (N/\gamma^2) v_t G^{(0)} + i\varepsilon_{\text{stab}}}$; set
 $G^{(1)} \leftarrow \frac{1}{z - \Sigma^{(0)}}$; set $n \leftarrow 1$;
repeat
 Compute
 $\Sigma^{(n)} \leftarrow \frac{1}{\gamma^2} \sum_{t=1}^T \frac{v_t}{1 - (N/\gamma^2) v_t G^{(n)} + i\varepsilon_{\text{stab}}}$;
 Set $\tilde{G}^{(n+1)} \leftarrow \frac{1}{z - \Sigma^{(n)}}$;
 $G^{(n+1)} \leftarrow (1 - \theta) G^{(n)} + \theta \tilde{G}^{(n+1)}$;
 $n \leftarrow n + 1$;
until $|G^{(n)} - G^{(n-1)}| < \varepsilon$;
 $G_N(z) \leftarrow G^{(n)}$; $\rho_{N,\eta}(x) \leftarrow -\frac{1}{\pi} \text{Im} G_N(z)$;

APPENDIX H: LINEARIZED RG FLOW IN THE SATURATING REGIME ($\alpha > 1/2$)

In the main text we define the RG Beta function in terms of the running coupling $g(T) := \gamma^{-2}(T)$, cf. Eq. (42). For $\alpha > 1/2$ the renormalization condition implies that $\gamma^2(T)$ saturates and hence $g(T)$ approaches a nonzero fixed point g_* . In this situation the Beta function $\beta_{\text{RG}}(g) = d \log g / d \log T$ necessarily tends to zero, and it no longer exposes the *rate* at which the coupling approaches its fixed point. To quantify this approach we linearize the flow around g_* and extract the corresponding decay exponent.

In the Wishart-type model, the running normalization $\gamma^2(T)$ is fixed by the renormalization condition on the first moment

$$m_1 = \frac{1}{\gamma^2(T)} \sum_{t=1}^T t^{-2\alpha} \stackrel{!}{=} m_1^*, \quad (\text{H1})$$

so that, up to the fixed constant m_1^* factor

$$\gamma^2(T) \propto \sum_{t=1}^T t^{-2\alpha}. \quad (\text{H2})$$

For $\alpha > 1/2$ the series converges and $\gamma^2(T)$ saturates as $T \rightarrow \infty$.

1. RG fixed-point approach

We define the fixed-point normalization and coupling,

$$\gamma_*^2 := \lim_{T \rightarrow \infty} \gamma^2(T) \propto \zeta(2\alpha), \quad g_* := \gamma_*^{-2}, \quad (\text{H3})$$

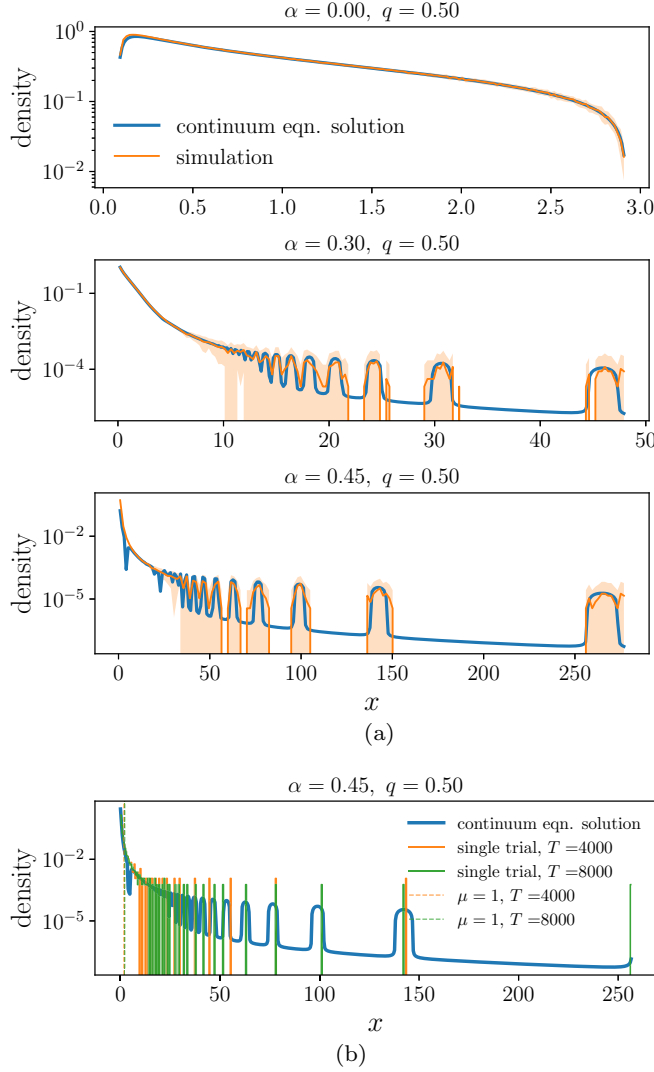


FIG. 9. Comparison of continuum equation solution and simulation. (a) Plots show the continuum limit solution (blue) and the simulation with mean and one standard deviation band (orange) for $\alpha = 0, 0.3$, and 0.45 . Simulation curves are obtained with $N = 4000$ and 40 independent trials. (b) Continuum limit solution for $\alpha = 0.45$ (blue) and single-trial simulations for $T = 4000$ (orange) and 8000 (green). Vertical lines show the sampling threshold beyond which single-trial smoothing is sampling-limited.

and measure the approach to the fixed point via the deviation in γ^2 ,

$$\delta(T) := \gamma_\star^2 - \gamma^2(T). \quad (\text{H4})$$

Using a tail estimate for the convergent series, we obtain the asymptotic scaling of the deviation

$$\delta(T) \propto \sum_{t=T+1}^{\infty} t^{-2\alpha} \sim \int_T^{\infty} t'^{-2\alpha} dt' = \frac{T^{1-2\alpha}}{2\alpha-1}. \quad (\text{H5})$$

Hence, the deviation decays as a power law,

$$\delta(T) \propto T^{-(2\alpha-1)}, \quad (\text{H6})$$

with exponent $2\alpha - 1 > 0$.

2. Linearized Beta function

A natural linearized Beta function is the logarithmic derivative of the deviation:

$$\beta_\delta := \frac{d \log \delta}{ds} = \frac{d \log \delta}{d \log T}. \quad (\text{H7})$$

Using Eq. (H5) we find

$$\beta_\delta \rightarrow 1 - 2\alpha < 0, \quad \alpha > 1/2. \quad (\text{H8})$$

Thus, in the saturating regime the flow is irrelevant in the RG sense: As the scale T increases the deviation from the fixed point decays as a power law, and under decimation (reducing the scale) the deviation grows with exponent $2\alpha - 1$. Equivalently, we may linearize the coupling itself. Since $g(T) = \gamma^{-2}(T)$ and $\gamma^2(T) = \gamma_\star^2 - \delta(T)$, we have

$$g(T) = \frac{1}{\gamma_\star^2 - \delta(T)} = g_\star + \frac{\delta(T)}{\gamma_\star^4} + O(\delta(T)^2), \quad (\text{H9})$$

so that $g(T) - g_\star \propto \delta(T)$ and therefore $g(T) - g_\star \propto T^{1-2\alpha}$ with the same exponent.

Thus, the linearized Beta function β_δ characterizes the approach to the fixed point through the deviation $\delta(T)$ [or, equivalently, $g(T) - g_\star$] and recovers the negative exponent $1 - 2\alpha$ in the saturating regime.

APPENDIX I: LARGE- N LIMIT AND THE CONTINUUM FIXED POINT

We take the large- N limit with fixed aspect ratio $q := N/T$ and pass from row indices $t = 1, \dots, T$ to the variable $y_t := t/T \in (0, 1]$ which becomes the continuous variable y for large T . For the row-heterogeneity, $v_t = t^{-2\alpha}$, we have

$$v_t = (Ty_t)^{-2\alpha} = T^{-2\alpha} y_t^{-2\alpha} \quad (\text{I1})$$

and further define the scaled row weights

$$\tilde{v}_t := \frac{T}{\gamma^2(T)} v_t = \frac{T^{1-2\alpha}}{\gamma^2(T)} y_t^{-2\alpha}. \quad (\text{I2})$$

In the continuum limit where $y_t \rightarrow y$, the analog expression is

$$\tilde{v}_T(y) := \frac{T^{1-2\alpha}}{\gamma^2(T)} y^{-2\alpha}. \quad (\text{I3})$$

For $\alpha \neq 1/2$ the pointwise limit exists for fixed $y \in (0, 1]$:

$$\tilde{v}(y) := \lim_{T \rightarrow \infty} \tilde{v}_T(y). \quad (\text{I4})$$

Starting from the self-consistent Eqs. (37) and (38), the sum giving the self-energy becomes an integral and the continuum closure takes the form

$$\bar{G}(z) = \frac{1}{z - \Sigma(z)}, \quad (\text{I5})$$

$$\Sigma(z) = \int_0^1 \frac{\tilde{v}(y)}{1 - q \tilde{v}(y) \bar{G}(z)} dy, \quad (\text{I6})$$

with the physical branch chosen by $\text{Im} \bar{G}(z) < 0$. Next, we treat the different regimes of α .

For $0 \leq \alpha < 1/2$ we have $\gamma^2(T) \sim \frac{T^{1-2\alpha}}{1-2\alpha}$, and hence

$$\tilde{v}(y) = (1 - 2\alpha) y^{-2\alpha}. \quad (\text{I7})$$

As a sanity check, for $\alpha = 0$ we have $\tilde{v} = 1$ and $\Sigma = 1/(1 - q\bar{G})$. Then Eq. (16) reproduces the self-consistent equation for the Marčenko-Pastur resolvent [12].

In the large- N continuum limit, the integral equation is derived at fixed $x = O(1)$ and for $y = t/T$ bounded away from 0. For $\alpha > 1/2$ one has $\tilde{v}(y) = 0$ for any fixed $y \in (0, 1]$, hence $\Sigma(z) = 0$ and $\bar{G}(z) = 1/z$ (the free resolvent) on the $O(1)$ spectral scale. The sector $t = O(1)$, or equivalently $y \sim 1/T$, forms a boundary layer that is not captured by the continuum limit and produces the outliers of size $O(N)$ that keep m_1 fixed.

Finally, the marginal case $\alpha = 1/2$ requires keeping the logarithmic T dependence. Since $\gamma^2(T) \sim \log T$, we have

$$\tilde{v}_t = \frac{1}{\log T} y_t^{-1}. \quad (18)$$

In particular, the continuum closure retains the lower cutoff $y_{\min} = 1/T$ inherited from $t \geq 1$ and reads

$$G_T(z) = \frac{1}{z - \Sigma_T(z)}, \quad (19)$$

$$\Sigma_T(z) = \int_{1/T}^1 \frac{\tilde{v}_T(y)}{1 - q\tilde{v}_T(y)G_T(z)} dy, \quad \tilde{v}_T(y) = \frac{1}{\log T} y^{-1}. \quad (110)$$

Thus, $\alpha = 1/2$ is marginal with logarithmic corrections. Although $\tilde{v}_T(y) \rightarrow 0$ pointwise for fixed $y \in (0, 1]$, the self-energy integral remains cutoff sensitive and can contribute at $O(1)$ through $y_{\min} = 1/T$ and the slow $1/\log T$ running. As a consequence, in the strict limit most eigenvalues accumulate near zero, while a vanishing fraction of eigenvalues remain away from zero on an $O(1)$ scale so that the mean level stays fixed.

In Fig. 9(a) we show a comparison of eigenvalue densities from simulated data with solutions of the large- N continuum integral equation for $\alpha < 1/2$. Overall, we find good agreement, but observe that the solutions of the integral equations tend to show some numerical instabilities leading to small local mismatches. The simulated eigenvalue bands at large x visible in the second and third subplots of Fig. 9(a) are produced by resonances of heavy rows with weights $v_t = t^{-2\alpha}$. Each band corresponds to a fixed row fraction $y = t/T$ and its x position is T independent [set by $1 - q\tilde{v}(y)\bar{G}(z) \approx 0$, cf. Eq. (16)], while the number of visible bands increases with T as the y grid refines. Each band carries a mass $O(1/T)$, and their Cauchy-smeared aggregate converges to the continuum bulk envelope (blue). In Fig. 9(b) we show that continuum eigenvalue bands are “underfilled” by single-trial simulation data when $\mu(x) := N\eta\rho(x) \lesssim 1$. Increasing η , averaging over independent trials, or making N very large fills the envelope without changing the bulk shape.

-
- [1] J. P. Cunningham and B. M. Yu, Dimensionality reduction for large-scale neural recordings, *Nat. Neurosci.* **17**, 1500 (2014).
 - [2] M. B. Eisen, P. T. Spellman, P. O. Brown, and D. Botstein, Cluster analysis and display of genome-wide expression patterns, *Proc. Natl. Acad. Sci. USA* **95**, 14863 (1998).
 - [3] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
 - [4] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74**, 47 (2002).
 - [5] W. Bialek, A. Cavagna, I. Giardina, T. Mora, E. Silvestri, M. Viale, and A. M. Walczak, Statistical mechanics for natural flocks of birds, *Proc. Natl. Acad. Sci. USA* **109**, 4786 (2012).
 - [6] E. P. Wigner, On the distribution of the roots of certain symmetric matrices, *Ann. Math.* **67**, 325 (1958).
 - [7] V. A. Marčenko and L. A. Pastur, Distribution of eigenvalues for some sets of random matrices, *Math. USSR-Sbornik* **1**, 457 (1967).
 - [8] H. Sompolinsky, A. Crisanti, and H. J. Sommers, Chaos in random neural networks, *Phys. Rev. Lett.* **61**, 259 (1988).
 - [9] F. Mastrogiuseppe and S. Ostojic, Linking connectivity, dynamics, and computations in low-rank recurrent neural networks, *Neuron* **99**, 609 (2018).
 - [10] Z. D. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. (Springer, Berlin, 2010).
 - [11] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices*, Cambridge Studies in Advanced Mathematics, Vol. 118 (Cambridge University Press, Cambridge, UK, 2010).
 - [12] M. Potters and J.-P. Bouchaud, *A First Course in Random Matrix Theory: For Physicists, Engineers and Data Scientists* (Cambridge University Press, Cambridge, UK, 2020).
 - [13] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications* (Cambridge University Press, Cambridge, UK, 2011).
 - [14] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters, Noise dressing of financial correlation matrices, *Phys. Rev. Lett.* **83**, 1467 (1999).
 - [15] P. Fleig and I. Nemenman, Statistical properties of large data sets with linear latent features, *Phys. Rev. E* **106**, 014102 (2022).
 - [16] L. Zdeborová and F. Krzakala, Statistical physics of inference: Thresholds and algorithms, *Adv. Phys.* **65**, 453 (2016).
 - [17] T. Lesieur, L. Miolane, M. Lelarge, F. Krzakala, and L. Zdeborová, in *Proceedings of the IEEE International Symposium on Information Theory (ISIT'17)* (IEEE, Los Alamitos, CA, 2017), pp. 511–515.
 - [18] C. A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel, *Commun. Math. Phys.* **159**, 151 (1994).
 - [19] L. P. Kadanoff, Scaling laws for Ising models near T_c , *Phys. Phys. Fiz.* **2**, 263 (1966).
 - [20] J. B. Kogut, An introduction to lattice gauge theory and spin systems, *Rev. Mod. Phys.* **51**, 659 (1979).
 - [21] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group* (CRC Press, Boca Raton, FL, 2018).
 - [22] K. G. Wilson and J. Kogut, The renormalization group and the ϵ expansion, *Phys. Rep.* **12**, 75 (1974).
 - [23] A. Atanasov, J. A. Zavattone-Veth, and C. Pehlevan, Scaling and renormalization in high-dimensional regression, [arXiv:2405.00592](https://arxiv.org/abs/2405.00592).
 - [24] G. Buzsáki and K. Mizuseki, The log-dynamic brain: how skewed distributions affect network operations, *Nat. Rev. Neurosci.* **15**, 264 (2014).

- [25] K. Rajan and L. F. Abbott, Eigenvalue spectra of random matrices for neural networks, *Phys. Rev. Lett.* **97**, 188104 (2006).
- [26] N. Masuda, M. A. Porter, and R. Lambiotte, Random walks and diffusion on networks, *Phys. Rep.* **716-717**, 1 (2017).
- [27] O. Alter, P. O. Brown, and D. Botstein, Singular value decomposition for genome-wide expression data processing and modeling, *Proc. Natl. Acad. Sci. USA* **97**, 10101 (2000).
- [28] M. D. Luecken and F. J. Theis, Current best practices in single-cell RNA-seq analysis: A tutorial, *Mol. Syst. Biol.* **15**, e8746 (2019).
- [29] C. G. Callan, Broken scale invariance in scalar field theory, *Phys. Rev. D* **2**, 1541 (1970).
- [30] K. Symanzik, Small distance behaviour in field theory and power counting, *Commun. Math. Phys.* **18**, 227 (1970).
- [31] O. H. Ajanki, L. Erdős, and T. Krüger, Universality for general Wigner-type matrices, *Probab. Theory Relat. Fields* **169**, 667 (2017).
- [32] O. H. Ajanki, L. Erdős, and T. Krüger, Stability of the matrix Dyson equation and random matrices with correlations, *Probab. Theory Relat. Fields* **173**, 293 (2019).
- [33] A. Knowles and J. Yin, Anisotropic local laws for random matrices, *Probab. Theory Relat. Fields* **169**, 257 (2017).
- [34] J. Alt, T. Krüger, and L. Erdős, The Dyson equation with linear self-energy: Spectral bands, edges and cusps, *Document. Math.* **25**, 1421 (2020).
- [35] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond*, World Scientific Lecture Notes in Physics, Vol. 9 (World Scientific, Singapore, 1987).
- [36] J. W. Silverstein and Z. D. Bai, On the empirical distribution of eigenvalues of a class of large dimensional random matrices, *J. Multivar. Anal.* **54**, 175 (1995).
- [37] R. B. Dozier and J. W. Silverstein, On the empirical distribution of eigenvalues of large dimensional information-plus-noise-type matrices, *J. Multivar. Anal.* **98**, 678 (2007).
- [38] R. Olfati-Saber, J. A. Fax, and R. M. Murray, Consensus and cooperation in networked multi-agent systems, *Proc. IEEE* **95**, 215 (2007).
- [39] F. Chung, L. Lu, and V. Vu, Spectra of random graphs with given expected degrees, *Proc. Natl. Acad. Sci. USA* **100**, 6313 (2003).
- [40] R. M. May, Will a large complex system be stable? *Nature (London)* **238**, 413 (1972).
- [41] P. Mehta and D. J. Schwab, An exact mapping between the variational renormalization group and deep learning, [arXiv:1410.3831](https://arxiv.org/abs/1410.3831).
- [42] M. Koch-Janusz and Z. Ringel, Mutual information, neural networks and the renormalization group, *Nat. Phys.* **14**, 578 (2018).
- [43] S. Bradde and W. Bialek, PCA meets RG, *J. Stat. Phys.* **167**, 462 (2017).
- [44] L. Meshulam, J. L. Gauthier, C. D. Brody, D. W. Tank, and W. Bialek, Coarse graining, fixed points, and scaling in a large population of neurons, *Phys. Rev. Lett.* **123**, 178103 (2019).
- [45] G. Nicoletti, S. Suweis, and A. Maritan, Scaling and criticality in a phenomenological renormalization group, *Phys. Rev. Res.* **2**, 023144 (2020).
- [46] A. Nguyen, D. J. Schwab, and V. Ngampruetikorn, Data coarse graining can improve model performance, [arXiv:2509.14498](https://arxiv.org/abs/2509.14498).
- [47] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (SIAM, Philadelphia, 2000).
- [48] T. M. Apostol, *Introduction to Analytic Number Theory*, Undergraduate Texts in Mathematics (Springer, New York, 1976).
- [49] M. S. Bartlett, An inverse matrix adjustment arising in discriminant analysis, *Ann. Math. Stat.* **22**, 107 (1951).