Title
Outlier eigenvalue fluctuations of perturbed iid matrices

Author
Rajagopalan, Anand

Publication Date
2015

University of California
Los Angeles

Outlier Eigenvalue Fluctuations of Perturbed IID Matrices

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

by

Anand Bharathwaj Rajagopalan

2015

Abstract of the Dissertation

Outlier Eigenvalue Fluctuations of Perturbed IID Matrices

by

Anand Bharathwaj Rajagopalan

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 2015

Professor Terence Tao, Chair

It is known that in various random matrix models, large perturbations create outlier eigenvalues which lie, asymptotically, in the complement of the support of the limiting spectral density. This thesis studies fluctuations of these outlier eigenvalues of iid matrices X_n under bounded rank and bounded operator norm perturbations A_n, namely the fluctuations $\lambda(\frac{X_n}{\sqrt{n}} + A_n) - \lambda(A_n)$. The perturbations A_n that we consider belong to a large class, where we allow for arbitrary Jordan types and almost minimal assumptions on the left and right eigenvectors. We obtain the joint convergence of the normalized asymptotic fluctuations of the outlier eigenvalues in this setting with a unified approach.

The dissertation of Anand Bharathwaj Rajagopalan is approved.

Rowan Killip

Dimitri Shlyakhtenko

Sudip Chakravarty

Terence Tao, Committee Chair

University of California, Los Angeles

2015

Table of Contents

List of Figures

Acknowledgments

First of all, I thank my advisor, Terence Tao, for his constant guidance, support, and feedback throughout the research process, and for suggesting my thesis topic. I also thank David Renfrew for the many helpful discussions.

I thank Maggie Albert, Martha Contreras, and Maida Bassili for their help with all matters administrative.

I thank my friends and housemates here who have contributed to an enjoyable experience, in particular I thank Ashay Burungale, Sam Miner, Pietro Carolino and Shagnik Das.

Finally, I thank my parents, Lakshmi Venkatesan and Raghu Rajagopalan, and my sister, Daksha Rajagopalan, for their love and support.

Vita

2009	B.S. (Mathematics) and B.S. (Physics), MIT, Cambridge, Massachusetts.
2011	M.A. (Mathematics), UCLA, Los Angeles, California.

CHAPTER 1

Introduction

1.1 Background

In this thesis, we study the fluctuations of outlier eigenvalues of additively perturbed iid matrices for a general class of perturbations. We begin by describing some models of random matrices, all of which are characterized by independence of their entries (up to certain restrictions).

One of the earliest studied random matrices models was introduced by John Wishart ([Wis28]) in the field of multivariate statistics. We define the model in a specialized setting, in order to simplify the presentation. Let $Y = (Y_n)_{n\geq 1}$ be an $m \times n$ matrix with $m/n \to \alpha \in (0, 1]$ such that the columns of Y_n are iid centered Gaussians with covariance matrix $\Sigma_p = I_p$. Letting

$$S_n := \frac{1}{n} Y_n Y_n^T, \tag{1.1}$$

we refer to $S := (S_n)_{n\geq 1}$ as a *Wishart ensemble.*

One of the main focuses of random matrix theory is to understand the eigenvalues (and singular values) of the various models, in the limit as the size of the matrix tends to infinity. Given a square random matrix X, let $\Lambda = \Lambda(X)$ denote the set of eigenvalues of X, or its *spectrum*. We then define the *empirical spectral measure(ESM)* of X_n by

$$\mu_{X_n} := \frac{1}{n} \sum_{\lambda \in \Lambda(X_n)} \delta_\lambda.$$

Thus the ESM is a random measure on the complex plane with delta masses (of size $1/n$) at the eigenvalues. One of the remarkable discoveries of random matrix theory is that for many ensembles, the ESM converges, under suitable normalization, to a *limiting spectral measure* and that this limiting measure is *universal* within a given class of matrices.

The first such result is due to Eugene Wigner ([Wig55]) for a Hermitian matrix model. We again define a specialized version of the ensemble. Let $W_n = (w_{ij})_{i,j\le n}$ be a sequence of symmetric matrices with the $\frac{n(n+1)}{2}$ real valued random variables $(\frac{1}{\sqrt{2}}w_{ii})_{i=1}^n$ and $(w_{ij})_{i\le j}$ being iid. Normalizing the *atom distribution* $w := w_{12}$ to have mean zero and variance 1, we refer to $W = (W_n)_{n\ge 1}$ as a *Wigner matrix*. The first proof of the semicircular law was for convergence in expectation; we state the almost sure version.

Theorem 1. *(Wigner Semicircular Law) Let W be a Wigner matrix. Then*

$$\mu_{\frac{W_n}{\sqrt{n}}} \Rightarrow \mu_{sc} \quad a.s.$$

with

$$d\mu_{sc}(x) := \frac{1}{2\pi}\sqrt{4-x^2}\,dx.$$

Thus the ESM of $W_n/\sqrt{n}$ converges in distribution to μ_{sc} almost surely. That is,

$$\mathbb{P}\left[\int f d\mu_{\frac{W_n}{\sqrt{n}}} \Rightarrow \int f d\mu_{\mathrm{sc}} \quad \text{for all } f \in \mathrm{BC}(\mathbb{R})\right] = 1,$$

where $\mathrm{BC}(\mathbb{R})$ denotes the set of bounded continuous functions on $\mathbb{R}$.

Finally, we define an *iid matrix*, which is the underlying model of our investigations.

Definition 1. *An* iid matrix *is a sequence $X = (X_n)_{n\ge 1}$ with X_n having (complex-valued) iid entries $(x_{i,j})_{1\le i,j\le n}$ and whose atom distribution $x = x_{1,1}$ satisfies $\mathbb{E}x = 0$ and $\mathbb{E}|x|^2 = 1$.*

We have, in analogy with the semicircular law, the circular law for iid matrices.

Theorem 2 (Circular law)**.** *For an iid matrix X, we have*

$$\mu_{\frac{X_n}{\sqrt{n}}} \Rightarrow \mu_C \ \ a.s.$$

with

$$d\mu_C := \frac{1}{\pi}\mathbb{1}_{\{z\in\mathbb{C}:|z|\le 1\}}dz.$$

Here dz denotes Lebesgue measure on the complex plane. The circular law has a long history and is the work of many authors (see [TV10] and references therein).

Once one has an understanding of the limiting spectrum, there are many other questions one can ask about the eigenvalues of a random matrix. One relatively recent line of research has focused on understanding the spectra of random matrix models under perturbations of various types. This is the context in which our study takes place and we now turn to exploring the relevant literature in this area.

The works of [BBP05] and [BS06b] were one of the first to study perturbed random matrix models. Consider the Wishart ensemble $S = (S_n)_{n\geq 1}$ as defined in (1.1). The limiting spectral measure is then given by the Marchenko-Pastur law (see [MP67]) $\mu_{\mathrm{MP}}(\alpha)$, defined by

$$d\mu_{\mathrm{MP}}(\alpha)(x) := \frac{1}{2\pi}\frac{\sqrt{(\alpha_+ - x)(x-\alpha_-)}}{\alpha x}\mathbb{1}_{[\alpha_-,\alpha_+]}dx$$

where $\alpha_\pm = (1 \pm \sqrt{\alpha})^2$.

Baik, Ben Arous and Péché consider a multiplicative perturbation of this model, with Σ_p having all but finitely many eigenvalues equal to 1. The resulting S_n (referred to as a spiked population model), while having the same limiting spectral measure, now has eigenvalues located outside $\mathrm{supp}(\mu_{\mathrm{MP}}(\alpha))$. These are the *outlier eigenvalues* of the perturbation. The central questions are then:

1. What are the asymptotic locations, θ_i of these outliers, λ_i?

2. What are their (normalized) *fluctuations*? Namely,

   (a) At what rate do we have the convergence $\lambda_i \to \theta_i$?

   (b) After normalizing, what is the limiting distribution of $\lambda_i - \theta_i$ (if it exists)?

In [BBP05], the limiting fluctuations of the outliers are shown to follow a *phase transition* depending on the largest eigenvalues of Σ_p; thus the limiting distribution changes as the largest eigenvalues of Σ_p change over a small scale.

We study the fluctuations of such outlier eigenvalues for the iid matrix ensemble, under a general class of perturbations (those having bounded rank and bounded operator norm).

Before presenting the details of this model and the outcome of our investigations, we first turn to the situation in the Wigner setting.

1.1.1 Outlier fluctuations of Wigner matrices

Following the work of [BBP05], the outlier eigenvalues of various perturbed matrix models have been studied. In this section, we review the literature of outlier eigenvalue fluctuations in the Hermitian setting. The works of [CDF09], [CDF12],[PRS13],[RS13],[KY13], and [KY14] build up to an essentially complete picture of the fluctuations for a certain class of perturbations.

We first state [CDF09, Theorem 2.2], which we have specialized to the case of real entries with unit variance for simplicity. Let W_n be a symmetri Wigner matrix as defined above, and let $A = A_n$ be the deterministic *perturbation matrix* given by $a_{ij} = \theta\delta_{(ij)=(11)}$. Let $M_n := \frac{W_n}{\sqrt{n}} + A_n$. It is known (see [Bai99]) that the limiting spectral measure of M_n is also given by the Wigner semicircle law with support $[-2, 2]$. Under the additional assumption that the atom distribution w is symmetric and satisfies a Poincaré inequality, Capitaine, Donati-Martin, and Féral show that M_n has an outlier eigenvalue λ_n which converges to $\theta + \frac{1}{\theta}$. Furthermore, they determine the limiting distribution of the normalized fluctuation of this outlier, namely

$$\sqrt{n}\left(1 - \frac{1}{\theta^2}\right)^{-1}\left(\lambda_n - \left(\theta + \frac{1}{\theta}\right)\right) \Rightarrow w + g \tag{1.2}$$

where g is a (real) centered Gaussian independent of w with variance

$$v_\theta = \frac{\mathbb{E}|w|^4 - 3}{\theta^2} + \frac{2}{\theta^2 - 1}.$$

We remark that the dependence of the limiting fluctuation on the law of the atom makes it *non-universal*. While this non-universality is also a feature of such limiting fluctuations in other models, it should be noted that in random matrix theory as a whole, non-universality is more an exception than the rule.

We next describe the generalizations of [CDF09, Theorem 2.2] in works that have followed. In [CDF12], the authors extend the result to finite rank matrices A_n of the same symmetry class as W_n, with eigenvectors belonging to two cases - local and delocal types, defined below. A *local* perturbation A_n is supported in a finite $K \times K$ box in the upper left corner. The fluctuations corresponding to an eigenvalue θ of multiplicity k are given by the eigenvalues of a $k \times k$ random matrix of the form $U^*(W_k + H_k)U$, where H_k is a Hermitian Gaussian matrix (with specified covariances) independent of W_k and U is a matrix of appropriate size that encodes the relevant eigenvectors of A. A *delocalized perturbation* matrix is one whose outlier eigenvectors u_i have uniformly small entries, thus $\max_i \|u_i\|_\infty \to 0$, where the maximum is over eigenvectors associated to the outlier eigenvalues. In this case, the authors shows that *universal* limiting fluctuations are obtained, which are eigenvalues of a random matrix of the form $U^* H_{k\times k} U$. Finally, we note that [CDF12] requires the technical condition that the eigenvectors of A_n corresponding to the outlier eigenvalues have supports of size $O(\sqrt{n})$.

In [PRS13] and [RS13], the assumptions that w is symmetric and satisfies a Poincaré inequality are removed. Instead, the authors show that it suffices to assume [1] $\mathbb{E}|w|^{4+\epsilon} < \infty$. Furthermore, the restriction on the support of the eigenvectors is also removed. In the delocalized case, the (suitably normalized) limiting fluctuations then become the eigenvalues of $U^*(M_3 + H)U$, where M_3 is a deterministic matrix that is zero on the diagonal and is a multiple of the third moment $\mathbb{E}[|w|^2 w]$ off the diagonal.

Finally, in [KY13] and [KY14], Knowles and Yin extend the model further in two directions. They allow for outliers θ_i that satisfy

$$|\theta_i - 1| \gtrapprox n^{-1/3},$$

and they do not require any hypothesis on the eigenvector of the perturbation. They require, however, that the atom distribution w has subexponential decay.

In the next section, we introduce prior results for iid matrices in the context of outliers which are relevant to our work. Then, in Section 1.3, we define the class of perturbations we study and state our theorem for the joint limiting distribution of the outlier fluctuations.

[1] In fact, they prove that it suffices to assume certain weaker Lindeberg conditions for the diagonal and off-diagonal entries

Finally, in Section 1.4, we look at some special cases of our results, and compare these with results from related works, namely from [RB13] and [BC14].

1.2 Outlier eigenvalues of iid matrices

Without perturbing an iid matrix X, one might still ask if there are outliers in the limit. It turns out that this is not the case, as we now proceed to show. Denoting the spectral radius of $X/\sqrt{n}$ by $\rho(X/\sqrt{n}) := \max_{\lambda \in \Lambda(X)} |\lambda|$, the circular law shows that $\limsup \rho(X/\sqrt{n}) \geq 1$ almost surely. The following is a complementary result (stated in [Tao13]) proven using the truncation method and the moment method (see e.g. [BY86]).

Theorem 3. *Let X_n be an iid matrix with atom distribution having bounded fourth moment. Then*

$$\rho\left(\frac{X}{\sqrt{n}}\right) = \lim_{l \to \infty} \left\| \left(\frac{X}{\sqrt{n}}\right)^l \right\|^{1/l}$$

converges to 1 almost surely as $n \to \infty$. Moreover, for $l \geq 1$, $\|(\frac{X}{\sqrt{n}})^l\|$ converges to $l+1$ almost surely as $n \to \infty$.

Now let $A = A_n$ be a deterministic matrix of rank $O(1)$ and operator norm $O(1)$. Next, let $\Theta = \Theta_n := \{\lambda \in \Lambda(A_n) : |\lambda| > 1\}$. In the interest of simplicity, we will state the following result assuming Θ is independent of n for n sufficiently large, though this can be relaxed. Finally, we let m_θ denote the multiplicity of θ. Then the following theorem due to [Tao13] (with generalizations to other models in [OR14], [RB13] and [BC14]) shows that outliers in the spectrum of $\frac{X}{\sqrt{n}} + A$ appear, in contrast to the situation in Theorem 3.

Theorem 4. *Let X be an iid matrix with bounded fourth moment and let A and Θ be as above. For each $\theta \in \Theta$ there exists*

$$\Lambda^\theta \subset \Lambda(\frac{X}{\sqrt{n}} + A)$$

with $|\Lambda^\theta| = m_\theta$ and for $\lambda \in \Lambda^\theta$,

$$\lambda \to \theta$$

almost surely.

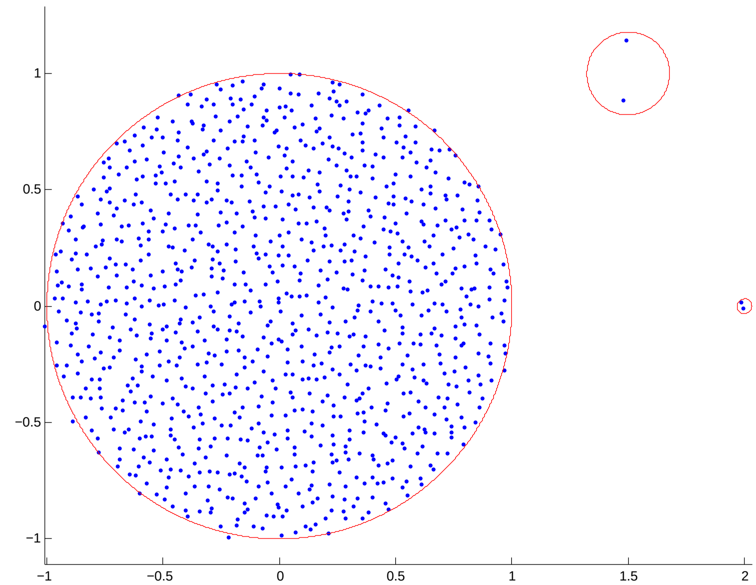


Figure 1.1: Eigenvalues of $X/\sqrt{n}+A$ with X having iid $\mathcal{N}(0,1)_{\mathbb{C}}$ entries, $A = 2I_2 \oplus J_{1.5+i,2} \oplus 0_{996}$ and $n = 1000$. The smaller circles have radii $n^{-1/2}$ and $n^{-1/4}$.

To illustrate Theorem 4, in Figure 1.1 we have plotted the eigenvalues of a perturbed Gaussian matrix $X/\sqrt{n} + A$, with x having distribution $\mathcal{N}(0,1)_{\mathbb{C}}$ and $n = 1000$. The two outliers near 2 correspond to the block $\left(\begin{smallmatrix} 2 & 0 \\ 0 & 2 \end{smallmatrix}\right)$ and the two outliers near $1.5+i$ are from the block $\left(\begin{smallmatrix} 1.5+i & 1 \\ 0 & 1.5+i \end{smallmatrix}\right)$ of A. Observe that the fluctuations from the Jordan block are larger; this phenomenon will be discussed later. It does not occur in the Wigner model with symmetric perturbations.

1.3 Model and statement of results

The objects we study are the fluctuations $\lambda - \theta$. More precisely, we obtain the joint limiting distribution of the (suitably normalized) fluctuations when A is allowed to have arbitrary Jordan type and left and right general eigenvectors satisfying a mild restriction. After introducing the main definition and theorem, we will discuss simpler special cases in Section 1.4.

We now define the perturbation matrices we will consider in this work, along with associated notation. To unify expressions involving mixed second moments, we will employ the following notational convention throughout this thesis. For any complex vector z, we let

$$z^{(d)} := \begin{cases} z & : d = 0 \\ \overline{z} & : d = 1 \end{cases} \tag{1.3}$$

where $\overline{z}$ denotes the (component-wise) conjugate of z. We will write z^T for the transpose of z and z^* for the conjugate transpose of z.

Definition 2. *A* perturbation matrix $A = (A_n)_{n\geq 1}$ *is a sequence of (complex) $n \times n$ matrices with rank $O(1)$ and operator norm $O(1)$. For $\theta \in \Theta = \{\theta \in \Lambda(A_n) : |\theta| > 1\}$, let J_θ be the Jordan block in the Jordan decomposition of A corresponding to θ with blocks written in nonincreasing order. We will assume that Θ and $(J_\theta)_{\theta\in\Theta}$ are independent of n for n sufficiently large. Let*

$$J_\theta = \bigoplus_{k=1}^{K_\theta} J_{\theta,k}^{m_{\theta,k}} \text{ where } J_{\theta,k} := \begin{pmatrix} \theta & 1 & & \\ & \theta & 1 & \\ & & \ddots & 1 \\ & & & \theta \end{pmatrix}$$

is the Jordan block of size k occuring with multiplicity $m_{\theta,k}$ in J_θ. To index the eigenvectors and generalized eigenvectors, we introduce the following notation. Let

$$I := \{(i, j, k, \theta) : i \in [k], j \in [m_{\theta,k}], k \in [K_\theta], \theta \in \Theta\}$$

and for $s \in I$, we write $s = (i_s, j_s, k_s, \theta_s)$. Let

$$I_\theta = \{s \in I : \theta_s = \theta\}.$$

For fixed j, k, θ, let $(v_s)_{i=1}^k$ be the generalized eigenvectors corresponding to the jth block of $J_{\theta,k}$, and let $v_{1,j,k}$ be the eigenvector for that block. Similarly define $(u_s^*)_{i=1}^k$ to be the generalized left eigenvectors with the $u_{(k,j,k)}^*$'s being the corresponding left eigenvectors. To index the left and right eigenvectors, we let

$$I_u^\theta := \{s \in I_\theta : i_s = k_s\}$$

and

$$I_v^\theta := \{t \in I_\theta : i_t = 1\}.$$

Finally, we let

$$I_2 := \bigcup_{\theta\in\Theta} I_u^\theta \times I_v^\theta \times \{\theta\}$$

and for $r \in I_2$, we write $r = (s_r, t_r, \theta_r)$. For $(s_i, t_i, \theta_i) \in I_2$, $i = 1, 2$, we assume that the limits of the following inner products exist and define, for $d_1, d_2 \in \{0, 1\}$, the scalars

$$U_{s_1,s_2}^{(d_1),(d_2)} := \lim_{n\to\infty} (u_{s_1})^{(d_1)*} \overline{(u_{s_2})}^{(d_2)}, \tag{1.4}$$

$$V_{t_1,t_2}^{(d_1),(d_2)} := \lim_{n\to\infty} (v_{t_1})^{(d_1)T} (v_{t_2})^{(d_2)}. \tag{1.5}$$

We also assume the following convergence and define $(G_r)_{r\in I_2}$ by

$$(u_{s_r}^* X v_{t_r})_{r\in I_2} \Rightarrow (G_r)_{r\in I_2}. \tag{1.6}$$

Lastly, we require the following technical assumption. Fix $\delta > 0$ and let

$$L = \bigcup_{r\in I_2} \{(i,j) \in [n]^2 : |u_{s_r,i} v_{t_r,j}| \geq n^{-1/4+\delta}\}.$$

Then we assume

$$\left(\sum_{(i,j)\in L} u_{s_r,i} x_{ij} v_{t_r,j} \right)_{r\in I_2} \Rightarrow (G_r^L)_{r\in I_2}. \tag{1.7}$$

Remark 1. *The eigenvectors satisfying the convergence criteria of (1.4)- (1.7) are quite general, and are allowed to be of* local, delocal *and* mixed *types (see Remark 4). These eigenvector requirements are similar to those of [KY13] and [KY14].*

We denote the Schur complement of A in the block matrix $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ by

$$\mathrm{SC}(A, \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)) := D - CA^{-1}B. \tag{1.8}$$

Recalling the notation of Theorem 4, we denote the elements of Λ^θ by λ_s^θ for $s \in I_\theta$. We now state our main theorem.

Theorem 5. *Let X be an iid matrix and A a perturbation matrix. We will assume the moment hypothesis $\mathbb{E}|x|^m < \infty$, with m defined as follows. First define $0 \leq c \leq 1$ through*

$$c = \sup\{c' \geq 0 : \max_{i\in[p]} \|u_i\|_\infty \|v_i\|_\infty \ll n^{-c'}\}. \tag{1.9}$$

Then fix $\epsilon > 0$ and set

$$m = \min(\max(2/c, 4), 8) + \epsilon. \tag{1.10}$$

Recalling (1.6), (1.4) and (1.5), we define the random variables $(F_r)_{r\in I_2}$ *by*

$$F_r := G_r + g_r, \tag{1.11}$$

where $(g_r)_{r\in I_2}$ *is a collection of centered complex Gaussians independent of* $(G_r)_{r\in I_2}$ *with mixed second moments specified by*

$$\mathbb{E}g_{r_1}^{(d_1)}g_{r_2}^{(d_2)} = \frac{(\mathbb{E}x^{(d_1)}x^{(d_2)})^2}{\theta_{r_1}\theta_{r_2} - \mathbb{E}x^{(d_1)}x^{(d_2)}} U_{s_{r_1},s_{r_2}}^{(d_1),(d_2)} V_{t_{r_1},t_{r_2}}^{(d_1),(d_2)}. \tag{1.12}$$

For $\theta \in \Theta$, *let* $F^\theta := (F_r)_{\theta_r=\theta}$ *be the* $I_u^\theta \times I_v^\theta$ *matrix of random variables and for* $k \in [K_\theta]$, *let*

$$F^{\theta,k} := \mathrm{SC}(F^\theta|_{\{(s,t):k_s,k_t\geq k+1\}}, F^\theta|_{\{(s,t):k_s,k_t\geq k\}})$$

be the $m_{\theta,k} \times m_{\theta,k}$ *matrix that is the Schur complement of the indicated submatrices of* F^θ. *Denote the eigenvalues of* $F^{\theta,k}$ *by* $(\tilde\lambda_{j,k}^\theta)_{j=1}^{m_{\theta,k}}$ *whose kth roots we denote*

$$\tilde f_{i,j,k}^\theta := (\zeta_k^i(\tilde\lambda_{j,k}^\theta)^{1/k})_{(i,j,k)\in I_\theta,\theta\in\Theta} \tag{1.13}$$

where $\zeta_k = e^{\frac{2\pi\sqrt{-1}}{k}}$. *Then for each* $\theta\in\Theta$, *we can label the eigenvalues in* Λ^θ *as* $(\lambda_{i,j,k}^\theta)_{(i,j,k)\in I_\theta}$ *such that the normalized outlier fluctuations*

$$f_{i,j,k}^\theta := n^{1/(2k)}\left(\lambda_{i,j,k}^\theta\left(\frac{X}{\sqrt{n}} + A\right) - \theta\right) \tag{1.14}$$

converge to $(\tilde f_{i,j,k}^\theta)_{\theta\in\Theta,(i,j,k)\in I_\theta}$ *in the following sense. Define the subgroup* S *of the permutation group* S_I *by*

$$\begin{aligned} S := &\{\pi \in S_I : \pi(s)_\theta = s_\theta, \pi(s)_k = s_k \text{ and} \\ &\pi(s)_j = \pi(t)_j \Leftrightarrow s_j = t_j \text{ for all } s,t\in I\}. \end{aligned}$$

Let $\mathrm{BC}(\mathbb{C}^I)^S$ *denote the set of bounded continuous functions on* $\mathbb{C}^I$ *invariant under the action of* S. *Then for* $f \in \mathrm{BC}(C^I)^S$, *and writing* $(\tilde f_l)_{l\in I}$ *for (1.13) and* $(f_l)_{l\in I}$ *for (1.14),*

$$\int f d\mu_{(f_l)_{l\in I}} \to \int f d\mu_{(\tilde f_l)_{l\in I}}.$$

Remark 2. *The moment hypothesis we require seems to be a technical limitation of the moment method that we have employed. While we need at most* $8+\epsilon$ *moments in all cases, we conjecture that* 4 *moments always suffice. In the* delocal *case with* $c = 1$ *(i.e.,* $\|u_i\|_\infty, \|v_i\|_\infty \ll 1/\sqrt{n}$, *we require* $4+\epsilon$ *moments which almost matches the conjectured optimal. On the other hand, under the assumption of* 4 *moments, [BC14] obtains the fluctuations of certain types of* local *matrices (with* $c = 0$*) as described in the next section.*

1.4 Discussion and related works

We now provide examples of different types of behavior for the fluctuations that illustrate Theorem 5. The first two examples are of rank 1 fluctuations.

(i) If A has a single non zero entry θ in the top left with $|\theta| > 1$, the limiting normalized fluctuation of the outlier is the law of $x + g$ where x is the atom distribution and g is a centered complex Gaussian with $\mathbb{E}g^2 = 0$ and $\mathbb{E}|g|^2 = \mathcal{N}(0, \frac{1}{|\theta|^2-1})_{\mathbb{C}}$. Note the non universality of the fluctuations, and their similarity to (1.2) from the Wigner case.

In Figure 1.2, we demonstrate this non-universality. Figures (a) and (c) are 500 samples of the normalized fluctuations $\sqrt{n}(\lambda_{out}(\frac{X}{\sqrt{n}} + A) - 2)$ of a single outlier with $n = 100$ and A given by $a_{i,j} = 2\delta_{(i,j)=(1,1)}$. In Figure (a), the atom distribution x is distributed uniformly over the square $[-l, l]^2 \subset \mathbb{C}$ with $l = \sqrt{3/2}$ so that $\mathbb{E}|x|^2 = 1$ (outlined in figure). In Figure (c), x is the standard complex normal $\mathcal{N}(0, 1)_{\mathbb{C}}$. Figures (b) and (d) are 500 samples from the corresponding limiting distributions as predicted by Theorem 5.

(ii) If $A = \theta v u^*$ is of rank 1 with $|\theta| > 1$ and $\|u\|_\infty \|v\|_\infty = o(1)$, then the normalized fluctuation $\sqrt{n}(\lambda - \theta)$ converges to the law of a centered complex Gaussian g_θ with

$$\mathbb{E}g_\theta^2 = \frac{|\theta|^2 \mathbb{E}x^2}{|\theta|^2 - \mathbb{E}x^2} \lim_{n\to\infty} u^* \overline{u} v^T v$$

and

$$\mathbb{E}|g_\theta|^2 = \frac{|\theta|^2}{|\theta|^2 - 1} \lim_{n\to\infty} u^* u v^* v.$$

In particular, if $\mathbb{E}x^2 = 0$ and A is normal (thus u and v are unit vectors), then g_θ is a circularly symmetric Gaussian with variance $\frac{|\theta|^2}{|\theta|^2-1}$.

(iii) Suppose $A = UDU^*$ is normal of rank k, with $\|u_i\|_\infty = o(1)$ for $i = 1, 2, \ldots, k$. For a fixed eigenvalue $\theta \in \Theta$ of multiplicity m, the covariance formula (1.12) reduces to

$$\mathbb{E}g_{ab}g_{cd} = \frac{\theta^2 \mathbb{E}x^2}{\theta^2 - \mathbb{E}x^2} \lim_{n\to\infty} u_a^* \overline{u_c} u_b^T u_d$$

and

$$\mathbb{E}g_{ab}\overline{g_{cd}} = \frac{\theta^2}{\theta^2-1}\lim_{n\to\infty} u_a^* u_c u_b^* u_c$$
$$= \frac{\theta^2}{\theta^2-1}\delta_{ac}\delta_{bd}$$

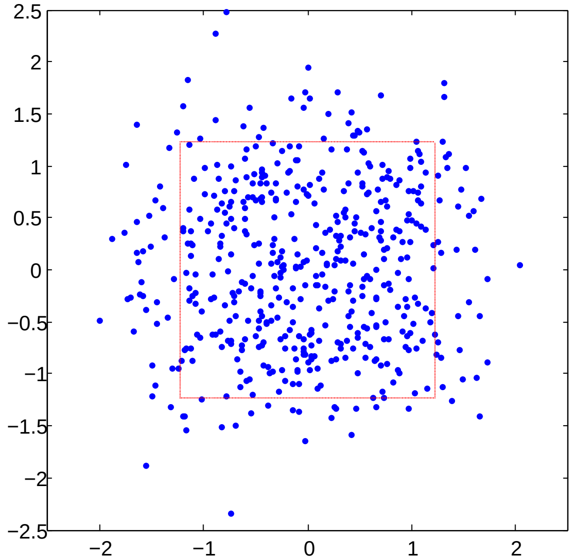
Note that fluctuations of different eigenvalues are still correlated in general. We obtain asymptotically independent fluctuations for distinct eigenvalues in the following cases.

(a) If A is real, $u_a^*\overline{u_c} = \delta_{ac}$, $u_b^T u_d = \delta_{bd}$ and the entries of $F^\theta = (g_{ab})_{a,b=1}^m$ are independent Gaussians. Depending on the Jordan structure J_θ, the normalized fluctuations converge to the appropriate roots of eigenvalues of Schur complements of submatrices of F^θ as specified in Theorem 5.

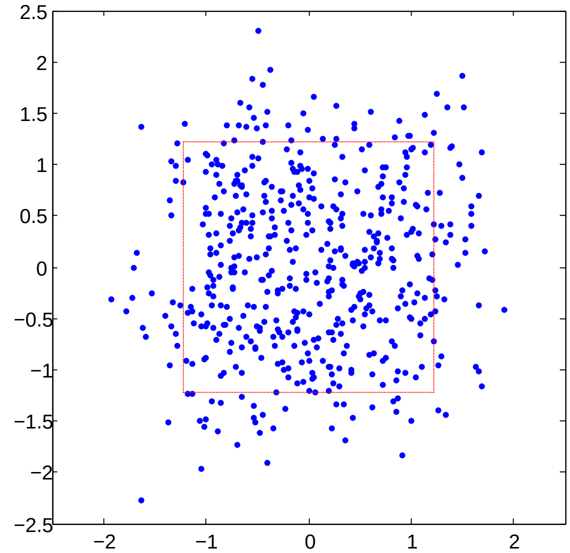
(b) If $\mathbb{E}x^2 = 0$, $F^\theta = (g_{ab})_{a,b=1}^m$ is a scaled complex Ginibre ensemble with atom distribution g satisfying $\mathbb{E}g = 0$, $\mathbb{E}g^2 = 0$ and $\mathbb{E}|g|^2 = \frac{|\theta^2|}{|\theta^2|-1}$. If we now suppose further that $J_\theta = \theta I_m$, then the m fluctuations associated to θ are given by the eigenvalues of the complex Ginibre ensemble specified above. By the circular law, they lie approximately uniformly in a disk of radius $\frac{|\theta|}{(|\theta|^2-1)^{\frac{1}{2}}}$ for large m.

(c) So far, the fluctuations have been of order $O(\frac{1}{\sqrt{n}})$. Suppose again that $\mathbb{E}x^2 = 0$ but that J_θ is a single Jordan block of size m. Then as remarked below Proposition 2, the m fluctuations scaled by $n^{1/(2m)}$ are given by $(e^{2\pi i j/m} g_\theta^{1/m})_{j=0}^{m-1}$ where $g_\theta = (F^\theta)_{m1}$ is the lower left entry of F^θ. Hence the fluctuations are distributed uniformly around a circle of radius $n^{-1/(2r)} g_\theta^{1/m}$. This dependence of the rate of convergence on the size of the Jordan block is illustrated by the outliers in Figure 1.1.

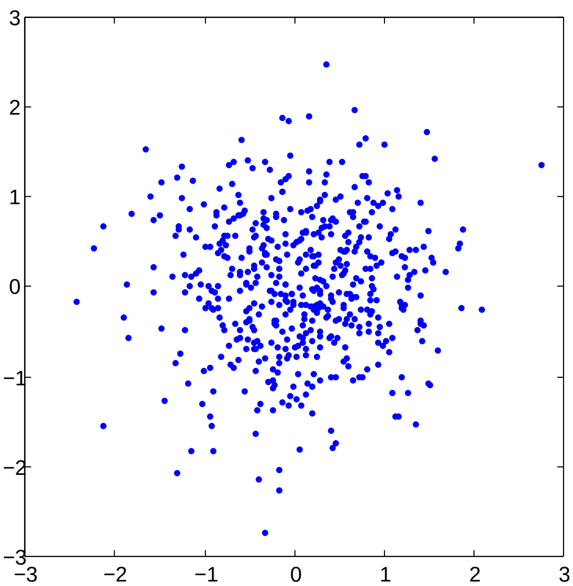
In [RB13], the outlier eigenvalues of perturbations of the single ring model are studied and their locations and limiting fluctuations are obtained ([RB13, Theorem 2.9]) for finite rank and finite operator norm perturbations of arbitrary Jordan type. Note that the special case of the Ginibre ensemble, which is an iid matrix, is contained in this model as well. Our approach to dealing with perturbations of various Jordan types is similar and relies



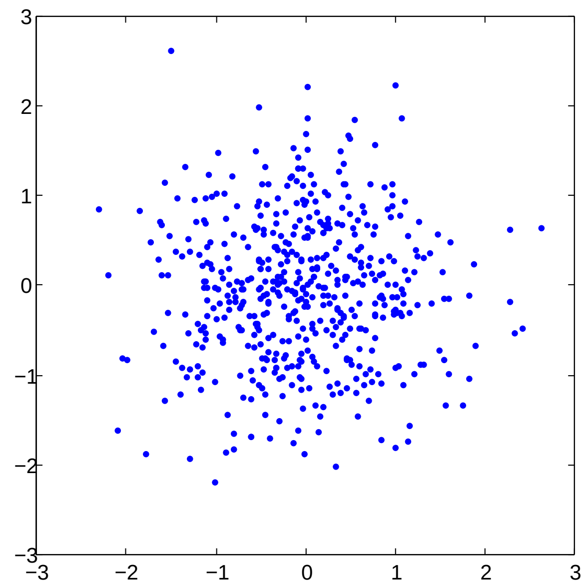
(a)



(b)



(c)



(d)

Figure 1.2: Comparison of empirical and theoretical outlier distributions for non-Gaussian and Gaussian atom distribution x.

on a deterministic perturbation result known as the Lidskii-Vishik-Lyusternik perturbation theorem (see [Lid66], [VL60], [MBO97] and references therein) which we have reproduced in Appendix A.

In [BC14], Bordenave and Captaine study asymptotic outlier locations and fluctuations for perturbed iid matrices. The perturbations considered there are of the form $A = A' + A''$ where A'' is of bounded rank and A' (with possibly unbounded rank) satisfies a well-conditioning property. In the case of local perturbations, where A has a finite nonzero block A'' at the top-left, [BC14, Theorems 1.7 and 1.8] obtain the limiting normalized outlier fluctuation when $A'' = \theta 1_{\mathrm{rk}(A)}$ and when $A'' = J_{\theta,\mathrm{rk}(A)}$ under the hypothesis of bounded fourth moments. In the case when $A'' = vu^*$ is of rank 1 and is delocalized ($\|u\|_\infty, \|v\|_\infty = O(n^{-1/2})$), they show that the outliers exhibit macroscopic fluctuations and demonstrate a convergence of these fluctuations to the zeros of a Gaussian analytic function. While this phenomenon does not occur with finite rank perturbations, some techniques of the proof are similar to the ones in our proof.

In the setting of finite rank perturbations of iid matrices, when Theorem 5 is specialized appropriately, our results coincide with [RB13, Theorem 2.9] for the Ginibre ensemble and with [BC14, Theorems 1.7 and 1.8] for local perturbations of the specified Jordan types. All other cases however, with X having a non-Gaussian atom distribution and A having general eigenvectors (see Remark 1), including the delocalized cases of (ii) and (iii), do not appear to have been explicitly addressed in the literature.

The main technical result of this thesis is Proposition 1 which we prove using the moment method. We require a bounded number of moments in all cases and are able to obtain the limiting fluctuations in a more general setting with a unified approach.

The organization of the rest of this thesis is as follows. In Chapter 2 we prove Proposition 1 which characterizes the joint asymptotic distribution of certain random variables arising from powers of X_n appearing in the Neumann series of $(X_n/\sqrt{n} - \lambda)^{-1}$. In Chapter 3 we prove Lemma 7, which determines the joint limiting distribution of random variables related to a normalized resolvent of X_n, namely of the form $\sqrt{n}u^*[(X_n/\sqrt{n} - \lambda)^{-1} + \lambda^{-1}]v$.

Using Lemma 7, Theorem 5 is proven in Chapter 4, with the help of Proposition 2 from Appendix A, a deterministic perturbation result needed to understand the effect of Jordan blocks in perturbations. Appendix B presents the truncation argument that allows us to prove Proposition 1 and Lemma 7 under weaker assumptions.

1.5 Notation

In this section we detail the notation used in this thesis.

Asymptotic notation: n will always be used to denote a parameter going to infinity and many quantities will be implicitly understood to depend on n. We will use the asymptotic notation $X = O(Y)$ and $X \ll Y$ to mean there is a constant C independent of n, but possibly dependent on other parameters, such that $X \leq CY$ for sufficiently large n. Similarly, we write $X = \Omega(Y)$ or $X \gg Y$ to mean for some C and sufficiently large n, $X \geq CY$. We write $X = o(Y)$ and $y = \omega(x)$ to mean $\lim_{n\to\infty} X/Y \to 0$.

Probability and combinatorics: For a sequence of events $E = E_n$, we say E occurs with high probability (w.h.p.) if $\mathbb{P}(E_n) = 1 - o(1)$. We use $\Rightarrow$ to denote convergence in distribution (and occasionally to denote logical implication) and $\to_P$ to denote convergence in probabiltiy. Finally, we write $[k]$ for $\{1, 2, \dots, k\}$.

Vector and matrix notation: $\|X\|$ will denote the operator norm of X, $\text{rk}(A)$ will denote the rank of A and $\langle v, u\rangle$ will denote the inner product $u^*v = \sum_{i=1}^n \overline{u_i} v_i$.

CHAPTER 2

A central limit theorem

2.1 Motivation and statement

To obtain the limiting fluctuations of the outliers in Theorem 5, we will have to derive the joint asymptotic distributions for certain bilinear averages of the recentered and normalized resolvent, namely for

$$S_\lambda^{u,v} := -\lambda\sqrt{n}u^*((X/\sqrt{n}-\lambda)^{-1}+\lambda^{-1})v, \tag{2.1}$$

with u and v ranging over the generalized eigenvectors of the perturbation matrix A. To this end, in this chapter we prove Proposition 1 which obtains the limiting joint distribution for a bounded number of terms of the Neumann series of (2.1). In Lemma 7 we will control the tail of (2.1), thus obtaining its limiting distribution.

Recall the notation introduced in (1.3) which we reproduce here for convenience. For any complex vector z, we let

$$z^{(d)} := \begin{cases} z & : d=0 \\ \overline{z} & : d=1 \end{cases}.$$

For $S \subset [n]\times[n]$, we define $X_S = (X_{ij}^S)$ through

$$X_{ij}^S = \delta_{(i,j)\in S}x_{ij}.$$

Proposition 1. *Let X be an iid matrix and $(u_i,v_i)_{i=1}^p = (u_i^{(n)},v_i^{(n)})_{i=1}^p$ be a sequence of vectors in $\mathbb{C}^n$. We assume the hypotheses of Theorem 5 with $(u_i,v_i)_{i=1}^p$ in the place of $(u_{s_r},v_{t_r})_{r\in I_2}$. Thus, in the place of (1.4) and (1.5), we assume the following limits and define the scalars*

$$C_{i_1,i_2}^{(d_1),(d_2)} := \lim_{n\to\infty} (u_{i_1})^{(d_1)*}\overline{(u_{i_2})}^{(d_2)}(v_{i_1})^{(d_1)T}(v_{i_2})^{(d_2)}. \tag{2.2}$$

We will assume $\mathbb{E}|x|^m < \infty$ *with* m *defined via* (1.9) *and* (1.10). *Define*

$$Z_{i,j} = Z_{i,j}^{(n)} := \sqrt{n}u_i^* \left(\frac{X_n}{\sqrt{n}}\right)^j v_i$$

where we have suppressed the n *dependence for* X_n, $u_i^{(n)}$ *and* $v_i^{(n)}$. *Also, for*

$$L := \bigcup_{i\in[p]} \{(k,l) \in [n]\times[n] : |u_{i,k}v_{i,l}| \geq n^{-1/4+\delta}\}$$

and $L^c := ([n]\times[n])\backslash L$, *define*

$$Z_{i,j}^L := \sqrt{n}u_i^* \left(\frac{X_L}{\sqrt{n}}\right)^j v_i$$

and

$$Z_{i,j}^{L^c} := \sqrt{n}u_i^* \left(\frac{X_{L^c}}{\sqrt{n}}\right)^j v_i.$$

For $j = 1$, *we will assume that the following joint convergences in distribution and define the independent families* $(G_{i,1}^L)_{i=1}^p$ *and* $(G_{i,1}^{L^c})_{i=1}^p$ *through*

$$(Z_{i,1}^L)_{i=1}^p \Rightarrow (G_{i,1}^L)_{i=1}^p$$

and

$$(Z_{i,1}^{L^c})_{i=1}^p \Rightarrow (G_{i,1}^{L^c})_{i=1}^p.$$

Also define $G_{i,1} := G_{i,1}^L + G_{i,1}^{L^c}$ *so that*

$$(Z_{i,1})_{i=1}^p \Rightarrow (G_{i,1})_{i=1}^p. \tag{2.3}$$

Then for any fixed $m \geq 1$, *the* pm *random variables* $(Z_{i,j})_{i=1,j=1}^{p,m}$ *converge jointly in distribution to the law of random variables* $(G_{i,j})_{i=1,j=1}^{p,m}$ *with* $(G_{i,j})_{i=1,j=2}^{p,m}$ *specified by*

(i) *The* $G_{i,j}$*'s are centered complex Gaussians for* $j \geq 2$ *with mixed second moments given by*

$$\mathbb{E}G_{i_1,j}^{(d_1)}G_{i_2,k}^{(d_2)} = \delta_{jk}(\mathbb{E}x^{(d_1)}x^{(d_2)})^j C_{i_1,i_2}^{(d_1),(d_2)}. \tag{2.4}$$

(ii) The collections of random variables $(G_{i,1})_{i=1}^{p}$ *and* $(G_{i,j})_{i=1,j=2}^{p,m}$ *are independent.*

Note in particular that for $j \neq k$, $Z_{i_1,j}$ *and* $Z_{i_2,k}$ *are asymptotically independent.*

Remark 3. *We note that the case* $p = 1$ *and* $c = 1$ *is a generalization of [Tao13, Section 4] to the complex case with weaker moment assumptions, and is a special case of [BC14, Theorems 6.3, 6.4].*

Remark 4. *The assumption of the joint convergence of* $Z_{i,1}^{L}$ *and* $Z_{i,1}^{L^c}$ *is satisfied under various conditions. We describe some of these below.*

(i) If each u_i *and* v_i *have finite support in* $[C]$ *independent of* n*, we have the case of a local perturbation and the* $G_{i,1}$*'s are finite linear combinations of the* $x_{i,j}$*'s.*

(ii) If each u_i *and* v_i *is uniformly delocalized in the sense that* $\|u_i\|_\infty = o(1)$ *and* $\|v_i\|_\infty = o(1)$ *for* $i \in [p]$*, then by the classical central limit theorem, the* $G_{i,1}$*'s are joint centered complex Gaussians with mixed second moments given by*

$$\mathbb{E}G_{i_1,1}^{(d_1)}G_{i_2,1}^{(d_2)} = \mathbb{E}x^{(d_1)}x^{(d_2)}C_{i_1,i_2}^{(d_1),(d_2)}.$$

(iii) Each u_i *and* v_j *can be allowed to have a local and a uniformly delocalized part. Namely, we suppose that for some* C *independent of* n *and all* $i \in [p]$, $\sup_{i>C}|u_i|, \sup_{i>C}|v_i| = o(1)$*. In this case, the* $G_{i,1}$*'s are a sum of a finite linear combination of the* $x_{i,j}$*'s and an independent Gaussian.*

(iv) Finally, we mention an example that is not contained in the above cases. Let $p = 1$*, fix* $0 < r < 1$ *and set* $u_{1,k} = v_{1,k} = r^k c_n$ *with* c_n *chosen such that* $u^*v = \theta := 2$ *say. Then* $G_{1,1}$ *is an infinite linear combination of the* $x_{i,j}$*'s with exponentially decreasing entries.*

2.2 Proof of Proposition 1

Instead of assuming (1.10), via a truncation argument presented in Appendix B, it suffices to prove Proposition 1 under the stronger assumption that the atom distribution x satisfies the bound $|x| \leq K := o(n^M)$ with $M = 2/m$ given by

$$M = \max(\min(c, 1/2), 1/4) - \epsilon, \tag{2.5}$$

with c defined by (1.9). Furthermore, by decreasing c slightly (and decreasing ϵ), we may assume

$$\max_{i\in[p]} \|u_i\|_\infty \|v_i\|_\infty \ll n^{-c}$$

instead. We will also assume without loss of generality that $(u_i, v_i)_{i\in[p]}$ are unit vectors.

In step 1, we show that $(Z_{i,1}^L)_{i=1}^p$ is asymptotically independent of

$$(Z_{i,1}^{L^c})_{i=1}^p \cup (Z_{i,j})_{i=1,j=2}^{p,m}.$$

In step 2, we derive the joint asymptotic distribution of $(Z_{i,1}^{L^c})_{i=1}^p \cup (Z_{i,j})_{i=1,j=2}^{p,m}$. A key part of the proof is contained in Lemma 5, whose proof we postpone to the end of this chapter.

Step 2 employs the moment method which, together with the truncation method (see Appendix B), contributes to the moment hypothesis. The moment hypothesis decays when the random variables $(Z_{i,1}^L)_{i=1}^p$ are dealt with using the moment method; thus we deal with them separately.

We will need

Lemma 1. *Let $A^{(n)} = (A_1^{(n)}, \dots, A_k^{(n)})$, $B^{(n)} = (B_1^{(n)}, \dots, B_k^{(n)})$ and $C^{(n)} = (C_1^{(n)}, \dots, C_l^{(n)})$ be sequences of complex vector valued random variables such that*

$$(A^{(n)}, C^{(n)}) \Rightarrow (A, C) \text{ and } B^{(n)} \to_P 0.$$

Then $(A^{(n)} + B^{(n)}, C^{(n)}) \Rightarrow (A, C)$. In particular, if $A^{(n)}$ and $C^{(n)}$ are independent, then $A^{(n)} + B^{(n)}$ and $C^{(n)}$ are asymptotically independent.

Proof. This follows from the Cramér-Wold device (see [Bil99, Chapter 1.7]) and appears in [Bil99, Exercise 1.4.2]. □

2.2.1 Step 1

For $j \geq 2$, define

$$Z'_{i,j} := n^{-(j-1)/2} u_i^* (X - X_L)^j v_i.$$

Note that $Z^L_{i_1,j_1}$ and Z'_{i_2,j_2} are functions of disjoint subsets of $\{x_{rs} : r, s \in [n]\}$ and hence, $(Z_{i,j})_{(i,j)\in D}$ and $(Z'_{i,j})_{(i,j)\in D^c}$ are independent. $Z^L_{i_1,j_1}$ and $Z^{L^c}_{i_2,j_2}$ are independent for the same reason.

By Lemma 1, it suffices to show that

$$E = E_{i,j} := n^{-(j-1)/2} u_i^* (X^j - (X - X_L)^j) v_i \to_P 0 \tag{2.6}$$

for $i \in [p]$ and $2 \leq j \leq m$.

We will need the following result.

Lemma 2. *Let u and v be unit vectors in $\mathbb{C}^n$ and X be an iid random matrix with atom distribution having mean 0, variance 1 and bounded fourth moment. Then*

$$\mathbb{E}\left| u^* \left(\frac{1}{\sqrt{n}} X\right)^k v \right|^2 = O\left(\frac{1}{n}\right) \tag{2.7}$$

for any fixed $k \geq 1$.

Remark 5. *Lemma 2 is a special case of Lemma 9 which establishes the same statement for k that is allowed to grow polynomially with n. We postpone the proof to Chapter 3, where the result is needed in full generality. We remark that Lemma 2 can also be found in [Tao13, Lemma 2.3].*

Fix $j \geq 2$ and let $\delta_n = \log n$ (any slowly growing function of n will suffice). By Lemma 2 and Markov's inequality, for any $k \geq 1$,

$$\bigcap_{m=1}^{M} \left\{ u_m^* \left(\frac{1}{\sqrt{n}} X\right)^k v_m \leq \frac{\delta_n}{\sqrt{n}} \right\} \tag{2.8}$$

occurs with high probability for any finite set of $2M$ unit vectors $(u_m)_{m=1}^M$ and $(v_m)_{m=1}^M$.

Recall that

$$L := \bigcup_{i\in[p]} \{(k,l) \in [n] \times [n] : |u_{i,k} v_{i,l}| \geq n^{-1/4+\delta}\}$$

where $\delta > 0$ is fixed. Since $|u_i|_2 = |v_i|_2 = 1$, we have $|L| \ll n^{1/2-2\delta}$. To control, $\|X_L\|$, we will need

Lemma 3. *Suppose* $S \subset A \times B$ *with* $\max(|A|, |B|) \leq m$. *Then* $\|X_S\| \leq O(\log n\sqrt{m})$ *w.h.p.*

Proof. Since $\|X_S\|$ is unchanged when restricting X_S to an $m \times m$ submatrix containing S, we may assume $m = n$. If $S = \emptyset$, Lemma 3 is a consequence of Theorem 3. Writing $X' := X_L - X_{L^c}$, we have

$$\|X_L\| \leq \frac{1}{2}(\|X\| + \|X'\|)$$

from the triangle inequality. If the atom distribution x is symmetric, applying Theorem 3 to X and X' yields the desired bound. To prove the lemma for general x, we will need a symmetrization argument from [Tao12, Section 2.3.2] that we reproduce here for convenience. Letting X'' be an independent copy of X', we have

$$\mathbb{E}[X' - X''|X'] = X'.$$

Since the operator norm is a convex function, we may apply Jensen's inequality to get

$$\|X'\| \leq \mathbb{E}[\|X' - X''\||X'].$$

Removing the conditioning on X', we have

$$\mathbb{E}\|X'\| \leq \mathbb{E}\|X' - X''\|.$$

Now $X' - X''$ has iid entries, so applying Theorem 3, we have

$$\begin{aligned}\mathbb{P}[\|X'\| \geq \log n\sqrt{n}] &\leq \frac{\mathbb{E}\|X'\|}{\log n\sqrt{n}} \\ &\leq \frac{\mathbb{E}\|X' - X''\|}{\log n\sqrt{n}} \\ &= o(1).\end{aligned}$$

□

Applying Lemma 3 with $m = n^{1/2-2\delta}$ gives

$$\|X_L\| \ll (\log n)n^{1/4-\delta} \text{ w.h.p.} \tag{2.9}$$

Now let

$$X^a := \begin{cases} X & : a = 0 \\ X_L & : a = 1 \end{cases}$$

Expanding (2.6), we have

$$\begin{aligned} |E| &\leq \sum_{a=1}^{j} \sum_{\substack{a_1,\cdots,a_j \in \{0,1\} \\ \sum a_i = a}} n^{-(j-1)/2} \left| u_i^* X^{a_1} \ldots X^{a_j} v_i \right| \\ &=: \sum_{a=1}^{k} E_a. \end{aligned}$$

For $a \geq 2$,

$$\begin{aligned} E_a &\ll \binom{j}{a} \left\| \frac{X}{\sqrt{n}} \right\|^{j-a} \left\| \frac{X_L}{\sqrt{n}} \right\|^{a-1} \|X_L\| \\ &= o(1) \text{ w.h.p.}, \end{aligned}$$

where we have used (2.9) and that $a \geq 2$.

To bound E_1, we have

$$\begin{aligned} E_1 &\leq \sum_{m=0}^{j-1} \left| u_i^* \left(\frac{X}{\sqrt{n}} \right)^m X_L \left(\frac{X}{\sqrt{n}} \right)^{j-1-m} v_i \right| \\ &\leq \sum_{(k,l) \in L} \sum_{m=0}^{j-1} |x_{kl}| \left| u_i^* \left(\frac{X}{\sqrt{n}} \right)^m e_k \right| \left| e_l^T \left(\frac{X}{\sqrt{n}} \right)^{j-1-m} v_i \right| \\ &\ll \frac{\delta_n}{\sqrt{n}} \sum_{(k,l) \in L} |x_{kl}| \text{ w.h.p.} \end{aligned}$$

Note that if $j \geq 2$, then either $m \geq 1$ or $j - 1 - m \geq 1$ for $0 \leq m \leq k - 1$. Hence the last line follows from (2.8).

Since $\mathbb{E}|x_{kl}| \leq 1$, $\delta_n = \log n$ and $|L| = O(n^{1/2-2\delta})$, we have $E_1 \to_P 0$ by Markov's inequality, and (2.6) follows.

2.2.2 Step 2

We first state and prove the complex version of Wick's theorem (also known as Isserlis' theorem, see [Iss18]) which will be needed later.

Lemma 4. (Complex Wick's theorem)

Let $(Z_1, Z_2, \dots, Z_n) = (X_1 + iY_1, \dots, X_n + iY_n)$ be a centered complex Gaussian vector. Thus the vector $(X_1, Y_1, \dots, X_n, Y_n)$ is multivariate normal. Then for any $I = (i_1, \dots, i_{2k}) \in [n]^{2k}$,

$$E\prod_{l=1}^{2k} Z_{i_l} = \sum_P \prod_{j=1}^{k} \mathbb{E}[Z_{i_{p_{2j-1}}} Z_{i_{p_{2j}}}]$$

where the sum is over all partitions $P = \bigcup_{j=1}^{k}\{p_{2j-1}, p_{2j}\}$ of $[2k]$ into pairs. Also, the left hand side is 0 if I has odd length.

Proof. Wick's theorem is the statement of the lemma for multivariate centered real Gaussians. The complex version follows by expanding both sides of the equation into real and imaginary parts and applying Wick's theorem. Let

$$W_i^a = \begin{cases} X_i & : a = 1 \\ iY_i & : a = 2 \end{cases}$$

Then

$$\mathbb{E}\prod_{l=1}^{2k} Z_{i_l} = \sum_{a_1,\dots,a_{2k}\in\{1,2\}} \prod_{l=1}^{2k} W_{i_l}^{a_l}$$

while

$$\sum_P \prod_{i=1}^{k} \mathbb{E}[Z_{p_{2j-1}} Z_{p_{2j}}] = \sum_P \sum_{a_1,\dots,a_{2k}\in\{1,2\}} \prod_{i=1}^{k} \mathbb{E}[W_{p_{2j-1}}^{a_{2j-1}} W_{p_{2j}}^{a_{2j}}].$$

Switching the sums and applying Wick's theorem to $\mathbb{E}\prod_{l=1}^{2k} W_{i_l}^{a_l}$ for each choice of the a_l's yields the result.

□

We now prove Proposition 1 for the collection of random variables $(Z_{i,1}^{L^c})_{i=1}^{p} \cup (Z_{i,j})_{i=1,j=2}^{p,m}$. This part of the proof employs the moment method in a similar way to those in [Tao13] and [BC14]. To avoid notational clutter on a first reading, one may set $p = 1$ to grasp the main ideas of the proof.

To handle the $j = 1$ case uniformly, in the proof we will abuse notation by writing $Z_{i,1}$

for $Z_{i,1}^{L^c}$ and $G_{i,1}$ for $G_{i,1}^{L^c}$. When $j=1$, we will denote X_{L^c} by X^j and finally, we define

$$C_{i_1,i_2}^{(d_1),(d_2)}(j) := \begin{cases} C_{i_1,i_2}^{(d_1),(d_2)} : j \geq 2 \\ \lim_{n\to\infty} \sum_{(k,l)\in L^c} \overline{(u_{i_1,k})}^{(d_1)} \overline{(u_{i_2,k})}^{(d_2)} (v_{i_1,l})^{(d_1)} (v_{i_2,l})^{(d_2)}. \end{cases} \tag{2.10}$$

By Carleman's theorem for the case of a complex vector of random variables (see e.g. [BS06a]), it suffices to show that the multivariate mixed moments converge. Namely,

$$\mathbb{E} \prod_{\substack{1\leq i\leq p \\ 1\leq j\leq m}} Z_{i,j}^{r_{i,j}} \overline{Z_{i,j}}^{s_{i,j}} = \mathbb{E} \prod_{\substack{1\leq i\leq p \\ 1\leq j\leq m}} G_{i,j}^{r_{i,j}} \overline{G_{i,j}^{s_{i,j}}} + o(1) \tag{2.11}$$

for $(r_{i,j})_{i=1,j=1}^{p,m}, (s_{i,j})_{i=1,j=1}^{p,m} \in \mathbb{N}^{pm}$.

Let $Q_1 := -\frac{1}{2}\sum_{i,j}(j-1)(r_{i,j}+s_{i,j})$. Then the left hand side of (2.11) is

$$n^{-Q_1}\mathbb{E} \prod_{\substack{1\leq i\leq p \\ 1\leq j\leq m}} (u_i^* X^j v_i)^{r_{i,j}} (u_i^T \overline{X}^j \overline{v_i})^{s_{i,j}}. \tag{2.12}$$

Expanding the product in (2.12) will yield terms corresponding to the union of directed paths on the vertex set $[n]$ with $\sum_i r_{i,j}+s_{i,j}$ of them having length j for each $1\leq j\leq m$. We first introduce notation in order to write (2.12) as a sum $n^{-Q_1}\sum_* W(F)$, with $*$ and $W(F)$ defined appropriately. Next, we reduce the sum to terms with paths having multiplicity two and disjoint interior vertices (see Lemma 5). Finally we apply the complex Wick theorem to obtain the proposition.

Let

$$S := \{(a,b,c,d) : a\in[p], b\in[m], d\in\{0,1\}, c\in[r_{a,b}] \text{ if } d=0 \text{ and } c\in[s_{a,b}] \text{ if } d=1\}$$

be the index set for the $Z_{i,j}$'s. For $s\in S$ we write $s=(s_a,s_b,s_c,s_d)$. Recalling (1.3), (2.12) can be written as

$$n^{-Q_1}\mathbb{E}\prod_{s\in S}(u_{s_a}^* X^{s_b} v_{s_a})^{(s_d)}. \tag{2.13}$$

We let

$$T := \{(s,e) : s\in S \text{ and } e\in[s_{b+1}]\}$$

be the index set of terms within the $Z_{i,j}$'s. For $t\in T$, we write

$$t = (t_s, t_e) = (t_a, t_b, t_c, t_d, t_e).$$

By a slight abuse of notation, we will write u_t for u_{t_a} and u_s for u_{s_a}. We denote the index set for terms in the expansion of (2.13) by

$$\mathcal{F}' := \{F : T \to [n] : t_b = 1 \Rightarrow (F(t,1), F(t,2)) \in L^c\}.$$

Finally for $s \in S$ and $F \in \mathcal{F}'$ let

$$W_s(F) := (u^*_{s,F(s,1)} v_{s,F(s,s_b+1)} \mathbb{1}_{[s_b \geq 2 \text{ or } (F(s,1),F(s,2)) \in L^c]})^{(s_d)} (\mathbb{E} \prod_{e=1}^{s_b} x_{F(s,e),F(s,e+1)})^{(s_d)} \tag{2.14}$$

$$=: W_{s,(u,v)}(F) W_{s,x}(F) \tag{2.15}$$

and set

$$W_{u,v}(F) := \prod_{s \in S} W_{s,(u,v)}(F), \tag{2.16}$$

$$W_x(F) := \prod_{s \in S} W_{s,x}(F)$$

and

$$W(F) := \prod_{s \in S} W_s(F). \tag{2.17}$$

Now we can write (2.13) as

$$n^{-Q_1} \mathbb{E} \prod_{s \in S} (u^*_s X^{s_b} v_s)^{(s_d)} = n^{-Q_1} \sum_{F \in \mathcal{F}'} W(F). \tag{2.18}$$

For each partition $\mathcal{T} = \{T_1, \ldots, T_q\}$ of T, set

$$\mathcal{F}_{\mathcal{T}} := \{F \in \mathcal{F} : \{F^{-1}(i) : i \in [n], F^{-1}(i) \neq \emptyset\} = \{T_1 \ldots, T_q\}\}$$

to be the set of terms F whose preimages induce the partition $\{T_1. \ldots, T_q\}$. We can now write

$$n^{-Q_1} \sum_{F \in \mathcal{F}} W(F) = n^{-Q_1} \sum_{\mathcal{T} = \{T_1, \ldots, T_q\}} \sum_{F \in \mathcal{F}_{\mathcal{T}}} W(F).$$

We now define notation for the edges of the graph induced by the terms F. First, let $E := \{(t,t') \in T^2 : t_s = t'_s, t'_e = t_e + 1\}$ and fix a partition $\mathcal{T} = \{T_1 \ldots, T_q\}$ of T. For $F \in \mathcal{F}_{\mathcal{T}}$ and $i, j \in [q] = [q(\mathcal{T})]$, let

$$E^{\mathcal{T}}_{i,j} := \{e = (t,t') \in E : t \in T_i \text{ and } t' \in T_j\}$$

and let

$$E_{\mathcal{T}} := \{E^{\mathcal{T}}_{i,j} : |E^{\mathcal{T}}_{i,j}| > 0\}.$$

Note that $(|e|)_{e\in E_{\mathcal{T}}}$ is independent of $F \in \mathcal{F}_{\mathcal{T}}$ and that

$$W_x(F) = \prod_{e\in E_{\mathcal{T}}} \mathbb{E}|x|^{|e|}. \tag{2.19}$$

Since $\mathbb{E}|x| = 0$, $W_x(F) = 0$ if $|e| = 0$ for any $e \in E_{\mathcal{T}}$. Thus defining

$$\mathcal{F} := \bigcup_{\substack{\mathcal{T} \text{ partition of } T: \\ |e|\geq 2 \forall e\in E_{\mathcal{T}}}} \mathcal{F}_{\mathcal{T}},$$

we have

$$n^{-Q_1} \sum_{F\in\mathcal{F}'} W(F) = n^{-Q_1} \sum_{F\in\mathcal{F}} W(F). \tag{2.20}$$

Each $F \in \mathcal{F}$ can be interpreted as a union of paths on $[n]$. More precisely, letting $T_s := \{t \in T : t_s = s\}$, we define $\pi_{F,s} := F|_{T_s}$ to be the path of F corresponding to term $s \in S$. The interior vertices of $\pi_{F,s}$ are defined to be $F(\{(s, e) : e = 2, 3, \ldots, s_b\})$.

Lemma 5. *Assume the hypotheses of Proposition 1 and recall the notation introduced above. Let $\mathcal{F}_0$ be the set of terms F such that each path $\pi_{F,s}$ for $s \in S$ has multiplicity 2 and different paths have disjoint interior vertices.*

Then

$$n^{-Q_1} \sum_{F\in\mathcal{F}} W(F) = n^{-Q_1} \sum_{F\in\mathcal{F}_0} W(F) + o(1).$$

We will postpone the proof of the lemma to the end of the chapter. Assuming the lemma, we now prove the proposition.

First suppose $\sum_i r_{i,j} + s_{i,j}$ is odd for some j. Then $\mathcal{F}_0$ is empty and the left-hand side of (2.11) is $o(1)$ which matches the right-hand side by the vanishing of odd mixed moments of a centered complex Gaussian. For the rest of the proof, we can thus assume that for each j, $\sum_i r_{i,j} + s_{i,j}$ is even.

We group the terms in $\mathcal{F}_0$ as follows. Let $S_j := \{s \in S : s_b = j\}$ and define $\mathcal{P}_j$ to be the set of unordered partitions of S_j into parts of size two. Note that by assumption, $|S_j|$ is even for all j.

For $F \in \mathcal{F}_0$, note by (2.14) and (2.17) that $W(F)$ does not depend on the interior points $\{f(s,e) : s \in S, e = 2, \ldots, s_b\}$. There are $\sum_{i,j}(j-1)(r_{i,j}+s_{i,j})$ such points which occur in pairs and can be chosen in n^{Q_1} ways.

For $F \in \mathcal{F}_0$ and $j \in [m]$, let $P_{F,j} \in \mathcal{P}_j$ be the partition of $\mathcal{P}_j$ induced by F. Then F satisfies the condition that for each part $\{p,q\} \in P_{F,j}$, $F(p,1) = F(q,1)$ and $F(p, p_b+1) = F(q, q_b+1)$.

Summing over the choices for interior points and F satisfying the above condition instead of summing over $F \in \mathcal{F}_0$ incurs an $o(1)$ error and we have

$$n^{-Q_1} \sum_{F \in \mathcal{F}_0} W(F) = \prod_{j=1}^{m} \sum_{P_j \in \mathcal{P}_j} \prod_{\{p,q\} \in P_j} \sum_{\substack{F(p,1)=F(q,1),\\ F(p,j+1)=F(q,j+1)\in[n]}} W_p(F) W_q(F) + o(1) \tag{2.21}$$

where, recalling (2.14),

$$W_p(F)W_q(F) = (\mathbb{E}x^{(p_d)}x^{(q_d)})^j \prod_{r \in \{p,q\}} \overline{u}^{(r_d)}_{r,F(r,1)} v^{(r_d)}_{r,F(r,j+1)} \mathbb{1}_{[j \geq 2 \text{ or } (F(r,1),F(r,2)) \in L^c]}.$$

Finally, using (2.2) and (2.10), (2.21) evaluates to

$$\prod_{j=1}^{m} \sum_{P_j \in \mathcal{P}_j} \prod_{\{p,q\} \in P_j} (\mathbb{E}x^{(p_d)}x^{(q_d)})^j C^{(p_d),(q_d)}_{p_a,q_a}(j) + o(1). \tag{2.22}$$

On the other hand, we let $\mathcal{P}$ be the set of partitions of S into pairs and for $s \in S$, we set $G_s := G^{(s_d)}_{s_a,s_b}$. Note that for $j \neq k$, $\mathbb{E}G_{i_1,j}G_{i_2,k} = 0$ and hence $G_{i_1,j}$ and $G_{i_2,k}$ are independent. Applying Wick's theorem to the right hand side of (2.11) gives

$$\begin{aligned}
\mathbb{E} \prod_{\substack{1 \leq i \leq p \\ 1 \leq j \leq m}} G^{r_{i,j}}_{i,j} \overline{G^{s_{i,j}}_{i,j}} &= \mathbb{E} \prod_{s \in S} G_s \\
&= \prod_{j=1}^{m} \mathbb{E} \prod_{s \in S_j} G_s \\
&= \prod_{j=1}^{m} \sum_{P_j \in \mathcal{P}_j} \prod_{\{p,q\} \in P_j} \mathbb{E} G_p G_q
\end{aligned}$$

where we have used Wick's theorem in the third line. Comparing (2.22) and (2.4) then concludes the proof of the proposition.

Note the following special cases of Proposition 1, where we write $G_{i,1}$ for $G_{i,1}^{L^c}$.

(i) If $\mathbb{E}x^2 = 0$, condition (2.4) becomes

$$\mathbb{E}G_{i_1,j}\overline{G_{i_2,k}} = \delta_{jk}C^{(0),(1)}(i_1,i_2) \tag{2.23}$$

and

$$\mathbb{E}G_{i_1,j}G_{i_2,k} = 0.$$

(ii) If we further assume that for $p = d^2$, the vectors $(u_i, v_i)_{i=1}^p$ are of the form $(u_a, u_b)_{a,b=1}^d$ with $(u_a)_{a=1}^d$ orthonormal, then (2.23) reduces to

$$\mathbb{E}G_{(a,b),j}\overline{G_{(c,d),k}} = \delta_{jk}\delta_{ab}\delta_{cd}. \tag{2.24}$$

2.3 Proof of Lemma 5

Fix a partition $\mathcal{T} = \{T_1, \ldots, T_q\}$ of T with $|e| \geq 2$ for every $e \in E_{\mathcal{T}}$. We first rewrite the sum $n^{-Q_1}\sum_{F\in\mathcal{F}_{\mathcal{T}}} W(F)$ as a product of terms over $j \in [q]$.

Define $T^1 := \{t \in T : t_e = 1\}$, $T^2 := \{t \in T : t_e = t_b + 1\}$, $T^3 := T\backslash(T^1 \cup T^2)$ and let $T_j^l := T_j \cap T^l$ for $l = 1, 2, 3$. For $t \in T$ and $i \in [n]$, define the vertex weights

$$w(t,i) := \begin{cases} |u_{t,i}| : t \in T^1 \\ |v_{t,i}| : t \in T^2 \\ n^{-1/2} : t \in T^3 \end{cases} .. \tag{2.25}$$

The $w(t,i)$'s account for the factors n^{-Q_1} and $W_{u,v}(F)$ in (2.16) and (2.18) respectively. Since $\mathbb{E}|x|^a \ll K^{(a-4)_+}$, using (2.19) we have

$$\begin{aligned} n^{-Q_1}\left|\sum_{F\in\mathcal{F}_{\{T_1,\ldots,T_q\}}} W(F)\right| &\leq \sum_{\substack{i_1,\ldots,i_q\in[n] \\ \text{distinct}}} \left(\prod_{j=1}^q \prod_{t\in T_j} w(t,i_j)\right) \prod_{e\in E_{\mathcal{T}}} K^{(|e|-4)_+} \\ &\leq \sum_{i_1,\ldots,i_q\in[n]} \left(\prod_{j=1}^q \prod_{t\in T_j} w(t,i_j)\right) \prod_{e\in E_{\mathcal{T}}} K^{(|e|-4)_+}. \end{aligned} \tag{2.26}$$

We would like to bound $\prod_{e\in E_{\mathcal{T}}} K^{(|e|-4)_+}$ by $\prod_{t\in T} K^*(t, i_{j(t)})$ for some suitably defined K^* in order to bound the right-hand side (2.26) by

$$\prod_{j=1}^{q} \sum_{i_1,\ldots,i_q\in[n]} \prod_{t\in T_j} w(t,i_j)K^*(t,i_j).$$

We do this first for the expression $\prod_{e\in E_{\mathcal{T}}} K^{|e|}$ in order to motivate some of the technical definitions. Fix $i_1,\ldots,i_q \in [n]$ and assume for $t \in T$ and $j \in [q]$ that $|u_{t,i_j}|, |v_{t,i_j}| \neq 0$. Recall the parameter $c \in [0,1]$ from (2.2). For $t \in T_j$, $t^1 \in T_j^1$ and $t^2 \in T_j^2$, define

$$K(t,i) := \begin{cases} |u_{t,i}^{-(1-\epsilon)}| & : t \in T^1 \\ |v_{t,i}^{-(1-\epsilon)}| & : t \in T^2 \\ \max(K^2 n^{-c(1-\epsilon)}, K) & : t \in T^3 \end{cases}. \tag{2.27}$$

We first show that

$$\prod_{e\in E_{\mathcal{T}}} K^{|e|} \ll \prod_{j=1}^{q}\prod_{t\in T_j} K(t,i_j). \tag{2.28}$$

Fix $s \in S$. Suppoes $s_b = 1$. Then for δ and ϵ sufficiently small,

$$\begin{aligned}\prod_{t\in T: t_s = s} K(t, i_{j(t)}) &\geq \min_{(k,l)\in L^c} |u_{t,k}v_{t,l}|^{-(1-\epsilon)} \\ &\gg n^{\max(1/4-\delta, c)(1-\epsilon)} \\ &\gg K. \end{aligned}\tag{2.29}$$

The last line follows from $M < \max(1/4, c)$ which is a consequence of (2.5), .

If $s_b \geq 2$,

$$\begin{aligned}\prod_{t\in T: t_s = s} K(t, i_{j(t)}) &\gg K^2 n^{c(1-\epsilon)} \left(\|u_t\|_\infty \|v_t\|_\infty\right)^{-(1-\epsilon)} K^{s_b-2} \\ &\gg K^{s_b}\end{aligned}\tag{2.30}$$

where we have used $\|u_t\|_\infty\|v_t\|_\infty \ll n^{-c}$. Using (2.29) and (2.30) and taking the product over $s \in S$ gives (2.28). We now define $K^*(t,i)$ in such a way that we have the analogous bound

$$\prod_{e\in E_{\mathcal{T}}} K^{(|e|-4)_+} \ll \prod_{j=1}^{q}\prod_{t\in T_j} K^*(t,i_j). \tag{2.31}$$

First, order the elements of $T_j^l = \{t_1^l, t_2^l, \dots, t_{|T_j^l|}^l\}$ arbitrarily for $l = 1, 2, 3$. We define the set $C_j \subset T_j$ by the following conditions.

(i) $t_k^3 \in C_j \iff k \leq 2$.

(ii) For $l = 1, 2$, $t_k^l \in C_j \iff k + |T_j^3| \leq 2$.

It is easy to verify that $|C_j \backslash T_j^1|, |C_j \backslash T_j^2| \leq 2$. We now define

$$K^*(t, i) := \begin{cases} 1 : t \in C_j \\ K(t, i) : \text{ otherwise } . \end{cases} .$$

We now prove (2.31). Fix $e \in E_{\mathcal{T}}$ and suppose $e \subset T_i \times T_j$. Define $e' \subset e$ by

$$e' := \{(s, t) \in e : s \in C_i \text{ or } t \in C_j\}.$$

Since $|C_i \backslash T_i^2|, |C_j \backslash T_j^1| \leq 2$, $|e'| \leq 4$ and we have

$$\prod_{e \in \mathcal{T}} K^{(|e|-4)_+} \leq \prod_{e \in \mathcal{T}} K^{|e \backslash e'|}.$$

It thus suffices to show

$$\prod_{e \in \mathcal{T}} K^{|e \backslash e'|} \leq \prod_{j=1}^{q} \prod_{t \in T_j} K^*(t, i_j).$$

As in the proof of (2.28), we fix $s \in S$. Let $C := \bigcup_{j \in [q]} C_j$ and define

$$e_s = \{((s, l), (s, l+1)) : 1 \leq l \leq s_b \text{ and } (s, l), (s, l+1) \notin C\}$$

and

$$v_s = \{(s, l) : (s, l) \notin C \text{ and } l = 2, 3, \dots, s_b\}.$$

Since $K(t, i) \geq 1$ for $t = (s, 1)$ and $t = (s, s_b + 1)$, it suffices to show

$$K^{|e_s|} \leq \prod_{t \in v_s} K(t, i).$$

If $|e_s| = s_b$, this follows from (2.29) and (2.30). Now suppose $|e_s| < s_b$. We first show that $|e_s| \leq |v_s|$. Choose l^* such that $(s, l^*) \in C$ and define the map $f : e_s \to v_s$ by

$$f((s, l), (s, l+1)) := \begin{cases} (s, l+1) : l \leq l^* - 2 \\ (s, l) : l \geq l^* + 1 \end{cases} .$$

We see that f is injective and hence $|e_s| \le |v_s|$. Since $K(t,i) \ge K$ for $t \in v_s$, we have

$$\begin{aligned} K^{|e_s|} &\le K^{|v_s|} \\ &\le \prod_{t \in v_s} K(t,i) \end{aligned}$$

completing the proof of (2.31).

We can now use (2.31) in (2.26) to write

$$\begin{aligned} \left| n^{-Q_1} \sum_{F \in \mathcal{F}_{\{T_1,\dots,T_q\}}} W(F) \right| &\le \sum_{i_1,\dots,i_q \in [n]} \prod_{j=1}^{q} \prod_{t \in T_j} w(t,i_j) K^*(t,i_j) \\ &= \prod_{j=1}^{q} \left(\sum_{i_j \in [n]} \prod_{t \in T_j} w(t,i_j) K^*(t,i_j) \right) \qquad (2.32) \\ &=: \prod_{j=1}^{q} W^*(T_j). \qquad (2.33) \end{aligned}$$

We now fix a part of $\mathcal{T}$, say T_1 and consider $W^*(T_1)$. To prove Lemma 5, it suffices to prove the following.

Lemma 6. *(i)* $W^*(T_1) = O(1)$

(ii) If $|T_1^3| \ge 1$, *then* $|W^*(T_1)| = o(1)$ *unless* $|T_1^3| = |T_1| = 2$.

(iii) $\prod_j W^*(T_j) = o(1)$ *unless* $|e| = 2$ *for every* $e \in E_{\mathcal{T}}$.

Proof. We first show that

$$w(t,i)K(t,i) = \begin{cases} O(1) : t \in T_1^1 \cup T_1^2 \\ o(1) : t \in T_1^3 \end{cases} \qquad (2.34)$$

using (2.25), (2.27) and (2.5). Suppose $t \in T^1$. Then $w(t,i)K(t,i) \le |u_i|^\epsilon = O(1)$. We have a similar bound for $t \in T^2$. Finally, if $t \in T^3$, then

$$w(t,i)K(t,i) = n^{-1/2} \max(K^2 n^{-c(1-\epsilon)}, K).$$

Since $K = o(n^M)$ and $M \le \min(1/2, c)$, we have the desired bound. This implies in particular that for any $D \subset C_1$,

$$W^*(T_1) \ll \sum_{i \in [n]} \prod_{t \in D} w(t,i). \qquad (2.35)$$

We prove Lemma 6.(ii) first. For u and v unit vectors in $\mathbb{C}^n$, we will need the estimate

$$\sum_{i\in[n]} |u_i|^\epsilon \ll O(n^{1-\epsilon/2}) \tag{2.36}$$

which follows from Hölder's inequality. Suppose $|T_1^3| = 1$. Then, since each edge has multiplicity at least 2, we must have $|T_1^1|, |T_1^2| \geq 1$. Applying (2.35) with $D = C_1 = \{t_1^1, t_1^2, t_1^3\}$, we have that for some (u, v),

$$\begin{aligned} W^*(T_1) &\ll \sum_{i=1}^n n^{-1/2}|u_i||v_i| \\ &\leq O(n^{-1/2}). \end{aligned}$$

If $|T_1^3| \geq 3$, then $C_1 = \{t_1^3, t_2^3\}$ and

$$\begin{aligned} W^*(T_1) &\ll \sum_{i\in[n]} n^{-1}K^*(t_3^3, i) \\ &= o(1) \end{aligned}$$

by (2.34). Finally, suppose $|T_1^3| = 2$. Suppose $|T_1^1| \geq 1$. Then from (2.36), we have

$$W^*(T_1) \ll \sum_{i\in[n]} n^{-1}|u_i|^\epsilon = o(1).$$

We have a similar estimate if $|T_1^2| \geq 1$. We conclude that if $|T_1^3| \geq 1$, $W^*(T_1) = o(1)$ unless $|T_1^3| = 2$ and $|T_1^1| = |T_1^2| = 0$, in which case $W^*(T_1) = O(1)$.

We now prove (iii). Assume first that e is an edge incident to distinct vertices, say $e \subset T_1 \times T_2$, and that $|e| \geq 3$. By (ii), we may assume $T_1^3 = T_2^3 = \emptyset$. Since $|T_1^1|, |T_2^2| \geq 3$, we may choose $(s_i, t_i) \in e$ for $i = 1, 2, 3$ where $s_i \in T_1^1$ and $t_i \in T_2^2$ and let $C_1 = \{s_1, s_2\}$ and $C_2 = \{t_1, t_2\}$. Then bounding $W^*(T_1)W^*(T_2)$ by the contribution from $(s_i, t_i)_{i=1}^3$, we have

$$\begin{aligned} W^*(T_1)W^*(T_2) &\ll \sum_{(i,j)\in L^c} \prod_{k=1}^2 |u_{s_k,i}v_{t_k,j}||u_{s_3,i}v_{t_3,j}|^\epsilon \\ &\leq \max_{(i,j)\in L^c} |u_{s_3,i}v_{t_3,j}|^\epsilon \sum_{i\in[n]} |u_{s_1,i}||u_{s_2,i}| \sum_{j\in[n]} |v_{t_1,j}||v_{t_2,j}| \\ &= o(1). \end{aligned}$$

We have a similar bound if e is a loop at say T_1.

To complete the proof of the lemma, it remains to prove (i) in the cases not covered by (ii) and (iii). Thus, set $|T_1^3| = 0$ and assume without loss of generality that $|T_1^1| \geq 2$. Then with $D = \{t_1^1, t_1^2\} =: \{s, t\}$ in (2.35) we have

$$\begin{aligned} W^*(T_1) &\ll \sum_{i \in [n]} |u_{s,i} u_{t,i}| \\ &= O(1). \end{aligned}$$

□

CHAPTER 3

Proof of Lemma 7

Recall the bilinear average of the normalized resolvent introduced in (2.1) in Chapter 2. In this chapter, we control the tail of its Neumann series and, with the help of Proposition 1, obtain the joint limiting distribution of such terms in Lemma 7. This is the main ingredient in the proof of Theorem 5 which is presented in the next chapter.

Lemma 7. *Fix complex numbers $\theta_1, \dots, \theta_a$ with $|\theta_j| > 1$ for $j \in [a]$ and suppose $\lambda_j = \lambda_{n,j} \to_P \theta_j$ as $n \to \infty$. Let $(u_i, v_i)_{i=1}^p$ be p pairs of vectors satisfying the hypotheses of Proposition 1. Let*

$$S_{i,j} := \sum_{k \geq 1} \frac{\sqrt{n} \left\langle (\frac{X}{\sqrt{n}})^k v_i, u_i \right\rangle}{\lambda_j^k} =: \sum_{k \geq 1} \frac{Z_{i,k}}{\lambda_j^k}.$$

Recall the definition of $(G_{i,1})_{i=1}^p$ from Proposition 1 and define centered complex Gaussians $(g_{i,j})_{i=1,j=1}^{p,a}$ independent of $(G_{i,1})_{i=1}^p$ with mixed second moments given by

$$\mathbb{E} g_{i,j}^{(d_1)} g_{i',j'}^{(d_1)} = \frac{(\mathbb{E} x^{(d_1)} x^{(d_2)})^2}{\theta_j \theta_{j'} (\theta_j \theta_{j'} - \mathbb{E} x^{(d_1)} x^{(d_2)})} U_{i,i'}^{(d_1),(d_2)} V_{i,i'}^{(d_1),(d_2)}. \tag{3.1}$$

Then

$$(S_{i,j})_{i=1,j=1}^{p,a} \Rightarrow (F_{i,j})_{i=1,j=1}^{p,a}$$

where

$$F_{i,j} := \frac{G_{i,1}}{\theta_j} + g_{i,j}. \tag{3.2}$$

To prove the lemma, we split $S_{i,j}$ into three sums as follows. Fix cutoffs $m > 0$ and $T_n = \log^2 n$ ($T_n = \omega(\log n)$ suffices) and define

$$\begin{aligned} S_{i,j} &= \sum_{k=1}^{m} \frac{Z_{i,k}}{\lambda_j^k} + \sum_{k=m+1}^{T_n} \frac{Z_{i,k}}{\lambda_j^k} + \sum_{k>T_n} \frac{Z_{i,k}}{\lambda_j^k} \\ &=: S_{i,j}^A + S_{i,j}^B + S_{i,j}^C. \end{aligned}$$

We define

$$T_{i,j}^A := \sum_{k=1}^{m} \frac{G_{i,k}}{\theta_j^k} \tag{3.3}$$

where the $G_{i,k}$ are defined as in the statement of Proposition 1. Note that $T_{i,j}^A$ is independent of n.

By Proposition 1 and the multivariate version of Slutsky's theorem (see [Bil99]),

$$((Z_{i,k}), (\lambda_j)) \Rightarrow ((G_{i,k}), (\theta_j)),$$

where the joint convergence is over all $i \in [p]$, $k \in [m]$ and $j \in [a]$. By the continuous mapping theorem, $(S_{i,j}^A) \Rightarrow (T_{i,j}^A)$ jointly for $i \in [p]$ and $j \in [a]$. By the definitions of $T_{i,j}^A$ in (3.3) and of $G_{i,k}$ in (2.3) and (2.4), and by inspecting (3.1) and (3.2), we see that

$$T_{i,j}^A \overset{m\to\infty}{\Longrightarrow} F_{i,j}$$

jointly.

To prove Lemma 7, it suffices to prove

Lemma 8. *(a)* $\lim_{m\to\infty} \lim_{n\to\infty} \mathbb{E}|S^B| = 0$ *and*

(b) $\lim_{m\to\infty} \lim_{n\to\infty} \mathbb{E}|S^C| = 0.$

where we have suppressed the i *and* j *dependence for* $S_{i,j}^B$ *and* $S_{i,j}^C$.

Define the event

$$E_n := \{|\lambda_{n,j} - \theta_j| < \delta_j := \frac{|\theta_j| - 1}{4} \text{ for all } j \in [a]\}. \tag{3.4}$$

By hypothesis $\mathbb{P}(E_n) = 1 - o(1)$ so it suffices to prove Lemma 7 (and hence Lemma 8) on E_n. In the following, we fix an index j and set $\delta := \frac{|\theta|-1}{4}$. Note that we have

$$|\lambda| > 1 + \frac{3}{4}(|\theta| - 1). \tag{3.5}$$

We prove Lemma 8b first.

Proof. Recall that on E_n, $|\lambda| > 1+3\delta$ (see (3.4)). By Theorem 3, with probability $1-o(1)$ $\rho(X/\sqrt{n}) < 1+\delta$ and we can choose l such that $\|(\frac{X}{\sqrt{n}})^l\|^{1/l} < 1+2\delta$. We may assume without loss of generality that these events occur on E_n. By submultiplicativity of the operator norm,

$$\begin{aligned}\left\|\left(\frac{X}{\sqrt{n}}\right)^k\right\| &\leq \left\|\left(\frac{X}{\sqrt{n}}\right)^l\right\|^{\lfloor \frac{k}{l} \rfloor} \max_{0\leq i<l}\left\|\left(\frac{X}{\sqrt{n}}\right)^i\right\| \\ &\leq O_l(1+2\delta)^k \text{w.h.p.}\end{aligned}$$

By the Cauchy-Schwarz inequality, we have

$$\begin{aligned}|S_C| &\leq \sum_{k>T_n} \frac{\sqrt{n}\left\|\left(\frac{1}{\sqrt{n}}X\right)^k\right\| |u|_2|v|_2}{|\lambda|^k} \\ &< O_l(\sqrt{n}) \sum_{k>T_n}\left(\frac{1+2\delta}{1+3\delta}\right)^k \\ &= o(1)\end{aligned}$$

where the last line follows from our choice of $T_n = \log^2 n$. □

To prove Lemma 8a, we will need

Lemma 9. *Let u and v be unit vectors in $\mathbb{C}^n$ and set*

$$Z_k := \sqrt{n}u^*\left(\frac{1}{\sqrt{n}}X\right)^k v.$$

Fix $\epsilon > 0$ and assume $|x| \leq K = O(n^{\frac{1-\epsilon}{2}})$. Then there exists $c = c(\epsilon) > 0$ such that for all $k \ll n^c$,

$$\mathbb{E}|Z_k|^2 = O(1). \tag{3.6}$$

Assuming Lemma 9 we prove Lemma 8a on E_n. Since $(\mathbb{E}|Z|)^2 \leq \mathbb{E}|Z|^2$, we have

$$\begin{aligned}\mathbb{E}|S^B| &\leq \sum_{k=m+1}^{T_n} \mathbb{E}\frac{|Z_k|}{|\lambda|^k} \\ &\ll |1+\frac{3}{4}(|\theta|-1)|^{-2m}\end{aligned}$$

where we have used Lemma 9 and (3.5) in the last line. Lemma 8(a) follows from letting $m \to \infty$.

Remark 6. *Note that by the truncation argument given in Appendix B, Lemma 9, and hence Lemma 8a, is valid under the moment hypothesis* $\mathbb{E}|x|^{4+\epsilon} < \infty$ *for any fixed* $\epsilon > 0$.

3.1 Proof of Lemma 9

In this section we prove Lemma 9.

Proof. It suffices to show

$$\mathbb{E}|u^* X^k v|^2 = O(n^{k-1}). \tag{3.7}$$

Let

$$T := \{(a,b) : a = 1,2, b = 0,1,\dots,k\},$$

$$T' := \{(a,b) \in T : b < k\}$$

and

$$E := \{((a,b),(a,b+1)) \in T^2 : b < k\}.$$

Let $T_P := T|_{a=1}$, $T_Q := T|_{a=2}$ and for $t \in T'$, set $t^s := (a, b+1)$. We will designate the terms in the expansion of (3.7) by

$$\mathcal{P}' := \{F : T \to [n]\}.$$

For $F \in \mathcal{P}'$, let $F_P := F|_{T_P}$ and $F_Q := F|_{T_Q}$. Let

$$W_{u,v}(F) := |u_{F(1,0)} u_{F(2,0)} v_{F(1,k)} v_{F(2,k)}|$$

and

$$W_x(F) := \mathbb{E}|\prod_{t \in T'} x_{F(t),F(t^s)}|.$$

Then we have

$$\mathbb{E}|u^* X^k v|^2 \le \sum_{F \in \mathcal{P}'} W_{u,v}(F) W_x(F). \tag{3.8}$$

For $F \in \mathcal{P}'$, let

$$E_F := \{(F(t), F(t^s)) \in [n]^2 : t \in T'\}.$$

denote the edges of F and let

$$\mathcal{E}_F := \{\{t \in T' : (F(t), F(t^s)) = (i,j)\} : (i,j) \in E^F\}.$$

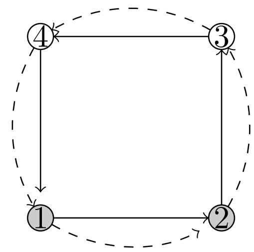


Figure 3.1: An example of $F \in \mathcal{P}$ with $k = 4$.

Then

$$W_x(F) = \prod_{e \in \mathcal{E}^F} \mathbb{E}|x|^{|e|}.$$

Noting that $\mathbb{E}|x| = 0$ and letting

$$\mathcal{P} := \{F \in \mathcal{P}' : |e| \geq 2 \text{ for all } e \in \mathcal{E}^F\},$$

we have

$$\mathbb{E}|u^* X^k v|^2 \leq \sum_{F \in \mathcal{P}} W_{u,v}(F) W_x(F). \tag{3.9}$$

Now, for a fixed $F \in \mathcal{P}$, let

$$V = V_F := \{F(t) : t \in T\}$$

be the set of vertices. For $v \in V$ let $m(v) = |F^{-1}(v)|$ denote its multiplicity. Let $d_{\text{in}}(v) := |\{x \in [n] : (x, v) \in E\}|$ and $d_{\text{out}}(v) := |\{x : (v, x) \in E\}|$ denote its indegree and outdegree. Finally, let $d(v) := d_{\text{in}}(v) + d_{\text{out}}(v)$ be the (total) degree of v.

Shown in Figure 3.1 is an example with $k = 4$ with the paths $(1, 2, 3, 4)$ and $(2, 3, 4, 1)$. Each vertex has indegree 2 and outdegree 2.

We will first determine the main term from $\mathcal{P}$ and its contribution to (3.9).

Lemma 10. *Suppose $F \in \mathcal{P}$. Then $|V| \leq k+1$ and that equality occurs only when $F_P = F_Q$ and $|F_P| = |F_Q| = k + 1$.*

Fix $v \in V_F$ and suppose $d(v) = 1$. Since each edge has multiplicity at least two, we have the following.

(i) If $d_{\text{in}}(v) = 1$, $v = F(1,0) = F(2,0)$.

(ii) If $d_{\text{out}}(v) = 1$, $v = F(1,k) = F(2,k)$

In particular, if two vertices of V have degree 1, then one has outdegree 1, the other has indegree 1 and the rest have both outdegree and indegree of at least 1. Since $|e| \geq 2$ for each $e \in \mathcal{E}_F$, we also have $|E_F| \leq k$. Thus

$$\begin{aligned} 2k \geq 2|E_F| &= \sum_{v \in V} d(v) \\ &\geq 1 + 1 + 2(|V| - 2) \\ &= 2(|V| - 1). \end{aligned}$$

Thus, $|V| \leq k + 1$ with equality occurring only when two of the vertices have degree 1 and the rest have degree 2. This proves the lemma.

We let $\mathcal{P}_{\text{main}} := \{F \in \mathcal{P} : |V_F| = k + 1\}$. We also let

$$\mathcal{P}'_0 := \{F \in \mathcal{P} : |V_F| = k, F_P = F_Q, F(1,0) = F(2,0) = F(1,k) = F(2,k)\}.$$

Then, the contribution of $\mathcal{P}_{\text{main}} \cup \mathcal{P}'_0$ to (3.9) is given by

$$\sum_{\substack{F(1,0)=F(2,0)\in[n] \\ F(1,k)=F(2,k)\in[n]}} |u_{F(1,0)}|^2 |u_{F(1,k)}|^2 n^{k-1} = n^{k-1}.$$

We partition the remainder of $\mathcal{P}$ in the following way. First let

$$T_1 := \{(1,0), (2,0), (1,k), (2,k)\} \subset T$$

be the terms corresponding to the starts and ends of the paths. For $t \geq 0$ and P a partition T_1 with $|P| \geq 2$ if $t = 0$, let

$$\mathcal{P}_{P,t} := \{F \in \mathcal{P} : |V_F| = k - t, F(s) = F(t) \Leftrightarrow s \sim_P t, s, t \in T_1\}.$$

Note that we exclude the trivial partition $P = \{T_1\}$ when $t = 0$ since $\mathcal{P}_{\{T_1\},0} = \mathcal{P}'_0$. We let $\mathcal{P}_0 = \bigcup_{P \neq \{T_1\}} \mathcal{P}_{P,0}$ and for $t > 0$, we let $\mathcal{P}_t = \bigcup_P \mathcal{P}_{P,t}$.

Lemma 11. *For $F \in \mathcal{P}_t$, $W_x(F) \ll K^{2t}$.*

Since $\mathbb{E}|x|^a \ll K^{(a-4)_+}$,

$$W_x(F) \le \prod_{e\in\mathcal{E}_F} \mathbb{E}|x|^{|e|} \ll \prod_{e\in\mathcal{E}_F} K^{(|e|-4)_+}.$$

It suffices to show that $\sum_{e\in\mathcal{E}_F}(|e|-4)_+ \le 2t$. Since at most one vertex has no outgoing edge, $|\mathcal{E}_F| \ge k-t-1$. Also the $|e|$'s satisfy $\sum_{e\in\mathcal{E}_F}|e| = 2k$ and $|e|\ge 2$. If $|e|\le 4$ for all $e\in\mathcal{E}_F$, there is nothing to prove. If $|e_1|\ge 4$ say, then

$$\begin{aligned}\sum_{e\in\mathcal{E}_F}(|e|-4)_+ &= |e_1|-4+\sum_{e\neq e_1}(|e|-4)_+\\ &\le \sum_{e\in\mathcal{E}_F}(|e|-2)-2\\ &\le 2k-2(k-t-1)-2 = 2t.\end{aligned}$$

We now turn to controlling $S_{p,t} := |\sum_{F\in\mathcal{P}_{P,t}} W_{u,v}(F)|$. To simplify notation, we will do this for the specific case $P = \{\{(1,0)\},\{(2,0)\},\{(1,k),(2,k)\}\}$. We can bound $S_{P,t}$ by

$$\sum_{\substack{i_1,i_2,i_3\in[n]\\ \text{distinct}}} |u_{i_1}u_{i_2}v_{i_3}^2||\{F\in\mathcal{P}_t : F(1,0)=i_1, F(2,0)=i_2, F(1,k)=F(2,k)=i_3\}|.$$

The cardinality of the last set is independent of the choice of indices i_1, i_2 and i_3, and in fact only depends on size of the partition P. We denote it by $N_{|P|}$. Removing the restriction to distinct indices and using $\sum_i |u_i| = O(\sqrt{n})$, we may bound the contribution as $nN_{|P|}$.

The case for a general partition is similar and we have the bound

$$S_{P,t} \le n^{c_P/2} N_{|P|}$$

where c_P is the number of singletons in the partition P. To determine $N_{|P|}$, we first choose the remaining vertices of V_F in $\binom{n}{k-t-|P|}$ ways. We let $N_2 = N_2(t)$ be the maximum number of ways to choose E_F, over P and V_F. Similarly, we let $N_3 = N_3(t)$ be the maximum number of ways to choose $\mathcal{E}_F$, over P, V_F and E_F. Since

$$\binom{n}{k-t-|P|} \le \frac{n^{k-t-|P|}k^{|P|}}{(k-t)!},$$

we have

$$S_{P,t} \le (n^{c_P/2-|P|}k^{|P|})\frac{n^{k-t}}{(k-t)!}N_2(t)N_3(t),$$

with $|P| \ge 2$ if $t = 0$. Considering the possibilities for P and setting

$$S_t := \sum_{P \text{ partition of } T_1} S_{P,t},$$

we have

$$S_t \ll \begin{cases} k^2 n^{k-3/2} N_2 N_3 / k! & : \quad t = 0 \\ k n^{k-t-1} N_2 N_3 / (k-t)! & : \quad t \ge 1 \end{cases} . \tag{3.10}$$

We now estimate $N_2 = N_2(t)$, the number of ways to choose the set of edges E_F for $F \in \mathcal{P}_t$. As observed earlier, at least $k - t - 1$ vertices have positive outdegree, and similarly for the indegree. We need to assign at most k oriented edges to the $k - t$ vertices such that these conditions are met. Recall $d_{\text{out}}(i)$ to be the outdegree of vertex i. We will allow for repetitions when choosing the edges to include graphs with less than k edges. Hence we may impose the constraint $\sum_{i=1}^{k-t} d_{\text{out}}(i) = k$. For at least $k - t - 1$ vertices, $d_{\text{out}}(i) \ge 1$. This gives[1] $\binom{k}{t+1}$ ways of choosing the outdegrees $(d_{\text{out}}(i))_{i=1}^{k-t}$. To assign the incoming edges of the vertices, we partition the k edges into $k - t$ nonempty parts $(E_i)_{i=1}^{k-t}$. We first choose $k-t$ edges to belong to the different E_i's and then we choose parts for each of the remaining edges. This can be done in at most $\binom{k}{t}(k-t)^t$ ways. Finally, we assign the $k-t$ parts to the vertices with positive indegree. If all $k - t$ vertices have incoming edges, there are at most $(k - t)!$ ways to assign each of them an E_i. Now suppose only $k - t - 1$ of the vertices have incoming edges. First, there are at most $(k - t)^4$ ways to choose 2 vertices and 2 parts, with one vertex being assigned both parts and the other having no incoming edges. Next, there are $(k - t - 2)!$ ways of assigning the remaining parts to the remaining vertices. Hence

$$\begin{aligned} N_2 &\le \binom{k}{t+1}\binom{k}{t}(k-t)^t((k-t)! + (k-t)^4(k-t-2)!) \\ &\le \frac{k^{2t+1}}{(t+1)!t!}k^t(k-t)!k^2 \\ &\le \frac{k^{3t+3}(k-t)!}{(t+1)!t!}. \end{aligned} \tag{3.11}$$

[1]This follows from the standard *stars and bars* combinatorial argument; see [Fel50].

We now estimate $N_3 = N_3(t)$, the number of ways of choosing $\mathcal{E}_F$ once V_F and E_F have been chosen. Since each vertex has at least one outgoing edge, the maximum outdegree of any vertex is at most $t+1$. On the other hand, since $d_{\text{out}}(1) + \ldots + d_{\text{out}}(k-t) \le k$, at least $\max(k-2t, 0)$ vertices have $d_{\text{out}}(i) = 1$. At least $\max(2k-4t, 0)$ legs start from these vertices so at most $4t$ legs begin at vertices with $d_{\text{out}}(i) > 1$. At each of these legs, we have at most $t+1$ choices to make when choosing the path. We thus have

$$N_3 \le (t+1)^{4t}, \tag{3.12}$$

which is independent of the chosen vertices and edges.

For $t = 0$, using (3.10), (3.11) and (3.12), we have

$$\begin{aligned} S_0 &\le n^{k-3/2} k^2 / k! N_2(0) N_3(0) \\ &\le k^5 n^{k-3/2}. \end{aligned}$$

Since $W_x(F) = O(1)$ for $t = 0$, the contribution to (3.9) is $o(n^{k-1})$.

For $t \ge 1$, we have

$$\begin{aligned} S_t K^{2t} &\ll n^{k-t-1} k N_2 N_3 K^{2t} / (k-t)! \\ &\le k^4 n^{k-1} \frac{k^{3t} (t+1)^{4t}}{n^{\epsilon t} (t+1)! t!} \\ &\le k^4 n^{k-1} \frac{k^{3t} (te)^{2t}}{n^{\epsilon t}}, \end{aligned}$$

where we have used the estimates $t! > \frac{t^t}{e^t}$ and $\frac{(t+1)^t}{t^t} \le e$. For $k = o(n^{\epsilon/5})$, the last expression is decreasing for $t \le k$ and bounding each term by the bound for the $t = 1$ term, we have

$$\sum_t S_t K^{2t} \ll k^7 n^{k-1-\epsilon} = o(n^{k-1})$$

for $k = o(n^{\epsilon/7})$.

□

CHAPTER 4

Proof of Theorem 5

Proof. We will work on the event

$$E = E_n = \{\rho(X) < 1+\epsilon, \lambda > 1+2\epsilon \text{ for all } \lambda \in \bigcup_{\theta\in\Theta} \Lambda^\theta\}$$

which occurs w.h.p. Fix $\theta \in \Theta$ and for $\lambda \in \Lambda^\theta$, let

$$R_\lambda := \left(\frac{X}{\sqrt{n}} - \lambda\right)^{-1}$$

denote the resolvent of $X/\sqrt{n}$. On E, $\lambda > \rho(X)$, so we may expand R_λ as a Neumann series

$$\begin{aligned} R_\lambda &= -\frac{1}{\lambda}\left(1 + \frac{1}{\sqrt{n}} \sum_{i\geq 1} \frac{X^i}{n^{(i-1)/2}\lambda^i}\right) \\ &=: -\frac{1}{\lambda}\left(1 + \frac{1}{\sqrt{n}} S_\lambda\right). \end{aligned}$$

We write the Jordan decomposition of A as $A = VJU^*$ where V (resp. U^*) is the $n \times \mathrm{rk}(A)$ (resp. $\mathrm{rk}(A) \times n$) matrix of generalized right (resp. left) eigenvectors of A associated to nonzero eigenvalues of A satisfying $U^*V = 1$ and J is the Jordan matrix of A restricted to nonzero eigenvalues with size $\mathrm{rk}(A) \times \mathrm{rk}(A)$. Starting with the eigenvalue equation $\det(\frac{X}{\sqrt{n}} + A - \lambda) = 0$ and using the determinant identity $\det(1+AB) = \det(1+BA)$, we have

$$\begin{aligned} \det\left(\frac{X}{\sqrt{n}} + A - \lambda\right) = 0 &\Rightarrow \det\left(1 + R_\lambda A\right) = 0 \\ &\Rightarrow \det\left(1 - \frac{1}{\lambda} U^* \left(1 + \frac{1}{\sqrt{n}} S_\lambda\right) VJ\right) = 0 \\ &\Rightarrow \det\left(-\lambda + J + \frac{1}{\sqrt{n}} U^* S_\lambda VJ\right) = 0. \end{aligned}$$

Let J_θ be the block matrix of J corresponding to eigenvalue θ and let U_θ^* and V_θ be the restrictions of U^* and V to the generalized left and right eigenvectors of θ respectively. Recall

Proposition 2 as well as the notation used therein. We apply Proposition 2 with $M = J_\theta$ and $P = P^\theta = \frac{1}{\sqrt{n}} U_\theta^* S_\lambda V_\theta J_\theta$.

First note that for each column indexed by $t \in I_v^\theta$, $J_\theta e_t = \theta e_t$, where e_t is the coordinate vector corresponding to t. Hence for $s \in I_u^\theta$ and $t \in I_v^\theta$,

$$P_{st}^\theta = \frac{1}{\sqrt{n}} \theta u_s^* S_\lambda v_t.$$

Observe that the moment assumption made in Theorem 5 guarantees the applicability of Lemma 7 to the collection

$$\{\sqrt{n} P_{st}^\theta : s \in I_u^\theta, t \in I_v^\theta, \theta \in \Theta\}.$$

By Lemma 7, $(\sqrt{n} P_{st}^\theta)_{s,t,\theta} \Rightarrow (F_r)_{r \in I_2}$ defined by (1.11), (1.6) and (1.12). Finally, applying Proposition 2 yields the procedure to determine the fluctuations as specified in Theorem 5. $\square$

Appendix A

Deterministic perturbations

In this appendix we state the deterministic perturbation result referred to in the proof of Theorem 5. It is originally attributed to Lidskii. See [MBO97] and references cited within. We remind the reader that we denote the Schur complement of A in the block matrix $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$, $D - CA^{-1}B$ by $SC(A, \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right))$.

Proposition 2. *Let M be a $d \times d$ deterministic matrix in Jordan form. For notational simplicity, we will assume M has a single eigenvalue θ. Let*

$$J_k := \begin{bmatrix} \theta & 1 & & \\ & \theta & 1 & \\ & & \ddots & 1 \\ & & & \theta \end{bmatrix}$$

denote the $k \times k$ Jordan block and write

$$M = \bigoplus_{k=1}^{K} J_k^{\oplus m_k}$$

Thus for each $k \in [K]$, M has m_k Jordan blocks J_k. Let P_n be a sequence of $d \times d$ perturbation matrices with entries of size $o(1)$. Then $M + P_n$ has spectrum

$$\Lambda(M + P_n) = \{\lambda_{k,m,i} : k \in [K], m \in [m_k], i \in [k]\}$$

with $\lambda_{k,m,i} \to \theta$ for all $k \in K$, $m \in [m_k]$ and $i \in [k]$. The fluctuations

$$f_{k,m,i} := \lambda_{k,m,i} - \theta$$

are given by the following procedure.

Let $c_k := \sum_{j=1}^k m_j$ and set $c := c_K$. Decompose $P = P_n$ into c^2 blocks $(B_{ij})_{i,j=1}^c$ with the c diagonal blocks $(B_{i,i})_{i=1}^c$ having sizes

$$1, \ldots, 1, 2. \ldots, 2, \ldots, K, \ldots K$$

with k occurring with multiplicity m_k. Let $k_i \times k_j$ denote the size of block $B_{i,j}$. This block decomposition is conformal with that of M induced by the J_k's. Let $R = R_n$ be the submatrix of P of size $c \times c$ with entries given by

$$R_{ij} = (B_{ij})_{k_i 1}.$$

Hence R is formed from the lower left elements of the blocks in the decomposition of P.

Let $E_k = R_{c_k \times c_k}$ be upper left submatrices of R and let $F_k := \mathrm{SC}(E_{k-1}, E_k)$ be the $m_k \times m_k$ Schur complement of E_{k-1} in E_k, where we set $F_1 := E_1$. Then, to leading order, the fluctuations $f_{k,m,i}$ are given by the k k-th roots of the m_k eigenvalues of F_k for each $k \in [K]$. If M has multiple eigenvalues, we apply the above procedure to each eigenvalue separately.

We remark on a few special cases of Proposition 2. We denote the entries of $P = P_n$ by p_{ij} and assume $p = O(\frac{1}{\sqrt{n}})$ (as will turn out to be the case in our applications).

1. Suppose $M = \mathrm{diag}(\theta_1, \ldots, \theta_d)$ is diagonal with distinct eigenvalues. Let λ_j denote the corresponding eigenvalues of $M + P$ in the sense that $\lambda_j \to \theta_j$ as $n \to \infty$ Then

$$f_j := \lambda_j - \theta_j = p_{jj}(1 + o(1)).$$

2. Suppose $M = \theta I_d$. Then $\{\sqrt{n}(\lambda_j - \theta)\}_{j=1}^d$ converge to the d eigenvalues of $\sqrt{n}P$.

3. Suppose $M = J_d(\theta)$. Then $\{n^{\frac{1}{2d}}(\lambda_j - \theta)\}_{j=1}^d$ converge to the d roots of $\sqrt{n}P_{k1}$.

We now present a special case of the proof of Theorem 2. The presentation below is adapted from the proof of [MBO97, Theorem 2.1] and illustrates the key features of Theorem 2. The proof begins with a change of variables to arrive at the appropriately normalized fluctuation.

One then pre- and post-multiplies the perturbed matrix by suitable diagonal matrices in order to isolate the leading order terms. After performing some elementary operations, evaluating the relevant determinant yields the result.

We set ϵ to be a small parameter, $\epsilon = n^{-1/2}$ is sufficient for our purposes. We consider a 5×5 matrix A in Jordan normal form, with a single eigenvalue θ defined by

$$A := \left[\begin{array}{cc|cc|c} \theta & 1 & & & \\ & \theta & & & \\ \hline & & \theta & 1 & \\ & & & \theta & \\ \hline & & & & \theta \end{array}\right] = J_2^{\oplus 2} \oplus J_1.$$

Let B be a fixed complex matrix of same size, thus $B := (b_{ij})_{1 \le i,j \le 5} \in \mathbb{C}^{5\times 5}$. Finally let $M = M(\epsilon) := A + \epsilon B$ be the perturbation matrix. For $\lambda = \lambda(\epsilon)$ an eigenvalue of M we have the eigenvalue equation

$$\det(\lambda - (A + \epsilon B)) = 0.$$

We first find the perturbations $\lambda - \theta$ to leading order in ϵ that correspond to the eigenvalues of $J_2^{\oplus 2}$. It turns out that these 4 fluctuations are of order $\epsilon^{1/2}$. Thus we make the change of variables

$$z = \epsilon^{1/2}, \quad \mu = \frac{\lambda - \theta}{z}$$

and write $\lambda - (A + \epsilon B) = \mu z + \theta - A - z^2 B =: P(\mu, z)$. The matrix $\mu z + \theta - A$ has entries μz on the diagonal and -1's in some of the entries of the superdiagonal. We wish to transform this to a matrix with μ's on the diagonal while preserving the property that the entries of $\mu z + \theta - A$ and $z^2 B$ are polynomial in z. This we accomplish by defining $L_2 := \operatorname{diag}(z^{-1}, z^{-2}, z^{-1}, z^{-2}, z^{-1})$, $R_2 = \operatorname{diag}(1, z, 1, z, 1)$ and setting $F_2 = L_2 P R_2$. We then have

$$F_2(\mu,z) = \left[\begin{array}{cc|cc|c} \mu & -1 & 0 & 0 & 0 \\ -b_{21} & \mu & -b_{23} & 0 & 0 \\ \hline 0 & 0 & \mu & -1 & 0 \\ -b_{41} & 0 & -b_{43} & \mu & -1 \\ \hline 0 & 0 & 0 & 0 & \mu \end{array}\right] + O(z)$$

$$=: G(\mu) + O(z).$$

To first order, the fluctuations μ are given by the roots of $\det(G_2(\mu))$. Denoting the columns of $G_2(\mu)$ by $(C_i)_{i=1}^5$ and performing the operations $C_1 \leftarrow C_1 + C_2$, $C_3 \leftarrow C_3 + C_4$, we arrive at the matrix

$$Q_2(\mu) := \left[\begin{array}{cc|cc|c} 0 & -1 & 0 & 0 & 0 \\ -b_{21}+\mu^2 & \mu & -b_{23} & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 0 \\ -b_{41} & 0 & -b_{43}+\mu^2 & \mu & 0 \\ \hline 0 & 0 & 0 & 0 & \mu \end{array}\right].$$

Let $E_1 := \left(\begin{smallmatrix} b_{21} & b_{23} \\ b_{41} & b_{43} \end{smallmatrix}\right)$; the notation is chosen to coincide with that in Theorem 2. Observing that rows 1, 3 and 5 have a single nonzero element in positions (12), (24) and (35) respectively, we have

$$\det(Q_2(\mu)) = \mu \det(\mu^2 - B_2)$$

and thus μ^2 is an eigenvalue of E_1, to first order.

The procedure of finding the perturbed eigenvalue arising from the singleton J_1 block is similar; we sketch the differences. We now make the change of variables

$$\mu = \frac{\lambda - \theta}{\epsilon}$$

and let

$$L_1 := \operatorname{diag}(1, \epsilon^{-1}, 1, \epsilon^{-1}, \epsilon^{-1}), \quad R_1 = I_5.$$

As before, we define

$$
\begin{aligned}
F_1(\mu,\epsilon) &:= L_1(\mu\epsilon + \theta - A - \epsilon B)R_1 \\
&= \left[\begin{array}{cc|cc|c}
0 & -1 & 0 & 0 & 0 \\
-b_{21} & \mu - b_{22} & -b_{23} & -b_{24} & -b_{25} \\
\hline
0 & 0 & 0 & -1 & 0 \\
-b_{41} & -b_{42} & -b_{43} & \mu - b_{44} & -b_{45} \\
\hline
-b_{51} & -b_{52} & -b_{53} & -b_{54} & \mu - b_{55}
\end{array}\right] + O(\epsilon) \\
&=: G_1(\mu) + O(\epsilon).
\end{aligned}
$$

Setting

$$
E_2 := \begin{bmatrix} b_{21} & b_{23} & b_{25} \\ b_{41} & b_{43} & b_{45} \\ b_{51} & b_{53} & b_{55} \end{bmatrix}.
$$

we have

$$
\det(G_1(\mu)) = \det\left(\left[\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mu \end{smallmatrix}\right] - E_2\right)
$$

and $\mu = \mathrm{SC}(E_1, E_2)$ to first order.

Appendix B

Truncation

In this appendix, we extend the results involving the moment method, namely Proposition 1 and Lemma 7 using a truncation argument (see [BS06a]). We consider the two assumptions

(i) $|x| \leq K = O(n^M)$.

(ii) $\mathbb{E}|x|^m < \infty$, $m = 2/M$.

We show that if Proposition 1 and Lemma 9 hold for (i) with $M < 1/2$, then they hold for (ii).

Suppose we have (ii) with $m > 4$, corresponding to $M = 2/m < 1/2$. We first show that the event

$$\{|x_{ij}| \leq n^M \text{ for all } i, j \in [n]\}$$

occurs w.h.p. Indeed, we have

$$\mathbb{P}\left[|x_{ij}| \geq n^M \text{ some } i, j \in [n]\right] \leq n^2 \mathbb{P}\left[|x|^m \geq n^2\right]. \tag{B.1}$$

Since $n^2 \mathbb{1}_{|x|^m \geq n^2} \leq |x|^m$ and $\mathbb{E}|x|^m < \infty$, the last expression converges to 0 by the dominated convergence theorem.

Now define the truncated random variables $\hat{x} := x\mathbb{1}_{|x| \leq n^M}$ and $\hat{X} = (\hat{X})_{ij}$ by $\hat{X}_{ij} := \hat{x}_{ij}$. While $\hat{x}$ is bounded, it no longer has mean zero. On the other hand, for n sufficiently large, we have

$$\begin{aligned}|\mathbb{E}\hat{x}| &\leq \mathbb{E}|x\mathbb{1}_{|x|\geq n^M}| \\ &\leq \frac{\mathbb{E}|x|^m\mathbb{1}_{|x|\geq n^M}}{n^{(m-1)M}} \\ &\ll n^{-(m-1)M} \\ &\leq n^{-3/2}. \end{aligned} \tag{B.2}$$

By Schur's test for the operator norm of a matrix, we have

$$\|\mathbb{E}\hat{X}\| = O(n^{-1/2}). \tag{B.3}$$

Now let $\tilde{x} := \hat{x} - \mathbb{E}\hat{x}$ and $\tilde{X} := \hat{X} - \mathbb{E}\hat{X}$ denote the truncated and centered random variables. By construction, $\mathbb{E}\tilde{x} = 0$. Furthermore,

$$\mathbb{E}|\tilde{x}|^2 = \mathbb{E}|\hat{x}|^2 - |\mathbb{E}\hat{x}|^2 \to \mathbb{E}|x|^2 = 1 \tag{B.4}$$

by (B.2) and dominated convergence. Given (B.4), it is easy to check that under (i), Proposition 1 is valid for $\tilde{X}$. Since $\mathbb{E}|\tilde{x}|^2 \leq \mathbb{E}|x|^2$, Lemma 7 also valid for $\tilde{x}$. To prove the validity of Proposition 1 and Lemma 7 for x under (ii), it suffices to prove the following.

Lemma 12. *Suppose $u = u_n$ and $v = v_n$ are unit vectors in $\mathbb{C}^n$. Then for every $\gamma > 0$, the event*

$$A_{n,\gamma} := \bigcup_{k\leq \log^2 n} \{|u^*\hat{X}^k v - u^*\tilde{X}^k v| > \gamma n^{(k-1)/2}\}$$

occurs w.h.p.

We first state a result that is a consequence of the proof in [BY86]. Following the notation of [BY86] we define $\delta := n^{M-1/2}$ so that $|\tilde{x}| \leq \delta\sqrt{n}$. Fix $z > k+1$ and p a positive integer. Then

$$\begin{aligned}\mathbb{P}\left[\left\|\left(\frac{1}{\sqrt{n}}\tilde{X}\right)^k\right\| \geq z\right] &\leq z^{-2p}n^{-pk}\mathbb{E}\,\mathrm{Tr}\left(\tilde{X}^k\left(\tilde{X}^k\right)^*\right)^p \\ &=: z^{-2p}n^{-pk}E_n.\end{aligned}$$

In [BY86](pg. 561), it is shown that

$$E_n \leq n^{kp+\frac{3}{2}} \sum_{l=1}^{pk} \binom{2kp}{2l} (k+1)^{2kp-2l+2p} (2kp) \left(\frac{6kp\delta^{1/6}}{\log \frac{\delta\sqrt{n}}{(2kp)^3}} \right)^{6kp-6l} \delta^{kp-l}.$$

In our application, $k \leq \log^2 n$ and choosing $p = \delta^{-1/7}$ say, we have

$$\frac{6kp\delta^{1/6}}{\log \frac{\delta\sqrt{n}}{(2kp)^3}} \to 0. \tag{B.5}$$

In fact, the left-hand side of (B.5) is less than 1 for $n \geq N(m)$.

For such n, following [BY86](pg. 562), it then follows that

$$z^{-2p} n^{-pk} E_n \leq \left((2kpn^2)^{1/p} (1 + (k+1)\delta^{1/2})^{2k} \left(\frac{k+1}{z} \right)^2 \right)^p.$$

Choosing $z = 3k$ say, for any $k \leq \log^2 n$ we have

$$\mathbb{P}\left[\left\| \left(\frac{1}{\sqrt{n}} \tilde{X} \right)^k \right\| \geq 3k \right] = O(e^{-n^c}) \tag{B.6}$$

for some $c = c(m) > 0$. We now turn to the proof of Lemma 12.

Proof. By (B.6), we may assume $\left\| \left(\frac{1}{\sqrt{n}} \tilde{X} \right)^k \right\| \leq 3k$ for all $k \leq \log^2 n$ which occurs w.h.p. We will need the crude bound

$$\sum_{\substack{a_1+\ldots+a_k=n \\ a_i \geq 0}} \prod_{i=1}^{k} a_i \leq n^{2k}. \tag{B.7}$$

We then have

$$\begin{aligned}
\frac{1}{n^{(k-1)/2}} |u^* \hat{X}^k v - u^* \tilde{X}^k v| &\leq n^{-(k-1)/2} \|(\tilde{X} + \mathbb{E}\hat{X})^k - \tilde{X}^k\| \\
&\leq \sum_{l=1}^{k} \frac{1}{n^{(l-1)/2}} \sum_{l'=0}^{l+1} \sum_{a_1+\ldots+a_{l'}=k-l} \prod_{i=1}^{l'} \left\| \left(\frac{\tilde{X}}{\sqrt{n}} \right)^{a_i} \right\| \|\mathbb{E}\tilde{X}\|^l \\
&\leq \sum_{k=1}^{l} \frac{1}{n^{l-1/2}} \sum_{l'=0}^{l+1} (3k^2)^{l'} \\
&\leq \sum_{l=1}^{k} \frac{l+1}{n^{(l-1)/2}} \left(\frac{3k^2}{\sqrt{n}} \right)^{l+1} = o(1),
\end{aligned}$$

where we have used (B.7) and (B.3) in the third line.

□

References

[Bai99] Z. D. Bai. "Methodologies in spectral analysis of large-dimensional random matrices, a review." *Statist. Sinica*, **9**(3):611–677, 1999.

[BBP05] J. Baik, G. Ben Arous, and S. Péché. "Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices." *Ann. Probab.*, **33**(5):1643–1697, 2005.

[BC14] C. Bordenave and M. Capitaine. "Outlier eigenvalues for deformed i.i.d. random matrices." *arXiv*, **math.PR**(1403.6001v2), 2014.

[Bil99] P. Billingsley. *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[BS06a] Z. D. Bai and J. Silverstein. *Spectral analysis of large dimensional random matrices.* Mathematics Monograph Series **2**. Science Press, Beijing, 2006.

[BS06b] J. Baik and J. W. Silverstein. "Eigenvalues of large sample covariance matrices of spiked population models." *J. Multivariate Anal.*, **97**(6):1382–1408, 2006.

[BY86] Z. D. Bai and Y. Q. Yin. "Limiting behavior of the norm of products of random matrices and two problems of Geman-Hwang." *Probab. Theory Relat. Fields*, **73**:555–569, 1986.

[CDF09] M. Capitaine, C. Donati-Martin, and D. Féral. "The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations." *Ann. Probab.*, **37**(1):1–47, 2009.

[CDF12] M. Capitaine, C. Donati-Martin, and D. Féral. "Central limit theorems for eigenvalues of deformations of Wigner matrices." *Ann. Inst. Henri Poincaré Probab. Stat.*, **48**(1):107–133, 2012.

[Fel50] W. Feller. *An Introduction to Probability Theory and Its Applications, Vol 1.* Wiley, 2nd ed edition, 1950.

[Iss18] L. Isserlis. "On a Formula for the Product-Moment Coefficient of any Order of a Normal Frequency Distribution in any Number of Variables." *Biometrika*, **12**(1/2):134–139, 1918.

[KY13] A. Knowles and J. Yin. "The isotropic semicircle law and deformation of Wigner matrices." *Comm. Pure Appl. Math.*, **66**(11):1663–1750, 2013.

[KY14] A. Knowles and J. Yin. "The outliers of a deformed Wigner matrix." *Ann. Probab.*, **42**(5):1980–2031, 2014.

[Lid66] V. B. Lidskiĭ. "On the theory of perturbations of nonselfadjoint operators." *Ž. Vyčisl. Mat. i Mat. Fiz.*, **6**(1):52–60, 1966.

[MBO97] J. Moro, J. V. Burke, and M. L. Overton. “On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure.” *Siam J. Matrix Anal. Appl.*, **18**(4):793–817, 1997.

[MP67] V. A. Marčenko and L. A. Pastur. “Distribution of eigenvalues for some sets of random matrices.” *Mathematics of the USSR-Sbornik*, **1**(4):457–483, 1967.

[OR14] S. O’Rourke and D. Renfrew. “Low rank perturbations of large elliptic random matrices.” *Electron. J. Probab.*, **19**:no. 43, 65, 2014.

[PRS13] A. Pizzo, D. Renfrew, and A. Soshnikov. “On finite rank deformations of Wigner matrices.” *Ann. Inst. Henri Poincaré Probab. Stat.*, **49**(1):64–94, 2013.

[RB13] J. Rochet and F. Benaych-Georges. “Outliers in the single ring theorem.” *arXiv*, **math.PR**(1308.3064v4), 2013.

[RS13] D. Renfrew and A. Soshnikov. “On finite rank deformations of Wigner matrices II: Delocalized perturbations.” *Random Matrices Theory Appl.*, **2**(1):1250015, 36, 2013.

[Tao12] T. Tao. *Topics in random matrix theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

[Tao13] T. Tao. “Outliers in the spectrum of iid matrices with bounded rank perturbations.” *Probab. Theory Relat. Fields*, **155**(1-2):231–263, 2013.

[TV10] T. Tao and V. Vu. “Random matrices: universality of ESDs and the circular law.” *Ann. Probab.*, **38**(5):2023–2065, 2010. With an appendix by Manjunath Krishnapur.

[VL60] M. I. Višik and L. A. Ljusternik. “Solution of some perturbation problems in the case of matrices and self-adjoint or non-selfadjoint differential equations. I.” *Russian Math. Surveys*, **15**(3):1–73, 1960.

[Wig55] E. Wigner. “Characteristic vectors of bordered matrices with infinite dimensions.” *Ann. of Math. (2)*, **62**:548–564, 1955.

[Wis28] J. Wishart. “The generalised product moment distribution in samples from a normal multivariate population.” *Biometrica*, **20A**(1-2):32–52, 1928.