Properties of networks with partially structured and partially random connectivity

Yashar Ahmadian,[1,2,*] Francesco Fumarola,[1] and Kenneth D. Miller[1,2]

[1]*Center for Theoretical Neuroscience, Department of Neuroscience,*
[2]*Swartz Program in Theoretical Neuroscience, and Kavli Institute for Brain Science, College of Physicians and Surgeons,*
Columbia University, New York, New York 10032, USA

Networks studied in many disciplines, including neuroscience and mathematical biology, have connectivity that may be stochastic about some underlying mean connectivity represented by a non-normal matrix. Furthermore, the stochasticity may not be independent and identically distributed (iid) across elements of the connectivity matrix. More generally, the problem of understanding the behavior of stochastic matrices with nontrivial mean structure and correlations arises in many settings. We address this by characterizing large random $N \times N$ matrices of the form $A = M + LJR$, where M, L, and R are arbitrary deterministic matrices and J is a random matrix of zero-mean iid elements. M can be non-normal, and L and R allow correlations that have separable dependence on row and column indices. We first provide a general formula for the eigenvalue density of A. For A non-normal, the eigenvalues do not suffice to specify the dynamics induced by A, so we also provide general formulas for the transient evolution of the magnitude of activity and frequency power spectrum in an N-dimensional linear dynamical system with a coupling matrix given by A. These quantities can also be thought of as characterizing the stability and the magnitude of the linear response of a nonlinear network to small perturbations about a fixed point. We derive these formulas and work them out analytically for some examples of M, L, and R motivated by neurobiological models. We also argue that the persistence as $N \to \infty$ of a finite number of randomly distributed outlying eigenvalues outside the support of the eigenvalue density of A, as previously observed, arises in regions of the complex plane Ω where there are nonzero singular values of $L^{-1}(z\mathbf{1} - M)R^{-1}$ (for $z \in \Omega$) that vanish as $N \to \infty$. When such singular values do not exist and L and R are equal to the identity, there is a correspondence in the normalized Frobenius norm (but not in the operator norm) between the support of the spectrum of A for J of norm σ and the σ pseudospectrum of M.

1. INTRODUCTION

Knowledge of the statistics of eigenvalues and eigenvectors of random matrices has applications in the modeling of phenomena relevant to a wide range of disciplines [1–3]. In many applications, however, the matrices of interest are not entirely random, but feature substantial deterministic structure. Furthermore, this structure, as well as the disorder on top of it, are, in general, described by non-normal matrices.

In neuroscience, for example, connections between neurons typically have restricted spatial range and show specificity with respect to neuronal type, location, and response properties. Experience-based synaptic plasticity, which underlies learning and memory, naturally gives rise to synaptic connectivity matrices that encode aspects of the statistical structure of the sensory environment, while containing significant randomness partly due to the inherent stochasticity of particular histories of sensory experience. Another simple example of structured neural connectivity is due to what is known as Dale's principle [4–6]: Neurons come in two main types, excitatory and inhibitory. This empirical principle imposes a certain structure on the synaptic connectivity matrix, forcing all elements in each column of the matrix, describing the synaptic projections of a certain neuron, to have the same sign. Particularly when the typical weight magnitude is much larger than typical differences between the magnitudes of excitatory and inhibitory weights, such a matrix can be extremely non-normal by some measures, much more so than a fully random matrix [7]. Similarly, biological knowledge imparts a great deal of structure to models of biochemical [8–11] or ecological networks [12–16], and matrices characterizing such interactions are typically non-normal. Yet our knowledge of connectivity or interactions is, at best, probabilistic. To describe realistic biological behavior, we must generalize from the behavior of a fixed, regular connectivity to the expected behavior of a typical sample from an appropriate connectivity ensemble.

Furthermore, non-normality can lead to important dynamical properties not seen for normal matrices [17]. In general, networks with a recurrent connectivity pattern described by a non-normal matrix can be described as having a hidden feedforward connectivity structure between orthogonal activity patterns, each of which can also excite or inhibit itself [7,18,19]. In neural networks such hidden feedforward connectivity arises from the natural separation of excitatory and inhibitory neurons, yielding so-called "balanced amplification" of patterns of activity without any dynamical slowing [7]. Underlying this is the phenomenon of "transient amplification": a small perturbation from a fixed point of a stable system with non-normal connectivity can lead to a large transient response over finite time [17]. Transient amplification also yields unexpected results in ecological networks [20–22] and has been conjectured to play a key role in many biochemical systems [23]. Networks that yield long hidden feedforward chains can also generate long time scales and provide a substrate for working memory [18,19]. Systems with non-normal connectivity can also exhibit pseudoresonance

———————
*Corresponding author: ya2005@columbia.edu

frequencies in their power spectrum at which the system responds strongly to external inputs, even though the external frequency is not close to any of the system's natural frequencies as determined by its eigenvalues [17].

While Hermitian random matrices and fully random non-Hermitian matrices with zero-mean, independent and identically distributed (iid) elements have been widely studied, there is a shortage of results on quantities of interest for non-normal matrices that fall between the two extremes of fully random or fully deterministic. A natural departure from a non-normal deterministic structure, described by a connectivity matrix M, is to additively perturb it with a fully random matrix J with zero-mean, iid elements. In many important examples, however, the strength of disorder (deviations from the mean structure) is not uniform and itself has some structure (e.g., for each connection it can depend on the types of the connected nodes or neurons). Moreover, the deviations of the strength of different connections or interactions from their average need not be independent. Hence, it is important to move beyond a simple iid deviation from the mean structure. Here we study ensembles of large $N \times N$ random matrices of the form $A = M + LJR$, where M, L, and R are arbitrary (M) or arbitrary invertible (L and R) deterministic matrices that are, in general, non-normal, and J is a completely random matrix with zero-mean iid elements of variance $1/N$. The matrix M is thus the average of A and describes average connectivity. Note that when L and R are diagonal, they specify variances that depend separably on the row and column of A, while when they are not diagonal, the elements of A are not statistically independent. As we show in Sec. II C 3, this form arises naturally, for example, in linearizations of dynamical systems involving simple classes of nonlinearities. This type of ensemble is also natural from the random matrix theory viewpoint, as it describes a classical fully random ensemble—an iid random matrix J—modified by the two basic algebraic operations of matrix multiplication and addition.

We study the eigenvalue distribution of such matrices, but also directly study the dynamics of a linear system of differential equations governed by such matrices. Specifically, for matrices of the above type, using the Feynman diagram technique in the large N limit (we follow the particular version of this method developed by Refs. [24,25]), we have derived a general formula for the density of their eigenvalues in the complex plane, which generalizes the well-known circular law for fully random matrices [26–30]. It also generalizes a result [31] obtained for the case where L and R are scalar multiples of $\mathbf{1}$, the N-dimensional identity matrix (the same result was obtained in [32] using the methods and language of free probability theory; the eigenvalue density for the case $L \propto R \propto \mathbf{1}$ and a *normal M* was also calculated in Ref. [24] in the limit $N \to \infty$, and that result was extended to finite N in Ref. [33]). Apart from generalization to arbitrary invertible L and R, we also provide a correct regularizing procedure for finding the support of the eigenvalue density in the limit $N \to \infty$ in certain highly non-normal cases of M; the naive interpretation of the formulas fails in these cases, which were not previously discussed. Furthermore, with the aim of studying dynamical signatures of non-normal connectivity, we focused on the dynamics directly, deriving general formulas for the magnitude of the response of the

system to a δ function pulse of input (which provides a measure of the time course of potential transient amplification), as well as the frequency power spectrum of the system's response to external time-dependent inputs.

These general results are presented in the next section. There we also present the explicit results of analytical or numerical calculations based on these general formulas for some specific examples of M, L, and R. Sections III and IV contain the detailed derivations of our general formulas for the eigenvalue density and the response magnitude formulas, respectively. Section V contains the detailed analytical calculations of these quantities for the specific examples presented in Sec. II, based on the general formulas. We conclude the paper in Sec. VI.

2. SUMMARY OF RESULTS

We study ensembles of large $N \times N$ random matrices of the form

$$A = M + LJR, \tag{2.1}$$

where M, L, and R are arbitrary (M) or arbitrary invertible (L and R) deterministic matrices [34] and J is a random matrix of iid elements with zero mean and variance $1/N$. Since J and therefore LJR have zero mean, M is the ensemble average of A. The random fluctuations of A around its average are given by the matrix LJR, which for general L and/or R has dependent and nonidentically distributed elements, due to the possible mixing and nonuniform scaling of the rows (columns) of the iid J by L (R).

We are first interested in the statistics of the eigenvalues of Eq. (2.1). While the statistics of the eigenvalues and eigenvectors of A are of interest in their own right, we also directly consider certain properties of the linear dynamical system,

$$\frac{d\mathbf{x}(t)}{dt} = -\gamma \mathbf{x}(t) + A\mathbf{x}(t) + \mathbf{I}(t), \tag{2.2}$$

for an N-dimensional state vector $\mathbf{x}(t)$, when A is a sample of the ensemble Eq. (2.1). Here γ is a scalar and $\mathbf{I}(t)$ is an external, time-dependent input. In studying this system, we generally assume that Eq. (2.2) is asymptotically stable. This means that M, L, R, and γ must be chosen such that for any typical realization of J, no eigenvalue of $-\gamma \mathbf{1} + M + LJR$ has a positive real part; this can normally be achieved, for example, by choosing a large enough $\gamma > 0$.

Using the diagrammatic technique in the noncrossing approximation, which is valid for large N, we have derived general formulas for several useful properties of such matrices involving their eigenvalues and eigenvectors (see Secs. III–IV for the details of the derivations and the definition of the noncrossing approximation). We present these results in this section. In our derivation of these results, we assume the random J belongs to the complex Ginibre ensemble [26], i.e., the distribution of the elements of J is complex Gaussian. However, we emphasize that universality theorems ensure that, for given M, L, and R, the obtained result for the eigenvalue density in the limit $N \to \infty$ will not depend on the exact choice of the distribution of the elements of J, beyond its first two moments, and extend to *any* iid J (including, e.g., J with

real binary or log-normal elements) whose elements have the same first two moments, i.e., zero mean and variance $1/N$; the universality of the eigenvalue density for general M, L, and R was established in Ref. [35], following earlier work on the universality of the circular law established and successively strengthened in Refs. [26–30]. Furthermore, empirically, from limited simulations, we have thus far found (but have not proved) such universal behavior to also hold for the other quantities we compute here (however, it is possible that universality for these quantities might require the existence of some higher moments beyond the second, as has been found for universality of certain other properties of random matrices; see, e.g., Ref. [36]). To demonstrate the universality of our results, we have used non-Gaussian and/or real J's in most of the numerical examples below.

Hereinafter, we adopt the following notations. For any matrix B, we denote its operator norm (its maximum singular value) by $\|B\|$ and we define its (normalized) Frobenius norm via

$$\|B\|_{\mathrm{F}}^2 \equiv \frac{1}{N} \sum_{ij} |B_{ij}|^2 = \frac{1}{N} \mathrm{Tr}(BB^\dagger) \qquad (2.3)$$

(equivalently, $\|B\|_{\mathrm{F}}$ is the root mean square of the singular values of B).

For general matrices, A and B,

$$\mathrm{tr}(A) \equiv \frac{1}{N} \mathrm{Tr}(A), \quad A^{-\dagger} \equiv (A^\dagger)^{-1},$$

$$\frac{1}{A} \equiv A^{-1}, \quad \frac{A}{B} \equiv AB^{-1},$$

and when adding a scalar to a matrix, it is implied that the scalar is multiplied by the appropriate identity matrix. We denote the identity matrix in any dimension (deduced from the context) by $\mathbf{1}$. For a complex variable $z = x + iy$, the Dirac δ function is defined by $\delta^2(z) \equiv \delta(x)\delta(y)$, and we define $\partial_{\bar{z}} \equiv \partial/\partial\bar{z} = (\partial/\partial x + i\partial/\partial y)/2$ and $\partial_z \equiv \partial/\partial z = (\partial/\partial x - i\partial/\partial y)/2$. For simplicity, we use the notation $f(z)$ [instead of $f(z,\bar{z})$] for general, nonholomorphic functions on the complex plane. We say a quantity is $O(f(N))$ [$\Theta(f(N))$] when, for large enough N, the absolute value of that quantity is bounded above [above *and* below] by a fixed positive multiple of $|f(N)|$. Finally, we say a quantity is $o(f(N))$ when its ratio to $|f(N)|$ vanishes as $N \to \infty$.

The only conditions we impose on M, L, and R are that $\|M\|_{\mathrm{F}}$, $\|L\|_{\mathrm{F}}$, $\|R\|_{\mathrm{F}}$, $\|L^{-1}MR^{-1}\|_{\mathrm{F}}$, and $\|(LR)^{-1}\|$ are bounded as $N \to \infty$. We use the bound on $\|(LR)^{-1}\|$ in Appendixes A and B; the Frobenius norm conditions are assumptions in the universality theorem of Ref. [35], which we use as discussed above. Finally, we assume that for all $z \in \mathbb{C}$, the distribution of the eigenvalues of $M_z M_z^\dagger$, where M_z is defined below in Eq. (2.6), tends to a limit distribution as $N \to \infty$. This last condition simply makes precise the requirement that M, L, and R are defined consistently as functions of N, such that a limit spectral density for $M + LJR$ is meaningful; in particular, it does not impose any further limits on the growth of the eigenvalues of $M_z M_z^\dagger$ with N, beyond the various norm bounds imposed above.

A. Spectral density

1. Summary of results

The density of the eigenvalues of $M + LJR$ in the complex plane for a realization of J (also known as the empirical spectral distribution) is defined by

$$\rho_J(z) = \frac{1}{N} \sum_\alpha \delta^2(z - \lambda_\alpha), \qquad (2.4)$$

where λ_α are the eigenvalues of $M + LJR$. It is known [35] that $\rho_J(z)$ is asymptotically self-averaging, in the sense that with probability one $\rho_J(z) - \rho(z)$ converges to zero (in the distributional sense) as $N \to \infty$, where $\rho(z) \equiv \langle \rho_J(z) \rangle_J$ is the ensemble average of $\rho_J(z)$. Thus, for large enough N, any typical realization of J yields an eigenvalue density $\rho_J(z)$ that is arbitrarily close to $\rho(z)$.

Our general result is that, for large N, with certain cautions and excluding certain special cases as described below [Eqs. (2.19) and (2.20) and preceding discussion], $\rho(z)$ is nonzero in the region of the complex plane satisfying

$$\mathrm{tr}\,[(M_z M_z^\dagger)^{-1}] \geqslant 1, \qquad (2.5)$$

where we defined

$$M_z \equiv L^{-1}(z - M)R^{-1}. \qquad (2.6)$$

Using the definition Eq. (2.3), we can also express Eq. (2.5) as

$$\left\| R\frac{1}{z - M}L \right\|_{\mathrm{F}} \geqslant 1. \qquad (2.7)$$

Inside this region, $\rho(z)$ is given by

$$\rho(z) = \frac{1}{\pi} \frac{\partial}{\partial\bar{z}}\,\mathrm{tr}\left[\frac{(RL)^{-1}M_z^\dagger}{M_z M_z^\dagger + g(z)^2}\right], \qquad (2.8)$$

where $g(z)$ is a real, scalar function found by solving

$$\mathrm{tr}\left[\frac{1}{M_z M_z^\dagger + g^2}\right] = 1, \qquad (2.9)$$

for g for each z. As a first example, for the well-known case of $M = 0$, $L = \mathbf{1}$, and $R = \sigma\mathbf{1}$, we have $M_z = z/\sigma$ and the circular law follows immediately from Eq. (2.5), which yields $\sigma^2/|z|^2 \geqslant 1$ or $|z| \leqslant \sigma$ for the support, and from Eqs. (2.8) and (2.9), which yield the uniform $\rho(z) = 1/(\pi\sigma^2)$ within that support. As we noted in the Introduction, formulas (2.5)–(2.9) generalize the results of Refs. [31] and [32] for the special case $L \propto R \propto \mathbf{1}$ to arbitrary invertible L and R (the eigenvalue density for the case $L \propto R \propto \mathbf{1}$ and a *normal* M was also calculated in Ref. [24]).

It is possible and illuminating to express Eqs. (2.7)–(2.9) exclusively in terms of the singular values of M_z, which we denote by $s_i(z)$ [we include possibly vanishing singular values among $s_i(z)$, so that we always have N of them]. First, noting that the squared singular values of M_z are the eigenvalues of the Hermitian $M_z M_z^\dagger$, we can evaluate the trace in Eq. (2.9) in the eigenbasis of the latter matrix, and rewrite this equation as

$$\frac{1}{N} \sum_{i=1}^N \frac{1}{s_i(z)^2 + g^2} = 1. \qquad (2.10)$$

Similarly, Eq. (2.5) can be equivalently rewritten as

$$\frac{1}{N} \sum_{i=1}^{N} s_i(z)^{-2} \geqslant 1. \tag{2.11}$$

As we prove at the end of Sec. III, Eq. (2.8) can also be written in a form that makes it explicit that the dependence of $\rho(z)$ on M, L, and R is *only* through the singular values of M_z and their derivatives with respect to z and $\bar{z}$. We have

$$\rho(z) = \frac{1}{\pi} \partial_{\bar{z}} \left\{ \frac{1}{N} \sum_{i=1}^{N} \frac{\partial_z [s_i(z)^2]}{s_i(z)^2 + g(z)^2} \right\}. \tag{2.12}$$

For the special case of $M = 0$ and general L and R, our formulas can be simplified considerably. The spectrum is isotropic around the origin in this case; i.e., $\rho(z)$ depends only on $r \equiv |z|$, and its support is a disk centered at the origin with radius

$$r_0 = \|RL\|_{\mathrm{F}} = \left[\frac{1}{N} \sum_{i=1}^{N} \sigma_i^2 \right]^{1/2}, \tag{2.13}$$

where σ_i are the singular values of RL [this follows from Eq. (2.11) by noting that for $M = 0$, the singular values of $M_z = z(RL)^{-1}$ are $s_i(z) = |z|/\sigma_i$]. Within this support the spectral density is given by

$$\rho(r) = -\frac{1}{2\pi r} \partial_r [g(r)^2], \tag{2.14}$$

where $g(r)^2 > 0$ is found by solving

$$1 = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\sigma_i^{-2} r^2 + g(r)^2}. \tag{2.15}$$

Integrating Eq. (2.14), we see that the proportion of eigenvalues lying a distance larger than r from the origin is, in this case, given by

$$n_>(r) = \begin{cases} g(r)^2 & (r < r_0), \\ 0 & (r \geqslant r_0). \end{cases} \tag{2.16}$$

In Sec. III we prove that the eigenvalue density, given by Eqs. (2.14) and (2.15), is always a decreasing function of $r = |z|$, i.e., for $r > 0$ its derivative with respect to r is strictly negative, as long as the limit distribution of the $\{\sigma_i\}$ as $N \to \infty$ has nonzero variance [otherwise, $\rho(z)$ is given by the circular law with radius Eq. (2.13)]. The values of spectral density at $r = 0$ and $r = r_0$ can be calculated explicitly for general L and R:

$$\rho(r = 0) = \frac{1}{\pi} \frac{1}{N} \sum_{i=1}^{N} \sigma_i^{-2}, \tag{2.17}$$

$$\rho(r = r_0) = \frac{1}{\pi} r_0^2 \left[\frac{1}{N} \sum_{i=1}^{N} \sigma_i^4 \right]^{-1} \leqslant \rho \quad (r = 0). \tag{2.18}$$

As noted above, certain cautions apply in using the above formulas for the eigenvalue density and its boundary [(2.5)–(2.9), or, equivalently, Eqs. (2.10)–(2.12), and for $M = 0$, Eqs. (2.14) and (2.15)]. We have written these formulas for finite N (assuming it is large). However, the noncrossing approximation used in deriving these formulas is only guaranteed

to yield the correct result for the eigenvalue density in the limit, i.e., $\lim_{N \to \infty} \rho(z)$ (see Appendix A); finite-size corrections obtained from Eqs. (2.5)–(2.9) are not, in general, correct, and $o(1)$ contributions to $g(z)^2$ or $\rho(z)$ obtained from Eqs. (2.9) and (2.8) should be discarded.

Furthermore, in general, the correct way of finding the support of $\lim_{N \to \infty} \rho(z)$ using Eq. (2.5) is by setting the left side of the inequality (2.5) to $\lim g^2 \to 0^+ \lim N \to \infty$ of the left side of Eq. (2.9), as discussed in Sec. III and Appendix A. However, in writing Eq. (2.5) we have simply set $g = 0$ in Eq. (2.9) and thus implicitly taken the limit $g^2 \to 0^+$ before the $N \to \infty$ limit. To correctly express the support, we must first define the function

$$\mathcal{K}(g,z) \equiv \lim_{N \to \infty} \mathrm{tr} \left[\frac{1}{M_z M_z^\dagger + g^2} \right]$$

$$= \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \left[\frac{1}{s_i(z)^2 + g^2} \right] \tag{2.19}$$

for fixed, strictly positive g, which serves to regularize the denominators in Eq. (2.19) for $s_i(z)$ which are zero or vanishing in the limit $N \to \infty$. The generally correct way of expressing Eq. (2.5) or Eq. (2.11) is then

$$\mathcal{K}(0^+,z) \equiv \lim_{g \to 0^+} \mathcal{K}(g,z) \geqslant 1. \tag{2.20}$$

Let us denote the the support of $\lim_{N \to \infty} \rho(z)$, given by Eq. (2.20), by $\mathcal{S}_{0^+}$ and the region specified by the limit $N \to \infty$ of Eq. (2.5) or Eq. (2.11) by $\mathcal{S}_0$. For many examples of M, L, and R, the limits $N \to \infty$ and $g \to 0^+$ commute everywhere and hence $\mathcal{S}_{0^+} = \mathcal{S}_0$. However, if there are z's at which some of the smallest $s_i(z)$ are either zero or vanish in the limit $N \to \infty$, the two limits may fail to commute, and the naive use of Eq. (2.5) can yield a region, $\mathcal{S}_0$, strictly larger than and containing $\mathcal{S}_{0^+}$, the correct support of $\lim_{N \to \infty} \rho(z)$. For example, at z's for which a $\Theta(1)$ number of $s_i(z)$ are zero or $o(1)$, these singular values do not make a contribution to $\mathcal{K}(g,z)$ for $g > 0$ [their contribution to the sum in Eq. (2.19) is $O(N^{-1})$] and hence to $\mathcal{K}(0^+,z)$, but if they vanish sufficiently fast as $N \to \infty$ they can make a nonzero contribution to the left side of Eq. (2.11); such z may fall within $\mathcal{S}_0$, but not within $\mathcal{S}_{0^+}$. For finite N, the $s_i(z)$ can vanish exactly when z coincides with an eigenvalue of M; thus, the above situation can, e.g., arise close to eigenvalues of M that are isolated and far from the rest of M's spectrum, so that they fall outside the support of $\lim_{N \to \infty} \rho(z)$. In such cases, the spectrum of $M + LJR$ will nonetheless typically also contain isolated eigenvalues [which do not contribute to $\lim_{N \to \infty} \rho(z)$] with effectively deterministic location, i.e., within $o(1)$ distance of corresponding isolated eigenvalues of M; examples of this phenomenon, for which $\mathcal{S}_0 - \mathcal{S}_{0^+}$ is not empty but has zero measure, have been studied in Refs. [37,38] (for symmetric matrices, outlier eigenvalues corresponding to eigenvalues of the mean matrix were first studied in Ref. [39]). For some choices of M, L, and R, however, a more interesting case can arise such that for z in a certain region of the complex plane with nonzero measure, all $s_i(z)$ are nonzero at finite N (hence, M has no eigenvalue there), but a few $s_i(z)$ are $o(1)$ and vanish sufficiently fast as $N \to \infty$; in

particular, when $L \propto R \propto \mathbf{1}$, this can occur for certain highly non-normal M [40]. In such cases, the noncommutation of the two limits can lead to a difference $\mathcal{S}_0 - \mathcal{S}_{0^+}$ with nonzero measure. In cases we have examined this signifies that there exists a finite, nonvanishing region outside the support of $\lim_{N\to\infty} \rho(z)$ (typically surrounding it) where, although $\rho(z)$ is $o(1)$, it nonetheless converges to zero sufficiently slowly that a $\Theta(1)$ number of "outlier" eigenvalues lie there (note that the vast majority of eigenvalues, i.e., $[1 - o(1)]N$ of them, lie within the support of the limit density). We discuss examples of this phenomenon in Sec. II C below; in one of the examples (discussed in Sec. II C 2), the existence of such outlier eigenvalues was first noted in Ref. [41], and their distribution was quantitatively characterized in Ref. [37]. However, the connection between such outlier eigenvalues and nonzero but $o(1)$ singular values of M_z, which arise, e.g., for highly non-normal M, were not noted before to the best of our knowledge.

We have observed in simulations (and also supported by [37]) that the distribution of these outliers remains random as $N \to \infty$, is, in general, less universal than $\lim_{N\to\infty} \rho(z)$ (e.g., it could depend on the choice of real vs complex ensembles for J), and has an average behavior that may not be correctly given by the noncrossing approximation.

2. Relationship to pseudospectra

Finally, we note a remarkable connection between our general result for the support of the spectrum Eq. (2.5) and the notion of pseudospectra, in the case in which the limits $g^2 \to 0^+$ and $N \to \infty$ commute [so that Eq. (2.5) correctly describes the support]. Pseudospectra are generalizations of eigenvalue spectra, which are particularly useful in the case of non-normal matrices (see Ref. [17] for a review). The eigenvalue spectrum of matrix M can be thought of as the set of points, z, in the complex plane where $(z - M)^{-1}$ is singular, i.e., it has infinite norm. Given a fixed choice of matrix norm, $\| \cdot \|$, the pseudospectrum of M at level σ, or its "σ pseudospectrum" in the given norm, is the set of points z for which $\|(z - M)^{-1}\| \geqslant \sigma^{-1}$ (thus, as $\sigma \to 0$ we recover the spectrum). For the specific choice of the operator norm (i.e., when $\|A\|$ is taken to be the maximum singular value of A), the σ pseudospectrum can equivalently be characterized as the set of points, z, for which there exists a matrix perturbation ΔM, with $\|\Delta M\| \leqslant \sigma$, such that z is in the eigenvalue spectrum of $M + \Delta M$ [17,42]. In words, in the operator norm, the σ pseudospectrum of M is the set to which its spectrum can be perturbed by adding to it arbitrary perturbations of size σ or smaller.

In our setting we can think of LJR as a perturbation of M. Let us focus on the case where L and R are proportional to the identity, i.e., we have $\Delta M = \sigma J$, with a positive scalar σ. Our result, Eq. (2.7), in this case reads $\|\sigma(z - M)^{-1}\|_{\mathrm{F}} \geqslant 1$ or $\|(z - M)^{-1}\|_{\mathrm{F}} \geqslant \sigma^{-1}$. In other words, as $N \to \infty$, the spectrum of $M + \sigma J$, for an iid random J with zero mean and variance $1/N$, is the σ pseudospectrum of M in the normalized Frobenius norm defined by Eq. (2.3). Interestingly, the perturbation, $\Delta M = \sigma J$, has normalized Frobenius norm σ as $N \to \infty$: This norm is $\sigma\sqrt{\sum_{ij} J_{ij}^2 / N}$, which, by the law of large numbers, converges to σ for large N. That is, as

$N \to \infty$, the spectrum in response to the random perturbation σJ, which has size σ (in normalized Frobenius norm), is the σ pseudospectrum of M in the normalized Frobenius norm.

This result sounds similar to the equivalence of the two definitions of pseudospectra for the operator norm which we noted above [one based on the norm of $(z - M)^{-1}$ and one based on the spectra of bounded perturbations], but it differs in two key respects. First, unlike in the case of the operator norm, the general equivalence of the two notions of pseudospectra noted above does *not* hold for the normalized Frobenius norm. Second, for the operator norm, it is *not*, in general, the case that the σ pseudospectrum of M is equivalent to the spectrum obtained from a single *random* perturbation of M of size σ, even in the limit $N \to \infty$ (although the spectra arising from such random perturbations are sometimes used as a "poor man's version" or approximation of the pseudospectra [17]). This can be seen as follows. The operator norm of the random iid perturbation, σJ, i.e., its maximum singular value, converges almost surely to 2σ as $N \to \infty$ [43]. Condition (2.7) for z to be in the spectrum under this random perturbation is $\|(z - M)^{-1}\|_{\mathrm{F}} \geqslant \sigma^{-1}$, or $\mathrm{rms}(\{s_i(z)^{-1}\}) \geqslant 1$ where the $s_i(z)$ are the singular values of $\frac{z-M}{\sigma}$ and $\mathrm{rms}(\{x_i\})$ represents the root mean square of the set of values $\{x_i\}$. This is not equivalent to the condition that z be in the 2σ pseudospectrum of M in the operator norm, i.e., that $\|(z - M)^{-1}\| \geqslant (2\sigma)^{-1}$ or $s_{\min}(z)^{-1} \geqslant \frac{1}{2}$, where $s_{\min}(z)$ is the minimum of the $s_i(z)$; in fact, noting that $s_{\min}(z)^{-1} \geqslant \mathrm{rms}(\{s_i(z)^{-1}\})$, it is easy to see that the spectrum under random iid perturbations with operator norm $\|\sigma J\| = 2\sigma$ is strictly a proper subset of the 2σ pseudospectrum in the operator norm. For example, for $M = 0$, the "poor man's 2σ pseudospectrum" in the limit $N \to \infty$ is a ball of radius σ about the origin (the circular law), while the true 2σ pseudospectra of the zero matrix is the ball of radius 2σ about the origin.

In sum, in the operator norm, the σ pseudospectrum of M for any N is equivalent to the set of points z for which *some* perturbation ΔM with $\|\Delta M\| \leqslant \sigma$ can be found such that z is in the spectrum of $M + \Delta M$ [17]. In the normalized Frobenius norm in the limit $N \to \infty$, however, the σ pseudospectrum of M is equivalent to the spectrum of $M + \Delta M$, where ΔM is any *random* perturbation with zero-mean iid elements with $\|\Delta M\|_{\mathrm{F}} = \sigma$. This statement for the normalized Frobenius norm holds when the two limits $N \to \infty$ and $g \to 0^+$ commute; when the two limits do not commute, the support of the spectral distribution of $M + \Delta M$ is a subset of the σ pseudospectrum of M in the normalized Frobenius norm.

B. Average norm squared and power spectrum

As we mentioned in the Introduction, an important phenomenon encountered in dynamics governed by non-normal matrices, as described by Eq. (2.2) with $\mathbf{I}(t) = 0$, is transient amplification in asymptotically stable systems. In any stable system, the size of the response to an initial perturbation eventually decays to zero, with an asymptotic rate set by the system's eigenvalues. In stable non-normal systems, however, after an initial perturbation, the size of the network activity, as measured, e.g., by its norm squared $\|\mathbf{x}(t)\|^2 = \mathbf{x}(t)^{\mathsf{T}}\mathbf{x}(t)$, can nonetheless exhibit transient, yet possibly large and

long-lasting growth, before it eventually decays to zero. By contrast, in stable normal systems, $\|\mathbf{x}(t)\|^2$ can only decrease with time. The strength and even the time scale of transient amplification are set by properties of the matrix A beyond its eigenvalues; they depend on the degree of non-normality of the matrix, as measured, e.g., by the degree of nonorthogonality of its eigenvectors, or, alternatively, by its hidden feedforward structure [see Eq. (2.34) for the latter's definition].

Non-normal systems can also exhibit pseudoresonances at frequencies that could be very different from their natural frequencies as determined by their eigenvalues; such pseudoresonances will be manifested in the frequency power spectrum of the response of the system to time-dependent inputs. $\|\mathbf{x}(t)\|^2$ and the power spectrum of response are examples of quantities that depend not only on the eigenvalues of $M + LJR$ but also on its eigenvectors.

Here we present a few closely related formulas for general M, L, and R. These include a formula for $\langle \|\mathbf{x}(t)\|^2 \rangle_J$, i.e., the ensemble average of the norm squared of the state vector, $\mathbf{x}(t)$, as it evolves under Eq. (2.2) with $\mathbf{I}(t) = 0$, as well as a formula for the ensemble average of the power spectrum of the response of the network to time-varying inputs. The results of this section are valid, and in the case of the power spectrum meaningful, when the system Eq. (2.2) is asymptotically stable. As we mentioned after Eq. (2.2), this means that M, L, R, and γ must be chosen such that for any typical realization of J, all eigenvalues of $-\gamma \mathbf{1} + M + LJR$ have negative real part. In particular, the entire support of the eigenvalue density of $M + LJR$, as determined by Eq. (2.5), must fall to the left of the vertical line of z's with real part γ; this is a necessary condition, but may not be sufficient either at finite N or in cases where an $O(1)$ number of eigenvalues remain outside this region of support even as $N \to \infty$.

First, we consider the time evolution of the squared norm, $\|\mathbf{x}(t)\|^2$, of the response of the system to an impulse input, $\mathbf{I}(t) = \mathbf{x}_0 \delta(t)$, at $t = 0$, before which we assume the system was in its stable fixed point $\mathbf{x} = 0$ [for $t > 0$ this is equivalent to the squared norm of the activity as it evolves according to Eq. (2.2) with $\mathbf{I}(t) = 0$, starting from the initial condition $\mathbf{x}(0) = \mathbf{x}_0$]. We provide a formula for the ensemble average of the more general quadratic function, $\mathbf{x}(t)^{\mathrm{T}} B \mathbf{x}(t)$, where B is any $N \times N$ symmetric matrix; the norm squared corresponds to $B = \mathbf{1}$. The result for general B, M, L, and R is given as a double inverse Fourier transform,

$$\langle \mathbf{x}(t)^{\mathrm{T}} B \mathbf{x}(t) x \rangle_J$$
$$= \iint \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} e^{it(\omega_1 - \omega_2)} \mathrm{Tr} \left[B \, \mathcal{C}^{\mathbf{x}}(\omega_1, \omega_2; \mathbf{x}_0 \mathbf{x}_0^{\mathrm{T}}) \right], \quad (2.21)$$

in terms of the $N \times N$ Fourier-domain "covariance matrix," $\mathcal{C}^{\mathbf{x}}(\omega_1, \omega_2; \mathbf{x}_0 \mathbf{x}_0^{\mathrm{T}}) \equiv \left\langle \tilde{\mathbf{x}}(\omega_1) \tilde{\mathbf{x}}(\omega_2)^{\dagger} \right\rangle_J$ [where $\tilde{\mathbf{x}}(\omega)$ is the Fourier transform of $\mathbf{x}(t)$]. The expression for the latter is given by

$$\mathcal{C}^{\mathbf{x}}(\omega_1, \omega_2; C^{\mathbf{I}}) = \mathcal{C}_0^{\mathbf{x}}(\omega_1, \omega_2; C^{\mathbf{I}}) + \Delta \mathcal{C}^{\mathbf{x}}(\omega_1, \omega_2; C^{\mathbf{I}}), \quad (2.22)$$

where

$$\mathcal{C}_0^{\mathbf{x}}(\omega_1, \omega_2; C^{\mathbf{I}}) \equiv \frac{1}{\gamma + i\omega_1 - M} C^{\mathbf{I}} \frac{1}{\gamma - i\omega_2 - M^{\dagger}} \quad (2.23)$$

yields the result obtained by ignoring the randomness in the connectivity (i.e., by setting $A = M$), and

$$\Delta \mathcal{C}^{\mathbf{x}}(\omega_1, \omega_2; C^{\mathbf{I}})$$
$$\equiv \frac{1}{\gamma + i\omega_1 - M} LL^{\dagger} \frac{1}{\gamma - i\omega_2 - M^{\dagger}}$$
$$\times \frac{\mathrm{tr} \left(R^{\dagger} R \frac{1}{\gamma + i\omega_1 - M} C^{\mathbf{I}} \frac{1}{\gamma - i\omega_2 - M^{\dagger}} \right)}{1 - \mathrm{tr} \left(R^{\dagger} R \frac{1}{\gamma + i\omega_1 - M} LL^{\dagger} \frac{1}{\gamma - i\omega_2 - M^{\dagger}} \right)} \quad (2.24)$$

is the contribution of the random part of connectivity LJR. For later use, we have provided these expressions for a general third argument in $\mathcal{C}^{\mathbf{x}}(\cdot, \cdot; \cdot)$; for use in Eq. (2.21) $C^{\mathbf{I}}$ must be substituted with $\mathbf{x}_0 \mathbf{x}_0^{\mathrm{T}}$. In the special case of $\langle \|\mathbf{x}(t)\|^2 \rangle_J$ corresponding to $B = \mathbf{1}$, and iid disorder ($L = \mathbf{1}$, $R = \sigma \mathbf{1}$), the contributions from Eqs. (2.23) and (2.24) can be more compactly combined into

$$\langle \|\mathbf{x}(t)\|^2 \rangle_J = \iint \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} e^{it(\omega_1 - \omega_2)}$$
$$\times \frac{\mathbf{x}_0^{\mathrm{T}} \frac{1}{\gamma - i\omega_2 - M^{\dagger}} \frac{1}{\gamma + i\omega_1 - M} \mathbf{x}_0}{1 - \sigma^2 \mathrm{tr} \left(\frac{1}{\gamma + i\omega_1 - M} \frac{1}{\gamma - i\omega_2 - M^{\dagger}} \right)} \quad (2.25)$$

[we used $\mathrm{Tr} \left(\frac{1}{z_1 - M} \mathbf{x}_0 \mathbf{x}_0^{\mathrm{T}} \frac{1}{z_2 - M^{\dagger}} \right) = \mathbf{x}_0^{\mathrm{T}} \frac{1}{z_2 - M^{\dagger}} \frac{1}{z_1 - M} \mathbf{x}_0$ to write the numerator in Eq. (2.25)].

Next we look at the power spectrum of the response of the system to a noisy input, $\mathbf{I}(t)$, that is temporally white, with zero mean and covariance

$$\overline{\mathbf{I}_i(t_1) \mathbf{I}_j(t_2)} = \delta(t_1 - t_2) C_{ij}^{\mathbf{I}}. \quad (2.26)$$

Here the bar indicates averaging over the input noise (or by ergodicity, over a long enough time). Our general result for the ensemble average of the matrix power spectrum of the response, which by definition is the Fourier transform of the steady-state response covariance,

$$C_{ij}^{\mathbf{x}}(\omega) \equiv \int d\tau \, e^{-i\omega\tau} \overline{\mathbf{x}_i(t+\tau) \mathbf{x}_j(t)}, \quad (2.27)$$

is given by

$$\langle C^{\mathbf{x}}(\omega) \rangle_J = C_0^{\mathbf{x}}(\omega) + \Delta C^{\mathbf{x}}(\omega). \quad (2.28)$$

Here we defined

$$C_0^{\mathbf{x}}(\omega) \equiv \mathcal{C}_0^{\mathbf{x}}(\omega, \omega; C^{\mathbf{I}}) \quad (2.29)$$

and

$$\Delta C^{\mathbf{x}}(\omega) \equiv \Delta \mathcal{C}^{\mathbf{x}}(\omega, \omega; C^{\mathbf{I}}) \quad (2.30)$$

as the power spectrum matrices obtained by ignoring the randomness in connectivity (i.e., by setting $A = M$) and the contribution of quenched randomness LJR, respectively.

A closely related quantity is the total power of the steady-state response of the system to a sinusoidal input $\mathbf{I}(t) = \mathbf{I}_0 \sqrt{2} \cos \omega t$ (the $\sqrt{2}$ serves to normalize the average power of $\sqrt{2} \cos \omega t$ to unity, so that the total power in the input is $\|\mathbf{I}_0\|^2$). For such an input, the steady-state activity, which we denote by $\mathbf{x}_\omega(t)$, is also sinusoidal (with a possible phase shift). By total power of the steady-state response we mean the time average of the squared norm of the activity, $\overline{\|\mathbf{x}_\omega(t)\|^2}$, where now the bar indicates temporal averaging [we call this

total power, because the squared norm sums the power in all components of $\mathbf{x}_\omega(t)$]. As in Eqs. (2.21)–(2.24), we present a formula for the ensemble average of the more general quantity $\overline{\mathbf{x}_\omega^{\mathsf{T}} B\, \mathbf{x}_\omega}$. We have

$$\langle \overline{\mathbf{x}_\omega^{\mathsf{T}} B\, \mathbf{x}_\omega} \rangle_J = \mathrm{Tr}\,[B\,\langle C^{\mathbf{x}}(\omega)\rangle_J], \tag{2.31}$$

where $\langle C^{\mathbf{x}}(\omega)\rangle_J$ is given by Eqs. (2.28)–(2.30) with $C^{\mathbf{I}}$ replaced with $\mathbf{I}_0 \mathbf{I}_0^{\mathsf{T}}$. For the special case of $B = \mathbf{1}$, corresponding to the total power of the response at frequency ω, using Eqs. (2.23) and (2.24) with $\omega_1 = \omega_2 = \omega$, this formula can be simplified into

$$\langle \overline{\|\mathbf{x}_\omega\|^2} \rangle_J = \left\| \frac{1}{z-M}\mathbf{I}_0 \right\|^2 + \frac{\left\| \frac{1}{z-M} L \right\|_{\mathrm{F}}^2 \; \left\| R \frac{1}{z-M}\mathbf{I}_0 \right\|^2}{1 - \left\| R \frac{1}{z-M} L \right\|_{\mathrm{F}}^2}, \tag{2.32}$$

where $z = \gamma + i\omega$, $\|\cdot\|$ denotes the vector norm, and $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm defined in Eq. (2.3). Finally, for the case that the random part of the matrix is iid, i.e., $L = \sigma \mathbf{1}$ and $R = \mathbf{1}$, we can further simplify Eq. (2.32) into

$$\langle \overline{\|\mathbf{x}_\omega\|^2} \rangle_J = \frac{\|(\gamma + i\omega - M)^{-1}\mathbf{I}_0\|^2}{1 - \sigma^2 \|(\gamma + i\omega - M)^{-1}\|_{\mathrm{F}}^2}. \tag{2.33}$$

The stability of the $\mathbf{x} = 0$ fixed point guarantees the positivity of the expressions Eqs. (2.32) and (2.33) for the power spectrum. This is true because, as we noted above, stability requires that the support of the eigenvalue density of A is entirely to the left of the vertical line $\mathrm{Re}(z) = \gamma$. By our result, Eq. (2.7), for that support, this can only be true if the denominators of the last terms in Eqs. (2.32) and (2.33) are positive, which guarantees the positivity of the full expressions.

Note that the first term in Eq. (2.32) and the numerator in Eq. (2.33) represent the power spectrum in the absence of randomness, i.e., if A in Eq. (2.2) is replaced with M. Thus, formulas (2.32) and (2.33) show that the correct average power spectrum is always strictly larger than the naive power spectrum obtained by assuming that random effects will "average out." Furthermore, due to the denominators of the last terms in Eqs. (2.32) and (2.33), the power spectrum will be larger for frequencies where the support of the eigenvalue density, Eq. (2.7), is closer to the vertical line with $\mathrm{Re}(z) = \gamma$. Similar, but less precise statements can also be made about the strength of transient amplification using formulas (2.21)–(2.25) for the squared norm of the impulse response. One measure of the strength of transient amplification up to time T is $\int_0^T \|\mathbf{x}(t)\|^2 dt$. Integrating formulas Eq. (2.21) (with $B = \mathbf{1}$) or Eq. (2.25) over t, one obtains formulas for $\int_0^T \|\mathbf{x}(t)\|^2 dt$ that are the same as Eqs. (2.21)–(2.25), except for the factor $e^{it(\omega_1 - \omega_2)}$ in the integrands of Eqs. (2.21) and (2.25) being replaced with $\frac{i[1 - e^{iT(\omega_1 - \omega_2 + i\epsilon)}]}{\omega_1 - \omega_2 + i\epsilon}$ (with $\epsilon \to 0^+$). Due to the denominator in this factor (for T sufficiently large the numerator is constant), the main contribution to the integrals over ω_1 and ω_2 should typically arise for $\omega_1 \approx \omega_2$. On the other hand, note that for $\omega_1 = \omega_2$ the denominators in Eqs. (2.24) and (2.25) reduce to the those in Eqs. (2.32) and (2.33), with the connection to the support of the spectral density noted above. Thus, this dominant contribution to $\int_0^T \|\mathbf{x}(t)\|^2 dt$ must be larger, the closer the support of the eigenvalue density, Eq. (2.7),
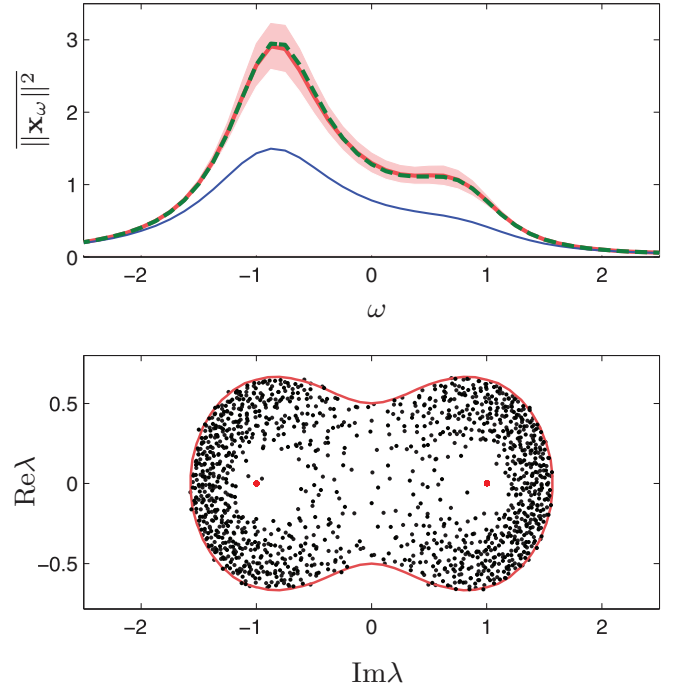

FIG. 1. (Color online) (Top) The total power spectrum of steady-state response $\overline{\|\mathbf{x}_\omega\|^2}$ as a function of input frequency ω, Eq. (2.33), for the system Eq. (2.2) with $A = M + \sigma J$ and M given by Eq. (2.35) with $w = 1$ and $\lambda_n = \pm i$ (with $+i$ and $-i$ alternating), respectively. Here $N = 700$, $\sigma = 0.5$, and $\gamma = 0.8$. The input was fed into the last component of $\mathbf{x}$ [the beginning of the feedforward chain characterized by Eq. (2.35)], which for the matrix M has natural frequency -1. That is, the input was $\mathbf{I}_0\sqrt{2}\cos \omega t$, where $\mathbf{I}_0$ was 1 for the last component and 0 for all other components. The green (thick dashed) curve is the ensemble average of the total power spectrum, $\langle \overline{\|\mathbf{x}_\omega\|^2}\rangle_J$, calculated numerically using the general formula Eq. (2.33), which is compared with an empirical average over 100 realizations of real Gaussian J (solid red line, mostly covered by the dashed green line). The pink (light gray) area shows the standard deviation among these 100 realizations around this average. The blue (thin) line shows the result when disorder, σJ, is ignored, i.e., A is replaced with its ensemble average M. (Bottom) The eigenvalue spectrum of $M + \sigma J$ (black dots). Red big dots at $\pm i$ show the eigenvalues of M. The red curve is the outer boundary of the eigenvalue spectrum of A as computed numerically using Eq. (2.5). The real and imaginary axes of the complex plane are interchanged, so that the frequency axis in the top panel can be matched with the imaginary part of the eigenvalues, i.e., the natural frequencies of Eq. (2.2).

is to the vertical line with $\mathrm{Re}(z) = \gamma$. This also suggests that, as in the case of the power spectrum, the strength of transient amplification would typically be underestimated if randomness of connectivity is ignored and only its ensemble average M is taken into account in solving Eq. (2.2).

Numerical simulations indicate that the quantities $\|\mathbf{x}(t)\|^2$ and $\overline{\|\mathbf{x}_\omega\|^2}$ are self-averaging in the large N limit; that is, for large N, $\|\mathbf{x}(t)\|^2$ or $\overline{\|\mathbf{x}_\omega\|^2}$ for any typical random realization of J will be very close to their ensemble averages, given by Eqs. (2.25) and (2.32), respectively, with the random deviations from these averages approaching zero as N goes to infinity (see Figs. 1, 3, and 8, below). This conclusion is also corroborated

by rough estimations (not shown) based on Feynman diagrams (the diagrammatic method is introduced in Secs. III and IV) of the variance of fluctuations of these quantities for different realizations of J.

Finally, we note that the general formulas presented in this section are valid only for cases where the initial condition, $\mathbf{x}_0$, or the input structure, $\mathbf{I}_0$ or $C^{\mathbf{I}}$, are chosen independently of the particular realization of the random matrix J (e.g., cases where $\mathbf{x}_0$ is itself random but independent of J, or when $\mathbf{x}_0$ is chosen based on properties of M, L, or R). In particular, our results do not apply to cases in which the initial condition or the input is tailored or optimized for the particular realization of the quenched randomness, J, in which case the true result could be significantly different from those given by the formulas of this section.

C. Some specific examples of M, R, and L

In this section we present the results of explicit calculations of the eigenvalue density, Eq. (2.8), the average squared norm of response to impulse, Eqs. (2.21) and (2.25), and the total power in response to sinusoidal input, Eq. (2.33), for specific examples of M, L, and R (the details of the calculations for the results presented here can be found in Sec. V). For many of the examples presented here, L and R are both proportional to the identity matrix; thus, in these examples the full matrix is of the form $M + \sigma J$, where $\sigma > 0$ determines the strength of disorder in the matrix. In Secs. II C 2 and II C 3, we also present examples with nontrivial L and/or R.

Any matrix, M, can be turned into an upper-triangular form by a unitary transformation; i.e.,

$$M = UTU^{\dagger}, \qquad (2.34)$$

where U is unitary and T is upper-triangular (i.e., $T_{ij} = 0$ if $i > j$) with its main diagonal consisting of the eigenvalues of M. The difference between non-normal and normal matrices is that for the latter, T can be taken to be strictly diagonal. Equation (2.34) is referred to as a Schur decomposition of M [44], and we refer to the orthogonal modes of activity represented by the columns of U as Schur modes. The Schur decomposition provides an intuitive way of characterizing the dynamical system Eq. (2.2). Rewriting Eq. (2.2), with J and $\mathbf{I}(t)$ set to zero, in the Schur basis by defining $\mathbf{y} = U^{\dagger}\mathbf{x}$ (i.e., y_i is the activity in the ith Schur mode), we obtain $\frac{d\mathbf{y}}{dt} = -\gamma \mathbf{y} + T\mathbf{y}$. We see that activity in the jth Schur mode provides an input to the equation for the ith mode only when $i \leqslant j$ (as $T_{ij} = 0$ for $i > j$). Thus, the coupling between modes is feedforward, going only from higher modes to lower ones, without any feedback. We refer to T_{ij}'s for $j > i$ as feedforward weights. As these vanish for normal matrices, we can say that a matrix is more non-normal the stronger its feedforward weights are.

Due to the invariance of the trace, the norm, and the adjoint operation under unitary transforms, our general formulas for the spectral density Eq. (2.8) and the average squared norm in time and frequency space, Eqs. (2.25) and (2.33), take the same form in any basis, so, in particular, we can work in the Schur basis of M. Hence, M can be replaced with T, provided L and R are also expressed in M's Schur basis and $\mathbf{x}_0$ or $\mathbf{I}_0$ are replaced with $U^{\dagger}\mathbf{x}_0$ or $U^{\dagger}\mathbf{I}_0$, respectively [45]. Thus, we

use the feedforward structure of the Schur decomposition to characterize the different examples we consider below. Our examples are chosen to demonstrate interesting features of non-normal matrices in the simplest possible settings.

1. Single feedforward chain of length N

In the first example, each and every Schur mode is only connected to its lower adjacent mode, forming a long feedforward chain of length N. For simplicity, we take all feedforward weights in this chain to have the same value w, so that

$$M = T = \begin{pmatrix} \lambda_1 & w & 0 & \cdots \\ 0 & \lambda_2 & w & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad (2.35)$$

or, more succinctly, $M_{nm} = w\,\delta_{n+1,m} + \lambda_n \delta_{nm}$.

Figure 1 shows the power spectrum of response (top panel) and the eigenvalue distribution (bottom panel) of $A = M + \sigma J$ for an example M of the form Eq. (2.35) with alternating imaginary eigenvalues, $\lambda_n = (-1)^{n+1}i$. The black dots in the bottom panel of Fig. 1 show the eigenvalues of A for one realization of J, scattered around the highly degenerate spectrum of M at $\pm i$ (red dots). The top panel shows the ensemble average of the total power spectrum of response, $\langle \overline{\|\mathbf{x}_\omega\|^2} \rangle_J$, of the system Eq. (2.2) to sinusoidal stimuli as given by our general formula Eq. (2.33) (green curve), showing that it perfectly matches the empirical average (red curve) over a set of 100 realizations of J [the latter was obtained by generating 100 realizations of J, calculating $\overline{\|\mathbf{x}_\omega\|^2}$ for each realization, which is given by the numerator of Eq. (2.33) with M replaced with $M + \sigma J$, and then averaging the results over the 100 realizations]. The pink (light gray) shading shows the standard deviation of the power spectrum over these 100 realizations. This will shrink to zero as N goes to infinity, so that for large N the power spectrum of any single realization of $A = M + \sigma J$ will lie very close to the ensemble average. The system (2.2) in the zero-disorder case, $\sigma = 0$, has two highly degenerate resonant frequencies (imaginary parts of the eigenvalues of M), $\omega_0^{\pm} = \pm 1$, leading to possible peaks in the power spectrum at these frequencies. The smaller the decay of these modes (in this case given by γ) is, i.e., the closer the eigenvalues of the combined matrix $-\gamma + M$ are to the imaginary axis, the sharper and stronger are the resonances. Comparing the zero-disorder power spectrum (blue curve) with that for $A = M + \sigma J$, we see that the disorder has led to strong but unequal amplification of the two resonances relative to the case without disorder. This is partly due to the disorder scattering some of the eigenvalues of $-\gamma + A$ much closer to the imaginary axis, creating larger resonances.

For M of the form (2.35) with all eigenvalues zero we have analytically calculated the eigenvalue density, Eq. (2.8), the magnitude of response to impulse Eq. (2.25), and the power spectrum Eq. (2.33). In this case, using Eq. (2.5) naively yields $|z| \leqslant \sqrt{|w|^2 + \sigma^2}$ for the support of the eigenvalue density. However, using the correct procedure, Eqs. (2.19) and (2.20), we find that this formula is only correct for $\sigma \geqslant |w|$, while for $\sigma < |w|$, the true support of the eigenvalue density in the limit
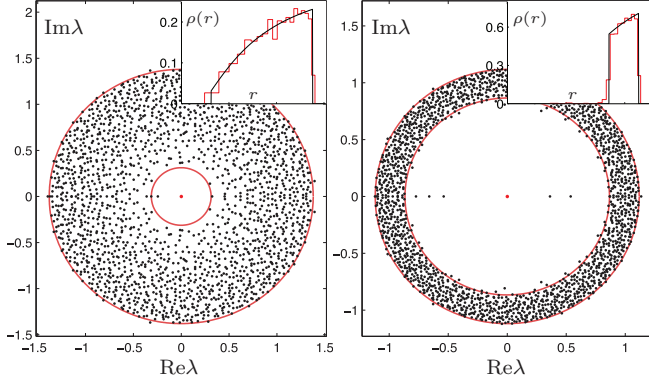
FIG. 2. (Color online) The eigenvalue spectra of $A = M + \sigma J$ for $N = 2000$ and M given by Eq. (2.35) with $\lambda_n = 0$, $w = 1$ for single realizations of real Gaussian J. $\sigma = 0.95$ and 0.5 in the left and rights panels, respectively. The red circles mark the circular boundaries of the spectral support given by Eq. (2.36). The insets show a comparison of the analytic formula Eq. (2.37) for the spectral density (black smooth trace) and histograms corresponding to the particular realization shown in the main plot (red jagged trace).

$N \to \infty$ is the annulus

$$\sqrt{|w|^2 - \sigma^2} \leqslant |z| \leqslant \sqrt{|w|^2 + \sigma^2} \qquad (2.36)$$

(this result was obtained in Ref. [31]). Within this support the eigenvalue density in either case is

$$\rho(z) = \frac{1}{\pi \sigma^2} \left[1 - \frac{|w|^2}{\sqrt{4|w|^2|z|^2 + \sigma^4}} \right]. \qquad (2.37)$$

Figure 2 demonstrates the close agreement of Eqs. (2.36) and (2.37) with the empirical spectrum of $M + \sigma J$ for a single realization of J, for $N = 2000$ and two different values of σ. The discrepancy between the results obtained by the naive use of Eq. (2.5) and Eq. (2.36) is due to the fact that for $|z| < |w|$, $M_z = (z - M)/\sigma$ has an exponentially small, $O(e^{-cN})$, singular value (see next paragraph) which makes the result of Eqs. (2.19) and (2.20) dependent on the order of the two limits $N \to \infty$ and $g \to 0^+$. As we discussed after Eq. (2.12), such a discrepancy can signify the existence of an $O(1)$ number of outlier eigenvalues outside the support of $\lim_{N \to \infty} \rho(z)$. Simulations show that this is the case for $|z| < \sqrt{|w|^2 - \sigma^2}$ (see Fig. 2).

The most striking aspect of these results is revealed in the limit $\sigma \to 0$. For $\sigma = 0$, the spectrum is that of M, which is concentrated at the origin. Remarkably, however, as seen from Eqs. (2.36) and (2.37), for very small but nonzero σ the bulk of the eigenvalues are concentrated in the narrow ring with modulus $|z| \approx |w|$. Thus, in the limit $N \to \infty$ the spectrum has a discontinuous jump at $\sigma = 0$. This is a consequence of the extreme non-normality of M, which manifests itself in the extreme sensitivity of its spectrum to small perturbations, which is well known (see Ref. [17], Chap. 7). The notion of pseudospectra quantifies this sensitivity: The (operator norm) ϵ pseudospectrum of M is the region of complex plane to which its spectrum can be perturbed by adding to M a matrix of operator norm no larger than ϵ. As we mentioned in Sec. II A, this is precisely

the set of complex values z for which $\|(z - M)^{-1}\| > \epsilon^{-1}$ [17], and therefore by the definition of the operator norm $\| \cdot \|$, the region in which $\|(z - M)^{-1}\|^{-1} = s_{\min}(z - M) < \epsilon$, where $s_{\min}(z - M)$ is the least singular value of $z - M$. As noted above, for $|z| < |w|$, $s_{\min}(z - M)$ is exponentially small: $s_{\min}(z - M) \leqslant |w| |\frac{z}{w}|^N$ [for a proof, see after Eq. (5.15) in Sec. V A]. Thus, the ϵ pseudospectrum of M contains the set of points z satisfying $|w| |\frac{z}{w}|^N < \epsilon$, i.e., the centered disk with radius $|w|(\frac{\epsilon}{|w|})^{1/N}$, which approaches $|w|$ as $N \to \infty$. In other words, for large enough N, any point $|z| < |w|$ is in the ϵ pseudospectrum for any *fixed* ϵ, no matter how small. It has been stated [17] that dense random perturbations, of the form σJ considered here, tend to trace out the entire ϵ pseudospectrum (where $\epsilon = \sigma \|J\| \approx 2\sigma$). Our result shows that, for $\epsilon, \sigma \ll |w|$, the spectrum of such perturbations traces out the ϵ pseudospectrum in quite an uneven fashion; the vast majority $[\Theta(N)]$ of the perturbed eigenvalues only trace out the boundary of the pseudospectrum, $|z| \approx |w|$, while only a few $[O(1)]$ eigenvalues lie in its interior. Thus, dense random perturbations can fail as a way of visualizing (operator norm based) pseudospectra.

We now turn to the dynamics. We have explicitly calculated the average evolution of the magnitude of $\mathbf{x}(t)$, Eq. (2.25), and the total power spectrum of steady-state response, Eq. (2.33), for the case where the initial condition is (or the input is fed into) the last Schur mode, i.e., the beginning of the feedforward chain: $\mathbf{x}_0 = (0, \dots, 0, 1)^{\mathsf{T}}$ [or $\mathbf{I}_0 \propto (0, \dots, 0, 1)^{\mathsf{T}}$]. For the evolution of the average norm squared, with the initial condition $\mathbf{x}_0 = (0, \dots, 0, 1)^{\mathsf{T}}$, we obtain

$$\langle \|\mathbf{x}(t)\|^2 \rangle_J = e^{-2\gamma t} I_0(2t\sqrt{|w|^2 + \sigma^2}) \quad (t \geqslant 0), \qquad (2.38)$$

where $I_\nu(x)$ is the νth modified Bessel function. Figure 3 plots the function Eq. (2.38) and compares it with the result obtained by ignoring the disorder (corresponding to $\sigma = 0$). The main difference between the two curves is the slower asymptotic decay of the $\sigma \neq 0$ result (green) compared with the zero-disorder case (purple). This is the result of the disorder spreading some of the eigenvalues of $-\gamma + A$ closer to the imaginary axis, creating modes with smaller decay. Importantly, in neither case do we see transient amplification. By contrast, in the $\sigma = 0$ and for small enough decay, i.e., for $\gamma < |w|$, the system Eq. (2.2) exhibits very strong transient amplification. In this case, starting from the initial condition $\mathbf{x}_0 = (0, \dots, 0, 1)^{\mathsf{T}}$, the solution for the $(N - n)$th Schur component is $x_{N-n}(t) = \frac{(wt)^n}{n!} e^{-\gamma t}$ (for $0 \leqslant n \leqslant N - 1$), which is maximized at $t = n/\gamma$ with a value $\max |x_{N-n}| \sim (\frac{|w|}{\gamma})^n$ for $n \gg 1$. Thus, up to time $t \sim N/\gamma$ the norm of the activity grows exponentially; $\|\mathbf{x}(t)\|^2 \gtrsim (\frac{|w|}{\gamma})^{2\gamma t}$ for $t \lesssim N/\gamma$. For larger times the activity reaches the end of the N-long feedforward chain and starts decaying to zero; asymptotically, $\|\mathbf{x}(t)\|^2 \sim e^{-2\gamma t}$ for $t \gg N/\gamma$. However, as we have seen, the spectrum of M is extremely sensitive to perturbations; even for very small but nonzero σ, the spectrum of $-\gamma \mathbf{1} + A$ has eigenvalues with a real part as large as $|w| - \gamma$. Therefore, in the limit $N \to \infty$, the system Eq. (2.2) is unstable for $|w| > \gamma$, as soon as $\sigma \neq 0$. Conversely, in the presence of disorder (even infinitesimally small disorder in the $N \to \infty$ limit), as long as the system is stable [which from Eq. (5.15) requires $\gamma > \sqrt{|w|^2 + \sigma^2}$] it exhibits no transient amplification for
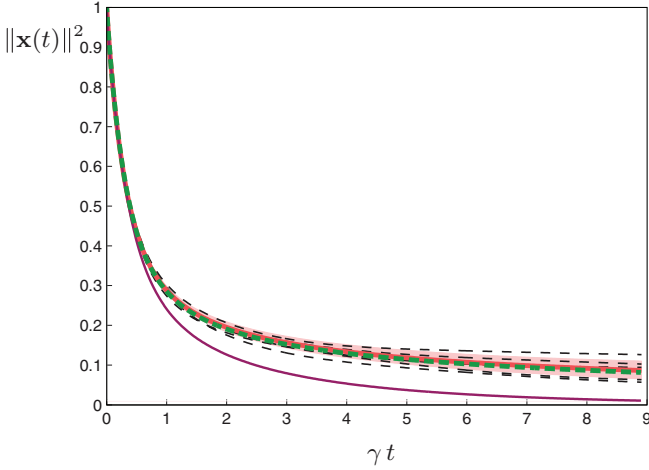
FIG. 3. (Color online) The norm squared of the response to impulse, $\|\mathbf{x}(t)\|^2$, of the system Eq. (2.2), for $A = M + \sigma J$, with binary J, and M given by Eq. (2.35) (with $\lambda_n = 0$) describing a N-long feedforward chain with uniform weights w. Here $w = 1$, $\sigma = 0.5$, $\gamma = 1.005\sqrt{\sigma^2 + w^2} \simeq 1.124$, and $N = 700$. The green (thick dashed) curve shows our result, Eq. (2.38), for the average squared impulse response, $\langle\|\mathbf{x}(t)\|^2\rangle_J$, which lies on top of the red (thick solid) curve showing the empirical average of $\|\mathbf{x}(t)\|^2$ over 100 realizations of binary J. The five thin dashed black curves show the result for five particular realizations of J, and the pink (light gray) area shows the standard deviation among the 100 realizations. The standard deviation shrinks to zero as $N \to \infty$, and $\|\mathbf{x}(t)\|^2$ for any realization lies close to its average for large N. For comparison the purple (thin, lowest) curve shows $\|\mathbf{x}(t)\|^2$ obtained by ignoring the effect of quenched disorder, i.e., by setting $A = M$.

the initial condition along the last Schur mode. Let us note, however, that as we mentioned after Eq. (2.33), Eq. (2.25) and hence Eq. (2.38) do not yield the correct answer when the direction of the impulse is optimized for the specific realization of the quenched disorder J; such disorder-tuned initial conditions can yield significant transient amplification even for the stable $\sigma \neq 0$ system.

Incidentally, we can also read the result for $M = 0$ from Eq. (2.38), by setting $w = 0$, obtaining $\langle\|\mathbf{x}(t)\|^2\rangle_J = e^{-2\gamma t} I_0(2\sigma t)$. Since in this case all directions are equivalent, this is the answer for the (normalized) initial condition along any direction, again as long as the direction is chosen independently of the specific realization of J.

Finally, the total power of response to a sinusoidal input with amplitude $\mathbf{I}_0 = (0, \ldots, 0, I_0)^\intercal$ is given by

$$\langle\overline{\|\mathbf{x}_\omega\|^2}\rangle_J = \frac{\|\mathbf{I}_0\|^2}{\omega^2 + \gamma^2 - |w|^2 - \sigma^2}. \quad (2.39)$$

The main effect of the disorder is to reduce the width of the resonance (the peak of $\langle\overline{\|\mathbf{x}_\omega\|^2}\rangle_J$ at $\omega = 0$) and increase its height. This is partly a consequence of the scattering of the eigenvalues of $-\gamma + A$ closer to the imaginary line by the disorder, creating modes with smaller decay.

2. Examples motivated by Dale's law: 1 or $N/2$ feedforward chains of length 2

In this section we consider examples motivated by Dale's law [4–6] in neurobiology. Dale's law is the observation (which holds generally but with some exceptions [46,47]) that individual neurons release the same neurotransmitter at all of their synapses. In the context of many theoretical papers including this one, it refers more specifically to the fact that an individual neuron either makes only excitatory synapses or only inhibitory synapses; that is, each column of the synaptic connectivity matrix has a fixed sign, positive for excitatory neurons and negative for inhibitory ones. We first consider two examples of connectivity matrices respecting Dale's law which take the form Eq. (2.1) with $L = \sigma^{-1}R = \mathbf{1}$ and a scalar σ. At the end of this section we consider an example with nontrivial L and R.

In the first example, we consider a matrix M, which, as we show, has a Schur form T that is composed of $N/2$ disjoint feedforward chains, each connecting only two modes (we assume N is even). For simplicity we focus on the case where all eigenvalues are zero. Thus, in the Schur basis we have

$$T = \begin{pmatrix} 0 & w_1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & w_2 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} = W \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad (2.40)$$

where we defined W to be the $N/2 \times N/2$ diagonal matrix of Schur weights $W = \mathrm{diag}(w_1, w_2, \ldots, w_{N/2})$. T in Eq. (2.40) arises as the Schur form of a mean matrix of the form

$$M = \frac{1}{2}\begin{pmatrix} K & -K \\ K & -K \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \otimes K, \quad (2.41)$$

where K is a *normal* (but otherwise arbitrary) $N/2 \times N/2$ matrix (note that M is nonetheless non-normal). The feedforward weights in Eq. (2.40) are then the eigenvalue of K. When K has only positive entries, matrices of the form Eq. (2.41) satisfy Dale's principle, and were studied in Ref. [7], in the context of networks of excitatory and inhibitory neurons. We imagine a grid of $N/2$ spatial positions, with an excitatory and an inhibitory neuron at each position. $\frac{1}{2}K$, a matrix with positive entries, describes the mean connectivity strength between spatial positions, which is taken to be identical regardless of whether the projecting or receiving neuron is excitatory or inhibitory. The sign of the weight, on the other hand, depends on the excitatory or the inhibitory nature of the projecting or presynaptic neuron; the first (last) $N/2$ columns of M represent the projections of the excitatory (inhibitory) neurons and are positive (negative). Since K is normal, it can be diagonalized by a unitary transform: $K = EWE^\dagger$, where W is as above, and $E = (\mathbf{e}_1, \mathbf{e}_2, \ldots)$ is the matrix of the orthonormal eigenvectors $\mathbf{e}_b$ of K, $b = 1, \ldots, N/2$ (with eigenvalues w_b). Then transforming to the basis

$$\left\{ \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{0} \\ \mathbf{e}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{e}_2 \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{0} \\ \mathbf{e}_2 \end{pmatrix}, \ldots, \begin{pmatrix} \mathbf{e}_{N/2} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{0} \\ \mathbf{e}_{N/2} \end{pmatrix} \right\}$$

(where $\mathbf{0}$ represents the $N/2$-dimensional vector of 0's) transforms the matrix to being block-diagonal with the 2×2

matrices

$$\frac{1}{2}\begin{pmatrix} w_b & -w_b \\ w_b & -w_b \end{pmatrix}, \quad b = 1, \ldots, N/2,$$

along the diagonal. The bth block becomes

$$\begin{pmatrix} 0 & w_b \\ 0 & 0 \end{pmatrix}$$

in its Schur basis

$$\left\{ \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbf{e}_b \\ \mathbf{e}_b \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbf{e}_b \\ -\mathbf{e}_b \end{pmatrix} \right\},$$

so the full matrix takes the form Eq. (2.40). Thus, the bth difference mode

$$\frac{1}{\sqrt{2}}\begin{pmatrix} \mathbf{e}_b \\ -\mathbf{e}_b \end{pmatrix}$$

feeds forward to the bth sum mode

$$\frac{1}{\sqrt{2}}\begin{pmatrix} \mathbf{e}_b \\ \mathbf{e}_b \end{pmatrix},$$

with weight w_b. This feedforward structure leads to a specific form of non-normal transient amplification, which the authors of Ref. [7] dubbed "balanced amplification"; small differences in the activity of excitatory and inhibitory modes feedforward to and cause possibly large transients in modes in which the excitatory and inhibitory activities are balanced.

Another interesting example of Dale's law is that in which M simply captures the differences between the mean inhibitory and mean excitatory synaptic strengths and between the numbers of excitatory and inhibitory neurons, with no other structure assumed (uniform mean connectivity), as studied in Ref. [41]. Thus, all excitatory projections have the same mean $\mu_E/\sqrt{N}$, and all inhibitory ones have the mean $-\mu_I/\sqrt{N}$. If we assume a fraction f of all neurons are excitatory, then we can write M as

$$M = \mathbf{u}\mathbf{v}^{\mathrm{T}}, \tag{2.42}$$

where $\mathbf{u} = N^{-1/2}(1, \ldots, 1)^{\mathrm{T}}$ is a unit vector, and the vector $\mathbf{v}$ has components $v_i = \mu_E$ or $v_i = -\mu_I$ for $i \leqslant fN$ and $i > fN$, respectively [for $f = 1/2$ and $\mu_E = \mu_I$, Eq. (2.42) is a special case of Eq. (2.41)]. The single-rank matrix M has only one nonzero eigenvalue given by $\mathbf{v} \cdot \mathbf{u} = \frac{1}{\sqrt{N}}\sum_i v_i$, with eigenvector $\mathbf{u}$. The case in which the excitatory and inhibitory weights are *balanced on average*, in the sense that $\sum_i v_i = 0$, is of particular interest; mathematically, it is in a sense the least symmetric and most non-normal case as $\mathbf{v} \cdot \mathbf{u} = 0$. In this case all eigenvalues of M are equal to zero. Furthermore, since in this case $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, we can readily read off the Schur decomposition of M from Eq. (2.42). The normalized Schur modes are given by $\mathbf{u}$, $\mathbf{v}/\|\mathbf{v}\|$, and $N - 2$ other unit vectors spanning the subspace orthogonal to both $\mathbf{u}$ and $\mathbf{v}$. All feedforward Schur weights are zero, except for one very large weight, equal to $\|\mathbf{v}\| \propto \sqrt{N}$, which feeds from $\mathbf{v}/\|\mathbf{v}\|$ to $\mathbf{u}$. Thus, the Schur representation of M has the form Eq. (2.40), with $w_1 = \|\mathbf{v}\| \equiv \mu\sqrt{N}$ and $w_{b \neq 1} = 0$, where we defined

$$\mu^2 \equiv \mathrm{tr}(M^\dagger M) = \|\mathbf{v}\|^2/N = f\mu_E^2 + (1-f)\mu_I^2. \tag{2.43}$$

Note that this is again a case of balanced amplification: Differences between excitatory and inhibitory activity, repre-

sented by $\mathbf{v}$, feedforward to balanced excitatory and inhibitory activity, represented by $\mathbf{u}$, with a very large weight. In the following we present results only for this balanced case of Eq. (2.42), which as just noted is a special case of Eqs. (2.40).

We start by presenting the results for the eigenvalue density. For general diagonal W in Eq. (2.40) [or equivalently, for general normal K in Eq. (2.41)], the eigenvalue density, $\rho(z)$, of $A = M + \sigma J$ is isotropic around the origin $z = 0$ and depends only on $r = |z|$. The spectral support is a disk centered at the origin. In cases in which all the weights w_b are $O(1)$, the radius of this disk can be found directly from Eq. (2.5), which yields

$$r_0 = \sigma\left[\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{\langle|w_b|^2\rangle_b}{2\sigma^2}}\right]^{1/2}. \tag{2.44}$$

Here $\langle|w_b|^2\rangle_b$ is the average of the squared feedforward weights over all blocks of Eq. (2.40); equivalently, $\langle|w_b|^2\rangle_b = 2\,\mathrm{tr}(M^\dagger M) \equiv 2\mu^2$. As long as some w_b are nonzero, r_0 is larger than the radius of the circular law, σ, with the difference an increasing function of $\langle|w_b|^2\rangle_b$; thus, the spreading of the spectrum of M (originally concentrated at the origin) after the random perturbation by σJ, is larger the more non-normal M is. In cases in which the feedforward weights of some of the 2×2 blocks of Eq. (2.40) grow without bound as $N \to \infty$, there is a corresponding singular value of $M_z \propto z - M$ for every such block which is nonzero for $z \neq 0$ but vanishes in the limit, scaling like $\sim \frac{|z|^2}{|w_b|}$, where w_b is the unbounded weight of that block [see Eq. (5.37) and its preceding paragraph]. [Note that, as stated after Eq. (2.3), we assume $\|M\|_{\mathrm{F}} = \mu = \sqrt{\langle|w_b|^2\rangle_b/2}$ is $O(1)$, so that at most $o(N)$ number of weights can be unbounded, and each can at most scale like $O(N^{1/2})$.]

In line with the general discussion after Eq. (2.12), in such cases the naive use of Eq. (2.5) may yield an area larger than the true support of $\lim_{N\to\infty}\rho(z)$; the correct support must be found by using Eqs. (2.19) and (2.20), which in this case can yield a support radius strictly smaller than Eq. (2.44).

We have calculated the explicit results for $\lim_{N\to\infty}\rho(z)$ for two specific examples of M with the Schur form Eq. (2.40). The first example belongs to the first case (bounded w_b's) where $\lim_{N\to\infty}\rho(z)$ is $\Theta(1)$ within the entire disk $r \leqslant r_0$, while the second belongs to the second case (unbounded w_b's) where the limit density is only nonzero in a proper subset of that disk.

In the first example, we take all the Schur weights in Eq. (2.40) to have the same value, which we denote by w. In this case, the eigenvalue density is given by $\rho(r) = \frac{1}{2\pi r}\frac{\partial n_<(r)}{\partial r}$, where $n_<(r)$ is the proportion of eigenvalues within a distance r from the origin and is given by

$$n_<(r) = \frac{r^2}{\sigma^2}\left[1 - \frac{|w|^2}{\sigma^2 + \sqrt{\sigma^4 + |w|^4 + 4|w|^2 r^2}}\right]. \tag{2.45}$$

$n_<(r)$ reaches unity exactly at $r = r_0$ given by Eq. (2.44), and $\rho(r)$ is $\Theta(1)$ for any smaller r. Figure 4 shows the close agreement of Eq. (2.45) with empirical results based on single binary realizations of J, for N as low as 60.

The second example is that of the balanced Eq. (2.42) with $\mathbf{u} \cdot \mathbf{v} = 0$. As we saw, all w_b are zero in this case except for
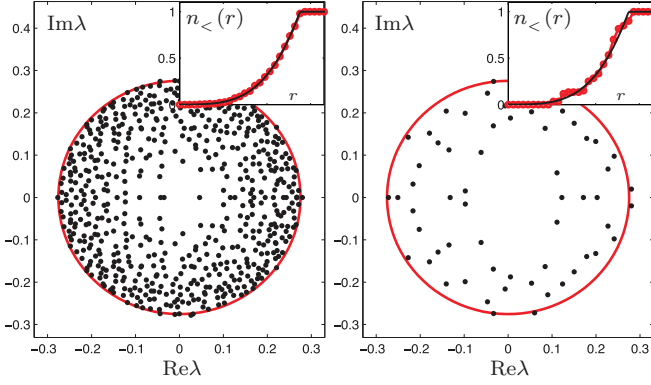
FIG. 4. (Color online) The eigenvalue spectra of $A = M + \sigma J$ for a binary J with $\sigma = 0.1$ and M given by Eq. (2.41) with $K = 1$ [corresponding to $w_b = 1$ for all the diagonal 2×2 blocks in Eq. (2.40)]. The main panels show the eigenvalues for single realizations of J, with $N = 600$ (left) and $N = 60$ (right). The red circles mark the boundaries of the spectral support, Eq. (2.44). Since A is real in this case, its eigenvalues are either exactly real, or come in complex conjugate pairs; the spectrum is symmetric under reflections about the real axis. However, such signatures of the reality of the matrix appear only as subleading corrections to the spectral density $\rho(z)$; they are finite-size effects which vanish as $N \to \infty$. The insets show a comparison of the analytic formula Eq. (2.45) (black curve) and the empirical result, based on the eigenvalues of the realizations in the main panels, for the proportion, $n_<(r)$, of eigenvalues lying within a radius r of the origin (red dots). The random fluctuations and the average bias of the empirical $n_<(r)$ are both already small for $N = 60$ and negligible for $N = 600$.

one very large, unbounded weight $w_1 = \mu \sqrt{N}$. As discussed above, in this case $M_z \propto z - M$ has an $o(1)$ smallest singular value, approximately given by $\frac{|z|^2}{\mu\sqrt{N}}$. Using Eqs. (2.19) and (2.20), we find that the support of $\lim_{N\to\infty} \rho(z)$ is the disk with radius σ [within the annulus $\sigma < |z| \leqslant r_0$ the eigenvalue density is $o(1)$], and solving Eqs. (2.8) and (2.9) for $|z| \leqslant \sigma$, we find that the spectral density is, in fact, identical with the circular law (the eigenvalue density for the $M = 0$ case), i.e.,

$$\rho(r) = \begin{cases} \frac{1}{\pi\sigma^2} + o(1) & (r < \sigma), \\ o(1) & (r > \sigma). \end{cases} \qquad (2.46)$$

It was shown in Refs. [37,48] that more generally, for any M of rank $o(N)$ and bounded $\|M\|_F$, the eigenvalue density of $A = M + \sigma J$ is given by the circular law in the limit $N \to \infty$. For single rank M (as in the present case) and a diagonal R, it was shown in Ref. [49] that the eigenvalue density of $M + JR$ agrees with that of JR as $N \to \infty$. In the present example, it was observed in Ref. [41] that even though the majority of the eigenvalues are distributed according to the circular law, there also exist a number of "outlier" eigenvalues spread outside the circle $|z| = \sigma$, which, unlike in the $M = 0$ case, may lie at a significant distance away from it (see Fig. 5). As we mentioned in Sec. II A, the noncrossing approximation cannot be trusted to correctly yield the $o(1)$ contributions to $\rho(z)$ by these outliers for $|z| > \sigma$. However, we found that if we ignore this warning and use Eqs. (2.8) and (2.9), keeping track of finite-size $o(1)$ contributions, we obtain results that agree surprisingly well
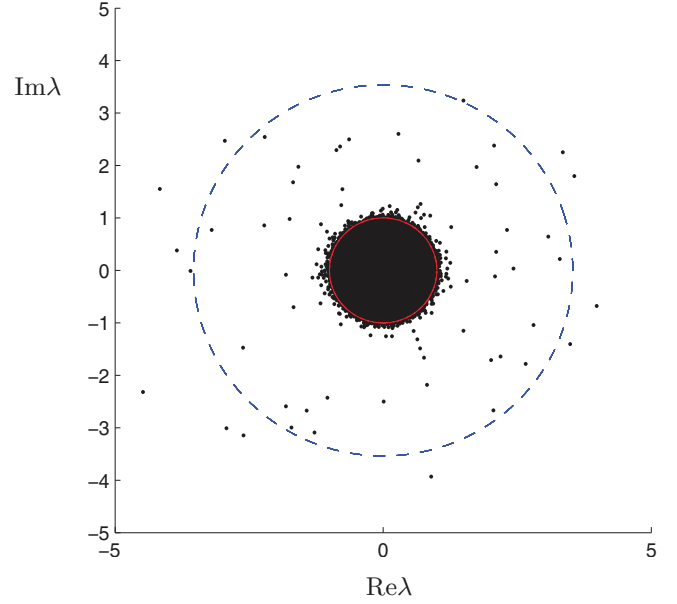

FIG. 5. (Color online) The eigenvalue spectra of $A = M + \sigma J$ for the M given by Eq. (2.42) in the balanced case, $\mathbf{v}^{\mathsf{T}}\mathbf{u} = 0$. Here $N = 800$, $\sigma = 1$, and $\mu = 12$ [see Eq. (2.43)]. The black dots are the superimposed eigenvalues of A for 20 different realizations of complex Gaussian J. The small red circle enclosing the vast majority of the eigenvalues has radius $\sigma = 1$, corresponding to the standard circular law Eq. (2.46). A $\Theta(N)$ number of eigenvalues lie within this circle. A $\Theta(\sqrt{N})$ number lie just outside of this circle in a thin boundary layer which shrinks to zero as $N \to \infty$. Finally, a $\Theta(1)$ number of eigenvalues lie at macroscopic distances outside the unit circle. The dashed blue circle shows radius r_0 given by Eq. (2.44); outliers can even lie outside this boundary.

(though not completely) with simulations. First, for the total number of outlier eigenvalues lying outside the circle $|z| = \sigma$ we obtain

$$N_>(\sigma) \equiv Nn_>(\sigma) = \sqrt{N} + O(1) \qquad (2.47)$$

[here we defined $n_>(r) = 1 - n_<(r)$ to be the proportion of eigenvalues lying outside the radius r]; see Fig. 6 for a comparison of Eq. (2.47) with simulations. The vast majority of the outlier eigenvalues counted in Eq. (2.47) lie in a narrow boundary layer immediately outside the circle $|z| = \sigma$, the width of which shrinks with growing N. In addition to these, however, there are a $\Theta(1)$ number of eigenvalues lying at macroscopic, $\Theta(1)$ distances outside the circle $|z| = \sigma$. Using Eqs. (2.8) and (2.9) we have calculated $N_>(r)$, the number of outlier eigenvalues lying outside radius r for $r > \sigma$. Figure 7 shows a plot of $N_>(r)$ and compares it with the results of simulations for different N. For roughly the inner half of the annulus $\sigma < |z| < r_0$, $N_>(r)$ agrees well with simulations, but as r increases it deviates significantly from the empirical averages. In particular, $N_>(r)$ calculated from Eqs. (2.8) and (2.9) vanishes at r_0 given by Eq. (2.44), while the empirical average of the number of outliers is nonzero well beyond r_0. Finally, we note that the distribution of these eigenvalues is not self-averaging and depends on the real vs complex nature of the random matrix J [37]. In the real case, their distribution has been recently characterized as that of the inverse roots of
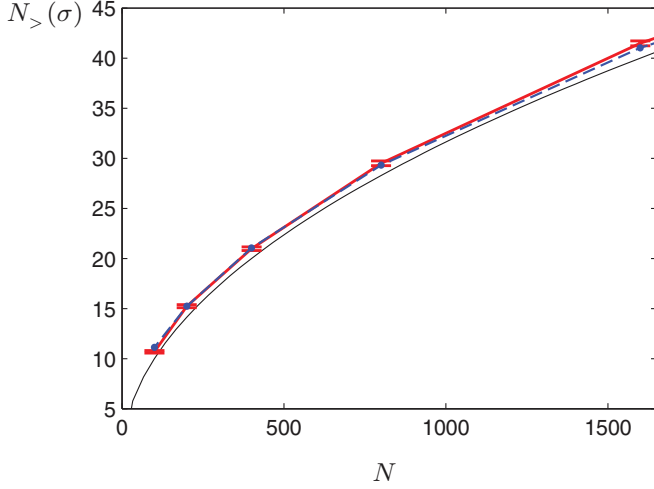
FIG. 6. (Color online) The number of eigenvalues of $M + \sigma J$, for the M given by Eq. (2.42), lying outside the circle of radius σ vs N (red line). Here $\sigma = 1, \mu = 12$, and $\mathbf{v}^{\mathrm{T}}\mathbf{u} = 0$. The numbers (red points connected by solid red lines) are obtained by numerically calculating the eigenvalues and counting the outliers for 200 realizations of J, and taking the average of the counts over all realizations, for $N = 100, 200, 400, 800, 1600$ (error bars show standard error of mean). The black dashed line plots $\sqrt{N}$ for comparison with our theoretical result Eq. (2.47); the (dashed) blue line which includes subleading corrections to $\sqrt{N}$, is obtained by numerically solving Eq. (5.42) and substituting the result in Eq. (5.43) [these formulas are, in turn, obtained from Eqs. (2.8) and (2.9) in Sec. V B].
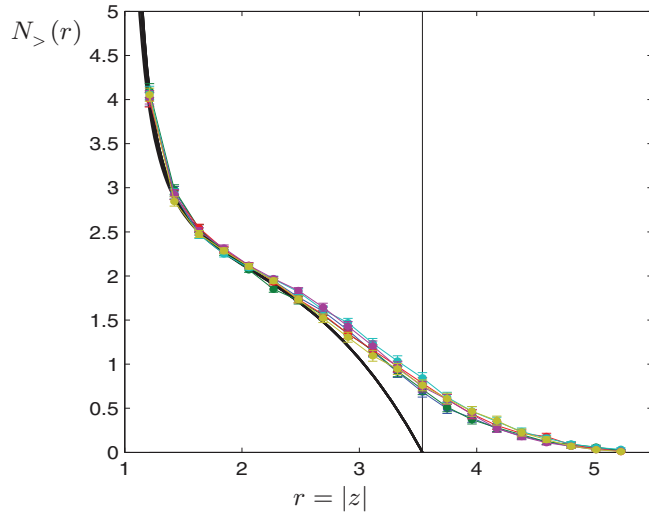

FIG. 7. (Color online) The number, $N_{>}(r)$, of outlier eigenvalues of $A = M + \sigma J$, for the M given by Eq. (2.42), lying farther from the origin than r, as a function of r. Here $\sigma = 1$, $\mu = 12$, and $\mathbf{v}^{\mathrm{T}}\mathbf{u} = 0$. The vertical line marks $|z| = r_0 \simeq 3.54$, where r_0 is given by Eq. (2.44). The colored (shades of gray) connected points are $N_{>}(r)$ for realizations of A, based on 200 samples of J, each color for a different N, for $N = 100, 200, 400, 800, 1600$, and 3200 (error bars show standard error of sample mean). Note the lack of scaling of $N_{>}(r)$ with N.
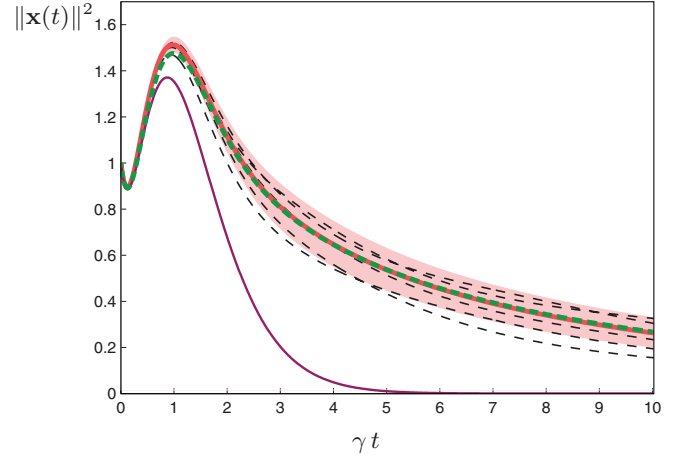

FIG. 8. (Color online) The squared norm of response to impulse, $\|\mathbf{x}(t)\|^2$, of the system Eq. (2.2), for $A = M + \sigma J$, with log-normal J, and M given by Eq. (2.40) describing $N/2$ doublet feedforward chains weights w_b. Here $w_a = \sqrt{\langle |w_b|^2 \rangle_b} = 3$, $\sigma = 0.4$, $\gamma = 1$, and $N = 1400$. The green (thick dashed) curve shows our result, Eqs. (2.48) and (2.49), for the average norm squared, which, except for a small window around its peak, lies on top of the red (thick solid) curve showing the empirical average of $\|\mathbf{x}(t)\|^2$ over 100 realizations of binary J. The five thin dashed black curves show the result for five particular realizations of J, and the pink (light gray) area shows the standard deviation among the 100 realizations. The standard deviation shrinks to zero as $N \to \infty$ and $\|\mathbf{x}(t)\|^2$ for any realization lies close to its average for large N. For comparison, the purple (thin, lowest) curve shows $\|\mathbf{x}(t)\|^2$ obtained by ignoring the effect of quenched disorder, i.e., by setting $A = M$.

a certain random power series with iid standard real Gaussian coefficients [37].

As for the dynamics, we have analytically calculated the magnitude of impulse response, Eq. (2.25), as well as the power-spectrum of steady-state response Eq. (2.33), for $A = M + \sigma J$ with M given by Eqs. (2.40) and (2.41) with general w_b, when the (impulse or sinusoidal) input feeds into the second Schur mode in one of the $N/2$ chains/blocks of Eq. (2.40); we denote the index for this block by a. For the average magnitude of impulse response we find

$$\langle \|\mathbf{x}(t)\|^2 \rangle_J = \left[\frac{1 + C_a}{2} I_0(2r_0 t) + \frac{1 - C_a}{2} J_0(2r_1 t) \right] e^{-2\gamma t},$$
(2.48)

where $J_0(x)$ [$I_0(x)$] is the [modified] Bessel function, r_0 is given by Eq. (2.44), we defined $r_1^2 = r_0^2 - \sigma^2$, and

$$C_a \equiv \frac{1 + 2|w_a|^2/\sigma^2}{\sqrt{1 + 2\langle |w_b|^2 \rangle_b/\sigma^2}},$$
(2.49)

with $\langle |w_b|^2 \rangle_b = 2\,\mathrm{tr}\,(M^\dagger M)$ denoting the average squared feedforward weight among all the blocks of Eq. (2.40). In Fig. 8 we plot Eq. (2.48) and compare it with the result obtained by ignoring the disorder (i.e., by setting $\sigma = 0$); in the latter case, the block a is decoupled from the rest of the network, and solving the 2×2 linear system governed by the matrix

$$\begin{pmatrix} -\gamma & w_a \\ 0 & -\gamma \end{pmatrix},$$

we obtain $\|\mathbf{x}(t)\|^2 = (1 + w_a^2 t^2)e^{-2\gamma t}$. From the figure, we see that the $\sigma \neq 0$ result (green) has a slower asymptotic decay compared with the zero-disorder case (purple); this is due to the disorder having spread some eigenvalues closer to the imaginary axis, creating modes with smaller decay, along with the fact that the coupling between the 2×2 blocks induced by the disorder ensures that these more slowly decaying modes will be activated. Indeed, for large t, $\|\mathbf{x}(t)\|$ decays like $e^{-\gamma t}$ when $\sigma = 0$, while in the $\sigma > 0$ case, based on Eq. (2.44) it must decay like $e^{-(\gamma-r_0)t}$, i.e., by a rate set by the largest real part of the spectrum shifted by $-\gamma$ [this is indeed what we obtain from Eq. (2.48) using the asymptotics of Bessel functions]. In addition, both curves exhibit transient amplification where the magnitude of activity initially grows to a maximum, before it decays asymptotically to zero. The $\sigma \neq 0$ curve shows larger and longer transient amplification, which is most likely attributable both to the eigenvalues being closer to the $\text{Re}(z) = \gamma$ line and to augmented non-normal effects (e.g., larger effective feedforward weights, or longer chains).

We also mention that, as in our previous examples, if the input direction is optimized for the particular realization of J, significantly larger transient amplification may be achieved.

Finally, the total power spectrum of response to a sinusoidal input, Eq. (2.32), is given by the explicit formula

$$\langle \overline{\|\mathbf{x}_\omega\|^2} \rangle_J = \frac{\omega^2 + \gamma^2 + |w_a|^2}{(\omega^2 + \gamma^2)^2 - \sigma^2(\omega^2 + \gamma^2 + \mu^2)} \|\mathbf{I}_0\|^2, \quad (2.50)$$

where $\mu^2 \equiv \text{tr}(M^\dagger M) = \langle |w_b|^2 \rangle_b/2$ and, as noted above, the direction of $\mathbf{I}_0$ is that of the second Schur mode in block a.

The example Eq. (2.42) motivated by Dale's law with neurons of either excitatory or inhibitory types, can be generalized to a network of neurons belonging to one of C different types (these could be subtypes of excitatory or inhibitory neurons), in which not only the mean but also the variance of connection strengths depends on the pre- and postsynaptic types. When this dependence is factorizable, in a way we now describe, the connectivity matrix of such a network will be of the form Eq. (2.1) with nontrivial L and R. Let $c(i) \in \{1, \ldots, C\}$ denote the type of neuron i, and let f_c denote the fraction of neurons of type c (so $\sum_{c=1}^C f_c = 1$); we assume C and f_c are all $\Theta(1)$. Assume further that each synaptic weight is a product of a pre- and a postsynaptic factor and that in each synapse these factors are chosen independently from the same distribution, except for a deterministic sign and overall scale that depend only on the type of the pre- and postsynaptic neurons, respectively. Thus, if A_{ij} denotes the weight of the synaptic projection from neuron j to neuron i, we have

$$A_{ij} = \frac{1}{\sqrt{N}}(l_{c(i)}x_{ij})(r_{c(j)}y_{ij}), \quad (2.51)$$

where x_{ij}'s and y_{ij} are positive random variables chosen iid from the distributions $P_x(x)$ and $P_y(y)$, respectively. Here l_c and r_c determine the sign and the scale (apart from the overall $\frac{1}{\sqrt{N}}$) of the pre- and postsynaptic factors of the neurons in cluster c, respectively. Note that when all l_c are positive, A_{ij} satisfies Dale's law. By absorbing appropriate constants into l_c's and r_c's we can assume that $\text{Var}[xy] =$

$\langle x^2 \rangle \langle y^2 \rangle - \langle x \rangle^2 \langle y \rangle^2 = 1$. Then it is easy to see that A can be cast in the form Eq. (2.1) with

$$L_{ij} = l_{c(i)}\delta_{ij}, \quad (2.52)$$

$$R_{ij} = r_{c(i)}\delta_{ij}, \quad (2.53)$$

$$J_{ij} = \frac{1}{\sqrt{N}}(x_{ij}y_{ij} - \xi), \quad (2.54)$$

$$M = s L \mathbf{u}\mathbf{u}^\mathsf{T} R, \quad (2.55)$$

where $\mathbf{u}$ is the unit vector $\frac{1}{\sqrt{N}}(1, \ldots, 1)^\mathsf{T}$,

$$s \equiv \xi \sqrt{N}, \quad (2.56)$$

and $\xi \equiv \langle x \rangle \langle y \rangle$ is dimensionless and $\Theta(1)$ [note that J, given by Eq. (2.54), indeed has iid elements with zero mean and variance N^{-1}]. Being single rank, M has $N - 1$ zero eigenvalues; its only (potentially) non-null eigenvector is $L\mathbf{u}$, with a generically large eigenvalue

$$\lambda_M = s\mathbf{u}^\mathsf{T} RL\mathbf{u} = s\frac{1}{N}\sum_{i=1}^N r_{c(i)}l_{c(i)} = \xi\sqrt{N}\langle\sigma_c\rangle_c, \quad (2.57)$$

where we defined

$$\sigma_c \equiv l_c r_c, \quad (2.58)$$

$$\langle X_c \rangle_c \equiv \sum_{c=1}^C f_c X_c. \quad (2.59)$$

As for the example Eq. (2.42), we focus on the balanced case in which $\lambda_M \propto \langle\sigma_c\rangle_c = 0$. From Eq. (2.55), $M = \tilde{\mathbf{u}}\tilde{\mathbf{v}}^\mathsf{T}$ with $\tilde{\mathbf{u}} = L\mathbf{u}$ and $\tilde{\mathbf{v}} = sR\mathbf{u}$. The balanced condition is equivalent to $\tilde{\mathbf{u}} \cdot \tilde{\mathbf{v}} = 0$ [see Eq. (2.57)]. Thus, similar to Eq. (2.42), the Schur representation of M has the form (2.40), with $w_1 = \|\tilde{\mathbf{u}}\|\|\tilde{\mathbf{v}}\|$ and $w_b = 0$ for $b > 1$.

In Sec. V C we prove that, as for Eq. (2.42), for the ensemble Eqs. (2.52)–(2.55) the limit of the eigenvalue distribution, $\lim_{N\to\infty} \rho(z)$, is also not affected by the nonzero mean matrix Eq. (2.55); hence, we can obtain $\lim_{N\to\infty} \rho(z)$ for that example by safely setting M to zero, and using formulas, Eqs. (2.13)–(2.16), with L and R given by Eqs. (2.52) and (2.53). Thus, $\lim_{N\to\infty} \rho(z)$ is isotropic and its support is the disk with radius

$$r_0 = \|RL\|_F = \sqrt{\langle\sigma_c^2\rangle_c}. \quad (2.60)$$

As in the previous example, when the balance condition $\langle\sigma_c\rangle_c = 0$ holds, use of the naive formula Eq. (2.5) with $M = \tilde{\mathbf{u}}\tilde{\mathbf{v}}^\mathsf{T}$ would have yielded

$$\tilde{r}_0 = r_0\left[\frac{1}{2} + \sqrt{\frac{1}{4} + \xi^2}\right]^{1/2}, \quad (2.61)$$

which is larger than the correct result Eq. (2.60). As discussed above, this result is not correct, but it indicates the existence of $\Theta(1)$ number of outlier eigenvalues lying outside the boundary of $\lim_{N\to\infty} \rho(z)$ given by Eq. (2.60). For $r < r_0$, the $N \to \infty$ limit of the proportion, $n_>(r)$, of eigenvalues lying farther than

distance r of the origin is given by $g^2(r)$ which is found by solving Eq. (2.15), or, equivalently,

$$\left\langle \frac{1}{g^2 + \sigma_c^{-2} r^2} \right\rangle_c = 1. \tag{2.62}$$

The results, Eqs. (2.17) and (2.18), also hold, wherein the normalized sums over i can be replaced with appropriate averages $\langle \cdot \rangle_c$. In the case of two neuronal types a closed solution can be obtained for $n_>(r)$ and $\rho(r)$. Identifying the two types with excitatory and inhibitory neurons, and assuming that $l_c = 1$, $\sigma_E \equiv \sigma_1 > 0$, and $\sigma_I \equiv \sigma_2 < 0$ (we use E and I as indices instead of $c = 1$ and 2, respectively) the ensemble Eqs. (2.52)–(2.55) describes a synaptic connectivity matrix in which all excitatory (inhibitory) connections are iid with mean $\xi \sigma_E N^{-\frac{1}{2}}$ ($-\xi |\sigma_I| N^{-\frac{1}{2}}$) and variance $\sigma_E^2 N^{-1}$ ($\sigma_I^2 N^{-1}$). In this case, Eq. (2.62) yields a quadratic equation. Differentiating the solution of that equation with respect to r^2 we obtain the explicit result

$$\rho(r) = \frac{\sigma_E^{-2} + \sigma_I^{-2}}{2\pi} \left\{ 1 - \frac{\frac{(\sigma_E^{-2} + \sigma_I^{-2}) r^2 - 1}{2} + \frac{r_0^2 - 2r^2}{\sigma_E^2 + \sigma_I^2}}{\sqrt{\frac{\left[(\sigma_E^{-2} + \sigma_I^{-2}) r^2 - 1 \right]^2}{4} + \frac{r^2 (r_0^2 - r^2)}{(\sigma_E \sigma_I)^2}}} \right\}. \tag{2.63}$$

This result was first obtained (in a less simplified form) in Ref. [41]. Figure 9 shows two examples of spectra for single realizations of matrices of the form Eq. (2.51), with three neural types ($C = 3$), where x_{ij} and y_{ij}, and hence J_{ij}, have log-normal distributions. The insets compare $n_>(r)$ based on the numerically calculated eigenvalues, with those found by solving Eq. (2.62). In the right panel, the normally distributed $\log J_{ij}$ have a higher standard deviation, and hence the distribution of J_{ij} has a heavier tail. The right panel's inset demonstrates that the convergence to the universal, $N \to \infty$ limit can be considerably slow when the distribution of J_{ij} is heavy tailed.

3. Linearizations of nonlinear neural and ecological networks

In neuroscience applications, Eq. (2.2) can arise as a linearization of nonlinear firing rate equations for a recurrent neural network of N neurons, around some stationary background. The nonlinear dynamical equations for the evolution of the network activity typically take the form [50]

$$T \frac{d\mathbf{v}(t)}{dt} = -\mathbf{v}(t) + W f(\mathbf{v}(t)) + \mathbf{I}^v(t). \tag{2.64}$$

Here $\mathbf{v}(t)$ is the vector of state variables of all neurons at time t; its ith component, $v_i(t)$, is commonly thought of as the voltage of the ith neuron, or the total synaptic input it receives. $f(\cdot)$ is the neuronal nonlinear input-output function, which is imposed element by element on its vector argument, with $f(\mathbf{v})_i \equiv f(v_i)$ giving the output, i.e., the firing rate, of neuron i; $\mathbf{I}^v(t)$ is the external input vector; $T = \mathrm{diag}(\tau_1, \tau_2, \ldots, \tau_N)$ is a $N \times N$ diagonal matrix whose diagonal elements are the positive time constants of the neurons (hence T is invertible); and W is the $N \times N$ synaptic connectivity matrix.

Suppose that for a constant external input, $\mathbf{I}_*^v$, Eq. (2.64) has a fixed point $\mathbf{v}_*$. Then, given a small perturbation in the
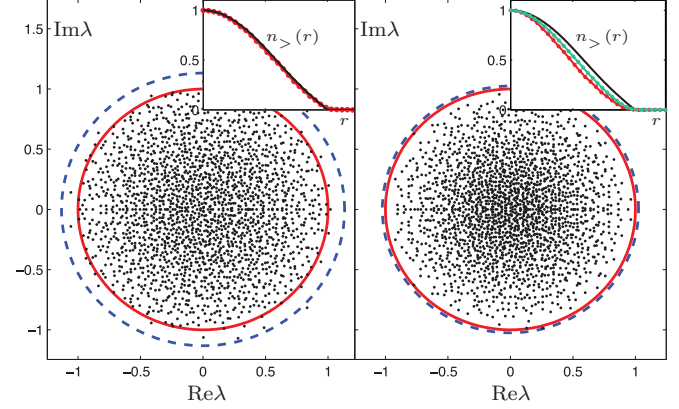

FIG. 9. (Color online) The eigenvalue spectra of $A = M + LJR$ with M, L, and R given by Eqs. (2.52)–(2.55) with neurons belonging to one of three different types ($C = 3$). The main panels show the eigenvalues for two particular realizations of J. In both panels, $N = 2000$, $f_1 = 0.6$, $f_2 = f_3 = 0.2$, $l_c = 1$, $\sigma_1 = r_1 = 0.76$, $\sigma_2 = r_2 = -0.57$, $\sigma_3 = r_3 = -1.71$ (so $\langle \sigma_c \rangle_c = 0$ and $r_0^2 = \langle \sigma_c^2 \rangle_c = 1$), and J_{ij} had real entries with log-normal distribution; in the left (right) panel, the normally distributed $\log_{10} J_{ij}$ had standard deviation 0.5 (0.75). The solid red circles mark the boundaries of the spectral support as given by Eq. (2.60), and the dashed blue circles show the radii given by Eq. (2.61). The insets compare $n_>(r)$ based on the numerically calculated eigenvalues shown in the main panels (connected red dots), with that found by solving Eq. (2.62) (black curve). In the right panel's inset we have also plotted (green connected dots lying slightly above the red connected circles) the empirically calculated $n_>(r)$ for a single realization with the same ensemble parameters, but with $N = 8000$; the convergence to the universal limit at $N \to \infty$ is significantly slower in the right panel in which the distribution of J_{ij} had a considerably heavier tail.

input, $\mathbf{I}^v(t) = \mathbf{I}_*^v + \delta \mathbf{I}^v(t)$, we can write $\mathbf{v}(t) = \mathbf{v}_* + \mathbf{x}(t)$, and linearize the dynamics around the fixed point by expanding Eq. (2.64) to first order in $\mathbf{x}(t)$ and $\delta \mathbf{I}^v(t)$. This yields the set of linear differential equations

$$T \frac{d\mathbf{x}(t)}{dt} = -\mathbf{x}(t) + W \mathbf{\Phi} \, \mathbf{x}(t) + \delta \mathbf{I}^v(t), \tag{2.65}$$

for the (small) deviations, where we defined the diagonal Jacobian

$$\mathbf{\Phi} = \mathrm{diag}[f'(\mathbf{v}_*)]. \tag{2.66}$$

Now suppose that the original connectivity matrix can be written as $W = \langle W \rangle + \delta W$, with a quenched disorder part that is an iid random matrix: $\delta W = \sigma J$. Then multiplying Eq. (2.65) by T^{-1}, we can convert Eq. (2.65) into the form Eq. (2.2), with $\gamma = 0$, and $A = M + LJR$, with

$$M = T^{-1}(-\mathbf{1} + \langle W \rangle \mathbf{\Phi}), \tag{2.67}$$

$$L = T^{-1}, \tag{2.68}$$

$$R = \sigma \mathbf{\Phi}, \tag{2.69}$$

and input

$$\mathbf{I}(t) = T^{-1} \delta \mathbf{I}^v(t). \tag{2.70}$$

This observation is not limited to neuroscience applications and can also apply to many other frameworks, e.g., those used in mathematical biology. Generalized Lotka-Volterra (GLV) equations [52] used in modeling the dynamics of food webs provide an example. Let $\boldsymbol{n}(t) = [n_1(t), \ldots, n_N(t)]^{\mathrm{T}}$ denote the vector of population sizes of N species. The GLV equations take the form $\frac{dn_i}{dt} = n_i(r_i + \sum_j W_{ij}n_j)$ or

$$\frac{d\boldsymbol{n}}{dt} = \mathrm{diag}(\boldsymbol{r} + W\boldsymbol{n})\boldsymbol{n}, \qquad (2.71)$$

where $r_i > 0$ are the species' intrinsic growth rates and W is the interaction matrix. Linearizing Eq. (2.71) around a fixed point, $\boldsymbol{n}_*$, yields again a linear system of the form Eq. (2.2), with $\gamma = \mathbf{I}(t) = 0$. Starting with the same simple model $W = \langle W \rangle + \sigma J$, we find that A can be written in the form Eq. (2.1), with

$$R = \sigma \mathbf{1}, \quad L = \mathrm{diag}(\boldsymbol{n}_*), \qquad (2.72)$$

$$M = \mathrm{diag}(\boldsymbol{r} + W\boldsymbol{n}_*) + L\langle W \rangle. \qquad (2.73)$$

Note that if no species is extinct in the fixed point, i.e., if all $n_{i*} > 0$, then $M = L\langle W \rangle$.

Assuming that the linear systems thus obtained, i.e., the fixed points $\boldsymbol{v}_*$ or $\boldsymbol{n}_*$, are stable, we can therefore think of our results for $\|\mathbf{x}(t)\|^2$ and $\|\mathbf{x}_\omega\|^2$ as characterizing the temporal evolution and the spectral properties of the linear response of the *nonlinear* system Eq. (2.64) [Eq. (2.71)] in its fixed point $\boldsymbol{v}_*$ ($\boldsymbol{n}_*$) to perturbations.

The necessary and sufficient condition for the stability of a fixed point (without any change in the external input) is that all eigenvalues of the corresponding A have negative real parts. Our formula for the boundary of the eigenvalue distribution, Eq. (2.5), can be applied in these cases to map out the region in parameter space (parameters here mean the time constants or intrinsic growth rates in T or $\boldsymbol{r}$, or the connectivity parameters determining the random ensemble for W, i.e., σ and the parameters of $\langle W \rangle$) in which a particular fixed point is stable. Recently, our general formula Eq. (2.5) was used in this way by colleagues [53] to determine the phase diagram of a clustered network of neurons, in which intracluster connectivity is large, but intercluster connectivity is random and weak. Because of the strong intracluster connectivity, each cluster behaves as a unit with a single self-coupling a. Letting the random intercluster couplings between N clusters have zero mean and variance g^2/N, their analysis starts from the equation

$$\frac{d\boldsymbol{v}(t)}{dt} = -\boldsymbol{v}(t) + a \tanh[\boldsymbol{v}(t)] + g J \tanh[\boldsymbol{v}(t)], \qquad (2.74)$$

where J is an iid random matrix as above. Here $\boldsymbol{v}$ is a vector whose ith component is the mean voltage of *cluster i*, while the nonlinear function $\tanh[\boldsymbol{v}(t)]$ (with the hyperbolic tangent acting componentwise) represents the vector of mean firing rates of the clusters. The analysis of Ref. [53] shows that there is a region of the phase plane (a,g) where the self-connectivity, a, is excitatory and sufficiently strong, in which the system eventually relaxes to nonzero random attractor fixed points $\boldsymbol{v}_*$; for smaller values of a, the dynamics is chaotic (chaos in the $a = 0$ case was established in Ref. [54]). The form of these fixed points [the distribution of the elements of $\boldsymbol{v}_*$ as $N \to \infty$
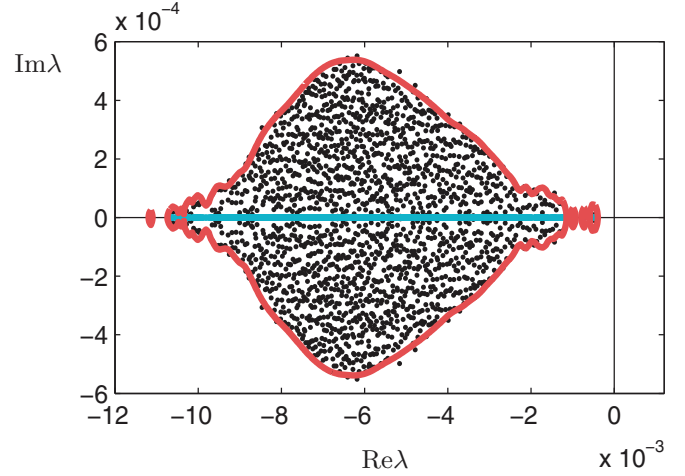

FIG. 10. (Color online) The eigenvalues (black dots) of $A = M + JR$, with M and R given by Eqs. (2.75) and (2.76) with $g = 0.01$, $a = 1.02$, and $N = 2000$. This matrix governs the dynamics of small perturbations away from a nontrivial random fixed point in a clustered network of neurons [see Eq. (2.74)], studied in Ref. [53]. The cyan dots on the real line are the eigenvalues of M, and the red curve is the boundary of support of the eigenvalue distribution, as calculated numerically from Eq. (2.5).

for a given (a,g)] can be obtained using mean-field theory, and the linearization about $\boldsymbol{v}_*$ leads to an equation in the form of Eq. (2.2), with $A = M + JR$, where M and R are the diagonal matrices

$$M = \mathrm{diag}[-1 + a \tanh'(\boldsymbol{v}_*)], \qquad (2.75)$$

$$R = \mathrm{diag}[g \tanh'(\boldsymbol{v}_*)]. \qquad (2.76)$$

Given this form, it can be shown that the fixed point $\boldsymbol{v}_*$ is stable if $z = 0$ is outside and to the right of the spectrum of the Jacobian matrix of the linearization, A. The mean field solution for $\boldsymbol{v}_*$ determines the statistics of the elements of $R^2 M^{-2}$ for a given (a,g). From these it can be determined if $z = 0$ is outside the spectrum using our formula for the boundary of spectrum, Eq. (2.5), which yields the requirement $\mathrm{tr}\left(\frac{R^2}{M^2}\right) < 1$. In this way, the region of stability of the fixed points in the (a,g) plane can be mapped (see Ref. [53] for the results and a complete discussion of the analysis outlined here). Figure 10 shows a numerical example of the eigenvalue distribution for A for a given (a,g) and the superimposed boundary calculated using Eq. (2.5).

In closing we note a potential caveat in the applicability of our formulas to the linearization analysis of systems like Eqs. (2.65) and (2.71). We have derived the general formulas of Secs. II A and II B assuming that M, L, and R are independent of J. However, M and R as given by Eqs. (2.67) and (2.69) [or M and L in Eqs. (2.72) and (2.73)] depend on J via their dependence on $\boldsymbol{v}_*$ ($\boldsymbol{n}_*$). However, in our experience this dependence is often too weak and indirect to render our formulas inapplicable; an example is provided by the excellent agreement of the empirical spectrum and the red boundary given by our formula in Fig. 10, which also held for other parameter choices of the model of Ref. [53].

3. DERIVATION OF THE FORMULA FOR THE SPECTRAL DENSITY

In this section we derive the formulas, Eqs. (2.5)–(2.8), for the average spectral density, $\rho(z)$, of random matrices of the form $A = M + LJR$, where M, L, and R are deterministic matrices, and J is random with iid elements of zero mean and variance $1/N$. We use the Hermitianized diagrammatic method developed in Refs. [24,25] (and reviewed in Ref. [55]), which we recapitulate here for completeness. As mentioned in Sec. II, the spectral density is self-averaging for large N. Furthermore, as established in Ref. [35], it is also universal in the large N limit, in the sense that it is independent of the details of the distribution of the elements of J as long its mean and variance are as stated. The same universality theorem also ensures that the real or complex nature of J does not, by itself, affect $\rho(z)$ to leading order. Therefore, for simplicity we consider the case where J is a zero-mean complex Gaussian random matrix with $\langle J_{ab}J_{cd}\rangle = 0$, and

$$\langle J_{ab}J_{cd}^*\rangle = \frac{1}{N}\delta_{ac}\delta_{bd}. \qquad (3.1)$$

Thus, $\langle|J_{ab}|^2\rangle = \frac{1}{N}$, and all other first and second moments of J (including $\langle J_{ab}^2\rangle$) vanish. The measure on J can be written as

$$d\mu(J) \propto e^{-N\mathrm{Tr}(JJ^\dagger)}\prod_{ab}d\mathrm{Im}J_{ab}d\mathrm{Re}J_{ab}. \qquad (3.2)$$

In this form, and by the invariance of the trace, it is clear that the measure is symmetric with respect to the group $U(N) \otimes U(N)$, acting on J by $J \mapsto UJV^\dagger$, where U and V are arbitrary $N \times N$ unitary matrices.

For a particular realization of J, we define the "Green's function" $G(z; J)$ by

$$G(z;J) \equiv \frac{1}{M_z - J}, \qquad (3.3)$$

where $M_z = L^{-1}(z - M)R^{-1}$ [Eq. (2.6)]. In the case $L, R \propto \mathbf{1}$, $G(z; J)$ will be proportional to the resolvent of A, $\frac{1}{z-A}$. More generally, we have

$$\frac{1}{z - A} = R^{-1}G(z;J)L^{-1}. \qquad (3.4)$$

Following Ref. [24], we use the identity

$$\delta^2(z) = \frac{1}{\pi}\partial_{\bar{z}}\partial_z \ln|z|^2 = \frac{1}{\pi}\partial_{\bar{z}}\left(\frac{1}{z}\right), \qquad (3.5)$$

where the first identity follows by noting that $4\partial_{\bar{z}}\partial_z = \nabla^2$, where ∇^2 is the two-dimensional (2D) Laplacian, and recalling from electrostatics that the solution of Poisson's equation for a point charge at origin, i.e., $\nabla^2\phi(z) = 4\pi\delta^2(z)$, in 2D is given by the potential field $\phi(z) = \ln|z|^2$; the second identity follows from $\partial_z \ln|z|^2 = \partial_z(\ln z + \ln\bar{z}) = \frac{1}{z} + 0$. Using Eq. (3.5), we can write the empirical spectral density, defined in Eq. (2.4), as

$$\rho_J(z) = \frac{1}{\pi}\partial_{\bar{z}}\frac{1}{N}\sum_\alpha \frac{1}{z - \lambda_\alpha} = \frac{1}{\pi}\partial_{\bar{z}}\,\mathrm{tr}\,\frac{1}{z - A}. \qquad (3.6)$$

Performing the ensemble average we obtain

$$\rho(z) \equiv \langle\rho_J(z)\rangle_J = \frac{1}{\pi}\partial_{\bar{z}}\,\mathrm{tr}\,[(RL)^{-1}\langle G(z;J)\rangle_J], \qquad (3.7)$$

where we used Eq. (3.4) and the linearity and cyclicity of the trace. Thus, to calculate $\rho(z)$, our task boils down to calculating $\langle G(z;J)\rangle_J$.

The diagrammatic technique provides a method for calculating averages of products of $G(z; J)$'s. However, this method in its standard form relies on A being a Hermitian matrix. It starts by an expansion of $G(z; J)$ in powers of J, which is only valid when z is far enough from the spectrum of A, i.e., away from the points in which we are most interested. For Hermitian matrices, this is no problem as the spectrum is confined to the real line, and therefore $G(z; J)$ and $\langle G(z; J)\rangle_J$ will be analytic outside the real line. Thus, one can use the expansion for z far away outside the real line, perform the averaging over J, and sum up the most dominant contributions to obtain a result analytic in z. This result can then be analytically continued to z arbitrarily close to the spectrum on the real line, yielding information about the spectrum. All this would seemingly fail in the case of a non-normal (and in particular non-Hermitian) A, with eigenvalues that, in general, cover a 2D region in the complex plane. However, using a trick introduced by Ref. [24], we can turn this problem to an auxiliary problem of averaging the Green's functions for a Hermitian matrix. By doubling the degrees of freedom, one defines a z-dependent $2N \times 2N$ Hermitian "Hamiltonian,"

$$H(z) \equiv \begin{pmatrix} 0 & M_z - J \\ M_z^\dagger - J^\dagger & 0 \end{pmatrix}, \qquad (3.8)$$

and the corresponding $2N \times 2N$ resolvent matrix or Green's function depending on a new complex variable η:

$$\boldsymbol{G}(\eta,z;J) \equiv [\eta - H(z)]^{-1} \qquad (3.9)$$
$$= -\begin{pmatrix} \frac{\eta}{(M_z-J)(M_z-J)^\dagger-\eta^2} & \frac{M_z-J}{(M_z-J)^\dagger(M_z-J)-\eta^2} \\ \frac{(M_z-J)^\dagger}{(M_z-J)(M_z-J)^\dagger-\eta^2} & \frac{\eta}{(M_z-J)^\dagger(M_z-J)-\eta^2} \end{pmatrix}.$$

For $\eta \to i0$, we see that

$$\boldsymbol{G}(0,z;J) = -\begin{pmatrix} 0 & (M_z - J)^{-\dagger} \\ (M_z - J)^{-1} & 0 \end{pmatrix}, \qquad (3.10)$$

and thus from Eq. (3.3), for any realization of J,

$$G(z;J) = -\lim_{\eta\to i0}\boldsymbol{G}^{21}(\eta,z;J). \qquad (3.11)$$

Here we have used the notation

$$\boldsymbol{G}(\eta,z;J) = \begin{pmatrix} \boldsymbol{G}^{11}(\eta,z;J) & \boldsymbol{G}^{12}(\eta,z;J) \\ \boldsymbol{G}^{21}(\eta,z;J) & \boldsymbol{G}^{22}(\eta,z;J) \end{pmatrix}, \qquad (3.12)$$

where $\boldsymbol{G}^{\alpha\beta}$ (with $\alpha,\beta \in \{1,2\}$) are $N \times N$ matrices, forming the four blocks of $\boldsymbol{G}$. We have written the limit in Eq. (3.11) as $\eta \to i0$ to emphasize that until the end of our calculations η is to retain a nonzero imaginary part, which serves to regularize the denominators in Eq. (3.9); cf. the discussion after Eq. (3.35). We carry out a perturbation expansion in powers of J, so we decompose the Hamiltonian according to

$$H(z) = H_0(z) - J, \qquad (3.13)$$

$$J \equiv \begin{pmatrix} 0 & J \\ J^\dagger & 0 \end{pmatrix}, \quad H_0(z) \equiv \begin{pmatrix} 0 & M_z \\ M_z^\dagger & 0 \end{pmatrix}. \quad (3.14)$$

We sometimes use a tensor product notation to denote matrices in this doubled up space, e.g., writing $J = \sigma^+ \otimes J + \sigma^- \otimes J^\dagger$, where we defined the 2×2 matrices

$$\sigma^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \sigma^- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \quad (3.15)$$

By a slight abuse of notation we also denote $2N \times 2N$ matrices $\sigma^\pm \otimes \mathbf{1}_{N \times N}$ by $\sigma^\pm$, and we denote the identity matrix in any space by $\mathbf{1}$. From Eqs. (3.11) we obtain $\mathrm{tr}\,[(RL)^{-1}G(z;J)] = -\mathrm{tr}\,\{[\sigma^+ \otimes (RL)^{-1}]\, \boldsymbol{G}(i0^+,z;J)\}$, and from Eq. (3.7)

$$\rho(z) = -\lim_{\eta \to i0} \frac{1}{\pi} \partial_{\bar{z}} \,\mathrm{tr}\,\{[\sigma^+ \otimes (RL)^{-1}]\mathcal{G}(\eta,z)\} \quad (3.16)$$

$$= -\lim_{\eta \to i0} \frac{1}{\pi} \partial_{\bar{z}} \,\mathrm{tr}\,[(RL)^{-1}\mathcal{G}^{21}(\eta,z)], \quad (3.17)$$

where we defined

$$\mathcal{G}(\eta,z) \equiv \langle \boldsymbol{G}(\eta,z;J)\rangle_J. \quad (3.18)$$

Having expressed $\rho(z)$ in terms of the ensemble average of the Green's function for a Hermitian matrix, we now develop the diagrammatic method for calculating ensemble averages of products of $\boldsymbol{G}(\eta,z;J)$ [including $\mathcal{G}(\eta,z)$]. Note that, being the Green's function of a Hermitian matrix, $\boldsymbol{G}(\eta,z;J)$ and hence $\mathcal{G}(\eta,z) = \langle \boldsymbol{G}(\eta,z;J)\rangle_J$ are analytic functions of η for η outside the real line, and therefore analytic continuation can be used to take the limit $\eta \to i0$ after obtaining the average over J for η sufficiently away from the real line.

We denote the elements of a generic $2N \times 2N$ matrix A by $A_{ab}^{\alpha\beta}$, where the greek indices range in $\{1,2\}$ and the latin indices range in $\{1, \ldots, N\}$. Using this notation, the definition Eq. (3.14), and Eq. (3.1), we can write the covariance for the components of J as

$$\langle J_{ab}^{\alpha\beta} J_{cd}^{\gamma\delta}\rangle_J = \frac{1}{N} \delta_{ad}\delta_{bc}(\sigma_{\alpha\beta}^+ \sigma_{\gamma\delta}^- + \sigma_{\alpha\beta}^- \sigma_{\gamma\delta}^+) \quad (3.19)$$

(the terms proportional to $\sigma^+\sigma^+$ and $\sigma^-\sigma^-$ involve $\langle J_{ab}J_{cd}\rangle$, or its complex conjugate, which vanish for the complex Gaussian ensemble). It will be more handy to rewrite the parentheses on the right side of Eq. (3.19) as $\pi_{\alpha\delta}^1 \pi_{\gamma\beta}^2 + \pi_{\alpha\delta}^2 \pi_{\gamma\beta}^1$, where

$$\pi^1 \equiv \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \pi^2 \equiv \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad (3.20)$$

yielding

$$\langle J_{ab}^{\alpha\beta} J_{cd}^{\gamma\delta}\rangle_J = \frac{1}{N} \sum_{r=1}^{2} \left(\pi_{\alpha\delta}^r \delta_{ad}\right)\left(\pi_{\gamma\beta}^{3-r} \delta_{cb}\right). \quad (3.21)$$

Also, since J_{ab} have zero mean, we have $\langle J\rangle_J = 0$.

The starting point of the diagrammatic method is the perturbation expansion of $\boldsymbol{G}(\eta,z;J) = [\eta - H_0(z) - J]^{-1}$ in powers of J,

$$\boldsymbol{G}(\eta,z;J) = \boldsymbol{G}(\eta,z;0) \sum_{n=0}^{\infty} [J\boldsymbol{G}(\eta,z;0)]^n, \quad (3.22)$$
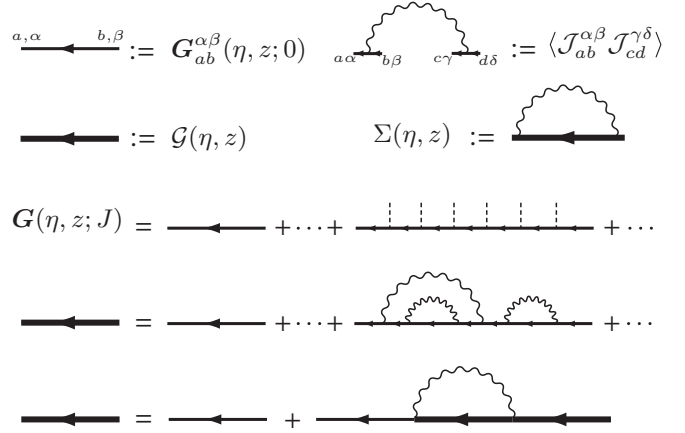

FIG. 11. The first two lines define different elements of Feynman diagrams: the Green's function for $J = 0$ (zero disorder), $\boldsymbol{G}_{ab}^{\alpha\beta}(\eta,z;0)$, the covariance of two J elements, the ensemble averaged Green's function, $\mathcal{G}(\eta,z) \equiv \langle \boldsymbol{G}(\eta,z;J)\rangle_J$, and the self-energy $\Sigma(\eta,z)$, Eq. (3.24) [the matrix indices for $\mathcal{G}(\eta,z)$ and $\Sigma(\eta,z)$ are arranged as for $\boldsymbol{G}_{ab}^{\alpha\beta}(\eta,z;0)$]. The third line is the diagrammatic representation of the expansion Eq. (3.22) of $\boldsymbol{G}(\eta,z;J)$ before averaging over J, where the J's are represented by dashed lines. Averaging over Eq. (3.2) is performed by pairing all J's and connecting them with the wavy lines representing $\langle JJ\rangle$. In the large N limit, the contribution of crossing pairings is suppressed by negative powers of N; the sum of all noncrossing diagrams, shown on the fourth line, yields the leading contribution to $\mathcal{G}(\eta,z)$ for large N. The last line shows the diagrammatic representation of Eq. (3.23), which if iterated generates all the noncrossing diagrams. Alternatively, $\mathcal{G}(\eta,z)$ can be found by solving this self-consistent equation directly.

where $\boldsymbol{G}(\eta,z;0)$ is given by Eq. (3.9) with the J's set to zero. This equation is represented diagrammatically in the third line of Fig. 11; the thin arrows defined in the first line of the figure represents $\boldsymbol{G}(\eta,z;0)$, and the dashed lines represent a power of J before ensemble averaging. To obtain the average resolvent, $\mathcal{G}(\eta,z)$, we then average Eq. (3.22), term by term, with respect to the ensemble Eq. (3.2). Since the measure is Gaussian with zero mean, according to Wick's formula, the average of *each* term of Eq. (3.22) involving n factors of J is given by a sum over the contributions of all possible complete pairings of the J's in that term [in particular, since $\langle J\rangle_J = 0$, terms in Eq. (3.22) with odd powers of J vanish after averaging]. Each pairing can be represented as a Feynman diagram, as shown in Fig. 11, the first two lines of which define the diagram elements. For example, the last diagram in the fourth line of Fig. 11 shows one possible pairing of the term in Eq. (3.22) corresponding to $n = 6$. The contribution of each pairing diagram is given by a product of factors, one per each pair, given by Eq. (3.21) (represented by wavy lines) with the right indices for that pair, as well as the factors of $\boldsymbol{G}(\eta,z;0)$ (represented by thin arrows), with all the intervening greek and latin matrix indices summed over their proper ranges. We show in Appendix A that for Im $\eta \neq 0$, and so long as $\|(RL)^{-1}\|$ remains bounded as $N \to \infty$, only noncrossing pairings need to be retained in the large N limit, as crossing pairings are suppressed by inverse powers of N and do not contribute in the limit (a pairing diagram is noncrossing if it can be drawn on a plane, with the wavy lines drawn only on the half-plane

above the straight arrow line, without any wavy lines crossing). As the last two lines of Fig. 11 demonstrate, all noncrossing diagrams can be generated by iterating the equation

$$\mathcal{G}(\eta,z) = \boldsymbol{G}(\eta,z;0) + \boldsymbol{G}(\eta,z;0)\Sigma(\eta,z)\mathcal{G}(\eta,z), \qquad (3.23)$$

starting from $\mathcal{G}_{(0)}(\eta,z) = \boldsymbol{G}(\eta,z;0)$. This equation is represented diagrammatically in the last line of Fig. 11, with the "self-energy" matrix, $\Sigma(\eta,z)$, defined by the diagram in the second line of that figure, i.e.,

$$\Sigma(\eta,z) \equiv \langle J\mathcal{G}(\eta,z)J \rangle_J . \qquad (3.24)$$

Using Eq. (3.21) we obtain

$$\Sigma_{ad}^{\alpha\delta}(\eta,z) = \delta_{ad} \sum_{r=1}^{2} \pi_{\alpha\delta}^{r} \frac{1}{N} \operatorname{Tr}\left[\pi^{3-r}\mathcal{G}(\eta,z)\right], \qquad (3.25)$$

which using Eq. (3.20) we can write as

$$\Sigma(\eta,z) = \begin{pmatrix} -ig_2(\eta,z)\mathbf{1} & 0 \\ 0 & -ig_1(\eta,z)\mathbf{1} \end{pmatrix}, \qquad (3.26)$$

where we defined the scalar functions

$$g_\alpha(\eta,z) \equiv i \operatorname{tr} \mathcal{G}^{\alpha\alpha}(\eta,z). \qquad (3.27)$$

Using Eq. (3.26), we can solve Eqs. (3.23)–(3.26) for $\mathcal{G}(\eta,z)$ at once, in terms of $g_\alpha(\eta,z)$, and then use Eq. (3.27) to obtain a self-consistency equation, which can be solved for $g_\alpha(\eta,z)$. To this end, we multiply Eq. (3.23) by $\boldsymbol{G}^{-1}(\eta,z;0)$ on the left and by $\mathcal{G}^{-1}(\eta,z)$ on the right to obtain

$$\mathcal{G}(\eta,z) = \left[\boldsymbol{G}^{-1}(\eta,z;0) - \Sigma(\eta,z)\right]^{-1}$$
$$= [\eta - H_0(z) - \Sigma(\eta,z)]^{-1}. \qquad (3.28)$$

Using this expression with Eqs. (3.14) and (3.26), it can be easily checked that

$$\mathcal{G}(\eta,z) = -\begin{pmatrix} (\eta+ig_1)K_1^{-1} & (z-M)K_2^{-1} \\ (\bar{z}-M^\dagger)K_1^{-1} & (\eta+ig_2)K_2^{-1} \end{pmatrix}, \qquad (3.29)$$

where $K_1 \equiv M_z M_z^\dagger + (g_1 - i\eta)(g_2 - i\eta)$ and $K_2 \equiv M_z^\dagger M_z + (g_1 - i\eta)(g_2 - i\eta)$, and we dropped the arguments of $g_\alpha(\eta,z)$ for succinctness. Imposing Eq. (3.27) we obtain the self-consistency equations

$$g_1 = (g_1 - i\eta) \operatorname{tr}\left(K_1^{-1}\right), \qquad (3.30)$$

$$g_2 = (g_2 - i\eta) \operatorname{tr}\left(K_2^{-1}\right). \qquad (3.31)$$

Before solving these equations for g_1 and g_2, we first show that $\operatorname{tr}(K_1^{-1}) = \operatorname{tr}(K_2^{-1})$. One way to see this is to use the singular value decomposition (SVD) of M_z in the form

$$M_z = U_z S_z V_z^\dagger, \qquad (3.32)$$

where S_z is a non-negative diagonal matrix with the singular values of M, $s_i(z)$ $(i = 1,\dots,N)$, on the diagonal, and U_z and V_z are unitary matrices [as in Sec. II A we include possibly vanishing singular values among $s_i(z)$, so that S_z, U_z, and V_z are always $N \times N$ matrices]. Using the invariance of trace under similarity transforms, we obtain $\operatorname{tr}(K_1^{-1}) = \operatorname{tr}(K_2^{-1}) = \operatorname{tr}[S_z^2 + (g_1 - i\eta)(g_2 - i\eta)]^{-1}$. Given this equality, it is not

hard to see that Eqs. (3.30) cannot be simultaneously satisfied unless $g_1(\eta,z) = g_2(\eta,z) \equiv g(\eta,z)$, with $g(\eta,z)$ satisfying

$$g = (g - i\eta) \operatorname{tr}\left[\frac{1}{S_z^2 + (g - i\eta)^2}\right], \qquad (3.33)$$

or as written in the original basis,

$$g = (g - i\eta)\operatorname{tr}\left[\frac{1}{M_z M_z^\dagger + (g - i\eta)^2}\right]. \qquad (3.34)$$

Noting from Eqs. (3.26), that the self-energy is thus proportional to the $2N \times 2N$ identity matrix, from Eqs. (3.28) and (3.9) (for $J = 0$) we obtain

$$\mathcal{G}(\eta,z) = \boldsymbol{G}(\eta + ig(\eta,z),z;0)$$
$$= -\begin{pmatrix} \frac{i\gamma}{M_z M_z^\dagger + \gamma^2} & \frac{M_z}{M_z^\dagger M_z + \gamma^2} \\ \frac{M_z^\dagger}{M_z M_z^\dagger + \gamma^2} & \frac{i\gamma}{M_z^\dagger M_z + \gamma^2} \end{pmatrix}, \qquad (3.35)$$

where $\gamma \equiv g(\eta,z) - i\eta$.

According to Eq. (3.11), for our case of interest we must solve Eq. (3.34) in the limit $\eta \to i0$. Note, however, that, as shown in Appendix A, the noncrossing approximation is, in general, guaranteed to work only for $\operatorname{Im} \eta \neq 0$; hence, the limit $\eta \to i0$ must be taken after the limit $N \to \infty$ (as already pointed out in Sec. II, taking the limits in this order is important in cases where some of the singular values in S_z vanish in the limit $N \to \infty$). For our purposes, it suffices to let $\eta = i\epsilon$ for some real positive ϵ and take the limit $\epsilon \to 0^+$ at the end. In this case one must seek a positive solution for $g(i\epsilon,z)$ in Eq. (3.34); this is because, by definition, $g(\eta,z) = i\operatorname{tr}\mathcal{G}_{11}(\eta,z) = \langle \operatorname{tr} i\boldsymbol{G}_{11}(\eta,z;J) \rangle_J$ and from Eq. (3.9) we obtain $g(i\epsilon,z) = \langle \operatorname{tr}\frac{\epsilon}{(M_z - J)(M_z - J)^\dagger + \epsilon^2} \rangle_J$, which for $\epsilon > 0$ is the ensemble average of the trace of a positive definite matrix and hence positive. Taking the limit $N \to \infty$ while keeping ϵ (and hence $\epsilon + g$) positive and nonzero, we define

$$\mathcal{K}(\gamma,z) \equiv \lim_{N \to \infty} \operatorname{tr}\left[\frac{1}{M_z M_z^\dagger + \gamma^2}\right] = \lim_{N \to \infty} \operatorname{tr}\left[\frac{1}{S_z^2 + \gamma^2}\right] \qquad (3.36)$$

for $\gamma = g + \epsilon > 0$. We can then rewrite Eq. (3.34) as

$$\gamma(1 - \mathcal{K}(\gamma,z)) = \epsilon, \qquad (3.37)$$

with $\gamma = g + \epsilon$. Since ϵ and $\gamma = g + \epsilon$ are positive, it follows that $1 - \mathcal{K}(\gamma,z)$ must also be positive. In the limit $\epsilon \to 0^+$ there are two possible situations: (1) $g,\gamma \to 0^+$, in which case we must have

$$\lim_{\gamma \to 0^+} \mathcal{K}(\gamma,z) < 1, \qquad (3.38)$$

or $\lim_{\gamma \to 0^+} \mathcal{K}(\gamma,z) = 1$, or (2) the solution for g stays finite and positive in the limit, while $\mathcal{K}(\gamma,z) \to 1^-$ as $\gamma \to g^+$. Thus, in the second case $g(z) \equiv \lim_{\epsilon \to 0^+} g(\epsilon,z)$ must satisfy $\mathcal{K}(g(z),z) = 1$, i.e.,

$$1 = \lim_{N \to \infty} \operatorname{tr}\left[\frac{1}{S_z^2 + g(z)^2}\right]. \qquad (3.39)$$

Note further that since $\mathcal{K}(\gamma,z)$ is a decreasing function of γ, in the second case we have $\mathcal{K}(0^+,z) \geqslant \mathcal{K}(g(z),z) = 1$, i.e.,

$$\lim_{\gamma \to 0^+} \mathcal{K}(\gamma,z) \geqslant 1. \tag{3.40}$$

Thus, the two possible solutions are realized in complimentary regions (with a shared boundary) of the complex plane for z, respectively, given by Eqs. (3.38) and (3.40).

Let us substitute the $g(z) = 0$ solution for the case (3.38) in Eq. (3.35) and naively set $\eta = i\epsilon$ (and thus γ) to zero, to obtain

$$\mathcal{G}(\eta = i0^+,z) = \boldsymbol{G}(\eta = i0^+,z;0). \tag{3.41}$$

From Eqs. (3.10) and (3.11), this solution yields $\langle G(z;J) \rangle_J = -\mathcal{G}^{21}(\eta = i0^+,z) = M_z^{-1} = R(z - M)^{-1}L$, which is analytic outside the spectrum of M. Hence, from Eq. (3.7), it yields $\rho(z) = 0$, at least outside the spectrum of M; a more careful analysis presented in Appendix B, in which we correctly take the limit $N \to \infty$ in Eq. (3.17) before taking $\epsilon \to 0^+$, confirms that in the region Eq. (3.38), $\lim_{N\to\infty} \rho(z)$ always vanishes.

We conclude that the support of $\lim_{N\to\infty} \rho(z)$ is where Eq. (3.40) holds [which is Eq. (2.20) of Sec. II]; here $g(z)$ is to be found by solving Eq. (3.39), or equivalently Eq. (2.9) or Eq. (2.10). In this region, we obtain $\rho(z)$ by substituting Eq. (3.35), with the solution of Eq. (3.39), into Eqs. (3.17). This yields Eq. (2.8), which we rewrite here as

$$\rho(z) = \frac{1}{\pi} \partial_{\bar{z}} E(z), \tag{3.42}$$

$$E(z) \equiv \mathrm{tr}\left[\frac{(RL)^{-1}M_z^\dagger}{M_z M_z^\dagger + g(z)^2}\right], \tag{3.43}$$

with $g(z)$ given by Eq. (3.39), or equivalently Eq. (2.9).

We now obtain an alternative expression for $\rho(z)$, equivalent to Eqs. (3.42) and (3.43), which explicitly shows that it depends only on the singular values of M_z. Noting that, from Eq. (2.6), $\partial_z(M_z M_z^\dagger) = (RL)^{-1}M_z^\dagger$, we can write Eq. (3.43) as

$$E(z) = \mathrm{tr}\left[\frac{\partial_z(M_z M_z^\dagger)}{M_z M_z^\dagger + g(z)^2}\right]. \tag{3.44}$$

On the other hand, we have

$$\partial_z \mathrm{tr}\, \ln[M_z M_z^\dagger + g(z)^2] = \mathrm{tr}\left[\frac{\partial_z(M_z M_z^\dagger + g^2(z))}{M_z M_z^\dagger + g(z)^2}\right],$$
$$= E(z) + \partial_z[g^2(z)], \tag{3.45}$$

where to write the last term we used Eq. (2.9). Thus, we obtain

$$E(z) = \partial_z \varphi(z), \tag{3.46}$$

$$\varphi(z) \equiv -g^2(z) + \mathrm{tr}\, \ln[M_z M_z^\dagger + g(z)^2], \tag{3.47}$$

or using the SVD, Eq. (3.32),

$$\varphi(z) = -g^2(z) + \mathrm{tr}\, \ln\left[S_z^2 + g(z)^2\right] \tag{3.48}$$

$$= -g(z)^2 + \frac{1}{N}\sum_{i=1}^{N} \ln\left[s_i(z)^2 + g(z)^2\right]. \tag{3.49}$$

Finally, substituting Eq. (3.49) in Eq. (3.46), and using Eq. (2.10), we obtain Eq. (2.12).

For the special case of $M = 0$, we have $M_z = z(RL)^{-1}$. If we let σ_i to be the singular values of RL, then the singular values of M_z will be given by $s_i(z) = |z|\sigma_i^{-1}$. Substituting this in Eq. (2.10) and multiplying both sides by $r^2 = |z|^2$, we obtain Eq. (2.15). We see immediately that $g(z)$, $\varphi(z)$, and $\rho(z)$ depend only on the radius $r = |z|$. Similarly, we can rewrite Eq. (3.49) as

$$\varphi(r) = -g(r)^2 + \frac{1}{N}\sum_{i=1}^{N} \ln\left[r^2\sigma_i^{-2} + g(r)^2\right]. \tag{3.50}$$

To find the spectral radius (boundary of the spectrum) r_0 we have to solve Eq. (2.15) for r, setting $g(r) = 0$. This yields $r_0^2 = \frac{1}{N}\sum_{i=1}^{N} \sigma_i^2 = \|RL\|_{\mathrm{F}}^2$, yielding Eq. (2.13). Let us define the proportion of eigenvalues lying outside a radius r from the origin by $n_>(r)$. To obtain Eqs. (2.14) and (2.16), first note that

$$\rho(r) = \frac{1}{\pi} \partial_{\bar{z}} \partial_z\, \varphi(z) = \frac{1}{4\pi} \nabla^2 \varphi(z) = \frac{1}{4\pi r} \frac{\partial}{\partial r}[r\partial_r\varphi(r)], \tag{3.51}$$

where we used the expression of Laplacian, $\nabla^2 = \partial_x^2 + \partial_y^2$, in 2D polar coordinates in the last equality. Using this with the definition $n_>(r) = 2\pi \int_r^\infty \rho(r)r\,dr$, we obtain $n_>(r) = \left[\frac{r}{2}\partial_r\varphi(r)\right]_r^\infty$. For the limit at $r \to \infty$, note that for $r > r_0$, $g(r) = 0$ and we have $\varphi(r) = \frac{1}{N}\sum_{i=1}^{N} \ln(r^2\sigma_i^{-2}) = 2\ln r - \frac{2}{N}\ln \det(RL)$, and hence $\frac{r}{2}\partial_r\varphi(r) \to 1$ as $r \to \infty$. Thus, we obtain

$$n_>(r) = 1 - \frac{r}{2}\partial_r\varphi(r). \tag{3.52}$$

Differentiating Eq. (3.50) and using Eq. (2.15) we obtain

$$\partial_r\varphi(r) = 2r\frac{1}{N}\sum_{i=1}^{N} \frac{1}{r^2 + \sigma_i^2 g(r)^2}, \tag{3.53}$$

and

$$n_>(r) = 1 - r^2\frac{1}{N}\sum_{i=1}^{N} \frac{1}{r^2 + \sigma_i^2 g(r)^2} \tag{3.54}$$

$$= g(r)^2\frac{1}{N}\sum_{i=1}^{N} \frac{\sigma_i^2}{r^2 + \sigma_i^2 g(r)^2}. \tag{3.55}$$

Using Eq. (2.15) once again we obtain Eq. (2.16). Finally, using the latter together with Eqs. (3.51) and (3.52) yields Eq. (2.14).

We will prove further general properties for the eigenvalue density for $M = 0$. Let us first define

$$I_{n,k}(g,r) \equiv \left\langle \frac{\sigma^{-k}}{(g^2 + \sigma^{-2}r^2)^n} \right\rangle_\sigma \tag{3.56}$$

and

$$\langle f(\sigma) \rangle_\sigma \equiv \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} f(\sigma_i). \tag{3.57}$$

[We assume σ_i have a limit density, $\rho_\sigma(\sigma)$, such that $\langle f(\sigma) \rangle_\sigma = \int_0^\infty f(\sigma)\rho_\sigma(\sigma)d\sigma$ is well defined for $f(\sigma)$ with sufficiently

fast decay at infinity. Note that since we assumed that $\|(RL)^{-1}\| = (\min_i \sigma_i)^{-1} = O(1)$, this density has no measure at $\sigma = 0$ and hence the averages in Eq. (3.56) are nonsingular for $n, k \geqslant 0$. Also, $\langle f(\sigma) \rangle_\sigma$ is finite as long as $f(\sigma) = O(\sigma^2)$ as $\sigma \to \infty$, as we are assuming that the $\|RL\|_F = O(1)$ and $\lim_{N \to \infty} \|RL\|_F^2 = \langle \sigma^2 \rangle_\sigma$.] First, we obtain general expressions for $\rho(r = 0)$ and $\rho(r = r_0)$, with r_0 given by Eq. (2.13). From Eq. (2.14), $\pi \rho(r) = -\frac{\partial n_>(r)}{\partial (r^2)}$, which using Eq. (2.15), reexpressed as $I_{1,0}(g, r) = 1$, we can write as

$$\pi \rho(r) = \frac{\frac{\partial I_{1,0}}{\partial (r^2)}}{\frac{\partial I_{1,0}}{\partial (g^2)}} = \frac{I_{2,2}(g, r)}{I_{2,0}(g, r)}. \qquad (3.58)$$

Using the facts that at $r = 0$, $g = 1$ and at $r = r_0$, $g = 0$ we obtain

$$\rho(r = 0) = \frac{1}{\pi} \frac{I_{2,2}(1, 0)}{I_{2,0}(1, 0)} = \frac{1}{\pi} \langle \sigma^{-2} \rangle_\sigma, \qquad (3.59)$$

$$\rho(r = r_0) = \frac{1}{\pi} \frac{I_{2,2}(0, r_0)}{I_{2,0}(0, r_0)} = \frac{1}{\pi} \frac{\langle \sigma^2 \rangle_\sigma}{\langle \sigma^4 \rangle_\sigma}. \qquad (3.60)$$

Using the fact that σ^4 and σ^{-2} are anticorrelated and that $\sigma^2 = \sigma^4 \sigma^{-2}$, we see that $\langle \sigma^2 \rangle_\sigma \leqslant \langle \sigma^4 \rangle_\sigma \langle \sigma^{-2} \rangle_\sigma$ or

$$\rho(r = r_0) \leqslant \rho(r = 0), \qquad (3.61)$$

with equality if and only if $\rho(\sigma)$ is deterministic, i.e., a δ function. This can happen if all but an $o(1)$ fraction of the σ_i's have the same limit as $N \to \infty$; in that case the eigenvalue distribution is given by the circular law. More generally, we can prove that $\rho(r)$ is a decreasing function of r for any choice of L and R (with $M = 0$). Using $\frac{d\rho(r)}{dr} = 2r \frac{d\rho(r)}{d(r^2)}$, and Eq. (3.58) we obtain

$$\frac{d\rho(r)}{dr} = 2r \frac{\frac{dI_{2,2}}{d(r^2)} I_{2,0} - I_{2,2} \frac{dI_{2,0}}{d(r^2)}}{I_{2,0}^2}, \qquad (3.62)$$

and using $\frac{d}{d(r^2)} = \frac{\partial}{\partial (r^2)} + \frac{\partial (g^2)}{\partial (r^2)} \frac{\partial}{\partial (g^2)} = \frac{\partial}{\partial (r^2)} - \pi \rho(r) \frac{\partial}{\partial (g^2)}$ and $\frac{\partial I_{n,k}}{\partial (g^2)} = -n I_{n+1,k}$ and $\frac{\partial I_{n,k}}{\partial (r^2)} = -n I_{n+1,k+2}$ [we drop the explicit (g, r) dependence of $I_{n,k}$'s when convenient], we find

$$\frac{d\rho(r)}{dr} = -4r \frac{I_{2,0}^2 I_{3,4} - 2 I_{2,2} I_{2,0} I_{3,2} + I_{2,2}^2 I_{3,0}}{I_{2,0}^3}. \qquad (3.63)$$

Defining

$$\langle f(\sigma) \rangle'_\sigma \equiv \frac{\left\langle \frac{f(\sigma)}{(g^2 + \sigma^{-2} r^2)^2} \right\rangle_\sigma}{\left\langle \frac{1}{(g^2 + \sigma^{-2} r^2)^2} \right\rangle_\sigma}, \qquad (3.64)$$

[$\langle f(\sigma) \rangle'_\sigma$ is a bona fide expectation operator] we can write

$$\frac{d\rho(r)}{dr} = -4r \left[\left\langle \frac{\sigma^{-4}}{g^2 + \sigma^{-2} r^2} \right\rangle'_\sigma - 2 \left\langle \frac{\sigma^{-2}}{g^2 + \sigma^{-2} r^2} \right\rangle'_\sigma \langle \sigma^{-2} \rangle'_\sigma \right.$$

$$\left. + \left\langle \frac{1}{g^2 + \sigma^{-2} r^2} \right\rangle'_\sigma \langle \sigma^{-2} \rangle'^2_\sigma \right] \qquad (3.65)$$

or

$$\frac{d\rho(r)}{dr} = -4r \left\{ \mathrm{Cov}' \left[\frac{\sigma^{-2}}{g^2 + \sigma^{-2} r^2}, \sigma^{-2} \right] \right.$$

$$\left. - \langle \sigma^{-2} \rangle'_\sigma \mathrm{Cov}' \left[\frac{1}{g^2 + \sigma^{-2} r^2}, \sigma^{-2} \right] \right\}, \qquad (3.66)$$

where $\mathrm{Cov}'[f, g] \equiv \langle fg \rangle'_\sigma - \langle f \rangle'_\sigma \langle g \rangle'_\sigma$ is the covariance under $\langle \cdot \rangle'_\sigma$. Now since $\frac{\sigma^{-2}}{g^2 + \sigma^{-2} r^2}$ and σ^{-2} are both strictly decreasing functions of σ (since $g > 0$ for $r < r_0$), while $\frac{1}{g^2 + \sigma^{-2} r^2}$ is a strictly increasing function of σ (for $r > 0$), the first covariance on the right-hand side of Eq. (3.66) is positive, while the second one is negative, and, therefore,

$$\frac{d\rho(r)}{dr} \leqslant 0. \qquad (3.67)$$

This slope is zero at $r = 0$ and strictly negative for $r > 0$ as long as $\mathrm{Var}[\sigma] > 0$ (again when $\mathrm{Var}[\sigma] = 0$ we obtain the circular law). At $r = r_0$ we obtain

$$\rho'(r_0) = -\frac{4}{r_0} \langle \sigma^{-2} \rangle'_\sigma (\langle \sigma^{-2} \rangle'_\sigma \langle \sigma^2 \rangle'_\sigma - 1)$$

$$= -\frac{4}{r_0} \frac{\langle \sigma^2 \rangle_\sigma}{\langle \sigma^4 \rangle_\sigma^3} \left(\langle \sigma^2 \rangle_\sigma \langle \sigma^6 \rangle_\sigma - \langle \sigma^4 \rangle_\sigma^2 \right). \qquad (3.68)$$

The curvature of $\rho(r)$ at zero can also be evaluated by taking the limit $r \to 0$ of the bracket in Eq. (3.66), noting that $g \to 1$ as $r \to 0$. We obtain

$$\rho''(r = 0) = -4 \mathrm{Var}'[\sigma^{-2}] = -4 \frac{\mathrm{Var}[\sigma^2]}{\langle \sigma^4 \rangle_\sigma^2} \leqslant 0. \qquad (3.69)$$

4. DERIVATION OF THE FORMULA FOR THE AVERAGE NORM SQUARED

In this section, we focus on the dynamics governed by the matrix $A = M + LJR$, according to Eq. (2.2), and derive the general formulas presented in Sec. II B. We first consider the system's response to an impulse input, $\mathbf{I}(t) = \mathbf{x}_0 \delta(t)$, at $t = 0$, before which we assume the system was at rest in its fixed point $\mathbf{x} = 0$. We assume $\mathbf{x} = 0$ is a stable fixed point, i.e., all eigenvalues of $-\gamma \mathbf{1} + A$ have negative real parts, or equivalently, all eigenvalues of A have real parts less than γ [more precisely, we assume that as $N \to \infty$, this will be the case almost surely, i.e., for any typical realization of J; in particular, the vertical line of z's with real part γ must be to the right of the support of $\rho(z)$, the average eigenvalue density for A, as found by solving Eq. (2.5)]. This means that $\mathbf{x}(t)$ decays exponentially as $t \to \infty$, and therefore its Fourier transform, $\tilde{\mathbf{x}}(\omega) \equiv \int_{-\infty}^{\infty} e^{-i\omega t} \mathbf{x}(t) dt = \int_0^{\infty} e^{-i\omega t} \mathbf{x}(t) dt$, is well defined. Fourier transformation of Eq. (2.2) with $\mathbf{I}(t) = \mathbf{x}_0 \delta(t)$ yields $i\omega \tilde{\mathbf{x}}(\omega) = (-\gamma + A) \tilde{\mathbf{x}}(\omega) + \mathbf{x}_0$. Solving algebraically for $\tilde{\mathbf{x}}(\omega)$, we obtain $\tilde{\mathbf{x}}(\omega) = (\gamma + i\omega - A)^{-1} \mathbf{x}_0$, or using Eqs. (3.3) and (3.4), $\tilde{\mathbf{x}}(\omega) = R^{-1} G(\gamma + i\omega; J) L^{-1} \mathbf{x}_0$. The inverse Fourier transform, $\mathbf{x}(t) = \int_{-\infty}^{\infty} e^{it\omega} \tilde{\mathbf{x}}(\omega) \frac{d\omega}{2\pi}$, then yields

$$\mathbf{x}(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{it\omega} R^{-1} G(\gamma + i\omega; J) L^{-1} \mathbf{x}_0. \qquad (4.1)$$

Our goal is to study the statistics of $\mathbf{x}(t)$ (e.g., its moments) under the distribution Eq. (3.2). Equation (4.1) allows us to reduce this task to the calculation of various moments

of $G(z;J)$ and its adjoint, and these can be found using the diagrammatic technique. Note that, in general, these moments involve not only the statistics of the eigenvalues, but also that of the eigenvectors of $A = M + LJR$; this can be seen from the spectral representation $R^{-1}G(z;J)L^{-1} = (z - A)^{-1} = V(z - \Lambda)^{-1}V^{-1}$, where Λ is a diagonal matrix of the eigenvalues of A, and V is the matrix whose columns are the eigenvectors of A. Here we look at the simplest interesting statistic involving the eigenvectors: the average square norm of the state vector, namely, $\langle \|\mathbf{x}(t)\|^2 \rangle_J$. As we discussed in Sec. II B, its study is also motivated by the fact that transient amplification due to non-normality of A manifests itself in the transient growth of $\|\mathbf{x}(t)\|^2 = \mathbf{x}(t)^\mathrm{T}\mathbf{x}(t)$. With a slight generalization, we derive a formula for the average of a general quadratic function $\mathbf{x}(t)^\mathrm{T}B\mathbf{x}(t)$, where B is any symmetric matrix; the norm squared corresponds to $B = \mathbf{1}$. Using $\mathbf{x}(t)^\mathrm{T} = \mathbf{x}(t)^\dagger$ [$\mathbf{x}(t)$ is real], the identity $\mathbf{x}^\dagger B\mathbf{x} = \mathrm{Tr}(B\mathbf{x}\mathbf{x}^\dagger)$, and Eq. (4.1), we obtain

$$\mathbf{x}(t)^\mathrm{T}B\mathbf{x}(t) = \iint \frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}e^{it(\omega_1-\omega_2)}$$
$$\times \mathrm{Tr}\left[B_R\,G(\gamma+i\omega_1;J)C_L\,G^\dagger(\gamma+i\omega_2;J)\right],$$
$$(4.2)$$

where we defined $C_L \equiv L^{-1}\mathbf{x}_0\mathbf{x}_0^\mathrm{T}L^{-\dagger}$ and $B_R \equiv R^{-\dagger}BR^{-1}$. Using Eq. (3.11) and $G^\dagger(z;J) = -\lim_{\eta\to i0^+}\mathbf{G}^{12}(\eta,z;J)$, and the 2×2 matrices π^r defined in Eq. (3.20), we can rewrite the trace in Eq. (4.2) as $\mathrm{Tr}[\pi^2 \otimes B_R\,\mathbf{G}(0,z_1;J)\pi^1 \otimes C_L\,\mathbf{G}(0,z_2;J)]$, with $z_i = \gamma + i\omega_i$, where now the trace is performed over $2N \times 2N$ matrices. Averaging over J, we then obtain

$$\langle\mathbf{x}(t)^\mathrm{T}B\mathbf{x}(t)\rangle_J = \iint \frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}e^{it(\omega_1-\omega_2)}$$
$$\times \mathcal{F}\left(\gamma+i\omega_1,\gamma+i\omega_2;B,\mathbf{x}_0\mathbf{x}_0^\mathrm{T}\right),\quad(4.3)$$

where, for general matrix arguments B and C, we define

$$\mathcal{F}(z_1,z_2;B,C) \equiv \langle\mathrm{Tr}[\mathbf{B}\,\mathbf{G}(0,z_1;J)\mathbf{C}\,\mathbf{G}(0,z_2;J)]\rangle_J.\quad(4.4)$$

with

$$\mathbf{B} \equiv \pi^2 \otimes B_R, \quad B_R \equiv R^{-\dagger}BR^{-1},\quad(4.5)$$

$$\mathbf{C} \equiv \pi^1 \otimes C_L, \quad C_L \equiv L^{-1}CL^{-\dagger}.\quad(4.6)$$

Before proceeding to the calculation of $\mathcal{F}(z_1,z_2;B,C)$ using the diagrammatic technique, we also express the other quantities presented in Sec. II B in terms of $\mathcal{F}(\gamma+i\omega,\gamma+i\omega;B,C)$, with appropriate B's and C's. First we obtain the desired expression for the matrix power spectrum, Eq. (2.27), of the steady-state response to a temporally white noisy input, $\mathbf{I}(t)$, with covariance Eq. (2.26). Using the Fourier transform of Eq. (2.2), and following similar steps to those leading to Eq. (4.1), we can write the steady-state solution for $\mathbf{x}(t)$ as in Eq. (4.1) with $\mathbf{x}_0$ replaced with the Fourier transform of the input, $\tilde{\mathbf{I}}(\omega)$. Using this and exploiting $\mathbf{x}_j(t_2) = \mathbf{x}_j^*(t_2)$ we can write (after averaging over the input noise)

$$\overline{\mathbf{x}_i(t_1)\mathbf{x}_j(t_2)} = \iint \frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}e^{it_1\omega_1-it_2\omega_2}K_{ij}(\omega_1,\omega_2),\quad(4.7)$$

where the Fourier-domain covariance matrix, $K(\omega_1,\omega_2) \equiv \overline{\tilde{\mathbf{x}}(\omega_1)\tilde{\mathbf{x}}(\omega_2)^\dagger}$, is given by

$$K(\omega_1,\omega_2) \equiv R^{-1}G(\gamma+i\omega_1;J)$$
$$\times L^{-1}C^\mathbf{I}(\omega_1,\omega_2)L^{-\dagger}G^\dagger(\gamma+i\omega_2;J)R^{-\dagger}.\quad(4.8)$$

Here the bars indicate averaging over the input noise distribution, and we defined $C^\mathbf{I}(\omega_1,\omega_2) \equiv \overline{\tilde{\mathbf{I}}(\omega_1)\tilde{\mathbf{I}}(\omega_2)^\dagger}$. On the other hand, the Fourier transform of Eq. (2.26) yields

$$C^\mathbf{I}(\omega_1,\omega_2) \equiv \overline{\tilde{\mathbf{I}}(\omega_1)\tilde{\mathbf{I}}(\omega_2)^\dagger} = 2\pi\delta(\omega_1-\omega_2)C^\mathbf{I},\quad(4.9)$$

where we also exploited $\tilde{\mathbf{I}}_j^*(\omega) = \tilde{\mathbf{I}}_j(-\omega)$ for a real $\mathbf{I}(t)$. Substituting Eq. (4.9) into Eqs. (4.7) and (4.8) we obtain

$$\overline{\mathbf{x}_i(t_1)\mathbf{x}_j(t_2)} = \int \frac{d\omega}{2\pi}e^{i\omega(t_1-t_2)}C_{ij}^\mathbf{x}(\omega),\quad(4.10)$$

where

$$C^\mathbf{x}(\omega) = R^{-1}G(\gamma+i\omega;J)L^{-1}C^\mathbf{I}L^{-\dagger}G^\dagger(\gamma+i\omega;J)R^{-\dagger}.\quad(4.11)$$

Noting that Eq. (4.10) expresses the covariance of the response as an inverse Fourier transform, we see that $C^\mathbf{x}(\omega)$ is indeed the power spectrum of the response, as defined in Eq. (2.27). Finally note that the element, C_{ij}, of any matrix can be expressed as $\mathrm{Tr}(\mathbf{e}_j\mathbf{e}_i^\mathrm{T}C)$, where $\mathbf{e}_i$ are the unit basis vectors (i.e., vectors whose ath component is δ_{ia}). Using this trick with Eq. (4.11), and following the steps leading from Eq. (4.2) to Eq. (4.3), we see that after ensemble averaging, $\langle C_{ij}^\mathbf{x}(\omega)\rangle_J$ can be written in the form

$$\langle C_{ij}^\mathbf{x}(\omega)\rangle_J = \mathcal{F}(\gamma+i\omega,\gamma+i\omega;\mathbf{e}_j\mathbf{e}_i^\mathrm{T},C^\mathbf{I}),\quad(4.12)$$

where $\mathcal{F}$ was defined by Eqs. (4.4)–(4.6).

Next, consider the system Eq. (2.2) being driven by a sinusoidal input $\mathbf{I}(t) = \mathbf{I}_0\sqrt{2}\cos\omega t$ [the factor of $\sqrt{2}$ serves to normalize the time average of $(\sqrt{2}\cos\omega t)^2$ to one], and consider the steady-state response, which will also oscillate at frequency ω. Decomposing the input, $\mathbf{I}(t)$, and the steady-state response, $\mathbf{x}_\omega(t)$, into their positive and negative frequency components (proportional to $e^{i\omega t}$ and $e^{-i\omega t}$, respectively), from Eq. (2.2) we obtain

$$\mathbf{x}_\omega(t) = \sqrt{2}R^{-1}\mathrm{Re}[e^{i\omega t}G(\gamma+i\omega;J)]L^{-1}\mathbf{I}_0.\quad(4.13)$$

Thus, the norm squared of the steady-state response, $\|\mathbf{x}(t)\|^2 = \mathbf{x}(t)^\dagger\mathbf{x}(t)$, will have a zero frequency component, plus components oscillating at $\pm 2\omega$. Averaging over time kills the latter, leaving the zero frequency component intact, yielding

$$\overline{\mathbf{x}_\omega(t)^\mathrm{T}\mathbf{x}_\omega(t)} = \mathbf{I}_0^\mathrm{T}L^{-1}G^\dagger(z;J)R^{-\dagger}R^{-1}G(z;J)L^{-1}\mathbf{I}_0$$
$$= \mathrm{Tr}[R^{-\dagger}R^{-1}G(z;J)\rho_\mathbf{I}G^\dagger(z;J)],\quad(4.14)$$

where $z = \gamma + i\omega$, the bar indicates temporal averaging, and we defined $\rho_\mathbf{I} \equiv L^{-1}\mathbf{I}_0\mathbf{I}_0^\mathrm{T}L^{-\dagger}$. Generalizing to $\overline{\mathbf{x}_\omega(t)^\mathrm{T}B\mathbf{x}_\omega(t)}$, averaging over the ensemble, and following the steps leading from Eq. (4.2) to Eq. (4.3), we obtain

$$\langle\overline{\mathbf{x}_\omega(t)^\mathrm{T}B\mathbf{x}_\omega(t)}\rangle_J = \mathcal{F}(\gamma+i\omega,\gamma+i\omega;B,\mathbf{I}_0\mathbf{I}_0^\mathrm{T}),\quad(4.15)$$

where $\mathcal{F}$ is given by Eqs. (4.4)–(4.6). Comparing Eq. (4.15) with Eq. (4.12), we also obtain

$$\langle\overline{\mathbf{x}_\omega(t)^\mathrm{T}B\mathbf{x}_\omega(t)}\rangle_J = \mathrm{Tr}[B\langle C^\mathbf{x}(\omega)\rangle_J],\quad(4.16)$$
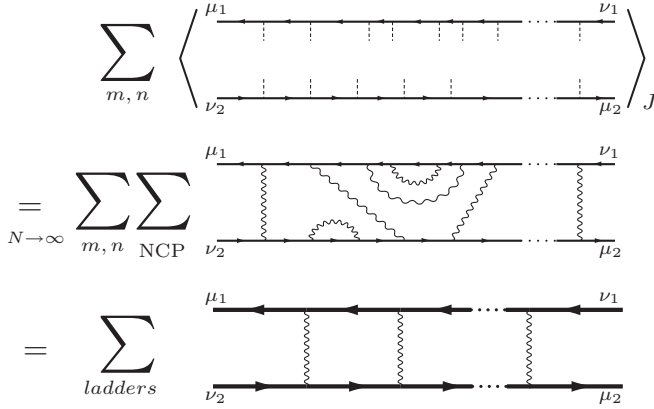
FIG. 12. Contributions to Eq. (4.17) in the noncrossing approximation. The first line shows Eq. (4.17) written using the expansion Eq. (3.22). The diagram shows the contribution of the mth and nth terms in the expansion for two Green's functions, respectively. Thus, the top (bottom) solid line contains m (n) factors of J, shown by dashed lines. In the large N limit, averaging each summand over J boils down to summing all noncrossing pairings (NCPs) of the dashed lines. The second row shows a specific noncrossing pairing for the diagram shown in the first line. Finally, summing over all m and n and all NCPs is equivalent to replacing all solid lines [representing $\boldsymbol{G}(\eta_i, z_i; J = 0)$] with thick solid lines representing the noncrossing average Green's function, $\mathcal{G}(\eta_i, z_i)$ [calculated according to Eqs. (3.26)–(3.28)], and summing over all NCPs with *every* pairing connecting the straight lines on top and bottom (and not each to itself). This procedure yields the ladder diagrams, the sum over which is shown in the third line.

which is Eq. (2.31) of Sec. II, it being understood that $C^{\mathbf{I}}$ in Eq. (4.12) is replaced with $\mathbf{I}_0 \mathbf{I}_0^{\mathsf{T}}$, as in Eq. (4.15).

Now that we have expressed all our quantities of interest in terms of the kernel $\mathcal{F}$ as defined in Eq. (4.4), our task boils down to performing the average over J in Eq. (4.4) to obtain a closed formula for $\mathcal{F}$ with general arguments B and C. To this end, we now proceed to calculate the more general object,

$$F_{\mu_1\nu_2;\mu_2\nu_1}(1;2) \equiv \langle \boldsymbol{G}_{\mu_1\nu_1}(1;J)\boldsymbol{G}_{\mu_2\nu_2}(2;J)\rangle_J, \quad (4.17)$$

using the diagrammatic technique. Here we adopted the abbreviated notation $(1) \equiv (\eta_1, z_1)$ and $(2) \equiv (\eta_2, z_2)$ for the function arguments and $\mu_i = (\alpha_i, a_i)$ (similarly for ν_i) for indices in the $2N$-dimensional space (as in Sec. III, $\alpha, \beta, \ldots$, and $a, b, \ldots$, denote indices in the 2 and N-dimensional spaces, respectively). Once we have calculated $F_{\mu_1\nu_2;\mu_2\nu_1}(1;2)$, we can obtain $\mathcal{F}(z_1, z_2; B, C)$, with the appropriate B and C, via

$$\mathcal{F}(z_1, z_2; B, C) = \boldsymbol{B}_{\nu_2\mu_1} F_{\mu_1\nu_2;\mu_2\nu_1}(0, z_1; 0, z_2) \boldsymbol{C}_{\nu_1\mu_2}, \quad (4.18)$$

where all indices are summed over and B and C were defined in Eqs. (4.5) and (4.6).

As before, we start by using the expansion Eq. (3.22) for the two Green's functions in Eq. (4.17). This is shown diagrammatically in the first line of Fig. 12, for the contribution of mth and nth terms in the expansion of the first and second Green's function, respectively. As before, for large N, averaging over J entails summing the contribution of all noncrossing pairings. This is indicated in the second line of Fig. 12. Finally, the third line of Fig. 12 shows that summing
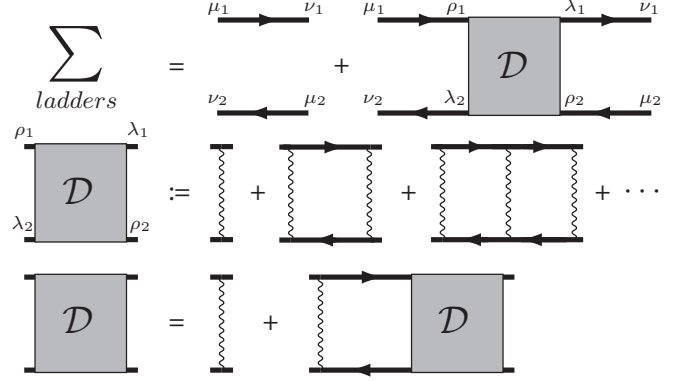

FIG. 13. The first row is the diagrammatic representation of Eqs. (4.19)–(4.21). In the last term, ρ's and λ's are summed over. It shows the sum of all ladder diagrams contributing to Eq. (4.17) (i.e., the last line of Fig. 12) in terms of $\mathcal{D}$, which is defined in the second row. The first term on the right side of the first row equation (the ladder with zero rungs) is the disconnected average Eq. (4.20); it corresponds to taking the average of each Green's function in Eq. (4.17) separately and then multiplying. The last row shows an iterative form of the equation in the second row, which can be solved to give the expression Eqs. (4.23) and (4.26) for $\mathcal{D}$.

over all m's and n's and all noncrossing pairings is equivalent to replacing all solid lines with thick solid lines representing the average Green's function in the noncrossing approximation, $\mathcal{G}(\eta_i, z_i)$ [defined diagrammatically in the third line of Fig. 11, and given by Eq. (3.35) as we found in the previous section], and summing over all noncrossing pairings with *every* pairing connecting the thick arrow lines on top and bottom (and not each to itself). This procedure yields a sum over all ladder diagrams with different numbers of rungs, as shown in the third line of Fig. 12.

As shown in the first row of Fig. 13, the sum of all ladder diagrams can be written as a sum

$$F = F^0 + F^D, \quad (4.19)$$

where

$$F^0_{\mu_1\nu_2;\mu_2\nu_1}(1;2) \equiv \mathcal{G}_{\mu_1\nu_1}(1)\mathcal{G}_{\mu_2\nu_2}(2) \quad (4.20)$$

is the disconnected average of the two Green's functions, and $F^D_{\mu_1\nu_2;\mu_2\nu_1}(1;2)$ is the sum of ladder diagrams in which the two Green's functions are connected by at least one wavy line. The latter can be written in the form

$$F^D_{\mu_1\nu_2;\mu_2\nu_1}(1;2)$$
$$\equiv \mathcal{G}_{\lambda_2\nu_2}(2)\mathcal{G}_{\mu_1\rho_1}(1)\mathcal{D}_{\rho_1\lambda_2;\rho_2\lambda_1}(1;2)\mathcal{G}_{\lambda_1\nu_1}(1)\mathcal{G}_{\mu_2\rho_2}(2), \quad (4.21)$$

where all repeated indices are summed over and $\mathcal{D}$ is given by the sum of all diagrams in the second row of Fig. 13.

To calculate $\mathcal{D}$, it helps to first rewrite Eq. (3.21) as

$$\langle J_{ab}^{\alpha\beta} J_{cd}^{\gamma\delta}\rangle_J = \frac{1}{N}\sum_{r,s=1}^{2}\left(\pi_{\alpha\delta}^r \delta_{ad}\right)\sigma_{rs}^1\left(\pi_{\gamma\beta}^s \delta_{cb}\right), \quad (4.22)$$

where

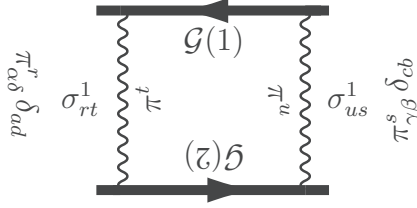$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

FIG. 14. The contribution to $\mathcal{D}_{ad;cb}^{\alpha\delta;\gamma\beta}(1;2)$ from the second term in the series shown in the second row of Fig. 13, in more detail. The covariance of J in the form Eq. (4.22) is used to write this expression in a more manageable form. The repeated indices, r,t,u,s, are summed over 1 and 2. The matrices inside the loop multiply each other in cyclic order, giving rise to the trace $\text{Tr}\,(\mathcal{G}(2)\pi^t\mathcal{G}(1)\pi^u)$. The whole diagram gives $\frac{1}{N}\sum_{rs}(\pi^r\otimes\mathbf{1})_{ad}^{\alpha\delta}\left[\sigma^1\Pi^D\sigma^1\right]_{rs}(\pi^s\otimes\mathbf{1})_{cb}^{\gamma\beta}$ where the "polarization matrix" Π_{tu}^D was defined in Eq. (4.25).

is the first Pauli matrix. This helps us because in the expansion of Fig. 13, the two factors in Eq. (3.21) involving π^r and π^s decouple and get absorbed in adjacent loops, or contribute to form factors in the left or right ends of the ladder diagrams. This is demonstrated in Fig. 14 for the second term in the series expansion of $\mathcal{D}$ shown in the second line of Fig. 13. Extending this similarly to all the terms in that expansion, we obtain

$$
\begin{aligned}
\mathcal{D}_{\mu\rho;\lambda\nu}(1;2) &= \mathcal{D}_{ad;cb}^{\alpha\delta;\gamma\beta}(1;2) \\
&= \frac{1}{N}\sum_{r,s=1}^{2}\left(\pi_{\alpha\delta}^r\delta_{ad}\right)D_{rs}(1;2)\left(\pi_{\gamma\beta}^s\delta_{cb}\right), \quad (4.23)
\end{aligned}
$$

where $\mu=(\alpha,a)$, $\nu=(\beta,b)$, $\lambda=(\gamma,c)$, $\rho=(\delta,d)$, and we defined the 2×2 matrices,

$$
D(1;2)\equiv\sigma^1+\sigma^1\Pi^D\sigma^1+\cdots=\sigma^1\sum_{n=0}^{\infty}(\Pi^D\sigma^1)^n, \quad (4.24)
$$

and the "polarization matrix",

$$
\Pi_{rs}^D(1;2)\equiv\text{tr}\,[\pi^r\mathcal{G}(1)\pi^s\mathcal{G}(2)]=\text{tr}\,[\mathcal{G}^{rs}(1)\mathcal{G}^{sr}(2)]. \quad (4.25)
$$

Here, as before, the trace was performed over the $2N$-dimensional space, and we used Eq. (3.20) to write the last form of Π^D. Summing the geometric series in Eq. (4.24) we obtain

$$
D(1;2)=\sigma^1[\mathbf{1}_{2\times 2}-\Pi^D(1;2)\sigma^1]^{-1}. \quad (4.26)
$$

The 2×2 matrix inversion yields

$$
\begin{aligned}
D(1;2)=&\frac{1}{(1-\Pi_{12}^D)(1-\Pi_{21}^D)-\Pi_{11}^D\Pi_{22}^D} \\
&\times\begin{pmatrix}\Pi_{22}^D & 1-\Pi_{12}^D \\ 1-\Pi_{21}^D & \Pi_{11}^D\end{pmatrix}, \quad (4.27)
\end{aligned}
$$

where all Π^D's have arguments $(1;2)=(\eta_1,z_1;\eta_2,z_2)$ which were suppressed for clarity.

Going back to Eq. (4.18), we can also break up $\mathcal{F}(z_1,z_2;B,C)$ into a disconnected part and a connected part mirroring the decomposition Eqs. (4.19)–(4.21),

$$
\mathcal{F}(z_1,z_2;B,C)=\mathcal{F}^0(z_1,z_2;B,C)+\Delta\mathcal{F}(z_1,z_2;B,C), \quad (4.28)
$$

where $\mathcal{F}^0(z_1,z_2;B,C)$ and $\Delta\mathcal{F}(z_1,z_2;B,C)$ are defined as in Eq. (4.18), but with $F_{\mu_1\nu_2;\mu_2\nu_1}$ on the right side replaced with $F_{\mu_1\nu_2;\mu_2\nu_1}^0$ and $F_{\mu_1\nu_2;\mu_2\nu_1}^D$, respectively. Using Eqs. (4.20) and (4.21) and (4.23), we then obtain

$$
\begin{aligned}
\mathcal{F}^0(z_1,z_2;B,C) &= \text{Tr}\,[\boldsymbol{B}\mathcal{G}(0,z_1)\boldsymbol{C}\mathcal{G}(0,z_2)] \\
&= \text{Tr}\,[B_R\mathcal{G}^{21}(0,z_1)C_L\mathcal{G}^{12}(0,z_2)], \quad (4.29)
\end{aligned}
$$

and

$$
\begin{aligned}
&\Delta\mathcal{F}(z_1,z_2;B,C) \\
&= \frac{1}{N}\sum_{r,s}\text{Tr}\,[B_R\mathcal{G}^{2r}(0,z_1)\mathcal{G}^{r2}(0,z_2)] \\
&\quad\times D_{rs}(0,z_1;0,z_2)\,\text{Tr}\,[\mathcal{G}^{s1}(0,z_1)C_L\mathcal{G}^{1s}(0,z_2)], \quad (4.30)
\end{aligned}
$$

where r and s are summed over $\{1,2\}$.

According to Eq. (4.3), we are interested in $z_i=\gamma+i\omega_i$ ($i=1,2$) for arbitrary real ω_i. As we mentioned before Eq. (4.1), these trace a vertical line in the complex plane that is entirely to the right of the support of the average eigenvalue density, $\rho(z)$, of A, i.e., they are in the region where the the valid solution of Eq. (3.34) is the trivial $g(0,z)=0$. In this case, we have Eq. (3.41), and for $\eta\to i0^+$, from Eq. (3.10) (replacing A with M, corresponding to $J=0$) we have

$$
\mathcal{G}(0,z_i)=-\begin{pmatrix}0 & M_{z_i}^{-\dagger}\\ M_{z_i}^{-1} & 0\end{pmatrix}. \quad (4.31)
$$

Using this in Eqs. (4.29) and (4.30), we obtain

$$
\mathcal{F}^0(z_1,z_2;B,C)=\text{Tr}\,\left(B_R M_{z_1}^{-1}C_L M_{z_2}^{-\dagger}\right), \quad (4.32)
$$

and

$$
\begin{aligned}
&\Delta\mathcal{F}(z_1,z_2;B,C) \\
&= \text{tr}\,[B_R\mathcal{G}^{21}(0,z_1)\mathcal{G}^{12}(0,z_2)]D_{12}(0,z_1;0,z_2) \\
&\quad\times\text{Tr}\,[\mathcal{G}^{21}(0,z_1)C_L\mathcal{G}^{12}(0,z_2)]. \quad (4.33)
\end{aligned}
$$

Using the definitions Eq. (2.6) and Eqs. (4.5) and (4.6) we can simplify Eq. (4.32) to

$$
\mathcal{F}^0(z_1,z_2;B,C)=\text{Tr}\,\left(B\frac{1}{z_1-M}C\frac{1}{\bar{z}_2-M^\dagger}\right). \quad (4.34)
$$

From Eqs. (4.25) and (4.31) we see that (for z_i of interest and for η_i going to zero) $\Pi^{rr}=0$ and $\Pi^{12}=\Pi^{21}$, and from Eq. (4.27) we obtain

$$
\begin{aligned}
D_{12}(0,z_1;0,z_2) &= \frac{1}{1-\Pi_{21}^D(0,z_1;0,z_2)} \\
&= \frac{1}{1-\text{tr}\,[\mathcal{G}^{21}(0,z_1)\mathcal{G}^{12}(0,z_2)]}. \quad (4.35)
\end{aligned}
$$

Substituting this in Eq. (4.33) and using Eq. (4.31) once again, we finally obtain

$$
\Delta\mathcal{F}(z_1,z_2;B,C)=\frac{\text{tr}\,\left(B_R M_{z_1}^{-1}M_{z_2}^{-\dagger}\right)\text{Tr}\,\left(M_{z_1}^{-1}C_L M_{z_2}^{-\dagger}\right)}{1-\text{tr}\,\left(M_{z_1}^{-1}M_{z_2}^{-\dagger}\right)},
$$

and after simplification using Eqs. (2.6) and (4.5) and (4.6),

$$
\Delta\mathcal{F}(z_1,z_2;B,C)
$$
$$
= \frac{\operatorname{tr}\left(B\,\frac{1}{z_1-M}LL^\dagger\,\frac{1}{\bar{z}_2-M^\dagger}\right)\operatorname{Tr}\left(R^\dagger R\,\frac{1}{z_1-M}C\,\frac{1}{\bar{z}_2-M^\dagger}\right)}{1-\operatorname{tr}\left(R^\dagger R\,\frac{1}{z_1-M}LL^\dagger\,\frac{1}{\bar{z}_2-M^\dagger}\right)}. \quad (4.36)
$$

The general formulas of Sec. II B readily follow. Equations (2.21)–(2.24), with $C^{\mathbf{I}}$ replaced with $\mathbf{x}_0\mathbf{x}_0^{\mathrm{T}}$, for the case of response to an impulse input follow from Eqs. (4.3), (4.28), (4.34), and (4.36), respectively. Equations (2.28)–(2.30) [with $\mathcal{C}_0^{\mathbf{x}}$ and $\Delta\mathcal{C}^{\mathbf{x}}$ defined in Eqs. (2.23) and (2.24)] for the power spectrum of the response to a temporally white noisy input, are similarly obtained from Eq. (4.12) by using Eqs. (4.28), (4.34), and (4.36), after setting $B=\mathbf{e}_j\mathbf{e}_i^{\mathrm{T}}$, $C=C^{\mathbf{I}}$, and $z_1=z_2=\gamma+i\omega$ [with the traces involving $B=\mathbf{e}_j\mathbf{e}_i^{\mathrm{T}}$ turned into matrices in Eqs. (2.28)–(2.30), using $\operatorname{Tr}(\mathbf{e}_j\mathbf{e}_i^{\mathrm{T}}X)=X_{ij}$]. The result Eq. (2.31) for the steady-state response to a sinusoidal input was already derived in Eq. (4.16).

We see that according to Eqs. (4.3) and (4.28),

$$
\left\langle\mathbf{x}(t)^{\mathrm{T}}B\mathbf{x}(t)\right\rangle_J = [\mathbf{x}(t)^{\mathrm{T}}B\mathbf{x}(t)]_{J=0}+\Delta f_B(t), \quad (4.37)
$$

where the two terms on the right-hand side are obtained by replacing $\mathcal{F}(\cdot,\cdot;B)$ in Eq. (4.3) with Eqs. (4.34) and (4.36), respectively. The integrals over ω_1 and ω_2 decouple for the first term yielding the expected result for $J=0$,

$$
[\mathbf{x}(t)^{\mathrm{T}}B\mathbf{x}(t)]_{J=0} = e^{-2\gamma t}\operatorname{Tr}\left(Be^{tM}\mathbf{x}_0\mathbf{x}_0^{\mathrm{T}}e^{tM^\dagger}\right)
$$
$$
= \mathbf{x}_0^{\mathrm{T}}e^{t(-\gamma+M)^\dagger}Be^{t(-\gamma+M)}\mathbf{x}_0. \quad (4.38)
$$

Unlike the $J=0$ contribution, it is not possible to perform the double Fourier transform, Eq. (4.3), needed for obtaining $\Delta f_B(t)$ for arbitrary M, L, and R. In the next section, we will analytically calculate this for some special examples of M, with L and R proportional to the identity matrix (i.e., for iid quenched randomness).

5. CALCULATIONS FOR SPECIFIC EXAMPLES OF M

In this section we give the detailed calculations of the explicit expressions for the spectral density Eq. (2.8), the power spectrum Eq. (2.31), and the average squared norm Eqs. (2.21) and (2.25), for the specific examples of M, L, and R presented in Sec. II C.

In the examples worked out in Secs. V A and V B, both R and L are proportional to the identity matrix; we take $L=\mathbf{1}$ and $R=\sigma\mathbf{1}$. Furthermore, for such examples we do the calculations by choosing the unit of time such that $\sigma=1$ [notice that given Eq. (2.2), the elements of A and M have dimensions of frequency]; then at the end of our calculations using the replacements $t\to t\sigma$, $z\to z/\sigma$, $\gamma\to\gamma/\sigma$, $M\to M/\sigma$, and $\rho\to\sigma^2\rho$ (with the latter applying to both the eigenvalue density and the power spectral density), we obtain the result for general σ. The eigenvalue density and the norm squared $\|\mathbf{x}\|^2$ are invariant with respect to unitary transforms, and, for L and R proportional to the identity, so is the distribution of the random part of A, Eq. (3.2). Thus, by effecting a unitary transform $M\to U^\dagger MU$, we can assume M is already in its Schur form, Eq. (2.40), without loss of generality.

A. Single feedforward chain of length N: $M_{ij}=w\,\delta_{i+1,j}-\gamma\,\delta_{ij}$

We start with the example in Sec. II C 1, where M is

$$
M=T=\begin{pmatrix}0 & w & 0 & \cdots\\ 0 & 0 & w & \cdots\\ \vdots & \vdots & \vdots & \ddots\end{pmatrix} \quad (5.1)
$$

or $M_{ij}=w\,\delta_{i+1,j}$. First we calculate the eigenvalue density. According to Eqs. (2.8) and (2.9), in order to calculate the spectral density, we need to calculate first the inverse of $M_zM_z^\dagger+g^2=(z-M)(z-M)^\dagger+g^2$ (remember that we have set $\sigma=1$, as we explained in the beginning of the section). To this end, notice that $K_{ij}\equiv[(z-M)(z-M)^\dagger]_{ij}=Q_{ij}-|w|^2\delta_{iN}\delta_{jN}$, where

$$
Q_{ij}\equiv(|z|^2+|w|^2)\delta_{ij}-w\bar{z}\delta_{i+1,j}-\bar{w}z\delta_{i,j+1}. \quad (5.2)
$$

As the difference $(z-M)(z-M)^\dagger-Q=-|w|^2\mathbf{e}_N\mathbf{e}_N^{\mathrm{T}}$ is single rank, we can use the Woodbury formula for matrix inversion to write

$$
\frac{1}{K+g^2}=\frac{1}{Q+g^2}+\frac{1}{Q+g^2}\mathbf{e}_N\mathbf{e}_N^{\mathrm{T}}\frac{1}{Q+g^2}
$$
$$
\times\left[\frac{1}{|w|^{-2}-\mathbf{e}_N^{\mathrm{T}}(Q+g^2)^{-1}\mathbf{e}_N}\right], \quad (5.3)
$$

where $\mathbf{e}_N^{\mathrm{T}}=(0,\ldots,0,1)$. [The only conditions for the validity of Eq. (5.3) is that the factor in parentheses is not singular, i.e., $\mathbf{e}_N^{\mathrm{T}}(Q+g^2)^{-1}\mathbf{e}_N\neq|w|^{-2}$; we consider the validity of this condition below.] Since Q is Toeplitz and Hermitian, it can be diagonalized easily. Using standard methods, we find that the eigenvalues and eigenvectors of Q, satisfying $Q\boldsymbol{v}_n=\lambda_n\boldsymbol{v}_n$, are given by

$$
\lambda_n=||z|-|w|e^{i\phi_n}|^2, \quad \phi_n\equiv\frac{\pi n}{N+1}, \quad (5.4)
$$

$$
v_n^j=\sqrt{\frac{2}{N+1}}\left(\frac{\bar{w}z}{w\bar{z}}\right)^{j/2}\sin\phi_n j, \quad (5.5)
$$

for $n=1,\ldots,N$. The eigenvectors are orthonormal $\boldsymbol{v}_n^\dagger\boldsymbol{v}_m=\delta_{nm}$ and we have the spectral representation

$$
\frac{1}{Q+g^2}=\sum_{n=1}^{N}\boldsymbol{v}_n\frac{1}{\lambda_n+g^2}\boldsymbol{v}_n^\dagger. \quad (5.6)
$$

Using Eqs. (5.4)–(5.6) in Eq. (5.3) we obtain

$$
\operatorname{tr}\frac{1}{K+g^2}=I_1(g,z)+\frac{1}{N}\frac{-\frac{\partial I_2(g,z)}{\partial g^2}}{|w|^{-2}-I_2(g,z)}, \quad (5.7)
$$

where we defined

$$
I_1(g,z)\equiv\operatorname{tr}\frac{1}{Q+g^2}=\frac{1}{N}\sum_{n=1}^{N}\frac{1}{\lambda_n+g^2}, \quad (5.8)
$$

$$
I_2(g,z)\equiv\mathbf{e}_N^{\mathrm{T}}\frac{1}{Q+g^2}\mathbf{e}_N=\frac{1}{N+1}\sum_{n=1}^{N}\frac{2\sin^2\phi_n}{\lambda_n+g^2}. \quad (5.9)
$$

In writing the numerator of the last term in Eq. (5.7), we used Eqs. (5.5) and (5.6) to write $\operatorname{Tr}(\frac{1}{Q+g^2}\mathbf{e}_N\mathbf{e}_N^{\mathrm{T}}\frac{1}{Q+g^2})=\|\frac{1}{Q+g^2}\mathbf{e}_N\|^2=\frac{1}{N+1}\sum_{n=1}^{N}\frac{2\sin^2\phi_n}{(\lambda_n+g^2)^2}=-\frac{\partial I_2(g,z)}{\partial g^2}$. In the $N\to\infty$

limit, the sums in Eqs. (5.8) and (5.9) can be approximated by the integrals

$$I_1(g,z) = \int_0^{2\pi} \frac{1}{||z|-|w|e^{i\phi}|^2 + g^2} \frac{d\phi}{2\pi}, \quad (5.10)$$

$$I_2(g,z) = \int_0^{2\pi} \frac{2\sin^2\phi}{||z|-|w|e^{i\phi}|^2 + g^2} \frac{d\phi}{2\pi}. \quad (5.11)$$

Some elementary contour integration then yields

$$I_1(g,z) = [(|z|^2 + |w|^2 + g^2)^2 - 4|w|^2|z|^2]^{-1/2}, \quad (5.12)$$

$$I_2(g,z) = \frac{|z|^2 + |w|^2 + g^2 - I_1(g,z)^{-1}}{2|w|^2|z|^2}. \quad (5.13)$$

In particular, we see that $I_2(0,z) = \min(|w|^{-2},|z|^{-2})$, so that the condition for the validity of Eq. (5.3) would be violated for $|z| < |w|$, if g turns out to be zero. However, note that $I_2(g,z)$ is a decreasing function of g^2, so for finite $g^2 > 0$, the denominator in Eq. (5.7) is always positive (as is its numerator, for the same reason). Thus, if we follow the correct procedure of Eqs. (2.19) and (2.20), taking the $N \to \infty$ limit before sending g^2 to zero, we are justified in using Eqs. (5.3) and (5.7). Furthermore, for $g^2 > 0$ the second term in Eq. (5.7) is $O(N^{-1})$ and should be neglected. Solving Eq. (2.9) [with left-hand side correctly interpreted as Eq. (2.19)], which now takes the form $I_1(g,z) = 1$, yields

$$g(z)^2 = -|z|^2 - |w|^2 + \sqrt{4|w|^2|z|^2 - 1}. \quad (5.14)$$

This is positive if and only if

$$\sqrt{|w|^2 - 1} \leqslant |z| \leqslant \sqrt{|w|^2 + 1}, \quad (5.15)$$

which after the proper rescaling yields Eq. (2.36) for general σ. Note that Eq. (5.15) is precisely the region given by Eq. (2.20), which in the present case reads $I_1(0,z) \geqslant 1$. It is instructive to compare this result with what we would obtain by naively using Eq. (2.5), i.e., $\mathrm{tr}(K^{-1}) \geqslant 1$, wherein g is set to zero before taking the $N \to \infty$ limit; as we now show, that only yields the right inequality in Eq. (5.15). To see this, first note that for $|w| > |z|$, we can use Eq. (5.7) even for $g^2 = 0$ (since the denominator of the last term does not vanish), which yields $\mathrm{tr}\,[(M_z M_z^\dagger)^{-1}] = \mathrm{tr}\,(K^{-1}) = I_1(0,z) + o(1)$, and by Eq. (2.5), the right inequality in Eq. (5.15). For $|z| < |w|$, however, we cannot set $g = 0$ in Eq. (5.7). In fact, when $|z| < |w|$, the matrix $z - M$ has an exponentially small singular value; to see this, note that the vector $\mathbf{u}$ with components $u_i = (\frac{z}{w})^{i-1}$ satisfies $(z - M)\mathbf{u} = w(\frac{z}{w})^N \mathbf{e}_N$, so that $\|(z - M)\mathbf{u}\| = |w||\frac{z}{w}|^N$, and since $s_{\min}(z - M) \leqslant \frac{\|(z-M)\mathbf{u}\|}{\|\mathbf{u}\|}$ and $\|\mathbf{u}\| \geqslant 1$, it follows that $s_{\min}(z - M) \leqslant |w||\frac{z}{w}|^N$, which is $O(e^{-cN})$ for $|z| < |w|$. For large enough N, this singular value *alone* suffices to make Eq. (2.11) [equivalent to Eq. (2.5)] hold for *any* $|z| < |w|$, as $\frac{1}{N} s_{\min}(z)^{-2}$ diverges despite its $\frac{1}{N}$ prefactor.

Let us now calculate the eigenvalue density in the annulus Eq. (5.15). In order to use Eq. (2.8), we first calculate

$$\mathrm{tr}\,\frac{\bar{z}-M^\dagger}{K+g(z)^2} = \bar{z} - \mathrm{tr}\,\frac{M^\dagger}{K+g(z)^2}, \quad (5.16)$$

where we used Eq. (2.9) to write the last expression. To obtain $\mathrm{tr}\,\frac{M^\dagger}{K+g(z)^2}$, we again use Eq. (5.3). In the region Eq. (5.15), the contribution of the second term in Eq. (5.3) is again suppressed by $1/N$, and from Eqs. (5.1) and (5.6) we have $\mathrm{tr}\,M^\dagger[Q + g^2]^{-1} = \bar{w}\frac{1}{N}\sum_{n=1}^N (\sum_{j=1}^{N-1} v_n^j \bar{v}_n^{j+1})\frac{1}{\lambda_n + g^2}$. A straightforward calculation using Eq. (5.5) (and the orthonormality of $\boldsymbol{v}_n$) yields $\sum_{j=1}^{N-1} v_n^j \bar{v}_n^{j+1} = \left(\frac{w\bar{z}}{\bar{w}z}\right)^{1/2}\cos\phi_n$. Using this and approximating the sum over n with an integral, we obtain

$$\mathrm{tr}\left[\frac{M^\dagger}{Q+g(z)^2}\right]$$

$$\approx \frac{|w||z|}{z}\int_0^{2\pi} \frac{\cos\phi}{||z|-|w|e^{i\phi}|^2 + g(z)^2}\frac{d\phi}{2\pi},$$

$$= \frac{1}{2z}\{[|z|^2 + |w|^2 + g(z)^2]I_1(g(z),z) - 1\}. \quad (5.17)$$

Using Eq. (5.17) with $I_1(g(z),z) = 1$ [true in the region Eq. (5.15)], differentiating Eq. (5.16) with respect to $\bar{z}$, and substituting in Eq. (2.8), we finally obtain

$$\rho(z) = \frac{1}{\pi}\left[1 - \frac{|w|^2}{\sqrt{4|w|^2|z|^2 + 1}}\right] \quad (5.18)$$

for z in the region Eq. (5.15). After the proper rescaling this yields Eq. (2.37).

We now turn to the calculation of $\langle \|\mathbf{x}(t)\|^2 \rangle_J$, using (2.25). To calculate the trace in the denominator of Eq. (2.25), first note that for Eq. (5.1) the expansion $(z - M)^{-1} = \sum_{n=0}^{N-1} \frac{M^n}{z^{n+1}}$ terminates and is exact, yielding

$$\left[\frac{1}{z-M}\right]_{i,j} = \frac{1}{z}\left(\frac{w}{z}\right)^{j-i} \quad (5.19)$$

for $j \geqslant i$, and zero otherwise. Turning the sums in the trace into a sum over the nonzero diagonals of Eq. (5.19) we obtain

$$\mathrm{tr}\left(\frac{1}{\bar{z}_2-M^\dagger}\frac{1}{z_1-M}\right) = \frac{1}{\bar{z}_2 z_1}\sum_{n=0}^{N-1}\left(1 - \frac{n}{N}\right)q^n, \quad (5.20)$$

where $q \equiv |w|^2/(\bar{z}_2 z_1)$ and $z_i = \gamma + i\omega_i$. The condition of stability of Eq. (2.2) requires the entire spectrum of $-\gamma\mathbf{1} + A = -\gamma\mathbf{1} + M + J$ to be to the left of the imaginary axis. By Eq. (5.15), this requires $\gamma > \sqrt{|w|^2 + 1} > |w|$. It follows that $|q| < 1$, and therefore the geometric series Eq. (5.20) converges as $N \to \infty$. Summing the series and retaining terms of leading order as $N \to \infty$, we obtain

$$\mathrm{tr}\left(\frac{1}{\bar{z}_2-M^\dagger}\frac{1}{z_1-M}\right) = \frac{1}{\bar{z}_2 z_1 - |w|^2}. \quad (5.21)$$

If we set the initial condition $\mathbf{x}_0$ in Eqs. (2.25) [or the input amplitude $\mathbf{I}_0$ in Eq. (2.31)] to $\mathbf{e}_N = (0,\ldots,0,1)^\top$ and use Eq. (5.19), we find that the numerator in Eq. (2.25) is also given by the right hand side of Eq. (5.21). Using this and Eqs. (5.21), we obtain

$$F(z_1,z_2) = \frac{1}{\bar{z}_2 z_1 - |w|^2 - 1} \quad (5.22)$$

for the integrand of Eq. (2.25), which we denoted by $F(z_1,z_2)$, with $z_i = \gamma + i\omega_i$ ($i = 1,2$). By comparing the integrand of

Eq. (2.25) with the right hand side of Eq. (2.33), we see that to obtain the total power spectrum for the input amplitude $\mathbf{I}_0 = I_0(0,\ldots,0,1)^{\mathsf{T}}$, we need to multiply Eq. (5.22) by $I_0^2 = \|\mathbf{I}_0\|^2$ and substitute $z_1 = z_2 = \gamma + i\omega$. With the proper rescaling, this yields Eq. (2.39) for general σ. To obtain the formula for $\langle\|\mathbf{x}(t)\|^2\rangle_J$, we substitute Eq. (5.22) with $z_i = \gamma + i\omega_i$ for the integrand of Eq. (2.25). Changing the integration variables by $\omega_1 = \Omega + \omega/2$ and $\omega_2 = \Omega - \omega/2$, we obtain

$$
\begin{aligned}
\langle\|\mathbf{x}(t)\|^2\rangle_J &= \int \frac{d\omega}{2\pi} e^{it\omega} \int \frac{d\Omega}{2\pi} \frac{1}{\Omega^2 + (\gamma + i\omega/2)^2 - |w|^2 - 1} \\
&= \frac{1}{2} \int \frac{d\omega}{2\pi} \frac{e^{it\omega}}{\sqrt{(\gamma + i\omega/2)^2 - |w|^2 - 1}}.
\end{aligned} \tag{5.23}
$$

Finally consulting a table of Laplace transforms [56], we obtain

$$
\langle\|\mathbf{x}(t)\|^2\rangle_J = e^{-2\gamma t} I_0(2t\sqrt{|w|^2 + 1}), \quad (t \geqslant 0), \tag{5.24}
$$

where $I_0(x)$ is the 0th modified Bessel function. Implementing the rescalings $t \to t\sigma$, $\gamma \to \gamma/\sigma$, and $w \to w/\sigma$, we obtain Eq. (2.38).

B. $N/2$ feedforward chains of length 2

Here we carry out the explicit calculations for the example of Sec. II C 2, where M is given by Eq. (2.40) (without loss of generality, we assume M has its Schur form), using formulas (2.8) and (2.9) for the spectral density and Eq. (2.25) for $\langle\|\mathbf{x}(t)\|^2\rangle_J$. First we calculate the eigenvalue density. From Eq. (2.40), $K \equiv M_z M_z^\dagger = (z - M)(z - M)^\dagger$ [we are setting $L = R = \mathbf{1}$ in Eq. (2.6); see the comments at the beginning of this section] is a block-diagonal matrix with 2×2 diagonal blocks, with the bth block ($b = 1,\ldots,N/2$) given by

$$
\begin{pmatrix} z & -w_b \\ 0 & z \end{pmatrix} \begin{pmatrix} \bar{z} & 0 \\ -\bar{w}_b & \bar{z} \end{pmatrix} = \begin{pmatrix} |z|^2 + |w_b|^2 & -w_b\bar{z} \\ -\bar{w}_b z & |z|^2 \end{pmatrix}, \tag{5.25}
$$

where w_b is the corresponding Schur weight in Eq. (2.40). Likewise, $(K + g^2)^{-1}$ whose trace appears in Eqs. (2.9) is given by a block-diagonal matrix with diagonal blocks

$$
\frac{1}{(|z|^2 + g^2)^2 + |w_b|^2 g^2} \begin{pmatrix} |z|^2 + g^2 & w_b\bar{z} \\ \bar{w}_b z & |z|^2 + g^2 + |w_b|^2 \end{pmatrix}.
$$

Taking the normalized trace, we thus obtain

$$
\operatorname{tr}(K + g^2)^{-1} = \left\langle \frac{|z|^2 + g^2 + \frac{1}{2}|w_b|^2}{(|z|^2 + g^2)^2 + |w_b|^2 g^2} \right\rangle_b, \tag{5.26}
$$

where $\langle\cdot\rangle_b$ means averaging over the $N/2$ blocks, i.e., $\langle f(w_b)\rangle_b \equiv \frac{1}{N/2} \sum_{b=1}^{N/2} f(w_b)$.

Let us first calculate the support boundary of $\rho(z)$. As discussed in Sec. II A, when (for $|z| \neq 0$) all singular values of $M_z = z - M$ are bounded from below as $N \to \infty$, the support is correctly given by Eq. (2.5) [we discuss cases in which some $s_i(z)$ are $o(1)$ further below]. Setting $g = 0$ in Eq. (5.26), and substituting in Eq. (2.5), this yields

$$
1 \leqslant \operatorname{tr} K^{-1} = |z|^{-2} + \mu^2 |z|^{-4}, \tag{5.27}
$$

where we defined $\mu^2 = \frac{1}{2}\langle|w_b|^2\rangle_b = \operatorname{tr}(M^\dagger M)$. It follows that the support is the disk $|z| \leqslant r_0$, where

$$
r_0^2 = \frac{1}{2} + \sqrt{\frac{1}{4} + \mu^2}. \tag{5.28}
$$

The replacements $\mu \to \mu/\sigma$ and $r_0 \to r_0/\sigma$ then yield Eq. (2.44).

From Eqs. (2.9) and (5.26), within the support, $g^2(z)$ is found by solving the equation

$$
\operatorname{tr} \frac{1}{K + g^2} = \left\langle \frac{|z|^2 + g^2 + \frac{1}{2}|w_b|^2}{(|z|^2 + g^2)^2 + |w_b|^2 g^2} \right\rangle_b = 1, \tag{5.29}
$$

while for $|z| > r_0$ we have $g(z) = 0$. It is clear from Eq. (5.29) that $g^2(z)$ depends on z and $\bar{z}$ only through $|z| \equiv r$.

From Eq. (2.8), within its support the eigenvalue density is given by

$$
\begin{aligned}
\pi\rho(z) &= \frac{\partial}{\partial\bar{z}} \operatorname{tr}[M_z^\dagger(K + g^2)^{-1}] \\
&= 1 - \frac{\partial}{\partial\bar{z}} \operatorname{tr}[M^\dagger(K + g^2)^{-1}], \tag{5.30}
\end{aligned}
$$

where we are now using the short-hand $g^2 = g^2(|z|)$ [the solution of Eq. (5.29)], and in writing the second line we used $M_z^\dagger = \bar{z} - M^\dagger$ and Eq. (5.29). From Eqs. (2.40) and (5.26) we see that $M^\dagger(K + g^2)^{-1}$ has the same block-diagonal structure as Eq. (2.40), and a short calculation shows that $\operatorname{tr}[M^\dagger(K + g^2)^{-1}] = \bar{z} I_3(|z|)$, where we defined

$$
I_3(r) \equiv \left\langle \frac{\frac{1}{2}|w_b|^2}{[r^2 + g(r)^2]^2 + |w_b|^2 g(r)^2} \right\rangle_b. \tag{5.31}
$$

$I_3(r)$ is manifestly positive (assuming some w_b are nonzero), while when $g^2 > 0$, from Eq. (5.29) we have $I_3(r) \leqslant 1$, and thus,

$$
0 < I_3(r) \leqslant 1. \tag{5.32}
$$

Replacing this in Eq. (5.30), and using $\bar{z}\frac{\partial f(|z|)}{\partial\bar{z}} = 2r\frac{\partial f(r)}{\partial r}\big|_{r=|z|}$, we obtain

$$
\pi\rho(z) = \frac{1}{2r} \frac{\partial}{\partial r}[r^2 - r^2 I_3(r)] \tag{5.33}
$$

for $r = |z| \leqslant r_0$, and zero otherwise; the spectral density is rotationally symmetric and depends only on $r = |z|$. The advantage of writing the density as a complete derivative is that it can be immediately integrated to yield $n_<(r)$, the proportion of eigenvalues with modulus smaller than some radius r. We have $n_<(r) = 2\pi \int_0^r \rho(r')r'dr'$, which, upon substitution of Eq. (5.33), yields

$$
n_<(r) = r^2[1 - I_3(r)] \quad (r \leqslant r_0). \tag{5.34}
$$

Likewise, we define $n_>(r) \equiv 1 - n_<(r)$ to be the proportion of eigenvalues with modulus larger than r. From these definitions we have

$$
\rho(r) = \frac{1}{2\pi r} \frac{\partial n_<(r)}{\partial r} = -\frac{1}{2\pi r} \frac{\partial n_>(r)}{\partial r}, \tag{5.35}
$$

and from Eqs. (5.34) and $n_>(r) = 1 - n_<(r)$, after some manipulation exploiting Eq. (5.29), we obtain

$$
n_>(r) = g(r)^2[1 + I_3(r)]. \tag{5.36}
$$

We see that beyond the radius r at which g^2 vanishes (which when all w_b's are bounded is $r = r_0$), $n_>(r)$ and $\rho(r) = -\frac{1}{2\pi r}\frac{\partial n_>}{\partial r}$ vanish identically, while for smaller r they are positive.

In cases in which some w_b grow without bound as $N \to \infty$, some singular values, $s_i(z)$, of $M_z = z - M$ are $o(1)$, and more care is needed. First, to see this, note that by definition $s_i(z)^2$ are the eigenvalues of the block-diagonal $K = M_z M_z^\dagger$; thus, they come in pairs composed of the eigenvalues of K's 2×2 blocks, given by Eq. (5.25). We denote the pair of eigenvalues corresponding to block b by $s_{b,\pm}(z)^2$, with the plus and minus subscripts denoting the larger and smaller singular values, respectively. The sum $s_{b+}(z)^2 + s_{b-}(z)^2$ and the product $s_{b+}(z)^2 s_{b-}(z)^2$ are given by the trace and determinant of Eq. (5.25), i.e., by $|w_b|^2 + 2|z|^2$ and $|z|^4$, respectively. It follows that for blocks where the feedforward weight w_b is $O(1)$, both $s_{b,\pm}(z)$ will be $\Theta(1)$ for $|z| \neq 0$, while for blocks in which $w_b \to \infty$ as $N \to \infty$, we have

$$s_{b+}^2(z) = |w_b|^2 + O(1) \to \infty, \qquad (5.37)$$

$$s_{b-}^2(z) \approx \frac{|z|^4}{|w_b|^2} = o(1). \qquad (5.38)$$

[Note that as stated after Eq. (2.3) we assume $\|M\|_F^2 = \mu^2 = \langle |w_b|^2 \rangle_b / 2$ is $O(1)$, so that at most $o(N)$ number of weights can be unbounded, and each such w_b can at most be $O(\sqrt{N})$.] If all the w_b are $O(1)$, and hence all singular values are $\Theta(1)$ (for $|z| \neq 0$), Eq. (5.28) yields the correct support radius as noted above, and for $r \leqslant r_0$, Eq. (5.29) yields a $\Theta(1)$ solution for $g(r)^2$, which leads to a $\Theta(1)$ solution for $n_>(r)$ and $\rho(r)$ via Eqs. (5.36) and (5.35). In cases in which some w_b are unbounded, however, Eq. (5.28) [derived from Eq. (2.5)] may not yield the correct support boundary. Such cases are examples of the highly non-normal cases mentioned in the general discussion after Eq. (2.12), for which the support of $\lim_{N \to \infty} \rho(z)$ must be found by using Eqs. (2.19) and (2.20). This is equivalent to solving Eq. (5.29) after the limit $N \to \infty$ is taken (assuming $g^2 > 0$) and then finding where the solution for $g^2(|z|)$ vanishes, which yields the correct support radius. From Eq. (5.36) this is indeed the radius at which $\lim_{N \to \infty} n_>(r)$ and hence $\lim_{N \to \infty} \rho(r)$ vanish as well. This radius is, in general, smaller than r_0 as given by Eq. (5.28).

We now calculate $\rho(z)$ for two specific examples of M from each group. The first example is that of equal and $O(1)$ feedforward weights in all blocks, which we denote by w [in terms of Eq. (2.41), this case corresponds to $K = w\mathbf{1}$]. Here we can drop the block averages in Eqs. (5.29) and (5.31), replacing w_b with w. Solving Eqs. (5.29) for $g^2(|z| = r)$ we find

$$g^2(r) = \frac{1}{2} - \frac{w^2}{2} - r^2 + \frac{1}{2}\sqrt{1 + w^4 + 4w^2 r^2}. \quad (5.39)$$

Substituting this into Eq. (5.31) and Eq. (5.34) yields

$$n_<(r) = r^2 - \frac{w^2 r^2}{1 + \sqrt{1 + w^4 + 4w^2 r^2}}. \qquad (5.40)$$

The replacements $w \to w/\sigma$ and $r \to r/\sigma$ then yield Eq. (2.45) for general σ, and $\rho(r)$ can be calculated using Eq. (5.35).

The second case is that of Eq. (2.42). In this case only one of the blocks has a nonzero Schur weight given by $|w_1|^2 = \mathrm{Tr}(M^\dagger M) = N\mu^2 = O(N)$, where $\mu = O(1)$ is given by Eq. (2.43). Equation (5.29) now yields

$$1 = \frac{1}{r^2 + g^2} + \frac{\mu^2}{(r^2 + g^2)^2 + N\mu^2 g^2}\frac{r^2 - g^2}{r^2 + g^2}, \quad (5.41)$$

or

$$\frac{r^2 + g^2 - 1}{r^2 - g^2} = \frac{\mu^2}{(r^2 + g^2)^2 + N\mu^2 g^2}. \qquad (5.42)$$

The right hand side of this last equation is $I_3(r)$, as follows from Eq. (5.31); thus, using Eq. (5.42) we can rewrite Eq. (5.36) as

$$n_>(r) = g^2(r)\frac{2r^2 - 1}{r^2 - g^2(r)}. \qquad (5.43)$$

Let us now solve Eq. (5.42) to find $g(r)^2$. As noted above, and in accordance with the general prescription given after Eq. (2.12), for the purpose of obtaining $\lim_{N \to \infty} \rho(z)$ we have to first take the $N \to \infty$ limit in Eq. (5.41), keeping $g^2 > 0$ fixed, and only then solve for g^2. Doing so makes the last term in Eq. (5.41) vanish, and we obtain $g^2(r) = 1 - r^2$. This is positive for $r \leqslant 1$ and vanishes at $r = 1$, the correct support radius of $\lim_{N \to \infty} \rho(z)$, which is strictly smaller than r_0 given by Eq. (5.28). From Eq. (5.43) we obtain $n_>(r) = g^2(r) = 1 - r^2$. It then follows from Eq. (5.35) that the $N \to \infty$ limit of the eigenvalue density is identical with the circular law (the result for the $M = 0$), i.e., $\lim_{N \to \infty} \rho(r) = \frac{1}{\pi}$ for $r \leqslant 1$ and zero otherwise. With the correct scaling, this yields Eq. (2.46).

Contrary to the general prescription given after Eq. (2.12), we now solve equations Eqs. (5.29), (5.31), and (5.36) for $r > 1$, without taking the limit $N \to \infty$ first. As we will see, the obtained solution for $g(r)^2$, and by Eqs. (5.32) and (5.36) therefore the solutions for $n_>(r)$ and $\rho(r)$, will be nonzero but $o(1)$ for $1 < r \leqslant r_0$. As discussed in Sec. II C 2, these finite-size corrections, which, in general, are not trustworthy, in the present case are in surprisingly good agreement with simulations for some range of r's beyond $r_{\Theta(1)}$, but deviate from the true $n_>(r)$ for larger r (see Fig. 7). At finite N, it can indeed be checked that Eq. (5.29) has a positive solution for g^2 if and only if $r < r_0$, with r_0 given by Eq. (5.28). Simplifying Eq. (5.42) yields a cubic equation in g^2. However, it turns out that ignoring the cubic term in g^2 is harmless for large N; the quadratic approximation has the positive solution

$$g^2(r) = \sqrt{\left[\frac{1 - r^2}{2}\right]^2 + \frac{r^2(r^2 + \mu^2) - r^6}{\mu^2 N}} + \frac{1 - r^2}{2}, \quad (5.44)$$

and for all $r < r_0$, corrections to Eq. (5.44) when the cubic term is reinstated decay faster than the leading contribution from Eq. (5.44) [nevertheless, we numerically solved the full cubic equation (5.29) to obtain the black curve in Fig. 7, and the blue trace in Fig. 6]. First, analyzing Eq. (5.44) we see that $g^2(r)$ is indeed $\Theta(1)$ only for $r < 1$, where, as we already

found, $g^2(r) = 1 - r^2 + o(1)$. Furthermore, for a fixed $r > 1$ (such that $r - 1$ does not vanish as $N \to \infty$), the solution for $g(r)$ is $O(N^{-1})$. Thus, from Eq. (5.43) we see that $Nn_>(r)$, i.e., the total number of eigenvalues with modulus larger than r, for $1 < r < r_0$ [and $r - 1 = \Theta(1)$] is only $O(1)$; the solution for $Nn_>(r)$ is shown in Fig. 7. Correspondingly, from Eqs. (5.43) and (5.35) we see that $\rho(r)$ is $o(1)$ in this region and vanishes in the limit $N \to \infty$, as already found. Now let us calculate the total number of eigenvalues lying outside the circle $|z| = 1$. This is given by $Nn_>(1)$. From Eq. (5.44) we find $g^2(1) = \frac{1}{\sqrt{N}}$, and substituting in Eq. (5.43) we obtain

$$N_>(1) \equiv Nn_>(1) = \sqrt{N} + O(1). \tag{5.45}$$

With the proper rescaling this yields Eq. (2.47) for general σ. Note that, according to Eq. (5.44), $g(r)$ [and hence $n_>(r)$] remains $\Theta(N^{-1/2})$ [as opposed to $O(N^{-1})$] in a thin boundary layer outside of width $\Theta(N^{-1/2})$ just outside of the circle $|z| = 1$.

We now work out the formula for $\langle\|\mathbf{x}(t)\|^2\rangle_J$, Eqs. (2.21)–(2.25), when the initial condition $\mathbf{x}_0$ is the second Schur vector in block $b = a$, which we denote by $\mathbf{e}_{a2}$; in the Schur representation, Eq. (2.40), we have $\mathbf{e}_{a2} = (0,1)^T$ (we only write the components of $\mathbf{e}_{a2}$ in block a). To calculate the numerator in Eq. (2.25), we first calculate $(z - T_a)^{-1}\mathbf{e}_{a2}$, where

$$T_a = \begin{pmatrix} 0 & w_a \\ 0 & 0 \end{pmatrix}$$

denotes the ath diagonal 2×2 block of Eq. (2.40). Since $T_a^2 = 0$, we have $(z - T_a)^{-1} = z^{-1} + z^{-2}T_a$, which yields $\mathbf{v}_a(z) \equiv (z - T_a)^{-1}\mathbf{e}_{a2} = (w_a z^{-2}, z^{-1})^T$. We thus obtain

$$\mathbf{x}_0^T \frac{1}{\bar{z}_2 - M^\dagger} \frac{1}{z_1 - M}\mathbf{x}_0 = \mathbf{v}_a(z_2)^\dagger \mathbf{v}_a(z_1) = \frac{1}{z_1 \bar{z}_2} + \frac{|w_a|^2}{z_1^2 \bar{z}_2^2}. \tag{5.46}$$

On the other hand, we have

$$\mathrm{tr}\, \frac{1}{\bar{z}_2 - M^\dagger} \frac{1}{z_1 - M}$$
$$= \left\langle \frac{1}{2} \mathrm{Tr}_{2\times2} (\bar{z}_2 - T_b^\dagger)^{-1}(z_1 - T_b)^{-1} \right\rangle_b$$
$$= \frac{1}{z_1 \bar{z}_2} + \frac{\langle|w_b|^2\rangle_b/2}{z_1^2 \bar{z}_2^2}. \tag{5.47}$$

Substituting Eqs. (5.46) and (5.47) in Eq. (2.25), we obtain

$$F(z_1, z_2) = \frac{z_1 \bar{z}_2 + |w_a|^2}{(z_1 \bar{z}_2)^2 - (z_1 \bar{z}_2 + \mu^2)}, \tag{5.48}$$

where we used $\mu^2 = \langle|w_b|^2\rangle_b/2$, and we denoted the integrand of Eq. (2.25) by $F(z_1, z_2)$ with $z_i = \gamma + i\omega_i$ ($i = 1,2$). By comparing the integrand of Eq. (2.25) with Eq. (2.33), we see that substituting $z_1 = z_2 = \gamma + i\omega$ into Eq. (5.48) yields the total power spectrum, $\langle\|\mathbf{x}_\omega\|^2\rangle_J$. After the proper rescalings, this yields Eq. (2.50) for general σ. To obtain $\langle\|\mathbf{x}(t)\|^2\rangle_J$, on the other hand, we should substitute Eq. (5.48) into Eq. (2.21) with $z_i = \gamma + i\omega_i$. Let us use the change of variables $\omega_1 = \Omega + \omega$ and $\omega_2 = \Omega - \omega$. Then we have $z_1 \bar{z}_2 = \Omega^2 + (\gamma + i\omega)^2$, and

from Eq. (2.21) we obtain

$$\langle\|\mathbf{x}(t)\|^2\rangle_J = \int \frac{d\omega}{2\pi} e^{2it\omega} f_a(\gamma + i\omega), \tag{5.49}$$

where we defined

$$f_a(u) \equiv 2 \int \frac{d\Omega}{2\pi} \frac{\Omega^2 + u^2 + |w_a|^2}{(\Omega^2 + u^2)^2 - (\Omega^2 + u^2 + \mu^2)}. \tag{5.50}$$

Let us rewrite the integrand in Eq. (5.50) as

$$\frac{\Omega^2 + u^2 + |w_a|^2}{(\Omega^2 + u^2 - r_0^2)(\Omega^2 + u^2 + r_1^2)}$$
$$= \frac{\Omega^2 + u^2 + |w_a|^2}{r_0^2 + r_1^2} \left[\frac{1}{\Omega^2 + u^2 - r_0^2} - \frac{1}{\Omega^2 + u^2 + r_1^2} \right], \tag{5.51}$$

where r_0^2 was defined in Eq. (5.28) and

$$r_1^2 \equiv r_0^2 - 1 \geqslant 0. \tag{5.52}$$

One can calculate the integral over Ω in Eq. (5.50) by contour integration, closing the contour, say, in the upper half of complex plane. The poles of the first and the second terms on the second line of Eq. (5.51) are located at $\Omega_{0,\pm} = \pm i\sqrt{u^2 - r_0^2}$ and $\Omega_{1,\pm} = \pm i\sqrt{u^2 + r_1^2}$, respectively. For $u = \gamma + i\omega$ ($\gamma > 0$) the roots falling in the upper half plane are $\Omega_{0,+}$ and $\Omega_{1,+}$, independently of ω. From their residues we obtain

$$f_a(u) = \frac{1}{r_0^2 + r_1^2} \left[\frac{r_0^2 + |w_a|^2}{\sqrt{u^2 - r_0^2}} + \frac{r_1^2 - |w_a|^2}{\sqrt{u^2 + r_1^2}} \right]. \tag{5.53}$$

The integral of Eq. (5.53) in Eq. (5.49) is essentially the inverse Laplace transform of Eq. (5.53). Consulting a table of Laplace transforms [56] yields

$$\langle\|\mathbf{x}(t)\|^2\rangle_J$$
$$= e^{-2\gamma t} \left[\frac{r_0^2 + |w_a|^2}{r_0^2 + r_1^2} I_0(2r_0 t) + \frac{r_1^2 - |w_a|^2}{r_0^2 + r_1^2} J_0(2r_1 t) \right], \tag{5.54}$$

where $J_0(x)$ [$I_0(x)$] is the 0th Bessel function (modified Bessel function). From Eqs. (5.28) and (5.52) it follows that $r_0^2 + r_1^2 = \sqrt{1 + 4\mu^2}$, and using $\mu^2 = \langle|w_b|^2\rangle_b/2$ once again, we obtain

$$\langle\|\mathbf{x}(t)\|^2\rangle_J = \left[\frac{1 + C_a}{2} I_0(2r_0 t) + \frac{1 - C_a}{2} J_0(2r_1 t) \right] e^{-2\gamma t}, \tag{5.55}$$

where we defined

$$C_a \equiv \frac{1 + 2|w_a|^2}{\sqrt{1 + 2\langle|w_b|^2\rangle_b}}. \tag{5.56}$$

Effecting the proper rescalings we obtain the result for general σ, Eqs. (2.48) and (2.49).

C. Network with different neural types and independent, factorizable weights

Here we carry out the explicit calculations for the network with C neural types presented Sec. II C 2, with M, L, and R

given by Eqs. (2.52)–(2.55). From Eqs. (2.6) and (2.52)–(2.55) we obtain $M_z = z(RL)^{-1} - s\mathbf{u}\mathbf{u}^{\mathrm{T}}$ and

$$M_z M_z^\dagger = |z|^2 (RL)^{-2} - zs\mathbf{v}\mathbf{u}^{\mathrm{T}} - \bar{z}s\mathbf{u}\mathbf{v}^{\mathrm{T}} + s^2\mathbf{u}\mathbf{u}^{\mathrm{T}}, \quad (5.57)$$

where we defined $\mathbf{v} \equiv (RL)^{-1}\mathbf{u}$. Using the Woodbury matrix identity we can write

$$\frac{1}{g^2 + M_z M_z^\dagger} = Q - QU\frac{1}{D^{-1} + U^\dagger QU}U^\dagger Q, \quad (5.58)$$

where we defined the $N \times 2$ matrix $U = (\mathbf{u}, \mathbf{v})$, and

$$Q \equiv \frac{1}{g^2 + |z|^2(RL)^{-2}}, \quad (5.59)$$

$$D \equiv \begin{pmatrix} s^2 & -\bar{z}s \\ -zs & 0 \end{pmatrix}. \quad (5.60)$$

We argue that for $g > 0$, $\mathrm{tr}\,(g^2 + M_z M_z^\dagger)^{-1} = \mathrm{tr}\,Q$, up to $o(1)$ corrections. From Eq. (5.58), for the remainder $\Delta(g,z) \equiv \mathrm{tr}\,(g^2 + M_z M_z^\dagger)^{-1} - \mathrm{tr}\,Q$, we obtain

$$\Delta(g,z) = -\frac{1}{N}\mathrm{Tr}\left[\frac{U^\dagger Q^2 U}{D^{-1} + U^\dagger QU}\right], \quad (5.61)$$

where the trace is now over 2×2 matrices. We have

$$D^{-1} = -\begin{pmatrix} 0 & (zs)^{-1} \\ (\bar{z}s)^{-1} & |z|^{-2} \end{pmatrix} \quad (5.62)$$

and for $n = 1,2$ we obtain

$$U^\dagger Q^n U = \begin{pmatrix} \mathbf{u}^\dagger Q^n \mathbf{u} & \mathbf{u}^\dagger Q^n \mathbf{v} \\ \mathbf{u}^\dagger Q^n \mathbf{v} & \mathbf{v}^\dagger Q^n \mathbf{v} \end{pmatrix} = \begin{pmatrix} I_{n,0} & I_{n,1} \\ I_{n,1} & I_{n,2} \end{pmatrix}, \quad (5.63)$$

where

$$I_{n,k}(g,z) \equiv \frac{1}{N}\sum_{i=1}^N \frac{(l_{c(i)}r_{c(i)})^{-k}}{[g^2 + |z|^2(l_{c(i)}r_{c(i)})^{-2}]^n} \quad (5.64)$$

$$= \left\langle \frac{\sigma_c^{-k}}{(g^2 + \sigma_c^{-2}|z|^2)^n} \right\rangle_c \quad (5.65)$$

and we are using the notation Eq. (2.59) (we drop the explicit g and z dependence of $I_{n,k}$ when convenient). Note that all $I_{n,k}(g,z)$ are $O(1)$ and for even k are positive. Inverting $D^{-1} + U^\dagger QU$ we obtain

$$\Delta(g,z) = \frac{1}{N}\frac{T(g,z)}{-\det(D^{-1} + U^\dagger QU)}, \quad (5.66)$$

where

$T(g,z)$
$$\equiv \mathrm{Tr}\left[\begin{pmatrix} I_{2,0} & I_{2,1} \\ I_{2,1} & I_{2,2} \end{pmatrix}\begin{pmatrix} I_{1,2} - |z|^{-2} & (zs)^{-1} - I_{1,1} \\ (\bar{z}s)^{-1} - I_{1,1} & I_{1,0} \end{pmatrix}\right]$$
$$= I_{2,2}I_{1,0} + I_{2,0}\left(I_{1,2} - \frac{1}{|z|^2}\right) - 2I_{2,1}\left(I_{1,1} - \frac{1}{s\mathrm{Re}\,z}\right)$$

and

$$-\det(D^{-1} + U^\dagger QU) = I_{1,0}\left(\frac{1}{|z|^2} - I_{1,2}\right) + \left|I_{1,1} - \frac{1}{sz}\right|^2$$
$$= \frac{g^2}{|z|^2}I_{1,0}^2 + \left|I_{1,1} - \frac{1}{sz}\right|^2. \quad (5.67)$$

We see that both $T(g,z)$ and $-\det(D^{-1} + U^\dagger QU)$ are $O(1)$ [to obtain their limits as $N \to \infty$ we can set $s^{-1} = O(N^{-1/2})$ equal to zero] and since $-\det(D^{-1} + U^\dagger QU) \geqslant \frac{g^2}{|z|^2}I_{1,0}^2$ and $I_{1,0}^2 > 0$, we see that for $g > 0$, the denominator in Eq. (5.66) is bounded away from zero, and hence $\Delta(g,z) = O(N^{-1})$ and can be safely ignored for $g > 0$.

We thus use $\mathrm{tr}\,(g^2 + M_z M_z^\dagger)^{-1} = \mathrm{tr}\,Q + o(1)$. From Eq. (5.59) we obtain $\mathrm{tr}\,Q = I_{1,0}(g,z)$, and hence from Eqs. (2.19),

$$\mathcal{K}(g,z) = \lim_{N\to\infty}\mathrm{tr}\,Q = \left\langle \frac{1}{g^2 + \sigma_c^{-2}r^2} \right\rangle_c, \quad (5.68)$$

where $r \equiv |z|$. Note that the approximation $\mathrm{tr}\,(g^2 + M_z M_z^\dagger)^{-1} = \mathrm{tr}\,Q$ is equivalent to using $M_z M_z^\dagger = |z|^2(RL)^{-2}$ instead of the full expression Eq. (5.57) and hence to setting $M = 0$. Accordingly, the support of the eigenvalue distribution is given by Eq. (2.13), or equivalently by Eq. (2.60), and within this support, g^2 depends only on $|z| = r$ and is found by solving Eq. (2.15) or, equivalently, Eq. (2.62). Similar considerations show that in using Eq. (2.8) to obtain $\lim_{N\to\infty}\rho(z)$ we can set $M = 0$, yielding an isotropic eigenvalue density. From Eqs. (2.14)–(2.16), the proportion, $n_>$, of eigenvalues lying at a distance larger than r is equal to $g^2(r)$, which is found by solving Eq. (2.62). The results Eqs. (2.17) and (2.18) also hold, wherein the normalized sums over i can be replaced with appropriate averages $\langle\cdot\rangle_c$.

Let us now go back to the expression for $\Delta(g,z)$, and consider the case $g = 0$. In this case,

$$I_{n,k}(0,z) = |z|^{-2n}\langle\sigma_c^{2n-k}\rangle_c, \quad (5.69)$$

and we obtain

$$T(g,z) = |z|^{-6}(\langle\sigma_c^2\rangle_c^2 - 2\langle\sigma_c^3\rangle_c\langle\sigma_c\rangle_c) + 2\langle\sigma_c^3\rangle_c\frac{s^{-1}}{|z|^4\mathrm{Re}\,z} \quad (5.70)$$

and

$$-\det(D^{-1} + U^\dagger QU) = \left||z|^{-2}\langle\sigma_c\rangle_c - \frac{s^{-1}}{z}\right|^2. \quad (5.71)$$

In the special case in which $\langle\sigma_c\rangle_c = 0$ [this corresponds to the special case of the example Eq. (2.42) with $f\mu_E - (1-f)\mu_I \propto \mathbf{u}\cdot\mathbf{v} = 0$, which we considered above], the determinant will have a vanishing limit as $N \to \infty$ (or $s^{-1} \to 0$). This leads to a finite limit for $\Delta(0,z)$ and we obtain

$$\Delta(0,z) = \frac{s^2\langle\sigma_c^2\rangle_c^2}{N|z|^4} = \frac{\xi^2\langle\sigma_c^2\rangle_c^2}{|z|^4} \quad (\langle\sigma_c\rangle_c = 0). \quad (5.72)$$

Adding this to $\mathrm{tr}\,Q$ in the right side of Eq. (5.68), and using the naive formula Eq. (2.5) or $\mathcal{K}(0,z) = 1$ for the spectral boundary, we would have obtained the equation

$$1 = \frac{\langle\sigma_c^2\rangle_c}{r^2} + \frac{\xi^2\langle\sigma_c^2\rangle_c^2}{r^4}. \quad (5.73)$$

This, in turn, yields the radius Eq. (2.61), which is larger than the true boundary of the support of $\lim_{N\to\infty}\rho(z)$ given by Eq. (2.60).

6. CONCLUSIONS

We have provided a general formula for the eigenvalue density of partly random matrices, i.e., matrices with general mean and nontrivial covariance structure. General formulas have also been derived for the magnitude of impulse response and frequency power spectrum in an N-dimensional linear dynamical system with a coupling given by such partly random matrices. Our theory makes no requirement on the normality of matrices; its applications include, therefore, the stability and linear response analysis of neural circuits, whose linearized dynamics is always non-normal. We have demonstrated our theory by tackling analytically two specific neural circuits: a feedforward chain of length N and a set of randomly coupled feedforward subchains of length 2. A connection has also been revealed between the eigenvalue spectra of dense random matrix perturbations and the theory of pseudospectra.

The noncrossing diagrammatic method can be used to calculate other quantities of interest for matrix ensembles of the form $A = M + LJR$, considered here as well; possible examples are direct statistics of eigenvectors [57], or the correlations of the random fluctuations of the eigenvalue density $\langle \delta\rho_J(z)\delta\rho_J(z+w)\rangle_J$ for macroscopic w [i.e., for $|w| = \Theta(1)$]. On the other hand, quantities such as the microscopic structure of $\langle \delta\rho_J(z)\delta\rho_J(z+w)\rangle_J$, e.g., for $|w| = \Theta(N^{-1/2})$ with z inside the support, which could be of interest in the study of eigenvalue repulsion are not accessible to the noncrossing approximation. This is also the case, in general, for the statistics of the "outlier" eigenvalues that we discussed after Eq. (2.12) and in the examples of Secs. II C 1 and II C 2, which may be of importance in practical applications. The calculation of such quantities is possible, for example, by using the replica technique (see, e.g., Ref. [58]).

Finally, there are important forms of disorder which are not covered by the general ensemble $A = M + LJR$ with iid, and hence dense, J. Examples of relevance to neuroscientific applications include sparse A [59–61] [note that, e.g., binary matrices with probability of a nonzero weight, p, which is small but $\Theta(1)$ as $N \to \infty$ are covered by our formulas; by "sparse" disorder we refer, e.g., to the case $p = o(1)$] or more general structure of correlations between the elements of A [in the ensemble considered in this article, and for real J, the covariance $\langle \delta A_{ij}\delta A_{i'j'}\rangle_J = (LL^{\mathsf{T}})_{ii'}(R^{\mathsf{T}}R)_{jj'}$ is single rank]; the latter is of importance in considering networks with local topologies where, e.g., the matrix A has a banded structure. Generalization to other forms of random disorder is thus an important direction for future research.

APPENDIX A: VALIDITY OF THE NONCROSSING APPROXIMATION

In this appendix we give the justification for the noncrossing approximation used in Secs. III and IV. That is, we show that the only diagrams not suppressed by inverse powers of N are the noncrossing diagrams. We limit our discussion to the case of the eigenvalue density considered in Sec. III, but the generalization to the quantities calculated in Sec. IV is straightforward. As explained after Eq. (3.22), averaging of $\boldsymbol{G}(\eta,z;J)$ over the disorder J involves summing over all complete pairings of the factors of J in every term of the expansion, Eq. (3.22), with each pairing of each term represented by a diagram, as shown in Fig. 11. Each such diagram is composed of a solid directed line [each segment of which represents a factor of $\boldsymbol{G}_{ab}^{\alpha\beta}(\eta,z;0)$], with a number of wavy lines [each representing the expression Eq. (3.21), with different indices] connecting different points on the solid arrow line, and all the internal matrix indices summed over. For the purpose of calculating the eigenvalue density, according to Eqs. (3.16)–(3.18), what we need to calculate is actually $\mathrm{tr}\,[\sigma^+(RL)^{-1}\,\langle \boldsymbol{G}(\eta,z;J)\rangle_J]$; thus, we can imagine the solid arrow making a loop by closing in on itself sandwiching $\sigma^+ \otimes (RL)^{-1}$ (see Fig. 15).

Given the structure of the Kronecker δ's in Eq. (3.21), it is more convenient for our purpose here, however, to think of each diagram as a number of "orbits," each formed by starting somewhere on the solid line and moving on it always along its arrow until the next wavy line is encountered, whereby we leave the solid line, continuing on the wavy line *without* crossing it (because Eq. (3.21) is composed of two Kronecker δ's, one for each side of the wavy line, enforcing index identification at the corresponding ends on each side [62]) and return somewhere else on the solid line, continuing as before until we reach the initial point (see Fig. 15). As we go around this orbit, for each solid line traversed we write down, from right to left, a $\boldsymbol{G}(\eta,z;0)$ and for each wavy line a π^{r_i} [see Eqs. (3.20)] where i is the index of the wavy line. Because all matrix indices are summed over, such adjacent factors multiply like matrices, and since the orbit forms a loop, in the end we obtain the *trace* of the matrix product thus obtained. [This recipe for assigning the contribution of each orbit accounts for the Kronecker δ's and π^r's in Eq. (3.21), but not for the factor $\frac{1}{N}$ and the sum over r's; we account for these, at the end, after Eq. (A2).] A generic orbit, which we refer to as internal, closes on itself after traversing, say, m wavy lines sandwiching m Green's functions [e.g., the orbits labeled 2 and 3 in panel (a) of Fig. 15] and thus contributes a trace of the form

$$I_{m,\boldsymbol{r}} \equiv \mathrm{Tr}\,[\boldsymbol{G}(\eta,z;0)\pi^{r_{i_1}} \cdots \boldsymbol{G}(\eta,z;0)\pi^{r_{i_m}}], \qquad \text{(A1)}$$

where $\boldsymbol{r}$ is short-hand for $\{r_{i_1}, \ldots, r_{i_m}\}$, and i_k are the indices of the wavy lines traversed in the orbit. In every diagram, there is also exactly one orbit (e.g., the orbits labeled 1 in both panels of Fig. 15) which, in addition, includes the factor $\sigma^+(RL)^{-1}$ sandwiched between the two external Green's functions. This orbit, which we call the external orbit, contributes a trace of
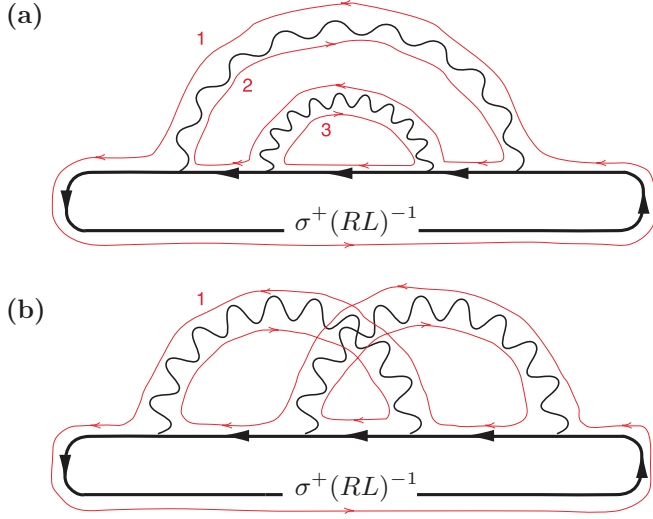
(a)

(b)

FIG. 15. (Color online) The orbits (shown by thin red paths) for two diagrams for the spectral density in a complex J ensemble. The noncrossing diagram on top has three orbits: orbit (1) is the external orbit connecting the two ends of the Green's function, while orbits (2) and (3) are the internal orbits. As in Eqs. (A1) and (A2), they contribute $\mathrm{tr}\,[\sigma^+ \boldsymbol{G}(\eta,z;0)\pi^{r_1}\boldsymbol{G}(\eta,z;0)]$, $\mathrm{Tr}\,[\pi^{3-r_1}\boldsymbol{G}(\eta,z;0)\pi^{r_2}\boldsymbol{G}(\eta,z;0)]$ and $\mathrm{Tr}\,[\pi^{3-r_2}\boldsymbol{G}(\eta,z;0)]$, respectively, with r_1 and r_2 summed over 1 and 2 [cf. Eq. (3.21)]. The trace contributed by each of the three orbits is $O(N)$, which, when combined with the three factors of $1/N$ accounting for the two wavy lines and the normalization of the external orbit's trace, yield an $O(1)$ expression for this diagram. By contrast, the crossing diagram on the right has no internal orbits. Its only external orbit contributes $\mathrm{Tr}\,[\sigma^+ \boldsymbol{G}(\eta,z;0)\pi^{r_2}\boldsymbol{G}(\eta,z;0)\pi^{3-r_1}\boldsymbol{G}(\eta,z;0)\pi^{3-r_2}$ $\boldsymbol{G}(\eta,z;0)\pi^{r_1}\boldsymbol{G}(\eta,z;0)]$ which after normalization is $O(1)$. Accounting for two factors of $1/N$ coming from the wavy lines, we then see that this crossing diagram is $O(N^{-2})$ and hence is suppressed as $N \to \infty$.

the form

$$E_{n,\tilde{r}} \equiv \mathrm{Tr}\,[\sigma^+(RL)^{-1}\boldsymbol{G}(\eta,z;0)\pi^{r_{j_1}}\cdots\pi^{r_{j_n}}\boldsymbol{G}(\eta,z;0)], \quad (\text{A2})$$

where n is the number of wavy lines the orbit traverses, $\tilde{r}$ is short for $\{r_{j_1},\ldots,r_{j_n}\}$, and j_k are the indices of the wavy lines traversed in this orbit [in writing Eq. (A2) we dropped the $\frac{1}{N}$ that normalizes the trace in Eqs. (3.16), but we account for it below]. For succinctness, in Eqs. (A1) and (A2) we suppressed the arguments (η,z) for $I_{m,r}$ and $E_{n,\tilde{r}}$ on which they depend. The full expression for the diagram is obtained by multiplying all such trace factors contributed by every orbit in the diagram, as well as a factor of N^{-w-1}, where w is the number of wavy lines in the diagrams, to account for the N^{-1} in Eq. (3.21) for each wavy line, as well as the extra N^{-1}, which normalizes the trace in the external orbit Eq. (A2), as dictated by Eq. (3.16). The obtained expression is finally summed over all the r indices corresponding to each wavy line, as required by Eq. (3.21).

The justification for the noncrossing approximation is based on the claim that each trace contributed by an orbit (external or internal) as in Eqs. (A1) and (A2) is $O(N)$, irrespective of η, z, m, or $\boldsymbol{r}$. We provide justification for this claim below. However, accepting it as true, we see that any diagram's scaling

with N solely depends on the number of orbits and wavy lines it contains. A well-known topological argument then shows that the contributions of crossing diagrams are suppressed by inverse powers of N [63,64]; for completeness, we summarize this argument here. First note that, assuming the claim, any diagram will yield an expression that is $O(N^\alpha)$ with

$$\alpha = f - w - 1, \quad (\text{A3})$$

where f is the number of orbits in the diagram (the sum over at most 2^w possible configurations of r_i does not contribute to the scaling with N). Let V denote the total number of vertices in the diagram [i.e., the number of intersections of wavy lines and the solid line, plus an extra one representing the insertion of $\sigma^+(RL)^{-1}$ in the solid line loop] and let E denote its total number of edges, i.e., $E \equiv w + s$, where s is the number of solid line segments ($s = 5$ in both panels of Fig. 15). It is easy to see that $V = s$. Thus, we have $E - V = w$. Formally defining the number of "faces" in the diagram by $F \equiv f + 1$, and its "Euler characteristic" by

$$\chi \equiv F - E + V, \quad (\text{A4})$$

we then find that $\chi = F - (E - V) = f + 1 - w$. From Eq. (A3) we then obtain

$$\alpha = \chi - 2. \quad (\text{A5})$$

Thus, the contribution of a diagram is $O(N^\alpha)$, with α determined solely by the diagram's formal "Euler characteristic" via Eq. (A5). It can be shown that a diagram with F formal "faces" and a formal "Euler characteristic" χ as defined above, can be drawn on (embedded in) a 2D oriented surface with Euler characteristic χ, such that no edges (solid or wavy) cross to create new vertices and each face created on the surface by its partitioning by the drawn diagram (a) is topologically a disk and (b) has a one-to-one correspondence with and is encircled by an orbit in the diagram, where we now count among the orbits, also the loop formed by the solid arrow line. Thus, the number of faces on the surface is indeed $F = f + 1$, and the χ, as defined above for the diagram, indeed agrees with the Euler characteristic of the surface, as conventionally defined. Topologically, such a surface is a generalized torus with g holes, satisfying $\chi = 2 - 2g$; the surface with zero holes is the sphere, or after decompactification, the plane [e.g., the diagram in panel (b) of Fig. 15 can be drawn in this manner on a torus]. We thus see that

$$\alpha = -2g, \quad (\text{A6})$$

and therefore the only diagrams that are not suppressed by inverse powers of N are those that can be drawn, as described above, on the plane. Since we took the area enclosed by the solid arrow line loop as a face by itself, this means that the diagram should be drawable with the wavy lines remaining outside this area (in order not to partition it into several faces) without crossing each other; this is the precise definition of the diagram being noncrossing [65].

Let us now go back to justifying the claim that the traces contributed by the orbits, as in Eqs. (A1) and (A2), are $O(N)$. For this purpose we make use of the SVD of M_z introduced in

Eq. (3.32). Defining the unitary matrix

$$\mathcal{U}_z \equiv \begin{pmatrix} U_z & 0 \\ 0 & V_z \end{pmatrix}, \tag{A7}$$

and using Eq. (3.32), we can write $H_0(z)$, defined in Eq. (3.14), as

$$H_0(z) = \mathcal{U}_z \tilde{H}_0(z) \mathcal{U}_z^\dagger, \tag{A8}$$

where

$$\tilde{H}_0(z) \equiv \begin{pmatrix} 0 & S_z \\ S_z & 0 \end{pmatrix}. \tag{A9}$$

Let us also define $\tilde{\bm{G}}(\eta,z;0) \equiv \mathcal{U}_z^\dagger \bm{G}(\eta,z;0)\mathcal{U}_z$, such that

$$\bm{G}(\eta,z;0) = \mathcal{U}_z \tilde{\bm{G}}(\eta,z;0)\mathcal{U}_z^\dagger. \tag{A10}$$

Then using the definition $\bm{G}(\eta,z;0) = [\eta - H_0(z)]^{-1}$, we see that

$$\tilde{\bm{G}}(\eta,z;0) = \frac{1}{\eta - \tilde{H}_0(z)} = \begin{pmatrix} \frac{\eta}{\eta^2-S_z^2} & \frac{S_z}{\eta^2-S_z^2} \\ \frac{S_z}{\eta^2-S_z^2} & \frac{\eta}{\eta^2-S_z^2} \end{pmatrix}, \tag{A11}$$

where we used Eq. (A9) to write the last equality. Given the block-diagonal nature of Eq. (A7) and the definitions, Eq. (3.20), we also have

$$\pi^r = \mathcal{U}_z \pi^r \mathcal{U}_z^\dagger. \tag{A12}$$

We now substitute $\bm{G}(\eta,z;0)$ and π^{r_i} in Eqs. (A1) and (A2) with the right hand sides of Eqs. (A10) and (A12), respectively. After canceling the $\mathcal{U}_z$'s we obtain

$$I_{m,\bm{r}} = \mathrm{Tr}\,[\tilde{\bm{G}}(\eta,z;0)\pi^{r_{i_1}} \cdots \tilde{\bm{G}}(\eta,z;0)\pi^{r_{i_m}}], \tag{A13}$$

$$E_{n,\tilde{\bm{r}}} = \mathrm{Tr}\,[\sigma^+ A(z)\tilde{\bm{G}}(\eta,z;0)\pi^{r_{j_1}} \cdots \pi^{r_{j_n}}\tilde{\bm{G}}(\eta,z;0)], \tag{A14}$$

where we defined

$$A(z) \equiv U_z^\dagger (RL)^{-1} V_z, \tag{A15}$$

such that $\mathcal{U}_z^\dagger[\sigma^+ \otimes (RL)^{-1}]\mathcal{U}_z = \sigma^+ \otimes A(z) \equiv \sigma^+ A(z)$. For the internal orbits, we see from Eq. (A11) that each $\tilde{\bm{G}}(\eta,z;0)$, depending on whether it is sandwiched between the same projectors π^r, or between two opposite projectors, π^r and π^{3-r}, contributes a *diagonal* factor equal to $\eta/(\eta^2 - S_z^2)$ or $S_z/(\eta^2 - S_z^2)$, respectively. Thus, for any configuration of r_i's, if the number of Green's functions sandwiched the second way is k ($1 \leqslant k \leqslant m$), we obtain

$$I_{m,\bm{r}}(\eta,z) = \sum_{i=1}^N \frac{\eta^{m-k}s_i(z)^k}{[\eta^2 - s_i(z)^2]^m} \quad (1 \leqslant k \leqslant m), \tag{A16}$$

for the internal orbits [in particular, we see that the sole dependence of $I_{m,\bm{r}}(\eta,z)$ on $\bm{r}$ is via the number k]. We therefore have

$$|I_{m,\bm{r}}(\eta,z)| \leqslant N \max_i \left| \frac{\eta^{m-k}s_i(z)^k}{[\eta^2 - s_i(z)^2]^m} \right|. \tag{A17}$$

When the imaginary part of η is nonzero, the denominator on the right hand side of Eq. (A17) cannot vanish for any value of $s_i(z)$ [while as $\eta \to i0$, which is the limit we have to take after summing up the relevant diagrammatic series, $s_i(z)$, that approach zero as N grows can make this expression unbounded

as $N \to \infty$]. Assuming $\mathrm{Im}\,\eta > 0$, it will be sufficient for our purposes to substitute Eq. (A17) with the weaker bound,

$$|I_{m,\bm{r}}(\eta,z)| \leqslant N \max_s \left| \frac{\eta^{m-k}s^k}{(\eta^2 - s^2)^m} \right| \quad (\mathrm{Im}\,\eta > 0), \tag{A18}$$

where now the maximum is taken for s ranging over the whole $[0,\infty)$. Since $\mathrm{Im}\,\eta > 0$ the expression has no singularities at finite real s, and since $2m > k$, it cannot diverge as $s \to \infty$ either; thus, it has a finite maximum independent of N. More precisely, it is easy to show that $\max_s |\frac{\eta^{m-k}s^k}{(\eta^2-s^2)^m}| \leqslant [\frac{\sqrt{2}}{|\mathrm{Im}\,\eta|}]^m$, irrespective of k as long as $1 \leqslant k \leqslant m$, yielding

$$|I_{m,\bm{r}}(\eta,z)| \leqslant N \left[\frac{\sqrt{2}}{|\mathrm{Im}\,\eta|} \right]^m \quad (\mathrm{Im}\,\eta > 0). \tag{A19}$$

Similarly, the trace for the external orbit can be written in the new basis Eq. (A11) as

$$E_{n,\tilde{\bm{r}}}(\eta,z) = \sum_{i=1}^N A_{ii}(z) \frac{\eta^{n-\tilde{k}}s_i(z)^{\tilde{k}}}{[\eta^2 - s_i(z)^2]^n} \quad (1 \leqslant \tilde{k} \leqslant n), \tag{A20}$$

where $\tilde{k}$ is the number of Green's functions in Eq. (A2) sandwiched between two π^r's with different superscripts; this convention works correctly for the external orbit as well, if we account for the presence of σ^+ by imagining a π^2 (π^1) to the left (right) of the leftmost (rightmost) Green's function. From Eq. (A15), we can write $A_{ii}(z) = \mathbf{u}_i(z)^\dagger (RL)^{-1}\mathbf{v}_i(z)$, where we defined the vectors $\mathbf{u}_i(z)$ and $\mathbf{v}_i(z)$ to be the ith column of U_z and V_z, respectively. By the Cauchy-Schwartz inequality we then have

$$|A_{ii}(z)| \leqslant \|\mathbf{u}_i(z)\|\|(RL)^{-1}\mathbf{v}_i(z)\|$$
$$\leqslant \|\mathbf{u}_i(z)\|\|\mathbf{v}_i(z)\|\|(RL)^{-1}\|, \tag{A21}$$

where $\|(RL)^{-1}\|$ is the operator norm, or the maximum singular value, of $(RL)^{-1}$. However, since U_z and V_z are unitary matrices, $\mathbf{u}_i(z)$ and $\mathbf{v}_i(z)$ are unit vectors, and we obtain

$$|A_{ii}(z)| \leqslant \|(RL)^{-1}\|. \tag{A22}$$

Going back to Eq. (A20), this yields the bound

$$|E_{n,\tilde{\bm{r}}}(\eta,z)| \leqslant N\|(RL)^{-1}\| \max_i \left| \frac{\eta^{n-\tilde{k}}s_i(z)^{\tilde{k}}}{[\eta^2 - s_i(z)^2]^n} \right|. \tag{A23}$$

The only difference with the inequality for $I_{m,\bm{r}}$ is the factor $\|(RL)^{-1}\|$. Repeating the same argument as for the internal traces, we therefore see that

$$|E_{n,\tilde{\bm{r}}}(\eta,z)| \leqslant N \left[\frac{\sqrt{2}}{|\mathrm{Im}\,\eta|} \right]^n \|(RL)^{-1}\| \quad (\mathrm{Im}\,\eta > 0), \tag{A24}$$

and thus a sufficient condition for $E_{n,\tilde{\bm{r}}}$ to be $O(N)$ for $\mathrm{Im}\,\eta > 0$, is that $\|(RL)^{-1}\|$ remains bounded as $N \to \infty$, i.e.,

$$\|(RL)^{-1}\| = O(1). \tag{A25}$$

Combining Eqs. (A19) and (A24), and given the prescription after Eq. (A2), we can bound the absolute value of the contribution of a diagram with genus g (or g crossings), w wavy lines, and s solid lines, by $2^w[\frac{\sqrt{2}}{|\mathrm{Im}\,\eta|}]^s \|(RL)^{-1}\|N^{-2g}$ [the power of s is obtained by noting that the powers of m and n in the bounds Eqs. (A19) and (A24), when summed over all

orbits must equal s, since every Green's function or solid line appears in exactly one orbit]. Hence, for a fixed, nonzero Im η, the contribution of crossing diagrams (i.e., those with $g \geqslant 1$) goes to zero as $N \to \infty$.

Thus, if we take the limit $N \to \infty$ before the limit $\eta \to i0^+$, ignoring the crossing diagrams is safe, and the expression for $\rho(z)$ obtained from Eq. (3.16) after analytic continuation of $\mathrm{tr}\,[\sigma^+(RL)^{-1}\mathcal{G}(\eta,z)]$ to $\eta = i0$, with $\mathcal{G}(\eta,z)$ given by the contribution of noncrossing diagrams to $\langle \boldsymbol{G}(\eta,z;J)\rangle_J$, gives the correct result for $\lim_{N\to\infty} \rho(z)$. We mention that when the smallest singular value $s_i(z)$ remains bounded away from zero as $N \to \infty$, even at $\eta = 0$ the traces Eqs. (A16) and (A20) are $O(N)$, as is not hard to check, justifying the noncrossing approximation at $\eta = 0$. Thus, it is only when some $s_i(z)$ are $o(1)$ that it becomes important to send η to $i0^+$ only after the limit $N \to \infty$ has been taken. In particular, in such cases, applying the limit $\eta \to i0^+$ to the results obtained using the noncrossing approximation before taking the limit $N \to \infty$ may yield finite-size contributions to $\lim_{N\to\infty} \rho(z)$, which, in general, may yield incorrect subleading corrections.

APPENDIX B: $\rho(z)$ VANISHES IN THE REGION EQ. (3.38)

In this appendix we prove more rigorously that in the region Eq. (3.38), the eigenvalue density vanishes. More precisely, we prove that $\rho(z) \equiv \lim_{\epsilon \to 0^+} \lim_{N\to\infty} \rho_N(z,\epsilon) = 0$, where

$$\rho_N(z,\epsilon) \equiv \frac{1}{\pi} \frac{\partial}{\partial \bar{z}} \, \mathrm{tr}\left[\frac{(RL)^{-1} M_z^\dagger}{M_z M_z^\dagger + \gamma^2} \right] \quad \text{(B1)}$$

is obtained by substituting Eq. (3.35) into Eqs. (3.17). Here $\gamma = g(z,\epsilon) + \epsilon$ is the solution of Eq. (3.37), which, as we argued in Sec. III, vanishes as $\epsilon \to 0^+$ when z is in the region Eq. (3.38) [note that since Eq. (3.37) is defined in the limit $N \to \infty$, γ has no dependence on N]. Recall that for $\epsilon > 0$, $g(z,\epsilon)$ is positive and therefore $\gamma > \epsilon > 0$. Expanding the derivative in Eq. (B1) we obtain

$$\pi\rho_N(z,\epsilon) = \mathrm{tr}\left[\frac{(RL)^{-1}(RL)^{-\dagger}}{M_z M_z^\dagger + \gamma^2} \right]$$
$$- \mathrm{tr}\left[(RL)^{-1} M_z^\dagger \frac{1}{M_z M_z^\dagger + \gamma^2} \right.$$
$$\left. \times M_z(RL)^{-\dagger} \frac{1}{M_z M_z^\dagger + \gamma^2} \right]$$
$$- \mathrm{tr}\left[\frac{(RL)^{-1} M_z^\dagger}{M_z M_z^\dagger + \gamma^2} \frac{1}{M_z M_z^\dagger + \gamma^2} \right] \partial_{\bar{z}}(\gamma^2)$$
$$= \mathrm{tr}\left[\frac{(RL)^{-1} Q (RL)^{-\dagger}}{M_z M_z^\dagger + \gamma^2} \right]$$
$$- \mathrm{tr}\left[\frac{(RL)^{-1} M_z^\dagger}{M_z M_z^\dagger + \gamma^2} \frac{1}{M_z M_z^\dagger + \gamma^2} \right] \partial_{\bar{z}}(\gamma^2), \quad \text{(B2)}$$

where we defined $Q = \mathbf{1} - M_z^\dagger \frac{1}{M_z M_z^\dagger + \gamma^2} M_z$ (we suppress the explicit dependence of γ on z for simplicity). By the Woodbury matrix identity $Q = \frac{\gamma^2}{M_z^\dagger M_z + \gamma^2}$, which upon substitution in

Eq. (B2) yields

$$\pi\rho_N(z,\epsilon) = \mathrm{tr}\left[\frac{(RL)^{-1}}{M_z^\dagger M_z + \gamma^2} \frac{(RL)^{-\dagger}}{M_z M_z^\dagger + \gamma^2} \right] \gamma^2$$
$$- \mathrm{tr}\left[\frac{(RL)^{-1} M_z^\dagger}{M_z M_z^\dagger + \gamma^2} \frac{1}{M_z M_z^\dagger + \gamma^2} \right] \partial_{\bar{z}}(\gamma^2). \quad \text{(B3)}$$

Differentiating Eq. (3.37) with respect to $\bar{z}$ yields

$$-\partial_{\bar{z}}(\gamma^2) = \frac{-2\gamma^2 \partial_{\bar{z}}\mathcal{K}}{1 - \mathcal{K} - 2\gamma^2 \partial_{\gamma^2}\mathcal{K}}, \quad \text{(B4)}$$

with the partial derivatives of $\mathcal{K}(\gamma,z)$ given by

$$-\partial_{\gamma^2}\mathcal{K} = \mathcal{T}_\infty^>(\gamma) \equiv \lim_{N\to\infty} \mathcal{T}_N^>(\gamma),$$
$$-\partial_{\bar{z}}\mathcal{K} = \mathcal{V}_\infty(\gamma) \equiv \lim_{N\to\infty} \mathcal{V}_N(\gamma), \quad \text{(B5)}$$

where we defined

$$\mathcal{T}_N^>(\gamma) \equiv \mathrm{tr}\left[\frac{1}{(M_z M_z^\dagger + \gamma^2)^2} \right], \quad \text{(B6)}$$

$$\mathcal{V}_N(\gamma) \equiv \mathrm{tr}\left[\frac{1}{M_z M_z^\dagger + \gamma^2} \frac{M_z(RL)^{-\dagger}}{M_z M_z^\dagger + \gamma^2} \right]. \quad \text{(B7)}$$

We thus obtain

$$\pi\rho_N(z,\epsilon) = \gamma^2 \mathcal{T}_N(\gamma) + \frac{2\gamma^2 \mathcal{V}_\infty(\gamma) \mathcal{V}_N(\gamma)^*}{1 - \mathcal{K}(\gamma) + 2\gamma^2 \mathcal{T}_\infty^>(\gamma)}, \quad \text{(B8)}$$

where we defined

$$\mathcal{T}_N(\gamma) \equiv \mathrm{tr}\left[\frac{(RL)^{-1}}{M_z^\dagger M_z + \gamma^2} \frac{(RL)^{-\dagger}}{M_z M_z^\dagger + \gamma^2} \right]. \quad \text{(B9)}$$

Having eliminated derivatives of γ, we now simply need to show that $\lim_{\gamma\to 0^+} \lim_{N\to\infty}$ of the right side of Eq. (B8) vanishes for z is in the region Eq. (3.38) [where $\gamma = 0^+$ is the solution of Eq. (3.37) as $\epsilon \to 0^+$].

We will start by bounding the traces $\mathcal{T}_N(\gamma)$ and $\mathcal{V}_N(\gamma)$ in Eq. (B8). For $\mathcal{V}_N(\gamma)$ we use the SVD Eq. (3.32),

$$|\mathcal{V}_N(\gamma)| = \left| \mathrm{tr}\left[\frac{(RL)^{-1} M_z^\dagger}{M_z M_z^\dagger + \gamma^2} \frac{1}{M_z M_z^\dagger + \gamma^2} \right] \right|$$
$$= \mathrm{tr}\left[U_z^\dagger (RL)^{-1} V_z \frac{S_z}{(S_z^2 + \gamma^2)^2} \right]$$
$$\leqslant \|(RL)^{-1}\| \mathrm{tr}\left[\frac{S_z}{(S_z^2 + \gamma^2)^2} \right] \quad \text{(B10)}$$

[where in the last line we used Eq. (A22)], i.e.,

$$|\mathcal{V}_N(\gamma)| \leqslant \|(RL)^{-1}\| \mathcal{V}_N^>(\gamma), \quad \text{(B11)}$$

where we defined

$$\mathcal{V}_N^>(\gamma) \equiv \frac{1}{N} \sum_{i=0}^N \frac{s_i(z)}{[s_i(z)^2 + \gamma^2]^2}. \quad \text{(B12)}$$

Taking the limit $N \to \infty$, we obtain from Eq. (B11)

$$|\mathcal{V}_\infty(\gamma)| \leqslant C \mathcal{V}_\infty^>(\gamma), \quad \text{(B13)}$$

where $\mathcal{V}_\infty^>(\gamma) \equiv \lim_{N\to\infty} \mathcal{V}_N^>(\gamma)$, and C is an upper bound on $\|(LR)^{-1}\|$ [which we have assumed is $O(1)$ as $N \to \infty$]. To bound $\mathcal{T}_N(\gamma)$, we use the inequality

$$|\text{tr}\,(ABCD)| \leqslant \|A\|\|C\|\text{tr}\,(BB^\dagger)^{\frac{1}{2}}\text{tr}\,(DD^\dagger)^{\frac{1}{2}}. \qquad \text{(B14)}$$

This can be derived by first using the Cauchy-Schwartz inequality, $|\text{tr}\,(AB)|^2 \leqslant \text{tr}\,(AA^\dagger)\text{tr}\,(BB^\dagger)$, and then using the inequality $|\text{tr}\,(AB)| \leqslant \|B\|\text{tr}\,(A)$, valid for positive semidefinite A (which in turn follows from the definition of $\|B\|$ after unitary diagonalization of A). Using (B14) we obtain

$$\left|\text{tr}\left[\frac{(RL)^{-1}}{M_z^\dagger M_z + \gamma^2}\frac{(RL)^{-\dagger}}{M_z M_z^\dagger + \gamma^2}\right]\right|$$
$$\leqslant \|(RL)^{-1}\|^2\,\text{tr}\left[\left(\frac{1}{M_z M_z^\dagger + \gamma^2}\right)^2\right] \qquad \text{(B15)}$$

or

$$|\mathcal{T}_N(\gamma)| \leqslant \|(RL)^{-1}\|^2 \mathcal{T}_N^>(\gamma). \qquad \text{(B16)}$$

Using the inequalities (B11), (B13), and (B16) in Eq. (B8) we obtain

$$\pi|\rho_N(z,\epsilon)| \leqslant C^2\left[\gamma^2\mathcal{T}_N^>(\gamma) + \frac{2\gamma^2\mathcal{V}_\infty^>(\gamma)\mathcal{V}_N^>(\gamma)}{1 - \mathcal{K}(\gamma) + 2\gamma^2\mathcal{T}_\infty^>(\gamma)}\right]. \qquad \text{(B17)}$$

Taking the $N \to \infty$ limit (while keeping γ finite), and defining $\rho(z,\epsilon) \equiv \lim_{N\to\infty}\rho_N(z,\epsilon)$, we obtain

$$\pi\rho(z,\epsilon) \leqslant C^2\left\{\gamma^2\mathcal{T}_\infty^>(\gamma) + \frac{2[\gamma\mathcal{V}_\infty^>(\gamma)]^2}{1 - \mathcal{K}(\gamma) + 2\gamma^2\mathcal{T}_\infty^>(\gamma)}\right\}, \qquad \text{(B18)}$$

where we defined

$$\mathcal{T}_\infty^>(\gamma) \equiv \lim_{N\to\infty}\frac{1}{N}\sum_{i=0}^{N}\frac{1}{[s_i(z)^2 + \gamma^2]^2}, \qquad \text{(B19)}$$

$$\mathcal{V}_\infty^>(\gamma) \equiv \lim_{N\to\infty}\frac{1}{N}\sum_{i=0}^{N}\frac{s_i(z)}{[s_i(z)^2 + \gamma^2]^2}. \qquad \text{(B20)}$$

Thus, to show that $\lim_{\epsilon\to 0^+}\rho(z,\epsilon) = 0$, it suffices to show that $\gamma^2\mathcal{T}_\infty^>(\gamma)$ and $\gamma\mathcal{V}_\infty^>(\gamma)$ vanish as $\gamma \to 0^+$ [since z is in the region Eq. (3.38), $1 - \mathcal{K}(\gamma)$ and hence the denominator in the

last term in Eq. (B18) remains positive as $\gamma \to 0^+$]. Let us rewrite Eq. (B19) as

$$\mathcal{T}_\infty^>(\gamma) = \int_0^\infty\frac{\rho_s(s;z)ds}{(s^2 + \gamma^2)^2}, \qquad \text{(B21)}$$

$$\mathcal{V}_\infty^>(\gamma) = \int_0^\infty\frac{s\rho_s(s;z)ds}{(s^2 + \gamma^2)^2}, \qquad \text{(B22)}$$

where we defined

$$\rho_s(s;z) = \lim_{N\to\infty}\frac{1}{N}\sum_{i=0}^{N}\delta[s - s_i(z)] \qquad \text{(B23)}$$

as the limit of the density of the singular values of M_z [66]. Note that contributions to $\mathcal{T}_\infty^>(\gamma)$ and $\mathcal{V}_\infty^>(\gamma)$ from integration on $[s_0,\infty)$ for any fixed, nonzero s_0 remain finite as $\gamma \to 0^+$; only singular contributions arising from the region $s = O(\gamma) \ll 1$ can contribute to $\gamma^2\mathcal{T}_\infty^>(\gamma)$ and $\gamma\mathcal{V}_\infty^>(\gamma)$ as $\gamma \to 0^+$. Thus, we only need concern ourselves with the portion of integrals from 0 to some arbitrary small, but fixed s_0, and show that $\gamma^2\int_0^{s_0}\frac{\rho_s(s;z)ds}{(s^2+\gamma^2)^2}$ and $\gamma\int_0^{s_0}\frac{s\rho_s(s;z)ds}{(s^2+\gamma^2)^2}$ vanish as $\gamma \to 0^+$.

Let us first consider the situation similar to that in the two examples Eqs. (5.1) and (2.42). For those examples, there is a region of z outside Eq. (3.38), where a single [more generally $O(1)$] singular value $s_i(z)$ vanishes as $N \to \infty$, while all the other $s_i(z)$ remain bounded from below. However, an $O(1)$ set of (vanishing) singular values does not contribute to the density Eq. (B23) and since the other $s_i(z)$ are bounded from below, there is an s_0 below which $\rho_s(s;z)$ identically vanishes. So the claim is clearly true for such cases. More generally, we exploit the fact that z is in the region Eq. (3.38), so that

$$\lim_{\gamma\to 0^+}\int_0^\infty\frac{\rho_s(s;z)ds}{s^2 + \gamma^2} < 1. \qquad \text{(B24)}$$

We conclude that as $s \to 0^+$ the density, $\rho_s(s;z)$, must vanish at least as fast as s^α; i.e., it must be $O(s^\alpha)$, for some $\alpha > 1$. Otherwise, the integral in Eq. (B24) diverges in the limit. Let us therefore choose s_0 to be small enough such that for $s \leqslant s_0$, $\rho(s;z) < cs^\alpha$ for some constant c and $\alpha > 1$. It is then an elementary exercise to show that $\gamma^2\int_0^{s_0}\frac{s^\alpha ds}{(s^2+\gamma^2)^2}$ and $\gamma\int_0^{s_0}\frac{s^{\alpha+1}ds}{(s^2+\gamma^2)^2}$ are $O(\gamma^{\min(2,\alpha-1)})$ and $O(\gamma^{\min(1,\alpha-1)})$, respectively, as $\gamma \to 0^+$, and since $\alpha > 1$, they both vanish in the limit, proving the claim.

[1] M. L. Mehta, *Random Matrices* (Academic Press, San Diego, 2004).

[2] Z. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices* (Science Press, Beijing, 2006).

[3] T. Guhr, A. Müller-Groeling, and H. A. Weidenmüller, Phys. Rep. **299**, 189 (1998).

[4] H. Dale, Proc. R. Soc. Med. **28**, 319 (1935).

[5] J. C. Eccles, P. Fatt, and K. Koketsu, J. Physiol. **126**, 524 (1954).

[6] P. Strata and R. Harvey, Brain. Res. Bull. **50**, 349 (1999).

[7] B. K. Murphy and K. D. Miller, Neuron **61**, 635 (2009).

[8] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabasi, Nature (London) **407**, 651 (2000).

[9] A. L. Barabasi and Z. N. Oltvai, Nat. Rev. Genet. **5**, 101 (2004).

[10] X. Zhu, M. Gerstein, and M. Snyder, Genes Dev. **21**, 1010 (2007).

[11] M. Vidal, M. E. Cusick, and A.-L. Barabási, Cell **144**, 986 (2011).

[12] R. M. May, Nature (London) **238**, 413 (1972).

[13] J. Camacho, R. Guimerà, and L. A. Nunes Amaral, Phys. Rev. Lett. **88**, 228102 (2002).

[14] F. S. Valdovinos, R. Ramos-Jiliberto, L. Garay-Narváez, P. Urbani, and J. A. Dunne, Ecol. Lett. **13**, 1546 (2010).

[15] J. E. Vermaat, J. A. Dunne, and A. J. Gilbert, Ecology **90**, 278 (2009).

[16] R. Guimerà, D. B. Stouffer, M. Sales-Pardo, E. A. Leicht, M. E. J. Newman, and L. A. N. Amaral, Ecology **91**, 2941 (2010).

[17] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra* (Princeton University Press, Princeton, NJ, 2005).

[18] S. Ganguli, D. Huh, and H. Sompolinsky, Proc. Natl. Acad. Sci. U.S.A. **105**, 18970 (2008).

[19] M. S. Goldman, Neuron **61**, 621 (2009).

[20] M. G. Neubert and H. Caswell, Ecology **78**, 653 (1997).

[21] X. Chen and J. E. Cohen, Proc. Biol. Sci. **268**, 869 (2001).

[22] S. Tang and S. Allesina, Front. Ecol. Evol. **2** (2014).

[23] J. H. McCoy, New J. Phys. **15**, 113036 (2013).

[24] J. Feinberg and A. Zee, Nucl. Phys. B **504**, 579 (1997).

[25] J. Feinberg and A. Zee, Nucl. Phys. B **501**, 643 (1997).

[26] J. Ginibre, J. Math. Phys. **6**, 440 (1965).

[27] V. L. Girko, Theory Probab. Its Appl. (Engl. Transl.) **29**, 694 (1984).

[28] Z. D. Bai, Ann. Probab. **25**, 494 (1997).

[29] T. Tao and V. Vu, Commun. Contemp. Math. **10**, 261 (2008).

[30] F. Götze and A. Tikhomirov, Ann. Probab. **38**, 1444 (2010).

[31] B. Khoruzhenko, J. Phys. A: Math. Gen. **29**, L165 (1996).

[32] P. Biane and F. Lehner, Colloq. Math. **90**, 181 (2001).

[33] S. Hikami and R. Pnini, J. Phys. A: Math. Gen. **31**, L587 (1998).

[34] Since we present results for the limit of the eigenvalue density, etc., as $N \to \infty$, M, L, R, and J must each be more precisely understood as an infinite sequence of matrices dependent on N.

[35] T. Tao, V. Vu, and M. Krishnapur, Ann. Probab. **38**, 2023 (2010).

[36] T. Tao and V. Vu, Acta Math. **206**, 127 (2011).

[37] T. Tao, Probab. Theory Relat. Fields **155**, 231 (2013).

[38] S. O'Rourke and D. Renfrew, Electron. J. Probab. **19**, 1 (2014).

[39] S. F. Edwards and R. C. Jones, J. Phys. A: Math. Gen. **9**, 1595 (1976).

[40] The designation "highly non-normal" can be motivated, when L and R are proportional to the identity matrix, as follows. Let us denote the (operator norm based) ϵ-pseudospectrum of M, i.e., the region of z's over which $\|(z - M)^{-1}\| > \epsilon^{-1}$, by $\Sigma_\epsilon(M)$. For fixed N, the true spectrum of M, which we denote by $\Sigma(M)$, is the set of points over which the smallest singular value of $(z - M)$ is exactly zero and hence $\|(z - M)^{-1}\| = \infty$. For finite N, $\lim_{\epsilon \to 0^+} \Sigma_\epsilon(M) = \Sigma(M)$ for any M. However, for non-normal M this approach could be much slower than in the normal case (see our discussion in Sec. II A 2, and the book [17] for a complete discussion of pseudospectra and their relationship with non-normality). Now suppose that, as in the atypical cases under discussion, in a finite region of the complex plane the smallest singular value of M_z is nonzero for finite N, but vanishes in the limit $N \to \infty$. This means that the operator norm of $(z - M)^{-1} \propto M_z^{-1}$ is finite over such a region but goes to infinity as $N \to \infty$. Hence, if we define $\Sigma_\epsilon^\infty(M) \equiv \lim_{N \to \infty} \Sigma_\epsilon(M)$ and $\Sigma^\infty(M) \equiv \lim_{N \to \infty} \Sigma(M)$, we see that in such cases $\lim_{\epsilon \to 0^+} \Sigma_\epsilon^\infty(M) \neq \Sigma^\infty(M)$ [or equivalently, $\lim_{\epsilon \to 0^+} \lim_{N \to \infty} \Sigma_\epsilon(M) \neq \lim_{N \to \infty} \lim_{\epsilon \to 0^+} \Sigma_\epsilon(M)$]. More generally but less precisely, this indicates that at finite but large N, the ϵ pseudospectra of such matrices can cover a significantly broader region than the spectrum even for very small ϵ, indicating extreme non-normality.

[41] K. Rajan and L. F. Abbott, Phys. Rev. Lett. **97**, 188104 (2006).

[42] This equivalence is true more generally for any matrix norm derived from a general vector norm; see Ref. [17] for a proof.

[43] Y. Yin, Z. Bai, and P. Krishnaiah, Probab. Theory Relat. Fields **78**, 509 (1988).

[44] R. A. Horn and R. G. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, UK, 1990).

[45] The unitary invariance of these formulas is, in turn, a consequence of the invariance of both the corresponding quantities $[\rho(z)$ and $\|\mathbf{x}(t)\|^2]$, as well as the statistical ensemble for J, Eq. (3.2), and hence that of LJR when $L \propto R \propto \mathbf{1}$, under unitary transforms like Eq. (2.34).

[46] P. Jonas, J. Bischofberger, and J. Sandkühler, Science **281**, 419 (1998).

[47] D. H. Root, C. A. Mejias-Aponte, S. Zhang, H.-L. Wang, A. F. Hoffman, C. R. Lupica, and M. Morales, Nat. Neurosci. **17**, 1543 (2014).

[48] D. Chafaï, J. Theor. Probab. **23**, 945 (2010).

[49] Y. Wei, Phys. Rev. E **85**, 066116 (2012).

[50] It is also common to write the firing rate equations in the different form, $T \frac{dr(t)}{dt} = -r(t) + f[Wr(t) + I^r(t)]$. At least in the case where all neurons have equal time constants, i.e., $T \propto \mathbf{1}$, the two formulations are equivalent and are related by the change of variable $v = Wr + I^r$ [51].

[51] K. D. Miller and F. Fumarola, Neural Comput. **24**, 25 (2012).

[52] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, UK, 1998).

[53] M. Stern, H. Sompolinsky, and L. F. Abbott, Phys. Rev. E **90**, 062710 (2014).

[54] H. Sompolinsky, A. Crisanti, and H. J. Sommers, Phys. Rev. Lett. **61**, 259 (1988).

[55] J. Feinberg, J. Phys. A: Math. Gen. **39**, 10029 (2006).

[56] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Wiley-Interscience, New York, 1970).

[57] B. Mehlig and J. T. Chalker, J. Math. Phys. **41**, 3233 (2000).

[58] S. M. Nishigaki and A. Kamenev, J. Phys. A: Math. Gen. **35**, 4571 (2002).

[59] T. Rogers and I. P. Castillo, Phys. Rev. E **79**, 012101 (2009).

[60] F. Slanina, Phys. Rev. E **83**, 011118 (2011).

[61] I. Neri and F. L. Metz, Phys. Rev. Lett. **109**, 030602 (2012).

[62] This structure is a consequence of using a complex ensemble for J, for which the covariances $\langle J_{ab} J_{cd} \rangle$ vanish. For the real Gaussian ensemble, by contrast, the latter do not vanish; in this case Eq. (3.21) becomes $\langle J_{ab}^{\alpha\beta} J_{cd}^{\gamma\delta} \rangle = \frac{1}{N} [\delta_{ad}\delta_{bc} (\sigma_{\alpha\beta}^+ \sigma_{\gamma\delta}^- + \sigma_{\alpha\beta}^- \sigma_{\gamma\delta}^+) + \delta_{ac}\delta_{bd} (\sigma_{\alpha\beta}^+ \sigma_{\gamma\delta}^- + \sigma_{\alpha\beta}^- \sigma_{\gamma\delta}^-)] = \frac{1}{N} \sum_{r=1}^2 [(\pi_{\alpha\beta}^r \delta_{ad})(\pi_{\beta\gamma}^{3-r} \delta_{bc}) + (\pi_{\alpha\gamma}^r \delta_{ac})(\pi_{\beta\delta}^{3-r} \delta_{bd})]$.

[63] G. t. Hooft, Nucl. Phys. B **72**, 461 (1974).

[64] E. Brézin, C. Itzykson, G. Parisi, and J. B. Zuber, Commun. Math. Phys. **59**, 35 (1978).

[65] Notice that this is a more restrictive property than planarity of the diagram; for example, the graph in panel (b) of Fig. 15 is planar, as one of the wavy lines can be drawn inside the solid loop without crossing any other line, but it is not noncrossing as defined here.

[66] More precisely, we only need to define the limit Eq. (B23) in the sense of distributions, i.e., such that for any regular test function, $f(s^2)$, bounded at infinity and regular everywhere, including at $s^2 \to 0^+$, we have $\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^N f[s_i(z)^2] = \int_0^\infty f(s^2)\rho_s(s; z)ds$. We do not assume any smooth form for $\rho_s(s; z)$; in particular, $\rho_s(s; z)$ may have δ function singularities when an $O(N)$ singular values converge to the same value as $N \to \infty$. Also note that this assumption does not forbid the possibility that some $s_i(z)$ diverge as $N \to \infty$; our requirement that $\|M\|_F$ remain bounded automatically guarantees that these will not be numerous enough to contribute to $\rho_s(s; z)$ at infinity.