New Results on the Spectral Density of Random Matrices

Timothy Rogers

Department of Mathematics,
King's College London, July 2010

A thesis submitted for the degree of
Doctor of Philosophy: Applied Mathematics

Abstract

This thesis presents some new results and novel techniques in the study of the spectral density of random matrices.

Spectral density is a central object of interest in random matrix theory, capturing the large-scale statistical behaviour of eigenvalues of random matrices. For certain random matrix ensembles the spectral density will converge, as the matrix dimension grows, to a well-known limit. Examples of this include Wigner's famous semi-circle law and Girko's elliptic law. Apart from these, and a few other cases, little else is known. Two factors in particular can complicate the analysis enormously - the introduction of sparsity (that is, many entries being zero) and the breaking of Hermitian symmetry. The results presented in this thesis extend the class of random matrix ensembles for which the limiting spectral density can be computed to include various sparse and non-Hermitian ensembles.

Sparse random matrices are studied through a mathematical analogy with statistical mechanics. Employing the cavity method, a technique from the field of disordered systems, a system of equations is derived which characterises the spectral density of a given sparse matrix. Analytical and numerical solutions to these equations are explored, as well as the ensemble average for various sparse random matrix ensembles.

For the case of non-Hermitian random matrices, a result is presented giving a method

for exactly calculating the spectral density of matrices under a particular type of perturbation. The techniques of the so-called simple approach to Hermitian random matrix theory are then adapted for the non-Hermitian case, yielding novel results for the spectral density of sums and products of deterministic and random matrices.

Acknowledgements

I am deeply grateful to all those people who have encouraged my academic development, providing collaboration, discussion, advice, inspiration and promotion. First among these is my supervisor, Isaac Pérez Castillo. The list must also include Reimer Kühn, Koujin Takeda, Conrad Pérez Vicente, Ton Coolen, Nick Trefethen, Raj Rao, Brian Davies, Gernot Akemann, Adriano Barra, Francesco Massucci, Matt Urry, Johannes van Baardewijk and the other members of the Disordered Systems group at King's College London.

On a personal note, I would like to thank my family and friends for their support and patience over the last three years, particularly Vanessa, Mum and Dad, and the King's College London Mountaineering Club.

List of publications

Cavity approach to the spectral density of sparse symmetric random matrices
T. Rogers, I. Pérez Castillo, R. Kühn and K. Takeda,
Physical Review E **78** 031116 (2008)

Cavity approach to the spectral density of non-Hermitian sparse matrices
T. Rogers and I. Pérez Castillo, Physical Review E **79** 012101 (2009)

Spectral density of random graphs with topological constraints
T. Rogers, I. Pérez Castillo, C. Pérez Vicente, and K. Takeda,
Journal of Physics A **43** 195002 (2010)

Universal sum and product rules for random matrices
T. Rogers, arXiv preprint 0912.2499 (2009)
(accepted to Journal of Mathematical Physics)

Contents

1

Introduction

1.1 Random matrix theory

Although random matrices had been considered earlier by Wishart and contemporaries in the context of multivariate statistics [Wis28], it was the work of the Hungarian physicist Eugene Wigner in the 1950's that lay the foundations for the vast field of study now known as Random Matrix Theory (RMT). Seeking to understand the excitation spectra of heavy nuclei, whose interactions are sufficiently complicated cause an analytical impasse, Wigner proposed to instead study a model in which the Hamiltonian is drawn at random (whilst maintaining the appropriate symmetries). This paradigm shift was described by Dyson, a great early proponent of RMT, as "a new kind of statistical mechanics in which we renounce the exact knowledge not of the state of the system but of the nature of the system itself" [Dys62a].

Over the last half century, RMT has established itself as a cornerstone of modern theoretical physics, with problems in quantum physics [GMGW98], finance [PBL05] and number theory [KS03] amongst its innumerable applications. Thorough treatments of the fundamentals and history of the subject are given in the book of Mehta [Meh91] and reviews of Forrester *et al* [FSV03] and Edelman and Rao [EN05].

1.1.1 Random matrix ensembles and eigenvalue distributions

A random matrix ensemble is defined by the joint probability density function (JPDF) P of the matrix entries, usually given for square matrices of arbitrary size N. Our main interest will be in behaviour in the limit $N \to \infty$, and we assume the dependence of P upon N to be understood.

Very often, research in RMT is concerned with ensembles of matrices endowed with certain symmetry properties, in particular the body of work on Hermitian matrices vastly outweighs that on the more general non-Hermitian case. The initial focus of this section will be on Hermitian RMT, with discussion of the non-Hermitian case saved for the end of the section.

The most famous and well-studied Hermitian random matrix ensembles are the Gaussian orthogonal (GOE), unitary (GUE) and symplectic (GSE) ensembles. Their names are derived from Dyson's categorisation of random matrix ensembles by symmetry class [Dys62b], corresponding to the possible symmetries of a random Hamiltonian[1].

As an example, let us consider the orthogonal case. GOE matrices are real and symmetric, with JPDF of entries

$$P(A) = (\pi)^{-N(N-1)/4}(2\pi)^{-N/2} \exp\left(-\frac{1}{4}\mathrm{Tr}\, AA^T\right). \tag{1.1}$$

Since, for real and symmetric matrices, $\mathrm{Tr}\, AA^T = \sum_{i,j=1}^{N} A_{ij}^2$, we see that the probability density function factorises, and entries of A are independent Gaussian random variables. It is also clear to see that this density is invariant under orthogonal transformations $A \mapsto OAO^T$, from which the name GOE is derived. Similarly, the GUE/GSE ensembles have complex/quaternionic entries, with JPDFs invariant under unitary/symplectic transformations.

Beyond these three ensembles, many other random matrix models have been well

[1] See [Zir10] for a modern review of Dyson's 'threefold way'.

studied over the years, with notable examples including Dyson's (non-Hermitian) circular ensembles, the chiral ensembles of quantum chromodynamics (QCD) and the Wishart-type ensembles of multivariate statistics. Other types of ensemble appear on an ad-hoc, model specific basis, such as those involved in studying phenomena in sparsely interacting systems. A large part of this thesis will be concerned with these sparse ensembles.

From its origins in the study of the energy levels of heavy nuclei, a primary concern of RMT is to understand the behaviour of the eigenvalues of a given random matrix ensemble. For the Gaussian ensembles, one can in fact compute the JPDF of eigenvalues for arbitrary matrix size. Continuing with the GOE as an example, changing variables in (1.1) and integrating over unused degrees of freedom (see [Meh91] for details) leads to the JPDF for the eigenvalues of a GOE matrix A,

$$P\left(\lambda_1, \ldots, \lambda_N\right) = \frac{1}{Z_N} \exp\left(-\frac{1}{2}\sum_{i=1}^{N} \lambda_i^2\right) \prod_{i<j} \left|\lambda_i - \lambda_j\right| \ ,$$

where Z_N is a normalising constant. Similar expressions hold for the GUE and GSE, and all the information about the statistical behaviour of the eigenvalues of these ensembles is contained within their JPDF.

Apart from the Gaussian ensembles and some other simple cases however, the full JPDF of eigenvalues of a general random matrix ensemble is likely to be extremely difficult to compute. All is not lost though if one is interested (as we very often are) in behaviour as $N \to \infty$. In this limit, the JPDF of eigenvalues is disregarded in favour of quantities such as the overall density of eigenvalues and correlations between them.

Eigenvalue correlations are often regarded as particularly important, both to theory and application, and have been studied in depth over the years. The main developments in this area are covered in [Meh91], though it is worth making special mention of the curious fact that the form of the eigenvalue correlations is universal across many different random matrix ensembles, often depending only upon symme-

try class, see for example [KHKZ95, BHZ95, DZ96]. Interestingly, the same correlations have also been observed in areas seemingly unrelated to RMT, for example in the distribution of zeros of the Riemann zeta function [Mon73].

Important as eigenvalue correlations are, there is still much to be said about the broader features of the distribution eigenvalues.

1.1.2 Spectral density

The focus of this thesis will be the macroscopic distribution of eigenvalues, captured in the spectral density. For an $N \times N$ matrix X, we label the eigenvalues of the X as $\{\lambda_i^{(X)}\}_{i=1}^N$, and define the following discrete measure known as the empirical spectral density of X:

$$\varrho(\lambda; X) = \frac{1}{N} \sum_{i=1}^{N} \delta\left(\lambda - \lambda_i^{(X)}\right) ,$$

where δ here is a Dirac delta. Generally, the empirical spectral density is a probability measure over the complex plane, however, if the eigenvalues of X are known to be confined to a certain subset of the complex plane then it is convenient to treat the empirical spectral density as a measure on that subset. In particular, all eigenvalues of Hermitian matrices are real, and the spectral density thus defines a real probability measure.

If the matrix X is drawn from some random matrix ensemble, then the empirical spectral density is a random probability measure. For a specified random matrix ensemble it becomes pertinent to ask if we can deduce anything about the empirical spectral densities of random matrices from this ensemble, particularly in the limit $N \to \infty$.

For certain ensembles, the answer is well known. The average empirical spectral

density can be expressed in terms of the JPDF of (unordered) eigenvalues by

$$\begin{aligned}\mathbb{E}\,\varrho(\lambda; A_N) &= \int P(\lambda_1, \ldots, \lambda_N)\,\frac{1}{N}\sum_{i=1}^{N}\delta(\lambda - \lambda_i)\; d\lambda_1 \cdots d\lambda_N \\ &= \int P(\lambda, \lambda_2 \ldots, \lambda_N)\; d\lambda_2 \cdots d\lambda_N\,,\end{aligned}$$

where here and hereafter we use $\mathbb{E}$ to denote the ensemble average, that is, the average over the JPDF of entries. For the Gaussian ensembles, this integral can in fact be computed exactly for any N [Meh91], though the calculation is rather involved. A simpler analysis is possible in the large N limit, for instance by analogy with a Coulomb gas as also given in [Meh91]. Again taking the GOE as an example, one obtains the expression

$$\mathbb{E}\,\varrho(\lambda; A_N) \simeq \frac{1}{2N\pi}\sqrt{4N - \lambda^2}\,\mathbb{I}_{[-2\sqrt{N}, 2\sqrt{N}]}(\lambda)\,, \tag{1.2}$$

where here and hereafter, $\mathbb{I}_\Omega$ denotes the indicator function for the set Ω and we use the symbol $\simeq$ to denote the dominant term in the limit $N \to \infty$. This semi-circular density was observed by Wigner, who in fact demonstrated it to hold not just for the GOE but also for many other random matrix ensembles. Specifically, Wigner showed the following [Wig58]:

Theorem (Wigner's Law)**.** *Let $\{A_N\}$ be a sequence of $N \times N$ random matrices such that for each N the entries of A_N are independent random variables of unit variance, drawn from symmetric distributions with bounded moments. Then for fixed λ,*

$$\lim_{N\to\infty} \mathbb{E}\,\varrho\left(\lambda; A_N/\sqrt{N}\right) = \frac{1}{2\pi}\sqrt{4 - \lambda^2}\,\mathbb{I}_{[-2,2]}(\lambda)\,. \tag{1.3}$$

This result can be thought of as an analogue of the central limit theorem, in that an ever increasing number of random variables of an unknown type combine to produce a known deterministic limit. Matrices satisfying the conditions of Wigner's law are known as Wigner-class. In fact, the conditions given above can be considerably relaxed, with the most general form given by a specification similar to the Lindeberg

condition for the central limit theorem [Pas72]. Moreover, the pointwise in λ convergence of the average of the spectral density stated above may be upgraded to strong convergence in the sense of probability measures.

Winger's original proof proceeded via an analysis of the moments of the spectral density, though other approaches are possible. For a review of the techniques applied to this and related problems, see [Bai99].

1.1.3 The 'simple approach'

One method for computing the spectral density of certain random matrices (including Wigner-class matrices and various functions thereof) will turn out to be particularly relevant to the work contained in this thesis. Techniques introduced by Marčenko and Pastur [MP67] to compute the spectral density of Wishart-type matrices were developed by Pastur and collaborators over many years to form what became known as the simple approach [Pas72, KP93, KKP96, Pas98].

To introduce the method, we first require a few definitions. For a fixed $N \times N$ matrix X, the resolvent R is given by

$$R(z; X) = (X - z)^{-1} . \tag{1.4}$$

This matrix-valued function is defined for all complex numbers z outside of the spectrum of X (known as the resolvent set). We will refer to the normalised trace of the resolvent as the Green's function,

$$G(z; X) = \frac{1}{N} \mathrm{Tr}\, R(z; X) .$$

The Green's function is related to the spectral density of X by the formula

$$G(z; X) = \int \frac{1}{\mu - z} \, \varrho(\mu; X) \, d\mu , \tag{1.5}$$

where the integral runs over the domain of the spectral density (for example $\mathbb{R}$ in the case of Hermitian matrices), proof of this fact is given in Appendix B. In probability theory, the RHS of the above expression is known as the Stieltjes (or Cauchy-Stieltjes) transform of the density ϱ. For Hermitian X, several useful properties are known:

1. The Green's function is a closed analytic map on $\mathbb{C}^+ = \{z : \operatorname{Im} z > 0\}$

2. The empirical spectral density can be recovered from the Green's function by the inverse Stieltjes transform

$$\varrho(\lambda; X) = \lim_{\varepsilon \searrow 0} \frac{1}{\pi} \operatorname{Im} G(\lambda + i\varepsilon; X) \tag{1.6}$$

3. Neglecting the limit $\varepsilon \to 0$ in the above, one obtains a regularised form of the empirical spectral density, in which each delta peak is replaced by a Lorentzian (i.e. a Cauchy probability density) with width parameter ε:

$$\begin{aligned} \varrho_\varepsilon(\lambda; X) &= \frac{1}{\pi} \operatorname{Im} G(\lambda + i\varepsilon; X) \\ &= \frac{1}{\pi} \int \frac{\varepsilon}{\varepsilon^2 + (\lambda - \mu)^2} \, \varrho(\mu; A) \, d\mu \\ &= \frac{1}{\pi N} \sum_{i=1}^{N} \frac{\varepsilon}{\varepsilon^2 + |\lambda_i^{(X)} - \lambda|^2} \end{aligned} \tag{1.7}$$

4. Suppose random matrices $\{X_N\}$ have Green's functions which converge in probability (resp. almost surely) to a function $G(z)$ which is the Stieltjes transform of a probability measure $\rho(\lambda)$. Then the spectral densities $\varrho(\lambda; X_N)$ converge weakly (resp. strongly) to $\rho(\lambda)$

In the work of Pastur and collaborators, straightforward and robust techniques were developed for studying the Green's functions of random matrices in the large N limit, from which the limiting spectral densities can be deduced. Though this method was first applied to Wishart-type matrices in the derivation of the famous Marčenko-Pastur law [MP67], for our purposes it will be instructive to present a brief sketch of

the proof of Wigner's semicircle law using the simple approach, following the notes [BdMK01].

Let $\{A_N\}$ be a sequence of $N \times N$ random matrices such that for each N the entries of A_N are independent random variables of unit variance, drawn from symmetric distributions with bounded moments. We wish to prove the convergence of the spectral densities of $A_N/\sqrt{N}$ to Wigner's law.

Using the shothands $R_N = R(z; A_N/\sqrt{N})$ and $G_N = G(z; A_N/\sqrt{N})$, we begin with the easily checked identity

$$zG_N = \frac{1}{N^{3/2}} \mathrm{Tr}\left[R_N\, A_N\right] - 1\,. \tag{1.8}$$

With straightforward arguments it is possible to establish both that the variance of the Green's function G_N vanishes as $N \to \infty$, a property we will refer to as self-averaging, and that

$$\mathbb{E}\,\frac{1}{N^{3/2}}\,\mathrm{Tr}\big[\,R_N\, A_N\big] + \mathbb{E}\,G_N^2 \to 0 \quad \text{as} \quad N \to \infty\,.$$

Together with (1.8), these facts are enough to obtain the convergence in probability of $G(z; A_N/\sqrt{N})$ to a non-random limit $G(z)$, which must satisfy

$$zG(z) = -G(z)^2 - 1\,. \tag{1.9}$$

Solving for $G(z) \in \mathbb{C}^+$ and applying the inverse Stieltjes transform, one obtains Wigner's semicircular law.

Much more can be determined using this technique. For example, suppose the matrices $\{D_N\}$ are Hermitian and non-random with a known limiting spectral density ρ_D. Then the limiting Green's function $G(z)$ of the sum $A_N + D_N$ is given by the fixed point of what is known as Pastur's equation [Pas72],

$$G(z) = \int \frac{1}{\mu - z - G(z)}\,\rho_D(\mu)\,d\mu\,.$$

1.1.4 Non-Hermitian random matrix theory

As mentioned earlier, non-Hermitian random matrices, with their generally complex eigenvalues, have received a great deal less attention over the years than their Hermitian counterparts. The reasons for this are two-fold: firstly, for the vast majority of physical applications, a Hermitian matrix provides the correct model; secondly, many useful mathematical techniques of RMT (including the simple approach) rely heavily on the symmetry properties of Hermitian matrices and quite simply do not work in their absence.

These drawbacks have not entirely stymied research into non-Hermitian random matrices however, and the list of applications is steadily growing: important examples of applications in physics include the works of Hatano and Nelson on vortex pinning [HN96] and Stephanov on QCD with chemical potential [Ste96].

Once again, the most well understood non-Hermitian random matrix ensembles are Gaussian. Following the proposal of Ginibre [Gin65], we consider matrices in which all the entries are independent identically distributed Gaussian random variables. Referred to as the Ginibre orthogonal (GinOE), unitary (GinUE) and symplectic (GinSE) ensembles, these models are the natural non-Hermitian analogue of the Gaussian ensembles. As in the case of the Hermitian Gaussian ensembles, the JPDF of eigenvalues can be calculated exactly for Ginibre's ensembles. For example, in the GinUE case we have

$$P(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z_N} \exp\left(-\sum_{i=1}^{N} |\lambda_i|^2 \right) \prod_{i<j} |\lambda_i - \lambda_j|^2 \ ,$$

where Z_N is again the required normalising constant, but this time $\lambda_1, \ldots, \lambda_N$ are complex. This formula leads to the following limiting spectral density for suitably normalised GinUE matrices $A_N/\sqrt{N}$,

$$\lim_{N\to\infty} \mathbb{E}\, \varrho\left(\lambda; A_N/\sqrt{N}\right) = \frac{1}{\pi} \mathbb{I}_{[0,1]}\big(|\lambda|\big) \ .$$

In light of the universality of Wigner's law, it was conjectured that this circular distri-

bution should hold universally for random matrices with independent entries. Known as the Circular Law, this conjecture remained an open problem for many years. Important progress was made by Girko [Gir90] and Bai [Bai97], but the strongest and most general version was proved only recently by Tao and Vu [TV08, TV09].

Theorem (Circular Law)**.** *Let $\{A_N\}$ be a sequence of $N \times N$ random matrices such that for each N the entries of A_N are independent random variables of zero mean and unit variance. Then as $N \to \infty$ the spectral densities $\varrho(\lambda; A_N/\sqrt{N})$ converge strongly to the uniform density on the unit disc.*

A generalisation of this law to elliptic distributions has been observed [Gir90, SCSS88], interpolating between the circular and semi-circular densities, and in the 'almost Hermitian' scaling limit these matrices display different behaviour again [FKS97, FS03].

The methods used in Tao and Vu's proof of the Circular Law, as well as those applied in other areas of non-Hermitian RMT, do not immediately resemble those from the Hermitian theory. For non-Hermitian matrices, the Green's function does not provide the same useful regularisation of the spectral density as was exploited in the Hermitian case. Because of this, techniques used in the study of the spectral density of non-Hermitian matrices often involve alternative methods of regularisation, examples include the integral transforms used in [Gir90, Bai97, TV08, TV09] and an analogy with electrostatic potential introduced in [SCSS88] and used frequently thereafter [Kho96, Cha10, GT07a, GT07b, PZ07].

Very often, such methods rely upon the study of Hermitian proxies, of which there are several equally good (and often equivalent) choices. In the series of papers [FZ97a, FZ97b, FSZ01] Feinberg and Zee worked with $2N \times 2N$ block matrices of the form

$$H_\varepsilon(\lambda; X) = \begin{pmatrix} -i\varepsilon & (X-\lambda) \\ (X-\lambda)^\dagger & -i\varepsilon \end{pmatrix}, \tag{1.10}$$

a process they christened 'Hermitization'. Around the same time, Janik, Nowak and

collaborators proposed a similar block extension techniques, obtaining a generalisation of the Green's function with a quaternionic structure [JNPZ97, JNP$^+$97].

1.2 Complex networks

In recent years there has been an explosion of interest from all over the scientific community in the properties of networks. From social science [Sco88] to finance [NS97] and biology [CN85], it is hoped that analysis of the network structures underlying complex real world systems may provide insight into the properties and behaviour of those systems. Well-known examples include scientific collaboration networks, the internet and world wide web and protein interaction networks. Each of these systems can be modelled as a large collection of agents linked together in pairs to form a network (for example, a pair of scientists are linked if they have co-authored a paper).

To treat complex networks mathematically we turn to graph theory and in particular to the study of large random graphs. Much like the heavy nuclei deemed complex enough to warrant a random matrix model, the modern approach to complex networks is to treat them as typical objects from a random ensemble. The book of Dorogovtsev and Mendes [DM02] and reviews of Barabási [AB02] and Newman [New03] all provide introductions to the subject, though the field is expanding in all directions with novel techniques and applications appearing frequently.

1.2.1 Graph theory definitions

A graph is a pair $G = (V, E)$ of sets, such that $E \subset V \times V$. We refer to the elements of V as vertices, and those of E as edges. Unless otherwise stated, all graphs we work with are assumed to be simple, meaning that the edge-sets satisfy the rules $(i, j) \in E \Leftrightarrow (j, i) \in E$ and $(i, i) \notin E$ for all $i, j \in V$. Occasionally we will consider

directed graphs, in which there exist edges $(i, j) \in E$ such that $(j, i) \notin E$.

We will usually take $V = \{1, \ldots, N\}$, and refer to N as the size of the graph. A pair of vertices $i, j \in V$ are said to be neighbours if they are joined by an edge $(i, j) \in E$. The collection of all neighbours of a vertex $i \in V$ is refereed to as its neighbourhood and is denoted ∂i, and the number of neighbours of a vertex $i \in V$ is known as the degree. A glossary of these and all other graph theory terms used in the thesis is provided in Appendix A.

All the information about a graph $G = (V, E)$ is encoded in the adjacency matrix C, an $N \times N$ matrix whose entries are given by

$$C_{ij} = \begin{cases} 1 & \text{if} (i, j) \in E \\ 0 & \text{otherwise.} \end{cases}$$

Some properties of the graph are easily recovered from simple operations on the adjacency matrix, for example, the degree of vertex $i \in V$ is given by

$$k_i(C) = \sum_{j=1}^{N} C_{ij} \, .$$

From the definition we see that the adjacency matrix of a simple graph is real and symmetric and hence has N real eigenvalues. Investigation of the eigenvalues of adjacency matrices (and other related matrices such as the graph Laplacian) forms a field of study in its own right, known as spectral graph theory [Chu97].

The relationship between the topology of a graph and its spectral density is detailed, for example: the graph is bipartite if and only if the spectral density is symmetric; the mean degree of the graph is given by the variance of the spectral density; the maximum degree of a vertex in the graph is at least as big as supremum of the support of the spectral density. The spectral density also has important consequences for the behaviour of random processes defined on the graph, and has been used to characterise the robustness of the graph under attack [WTD$^+$08].

1.2.2 Random graphs

The study of random graphs first appeared in mathematics at around the same time that random matrix theory was gaining momentum amongst the physics community, though early work in the field was purely mathematical and the idea of modelling real world systems using random graphs is perhaps more recent.

Formally speaking, the notion of a random graph reduces to defining a probability measure on the set of all graphs on a given vertex set V; a random graph can then be thought of as a random variable with respect to that measure. For our purposes, it will be much easier to phrase the problem in terms of random matrices by introducing a joint probability density P of the entries of an $N \times N$ adjacency matrix.

The first and simplest example of a random graph was proposed by Gilbert [Gil59], in which each edge is independently chosen to be present or absent with fixed probability p. In this case, the joint probability density of entries of the adjacency matrix is given by[2]

$$P(C) = \prod_{i<j} \left(p\, \delta_{C_{ij},1} + (1-p)\, \delta_{C_{ij},0} \right) ,$$

where here δ is the Kronecker delta. The introduction of this and the closely related model of Erdős and Rényi [ER59] generated enormous interest in random graphs in the combinatorial and probabilistic communities, both as an ingenious tool for creating non-constructive proofs in extremal graph theory, and as fascinating objects of study in their own right [Bol01].

As with random matrices, we will be principally concerned with behaviour in the limit $N \to \infty$ and again we assume the dependence of P (and other related quantities) on N to be understood. As the size of the graphs grow, talking about an entire graph is no longer practical and one may instead wish to extract some statistical information to quantify some aspect of its structure. An important example of such an

[2] By symmetry only the JPDF of entries above the diagonal need be specified.

object is the degree distribution, giving the probability distribution of the degree of a randomly selected vertex. For a specified random graph ensemble, we define the degree distribution as

$$p(k) = \lim_{N\to\infty} \mathbb{E}\, \frac{1}{N} \sum_{i=1}^{N} \delta_{k,k_i(C)} \,,$$

where here $\mathbb{E}$ is used to denote the average over the JPDF of entries of the adjacency matrix. In the case of Gilbert random graphs, it is easy to see that the degree distribution of a finite sized graph is binomial and does not have a limit as $N \to \infty$ if p is fixed. However, taking $p = c/N$, for fixed c, one obtains a random graph ensemble whose degree distribution converges to the Poisson distribution with mean c in the limit $N \to \infty$. Known to the physics community as the Poissonian random graph ensemble, this is one of the simplest random graph ensembles with non-trivial behaviour in the large N limit, and has been extensively studied (see [Bol01] and references therein).

Looking for more detailed information about a large graph (or random graph ensemble), one could consider further statistics such as the likelyhood of finding a connected pair of vertices with degree k and k'. Precisely, we define the degree-degree correlation

$$P(k,k') = \lim_{N\to\infty} \mathbb{E}\, \frac{1}{Nc} \sum_{i,j=1}^{N} \delta_{k,k_i(C)} \delta_{k',k_j(C)} C_{ij} \,.$$

For the Poissonian random graph ensemble, this quantity has the simple factorised form $P(k,k') = (kp(k)/c)(k'p(k')/c)$.

In the large N limit, Poissonian random graphs share with random matrices the property that, despite depending upon an ever-increasing number of random variables, rather a lot can be said with certainty. For example, if $c > 1$ then as the graph grows a giant connected component forms, containing an extensive number of vertices. A particular feature of Poissonian random graphs which will be of central importance to us is the emergence of a tree-like structure in the large N limit. We say that a random graph ensemble is tree-like if a ball of fixed radius about a randomly selected vertex is

a tree with probability approaching one as $N \to \infty$. This property also holds for random regular graphs [Wor99] and random graphs with specified degree distribution [MR95]. As we will see later, the tree-like nature of certain random graph ensembles can provide important simplifications to the analysis of random processes on such graphs.

1.2.3 Modelling complex networks with random graphs

Empirical studies have provided a wealth of information about the structural properties of the networks we might hope to model mathematically. Remarkably, certain features appear in a multitude of different networks. An investigation of the internet [FFF99] in the late 1990s found degree distributions with power-law tails, suggesting a scale-free topology; similar degree distributions have been frequently observed in other networks, including the world-wide-web (which is directed) [BKM$^+$00] and semantic networks [ST05]. Another common feature of real-world networks is known as the small-world phenomenon, describing the property that the typical distance between randomly selected vertices is very much less than in a Poissonian random graph. The small-world phenomenon is practically ubiquitous, having been observed in everything from social interaction [Mil67] to gene regulation [PVSW05].

A well-known heuristic random graph model of real-world networks was proposed by Barabási and Albert [BA99]. In their model a random graph is grown from a small initial seed via a mechanism of expansion and preferential attachment, in which new vertices are added to the graph and edges drawn randomly with a bias towards vertices of higher degree. Though based on very simple assumptions, the model shows many of the key features of real-world networks and has been extensively studied. Much is known about various properties of the model, including typical path lengths, degree correlations and clustering coefficient [AB02]. The spectral density of Barabási-Albert random graphs has been studied numerically [FDBV01] and

shown to be markedly different to that of other random graph ensembles.

Despite the success of heuristic models like that of Barabási and Albert in capturing some common aspects of complex networks, a given real-world complex network may not be a perfect match for any existing heuristic model. It could be argued then that a more phenomenological approach is needed, in which certain statistics of a real-world network can be measured and incorporated into a random graph model. For example, one could consider random graphs with the degree distribution specified, or more involved statistics such as degree-degree correlations and further constraints [BCV08, ACF$^+$09].

Later in the thesis we will analyse the spectral density of such topologically constrained ensembles.

1.2.4 Generating random graphs

In order to perform numerical experiments and compare with the real-world complex networks one is hoping to model, it is crucial to be able to generate random graphs from a given ensemble. Similarly, for our purposes we would like to generate random adjacency matrices corresponding to a random graph ensemble, so as to test our theoretical results for the spectral density. In the simple case of the Poissonian random graph, this is easily achieved by independently choosing each edge to be present or absent with probability c/N. For more complicated or 'realistic' random graph ensembles, the problem is much harder and has been studied by a number of authors, examples include [MW90, BDML06, BD06, ACF$^+$09], to name but a few.

To give a flavour for the techniques involved we describe an algorithm for generating random regular graphs, which we will employ later in numerical experiments. To generate a random graph of size N with each degree equal to a fixed value k, Steger and Wormwald proposed [SW99] the following algorithm:

Algorithm 1

1. Set $V = \{1, \ldots, N\}$, $E = \emptyset$ and $\hat{k} = (k, \ldots, k)$; an N-tuple with each entry k
2. Randomly select distinct non-neighbouring vertices $i, j \in V$ with probability proportional to $\hat{k}_i \hat{k}_j$
3. Add the edge (i, j) to E, reduce $\hat{k}_i$ and $\hat{k}_j$ by 1, and repeat step 2 until no more edges can be added
4. If $\hat{k} \neq 0$ report failure, otherwise output the graph $G = (V, E)$

This algorithm selects approximately uniformly at random from the set of all k-regular graphs of size N, in the sense that the deviation from the uniform distribution is vanishing in the limit $N \to \infty$. Moreover, the expected running time is significantly less that other approaches, being of the order Nk^2 [SW99].

To generate random graphs with a specified degree sequence $\bar{k} = (k_1, \ldots, k_N)$ and degree-degree correlation function $P(k, k')$ is a rather more difficult task. We suggest the following heuristic adaptation of the Steger-Wormald algorithm:

Algorithm 2

1. Set $V = \{1, \ldots, N\}$, $E = \emptyset$ and $\hat{k} = \bar{k}$
2. Randomly select distinct non-neighbouring vertices $i, j \in V$ with probability proportional to $P(k_i, k_j)\hat{k}_i \hat{k}_j$
3. Add the edge (i, j) to E, reduce $\hat{k}_i$ and $\hat{k}_j$ by 1, and repeat step 2 until no more edges can be added
4. If $\hat{k} \neq 0$ report failure, otherwise output the graph $G = (V, E)$

A mathematical analysis of the performance of this algorithm is yet to be undertaken, though experimental results are very encouraging.

1.3 Disordered systems

There are several common themes occurring in the fields of random matrices and random graphs, and nowhere is this more apparent than in the study of condensed matter physics, particularly disordered systems and spin-glasses.

It has been argued that the field of disordered systems has its origins in the vastly influential work of Anderson on electron conductance in lattices with random impurities [And58], in which the phenomenon now known as Anderson localisation was first discovered. In the subsequent decades many different models of dynamical systems with random impurities have be put to use as tools for understanding the behaviour of anything from the functions of the brain [CKS05] to the stock market [Coo07].

1.3.1 Spin-glasses and the replica trick

An important development in the theory of disordered systems came in 1975 with Edwards and Anderson's simple model of certain dilute magnetic alloys featuring unusual critical magnetic behaviour, known as spin-glasses [EA75].

In their model, a system of N dynamical variables (spins) $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)$ are arranged in a lattice, with interactions between neighbouring spins σ_i and σ_j mediated by a bond of strength J_{ij}. The bond strengths are assumed to be independent Gaussian random variables, which are taken to be a fixed realisation of a random sample, so-called quenched disorder. Computation of the disorder averaged free energy $\mathcal{F}$ for

this system reduces to the evaluation of the expression

$$-\frac{1}{\beta}\mathcal{F} = \mathbb{E}\,\log \mathcal{Z}(J)\,, \tag{1.11}$$

where β is the inverse temperature, $\mathbb{E}$ is as usual the disorder average (i.e. over the J_{ij}), and $\mathcal{Z}$ is the partition function

$$\mathcal{Z}(J) = \int \left(\prod_{i=1}^{N} d\sigma_i\right) \exp\left\{\sum_{i<j} C_{ij} J_{ij} \sigma_i \cdot \sigma_j\right\}\,, \tag{1.12}$$

with C being the adjacency matrix of the lattice.

Though the partition function is different for every model, the desire to evaluate expressions of the form (1.11) is a common problem in the theory of disordered systems. In tackling the average of the logarithm in this equation, Edwards and Anderson introduced the replica trick, which was to become a staple tool of researchers in disordered systems for the next thirty years. The essence of the trick is to exchange the order of the average and logarithm through the use of the identity

$$\mathbb{E}\,\log\,\mathcal{Z}(J) = \lim_{n\to 0} \frac{1}{n} \log\,\mathbb{E}\left(\mathcal{Z}(J)\right)^n.$$

To make progress with this expression, the dummy variable n is treated as integer (so that the problem becomes one of n replicas of the original system, each with the same disorder) and the disorder average is calculated in the limit $N \to \infty$.

There are many subtleties of the method, and it has only been made rigorous for a small number of models, for example in the Hopfield model with certain parameter values [Tal98], though the practical successes of the approach are numerous, see for instance [MPV87]. The replica method has also been applied rigorously to the Anderson model [KMP86].

Just a year after its introduction to the problem of spin-glasses, Edwards and Jones applied the replica trick to random matrix theory [EJ76], re-deriving Wigner's semicircular law. Replicas have since been applied frequently to problems in RMT, see

[Kan01] for a review. Unsuccessful attempts to compute the two-point correlation function using replica prompted criticism of the method from Verbaarschot and Zirnbauer [VZ85, Zir99], who argue that it is superseded (at least in RMT) by the methods of supersymmetry. The supersymmetric method has been used extensively in RMT [MF91, FS95, FKS97, GMGW98, Mir00, FS03], for a review of its application to disordered systems see [Efe06]. Restricted to the purpose of computing the spectral density however, the methods of supersymmetry and replica have been demonstrated to be equivalent [FM91].

1.3.2 The cavity method

At its core, the replica method relies upon the hope that the disorder average of the replicated partition function will be tractable, and that the limit $n \to 0$ is then possible and makes physical sense. It has been found to be the case, however, that even for some simple models the expressions resulting from the disorder average are sufficiently complicated to necessitate a rather protracted analysis. For certain models then, it is highly desirable to have a technique to compute thermodynamic quantities without taking the disorder average so early in the analysis, if at all.

The cavity method [MPV86] offers an alternative approach to replicas, by seeking to exploit the topological structure of the underlying network. Suppose one is interested in a generic disordered system in which the dynamical variables reside on the vertices of a graph $G = (V, E)$ and interact in pairs according to the edges of the graph. Following [YFW03], we can model this situation in a fairly general way by considering a vector of spins $\boldsymbol{\sigma}$ with JPDF $P(\boldsymbol{\sigma})$, which factorises into terms $\{\psi_{ij}\}$ associated to the edges of G and $\{\phi_i\}$ associated to the vertices; we assume the form

$$P(\boldsymbol{\sigma}) = \frac{1}{\mathcal{Z}} \prod_{(i,j)\in E} \psi_{ij}(\sigma_i, \sigma_j) \prod_{i \in V} \phi_i(\sigma_i)\,. \tag{1.13}$$

Now consider the same system with the vertex i removed. We write $G^{(i)} = (V^{(i)}, E^{(i)})$

for the subgraph of G induced by the deletion of vertex i, the so-called cavity graph. This superscript notation will also be used for generic objects associated with the cavity graph. In particular, if $\boldsymbol{\sigma}^{(i)}$ is the spin vector with the i^{th} component removed, the JPDF of spins on the cavity graph is given by

$$P^{(i)}(\boldsymbol{\sigma}^{(i)}) = \frac{1}{\mathcal{Z}^{(i)}} \prod_{(j,k)\in E^{(i)}} \psi_{jk}(\sigma_j, \sigma_k) \prod_{j\in V^{(i)}} \phi_j(\sigma_j)\,. \tag{1.14}$$

Writing $P^{(i)}_{\partial i}(\boldsymbol{\sigma}_{\partial i})$ for the joint marginal of spins on the cavity graph whose vertices were neighbours of i in the original graph, it is a straightforward consequence of the form of (1.13) that the single-spin marginals are given by

$$P_i(\sigma_i) = \frac{1}{\mathcal{Z}_i} \int \Big[\prod_{j\in\partial i} d\sigma_j \Big] P^{(i)}_{\partial i}(\boldsymbol{\sigma}_{\partial i}) \left(\phi_i(\sigma_i) \prod_{j\in\partial i} \psi_{ij}(\sigma_i, \sigma_j) \right)\,, \tag{1.15}$$

where $\mathcal{Z}_i = \mathcal{Z}/\mathcal{Z}^{(i)}$. As mentioned earlier, a common feature of many of the most frequently studied random graph ensembles is a tree-like structure in the large N limit. The central idea of the cavity approach is to exploit this structure in order to approximately compute the distributions $P^{(i)}_{\partial i}(\boldsymbol{\sigma}_{\partial i})$, and hence give an approximation to the true marginal distribution at any given vertex.

Suppose that for some $i \in V$ and $j \in \partial i$ we wish to compute the cavity distribution $P^{(j)}_i(\sigma_i)$ of the spin at vertex i in the cavity graph $G^{(j)}$. Let us assume that G is indeed a tree. Removing the vertex j from (1.15) we reach the equation

$$P^{(j)}_i(\sigma_i) = \frac{1}{\mathcal{Z}^{(j)}} \int \Big[\prod_{l\in\partial i\setminus j} d\sigma_l \Big] P^{(i)(j)}_{\partial i\setminus j}(\sigma_{\partial i\setminus j}) \left(\phi_i(\sigma_i) \prod_{l\in\partial i\setminus j} \psi_{il}(\sigma_i, \sigma_l) \right)\,, \tag{1.16}$$

with $\mathcal{Z}^{(j)}_i = \mathcal{Z}^{(i)}/\mathcal{Z}^{(i)(j)}$. Now, since G is a tree, each vertex in $\partial i \setminus j$ resides in a different connected component of the cavity graph $G^{(i)}$, both to each other and to j itself - see Figure 1.1. We can therefore conclude that

$$P^{(i)(j)}_{\partial i\setminus j}(\boldsymbol{\sigma}_{\partial i\setminus j}) = \prod_{l\in\partial i\setminus j} P^{(i)}_l(\sigma_l)\,, \tag{1.17}$$

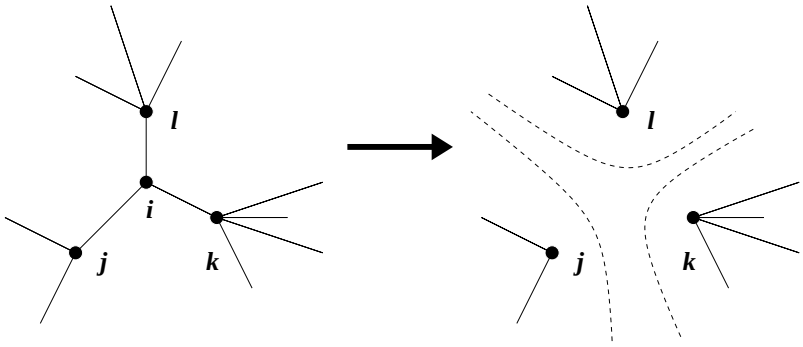


Figure 1.1: Left - part of a tree-like graph G showing the neighbourhood of a vertex i. Right - in the cavity graph $G^{(i)}$, the spins residing on vertices j, k and l are approximately independent.

and thus (1.16) simplifies significantly to

$$P_i^{(j)}(\sigma_i) = \frac{1}{\mathcal{Z}_i^{(j)}}\,\phi_i(\sigma_i) \prod_{l\in\partial i\setminus j} \left(\int d\sigma_l\, P_l^{(i)}(\sigma_l)\, \psi_{il}(\sigma_i,\sigma_l) \right) . \tag{1.18}$$

Similarly, (1.15) becomes

$$P_i(\sigma_i) = \frac{1}{\mathcal{Z}_i}\,\phi_i(\sigma_i) \prod_{j\in\partial i} \left(\int d\sigma_j\, P_j^{(i)}(\sigma_j)\, \psi_{ij}(\sigma_i,\sigma_j) \right) . \tag{1.19}$$

For a system on given graph $G = (V, E)$, there are $2|E|$ equations of the type (1.18), which together define a self-consistent scheme for the cavity distributions $\{P_i^{(j)}\}$. If these equations can be solved, then the true marginal distributions may be computed using (1.19).

Obtaining a solution to the cavity equations for a given graph can potentially be very difficult. The tractability of the problem is greatly improved in the event that the cavity distributions can be parameterised (perhaps approximately) by a finite set of parameters; the cavity equations should then reduce to a set of consistency equations on those parameters, which can, for example, be solved numerically using a belief propagation algorithm.

Cavity equations of the type (1.18) may of course be written down for any graph, tree

or not. How well the solutions (if indeed any can be found) approximate the correct distributions then depends on the degree to which correlations between spins in the cavity graph are present and influence the marginal at a given vertex. One situation in which the cavity approach is known to work well is in large random graphs drawn from a tree-like ensemble, for example, recent work proving the asymptotic correctness of the cavity approach for the Ising model on locally tree-like graphs was presented in [DM08].

The cavity method has been successfully applied to the Anderson model both explicitly in [CGMM$^+$05] and [AF06], and in all but name in [ACTA73]. A large part of this thesis will consider the application of the method to general sparse random matrix ensembles.

1.4 Thesis outline

The main body of the thesis is subdivided thus:

Chapter 2: Spectral density of sparse Hermitian random matrices

A mathematical analogy is introduced in Chapter 2, rephrasing the problem of determining the spectral density of Hermitian random matrices in the language of statistical mechanics. After a brief review of past work on the topic, the cavity method is applied to derive a set of equations from which an approximation to the spectral density of a given sparse matrix may be recovered. The equations are studied both analytically, re-deriving known results, and by numerical methods both for single instances and in the ensemble average for Poissonian graphs.

The main results of this chapter were published in the article [RPKT08].

Chapter 3: Beyond Poissonian graphs

The class of ensembles for which the average spectral density may be computed is expanded in Chapter 3, firstly through the consideration of graphs with correlations between vertex degrees and later through block matrix models for sparse covariance matrices and graphs with a community structure. A number of examples, both analytic and numerical are presented and the relationship between the statistics of the graph and the resulting spectral density is discussed.

The articles [RPKT08] and in particular [RPVT10] contain many of the results of this chapter.

Chapter 4: Spectral density of sparse non-Hermitian random matrices

In Chapter 4, techniques are introduced allowing the application of the cavity method to the problem of determining the spectral density of sparse non-Hermitian random matrices. Analytically, Girko's Elliptic Law is re-derived, along with an apparently new result for random directed regular graphs. Numerical simulations are presented for sparse ensembles breaking the universality class of the Circular Law.

Most of the results in this chapter appear in the paper [RP09].

Chapter 5: Universal sum and product rules for random matrices

In a departure from earlier chapters, the work presented in Chapter 5 concerns universality in the spectral density of full random matrices, both Hermitian and non-Hermitian. A theory is developed which marries the techniques of the simple approach, as discussed earlier in this introduction, to a quaternionic Green's function. Universal results are obtained for the spectral density of random matrices with independent entries when summed or multiplied with non-random matrices.

This chapter is closely based on the article [Rog09].

2

Spectral density of sparse Hermitian random matrices

2.1 Statistical mechanics analogy

The aim of the present chapter is to apply the cavity method as outlined in the introduction to the problem of determining the spectral density of sparse Hermitian random matrices. The first step towards this goal is to map the problem into one phrased in the language of statistical mechanics, to which the cavity method can then be applied.

The link between the spectral density of a matrix and its Green's function was discussed in the introduction. Indeed, in light of the Stieltjes transformation (1.5) and inverse (1.6), to compute the spectral density, it is sufficient to determine the Green's function. We follow the standard route of Edwards and Jones [EJ76] in expressing the Green's function in terms of a multidimensional Gaussian integral. Let A be an arbitrary $N \times N$ Hermitian matrix. We start from the simple identity (proof is given in Appendix B)

$$G(z;A) = \frac{1}{N}\frac{\partial}{\partial z}\log\frac{1}{\det(A-z)}\,. \tag{2.1}$$

For $z \in \mathbb{C}^+$, the eigenvalues of the matrix $i(A - z)$ all have positive real part and thus the determinant can be expressed as a Fresnel-type Gaussian integral

$$\frac{1}{\det(A-z)} = \left(\frac{i}{\pi}\right)^N \int \exp\left(-i\,\boldsymbol{x}^\dagger(A-z)\boldsymbol{x}\right)\, d\boldsymbol{x}\,, \tag{2.2}$$

where $\boldsymbol{x}$ is an N-vector of complex variables and we introduce the shorthand

$$d\boldsymbol{x} = \prod_{i=1}^{N} d\mathrm{Re}x_i\, d\mathrm{Im}x_i\,.$$

The proof of this identity is simple. Since A is Hermitian, we may diagonalise by a unitary transform, writing $i(A - z) = UDU^{-1}$, the identity then follows from the change of variables $\boldsymbol{x} \mapsto U^{-1}\boldsymbol{x}$ and normal Gaussian integration.

Ignoring the imaginary factors, the expression (2.2) is reminiscent of the partition function of a disordered system with continuous dynamical variables, in particular the Gaussian ferromagnetic model of [BK52]. We make this analogy with statistical mechanics concrete by introducing an object resembling a probability density function. Let

$$P(\boldsymbol{x}) = \frac{1}{\mathcal{Z}(z;A)} \exp\left(-i\sum_{i,j=1}^{N} \overline{x}_i (A-z)_{ij} x_j\right)\,, \tag{2.3}$$

where normalisation $\int P(\boldsymbol{x})\, d\boldsymbol{x} = 1$ is enforced through

$$\begin{aligned}\mathcal{Z}(z;A) &= \int \exp\left(-i\sum_{i,j=1}^{N} \overline{x}_i (A-z)_{ij} x_j\right)\, d\boldsymbol{x} \\ &= \left(\frac{i}{\pi}\right)^{-N} \frac{1}{\det(A-z)}\,.\end{aligned} \tag{2.4}$$

Though the function P can take complex values, and hence is manifestly not a probability density, it is sufficiently similar in form to a typical JPDF arising in disordered systems that it can be treated with the same mathematical tools. For convenience, we will switch to the language of statistical mechanics; referring to the integration variables x_i as spins with JPDF $P(\boldsymbol{x})$ and partition function $\mathcal{Z}$. For the sake of simplicity, the dependence of P upon A and z is not made explicit in the notation.

Plugging (2.4) into (2.1), we see that the Green's function of A is related to the partition function simply by

$$G(z;A) = \frac{1}{N}\frac{\partial}{\partial z}\log \mathcal{Z}(z;A)\,, \tag{2.5}$$

and thus the problem of determining the spectral density of A is mapped onto one of computing the partition function $\mathcal{Z}(z;A)$.

Suppose now that the matrix A was drawn from a Hermitian random matrix ensemble. To compute the ensemble average of the spectral density using the above formalism one is required to evaluate the ensemble average of equation (2.5),

$$\mathbb{E}\, G(z;A) = \frac{1}{N}\frac{\partial}{\partial z}\mathbb{E}\,\log \mathcal{Z}(z;A)\,. \tag{2.6}$$

Written in this way, the expressions (2.4) and (2.6) are clearly comparable to equations (1.12) and (1.11), to which Edwards and Anderson applied the replica trick [EA75].

2.2 Past approaches

In 1976, just a year after the introduction of the replica trick, Edwards and Jones applied the same method to the statistical mechanics analogy for spectral density [EJ76]. In that work Wigner's Law for real symmetric random matrices with independent Gaussian entries of zero mean and variance J/N was re-derived, as well as a similar result for the non-zero mean case. This work lay the foundations for numerous further applications of the replica trick to RMT.

Years later, motivated by applications to the behaviour of dilute spin systems, Rodgers and Bray [RB88] investigated the spectral density of a random matrix model with an intensive number of non-zero entries. They studied matrices in which each entry is drawn at random from the distribution

$$p(A_{ij}) = \frac{c}{N}\left(\frac{1}{2}\delta_{A_{ij},1} + \frac{1}{2}\delta_{A_{ij},-1}\right) + \left(1 - \frac{c}{N}\right)\delta_{A_{ij},0}\,.$$

Applying the now well-established replica method, a set of non-linear functional equations were derived which specify the ensemble average spectral density for these types of matrices in the large N limit. A tractable solution to these equations would not be found for many years, though several facts about the spectral density were established, including adherence to Wigner's Law in the limit $c \to \infty$ and the presence of Lifschitz tails extending beyond the support of the semi-circle.

Since then, the spectral density of sparse matrices, and the model of Rodgers and Bray in particular, has been studied in the statistical mechanics formalism by a number of different authors, including [RD90, MF91, FM91, BM99, SC02, Dea02, Küh08]. A typical calculation for the ensemble average replicated partition function might arrive at an expression equivalent to the following functional integral over order parameters ϕ and ψ:

$$\mathbb{E}\left(\mathcal{Z}(z;A)\right)^n = \int \{\mathcal{D}\phi\mathcal{D}\psi\}\, e^{N\,\Phi(\phi,\psi)}\,, \tag{2.7}$$

where

$$\begin{aligned}\Phi(\phi,\psi) = \frac{c}{2}\int \phi(\boldsymbol{x})\phi(\boldsymbol{y})(e^{-i\boldsymbol{x}\cdot\boldsymbol{y}}-1)\,d\boldsymbol{x}\,d\boldsymbol{y} \\ - i\int \phi(\boldsymbol{x})\psi(\boldsymbol{x})\,d\boldsymbol{x} + \log\int e^{iz\boldsymbol{x}\cdot\boldsymbol{x}+i\psi(\boldsymbol{x})}\,d\boldsymbol{x}\,.\end{aligned} \tag{2.8}$$

From here, one might hope to evaluate (2.7) in the limit $N \to \infty$ by steepest descent. This leads to an extremisation problem for the free energy (2.8) which was described by Biroli and Monasson as 'hopeless' [BM99].

Some progress was possible through the introduction of schemes such as the effective medium (EMA) and single defect (SDA) approximations [BM99, SC02], which provide a reasonable fit in some parts of the spectrum for intermediate values of c but do not describe the detail of the spectral density well for low c. The methods of supersymmetry were applied in [RD90, MF91], though equivalence with replica (in the calculation of spectral density, though not in general) was quickly realised [FM91].

An important advance was made in [Dea02], and later independently in [Küh08], in

which it was realised that the order parameters ϕ and ψ should be expressed as a supersposition of Gaussians. In this way, the solution to the previously intractable extremisation problem can be expressed exactly through a single distributional recursion equation for the parameters of these Gaussians.

Work along lines other than the statistical mechanics analogy is of course also possible. Particularly worth mentioning are the approaches of Dorogovtsev *et al* using walk generating functions [DGMS03, DGMS04] and Khorunzhy *et al* [KSV04] and Bordenave and Lelarge [BL10] using moment and resolvent methods.

In what follows, we will revisit the problem of determining the spectral density of sparse random matrices, this time applying the cavity method.

2.3 The cavity method

Consider a sparse Hermitian matrix A, specified by the adjacency matrix C of a graph $G = (V, E)$ and a bond strength matrix J. The entries of A are given by

$$A_{ij} = C_{ij} J_{ij} \, .$$

We plan to use the cavity method to compute an approximation to the spectral density of A. Performing the derivative in (2.5), we see that

$$\begin{aligned} G(z; A) &= \frac{1}{N} \frac{1}{\mathcal{Z}(z; A)} \int \frac{\partial}{\partial z} \exp\left(-i \sum_{i,j=1}^{N} \overline{x}_i (A - z)_{ij} x_j \right) d\boldsymbol{x} \\ &= i \int P(\boldsymbol{x}) \left(\frac{1}{N} \sum_{i=1}^{N} |x_i|^2 \right) d\boldsymbol{x} \\ &= i \, \frac{1}{N} \sum_{i=1}^{N} \left\langle |x_i|^2 \right\rangle \, , \end{aligned} \tag{2.9}$$

where $\langle \cdots \rangle$ denotes the average over the JPDF of spins in the statistical mechanics analogy. To determine the Green's function, it is therefore enough to compute the

single-spin marginal distributions of the system.

Ignoring the fact that the function P is not a probability density, we note that it is of the correct form to apply the method. The starting point of the discussion of the cavity method in the introduction was the JPDF

$$P(\boldsymbol{x}) = \frac{1}{\mathcal{Z}(z;A)} \prod_{(i,j)\in E} \psi_{ij}(x_i, x_j) \prod_{i\in V} \phi_i(x_i)\,,$$

which can be obtained in this case by putting

$$\phi_i(x_i) = \exp\left(iz|x_i|^2\right),$$

and

$$\psi_{ij}(x_i, x_j) = \exp\Big(- i(J_{ij}x_i\overline{x}_j + \overline{J_{ij}}x_j\overline{x}_i)\Big)\,.$$

We intend to apply the cavity method in this setting, approximating the marginal distributions $P_i^{(j)}$ and P_i in a general sparse graph G by those obtained under the assumption that G is a tree. To this end, we introduce distributions $\widetilde{P}_i^{(j)}$ and $\widetilde{P}_i$ which we take to satisfy the cavity equations (1.18) and (1.19) exactly. In the present case, this yields

$$\widetilde{P}_i(x_i) = \frac{1}{\mathcal{Z}_i}\, e^{iz|x_i|^2} \prod_{j\in\partial i} \left(\int \widetilde{P}_j^{(i)}(x_j)\, e^{-i(J_{ij}x_i\overline{x}_j + \overline{J_{ij}}x_j\overline{x}_i)}\, dx_j \right),$$

and

$$\widetilde{P}_i^{(j)}(x_i) = \frac{1}{\mathcal{Z}_i^{(j)}}\, e^{iz|x_i|^2} \prod_{l\in\partial i\setminus j} \left(\int \widetilde{P}_l^{(i)}(x_l)\, e^{-i(J_{il}x_i\overline{x}_l + \overline{J_{il}}x_l\overline{x}_i)}\, dx_l \right). \tag{2.10}$$

Further progress can be made by moving to a parameter dependent representation for the cavity distributions $\widetilde{P}_i^{(j)}$. This is always possible (see [SPH05] for a general derivation), however, for general systems it is quite possible that an infinite number of parameters may be required. Fortunately, in the present case it is straightforward to see that they a single complex parameter will suffice to characterise each single-spin marginal. Integral P over $N-1$ variables, one finds the marginal at i to be parameterised by the i^{th} entry of the diagonal of the resolvent $R(z;A)$, specifically,

$$P_i(x_i) = \frac{i}{\pi R(z;A)_{ii}} \exp\left(-\frac{i}{R(z;A)_{ii}}|x_i|^2 \right). \tag{2.11}$$

In light of this, and the Gaussian form of the cavity equations (2.10), we may surmise that the same form of parameterisation will hold for the approximate single-spin marginals $\widetilde{P}_i^{(j)}$ and $\widetilde{P}_i$. We introduce parameters Δ_i and $\Delta_i^{(j)}$, writing

$$\widetilde{P}_i(x_i) = \frac{i}{\pi\Delta_i} \exp\left(-\frac{i}{\Delta_i}|x_i|^2\right) ,$$

$$\widetilde{P}_i^{(j)}(x_i) = \frac{i}{\pi\Delta_i^{(j)}} \exp\left(-\frac{i}{\Delta_i^{(j)}}|x_i|^2\right) . \tag{2.12}$$

For normalisation, we require that each, Δ_i and $\Delta_i^{(j)}$ be confined to the upper half of the complex plane. These numbers of course depend upon both the matrix A and the complex number z, though again we choose not make this dependence explicit in the notation. Comparison with (2.11), suggests that Δ_i should be interpreted as an approximation to $R(z;A)_{ii}$, with equality in the case that G is a tree.

Returning this parameterisation to the cavity equations (2.10) and completing the integration yields

$$\begin{aligned}\widetilde{P}_i^{(j)}(x_i) &= \frac{i}{\pi\Delta_i^{(j)}} e^{iz|x_i|^2} \prod_{l\in\partial i\setminus j} \left(\frac{i}{\pi\Delta_l^{(i)}} \int e^{-i\frac{|x_l|^2}{\Delta_l^{(i)}} - i(J_{il}x_i\overline{x}_l + \overline{J_{il}}x_l\overline{x}_i)}\, dx_l\right) \\ &= \frac{i}{\pi\Delta_i^{(j)}} \exp\left\{-i|x_i|^2\left(-z - \sum_{l\in\partial i\setminus j}\Delta_l^{(i)}|J_{il}|^2\right)\right\} .\end{aligned}$$

The normalisation condition $\int \widetilde{P}_i^{(j)}(x_i)\, dx_i = 1$ then gives

$$\Delta_i^{(j)} = \left(-z - \sum_{l\in\partial i\setminus j}\Delta_l^{(i)}|J_{il}|^2\right)^{-1} . \tag{2.13}$$

Similarly for Δ_i, one obtains

$$\Delta_i = \left(-z - \sum_{j\in\partial i}\Delta_j^{(i)}|J_{ij}|^2\right)^{-1} . \tag{2.14}$$

Setting $z = \lambda + i\varepsilon$ for some ε, if a solution set $\{\Delta_i^{(j)}\}$ to the system (2.13) can be found, then we obtain from (2.14) and the Stieltjes inversion formula (1.6), an approximation

to the regularised spectral density of A,

$$\widetilde{\varrho}_{\varepsilon}(\lambda; A) = \frac{1}{\pi N} \operatorname{Im} \sum_{i=1}^{N} \Delta_i \,. \tag{2.15}$$

In this way, the problem of computing an approximation to the spectral density to is reduced to that of solving the system (2.13). As we will see, a unique solution to (2.13) can always be found, by hand in some cases and by computer in general.

Recursion equations of the form (2.13) are not themselves new to the random matrix literature. The cavity method, or other essentially identical techniques, have been applied by several authors to the spectral density of the Anderson model on various graphs [ACTA73, CGMM$^+$05, BAT04, AF06]. For sparse matrices associated to tree-like graphs, Dorogovtsev *et al* [DGMS03, DGMS04] have also derived an equivalent set of recursion equations to those given above, this time through consideration of a generating function for walks in the graph.

In fact, it is possible to derive the above cavity equations by elementary algebra (see, for example, [FHS07] for the Anderson model and [BL10] for general sparse graphs), without using a statistical mechanics analogy, walk generating function or any other such device. We will use this method to derive a generalised version of the equations in the next chapter.

2.4 Analytically solvable cases

For small trees, the system of cavity equations given by (2.13) can be solved by hand, allowing the computation of an exact formula for the spectral density. For larger, more complex graphs, solution by hand becomes too difficult and one might choose to procede either with the aid of a computer (which will be covered in the next section) or by seeking some way to reduce the number of equations to a more manageable level. Such simplification is most easily achieved if the random graph ensemble from which

the matrix A is drawn exhibits a type of translational invariance, either statistical or exact. We explore a pair of simple examples.

2.4.1 The fully connected limit - Wigner's Law

The cavity equations derived above are in principle valid only for trees or tree-like graphs in the large N limit. In the study of spin-glasses, a technique due to Thouless, Anderson and Palmer (known as the TAP approach) predates the cavity method and can be thought of as its natural analogue for fully connected systems. There is a standard method by which the results of the TAP approach can be derived from the cavity equations, which we describe now for the case of Wigner random matrices[1].

Let A_N be a Hermitian matrix with independent entries of zero mean and variance $1/N$. In this model, although the underlying graph is fully connected, the cavity equations are still valid in the large N limit since contributions to the local field at a vertex i from pairs of neighbours are vanishing. We also have approximately negligible difference between the true and cavity parameters, $\Delta_i^{(j)} = \Delta_i + \mathcal{O}(1/N)$. Defining the average

$$\Delta = \frac{1}{N}\sum_{i=1}^{N} \Delta_i \,,$$

we have, as $N \to \infty$,

$$\sum_{l \neq i} |A_{il}|^2 \Delta_l^{(i)} \to \Delta \,.$$

The system of equations defined in (2.13) and (2.14) then reduce to the single equation

$$\Delta = (-z - \Delta)^{-1} \,, \tag{2.16}$$

which rearranges to give exactly the equation (1.9) for the Green's function of Wigner matrices, derived in Chapter 1 using the simple approach. In this way we see that

[1] In fact, the derivation here is somewhat simpler than in typical disordered systems, as the difference between true and cavity fields turns out to be small enough that no Onsager reaction term occurs.

Wigner's law can be recovered from the cavity equations in the fully connected limit, also revealing a connection between the cavity method and simple approach.

2.4.2 Random regular graphs - McKay's Law

Let us consider the adjacency matrix of a k-regular graph on N vertices. The cavity equations for such a graph can be solved by hand if we assume translational invariance of the cavity fields, that is, we put $\Delta_i^{(j)} \equiv \Delta$ for all $i \in V$ and $j \in \partial i$. We now have a single equation reading

$$\Delta = \frac{1}{-z-(k-1)\Delta}\,,$$

which is easily solved to give

$$\Delta = \frac{-z \pm \sqrt{z^2-4(k-1)}}{2(k-1)}\,,$$

where the branch of the square root is determined by the condition $\Delta \in \mathbb{C}^+$.

We will show later that the cavity equations for a finite graph always have a unique solution. Here, this fact allows us to conclude that all k-regular graphs share the above solution, regardless of whether or not they are transitive.

Taking the limit $\varepsilon \to 0$ we have $\widetilde{\varrho}_\varepsilon(\lambda; A) \to \rho_k(\lambda)$, where

$$\rho_k(\lambda) = \frac{k\sqrt{4(k-1)-\lambda^2}}{2\pi(k^2-\lambda^2)}\,.$$

This distribution was found by McKay in [McK81], in which he proved that for adjacency matrices A of graphs chosen uniformly at random from the set of k-regular graphs on N vertices one has the convergence of $\varrho(\lambda; A)$ to $\rho_k(\lambda)$ as $N \to \infty$.

2.5 Numerical solution

2.5.1 Belief propagation

To solve a large system of cavity equations such as (2.13) numerically, it is common to employ a simple iterative approach, known in this context as belief propagation. In such a scheme the number $\Delta_i^{(j)}$ is interpreted as a message sent from vertex i to vertex j about the contribution made by i to the local field at j.

One implementation of a belief propagation algorithm is as follows: starting with an arbitrary list of initial guesses $\left\{\Delta_i^{(j)}[1]\right\}_{i\in V,\, j\in\partial i}$, one repeatedly applies the update equation[2]

$$\Delta_i^{(j)}[n] = \left(-z - \sum_{l\in\partial i\setminus j} \Delta_l^{(i)}[n-1]|J_{il}|^2\right)^{-1} \tag{2.17}$$

until a fixed point $\left\{\Delta_i^{(j)}\right\}_{i\in V,\, j\in\partial i}$ is reached such that

$$\Delta_i^{(j)} = \Delta_i^{(j)}[n] = \Delta_i^{(j)}[n-1]\,.$$

For a finite tree, it is easy to see that this happens in a number of steps equal to the diameter of the tree. The spectral density can then be recovered by the equations (2.14) and (2.15).

The same process can be applied to graphs which are not trees, though now it is not guaranteed to find a solution in finite time, making it necessary to halt the program at some point, for example after a pre-determined level of convergence has been reached. This approach has been used frequently in the statistical mechanics literature to perform numerical simulations based on the cavity method.

[2] Alternatively, one could choose to update the messages one at a time rather than in parallel, which even easier to program.

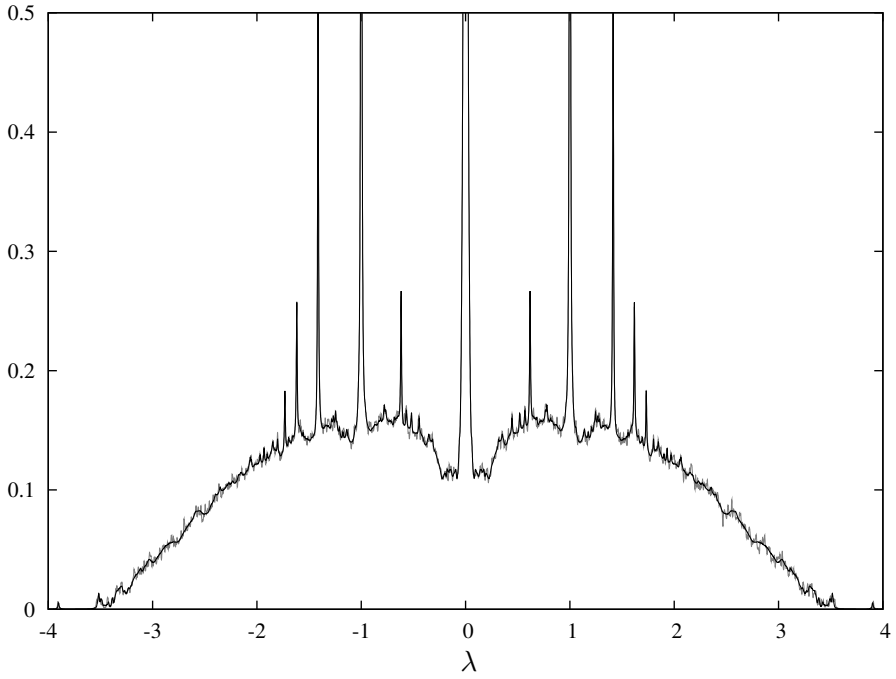


Figure 2.1: Grey - the regularised empirical spectral $\varrho(\lambda; A)$ density of a single Poissonian random graph of size $N = 10^4$, with average degree $c = 2$ at $\varepsilon = 0.005$. Black - the result $\widetilde{\varrho}_\varepsilon(\lambda; A)$ of the cavity equations solved by belief propagation on the same graph.

Figure 2.1 shows the results of a belief propagation algorithm using the update rule (2.17) applied to a single Poissonian random graph of size $N = 10^4$, with average degree $c = 2$, and regularising parameter $\varepsilon = 0.005$. The empirical regularised spectral density was obtained for the same graph by numerically determining the eigenvalues $\lambda_1^{(A)}, \ldots, \lambda_N^{(A)}$ and applying formula (1.7), that is,

$$\varrho_\varepsilon(\lambda; A) = \frac{1}{\pi N} \sum_{i=1}^{N} \frac{\varepsilon}{\varepsilon^2 + |\lambda - \lambda^{(A)}|^2} \,.$$

A close up of some detail from Figure 2.1 is shown in Figure 2.2. The results of the algorithm are very strong on the scale of the whole spectrum, though there are significant deviations on smaller scales.

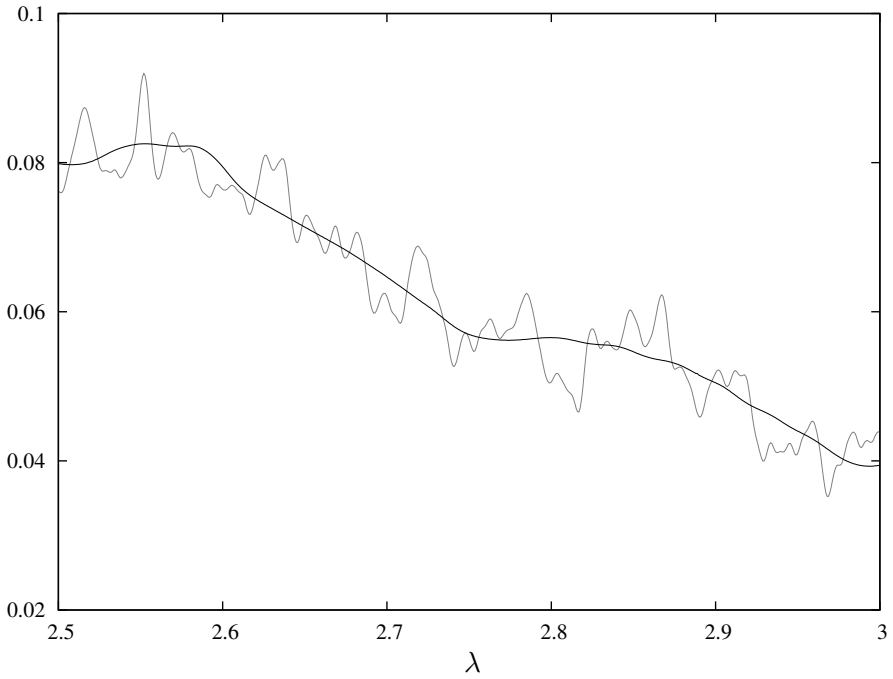


Figure 2.2: Close up detail of Figure 2.1. Discrepancies between the empirical $\varrho(\lambda; A)$ regularised spectral density (grey) and result $\widetilde{\varrho}_\varepsilon(\lambda; A)$ of the cavity equations (black) is clear on this scale.

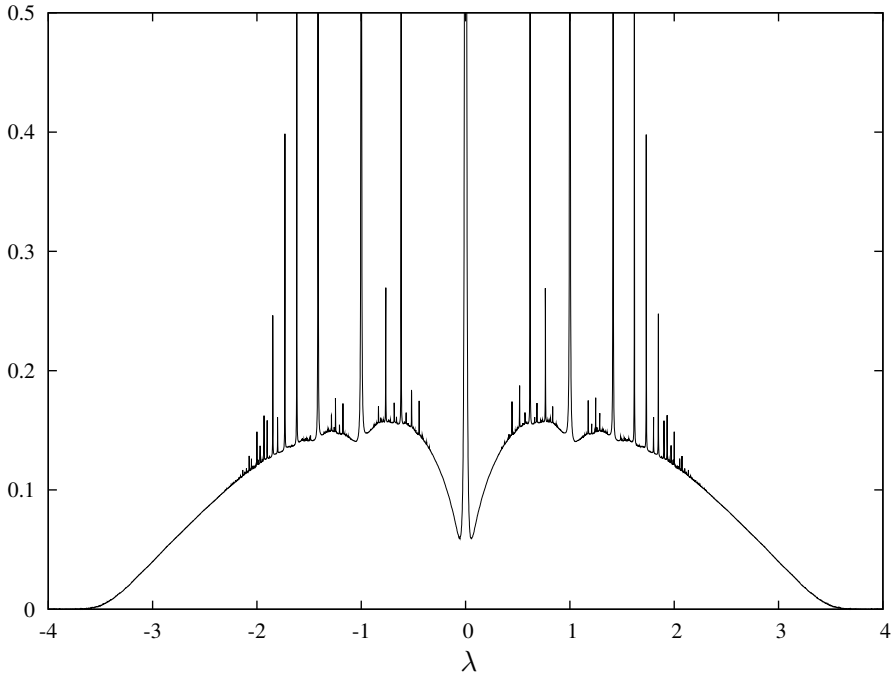


Figure 2.3: The result $\varrho_\varepsilon(\lambda; A)$ of the cavity equations solved by belief propagation on a graph Poissonian random graph of size $N = 10^7$, with average degree $c = 2$ at $\varepsilon = 0.005$.

The matrix size of $N = 10^4$ used in the production of figures 2.1 and 2.2 is close to the current practical size limit for numerical diagonalisation. By comparison, an efficient implementation of belief propagation can handle matrices many orders of magnitude larger using the same hardware. Figure 2.3 shows the results of the belief propagation algorithm on a graph of size $N = 10^7$.

2.5.2 Convergence to a unique solution

For belief propagation algorithms applied to general inference problems, there is no guarantee of the existence or uniqueness of solutions to the message passing equations, or of the hoped convergence of the algorithm. Some systems, such as certain

spin-glasses near a critical temperature, exhibit many solutions. In the present case however, convergence to a unique solution is a proveable fact.

For a given graph G with edge weights J and fixed $z = \lambda + i\varepsilon$, the update rule (2.17) may be written in terms of a closed map F defined on $(\mathbb{C}^+)^n$ (where $n = 2|E|$). Using boldface fonts for vectors in $(\mathbb{C}^+)^n$, we have

$$\boldsymbol{\Delta}[n] = F\big(\boldsymbol{\Delta}[n-1]\big)\,,$$

where the function F is defined for $\boldsymbol{\Gamma} \in (\mathbb{C}^+)^n$ by

$$F(\boldsymbol{\Gamma})_i^{(j)} = \left(-\lambda - i\varepsilon - \sum_{l \in \partial i \setminus j} \Gamma_l^{(i)} |J_{il}|^2\right)^{-1}. \tag{2.18}$$

To guarantee the convergence of the belief propagation algorithm to a unique solution, it is necessary and sufficient to prove the following:

Theorem 1. *The function F defined in (2.18) has a unique and globally asymptotically stable fixed point.*

Proof. First we note that by Brouwer's fixed point theorem, F certainly has at least one fixed point in $(\mathbb{C}^+ \cup \{\infty\})^n$.

Next we observe that the image of $F^2 = F \circ F$ is bounded away from ∞ and $\mathbb{R}$; for each i and j

$$\sup_{\boldsymbol{\Gamma} \in \mathbb{C}^{+n}} \left|F(\boldsymbol{\Gamma})_i^{(j)}\right| = \frac{1}{\varepsilon}\,,$$

and thus

$$\sup_{\boldsymbol{\Gamma} \in \mathbb{C}^{+n}} \left|F^2(\boldsymbol{\Gamma})_i^{(j)}\right| = \frac{1}{\varepsilon}\,, \quad \text{and} \quad \inf_{\boldsymbol{\Gamma} \in \mathbb{C}^{+n}} \left(\operatorname{Im}\left[F^2(\boldsymbol{\Gamma})_i^{(j)}\right]\right) > \kappa\,, \tag{2.19}$$

for some $\kappa > 0$. So any fixed point of F must be contained in the region Ω^n, where $\Omega = \big\{\Gamma \in \mathbb{C}^+ : \operatorname{Im}\Gamma > \kappa,\, |\Gamma| < \varepsilon^{-1}\big\}$. Take the fixed point implied by Brouwer and name it $\boldsymbol{\Delta}$. To complete the proof, we will show that for all $\boldsymbol{\Gamma}$ we have $F^{2m}(\boldsymbol{\Gamma}) \to \boldsymbol{\Delta}$ as $m \to \infty$.

This is most easily achieved by mapping to the polydisc D^n, where D is the unit disc in the complex plane. Introduce the following conformal maps:

$$c : \mathbb{C}^+ \to D\,, \quad c(z) = \frac{z-i}{z+i}\,,$$

$$t_i^{(j)} : D \to D\,, \quad t_i^{(j)}(z) = \frac{z - c(\Delta_i^{(j)})}{1 - \overline{c(\Delta_i^{(j)})}z}\,.$$

The composition $t_i^{(j)} \circ c$ maps the upper half-plane to the unit disc in such a way that the point $\Delta_i^{(j)}$ is mapped to 0. Define maps $C : \mathbb{C}^{+n} \to D^n$ and $T : D^n \to D^n$ by

$$C(\boldsymbol{z})_i^{(j)} = c(z_i^{(j)})\,, \quad \text{and} \quad T(\boldsymbol{z})_i^{(j)} = t(z_i^{(j)})\,.$$

We can now consider the composition $G = T \circ C \circ F^2 \circ C^{-1} \circ T^{-1}$, which maps D^n to itself and fixes 0. Moreover, thanks to the bounds (2.19) established on F^2, there must exist a number $\mu \in (0,1)$, such that

$$\sup_{\boldsymbol{z} \in D^n} \left(\max_{i,j} \left| G(\boldsymbol{z})_i^{(j)} \right| \right) < \mu\,.$$

The function G also satisfies a generalisation of the Cauchy-Riemann equations, in that each component is a holomorphic function of each other component, allowing us to apply a result from the theory of complex functions of several variables. Specifically, a generalisation of the Schwarz lemma [GK03] gives that for any $\boldsymbol{z} \in D^n$

$$\max_{i,j} \left| G(\boldsymbol{z})_i^{(j)} \right| \leq \mu \max_{i,j} \left| z_i^{(j)} \right|\,.$$

Since $\mu < 1$, we obtain $G^m(\boldsymbol{z}) \to 0$ as $m \to \infty$ for all $\boldsymbol{z} \in D^n$, and hence $F^{2m}(\boldsymbol{\Gamma}) \to \boldsymbol{\Delta}$ for all $\boldsymbol{\Gamma} \in \mathbb{C}^{+n}$. □

The above result has implications beyond confirming the algorithmic performance of belief propagation in this case. In particular, we may now conclude the correctness of the cavity equations for tree-like graphs in the limit $N \to \infty$. Suppose that the ball $B_n(i)$ of radius n around a vertex i is a tree and consider the following set of initial conditions for the belief propagation algorithm: for all $l \in V$ and all $j \in \partial l$,

$$\Delta_l^{(j)}[1] = (A^{(j)} - z)_{ll}^{-1}\,.$$

After n steps, inspection of (2.17) shows that (since $B_n(i)$ is a tree) $\Delta_i[n] = (A - z)^{-1}_{ii}$. But the convergence of the belief propagation algorithm implies that $|\Delta_i[n] - \Delta_i| \to 0$ as $n \to \infty$. For tree-like graph ensembles, balls of arbitrary radius $1 \ll n \ll N$ converge to trees with probability 1 in the limit $N \to \infty$, and we can thus conclude that the results of the cavity method are exact in this limit.

2.5.3 An exact solution to a different problem

It is a well-known fact in numerical analysis that the computation of the eigenvalue spectrum of a non-normal matrix can be highly sensitive to rounding errors, with the outputted numbers sometimes very far from the true eigenvalues. What is more, it can be shown that these 'approximate' eigenvalues are in fact the true eigenvalues of a small perturbation of the original matrix.

In light of this, it becomes natural to ask if, for graphs with cycles, the 'approximate' values for the diagonal entries of the resolvent determined by the cavity method are the correct values for some closely related system. To help answer this question, we introduce the notion of a walk in a graph. For simplicity, we assume below that G is a finite connected graph which is not a tree, and we are interested only in the spectral density of the adjacency matrix A. The arguments are easily generalised to other cases.

A walk $\boldsymbol{w} = (w_0, \ldots, w_n)$ of length $n \geq 0$ in a graph $G = (V, E)$ is an ordered collection of vertices $\{w_i\}_{i=0}^n \subset V$ such that for each $i = 1, \ldots, n$, there is an edge in G between vertices w_{i-1} and w_i. For walks $\boldsymbol{w} = (w_0, \ldots, w_n)$ of length $n > 0$, we define $\boldsymbol{w}^{(-)} = (w_0, \ldots, w_{n-1})$. A non-backtracking walk is a walk in which for each $i = 1, \ldots, n-1$ we have $w_{i-1} \neq w_{i+1}$. Write $\mathcal{W}$ for the collection of all non-backtracking walks in G.

Consider the graph $\mathcal{G} = (\mathcal{W}, \mathcal{E})$, with edge set defined by the constraint

$$(\boldsymbol{w}, \boldsymbol{v}) \in \mathcal{E} \quad \Leftrightarrow \quad \boldsymbol{v} = \boldsymbol{w}^{(-)} \quad \text{or} \quad \boldsymbol{v}^{(-)} = \boldsymbol{w}\,.$$

The structure of $\mathcal{G}$ may be summarised by the following facts:

1. $\mathcal{G}$ has N connected components, one for each vertex in $i \in V$; we refer to the zero-length walk (i) as the root of the corresponding component
2. The components of $\mathcal{G}$ are infinite trees, isomorphic to each other
3. For each $\boldsymbol{w} = (w_0, \ldots, w_n)$, there is an automorphism of $\mathcal{G}$ which maps $\boldsymbol{w} \mapsto w_n$
4. (i, j) is a neighbour of (i) in $\mathcal{G}$ if and only if j is a neighbour of i in G

A simple example of the relationship between graph G and tree $\mathcal{G}$ of non-backtracking walks is shown in Figure 2.4.

Write $\mathcal{A}$ for the (infinite) adjacency matrix of $\mathcal{G}$, whose action on $\boldsymbol{x} \in l^2(\mathcal{W})$ is defined by

$$(\mathcal{A}\boldsymbol{x})_{\boldsymbol{w}} = \sum_{\boldsymbol{v} \in \partial \boldsymbol{w}} \boldsymbol{x}_{\boldsymbol{v}}\,.$$

Let $\mu(\lambda; \mathcal{A})$ be the resolution of identity for $\mathcal{A}$ and $\mathcal{R}(z; \mathcal{A})$ the resolvent operator. The entries of μ are known as the spectral measures of the operator $\mathcal{A}$, and can be calculated from the resolvent. The two are related by

$$\mathcal{R}(z; \mathcal{A}) = \int \frac{1}{\lambda - z} \mu(\lambda; \mathcal{A})\, d\lambda\,,$$

and the inverse operation

$$\mu(\lambda; \mathcal{A}) = \lim_{\varepsilon \to 0} \mu_\varepsilon(\lambda; \mathcal{A})\,,$$

where

$$\Big[\mu_\varepsilon(\lambda; \mathcal{A})\Big]_{\boldsymbol{v}\boldsymbol{w}} = \frac{1}{\pi} \operatorname{Im} \Big[\mathcal{R}(\lambda + i\varepsilon; \mathcal{A})\Big]_{\boldsymbol{v}\boldsymbol{w}}\,.$$

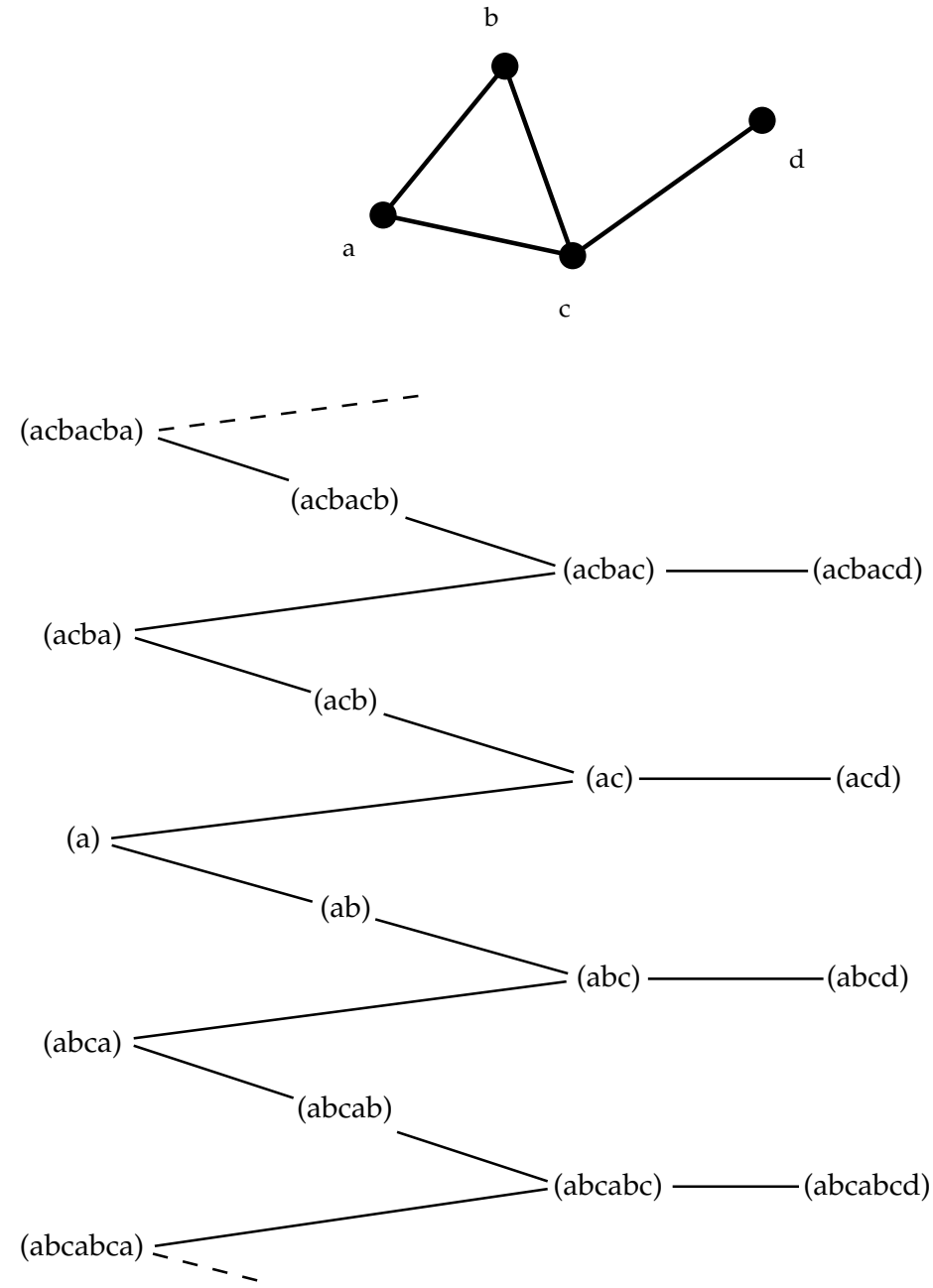


Figure 2.4: Top - A small graph G containing a cycle. Bottom - part of the (a)-rooted component of the infinite tree $\mathcal{G}$ of non-backtracking walks on G. Each vertex in G gives rise to the equivalence class of all vertices in $\mathcal{G}$ with that vertex as the endpoint of the walk, shown here by the four columns.

Following arguments similar to those in [GM88, DGMS03, MJW06], it is easy to establish recursion relations for the entries of the resolvent of $\mathcal{G}$:

$$\Big[\mathcal{R}(\lambda;\mathcal{A})\Big]_{\boldsymbol{ww}} = \left(-z - \sum_{\boldsymbol{u}\in\partial\boldsymbol{w}} \Big[\mathcal{R}(\lambda;\mathcal{A})\Big]^{(\boldsymbol{w})}_{\boldsymbol{uu}}\right)^{-1}, \tag{2.20}$$

and

$$\Big[\mathcal{R}(\lambda;\mathcal{A})\Big]^{(\boldsymbol{v})}_{\boldsymbol{ww}} = \left(-z - \sum_{\boldsymbol{u}\in\partial\boldsymbol{w}\setminus\boldsymbol{v}} \Big[\mathcal{R}(\lambda;\mathcal{A})\Big]^{(\boldsymbol{w})}_{\boldsymbol{uu}}\right)^{-1}. \tag{2.21}$$

Thanks to property 3 above, we have for all walks $\boldsymbol{v}$, $\boldsymbol{w}$, with end vertices i and j, respectively,

$$\Big[\mathcal{R}(\lambda;\mathcal{A})\Big]^{(\boldsymbol{w})}_{\boldsymbol{vv}} = \Big[\mathcal{R}(\lambda;\mathcal{A})\Big]^{((j))}_{(i)(i)}, \quad \text{and} \quad \Big[\mathcal{R}(\lambda;\mathcal{A})\Big]_{\boldsymbol{vv}} = \Big[\mathcal{R}(\lambda;\mathcal{A})\Big]_{(i)(i)}.$$

From this and property 4, we can deduce that the recursion relations (2.20) and (2.21) defined on $\mathcal{G}$ are precisely the same as the cavity equations (2.13) and (2.14) for G. Thus, if $i \in V$ is the end vertex of $\boldsymbol{w} \in \mathcal{W}$, then

$$\Big[\mathcal{R}(\lambda;\mathcal{A})\Big]_{\boldsymbol{ww}} = \Delta_i\,,$$

and we establish the identity

$$\widetilde{\varrho}_\varepsilon(\lambda;A) = \frac{1}{N}\sum_{i=1}^{N} \Big[\mu_\varepsilon(\lambda;\mathcal{A})\Big]_{(i)(i)}.$$

This result tells us that by solving the cavity equations in graphs containing cycles, we are in fact computing the spectral measures at the roots of the components of $\mathcal{G}$. The approximation to the spectral density obtained from the cavity method is nothing but the arithmetic mean of the spectral measures. This fact accounts for the discrepancy seen in Figure 2.2: the true spectral density of the finite matrix A is composed entirely of delta peaks, whereas the spectral measures of the infinite operator $\mathcal{A}$ may have continuous parts.

2.6 Ensemble average for Poissonian graphs

2.6.1 A distributional recursion equation

Although the cavity equations are derived for a fixed finite graph, it is possible to perform the ensemble average in the limit $N \to \infty$. The calculation requires a certain assumption on the solutions of the cavity equations, as well as the computation of some statistical properties of the graphs. For now, we consider only Poissonian random graphs with all $J_{ij} = 1$ (in the next chapter we will introduce more complicated ensembles). Fix $z \in \mathbb{C}^+$ and introduce the following distribution

$$\pi(\Delta) = \mathbb{E}\left(\delta\big(\Delta - \Delta_i^{(j)}\big) \,\middle|\, C_{ij} = 1\right),$$

where $\{\Delta_i^{(j)}\}$ is the unique solution of the cavity equations for the graph C, and the average is over the random graph ensemble. The cavity equations can now be used to relate average value of $\Delta_i^{(j)}$ the average on neighbouring vertices. To achieve this, we sum over all possible values of the degree of i and all possible choices of that many neighbours:

$$\begin{aligned}
\pi(\Delta) &= \mathbb{E}\left\{\frac{1}{Nc}\sum_{i,j=1}^{N} C_{ij}\delta\big(\Delta - \Delta_i^{(j)}\big)\right\} \\
&= \mathbb{E}\left\{\sum_k \frac{k}{Nc}\sum_{i=1}^{N}\delta_{k,k_i(C)}\sum_{j<l_1<\cdots<l_{k-1}} C_{ij}C_{il_1}\cdots C_{il_{k-1}}\,\delta\left(\Delta + \frac{1}{z+\sum_{\alpha=1}^{k-1}\Delta_{l_\alpha}^{(i)}}\right)\right\} \\
&= \sum_k \frac{k}{c}\int\Big[\prod_{\alpha=1}^{k-1} d\Delta_\alpha\Big]\,\Xi(\Delta_1,\ldots,\Delta_{k-1})\,\delta\left(\Delta + \frac{1}{z+\sum_{\alpha=1}^{k-1}\Delta_\alpha}\right)
\end{aligned} \tag{2.22}$$

where

$$\Xi(\Delta_1,\ldots,\Delta_{k-1}) = \sum_{j<l_1<\cdots<l_{k-1}} \mathbb{E}\left\{C_{ij}C_{il_1}\cdots C_{il_{k-1}}\delta_{k,k_i(C)}\prod_{\alpha=1}^{k-1}\delta\big(\Delta_\alpha - \Delta_{l_\alpha}^{(i)}\big)\right\}.$$

To enable us to evaluate Ξ, we make an assumption of independence for the $\{\Delta_{l_\alpha}^{(i)}\}$. The Poissonian random graph ensemble is known to be tree-like in the large N limit,

meaning that for sufficiently large N, in the absence of a vertex i the distance between any two of it's neighbours $l_1, l_2 \in \partial i$ is typically either infinite or of order very much larger than the number of steps required for the convergence of the belief propagation algorithm. The result of this is that we may take the $\{\Delta_{l_\alpha}^{(i)}\}$ to depend only upon their local environment and not on each other. Moreover, the Poissonian random graph ensemble lacks any correlation between the degrees of neighbouring vertices, making the $\{\Delta_{l_\alpha}^{(i)}\}$ independent of $k_i(C)$, and thus each other, allowing us to write as $N \to \infty$

$$\begin{aligned} &\mathbb{E}\left\{ C_{ij}C_{il_1}\cdots C_{il_{k-1}}\delta_{k,k_i(C)} \prod_{\alpha=1}^{k-1} \delta\big(\Delta_\alpha - \Delta_{l_\alpha}^{(i)}\big) \right\} \\ &\simeq \mathbb{E}\Big\{ C_{ij}C_{il_1}\cdots C_{il_{k-1}}\delta_{k,k_i(C)} \Big\} \prod_{\alpha=1}^{k-1} \mathbb{E}\left(\delta\big(\Delta_\alpha - \Delta_{l_\alpha}^{(i)}\big) \,\Big|\, C_{il_\alpha} = 1 \right) . \end{aligned} \tag{2.23}$$

To complete the computation of the ensemble average, we introduce a generating function with fields $\boldsymbol{h} = \{h_{ij}\}$, associated to the edges of the graph, let

$$Z(\boldsymbol{h}) = \mathbb{E}\left(\delta_{k,k_i(C)} \prod_{i<j} e^{h_{ij}C_{ij}} \right) ,$$

so that

$$\mathbb{E}\Big\{ C_{ij_1}\cdots C_{ij_k}\delta_{k,k_i(C)} \Big\} = \frac{\partial}{\partial h_{ij_1}} \cdots \frac{\partial}{\partial h_{ij_k}} Z(\boldsymbol{h}) \bigg|_{\boldsymbol{h}=0} . \tag{2.24}$$

For Poissonian random graphs, we have

$$P(C) = \prod_{i<j} \Big[\frac{c}{N}\delta_{C_{ij},1} + \Big(1 - \frac{c}{N}\Big)\delta_{C_{ij},0} \Big] ,$$

and thus, taking a Fourier representation for the degree constraint,

$$\begin{aligned} Z(\boldsymbol{h}) &= \sum_C P(C) \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\left\{ iw\left(k - \sum_{l=1}^{N} C_{il} \right) + \sum_{l<j} h_{jl}C_{jl} \right\} dw \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{iwk} \prod_{l=1}^{N} \Big[1 - \frac{c}{N} + \frac{c}{N} e^{h_{il} - iw} \Big] \prod_{j<l\neq i} \Big[1 - \frac{c}{N} + \frac{c}{N} e^{h_{jl}} \Big] \, dw \end{aligned}$$

Applying (2.24), we obtain

$$\begin{aligned} &\sum_{j_1<\ldots<j_k} \mathbb{E}\Big\{ C_{ij_1}\cdots C_{ij_k}\delta_{k,k_i(C)} \Big\} \\ &= \binom{N}{k} \Big(\frac{c}{N}\Big)^k \frac{1}{2\pi} \int_{-\pi}^{\pi} \Big(1 - \frac{c}{N} + \frac{c}{N} e^{-iw} \Big)^{N-k} \, dw \longrightarrow \frac{c^k\, e^{-c}}{k!} , \end{aligned}$$

which is simply the limiting Poissonian degree distribution $p(k)$. Returning this and (2.23) to the calculation started in (2.22) we reach

$$\pi(\Delta) = \sum_k \frac{k\,p(k)}{c} \int \Big[\prod_{\alpha=1}^{k-1} d\Delta_\alpha \pi(\Delta_\alpha) \Big] \delta\left(\Delta + \frac{1}{z + \sum_{\alpha=1}^{k-1} \Delta_\alpha} \right) . \tag{2.25}$$

In this way, we have obtained from the cavity equations a self-consistency relation for the distribution of solutions $\Delta_i^{(j)}$, which is valid in the large N limit for Poissonian random graphs. To compute the ensemble average spectral density, one may introduce another density

$$\widehat{\pi}(\Delta) = \mathbb{E}\, \delta\big(\Delta - \Delta_i\big) ,$$

which specifies the ensemble average spectral density by

$$\rho(\lambda) = \lim_{\varepsilon \searrow 0} \frac{1}{\pi} \operatorname{Im} \int \widehat{\pi}(\Delta) \Delta \, d\Delta .$$

A very similar calculation to the one given above determines $\widehat{\pi}$ in terms of π by

$$\widehat{\pi}(\Delta) = \sum_k p(k) \int \Big[\prod_{\alpha=1}^{k} d\Delta_\alpha \pi(\Delta_\alpha) \Big] \delta\left(\Delta + \frac{1}{z + \sum_{\alpha=1}^{k} \Delta_\alpha} \right) . \tag{2.26}$$

To reconnect with other methods, we note that equations equivalent to (2.25) were derived using replicas in [Dea02] and [Küh08]. The same result was also obtained by techniques more closely related to the cavity method in [BL10], in which the existence of a unique density π satisfying (2.25) was also proved. The effective medium approximation as presented by Dorogovtsev *et al* [DGMS03, DGMS04] is recovered from (2.25) by making the simplifying assumption $\pi(\Delta) = \delta(\Delta - \Delta_{\mathrm{EMA}})$.

2.6.2 Numerical solution by population dynamics

For the Poissonian random graph ensemble at least, and probably many others in general, the solution to (2.25) is unlikely to have a simple form. For example, it was proved by Golinelli [Gol03] that it must contain contributions from a collection of delta peaks dense in $\mathbb{R}$. We seek to make further progress via numerical methods.

Just as the cavity equations for a single instance can be solved numerically by iteration, so too can the distributional recursion equation (2.25). The idea behind the method is to approximate the distribution π by a large empirical sample, which is repeatedly updated at random.

Suppose k is a random variable drawn from the distribution with law $kp(k)/c$ and $\Delta_1, \ldots, \Delta_{k-1}$ are independent random variables drawn from π, then equation (2.25) implies that the dependent random variable

$$\Delta = \left(-z - \sum_{\alpha=1}^{k-1} \Delta_\alpha \right)^{-1} \tag{2.27}$$

also has probability density π. Starting with a large sample $\{\Delta_1, \ldots, \Delta_M\}$, the population dynamics algorithm works by iteratively replacing each member of the sample with a new value drawn at random according to (2.27). Once an approximate sample from π is obtained, an approximate sample from $\widehat{\pi}$ may be generated by the same method. A simple implementation may run as follows:

Algorithm 3

1. Generate an approximate sample from π:
   (a) Initialise a population of M numbers $\Delta_1, \ldots, \Delta_M \in \mathbb{C}^+$
   (b) Repeat for a predetermined number of steps:
   For each $m \in \{1, \ldots, M\}$
      i. Select a degree k at random according to $kp(k)/c$
      ii. Select uniformly at random $k-1$ members of the population and label them $\Delta_1, \ldots, \Delta_{k-1}$
      iii. Replace
      $$\Delta_m = \left(-\lambda - i\varepsilon - \sum_{\alpha=1}^{k-1} \Delta_\alpha \right)^{-1}$$

2. Generate an approximate sample from $\widehat{\pi}$:

   (a) Initialise a population of M numbers $\widehat{\Delta}_1, \ldots, \widehat{\Delta}_M \in \mathbb{C}^+$

   (b) For each $m \in \{1, \ldots, M\}$

      i. Select a degree k at random according to $p(k)$

      ii. Select uniformly at random k members of the sample generated in step 1 and label them $\Delta_1, \ldots, \Delta_k$

      iii. Put
$$\widehat{\Delta}_m = \left(-\lambda - i\varepsilon - \sum_{\alpha=1}^{k} \Delta_\alpha \right)^{-1}$$

3. Output an approximation to the spectral density at λ:
$$\rho_\varepsilon(\lambda) = \frac{1}{M\pi} \operatorname{Im} \sum_{m=1}^{M} \widehat{\Delta}_m \, .$$

It is also possible to extract an estimate of the unregularised spectral density (provided the point λ does not coincide with the location of a delta peak) by, for example, sending $\varepsilon \to 0$ during the course of the algorithm.

Figure 2.5 shows both the regularised and unregularised results of population dynamics for the Poissonian random graph ensemble of average degree $c = 2$. A population of size $M = 5 \times 10^4$ was used and the output was averaged over 2000 iterations. At $\varepsilon = 0$ only the non-singular part of the density is visible, whilst the inclusion of a finite regulariser makes visible some of the dense collection of delta peaks.

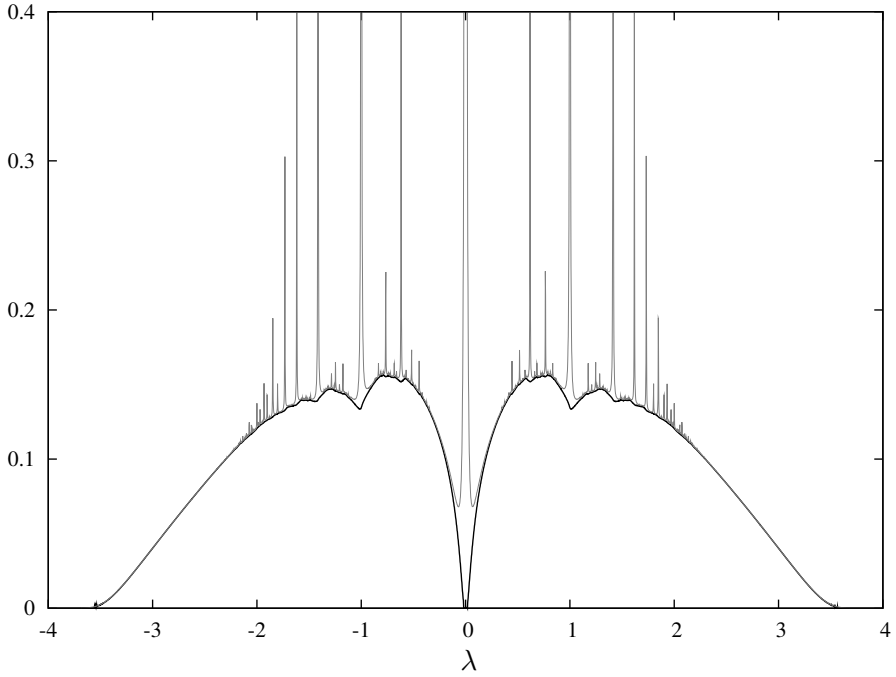


Figure 2.5: Results of population dynamics at $\varepsilon = 0$ (black) and $\varepsilon = 0.0016$ (grey) for the Poissonian random graph ensemble of average degree $c = 2$.

3

Beyond Poissonian graphs

3.1 Graphs with generalised degree correlations

3.1.1 Ensemble average spectral density

In the investigation of spectral density, the most commonly studied random graph ensembles are those in which the degrees of neighbouring vertices become independent in the limit $N \to \infty$. This class includes the Poissonian random graphs considered in the previous chapter and references [RB88, BG01, BM99, NT07, NR08, SC02, Dea02], and graphs with a specified degree distribution [Küh08, BL10]. Unfortunately, these simple ensembles may not provide at all realistic models for real-world complex networks, which can feature detailed correlations between neighbouring and non-neighbouring vertices. A few results have been obtained for the spectral density of random graph ensembles with more complex topologies, such as an approximation scheme for graphs with degree correlations [DGMS03] and numerical investigations of some other ensembles [FDBV01]. However, beyond these examples there is little further research.

In the previous chapter, we saw how the cavity equations derived for single instances

of tree-like graphs may be used as a starting point for the calculation of a closed expression characterising the ensemble averaged spectral density of Poissonian random graphs. In the present subsection, we will show that the same cavity equations may be combined with a more detailed analysis of local graph statistics to determine the mean spectral density of random graph ensembles featuring much more complex degree statistics. A concrete example of such an ensemble is introduced and analysed subsequently, and the theoretical results are compared with numerical simulations later in the section.

To capture the effect of increasingly complex graph topologies, we base our analysis on a generalised notion of degree, constructed by enumeration of walks. Define the l-th generalised degree of a vertex i in a graph with adjacency matrix C, denoted as $k_i^{(l)}(C)$, to be the total number of walks of length l starting at vertex i. In terms of the connectivity matrix, this is given by the recursive definition

$$k_i^{(0)}(C) = 1\,, \quad k_i^{(l)}(C) = \sum_{j=1}^{N} C_{ij} k_j^{(l-1)}(C) \quad \text{for } l \geq 1\,. \tag{3.1}$$

We consider generalised degrees up to some fixed radius L, introducing the L-vector $\boldsymbol{k}_i(C) = (k_i^{(1)}(C), \cdots, k_i^{(L)}(C))$. The first component of $\boldsymbol{k}_i(C)$ is simply the degree of vertex i, and so we will usually refer to it as $k_i(C)$, rather than $k_i^{(1)}(C)$. Importantly, it follows from the definition (3.1) that the generalised degree of a particular vertex contains information about the generalised degrees of its neighbours, thus providing a device with which to study the effect of correlations between vertex degrees.

The calculation of the ensemble average spectral density of Poissonian random graphs presented in the previous chapter was based on an analysis of the distribution

$$\pi(\Delta) = \mathbb{E}\left(\delta\big(\Delta - \Delta_i^{(j)}\big) \,\middle|\, C_{ij} = 1\right)\,, \tag{3.2}$$

where the $\{\Delta_i^{(j)}\}$ are the solutions to the cavity equations (2.13). That the final expression for the spectral density is written entirely in terms of this quantity implies that it carries enough information to characterise any effects of structure of the graph,

although the self-consistency equation derived (2.25) depends only upon the degree distribution $p(k)$.

To study graphs featuring correlations between generalised degrees, a quantity expressed in terms of the degree distribution alone will no longer be sufficient since additional structure is present. The relevant statistical properties may instead be captured in the joint distribution

$$\begin{aligned} &P\left(\{\boldsymbol{d}_t\}_{t=1}^k, \boldsymbol{k}\right) \\ &= \lim_{N\to\infty} \mathbb{E}\left(\frac{k}{cN}\sum_{i=1}^{N} \delta_{\boldsymbol{k}_i(C),\boldsymbol{k}} \sum_{j_1<\cdots<j_k} \delta_{\boldsymbol{k}_{j_1}(C),\boldsymbol{d}_1}\cdots\delta_{\boldsymbol{k}_{j_k}(C),\boldsymbol{d}_k} C_{ij_1}\cdots C_{ij_k}\right), \end{aligned} \tag{3.3}$$

describing the probability of finding a vertex whose neighbour has generalised degree $\boldsymbol{k}$ and is connected to k vertices with generalised degrees $\{\boldsymbol{d}_t\}_{t=1}^k$. Various marginals of (3.3) will play important roles in the forthcoming analysis, namely

1. Summing over all $\{\boldsymbol{d}_t\}_{t=1}^k$, one obtains the marginal

$$P(\boldsymbol{k}) = \lim_{N\to\infty} \mathbb{E}\left(\frac{k}{cN}\sum_{i=1}^{N} \delta_{\boldsymbol{k}_i(C),\boldsymbol{k}} C_{ij}\right) = \frac{kp(\boldsymbol{k})}{c},$$

   where $p(\boldsymbol{k}) = \lim_{N\to\infty}\mathbb{E}\,(1/N)\sum_{i=1}^{N}\delta_{\boldsymbol{k}_i(C),\boldsymbol{k}}C_{ij}$ is the generalised degree distribution

2. Summing over all but one of $\{\boldsymbol{d}_t\}_{t=1}^k$, we find the probability of an edge joining vertices of generalised degrees $\boldsymbol{k}$ and $\boldsymbol{d}$,

$$P(\boldsymbol{k},\boldsymbol{d}) = \lim_{N\to\infty} \mathbb{E}\left(\frac{1}{cN}\sum_{i=1}^{N} \delta_{\boldsymbol{k}_i(C),\boldsymbol{k}}\delta_{\boldsymbol{k}_j(C),\boldsymbol{d}} C_{ij}\right),$$

   and the conditional distribution $P(\boldsymbol{d}|\boldsymbol{k}) = P(\boldsymbol{d},\boldsymbol{k})/P(\boldsymbol{k})$

3. Important roles are also played by the conditional distributions

$$P\big(\{\boldsymbol{d}_t\}_{t=1}^k|\boldsymbol{k}\big) = P\left(\{\boldsymbol{d}_t\}_{t=1}^k, \boldsymbol{k}\right)/P(\boldsymbol{k}),$$

and

$$P\big(\{\boldsymbol{d}_t\}_{t=1}^{k-1}|\boldsymbol{k},\boldsymbol{d}\big) = P\left(\{\boldsymbol{d}_t\}_{t=1}^{k},\boldsymbol{k}\right)/P(\boldsymbol{d},\boldsymbol{k})\,.$$

Suppose we wish to compute the ensemble average spectral density of tree-like random graphs for which we know the distribution $P\left(\{\boldsymbol{d}_t\}_{t=1}^{k},\boldsymbol{k}\right)$. In a general case, we can no longer expect to find a closed expression for the distribution of all cavity variables (3.2) which will capture the full detail of the spectral density for such ensembles. Instead we work with a collection of distributions, describing the statistics of cavity variables associated to pairs of vertices with specified generalised degrees. For each $\boldsymbol{k}$ and $\boldsymbol{d}$ introduce

$$\begin{aligned}\pi(\Delta|\boldsymbol{k},\boldsymbol{d}) &= \mathbb{E}\left(\delta\big(\Delta-\Delta_i^{(j)}\big)\,\middle|\,C_{ij}=1,\,\boldsymbol{k}_i(C)=\boldsymbol{k},\,\boldsymbol{k}_j(C)=\boldsymbol{d}\right)\\ &= \frac{1}{P(\boldsymbol{k},\boldsymbol{d})}\mathbb{E}\left(\frac{1}{cN}\sum_{i,j=1}^{N}C_{ij}\delta_{\boldsymbol{k},\boldsymbol{k}_i(C)}\delta_{\boldsymbol{d},\boldsymbol{k}_j(C)}\delta\big(\Delta-\Delta_i^{(j)}\big)\right)\,.\end{aligned}$$

By summing over all possible configurations of vertices neighbouring a vertex of given generalised degree, we may use the cavity equations (2.13) to write

$$\begin{aligned}\pi(\Delta|\boldsymbol{k},\boldsymbol{d}) = \sum_{\boldsymbol{d}_1\ldots,\boldsymbol{d}_{k-1}}\int\left[\prod_{\alpha=1}^{k-1}d\Delta_\alpha\right]\Xi(\Delta_1,\ldots,\Delta_{k-1},\boldsymbol{d}_1,\ldots,\boldsymbol{d}_{k-1})\\ \times\,\delta\left(\Delta+\frac{1}{z+\sum_{\alpha=1}^{k-1}\Delta_\alpha}\right)\,,\end{aligned}\tag{3.4}$$

where

$$\begin{aligned}&\Xi(\Delta_1,\ldots,\Delta_{k-1},\boldsymbol{d}_1,\ldots,\boldsymbol{d}_{k-1})\\ &= \frac{1}{P(\boldsymbol{k},\boldsymbol{d})}\mathbb{E}\Bigg(\frac{k}{cN}\sum_{i=1}^{N}\delta_{\boldsymbol{k},\boldsymbol{k}_i(C)}\\ &\qquad\sum_{j<l_1<\ldots<l_{k-1}}C_{ij}\delta_{\boldsymbol{d},\boldsymbol{k}_j(C)}\prod_{\alpha=1}^{k-1}\Big[C_{il_\alpha}\delta_{\boldsymbol{d}_\alpha,\boldsymbol{k}_{l_\alpha}(C)}\delta\big(\Delta_\alpha-\Delta_{l_\alpha}^{(i)}\big)\Big]\Bigg)\,.\end{aligned}$$

As in the calculation of the ensemble average spectral density for Poissonian graphs, to evaluate the quantity Ξ we make an assumption on the behaviour of the cavity variables $\{\Delta_l^{(i)}\}_{l\in\partial i}$. In that calculation, it was assumed that all $\Delta_l^{(i)}$ would be independent. For more general ensembles, such a strong assumption is not generally

correct: $\Delta_l^{(i)}$ depends on the degree of vertex l, which can be strongly constrained by the generalised degree of i and its other neighbours. Provided, however, that we are dealing with a tree-like ensemble, we are able make the weaker assumption that the $\{\Delta_l^{(i)}\}_{l\in\partial i}$ are independent when conditioned upon the generalised degree $\boldsymbol{k}_i(C)$ of vertex i. We write, for large N,

$$\begin{aligned}
&\Xi(\Delta_1,\ldots,\Delta_{k-1},\boldsymbol{d}_1,\ldots,\boldsymbol{d}_{k-1})\\
&\quad\simeq \mathbb{E}\left(\frac{k}{cN}\sum_{i=1}^{N}\delta_{\boldsymbol{k},\boldsymbol{k}_i(C)}\sum_{j<l_1<\ldots<l_{k-1}}C_{ij}\delta_{\boldsymbol{d},\boldsymbol{k}_j(C)}\prod_{\alpha=1}^{k-1}\left[C_{il_\alpha}\delta_{\boldsymbol{d}_\alpha,\boldsymbol{k}_{l_\alpha}(C)}\right]\right)\\
&\qquad\times\prod_{\alpha=1}^{k-1}\mathbb{E}\left(\delta\big(\Delta_\alpha-\Delta_{l_\alpha}^{(i)}\big)\,\Big|\,C_{il_\alpha}=1,\,\boldsymbol{k}_i(C)=\boldsymbol{k},\,\boldsymbol{k}_{l_\alpha}(C)=\boldsymbol{d}\right)\\
&\quad\simeq P(\{\boldsymbol{d}_\alpha\}_{\alpha=1}^{k-1}|\boldsymbol{k},\boldsymbol{d})\prod_{\alpha=1}^{k-1}\pi(\Delta_\alpha|\boldsymbol{d}_\alpha,\boldsymbol{k})\,.
\end{aligned}$$

And thus from (3.4), we obtain for each $\boldsymbol{k}$ and $\boldsymbol{d}$ the relation

$$\begin{aligned}
&\pi(\Delta|\boldsymbol{k},\boldsymbol{d})=\\
&\sum_{\boldsymbol{d}_1\ldots,\boldsymbol{d}_{k-1}}P(\{\boldsymbol{d}_\alpha\}_{\alpha=1}^{k-1}|\boldsymbol{k},\boldsymbol{d})\int\left[\prod_{\alpha=1}^{k-1}d\Delta_\alpha\pi(\Delta_\alpha|\boldsymbol{d}_\alpha,\boldsymbol{k})\right]\delta\left(\Delta+\frac{1}{z+\sum_{\alpha=1}^{k-1}\Delta_\alpha}\right)\,.
\end{aligned}\tag{3.5}$$

Equation (3.5) describes a self-consistency relation for the collection of distributions $\{\pi(\Delta|\boldsymbol{k},\boldsymbol{d})\}_{\boldsymbol{k},\boldsymbol{d}}$. In possession of distributions satisfying this relation, one may recover the spectral density by performing a very similar calculation to that above, this time decomposing into a sum over the possible constrained generalised degrees

$$\rho(\lambda)=\lim_{\varepsilon\searrow 0}\frac{1}{\pi}\mathrm{Im}\sum_{\boldsymbol{k}}p(\boldsymbol{k})\int\widehat{\pi}(\Delta|\boldsymbol{k})\,d\Delta\,,\tag{3.6}$$

where $\widehat{\pi}(\Delta|\boldsymbol{k})$ is given by

$$\begin{aligned}
&\widehat{\pi}(\Delta|\boldsymbol{k})=\\
&\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_k}P(\{\boldsymbol{d}_\alpha\}_{\alpha=1}^{k}|\boldsymbol{k})\int\left[\prod_{\alpha=1}^{k}d\Delta_\alpha\pi(\Delta_\alpha|\boldsymbol{d}_\alpha,\boldsymbol{k})\right]\delta\left(\Delta+\frac{1}{z+\sum_{\alpha=1}^{k}\Delta_\alpha}\right)\,.
\end{aligned}$$

In this way, the mean spectral density of tree-like graph ensembles with specified correlations between generalised degrees can be computed by finding distributions satisfying equation (3.5).

At first sight, the above equations are far more complicated than the corresponding result for Poissonian random graphs, indeed we must now work with a potentially infinite collection of distributions. To connect with the earlier result, simplifying assumptions about the generalised degree statistics can be made in order to reduce the above equations to a more manageable form. For example, consider the case in which we take $L = 1$ (that is, we consider only ordinary degrees) and furthermore assume that degrees of next-nearest neighbours are correlated only through the central vertex, i.e.

$$P\Big(\{d_t\}_{t=1}^{k}|k\Big) = \prod_{t=1}^{k} P(d_t|k)\,, \quad \text{and} \quad P\Big(\{d_t\}_{t=1}^{k-1}|d,k\Big) = \prod_{t=1}^{k-1} P(d_t|k)\,. \tag{3.7}$$

In this case, writing $\pi(\Delta|k) = \sum_d P(d|k)\pi(\Delta|d,k)$, we obtain from (3.5)

$$\pi(\Delta|k) = \sum_d P(d|k) \int \left[\prod_{\alpha=1}^{d-1} d\Delta_\alpha \pi(\Delta_\alpha|d)\right] \delta\left(\Delta + \frac{1}{z + \sum_{\alpha=1}^{d-1}\Delta_\alpha}\right)\,, \tag{3.8}$$

and the spectral density of such an ensemble is then given by

$$\rho(\lambda) = \lim_{\varepsilon \searrow 0} \frac{1}{\pi} \mathrm{Im} \int \widehat{\pi}(\Delta)\, d\Delta\,, \tag{3.9}$$

where

$$\widehat{\pi}(\Delta) = \sum_k p(k) \int \left[\prod_{\alpha=1}^{k} d\Delta_\alpha \pi(\Delta_\alpha|k)\right] \delta\left(\Delta + \frac{1}{z + \sum_{\alpha=1}^{k}\Delta_\alpha}\right)\,. \tag{3.10}$$

Further simplification occurs if we remove all degree correlations entirely, putting $P(d|k) = P(d)$. In this case we find $\pi(\Delta|k) \equiv \pi(\Delta)$, and we recover the equations (2.25) and (2.26), derived in the previous chapter for Poisson graphs, which we now see hold for graphs with arbitrary degree distributions in the absence of degree correlations.

3.1.2 The constrained generalised degree ensemble

The calculation above provides a general procedure by which an approximation to the spectral density of any random graph ensemble may be computed from knowledge

of the statistics of generalised degrees. In the absence of additional correlations not captured in the joint distribution (3.3), this scheme is exact in the $N \to \infty$ limit. For definiteness, it is desirable to have a concrete example of a random graph ensemble for which this holds.

In [BCV08], an ensemble was introduced in which the generalised degrees of each vertex are predetermined, and edges are drawn with a bias designed to induce certain correlations between generalised degrees. This so-called constrained generalised degree ensemble will provide the example we seek. Following that work, we define the weight function

$$W(C) = \prod_{i<j} \left[\frac{c}{N} Q(\boldsymbol{k}_i, \boldsymbol{k}_j) \delta_{C_{ij},1} + \left(1 - \frac{c}{N} Q(\boldsymbol{k}_i, \boldsymbol{k}_j)\right) \delta_{C_{ij},0} \right] \prod_{i=1}^{N} \delta_{\boldsymbol{k}_i(C),\boldsymbol{k}_i} \,, \tag{3.11}$$

where $\{\boldsymbol{k}_i\}_{i=1}^N$ are taken to be arbitrary, c is the average $c = (1/N)\sum_{i=1}^N k_i$, and $Q(\boldsymbol{k}, \boldsymbol{d})$ is a symmetric, non-negative function. After normalisation, this weight specifies a probability density for adjacency matrices via

$$P(C) = \frac{W(C)}{\sum_{C'} W(C')} \,.$$

Notice that for each vertex i, the generalised degree $\boldsymbol{k}_i(C)$ is constrained by (3.11) to be precisely $\boldsymbol{k}_i$. For this reason we also refer to the $\{\boldsymbol{k}_i\}_{i=1}^N$ as generalised degrees and will often drop from the notation the dependence of $\boldsymbol{k}_i(C)$ upon C. By appropriately specifying the function Q, it is possible to manipulate the frequency of edges between vertices with specified generalised degrees, thereby introducing correlations between the degrees of nearby vertices.

This ensemble is specifically designed to be amenable to analysis in the large N limit, whilst allowing the degree statistics of the graphs to be tuned by choosing the generalised degrees $\{\boldsymbol{k}_i\}$ and the function $Q(\boldsymbol{k}, \boldsymbol{d})$. This was demonstrated in [BCV08] with the computation of the entropy of the ensemble, and in [VC08] the Ising model is analysed on such graphs in the simpler case of $L = 1$.

To apply the previous calculation of the spectral density of graphs with correlated generalised degrees to ensembles defined by (3.11), it is necessary to first determine that such ensembles are indeed tree-like in the large N limit, and then to compute the distribution $P\left(\{\boldsymbol{d}_t\}_{t=1}^k, \boldsymbol{k}\right)$ and its relevant marginals.

Since the generalised degrees are constrained to certain values, we take the generalised degree distribution $p(\boldsymbol{k})$ to be given deterministically by

$$p(\boldsymbol{k}) = \frac{1}{N}\sum_{i=1}^{N} \delta_{\boldsymbol{k},\boldsymbol{k}_i} \,,$$

which we assume is suitably well-behaved in the limit $N \to \infty$ that the dependence upon N may be suppressed in the notation.

Statistics of cycles

We begin by computing the asymptotic mean number of cycles. For $n \geq 3$ define

$$\mathcal{C}(n) = \lim_{N\to\infty} \mathbb{E}\left(\frac{1}{2n}\sum_{\substack{v_1,\cdots,v_n \\ \text{distinct}}} C_{v_1v_2}C_{v_2v_3}\cdots C_{v_{n-1}v_n}C_{v_nv_1}\right) . \tag{3.12}$$

We will evaluate the $N \to \infty$ limit here in two stages, first performing a saddle-point calculation to determine $\mathbb{E}\left(C_{v_1v_2}\cdots C_{v_nv_1}\right)$ up to leading order in N, before returning the result to (3.12) and completing the calculation of the limit. The introduction of a generating function, as employed in the previous chapter for the simple case of Poissonian graphs, will again prove useful. Introduce fields $\boldsymbol{h} = \{h_{ij}\}$, which are taken to be symmetric (i.e. $h_{ij} = h_{ji}$), and let

$$Z(\boldsymbol{h}) = \sum_{C} W(C) \prod_{i<j} e^{h_{ij}C_{ij}} \,, \tag{3.13}$$

so that

$$\mathbb{E}\left(C_{v_1v_2}\cdots C_{v_nv_1}\right) = \frac{1}{Z(\boldsymbol{0})}\left[\frac{\partial}{\partial h_{v_1v_2}}\cdots\frac{\partial}{\partial h_{v_nv_1}} Z(\boldsymbol{h})\right]_{\boldsymbol{h}=\boldsymbol{0}} . \tag{3.14}$$

It is helpful to write the Kronecker delta degree constraints appearing in (3.11) in a Fourier representation. Introducing N vectors $\boldsymbol{\omega}_i$ of L Fourier variables each, the definition of generalised degree, given in equation (3.1), is enforced by

$$\delta_{\boldsymbol{k}_i(C),\boldsymbol{k}_i} = \int_{-\pi}^{\pi} \frac{d\boldsymbol{\omega}_i}{(2\pi)^L} \exp\left(i\boldsymbol{\omega}_i \cdot \boldsymbol{k}_i - i \sum_{l=1}^{L} \omega_i^{(l)} \sum_{j=1}^{N} C_{ij} k_j^{(l-1)} \right) .$$

In this way, the weighted sum over all C in (3.13) is easily performed to gives for large N,

$$\begin{aligned} Z(\boldsymbol{h}) &= \int_{-\pi}^{\pi} \left[\prod_{i=1}^{N} \frac{d\boldsymbol{\omega}_i}{(2\pi)^L} \right] \left(\prod_{i=1}^{N} e^{i\boldsymbol{\omega}_i \cdot \boldsymbol{k}_i} \right) \prod_{i<j} \left\{ 1 - \frac{c}{N} Q(\boldsymbol{k}_i, \boldsymbol{k}_j) \right. \\ &\qquad \left. + \frac{c}{N} Q(\boldsymbol{k}_i, \boldsymbol{k}_j) \exp\left(-i \sum_{l=1}^{L} \left(\omega_i^{(l)} k_j^{(l-1)} + \omega_j^{(l)} k_i^{(l-1)} \right) + h_{ij} \right) \right\} \\ &\simeq \int_{-\pi}^{\pi} \left[\prod_{i=1}^{N} \frac{d\boldsymbol{\omega}_i}{(2\pi)^L} \right] \left(\prod_{i=1}^{N} e^{i\boldsymbol{\omega}_i \cdot \boldsymbol{k}_i} \right) \\ &\qquad \exp\left\{ \frac{c}{2N} \sum_{i,j=1}^{N} Q(\boldsymbol{k}_i, \boldsymbol{k}_j) \left(e^{-i \sum_{l=1}^{L} \left(\omega_i^{(l)} k_j^{(l-1)} + \omega_j^{(l)} k_i^{(l-1)} \right) + h_{ij}} - 1 \right) \right\} . \end{aligned} \tag{3.15}$$

We first examine the normalisation constant $Z(\mathbf{0})$. The double sum over vertex indices above is decoupled through the introduction of the order parameter

$$\varphi(\boldsymbol{k}, \boldsymbol{d}) = \frac{1}{N} \sum_{i=1}^{N} \delta_{\boldsymbol{k}_i, \boldsymbol{k}} \exp\left(-i \sum_{l=1}^{L} \omega_i^{(l)} d^{(l-1)} \right) , \tag{3.16}$$

giving

$$\sum_{i,j=1}^{N} Q(\boldsymbol{k}_i, \boldsymbol{k}_j)\, e^{-i \sum_{l=1}^{L} \left(\omega_i^{(l)} k_j^{(l-1)} + \omega_j^{(l)} k_i^{(l-1)} \right)} = N^2 \sum_{\boldsymbol{k},\boldsymbol{d}} Q(\boldsymbol{k}, \boldsymbol{d}) \varphi(\boldsymbol{k}, \boldsymbol{d}) \varphi(\boldsymbol{d}, \boldsymbol{k}) .$$

The definition (3.16) is enforced by a functional Dirac delta, which we express in terms of a conjugate order parameter ψ. All together, we may write

$$Z(\mathbf{0}) = \int \{ \mathcal{D}\varphi \mathcal{D}\psi \} \exp\Big(N \mathcal{F}(\varphi, \psi) \Big) ,$$

where

$$\begin{aligned} \mathcal{F}(\varphi, \psi) &= \frac{c}{2} \sum_{\boldsymbol{k},\boldsymbol{d}} \Big(Q(\boldsymbol{k}, \boldsymbol{d}) \Big[\varphi(\boldsymbol{k}, \boldsymbol{d}) \varphi(\boldsymbol{d}, \boldsymbol{k}) - p(\boldsymbol{k}) p(\boldsymbol{d}) \Big] - 2\psi(\boldsymbol{k}, \boldsymbol{d}) \varphi(\boldsymbol{k}, \boldsymbol{d}) \Big) \\ &\quad + \frac{1}{N} \log \prod_{i=1}^{N} \int_{-\pi}^{\pi} \frac{d\boldsymbol{\omega}_i}{(2\pi)^L} \exp\left\{ i\boldsymbol{\omega}_i \cdot \boldsymbol{k}_i + c \sum_{\boldsymbol{d}} \psi(\boldsymbol{k}_i, \boldsymbol{d})\, e^{-i \sum_{l=1}^{L} \omega_i^{(l)} d^{(l-1)}} \right\} . \end{aligned} \tag{3.17}$$

The integration over Fourier variables in the last line can be performed explicitly by expanding the exponential in Taylor series,

$$\begin{aligned}
&\int_{-\pi}^{\pi} \frac{d\boldsymbol{\omega}_i}{(2\pi)^L} \exp\left(i\boldsymbol{\omega}_i \cdot \boldsymbol{k}_i + c \sum_{\boldsymbol{d}} \psi(\boldsymbol{k}_i, \boldsymbol{d})\, e^{-i\sum_{l=1}^{L} \omega_i^{(l)} d^{(l-1)}} \right) \\
&= \sum_{s=0}^{\infty} \frac{c^s}{s!} \int_{-\pi}^{\pi} \frac{d\boldsymbol{\omega}_i}{(2\pi)^L} \exp\left(i\boldsymbol{\omega}_i \cdot \boldsymbol{k}_i\right) \left(\sum_{\boldsymbol{d}} \psi(\boldsymbol{k}_i, \boldsymbol{d})\, e^{-i\sum_{l=1}^{L} \omega_i^{(l)} d^{(l-1)}} \right)^s \\
&= \sum_{s=0}^{\infty} \frac{c^s}{s!} \sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_s} \psi(\boldsymbol{k}_i, \boldsymbol{d}_1) \cdots \psi(\boldsymbol{k}_i, \boldsymbol{d}_s) \prod_{l=1}^{L} \int_{-\pi}^{\pi} \frac{d\boldsymbol{\omega}_i}{2\pi} e^{i\omega_i^{(l)} \left(k_i^{(l)} - \sum_{\alpha=1}^{s} d_\alpha^{(l-1)} \right)} \\
&= \frac{c^{k_i}}{k_i!} \sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_{k_i}} \psi(\boldsymbol{k}_i, \boldsymbol{d}_1) \cdots \psi(\boldsymbol{k}_i, \boldsymbol{d}_{k_i}) \prod_{l=1}^{L} \delta_{k_i^{(l)}, \sum_{r=1}^{k_i} d_r^{(l-1)}} \,.
\end{aligned} \tag{3.18}$$

Here, the Kronecker delta corresponding to $l = 1$ forces $s = k_i$, which results in the cancellation of the sum over s. To simplify the equations a little, it will be convenient to introduce a shorthand notation for the constraint. Let

$$\mathbb{I}_{\boldsymbol{k}}\left(\{\boldsymbol{d}_r\}_{r=1}^{k}\right) = \prod_{l=1}^{L} \delta_{k^{(l)}, \sum_{r=1}^{k} d_r^{(l-1)}} \,.$$

Returning this and (3.18) to equation (3.17), we may write

$$\begin{aligned}
\mathcal{F}(\varphi, \psi) =& \frac{c}{2} \sum_{\boldsymbol{k},\boldsymbol{d}} \left(Q(\boldsymbol{k}, \boldsymbol{d}) \Big[\varphi(\boldsymbol{k}, \boldsymbol{d}) \varphi(\boldsymbol{d}, \boldsymbol{k}) - p(\boldsymbol{k}) p(\boldsymbol{d}) \Big] - 2\psi(\boldsymbol{k}, \boldsymbol{d}) \varphi(\boldsymbol{k}, \boldsymbol{d}) \right) \\
&+ \sum_{\boldsymbol{k}} p(\boldsymbol{k}) \log \frac{c^k}{k!} \sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_k} \psi(\boldsymbol{k}, \boldsymbol{d}_1) \cdots \psi(\boldsymbol{k}, \boldsymbol{d}_k)\, \mathbb{I}_{\boldsymbol{k}}\left(\{\boldsymbol{d}_r\}_{r=1}^{k}\right) .
\end{aligned}$$

We move on now to the calculation of $\mathbb{E}\left(C_{v_1 v_2} \cdots C_{v_n v_1}\right)$. For simplicity we enforce the periodic boundary conditions $v_{n+1} = v_1$ and $v_0 = v_n$, then applying the deriva-

tives in (3.14) to equation (3.15), give, for large N,

$$\begin{aligned}
&\mathbb{E}\left(C_{v_1v_2}\cdots C_{v_nv_1}\right)\\
&\quad\simeq \frac{1}{Z(\mathbf{0})}\left[\prod_{t=1}^{n}\frac{c}{N}Q(\boldsymbol{k}_{v_t},\boldsymbol{k}_{v_{t+1}})\right]\int_{-\pi}^{\pi}\left[\prod_{i=1}^{N}\frac{d\boldsymbol{\omega}_i}{(2\pi)^L}\right]\left(\prod_{i=1}^{N}e^{i\boldsymbol{\omega}_i\cdot\boldsymbol{k}_i}\right)\\
&\qquad\exp\left\{\frac{c}{2N}\sum_{i,j=1}^{N}Q(\boldsymbol{k}_i,\boldsymbol{k}_j)\left(e^{-i\sum_{l=1}^{L}\left(\omega_i^{(l)}k_j^{(l-1)}+\omega_j^{(l)}k_i^{(l-1)}\right)}-1\right)\right\}\\
&\qquad\exp\left\{-i\sum_{t=1}^{n}\sum_{l=1}^{L}\left(\omega_{v_t}^{(l)}k_{v_{t+1}}^{(l-1)}+\omega_{v_{t+1}}^{(l)}k_{v_t}^{(l-1)}\right)\right\}\\
&\quad=\left[\prod_{t=1}^{n}\frac{c}{N}Q(\boldsymbol{k}_{v_t},\boldsymbol{k}_{v_{t+1}})\right]\frac{\int\{\mathcal{D}\varphi\mathcal{D}\psi\}\,\exp\left(N\mathcal{F}_\circ(\varphi,\psi)\right)}{\int\{\mathcal{D}\varphi\mathcal{D}\psi\}\,\exp\left(N\mathcal{F}(\varphi,\psi)\right)},
\end{aligned}\tag{3.19}$$

where

$$\begin{aligned}
\mathcal{F}_\circ(\varphi,\psi)=\frac{c}{2}\sum_{\boldsymbol{k},\boldsymbol{d}}&\left(Q(\boldsymbol{k},\boldsymbol{d})\Big[\varphi(\boldsymbol{k},\boldsymbol{d})\varphi(\boldsymbol{d},\boldsymbol{k})-p(\boldsymbol{k})p(\boldsymbol{d})\Big]-2\psi(\boldsymbol{k},\boldsymbol{d})\varphi(\boldsymbol{k},\boldsymbol{d})\right)\\
&+\frac{1}{N}\log\int_{-\pi}^{\pi}\left[\prod_{i=1}^{N}\frac{d\boldsymbol{\omega}_i}{(2\pi)^L}\right]\exp\left\{-i\sum_{t=1}^{n}\sum_{l=1}^{L}\left(\omega_{v_t}^{(l)}k_{v_{t+1}}^{(l-1)}+\omega_{v_{t+1}}^{(l)}k_{v_t}^{(l-1)}\right)\right\}\\
&\qquad\exp\left\{\sum_{i=1}^{N}\left(i\boldsymbol{\omega}_i\cdot\boldsymbol{k}_i+c\sum_{\boldsymbol{d}}\psi(\boldsymbol{k}_i,\boldsymbol{d})\,e^{-i\sum_{l=1}^{L}\omega_i^{(l)}d^{(l-1)}}\right)\right\}.
\end{aligned}$$

For the integration over Fourier variables here, we distinguish two cases: if $i\neq v_1,\ldots,v_n$ we recover expression (3.18), whereas if $i=v_t$ for some t, the same technique gives

$$\begin{aligned}
&\int_{-\pi}^{\pi}\frac{d\boldsymbol{\omega}_{v_t}}{(2\pi)^L}e^{i\sum_{l=1}^{L}\omega_{v_t}^{(l)}\left(k_{v_t}^{(l)}-k_{v_{t+1}}^{(l-1)}-k_{v_{t-1}}^{(l-1)}\right)+c\sum_{\boldsymbol{d}}\psi(\boldsymbol{k}_{v_t},\boldsymbol{d})\,e^{-i\sum_{l=1}^{L}\omega_{v_t}^{(l)}d^{(l-1)}}}\\
&=\frac{c^{(k_{v_t}-2)}}{(k_{v_t}-2)!}\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_{k_{v_t}-2}}\psi(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi(\boldsymbol{k},\boldsymbol{d}_{k_{v_t}-2})\,\mathbb{I}_{\boldsymbol{k}}\big(\boldsymbol{k}_{v_t+1},\boldsymbol{k}_{v_t-1},\{\boldsymbol{d}_r\}_{r=1}^{k_{v_t}-2}\big).
\end{aligned}$$

Together, we reach

$$\begin{aligned}
\mathcal{F}_\circ(\varphi,\psi)=\mathcal{F}(\varphi,\psi)+\frac{1}{N}\log\prod_{t=1}^{n}\Bigg[&\frac{k_{v_t}(k_{v_t}-1)}{c^2}\\
&\times\frac{\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_{k_{v_t}-2}}\psi(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi(\boldsymbol{k},\boldsymbol{d}_{k_{v_t}-2})\,\mathbb{I}_{\boldsymbol{k}}\big(\boldsymbol{k}_{v_t+1},\boldsymbol{k}_{v_t-1},\{\boldsymbol{d}_r\}_{r=1}^{k_{v_t}-2}\big)}{\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_k}\psi(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi(\boldsymbol{k},\boldsymbol{d}_k)\,\mathbb{I}_{\boldsymbol{k}}\big(\{\boldsymbol{d}_r\}_{r=1}^{k}\big)}\Bigg],
\end{aligned}$$

and returning this to (3.19) yields

$$\begin{aligned}
&\mathbb{E}\left(C_{v_1v_2}\cdots C_{v_nv_1}\right)\\
&=\left\langle \prod_{t=1}^{n}\frac{k_{v_t}(k_{v_t}-1)}{cN}\,Q(\boldsymbol{k}_{v_t},\boldsymbol{k}_{v_{t+1}})\right.\\
&\left.\times\frac{\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_{k_{v_t}-2}}\psi(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi(\boldsymbol{k},\boldsymbol{d}_{k_{v_t}-2})\,\mathbb{I}_{\boldsymbol{k}}\big(\boldsymbol{k}_{v_t+1},\boldsymbol{k}_{v_t-1},\{\boldsymbol{d}_r\}_{r=1}^{k_{v_t}-2}\big)}{\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_k}\psi(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi(\boldsymbol{k},\boldsymbol{d}_k)\,\mathbb{I}_{\boldsymbol{k}}\big(\{\boldsymbol{d}_r\}_{r=1}^{k}\big)}\right\rangle_{\mathcal{F}(\varphi,\psi)},
\end{aligned}$$

where the measure is defined by

$$\langle\cdots\rangle_{\mathcal{F}(\varphi,\psi)}=\frac{\displaystyle\int\{\mathcal{D}\varphi\mathcal{D}\psi\}\,\exp\Big(N\mathcal{F}(\varphi,\psi)\Big)(\cdots)}{\displaystyle\int\{\mathcal{D}\varphi\mathcal{D}\psi\}\,\exp\Big(N\mathcal{F}(\varphi,\psi)\Big)}\,. \tag{3.20}$$

In the limit $N\to\infty$, this measure approaches a functional Dirac delta centred at the saddle point $(\varphi_\star,\psi_\star)$ of $\mathcal{F}(\varphi,\psi)$. Extremising $\mathcal{F}(\varphi,\psi)$ with respect to φ and ψ, we obtain the saddle-point equations

$$\begin{aligned}
\psi_\star(\boldsymbol{k},\boldsymbol{d})&=Q(\boldsymbol{k},\boldsymbol{d})\varphi_\star(\boldsymbol{d},\boldsymbol{k})\,,\\
\varphi_\star(\boldsymbol{k},\boldsymbol{d})&=\frac{k\,p(\boldsymbol{k})}{c}\frac{\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_{k-1}}\psi_\star(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi_\star(\boldsymbol{k},\boldsymbol{d}_{k-1})\,\mathbb{I}_{\boldsymbol{k}}\big(\boldsymbol{d},\{\boldsymbol{d}_r\}_{r=1}^{k-1}\big)}{\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_k}\psi_\star(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi_\star(\boldsymbol{k},\boldsymbol{d}_k)\,\mathbb{I}_{\boldsymbol{k}}\big(\{\boldsymbol{d}_r\}_{r=1}^{k}\big)}\,.
\end{aligned}$$

The asymptotic mean number of cycles of length n is then given in terms of $\psi_\star$ by

$$\begin{aligned}
\mathcal{C}(n)=\lim_{N\to\infty}\frac{1}{2n}\sum_{\substack{v_1,\cdots,v_n\\ \text{distinct}}}\prod_{t=1}^{n}\Bigg[&\frac{k_{v_t}(k_{v_t}-1)}{cN}\,Q(\boldsymbol{k}_{v_t},\boldsymbol{k}_{v_{t+1}})\\
&\times\frac{\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_{k_{v_t}-2}}\psi_\star(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi_\star(\boldsymbol{k},\boldsymbol{d}_{k_{v_t}-2})\,\mathbb{I}_{\boldsymbol{k}}\big(\boldsymbol{k}_{v_t+1},\boldsymbol{k}_{v_t-1},\{\boldsymbol{d}_r\}_{r=1}^{k_{v_t}-2}\big)}{\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_k}\psi_\star(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi_\star(\boldsymbol{k},\boldsymbol{d}_k)\,\mathbb{I}_{\boldsymbol{k}}\big(\{\boldsymbol{d}_r\}_{r=1}^{k}\big)}\Bigg].
\end{aligned} \tag{3.22}$$

Note that there are $N!/(N-n)!$ terms in the sum over ordered collections of n distinct vertices, and each term is of order N^{-n}, so we can conclude that the asymptotic mean number of cycles of length n is an order 1 quantity, regardless of the detail of Q and $p(\boldsymbol{k})$ (provided its first two moments are finite).

For general L and Q, solution of the saddle point equations (3.21) is likely to be a cumbersome task, making exact evaluation of $\mathcal{C}(n)$ impractical. We are, however, able

to obtain exact results for some special cases which offer significant simplification. For instance, if we are interested in constraining only degrees, setting $L = 1$ yields $\mathbb{I}_{\boldsymbol{k}} \equiv 1$ in all of the above. Hence, introducing $\Psi(k) = \sum_d \psi_\star(k, d)$, the saddle point equations (3.21) reduce to

$$\Psi(k) = \sum_d Q(k, d) \frac{d\, p(d)}{c} \big(\Psi(d)\big)^{-1}, \tag{3.23}$$

and (3.22) becomes

$$\mathcal{C}(n) = \lim_{N\to\infty} \frac{1}{2n} \sum_{\substack{v_1,\cdots,v_n \\ \text{distinct}}} \prod_{t=1}^{n} \frac{k_{v_t}(k_{v_t} - 1) Q(k_{v_t}, k_{v_{t+1}})}{c\, N\, \Psi(k_{v_{t-1}}) \Psi(k_{v_{t+1}})}. \tag{3.24}$$

Further simplification occurs in the event that the function Q is separable, in which case it is easily determined from (3.23) that $Q(k, d) = \Psi(k)\Psi(d)$. Assuming the degrees of any collection of n vertices to be approximately independent for large N, we obtain from (3.24) the explicit formula

$$\mathcal{C}(n) = \frac{1}{2n} \left(\sum_k \frac{k\, p(k)}{c} (k - 1) \right)^n.$$

This result was derived for the special case $n = 3$ in [BJK04] using techniques which are easily applied to recover the same result as above for general n.

Statistics of stars

The same techniques may be employed in the calculation of the joint distribution

$$P\left(\{\boldsymbol{d}_t\}_{t=1}^k, \boldsymbol{k}\right) = \lim_{N\to\infty} \mathbb{E}\left(\frac{k}{cN} \sum_{i=1}^{N} \delta_{\boldsymbol{k}_i(C), \boldsymbol{k}} \sum_{j_1 < \cdots < j_k} \delta_{\boldsymbol{k}_{j_1}, \boldsymbol{d}_1} \cdots \delta_{\boldsymbol{k}_{j_k}, \boldsymbol{d}_k} C_{ij_1} \cdots C_{ij_k} \right)$$

for the constrained generalised degree ensemble. Proceeding as above, one arrives at

$$\begin{aligned}
&\mathbb{E}\left(C_{ij_1}\cdots C_{ij_{k_i}}\right)\\
&= k_i!\Bigg\langle \frac{\mathbb{I}_{\boldsymbol{k}_i}\big(\{\boldsymbol{k}_{j_t}\}_{t=1}^{k_i}\big)}{\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_{k_i}}\psi(\boldsymbol{k}_i,\boldsymbol{d}_1)\cdots\psi(\boldsymbol{k}_i,\boldsymbol{d}_{k_i})\mathbb{I}_{\boldsymbol{k}_i}\big(\{\boldsymbol{d}_r\}_{r=1}^{k_i}\big)}\prod_{t=1}^{k_i}\Bigg[\frac{k_{j_t}Q(\boldsymbol{k}_i,\boldsymbol{k}_{j_t})}{cN}\\
&\times\frac{\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_{k_{j_t}-1}}\psi(\boldsymbol{k}_{j_t},\boldsymbol{d}_1)\cdots\psi(\boldsymbol{k}_{j_t},\boldsymbol{d}_{k_{j_t}-1})\,\mathbb{I}_{\boldsymbol{k}_{j_t}}\big(\boldsymbol{k}_i,\{\boldsymbol{d}_r\}_{r=1}^{k_{j_t}-1}\big)}{\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_{k_{j_t}}}\psi(\boldsymbol{k}_{j_t},\boldsymbol{d}_1)\cdots\psi(\boldsymbol{k}_{j_t},\boldsymbol{d}_{k_{j_t}})\,\mathbb{I}_{\boldsymbol{k}_{j_t}}\big(\{\boldsymbol{d}_r\}_{r=1}^{k_{j_t}}\big)}\Bigg]\Bigg\rangle_{\mathcal{F}(\varphi,\psi)},
\end{aligned}$$

where the measure is the same as found previously in (3.20). At the saddle point $(\varphi_\star,\psi_\star)$, we may use the above to evaluate (3.3), obtaining

$$P\left(\{\boldsymbol{d}_t\}_{t=1}^{k},\boldsymbol{k}\right)=P(\boldsymbol{k})\frac{\psi_\star(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi_\star(\boldsymbol{k},\boldsymbol{d}_k)\mathbb{I}_{\boldsymbol{k}}\big(\{\boldsymbol{d}_t\}_{t=1}^{k}\big)}{\sum_{\boldsymbol{d}'_1,\ldots,\boldsymbol{d}'_k}\psi_\star(\boldsymbol{k},\boldsymbol{d}'_1)\cdots\psi_\star(\boldsymbol{k},\boldsymbol{d}'_k)\mathbb{I}_{\boldsymbol{k}}\big(\{\boldsymbol{d}'_r\}_{r=1}^{k}\big)}\,.$$

For the various marginals appearing in the computation of the ensemble average spectral density, we obtain

$$P(\boldsymbol{d},\boldsymbol{k})=Q(\boldsymbol{k},\boldsymbol{d})\varphi_\star(\boldsymbol{k},\boldsymbol{d})\varphi_\star(\boldsymbol{d},\boldsymbol{k}),$$

also

$$P\big(\{\boldsymbol{d}_t\}_{t=1}^{k}|\boldsymbol{k}\big)=\frac{\psi_\star(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi_\star(\boldsymbol{k},\boldsymbol{d}_k)\mathbb{I}_{\boldsymbol{k}}\left(\{\boldsymbol{d}_t\}_{t=1}^{k}\right)}{\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_k}\psi_\star(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi_\star(\boldsymbol{k},\boldsymbol{d}_k)\mathbb{I}_{\boldsymbol{k}}\left(\{\boldsymbol{d}_t\}_{t=1}^{k}\right)},$$

and similarly

$$P\big(\{\boldsymbol{d}_t\}_{t=1}^{k-1}|\boldsymbol{k},\boldsymbol{d}\big)=\frac{\psi_\star(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi_\star(\boldsymbol{k},\boldsymbol{d}_{k-1})\mathbb{I}_{\boldsymbol{k}}\left(\boldsymbol{d},\{\boldsymbol{d}_t\}_{t=1}^{k-1}\right)}{\sum_{\boldsymbol{d}_1,\ldots,\boldsymbol{d}_{k-1}}\psi_\star(\boldsymbol{k},\boldsymbol{d}_1)\cdots\psi_\star(\boldsymbol{k},\boldsymbol{d}_{k-1})\mathbb{I}_{\boldsymbol{k}}\left(\boldsymbol{d},\{\boldsymbol{d}_t\}_{t=1}^{k-1}\right)}\,.$$

We now have access to all the ingredients necessary to compute the spectral density for the constrained generalised degree ensemble. Having determined that the graphs produced are indeed tree-like, and derived expressions for the generalised degree statistics, equations (3.5) and (3.6) may be applied to determine the spectral density. It is worth mentioning that this task may also be accomplished using the replica method, as in [RPVT10], with identical results.

We also point out that considerable simplification is again possible in the case $L=1$. In this circumstance we find

$$P(d,k)=\left(\frac{k\,p(k)}{c}\right)\left(\frac{d\,p(d)}{c}\right)\frac{Q(k,d)}{\Psi(k)\Psi(d)}\,,$$

and

$$P(d|k) = \left(\frac{d\,p(d)}{c}\right) \frac{Q(k,d)}{\Psi(k)\Psi(d)}\,, \tag{3.25}$$

where Ψ satisfies the simplified saddle point equation (3.23). Moreover, the star distributions now factorise as in (3.7), and the spectral density is determined by the simplified equations (3.8) and (3.9).

3.1.3 Numerical simulations

For general ensembles, solving the set of distributional recursion relations specified in (3.5) by hand may well be out of the question, however we can still extract meaningful results through the use of numerical simulations. The population dynamics algorithm (Algorithm 3) described in the previous chapter can easily be generalised to approximately generate large samples from the distributions described by equations (3.5) and (3.8), which can then be used to predict the ensemble average spectral density via equation (3.6) or (3.9).

To assess the results obtained it will be desirable to compare with data gathered from numerical diagonalisation of randomly generated graphs from ensembles with the same degree statistics. In the case $L = 1$, this can be achieved through the use of Algorithm 2, given in the introduction. That algorithm takes as its input a degree sequence $\bar{k}$ and a degree correlation function $P(k,d)$, however, our results are given in terms of the degree distribution $p(k)$. We therefore need to generate degree sequences which are compatible with $p(k)$. We discuss two possibilities:

Random degrees:

For each instance, the degrees are chosen independently randomly from $p(k)$. There is a chance that no graph can be generated that exactly fits the resulting degree sequence $\overline{k}$, in this case, the degree sequence is said to be non-graphical.

To deal with this, one may choose to check the graphicality of $\overline{k}$ before generating graphs. There are various ways to check graphicality, for instance, a theorem of Erdős and Gallai states that a degree sequence $\overline{k}$ with $k_1 \geq \cdots \geq k_N$ and $\sum_{i=1}^{N} k_i$ even is graphical if and only if, for all $n = 1, ..., N$

$$\sum_{i=1}^{n} k_i \leq n(n-1) + \sum_{i=n+1}^{N} \min\{k_i, n\}\,.$$

See [BD06] for a discussion on graphicality and the generation of random graphs.

Static degrees:

Select a set of positive integers $\{N, N_1, N_2, \ldots\}$ such that $N = \sum_k N_k$ and for each k, N_k/N is approximately $p(k)$. We then generate random degree sequences with N_1 nodes of degree one, N_2 vertices of degree 2, and so on.

Although no mathematical analysis of Algorithm 2 has been conducted, experimental evidence suggests it yields graphs with the desired properties and produces almost no failures, provided one selects the appropriate method of generating degree sequences.

Example: power-law degree distribution

Let us consider a simple random graph model with correlated degrees and a power-law degree distribution, as is often observed in empirical studies of real-world complex networks. We choose $P(k, d) \propto \tau^{kd}$, where $\tau < 1$, with a maximum value for k being $k_{\max}$. In the limits $k_{\max} \to \infty$ and $\tau \to 1$, this choice results in a power-law degree distribution with exponent 2, though for the purpose of simulations, we will take $\tau = 0.999$, and keep $k_{\max}$ finite.

Using Algorithm 2, 500 random graphs of size $N = 2000$ were generated with the desired degree statistics. For this value of N, is it necessary to take a rather low maximum degree $k_{\max} = 45\,(< \sqrt{N}\,)$; any larger and additional correlations occur be-

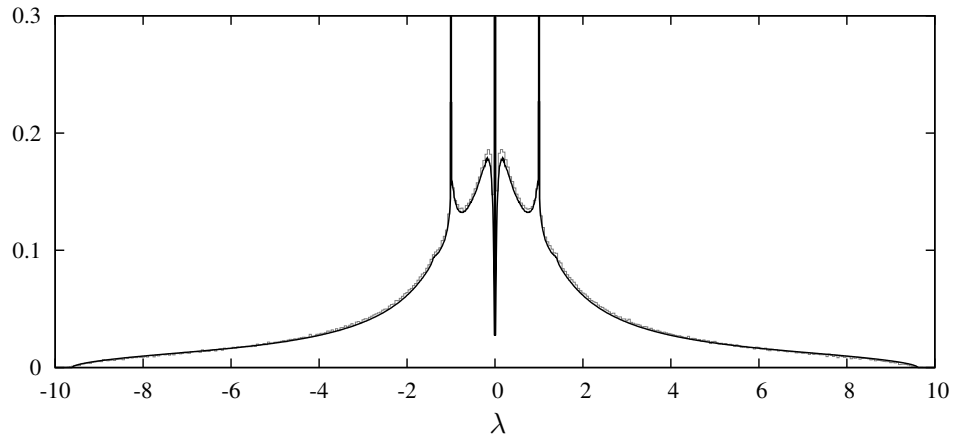


Figure 3.1: Comparison of the results of population dynamics (black line) and direct diagonalisation (grey histogram) for the choice $P(k,d) \propto \tau^{kd}$, where $\tau = 0.999$ and $k_{\max} = 45$. Algorithm 2 was used to generate 500 graphs of size $N = 2000$, whose eigenvalues were used to construct the histogram.

tween high degree vertices [CBPS05], and the failure rate becomes unacceptable. Alternatively, one could have larger values of $k_{\max}$ by increasing the graph size, but that would make the numerical diagonalisation computationally expensive. The degree sequences were generated randomly and checked for graphicality before generating the graphs.

Figure 3.1 shows a comparison between the results of population dynamics and a histogram of eigenvalues from 500 random graphs of size $N = 2000$.

Some salient features of the spectral density can be explained in terms of the underlying graph structure. First we note that the spectral density for $k_{\max} = 45$ and $\tau = 0.999$ has a bounded support, in fact, a simple application of the Perron-Frobenius theorem shows that the support of the spectral density is bounded by $\pm k_{\max}$ (although here this bound is clearly not at all sharp). The picture changes dramatically in the limits

$k_{\max} \to \infty$ and $\tau \to 1$. The mean degree can be calculated for $k_{\max} = \infty$,

$$c = \frac{\sum_{k,d=1}^{\infty} \tau^{kd}}{\sum_{k,d=1}^{\infty} \frac{1}{k}\tau^{kd}} = \frac{\sum_{k=1}^{\infty} \frac{1}{1-\tau^k}}{\sum_{k=1}^{\infty} \ln\left(\frac{1}{1-\tau^k}\right)},$$

and simple integral bounds show that $c \to \infty$ as $\tau \to 1$. We can thus conclude that (since the variance of the spectral density is given by the mean of the degree distribution) the spectral density will be heavy-tailed in these limits.

Reasoning further, one may suppose that, for vertices of very high degree, it may be sufficient to consider only the mean behaviour of the neighbouring vertices. Taking the approximation $\varphi(\Delta|k) \simeq \delta(\Delta - \Delta_k)$, we obtain from (3.8) the correlated form of the effective medium approximation (EMA) derived in [DGMS03],

$$\Delta_k = \sum_d P(d|k) \frac{1}{-z - \sum_{t=1}^{k-1} \Delta_d}.$$

As shown in [DGMS03], this equation can be used to compute the approximate behaviour of the tails of the spectral density, specifically, for very large $|\lambda|$ we have $\rho(\lambda) \simeq 2k_\lambda p(k_\lambda)/|\lambda|$, where $k_\lambda = \lambda^2 + \mathcal{O}(1)$. The contribution to the spectral density made by vertices with high degree was analysed more rigorously in [MP02]; stated roughly, the largest eigenvalues of graphs with heavy-tailed degree distributions occur close to the square roots of the largest degrees.

Figure 3.2 shows the results of population dynamics simulations in the tail of the density for a much larger maximum degree of $k_{\max} = 400$. The approximate curve given by the EMA gives a reasonable fit with the result of the simulation, and, as expected, the density drops dramatically shortly after $\sqrt{k_{\max}} = 20$. The contributions to the density coming from high degree vertices can be isolated in the output of the population dynamics algorithm; several such contributions have been included in Figure 3.2, each of which exhibits sharp peak close to $\sqrt{k}$.

The other main feature of the density shown in Figure 3.1 is the presence of Dirac delta peaks at -1, 0 and 1. The distributions $p(k)$ and $P(k|d)$ can be used to give

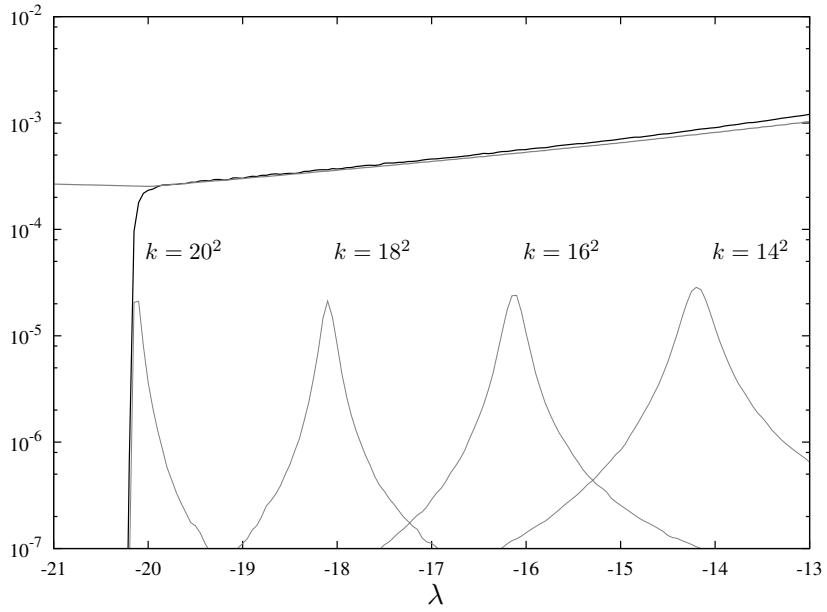


Figure 3.2: Detail of the tail of the spectral density for the choice $P(k, d) \propto \tau^{kd}$, where $\tau = 0.999$ and $k_{\max} = 400$. The black line is the full result from population dynamics, with the labelled thin grey lines being isolated contributions from vertices of high degree. The thick grey line shows an estimate for the tail derived from the effective medium approximation.

lower bounds for the weight of these peaks by computing the probability of specific configurations of vertices in the graph. For example, every connected pair of vertices of degree 1 contributes a pair of ± 1 eigenvalues to the spectrum. For the choice $k_{\max} = 45$ and $\tau = 0.999$, a lower bound for the weight of the peaks at ± 1 can be found by $p(1)P(1|1)/2 = 0.0031$. This compares well with the data from direct diagonalisation, giving a weight to the peaks of 0.0037 each. In the limits $k_{\max} \to \infty$ and $\tau \to 1$, we find

$$p(1)P(1|1) = \tau \left(\sum_{k=1}^{\infty} \ln \left(\frac{1}{1-\tau^k} \right) \right)^{-1} \to 0 \quad \text{as} \quad \tau \to 1 \,,$$

so this contribution to the density is vanishing in these limits. A lower bound for the weight of the peak at zero can be found in terms of the size, K, of largest collection of mutually disconnected vertices, specifically, the weight is at least $2K/N - 1$. A crude estimate for K can be constructed by counting the frequency of vertices connected only to vertices of higher degree. For the choice $k_{\max} = 45$ and $\tau = 0.999$ we obtain

$$2 \sum_{k=1}^{k_{\max}} p(k) \left(\sum_{d=k+1}^{k_{\max}} P(d|k) \right)^{k} - 1 = 0.1041 \,.$$

This time the data from direct diagonalisation gives a weight to the peak of 0.1828, so whilst the bound is certainly not sharp, it is at least the right order of magnitude.

Example: levels of approximation

Suppose that ones knowledge of a graph ensemble is limited solely to a set of statistical properties captured by, for instance, the degree distribution $p(k)$ and the conditional distribution $P(k|d)$. Whilst in a few cases such quantities suffice to fully characterise the graph ensemble, this is not generally true. In the following, a simple example is used to illustrate the effects on the spectral density caused by reducing the amount of information known about the underlying graph ensemble.

Put $L = 2$ and consider a graph ensemble with generalised degrees taken from the set

$\{\kappa_2, \kappa_3, \kappa_4\}$, where

$$\kappa_2 = \begin{pmatrix} 2 \\ 7 \end{pmatrix}, \quad \kappa_3 = \begin{pmatrix} 3 \\ 6 \end{pmatrix}, \quad \kappa_4 = \begin{pmatrix} 4 \\ 8 \end{pmatrix}. \tag{3.26}$$

This ensemble is composed of graphs with vertices of degrees 2, 3 and 4. Moreover, those vertices of degree 2 must be connected to one vertex of degree 3 and one of degree 4, and those of degrees 3 and 4 must be connected to vertices of degree 2 only. A portion of such a graph is shown in the leftmost part of Figure 3.1.3.

Simple counting arguments give the following expressions for the degree statistics:

$$p(k) = \frac{12}{19}\delta_{k,2} + \frac{4}{19}\delta_{k,3} + \frac{3}{19}\delta_{k,4}\,,$$

$$P(k|d) = \delta_{k,2}\delta_{d,4} + \delta_{k,2}\delta_{d,3} + \left(\frac{1}{2}\delta_{k,3} + \frac{1}{2}\delta_{k,4}\right)\delta_{d,2}\,,$$

knowledge of which does not fully characterise the graph ensemble. The conditional distribution $P\left(\{\boldsymbol{d}_t\}_{t=1}^{k-1}|\boldsymbol{k}, \boldsymbol{d}\right)$ is more restrictive however;

$$\begin{aligned} &P(\kappa_3|\kappa_2, \kappa_4) = 1\,, \quad P(\kappa_4|\kappa_2, \kappa_3) = 1\,, \\ &P(\kappa_2, \kappa_2|\kappa_3, \kappa_3) = 1\,, \quad P(\kappa_2, \kappa_2, \kappa_2|\kappa_4, \kappa_2) = 1\,, \end{aligned} \tag{3.28}$$

or zero otherwise. The deterministic form of the distributions (3.28) mean that the equations (3.5) may in this case be solved by hand, in particular, we have $\pi(\Delta|\kappa_a, \kappa_b) = \delta(\Delta - \Delta_{a,b})$, where the $\{\Delta_{a,b}\}$ satisfy

$$\begin{aligned} \Delta_{2,3} &= \frac{1}{-z - \Delta_{4,2}}\,, \ \Delta_{2,4} = \frac{1}{-z - \Delta_{3,2}}\,, \\ \Delta_{3,2} &= \frac{1}{-z - 2\Delta_{2,3}}\,, \ \Delta_{4,2} = \frac{1}{-z - 3\Delta_{2,4}}\,. \end{aligned} \tag{3.29}$$

To find an expression for the spectral density we are required to compute the joint distribution $P(\{\boldsymbol{d}_t\}_{t=1}^{k}|\boldsymbol{k})$, which in this case reads

$$\begin{aligned} &P(\kappa_3, \kappa_4|\kappa_2) = \tfrac{1}{2}\,, \quad P(\kappa_4, \kappa_3|\kappa_2) = \tfrac{1}{2}\,, \\ &P(\kappa_2, \kappa_2, \kappa_2|\kappa_3) = 1\,, \quad P(\kappa_2, \kappa_2, \kappa_2, \kappa_2|\kappa_4) = 1\,, \end{aligned}$$

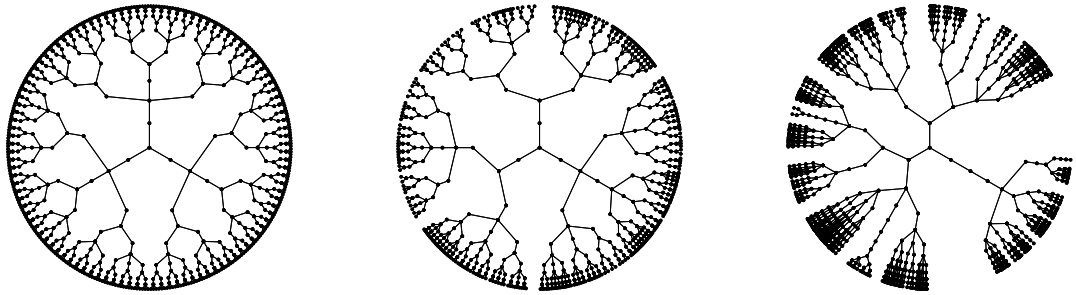


Figure 3.3: Left - a typical neighbourhood of a vertex of degree 3 in a large random graph specified by (3.26). Middle - a neighbourhood of a vertex of degree 3 in a random graph with degree distribution (3.27) and $P(k|d)$ given by (3.27). Right - a neighbourhood of a vertex of degree 3 in an uncorrelated random graph with degree distribution (3.27).

or zero otherwise. This yields

$$\rho(\lambda) = -\lim_{\varepsilon \searrow 0} \frac{1}{19\pi} \mathrm{Im} \left[\frac{12}{\lambda + i\varepsilon + \Delta_{3,2} + \Delta_{4,2}} + \frac{4}{\lambda + i\varepsilon + 3\Delta_{2,3}} + \frac{3}{\lambda + i\varepsilon + 4\Delta_{2,4}} \right].$$

After solving (3.29), plugging the solutions into the above equation for $\rho(\lambda)$ and carefully analysing the poles, we may write

$$\begin{aligned} \rho(\lambda) = \frac{5}{19}\delta(\lambda) + \frac{1}{19}\delta(\lambda + \sqrt{3}) + \frac{1}{19}\delta(\lambda - \sqrt{3}) \\ + \frac{12|2\lambda^2 - 7|\sqrt{-25 - \lambda^2(-7 + \lambda^2)(14 - 7\lambda^2 + \lambda^4)}}{19\pi|\lambda(\lambda^2 - 4)(\lambda^2 - 7)(\lambda^2 - 3)|} \mathbb{I}_D(|\lambda|) \end{aligned} \quad (3.30)$$

with the domain $D = [\lambda_+^-, \lambda_-^-] \cup [\lambda_-^+, \lambda_+^+]$ and for $\alpha, \beta \in \{-, +\}$,

$$\lambda_\alpha^\beta = \frac{1}{2}\sqrt{14 + \beta 2\sqrt{21 + \alpha 8\sqrt{6}}}\,. \quad (3.31)$$

This example was chosen specifically to keep the local structure of the graphs deterministic and thus make equations (3.5) exactly solvable in the case $L = 2$. If we reduce the statistical information we have about the ensemble, either by taking $L = 1$, or further by assuming the degrees to be uncorrelated, the local graph structure becomes

random, and the resulting expressions are no longer exactly solvable. Figure 3.1.3 illustrates the randomising effect of these simplifications.

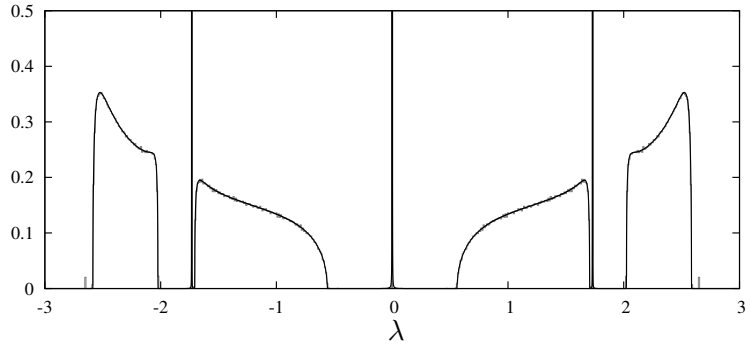
Population dynamics algorithms were used to compute the spectral density of random graphs with degree distribution (3.27) and degree correlations (3.27), as well as uncorrelated random graphs with degree distribution (3.27). The results of these simulations are shown in Figures 3.4(b) and 3.4(c), alongside histograms of the eigenvalues of 1000 graphs of size $N = 1900$, generated using Algorithm 2. Figure 3.4(a) shows the exact spectral density given by (3.30), plotted alongside a histogram of eigenvalues obtained from randomly generated graphs of that type.

Algorithm 2 is not sufficient to generate random graphs specified by (3.26). Instead the following simple procedure may be used, with satisfactory results

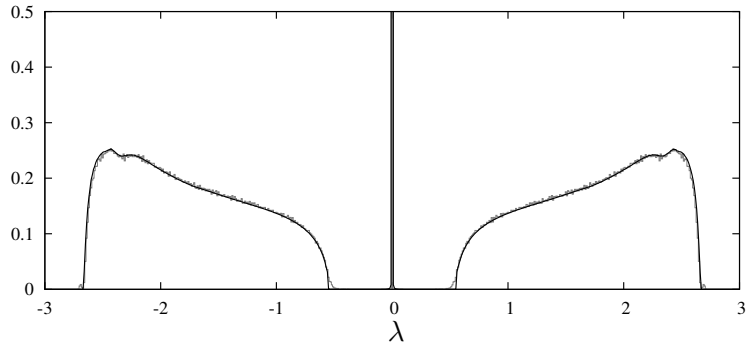
Algorithm 4

1. Begin with three groups of vertices A, B and C of sizes 1200, 400 and 300
2. For each $a \in A$
   i) Randomly choose $b \in B$ and $c \in C$
   ii) If the degree of b is less than 3, and the degree of c is less than 4, connect a to b and a to c, otherwise return to i)

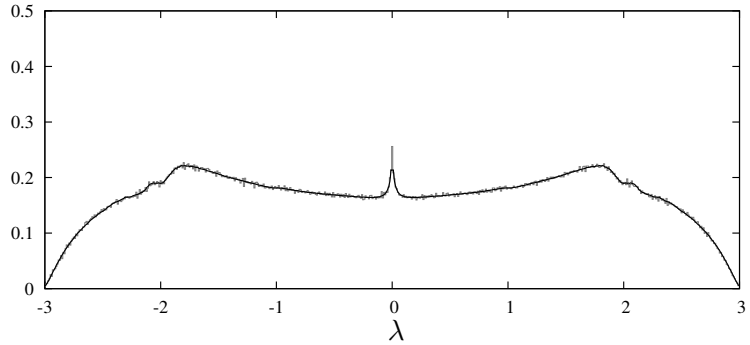
In Figure 3.4, the effects of reducing knowledge of an ensemble (and hence increasing randomness) are clearly visible. In addition to a general smoothing effect, which one might expect, the most striking feature is the appearance of gaps in the spectral density. When the ensemble is fully specified by (3.26), the continuous part of the spectral density is divided into four disjoint components. When one specifies only the degree distribution and degree correlations, the number of components reduces to two, and when only the degree distribution is know, there is no gap in the density at all.



(a) The spectral density for graphs specified by (3.26).



(b) The spectral density for graphs with degree distribution (3.27) and $P(k|d)$ given by (3.27).



(c) The spectral density for graphs with degree distribution (3.27).

Figure 3.4: Comparison theoretical results (black lines) and direct diagonalisation (grey histograms) for the different levels of approximation to the ensemble specified by (3.26). In plot (a), the black line shows the exact result for the spectral density, given by (3.30), in plots (b) and (c) the result of population dynamics is shown. Algorithm 4 was used to generate graphs for plot (a), and Algorithm 2 for (b) and (c). In each case, 1000 graphs of size $N = 1900$ were generated and diagonalised.

This phenomenon can be heuristically explained in terms of periodicity in the random graph ensembles. A walker moving away from a central vertex in the original ensemble will repeatedly visit vertices of degree sequence $\{\ldots, 2, 4, 2, 3, 2, 4, \ldots\}$, which has period four. In the calculation of the spectral density, this fact in manifested in the equations (3.29), moreover, it is known that periodicity in recursive models is directly related to gaps in the spectral density [VM94]. In the case that the degree correlations are specified, whilst the degree sequence of a walk is now random and therefore not strictly periodic, it is still true that every other vertex visited will have degree 2 (see Figure 3.1.3). This repetition with period two appears to be enough to split the spectral density into two components. In the last case, where only the degree distribution is known, the sequence is fully random without any kind of deterministic repetition (see Figure 3.1.3), and the resulting spectral density has no gap.

3.2 Block matrix models

3.2.1 General calculation

All of the random graph ensembles considered thus far share the common feature that cycles of all lengths are asymptotically rare, indeed the analysis has been based upon this very assumption. An extremely simple generalisation (with some trade off in the difficulty of computation) to certain random graph ensembles featuring an extensive number cycles is possible through the use of block matrix models.

Let $G = (V, E)$ be a tree on N vertices. To each vertex $i \in V$ we associate an $n \times n$ Hermitian matrix A_i, and to each edge $(i, j) \in E$ an $n \times n$ matrix B_{ij}. A Hermitian $nN \times nN$ block matrix $\boldsymbol{A}$, composed of N^2 blocks of size $n \times n$, can be constructed by

defining the i^{th}, j^{th} block to be

$$\boldsymbol{A}_{ij} = \begin{cases} A_i & \text{if} \quad i = j \\ B_{ij} & \text{if} \quad (i,j) \in E \quad \text{and} \quad i < j \\ B_{ij}^\dagger & \text{if} \quad (i,j) \in E \quad \text{and} \quad i > j \\ 0 & \text{otherwise.} \end{cases} \tag{3.32}$$

To compute the spectral density of these block matrix models via the cavity method, we will demonstrate an alternative algebraic derivation of the cavity equations that does not make any analogy with statistical mechanics. Writing $\boldsymbol{A}^{(i)}$ for the $n(N-1) \times n(N-1)$ matrix formed from $\boldsymbol{A}$ by the deletion of all blocks with at least one index i, and using $\cong$ for matrices similar by permutation of blocks, we may write

$$\boldsymbol{A} \cong \left(\begin{array}{c|c} \boldsymbol{A}_{ii} & \boldsymbol{A}_{ij_1} \cdots \boldsymbol{A}_{ij_{k_i}} \; 0 \cdots 0 \\ \hline \boldsymbol{A}_{ij_1}^\dagger & \\ \vdots & \\ \boldsymbol{A}_{ij_{k_i}}^\dagger & \boldsymbol{A}^{(i)} \\ 0 & \\ \vdots & \\ 0 & \end{array}\right) \tag{3.33}$$

where $j_1, \ldots, j_{k_i}$ are the neighbours of i in G. Crucially, since G is a tree, $\boldsymbol{A}^{(i)}$ has a particular structure which can be exploited in the evaluation of the resolvent. We may

write

$$\boldsymbol{A}^{(i)} \cong \left(\begin{array}{ccc|ccc} \boldsymbol{A}_{j_1 j_1} & & 0 & \boldsymbol{C}_{j_1} & & 0 \\ & \ddots & & & \ddots & \\ 0 & & \boldsymbol{A}_{j_{k_i} j_{k_i}} & 0 & & \boldsymbol{C}_{j_{k_i}} \\ \hline \boldsymbol{C}_{j_1}^\dagger & & 0 & \boldsymbol{D}_{j_1} & & 0 \\ & \ddots & & & \ddots & \\ 0 & & \boldsymbol{C}_{j_{k_i}}^\dagger & 0 & & \boldsymbol{D}_{j_{k_i}} \end{array}\right) \tag{3.34}$$

$$\cong \begin{pmatrix} \boldsymbol{A}_{j_1 j_1} & \boldsymbol{C}_{j_1} & & & 0 & \\ \boldsymbol{C}_{j_1}^\dagger & \boldsymbol{D}_{j_1} & & & & \\ & & \ddots & & & \\ & & & \boldsymbol{A}_{j_{k_i} j_{k_i}} & \boldsymbol{C}_{j_{k_i}} & \\ & 0 & & \boldsymbol{C}_{j_{k_i}}^\dagger & \boldsymbol{D}_{j_{k_i}} & \end{pmatrix},$$

where the $n \times n$ blocks $\boldsymbol{C}_j$ and $\boldsymbol{D}_j$ are given by the appropriate parts of $\boldsymbol{A}^{(i)}$. Using the above and Schur's formula for the block inverse (given in Appendix B), we arrive at a formula for the i^{th} diagonal block of the resolvent,

$$\left[\left(\boldsymbol{A} - z\right)^{-1}\right]_{ii} = \left(\boldsymbol{A}_{ii} - z - \sum_{j \in \partial i} \boldsymbol{A}_{ij} \left[\left(\boldsymbol{A}^{(i)} - z\right)^{-1}\right]_{jj} \boldsymbol{A}_{ij}^\dagger\right)^{-1}. \tag{3.35}$$

Repeating this calculation for the matrix $\boldsymbol{A}^{(j)}$ yields

$$\left[\left(\boldsymbol{A}^{(j)} - z\right)^{-1}\right]_{ii} = \left(\boldsymbol{A}_{ii} - z - \sum_{l \in \partial i \setminus j} \boldsymbol{A}_{il} \left[\left(\boldsymbol{A}^{(i)(j)} - z\right)^{-1}\right]_{ll} \boldsymbol{A}_{il}^\dagger\right)^{-1},$$

however, given the diagonal block structure of $\boldsymbol{A}^{(i)}$ shown in (3.34), we may conclude $\left[\left(\boldsymbol{A}^{(i)(j)} - z\right)^{-1}\right]_{ll} = \left[\left(\boldsymbol{A}^{(i)} - z\right)^{-1}\right]_{ll}$, and hence

$$\left[\left(\boldsymbol{A}^{(j)} - z\right)^{-1}\right]_{ii} = \left(\boldsymbol{A}_{ii} - z - \sum_{l \in \partial i \setminus j} \boldsymbol{A}_{il} \left[\left(\boldsymbol{A}^{(i)} - z\right)^{-1}\right]_{ll} \boldsymbol{A}_{il}^\dagger\right)^{-1}. \tag{3.36}$$

In the special case $n = 1$, it is immediate to see that the equation (3.36) reduces to the familiar cavity equations (2.13) derived in the previous chapter. For general n, the

spectral density of A can be determined from (3.36) and (3.35) by

$$\varrho_\varepsilon(\lambda; A) = \frac{1}{\pi n N} \operatorname{Im} \sum_{i=1}^{N} \operatorname{Tr} \left[\left(\boldsymbol{A} - \lambda - i\varepsilon \right)^{-1} \right]_{ii} .$$

As in the earlier analysis, if G is a not a tree exactly but simply tree-like, the above expression yields an approximation to the spectral density, whose relation to the true density is the same as in the $n = 1$ case.

3.2.2 Graphs with communities

We can use this block matrix construction to study a simple model of random graphs with community structure. For this model, we take G to be Poissonian random graph with average degree c, whilst the blocks A_i and B_{ij} are taken to be iid random variables with densities $\mu(A)$ and $\nu(B)$, respectively. For ease of calculation, we take ν to satisfy $\nu(B) = \nu(B^T)$. More complicated models, incorporating features such as constrained generalised degrees or communities of varying size are of course possible, however the present simple model is sufficient to study some of the strongest effects of community structure.

The ensemble average for this model is performed easily by applying the calculation discussed in Chapter 1 to the equation (3.36) above, yielding a self-consistency equation for a distribution π defined on $n \times n$ matrices $\boldsymbol{\Delta}$,

$$\begin{aligned} \pi(\boldsymbol{\Delta}) = \sum_k \frac{k\, p(k)}{c} \int \Big[\prod_{\alpha=1}^{k-1} d\boldsymbol{\Delta}_\alpha dA_\alpha dB_\alpha \pi(\boldsymbol{\Delta}_\alpha) \mu(A) \nu(B) \Big] \\ \delta \left(\boldsymbol{\Delta} - \left(A - z - \sum_{\alpha=1}^{k-1} B_\alpha \boldsymbol{\Delta}_\alpha B_\alpha^T \right)^{-1} \right) . \end{aligned} \tag{3.37}$$

The spectral density is then given by

$$\rho(\lambda) = \lim_{\varepsilon \searrow 0} \frac{1}{n\pi} \operatorname{Im} \int \widehat{\pi}(\boldsymbol{\Delta}) \operatorname{Tr} \boldsymbol{\Delta}\, d\boldsymbol{\Delta} ,$$

where

$$\widehat{\pi}(\mathbf{\Delta}) = \sum_k p(k) \int \Big[\prod_{\alpha=1}^{k} d\mathbf{\Delta}_\alpha dA_\alpha dB_\alpha \pi(\mathbf{\Delta}_\alpha) \mu(A) \nu(B) \Big] \\ \delta\left(\mathbf{\Delta} - \left(A - z - \sum_{\alpha=1}^{k} B_\alpha \mathbf{\Delta}_\alpha B_\alpha^T \right)^{-1} \right) . \tag{3.38}$$

For relatively small block size n, the equation (3.37) may once again be solved numerically using a population dynamics algorithm, the only change being to deal with populations of $n \times n$ matrices, rather than single variables.

For a simple example, consider a graph with communities given by the complete graph on n vertices, connected in a Poissonian random graph of average degree c, in which connected communities are joined by a single randomly drawn edge. These choices correspond to $\nu(B) = n^{-2} \sum_{a,b} \delta_{B_{ab},1} \prod_{(c,d)\neq(a,b)} \delta_{B_{cd},0}$, and $\mu(A) = \big(\prod_a \delta_{A_{aa},0} \big) \big(\prod_{a\neq b} \delta_{A_{ab},1} \big)$. Taking $n = 5$ and $c = 5$, population dynamics was used to solve (3.37) for this ensemble, the spectral density was then computed using (3.38). To compare with the results of direct diagonalisation, 1000 graphs of size $N = 5000$ were generated. A histogram of their eigenvalues, alongside the result of population dynamics, is shown in Figure 3.5.

As discussed in the previous chapter, in the limit $c \to \infty$ the spectral density of a Poissonian random graph with average degree c, and edges of weight $1/\sqrt{c}$, converges to a semi-circular distribution. We can compute a generalisation of this result for the community ensemble considered here through a similar treatment of the self-consistency equation (3.37) to that used earlier to derive Wigner's Law. It is necessary to re-weight the edges between communities, in order to keep the spectral density bounded as $c \to \infty$: we put $\nu(B) = n^{-2} \sum_{a,b} \delta_{B_{ab},\sqrt{n/c}} \prod_{(c,d)\neq(a,b)} \delta_{B_{cd},0}$. Expanding (3.37) in $1/c$ for this case, and keeping only the terms relevant in the $c \to \infty$ limit, we obtain an expression for the mean $\mathbb{E}(\mathbf{\Delta})$,

$$\mathbb{E}(\mathbf{\Delta}) = \Big(K_n - \lambda - i\varepsilon - c\,\mathbb{E}\big(B\mathbf{\Delta}B^T\big) \Big)^{-1} ,$$

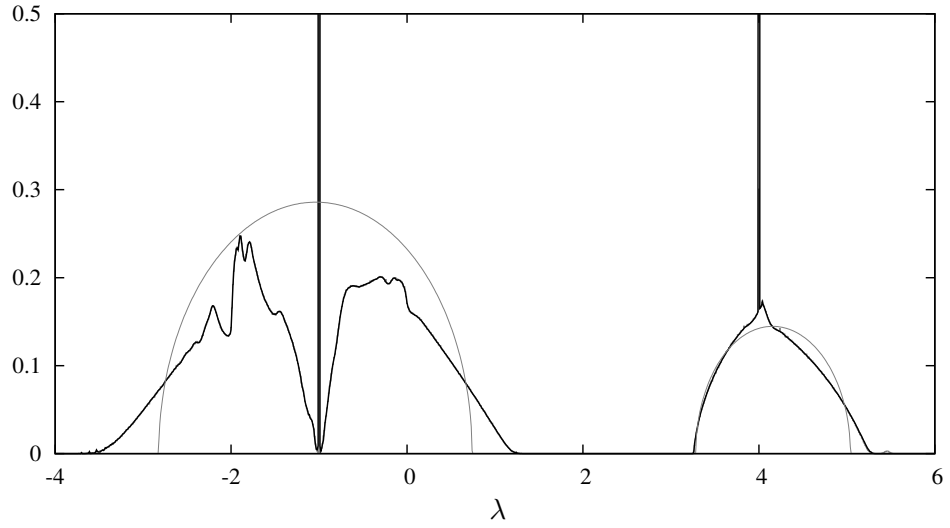


Figure 3.5: Comparison of the results of population dynamics (black line) and direct diagonalisation (grey histogram) for the community structure ensemble described in the text. The grey curve shows the high connectivity limit for this ensemble.

where K_n is the connectivity matrix of the complete graph on n vertices. For the above choice of $\nu(B)$, we have $\mathbb{E}(B\mathbf{\Delta}B^T) = \Delta I_n/cn$, where $\Delta = \operatorname{Tr}\mathbb{E}(\mathbf{\Delta})/n$. Diagonalising K_n, we obtain a cubic equation for Δ,

$$\Delta = \left(\frac{n-1}{n}\right)\frac{1}{-1-\lambda-i\varepsilon-\Delta} + \left(\frac{1}{n}\right)\frac{1}{n-1-\lambda-i\varepsilon-\Delta}\,. \tag{3.39}$$

Taking Δ^+ to be the solution with positive real part, the $c\to\infty$ limit of the spectral density is given by $\rho(\lambda) = \lim_{\varepsilon\searrow 0}\operatorname{Im}\Delta^+/\pi$. In the case $n=5$, we solve (3.39) to give the following expression:

$$\rho(\lambda) = \frac{\sqrt{3}}{2\pi}\left|u(\lambda) - \frac{\lambda^2-3\lambda+18}{9u(\lambda)}\right|\mathbb{I}_D(\lambda)\,,$$

where

$$u(\lambda) = \left|\frac{1}{27}\lambda^3 - \frac{1}{6}\lambda^2 - 2\lambda + 4 + \frac{1}{18}\sqrt{3d(\lambda)}\right|^{1/3},$$

and $d(\lambda) = -25\lambda^4 + 156\lambda^3 + 72\lambda^2 - 1296\lambda + 864$. The domain is given by $D = \{\lambda \,|\, d(\lambda) > 0\}$. A plot of this curve is included in Figure 3.5.

3.2.3 Sparse covariance matrices

In the previous subsections a simple model with an extensive number of 'small' blocks was considered. The reverse arrangement, consisting of a small number of blocks of extensive size, has also been studied in the context of the spectral density of graphs with communities [EK09].

Block matrix models have further applications, including the study of the spectral density of non-Hermitian random matrices (to be explored in later chapters) and the spectral density of sample covariance matrices, which we consider now.

Let A be an $N \times M$ matrix, whose columns are N-dimensional zero-mean sample vectors; the sample covariance matrix is proportional to the $N \times N$ matrix $AA^\dagger$. Eigendecomposition of the sample covariance matrix provides useful information about the distribution of samples, and it is natural to study bulk properties such as the spectral density. For matrices A with independent entries and fixed ratio M/N, the spectral density of $AA^\dagger$ in the limit $N \to \infty$ is given by the Marčenko-Pastur Law [MP67]. Moreover, just as Wigner's Law breaks down with the introduction of sparsity, so to does the Marčenko-Pastur Law for sample covariance matrices.

The case of sparse sample covariance matrices with independent entries has been considered before using the replica method [NT07], with similar approximation schemes to the EMA for symmetric matrices proposed. In the present subsection, we will apply the cavity method to a simple block matrix model, allowing us to compute a close approximation to the spectral density of a sample covariance matrix A described by tree-like graph.

Suppose the locations of the non-zero entries of an $N \times M$ sparse matrix A are encoded in a Boolean matrix C. Whilst in general C will not be the adjacency matrix of a graph,

the $(N+M)\times(N+M)$ block matrix

$$\begin{pmatrix} 0 & C \\ C^T & 0 \end{pmatrix} \tag{3.40}$$

describes a sparse bipartite graph on groups of vertices of size N and M. Taking this idea further, if we define

$$B(z;A) = \begin{pmatrix} -z & A \\ A^\dagger & -I \end{pmatrix},$$

then the Green's function of $AA^\dagger$ is given by the trace of the top left hand block of $B(z;A)$,

$$G(z;AA^\dagger) = \frac{1}{N}\sum_{i=1}^{N} B(z;A)^{-1}_{ii}\,.$$

The cavity method can be applied to $B(z;A)$, using the bipartite graph described by (3.40), to yield equations

$$\Delta_i^{(j)} = \begin{cases} \left(-z - \sum_{l\in\partial i\setminus j} |A_{il}|^2 \Delta_l^{(i)}\right)^{-1} & \text{if} \quad i\in\{1,\dots,N\} \\ \\ \left(-1 - \sum_{l\in\partial i\setminus j} |A_{li}|^2 \Delta_l^{(i)}\right)^{-1} & \text{if} \quad i\in\{N+1,\dots,N+M\}\,. \end{cases} \tag{3.41}$$

Since there are no edges either between vertices $i,j\in\{1,\dots,N\}$ or $i,j\in\{N+1,\dots,N+M\}$, the above can be rewritten in terms of only those $\Delta_i^{(j)}$ with $i\in\{1,\dots,N\}$ and $j\in\{N+1,\dots,N+M\}$,

$$\Delta_i^{(j)} = \left(-z + \sum_{l\in\partial i\setminus j} |A_{il}|^2 \left(1 + \sum_{m\in\partial l\setminus i} |A_{ml}|^2 \Delta_m^{(l)}\right)^{-1}\right)^{-1}. \tag{3.42}$$

An approximation to the spectral density can be recovered from a solution to (3.42) by

$$\varrho_\varepsilon(\lambda;A^\dagger A) = \frac{1}{\pi}\,\mathrm{Im}\sum_{i=1}^{N}\left(-\lambda - i\varepsilon + \sum_{l\in\partial i} |A_{il}|^2 \left(1 + \sum_{m\in\partial l\setminus i} |A_{ml}|^2 \Delta_m^{(l)}\right)^{-1}\right)^{-1}. \tag{3.43}$$

For single instances, the system given by either (3.41) or (3.42) can easily be solved numerically using a belief propagation algorithm like that described in the previous

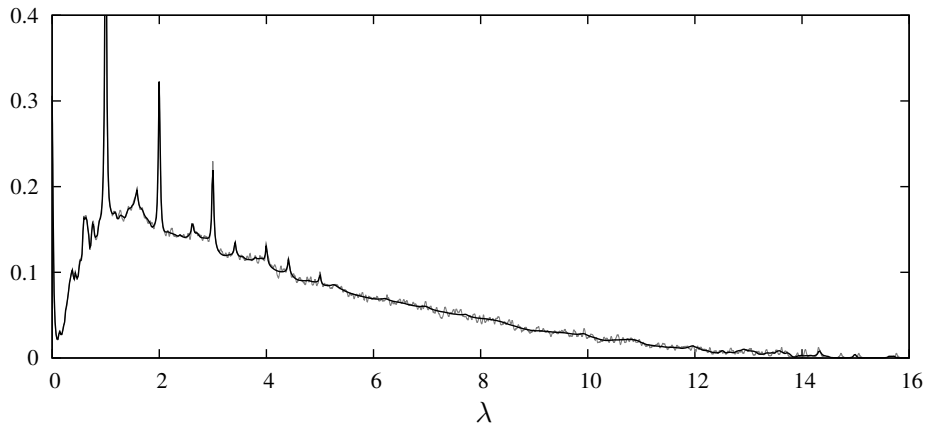


Figure 3.6: Grey: the regularised empirical spectral density of a single sparse sample matrix of Bernoulli random variable of size 3000×10000, with an average row sum of $c = 12$, at $\varepsilon = 0.05$. Black: the result of the cavity equations solved by belief propagation on the same graph.

chapter for general sparse matrices. Figure 3.6 shows the results of a belief propagation algorithm using equations (3.41) applied to a single sparse sample matrix of size 3000×10000. The matrix entries were taken to be Bernoulli random variables with the average number of ones per row given by $c = 12$. A regularising parameter of $\varepsilon = 0.05$ was used. The empirical regularised spectral density for the same graph is provided for comparison.

In the ensemble average, the equation (3.42) can easily be reinterpreted as a distribution recursion relation using the same techniques as discussed in previous sections. The resulting expressions may, as usual, be tackled numerically using population dynamics. Figure 3.7 shows the results of population dynamics, both regularised and unregularised, for Bernoulli sample with an average row sum of $c = 6$.

In 1967 Marčenko and Pastur [MP67] proved a result for covariance matrices analogous to Wigner's Law for matrices with independent entries:

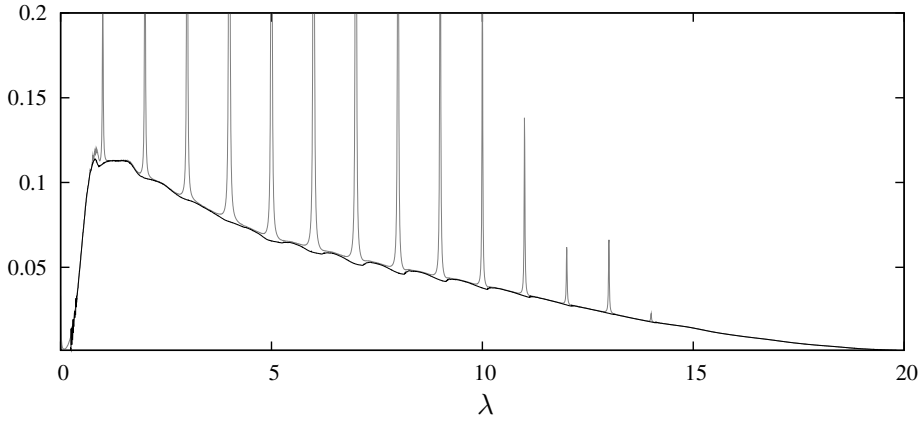


Figure 3.7: Results of population dynamics at $\varepsilon = 0$ (black) and $\varepsilon = 0.05$ (grey) for sparse sample matrices of of Bernoulli random variable with average row sum $c = 6$.

Theorem (The Marčenko-Pastur Law)**.** *Let $\{A_N\}$ be a sequence of $N \times M$ random matrices, where $M/N = \alpha$, a constant. Suppose that for each N the entries of A_N are independent random variables of unit variance, drawn from symmetric distributions with bounded moments. Then*

$$\lim_{N\to\infty} \mathbb{E}\, \varrho\left(\lambda; A_N^\dagger A_N/N\right) = \frac{\sqrt{2(\lambda+\alpha) - (\lambda-\alpha)^2 - 1}}{2\pi\lambda}\, \mathbb{I}_S(\lambda) + (1-\alpha)^+ \delta\left(\lambda\right) , \quad (3.44)$$

where $S = \left((1-\sqrt{\alpha})^2, (1+\sqrt{\alpha})^2\right)$, and $x^+ = \max\{x, 0\}$.

Applying the same techniques to the present situation as were used earlier in deriving Wigner's Law and the high degree limit for the community model, we consider fully connected graphs in which the A_{ij} are independent random variables with variance $1/N$, recovering from (3.42) and (3.43)

$$\rho(\lambda) = \frac{1}{\pi} \operatorname{Im} \Delta\,, \quad \text{where } \Delta \in \mathbb{C}^+ \text{ satisfies } \quad \frac{1}{\Delta} = \frac{\alpha}{1+\Delta} - z\,.$$

Solving the above immediately yields the Marčenko-Pastur law (3.44).

4

Spectral density of sparse non-Hermitian random matrices

4.1 Electrostatic analogy

In this chapter, we extend the work of the previous chapters in determining the spectral density of sparse Hermitian random matrices to the study of sparse non-Hermitian ensembles.

The central object in our work on the spectral density of Hermitian matrices was the Green's function

$$G(z;A) = \frac{1}{N}\mathrm{Tr}\,(A-z)^{-1}\,.$$

Analytic in z away from the real line (to which the eigenvalues of Hermitian matrices are confined), the Green's function provides an N-independent regularisation of the spectral density of the Hermitian matrix A, given by equation (1.7). At points $z = \lambda + i\varepsilon$, for ε fixed and strictly positive, the Green's function can be written in terms of a (convergent) Gaussian integral, which formed the basis for our application of the cavity method. Moreover, the limits $N \to \infty$ and $\varepsilon \to 0$ are freely exchanged, allowing the limiting spectral density to be inferred from the limit of the Green's function

evaluated away from the real line.

For non-Hermitian matrices, this is not the case. Though the Green's function of a non-Hermitian matrix is still analytic away from the spectrum, the likely presence of complex eigenvalues makes invalidates equation (1.7) and forces us to abandon the notion of an imaginary regulariser. To apply the cavity method (or related techniques) with the same success as in the Hermitian case, it will therefore first be necessary to find an alternative formalism with which to tackle the problem.

A number of approaches to the study of non-Hermitian spectral density have been proposed over recent decades, including Girko's various integral transformations [Gir90], Feinberg and Zee's simple 'Hermitization' technique [FZ97a, FZ97b, FSZ01] and the quaternionic Green's function of Janik, Nowak and collaborators [JNPZ97, JNP$^+$97, JN04, JN06, GJ07].

For the purposes of the present chapter, we will concentrate on an analogy with electrostatic potential introduced in 1988 by Sommers *et al* [SCSS88], and used frequently thereafter (for example in [Kho96, SCT96, Ste96, PZ07, GT07a]). The starting point for the approach is the identity

$$\varrho(\lambda; A) = -\frac{1}{\pi}\frac{\partial}{\partial\overline{\lambda}}G(\lambda; A)\,, \tag{4.1}$$

which holds for any matrix A and is easily proved by integration (see Appendix B). Outside of the spectrum of A, we also have

$$\frac{\partial}{\partial\lambda}\log\det\Big[(A-\lambda)^\dagger(A-\lambda)\Big] = -\mathrm{Tr}(A-\lambda)^{-1}\,.$$

The addition of a small real regulariser $\varepsilon > 0$, to keep argument of the logarithm strictly positive, allows the combination these identities, resulting in a regularised form of the spectral density of the non-Hermitian matrix A: we define the function

$$\varrho_\varepsilon(\lambda; A) = \frac{1}{\pi N}\frac{\partial}{\partial\overline{\lambda}}\frac{\partial}{\partial\lambda}\log\det\Big[(A-\lambda)^\dagger(A-\lambda)+\varepsilon^2\Big]\,, \tag{4.2}$$

and then $\varrho(\lambda; A) = \lim_{\varepsilon\searrow 0}\varrho_\varepsilon(\lambda; A)$.

The same notation is used here as was employed earlier for the regularised spectral density of Hermitian matrices, though the expressions are not the same. From this point forward, definition (4.2) is taken to supersede the earlier usage.

In the electrostatic analogy, $\varrho_\varepsilon(\lambda; A)$ is thought of as the charge density corresponding to the potential $\Phi(\lambda) = N^{-1} \log\det\left[(A-\lambda)^\dagger(A-\lambda) + \varepsilon^2\right]$.

To simplify the later analysis, we choose to separate the $(A-\lambda)$ and $(A-\lambda)^\dagger$ terms using the same block matrix trick as was employed for sparse sample covariance matrices in the previous chapter. Introduce the $2N \times 2N$ matrix

$$H_\varepsilon(\lambda; A) = \begin{pmatrix} -i\varepsilon & (A-\lambda) \\ (A-\lambda)^\dagger & -i\varepsilon \end{pmatrix} . \tag{4.3}$$

This is precisely the 'Hermitized' form of $(A-\lambda)$, as introduced by Feinberg and Zee [FZ97a, FZ97b, FSZ01]. Using the formula for the determinant of a block matrix, it is easy to compute

$$\det\left[iH_\varepsilon(\lambda; A)\right] = \det\left[(A-\lambda)^\dagger(A-\lambda) + \varepsilon^2\right] , \tag{4.4}$$

and thus

$$\varrho_\varepsilon(\lambda; A) = \frac{1}{\pi N} \frac{\partial}{\partial\overline{\lambda}} \frac{\partial}{\partial\lambda} \log\det\left[iH_\varepsilon(\lambda; A)\right] . \tag{4.5}$$

Following the same steps as used already for Hermitian matrices, we arrange to rewrite (4.5) in terms of an object resembling the partition function of a disordered system. As always, the use of statistical mechanics terminology in this setting is for illustrative purposes only as this construction does not describe a true physical system. Noting that the eigenvalues of $iH_\varepsilon(\lambda; A)$ all have strictly positive real part, we introduce $2N$ integration variables, organised into N-vectors $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$, so that

$$\frac{1}{\det\left[iH_\varepsilon(\lambda; A)\right]} = \left(\frac{1}{\pi}\right)^{2N} \int \exp\left\{ -i \begin{pmatrix} \boldsymbol{\phi}^\dagger & \boldsymbol{\psi}^\dagger \end{pmatrix} H_\varepsilon(\lambda; A) \begin{pmatrix} \boldsymbol{\phi} \\ \boldsymbol{\psi} \end{pmatrix} \right\} d\boldsymbol{\phi} d\boldsymbol{\psi} .$$

With a modest amount of foresight, we group the integration variables into pairs. For

$i = 1, \ldots, N$, write

$$\chi_i = \begin{pmatrix} \phi_i \\ \psi_i \end{pmatrix} .$$

For the rest of the analysis, we treat these pairs of integration variables as vector valued spins, interacting via the Hamiltonian

$$\mathcal{H}_\varepsilon(\boldsymbol{\chi}, \lambda; A) = \sum_{i=1}^{N} \chi_i^\dagger \begin{pmatrix} \varepsilon & -i\lambda \\ -i\overline{\lambda} & \varepsilon \end{pmatrix} \chi_i + i \sum_{i,j} \chi_i^\dagger \begin{pmatrix} 0 & A_{ij} \\ \overline{A_{ij}} & 0 \end{pmatrix} \chi_j \, . \tag{4.6}$$

The JPDF of spins for this system is given by

$$P(\boldsymbol{\chi}) = \frac{1}{\mathcal{Z}(\lambda; A)} \exp\Big\{ - \mathcal{H}_\varepsilon(\boldsymbol{\chi}, \lambda; A) \Big\} , \tag{4.7}$$

with $\mathcal{Z}(\lambda; A)$ being the partition function

$$\mathcal{Z}(\lambda; A) = \int \exp\Big\{ - \mathcal{H}_\varepsilon(\boldsymbol{\chi}, \lambda; A) \Big\} \, d\boldsymbol{\chi} \, .$$

Applying the first of the partial derivatives appearing in (4.5), we find that the regularised spectral density may be written in terms of a sum of local quantities,

$$\varrho_\varepsilon(\lambda; A) = -\frac{i}{2\pi N} \sum_{i=1}^{N} \frac{\partial}{\partial \overline{\lambda}} \left\langle \chi_i^\dagger (\sigma_x + i\sigma_y) \chi_i \right\rangle , \tag{4.8}$$

where $\langle \cdots \rangle$ denotes the thermal average over the system defined by (4.7), and σ_x and σ_y are standard Pauli matrices (see equation C.2 in Appendix C for a definition).

In this way the problem of determining the spectral density of the non-Hermitian matrix A may be phrased in the language of statistical mechanics, in a form which the cavity method may be applied.

Before moving on to the main calculation, it is worth discussing some of the subtleties surrounding the relationship between the regularised and unregularised spectral densities, which is far from clear at this point.

4.2 The perturbation formula

The regularised spectral density of a non-Hermitian matrix, as defined in equation (4.2), is not so easily exploited to obtain rigorous results for the unregularised density as we might expect from the Hermitian case. Although it is certainly true that for any fixed, finite size matrix A, we have

$$\varrho(\lambda; A) = \lim_{\varepsilon \searrow 0} \varrho_\varepsilon(\lambda; A)\,,$$

there is in general no simple convolution identity such as (1.7)[1]. As a result, even when the regularised spectral density of a random matrix ensemble is accessible to study in the large N limit, obtaining an expression for the unregularised density is not straightforward, with a naive interchange of the limits $N \to \infty$ and $\varepsilon \to 0$ hard to justify.

For a concrete example, consider $N \times N$ Jacobi matrices X_N with entries $[X_N]_{ij} = \delta_{i,j+1}$. For each N, the spectral density of X_N is simply given by a point mass at the origin, however taking first $N \to \infty$ and subsequently $\varepsilon \to 0$, one finds the convergence of $\varrho_\varepsilon(\lambda; X_N)$ to the uniform density on the unit circle.

This problem is not unique to the formalism used here and is shared by many different approaches to non-Hermitian RMT. A considerable amount of effort has gone into attempts to circumvent this difficultly, mostly in connection with the famous problem of proving the Circular Law (for example in [Bai97, GT07a, TV09]). The usual approach taken in justifying the exchange of limits needed is to prove bounds on the least singular values of the matrices involved, though this has the obvious drawback that it must be completed on an ad-hoc basis for each ensemble. Moreover, this method is a stronger tool than is necessary, in the sense that examples exist for which no such bounds are possible and yet the exchange of limits can be observed to be correct.

[1] Although for normal matrices (which include Hermitian matrices) such an expression does exist.

In the present chapter no attempt will be made to explicitly tackle this problem, however, we are able to offer a remarkable relation between the regularised spectral density of a non-Hermitian matrix and the mean spectral density of the same matrix under a particular type of random perturbation[2].

Theorem 2. *Let X be an arbitrary $N \times N$ matrix and ε a strictly positive real number. Suppose A and B are random $N \times N$ matrices, with independent standard complex Gaussian entries, then*

$$\mathbb{E}\, \varrho(\lambda; X + \varepsilon A B^{-1}) = \varrho_\varepsilon(\lambda; X)\,.$$

Proof. For the case $N = 1$, the ratio of two standard complex Gaussian random variables takes the uniform density on the Riemann sphere. One proof of this fact comes from the observation that the density generated is invariant under a class of Möbius transformations. We generalise this idea to matrices. Begin by noting that, for an arbitrary $N \times N$ matrix X,

$$\varrho_\varepsilon(\lambda; X) = -\frac{1}{N}\mathrm{Tr}\frac{\partial}{\partial\overline{\lambda}}\Big((X-\lambda)^\dagger(X-\lambda) + \varepsilon^2\Big)^{-1}(X-\lambda)^\dagger\,,$$

and

$$\varrho(\lambda; X + \varepsilon A B^{-1}) = -\frac{1}{N}\mathrm{Tr}\frac{\partial}{\partial\overline{\lambda}}\big(\varepsilon A B^{-1} + X - \lambda\big)^{-1}\,.$$

Theorem 2 will then follow from the stronger claim that for any matrix Y

$$\mathbb{E}\,\big(\varepsilon A B^{-1} + Y\big)^{-1} = \big(Y^\dagger Y + \varepsilon^2\big)^{-1} Y^\dagger\,.$$

A little rearrangement leads to the equivalent statement $\mathbb{E}\, f(AB^{-1}) = 0$, where we have introduced the matrix Möbius transformation f, given by

$$f(Z) = \big(Y^\dagger Z - \varepsilon\big)\big(\varepsilon Z + Y\big)^{-1}\,.$$

[2]In fact, this result was found independently by Haagerup nearly a decade ago; it appears in handwritten notes distributed at MSRI in 2001, though the proof is substantially different.

Notice that if $\begin{pmatrix} A \\ B \end{pmatrix} \mapsto AB^{-1}$, then $F \begin{pmatrix} A \\ B \end{pmatrix} \mapsto f(AB^{-1})$, where F is the $2N \times 2N$ block matrix

$$F = \begin{pmatrix} Y^\dagger & -\varepsilon \\ \varepsilon & Y \end{pmatrix}.$$

It will be useful to 'normalise' f, introducing

$$\tilde{f}(Z) = \left(Y^\dagger Y + \varepsilon^2\right)^{-1/2} \left(Y^\dagger Z - \varepsilon\right) \left(\varepsilon Z + Y\right)^{-1} \left(YY^\dagger + \varepsilon^2\right)^{1/2}.$$

Then the normalised form of F is given by $\tilde{F} \in \mathrm{SU}(2N)$,

$$\tilde{F} = \begin{pmatrix} Y^\dagger Y + \varepsilon^2 & 0 \\ 0 & YY^\dagger + \varepsilon^2 \end{pmatrix}^{-1/2} \begin{pmatrix} Y^\dagger & -\varepsilon \\ \varepsilon & Y \end{pmatrix}.$$

Now, let A and B be independent complex Gaussian matrices, with JPDF

$$p \begin{pmatrix} A \\ B \end{pmatrix} = \frac{1}{\pi^{2N^2}} \exp\left\{ -\mathrm{Tr} \left[\begin{pmatrix} A^\dagger & B^\dagger \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} \right] \right\},$$

and let $\begin{pmatrix} A' \\ B' \end{pmatrix} = \tilde{F} \begin{pmatrix} A \\ B \end{pmatrix}$. We perform a change of variables to find that the JPDF of A' and B' is given by

$$\frac{1}{\pi^{2N^2}} \exp\left\{ -\mathrm{Tr} \left[\begin{pmatrix} A^\dagger & B^\dagger \end{pmatrix} \left(\tilde{F}^\dagger \tilde{F}\right)^{-1} \begin{pmatrix} A \\ B \end{pmatrix} \right] \right\} \left| \det \tilde{F}^{-1} \right| = p \begin{pmatrix} A \\ B \end{pmatrix}.$$

Thus the density p is invariant under multiplication with $\tilde{F}$ and we can conclude that the distribution of the random variable AB^{-1} is invariant under $\tilde{f}$.

So to prove that $\mathbb{E} f(AB^{-1}) = 0$, it will suffice to show that $\mathbb{E} AB^{-1} = 0$. But A and B are independent, $\mathbb{E} A = 0$, and it can be shown (see [Ede88]) that $\mathbb{E}\|B^{-1}\| = \sqrt{N\pi} < \infty$, so we are done. □

Theorem 2 is the non-Hermitian analogue of the identity (1.7). In that expression, the regularised spectral density is related to the mean spectral density under perturbation of λ, whilst in the non-Hermitian case we must perturb the whole matrix. This

result also offers a strong heuristic argument in favour of the interchange of limits $N \to \infty$ and $\varepsilon \to 0$ for random matrices: if we are dealing with very large and fully random matrices, the addition of an infinitesimal random perturbation should not change the spectral density in a discontinuous fashion. It is reasonable to hope that this argument could be made rigorous, perhaps allowing for simpler and more widely applicable techniques than the traditional analysis of least singular values, though this possibility is yet to be fully explored.

4.3 The cavity method

Having introduced the electrostatic analogy and discussed some of the subtleties and limitations inherent in the study of non-Hermitian random matrices, we move on now to the calculation of the spectral density of sparse non-Hermitian random matrices via the cavity method.

Let C be the connectivity matrix of a tree-like graph $G = (V, E)$, and let J be an arbitrary non-Hermitian matrix. Define the matrix A by setting

$$A_{ij} = C_{ij} J_{ij} \, .$$

Although C is real and symmetric, the construction above also includes directed graphs by simply setting, for example, $J_{ij} = 1$ for a directed edge from i to j and $J_{ij} = 0$ otherwise. For simplicity, we study matrices with no diagonal terms, though these are very easily added to the theory if necessary. It will be useful to decompose J into Hermitian and anti-Hermitian parts; we write

$$J^x = \frac{1}{2}(J + J^\dagger) \, , \quad \text{and} \quad J^y = \frac{1}{2i}(J - J^\dagger) \, ,$$

so that $J = J^x + iJ^y$.

We will apply the cavity method to the formalism derived earlier, expressing the regularised spectral density of A in terms of the thermal average over the system defined

by (4.7). The first step in our analysis is to observe that the Hamiltonian defined in (4.6) may be split into contributions coming from single spins and pairs of linked spins,

$$\mathcal{H}_\varepsilon(\boldsymbol{\chi}, \lambda; A) = \sum_{i=1}^{N} \mathcal{H}_i(\chi_i) + \sum_{i<j} C_{ij} \mathcal{H}_{ij}(\chi_i, \chi_j)\,,$$

where

$$\mathcal{H}_i(\chi_i) = \chi_i^\dagger \big(\varepsilon I_2 - ix\sigma_x + iy\sigma_y\big)\chi_i\,,$$

and

$$\mathcal{H}_{ij}(\chi_i, \chi_j) = i\chi_i^\dagger \big(J_{ij}^x \sigma_x - J_{ij}^y \sigma_y\big)\chi_j + i\chi_j^\dagger \big(J_{ji}^x \sigma_x - J_{ji}^y \sigma_y\big)\chi_i\,.$$

Following the same steps as in earlier applications of the cavity method, we find for tree-like graphs that the marginal at vertex i is recovered from the cavity marginals of the neighbours by

$$P_i(\chi_i) = \frac{1}{Z_i} e^{-\mathcal{H}_i(\chi_i)} \int e^{-\sum_{l\in\partial i} \mathcal{H}_{il}(\chi_i,\chi_l)} \prod_{l\in\partial i} P_l^{(i)}(\chi_l)\, d\chi_l\,,$$

and the cavity marginals are themselves related by the self-consistency equation

$$P_i^{(j)}(\chi_i) = \frac{1}{Z_i^{(j)}} e^{-\mathcal{H}_i(\chi_i)} \int e^{-\sum_{l\in\partial i\setminus j} \mathcal{H}_{il}(\chi_i,\chi_l)} \prod_{l\in\partial i\setminus j} P_l^{(i)}(\chi_l)\, d\chi_l\,, \tag{4.9}$$

where Z_i and $Z_i^{(j)}$ are normalising constants.

Once again, further progress is possible by taking a parameter dependent representation for the marginals. For each $i \in V$ and $l \in \partial i$, the cavity marginal at l in the absence of i may be parameterised in terms of a 2×2 matrix $\boldsymbol{\Delta}_l^{(i)}$, describing the covariance the components of the spin χ_l:

$$P_l^{(i)}(\chi_l) = \frac{1}{\pi^2 \det\big(i\boldsymbol{\Delta}_l^{(i)}\big)} \exp\Big(-i\chi_l^\dagger \big(\boldsymbol{\Delta}_l^{(i)}\big)^{-1}\chi_l\Big)\,, \tag{4.10}$$

The true marginal at i may be written is the same way in terms of a matrix $\boldsymbol{\Delta}_i$ as

$$P_i(\chi_i) = \frac{1}{\pi^2 \det\big(i\boldsymbol{\Delta}_i\big)} \exp\Big(-i\chi_l^\dagger \boldsymbol{\Delta}_i^{-1}\chi_l\Big)\,. \tag{4.11}$$

Examination of the matrix $H_\varepsilon(\lambda; A)$ shows that the cavity variables $\{\mathbf{\Delta}_i^{(j)}\}$ may be conveniently expressed in terms of Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ and four real variables $s_i^{(j)}, t_i^{(j)}, u_i^{(j)}, v_i^{(j)} \in \mathbb{R}$:

$$\mathbf{\Delta}_i^{(j)} = is_i^{(j)} I_2 + it_i^{(j)}\sigma_z + u_i^{(j)}\sigma_x + v_i^{(j)}\sigma_y \,,$$

with the additional constraint that $s_i^{(j)} > t_i^{(j)} > 0$. Similarly, the true covariance matrices $\{\mathbf{\Delta}_i\}$ are written in the same way in terms of variables $s_i, t_i, u_i, v_i \in \mathbb{R}$. If the marginal distribution of the spin χ_i given by (4.11), and $\mathbf{\Delta}_i$ is parameterised as mentioned, we find that the covariance of the components ϕ_i and ψ_i has the simple form

$$\langle \overline{\phi}_i \psi_i \rangle = u_i + iv_i \,.$$

Equation (4.8) for the spectral density thus becomes

$$\varrho_\varepsilon(\lambda; A) = -\frac{1}{\pi N} \sum_{i=1}^{N} \frac{\partial}{\partial \overline{\lambda}} (u_i + iv_i) \,. \tag{4.12}$$

Using the parameterisation (4.10), we have for each $i \in V$ and $l \in \partial i$,

$$\int e^{-\mathcal{H}_{il}(\chi_i, \chi_l)} P_l^{(i)}(\chi_l)\, d\chi_l = \exp\left(i\chi_i^\dagger \big(J_{il}^x \sigma_x - J_{il}^y \sigma_y\big) \mathbf{\Delta}_l^{(i)} \big(J_{li}^x \sigma_x - J_{li}^y \sigma_y\big) \chi_i \right) ,$$

and thus the cavity equations (4.9) reduce to the following relation between matrices:

$$\begin{aligned} \mathbf{\Delta}_i^{(j)} = \Bigg[& -i\varepsilon I_2 - x\sigma_x + y\sigma_y \\ & - \sum_{l \in \partial i \setminus j} \big(J_{il}^x \sigma_x - J_{il}^y \sigma_y\big) \mathbf{\Delta}_l^{(i)} \big(J_{li}^x \sigma_x - J_{li}^y \sigma_y\big) \Bigg]^{-1} . \end{aligned} \tag{4.13}$$

Similarly, the true marginals are given by

$$\begin{aligned} \mathbf{\Delta}_i = \Bigg[& -i\varepsilon I_2 - x\sigma_x + y\sigma_y \\ & - \sum_{l \in \partial i} \big(J_{il}^x \sigma_x - J_{il}^y \sigma_y\big) \mathbf{\Delta}_l^{(i)} \big(J_{li}^x \sigma_x - J_{li}^y \sigma_y\big) \Bigg]^{-1} . \end{aligned} \tag{4.14}$$

Equations (4.13) and (4.14) comprise the main result of this chapter. For a given matrix associated to a tree-like graph, solution of (4.13) yields, though (4.14) and (4.12), a close approximation to the regularised spectral density.

These equations are noticeably similar to those derived in both the original case of sparse Hermitian matrices, and, even more closely, to those of the block matrix models discussed in the previous chapter. In fact, it is possible to simply rearrange $H_\varepsilon(\lambda; A)$ into a convenient block form, from which equation (3.36) will yield (4.13) directly.

4.4 Analytically solvable cases

4.4.1 The fully connected limit - Girko's Elliptic Law

The Circular Law given in the first chapter states that the spectral density of random matrices with independent entries converges, after suitable normalisation, to the uniform density on the unit disc. A generalisation of this result to ensembles with independent Hermitian and anti-Hermitian parts appears in [SCSS88, Gir90], and is sometimes referred to as Girko's Elliptic Law. Just as Wigner's Law was found in the fully connected limit of the cavity equations for Hermitian matrices, the Circular Law and Girko's Elliptic Law may be recovered from equations (4.13) using the same techniques.

Let G be a complete graph and suppose that the entries of J^x and J^y are random variables satisfying, for each i and j

$$\mathbb{E}\left(|J^x_{ij}|^2\right) = 1/N\,, \quad \mathbb{E}\left(|J^y_{ij}|^2\right) = 1/N\,, \quad \text{and} \quad \mathbb{E}\left(J^x_{ij} J^y_{ij}\right) = \tau\,.$$

The parameter $\tau \in [0, 1]$ controls the degree to which the resulting matrix is Hermitian; at $\tau = 1$ it is completely Hermitian and obeys Wigner's Law, at $\tau = 0$ all correlations vanish and the Circular Law holds.

Following the same steps as in previous calculations of the full connected limit, the above correlations transform (4.13) to give an equation for the mean single-spin co-

variance matrix $\mathbf{\Delta}$:

$$\mathbf{\Delta}^{-1} = -x\sigma_x + y\sigma_y - \frac{1}{2}(1+\tau)\sigma_x\mathbf{\Delta}\sigma_x - \frac{1}{2}(1-\tau)\sigma_y\mathbf{\Delta}\sigma_y\,. \tag{4.15}$$

Note that we have removed the regularising parameter, setting $\varepsilon = 0$. As mentioned in the discussion earlier in this chapter, it is not straightforward to provide a rigorous justification for this step. Solving (4.15), we find two regimes:

$$\mathbf{\Delta} = \sqrt{1 - \left(\frac{x}{1+\tau}\right)^2 - \left(\frac{y}{1-\tau}\right)^2}\, I_2 - \frac{x}{1+\tau}\sigma_x + \frac{y}{1-\tau}\sigma_y\,,$$

inside the ellipse defined by $\left(\frac{x}{1+\tau}\right)^2 + \left(\frac{y}{1-\tau}\right)^2 \leq 1$, and

$$\mathbf{\Delta} = -\beta(x\sigma_x - y\sigma_y)\,,$$

outside, where β satisfies $\tau(x^2+y^2)\beta^2 - (x^2+y^2)\beta + 1 = 0$. From this and equation (4.12), the spectral density in this case is found to be the uniform density on the aforementioned ellipse,

$$\rho(\lambda) = \begin{cases} \frac{1}{\pi(1-\tau^2)} & \text{when} \quad \left(\frac{x}{\tau+1}\right)^2 + \left(\frac{y}{\tau-1}\right)^2 < 1 \\ 0 & \text{otherwise.} \end{cases}$$

This is precisely the Girko's Elliptic Law. The Circular Law occurs as a special case found by setting $\tau = 0$.

4.4.2 Directed random regular graphs

One of the first applications of the cavity equations for Hermitian matrices given in Chapter 2 was to derive McKay's Law for the spectral density of random regular graphs. There is a nice generalisation of this result to directed graphs, which seemingly has not yet appeared in print[3].

[3]Though it was found independently and by the same method by Bordenave.

Suppose we choose uniformly randomly from the collection of all directed graphs on N vertices in which each vertex has both in $(+)$ and out $(-)$ degree k. Assuming as we did in the case of McKay's Law that the resulting graphs are tree-like and translationally invariant in the large N limit, we write $\mathbf{\Delta}_i \equiv \mathbf{\Delta}$ for all i and $\mathbf{\Delta}_i^{(j)} \equiv \mathbf{\Delta}^{(+)}$ if there is an edge from i to j, $\mathbf{\Delta}_i^{(j)} \equiv \mathbf{\Delta}^{(-)}$ in the case of an edge from j to i. From (4.14) we deduce

$$\mathbf{\Delta} = \Bigg[-x\sigma_x + y\sigma_y - k\left(\frac{1}{2}\sigma_x - \frac{1}{2i}\sigma_y\right)\mathbf{\Delta}^{(\pm)}\left(\frac{1}{2}\sigma_x + \frac{1}{2i}\sigma_y\right) - k\left(\frac{1}{2}\sigma_x + \frac{1}{2i}\sigma_y\right)\mathbf{\Delta}^{(\mp)}\left(\frac{1}{2}\sigma_x - \frac{1}{2i}\sigma_y\right)\Bigg]^{-1}$$

where the cavity equations (4.13) provide

$$\mathbf{\Delta}^{(\pm)} = \Bigg[-x\sigma_x + y\sigma_y - k\left(\frac{1}{2}\sigma_x - \frac{1}{2i}\sigma_y\right)\mathbf{\Delta}^{(\pm)}\left(\frac{1}{2}\sigma_x + \frac{1}{2i}\sigma_y\right) - (k-1)\left(\frac{1}{2}\sigma_x + \frac{1}{2i}\sigma_y\right)\mathbf{\Delta}^{(\mp)}\left(\frac{1}{2}\sigma_x - \frac{1}{2i}\sigma_y\right)\Bigg]^{-1}$$

and

$$\mathbf{\Delta}^{(\mp)} = \Bigg[-x\sigma_x + y\sigma_y - (k-1)\left(\frac{1}{2}\sigma_x - \frac{1}{2i}\sigma_y\right)\mathbf{\Delta}^{(\pm)}\left(\frac{1}{2}\sigma_x + \frac{1}{2i}\sigma_y\right) - k\left(\frac{1}{2}\sigma_x + \frac{1}{2i}\sigma_y\right)\mathbf{\Delta}^{(\mp)}\left(\frac{1}{2}\sigma_x - \frac{1}{2i}\sigma_y\right)\Bigg]^{-1}$$

Solving, we find

$$\mathbf{\Delta}^{(\pm)} = i\frac{2k-1}{2k}\sqrt{\frac{k-|\lambda|^2}{k(k-1)}}I_2 \pm i\frac{1}{k}\sqrt{\frac{k-|\lambda|^2}{k(k-1)}}\sigma_z - \frac{x}{k}\sigma_x - \frac{y}{k}\sigma_y\,.$$

in the region $|\lambda|^2 < k^2$ and $\mathbf{\Delta}^{(\pm)} = -(x\sigma_x + y\sigma_y)/(x^2+y^2)$ outside. In turn this gives

$$\mathbf{\Delta} = \begin{cases} i\dfrac{\sqrt{k(k-1)(k-|\lambda|^2)}}{k^2-\lambda^2}I_2 - \dfrac{(k-1)}{k^2-|\lambda|^2}\big(x\sigma_x + y\sigma_y\big) & \text{if} \quad |\lambda|^2 < k^2 \\ -\dfrac{1}{x^2+y^2}\big(x\sigma_x + y\sigma_y\big) & \text{otherwise.} \end{cases}$$

And thus we find that the asymptotic mean spectral density of k-in k-out random directed graphs is given by the simple formula

$$\rho(\lambda) = \begin{cases} \dfrac{k-1}{\pi}\left(\dfrac{k}{k^2-|\lambda|^2}\right)^2 & \text{if} \quad |\lambda|^2 < k^2 \\ 0 & \text{otherwise.} \end{cases}$$

4.5 Numerical simulations

To obtain an estimate for the spectral density of a given matrix from the solution of the cavity equations (4.13), it is necessary to perform the anti-holomorphic derivative appearing in equation (4.12). This is simple if, as in the examples above, one has succeeded in solving the cavity equations algebraically, however for numerical applications further work will be necessary to avoid having to take a numerical approximation to the derivative.

For each $i \in V$ and $j \in \partial i$, introduce new variables $\mathbf{\Gamma}_i$ and $\mathbf{\Gamma}_i^{(j)}$ for the value of the derivatives $\partial \mathbf{\Delta}_i / \partial \overline{\lambda}$ and $\partial \mathbf{\Delta}_i^{(j)} / \partial \overline{\lambda}$ at the point $\lambda = x + iy$. Applying the derivative at the level of equations (4.13) and (4.14), we find a set of consistency equations:

$$\mathbf{\Gamma}_i^{(j)} = \mathbf{\Delta}_i^{(j)} \Big[-\frac{1}{2}(\sigma_x - i\sigma_y) - \sum_{l \in \partial i \setminus j} \big(J_{il}^x \sigma_x - J_{il}^y \sigma_y\big) \mathbf{\Gamma}_i^{(l)} \big(J_{li}^x \sigma_x - J_{li}^y \sigma_y\big) \Big] \mathbf{\Delta}_i^{(j)} \,, \quad (4.16)$$

from which the derivative of the true $\{\mathbf{\Delta}_i\}$ is given by

$$\mathbf{\Gamma}_i = \mathbf{\Delta}_i \Big[-\frac{1}{2}(\sigma_x - i\sigma_y) - \sum_{l \in \partial i} \big(J_{il}^x \sigma_x - J_{il}^y \sigma_y\big) \mathbf{\Gamma}_i^{(l)} \big(J_{li}^x \sigma_x - J_{li}^y \sigma_y\big) \Big] \mathbf{\Delta}_i \,. \quad (4.17)$$

For a given matrix, equations (4.16) may be solved simultaneously with the cavity equations (4.13) numerically by belief propagation. Finally, rewriting (4.12), we find that the spectral density may then be recovered from the result of the equations (4.17) by summing over the lower-left entries of the $\mathbf{\Gamma}_i$:

$$\varrho_\varepsilon(\lambda; A) = -\frac{1}{\pi N} \sum_{i=1}^{N} \big[\mathbf{\Gamma}_i\big]_{2,1} \,. \quad (4.18)$$

The results from belief propagation and comparison with those of numerical diagonalisation are presented for two cases: in Figure 4.1 symmetrically connected Poissonian random graphs with average connectivity c and with asymmetric Gaussian edge weights with zero mean and variance $1/c$ and in Figure 4.2 asymmetrically connected Poissonian graphs with edge weights drawn uniformly from the circle of radius $1/\sqrt{c}$. For both cases the ensemble average has been estimated by averaging over a sample of results from 1000 random matrices for belief propagation and 10^5 for numerical diagonalisation[4].

Notice that the ensembles in both cases satisfy the conditions for the Circular Law in the limit $c \to \infty$. However, it is evident from the figures that, for finite c, they have spectral densities dramatically different both from each other and from the limiting case of the Circular Law. Apart from small discrepancies near the boundaries due to the discretisation the histogram introduces, the comparison shows excellent agreement.

It is of course possible to perform the ensemble average at an analytic level as seen earlier for Hermitian matrices. The end product will be a set of equations similar to, for example, (2.25) and (2.26), allowing for numerical solution by population dynamics. The major difference in moving from Hermitian to non-Hermitian is simply that distributions of both cavity variables $\mathbf{\Delta}$, and their derivatives $\mathbf{\Gamma} = \partial\mathbf{\Delta}/\partial\overline{\lambda}$ must be considered simultaneously.

[4] A larger sample size is needed for the diagonalisation results than for the data coming from belief propagation in order to construct a smooth 3D histogram.

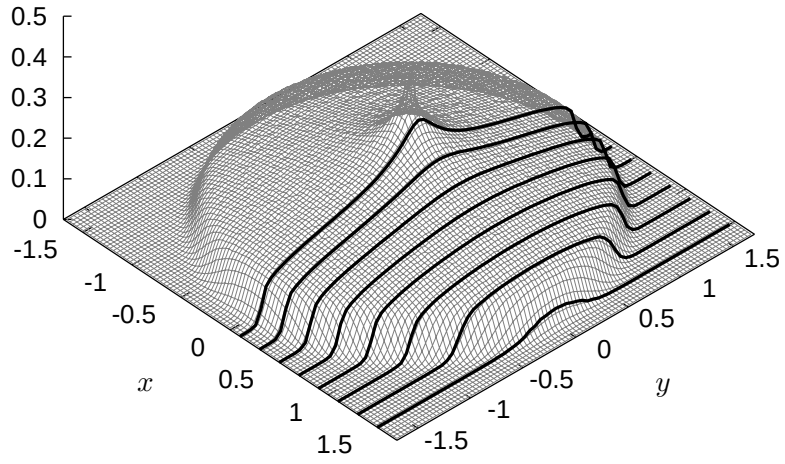


Figure 4.1: Spectral density of symmetric Poissonian graphs with asymmetric Gaussian edge weights and average connectivity $c = 10$. The grey grid is a histogram of the eigenvalues of 10^5 samples, the black lines are the result of the cavity equations, averaged over 1000 samples.

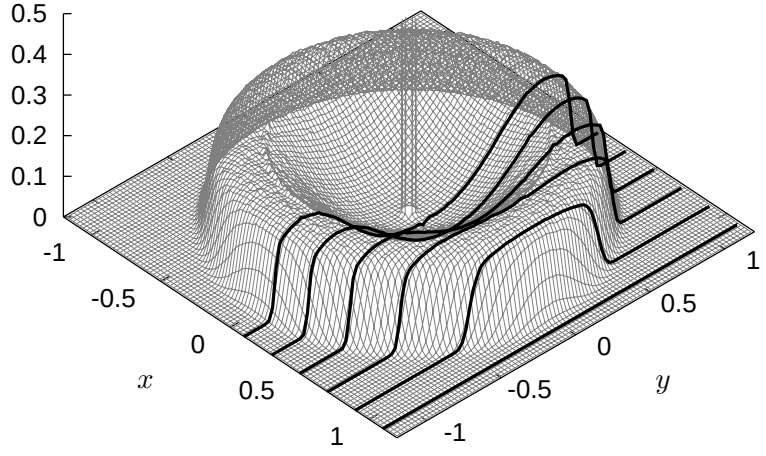


Figure 4.2: Spectral density of asymmetric Poissonian graphs with unitary edge weights and average connectivity $c = 2$. The grey grid is a histogram of the eigenvalues of 10^5 samples, the black lines are the result of the cavity equations, averaged over 1000 samples.

5

Universal sum and product rules for random matrices

5.1 Universality of spectral density

The concept of universality is of central importance to research in RMT. It is an interesting feature of the theory that many key emergent properties of random matrix ensembles are universal in the sense that they are determined entirely by the broader characteristics of the ensemble and do not depend on the exact details of the JPDF of entries. For example, in the study of level spacing (an important topic in RMT not considered in this thesis) it has been found that random matrices of the same symmetry class will exhibit the same level spacing distribution, regardless of the exact details of the ensemble.

In the study of spectral density, there are three main universality results; Wigner's Law, the Marčenko-Pastur Law and the Circular Law, each describing the limiting spectral density for certain random matrix ensembles with independent entries, under rather broad conditions. The focus of the thesis thus far has been to obtain results on the spectral density of sparse random matrix ensembles, which usually break away

from the universality classes of each of these results. The figures presented at the end of the last chapter are a perfect example, showing results from a pair of ensembles which will conform to the Circular Law in the fully connected limit, yet have radically different spectral densities in the sparse regime.

Though the cavity equations found in the previous chapters were derived for sparse random matrices associated to tree-like graphs, we have seen that all three of the universality theorems mentioned above may be easily deduced from the cavity equations when analysed in the fully connected limit. Interestingly, in the case of both Wigner's Law and the Marčenko-Pastur Law, the cavity equations transform in the fully connected limit into a single quadratic equation for the Green's function; precisely the same equation as is derived when applying Pastur's so-called simple approach to RMT discussed in the introduction (for example, compare equation (1.9) from the simple approach to Wigner's Law with equation (2.16) coming from the cavity method). For the derivation of Girko's Elliptic Law in the last chapter, a very similar equation was found (4.15), though in the random matrix literature there does not appear to be a corresponding version of the simple approach with which to study non-Hermitian matrices.

For this final part of the thesis, we will take a different direction to the work of previous chapters, seeking to develop a non-Hermitian analogue of the simple approach, with which universal results about the spectral density of non-Hermitian random matrices may be obtained. In Hermitian RMT, the techniques of the simple approach have utility beyond providing straightforward proofs of the Wigner and Marčenko-Pastur Laws, in particular they have been used to determine the limiting spectral density of Hermitian random matrices under deterministic additive perturbation, captured by the Pastur equation [Pas72]. In what follows, techniques will be developed enabling the establishment universal results for the spectral density of random matrices with independent entries, whether Hermitian or not, when summed or multiplied with deterministic matrices.

5.2 The Spherical Law

Before beginning the development of a simple approach to non-Hermitian RMT, we pause a moment to conjecture a new universal law for the spectral density of a certain class of random matrix.

As a corollary of Theorem 2 we see that if A and B are random matrices of independent complex Gaussian entries, the mean spectral density of the matrices AB^{-1} is, regardless of their size, the uniform distribution on the Riemann Sphere[1]. In light of the known universality results for the spectral density of random matrices, we are immediately prompted to ask if this spherical density occurs in the limit $N \to \infty$ for any distribution of the entries of A_N and B_N, in an analogue of the Circular Law:

Conjecture (The Spherical Law)**.** *Let $\{A_N\}$ and $\{B_N\}$ be sequences of matrices of independent complex random variables of zero mean and unit variance. Then the spectral densities of the matrices $A_N B_N^{-1}$ converge to the uniform density on the Riemann sphere.*

Figure 5.2 shows the eigenvalues of a single random matrix AB^{-1} of size $N = 10^4$, where A and B were randomly filled with ones and zeros. The eigenvalues have been projected onto the Riemann sphere, where they exhibit a highly uniform distribution, supporting the claim of the Spherical Law conjecture.

This phenomenon was also noticed by Forrester and Mays [FM09] and can be easily derived in a non-rigorous fashion using the techniques developed in this chapter, however a full proof is likely to require a more in-depth analysis.

[1] In fact, the full JPDF of eigenvalues for matrices of this type was found recently in [Kri09].

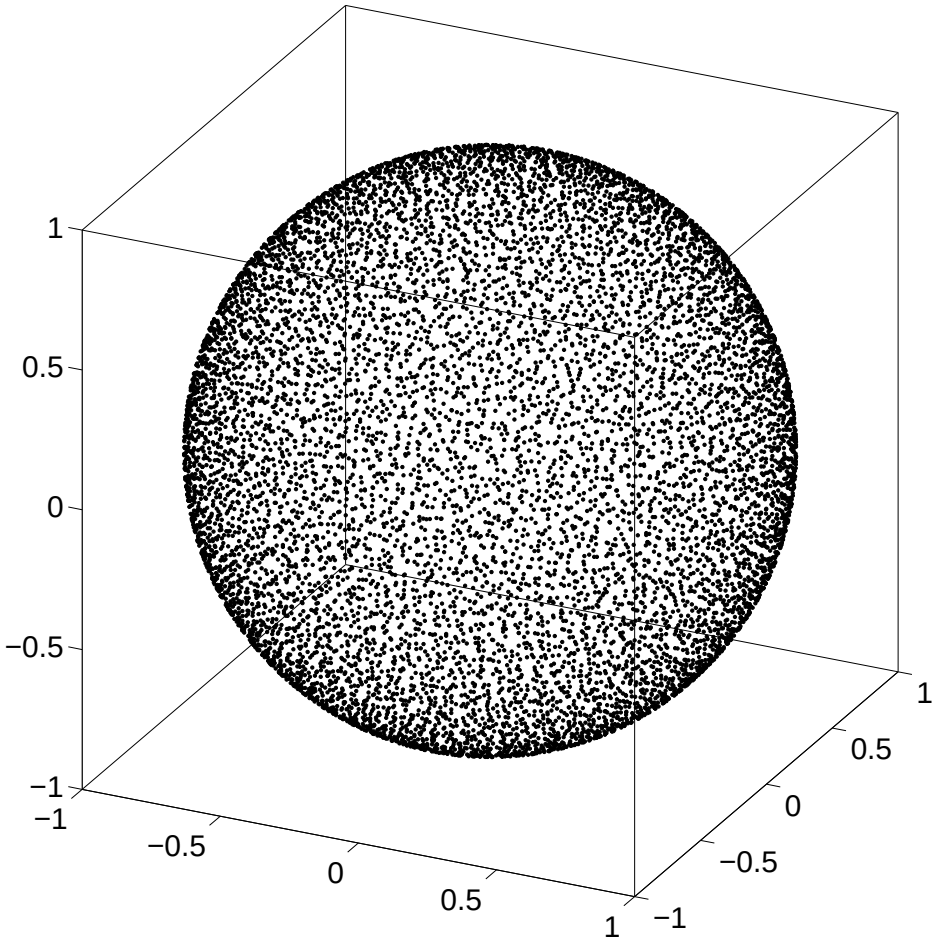


Figure 5.1: Eigenvalues of a single $10^4 \times 10^4$ matrix of the form AB^{-1}, where A and B were randomly filled with ones and zeros. The eigenvalues are mapped onto the Riemann sphere by inverse stereographic projection.

5.3 The quaternionic Green's function

One possible reason for the lack of techniques in non-Hermitian RMT analogous to the simple approach for Hermitian RMT is that this method is based on analysis of the Green's function; an object which we observe to have a radically different role in non-Hermitian RMT, as discussed at the start of the previous chapter. In tackling this problem, we will look to the work of Janik, Nowak and collaborators on a quaternionic generalisation of the Green's function (formally equivalent to Feinberg and Zee's 'Hermitization' trick), which will turn out to replicate in the non-Hermitian setting many of the useful properties of Green's functions of Hermitian matrices.

Introduced in [JNPZ97, JNP$^+$97], the quaternionic Green's function offers a convenient formalism with which to study the limiting spectral density of non-Hermitian random matrices. In these and subsequent works [JN04, JN06], the application of free probability theory to this quaternionic formalism has yielded many interesting results, including for sums of unitary random matrices [GJ07] and infinite products of large random matrices [GNJJN03], though the question of universality has not been considered in any detail.

The idea underlying the notion of a quaternionic Green's function is straightforward: since the (real) spectral density of Hermitian matrices may be studied through analysis of the Green's function in the complex plane, one might hope that the (complex) spectral densities of non-Hermitian matrices could be studied through a Green's function acting on a higher dimensional space such as the quaternions. In practice, this elegant concept is realised through block matrix techniques with a very similar feel to those of Feinberg and Zee.

5.3.1 Notation

We will be doubling the size of our matrices. To simplify the formulas, we introduce the following notation:

Let X be an $N \times N$ matrix. Denote by $\boldsymbol{X}$ the $2N \times 2N$ matrix composed of N^2 blocks of size 2×2 whose $i^{\text{th}}, j^{\text{th}}$ block is given by

$$\boldsymbol{X}_{ij} = \begin{pmatrix} X_{ij} & 0 \\ 0 & \overline{X_{ji}} \end{pmatrix} .$$

We will always use subscript Roman indices to refer to the 2×2 blocks of boldface matrices, rather than the individual entries.

Let a and b be complex numbers and define the quaternion $q = a + bj$, where j is a quaternionic basis element. Some basic properties of quaternions are discussed in Appendix C. To work with quaternions and our size-doubled matrices together, we introduce the 2×2 matrix

$$\boldsymbol{j} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} ,$$

and identify the symbol $\boldsymbol{q}$ with the matrix representation of q:

$$q \mapsto \boldsymbol{q} = \boldsymbol{a} + \boldsymbol{bj} = \begin{pmatrix} a & ib \\ i\overline{b} & \overline{a} \end{pmatrix} . \tag{5.1}$$

When $\boldsymbol{q}$ is a 2×2 matrix, and $\boldsymbol{X}$ a $2N \times 2N$ matrix, we use the shorthands

$$\boldsymbol{qX} = (\boldsymbol{q} \otimes I_N)\boldsymbol{X} , \quad \text{and} \quad (\boldsymbol{X} + \boldsymbol{q}) = \boldsymbol{X} + \boldsymbol{q} \otimes I_N .$$

In addition to the usual operations, we define an elementwise product for quaternions

$$(a + bj) \cdot (c + dj) = ac + bdj .$$

Note that the matrix representation of an elementwise product of quaternions is not the same as the usual elementwise product of the matrices, in fact we use

$$\begin{pmatrix} a & ib \\ i\overline{b} & \overline{a} \end{pmatrix} \cdot \begin{pmatrix} c & id \\ i\overline{d} & \overline{c} \end{pmatrix} = \begin{pmatrix} ac & ibd \\ i\overline{bd} & \overline{ac} \end{pmatrix} .$$

5.3.2 Definition and basic properties

Let X be an $N \times N$ matrix, λ a complex variable and ε a strictly positive real number. Putting $q = \lambda + \varepsilon j$, we define the $2N \times 2N$ 'resolvent'

$$\boldsymbol{\mathcal{R}}(q; X) = (\boldsymbol{X} - \boldsymbol{q})^{-1} \, .$$

To connect with other approaches, note that there exists a permutation matrix P such that

$$(\boldsymbol{X} - \boldsymbol{q}) = P \begin{pmatrix} 0 & I_N \\ I_N & 0 \end{pmatrix} H_\varepsilon(\lambda; X) \, P^{-1} \, ,$$

where $H_\varepsilon(\lambda; X)$ is the 'Hermitized' block matrix given in (4.3). The quaternionic Green's function of X at q is then defined to be the quaternion $\mathcal{G}(q; X)$ with matrix representation

$$\boldsymbol{\mathcal{G}}(q; X) = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\mathcal{R}}(q; X)_{ii} \, .$$

It is straightforward to see that for all X and q, the 2×2 matrix $\boldsymbol{\mathcal{G}}(q; X)$ satisfies the symmetries of (5.1) and thus is indeed the matrix representation of a quaternion. Without the hypercomplex part, the quaternionic Green's function agrees with the usual Green's function, that is,

$$\mathcal{G}(\lambda + 0j; X) = G(\lambda; X) + 0j \, .$$

Adding a positive regulariser $\varepsilon > 0$, we find that the regularised spectral density for non-Hermitian matrices introduced in the previous chapter (4.2) may be recovered from the quaternionic Green's function simply by

$$\varrho_\varepsilon(\lambda; X) = -\frac{1}{\pi} \mathrm{Re} \, \frac{\partial}{\partial \overline{\lambda}} \, \mathcal{G}(\lambda + \varepsilon j; X) \, .$$

This formula is in close analogy with the relationship between the ordinary Green's function and regularised spectral density for Hermitian matrices, given by (1.7). Moreover, in light of Theorem 2, computation of the quaternionic Green's function for some ε offers access to the spectral density of the underlying matrix after a size-ε random spherical perturbation.

5.4 Main results

Application of the techniques of the simple approach to the quaternionic Green's function will allow for the derivation of equations describing the spectral density of both Hermitian and non-Hermitian random matrices with independent entries when summed or multiplied with deterministic matrices. These results will be universal in the sense that they hold under only relatively mild conditions on the moments of the distributions of the entries.

Suppose we are in possession of an infinite array of complex random variables $\{\xi_{ij}\}$, with joint probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We assume the ξ_{ij} to have the properties:

A1) $\mathbb{E}\xi_{ij} = 0$ for all i, j

A2) $\mathbb{E}|\xi_{ij}|^2 = 1$ for all i, j

A3) There exists a finite constant C_ξ such that $\mathbb{E}|\xi_{ij}|^3 < C_\xi$ for all i, j

A4) All ξ_{ij} are independent, except for the covariance $\mathbb{E}\xi_{ij}\xi_{ji} = \tau$, where $\tau \in [0, 1]$.

A normalised $N \times N$ random matrix A_N can then be constructed by taking

$$\left(A_N\right)_{ij} = \frac{1}{\sqrt{N}}\xi_{ij}\,. \tag{5.2}$$

The parameter τ here is the same as introduced earlier in the derivation of Girko's Elliptic Law, controlling the degree of Hermiticity of A_N. In our calculations, τ will only appear as the real part of the quaternion $t = \tau + j$.

The assumptions A1-A4 above are by no means the weakest conditions for which the results we derive will hold, however they sufficiently broad to give a flavour of the strength of the results, without complicating the proofs unnecessarily.

In the forthcoming results, we will characterise the quaternionic Green's functions of the sum or product of such matrices with deterministic matrices $\{D_N\}$ satisfying some or all of the assumptions

D1) The quaternionic Green's functions of $\{D_N\}$ converge to the limit $\mathcal{G}_D$

D2) The quaternionic Green's functions of $\{D_N^{-1}\}$ converge to the limit $\mathcal{G}_{D^{-1}}$

D3) There exists a constant $d \in [0, \infty)$ such that $\sup \|D_N\| < d$ and $\sup \|D_N^{-1}\| < d$.

The last point is a technical assumption made for the sake of simplicity and, as we note in a later example, may not be strictly necessary.

Theorem 3 (Sum Rule)**.** *Let $\{A_N\}$ be a sequence of random matrices given by (5.2) and let $\{D_N\}$ be a sequence of deterministic matrices satisfying D1. Fix a quaternion $q = \lambda + \varepsilon j$, where $\varepsilon > 1$. Then $\mathcal{G}(q; D_N + A_N) \xrightarrow{\mathbb{P}} \mathcal{G}(q)$, where $\mathcal{G}(q)$ satisfies*

$$\mathcal{G}(q) = \mathcal{G}_D\Big(q + t \cdot \mathcal{G}(q)\Big), \tag{5.3}$$

Theorem 3 is a straightforward generalisation of the Pastur equation for the sum of deterministic and random Hermitian matrices, indeed at $\tau = 1$ and $\varepsilon = 0$, equation (5.3) precisely *is* the Pastur equation. For the case $\tau = 0$, an equivalent result has already been found by Khoruzhenko using techniques based on the electrostatic analogy [Kho96].

Theorem 4 (Product Rule)**.** *Let $\{A_N\}$ be a sequence of random matrices given by (5.2) and let $\{D_N\}$ be a sequence of deterministic matrices satisfying D1-D3. Fix a quaternion $q = \lambda + \varepsilon j$, where $\varepsilon > 2d$. Then*

$$\mathcal{G}(q; D_N A_N) \xrightarrow{\mathbb{P}} \mathcal{G}(q) = -\big(t \cdot \widetilde{\mathcal{G}}\big)^{-1} \mathcal{G}_D\Big(-q\,\big(t \cdot \widetilde{\mathcal{G}}\big)^{-1}\Big), \tag{5.4}$$

where $\widetilde{\mathcal{G}}$ satisfies

$$\widetilde{\mathcal{G}} = -q^{-1} \mathcal{G}_{D^{-1}}\Big(-\big(t \cdot \widetilde{\mathcal{G}}\big)\, q^{-1}\Big). \tag{5.5}$$

Unlike Theorem 3, this result is not related to any in Hermitian RMT for the simple reason that the space of Hermitian matrices is not closed under multiplication. The spectral density of certain Hermitian random matrix products have been considered on occasion however, for example in [BMJ07] and references therein. Before the proof of these results is presented, we give a number of simple examples of their application.

5.5 Examples

The statement of the sum and product rules given above concerns the behaviour of the quaternionic Green's function in the limit $N \to \infty$, for a fixed regulariser ε, which is taken to be large. In light of Theorem 2 we are in effect computing the limiting spectral density of matrices under a large perturbation. However, as the following examples will demonstrate, accurate predictions about the spectral densities of sums and products of matrices satisfying conditions A1-A4 and D1-D3 can be found by naively taking $\varepsilon = 0$ in the final equations.

Girko's Elliptic Law

This well-known result occurs naturally. Taking either $D_N = 0$ in the sum rule, or $D_N = I_N$ in the product rule, gives $\mathcal{G}(q; A_N) \xrightarrow{\mathbb{P}} \mathcal{G}(q)$, where

$$\mathcal{G}(q) = -\Big(q + t \cdot \mathcal{G}(q)\Big)^{-1}.$$

Writing $\mathcal{G}(q) = \alpha + \beta j$, we send $\varepsilon \to 0$ and assume that β stays strictly positive, obtaining

$$(\alpha + \beta j)(\lambda + \tau\alpha + \beta j) + 1 = 0\,. \tag{5.6}$$

This equation is simply a more compact expression of (4.15), found in the fully connected limit of the cavity approach. The support of the spectral density is restricted

to the region allowing a solution with $\beta > 0$. The hypercomplex part (that is, the coefficient of j) in (5.6) gives

$$\alpha = -\left(\frac{x}{\tau+1} + i\frac{y}{\tau-1}\right),$$

where $\lambda = x + iy$, and the complex part gives

$$\beta = \sqrt{1 - \left(\frac{x}{\tau+1}\right)^2 - \left(\frac{y}{\tau-1}\right)^2}.$$

The condition $\beta > 0$ determines the elliptic support, and taking an anti-holomorphic derivative yields the spectral density inside that region:

$$\rho(\lambda) = \begin{cases} \frac{1}{\pi(1-\tau^2)} & \text{when} \quad \left(\frac{x}{\tau+1}\right)^2 + \left(\frac{y}{\tau-1}\right)^2 < 1 \\ 0 & \text{otherwise.} \end{cases}$$

A matrix model in QCD

Let us consider the following random matrix model for the Dirac operator in QCD with finite chemical potential, considered in [Ste96],

$$\begin{pmatrix} 0 & iA_N + \mu \\ iA_N^\dagger + \mu & 0 \end{pmatrix}, \tag{5.7}$$

where A_N is drawn from the Ginibre Unitary Ensemble and $\mu > 0$ is the chemical potential. Following similar lines to [JN04], we consider a simpler model of the form $A_N + iM_N$, where A_N now is taken to be a Hermitian GUE matrix and

$$M_N = \begin{pmatrix} 0 & \mu \\ \mu & 0 \end{pmatrix} \otimes I_{N/2}\,.$$

We will show that, after multiplication by i, the $N \to \infty$ limit of the spectral density of (5.7) is the same as that of this simpler model. For this construction, the sum rule states that $\mathcal{G}(q; A_N + iM_N) \xrightarrow{\mathbb{P}} \alpha + \beta j$, where equation (5.3) now reads

$$\alpha + \beta j = \frac{1}{2}\Big(i\mu - q - \alpha - \beta j\Big)^{-1} - \frac{1}{2}\Big(i\mu + q + \alpha + \beta j\Big)^{-1}. \tag{5.8}$$

Assuming the limit $\varepsilon \to 0$, we write $q = \lambda$. The support of spectral density is then given by the region allowing a solution of (5.8) with $\beta > 0$, in this case determined by the condition

$$1 - \frac{1}{4}x^2 - \frac{1}{4}y^2\Big(1 + 2\big(y^2 - \mu^2\big)\Big)^2\big(y^2 - \mu^2\big)^{-2} + \big(y^2 - \mu^2\big) > 0\,.$$

Inside this region, one may solve for α and take the anti-holomorphic derivative to determine the spectral density

$$\rho(\lambda) = \frac{1}{4\pi}\left(\frac{y^2 + \mu^2}{(y^2 - \mu^2)^2} - 1\right)\,. \tag{5.9}$$

As one would hope, this is precisely the density first recovered by Stephanov in [Ste96], rotated by $\pi/2$. In that work, only the mean of the spectral density was computed, however the sum rule tells us that the quaternionic Green's function converges in probability as $N \to \infty$, suggesting weak convergence in spectral density. Moreover, only Gaussian distributed random matrices were considered in [Ste96], whereas the sum rule predicts the limiting density to be universal in the sense that it is independent of distribution of the entries of A_N. An analogue of the construction (5.7) for orthogonal and symplectic ensembles has been studied in [HOV97], where, interestingly, the limiting densities were found to be different to those for the unitary ensemble.

Perturbations of Jordan matrices

Consider $N \times N$ Jordan matrices X_N with entries $[X_N]_{ij} = \delta_{i,j+1}$. Let A_N, B_N and C_N be $N \times N$ matrices of complex Gaussian random variables. For parameters $\varepsilon \in (0, \infty)$ and $\delta \in [0, \infty)$ we consider random perturbations of J_N of the form

$$X_N = J_N + \delta A_N + \varepsilon B_N C_N^{-1}.$$

Perturbations of Jordan matrices have been studied extensively in the past as they provide a simple model with which to study the behaviour of degenerate eigenvalues

under perturbation. In [DH09], perturbations of the type given above with $\varepsilon = 0$ were considered, and it was shown that with high probability the majority of eigenvalues of $J_N + \delta A_N$ lie close to the circle around the origin of radius $\delta^{1/N}$. In the limit $N \to \infty$, the eigenvalues are pushed towards the unit circle, where one would expect to find a uniform distribution in the limit $\delta \to 0$.

To access the spectral density of perturbations of J_N, the inclusion of additional perturbations of the type $\varepsilon B_N C_N^{-1}$ allows the use of Theorem 2 and thus the sum rule. Averaging over B_N and C_N, and taking the limit $N \to \infty$, we have the weak convergence

$$\mathbb{E}_{B_N, C_N}\Big[\varrho\left(\lambda; X_N\right)\Big] \to \rho(\lambda; \delta, \varepsilon) = -\frac{1}{\pi}\mathrm{Re}\,\frac{\partial}{\partial\overline{\lambda}}\,\mathcal{G}(\lambda + \varepsilon j; \delta)\,,$$

where $\mathcal{G}(\lambda + \varepsilon j; \delta)$ is determined by the sum rule. The full expression is somewhat complicated and not especially enlightening, though taking one or other perturbation to zero, we obtain the results:

$$\lim_{\varepsilon \to 0} \rho(\lambda; \delta, \varepsilon) = \frac{1}{\pi\delta^2}\left(1 - \frac{1}{\sqrt{4|\lambda|^2 + \delta^4}}\right)\mathbb{I}_{[1-\delta^2, 1+\delta^2]}(|\lambda|^2)\,, \tag{5.10}$$

and

$$\lim_{\delta \to 0} \rho(\lambda; \delta, \varepsilon) = \frac{\left(1 + |\lambda|^2 + \varepsilon^2\right)\varepsilon^2}{\pi\left(\left(1 + |\lambda|^2 + \varepsilon^2\right)^2 - 4|\lambda|^2\right)^{3/2}}\,. \tag{5.11}$$

These expressions agree as $\varepsilon \to 0$ and $\delta \to 0$, converging to the uniform distribution on the unit circle. Figure 5.2 shows cross-sections of these densities along the real line.

As mentioned earlier, a result equivalent to that of the sum rule found in this chapter was derived in [Kho96]. The example of the Jordan matrix was also considered in that paper and the support of the density was computed for the $\varepsilon = 0$ case, given here by the indicator function in (5.10).

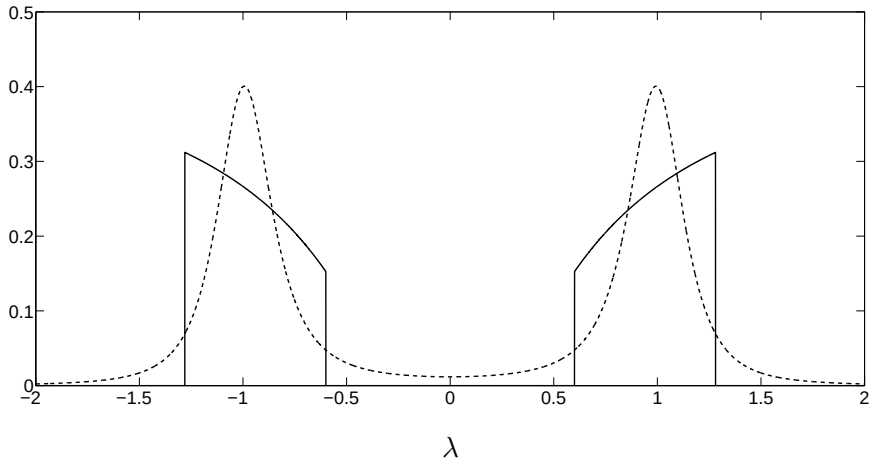


Figure 5.2: Solid line: a cross-section through the density $\lim_{\varepsilon\to 0}\ \rho(\lambda;\delta,\varepsilon)$ at $\delta = 0.8$. Dashed line: a cross-section through the density $\lim_{\delta\to 0}\ \rho(\lambda;\delta,\varepsilon)$ at $\varepsilon = 0.2$.

A product of random matrices

We compute the limiting spectral density for the product matrix $D_N A_N$, where the A_N are given by (5.2) with $\tau = 0$, and D_N is a diagonal matrix with entries D_{ii} drawn independently from the standard Cauchy distribution[2]. With this choice of D_N, we have the limits

$$\mathcal{G}(q; D_N) \to \mathcal{G}_D(q) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{1+r^2}(r-q)^{-1}dr\,,$$

and $\mathcal{G}(q; D_N^{-1}) \to \mathcal{G}_D(q)$, as $N \to \infty$. As before, we assume the $\varepsilon \to 0$ limit, taking $q = \lambda$. Then $\widetilde{\mathcal{G}}(q) = \widetilde{\alpha} + \widetilde{\beta} j$, where the product rule (5.5) gives

$$\widetilde{\alpha} + \widetilde{\beta} j = -\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{1+r^2}\left(\frac{\lambda}{r} + \widetilde{\beta} j\right)^{-1} dr\,.$$

Performing the integral and solving for $\widetilde{\beta}$, we obtain $\widetilde{\beta} = \left(\sqrt{|\lambda|^2+4} - |\lambda|\right)/2$. Returning to (5.4), we reach

$$\mathcal{G}(q; D_N A_N) \xrightarrow{\mathbb{P}} -\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{1+r^2}\Big(\lambda + r\widetilde{\beta} j\Big)^{-1} dr = \frac{2\overline{\lambda}}{|\lambda|^2 + |\lambda|\sqrt{|\lambda|^2+4}}\,,$$

[2]Notice that although this choice fails the technical assumption D3, the result still appears to hold.

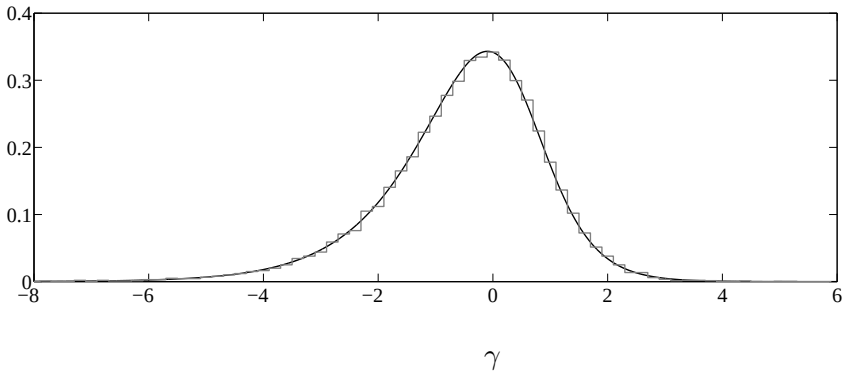


Figure 5.3: Grey histogram: the log-moduli of eigenvalues of a single random matrix $D_N A_N$ of size $N = 10^4$, where the entries of A_N are independent complex Gaussians with variance $1/N$, and D_N is a diagonal matrix of Cauchy random variables. Black line: the density $\nu(\gamma)$ predicted by the product rule (Theorem 4), given in this case by equation (5.12) in the text.

and finally an expression for the limiting spectral density,

$$\rho(\lambda) = \frac{1}{\pi}\left(\frac{1}{|\lambda|^2 + |\lambda|\sqrt{|\lambda|^2+4}} - \frac{1}{|\lambda|^2 + |\lambda|\sqrt{|\lambda|^2+4}+4}\right).$$

To provide an effective comparison with numerical data, we change variables to $\gamma = \log|\lambda|$, whose distribution is given by the pdf

$$\nu(\gamma) = 2\frac{e^{\gamma}\sqrt{e^{2\gamma}+4} - e^{2\gamma}}{e^{\gamma}\sqrt{e^{2\gamma}+4} + e^{2\gamma}+4}\,. \tag{5.12}$$

Figure 5.5 shows a histogram of the log-moduli of the eigenvalues of a single such random matrix of size $N = 10^4$ alongside the predicted density $\nu(\gamma)$.

5.6 Proof of the main results

5.6.1 Preliminaries for the use of the quaternionic Green's function

For the proofs of Theorems 3 and 2, a number of standard tools will be of repeated use.

The resolvent identity. For $N \times N$ matrices A and B, and quaternion q, we have

$$\boldsymbol{\mathcal{R}}(q;A) - \boldsymbol{\mathcal{R}}(q;B) = \boldsymbol{\mathcal{R}}(q;A)(\boldsymbol{B} - \boldsymbol{A})\boldsymbol{\mathcal{R}}(q;B)\,. \tag{5.13}$$

This is a consequence of the more general expression for any $\boldsymbol{X}$ and $\boldsymbol{Y}$,

$$\boldsymbol{X}^{-1} - \boldsymbol{Y}^{-1} = \boldsymbol{X}^{-1}(\boldsymbol{Y} - \boldsymbol{X})\boldsymbol{Y}^{-1}\,.$$

The resolvent bound. For any matrix A and quaternion $q = \lambda + \varepsilon j$, we have the following bound on the norm of the resolvent and its blocks

$$\left\|\boldsymbol{\mathcal{R}}(q;A)_{ij}\right\| \leq \left\|\boldsymbol{\mathcal{R}}(q;A)\right\| \leq \frac{1}{\varepsilon}\,, \tag{5.14}$$

where $\|\cdot\|$ is the spectral norm. Again, this is a special case of a more general result: for any matrices X and Y, with Y invertible, we have

$$\left\|(\boldsymbol{X} + \boldsymbol{Y}\boldsymbol{j})^{-1}_{ij}\right\| \leq \left\|(\boldsymbol{X} + \boldsymbol{Y}\boldsymbol{j})^{-1}\right\| \leq \frac{1}{x+y}\,,$$

where x and y are the smallest singular values of X and Y.

The Cauchy-Schwarz inequality. Let $\{\boldsymbol{x}_i\}_{i=1}^N$ and $\{\boldsymbol{y}_i\}_{i=1}^N$ be collections of 2×2 matrices. We have the following analogue of the Cauchy-Schwarz inequality:

$$\left\|\sum_{i=1}^N \boldsymbol{x}_i \boldsymbol{y}_i^\dagger\right\|^2 \leq \left\|\sum_{i=1}^N \boldsymbol{x}_i \boldsymbol{x}_i^\dagger\right\| \left\|\sum_{i=1}^N \boldsymbol{y}_i \boldsymbol{y}_i^\dagger\right\|\,. \tag{5.15}$$

Integration by parts. Let $f : \mathbb{C}^n \to \mathbb{C}$ be a continuous and differentiable (but not generally analytic) function, with bounded second order partial derivatives,

$$\sup_{\vec{x},i,j} \left\{ \left| \frac{\partial}{\partial x_i \partial x_j} f(\vec{x}) \right|, \left| \frac{\partial}{\partial x_i \partial \overline{x_j}} f(\vec{x}) \right|, \left| \frac{\partial}{\partial \overline{x_i} \partial \overline{x_j}} f(\vec{x}) \right| \right\} < C_f \,.$$

Then if $F = f(\xi_{11}, \ldots, \xi_{NN})$, we have

$$\mathbb{E}\xi_{ij} F = \mathbb{E}\frac{\partial F}{\partial \overline{\xi_{ij}}} + \tau \mathbb{E}\frac{\partial F}{\partial \xi_{ji}} + \mathbb{E}\xi_{ij}^2 \, \mathbb{E}\frac{\partial F}{\partial \xi_{ij}} + \mathbb{E}\xi_{ij}\overline{\xi_{ji}} \, \mathbb{E}\frac{\partial F}{\partial \overline{\xi_{ji}}} + k \,, \tag{5.16}$$

where $|k|$ is bounded by some constant depending on C_ξ and C_f. Proof is by Taylor's Theorem.

In addition to these general facts, the main part of the work in proving Theorems 3 and 4 comes down to the application of two central results.

Lemma 1. *Let A_N be a random $N \times N$ matrix given by (5.2) and let X, Y and Z be arbitrary deterministic matrices of the same size, with Z invertible. Define*

$$\boldsymbol{R} = (\boldsymbol{X}\boldsymbol{A}_N + \boldsymbol{Y} + \boldsymbol{Z}\boldsymbol{j})^{-1} \,, \quad \textit{and} \quad \boldsymbol{G} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{R}_{ii} \,.$$

Then

$$\mathbb{E}\left\| \boldsymbol{G} - \mathbb{E}\boldsymbol{G} \right\|^2 < CN^{-1} \,,$$

where C is a constant depending on X, Y and Z.

This is simply a slightly more general version of the statement that the quaternionic Green's functions of the matrices we are interested in are self-averaging in the limit $N \to \infty$. This property is crucial if we are to extract useful information about the behaviour in that limit. The proof is based on an elegant martingale technique from a paper on spin-glasses [CH06].

Proof. Let $P = \{(i,j) : 1 \leq i \leq j \leq N\}$. We label these pairs by introducing the bijective numbering $(i,j) \leftrightarrow p$ where $p \in \{1, ..., |P|\}$. The names (i,j) and p will be

used interchangeably; the meaning should be clear from the context. Let $\mathcal{F}_0 = \mathcal{F}$ and recursively define the sub-σ-algebras $\mathcal{F}_p = \sigma\{\mathcal{F}_{p-1}, \xi_{ij}, \xi_{ji}\}$. Introduce the martingale

$$\Delta_p = \mathbb{E}\left(\boldsymbol{G} \,\middle|\, \mathcal{F}_p\right) - \mathbb{E}\left(\boldsymbol{G} \,\middle|\, \mathcal{F}_{p-1}\right) ,$$

so that

$$\sum_{p=1}^{|P|} \Delta_p = \boldsymbol{G} - \mathbb{E}\boldsymbol{G} .$$

We plan to bound each Δ_p, to do so, we consider the fictitious situation in which the blocks $\left(\boldsymbol{A}_N\right)_{ij}$ and $\left(\boldsymbol{A}_N\right)_{ji}$ are removed. Write $A_N^{(ij)}$ for the matrix obtained from A_N by setting $\xi_{ij} = \xi_{ji} = 0$. Introduce

$$\boldsymbol{R}^{(ij)} = \left(\boldsymbol{X}\boldsymbol{A}_N^{(ij)} + \boldsymbol{Y} + \boldsymbol{Z}\boldsymbol{j}\right)^{-1} , \quad \text{and} \quad \boldsymbol{G}^{(ij)} = \frac{1}{N}\sum_{k=1}^{N} \boldsymbol{R}_{kk}^{(ij)} .$$

The resolvent identity (5.13) provides

$$\boldsymbol{R} = \boldsymbol{R}^{(ij)} - \boldsymbol{R}\boldsymbol{X}\left(\boldsymbol{A}_N - \boldsymbol{A}_N^{(ij)}\right)\boldsymbol{R}^{(ij)}$$

and thus

$$\boldsymbol{G} = \boldsymbol{G}^{(ij)} - \frac{1}{N}\boldsymbol{K}_{ij} ,$$

where the error term $\boldsymbol{K}_{ij}$ is given by

$$\begin{aligned}
\boldsymbol{K}_{ij} &= \sum_{k=1}^{N} \left(\boldsymbol{R}\boldsymbol{X}\left(\boldsymbol{A}_N - \boldsymbol{A}_N^{(ij)}\right)\boldsymbol{R}^{(ij)}\right)_{kk} \\
&= \sum_{k=1}^{N} \left(\left(\boldsymbol{R}\boldsymbol{X}\right)_{ki}\left(\boldsymbol{A}_N\right)_{ij}\boldsymbol{R}_{jk}^{(ij)} + \left(\boldsymbol{R}\boldsymbol{X}\right)_{kj}\left(\boldsymbol{A}_N\right)_{ji}\boldsymbol{R}_{ik}^{(ij)}\right) .
\end{aligned}$$

The Cauchy-Schwarz inequality (5.15) provides a bound for $\boldsymbol{K}_{ij}$, since, for example

$$\begin{aligned}
&\left\| \sum_{k=1}^{N} \left(\boldsymbol{R}\boldsymbol{X}\right)_{ki}\left(\boldsymbol{A}_N\right)_{ij}\boldsymbol{R}_{jk}^{(ij)} \right\|^2 \\
&\qquad \le \left\| \sum_{k=1}^{N} \left(\boldsymbol{R}\boldsymbol{X}\right)_{ki}\left(\boldsymbol{R}\boldsymbol{X}\right)_{ki}^{\dagger} \right\| \left\| \left(\boldsymbol{A}_N\right)_{ij}\left(\boldsymbol{A}_N\right)_{ij}^{\dagger} \right\| \left\| \sum_{k=1}^{N} \boldsymbol{R}_{jk}^{(ij)}\boldsymbol{R}_{jk}^{(ij)\dagger} \right\| \\
&\qquad = \left\| \left(\boldsymbol{X}^T\boldsymbol{R}^T\boldsymbol{X}^{\dagger}\boldsymbol{R}^{\dagger}\right)_{ii} \right\| \left\| \left(\boldsymbol{A}_N\right)_{ij}\left(\boldsymbol{A}_N\right)_{ij}^{\dagger} \right\| \left\| \left(\boldsymbol{R}^{(ij)}\boldsymbol{R}^{(ij)\dagger}\right)_{ii} \right\| \\
&\qquad\qquad \le \frac{CM_{ij}^2}{N} ,
\end{aligned}$$

where $M_{ij} = \max\{|\xi_{ij}|, |\xi_{ji}|\}$, and C is the constant coming from the resolvent bound (5.14). We can conclude $\|\boldsymbol{K}_{ij}\| < C_K M_{ij} N^{-1/2}$, for some constant C_K, and thus

$$\begin{aligned}\big\|\Delta_p\big\| &= \Big\|\mathbb{E}\left(\boldsymbol{G} \,\big|\, \mathcal{F}_p\right) - \mathbb{E}\left(\boldsymbol{G} \,\big|\, \mathcal{F}_{p-1}\right)\Big\| \\ &= \Big\|\mathbb{E}\left(\boldsymbol{E}_{ij} \,\big|\, \mathcal{F}_p\right) - \mathbb{E}\left(\boldsymbol{E}_{ij} \,\big|\, \mathcal{F}_{p-1}\right)\Big\| \\ &\leq C_K N^{-3/2}\big(M_{ij} - \mathbb{E}M_{ij}\big)\,.\end{aligned}$$

Burkholder's inequality then gives, for some constant C_Δ,

$$\begin{aligned}\mathbb{E}\Big\|\boldsymbol{G} - \mathbb{E}\boldsymbol{G}\Big\|^3 = \mathbb{E}\left\|\sum_{p=1}^{|P|}\Delta_p\right\|^3 &\leq C_\Delta\,\mathbb{E}\left(\sum_{p=1}^{|P|}\big\|\Delta_p\big\|^2\right)^{3/2} \\ &\leq C_\Delta C_K^3 N^{-9/2}\,\mathbb{E}\left(\sum_{i\leq j}\big(M_{ij} - \mathbb{E}M_{ij}\big)^2\right)^{3/2} \\ &< 4C_\Delta C_K^3 C_M N^{-3/2}\,,\end{aligned}$$

where $C_M < 2C_\xi + 2$ is a bound for $\mathbb{E}M_{ij}^3$, and we have repeatedly used Jensen's inequality. The desired bound is then given by

$$\mathbb{E}\Big\|\boldsymbol{G} - \mathbb{E}\boldsymbol{G}\Big\|^2 < CN^{-1}\,,$$

where $C = (4C_\Delta C_K^3 C_M)^{2/3}$. □

With self-averaging established, we next require a mechanism by which we can convert the general statement of the resolvent identity (5.13) into an equation for the mean of the quaternionic Green's function.

Lemma 2. *Fix a quaternion $q = \lambda + \varepsilon j$, with $\varepsilon > 1$. Let A_N be a random $N \times N$ matrix given by (5.2) and let X be an arbitrary deterministic matrix of the same size. Define*

$$\boldsymbol{R} = \big(\boldsymbol{X}\boldsymbol{A}_N - \boldsymbol{q}\big)^{-1} \quad \textit{and} \quad \widetilde{\boldsymbol{G}} = N^{-1}\sum_{i=1}^{N}\big(\boldsymbol{R}\boldsymbol{X}\big)_{ii}\,.$$

Then

$$\mathbb{E}\big(\boldsymbol{A}_N\boldsymbol{R}\big) = -\big(\mathbb{E}(\boldsymbol{t}\cdot\widetilde{\boldsymbol{G}})\big)(\mathbb{E}\boldsymbol{R}) + \boldsymbol{K}_N\,.$$

where $t = \tau + j$, and $\|\boldsymbol{K}_N\| \to 0$ as $N \to \infty$.

Proof. We compute a generic block

$$\mathbb{E}\big(\boldsymbol{A}_N\boldsymbol{R}\big)_{ik} = \frac{1}{\sqrt{N}}\mathbb{E}\sum_{j=1}^{N}\begin{pmatrix}\xi_{ij} & 0\\ 0 & \overline{\xi_{ji}}\end{pmatrix}\boldsymbol{R}_{jk}\,. \tag{5.17}$$

Notice that the entries of $\boldsymbol{R}$ depend continuously upon the ξ, and the resolvent bound gives a bound on the first and higher order derivatives. We are thus able to apply the integration by parts formula (5.16). The derivatives are given by

$$\frac{\partial \boldsymbol{R}}{\partial\overline{\xi_{\alpha\beta}}} = -\frac{1}{\sqrt{N}}\boldsymbol{R}\boldsymbol{X}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\mathbf{1}_{\beta\alpha}\boldsymbol{R}$$

and

$$\frac{\partial \boldsymbol{R}}{\partial\xi_{\alpha\beta}} = -\frac{1}{\sqrt{N}}\boldsymbol{R}\boldsymbol{X}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\mathbf{1}_{\alpha\beta}\boldsymbol{R}\,,$$

where $\mathbf{1}_{\alpha\beta}$ is the $2N\times 2N$ block matrix containing a copy of I_2 in block (α,β) and zeros elsewhere. Applying this to (5.17), we thus find, after some tedious algebra,

$$\begin{aligned}\mathbb{E}\big(\boldsymbol{A}_N\boldsymbol{R}\big)_{ik} &= -\frac{1}{N}\sum_{j=1}^{N}\Big[\Big(\boldsymbol{t}\cdot\big(\boldsymbol{R}\boldsymbol{X}\big)_{jj}\Big)\boldsymbol{R}_{ik} + \Big(\boldsymbol{s}_{ij}\cdot\big(\boldsymbol{R}\boldsymbol{X}\big)_{ji}\Big)\boldsymbol{R}_{jk}\Big]\\ &= -\mathbb{E}\big((\boldsymbol{t}\cdot\widetilde{\boldsymbol{G}})\boldsymbol{R}_{ik}\,\big) - \frac{1}{N}\sum_{j=1}^{N}\Big(\boldsymbol{s}_{ij}\cdot\big(\boldsymbol{R}\boldsymbol{X}\big)_{ji}\Big)\boldsymbol{R}_{jk}\,,\end{aligned}$$

where $s_{ij} = \mathbb{E}\xi_{ij}^2 + (\mathbb{E}\xi_{ij}\overline{\xi_{ji}})j$. Notice that assumptions A1-A4 imply a universal bound for $|s_{ij}|$, and the Cauchy-Schwarz inequality together with the resolvent bound give a constant C such that

$$\frac{1}{N}\left\|\sum_{j=1}^{N}\Big(\boldsymbol{s}_{ij}\cdot\big(\boldsymbol{R}\boldsymbol{X}\big)_{ji}\Big)\boldsymbol{R}_{jk}\right\| < CN^{-1}\,.$$

Finally, we split the expectation

$$\Big\|\mathbb{E}\big((\boldsymbol{t}\cdot\widetilde{\boldsymbol{G}})\boldsymbol{R}\big) - \big(\mathbb{E}(\boldsymbol{t}\cdot\widetilde{\boldsymbol{G}})\big)(\mathbb{E}\boldsymbol{R})\Big\| \to 0 \quad \text{as} \quad N\to\infty,$$

since $\mathbb{E}\big\|\widetilde{\boldsymbol{G}} - \mathbb{E}\widetilde{\boldsymbol{G}}\big\|^2 \to 0$ by Lemma 1, $|t|<\infty$ and $\|\boldsymbol{R}\|$ is bounded. □

5.6.2 Proof of sum and product rules

Proof of Theorem 3. Fix $q = \lambda + \varepsilon j$ with $\varepsilon > 1$. We write the shorthands $\boldsymbol{R}_N = \boldsymbol{\mathcal{R}}(q; D_N + A_N)$, $\boldsymbol{G}_N = \boldsymbol{\mathcal{G}}(q; D_N + A_N)$ and $\mathcal{G}_N(q) = \mathcal{G}(q; D_N + A_N)$. Now, applying the resolvent identity and Lemma 2, we obtain

$$\begin{aligned}\mathbb{E}\boldsymbol{R}_N - (\boldsymbol{D}_N - \boldsymbol{q})^{-1} &= -(\boldsymbol{D}_N - \boldsymbol{q})^{-1}\mathbb{E}\boldsymbol{A}_N\boldsymbol{R}_N \\ &= (\boldsymbol{D}_N - \boldsymbol{q})^{-1}\big(\mathbb{E}(\boldsymbol{t}\cdot\boldsymbol{G}_N)\big)(\mathbb{E}\boldsymbol{R}_N) + \boldsymbol{K}_N\,,\end{aligned}$$

where $\|\boldsymbol{K}_N\| \to 0$ as $N \to \infty$. Rearranging, we have

$$\mathbb{E}\boldsymbol{R}_N = \Big(\boldsymbol{D}_N - \boldsymbol{q} - \boldsymbol{t}\cdot\mathbb{E}\boldsymbol{G}_N\Big)^{-1} + \boldsymbol{K}'_N\,,$$

where the resolvent bound gives $\|\boldsymbol{K}'_N\| \to 0$ as $N \to \infty$ also. Summing over the diagonal blocks, we deduce

$$\mathbb{E}\mathcal{G}_N(q) = \mathcal{G}\Big(q + t\cdot\mathbb{E}\mathcal{G}_N(q); D_N\Big) + k_N\,, \tag{5.18}$$

with $|k_N| \to 0$ as $N \to \infty$. Define the functions

$$f_N(g) = \mathcal{G}_{D_N}\Big(q + t\cdot g\Big) + k_N\,, \qquad f(g) = \mathcal{G}_D\Big(q + t\cdot g\Big)\,.$$

Since $\varepsilon > 1$, the resolvent bound gives that each f_N is a contraction with parameter ε^{-1}, and thus the pointwise limit f is also; it is therefore continuous, with a unique fixed point which we call $\mathcal{G}(q)$. Finally, Lemma 1 gives $\mathbb{E}|\mathcal{G}_N(q) - \mathcal{G}_N(q)|^2 \to 0$ as $N \to \infty$, and we conclude from (5.18) and Tchebychev's inequality that

$$\mathcal{G}_N(q) \xrightarrow{\mathbb{P}} \mathcal{G}(q)\,.$$

□

Proof of Theorem 4. The proof is very similar to that of Theorem 3, so we give only the main steps of the derivation.

Fix $q = \lambda + \varepsilon j$ with $\varepsilon > 2d$. Again we write the shorthands $\boldsymbol{R}_N = \boldsymbol{\mathcal{R}}(q; D_N A_N)$ and $\mathcal{G}_N(q) = \mathcal{G}(q; D_N A_N)$. Let $\widetilde{\boldsymbol{G}}_N = N^{-1} \sum_{i=1}^{N} (\boldsymbol{R}_N \boldsymbol{D}_N)_{ii}$. Applying the resolvent identity and Lemma 2, we obtain

$$\begin{aligned} \mathbb{E}\boldsymbol{R}_N + \boldsymbol{q}^{-1} &= \boldsymbol{q}^{-1} \boldsymbol{D}_N \mathbb{E}\boldsymbol{A}_N \boldsymbol{R}_N \\ &= -\boldsymbol{q}^{-1} \boldsymbol{D}_N \big(\mathbb{E}(\boldsymbol{t} \cdot \widetilde{\boldsymbol{G}}_N)\big)(\mathbb{E}\boldsymbol{R}_N) + \boldsymbol{K}_N, \end{aligned}$$

Rearranging, we have

$$\mathbb{E}\boldsymbol{R}_N = -\big(\mathbb{E}(\boldsymbol{t} \cdot \widetilde{\boldsymbol{G}}_N)\big)^{-1} \Big(\boldsymbol{D}_N + \boldsymbol{q} \big(\mathbb{E}(\boldsymbol{t} \cdot \widetilde{\boldsymbol{G}}_N)\big)^{-1} \Big)^{-1} + \boldsymbol{K}'_N \,,$$

where the resolvent bound gives $\|\boldsymbol{K}'_N\| \to 0$ as $N \to \infty$ also (it is here that assumption D3 and the requirement $\varepsilon > 2d$ are used). Summing over the diagonal blocks, we deduce

$$\mathbb{E}\mathcal{G}_N(q) = -\big(\mathbb{E}(t \cdot \widetilde{\mathcal{G}}_N)\big)^{-1} \mathcal{G}\Big(- q \big(\mathbb{E}(t \cdot \widetilde{\mathcal{G}}_N)\big)^{-1}; D_N \Big) + k_N \,, \tag{5.19}$$

where $\widetilde{\mathcal{G}}_N$ is the quaternion with matrix representation $\widetilde{\boldsymbol{G}}_N$, and $|k_N| \to 0$ as $N \to \infty$. The equation for $\widetilde{\boldsymbol{G}}$ is found similarly; let $\widetilde{\boldsymbol{R}}_N = \boldsymbol{R}_N \boldsymbol{D}_N$, then

$$\begin{aligned} \mathbb{E}\widetilde{\boldsymbol{R}}_N + \boldsymbol{q}^{-1} D_N &= \boldsymbol{q}^{-1} \boldsymbol{D}_N \mathbb{E}\boldsymbol{A}_N \boldsymbol{R}_N \boldsymbol{D}_N \\ &= -\boldsymbol{q}^{-1} \boldsymbol{D}_N \big(\mathbb{E}(\boldsymbol{t} \cdot \widetilde{\boldsymbol{G}}_N)\big)(\mathbb{E}\widetilde{\boldsymbol{R}}_N) + \widetilde{\boldsymbol{K}}_N \,. \end{aligned}$$

and thus

$$\mathbb{E}\widetilde{\mathcal{G}}_N = -q^{-1} \mathcal{G}\left(-\big(t \cdot \mathbb{E}\widetilde{\mathcal{G}}_N\big) q^{-1}; D_N^{-1} \right) + \widetilde{k}_N \,, \tag{5.20}$$

where as usual $|\widetilde{k}_N| \to 0$ as $N \to \infty$. The proof is completed in the same fashion as Theorem 3, with equations (5.19) and (5.20) providing (5.4) and (5.5), respectively. $\square$

6

Conclusions and outlook

Several new results have been presented in this thesis on the spectral density of random matrices, the majority of which have appeared in print in the articles [RPKT08, RP09, Rog09, RPVT10]. In this brief final chapter we recap the main developments of the thesis and discuss some limitations of the work and potential for further research.

Sparse Hermitian random matrices associated to tree-like random graphs were considered in Chapter 2. The statistical mechanics analogy for the spectral density was revisited using the cavity method, resulting in a collection of equations whose solution provides a close approximation to the Green's function of a given matrix. The success of the approach is evidenced by the recovery of known results of Wigner and McKay, as well as the results of numerical simulations. Efficient numerical solution of the cavity equations by a belief propagation algorithm was demonstrated for sparse matrices several orders of magnitude larger than is possible by numerical diagonalisation on the same hardware. The ensemble average for Poissonian random graphs was computed, in which the cavity equations are transformed into a self consistency equation on a probability density of cavity variables. Again, efficient numerical solution is possible, this time through a population dynamics algorithm.

In the next chapter, the ensemble average spectral density was computed for some

more complex random graph ensembles. Ensembles featuring correlations in the degrees of neighbouring and non-neighbouring vertices were studied using the device of generalised degrees, with the ensemble average applied to cavity equations in this case yielding consistency equations for a collection of probability densities. The constrained generalised degree ensemble of [BCV08] was introduced as a concrete example of a tree-like random graph ensemble for which the average spectral density could be calculated in this way, and the appropriate generalised degree statistics were computed. The combination of the cavity method with block matrix models was also presented in Chapter 3, with applications demonstrated to random graphs with a community structure and sparse sample covariance matrices. Numerical and analytical examples were presented in all the above mentioned cases, along with discussion on relationship between graph structure and the features of the spectral density.

The results of Chapters 2 and 3 comprise a significant theoretical and practical advance, providing the opportunity to obtain qualitative information about the spectral density of sparse random matrix ensembles which could not realistically be obtained through direct calculation or simulation.

There are, of course, limitations in the method and its results. From the standpoint of interest in complex real-world networks, the chief drawback to an analysis based on the cavity method is the reliance upon the assumption that cycles in the underlying graph are long enough and rare enough to be discounted from the calculations. There is no evidence suggesting this simplifying assumption should hold for real-world graphs. Although the simple model of graphs with a community structure considered in Chapter 3 is capable of producing graphs with cycles of any length, it should be recognised that this is something of a trick, as the analysis remains dependent upon the graph describing inter-community connections being tree-like. Several possible methods of correcting the cavity method for the presence of cycles have been proposed recently in the statistical mechanics literature [MR05, PS06, CC06], and it is reasonable to hope that one or more of these might be usefully applied to the question

of spectral density.

The second major limitation of the techniques applied in these chapters is that they appear to be useful only in the consideration of the bulk of the spectral density, lacking the power necessary to probe details on smaller length scales. This is somewhat disheartening, as much important and interesting research in RMT is concerned with phenomena observable on smaller scales, such as repulsion between eigenvalues observed in level spacing and correlations. One possible analysis of the problem is that the approximation of the true graph by an infinite tree discussed in Section 2.4 induces a smoothing of the spectral density clearly visible on scales of order $1/N$, for example as seen in Fig 2.2. In another interpretation, we note that the cavity and replica methods are long known to produce identical results in the ensemble average, in which case the difficultly experienced here with smaller length scales may be thought of as a natural companion to the failings of the replica approach to spectral density in the same context [VZ85]. Whether or not modifications to the cavity method are possible to overcome this drawback remains an open question.

In the forth chapter, the cavity method was applied to sparse non-Hermitian random matrices. The electrostatic analogy and Hermitization techniques formed the starting point for the analysis, allowing the problem of determining the spectral density of non-Hermitian random matrices to be rephrased in a form to which the cavity method could be applied. Again, a simple collection of equations were derived whose solution provides a regularised approximation to the spectral density of a given matrix. These equations are exactly solvable in some cases, for example yielding Girko's Elliptic Law in the fully connected limit, as well as an apparently new result on the spectral density of random regular directed graphs. A little more work was necessary to implement numerical simulations, with a further set of consistency equations for the partial derivatives derived. The results of an efficient belief propagation algorithm were presented for two simple ensembles.

The results of Chapter 5 continued the theme of non-Hermitian RMT, though this time departing from the previous focus on sparseness to consider full random matrices with independent entries. Techniques analogous to the simple approach to Hermitian RMT were developed, based around a quaternionic generalisation of the Green's function. These methods were used to establish rigorous universal results for random matrices with independent entries (whether Hermitian or not) when summed or multiplied with deterministic matrices. Several examples were presented in which a mix of new and previously known results were derived. The Spherical Law was also presented in this chapter, conjecturing a new universal law for the spectral density of a certain class of random matrices.

Whilst the numerical and analytical examples given in Chapters 4 and 5 are just as successful as those of the earlier chapters in matching numerical experiments and previously known results, there is an important subtlety which, at present, prevents the establishment of rigorous results about the unregularised spectral density using these techniques. As discussed extensively at the start of Chapter 4, the regularisation of the spectral density exploited in both chapters is not as simple as that used in the Hermitian case, and the proposed exchange of the limits $N \to \infty$ and $\varepsilon \to 0$ is by no means easy to justify. In past work, this problem has been carefully circumvented by establishing results to bound the singular values of matrices of the type $(A - \lambda)$, where A is the matrix whose spectral density is under examination. Often, the techniques used follow the recent work of Rudelson (e.g. [Rud08]), and there exists similar work on products of matrices [Ver08] which gives hope for a proof that the sum and product rules for the quaternionic Green's given in Chapter 5 translate to results for the limiting spectral density.

As well as being interesting in its own right, the expression of the regularised spectral density in terms of a random perturbation given in Theorem 2 provides a believable (if not yet rigorous) argument justifying this problematic exchange of limits, since it is reasonable to expect that the spectral density of full random matrices ought to

be somehow continuous with respect to small random perturbations. In fact, this approach is not new, with similar ideas appearing the in work of Haagerup and others in free probability (see [Śni02] and references therein). It would be highly desirable to clarify the connections between this research and the techniques used in Chapters 4 and 5 of this thesis, perhaps providing a more straightforward and widely applicable approach to the problem of determining the limiting spectral density of ensembles of non-Hermitian random matrices. In another direction, the Spherical Law conjecture also provides an interesting direction of future research, a proof of which, by any method, would be a worthy goal.

A

Graph theory glossary

Let $G = (V, E)$ be a graph with adjacency matrix C. We define (in alphabetical order):

Automorphism A bijective function $f : V \to V$ is an automorphism if it satisfies $(f(i), f(j)) \in E \Leftrightarrow (i, j) \in E$.

Ball The ball of radius r around vertex i is the subgraph induced by those vertices of distance no greater than r from i.

Connected A graph is connected if there exists a path between any pair of vertices.

Cycle A cycle is a path of length at least 3, starting and ending on the same vertex.

Degree The number of vertices in the neighbourhood of a vertex $i \in V$ is known as the degree of i, denoted $k_i(C)$.

Degree sequence The N-vector of all degrees $\bar{k}(C) = (k_1(C), \ldots, k_N(C))$ is known as the degree sequence.

Distance The distance between vertices i and j is the length of the shortest path starting at i and ending at j.

Diameter The diameter of G is the maximum distance between vertices in V.

Induced Subgraph Given a subset $V' \subset V$, the induced subgraph of G on V' is the graph $G' = (V', E')$ where $(i, j) \in E' \Leftrightarrow (i, j) \in E$.

Neighbours The vertices $i, j \in V$, where $i \neq j$, are neighbours if and only if $(i, j) \in E$.

Neighbourhood The neighbourhood of the vertex $i \in V$ is the set of all its neighbours, we denote this set ∂i.

Path A path is a walk $(w_0, \ldots, w_n)$ such that $w_0, \ldots, w_{n-1}$ are distinct.

Regular A graph is said to be k-regular if each vertex has degree k

Transitive A graph is transitive if, for any pair of vertices i and j, there is an automorphism of G mapping i to j (in lay terms, all vertices are equivalent). Alternatively, the connectivity matrix C is preserved by any simultaneous permutation of rows and columns.

Tree If G is connected and no path in G is a cycle, then we say G is a tree.

Walk A walk $\boldsymbol{w} = (w_0, \ldots, w_n)$ of length $n \geq 0$ in a graph $G = (V, E)$ is an ordered collection of vertices $\{w_i\}_{i=0}^{n} \subset V$ such that for each $i = 1, \ldots, n$, there is an edge in G between vertices w_{i-1} and w_i.

B

Matrix identities

B.1 Green's function relations

Claim. *Let A be an $N \times N$ matrix with Green's function $G(z; A)$ and spectral density $\varrho(\mu; A)$, then for all z*

$$G(z; A) = \int \frac{1}{\mu - z} \varrho(\mu; A)\, d\mu\,.$$

Proof. Applying Schur's decomposition to the arbitrary matrix A, we obtain $A = UTU^{-1}$, where U is unitary and T upper triangular, with diagonal entries equal to the eigenvalues of A. From the definition of the Green's function we have

$$\begin{aligned} G(z; A) &= \frac{1}{N}\mathrm{Tr}\Big[(A - z)^{-1}\Big] \\ &= \frac{1}{N}\mathrm{Tr}\Big[\big(U(T - z)U^{-1}\big)^{-1}\Big] \\ &= \frac{1}{N}\mathrm{Tr}\Big[U(T - z)^{-1}U^{-1}\Big] \\ &= \frac{1}{N}\mathrm{Tr}(T - z)^{-1}\,, \end{aligned}$$

where the last line follows from the invariance of trace under similarity transformations. Next we note that, since $(T - z)$ is upper triangular, $(T - z)^{-1}_{ii} = (T_{ii} - z)^{-1}$,

and the claim follows straightforwardly:

$$\begin{aligned} G(z;A) &= \frac{1}{N}\mathrm{Tr}(T-z)^{-1} \\ &= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{T_{ii}-z} \\ &= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i^{(A)}-z} \\ &= \frac{1}{N}\sum_{i=1}^{N}\int\frac{1}{\mu-z}\delta\left(\lambda_i^{(A)}-\mu\right)\,d\mu \\ &= \int\frac{1}{\mu-z}\varrho(\mu;A)\,d\mu\,. \end{aligned}$$

□

Claim. *Let A be an $N\times N$ matrix with Green's function $G(z;A)$, then*

$$G(z;A) = \frac{1}{N}\frac{\partial}{\partial z}\log\frac{1}{\det(A-z)}\,.$$

Proof. From the general fact that the determinant of a matrix is given by the product of its eigenvalues, we see that $\det(A-z)=\prod_{i=1}^{N}(\lambda_i^{(A)}-z)$, and the result is immediate

$$\begin{aligned} \frac{1}{N}\frac{\partial}{\partial z}\log\frac{1}{\det(A-z)} &= -\frac{1}{N}\frac{\partial}{\partial z}\log\prod_{i=1}^{N}(\lambda_i^{(A)}-z) \\ &= -\frac{1}{N}\sum_{i=1}^{N}\frac{\partial}{\partial z}\log(\lambda_i^{(A)}-z) \\ &= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i^{(A)}-z} \\ &= G(z;A)\,. \end{aligned}$$

□

Claim. *Let A be an $N\times N$ matrix, and λ a complex variable, then*

$$\varrho(\lambda;A) = -\frac{1}{\pi}\frac{\partial}{\partial\overline{\lambda}}G(\lambda;A)\,.$$

Proof. Introduce the generalised function

$$D_\mu(\lambda) = -\frac{1}{\pi}\frac{\partial}{\partial\overline{\lambda}}(\mu-\lambda)^{-1}\,.$$

To prove the claim, we must identify $D_\mu(\lambda)$ as the Dirac delta $\delta(\mu-\lambda)$. It will suffice to show the following:

1. $D_\mu(\lambda) = 0$ for all $\lambda \neq \mu$
2. For any $a \in \mathbb{R}^+$, over the μ-centred square $S(a) = \{x+iy : x, y \in [\mu-a, \mu+a]\}$, we have
$$\int_{S(a)} D_\mu(\lambda)\, d\lambda = 1$$

For point 1, let $\lambda = x+iy$ and $\mu = u+iv$, assuming $\lambda \neq \mu$ the derivate may be explicitly calculated:

$$\begin{aligned}\frac{\partial}{\partial\overline{\lambda}}(\mu-\lambda)^{-1} &= \frac{1}{2}\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)\frac{1}{(u-x)+i(v-y)}\\ &= \frac{1}{2\big((u-x)+i(v-y)\big)^2}+i\left(\frac{i}{2\big((u-x)+i(v-y)\big)^2}\right)=0\,.\end{aligned}$$

For point 2, the change of variables $\tilde{x} = x-u$, $\tilde{y} = y-v$ allows us to compute

$$\begin{aligned}\int_{S(a)} D_\mu(\lambda)\, d\lambda &= -\frac{1}{2\pi}\int_{u-a}^{u+a}\int_{v-a}^{v+a}\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)\frac{1}{(u-x)+i(v-y)}\,dx\,dy\\ &= \frac{1}{2\pi}\int_{-a}^{a}\int_{-a}^{a}\left(\frac{\partial}{\partial\tilde{x}}+i\frac{\partial}{\partial\tilde{y}}\right)\frac{1}{\tilde{x}+i\tilde{y}}\,d\tilde{x}\,d\tilde{y}\\ &= \frac{1}{2\pi}\int_{-a}^{a}\left[\frac{1}{\tilde{x}+i\tilde{y}}\right]_{\tilde{x}=-a}^{\tilde{x}=a}d\tilde{y}+i\frac{1}{2\pi}\int_{-a}^{a}\left[\frac{1}{\tilde{x}+i\tilde{y}}\right]_{\tilde{y}=-a}^{\tilde{y}=a}d\tilde{x}\\ &= \frac{1}{\pi}\int_{-a}^{a}\frac{2a}{a^2+\tilde{x}^2}\,d\tilde{x}\\ &= 1\end{aligned}$$

□

B.2 Block matrix formulae

Let X be a block matrix consisting of four $N \times N$ blocks A, B, C, D arranged as follows:

$$\boldsymbol{X} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} .$$

The determinant and inverse of $\boldsymbol{X}$ may be written in terms of its blocks with the aid of the Schur complements

$$Y = A - BD^{-1}C\,, \quad \text{and} \quad Z = D - CA^{-1}B\,.$$

Specifically, we have

$$\det(X) = \det(D)\det(Y) = \det(A)\det(Z)\,,$$

and

$$\boldsymbol{X}^{-1} = \begin{pmatrix} Y^{-1} & -A^{-1}BZ^{-1} \\ -D^{-1}CY^{-1} & Z^{-1} \end{pmatrix} .$$

These, and a great many other matrix identities may be found in the Matrix Cookbook [PP08].

C

Quaternion algebra

C.1 Definitions

The real[1] quaternion algebra is a non-commutative extension of the complex numbers. Introduce basis elements i, j and k satisfying the relation

$$i^2 = j^2 = k^2 = ijk = -1\,. \tag{C.1}$$

A quaternion q is specified by a linear combination $q = \alpha + \beta i + \gamma j + \delta k$, where $\alpha, \beta, \gamma, \delta \in \mathbb{R}$. Note that from (C.1) we have $k = ij$, and thus a generic quaternion may just as well be specified by a pair of complex numbers $a, b \in \mathbb{C}$ by $q = a + bj$. It is this more compact representation that is used in the main body of the thesis.

The norm (or absolute value) of a quaternion $q = \alpha + \beta i + \gamma j + \delta k$ is defined to be

$$|q| = \sqrt{\alpha^2 + \beta^2 + \gamma^2 + \delta^2}\,.$$

Note (by setting $\gamma = 0$ and $\delta = 0$ in the above) that this norm agrees with the usual one for complex numbers and hence there is no confusion in using the same notation.

[1] Occasionally, quaternion algebras are defined over fields other than $\mathbb{R}$ and the name 'real quaternion' may be used to avoid confusion.

C.2 Matrix representation

Operations on quaternions have a close relation to those of matrices, in fact, there is an isomorphism between the algebra of quaternions and a certain group of 2×2 matrices. For the generic quaternion $q = a + ib$, where $a, b \in \mathbb{C}$ we introduce the matrix representation

$$M(q) = \begin{pmatrix} a & ib \\ i\overline{b} & \overline{a} \end{pmatrix} .$$

It is straightforward to see that the map M is an isomorphism. Moreover,

$$M(q)M(q)^{\dagger} = \begin{pmatrix} |q|^2 & 0 \\ 0 & |q|^2 \end{pmatrix} ,$$

and thus $|q| = \|M(q)\|$, where $\|\cdot\|$ denotes the spectral norm (that is, the largest singular value). The matrix representation defined here is introduced and used extensively in Chapter 5. There, for ease of notation, $M(q)$ is identified by $\boldsymbol{q}$.

C.3 Relation to Pauli matrices

The Pauli matrices are

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad \text{and} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{C.2}$$

Consider a generic quaternion $q = \alpha + \beta i + \gamma j + \delta k$, where $\alpha, \beta, \gamma, \delta \in \mathbb{R}$. The matrix representation of q may be written in terms of certain products of Pauli matrices:

$$M(q) = \begin{pmatrix} \alpha + i\beta & i(\gamma + i\delta) \\ i(\gamma - i\delta) & \alpha - i\beta \end{pmatrix} = \alpha I + \beta\sigma_x\sigma_y + \gamma\sigma_y\sigma_z + \delta\sigma_x\sigma_z \, .$$

Thus the algebra of quaternions is isomorphic to that generated by real linear combinations of the matrices $I, (\sigma_x\sigma_y), (\sigma_y\sigma_z)$ and $(\sigma_x\sigma_z)$.

Bibliography

[AB02] R. Albert and A.L. Barabási, *Statistical mechanics of complex networks*, Reviews of Modern Physics **74** (2002), 47.

[ACF+09] A. Annibale, A.C.C. Coolen, L.P. Fernandes, F. Fraternali, and J. Kleinjung, *Tailored graph ensembles as proxies or null models for real networks I: Tools for quantifying structure*, arXiv preprint 0908.1759 (2009).

[ACTA73] R. Abou-Chacra, D. J. Thouless, and P. W. Anderson, *A selfconsistent theory of localization*, Journal of Physics C **6** (1973), 1734.

[AF06] A. Alvermann and H. Fehske, *Stochastic Green's function approach to disordered systems*, Journal of Physics: Conference Series, vol. 35, Institute of Physics Publishing, 2006, p. 145.

[And58] P.W. Anderson, *Absence of diffusion in certain random lattices*, Physical Review **109** (1958), 1492.

[BA99] A.L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science **286** (1999), 509.

[Bai97] Z.D. Bai, *Circular law*, Annals of Probability **25** (1997), 494.

[Bai99] ———, *Methodologies in spectral analysis of large dimensional random matrices*, Statistica Sinica **9** (1999), 611.

[BAT04] F. X. Bronold, A. Alvermann, and F. Taylor, *Anderson localization in strongly coupled disordered electron-phonon systems*, Philosophical Magazine **84** (2004), 673.

[BCV08] G. Bianconi, A.C.C. Coolen, and C.J. Pérez Vicente, *Entropies of complex networks with hierarchically constrained topologies*, Physical Review E **78** (2008), 016114.

[BD06] J. Blitzstein and P. Diaconis, *A sequential importance sampling algorithm for generating random graphs with prescribed degrees*, Preprint (2006).

[BdMK01] A. Boutet de Monvel and A.M. Khorunzhy, *Some elementary results around the Wigner semicircle law*, 2001, (unpublished notes available online at http://www.physik.uni-bielefeld.de/bibos/preprints/01-03-035.pdf).

[BDML06] T. Britton, M. Deijfen, and A. Martin-Löf, *Generating simple random graphs with prescribed degree distribution*, Journal of Statistical Physics **124** (2006), 1377.

[BG01] M. Bauer and O. Golinelli, *Random incidence matrices: moments of the spectral density*, Journal of Statistical Physics **103** (2001), 301.

[BHZ95] E. Brézin, S. Hikami, and A. Zee, *Universal correlations for deterministic plus random Hamiltonians*, Physical Review E **51** (1995), 5442.

[BJK04] Z. Burda, J. Jurkiewicz, and A. Krzywicki, *Perturbing general uncorrelated networks*, Physical Review E **70** (2004), 026106.

[BK52] T.H. Berlin and M. Kac, *The spherical model of a ferromagnet*, Physical Review **86** (1952), 821.

[BKM$^+$00] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, *Graph structure in the web*, Computer Networks **33** (2000), 309.

[BL10] C. Bordenave and M. Lelarge, *Resolvent of large random graphs*, Random Structures and Algorithms (2010).

[BM99] G. Biroli and R. Monasson, *A single defect approximation for localized states on random lattices*, Journal of Physics A **32** (1999), L255.

[BMJ07] Z.D. Bai, B. Miao, and B. Jin, *On limit theorem for the eigenvalues of product of two random matrices*, Journal of Multivariate Analysis **98** (2007), 76.

[Bol01] B. Bollobás, *Random graphs*, Cambridge University Press, 2001.

[CBPS05] M. Catanzaro, M. Boguñá, and R. Pastor-Satorras, *Generation of uncorrelated random scale-free networks*, Physical Review E **71** (2005), 27103.

[CC06] M. Chertkov and V. Y. Chernyak, *Loop series for discrete statistical models on graphs*, Journal of Statistical Mechanics **2006** (2006), P06009.

[CGMM$^+$05] S. Ciliberti, T. S. Grigera, V. Martín-Mayor, G. Parisi, and P. Verrocchio, *Anderson localization in Euclidean random matrices*, Physical Review B **71** (2005), 153104.

[CH06] P. Carmona and Y. Hu, *Universality in Sherrington-Kirkpatrick's spin glass model*, Annales de l'Institut Henry Poincare **42** (2006), 215.

[Cha10] D. Chafai, *Circular law for non-central random matrices*, Journal of Theoretical Probability (2010).

[Chu97] F.R.K. Chung, *Spectral graph theory*, American Mathematical Society, 1997.

[CKS05] A.C.C. Coolen, R. Kühn, and P. Sollich, *Theory of neural information processing systems*, Oxford University Press, USA, 2005.

[CN85] JE Cohen and CM Newman, *A stochastic theory of community food webs i: Models and aggregated data*, Proceedings of the Royal Society of London **224** (1985), 421.

[Coo07] A.C.C. Coolen, *The mathematical theory of minority games: statistical mechanics of interacting agents*, Oxford University Press, USA, 2007.

[Dea02] D.S. Dean, *An approximation scheme for the density of states of the Laplacian on random graphs*, Journal of Physics A **35** (2002), L153.

[DGMS03] S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes, and A.N. Samukhin, *Spectra of complex networks*, Physical Review E **68** (2003), 046109.

[DGMS04] S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes, and A.N. Samukhin, *Random networks: eigenvalue spectra*, Physica A **338** (2004), 76.

[DH09] EB Davies and M. Hager, *Perturbations of Jordan matrices*, Journal of Approximation Theory **156** (2009), 82.

[DM02] S.N. Dorogovtsev and J.F.F. Mendes, *Evolution of networks*, Advances in Physics **51** (2002), 1079.

[DM08] A. Dembo and A. Montanari, *Ising models on locally tree-like graphs*, arXiv preprint 0804.4726 (2008).

[Dys62a] F.J. Dyson, *Statistical theory of the energy levels of complex systems I*, Journal of Mathematical Physics **3** (1962), 140.

[Dys62b] F.J. Dyson, *The threefold way: Algebraic structure of symmetry groups and ensembles in quantum mechanics*, Journal of Mathematical Physics **3** (1962), 1199.

[DZ96] J. D'Anna and A. Zee, *Correlations between eigenvalues of large random matrices with independent entries*, Physical Review E **53** (1996), 1399.

[EA75] S.F. Edwards and P.W. Anderson, *Theory of spin glasses*, Journal of Physics F **5** (1975), 965.

[Ede88] A. Edelman, *Eigenvalues and condition numbers of random matrices*, SIAM Journal of Matrix Analysis A **9** (1988), 543.

[Efe06] K.B. Efetov, *Random matrices and supersymmetry in disordered systems*, arXiv preprint cond-mat/0502322v1 (2006).

[EJ76] S.F. Edwards and R.C. Jones, *The eigenvalue spectrum of a large symmetric random matrix*, Journal of Physics A **9** (1976), 1595.

[EK09] G. Ergün and R. Kühn, *Spectra of modular random graphs*, Journal of Physics A **42** (2009), 395001.

[EN05] A. Edelman and Rao N.R., *Random matrix theory*, Acta Numerica **14** (2005), 233.

[ER59] P. Erdős and A. Rényi, *On random graphs*, Publicationes Mathematicae Debrecen **6** (1959), 156.

[FDBV01] I.J. Farkas, I. Derényi, A.L. Barabási, and T. Vicsek, *Spectra of 'real-world' graphs: Beyond the semicircle law*, Physical Review E **64** (2001), 26704.

[FFF99] M. Faloutsos, P. Faloutsos, and C. Faloutsos, *On power-law relationships of the internet topology*, Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication, ACM, 1999, p. 262.

[FHS07] R. Froese, D. Hasler, and W. Spitzer, *Absolutely continuous spectrum for the Anderson model on a tree: A geometric proof of Klein's theorem*, Communications in Mathematical Physics **269** (2007), 239.

[FKS97] Y.V. Fyodorov, B.A. Khoruzhenko, and H.J. Sommers, *Almost-Hermitian random matrices: eigenvalue density in the complex plane*, Physics Letters A **226** (1997), 46.

[FM91] Y.V. Fyodorov and A.D. Mirlin, *On the density of states of sparse random matrices*, Journal of Physics A **24** (1991), 2219.

[FM09] P.J. Forrester and A. Mays, *Pfaffian point process for the Gaussian real generalised eigenvalue problem*, arXiv preprint 0910.2531 (2009).

[FS95] Y.V. Fyodorov and H.J. Sommers, *Universality of "level curvature" distribution for large random matrices: systematic analytical approaches*, Zeitschrift fur Physik B **99** (1995), 123.

[FS03] Y.V. Fyodorov and H.J. Sommers, *Random matrices close to Hermitian or unitary: overview of methods and results*, Journal of Physics A **36** (2003), 3303.

[FSV03] P.J. Forrester, N.C. Snaith, and J.J.M. Verbaarschot, *Developments in random matrix theory*, Journal of Physics A **36** (2003), R1.

[FSZ01] J. Feinberg, R. Scalettar, and A. Zee, *"Single ring theorem" and the disk-annulus phase transition*, Journal of Mathematical Physics **42** (2001), 5718.

[FZ97a] J. Feinberg and A. Zee, *Non-Gaussian non-Hermitean random matrix theory: phase transitions and addition formalism*, Nuclear Physics B **501** (1997), 643.

[FZ97b] ———, *Non-Hermitean random matrix theory: method of Hermitization*, Nuclear Physics B **504** (1997), 579.

[Gil59] E.N Gilbert, *Random graphs*, Annals of Mathematical Statistics **30** (1959), 1141.

[Gin65] J. Ginibre, *Statistical ensembles of complex, quaternion, and real matrices*, Journal of Mathematical Physics **6** (1965), 440.

[Gir90] V. L. Girko, *Theory of random determinants*, Kulwer Academic, 1990.

[GJ07] A.T. Goerlich and A. Jarosz, *Addition of free unitary random matrices*, arXiv preprint math-ph/0408019 (2007).

[GK03] I.R. Graham and G. Kohr, *Geometric function theory in one and higher dimensions*, CRC, 2003.

[GM88] C.D. Godsil and B. Mohar, *Walk generating functions and spectral measures of infinite graphs*, Linear Algebra and its Applications **107** (1988), 191.

[GMGW98] T. Guhr, A. Müller-Groeling, and H.A. Weidenmüller, *Random matrix theories in quantum physics: Common concepts*, Physics Reports **299** (1998), 190.

[GNJJN03] E. Gudowska-Nowak, R.A. Janik, J. Jurkiewicz, and M.A. Nowak, *Infinite products of large random matrices and matrix-valued diffusion*, Nuclear Physics B **670** (2003), 479.

[Gol03] O. Golinelli, *Statistics of delta peaks in the spectral density of large random trees*, arXiv preprint cond-mat/0301437 (2003).

[GT07a] F. Gotze and A. Tikhomirov, *The circular law for random matrices*, arXiv preprint 0709.3995 (2007).

[GT07b] ———, *On the circular law*, arXiv preprint 0702386 (2007).

[HN96] N. Hatano and D.R. Nelson, *Localization transitions in non-Hermitian quantum mechanics*, Physical Review Letters **77** (1996), 570.

[HOV97] M.A. Halasz, J.C. Osborn, and J.J.M. Verbaarschot, *Random matrix triality at nonzero chemical potential*, Physical Review D **56** (1997), 7059.

[JN04] A. Jarosz and M.A. Nowak, *A novel approach to non-Hermitian random matrix models*, arXiv preprint math-ph/0402057 (2004).

[JN06] ———, *Random Hermitian versus random non-Hermitian operators - unexpected links*, Journal of Physics A **39** (2006), 10107.

[JNP+97] R.A. Janik, M.A. Nowak, G. Papp, J. Wambach, and I. Zahed, *Non-Hermitian random matrix models: Free random variable approach*, Physical Review E **55** (1997), 4100.

[JNPZ97] R.A. Janik, M.A. Nowak, G. Papp, and I. Zahed, *New developments in non-Hermitian random matrix models*, New Developments in Quantum Field Theory **366** (1997), 297.

[Kan01] E. Kanzieper, *Random matrices and the replica method*, Nuclear Physics B **596** (2001).

[KHKZ95] T.S. Kobayakawa, Y. Hatsugai, M. Kohmoto, and A. Zee, *Universal behavior of correlations between eigenvalues of random matrices*, Physical Review E **51** (1995), 5365.

[Kho96] B. Khoruzhenko, *Large-N eigenvalue distribution of randomly perturbed asymmetric matrices*, Journal of Physics A **29** (1996), L165.

[KKP96] A.M. Khorunzhy, B.A. Khoruzhenko, and L.A. Pastur, *Asymptotic properties of large random matrices with independent entries*, Journal of Mathematical Physics **37** (1996), 5033.

[KMP86] A. Klein, F. Martinelli, and J.F. Perez, *A rigorous replica trick approach to Anderson localization in one dimension*, Communications in Mathematical Physics **106** (1986), 623.

[KP93] A.M. Khorunzhy and L.A. Pastur, *Limits of infinite interaction radius, dimensionality and the number of components for random operators with off-diagonal randomness*, Communications in Mathematical Physics **153** (1993), 605.

[Kri09] M. Krishnapur, *From random matrices to random analytic functions*, Annals of Probability **37** (2009), 314.

[KS03] J.P. Keating and N.C. Snaith, *Random matrices and L-functions*, Journal of Physics A **36** (2003), 2859.

[KSV04] A.M. Khorunzhy, M. Shcherbina, and M. Vengerovsky, *Eigenvalue distribution of large weighted random graphs*, Journal of Mathematical Physics **45** (2004), 1648.

[Küh08] R. Kühn, *Spectra of sparse random matrices*, Journal of Physics A **41** (2008), 295002.

[McK81] B.D. McKay, *Expected eigenvalue distribution of a large regular graph*, Linear Algebra and its Applications **40** (1981), 203.

[Meh91] M.L. Mehta, *Random matrices*, New York: Academic, 1991.

[MF91] A.D. Mirlin and Y.V. Fyodorov, *Universality of level correlation function of sparse random matrices*, Journal of Physics A **24** (1991), 2273.

[Mil67] S. Milgram, *The small world problem*, Psychology Today **2** (1967), 60.

[Mir00] A.D. Mirlin, *Statistics of energy levels and eigenfunctions in disordered and chaotic systems: Supersymmetry approach*, arXiv preprint cond-mat/0006421 (2000).

[MJW06] D.M. Malioutov, J.K. Johnson, and A.S. Willsky, *Walk-sums and belief propagation in Gaussian graphical models*, Journal of Machine Learning Research **7** (2006), 2031.

[Mon73] H.L. Montgomery, *The pair correlation of the zeros of the Riemann zeta function*, Proceedings of Symposia in Pure Mathematics **24** (1973), 181.

[MP67] V.A. Marčenko and L.A. Pastur, *Distribution of eigenvalues for some sets of random matrices*, Sbornik: Mathematics **1** (1967), 457.

[MP02] M. Mihail and C. Papadimitriou, *On the eigenvalue power law*, Lecture Notes in Computer Science **254** (2002), 2483.

[MPV86] M. Mezard, G. Parisi, and M. Virasoro, *SK model: The replica solution without replicas*, Europhysics Letters **2** (1986), 77.

[MPV87] M. Mezard, G. Parisi, and M. Virasoro, *Spin glass theory and beyond*, World Scientific Publishing Company, 1987.

[MR95] M. Molloy and B.A. Reed, *A critical point for random graphs with a given degree sequence*, Random Structures and Algorithms **6** (1995), 161.

[MR05] A. Montanari and T. Rizzo, *How to compute loop corrections to the Bethe approximation*, Journal of Statistical Mechanics **10** (2005), 10011.

[MW90] B.D. McKay and N.C. Wormald, *Uniform generation of random regular graphs of moderate degree*, Journal of Algorithms **11** (1990), 52.

[New03] M.E.J. Newman, *The structure and function of complex networks*, arXiv preprint cond-mat/0303516 (2003).

[NR08] T. Nagao and G.J. Rodgers, *Spectral density of complex networks with a finite mean degree*, Journal of Physics A **41** (2008), 265002.

[NS97] A. Nagurney and S. Siokos, *Financial networks: Statics and dynamics*, Springer, 1997.

[NT07] T. Nagao and T. Tanaka, *Spectral density of sparse sample covariance matrices*, Journal of Physics A **40** (2007), 4973.

[Pas72] L. Pastur, *On the spectrum of random matrices*, Theoretical and Mathematical Physics **10** (1972), 67.

[Pas98] ———, *A simple approach to global regime of random matrix theory*, Mathematical Results in Statistical Mechanics, Marseilles (1998), 429.

[PBL05] M. Potters, J.P. Bouchaud, and L. Laloux, *Financial applications of random matrix theory: old laces and new pieces*, arXiv preprint physics/0507111 (2005).

[PP08] K.B. Petersen and M.S. Pedersen, *The matrix cookbook*, 2008.

[PS06] G. Parisi and F. Slanina, *Loop expansion around the Bethe-Peierls approximation for lattice models*, Journal of Statistical Mechanics **2006** (2006), L02003.

[PVSW05] A.P. Potapov, N. Voss, N. Sasse, and E. Wingender, *Topology of mammalian transcription networks*, Genome Informatics Series **16** (2005), 270.

[PZ07] G. Pan and W. Zhou, *Circular law, extreme singular values and potential theory*, arXiv preprint 0705.3773 (2007).

[RB88] G.J. Rodgers and A.J. Bray, *Density of states of a sparse random matrix*, Physical Review B **37** (1988), 3557.

[RD90] G.J. Rodgers and C. De Dominicis, *Density of states of sparse random matrices*, Journal of Physics A **23** (1990), 1567.

[Rog09] T. Rogers, *Universal sum and product rules for random matrices*, arXiv preprint 0912.2499 (2009).

[RP09] T. Rogers and I. Pérez Castillo, *Cavity approach to the spectral density of non-Hermitian sparse matrices*, Physical Review E **79** (2009), 012101.

[RPKT08] T. Rogers, I. Pérez Castillo, R. Kühn, and K. Takeda, *Cavity approach to the spectral density of sparse symmetric random matrices*, Physical Review E **78** (2008), 031116.

[RPVT10] T. Rogers, I. Pérez Castillo, C. Pérez Vicente, and K. Takeda, *Spectral density of random graphs with topological constraints*, Journal of Physics A **43** (2010), 195002.

[Rud08] M. Rudelson, *Invertibility of random matrices: norm of the inverse*, Annals of Mathematics **168** (2008), 575.

[SC02] G. Semerjian and L.F. Cugliandolo, *Sparse random matrices: the eigenvalue spectrum revisited*, Journal of Physics A **35** (2002), 4837.

[Sco88] J. Scott, *Social network analysis*, Sociology **22** (1988), 109.

[SCSS88] H.J. Sommers, A. Crisanti, H. Sompolinsky, and Y. Stein, *Spectrum of large random asymmetric matrices*, Physical Review Letters **60** (1988), 1895.

[SCT96] D.A. Stariolo, E.M.F. Curado, and F.A. Tamarit, *Distribution of eigenvalues of ensembles of asymmetrically diluted hopfield matrices*, Journal of Physics A **29** (1996), 4733.

[Śni02] P. Śniady, *Random regularization of Brown spectral measure*, Journal of Functional Analysis **193** (2002), 291.

[SPH05] N.S. Skantzos, I. Pérez Castillo, and J.P.L. Hatchett, *Cavity approach for real variables on diluted graphs and application to synchronization in small-world lattices*, Physical Review E **72** (2005), 66127.

[ST05] M. Steyvers and J.B. Tenenbaum, *The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth*, Cognitive Science **29** (2005), 41.

[Ste96] M.A. Stephanov, *Random matrix model of QCD at finite density and the nature of the quenched limit*, Physical Review Letters **76** (1996), 4472.

[SW99] A. Steger and N.C. Wormald, *Generating random regular graphs quickly*, Combinatorics, Probability and Computing **8** (1999), 377.

[Tal98] M. Talagrand, *Rigorous results for the Hopfield model with many patterns*, Probability Theory and Related Fields **110** (1998), 177.

[TV08] T. Tao and V. Vu, *Random matrices: The circular law*, Communications in Contemporary Mathematics **10** (2008), 261.

[TV09] T. Tao and V. Vu, *From the Littlewood-Offord problem to the circular law: Universality of the spectral distribution of random matrices*, Bulletin of the American Mathematical Society **46** (2009), 377.

[VC08] C.J. Pérez Vicente and A.C.C. Coolen, *Spin models on random graphs with controlled topologies beyond degree constraints*, Journal of Physics A **41** (2008), 255003.

[Ver08] T. Vershynin, *Spectral norm of products of random and deterministic matrices*, arXiv preprint 0812.2432 (2008).

[VM94] V.S. Viswanath and G. Müller, *The recursion method: Application to many-body dynamics*, Springer, 1994.

[VZ85] J.J.M. Verbaarschot and M.R. Zirnbauer, *Critique of the replica trick*, Journal of Physics A **18** (1985), 1093.

[Wig58] E.P. Wigner, *On the distribution of the roots of certain symmetric matrices*, Annals of Mathematics **67** (1958), 325.

[Wis28] J. Wishart, *The generalised product moment distribution in samples from a normal multivariate population*, Biometrika **20** (1928), 32.

[Wor99] N.C. Wormald, *Models of random regular graphs*, Surveys in combinatorics **276** (1999), 239.

[WTD+08] J. Wu, Y. Tan, H. Deng, Y. Li, B. Liu, and X. Lv, *Spectral measure of robustness in complex networks*, arXiv preprint 0802.2564 (2008).

[YFW03] J.S. Yedidia, W.T. Freeman, and Y. Weiss, *Understanding belief propagation and its generalizations*, Exploring artificial intelligence in the new millennium **8** (2003), 236.

[Zir99] M.R. Zirnbauer, *Another critique of the replica trick*, arXiv preprint cond-mat/9903338 (1999).

[Zir10] M.R. Zirnbauer, *Symmetry classes*, arXiv preprint 1001.0722 (2010).