Spectrum of Random Band Matrices

By

Indrajit Jana

M.S. (Indian Institute of Science) 2012

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Alexander Soshnikov

Debashis Paul

Craig Tracy

Committee in Charge

2017

To *peace.*

Contents

Indrajit Jana
June 2017
Mathematics

Spectrum of Random Band Matrices

Abstract

Starting early twentieth century, random matrices played a crucial role in Physics, Statistics, Engineering and many other branches of science. In this thesis, I have considered spectral properties of random band matrices. It was known that the limiting spectral distribution of the symmetric random band matrices follow the *Semicircle law* after proper scaling. In this thesis, I have considered the fluctuations of the linear eigenvalue statistics of random band matrices. In other words, I have considered the fluctuations of the empirical spectral distribution of the symmetric random band matrices around the Semicircle distribution.

In the second part of the thesis, I have considered the limiting spectral distribution of the singular values of random band matrices plus a deterministic band noise matrix. And finally, I conjectured that the limiting spectral distribution of general random band matrices follow the *Circular law*.

Acknowledgments

I wish to express my gratitude to Prof. Alexander Soshnikov, my advisor, for his guidance during my PhD. I am grateful to him for his encouragement, patience, and many valuable mathematical and non-mathematical discussions.

I would like to thank the referees of our papers for their constructive comments. I am thankful to Prof. Debashis Paul and Prof. Craig Tracy for going through this thesis and making useful comments which have led to a significant improvement in the presentation. I am thankful to all those who contributed to develop this subject. And finally, I would like to thank my family and friends without whom this would be impossible.

CHAPTER 0

Notations

For convenience, we list the notations and abbreviations which are used in this writing.

(1) δ_x: *Degenerate probability measure at* x *i.e.,* $\delta_x(y) = \mathbf{1}_{\{x=y\}}$

(2) $|S|$: *Depending on the context, cardinality of the set* S *or magnitude of the complex number* S

(3) S° : *Interior of the set* S OR *If* S *is a random variable then* $S^\circ = S - \mathbb{E}[S]$

(4) $\bar{S}$ *Closure of the set* S

(5) Ω : *The underlying probability space*

(6) $\|X\|_q = [\mathbb{E}|X|^q]^{1/q}$ for a random variable X and $1 \le q < \infty$

(7) CLT: *Central Limit Theorem*

(8) ESD: *Empirical spectral distribution*

(9) i.i.d.: *Independently and identically distributed random variables*

(10) $\Re(z)$: *Real part of the complex number* z

(11) $\Im(z)$: *Imaginary part of the complex number* z

(12) $\mathbb{C}_+ := \{z \in \mathbb{C} : \Im(z) > 0\}$

(13) $\mathbb{R}_+ := \{x \in \mathbb{R} : x > 0\}$

(14) SLLN: *Strong law of large numbers*

(15) WLLN: *Weak law of large numbers*

(16) a.s.: *Almost sure*

(17) pdf: *Probability density function*

(18) M^T: *Transpose of the matrix* M

(19) M^*: *Complex conjugate transpose of the matrix* M

(20) $a_n = O(b_n)$: *If* $\left|\frac{a_n}{b_n}\right| \le C$, $\forall\ n$, *for some positive constant* C

(21) $a_n << b_n$ or $a_n = o(b_n)$: *If* $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$ *for two sequences* $\{a_n\}_n, \{b_n\}_n$ *of real numbers*

CHAPTER 1

Introduction

Random Matrix Theory was developed from several different sources in the early 20th century. It is used as an important mathematical tool in various fields namely, Mathematics, Physics, wireless communication engineering etc. One of the earliest example of a random matrix was appeared in the study of sample covariance estimation which was done by John Wishart [**Wis28**]. In the early 1950s, Wigner introduced random matrix ensemble to study the energy spectra of heavy atoms undergoing slow nuclear reactions.

In 1970s, a connection between Random Matrix Theory and number theory was found. The Riemann hypothesis says that the nontrivial zeros of the Riemann zeta function lie on the line $\frac{1}{2} + iE$ with $-\infty < E < \infty$. Assuming the Riemann hypothesis, Montgomery [**Mon73**] calculated the asymptotic two point correlation of these zeros, which turns out to be same as the two-lavel correlation function of the unitary random matrix ensemble.

Random matrices are also used to model wireless channels. A random matrix model of CDMA networks can be found in [**TV04, VS99**].

Definition 1.0.1 (Empirical Spectral Distribution). Let M be an $n \times n$ matrix, and $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of M. Then

$$\mu_M = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i} \tag{1.1}$$

is called the Empirical Spectral Distribution (ESD) of the matrix M.

Limiting behaviour of μ_n for different types of random matrices is an important class of problems in Random Matrix Theory. To discuss about limiting behaviours of μ_ns, we need to define a notion of convergence in the space of measures.

Definition 1.0.2 (Weak convergence). Let $\{\mu_n\}_n$ be a sequence of measures on a complete separable topological space $\mathcal{S}$. μ_n is said to converge weakly to a measure μ if $\int f \, d\mu_n \to \int f \, d\mu$ as $n \to \infty$ for all bounded continuous functions $f : \mathcal{S} \to \mathbb{R}$.

The limiting ESD of properly scaled random Hermitian matrices follow the Semicircle law. The theorem is formulated below.

Theorem 1.0.3 (Semicircle law). *Let $X = (x_{ij})_{n\times n}$ be an $n \times n$ random Hermitian matrix such that x_{ij} are i.i.d. random variables and $\mathbb{E}[x_{ij}] = 0$, $\mathbb{E}[|x_{ij}|^2] = \sigma^2$ for all i, j. Then the ESD of $\frac{1}{\sqrt{n}}X$ converges a.s. to the Semicircle distribution in the weak topology of measures, where the pdf ρ_{sc} of the Semicircle distribution is given by*

$$\rho_{sc}(x) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}\,\mathbf{1}_{\{|x|\le 2\sigma\}}(x).$$

The first proof dates back to Wigner [**Wig55**, **Wig58**], where he assumed all finite moments and the convergence was in probability. There was many improvements thereafter [**Arn71**,**Bai99**].

Let X be an $n \times n$ random matrix with i.i.d. entries, then XX^* is a non-negative definite symmetric matrix. Therefore all the eigenvalues of XX^* are non-negative real numbers. So, it is obvious that the limiting ESD of XX^* can not be the Semicircle distribution. The limiting ESD of such matrices was found by Marčenko and Pastur [**MP67**].

Theorem 1.0.4 (Marčenko Pastur law). *Let $X = (x_{ij})_{m\times n}$ be an $m \times n$ random matrix with i.i.d. real valued random variables such that $\mathbb{E}[x_{ij}] = 0$, $\mathbb{E}[x_{ij}^2] = \sigma^2$ for all i, j. Further assume that $\frac{m}{n} \to \gamma \in (0, \infty)$. Then the ESD of $\frac{1}{m}XX^T$ converges to the probability distribution with pdf f_γ, wea kly, in probability, where f_γ is given by*

$$f_\gamma(x) = \left(1 - \frac{1}{\gamma}\right)\mathbf{1}_{\{x=0,\gamma>1\}}(x) + \frac{\sqrt{(x-\gamma_-)(\gamma_+ - x)}}{2\pi\sigma^2\gamma x}\mathbf{1}_{[\gamma_-,\gamma_+]}(x),$$

where $\gamma_\pm = \sigma^2(1 \pm \sqrt{\gamma})^2$.

The eigenvalues of random matrices X with iid entries are distributed over the complex plane. The limiting spectral distribution of such matrices are uniformly distributed over a disk on the complex plane centred at origin [**TVK10**].

THEOREM 1.0.5. *Let* $X = (x_{ij})_{n\times n}$ *be an* $n \times n$ *random matrix with i.i.d complex random variables such that* $\mathbb{E}[x_{ij}] = 0$ *and* $\mathbb{E}[|x_{ij}|^2] = 1$ *for all* i, j. *Then the ESD of* $\frac{1}{\sqrt{n}}X$ *converges weakly and a.s. to the uniform probability measure on the unit disk on the complex plane.*

Classically, the Semicircle law was proved using the moment method. An alternative approach to find the limiting ESD is the Stieltjes transform method. The Stieltjes transform of a probability measure μ supported on the real line is defined by

$$m(z) = \int \frac{d\mu(\lambda)}{\lambda - z}, \quad z \in \mathbb{C}_+.$$

The Stieltjes transform characterizes a probability measure uniquely. If $\{m_n\}_n$ is a sequence of Stieltjes transforms of a sequence of probability measures $\{\mu_n\}_n$ such that $m_n(z) \to m(z)$ for all $z \in \mathbb{C}_+$, then $\mu_n \to \mu$ in the week topology. This approach was used to prove both the Semicircle law and the Marčenko-Pastur law. However, this method is not useful for proving the Circular law [**BS10**]. It can be shown that the Stieltjes transform of the Semicircle law is given by

$$m(z) = \frac{-z + \sqrt{z^2 - 4}}{2}, \quad z \in \mathbb{C}_+, \tag{1.2}$$

and that of the Marčenko-Pastur law with parameter γ satisfies

$$m(z)(1 + \sigma^2 m(z))z + (1 - \gamma)\sigma^2 m(z) + 1 = 0, \tag{1.3}$$

where σ^2 is the variance of the entries of the random matrices [**MP67**].

1.1. Random Band Matrices

A special kind of random matrix ensemble is a random band matrix. In 1955, Wigner studied the matrices H of the form $H = K + V$, where K is an $n \times n$ diagonal matrix consisting of $\cdots -2, -1, 0, 1, 2, \cdots$, and V is an $n \times n$ symmetric sign matrix having non vanishing elements only up to a distance b_n from the main diagonal. Such a matrix H was called as bordered matrix [**Wig55**, **Wig57**]. Random band matrices can be used to model an interacting particle system where a particle interacts with it's neighbours only upto a certain distance.

A treatment of random band matrix was done by G. Casati et al. [**CMI90**, **CIM91**] in the context of Quantum Chaos. They studied $n \times n$ symmetric random band matrices of bandwidth

b_n, where b_n grows with n. In 1992, Molchanov et al. proved the Semicircle law for random band matrices [**MPK92**]. In 1991, Fyodorov and Mirlin had shown that $\frac{b_n^2}{n}$ is a crucial parameter for the spectral properties of random band matrices [**FM91**, **MFD**$^+$**96**].

The Semicircle law is also true for random band matrix and was proved in early 90's [**BMP91**, **MPK92**]. Convergence of ESD random matrices is analogous to the WLLN and SLLN in the classical probability. After SLLN or WLLN, the next forward step is the study of CLTs. Various CLTs were proven for independent, weakly dependent sequence of random variables. But in Random Matrix Theory, the eigenvalues of a random matrix are not independent. In fact, they are highly correlated. Which makes study of CLT for linear eigenvalue statistics more interesting. In Chapter 2, we discuss about the CLT of linear eigenvalue statistics for Hermitian random band matrices. Then in the Chapter 3 we prove the the Marčenko-Pastur law for random band matrices and in the Chapter 4 we provide the numerical evidence that the Circular law is true for random band matrices.

In general, in the study of random matrices, the underlying distributions of the entries of the random matrix are unknown. Many random matrix results are universal, and they do not depend on the underlying distributions. However, there are some regimes where universality of the results may fail, and this research makes an attempt to identify such regimes.

CHAPTER 2

Central Limit Theorem for Linear Eigenvalue Statistics

In this Chapter, we deal with the CLT for the eigenvalue statistics of band random matrices. We take the approach of M. Shcherbina in [**Shc11**] to establish the CLT for band matrices with bandwidth b_n where $b_n \to \infty$ as $n \to \infty$. We give an alternative proof of Li and Soshnikov [**LS13**] result on CLT of band matrices when $\sqrt{n} << b_n << n$. We have given some simulation results in Section 2.3, which ensure that the CLT for band matrices will also hold if $\sqrt{n}/b_n \nrightarrow 0$ and $b_n \to \infty$.

Now we define our model. Let us define the (circular) distance function $d_n : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ as

$$d_n(j,k) := \min\{|j-k|, n - |j-k|\},$$

and the index sets $I_n, I_n^+ \subset \mathbb{N} \times \mathbb{N}$, $I_1 \subset \mathbb{N}$ as

(2.1)

$$I_n := \{(j,k) : d_n(j,k) \le b_n\}, \quad I_n^+ = \{(j,k) : (j,k) \in I_n, j \le k\}, \quad I_1 = \{1 < j \le n : (1,j) \in I_n\}$$

where $\{b_n\}$ is a sequence of positive integers such that $b_n \to \infty$ as $n \to \infty$.

Define a real symmetric random band matrix $M = (m_{jk})_{n \times n}$ of bandwidth b_n as

$$m_{jk} = m_{kj} = \begin{cases} b_n^{-1/2} w_{jk} & \text{if } d_n(j,k) \le b_n \\ 0 & \text{otherwise,} \end{cases} \tag{2.2}$$

where $\{w_{ii}\}$ and $\{w_{jk}\}_{j \neq k, (j,k) \in I_n^+}$ are two sets of i.i.d. real random variables with

$$\mathbb{E}[w_{jk}] = 0, \quad \mathbb{E}[w_{jk}^2] = \begin{cases} 1 & \text{if } j \neq k \\ \sigma^2 & \text{if } j = k. \end{cases} \tag{2.3}$$

Here $\{w_{jk}\}$ may depend on n, but we suppress it when there is no confusion. Let $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ be the eigenvalues of the random band matrix M. Define the linear eigenvalue statistic of the eigenvalues of M as

$$\mathcal{N}_n(\phi) = \sum_{i=1}^{n} \phi(\lambda_i), \tag{2.4}$$

and the normalized eigenvalue statistic of the matrix M as

$$\mathcal{M}_n(\phi) = \sqrt{\frac{b_n}{n}} \mathcal{N}_n(\phi), \tag{2.5}$$

where ϕ is a test function.

2.1. Main Results

THEOREM 2.1.1. *Let M be a real symmetric random band matrix as defined in* (2.2), *and b_n be a sequence of integers satisfying $\sqrt{n} << b_n << n$. Assume the following:*

(i) w_{jk} satisfies the Poincaré inequality with constant $m > 0$ not depending on j, k, n i.e., for any continuously differentiable function f,

$$Var(f(w_{jk})) \leq \frac{1}{m} \mathbb{E}\left[\left| f'(w_{jk}) \right|^2 \right].$$

(ii) $\mathbb{E}[w_{jk}^4] = \mu_4$ for all $j \neq k$ and $d_n(j,k) \leq b_n$.

(iii) $\phi : \mathbb{R} \to \mathbb{R}$ is a test function in the Sobolev space H^s i.e., $\|\phi\|_s < \infty$, where

$$\|\phi\|_s^2 = \int_{\mathbb{R}} (1 + 2|t|)^{2s} |\hat{\phi}(t)|^2 \, dt,$$
$$\hat{\phi}(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-it\lambda} \phi(\lambda) \, d\lambda,$$

and $s > 5/2$.

Then the centred normalized eigenvalue statistic $\mathcal{M}^\circ(\phi) = \mathcal{M}_n(\phi) - \mathbb{E}[\mathcal{M}_n(\phi)]$ converges in distribution to the Gaussian random variable with mean zero and variance given by

$$V(\phi) = \frac{\kappa_4}{16\pi^2} \left(\int_{-2\sqrt{2}}^{2\sqrt{2}} \frac{4 - \mu^2}{\sqrt{8 - \mu^2}} \phi(\mu) \, d\mu \right)^2 + \frac{\sigma^2}{16\pi^2} \left(\int_{-2\sqrt{2}}^{2\sqrt{2}} \frac{\mu \phi(\mu)}{\sqrt{8 - \mu^2}} \, d\mu \right)^2$$
$$+ \int_{-2\sqrt{2}}^{2\sqrt{2}} \int_{-2\sqrt{2}}^{2\sqrt{2}} \sqrt{(8 - x^2)(8 - y^2)} F(x, y)$$

$$\times \int_{-2\sqrt{2}}^{2\sqrt{2}} \int_{-2\sqrt{2}}^{2\sqrt{2}} \frac{\mu_1 \phi(\mu_1)}{(x-\mu_1)\sqrt{8-\mu_1^2}} \frac{\mu_2 \phi(\mu_2)}{(x-\mu_2)^2\sqrt{8-\mu_2^2}} d\mu_1 d\mu_2 \; dxdy,$$

where for $x \neq y$

$$F(x,y) = 2\int_{-\infty}^{\infty} \frac{(s^3 \sin s - s \sin^3 s)}{2(s^2-\sin^2 s)^2 - (s^3 \sin s + s\sin^3 s)xy + s^2 \sin^2 s(x^2+y^2)} ds,$$

and κ_4 is the fourth cumulant of the off-diagonal entries, i.e., $\kappa_4 = \mu_4 - 3$.

2.2. Proof of the Theorem 2.1.1

We follow the approach taken by M. Shcherbina in [**Shc11**] for full (Wigner) matrix. This approach is based on two main ideas. The first ingredient is stated in the following proposition which gives a bound on the variance of linear eigenvalue statistics with a sufficiently smooth test function in term of the variance of the trace of the resolvent of a random matrix. For a proof of this result see [**Shc11**, **ORS13**]. In what follows, we denote $X^\circ = X - \mathbb{E}[X]$ for any random variable X.

PROPOSITION 2.2.1. *Let M be an $n \times n$ real symmetric random matrix and $\mathcal{N}_n(\phi)$ be a linear eigenvalue statistic of its eigenvalue as in* (2.4)*. Then for any $s > 0$ we have*

$$Var[\mathcal{N}_n(\phi)] \leq C_s \|\phi\|_s^2 \int_0^\infty dy\; e^{-y} y^{2s-1} \int_{-\infty}^{\infty} Var[Tr(G(x+iy))]\; dx,$$

where C_s is a constant depends only on s, and $G(z) = (M - zI)^{-1}$, is the resolvent of the matrix M.

The second ingredient of this approach is to use the martingale difference technique to provide a good bound on $\mathrm{Var}(\gamma_n)$ where γ_n is the trace of the resolvent of a matrix. The following proposition gives that bound.

PROPOSITION 2.2.2. *Consider symmetric band matrix M defined in* (2.2) *and assume* (2.3) *is satisfied. Then for some $C > 0$ not depending on z, n we have*

$$Var\{\gamma_n\} \leq \frac{Cn}{b_n}\left(y^{-2} + y^{-4}\right)\left(\max\left\{y, |x| - \frac{2}{y}\right\}\right)^{-2} \tag{2.6}$$

where $\gamma_n = Tr(M - zI)^{-1} = Tr(G)$ and $z = x + iy$, $y > 0$.

We prove this result in Chapter A . Now we outline the proof of Theorem 2.1.1

Proof of Theorem 2.1.1: By Lévy's continuity theorem, it suffices to show that if

$$Z_n(x) = \mathbb{E}[e_n(x)], \quad e_n(x) = e^{ix\mathcal{M}_n^\circ(\phi)} \tag{2.7}$$

then for each $x \in \mathbb{R}$

$$\lim_{n\to\infty} Z_n(x) = \exp\left[-\frac{x^2 V(\phi)}{2}\right],$$

where $V(\phi)$ as in Theorem 2.1.1. For any test function $\phi \in H^s$, define

$$\phi_\eta = P_\eta * \phi,$$

where P_η is the Poisson kernel given by

$$P_\eta(x) = \frac{\eta}{\pi(x^2+\eta^2)}.$$

We know that ϕ_η approximates ϕ in the H^s norm i.e.,

$$\lim_{\eta\to 0} \|\phi - \phi_\eta\|_s = 0. \tag{2.8}$$

For the moment, we denote the characteristic function defined in (2.7), by $Z_n(\phi)$ (to make its dependence on ϕ clear). Then we have

$$\lim_{n\to\infty} Z_n(\phi) = \lim_{\eta\downarrow 0}\lim_{n\to\infty} (Z_n(\phi) - Z_n(\phi_\eta)) + \lim_{\eta\downarrow 0}\lim_{n\to\infty} Z_n(\phi_\eta).$$

Now using the Proposition 2.2.1 and (2.8), we shall show that

$$\lim_{\eta\downarrow 0}\lim_{n\to\infty} (Z_n(\phi) - Z_n(\phi_\eta)) = 0. \tag{2.9}$$

and then

$$\lim_{n\to\infty} Z_n(\phi) = \lim_{\eta\downarrow 0}\lim_{n\to\infty} Z_n(\phi_\eta).$$

Hence it suffices to find the limit of

$$Z_{\eta,n} := Z_n(\phi_\eta) = \mathbb{E}\left[e_{\eta,n}(x)\right] \tag{2.10}$$

with

$$e_{\eta,n}(x) = \exp\left[ix\mathcal{M}_n^\circ(\phi_\eta)\right]$$

as $n \to \infty$ and $\eta \downarrow 0$ uniformly in n. Proofs of (2.9) and (2.10) are given in the next two subsections and that will complete the proof of this theorem. □

2.2.1. Proof of equation (2.9)**.** First observe that

$$|Z_n(\phi) - Z_n(\phi_\eta)|^2 \leq 2|x|^2 \mathrm{Var}\left[\mathcal{M}_n(\phi) - \mathcal{M}_n(\phi_\eta)\right] \leq 2|x|^2 \frac{b_n}{n} \mathrm{Var}\left[\mathcal{N}_n(\phi) - \mathcal{N}_n(\phi_\eta)\right]. \tag{2.11}$$

Now, in view of Proposition 2.2.1, to bound $\mathrm{Var}\left[\mathcal{N}_n(\phi) - \mathcal{N}_n(\phi_\eta)\right]$ we need to estimate

$$\int_{-\infty}^{\infty} \mathrm{Var}\left(\gamma_n(x+iy)\right)\, dx,$$

where $\gamma_n(x+iy) = \mathrm{Tr}(G(x+iy))$ and $G(z) = (M - zI)^{-1}$. We estimate that for $y > 0$

$$\begin{aligned}\int_{-\infty}^{\infty} \left(\max\left\{y, |x| - \frac{2}{y}\right\}\right)^{-2} dx &\leq \int_{||x|-2/y|<y} \frac{1}{y^2}\, dx + \int_{||x|-2/y|\geq y} (x - 2/y)^{-2}\, dx \\ &\leq \frac{10}{y} + 10y\end{aligned}$$

Using the above estimate and (2.6), we have

$$\begin{aligned}\int_0^{\infty} dy\, e^{-y} y^{2s-1} \int_{-\infty}^{\infty} \mathrm{Var}(\gamma_n)\, dx &\leq \frac{C'}{b_n} \int_0^{\infty} e^{-y} y^{2s-1} 4n \left(\frac{1}{y} + y\right)\left(\frac{1}{y^2} + \frac{1}{y^4}\right) dy \\ &= C\frac{n}{b_n} \int_0^{\infty} e^{-y} \left(2y^{2s-3-1} + y^{2s-1-1} + y^{2s-5-1}\right) dy \\ &= C\frac{n}{b_n} \left(\Gamma(2s-3) + \Gamma(2s-1) + \Gamma(2s-5)\right). \end{aligned} \tag{2.12}$$

If we take

$$s = \frac{5}{2} + \epsilon, \quad \epsilon > 0$$

then $\Gamma(2s-3) = \Gamma(2+2\epsilon)$, $\Gamma(2s-1) = \Gamma(4+2\epsilon)$, and $\Gamma(2s-5) = \Gamma(2\epsilon)$. By Proposition 2.2.1, and (2.12), we have

$$\mathrm{Var}\left(\mathcal{N}_n(\phi) - \mathcal{N}_n(\phi_\eta)\right) \le C(\epsilon)\frac{n}{b_n}\|\phi - \phi_\eta\|_s.$$

Using the above estimate and (2.11), we have

$$\begin{aligned} |Z_n(\phi) - Z_n(\phi_\eta)|^2 &\le 2|x|^2\frac{b_n}{n}\cdot C(\epsilon)\frac{n}{b_n}\|\phi - \phi_\eta\|_s \\ &= 2C(\epsilon)|x|^2\|\phi - \phi_\eta\|_s \\ &\to 0 \quad \text{as } \eta \to 0. \end{aligned}$$

The last limit follows from the equation (2.8). This completes the proof of (2.9).

2.2.2. Finding the limit of the characteristic function (2.10)**.** We will be using the Lemma A.0.1 and Lemma A.0.2 from Appendix in the proof of (2.10). Let us denote the averaging with respect to $\{w_{1i}; 1 \le i \le n\}$ by $\mathbb{E}_1$.

Proof of (2.10): Using the dominated convergence theorem we have

$$\begin{aligned} \frac{d}{dx}Z_n(\phi_\eta) &= \frac{d}{dx}\mathbb{E}\left[e_{\eta,n}(x)\right] \\ &= \frac{d}{dx}\mathbb{E}\left[\exp\left(ix\sqrt{\frac{b_n}{n}}\mathcal{N}_n^\circ(\phi_\eta)\right)\right] \\ &= \mathbb{E}\left[i\sqrt{\frac{b_n}{n}}\mathcal{N}_n^\circ(\phi_\eta)e_{\eta,n}(x)\right]. \end{aligned}$$

Since by construction $\phi_\eta = P_\eta * \phi$, we have

$$\mathcal{N}_n^\circ(\phi_\eta) = \frac{1}{\pi}\int_{-\infty}^{\infty}\phi(\mu)\Im\gamma_n^\circ(z_\mu)\,d\mu, \quad \text{where } z_\mu = \mu + i\eta.$$

Hereinafter, we use the finiteness of $\int_{\mathbb{R}}|\phi(\mu)|\,d\mu$ for $\phi \in H^s, s > \frac{1}{2}$, when changing the order of integration. For notational convenience, from now on we will denote $e_{\eta,n}(x)$ by $e(x)$. Therefore

$$\frac{d}{dx}Z_n(\phi_\eta) = \mathbb{E}\left[i\sqrt{\frac{b_n}{n}}e(x)\frac{1}{\pi}\int_{-\infty}^{\infty}\phi(\mu)\Im\gamma_n^\circ(z_\mu)\,d\mu\right]$$

$$= \frac{1}{2\pi}\sqrt{\frac{b_n}{n}}\int_{-\infty}^{\infty}\phi(\mu)\mathbb{E}\left[e(x)\mathrm{Tr}\left(G^{\circ}(z_\mu)-G^{\circ}(\bar{z}_\mu)\right)\right]\,d\mu$$

$$= \frac{1}{2\pi}\sqrt{\frac{b_n}{n}}\int_{-\infty}^{\infty}\phi(\mu)\left(Y_n(z_\mu,x)-Y_n(\bar{z}_\mu,x)\right)\,d\mu,$$

where

$$\begin{aligned} Y_n(z,x) &= \mathbb{E}\left[e(x)\mathrm{Tr}\left(G^{\circ}(z)\right)\right] \\ &= \mathbb{E}\left[e^{\circ}(x)\mathrm{Tr}(G(z))\right] \\ &= n\mathbb{E}\left[G_{11}(z)e^{\circ}(x)\right] \\ &= -n\mathbb{E}\left[\left(A^{-1}\right)^{\circ}e_1(x)\right]-n\mathbb{E}\left[\left(A^{-1}\right)^{\circ}\left(e(x)-e_1(x)\right)\right], \end{aligned} \tag{2.13}$$

$$e_1(x) = \exp\left[ix\sqrt{\frac{b_n}{n}}\left(\mathcal{N}_{n-1}(\phi_\eta)\right)^{\circ}\right],$$

$$\left(\mathcal{N}_{n-1}(\phi_\eta)\right)^{\circ} = \frac{1}{\pi}\int_{-\infty}^{\infty}\phi(\mu)\Im\left(\gamma_{n-1}(z)\right)^{\circ}\,d\mu,$$

$$\gamma_{n-1}(z) = \mathrm{Tr}G^{(1)}(z),$$

$$A(z) = z-\frac{1}{\sqrt{b_n}}w_{11}+\left\langle G^{(1)}m^{(1)},m^{(1)}\right\rangle, \tag{2.14}$$

$$m^{(1)} = \frac{1}{\sqrt{b_n}}(w_{12},w_{13},\ldots,w_{1n})^T, \tag{2.15}$$

$$G^{(1)}(z) = \left(G^{(1)}_{ij}(z)\right)_{i,j=2}^{n} = (M^{(1)}-zI)^{-1}, \tag{2.16}$$

and $M^{(1)}$ is the main bottom $(n-1)\times(n-1)$ minor of M. In the above notation $\langle\cdot,\cdot\rangle$ represents the inner product of two complex vectors, i.e., $\langle x,y\rangle = \bar{y}^Tx$ for $x,y\in\mathbb{C}^{n-1}$. Equation (2.13) follows from the Schur complement lemma, which says that

$$G_{11}(z) = \frac{1}{\frac{1}{\sqrt{b_n}}w_{11}-z-\left\langle G^{(1)}m^{(1)},m^{(1)}\right\rangle} = -\frac{1}{A(z)}. \tag{2.17}$$

Now we rewrite

$$\sqrt{\frac{b_n}{n}}Y_n(z,x) = -\sqrt{nb_n}\mathbb{E}\left[\left(A^{-1}\right)^{\circ}e_1(x)\right]-\sqrt{nb_n}\mathbb{E}\left[\left(A^{-1}\right)^{\circ}\left(e(x)-e_1(x)\right)\right]$$

$$=: T_1 + T_2. \tag{2.18}$$

Using Taylor expansion we have

$$A^{-1} = \frac{1}{\mathbb{E}[A]} - \frac{A^\circ}{(\mathbb{E}[A])^2} + \frac{(A^\circ)^2}{(\mathbb{E}[A])^3} - \frac{(A^\circ)^3}{(\mathbb{E}[A])^4} + \frac{(A^\circ)^4}{A(\mathbb{E}[A])^4}. \tag{2.19}$$

Therefore, we can estimate

$$\begin{aligned} T_1 &= -\sqrt{nb_n}\mathbb{E}\left[\left(A^{-1}\right)^\circ e_1(x)\right] \\ &= -\sqrt{nb_n}\mathbb{E}\left[\left(A^{-1}\right) e_1^\circ(x)\right] \\ &= -\sqrt{nb_n}\mathbb{E}\left[\left(\frac{1}{\mathbb{E}[A]} - \frac{A^\circ}{(\mathbb{E}[A])^2} + \frac{(A^\circ)^2}{(\mathbb{E}[A])^3} - \frac{(A^\circ)^3}{(\mathbb{E}[A])^4} + \frac{(A^\circ)^4}{A(\mathbb{E}[A])^4}\right) e_1^\circ(x)\right] \\ &= \sqrt{nb_n}\mathbb{E}\left[\left(\frac{A^\circ}{(\mathbb{E}[A])^2} - \frac{(A^\circ)^2}{(\mathbb{E}[A])^3}\right) e_1^\circ(x)\right] + \sqrt{nb_n}\mathbb{E}\left[\left(\frac{(A^\circ)^3}{(\mathbb{E}[A])^4} - \frac{(A^\circ)^4}{A(\mathbb{E}[A])^4}\right) e_1^\circ(x)\right]. \end{aligned} \tag{2.20}$$

Now we shall estimate each term individually. First of all, since M is a real symmetric matrix we have

$$\|G(z)\| \le \frac{1}{|\Im z|}, \tag{2.21}$$

and, in particular, $1/|A| \le 1/|\Im z|$. It can also be checked that $1/|\mathbb{E}[A]| \le 1/|\Im z|$. Hereinafter $\|X\|$ is the spectral norm of a matrix X. Using the above equation (2.21) and the estimates (A.8), (A.10), we have

$$\left|\sqrt{nb_n}\mathbb{E}\left[\frac{(A^\circ)^4}{A(\mathbb{E}[A])^4}e_1^\circ(x)\right]\right| \le \frac{\sqrt{nb_n}}{|\Im z|^5}\mathbb{E}\left[|(A^\circ)^4|\right] = \frac{\sqrt{nb_n}}{|\Im z|^5}O(b_n^{-2}) = O\left(\sqrt{\frac{n}{b_n^3}}\right) = o(1),$$

$$\left|\sqrt{nb_n}\mathbb{E}\left[\frac{(A^\circ)^3}{(\mathbb{E}[A])^4}e_1^\circ(x)\right]\right| \le \frac{\sqrt{nb_n}}{|\Im z|^4}\mathbb{E}\left[|(A^\circ)^3|\right] = \frac{\sqrt{nb_n}}{|\Im z|^4}O(b_n^{-3/2}) = O\left(\sqrt{\frac{n}{b_n^2}}\right) = o(1),$$

$$\begin{aligned} \left|\sqrt{nb_n}\mathbb{E}\left[\frac{(A^\circ)^2}{(\mathbb{E}[A])^3}e_1^\circ(x)\right]\right| &\le \frac{\sqrt{nb_n}}{|\Im z|^3}\left|\mathbb{E}\left[e_1^\circ(x)\mathbb{E}_1\left[(A^\circ)^2\right]\right]\right| \\ &\le C\sqrt{\frac{n}{b_n}}\left|\mathbb{E}\left[e_1^\circ(x)\left(b_n\mathbb{E}_1(A^\circ)^2\right)\right]\right| \end{aligned}$$

$$\begin{aligned}
&\leq C\sqrt{\frac{n}{b_n}}\left[\mathrm{Var}(e_1^\circ(x))\right]^{1/2}\left[\mathrm{Var}\left(b_n\mathbb{E}_1(A^\circ)^2\right)\right]^{1/2}\\
&\leq C\sqrt{\frac{n}{b_n}}O(b_n^{-1/2})\\
&= O\left(\sqrt{\frac{n}{b_n^2}}\right) = o(1),\ \text{as } n\to\infty,
\end{aligned}$$

Therefore, we have

$$T_1 = \frac{\sqrt{nb_n}}{(\mathbb{E}[A])^2}\mathbb{E}\left[A^\circ e_1^\circ(x)\right] + O\left(\sqrt{\frac{n}{b_n^2}}\right) = \frac{\sqrt{nb_n}}{(\mathbb{E}[A])^2}\mathbb{E}\left[e_1^\circ(x)\mathbb{E}_1(A^\circ)\right] + O\left(\sqrt{\frac{n}{b_n^2}}\right).$$

Now

$$A^\circ = -\frac{1}{\sqrt{b_n}}w_{11} + \frac{1}{b_n}\sum_{\substack{i\neq j\\ i,j\in I_1}} G_{ij}^{(1)}w_{1i}w_{1j} + \frac{1}{b_n}\sum_{i\in I_1}\left(G_{ii}^{(1)}w_{1i}^2 - \mathbb{E}[G_{ii}^{(1)}]\right),$$

where $I_1 = \{1 < j \leq n : (1,j)\in I_n\}$. Therefore,

$$\mathbb{E}_1[A^\circ(z)] = \frac{1}{b_n}\sum_{i\in I_1}\left(G_{ii}^{(1)} - \mathbb{E}[G_{ii}^{(1)}]\right)$$

and hence

$$\begin{aligned}
T_1 &= \frac{\sqrt{nb_n}}{(\mathbb{E}[A])^2}\mathbb{E}\left[e_1^\circ(x)\mathbb{E}_1(A^\circ)\right] + O\left(\sqrt{\frac{n}{b_n^2}}\right)\\
&= \frac{\sqrt{nb_n}}{(\mathbb{E}[A])^2}\mathbb{E}\left[e_1^\circ(x)\frac{1}{b_n}\sum_{i\in I_1}(G_{ii}^{(1)} - \mathbb{E}[G_{ii}^{(1)}])\right] + O\left(\sqrt{\frac{n}{b_n^2}}\right)\\
&= \frac{\sqrt{nb_n}}{(\mathbb{E}[A])^2}\,2\mathbb{E}\left[(G_{22}^{(1)})^\circ e_1^\circ(x)\right] + O\left(\sqrt{\frac{n}{b_n^2}}\right)\\
&= \frac{\sqrt{nb_n}}{(\mathbb{E}[A])^2}\,\frac{2}{n}\,\mathbb{E}[\gamma_{n-1}^\circ e_1^\circ(x)] + O\left(\sqrt{\frac{n}{b_n^2}}\right)\\
&= \sqrt{\frac{b_n}{n}}\frac{2}{(\mathbb{E}[A])^2}\mathbb{E}\left[\gamma_{n-1}^\circ e_1(x)\right] + O\left(\sqrt{\frac{n}{b_n^2}}\right). \qquad (2.22)
\end{aligned}$$

Hereinafter, all bounds (implicitly) depending on z hold uniformly on the set $\{\mu + i\eta : \mu\in\mathbb{R}\}$ for any given $\eta > 0$. Now

$$\left|\mathbb{E}\left[\gamma_{n-1}^\circ e_1(x)\right] - \mathbb{E}\left[\gamma_n^\circ e(x)\right]\right| = \left|\mathbb{E}\left[\gamma_{n-1}^\circ e_1(x)\right] - \mathbb{E}\left[\gamma_n^\circ e_1(x)\right] + \mathbb{E}\left[\gamma_n^\circ e_1(x)\right] - \mathbb{E}\left[\gamma_n^\circ e(x)\right]\right|$$

$$\leq \left(\mathbb{E}\left[\left|\gamma_{n-1}^\circ - \gamma_n^\circ\right|^4\right]\right)^{1/4} + \left|\mathbb{E}\left[\gamma_n^\circ(e_1(x) - e(x))\right]\right|$$
$$= O(b_n^{-1/2}) + \left|\mathbb{E}\left[\gamma_n^\circ(e_1(x) - e(x))\right]\right|.$$

The last equality follows from (A.11). We estimate

$$\begin{aligned}
e(x) - e_1(x) &= \exp\left[ix\sqrt{\frac{b_n}{n}}\mathcal{N}_n^\circ(\phi_\eta)\right] - \exp\left[ix\sqrt{\frac{b_n}{n}}\mathcal{N}_{n-1}^\circ(\phi_\eta)\right] \\
&= \left(\exp\left[ix\sqrt{\frac{b_n}{n}}\mathcal{N}_n^\circ(\phi_\eta) - ix\sqrt{\frac{b_n}{n}}\mathcal{N}_{n-1}^\circ(\phi_\eta)\right] - 1\right)e_1(x) \\
&= ix\sqrt{\frac{b_n}{n}}\left(\mathcal{N}_n^\circ(\phi_\eta) - \mathcal{N}_{n-1}^\circ(\phi_\eta)\right)e_1(x) + \frac{b_n}{n}O\left(x^2\left(\mathcal{N}_n^\circ(\phi_\eta) - \mathcal{N}_{n-1}^\circ(\phi_\eta)\right)^2 e_1(x)\right) \\
&= \frac{ix}{\pi}\sqrt{\frac{b_n}{n}}\int_{-\infty}^{\infty}\left[\phi(\mu)\Im\left(\gamma_n^\circ - \gamma_{n-1}^\circ\right)e_1(x) + \sqrt{\frac{b_n}{n}}\phi(\mu)O(\gamma_n^\circ - \gamma_{n-1}^\circ)^2\right]d\mu. \qquad (2.23)
\end{aligned}$$

Therefore

$$\begin{aligned}
&\mathbb{E}\left[\gamma_n^\circ(e(x) - e_1(x))\right] \\
&= \mathbb{E}\left[\frac{ix}{\pi}\sqrt{\frac{b_n}{n}}\int_{-\infty}^{\infty}\phi(\mu)\left[\Im\left(\gamma_n^\circ - \gamma_{n-1}^\circ\right)e_1(x)\gamma_n^\circ + \sqrt{\frac{b_n}{n}}\gamma_n^\circ O\left(\gamma_n^\circ - \gamma_{n-1}^\circ\right)^2\right]d\mu\right].
\end{aligned}$$

Using estimates (2.6) and (A.11), we have

$$\left|\mathbb{E}\left[\Im\left(\gamma_n^\circ - \gamma_{n-1}^\circ\right)e_1(x)\gamma_n^\circ\right]\right| \leq \left(\mathbb{E}[|\gamma_n^\circ|^2]\right)^{1/2}\left(\mathbb{E}\left[\left|e_1(x)\Im\left(\gamma_n^\circ - \gamma_{n-1}^\circ\right)\right|^2\right]\right)^{1/2} = O\left(\sqrt{\frac{n}{b_n}}\sqrt{\frac{1}{b_n}}\right).$$

Similarly,

$$\mathbb{E}\left[\gamma_n^\circ O\left(\gamma_n^\circ - \gamma_{n-1}^\circ\right)^2\right] = O\left(\sqrt{\frac{n}{b_n}}\frac{1}{b_n}\right).$$

Therefore,

$$\left|\mathbb{E}\left[\gamma_{n-1}^\circ e_1(x)\right] - \mathbb{E}\left[\gamma_n^\circ e(x)\right]\right| = O\left(\frac{1}{\sqrt{b_n}}\right).$$

From the equation (2.22) and the above estimates we have

$$T_1 = \sqrt{\frac{b_n}{n}}\frac{2}{(\mathbb{E}[A])^2}\mathbb{E}\left[\gamma_{n-1}^\circ e_1(x)\right] + O\left(\frac{\sqrt{n}}{b_n}\right)$$

$$= \sqrt{\frac{b_n}{n}} \frac{2}{(\mathbb{E}[A])^2} \left[\mathbb{E}[\gamma_n^\circ e(x)] + O\left(\frac{1}{\sqrt{b_n}}\right)\right] + O\left(\frac{\sqrt{n}}{b_n}\right)$$

$$= \sqrt{\frac{b_n}{n}} \frac{2}{(\mathbb{E}[A])^2} Y_n(z,x) + O\left(\frac{\sqrt{n}}{b_n}\right). \tag{2.24}$$

Now consider T_2. Using (2.23) and (A.1) we have

$$\begin{aligned}
T_2 &= -\sqrt{nb_n}\mathbb{E}\left[(A^{-1})^\circ(e(x) - e_1(x))\right] \\
&= -\frac{ixb_n}{\pi}\mathbb{E}\left[(A^{-1})^\circ \int_{-\infty}^{\infty} \phi(\mu)\Im\left(\gamma_n^\circ - \gamma_{n-1}^\circ\right) e_1(x)\, d\mu\right] \\
&- \frac{1}{\pi}\sqrt{\frac{b_n^3}{n}}\mathbb{E}\left[(A^{-1})^\circ \int_{-\infty}^{\infty} \phi(\mu) O(\gamma_n^\circ - \gamma_{n-1}^\circ)^2\, d\mu\right] \\
&= -\frac{ixb_n}{\pi}\int_{-\infty}^{\infty} \phi(\mu)\mathbb{E}\left[e_1(x)(A^{-1})^\circ\Im(\gamma_n^\circ - \gamma_{n-1}^\circ)\right]\, d\mu + \sqrt{\frac{b_n^3}{n}} O\left(\frac{1}{b_n}\right) \\
&= -\frac{ixb_n}{\pi}\int_{-\infty}^{\infty} \phi(\mu)\mathbb{E}\left[e_1(x)(A^{-1})^\circ\Im\left(\gamma_n^\circ - \gamma_{n-1}^\circ\right)\right]\, d\mu + O\left(\sqrt{\frac{b_n}{n}}\right) \\
&= -\frac{ixb_n}{\pi}\int_{-\infty}^{\infty} \phi(\mu)\mathbb{E}\left[e_1(x)(A^{-1})^\circ\Im\left(\gamma_n - \gamma_{n-1}\right)^\circ\right]\, d\mu + O\left(\sqrt{\frac{b_n}{n}}\right) \\
&= \frac{ixb_n}{\pi}\int_{-\infty}^{\infty} \phi(\mu)\mathbb{E}\left[e_1(x)(A^{-1})^\circ(z)\Im\left(\frac{1 + B(z_\mu)}{A(z_\mu)}\right)^\circ\right]\, d\mu + O\left(\sqrt{\frac{b_n}{n}}\right) \\
&= T_{21} - T_{22} + O\left(\sqrt{\frac{b_n}{n}}\right),
\end{aligned}$$

where $B(z) = \left\langle G^{(1)}G^{(1)}m^{(1)}, m^{(1)}\right\rangle$ and

$$T_{21} = \frac{xb_n}{2\pi}\int_{-\infty}^{\infty} \phi(\mu)\mathbb{E}\left[e_1(x)(A^{-1})^\circ(z)\left(\frac{1 + B(z_\mu)}{A(z_\mu)}\right)^\circ\right] d\mu,$$

$$T_{22} = \frac{xb_n}{2\pi}\int_{-\infty}^{\infty} \phi(\mu)\mathbb{E}\left[e_1(x)(A^{-1})^\circ(z)\left(\frac{\overline{1 + B(z_\mu)}}{\overline{A(z_\mu)}}\right)^\circ\right] d\mu.$$

Using $\Im\left\langle G^{(1)}m^{(1)}, m^{(1)}\right\rangle = \Im z\left\langle G^{(1)*}G^{(1)}m^{(1)}, m^{(1)}\right\rangle$, it can be easily verified that

$$\left|\frac{B(z)}{A(z)}\right| \le \frac{1}{|\Im z|},\ \frac{1}{|\mathbb{E}[A(z)]|} \le \frac{1}{|\Im z|},\ \text{and } |\mathbb{E}[B(z)]| \le \frac{2}{|\Im z|^2}. \tag{2.25}$$

Applying $A^{-1} = \frac{1}{\mathbb{E}[A]} - \frac{A^\circ}{(\mathbb{E}[A])^2} + \frac{(A^\circ)^2}{A(\mathbb{E}[A])^2}$ to $A^{-1}(z),\ A^{-1}(z_\mu)$ and using (A.8), we get

$$
\begin{aligned}
&b_n\mathbb{E}\left[e_1(x)(A^{-1})^\circ(z)\left(\frac{1+B(z_\mu)}{A(z_\mu)}\right)^\circ\right]\\
&= b_n\mathbb{E}\left[e_1(x)\left\{\frac{A^\circ(z)}{\mathbb{E}^2[A(z)]}\left(-\frac{B^\circ(z_\mu)}{\mathbb{E}[A(z_\mu)]}+\frac{(1+B(z_\mu))A^\circ(z_\mu)}{\mathbb{E}^2[A(z_\mu)]}-\frac{\mathbb{E}[B(z_\mu)A^\circ(z_\mu)]}{\mathbb{E}^2[A(z_\mu)]}\right)\right\}\right]+O(b_n^{-1/2})\\
&= b_n\mathbb{E}\left[e_1(x)\left\{\frac{A^\circ(z)}{\mathbb{E}^2[A(z)]}\left(-\frac{B^\circ(z_\mu)}{\mathbb{E}[A(z_\mu)]}+\frac{(1+\mathbb{E}[B(z_\mu)])A^\circ(z_\mu)}{\mathbb{E}^2[A(z_\mu)]}+\frac{(B^\circ(z_\mu)A^\circ(z_\mu))^\circ}{\mathbb{E}^2[A(z_\mu)]}\right)\right\}\right]+O(b_n^{-1/2})\\
&= \frac{(1+\mathbb{E}[B(z_\mu)])}{\mathbb{E}^2[A(z)]\mathbb{E}^2[A(z_\mu)]}\mathbb{E}\left[e_1(x)b_nA^\circ(z)A^\circ(z_\mu)\right]-\frac{\mathbb{E}\left[e_1(x)b_nA^\circ(z)B^\circ(z_\mu)\right]}{\mathbb{E}^2[A(z)]\mathbb{E}[A(z_\mu)]}+O(b_n^{-1/2})\\
&= \frac{(1+\mathbb{E}[B(z_\mu)])}{\mathbb{E}^2[A(z)]\mathbb{E}^2[A(z_\mu)]}\mathbb{E}\left[e_1(x)\mathbb{E}_1\left(b_nA^\circ(z)A^\circ(z_\mu)\right)\right]-\frac{\mathbb{E}\left[e_1(x)\mathbb{E}_1\left[b_nA^\circ(z)B^\circ(z_\mu)\right]\right]}{\mathbb{E}^2[A(z)]\mathbb{E}[A(z_\mu)]}+O(b_n^{-1/2}).
\end{aligned}
$$

Using (A.10), from the last expression we get

$$
\begin{aligned}
&b_n\mathbb{E}\left[e_1(x)(A^{-1})^\circ(z)\left(\frac{1+B(z_\mu)}{A(z_\mu)}\right)^\circ\right]\\
&= \frac{(1+\mathbb{E}B(z_\mu))}{\mathbb{E}^2[A(z)]\mathbb{E}^2[A(z_\mu)]}\mathbb{E}[e_1(x)]\mathbb{E}\left[b_nA^\circ(z)A^\circ(z_\mu)\right]-\frac{\mathbb{E}[e_1(x)]\mathbb{E}\left[b_nA^\circ(z)B^\circ(z_\mu)\right]}{\mathbb{E}^2[A(z)]\mathbb{E}[A(z_\mu)]}+O(b_n^{-1/2}).
\end{aligned}
\tag{2.26}
$$

Define

$$
D_n(z,z_\mu)=\frac{(1+\mathbb{E}[B(z_\mu)])\mathbb{E}[b_n\mathbb{E}_1\left\{A^\circ(z)A^\circ(z_\mu)\right\}]}{\mathbb{E}^2[A(z)]\mathbb{E}^2[A(z_\mu)]}-\frac{\mathbb{E}[b_n\mathbb{E}_1\{A^\circ(z)B^\circ(z_\mu)\}]}{\mathbb{E}^2[A(z)]\mathbb{E}[A(z_\mu)]}.
$$

Also, using (2.23) and (A.11), we have

$$
\begin{aligned}
&\mathbb{E}[e(x)]-\mathbb{E}[e_1(x)]\\
&=\mathbb{E}\left[\frac{ix}{\pi}\sqrt{\frac{b_n}{n}}\int_{-\infty}^{\infty}\phi(\mu)\Im\left(\gamma_n^\circ-\gamma_{n-1}^\circ\right)e_1(x)\,d\mu+\frac{b_n}{n}x^2\int_{-\infty}^{\infty}\phi(\mu)O(\gamma_n^\circ-\gamma_{n-1}^\circ)^2\,d\mu\right]\\
&=O(n^{-1/2})+O(n^{-1}).
\end{aligned}
$$

Therefore

$$
\mathbb{E}[e_1(x)]=Z_n(\phi_\eta)+O(n^{-1/2}).
\tag{2.27}
$$

Combining (2.18), (2.24), (2.26), and (2.27), we get

$$
\sqrt{\frac{b_n}{n}}Y_n(z,x)=T_1+T_2
$$

$$
\begin{aligned}
&= \frac{2}{\mathbb{E}^2[A]}\sqrt{\frac{b_n}{n}}Y_n(z,x) + \frac{x}{2\pi}\mathbb{E}[e_1(x)]\int_{-\infty}^{\infty}[D_n(z,z_\mu) - D_n(z,\bar{z}_\mu)]\phi(\mu)\,d\mu + O(b_n^{-1/2}) \\
&= \frac{2}{\mathbb{E}^2[A]}\sqrt{\frac{b_n}{n}}Y_n(z,x) + \frac{x}{2\pi}Z_n(\phi_\eta)\int_{-\infty}^{\infty}[D_n(z,z_\mu) - D_n(z,\bar{z}_\mu)]\phi(\mu)\,d\mu + o(1) \\
&\approx 2f^2(z)\tilde{Y}_n(z,x) + \frac{x}{2\pi}Z_n(\phi_\eta)\int_{-\infty}^{\infty}[D_n(z,z_\mu) - D_n(z,\bar{z}_\mu)]\phi(\mu)\,d\mu + o(1),
\end{aligned}
$$

where $\tilde{Y}_n(z,x) = \sqrt{\frac{b_n}{n}}Y_n(z,x)$. Therefore,

$$
\tilde{Y}_n(z,x) = Z_n(\phi_\eta)\frac{x}{2\pi}\int_{-\infty}^{\infty}\left(C_n(z,z_\mu) - C_n(z,\bar{z}_\mu)\right)\phi(\mu)\,d\mu + o(1)
$$

uniformly in z with $\Im z = \eta$, where $C_n(z,z_\mu) = \frac{D_n(z,z_\mu)}{1-2f^2(z)}$ and $f(z)$ is given in (A.12). Hence

$$
\begin{aligned}
\frac{d}{dx}Z_n(\phi_\eta) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi(\mu)\left(\tilde{Y}_n(z_\mu,x) - \tilde{Y}_n(\bar{z}_\mu,x)\right)\,d\mu \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi(\mu_1)\left[\frac{x}{2\pi}Z_n(\phi_\eta)\int_{-\infty}^{\infty}\phi(\mu_2)\left(C_n(z_{\mu_1},z_{\mu_2}) - C_n(z_{\mu_1},\bar{z}_{\mu_2})\right)d\mu_2\right. \\
&\left.-\frac{x}{2\pi}Z_n(\phi_\eta)\int_{-\infty}^{\infty}\phi(\mu_2)\left(C_n(\bar{z}_{\mu_1},z_{\mu_2}) - C_n(\bar{z}_{\mu_1},\bar{z}_{\mu_2})\right)d\mu_2\right]d\mu_1 + o(1) \\
&= -\frac{x}{4\pi^2}Z_n(\phi_\eta)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(\mu_1)\phi(\mu_2)\left[C_n(z_{\mu_1},\bar{z}_{\mu_2}) + C_n(\bar{z}_{\mu_1},z_{\mu_2})\right. \\
&\left.-C_n(z_{\mu_1},z_{\mu_2}) - C_n(\bar{z}_{\mu_1},\bar{z}_{\mu_2})\right]d\mu_2 d\mu_1 + o(1) \\
&= -xZ_n(\phi_\eta)V_n(\phi,\eta) + o(1).
\end{aligned}
$$

To find the limit of $V_n(\phi,\eta)$, we shall calculate limit of $[C_n(z_{\mu_1},\bar{z}_{\mu_2}) + C_n(\bar{z}_{\mu_1},z_{\mu_2}) - C_n(z_{\mu_1},z_{\mu_2}) - C_n(\bar{z}_{\mu_1},\bar{z}_{\mu_2})]$ as $n \to \infty$. Using (A.5) and (A.6),

$$
\begin{aligned}
D_n(z,z_\mu) &= \frac{(1+\mathbb{E}[B(z_\mu)])\mathbb{E}[b_n\mathbb{E}_1\{A^\circ(z)A^\circ(z_\mu)\}]}{\mathbb{E}^2[A(z)]\mathbb{E}^2[A(z_\mu)]} - \frac{\mathbb{E}[b_n\mathbb{E}_1\{A^\circ(z)B^\circ(z_\mu)\}]}{\mathbb{E}^2[A(z)]\mathbb{E}[A(z_\mu)]} \\
&= \frac{1+\mathbb{E}[B(z_\mu)]}{\mathbb{E}^2[A(z)]\mathbb{E}^2[A(z_\mu)]}\mathbb{E}\Big[\frac{2}{b_n}\sum_{i,j\in I_1}G_{ij}^{(1)}(z)G_{ij}^{(1)}(z_\mu) + \sigma^2 + \frac{\kappa_4}{b_n}\sum_{i\in I_1}G_{ii}^{(1)}(z)G_{ii}^{(1)}(z_\mu) \\
&+ \frac{1}{b_n}\widetilde{\gamma_{n-1}}(z)\widetilde{\gamma_{n-1}}(z_\mu)])\Big] - \frac{1}{\mathbb{E}^2[A(z)]\mathbb{E}[A(z_\mu)]}\frac{d}{dz_\mu}\mathbb{E}\Big[\frac{2}{b_n}\sum_{i,j\in I_1}G_{ij}^{(1)}(z)G_{ij}^{(1)}(z_\mu) \\
&+ \sigma^2 + \frac{\kappa_4}{b_n}\sum_{i\in I_1}G_{ii}^{(1)}(z)G_{ii}^{(1)}(z_\mu) + \frac{1}{b_n}\widetilde{\gamma_{n-1}}(z)\widetilde{\gamma_{n-1}}(z_\mu)])\Big].
\end{aligned}
$$

Now using (A.7), we get

$$\left|\mathbb{E}\left[\frac{1}{b_n}\widetilde{\gamma_{n-1}}(z)\widetilde{\gamma_{n-1}}(z_\mu)]\right)\right| \le \frac{1}{b_n}\sqrt{\mathrm{Var}\sum_{i\in I_1} G_{ii}^{(1)}}\sqrt{\mathrm{Var}\sum_{i\in I_1} G_{ii}^{(1)}} = O\left(\frac{1}{b_n}\right).$$

Letting $n\to\infty$, using (A.13) we have

$$\lim_{n\to\infty} D_n(z,z_\mu) = f^2(z)f^2(z_\mu)(1+2f'(z_\mu))\left[\lim_{n\to\infty}\mathbb{E}[T_n] + \sigma^2 + \kappa_4 \lim_{n\to\infty}\frac{1}{b_n}\sum_{i\in I_1}\mathbb{E}\left[G_{ii}^{(1)}(z)G_{ii}^{(1)}(z_\mu)\right]\right]$$

$$+ f^2(z)f(z_\mu)\frac{d}{dz_\mu}\left[\lim_{n\to\infty}\mathbb{E}[T_n] + \kappa_4 \lim_{n\to\infty}\frac{1}{b_n}\sum_{i\in I_1}\mathbb{E}\left[G_{ii}^{(1)}(z)G_{ii}^{(1)}(z_\mu)\right]\right], \tag{2.28}$$

where

$$T_n = \frac{2}{b_n}\sum_{i,j\in I_1} G_{ij}^{(1)}(z)G_{ij}^{(1)}(z_\mu).$$

Since $\mathrm{Var}(G_{ii}) = O(1/b_n)$ (see (A.7)), we have

$$\lim_{n\to\infty}\frac{1}{b_n}\sum_{i\in I_1}\mathbb{E}\left[G_{ii}^{(1)}(z)G_{ii}^{(1)}(z_\mu)\right] = \lim_{n\to\infty}\frac{1}{b_n}\sum_{i\in I_1}\mathbb{E}\left[G_{ii}^{(1)}(z)\right]\mathbb{E}\left[G_{ii}^{(1)}(z_\mu)\right] = 2f(z)f(z_\mu).$$

We shall show in the Appendix (A.0.1) that

$$\lim_{n\to\infty}\mathbb{E}[T_n] = \frac{1}{4\pi^3}\int_{-2\sqrt{2}}^{2\sqrt{2}}\int_{-2\sqrt{2}}^{2\sqrt{2}}\frac{\sqrt{8-x^2}\sqrt{8-y^2}}{(x-z)(y-z_\mu)}F(x,y)\mathbf{1}_{\{x\neq y\}}\,dxdy, \tag{2.29}$$

where

$$F(x,y) = 2\int_{-\infty}^{\infty}\frac{u-u^3}{2(1-u^2)^2+u^2(x^2+y^2)-u(1+u^2)xy}\,ds,$$

where $u = \frac{\sin s}{s}$. Therefore,

$$\begin{aligned}\lim_{n\to\infty} C_n(z_{\mu_1},z_{\mu_2}) = \frac{1}{1-2f^2(z_{\mu_1})}\Big[&f^2(z_{\mu_1})f^2(z_{\mu_2})(1+2f'(z_{\mu_2}))\lim_{n\to\infty}\mathbb{E}[T_n]\\ &+ f^2(z_{\mu_1})f(z_{\mu_2})\lim_{n\to\infty}\frac{d}{dz_{\mu_2}}\mathbb{E}[T_n] + \sigma^2 f^2(z_{\mu_1})f^2(z_{\mu_2})(1+2f'(z_{\mu_2}))\\ &+ 2\kappa_4\Big\{f^3(z_{\mu_1})f^3(z_{\mu_2})(1+2f'(z_{\mu_2})) + f^3(z_{\mu_1})f(z_{\mu_2})f'(z_{\mu_2})\Big\}\Big].\end{aligned}$$

Hence

$$\begin{aligned} V(\phi) &= \lim_{\eta \downarrow 0} \lim_{n\to\infty} V_n(\phi,\eta) \\ &= \frac{\kappa_4}{16\pi^2}\left(\int_{-2\sqrt{2}}^{2\sqrt{2}} \frac{4-\mu^2}{\sqrt{8-\mu^2}}\phi(\mu)\, d\mu\right)^2 + \frac{\sigma^2}{16\pi^2}\left(\int_{-2\sqrt{2}}^{2\sqrt{2}} \frac{\mu\phi(\mu)}{\sqrt{8-\mu^2}}\, d\mu\right)^2 \\ &+ \int_{-2\sqrt{2}}^{2\sqrt{2}}\int_{-2\sqrt{2}}^{2\sqrt{2}} \sqrt{(8-x^2)(8-y^2)}F(x,y) \\ &\times \int_{-2\sqrt{2}}^{2\sqrt{2}}\int_{-2\sqrt{2}}^{2\sqrt{2}} \frac{\mu_1\phi(\mu_1)}{(x-\mu_1)\sqrt{8-\mu_1^2}} \frac{\mu_2\phi(\mu_2)}{(x-\mu_2)^2\sqrt{8-\mu_2^2}} d\mu_1 d\mu_2\, dxdy. \end{aligned}$$

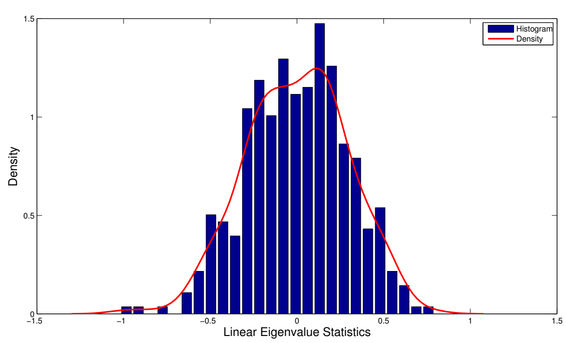
This completes the proof of (2.10) and the proof of Theorem 2.1.1. □

Recent development

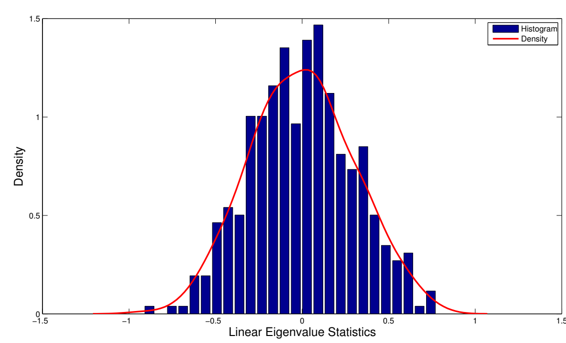
After submission of our result, M. Shcherbina [**Shc15**] improved our result by removing the restriction $b_n >> \sqrt{n}$ and proved it for all b_n which satisfies $b_n \to \infty$ and $\frac{b_n}{n} \to 0$ as $n \to \infty$.

2.3. Numerical simulations

Numerical simulations show that the CLT is true for any bandwidth b_n as long as $b_n \to \infty$. Here is what we found in MATLAB simulations.

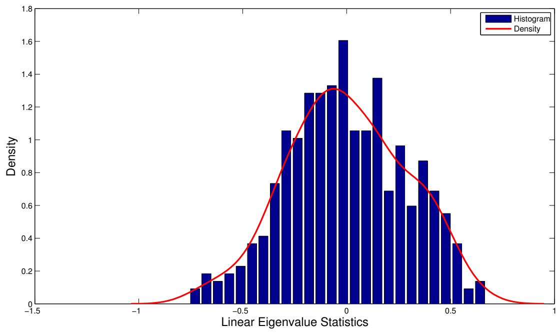


(a) $n = 2000$, $b_n = n^{0.2}$.
Fourth moment/(variance)^2=2.92

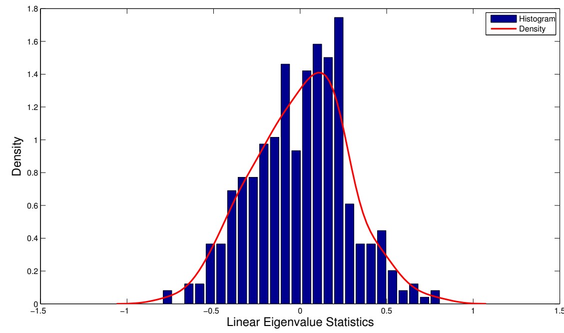


(b) $n = 2000$, $b_n = n^{0.4}$.
Fourth moment/(variance)^2=2.71

In the following example we had taken a different test function.

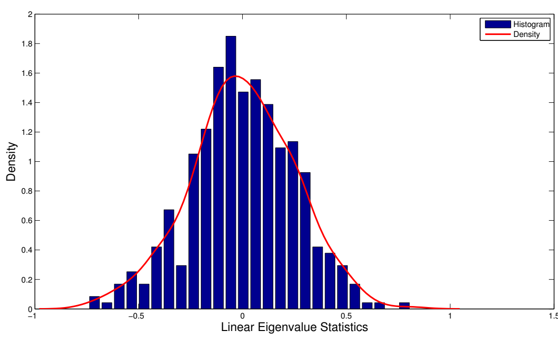


(a) $n = 2000$, $b_n = n^{0.6}$.
Fourth moment/(variance)^2=2.57

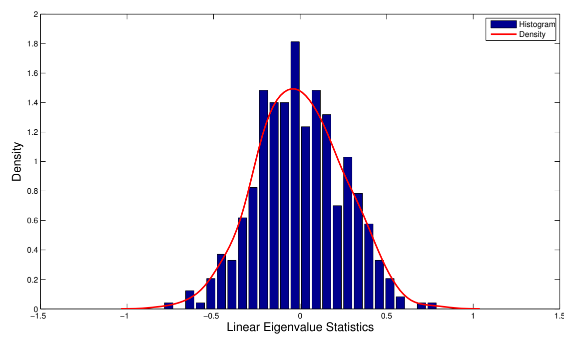


(b) $n = 2000$, $b_n = n^{0.8}$.
Fourth moment/(variance)^2=2.91

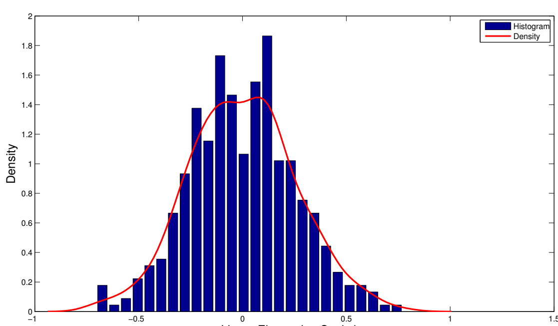
FIGURE 2.2. The eigenvalue statistics was sampled 400 times. The test function was $\phi(x) = \sqrt{16 - x^2}$.



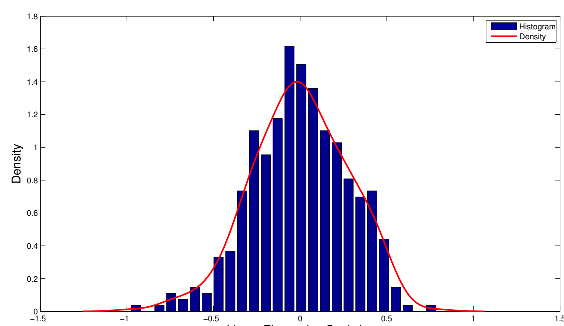
(a) $n = 2000$, $b_n = n^{0.2}$.
Fourth moment/(variance)^2=3.08



(b) $n = 2000$, $b_n = n^{0.4}$.
Fourth moment/(variance)^2=2.94



(c) $n = 2000$, $b_n = n^{0.6}$.
Fourth moment/(variance)^2=3.00



(d) $n = 2000$, $b_n = n^{0.8}$.
Fourth moment/(variance)^2=3.08

FIGURE 2.3. The eigenvalue statistics was sampled 400 times. The test function was $\phi(x) = e^{-x^2}$.

CHAPTER 3

Empirical Spectral Distribution of Singular Values

In this Chapter, we shall prove the Marčenko-Pastur law for random band matrices. Ginibre [**Gin65**] showed that if $M = \frac{1}{\sqrt{n}}X$, where x_{ij}, the entries of X, are i.i.d. complex normal variables, then the joint density of $\lambda_1, \ldots, \lambda_n$ is given by

$$f(\lambda_1, \ldots, \lambda_n) = c_n \prod_{i<j} |\lambda_i - \lambda_j|^2 \prod_{i=1}^{n} e^{-n|\lambda_i|^2},$$

where c_n is the normalizing constant. Using this, Mehta [**Meh67**] showed that the ESD μ_n of M converges to the uniform distribution on the unit disk. Later on, Girko [**Gir84, Gir85**] and Bai [**B$^+$97**] proved the result under more relaxed assumptions, namely under the assumption that $\mathbb{E}|X_{ij}|^6 < \infty$. Proving the result only under second moment assumption was open until Tao and Vu [**TV08, TVK10**].

Following the method used by Girko, and Bai, the real part of the Stieltjes transform $m_n(z) := \frac{1}{n}\sum_{i=1}^{n} \frac{1}{\lambda_i - z}$ can be written as

$$\begin{aligned} m_{nr}(z) &:= \Re(m_n(z)) \\ &= \frac{1}{n}\sum_{i=1}^{n} \frac{\Re(\lambda_i - z)}{|\lambda_i - z|^2} \\ &= -\frac{1}{2}\frac{\partial}{\partial \Re(z)} \int_0^\infty \log x \nu_n(dx, z), \end{aligned}$$

where $\nu_n(\cdot, z)$ is the ESD of $(\frac{1}{\sqrt{n}}X - zI)(\frac{1}{\sqrt{n}}X - zI)^*$. Secondly the characteristic function of $\frac{1}{\sqrt{n}}X$ satisfies [**Gir85**, section 1]

$$\int\int e^{i(ux+vy)}\mu_n(dx, dy) = \frac{u^2+v^2}{i4\pi u} \int\int \frac{\partial}{\partial s}\left[\int_0^\infty \log x \nu_n(dx, z)\right] e^{i(us+vt)}\, dtds,$$

for any $uv \neq 0$, and where $z = s + it$.

Thus, finding the limiting behaviour of $\nu_n(\cdot, z)$ is an essential ingredient in finding the limiting behaviour of $\mu_n(\cdot, \cdot)$. In this Chapter, we will focus on finding the limiting behaviour of $\nu_n(\cdot, z)$ for random band matrices so that it can be used for finding the limiting behaviour of $\mu_n(\cdot, \cdot)$ for random band matrices.

We consider the limiting ESD of matrices of the form $\frac{1}{2b_n+1}(R+X)(R+X)^*$, where X is an $n \times n$ band matrix of bandwidth b_n and R is a non random band matrix. Silverstein, Bai, and Dozier [**Sil95**, **SB95**, **DS07**] considered the ESD of $\frac{1}{n}(R+X)(R+X)^*$ type of matrices where X was $m \times n$ rectangular matrix with i.i.d. entries, R was a matrix independent of X, and the ratio $\frac{m}{n} \to c \in (0, \infty)$.

This Chapter is organized in the following way; in the Section 3.1, we formulate the band matrix model and state the main results. In Section 3.2, we give the main idea of the proof. In Section 3.6, we prove two concentration results which are the main ingredients of the proof. We shall use the tools from the Chapter A .

3.1. Main Results

Definition 3.1.1 (Periodic band matrix). An $n \times n$ matrix $M = (m_{ij})_{n \times n}$ is called a periodic band matrix of bandwidth b_n if $m_{ij} = 0$ whenever $b_n < |i-j| < n - b_n$.

M is called a non-periodic band matrix of bandwidth b_n if $m_{ij} = 0$ whenever $b_n < |i-j|$.

Notice that in case of a periodic band matrix, the maximum number of non-zero elements in each row is $2b_n + 1$. On the other hand, in case of a non-periodic band matrix, the number of non-zero elements in a row depends on the index of the row. For example, in the first row there are at most $b_n + 1$ non-zero elements. And in the $(b_n + 1)$th row there are at most $2b_n + 1$ many non-zero elements. In general, the ith row of a non-periodic band matrix has at most $b_n + i\mathbf{1}_{\{i \le b_n+1\}} + (b_n+1)\mathbf{1}_{\{b_n+1<i<n-b_n\}} + (n+1-i)\mathbf{1}_{\{i \ge n-b_n\}}$ many non-zero elements. In any case, the maximum number of non-zero elements is $O(b_n)$. In this context, let us define some index sets.

Let $M = (m_{ij})_{n\times n}$ be a random band matrix (periodic or non-periodic), then we define

$$(3.1)\qquad \begin{aligned} I_j &= \{1 \le k \le n : m_{jk} \text{ are not identically zero}\},\\ I'_k &= \{1 \le j \le n : m_{jk} \text{ are not identically zero}\}. \end{aligned}$$

Notice that in case of periodic band matrices, $|I_j| = 2b_n + 1 = |I'_k|$ for all j and k. Now we proceed to our main results.

Let $X = (x_{ij})_{n\times n}$ be an $n \times n$ periodic band matrix of bandwidth b_n, where $b_n \to \infty$ as $n \to \infty$. Let R be a sequence of $n \times n$ deterministic periodic band matrices of bandwidth b_n. Let us denote $c_n := 2b_n + 1$ and μ_M be the ESD of M. Assume that

$$(3.2)\qquad \begin{aligned} &(a)\ \mu_{\frac{1}{c_n}RR^*} \to H, \text{ for some non random probability distribution } H\\ &(b)\ \{x_{jk}:\ k \in I_j,\ 1 \le j \le n\} \text{ is an i.i.d. set of random variables,}\\ &(c)\ \mathbb{E}[x_{11}] = 0, \mathbb{E}[|x_{11}|^2] = 1,\\ &\text{and define } (d)\ Y = \frac{1}{\sqrt{c_n}}(R + \sigma X), \text{ where } \sigma > 0 \text{ is fixed.} \end{aligned}$$

For technical convenience, we assume the band matrices X and R are periodic. However, the following results can easily be extended to the case when the band matrix is not periodic. We will discuss it in the Section 3.5.

Let M be an $n \times n$ matrix. For convenience, let us introduce the following notation

$$\{\lambda_i(M) : 1 \le i \le n\} = \text{eigenvalues of } M,$$

$$m_j := (m_{1j}, m_{2j}, \ldots, m_{nj})^T$$

It is easy to see that $MM^* = \sum_{j=1}^{n} m_j m_j^*$.

Definition 3.1.2 (Poincaré inequality). Let X be a $\mathbb{R}^k$ valued random variable with probability measure μ. The probability measure μ is said to satisfy the Poincaré inequality with constant $m > 0$, if for all continuously differentiable functions $f : \mathbb{R}^k \to \mathbb{R}$,

$$\mathrm{Var}(f(X)) \le \frac{1}{m}\mathbb{E}(|\nabla f(X)|^2).$$

It can be shown that if μ satisfies the Poincaré inequality with constant m, then $\mu \otimes \mu$ also satisfies the Poincaré inequality with the same constant m [**GZ01**, Theorem 2.5]. It can also be shown that if μ satisfies Poincaré inequality and $f : \mathbb{R}^k \to \mathbb{R}$ is a continuously differentiable function then

$$\mathbb{P}_\mu \left(|f - \mathbb{E}_\mu(f)| > t\right) \leq 2K \exp\left(-\frac{\sqrt{m}}{\sqrt{2}\|\|\nabla f\|_2\|_\infty} t\right), \tag{3.3}$$

where $K = -\sum_{i\geq 0} 2^i \log(1 - 2^{-2i-1})$ [**AGZ10**, Lemma 4.4.3]. For example, Gaussian random variables satisfy the Poincaré inequality. Below, we formulate the main theorems of this Chapter.

THEOREM 3.1.3. *Let Y be the same as* (3.2). *In addition to the existing assumption, assume that*

(i) $(\log n)^2 = O(c_n)$

(ii) H is compactly supported

(iii) Both $\Re(x_{ij})$ and $\Im(x_{ij})$ satisfy the Poincaré inequality with constant m.

Then $\mathbb{E}|m_n(z) - m(z)|^2 \to 0$ uniformly on the compact subsets of $z \in \{z : \Im(z) > \eta\}$ for fixed $\eta > 0$, where $m_n(z) = \frac{1}{n}\sum_{i=1}^n (\lambda_i(YY^*) - z)^{-1}$ is the empirical Stieltjes transform of YY^*, and $m(z) = \int_{\mathbb{R}} \frac{d\mu(x)}{x-z}$ is non-random. In particular, the ESD of YY^* converges in L^2 norm. In addition, $m(z)$ satisfies

$$m(z) = \int_{\mathbb{R}} \frac{dH(t)}{\frac{t}{1+\sigma^2 m(z)} - (1 + \sigma^2 m(z))z} \quad \text{for any } z \in \mathbb{C}^+. \tag{3.4}$$

In particular, the above result is true for standard Gaussian random variables. The Poincaré inequality in the Theorem 3.1.3 simplifies the proof significantly. A similar result can also be obtained without using the Poincaré inequality. However in that case, we prove the theorem under the assumption that the bandwidth grows sufficiently faster. The theorem is formulated below.

THEOREM 3.1.4. *Let Y be the band matrix as defined in* (3.2). *In addition to the existing assumption, assume that*

(i) $\dfrac{n}{c_n^2} \to 0,$

(ii) H is compactly supported

(iii) $\mathbb{E}[|x_{11}|^{4p}] < \infty$, for some $p \in \mathbb{N}$.

Then $\mathbb{E}|m_n(z) - m(z)|^{2p} \to 0$ uniformly on compact subsets of $z \in \{z : \Im(z) > \eta\}$ for fixed $\eta > 0$, and the Stieltjes transform of μ satisfies (3.4)*.*

REMARK 3.1.5. If $c_n = n^\alpha$, where $\alpha > 0$, then the convergence in Theorem 3.1.3 is almost sure. And if $c_n = n^\beta$ where $\beta = \frac{1}{2} + \frac{1}{2p}$, then the convergence in Theorem 3.1.4 is almost sure. We will discuss about it at the ends of the Sections 3.2 and 3.3 respectively.

Notice that if we take $R = 0$ and $\sigma = 1$, then H is supported only at the real number 0. In that case, (3.4) simplifies to

$$m(z)(1 + m(z))z + 1 = 0,$$

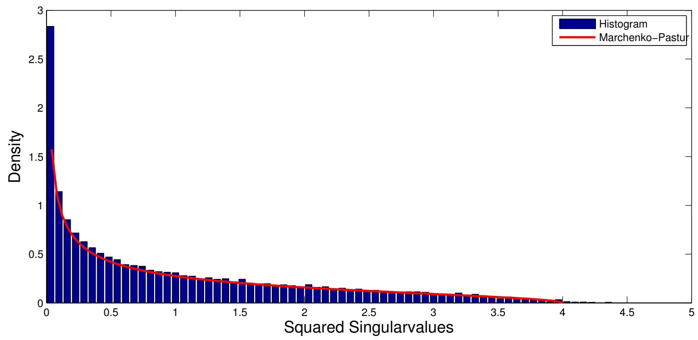
which is same as the equation (1.3) for $\gamma = 1$. The figure 3.1 shows some numerical evidence of Marčenko-Pastur limit law for random band matrices with i.i.d standard Gaussian entries.

Proof of the Theorem 3.1.4 contains the main idea of the proof of both of the theorems. Main structure of the proof is similar to the method described in [**DS07**]. However in case of band matrices, we need to prove a generalised version of the Lemma 3.1 in [**DS07**], which is proved in the Propositions 3.6.1 and 3.6.3. In addition, Lemma 3.6.2 gives a large deviation estimate of the norm of a random band matrix.

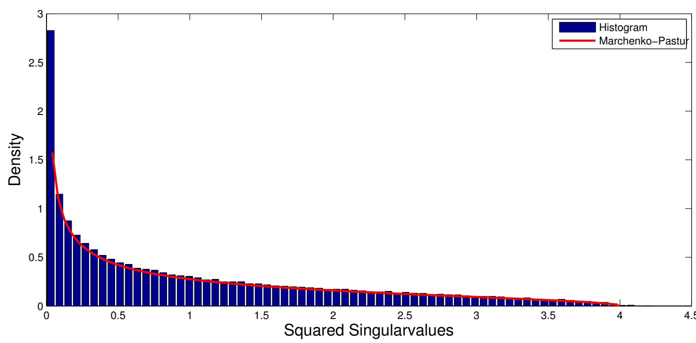
Also, it is not necessary for H to be compactly supported. We can truncate r_{ij} at a threshold of $\log(c_n)$ and have the same result as the Theorems 3.1.3 and 3.1.4. However in that case, we need the bandwidth c_n to grow a little faster; $\log(c_n)$ times faster than the existing rate of divergence. We will discuss it in the Section 3.4.

3.2. Proof of the Theorem 3.1.4

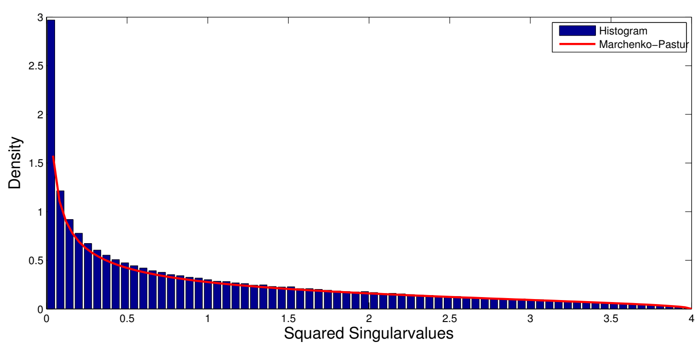
Let us define the empirical Stieltjes transform of YY^* as $m_n = \frac{1}{n}\sum_{i=1}^{n}(\lambda_i(YY^*) - z)^{-1}$. It is clear from the context that m_n depends on z. But, we omit it hereafter to avoid unnecessary



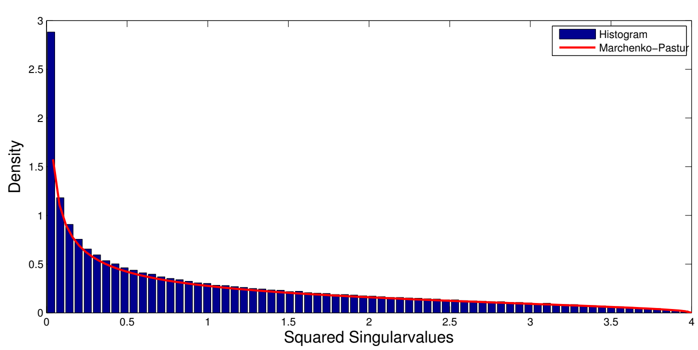
(a) Bandwidth $= 7000^{0.3}$; Non-Periodic band matrix



(b) Bandwidth $= 7000^{0.5}$; Periodic band matrix



(c) Bandwidth $= 7000^{0.7}$; Non-periodic band matrix



(d) Bandwidth $= 7000^{0.9}$; Periodic band matrix

FIGURE 3.1. The above simulations are done with 7000×7000 random band matrices with i.i.d. standard Gaussian entries and varying bandwidths.

notational complexity. Let $m : \mathbb{C}_+ \to \mathbb{C}$ be a complex analytic function and $\sigma > 0$. Define

$$f_{z,\sigma}(m,t) = \left[\frac{t}{1+\sigma^2 m} - (1+\sigma^2 m)z\right]^{-1}, \quad z \in \mathbb{C}_+, t \in \mathbb{R}. \tag{3.5}$$

Proof of this theorem is organized as follows;

(i) $m_n = \int f_{z,\sigma}(m_n,t)\, d\mu_{RR^*/c_n}(t) + \alpha_n$ such that $\|\alpha_n\|_{2p} \to 0$. Hereafter, $\|X\|_q := [\mathbb{E}|X|^q]^{1/q}$.

(ii) For a sufficiently large $\Im(z)$, $\{m_n\}_n$ is Cauchy in L^{2p}. Therefore, there exists m such that $\|m_n - m\|_{2p} \to 0$ as $n \to \infty$.

(iii) $\left\|\int f_{z,\sigma}(m_n,t)\, d\mu_{RR^*/c_n}(t) - \int f_{z,\sigma}(m,t) dH(t)\right\|_{2p} \to 0$.

(iv) Therefore, we have

$$\left\|m - \int f_{z,\sigma}(m,t)\, dH(t)\right\|_{2p} \le \|m_n - m\|_{2p} + \left\|\int f_{z,\sigma}(m,t)\, d\mu_{RR^*/c_n}(t) - \int f_{z,\sigma}(m,t)\, dH(t)\right\|_{2p} + \|\alpha_n\|_{2p} \to 0.$$

As a result, we have $m = \int f_{z,\sigma}(m,t)\, dH(t)$ almost surely.

(v) Solution to the equation $m = \int f_{z,\sigma}(m,t)\, dH(t)$ is unique. Since solution to this integral equation is unique, m is non-random.

The above mentioned method is summarized from [**DS07**]. However, the band structure plays a role in the technical parts of the proof of $\|\alpha_n\|_{2p} \to 0$.

We introduce the following notations which will be used in the proof of the theorem.

$$\begin{aligned}
A &= \frac{RR^*}{c_n(1+\sigma^2 m_n)} - \sigma^2 z m_n I \\
B &= A - zI \\
C &= YY^* - zI \\
C_j &= C - y_j y_j^* \\
m_n^{(j)} &= \frac{1}{n}\sum_{i=1}^{n}\left[\lambda_i(YY^* - y_j y_j^*) - zI\right]^{-1} = \frac{1}{n}\sum_{i=1}^{n}(\lambda_i(C_j))^{-1} \\
A_j &= \frac{RR^*}{c_n(1+\sigma^2 m_n^{(j)})} - \sigma^2 z m_n^{(j)} I \\
B_j &= A_j - zI.
\end{aligned} \tag{3.6}$$

Since $YY^* = \sum_{j=1}^n y_j y_j^*$, we observe that A_j, B_j, C_j are independent of y_j. This will become useful in several estimates; in particular, in the proof of Proposition 3.6.1.

REMARK 3.2.1. We notice that the eigenvalues of B are given by $\frac{\lambda_i}{1+\sigma^2 m_n} - (1+\sigma^2 m_n)z$, where λ_is are the eigenvalues of $\frac{1}{c_n}RR^*$. Therefore, $\int_{\mathbb{R}} \frac{dH(t)}{\frac{t}{1+\sigma^2 m} - (1+\sigma^2 m)z}$ can be thought of as $\frac{1}{n}\mathrm{tr}B^{-1}$ for large n. So heuristically, proving the theorem is same as showing that $\frac{1}{n}\mathrm{tr}B^{-1} - m_n \to 0$ as $n \to \infty$.

Let us define

$$\alpha_n = m_n - \frac{1}{n}\mathrm{tr}B^{-1}.$$

It is easy to see that $m_n = \int f_{z,\sigma}(m_n, t)\; d\mu_{RR^*/c_n}(t) + \alpha_n$, where $f_{z,\sigma}(m,t)$ is defined in (3.5). We first show that $\|\alpha_n\|_{2p} \to 0$ as $n \to \infty$.

Using the definition (3.6) and Lemma A.0.6, we have

$$\begin{aligned} I + zC^{-1} &= YY^*C^{-1} \\ &= \sum_{j=1}^n y_j y_j^* C^{-1} \\ &= \sum_{j=1}^n y_j \frac{y_j^* C_j^{-1}}{1 + y_j^* C_j^{-1} y_j}. \end{aligned}$$

Taking trace, and dividing by n on the both sides, we obtain

$$\begin{aligned} zm_n &= \frac{1}{n}\sum_{j=1}^n \frac{y_j^* C_j^{-1} y_j}{1 + y_j C_j^{-1} y_j^*} - 1 \\ &= -\frac{1}{n}\sum_{j=1}^n \frac{1}{1 + y_j^* C_j^{-1} y_j}. \end{aligned} \tag{3.7}$$

Using the resolvent identity,

$$\begin{aligned} B^{-1} - C^{-1} &= B^{-1}(YY^* - A)C^{-1} \\ &= \frac{1}{c_n}B^{-1}\left[RR^* + \sigma RX^* + \sigma XR^* + \sigma^2 XX^* - \frac{1}{1+\sigma^2 m_n}RR^* + c_n\sigma^2 z m_n\right]C^{-1} \end{aligned}$$

$$= \frac{1}{c_n}\sum_{j=1}^{n} B^{-1}\left[\frac{\sigma^2 m_n}{1+\sigma^2 m_n} r_j r_j^* + \sigma r_j x_j^* + \sigma x_j r_j^* + \sigma^2 x_j x_j^* - \frac{c_n}{n}\frac{1}{1+y_j^* C_j^{-1} y_j}\sigma^2\right] C^{-1}.$$

Taking the trace, dividing by n, and using (3.7), we have

$$\begin{aligned}
\frac{1}{n}\mathrm{tr}B^{-1} - m_n &= \frac{1}{n}\sum_{j=1}^{n}\left[\frac{\sigma^2 m_n}{1+\sigma^2 m_n}\frac{1}{c_n} r_j^* C^{-1}B^{-1} r_j + \frac{1}{c_n}\sigma x_j^* C^{-1}B^{-1} r_j + \frac{1}{c_n}\sigma r_j^* C^{-1}B^{-1} x_j\right.\\
&\quad \left. + \frac{1}{c_n}\sigma^2 x_j^* C^{-1}B^{-1} x_j - \frac{1}{1+y_j^* C_j^{-1} y_j}\frac{1}{n}\sigma^2 \mathrm{tr}C^{-1}B^{-1}\right]\\
&\equiv \frac{1}{n}\sum_{j=1}^{n}\left[T_{1,j}+T_{2,j}+T_{3,j}+T_{4,j}+T_{5,j}\right]. \qquad (3.8)
\end{aligned}$$

For convenience of writing $T_{i,j}$s, let us introduce some notations

$$\begin{aligned}
\rho_j &= \frac{1}{c_n} r_j^* C_j^{-1} r_j, \quad \omega_j = \frac{1}{c_n}\sigma^2 x_j^* C_j^{-1} x_j,\\
\beta_j &= \frac{1}{c_n}\sigma r_j^* C_j^{-1} x_j, \quad \gamma_j = \frac{1}{c_n}\sigma x_j^* C_j^{-1} r_j,\\
\hat{\rho}_j &= \frac{1}{c_n} r_j^* C_j^{-1} B^{-1} r_j, \quad \hat{\omega}_j = \frac{1}{c_n}\sigma^2 x_j^* C_j^{-1} B^{-1} x_j, \qquad (3.9)\\
\hat{\beta}_j &= \frac{1}{c_n}\sigma r_j^* C_j^{-1} B^{-1} x_j, \quad \hat{\gamma}_j = \frac{1}{c_n}\sigma x_j^* C_j^{-1} B^{-1} r_j,\\
\alpha_j &= 1 + \frac{1}{c_n}(r_j+\sigma x_j)^* C_j^{-1}(r_j+\sigma x_j) = 1+\rho_j+\beta_j+\gamma_j+\omega_j.
\end{aligned}$$

Using Lemma A.0.6 for $C = C_j + y_j y_j^* = C_j + \frac{1}{c_n}(r_j+\sigma x_j)(r_j+\sigma x_j)^*$ and the above notations, we can compute

$$\begin{aligned}
T_{1,j} &= \frac{1}{c_n}\frac{\sigma^2 m_n}{1+\sigma^2 m_n}\left[r_j^* C_j^{-1}B^{-1} r_j - \frac{1}{\alpha_j} r_j^* C_j^{-1} y_j y_j^* C_j^{-1} B^{-1} r_j\right]\\
&= \frac{1}{c_n\alpha_j}\frac{\sigma^2 m_n}{1+\sigma^2 m_n}\left[\alpha_j r_j^* C_j^{-1}B^{-1} r_j - \frac{1}{c_n} r_j^* C_j^{-1}(r_j r_j^* + \sigma r_j x_j^* + \sigma x_j r_j^* + \sigma^2 x_j x_j^*) C_j^{-1} B^{-1} r_j\right]\\
&= \frac{1}{\alpha_j}\frac{\sigma^2 m_n}{1+\sigma^2 m_n}\left[\alpha_j\hat{\rho}_j - (\rho_j\hat{\rho}_j + \rho_j\hat{\gamma}_j + \beta_j\hat{\rho}_j + \beta_j\hat{\gamma}_j)\right]\\
&= \frac{1}{\alpha_j}\frac{\sigma^2 m_n}{1+\sigma^2 m_n}[(1+\gamma_j+\omega_j)\hat{\rho}_j - (\rho_j+\beta_j)\hat{\gamma}_j].
\end{aligned}$$

Similarly,

$$T_{2j} = \frac{1}{\alpha_j}[(1+\rho_j+\beta_j)\hat{\gamma}_j - (\gamma_j+\omega_j)\hat{\rho}_j],$$
$$T_{3,j} = \frac{1}{\alpha_j}[(1+\gamma_j+\omega_j)\hat{\beta}_j - (\rho_j+\beta_j)\hat{\omega}_j],$$
$$T_{4,j} = \frac{1}{\alpha_j}[(1+\rho_j+\beta_j)\hat{\omega}_j - (\gamma_j+\omega_j)\hat{\beta}_j],$$

and,

$$T_{5,j} = -\frac{1}{\alpha_j}\frac{1}{n}\sigma^2 \mathrm{tr} C^{-1}B^{-1}.$$

Using the equations (3.7) and (3.8) and the above expressions, we can write

$$\begin{aligned}\frac{1}{n}\mathrm{tr}B^{-1} - m_n = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\alpha_j}\Bigg[&\frac{1}{1+\sigma^2 m_n}(\sigma^2 m_n - \gamma_j - \omega_j)\hat{\rho}_j \\ &+ \frac{1}{1+\sigma^2 m_n}(1+\rho_j+\beta_j+\sigma^2 m_n)\hat{\gamma}_j + \hat{\beta}_j + \hat{\omega}_j - \frac{1}{n}\sigma^2\mathrm{tr}C^{-1}B^{-1}\Bigg].\end{aligned} \tag{3.10}$$

We would like to show that the above quantity converges to zero in L^{2p} as $n\to\infty$. So, we start listing up some basic observations now.

Since x_{ij} are i.i.d. and $\mathbb{E}[|x_{ij}|^2]=1$, by the SLLN,

$$\frac{1}{nc_n}\mathrm{tr}XX^* = \frac{1}{nc_n}\sum_{|i-j|\le b_n}|x_{ij}|^2 \stackrel{a.s.}{\to} 1.$$

So, $\mu_{\frac{1}{c_n}XX^*}$ is almost surely tight. Using the condition (3.2)(a) and Lemma A.0.5 we conclude that μ_{YY^*} is almost surely tight. Therefore,

$$\delta := \inf_n \int \frac{1}{|\lambda - z|^2}\, d\mu_{YY^*}(\lambda) > 0.$$

As a result, for any $z\in\mathbb{C}^+$, we have

$$\begin{aligned}\Im(zm_n) &= \int \frac{\lambda\Im(z)}{|\lambda-z|^2}\, d\mu_{YY^*}(\lambda) \ge 0, \\ \Im(m_n) &= \int \frac{\Im(z)}{|\lambda-z|^2}\, d\mu_{YY^*}(\lambda) \ge \Im(z)\delta > 0.\end{aligned} \tag{3.11}$$

Let $z =\in \mathbb{C}^+ := \{z \in \mathbb{C} : \Im(z) > 0\}$, where $\Im(z)$ stands for the imaginary part of z. For any Hermitian matrix M, $\|(M - zI)^{-1}\| \leq \frac{1}{\Im(z)}$. Therefore,

$$\|C^{-1}\| \leq \frac{1}{\Im(z)}, \quad \|C_j^{-1}\| \leq \frac{1}{\Im(z)}. \tag{3.12}$$

We also have a similar bound for B^{-1}. If λ is an eigenvalue of $\frac{1}{c_n}RR^*$, then $\lambda(B) := \frac{1}{1+\sigma^2 m_n}\lambda - (1 + \sigma^2 m_n)z$ is the corresponding eigenvalue of B. So,

$$|\lambda(B)| \geq |\Im\lambda(B)| = \left|\frac{\sigma^2\Im(m_n)}{|1 + \sigma^2 m_n|^2}\lambda + \sigma^2\Im(zm_n) + \Im(z)\right| \geq \Im(z), \tag{3.13}$$

where the last inequality follows from (3.11).

We can do the similar calculation for B_j. As a result, we have

$$\|B^{-1}\| \leq \frac{1}{\Im(z)}, \quad \|B_j^{-1}\| \leq \frac{1}{\Im(z)}. \tag{3.14}$$

Secondly, we would like to estimate the effect of rank one perturbation on C and B. More precisely, we would like to estimate $C^{-1} - C_j^{-1}$ and $B^{-1} - B_j^{-1}$. Using the Lemma A.0.7, we have

$$\begin{aligned} &\left|\operatorname{tr}(C^{-1} - C_j^{-1})\right| \leq \frac{1}{|\Im(z)|}, \\ &\left|m_n - m_n^{(j)}\right| = \frac{1}{n}\left|\operatorname{tr}(C^{-1} - C_j^{-1})\right| \leq \frac{1}{n|\Im(z)|}. \end{aligned} \tag{3.15}$$

Using the estimates (3.11) for $z \in \mathbb{C}^+$, we have

$$|1 + \sigma^2 m_n| = \frac{|z + \sigma^2 z m_n|}{|z|} \geq \frac{1}{|z|}|\Im(z) + \sigma^2\Im(zm_n)| \geq \frac{\Im(z)}{|z|}.$$

Similarly, we also have $|1 + \sigma^2 m_n^{(j)}| \geq \frac{\Im(z)}{|z|}$ for $z \in \mathbb{C}^+$.

Therefore, using the estimates (3.14), (3.15) and the estimate of $\|RR^*\|$ from subsection 3.2.1, we have

$$\begin{aligned} \|B^{-1} - B_j^{-1}\| &= \|B^{-1}(B_j - B)B_j^{-1}\| \\ &\leq \frac{1}{|\Im(z)|^2}\|B_j - B\| \\ &= |m_n - m_n^{(j)}|\frac{\sigma^2}{|\Im(z)|^2}\left\|\frac{1}{c_n(1 + \sigma^2 m_n)(1 + \sigma^2 m_n^{(j)})}RR^* + zI\right\| \leq \frac{K\sigma^2}{n}. \end{aligned} \tag{3.16}$$

Here and in the following estimates, $K > 0$ is a constant that depends only on $p, \Im(z)$ and the moments of x_{ij}.

Now we start estimating several components of the equation (3.10).

3.2.1. Estimates of $\hat{\rho}_j$ and ρ_j. According to our assumptions, we have $\mu_{\frac{1}{c_n}RR^*} \to H$, where H is compactly supported. Then there exists $K > 0$ such that

$$\|r_j\|^2 = \|r_j r_j^*\| \leq \|RR^*\| \leq Kc_n. \tag{3.17}$$

Using the estimates (3.12) and (3.14), we have

$$|\hat{\rho}_j| \leq Kc_n, \quad |\rho_j| \leq Kc_n,$$

where $K > 0$ is a constant which depends only on the imaginary part of z.

3.2.2. Estimates of $\gamma_j, \beta_j, \hat{\gamma}_j$ and $\hat{\beta}_j$. Using Proposition 3.6.1 and equations (3.12),(3.14), (3.17), we have

$$\begin{aligned}
\mathbb{E}[|\gamma_j|^{4p}] &= \frac{1}{c_n^{4p}}\mathbb{E}\left|x_j^* C_j^{-1} r_j r_j^* (C_j^{-1})^* x_j\right|^{2p} \\
&\leq \frac{K}{c_n^{4p}}\mathbb{E}\left|x_j^* C_j^{-1} r_j r_j^* (C_j^{-1})^* x_j - \frac{c_n}{n}\mathrm{tr}(C_j^{-1} r_j r_j^* (C_j^{-1})^*)\right|^{2p} + \frac{K}{c_n^{2p} n^{2p}}\mathbb{E}\left|r_j^* C_j^{-1} C_j^{-1*} r_j\right|^{2p} \\
&\leq \frac{Kn^p}{c_n^{4p}}\|r_j r_j^*\|^{2p} + \frac{K}{c_n^{2p} n^{2p} |\Im(z)|^{4p}}\|r_j\|^{4p} \leq \frac{Kn^p}{c_n^{2p}} + \frac{1}{n^{2p}|\Im(z)|^{4p}} \leq \frac{Kn^p}{c_n^{2p}}.
\end{aligned}$$

Similarly, we can show that

$$\mathbb{E}[|\beta_j|^{4p}] \leq \frac{Kn^p}{c_n^{2p}}.$$

Notice that there are c_n many non-trivial elements in the vector x_j and $\mathbb{E}[|x_{11}|^2]$=1. Therefore, $\mathbb{E}\|x_j\|^2 = c_n$. Similarly,

$$\mathbb{E}\|x_j\|^{2p} \leq Kc_n^p.$$

To estimate $\hat{\gamma}_j$, we are going to use Proposition 3.6.1, and equations (3.12),(3.14) (3.17), (3.16).

$$\mathbb{E}\,|\hat{\gamma}_j|^{4p} \leq \frac{K}{c_n^{4p}}\mathbb{E}\left|x_j^* C_j^{-1} B_j^{-1} r_j\right|^{4p} + \frac{K}{c_n^{4p}}\mathbb{E}\left|x_j^* C_j^{-1}(B^{-1} - B_j^{-1}) r_j\right|^{4p}$$

$$
\begin{aligned}
&= \frac{K}{c_n^{4p}}\mathbb{E}\left|x_j^* C_j^{-1} B_j^{-1} r_j r_j^* B_j^{-1*} C_j^{-1*} x_j\right|^{2p} + \frac{K c_n^{2p} c_n^{2p}}{(nc_n)^{4p}} \\
&\le \frac{K}{c_n^{4p}}\mathbb{E}\left|x_j^* C_j^{-1} B_j^{-1} r_j r_j^* B_j^{-1*} C_j^{-1*} x_j - \frac{c_n}{n}\mathrm{tr}(C_j^{-1} B_j^{-1} r_j r_j^* B_j^{-1*} C_j^{-1*})\right|^{2p} \\
&+ \frac{K}{c_n^{2p} n^{2p}}\mathbb{E}\left|\mathrm{tr}(C_j^{-1} B_j^{-1} r_j r_j^* B_j^{-1*} C_j^{-1*})\right|^{2p} + \frac{K}{n^{4p}} \\
&\le \frac{Kn^p}{c_n^{2p}} + \frac{K}{n^{2p}} + \frac{K}{n^{4p}} \le \frac{Kn^p}{c_n^{2p}}.
\end{aligned}
$$

Similarly,

$$
\mathbb{E}[|\hat{\beta}_j|^{4p}] \le \frac{Kn^p}{c_n^{2p}}.
$$

3.2.3. Estimates of ω_j and $\hat{\omega}_j$. Using the Proposition 3.6.1, Lemma A.0.7 and the estimates (3.12), (3.14), (3.15), (3.16), we can write

$$
\begin{aligned}
&\frac{1}{\sigma^{4p}}\mathbb{E}\left|\hat{\omega}_j - \frac{\sigma^2}{n}\mathrm{tr}C^{-1}B^{-1}\right|^{2p} \\
&= \frac{1}{\sigma^{4p}}\mathbb{E}\left|\frac{1}{c_n}\sigma^2 x_j^* C_j^{-1} B^{-1} x_j - \frac{\sigma^2}{n}\mathrm{tr}C^{-1}B^{-1}\right|^{2p} \\
&\le \frac{K}{c_n^{2p}}\mathbb{E}\left|x_j^* C_j^{-1}(B^{-1} - B_j^{-1}) x_j\right|^{2p} + \frac{K}{c_n^{2p}}\mathbb{E}\left|x_j^* C_j^{-1} B_j^{-1} x_j - \frac{c_n}{n}\mathrm{tr}C_j^{-1}B_j^{-1}\right|^{2p} \\
&+ \frac{K}{n^{2p}}\mathbb{E}\left|\mathrm{tr}(C^{-1} - C_j^{-1})B^{-1}\right|^{2p} + \frac{K}{n^{2p}}\mathbb{E}\left|\mathrm{tr}C_j^{-1}(B^{-1} - B_j^{-1})\right|^{2p} \\
&\le \frac{K}{c_n^{2p} n^{2p}}\mathbb{E}\|x_j\|^{2p} + \frac{Kn^p}{c_n^{2p}} + \frac{K}{n^{2p}} + \frac{K}{n^{2p}} \le \frac{Kn^p}{c_n^{2p}}.
\end{aligned}
$$

Similarly, it can be shown that

$$
\frac{1}{\sigma^{4p}}\mathbb{E}\left|\omega_j - \sigma^2 m_n\right|^{2p} = \frac{1}{\sigma^{4p}}\mathbb{E}\left|\omega_j - \frac{\sigma^2}{n}\mathrm{tr}C^{-1}\right|^{2p} \le \frac{Kn^p}{c_n^{2p}}.
$$

The above completes the estimates of the main components of (3.10). Finally, we notice that if $z \in \mathbb{C}^+$, then $\Im(z y_j^*(C_j - zI)^{-1} y_j) \ge 0$. As a result, we have $|z\alpha_j| \ge |\Im(z)|$.

Plugging in all the above estimates into (3.10), we obtain

$$
\mathbb{E}\left|\frac{1}{n}\mathrm{tr}B^{-1} - m_n\right|^{2p} \le \frac{1}{n}\sum_{j=1}^{n}\frac{Kn^p}{c_n^{2p}} = \frac{Kn^p}{c_n^{2p}} \to 0. \tag{3.18}
$$

This completes the proof of $\|\alpha_n\|_{2p} \to 0$. Now we would like to show that $\|m_n - m_l\|_{2p} \to 0$ as $n, l \to \infty$. First of all, we notice that

$$\begin{aligned}
f_{z,\sigma}(m_n,t) - f_{z,\sigma}(m_l,t) &= \left[\frac{1}{\frac{t}{1+\sigma^2 m_n} - (1+\sigma^2 m_n)z} - \frac{1}{\frac{t}{1+\sigma^2 m_l} - (1+\sigma^2 m_l)z}\right] \\
&= (m_n - m_l)\left[\frac{\frac{\sigma^2 t}{(1+\sigma^2 m_n)(1+\sigma^2 m_l)} + z\sigma^2}{\left(\frac{t}{1+\sigma^2 m_n} - (1+\sigma^2 m_n)z\right)\left(\frac{t}{1+\sigma^2 m_l} - (1+\sigma^2 m_l)z\right)}\right] \\
&= (m_n - m_l) f_{z,\sigma}(m_n,t) f_{z,\sigma}(m_l,t)\left[\frac{\sigma^2 t}{(1+\sigma^2 m_n)(1+\sigma^2 m_l)} + z\sigma^2\right]. \qquad (3.19)
\end{aligned}$$

Moreover from (3.11), we know that if $\Im(z) > 0$ then $\Im(zm_n) \geq 0, \Im(m_n) \geq \delta\Im(z) > 0$. Similarly, $\Im(zm) \geq 0$ and $\Im(m) \geq \delta\Im(z) > 0$. From (3.13), we have $|f_{z,\sigma}(m_{n_k},t)| < \frac{1}{\Im(z)}$ for all $t \geq 0$. Similarly, $|f_{z,\sigma}(m,t)| < \frac{1}{\Im(z)}$ for all $t \geq 0$. Therefore,

$$\begin{aligned}
|f_{z,\sigma}(m_n,t) - f_{z,\sigma}(m,t)| &\leq \sigma^2 |m_n - m_l|\left[\frac{|z|}{\Im(z)^2} + \left|\frac{f_{z,\sigma}(m_n,t)}{1+\sigma^2 m_n}\right|\left|\frac{t f_{z,\sigma}(m_l,t)}{1+\sigma^2 m_l}\right|\right] \\
&\leq \sigma^2 |m_n - m_l|\left[\frac{|z|}{\Im(z)^2} + \frac{1}{\sigma^2\delta\Im(z)^2}\left|\frac{t f_{z,\sigma}(m_l,t)}{1+\sigma^2 m_l}\right|\right] \\
&\leq \sigma^2 |m_n - m_l|\left[\frac{|z|}{\Im(z)^2} + \frac{1}{\sigma^2\delta\Im(z)^2}\left|1 + (1+\sigma^2 m_l) z f_{z,\sigma}(m_l,t)\right|\right] \\
&\leq \sigma^2 |m_n - m_l|\left[\frac{|z|}{\Im(z)^2} + \frac{1}{\sigma^2\delta\Im(z)^2}\left\{1 + \frac{\left(1 + \frac{\sigma^2}{\Im(z)}\right)|z|}{\Im(z)}\right\}\right].
\end{aligned}$$

Let us define $E := \{z \in \mathbb{C}_+ : \Re(z) \leq 1, \Im(z) > \eta\}$, where $\eta > 0$ is sufficiently large such that from the above we have

$$|f_{z,\sigma}(m_n,t) - f_{z,\sigma}(m,t)| \leq \phi|m_n - m_l|, \;\; \phi < \frac{1}{2}. \qquad (3.20)$$

Thus using the fact that $m_n = \int f_{z,\sigma}(m_n,t)\, d\mu_{RR^*/c_n}(t) + \alpha_n$, we have $\|m_n - m_l\|_{2p} \leq 2\phi\|m_n - m_l\|_{2p} + \|\alpha_n\|_{2p} + \|\alpha_l\|_{2p}$. Since $\phi < 1/2$ and $\|\alpha_k\|_{2p} \to 0$ as $k \to \infty$, we have $\|m_n - m_l\|_{2p} \to 0$ as $n, l \to \infty$. Since $L^{2p}(\Omega)$ is complete, there exits $m : \mathbb{C}_+ \to \mathbb{C}$ such that $\|m_n(z) - m(z)\|_{2p} \to 0$ for each $z \in E$. Thus for each $z \in E$, we can find a subsequence $\{m_{n_k(z)}\}_{n_k(z)}$ such that $m_{n_k(z)}(z) \to m(z)$ a.s.. Let $\{z_i\} \subset E$ be a sequence such that $z_i \to z^* \in E^\circ$. Then we can construct a subsequence $\{m_{n_j}\}$ such that $m_{n_j}(z_i) \to m(z_i)$ a.s. for all z_i. Since $|m_n(z)| < \frac{1}{\Im(z)}$ for all n, by Vitali-Porter theorem, we may conclude that $m_{n_j}(z) \to m(z)$ a.s. and uniformly on the compact sets of $\mathbb{C}_+$, where $m(z)$ is an analytic function on $\mathbb{C}_+$ a.s..

Now, we show that the solution of (3.4) is unique. Suppose $m_1, m_2 : \mathbb{C}_+ \to \mathbb{C}$ be two complex analytic functions such that $\Im(zm_1) > 0$ and $\Im(zm_2) > 0$ and they satisfy (3.4). From (3.20), we know that for sufficiently large $\Im(z) > 0$

$$|f_{z,\sigma}(m_1, t) - f_{z,\sigma}(m_2, t)| = \phi|m_1 - m_2|, \;\; \phi < \frac{1}{2}$$

Therefore if both m_1 and m_2 satisfies (3.4), then we have $|m_1 - m_2| \leq \frac{1}{2}|m_1 - m_2|$ for $z \in E$. Therefore $m_1 = m_2$ for $z \in E$. But since m_1, m_2 are analytic functions on $\mathbb{C}_+$, we have $m_1 = m_2$ on $\mathbb{C}_+$.

This completes the proof of the Theorem 3.1.4.

From the estimate (3.18), We see that if $c_n = n^\beta$, where $\beta > \frac{1}{2} + \frac{1}{2p}$, then $\sum_{n=1}^{\infty} \frac{n^p}{c_n^{2p}} < \infty$. Therefore, by Borel-Cantelli lemma, we can conclude that $\frac{1}{n}\mathrm{tr}B^{-1} - m_n \to 0$ almost surely.

3.3. Proof of the Theorem 3.1.3

Proof of the Theorem 3.1.3 is exactly same as the proof of Theorem 3.1.4. We notice that we obtained the bound $O\left(\frac{n^p}{c_n^{2p}}\right)$ because of the proposition 3.6.1. Therefore, while estimating the bounds of several components of equation (3.10), instead of using the proposition 3.6.1, we will use the Proposition 3.6.3. And by doing so we can obtain that $\mathbb{E}\left|\frac{1}{n}\mathrm{tr}B^{-1} - m_n\right|^2 = O(1/c_n)$. Which will conclude the Theorem 3.1.3.

To prove the almost sure convergence, we can truncate all the entries of the matrix X by $6\sqrt{\frac{2}{m}}\log n$. Let us denote that truncated matrix as $\tilde{X}$. Since x_{ij}s satisfy the Poincaré inequality, from (3.3) we have

$$\mathbb{P}\left(|x_{ij}| > t\right) \leq 2K \exp\left(-\sqrt{\frac{m}{2}}t\right).$$

Therefore,

$$\mathbb{P}\left(X \neq \tilde{X}\right) \leq 2Kn^2 \exp\left(-6\log n\right) \leq \frac{K'}{n^4}.$$

Now using the second part of Proposition 3.6.3 and following the same method as described in Section 3.2, we have

$$\mathbb{E}\left[\left|\frac{1}{n}\mathrm{tr}B^{-1} - m_n\right|^{2l} \mathbf{1}_{\{X=\tilde{X}\}}\right] \leq K\frac{(\log n)^{2l}}{c_n^l}.$$

Since $\left|\frac{1}{n}\mathrm{tr}B^{-1}\right|, |m_n| \leq |\Im z|^{-1}$, we have

$$\mathbb{E}\left[\left|\frac{1}{n}\mathrm{tr}B^{-1} - m_n\right|^{2l}\right] \leq K\frac{(\log n)^{2l}}{c_n^l} + \frac{K}{|\Im z|^{2l} n^4}.$$

If $c_n = n^\alpha$, $\alpha > 0$, then taking l large enough and using the Borel-Cantelli lemma we may conclude the almost sure convergence.

3.4. Truncation of R

In several estimates, it was convenient when we have bounded r_{ij}. However, we can achieve the same by properly truncating the random variables. Below we have described the truncation method by following the same procedure as described in [**DS07**].

Let $\frac{1}{\sqrt{c_n}}R = USV$ be the singular value decomposition of R, where $S = diag[s_1, \ldots, s_n]$ are the singular values of R and U, V are orthonormal matrices. Let us construct a diagonal matrix S_α as $S_\alpha = diag[s_1\mathbf{1}(s_1 \leq \alpha), \ldots, s_n\mathbf{1}(s_n \leq \alpha)]$ and consider the matrices $R_\alpha = US_\alpha V$, $Y_\alpha = \frac{1}{\sqrt{c_n}}(R_\alpha + \sigma X)$. Then by Lemma A.0.9 we have

$$\begin{aligned}
\|\mu_{YY^*} - \mu_{Y_\alpha Y_\alpha^*}\| &\leq \frac{2}{n}\mathrm{rank}\left(\frac{R}{\sqrt{c_n}} - \frac{R_\alpha}{\sqrt{c_n}}\right) \\
&= \frac{2}{n}\sum_{i=1}^{n} \mathbf{1}(s_i > \alpha) \\
&= 2H(\alpha^2, \infty),
\end{aligned}$$

If we take $\alpha^2 \to \infty$ for example $\alpha = \log(c_n)$ then $\|\mu_{YY^*} - \mu_{Y_\alpha Y_\alpha^*}\| \to 0$. So without loss of generality, we can assume that $\|r_j\|^2 \leq \|RR^*\| \leq c_n \log(c_n)$. In that case, we have

$$\|r_j\|^2 = \|r_j r_j^*\| \leq \|RR^*\| \leq c_n \log(c_n).$$

So using the estimates (3.12) and (3.14), we have

$$|\hat{\rho}_j| \leq K c_n \log(c_n), \quad |\rho_j| \leq K c_n \log(c_n),$$

where $K > 0$ is a constant which depends only on the imaginary part of z. Similarly, all the places in the proof of Theorem 3.1.4 we can replace the estimates $|r_j r_j^*| \leq K c_n$ by the estimates $|r_j r_j^*| \leq K c_n \log(c_n)$.

3.5. Extension of the results to non-periodic band matrices

The result can easily be extended to non-periodic band matrices. We observe that for the purpose of our proof, the main difference between a periodic and a non-periodic band matrix is the number of elements in certain rows. In the case of a periodic band matrix the number of non-trivial elements in any row is $|I_j| = 2b_n + 1 = c_n$ which is fixed for any $1 \leq j \leq n$. Therefore, in the definition (3.9) we divide by c_n. For a non-periodic band matrix $|I_j| = b_n + i\mathbf{1}_{\{i \leq b_n+1\}} + (b_n + 1)\mathbf{1}_{\{b_n+1 < i < n-b_n\}}(n+1-i)\mathbf{1}_{\{i \geq n-b_n\}} = O(b_n)$. Once in the definition (3.9) and in the Proposition 3.6.1, Proposition 3.6.3 if we replace c_n by $|I_j|$, everything works out as before.

3.6. Two concentration results

In this Section we list two main concentration results which are used in the proofs of the Theorems 3.1.3, 3.1.4.

Proposition 3.6.1. *Let M be one of $C_j^{-1}, C_j^{-1}B_j^{-1}$, and N be one of $C_j^{-1} r_j r_j^* C_j^{-1*}$ or $C_j^{-1} B_j^{-1} r_j r_j^* B^{-1*} C_j^{-1*}$. Let x_j be the jth column of X as defined in Theorem 3.1.4. Let us also assume that $\mathbb{E}|x_{11}|^{4l} < \infty$. Then for any $l \in \mathbb{N}$,*

$$\mathbb{E}\left| x_j^* M x_j - \frac{c_n}{n} trM \right|^{2l} \leq K n^l$$
$$\mathbb{E}\left| x_j^* N x_j - \frac{c_n}{n} trN \right|^{2l} \leq K n^l \|r_j r_j^*\|^{2l}$$

where $K > 0$ is a constant that depends on l, $\Im(z)$, and the moments of x_j, but not on n.

Proof. From the estimates (3.12) and (3.14) we know that $\|C_j^{-1}\| \leq 1/|\Im(z)|$ and $\|B_j^{-1}\| \leq 1/|\Im(z)|$. So, for convenience of writing the proof, let us assume that $\|M\| \leq 1$ and $\|N\| \leq \|r_j r_j^*\|$. Also without loss of generality, we can assume that $j = 1$, and recall the definition of I_j from (3.1). We can write $M = P + iQ$, where P and Q are the real and imaginary parts of M respectively. Then we can write

$$\mathbb{E}\left|x_j^* M x_j - \frac{c_n}{n}\mathrm{tr}M\right|^{2l} \leq 2^{2l-1}\left|x_1^* P x_1 - \frac{c_n}{n}\mathrm{tr}P\right|^{2l} + 2^{2l-1}\mathbb{E}\left|x_1^* Q x_1 - \frac{c_n}{n}\mathrm{tr}Q\right|^{2l}.$$

We can write the first part as

$$\begin{aligned}
\left|x_1^* P x_1 - \frac{c_n}{n}\mathrm{tr}P\right|^{2l} &= \left|x_1^* P x_1 - \sum_{k\in I_1} P_{kk} + \sum_{k\in I_1} P_{kk} - \frac{c_n}{n}\mathrm{tr}P\right|^{2l} \\
&\leq 3^{2l-1}\mathbb{E}\left[\sum_{k\in I_1}(|x_{1k}|^2 - 1)P_{kk}\right]^{2l} + 3^{2l-1}\mathbb{E}\left[\sum_{\substack{i\neq j\\ i,j\in I_1}} P_{ij}\overline{x_{1i}}x_{1j}\right]^{2l} \\
&+ 3^{2l-1}\left|\sum_{k\in I_1} P_{kk} - \frac{c_n}{n}\mathrm{tr}P\right|^{2l} \\
&=: 3^{2l-1}(S_1 + S_2 + S_3).
\end{aligned}$$

Following the same procedure as in [**SB95**], we can estimate the first part. Note that $\|P^m\| \leq \|P\|^m \leq \|M\|^m \leq 1$ for any $m \in \mathbb{N}$. In the expansion of $\left[\sum_{k\in I_1}(|x_{1k}|^2 - 1)P_{kk}\right]^{2l}$, the maximum contribution (in terms of c_n) will come from the terms like

$$\sum_{k_1,\ldots,k_l\in I_1}(|x_{1k_1}|^2 - 1)^2 \cdots (|x_{1k_l}|^2 - 1)^2 (P_{i_1 i_1}\cdots P_{i_l i_l})^2,$$

when all $i_1, \ldots, i_l$ are distinct. Note that $(P_{i_1 i_1}\cdots P_{i_l i_l})^2 \leq 1$. Consequently, expectation of the above term is bounded by Kc_n^l, where K depends only on the fourth moment of x_{ij}. Therefore,

$$S_1 = \mathbb{E}\left[\sum_{k\in I_1}(|x_{1k}|^2 - 1)P_{kk}\right]^{2l} \leq Kc_n^l,$$

where K depends only on l and the moments of x_{ij}.

Since $C_1^{-1}, C_1^{-1}B_1^{-1}, C_1^{-1}r_1r_1^*C_1^{-1*}$ or $C_1^{-1}B_1^{-1}r_1r_1^*B^{-1*}C_1^{-1*}$ are independent of x_1, for the second sum we have

$$\sum_{\substack{i_1\neq j_1,\ldots,i_{2l}\neq j_{2l}\\ i_1,j_1,\ldots,i_{2l},j_{2l}\in I_1}} \mathbb{E}[P_{i_1j_1}\cdots P_{i_{2l}j_{2l}}]\mathbb{E}[\overline{x_{1i_1}}x_{1j_1}\cdots\overline{x_{1i_{2l}}}x_{1j_{2l}}].$$

The expectation will be zero if a term appears only once and the maximum contribution (in terms of c_n) will come from the case when each of x_{1j} and $\overline{x_{1j}}$ appears only twice. In that case, the contribution is

$$\sum_{\substack{i_1\neq j_1\\ i_1,j_1\in I_1}} P_{i_1j_1}^2\cdots\sum_{\substack{i_l\neq j_l\\ i_l,j_l\in I_1}} P_{i_lj_l}^2 \le c_n^l,$$

where the last inequality follows from the fact that $\sum_{i,j\in I_1} P_{ij}^2 = \mathrm{tr}(LPL^TLP^TL^T) \le c_n$, where $L_{c_n\times n}$ is the projection matrix onto the co-ordinates indexed by I_1. As a result, we have

$$S_2 = \mathbb{E}\left[\sum_{\substack{i\neq j\\ i,j\in I_1}} P_{ij}\overline{x_{1i}}x_{1j}\right]^{2l} \le Kc_n^l,$$

where K depends only on l and the moments of x_{ij}.

To estimate the S_3, we can write it as

$$S_3 = \left|\sum_{k\in I_1} P_{kk} - \frac{c_n}{n}\mathrm{tr}P\right|^{2l} = 2^{2l-1}\left|\sum_{k\in I_1} P_{kk} - \mathbb{E}\sum_{k\in I_1} P_{kk}\right|^{2l} + 2^{2l-1}\left|\mathbb{E}\sum_{k\in I_1} P_{kk} - \frac{c_n}{n}\mathrm{tr}P\right|^{2l}.$$

Since $|P_{kk} - \mathbb{E}[P_{kk}]| \le |(C_1^{-1})_{kk} - \mathbb{E}[(C_1^{-1})_{kk}]|$, from Lemma A.0.11 we have exponential tail bound on

$\left|\sum_{k\in I_1} P_{kk} - \mathbb{E}\sum_{k\in I_1} P_{kk}\right|$. As a result

$$\mathbb{E}\left|\sum_{k\in I_1} P_{kk} - \mathbb{E}\sum_{k\in I_1} P_{kk}\right|^{2l} \le Kn^l, \tag{3.21}$$

where K depends only on l.

Since x_{ij} are i.i.d., for any choice of M we have $\mathbb{E}[m_{11}] = \mathbb{E}[m_{ii}]$. Which implies that $\mathbb{E}[\sum_{k\in I_1} P_{kk}] = \frac{c_n}{n}\mathbb{E}[\mathrm{tr}P]$. Therefore, from Lemma A.0.11, we have

$$\begin{aligned}\left|\mathbb{E}\sum_{k\in I_1} P_{kk} - \frac{c_n}{n}\mathrm{tr}P\right|^{2l} &= \frac{c_n^{2l}}{n^{2l}}\left|\mathbb{E}[\mathrm{tr}P] - \mathrm{tr}P\right|^{2l}\\ &\leq K\frac{c_n^{2l}}{n^l} \leq Kc_n^l,\end{aligned}$$

where K depends only on l. Hence we have

$$S_3 \leq K(n^l + c_n^l).$$

Combining all the above estimates we have

$$\mathbb{E}\left|x_1^*Px_1 - \frac{c_n}{n}\mathrm{tr}P\right|^{2l} \leq Kn^l$$

Repeating the above computation we can do the same estimate $\mathbb{E}\left|x_1^*Qx_1 - \frac{c_n}{n}\mathrm{tr}Q\right|^{2l} \leq Kn^l$. This completes the proof.

□

LEMMA 3.6.2 (Norm of a random band matrix). *Let X be same as in the definitions* (3.2)*, x_{ij} satisfy the Poincaré inequality with constant m, and $c_n > (\log n)^2$. Then $\mathbb{E}\|XX^*\| \leq Kc_n^2$ for some universal constant K which may depend on the Poincaré constant m. In particular, if the limiting ESD of $\frac{1}{c_n}RR^*$ i.e., H is compactly supported then $\mathbb{E}\|YY^*\| \leq Kc_n$.*

PROOF. We will follow the method described in [**Tro15**, **MJC**$^+$**14**, **Tro12**] and the references therein. The analysis becomes somewhat easier if we assume that all non-zero entries of X are standard Gaussian random variables. However, the Gaussian case contains the main idea of the analysis.

Case I (x_{jk} are standard Gaussian random variables): Using the Markov's inequality, we have

$$\mathbb{P}\left(\frac{1}{c_n}\|XX^*\| > t\right) \leq e^{-t}\mathbb{E}\left[\exp\left(\frac{1}{c_n}\|XX^*\|\right)\right] \leq e^{-t}\mathbb{E}\left[\operatorname{tr}\exp\left(\frac{1}{c_n}XX^*\right)\right].$$

To estimate the right hand side, we will use the Lieb's theorem [**Lie73**, Theorem 6]. Let H be any $n \times n$ fixed Hermitian matrix. From Lieb's theorem, we know that the function $f(A) = \operatorname{tr}\exp(H + \log A)$ is a concave function on the convex cone of $n \times n$ positive definite Hermitian matrices.

Let us write $\frac{1}{c_n}XX^* = \sum_{k=1}^n x_k x_k^*$, where x_k is the kth column vector of $X/\sqrt{c_n}$. Then using the Lieb's theorem and Jensen's inequality, we have

$$\begin{aligned}\mathbb{E}\left[\operatorname{tr}\exp\left(\frac{1}{c_n}XX^*\right)\Bigg| x_1,\dots,x_{n-1}\right] &= \mathbb{E}\left[\operatorname{tr}\exp\left(\frac{1}{c_n}\sum_{k=1}^{n-1}x_kx_k^* + \log\exp\left(\frac{1}{c_n}x_nx_n^*\right)\right)\Bigg| x_1,\dots,x_{n-1}\right]\\ &\leq \operatorname{tr}\exp\left[\frac{1}{c_n}\sum_{k=1}^{n-1}x_kx_k^* + \log\mathbb{E}\exp\left(\frac{1}{c_n}x_nx_n^*\right)\right].\end{aligned}$$

Proceeding in this way, we obtain

$$\mathbb{E}\left[\operatorname{tr}\exp\left(\frac{1}{c_n}XX^*\right)\right] \leq \operatorname{tr}\exp\left[\sum_{k=1}^{n}\log\mathbb{E}\exp\left(\frac{1}{c_n}x_kx_k^*\right)\right].$$

Therefore,

$$\mathbb{P}\left(\frac{1}{c_n}\|XX^*\| > t\right) \leq e^{-t}\operatorname{tr}\exp\left[\sum_{k=1}^{n}\log\mathbb{E}\exp\left(\frac{1}{c_n}x_kx_k^*\right)\right]. \tag{3.22}$$

It is easy to see that

$$\begin{aligned}\exp\left(\frac{1}{c_n}x_kx_k^*\right) &= I + \left(\sum_{l=1}^{\infty}\frac{1}{l!c_n^l}\|x_k\|^{2(l-1)}\right)x_kx_k^*\\ &= I + \frac{e^{\|x_k\|^2/c_n}-1}{\|x_k\|^2}x_kx_k^*\\ &\preceq I + \frac{1}{c_n}e^{\|x_k\|^2/c_n}x_kx_k^*,\end{aligned}$$

where $A \preceq B$ denotes that $(B - A)$ is positive semi-definite. Since $\{x_{jk}\}_{1 \le k \le n,\, j \in I_k'}$ are independent standard Gaussian random variables, we have

$$\mathbb{E}\left[e^{\|x_k\|^2/c_n} x_{jk}\bar{x}_{lk}\right] = 0, \quad \text{if } j \neq l$$
$$\mathbb{E}\left[e^{\|x_k\|^2/c_n} |x_{jk}|^2\right] = \left(1 - \frac{1}{c_n}\right)^{-(c_n+1)}$$

As a result,

$$\operatorname{tr}\exp\left[\sum_{k=1}^{n} \log \mathbb{E}\exp\left(\frac{1}{c_n} x_k x_k^*\right)\right] \le n\left(1 + \frac{e}{c_n}\right)^{c_n}$$

Substituting this estimate in (3.22), we have

$$\mathbb{P}\left(\frac{1}{c_n}\|XX^*\| > t + \log n\right) \le e^e n e^{-(t+\log n)} = e^e e^{-t}. \tag{3.23}$$

As a result,

$$\begin{aligned}\frac{1}{c_n}\mathbb{E}[\|XX^*\|] &= \int_0^\infty \mathbb{P}\left(\frac{1}{c_n}\|XX^*\| > u\right)\, du \\ &\le \int_0^{\log n} du + \int_0^\infty \mathbb{P}\left(\frac{1}{c_n}\|XX^*\| > t + \log n\right)\, dt \\ &\le \log n + e^e \le K c_n.\end{aligned}$$

This completes the proof.

Case II (x_{jk}s satisfy the Poincaré inequality): First of all, let us write the random matrix X as $X = X_1 + iX_2$, where X_1 and X_2 are the real and imaginary parts of X respectively. Since $\|X\| \le \|X_1\| + \|X_2\|$, it is enough to estimate $\|X_1\|$ and $\|X_2\|$ separately. In other words, without loss of generality, we can assume that x_{ij} are real valued random variables.

Let us construct the matrix

$$\tilde{X} = \begin{bmatrix} O & X \\ X & O \end{bmatrix}.$$

It is easy to see that $\|\tilde{X}\| = \|X\|$. Therefore, it is enough to bound $\mathbb{E}\|\tilde{X}\|^2$.

We can write $\tilde{X}$ as

$$\tilde{X} = \sum_{i=1}^{n} \sum_{j \in I_i} x_{ij}(E_{i,n+j} + E_{n+j,i}),$$

where E_{ij} is a $2n \times 2n$ matrix with all 0 entries except 1 at the (i,j)th position. Proceeding in the same way as case I, we may write

$$\mathbb{P}\left(\frac{1}{\sqrt{c_n}}\|\tilde{X}\| > t\right) \le e^{-t}\mathrm{tr}\exp\left[\sum_{i=1}^{n}\sum_{j\in I_i}\log\mathbb{E}\exp\left(\frac{1}{\sqrt{c_n}}x_{ij}(E_{i,n+j}+E_{n+j,i})\right)\right]. \tag{3.24}$$

Let us consider the 2×2 matrix $H = \begin{bmatrix} 0 & \gamma \\ \gamma & 0 \end{bmatrix}$, where γ is a real valued random variable. We can decompose H as

$$H = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \gamma & 0 \\ 0 & -\gamma \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

As a result,

$$\exp(H) = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} e^{\gamma} & 0 \\ 0 & e^{-\gamma} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Therefore,

$$\begin{aligned}\log\mathbb{E}[\exp(H)] &= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \log\mathbb{E}e^{\gamma} & 0 \\ 0 & \log\mathbb{E}e^{-\gamma} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} \log[\mathbb{E}e^{\gamma}\mathbb{E}e^{-\gamma}] & \log[\mathbb{E}e^{\gamma}/\mathbb{E}e^{-\gamma}] \\ \log[\mathbb{E}e^{\gamma}/\mathbb{E}e^{-\gamma}] & \log[\mathbb{E}e^{\gamma}\mathbb{E}e^{-\gamma}] \end{bmatrix}.\end{aligned}$$

Since x_{ij}s are i.i.d., let us assume that all x_{ij} have the same probability distribution as a real valued random variable γ. Then proceeding as above, we can see that

$$\log\mathbb{E}\exp\left(\frac{1}{\sqrt{c_n}}x_{ij}(E_{i,n+j}+E_{n+j,i})\right) = \frac{1}{2}\log[\mathbb{E}e^{\gamma/\sqrt{c_n}}\mathbb{E}e^{-\gamma/\sqrt{c_n}}](E_{ii}+E_{n+j,n+j})$$

$$+\frac{1}{2}\log[\mathbb{E}e^{\gamma/\sqrt{c_n}}/\mathbb{E}e^{-\gamma/\sqrt{c_n}}](E_{i,n+j}+E_{n+j,i}).$$

Therefore,

$$\sum_{i=1}^{n}\sum_{j\in I_i}\log\mathbb{E}\exp\left(\frac{1}{\sqrt{c_n}}x_{ij}(E_{i,n+j}+E_{n+j,i})\right)=\frac{c_n}{2}\log[\mathbb{E}e^{\gamma/\sqrt{c_n}}\mathbb{E}e^{-\gamma/\sqrt{c_n}}]\ I$$
$$+\frac{1}{2}\log[\mathbb{E}e^{\gamma/\sqrt{c_n}}/\mathbb{E}e^{-\gamma/\sqrt{c_n}}]\sum_{i=1}^{n}\sum_{j\in I_i}(E_{i,j+n}+E_{j+n,i}).$$

From the Golden–Thompson inequality, if A and B are two $d\times d$ real symmetric matrices then $\mathrm{tr}e^{A+B}\le\mathrm{tr}(e^Ae^B)$.

In our case, let us take

$$A=\frac{c_n}{2}\log[\mathbb{E}e^{\gamma/\sqrt{c_n}}\mathbb{E}e^{-\gamma/\sqrt{c_n}}]\ I$$
$$B=\frac{1}{2}\log[\mathbb{E}e^{\gamma/\sqrt{c_n}}/\mathbb{E}e^{-\gamma/\sqrt{c_n}}]\sum_{i=1}^{n}\sum_{j\in I_i}(E_{i,j+n}+E_{j+n,i}).$$

Then

$$e^A=[\mathbb{E}e^{\gamma/\sqrt{c_n}}\mathbb{E}e^{-\gamma/\sqrt{c_n}}]^{c_n/2}\ I.$$

$$\begin{aligned}&\mathrm{tr}\exp\left[\sum_{i=1}^{n}\sum_{j\in I_i}\log\mathbb{E}\exp\left(\frac{1}{\sqrt{c_n}}x_{ij}(E_{i,n+j}+E_{n+j,i})\right)\right]\\&\le\mathrm{tr}\left[\left\{[\mathbb{E}e^{\gamma/\sqrt{c_n}}\mathbb{E}e^{-\gamma/\sqrt{c_n}}]^{nc_n/2}\right\}e^B\right]\\&\le\left\{[\mathbb{E}e^{\gamma/\sqrt{c_n}}\mathbb{E}e^{-\gamma/\sqrt{c_n}}]^{nc_n/2}\right\}\ ne^{\|B\|}.\end{aligned}$$

It is not difficult to see that $\left\|\sum_{i=1}^{n}\sum_{j\in I_i}(E_{i,j+n}+E_{j+n,i})\right\|\le c_n$. Combining all the estimates and plugging them in (3.24) we obtain

$$\begin{aligned}\mathbb{P}\left(\frac{1}{\sqrt{c_n}}\|\tilde{X}\|>t\right)&\le ne^{-t}[\mathbb{E}e^{\gamma/\sqrt{c_n}}\mathbb{E}e^{-\gamma/\sqrt{c_n}}]^{c_n/2}[\mathbb{E}e^{\gamma/\sqrt{c_n}}/\mathbb{E}e^{-\gamma/\sqrt{c_n}}]^{c_n/2}\\&=ne^{-t}\left\{\mathbb{E}e^{\gamma/\sqrt{c_n}}\right\}^{c_n}.\end{aligned}$$

From the concentration estimate (3.3) we have that $\mathbb{P}(|\gamma| > t) \leq \exp\{-t\sqrt{m}/\sqrt{2}\}$

$$\begin{aligned}
\mathbb{E}[e^{\gamma/\sqrt{c_n}}] &= \int_0^\infty \mathbb{P}\left(\frac{\gamma}{\sqrt{c_n}} > \log t\right)\, dt \\
&\leq \int_0^1 \mathbb{P}\left(\gamma > \sqrt{c_n}\log t\right)\, dt + \int_1^\infty \mathbb{P}\left(\gamma > \sqrt{c_n}\log t\right)\, dt \\
&\leq 1 + \int_1^\infty t^{-\sqrt{mc_n}/\sqrt{2}}\, dt \\
&= 1 + \left(\sqrt{\frac{mc_n}{2}} - 1\right)^{-1}.
\end{aligned}$$

As a result,

$$\mathbb{P}\left(\frac{1}{\sqrt{c_n}}\|\tilde{X}\| > t\right) \leq ne^{-t}e^{\sqrt{2c_n}/\sqrt{m}}.$$

Therefore,

$$\frac{1}{c_n}\mathbb{E}\|\tilde{X}\|^2 \leq (\log n + \sqrt{2c_n}/\sqrt{m})^2 \leq Kc_n.$$

□

PROPOSITION 3.6.3. *Let M be one of $C_j^{-1}, C_j^{-1}B_j^{-1}, C_j^{-1}r_jr_j^*C_j^{-1*}$ or $C_j^{-1}B_j^{-1}r_jr_j^*B^{-1*}C_j^{-1*}$, and x_j be the jth column of X. In addition, let us also assume that the random variables x_{ij} satisfy the Poincaré inequality with constant m, and $c_n > (\log n)^2$. Then we have*

$$\mathbb{E}\left|x_j^*Mx_j - \frac{c_n}{n}trM\right|^2 \leq Kc_n,$$

where $K > 0$ is a constant depends on $\Im(z)$, σ, and the Poincaré constant m. Moreover if the entries of of the matrix X are bounded by $6\sqrt{\frac{2}{m}}\log n$, then

$$\mathbb{E}\left|x_j^*Mx_j - \frac{c_n}{n}trM\right|^{2l} \leq Kc_n^l(\log n)^{2l},$$

$K > 0$ depends on l, $\Im(z)$, σ, and the Poincaré constant m.

PROOF. Let us first prove this for $M = C_j^{-1} = (YY^* - y_jy_j^* - zI)^{-1}$. Since x_{ij} satisfy the Poincaré inequality, they have exponential tails and consequently they have all moments. As a result we can repeat the same proof of Proposition 3.6.1. However, notice that in Proposition 3.6.1

we are getting the order n^l instead of c_n^l solely because of the estimate (3.21). So, it boils down to obtain an estimate of $O(c_n)$ for (3.21) when x_{ij} satisfy Poincaré inequality.

Since x_{ij} satisfy the Poincaré inequality we can write

$$\operatorname{Var}\left(\sum_{p\in I_j} M_{pp}\right) \le \frac{1}{\kappa}\sum_{s,t}\mathbb{E}\left|\sum_{p\in I_j}\frac{\partial M_{pp}}{\partial x_{st}}\right|^2 + \frac{1}{\kappa}\sum_{s,t}\mathbb{E}\left|\sum_{p\in I_j}\frac{\partial M_{pp}}{\partial \bar{x}_{st}}\right|^2,$$

where $\kappa > 0$ is the constant of Poincaré inequality. Let $m_{kl} := \sum_{i\neq j} y_{ki}\bar{y}_{li} = \frac{1}{c_n}\sum_{i\neq j}(r_{ki} + \sigma x_{ki})(\bar{r}_{li} + \sigma\bar{x}_{li})$ be the klth entry of $YY^* - y_j y_j$. It is very easy to compute, and done in the literature in past [**JSS16**, and references therein], that

$$\frac{\partial M_{pp}}{\partial m_{kl}} = -\frac{1}{1+\delta_{kl}}\left[M_{pk}M_{lp} + M_{pl}M_{kp}\right] = -\frac{2}{1+\delta_{kl}}M_{kp}M_{pl}.$$

Now it is easy to see that

$$\frac{\partial m_{kl}}{\partial \bar{x}_{st}} = \frac{\sigma}{c_n}\sum_{i\neq j}\delta_{ks}\delta_{it}(r_{li} + \sigma x_{li}) = \frac{\sigma}{c_n}\delta_{ks}(r_{lt} + \sigma x_{lt})\mathbf{1}_{\{t\neq j\}}.$$

Consequently,

$$\begin{aligned}
\sum_{p\in I_j}\frac{\partial M_{pp}}{\partial \bar{x}_{st}} &= -\frac{\sigma}{c_n}\sum_{p\in I_j}\sum_{k,l}\frac{2\delta_{ks}}{1+\delta_{kl}}M_{kp}M_{pl}[r_{lt} + \sigma x_{lt}]\mathbf{1}_{\{t\neq j\}} \\
&= -\frac{\sigma}{c_n}\sum_{p\in I_j}\sum_{l}\frac{2}{1+\delta_{sl}}M_{sp}M_{pl}[r_{lt} + \sigma x_{lt}]\mathbf{1}_{\{t\neq j\}} \\
&= -\frac{\sigma}{c_n}\sum_{l}(\tilde{M}_j)_{sl}[r_{lt} + \sigma x_{lt}]\mathbf{1}_{\{t\neq j\}} \\
&= -\frac{\sigma}{\sqrt{c_n}}(\tilde{M}_j Y_j)_{st},
\end{aligned}$$

where $(\tilde{M}_j)_{sl} = \frac{1}{1+\delta_{sl}}\sum_{p\in I_j} M_{sp}M_{pl}$, and Y_j is the matrix Y with jth column replaced by zeros.

Let us construct a matrix $(\hat{M}_j)_{n\times c_n}$ from M by removing all the columns except the ones indexed by I_j. For example, $\hat{M}_1$ is the matrix obtained from M by removing $(n - c_n)$ (i.e., $n - 2b_n - 1$) many columns of M indexed by $b_n + 2, b_n + 3, \ldots, n - (b_n + 1)$. Clearly $\tilde{M}_j = \hat{M}_j\hat{M}_j^T$ (the diagonals are divided by 2). Therefore, $\operatorname{rank}(\tilde{M}_j) \leq c_n$. As a result

$$\sum_{s,t} \mathbb{E}\left|\sum_{p\in I_j} \frac{\partial M_{pp}}{\partial \bar{x}_{st}}\right|^2 \leq \frac{\sigma^2}{c_n}\mathbb{E}\mathrm{tr}(\tilde{M}_j Y_j Y_j^* \tilde{M}_j^*) \leq \sigma^2 \mathbb{E}[\|\tilde{M}_j\|^2 \|Y_j Y_j^*\|] \leq \frac{\sigma^2}{|\Im(z)|^4}\mathbb{E}[\|Y_j Y_j^*\|], \tag{3.25}$$

where in the last inequality we have used the fact that $\|\hat{M}_j\| \leq 1/|\Im(z)|$. Consequently using the Lemma 3.6.2, we have

$$\sum_{s,t} \mathbb{E}\left|\sum_{p\in I_j} \frac{\partial M_{pp}}{\partial \bar{x}_{st}}\right|^2 \leq K c_n$$

Repeating the above calculations for $\sum_{s,t} \mathbb{E}\left|\sum_{p\in I_j} \frac{\partial M_{pp}}{\partial x_{st}}\right|^2$ we can obtain the same bounds. Hence the result follows for $M = C_j^{-1}$.

Since $\|B_j^{-1}\| \leq 1/|\Im(z)|$ and $\|r_j r_j^*\| \leq K c_n$, the result follows for $C_j^{-1}B_j^{-1}$, $C_j^{-1} r_j r_j^* C_j^{-1*}$, $C_j^{-1}B_j^{-1} r_j r_j^* B^{-1*} C_j^{-1*}$ too.

To prove the second part, we invoke the equation (3.3). Using (3.3), we have

$$\mathbb{P}\left(\left|\sum_{k\in I_j} M_{kk} - \mathbb{E}\sum_{k\in I_j} M_{kk}\right| > t\right) \leq 2K\exp\left(-\frac{\sqrt{m}}{\sqrt{2}\|\|\nabla \sum_{k\in I_j} M_{kk}\|_2\|_\infty} t\right).$$

From the equation (3.25), we have

$$\left\|\nabla \sum_{k\in I_j} M_{kk}\right\|_2^2 \leq \frac{2\sigma^2}{|\Im z|^4}\|Y_j Y_j^*\|.$$

Since all the entries of X are bounded by $6\sqrt{\frac{2}{m}}\log n$, we have $\|XX^*\| \leq K c_n^2 (\log n)^2$. And we know that $\|RR^*\| \leq K c_n$ for large n. Therefore, $\|YY^*\| \leq K c_n (\log n)^2$. We can get the same bound for $\|Y_j Y_j^*\|$. As a result,

$$\mathbb{P}\left(\left|\sum_{k\in I_j} M_{kk} - \mathbb{E}\sum_{k\in I_j} M_{kk}\right| > t\right) \leq 2K\exp\left(-\frac{\sqrt{m}}{K'\sqrt{2c_n}\log n} t\right).$$

Which implies that

$$\left|\sum_{k\in I_j} M_{kk} - \mathbb{E}\sum_{k\in I_j} M_{kk}\right|^{2l} \leq Kc_n^l(\log n)^{2l}.$$

Plugging this in (3.21), and following the same procedure as in Proposition 3.6.1, we have the result.

Observe that the second result of this Proposition is somewhat stronger than the first result, as it leads to the almost sure convergence (see Section 3.3) and it does not need the help of Lemma 3.6.2. However, the method used in Lemma 3.6.2 is interesting by itself. So we keep it. □

CHAPTER 4

Other Related Problems

In the Chapter 1, we have discussed about the limiting spectral distribution of Wigner random matrices and Hermitian random band matrices with i.i.d. entries. In both the cases, it turned out to be the Semicircle law. In this Chapter, we will briefly discuss about the limiting spectral distribution of random matrices with correlated entries.

4.1. Elliptic Law

4.1.1. Full Matrices. The Semicircle law holds for symmetric or Hermitian random matrices. However, if the ijth and jith entries of a random band matrix are correlated but not identically equal then the limiting spectral distribution of those matrices is no longer the Semicircle law. In 2014, Nguyen, and O'Rourke [**NO14**] proved the following result.

THEOREM 4.1.1. *Let $X = (x_{ij})_{n\times n}$ be a sequence of $n \times n$ random matrices such that*

(i) (x_{ij}, x_{ji}) are i.i.d. random vectors, and x_{ii} are i.i.d. random variables with mean zero and finite variance.

(ii) $\mathbb{E}[x_{12}] = 0 = \mathbb{E}[x_{21}]$, $\mathbb{E}[|x_{12}|^2] = 1 = \mathbb{E}[|x_{21}|^2]$,

(iii) $\mathbb{E}[\Re(x_{12})^2] = \mu = \mathbb{E}[\Re(x_{21})^2]$, $\mathbb{E}[\Im(x_{12})^2] = 1-\mu = \mathbb{E}[\Im(x_{21})^2]$,

(iv) $\mathbb{E}[\Re(x_{12})\Re(x_{21})] = \mu\rho$, $\mathbb{E}[\Im(x_{12})\Im(x_{21})] = -(1-\mu)\rho$,

(v) $\mathbb{E}[\Re(x_{12})\Im(x_{21})] = 0 = \mathbb{E}[\Im(x_{12})\Re(x_{21})]$.

Let $0 \le \mu \le 1$ and $-1 < \rho < 1$ be given, and F be a sequence of deterministic matrices with $rank(F) = o(n)$ and $\sup_n \frac{1}{n^2} tr(FF^*) < \infty$. Then the ESD of $\frac{1}{\sqrt{n}}(X+F)$ converges almost surely to the elliptic law with parameter ρ, which is given by the following pdf

$$\mathcal{E}_\rho(z) = \begin{cases} 1 & \text{if } \frac{\Re(z)^2}{(1+\rho)^2} + \frac{\Im(z)^2}{(1-\rho)^2} \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

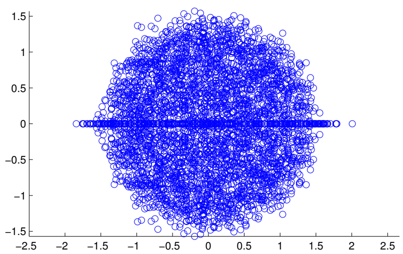
In 2012, Naumov [**Nau12**] proved the above result under finite fourth moment assumption and for $\mu = 1$ i.e., for the real valued random matrices. Also in 1985, Girko [**Gir86**] proved this result for real valued random matrices under the assumption that the entries have probability density functions.

4.1.2. Band Matrices. In case of random band matrices, the Elliptic law seems to be true. Let $X = (x_{ij})_{n\times n}$ be a random band matrix with bandwidth b_n. Let $\mathbb{E}[x_{ij}] = 0 = \mathbb{E}[x_{ji}]$, $\mathbb{E}[|x_{ij}|^2] = 1 = \mathbb{E}[|x_{ji}|^2]$, and $\mathbb{E}[x_{ij}x_{ji}] = \rho$ for all $i \neq j$ and $\min\{|i-j|, n-|i-j|\} \leq b_n$. In the figure 4.1, we have done some numerical simulations with random band matrices with different bandwidths and $(x_{ij}, x_{ji}) \overset{i.i.d}{\sim} \mathcal{N}_2(\mathbf{0}, \rho)$, where $\mathcal{N}_2$ denotes the standard bivariate Gaussian random variables with correlation coefficient ρ. We suspect that when the bandwidth $b_n >> \sqrt{n}$, the elliptic law is true for random band matrices. But when $b_n << \sqrt{n}$, there are significantly many real eigenvalues. However, the nonreal eigenvalues still follow the elliptic law. The figures shown in 4.1 are for periodic band matrices. Corresponding figures for non-periodic band matrices look the same and are not included here.

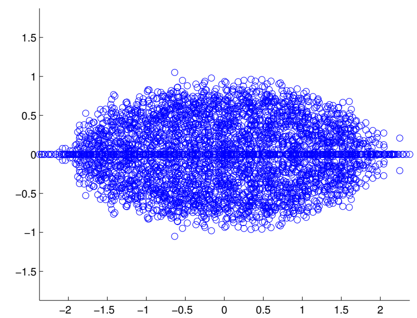
4.2. Localization-delocalization of the eigenvectors

It was proposed by Casati, Molinari, and Izrailev [**CMI90**] that the eigenvectors of a random band matrix of bandwidth b_n are localized if $b_n << \sqrt{n}$ and delocalized otherwise. Here localization indicates the the number of contributing components in an eigenvector. For example, the eigenvectors of a diagonal matrix is highly localized. Because the main contribution of such vectors come only from one component. On the other hand, if all components of a vector are the same then the vector is highly delocalized. It was proved by Schenker [**Sch09**] that the eigenvectors are localized if $b_n << n^{1/8}$, and Erdös et al. [**EKYY13**] proved that eigenvectors are delocalized when the bandwidth $b_n >> n^{4/5}$. Recently, Olver and Swan [**OS17**] have found that a second order phase transition of localization-delocalization occurs at $b_n \asymp \frac{2}{5}N$.

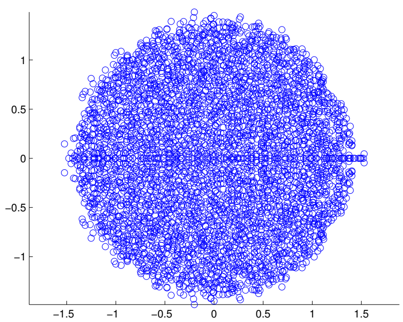
We have done some MATLAB simulations to see the behavior of the eigenvector of band matrices of several bandwidths. we took a $n \times n$ band matrix of different bandwidths. For example, let us say the band width is $n^{0.3}$. Then we looked at the normalized eigen vectors of that matrix. For each eigenvector we counted the number of components whose absolute values are more than



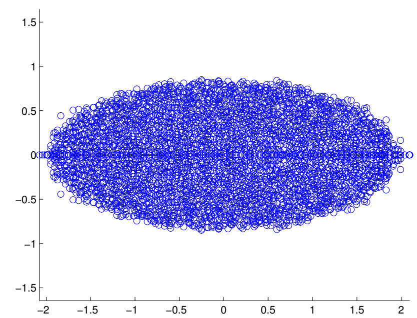
(a) Bandwidth $(b_n) = n^{0.3}$, Covariance $(\rho) = 0$.



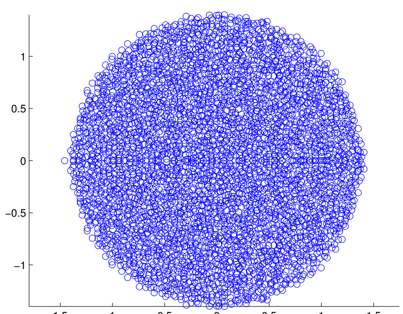
(b) Bandwidth $(b_n) = n^{0.3}$, Covariance $(\rho) = 0.4$.



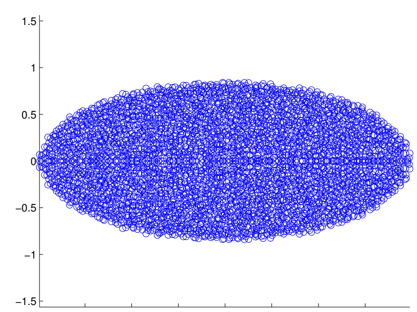
(c) Bandwidth $(b_n) = n^{0.5}$, Covariance $(\rho) = 0$.



(d) Bandwidth $(b_n) = n^{0.5}$, Covariance $(\rho) = 0.4$.



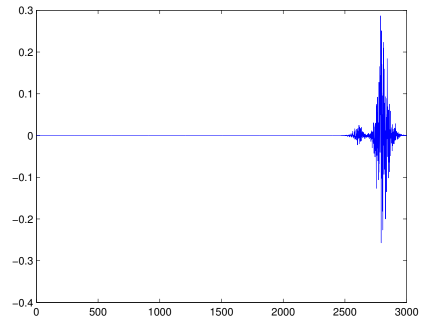
(e) Bandwidth $(b_n) = n^{0.7}$, Covariance $(\rho) = 0$.



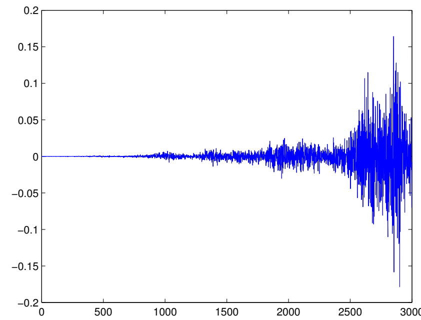
(f) Bandwidth $(b_n) = n^{0.7}$, Covariance $(\rho) = 0.4$.

FIGURE 4.1. ESDs of random band matrices with size 4000×4000.

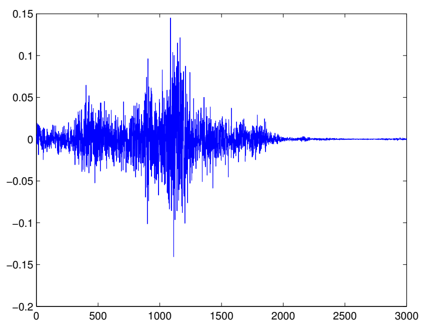
$\frac{1}{n^{error}}$. If less than $\frac{n}{2}$ components are bigger than $\frac{1}{n^{error}}$, then we call that eigenvector is localized. Figure 4.2 shows some such simulations.



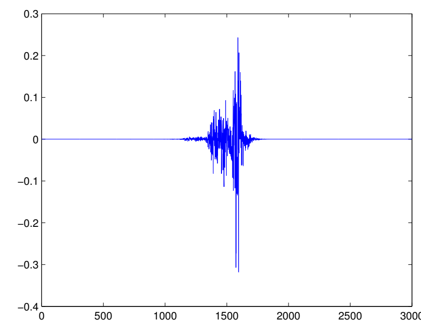
(a) 6th eigenvector; localized



(b) 1200th eigenvector; delocalized



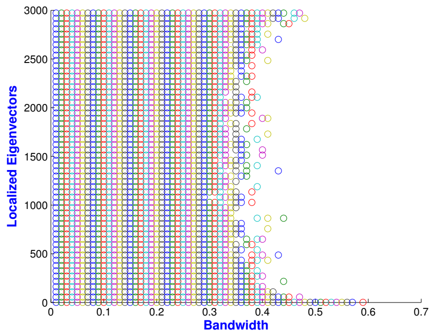
(c) 1800th eigenvector; delocalized



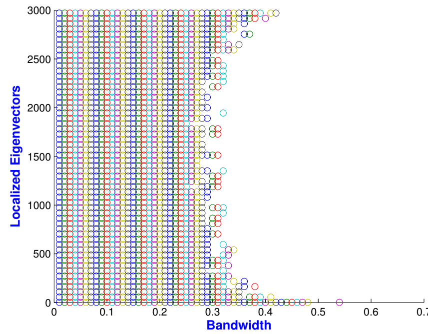
(d) 2970th eigenvector; localized

FIGURE 4.2. 3000×3000 band matrix of bandwidth $3000^{0.4}$.

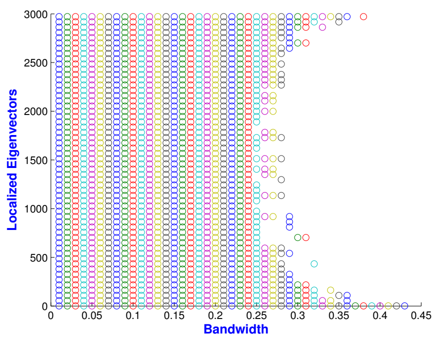
The simulations show even some more interesting facts. It seems that the eigenvectors corresponding to the extreme (smallest or largest) eigenvalues tend to be more localized than the eigenvectors corresponding to the bulk eigenvalues. For example, we looked at the eigenvectors of a 3000×3000 random band matrix with varying bandwidth and with i.i.d. standard Gaussian entries. Let M be a $n \times n$ random band matrix with band width n^d and with i.i.d. standard Gaussian entries. It has n many eigenvectors. I mark a vector by a colored dot if it is found to be localized with respect to the error threshold $1/n^{error}$. In the figure 4.3, d is plotted along the x-axis and the localized eigenvectors corresponding to the random band matrix of bandwidth n^d is plotted along the y-axis. The error threshold in the following figures are respectively $\frac{1}{n}$, $\frac{1}{n^2}$, $\frac{1}{n^3}$ and $\frac{1}{n^4}$. It is clear that all the eigenvectors are localized up to certain bandwidth (around $n^{0.35}$ in the first figure). Then the eigenvectors corresponding to the bulk eigenvalues start delocalizing.



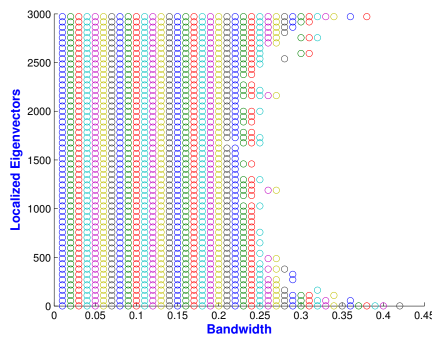
(a) Error threshold=$\frac{1}{n}$



(b) Error threshold=$\frac{1}{n^2}$



(c) Error threshold=$\frac{1}{n^3}$



(d) Error threshold=$\frac{1}{n^4}$

FIGURE 4.3. The colored dots indicate the localized eigenvectors.

APPENDIX A

Auxiliary Results

PROOF OF PROPOSITION 2.2.2: Let us denote the averaging with respect to $\{w_{ij}; 1 \le i \le k \text{ or } 1 \le j \le n\}$ by $\mathbb{E}_{\le k}$ and the averaging with respect to $\{w_{kj}; 1 \le j \le n\}$ by $\mathbb{E}_k$. Using the martingale difference technique (see [**DFJ68**]), we have

$$\begin{aligned}\mathrm{Var}\{\gamma_n\} &\le \sum_{k=1}^{n} \mathbb{E}\left[|\mathbb{E}_{\le k-1}[\gamma_n] - \mathbb{E}_{\le k}[\gamma_n]|^2\right] \\ &= \sum_{k=1}^{n} \mathbb{E}\left[|\mathbb{E}_{\le k-1}\left[\gamma_n - \mathbb{E}_{\le k}[\gamma_n]\right]|^2\right] \\ &\le \sum_{k=1}^{n} \mathbb{E}\left[\mathbb{E}_{\le k-1}\left|\gamma_n - \mathbb{E}_{k}[\gamma_n]\right|^2\right] \\ &= \sum_{k=1}^{n} \mathbb{E}\left[|\gamma_n - \mathbb{E}_{k}[\gamma_n]|^2\right].\end{aligned}$$

Note that

$$\begin{aligned}\mathbb{E}\left[|\gamma_n - \mathbb{E}_1[\gamma_n]|^2\right] &= \mathbb{E}\left[|\mathrm{Tr}(G) - \mathbb{E}_1[\mathrm{Tr}(G)]|^2\right] \\ &= \mathbb{E}\left[\left|\mathrm{Tr}(G) - \mathbb{E}_1[\mathrm{Tr}(G)] + \mathrm{Tr}(G^{(1)}) - \mathrm{Tr}(G^{(1)})\right|^2\right] \\ &= \mathbb{E}\left[\left|\mathrm{Tr}(G - G^{(1)}) - \mathbb{E}_1\left[\mathrm{Tr}(G - G^{(1)})\right]\right|^2\right].\end{aligned}$$

From (A.3) we have

$$\mathrm{Tr}(G - G^{(1)}) = -\frac{1 + B(z)}{A(z)} \tag{A.1}$$

where $A(z) = -G_{11}^{-1}$, $B(z) = \left\langle G^{(1)} G^{(1)} m^{(1)}, m^{(1)} \right\rangle$, and $G^{(1)}$ is defined in (2.16), and $m^{(1)} = \frac{1}{\sqrt{b_n}}(w_{12}, w_{13}, \ldots, w_{1n})^T$. Indeed,

$$\mathbb{E}\left[|\gamma_n - \mathbb{E}_1[\gamma_n]|^2\right] \le \mathbb{E}\left[\left|\frac{1+B(z)}{A(z)} - \mathbb{E}_1\left[\frac{1+B(z)}{A(z)}\right]\right|^2\right]$$
$$\le 2\mathbb{E}\left[\left|\frac{1}{A(z)} - \mathbb{E}_1\left[\frac{1}{A(z)}\right]\right|^2\right] + 2\mathbb{E}\left[\left|\frac{B(z)}{A(z)} - \mathbb{E}_1\left[\frac{B(z)}{A(z)}\right]\right|^2\right].$$

Now, by (2.21) and (2.25),

$$\mathbb{E}_1\left[\left|\frac{B(z)}{A(z)} - \mathbb{E}_1\left[\frac{B(z)}{A(z)}\right]\right|^2\right] \le \mathbb{E}_1\left[\left|\frac{B(z)}{A(z)} - \frac{\mathbb{E}_1[B(z)]}{E_1[A(z)]}\right|^2\right]$$
$$\le \mathbb{E}_1\left[\left|\frac{B_1^\circ}{\mathbb{E}_1[A]} - \frac{A_1^\circ}{\mathbb{E}_1[A]}\frac{B}{A}\right|^2\right]$$
$$\le 2\mathbb{E}_1\left[\left|\frac{B_1^\circ}{\mathbb{E}_1[A]}\right|^2\right] + \frac{2}{|\Im z|^2}\mathbb{E}_1\left[\left|\frac{A_1^\circ}{\mathbb{E}_1[A]}\right|^2\right],$$

where $A_1^\circ = A - \mathbb{E}_1[A]$. So it is enough to estimate $\mathbb{E}_1\left[\left|\frac{A_1^\circ}{\mathbb{E}_1[A]}\right|^2\right]$ and $\mathbb{E}_1\left[\left|\frac{B_1^\circ}{\mathbb{E}_1[A]}\right|^2\right]$. Note that

$$A = z - \frac{1}{\sqrt{b_n}} w_{11} + \left\langle G^{(1)} m^{(1)}, m^{(1)} \right\rangle$$
$$A_1^\circ = -\frac{1}{\sqrt{b_n}} w_{11} + \frac{1}{b_n}\sum_{\substack{i \ne j \\ i,j \in I_1}} G_{ij}^{(1)} w_{1i} w_{1j} + \frac{1}{b_n}\sum_{i \in I_1} G_{ii}^{(1)} (w_{1i}^2)^\circ.$$

Therefore,

$$\mathbb{E}_1\left[|A_1^\circ|^2\right] = \mathbb{E}_1\left[\frac{1}{b_n} w_{11}^2 + \frac{1}{b_n^2}\sum_{\substack{i \ne j \\ i,j \in I_1}} G_{ij}^{(1)} w_{1i} w_{1j} \sum_{\substack{k \ne l \\ k,l \in I_1}} \overline{G_{kl}^{(1)}} w_{1k} w_{1l} + \frac{1}{b_n^2}\sum_{i \in I_1} G_{ii}^{(1)} (w_{1i}^2)^\circ \sum_{l \in I_1} \overline{G_{ll}^{(1)}} (w_{1l}^2)^\circ\right]$$
$$= \frac{\sigma^2}{b_n} + \frac{2}{b_n^2}\sum_{\substack{i \ne j \\ i,j \in I_1}} |G_{ij}^{(1)}|^2 + \frac{\mu_4 - 1}{b_n^2}\sum_{i \in I_1} |G_{ii}^{(1)}|^2$$
$$\le \frac{\sigma^2}{b_n} + \frac{2}{b_n^2}\frac{2b_n}{|\Im z|^2} + \frac{\mu_4 - 1}{b_n^2}\frac{2b_n}{|\Im z|^2}$$
$$\le \frac{1}{b_n}\left(\sigma^2 + \frac{2 + 2\mu_4}{|\Im z|^2}\right). \tag{A.2}$$

Now, we want to estimate $\mathbb{E}_1\left[|B_1^\circ|^2\right]$, where $B = \langle G^{(1)}G^{(1)}m^{(1)}, m^{(1)}\rangle = \langle H^{(1)}m^{(1)}, m^{(1)}\rangle$, and $B_1^\circ = B - \mathbb{E}_1[B]$. Therefore,

$$\mathbb{E}_1[B] = \frac{1}{b_n}\sum_{i\in I_1} H_{ii}^{(1)} = \frac{1}{b_n}\sum_{i\in I_1}\sum_{j=2}^{n}\left(G_{ij}^{(1)}\right)^2,$$

and

$$B_1^\circ = \frac{1}{b_n}\sum_{\substack{i\neq j\\ i,j\in I_1}} H_{ij}^{(1)} w_{1i}w_{1j} + \frac{1}{b_n}\sum_{i\in I_1} H_{ii}^{(1)}\left(w_{1i}^2\right)^\circ.$$

Let us call $C_0 = \mathbb{E}\left[(w_{1i}^2)^\circ\right]^2$. Then

$$\begin{aligned}
\mathbb{E}_1[|B_1^\circ|^2] &= \frac{1}{b_n^2}\sum_{\substack{i\neq j\\ i,j\in I_1}} |H_{ij}^{(1)}|^2 + \frac{C_0}{b_n^2}\sum_{i\in I_1}|H_{ii}^{(1)}|^2 \\
&= \frac{1}{b_n^2}\sum_{\substack{i\neq j\\ i,j\in I_1}}\left|\sum_{k=2}^{n} G_{ik}^{(1)}G_{kj}^{(1)}\right|^2 + \frac{C_0}{b_n^2}\sum_{i\in I_1}\left|\sum_{k=2}^{n} G_{ik}^{(1)}G_{ki}^{(1)}\right|^2 \\
&\leq \frac{1}{b_n^2}\sum_{i\in I_1}\|(G^{(1)})^2\|^2 + \frac{C_0}{b_n^2}2b_n\|(G^{(1)})^2\|^2 \\
&= \frac{2}{b_n}\frac{1}{|\Im z|^4} + \frac{2C_0}{b_n}\frac{1}{|\Im z|^4} \\
&= \frac{2(1+C_0)}{b_n|\Im z|^4}.
\end{aligned}$$

We also have

$$\mathbb{E}\left[\left|\frac{1}{A(z)} - \mathbb{E}_1\left[\frac{1}{A(z)}\right]\right|^2\right] \leq \frac{1}{|\Im z|^2}\mathbb{E}\left[\left|\frac{A_1^\circ}{\mathbb{E}_1[A]}\right|^2\right]$$

Note that $\mathbb{E}_1[A] = z + \frac{1}{b_n}\sum_{i\in I_1} G_{ii}^{(1)}$. Since $\Im G_{ii}^{(1)} > 0$, we have $|\mathbb{E}_1[A]| \geq |\Im A| \geq y$. Also, we know that $|G_{ii}^{(1)}| \leq 1/|\Im z| = 1/y$. Therefore $|\mathbb{E}_1[A]| \geq |x| - \frac{2}{y}$. Combining these we have

$$|\mathbb{E}_1[A]| > \max\left\{y, |x| - \frac{2}{y}\right\}.$$

Therefore,

$$\mathbb{E}\left[|\gamma_n - \mathbb{E}_1[\gamma_n]|^2\right] \le C_1 \frac{2(1+C_0)}{b_n |\Im z|^4} |\mathbb{E}_1[A]|^{-2} + \frac{C_2}{b_n}\left(\sigma^2 + \frac{2+2\mu_4}{|\Im z|^2}\right)\frac{|\mathbb{E}_1[A]|^{-2}}{|\Im z|^2}$$
$$\le \frac{C}{b_n}\left(\frac{1}{|\Im z|^2} + \frac{1}{|\Im z|^4}\right)|\mathbb{E}_1[A]|^{-2},$$

for some $C_1, C_2, C > 0$ not depending on z, n. This implies

$$\mathrm{Var}(\gamma_n) \le \frac{Cn}{b_n}\left(\frac{1}{|\Im z|^2} + \frac{1}{|\Im z|^4}\right)\left(\max\left\{y, |x| - \frac{2}{y}\right\}\right)^{-2}.$$

This completes the proof of proposition 2.2.2. □

Now, we proceed to the proofs of the asymptotic estimates. All the asymptotic estimates listed in Lemma A.0.1 and Lemma A.0.2 hold uniformly in the set $\{z \in \mathbb{C} : |\Im z| \ge \eta\}$ for any given $\eta > 0$.

Lemma A.0.1. *Let M be an $n \times n$ symmetric band matrix as defined in (2.2) which satisfies (2.3) and $\mathbb{E}[|w_{ij}|^8]$ is uniformly bounded. Then*

(i)

(A.3)
$$G_{ii}^{(1)} - G_{ii} = \frac{1}{A(z)}\left(G^{(1)} m^{(1)}\right)_i^2 = \frac{1}{A(z)}\left(\frac{1}{\sqrt{b_n}}\sum_{j\in I_1} G_{ij}^{(1)} w_{1j}\right)^2$$

where $2 \le i \le n$, $A(z)$, $m^{(1)}$ and $G^{(1)}$ are as defined in (2.14), (2.15) and (2.16).

(ii) $\left|\mathbb{E}\left[G_{ii}^{(1)}(z)\right] - \mathbb{E}[G_{ii}(z)]\right| = O\left(\frac{1}{b_n}\right)$.

(iii)

(A.4)
$$\mathbb{E}[|G_{12}|^2] = O\left(\frac{1}{b_n}\right)\frac{1}{|\Im z|^6},\ \mathbb{E}[|G_{12}|^4] = O\left(\frac{1}{b_n^2}\right)\frac{1}{|\Im z|^{12}} \text{ and } \mathbb{E}[|G_{12}|^8] = O\left(\frac{1}{b_n^4}\right)\frac{1}{|\Im z|^{24}}.$$

(iv) Let us denote the averaging with respect to $\{w_{1i}\}_{1\le i\le n}$ by $\mathbb{E}_1$. Then

(A.5)
$$b_n \mathbb{E}_1\left[A^\circ(z_1)A^\circ(z_2)\right] = \sigma^2 + \frac{2}{b_n}\sum_{i,j\in I_1} G_{ij}^{(1)}(z_1)G_{ij}^{(1)}(z_2) + \frac{\kappa_4}{b_n}\sum_{i\in I_1} G_{ii}^{(1)}(z_1)G_{ii}^{(1)}(z_2) + \frac{1}{b_n}\widetilde{\gamma_{n-1}}(z_1)\widetilde{\gamma_{n-1}}(z_2)$$

where $\widetilde{\gamma_{n-1}}(z) = \sum_{i\in I_1}\left(G_{ii}^{(1)} - \mathbb{E}[G_{ii}^{(1)}(z)]\right)$ and $I_1 = \{1 < i \le n : (1,i) \in I_n\}$.

(v)

$$\mathbb{E}_1\left[A^\circ(z_1)B^\circ(z_2)\right] = \frac{d}{dz_2}\mathbb{E}_1\left[A^\circ(z_1)A^\circ(z_2)\right] \text{ where } B(z_2) = \left\langle G^{(1)}(z_2)G^{(1)}(z_2)m^{(1)}, m^{(1)}\right\rangle. \tag{A.6}$$

LEMMA A.0.2. *Let M be an $n \times n$ symmetric band matrix as defined in* (2.2) *which satisfies* (2.3)*. Also assume that the probability distribution of w_{jk} satisfies the Poincaré inequality with some uniform constant m which does not depend on n, j, k. Then*

(i)

$$Var\Big(\sum_{(1,i)\in I_n} G_{ii}\Big) = O(1) \quad \text{and} \quad Var(G_{11}(z)) = O\left(\frac{1}{b_n}\right). \tag{A.7}$$

(ii)

$$\mathbb{E}\left[|A^\circ|^4\right] = O\left(\frac{1}{b_n^2}\right), \qquad \mathbb{E}\left[|A^\circ|^3\right] = O\left(\frac{1}{b_n^{3/2}}\right) \tag{A.8}$$

$$\mathbb{E}\left[|B^\circ|^4\right] = O\left(\frac{1}{b_n^2}\right) \tag{A.9}$$

(iii)

$$Var\left\{b_n\mathbb{E}_1\left[A^\circ(z_1)A^\circ(z_2)\right]\right\} = O\left(\frac{1}{b_n}\right) \text{ and } Var\left\{b_n\mathbb{E}_1\left[A^\circ(z_1)B^\circ(z_2)\right]\right\} = O\left(\frac{1}{b_n}\right) \tag{A.10}$$

(iv)

$$\mathbb{E}\left[\left|\gamma_{n-1}^\circ(z) - \gamma_n^\circ(z)\right|^4\right] = O\left(\frac{1}{b_n^2}\right) \quad \text{and} \quad \mathbb{E}\left[|\gamma_n^\circ|^4\right] = O\left(\frac{n^2}{b_n^2}\right). \tag{A.11}$$

(v)

$$\frac{1}{n}\mathbb{E}\left[TrG(z)\right] = f(z) + O\left(\frac{1}{|\Im z|^6 b_n}\right) \text{ where } f(z) = \frac{1}{4}\left(-z + \sqrt{z^2-8}\right). \tag{A.12}$$

(vi)

$$(\mathbb{E}[A(z)])^{-1} = -f(z) + O(b_n^{-1}) \quad \text{and} \quad \mathbb{E}[B(z)] = 2f'(z) + O(b_n^{-1}). \tag{A.13}$$

PROOF OF LEMMA A.0.1: **Proof of (i):** Suppose $(X_1, X_2, \ldots, X_n)$ is a n dimensional normal random vector with a positive definite covariance matrix A^{-1} and a mean $A^{-1}\underline{h}$, where $\underline{h} \in \mathbb{R}^n$.

Then we have

$$\int \exp\left[-\frac{1}{2}\langle A\underline{x},\underline{x}\rangle + \langle \underline{h},\underline{x}\rangle\right]\ d\underline{x} = (2\pi)^{n/2}|\det A|^{-1/2}\exp\left[\frac{1}{2}\langle A^{-1}\underline{h},\underline{h}\rangle\right], \tag{A.14}$$

$$\frac{\int x_i x_j \exp\left[-\frac{1}{2}\langle A\underline{x},\underline{x}\rangle + \langle \underline{h},\underline{x}\rangle\right]\ d\underline{x}}{\int \exp\left[-\frac{1}{2}\langle A\underline{x},\underline{x}\rangle + \langle \underline{h},\underline{x}\rangle\right]\ d\underline{x}} = (A^{-1})_{ij} + (A^{-1}\underline{h})_i(A^{-1}\underline{h})_j. \tag{A.15}$$

where $\underline{x} = (x_1, x_2, \ldots, x_n)^T$. In particular, for $\underline{h} = 0$,

$$(A^{-1})_{ij} = \frac{\int x_i x_j \exp[-\frac{1}{2}\langle A\underline{x},\underline{x}\rangle]\ d\underline{x}}{\int \exp[-\frac{1}{2}\langle A\underline{x},\underline{x}\rangle]\ d\underline{x}}. \tag{A.16}$$

Now, doing the integrations in (A.16) with respect to all variables except x_1, and using (A.14) we get

$$\begin{aligned}\int \exp[-\frac{1}{2}\langle A\underline{x},\underline{x}\rangle]\ d\underline{x} &= \int \exp[-\frac{a_{11}x_1^2}{2}] \int \exp[-\frac{1}{2}\langle A_1\underline{x}^{(1)},\underline{x}^{(1)}\rangle - \langle x_1\underline{a_1},\underline{x}^{(1)}\rangle]\ d\underline{x} \\ &= \frac{(2\pi)^{\frac{n-1}{2}}}{|\det A_1|^{1/2}} \int \exp[-\frac{x_1^2}{2}(a_{11} - \langle A_1^{-1}\underline{a_1},\underline{a_1}\rangle)]\ dx_1\end{aligned}$$

where $\underline{x}^{(1)} = (x_2, x_3, \ldots, x_n)^T$, $\underline{a_1} = (a_{12}, a_{13}, \ldots, a_{1n})^T$ and $A_1 = ((A_1)_{ij})_{i,j=2}^n$ is the $(n-1)\times(n-1)$ matrix obtained from A after removing first row and first column, and for $i, j \neq 1$, using (A.15) and (A.14) we get

$$\begin{aligned}&\int x_i x_j \exp[-\frac{1}{2}\langle A\underline{x},\underline{x}\rangle]\ d\underline{x} \\ &= \int \exp[-\frac{a_{11}x_1^2}{2}] \int x_i x_j \exp[-\frac{1}{2}\langle A_1\underline{x}^{(1)},\underline{x}^{(1)}\rangle - \langle x_1\underline{a_1},\underline{x}^{(1)}\rangle]\ d\underline{x}^{(1)}dx_1 \\ &= \int \exp[-\frac{a_{11}x_1^2}{2}]\ [(A_1^{-1})_{ij} + x_1^2(A_1^{-1}\underline{a_1})_i(A_1^{-1}\underline{a_1})_j] \int \exp[-\frac{1}{2}\langle A_1\underline{x}^{(1)},\underline{x}^{(1)}\rangle - \langle x_1\underline{a_1},\underline{x}^{(1)}\rangle]\ d\underline{x}^{(1)}dx_1 \\ &= \frac{(2\pi)^{\frac{n-1}{2}}}{|\det A_1|^{1/2}} \int [(A_1^{-1})_{ij} + x_1^2(A_1^{-1}\underline{a_1})_i(A_1^{-1}\underline{a_1})_j]\ \exp[-\frac{x_1^2}{2}(a_{11} - \langle A_1^{-1}\underline{a_1},\underline{a_1}\rangle)]\ dx_1.\end{aligned}$$

Therefore, from (A.16) we get

$$\begin{aligned}(A^{-1})_{ij} &= (A_1^{-1})_{ij} + (A_1^{-1}\underline{a_1})_i(A_1^{-1}\underline{a_1})_j\ \frac{\int x_1^2 \exp[-\frac{x_1^2}{2}(a_{11} - \langle A_1^{-1}\underline{a_1},\underline{a_1}\rangle)]\ dx_1}{\int \exp[-\frac{x_1^2}{2}(a_{11} - \langle A_1^{-1}\underline{a_1},\underline{a_1}\rangle)]\ dx_1} \\ &= (A_1^{-1})_{ij} + \frac{(A_1^{-1}\underline{a_1})_i(A_1^{-1}\underline{a_1})_j}{a_{11} - \langle A_1^{-1}\underline{a_1},\underline{a_1}\rangle}.\end{aligned}$$

Applying the above formula for $A = (M - zI)$, where $z \in \mathbb{R}$, $|z| > \|M\|$, we obtain

$$G_{ij} = G_{ij}^{(1)} + \frac{(G^{(1)}m^{(1)})_i (G^{(1)}m^{(1)})_j}{\frac{w_{11}}{\sqrt{b_n}} - z - \langle G^{(1)}m^{(1)}, m^{(1)} \rangle}, \quad i, j \geq 2,$$

where $m^{(1)}$, $G^{(1)}$ are as defined in (2.15), (2.16) respectively. From the above formula we obtain

$$G_{ii} - G_{ii}^{(1)} = -\frac{(G^{(1)}m^{(1)})_i^2}{A(z)}, \quad \text{for all } \ 2 \leq i \leq n,$$

where $A(z)$ is as defined in (2.14). The above is true for all $z \in \mathbb{R}$ such that $|z| > \|M\|$. By analytic continuity one can extend it to the whole complex plane. This completes the proof.

Proof of (ii): Recall $I_1 = \{1 < i \leq n : (1, i) \in I_n\}$. Now, using (A.3) and (2.21) we have

$$\begin{aligned}
\left| \mathbb{E}\left[G_{ii}^{(1)}(z) \right] - \mathbb{E}[G_{ii}(z)] \right| &= \left| \mathbb{E}\left[\frac{1}{A} \left(G^{(1)}m^{(1)} \right)_i^2 \right] \right| \\
&= \left| \mathbb{E}\left[\frac{1}{A} \left(\frac{1}{\sqrt{b_n}} \sum_{j \in I_1} G_{ij}^{(1)} w_{1j} \right)^2 \right] \right| \\
&\leq \frac{1}{b_n} \frac{1}{|\Im z|} \mathbb{E}\left[\left| \sum_{j \in I_1} G_{ij}^{(1)} w_{1j} \right|^2 \right] \\
&\leq \frac{1}{b_n |\Im z|} \mathbb{E}\left[\sum_{j \in I_1} |G_{ij}^{(1)}|^2 w_{1j}^2 + \sum_{j_1 \neq j_2 \in I_1} G_{ij_1}^{(1)} \overline{G_{ij_2}^{(1)}} w_{1j_1} w_{1j_2} \right] \\
&= \frac{1}{b_n |\Im z|} \mathbb{E}\mathbb{E}_1 \left[\sum_{j \in I_1} |G_{ij}^{(1)}|^2 w_{1j}^2 + \sum_{j_1 \neq j_2 \in I_1} G_{ij_1}^{(1)} \overline{G_{ij_2}^{(1)}} w_{1j_1} w_{1j_2} \right] \\
&= \frac{1}{b_n |\Im z|} \mathbb{E}\left[\sum_{j \in I_1} |G_{1j}^{(1)}|^2 \right] \\
&\leq \frac{1}{b_n |\Im z|} \mathbb{E}\|G^{(1)}\|^2 \leq \frac{1}{b_n |\Im z|^3}.
\end{aligned}$$

Proof of (iii): Using the resolvent formula given in [**Erd11**], we have

$$G_{12} = -G_{22} G_{11}^{(2)} K_{12}^{(12)},$$

where $G^{(2)}$ is the resolvent of the $(n-1)\times(n-1)$ minor obtained by removing the kth row and kth column from the matrix M, $K_{12}^{(12)} = m_{12} - m_{(1)}G^{(12)}m_{(2)}$, $m_{(1)} = \frac{1}{\sqrt{b_n}}(w_{13}, w_{14}, \ldots, w_{1n})$, $m_{(2)} = \frac{1}{\sqrt{b_n}}(w_{23}, w_{24}, \ldots, w_{2n})^T$, $G^{(ij)} = \left(M^{(ij)} - zI\right)^{-1}$, and $M^{(ij)}$ is $(n-2)\times(n-2)$ matrix obtained from M after removing ith and jth rows and columns. Therefore,

$$\begin{aligned}
\mathbb{E}[|G_{12}|^2] &= \mathbb{E}\left[\left|G_{22}G_{11}^{(2)}K_{12}^{(12)}\right|^2\right] \\
&\leq \frac{1}{|\Im z|^2}\frac{1}{|\Im z|^2}\mathbb{E}\left[\left|m_{12} - m_{(1)}G^{(12)}m_{(2)}\right|^2\right] \\
&= \frac{1}{|\Im z|^4}\mathbb{E}\left[\left|\frac{w_{12}}{\sqrt{b_n}} - \frac{1}{b_n}\sum_{\substack{(1,i),(2,j)\in I_n \\ i,j\neq 1,2}} G_{ij}^{(12)} w_{1i}w_{2j}\right|^2\right] \\
&\leq \frac{1}{|\Im z|^4}\mathbb{E}\mathbb{E}_{\leq 2}\left[\frac{w_{12}^2}{b_n} + \frac{1}{b_n^2}\sum_{\substack{(1,i),(2,j)\in I_n \\ i,j\neq 1,2}} |G_{ij}^{(12)}|^2 w_{1i}^2 w_{2j}^2\right] \\
&\leq \frac{1}{|\Im z|^4}\mathbb{E}\left[\frac{1}{b_n} + \frac{1}{b_n^2}\sum_{i,j}|G_{ij}^{(12)}|^2\mathbb{E}_{\leq 2}[w_{1i}^2]\mathbb{E}_{\leq 2}[w_{2j}^2]\right] \\
&\leq \frac{1}{|\Im z|^4}\mathbb{E}\left[\frac{1}{b_n} + \frac{1}{b_n^2}\frac{b_n}{|\Im z|^2}\right] = O\left(\frac{1}{b_n}\right)\frac{1}{|\Im z|^6},
\end{aligned}$$

where $E_{\leq 2}$ is the averaging with respect to the first two rows and columns. Similarly, we can prove that $\mathbb{E}[|G_{12}|^4] = O\left(\frac{1}{b_n^2}\right)\frac{1}{|\Im z|^{12}}$, and $\mathbb{E}[|G_{12}|^8] = O\left(\frac{1}{b_n^4}\right)\frac{1}{|\Im z|^{24}}$.

Proof of (iv): We know that

$$A(z_1) = z_1 - \frac{w_{11}}{\sqrt{b_n}} + \left\langle G^{(1)}m^{(1)}, m^{(1)}\right\rangle,$$

$$\text{and}\quad A^\circ(z_1) = -\frac{w_{11}}{\sqrt{b_n}} + \frac{1}{b_n}\sum_{i\neq j\in I_1} G_{ij}^{(1)}w_{1i}w_{1j} + \frac{1}{b_n}\sum_{i\in I_1} G_{ii}^{(1)}w_{1i}^2 - \frac{1}{b_n}\sum_{i\in I_1}\mathbb{E}[G_{ii}^{(1)}].$$

Now, we can estimate

$b_n\mathbb{E}_1\left[A^\circ(z_1)A^\circ(z_2)\right]$

$$
= \sigma^2 + \frac{1}{b_n}\mathbb{E}_1\left[\sum_{\substack{i_1\neq j_1\in I_1\\ i_2\neq j_2\in I_1}} G^{(1)}_{i_1j_1}(z_1)w_{1i_1}w_{1j_1}G^{(1)}_{i_2j_2}(z_2)w_{1i_2}w_{1j_2}\right] + \frac{1}{b_n}\mathbb{E}_1\left[\sum_{i,j\in I_1} G^{(1)}_{ii}(z_1)G^{(1)}_{jj}(z_2)w_{1i}^2w_{1j}^2\right]
$$

$$
- \frac{1}{b_n}\mathbb{E}\left[\sum_{i\in I_1} G^{(1)}_{ii}(z_2)\right]\mathbb{E}_1\left[\sum_{i\in I_1} G^{(1)}_{ii}(z_1)w_{1i}^2\right] - \frac{1}{b_n}\mathbb{E}\left[\sum_{i\in I_1} G^{(1)}_{ii}(z_1)\right]\mathbb{E}_1\left[\sum_{i\in I_1} G^{(1)}_{ii}(z_2)w_{1i}^2\right]
$$

$$
+ \frac{1}{b_n}\mathbb{E}\left[\sum_{i\in I_1} G^{(1)}_{ii}(z_1)\right]\mathbb{E}\left[\sum_{i\in I_1} G^{(1)}_{ii}(z_2)\right]
$$

$$
= \sigma^2 + \frac{2}{b_n}\sum_{i\neq j\in I_1} G^{(1)}_{ij}(z_1)G^{(1)}_{ij}(z_2) + \frac{1}{b_n}\sum_{i\neq j\in I_1} G^{(1)}_{ii}(z_1)G^{(1)}_{jj}(z_2) + \frac{\mu_4}{b_n}\sum_{i\in I_1} G^{(1)}_{ii}(z_1)G^{(1)}_{ii}(z_2)
$$

$$
+ \frac{1}{b_n}\widetilde{\gamma_{n-1}}(z_1)\widetilde{\gamma_{n-1}}(z_2) - \frac{1}{b_n}\left(\sum_{i\in I_1} G^{(1)}_{ii}(z_1)\right)\left(\sum_{i\in I_1} G^{(1)}_{ii}(z_2)\right)
$$

$$
= \sigma^2 + \frac{2}{b_n}\sum_{i,j\in I_1} G^{(1)}_{ij}(z_1)G^{(1)}_{ij}(z_2) + \frac{\mu_4}{b_n}\sum_{i\in I_1} G^{(1)}_{ii}(z_1)G^{(1)}_{ii}(z_2) - \frac{3}{b_n}\sum_{i\in I_1} G^{(1)}_{ii}(z_1)G^{(1)}_{ii}(z_2)
$$

$$
+ \frac{1}{b_n}\widetilde{\gamma_{n-1}}(z_1)\widetilde{\gamma_{n-1}}(z_2)
$$

$$
= \sigma^2 + \frac{2}{b_n}\sum_{i,j\in I_1} G^{(1)}_{ij}(z_1)G^{(1)}_{ij}(z_2) + \frac{\kappa_4}{b_n}\sum_{i\in I_1} G^{(1)}_{ii}(z_1)G^{(1)}_{ii}(z_2) + \frac{1}{b_n}\widetilde{\gamma_{n-1}}(z_1)\widetilde{\gamma_{n-1}}(z_2),
$$

where $\kappa_4 = \mu_4 - 3$.

Proof of (v): Observe that

$$
B(z_2) = \left\langle G^{(1)}G^{(1)}m^{(1)}, m^{(1)}\right\rangle = \frac{1}{b_n}\sum_{i,j\in I_1}\left(G^{(1)}G^{(1)}\right)_{ij} w_{1i}w_{1j} = \frac{1}{b_n}\sum_{i,j\in I_1}\sum_{k=2}^{n} G^{(1)}_{ik}G^{(1)}_{kj}w_{1i}w_{1j},
$$

and

$$
\frac{d}{dz_2}G^{(1)}_{ij}(z_2) = \left(G^{(1)}(z_2)G^{(1)}(z_2)\right)_{ij} = \sum_{k=2}^{n} G^{(1)}_{ik}(z_2)G^{(1)}_{kj}(z_2).
$$

Now, proceed as in **(iv)** and use the above facts to prove the result. Here we skip the details.

□

Proof of Lemma A.0.2: **Proof of (i):** Since w_{jk} satisfies the Poincaré inequality with constant m and the Poincaré inequality tensorises, the joint distribution of $\{w_{jk}\}_{(j,k)\in I_n^+}$ on $\mathbb{R}^{n(b_n+1)}$

satisfies the Poincaré inequality with same constant m. Therefore we have

$$\mathrm{Var}\left(\Phi\left(\{w_{jk}\}_{(j,k)\in I_n^+}\right)\right) \leq \frac{1}{m}\sum_{(j,k)\in I_n^+}\mathbb{E}\left[\left|\frac{\partial\Phi}{\partial w_{jk}}\right|^2\right],$$

for any continuously differentiable function Φ. Therefore,

$$\begin{aligned}
\text{(A.17)}\qquad \mathrm{Var}\left(\sum_{(1,i)\in I_n} G_{ii}\right) &\leq \frac{1}{m}\sum_{(j,k)\in I_n^+}\mathbb{E}\left[\left|\frac{\partial}{\partial w_{jk}}\sum_{(1,i)\in I_n} G_{ii}\right|^2\right] \\
&\leq \frac{4}{mb_n}\sum_{(j,k)\in I_n^+}\mathbb{E}\left[\left|\sum_{(1,i)\in I_n} G_{ij}G_{ki}\right|^2\right] \\
&= \frac{4}{mb_n}\sum_{(j,k)\in I_n^+}\mathbb{E}\left[|\alpha_{kj}|^2\right] \quad \text{where } \alpha_{kj} = \sum_{(1,i)\in I_n} G_{ki}G_{ij} \\
&\leq \frac{4}{mb_n}\sum_{j,k=1}^{n}\mathbb{E}\left[|\alpha_{kj}|^2\right] \\
&= \frac{4}{mb_n}\mathbb{E}\left[\|VV^T\|_{Fb}^2\right] \\
&= \frac{4}{mb_n}\mathbb{E}\left[\sum_{i=1}^{n}|\beta_i|^2\right],
\end{aligned}$$

where

$$V = \begin{bmatrix} G_{11} & G_{12} & \cdots & G_{1k_n} & 0 & \cdots & 0 \\ G_{21} & G_{22} & \cdots & G_{2k_n} & 0 & \cdots & 0 \\ & & & \vdots & & & \\ G_{n1} & G_{n2} & \cdots & G_{nk_n} & 0 & \cdots & 0 \end{bmatrix}_{n\times n}$$

and $\|\cdot\|_{Fb}$ stands for the Frobenius norm, and β_is are the eigenvalues of VV^T. Here, we denote the set $\{i : (1,i) \in I_n\}$ by $\{1, 2, \ldots, k_n\}$. Observe that $k_n = 2b_n + 1$. Since $rank(VV^T) \leq k_n = O(b_n)$, we have $|\{i : \beta_i \neq 0\}| \leq k_n = O(b_n)$. Also we know that $\|V\| \leq \|G\|$. Therefore,

$$|\beta_i|^2 \leq \|VV^T\|^2 \leq \|G\|^4 \leq \frac{1}{|\Im z|^4}.$$

Consequently, we have

$$\text{Var}\left(\sum_{(1,i)\in I_n} G_{ii}\right) \le \frac{4}{mb_n}\mathbb{E}\left[\sum_{i=1}^{n}|\beta_i|^2\right] \le \frac{4}{mb_n}\frac{O(b_n)}{|\Im z|^4} = O(1). \tag{A.18}$$

This completes proof of first part of (A.7).

Recall the definition of A from (2.14), $A = z - \frac{1}{\sqrt{b_n}}w_{11} + \left(G^{(1)}m^{(1)}, m^{(1)}\right)$. Then

$$\begin{aligned} A^\circ &= A - \mathbb{E}[A] \\ &= -\frac{1}{\sqrt{b_n}}w_{11} + \frac{1}{b_n}\sum_{\substack{i\neq j\\ i,j\in I_1}} G_{ij}^{(1)}w_{1i}w_{1j} + \frac{1}{b_n}\sum_{i\in I_1}\left(G_{ii}^{(1)}w_{1i}^2 - \mathbb{E}[G_{ii}^{(1)}]\right), \end{aligned}$$

Consider

$$\begin{aligned} A_1^\circ &= A - \mathbb{E}_1[A] \\ &= -\frac{1}{\sqrt{b_n}}w_{11} + \frac{1}{b_n}\sum_{\substack{i\neq j\\ i,j\in I_1}} G_{ij}^{(1)}w_{1i}w_{1j} + \frac{1}{b_n}\sum_{i\in I_1}\left(G_{ii}^{(1)}w_{1i}^2 - G_{ii}^{(1)}\right). \end{aligned} \tag{A.19}$$

So we have

$$A^\circ - A_1^\circ = \frac{1}{b_n}\sum_{i\in I_1}\left(G_{ii}^{(1)} - \mathbb{E}\left[G_{ii}^{(1)}\right]\right) =: \frac{1}{b_n}\widetilde{\gamma_{n-1}}. \tag{A.20}$$

Hence

$$\mathbb{E}[|A^\circ|^2] = \mathbb{E}[|A_1^\circ + b_n^{-1}\widetilde{\gamma_{n-1}}|^2] \le 2\left[\mathbb{E}[|A_1^\circ|^2] + \frac{1}{b_n^2}\mathbb{E}[|\widetilde{\gamma_{n-1}}|^2]\right].$$

From (A.2), we know that $\mathbb{E}[|A_1^\circ|^2] = O\left(\frac{1}{b_n}\right)$ and from (A.18), We have $\mathbb{E}[|\widetilde{\gamma_{n-1}}|^2] = O(1)$. Combining these two facts and using (2.17), we have

$$\text{Var}(G_{11}(z)) = \mathbb{E}\left|\frac{1}{A} - \mathbb{E}\frac{1}{A}\right|^2 \le \mathbb{E}\left|\frac{1}{A} - \frac{1}{\mathbb{E}A}\right|^2 = \mathbb{E}\left|\frac{A^\circ}{A\mathbb{E}A}\right|^2 = O\left(\frac{1}{b_n}\right).$$

This completes the proof of second part.

Proof of (ii): *Proof of* (A.8)*:* Recall from (A.19)

$$A_1^\circ = -\frac{w_{11}}{\sqrt{b_n}} + \frac{1}{b_n}\sum_{i\neq j\in I_1} G_{ij}^{(1)} w_{1i}w_{1j} + \frac{1}{b_n}\sum_{i\in I_1} G_{ii}^{(1)}(w_{1i}^2)^\circ =: T_1 + T_2 + T_3. \tag{A.21}$$

We have $\mathbb{E}[|T_1|^4] = O\left(\frac{1}{b_n^2}\right)$. Now

$$\mathbb{E}\left[|T_2|^4\right] = \frac{1}{b_n^4}\mathbb{E}\left[\sum_{i\neq j,k\neq l,p\neq q,s\neq t\in I_1} G_{ij}^{(1)}\overline{G_{kl}^{(1)}}G_{pq}^{(1)}\overline{G_{st}^{(1)}} w_{1i}w_{1j}w_{1k}w_{1l}w_{1p}w_{1q}w_{1s}w_{1t}\right].$$

We use the similar technique as the moment method in the proof of the Semicircle Law. In the above sum of expectations, we have nonzero terms if the indices of w_{1m}'s match in a certain way. Non zero contribution to $\mathbb{E}[|T_2|^4]$ come from the two types of matches.

FIGURE A.1. Type I matching

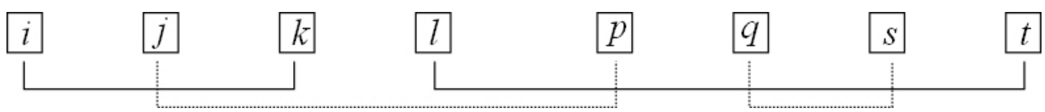


FIGURE A.2. Type II matching

Type I: Contribution from this kind of matching is

$$\begin{aligned}
\frac{1}{b_n^4}\mathbb{E}\left[\sum_{i\neq j,p\neq k\,\in I_1} |G_{ij}^{(1)}|^2|G_{pq}^{(1)}|^2 w_{1i}^2w_{1j}^2w_{1p}^2w_{1q}^2\right] &= \frac{1}{b_n^4}\mathbb{E}\mathbb{E}_1\left[\sum_{i\neq j,p\neq k\,\in I_1} |G_{ij}^{(1)}|^2|G_{pq}^{(1)}|^2 w_{1i}^2w_{1j}^2w_{1p}^2w_{1q}^2\right] \\
&= \frac{1}{b_n^4}\sum_{i\neq j}\sum_{p\neq q}\mathbb{E}\left[|G_{ij}^{(1)}|^2|G_{pq}^{(1)}|^2\right] \\
&\leq \frac{1}{b_n^4}\sum_{i\neq j}\sum_{p\neq q}\sqrt{\mathbb{E}\left[\left|G_{ij}^{(1)}\right|^4\right]\mathbb{E}\left[\left|G_{pq}^{(1)}\right|^4\right]} \\
&= \frac{1}{b_n^4}\sum_{i\neq j}\sum_{p\neq q} O\left(\frac{1}{b_n^2}\right) \quad \text{(using (A.4))} \\
&= O\left(\frac{1}{b_n^2}\right).
\end{aligned}$$

Type II: Similarly, contribution from the type II matching is

$$\frac{1}{b_n^4}\mathbb{E}\left[\sum_{i\neq j}\sum_{q\neq l\in I_1} G_{ij}^{(1)}\overline{G_{il}^{(1)}}G_{jq}^{(1)}\overline{G_{ql}^{(1)}}w_{1i}^2w_{1j}^2w_{1q}^2w_{1l}^2\right] = O\left(\frac{1}{b_n^2}\right).$$

Similarly, $\mathbb{E}[|T_3|^4] = O\left(\frac{1}{b_n^2}\right)$. Hence

$$\mathbb{E}[|A_1^\circ|^4] = O\left(\frac{1}{b_n^2}\right). \tag{A.22}$$

Using Lemma 4.4.3. from [**AGZ10**] with the help of the Poincaré inequality, we have $\mathbb{E}\left[\left|\widetilde{\gamma_{n-1}}\right|^4\right] \leq C\|\|\nabla\widetilde{\gamma_{n-1}}\|_2\|_\infty^4$, where C is a constant depends only on the constant m of the Poincaré inequality. Following the arguments given at the right side of (A.17) onward and (A.18), one can show that $\|\nabla\widetilde{\gamma_{n-1}}\|_2 \leq \frac{C}{|\Im z|^4}$, where C depends only on m. Hence $\mathbb{E}\left[\left|\widetilde{\gamma_{n-1}}\right|^4\right] = O(1)$. Consequently, using relation (A.20) and (A.22), we have $\mathbb{E}[|A^\circ|^4] = O\left(\frac{1}{b_n^2}\right)$. Then $\mathbb{E}[|A^\circ|^3] \leq \left(\mathbb{E}[|A^\circ|^4]\right)^{3/4} = O\left(\frac{1}{b_n^{3/2}}\right)$.

Proof of (A.9)*:* First we write B as

$$B = \left\langle G^{(1)}G^{(1)}m^{(1)}, m^{(1)}\right\rangle = \left\langle H^{(1)}m^{(1)}, m^{(1)}\right\rangle = \frac{1}{b_n}\sum_{i,j\in I_1} H_{ij}^{(1)}w_{1i}w_{1j},$$

where $H^{(1)} = G^{(1)}G^{(1)}$. Define

$$B_1^\circ := \frac{1}{b_n}\sum_{i\neq j\in I_1} H_{ij}^{(1)}w_{1i}w_{1j} + \frac{1}{b_n}\sum_{i\in I_1} H_{ii}^{(1)}(w_{1i}^2)^\circ.$$

Then we can write

$$\begin{aligned}
B^\circ &= B - \mathbb{E}[B]\\
&= \frac{1}{b_n}\sum_{i\neq j\in I_1} H_{ij}w_{1i}w_{1j} + \frac{1}{b_n}\sum_{i\in I_1}\left[H_{ii}^{(1)}w_{1i}^2 - \mathbb{E}[H_{ii}^{(1)}]\right]\\
&= B_1^\circ + \frac{1}{b_n}\sum_{i\in I_1}\left(H_{ii}^{(1)} - \mathbb{E}[H_{ii}^{(1)}]\right)\\
&= B_1^\circ + \frac{1}{b_n}\overline{\gamma_{n-1}},
\end{aligned}$$

where

$$\overline{\gamma_{n-1}}(z) = \sum_{i \in I_1} \left(H_{ii}^{(1)} - \mathbb{E}[H_{ii}^{(1)}] \right) = \sum_{i \in I_1} \sum_{j=2}^{n} \left(G_{ij}^{(1)} G_{ji}^{(1)} - \mathbb{E}\left[G_{ij}^{(1)} G_{ji}^{(1)} \right] \right) = \frac{d}{dz} \widetilde{\gamma_{n-1}}(z).$$

Proceeding as in the estimate of $\mathbb{E}[|A_1^{\circ}|^4]$, we can show

$$\mathbb{E}[|B_1^{\circ}|^4] = O\left(\frac{1}{b_n^2}\right). \tag{A.23}$$

We have shown that $\mathbb{E}[|\widetilde{\gamma_{n-1}}(z)|^4] = O(1)$. Using this fact and Cauchy's theorem we have $\mathbb{E}[|\overline{\gamma_{n-1}}(z)|^4] = O(1)$. Hence we have the result.

Proof of (iii):

$$\begin{aligned} &\operatorname{Var}\left\{ b_n \mathbb{E}_1 \left[A^{\circ}(z_1) A^{\circ}(z_2) \right] \right\} \\ &= \operatorname{Var}(T_1) + \operatorname{Var}(T_2) + \operatorname{Var}(T_3) + 2\operatorname{Cov}(T_1, T_2) + 2\operatorname{Cov}(T_2, T_3) + 2\operatorname{Cov}(T_3, T_1), \end{aligned}$$

where

$$T_1 = \frac{2}{b_n} \sum_{i,j \in I_1} G_{ij}^{(1)}(z_1) G_{ij}^{(1)}(z_2),\ T_2 = \frac{\kappa_4}{b_n} \sum_{i \in I_1} G_{ii}^{(1)}(z_1) G_{ii}^{(1)}(z_2) \text{ and } T_3 = \frac{1}{b_n} \widetilde{\gamma_{n-1}}(z_1) \widetilde{\gamma_{n-1}}(z_2).$$

Now, $\operatorname{Var}(T_2) = \frac{\kappa_4^2}{b_n^2} \operatorname{Var}\left\{ \sum_{i \in I_1} G_{ii}^{(1)}(z_1) G_{ii}^{(1)}(z_2) \right\}$ and

$$\begin{aligned} \operatorname{Var}\left\{ G_{ii}^{(1)}(z_1) G_{ii}^{(1)}(z_2) \right\} &= \mathbb{E}\left| G_{ii}^{(1)}(z_1) G_{ii}^{(1)}(z_2) - \mathbb{E}[G_{ii}^{(1)}(z_1) G_{ii}^{(1)}(z_2)] \right|^2 \\ &\leq \frac{2}{|\Im z_1|^2} \operatorname{Var}\left(G_{ii}^{(1)}(z_2) \right) + \frac{2}{|\Im z_2|^2} \operatorname{Var}\left(G_{ii}^{(1)}(z_1) \right) \\ &= \left(\frac{1}{|\Im z_1|^2} + \frac{1}{|\Im z_2|^2} \right) O\left(\frac{1}{b_n} \right). \end{aligned}$$

Therefore,

$$\operatorname{Var}(T_2) \leq \frac{\kappa_4^2}{b_n^2} \left(b_n O\left(\frac{1}{b_n} \right) + b_n^2 O\left(\frac{1}{b_n} \right) \right) = O\left(\frac{1}{b_n} \right).$$

Now

$$\begin{aligned}
\mathrm{Var}(T_3) &\le \frac{1}{b_n^2}\mathrm{Var}\left(\widetilde{\gamma_{n-1}}(z_1)\widetilde{\gamma_{n-1}}(z_2)\right)\\
&\le \frac{1}{b_n^2}\mathbb{E}\left[|\widetilde{\gamma_{n-1}}(z_1)|^2|\widetilde{\gamma_{n-1}}(z_2)|^2\right]\\
&\le \frac{1}{b_n^2}\sqrt{\mathbb{E}\left[|\widetilde{\gamma_{n-1}}(z_1)|^4\right]}\sqrt{\mathbb{E}\left[|\widetilde{\gamma_{n-1}}(z_2)|^4\right]}\\
&= \frac{1}{b_n^2}O(1).
\end{aligned}$$

Last equality holds, since $\mathbb{E}\left[|\widetilde{\gamma_{n-1}}(z_1)|^4\right] = O(1)$. And finally

$$\mathrm{Var}(T_1) = \frac{4}{b_n^2}\mathrm{Var}\left(\sum_{i,j\in I_1} G_{ij}^{(1)}(z_1)G_{ij}^{(1)}(z_2)\right).$$

Now, using the Poincaré inequality

$$\begin{aligned}
&\mathrm{Var}\left(\sum_{i,j\in I_1} G_{ij}^{(1)}(z_1)G_{ij}^{(1)}(z_2)\right)\\
&\le \frac{1}{m}\sum_{(s,t)\in I_n^+}\mathbb{E}\left[\left|\frac{\partial}{\partial w_{st}}\sum_{i,j\in I_1} G_{ij}^{(1)}(z_1)G_{ij}^{(1)}(z_2)\right|^2\right]\\
&\le \frac{1}{mb_n}\sum_{(s,t)\in I_n^+}\mathbb{E}\left[\left|\sum_{i,j\in I_1} G_{is}^{(1)}(z_1)G_{tj}^{(1)}(z_1)G_{ij}^{(1)}(z_2) + G_{ij}^{(1)}(z_1)G_{is}^{(1)}(z_2)G_{tj}^{(1)}(z_2)\right|^2\right]\\
&\le \frac{2}{mb_n}\sum_{(s,t)\in I_n^+}\mathbb{E}\left[\left|\sum_{i,j\in I_1} G_{is}^{(1)}(z_1)G_{tj}^{(1)}(z_1)G_{ij}^{(1)}(z_2)\right|^2\right]\\
&+ \frac{2}{mb_n}\sum_{(s,t)\in I_n^+}\mathbb{E}\left[\left|\sum_{i,j\in I_1} G_{ij}^{(1)}(z_1)G_{is}^{(1)}(z_2)G_{tj}^{(1)}(z_2)\right|^2\right]\\
&=: I_1 + I_2.
\end{aligned}$$

We estimate

$$I_1 = \frac{2}{mb_n}\sum_{(s,t)\in I_n^+}\mathbb{E}\left[\left|\sum_{i,j\in I_1} G_{is}^{(1)}(z_1)G_{tj}^{(1)}(z_1)G_{ij}^{(1)}(z_2)\right|^2\right]$$

$$
\begin{aligned}
&= \frac{2}{mb_n} \sum_{(s,t)\in I_n^+} \mathbb{E}\left[\left|\sum_{i\in I_1} G_{is}^{(1)}(z_1)G_{ti}^{(1)}(z_1,z_2)\right|^2\right] \\
&= \frac{2}{mb_n} \sum_{(s,t)\in I_n^+} \mathbb{E}\left[\left|G_{st}^{(1)}(z_1,z_2,z_1)\right|^2\right] \\
&\le \frac{2}{mb_n}\mathbb{E}\left[\sum_{s,t=1}^{n}\left|G_{st}^{(1)}(z_1,z_2,z_1)\right|^2\right] \\
&= \frac{2}{mb_n}\mathbb{E}\left[\|A\|_{Fb}^2\right] \\
&= \frac{2}{mb_n}\mathbb{E}\left[\sum_{i=1}^{n}\beta_i^2\right] \\
&\le \frac{C(z_1,z_2)}{mb_n}O(b_n) = O(1),
\end{aligned}
$$

where $\|\cdot\|_{Fb}$ is the Frobenius norm, β_i are the eigenvalues of VV^*, and V is the following matrix

$$
V_{n\times n} = \begin{bmatrix} G_{11}^{(1)}(z_1) & G_{12}^{(1)}(z_1) & \cdots & G_{1k_n}^{(1)}(z_1) \\ G_{21}^{(1)}(z_1) & G_{22}^{(1)}(z_1) & \cdots & G_{2k_n}^{(1)}(z_1) \\ \vdots & \vdots & & \vdots \\ G_{n1}^{(1)}(z_1) & G_{n2}^{(1)}(z_1) & \cdots & G_{nk_n}^{(1)}(z_1) \end{bmatrix}_{n\times k_n} \begin{bmatrix} G_{11}^{(1)}(z_2) & G_{12}^{(1)}(z_2) & \cdots & G_{1k_n}^{(1)}(z_2) \\ G_{21}^{(1)}(z_2) & G_{22}^{(1)}(z_2) & \cdots & G_{2k_n}^{(1)}(z_2) \\ \vdots & \vdots & & \vdots \\ G_{k_n1}^{(1)}(z_2) & G_{k_n2}^{(1)}(z_2) & \cdots & G_{k_nk_n}^{(1)}(z_2) \end{bmatrix}_{k_n\times k_n}
$$

$$
\times \begin{bmatrix} G_{11}^{(1)}(z_1) & G_{12}^{(1)}(z_1) & \cdots & G_{1n}^{(1)}(z_1) \\ G_{21}^{(1)}(z_1) & G_{22}^{(1)}(z_1) & \cdots & G_{2n}^{(1)}(z_1) \\ \vdots & \vdots & & \vdots \\ G_{k_n1}^{(1)}(z_1) & G_{t2}^{(1)}(z_1) & \cdots & G_{k_nn}^{(1)}(z_1) \end{bmatrix}_{k_n\times n}.
$$

Here we denoted the elements of set I_1 as $I_1 = \{1, 2, \ldots, k_n\}$. Observe that $k_n = 2b_n$. Rank of $V \le k_n = O(b_n)$. This implies

$$
\sum_{i=1}^{n}\beta_i^2 \le k_n C(z_1,z_2) = O(b_n)C(z_1,z_2).
$$

Therefore, $\mathrm{Var}(T_1) = O\left(\frac{1}{b_n^2}\right)$, and hence $\mathrm{Var}\{b_n\mathbb{E}_1\left[A^\circ(z_1)A^\circ(z_2)\right]\} = O\left(\frac{1}{b_n}\right)$.

Second part of **(iii)** follows from the following two facts with the help of Cauchy's theorem.

$$b_n \mathbb{E}_1 \left[A^\circ(z_1) B^\circ(z_2)\right] = b_n \frac{d}{dz_2} \mathbb{E}_1 \left[A^\circ(z_1) A^\circ(z_2)\right]$$
$$\text{and} \quad \mathrm{Var}\left\{b_n \mathbb{E}_1 \left\{A^\circ(z_1) A^\circ(z_2)\right\}\right\} = O\left(\frac{1}{b_n}\right).$$

Here we skip the details.

Proof of (iv): Using (2.25) and (A.1), and proceeding as the proof of proposition 2.2.2,

$$\begin{aligned}
\mathbb{E}\left[\left|\gamma^\circ_{n-1}(z) - \gamma^\circ_n(z)\right|^4\right] &= \mathbb{E}\left[\left|\left(\mathrm{Tr}G^{(1)}(z) - \mathbb{E}[\mathrm{Tr}G^{(1)}(z)]\right) - (\mathrm{Tr}G(z) - \mathbb{E}[\mathrm{Tr}G(z)])\right|^4\right] \\
&= \mathbb{E}\left[\left|\frac{1+B(z)}{A(z)} - \mathbb{E}\left[\frac{1+B(z)}{A(z)}\right]\right|^4\right] \\
&\leq \frac{C}{|\Im z|^8}\left[\mathbb{E}\left[|A^\circ|^4\right] + \mathbb{E}\left[|B^\circ|^4\right] + \mathbb{E}\left[|A^\circ|^4\right]\right] = O\left(\frac{1}{b_n^2}\right).
\end{aligned}$$

The last equality follows from the estimates (A.8) and (A.9).

Using martingale differences as in the proof of Proposition 2.2.2,

$$\mathbb{E}\left[|\gamma^\circ_n|^4\right] \leq Cn \sum_{k=1}^{n} \mathbb{E}\left[|\gamma_n - \mathbb{E}_k[\gamma_n]|^4\right].$$

Consider for $k = 1$, others will be similar.

$$\begin{aligned}
\mathbb{E}\left[|\gamma_n - \mathbb{E}_1[\gamma_n]|^4\right] &= \mathbb{E}\left[\left|\mathrm{Tr}(G - G^{(1)}) - \mathbb{E}_1\left[\mathrm{Tr}(G - G^{(1)})\right]\right|^4\right] \\
&= \mathbb{E}\left[\left|\frac{1+B(z)}{A(z)} - \mathbb{E}_1\left[\frac{1+B(z)}{A(z)}\right]\right|^4\right] \\
&\leq C_1(z)\mathbb{E}[|A_1^\circ|^4] + C_2(z)\mathbb{E}[|B_1^\circ|^4] \\
&= O\left(\frac{1}{b_n^2}\right).
\end{aligned}$$

The last equality follows from (A.22) and (A.23). Hence we have the result.

Proof of (v): Using resolvent identity,

$$(X_2 - zI)^{-1} = (X_1 - zI)^{-1} + (X_1 - zI)^{-1}(X_1 - X_2)(X_2 - zI)^{-1},$$

we have

$$zG_{11}(z) = -1 + \sum_{(1,k)\in I_n} m_{1k}G_{k1}, \tag{A.24}$$

where I_n is defined in (2.1) and m_{ij}s are defined in (2.2). Now, to analyse the terms $\mathbb{E}[m_{1k}G_{k1}]$, we use the following (see eg. [**LP09**]): Given ξ, a real valued random variable with $p+2$ finite moments, and ϕ, a function from $\mathbb{C} \to \mathbb{R}$ with $p+1$ continuous and bounded derivatives then:

$$\mathbb{E}[\xi\phi(\xi)] = \sum_{a=0}^{p} \frac{\kappa_{a+1}}{a!}\mathbb{E}\left[\phi^{(a)}(\xi)\right] + \epsilon_{p+1} \tag{A.25}$$

where κ_a is the a-th cumulant of ξ, $|\epsilon_{p+1}| \le C\sup_t |\phi^{(p+1)}(t)|\mathbb{E}[|\xi|^{p+2}]$ and C depends only on p. Since $f_n(z) = \frac{1}{n}\mathbb{E}[\mathrm{Tr}G(z)] = \mathbb{E}[G_{11}(z)]$, using (A.24) and (A.25) we get

$$zf_n(z) = -1 + \sum_{(1,k)\in I_n} \mathbb{E}[m_{1k}G_{k1}] = -1 - \sum_{k\in I_1} \frac{1}{b_n}\mathbb{E}\left[G_{k1}^2 + G_{kk}G_{11}\right] + r_n, \tag{A.26}$$

where r_n contains the third cumulant term corresponding to $p=2$ in (A.25) for $k \neq 1$, and the error terms due to the truncation of the decoupling formula (A.25) at $p=2$ for $k \neq 1$ and at $p=0$ for $k=1$. We write (A.26)

$$\begin{aligned} zf_n(z) &= -1 - \frac{1}{b_n}\mathbb{E}[G_{11}]\mathbb{E}\left[\sum_{k\in I_1} G_{kk}\right] - \frac{1}{b_n}\mathrm{Cov}\left(G_{11}, \sum_{k\in I_1} G_{kk}\right) - \frac{1}{b_n}\mathbb{E}\left[\sum_{k\in I_1} G_{k1}^2\right] + r_n \\ &= -1 - f_n(z)(2f_n(z)) - \frac{1}{b_n}\mathrm{Cov}\left(G_{11}, \sum_{k\in I_1} G_{kk}\right) - \frac{1}{b_n}\mathbb{E}\left[\sum_{k\in I_1} G_{k1}^2\right] + r_n. \end{aligned}$$

Now, by the Cauchy-Schwarz inequality and (A.18) we get

$$\begin{aligned} \frac{1}{b_n}\left|\mathrm{Cov}\left(G_{11}, \sum_{k\in I_1} G_{kk}\right)\right| &\le \frac{1}{b_n}\sqrt{\mathrm{Var}(G_{11})}\sqrt{\mathrm{Var}\left(\sum_{k\in I_1} G_{kk}\right)} \\ &\le \frac{1}{b_n}2|\Im z|^{-1}\sqrt{O(|\Im z|^{-4})} \\ &= O\left(\frac{1}{b_n|\Im z|^3}\right). \end{aligned}$$

Also notice that

$$\frac{1}{b_n}\left|\mathbb{E}\left(\sum_{k\in I_1} G_{k1}^2\right)\right| \le \frac{1}{b_n}|\Im z|^{-2}.$$

We claim $r_n = O\left(\frac{1}{b_n|\Im z|^4}\right)$. To prove this, observe that the third cumulant term gives

$$\frac{\kappa_3}{2b_n^{3/2}}\mathbb{E}\left[\sum_{k\in I_1} 2(G_{1k})^3 + 6G_{11}G_{1k}G_{kk}\right] \tag{A.27}$$

Since

$$\sum_{k\in I_1} |G_{1k}|^2 \le \|G\|^2 \le |\Im z|^{-2} \quad \text{and} \quad |G_{ij}| \le |\Im z|^{-1},$$

we conclude that the third cumulant term contributes $O\left(\frac{1}{b_n|\Im z|^3}\right)$ to r_n. In a similar manner, the error due to truncation of decoupling formula (A.25) at $p = 2$ is $O\left(\frac{1}{b_n|\Im z|^4}\right)$. Similarly, the error term due to truncation of decoupling formula at $p = 0$ for $k = 1$ is $O\left(\frac{1}{b_n|\Im z|^2}\right)$. Thus the claim is proved. Hence

$$zf_n(z) = -1 - 2f_n^2(z) + O\left(\frac{1}{b_n|\Im z|^4}\right) \quad \text{for } z \in \mathbb{C}\backslash\mathbb{R}.$$

Now, following similar argument given in the proof of (3.1) in [**PRS12**], one can show that

$$|f_n(z) - f(z)| \le O\left(\frac{1}{b_n|\Im z|^6}\right)$$

where $f(z) = \frac{1}{4}(-z + \sqrt{z^2 - 8})$.

Proof of (vi): Recall $A(z) = z - b_n^{-1/2}w_{11} + b_n^{-1}\sum_{i,j\in I_1} G_{ij}^{(1)} w_{1i}w_{1j}$. Now, using (A.12) with G replaced by $G^{(1)}$, we have

$$(\mathbb{E}[A(z)])^{-1} = \frac{1}{z + b_n^{-1}\sum_{j\in I_1}\mathbb{E}[G_{jj}^{(1)}]} = \frac{1}{z + 2f_n(z)} = (z + 2f(z))^{-1} + O(b_n^{-1}) = -f(z) + O(b_n^{-1})$$

Hence $(\mathbb{E}[A(z)])^{-1} = -f(z) + O(b_n^{-1})$. To prove the second part, observe that

$$\mathbb{E}[B(z)] = \frac{1}{b_n}\mathbb{E}\left[\sum_{i,j\in I_1}(G^{(1)}G^{(1)})_{ij}w_{1i}w_{1j}\right] = \frac{1}{b_n}\mathbb{E}\left[\sum_{i\in I_1}(G^{(1)}G^{(1)})_{ii}\right]$$

$$= \frac{1}{b_n}\mathbb{E}\left[\sum_{i\in I_1}\sum_{k=2}^{n} G_{ik}^{(1)}G_{ki}^{(1)}\right] = \frac{1}{b_n}\sum_{i\in I_1}\frac{d}{dz}G_{ii}^{(1)}$$

Again using (A.12) and Cauchy's integral formula, we have

$$\mathbb{E}[B(z)] = \frac{d}{dz}(2f_n(z)) = 2f'(z) + O(b_n^{-1}).$$

This completes the proof of Lemma A.0.2. □

A.0.1. Proof of (2.29):

PROOF. We have to find the limit of

$$\mathbb{E}[T_n] = \frac{2}{b_n}\mathbb{E}\left[\sum_{i,j\in I_1} G_{ij}^{(1)}(z)G_{ij}^{(1)}(z_\mu)\right]$$

as $n\to\infty$, where $I_1 = \{2 \le i \le n : (1,i) \in I_n\}$. Let $f,g \in C_b(\mathbb{R})$. Define a bilinear form on $C_b(\mathbb{R})$ as

$$\langle f,g\rangle_n = \frac{1}{b_n}\mathbb{E}\left[\sum_{i,j\in I_1} f(M)_{ij}g(M)_{ji}\right]. \tag{A.28}$$

Then $\mathbb{E}[T_n] = \langle h(M), h_\mu(M)\rangle_n$, where $h(x) = (x-z)^{-1}$ and $h_\mu(x) = (x-z_\mu)^{-1}$.

LEMMA A.0.3. *For $f,g\in C_b(\mathbb{R})$ the limit $\langle f,g\rangle = \lim_{n\to\infty}\langle f,g\rangle_n$ exists.*

PROOF. The idea of the proof is similar to the proof of Lemma 3.11 of [**LS13**]. First we prove this result for monomials. Although monomials are unbounded, still (A.28) makes sense for all n, since all moments of the entries of M are finite. Consider $f(x) = x^l$ and $g(x) = x^m$ where $l,m\in\mathbb{N}$. Then

$$\langle x^l, x^m\rangle_n = \frac{1}{b_n^{1+(l+m)/2}} \sum_{\substack{(i_0,i_1),(i_1,i_2),\ldots,(i_{l+m-1},i_0)\in I_n \\ i_o,i_l\in I_1}} \mathbb{E}\left[w_{i_0i_1}w_{i_1i_2}\ldots w_{i_{l+m-1}i_0}\right]$$

If $(l+m)$ is odd then $\langle x^l, x^m\rangle_n \to 0$ using independence of matrix entries and $\mathbb{E}(w_{ij}) = 0$, and order counting of independent vertices. The argument is similar to the combinatorial argument given in the proof of Wigner semicircular law (see [**AGZ10**]). We leave it for the reader.

Now, we assume $l+m$ is even. Then

$$\begin{aligned}
\langle x^l, x^m\rangle_n &= \frac{1}{b_n^{1+(l+m)/2}} \sum_{\substack{(i_0,i_1),(i_1,i_2),\ldots,(i_{l+m-1},i_0)\in I_n \\ i_o,i_l\in I_1}} \mathbb{E}\left[w_{i_0i_1}w_{i_1i_2}\ldots w_{i_{l+m-1}i_0}\right] \\
&= \frac{1}{b_n^{1+(l+m)/2}} \sum_{\substack{(i_0,i_1),(i_1,i_2),\ldots,(i_{l+m-1},i_0)\in I_n \\ i_o,i_l\in I_1}} \mathbb{E}\left[w_{1i_0}w_{i_0i_1}w_{i_1i_2}\ldots w_{i_{l+m-1}i_0}w_{i_01}\right] + O(b_n^{-1}) \\
&= \frac{1}{b_n^{1+(l+m)/2}} \sum_{\substack{(i_0,i_1),(i_1,i_2),\ldots,(i_{l+m-1},i_0)\in I_n \\ (1,i_o),(1,i_l)\in I_n}} \mathbb{E}\left[w_{1i_0}w_{i_0i_1}w_{i_1i_2}\ldots w_{i_{l+m-1}i_0}w_{i_01}\right] + O(b_n^{-1})
\end{aligned} \tag{A.29}$$

The second last equality in (A.29) holds due to order calculation of independent vertices and independence of matrix entries. Now, define for $k = 1, 2, \ldots, l+m$,

$$x_k = \begin{cases} i_k - i_{k-1} & \text{if } \ |i_k - i_{k-1}| \le b_n \\ (i_k - i_{k-1}) - n & \text{if } \ i_k - i_{k-1} > b_n \\ n + (i_k - i_{k-1}) & \text{if } \ i_k - i_{k-1} < -b_n \end{cases} \quad \text{with } i_{l+m} = i_0, \text{ and}$$

$$x_0 = \begin{cases} i_0 - 1 & \text{if } \ |i_0 - 1| \le b_n \\ (i_0 - 1) - n & \text{if } \ i_0 - 1 > b_n \end{cases} \quad \text{and} \quad x_{l+m+1} = \begin{cases} 1 - i_0 & \text{if } \ |1 - i_0| \le b_n \\ n + (1 - i_0) & \text{if } \ 1 - i_0 < -b_n. \end{cases}$$

Note, $x_0 = -x_{l+m+1}$. Since l, m are fixed and $b_n \to \infty$, for large n the restrictions $\{(i_0, i_1), (i_1, i_2), \ldots, (i_{l+m-1}, i_0) \in I_n$ and $(1, i_0), (1, i_l) \in I_n\}$ are equivalent to $\{|x_0|, |x_1|, \ldots, |x_{l+m}| \le b_n,\ x_0 + x_1 + \cdots + x_{l+m} + x_{l+m+1} = 0$ and $|x_0 + x_1 + \cdots + x_l| \le b_n\}$. Also observe that $x_0 + x_1 + \cdots + x_{l+m} + x_{l+m+1} = 0$ is same as $x_1 + \cdots + x_{l+m} = 0$ since $x_0 = -x_{l+m+1}$. Therefore for large n

$$\langle x^l, x^m\rangle_n = \frac{1}{b_n^{1+(l+m)/2}} \sum_{\substack{x_1+\cdots+x_{l+m}=0 \\ |x_i|\le b_n, 0\le i\le l+m,\ |x_0+x_1+\cdots+x_l|\le b_n}} \mathbb{E}\left[w_{1i_0}w_{i_0i_1}w_{i_1i_2}\ldots w_{i_{l+m-1}i_0}w_{i_01}\right) + O(b_n^{-1}).$$

Without loss of generality, we assume that $l \le m$. Each $\{i_0, i_1, i_2, \ldots, i_{l+m-1}, i_0\}$ is a closed path such that distance between the end points of each edge is bounded by b_n. As in the proof of Wigner semicircular law only the paths whose edges are pair matched contributes to the limit, here also, only such paths contribute to the limit. And contribution of each path is $\mathbb{E}(w_{1i_0}w_{i_0i_1}\dots w_{i_{l+m-1}i_0}w_{i_01}) = 1$ since $\mathbb{E}(w_{ij}^2) = 1$. Each such path corresponds to a Dyck path of length $(l+m)$. Recall that a Dyck path $(S(0), S(1), \dots, S(l+m))$ of length $(l+m)$ satisfies (see [**AGZ10**])

$$S(0) = S(l{+}m) = 0,\ S(1), S(2), \dots, S(l{+}m{-}1) \geq 0 \text{ and } |S(i{+}1){-}S(i)| = 1, \text{ for } i = 0, 1, \dots, l{+}m{-}1.$$

Specifically, $S(t+1) - S(t) = 1$ if the non-oriented edge (i_t, i_{t+1}) appears in $\{i_0, i_1, \dots, i_{l+m-1}, i_0\}$ for the first time and $S(t+1) - S(t) = -1$ if the edge (i_t, i_{t+1}) appears in $\{i_0, i_1, \dots, i_{l+m-1}, i_0\}$ for the second time.

Here each Dyck path does not give equal contribution to the limit due to the condition that $(1, i_l) \in I_n$ and in terms of x_i, which is same as $|x_0 + x_1 + \cdots + x_l| \leq b_n$. We have to take into account this condition. Suppose $S(l) = k$, $0 \leq k \leq l$. Then during the first l steps of the path $\{i_0, i_1, \dots, i_{l+m-1}, i_0\}$, k edges appear only once and $(l-k)/2$ edges appear twice. The edges appearing twice, the corresponding two number x_i have same absolute value but with different sign. We rename the remaining k numbers x_i which appear only once as $y_1, y_2, \dots, y_k$ (according to their order of appearance) and x_0 as y_0. So the condition $|x_0 + x_1 + \ldots + x_l| \leq b_n$ reduces to $|y_0 + y_1 + \ldots + y_k| \leq b_n$. Therefore

$$\begin{aligned}\langle x^l, x^m\rangle_n &= \frac{1}{b_n^{1+(l+m)/2}} \sum_{k=0}^{l} |\{\text{Dyck path of length } l+m \text{ with } S(l) = k\}| \\ &\quad \times |\{|y_0| \leq b_n, |y_1| \leq b_n, \dots, |y_k| \leq b_n, \dots, |y_{l+m}| \leq b_n, |y_0 + y_1 + \cdots + y_k| \leq b_n\}| + O(b_n^{-1}).\end{aligned}$$

and

$$\begin{aligned}&\langle x^l, x^m\rangle \\ &= \lim_{n\to\infty} \langle x^l, x^m\rangle_n \\ &= (\sqrt{2})^{l+m+2} \sum_{k=0}^{l} |\{\text{Dyck path of length } l+m \text{ with } S(l) = k\}| \\ &\times \text{Vol}\{|t_0| \leq 1/2, |t_1| \leq 1/2, \dots, |t_{\frac{l+m}{2}}| \leq 1/2, |t_0 + t_1 + \cdots + t_k| \leq 1/2\}\end{aligned}$$

$$= (\sqrt{2})^{l+m+2} \sum_{k=0}^{l} |\{\text{Dyck path of length } l+m \text{ with } S(l)=k\}| \times P(|T_0 + T_1 + \cdots + T_k| \leq 1/2),$$

where $T_0, T_1, \ldots, T_{\frac{l+m}{2}}$ are independent random variables uniformly distributed on $[-1/2, 1/2]$. Let $S_{k+1} = T_0 + T_1 + \ldots + T_k$. Then

$$\mathbb{E}\left[e^{ixS_{k+1}}\right] = \left(\mathbb{E}[e^{ixT_0}]\right)^{k+1} = \left(\frac{\sin x/2}{x/2}\right)^{k+1}.$$

Using inversion formula, the density of S_{k+1} is given by

$$f_{k+1}(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ixs} \left(\frac{\sin x/2}{x/2}\right)^{k+1} dx.$$

Now,

$$\gamma_{k+1} := P(|S_{k+1}| \leq 1/2) = \int_{-1/2}^{1/2} f_{k+1}(s) ds = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\frac{\sin x/2}{x/2}\right)^{k+2} dx = f_{k+2}(0),$$

using [**Cra99**] we get exact formula of γ_{k+1}:

$$\text{(A.30)} \qquad \gamma_{k+1} = \begin{cases} \frac{1}{(k+1)!} \sum_{s=0}^{(k+1)/2} (-1)^s \binom{k+2}{s} \left(\frac{k+1}{2} - s + \frac{1}{2}\right)^{k+1} & \text{if } \quad k+1 \text{ even} \\ \frac{1}{(k+1)!} \sum_{s=0}^{k/2} (-1)^s \binom{k+2}{s} \left(\frac{k+1}{2} - s + \frac{1}{2}\right)^{k+1} & \text{if } \quad k+1 \text{ odd.} \end{cases}$$

The number of Dyck path of length $l+m$ with $S(l) = k$ is

$$\text{(A.31)} \qquad \left[\binom{l}{\frac{l-k}{2}} - \binom{l}{\frac{l-k-2}{2}}\right] \times \left[\binom{m}{\frac{m-k}{2}} - \binom{m}{\frac{m-k-2}{2}}\right] = \frac{(k+1)^2}{(l+1)(m+1)} \binom{l+1}{\frac{l+k+2}{2}} \binom{m+1}{\frac{m+k+2}{2}}.$$

Hence from (A.30) and (A.31), we get

$$\langle x^l, x^m \rangle = (\sqrt{2})^{l+m+2} C_{l,m}$$

where $C_{l,m} = 0$ if $(l+m)$ is odd and

$$\begin{aligned} C_{l,m} &= \sum_{k=0}^{l} \frac{(k+1)^2}{(l+1)(m+1)} \binom{l+1}{\frac{l+k+2}{2}} \binom{m+1}{\frac{m+k+2}{2}} \gamma_{k+1} \\ &= \begin{cases} \sum_{k=0}^{l/2} \frac{(2k+1)^2}{(l+1)(m+1)} \binom{l+1}{\frac{l-2k}{2}} \binom{m+1}{\frac{m-2k}{2}} \gamma_{2k+1} & \text{if } \quad l \text{ even} \\ \sum_{k=0}^{(l-1)/2} \frac{(2k+2)^2}{(l+1)(m+1)} \binom{l+1}{\frac{l-2k-1}{2}} \binom{m+1}{\frac{m-2k-1}{2}} \gamma_{2k+2} & \text{if } \quad l \text{ odd} \end{cases} \end{aligned}$$

if $(l+m)$ is even and $l \le m$, otherwise, $C_{l,m} = C_{m,l}$. If f, g are polynomials, $f(x) = \sum_{i=0}^{p} a_i x^i,\ g(x) = \sum_{i=0}^{q} b_i x^i$, then by linearity

$$\langle f, g\rangle = \sum_{i=0}^{p}\sum_{j=0}^{q} a_i b_j (\sqrt{2})^{i+j+2} C_{i,j}. \tag{A.32}$$

For general bounded continuous functions f, g, to show that $\langle f, g\rangle$ exists we have to use the Stone-Weierstrass theorem to approximate f, g by appropriate polynomial and then (A.32). The argument is similar to the argument given in the proof of Lemma 3.11 of [**LS13**]. We skip the details. □

In the next lemma we diagonalize the bilinear form $\langle f, g\rangle$.

LEMMA A.0.4. *Let $\{U_n(x)\}_{n\ge 0}$ be the rescaled Chebyshev polynomial of the second kind on $[-2\sqrt{2}, 2\sqrt{2}]$,*

$$U_n(x) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \binom{n-k}{k} \left(\frac{x}{\sqrt{2}}\right)^{n-2k}.$$

Then $\{U_n(x)\}$ are orthogonal with respect to the bilinear form (A.32)*, that is,*

$$\langle U_n, U_m\rangle = 2\delta_{nm}\gamma_{n+1}, \tag{A.33}$$

where γ_{n+1} is defined in (A.30)*.*

PROOF. The proof of this lemma is similar to the proof of Lemma 3.12 of [**LS13**]. For sake of completeness we outline it here. Since $\langle x^l, x^m\rangle = 0$ if $l+m$ is odd, from linearity $\langle U_l, U_m\rangle = 0$ if $l+m$ is odd. We are left to compute $\langle U_{2n}, U_{2m}\rangle$ and $\langle U_{2n+1}, U_{2m+1}\rangle$. We first compute $\langle x^{2l}, U_{2n}\rangle$ and $\langle x^{2l+1}, U_{2n+1}\rangle$ for $l = 0, 1, \ldots, n$.

$$\begin{aligned}
\langle x^{2l}, U_{2n}\rangle &= (\sqrt{2})^{2l+2} \sum_{k=0}^{n} (-1)^k \binom{2n-k}{k} C_{2l,2n-2k} \\
&= (\sqrt{2})^{2l+2} \Bigg[\sum_{k=0}^{n-l} (-1)^k \binom{2n-k}{k} \sum_{t=0}^{l} \frac{(2t+1)^2}{(2l+1)(2n-2k+1)} \binom{2l+1}{l-t}\binom{2n-2k+1}{n-k-t}\gamma_{2t+1} \\
&\quad + \sum_{k=n-l+1}^{n} (-1)^k \binom{2n-k}{k} \sum_{t=0}^{n-k} \frac{(2t+1)^2}{(2l+1)(2n-2k+1)} \binom{2l+1}{l-t}\binom{2n-2k+1}{n-k-t}\gamma_{2t+1} \Bigg]
\end{aligned}$$

$$= (\sqrt{2})^{2l+2} \sum_{t=0}^{l} \frac{(2t+1)^2}{2l+1} \binom{2l+1}{l-t} \left[\sum_{k=0}^{n-t} \frac{(-1)^k (2n-k)!}{k!(n-k-t)!(n-k+t+1)!} \right] \gamma_{2t+1}$$
$$= (\sqrt{2})^{2l+2} \sum_{t=0}^{l} \frac{(2t+1)^2}{2l+1} \binom{2l+1}{l-t} G_1(n,t) \gamma_{2t+1},$$

where

$$G_1(n,t) = \sum_{k=0}^{n-t} \frac{(-1)^k (2n-k)!}{k!(n-k-t)!(n-k+t+1)!}.$$

Similarly,

$$\langle x^{2l+1}, U_{2n+1} \rangle = (\sqrt{2})^{2l+3} \sum_{t=0}^{l} \frac{(2t+2)^2}{2l+2} \binom{2l+2}{l-t} \left[\sum_{k=0}^{n-t} \frac{(-1)^k (2n+1-k)!}{k!(n-k-t)!(n-k+t+2)!} \right] \gamma_{2t+2}$$
$$= (\sqrt{2})^{2l+3} \sum_{t=0}^{l} \frac{(2t+2)^2}{2l+2} \binom{2l+2}{l-t} G_2(n,t) \gamma_{2t+2},$$

where

$$G_2(n,t) = \sum_{k=0}^{n-t} \frac{(-1)^k (2n+1-k)!}{k!(n-k-t)!(n-k+t+2)!}.$$

$G_1(n,t)$ and $G_2(n,t)$ can be written in terms of hypergeometric function as follows:

$$G_1(n,t) = \frac{(2n)!}{(n-t)!(n+t+1)!} \; {}_2F_1 \left({-(n-t), -(n+t+1) \atop -2n} ; \; 1 \right)$$
$$G_2(n,t) = \frac{(2n+1)!}{(n-t)!(n+t+2)!} \; {}_2F_1 \left({-(n-t), -(n+t+2) \atop -2n-1} ; \; 1 \right)$$

where ${}_2F_1$ is a hypergeometric function. By the Chu-Vandermonde identity (see [**AAR99**]), we have

$${}_2F_1 \left({-(n-t), -(n+t+1) \atop -2n} ; \; 1 \right) = \frac{(-n+t+1)_{n-t}}{(-2n)_{n-t}},$$
$${}_2F_1 \left({-(n-t), -(n+t+2) \atop -2n-1} ; \; 1 \right) = \frac{(-n+t+1)_{n-t}}{(-2n-1)_{n-t}},$$

where $(a)_n = a(a+1)\cdots(a+n-1)$. Since

$$(-n+t+1)_{n-t} = \begin{cases} 0 & \text{if} \quad t = 0, 1, \cdots, n-1 \\ 1 & \text{if} \quad t = n \end{cases}$$

we have $G_1(n,t) = 0$, $G_2(n,t) = 0$ for $t = 0, 1, \ldots, n-1$ and $G_1(n,n) = 1/(2n+1)$, $G_2(n,n) = 1/(2n+2)$. Therefore, $\langle x^{2l}, U_{2n}\rangle = 0$ for $0 \le l \le n-1$ and

$$\langle x^{2n}, U_{2n}\rangle = (\sqrt{2})^{2n+2}\gamma_{2n+1}.$$

Similarly, $\langle x^{2l+1}, U_{2n+1}\rangle = 0$ for $0 \le l \le n-1$ and

$$\langle x^{2n+1}, U_{2n+1}\rangle = (\sqrt{2})^{2n+3}\gamma_{2n+2}.$$

Therefore

$$\langle U_{2n}, U_{2n}\rangle = 2\gamma_{2n+1} \text{ and } \langle U_{2n+1}, U_{2n+1}\rangle = 2\gamma_{2n+2}.$$

This completes the proof of the lemma. □

Now, we complete the proof of (2.29). For $f, g \in C_b(\mathbb{R})$, if

$$f_k = \frac{1}{4\pi}\int_{-2\sqrt{2}}^{2\sqrt{2}} f(x)U_k(x)\sqrt{8-x^2}dx, \quad g_k = \frac{1}{4\pi}\int_{-2\sqrt{2}}^{2\sqrt{2}} g(x)U_k(x)\sqrt{8-x^2}dx,$$

then

$$\begin{aligned} \text{(A.34)} \qquad \langle f, g\rangle &= \sum_{k=0}^{\infty} f_k g_k 2\gamma_{k+1} \\ &= \frac{1}{8\pi^3}\int_{-2\sqrt{2}}^{2\sqrt{2}}\int_{-2\sqrt{2}}^{2\sqrt{2}} f(x)g(y)\sqrt{8-x^2}\sqrt{8-y^2}\left[\pi\sum_{k=0}^{\infty} U_k(x)U_k(y)\gamma_{k+1}\right]dxdy \\ &= \frac{1}{8\pi^3}\int_{-2\sqrt{2}}^{2\sqrt{2}}\int_{-2\sqrt{2}}^{2\sqrt{2}} f(x)g(y)\sqrt{8-x^2}\sqrt{8-y^2}F(x,y)dxdy \end{aligned}$$

where

$$\text{(A.35)} \qquad F(x,y) = \pi\sum_{k=0}^{\infty} U_k(x)U_k(y)\gamma_{k+1} = 2\int_{-\infty}^{\infty}\frac{z-z^3}{2(1-z^2)^2 + z^2(x^2+y^2) - z(1+z^2)xy}ds$$

with $z = \frac{\sin s}{s}$. Now, (A.34) holds due to (A.33) and orthogonality of Chebyshev polynomial with respect to the Wigner semicircular law, that is,

$$\int_{-2\sqrt{2}}^{2\sqrt{2}} U_n(x)U_m(x)\frac{1}{4\pi}\sqrt{8-x^2}dx = \delta_{mn}.$$

And (A.35) is a straightforward consequence of the Fourier analysis using the following fact

$$U_n(x) = \frac{\sin[(n+1)\theta]}{\sin\theta},\ x = 2\sqrt{2}\cos\theta.$$

This completes the proof of Proof of (2.29). □

LEMMA A.0.5 (Lemma 2.3, [**SB95**]). *Let P, Q be two rectangular matrices of the same size. Then for any $x, y \geq 0$,*

$$\mu_{(P+Q)(P+Q)^*}(x+y,\infty) \leq \mu_{PP^*}(x,\infty) + \mu_{QQ^*}(y,\infty)$$

PROOF. It follows directly from the Cauchy interlacing theorem. □

LEMMA A.0.6 (Sherman-Morrison formula). *Let $P_{n\times n}$ and $(P + vv^*)$ be invertible matrices, where $v \in \mathbb{C}^n$. Then we have*

$$(P+vv^*)^{-1} = P^{-1} - \frac{P^{-1}vv^*P^{-1}}{1+v^*P^{-1}v}.$$

In particular,

$$v^*(P+vv^*)^{-1} = \frac{v^*P^{-1}}{1+v^*P^{-1}v}.$$

PROOF. Using the resolvent identity, we have

$$P^{-1} - (P+vv^*)^{-1} = P^{-1}vv^*(P+vv^*)^{-1}.$$

Secondly since $v^*P^{-1}v$ is a scalar, we have

$$\begin{aligned}(v^*P^{-1}v)P^{-1}vv^*(P+vv^*)^{-1} &= P^{-1}v(v^*P^{-1}v)v^*(P+vv^*)^{-1} \\ &= (P^{-1}vv^*P^{-1})vv^*(P+vv^*)^{-1}\end{aligned}$$

$$= (P^{-1}vv^*P^{-1})[I - P(P + vv^*)^{-1}]$$

$$= P^{-1}vv^*P^{-1} - P^{-1}vv^*(P + vv^*)^{-1}.$$

Which implies that $(1 + v^*P^{-1}v)P^{-1}vv^*(P + vv^*)^{-1} = P^{-1}vv^*P^{-1}$. This completes the proof. □

LEMMA A.0.7 (Lemma 2.6, [**SB95**]). *Let P, Q be $n \times n$ matrices such that Q is Hermitian. Then for any $r \in \mathbb{C}^n$ and $z = E + i\eta \in \mathbb{C}^+$ we have*

$$\left| tr\left((Q - zI)^{-1} - (Q + rr^* - zI)^{-1}\right) P\right| = \left|\frac{r^*(Q - zI)^{-1}P(Q - zI)^{-1}r}{1 + r^*(Q - zI)^{-1}r}\right| \leq \frac{\|P\|}{\eta}.$$

PROOF. If C, D are two invertible matrices, then the resolvent identity tells us that $C^{-1} - D^{-1} = C^{-1}(D - C)D^{-1}$. Applying this result, we have

$$\begin{aligned}\operatorname{tr}\left((Q - zI)^{-1} - (Q + rr^* - zI)^{-1}\right) P &= \operatorname{tr}\left[(Q - zI)^{-1}rr^*(Q + rr^* - zI)^{-1}P\right] \\ &= r^*(Q + rr^* - zI)^{-1}P(Q - zI)^{-1}r.\end{aligned}$$

Once again using the resolvent identity, we have

$$\begin{aligned}&[1 + r^*(Q - zI)^{-1}r]r^*(Q + rr^* - zI)^{-1}P(Q - zI)^{-1}r \\ &= r^*(Q + rr^* - zI)^{-1}P(Q - zI)^{-1}r \\ &+ r^*(Q - zI)^{-1}rr^*(Q + rr^* - zI)^{-1}P(Q - zI)^{-1}r \\ &= r^*[(Q + rr^* - zI)^{-1} + (Q - zI)^{-1}rr^*(Q + rr^* - zI)^{-1}] \\ &\times P(Q - zI)^{-1}r \\ &= r^*(Q - zI)^{-1}P(Q - zI)^{-1}r.\end{aligned}$$

Combining the above two equations, we obtain

$$\left|\operatorname{tr}\left((Q - zI)^{-1} - (Q + rr^* - zI)^{-1}\right) P\right| = \left|\frac{r^*(Q - zI)^{-1}P(Q - zI)^{-1}r}{1 + r^*(Q - zI)^{-1}r}\right|.$$

Let $\{\lambda_i\}_{1\le i\le n}$ be the eigenvalues of Q, and $U = \sum_{i=1}^n u_i u_i^*$ be a matrix formed by the orthonormal eigenvectors of Q. Then

$$\begin{aligned}|r^*(Q-zI)^{-1}P(Q-zI)^{-1}r| &\le \|P\| \|(Q-zI)^{-1}r\|^2 \\ &= \|P\| \sum_{i=1}^n \frac{|u_i^* r|^2}{|\lambda_i - z|^2}.\end{aligned}$$

Secondly,

$$\begin{aligned}|1 + r^*(Q-zI)^{-1}r| &\ge |\Im(r^*(Q-zI)^{-1}r)| \\ &= \left|\Im\left(\sum_{i=1}^n \frac{|u_i^* r|^2}{\lambda_i - z}\right)\right| \\ &= \eta \sum_{i=1}^n \frac{|u_i^* r|}{|\lambda_i - z|^2}.\end{aligned}$$

Combining the last two estimates, we obtain

$$\left|\mathrm{tr}\left((Q-zI)^{-1} - (Q + rr^* - zI)^{-1}\right)P\right| = \left|\frac{r^*(Q-zI)^{-1}P(Q-zI)^{-1}r}{1 + r^*(Q-zI)^{-1}r}\right| \le \frac{\|P\|}{\eta}.$$

□

Lemma A.0.8 ([**Azu67**], Lemma 1). *Let $\{X_n\}_n$ be a sequence of random variables such that $|X_n| \le K_n$ almost surely, and $\mathbb{E}[X_{i_1}X_{i_2}\ldots X_{i_k}] = 0$ for all $k \in \mathbb{N}$, $i_1 < i_2 < \cdots < i_k$. Then for every $\lambda \in \mathbb{R}$ we have*

$$\mathbb{E}\left[\exp\left\{\lambda \sum_{i=1}^n X_i\right\}\right] \le \exp\left\{\frac{\lambda^2}{2}\sum_{i=1}^n K_i^2\right\}.$$

In particular, for any $t > 0$ we have

$$\mathbb{P}\left(\left|\sum_{i=1}^n X_i\right| > t\right) \le 2\exp\left\{-\frac{t^2}{2\sum_{i=1}^n K_i^2}\right\}.$$

Proof. Since e^x is a convex function, and $|X_i| \le K_i$, we have

$$\exp\{\lambda X_i\} = \exp\left\{\lambda\left(\frac{K_i + X_i}{2}\right) + (-\lambda)\left(\frac{K_i - X_i}{2}\right)\right\}$$

$$
\begin{aligned}
&= \exp\left\{\lambda K_i \left(\frac{1+X_i/K_i}{2}\right) + (-\lambda K_i)\left(\frac{1-X_i/K_i}{2}\right)\right\} \\
&\le \frac{1}{2}(1+X_i/K_i)\exp\{\lambda K_i\} + \frac{1}{2}(1-X_i/K_i)\exp\{-\lambda K_i\} \\
&= \cosh(\lambda K_i) + \frac{X_i}{K_i}\sinh(\lambda K_i) \\
&\le \exp\{\lambda^2 K_i^2/2\} + \frac{X_i}{K_i}\sinh(\lambda K_i) \quad \text{a.s.,}
\end{aligned}
$$

where the last inequality follows from the fact that $\cosh(x) \le \exp\{x^2/2\}$. Therefore using the fact $\mathbb{E}[X_{i_1}X_{i_2}\dots X_{i_k}] = 0$ for all $k \in \mathbb{N}$, $i_1 < i_2 < \cdots < i_k$, we have

$$
\mathbb{E}\left[\exp\left\{\lambda\sum_{i=1}^n X_i\right\}\right] \le \prod_{i=1}^n \exp\{\lambda^2 K_i^2/2\} = \exp\left\{\frac{\lambda^2}{2}\sum_{i=1}^n K_i^2\right\}.
$$

□

LEMMA A.0.9. *Let P, Q be two $n \times n$ matrices, then*

$$
\|\mu_{PP^*} - \mu_{QQ^*}\| \le \frac{2}{n} rank(P-Q).
$$

PROOF. By Cauchy's interlacing property,

$$
\begin{aligned}
\|\mu_{PP^*} - \mu_{QQ^*}\| &\le \frac{1}{n}\operatorname{rank}(PP^* - QQ^*) \\
&\le \frac{1}{n}\operatorname{rank}((P-Q)P^*) + \frac{1}{n}\operatorname{rank}(Q(P-Q)^*) \\
&\le \frac{2}{n}\operatorname{rank}(P-Q).
\end{aligned}
$$

□

LEMMA A.0.10 ([**BG13**], Lemma C.3). *Let P and Q be $n \times n$ Hermition matrices, and $I \subset \{1, 2, \dots, n\}$, then*

$$
\left|\sum_{k\in I}(P-zI)^{-1}_{kk} - \sum_{k\in I}(Q-zI)^{-1}_{kk}\right| \le \frac{2}{\Im(z)} rank(P-Q).
$$

PROOF. Using the resolvent identity, we have

$$
(P-zI)^{-1} - (Q-zI)^{-1} = (P-zI)^{-1}(Q-P)(Q-zI)^{-1}.
$$

Therefore $r := \text{rank}((P-zI)^{-1} - (Q-zI)^{-1}) \le \text{rank}(P-Q)$. Using the singular value decomposition, we can write

$$(P-zI)^{-1} - (Q-zI)^{-1} = \sum_{i=1}^{r} s_i u_i v_i^*,$$

where $\{u_1, \ldots, u_r\}$ and $\{v_1, \ldots, v_r\}$ are two sets of orthonormal vectors, and $s_1, \ldots, s_r$ are the at most r non zero singular values of $(P-zI)^{-1} - (Q-zI)^{-1}$. As a result, we can write

$$(P-zI)^{-1}_{kk} - (Q-zI)^{-1}_{kk} = \sum_{i=1}^{r} s_i \langle u_i, e_k\rangle \langle v_i, e_k\rangle.$$

Using Cauchy-Schwarz inequality,

$$\begin{aligned}
\sum_{k\in I}(P-zI)^{-1}_{kk} - \sum_{k\in I}(Q-zI)^{-1}_{kk} &= \sum_{i=1}^{r} s_i \sum_{k\in I} \langle u_i, e_k\rangle \langle v_i, e_k\rangle \\
&\le \sum_{i=1}^{r} s_i \sqrt{\sum_{k\in I} |\langle u_i, e_k\rangle|^2} \sqrt{\sum_{k\in I} |\langle v_i, e_k\rangle|^2} \\
&\le \sum_{i=1}^{r} s_i \le \frac{2r}{\Im(z)} \le \frac{2}{\Im(z)} \text{rank}(P-Q),
\end{aligned}$$

where the second last inequality follows from the fact that $s_i \le \|(P-zI)^{-1} - (Q-zI)^{-1}\| \le 2/\Im(z)$ for all $1 \le i \le r$. □

Lemma A.0.11. *Let C_j and B_j be same as defined in* (3.6), *r_j be the jth column of R, and $I_j \subset \{1, 2, \ldots, n\}$ be same as* (3.1). *Then*

$$\mathbb{P}\left(\left|\sum_{k\in I_j}(C_j^{-1})_{kk} - \mathbb{E}\sum_{k\in I_j}(C_j^{-1})_{kk}\right| > t\right) \le 2\exp\left\{-\frac{\Im(z)^2 t^2}{32n}\right\}$$

$$\mathbb{P}\left(\left|\sum_{k\in I_j}(C_j^{-1}B_j^{-1})_{kk} - \mathbb{E}\sum_{k\in I_j}(C_j^{-1}B_j^{-1})_{kk}\right| > t\right) \le 2\exp\left\{-\frac{\Im(z)^2 t^2}{32n}\right\}$$

$$\mathbb{P}\left(\left|\sum_{k\in I_j}(C_j^{-1}r_j r_j^* C_j^{-1*})_{kk} - \mathbb{E}\sum_{k\in I_j}(C_j^{-1}r_j r_j^* C_j^{-1*})_{kk}\right| > t\right) \le 2\exp\left\{-\frac{\Im(z)^2 t^2}{32n}\right\}$$

$$\mathbb{P}\left(\left|\sum_{k\in I_j}(C_j^{-1}B_j^{-1}r_j r_j^* B^{-1*} C_j^{-1*})_{kk} - \mathbb{E}\sum_{k\in I_j}(C_j^{-1}B_j^{-1}r_j r_j^* B^{-1*} C_j^{-1*})_{kk}\right| > t\right) \le 2\exp\left\{-\frac{\Im(z)^2 t^2}{32n}\right\}.$$

PROOF. Let $\mathcal{F}_l = \sigma\{y_1, \ldots, y_l\}$ be the σ-algebra generated by the column vectors $y_1, \ldots, y_l$. Then, we can write

$$\sum_{k\in I_j}(C_j^{-1})_{kk} - \mathbb{E}\sum_{k\in I_j}(C_j^{-1})_{kk} = \sum_{l=1}^{n}\left[\mathbb{E}\left\{\sum_{k\in I_j}(C_j^{-1})_{kk}\middle|\mathcal{F}_l\right\} - \mathbb{E}\left\{\sum_{k\in I_j}(C_j^{-1})_{kk}\middle|\mathcal{F}_{l-1}\right\}\right].$$

Notice that for any two matrices P, Q, we have $\mathrm{rank}(PP^* - QQ^*) \leq 2\mathrm{rank}(P - Q)$ (from Lemma A.0.9). Therefore, using the Lemma A.0.10 and Lemma A.0.8, we can conclude the result. The remaining three equations can also be proved in the same way. □

Bibliography

[AAR99] G. E. Andrews, R. Askey, and R. Roy, *Special functions, volume 71 of encyclopedia of mathematics and its applications*, 1999.

[AGZ10] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An introduction to random matrices*, no. 118, Cambridge University Press, 2010.

[Arn71] L. Arnold, *On wigner's semicircle law for the eigenvalues of random matrices*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **19** (1971), no. 3, 191–198 (English).

[Azu67] K. Azuma, *Weighted sums of certain dependent random variables*, Tohoku Mathematical Journal, Second Series **19** (1967), no. 3, 357–367.

[B$^+$97] Z. Bai et al., *Circular law*, The Annals of Probability **25** (1997), no. 1, 494–529.

[Bai99] Z. D. Bai, *Methodologies in spectral analysis of large dimensional random matrices, a review*, Statistica Sinica **9** (1999), no. 3, 611–662.

[BG13] C. Bordenave and A. Guionnet, *Localization and delocalization of eigenvectors for heavy-tailed random matrices*, Probability Theory and Related Fields **157** (2013), no. 3-4, 885–953.

[BMP91] L. V. Bogachev, S. A. Molchanov, and L. A. Pastur, *On the level density of random band matrices*, Mathematical Notes **50** (1991), no. 6, 1232–1242.

[BS10] Z. Bai and J. W. Silverstein, *Spectral analysis of large dimensional random matrices*, vol. 20, Springer, 2010.

[CIM91] G. Casati, F. Izrailev, and L. Molinari, *Scaling properties of the eigenvalue spacing distribution for band random matrices*, Journal of Physics A: Mathematical and General **24** (1991), no. 20, 4755.

[CMI90] G. Casati, L. Molinari, and F. Izrailev, *Scaling properties of band random matrices*, Physical review letters **64** (1990), no. 16, 1851–1854.

[Cra99] H. Cramér, *Mathematical methods of statistics*, vol. 9, Princeton University Press, 1999.

[DFJ68] S. Dharmadhikari, V. Fabian, and K. Jogdeo, *Bounds on the moments of martingales*, The Annals of Mathematical Statistics (1968), 1719–1723.

[DS07] R. B. Dozier and J. W. Silverstein, *On the empirical distribution of eigenvalues of large dimensional information-plus-noise-type matrices*, Journal of Multivariate Analysis **98** (2007), no. 4, 678–694.

[EKYY13] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin, *Delocalization and diffusion profile for random band matrices*, Communications in Mathematical Physics **323** (2013), no. 1, 367–416.

[Erd11] L. Erdős, *Universality of wigner random matrices: a survey of recent results*, Russian Mathematical Surveys **66** (2011), no. 3, 507.

[FM91] Y. V. Fyodorov and A. D. Mirlin, *Scaling properties of localization in random band matrices: a σ-model approach*, Physical review letters **67** (1991), no. 18, 2405.

[Gin65] J. Ginibre, *Statistical ensembles of complex, quaternion, and real matrices*, Journal of Mathematical Physics **6** (1965), no. 3, 440–449.

[Gir84] V. L. Girko, *Circular law*, Teoriya Veroyatnostei i ee Primeneniya **29** (1984), no. 4, 669–679.

[Gir85] ———, *Circular law*, Theory of Probability & Its Applications **29** (1985), no. 4, 694–706.

[Gir86] V. L. Girko, *Elliptic law*, Theory of Probability & Its Applications **30** (1986), no. 4, 677–690, http://dx.doi.org/10.1137/1130089.

[GZ01] A. Guionnet and B. Zegarlinksi, *Lectures on logarithmic sobolev inequalities*, Springer, 1801.

[JS16] I. Jana and A. Soshnikov, *Esd of singular values of random band matrices; marchenko-pastur law and more*, arXiv preprint arXiv:1610.02153 (2016).

[JSS16] I. Jana, K. Saha, and A. Soshnikov, *Fluctuations of linear eigenvalue statistics of random band matrices*, Theory of Probability & Its Applications **60** (2016), no. 3, 407–443, http://dx.doi.org/10.1137/S0040585X97T987788.

[Lie73] E. H. Lieb, *Convex trace functions and the wigner-yanase-dyson conjecture*, Advances in Mathematics **11** (1973), no. 3, 267–288.

[LP09] A. Lytova and L. Pastur, *Central limit theorem for linear eigenvalue statistics of random matrices with independent entries*, The Annals of Probability **37** (2009), no. 5, 1778–1840.

[LS13] L. Li and A. Soshnikov, *Central limit theorem for linear statistics of eigenvalues of band random matrices*, Random Matrices: Theory and Applications **2** (2013), no. 04.

[Meh67] M. Mehta, *Random matrices and the statistical theory of energy levels, acad*, Press, New York (1967).

[MFD+96] A. D. Mirlin, Y. V. Fyodorov, F.-M. Dittes, J. Quezada, and T. H. Seligman, *Transition from localized to extended eigenstates in the ensemble of power-law random banded matrices*, Physical Review E **54** (1996), no. 4, 3221.

[MJC+14] L. Mackey, M. I. Jordan, R. Y. Chen, B. Farrell, J. A. Tropp, et al., *Matrix concentration inequalities via the method of exchangeable pairs*, The Annals of Probability **42** (2014), no. 3, 906–945.

[Mon73] H. L. Montgomery, *The pair correlation of zeros of the zeta function*, Proc. Symp. Pure Math, vol. 24, 1973, pp. 181–193.

[MP67] V. A. Marčenko and L. A. Pastur, *Distribution of eigenvalues for some sets of random matrices*, Mathematics of the USSR-Sbornik **1** (1967), no. 4, 457.

[MPK92] S. A. Molchanov, L. A. Pastur, and A. Khorunzhii, *Limiting eigenvalue distribution for band random matrices*, Theoretical and Mathematical Physics **90** (1992), no. 2, 108–118.

[Nau12] A. Naumov, *Elliptic law for real random matrices*, arXiv preprint arXiv:1201.1639 (2012).

[NO14] H. H. Nguyen and S. O'Rourke, *The elliptic law*, International Mathematics Research Notices (2014), rnu174.

[ORS13] S. O'Rourke, D. Renfrew, and A. Soshnikov, *On fluctuations of matrix entries of regular functions of wigner matrices with non-identically distributed entries*, Journal of Theoretical Probability **26** (2013), no. 3, 750–780.

[OS17] S. Olver and A. Swan, *Evidence of the poisson/gaudin–mehta phase transition for banded matrices on global scales*, arXiv preprint arXiv:1703.06985 (2017).

[PRS12] A. Pizzo, D. Renfrew, and A. Soshnikov, *Fluctuations of matrix entries of regular functions of wigner matrices*, Journal of Statistical Physics **146** (2012), no. 3, 550–591.

[SB95] J. W. Silverstein and Z. Bai, *On the empirical distribution of eigenvalues of a class of large dimensional random matrices*, Journal of Multivariate analysis **54** (1995), no. 2, 175–192.

[Sch09] J. Schenker, *Eigenvector localization for random band matrices with power law band width*, Communications in Mathematical Physics **290** (2009), no. 3, 1065–1097.

[Shc11] M. Shcherbina, *Central limit theorem for linear eigenvalue statistics of the wigner and sample covariance random matrices*, Journal of Mathematical Physics, Analysis, Geometry **7** (2011), no. 2, 176–192.

[Shc15] M. Shcherbina, *On fluctuations of eigenvalues of random band matrices*, Journal of Statistical Physics **161** (2015), no. 1, 73–90.

[Sil95] J. W. Silverstein, *Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices*, Journal of Multivariate Analysis **55** (1995), no. 2, 331–339.

[Tro12] J. A. Tropp, *User-friendly tail bounds for sums of random matrices*, Foundations of computational mathematics **12** (2012), no. 4, 389–434.

[Tro15] ———, *An introduction to matrix concentration inequalities*, arXiv preprint arXiv:1501.01571 (2015).

[TV04] A. M. Tulino and S. Verdú, *Random matrix theory and wireless communications*, vol. 1, Now Publishers Inc, 2004.

[TV08] T. Tao and V. Vu, *Random matrices: the circular law*, Communications in Contemporary Mathematics **10** (2008), no. 02, 261–307.

[TVK10] T. Tao, V. Vu, and M. Krishnapur, *Random matrices: universality of esds and the circular law*, The Annals of Probability **38** (2010), no. 5, 2023–2065.

[VS99] S. Verdú and S. Shamai, *Spectral efficiency of cdma with random spreading*, Information Theory, IEEE Transactions on **45** (1999), no. 2, 622–640.

[Wig55] E. P. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions*, Annals of Mathematics **62** (1955), no. 3, pp. 548–564 (English).

[Wig57] ———, *Characteristics vectors of bordered matrices with infinite dimensions ii*, Annals of Mathematics **65** (1957), no. 2, pp. 203–207 (English).

[Wig58] ———, *On the distribution of the roots of certain symmetric matrices*, Annals of Mathematics **67** (1958), no. 2, pp. 325–327 (English).

[Wis28] J. Wishart, *The generalised product moment distribution in samples from a normal multivariate population*, Biometrika (1928), 32–52.