UNIVERSITY OF CALIFORNIA SAN DIEGO

Dynamics and Information Processing in Recurrent Networks

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Physics

by

Alexander Kuczala

Committee in charge:

Professor Tatyana O. Sharpee, Chair
Professor Massimo Vergassola, Co-Chair
Professor Daniel Arovas
Professor Terrence J. Sejnowski
Professor Charles F. Stevens
Professor Congjun Wu

2019

The dissertation of Alexander Kuczala is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Co-Chair

_____
Chair

University of California San Diego

2019

EPIGRAPH

The most merciful thing in the world, I think, is the inability of the human mind to correlate all its contents.

—H.P. Lovecraft

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank Tatyana Sharpee for her guidance and support as advisor, and my committee members Massimo Vergassola, Terry Sejnowski, Chuck Stevens, Dan Arovas, and Congjun Wu. I would also like to thank Jack Berkowitz for many discussions, especially those concerning dubious integrals and non-parametric estimators. I am grateful to my friends for encouraging me to believe in them who believed in me, and I grateful to my family for their support.

Chapter 2, in part, is based on material contained in Kuczala and Sharpee. *Eigenvalue spectra of large correlated random matrices.* Physical Review E, 94(5), 05010, 2016. The dissertation author was the primary investigator and author of this paper.

Chapter 2 and Chapter 3, in part, are currently being prepared for submission for publication of the material. Kuczala and Sharpee. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in part, is currently being prepared for submission for publication of the material. Kuczala and Sharpee. The dissertation author was the primary investigator and author of this paper.

Chapter 5, in part, is currently being prepared for submission for publication of the material. Kuczala and Sharpee. The dissertation author was the primary investigator and author of this paper.

VITA

2012	B. S. in Physics *magna cum laude*, University of Rochester
2012	B. S. in Mathematics *magna cum laude*, University of Rochester
2012-2016	Graduate Teaching Assistant, University of California San Diego
2019	Ph. D. in Physics, University of California San Diego

PUBLICATIONS

Malenda, R.F., Price, T.J., Stevens, J., Uppalapati, S.L., Fragale, A., Weiser, P.M., Kuczala, A., Talbi, D. and Hickman, A.P., 2015. *Theoretical calculations of rotationally inelastic collisions of He with NaK (A 1+): Transfer of population, orientation, and alignment.* The Journal of chemical physics, 142(22), p.224301.

Lombardini, R., Acevedo, R., Kuczala, A., Keys, K.P., Goodrich, C.P. and Johnson, B.R., 2016. *Higher-order wavelet reconstruction/differentiation filters and Gibbs phenomena.* Journal of Computational Physics, 305, pp.244-262.

Kuczala, A. and Sharpee, T.O., 2016. *Eigenvalue spectra of large correlated random matrices.* Physical Review E, 94(5), p.050101.

ABSTRACT OF THE DISSERTATION

Dynamics and Information Processing in Recurrent Networks

by

Alexander Kuczala

Doctor of Philosophy in Physics

University of California San Diego, 2019

Professor Tatyana O. Sharpee, Chair
Professor Massimo Vergassola, Co-Chair

Random recurrent networks facilitate the tractable analysis of large networks. The spectrum of the connectivity matrix, determined analytically by random matrix techniques, determines the network's linear dynamics as well as the stability of the nonlinear dynamics. Knowledge of the onset of chaos helps determine the network's computational capabilities and memory capacity. However, fully homogeneous random networks lack the non-trivial structures found in real world networks, such as cell types and plasticity induced correlations in neural networks. We address this deficiency by investigating the impact of correlations between forward and reverse connections, which may depend on the neuronal type. Using

random matrix theory, we derive a formula that efficiently computes the eigenvalue spectrum of large random matrices with block-structured correlations. The inclusion of structured correlations distorts the eigenvalue distribution in a nontrivial way; the distribution is neither a circle nor an ellipse. We find that layered networks with strong interlayer correlations have gapped spectra. For antisymmetric layered networks, oscillatory modes dominate the linear dynamics.

We analyze the effect of structured correlations on the nonlinear dynamics of rate networks by developing a set of dynamical mean field equations applicable for large system sizes. We find that the power spectrum of strongly antisymmetric bipartite networks peaks at nonzero frequency, miming the gap present in the eigenvalue distribution.

Heterogeneous connection statistics facilitate the presence of strongly feed-forward connections in addition to recurrent ones, both of which promote signal amplification. We investigate the role of feed-forward amplification in i.i.d. block-structured networks by computing the Fisher information of past input perturbations. We apply this result to find the optimal architecture for information retention in two populations, under energy constraints. We find that this architecture is both strongly feed-forward and recurrent, with the respective strengths of these connections depending on the available synaptic gain.

Finally, we assess the ability of rate networks to dynamically approximate the dominant mode of a random symmetric matrix. Given an initial estimate of the eigenvector as input, we find that there is an optimal processing time and synaptic gain strength depending on the dimensionality and quality of the initial estimate.

Chapter 1

Introduction

Random recurrent networks pragmatically characterize the dynamics and computational capabilities of large networks, permitting tractable analysis. These networks contain a plentitude of feedback connections that allow "memories" of input signals to be stored for long periods of time. Recurrent networks iteratively transform input signals, as well as generate spontaneous activity in the absence of input, supplying a high-dimensional *reservoir* of signals that can be employed in computational tasks [30, 36, 14, 63, 24]. However, the strength of these connections strongly impact the network's information processing capabilities. If the feedback connections are too weak, signals in the network decay too rapidly, whereas if the connections are too strong, echoes of past signals pollute the network with noise. It has been proposed that large random networks and other systems with random interactions maximize information processing at the *edge of chaos* [40, 63, 49, 69, 14, 34, 55, 41], although the scope of applicability of this principle is disputed [12, 47, 11, 56]. These systems undergo a sharp transition from non-chaotic to chaotic behavior as the strength of connections is increased. It is near this transition that large disordered networks typically have the greatest information capacity.

Another key network property instrumental in information retention is *non-normality*.

Networks with non-normal connections have hidden feed-forward structure, which amplifies and propagates signals within the network. Non-normal connectivity can arise, for example, from balanced excitation and inhibition [50, 1], from explicit feed-forward connectivity [29], or from more subtle architectures [70, 1]. Introducing heterogeneity into the connections of random networks allows us to investigate the impact of non-normal structure on information processing.

Our prototypical network throughout this dissertation is the the firing rate model of neural networks. In this model, the activities $x_i(t)$ of the N neurons obey

$$\dot{x}_i(t) = -x_i(t) + \sum_{j=1}^{N} J_{ij}\phi[x_j(t)] \tag{1.1}$$

The function $\phi(x)$ is a saturating nonlinearity that maps neuron activities to a finite range of firing rates. We will frequently choose $\phi(x) = \tanh(x)$, which saturates to ± 1. The term $-x_i(t)$ acts as a capacitance, causing each node to decay exponentially in the absence of input from other neurons. Each element J_{ij} of the $N \times N$ connectivity matrix describes the weight from the firing rate of node j to the node i. These weights can be positive or negative, and are drawn randomly from some probability distribution $\mathrm{P}(J)$. This assumption lends itself well to analysis. When the number of neurons N is large, the exact choice of connectivity J does not impact the macroscopic behavior of the network – a well-known result from random matrix theory that is a non-abelian generalization of the central limit theorem [66].

To gain some intuition about the network dynamics, we first consider the linear case $\phi(x) = x$, where

$$\dot{x}_i(t) = -x_i(t) + \sum_{j=1}^{N} J_{ij}x_j(t). \tag{1.2}$$

The solution of this system is $\mathbf{x}(t) = e^{(\mathbf{J}-\mathbf{1})t}\mathbf{x}(0)$, which, in the eigenbasis $\{\mathbf{e}_k\}$ corresponding

to the eigenvalues $\{\lambda_k\}$ of J, is $x_k(t) = e^{(\lambda_k - 1)t} x_k(0)$. This system is stable only if the rightmost eigenvalue[1] of J is less than one, in which case the network activity decays to zero. When the system is unstable, the network activity blows up exponentially.

In the simplest case, the connections J_{ij} are all drawn independently and identically (i.i.d.) from a Gaussian distribution with mean zero and variance g^2/N. The gain g is a parameter that controls the strength of the connections. Remarkably, for large N, the eigenvalues of J lie uniformly in a disk of radius g in the complex plane [33]. See Fig. 1.1. This means that, for large networks, the rightmost eigenvalue is g, so that the system is stable for $g < 1$ and unstable for $g > 1$.

In the nonlinear case where $\phi(x) = \tanh(x)$, the saturating nonlinearity prevents the network activity from blowing up when $g > 1$. The Jacobian is

$$\frac{\partial \dot{x}_i}{\partial x_j} = -\delta_{ij} + J_{ij}\phi'(x_j). \tag{1.3}$$

For $g < 1$, the system is everywhere contracting, and $x = 0$ is the unique fixed point[2]. This is known as the silent or quiescent phase, where all activity decays to zero, as in the linear case (Fig. 1.1). When $g = 1$, the $x = 0$ fixed point becomes unstable, and the number of fixed points (most of which are saddles) proliferates as g increases [72]. This is a distinctive feature of chaotic dynamics [21]. Indeed, well-known dynamical mean field methods can be used to show that for $g > 1$, the Lyapunov exponent is positive, and consequently the network dynamics are chaotic [59, 23].

However, the assumption that connections are independent is limiting, and does not address the affect of correlations. Most notably, correlations between connection strengths arise as the result of plasticity, where connections are modified depending on node activity and network input. One of the predominant effects of plasticity is that it

[1] The eigenvalue with greatest real part, also known as the spectral abscissa

[2] Banach fixed point theorem

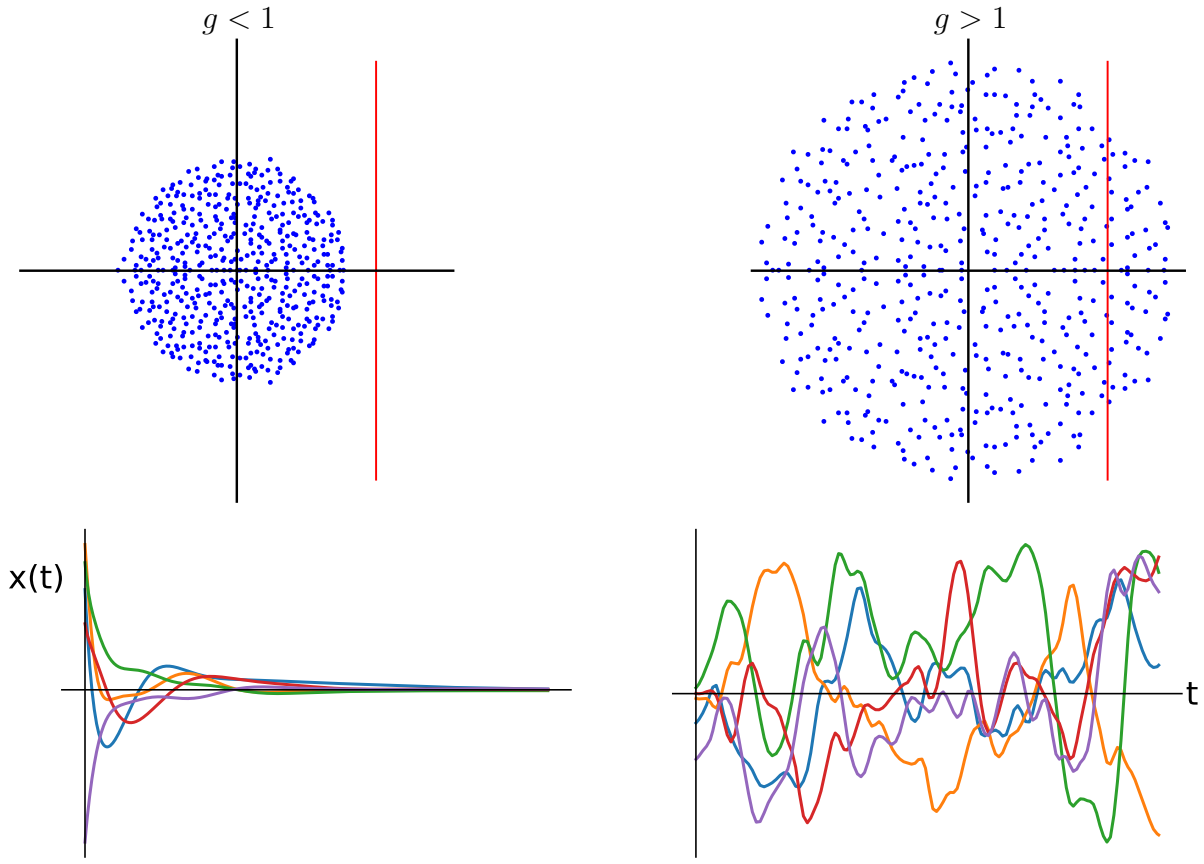


Figure 1.1: For a large, homogeneous random network, the eigenvalue spectrum of g lies in a disk of radius g on the complex plane, so that the rightmost eigenvalue is g. When $g < 1$, the network activity decays to zero, and for $g > 1$, the network activity is chaotic

induces correlations between forward and reverse connections [31, 46]. That is, the degree to which node i affects node j is correlated with the strength of the reverse connection from node j to node i. We also wish to account for macroscopic heterogeneity in our network models. Heterogeneity can arise, for example, as a result of distinct cell types in the brain, which may be selective in their connection strengths and correlations to other cell types.

In the first part of this dissertation, we account for both heterogeneity and correlations in network dynamics. In Chapter 2, we use random matrix theory to compute the eigenvalue distribution of large matrices with structured correlations. We find that the combined effect of correlations and heterogeneity yields exotic eigenspectra not found in random matrices with either property alone. Furthermore, the structure of these eigenspectra are difficult to characterize succinctly; there are a wide variety of shapes and behaviors that depend on the relative strengths of the correlations and variances. We analyze special cases that exemplify these unique features.

In Chapter 3, we study the mean field dynamics of rate networks with block-structured correlations. Using dynamical mean field theory (DMFT), we derive a set of self-consistent equations that describe the equilibrium dynamics. For finite-sized networks, this description is complicated by the presence of nonzero fixed points, whose presence or absence is difficult to determine analytically.

In the second part of this work, we assess the ability of structured recurrent networks to store information and perform computations. In Chapter 4, we extend the work of Toyoizumi and Abbott [69] to compute the information that heterogeneous networks retain about input perturbations. We find that the integrated Fisher memory curve [29] is maximized when the structure matrix, describing the mean strength of connections between populations, is highly non-normal. In the two-population case, networks with a combination of strong feed-forward and recurrent connections maximize information retention. Finally, in Chapter 5, we assess the ability of symmetric networks to perform

error correction. In particular, we determine the conditions under which an ensemble of random symmetric networks can improve the estimate of the eigenvector corresponding to the rightmost eigenvalue of the connectivity matrix.

Chapter 2

Random matrix theory

Random matrices serve as a useful tool for analyzing the stability and dynamics of a variety of networks, from neuroscience [59, 52, 5, 4] and genetic circuits [2] to ecology [44, 6]. The eigenspectra of random matrices also help determine solutions to problems in nuclear [73] and condensed matter physics [8, 60] as well as in data compression [42, 19]. In particular, the rightmost eigenvalue (the eigenvalue with largest real component) determines the stability of the system's linear dynamics and onset of chaos of the nonlinear dynamics. Knowledge of the onset of chaos is also useful for determining the network's computational capabilities [63, 14] as well as the network's response to inputs [53].

However, most of these results do not address an important feature of biological circuits where connection strengths are correlated [31, 61, 46]. While correlated Hermitian ensembles have received some attention, [71, 57, 39, 18], results about correlated non-Hermitian ensembles are scarce [58, 54].

2.1 Hermitian random matrix theory

By means of introduction, we first consider the case where J is a complex $N \times N$ hermitian matrix, distributed according to

$$P(J) = \frac{1}{Z} e^{-N \mathrm{tr} V(J)} \tag{2.1}$$

where $Z = \int dJ P(J)$ is the partition function. Since J is hermitian, it has only N^2 independent elements, whereas a general complex matrix has $2N^2$ independent components. This allows us to write the integral measure as dJ, as opposed to $dJ d\bar{J}$. A nice way to see this is to transform the integral to an integral over the real eigenvalues of J.

If we choose a quadratic potential $V(J) = \frac{1}{2g^2} J^2$, then the independent elements of J obey a Gaussian distribution with zero mean and variance g^2/N. Specifically,

$$\langle J_{ij} \rangle = 0 \tag{2.2}$$

$$\langle J_{ij} \bar{J}_{kl} \rangle = \frac{g^2}{N} \delta_{il} \delta_{jk} \tag{2.3}$$

where $\langle \cdot \rangle$ denotes the expectation value over $P(J)$. We can study the distribution of eigenvalues of J with the Green's function (also called the resolvent) [76]

$$G(z) = \frac{1}{N} \left\langle \sum_{i=1}^{N} \frac{1}{z - \lambda_i} \right\rangle \tag{2.4}$$

where λ_i are the eigenvalues of J. This function has poles at the eigenvalues of J. Since J is hermitian, its eigenvalues will lie on the real axis. The density of eigenvalues $\rho(x)$ can be determined by taking the imaginary part of $G(z)$. We can see this as follows [76]: on the real axis $z = x$,

$$\frac{1}{x + i\epsilon} = \frac{x}{x^2 + \epsilon^2} - \frac{i\epsilon}{x^2 + \epsilon^2} \tag{2.5}$$

For any $\epsilon > 0$,

$$\int dx \frac{i\epsilon}{x^2 + \epsilon^2} = -i\pi \tag{2.6}$$

The width of this function decreases with ϵ, approaching a delta function as $\epsilon \to 0$. Taking this limit,

$$\lim_{\epsilon \to 0} \frac{1}{x + i\epsilon} = \frac{1}{x} - i\pi\delta(x) \tag{2.7}$$

Thus we can see that

$$\begin{aligned} \rho(x) &= \lim_{\epsilon \to 0} \frac{1}{\pi} \mathrm{Im} G(x - i\epsilon) \\ &= \lim_{\epsilon \to 0} \frac{1}{N\pi} \left\langle \sum_i \mathrm{Im} \frac{1}{x - i\epsilon - \lambda_i} \right\rangle \\ &= \left\langle \frac{1}{N} \sum_i \delta(x - \lambda_i) \right\rangle \end{aligned} \tag{2.8}$$

as desired. The Green's function $G(z)$ will have branch cuts or poles precisely on the support of $\rho(x)$. Wherever the expected spectrum has an isolated eigenvalue, $G(z)$ has a pole, and wherever the expected spectrum is continuous, $G(z)$ has a branch cut (which can be thought of as a continuous set of poles). Everywhere else, $G(z)$ is holomorphic. This means that we may compute $G(z)$ analytically outside the support of the spectrum, and take the limit to compute $\rho(x)$.

We can calculate $G(z)$ by expanding in powers of $1/z$ [76]. This expansion is valid for all points outside the eigenvalue distribution, where $G(z)$ is holomorphic. We find

$$G(z) = \frac{1}{N} \mathrm{tr} \left\langle \frac{1}{z - J} \right\rangle = \sum_{n=0}^{\infty} \mathrm{tr} \left\langle \frac{1}{z} \left(\frac{1}{z} J \right)^n \right\rangle \tag{2.9}$$

We make this calculation tractable by expressing $G(z)$ in terms of diagrams [76, 16]. First

we define the matrix Green's function

$$\mathcal{G} = \left\langle \frac{1}{z - J} \right\rangle \tag{2.10}$$

from which we can recover $G(z) = \frac{1}{N}\text{tr}\mathcal{G}(z)$. Then we have the matrix series

$$\mathcal{G}(z) = \sum_{n=0}^{\infty} \left\langle \frac{1}{z} \left(\frac{1}{z} J \right)^n \right\rangle = \mathcal{G}_0 + \langle \mathcal{G}_0 J \mathcal{G}_0 J \mathcal{G}_0 \rangle + \langle \mathcal{G}_0 J \mathcal{G}_0 J \mathcal{G}_0 J \mathcal{G}_0 J \mathcal{G}_0 \rangle + \dots \tag{2.11}$$

Here, $(\mathcal{G}_0)_{ij} \equiv z^{-1}\delta_{ij}$, and the odd terms vanish since $\langle J \rangle = 0$. Furthermore, since the distribution over J is Gaussian, each term in the sum reduces to the Wick contraction of n factors of J.

We represent $\mathcal{G}_0$ by a single directed line carrying one index, and the correlator $\langle J_{ij} J_{kl} \rangle = \frac{g^2}{N}\delta_{il}\delta_{jk}$ by a double line carrying two sets of indices (Fig. 2.1) [28, 38, 27, 16]. In analogy to QCD, these are called 'quark' and 'gluon' propagators, respectively. Each line indicates the action of the delta function, which is to set its first index equal to its second. Indices are summed at each connecting vertex. The nth term in the series (2.11) is the sum of all diagrams with n vertices. As an example, the quartic term in the $\mathcal{G}(z)$ expansion is

$$\frac{1}{z} \sum_{jkl} \left\langle \frac{1}{z} J_{ij} \frac{1}{z} J_{jk} \frac{1}{z} J_{kl} \frac{1}{z} J_{lm} \frac{1}{z} \right\rangle \tag{2.12}$$

Wick contracting, we get

$$\begin{aligned} &\frac{1}{z^5} \sum_{jkl} \left[\langle J_{ij} J_{jk} \rangle \langle J_{kl} J_{lm} \rangle + \langle J_{ij} J_{kl} \rangle \langle J_{jk} J_{lm} \rangle + \langle J_{ij} J_{lm} \rangle \langle J_{jk} J_{kl} \rangle \right] \\ &= \frac{g^4}{N^2 z^5} \sum_{jkl} \left[\delta_{ik}\delta_{jj}\delta_{km}\delta_{ll} + \delta_{il}\delta_{jk}\delta_{jm}\delta_{kl} + \delta_{im}\delta_{jl}\delta_{jl}\delta_{kk} \right] \\ &= \frac{g^4}{N^2 z^5} \left[N^2 + 1 + N^2 \right] \delta_{im} \end{aligned} \tag{2.13}$$

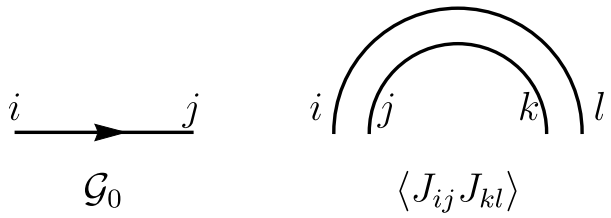


Figure 2.1: Propagators in diagrammatic expansion of (2.11)

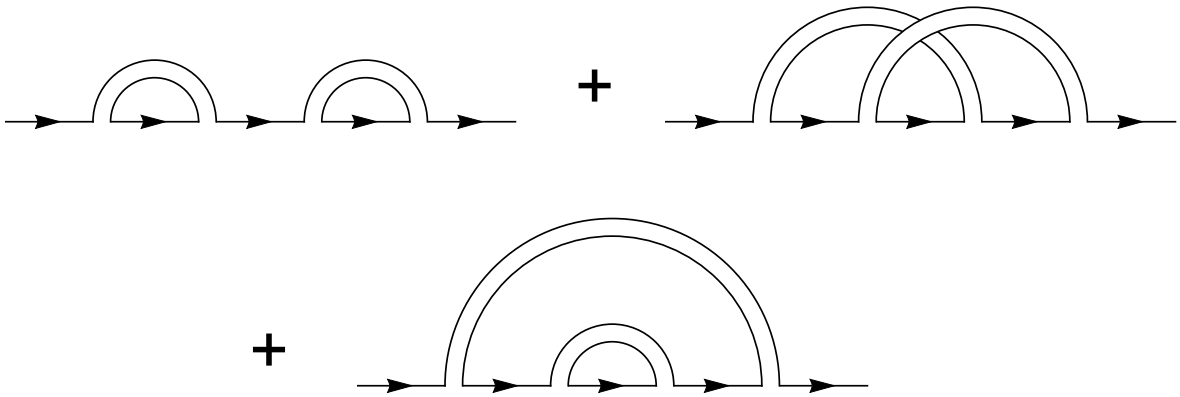


Figure 2.2: Fourth order diagrams in (2.13). The second term, which is nonplanar, is subleading in $1/N$.

Compare this explicit calculation with the diagrammatic representation (Fig. 2.2). For every vertex, include a factor of g^2/N, for every propagator $\mathcal{G}_0$, include a factor of z^{-1}, and for every loop, include a factor of N. This immediately yields the result above. For large N we can drop all but the $\mathcal{O}(1)$ terms. The neglected terms correspond to diagrams which crossing lines, and only "planar" diagrams remain (see Fig 2.3). Planar diagrams are precisely those that can be embedded in the plane with no intersections. There is a very nice topological proof of this due to t'Hooft [64, 27].

Although we have reduced the number of diagrams, there are still infinitely many planar diagrams to sum over. We can solve this problem by resumming the series, a technique well-known in quantum field theory [76]. In order to do this explicitly, we exploit the recursive nature of the diagrams. Looking at a selection of diagrams in the series, (Fig. 2.3), we see that we can generate every possible diagram D by starting with $\mathcal{G}_0$

and recursively joining diagrams ($D = D_1 D_2$) and placing J propagators (*rainbows*) over diagrams ($D = z^{-1}\langle JD_1J\rangle z^{-1}$). It is clear that every diagram D is of the form

$$\begin{aligned} D &= \mathcal{G}_0\langle JD_1J\rangle\mathcal{G}_0\langle JD_2J\rangle\mathcal{G}_0\cdots\langle JD_kJ\rangle\mathcal{G}_0 \\ &= \mathcal{G}_0\prod_{i=1}^{k}\langle JD_iJ\rangle\mathcal{G}_0 \end{aligned} \tag{2.14}$$

for some sequence of diagrams $D_1, \ldots, D_k$, $k \geq 0$. Since $\mathcal{G}$ is the sum of all such diagrams,

$$\mathcal{G} = \prod_{k=0}^{\infty}\sum_{D_1,\ldots D_k}\mathcal{G}_0\prod_{i=1}^{k}\langle JD_iJ\rangle\mathcal{G}_0 \tag{2.15}$$

where the sum is over all sequences of k diagrams. Now, by linearity,

$$\begin{aligned} \mathcal{G} &= \mathcal{G}_0\prod_{k=0}^{\infty}\prod_{i=1}^{k}\left\langle J\left(\sum_{D_i}D_i\right)J\right\rangle\mathcal{G}_0 \\ &= \mathcal{G}_0\prod_{k=0}^{\infty}\prod_{i=1}^{k}\langle J\mathcal{G}J\rangle\mathcal{G}_0 \end{aligned} \tag{2.16}$$

since the sum of all diagrams is just $\mathcal{G}$. Finally,

$$\begin{aligned} \mathcal{G} &= \mathcal{G}_0\prod_{k=0}^{\infty}(\langle J\mathcal{G}J\rangle\mathcal{G}_0)^k \\ &= \frac{1}{z-\langle J\mathcal{G}J\rangle} \\ &\equiv \frac{1}{z-\Sigma} \end{aligned} \tag{2.17}$$

$\Sigma \equiv \langle J\mathcal{G}J\rangle$ is known as the one-point irreducible self-energy, and has a simple form in this case:

$$\Sigma_{il} = \sum_{jk}\langle J_{ij}\mathcal{G}_{jk}J_{kl}\rangle = \frac{g^2}{N}\delta_{il}\mathrm{tr}\mathcal{G} = g^2\delta_{ij}G(z) \tag{2.18}$$

So we see that Σ is proportional to the identity. Tracing both sides of Eq. 2.17, we get an equation for $G(z)$:

$$G(z) = \frac{1}{z - g^2 G(z)} \tag{2.19}$$

This gives us the quadratic equation

$$g^2 G^2 - zG + 1 = 0 \tag{2.20}$$

with solutions

$$G(z) = \frac{1}{2g^2}\left[z \pm \sqrt{z^2 - 4g^2}\right] \tag{2.21}$$

We can determine which root to take by noting that $G(z) \to 1/z$ for large z. Since we have $G(z) \to \frac{1}{2g^2}\left[z \pm (z - \frac{2g^2}{z})\right]$, we must take the minus sign.

$$G(\lambda) = \begin{cases} \frac{1}{2g^2}\left[\lambda - \sqrt{\lambda^2 - 4g^2}\right] & \text{for } |\lambda| > 2g \\ \frac{1}{2g^2}\left[\lambda - i\sqrt{4g^2 - \lambda^2}\right] & \text{for } |\lambda| < 2g \end{cases} \tag{2.22}$$

Taking the negative imaginary part of $G(x)$ as in Eq. 2.8, we get

$$\rho(\lambda) = \begin{cases} 0 & \text{for } |\lambda| > 2g \\ \frac{1}{2\pi g^2}\sqrt{(2g)^2 - \lambda^2} & \text{for } |\lambda| < 2g \end{cases} \tag{2.23}$$

and recover Wigner's semicircle law [73]. Remarkably, the semicircle law holds for any hermitian random matrix ensemble where the $2N$ components $J_{ij}, i > j$ are independent [51, 67]. So long as the matrix elements have zero mean and variance g^2/N, the eigenvalues will obey the above distribution as $N \to \infty$. The matrix elements can be real, complex, discrete, or sparse. This property, known as universality, is the matrix analogue of the central limit theorem [66]. Other properties of the matrix distribution, such as eigenvalue

correlations and the eigenvector distribution, are also universal for large N [67, 68]. Non-hermitian matrix ensembles have also been shown to be universal [51, 10, 65, 3]. A well-known non-hermitian matrix ensemble is the so-called Ginibre ensemble, in which the matrix elements J_{ij} are i.i.d. with mean zero and variance $1/N$. In this case, the expected eigenvalue distribution for large N lies uniformly on the unit circle in the complex plane, known as the circular law [32, 33].

The advantage of the diagrammatic formalism is its extensibility beyond Gaussian ensembles with zero mean. A. Zee and coauthors applied this formalism to a myriad of generalizations, including general polynomial potentials $V(J)$ [16], ensembles with nonzero mean [15], sums of random matrices [75], and non-hermitian random matrices [27]. The diagrammatic approach can also be used to compute eigenvalue correlations [16, 15].

2.2 Random matrices with structured correlations

We now analyze the case where J is non-hermitian, with matrix elements that are correlated and not identically distributed. We allow J to have correlations between forward and reverse connections that depend on the identity of the input j and output i node. Specifically, the matix elements J_{ij} have moments

$$\begin{aligned} \langle J_{ij} \rangle &= 0 \\ \langle J_{ij}^2 \rangle &= g_{ij}^2/N \\ \langle J_{ij} J_{ji} \rangle &= \frac{1}{N} \tau_{ij} g_{ij} g_{ji} \end{aligned} \tag{2.24}$$

The gain matrix g_{ij} has positive elements. Correlation values τ_{ij} are symmetric in i, j, with $|\tau_{ij}| \leq 1$, and denote the degree of correlation between forward j, i and reverse i, j connections in the corresponding random network. However, to simplify the analysis of the eigenvalue distribution, we choose the distribution over J to be complex gaussian. Although

this may seem like a restrictive and unrealistic choice, we appeal to the universality principle of random matrix distributions. Strictly, there is at present no proof that random matrix distributions with the above moments have universal limiting eigenvalue distributions. However, we have significant numerical evidence to suggest that universality holds in our case.

We choose the complex gaussian matrix distribution to have real-valued variances

$$\langle J_{ij}\bar{J}_{ij}\rangle = \frac{1}{N}g_{ij}^2, \tag{2.25}$$

and real-valued covariances

$$\langle J_{ij}J_{ji}\rangle = \frac{1}{N}\tau_{ij}g_{ij}g_{ji}. \tag{2.26}$$

All other-second order correlations vanish. Explicitly,

$$\begin{aligned}\langle J_{ij}\bar{J}_{kl}\rangle &= \frac{1}{N}\delta_{ik}\delta_{jl}g_{ij}g_{kl}\\ \langle J_{ij}J_{kl}\rangle &= \frac{1}{N}\delta_{il}\delta_{jk}\sqrt{\tau_{ij}}\sqrt{\tau_{kl}}g_{ij}g_{kl}\end{aligned} \tag{2.27}$$

where $\sqrt{\cdot}$ denotes the principal square root.

To outline the steps of the derivation, we will first seek the expected density of eigenvalues of J for large N by writing the density in terms of the Green's function G. While G is analytic for Hermitian matrices, G is generally non-analytic for non-Hermitian matrices, so we cannot directly apply the diagrammatic method. We therefore relate G to the analytic Green's function of a Hermitian random matrix H, which we compute with standard diagrammatic techniques. We derive a set of self-consistent equations for G when the gain matrix g_{ij} is a continuous function in the limit $N \to \infty$, and when g_{ij} is block-structured. Finally, we apply our method to examples and compare the results to empirical eigenvalue distributions obtained by exact diagonalization of realizations of J.

We start by writing the expected density of eigenvalues of J in the complex plane as

$$\rho(x,y) = \left\langle \frac{1}{N}\sum_k \delta(x - \mathrm{Re}\lambda_k)\delta(y - \mathrm{Im}\lambda_k) \right\rangle. \tag{2.28}$$

where $\langle\cdot\rangle$ indicates an average over realizations of J. Defining $\partial = (\partial_x - i\partial_y)/2$ and $\bar{\partial} = (\partial_x + i\partial_y)/2$, and using the identity $\bar{\partial}\frac{1}{x+iy} = \pi\delta(x)\delta(y)$ [1], we can write the density (2.28) in terms of the Green's function

$$G(z,\bar{z}) \equiv \left\langle \frac{1}{N}\mathrm{tr}\frac{1}{z-J} \right\rangle \tag{2.29}$$

as

$$\rho(x,y) = \frac{1}{\pi}\bar{\partial}G(z,\bar{z}). \tag{2.30}$$

By way of analogy, we can think of the eigenvalues as electric charges on the plane, and the Green's function as the electric field of these charges, with the real and imaginary parts of $G(z)$ as the vector components. From this perspective, Eq. 2.30 is simply Gauss's law, and the identity $\bar{\partial}\frac{1}{x+iy} = \pi\delta(x)\delta(y)$ is Gauss's law for a point charge.

Since J is non-Hermitian, the eigenvalues of J will in general lie in some region of the complex plane. As we have seen, Girko's circular law states that if the elements of J are independently and identically distributed with variances g^2/N, then the eigenvalues lie in a disk of radius g [32]. The Green's function is therefore not in general holomorphic, and we cannot expand in powers of $1/z$ as required for the diagrammatic expansion. Following [28], we can find the Green's function by solving a related Hermitian random matrix problem,

[1] This relation follows from the solution $\bar{\partial}\partial\log z = \pi\delta(x)\delta(y)$ of Poisson's equation in two dimensions. See also [28].

to which we can apply the diagrammatic approach. Define the $2N \times 2N$ Hermitian matrix

$$H = \begin{bmatrix} 0 & J - z \\ (J-z)^\dagger & 0 \end{bmatrix}. \tag{2.31}$$

The matrix Green's function for H is

$$\mathcal{G}(\eta, z, \bar{z}) = \left\langle \frac{1}{\eta - H} \right\rangle, \tag{2.32}$$

where we think of the eigenvalues of H as lying on the complex plane η. Since H is Hermitian, these eigenvalues will lie on the real axis, and $\mathcal{G}$ is holomorphic in η except for cuts on the real axis. Once $\mathcal{G}$ is computed, we obtain the original Green's function G from $\mathcal{G}$ by extracting the lower left matrix block and taking the limit $\eta \to i0^+$:

$$\mathcal{G}(\eta = 0, z, \bar{z}) = \left\langle \begin{bmatrix} 0 & \frac{1}{(z-J)^\dagger} \\ \frac{1}{z-J} & 0 \end{bmatrix} \right\rangle, \tag{2.33}$$

yielding Eq. (2.29):

$$G(z, \bar{z}) = \frac{1}{N} \mathrm{tr} \mathcal{G}^{21}(\eta = 0, z, \bar{z}). \tag{2.34}$$

Here, $\mathcal{G}^{21}$ is the lower left block of $\mathcal{G}$. To compute $\mathcal{G}$ (2.32), we first rewrite $\eta - H = \mathcal{G}_0^{-1} - \mathcal{J}$ with

$$\mathcal{G}_0^{-1} \equiv \begin{bmatrix} \eta & z \\ \bar{z} & \eta \end{bmatrix} \quad \text{and} \quad \mathcal{J} \equiv \begin{bmatrix} 0 & J \\ J^\dagger & 0 \end{bmatrix}, \tag{2.35}$$

so that the random part $\mathcal{J}$ has zero mean. Note that $\mathcal{G}_0$ is equal to $\mathcal{G}$ with $J = 0$. We expand $\mathcal{G}$ in $\mathcal{G}_0$ as follows:

$$\mathcal{G} = \sum_{n=0}^{\infty} \mathcal{G}_0 \langle (\mathcal{J}\mathcal{G}_0)^n \rangle = \mathcal{G}_0 + \langle \mathcal{G}_0 \mathcal{J} \mathcal{G}_0 \mathcal{J} \mathcal{G}_0 \rangle + \dots \tag{2.36}$$

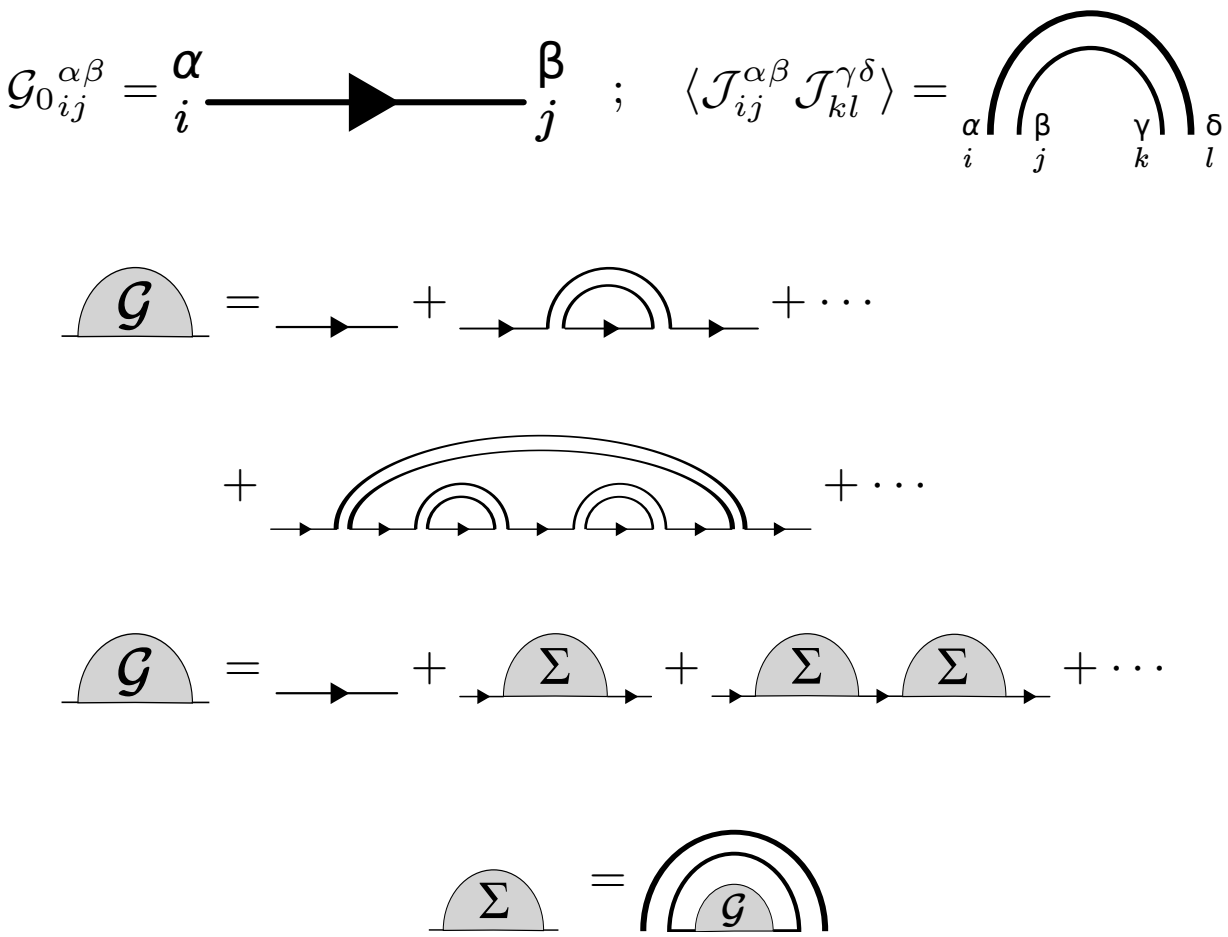


Figure 2.3: Diagrams used in the expansion of the Green's function $\mathcal{G}$ (2.36). $\mathcal{G}$ is the sum of all planar diagrams in the large N limit. $\mathcal{G}$ can be re-summed in terms of the self-energy matrix Σ. In the large N limit, Σ consists of all diagrams nested under a double line (2.38).

Here, the odd terms vanish since $\langle \mathcal{J} \rangle = 0$. Since the distribution over $\mathcal{J}$ is Gaussian, each term in the sum reduces to the Wick contraction of n factors of $\mathcal{J}$. We therefore use the diagrammatic technique [17, 16] to represent each term in the sum. We denote the N node indices by roman letters $i = 1, \ldots N$ and index the blocks by Greek letters $\alpha = 1, 2$. We represent $\mathcal{G}_0$ by a single directed line carrying one set of indices, and the correlator $\langle \mathcal{J}\mathcal{J} \rangle$ by a double line carrying two sets of indices (Fig. 2.3) [28, 38, 27]. Indices are summed at each connecting vertex. The nth term in $\mathcal{G}$ is the sum of all diagrams with n vertices. In the large N limit diagrams which have crossing lines vanish, and only "planar" diagrams

remain[2] [64, 1]. This greatly simplifies the sum, since the only allowed diagrams are nested 'rainbow diagrams' such as those depicted in Fig. 2.3. This allows us to evaluate (2.36) by performing a resummation of $\mathcal{G}$ in terms of the 'self-energy' matrix Σ:

$$\mathcal{G} = \sum_{n=0}^{\infty} \mathcal{G}_0(\Sigma\mathcal{G}_0)^n = \left(\frac{1}{\mathcal{G}_0^{-1} - \Sigma}\right). \tag{2.37}$$

In the planar limit, the self-energy matrix is

$$\Sigma = \langle \mathcal{J}\mathcal{G}\mathcal{J} \rangle, \tag{2.38}$$

encoding the nested 'rainbow' structure of the diagrams [16]. This is depicted diagrammatically in Fig. 2.3.

In block form, Eq. (2.37) is

$$\mathcal{G} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \eta - \Sigma^{11} & z - \Sigma^{12} \\ \bar{z} - \Sigma^{21} & \eta - \Sigma^{22} \end{bmatrix}^{-1}. \tag{2.39}$$

and Eq. (2.38) is

$$\Sigma = \begin{bmatrix} \Sigma^{11} & \Sigma^{12} \\ \Sigma^{21} & \Sigma^{22} \end{bmatrix} = \left\langle \begin{bmatrix} JDJ^\dagger & JCJ \\ J^\dagger B J^\dagger & J^\dagger A J \end{bmatrix} \right\rangle \tag{2.40}$$

where we have denoted the blocks of $\mathcal{G}$ as A, B, C and D. Substituting (2.40) into (2.39) will give us self-consistent equations for the blocks of $\mathcal{G}$.

Equations (2.39) and (2.40) describe the eigenvalue distribution in the general case, with or without correlations.

[2]It is worth noting that since the correlators (2.25, 2.26) are not proportional to the identity as in the i.i.d. case, loops produce a weighted trace weighted by elements of g_{ij}. However, assuming all elements of g_{ij} are of $O(1)$, the weighted trace is of $O(N)$ as in the i.i.d. case.

2.2.1 No correlations

Before analyzing the impact of correlations on the eigenvalue distribution, we first check that this result reproduces previous results obtained in the absence of correlations. When elements of J are independently distributed, the covariances (2.26) vanish. In this case we find[3]

$$\begin{aligned}\Sigma^{11}_{il} &= \sum_{j,k}\langle J_{ij}D_{jk}J^\dagger_{kl}\rangle = \frac{1}{N}\delta_{il}\sum_j g_{ij}g_{lj}D_{jj},\\ \Sigma^{22}_{il} &= \sum_{j,k}\langle J^\dagger_{ij}A_{jk}J_{kl}\rangle = \frac{1}{N}\delta_{il}\sum_j g_{ji}g_{jl}A_{jj},\end{aligned} \tag{2.41}$$

and $\Sigma^{12} = \Sigma^{21} = 0$. This means that the matrix Σ is diagonal. Then, since each block on the RHS of Eq. (2.39) is diagonal, each block of $\mathcal{G}$ is also diagonal. Inverting the RHS and equating matrix elements yields

$$A_{ii} = \frac{\eta - \frac{1}{N}\sum_j A_{jj}g^2_{ji}}{q_i(\eta,|z|)}, \quad D_{ii} = \frac{\eta - \frac{1}{N}\sum_j g^2_{ij}D_{jj}}{q_i(\eta,|z|)}, \tag{2.42}$$

$$C_{ii} = \bar{z}/q_i(\eta,|z|), \tag{2.43}$$

where

$$q_i(\eta,|z|) = (\eta - \frac{1}{N}\sum_j A_{jj}g^2_{ji})(\eta - \frac{1}{N}\sum_j g^2_{ij}D_{jj}) - |z|^2. \tag{2.44}$$

Writing out the blocks of $\mathcal{G}$ in Eq. (2.32),

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \left\langle \begin{bmatrix} \frac{\eta}{\eta^2-(J-z)(J-z)^\dagger} & \frac{J-z}{\eta^2-(J-z)^\dagger(J-z)} \\ \frac{(J-z)^\dagger}{\eta^2-(J-z)(J-z)^\dagger} & \frac{\eta}{\eta^2-(J-z)^\dagger(J-z)} \end{bmatrix} \right\rangle, \tag{2.45}$$

[3]We furthermore demand that g_{ij} converges to a uniformly bounded continuous function $g(i/N, j/N)$ on the unit square as $N \to \infty$, excepting discontinuities on a set of measure zero (see [4]).

and rewriting $\eta = i\epsilon$, with $\epsilon > 0$, we see that blocks A and D are positive definite matrices multiplied by $-i$. We therefore define $a_j \equiv iA_{jj}$ and $d_j \equiv iD_{jj}$, where a_i and d_i are positive real numbers. We also define $c_j = C_{jj}$. This allows us to rewrite (2.42) and (2.43) as

$$a_i = \hat{a}_i/q_i, \quad d_i = \hat{d}_i/q_i, \quad c_i = \bar{z}/q_i(\epsilon, |z|) \tag{2.46}$$

with $q_i(\epsilon, r) \equiv -q_i(\eta, |z|) = \hat{a}_i\hat{d}_i + r^2$ and

$$\hat{a}_i \equiv \epsilon + \frac{1}{N}\sum_j a_j g_{ji}^2, \quad \hat{d}_i \equiv \epsilon + \frac{1}{N}\sum_j g_{ij}^2 d_j, \tag{2.47}$$

where $r = |z|$. We now have a set of $2N$ self-consistent equations (2.46) for the elements a_i and d_i of the Green's function $\mathcal{G}$. These can be solved numerically with $\epsilon = 0$ (or ϵ set to a small value if many elements g_{ij} are also small). Once the a_i and d_i are found, the c_i can be computed and used to find the original Green's function G with Eq. (2.34), since the trace of $\mathcal{G}^{21} \equiv C$ is the sum of the coefficients $c_i \equiv C_{ii}$. Note that since $c_j = re^{-i\theta}/q_j(\epsilon, r)$ in polar coordinates, $|c_j|$ depends only on r. This allows us to rewrite Eq. (2.30) as a function of r only:

$$\rho(r) = \frac{1}{2\pi N}\sum_j \left(\frac{\partial |c_j|}{\partial r} + \frac{|c_j|}{r}\right). \tag{2.48}$$

The resulting eigenvalue distribution has support on the disk with radius $r = \sqrt{\lambda_1(K)}$, where $\lambda_1(K)$ is the largest eigenvalue of the matrix $K_{ij} \equiv g_{ij}^2/N$ (see Appendix A). This result was previously presented in [3] and [4], and a similar argument was used in [1] for the case of matrices with non-zero mean. However, previous analyses do not hold when J has covariant elements.

2.2.2 Including correlations

We now allow J to have correlated elements across its diagonal (Eq. 2.26). Then Σ^{12} and $\Sigma^{21} \neq 0$, yielding a new expression for c:

$$c_i = \hat{c}_i / q_i(\epsilon, z, \bar{z}), \quad \hat{c}_i \equiv \bar{z} - \frac{1}{N} \sum_j \tau_{ij} g_{ij} g_{ji} \bar{c}_j, \tag{2.49}$$

where now $q_i = \hat{a}_i \hat{d}_i + |\hat{c}_i|^2$, $b_i = \bar{c}_i$. The τ_{ij} denote the degree of correlation between i and j as in Eq. (2.26). In this case, the eigenvalue density has the more general form

$$\rho(x, y) = \frac{1}{\pi} \bar{\partial} G(z, \bar{z}) = \frac{1}{N\pi} \bar{\partial} \sum_{j=1}^{N} c_j(z, \bar{z}). \tag{2.50}$$

The density ρ depends on x and y in a nontrivial way, and the support of the distribution is neither circular nor elliptical. The boundary of the eigenvalue distribution now satisfies (see Appendix A for a derivation):

$$\lambda_1(K(z)) = 1, \quad K_{ij}(z) = \frac{1}{N} |c_i(z)|^2 g_{ij}^2, \tag{2.51}$$

where the complex-valued $c_i(z)$ are now given by the self-consistent equations

$$c_i(z) = (z - \sum_j \tau_{ij} g_{ij} g_{ji} c_j(z))^{-1}. \tag{2.52}$$

Now, to obtain the boundary, it is necessary to simultaneously solve (2.51) and (2.52) for each boundary point. For example, we can set $z = re^{i\theta}$ and solve the above for r for each θ. Note that these expressions reduce to the circularly symmetric case when $\tau_{ij} = 0$.

2.2.3 Block-structured correlations

We now consider the special case for which the gain matrix g_{ij} is block structured. Block structured matrices describe networks with nodes partitioned into subgroups, for example neural networks with cell-type-specific connectivity [5], or networks of ecological communities [54]. Suppose the nodes of the network are grouped into M populations of size $f_m N$, for $m = 1 \dots M$ and that J is block structured so that the gain $g^2_{m_i n_j} = g^2_{mn}$ and correlations $\tau_{m_i n_j} = \tau_{mn}$ depend only on the population indices m and n of the output and input nodes i and j, respectively. This allows us to sum (2.46) and (2.49) over each population. Let $N_m \equiv N \sum_{n=1}^{m} f_n$. Then define[4]

$$a_m \equiv \frac{1}{N f_m} \sum_{i=N_m+1}^{N_m} a_i, \tag{2.53}$$

and define c_m and d_m similarly. Then $q_m \equiv q_i$ depends only on the population index, and now we have

$$a_m = \hat{a}_m / q_m, \quad d_m = \hat{d}_m / q_m, \quad c_m = \hat{c}_m / q_m, \tag{2.54}$$

and $q_m = \hat{a}_m \hat{d}_m + |\hat{c}_m|^2$, with

$$\hat{a}_m = \epsilon + \sum_{n=1}^{M} f_n a_n g^2_{nm}, \quad \hat{d}_m = \epsilon + \sum_{n=1}^{M} g^2_{mn} f_n d_n, \tag{2.55}$$

$$\hat{c}_m = \bar{z} - \sum_{n=1}^{M} \tau_{mn} g_{mn} g_{nm} f_n \bar{c}_n. \tag{2.56}$$

[4]These sums converge, since the Green's function for H, $\frac{1}{N}\mathrm{tr}\mathcal{G} = \sum_i (A_{ii} + D_{ii})$ and Eq. (2.50) must converge

Now the dependence on N is removed, and we need only solve $3M$ self-consistent equations. The eigenvalue density is now

$$\rho(x,y) = \frac{1}{\pi}\bar{\partial}\sum_m f_m c_m(z,\bar{z}). \tag{2.57}$$

The boundary of the distribution satisfies the M equations

$$\lambda_1(K(z)) = 1, \quad K_{mn}(z) = |c_m(z)|^2 g_{mn}^2 f_n \tag{2.58}$$

and

$$c_m(z) = (z - \sum_n \tau_{mn} g_{mn} g_{nm} f_n c_n(z))^{-1} \tag{2.59}$$

similar to (2.51), (2.52). When $\tau_{mn} = 0$, the distribution has boundary $|z| = \sqrt{\lambda_1(K)}$, where $\lambda_1(K)$ is the largest eigenvalue of the matrix $K_{mn} \equiv g_{mn}^2 f_n$ [5].

2.2.4 Simple examples

We start by deriving Girko's circular law in the case where J is homogeneous and i.i.d. In this case self-consistent equations are simply

$$a = \frac{1}{q}(\epsilon + g^2 a), \quad d = \frac{1}{q}(\epsilon + g^2 d) \tag{2.60}$$

$$q = (\epsilon + g^2 a)(\epsilon + g^2 d) + r^2 \tag{2.61}$$

By symmetry, $a = d$. Furthermore, setting $\epsilon = 0$ yields the cubic equation

$$a(a^2 + r^2) = g^2 a \tag{2.62}$$

which has solutions $a = 0$ and $g^4a^2 + r^2 = g^2$. We also have

$$c = \frac{\bar{z}}{g^4a^2 + r^2} = \begin{cases} g^{-2}\bar{z} & a > 0 \\ z^{-1} & a = 0 \end{cases} \tag{2.63}$$

$\rho(z, \bar{z}) \neq 0$ precisely where $\bar{\partial}c \neq 0$, so we see that the $a > 0$ solution holds on the support of the eigenvalue distribution, and the $a = 0$ solution holds outside the support of the distribution. Further, the Green's function $G = c = 1/z$ for large z. These two solutions match when $r = g$ (both a and c are continuous). This means that the boundary of the eigenvalue distribution must be circular with radius g. Finally, from (2.48), we have

$$\rho = \frac{1}{\pi}\bar{\partial}c = \begin{cases} \frac{1}{\pi g^2} & r > g \\ 0 & r < g \end{cases} \tag{2.64}$$

which is Girko's circular law.

Next, we compute the density and boundary of the eigenvalue distribution for a homogeneous matrix with correlation $\tau \neq 0$. Again, $a = d$ by symmetry, but a now has solutions $a = 0$ and $g^4a^2 + |\hat{c}|^2 = g^2$. The corresponding solutions for c are

$$c = \frac{\bar{z} - \tau g^2 \bar{c}}{g^4a^2 + |\hat{c}|^2} = \begin{cases} \frac{\bar{z} - \tau z}{g^2(1 - |\tau|^2)} & a > 0 \\ \frac{z \pm (z^2 - 4g^2\tau)^{1/2}}{2g^2\tau} & a = 0 \end{cases} \tag{2.65}$$

The boundary of the eigenvalue distribution satisfies (2.51). We can verify that the spectrum is elliptical by setting $z = z_{el} \equiv (1 + \tau)\cos\theta + (1 - \tau)\sin\theta$. Since the solutions for c match at the boundary, we can substitute this value into the $a > 0$ solution[5] to yield $c(z_{el}, \bar{z_{el}})$

[5] To evaluate $c(z_{el})$ for $a = 0$, we must choose the square root so that the boundary condition $c(z) \to 1/z$ is satisfied for $z \to \infty$. This is necessary in the more complex examples explored in 2.3.

distribution is uniform inside the ellipse.

Another simple example is the case where g_{mn} and τ_{mn} are block diagonal. In this case the eigenvalue distribution is simply a weighted sum over the eigenvalue distributions of each block. This is the simplest example of a matrix that need not have a circular or ellipsoidal eigenvalue distribution. For example, if we have two diagonal blocks with correlations of opposite signs, the positively correlated block has a spectrum supported on an ellipse stretched across the real axis, whereas the negatively correlated block will has a spectrum supported on an ellipse stretched across the imaginary axis. The combined distibution has support on the union of these two ellipses.

2.2.5 Numerical tests

To verify our results numerically, we consider a network with $M = 3$ populations, with relative population sizes $f = (1/6, 1/3, 1/2)$, and

$$g_{mn}^2 = \begin{bmatrix} .54 & .83 & .65 \\ .95 & .46 & .01 \\ .72 & .59 & .55 \end{bmatrix}, \quad \tau_{mn} = \begin{bmatrix} .5 & -.2 & .9 \\ -.2 & .3 & .1 \\ .9 & .1 & -.6 \end{bmatrix}. \tag{2.66}$$

We iteratively solved the self-consistent Eqs. (2.54) for a grid of points on the complex plane and approximated the eigenvalue distribution using finite differences, shown in Fig. 2.4(a). We compare this distribution with eigenvalue histograms generated by exact diagonalization of 1000 realizations of J. We find that realizations of J with complex elements agree with our result (Fig. 2.4(c,e,f)). Removing the correlations (2.26) from realizations of J yields a circular distribution (Fig. 2.4(b)). Notably, we find that including these correlations distorts the eigenvalue distribution in a nontrivial way: the distribution is in general neither a circle nor an ellipse. Furthermore, we find using Eqs. (A.4),(A.2) that the rightmost eigenvalue of the distribution has moved from ~ 0.713 to ~ 0.890, so that the corresponding

linear system (1.2) becomes more unstable.

For any finite N, J has non-universal features that disappear as $N \to \infty$ (Fig. 2.7). In particular, the matrix J with real elements will have a higher density of eigenvalues on the real axis (Fig. 2.4(d)). However, we find that the proportion of eigenvalues on the real axis drops off as $1/\sqrt{N}$ (Fig. 2.6), as anticipated for large N [26].

To demonstrate that our technique applies to situations where the variance and covariance depend continuously on the node indices i, j, we consider a neural network inspired by connectivity around pinwheels in the visual cortex [50, 74]. The neurons are arranged on a square grid on the unit square and assigned orientations based on their position, shown in Fig. 2.5(a). For neurons i and j with positions $\mathbf{r}_i$ and $\mathbf{r}_j$, the gain is

$$g_{ij} = g_0 \exp\left[-|\mathbf{r}_i - \mathbf{r}_j|^2/w_r^2 - \Delta\theta^2(\mathbf{r}_i, \mathbf{r}_j)/w_\theta^2\right], \tag{2.67}$$

where $\Delta\theta(\mathbf{r}_i, \mathbf{r}_j)$ denotes the difference in orientation of neurons at $\mathbf{r}_i$ and $\mathbf{r}_j$. We choose the covariance to be proportional to the gain: $\tau_{ij} = \tau_0 g_{ij}$. In this example, $w_r = 0.2$, $w_\theta = 20°$, $g_0 = 1$, and $\tau_0 = 0.8$. The gain matrix for a grid of 16×16 neuron populations is shown in Fig. 2.5(b). This grid size requires us to solve $N = 256$ self-consistent equations to determine the eigenvalue density. For comparison, we generated 1000 realizations of J with $N = 2048$; to mitigate finite-N effects [4], we used block structured matrices with 16×16 populations, with 8 nodes in each population. We find that our result closely matches the empirical distribution (Fig. 2.5(c-e)). Increasing the grid size to 32×32 and 64×64 did not appreciably change the resulting eigenvalue distribution, indicating that the current resolution is sufficient. Finally, using Eqs. (2.51-2.52), we find that including correlations moves the rightmost eigenvalue from 0.24 to 0.41, decreasing the stability of the system.

In conclusion, we have adapted the diagrammatic technique to study correlated connectivity matrices that are not independently or identically distributed, and relevant to

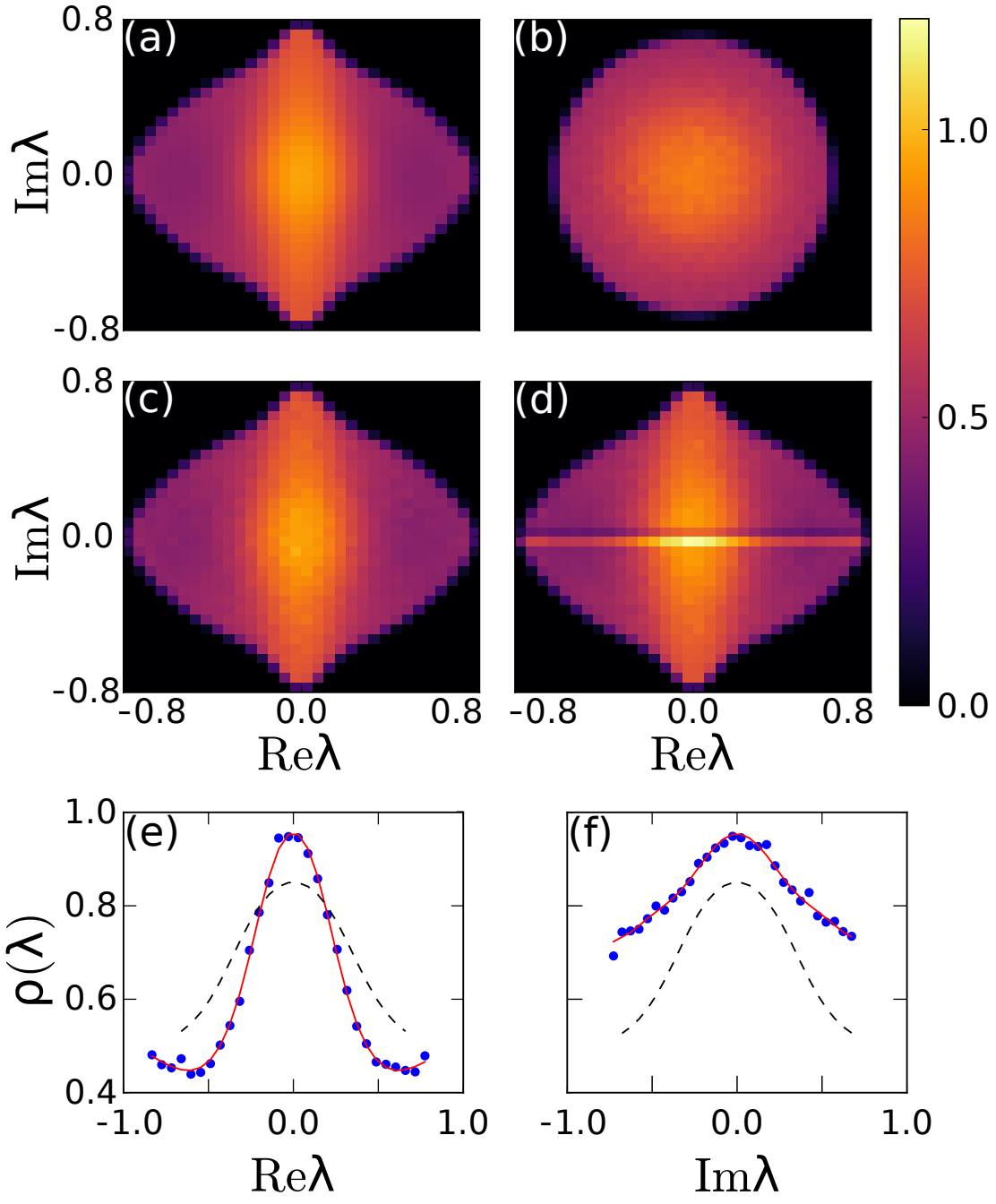


Figure 2.4: Eigenvalue density for block structured J with gain and covariance given by (2.66). (a) Eigenvalue density calculated from self-consistent equations (2.54). (b) Empirical histogram of eigenvalues from exact diagonalization of realizations of J with independent elements. The empirical histogram for J with covariance is shown with complex (c) and real (d) entries. (e,f) Cross sections of the density along the real (e) and imaginary (f) axes, showing the theoretical result (solid red line), the complex-valued empirical result (blue dots), and the distribution with no covariance (dashed curve).

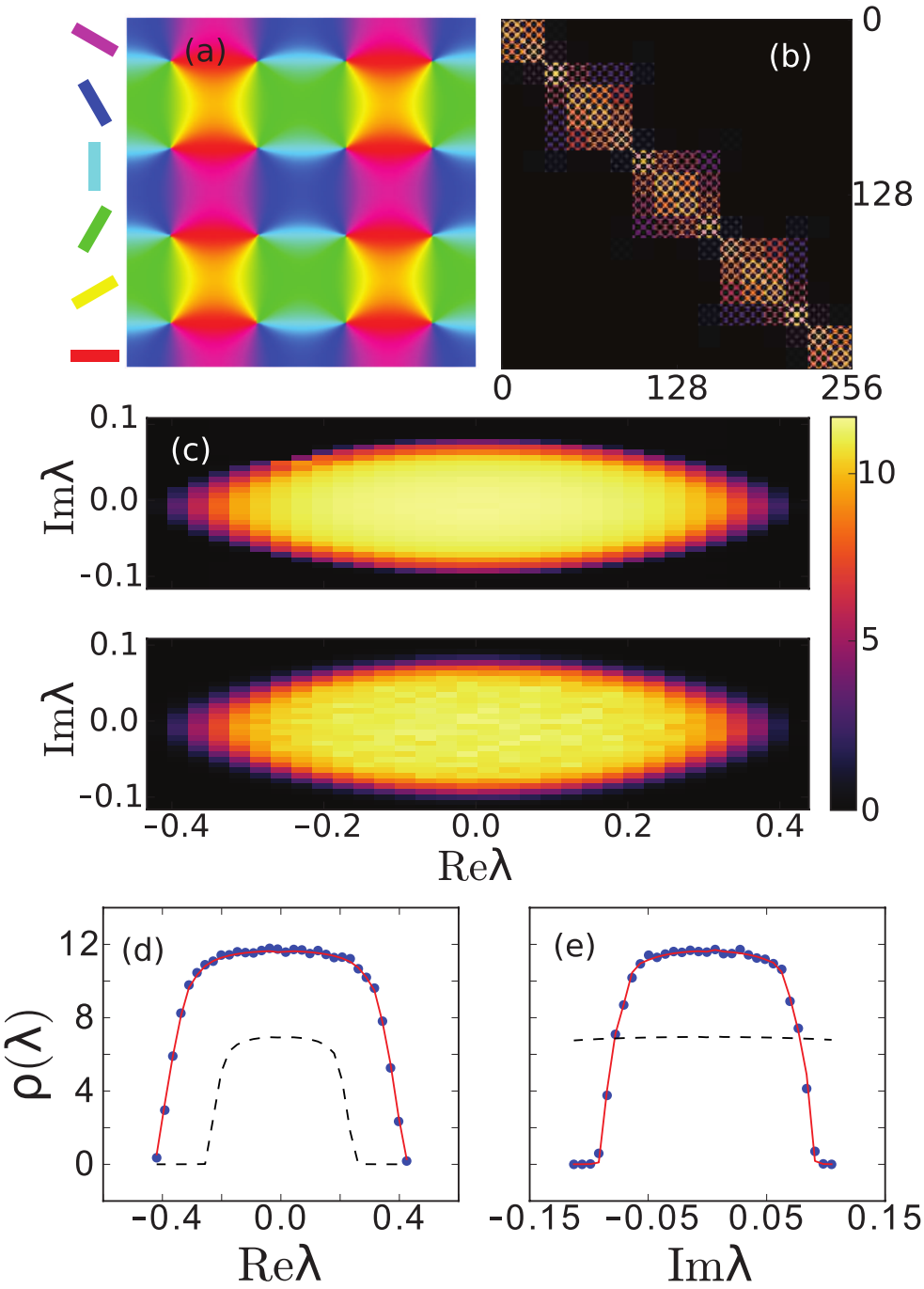


Figure 2.5: Analysis of eigenvalue distribution with continuously varying gain (2.67). (a) Orientation map of neurons. (b) Gain matrix g_{ij}. (c) Eigenvalue density calculated from self-consistent equations (top) and from realizations of J (bottom). Density cross sections along the real (d) and imaginary (e) axes, plotted as in Fig. 2.4(e).

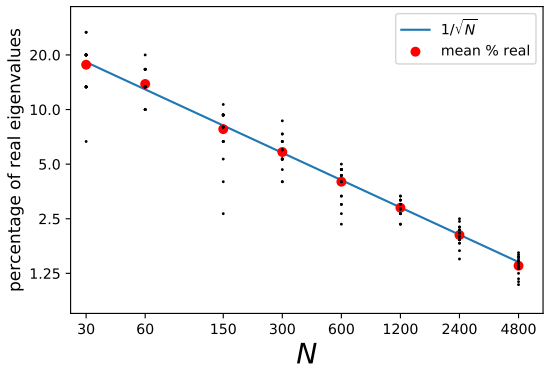


Figure 2.6: For real-valued matrices, the percentage of real eigenvalues drops as $\mathcal{O}(1/\sqrt{N})$. We plot these percentages (black) as a function of N for different realizations of the block structured matrix (2.66). Averages for a given size N are shown in red. We have plotted $1/\sqrt{N}$ in blue for comparison.

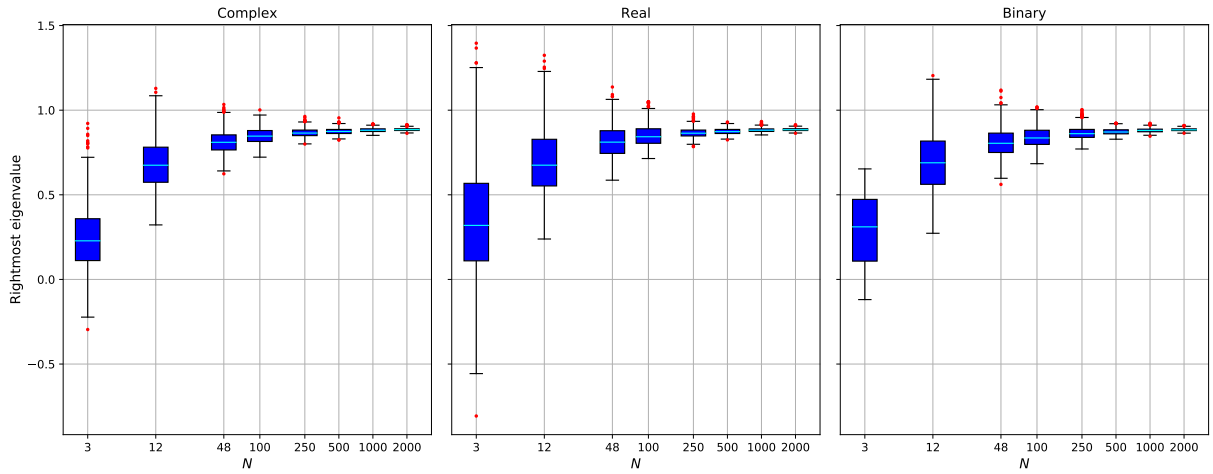


Figure 2.7: Distribution of rightmost eigenvalues for realizations of the block stuctured matrix (2.66), with complex, real, and binary matrix elements. Fluctuations of the rightmost eigenvalue decrease as N increases, and approach the same value for large N, evidence of universality.

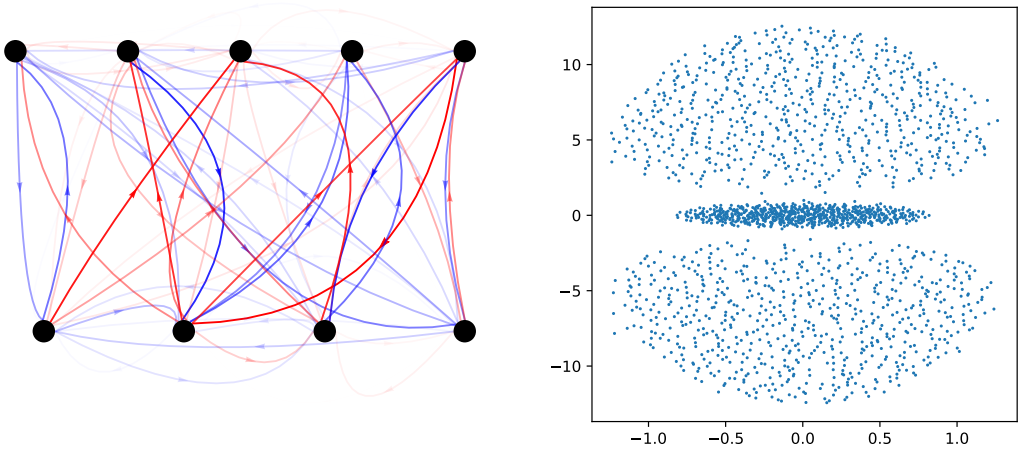


Figure 2.8: Left: Sketch of two-layer network (2.68). Right: Eigenvalue distribution for $N = 2000, g = 10, \beta = 0.1, \tau = -0.8, f_1 = 1/3$

biological circuits. The results indicate that the presence of correlations can dramatically influence the network stability and dynamics. The correlation structure is determined by plasticity rules, which act locally on connections between nodes [31, 46]. The presented analytical framework therefore makes it possible to evaluate the impact of local plasticity rules on global network activity.

2.3 Examples

We can gain further insight into the variety and unique features that block-structured correlations permit by considering simple examples. Of particular interest are layered networks, with strong correlations between layers. We will first analyze a two layer network in detail.

2.3.1 Two layer network

Consider a network of two populations, with population sizes $f_1 N$ and $f_2 N$, where the gain matrix g_{mn} and correlation coefficient matrix τ_{mn} are

$$g_{mn} = g \begin{bmatrix} \beta & 1 \\ 1 & \beta \end{bmatrix}, \quad \tau_{mn} = \begin{bmatrix} 0 & \tau \\ \tau & 0 \end{bmatrix} \tag{2.68}$$

We assume $\tau \neq 0$. When $\beta = 0$ the network is bipartite. Small values of β introduce weak, uncorrelated intra-layer connections (Fig. 2.8, left). Under certain choices of β and τ, (in particular when β is small and τ is close to 1), the eigenvalue distribution of J has three disconnected components (Fig. 2.8, right). This behavior is especially significant for $\tau < 0$: it is possible for the rightmost eigenvalue of the distribution to have a large imaginary component, leading to strongly oscillatory behavior in the network (see Fig. 3.3). This is in stark contrast to block random matrices without covariance, which are always circularly symmetric, and therefore have rightmost eigenvalues with small imaginary part.

This class of networks also has tractable implicit analytic expressions for the boundary of the eigenvalue distribution, given by the matrix K and complex coefficients c_m defined in Eqs. 2.58 and 2.59, respectively. Substituting Eq. 2.68 into Eq. 2.59, we get the quadratic system of equations

$$\begin{aligned} (z - \tau f_2 g^2 c_2(z)) c_1(z) &= 1 \\ (z - \tau f_1 g^2 c_1(z)) c_2(z) &= 1 \end{aligned} \tag{2.69}$$

Since g and τ amount to overall scalings and rotations of the Green's function (which will become clear below), we work with the rescaled variables $\zeta = \tau^{-1/2} g^{-1} z$ and $\chi_m(\zeta) = \tau^{1/2} g c_m(z)$, so that $c_m(z) = \tau^{-1/2} g^{-1} \chi_m(\tau^{-1/2} g^{-1} z)$. Then our system of equations reduces

to

$$\begin{aligned}(\zeta - f_2\chi_2(\zeta))\chi_1(\zeta) &= 1 \\ (\zeta - f_1\chi_1(\zeta))\chi_2(\zeta) &= 1\end{aligned} \tag{2.70}$$

with two solutions

$$\begin{aligned}\chi_1(\zeta) &= \frac{1}{2f_1\zeta}\left((f_1 - f_2) + \zeta^2 \pm \sqrt{\Delta(\zeta)}\right) \\ \chi_2(\zeta) &= \frac{1}{2f_2\zeta}\left((f_2 - f_1) + \zeta^2 \pm \sqrt{\Delta(\zeta)}\right)\end{aligned} \tag{2.71}$$

where the discriminant

$$\Delta(\zeta) = (\zeta^2 - 1)^2 + (f_1 - f_2)^2 - 1, \tag{2.72}$$

plotted in the top panel of Fig. 2.9. The principal square root in Eq. 2.71 has four branch cuts, shown in the middle panel of Fig. 2.9. Two of these cuts lie along the real axis, from the 'inside' branch points $\pm r_{in}$ to the 'outside' branch points $\pm r_{out}$, where

$$\begin{aligned}r_{in} &= |f_1^{1/2} - f_2^{1/2}| \\ r_{out} &= |f_1^{1/2} + f_2^{1/2}|.\end{aligned} \tag{2.73}$$

The two remaining cuts lie along the hyperbola $(\mathrm{Re}\zeta)^2 - (\mathrm{Im}\zeta)^2 = 1$. We must choose a branch of $\sqrt{\Delta(\zeta)}$ that satisfies the boundary condition $G(z) = f_1 c_1(z) + f_2 c_2(z) \to 1/z$ as $z \to \infty$. Furthermore, recall that in using Eq. 2.59, our expression for the Green's function is only accurate for z outside the eigenvalue distribution. Since $\rho(z, \bar{z}) \propto \bar{\partial} G(z, \bar{z}) = 0$ (Eq. 2.30), G must be analytic outside the eigenvalue distribution. We therefore must choose

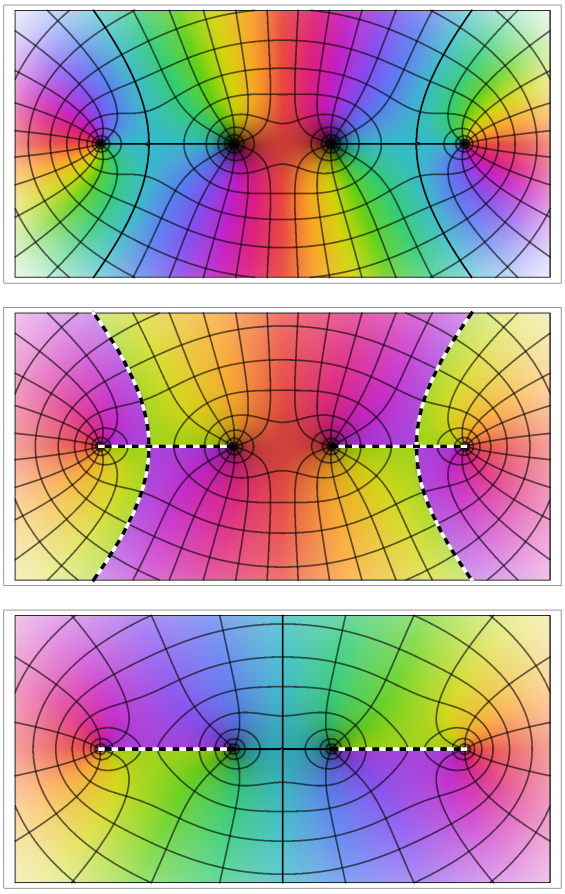


Figure 2.9: Plots of the discriminant $\Delta(z)$ (top), the principal root $\sqrt{\Delta(z)}$ (middle), and root with compact branch cuts (bottom) over the complex plane. Branch cuts are indicated by dashed lines. Constant lines of complex modulus and argument are indicated by black contour lines. The complex argument is indicated by hue.

our branch cuts to be compact. First, we factor $\Delta(\zeta)$ into its root pairs r_{in} and r_{out}:

$$\Delta(\zeta) = (\zeta^2 - r_{in}^2)(\zeta^2 - r_{out}^2) \tag{2.74}$$

By pulling factors out of the square root, we can eliminate the hyperbolic branch cuts:

$$\sqrt{\Delta(\zeta)} = (\zeta^2 - r_{in}^2)\sqrt{\frac{\zeta^2 - r_{out}^2}{\zeta^2 - r_{in}^2}} \tag{2.75}$$

where again the square root indicates the principal square root. Finally, in order to satisfy the $z \to \infty$ boundary condition on G, we choose the negative root in (2.71).

Let us now turn our attention to the eigenvalue equation (2.58). The matrix K has the form

$$K(z) = \begin{bmatrix} |c_1|^2 f_1 \beta & |c_1|^2 f_2 \\ |c_2|^2 f_1 & |c_2|^2 f_2 \beta \end{bmatrix} = \frac{1}{|\tau|} \begin{bmatrix} |\chi_1(\zeta)|^2 f_1 \beta^2 & |\chi_1(\zeta)|^2 f_2 \\ |\chi_2(\zeta)|^2 f_1 & |\chi_2(\zeta)|^2 f_2 \beta^2 \end{bmatrix} \tag{2.76}$$

The Perron-Frobenius eigenvalue of K is then given by

$$\lambda_1(K) = \frac{1}{2}\beta^2(\gamma_1 + \gamma_2) + \frac{1}{2}\sqrt{\beta^4(\gamma_1 - \gamma_2)^2 + 4\gamma_1\gamma_2} \tag{2.77}$$

where $\gamma_m \equiv |\tau|^{-1} f_m |\chi_m(\zeta)|^2$. Requiring that $\lambda_1(K) = 1$ yields the boundary condition[6]

$$\beta^2(\gamma_1 + \gamma_2) + (1 - \beta^4)\gamma_1\gamma_2 = 1 \tag{2.78}$$

Note that this depends only on the sum and product of γ_1 and γ_2. When $\beta = 0$, this reduces to $\gamma_1\gamma_2 = 1$, and when $\beta = 1$, to $\gamma_1 + \gamma_2 = 1$. Finally, from (2.71), we get somewhat

[6] For an arbitary 2×2 g_{mn}, the boundary satisfies $g_{11}^2\gamma_1 + g_{22}^2\gamma_2 + \det[g_{mn}^2]\gamma_1\gamma_2 = 1$

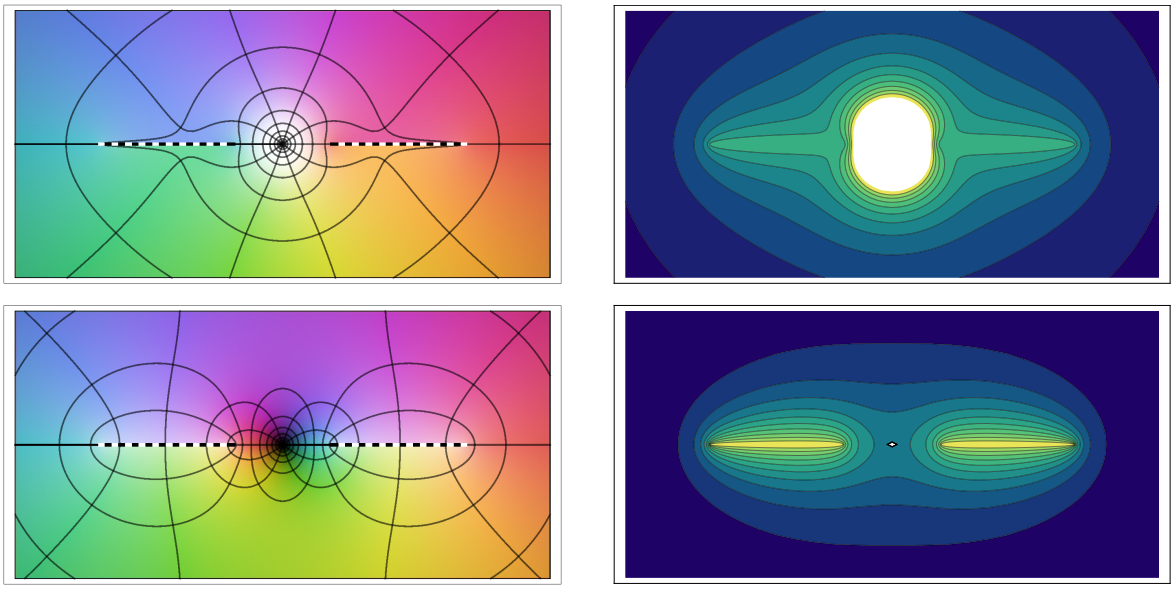


Figure 2.10: Functions determining eigenvalue support of two-layer network. Left: Plots of χ_1 (top) and χ_2 (bottom) in the complex plane. Right: Contour plots of the real-valued functions $\gamma_1 + \gamma_2$ (top), and $\gamma_1\gamma_2$ (bottom). The white regions indicate large values near the pole at $z = 0$.

compact expressions for the sum and product of γ_1 and γ_2:

$$\begin{aligned} \gamma_1 + \gamma_2 &= \frac{1}{4|\tau|f_1 f_2|\zeta|^2}\left[(f_1 - f_2)^2(1 - 2\mathrm{Re}[q(\zeta)]) + |q(\zeta)|^2\right] \\ \gamma_1\gamma_2 &= \left|\frac{(f_1 - f_2)^2 - q^2(\zeta)}{4\tau\sqrt{f_1 f_2}\zeta^2}\right|^2 \end{aligned} \tag{2.79}$$

where we have defined $q(\zeta) = \zeta^2 - \sqrt{\Delta(\zeta)}$, with the choice of square root as above. These functions are shown in Fig 2.10. Depending on the values of β^2, τ and f_1, these yield different boundaries for the eigenvalue distribution[7] (Fig. 2.11). Eq. (2.78) either defines three contours, one around each branch cut and one around the pole at the origin, or one contour encompassing all three non-holomorphic features. When $f_1 = f_2$, the distribution is an ellipse.

We can exactly compute the Green's function $G = f_1 c_1 + f_2 c_2$ outside the eigenvalue

[7]The parameter g trivially scales the distribution.

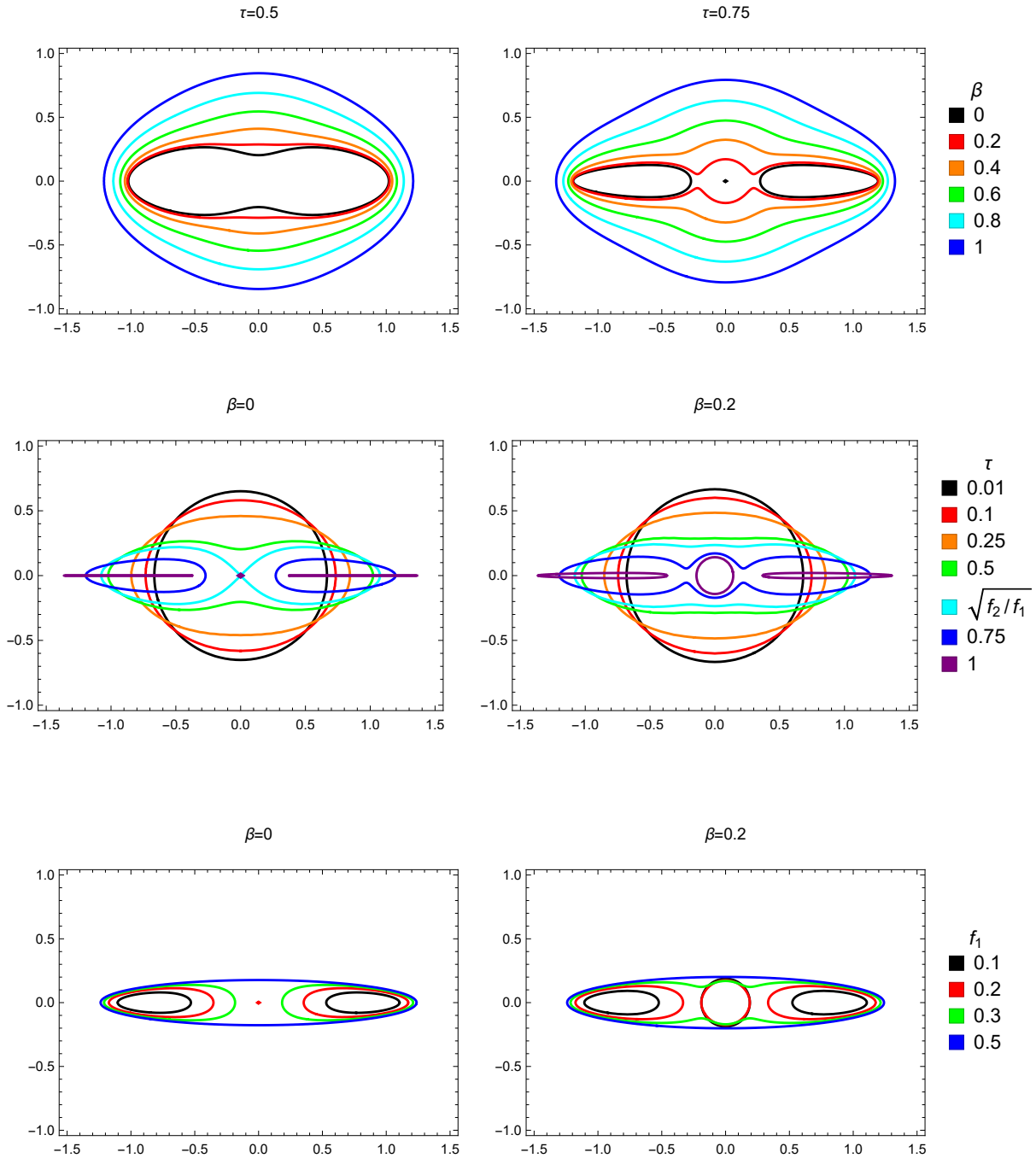


Figure 2.11: Support of eigenvalue distribution for the two-layer network, as a function of β (top), τ (middle), and f_1 (bottom). The top and middle figures have $f_1 = 0.75$, and the bottom figures have $\tau = 0.75$

distribution, yielding

$$\begin{aligned}\tau^{1/2} g G(\tau^{1/2} g \zeta) &= f_1 \chi_1(\zeta) + f_2 \chi_2(\zeta) \\ &= \zeta - \frac{1}{\zeta}(\zeta^2 - r_{in}^2)\sqrt{\frac{\zeta^2 - r_{out}^2}{\zeta^2 - r_{in}^2}}\end{aligned} \tag{2.80}$$

Changing the sign of τ swaps the real and imaginary axes (and allowing τ to have an arbitrary phase rotates the distribution) .

In the case where the eigenvalue distribution has disconnected components, we can determine the fraction of eigenvalues in each piece. This is given by

$$\int_R dx dy \rho(x, y) = \frac{1}{2\pi i} \oint_{\partial R} dz G(z) \tag{2.81}$$

where the region R contains the component of interest, and the boundary ∂R is outside the support of the eigenvalue distribution, where G is holomorphic. It is easiest to find the fraction of eigenvalues in the center component. Computing the residue around the simple pole at $z = 0$ yields the fraction $|f_1 - f_2|$ for the central component. The other two components therefore each contain a fraction $\frac{1}{2}\min(f_1, f_2)$ of the eigenvalues.

For the special case $\beta = 0$, when the network is bipartite, we can make more definitive statements. In this case, the matrix J is not full rank. With probability one, the first $f_1 N$ columns of J have rank $f_2 N$, and the remaining $f_2 N$ columns of J also have rank $f_2 N$. Thus J has only rank $2 f_2 N$, so it must have $N - 2 f_2 N = (f_1 - f_2)N$ zero eigenvalues. The eigenvalue distribution therefore has a delta function at $z = 0$ containing a fraction $|f_1 - f_2|$ of the eigenvalues, consistent with our result above.

We can also write down a condition for whether or not the distribution has disconnected components. First, as noted before, the boundary in this case is simply $\gamma_1 \gamma_2 = 1$.

For $\zeta \neq 0$, this simplifies to

$$|1 - q(\zeta)| = 2|\tau|\sqrt{f_1 f_2} \tag{2.82}$$

Expanding the holomorphic function on the LHS around $\zeta = 0$:

$$1 - q(\zeta) \approx (|1 - |f_1 - f_2|) \left(1 + \frac{\zeta^2}{|f_1 - f_2|}\right) + \mathcal{O}(\zeta^4) \tag{2.83}$$

We see that this function is a saddle point near $\zeta = 0$. If the boundary condition (2.82) is satisfied at $\zeta = 0$, then the boundary will form an 'X' near $\zeta = 0$; the eigenvalue distribution will be at the cusp of splitting into disconnected pieces (see Fig. 2.11). Setting $z = 0$ in (2.82) yields the condition

$$|\tau|^2 = \min\left(\frac{f_1}{f_2}, \frac{f_2}{f_1}\right) \tag{2.84}$$

2.3.2 Three layer network

Now consider a three layer network with

$$g_{mn} = \begin{bmatrix} g_d & g_{fb} & 0 \\ g_{ff} & g_d & g_{fb} \\ 0 & g_{ff} & g_d \end{bmatrix} \quad \text{and} \quad \tau_{mn} = \begin{bmatrix} 0 & \tau & 0 \\ \tau & 0 & \tau \\ 0 & \tau & 0 \end{bmatrix} \tag{2.85}$$

and populations Nf_1, Nf_2, Nf_3. Transforming $\zeta = (\tau g_{fb} g_{ff})^{-1/2} z$, $\chi_m(\zeta) = (\tau g_{fb} g_{ff})^{1/2} c_m(z)$, and solving for $\chi_m(\zeta)$, we get

$$\begin{aligned} \chi_1(\zeta) &= \frac{1}{2(f_1 + f_3)\zeta}\left((f_1 + f_3 - f_2) + \zeta^2 \pm \sqrt{\Delta(\zeta)}\right) \\ \chi_2(\zeta) &= \frac{1}{2f_2\zeta}\left(-(f_1 + f_3 - f_2) + \zeta^2 \pm \sqrt{\Delta(\zeta)}\right) \\ \chi_3(\zeta) &= \chi_1(\zeta) \end{aligned} \tag{2.86}$$

where

$$\Delta(\zeta) = (\zeta^2 - 1)^2 + (f_1 + f_3 - f_2)^2 - 1. \tag{2.87}$$

χ_1 and χ_2 have the same forms as 2.71, with $f_1 \to f_1 + f_3$. This means that the boundary of the eigenvalue distribution for a symmetric 3-layer can be reduced to that of a 2-layer network. More complex distributions (such distributions with 5 disconnected pieces) can be obtained by introducing asymmetry between the connections $1 \leftrightarrow 2$ and $2 \leftrightarrow 3$.

2.3.3 Four layer network

Next, we consider a symmetric four layer network, with g_{mn} and τ_{mn} defined analogously to above, and $f_m = (f_1, f_2, f_2, f_1)$. The c_m now satisfy cubic polynomials. When $z \neq 0$, $\chi_1 = \chi_4$ and $\chi_2 = \chi_3$. The discriminant $\Delta(\zeta)$ is[8]

$$\Delta(\zeta) = \zeta^4 \left[4(f_1 - f_2)^3 - (8f_1 + 20f_1f_2 - {f_2}^2)\zeta^2 + 4f_1\zeta^4\right] \tag{2.88}$$

The roots of this polynomial tell us the endpoints of the branch cuts in the expression for χ_m. These roots are shown in Fig. 2.12a as a function of f_1. $\Delta(z)$ has a four roots at $z = 0$ for all f_1 and four other roots on the real or imaginary axes. Disconnected components in the eigenvalue distribution appear only for $f_1 > 1/4$, which then shrink as f_1 increases. By computing the residue of $G(z)$ around $z = 0$, we find[9] that the proportion of eigenvalues in the center component is $2(f_1 - f_2)$. The eigenvalue equation for K yields a tractable constraint on the boundary:

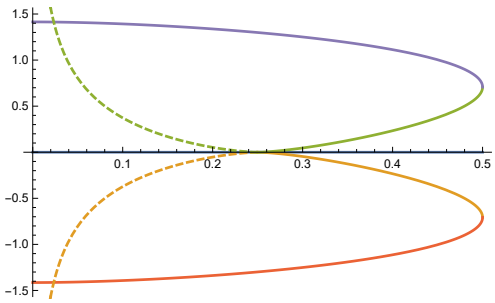
$$\gamma_2 \left(g_{ff}g_{bf} + g_d^2\right) - \gamma_1\gamma_2 \left(g_d^2 g_{ff} g_{bf} - g_{ff}^2 g_{bf}^2 + g_d^4\right) + \gamma_1 g_d^2 = 1 \tag{2.89}$$

Consider the same four layer network, but now with alternating fractions $f_m =$

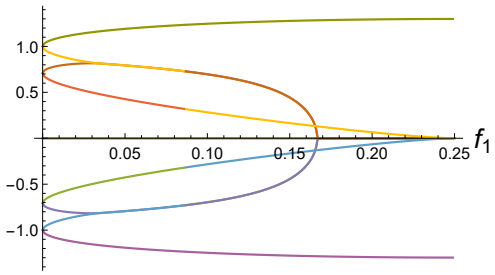
[8] This is the discriminant for χ_1 and χ_4. The discriminant for χ_2, χ_3 is identical, times a factor of f_2^2.
[9] This is left as an exercise to the reader.

(f_1, f_2, f_1, f_2). $\Delta(z)$ is now a complicated 10th order polynomial of the form $\Delta(z) = z^2(a_0 + a_1 z^2 + a_2 z^4 + a_3 z^6 + a_4 z^8)$, whose solutions, while analytic, are very long and complicated. For small values of f_1 or f_2, the eigenvalue distribution has five disconnected components (Figs 2.12b,c). We can painstakingly calculate the residue of G around $z = 0$ to yield the proportion of eigenvalues in the center component: $2|f_1 - f_2|$.

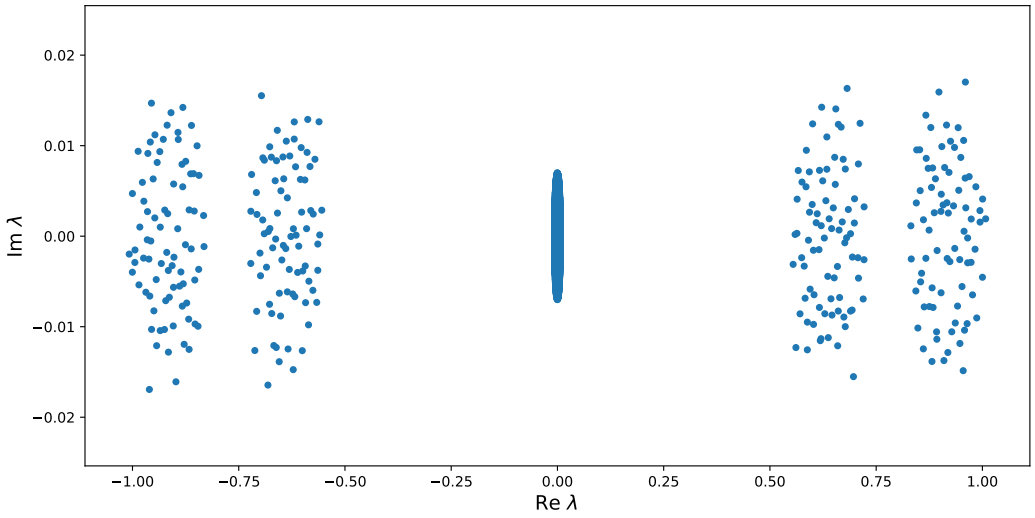
This chapter, in part, is based on material contained in Kuczala and Sharpee. *Eigenvalue spectra of large correlated random matrices.* Physical Review E, 94(5), 05010, 2016. Part of this chapter is currently being prepared for submission for publication of the material. Kuczala and Sharpee. The dissertation author was the primary investigator and author of these papers.



(a) Roots of the discriminant as a function of f_1 for symmetric layers. Imaginary roots are shown as dashed lines.



(b) Roots of the discriminant for alternating layers. Each pair of disconnected components merges into one at $f_1 \sim 0.035$. This behavior is symmetric in f_1 and f_2.



(c) Eigenvalues of a four-layer alternating network with $f_1 = 0.01$, and $N = 8000$. Note that only 320 of these eigenvalues are outside the central disk.

Figure 2.12: Four layer network

Chapter 3

Mean field dynamics

The dynamics of the rate equations

$$\dot{x_i}(t) = -x_i(t) + \sum_j J_{ij}\phi[x_j(t)] \tag{3.1}$$

can only be coarsely characterized by the eigenvalues of the connectivity matrix J. Strictly, the eigenvalues only tell us the stability of the zero fixed point. Dynamical mean field theory offers us a more quantitative characterization of the dynamics. In equilibrium, as the number of nodes $N \to \infty$, the activities $x_i(t)$ become statistically independent, and obey simple stochastic differential equations. For example, when the connections J_{ij} are i.i.d., the activities obey independent and identical equations

$$\dot{x}_i(t) = -x_i(t) + \eta_i(t) \tag{3.2}$$

where $\eta_i(t)$ is Gaussian noise with zero mean. The $\eta_i(t)$ have nontrivial temporal correlations [59, 23]

$$\langle \eta_i(t)\eta_i(t')\rangle = g^2\langle S_i(t)S_i(t')\rangle \tag{3.3}$$

determined by the gain g and nonlinearity ϕ. Here the expectation value is over realizations of $\eta_i(t)$, and $S_i(t) \equiv \phi[x_i(t)]$ The correlation function can be determined self-consistently, as we will see below. In equilibrium, the correlation function depends only on $\tau \equiv t - s$. For $g < 1$, equilibrium is achieved only when all $x_i(t) = 0$, so the correlation function is zero. More interestingly, for $g > 1$ in the chaotic phase, the correlation decreases monotonically with $|\tau|$. Intuitively, trajectories become more scrambled as time goes on due to the chaotic nature of the dynamics. As g increases, the width of the correlation function decreases, but its amplitude increases. Again, this is an intuitive consequence of the network dynamics becoming more chaotic for larger g. As $g \to 1^+$, the width of the correlation function becomes infinite. We will later see in Chapter 4 that this is where the network can best preserve information about its inputs [69]. Our goal in this chapter will be to extend dynamical mean field theory to networks where J has block-structured correlations.

3.1 Single population

We first extend the mean field theory above to a homogeneous network with nonzero correlations

$$\begin{aligned} \langle J_{ij}^2 \rangle &= g^2/N \\ \langle J_{ij} J_{ji} \rangle &= \rho g^2/N \end{aligned} \tag{3.4}$$

In this chapter, we denote the correlation coefficient as ρ, where $0 \leq \rho \leq 1$ as usual. As mentioned before, random matrix theory yields a coarse prediction of the dynamics. Since the eigenvalues of J lie on an ellipse with major radius $g(1 + \rho)$, positive correlations destabilize the zero fixed point, whereas negative correlations will stabilize it. We expect that the network transitions to chaos at $g = 1/(1 + \rho)$.

Using dynamical mean field theory (see Appendix B.1 and [43]), we can show that

each $x_i(t)$ independently obeys the mean field equation

$$\dot{x}(t) = -x(t) + g^2\rho \int dt' R(t,t')S(t') + \eta(t) \tag{3.5}$$

where $R(t,t')$ is the response function and $\eta(t)$ is Gaussian with

$$\langle \eta(t) \rangle = 0 \tag{3.6}$$

$$\langle \eta(t)\eta(t') \rangle = g^2 C(t,t') = g^2 \langle S(t)S(t') \rangle \tag{3.7}$$

as in the uncorrelated case. The effect of the correlations, then, is to introduce a convolution term with the response function, which has a 'memory' of past firing rates. Assuming $R(t,t') > 0$, and noting that $S(t)$ has the same sign as $x(t)$, the convolution term competes with the decay term$-x(t)$ for $\rho > 0$, and enhances it for $\rho < 0$. This tends to slow the dynamics for $\rho > 0$, and speed up the dynamics (and decay rate) for $\rho < 0$.

Note that the response function has $R(t,t') = 0$ for $t < t'$ due to causality, and as we shall see, decays exponentially. To determine $R(t,t')$ we use the relation (see B.22)

$$\langle S(t)\eta(t') \rangle = g^2 \int dt'' C(t',t'')R(t,t'') \tag{3.8}$$

Convolving (3.8) with the convolutional inverse[1] $C^{-1}(t,t')$ of $C(t,t')$, we get

$$\begin{aligned} \int dt' C^{-1}(s,t') \langle S(t)\eta(t') \rangle &= g^2 \int dt' dt'' C^{-1}(s,t')C(t',t'')R(t,t'') \\ &= g^2 \int dt'' \delta(s-t'')R(t,t'') \\ &= g^2 R(t,s) \end{aligned} \tag{3.9}$$

[1] $C^{-1}(t,t')$ is defined by $\int dt' C^{-1}(s,t')C(t',t'') = \delta(s-t'')$.

so that

$$R(t,s) = \frac{1}{g^2}\int dt' C^{-1}(s,t')\langle S(t)\eta(t')\rangle \tag{3.10}$$

At equilibrium, the correlation function and response function are functions only of $\tau = t-s$. Defining[2]

$$R(\tau) \equiv \int dtds R(t,s)\delta(t-s-\tau), \tag{3.11}$$

Eq. 3.10 becomes

$$R(\tau) = \frac{1}{g^2}\int dtdt' \langle S(t)\eta(t')\rangle C^{-1}(t-t'-\tau) \tag{3.12}$$

which has Fourier components

$$R(\omega) = \frac{1}{g^2}\langle S(\omega)\eta(-\omega)\rangle C^{-1}(-\omega) \tag{3.13}$$

3.2 Multiple populations

Now we consider the case where the network has M populations with sizes $f_m N$, and J has block-structured correlations

$$\begin{aligned} \langle J_{ij}^2\rangle &= \frac{1}{N} g_{m(i)n(j)}^2 \\ \langle J_{ij}J_{ji}\rangle &= \frac{1}{N}\rho_{mn} g_{mn} g_{nm} \end{aligned} \tag{3.14}$$

In the mean field limit, the activities $x_i(t)$ of each neuron are all independent and obey statistics determined by the neuron's population identity m. In Appendix B.2, we derive the M mean field equations

$$\dot{x}_m(t) = -x_m(t) + \sum_n T_{mn}\int dt' R_n(t,t')S_m(t') + \eta_m(t) \tag{3.15}$$

[2]With this definition of $R(\tau)$, $\tau > 0$ denotes a time τ in the *past*. We expect $R(\tau)$ to decrease with τ as the network loses memory of its past.

where the $\eta_m(t)$ are independently Gaussian with correlation functions

$$C_m^\eta(t,s) \equiv \langle \eta_m(t)\eta_m(s)\rangle = \sum_n G_{mn}\langle S_n(t)S_n(s)\rangle \tag{3.16}$$

For simplicity we have introduced the shorthand

$$G_{mn} \equiv g_{mn}^2 f_n, \quad T_{mn} \equiv \rho_{mn} f_n g_{mn} g_{nm} \tag{3.17}$$

The response functions $R_m(t,s)$ are determined analogously to Eq. 3.10:

$$R_m(t,s) = \int dt' [C_m^\eta]^{-1}(s,t')\langle S_m(t)\eta_m(t')\rangle \tag{3.18}$$

where $[C_m^\eta]^{-1}(t,t')$ denotes the convolutional inverse of $C_m^\eta(t,t')$. When $\tau_{mn} = 0$, the mean field equations reduce to $\dot{x}_m(t) = -x_m(t) + \eta_m(t)$ as in [5].

3.3 Simulations

We now assess the accuracy of dynamical mean field theory (DMFT) by comparing numerical simulations of the exact dynamics (3.1) to simulations of the DMFT equations. For the simulations of the exact dynamics, we compute trajectories $x(t)$ with random initial conditions and realizations of J. We first run the dynamics for sufficient time for the networks to approximately reach equilibrium, and then simulate the dynamics for an additonal time T. Then we Fourier transform to average over time and different trajectories to compute $C(\tau)$ using Eq. 3.7. For the simulations of the DMFT dynamics, we use an iterative approach to compute $C(\tau)$ and $R(\tau)$. We first select an ansatz for $C(\tau)$ and $R(\tau)$. We find that a Gaussian function is a reasonable ansatz for $C(\tau)$, and that a decaying exponential convolved with a step function is a reasonable ansatz for $R(\tau)$. We then

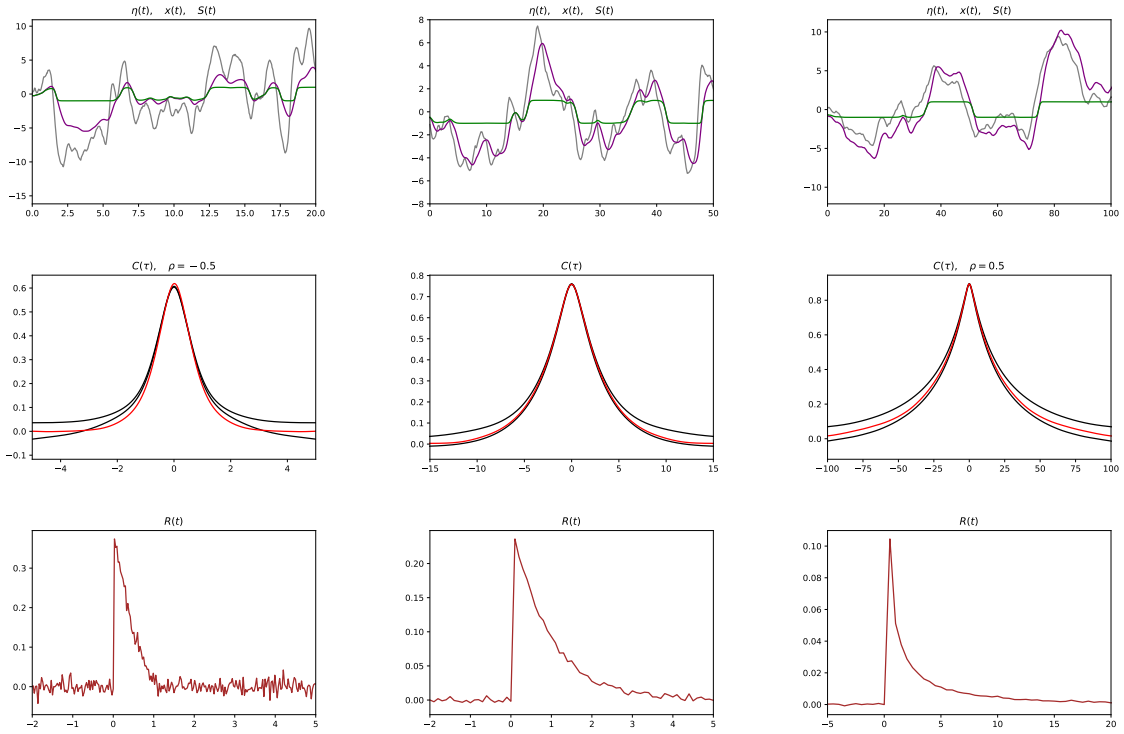


Figure 3.1: Comparison of DMFT dynamics and exact numerical simulations. Note the difference in time scales for $\rho < 0$, $\rho = 0$, and $\rho > 0$. Top: Traces of $\eta(t)$, $x(t)$, and $S(t)$ (gray, purple, and green, respectively) from an exemplar sample. Middle: Correlation functions $C(\tau)$. DMFT estimate shown in red, and $\pm$ standard deviation of $C(\tau)$ for exact simulations shown in black.

generate samples $\eta(t)$ with time correlations given by the ansatz for $g^2C(\tau)$ (Eq. 3.7). Next, we compute $x(t)$ and $S(t)$ using the mean field equation (Eq. 3.5). This can be accomplished either with fourier transforms or integration methods, though we find the latter to be more accurate in practice. Once we have samples $x(t)$ and $S(t)$, we use $S(t)$ to compute $C(\tau)$, and then Eq. 3.13 to compute $R(\tau)$. The results of simulations are shown in Fig. 3.1. We see here that compared to $\rho = 0$, $\rho > 0$ increases the width of the correlation function, whereas $\rho < 0$ decreases it, in corroboration with the results of [43].

We employ an analogous algorithm for block-structured networks by simulating the mean field equations 3.15 and iterating Eqs. 3.16 and 3.18. In Fig. 3.2, we plot results for the two-layer network explored in 2.3.1. When ρ is sufficiently negative, the correlation

function exhibits oscillations, as suggested by the eigenvalue distribution analyzed in 2.3.1. In Fig. 3.3, we compare the eigenvalue distribution, correlation function, and power spectrum of two such networks. The behavior of the power spectrum qualitatively matches the eigenvalue distribution, indicating that the gap in the eigenvalue distribution has a significant impact on the modes of the network dynamics.

The proposed algorithm is often much slower than simulating the exact dynamics, even for large values of N. A likely better approach is to derive a set of differential equations for the correlation functions and response functions (see [43, 20]), and avoid simulating the differential DMFT equations altogether, if possible.

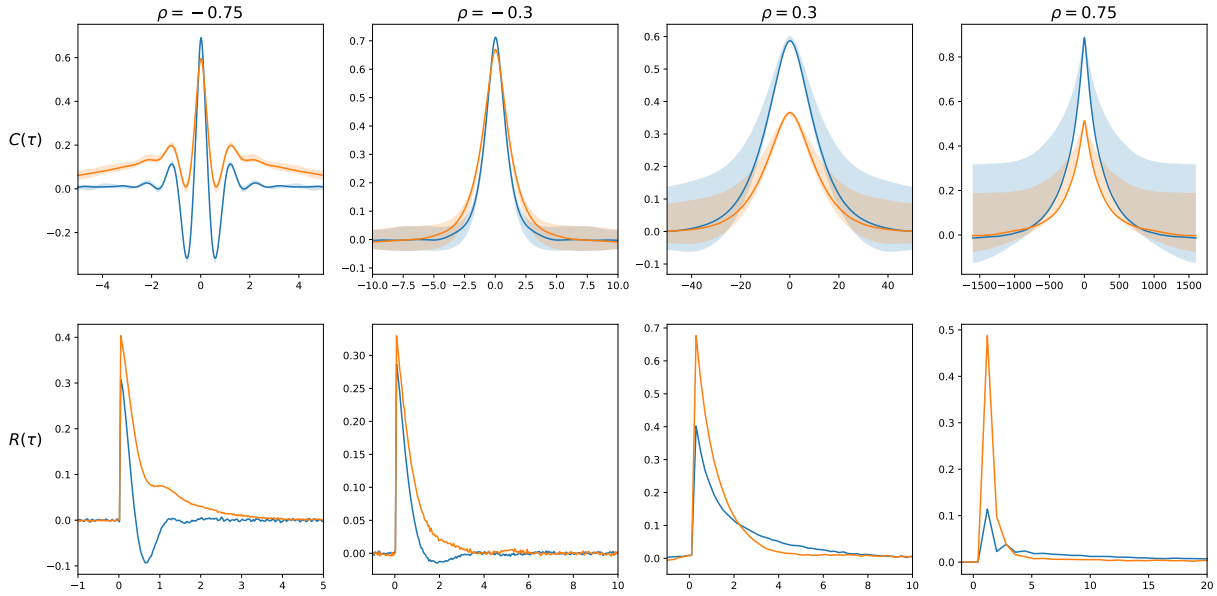


Figure 3.2: Comparison of DMFT solutions to numerical simulations for the two-layer network (Eq. 2.68, with $\rho = \tau$). Blue and orange denote the first and second populations, respectively. The shaded regions indicate +/- one standard deviation of exact simulations. From left to right, we have set $g = 16, 8, 2, 2$. In all cases, $\beta = 0.2$ and $f_1 = 1/4$.

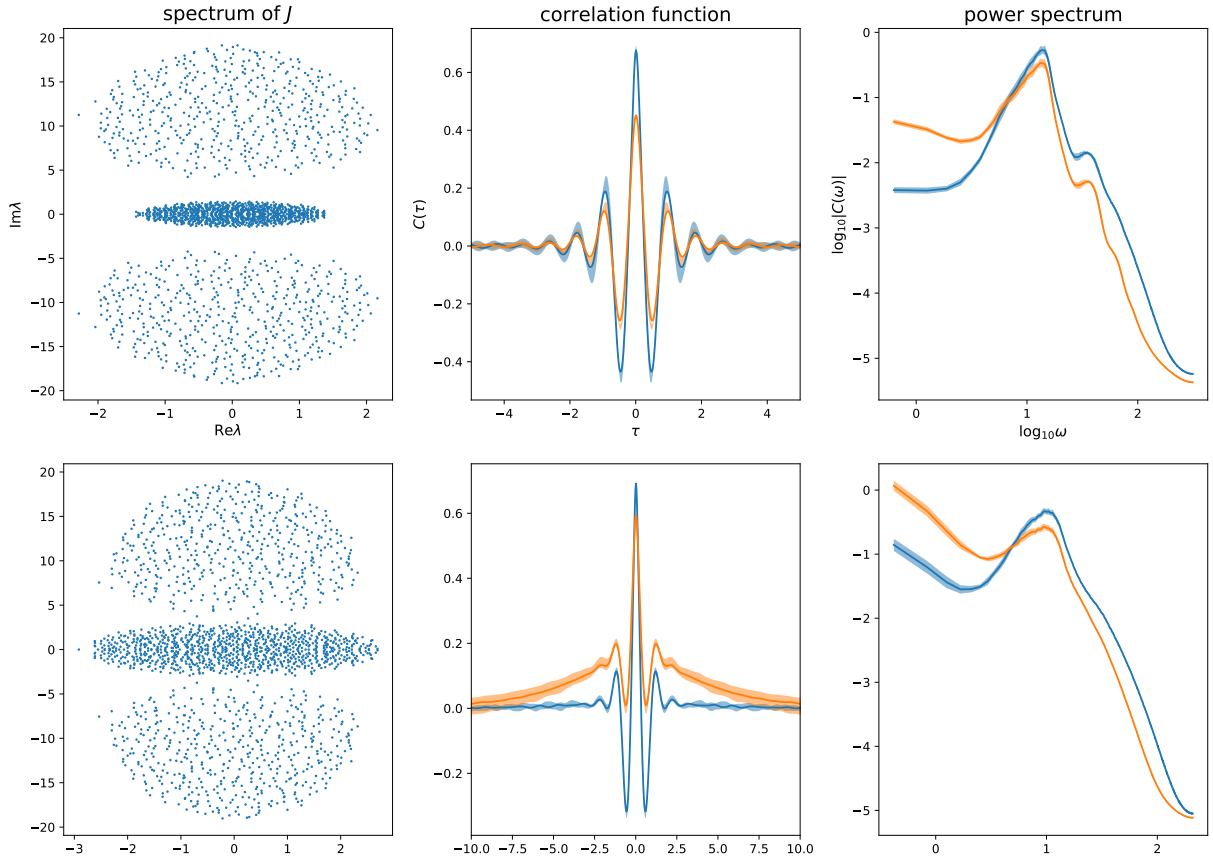


Figure 3.3: Examples of two-layer networks with a spectral gap (left). These networks have strong oscillations in the correlation function $C(\tau)$ (center). The power spectrum (right) peaks at a nonzero frequency. The top network has $\beta = 0.1$, and the bottom network has $\beta = 0.2$. In the first case, the rightmost eigenvalues lie in the high frequency components of the distribution, whereas in the second case, the rightmost eigenvalues lie in the central low frequency component. Consequently, the $\beta = 0.1$ power spectrum peaks at nonzero frequency, whereas the $\beta = 0.2$ power spectrum peaks more strongly at zero frequency.

3.4 Nonzero stable fixed points

The DMFT description of the network dynamics is valid only at equilibrium for $N \to \infty$. However, for finite N, the network dynamics may reach a fixed point. For simplicity, we discuss fixed points in single population networks. We have already seen that large networks decay to zero when $g(1 + \rho) < 1$. However, there is a finite probability that even large networks will reach a nonzero fixed point. The probability of reaching a fixed point increases with the correlation ρ, reaching certainty at $\rho = 1$, when J is symmetric. It

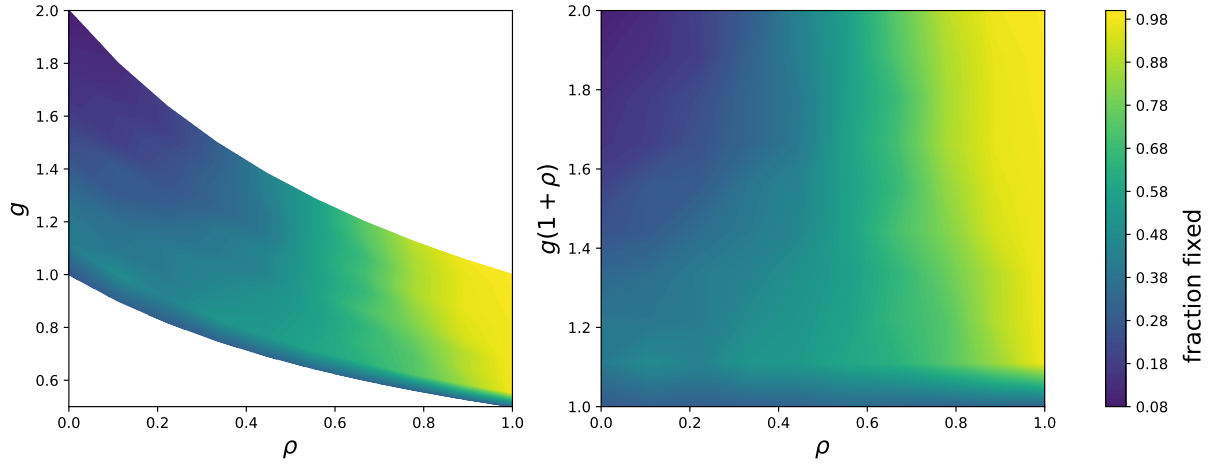


Figure 3.4: Fraction of 1000 networks reaching a nonzero stable fixed point for $N = 2000$ with a time limit of $T = 2000$ and $\epsilon = 10^{-8}$

is difficult to analytically compute the probability of a network converging to a nonzero fixed point, even for finite N. We therefore numerically approximate this probability by brute force simulation. For a given N, we choose a maximum network run time T and fixed point threshold ϵ. If a given network satisfies $\frac{1}{g^2 N}|\dot{x}(t^*)|^2 < \epsilon^2$ for $t^* < T$, we say that the network has reached a fixed point. We repeat this process for a number of initial conditions and realizations of J then count the fraction of networks that meet this criterion to estimate the probability. Fig. 3.4 shows the dependence of this probability on g and τ for $N = 2000$. We have chosen $T = 2000$ because the majority of fixed points are reached well before this time (Fig. 3.5).

Chapter 3, in part, is currently being prepared for submission for publication of the material. Kuczala and Sharpee. The dissertation author was the primary investigator and author of this paper.

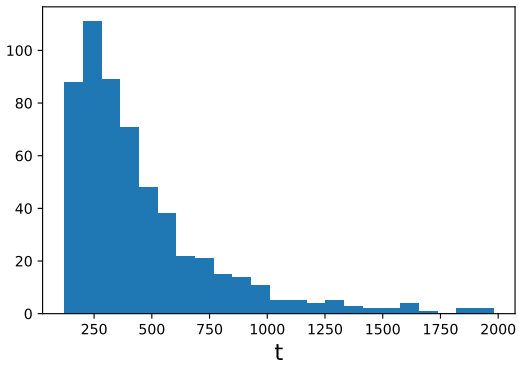


Figure 3.5: Histogram of fixed point times for $N = 2000$, $T = 2000$, $\epsilon = 10^{-8}$, $g = 1.56$ and $\tau = 0.55$

Chapter 4

Memory capacity of block-structured networks

There is significant evidence that the information processing capabilities of homogeneous random networks is maximized at the edge of chaos [69, 14, 34, 63]. However, in structured networks, non-normal and feed-forward structure plays an important role in retention and amplification of inputs [29, 50, 1, 70]. Due to their self-averaging properties, the effective structure of homogeneous random networks is highly constrained and cannot fully take advantage of feed-forward connections. In this chapter we analyze the ability of a block-structured network to retain information about its inputs. We compute the Fisher information about an input perturbation given a noisy readout of the network activity, and find a general expression dependent on the block structure. Finally, we maximize the Fisher information for two-population networks with constraints on the population size mean synaptic gain.

We consider discrete dynamics of N neurons, given by

$$\begin{aligned} x_i(t) &= \sum_{j=1}^{N} J_{ij} S_j(t) + \xi_i(t), \\ S_j(t) &= \phi\left[w_j I(t-1) + x_j(t-1)\right]. \end{aligned} \tag{4.1}$$

As before,the $x_i(t)$ and $S_i(t)$ denote the activity and firing rate of the ith neuron, respectively. The function ϕ is a saturating nonlinearity with $\phi'(0) = 1$. $I(t)$ is a scalar external input to the network, and the w_i are input weights to each neuron. The network has internal white Gaussian noise $\eta_i(t)$ with $\langle \xi_i(t)\xi_j(s)\rangle = \sigma^2 \delta_{ij}\delta_{ts}$.

In previous work [69], the connectivity matrix J was taken to be a homogeneous i.i.d. Gaussian random matrix. In order to introduce heterogeneity into the network, we consider M cell types with population sizes $f_1 N \cdots f_m N$, where $\sum_m f_m = 1$. We allow the variance of connections to depend on the input and output cell type m and n. Then the connectivity matrix J has block-structured variances

$$\langle J_{ij}^2 \rangle_J = \frac{1}{N} g_{m(i)n(j)}^2. \tag{4.2}$$

where $g_{mn} \geq 0$ is the population-dependent gain, and $\langle \cdot \rangle_J$ denotes the expectation value over realizations of J. For simplicity, we take $\langle J_{ij} \rangle_J = 0$, and neglect correlations between matrix elements.

For large N, the dynamics of (4.1) are well approximated by dynamical mean field theory [5, 69, 48]. We can straightforwardly adapt the results of Chapter 3 to the discrete case. In the limit $N \to \infty$, the neuron activities $x_i(t) = \eta_i(t)$ are independent and Gaussian with zero mean. The statistics are completely described by the correlation

function $q_m(t,s) \equiv \langle x_m(t) x_n(s) \rangle$, given by

$$\begin{aligned} q_m(t,s) &= \sigma^2 \delta_{ts} + \sum_n g_{mn}^2 f_n \langle S_n(t) S_n(s) \rangle_n \\ &= \sigma^2 \delta_{ts} + \sum_n G_{mn} C_n(t,s) \end{aligned} \tag{4.3}$$

where we have defined

$$\begin{aligned} C_n(t,s) &\equiv \langle S_n(t) S_n(s) \rangle_n \\ G_{mn} &\equiv g_{mn}^2 f_n \end{aligned} \tag{4.4}$$

and $\langle \cdot \rangle_n$ denotes expectation over the Gaussian variables $x_n(t)$. Compare Eq. 3.16.

When the input is constant, the system reaches equilibrium, and the correlation function $q_m(t,s) \propto q_m \delta_{t,s}$. q_m is the variance of $x_m(t)$, satisfying the self consistent equations

$$q_m = \sigma^2 + \sum_n G_{mn} \langle S_n^2 \rangle_n \tag{4.5}$$

and can be solved numerically by iteration. When the internal noise σ is small, the spectral radius of G, $\rho(G)$, determines the macroscopic behavior of the network. As in the continuous case, the network transitions from silent to chaotic activity at $\rho(G) \to 1$ as $\sigma \to 0$ [4, 5, 3, 48]. For finite σ, the noise suppresses the onset of chaos, and the transition is pushed to larger values of g [48, 53].

4.1 M populations

We now investigate how much information about an input perturbation at time t_0 can be read out of the network at time t. The input $I(t) = \theta \delta_{t,t_0}$ is an infinitesimal pulse with amplitude θ at time t_0, where we will take the limit $\theta \to 0$. We consider noisy

readouts of the network activity

$$v_i(t) = w_{m(i)} I(t) + x_i(t) + \chi_i(t) \tag{4.6}$$

from $K \sim \mathcal{O}(\sqrt{N})$ nodes of the network, preserving the relative fractions f_m from each population. The last term is i.i.d. Gaussian white noise with variance σ^2_{obs}. The input weights w_m depend only on the population identity m, and we normalize $\sum_m w_m^2 = 1$. The Fisher information that the readout contains about θ at time t is

$$\mathcal{I}(\theta, t) = \left\langle \frac{\partial}{\partial\theta} \frac{\partial}{\partial\theta} \log p(\mathbf{v}(t)|\theta) \right\rangle_{p(\mathbf{v}(t)|\theta)}, \tag{4.7}$$

known as the Fisher memory curve (FMC) [29]. For a fixed realization of J, $p(\mathbf{v}|\theta)$ is Gaussian. Assuming the covariance of the distribution is independent of θ, the FMC has the well-known form [69, 29]

$$\mathcal{I}(\theta, t) = \sum_{ij} \left\langle \frac{\partial \mu_i(t)}{\partial\theta} \Sigma^{-1}_{ij}(t) \frac{\partial \mu_j(t)}{\partial\theta} \right\rangle_J \tag{4.8}$$

where

$$\mu_i(t) \equiv \langle v_i(t)|J\rangle \tag{4.9}$$

$$\Sigma_{ij}(t) \equiv \mathrm{Cov}[v_i(t), v_j(t)|J] \tag{4.10}$$

are the sufficient statistics of the $v_i(t)$ averaged over different trajectories (replicas) for a fixed network realization J. From Appendix B of [69], the replica-averaged covariance Σ deviates from the network averaged covariance only by $\mathcal{O}(N^{-1/2})$:

$$\Sigma_{ij}(t) = \langle \Sigma_{ij}(t) \rangle_J + \mathcal{O}(N^{-1/2}) \tag{4.11}$$

The $\mathcal{O}(N^{-1/2})$ term introduces corrections of order $\sqrt{K/N}$ to the fisher information. So long as $K \sim \mathcal{O}(\sqrt{N})$ or less, these corrections are negligible.

Assuming Σ is approximately θ independent as before, the mean field equilibrium statistics (4.5) yield

$$\langle \Sigma_{ij}(t) \rangle_J = \delta_{ij}(\sigma_{obs}^2 + q_n) \tag{4.12}$$

which is diagonal, meaning that the FMC is the weighted sum of squared signal-to-noise ratios from each population

$$\mathcal{I}(\theta, t) = K \sum_m \frac{f_m}{\sigma_{obs}^2 + q_m} \left\langle \left(\frac{\partial \mu_m(t)}{\partial \theta} \right)^2 \right\rangle_J \tag{4.13}$$

In the limit of vanishingly small input perturbations $\theta = 0$, we obtain a tractable expression for $\langle (\partial \mu_m / \partial \theta)^2 \rangle$ using replicas (see Appendix C). In this case, the FMC is

$$\mathcal{I}(t \geq t_0) = K \sum_{mn} \frac{f_m}{\sigma_{obs}^2 + q_m} (\mathbf{M}^{t-t_0})_{mn} w_n^2 \tag{4.14}$$

The matrix $M_{mn} \equiv G_{mn} \langle \phi' \rangle_n^2$ describes the decay or amplification of information in each population. When $\mathbf{M}$ is highly non-normal, the sequence $\|\mathbf{M}^{t-t_0}\|$ exhibits transient amplification [70], even though $\rho(M) < 1$. For non-normal $\mathbf{M}$, the spectral radius only asymptotically describes the decay rate of the sequence (Fig. 4.2).

Since $v(t)$ is uncorrelated in time for $\theta = 0$, the probability distribution over the entire readout $p(\{v(t_0), \ldots, v(\infty)\}|\theta)$ factorizes across time, and we sum to compute the total signal-to-noise ratio (SNR):

$$\begin{aligned} \mathcal{I} &= \sum_{t=t_0}^{\infty} \mathcal{I}(\theta, t) \\ &= K \sum_{mn} \frac{f_m}{\sigma_{obs}^2 + q_m} [(\mathbf{1} - \mathbf{M})^{-1}]_{mn} w_n^2 \end{aligned} \tag{4.15}$$

This depends on the input weights w, but we can easily determine the best case using the l_1 norm, yielding

$$\max_w \mathcal{I} = K \left\| \frac{f_m}{\sigma_{obs}^2 + q_m} \left[(\mathbf{1} - \mathbf{M})^{-1} \right]_{mn} \right\|_1 \tag{4.16}$$

where $||A_{mn}||_1$ denotes the l_1 operator norm of a matrix $\mathbf{A}$. From (C.8), the spectral radius $\rho(\mathbf{M}) < 1$, and the SNR is maximized when the spectral radius $\rho(\mathbf{M}) \approx 1$. When $\sigma < \sigma_{obs}$ is small, this occurs close to the transition to chaos.

The behavior of $\mathcal{I}$ away from the critical regions depends on features of M other than the eigenvalues. In particular, the SNR is generally larger the less normal the matrix M is. The resolvent $\|(z - M)^{-1}\|$ diverges faster near eigenvalues for strongly non-normal matrices, and is closely related to the phenomenon of transient amplification [70].

For $M = 1$, this expression reduces to the single-population case [69]:

$$\begin{aligned} \mathcal{I}_{single} &= K \frac{1}{\sigma_{obs}^2 + q} \frac{1}{1 - g^2 \langle \phi' \rangle^2} \\ q &= g^2 \langle S^2 \rangle \end{aligned} \tag{4.17}$$

For $\sigma_{obs} > \sigma$, $\mathcal{I}_{single}$ is maximized for some $g^* > 1$ (see Fig. 4.1). When either the internal noise $\sigma \to 0$ or $\sigma \ll \sigma_{obs}$, the optimal gain $g^* \to 1$.

4.2 Two populations

We now consider the problem of determining the optimal network architecture (choices of g_{mn} and f_m) for a network with two populations, under particular constraints. First, we put an upper bound 'budget' on the mean synaptic gain $\bar{g} \equiv \left(\sum_{mn} f_m g_{mn}^2 f_n \right)^{1/2} < \bar{g}_{max}$. Secondly, we place a lower limit $f_{min} N$ on the number of neurons in a given population, such that the mean field limit is valid to a desired accuracy for a given N. For example, if $N = 1000$, one may choose $f_{min} = 0.1$ so that each population has at least 100 neurons and

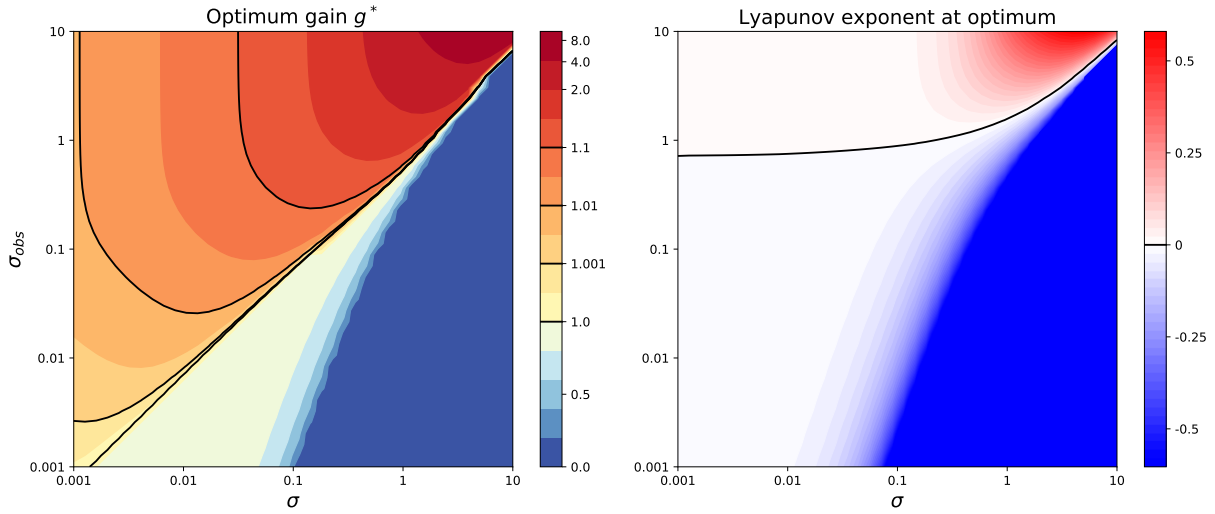


Figure 4.1: Optimal gain g^* and Lyapunov exponent for a single-population network, as a function of internal noise σ and observation noise σ_{obs}. When $\sigma < \sigma_{obs}$, $g^* > 1$. For small values of σ, the total SNR is maximized when the Lyapunov exponent ≈ 0, close to the edge of chaos.

the largest fluctuations are of order $1/\sqrt{f_{min}N} = 1/10$. For simplicity, we restrict $\sigma < \sigma_{obs}$ and $\sigma \ll 1$.

To determine the form of the optimal network, we first numerically maximize (4.16) with respect to the G_{mn} and f_m with the above constraints. We find that optimal matrices $\mathbf{G}$ have at least one vanishing off-diagonal element, corresponding to a feedforward architecture. We can parameterize these networks in terms of f_1, $\bar{g}$ and the eigenvalues λ_1, λ_2 of $\mathbf{G}$ as follows:

$$\mathbf{G}_{FF} = \begin{bmatrix} \lambda_1 & 0 \\ G_{21} & \lambda_2 \end{bmatrix} \tag{4.18}$$
$$G_{21} = f_2^{-1}(\bar{g}^2 - f_1\lambda_1 - f_2\lambda_2)$$

where $G_{21} \geq 0$ requires that $f_1\lambda_1 + f_2\lambda_2 \leq \bar{g}^2$. In this case, the equilibrium variances q_m

satisfy

$$\begin{aligned} q_1 &= \sigma^2 + \lambda_1 \langle S^2 \rangle_1 \\ q_2 &= \sigma^2 + \lambda_2 \langle S^2 \rangle_2 + G_{21} \langle S^2 \rangle_1 \end{aligned} \tag{4.19}$$

The best case SNR is

$$\begin{aligned} \max_w \mathcal{I} &= \max(\mathcal{I}_1, \mathcal{I}_2) \\ \mathcal{I}_1 &= \frac{f_1}{\Delta_1(\sigma_{obs}^2 + q_1)} + \frac{f_2 \langle \phi' \rangle_1^2 G_{21}}{\Delta_1 \Delta_2 (\sigma_{obs}^2 + q_2)} \\ \mathcal{I}_2 &= \frac{f_2}{\Delta_2(\sigma_{obs}^2 + q_2)} \end{aligned} \tag{4.20}$$

where $\Delta_m \equiv 1 - M_{mm} = 1 - \langle \phi' \rangle_m^2 \lambda_m$. $\mathcal{I}_1$ and $\mathcal{I}_2$ denote the SNR (4.15) for $w = (1, 0)$ and $w = (0, 1)$, respectively. In the first case, the input is fed into the first population, which has feed-forward connections to the second. In the second case, the input is fed into the second population, which cannot propagate into the first. We therefore expect that $\max_w \mathcal{I} = \mathcal{I}_1$ in situations where the feed-forward connections can be leveraged to maximize information. In all cases, the SNR is maximized when the mean gain is saturated, $\bar{g} = \bar{g}_{max}$ and increases monotonically with $\bar{g}$ (Fig. 4.3). Notably, the SNR of the feed-forward network equals the SNR of the single-population network at some $\bar{g} = g_c$ slightly less[1] than the single-population maximum at g^*. The optimal choices of the network parameters λ_1, λ_2 and f_1 vary with $\bar{g}$. Notably, $\lambda_2 = 0$ for $\bar{g} < g_c$, and $\lambda_2 > 0$ for $\bar{g} > g_c$, meaning that the second population has no recurrent connections for $\bar{g} < g_c$. Further, for most values of $\bar{g}$, $f_1 = f_{min}$, meaning that the network has divergent feed-forward connections. We discuss these varying behaviors below.

For sufficiently small mean gain $\bar{g} < \bar{g}_s$, the feed-forward connections do not provide any benefit, and $\mathcal{I}_2 > \mathcal{I}_1$. In fact, the single-population network with gain $\bar{g}$ outperforms

[1] As $\sigma \to 0$, $g_c \to g^* \to 1$.

all two-population networks with $\bar{g} < \bar{g}_s$. For $\bar{g}_s < \bar{g} < g_c$, the second population has no recurrent connections ($\lambda_2 = 0$), and is simply a noisy readout of the first population. At $\bar{g} = g_c$, the memory capacity of the feed-forward network approximately equals the memory capacity of the single population network at its maximum, and $\mathcal{I}_1 \to \mathcal{I}_{single}$ as $f_{min} \to 0$. The feed-forward network becomes convergent as $f_1 \to 1 - f_{min}$ as $\bar{g}$ approaches g_c from either side. As a consequence, $\mathcal{I}_1$ becomes insensitive to λ_2 (Fig. 4.4). For $\bar{g} > g_c$, both λ_1 and λ_2 are nonzero, meaning that there are recurrent dynamics in both populations. The strength of the recurrent connections controls the eigenvalues of M, the largest of which sets the spectral radius $\rho(M)$, which in turn determines the asymptotic rate of decay of the FMC. Secondly, the feed-forward connections increase monotonically for large $\bar{g}$, causing the transient amplification to rapidly increase with $\bar{g}$ (Fig. 4.2). This can be understood in the context of non-normality. $\mathbf{G}_{FF}$ is in Schur form, and one measure of its departure from normality is the relative size of the off-diagonal element [70]. Alternatively, we note that the angle ψ between the eigenvectors of $\mathbf{G}_{FF}$ is given by

$$\cos\psi = \frac{G_{21}}{\sqrt{G_{21}^2 + (\lambda_1 - \lambda_2)^2}} \tag{4.21}$$

The eigenvectors are perpendicular for $G_{21} = 0$, when G is diagonal, and the eigenvectors become increasingly parallel as G_{21} increases.

Chapter 4, in part, is currently being prepared for submission for publication of the material. Kuczala and Sharpee. The dissertation author was the primary investigator and author of this paper.

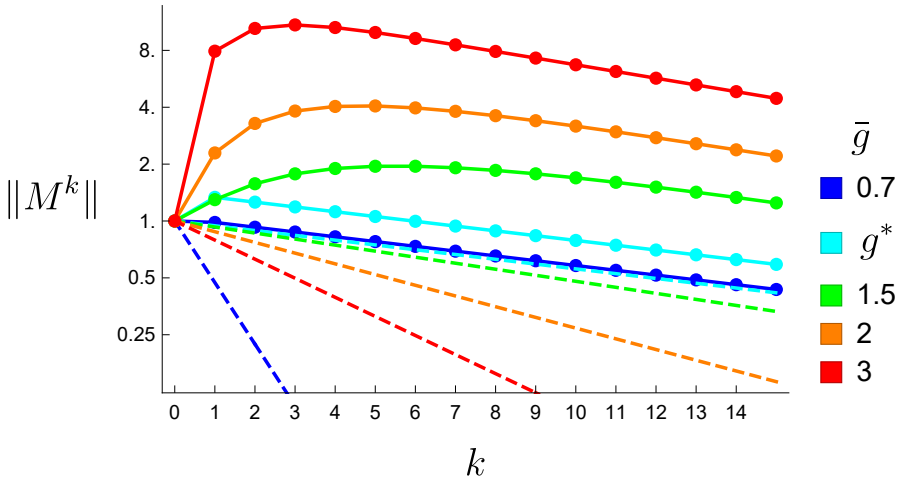


Figure 4.2: Amplification profiles of the matrix **M**. The dashed lines indicate the amplification profiles of single population networks with $g = \bar{g}$. For $\bar{g} > g_c$, **M** has significant transient amplification. Note that the rate of transient amplification is faster for larger $\bar{g}$, and that the rate of asymptotic decay is largely independent of $\bar{g}$.

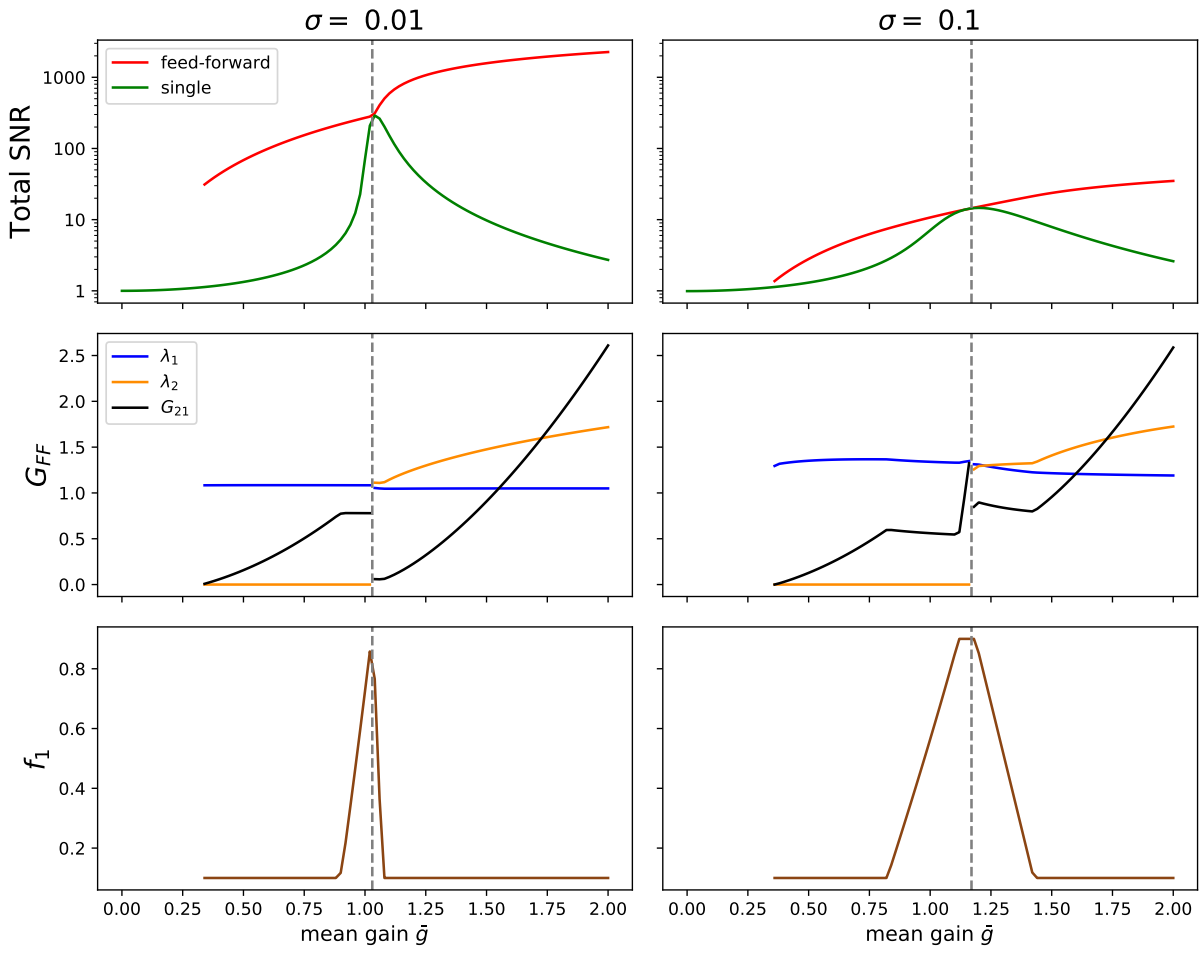


Figure 4.3: Best case fisher information and optimal parameters as a function of $\bar{g}$, with $f_{min} = 0.1$ and $\sigma_{obs} = 1$.

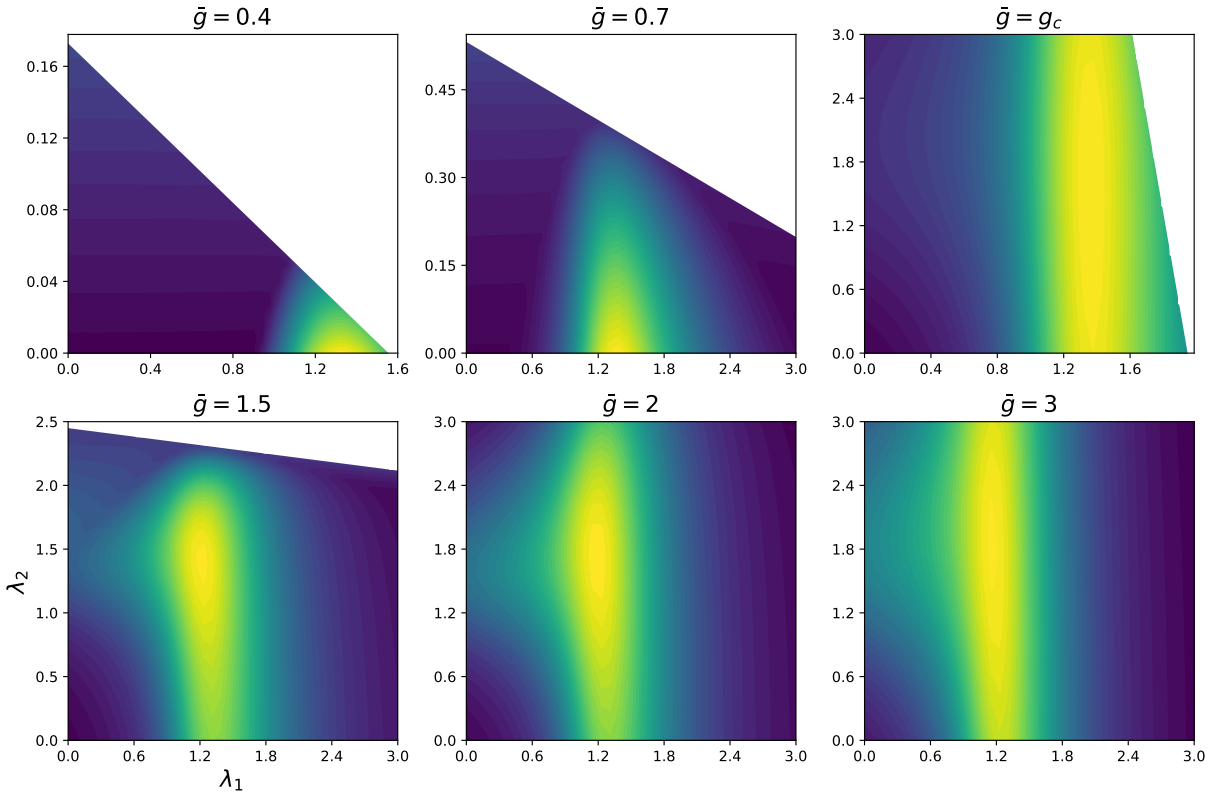


Figure 4.4: Total SNR of the feed-forward network (Fig. 4.3, $\sigma = 0.1$), as a function of λ_1 and λ_2 for select values of $\bar{g}$. For each plot we have chosen the value of f_1 shown in (4.3). Parts of the plot region are excluded according to the condition $f_1\lambda_1 + f_2\lambda_2 < 1$. For all values of $\bar{g}$, the SNR is sensitive to the value of λ_1. While the SNR is sharply peaked at $\lambda_2 = 0$ for $\bar{g} < g_c$, the SNR becomes less sensitive to λ_2 as $\bar{g} \to g_c$. For $\bar{g} > g_c$, the SNR is initially peaked around a nonzero value of λ_2, but becomes less strongly peaked as $\bar{g}$ increases.

Chapter 5

Error correction in symmetric networks

Here we study the optimal properties of symmetric RNNs for input reconstruction. In the quadratic decoding model of cortical circuits [13], the sufficient statistics of stimuli are encoded in the connection weights of a recurrent network. The matrix of weights are capable of encoding multiple ambiguous stimuli, with the most prescient stimulus encoded in the dominant eigenvector. Given an initial estimate of this stimulus from a linear decoder, we expect that the dynamics of the recurrent network will converge to the dominant eigenvector. By analogy, the power iteration method for finding the dominant eigenvector works by iteratively multiplying some initial vector by the matrix, and normalizing the resulting vector. The rate dynamics performs similar operations: the network connections perform matrix multiplication, and the nonlinearity normalizes the vector. We would like to optimize an ensemble of symmetric rate networks for this purpose.

We consider an ensemble of networks with normally distributed symmetric $N \times N$

connectivity matrices J, with

$$\langle J_{ij} \rangle = 0 \tag{5.1}$$

$$\langle J_{ij}^2 \rangle = (1 + \delta_{ij})/N \tag{5.2}$$

Here $\langle \cdot \rangle$ denotes an expectation value over the Gaussian orthogonal ensemble. Requiring J to be symmetric implies that its eigenvalues are real and its eigenvectors are orthogonal. Since J is real, the eigenvectors can also be chosen to be real-valued. Another consequence, as we will see below, is that in appropriate limits, the network dynamics can be described by a Lyapunov function, which guarantees that the network activity will asymptotically approach a fixed point. In this setting, it is meaningful to ask under what conditions the network activity approaches the eigenvector $\mathbf{e}_1$, corresponding to the rightmost eigenvalue [1] λ_1 of J. More specifically, we want the network activity to be an approximate scalar multiple of $\mathbf{e}_1$, with a choice of sign.

The network dynamics is given by

$$\frac{d\mathbf{x}}{dt} = -\mathbf{x}(t) + J\phi[g\mathbf{x}(t)] + \boldsymbol{I}(t) + \sigma_t \boldsymbol{\xi}(t) \tag{5.3}$$

where $\mathbf{x}(t)$ is the N-dimensional activity vector, and $\phi(x)$ is a saturating nonlinearity applied element-wise. We choose $\phi(x) = \tanh(x)$ throughout, but our results can be easily generalized to any saturating nonlinearity with $0 \leq \phi'(x) \leq \phi'(0) = 1$. For example, our results also qualitatively hold for a piecewise linear function. $\boldsymbol{\xi}(t)$ is Gaussian white noise with $\langle \xi_i(t) \rangle_{\boldsymbol{\xi}} = 0$ and $\sigma_t^2 \langle \xi_i(t) \xi_j(t') \rangle_{\boldsymbol{\xi}} = \sigma_t^2 \delta_{ij} \delta(t - t')$. We take σ_t to be small compared to the input $\boldsymbol{I}(t)$; in particular we will choose $\sigma_t = 0.01$ unless otherwise specified. This noise prevents the network activity from accurately approximating $\mathbf{e}_1$ for arbitrarily small $|x|$.

[1] In contrast to the power iteration algorithm, which converges to the eigenvector corresponding to eigenvalue with largest absolute value (the spectral radius), the recurrent networks considered here do not converge to eigenvectors with negative eigenvalue, even for those with maximum absolute value.

The rate dynamics described in Eq. 5.3 has a few crucial differences from the rate networks considered in previous chapters. Firstly, the variance of J is fixed, and g instead controls the steepness of the nonlinearity. This amounts to rescaling the activity vector $\mathbf{x}$ in previous chapters by g. One consequence of this choice is that the norm of the fixed points of the dynamics saturates instead of diverging for large g. We include a pulsed input $\boldsymbol{I}(t)$ of length Δt, the form

$$\boldsymbol{I}(t) = a\widetilde{\mathbf{e}_1}\Theta(\Delta t - t) \tag{5.4}$$

for $t > 0$ and $\Delta t < T$, where

$$\widetilde{\mathbf{e}_1} = (\mathbf{e}_1 + \sigma\boldsymbol{\eta})/(1 + \sigma^2) \tag{5.5}$$

is an estimate of the eigenvector $\mathbf{e}_1$. $\boldsymbol{\eta}$ is a static i.i.d. Gaussian vector with $\langle \sigma^2 |\boldsymbol{\eta}|^2 \rangle = \sigma^2$, which serves to corrupt the input. The constant $\sigma > 0$ controls the size of the corruption noise. Averaged over $\boldsymbol{\eta}$, the pulse has amplitude $a > 0$. Although both $+\mathbf{e}_1$ and $-\mathbf{e}_1$ correspond to the same eigenvector, the input introduces a bias towards $+\mathbf{e}_1$. We set the initial condition $\mathbf{x}(0) = \boldsymbol{I}(0)$.

For a given readout time T, we there is an optimal choice of gain g and input size a for which the network activity $x(T)$ most closely converges to $\mathbf{e}_1$ on average. In other words, $\mathbf{x}(T) \approx c\mathbf{e}_1$ for some constant $c > 0$. We measure the deviation of $\mathbf{x}$ from $\mathbf{e}_1$ with the l_2 norm

$$\begin{aligned} d(\mathbf{x}) = \frac{1}{4}\left|\frac{\mathbf{x}}{|\mathbf{x}|} - \mathbf{e}_1\right|^2 &= \frac{1}{2}\left(1 - \frac{\mathbf{x}\cdot\mathbf{e}_1}{|\mathbf{x}|}\right) \\ &= \frac{1}{2}(1 - \cos(\theta)) \end{aligned} \tag{5.6}$$

where θ is the angle between $\mathbf{e}_1$ and $\mathbf{x}$. We quantify the performance of the network in

terms of its ability to improve the estimate $\widetilde{\mathbf{e}_1}$. To that end we define the performance metric

$$m(x, \widetilde{\mathbf{e}_1}) = \log_{10} \frac{d(\widetilde{\mathbf{e}_1})}{d(x)} \tag{5.7}$$

This measures the number of order of magnitudes by which the network improves the estimate: $m > 0$ and $m < 0$ indicate improvement and diminishment of the estimate, respectively.

5.1 Linear fixed points

Let us first consider the time-independent case with $\sigma_t = 0$, and an infinite pulse width $\Delta t \to \infty$. A fixed point $\mathbf{x}_f$ satisfies

$$\mathbf{x}_f = J\phi(\mathbf{x}_f) + \boldsymbol{I} \tag{5.8}$$

which in the orthonormal eigenbasis $\{\mathbf{e}_k\}$ of J is

$$x_k = \lambda_k S_k + b\delta_{1k} + \sigma\eta_k \tag{5.9}$$

where v_k denotes $\mathbf{e}_k \cdot \boldsymbol{v}$ for any vector $\boldsymbol{v}$ and $\boldsymbol{S} = \phi(\mathbf{x})$. For linear networks, with $\phi(\mathbf{x}) = \mathbf{x}$, we can solve for x_k:

$$x_k = \frac{b\delta_{1k} + \sigma\eta_k}{1 - \lambda_k} \tag{5.10}$$

This fixed point is stable only when all $\lambda_k < 1$. We see that the network amplifies modes with eigenvalues close to 1. When the network has any $\lambda_k > 1$, the corresponding x_k diverge exponentially, with the fastest growing modes corresponding to the largest λ_k. In either case, the network evolves towards an approximation of the dominant eigenvector.

5.2 Lyapunov function

In the nonlinear case, the presence of the saturating function ϕ prevents the network activity from diverging. Since J is symmetric, for a constant input $\boldsymbol{I}$ the network evolution is described by a Lyapunov function. For simplicity, we first set $\boldsymbol{I} = 0$, in which case the Lyapunov function is

$$E(\boldsymbol{S}) = -\frac{g}{2}\sum_{ij} J_{ij}S_iS_j + \sum_i \int_0^{S_i} \phi^{-1}(S)dS \tag{5.11}$$

where here $S_i \equiv \phi(gx_i)$. The minima of E correspond to stable fixed points of the network dynamics. Further, we have $\dot{\boldsymbol{x}}(t) = -g\frac{\partial E}{\partial \boldsymbol{S}}$. Note that since $S_i \in (-1, 1)$, E has an infinite barrier near the edges of the unit hypercube $(-1, 1)^N$. This holds for any $\phi(x)$ that saturates to ± 1. Near $\boldsymbol{S} = 0$, the dynamics is approximately linear, with

$$E_{linear} = \frac{1}{2}\sum_{ij}(-gJ_{ij} + \delta_{ij})S_iS_j \tag{5.12}$$

When $g\lambda_1 > 1$, the origin is a saddle point with unstable axes along all eigenvectors $\mathbf{e}_k$ of J with $g\lambda_k > 1$. The Hessian of $E(\boldsymbol{S})$ is

$$\frac{\partial^2 E}{\partial S_i S_j} = -gJ_{ij} + \frac{\partial}{\partial S_i}\phi^{-1}(S_j) \geq -gJ_{ij} + \delta_{ij} = \frac{\partial^2 E_{linear}}{\partial S_i S_j}, \tag{5.13}$$

noting that[2] $\frac{\partial}{\partial S}\phi^{-1}(S) \geq 1$. In the linear case the Hessian is positive definite for $g\lambda_1 < 1$. The inequality (5.13) implies that the Hessian for the nonlinear case is also positive semidefinite, so that the Lyapunov function is convex. Therefore there is only one global minimum for $g\lambda_1 < 1$.

When $g\lambda_1 > 1$, the saddle point potential and nonlinear potential compete, yielding

[2]For $\phi(x) = \tanh(x)$, $\frac{\partial}{\partial S}\phi^{-1}(S) = (1 - S^2)^{-1}$

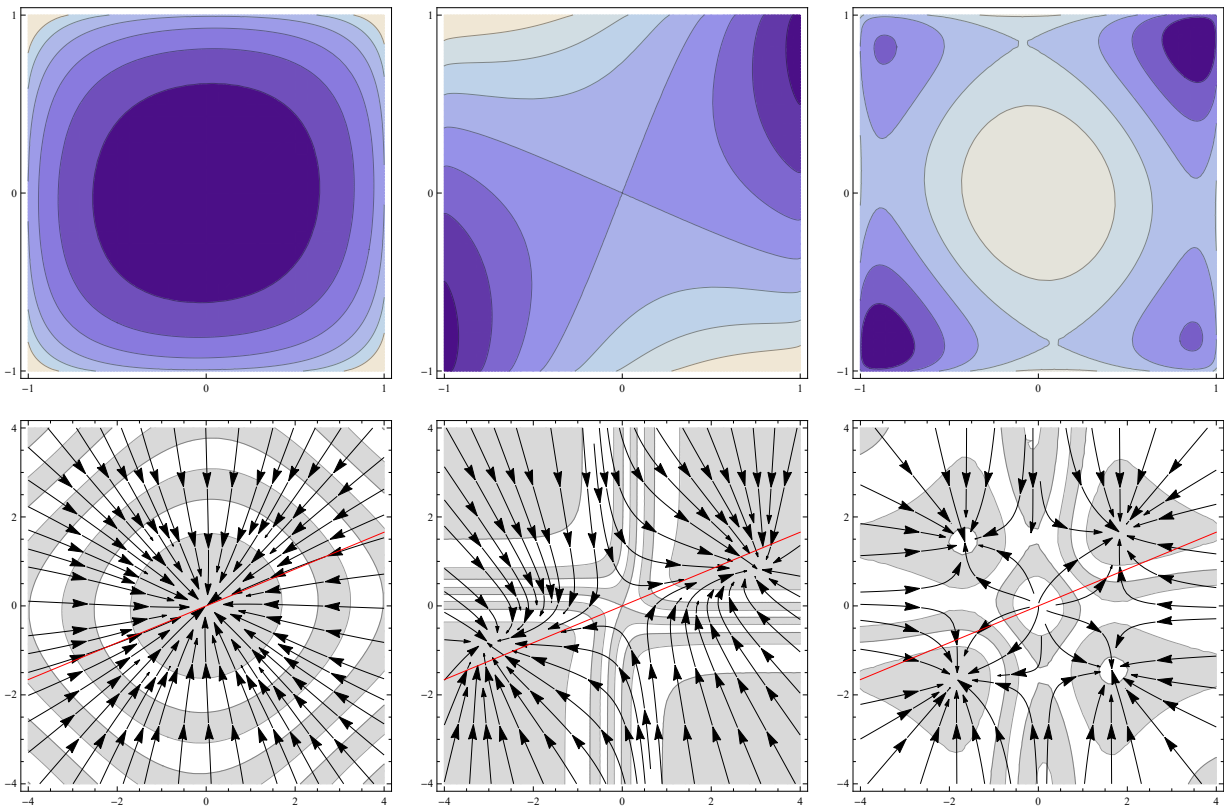


Figure 5.1: Top: Contours of the lyapunov function as a function of $\boldsymbol{S}$ for examples of $N = 2$ networks. Bottom: Associated dynamics as a function of $\mathbf{x}$. Red line indicates $\pm\mathbf{e}_1$

2^m minima, where m is the number of unstable eigenvalues. As g increases, these minima are pushed closer to the boundary of the hypercube [35]. As an example, for $N = 2$, E has either one, two, or four minima depending on the eigenvalues of J (Figure 5.1). When $\lambda_1 \approx 1$, there is a line of approximate fixed points along $\mathbf{e}_1$ near the origin, where the dynamics is slow.

For a fixed g, the network ensemble over realizations of J will generally contain a mixture of stable and unstable networks. As $N \to \infty$, the distribution of eigenvalues approaches the the semicircular distribution [73], with support on $[-2, 2]$. In this limit, the linear network is stable for $g < 1/2$ and unstable for $g > 1/2$. For finite N, the rightmost eigenvalue fluctuates depending on the particular realization of J. The transition from unstable to stable dynamics becomes sharper for larger N. As N decreases, the transition

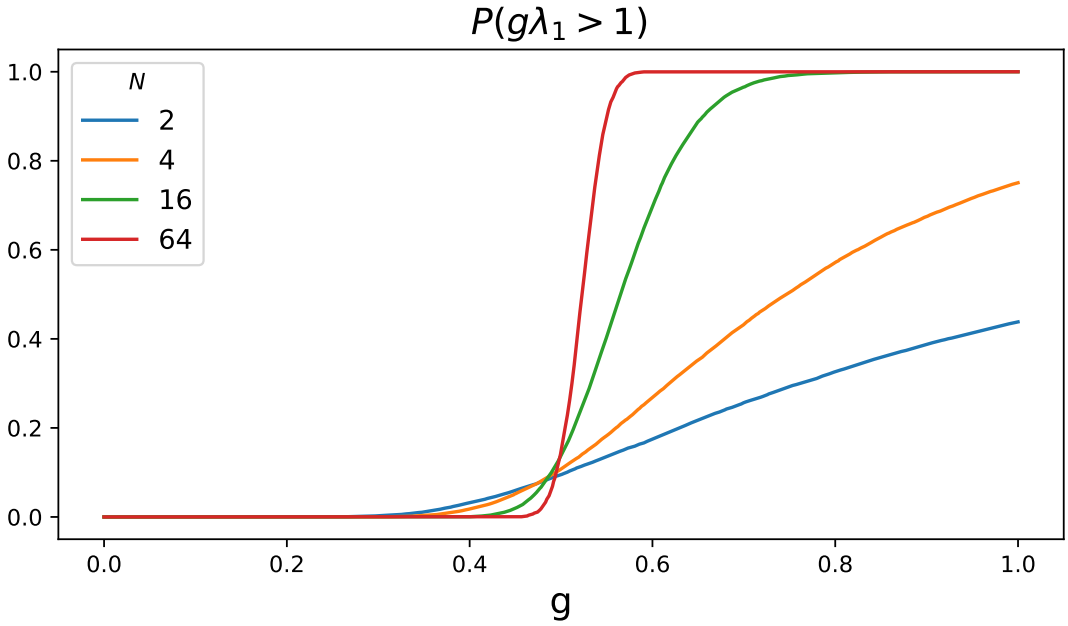


Figure 5.2: Fraction of networks with $g\lambda_1 > 1$ as a function of g for different network dimensions N

becomes more gradual and shifted to larger values of g (Fig. 5.2).

A nonzero constant input biases the potential towards direction of the input. When the origin is unstable, a sufficiently strong input eliminates minima opposite the input direction. In the case of a pulsed input as in 5.4, the network dynamics follows the gradient of $E(\boldsymbol{S}, a\widetilde{\mathbf{e}_1})$ for a time Δt before switching to the dynamics given by $E(\boldsymbol{S}, 0)$. This means that the dynamics approaches two different fixed points during its dynamics. If Δt is sufficiently long, the network will asymptotically reach a minimum of $E(\boldsymbol{S}, a\widetilde{\mathbf{e}_1})$ before approaching a minimum of $E(\boldsymbol{S}, 0)$.

5.3 Role of the nonlinearity in fixed point deviation

The stabilization of the dynamics incurred by the nonlinearity introduces a cost: the nonlinear component of E distorts the quadratic form away from the origin. While the system will flow towards unstable eigenvalues near the origin, this flow is distorted when

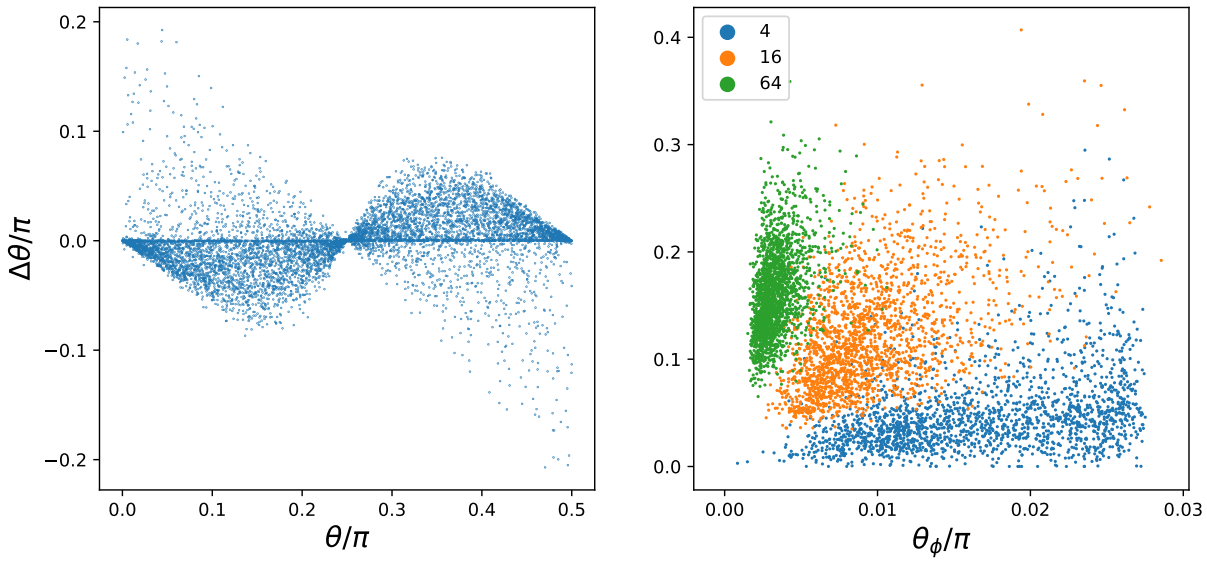


Figure 5.3: Left: Angular deviation dependence on $\mathbf{e}_1$ orientation θ for different realizations of J for $N = 2$. The scatter is due to fluctuations in the eigenvalues. Right: Deflection angles for $N = 4, 16, 64$ and $g = 2$ as a function of θ_ϕ. In both plots, there are a small number of outliers not shown.

the nonlinear term becomes significant. This distortion effected by the nonlinearity can be best understood by considering the $N = 2$ case. In two dimensions, the eigenvector $\mathbf{e}_1$ lies on the unit circle, which we parameterize by the angle θ. The competition between the quadratic potential and the nonlinear potential causes the stable fixed points of the network to deviate from θ by an angle $\Delta\theta$. We plot this deviation for $I = 0$ as a function of θ in Fig. 5.3, for $g = 2$ and different realizations of J. The average deviation is minimized when $\mathbf{e}_1$ and $\phi(\mathbf{e}_1)$ are parallel, and the largest average deviation occurs near $\theta \approx \pi/8$. This behavior is a consequence of the component-wise application of the nonlinearity. The distortion increases with g, which controls the steepness of the nonlinearity.

For general N, the eigenvector matrix e_{ij} of J is uniformly (Haar) distributed on $O(N)$, and $\mathbf{e}_1$ is drawn from the matrix ensemble are uniformly distributed on the $N - 1$ sphere [45]. In analogy to the $N = 2$ case, the smallest angular deviations $\Delta\theta$ occur when the angle θ_ϕ between $\phi(\mathbf{e}_1)$ and $\mathbf{e}_1$ is small. See Fig. 5.3. In general, the angular deviation

depends on the orientation of all the eigenvectors, but the orientation of $\mathbf{e}_1$ places an upper and lower bound on $\Delta\theta$. The deviation is zero only when $\mathbf{e}_1$ has k components $e_{i1} = 0$, and $N - k$ components $e_{i1} = \pm 1/\sqrt{k}$; in other words, when $\mathbf{e}_1$ is aligned with a vertex, edge, face, etc of the hypercube.

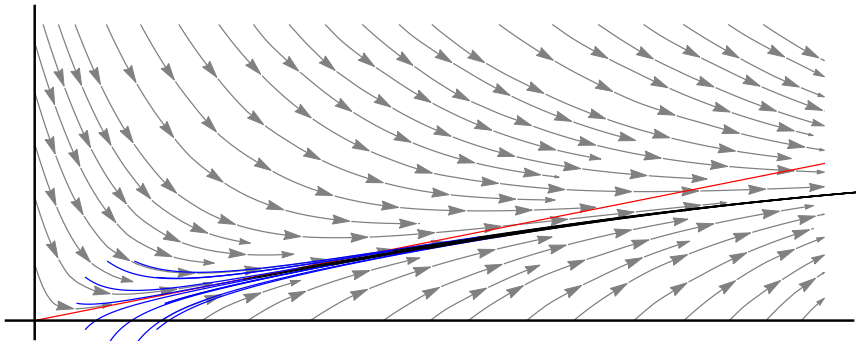
5.4 Numerical simulations and optimization

We compute the optimal performance, averaged over realizations of J and $\boldsymbol{\eta}$, as a function of the network parameters g, a, Δt, and the 'processing time' $T_0 = T - \Delta t$. Surprisingly, the optimal performance is approximately independent of the pulse length Δt (see Fig. 5.9 for a demonstrative example). In the interest of choosing solutions that minimize the overall time T, We therefore only consider the case $\Delta t = 0$ in what follows, where the estimate of $\mathbf{e}_1$ appears only in the initial condition $\mathbf{x}(0)$.

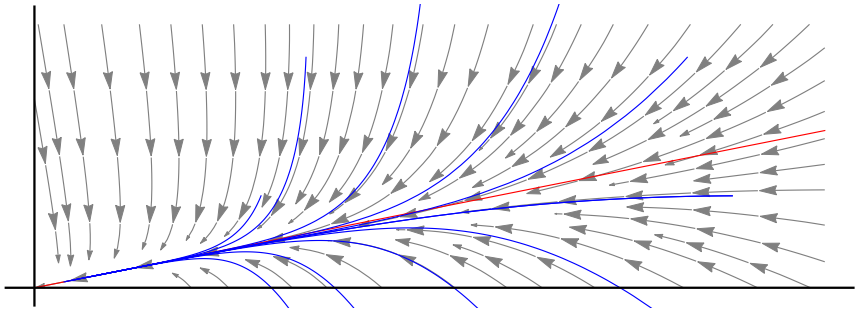
First, we investigate the dependence of the optimal parameters on the processing time T_0. In Fig. 5.5, we plot the optimal performance $\langle \mathcal{P} \rangle$ as a function of T_0. The optimal values of $g^*(T_0)$ and $a^*(T_0)$ are shown in Figs 5.6, 5.7. We now characterize the optimal strategies for small and large T_0.

For small T_0, the optimal strategy has large gain g and small input magnitude a. A large value of g means that the nonlinearity is sharp, and as a result, the fixed points are highly distorted away from $\mathbf{e}_1$. However, this distortion is mitigated by using small input weights a. This means that the network activity starts in the linear regime. Further, the large value of g has another advantage: it speeds up the dynamics. With such a short processing time, it is advantageous to run the dynamics for as long as it takes to span the approximately linear region, and then read out $\mathbf{x}$ before the vector field deviates from $\mathbf{e}_1$ (Fig. 5.4a).

For large T_0, the optimal gain g is slightly above 1/2, and the input magnitude



(a) Optimal strategy for small processing time T_0



(b) Optimal strategy for large processing time T_0

Figure 5.4: Optimal strategies for small and large processing times

a is large. For large a, the initial conditions have large components. When g is close to $1/2$, the network is close to its transition between having a single fixed point at zero for $g\lambda_1 < 1$, and having multiple nonzero fixed points for $g\lambda_1 > 1$. Near the origin, along $\mathbf{e}_1$, the dynamics is slow, and this strategy utilizes the large processing time to evolve along this line (Fig. 5.4b).

From Fig. 5.5, we see that intermediate values of T_0 yield maximum performance, meaning that there is a limit to the performance benefits brought on by additional processing time. For stable networks, with a single stable fixed point the origin, the performance decreases if $|\mathbf{x}|$ is of order σ_t, because it becomes corrupted by the stochastic noise $\boldsymbol{\xi}$. This limit is reached when a is of order $\sigma_t \log T_0$. a must therefore increase exponentially with T_0 to avoid stochastic corruption. But for large a this has no benefit: far from the origin, the nonlinearity is saturated, and the network dynamics is highly distorted.

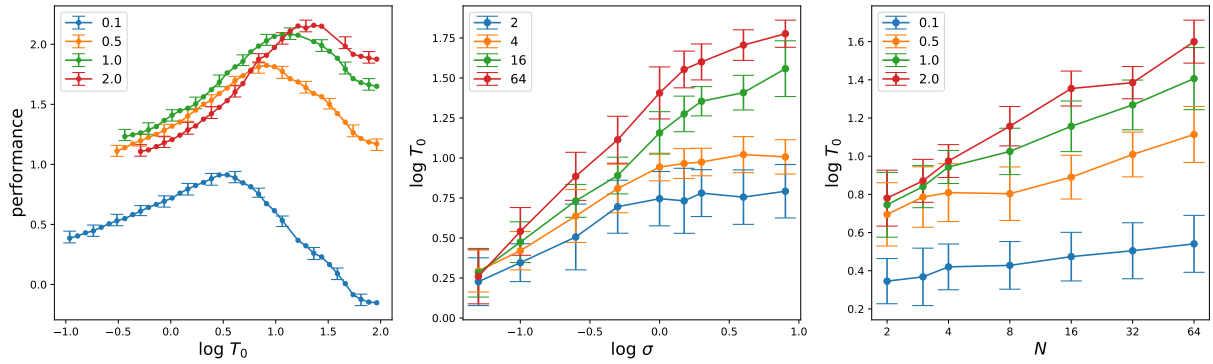


Figure 5.5: Left : performance as a function of T_0 for $N = 16$ and various values of σ. Right : Optimal value of T_0 for each σ, N

Now we discuss the optimal processing time T_0. The optimal processing time increases with the corruption level σ (see Fig. 5.5). For larger σ, the initial conditions are more scattered around $\mathbf{e}_1$, so it takes longer on average to evolve those points towards $\mathbf{e}_1$. We also find that T_0 increases with N, possibly because it takes longer for the dynamics to evolve the same distance for larger N. Fig. 5.6 indicates that large N ensembles have an approximately equal mixture of stable and unstable networks, whereas smaller networks have a greater proportion of stable networks.

The performance is a non-monotonic function of σ, with the best performance at $\sigma \sim \mathcal{O}(1)$. We can understand this by recalling that the performance is defined relative to the accuracy of the input. For small σ, the estimate of $\mathbf{e}_1$ is already accurate sufficiently accurate, and it is difficult for the network to improve on this estimate. For large σ, the estimate is poor, and the network has difficulty evolving towards $\mathbf{e}_1$. In particular, the network may start in the basin of attraction of another eigenvector $\mathbf{e}_k$ with eigenvalue $\lambda_k > 1$, or it may start in the basin of attraction for $-\mathbf{e}_1$. For intermediate σ, the network maximizes its usefulness towards improving the estimate.

Chapter 5, in part, is currently being prepared for submission for publication of the material. Kuczala and Sharpee. The dissertation author was the primary investigator and author of this paper.

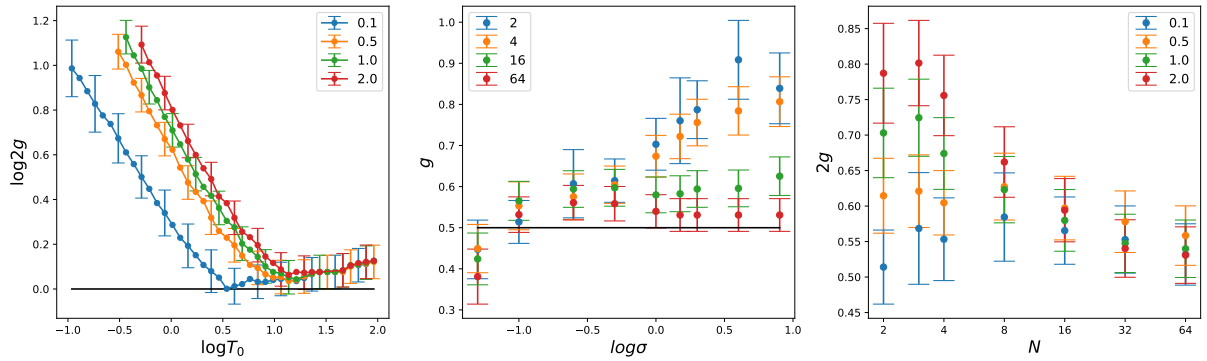


Figure 5.6: Left: Optimal gain g^* for $N = 16$ for various values of σ. Middle, Right: optimal value of g for each σ, N.

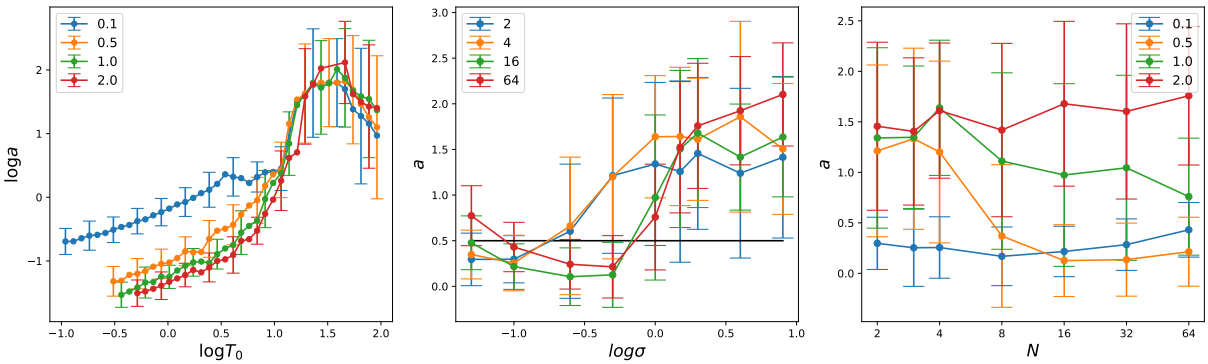


Figure 5.7: Left: Optimal input amplitude a^* for $N = 16$ for various values of σ. Middle, Right: optimal a^* for each σ, N.

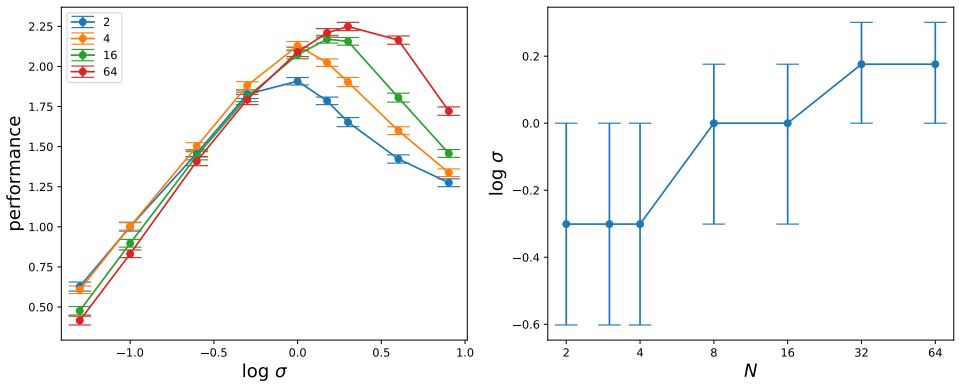


Figure 5.8: Left: best performance for each N as a function of σ. Right: Optimal value of σ for each N

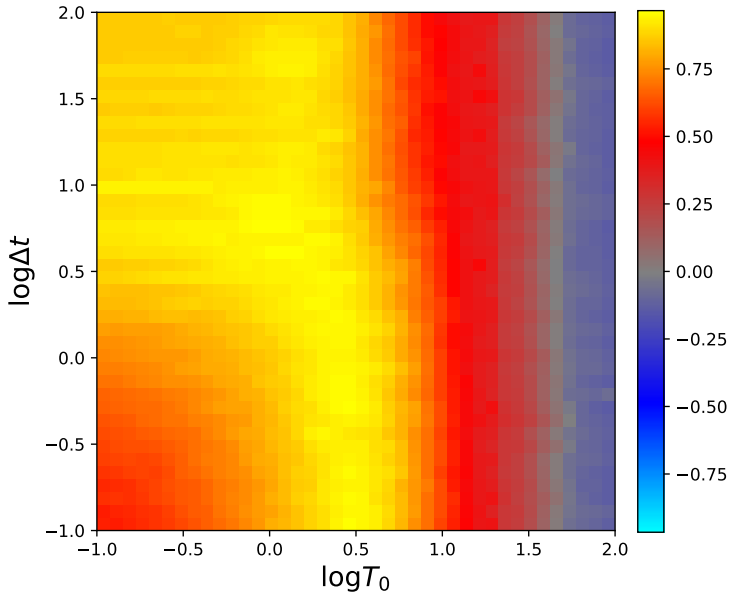


Figure 5.9: Peformance as a function of T_0 and Δt for $N = 16, \sigma = 0.1$. For any Δt, there is a choice of T_0 which yields comparable performance. We may therefore choose $\Delta t = 0$.

Chapter 6

Conclusions

We conclude by discussing potential extensions of the work covered in this dissertation, and then reflecting on the applicability of physics principles to new domains of nature.

Our results on correlated matrix ensembles (2.2) can be extended to more general correlation structures, such as correlations between arbitrary blocks, clusters, or motifs, by summing diagrams with higher-order vertices. Our results can also be extended to the case of nonzero mean as in [1], however, the number of self-consistent equations scales with N rather than M, and we have not yet found a way to simplify these equations[1]. The diagrammatic technique can be used to study further quantities of interest such as eigenvalue correlations [38], eigenvector correlations [22, 37], and linear dynamics not captured by the eigenvalues [1].

It would be interesting to strengthen the relationship between the presence of oscillations in the nonlinear dynamics and the gap of the eigenvalue spectrum (Fig. 3.3). It is likely straightforward to extend the approach of [43] to multiple populations – analyzing small oscillations in the quiescent phase, where the dynamics is approximately linear.

[1] There is a nontrivial matrix inversion in Eq. 2.39.

The appearance of non-zero fixed points has thus far eluded analytic treatment. In principle, the number of fixed points can be calculated using the Kac-Rice formula [72, 9]. To this end, we attempted a mean-field description of the distribution of fixed points, inspired by Merav, Sompolinsky and Abbott [62], but this formulation did not accurately capture the fixed point statistics.

The results of Chapter 4 on the Fisher memory curve for block-structured networks can likely be straightforwardly applied to the continuous rate dynamics (3.1) and more general input signals [56]. It would be interesting to determine if oscillations facilitate or inhibit information retention in strongly bipartite networks. However, the generalization of this calculation to connectivity matrices with block-structured correlations is likely nontrivial. In this case, the distribution of activities $x_i(t)$ is no longer Gaussian, and consequently the Fisher information is not necessarily tractable.

There are a number potential extensions of the results of Chapter 5. One such extension is studying the network's response to changes in connections and its ability to adapt to the evolution of the dominant eigenvector. When the elements of the connectivity matrix undergo Brownian motion, the eigenvalues evolve according to the Brownian motion of a Coulomb gas [25]. The behavior of the eigenvectors is also known [7].

The qualitative behavior of the optimal network parameters $g, a, T, \Delta t$ holds also for piecewise linear firing rates. A Fokker-Planck treatment of the dynamics of the network ensemble may be tractable in this case. However a fundamental difficulty arises from the tension between the component-wise application of the nonlinearity and the spherically symmetric distribution of the dominant eigenvector. We would also like to generalize the results of Chapter 5 to asymmetric connectivity matrices J. When J is asymmetric, the eigenvectors are no longer real; it is no longer generally meaningful for the network activity to converge to the dominant eigenvector. Instead, we may ask when the network activity most closely converges to the subspace corresponding to the rightmost eigenvalue pair.

Taking a broader view, I frequently wonder to what extent the ideals held dear in physics regarding phase transitions, criticality, and universality hold true in the wider world, whose complexity greatly exceeds that of the highly symmetric systems commonly considered by physicists. Dynamical mean field theory and statistical mechanics are chiefly concerned with equilibrium systems, yet the living and nonliving systems that we would like to think of as performing "computations" are far from equilibrium. To what degree can we maximize the entropy of our descriptions of these systems, and appropriately apply equilibrium assumptions? To what extent can we speak generally of non-equilibrium systems? What aspects of nature and computation are addressable and intelligible only on a case-by-case basis?

Appendix A

Deriving the boundary of the eigenvalue distribution

Here we derive the self-consistent equations (2.51, 2.52) for the boundary of the eigenvalue distribution.

A.1 Boundary of eigenvalue distribution without covariance

We first show that the eigenvalue density (2.48) for J with independent elements ($\tau_{ij} = 0$) has support on the disk with radius $R = \sqrt{\lambda_1(K)}$, where $\lambda_1(K)$ is the largest eigenvalue of the matrix $K_{ij} \equiv g_{ij}^2/N$. There are two solutions to the self-consistent equations (2.46) in the limit $\epsilon \to 0$: a trivial solution, with all $a_i = d_i = 0$, and a non-trivial solution, with all $a_i, d_i > 0$ [1]. The trivial solution corresponds to the region where $\rho(r) = 0$ [28, 27]. Indeed, we see that when $a_i = d_i = 0$, all $q_i = r^2$. Then by (2.46) $c_i = 1/z$, and therefore $\rho(r) = 0$ by (2.48).

[1] It is not hard to show that if just one a_i or d_i is zero, all are zero

Now consider the region where $\rho \neq 0$, where all the a_i and d_i are nonzero. Then, combining (2.46) and (2.47) for d_i in the $\epsilon \to 0$ limit yields

$$q_i d_i = \frac{1}{N} \sum_j g_{ij} d_j. \tag{A.1}$$

We determine the radius R of the boundary by finding where the two solutions match. Assuming continuity of the a_i and d_i, then as $d_i \to 0^+$ as we approach the boundary, all the $q_i \to R^2$. Then, in the limit, (A.1) indicates that d is an eigenvector of $K_{ij} = g_{ij}^2/N$ with eigenvalue R^2. Furthermore, since K and d have only positive entries, R^2 must be the largest eigenvalue $\lambda_1(K)$ of K by the Perron-Frobenius theorem. Thus, the boundary of the eigenvalue distribution has radius $R = \sqrt{\lambda_1(K)}$. A nearly identical argument shows $K_{mn} = g_{mn}^2 f_n$ for the block structured case. This result was previously presented in [3] and [4], and a similar argument was used in [1] for the case of matrices with non-zero mean. However, previous analyses do not hold when J has covariant elements.

A.2 Boundary of eigenvalue distribution with covariance

Now we show that when $\tau_{ij} \neq 0$, the boundary of the eigenvalue distribution satisfies (2.51) and (2.52). Again, we have $a_i, d_i \neq 0$ on the support of the eigenvalue distribution, and $a_i = d_i = 0$ otherwise. Plugging the trivial solution into (2.49), the c_i now satisfy

$$c_i = (z - \sum_j \tau_{ij} g_{ij} g_{ji} c_j)^{-1}. \tag{A.2}$$

Now, approaching the boundary from the inside as before, in the limit $d_i \to 0^+$,

$$d_i = \sum_j |c_i|^2 g_{ij}^2 d_j. \tag{A.3}$$

where the c_i satisfy (A.2) in the limit. Since all the $d_i > 0$, this means that d is the Perron-Frobenius eigenvector of the matrix $K_{ij} = |c_i|^2 g_{ij}^2$ with eigenvalue 1. This means that the points z on the boundary satisfy

$$\lambda_1(K) = 1 \tag{A.4}$$

where $\lambda_1(K)$ is the largest modulus eigenvalue of K. Together, (A.2) and (A.4) determine the points z that lie on the boundary of the eigenvalue distribution. We have found that these equations can be solved efficiently as follows: First we write $z = re^{i\theta}$ and fix θ. Then, to find the r satisfying (A.4), we use a root finding algorithm: at each step of the root finding algorithm, we iterate (A.2) to find the $c_i(z)$.

Appendix B

Dynamical mean field theory

Here, we derive mean field equations for the firing rate network (3.1) using dynamical mean field theory (DMFT) [20]. See [23, 43] for similar derivations of the single population case.

B.1 Single population

We start by writing the equations of motion (3.1) as a path integral:

$$Z = \int \prod_i Dx_i(t)\delta\left[(\partial_t + 1)x_i(t) - \sum_j J_{ij}S_j(t)\right] \tag{B.1}$$

where the delta function merely picks out the $x(t)$ that satisfy 3.1. Here we have defined $S_i(t) \equiv \phi[x_i(t)]$. Now we plug in the Fourier representation of the δ function:

$$Z = \int \prod_i Dx_i(t)D\hat{x}_i(t) \exp\left[-\int dt\hat{x}_i(t)\left((\partial_t + 1)x_i(t) - \sum_j J_{ij}S_j(t)\right)\right] \tag{B.2}$$

Here, the $\hat{x}_i$ are imaginary-valued. We will follow the convention throughout this section that hatted variables are imaginary-valued. The integral over t is analogous to

summing over all variables in a multidimensional Fourier transform (these are 'indexed' over t). Throughout, we will drop constant normalization factors.

Now we average over the quenched disorder J_{ij}. In static systems, it would be necessary to introduce replicas at this point, but they are not necessary in dynamical mean field theory. One way to think about this is that the time variable acts like a replica [20].

The average $\bar{Z}$ over J is

$$\bar{Z} = \int \prod_{ij} dJ_{ij} \exp\left(-\frac{N}{2g^2(1-\rho^2)}\mathrm{tr}(JJ^T - \rho J^2)\right) Z(J) \tag{B.3}$$

This is just a Gaussian integral over the J_{ij}, yielding

$$\begin{aligned}
\bar{Z} &= \int \left(\prod_i Dx_i(t) D\hat{x}_i(t)\right) e^{-\mathcal{S}_0} \\
&\times \exp\left[\frac{g^2}{2N}\sum_{ijkl}\left(\int dt \hat{x}_i(t) S_j(t)\right)(\delta_{ik}\delta_{jl} + \rho\delta_{il}\delta_{jk})\left(\int dt' \hat{x}_k(t') S_l(t')\right)\right] \\
&= \int Dx D\hat{x} e^{-\mathcal{S}_0} \exp\left[\frac{g^2}{2N}\int dt dt' \sum_{ij}\left(\hat{x}_i(t) S_j(t) \hat{x}_i(t) S_j(t') + \rho \hat{x}_i(t) S_j(t) \hat{x}_j(t') S_i(t')\right)\right]
\end{aligned} \tag{B.4}$$

with $\mathcal{S}_0 = \sum_j \int dt \hat{x}_j(t)(\partial_t + 1) x_j(t)$.

The average over disorder has introduced a four-body interaction. We write these more compactly by introducing the notation

$$a \cdot b \equiv \sum_j a_j(t) b_j(t') \tag{B.5}$$

reminiscient of replica theory, so we can write $\bar{Z}$ as

$$\bar{Z} = \int Dx D\hat{x} e^{-\mathcal{S}_0} \exp\left[\frac{g^2}{2N}\int dt dt' \big((\hat{x}\cdot\hat{x})(S\cdot S) + \rho(\hat{x}\cdot S)(S\cdot\hat{x})\big)\right] \tag{B.6}$$

Now we introduce the order parameters $Q_1(t,t'), Q_2(t,t'), Q_3(t,t'), Q_4(t,t')$ into the path integral with δ functions:

$$\bar{Z} = \int DxD\hat{x}DQe^{-\mathcal{S}_0}\delta(NQ_1 - \hat{x}\cdot\hat{x})\delta(NQ_2 - S\cdot S)\delta(NQ_3 - \hat{x}\cdot S)\delta(NQ_4 - S\cdot\hat{x}) \\ \times \exp\left[\frac{Ng^2}{2}\int dtdt'(Q_1Q_2 + \rho Q_3Q_4)\right] \tag{B.7}$$

Each δ function can be written as a Fourier integral as follows:

$$\int DQ_k\delta(NQ_k - a\cdot b) \propto \int DQ_kD\hat{Q}_k \exp\left[\int dtdt'\hat{Q}_k(NQ_k - a\cdot b)\right] \tag{B.8}$$

This lets us write the partition function in the form

$$\bar{Z} = DQD\hat{Q}e^{-NF(Q,\hat{Q})} \tag{B.9}$$

where

$$\begin{aligned}
\mathcal{F} &= \mathcal{S}_1(Q) + \mathcal{S}_2(Q,\hat{Q}) + \mathcal{S}_3(\hat{Q}) \\
\mathcal{S}_1 &= -\frac{g^2}{2}\int dtdt'(Q_1Q_2 + \rho Q_3Q_4) \\
\mathcal{S}_2 &= \int dtdt'\sum_{k=1}^{4}\hat{Q}_kQ_k \\
\mathcal{S}_3 &= -\frac{1}{N}\log\int DxD\hat{x}\exp\left[-\mathcal{S}_0(x,\hat{x}) - \mathcal{S}_4(\hat{Q},x,\hat{x})\right] \\
\mathcal{S}_4 &= -\int dtdt'\left[\hat{Q}_1(\hat{x}\cdot\hat{x}) + \hat{Q}_2(S\cdot S) + \hat{Q}_3(\hat{x}\cdot S) + \hat{Q}_4(S\cdot\hat{x})\right]
\end{aligned} \tag{B.10}$$

Now, for large N we can integrate over the Q_k and $\hat{Q}_k$ by saddle point. Extremizing the

action $\mathcal{F}$ with respect to the Q_k and $\hat{Q}_k$ in (B.10)

$$\hat{Q}_1 = \frac{g^2}{2}\ Q_2; \qquad Q_1 = \frac{1}{N}\langle \hat{x}\cdot\hat{x}\rangle_{x,\hat{x}} \tag{B.11}$$

$$\hat{Q}_2 = \frac{g^2}{2}\ Q_1; \qquad Q_2 = \frac{1}{N}\langle S\cdot S\rangle_{x,\hat{x}} \tag{B.12}$$

$$\hat{Q}_3 = \frac{g^2\rho}{2} Q_4; \qquad Q_3 = \frac{1}{N}\langle \hat{x}\cdot S\rangle_{x,\hat{x}} \tag{B.13}$$

$$\hat{Q}_4 = \frac{g^2\rho}{2} Q_3; \qquad Q_4 = \frac{1}{N}\langle S\cdot \hat{x}\rangle_{x,\hat{x}} \tag{B.14}$$

Substituting these values in for the $\hat{Q}_k$ and Q_k, and dropping overall factors independent of x and $\hat{x}$ gives us

$$\begin{aligned}\bar{Z} &= \int DxD\hat{x}e^{-\mathcal{S}_0} \\ &\times \exp\left[-\int dtdt'\frac{g^2}{2}\sum_k \left(C(t,t')\hat{x}_k(t)\hat{x}_k(t') + \rho R(t,t')\hat{x}_k(t)S_k(t') + \rho R(t',t)S_k(t)\hat{x}_k(t')\right)\right]\end{aligned} \tag{B.15}$$

where we have defined [20]

$$\begin{aligned}\hat{C}(t,t') &\equiv \frac{1}{N}\sum_i \langle \hat{x}_i(t)\hat{x}_i(t')\rangle_{x,\hat{x}} = 0 \\ C(t,t') &\equiv \frac{1}{N}\sum_i \langle S_i(t)S_i(t')\rangle_{x,\hat{x}} \\ R(t',t) &\equiv \frac{1}{N}\sum_i \langle \hat{x}_i(t)S_i(t')\rangle_{x,\hat{x}}\end{aligned} \tag{B.16}$$

The last two terms in the sum are symmetric with respect to t and t', so they can be combined. Furthermore, the sum factors out of the exponential as a product, allowing us to factor $\bar{Z} = (Z_{MF})^N$ and consider the partition function Z_{MF} over the mean fields $x(t)$

and $\hat{x}(t)$:

$$Z_{MF} = \int DxD\hat{x}e^{-\mathcal{S}_{0,MF}} \exp\left[-g^2 \int dtdt' \left(\frac{1}{2}C(t,t')\hat{x}(t)\hat{x}(t') + \rho R(t,t')\hat{x}(t')S(t')\right)\right] \tag{B.17}$$

Now we couple $\hat{x}$ to a new variable η via a Hubbard transformation:

$$\begin{aligned} Z_{MF} = \int DxD\hat{x}D\eta \exp\Bigg\{ & \int dt\,[-\hat{x}(t)(1+\partial_t)x(t) + \hat{x}(t)\eta(t)] \\ & - \int dtdt' \left[\frac{1}{2g^2}\eta(t)C^{-1}(t,t')\eta(t') + g^2\rho R(t,t')\hat{x}(t)S(t')\right]\Bigg\} \end{aligned} \tag{B.18}$$

and finally undo the Fourier transformation over $\hat{x}$ to get back a delta function:

$$Z_{MF} = \int \prod_i DxD\eta\delta\left[(\partial_t + 1)x(t) - g^2\rho \int dtR(t,t')S(t') - \eta(t)\right] P(\eta) \tag{B.19}$$

where

$$P(\eta) = \exp\left[-\frac{1}{2}\int dtdt'\eta(t)\frac{1}{g^2}C^{-1}(t,t')\eta(t')\right] \tag{B.20}$$

In other words, $\eta(t)$ is a gaussian random variable with $\langle\eta(t)\rangle = 0$ and $\langle\eta(t)\eta(t')\rangle = g^2C(t,t')$, and $x(t)$ satisfies the stochastic differential equation

$$\dot{x}(t) = -x(t) + g^2\rho\int dtR(t,t')S(t') + \eta(t) \tag{B.21}$$

Introducing a generating functional, we have the useful relation [20]

$$\begin{aligned}
\langle S(t)\eta(t')\rangle &= \left\langle S(t)\frac{\partial}{\partial j(t')}\exp\left[\int dt''\eta(t'')j(t'')\right]\right\rangle\Bigg|_{j=0} \\
&= \left\langle S(t)\frac{\partial}{\partial j(t')}\exp\left[\int dt''\eta(t'')j(t'')\right]\right\rangle\Bigg|_{j=0} \\
&= \left\langle S(t)\frac{\partial}{\partial j(t')}\exp\left[g^2(jCj+jC\hat{x}+\hat{x}Cj)\right]\right\rangle\Bigg|_{j=0} \\
&= g^2\left\langle S(t)\int dt''C(t',t'')\hat{x}(t'')\right\rangle \\
&= g^2\int dt''C(t',t'')R(t,t'')
\end{aligned} \tag{B.22}$$

where we have integrated over η in the third line, and used the shorthand $jCj \equiv \int dtdt' j(t)C(t,t')j(t')$

B.2 Multiple populations

Now we consider the mean field dynamics of a block-structured network with covariance (2.27). The calculation proceeds similarly to the single-population case. A notable difference is that instead of defining $(a \cdot b)$ according to (B.5), we define

$$(a \cdot b)_m \equiv \sum_{i\in m} a_i(t)b_i(t') \tag{B.23}$$

Then introducing $4M$ order parameters $Q_\cdot^m$ using delta functions of the form $\delta(Nf_mQ_k^m - (a\cdot b)_{mk})$, expressions such as

$$Q_2^m(t,t') = \frac{1}{f_mN}\sum_{i\in m}\langle S_i(t)S_i(t')\rangle_{x,\hat{x}} \tag{B.24}$$

we write the partition function in the form $\bar{Z} = DQD\hat{Q}e^{-NF(Q,\hat{Q})}$, where

$$\begin{aligned}
\mathcal{F} &= \mathcal{S}_1(Q) + \mathcal{S}_2(Q,\hat{Q}) + \mathcal{S}_3(\hat{Q}) \\
\mathcal{S}_1 &= -\frac{1}{2}\int dtdt' \sum_{mn} f_m(G_{mn}Q_1^mQ_2^n + T_{mn}Q_3^mQ_4^n) \\
\mathcal{S}_2 &= \int dtdt' \sum_m f_m \sum_{k=1}^4 \hat{Q}_k^m Q_k^m \\
\mathcal{S}_3 &= -\frac{1}{N}\log\int DxD\hat{x}\exp\left[-\mathcal{S}_0(x,\hat{x}) - \mathcal{S}_4(\hat{Q},x,\hat{x})\right] \\
\mathcal{S}_4 &= -\int dtdt' \sum_m \left[\hat{Q}_1^m(\hat{x}\cdot\hat{x})_m + \hat{Q}_2^m(S\cdot S)_m + \hat{Q}_3^m(\hat{x}\cdot S)_m + \hat{Q}_4^m(S\cdot\hat{x})_m\right]
\end{aligned} \tag{B.25}$$

and

$$G_{mn} = g_{mn}^2 f_n, \quad T_{mn} = g_{mn}g_{nm}\tau_{mn}f_n \tag{B.26}$$

Extremizing F yields the conditions

$$\begin{aligned}
\hat{Q}_1^m &= \frac{1}{2}\sum_n G_{mn}Q_2^n; \qquad Q_1^m = \frac{1}{f_m N}\langle(\hat{x}\cdot\hat{x})_m\rangle_{x,\hat{x}} \\
\hat{Q}_2^m &= \frac{1}{2}\sum_n G_{mn}Q_1^n; \qquad Q_2^m = \frac{1}{f_m N}\langle(S\cdot S)_m\rangle_{x,\hat{x}} \\
\hat{Q}_3^m &= \frac{1}{2}\sum_n T_{mn}\,Q_4^n; \qquad Q_3^m = \frac{1}{f_m N}\langle(\hat{x}\cdot S)_m\rangle_{x,\hat{x}} \\
\hat{Q}_4^m &= \frac{1}{2}\sum_n T_{mn}\,Q_3^n; \qquad Q_4^m = \frac{1}{f_m N}\langle(S\cdot\hat{x})_m\rangle_{x,\hat{x}}
\end{aligned}$$

Defining

$$\begin{aligned}
\hat{C}_m(t,t') &\equiv \frac{1}{f_m N}\sum_{i\in m}\langle\hat{x}_i(t)\hat{x}_i(t')\rangle_{x,\hat{x}} \stackrel{!}{=} 0 \\
C_m(t,t') &\equiv \frac{1}{f_m N}\sum_{i\in m}\langle S_i(t)S_i(t')\rangle_{x,\hat{x}} \\
R_m(t,t') &\equiv \frac{1}{f_m N}\sum_{i\in m}\langle S_i(t)\hat{x}_i(t')\rangle_{x,\hat{x}}
\end{aligned} \tag{B.27}$$

the partition function now has the form

$$\bar{Z} = \int DxD\hat{x} \exp\left[\int dtdt' \sum_m \sum_{i\in m} \left(\frac{1}{2}\left(\sum_n G_{mn}C_n(t,t')\right)\hat{x}_i(t)\hat{x}_i(t') + \left(\sum_n T_{mn}R_n(t,t')\right)\hat{x}_i(t)S_i(t')\right)\right] \tag{B.28}$$

which can be factored as $\bar{Z} = \prod_m (Z_m)^{f_m N}$, where

$$Z_m = \int Dx_m D\hat{x}_m \exp\left[\int dtdt' \left(\frac{1}{2}\left(\sum_n G_{mn}C_n(t,t')\right)\hat{x}_m(t)\hat{x}_m(t') + \left(\sum_n T_{mn}R_n(t,t')\right)\hat{x}_m(t)S_m(t')\right)\right] \tag{B.29}$$

Using Hubbard and Fourier transformations in analogy to the single population case yields the mean field equations

$$\dot{x}_m(t) = -x_m(t) + \sum_n T_{mn} \int dt' R_n(t,t')S_m(t') + \eta_m(t) \tag{B.30}$$

$$\langle \eta_m(t)\eta_n(s)\rangle = \delta_{mn}C_m^\eta(t,s) \tag{B.31}$$

$$C_m^\eta(t,s) \equiv \sum_p G_{mp}C_p(t,s) \tag{B.32}$$

$$R_m(t,s) = \int dt' [C_m^\eta]^{-1}(s,t')\langle S_m(t)\eta_m(t')\rangle \tag{B.33}$$

where the $\eta_m(t)$ are gaussian, and $[C_m^\eta]^{-1}(t,t')$ is the convolutional inverse of $C_m^\eta(t,t')$

When $\tau_{mn} = 0$, the mean field equations reduce to $\dot{x}_m(t) = -x_m(t) + \eta_m(t)$ as in [5].

Appendix C

Mean field calculation of signal amplification

Here, we compute the evolution of the mean squared signal amplitude appearing in Eq. 4.13 of Chapter 4. From (4.6),

$$\begin{aligned}
\left\langle \left(\frac{\partial \mu_m(t)}{\partial \theta} \right)^2 \right\rangle_J &= \left\langle \left[\frac{\partial}{\partial \theta} \langle v_{m(i)}(t)|J\rangle \right]^2 \right\rangle_J \\
&= \left\langle \frac{\partial}{\partial \theta^a} \langle v_m^a(t)|J\rangle \frac{\partial}{\partial \theta^b} \langle v_m^b(t)|J\rangle \right\rangle_J \\
&= \frac{\partial^2}{\partial \theta^a \partial \theta^b} \langle v_m^a(t) v_m^b(t)\rangle \\
&= \frac{\partial^2}{\partial \theta^a \partial \theta^b} (w_m^2 \theta^a \theta^b \delta_{t,t_0} + q_m^{ab}(t))
\end{aligned} \tag{C.1}$$

where a and b denote replicas and $q_m^{ab}(t,s) \equiv \langle x_m^a(t) x_n^b(s) \rangle$ is given from mean field theory by

$$\begin{aligned} q_m^{ab}(t,s) &= \sigma^2 \delta_{ab} \delta_{ts} + \sum_n g_{mn}^2 f_n \langle S_n^a(t) S_n^b(s) \rangle_n \\ &= \sum_n G_{mn} C_n^{ab}(t,s) \end{aligned} \tag{C.2}$$

where

$$C_n^{ab}(t,s) \equiv \langle S_n^a(t) S_n^b(s) \rangle_n \tag{C.3}$$

From [69], for normally distributed random variables z^a with $\langle z^a z^b \rangle_z = q^{ab}$,

$$\begin{aligned} \partial^c \langle \phi^a \phi^b \rangle_z &= \vec{a} \cdot \partial^c \vec{\Theta} \\ \vec{a} &\equiv (\langle \phi'^a \phi^b \rangle_z, \langle \phi^a \phi'^b \rangle_z, \langle \phi''^a \phi^b \rangle_z, \langle \phi^a \phi''^b \rangle_z, \langle \phi'^a \phi'^b \rangle_z) \\ \vec{\Theta} &\equiv (w\theta^a, w\theta^b, q^{aa}/2, q^{bb}/2, q^{ab}) \end{aligned} \tag{C.4}$$

where we have denoted $\partial^c \equiv \frac{\partial}{\partial \theta^c}$ and $\phi^a \equiv \phi(w\theta^a + z^a)$ With $\theta = 0$, the expectation values of odd functions vanish, and

$$\partial^c \langle \phi^a \phi^b \rangle_z = \langle \phi''^a \phi^b \rangle_z \frac{\partial^c q^{aa}}{2} + \langle \phi^a \phi''^b \rangle_z \frac{\partial^c q^{bb}}{2} + \langle \phi'^a \phi'^b \rangle_z \partial^c q^{ab} \tag{C.5}$$

This yields a recurrence relation for the derivative of order parameters

$$\begin{aligned} \partial^a C_m^{aa}(t+1) &= \sum_n G_{mn} \partial^a C_n^{aa}(t) \left[\langle \phi''^a \phi^a \rangle_m + \langle \phi'^a \phi'^a \rangle_m \right] \\ \partial^c C_m^{ab}(t+1) &= \sum_n G_{mn} \Bigg[\langle \phi^a \phi''^a \rangle_m \frac{\partial^c C_n^{aa}(t)}{2} \\ &\quad + \langle \phi^b \phi''^b \rangle_m \frac{\partial^c C_n^{bb}(t)}{2} + \langle \phi'^a \phi'^b \rangle_m \partial^c C_n^{ab}(t) \Bigg] \end{aligned} \tag{C.6}$$

At equilibrium, $\partial^a C^{aa}$, $\partial^a C^{ab}$, and $\partial^b C^{ab}$ vanish if the following conditions are satisfied:

$$\rho\left(\left[\langle\phi''^a\phi^a\rangle_m + \langle\phi'^a\phi'^a\rangle_m\right]G_{mn}\right) < 1 \tag{C.7}$$

$$\rho(\langle\phi'^a\phi'^b\rangle_m G_{mn}) < 1 \tag{C.8}$$

where $\rho(A_{mn})$ denotes the spectral radius of matrix A. Assuming that these conditions are met, [69] also tell us that

$$\partial^a\partial^b\langle\phi^a\phi^b\rangle_z = \langle\phi'^a\phi'^b\rangle_z(\partial^a\partial^b q_t^{ab} + w^2) \tag{C.9}$$

so that (C.3)

$$\partial^a\partial^b C_m^{ab}(t+1) = \langle\phi'^a\phi'^b\rangle_m(\partial_k^a\partial_l^b q_m^{ab}(t+1) + w_m^2\delta_{t,t_0}) \tag{C.10}$$

Using (C.2), this yields a recurrence relation for the time evolution of $\partial^a\partial^b q_m^{ab}(t+1)$:

$$\partial^a\partial^b q_m^{ab}(t+1) = \sum_n G_{mn}\langle\phi'^a\phi'^b\rangle_n(\partial^a\partial^b q_n^{ab}(t) + w_n^2\delta_{t,t_0}) \tag{C.11}$$

Note that for $t < t_0$, the pulse has not yet started, so the derivative is zero. Then for $t = t_0$, we have

$$\partial^a\partial^b q_m^{ab}(t_0+1) = \sum_n G_{mn}\langle\phi'^a\phi'^b\rangle_n w_n^2 \tag{C.12}$$

Each succesive iteration for $t > t_0$ amounts to matrix multiplication by $M_{mn} \equiv G_{mn}\langle\phi'^a\phi'^b\rangle_n$. We obtain

$$\partial^a\partial^b q_{m,t+1}^{ab} = \sum_n (M^{t-t_0+1})_{mn} w_n^2\Theta(t-t_0) \tag{C.13}$$

and finally

$$\left\langle\left(\frac{\partial\mu_m(t)}{\partial\theta}\right)^2\right\rangle_J = \sum_n (\mathbf{M}^{t-t_0})_{mn} w_n^2\Theta(t-t_0-1) \tag{C.14}$$

Bibliography

[1] Yashar Ahmadian, Francesco Fumarola, and Kenneth D Miller. Properties of networks with partially structured and partially random connectivity. *Physical Review E*, 91(1):012820, 2015.

[2] Maximino Aldana, Enrique Balleza, Stuart Kauffman, and Osbaldo Resendiz. Robustness and evolvability in genetic regulatory networks. *Journal of Theoretical biology*, 245(3):433–448, 2007.

[3] Johnatan Aljadeff, David Renfrew, and Merav Stern. Eigenvalues of block structured asymmetric random matrices. *Journal of Mathematical Physics*, 56(10):103502, 2015.

[4] Johnatan Aljadeff, David Renfrew, Marina Vegué, and Tatyana O. Sharpee. Low-dimensional dynamics of structured random networks. *Phys. Rev. E*, 93:022302, Feb 2016.

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

[6] Stefano Allesina, Jacopo Grilli, György Barabás, Si Tang, Johnatan Aljadeff, and Amos Maritan. Predicting the stability of large structured food webs. *Nature communications*, 6, 2015.

[7] Romain Allez and Jean-Philippe Bouchaud. Eigenvector dynamics under free addition. *Random Matrices: Theory and Applications*, 03(03):1450010, 2014.

[8] Philip W Anderson. Absence of diffusion in certain random lattices. *Physical review*, 109(5):1492, 1958.

[9] Jean-Marc Azaïs and Mario Wschebor. *Level sets and extrema of random processes and fields*. John Wiley & Sons, 2009.

[10] ZD Bai. Circular law. *The Annals of Probability*, 25(1):494–529, 1997.

[11] Lionel Barnett, Joseph T Lizier, Michael Harré, Anil K Seth, and Terry Bossomaier. Information flow in a kinetic ising model peaks in the disordered phase. *Physical review letters*, 111(17):177203, 2013.

[12] John Beggs and Nicholas Timme. Being critical of criticality in the brain. *Frontiers in Physiology*, 3:163, 2012.

[13] John Berkowitz and Tatyana Sharpee. Decoding neural responses with minimal information loss. *bioRxiv*, page 273854, 2018.

[14] Nils Bertschinger and Thomas Natschläger. Real-time computation at the edge of chaos in recurrent neural networks. *Neural Comp.*, 16(7):1413–1436, 2004.

[15] E Brézin, S Hikami, and A Zee. Universal correlations for deterministic plus random hamiltonians. *Physical Review E*, 51(6):5442, 1995.

[16] E Brézin and A Zee. Correlation functions in disordered systems. *Physical Review E*, 49(4):2588, 1994.

[17] Edouard Brézin, Claude Itzykson, Giorgio Parisi, and Jean-Bernard Zuber. Planar diagrams. *Communications in Mathematical Physics*, 59(1):35–51, 1978.

[18] Zdzisław Burda, Jerzy Jurkiewicz, and Bartłomiej Wacław. Spectral moments of correlated wishart matrices. *Physical Review E*, 71(2):026111, 2005.

[19] Emmanuel J Candes and Terence Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE transactions on information theory*, 52(12):5406–5425, 2006.

[20] Tommaso Castellani and Andrea Cavagna. Spin-glass theory for pedestrians. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(05):P05012, 2005.

[21] Bruno Cessac. Increase in complexity in random neural networks. *Journal de Physique I*, 5(3):409–432, 1995.

[22] J.T. Chalker and B Mehlig. Eigenvector statistics in non-hermitian random matrix ensembles. *Physical Review Letters*, 81(16):3367, 1998.

[23] A Crisanti and H Sompolinsky. Path integral approach to random neural networks. *Physical Review E*, 98(6):062120, 2018.

[24] Brian DePasquale, Christopher J Cueva, Kanaka Rajan, G Sean Escola, and LF Abbott. full-force: A target-based method for training recurrent networks. *PloS one*, 13(2):e0191527, 2018.

[25] Freeman J Dyson. A brownian-motion model for the eigenvalues of a random matrix. *Journal of Mathematical Physics*, 3(6):1191–1198, 1962.

[26] Alan Edelman, Eric Kostlan, and Michael Shub. How many eigenvalues of a random matrix are real? *Journal of the American Mathematical Society*, 7(1):247–267, 1994.

[27] Joshua Feinberg. Non-hermitian random matrix theory: summation of planar diagrams, the single-ringtheorem and the disc–annulus phase transition. *Journal of Physics A: Mathematical and General*, 39(32):10029, 2006.

[28] Joshua Feinberg and A Zee. Non-hermitian random matrix theory: method of hermitian reduction. *Nuclear Physics B*, 504(3):579–608, 1997.

[29] Surya Ganguli, Dongsung Huh, and Haim Sompolinsky. Memory traces in dynamical systems. *Proceedings of the National Academy of Sciences*, 105(48):18970–18975, 2008.

[30] Felix A Gers, Jürgen Schmidhuber, and Fred Cummins. Learning to forget: Continual prediction with lstm. 1999.

[31] Matthieu Gilson, Anthony N. Burkitt, David B. Grayden, Doreen A. Thomas, and J. Leo van Hemmen. Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks iv. *Biological Cybernetics*, 101(5):427–444, 2009.

[32] Jean Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *Journal of Mathematical Physics*, 6(3):440–449, 1965.

[33] VL Girko. Circle law. *Theory Probab. Appl*, 29:694–706, 1984.

[34] Clayton Haldeman and John M Beggs. Critical branching captures activity in living neural networks and maximizes the number of metastable states. *Physical review letters*, 94(5):058101, 2005.

[35] John J Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the national academy of sciences*, 81(10):3088–3092, 1984.

[36] Herbert Jaeger. The echo state approach to analysing and training recurrent neural networks-with an erratum note. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, 148(34):13, 2001.

[37] Romuald A Janik, Wolfgang Nörenberg, Maciej A Nowak, Gábor Papp, and Ismail Zahed. Correlations of eigenvectors for non-hermitian random-matrix models. *Physical Review E*, 60(3):2699, 1999.

[38] Romuald A Janik, Maciej A Nowak, Gábor Papp, and Ismail Zahed. Non-hermitian random matrix models. *Nuclear Physics B*, 501(3):603–642, 1997.

[39] AM Khorunzhii. Eigenvalue distribution of large random matrices with correlated entries. , , , 3(1/2):80–101, 1996.

[40] Chris G Langton. Computation at the edge of chaos: phase transitions and emergent computation. *Physica D: Nonlinear Phenomena*, 42(1-3):12–37, 1990.

[41] Lorenzo Livi, Filippo Maria Bianchi, and Cesare Alippi. Determination of the edge of criticality in echo state networks through fisher information maximization. *IEEE transactions on neural networks and learning systems*, 29(3):706–717, 2017.

[42] Satya N Majumdar and Massimo Vergassola. Large deviations of the maximum eigenvalue for wishart and gaussian random matrices. *Physical review letters*, 102(6):060601, 2009.

[43] Daniel Martí, Nicolas Brunel, and Srdjan Ostojic. Correlations between synapses in pairs of neurons slow down dynamics in randomly connected neural networks. *Phys. Rev. E*, 97:062314, Jun 2018.

[44] Robert M May. Will a large complex system be stable? *Nature*, 238:413–414, 1972.

[45] Elizabeth Meckes. Approximation of projections of random vectors. *Journal of Theoretical Probability*, 25(2):333–352, 2012.

[46] Daniel Miner and Jochen Triesch. Plasticity-driven self-organization under topological constraints accounts for non-random features of cortical synaptic wiring. *PLoS computational biology*, 12(2):e1002581, 2016.

[47] Melanie Mitchell, Peter Hraber, and James P Crutchfield. Revisiting the edge of chaos: Evolving cellular automata to perform computations. *arXiv preprint adap-org/9303003*, 1993.

[48] L Molgedey, J Schuchhardt, and HG Schuster. Suppressing chaos in neural networks by noise. *Phys. Rev. Lett.*, 69:3717–3719, 1992.

[49] Thierry Mora and William Bialek. Are biological systems poised at criticality? *Journal of Statistical Physics*, 144(2):268–302, 2011.

[50] Brendan K Murphy and Kenneth D Miller. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron*, 61(4):635–648, 2009.

[51] Leonid A Pastur. On the spectrum of random matrices. *Theoretical and Mathematical Physics*, 10(1):67–74, 1972.

[52] Kanaka Rajan and L F Abbott. Eigenvalue spectra of random matrices for neural networks. *Phys. Rev. Lett.*, 97(18):188104, 2006.

[53] Kanaka Rajan, L F Abbott, and Haim Sompolinsky. Stimulus-dependent suppression of chaos in recurrent neural networks. *Phys. Rev. E*, 82(1):011903, 2010.

[54] Timothy Rogers, Stefano Allesina, and Jacopo Grilli. Modularity and stability in ecological communities. *Nature Communications*, 7, 2016.

[55] Andrew M Saxe, James L McClelland, and Surya Ganguli. Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. *arXiv preprint arXiv:1312.6120*, 2013.

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

[57] Pragya Shukla. Random matrices with correlated elements: A model for disorder with interactions. *Physical Review E*, 71(2):026226, 2005.

[58] HJ Sommers, A Crisanti, Haim Sompolinsky, and Y Stein. Spectrum of large random asymmetric matrices. *Physical review letters*, 60(19):1895, 1988.

[59] H Sompolinsky, A Crisanti, and H J Sommers. Chaos in random neural networks. *Phys. Rev. Lett.*, 61:259–262, 1988.

[60] Haim Sompolinsky and Annette Zippelius. Relaxational dynamics of the edwards-anderson model and the mean-field theory of spin-glasses. *Phys Rev. B*, 25(11):6860, 1982.

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

[62] Merav Stern, Haim Sompolinsky, and LF Abbott. Dynamics of random neural networks with bistable units. *Physical Review E*, 90(6):062710, 2014.

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

[64] Gerardus t Hooft. A planar diagram theory for strong interactions. *Nucl. Phys. B*, 72(CERN-TH-1786):461–473, 1973.

[65] T. Tao, Van Vu, and Manjunath Krishnapur. Random matrices: Universality of esds and the circular law. *The Annals of Probability*, 38(5):2023–2065, 2010.

[66] Terence Tao. E pluribus unum: from complexity, universality. *Daedalus*, 141(3):23–34, 2012.

[67] Terence Tao and Van Vu. Random matrices: universality of local eigenvalue statistics. *Acta mathematica*, 206(1):127–204, 2011.

[68] Terence Tao and Van Vu. Random matrices: Universal properties of eigenvectors. *Random Matrices: Theory and Applications*, 1(01):1150001, 2012.

[69] Taro Toyoizumi and L F Abbott. Beyond the edge of chaos: Amplification and temporal integration by recurrent networks in the chaotic regime. *Phys. Rev. E*, 84(5):051908, 2011.

[70] Lloyd N Trefethen and Mark Embree. *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*. Princeton University Press, 2005.

[71] Vinayak and Akhilesh Pandey. Correlated wishart ensembles and chaotic time series. *Phys. Rev. E*, 81:036202, Mar 2010.

[72] Gilles Wainrib and Jonathan Touboul. Topological and dynamical complexity of random neural networks. *Physical review letters*, 110(11):118101, 2013.

[73] Eugene P Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, pages 325–327, 1958.

[74] Fred Wolf. Symmetry, multistability, and long-range interactions in brain development. *Physical review letters*, 95(20):208701, 2005.

[75] Anthony Zee. Law of addition in random matrix theory. *Nuclear Physics B*, 474(3):726–744, 1996.

[76] Anthony Zee. *Quantum field theory in a nutshell*, volume 7. Princeton university press, 2010.