Properties of networks with partially structured and partially random connectivity

Yashar Ahmadian,1,2 Francesco Fumarola,1 and Kenneth D. Miller1,2

1Center for Theoretical Neuroscience, Department of Neuroscience,
2Swartz Program in Theoretical Neuroscience, and Kavli Institute for Brain Science,
College of Physicians and Surgeons, Columbia University, NY, NY 10032
(Dated: May 19, 2014)

Networks studied in many disciplines, including neuroscience, have connectivity that may be stochastic about some underlying mean connectivity represented by a nonnormal matrix. Furthermore the stochasticity may not be i.i.d. 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 $A = M + LJ_{R}$, where $M$, $L$ and $R$ are arbitrary deterministic matrices and $J$ is a random matrix of zero-mean independent and identically distributed elements. $M$ can be nonnormal, 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 $A$ nonnormal, the eigenvalues do not suffice to specify the dynamics induced by $A$, so we also provide general formulae for the transient evolution of the magnitude 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 formulae 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 $\Omega$ where there are nonzero singular values of $L^{-1}(zI - M)R^{-1}$ (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 $\sigma$ and the $\sigma$-pseudospectrum of $M$.

PACS numbers: 87.18.Sn, 87.19.L-, 02.10.Yn, 87.19.lj, 89.75.-k

I. 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 nonnormal matrices.

In neuroscience, for example, connections between neurons typically have restricted spatial range and show specificity with respect to neuronal type 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, 5]: 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 nonnormal by some measures, much more so than a fully random matrix [6].

In general, neural networks with a recurrent connectivity pattern described by a nonnormal matrix can be described as having a hidden feedforward connectivity structure between orthogonal activity patterns, each of which can also excite or inhibit itself [6–8]. In particular, such networks arising from the natural separation of excitatory and inhibitory neurons yield so-called “balanced amplification” of patterns of activity without any dynamical slowing [9]. Underlying this is the phenomenon of “transient amplification”: a small perturbation from a fixed point of a stable system with nonnormal connectivity can lead to a large transient response over finite time [9]. Networks that yield long hidden feedforward chains can also generate long time scales and provide a substrate for working memory [1, 8]. Systems with nonnormal connectivity can also exhibit pseudo-resonance 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 [9].

*Corresponding author.
ya2005@columbia.edu
As our knowledge of connectivity is at best probabilistic, to describe realistic biological behavior, we must generalize from the nonnormal behavior of a fixed, regular connectivity to the expected behavior of a typical sample from an appropriate connectivity ensemble. 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 in between the two extremes of fully random or fully deterministic. 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 nonnormal, 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. Also note that when $L$ and $R$ are not diagonal, the elements of matrix $A$ are not statistically independent. 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. [10, 11]), 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. It also generalizes a result [12] obtained for the case where $L$ and $R$ are scalar multiples of 1, the $N$-dimensional identity matrix (the same result was obtained in [13] using the methods and language of free probability theory; the eigenvalue density for the case $L \propto R \propto 1$ and a normal $M$ was also calculated in Ref. [10]). 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 nonnormal cases of $M$; the naive interpretation of the formulae fails in these cases, which were not previously discussed. Furthermore, with the aim of studying dynamical signatures of nonnormal connectivity, we focused on the dynamics directly, deriving general formulae for the magnitude of the response of the system to a delta 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 formulae for some specific examples of $M$, $L$ and $R$. Sections [III] and [IV] contain the detailed derivations of our general formulae for the eigenvalue density and the response magnitude formulae, respectively. Section [V] contains the detailed analytical calculations of these quantities for the specific examples presented in Sec. [II] based on the general formulae. We conclude the paper in Sec. [VI].

**II. SUMMARY OF RESULTS**

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

$$A = M + LJR,$$

(2.1)

where $M$, $L$ and $R$ are general ($M$) or arbitrary invertible ($L$ and $R$) deterministic matrices, and $J$ is a random matrix of independent and identically distributed (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 non-identically distributed elements, due to the possible mixing and non-uniform scaling of the rows (columns) of the iid $J$ by $L$ ($R$).

We are firstly 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

$$\frac{dx(t)}{dt} = -\gamma x(t) + Ax(t) + I(t),$$

(2.2)

for an $N$-dimensional state vector $x(t)$, when $A$ is a sample of the ensemble Eq. (2.1). Here, $\gamma$ is a scalar and $I(t)$ is an external, time-dependent input. In studying this system, we generally assume that Eq. (2.2) is asymptotically stable. This means that $M$, $L$ and $R$ and $\gamma$ must be chosen such that for any typical realization of $J$, no eigenvalue of $-\gamma 1 + M + LJR$ 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 non-crossing approximation, which is valid for large $N$, we have derived general formulae for several useful properties of such matrices involving their eigenvalues and eigenvectors (see Sec. [III] [V] for the details of the derivations and the definition of the non-crossing approximation). We present these results in this section. In our derivation of these results, we assume the distribution of the elements of $J$ is complex Gaussian. However, we emphasize that universality theorems ensure that, for fixed $M$, $L$ and $R$, and for large $N$, the obtained result for the eigenvalue density 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. [13], following earlier work on the universality of the circular law established and successively strengthened in Refs. [15–18]. 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. [10]). 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

$$\|B\|_F^2 \equiv \frac{1}{N} \sum_{ij} |B_{ij}|^2 = \frac{1}{N} \text{Tr}(BB^\dagger)$$

(2.3)

(equivalently, $\|B\|_F$ is the root mean square of the singular values of $B$). For general matrices, $A$ and $B$,

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

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 $1$. For a complex variable $z = x + iy$, the Dirac delta function is defined by $\delta^2(z) \equiv \delta(x)\delta(y)$, and we define $\partial_z \equiv \partial/\partial x + i\partial/\partial y)/2$, and $\partial_{\bar{z}} \equiv \partial/\partial x - i\partial/\partial y)/2$. 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 $O(f(N))$ (resp. $\Theta(f(N))$) when, for large enough $N$, the absolute value of that quantity is bounded above (resp. 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 $\|M\|_F$, $\|L\|_F$, $\|R\|_F$, $\|LM^{-1}R^{-1}\|_F$, and $\|LR^{-1}\|$ are bounded as $N \to \infty$. We use the bound on $\|LR^{-1}\|$ in Appendices A and B. The Frobenius norm conditions are assumptions in the universality theorem of Ref. [14] which we use as discussed above. Finally, we assume that for all $z \in \mathbb{C}$, the distribution of the eigenvalues of $M_zM_1^\dagger$, 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 $M_zM_1^\dagger$ 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

$$\rho_J(z) = \frac{1}{N} \sum_{s} \delta^2(z - \lambda_s),$$

(2.4)

where $\lambda_s$ are the eigenvalues of $M + LJR$. It is known [13] 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 $\rho(z) \equiv \langle \rho_J(z) \rangle_z$ 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 $N$, $\rho(z)$ is nonzero in the region of complex plane satisfying

$$\text{tr}[(M_zM_1^\dagger)^{-1}] \geq 1$$

(2.5)

where we defined

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

(2.6)

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

$$\left\| R \frac{1}{z - M} L \right\|_F \geq 1.$$  

(2.7)

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

$$\rho(z) = \frac{1}{\pi} \frac{\partial}{\partial z} \text{tr} \left[ \left( RL^{-1}M_1^\dagger \right)^{-1} \frac{M_zM_1^\dagger + g(z)^2}{M_zM_1^\dagger + g(z)^2} \right],$$  

(2.8)

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

$$\text{tr} \left[ \frac{1}{M_zM_1^\dagger + g^2} \right] = 1,$$  

(2.9)

for $g$ for each $z$. As a first example, for the well-known case of $M = 0$, $L = 1$, and $R = \sigma 1$, we have $M_z = z/\sigma$ and the circular law follows immediately from Eq. (2.5), which yields $\sigma^2/|z|^2 \geq 1$ or $|z| \leq \sigma$ for the support, and from Eqs. (2.8)–(2.9) which yield the uniform $\rho(z) = 1/(\pi \sigma^2)$ within that support. As we noted in the introduction, formulae (2.5)–(2.9) generalize the results of Refs. [12] and [13] for the special case $L \propto R \propto 1$ to arbitrary invertible $L$ and $R$ (the eigenvalue density for the case $L \propto R \propto 1$ and a normal $M$ was also calculated in Ref. [10]).

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 $M_zM_1^\dagger$, we can evaluate the trace in Eq. (2.9) in the eigen-basis of the latter matrix, and rewrite this equation as

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

(2.10)
Similarly, Eq. (2.5) can be equivalently rewritten as

\[
\frac{1}{N} \sum_{i=1}^{N} s_i(z)^{-2} \geq 1. \tag{2.11}
\]

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 \( z \). We have

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

A note of caution is called for in using the formulas (2.5)–(2.9) (or equivalently Eqs. (2.10)–(2.12)). We have written these formulas for finite \( N \) (assuming it is large). However, the non-crossing 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 \( g^2 \to 0^+ \lim N \to \infty \) 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

\[ K(g, z) \equiv \lim_{N \to \infty} \frac{1}{M_z M_z^T + g^2}, \tag{2.13} \]

for fixed, strictly positive \( g \), which serves to regularize the denominators in Eq. (2.13) 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

\[ \lim_{g \to 0^+} K(g, z) \geq 1. \tag{2.14} \]

For most examples of \( M \), the two limits commute and can be taken in either order. However, the two limits may fail to commute if one or more of the smallest \( s_i(z) \) are nonzero but vanish as \( N \to \infty \). This can occur for some choices of \( M, L \) and \( R \), and in particular when \( L \propto R \propto 1 \) can occur for certain highly nonnormal \( M \) [38]. In such cases, the naive use of Eq. (2.5) yields a region larger than and containing the support of \( \lim_{N \to \infty} \rho(z) \) given by Eq. (2.14). In cases we have examined this signifies that there exists a finite, non-vanishing 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 will discuss examples of this phenomenon in Sec. III C below; in one of the examples (discussed in Sec. III C2), the existence of such outlier eigenvalues was first noted in Ref. [29], and their distribution was more quantitatively characterized in Ref. [21]. However, the connection between such outlier eigenvalues and vanishingly small singular values of \( M_{z} \), which arise, e.g., for highly nonnormal \( M \), were not noted before to the best of our knowledge. We have observed in simulations (and also supported by [21]) that the distribution of these outliers is typically not self-averaging as \( N \to \infty \), 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 its behavior may not be correctly given by the non-crossing approximation.

Finally, we note that 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

\[ r_0 = \|RL\|_F = \left[ \frac{1}{N} \sum_{i=1}^{N} \sigma_i^2 \right]^{1/2}, \tag{2.15} \]

where \( \sigma_i \) 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

\[ \rho(r) = -\frac{1}{2\pi r} \partial_r (g(r)^2), \tag{2.16} \]

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

\[ r^2 = \frac{1}{N} \sum_{i=1}^{N} \frac{\sigma_i^{-2}}{1 + g(r)^2}. \tag{2.17} \]

Integrating Eq. (2.16), we see that the proportion of eigenvalues lying a distance larger than \( r \) from the origin is, in this case, given by

\[ n_>(r) = \begin{cases} \frac{g(r)^2}{2\pi r} & (r < r_0) \\ 0 & (r \geq r_0). \end{cases} \tag{2.18} \]

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 nonnormal matrices (see Ref. [6] 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 \(\sigma\), or its \(\sigma\)-pseudospectrum in the given norm, is the set of points \(z\) for which \(|(z - M)^{-1}|| \geq \sigma^{-1}\) (thus as \(\sigma \rightarrow 0\) 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 \(\sigma\)-pseudospectrum can equivalently be characterized as the set of points, \(z\), for which there exists a matrix perturbation \(\Delta M\), with \(|\Delta M| \leq \sigma\), such that \(z\) is in the eigenvalue spectrum of \(M + \Delta M\) [9].

In the operator norm, \(\sigma\)-pseudospectrum of \(M\) is the set to which its spectrum can be perturbed by adding to it arbitrary perturbations of size \(\sigma\) 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 \(\sigma\). Our result Eq. \((2.7)\), in this case reads \(|(z - M^{-1})|_F \geq 1\) or \(|(z - M)^{-1}|_F \geq \sigma^{-1}\). In other words, as \(N \rightarrow \infty\), the spectrum of \(M + \sigma J\), for an iid random \(J\) with zero mean and variance \(1/N\), is the \(\sigma\)-pseudospectrum of \(M\) in the normalized Frobenius norm defined by Eq. \((2.3)\). Interestingly, the perturbation, \(\Delta M = \sigma J\), has normalized Frobenius norm \(\sigma\) as \(N \rightarrow \infty\): this norm is \(\sigma \sqrt{\sum_{ij} J^2_{ij}/N}\), which, by the law of large numbers, converges to \(\sigma\) for large \(N\). That is, as \(N \rightarrow \infty\), in the normalized Frobenious norm, the spectrum in response to the random perturbation \(\sigma J\), which has size \(\sigma\), is the \(\sigma\)-pseudospectrum of \(M\).

This result sounds similar to the equivalence for the operator norm of the two definitions of pseudospectra 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 equivalence stated above does not in general hold for the normalized Frobenius norm. Second, for the operator norm, it is not in general the case that the \(\sigma\)-pseudospectrum of \(M\) is equivalent to the spectrum obtained from random perturbations of \(M\) of size \(\sigma\), even in the limit \(N \rightarrow \infty\) (although the spectra arising from such random perturbations are sometimes used as a “poor man’s version” or approximation of the pseudospectra [9]). This can be seen as follows. The operator norm of the random iid perturbation, \(\sigma J\), i.e. its maximum singular value, converges almost surely to \(2\sigma\) as \(N \rightarrow \infty\) \((2.2)\). Condition \((2.7)\) for \(z\) to be in the spectrum under this random perturbation is \(|(z - M^{-1})|_F \geq \sigma^{-1}\), or \(\text{rms}\left\{s_i(z)^{-1}\right\} \geq 1\) where the \(s_i(z)\) are the singular values of \(z - M\) and \(\text{rms}\{x_i\}\) represents the root-mean-square of the set of values \(\{x_i\}\). This is not equivalent to the condition that \(z\) be in the \(2\sigma\)-pseudospectrum of \(M\) in the operator norm, i.e. that \(|(z - M)^{-1}| \geq (2\sigma)^{-1}\) or \(s_{\text{min}}(z)^{-1} \geq \frac{1}{\sigma}\), where \(s_{\text{min}}(z)\) is the minimum of the \(s_i(z)\); in fact, noting that \(s_{\text{min}}(z)^{-1} \geq \text{rms}\{s_i(z)^{-1}\}\), 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 \(2\sigma\)-pseudospectrum in the operator norm. For example, for \(M = 0\), the “poor man’s \(2\sigma\)-pseudospectrum” in the limit \(N \rightarrow \infty\) is a ball of radius \(\sigma\) about the origin (the circular law), while the true \(2\sigma\)-pseudospectra of the zero matrix is the ball of radius \(2\sigma\) about the origin.

In sum, in the operator norm, the \(\sigma\)-pseudospectrum of \(M\) is equivalent to the set of points \(z\) for which there exists some perturbation \(\Delta M\) with \(|\Delta M| \leq \sigma\) such that \(z\) is in the spectrum of \(M + \Delta M\) [9]. In the normalized Frobenius norm in the limit \(N \rightarrow \infty\), when the limits \(\sigma^2 \rightarrow 0^+\) and \(N \rightarrow \infty\) commute, the \(\sigma\)-pseudospectrum of \(M\) is equivalent to the set of points \(z\) that are in the spectrum of \(M + \Delta M\) where \(\Delta M\) is any random perturbation with zero-mean iid elements with \(|\Delta M|_F = \sigma\); when the two limits do not commute, the support of \(M + \Delta M\) is a subset of the \(\sigma\)-pseudospectrum of \(M\).

### 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 \(I(t) = 0\), 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 \(|x(t)|_2 = x(t)^*x(t)|\), can nevertheless exhibit transient, yet possibly large and long-lasting growth, before it eventually decays to zero. By contrast, in stable normal systems, \(|x(t)|_2^2\) 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 non-orthogonality of its eigenvectors, or alternatively by its hidden feedforward structure (see Eq. \((2.32)\) for the latter’s definition).

Non-normal systems can also exhibit pseudo-resonances at frequencies that could be very different from their natural frequencies as determined by their eigenvalues; such pseudo-resonances will be manifested in the frequency power spectrum of the response of the system to time dependent inputs. \(|x(t)|^2\) 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 \(|\langle |x(t)|^2| \rangle\), i.e. the ensemble average of the norm squared of the state vector, \(x(t)\), as it evolves under Eq. \((2.2)\) with \(I(t) = 0\), 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 $\gamma$ must be chosen such that for any typical realization of $J$, all eigenvalues of $-\gamma I + M + LJR$ have negative real part. In particular, the entire support of the eigenvalue density of $M + LJR$, as determined by Eq. (2.24), must fall to the left of the vertical line of $\gamma$'s with real part $\tau$; 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, consider the time evolution of the squared norm, $\|\!\| x(t) \|\!\|^2$, of the response of the system to an impulse input, $I(t) = x_0 \delta(t)$, at $t = 0$, before which we assume the system was in its stable fixed point $x = 0$ (for $t > 0$ this is equivalent to the squared norm of the activity as it evolves according to Eq. (2.2) with $I(t) = 0$, starting from the initial condition $x(0) = x_0$). We provide a formula for the ensemble average of the more general quadratic function, $x(t)^T B x(t)$, where $B$ is any $N \times N$ symmetric matrix; the norm squared corresponds to $B = 1$. The result for general $B$, $M$, $L$ and $R$ is given as a double inverse Fourier transform

$$
\langle x(t)^T B x(t) \rangle = \iint \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} e^{i(\omega_1 - \omega_2)t} \text{Tr}[B \mathcal{C}^x(\omega_1, \omega_2; x_0 x_0^T)]
$$

in terms of the $N \times N$ Fourier-domain “covariance matrix,” $\mathcal{C}^x(\omega_1, \omega_2; x_0 x_0^T) \equiv \langle \hat{x}(\omega_1) \hat{x}(\omega_2) \rangle_B$ (where $\hat{x}(\omega)$ is the Fourier transform of $x(t)$). The expression for the latter is given by

$$
\mathcal{C}^x(\omega_1, \omega_2; C^4) = \mathcal{C}_0^x(\omega_1, \omega_2; C^4) + \Delta \mathcal{C}^x(\omega_1, \omega_2; C^4),
$$

(2.20)

where

$$
\mathcal{C}_0^x(\omega_1, \omega_2; C^4) \equiv \frac{1}{\gamma + i\omega_1 - M} C^4 \left( \frac{1}{\gamma - i\omega_2 - M} \right) \equiv \frac{1}{\gamma + i\omega_1 - M} C^4 \left( \frac{1}{\gamma - i\omega_2 - M} \right),
$$

(2.21)

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

$$
\Delta \mathcal{C}^x(\omega_1, \omega_2; C^4) \equiv \frac{1}{\gamma + i\omega_1 - M} \frac{1}{\gamma - i\omega_2 - M} \left( \text{tr} \left( R^4 R \frac{1}{\gamma + i\omega_1 - M} C^4 \left( \frac{1}{\gamma - i\omega_2 - M} \right) \right) - \delta_{\omega_1, \omega_2} \delta_{\omega_1, \omega_2} \right),
$$

(2.22)

is the contribution of the random part of connectivity $LJR$. For later use, we have provided these expressions for a general third argument in $\mathcal{C}^x(\cdot, \cdot, \cdot)$; for use in Eq. (2.19) $C^4$ must be substituted with $x_0 x_0^T$. In the special case of $\|x(t)\|^2$, corresponding to $B = 1$, and iid disorder ($L = 1$, $R = \sigma I$), the contributions from Eqs. (2.21)–(2.22) can be more compactly combined into

$$
\langle \|x(t)\|^2 \rangle_j = \iint \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} e^{i(\omega_1 - \omega_2)t} \frac{x_0^T \frac{1}{\gamma - i\omega_2 - M} \frac{1}{\gamma + i\omega_1 - M} x_0}{1 - \sigma^2 \text{tr} \left( \frac{1}{\gamma + i\omega_1 - M} \frac{1}{\gamma - i\omega_2 - M} \right)}
$$

(2.23)

(we used $x_0^T \frac{1}{\gamma - i\omega_2 - M} \frac{1}{\gamma + i\omega_1 - M} x_0$ to write the numerator in Eq. (2.23)).

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

$$
\mathcal{I}_I(t_1) \mathcal{I}_I(t_2) = \delta(t_1 - t_2) C^4_{ij},
$$

(2.24)

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,

$$
C^x_{ij}(\omega) \equiv \langle x_i(t + \tau) x_j(t) \rangle \equiv \int d\tau e^{-i\omega \tau} \mathcal{I}_I(t + \tau) \mathcal{I}_I(t),
$$

(2.25)

given by

$$
\langle C^x(\omega) \rangle_J \equiv C^x_0(\omega) + \Delta C^x(\omega),
$$

(2.26)

Here we defined

$$
C^x_0(\omega) \equiv C^x_0(\omega, \omega; C^4)
$$

(2.27)

and

$$
\Delta C^x(\omega) \equiv \Delta C^x(\omega, \omega; C^4)
$$

(2.28)

are 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 $I(t) = I_0 \sqrt{2} \cos \omega t$ (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 $\|I_0\|^2$). For such an input, the steady-state activity, which we denote by $x_\omega(t)$, 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, $\|x_\omega(t)\|^2$, where now the bar indicates temporal averaging (we call this total power, because the squared norm sums the power in all components of $x_\omega(t)$). As in Eqs. (2.19)–(2.22) we present a formula for the ensemble average of the more general quantity $x_\omega^T B x_\omega$. We have

$$
\langle x_\omega^T B x_\omega \rangle_J = \text{Tr}(B \langle C^x(\omega) \rangle_J),
$$

(2.29)

where $\langle C^x(\omega) \rangle_J$ is given by Eqs. (2.26)–(2.28) with $C^4$ replaced by $I_0 I_0^T$. For the special case of $B = 1$, corresponding to the total power of the response at frequency $\omega$, using Eq. (2.21)–(2.22) with $\omega_1 = \omega_2 = \omega$, this formula can be simplified into

$$
\langle \|x_\omega\|^2 \rangle_J = \frac{1}{\|z - M\|_F^2} \left( \|z - M\|^2 \right) \frac{1}{\|z - M\|_F^2} \left( \|z - M\|^2 \right),
$$

(2.30)
where $z = \gamma + i\omega$, $\| \cdot \|_F$ denotes the Frobenius norm defined in Eq. (2.3). Finally, for the case that the random part of the matrix is iid, i.e., $L = \sigma I$ and $R = I$, we can further simplify Eq. (2.30) into

$$
\langle \|x_i\|^2 \rangle = \frac{\|(\gamma + i\omega - M)^{-1}I_0\|^2_2}{1 - \sigma^2 \|(\gamma + i\omega - M)^{-1}\|^2_2}. \tag{2.31}
$$

The stability of the $x = 0$ fixed point guarantees the positivity of the expressions Eqs. (2.30) and (2.31) 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 $\text{Re}(z) = \gamma$. By our result Eq. (2.27) for that support, this can only be true if the denominators of the last terms in Eq. (2.30) are positive, which guarantees the positivity of the full expressions. Also note how due to these denominators, the power spectrum will be larger for frequencies where the support of the eigenvalue density is closer to the vertical line with $\text{Re}(z) = \gamma$.

Numerical simulations indicate that the quantities $\|x(t)\|^2$ and $\|x_{\infty}\|^2$ are self-averaging in the large $N$ limit; that is, for large $N$, $\|x(t)\|^2$ or $\|x_{\infty}\|^2$ for any typical random realization of $J$ will be very close to their ensemble averages, given by Eq. (2.23) and Eq. (2.30) respectively, with the random deviations from these averages approaching zero as $N$ goes to infinity (see Figs. 1 and 3 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) for random fluctuations of these quantities for different realizations of $J$.

Finally, we note that the general formulae presented in this section are valid only for cases where the initial condition, $x_0$, or the input structure, $I_0$ or $C^I$, are chosen independently of the particular realization of the random matrix $J$ (e.g., cases where $x_0$ is itself random but independent of $J$, or when $x_0$ 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 formulae of this section.

C. Some specific examples of $M$, for $R, L \propto 1$

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. (2.23) and (2.23), and the total power in response to sinusoidal input Eq. (2.31), for specific examples of $M$, with $L$ and $R$ proportional to the identity matrix; without loss of generality we take $L = 1$ and $R = \sigma I$. Thus in these examples the full matrix is of the form $M + \sigma J$ where $\sigma$ determines the strength of disorder in the matrix. The details of the calculations for the results presented here can be found in Sec. V.

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

$$
M = UTU^\dagger, \tag{2.32}
$$

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 nonnormal and normal matrices is that for the latter, $T$ can be taken to be strictly diagonal. Equation (2.32) is referred to as a Schur decomposition of $M$ [23], 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 $I(t)$ set to zero, in the Schur basis by defining $y = UTx$ (i.e., $y_i$ is the activity in the $i$-th Schur mode), we obtain $\frac{dy}{dt} = -\gamma y + Ty$. We see that activity in the $j$-th Schur mode provides an input to the equation for the $i$-th mode only when $i \leq j$ (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 a matrix is more nonnormal the stronger its feedforward weights are.

Due to the invariance of the trace, the norm, and the adjoint operation under unitary transforms, our general formulae for the spectral density Eq. (2.8) and the average squared norm in time and frequency space, Eqs. (2.23) and (2.31), take the same form in the Schur space, i.e. $M$ can be replaced in them by $T$, and $x_0$ or $I_0$ by $U^\dagger x_0$ or $U^\dagger I_0$, respectively [40]. Thus we use the feedforward structure of the Schur decomposition to characterize the different examples we consider below. Our examples are picked to demonstrate interesting features of nonnormal 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

$$
M = T = \begin{pmatrix}
\lambda_1 & w & 0 & \cdots \\
0 & \lambda_2 & w & \cdots \\
\vdots & \vdots & \ddots & \ddots 
\end{pmatrix}, \tag{2.33}
$$

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. (2.33) with alternating imaginary eigenvalues, $\lambda_n = (-1)^{n+1} i$. The black dots in the bottom panel of Fig. 1 show the
FIG. 1: Top panel: the total power spectrum of steady state response $|x_\omega|^2$ as a function of input frequency $\omega$ Eq. (2.33), for the system Eq. (2.2) with $A = M + \sigma J$, and $M$ given by Eq. (2.33) 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 $x$ (the beginning of the feedforward chain characterized by Eq. (2.33)), which for the matrix $M$ has natural frequency $-1$. That is, the input was $I_0 \sqrt{2} \cos \omega t$, where $I_0$ was 1 for the last component and 0 for all other components. The green curve is the ensemble average of the total power spectrum, $\langle |x_\omega|^2 \rangle$, calculated numerically using the general formula Eq. (2.31), which is compared with an empirical average over 100 realizations of real Gaussian $J$ (red line, mostly covered by the green line). The pink area shows the standard deviation among these 100 realizations around this average. The blue line shows the result when disorder, $\sigma J$, is ignored, i.e. $A$ is replaced by its ensemble average $M$. Bottom panel: the eigenvalue spectrum of $M + \sigma J$ (black dots). Red dots show the eigenvalues of $M$ at $\pm i$. 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).

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, $\langle |x_\omega|^2 \rangle$, of the system Eq. (2.2) to sinusoidal stimuli as given by our general formula Eq. (2.31) (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 $|x_\omega|^2$ for each realization, which is given by the numerator of Eq. (2.31) with $M$ replaced by $M + \sigma J$, and then averaging the results over the 100 realizations). The pink shading shows

FIG. 2: The eigenvalue spectra of $A = M + \sigma J$ for $N = 2000$ and $M$ given by Eq. (2.33) with $\lambda_n = 0$, $w = 1$ for single realizations of real Gaussian $J$. $\sigma = 0.95$ and $0.5$ in the left and right panels, respectively. The red circles mark the circular boundaries of the spectral support given by Eq. (2.34). The insets show a comparison of the analytic formula Eq. (2.35) for the spectral density (black trace) and histograms corresponding to the particular realization shown in the main plot (red trace).

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 $\gamma$) 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 (2.33) with all eigenvalues zero we have analytically calculated the eigenvalue density, Eq. (2.8), the magnitude of response to impulse Eq. (2.23), and the power-spectrum Eq. (2.31). In this case, using Eq. (2.5) naively yields $|z| \leq \sqrt{|w|^2 + \sigma^2}$ for the support of the eigenvalue density. However, using the correct procedure, Eqs. (2.13)–(2.14), we find that this formula is only correct for $\sigma \geq |w|$, while for $\sigma < |w|$, the true support of the eigenvalue density in the limit $N \to \infty$ is the annulus

$$\sqrt{|w|^2 - \sigma^2} \leq |z| \leq \sqrt{|w|^2 + \sigma^2}, \quad (2.34)$$

(this result was obtained in Ref. [12]). Within this sup-
port the eigenvalue density in either case is
\[
\rho(z) = \frac{1}{\pi \sigma^2} \left[ 1 - \frac{|w|^2}{\sqrt{4|w|^2 |z|^2 + \sigma^2}} \right]. \tag{2.35}
\]

Figure 2 demonstrates the close agreement of Eqs. (2.34)–(2.35) with the empirical spectrum of \( M + \sigma J \) for a single realization of \( J \), for \( N = 2000 \) and two different values of \( \sigma \). The discrepancy between the results obtained by the naive use of Eq. (2.3) and Eq. (2.34) is due to the fact that for \( |z| < |w| \), \( M = 0 \)/\( \sigma \) has an exponentially small, \( O(e^{-cN}) \), singular value (see next paragraph) which makes the result of Eqs. (2.13)–(2.14) 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. (2.34)–(2.35), for very small but nonzero \( \sigma \) 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. [9], Ch. 7). The notion of pseudospectrum quantifies this sensitivity: the (operator norm) \( \epsilon \)-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 \( \epsilon \). As we mentioned in Sec. 1[A], this is precisely the set of complex values \( z \) for which \( \| (z - M)^{-1} \| > 1 - 1/e \) [9], 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: \( s_{\min}(z - M) \leq |w| |z|^{-N} \) (for a proof see after Eq. (5.15) in Sec. 5[A]). Thus the \( \epsilon \)-pseudospectrum of \( M \) contains the set of points \( z \) satisfying \( |w| |z|^{-1/N} < \epsilon \), i.e. the centered disk with radius \( |w| (\epsilon / |w|)^{1/N} \) which approaches \( |w| \) as \( N \to \infty \). In other words, for large enough \( N \), any point \( |z| < |w| \) is in the \( \epsilon \)-pseudospectrum for any fixed \( \epsilon \), no matter how small. It has been stated [9] that dense random perturbations, of the form \( \sigma J \) considered here, tend to trace out the entire \( \epsilon \)-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 \( \epsilon \)-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 \( x(t) \), Eq. (2.22), and the total power spectrum of steady-state response, Eq. (2.31), 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: \( x_0 = (0, \ldots, 0, 1)^T \) (or \( I_0 \propto (0, \ldots, 0, 1)^T \)). For the evolution of the average norm squared, with the initial condition \( x_0 = (0, \ldots, 0, 1)^T \), we obtain
\[
\langle \| x(t) \|^2 \rangle_{J} = e^{-2\gamma t} I_0 (2t \sqrt{|w|^2 + \sigma^2}), \quad (t > 0) \tag{2.36}
\]
where \( I_\nu(x) \) is the \( \nu \)-th modified Bessel function. Figure 3 plots the function Eq. (2.36) 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 \( x_0 = (0, \ldots, 0, 1)^T \), the solution for the \((N-n)\)-th Schur component is \( x_{N-n}(t) = (\text{grad}(\epsilon/|w|)^n) e^{-\gamma t} \) (for \( 0 \leq n \leq N-1 \)), which is maximized at \( t = n/\gamma \).
with a value \( \max |x_{N-n}| \sim \left( \frac{|w|}{\sqrt{1+\gamma}} \right)^n \) for \( n \gg 1 \). Thus up to time \( t \sim N/\gamma \) the norm of the activity grows exponentially; \( \|x(t)\|^2 \gtrsim \left( \frac{|w|}{\sqrt{1+\gamma}} \right)^{2t} \) 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 \( \|x(t)\|^2 \sim e^{-2t} \) 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 \( \sigma \), the 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 the initial condition along the last Schur mode. Let us note, however, that as we mentioned after Eq. (2.31), Eq. (2.23) and hence Eq. (2.36) 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 \( \sigma = 0 \) system.

Incidentally, we can also read the result for \( M = 0 \) from Eq. (2.36), by setting \( w = 0 \), obtaining \( \langle \|x(t)\|^2 \rangle_J = e^{-2t} I_0(2\sigma t) \). 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 \( I_0 = (0, \ldots, 0, I_0)^{\top} \) is given by

\[
\langle \|x\|^2 \rangle_J = \frac{I_0^2}{\omega^2 + \gamma^2 - |w|^2 - \sigma^2}.
\]

The main effect of the disorder is to reduce the width of the resonance (the peak of \( \langle \|x\|^2 \rangle_J \) 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 \([1, 5]\) in neurobiology. Dale’s law is the observation (which holds generally but with some exceptions \([24]\)) 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.

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

\[
T = \begin{pmatrix}
0 & w_1 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & \cdots \\
0 & 0 & w_2 & 0 & \cdots \\
0 & 0 & 0 & 0 & \cdots \\
& \cdots & & & \cdots \\
\end{pmatrix} = W \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}
\]

where we defined \( W \) to be the \( N/2 \times N/2 \) diagonal matrix of Schur weights \( W = \text{diag}(w_1, w_2, \ldots, w_{N/2}) \). In Eq. (2.38) arises as the Schur form of a mean matrix of the form

\[
M = \frac{1}{2} \begin{pmatrix} K & -K \\ K & -K \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \otimes K
\]

where \( K \) is a normal (but otherwise arbitrary) \( N/2 \times N/2 \) matrix (note that \( M \) is nonetheless nonnormal). The feedforward weights in Eq. (2.38) then are the eigenvalue of \( K \). When \( K \) has only positive entries, matrices of the form Eq. (2.39) satisfy Dale’s principle, and were studied in Ref. [6], 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 \( E = (e_1, e_2, \ldots) \) is the matrix of the orthonormal eigenvectors \( e_b \) of \( K \), \( b = 1, \ldots, N/2 \) (with eigenvalues \( w_b \)). Then transforming to the basis \( \{ (e_1, 0), (0, e_1), (e_2, 0), (0, e_2), \ldots, (e_{N/2}, 0), (0, e_{N/2}) \} \) (where \( 0 \) represents the \( N/2 \)-dimensional vector of \( 0 \)'s) transforms the matrix to being block-diagonal with the \( 2 \times 2 \) matrices \( \frac{1}{2} \begin{pmatrix} w_b & -w_b \\ w_b & -w_b \end{pmatrix}, b = 1, \ldots, N/2, \) along the diagonal. The \( b^{th} \) block becomes \( \begin{pmatrix} 0 & w_b \\ 0 & 0 \end{pmatrix} \) in its Schur basis \( \{ \frac{1}{\sqrt{2}} (e_b, e_b), \frac{1}{\sqrt{2}} (e_b, -e_b) \} \), so the full matrix takes the form Eq. (2.38). Thus, the \( b \)-th difference mode \( \frac{1}{\sqrt{2}} (e_b, e_b) \) feeds forward to the \( b \)-th sum mode \( \frac{1}{\sqrt{2}} (e_b, -e_b) \) with weight \( w_b \). This feedforward structure leads to a specific form of nonnormal transient amplification, which the authors of Ref. [6] 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. [20]. Thus, all excitatory neurons, with no other structure assumed (uniform mean inhibitory and mean excitatory synaptic strengths \( \mu_E + \mu_I \)), cause possibly large transients in modes in which the excitatory and inhibitory activities are balanced.

\[ M = uv^T \]  
(2.40)

where \( u = N^{-1/2}(1, 1, 1)^T \) is a unit vector, and the vector \( v \) has components \( v_i = \mu_E \) or \( v_i = -\mu_I \) for \( i \leq fN \) and \( i > fN \), respectively (for \( f = 1/2 \) and \( \mu_E = \mu_I \), Eq. (2.40) is a special case of Eq. (2.39)). The single-rank matrix \( M \) has only one non-zero eigenvalue given by \( v \cdot u = \frac{1}{\sqrt{N}} \sum_i v_i \), with eigenvector \( u \). 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 nonnormal case as \( v \cdot u = 0 \). In this case all eigenvalues of \( M \) are equal to zero. Furthermore, since in this case \( u \) and \( v \) are orthogonal, we can readily read off the Schur decomposition of \( M \) from Eq. (2.40). The normalized Schur modes are given by \( u, vv/\|v\|, \) and \( N-2 \) other unit vectors spanning the subspace orthogonal to both \( u \) and \( v \). All feedforward Schur weights are zero, except for one very large weight, equal to \( \|v\| \propto \sqrt{N} \), which feeds from \( v/\|v\| \) to \( u \). Thus the Schur representation of \( M \) has the form Eq. (2.38) with \( w_1 = \|v\| \equiv \mu \sqrt{N} \) and \( w_{b \neq 1} = 0 \), where we defined

\[ \mu^2 = \text{tr}(M^\dagger M) = \|v\|^2/N = f^2 \mu^2 + (1 - f) \mu^2. \]  
(2.41)

Note that this is again a case of balanced amplification: differences between excitatory and inhibitory activity, represented by \( v \), feed forward to balanced excitatory and inhibitory activity, represented by \( u \), with a very large weight. In the following we present results only for this balanced case of Eq. (2.40), which as just noted is a special case of Eqs. (2.38).

We start by presenting the results for the eigenvalue density. For general diagonal \( W \) in Eq. (2.38) (or equivalently, for general normal \( K \) in Eq. (2.39)), the eigenvalue density \( \rho(z) \) 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 \( w_i \) are \( O(1) \), the radius of this disk can be found directly from Eq. (2.5), which yields

\[ r_0 = \sigma \left[ \frac{1}{2} + \frac{1}{4} + \frac{\|w_b\|^2}{2\sigma^2} \right]^{1/2}. \]  
(2.42)

Here, \( \langle |w_b|^2 \rangle_b \) is the average of the squared feedforward weights over all blocks of Eq. (2.38), equivalently, in terms of the representation in Eq. (2.39), we have \( \langle |w_b|^2 \rangle_b = 2 \text{tr}(M^\dagger M) \equiv 2\mu^2 \). As long as some \( w_b \) are nonzero, \( r_0 \) is larger than the radius of the circular law, \( \sigma \), 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 \( \sigma J \) is larger the more nonnormal \( M \) is. In cases in which the feedforward weights of some of the \( 2 \times 2 \) blocks of Eq. (2.38) grow without bound as \( N \to \infty \), there is a corresponding singular value of \( M \propto z^{-1} - M \) for every such block which is nonzero for \( z \neq 0 \) but vanishes in the limit, scaling like \( \sim \frac{|z|}{|w_b|} \) where \( w_b \) is the unbounded weight of that block (see Eq. (2.37) and its preceding paragraph). (Note that as stated after Eq. (2.3) we assume \( \|M\|_F = \mu = \sqrt{\langle |w_b|^2 \rangle_b} / \sqrt{2} \) 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. (2.13) (2.14), which in this case can yield a support radius strictly smaller than Eq. (2.42). We have

FIG. 4: The eigenvalue spectra of \( A = M + \sigma J \) for a binary \( J \) with \( \sigma = 0.1 \) and \( M \) given by Eq. (2.39) with \( K = 1 \). (corresponding to \( w_b = 1 \) for all the diagonal \( 2 \times 2 \) blocks in Eq. (2.38)). The main panels show the eigenvalues for two particular realizations of \( J \), with \( N = 600 \) (left) and \( N = 60 \) (right). The red curves mark the circular boundaries of the spectral support as given by Eq. (2.42). Since the matrix \( A \) 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. (2.43) (black trace) and the empirical result (based on the eigenvalues of the realizations shown in the main panels) for the proportion, \( n_e(r) \), of eigenvalues lying further than a radius \( r \) from the origin. The insets show that the random fluctuations in the empirical value of \( n_e(r) \) as well as its average bias are already quite small for \( N = 60 \), and negligible for \( N = 600 \).
calculated the explicit results for $\lim_{N \to \infty} \rho(z)$ for two specific examples of $M$ with the Schur form Eq. (2.38). The first example belongs to the first case (bounded $w_b$’s) where $\lim_{N \to \infty} \rho(z)$ is $O(1)$ within the entire disk $r \leq r_0$, 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. (2.38) to have the same value, which we denote by $w$. In this case, the eigenvalue density is given by $\rho(r) = \frac{1}{2\pi} \frac{\partial n_{\sigma}(r)}{\partial r}$, where $n_{\sigma}(r)$ is the proportion of eigenvalues within a distance $r$ from the origin and is given by

$$n_{\sigma}(r) = \frac{r^2}{|w|^2} \left[ 1 - \frac{|w|^2}{\sigma^2 + \sqrt{\sigma^2 + |w|^4} |w|^2 r^2} \right]. \tag{2.43}$$

$n_{\sigma}(r)$ reaches unity exactly at $r = r_0$ given by Eq. (2.42), and $\rho(r)$ is $O(1)$ for any smaller $r$. Figure 1 shows the close agreement of Eq. (2.43) 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. (2.40) with $u \cdot v = 0$. As we saw, all $w_b$ are zero in this case except for one very large, unbounded weight $w_1 = \mu \sqrt{N}$. As discussed above, in this case $z - M$ has an $o(1)$ smallest singular value, approximately given by $\frac{|z|^2}{\mu \sqrt{N}}$.

Using Eqs. (2.13)–(2.14), we find that the support of $\lim_{N \to \infty} \rho(z)$ is the disc with radius $\sigma$ (within the an-
nulls \( \sigma < |z| \leq r_0 \) the eigenvalue density is \( o(1) \), and solving Eqs. \((2.8)\)–\((2.9)\) for \( |z| \leq \sigma \), we find that the spectral density is in fact identical with the circular law (the eigenvalue density for the \( M = 0 \) case), i.e.

\[
\rho(r) = \begin{cases} 
\frac{1}{\pi\sigma} + o(1), & (r < \sigma) \\
o(1) & (r > \sigma).
\end{cases}
\]

It was shown in Refs. \cite{21,25} 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. \cite{26} that the eigenvalue density of \( M + J + R \) agrees with that of \( J R \) as \( N \to \infty \). In the present example, it was observed in Ref. \cite{20} 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. 3). As we mentioned in Sec. II A, the non-crossing 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. \((2.8)\)–\((2.9)\), keeping track of finite-size, \( o(1) \) contributions, we obtain results that agree surprisingly well (though not completely) with simulations. First, for the total number of outlier eigenvalues lying outside the circle \( |z| = \sigma \) we obtain

\[
N_+(\sigma) \equiv N n_+(\sigma) = \sqrt{N} + O(1) \tag{2.45}
\]

(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. \((2.45)\) with simulations. The vast majority of the outlier eigenvalues counted in Eq. \((2.45)\) 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. \((2.8)\)–\((2.9)\) 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. \((2.8)\)–\((2.9)\) vanishes at \( r_0 \) given by Eq. \((2.42)\), 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 \( J \) \cite{21}.

As for the dynamics, we have analytically calculated the magnitude of impulse response, Eq. \((2.23)\), as well as the power-spectrum of steady-state response Eq. \((2.31)\), for Eqs. \((2.38)\)–\((2.39)\) 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. \((2.38)\); we denote the index for this block by \( a \). For the average magnitude of impulse response we find

\[
\langle ||x(\tau)||^2 \rangle = \frac{1 + C_a}{2} I_0(2r_0 t) + \frac{1 - C_a}{2} J_0(2r_1 t) e^{-2\gamma t} \tag{2.46}
\]

where \( J_0(x) \) (\( I_0(x) \)) is the (modified) Bessel function, \( r_0 \) is given by Eq. \((2.42)\), we defined \( r_1^2 = r_0^2 - \sigma^2 \), and

\[
C_a = \frac{1 + 2|w_a|^2/\sigma^2}{\sqrt{1 + 2(|w_b|^2)/\sigma^2}} \tag{2.47}
\]

with \( \langle |w_b|^2 \rangle_b = 2 \text{tr}(M^\dagger M) \) denoting the average squared feedforward weight among all the blocks of Eq. \((2.38)\). In Fig. 8 we plot Eq. \((2.46)\) 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 \( 2 \times 2 \) linear system governed by the matrix \( \left( \begin{array}{cc} -\omega_a & 0 \\ 0 & -\gamma \end{array} \right) \), we obtain

\[
\langle ||x(t)||^2 \rangle = (1 + w_a^2 t^2) e^{-2\gamma t}. \tag{2.48}
\]

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 \( 2 \times 2 \) blocks induced by the disorder insure that these more slowly decaying modes will be activated. Indeed, for large \( t \), \( ||x(t)|| \) decays like \( e^{-\gamma t} \) when \( \sigma = 0 \), while in the \( \sigma > 0 \) case, based on Eq. \((2.42)\) it must decay like \( e^{-(\gamma - \sigma) t} \), 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. \((2.46)\) 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 nonlinear 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. \((2.30)\), is given by the explicit formula

\[
\langle ||x_{\omega}||^2 \rangle_J = \frac{\omega^2 + \gamma^2 + |w_a|^2}{\omega^2 + \gamma^2 + \mu^2} ||I_0||^2. \tag{2.48}
\]

where \( \mu^2 \equiv \text{tr}(M^\dagger M) = \langle |w_b|^2 \rangle_b / 2 \) and, as noted above, the direction of \( I_0 \) is that of the second Schur mode in block \( a \).
diagonal matrix whose diagonal elements are the positive.

Eq. (2.49) has a fixed point by expanding Eq. (2.49) to first order in $x$.

\[ v = \text{external input vector}; \]

$i.e.$ output function, which is imposed element-by-element on the vector argument, with $f(v)_i \equiv f(v_i)$ giving the output, $i.e.$ the firing rate, of neuron $i$; \( \Gamma'(t) \) is the external input vector; $T = \text{diag}(\tau_1, \tau_2, \cdots, \tau_N)$ 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, \( \Gamma' \), Eq. (2.49) has a fixed point $v_*$. Then, given a small perturbation in the input, \( \Gamma'(t) = \Gamma' + \delta \Gamma'(t) \), we can write $v(t) = v_* + x(t)$, and linearize the dynamics around the fixed point by expanding Eq. (2.49) to first order in $x(t)$ and $\delta \Gamma'(t)$. This yields the set of linear differential equations

\[ T \frac{dx(t)}{dt} = -x(t) + \Phi x(t) + \delta \Gamma'(t), \]  

where $\Phi = \text{diag}(f'(v_*))$. (2.50)

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. (2.50) by $T^{-1}$, we can convert Eq. (2.50) into the form Eq. (2.2) with $\gamma = 0$ and $A = M + LR$ with

\[ M = T^{-1}(-1 + \langle W \rangle \Phi) \]  

\[ L = T^{-1} \]  

\[ R = \sigma \Phi \]  

and input

\[ I(t) = T^{-1} \delta \Gamma'(t). \]  

Assuming this linear system, $i.e.$ the fixed point $v_*$, is stable, we can thus think of our results for $\|x(t)\|^2$ and $\|x_0\|^2$ as characterizing the temporal evolution and the spectral properties of the linear response of the nonlinear system Eq. (2.49) in its fixed point $v_*$ to perturbations.

The necessary and sufficient condition for the stability of the fixed point (without any change in the external input) is that all eigenvalues of $A$ have negative real parts.
Our formula for the boundary of the eigenvalue distribution, Eq. (2.5), when applied in this particular case, can then be used to map out the region in parameter space (parameters here mean the time constants in $W$, \textit{i.e.} $\sigma$ and the parameters of $(W)$) in which a particular fixed point is stable. Recently our general formula Eq. (2.5) was used in this way by colleagues \cite{27} to determine the phase diagram of a clustered network of neurons, in which intra-cluster connectivity is large, but inter-cluster connectivity is random and weak. Because of the strong intra-cluster connectivity, each cluster behaves as a unit with a single self-coupling $a$. Letting the random inter-cluster couplings between $N$ clusters have zero mean and variance $g^2/N$, their analysis starts from the equation

$$\frac{d\mathbf{v}(t)}{dt} = -\mathbf{v}(t) + a\tanh(\mathbf{v}(t)) + g\mathbf{J}\tanh(\mathbf{v}(t))$$

(2.56)

where $\mathbf{J}$ is an iid random matrix as above. Here, $\mathbf{v}$ is a vector whose $i$-th component is the mean voltage of \textit{cluster} $i$, while the nonlinear function $\tanh(\mathbf{v}(t))$ (with the hyperbolic tangent acting component-wise) represents the vector of mean firing rates of the clusters. The analysis of Ref. \cite{27} 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 non-zero random attractor fixed points $\mathbf{v}_*$; for smaller values of $a$, the dynamics is chaotic (chaos in the $a = 0$ case was established in Ref. \cite{28}). The form of these fixed points (the distribution of the elements of $\mathbf{v}_*$ as $N \to \infty$ for a given $(a, g)$) can be obtained using mean-field theory, and the linearization about $\mathbf{v}_*$ leads to an equation in the form of Eq. (2.2), with $A = M + JR$, where $M$ and $R$ are the diagonal matrices

$$M = \text{diag}(-1 + a \tanh(\mathbf{v}_*))$$

(2.57)

$$R = \text{diag}(g \tanh(\mathbf{v}_*))$$

(2.58)

Given this form, it can be shown that the fixed point $\mathbf{v}_*$ 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 $\mathbf{v}_*$ determines the statistics of the elements of $R^2M^{-2}$ 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 $\text{tr}(R^2M^{-2}) < 1$. In this way, the region of stability of the fixed points in the $(a, g)$ plane can be mapped (see Ref. \cite{27} for the results, and a complete discussion of the analysis outlined here). Figure Eq. (9) shows a numerical example of the eigenvalue distribution for $A$ for a given $(a, g)$ and the superimposed boundary calculated using Eq. (2.5).

### III. Derivation of the Formula for the Spectral Density

In this section we will derive the formulae Eqs. (2.5)–(2.8) for the average spectral density, $\rho(z)$, of random matrices of the form $A = M + JR$ where $M$, $L$ and $R$ are deterministic matrices, and $J$ is random with iid elements of zero mean and variance $1/N$. We will use the Hermitianized diagrammatic method developed in Refs. \cite{10,11} (and reviewed in Ref. \cite{29}), which we will recapitulate here for completeness. As mentioned in Sec. II, the spectral density is self-averaging for large $N$. Furthermore, as established in Ref. \cite{14}, 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

$$\langle J_{ab}J^*_{cd}\rangle = \frac{1}{N} \delta_{ac}\delta_{bd}.$$  

(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

$$d\mu(J) = e^{-NT(\text{Tr}(JJ^T))} \prod_{ab} d\text{Im}J_{ab} d\text{Re}J_{ab}.$$  

(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

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

(3.3)

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

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

(3.4)

Following Ref. \cite{10}, we will use the identity

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

(3.5)

where the first identity follows by noting that $4\partial_z \partial_{\bar{z}} = \nabla^2$, where $\nabla^2$ is the 2-D Laplacian, and recalling from electrostatics that the solution of Poisson’s equation for a point charge at origin, \textit{i.e.} $\nabla^2 \phi(z) = 4\pi \delta(z)$, in 2-D is given by the potential field $\phi(z) = \ln |z|^2$; the second identity follows from $\partial_z \ln |z|^2 = \partial_z (\ln z + \ln \bar{z}) = \frac{1}{z} + \frac{1}{\bar{z}}$. 

\[15\]
0. Using Eq. (3.5) we can write the empirical spectral density, defined in Eq. (2.4), as
\[
\rho_J(z) = \frac{1}{\pi} \frac{1}{N} \sum_{\alpha} \frac{1}{z - \lambda_\alpha} = \frac{1}{\pi} \frac{1}{z - A},
\]
(3.6)
Performing the ensemble average we obtain
\[
\rho(z) \equiv \langle \rho_J(z) \rangle = \frac{1}{\pi} \frac{1}{z - A} \operatorname{tr}[(RL)^{-1} \langle G(z); J \rangle],
\]
(3.7)
where we used Eq. (3.3), and the linearity and cyclicity of the trace. Thus, to calculate \(\rho(z)\), our task boils down to calculating \(\langle G(z); J \rangle\).

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 we are most interested in. For Hermitian matrices, this is no problem as the spectrum is confined to the real line, and therefore \(G(z); J\) and \(\langle G(z); J \rangle\) 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 nonnormal (and in particular non-Hermitian) \(A\), with eigenvalues that in general cover a two dimensional region in the complex plane. However, using a trick introduced by Ref. [10], 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”
\[
H(z) \equiv \begin{pmatrix} 0 & M_z - J \\ M_z^+ - J^+ & 0 \end{pmatrix},
\]
(3.8)
and the corresponding \(2N \times 2N\) Hermitian matrices
\[
G(\eta, z; J) \equiv (\eta - H(z))^{-1}
\]
(3.9)
\[
= -\begin{pmatrix} (M_z - J)^{-1} & (M_z - J)^{-1} \\ (M_z - J)^{-1} & (M_z - J)^{-1} \end{pmatrix}.
\]
For \(\eta \to i0\), we see that
\[
G(0, z; J) = \begin{pmatrix} 0 & (M_z - J)^{-1} \\ (M_z - J)^{-1} & 0 \end{pmatrix}.
\]
(3.10)
and thus from Eq. (3.3), for any realization of \(J\)
\[
G(z; J) = -\lim_{\eta \to i0} G^{21}(\eta, z; J).
\]
(3.11)
Here, we have used the notation
\[
G(\eta, z; J) = \begin{pmatrix} G^{11}(\eta, z; J) & G^{12}(\eta, z; J) \\ G^{21}(\eta, z; J) & G^{22}(\eta, z; J) \end{pmatrix},
\]
(3.12)
where \(G^{\alpha\beta}\) (with \(\alpha, \beta \in \{1, 2\}\)) are \(N \times N\) matrices, forming the four blocks of \(G\). We have written the limit in Eq. (3.11) as \(\eta \to i0\) to emphasize that until the end of our calculations \(\eta\) is to retain a nonzero imaginary part, which serves to regularize the denominators in Eq. (3.9); c.f. the discussion after Eq. (3.35). We will be carrying out a perturbation expansion in powers of \(J\), so we decompose the Hamiltonian according to
\[
H(z) = H_0(z) - J,
\]
(3.13)
\[
J \equiv \begin{pmatrix} 0 & J^+ \\ J^+ & 0 \end{pmatrix}, \quad H_0(z) \equiv \begin{pmatrix} 0 & M_z \\ M_z^+ & 0 \end{pmatrix}.
\]
(3.14)
We will 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 \(2 \times 2\) matrices
\[
\sigma^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \sigma^- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.
\]
(3.15)
By a slight abuse of notation we also denote \(2N \times 2N\) matrices \(\sigma^\pm \otimes I_{N \times N}\) by \(\sigma^\pm\), and we will denote the identity matrix in any space by \(I\). From Eqs. (3.11) we obtain
\[
\operatorname{tr}[(RL)^{-1}G(z; J)] = -\operatorname{tr}[(\sigma^+ (RL)^{-1}G(i0^+, z; J)],
\]
and from Eq. (3.7)
\[
\rho(z) = -\lim_{\eta \to i0} \frac{1}{\pi} \bar{\partial} \operatorname{tr}(\sigma^+ (RL)^{-1}G(\eta, z)),
\]
(3.16)
\[
= -\lim_{\eta \to i0} \frac{1}{\pi} \bar{\partial} \operatorname{tr}((RL)^{-1} G^{21}(\eta, z)),
\]
(3.17)
where we defined
\[
G(\eta, z) \equiv \langle G(\eta, z; J) \rangle.
\]
(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 \(G(\eta, z; J)\) (including \(G(\eta, z)\)). Note that, being the Green’s function of a Hermitian matrix, \(G(\eta, z; J)\) and hence \(\bar{G}(\eta, z; J)\) are analytic functions of \(\eta\) for \(\eta\) outside the real line, and therefore analytic continuation can be used to take the limit \(\eta \to i0\) after obtaining the average over \(\eta\) sufficiently away from the real line.

We will denote the elements of a generic \(2N \times 2N\) matrix \(A\) by \(A_{ab}^{\gamma\delta}\), 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
\[
\langle J_{ab}^{\gamma\delta} J_{cd}^{\gamma_3 \delta_3} \rangle_J = \frac{1}{N^2} \delta_{ad} \delta_{bc} (\sigma^+_{\alpha\beta} \sigma^-_{\gamma\delta} + \sigma^-_{\alpha\beta} \sigma^+_{\gamma\delta}),
\]
(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 parenthesis on the right side of Eq. (3.19) as
\[
\pi^1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \pi^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},
\]
(3.20)
where \(J\) in powers of \(J\).

Also, since \(J\) have zero mean, we have \(\langle J \rangle = 0\). The starting point of the diagrammatic method is the perturbation expansion of \(G(\eta, z; J) = (\eta - H_0(z) - J)^{-1}\) in powers of \(J\),

\[
G(\eta, z; J) = G(\eta, z; 0) + \sum_{n=0}^{\infty} [J G(\eta, z; 0)]^n
\]  

(3.22)

where \(G(\eta, z; 0)\) is given by Eq. (3.9) with the \(J\)’s set to zero. This equation is represented diagrammatically in the third line of Fig. 10 the thin arrows defined in the first line of the figure represents \(G(\eta, z; 0)\), and the dashed lines represent a power of \(J\) before ensemble averaging. To obtain the average resolvent, \(G(\eta, z)\), we then average Eq. (3.22) term by term, with respect to the ensemble Eq. (3.12). 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 \(\langle J \rangle = 0\), 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. 10 the first two lines of which define the diagram elements. For example, the last diagram in the fourth line of Fig. 10 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 \(G(\eta, z; 0)\) (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 \(\| (RL)\| \) remains bounded as \(N \to \infty\), only non-crossing 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, As the last two lines of Fig. 10 demonstrate, all non-crossing diagrams can be generated by iterating the equation

\[
G(\eta, z) = G(\eta, z; 0) + G(\eta, z; 0) \Sigma(\eta, z) G(\eta, z),
\]  

(3.23)

starting from \(G(0)(\eta, z) = G(\eta, z; 0)\). This equation is represented diagrammatically in the last line of Fig. 10 with the “self-energy” matrix, \(\Sigma(\eta, z)\), defined by the diagram in the second line of that figure, i.e.

\[
\Sigma(\eta, z) \equiv \langle \mathcal{J} G(\eta, z) \mathcal{J} \rangle.
\]  

(3.24)

Using Eq. (3.21) we obtain

\[
\Sigma_{ab}^{cd}(\eta, z) = \delta_{ad} \sum_{r=1}^{2} \pi_{r}^{ab} \frac{1}{\bar{N}} \operatorname{Tr}(\pi^{3-r} G(\eta, z)),
\]  

(3.25)

which using Eq. (3.20) we can write as

\[
\Sigma(\eta, z) = \begin{pmatrix}
-i g_2(\eta, z) & 0 \\
0 & -i g_1(\eta, z)
\end{pmatrix},
\]  

(3.26)

where we defined the scalar functions

\[
g_\alpha(\eta, z) \equiv i \operatorname{tr} G^{\alpha \alpha}(\eta, z).
\]  

(3.27)

Using Eq. (3.26) we can solve Eqs. (3.23)–(3.26) for \(G(\eta, z)\) 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 \(G^{-1}(\eta, z; 0)\) on the left, and by \(G^{-1}(\eta, z)\) on the right, to obtain

\[
G(\eta, z) = [G^{-1}(\eta, z; 0) - \Sigma(\eta, z)]^{-1}
= [\eta - H_0(z) - \Sigma(\eta, z)]^{-1}.
\]  

(3.28)

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

\[
G(\eta, z) = - \left( (\eta + ig_1) K_1^{-1} (\bar{z} - M) K_2^{-1} \right).
\]  

(3.29)
where $K_1 \equiv M_1 M_1^\dagger + (g_1 - i\eta)(g_2 - i\eta)$ and $K_2 \equiv M_2 M_2^\dagger + (g_1 - i\eta)(g_2 - i\eta)$, and we dropped the arguments of $g_0(\eta, z)$ for succinctness. Imposing Eq. (3.27) we obtain the self-consistency equations

$$
g_1 = (g_1 - i\eta) \text{tr}(K_1^{-1}),$$

$$
g_2 = (g_2 - i\eta) \text{tr}(K_2^{-1}).$$

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

$$M_z = U_z S_z V_z^\dagger,$$

where $S_z$ is a nonnegative diagonal matrix with the singular values of $M$, $s_i(z)$ ($i = 1, \cdots, N$), 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 $\text{tr}(K_1^{-1}) = \text{tr}(K_2^{-1}) = \text{tr}(S_z^{-2} + (g_1 - i\eta)(g_2 - i\eta))^{-1}$. 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

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

or as written in the original basis

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

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

$$G(\eta, z) = \mathcal{G}(\eta + ig(\eta, z), z; 0)$$

$$= -\left(\frac{\gamma}{M_1 M_1^\dagger + \gamma^2} \frac{\gamma}{M_2 M_2^\dagger + \gamma^2} \frac{\gamma}{M_z M_z^\dagger + \gamma^2}\right)$$

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 $\eta \to i\epsilon$. Note, however, that as shown in Appendix A the non-crossing approximation is in general guaranteed to work only for $\text{Im} \eta \neq 0$; hence the limit $\eta \to i\epsilon$ 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 $\epsilon$, 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, $g(\eta, z) = \text{tr} \mathcal{G}_1(\eta, z) = \text{tr} (M_z J_1)^\dagger = \gamma J_1$, or equivalently $g(\eta, z) = \text{tr} (M_z J_1 M_z J_1)^\dagger$, 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 $\epsilon$ (and hence $\epsilon + g$) positive and nonzero, we define

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

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

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

with $\gamma = g + \epsilon$. Since $\epsilon$ 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) $g(\gamma, \epsilon) \to 0^+$, in which case we must have

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

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

$$1 = \lim_{N \to \infty} \text{tr} \left[\frac{1}{S_z^{-2} + \epsilon + g(\gamma, \epsilon)^2}\right].$$

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

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

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 in Eq. (3.38), and naively set $\eta = i\epsilon$ and $\gamma$ to zero, to obtain

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

From Eqs. (3.10) and (3.11), this solution yields $\langle G(z; J) \rangle = -G^{21}(\eta = i0^+, z) = M_2^{-1} R(z - M_1)^{-1} L$ 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. (2.14) of Sec. II); here $g(z)$ is to be found by solving Eq. (3.39), and 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

\[ \rho(z) = \frac{1}{\pi} \partial_z E(z) \]  
(3.42)

\[ E(z) \equiv \text{tr} \left[ \frac{(RL)^{-1}M^\dagger_z}{M_z M^\dagger_z + g(z)^2} \right], \]  
(3.43)

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

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

\[ E(z) = \text{tr} \left[ \frac{\partial_z(M_z M^\dagger_z)}{M_z M^\dagger_z + g(z)^2} \right]. \]  
(3.44)

On the other hand, we have

\[ \partial_z \text{tr} [M_z M^\dagger_z + g(z)^2] = \left[ \frac{\partial_z(M_z M^\dagger_z + g(z)^2)}{M_z M^\dagger_z + g(z)^2} \right], \]

\[ = E(z) + \partial_z(g^2(z)), \]  
(3.45)

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

\[ E(z) = \partial_z \varphi(z), \]  
(3.46)

\[ \varphi(z) \equiv -g^2(z) + \text{tr} \ln [M_z M^\dagger_z + g(z)^2], \]  
(3.47)

or using the SVD, Eq. (3.32),

\[ \varphi(z) = -g^2(z) + \text{tr} \ln [s_z^2 + g(z)^2], \]  
(3.48)

\[ = -g(z)^2 + \frac{1}{N} \sum_{i=1}^{N} \ln [s_i(z)^2 + g(z)^2]. \]  
(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 \( \sigma_i \) 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. (2.17). 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

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

To find the spectral radius (boundary of the spectrum) \( r_0 \) we have to solve Eq. (2.17) for \( r \), setting \( g(r) = 0 \). This yields \( r_0^2 = \frac{1}{N} \sum_{i=1}^{N} \sigma_i^{-2} = ||RL||^2_{F} \), yielding Eq. (2.15).

Let us define the proportion of eigenvalues lying outside a radius \( r \) from the origin by \( n_>(r) \). To obtain Eqs. (2.16) and (2.18), first note that

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

where we used the expression of Laplacian, \( \nabla^2 = \partial_{rr} + \frac{1}{r} \partial_r \), in 2-D polar coordinates in the last equality. Using this with the definition \( n_>(r) = 2\pi \int_{r_0}^{\infty} \rho(r) r^2 dr \), we obtain

\[ n_>(r) = \left[ \frac{2}{r} \partial_r \varphi(r) \right]_{r_0}^{r}. \]  
(3.52)

For the limit at \( r \to \infty \), note that for \( r > r_0 \), \( g(r) = 0 \) and we have \( \varphi(r) = \frac{1}{N} \sum_{i=1}^{N} \ln (r^2\sigma_i^{-2}) = 2 \ln r - \frac{2}{r} \ln \det(RL) \), and hence \( \frac{2}{r} \partial_r \varphi(r) \to 1 \) as \( r \to \infty \). Thus we obtain

\[ n_>(r) = 1 - \frac{r}{2} \partial_r \varphi(r). \]  
(3.53)

Differentiating Eq. (3.50) and using Eq. (2.17) we obtain

\[ \partial_r \varphi(r) = 2r \frac{1}{N} \sum_{i=1}^{N} \frac{1}{r^2 + \sigma_i^2 g(r)^2}, \]  
(3.54)

and

\[ n_>(r) = 1 - r^2 \frac{1}{N} \sum_{i=1}^{N} \frac{1}{r^2 + \sigma_i^2 g(r)^2}. \]  
(3.55)

Using Eq. (2.17) once again we obtain Eq. (2.18). Finally, using the latter together with Eqs. (3.51)–(3.52) yields Eq. (2.16).

IV. 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 formula presented in Sec. 11B. We will first consider the system’s response to an impulse input, \( I(t) = x_0 \delta(t) \), at \( t = 0 \), before which we assume the system was at rest in its fixed point \( x = 0 \). We assume \( x = 0 \) is a stable fixed point, i.e. all eigenvalues of \(-\gamma + A\) have negative real parts, or equivalently, all eigenvalues of \( A \) have real parts less than \( \gamma \) (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 \( \gamma \) must be to the right of the support of \( \rho(z) \), the average eigenvalue density for \( A \), as found by solving Eq. (4.1)). This means that \( x(t) \) decays exponentially as \( t \to \infty \), and therefore its Fourier transform, \( \hat{x}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} x(t) dt = \int_{-\infty}^{\infty} e^{-i\omega t} \hat{x}(t) dt \), is well-defined.

Fourier transformation of Eq. (2.2) yields \( I(t) = x_0 \delta(t) \) yields \( i\omega \hat{x}(\omega) = (-\gamma + i\omega) \hat{x}(\omega) + x_0 \hat{x}(\omega) \). Solving algebraically for \( \hat{x}(\omega) \), we obtain \( \hat{x}(\omega) = (\gamma + i\omega - A)^{-1} x_0 \), or using Eqs. (3.3)–(3.4), \( \hat{x}(\omega) = R^{-1}(\gamma + i\omega + J) L^{-1} x_0 \). The inverse Fourier transform, \( x(t) = \int_{-\infty}^{\infty} e^{i\omega t} \hat{x}(\omega) \frac{d\omega}{2\pi} \), then yields

\[ x(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{i\omega t} R^{-1}(\gamma + i\omega + J)L^{-1} x_0. \]  
(4.1)
Our goal is to study the statistics of $x(t)$ (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 $R^{-1}G(z;J)L^{-1} = (z - A)^{-1} = V(z - A)^{-1}V^{-1}$ where $A$ is a diagonal matrix of the eigenvalues of $A$, and $V$ is the matrix whose columns are the eigenvectors of $A$. Here we will look at the simplest interesting statistic involving the eigenvectors: the average square norm of the state vector, namely, $⟨∥x(t)∥^2⟩$. As we discussed in Sec. 1B, its study is also motivated by the fact that transient amplification due to nonnormality of $A$ manifests itself in the transient growth of $∥x(t)∥^2 = x(t)^T x(t)$. With a slight generalization, we derive a formula for the average of a general quadratic function $x(t)^T B x(t)$ where $B$ is any symmetric matrix; the norm squared corresponds to $B = 1$. Using, $x(t)^T = x(t)$( $x(t)$ is real), the identity $x^T B x = Tr (B x^T x)$, and Eq. (4.1), we obtain

$$x(t)^T B x(t) = \int \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} e^{i(t_1 - t_2)\omega_2} C_{ij}(\gamma + i\omega_2; J),$$

(4.2)

where we defined $C_{ij} \equiv L^{-1} x_0^T x_1^T \pi L^{-1}$ and $R_{ij} \equiv R^{-1} B R^{-1}$. Using Eq. (3.11) and $G(z;J) = -\lim_{\eta \to 0^+} G^{ij}(\eta, z; J)$, and the $2 \times 2$ matrices $\pi$ defined in Eq. (3.20), we can rewrite the trace in Eq. (4.2) as $Tr(\pi^T x_0^T B x_0)$, where $\pi(z) \equiv \gamma + i\omega_1$, where now the trace is performed over $2N \times 2N$ matrices. Averaging over $J$ we then obtain

$$\langle x(t)^T B x(t) \rangle_J = \int \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} e^{i(t_1 - t_2)\omega_2} F(\gamma + i\omega_1, \gamma + i\omega_2; B, x_0^T x_0^T),$$

(4.3)

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

$$F(z_1, z_2; B, C) = \langle Tr(B G(0, z_1; J) C G(0, z_2; J)) \rangle_{ij}. $$

(4.4)

with

$$B \equiv \pi^T \otimes B_{\pi}, \quad B_{\pi} \equiv R^{-1} B R^{-1},$$

$$C \equiv \pi^T \otimes C_{\pi}, \quad C_{\pi} \equiv L^{-1} C L^{-1}. $$

(4.5)

Before proceeding to the calculation of $F(z_1, z_2; B, C)$ using the diagrammatic technique, we will also express the other quantities presented in Sec. 1B in terms of $F(\gamma + i\omega_1, \gamma + i\omega_2; B, C)$, with appropriate $B$'s and $C$'s. First, we obtain the desired expression for the matrix power spectrum, Eq. (2.25) of the steady-state response to a temporally white noisy input, $I(t)$, with covariance Eq. (2.24). 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 $x(t)$ as in Eq. (4.1) with $x_0$ replaced by the Fourier transform of the input, $I(\omega)$. Using this and exploiting $x_j(t_2) = x^\dagger_j(t_2)$ we can write (after averaging over the input noise)

$$\overline{x_i(t_1)x_j(t_2)} = \int \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} e^{i(t_1 - t_2)\omega_2} K_{ij}(\omega_1, \omega_2)$$

(4.7)

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

$$K(\omega_1, \omega_2) \equiv R^{-1} G(\gamma + i\omega_1; J)L^{-1} C_{\pi} C^T \overline{G(\gamma + i\omega_2; J)} R^{-1}.$$  

(4.8)

Here, the bars indicate averaging over the input noise distribution, and we defined $C_{\pi}(\omega_1, \omega_2) \equiv \overline{I(\omega_1)I(\omega_2)^T}$. On the other hand, the Fourier transform of Eq. (2.24) yields

$$\overline{x_i(t_1)x_j(t_2)} = \int \frac{d\omega_1}{2\pi} \int \frac{d\omega_2}{2\pi} e^{i(t_1 - t_2)\omega_2} C^\pi_{ij}(\omega),$$

(4.10)

where

$$C^\pi(\omega) =$$

(4.11)

$$R^{-1} G(\gamma + i\omega_1; J)L^{-1} C^T \overline{G(\gamma + i\omega_2; J)} R^{-1}.$$  

(4.13)

Noting that Eq. (4.10) expresses the covariance of the response as an inverse Fourier transform, we see that $C^\pi(\omega)$ is indeed the power spectrum of the response, as defined in Eq. (2.25). Finally note that the element, $C_{ij}$, of any matrix can be expressed as $Tr(e_i^T e_j^C)$, where $e_i$ are the unit basis vectors (i.e., vectors whose $a$-th component is $\delta_{ia}$). Using this trick with Eq. (4.11), and following the steps leading from Eq. (4.12) to Eq. (4.13), we see that after ensemble averaging, $\langle C_{ij}^\pi(\omega) \rangle_{ij}$ can be written in the form

$$\langle C_{ij}^\pi(\omega) \rangle_{ij} = F(\gamma + i\omega_1, \gamma + i\omega_2; e_i^T e_j^C, C^T)$$

(4.12)

where $F$ was defined by Eqs. (4.4)–(4.6).

Next, consider the system Eq. (2.2) being driven by a sinusoidal input $I(t) = I_0 \sqrt{2} \cos \omega t$ (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 $\omega$. Decomposing the input, $I(t)$, and the steady-state response, $x_0(t)$, 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

$$x_0(t) = \sqrt{2} R^{-1} Re[e^{i\omega t} G(\gamma + i\omega_1; J)L^{-1} I_0].$$

(4.13)

Thus the norm squared of the steady state response, $∥x(t)∥^2 = x(t)^T x(t)$, will have a zero frequency component, plus components oscillating at $±2\omega$. Averaging
over time kills the latter, leaving the zero frequency component intact, yielding

$$\bar{x}_\omega(t) \bar{x}_\omega(t) = \mathbf{I}_0 \mathbf{L}^{-1} \mathbf{G}(z; J) \mathbf{R}^{-1} \mathbf{G}(z; J) \mathbf{L}^{-1} \mathbf{I}_0$$

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

$$\langle \bar{x}_\omega(t)^\dagger \mathbf{B} \bar{x}_\omega(t) \rangle_j = \mathcal{F}(\gamma + i \omega, \gamma + i \omega; B, \mathbf{I}_0 \mathbf{I}_0)$$

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

$$\langle \bar{x}_\omega(t)^\dagger \mathbf{B} \bar{x}_\omega(t) \rangle_j = \text{Tr}(B \langle C^x(\omega) \rangle_j)$$

which is Eq. (2.29) of Sec. III it being understood that $C^x$ in Eq. (4.12) is replaced by $\mathbf{I}_0 \mathbf{I}_0$ 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

$$F_{\mu_1 \nu_2 ; \mu_2 \nu_1} (1; 2) \equiv \langle \mathbf{G}_{\mu_1 \nu_1}(1; J) \mathbf{G}_{\mu_2 \nu_2}(2; J) \rangle_j$$

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 $\mu_i = (\alpha_i, \alpha_i)$ (similarly for $\nu_i$) 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

$$\mathcal{F}(z_1, z_2; B, C) = B_{\mu_2 \mu_1} F_{\mu_1 \nu_2 ; \mu_2 \nu_1} (0, z_1; 0, z_2) C_{\nu_1 \mu_2},$$

where all indices are summed over, and $B$ and $C$ were defined in Eqs. (4.5)–(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. 11 for the contribution of $m$-th and $n$-th terms in the expansion of the first and the second Green’s function, respectively. As before, for large $N$, averaging over $J$ entails summing the contribution of all non-crossing pairings. This is indicated in the second line of Fig. 11. Finally, the third line of Fig. 11 shows that summing over all $m$’s, and $n$’s and all non-crossing pairings, is equivalent to replacing all solid lines with thick solid lines representing the average Green’s function, $\mathbf{G}(\eta, z)$ (calculated according to Eqs. (3.26)–(3.29), and summing over all NCP’s with $\mathbf{l}$ 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.
The series expansion of \( \sigma \) and \( \rho \) are summed over. It shows the sum of all ladder diagram contributing to Eq. (4.17) (i.e. the last line of Fig. 11) in terms of \( \mathcal{D} \), which is defined in the second row. The first term on the right side of the first row equation is the disconnected average Eq. (4.20), which 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 \( \mathcal{D} \).

where \( \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \) is the first Pauli matrix. This helps us because in the expansion of Fig. 12, the two factors in Eq. (3.21) involving \( \pi^r \) and \( \pi^s \) 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. 13 for the second term in the series expansion of \( \mathcal{D} \) shown in the second line of Fig. 12. Extending this similarly to all the terms in that expansion, we obtain

\[
\mathcal{D}_{\mu \rho \lambda \nu}(1; 2) = \mathcal{D}_{adcb}^{\alpha \lambda \gamma \delta}(1; 2)
\]

\[
= \frac{1}{\mathcal{N}} \sum_{r,s=1}^{2} \left( \pi_{\alpha \beta} \delta_{ad} \right) D_{rs}(1; 2) \left( \pi_{\nu \delta} \delta_{cb} \right), \tag{4.23}
\]

where \( \mu = (\alpha, \alpha), \nu = (\beta, \beta), \lambda = (\gamma, \gamma), \rho = (\delta, \delta) \) and we defined the 2 \( \times \) 2 matrices

\[
D(1; 2) \equiv \sigma^1 + \sigma^1 \Pi^0 \sigma^1 + \cdots = \sigma^1 \sum_{n=0}^{\infty} (\Pi^0 \sigma^1)^n \tag{4.24}
\]

and the “polarization matrix” for the diffuson

\[
\Pi^0_{rs}(1; 2) \equiv \text{tr}(\pi^r \mathcal{G}(1) \pi^s \mathcal{G}(2)) = \text{tr}(\mathcal{G}^{\pi^r}(1) \mathcal{G}^{\pi^s}(2)). \tag{4.25}
\]

Here, as before, with the trace performed over the 2\( N \)-dimensional space, and we used Eq. (4.20) to write the last term of \( \Pi^0 \). Summing the geometric series in Eq. (4.24) we obtain

\[
D(1; 2) = \sigma^1 \left( 1_{2 \times 2} - \Pi^0(1; 2) \sigma^1 \right)^{-1}. \tag{4.26}
\]

FIG. 12: The first row is the diagrammatic representation of Eqs. (4.19)–(4.21). In the last term, \( \rho \)'s and \( \lambda \)'s are summed over. It shows the sum of all ladder diagram contributing to Eq. (4.17) (i.e. the last line of Fig. 11) in terms of \( \mathcal{D} \), which is defined in the second row. The first term on the right side of the first row equation is the disconnected average Eq. (4.20), which 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 \( \mathcal{D} \).

\[\text{FIG. 13: The contribution to } \mathcal{D}^{\alpha \lambda \gamma \delta}(1; 2) \text{ from the second term in the series expansion shown in the second row of Fig. 12 in more detail.} \]

\[\text{The } 2 \times 2 \text{ matrix inversion yields}
\]

\[
D(1; 2) = \begin{pmatrix}
\Pi_{22} & 1 - \Pi_{22} \\
1 - \Pi_{11} & \Pi_{22}
\end{pmatrix}^{-1} = \frac{1}{1 - \Pi_{11} \Pi_{22} - \Pi_{11} \Pi_{22} \Pi_{22}},
\]

where all \( \Pi^0 \)'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):

\[
\mathcal{F}(z_1, z_2; B, C) = \mathcal{F}^0(z_1, z_2; B, C) + \Delta \mathcal{F}(z_1, z_2; B, C),
\]

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 \( F_{\mu_j \nu_j \mu_j \nu_j} \) on the right side replaced by \( F^0_{\mu_j \nu_j \mu_j \nu_j} \) and \( F_{\mu_j \nu_j \mu_j \nu_j} \), respectively. Using Eqs. (4.20)–(4.22) and (4.23), we then obtain

\[
\mathcal{F}^0(z_1, z_2; B, C) = \text{Tr} \left( B \mathcal{G}(0, z_1) C \mathcal{G}(0, z_2) \right) \tag{4.29}
\]

\[
= \text{Tr} \left( B \mathcal{G}^{12}(0, z_1) C \mathcal{G}(0, z_2) \right)
\]

and

\[
\Delta \mathcal{F}(z_1, z_2; B, C) = \frac{1}{\mathcal{N}} \sum_{r,s} \text{Tr} \left( B \mathcal{G}^{2r}(0, z_1) C \mathcal{G}^{2s}(0, z_2) \right) \times D_{rs}(0, z_1; 0, z_2) \mathcal{G}^{s1}(0, z_1) C_L \mathcal{G}^{1s}(0, z_2)), \tag{4.30}
\]

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 \( \omega_i \). 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 valid solution of Eq. (3.34) is the trivial \( g(0, z) = 0 \). In this case, we have Eq. (3.41), and for \( \eta \rightarrow i \omega \), from Eq. (3.10) (replacing \( A \) with \( M \), corresponding to \( J = 0 \)) we have

\[
\mathcal{G}(0, z_i) = -\begin{pmatrix}
0 & M_{z_i}^{-1} \\
M_{z_i} & 0
\end{pmatrix}. \tag{4.31}
\]
Using this in Eqs. (4.29)–(4.30) we obtain
\[ F^0(z_1, z_2; B, C) = \text{Tr}(B_n M^{-1}_{z_1} C_k M^{-1}_{z_2}) , \] (4.32)
and
\[ \Delta F(z_1, z_2; B, C) = \text{tr}(B_n G^{21}(0, z_1) G^{12}(0, z_2)) \times \]
\[ D_{12}(0, z_1, 0, z_2) \text{Tr}(G^{21}(0, z_1) C_k G^{12}(0, z_2)). \] (4.33)
Using the definitions Eq. (2.6) and Eqs. (4.5)–(4.6) we can simplify Eq. (4.32) to
\[ F^0(z_1, z_2; B, C) = \text{Tr}\left( B \frac{1}{z_1 - M} C \frac{1}{z_2 - M} \right) . \] (4.34)
From Eqs. (4.25) and (4.31) we see that (for \( z_i \) of interest and for \( \eta_i \) going to zero) \( \Pi^{rr} = 0 \) and \( \Pi^{12} = \Pi^{21} \), and from Eq. (4.27) we obtain
\[ D_{12}(0, z_1, 0, z_2) = \frac{1}{1 - \Pi^{21}_i (0, z_1, 0, z_2)} \]
\[ = \frac{1}{1 - \text{tr}(G^{21}(0, z_1) G^{12}(0, z_2))}. \] (4.35)
Substituting this in Eq. (4.33) and using Eq. (4.31) once again, we finally obtain
\[ \Delta F(z_1, z_2; B, C) = \frac{\text{tr}(B_n M^{-1}_{z_1} M^{-1}_{z_2}) \text{Tr}(M^{-1}_{z_1} C_k M^{-1}_{z_2})}{1 - \text{tr}(M^{-1}_{z_1} M^{-1}_{z_2})} \]
and after simplification using Eqs. (4.5)–(4.6),
\[ \Delta F(z_1, z_2; B, C) = \frac{\text{tr}\left( B \frac{1}{z_1 - M} L L \frac{1}{z_2 - M} \right) \text{Tr}\left( R \frac{1}{z_1 - M} C \frac{1}{z_2 - M} \right)}{1 - \text{tr}\left( R \frac{1}{z_1 - M} L L \frac{1}{z_2 - M} \right)}. \] (4.36)

The general formulae of Sec. II B readily follow. Equations (2.19)–(2.22), with \( C^4 \) replaced by \( x_0 x_0^\dagger \), for the case of response to an impulse input follow from Eqs. (4.3)–(4.4), and (4.6), respectively. Equations (2.26)–(2.28) (with \( C^0 \) and \( \Delta C^0 \) defined in Eqs. (2.21)–(2.22)) 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 Eq. (4.36), after setting \( B = e_i e_i^\dagger, C = C^4 \) and \( z_1 = z_2 = \gamma + \omega \) (with the traces involving \( B = e_i e_i^\dagger \) turned into matrices in Eqs. (2.26)–(2.28), using \( \text{Tr}\left( e_i e_i^\dagger X \right) = X_{ij} \)). The result Eq. (2.29) 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)
\[ \langle x(t)^\dagger B x(t) \rangle_j = \left[ x(t)^\dagger B x(t) \right]_{j=0} + \Delta f_B(t), \] (4.37)
where the two terms on the right hand side are obtained by replacing \( F(\cdot; B) \) in Eq. (4.3) with Eqs. (4.34) and Eq. (4.36), respectively. The integrals over \( \omega_1 \) and \( \omega_2 \) decouple for the first term yielding the expected result for \( J = 0 \),
\[ \left[ x(t)^\dagger B x(t) \right]_{j=0} = e^{-2\gamma t} \text{Tr}(B e^{\gamma T} x_0 x_0^\dagger e^{\gamma T}), \]
\[ = x_0^\dagger e^{\gamma T} B e^{\gamma T} x_0. \] (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).

V. CALCULATIONS FOR SPECIFIC EXAMPLES OF \( M \)

In this section we give the detailed calculations for the explicit expressions for the spectral density Eq. (2.8), the power spectrum Eq. (2.29), and the average squared norm Eqs. (2.19) and (2.23), for the specific examples of \( M \), presented in Sec. II C. For these examples both \( R \) and \( L \) are proportional to the identity matrix; without loss of generality we take \( L = 1 \) and \( R = 1 I \).

The eigenvalue density and the norm squared \( ||x||^2 \) 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 \( M \rightarrow U^{\dagger} M U \), we can assume \( M \) is already in its Schur form Eq. (2.38) without loss of generality. Furthermore, we will 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 \( t \rightarrow t, z \rightarrow z/\sigma, \gamma \rightarrow \gamma/\sigma, M \rightarrow M/\sigma \), and \( \rho \rightarrow \sigma^2 \rho \) (with the latter applying to both the eigenvalue density and the power spectral density), we obtain the result for general \( \sigma \).

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 C1 where \( M \) is
\[ M = T = \begin{pmatrix} 0 & w & 0 & \cdots \\ 0 & 0 & w & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \] (5.1)
or \( M_{ij} = w \delta_{i+1,j} \). First we calculate the eigenvalue density. According to Eqs. (2.8)–(2.9), in order to calculate the spectral density, we need to calculate first the inverse of \( M \), \( M^{-1} + g^2 = (z - M)(z - M)^\dagger + g^2 \) (remember that we have set \( \sigma = 1 \), as we explained in the beginning of the section). To this end, notice that \( K_{ij} \equiv [(z - M)(z - M)^\dagger]_{ij} = Q_{ij} - |w|^2 \delta_{iN} \delta_{jN} \)
\[ Q_{ij} \equiv (|z|^2 + |w|^2) \delta_{ij} - w \overline{z} \delta_{i+1,j} - \overline{w} z \delta_{i,j+1}. \] (5.2)
As the difference \((z - M)(z - M)^\dagger - Q = -|w|e_\gamma e_\gamma^\dagger\) is single rank, we can use the Woodbury formula for matrix inversion to write

\[
\frac{1}{K + g^2} = \frac{1}{Q + g^2} + \frac{1}{Q + g^2} e_N e_N^\dagger \frac{1}{Q + g^2} \left( \frac{\lambda_N}{(Q + g^2)^{-1}} e_N \right),
\]

(5.3)

where \(e_N^\dagger = (0, \ldots, 0, 1)\). (The only conditions for the validity of Eq. (5.3) is that the factor in parenthesis is not singular, i.e. \(e_N (Q + g^2)^{-1} e_N \neq |w|^{-2}; we consider the validity of this condition below.) Since \(Q = \text{Toeplitz and Hermitian}, it can be diagonalized easily. Using standard methods, we find that the eigenvalues and eigenvectors of \(Q\), satisfying \(Qv_n = \lambda_n v_n\), are given by

\[
\lambda_n = \left| z - |w|e^{i\phi_n} \right|^2, \quad \phi_n \equiv \frac{\pi n}{N + 1}
\]

(5.4)

\[
v_n^\dagger = \sqrt{\frac{2}{N + 1}} \left( \bar{w}z \right)^{j/2} \sin \phi_n j
\]

(5.5)

for \(n = 1, \ldots, N\). The eigenvectors are orthonormal \(v_n v_m = \delta_{nm}\), and we have the spectral representation

\[
\frac{1}{Q + g^2} = \sum_{n=1}^N v_n \lambda_n + g^2 v_n^\dagger.
\]

(5.6)

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

\[
\text{tr} \frac{1}{K + g^2} = I_1(g, z) + \frac{1}{N} \frac{-\partial I_2(g, z)}{\partial g},
\]

(5.7)

where we defined

\[
I_1(g, z) = \text{tr} \left( \frac{1}{Q + g^2} \right), \quad I_2(g, z) = e_N^\dagger \frac{1}{Q + g^2} e_N = \frac{1}{N + 1} \sum_{n=1}^N \frac{2\sin^2 \phi_n}{\lambda_n + g^2}.
\]

(5.8)

(5.9)

In writing the numerator of the last term in Eq. (5.7), we used Eqs. (5.5)–(5.6) to write \(\text{tr} \left( \frac{1}{Q + g^2} \right) = \| \frac{1}{Q + g^2} e_N \|_2^2 = \frac{1}{N + 1} \sum_{n=1}^N \frac{2\sin^2 \phi_n}{\lambda_n + g^2}\). In the \(N \to \infty\) limit, the sums in Eqs. (5.8)–(5.9) can be approximated by the integrals

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

(5.10)

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

(5.11)

Some elementary contour integral then yields

\[
I_1(g, z) = \left( \frac{|z|^2 + |w|^2 + g^2}{2|w|^2} - 4|w|^2 \frac{|z|^2}{|w|^2} \right)^{-1/2},
\]

(5.12)

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

(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 Eq. (2.13)–(2.14), 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. (2.13)), which now takes the form \(I_1(g, z) = 1\), yields

\[
g(z)^2 = -|z|^2 - |w|^2 + 4|w|^2 |z|^2 - 1.
\]

(5.14)

This is positive if and only if

\[
\sqrt{|w|^2 - 1} \leq |z| \leq \sqrt{|w|^2 + 1},
\]

(5.15)

which after the proper rescaling yields Eq. (2.34) for general \(\sigma\). Note that Eq. (5.15) is precisely the region given by Eq. (2.14), which in the present case reads \(I_1(0, z) \geq 1\). It is instructive to compare this result with what we would obtain by naively using Eq. (2.5), i.e. \(\text{tr} (K^{-1}) \geq 1\), 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 \(\text{tr} (Q^{-1}) = \text{tr} (K^{-1}) = I_1(0, z) + o(1)\), 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 \(u = (z_w)^{-1}\) satisfies \((z - M)u = w(z_w)^N e_N\), so that \(||z - M||u|| = ||z|| ||w||^N\), and since \(\text{min}(z - M) \leq ||z - M||u|| \) and \(||u|| \geq 1\), it follows that \(\text{min}(z - M) \leq ||u|| z_w^{-N}\) which is \(O(e^{-N})\) for \(|z| < |w|\). For large enough \(N\), this singular value \(alone\) suffices to make Eq. (5.11) (equivalent to Eq. (5.5)) hold for any \(|z| < |w|\), as \(N\text{min}(z^{-2})\) diverges despite its \(1\) prefactor.

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

\[
\text{tr} \left( \frac{z - M}{K + g(z)} \right)^2 = \frac{z - \text{tr} M^\dagger}{K + g(z)}
\]

(5.16)

where we used Eq. (2.9) to write the last expression. To obtain \(\text{tr} M^\dagger (Q + g^2)^{-1}\), we will again use Eq. (5.3). In the region Eq. (5.15), the contribution of the second term in Eq. (5.8) is again suppressed by \(1/N\) and from Eq. (5.1) and (5.6) we have \(\text{tr} M^\dagger (Q + g^2)^{-1} = \frac{1}{N} \sum_{n=1}^N \sum_{j=1}^{N-1} v_n^2 \bar{v}_n^{j+1} \lambda^{-1}_n + g^2\). A straightforward calculation using Eq. (5.5) (and the orthonormality of
that to obtain the total power spectrum for the input amplitude $I_0 = I_0(0, \ldots, 0, 1)^T$, we need to multiply Eq. (5.22) by $I_0^2 = \|I_0\|^2$, and substitute $z_1 = z_2 = \gamma + i\omega$. With the proper rescaling, this yields Eq. (2.37) for general $\sigma$. To obtain the formula for $\langle \|x(t)\|^2 \rangle_j$, we substitute Eq. (5.22) with $z_i = \gamma + i\omega_i$ for the integrand of Eq. (2.23). Changing the integration variables with $\omega_1 = \Omega + \omega/2$ and $\omega_2 = \Omega - \omega/2$, we obtain

$$\langle \|x(t)\|^2 \rangle_j = \frac{1}{2} \int \frac{d\omega}{2\pi} e^{-it\omega} \int \frac{d\Omega}{2\pi} \Omega^2 + (\gamma + i\omega/2)^2 - |\omega|^2 - 1.$$  

(5.23)

Finally consulting a table of Laplace transforms [40], we obtain

$$\langle \|x(t)\|^2 \rangle_j = e^{-2\gamma t} I_0(2t\sqrt{|\omega|^2 + 1}), \quad (t \geq 0)$$

(5.24)

where $I_0(x)$ is the 0-th modified Bessel function. Implementing the rescalings $t \rightarrow \sigma t$, $\gamma \rightarrow \gamma/\sigma$ and $\omega \rightarrow \omega/\sigma$, we obtain Eq. (2.36).

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

Here we carry out the explicit calculations for the example of Sec. [I C.2] where $M$ is given by Eq. (2.38) (without loss of generality, we assume $M$ has its Schur form), using formulae (2.8)–(2.9) for the spectral density and Eq. (2.23) for $\langle \|x(t)\|^2 \rangle_j$. First we calculate the eigenvalue density. From Eq. (2.38), $K \equiv M_M^t = (z - M)(z - M)^t$ (we are setting $L = R = 1$ in Eq. (2.6); see the comments at the beginning of this section) is a block-diagonal matrix with $2 \times 2$ diagonal blocks, with the $b$-th block ($b = 1, \ldots, N/2$) given by

$$K(z) = \begin{pmatrix} z & -w_b \\ -\bar{w}_b & \bar{z} \end{pmatrix} = \begin{pmatrix} |z|^2 + |w_b|^2 & -w_b \bar{z} \\ -\bar{w}_b z & |z|^2 + g^2 + |w_b|^2 \end{pmatrix}, \quad (5.25)$$

where $w_b$ is the corresponding Schur weight in Eq. (2.38). Likewise, $(K + g^2)^{-1}$ whose trace appears in Eqs. (2.9) is given by a block-diagonal matrix with diagonal blocks

$$\int \frac{d|z|^2 + g^2 + |w_b|^2}{|z|^2 + g^2 + |w_b|^2} = \begin{pmatrix} |z|^2 + g^2 + \frac{1}{2} |w_b|^2 \\ |z|^2 + g^2 + |w_b|^2 \end{pmatrix}.$$  

(5.26)

Taking the normalized trace we thus obtain

$$\langle (K + g^2)^{-1} \rangle \equiv \langle f(w_b) \rangle_b \equiv \frac{1}{N/2} \sum_{b=1}^{N/2} \langle f(w_b) \rangle,$$

(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 compile the support boundary of $\rho(z)$. As discussed in Sec. [I A], when (for $|z| \neq 0$) all singular values of $M_z = z - M$ are bounded from below as $N \rightarrow \infty$, the support is correctly given by Eq. (2.5) (we will 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

$$1 \leq \text{tr} K^{-1} = |z|^{-2} + \mu^2 |z|^{-4}, \quad (5.27)$$

for the integrand of Eq. (2.23) which we denoted by $F(z_1, z_2)$, with $z_1 = \gamma + i\omega_1$, ($i = 1, 2$). By comparing the integrand of Eq. (2.23) with Eq. (2.31), we see that to obtain the total power spectrum for the input...
where we defined $\mu^2 = \frac{1}{2} \langle |w_b|^2 \rangle_b = \text{tr}(M^M M)$. It follows that the support is the disk $|z| \leq r_0$, where

$$r_0^2 \equiv \frac{1}{2} + \sqrt{\frac{1}{4} + \mu^2}. \quad (5.28)$$

The replacements $\mu \to \mu/\sigma$ and $r_0 \to r_0/\sigma$ then yield Eq. (2.42).

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

$$\text{tr} \frac{1}{K + g^2} = \left< \frac{|z|^2 + g^2 + \frac{1}{2}|w_b|^2}{(|z|^2 + g^2)^2 + |w_b|^2 g^2} \right>_b = 1, \quad (5.29)$$

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

$$\pi \rho(z) = \frac{\partial}{\partial z} \text{tr}\left[M^M (K + g^2)^{-1}\right] = 1 - \frac{\partial}{\partial z} \text{tr}\left[M^M (K + g^2)^{-1}\right], \quad (5.30)$$

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 $M^M = \bar{z} - M$ and Eq. (5.29). From Eqs. (2.38) and (5.26) we see that $M^M (K + g^2)^{-1}$ has the same block-diagonal structure as Eq. (2.38), and a short calculation shows that $\text{tr}\left[M^M (K + g^2)^{-1}\right] = \bar{z} I_3(|z|)$, where we defined

$$I_3(r) \equiv \left< \frac{1}{2} |w_b|^2 \right>_{b} - \frac{1}{2} \frac{\partial}{\partial r} \text{tr}\left[r^2 - r^2 I_3(r)\right]. \quad (5.31)$$

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

$$0 < I_3(r) \leq 1. \quad (5.32)$$

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

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

for $r = |z| \leq r_0$, 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 $n_\lesssim(r)$, the proportion of eigenvalues with modulus smaller than some radius $r$. We have

$$n_\lesssim(r) = 2\pi \int_0^r \rho(r')r'dr', \quad \text{upon substitution of Eq. (5.33)},$$

yields

$$n_\lesssim(r) = r^2 \left(1 - I_3(r)\right) \quad (r \leq r_0). \quad (5.34)$$

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

$$\rho(r) = \frac{1}{2\pi r} \frac{\partial n_\lesssim(r)}{\partial r} = -\frac{1}{2\pi r} \frac{\partial n_\gtrsim(r)}{\partial r}, \quad (5.35)$$

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

$$n_\gtrsim(r) = g(r)^2 (1 + I_3(r)). \quad (5.36)$$

We see that beyond the radius $r$ at which $g^2$ vanishes (which when all $w_b$’s are bounded is $r = r_0$), $n_\gtrsim(r)$ and $\rho(r) = -\frac{1}{2\pi r} \frac{\partial n_\gtrsim(r)}{\partial r}$ 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 $K = M_z M_z^M$; thus they come in pairs composed of the eigenvalues of $K$’s $2 \times 2$ 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 value, 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

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

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

(Note that as stated after Eq. (2.3) we assume $\|M\|^2_{b,\infty} = \mu^2 = \langle |w_b|^2 \rangle_b/2$ 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 $r \leq r_0$, Eq. (5.29) yields a $\Theta(1)$ solution for $g^2(r)$, which leads to a $\Theta(1)$ solution for $n_\lesssim(r)$ and $\rho(r)$ via Eqs. (5.36)–(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 nonnormal 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. (2.13)–(2.14). 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 $\lim_{N \to \infty} n_\gtrsim(r)$ 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. (2.33), this case corresponds to $K = wI$). 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

\[ n_>(r) = r^2 - \frac{w^2 r^2}{1 + \sqrt{1 + w^4 + 4w^2 r^2}}. \]

The replacements \( w \rightarrow w/\sigma \) and \( r \rightarrow r/\sigma \) then yield Eq. (2.49) for general \( \sigma \), and \( \rho(r) \) can be calculated using Eq. (5.35).

The second case is that of Eq. (2.40). In this case only one of the blocks has a nonzero Schur weight given by \( |w_1|^2 = \text{Tr} (M^2 M) = N \mu^2 = O(N) \), where \( \mu = O(1) \) is given by Eq. (2.41). Equation (5.29) now yields

\[ 1 = \frac{1}{r^2 + g^2} + \frac{\mu^2}{(r^2 + g^2)^2} \frac{r^2 - g^2}{r^2 + g^2}, \]

or

\[ r^2 + g^2 - 1 = \frac{\mu^2}{(r^2 + g^2)^2} + N \mu^2 g^2. \]

The right hand side of this last equation is \( I_2(r) \), as follows from Eq. (5.31); thus using Eq. (5.42) we can rewrite Eq. (5.36) as

\[ n_>(r) = g^2(r) \frac{2r^2 - 1}{r^2 - g^2(r)}. \]

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 \( \lim_{N \rightarrow \infty} \rho(z) \) we have to first take the \( N \rightarrow \infty \) 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 \( r \leq 1 \) and vanishes at \( r = 1 \), the correct support radius of \( \lim_{N \rightarrow \infty} \rho(z) \), 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 \( N \rightarrow \infty \) limit of the eigenvalue density is identical with the circular law (the result for the \( M = 0 \), i.e. \( \lim_{N \rightarrow \infty} \rho(r) = \frac{1}{\pi} \) for \( r \leq 1 \) and zero otherwise. With the correct scaling, this yields Eq. (2.44).

Contrary to the general prescription given after Eq. (2.12), we will now solve equations Eqs. (5.29), (5.31) and Eq. (5.36) for \( r > 1 \), without taking the limit \( N \rightarrow \infty \) 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 \( 1 < r \leq r_0 \). 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 \( r_0(1) \), 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

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

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. 4 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 \( N \rightarrow \infty \)), the solution for \( g(r) \) is \( O(N^{-1}) \). Thus from Eq. (5.43) we thus see that \( N_{n<} (r) \), i.e. the total number of eigenvalues with modulus larger than \( r \), for \( 0 < r \leq r_0 \) and \( r_1 = \Theta(1) \) is only \( O(1) \); the solution for \( N_{n<} (r) \) 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 \( N \rightarrow \infty \), as already found. Now let us calculate the total number of eigenvalues lying outside the circle \( |z| = 1 \). This is given by \( N_{n>} (1) \). From Eq. (5.44) we find \( g^2(1) = \frac{1}{\sqrt{N}} \), and substituting in Eq. (5.43) we obtain

\[ N_{n>} (1) = N_{n<} (1) = \sqrt{N} + o(1). \]
where we used $\mu^2 = \langle|w_b|^2\rangle / 2$, and we denoted the integrand of Eq. (2.23) by $F(z_2)$ with $z_i = \gamma + i\omega$, $(i=1,2)$. By comparing the integrand of Eq. (2.23) with Eq. (2.31), we see that substituting $z_1 = z_2 = \gamma + i\omega$ into Eq. (5.48) yields the total power spectrum, $\langle|x_0|^2\rangle_J$. After the proper rescalings, this yields Eq. (2.48) for general $\sigma$. To obtain $\langle|\mathbf{x}(t)|^2\rangle_J$, on the other hand, we should substitute Eq. (5.48) into Eq. (2.19) with $z_1 = \gamma + i\omega$. Let us use the change of variables $\omega_1 = \Omega + \omega$ and $\omega_2 = \Omega - \omega$. Then we have $z_1 \omega_1 = \Omega^2 + (\gamma + i\omega)^2$, and from Eq. (2.19) we obtain

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

where we defined

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

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

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

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

$$r_1^2 = r_0^2 - 1 \geq 0. \tag{5.52}$$

One can calculate the integral over $\Omega$ 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 $\omega$. 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. (5.51) in Eq. (5.49) is essentially the inverse Laplace transform of Eq. (5.53). Consulting a table of Laplace transforms [30] yields

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

where $J_0(x)$ ($I_0(x)$) is the 0-th 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

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

where we defined

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

Effecting the proper rescalings we obtain the result for general $\sigma$, Eqs. (2.46)–(2.47).

**VI. CONCLUSIONS**

We have provided a general formula for the eigenvalue density of partly random matrices, i.e. matrices with general mean and non-trivial covariance structure. General formulae 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 nonnormal. 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 non-crossing 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 [31], 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 non-crossing 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 Sec. 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. [32]).

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 $A$ [33]–[34] (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 formulae; 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 $\langle \delta A_{ij} \delta A_{ij'} \rangle = (\mathbb{L}^2)^{ii'}(R^2)_{jj'}$ is single rank); the latter is of importance in considering networks with local topologies where, e.g., the matrix $A$ has a banded struc-
ture. Generalization to other forms of random disorder is thus an important direction for future research.

Acknowledgments

We thank Larry Abbott and Merav Stern for helpful discussions. Y.A. was supported by the Kavli Institute for Brain Science Postdoctoral Fellowship and by the Swartz Program in Computational Neuroscience at Columbia University, supported by a gift from the Swartz Foundation. K.D.M. was supported by grant R01-EY11001 from the NIH and by the Gatsby Charitable Foundation through the Gatsby Initiative in Brain Circuitry at Columbia University.

Appendix A: Validity of the non-crossing approximation

In this appendix we will give the justification for the non-crossing approximation used in Sec. II and IV. That is, we will show that the only diagrams not suppressed by inverse powers of $N$ are the non-crossing diagrams. We will limit our discussion to the case of the eigenvalue density considered in Sec. II, but the generalization to the quantities calculated in IV is straightforward. As explained after Eq. (3.22), averaging of $G(\eta, z; J)$ 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. 10. Each such diagram is composed of a solid directed line (each segment of which represents a factor of $G_{i \sigma}^{n_\sigma}(\eta, z; 0)$), 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 $\text{tr}(\sigma^+ (RL)^{-1} G(\eta, z; J))$: thus we can imagine the solid arrow making a loop by closing-in on itself sandwiching $\sigma^+ \otimes (RL)^{-1}$ (see Fig. 14).

Given the structure of the Kronecker deltas 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 deltas, one for each side of the wavy line, enforcing index identification at the corresponding ends on each side [22] and return somewhere else on the solid line, continuing as before until we reach the initial point (see Fig. 14). As we go around this orbit, for each solid line traversed we write down, from right to left, a $G(\eta, z; 0)$ and for each wavy line a $\pi r_i$ (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 deltas and $\pi r_i$’s in Eq. (3.21), but not for the factor $\frac{1}{\pi}$ and the sum over $r_i$’s; we will 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. 14), and thus contributes a trace of the form

$$I_{m,r} = \text{Tr}(G(\eta, z; 0)\pi^{r_1} \cdots G(\eta, z; 0)\pi^{r_m}),$$

(A1)

where $r$ is short-hand for $\{r_1, \ldots, r_m\}$, 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. 14) 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 the form

\[ E_{n,\tilde{r}} = \text{Tr}(\sigma^+(RL)^{-1}G(\eta, z; 0)\pi^\tau \cdots \pi^\tau G(\eta, z; 0)) \]  

(A2)

where \( n \) is the number wavy lines the orbit traverses and \( \tilde{r} \) is short for \( \{r_{j1}, \ldots, r_{jn}\} \), 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 will account for it below). For succinctness, in Eqs. (A1)–(A2) we suppressed the arguments \( (\eta, z) \) 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 non-crossing approximation is based on the claim that each trace contributed by a orbit (external or internal) as in Eqs. (A1)–(A2) is \( O(N) \), irrespective of \( \eta, z, m \) or \( r \). We will 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 \( N \); for completeness we will summarize this argument here. First note that, assuming the claim, any diagram will yield an expression that is \( O(N^\alpha) \) with

\[ \alpha = f - w - 1, \]  

(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. 14). 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

\[ \chi \equiv F - E + V, \]  

(A4)

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

\[ \alpha = \chi - 2. \]  

(A5)

Thus the contribution of a diagram is \( O(N^{\alpha}) \), with \( \alpha \) 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” \( \chi \) as defined above, can be drawn on (embedded in) a two-dimensional oriented surface with Euler characteristic \( \chi \), 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 \( \chi \), 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 decompaction, the plane (e.g. the diagram in panel (b) of Fig. 14 can be drawn in this manner on a torus). We thus see that

\[ \alpha = -2g, \]  

(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 non-crossing [13].

Let us now go back to justifying the claim that the traces contributed by the orbits as in Eqs. (A1)–(A2) are \( O(N) \). For this purpose we will make use of the singular value decomposition of \( M_z \) introduced in Eq. (3.32). Defining the unitary matrix

\[ U_z \equiv \left( \begin{array}{cc} U_z & 0 \\ 0 & V_z \end{array} \right), \]  

(A7)

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

\[ H_0(z) = U_z \tilde{H}_0(z) U_z^\dagger, \]  

(A8)

where

\[ \tilde{H}_0(z) \equiv \left( \begin{array}{cc} 0 & S_z \\ S_z & 0 \end{array} \right). \]  

(A9)

Let us also define \( \tilde{G}(\eta, z; 0) \equiv U_z \tilde{G}(\eta, z; 0) U_z \), such that

\[ G(\eta, z; 0) = U_z \tilde{G}(\eta, z; 0) U_z^\dagger. \]  

(A10)

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

\[ \tilde{G}(\eta, z; 0) = \frac{1}{\eta - \tilde{H}_0(z)} = \left( \frac{\eta^{-1} - S_z^{-1}}{\eta^{-1} - S_z^{-1}} S_z^{-1} \right), \]  

(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

\[ \pi^r = U_z \pi^r U_z^\dagger. \]  

(A12)
We now substitute $G(\eta, z; 0)$ and $\pi^r$ in Eqs. (A1)–(A2) with the right hand sides of Eqs. (A10) and (A12), respectively. After canceling the $U_k$’s, we obtain
\[
I_{m,r} = \text{Tr}\left( \tilde{G}(\eta, z; 0) \pi^{r_1} \cdots \tilde{G}(\eta, z; 0) \pi^{r_m} \right), \quad (A13)
\]
\[
E_{n,\tilde{r}} = \text{Tr}\left( \sigma^+ A(z) \tilde{G}(\eta, z; 0) \pi^{r_1} \cdots \pi^{r_m} \tilde{G}(\eta, z; 0) \right) \quad (A14)
\]
where we defined
\[
A(\eta, z; 0) \equiv U_z^i (RL)^{-1} V_z, \quad (A15)
\]
such that $U_k \left[ \sigma^+ \otimes (RL)^{-1} \right] U_k = \sigma^+ \otimes A(\eta, z; 0)$. For the internal orbits, we see from Eq. (A11) that each $\tilde{G}(\eta, z; 0)$, depending on whether it is sandwiched between the same projectors $\pi^r$, or between two opposite projectors, $\pi^r$ and $\pi^{r'}$, 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 $k$ ($1 \leq k \leq m$), we obtain
\[
I_{m,r}(\eta, z) = \sum_{i=1}^{N} \eta^{m-k} n_i(z) \frac{\eta^{k}}{\eta^2 - s_i(z)^2} m, \quad (1 \leq k \leq m)
\]
(A16)
for the internal orbits (in particular, we see that the sole dependence of $I_{m,r}(\eta, z)$ on $r$ is via the number $k$). We therefore have
\[
|I_{m,r}(\eta, z)| \leq N \max_i \left| n_i(z) \frac{\eta^{m-k} s_i(z)^k}{\eta^2 - s_i(z)^2} m \right|, \quad (A17)
\]
When the imaginary part of $\eta$ is nonzero, the denominator in the right hand side of Eq. (A17) cannot vanish for any value of $s_i(z)$ (while as $\eta \rightarrow i0$, 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 $N \rightarrow \infty$). Assuming $\text{Im} \eta > 0$, it will be sufficient for our purposes to substitute Eq. (A17) with the weaker bound
\[
|I_{m,r}(\eta, z)| \leq N \max_s \left| \eta^{m-k} s_i(z)^k \left( \eta^2 - s_i(z)^2 \right)^m \right|, \quad (\text{Im} \eta > 0)
\]
(A18)
where now the maximum is taken for $s$ ranging over the whole $[0, \infty)$. Since $\text{Im} \eta > 0$ the expression has no singularities at finite real $s$, and since $2m > k$, it cannot diverge as $s \rightarrow \infty$ either; thus it has a finite maximum independent of $N$. More precisely, it is easy to show that
\[
\max_s \left| \eta^{m-k} s_i(z)^k \left( \eta^2 - s_i(z)^2 \right)^m \right| \leq \left| \frac{\sqrt{2}}{\text{Im} \eta} \right|^m, \quad (\text{irrespective of } k \text{ as long as } 1 \leq k \leq m)
\]
yielding
\[
|I_{m,r}(\eta, z)| \leq N \left| \frac{\sqrt{2}}{\text{Im} \eta} \right|^m, \quad (\text{Im} \eta > 0). \quad (A19)
\]
Similarly, the trace for the external orbit can be written in the new basis Eq. (A11) as
\[
E_{n,\tilde{r}}(\eta, z) = \sum_{i=1}^{N} A_{ii}(z) \eta^{n-k} n_i(z) \tilde{k}, \quad (1 \leq \tilde{k} \leq n), \quad (A20)
\]
where $\tilde{k}$ is the number of Green’s functions in Eq. (A2) sandwiched between two $\pi^r$’s with different superscripts; this convention works correctly for the external orbit as well, if we account for the presence of $\sigma^r$ by imagining a $\pi^2$ ($\pi^1$) to the left (right) of the leftmost (rightmost) Green’s function. From Eq. (A15), we can write $A_{ii}(z) = u_i(z) \dagger (RL)^{-1} v_i(z)$, where we defined the vectors $u_i(z)$ and $v_i(z)$ to be the $i$-th column of $U_z$ and $V_z$, respectively. By the Cauchy-Schwarz inequality we then have
\[
|A_{ii}(z)| \leq \|u_i(z)\| \|v_i(z)\| \|RL^{-1}\|, \quad (A21)
\]
where $\|RL^{-1}\|$ is the operator norm, or the maximum singular value, of $RL^{-1}$. But since $U_z$ and $V_z$ are unitary matrices, $u_i(z)$ and $v_i(z)$ are unit vectors, and we obtain
\[
|A_{ii}(z)| \leq \|RL^{-1}\|. \quad (A22)
\]
Going back to Eq. (A20), this yields the bound
\[
|E_{n,\tilde{r}}(\eta, z)| \leq N \|RL^{-1}\| \max_i \left| \eta^{n-k} s_i(z)^k \right| \left( \eta^2 - s_i(z)^2 \right)^m. \quad (A23)
\]
The only difference with the inequality for $I_{m,r}$ is the factor $\|RL^{-1}\|$. Repeating the same argument as for the internal traces, we therefore see that
\[
|E_{n,\tilde{r}}(\eta, z)| \leq N \left| \frac{\sqrt{2}}{\text{Im} \eta} \right|^n \|RL^{-1}\|, \quad (\text{Im} \eta > 0), \quad (A24)
\]
and thus a sufficient condition for $E_{n,\tilde{r}}$ to be $O(N)$ for $\text{Im} \eta > 0$, that is $\|RL^{-1}\|$ remains bounded as $N \rightarrow \infty$, i.e.
\[
\|RL^{-1}\| = O(1). \quad (A25)
\]
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), wavy lines, and $s$ solid lines, by $2^w \left| \frac{\sqrt{2}}{\text{Im} \eta} \right|^n \|RL^{-1}\||N^{2g}|$ (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 $\text{Im} \eta$, the contribution of crossing diagrams (i.e. those with $g \geq 1$) goes to zero as $N \rightarrow \infty$. Thus if we take the limit $N \rightarrow \infty$ before the limit $\eta \rightarrow i0^+$, ignoring the crossing diagrams is safe, and the expression for $\rho(z)$ obtained from Eq. (3.16), after analytic continuation of $\text{tr}(\sigma^+ (RL)^{-1} G(\eta, z))$ to $\eta = i0$, with $G(\eta, z)$ given by the contribution of non-crossing diagrams to $(G(\eta, z; J)_{ij})$, gives the correct result for $\lim_{N \rightarrow \infty} \rho(z)$. We mention that when the smallest singular value $s_i(z)$ remains bounded away from zero as $N \rightarrow \infty$, even at $\eta = 0$ the traces Eqs. (A16) and (A20) are $O(N)$, as is
not hard to check, justifying the non-crossing approximation at $\eta = 0$. Thus it is only when some $s_i(z)$ are $o(1)$ that it becomes important to send $\eta$ to $i0^+$ only after the limit $N \to \infty$ has been taken. In particular, in such cases, applying the limit $\eta \to 0^+$ to the results obtained using the non-crossing 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 $\rho(z) \equiv \lim_{z \to \infty} \lim_{N \to \infty} \rho_N(z, \epsilon) = 0$, where

$$
\rho_N(z, \epsilon) \equiv \frac{1}{\pi} \frac{\partial}{\partial z} \text{tr} \left( (RL)^{-1} M_z^{\dagger} I + \gamma^2 \right),
$$

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 $N \to \infty$, $\gamma$ has no dependence on $N$). Recall that for $\epsilon > 0$, $g(z, \epsilon)$ is positive and therefore $\gamma > 0$. Expanding the derivative in Eq. (B1) we obtain

$$
\pi \rho_N(z, \epsilon) = \text{tr} \left( (RL)^{-1} (RL)^{-1} \right) - \text{tr} \left( (RL)^{-1} M_z^{\dagger} I_M z (RL)^{-1} \right) - \text{tr} \left( (RL)^{-1} M_z^{\dagger} I_M z (RL)^{-1} \right) \frac{\partial z}{\partial \gamma^2} = \text{tr} \left( (RL)^{-1} Q (RL)^{-1} \right) - \text{tr} \left( (RL)^{-1} M_z^{\dagger} I_M z (RL)^{-1} \right) \frac{\partial z}{\partial \gamma^2} - \text{tr} \left( (RL)^{-1} M_z^{\dagger} I_M z (RL)^{-1} \right) \frac{\partial z}{\partial \gamma^2},
$$

where we defined $Q = 1 - M_z^{\dagger} I_M z (RL)^{-1}$ (we suppress the explicit dependence of $\gamma$ on $z$ for simplicity). By the Woodbury matrix identity $Q = \frac{\gamma^2}{M_z M_z^{\dagger} z}$, upon substitution in Eq. (B2) yields

$$
\pi \rho_N(z, \epsilon) = \text{tr} \left( (RL)^{-1} (RL)^{-1} \right) \gamma^2 - \text{tr} \left( (RL)^{-1} M_z^{\dagger} I_M z (RL)^{-1} \right) \frac{\partial z}{\partial \gamma^2}.
$$

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

$$
- \partial_{\bar{z}} \gamma^2 = \frac{-2\gamma^2 \partial_{\bar{z}} K}{1 - K - 2\gamma^2 \partial_{\bar{z}} K},
$$

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

$$
- \partial_{\bar{z}} K = T^\infty_\infty(\gamma) \equiv \lim_{N \to \infty} \frac{-\partial_{\bar{z}} K}{N} \quad \text{and} \quad - \partial_{\bar{z}} K = V^\infty_\infty(\gamma) \equiv \lim_{N \to \infty} \frac{-\partial_{\bar{z}} K}{N},
$$

where we defined

$$
T^\infty_\infty(\gamma) \equiv \text{tr} \left( \frac{1}{(M_z M_z^{\dagger} z + \gamma^2)^2} \right)
$$

$$
V^\infty_\infty(\gamma) \equiv \text{tr} \left( \frac{1}{(M_z M_z^{\dagger} z + \gamma^2)^2} \right).
$$

We thus obtain

$$
\pi \rho_N(z, \epsilon) = \gamma^2 T^\infty_\infty(\gamma) + \frac{2\gamma^2 V^\infty_\infty(\gamma) V^\infty_\infty(\gamma)^*}{1 - K(\gamma) + 2\gamma^2 T^\infty_\infty(\gamma)}
$$

where we defined

$$
T(\gamma) \equiv \text{tr} \left( (RL)^{-1} M_z^{\dagger} I_M z + \gamma^2 M_z M_z^{\dagger} z + \gamma^2 \right).
$$

Having eliminated derivatives of $\gamma$, we now simply need to show that $\lim_{z \to \infty} \lim_{N \to \infty}$ 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 $T(\gamma)$ and $V(\gamma)$ in Eq. (B8). For $V(\gamma)$ we use the singular value decomposition Eq. (3.32):

$$
|V(\gamma)| \leq \|LR\|^{-1} \|L\gamma\|_{(\gamma)} = \text{tr} \left( U^{\dagger} \gamma V^{\dagger} \gamma V \right) \leq \|\gamma\| \|\gamma V\|_{(\gamma)} \leq \|\gamma\| \|\gamma V\|_{(\gamma)}
$$

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

$$
|V_{xy}(\gamma)| \leq \|LR\|^{-1} \|V(\gamma)\|_{(\gamma)}
$$

where we defined

$$
V_{xy}(\gamma) \equiv \frac{1}{N} \sum_{i=0}^{N} \frac{s_i(z)}{(s_i(z) + \gamma^2)^2}.
$$

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

$$
|V(\gamma)| \leq \|LR\|^{-1} \|V(\gamma)\|_{(\gamma)}.
$$

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

$$
\left|\text{tr} \left[ \frac{(RL)^{-1}}{M_2^T M_2 + \gamma^2 M_2^T M_2 + \gamma^2} \right]\right| \leq \|RL\|^{-1} \left[ \frac{1}{M_2^T M_2 + \gamma^2} \right]^2 \quad (B15)
$$

or

$$
|T_N(\gamma)| \leq \|RL\|^{-1}\|T_N^\gamma(\gamma)\|. \quad (B16)
$$

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

$$
\pi|\rho_N(z, \epsilon)| \leq C^2 \left[ \gamma^2 T_N^\gamma(\gamma) + \frac{2 \gamma^2 V_N^\gamma(\gamma) V_N^\gamma(\gamma)}{1 - K(\gamma) + 2 \gamma^2 T_N^\gamma(\gamma)} \right] \quad (B17)
$$

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

$$
\pi \rho(z, \epsilon) \leq C^2 \left[ \gamma^2 T_N^\gamma(\gamma) + \frac{2 \gamma V_N^\gamma(\gamma)}{1 - \gamma^2 T_N(\gamma)} \right] \quad (B18)
$$

where we defined

$$
T_N^\gamma(\gamma) \equiv \lim_{N \to \infty} \frac{1}{N} \sum_{s_i(z) = 0}^{N} \frac{1}{(s_i(z)^2 + \gamma^2)^2}, \quad (B19)
$$

$$
V_N^\gamma(\gamma) \equiv \lim_{N \to \infty} \frac{1}{N} \sum_{s_i(z) = 0}^{N} \frac{s_i(z)}{(s_i(z)^2 + \gamma^2)^2}. \quad (B20)
$$

Thus, to show that $\lim_{\epsilon \to 0^+} \rho(z, \epsilon) = 0$, it suffices to show that $\gamma^2 T_N^\gamma(\gamma)$ and $\gamma V_N^\gamma(\gamma)$ vanish as $\gamma \to 0^+$ (since $z$ is in the region Eq. (3.38), $1 - K(\gamma)$ and hence the denominator in the last term in Eq. (B19) remains positive as $\gamma \to 0^+$). Let us rewrite Eq. (B19) as

$$
T_N^\gamma(\gamma) = \int_0^\infty \frac{\rho(s; z) ds}{(s^2 + \gamma^2)^2} \quad (B21)
$$

$$
V_N^\gamma(\gamma) = \int_0^\infty \frac{s \rho(s; z) ds}{(s^2 + \gamma^2)^2} \quad (B22)
$$

where we defined

$$
\rho_b(s; z) = \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N} \delta(s - s_i(z)) \quad (B23)
$$

as the limit of the density of the singular values of $M_N$. Note that contributions to $T_N^\gamma(\gamma)$ and $V_N^\gamma(\gamma)$ 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 $\gamma^2 T_N^\gamma(\gamma)$ and $\gamma V_N^\gamma(\gamma)$ 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 $\gamma^2 \int_0^{s_0} \rho_b(s; z) ds$ and $\gamma \int_0^{s_0} s \rho_b(s; z) ds$ vanish as $\gamma \to 0^+$. Let us first consider the situation similar to that in the two examples Eqs. (5.1) and (2.40). For those examples, there is a region of $z$ outside Eq. (3.35), 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. But 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 $\rho_b(s; z)$ 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

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

We conclude that as $s \to 0^+$ the density, $\rho_b(s; z)$, must vanish at least as fast as $s^\alpha$, 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 $s \leq s_0$, $\rho(s; z) < c s^\alpha$ for some constant $c$ and $\alpha > 1$. It is then an elementary exercise to show that $\gamma^2 \int_0^{s_0} \frac{s \rho_b(s; z) ds}{(s^2 + \gamma^2)^2}$ and $\gamma \int_0^{s_0} \frac{s^2 \rho_b(s; z) ds}{(s^2 + \gamma^2)^2}$ are $O(\gamma^{\text{min}(2,\alpha-1)})$ and $O(\gamma^{\text{min}(1,\alpha-1)})$, 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 (2004).
[2] Z. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, Science Press (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] P. Strata and R. Harvey, *Brain Res. Bull.* 50, 349 (1999).
[6] B. K. Murphy and K. D. Miller, *Neuron* 61, 635 (2009).
[7] S. Ganguli, D. Huh, and H. Sompolinsky, *Proceedings of the National Academy of Sciences* 105, 18970 (2008).
[8] M. S. Goldman, *Neuron* 61, 621 (2009).
[9] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra*, Princeton University Press (2005).
[10] J. Feinberg and A. Zee, *Nuclear Physics B* 504, 579 (1997).
[11] J. Feinberg and A. Zee, *Nuclear Physics B* 501, 643 (1997).
[12] B. Khoruzhenko, *Journal of Physics A: Mathematical and General* 29, L165 (1996).
[13] P. Biane and P. Lehner, *Colloq. Math.* 90, 181 (2001).
as in the atypical cases under discussion, in a finite region of the complex plane the smallest singular value of $M_\varepsilon$ is nonzero for finite $N$, but vanishes in the limit $N \to \infty$. This means that the operator norm of $(z - M)^{-1}$ is finite over such a region but goes to infinity as $N \to \infty$. Hence, if we define $\Sigma_\infty(M) \equiv \lim_{N \to \infty} \Sigma_i(M)$ and $\Sigma_\infty(M) \equiv \lim_{N \to \infty} \Sigma(M)$, we see that in such cases $\lim_{\varepsilon \to 0} \Sigma_\infty(M)^{-1} \neq \Sigma_\infty(M)$ (or equivalently, $\lim_{\varepsilon \to 0} \lim_{N \to \infty} \Sigma_i(M) \neq \lim_{N \to \infty} \lim_{\varepsilon \to 0} \Sigma_i(M)$). More generally but less precisely, this indicates that at finite but large $N$, the $\varepsilon$-pseudospectra of such matrices can cover a significantly broader region than the spectrum even for very small $\varepsilon$, indicating extreme nonnormality.

[39] This equivalence is true more generally for any matrix norm derived from a general vector norm; see Ref. [9] for a proof.

[40] The unitary invariance of these formulae is in turn a consequence of the invariance of both the corresponding quantities ($p(z)$ and $|x(t)|^2$), as well as the statistical ensemble for $J$, Eq. (3.2), and hence that of $LR$ when $L \propto R \propto 1$, under unitary transforms like Eq. (3.3).

[41] It is also common to write the firing rate equations in the different form, $T(t) \equiv -r(t) + f(Wr(t) + \Gamma(t))$. At least in the case where all neurons have equal time constants, i.e. $T \propto 1$, the two formulations are equivalent and are related by the change of variable $v \equiv Wr + \Gamma$ [37].

[42] This structure is a consequence of using a complex ensemble for $J$, for which the covariances $(J_{ab}, J_{cd})$ vanish. For the real Gaussian ensemble, by contrast, the latter do not vanish; in this case Eq. (3.21) becomes $(\sigma_{\alpha\beta}^a \sigma_{\gamma\delta}^b + \delta_{\alpha\beta} \delta_{\gamma\delta} + \sigma_{\alpha\gamma}^a \sigma_{\beta\delta}^b + \sigma_{\alpha\delta}^a \sigma_{\beta\gamma}^b) = \frac{1}{N} \sum_{r=1}^N (\pi^r_{\alpha\beta} \delta_{\gamma\delta} + \pi^r_{\alpha\gamma} \delta_{\beta\delta} + \pi^r_{\alpha\delta} \delta_{\beta\gamma} + \pi^r_{\gamma\delta} \delta_{\alpha\beta}).$

[43] Notice that this is a more restrictive property than planarity of the diagram; for example the graph in panel (b) of Fig. 14 is planar, as one of the wavy lines can be drawn inside the solid loop without crossing any other line, but it is non-crossing as defined here.

[44] More precisely, we only need to define the limit Eq. (1.23) 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 $\lim_{N \to \infty} \frac{1}{N} \sum_{s=1}^N f(s(z)^2) = \int_0^\infty f(s^2) \rho(s;z)ds$. We do not assume any smooth form for $\rho(s;z)$; in particular, $\rho(s;z)$ may have delta function singularities whenever 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 $\|M\|_\varepsilon$ remain bounded automatically guarantees that these will not be numerous enough to contribute to $\rho(s;z)$ at infinity.