Limiting spectral distribution of some patterned random matrices with independent entries

Priyanka Sen

Indian Statistical Institute

August 2023

Indian Statistical Institute

Doctoral Thesis

Limiting spectral distribution of some patterned random matrices with independent entries

---

Author:
Priyanka Sen

Supervisor:
Arup Bose

A thesis submitted to the Indian Statistical Institute
in partial fulfilment of the requirements for
the degree of
Doctor of Philosophy (in Mathematics)

Theoretical Statistics & Mathematics Unit

Indian Statistical Institute, Kolkata

August 2023

Dedicated to Baba

Acknowledgements

Undertaking this PhD has been a life-changing experience for me and it would not have been possible to do so without the help and advice from many people.

First and foremost, I would like to thank my thesis supervisor, Prof. Arup Bose for his patient guidance, encouragement and advice. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my queries so promptly. It is beyond words how his mere presence and attitude has evolved my perspective on several aspects of life.

I have been very fortunate to work with Prof. Koushik Saha and Prof. Arusharka Sen. I am thankful to them for allowing me to include our joint works in this thesis. I am immensely grateful to Prof. Koushik Saha for pointing me towards the direction of the problems tackled in thesis, and also his continuous support and encouragement during our joint works. The times I have spent at IIT, Bombay (in Sept, 2019 and Jan, 2020) have been the most formulative experiences of my life. Here, I would also like to mention Prof. Monika Bhattacharjee, with whom I got the opportunity to interact with both academically and informally while at IIT, Bombay and later in various occasions. Not only did she inspire me to try different problems in innovative ways, but also helped me see through some of my insecurities during this PhD venture. I am very grateful to her for being more like an elder sister to me.

I would like to thank Prof. Rudra Prasad Sarkar, Prof. Neena Gupta, Prof. Krishanu Maulik, Prof. Kingshook Biswas, Prof. Shashi Mohan Srivastava, Prof. Alok Goswami, Prof. Debashish Goswami, Prof. Mrinal Kanti Das, Prof. Mahuya Datta, Prof. Rajat Subhra Hazra and Prof. Antar Bandyopadhyay for teaching me various topics at ISI, Kolkata. I owe special thanks to Prof. Krishanu Maulik, Prof. Rajat Subhra Hazra and Prof. Antar Bandyopadhyay for introducing me to the world of probability and their encouragement towards pursuing research in probability. I would also like to express my gratitude to the wonderful teachers I had during my MSc at IIT, Guwahati. Here, I want to make special mention of Prof. Anjan K. Chakrabarty and Prof. Jitendriya Swain for being so open and welcoming in discussing different topics in mathematics. Thanks should also go to all my teachers during BSc, high school and primary education for their love and support.

One of the best parts of being a member of Stat-Math unit, ISI, Kolkata, is the group of people the unit constitutes of. Besides all the splendid professors, the unit hosts a magnificent group of research scholars and associates. I consider myself fortunate enough to have interacted with most of them and getting their point of view in several situations.

Here I would like to thank Biltu da, Sukrit da, Sumit da, Gopal da, Mallesham da for their invaluable advice and fun-filled conversations. Thanks to my batchmates, Mainak, Aritra, Debanil, Aparajita di and Anurag for being in the same boat with me and making the best of the journey together. I am also grateful for my juniors Parnashree, Bilol, Deborshi, Bodhisatta, Bishakh and Subhodeep for being such a great bunch of people and making my PhD journey much enjoyable. All of these people have put a smile on my face on the bad days of the journey. However, I owe the greatest thanks to Sukrit da in this regard, as he has always kept his door open and has lent me the pair of ears I needed in the most difficult times.

Most of my PhD life constituted of being in the hostel. So life would be incomplete without fellow hostellers who helped keep each others lives stress free and sane. I thank Bibhuti,Sankar, Mainak, Rakesh, Anjan, Sushanta, Sandeep, Manish, Joginder for making the hostel life so delightful. My sincere thanks to Bibhuti, Rakesh, Sankar and Mainak for all the long discussions we had about career and life in general. Here I would also like to express my heartfelt gratitude towards my very dear friends, Malay and Rituparna for their unconditional support and for keeping me sane throughout the whole process.

I would like to express my deepest gratitude to my parents for all their love, support and all the sacrifices they have made for me. My brother, although much younger than me, has been one of my biggest cheerleaders and companion. Finally, I want to thank my husband and best friend, Akash. He is my most enthusiastic cheerleader and without his sunny optimism, my days would have felt much longer and tiresome. This endeavour would have lost all its meaning without these people in my life. I am forever indebted to them.

Priyanka Sen
December 2022

Contents

Notations & Abbreviations

$\mathbb{N}$	The set of natural numbers
$\mathbb{Z}$	The set of integers
$\mathbb{R}$	The set of real numbers
$\mathbb{C}$	The set of complex numbers
$\operatorname{Tr} A$	Trace of matrix A
A^T	Transpose of a matrix A
$a_n \approx b_n$	$\frac{a_n}{b_n} \to 1$ as $n \to \infty$
$\lvert A \rvert$	Cardinality of the set A
$[k]$	The set $\{1, 2, \ldots, k\}$
$\mathcal{P}(k)$	Set of all partitions of $\{1, 2, \ldots, k\}$
$\lfloor a \rfloor$	Greatest integer less or equal to a
$F_n \xrightarrow{\mathcal{D}} F$	Distribution function F_n converges to F weakly (in distribution sense)
$X \sim F$	Random variable X has distribution F
$X \stackrel{\mathcal{D}}{=} Y$	Random variables X and Y have the same distribution
$\beta_k(F)$	kth moment of F
$\beta_k(\mu)$	kth moment of μ
$\lvert \sigma \rvert$	Number of blocks of a partition σ
F^A	ESD of the matrix A
$S(k)$	Set of all symmetric partitions of $\{1, 2, \ldots, k\}$
$S_b(k)$	Set of all symmetric partitions of $\{1, 2, \ldots, k\}$ with b many blocks
$E(k)$	Set of all even partitions of $\{1, 2, \ldots, k\}$
$E_b(k)$	Set of all even partitions of $\{1, 2, \ldots, k\}$ with b many blocks
$NC(k)$	Set of all non-crossing partitions of $\{1, 2, \ldots, k\}$
$NCE(k)$	Set of all non-crossing even partitions of $\{1, 2, \ldots, k\}$
$NC_2(k)$	Set of all non-crossing pair partitions of $\{1, 2, \ldots, k\}$

Chapter 1

Introduction

Random matrix theory (RMT) is the study of spectral properties of large dimensional random matrices. Data sets where both the dimension of the observations and the sample size is large, arise naturally in contemporary research, such as, genome sequence data in biology, online networks, financial portfolios, wireless communication etc. Johnstone and Titterington [2009] provides a few statistical and probabilistic challenges in dealing with such large data sets. RMT provides an efficient and tractable way to study and interpret these huge amounts of data. Also the theory has remarkable applications in nuclear physics and number theory. See Katz and Sarnak [1999], Conrey [2001], Keating and Snaith [2003], Firk and Miller [2009], Barrett et al. [2016]. More applications of RMT can be found in Mehta [2004], Forrester [2010], Akemann et al. [2011], Tao [2012], etc.

A random matrix is a matrix whose entries are random variables. Naturally the eigenvalues of random matrices are random too. Important information about these matrices are encoded in their eigenvalues. Thus, the study of various properties of the eigenvalues is at the heart of RMT. Through the last few decades, researchers have raised several questions about the asymptotic behaviour of the eigenvalues of random matrices. Some of them include, the maximum and minimum eigenvalue distribution, the local law, the bulk distribution, asymptotic results for different eigenvalue statistics, spectral (gap) statistics, the universality problem etc. In this thesis, we focus on the existence of the bulk distributions for various new models.

1.1 Empirical Spectral Distribution

Suppose A_n is an $n \times n$ real symmetric random matrix whose elements are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let the eigenvalues of A_n be $\lambda_1, \lambda_2, \ldots, \lambda_n$, which are real as A_n is symmetric. The *empirical spectral measure* of A_n, denoted by μ_{A_n}, is as follows:

$$\mu_{A_n} = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i},$$

where δ_x is the Dirac measure at x. The probability distribution function, F^{A_n}, known as the *empirical spectral distribution* (ESD) of A_n, is given by

$$F^{A_n}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(\lambda_i \leq x).$$

Clearly $F^{A_n}(\cdot)$ is a random distribution function and is a function of $\omega \in \Omega$. However, for convenience we often suppress this dependence. The expected empirical spectral distribution (EESD), denoted by $\mathbb{E}F^{A_n}$,

$$\mathbb{E}F^{A_n}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{P}(\lambda_i \leq x),$$

is also a distribution function and is non-random.

Definition 1.1.1. *(Convergence of ESD and EESD)*

(a) $\{F^{A_n}(\cdot)\}_n$ is said to converge weakly almost surely if there exists a distribution function $F(\cdot)$ such that for almost every $\omega \in \Omega$ (i.e., outside a null set) and for all continuity points x of F,

$$F^{A_n}(x) \to F(x) \quad \text{as } n \to \infty.$$

(b) The ESD of A_n converges to F in probability if for each $\epsilon > 0$, and at every continuity point x of F,

$$\mathbb{P}\big[|F^{A_n}(x) - F(x)| > \epsilon\big] \to 0 \quad \text{as } n \to \infty.$$

(c) The EESD of A_n converges weakly to $\tilde{F}$ if for all continuity points x of $\tilde{F}$,

$$\mathbb{E}F^{A_n}(x) \to \tilde{F}(x) \quad \textit{as } n \to \infty.$$

From this definition, we can clearly see, $(a) \Rightarrow (b)$. Further, the limiting distribution $\tilde{F}$ is always non-random as $\mathbb{E}F^{A_n}$ is so. The limits F and $\tilde{F}$ are identical when F is non-random. In general, F can be random.

In any case, any of these limits will be referred to as the *limiting spectral distribution* (LSD) of $\{A_n\}$ in this thesis.

Next, we describe the random matrices that we are going to study in this thesis and the questions we address regarding their LSDs.

1.2 Some symmetric patterned matrices of interest

We shall consider the Wigner, reverse circulant matrix, the symmetric circulant matrix, the symmetric Toeplitz matrix and the Hankel matrix where the entries of the matrices are independent for each fixed n but not necessarily identically distributed. Also we shall look into the LSD of $M_p M_p^T$ where M_p ($p \times n$ rectangular matrix) is one of the symmetric or asymetric versions of the aforementioned matrices. We shall also look at several modifications of these matrices.

Wigner matrix: An $n \times n$ Wigner matrix is defined as

$$W_n = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1(n-1)} & x_{1n} \\ x_{12} & x_{22} & x_{23} & \cdots & x_{2(n-1)} & x_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{1n} & x_{2n} & x_{3n} & \cdots & x_{(n-1)n} & x_{nn} \end{bmatrix}.$$

The matrix is symmetric and the (i, j)-th element of the matrix is x_{ij} $(i \leq j)$.

Reverse circulant matrix: An $n \times n$ reverse circulant matrix is defined as

$$R_n^{(s)} = \begin{pmatrix} x_0 & x_1 & x_2 & \cdots & x_{n-1} \\ x_1 & x_2 & x_3 & \cdots & x_0 \\ x_2 & x_3 & x_4 & \cdots & x_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n-1} & x_0 & x_1 & \cdots & x_{n-2} \end{pmatrix}.$$

For $1 \leq j \leq n-1$, its $(j+1)$-th row is obtained by giving its j-th row a left circular shift by one position. The matrix is symmetric and the (i, j)-th element of the matrix is $x_{(i+j-2) \bmod n}$.

Symmetric circulant matrix: An $n \times n$ symmetric circulant matrix is defined by

$$C_n^{(s)} = \begin{pmatrix} x_0 & x_1 & x_2 & \cdots & x_1 \\ x_1 & x_0 & x_1 & \cdots & x_2 \\ x_2 & x_1 & x_0 & \cdots & x_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_0 \end{pmatrix}.$$

The $(i,\ j)$-th element of the matrix is $x_{\frac{n}{2}-|\frac{n}{2}-|i-j||}$. Observe that for $j = 1, 2, \ldots, n-1$, its $(j+1)$-th row is obtained by giving its j-th row a right circular shift by one position.

Symmetric Toeplitz matrix: The symmetric Toeplitz matrix is defined as

$$T_n^{(s)} = \begin{pmatrix} x_0 & x_1 & x_2 & \cdots & x_{n-1} \\ x_1 & x_0 & x_1 & \cdots & x_{n-2} \\ x_2 & x_1 & x_0 & \cdots & x_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n-1} & x_{n-2} & x_{n-3} & \cdots & x_0 \end{pmatrix}.$$

The (i, j)-th element of the matrix is $x_{|i-j|}$. Note that the symmetric circulant matrix is a Toeplitz matrix with the added restriction that $x_{n-j} = x_j$.

Hankel matrix : An $n \times n$ Hankel matrix is defined as

$$H_n^{(s)} = \begin{pmatrix} x_2 & x_3 & x_4 & \cdots & x_{n+1} \\ x_3 & x_4 & x_5 & \cdots & x_{n+2} \\ x_4 & x_5 & x_6 & \cdots & x_{n+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n+1} & x_{n+2} & x_{n+3} & \cdots & x_{2n} \end{pmatrix}.$$

The (i, j)-th element of the matrix is $x_{(i+j)}$.

Now we describe the asymmetric ($p \times n$ rectangular) versions of the aforesaid matrices.

$$X_p = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2n} \\ x_{31} & x_{32} & x_{33} & \cdots & x_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & x_{p3} & \cdots & x_{pn} \end{bmatrix},$$

$$R_p = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_{n-1} \\ x_{-1} & x_2 & x_3 & \cdots & x_0 \\ x_{-2} & x_{-3} & x_4 & \cdots & x_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{-(p-1)\text{mod } n} & \cdots & & \cdots & x_{(p-2)\text{mod } n} \end{bmatrix},$$

$$C_p = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_{n-1} \\ x_1 & x_0 & x_1 & \cdots & x_{n-2} \\ x_2 & x_1 & x_0 & \cdots & x_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{(1-p)(\text{mod } n)} & \cdots & \cdots & \cdots & x_{(n-p)(\text{mod } n)} \end{bmatrix},$$

$$T_p = \begin{bmatrix} x_0 & x_{-1} & x_{-2} & \cdots & x_{1-n} \\ x_1 & x_0 & x_{-1} & \cdots & x_{2-n} \\ x_2 & x_1 & x_0 & \cdots & x_{3-n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{p-1} & x_{p-2} & x_{p-3} & \cdots & x_{p-n} \end{bmatrix},$$

$$H_p = \begin{bmatrix} x_2 & x_{-3} & x_{-4} & \cdots & x_{-(n+1)} \\ x_3 & x_4 & x_{-5} & \cdots & x_{-(n+2)} \\ x_4 & x_5 & x_6 & \cdots & x_{-(n+3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{p+1} & x_{p+2} & x_{p+3} & \cdots & x_{-(p+n)} \end{bmatrix}.$$

1.3 Motivation

Wigner matrix: The Wigner matrix, W_n, is one of the matrices that have been studied extensively since its emergence in Wigner [1955] in the study of statistical theory of energy levels of heavy nuclei. He began the study of these matrices with i.i.d. Gaussian entries, and established the convergence of the EESD of $\frac{1}{\sqrt{n}}W_n$ to the semi-circle distribution. Wigner [1958] relaxed some conditions on the entries of the matrix and proved the convergence of the EESD. The new sufficient conditions included independence of the entries, common variance and uniformly bounded moments. In Mehta and Gaudin [1960], the authors prescribed an exact expression for the joint probability of eigenvalues of the Wigner matrix with entries that are randomly and independently distributed with distributions invariant under unitary transformation. In the following decade, various subsequent works such as Grenander [1968], Arnold [1967], Arnold [1971], strengthened the notion of convergence from EESD to ESD in probability and in almost sure senses. Pastur [1972] proved the almost sure convergence of the ESD of $\frac{1}{\sqrt{n}}W_n$ with independent entries, common mean and variance under a special condition, called the *Lindeberg* condition. Bai [1999] provided a couple of further extensions of the Wigner matrix and proved the convergence of the ESD of $\frac{1}{\sqrt{n}}W_n$ to the semi-circle distribution almost surely. One direction of research about the LSD of the Wigner matrix constitutes relaxing further conditions on the entries of the matrix by letting the distribution of an entry to depend on its position in the matrix, and investigating the limiting distributions thereafter. Some results in this regard include Girko et al. [1994], Anderson and Zeitouni [2006], Erdős et al. [2012] and more recently Jin and Xie [2020], Zhu [2020]. These articles consider unequal variances of the entries (that depend on the position of the entries but remain unaltered with the size of the matrix) with different assumptions on the variances. Anderson and Zeitouni [2006] considered a *continuous variance profile* and Erdős et al. [2012] considered *discrete variance profiles* with the variances being *row stochastic*. In Jin and Xie [2020], Zhu [2020], the row stochastic condition of the variances is relaxed and the convergence of the ESD is shown under different assumptions on the variance profile. As the distribution of the (i, j)th entry now depends on i, j, the limiting distribution is often not the semi-circle distribution. Some more results that show limits beyond the semi-circle distribution, include universality of local laws as in Erdős et al. [2010], Tao and Vu [2011] and Ajanki et al. [2017]. However, in this thesis we will not discuss the local laws and will keep our focus on the LSD results.

Another problem of interest in the study of the LSD of a Wigner matrix is when the entries of the matrix have infinite moments or moments with exponential growth. In particular, Zakharevich [2006] studied the Wigner matrix with i.i.d. entries whose distribution is G_n for every fixed n, such that G_n has finite moments of all orders, and satisfies certain moment conditions (see Result 3.1.3). For these matrices, though each distribution G_n is light-tailed, as $n \to \infty$, G_n may converge to a heavy-tailed distribution. That leads to the problem of finding the LSD for matrices whose entries are i.i.d. with heavy tails. This problem was studied by Ben Arous and Guionnet [2008]. Another study that dealt with matrices whose entries have distribution dependent on the size of the matrix is Jung [2018], where the LSD of *Lévy-Khintchine random matrices* are investigated. The common theme in all of these works is that the entries are i.i.d. for every n and dependent on n, but the distribution does not alter with the position of the entries in the matrix. This theme is also present in yet another famous problem: the adjacency matrix of the Erdős- Rényi graphs. Introduced by Gilbert [1959], this model attracted huge attention after the celebrated work of Erdős and Rényi [1960] and hence the name of the graph. This is basically an undirected graph with n vertices in which every edge occurs with some probability, say, p_n, independent of the other edges. This model has been studied through the decades in various regimes, the two most popular ones being the *sparse regime*, when $np_n \to \lambda > 0$, and the *dense regime*, when

$np_n \to \infty$. For a detailed exposition, see Guionnet [2021]. Significant developments in studying the LSD of the adjacency matrix of a sparse Erdős- Rényi graph began with Bauer and Golinelli [2001]. Subsequent works by Khorunzhy et al. [2004], Liang et al. [2007], Bordenave et al. [2011] followed, finally culminating in the almost sure convergence of the ESD. Another variation of the Erdős- Rényi graphs is when the probability (p_n) of an edge being present or absent is dependent on which vertices that edge connects to. In that case, the graph is *inhomogeneous*. Some interesting works in this area include Bollobás et al. [2007], Benaych-Georges et al. [2019], Chakrabarty et al. [2022] and others. In this thesis, we will study the LSD of the adjacency matrix of some *sparse inhomogeneous* Erdős- Rényi graph.

Some more structured (subject to symmetry) models of the Wigner matrix include band, block and triangular matrices, and have been studied over the last few decades by Casati and Girko [1993a], Casati and Girko [1993b], Shlyakhtenko [1996], Anderson and Zeitouni [2006], Molchanov et al. [1992], Bolla [2004], Rashidi Far et al. [2008], Ding [2014], Zhu [2020] and Basu et al. [2012]. The common theme of all these results is that the distribution of the entries of the matrix is dependent on the position of the entries in the matrix.

The results mentioned above have been proved using different techniques in RMT, and are often very specific to the matrix model under consideration. Also the question about what happens to the LSD when the distribution of the entries depend on the position of the entries as well as the size of the matrix, remained unanswered generally. In this thesis, we address this question by considering a general model, and show that under certain moment conditions and independence of the entries, the almost sure convergence of the ESD of W_n holds (see Theorem 3.3.1). Also we show that most, if not all, results mentioned above can be brought under a common umbrella through our prescribed model. We describe the limiting moments via partitions, and find a new set of partitions called *special symmetric partitions* (see Section 3.2) that contribute positively to the limiting moments.

Sample Covariance matrix: The *Sample covariance matrix* S, is arguably one of the most important matrices in RMT. Suppose X_p is a $p \times n$ rectangular matrix with real entries $\{x_{ij,n} : 1 \leq i \leq p, 1 \leq j \leq n\}$, where $p = p(n)$. Then $S = X_p X_p^T$ is called the (unadjusted) Sample covariance matrix. There are two regimes that are generally considered in this case: $p/n \to y \in (0, \infty)$ as $n \to \infty$, i.e., p and n are comparable for large n and $p/n \to 0$. In this thesis, we focus on the first regime.

If the rows of X_p are i.i.d. Gaussian, then S is called a Wishart matrix. In this case the joint distribution of the eigenvalues was identified in Anderson [1963]. The first success in finding the LSD of $\frac{1}{n}S$ was by Marčenko and Pastur [1967] who found the limit, and which is now called the Marčenko-Pastur (MP) distribution. Further extensions of this result continued in the following decades by Grenander and Silverstein [1977], Wachter [1978], Yin [1986], Jonsson [1982], Bai [1999]. Bose and Sen [2008] gave a proof of the almost sure convergence of the ESD of $\frac{1}{n}S$, under the assumption that the entries are i.i.d. with second moment finite, via the moment method. Yin [1986], Lytova and Pastur [2009], Bai and Silverstein [2010], Hachem et al. [2006], Zhu [2020], Jin and Xie [2020] studied the ESD of $\frac{1}{n}S$ where the entries of X_p have independent distributions with unequal variances, with different moment conditions. However, these distributions do not alter with the size of the matrix.

LSD of sample covariance matrices where the distribution of entries of X_p are i.i.d. with heavy tails was studied in Belinschi et al. [2009]. They proved the almost sure convergence of the ESD of an appropriately scaled S matrix to a non-random heavy-tailed distribution, using truncation techniques. Thus just like the Wigner case, in this case too researchers became keen on studying the LSD when the distribution of the entries of X_p depend on the size of the matrix. Such results are found in Benaych-Georges and Cabanal-Duvillard [2012], Male [2017], Noiry [2018] and Zitelli [2022]. These articles spelt out the moments of the LSD via different techniques that involved various graphical and combinatorial structures.

As in the Wigner case, we consider the LSD of sample covariance matrices when the distribution of the entries depend on the position of the entries as well as the size of the matrix. We prove that under certain moment conditions and independence of the entries, the almost sure convergence of the ESD holds (see Theorem 5.2.1). We describe the limiting moments via the *special symmetric partitions*, that also appeared in case of the Wigner matrix. Also, we look into the relation between the LSD of the S matrix and the Wigner matrix when $p = n$. Many new results for S matrices with variance profile, band and triangular matrices follow as special cases of our theorem. These also make most, if not all, of the above results special cases of our theorem.

Other symmetric patterned matrices: Wigner and sample covariance matrices are definitely two of the most extensively studied matrices in RMT. However, some patterned matrices with more structures have also received considerable attention in the past few decades. These include the Toeplitz, Hankel and circulant matrices.

Any matrix of the form $((t_{i-j}))_{1\leq i,j\leq n}$ is called a Toeplitz matrix. It is symmetric when $t_{-k} = t_k$ for all k. When the dimension of the matrix is ∞ and $\sum_{k=1}^{\infty} |t_k|^2 < \infty$, then we get the Toeplitz operator on l^2, the space of square summable sequences. This non-random operator is well studied in mathematics. From the famous theorem of Grenander and Szegő [1984], we know that the LSD exists. The Toeplitz pattern appears in various places, such as in stationary processes, time series, and harmonic analysis. The *circulant* matrix plays a crucial role in studying the non-random Toeplitz operators and Toeplitz matrices of large dimension, see Grenander and Szegő [1984], and Gray [2006].

Another patterned matrix closely related to the Toeplitz matrix is the Hankel matrix. Any matrix of the form $((t_{i+j}))_{1\leq i,j\leq n}$ is called a Hankel matrix. Like the Toeplitz matrix, when the dimension of the matrix is ∞, under the assumption of square summability of $(t_k)_{k\geq 1}$, this defines the Hankel operator on l^2. The Hankel matrix finds application in time series, the Hamburger moment problem, and several other combinatorial problems.

The problem of finding the LSD of random Toeplitz and Hankel matrices was first proposed in Bai [1999]. As noticed before, the circulant matrices played a crucial role in understanding the behaviour of the non-random Toeplitz matrix. Thus one of the motivations to study the LSD of circulant matrices is to understand the same for Toeplitz and Hankel matrices. In this thesis, we will look at two specific models of the circulant matrices– the symmetric reverse circulant matrix, $R_n^{(s)}$ and the symmetric circulant matrix, $C_n^{(s)}$. The symmetric reverse circulant matrix first appeared in Bose and Mitra [2002] where the authors proved the convergence (in probability) of the ESD of $\frac{1}{\sqrt{n}}R_n^{(s)}$ with i.i.d. entries that have finite third moment. Then Hammond and Miller [2005] and Bryc et al. [2006] independently proved the almost sure convergence of $\frac{1}{\sqrt{n}}T_n^{(s)}$ under the assumption that the entries are i.i.d. with finite variance. Under the same assumptions, Bryc et al. [2006] also proved the almost sure convergence of the ESD of $\frac{1}{\sqrt{n}}H_n^{(s)}$. In Bose and Sen [2008], the authors revisited the problem of convergence of the ESDs of $\frac{R_n^{(s)}}{\sqrt{n}}, \frac{T_n^{(s)}}{\sqrt{n}}, \frac{H_n^{(s)}}{\sqrt{n}}$, and proved the almost sure convergence of the scaled matrices with scaling $\frac{1}{\sqrt{n}}$ using a unified treatment. Bose and Sen [2008] also studied the ESD of $\frac{1}{\sqrt{n}}C_n^{(s)}$ and proved that it converges weakly almost surely to the normal distribution. A common theme of all these results is that the entries of the matrices are i.i.d. with common mean and variance, or independent with common mean, variance and all other moments uniformly bounded. Some other results that added more structure to these patterned matrices included band matrices, triangular matrices and matrices with certain features such as the rows being palindromic, etc. See Kargin [2009], Basak and Bose [2011], Liu and Wang [2011], Popescu [2009], Basu et al. [2012], Jackson et al. [2012], Kologlu et al. [2013], Blackwell et al. [2021], Chen et al. [2021].

LSD of patterned matrices, $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$ when the entries are i.i.d. but the distribution changes with the size of the matrix, have not been studied as extensively. As we saw in case of the Wigner and the sample covariance matrices, one of the examples of such matrices are the sparse matrices. The LSD of sparse patterned matrices $(R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)})$ have been studied in Banerjee and Bose [2017] where the convergence of the EESD was proved. In this case, the weak limit of the ESDs are random. This is one of the major differences of these structured matrices from the Wigner matrix. For the sparse Wigner matrix, we know that the almost sure convergence of the ESD to a non-random probability distribution occurs. Another such example where the limit is random is when the entries of the matrices are heavy-tailed. In case of the Wigner matrix, the limiting distribution is symmetric and heavy-tailed, see Ben Arous and Guionnet [2008]. However, in case of the reverse circulant and symmetric circulant matrices with heavy-tailed inputs, Bose et al. [2011a] proved that the LSDs are random. We will see such contrast in LSD results between the Wigner matrix and $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$ throughout this thesis.

Singular values of patterned matrices: All the patterned matrices we have discussed so far are symmetric. But there are asymmetric rectangular $(p \times n)$ versions of these matrices that occur in several areas. A typical way of studying rectangular versions of these matrices is via their singular values. Suppose A_p is one of the eight rectangular matrices $R_p^{(s)}, R_p, C_p^{(s)}, C_p, T_p^{(s)}, T_p, H_p^{(s)}, H_p$. Bose et al. [2010] studied the ESD of $\frac{1}{n}S_A = \frac{1}{n}A_pA_p^T$ and concluded its almost sure convergence as $p/n \to y \in (0, \infty)$, under the assumption that the entries of the matrix A_p are independent with mean zero variance 1, and are either uniformly bounded or are identically distributed.

Similar to the Wigner and the sample covariance matrices, the question of what happens to the LSD of the reverse circulant, symmetric circulant, Toeplitz and Hankel matrices with independent entries where the distribution of an entry varies with the position of the entry and the size of the matrix, remained unanswered. We address this question and describe a general setting in which the distribution of the (i, j)th entry of $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$ can depend on i, j and n.

We show in Theorems 4.2.2–4.2.4 that the EESD of each of these matrices converges weakly to symmetric probability distributions under suitable conditions. Also we show that most, if not all, of the LSD results for these matrices mentioned above follow from our theorems. We describe the limiting moments via partitions and also relate the LSDs to the notion of independence and half independence in certain cases.

We also relax the assumptions for the rectangular asymmetric versions of these matrices and allow the distribution of the (i, j)th entry of A_p to depend on i, j and n. We show that under appropriate assumptions, the ESD of $S_A = A_p A_p^T$ converges (see Theorem 6.0.1). Further, most of the existing results follow from our Theorem 6.0.1. This theorem also gives rise to several new LSD results for matrices with variance profile, sparse, band and triangular matrices.

Remark 1.3.1. *In this thesis, we use the moment method to conclude the convergence of the ESDs and EESDs. However, using the moment method one cannot find the ESD of non-symmetric matrices. Finding the LSD of a non-symmetric random matrix is a much harder problem that requires more sophisticated mathematical tools. These tools have taken a long time to come to light and are still mostly applicable to particular matrices. The challenge posed by the proof of the convergence of the ESD of the $n \times n$ matrix X, with scaling $\frac{1}{\sqrt{n}}$, where all entries of X are i.i.d. with mean zero and variance 1 to the circular law, alludes to the level of difficulty in dealing with non-symmetric matrices in general. A series of works by Ginibre [1965], Edelman [1997], Girko [1984], Bai [1997], Girko [2004], Tao and Vu [2008], Pan and Zhou [2010] led to the gradual advancement before it was proved in its full generality by Tao and Vu [2010].*

This thesis presents a unified set up for dealing with the LSD of various symmetric matrices and singular values of some non-symmetric matrices. The problem of dealing with the ESD of non-symmetric matrices does not fit in with the rest of the thesis and is not addressed here.

1.4 Plan of the thesis

In Chapter 2, we describe the well known *moment method* and the *Stieltjes transform method* briefly. Then we describe several metrics to study the closeness and convergence of ESDs and EESDs. Next, we discuss some familiar notions associated with the moment method such as, *link function, circuits, words*, and their relation to *partitions* and limiting moment sequences. This forms a very crucial part of the techniques used in this thesis. Finally, we describe the standard concepts of *cumulants, free cumulants* and *half cumulants* that we will use in the next chapters.

In Chapter 3, we investigate LSD of the Wigner matrix described in Section 1.3. Under appropriate conditions, we find the LSD in Theorem 3.3.1 and describe a set of partitions, the *special symmetric* partitions, that play a crucial role in the limiting moments. Then we discuss some corollaries which relates our result to the existing ones. Finally we conclude the chapter with some simulations illustrating various distribution that occurs as limits. *This chapter is based on Bose, Saha, Sen and Sen [2022].*

In Chapter 4, we investigate the LSD of symmetric reverse circulant, symmetric circulant, Toeplitz and Hankel matrices described in Section 1.3. Under appropriate conditions, we find their LSDs in Theorems 4.2.2–4.2.4 via moment method and find that *symmetric* and *even partitions* play very crucial roles in the limiting moments. Then we describe a few corollaries and conclude the chapter with some simulations. *This chapter is based on Bose, Saha and Sen [2021].*

In Chapter 5, we study the sample covariance matrix described in Section 1.3. Under appropriate conditions, we find the LSD of the S matrix in Theorem 5.2.1 and describe the LSD via special symmetric partitions. Next, we find relation between the LSD of the Wigner matrix and the S matrix. We deduce a few corollaries that yield several existing as well as new LSD results for specific models. Finally we show the versatility of the LSDs via some simulations. *This chapter is based on Bose and Sen [2022].*

In Chapter 6, we study the LSD of $S_A = AA^T$ where A is one of the matrices $T^{(s)}, T, H^{(s)}$, $H, R^{(s)}, R, C^{(s)}$ and C as described in Section 1.3. In Theorem 6.0.1, under appropriate conditions, we find the LSDs via several lemmas and then deduce some existing results as well as new results as corollaries of that theorem. Finally, we conclude the chapter with a few simulations. *This chapter is based on Bose and Sen [2022].*

Chapter 2

Methodologies

Eigenvalues of a matrix can be thought of as continuous functions of entries of the matrix. In general, for large dimensional matrices, no closed form of these functions are known. So special methods are required to study them. Two of the most common methods that are used in case of symmetric matrices are: the *moment method* and the *Stieltjes transform method*. One depends on enumerative combinatorics, and the other involves more technical sophisticated mathematical tools. In this thesis, we only use the moment method to prove our LSD results. For sake of completeness, we also discuss the Stieltjes transform method briefly.

In Section 2.1 and 2.2, we describe these methods and some general lemmas that relate them to convergence of spectral distributions. Next, in Section 2.3, we describe two metrics on the set of all probability distributions and some corresponding inequalities. In Section 2.4, we describe the concept of link function, circuits and words; thereafter relating these concepts to partitions that play a very important role in the forthcoming chapters. Finally in Section 2.5, we describe three notions of independence, and the corresponding notion of cumulants associated to each of them.

2.1 The Moment method

The moment method is used to understand a random variable X via its moments, $\mathbb{E}[X^k]$ (provided all moments are finite). The *Moment Convergence Theorem* and the *trace-moment formula* help us to use the moment method for finding the LSD of random matrices.

Lemma 2.1.1 (Moment convergence Theorem)**.** *Suppose $\{X_n\}$ is a sequence of real valued random variables with distribution $\{F_n\}$ that satisfies the following conditions:*

(i) there exists a sequence $\{\gamma_k\}_{k\geq 1}$ such that for every $k \geq 1$,
$\mathbb{E}[X_n^k] = \int x^k \, dF_n(x) \to \gamma_k$ as $n \to \infty$.

(ii) there is a unique distribution F whose moments are $\{\gamma_k\}_{k\geq 1}$.

Then, F_n converges weakly to F.

A detailed proof of this lemma is available in Bose [2018](Lemma 1.2.1) and Bai and Silverstein [2010](Lemma B.1).

Verifying condition (ii) in Lemma 2.1.1 can often be challenging without any prior information on the sequence $\{\gamma_k\}$. In Riesz [1923] and Carleman [1926], the question about the uniqueness of the distributions are addressed. We present this in the following lemma:

Lemma 2.1.2. *Let $\{\gamma_k\}$ be the sequence of moments of a distribution F which satisfy either of the following conditions:*

(i) $\lim\limits_{k\to\infty} \frac{1}{k}\gamma_{2k}^{1/2k} < \infty$ (Riesz's condition)

(ii) $\sum\limits_{k=1}^{\infty} \gamma_{2k}^{-1/2k} = \infty$ (Carleman's condition).

Then, F is the unique distribution with the moment sequence $\{\gamma_k\}_{k\geq 1}$.

Trace-moment formula: The moments of the ESD and EESD of a symmetric $n \times n$ matrix A_n, are connected to the trace of the powers of A_n in the following way:

$$\begin{aligned}
\beta_k(F^{A_n}) &= \frac{1}{n}\sum_{i=1}^{n} \lambda_i^k(A_n) = \frac{1}{n}\,\mathrm{Tr}(A_n^k), \\
\beta_k(\mathbb{E}F^{A_n}) &= \mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^{n} \lambda_i^k(A_n)\Big] = \frac{1}{n}\mathbb{E}[\mathrm{Tr}(A_n^k)],
\end{aligned} \tag{2.1.1}$$

where $\beta_k(F)$ denotes the kth moment of the distribution F.

Now we present a general lemma that is very often used to prove the convergence of the ESD and EESD using the moment method (see Bai [1999], Bai and Silverstein [2010], Bose [2018]).

Lemma 2.1.3. *Suppose $\{A_n\}$ is any sequence of symmetric random matrices such that the following conditions hold:*

(i) For every $k \geq 1$, $\frac{1}{n}\mathbb{E}[\mathrm{Tr}(A_n)^k] \to \alpha_k$ as $n \to \infty$.

(ii) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^4}\mathbb{E}[\mathrm{Tr}(A_n^k) \ - \ \mathbb{E}(\mathrm{Tr}(A_n^k))]^4 < \infty$ for every $k \geq 1$.

(iii) The sequence $\{\alpha_k\}$ is the moment sequence of a unique probability measure μ (whose distribution is say, F).

Then μ_{A_n} converges to μ weakly almost surely.

Proof. Using (2.1.1) and Lemma 2.1.1 in conjunction with the conditions (i) and (iii), we find that $\mathbb{E}F^{A_n}$ converges weakly to F. Thus, although the moments of μ_{A_n} are random, (ii) along with the Borel-Cantelli Lemma implies that μ_{A_n} converges weakly almost surely to μ. $\square$

2.2 Stieltjes Transform

The Stieltjes transform of a probability measure μ, denoted by $s_\mu(\cdot)$ is given as follows:

$$s_\mu(z) = \int \frac{1}{\lambda - z}\, d\mu(z), \quad z \in \{x + \iota y; x \in \mathbb{R}, y > 0\}. \tag{2.2.1}$$

It is an essential object in studying probability distributions. The integral in (2.2.1) is always finite. Geronimo and Hill [2003] gave a rigorous description about the relationship between limits of Stieltjes transforms of probability distributions and weak convergence. A detailed proof of this result along with some properties of Stieltjes transform is available in Bai and Silverstein [2010](Section B.2.1).

Lemma 2.2.1. *(Geronimo and Hill [2003]) Suppose $\{\mathbb{P}_n\}$ is a sequence of probability measures on $\mathbb{R}$ with Stieltjes transform $\{s_n\}$. If $\lim\limits_{n\to\infty} s_n(z) = s(z)$ for all $z \in \mathbb{C}_+$, then* *there exists a probability measure $\mathbb{P}$ with Stieltjes transform $s_{\mathbb{P}} = s$ if and only if*

$$\lim_{y\to\infty} \iota y s(\iota y) = -1, \tag{2.2.2}$$

and then $\mathbb{P}_n$ converges weakly to $\mathbb{P}$.

Now let A_n be a real symmetric matrix with eigenvalues $\{\lambda_i\}_{1\le i\le n}$ and ESD F^{A_n}. Then the Stieltjes transform of F^{A_n} is given by

$$s_{F^{A_n}} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\lambda_i - z} = \frac{1}{n}\operatorname{Tr}((A_n - zI)^{-1}).$$

To conclude the convergence of the ESD, often it is first shown that $\mathbb{E}[s_{F^{A_n}}(z)] \to s(z)$, where $s(z)$ satisfies (2.2.2). This implies the convergence of the EESD. Then using some martingale convergence techniques, it is shown that $s_{F^{A_n}}(z) - \mathbb{E}[s_{F^{A_n}}(z)] \to 0$ almost surely for each $z \in \mathbb{C}_+$.

As mentioned before, we will not use the Stieltjes transform method in this thesis. We shall use the moment method for all our proofs.

2.3 Some metrics and inequalities

Weak convergence of probability distributions is metrizable. We shall discuss mainly two metrics in this regard– the d_2 *metric* and the *Lévy metric*.

d_2 **metric**: Let F and G be two distributions with finite second moment. Then the d_2 distance between them is defined as

$$d_2(F,G) = \left[\inf_{(X\sim F, Y\sim G)} \mathbb{E}[X-Y]^2\right]^{\frac{1}{2}}, \tag{2.3.1}$$

where $(X \sim F, Y \sim G)$ denotes that the joint distribution of (X,Y) is such that the marginal distributions of X and Y are F and G respectively.

It is well-known that if $d_2(F_n, F) \to 0$ as $n \to \infty$, then F_n converges to F in distribution, i.e., $F_n \xrightarrow{\mathcal{D}} F$ (see Sturm [2006], Villani [2008], Lott and Villani [2009], Bose [2018]).

Lévy metric: Let F and G be two distribution functions. Then the *Lévy distance* between F and G is given by

$$L(F,G) = \inf\{\epsilon : F(x-\epsilon) - \epsilon \leq G(x) \leq F(x+\epsilon) + \epsilon\}.$$

It is well-known that if $\{F_n\}$ and F are probability measures, then $L(F_n, F) \to 0$ as $n \to \infty$, implies $F_n \xrightarrow{\mathcal{D}} F$ (Billingsley [1968]).

The following inequalities concerning these metrics will be required.

Lemma 2.3.1. *Let A and B be $n \times n$ symmetric real matrices. Then*

$$d_2^2(F^A, F^B) \leq \frac{1}{n} \operatorname{Tr}(A-B)^2. \tag{2.3.2}$$

The idea of the proof is borrowed from Lemma 1.3.2 in Bose [2018].

Proof. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_n$ be the eigenvalues of A and B, respectively. Consider the joint distribution which puts mass $\frac{1}{n}$ at (λ_i, δ_i). Then the marginals are the ESDs of A and B. Thus, from (2.3.1), we have

$$d_2^2(F^A, F^B) \leq \frac{1}{n} \sum_{i=1}^{n} (\lambda_i - \delta_i)^2. \tag{2.3.3}$$

Now, using the *Hoffman-Wielandt inequality* (Hoffman and Wielandt [1953]), we obtain

$$d_2^2(F^A, F^B) \leq \frac{1}{n} \sum_{i=1}^{n} (\lambda_i - \delta_i)^2 \leq \frac{1}{n} \operatorname{Tr}(A-B)^2.$$

This completes the proof of the lemma. □

Lemma 2.3.2. *Let $\{F_i\}_{1\leq i\leq n}$ and $\{G_i\}_{1\leq i\leq n}$ be two sequences of distributions with finite second moment. Suppose $\{X_i\}_{1\leq i\leq n}$ and $\{Y_i\}_{1\leq i\leq n}$ are sequences of random variables such that $X_i \sim F_i$ and $Y_i \sim G_i$. Then*

$$d_2^2\Big(\frac{1}{n}\sum_{i=1}^{n} F_i, \frac{1}{n}\sum_{i=1}^{n} G_i\Big) \leq \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[(X_i - Y_i)^2]. \tag{2.3.4}$$

Proof. Let e_k denote the $k-$th canonical vector in $\mathbb{R}^n$ whose $k-$th coordinate is 1, and all other coordinates are zero. Let $Z_1, \ldots, Z_n$ be random variables independent of $X_i, Y_i, 1 \leq i \leq n$ such that $(Z_1, \ldots, Z_n) = e_k$, with probability $1/n$ for each $1 \leq k \leq n$.

Let

$$X = \sum_{i=1}^{n} Z_i X_i \quad \text{and } Y = \sum_{i=1}^{n} Z_i Y_i \tag{2.3.5}$$

Then $X \sim \frac{1}{n}\sum_{i=1}^{n} F_i$ and $Y \sim \frac{1}{n}\sum_{i=1}^{n} G_i$. Thus from the definition of d_2,

$$d_2^2\left(\frac{1}{n}\sum_{i=1}^{n} F_i, \frac{1}{n}\sum_{i=1}^{n} G_i\right) \leq \mathbb{E}[(X-Y)^2]. \tag{2.3.6}$$

Now by (2.3.5), it is clear that

$$\mathbb{E}[(X-Y)^2] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[(X_i - Y_i)^2]. \tag{2.3.7}$$

Hence from (2.3.6) and (2.3.7) the result follows. □

Now, we prove a lemma that helps us estimate the closeness of two EESDs.

Lemma 2.3.3. *Suppose A and B are two $n \times n$ real symmetric random matrices and F^A and F^B are their ESDs. Then*

$$d_2^2(\mathbb{E}F^A, \mathbb{E}F^B) \leq \frac{1}{n}\mathbb{E}[\operatorname{Tr}(A-B)^2]. \tag{2.3.8}$$

Proof. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_n$ be the eigenvalues of A and B, respectively. Now, for $x \in \mathbb{R}$

$$\begin{aligned} \mathbb{E}F^A(x) =& \frac{1}{n}\sum_{i=1}^{n} \mathbb{P}(\lambda_i \leq x) \\ \mathbb{E}F^B(x) =& \frac{1}{n}\sum_{i=1}^{n} \mathbb{P}(\delta_i \leq x). \end{aligned}$$

Thus, from Lemma 2.3.2, we have

$$d_2^2(\mathbb{E}F^A, \mathbb{E}F^B) \leq \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} (\lambda_i - \delta_i)^2\right].$$

Now, using the *Hoffman-Wielandt inequality* (Hoffman and Wielandt [1953]), we obtain

$$d_2^2(\mathbb{E}F^A, \mathbb{E}F^B) \leq \mathbb{E}[\frac{1}{n}\sum_{i=1}^{n}(\lambda_i - \delta_i)^2] \leq \frac{1}{n}\mathbb{E}[\mathrm{Tr}(A-B)^2].$$

□

Now we discuss some inequalities for the Lévy metric.

Lemma 2.3.4 (Theorem A.38, Bai and Silverstein [2010]). *Let λ_k and δ_k, $1 \leq k \leq n$ be two sets of real numbers, and F and $\bar{F}$ denote their empirical distributions, i.e., $F(x) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(\lambda_i \leq x)$ and $\bar{F}(x) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(\delta_i \leq x)$, $x \in \mathbb{R}$. Then for any $\alpha > 0$,*

$$L^{\alpha+1}(F, \bar{F}) \leq \min_{\pi} \frac{1}{n}\sum_{k=1}^{n}|\lambda_k - \delta_{\pi(k)}|^{\alpha} \tag{2.3.9}$$

where L is the Lévy distance and $\pi = (\pi(1), \ldots, \pi(n))$ is any permutation of $(1, 2, \ldots, n)$.

Lemma 2.3.5 (Theorem A.37, Bai and Silverstein [2010]). *Suppose A and B are real $p \times n$ matrices and λ_k and δ_k, $1 \leq k \leq p$ are the singular values of A and B arranged in descending order. Then,*

$$\sum_{k=1}^{\min(p,n)} |\lambda_k - \delta_k|^2 \leq \mathrm{Tr}[(A-B)(A-B)^T]. \tag{2.3.10}$$

Lemma 2.3.6. *Suppose A and B are real $p \times n$ matrices and F^{S_A} and F^{S_B} denote the ESDs of AA^T and BB^T respectively. Then the Lévy distance L, between the distributions F^{S_A} and F^{S_B} satisfies the following inequality:*

$$L^4(F^{S_A}, F^{S_B}) \leq \frac{2}{p^2}(\mathrm{Tr}(AA^T + BB^T))(\mathrm{Tr}[(A-B)(A-B)^T]). \tag{2.3.11}$$

The proof of this is immediate from Lemmas 2.3.4 and 2.3.5 (see Corollary A.42 in Bai and Silverstein [2010]).

Next, we prove a lemma that helps us estimate the closeness of EESDs of

$$S_A = AA^T \quad \text{and} \quad S_B = BB^T.$$

Lemma 2.3.7. *Suppose A and B are real $p \times n$ matrices and $\mathbb{E}F^{S_A}$ and $\mathbb{E}F^{S_B}$ denote the EESDs of AA^T and BB^T respectively. Then the Lévy distance L, between these distributions, satisfies the following inequality:*

$$L^4(\mathbb{E}F^{S_A}, \mathbb{E}F^{S_B}) \leq \frac{2}{p^2}(\mathbb{E}\,\mathrm{Tr}(AA^T + BB^T))(\mathbb{E}\,\mathrm{Tr}[(A-B)(A-B)^T]). \tag{2.3.12}$$

Proof. The idea of this proof is taken from Theorem A.38 of Bai and Silverstein [2010]. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$ and $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_p$ be the singular values of A and B, respectively. We will first prove that

$$L^2(\mathbb{E}F^{S_A}, \mathbb{E}F^{S_B}) \leq \mathbb{E}\big[\frac{1}{p}\sum_{i=1}^{p} \big|\lambda_i^2 - \delta_i^2\big|\big]. \tag{2.3.13}$$

Let $\epsilon > 0$ be such that $\epsilon^2 = \frac{1}{p}\sum_{i=1}^{p} \mathbb{E}\big[\big|\lambda_i^2 - \delta_i^2\big|\big]$. If $\epsilon^2 \geq 1$, (2.3.13) is trivially true. So suppose $\epsilon^2 < 1$. Fix $x \in \mathbb{R}$. Let

$$A(x) = \{i : i \leq p, \lambda_i^2 \leq x\} \tag{2.3.14}$$

$$B(x) = \{i : i \leq p, \delta_i^2 \leq x + \epsilon\} \tag{2.3.15}$$

$$M(x) = \big|A(x) \setminus B(x)\big|. \tag{2.3.16}$$

Then we have,

$$\begin{aligned} \mathbb{E}F^{S_A}(x) - \mathbb{E}F^{S_B}(x+\epsilon) =& \mathbb{E}\left[\frac{1}{p}\sum_{i=1}^{p}\big[\mathbf{1}_{[\lambda_i^2 \leq x]} - \mathbf{1}_{[\delta_i^2 \leq x+\epsilon]}\big]\right] \\ \leq& \frac{1}{p}\mathbb{E}M(x). \end{aligned} \tag{2.3.17}$$

For any x and each $i \in A(x) \setminus B(x)$, $|\lambda_i^2 - \delta_i^2| \geq \epsilon$. Hence from (2.3.17), we obtain

$$\mathbb{E}F^{S_A}(x) - \mathbb{E}F^{S_B}(x+\epsilon) \leq \frac{1}{p}\mathbb{E}\left[\sum_{i=1}^{p} \frac{|\lambda_i^2 - \delta_i^2|}{\epsilon}\right] = \epsilon. \tag{2.3.18}$$

In a similar manner we have that

$$\mathbb{E}F^{S_B}(x-\epsilon) - \mathbb{E}F^{S_A}(x) \leq \epsilon.$$

Hence we have $L(\mathbb{E}F^{S_A}, \mathbb{E}F^{S_B}) \leq \epsilon$ and thus (2.3.13) is proved.

Now, observe that

$$\begin{aligned}
L^2(\mathbb{E}F^{S_A}, \mathbb{E}F^{S_B}) &\leq \frac{1}{p}\mathbb{E}\Big[\sum_{i=1}^{p}\big|\lambda_i^2 - \delta_i^2\big|\Big] \\
&\leq \frac{1}{p}\mathbb{E}\Big[\big(\sum_{i=1}^{p}(\lambda_i+\delta_i)^2\big)^{1/2}\big(\sum_{i=1}^{p}|\lambda_i-\delta_i|^2\big)^{1/2}\Big] \\
&\leq \frac{1}{p}\Big(\mathbb{E}\big[\sum_{i=1}^{p}(\lambda_i+\delta_i)^2\big]\Big)^{1/2}\Big(\mathbb{E}\big[\sum_{i=1}^{p}|\lambda_i-\delta_i|^2\big]\Big)^{1/2} \\
&\leq \frac{1}{p}\Big(2\mathbb{E}\big[\sum_{i=1}^{p}(\lambda_i^2+\delta_i^2)\big]\Big)^{1/2}\Big(\mathbb{E}\big[\sum_{i=1}^{p}|\lambda_i-\delta_i|^2\big]\Big)^{1/2}.
\end{aligned}$$

Now using Lemma 2.3.5 on the second factor of the above inequality, we get (2.3.12). □

The next inequality between the spectral distributions, relates the ESD to the ranks of the matrices and is often called the **rank inequality**.

Lemma 2.3.8. *(Corollary A.43 in [Bai and Silverstein, 2010]) Let A and B be $n \times n$ real symmetric matrices. Then*

$$||F^A - F^B|| \leq \frac{1}{n} rank(A-B), \tag{2.3.19}$$

where $||f|| = \sup_x |f(x)|$ is the sup norm of f.

Lemma 2.3.9. *(Corollary A.44 in [Bai and Silverstein, 2010]) Let A and B be $p \times n$ symmetric matrices. Then*

$$||F^{S_A} - F^{S_B}|| \leq \frac{1}{p} rank(A-B), \tag{2.3.20}$$

where $||f|| = \sup_x |f(x)|$ is the sup norm of f.

2.4 Some notation and preliminaries

In this section we describe some more notions related to the moment method. The ideas and concepts below are discussed in more detail in Bose [2018]. Here we discuss only what is needed in the upcoming chapters.

Let $[k]$ denote the set $\{1, 2, \ldots, k\}$ and $\mathcal{P}(k)$ denote the set of all partitions of $[k]$.

Multiplicative extension: Suppose $\{c_k\}$ is any sequence of numbers. Its multiplicative extension is defined on the set $\mathcal{P}(k)$, $k \geq 1$ as follows. For any partition σ of $[k]$, define

$$c_\sigma = \prod c_{|V|},$$

where $|A|$ denotes the cardinality of a set A, and the product is taken over all blocks V of the partition σ.

Link functions: All the patterned matrices that are being discussed in this thesis are constructed from a sequence or bi-sequence of variables $\{x_{i,n}\}$ or $\{x_{ij,n}\}$, called the *input sequence*. Let

$$L_n : \{1, 2, \ldots, n\}^2 \longrightarrow \mathbb{Z}^d, \;\; n \geq 1, d = 1 \;\; \text{or } 2$$

be a sequence of functions, called *link functions*. Often we write $L_n = L$ for convenience. Then a patterned matrix, A_n, can be described as follows:

$$A_n = ((x_{L_n(i,j)})).$$

If L is symmetric, then we have a symmetric matrix. Here we give the link functions for the matrices that are being dealt with in this thesis:

(i) Wigner matrix (W_n): $L(i, j) = (\min(i, j), \max(i, j))$. Let us denote this link function L_W.

(ii) Symmetric reverse circulant ($R_n^{(s)}$): $L(i, j) = (i + j - 2)(\text{mod } n), 1 \leq, i, j \leq n$. Let us denote this link function as $L_{R^{(s)}}$.

(iii) Asymmetric reverse circulant (R_p):

$$L(i, j) = \begin{cases} (i + j - 2)(\text{mod } n) & i \leq j, \\ -[(i + j - 2)(\text{mod } n)] & i > j. \end{cases}.$$

Let us denote this link function as L_R.

(iv) Symmetric circulant ($C_n^{(s)}$): $L(i, j) = n/2 - |n/2 - |i - j||, 1 \leq, i, j \leq n$. Let us denote this link function as L_{SC}.

(v) Circulant (C_p): $L(i,j) = (j-i)(\text{mod } n)$. Let us denote this link function as L_C.

(vi) Symmetric Toeplitz ($T_n^{(s)}$): $L(i,j) = |i-j|, 1 \leq, i,j \leq n$. Let us denote this link function as $L_{T^{(s)}}$.

(vii) Asymmetric Toeplitz (T_p): $L(i,j) = i-j$. Let us denote this link function as L_T.

(viii) Symmetric Hankel ($H_n^{(s)}$): $L(i,j) = i+j$. Let us denote this link function as $L_{H^{(s)}}$.

(ix) Asymmetric Hankel (H_p):

$$L(i,j) = \begin{cases} (i+j) & i \geq j, \\ -(i+j) & i < j. \end{cases}.$$

Let us denote this link function as L_H.

(x) Sample covariance matrix (S): The relevant link function for the sample covariance matrix is given by a pair of functions as follows:

$$L_1(i,j) = (i,j) \quad \text{and} \quad L_2(i,j) = (j,i).$$

We shall see later in Section 5.3, how this pair of link functions help in describing the S matrix.

Now we define a characteristic of the link functions that counts the maximum number of repetitions of a specific variable in a row or coloumn.

Define for a link function L,

$$\Delta(L) = \sup_n \sup_{t \in \mathbb{Z}} \sup_{1 \leq k \leq n} \left|\{l : 1 \leq l \leq n, L(k,l) = t\}\right|. \tag{2.4.1}$$

For the matrices mentioned in the thesis, we have $\Delta(L) < \infty$.

Observe that $\Delta(L) = 1$ for symmetric reverse circulant and symmetric Hankel matrices, and $\Delta(L) = 2$ for Wigner, symmetric circulant and symmetric Toeplitz matrices.

From Lemma 2.1.1, it is evident that in order to use the moment method to study the LSD of symmetric matrices, the computation of the moments play a very crucial part. When the matrix under consideration is patterned, i.e., can be written as $A_n = ((x_{L(i,j)}))$, for some link function L, then from (2.1.1), we have,

$$\beta_k(F^{A_n}) = \frac{1}{n} \sum_{1 \le i_1, \ldots, i_k \le n} x_{L(i_1,i_2)} x_{L(i_2,i_3)} \cdots x_{L(i_k,i_1)}, \tag{2.4.2}$$

$$\beta_k(\mathbb{E}F^{A_n}) = \frac{1}{n} \sum_{1 \le i_1, \ldots, i_k \le n} \mathbb{E}[x_{L(i_1,i_2)} x_{L(i_2,i_3)} \cdots x_{L(i_k,i_1)}]. \tag{2.4.3}$$

In order to keep track of these sums, we make use of the notion of *Circuits and Words.*

Circuit: For a fixed n, a *circuit* π is a function $\pi : \{0, 1, 2, \ldots, k\} \longrightarrow \{1, 2, 3, \ldots, n\}$ with $\pi(0) = \pi(k)$. We say that the *length* of π is k and denote it by $\ell(\pi)$. Suppose A_n is any patterned matrix with link function L, that is, $A_n = \big((x_{L(i,j)})\big)$. Then from (2.1.1) and (2.4.2), using circuits, we can express the trace of A_n^k as

$$\mathrm{Tr}(A_n^k) = \sum_{\pi:\ell(\pi)=k} x_{L(\pi(0),\pi(1))} x_{L(\pi(1),\pi(2))} \cdots x_{L(\pi(k-1),\pi(k))} = \sum_{\pi:\ell(\pi)=k} X_\pi, \tag{2.4.4}$$

where $X_\pi = x_{L(\pi(0),\pi(1))} x_{L(\pi(1),\pi(2))} \cdots x_{L(\pi(k-1),\pi(k))}$. For any π, the values $L(\pi(i-1), \pi(i))$ will be called $L - values$ or *edges*. When an edge appears more than once in a circuit π, then π is called *matched*. Any m circuits $\pi_1, \pi_2, \ldots, \pi_m$ are said to be *jointly-matched* if each edge occurs at least twice across all circuits. They are said to be *cross-matched* if each circuit has an edge which occurs in at least one of the other circuits.

Equivalence of Circuits, Words: The circuits can be classified into groups via an equivalence relation. Circuits π_1 and π_2 are *equivalent* if and only if for all $1 \le i, j \le k$,

$$L(\pi_1(i-1), \pi_1(i)) = L(\pi_1(j-1), \pi_1(j)) \Longleftrightarrow L(\pi_2(i-1), \pi_2(i)) = L(\pi_2(j-1), \pi_2(j)).$$

It is easy to see that the above indeed is an equivalence relation on $\{\pi : \ell(\pi) = k\}$. This implies that $\{\pi : \ell(\pi) = k\}$ can be divided into equivalence classes. An equivalence class of circuits can be indexed by an element of $\mathcal{P}(k)$, by identifying the positions where the edges match as one block of a partition of $[k]$. For example, the partition

$\{\{1,3\},\{2,4,5\},\{6\}\}$ of $[6]$ corresponds to the equivalent class

$$\big\{\pi : \ell(\pi) = 6 \text{ and } L(\pi(0),\pi(1)) = L(\pi(2),\pi(3)),$$
$$L(\pi(1),\pi(2)) = L(\pi(3),\pi(4)) = L(\pi(4),\pi(5))\big\}.$$

Further, an element of $\mathcal{P}(k)$ can be identified with a *word* of length k of letters where the first appearance of each letter is in *alphabetical order*. Given a partition, we represent the integers of the same partition block by the same letter. For example, the partition $\{\{1,3\},\{2,4,5\},\{6\}\}$ of $[6]$ corresponds to the word *ababbc*. On the other hand, the word *aabccba* represents the partition $\{\{1,2,7\},\{3,6\},\{4,5\}\}$ of $[7]$. A typical word will be denoted by $\boldsymbol{\omega}$ and its i-th letter as $\boldsymbol{\omega}[i]$. For example, for the word $\boldsymbol{\omega} = abbacac$, the partition is $\{\{1,4,6\},\{2,3\},\{5,7\}\}$, $\boldsymbol{\omega}[1] = \boldsymbol{\omega}[4] = \boldsymbol{\omega}[6] = a, \boldsymbol{\omega}[2] = \boldsymbol{\omega}[3] = b, \boldsymbol{\omega}[5] = \boldsymbol{\omega}[7] = c$. When a letter x appears at the ith position, i.e., $\boldsymbol{\omega}[i] = x$, the coordinates of any circuit *pi* associated to this x is $(\pi(i-1),\pi(i))$ and we will say x appears at $(\pi(i-1),\pi(i))$.

To count the number of circuits efficiently, we need to find the cardinality of the equivalence classes arising out of the different link functions.

The class $\Pi(\boldsymbol{\omega})$: Attached to any word $\boldsymbol{\omega}$ and any link function L, is an equivalence class of circuits $\Pi(\boldsymbol{\omega})$:

$$\Pi(\boldsymbol{\omega}) = \Big\{\pi : \boldsymbol{\omega}[i] = \boldsymbol{\omega}[j] \Leftrightarrow L(\pi(i-1),\pi(i)) = L(\pi(j-1),\pi(j))\Big\}.$$

This implies that for any word $\boldsymbol{\omega}$ of length k, the cardinality of $\Pi(\boldsymbol{\omega})$ is

$$\big|\Pi(\boldsymbol{\omega})\big| = \big|\big\{\big(\pi(0),\pi(1),\ldots,\pi(k)\big) : 1 \le \pi(i) \le n \text{ for } i = 0,1,\ldots,k,\ \pi(0) = \pi(k),$$
$$L(\pi(i-1),\pi(i)) = L(\pi(j-1),\pi(j)) \text{ if and only if } \boldsymbol{\omega}[i] = \boldsymbol{\omega}[j]\big\}\big|. \quad (2.4.5)$$

Vertex and generating vertex: Any $\pi(i)$ of a circuit π will be called a *vertex*. It is a *generating vertex* if $i = 0$ or $\boldsymbol{\omega}[i]$ is the first occurrence of a letter in the word $\boldsymbol{\omega}$ corresponding to π. All other vertices are *non-generating*. Note that the vertices and their values are both denoted by $\pi(\cdot)$. However, when we talk about a vertex $\pi(i)$, we mean it's a variable that can take at most n number of values. We will see, how many such choices we can make is based on whether the vertex is generating or non-generating.

Observe that as $\Delta(L) < \infty$ (see (2.4.1)), the circuits corresponding to a word $\boldsymbol{\omega}$ are completely determined by the generating vertices upto finitely many choices: once we have determined the generating vertices, there are only finitely many choices for the non-generating vertices. In particular, for the matrices considered here, the non-generating vertices are linear combinations of the generating vertices due to the structure of the link functions. (Note that here *any such linear combination has coefficients from* $\mathbb{Z}$.) This can be shown by induction. Since there is an indispensable dependence on the link function, the set of generating vertices and the number of ways they can be chosen varies from one matrix to another.

Free choice of generating vertices: Using the notion of words, the problem of counting circuits in (2.4.4) boils down to finding $|\Pi(\boldsymbol{\omega})|$. From (2.4.5), observe that the growth of $|\Pi(\boldsymbol{\omega})|$ is determined by the number of distinct generating vertices, and whether or not they have *free choices*. By *free choice of the generating vertices*, it is meant that the numerical values of each of the generating vertices (see the discussion on generating and non-generating vertices above) can be chosen independently of the choice of the other generating vertices, except that they cannot be equal to each other. For example, consider the Wigner matrix and the word $aabb$, then $\pi(0), \pi(1)$ and $\pi(3)$ are the generating vertices. By free choice, it is meant, the numerical values of $\pi(0)$ can be chosen in n ways, those of $\pi(1)$ can be chosen in $(n-1)$ ways and those of $\pi(3)$ can be chosen in $(n-2)$ ways.

The first vertex $\pi(0)$ is always generating, and after that there is one generating vertex for each new letter in $\boldsymbol{\omega}$. So, if $\boldsymbol{\omega}$ has b distinct letters then the number of generating vertices is $(b+1)$. Note that the growth of $|\Pi(\boldsymbol{\omega})|$ is determined by the number of generating vertices that can be chosen freely. For some words, depending on the link function, some of these vertices may not be chosen freely, that is some of the generating vertices might be a linear combination (with coefficients from $\mathbb{Z}$) of the other generating vertices. For all the matrices in this thesis, since $\Delta(L) < \infty$, we can conclude that

$$|\Pi(\boldsymbol{\omega})| = \mathcal{O}(n^{b+1}) \quad \text{whenever } \omega \text{ has } b \text{ distinct letters.} \tag{2.4.6}$$

Note that $|\Pi(\boldsymbol{\omega})| \neq n^{b+1}$. This is because some trivial cases such as all vertices being equal, needs to be discarded. However, such cases are negligible compared to n^{b+1}.

If all $(b+1)$ distinct generating vertices of $\boldsymbol{\omega}$ can be chosen freely, then $\lim_{n\to\infty} \frac{|\Pi(\boldsymbol{\omega})|}{n^{b+1}} > 0$.

A great deal of our proofs revolve around finding which words contribute positively to the limiting moments. That is determined by the number of ways the circuits can be chosen subject to the equivalence relation arising out of the link functions. This is where $\lim_{n\to\infty} \frac{|\Pi(\boldsymbol{\omega})|}{n^{b+1}}$ will play a key role.

2.5 Cumulants, free cumulants and half cumulants

Cumulants, free cumulants and half cumulants arise very naturally while discussing *classical, free and half independence* of different variables. Below we briefly discuss these concepts as they will arise while investigating the LSD of some matrices in the forthcoming chapters. Detailed development of these concepts can be found in Nica and Speicher [2006], Banica et al. [2012], Bose et al. [2011b], Novak [2014], and Bose [2021].

2.5.1 Cumulants and classical independence

Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and X is a real valued random variable. Then the *moment generating function* (mfg) of X is defined as

$$M_X(t) = \mathbb{E}[\exp(tX)],$$

provided the expectation exist for all $t \in [-\delta, \delta]$ with $\delta > 0$.

Moments of X can be determined from the mgf of X via its derivatives at 0 as follows:

$$\mathbb{E}[X^k] = \frac{d^k M_X(t)}{dt^k}\bigg|_{t=0}, \quad k \geq 1. \tag{2.5.1}$$

The *cumulant generating function* (cgf) of X (when X has mgf $M_X(t)$) is defined as

$$C_X(t) = \ln[M_X(t)].$$

The *cumulants* are defined as the derivative of the cgf at 0,

$$C_k(X) = \frac{d^k C_X(t)}{dt^k}\bigg|_{t=0},\ k \geq 1.$$

It is easy to see that the first two cumulants are the mean and variance of X.

Relation between moments and cumulants: It is possible to determine the moments if the cumulants of a random variable are known and vice versa, with the help of partitions and *Möbius function* (T. N. Thiele [1889], Hald [2000], Novak [2014], Bose [2021]). The formula for deriving moments from cumulants is relatively straightforward. We state this formula now.

Recall multiplicative extension of a sequence of numbers from Section 2.4. Let X be a random variable with moment and cumulant sequences $\{m_k\}_{k\geq 1}$ and $\{c_k\}_{k\geq 1}$, respectively. Then,

$$m_k = \sum_{\pi \in \mathcal{P}(k)} c_\pi,\ k \geq 1. \tag{2.5.2}$$

Relation between cumulants and independence: Suppose $\{X_i, 1 \leq i \leq n\}$ are random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. They are said to be (classically) independent if

$$\mathbb{P}(X_i \in B_i, 1 \leq i \leq n) = \prod_{i=1}^{n} \mathbb{P}(X_i \in B_i) \quad \text{for all Borel sets } B_1, B_2, \ldots, B_n.$$

The definitions of mgf and cgf can be extended to the joint distribution of $\{X_i\}_{1\leq i\leq n}$.

The *joint moment generating function* (mgf) is defined as

$$M_{X_1,\ldots,X_n}(t_1,\ldots,t_n) = \mathbb{E}\Big[\exp\big(\sum_{i=1}^{n} t_i X_i\big)\Big],\ t_1\ldots,t_n \in \mathbb{R},$$

provided the expectation exist for all $(t_1\ldots,t_n)$ in a neighbourhood of the origin in $\mathbb{R}^n$.

The joint moments are determined from the mgf as follows:

$$\mathbb{E}[X_1^{k_1}\cdots X_n^{k_n}] = \frac{\partial^{k_1+\cdots+k_n} M_{X_1,\ldots,X_n}(t_1,\ldots,t_n)}{\partial t_1^{k_1}\cdots \partial t_n^{k_n}}\bigg|_{t_1=0,\ldots,t_n=0},\ k \geq 1. \tag{2.5.3}$$

Lemma 2.5.1. *Let the mgf $M_{X_1,\ldots,X_n}(t_1,\ldots,t_n)$ of $\{X_i, 1 \le i \le n\}$ exist in a neighbourhood of the origin in $\mathbb{R}^n$. Then $\{X_i, 1 \le i \le n\}$ are independent if and only if*

$$M_{X_1,\ldots,X_n}(t_1,\ldots,t_n) = \prod_{i=1}^{n} M_{X_i}(t_i), \tag{2.5.4}$$

for all $(t_1 \ldots, t_n)$ in a neighbourhood of the origin in $\mathbb{R}^n$.

If $M_{X_1,\ldots,X_n}$ is finite in a a neighbourhood of the origin in $\mathbb{R}^n$, the *joint cumulant generating function* (cgf) is defined as

$$C_{X_1,\ldots,X_n}(t_1,\ldots,t_n) = \ln[M_{X_1,\ldots,X_n}(t_1,\ldots,t_n)],\ t_1 \ldots, t_n \in \mathbb{R}.$$

The joint cumulants are determined from the cgf as follows:

$$c_{k_1,\ldots,k_n}(X_1,\ldots,X_n) = \frac{\partial^{k_1+\cdots+k_n} C_{X_1,\ldots,X_n}(t_1,\ldots,t_n)}{\partial t_1^{k_1}\cdots\partial t_n^{k_n}}\bigg|_{t_1=0,\ldots,t_n=0},\ k \ge 1. \tag{2.5.5}$$

$c_{k_1,\ldots,k_n}(X_1,\ldots,X_n)$ are the cumulants of $\{X_i, 1 \le i \le n\}$. If $k_j \neq 0$ for at least two indices, then it is called a mixed cumulant of $\{X_i, 1 \le i \le n\}$.

Just as we saw in case of a single variable, the joint moments and joint cumulants of $X_1,\ldots,X_n$ as described in (2.5.3) and (2.5.5) are related as in (2.5.2), where the multiplicative extension of $c_{k_1,\ldots,k_n}(X_1,\ldots,X_n)$ on $\mathcal{P}(k_1+\cdots+k_n)$ is defined appropriately. In this thesis, our focus will be on the single variable case, so we omit the details for the moment-cumulant relation in the multivariate situation.

Lemma 2.5.2. *Let the mgf $M_{X_1,\ldots,X_n}(t_1,\ldots,t_n)$ of $\{X_i, 1 \le i \le n\}$ exist in a neighbourhood of the origin in $\mathbb{R}^n$. Then $\{X_i, 1 \le i \le n\}$ are independent if and only if*

$$C_{X_1,\ldots,X_n}(t_1,\ldots,t_n) = \sum_{i=1}^{n} C_{X_i}(t_i), \tag{2.5.6}$$

for all $(t_1 \ldots, t_n)$ in a neighbourhood of the origin in $\mathbb{R}^n$.

Lemma 2.5.2 essentially states that the variables $\{X_i, 1 \le i \le n\}$ are independent if and only if *all mixed cumulants are zero.*

2.5.2 Free cumulants and free independence

The concept of free cumulants and free independence are defined within Free Probability theory. This was introduced by Voiculescu in 1985 while he investigated von Neumann algebras of free groups. Later Voiculescu [1991] established the connection between free independence and large random matrices. Voiculescu et al. [1992], Anderson et al. [2010], Nica and Speicher [2006], Mingo and Speicher [2017], [Bose, 2021] are some references that deal with this theory in details. Here, we will describe only those notions that we will require later in the thesis. The notation that we use below are in compliance with Nica and Speicher [2006].

Non-commutative probability and free independence are the main constituents of Free probability. We begin by describing *Non-commutative probability spaces* and its connection to random variables, random matrices and ESD.

Non-commutative $*-$Probability Spaces and moments: A non-commutative probability space $(\mathcal{A}, \phi)$ is a unital $*-$algebra $\mathcal{A}$ over $\mathbb{C}$ with a linear functional $\phi : \mathcal{A} \to \mathbb{C}$ such that $\phi(1_{\mathcal{A}}) = 1$, where $1_{\mathcal{A}}$ denotes the unit (identity) of $\mathcal{A}$. The function ϕ is often called a *state* of the algebra $\mathcal{A}$.

The elements of $\mathcal{A}$ are called non-commutative variables (analogs of random variables in classical probability). If $a \in \mathcal{A}$ is such that $a = a^*$, then a is called a *self-adjoint* variable, and if $aa^* = a^*a$, then a is called *normal*. All self-adjoint variables are normal but the converse is not true. Also observe that ϕ is really the analog of the expectation operator for classical random variables. The state ϕ is said to be

$$\begin{aligned} &\textit{tracial} \quad \text{if} \quad \phi(ab) = \phi(ba), \quad \text{for all} \quad a, b \in \mathcal{A}, \\ &\textit{positive} \quad \text{if} \quad \phi(aa^*) \geq 0, \quad \text{for all} \quad a \in \mathcal{A}. \end{aligned}$$

Observe that the set of all $n \times n$ real matrices, $\mathcal{M}_n(\mathbb{R})$, with the state $\frac{1}{n}\operatorname{Tr}$, is a Non-commutative $*-$Probability Space where the $*-$operation is taking transpose. Also note that the state $\frac{1}{n}\operatorname{Tr}$ is tracial and positive.

Moments: Let a be an element of a non-commutative $*-$probability space $(\mathcal{A}, \phi)$. Then $\{\phi(a^{\epsilon_1} a^{\epsilon_2} \cdots a^{\epsilon_k}), \epsilon_i \in \{1, *\}, k \geq 1\}$ are called the $*-$moments of a. If a is self-adjoint, then $\{\phi(a^k)\}_{k\geq 1}$ are called the moments of a.

It is to be noted that as we are in the non-commutative setting, the above moments need not define any probability distribution. However the existence of a probability distribution corresponding to a normal variable in $(\mathcal{A}, \phi)$ can be assured as follows. Suppose a is normal and there is a unique measure μ on $\mathbb{C}$ such that for all $m, n \in \mathbb{N}$,

$$\phi(a^m a^{*n}) = \int z^m \bar{z}^n \; d\mu(z). \tag{2.5.7}$$

Then μ is called the probability measure of a.

Conversely, given a probability distribution μ on Ω with all moments finite, we can define a $*-$probability space $(\mathcal{A}, \phi)$, with $\mathbf{A} = \cap_{1\leq p<\infty} L^p(\Omega, \mu)$, such that ϕ is defined as follows:

$$\phi(a) = \int a(\omega) \; d\mu(\omega), a \in \mathcal{A}. \tag{2.5.8}$$

Thus for any probability measure with finite moments, we get a $*-$probability space and variables corresponding to it, see [Nica and Speicher, 2006]. This will help us later to describe *free cumulants* and *half cumulants* of probability distributions.

The definition of moments can be extended naturally to any collection of variables $\{a_i, 1 \leq i \leq n\}$. Then $\{\phi\big(\Pi(a_i, a_i^*) : 1 \leq i \leq n\big), \Pi$ is a finite degree monomial$\}$ are the joint $*-$moments of $\{a_i, 1 \leq i \leq n\}$.

As we saw in case of (classical) cumulants, *free cumulants* can also be expressed via moments of the variables. Towards that, we first need to describe the notions of *Möbius functions* and *non-crossing* partitions.

Non-crossing partitions($NC(k)$): Let $\pi \in \mathcal{P}(k)$. If two elements p and q belong to the same block, we write $p \sim_\pi q$. π is *crossing* if there exist $p_1, p_2, q_1, q_2 \in [k]$ with $p_1 < q_1 < p_2 < q_2$, such that $p_1 \sim_\pi q_1$ and $p_2 \sim_\pi q_2$ but $\{p_1, p_2, q_1, q_2\}$ are not in the same block of π. If π is not crossing then, it is called *non-crossing* and we write $\pi \in NC(k)$.

For example, $\{\{1,4,5\},\{2,3\}\} \in NC(5)$ but $\{\{1,4\},\{2,3,5\}\} \notin NC(5)$. Also $1_k = \{1,2,\ldots,k\} \in NC(k)$.

It can be easily verified that $NC(k)$ is a lattice with the reverse refinement partial order $\leq$ (Nica and Speicher [2006]). Let the set of *intervals* of $NC(k)$ be denoted by

$$NC^{(2)}(k) = \{(\pi,\sigma), \pi,\sigma \in NC(k), \pi \leq \sigma\}.$$

Mobius function: We first describe the *Zeta function*, ξ. It is defined by

$$\xi(\pi,\sigma) = \begin{cases} 1, & (\pi,\sigma) \in NC^{(2)}(k), \\ 0 & \text{otherwise.} \end{cases}$$

For two complex-valued functions $F, G : NC^{(2)}(k) \to \mathbb{C}$, their *convolution* $F \odot G : NC^{(2)}(k) \to \mathbb{C}$ is given by

$$(F \odot G)(\pi,\sigma) = \sum_{\substack{\tau \in NC(k) \\ \pi \leq \tau \leq \sigma}} F(\pi,\tau)G(\tau,\sigma).$$

F is said to be invertible if there exists a unique G such that

$$(F \odot G)(\pi,\sigma) = (G \odot F)(\pi,\sigma) = \mathbf{1}_{(\pi=\sigma)}.$$

The *Möbius function*, $\mu : NC^{(2)}(k) \to \mathbb{C}$ is the inverse of ξ with respect to $\odot$. This inverse always exists.

Multiplicative extension of a moment sequence: Let $(\mathcal{A},\phi)$ be a non-commutative probability space. A sequence of multilinear functionals $\{\pi_n\}_{n\geq 1}$ can be defined on $\mathcal{A}^n$ as follows:

$$\phi_n(a_1,a_2,\ldots,a_n) = \phi(a_1 a_2 \cdots a_n),\ a_1,\ldots,a_n \in \mathcal{A}. \tag{2.5.9}$$

Recall the multiplicative extension of a sequence from Section 2.4. Similarly, the sequence of functionals $\{\pi_n\}_{n\geq 1}$ can be extended multiplicatively to

$\{\phi_\pi, \pi \in NC(n), n \geq 1\}$. If $\pi = \{V_1, V_2, \ldots, V_r\} \in NC(n)$, then

$$\phi_\pi[a_1, a_2, \ldots, a_n] = \prod_{i=1}^{r} \phi(V_i)[a_1, a_2, \ldots, a_n], \tag{2.5.10}$$

where for $V = \{i_1 < i_2 < \cdots < i_s\}$,

$$\phi(V)[a_1, a_2, \ldots, a_n] = \phi_s(a_{i_1}, a_{i_2}, \ldots, a_{i_s}) = \phi(a_{i_1} a_{i_2} \cdots a_{i_s}).$$

Free Cumulants: The *free cumulant* of $\{a_i, 1 \leq i \leq n\}$ of order n is defined as

$$\kappa_n(a_1, a_2, \ldots, a_n) = \sum_{\pi \in NC(n)} \phi_\pi[a_1, a_2, \ldots, a_n] \mu(\pi, 1_n), \tag{2.5.11}$$

where μ is the Möbius function on $NC^{(2)}(n)$. Suppose for some $i \neq j$, $a_i \neq a_j$ or a_j^*. Then $\kappa_n(a_1, a_2, \ldots, a_n)$ is called a *mixed cumulant* of order n. If $n = 1$ and $a_1 = a \in \mathcal{A}$ is self-adjoint, then the nth cumulant of a is given by

$$\kappa_n(a) = \kappa_n(a, a, \ldots, a).$$

Relation between moments and free cumulants: We saw in the classical case in Section 2.5.1 that it is possible to determine the moments if the cumulants of a variable are known and vice versa. Here too, the formula for deriving moments from free cumulants is relatively simple and will be used.

Let $a, a_1, a_2, \ldots, a_n \in \mathcal{A}$. We define the multiplicative extension of $\{\kappa_n(a_1, a_2, \ldots, a_n), n \geq 1\}$ on $NC(n)$ as in (2.5.10). If $\pi = \{V_1, \ldots, V_r\} \in NC(n)$, then

$$\kappa_\pi[a_1, a_2, \ldots, a_n] = \prod_{i=1}^{r} \kappa(V_i)[a_1, a_2, \ldots, a_n], \tag{2.5.12}$$

where for $V = \{i_1 < i_2 < \cdots < i_s\}$,

$$\kappa(V)[a_1, a_2, \ldots, a_n] = \kappa_s(a_{i_1}, a_{i_2}, \ldots, a_{i_s}), \text{ (see (2.5.11))}.$$

Then using the Möbius function, it can be shown that

$$\phi(a_1a_2\cdots a_n) = \sum_{\pi\in NC(n)} \kappa_\pi[a_1, a_2, \ldots, a_n], \tag{2.5.13}$$

$$\phi(a^n) = \sum_{\pi\in NC(n)} \kappa_\pi[a, a, \ldots, a], \tag{2.5.14}$$

where $\{\kappa_\pi, \pi \in NC(n), n \geq 1\}$ on the r.h.s. of (2.5.13) and (2.5.14) is the multiplicative extension of $\{\kappa_n(a_1, a_2, \ldots, a_n), n \geq 1\}$ or $\{\kappa_n(a), n \geq 1\}$.

Note that when the variable is self-adjoint and has a probability law μ corresponding to it in the sense of (2.5.7), then the free cumulants of μ are defined to be the free cumulants of a.

Relation between free cumulants and free independence (Nica and Speicher [2006]): Let $(\mathcal{A}, \phi)$ be a non-commutative probability space. Let $\{\mathcal{A}_i, i \in I\}$ be unital sub-algebras of $\mathcal{A}$. Then these sub-algebras are called *freely independent* if

$$\phi(a_1a_2\cdots a_n) = 0,$$

whenever the following holds :

(i) $n \in \mathbb{N}$,

(ii) $a_i \in \mathcal{A}_{j(i)}, j(i) \in I$,

(ii) $\phi(a_i) = 0$ for all $1 \leq i \leq n$,

(iii) $j(1) \neq j(2), j(2) \neq j(3), \ldots, j(n-1) \neq j(n)$.

Let $a_1, a_2, \ldots, a_n \in \mathcal{A}$. Then $\{a_i, 1 \leq i \leq n\}$ are *freely independent variables* if the unital algebras generated by $a_i, 1 \leq i \leq n$, $\mathcal{A}_i = \text{alg}(a_i, 1)$, are freely independent.

Lemma 2.5.3. *(Nica and Speicher [2006])(Mixed free cumulants are zero for free variables) Let $(\mathcal{A}, \phi)$ be a non-commutative probability space and let $\mathcal{A}_i, i \in I$ be unital subalgebras of $\mathcal{A}$ and $\{\kappa_n, n \geq 1\}$ be the sequence of free cumulants. Then $\{\mathcal{A}_i\}_{1\leq i\leq n}$ are freely independent if and only if for all $n \geq 2$, and for all $a_i \in \mathcal{A}_{j(i)}, i = 1, 2, \ldots n$*

with $j(1),\ldots,j(n) \in I$, *we have* $\kappa_n(a_1,\ldots,a_n) = 0$ *whenever there exists* $1 \le s,t \le n$ *with* $j(s) \ne j(t)$.

2.5.3 Half cumulants and half independence

In the previous section we saw free independence defined in the non-commutative set up. Another interesting notion of independence called, *half independence*, arises in the non-commutative setting. This has been dealt with in details in Banica et al. [2012] and Bose et al. [2011b]. Here we will only briefly discuss these notions. They will be helpful in identifying some of the LSDs in the forthcoming chapters.

From this point, we assume that all variables are self-adjoint. We will need the notion of *independence of variables in algebras* in order to describe *half independence.*

Just like free independence, *independence of variables in algebras* are defined via independence of unital algebras generated by each of the variables.

Let $(\mathcal{A},\phi)$ be a non-commutative probability space. Let $\{\mathcal{A}_i, i \in I\}$ be unital sub-algebras of $\mathcal{A}$. Then the sub-algebras $\{\mathcal{A}_i\, i \in I\}$ are called *independent* if they commute, and for all $a_i \in \mathcal{A}_{j(i)}(j(i) \in I)$,

$$\phi(a_1a_2\cdots a_n) = \phi(a_1)\phi(a_2)\cdots\phi(a_n) \tag{2.5.15}$$

whenever $k \ne l \Rightarrow j(k) \ne j(l)$.

Let $a_1,a_2,\ldots,a_n \in \mathcal{A}$. Then $\{a_i, 1 \le i \le n\}$ are said to be *independent variables* if $\{\mathcal{A}_i = \mathrm{alg}(a_i,1), 1 \le i \le n\}$, are independent.

Half commuting elements: Let $(\mathcal{A},\phi)$ be a non-commutative probability space and $\{a_i, i \in I\}$ be a collection of variables in $\mathcal{A}$. Then $\{a_i, i \in I\}$ are said to *half commute* if

$$a_ia_ja_k = a_ka_ja_i, \quad \text{for all } i,j,k \in I.$$

To define *half independence*, two other terminology that are required are *unbalanced variables* with respect to $\{a_i, i \in I\}$ and *independence of variables in an algebra.*

A variable $a = a_{i_1}a_{i_2}\cdots a_{i_k}$, $i_1, i_2, \ldots, i_k \in I$ is said to be *balanced* with respect to $\{a_i, i \in I\}$, if each variable a_i appears equal numer of times in odd and even positions of a. If a is not balanced, then it is called *unbalanced*.

Half independence(Banica et al. [2012]): Half commuting elements $\{a_i, i \in I\}$ are said to be *half independent* if the following is true:

(i) $\{a_i^2, i \in I\}$ are independent (see (2.5.15)),

(ii) if $a = a_{i_1}a_{i_2}\cdots a_{i_k}$ is unbalanced with respect to $\{a_i, i \in I\}$, then $\phi(a) = 0$.

Observe that for half independent variables, the odd moments are always 0.

From Section 2.5.1 and 2.5.2, it can be seen that the notions of cumulants and free cumulants are intimately tied to the set of all partitions and the set of all non-crossing partitions. The corresponding set of partitions for *half cumulants* are the *symmetric partitions*. We first describe this.

Symmetric partition: A partition $\sigma = \{V_1, V_2, \ldots, V_j\}$ of $[k]$ is said to be a symmetric partition if each V_i has the same number of odd and even integers. The set of all symmetric partitions of $[k]$ is denoted by $S(k)$ and the set of all symmetric partitions of $[k]$ with b blocks is denoted by $S_b(k)$.

For example, $\{\{1,4\},\{2,3,5,6\},\{7,10\},\{8,9\}\} \in S_4(10)$ is a symmetric partition of $[10]$, but $\{\{1,3\},\{2,4,5,6\},\{7,10\},\{8,9\}\} \notin S(10)$. In particular, all blocks of $\pi \in S(k)$ must be of even size.

Even partition: A partition $\sigma = \{V_1, V_2, \ldots, V_j\}$ of $[k]$ is said to be an even partition if each V_i has an even number of integers. The set of all even partitions of $[k]$ is denoted by $E(k)$ and the set of all even partitions of $[k]$ with b blocks is denoted by $E_b(k)$.

For example, $\{\{1,3\},\{2,4,5,6\},\{7,10\},\{8,9\}\} \in E_4(10)$ is an even partition of $[10]$. Observe that if k is odd, then $[k]$ does not have an even partition.

For any $k \geq 1$, define $NCE(2k) = NC(2k) \cap E(2k)$, i.e., $NCE(2k)$ is the set of all non-crossing partitions whose blocks are all of even size. It can be easily seen that

$$NCE(2k) \subset S(2k) \subset E(2k) \tag{2.5.16}$$

Recall the notion of multiplicative extension on $NC(k)$ from Section 2.5.2. Multiplicative extension can similarly be defined on $S(k)$. We omit the details.

Half cumulant(Bose et al. [2011b]): Let $(\mathcal{A}, \phi)$ be a non-commutative probability space and let $\{a_i, i \in I\}$ be variables in $\mathcal{A}$ such that $\phi(a_{i_1}a_{i_2}\cdots a_{i_k}) = 0$ whenever k is odd. Then the half cumulants $\{r_k\}$ of $\{a_i, i \in I\}$ are defined by the following recursion:

$$\phi(a_{i_1}a_{i_2}\cdots a_{i_k}) = \sum_{\pi \in S(k)} r_\pi(a_{i_1}, a_{i_2}, \ldots, a_{i_k}), \quad \text{for any } i_1, i_2, \ldots, i_k \in I. \tag{2.5.17}$$

If $a \in \mathcal{A}$ is self-adjoint, then the half cumulants $\{r_k(a)\}_{k\geq 1}$ of a are given by

$$r_k(a) = r_k(a, a, \ldots, a). \tag{2.5.18}$$

Note that when the variable is self-adjoint and has a probability law μ corresponding to it in the sense of (2.5.7), then the half cumulants of μ are defined to be the half cumulants of a.

Observe that from the definition of half cumulants $r_{2k+1}(a) = 0$ for all $k \geq 0$.

For example, consider the *standard symmetrized Rayleigh* distribution $\mathcal{R}$, that has density

$$f(x) = |x| \exp(-x^2), \quad x \in \mathbb{R}.$$

The moments $\beta_k(\mathcal{R})$ of $\mathcal{R}$ are given by

$$\beta_k(\mathcal{R}) = \begin{cases} 0 & \text{if } k \text{ is odd,} \\ k! & \text{if } k \text{ is even.} \end{cases}$$

Recall that $S_b(2k)$ is the set of all symmetric partititions with b blocks. Now $k! = \sum_{\pi \in S_k(2k)} 1$. A proof of this fact is available in Lemma 2.3.1 in Bose [2018].

Suppose a is a (self-adjoint) variable whose probability distribution is $\mathcal{R}$. Then $\phi(a^k) = \beta_k(\mathcal{R})$. Therefore the half cumulants of a are given by

$$r_2(a) = 1, r_{2n}(a) = 0, n \geq 2.$$

Lemma 2.5.4. *(Theorem 1, Bose et al. [2011b]) Suppose $\{a_i, 1 \leq i \leq n\}$ is a sequence of self-adjoint half commuting variables in $(\mathcal{A}, \phi)$ and $a = a_{i_1}a_{i_2}\cdots a_{i_{2k}}$ is such that a_i*

occurs s_i times in a. Then the following hold true:

(a) *if $\{a_i, 1 \leq i \leq n\}$ are half independent and a is balanced with respect to $\{a_i, 1 \leq i \leq n\}$, then $r_{2k}(a_{i_1}, a_{i_2}, \ldots, a_{i_{2k}}) = 0$ whenever there is $1 \leq i, j \leq n$ with $s_i, s_j \geq 2$;*

(b) *if $\phi(a_{i_1} a_{i_2} \cdots a_{i_k}) = 0$ whenever k is odd and $r_{2k}(a_{i_1}, a_{i_2}, \ldots, a_{i_{2k}}) = 0$ whenever $s_i, s_j \geq 1$ for some $1 \leq i, j \leq n$, then $\{a_i, 1 \leq i \leq n\}$ are half independent.*

Chapter 3

Wigner matrices

The symmetric matrix that has received the most attention in Random Matrix Theory is the *Wigner matrix*. Since the seminal work of Wigner [1955], this matrix has gained increasing importance and there has been various studies about the LSD of the Wigner matrix and its different variations. Classically a Wigner matrix W_n is a *symmetric matrix* defined as

$$W_n = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1(n-1)} & x_{1n} \\ x_{12} & x_{22} & x_{23} & \cdots & x_{2(n-1)} & x_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{1n} & x_{2n} & x_{3n} & \cdots & x_{(n-1)n} & x_{nn} \end{bmatrix}.$$

In Section 3.1, we describe a few LSD results that already exist in the literature for Wigner matrices. These are closely related to the main result of this chapter that is described in Section 3.3 (see Theorem 3.3.1). Before that, in Section 3.2, we give the detailed description of the *special symmetric words* that play a crucial role in the LSD. Next, in Section 3.5, we discuss how the results described in Section 3.1 can be concluded from Theorem 3.3.1. We conclude the chapter with some simulations that show the various distributions that can appear as LSD (see Section 3.5.7). *This chapter is based on Bose, Saha, Sen and Sen [2022].*

3.1 Review of existing literature

Initially, Wigner [1955] considered the entries of W_n to be i.i.d. real Gaussian, and showed that the EESD of $\frac{1}{\sqrt{n}}W_n$ converges weakly to the *semicircle distribution*. In Wigner [1958], a more general version was proved where the sub-diagonal entries are independent with variance σ^2, and the diagonal entries have variance $2\sigma^2$, and every other higher order moments are uniformly bounded, i.e., $\mathbb{E}[x_{ij}^k] \leq M_k$ where $\{x_{ij}; i \leq j\}$ are the entries of W_n. These results have been extended by Grenander [1968] who proved the convergence of the ESD to the semicircle law in probability, and by Arnold [1967], Arnold [1971] where the convergence of the ESD is shown to be in the almost sure sense. Some subsequent studies include Pastur [1972] and Pastur [1973] with independent mean zero variance 1 entries that satisfy a *Lindeberg-type* condition, also known as *Pastur's condition*

$$\lim_{n\to\infty} \frac{1}{\eta^2 n^2} \sum_{i,j} \mathbb{E}\big[x_{ij}^2 \mathbf{1}_{[|x_{ij}|>\eta\sqrt{n}]}\big] = 0 \quad \text{for every } \eta > 0. \tag{3.1.1}$$

Later Bai [1999] provided two extensions of the model – one that allows the sub-diagonal entries to be complex i.i.d. with variance 1 and the other where the sub-diagonal entries are independent with common mean 0, variance 1 and satisfy (3.1.1). Bai [1999] showed that in either case the ESD of $\frac{1}{\sqrt{n}}W_n$ converges weakly almost surely to the semicircle law.

We recall the most widely known result in the fully i.i.d. regime.

Result 3.1.1. ***(Standardized fully i.i.d. entries)*** *Suppose that the entries $\{x_{i,j}\,; 1 \leq i \leq j \leq n\}$ of W_n are i.i.d. with mean 0 and variance 1. Then, as $n \to \infty$, the almost sure LSD of $W_n/\sqrt{n}$ is the standard semicircular distribution. This distribution, say μ_s, has the following density*

$$f(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2} & \textit{if } x \in [-2,\ 2], \\ 0 & \textit{otherwise.} \end{cases}$$

The moments $\beta_k(\mu_s)$ of μ_s are given by

$$\beta_k(\mu_s) = \begin{cases} 0 & \text{if } k \text{ is odd,} \\ \frac{2}{(k+2)}\binom{k}{\frac{k}{2}} = |NC_2(k)| & \text{if } k \text{ is even.} \end{cases}$$

Here for any k, $\frac{1}{(k+1)}\binom{2k}{k} = C_{2k}$ is the $2k$th *Catalan number* and $NC_2(2k)$ denotes the set of non-crossing pair-partitions of $\{1, 2, \ldots, 2k\}$. A detailed proof and a short history of the precursors of this result is available in Section 2.1 of Bose [2018]. As the moments of μ_s are given via the non-crossing pair partitions as in the equation above, its free cumulants (see Section 2.5.2) are $\kappa_1 = 0, \kappa_2 = 1$ and $\kappa_n = 0$ for all $n \geq 3$.

Heavy-tailed entries: A natural extension of the above model is made by considering the case where the entries are i.i.d. with a *heavy-tailed* distribution.

Result 3.1.2. *(Ben Arous and Guionnet [2008]) Suppose the entries $\{x_{ij}\ ; i \leq j\}$ of W_n are i.i.d. and satisfy $P\{|x_{ij}| > u\} = u^{-\alpha}L(u)$ as $u \to \infty$ where $L(\cdot)$ is slowly varying and $\alpha \in (0, 2)$. Also $\lim\limits_{u\to\infty} \dfrac{\mathbb{P}[x_{ij} > u]}{\mathbb{P}[|x_{ij}| > u]} = \theta \in [0, 1]$. Let $a_n = \inf\{u : P[|x_{ij}| > u] \leq 1/n\}$. Then the ESD of $\frac{W_n}{a_n}$ converges to a probability measure μ_α in probability.*

Their method of proof was the following. They first showed the convergence of the ESD of the truncated matrices (where each entry of the matrix is truncated at a number) to a non-random symmetric distribution, which then converges weakly in the space of distributions as the truncation level goes to infinity. We note from various articles relating to the heavy-tailed case, for example, Ben Arous and Guionnet [2008], Belinschi et al. [2009], Benaych-Georges et al. [2014], Male [2017], that in dealing with the heavy-tailed case, the entries of the matrix are often truncated where the truncation depends on n, and hence the truncated matrix has entries whose distribution depends on n.

Size dependent entries: Wigner matrices where the distribution of the entries are dependent on the size of the matrices has been considered by Zakharevich [2006]. The primary motivation of her result was to consider entries that are i.i.d. with distribution G_n which may be all light-tailed, but as $n \to \infty$, G_n may converge to a heavy-tailed distribution. These are called *Wigner matrices with exploding moments.*

Result 3.1.3. *(Theorem 1, Zakharevich [2006]) Suppose $\{G_k\}$ is a sequence of probability distributions each of which has mean zero and all moments finite. Let $\mu_n(k)$ be the kth moment of G_n. Suppose that for each n, the entries of the Wigner matrix W_n are i.i.d. G_n. Let*

$$\lim_{n\to\infty} \frac{\mu_n(k)}{n^{k/2-1}\mu_n(2)^{k/2}} = g_k, \quad \textit{say, exists for all} \ \ k \geq 1.$$

Then the ESD of $\frac{W_n}{\sqrt{n\mu_n(2)}}$ *converges in probability to a distribution* μ_{zak} *say, that depends only on the sequence* $\{g_{2k}\}$*.*

Note that $g_2 = 1$. It is known that μ_{zak} is a semicircular distribution if and only if $g_{2k} = 0$ for all $k \geq 2$. Further, if the $\{G_k\}_{k\geq 1}$ are identical, then clearly $g_{2k} = 0$ when $k \neq 2$, and we recover the semicircular distribution as in Result 3.1.1. Zakharevich gave a description of the moments of μ_{zak} in terms of certain *trees*. Another study that dealt with such matrices is Jung [2018], where the author describes the *Lévy-Khintchine ensemble* (the entries of the matrix being i.i.d. for every fixed n satisfying certain moment conditions) and concludes the convergence of their ESDs via local weak convergence of associated graphs.

Sparse homogeneous Erdős-Rényi graphs: A very well-known random matrix model where the entries are identically distributed for a fixed dimension, but the distribution changes with the dimension, is the adjacency matrix of the homogenous Erdős-Rényi graph. For a homogeneous Erdös-Rényi graph with n vertices, an edge between vertices i and j occur with probability p_n. This is called the *homogeneous* model as the edge probability does not depend on i, j. The adjacency matrix of homogeneous Erdös-Rényi graph is a Wigner matrix with entries following $Ber(p_n)$ distribution. This matrix has been studied under various regimes over the last few decades. One of the regimes is the sparse regime, where the average degree or average connectivity np_n tends to a constant as n increases to ∞. In this thesis, we focus our study on the sparse regime.

One of the earlier papers that dealt with the ESD of the adjacency matrix of the homogenous Erdős-Rényi graph in the sparse regime with mathematical rigour was by Bauer and Golinelli [2001]. They showed that the EESD converges weakly to a non-symmetric distribution that is not the semicircle law. They gave a description of the limiting moments via certain *trees*, which are not same as the trees in Zakharevich [2006].

Result 3.1.4. *(Bauer and Golinelli [2001]) Suppose that for each fixed* n*,* M_n *is the adjacency matrix of a simple(without loops or multiple edges) sparse homogeneous Erdös-Rényi graph with* n *vertices and edge probability* p_n *such that* $np_n \to \lambda > 0$*. Then the EESD of* M_n *converges weakly to a probability measure* μ_{bg} *say, which is symmetric* *about 0. The moments of μ_{bg} are given by*

$$\beta_{2k}(\mu_{bg}) = \sum_{b=1}^{k} I_{k,b}\lambda^b. \tag{3.1.2}$$

Here $I_{k,b}$ is the number of normalised $2k$-plets of the form $(i_1, i_2, \ldots, i_{2k})$, associated to certain trees with b edges.

Note that the $2k$th moment of the ESD of W_n can be written as

$$\beta_{2k}(\mathbb{E}F^{W_n}) = \frac{1}{n} \sum_{1 \le i_1, \ldots, i_{2k} \le n} \mathbb{E}[x_{i_1 i_2} x_{i_2 i_3} \cdots x_{i_{2k} i_1}],$$

thus giving rise to the $2k$-plets and the corresponding closed walk on a tree. A detailed construction of the trees can be found in Section 3.4 in Bauer and Golinelli [2001]. However, no nice description of μ_{bg} is available, but it is easy to check that μ_{bg} is not a semicircular distribution.

There has been subsequent works on this topic, for example, by Khorunzhy et al. [2004], Ding and Jiang [2010], Bordenave and Lelarge [2010b], Bordenave et al. [2017], Bordenave and Lelarge [2010a] that have strengthened Result 3.1.4 by proving the almost sure convergence, studying the empirical distribution via local weak convergence of sequences of graphs, as well as studying the atom found at 0 in the LSD.

Sparse inhomogeneous Erdős-Rényi graphs: If the probability of an edge occurring between two vertices i and j does depend on i, j, it is called an *inhomogeneous* graph model. There has also been developments in the inhomogenous Erdős-Rényi graphs in the sparse regime. Some interesting inhomogeneous models have been considered in Liang et al. [2007] and Bollobás et al. [2007]. In Bollobás et al. [2007], the authors define a model of Erdős-Rényi graphs that is sufficient to include many of the specific models considered previously. In their model, the *uniformly grown random graph on* c/j denoted by $G_n^{1/j}(c)$, the probability that there is an edge between the vertices i and j, is given as $p_{ij} = \min\{c/\max\{i,j\}, 1\}$. They study the phase transition and related results about $G_n^{1/j}(c)$. Further, they mention that Erdős-Rényi graphs $G(n, \kappa)$ where κ is a suitably chosen function can be dealt with in a very similar manner as $G_n^{1/j}(c)$.

Matrices with variance profile: The dependence of the distribution of an entry on its position also arise when the matrices have a *variance profile*, i.e., when the entries of the matrix are of the form $\{\sigma_{ij}x_{ij}; i \leq j\}$ or $\{\sigma(i/n, j/n)x_{ij}; i \leq j\}$, with $(\sigma_{ij})_{i \leq j}$ or $\sigma(x, y)$ being an appropriately chosen sequence or function, (see Definition 3.5.15 and 3.5.16). Some results that deal with such Wigner matrices have been derived in Anderson and Zeitouni [2006] and Lytova and Pastur [2009]. Under the assumption that the entries have common mean and variance, or have different variances but each column of the variance profile is stochastic, i.e., $\frac{1}{n}\sum_i \sigma_{ij}^2 \to 1$ or $\int \sigma^2(x, y)\, dx = 1$, it has been shown that the ESD of $\frac{1}{\sqrt{n}}W_n$ converges weakly almost surely to the semicircle distribution.

Wigner matrices with more complicated variance profile have been studied recently by Zhu [2020] using *graphon sequences*. We shall briefly describe graphons for our purposes, and a more detailed study can be found in Lovász [2012]. Using the notion of convergence of graphons and graph homomorphism densities, the author proves that the ESD of $\frac{1}{\sqrt{n}}W_n$ converges weakly almost surely to a symmetric probability distribution whose moments are described as limits of homomorphism densities on certain trees. He also provides sufficient conditions for the LSD to be semicircle. For describing Zhu's result we first define the following:

Graphons: For every $n \geq 1$, divide the interval $[0, 1]$ into n non-overlapping subintervals $I_1, \ldots, I_n$ of length $\frac{1}{n}$ each. Let $I_1 = [0, \frac{1}{n}]$, $I_i = (\frac{i-1}{n}, \frac{i}{n}]$, $i \geq 2$ and $H_n(x, y) = \frac{\sigma_{ij}^2}{n}$ if $(x, y) \in I_i \times I_j$. This defines a sequence of functions (*graphons*) H_n on $[0,\ 1]^2$.

Homomorphism density: Consider all finite multigraphs $G = (V, E)$ without loops with vertex set $V = \{1, \ldots, n\}$ and edge set E. Let $t(G, H_n)$ denote the *homomorphism density*

$$t(G, H_n) = \int_{[0,1]^{|V|}} \prod_{i\ j \in E} H_n(x_i, x_j) \prod_{i \in V} dx_i. \tag{3.1.3}$$

Result 3.1.5. *(Theorem 3.2, Zhu [2020]) Suppose the entries $\{a_{ij} : 1 \leq i \leq j \leq n\}$ of the Wigner matrix W_n are independent and have mean zero. Let $\mathbb{E}[a_{ij}^2] = \sigma_{ij}^2$. Assume that*

$$\sup_{ij} \sigma_{ij}^2 \leq B < \infty, \tag{3.1.4}$$

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{1\le i,j\le n}\mathbb{E}\big[a_{ij}^2\mathbf{1}_{[|a_{ij}|\ge\eta\sqrt{n}]}\big]=0 \quad \text{for any constant } \eta>0 \text{ (Lindeberg's condition).} \tag{3.1.5}$$

Suppose that,

$$\lim t(T,H_n) \text{ exists for every finite tree } T. \tag{3.1.6}$$

Then under the conditions (3.1.4), (3.1.5) and (3.1.6), the LSD of $W_n/\sqrt{n}$, say μ_{zhu} exists almost surely. The odd moments of μ_{zhu} are 0, and its $2k$th moment is given by

$$\beta_{2k}(\mu_{zhu})=\sum_T \lim t(T,H_n),\ \ k\ge 1, \tag{3.1.7}$$

where the sum is over all rooted trees T each with $k+1$ vertices.

Other Wigner matrices: Some other variations of Wigner matrices where the distributions of the entries depend on their positions in the matrix, include *band matrices* (Casati and Girko [1993a], Casati and Girko [1993b], Anderson and Zeitouni [2006], Molchanov et al. [1992]), *block matrices* (Bolla [2004], Ding [2014], Zhu [2020]) and *triangular matrices* (Basu et al. [2012]).

Band Wigner matrices: Band matrices are matrices whose non-zero entries form a band like structure. These matrices appear naturally in physics (see Casati et al. [1979] and Chirikov [1985]). When the bandwidth m_n of a banded Wigner matrix, say, W_n^B, is such that $m_n/n\to\alpha>0$, a constant, then Casati and Girko [1993b] proved that the ESD of $\frac{1}{\sqrt{n}}W_n^B$ converges in probability to a semicircle law.

Result 3.1.6. *(Theorem 2, Casati and Girko [1993b]) Suppose the entries $\{x_{ij};|i-j|<m_n\}$ of W_n^B are such that $\{x_{ij}\}$ are independent with mean zero and variance $\sigma^2<\infty$. All other entries of W_n^B are zero. Also assume that*

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{i,j:|i-j|\le m_n}\mathbb{E}\big[x_{ij}^2\mathbf{1}_{[|x_{ij}|\ge\eta\sqrt{n}]}\big]=0 \quad \text{for any constant } \eta>0 \text{ (Lindeberg's condition).}$$

Then the ESD of $\frac{1}{\sqrt{n}}W_n^B$ converges in probability to a semicircle law with parameters determined by α and σ.

Triangular Wigner matrices: Triangular matrices have gained importance since their consideration in Dykema and Haagerup [2004], who considered the singular values of (asymmetric) triangular Wigner matrices with Gaussian entries. Later in Basu et al. [2012], the authors concluded LSD results for symmetric triangular Wigner matrix. We shall give the formal description of such matrices in Section 3.5. Here we state the LSD result for symmetric triangular Wigner matrices, W_n^u from Basu et al. [2012].

Result 3.1.7. *(Theorem 2.2, Basu et al. [2012]) Suppose the non-zero entries $\{x_{ij}; (i+j) \leq (n+1)\}$ of W_n^u are i.i.d. with mean zero and variance 1. Then the LSD of $\frac{1}{\sqrt{n}}W_n^u$ exists almost surely and is symmetric about zero.*

Block Wigner matrices: Random block matrices with finite number of rectangular blocks have been studied in Bolla [2004], Ding [2014] and Zhu [2020]. In Zhu [2020], the graphon approach has been used to prove results on random block matrices. We state their result here.

Result 3.1.8. *Suppose W_n' is a block matrix with d^2 rectangular blocks $W_n'^{(m,l)}, 1 \leq m \leq l \leq d$ each of size $n_m \times n_l$ consisting of independent entries (modulo symmetry). Assume that the entries $\{a_{ij}\}$, $1 \leq i, j \leq n$ are centered, and all entries are independent (modulo symmetry) with*

$$\mathbb{E}\left[a_{ij}^2\right] = s_{ml} \quad \textit{whenever } a_{ij} \quad \textit{is in the } (m,l)-\textit{th block}, \tag{3.1.8}$$

$$\sup_{m,l} s_{ml} \leq C \quad \textit{for some constant } C > 0, \tag{3.1.9}$$

and a_{ij} satisfies (3.1.5). Also suppose that $\lim_{n\to\infty} \frac{n_m}{n} = \alpha_m$ as $n \to \infty$. Then the ESD of $\frac{1}{\sqrt{n}}W_n'$ converges weakly almost surely to a non-random symmetric probability distribution, say, μ_b.

Our results: One of our goals here is to understand how far the above results can be brought under one umbrella. In Theorem 3.3.1 we formulate a general version where the entries of W_n are assumed to be independent but they need not be identically distributed, and their distribution may change with n, i.e., the distribution of the entries can depend on their positions as well as the size of the matrix. This yields the Results 3.1.1–3.1.8 as special cases. Moreover, results from certain non-homogeneous Erdős-Rényi graphs, matrices with variance profile, certain band matrices, triangular matrices and block matrices are also included.

It is known that the semicircular distribution in Result 3.1.1 sits on the set $NC_2(2k)$ in the sense that each such partition contributes one to the $2k$th moment of the semicircular distribution and any other partition contributes zero. The limit distributions $\mu_\alpha, \mu_{zak}, \mu_{bg}$ and μ_{zhu} in Results 3.1.2–3.1.5 respectively, cannot be described via $NC_2(2k)$ except in special cases. In Theorem 3.3.1 we identify a special class of partitions, called *special symmetric partitions* of $\{1, 2, \ldots, 2k\}$, which contribute to the limit moments. This class has a more complex structure, and includes $NC_2(2k)$ as well as many other crossing and non-crossing partitions. We shall see that the contributions to the moments of the LSD in general vary across the partitions, and depend on specific moment properties of the entries.

3.2 Special Symmetric Partitions and Coloured Rooted Trees

Definition 3.2.1. *(Special Symmetric Partition) Let $\sigma \in \mathcal{P}(k)$ and let $V_1, V_2, \ldots$ be the blocks of the partition, arranged in ascending order of their smallest elements. This partition is said to be special symmetric if*

between any two successive elements of any block there are even number of elements from any other block. (a)

Note that by the stipulated condition, each block is of even size and hence k is necessarily even. We denote the set of special symmetric partitions of $\{1, 2, \ldots, 2k\}$ by $SS(2k)$. We denote by $SS_b(2k)$ the subset of $SS(2k)$ where the partitions have b blocks. Observe that $b \leq k$ always.

For example, the partition $\{\{1, 4, 5, 8\}, \{2, 3, 6, 7\}, \{9, 10\}\}$ of $[10]$ belongs to $SS_3(10)$. The corresponding special symmetric word is abbaabbacc.

The one-block partition of $\{1, 2, \ldots, 2k\}$ is always in $SS(2k)$. It is easy to check that every $\pi \in NC_2(2k)$ is in $SS(2k)$. In fact $SS_k(2k) = NC_2(2k)$, i.e., every special symmetric pair-partition is a non-crossing pair-partition. Moreover, $SS(2k) = NC(2k)$ for $1 \leq k \leq 3$. However, when $k \geq 4$, there are special symmetric partitions that are either crossing or not paired. For example the partition of $[8]$ with blocks $\{1, 2, 5, 6\}$ and $\{3, 4, 7, 8\}$ is a special symmetric partition but is crossing.

Special symmetric words: Recall in Section 2.4, we saw any partition corresponds to a *word*. Hence each special symmetric partition also correspond to a special symmetric word. Special symmetric words are words such that between any two successive appearances of the same letter there are even number of other letters. We will denote the set of these words as $SS(2k)$ and the set of all special symmetric words with b distinct letters as $SS_b(2k)$. For example, $aabbaabb \in SS_2(8)$ and $abcabc$ is not special symmetric.

Now we state a lemma which describes some properties of $SS(2k)$ that are direct consequences from Definition 3.2.1. So we skip the proof.

First we define the notion of a *pure block* which is different from the notion of block.

Definition 3.2.2. *(**Pure block**) Any string of length m $(m > 1)$ of the same letter in a word ω is called a pure block of size m.*

For instance, in $aabbaabbbb$, a occurs in two pure blocks of size 2 each and b occurs in two pure blocks of sizes 2 and 4. The block sizes of a and b is 4 and 6, respectively.

Lemma 3.2.3. *For any special symmetric word,*

(i) each letter appears even number of times,

(ii) the last letter appears in pure even blocks,

(iii) between any two successive appearances of the same letter each of the other letters that do appear, appear equal number of times in an odd and an even position.

Next, we present some more details $SS(2k)$.

A subset of consecutive natural numbers $\{i, i + 1, \ldots, i + h\}$ is called an *interval* in $\mathbb{N}$ and denoted by $[i, i + h]$. For example, $[k]$ and $\{2, 3, 4, 5\} = [2, 5]$ are intervals in $\mathbb{N}$. We define an order on disjoint subsets of $\mathbb{N}$ as follows. For $V_1, V_2 \subset \mathbb{N}$ and $V_1 \cap V_2 = \emptyset$, we say $V_1 < V_2$ if $\min\{j : j \in V_1\} < \min\{j : j \in V_2\}$. For $\sigma \in \mathcal{P}(2k)$, we write σ as $\{V_1, V_2, \ldots, V_r\}$ where $V_1, V_2, \ldots, V_r$ are the blocks of the partition σ and $V_1 < V_2 < \cdots < V_r$. Then we can describe $SS(2k)$ as follows:

(a) For every $k \in \mathbb{N}$, the single block partition $[2k]$ is an element of $SS(2k)$. For $k = 1$, this is the only element of $SS(2)$.

(b) for $k > 1$, a partition $\sigma = \{V_1 < V_2 < \cdots < V_r\}$ of $[2k]$ belongs to $SS(2k)$ if and only if

(i) The *last* block V_r is a union of even sized intervals, say $V_r = I_1 \cup I_2 \cup \cdots \cup I_l$.

(ii) $|\{[\min\ I_j, \max\ I_{j+1}] \setminus (I_j \cup I_{j+1})\} \cap V_i|$ is even for $1 \le i \le r-1$ and $1 \le j \le l$.

(iii) $\sigma \setminus V_r$ can be realized as a special symmetric partition of $[2k] \setminus V_r$.

Remark 3.2.4. *The set $SS(2k)$ does not form a lattice under the inclusion ordering. For example, the partitions $\{\{1,2\},\{3,6\},\{4,5\}\}$ and $\{\{1,6\},\{2,3,4,5\}\}$ both belong to $SS(6)$. However, the largest partition which are smaller than both satisfies*

$$\{\{1,2\},\{3,6\},\{4,5\}\} \wedge \{\{1,6\},\{2,3,4,5\}\} = \{\{1\},\{2\},\{3\},\{4,5\},\{6\}\} \notin SS(6).$$

A graphical approach: Special symmetric partitions (words) can also be shown to be in one-one correspondence with certain *colored rooted ordered trees.*

A colored rooted ordered tree is a graph with no cycles, with one distinguished vertex as the root, and each vertex has a colour that signifies certain properties, where the colours of the vertices are ordered. For example, suppose $\{a_0, a_1, a_2, a_3\}$ is the set of colours with the ordering according to the ordering of their indices, i.e., $a_0 < a_1 < a_2 < a_3$ and T is a coloured rooted ordered tree. Then, the first appearance of the colours occur in ascending order. It is also known as a coloured plane tree. For more on this one can see Stanley [2012].

The moments of the LSD has been described by some authors such as Zakharevich [2006], Zhu [2020] in their works as count of the number of trees or homomorphism densities on them. The next lemma claims $SS(2k)$ is in one-to-one correspondence with certain colored rooted ordered trees. This description gives us some more insight and a different combinatorial perspective for this set of partitions.

Lemma 3.2.5. *Let $\boldsymbol{\omega}$ be a special symmetric word of length $2k$ with b distinct letters, i.e., $\boldsymbol{\omega} \in SS_b(2k)$. Also suppose that each letter appears $2k_1, 2k_2, \ldots, 2k_b$ times respectively in $\boldsymbol{\omega}$. Then there is a coloured rooted ordered tree corresponding to $\boldsymbol{\omega}$ with $(k+1)$ vertices and $(b+1)$ distinct colours $a_0, a_1, \ldots, a_b$ with the following properties:*

(a) *The root is of colour a_0 and there are exactly k_i vertices of colour a_i, $1 \leq i \leq b$.*

(b) *If two vertices are of same colour then their parents are also of same colour.*

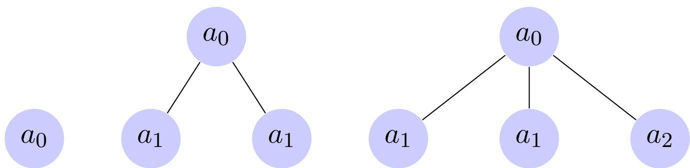
(c) *Vertices with the same colour are at the same distance from the root.*

Also, for every such tree with $k+1$ vertices and $b+1$ distinct colours, there is a unique word that belongs to $SS_b(2k)$ and vice-versa.

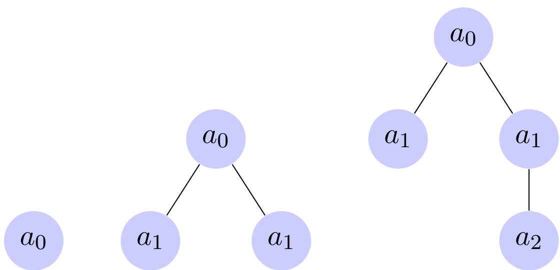
In particular, the Catalan number C_{2k} counts the number of coloured rooted trees with $k+1$ vetrices, each with a distinct colour.

Proof. $SS_b(2k) \to$ **coloured rooted ordered tree with** $(k+1)$ **vertices and** $(b+1)$ **colours**. Suppose $\boldsymbol{\omega} \in SS_b(2k)$ such that each letter appears $2k_1, 2k_2, \ldots, 2k_b$ times respectively. Let $i_1, i_2, \ldots, i_b$ be the positions where the distinct letters made their first appearance in $\boldsymbol{\omega}$. We choose $(b+1)$ distinct colours $a_0, a_1, a_2, \ldots, a_b$ where $a_0 < a_1 < \cdots < a_b$. Now we begin constructing a tree from left to right.

Construction of the tree: Create a root and colour it a_0. The first letter in $\boldsymbol{\omega}$, say, a, has coordinates $(\pi(0), \pi(1))$. For this first appearance of a, we create a child of the root and colour it a_1. For every odd appearance of the letter we put a new child of colour a_1 and for every even appearance, we traverse back to the root. If we get a new letter we have a new child of colour a_2. Now we describe the construction with a couple of examples. Suppose the word is $aaaabb$. Then the tree is constructed as follows:



If the word is $aaabba$, then the tree is constructed as follows:



Continuing this process, we construct the tree until we exhaust all letters of $\boldsymbol{\omega}$. Note that in this construction, all the double edges between two nodes of colours a_s and a_t $(0 \leq s, t \leq b)$ are counted as single edges. This gives us a tree with $(k+1)$ vertices, $(b+1)$ colours, and a root of colour a_0.

Verification that this tree satisfies Properties (a), (b) and (c): Now if x is the jth distinct letter of $\boldsymbol{\omega}$, then it appears $2k_j$ times in $\boldsymbol{\omega}$ out of which every odd time we have a child of colour a_j. So there are k_j vertices of colour a_j. So Property (a) holds for this tree.

Suppose (b) does not hold for the tree constructed. Then there are two vertices of the same colour say, a_m such that their parents are of different colours, say a_{p_1} and a_{p_2} $(p_1 \neq p_2)$. Suppose that the node coloured a_{p_1} appears to the left of the node coloured a_{p_2}. Note that by the construction, we can only get a new node when that particular letter is appearing for the ith time, i being odd. So, we have got the node of colour a_m as a child of a node coloured a_{p_1}, and again another node of colour a_m as a child of a node coloured a_{p_2} when the mth letter, say, x, is appearing for the t_1th and t_2th time, where t_1, t_2 are odd. Therefore in between these two x's there is at least one occurrence where x appears for the ith time, i being even. It can be easily seen that without loss, we can assume that the number of such occurrences is one. Then, clearly by the construction when this x (which is appearing in between the two xs for which we have got the two nodes of colour a_m, with different coloured parent nodes) occurs, we traverse back to the node of colour a_{p_1}, having traversed the subtree starting at the node coloured a_{p_1}. Then getting another node of colour a_m as a child of a_{p_2} after that, implies that the p_2th letter has appeared an odd number of times in between the two x's. This is a contradiction to the fact that $\boldsymbol{\omega} \in SS_b(2k)$. Thus such a thing cannot occur.

Hence Property (b) holds true for a tree constructed above.

Next we verify that Property (c) holds for the tress constructed. Observe that, the root only has colour a_0, and vertices of colour a_1 can appear only as chlidren of the root according to the construction of the tree. Suppose, Property (c) is true for all colours a_i where $1 \leq i \leq j-1$. Now there are k_j vertices of the colour a_j. By the construction, we see that either all of these k_j vertices appear as children of the root or, they appear as children of the vertices of the colour a_t where $t < j$. If all k_j vertices are children of the root, we have nothing to prove, and (c) holds for the colour a_j. In the other case, all of these k_j vertices appear as children of the vertices of the colour a_t, and as $t < j$, all vertices of colour a_t are at the same distance from the root. Therefore, all vertices of colour a_j also are at the same distance from the root. Hence by induction, we have that (c) is true for the tree corresponding to $\boldsymbol{\omega}$. Thus for any word $\boldsymbol{\omega} \in SS_b(2k)$, the $(k+1)$ vertices and $(b+1)$ distinct colours, satisfy Properties (a), (b) and (c).

Further, for any two distinct words $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$ in $SS_b(2k)$, the above process of construction yields two distinct coloured rooted trees with $(k+1)$ vertices and $(b+1)$ distinct colours each with Properties (a), (b) and (c).

Coloured rooted ordered tree with $(k+1)$ vertices and $(b+1)$ colours $\rightarrow SS_b(2k)$. Now, suppose we have a coloured rooted ordered tree with $(k+1)$ vertices and $(b+1)$ distinct colours with properties (a), (b) and (c). We need to show that there is a word in $SS_b(2k)$ corresponding to this tree.

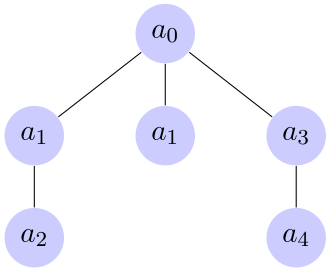
Construction of the word $\boldsymbol{\omega}$: Suppose $a_0, a_1, a_2, \ldots, a_b$ are the distinct colours of the nodes. By (c), there is no other node of the same colour as the root. Suppose the colour of the root is a_0. As there are b distinct colours left, we can associate to each colour a_j the jth distinct letter of the word as follows.

We traverse the tree from left to right in the depth-first way, starting at the root. For every step downward, when we get a vertex of colour a_j, we add the jth distinct letter to the word, and for every step upward to a vertex of colour a_t, we add the tth distinct letter to the word. We repeat this process for all the branches of the tree, left to right.

The first vertex appearing after the root is of colour a_1, which creates the letter a.

If there are no further children of this vertex, then we come back to a_0 (root,) and add the letter a to obtain the partial word aa. We then move to the next right branch. We describe this with the help of an example.

Consider the following tree.



Traversing from the root to the leftmost node of colour a_1, we get the first letter of the word, say, a. From there we traverse (depth first) to the node coloured a_2 and get the second letter, say b. As this node has no further children we traverse back to its parent node. By the construction, this implies that we get b for the second time. Hence the partial word at this point is abb. Then we traverse back to the root, which adds another a to the word. After this, we traverse down to the second child of the root coloured a_1, and (since this node has no further children) back. The word that we have at this point is $abbaaa$. Then we traverse down to the child of the root coloured a_3, that adds the third new letter, say c, to the word. Next, we traverse to its child, which is of a new colour a_4. Hence we get a new letter, say d in the word. The word constructed at this point is $abbaaacd$. Now as the last node has no children, we traverse back to the root via the node of colour a_3. Thus the word we finally get from the above tree is $abbaaacddc$.

Verification of $\boldsymbol{\omega} \in \boldsymbol{SS_b(2k)}$: We now wish to verify that $\boldsymbol{\omega} \in SS_b(2k)$.

Consider two successive appearances of the same letter say, x. If they are side by side, there is nothing to verify. Now suppose they have some other letters in between. Suppose x is the jth distinct letter of $\boldsymbol{\omega}$. Then there are two ways in which we had the first x: (i) while we were going down to the vertex of colour a_j which has a further child or, (ii) while we were coming upward from a vertex of colour a_j.

In Case (i), during the construction, we had gone down that branch and had reached the end of this sub-tree, and each time added a letter to the word for each vertex. Having reached the end, we had started coming upward, and had added those letters in the reverse order to the word, to reach the vertex of colour a_j we started with. Because of Property (c), we cannot get the next x before we reach this vertex. Thus, in between these two successive x's, each letter has been added an even number of times.

In Case (ii), we have got the first x while coming upward from a vertex of colour a_j (to a vertex of color a_t say). Observe that from the Properties (b) and (c) of the tree, the next x can only occur while going downward to another vertex of colour a_j, whose parent must have the colour a_t. By (c), same coloured vertices occur only at the same level of the tree. So to come to this next x, we have to keep traversing the tree from left to right until we reach a vertex of colour a_j from a_t. In this process we pass each intermediate vertex exactly two times, once going up and once going down. Hence each letter can appear only an even number of times in between these two successive x's.

Therefore, $\boldsymbol{\omega} \in SS_b(2k)$. Finally, by construction, it is also clear that different colored rooted ordered trees yield different words.

We now argue the validity of the last claim in the lemma. It is clear from above construction that each Catalan word, i.e., words in $NC_2(2k)$ corresponds to a coloured rooted ordered tree with $k+1$ vertices and $k+1$ distinct colours. Hence the Catalan number counts the number of colored rooted ordered trees with $k+1$ vertices, each vertex having a distinct colour.

This completes the proof of the lemma. $\square$

Remark 3.2.6. *Recall that $NC_2(2k) \subset SS(2k)$. In fact, $SS_k(2k) = NC_2(2k)$. From Lemma 3.2.5, corresponding to each word in $SS_k(2k)$, there is a colored rooted ordered tree with $(k+1)$ vertices and $(k+1)$ colours satisfying Properties (a), (b) and (c). As the number of colours and vertices are equal, each vertex in the tree must be coloured differently. One such coloured rooted ordered tree where each vertex is of different colour can be easily identified with a rooted ordered tree with no colouring. Thus all Catalan words of length $2k$ can be actually described by rooted trees with $(k+1)$ vetrices. Such trees were used to compute the moments of the LSD in Zhu [2020].*

3.3 Main Results

In this section we describe our main LSD result. First we introduce a set of assumptions on the entries $\{x_{ij,n}\}$ of W_n. We drop the suffix n for convenience wherever there is no scope for confusion.

Assumption A $\{g_{k,n}\}_{k\geq 1}$ is a sequence of non-negative, symmetric, bounded Riemann integrable functions on $[0,1]^2$. There exists a sequence $\{r_n\}$ with $r_n \in [0,\infty]$ such that

(i) For each $k \in \mathbb{N}$,

$$n\,\mathbb{E}\left[x_{ij}^{2k}\mathbf{1}_{\{|x_{ij}|\leq r_n\}}\right] = g_{2k,n}\left(\frac{i}{n},\frac{j}{n}\right) \quad \text{for} \ \ 1\leq i\leq j\leq n, \tag{3.3.1}$$

$$\lim_{n\to\infty} n^{\alpha} \sup_{1\leq i\leq j\leq n} \mathbb{E}\left[x_{ij}^{2k-1}\mathbf{1}_{\{|x_{ij}|\leq r_n\}}\right] = 0 \ \ \text{for any}\ \alpha<1. \tag{3.3.2}$$

(ii) The functions $g_{2k,n}, n\geq 1$ converges uniformly to g_{2k} for all $k\geq 1$.

(iii) Let $M_{2k} = \|g_{2k}\|$ (where $\|\cdot\|$ denotes the sup norm) and $M_{2k-1}=0$ for all $k\geq 1$. Then, $\alpha_{2k} = \sum_{\sigma\in\mathcal{P}(2k)} M_\sigma$ satisfy *Carleman's condition*,

$$\sum_{k=1}^{\infty} \alpha_{2k}^{-\frac{1}{2k}} = \infty.$$

Note that the odd sequence of functions, i.e., g_{2k-1} do not make an appearence due to condition (3.3.2). These assumptions hold for most, if not all, models discussed in Section 3.1. We shall have a more detailed discussion on this in Section 3.5. Now we state the main result of this chapter.

Theorem 3.3.1. *Let $W_n = (x_{ij,n})_{1\leq i\leq j\leq n}$ be the $n\times n$ symmetric matrix where $\{x_{ij,n}\,; 1\leq i\leq j\leq n\}$ are independent and satisfy Assumption A. Let $y_{ij,n} = x_{ij}\mathbf{1}_{\{|x_{ij}|\leq r_n\}}$ and $Z_n = (y_{ij,n})_{1\leq i\leq j\leq n}$. Then*

(a) the ESD of Z_n converges weakly almost surely to μ' say, whose odd moments are zero and even moments are determined by the functions g_{2k}, $k\geq 1$ as follows:

$$\beta_{2k}(\mu') = \sum_{b=1}^{k} \sum_{\sigma\in SS_b(2k)} \int_{[0,1]^{b+1}} \prod_{j=1}^{b} g_{|V_j|}(x_{t_j}, x_{l_j}) \prod_{j=0}^{b} dx_{l_j} \tag{3.3.3}$$

where $\prod_{j=0}^{b} dx_{l_j}$ (with $l_0 = t_1 = 0$) denotes the $(b+1)-$dimensional Lebesgue measure on $[0,1]^{b+1}$ and (t_j, l_j) are indices corresponding to non-generating and generating vertices of each $\sigma\in SS_b(2k)$.

(b) Further if

$$\frac{1}{n}\sum_{i,j} x_{ij}^2\mathbf{1}_{\{|x_{ij}|>r_n\}} \to 0, \ \ \textit{almost surely (respectively in probability)}, \tag{3.3.4}$$

then the ESD of W_n converges weakly to μ' almost surely (respectively in probability).

In particular μ' is semicircular if and only if $g_{2k} = 0$ for all $k > 1$, and $\int_0^1 g_2(x, y)\, dy$ is a constant.

Note: It is to be noted that in (3.3.3), as the t_js correspond to the non-generating vertex for the jth block (letter), $t_j = l_i$ for some $i < j$. Hence the integral is on $\{l_j\}$ variables. We shall describe this in details in the proof of the theorem.

Remark 3.3.2. *If $\{x_{ij}\}$ has all moments finite then choosing $r_n = \infty$, we have $Z_n = W_n$, and then the theorem concludes that the ESD of W_n converges almost surely to μ'.*

Recall that the (standard) semicircle law in Result 3.1.1 has the support $[-2, 2]$. In contrast the limit μ' in Theorem 3.3.1 can have unbounded support.

Proposition 3.3.3. *(**Unbounded support**) Let $f_{2m}(x) = \int_{[0,1]} g_{2m}(x, y)\, dy$ for each $m \geq 1$. Suppose that there exist an $m > 1$ such that $\inf_{t\geq 1} \int_{[0,1]} \left(\frac{f_{2m}(x)}{m!}\right)^t dx = \alpha > 0$. Then the LSD μ' in Theorem 3.3.1 has unbounded support.*

In particular, when $g_{2k} \equiv c_{2k}$ for all $k \geq 1$, and if for some $m > 1$, $c_{2m} > \delta > 0$, then the LSD μ' has unbounded support.

Remark 3.3.4. *The convergence in the truncation condition, (3.3.4), can occur either in probability and not almost surely. For instance, let $X_2, \ldots, X_n, \ldots$ be a sequence of independent random variables. For each $n \geq 2$,*

$$\mathbb{P}[X_n = 1] = 1 - \frac{1}{n^2 \ln n} \tag{3.3.5}$$

$$\mathbb{P}[X_n = n^{1/2}] = \frac{1}{n^2 \ln n}. \tag{3.3.6}$$

Let $\mathbb{P}[X_1 = 0] = 1$. Let $x_{ij,n} \stackrel{\mathcal{D}}{=} X_n$ for every fixed n. Then the convergence in (3.3.4) is in probability and not almost surely.

3.4 Proof of Theorem 3.3.1

We first present a few lemmas that lead to the proof of Theorem 3.3.1. In Lemmas 3.4.2 and 3.4.6, we identify the set of words that can possibly contribute to the limiting moments. Then in Lemma 3.4.7, we prove a combinatorial inequality that helps us to obtain the almost sure convergence of the ESD. Finally we prove Theorem 3.3.1.

First, recall the notion of link function, words and circuits from Section 2.4. Recall the link function L_W for the Wigner matrix from Section 2.4. If for i, j, $\boldsymbol{\omega}[i] = \boldsymbol{\omega}[j]$, then $\big(\pi(i-1), \pi(i)\big) = \big(\pi(j-1), \pi(j)\big)$ as *unordered pairs.*

Definition 3.4.1. *(($C1$) and ($C2$) constraints) Suppose that for a word $\boldsymbol{\omega}$ arising from the link function L_W and $\boldsymbol{\omega}[i] = \boldsymbol{\omega}[j]$. Then for any n, ($C1$) and ($C2$) constraints are defined on any $\pi \in \Pi_n(\boldsymbol{\omega})$ as ordered pairs*

$$\big(\pi(i-1), \pi(i)\big) = \begin{cases} \big(\pi(j-1), \pi(j)\big) & \textit{((C1) constraint) or,} \\ \big(\pi(j), \pi(j-1)\big) & \textit{((C2) constraint).} \end{cases} \tag{3.4.1}$$

For a word $\boldsymbol{\omega}$ with b distinct letters, suppose $i_1, i_2, \ldots, i_b$ are the positions where new letters made their first appearances in $\boldsymbol{\omega}$. Denote by $\mathcal{E}_{i_j}$ the partition block where $\pi(i_j)$ belongs. Also let $\mathcal{E}_0$ denote the partition block where $\pi(0)$ belongs. $\pi(i_j)$ will be said to be the representative of the partition block $\mathcal{E}_{i_j}$. Note that any two such blocks are either equal or disjoint. For example for the word *abcabc*, $\mathcal{E}_0 = \{\pi(0), \pi(3), \pi(6)\}$, $\mathcal{E}_1 = \{\pi(1), \pi(4)\}$, $\mathcal{E}_2 = \{\pi(2), \pi(5)\}$ and $\mathcal{E}_3 = \mathcal{E}_0$.

Note that if the number of partition blocks is exactly equal to $b+1$ then

$$\lim_{n\to\infty} \frac{|\Pi(\boldsymbol{\omega})|}{n^{b+1}} = 1. \tag{3.4.2}$$

On the other hand, if the number of partitions blocks is strictly less than $b+1$ then $|\Pi(\boldsymbol{\omega})| \le n^b$ and hence

$$\lim_{n\to\infty} \frac{|\Pi(\boldsymbol{\omega})|}{n^{b+1}} = 0. \tag{3.4.3}$$

We shall now investigate for which words either of (3.4.2) or (3.4.3) happens. This is given in Lemmas 3.4.2 and 3.4.6.

Identification of words that may contribute: In the following two lemmas, we find out the words that contribute to the moments of the LSD in Theorem 3.3.1.

Lemma 3.4.2. *Let $\boldsymbol{\omega}$ have b distinct letters and satisfy*

$$\frac{|\Pi(\boldsymbol{\omega})|}{n^{b+1}} \to 1 \quad \text{as } n \to \infty. \tag{3.4.4}$$

Then $\boldsymbol{\omega}$ is a special symmetric word (as in Definition 3.2.1).

Proof. From the discussion before the lemma, there are $(b+1)$ partition blocks, $\mathcal{E}_0, \mathcal{E}_{i_1}, \ldots, \mathcal{E}_{i_b}$.

Fix any letter x in $\boldsymbol{\omega}$. Then there exists $t \in \{1, 2, \ldots, b\}$ such that x first appears at the i_tth position whose coordinates in terms of circuits are $(\pi(i_t - 1), \pi(i_t))$. Then we will say that x is *associated* with $\pi(i_t)$ and the partition block $\mathcal{E}_{i_t}$. Also we can say that $\pi(i_t - 1) \in \mathcal{E}_{i_s}$ for some $s < t$.

To show that $\boldsymbol{\omega} \in SS_b(2k)$, we shall make use of the following claims.

Claim 3.4.3. *Suppose $\boldsymbol{\omega}$ satisfies (3.4.4). Then, the successive appearances of the same letter obey the $(C2)$ constraint.*

Claim 3.4.4. *Suppose $\boldsymbol{\omega}$ satisfies (3.4.4). Suppose $x_1, x_2, \ldots, x_m$ appear in between two successive appearances of the same letter, say y. Then, all x_i cannot be distinct.*

We shall prove these claims first.

Proof of Claim 3.4.3. We show this first for the last new letter of $\boldsymbol{\omega}$ and then for the others. We will show that the last new letter of $\boldsymbol{\omega}$ appears in pure blocks of even size.

Suppose z is the last new letter of $\boldsymbol{\omega}$ and appears at $(\pi(i_b - 1), \pi(i_b))$th position in $\boldsymbol{\omega}$ for the *first time*. Suppose this z is followed by $x \neq z$. So x is an old letter, say the jth new letter of $\boldsymbol{\omega}$ where $j < b$. Then clearly either
(a) $\pi(i_b) = \pi(i_j)$, that is, $\pi(i_b) \in \mathcal{E}_{i_j}$, or
(b) $\pi(i_b) = \pi(i_j - 1)$, that is, $\pi(i_b) \in \mathcal{E}_{i_t}$ (where $\pi(i_j - 1) \in \mathcal{E}_{i_t}$) for some $t < j$.
In any case, $\mathcal{E}_{i_b}$ coincides with $\mathcal{E}_{i_m}$ for some $m \neq b$. As a result $|\Pi(\boldsymbol{\omega})| \leq n^b$ and $\boldsymbol{\omega}$ does not satisfy (3.4.4). So, z has to be followed by itself, that is, it appears in a pure block.

Now suppose, this pure block size is odd. Consider the last z that appears in this block at $(\pi(i_b+s-1), \pi(i_b+s))$ where s is even. Then

$$\pi(i_b-1) = \pi(i_b+1) = \pi(i_b+3) = \cdots = \pi(i_b+s-1),$$
$$\pi(i_b) = \pi(i_b+2) = \pi(i_b+4) = \cdots = \pi(i_b+s).$$

Therefore $\pi(i_b+s) \in \mathcal{E}_{i_b}$. If this z is followed by an old letter then by the same arguments as given earlier, we arrive at a contradiction. Hence this block size has to be even and s is odd.

If z does not appear elsewhere, we are done. So suppose after this block, z appears again next at the $(\pi(t-1), \pi(t))$th position where $t-(i_b+s) > 1$. Then there are two possibilities: either (a) $\{\pi(t-1) = \pi(i_b), \pi(t) = \pi(i_b-1)\}$ or (b) $\{\pi(t-1) = \pi(i_b-1), \pi(t) = \pi(i_b)\}$. As an old letter appears at the $(\pi(t-2), \pi(t-1))$th position, if (a) happens, then by the argument as before, $\mathcal{E}_{i_b}$ coincides with one of the other partition blocks $\mathcal{E}_0 = \mathcal{E}_{i_0}, \ldots, \mathcal{E}_{i_{b-1}}$. This contradicts (3.4.4). So, the only possibility is (b), i.e., these two z's satisfy the $(C2)$ constraint. Now we are back to the same situation as that we had for the first appearance of z in the first block of z's. Repeating that argument, it follows that this z is also followed by an odd number of z. Same argument holds for all blocks. This shows that successive appearances of two z's obey the $(C2)$ constraint and z appears in even pure blocks.

Now we drop all these z's and consider the reduced word $\boldsymbol{\omega}'$. We have already seen that successive appearances of two z's obey the $(C2)$ constraint and z appears in even pure blocks. Thus $\pi(i_b)$ is the only generating vertex that has been dropped in the process of reduction. Also all the vertices in the partition block $\mathcal{E}_{i_b}$ are dropped. Therefore $\boldsymbol{\omega}'$ has b distinct partition blocks (as only one of the partition block is reduced), $\mathcal{E}_{i_0}, \mathcal{E}_{i_1}, \mathcal{E}_{i_2}, \ldots, \mathcal{E}_{i_{b-1}}$.

Now let y be the last new letter in $\boldsymbol{\omega}'$. As $\boldsymbol{\omega}'$ satisfies $\lim_{n\to\infty} \frac{|\Pi(\boldsymbol{\omega}')|}{n^b} = 1$, the argument given earlier for z can be repeated to show that y must appear in pure even blocks and any two successive appearances of y must obey the $(C2)$ constraint.

Clearly this argument can be repeated sequentially to complete the proof of the claim.

Proof of Claim 3.4.4. Let y be any letter of $\boldsymbol{\omega}$ with successive occurrences at $(\pi(s), \pi(s+1))$ and $(\pi(s+m+1), \pi(s+m+2))$th positions. Let $x_1, x_2, \ldots, x_m$ be the letters in between these y's, all distinct. Note that x_i appears at $(\pi(s+i), \pi(s+i+1))$th position. Suppose $\pi(s+1) \in \mathcal{E}_{i_j}$. Now there are two cases:

Case 1: Suppose $\pi(s+2)$ belongs to the partition block associated to x_1. As x_1 and x_2 are distinct, $\pi(s+2)$ does not belong to the partition block associated to x_2. Therefore, $\pi(s+3)$ belongs to the partition block associated to x_2. Again as x_3 is different from x_2, $\pi(s+4)$ belongs to the partition block associated to x_3. Therefore, repeating this argument for all x_i's, we see that for each $1 \leq i \leq m$, $\pi(s+i+1)$ belongs to the partition block associated to x_i. Now we know that $\pi(s+1) = \pi(s+m+1)$ since the two appearances of y are in $(C2)$. Therefore $\pi(s+m+1) \in \mathcal{E}_{i_j}$ for some j. Also, $\pi(s+m+1)$ belongs to the partition block associated to x_m. Thus the partition block associated to x_m coincides with $\mathcal{E}_{i_j}$. Now for $\boldsymbol{\omega}$ to satisfy (3.4.4), x_m must be the jth new letter of $\boldsymbol{\omega}$. Therefore, $\pi(s+m+1) = \pi(i_j)$ and $\pi(s+m) = \pi(i_j - 1)$. Hence we have $\pi(s+m) \in \mathcal{E}_{i_{t_1}}$, where $t_1 < j$. Proceeding in this manner we have

$$
\begin{aligned}
&\pi(s+m-1) \in \mathcal{E}_{i_{t_2}} \quad \text{where} \quad t_2 < t_1; \\
&\pi(s+m-2) \in \mathcal{E}_{i_{t_3}} \quad \text{where} \quad t_3 < t_2; \\
&\;\cdots \;\cdots \;\cdots \;\cdots \;\cdots \;\cdots \;\cdots \;\cdots; \\
&\pi(s+2) \in \mathcal{E}_{i_{t_{m-1}}} \quad \text{where} \quad t_{m-1} < t_{m-2}; \\
&\pi(s+1) \in \mathcal{E}_{i_{t_m}} \quad \text{where} \quad t_m < t_{m-1} < j.
\end{aligned}
$$

This shows that $\mathcal{E}_{i_j}$ and $\mathcal{E}_{i_{t_m}}$ coincide for some $t_m \neq j$. So in this case, $|\Pi(\omega)| \leq n^b$, which contradicts (3.4.4).

Case 2: Suppose $\pi(s+1)$ belongs to the partition block associated to x_1. Now we know that $\pi(s+1) = \pi(s+m+1)$ since the two appearances of y are in $(C2)$. As x_1 and x_m are distinct, $\pi(s+m+1)$ does not belong to the partition block associated to x_m. Thus, $\pi(s+m)$ belongs to the partition block associated to x_m. So, $\pi(s+m-1)$ belongs to the partition block associated to x_{m-1}. Therefore, we see that for each $1 \leq i \leq m$, $\pi(s+i)$ belongs to the partition block associated to x_i.

We have observed that $\pi(s+m)$ belongs to the partition block associated to x_m. We know that $\pi(s+m) = \pi(i_{t_1})$ for some $t_1 \in \{1, 2, \ldots, b\}$. Then, $\pi(s+m+1) = \pi(i_{t_1} - 1)$. Now recall that $\pi(s+1) = \pi(s+m+1)$ and $\pi(s+1) \in \mathcal{E}_{i_j}$. So, $\pi(s+m+1) \in \mathcal{E}_{i_j}$. Therefore, $\pi(i_{t_1} - 1) \in \mathcal{E}_{i_j}$. For the partition block associated to each letter to be distinct, we must have, $\pi(s+m) \in \mathcal{E}_{i_{t_1}}$ where $t_1 > j$. Proceeding in this manner we have

$$\begin{aligned}
&\pi(s+m-1) \in \mathcal{E}_{i_{t_2}} \quad \text{where} \quad t_2 > t_1;\\
&\pi(s+m-2) \in \mathcal{E}_{i_{t_3}} \quad \text{where} \quad t_3 > t_2;\\
&\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots;\\
&\pi(s+2) \in \mathcal{E}_{i_{t_{m-1}}} \quad \text{where} \quad t_{m-1} > t_{m-2};\\
&\pi(s+1) \in \mathcal{E}_{i_{t_m}} \quad \text{where} \quad t_m > t_{m-1} > j.
\end{aligned}$$

This shows that $\mathcal{E}_{i_j} \cap \mathcal{E}_{i_{t_m}} \neq \phi$ for some $t_m \neq j$. But then $|\Pi(\omega)| \leq n^b$ and it contradicts (3.4.4). This completes the proof of Claim 3.4.4.

Now we prove the lemma using Claims 3.4.3 and 3.4.4.

Let the letter x occur sucessively at $(\pi(s), \pi(s+1))$ and $(\pi(s+m+1), \pi(s+m+2))$th positions and let us focus on the string of letters $z_1, z_2, \ldots, z_m$ in between two successive appearances of x. We shall show that m is even, and each distinct z_j appears an even number of times among these m places.

By Claim 3.4.4, there exists j such that z_j has appeared at least twice. For any such letter z_j, let p_j be the second last position of its appearance (in between the two x's considered). Let $p = \max\{p_l : z_l \text{ has appeared more than once in between the two } x's\}$. There is some z_i such that $p_i = p$. That is, z_i appears for the second last time at the pth position. Consider the z_i at the pth position and the last z_i (in between the two x's considered). By our choice of p, no letter can appear more than once in between. Now we invoke Claim 3.4.4 to conclude that there are no other letters in between these z_i's. So these two z_i's appear in consecutive positions. Now we drop this pair of z_i's. If this z_i does not appear elsewhere in the word then the partition block corresponding to z_i is dropped, and for the reduced word $\boldsymbol{\omega}'$ with $(b-1)$ distinct letters, $\lim_{n\to\infty} \frac{|\Pi(\boldsymbol{\omega}')|}{n^b} = 1$. So we repeat the process.

If this z_i appears somewhere else in $\boldsymbol{\omega}$, then the reduced word $\boldsymbol{\omega}'$ still has b distinct letters and $\lim_{n\to\infty} \frac{|\Pi(\boldsymbol{\omega}')|}{n^{b+1}} = 1$. This is because none of the partition blocks have been dropped or have coincided in the process of dropping the two z_i's, as this pair of z_i are in a $(C2)$ relation (by Claim 3.4.3), and as a result only one member of a partition block has been omitted. So here again $\boldsymbol{\omega}'$ retains the properties of $\boldsymbol{\omega}$ and we may repeat the process. Continuing this process, either there are no letters left in between the x's or all letters that remain are distinct. But the latter is not possible due to Claim 3.4.4.

Hence, we conclude that $\boldsymbol{\omega}$ is a special symmetric word. $\square$

The following corollary follows from Lemma 3.4.2.

Corollary 3.4.5. *Suppose the word $\boldsymbol{\omega}$ with b distinct letters does not belong to $SS(2k)$ for any k. Then $|\Pi(\boldsymbol{\omega})| \leq n^b$. Hence $\lim_{n\to\infty} \frac{1}{n^{b+1}}|\Pi(\boldsymbol{\omega})| = 0$.*

We now establish the converse of Lemma 3.4.2. This will identify the words which contribute to the limiting spectral distribution of W_n.

Lemma 3.4.6. *Suppose $\boldsymbol{\omega} \in SS(2k)$ has b distinct letters. Then, $\lim_{n\to\infty} \frac{1}{n^{b+1}}|\Pi(\boldsymbol{\omega})| = 1$.*

Proof. We use induction on the number of distinct letters to prove this lemma. First note that the length of the word is even.

When $b = 1$, all letters are identical. In this case $\pi(0)$ and $\pi(1)$ can be chosen in $n \times (n-1) \approx n^2$ ways. Given these vertices, all other vertices have exactly one choice each. Therefore, $|\Pi(\boldsymbol{\omega})| \approx n^2$ and hence the result is true for $b = 1$.

Now suppose, that the result is true for words with $(b-1)$ distinct letters. We shall prove the result for all words with b distinct letters.

As we saw in the proof of Claim 3.4.3, the last new letter of the word, say z, appears in pure even blocks. Drop these blocks. Now the reduced word $\boldsymbol{\omega}'$ has $(b-1)$ distinct letters. For the reduced word $\boldsymbol{\omega}'$, property (a) of Definition 3.2.1 is clearly preserved. Thus we are able to apply our induction hypothesis on $\boldsymbol{\omega}'$.

After we have dropped the z's from $\boldsymbol{\omega}$, y's have to appear in pure even blocks in $\boldsymbol{\omega}'$. Therefore, $\boldsymbol{\omega}'$ satisfies property (a) of Definition 3.2.1 again. So, by induction hypothesis, $\lim_{n\to\infty} \frac{|\Pi(\boldsymbol{\omega}')|}{n^b} = 1$. The generating vertex of z in $\boldsymbol{\omega}$ has $\approx n$ choices. Therefore, $\lim_{n\to\infty} \frac{1}{n^{b+1}}|\Pi(\boldsymbol{\omega})| = 1$. This completes the proof of the lemma. $\square$

Handling almost sure convergence: As mentioned earlier, we shall use moment method to prove the almost sure convergence of the ESD and hence shall take help of Lemma 2.1.3. To verify the fourth moment condition, consider four circuits π_1, π_2, π_3 and π_4 that are cross-matched. Suppose we put a new letter wherever a new edge (or $L-$value) appears across all the circuits π_1, π_2, π_3 and π_4. As the circuits are cross-matched, out of the $4k$ places across $(\pi_i)_{1\leq i\leq 4}$, there can be at most $2k$ distinct edges or distinct letters. It suffices to have a bound on the cardinality of the following:

$$Q_{k,4}^b = \{(\pi_1, \pi_2, \pi_3, \pi_4) : \ell(\pi_i) = k; \pi_i, 1 \leq i \leq 4 \text{ jointly- and cross-matched with}$$
$$b \text{ distinct edges or } b \text{ distinct letters across all } (\pi_i)_{1\leq i\leq 4}\}.$$

Note: In $Q_{k,4}^b$ the quadruple of circuits have b distinct edges and hence b distinct letters.

Lemma 3.4.7. *There exists a constant C, such that,*

$$|Q_{k,4}^b| \leq C\ n^{b+2}\ . \tag{3.4.5}$$

Proof. First observe that for any circuit π, if we set aside the first vertex $\pi(0)$, then the number of choices for the generating vertices is $\approx n^b$, where b is the number of distinct letters in π. Once all the generating vertices have been chosen, the number of choices for the non-generating vertices is at most one. This observation will be used repeatedly.

Consider all circuits $(\pi_1, \pi_2, \pi_3, \pi_4)$ of length k which are jointly-matched and cross-matched with b distinct letters. Let the number of new distinct letters appearing in π_i be k_i, $i \in \{1, 2, 3, 4\}$. So clearly, $k_1 + k_2 + k_3 + k_4 = b$. We begin with the circuit π_1 and its vertices can be chosen in at most n^{k_1+1} ways. Now, since the circuits are cross-matched, there is another circuit with which π_1 shares a letter. So we have the following three cases:

Case 1: π_1 shares a letter with only one of the circuits, say π_2. Then, without loss of generality (since we are dealing with circuits) we can assume that π_2 begins with the letter it shares with π_1. Thus, $\pi_2(0)$ and $\pi_2(1)$ both cannot be chosen freely. Hence, having chosen the generating vertices of π_1, choosing from left to right, the vertices of π_2 can be chosen in at most n^{k_2} ways. The generating vertices of π_3 can be chosen in at most n^{k_3+1} ways. Now, since π_4 does not share any letter with π_1, it must share at least one letter with either π_2 or π_3. Again, we can assume that π_4 begins with this letter. Thus, having chosen the generating vertices of π_1, π_2 and π_3, the vertices $\pi_4(0)$ and $\pi_4(1)$ cannot be chosen freely, and the number of choices for all generating vertices of π_4 is at most n^{k_4}. Therefore, all vertices of $(\pi_1, \pi_2, \pi_3, \pi_4)$ can be chosen in at most $n^{(k_1+k_2+k_3+k_4)+2} = n^{b+2}$ ways.

Case 2: π_1 shares a letter with exactly two circuits, say π_2 and π_3. Then again we can assume that π_2 and π_3 begin with the letters that they share with π_1. Thus, $\pi_2(0), \pi_2(1), \pi_3(0), \pi_3(1)$ cannot be chosen freely. Hence, choosing from left to right, the generating vertices of $\pi_1(j)$, $\pi_2(j)$, $\pi_3(j)$ can be chosen in at most in $n^{(k_1+1)+k_2+k_3}$ ways. Now, π_4 shares a letter with either π_2 or π_3. Again, we can assume that π_4 begins with this letter. Hence $\pi_4(0)$ and $\pi_4(1)$ cannot be chosen freely, and so the generating vertices of $\pi_4(j)$ can be chosen in at most n^{k_4} ways. Therefore, all vertices of $(\pi_1, \pi_2, \pi_3, \pi_4)$ can be chosen in at most $n^{(k_1+k_2+k_3+k_4)+2} = n^{b+2}$ ways.

Case 3: π_1 shares a letter with all the other three circuits. Then, arguing as in the other cases, the number of choices is now at most $n^{(k_1+1)+k_2+k_3+k_4} = n^{b+1}$.

This completes the proof of the lemma. $\square$

Proof of Theorem 3.3.1. We break the proof into six steps.

Step 1 (Reduction to the case where all the entries of Z_n have mean 0): To see this, consider the matrix $\widetilde{Z}_n$ whose entries are $(y_{ij} - \mathbb{E}y_{ij})$. Clearly the entries of $\widetilde{Z}_n$ have mean 0. Now

$$n\ \mathbb{E}[(y_{ij} - \mathbb{E}y_{ij})^{2k}] = n\ \mathbb{E}[y_{ij}^{2k}] + n \sum_{t=0}^{2k-1} \binom{2k}{t} \mathbb{E}[y_{ij}^t]\ (\mathbb{E}y_{ij})^{2k-t}. \tag{3.4.6}$$

The first term of the r.h.s. is equal to $g_{2k}(i/n, j/n)$ by (3.3.1). For the second term we argue as follows:

$$\begin{aligned}\text{For } t \neq 2k-1, \quad n\ \mathbb{E}[y_{ij}^t]\ (\mathbb{E}y_{ij})^{2k-t} &= (n^{\frac{1}{2k-t}}\ \mathbb{E}y_{ij})^{2k-t}\ \mathbb{E}[y_{ij}^t] \\ &\stackrel{n\to\infty}{\longrightarrow} 0, \quad \text{uniformly} \quad \text{by Condition (3.3.2).}\end{aligned}$$

$$\text{For } t = 2k-1, \quad n\,\mathbb{E}[y_{ij}^{2k-1}]\,\mathbb{E}y_{ij} = (\sqrt{n}\,\mathbb{E}[y_{ij}^{2k-1}])\,(\sqrt{n}\,\mathbb{E}y_{ij}) \xrightarrow{n\to\infty} 0, \quad \text{uniformly} \quad \text{by Condition (3.3.2)}.$$

Hence from (3.4.6), we see Condition (3.3.1) is true for the matrix $\widetilde{Z}_n$ with a modified sequence $\tilde{g}_{2k,n}$ that still converges uniformly to g_{2k}. However, for ease of notation, we will continue to call this sequence of functions as $g_{2k,n}$. Similarly we can show that (3.3.2) is true for $\widetilde{Z}_n$. Hence, Assumption A holds for the matrix $\widetilde{Z}_n$.

Now observe that

$$\begin{aligned} d_2^2(\mu_{Z_n}, \mu_{\widetilde{Z}_n}) &\leq \frac{1}{n}\sum_{i,j}(\mathbb{E}y_{ij})^2 \\ &\leq n\,(\sup_{i,j}\,\mathbb{E}y_{ij})^2. \\ &= (\sup_{i,j}\,\sqrt{n}\,\mathbb{E}y_{ij})^2 \xrightarrow{n\to\infty} 0, \qquad \text{by Condition (3.3.2)}. \end{aligned}$$

Hence the LSD of Z_n and $\widetilde{Z}_n$ are almost surely same. Hence we can assume the entries have mean zero.

Now we shall prove part (a) of Theorem 3.3.1 by verifying Conditions (i), (ii) and (iii) of Lemma 2.1.3 using Assumption A and a few other observations made earlier.

Step 2 (Verification of the fourth moment condition, (ii) of Lemma 2.1.3 for Z_n): We show that

$$\frac{1}{n^4}\mathbb{E}[\mathrm{Tr}(Z_n^k) \;-\; \mathbb{E}\,(\mathrm{Tr}(Z_n^k))]^4 \;= \mathcal{O}(n^{-\frac{3}{2}}). \tag{3.4.7}$$

Observe that

$$\frac{1}{n^4}\mathbb{E}[\mathrm{Tr}(Z_n^k) \;-\; \mathbb{E}(\mathrm{Tr}(Z_n^k))]^4 \;= \frac{1}{n^4}\sum_{\pi_1,\pi_2,\pi_3,\pi_4}\mathbb{E}[\Pi_{i=1}^4(Y_{\pi_i} \;-\; \mathbb{E}Y_{\pi_i})]. \tag{3.4.8}$$

If $(\pi_1,\pi_2,\pi_3,\pi_4)$ are not jointly-matched, then one of the circuits has a letter that does not appear elsewhere. Hence by independence, and mean zero assumption, we have $\mathbb{E}[\Pi_{i=1}^4(Y_{\pi_i} \;-\; \mathbb{E}Y_{\pi_i})] = 0$.

Now, if $(\pi_1, \pi_2, \pi_3, \pi_4)$ are not cross-matched, then one of the circuits say π_j is only self-matched. Then, we have $\mathbb{E}[Y_{\pi_j} - \mathbb{E}Y_{\pi_j}] = 0$. So, again $\mathbb{E}[\Pi_{i=1}^4 (Y_{\pi_i} - \mathbb{E}Y_{\pi_i})] = 0$. Thus, we need to consider only circuits $(\pi_1, \pi_2, \pi_3, \pi_4)$ that are jointly- and cross-matched. Also observe that it is enough to prove the bound for $\frac{1}{n^4} \sum_{\pi_1,\pi_2,\pi_3,\pi_4} \mathbb{E}[\Pi_{i=1}^4 Y_{\pi_i}]$. This is because the other terms have more factors and as a result have more $\frac{1}{n^\delta}, \delta > 0$ in the denominator with bounded terms in the numerator, making them significantly smaller than $\frac{1}{n^4} \sum_{\pi_1,\pi_2,\pi_3,\pi_4} \mathbb{E}[\Pi_{i=1}^4 Y_{\pi_i}]$. Now suppose π_i has k_i new distinct letters (that have not appeared in the circuits $\pi_j, j < i$) for each $1 \le i \le 4$ where $k_1 + k_2 + k_3 + k_4 = b$. Suppose the jth letter appears s_j times across $\pi_1, \pi_2, \pi_3, \pi_4$. Now the s_i's might be odd or even. Without loss of generality assume that among the s_i's there are $s_{i_1}, s_{i_2}, \ldots, s_{i_{b_1}}$ which are even and $s_{i_{b_1+1}}, s_{i_{b_1+2}}, \ldots, s_{i_{b_2}}$ which are odd where b_1 and b_2 are any two numbers adding up to b. Then, each term in the sum $\frac{1}{n^4} \sum_{\pi_1,\pi_2,\pi_3,\pi_4} \mathbb{E}[\Pi_{i=1}^4 Y_{\pi_i}]$ is

$$\frac{1}{n^4} n^{-b_1} n^{-(b_2 - \frac{1}{2})} \prod_{j=1}^{b_1} g_{s_{i_j},n}(\pi(i_j - 1)/n, \pi(i_j)/n) \prod_{m=b_1+1}^{b_1+b_2} n^{\frac{b_2-(1-1/2)}{b_2}} \mathbb{E}[y_{\pi(i_m-1)\pi(i_m)}^{s_{i_m}}].$$

Note that $g_{s_{i_j},n} \to g_{s_{i_j}}$ for all $1 \le j \le b_1$. Hence the sequence $\|g_{s_{i_j},n}\|$ is bounded by a constant M_j. Also as $\frac{b_2-(1-1/2)}{b_2} < 1$, by (3.3.2), we have $n^{\frac{b_2-(1-1/2)}{b_2}} \mathbb{E}[y_{\pi(i_m-1)\pi(i_m)}^{s_{i_m}}]$ is bounded by 1 for n large when $b_1 + 1 \le m \le b_1 + b_2$. Let

$$M' = \max_{b_1+b_2=b} \{M_t, 1 : 1 \le t \le b_1\} \quad \text{and} \quad M_0' = \max\{{M'}^b : 1 \le b \le 2k\}.$$

By Lemma 3.4.7, we have the total number of such circuits is of the order of n^{b+2}. Therefore we have

$$\begin{aligned} \frac{1}{n^4} \mathbb{E}[\mathrm{Tr}(Z_n^k) - \mathbb{E}(\mathrm{Tr}(Z_n^k))]^4 &\le M_0' \sum_{b=1}^{2k} \frac{1}{n^{b+3\frac{1}{2}}} n^{b+2} \\ &= \mathcal{O}(n^{-\frac{3}{2}}). \end{aligned}$$

This completes the proof of (3.4.7).

By Lemma 2.1.3 and (3.4.7), it is now enough to show that for every $k \ge 1$, $\lim_{n\to\infty} \frac{1}{n} \mathbb{E}[\mathrm{Tr}(Z_n)^k]$ exists and is given by $\beta_k(\mu')$ for each $k \ge 1$.

Step 3 (Verification of first moment condition, (i) of Lemma 2.1.3 for Z_n): Note that from (2.4.4) and using the fact that $\mathbb{E}(y_{ij}) = 0$, we have

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\operatorname{Tr}(Z_n)^k] &= \lim_{n\to\infty} \frac{1}{n} \sum_{\pi:\ell(\pi)=k} \mathbb{E}[Y_\pi] \\ &= \lim_{n\to\infty} \sum_{b=1}^{k} \frac{1}{n} \sum_{\substack{\omega \text{ matched} \\ \text{with b distinct letters}}} \sum_{\pi\in\Pi(\boldsymbol{\omega})} \mathbb{E}(Y_\pi) \qquad (3.4.9)\end{aligned}$$

Let $\boldsymbol{\omega}$ be a matched word with b distinct letters and let $\pi \in \Pi(\boldsymbol{\omega})$. Suppose the first appearance of the letters of $\boldsymbol{\omega}$ are at the $i_1, i_2, \ldots, i_b$ positions. Thus, the jth new letter appears at the $(\pi(i_j - 1), \pi(i_j))$th position for the first time. Recall the partition blocks $\mathcal{E}_0, \mathcal{E}_{i_1}, \mathcal{E}_{i_2}, \ldots, \mathcal{E}_{i_b}$ as described after Definition 3.4.1.

From the proofs of Lemma 3.4.2 and Lemma 3.4.6, we know that all the $\mathcal{E}_{i_j}$ $(0 \le j \le b)$ are distinct if and only if $\boldsymbol{\omega} \in SS_b(2k)$.

So write (3.4.9) as

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\operatorname{Tr}(Z_n)^k] &= \lim_{n\to\infty} \sum_{b=1}^{k} \Big[\frac{1}{n} \sum_{\omega\in SS_b(k)} \sum_{\pi\in\Pi(\boldsymbol{\omega})} \mathbb{E}(Y_\pi) + \frac{1}{n} \sum_{\substack{\omega\notin SS(k) \\ \boldsymbol{\omega} \text{ with b distinct letters}}} \sum_{\pi\in\Pi(\boldsymbol{\omega})} \mathbb{E}(Y_\pi)\Big]. \\ &= T_1 + T_2. \qquad (3.4.10)\end{aligned}$$

Clearly T_1 is the term involving all the special symmetric partitions. This will be shown to contribute positively to the limit. The sum of contributions of all other partitions is T_2 and will be shown to go to 0 as $n \to \infty$.

For each $j \in \{1, 2, \ldots, b\}$ denote $(\pi(i_j - 1), \pi(i_j))$ as (t_j, l_j). Clearly $t_1 = \pi(0)$ and $l_1 = \pi(1)$. It is easy to see that each distinct (t_j, l_j) corresponds to each distinct letter in $\boldsymbol{\omega}$. Let S be the set of all generating vertices, i.e., set of representatives of each of the distinct $\mathcal{E}_{l_j}$'s and $\mathcal{E}_{t_1}$ (see page 61). Clearly, by (2.4.6), $|S| \le (b+1)$.

Let $\boldsymbol{\omega} \in SS_b(k)$. Then by Lemma 3.4.6, $|S| = b + 1$. Suppose the jth new letter appear s_j times in $\boldsymbol{\omega}$. Clearly all the s_j are even. So the total contribution of this $\boldsymbol{\omega}$ to

T_1 in (3.4.10) is as follows:

$$\frac{1}{n}\sum_{\substack{(t_j,l_j)\\1\le j\le b}}\prod_{j=1}^{b}\mathbb{E}\Big[y_{t_jl_j}^{s_j}\Big]=\frac{1}{n^{b+1}}\sum_{\substack{(t_j,l_j)\\1\le j\le b}}\prod_{j=1}^{b}g_{s_j,n}(t_j/n,l_j/n). \tag{3.4.11}$$

Next observe that if a sequence of bounded Riemann integrable functions, say f_n converges uniformly to a function f and a sequence of finite measures, say ν_n converges weakly to a finite measure ν, then

$$\int f_n\ d\nu_n\to\int f\ d\nu.$$

From this observation it is clear that for any sequence of bounded Riemann integrable function $f_n(x_1,x_2,\ldots,x_{b+1})$ on $[0,1]^{b+1}$, that converges uniformly to $f(x_1,x_2,\ldots,x_{b+1})$, as $n\to\infty$,

$$\frac{1}{n^{b+1}}\sum_{j_1,\ldots,j_{b+1}=1}^{n}f_n(j_1/n,j_2,n,\ldots,j_{b+1}/n)\to\int_{[0,1]^{b+1}}f(x_1,x_2,\ldots,x_{b+1})\,dx_1dx_2\cdots dx_{b+1}.$$

Now since $|S|=b+1$ for $\boldsymbol{\omega}\in SS_b(2k)$, as $n\to\infty$, the expression (3.4.11) converges to

$$\int_{[0,1]^{|S|}}\prod_{j=1}^{b}g_{s_j}\left(x_{t_j},x_{l_j}\right)\,dx_S, \tag{3.4.12}$$

where $dx_S=\prod_{l_j\in S}dx_{l_j}$ (with $l_0=t_1=0$) denotes the $|S|-$dimensional Lebesgue measure on $[0,1]^{|S|}$.

We split the investigation of T_2 into two cases.

Case 1. Suppose $\boldsymbol{\omega}$ is an even word with b distinct letters, but is not special symmetric. Then the contribution to T_2 of (3.4.10) can be calculated as in (3.4.11). But now note that $|S|\le b$. Hence in this case as $n\to\infty$, the contribution of this word $\boldsymbol{\omega}$ is 0.

Case 2. $\boldsymbol{\omega}\notin E(2k)$. Suppose $\boldsymbol{\omega}$ contains b_1 distinct letters each of which appear an even number of times and b_2 number of distinct letters that appear an odd number of times, and $b=b_1+b_2$. Without loss of generality, we can assume for each $\pi\in\Pi(\boldsymbol{\omega})$, s_{j_p}, $1\le p\le b_1$ to be even and s_{j_q}, $b_1+1\le q\le b_1+b_2$ to be odd. Hence the contribution of this $\boldsymbol{\omega}$ to T_2 in (3.4.10) is as follows:

$$\begin{aligned}
&\frac{1}{n}n^{-b_1}n^{-(b_2-\frac{1}{2})}\sum_{\substack{(t_j,l_j)\\1\le j\le b}}\prod_{p=1}^{b_1} g_{s_{j_p},n}\big(t_{j_p}/n, l_{j_p}/n\big)\prod_{m=b_1+1}^{b_1+b_2} n^{\frac{b_2-(1-1/2)}{b_2}}\mathbb{E}\Big[y_{t_{im}l_{jm}}^{s_{jm}}\Big]\\
&=\frac{1}{n^{b_1+b_2+\frac{1}{2}}}\sum_{\substack{(t_j,l_j)\\1\le j\le b}}\prod_{p=1}^{b_1} g_{s_{j_p},n}\big(t_{j_p}/n, l_{j_p}/n\big)\prod_{m=b_1+1}^{b_1+b_2} n^{\frac{b_2-(1-1/2)}{b_2}}\mathbb{E}\Big[y_{t_{im}l_{jm}}^{s_{jm}}\Big]. \qquad (3.4.13)
\end{aligned}$$

For n large, $n^{\frac{b_2-(1-1/2)}{b_2}}\mathbb{E}[y_{t_{im}l_{jm}}^{s_{jm}}] < 1$ for any $b_1+1 \le m \le b_1+b_2$. Now as $|S| \le b$, (3.4.13) contributes 0 as $n \to \infty$.

For any partition $\sigma \in SS_b(2k)$ let $\{V_1, \ldots, V_b\}$ be its partition blocks. Then from (3.4.10), (3.4.12) and the above argument, we have

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}[\mathrm{Tr}(Z_n)^{2k}] = \sum_{b=1}^{k}\sum_{\sigma\in SS_b(2k)}\int_{[0,1]^{|S|}}\prod_{j=1}^{b} g_{|V_j|}(x_{t_j}, x_{l_j})\, dx_S, \qquad (3.4.14)$$

where $dx_S = \prod_{i\in S} dx_i$ denotes the $|S|-$dimensional Lebesgue measure on $[0,1]^{|S|}$.

We also note that $\lim_{n\to\infty}\frac{1}{n}\mathbb{E}[\mathrm{Tr}(Z_n)^{2k+1}] = 0$ for any $k \ge 0$. This completes the proof of the first moment condition.

Step 4 (Uniqueness of the LSD): We have obtained

$$\gamma_{2k} = \lim_{n\to\infty}\frac{1}{n}\mathbb{E}[\mathrm{Tr}(Z_n)^{2k}] \le \sum_{\sigma\in SS(2k)} M_\sigma \le \sum_{\sigma\in\mathcal{P}(2k)} M_\sigma = \alpha_{2k}.$$

As $\{\alpha_{2k}\}$ satisfies Carleman's condition, $\{\gamma_{2k}\}$ also does so. Now using Lemma 2.1.3, we see that there exists a measure μ' with moment sequence $\{\gamma_{2k}\}$ such that μ_{Z_n} converges weakly almost surely to μ'.

This completes the proof of Part (a).

Step 5 (Proof of Part (b)): To see this, observe that (see Lemma 2.3.1)

$$d_2^2(\mu_{W_n}, \mu_{Z_n}) \le \frac{1}{n}\sum_{i,j} x_{ij}^2\mathbf{1}_{[|x_{ij}|>r_n]}. \qquad (3.4.15)$$

Now if this $\{r_n\}$ also satisfies Condition (3.3.4), then using (3.4.15) and (a), we can say that the ESD of W_n converges to μ' almost surely (respectively in probability).

This proves Part (b).

Step 6 (Semicircularity): To complete the proof, it remains to verify the condition for the limit to be semicircular. Recall that any word of length $2k$ is *Catalan* if it is pair-matched and at the same time non-crossing. This corresponds to the partitions in $SS_k(2k) = NC_2(2k)$.

First note that if $g_{2k} = 0$ for $k > 1$, then by (3.4.14) only the Catalan words contribute to the sum. Let ω be a Catalan word with the last new letter appearing at the $(\pi(i_k - 1), \pi(i_k))$th position. Then the generating vertex for that letter is $\pi(i_k)$. Therefore, going by the notation mentioned in the beginning of this proof and as well as those in the proof of Lemma 3.4.6, i_k does not appear in any of the pairs (t_j, l_j) for $j < k$. So, the contribution $p(\omega)$ say, of ω to γ_{2k} is as follows:

$$p(\omega) = \int_{[0,1]^{k+1}} g_2(x_{t_1}, x_{l_1}) \cdots g_2(x_{t_k}, x_{l_k})\ dx_S. \tag{3.4.16}$$

Since x_{l_k} does not appear in any of the other factors of the integrand, we can integrate w.r.t x_{l_k} to get

$$\begin{aligned} p(\omega) &= \int_{[0,1]^k} g_2(x_{t_1}, x_{l_1}) \cdots \int_{[0,1]} g_2(x_{t_k}, x_{l_k})\ dx_{i_k}\ dx_{S\backslash\{x_{l_k}\}} \\ &= c \int_{[0,1]^k} g_2(x_{t_1}, x_{l_1}) \cdots g_2(x_{t_{k-1}}, x_{l_{k-1}})\ dx_{S\backslash\{x_{l_k}\}}, \quad \text{as} \int_{[0,1]} g_2(x,y)\ dy\ = c. \end{aligned}$$

Now dropping the last new letter from the word ω, the reduced word ω' is also a Catalan word. Hence following the same argument, we have that (3.4.16) becomes c^k which is independent of ω. Hence for any Catalan word the contribution to γ_{2k} is same. Therefore, the limit is semicircular.

Now suppose that the limit is semicircular, and without loss assume that this semicircular has variance 1. Note that then the fourth moment equals 2. Define

$$f(x) = \int_0^1 g_2(x,y)dy.$$

Then

$$\gamma_2 = \int_0^1 f(x)dx = 1.$$

Then from equation (3.4.12), the fourth moment is given by

$$\begin{aligned}
2 &= \gamma_4 \\
&= \int_{[0,1]^2} g_4(x,y)dxdy + 2\int_{[0,1]^3} g_2(x_1,x_2)g_2(x_1,x_3)dx_1dx_2dx_3 \\
&= \int_{[0,1]^2} g_4(x,y)dxdy + 2\int_0^1 f^2(x)dx \\
&\geq \int_{[0,1]^2} g_4(x,y)dxdy + 2\Big(\int_0^1 f(x)dx\Big)^2 \\
&\geq \int_{[0,1]^2} g_4(x,y)dxdy + 2.
\end{aligned}$$

Clearly then from the above $f(x) = 1$ for all x, and $g_4 \equiv 0$.

Now we shall use induction on k to prove that $g_{2k} \equiv 0$ for all $k \geq 1$. Suppose $g_{2r} \equiv 0$ for all $r \leq k$. As the limit is semicircular, we have

$$\begin{aligned}
C_{2(k+1)} &= \gamma_{2k+2} \\
&= \int_{[0,1]^2} g_{2k+2}(x,y)dxdy + \sum_{SS_k(2k+2)} \int_{[0,1]^{k+2}} \prod_{j=1}^{b} g_2(x_{t_j}, x_{l_j}) \prod_{j=0}^{b} x_{l_j} \\
&= \int_{[0,1]^2} g_{2k+2}(x,y)dxdy + C_{2(k+1)} \int_0^1 f^{k+1}(x)dx \\
&\geq \int_{[0,1]^2} g_{2k+2}(x,y)dxdy + C_{2(k+1)}\Big(\int_0^1 f(x)dx\Big)^{k+1} \\
&\geq \int_{[0,1]^2} g_{2k+2}(x,y)dxdy + C_{2(k+1)}.
\end{aligned}$$

The last equality above occurs as all other words involves g_{2r} for some $r \leq k$ and hence contributes 0. Now note that from earlier calculations, since $f \equiv 1$, the contribution of each Catalan word (of any order) equals 1. Since this contribution already gives the moments of the semicircular distribution, the contribution from other words vanish. As a consequence, $g_{2k} = 0$ for all $k > 1$. The proof of the theorem is now complete. □

Remark 3.4.8. *When we look at the case where $g_{2k} = 0$ for all $k > 1$, we know from (3.4.14) that only Catalan words contribute to the sum. But without any further condition on g_2, the contribution of different Catalan words may be unequal. For example, let $\omega_1 = aabbcc$ and $\omega_2 = abccba$ be two Catalan words of length 6. By (3.4.12), the* *contribution for ω_1 is*

$$p(\omega_1) = \int_{[0,1]^4} g_2(x_1, x_2) g_2(x_1, x_3) g_2(x_1, x_4) \prod_{i=1}^{4} dx_i,$$

while the contribution for ω_2 is

$$p(\omega_2) = \int_{[0,1]^4} g_2(x_1, x_2) g_2(x_2, x_3) g_2(x_3, x_4) \prod_{i=1}^{4} dx_i.$$

Obviously $p(\omega_1) \neq p(\omega_2)$ in general. Also, it can be verified that under the assumption that $g_{2k} = 0$ for all $k > 1$, the condition $\int_{[0,1]} g_2(x, y)\, dy$ is constant is necessary for the limit to be semicircular.

Remark 3.4.9. *Assumption A(ii) that the sequence of functions $g_{2k,n}$ converges to g_{2k} uniformly for all $k \geq 1$, can be weakened slightly. For details see Corollary 3.5.16.*

Proof of Proposition 3.3.3 (Unbounded support). Consider $k = mt$ for some $t \geq 1$. Then from (3.4.14), we have

$$\beta_{2k}(\mu') = \sum_{b=1}^{k} \sum_{\sigma \in SS_b(2k)} \int_{[0,1]^{b+1}} \prod_{j=1}^{b} g_{|V_j|}(x_{t_j}, x_{l_j}) \prod_{i \in S} dx_i. \tag{3.4.17}$$

Recall that σ in the above expression could be described as a word in $SS_b(2k)$, and hence has $(b+1)$ distinct generating vertices. Let us focus on words $\omega \in SS_t(2k)$ with t distinct letters, and where each letter appears $2m$ times in pure even blocks. Therefore as $n \to \infty$, the contribution of ω in the limiting moment is (see (3.4.12)):

$$\int_{[0,1]^{t+1}} g_{2m}(x_0, x_1) g_{2m}(x_0, x_2) \cdots g_{2m}(x_0, x_t)\ dx_0 dx_1 \cdots dx_t = \int_{[0,1]} \left(f_{2m}(x_0)\right)^t\ dx_0. \tag{3.4.18}$$

Now, there are t distinct letters each appearing $2m$ times in pure even blocks in ω. A pure even block is at least a string of length 2 of the same letters (see Definition 3.2.2). Thus there are mt pure blocks across the different letters, and m pure blocks of size 2 for each letter (note that a pure block of size $2s$ for any letter can be thought of as s pure blocks of size 2 of that letter) in ω. To count the number of words, we will use the following argument:

Among the mt places for pure blocks of size 2, pure blocks of size 2 of the first letter can come in any of the m places. For pure blocks of size 2 of the second letter, m places among the remaining $(mt-m)$ places have to be chosen, which can be done in $\binom{mt-m}{m}$ ways. Carrying on this way, for the last letter, there is only $\binom{m}{m}$ way in which the m pure blocks of size 2 can appear in a word as described above. Therefore the number of such words ω is

$$\frac{1}{t!}\binom{mt}{m}\binom{mt-m}{m}\cdots\binom{m}{m} = \frac{1}{t!}\frac{(mt)!}{(m!)^t}. \tag{3.4.19}$$

Since the integrand in (3.4.17) is non-negative, using (3.4.18) and (3.4.19), we have

$$\begin{aligned}\beta_{2k}(\mu') &> \frac{1}{t!}\frac{(mt)!}{(m!)^t}\int_{[0,1]}\left(f_{2m}(x_0)\right)^t\,dx_0.\\ &= \frac{(mt)!}{t!}\int_{[0,1]}\left(\frac{f_{2m}(x_0)}{m!}\right)^t\,dx_0\\ &> \alpha\frac{(mt)!}{t!},\quad k=mt.\end{aligned}$$

Therefore for t sufficiently large (with $k=mt$),

$$(\beta_{2k}(\mu'))^{1/2k} > K\ t^{(\frac{1}{2}-\frac{1}{2m})} \quad \text{for some constant } K>0.$$

As a consequence, $(\beta_{2k}(\mu'))^{1/2k} \to \infty$ as $m>1$ and $k=mt\to\infty$. Hence μ' has unbounded support. □

3.5 Some Corollaries

In this section, we present a few corollaries that follow from Theorem 3.3.1. In particular, we deduce Results 3.1.1-3.1.8 from Theorem 3.3.1. We also discuss the convergence in some other models that can be deduced using Theorem 3.3.1.

3.5.1 Fully i.i.d. entries

Corollary 3.5.1. *Result 3.1.1 follows from Theorem 3.3.1.*

Proof. Suppose $W_n = (x_{ij}/\sqrt{n})$, where $\{x_{ij}\}$ are i.i.d. with distribution F, mean zero and variance one.

Let $r_n = n^{-1/3}$. Then $r_n\sqrt{n} \to \infty$ as $n \to \infty$ and

$$\lim_{n\to\infty} n\,\mathbb{E}\left[\left(\frac{x_{ij}}{\sqrt{n}}\right)^2 \mathbf{1}_{[|x_{ij}/\sqrt{n}|\le r_n]}\right] \;=\; 1 = C_2. \tag{3.5.1}$$

Also, for any $k > 2$,

$$\begin{aligned}
n\,\mathbb{E}\left[\left(\frac{x_{ij}}{\sqrt{n}}\right)^k \mathbf{1}_{[|x_{ij}/\sqrt{n}|\le r_n]}\right] &= n\,\mathbb{E}\big[(x_{11}/\sqrt{n})^{(k-2)}\;(x_{11}/\sqrt{n})^2\mathbf{1}_{[|x_{11}|\le r_n\sqrt{n}]}\big]\\
&\le n\frac{r_n^{(k-2)}}{n}\mathbb{E}\big[{x_{11}}^2\mathbf{1}_{[|x_{11}|\le r_n\sqrt{n}]}\big]\\
&\le r_n^{(k-2)}\\
&= (n^{-\frac{1}{3}})^{k-2} \;\to\; 0 \quad \text{as } n\to\infty.
\end{aligned} \tag{3.5.2}$$

(3.5.1) and (3.5.2) implies that $\{x_{ij}/\sqrt{n}\}$ satisfies (3.3.1) and (3.3.2). Now for any $t > 0$,

$$\begin{aligned}
\frac{1}{n}\sum_{i,j}\;(x_{ij}/\sqrt{n})^2[\mathbf{1}_{[|x_{ij}/\sqrt{n}|>r_n]}] &= \frac{1}{n^2}\sum_{i,j}\;x_{ij}^2[\mathbf{1}_{[|x_{ij}|>r_n\sqrt{n}]}]\\
&\le \frac{1}{n^2}\sum_{i,j}\;x_{ij}^2[\mathbf{1}_{[|x_{ij}|>t]}] \quad \text{for all large } n,\\
&\xrightarrow{a.s.}\;\mathbb{E}\big[x_{11}^2[\mathbf{1}_{[|x_{11}|>t]}]\big].
\end{aligned}$$

As $\mathbb{E}[x_{11}^2] = 1$, taking t to infinity, the left side converges to 0 almost surely.

So without loss we may assume that all moments of F are finite. Let G_n be the distribution of $X/\sqrt{n}$ for each n, where $X \sim F$. So the kth moment of G_n equals $\mu_n(k) = \frac{\beta_k(F)}{n^{k/2}}$ for $k \ge 1$. Thus

$$n\mu_n(k) = \begin{cases} \beta_2(F) = 1 & \text{if } k = 2,\\ \frac{\beta_k(F)}{n^{k/2-1}} & \text{if } k \ge 3.\end{cases}$$

As F has all moments finite, we have $g_2 \equiv 1$ and $g_{2k} \equiv 0$ for all $k > 1$.

Hence $W_n = (x_{ij}/\sqrt{n})_{1\le i,j\le n}$, where $x_{ij} = x_{ji},\; 1 \le i < j \le n$ satisfy the assumptions of Theorem 3.3.1. Therefore the ESD of W_n converges almost surely to μ whose moments are given by

$$\beta_{2k}(\mu) = \sum_{\sigma\in SS(2k)} g_\sigma.$$

Thus as in proof of Proposition 3.3.3

$$\beta_k(\mu) = \sum_{\sigma \in SS(2k)} g_\sigma = \sum_{\sigma \in NC_2(2k)} 1 = \frac{1}{k+1}\binom{2k}{k}. \tag{3.5.3}$$

Therefore for every k, the $2k$th moment is the kth Catalan number, and hence the limiting spectral distribution of $W_n = (x_{ij})/\sqrt{n}$ is the semicircular distribution. Thus we get Result 3.1.1 as a special case of our Theorem 3.3.1. □

3.5.2 Heavy-tailed Entries

Corollary 3.5.2. *Result 3.1.2 follows from Theorem 3.3.1.*

Proof. Suppose F is an α-stable distribution ($0 < \alpha < 2$), i.e., there exists a slowly varying function L such that

$$\mathbb{P}[|x_{ij}| \geq u] = \frac{L(u)}{u^\alpha}.$$

Now we consider $W_n = (x_{ij}/a_n)$ where $\{x_{ij}\}$ are i.i.d. with distribution F and

$$a_n = \inf\{u : \mathbb{P}[|x_{ij}| \geq u] \leq \frac{1}{n}\}.$$

Ben Arous and Guionnet [2008] proved the existence of LSD of W_n using the method of Stieltjes transform. We show how our theorem can be used to give a partially different proof. First let us define the d_1 metric.

Let f be a *Lipschitz function* on $\mathbb{R}$. Then $||f||_{\mathcal{L}}$ is the norm defined by

$$||f||_{\mathcal{L}} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x-y|} + \sup_x |f(x)|.$$

The d_1 distance between two real probability measures μ and ν is given by

$$d_1(\mu, \nu) = \sup_{||f||_{\mathcal{L}} \leq 1, f\uparrow} \left| \int f \, d\nu - \int f \, d\mu \right|, \tag{3.5.4}$$

where the supremum is taken over non-decreasing Lipschitz function f with $||f||_{\mathcal{L}} \leq 1$. It is known (Lemma 2.1 in Ben Arous and Guionnet [2008]) that d_1 metrizes weak convergence of distributions.

For any fixed positive constant say B, consider the matrix W_n^B whose entries are $\frac{x_{ij}}{a_n}\mathbf{1}_{[|x_{ij}|\le Ba_n]}$. Then we have the following:

(a) For every fixed $B \in \mathbb{N}$, W_n^B satisfies assumptions of Theorem 3.3.1 with $g_{2k} \equiv c_{2k,B}$ for every $k \geq 1$, where $c_{2k,B}$ are constants independent of n. Hence there exists a probability measure μ_B which is the weak limit of the ESD of W_n^B almost surely. That is, for each fixed B and $\epsilon > 0$, for n large enough,

$$d_1(\mu_B, \mathbb{E}[\mu_{W_n^B}]) \leq \epsilon/3. \tag{3.5.5}$$

(b) Using Theorem 2.2 in Ben Arous and Guionnet [2008], for every $\epsilon > 0$, there exists $B(\epsilon)$ and $\delta(\epsilon, B) > 0$ such that for n large enough

$$\mathbb{P}\big[d_1(\mu_{W_n}, \mu_{W_n^B}) > \epsilon\big] \leq \exp(-\delta(\epsilon, B)n). \tag{3.5.6}$$

Using the convexity of d_1 and (3.5.6), for every $\epsilon > 0$, there exists $B(\epsilon)$ and $\delta(\epsilon, B) > 0$ such that for n large enough

$$d_1\big(\mathbb{E}(\mu_{W_n}), \mathbb{E}(\mu_{W_n^B})\big) \leq \exp(-\delta(\epsilon, B)n). \tag{3.5.7}$$

Hence using (3.5.5) and (3.5.7), we have for B, B' large so that (3.5.7) holds,

$$\begin{aligned} d_1(\mu_B, \mu_{B'}) \leq{}& d_1(\mu_B, \mathbb{E}[\mu_{W_n^B}]) + d_1(\mu_{B'}, \mathbb{E}[\mu_{W_n^{B'}}]) + d_1\big(\mathbb{E}(\mu_{W_n}), \mathbb{E}(\mu_{W_n^B})\big) \\ &+ d_1\big(\mathbb{E}(\mu_{W_n}), \mathbb{E}(\mu_{W_n^{B'}})\big). \end{aligned} \tag{3.5.8}$$

Then letting $n \to \infty$ and B, B' sufficiently large we have that

$$d_1(\mu_B, \mu_{B'}) < \epsilon.$$

Thus $(\mu_B)_B$ is d_1- Cauchy. As the space of all distributions is complete with respect to this metric, μ_B converges to a probability measure $\tilde{\mu}$, say, as $B \to \infty$.

(c) Next observe that

$$\begin{aligned} \mathbb{P}\big[d_1(\mu_{W_n}, \tilde{\mu}) > \epsilon\big] \leq{}& \mathbb{P}\big[d_1(\mu_{W_n}, \mu_{W_n^B}) > \epsilon/3\big] + \mathbb{P}\big[d_1(\mu_{W_n^B}, \mu_B) > \epsilon/3\big] \\ &+ \mathbb{P}\big[d_1(\tilde{\mu}, \mu_B) > \epsilon/3\big]. \end{aligned}$$

Choosing B large enough, and then taking n to ∞, the r.h.s. of the above inequality can be made arbitrarily small.

In conclusion μ_{W_n} converges weakly to $\tilde{\mu}$ in probability. This yields Result 3.1.2. $\square$

3.5.3 General triangular i.i.d.

The next corollary states a LSD result about Wigner matrices with general triangular i.i.d. entries, where the entries of the matrix are i.i.d. for every n (size of the matrix), but are allowed to vary with n.

Corollary 3.5.3. *Suppose the entries $\{x_{ij}\ ; 1 \le i \le j \le n\}$ are i.i.d. for each n, and the following two conditions hold:*

(i) For each $k \in \mathbb{N}$,

$$c_{2k} = \lim_{n\to\infty} n\ \mathbb{E}\left[x_{ij}^{2k}\mathbf{1}_{\{|x_{ij}|\le r_n\}}\right] \quad \text{is finite,} \tag{3.5.9}$$

$$\lim_{n\to\infty} n^{\alpha}\ \mathbb{E}\left[x_{ij}^{2k-1}\mathbf{1}_{\{|x_{ij}|\le r_n\}}\right] = 0 \quad \text{for any} \quad \alpha < 1. \tag{3.5.10}$$

(ii) The sequence $\{0, c_2, 0, c_4, 0, \ldots\}$ is the cumulant sequence of a probability distribution G whose moment sequence $\{\beta_k\}$ satisfies Carleman's condition:

$$\sum_{k=1}^{\infty} \beta_{2k}^{-\frac{1}{2k}} = \infty.$$

Let $y_{ij} = x_{ij}\mathbf{1}_{[|x_{ij}|\le r_n]}$ and $Z_n = (y_{ij})_{1\le i\le j\le n}$ be symmetric. Then the ESD of Z_n converges almost surely to μ whose moments are given by

$$\beta_k(\mu) = \begin{cases} \sum\limits_{\sigma\in SS(k)} c_\sigma & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd.} \end{cases}$$

Further if

$$\frac{1}{n}\sum_{i,j} x_{ij}^2[\mathbf{1}_{\{|x_{ij}|>r_n\}}] \to 0, \ \text{almost surely (respectively in probability).}$$

then the ESD of W_n converges weakly to μ almost surely (respectively in probability),

Proof of Corollary 3.5.3. We know that $\{x_{ij}; 1 \leq i \leq j \leq n\}$ are i.i.d. for every fixed n. Then $\{y_{ij}; 1 \leq i \leq j \leq n\}$ are also i.i.d. for every fixed n. From condition (i) of the corollary, clearly (3.3.2) and (3.3.1) are satisfied with $g_{2k,n} \equiv c_{2k}$ on $[0,1]^2$. Therefore, $g_{2k} \equiv c_{2k}$ on $[0,1]^2$ and W_n satisfies condition (ii) of Assumption A. Having observed this, condition (ii) of the corollary implies condition (iii) of Assumption A. Thus, from Theorem 3.3.1, the ESD of Z_n converges to a probability measure μ.

From (3.4.14), we see that only the special symmetric words contribute to the limiting moment sequence. Also as $g_{2k} \equiv c_{2k}$ on $[0,1]^2$, the moments of μ are given by

$$\beta_k(\mu) = \begin{cases} \displaystyle\sum_{\sigma \in SS(k)} c_\sigma & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd.} \end{cases}$$

Now suppose further that $\{x_{ij,n}, i \leq j\}$ satisfies (3.3.4). Then by Theorem 3.3.1, the ESD of W_n converges to μ almost surely (respectively in probability). $\square$

Remark 3.5.4. *We want to find a class of non-trivial matrices whose almost sure LSD will be shown to exist from Corollary 3.5.3. This will be done with the help of an infintely divisible distribution.*

Suppose F is a symmetric infinitely divisible distribution with all moments finite and cumulant sequence $\{D_k\}_{k\geq 1}$. Then due to infinite divisibility, for every n, we can find i.i.d. random variables $\{y_{i,n}, 1 \leq i \leq n\}$ with distribution G_n, such that $\sum_{i=1}^{n} y_{i,n}$ converges in distribution to F (see Characterization 1 in Bose et al. [2002]). Moreover, it can be easily verified that the convergence holds if

$$n\mathbb{E}[y_{i,n}^k] \to D_k, \quad \textit{for every } k \geq 1. \tag{3.5.11}$$

Now let $\{x_{ij,n}, i \leq j\}$ be i.i.d. with distribution G_n, for every fixed n. That is, for $i \leq j$, consider $x_{ij,n} \stackrel{\mathcal{D}}{=} y_{i,n}$ for every fixed n. Thus, (3.5.11) holds for $x_{ij,n}$. Then from the above discussion, it is clear that these variables satisfy (3.5.9) and (3.5.10) with $r_n = \infty$ and $c_{2k} = D_{2k}$. Now if the moments of F satisfy Carleman's condition, then the variables $\{x_{ij,n}, i \leq j\}$ satisfy Conditions (i) and (ii) of Corollary 3.5.3. Thus the ESD of W_n with entries $\{x_{ij,n}, i \leq j\}$ converges almost surely to the symmetric probability distribution μ which is identified by $\{D_{2k}\}_{k\geq 1}$. This gives a class of non-trivial matrices *whose almost sure LSD exist. However since the moments of μ are given via $SS(2k)$, $\{D_{2k}\}_{k\geq 1}$ does not give the free or classical or half cumulants of μ.*

Remark 3.5.5. *If Condition (ii) of Corollary 3.5.3 is replaced by the condition that the sequence $\{0, c_2, 0, c_4, 0, \ldots\}$ is the cumulant sequence of a probability distribution G whose moment generating function has positive radius of convergence around 0, then the result still holds. To see this, suppose $X \sim G$. Suppose Y is a random variable whose moments are as as follows:*

$$\mathbb{E}(Y^{2k-1}) = 0 \quad \textit{and} \quad \mathbb{E}(Y^{2k}) = \sum_{\sigma \in SS(2k)} c_\sigma \qquad \textit{for each } k \geq 1.$$

Then observe that

$$\mathbb{E}(Y^{2k}) = \sum_{\sigma \in SS(2k)} c_\sigma \leq \sum_{\sigma \in E(2k)} c_\sigma = \mathbb{E}(X^{2k}), \tag{3.5.12}$$

and

$$0 \leq M_Y(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathbb{E}(Y^k) = \sum_{k=0}^{\infty} \frac{t^{2k}}{(2k)!}\mathbb{E}(Y^{2k}) \leq \sum_{k=0}^{\infty} \frac{t^{2k}}{(2k)!}\mathbb{E}(X^{2k}) = M_X(t).$$

So if $M_X(t)$ has a positive radius of convergence around 0, then $M_Y(t)$ has a positive radius of convergence around 0. This implies that the distribution of Y is uniquely determined by its moments, and everything else follows as in the proof of Theorem 3.3.1.

Corollary 3.5.6. *Result 3.1.3 follows from Corollary 3.5.3.*

Proof. In Zakharevich [2006] the entries of W_n are $\frac{x_{ij}}{\sqrt{n\mu_n(2)}}$ where x_{ij} are i.i.d. G_n for each n, which has mean zero, and all moments finite. With the additional condition on $\mu_n(k)$ assumed in Zakharevich [2006] that

$$\lim_{n\to\infty} \frac{\mu_n(k)}{n^{k/2-1}\mu_n(2)^{k/2}} = g_k \quad \text{exists for all } k \geq 1,$$

it is easy to see that W_n satisfies assumptions of Corollary 3.5.3. Hence by Corollary 3.5.3, the ESD of W_n converges to μ almost surely.

As mentioned earlier, Zakharevich [2006] used the moment method to calculate the limiting moments. Since Lemma 3.2.5 provides a description of $SS(2k)$ in terms of colored rooted trees, we can express the $2k$th moment of the LSD of W_n under the assumption of Corollary 3.5.3 in terms of these trees. Let $T_{2k}^b(k_1, k_2, \dots, k_b)$ denote the number of coloured rooted trees with $(k+1)$ vertices and $(b+1)$ distinct colours which satisfy Properties (a), (b) and (c) of Lemma 3.2.5.

Then, the $2k$th moment of the LSD of W_n is given by

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathrm{Tr}(W_n^{2k})] = \sum_{\sigma\in SS(2k)} c_\sigma = \sum_{b=1}^{k} \sum_{k_1+k_2+\cdots+k_b=k} T_{2k}^b(k_1, k_2, \dots, k_b) \prod_{i=1}^{b} c_{2k_i}.$$

This helps us obtain Result 3.1.3 from Corollary 3.5.3 and completes the proof. □

3.5.4 Sparse Matrices

In the next two Corollaries 3.5.7 and 3.5.8, we show how LSD results for the adjacency matrices of Erdös-Rényi graphs can be concluded from Thorem 3.3.1. As mentioned before, we consider two types of Erdös-Rényi graphs– *homogeneous* and *inhomogeneous*, in the sparse regime.

Corollary 3.5.7. *The ESD of M_n where M_n is the adjacency matrix in Result 3.1.4 converges weakly almost surely to a symmetric probability distribution μ whose odd moments are zero and even moments are given by*

$$\beta_{2k}(\mu) = \sum_{\sigma\in SS(2k)} \lambda^{|\sigma|} = \sum_{b=1}^{k} |SS_b(2k)|\lambda^b. \tag{3.5.13}$$

Proof. Observe that the independent entries of M_n, $x_{ij,n} \sim Ber(p_n)$ for $i < j$ with $np_n \to \lambda$; $x_{ii,n} = 0$ for all $1 \le i \le n$. It is easy to see that assumptions of Corollary 3.5.3 hold with $r_n = \infty$ and $c_k \equiv \lambda$ for all k. Hence by Corollary 3.5.3, the ESD of M_n converges almost surely to μ with odd moments zero, and even moments given by

$$\beta_{2k}(\mu) = \sum_{\sigma\in SS(2k)} c_\sigma = \sum_{\sigma\in SS(2k)} \lambda^{|\sigma|} = \sum_{b=1}^{k} |SS_b(2k)|\lambda^b. \tag{3.5.14}$$

□

Now since the limit of the EESD and almost sure limit of the ESD must be equal, the two expressions (3.5.13) and (3.1.2), must be identical. Since both expressions are polynomials in λ, we must have

$$|SS_b(2k)| = I_{k,b} \text{ for all } 1 \leq b \leq k.$$

Corollary 3.5.8. *(**Inhomogeneous**) Suppose there is a sequence of bounded Riemann integrable symmetric functions $p_n : [0,1]^2 \longrightarrow [0,1]$ such that np_n converges uniformly to a function p. Consider the Wigner matrix $W_n = (x_{ij})_{1\leq i\leq j\leq n}$, where $\{x_{ij},\ 1 \leq i \leq j \leq n\}$ are such that*

$$n\mathbb{E}[x_{ij}^{2k}] = p_n(i/n, j/n).$$

Then the ESD of W_n converges weakly almost surely to a non-random symmetric probability measure $\hat{\mu}$. Additionally, if $\int_0^1 p(x,y)\ dy = \lambda$, then $\hat{\mu} = \mu_{bg}$.

Proof of Corollary 3.5.8. It is easy to see that W_n satisfies Assumption A with $g_{2k,n} = p_n$ and $g_{2k} = p$ for all $k \geq 1$ and $r_n = \infty$.

Therefore, from Theorem 3.3.1, the ESD of W_n converges weakly almost surely to a non-random symmetric probability measure, $\hat{\mu}$ say.

Now suppose, $\int_0^1 p(x,y)\ dy = \lambda$. Then from (3.4.12), each word in $SS_b(2k)$ contributes λ^b to the $2k$th moment of $\hat{\mu}$. Hence this moment is given by

$$\beta_{2k}(\hat{\mu}) = \sum_{\pi\in SS(2k)} \lambda^{|\pi|} = \sum_{b=1}^{k} |SS_b(2k)|\lambda^b.$$

As these moments determine the distribution uniquely, we have $\hat{\mu} = \mu_{bg}$. $\square$

Remark 3.5.9. *We could of course start with numbers $p_{i,j,n} \in [0,1]$, $1 \leq i \leq j \leq n$, for each fixed n in Corollary 3.5.8. Then we can create a sequence of continuous functions p_n on $[0,1]^2$ such that*

$$p_n(i/n, j/n) = p_{i,j,n} \qquad \textit{for every } 1 \leq i \leq j \leq n.$$

One way of doing this is as follows: Consider the grid points $\{(i/n, j/n) : 1 \leq i, j \leq n\}$ on $[0,1]^2$ for every n.

Consider $(x,y) \in [0,1]^2$. Then there exist a unique pair (i,j) such that $\frac{(i-1)}{n} \leq x \leq \frac{i}{n}$ and $\frac{(j-1)}{n} \leq x \leq \frac{j}{n}$. Then (x,y) can be written uniquely as a convex combination of

$(x_1, y_1) = (\frac{i-1}{n}, \frac{j-1}{n}), (x_2, y_2) = (\frac{i}{n}, \frac{j-1}{n}), (x_3, y_3) = (\frac{i-1}{n}, \frac{j}{n}), (x_4, y_4) = (\frac{i}{n}, \frac{j}{n})$. *There exists* $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ *with* $0 \le \alpha_i \le 1$, $\alpha_1+\alpha_2+\alpha_3+\alpha_4 = 1$ *such that* $(x, y) = \sum_{i=1}^4 \alpha_i(x_i, y_i)$.

Now we can define the functions p_n *on* $[0,1]^2$ *as:*

$$p_n(x, y) = \alpha_1 p_n(x_1, y_1) + \alpha_2 p_n(x_2, y_2) + \alpha_3 p_n(x_3, y_3) + \alpha_4 p_n(x_4, y_4).$$

Now, assume that the functions np_n *converge uniformly to the function* p *on* $[0,1]^2$. *Then we can conclude the convergence of the ESD as discussed above.*

It can be verified that the condition $\int_0^1 p(x, y)\, dy = \lambda$ *is equivalent to the condition*

$$\sup_i \Big| \sum_{j=1}^n p_{ij,n} - \lambda \Big| \longrightarrow 0.$$

Remark 3.5.10. *One of the models for inhomogeneous Erdös-Rényi graphs in the sparse regime is the uniformly grown graph on* c/j, $G_n^{1/j}(c)$ *(see Bollobás et al. [2007]). In this case observe that,* $np_n \to p : [0,1]^2 \longrightarrow [0,1]$ *such that* $p(x, y) = c/\max(x, y)$. *Then we can see that from Theorem 3.3.1, following the above discussion, the ESD of the adjacency matrix of* $G_n^{1/j}(c)$ *converges weakly almost surely to a non-random symmetric probability measure, say* μ_c, *as* $n \to \infty$.

3.5.5 Matrices with variance profile

In the next two corollaries we describe results about *Wigner matrices with variance profile*, $(W_n, \cdot)$. We discuss two kinds of variance profiles– *discrete variance profile* and *continuous variance profile*.

Definition 3.5.11. *(a) Discrete variance profile: Suppose* $\{x_{ij,n}; i \le j\}$ *are i.i.d. random variables with mean zero and variance 1 and let* $\{\sigma_{ij} = \sigma_{ji}\}_{1 \le i,j \le n}$ *be uniformly bounded real numbers. Then the Wigner matrix with discrete variance profile* σ_d *is given by*

$$(W_n, \sigma_d) = \big((y_{ij,n} = \sigma_{ij} x_{ij,n})\big)_{1 \le i \le j \le n}. \tag{3.5.15}$$

(b) Continuous variance profile: Suppose $\{x_{ij,n}; i \le j\}$ *are i.i.d. random variables for every fixed* n, *and let* σ *be a symmetric bounded piecewise continuous function on*

$[0,1]^2$. *Then the Wigner matrix with continuous variance profile* σ_c, *is given by*

$$(W_n, \sigma_c) = \big((y_{ij,n} = \sigma(i/n, j/n)x_{ij,n})\big)_{1\le i\le j\le n}. \tag{3.5.16}$$

Wigner matrix, where the entries are independent but not identically distributed have been previously considered by Lytova and Pastur [2009], Bai and Silverstein [2010] etc. However, a common theme has been to consider the entries to have common mean and variance. Recently Zhu [2020], Jin and Xie [2020], have considered Wigner matrices with non-trivial discrete variance profile.

First we state and prove an LSD result for (W_n, σ_d) where $\sigma_d = \{\sigma_{ij}\}$ satisfy certain properties.

Corollary 3.5.12. ***(Discrete variance profile)*** *Consider the Wigner matrix* (W_n, σ_d) *with entries* $\{\frac{y_{ij}}{\sqrt{n}} : 1 \le i \le j \le n\}$ *where* $\{y_{ij,n} = y_{ij}\}$ *as in* (3.5.15). *Further assume that* $\{\sigma_{ij}\}$ *is uniformly bounded above by* $c > 0$ *and*

(i) the sequence $\{\sigma_{ij}\}$ *satisfies*

$$\sup_i \left| \frac{1}{n}\sum_{j=1}^{n} \sigma_{ij}^2 - 1 \right| \to 0 \qquad \text{as } n \to \infty. \tag{3.5.17}$$

(ii) The variables $\{x_{ij}\}$ *satisfy*

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{i,j} \mathbb{E}\big[x_{ij}^2 \mathbf{1}_{[|x_{ij}|>\eta\sqrt{n}]}\big] = 0, \quad \text{for every } \eta > 0. \tag{3.5.18}$$

Then the almost sure LSD of (W_n, σ_d) *is the semicircular distribution.*

Proof. We shall break the proof into three steps.

Step 1 (Truncation): In this step we will show that we can assume that the variables $\{x_{ij}\}$ are bounded by $\eta_n\sqrt{n}$, where $\{\eta_n\}_n$ is a sequence decreasing to 0. The idea of the proof of this step is borrowed from the proof of Theorem 2.9 in Bai and Silverstein [2010].

First observe that (3.5.18) is equivalent to

$$\lim_{n\to\infty} \frac{1}{\eta^2 n^2} \sum_{i,j} \mathbb{E}\big[x_{ij}^2 \mathbf{1}_{[|x_{ij}|>\eta\sqrt{n}]}\big] = 0, \quad \text{for every } \eta > 0. \tag{3.5.19}$$

Thus there is a sequence $\{\eta_n\}$ decreasing to 0 such that (3.5.19) is true with η_n in place of η. Now suppose Z_n is the matrix whose entries are $\{\frac{\tilde{y}_{ij}}{\sqrt{n}}\}$ where $\tilde{y}_{ij} = \sigma_{ij} x_{ij} \mathbf{1}_{[|x_{ij}|\leq \eta_n\sqrt{n}]}$. Then from (2.3.8), we have

$$\begin{aligned} ||F^{W_n} - F^{Z_n}|| &\leq \frac{1}{n} \operatorname{rank}(W_n - Z_n) \\ &\leq \frac{1}{n} \sum_{1\leq i,j\leq n} \mathbf{1}_{[|x_{ij}|>\eta_n\sqrt{n}]}. \end{aligned} \tag{3.5.20}$$

Now note that by (3.5.19),

$$\mathbb{E}\left[\frac{1}{n} \sum_{1\leq i,j\leq n} \mathbf{1}_{[|x_{ij}|>\eta_n\sqrt{n}]}\right] \leq \frac{1}{\eta^2 n^2} \sum_{i,j} \mathbb{E}\big[x_{ij}^2 \mathbf{1}_{[|x_{ij}|>\eta\sqrt{n}]}\big] \to 0. \tag{3.5.21}$$

Also,

$$\operatorname{Var}\left(\frac{1}{n} \sum_{1\leq i,j\leq n} \mathbf{1}_{[|x_{ij}|>\eta_n\sqrt{n}]}\right) \leq \frac{1}{\eta^2 n^3} \sum_{i,j} \mathbb{E}\big[x_{ij}^2 \mathbf{1}_{[|x_{ij}|>\eta\sqrt{n}]}\big] = o(1/n). \tag{3.5.22}$$

Applying *Bernstein's inequality* and (3.5.21) and (3.5.22), we have for $\epsilon > 0$, n can be chosen large enough so that

$$\mathbb{P}\left(\frac{1}{n} \sum_{1\leq i,j\leq n} \mathbf{1}_{[|x_{ij}|>\eta_n\sqrt{n}]} > \epsilon\right) \leq 2\exp(-\epsilon n). \tag{3.5.23}$$

Thus by Borel-Cantelli lemma, we get that the rhs of (3.5.20) goes to 0 almost surely. Also, using (3.5.17) and (3.5.18), for every η_n,

$$\sup_i \left| \frac{1}{n} \sum_j \mathbb{E}\left[x_{ij} \mathbf{1}_{[|x_{ij}|\leq\eta_n\sqrt{n}]} - \mathbb{E}\big[x_{ij} \mathbf{1}_{[|x_{ij}|\leq\eta_n\sqrt{n}]}\big]\right]^2 - 1 \right| \longrightarrow 0. \tag{3.5.24}$$

Hence, we see that the variables $\{x_{ij}\}$ associated to (W_n, σ_d) can be assumed to be bounded by $\eta_n\sqrt{n}$ for some sequence $\eta_n \downarrow 0$.

Step 2 (Convergence of the expected moments): In this step we will show that the EESD of (W_n, σ_d) converges weakly to the semicircle distribution.

For this we follow the arguments as in Step 3 of the proof of Theorem 3.3.1 and focus on finding the contribution of words to the limiting distribution.

First observe that for the word aa, the contribution to the moment sequence is 1. This is because

$$\Big|\frac{1}{n}\sum_{i_0,i_1}\frac{1}{n}\sigma^2_{i_0,i_1}-1\Big| \leq \sup_{i_0}\Big|\frac{1}{n}\sum_{i_1=1}^{n}\sigma^2_{i_0 i_1}-1\Big| \to 0 \quad \text{as } n\to\infty.$$

We shall prove by induction on the length of the word that, each Catalan word contributes 1 to the limit. Towards that, suppose all Catalan words of length $2(k-1)$ contribute 1 to the moment.

Now suppose ω is a Catalan word of length $2k$. Recall the notation used in the proof of Theorem 3.3.1. Using those notation, clearly $|S|=k+1$ and each distinct letter corresponds to the pair (t_j,l_j), $(1\leq j\leq k)$. As ω is Catalan, i_k appears only once in the sum. Therefore for this word ω,

$$\begin{aligned}\frac{1}{n^{k+1}}\sum_{t_0,l_1,\ldots,l_k}\prod_{j=1}^{k}\sigma^2_{t_j,l_j} &= \frac{1}{n^k}\sum_{t_0,l_1,\ldots,l_{k-1}}\prod_{j=1}^{k-1}\sigma^2_{t_j,l_j}\,\Big(\frac{1}{n}\sum_{l_k}{\sigma^2}_{t_k,l_k}-1\Big)\\ &\quad+\frac{1}{n^k}\sum_{t_0,l_1,\ldots,l_{k-1}}\prod_{j=1}^{k-1}\sigma^2_{t_j,l_j}. \qquad (3.5.25)\end{aligned}$$

The second term of the r.h.s. of the above equation goes to 1 as n goes to ∞ by the induction hypothesis. For the first term observe that the factor $\big(\frac{1}{n}\sum_{l_k}\sigma^2_{t_k,l_k}-1\big)\to 0$ by (3.5.17). Also note that as ω is Catalan, we can write the first term as

$$\frac{1}{n}\sum_{t_0}\prod_{j=1}^{k-1}\Big(\frac{1}{n}\sum_{l_j}\sigma^2_{t_j,l_j}\Big)\Big(\frac{1}{n}\sum_{l_k}\sigma^2_{t_k,l_k}-1\Big).$$

By (3.5.17), $\big(\frac{1}{n}\sum_{l_j}\sigma^2_{t_j,l_j}\big)$ is bounded for each $1\leq j\leq k-1$. Thus the first term of the r.h.s. of (3.5.25) goes to 0 as $n\to\infty$. So every Catalan word contributes 1 in the limit.

Now, suppose ω is a non-Catalan word with b distinct letters which appear $s_1,s_2,\ldots,s_b$ times. So, $|S|\leq b$, where S is the set of distinct generating vertices for ω. Then the contribution for this word is as follows:

$$\begin{aligned}
\frac{1}{n^{k+1}} \sum_{S} \prod_{j=1}^{b} \mathbb{E}[x_{t_j,l_j}^{s_j}] &\leq \frac{(\eta_n \sqrt{n})^{2k-2b}}{n^{k+1}} \sum_{S} \prod_{j=1}^{b} \sigma_{t_j,l_j}^2. \\
&= \frac{{\eta_n}^{2k-2b}}{n^{b+1}} \sum_{S} \prod_{j=1}^{b} \sigma_{t_j,l_j}^2. \\
&\leq \frac{{\eta_n}^{2k-2b+2}}{n^{b}} \sum_{S\setminus\{l_m\}} \prod_{\substack{j=1 \\ j\neq m}}^{b} \sigma_{t_j,l_j}^2 \qquad (\text{as } |S| \leq b)
\end{aligned}$$

As $|S| \leq b$ and $\left(\frac{1}{n}\sum_{l_j} \sigma_{t_j,l_j}^2\right)$ is bounded, the above quantity is of the order of ${\eta_n}^{2k-2b+2}$. Also since $b \leq k$ and $\eta_n \downarrow 0$, we see that such words do not contribute in the limit.

Therefore, $\lim\limits_{n\to\infty} \frac{1}{n}\mathbb{E}\big[\operatorname{Tr}(W_n, \sigma_d)^{2k}\big] = |NC_2(2k)|$ and $\lim\limits_{n\to\infty} \frac{1}{n}\mathbb{E}\big[\operatorname{Tr}(W_n, \sigma_d)^{2k+1}\big] = 0$. Hence the EESD of (W_n, σ_d) converges weakly to the semicircle distribution.

Step 3 (Fourth moment condition): Here we prove Condition (ii) of Lemma 2.1.3. Now, observe that as $|x_{ij}| \leq \eta_n\sqrt{n}$ where $\eta_n \downarrow 0$ and $|\sigma_{ij}| \leq c$, the fourth moment condition is very similar to that of Theorem 3.3.1. Just as in Step 2 of the proof of Theorem 3.3.1, it is enough to compute $\sum\limits_{\pi_1,\pi_2,\pi_3,\pi_4} \mathbb{E}[\prod\limits_{i=1}^{4} Y_{\pi_i}]$, where $(\pi_1, \pi_2, \pi_3, \pi_4)$ are jointly- and cross-matched. Now here too (3.4.5) is true. Again using the uniform boundedness of σ_{ij} by c and the fact that $|x_{ij}| \leq \eta_n\sqrt{n}$, we have that for each $(\pi_1, \pi_2, \pi_3, \pi_4)$ with b $(1 \leq b \leq 2k)$ distinct letters across $\pi_i, 1 \leq i \leq 4$,

$$\mathbb{E}[\prod_{i=1}^{4} Y_{\pi_i}] \leq \frac{\eta_n^{4k-2b}}{n^{b+4}} c^{2b}.$$

Now using Lemma 3.4.7, we have at most $\mathcal{O}(n^{b+2})$ such circuits $(\pi_1, \pi_2, \pi_3, \pi_4)$. Hence,

$$\frac{1}{n^4}\mathbb{E}\Big[\operatorname{Tr}(W_n, \sigma_d)^k \; - \; \mathbb{E}\big[\operatorname{Tr}(W_n, \sigma_d)^k\big]\Big]^4 = \mathcal{O}(n^{-2}).$$

Hence the fourth moment condition is verified.

Thus, the ESD of (W_n, σ_d) converges weakly almost surely to the semicircle distribution.

$\square$

Remark 3.5.13. *Theorem 1.1 in Jin and Xie [2020] states a similar result where condition (3.5.17) is replaced by* $\frac{1}{n}\sum_i \left|\frac{1}{n}\sum_{j=1}^{n}\sigma^2{}_{ij} - 1\right| \to 0$. *However, the proof of equation (2.6) there is not very clear.*

Corollary 3.5.14. *(**Continuous variance profile**) Consider the Wigner matrix* (W_n, σ_c) *with entries* $\{y_{ij,n}; 1 \le i \le j \le n\}$ *as in (3.5.16). Assume that the variables* $\{x_{ij}\}$ *satisfy conditions (i) and (ii) of Corollary 3.5.3. Then the ESD of* (W_n, σ_c) *converges weakly almost surely to a symmetric probability distribution* ν *whose* $2k$*th moment is determined by* σ *and* $\{C_{2m}\}_{2\le m\le 2k}$.

Proof. As $\{x_{ij}\}$ satisfy (3.5.9) and (3.5.10) with $r_n = \infty$, observe that (W_n, σ_c) satisfy Assumption A with $g_{2k}(x, y) = \sigma^{2k}(x, y)c_{2k}$ for all $(x, y) \in [0, 1]^2$ and $r_n = \infty$. Thus the ESD of (W_n, σ_c) converges weakly almost surely to a symmetric probability distribution ν. Now from Step 3 in Theorem 3.3.1, for each word in $SS_b(2k)$ with each distinct letter appearing $s_1, s_2, \ldots, s_b$ times, its contribution to the limiting moment is (see (3.4.12))

$$\int_{[0,1]^{b+1}} \prod_{j=1}^{b} \sigma^{s_j}(x_{t_j}, x_{l_j}) \prod_{i\in S} dx_i \prod_{j=1}^{b} c_{s_j},$$

where S is the set of distinct generating vertices for the word. Hence the $2k$th moment of ν is given as follows:

$$\beta_{2k}(\nu) = \sum_{b=1}^{k} \sum_{\pi\in SS_b(2k)} \int_{[0,1]^{b+1}} \prod_{j=1}^{b} \sigma^{s_j}(x_{t_j}, x_{l_j}) \prod_{i\in S} dx_i \prod_{j=1}^{b} c_{s_j}.$$

□

Remark 3.5.15. *Note that such a model was previously studied in Anderson and Zeitouni [2006]. There the authors considered the* $\{y_{ij,n} = \sigma(i/n, j/n)x_{ij,n}; i \le j\}$ *where* $\{x_{ij}; i \le j\}$ *are fully i.i.d. with mean zero and variance 1 with* $\int_0^1 \sigma^2(x, y)\, dy = 1$ *for every* x, *and proved that the ESD of* $\frac{1}{\sqrt{n}}W_n$ *converges weakly almost surely to the semi-circle law. Observe that this result can be concluded as discussed above with* $c_2 = 1$ *and* $c_{2k} = 0, k \ge 2$. *Hence the ESD result of Anderson and Zeitouni [2006] follows as a special case of Theorem 3.3.1.*

In the next corollary we describe a new result for Wigner matrices with non-trivial variance structures via graphons and homomorphism density. Recall graphons and homomorphism density from Section 3.1. We first generalize the concept of graphons, relating it to the higher moments of a sequence of random variables $\{x_{ij,n}; 1 \leq i \leq j \leq n\}$.

Define, for each k, a *graphon sequence* $M_{2k,n}$ that takes the value

$$n\mathbb{E}[x_{ij,n}^{2k}\mathbf{1}_{[|x_{ij,n}|\leq r_n]}] = g_{2k,n}(i/n, j/n) \quad \text{on } I_i \times I_j (1 \leq i, j \leq n),$$

where r_n is a sequence as given in Assumption A and $I_1 = (0, \frac{1}{n}], I_i = (\frac{i-1}{n}, \frac{i}{n}], 2 \leq i \leq n$.

Note that corresponding to each word in $SS(2k)$ with b distinct letters, we have a coloured rooted ordered tree as described in Lemma 3.2.5. Denote its vertex set by $V := \{0, \dots k\}$, enumerated by first appearances, depth first and left to right. Each vertex is painted with a colour from the colour set $C := \{a_0 < \cdots < a_b\}$, say. Let E be the edge set of T'. Observe that there can be many edges whose vertices have a fixed pair of colours a_i and a_j. Enumerate E as follows:

$$E = \cup_{0 \leq i < j \leq b} E(i, j)$$

$$E(i,j) = \big\{(v_1, v_2) \in E : v_1 < v_2 \text{ are coloured } a_i \text{ and } a_j \text{ respectively}\big\}, 0 \leq i < j \leq b.$$

We now extend the homomorphism density $t(T, H_n)$ in (3.1.3) to *generalized homomorphism density* for every coloured rooted ordered tree T' (i.e. every element of $SS_b(2k)$). Define

$$t(T', \{M_{2k,n}\}) = \int_{[0,1]^{b+1}} \prod_{\substack{(i,j)\in E \\ 0\leq i<j\leq b}} g_{2|E(i,j)|,n}(x_i, x_j) \prod_{0\leq i\leq b} dx_i. \tag{3.5.26}$$

Now we state the following corollary:

Corollary 3.5.16. *(**Generalized graphons**) Suppose W_n is the $n \times n$ Wigner matrix with independent entries $\{x_{ij,n}; i \leq j\}$ that satisfies (3.3.1), (3.3.2) and (3.3.4). Suppose*

$$\lim_{n\to\infty} t(T', \{M_{2k,n}\}) \quad \textit{exists for each coloured rooted tree } T'. \tag{3.5.27}$$

Then, the ESD of W_n converges weakly almost surely (or in probability as is the case in (3.3.4)) to a distribution whose odd moments are 0, and the $2k$th moment is given by

$$\sum_{T \text{ is a colored rooted tree}} \lim t(T, \{M_{2k,n}\}),$$

provided these moments determine a unique probability distribution.

Proof. We shall use Lemma 2.1.3 to prove this result. In order to verify the first moment condition, first observe that those words that do not belong to the set $SS(2k)$, do not contribute to the limiting $2k$th moment. Also from Lemma 3.2.5, we know that each word in $SS_b(2k)$ corresponds to a coloured rooted tree with b distinct colours. Hence the contribution for each such word (or tree) is $\lim_{n\to\infty} t(T, \{M_{2k,n}\})$. Thus we get the first moment condition.

The fourth moment condition can be verified in the same manner as in Step 2 of the proof of Theorem 3.3.1.

Finally, as these moments determine a unique probability distribution, we conclude that the ESD of W_n converges weakly almost surely (or in probability) to a symmetric distribution μ whose odd moments are 0 and the $2k$th moment is given by

$$\beta_{2k}(\mu) = \sum_{T \text{ is a colored rooted tree}} \lim t(T, \{M_{2k,n}\}).$$

□

Corollary 3.5.17. *Result 3.1.5 follows from Corollary 3.5.16.*

Proof. Let the matrix W_n be as defined in Result 3.1.5. As the variance σ_{ij}^2 are uniformly bounded and $\{a_{ij}\}$ satisfy (3.1.5), following the arguments in Step 1 in the proof of Corollary 3.5.12, we can assume that $|a_{ij}| \leq \eta_n \sqrt{n}$ where $\eta_n \downarrow 0$ as $n \to \infty$. This, with (3.1.6) implies that W_n satisfies the conditions of Corollary 3.5.16 with $\lim t(T', \{M_{2k,n}\}) = 0$ for all trees with less than $k+1$ colours. Now from Lemma 3.2.5, a colored rooted tree with $(k+1)$ vertices and $(k+1)$ colours means each vertex is of different colour. This tree may be then identified with a rooted tree with no colours, see Remark 3.2.6. Therefore we get,

$$\beta_{2k}(\mu_{zhu}) = \sum_{T} \lim t(T, H_n).$$

Thus, Result 3.1.5 follows from Corollary 3.5.16. Further, if $\lim \int M_{2,n}(x_1, x_2)\, dx_1 \to 1$, then the limit is semicircular. □

Remark 3.5.18. *Note that the uniform convergence of $\{g_{2k,n}\}$ to $\{g_{2k}\}$ and their integrability is a sufficient condition for (3.5.27). Also observe that under Assumption A,*

$$t(T', \{M_{2k,n}\}) \to \int_{[0,1]^{|V|}} \prod_{\substack{j=1 \\ (x_i,x_j)\in E}}^{b} g_{2E(x_i,x_j)}(x_i, x_j) \prod_{i\in V} dx_i,$$

for every coloured rooted tree T'. As seen in the proof of Theorem 3.3.1 and Lemma 3.2.5, these trees correspond to the words in $SS(2k)$, and thereby give rise to their contribution to the limiting moments.

3.5.6 Band, Block and Triangular matrices

Next we shall discuss results about *band, block* and *triangular* matrices. The previous works (Casati and Girko [1993a], Casati and Girko [1993b], Molchanov et al. [1992], Basu et al. [2012] Ding [2014], Zhu [2020]) that have dealt with these matrices have assumed that the distribution of the entries do not change with the size of the matrix. In the next few corollaries, we generalize the previous results regarding these matrices by allowing the distribution of the entries to vary with the size of the matrix.

In band matrices entries are non-zero only around the diagonal in the form of a band. As the dimension of the matrices increase, so does the number of non-zero elements around the diagonal.

Let m_n be a sequence of positive integers such that $m_n \to \infty$ and $m_n/n \to \alpha > 0$ as $n \to \infty$. There are two banding models—*periodic banding* and *non-periodic banding.*

Definition 3.5.19. ***(Band matrices)***

(a) ***Periodic banding**: W_n^b is the symmetric matrix with entries $y_{ij,n}$ where for $m_n \le n/2$,*

$$y_{ij,n} = \begin{cases} x_{ij,n} & \text{if } |i-j| \le m_n \quad \text{or } |i-j| \ge n - m_n, \\ 0 & \text{otherwise.} \end{cases} \tag{3.5.28}$$

(b) ***Non-periodic banding***: *W_n^B is the symmetric matrix with entries $y_{ij,n}$ where*

$$y_{ij,n} = \begin{cases} x_{ij,n} & \text{if } |i-j| \le m_n, \\ 0 & \text{otherwise.} \end{cases} \tag{3.5.29}$$

Corollary 3.5.20. *(**Periodic banding**) Suppose the random variables $\{x_{ij,n}\}$ in (3.5.28) satisfy Assumption A. Then the ESD of W_n^b converges weakly almost surely to a symmetric probability measure μ_α whose moments are determined by the functions $\{g_{2k}\}_{k\ge 1}$ and $\lim_{n\to\infty} m_n/n = \alpha \le 1/2$.*

Proof. For every n, define the function f_n on $[0,1]^2$ by

$$f_n(x,y) = \begin{cases} 1 & \text{if } |x-y| \le m_n/n \quad \text{or} \quad |x-y| \ge 1 - m_n/n, \\ 0 & \text{otherwise.} \end{cases} \tag{3.5.30}$$

Observe that the entries $y_{ij,n}$ of W_n^b can be written as $f_n(i/n, j/n)x_{ij,n}$.

As $|f_n| \le 1$, following Steps 2 and 4 of Theorem 3.3.1, the fourth moment condition and Carleman's condition follow immediately. Next observe that $\int_{[0,1]^2} f_n(x,y)\ dxdy$ converges to $\int_{[0,1]^2} f(x,y)\ dxdy$ where f is defined on $[0,1]^2$ as follows:

$$f(x,y) = \begin{cases} 1 & \text{if } |x-y| \le \alpha \quad \text{or } |x-y| \ge 1-\alpha, \\ 0 & \text{otherwise.} \end{cases} \tag{3.5.31}$$

So we have that $n\mathbb{E}[y_{ij,n}^{2k}] = f_n^{2k}(i/n,j/n)g_{2k,n}(i/n,j/n)$ and $f_n g_{2k,n}$ converges to fg_{2k}. Hence Condition (3.5.27) of the convergence of the generalized homomorphism densities holds. Now following Step 3 in Theorem 3.3.1, we get that only words in $SS(2k)$ contribute to the limiting moments. Moreover, for each word in $SS(2k)$ with b distinct letters, its contribution to the limiting moments is as follows:

$$\int_{[0,1]^{b+1}} \prod_{j=1}^{b} \left[g_{s_j}(x_{t_j}, x_{l_j}) \big[\mathbf{1}(|x_{t_j} - x_{l_j}| \ge 1-\alpha) + \mathbf{1}(|x_{t_j} - x_{l_j}| \le \alpha) \big] \right] dx_{t_1} dx_{l_1} \cdots dx_{l_b}.$$

The first moment condition follows similarly as Step 3 in the proof of Theorem 3.3.1. Hence the ESD of W_n^b converges weakly almost surely to a symmetric probability measure μ_α. $\square$

Remark 3.5.21. *Suppose $\{x_{ij,n}; i \leq j\}$ in the previous result are i.i.d for every fixed n and satisfies (3.3.1) and (3.3.2). Then from Corollary 3.5.20 we get that the ESD of W_n^b converges weakly almost surely to a symmetric probability measure μ_α whose $2k$th moment is given as follows:*

$$\beta_{2k}(\mu_\alpha) = \sum_{\pi \in SS(2k)} (2\alpha)^{|\pi|} c_\pi.$$

Corollary 3.5.22. *Result 3.1.6 follows from Corollary 3.5.20.*

Proof. Now if the entries of the matrix are $\{y_{ij,n}/\sqrt{n}\}$ where $\{y_{ij,n}\}$ are i.i.d. with finite mean and variance σ^2, then from Corollary 3.5.20 (see Remark 3.5.21), $c_2 = \sigma^2$ and $c_{2k} = 0$ for all $k \geq 2$. Therefore, the ESD of W_n^b converges to a symmetric probability measure μ_α whose $2k$th moment is given as follows:

$$\beta_{2k}(\mu_\alpha) = \sum_{\pi \in SS_k(2k)} (2\alpha)^{|\pi|} \sigma^{2k}.$$

As $SS_k(2k) = NC_2(2k)$, μ_α in this case is the semicircular distribution with variance $2\alpha\sigma^2$. Hence we can conclude the convergence of Theorem 4 of Casati and Girko [1993b] in the almost sure sense. □

Corollary 3.5.23. *(**Non-periodic banding**) Suppose the random variables $\{x_{ij,n}\}$ in (3.5.29) satisfy Assumption A. Then the ESD of W_n^B converges weakly almost surely to a symmetric probability measure μ_α whose moments are determined by the functions $\{g_{2k}\}_{k \geq 1}$ and $\alpha = \lim_{n \to \infty} m_n/n$.*

Proof. For every n, define the function f_n on $[0,1]^2$ by

$$f_n(x,y) = \begin{cases} 1 & \text{if } |x-y| \leq m_n/n, \\ 0 & \text{otherwise.} \end{cases} \tag{3.5.32}$$

Observe that the entries $y_{ij,n}$ of W_n^B can be written as $f_n(i/n, j/n)x_{ij,n}$.

As $|f_n| \leq 1$, following Steps 2 and 4 in the proof of Theorem 3.3.1, the fourth moment condition and Carleman's condition follow immediately. Next observe that $\int_{[0,1]^2} f_n(x,y)\, dxdy$ converges to $\int_{[0,1]^2} f(x,y)\, dxdy$ on $[0,1]^2$ where f is defined on $[0,1]^2$

as follows:

$$f(x,y) = \begin{cases} 1 & \text{if } |x-y| \leq \alpha, \\ 0 & \text{otherwise.} \end{cases} \tag{3.5.33}$$

So we have that $n\mathbb{E}[y_{ij,n}^{2k}] = f_n^{2k}(i/n, j/n) g_{2k,n}(i/n, j/n)$ and $f_n g_{2k,n}$ converges to $f g_{2k}$. Hence Condition (3.5.27) of the convergence of the generalized homomorphism densities holds. Now following Step 3 in the proof of Theorem 3.3.1, we get that only words in $SS(2k)$ contribute to the limiting moments. Moreover for each word in $SS(2k)$ with b distinct letters, its contribution to the limiting moment is as follows:

$$\int_{[0,1]^{b+1}} \prod_{j=1}^{b} g_{s_j}(x_{t_j}, x_{l_j}) \mathbf{1}(|x_{t_j} - x_{l_j}| \leq \alpha)\ dx_{t_1} dx_{l_1} \cdots dx_{l_b}.$$

So the first moment condition follows in the same way as Step 3 in the proof of Theorem 3.3.1. Hence the ESD of W_n^B converges weakly almost surely to a symmetric probability measure μ_α. □

Next we look into *triangular matrices.*

Definition 3.5.24. *(**Triangular Wigner**) The triangular Wigner matrix, denoted by W_n^u is the matrix whose entries $y_{i,n}$ are as follows:*

$$y_{ij,n} = \begin{cases} x_{ij,n} & \text{if } \ (i+j) \leq n+1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.5.34}$$

Corollary 3.5.25. *Suppose that the variables $\{x_{ij,n}; i \geq 0\}$ associated with the matrices W_n^u (as in (3.5.34)) are i.i.d. random variables with all moments finite for every fixed n, and satisfy Conditions (i) and (ii) of Corollary 3.5.3. Then the ESD of W_n^u converges weakly almost surely to some symmetric probability measure μ that depends on $\{c_{2k}\}_{k\geq 1}$.*

The proof of Corollary 3.5.25 follows in the same manner as that of Corollary 3.5.14 by considering $\sigma(x,y) = \mathbf{1}_{[x+y\leq 1]}$. So we omit the details.

Corollary 3.5.26. *Result 3.1.7 follows from Corollary 3.5.25.*

Proof. Observe that the entries of W_n^u are $\frac{y_{ij,n}}{\sqrt{n}}$ where $\{y_{ij,n}; i \geq 0\}$ are as in (3.5.34) and $\{x_{ij,n}; i \geq 0\}_{n\geq 1}$ are i.i.d. random variables with mean 0 and variance 1. Therefore, from Step 1 of the proof of Corollary 3.5.12, we can assume that the variables $\{x_{ij}\}$ are bounded. Hence the entries $\frac{y_{ij,n}}{\sqrt{n}}$ of W_n^u satisfy conditions of Corollary 3.5.25 with $c_2 = 1$ and $c_{2k} = 0$ for $k \geq 2$. Hence Corollary 3.5.25 implies that the ESD of W_n^u converges weakly almost surely to a non-random symmetric probability measure. □

We now discuss *block Wigner matrices* and prove a general result using the methods of proof of Theorem 3.3.1. We begin by defining the *Kronecker product of matrices.*

Definition 3.5.27. *If A and B are $m \times n$ and $p \times q$ matrices respectively, then the Kronecker product, $A \otimes B$ is the $pm \times qn$*

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & & \ddots & \vdots \\ a_{m1}B & a_{22}B & \dots & a_{mn}B \end{bmatrix}, \tag{3.5.35}$$

where a_{ij} is the (i,j)th entry of the matrix A.

Definition 3.5.28. *(**Block Wigner matrices**) Let E_{ml} be $d \times d$ elementary matrices with entry 1 at the (ml)th position and 0 otherwise. The symmetric matrix W'_n consisting of d^2 rectangular blocks, $W_n'^{(m,l)}, 1 \leq m, l \leq d$ is given by*

$$W'_n = \sum_{m,l} E_{ml} \otimes W_n'^{(m,l)}, \tag{3.5.36}$$

where $\otimes$ is the Kronecker product of matrices as in Definition 3.5.27. The blocks $W_n'^{(m,l)}, 1 \leq m \leq l \leq d$ are $n_m \times n_l$ rectangular random matrices with i.i.d. entries $\{x_{ij}\}$ inside each block but independent of the other blocks, subject to symmetry.

Corollary 3.5.29. *Let $\lim_{n\to\infty} \frac{n_m}{n} = \alpha_m > 0, 1 \leq m \leq d$. Suppose W'_n is the $n \times n$ symmetric random matrix as described in Definition 3.5.28. Also suppose the entries $x_{ij,n}$ associated to $W_n'^{(m,l)}$ in (3.5.36) have all moments finite and satisfy the following:*

(i) for each $k \in \mathbb{N}$ and $1 \le m \le l \le d$,

$$\infty > c_{2k}^{(m,l)} = \lim_{n\to\infty} n\ \mathbb{E}\left[x_{ij,n}^{2k}\right] \quad \textit{whenever } x_{ij,n} \quad \textit{is in the } (m,l)-\textit{th block}, \tag{3.5.37}$$

(ii) for each $k \in \mathbb{N}$ and all $1 \le i \le n_m, 1 \le j \le n_l$, $1 \le m \le l \le d$,

$$\lim_{n\to\infty} n^{\delta}\ \mathbb{E}\left[x_{ij,n}^{2k-1}\right] = 0 \quad \textit{for any} \quad \delta < 1. \tag{3.5.38}$$

Then the ESD of W_n' converges weakly almost surely to a symmetric probability measure $\tilde{\mu}$ whose moments are determined by $(c_{2k}^{(m,l)})_{k\ge1}$ and $(\alpha_m)_{m=1}^d$.

The proof is very similar to the proof of Theorem 3.3.1. We omit the details and give an outline for verification of only the first moment condition.

Let $c_{2k,n}^{(m,l)} = n\ \mathbb{E}\left[x_{ij,n}^{2k}\right]$ whenever $x_{ij,n}$ is in the $(m,l)-$th block. Also let $n_0 = 0$ and $\alpha_0 = 0$. Observe that in this case the sequence of functions $g_{2k,n}$ are given as follows:

$$g_{2k,n}(x,y) = c_{2k,n}^{(m,l)} \quad \text{when } (x,y) \in \left[\sum_{t=0}^{m-1} n_t/n, \sum_{t=0}^{m} n_t/n\right] \times \left[\sum_{t=0}^{l-1} n_t/n, \sum_{t=0}^{l} n_t/n\right].$$

This converges to g_{2k} defined as below:

$$g_{2k}(x,y) = c_{2k}^{(m,l)} \quad \text{when } (x,y) \in \left[\sum_{t=0}^{m-1} \alpha_t, \sum_{t=0}^{m} \alpha_t\right] \times \left[\sum_{t=0}^{l-1} \alpha_t, \sum_{t=0}^{l} \alpha_t\right].$$

Now following the same arguments as in Step 3 of the proof of Theorem 3.3.1 with the above function $(g_{2k})_{k\ge1}$ (that are determined by $(c_{2k}^{(m,l)})_{k\ge1}$ and $(\alpha_m)_{m=1}^d$), we have that the first moment condition holds.

Corollary 3.5.30. *Result 3.1.8 follows from Corollary 3.5.29.*

Proof. First note that just as in Step 1 in the proof of Corollary 3.5.12, the variables $\{a_{ij}\}$ can be assumed to be bounded. Then from (3.1.8) and (3.1.9), it can be shown easily that (3.5.37) and (3.5.38) hold. Thus the conditions of Corollary 3.5.29 are satisfied. Therefore there is a non-random symmetric probability distribution, say, μ_b which is the almost sure LSD of $\frac{1}{\sqrt{n}}W_n'$. Also it can be seen that only words with k distinct letters

($SS_k(2k)$) contribute to the limiting $2k$th moment. Also the words that do not belong to $SS(2k)$ contribute 0 in the limit. As $SS_k(2k)$ is the collection of all *Catalan words*, only the Catalan words contribute to the limiting moments. Hence we get Result 3.1.8 as a special case of Corollary 3.5.29. □

3.5.7 Simulations

Here are some simulations where the entries have different distributions. The LSDs are of course not universal. We demonstrate this phenomenon by considering standard normal entries with different distributions.

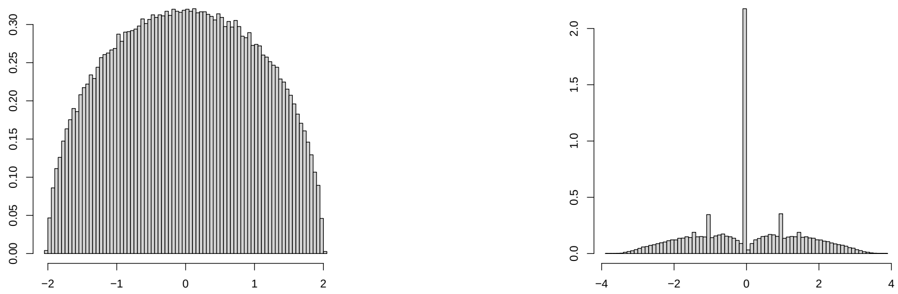


FIGURE 3.1: Histogram of the eigenvalues of W_n with entries i.i.d. $N(0,1)/\sqrt{n}$ (left) and Ber$(2/n)$ (right), for $n = 1000, 30$ replications.

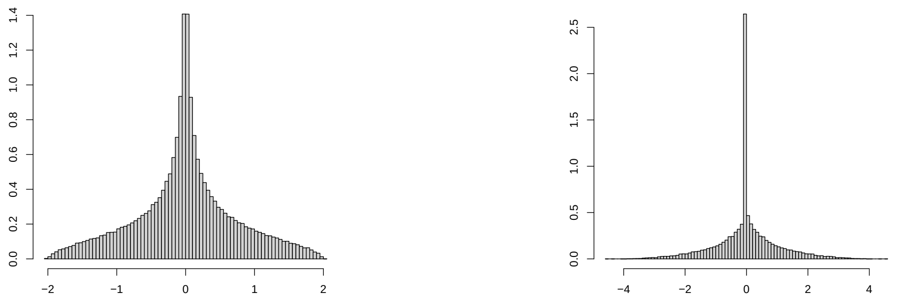


FIGURE 3.2: Histogram of the eigenvalues of W_n with entries $x_{ij}/\sqrt{n} = \frac{(i+j)^2}{2n^2}\frac{N(0,1)}{\sqrt{n}}$ (left) and $x_{ij} = \frac{(i+j)^2}{2n^2}\text{Ber}(2/n)$ (right) for $n = 1000, 30$ replications.

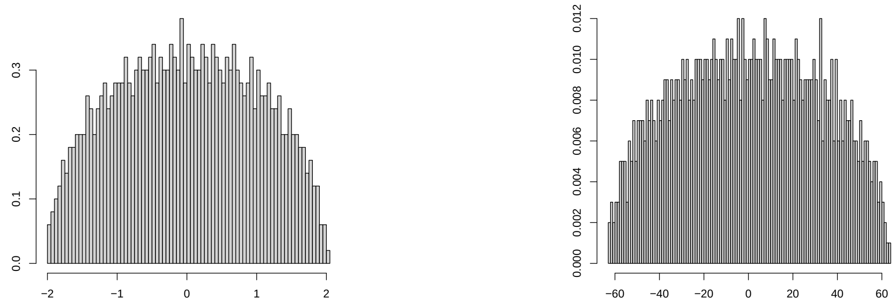


FIGURE 3.3: Histogram of the eigenvalues of (periodic) band Wigner matrices W_n^b with non-zero entries x_{ij} i.i.d. $\frac{N(0,1)}{\sqrt{n}}$ (left) and i.i.d. Ber$(3/n)$ (right) with $\alpha = 1/4$ for $n = 1000, 30$ replications.

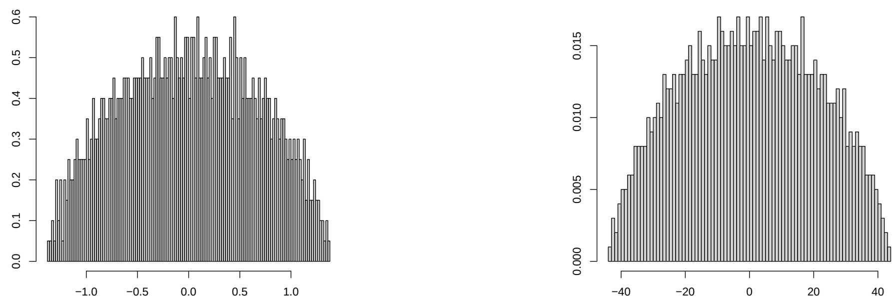


FIGURE 3.4: Histogram of the eigenvalues of (non-periodic) band Wigner matrices W_n^B with non-zero entries x_{ij} i.i.d. $\frac{N(0,1)}{\sqrt{n}}$ (left) and i.i.d. Ber$(3/n)$ (right) with $\alpha = 1/4$ for $n = 1000, 30$ replications.

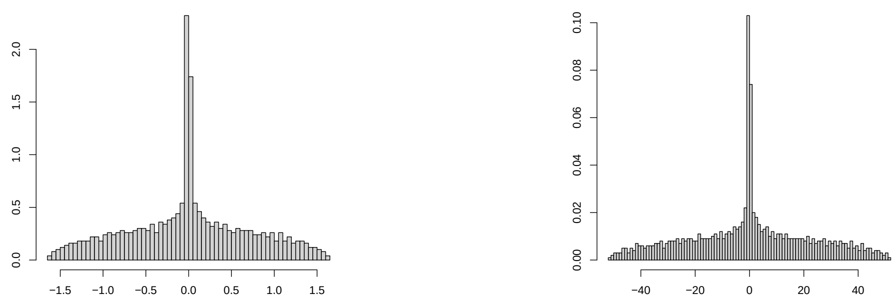


FIGURE 3.5: Histogram of the eigenvalues of triangular Wigner matrices W_n^u with non-zero entries x_{ij} i.i.d. $\frac{N(0,1)}{\sqrt{n}}$ (left) and i.i.d. Ber$(3/n)$ (right) for $n = 1000, 30$ replications.

Chapter 4

Other patterned matrices

In this chapter we shall look at the LSD of four other symmetric patterned matrices– the symmetric reverse circulant matrix $R^{(s)}$, the symmetric circulant matrix $C^{(s)}$, the symmetric Toeplitz matrix $T^{(s)}$, and the symmetric Hankel matrix, $H^{(s)}$:

$$R_n^{(s)} = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_{n-1} \\ x_1 & x_2 & x_3 & \cdots & x_0 \\ x_2 & x_3 & x_4 & \cdots & x_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n-1} & x_0 & x_1 & \cdots & x_{n-2} \end{bmatrix}, \quad C_n^{(s)} = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_1 \\ x_1 & x_0 & x_1 & \cdots & x_2 \\ x_2 & x_1 & x_0 & \cdots & x_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_0 \end{bmatrix}, \tag{4.0.1}$$

$$T_n^{(s)} = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_{n-1} \\ x_1 & x_0 & x_1 & \cdots & x_{n-2} \\ x_2 & x_1 & x_0 & \cdots & x_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n-1} & x_{n-2} & x_{n-3} & \cdots & x_0 \end{bmatrix}, \quad H_n^{(s)} = \begin{bmatrix} x_2 & x_3 & x_4 & \cdots & x_{n+1} \\ x_3 & x_4 & x_5 & \cdots & x_{n+2} \\ x_4 & x_5 & x_6 & \cdots & x_{n+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n+1} & x_{n+2} & x_{n+3} & \cdots & x_{2n} \end{bmatrix}. \tag{4.0.2}$$

In Section 4.1, we describe a few LSD results that already exist in the literature. These are closely related to the main results of this chapter that are described in Section 4.2. In Section 4.3, we give the detailed proofs of Theorems 4.2.2-4.2.4 using the moment method. Next, in Section 4.4, we discuss how the results of Section 4.1 can be obtained as special cases from Theorems 4.2.2-4.2.4. We conclude the chapter with some simulations and discussion on how these LSD results differ from those obtained in Chapter 3. *This chapter is based on Bose et al. [2021](Bose, Saha and Sen [2021]).*

4.1 Review of existing literature

We first recall some of the existing results about these matrices.

Fully i.i.d. entries: Previous works like, Bose and Mitra [2002], Hammond and Miller [2005], Bryc et al. [2006], Bose and Sen [2008] studied the spectral distribution of the above patterned matrices when the entries are fully i.i.d., i.e., the distribution of the entries does not change with the size of the matrix.

(i) Symmetric Reverse Circulant: $R^{(s)}$, displayed in (4.0.1), was first considered in Bose and Mitra [2002], where the authors showed the convergence of the ESD of $n^{-1/2}R_n^{(s)}$ to a non-random probability measure weakly in probability under the assumption that the entries are i.i.d. with finite third moment. Then Bose and Sen [2008] extended the result by proving almost sure convergence of the ESD. This is described in Result 4.1.1. A detailed proof is available in Bose [2018].

Result 4.1.1. *Suppose that the entries $\{x_n; n \geq 0\}$ are i.i.d. with mean 0 and variance 1. Then, as $n \to \infty$, the almost sure LSD of $\frac{1}{\sqrt{n}}R_n^{(s)}$ is the symmetrised Rayleigh distribution. This law, say $\mathcal{R}$, has the following density*

$$f(x) = |x| \exp(-x^2), \quad x \in \mathbb{R}.$$

The moments $\beta_k(\mathcal{R})$ of $\mathcal{R}$ are given by

$$\beta_k(\mathcal{R}) = \begin{cases} 0 & \text{if } k \text{ is odd}, \\ k! & \text{if } k \text{ is even}. \end{cases}$$

(ii) Symmetric Circulant: Recall $C_n^{(s)}$ defined in (4.0.1). The LSD of $n^{-1/2}C_n^{(s)}$ was first studied in Bose and Sen [2008]. It is worth mentioning that this matrix has an inherent connection to the palindromic Toeplitz matrix (Massey et al. [2007]). Theorem 2.4.2 in Bose [2018] presents a detailed proof of the following result.

Result 4.1.2. *Suppose that the entries $\{x_n; n \geq 0\}$ are i.i.d. with mean 0 and variance 1. Then, as $n \to \infty$, the almost sure LSD of $\frac{1}{\sqrt{n}} C_n^{(s)}$ is the standard normal distribution.*

(iii) Symmetric Toeplitz and Hankel: The study of the random Toeplitz and Hankel matrices, defined in (4.0.2), were initiated in a seminal paper by Bai [1999]. Hammond and Miller [2005] and Bryc et al. [2006] established the LSD of $n^{-1/2} T_n^{(s)}$. The LSD of $n^{-1/2} H_n^{(s)}$ was established by Bryc et al. [2006] and Liu and Wang [2011] with different techniques. We refer to Bose [2018] for a detailed proof of the following result on the LSD of Toeplitz and Hankel matrices.

Result 4.1.3. *Suppose that the entries $\{x_n; n \geq 0\}$ are i.i.d. with mean 0 and variance 1. Then, as $n \to \infty$, the almost sure LSD of $\frac{1}{\sqrt{n}} T_n^{(s)}$ and $\frac{1}{\sqrt{n}} H_n^{(s)}$ exist, say $\mathcal{L}_T$ and $\mathcal{L}_H$ respectively, and they are symmetric about 0.*

Sparse matrices: After having seen the fully i.i.d. case, it was natural to investigate cases where the distribution of the entries are i.i.d. but depend on the size of the matrix. One special case, in this regard, was investigated by Banerjee and Bose [2017], where they considered sparse patterned matrices and proves the convergence of the EESD. The almost sure convergence of the ESD does not occur in this case.

Result 4.1.4 (Theorem 3.1 and 3.2, Banerjee and Bose [2017])**.** *Suppose that for each fixed n, the entries of A_n (where A_n is any one of the above four matrices) are i.i.d. Bernoulli (p_n) with $np_n \to \lambda > 0$. Then the EESD of A_n converges weakly to a symmetric probability distribution ν_A.*

We now discuss some other variations of LSD results for these patterned matrices which include band matrices and triangular matrices.

Band matrices: Patterned matrices with banding have been studied in some previous works. Basak and Bose [2011] considered the LSD of the banded versions (Type I and Type II banding, see Definition 4.4.13) of the four matrices where the scaling depends on the number of non-zero entries in the matrices. Liu and Wang [2011] studied the band Toeplitz (Type I banding) and Hankel (Type II banding) matrices (see Definition 4.4.13), with a particular scaling, and proved that the ESD of these matrices converge weakly almost surely to symmetric probability distributions which depend on the proportion of non-zero entries.

Result 4.1.5. *(Basak and Bose [2011]) Let m_n be the bandwidth such that $m_n/n \to \alpha > 0$. Suppose the entries $\{x_i; i \leq m_n\}$ of the banded version A_n^b (with Type I banding) of A_n (where A_n is any one of the above four matrices), are i.i.d. with mean zero and variance 1. All other entries of A_n^b are zero.*

(i) If $m_n \leq n/2$, then the ESD of $\frac{1}{\sqrt{m_n}} C_n^{(s)b}$ converges weakly almost surely to the Normal distribution with mean 0 and variance 2.

(ii) If $m_n \leq n$, then the ESD of $\frac{1}{\sqrt{m_n}} R_n^{(s)b}$ converges weakly almost surely to the symmetrised Rayleigh distribution.

(iii) If $m_n \leq n$ (respectively $m_n \leq 2n$), then the ESD of $\frac{1}{\sqrt{m_n}} T_n^{(s)b}$ (and respectively $\frac{1}{\sqrt{m_n}} H_n^{(s)b}$) converge weakly almost surely to symmetric probability distributions.

Result 4.1.6. *(Basak and Bose [2011]) Let m_n be the bandwidth such that $m_n/n \to \alpha > 0$. Suppose the entries $\{x_i; i \leq m_n \text{ or } i \geq n - m_n\}$ or $\{x_i; i \leq m_n \text{ or } n - m_n \leq i \leq n + m_n\}$ of the banded version A_n^B (with Type II banding) of A_n (where A_n is any one of $R_n^{(s)}, T_n^{(s)}$ or $H_n^{(s)}$), are i.i.d. with mean zero and variance 1. All other entries of A_n^B are zero. Then*

(i) If $m_n \leq n/2$, then the ESD of $\frac{1}{\sqrt{2m_n}} R_n^{(s)B}$ converges weakly almost surely to the symmetrised Rayleigh distribution.

(ii) If $m_n \leq n/2$ (respectively $m_n \leq n$), then the ESD of $\frac{1}{\sqrt{m_n}} T_n^{(s)B}$ (and respectively $\frac{1}{\sqrt{2m_n}} H_n^{(s)B}$) converge weakly almost surely to symmetric probability distributions.

Result 4.1.7. *(Liu and Wang [2011]) Suppose $\{x_i; i \geq 0\}$ are independent with mean zero, variance 1 and have all moments uniformly bounded.*

(a) Let $T_n^{(s)b}$ be a real symmetric band Toeplitz matrix with bandwidth m_n such that $m_n/n \to \alpha \in (0, 1]$ and entries $\{x_i\}$. Then the ESD of the matrix $X_n^b = \frac{T_n^s b}{\sqrt{(2-\alpha)\alpha n}}$ converges weakly almost surely to a symmetric probability measure $\gamma_T(\alpha)$ whose even moments depend on α.

(b) Let $H_n^{(s)b}$ be a real symmetric band Hankel matrix with bandwidth m_n such that $m_n/n \to \alpha \in (0, 1]$ and entries $\{x_i\}$. Then the ESD of the matrix $Y_n^b = \frac{H_n^s b}{\sqrt{(2-\alpha)\alpha n}}$ converges weakly almost surely to a symmetric probability measure $\gamma_H(\alpha)$ whose even moments depend on α.

Triangular matrices: Triangular random matrices have gained importance since their consideration in Dykema and Haagerup [2004], where the authors considered the triangular Wigner matrices with Gaussian entries. Later LSD of symmetric triangular matrices with the above mentioned patterns was studied in Basu et al. [2012], where the authors proved the almost sure convergence of the ESD of appropriately scaled matrices. A generic triangular (symmetric) version A_n^u of a (symmetric) matrix A_n with link function L is of the form

$$A_n^u = \begin{bmatrix} x_{L(1,1)} & x_{L(1,2)} & x_{L(1,3)} & \cdots & x_{L(1,n)} \\ x_{L(2,1)} & x_{L(2,2)} & x_{L(2,3)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{L(n-1,1)} & x_{L(n-1,2)} & 0 & \cdots & 0 \\ x_{L(n,1)} & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Result 4.1.8 (Basu et al. [2012]). *Suppose the non-zero entries $\{x_{ij}; (i+j) \leq (n+1)\}$ of the symmetric triangular matrix, A_n^u, where A_n is either $C_n^{(s)}, T_n^{(s)}$ or $H_n^{(s)}$, are i.i.d. with mean zero and variance 1. Then the LSD of $\frac{1}{\sqrt{n}} A_n^u$ exists almost surely and is symmetric about zero.*

Our results: A common theme to be noted in most of the above results is that the distribution of an entry depends on its position in the matrix but remains unchanged with the size of the matrix, except for Result 4.1.4 (where the entries are i.i.d. only for every fixed n). In the previous chapter we have described a general result (Theorem 3.3.1) that tackles the case where the distribution of the (i,j)th entry depend on i, j and n. In this chapter, we prove LSD results for $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$ (see Theorems 4.2.2—4.2.4), with independent entries, under appropriate moment conditions. This yields Results 4.1.1—4.1.8 as special cases. Further, new results on matrices with variance profile, band and triangular patterned matrices emerge.

It is known that pair partitions play a very important role in the LSD of the above matrices with i.i.d. entries, see Results 4.1.1–4.1.3. For example, just like the moments of semicircle distribution are given via the set of all non-crossing pair partitions, the moments of the symmetrised Rayleigh distribution are given via the set of all symmetric pair partitions (see Section 2.5.3), and those of the standard normal distribution on the set of all pair partitions. We show that unlike the i.i.d. case, many other interesting partition classes are involved in the general independent case, see Theorem 4.2.2-4.2.4. Unlike the Wigner matrix, we shall see that in case of the symmetric reverse circulant and symmetric circulant matrices, the partitions with similar block structure (i.e., partitions that have same number of blocks with equal block sizes) contribute equally to the moments of the LSD. Hence the cumulants and half cumulants (see Sections 2.5.1 and 2.5.3) can be identified in these cases.

4.2 Main results

For the Wigner matrices the special symmetric partitions played a central role in the moments of the LSD. Two types of partitions of $[k]$ which play analogous role in the LSD of symmetric circulant (Toeplitz) and reverse circulant (Hankel) matrices, respectively are *even partitions* and *symmetric partitions* defined in Section 2.5.3.

Now suppose that the distribution of the entries for the nth matrix depends on n—they come from a triangular array of sequences $\{x_{i,n}; 0 \leq i \leq (n \text{ or } 2n)\}_{n\geq 1}$. To keep the notation simple, we shall often write x_i for $x_{i,n}$. We introduce the following set of assumptions on the entries. This is similar to Assumption A of Chapter 3.

Assumption B. Let $\{\tilde{g}_{k,n}; 0 \leq k \leq n\}$ be a sequence of bounded Riemann integrable functions on $[0,1]$. Suppose there exists a sequence $\{r_n\}$ with $r_n \in [0,\infty]$ such that

(i) for each $k \in \mathbb{N}$,

$$n\ \mathbb{E}\left[x_i^{2k}\mathbf{1}_{\{|x_i|\leq r_n\}}\right] = \tilde{g}_{2k,n}\Big(\frac{i}{n}\Big) \quad \text{for} \ \ 0 \leq i \leq n-1, \tag{4.2.1}$$

$$\lim_{n\to\infty} n^{\alpha} \sup_{0\leq i\leq n-1} \mathbb{E}\left[x_i^{2k-1}\mathbf{1}_{\{|x_i|\leq r_n\}}\right] = 0 \ \text{ for any } \alpha < 1. \tag{4.2.2}$$

(ii) The functions $\tilde{g}_{2k,n}$ converge uniformly to functions $\tilde{g}_{2k}$ for all $k \geq 1$.

(iii) Let $M_{2k} = \|\tilde{g}_{2k}\|$ (where $\|\cdot\|$ denotes the sup norm) and $M_{2k-1} = 0$ for all $k \geq 1$. Then, $\alpha_{2k} = \sum_{\sigma\in\mathcal{P}(2k)} M_\sigma$ satisfy *Carleman's condition,*

$$\sum_{k=1}^{\infty} \alpha_{2k}^{-\frac{1}{2k}} = \infty.$$

Remark 4.2.1. *Theorems 4.2.2—4.2.4 state the convergence of the EESD of the matrices,* $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}$ *and* $H_n^{(s)}$ *with independent entries that satisfy Assumption B. Here, the LSD of* $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}$ *and* $H_n^{(s)}$ *are random. The matrices,* $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}$ *and* $H_n^{(s)}$ *are more structured relative to the Wigner matrix and thus even with similar assumption on the entries, their LSD results vary widely from those of the latter. In the case when the entries of the matrices are all independent, it can be seen that there are* $\mathcal{O}(n)$ *number of random variables that constitute the matrix whereas the Wigner matrix is constituted by* $\frac{n(n+1)}{2}$ *random variables. Hence we have the variation in the results. Thus, although Assumption B is similar to Assumption A in Chapter 3, unlike the Wigner matrix (Theorem 3.3.1), where we show the almost sure convergence of the ESD, we can only conclude the convergence of the EESD for these matrices (Theorems 4.2.2—4.2.4).*

Now we state our theorems. Recall cumulants from Section 2.5.1 and half cumulants, symmetric and even partitions, $S(k)$ and $E(k)$ from Section 2.5.3.

Theorem 4.2.2. *Consider* $R_n^{(s)}$ *whose entries* $\{x_i; 0 \leq i < n\}$ *are independent and satisfy Assumption B. Let* Z_n *be the reverse circulant matrix with the entries* $y_i = x_i \mathbf{1}_{\{|x_i| \leq r_n\}}$. *Then*

(a) the EESD of Z_n *converges weakly to a symmetric probability measure* ν_R, *say. The moment sequence of* ν_R *is given by*

$$\beta_k(\nu_R) = \begin{cases} \sum_{\sigma \in S(k)} c_\sigma & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd,} \end{cases}$$

where $c_{2m} = \int_0^1 \tilde{g}_{2m}(t)\, dt, m \geq 1$. *Also* $\{c_{2m}\}_{m \geq 1}$ *is the half cumulant sequence of* ν_R. c_σ *is the multiplicative extension of* $\{c_{2m}\}_{m \geq 1}$.

In particular, if for every n, $\{x_{i,n}; 1 \leq i \leq n\}$ *are i.i.d., then the above holds with* $c_{2m} = \lim \tilde{g}_{2m,n}$ *(which are now constant functions).*

(b) Further if

$$\sum_{i=0}^{n-1} \mathbb{E}[x_i^2 \mathbf{1}_{\{|x_i| > r_n\}}] \to 0, \tag{4.2.3}$$

then the EESD of $R_n^{(s)}$ *converges weakly to* ν_R.

Next we deal with the symmetric circulant matrix with independent entries.

Theorem 4.2.3. *Consider $C_n^{(s)}$ whose entries $\{x_i; 0 \le i < n\}$ are independent, and satisfy Assumption B. Let Z_n be the $n \times n$ symmetric circulant matrix with entries $y_i = x_i \mathbf{1}_{\{|x_i| \le r_n\}}$. Then*

(a) *the EESD of Z_n converges weakly to a symmetric probability measure ν_C, and the moments of ν_C are given by*

$$\beta_k(\nu_C) = \begin{cases} \sum\limits_{\sigma \in E(k)} a_\sigma c_\sigma & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd,} \end{cases}$$

where $c_{2m} = 2\int_0^{\frac{1}{2}} \tilde{g}_{2m}(t)\ dt, m \ge 1$ are constants determined by the functions $\{\tilde{g}_{2k},\ k \ge 1\}$ and $a_{2n} = \frac{1}{2}\binom{2n}{n}, n \ge 1$. Also $\{0, a_2c_2, 0, a_4c_4, \ldots\}$ is the cumulant sequence of ν_C.

(b) *Further if*

$$\sum_i \mathbb{E}[x_i^2 \mathbf{1}_{\{|x_i| > r_n\}}] \to 0, \tag{4.2.4}$$

then the EESD of $C_n^{(s)}$ converges weakly to ν_C.

The following theorem is for the Topelitz and Hankel matrices with independent entries.

Theorem 4.2.4. *Consider $T_n^{(s)}$ (respectively, $H_n^{(s)}$) whose entries $\{x_i; 0 \le i < n\}$ are independent and satisfy Assumption B. Let Z_n be the $n \times n$ Toeplitz matrix (respectively, Hankel matrix) with entries $y_i = x_i \mathbf{1}_{\{|x_i| \le r_n\}}$. Then*

(a) *the EESD of Z_n converges weakly to a symmetric probability measure ν_T (respectively, ν_H) say, whose moment sequence is determined by the functions $\tilde{g}_{2k},\ k \ge 1$.*

(b) *Further if*

$$\sum_i \mathbb{E}[x_i^2 \mathbf{1}_{\{|x_i| > r_n\}}] \to 0, \tag{4.2.5}$$

then the EESD of $T_n^{(s)}$ (respectively, $H_n^{(s)}$) converges weakly to ν_T (respectively, ν_H) almost surely (or in probability).

Note that in the above theorems convergence of the ESD have not been claimed. We shall have more to say on this in Section 4.4.

Remark 4.2.5. *(i) The LSD of $R_n^{(s)}$ in Theorem 4.2.2 is the symmetrised Rayleigh distribution if and only if $\tilde{g}_{2k} = 0$ for all $k \geq 2$ and $c_2 = \int_0^1 \tilde{g}_2(t)\, dt = 1$.*

(ii) The LSD of $C_n^{(s)}$ in Theorem 4.2.3 is the standard normal distribution if and only if $c_2 = \int_0^1 \tilde{g}_2(t)\, dt = 1$ and $\tilde{g}_{2k} = 0$ for $k \geq 2$.

Remark 4.2.6. *The eigenvalues of the matrices $R^{(s)}$ and $C^{(s)}$ can be given via discrete Fourier transform of the input variables, see Bose and Mitra [2002]. We had previously tried the aprroach of writing the eigenvalues as a discrete Fourier transform as in Bose and Mitra [2002]. However, we still needed to compute certain groupings of terms in the sum and were not able to shorten the proofs effectively. Also the approach we have taken here helps us remain in the unified set up in the thesis as we would require to build this set up anyway for the Toeplitz and Hankel matrices. Hence, we have given the proofs in this unified approach throughout the thesis.*

4.3 Proofs of theorems

In Chapter 3, we have seen how first identifying the words that may contribute to the limiting moments led to the proof of Theorem 3.3.1. A similar approach is taken for proving the theorems in this chapter. As we have seen in the case of Wigner matrices, the existence of $\lim_{n\to\infty} \frac{|\Pi(\boldsymbol{\omega})|}{n^{b+1}}$ is intimately tied to the LSD, as this determines the words or partitions that possibly contribute positively to the limiting moments for each of the matrices. So, first we focus on finding $\lim_{n\to\infty} \frac{|\Pi(\boldsymbol{\omega})|}{n^{b+1}}$ for each of the four matrices.

4.3.1 Contributing words for the different matrices

In Lemmas 4.3.1–4.3.4, we shall identify which words can contribute positively to the limiting moments for each of the four matrices $R^{(s)}, C^{(s)}, T^{(s)}$ and $H^{(s)}$. Recall the notion of generating and non-generating vertices from Section 2.4. For a word with b

distinct letters, suppose the distinct letters appear for the first time at the positions $i_1, i_2, \ldots, i_b$. Then $\{\pi(i_j), 0 \le j \le b\}$ is the set of generating vertices and all other vertices $\pi(i)$, $i \neq i_j$ for any $j \in \{0, 1, \ldots, b\}$, is a linear combination of $\{\pi(i_j); i_j < i\}$ (here, linear combination is meant in the sense that the coefficients are allowed to be only integers). From (2.4.5), it is clear that to find $\lim_{n\to\infty} \frac{|\Pi(\boldsymbol{\omega})|}{n^{b+1}}$, we need to find in how many ways can the generating vertices be chosen freely (see discussion on generating and non-generating vertices and free choice of generating vertices in Section 2.4). Also recall the fact that if $\{\pi(i_j), 0 \le j \le b\}$ satisfy a non-trivial linear equation then one of the generating vertices has finitely many choices (no dependence on n) and in that case, $\lim_{n\to\infty} \frac{|\Pi(\boldsymbol{\omega})|}{n^{b+1}} \to 0$.

Now let us find $\lim_{n\to\infty} \frac{|\Pi(\boldsymbol{\omega})|}{n^{b+1}}$ for each of the four matrices.

Lemma 4.3.1. *(**Reverse Circulant Matrix**) For each word $\boldsymbol{\omega}$ with b distinct letters*

$$\lim_{n\to\infty} \frac{1}{n^{b+1}} |\Pi(\boldsymbol{\omega})| = \begin{cases} 1 & \text{if } \boldsymbol{\omega} \text{ is symmetric,} \\ 0 & \text{otherwise.} \end{cases} \tag{4.3.1}$$

Proof. Let $\pi \in \Pi_n[\omega]$. Let

$$t_i = \pi(i) + \pi(i-1) \text{ for } 1 \le i \le 2k. \tag{4.3.2}$$

Clearly, $\boldsymbol{\omega}[i] = \boldsymbol{\omega}[j]$ if and only if

$$(\pi(i-1) + \pi(i) - 2) \bmod n = (\pi(j-1) + \pi(j) - 2) \bmod n,$$

that is,

$$t_i - t_j = 0,\ n \text{ or } -n.$$

Now fix an $\boldsymbol{\omega}$ with b distinct letters. Suppose the b distinct letters appear for the first time at the positions $i_1, i_2, \ldots, i_b$. Clearly $t_{i_1} = t_1$. If the first letter appears again in the jth position, then

$$t_j = t_1 \pmod n.$$

Similarly, for every i, $1 \le i \le 2k$,

$$t_i = t_{i_j} \pmod n \quad \text{for some } j \in \{1, 2, \ldots, b\}. \tag{4.3.3}$$

We break the proof into three steps.

Step 1. $\{\pi(i_j), 0 \le j \le b\}$ (where $i_0 = 0$) can be chosen freely if and only if $\{\pi(0), t_{i_j}; 1 \le j \le b\}$ do not satisfy any non-trivial linear relation.

To see this, observe, for any $1 \le j \le b$

$$\pi(i_j) = t_{i_j} - t_{i_j - 1} + \cdots + (-1)^{i_j+1} t_1 + (-1)^{i_j} \pi(0). \tag{4.3.4}$$

Now, using the formula for t_i in (4.3.2), we have that all the t_is can be written as linear combinations (with coefficients from $\mathbb{Z}$) of $t_{i_m}, m < j$. Hence from (4.3.4), it is clear that if $\{\pi(i_j), 0 \le j \le b\}$ (where $i_0 = 0$) can be chosen freely then $\{\pi(0), t_{i_j}; 1 \le j \le b\}$ do not satisfy any non-trivial linear relation and vice versa.

This completes the proof of this step.

Step 2. If $\boldsymbol{\omega}$ is not symmetric, then $\lim_{n\to\infty} \frac{1}{n^{b+1}} |\Pi(\boldsymbol{\omega})| = 0$.

Suppose $\boldsymbol{\omega}$ is of length $2k$. Then for any corresponding circuit,

$$(t_1 + t_3 + \cdots + t_{2k-1}) - (t_2 + t_4 + \cdots + t_{2k}) = \pi(0) - \pi(2k) = 0. \tag{4.3.5}$$

Therefore, using (4.3.3), we see that there exists $\alpha_j \in \mathbb{Z}$ for all $1 \le j \le b$ such that

$$\pi(0) - \pi(2k) = \alpha_1 t_{i_1} + \alpha_2 t_{i_2} + \cdots + \alpha_b t_{i_b} = 0 \pmod{n}.$$

However, $|\pi(0) - \pi(2k)| \le (n-1)$. Therefore,

$$\alpha_1 t_{i_1} + \alpha_2 t_{i_2} + \cdots + \alpha_b t_{i_b} = 0.$$

Now for $\{\pi(0), t_{i_j}, 1 \le j \le b\}$ to be such that they do not satisfy a non-trivial linear relation, we must have $\alpha_j = 0$ for all $j \in \{1, 2, \ldots, b\}$. Thus for each j,

$$\big|\{l : l \text{ odd}, t_l = t_{i_j} \pmod{n}\}\big| = \big|\{l : l \text{ even}, t_l = t_{i_j} \pmod{n}\}\big|.$$

i.e., each letter in $\boldsymbol{\omega}$ appears equal number of times at odd and even places, and it is symmetric.

Now if the length of the word is odd, say $2k+1$, then

$$(t_1 + t_2 + \cdots + t_{2k+1}) - (t_2 + t_4 + \cdots + t_{2k}) = \pi(2k+1) + \pi(0) = 2\pi(0).$$

Now substituting t_i by t_{i_j}'s using (4.3.3) in the above equation, we see that $\pi(0), t_{i_1}, \ldots, t_{i_b}$ satisfy a non-trivial linear relation.

Therefore, from Step 1 and the discussion above, we have that, when the word is not symmetric, $\{\pi(i_j), 0 \le j \le b\}$ satisfy a non-trivial linear equation. Hence, at least one of the generating vertices has at most a finite number of choices as $n \to \infty$. As a result,

$$\lim_{n\to\infty} \frac{1}{n^{b+1}} |\Pi(\boldsymbol{\omega})| = 0 \quad \text{for any non-symmetric word } \boldsymbol{\omega}.$$

Step 3. Suppose $\boldsymbol{\omega}$ is symmetric with b distinct letters. Then $\pi(i_j)$ $(0 \le j \le b)$ can be chosen freely, and then the non-generating vertices have a unique choice.

Since the word is symmetric, from (4.3.2) and (4.3.3), we have, for each $1 \le j \le b$,

$$\big|\{l : l \text{ odd}, \, t_l = t_{i_j} \pmod n\}\big| = \big|\{l : l \text{ even}, \, t_l = t_{i_j} \pmod n\}\big|.$$

Thus,

$$\pi(0) - \pi(2k) = (t_1 + t_3 + \cdots + t_{2k-1}) - (t_2 + t_4 + \cdots + t_{2k}) = Mn \quad \text{for some} \quad M \in \mathbb{Z}.$$

However, $|\pi(0) - \pi(2k)| \le (n-1)$. Therefore, $M = 0$. Hence (4.3.5) (circuit condition) automatically holds for $\boldsymbol{\omega}$. Thus there is no additional constraint in choosing the generating vertices. Now we have to ensure that once the generating vertices have been chosen, the non-generating vertices have unique choices.

First we choose $\pi(0)$ and $\pi(1)$ freely, i.e., there are $n(n-1)$ choices for $\pi(0)$ and $\pi(1)$. Next, if $\pi(2)$ is a generating vertex, we choose it freely. If $\pi(2)$ is not a generating vertex, we consider the relation between t_1 and t_2. Clearly as $\pi(2)$ is not a generating vertex, $\boldsymbol{\omega}[2]$ is an old letter (i.e., a letter that has appeared earlier in the word), and hence $t_1 = t_2$ which implies $\pi(0) = \pi(2)$.

To prove that the other non-generating vertices can be chosen uniquely, we argue inductively from left to right. Suppose $\pi(i)$ is a non-generating vertex, and the non-generating vertices among $\pi(t)$ $(1 \leq t \leq i-1)$ have already been chosen uniquely. Suppose the jth distinct letter appears at the ith position. Then from (4.3.3) we have

$$\pi(i) = \pi(i_j - 1) + \pi(i_j) - \pi(i-1) \pmod n, \quad \text{for some } j \quad \text{where } i_j \leq i. \tag{4.3.6}$$

Therefore $\pi(i) = A - n, A, A + n$, where $A = \pi(i_j - 1) + \pi(i_j) - \pi(i-1)$. Observe that $-(n-1) \leq A \leq 2n$, and only one of the values $A - n, A, A + n$ can be between 0 and $n - 1$. Hence $\pi(i)$ can be determined *uniquely* from (4.3.6) as $\pi(i_j - 1), \pi(i_j), \pi(i-1)$ have already been determined.

As a consequence of the above three steps (4.3.1) is proved. $\square$

Lemma 4.3.2. *(**Symmetric circulant matrix**) For each word $\boldsymbol{\omega}$ with b distinct letters, let k_i be the number of times the ith distinct letter appeared in $\boldsymbol{\omega}$. Then,*

$$\lim_{n\to\infty} \frac{1}{n^{b+1}} |\Pi(\boldsymbol{\omega})| = \begin{cases} \prod_{i=1}^{b} \binom{k_i - 1}{\frac{k_i}{2}} & \text{if } \boldsymbol{\omega} \text{ is even,} \\ 0 & \text{otherwise.} \end{cases} \tag{4.3.7}$$

Proof. First note the following. Let

$$s_i = \pi(i) - \pi(i-1) \quad \text{for } 1 \leq i \leq 2k. \tag{4.3.8}$$

Clearly, $\boldsymbol{\omega}[i] = \boldsymbol{\omega}[j]$ if and only if $|n/2 - |s_i|| = |n/2 - |s_j||$. This is the same as saying, $s_i - s_j = 0,\ n$ or $-n$, or $s_i + s_j = 0,\ n$ or $-n$.

Now we fix an $\boldsymbol{\omega}$ with b distinct letters which appear at $i_1, i_2, \ldots, i_b$ positions for the first time. Clearly $s_{i_1} = s_1$. If the first letter also appears in the jth position, then

$$s_j = s_1 \pmod n \text{ or } s_j = -s_1 \pmod n.$$

Similarly, for every i, $1 \leq i \leq 2k$,

$$s_i = s_{i_j} \pmod n \quad \text{or} \quad s_i = -s_{i_j} \pmod n, \quad \text{for some } j \in \{1, 2, \ldots, b\}. \tag{4.3.9}$$

We break the proof into three steps.

Step 1. (Similar to Step 1 in the proof of Lemma 4.3.1) $\{\pi(i_j), 0 \leq j \leq b\}$ (where $i_0 = 0$) can be chosen freely if and only if $\{\pi(0), s_{i_j}; 1 \leq j \leq b\}$ do not satisfy any non-trivial linear relation.

To see this, observe that, for any $1 \leq j \leq b$

$$\pi(i_j) = s_{i_j} + s_{i_j-1} + \cdots + s_1 + \pi(0). \tag{4.3.10}$$

Now, using the formula for s_i in (4.3.8), we have that all the s_is can be written as linear combinations (with coefficients from $\mathbb{Z}$) of $s_{i_m}, m < j$. Hence from (4.3.10), it is clear that if $\{\pi(i_j), 0 \leq j \leq b\}$ (where $i_0 = 0$) can be chosen freely then $\{\pi(0), s_{i_j}; 1 \leq j \leq b\}$ do not satisfy any non-trivial linear relation and vice versa.

This completes the proof of this step.

Step 2. (Similar to Step 2 in the proof of Lemma 4.3.1) If $\boldsymbol{\omega}$ is not an even word, then $\lim_{n\to\infty} \frac{1}{n^{b+1}} |\Pi(\boldsymbol{\omega})| = 0$.

To see this, observe that, if the length of the word is k, then for any corresponding circuit π,

$$\sum_{i=1}^{k} s_i = \pi(0) - \pi(k) = 0. \tag{4.3.11}$$

Therefore, using (4.3.9), we see that for all $1 \leq j \leq b$, there exists $\alpha_j \in \mathbb{Z}$ such that

$$\pi(0) - \pi(k) = \alpha_1 s_{i_1} + \alpha_2 s_{i_2} + \cdots + \alpha_b s_{i_b} = 0 \ (\text{mod } n).$$

However, $|\pi(0) - \pi(k)| \leq (n-1)$. Thus,

$$\alpha_1 s_{i_1} + \alpha_2 s_{i_2} + \cdots + \alpha_b s_{i_b} = 0.$$

For $\{\pi(0), s_{i_j}, 1 \leq j \leq b\}$ to be such that they do not satisfy a non-trivial relation, we must have $\alpha_j = 0$ for all $j \in \{1, 2, \ldots, b\}$. Therefore for each j,

$$\left|\{l : s_l = s_{i_j} \ (\text{mod } n)\}\right| = \left|\{l : s_l = -s_{i_j} \ (\text{mod } n)\}\right|. \tag{4.3.12}$$

That is, each letter appears an even number of times, and the word is even. So, if the word is not even, then using Step 1 and the discussion above, it follows that $\{\pi(i_j), 0 \leq j \leq b\}$ satisfy a non-trivial linear relation. Hence, at least one of the generating vertices has at most finite number of choices. As a result, $\lim_{n\to\infty} \frac{1}{n^{b+1}}|\Pi(\boldsymbol{\omega})| = 0$ if $\boldsymbol{\omega}$ is not even.

Step 3. Suppose that $\boldsymbol{\omega}$ is an even word of length $2k$ with b distinct letters. Then $\pi(i_j)$ $(0 \leq j \leq b)$ can be chosen freely. Subsequently, the non-generating vertices can be chosen in $\displaystyle\prod_{i=1}^{b} \binom{k_i - 1}{\frac{k_i}{2}}$ ways.

Since the word is even, from (4.3.8) and (4.3.9), we have for each j,

$$\big|\{l : s_l = s_{i_j} \pmod n\}\big| = \big|\{l : s_l = -s_{i_j} \pmod n\}\big|.$$

Thus,

$$\pi(0) - \pi(2k) = \sum_{i=1}^{2k} s_i = Mn \quad \text{for some} \quad M \in \mathbb{Z}.$$

However, $|\pi(0) - \pi(2k)| \leq (n-1)$. Therefore, $M = 0$. Hence, the circuit condition (4.3.11) is automatically satisfied. Thus there is no additional constraint for choosing the generating vertices.

First we choose $\pi(0)$ and $\pi(1)$ freely. Next if $\pi(2)$ is a generating vertex, we choose it freely. Else we consider the relation between s_1 and s_2. Clearly, as $\pi(2)$ is not a generating vertex, the letter in the second position of $\boldsymbol{\omega}$ has appeared earlier, and hence either $s_1 = s_2$ or $s_2 = -s_1$. If the former is true then $\pi(2) = 2\pi(1) - \pi(0) \pm n$. Else $\pi(2) = \pi(0)$. In any case we see that there is only one value of $\pi(2)$ between 1 and n.

Now we shall show that all other non-generating vertices can be chosen uniquely *once a particular set of signs have been chosen in* (4.3.9) *subject to* (4.3.12). We argue inductively from left to right. Suppose $\pi(i)$ is a non-generating vertex, and the non-generating vertices among $\pi(l)$ $(1 \leq l \leq i-1)$ have been chosen uniquely. Suppose the jth distinct letter appears at the ith position. Then we know from (4.3.9) that

$$\pi(i) = \pm(\pi(i_j - 1) - \pi(i_j)) + \pi(i-1) \pmod n, \quad \text{for some } j \quad \text{where} \quad i_j \leq j. \tag{4.3.13}$$

Therefore $\pi(i) = A - n, A, A + n$, where

$$A = \pi(i_j - 1) - \pi(i_j) + \pi(i - 1) \quad \text{or} \quad A = \pi(i_j) - \pi(i_j - 1) + \pi(i - 1), \tag{4.3.14}$$

according as the sign chosen in (4.3.9) ($s_i + s_{i_j} = 0$ or $s_i - s_{i_j} = 0$).

Observe that in either cases, $-(n-1) \leq A \leq 2n$, and only one of the values $A - n, A, A+n$ can be between 1 and n. Hence $\pi(i)$ can be determined *uniquely* from (4.3.13) once the set of signs in equation (4.3.9) are fixed.

Now as $\boldsymbol{\omega}$ is an even word where each distinct letter appears $k_1, k_2, \ldots, k_b$ times (and hence each k_i is even), we first observe that by (4.3.12) there are total of $\prod_{i=1}^{b} \binom{k_i - 1}{\frac{k_i}{2}}$ set of equations available for determining the non-generating vertices, once the generating vertices are chosen. We now show that all the choices are allowed. Note that, from the link function we have $s_i + s_{i_j} = 0$ or $s_i - s_{i_j} = 0$ for all $i \in \{1, 2, \ldots, 2k\}$ and some $j \in \{1, 2, \ldots, b\}$. Now the jth letter appears k_j times in $\boldsymbol{\omega}$, and first time at the i_jth position in the word. Now, because of (4.3.12), $\big|\{l : s_l = -s_{i_j} \pmod n)\}\big| = k_j/2$. This can occur in $\binom{k_j-1}{k_j/2}$ ways. Further, we have seen from the above argument in the previous paragraph (see (4.3.14)) that any particular set of equations arising out of (4.3.9) determine the non-generating vertices uniquely.

As an example consider the word *abcabcabcabc*. For this word,

$$\begin{aligned}
&s_1 + s_4 = 0, \;\; s_1 - s_7 = 0, \;\; s_1 + s_1 = 0, \\
&s_2 + s_5 = 0, \;\; s_2 - s_8 = 0, \;\; s_2 + s_1 = 0, \\
&s_3 + s_6 = 0, \;\; s_3 - s_9 = 0, \;\; s_3 + s_1 = 0,
\end{aligned}$$

and

$$\begin{aligned}
&s_1 + s_4 = 0, \;\; s_1 - s_7 = 0, \;\; s_1 + s_1 = 0, \\
&s_2 + s_5 = 0, \;\; s_2 + s_8 = 0, \;\; s_2 - s_1 = 0, \\
&s_3 + s_6 = 0, \;\; s_3 - s_9 = 0, \;\; s_3 + s_1 = 0,
\end{aligned}$$

correspond to two different sets of equations. Each of these sets (4.3.9) determines the non-generating vertices uniquely.

Suppose for any word $\boldsymbol{\omega}$, we start with one particular choice of equations (4.3.9) of $\{s_i, 1 \leq i \leq 2k\}$. Consider another set of equations. We know that the equations differ only in sign. If the change of sign occurs first at the mth position (so that if $s_m = s_{i_j} \pmod n$ in the first set, then in the other set of equations we have, $s_m = -s_{i_j} \pmod n$), then for that m, the corresponding value of A in (4.3.14) for determining $\pi(m)$ changes because $\pi(m-1)$ remains unaltered and s_{i_j} changes sign in the expression. This changes the value of $\pi(m)$ obtained from the previous set of equations. Hence we conclude that each set of equations gives a distinct (unique) choice for $(\pi(0), \pi(1), \ldots, \pi(2k-1), \pi(2k))$. Therefore,

$$\lim_{n\to\infty} \frac{1}{n^{b+1}} |\Pi(\boldsymbol{\omega})| = \prod_{i=1}^{b} \binom{k_i - 1}{\frac{k_i}{2}} \quad \text{if } \boldsymbol{\omega} \text{ is even.}$$

This completes the proof of the lemma. □

Lemma 4.3.3. ***(Toeplitz Matrix)*** *Suppose $\boldsymbol{\omega}$ is a word with b distinct letters. Then $\lim_{n\to\infty} \frac{1}{n^{b+1}} |\Pi(\boldsymbol{\omega})| = \alpha(\boldsymbol{\omega}) > 0$ if and only if $\boldsymbol{\omega}$ is an even word.*

Proof. Let

$$s_i = \pi(i) - \pi(i-1) \quad \text{for } 1 \leq i \leq 2k.$$

Clearly, $\boldsymbol{\omega}[i] = \boldsymbol{\omega}[j]$ if and only if $|s_i| = |s_j|$, that is, $s_i - s_j = 0$ or $s_i + s_j = 0$.

Now we fix an $\boldsymbol{\omega}$ with b distinct letters. Suppose $i_1, i_2, \ldots, i_b$ are the positions where new letters made their first appearances, and let k_i be the number of times the ith distinct letter appeared in $\boldsymbol{\omega}$. Clearly $s_{i_1} = s_1$. If the first letter appears in the jth position, then $s_j = s_1$ or $s_j = -s_1$. Similarly, for every i, $1 \leq i \leq 2k$,

$$s_i = s_{i_j} \text{ or } s_i = -s_{i_j} \tag{4.3.15}$$

for some $j \in \{1, 2, \ldots, b\}$. We split the proof into a few steps.

Step 1. (Similar to Step 1 in the proof of Lemma 4.3.2) $\{\pi(i_j), 0 \leq j \leq b\}$ (where $i_0 = 0$) can be chosen freely if and only if $\{\pi(0), s_{i_j}; 1 \leq j \leq b\}$ do not satisfy any non-trivial linear relation.

This follows from the same argument as in Step 1 in the proof of Lemma 4.3.2.

Step 2. (Similar to Step 2 in the proof of Lemma 4.3.2) If $\boldsymbol{\omega}$ is not an even word, then, $\lim_{n\to\infty}\frac{1}{n^{b+1}}|\Pi(\boldsymbol{\omega})| = 0$.

To see this, observe that the circuit condition gives

$$\sum_{i=1}^{k} s_i = \pi(0) - \pi(k) = 0. \tag{4.3.16}$$

Therefore, using (4.3.15), we see that for all $1 \leq j \leq b$, there exists $\alpha_j \in \mathbb{Z}$ such that

$$\alpha_1 s_{i_1} + \alpha_2 s_{i_2} + \cdots + \alpha_b s_{i_b} = 0.$$

Since s_{i_j}'s do not satisfy any non-trivial linear relation, we must have $\alpha_j = 0$ for all $j \in \{1, 2, \ldots, b\}$. Therefore for each j,

$$\big|\{l : s_l = s_{i_j}\}\big| = \big|\{l : s_l = -s_{i_j}\}\big|. \tag{4.3.17}$$

So each letter appears an even number of times, and the word is even. So, if the word is not even, it follows that $\{\pi(i_j), 0 \leq j \leq b\}$ satisfy a non-trivial linear relation. Hence, at least one of the generating vertices has at most finite number of choices. Therefore

$$\lim_{n\to\infty}\frac{1}{n^{b+1}}|\Pi(\boldsymbol{\omega})| = 0 \ \text{ if } \boldsymbol{\omega} \text{ is not an even word.}$$

Step 3. Suppose $\boldsymbol{\omega}$ is an even word with b distinct letters. Using similar arguments as in Step 3 of the proof of Lemma 4.3.2, we can see that the circuit condition is automatically satisfied. In Lemma 4.3.2, we had obtained $\prod_{i=1}^{b}\binom{k_i-1}{\frac{k_i}{2}}$ different sets of linear combinations corresponding to each word $\boldsymbol{\omega}$. In this case too, we have $\prod_{i=1}^{b}\binom{k_i-1}{\frac{k_i}{2}}$ different sets of linear combinations corresponding to each word $\boldsymbol{\omega}$ arising from (4.3.15) and (4.3.17). However, unlike the symmetric circulant case, having chosen the generating vertices, each set of linear combinations here *does not* determine the non-generating vertices uniquely. We will see below a different phenomenon is observed in this case.

First we fix the generating vertices $\pi(i_j)$, $0 \leq j \leq b$. Consider $\pi(i)$. By (4.3.15),

$$\pi(i) = \pm(\pi(i_j) - \pi(i_j - 1)) + \pi(i-1) = \pm s_{i_j} + \pi(i-1) \ \text{ for some } j. \tag{4.3.18}$$

Let

$$v_i = \frac{\pi(i)}{n} \text{ for } 0 \le i \le 2k \quad \text{and} \quad u_j = \frac{s_j}{n} \text{ for } 1 \le j \le 2k.$$

Clearly, $\pi(i) = \pi(i-1) \pm s_{i_j}$ whenever the ith letter in $\boldsymbol{\omega}$ is same as the jth distinct letter that appeared first at the i_jth position. Therefore $v_i = v_{i-1} \pm u_{i_j}$. Also observe that $v_1 = \ v_0 + u_{i_1}$ where $u_{i_1} = u_1$.

Let

$$S = \{\pi(i_j) : 0 \le j \le b\} \quad \text{and} \quad S' = \{i : \ \pi(i) \notin S\}.$$

That is, S is the set of all distinct generating vertices and S' is the set of all *indices* of the non-generating vertices. We have the following claim.

Claim: For any $1 \le i \le 2k$, $v_i = v_0 + \sum_{j=1}^{i} \alpha_{ij} u_{i_j}$, where α_{ij} depends on the choice of the sign in (4.3.18).

We prove this by induction. We know that $\pi(1) \in S$. Clearly, $v_1 = s_1 + v_0$. Now either $\pi(2) \in S$ or $2 \in S'$. If $\pi(2) \in S$, then $v_2 = s_2 - v_1$ and $v_1 = u_1 - v_0$. Therefore $v_2 = u_2 - s_1 + u_0$. If $2 \in S'$, then $u_2 = \pm u_1$ and $v_2 = v_1 \pm u_1$. So either $v_2 = u_1 + v_0 + u_1$ or $v_2 = u_1 + v_0 - u_1 = v_0$. Hence the claim is true for $i = 2$.

Now we assume that the claim is true for all $j < i$, and try to prove it for i. Then either $\pi(i) \in S$ or $i \in S'$. If $\pi(i) \in S$, then

$$\begin{aligned} v_i &= u_i + v_{i-1} \\ &= u_i + v_0 + \sum_{j=1}^{i-1} \alpha_{(i-1)j} u_{i_j} \quad \text{(by induction hypothesis)} \\ &= v_0 + \sum_{j=1}^{i} \alpha_{ij} u_{i_j}, \end{aligned}$$

where $\alpha_{ii} = 1$. If $i \in S'$, then there exists j such that $i_j < i$ and $u_i = \pm u_{i_j}$. Then either $v_i = v_{i-1} + u_{i_j}$, or $v_i = v_{i-1} - u_{i_j}$. Hence

$$\text{either} \quad v_i = v_0 + \sum_{j=1}^{i-1} \alpha_{(i-1)j} u_{i_j} + u_{i_j}, \quad \text{or} \quad v_i = v_0 + \sum_{j=1}^{i-1} \alpha_{(i-1)j} u_{i_j} - u_{i_j}.$$

Therefore $v_i = v_0 + \sum_{j=1}^{i} \alpha_{ij} u_{i_j}$ where $\alpha_{ij} = \alpha_{(i-1)j} + 1$ or $\alpha_{(i-1)j} - 1$ (depending on the sign of the above equation). Thus the claim is proved.

Therefore for $1 \leq i \leq 2k$, we have

$$v_i = v_0 + LC^T_{i,u}(u_S),$$

where $LC^T_{i,u}(u_S)$ denotes a linear combination of $\{u_i : \pi(i) \in S\}$.

Also, for $1 \leq i \leq 2k$, we have

$$v_i = v_0 + LC^T_i(v_S),$$

where $LC^T_i(v_S)$ denotes a linear combination of $\{v_i : \pi(i) \in S\}$ arising from (4.3.15). As discussed before there are $\prod_{i=1}^{b} \binom{k_i - 1}{\frac{k_i}{2}}$ different sets of linear combinations corresponding to each word $\boldsymbol{\omega}$ due to the sign chosen in (4.3.15). Let us denote the collection of such different sets of linear combinations corresponding to a particular $\boldsymbol{\omega}$ as $LC^T_{\boldsymbol{\omega}}$. Now as $LC^T_{i,u}(u_S)$ and $LC^T_i(v_S)$ arising from the same signs are related to each other (as we have seen in the above claim), we can use $LC^T_{\boldsymbol{\omega}}$ as the collection of different sets of linear combinations of v_S as well as u_S corresponding to a particular $\boldsymbol{\omega}$.

Let $U_n = \{0, 1/n, \ldots, (n-1)/n\}$. From (2.4.5), it is easy to see that for $\boldsymbol{\omega}$ of length $2k$,

$$\begin{aligned}\big|\Pi(\boldsymbol{\omega})\big| = \big|\big\{&(v_0, v_1, \ldots, v_{2k}) : v_i \in U_n \text{ for } 0 \leq i \leq 2k, v_0 = v_{2k},\\ &L(v_{i-1}, v_i) = L(v_{j-1}, v_j) \text{ whenever } \boldsymbol{\omega}[i] = \boldsymbol{\omega}[j]\big\}\big|.\end{aligned}$$

Hence

$$\lim_{n\to\infty} \frac{1}{n^{b+1}} |\Pi(\boldsymbol{\omega})| = \sum_{LC^T_{i,u} \in LC^T_{\boldsymbol{\omega}}} \int_0^1 \int_0^1 \int_0^1 \cdots \int_0^1 \mathbf{1}(0 \leq v_0 + LC^T_i(v_S) \leq 1, \ \forall\ i \in S')\ dv_S, \tag{4.3.19}$$

where $dv_S = \prod_{j=0}^{b} dv_{i_j}$ denotes the $(b+1)$-dimensional Lebesgue measure, $v_{i_0} = v_0$ and $LC^T_{\boldsymbol{\omega}}$ is the collection of all the $\prod_{i=1}^{b} \binom{k_i - 1}{\frac{k_i}{2}}$ such different sets of linear combinations corresponding to $\boldsymbol{\omega}$.

As observed previously, choosing $v_{i_j}, 0 \leq j \leq b$ is equivalent to choosing v_0 and $u_{i_j}, 1 \leq j \leq b$. So

$$\lim_{n\to\infty} \frac{1}{n^{b+1}}|\Pi(\boldsymbol{\omega})| = \sum_{LC^T_{i,u} \in LC^T_{\boldsymbol{\omega}}} \int_0^1 \int_{-1}^1 \int_{-1}^1 \cdots \int_{-1}^1 \mathbf{1}(-1 \leq v_0 + LC^T_{i,u}(u_S) \leq 1, \forall\, i \in S')\, du_S, \tag{4.3.20}$$

where $du_S = \prod_{j=0}^{b} u_{i_j}$ denotes the $(b+1)$-dimensional Lebesgue measure on $[0,1]\times[-1,1]^b$ and $u_{i_0} = v_0$.

Suppose a particular set of linear combinations $LC^T_{i,u}$ is given, i.e., for $i \in S'$, $v_i = v_0 + \sum_{m=1}^{i} \alpha_{im} u_{i_m}$, and the values of $\alpha_{im}, 1 \leq i \leq 2k, 1 \leq j \leq b$ are known. Choose

$$C = \max\{|\alpha_{ij}| : 1 \leq j \leq b \text{ and } i \in S'\}.$$

Next we choose ϵ such that $Cb\epsilon < 1/2$. Now, let $|u_{i_j}| < \epsilon$ for $1 \leq j \leq b$ and $Cb\epsilon \leq v_0 \leq 1 - Cb\epsilon$. Then, for all $i \in S'$, $0 \leq v_0 + LC^T_{i,u}(u_S) \leq 1$. Thus,

$$\lim_{n\to\infty} \frac{1}{n^{b+1}}|\Pi(\boldsymbol{\omega})| = \alpha(\boldsymbol{\omega}) > 0 \text{ for any even word } \boldsymbol{\omega},$$

where $\alpha(\boldsymbol{\omega})$ is the sum of the integrals defined in (4.3.20).

This completes the proof of the lemma. $\square$

Lemma 4.3.4. ***(Hankel matrix)*** *Suppose $\boldsymbol{\omega}$ is a word with b distinct letters. Then $\lim_{n\to\infty} \frac{1}{n^{b+1}}|\Pi(\boldsymbol{\omega})| = \alpha(\boldsymbol{\omega}) > 0$ if and only if $\boldsymbol{\omega}$ is a symmetric word. Moreover, for every symmetric word $\boldsymbol{\omega}$, $0 < \alpha(\boldsymbol{\omega}) \leq 1$.*

Proof. Let

$$t_i = \pi(i) + \pi(i-1) \text{ for } 1 \leq i \leq 2k.$$

Clearly, $\boldsymbol{\omega}[i] = \boldsymbol{\omega}[j]$ if and only if $\pi(i-1) + \pi(i) = \pi(j-1) + \pi(j)$, that is, $t_i - t_j = 0$.

Now we fix an $\boldsymbol{\omega}$ with b distinct letters. Suppose $i_1, i_2, \ldots, i_b$ are the positions where new letters made their first appearances. Clearly $t_{i_1} = t_1$. If the first letter again appears at the jth position, then $t_j = t_1$. Similarly, for every i, $1 \le i \le 2k$,

$$t_i = t_{i_j} \quad \text{for some } j \in \{1, 2, \ldots, b\}. \tag{4.3.21}$$

Step 1. (Similar to Step 1 in the proof of Lemma 4.3.1) $\{\pi(i_j), 0 \le j \le b\}$ (where $i_0 = 0$) can be chosen freely if and only if $\{\pi(0), s_{i_j}; 1 \le j \le b\}$ do not satisfy any non-trivial linear relation.

This follows from the same argument as in Step 1 of the proof of Lemma 4.3.1.

Step 2. (Similar to Step 2 in the proof of Lemma 4.3.1) If $\boldsymbol{\omega}$ is not a symmetric word, then, $\lim_{n\to\infty} \frac{1}{n^{b+1}} |\Pi(\boldsymbol{\omega})| = 0$.

If the length of the word $\boldsymbol{\omega}$ is odd, say $2k+1$, then

$$\pi(2k+1) - \pi(0) = (t_1 + t_2 + \cdots + t_{2k+1}) - (t_2 + t_4 + \cdots + t_{2k}) - 2\pi(0) = 0.$$

Hence substituting t_i by t_{i_j}'s using (4.3.3) in the above equation, we see that $\pi(0), t_{i_1}, \ldots, t_{i_b}$ satisfy a non-trivial linear relation. Hence the length of the word cannot be odd.

Suppose the length of the word $\boldsymbol{\omega}$ is $2k$. Then

$$(t_1 + t_3 + \cdots + t_{2k-1}) - (t_2 + t_4 + \cdots + t_{2k}) = \pi(0) \ - \pi(2k) = 0.$$

Hence we have

$$(t_1 + t_3 + \cdots + t_{2k-1}) - (t_2 + t_4 + \cdots + t_{2k}) = 0. \tag{4.3.22}$$

Therefore, from (4.3.21), we see that there exists $\alpha_j \in \mathbb{Z}$ for all $1 \le j \le b$ such that

$$\alpha_1 t_{i_1} + \alpha_2 t_{i_2} + \cdots + \alpha_b t_{i_b} = 0.$$

Now as t_{i_j}'s do not satisfy any non-trivial linear relation, we must have $\alpha_j = 0$ for all $j \in \{1, 2, \ldots, b\}$. Thus, for each j,

$$\big|\{l : l \text{ odd and } t_l = t_{i_j}\}\big| = \big|\{l : l \text{ even and } t_l = t_{i_j}\}\big|.$$

Each letter appears equal number of times at odd and even places. Hence $\boldsymbol{\omega}$ is symmetric.

Therefore if $\boldsymbol{\omega}$ is not symmetric, at least one of the generating vertices has finitely many choices. Hence

$$\lim_{n\to\infty} \frac{1}{n^{b+1}} |\Pi(\boldsymbol{\omega})| = 0 \ \text{ if } \boldsymbol{\omega} \text{ is not symmetric.}$$

Step 3. We now show that, if $\boldsymbol{\omega}$ is a symmetric word with b distinct letters, then $\lim_{n\to\infty} \frac{1}{n^{b+1}} |\Pi(\boldsymbol{\omega})| = \alpha(\boldsymbol{\omega}) > 0$. Note that in case of the reverse circulant matrix (Lemma 4.3.1), for any symmetric word, once the generating vertices were chosen, each of the non-generating vertices had a unique choice. However, that is not the case here. In the following argument we see how each for symmetric word, the contribution is positive but less than or equal to 1.

First we fix the generating vertices $\pi(i_j),\ j = 0, 1, 2, \ldots, b$. Let

$$v_i = \frac{\pi(i)}{n} \text{ for } 0 \le i \le 2k, \quad S = \{\pi(i_j) : 0 \le j \le b\} \ \text{ and } \ S' = \{i : \pi(i) \notin S\}.$$

For $1 \le i \le 2k$, from the link function and the formula for t_i we have

$$v_i = LC_i^H(v_S), \tag{4.3.23}$$

where $LC_i^H(v_S)$ denotes a linear combination of $\{v_i : \pi(i) \in S\}$.

Let $U_n = \{0, 1/n, \ldots, (n-1)/n\}$. From (2.4.5), it is easy to see that for $\boldsymbol{\omega}$ of length $2k$,

$$\big|\Pi(\boldsymbol{\omega})\big| = \big|\big\{(v_0, v_1, \ldots, v_{2k}) : v_i \in U_n \text{ for } 0 \le i \le 2k, v_0 = v_{2k}, v_i = LC_i^H(v_S)\big\}\big|.$$

Transforming $v_i \mapsto y_i = v_i - \frac{1}{2}$, we get that

$$\big|\Pi(\boldsymbol{\omega})\big| = \big|\big\{(y_0, y_1, \ldots, y_{2k}) : y_i \in \{-1/2, -1/2 + 1/n, \ldots, -1/2 + (n-1)/n\} \text{ for } 0 \le i \le 2k,$$
$$y_0 = y_{2k} \text{ and } y_i = LC_i^H(y_S)\big\}\big|.$$

Hence

$$\lim_{n\to\infty} \frac{1}{n^{b+1}} \Pi(\boldsymbol{\omega}) = \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \cdots \int_{-1/2}^{1/2} \mathbf{1}(-1/2 \le LC_i^H(y_S) \le 1/2,\ \forall\, i \in S')\, dy_S, \tag{4.3.24}$$

where $dy_S = \prod_{j=0}^{b} dy_{i_j}$ denotes the $(b+1)$-dimensional Lebesgue measure on $[-\frac{1}{2}, \frac{1}{2}]^{b+1}$.

We now need to show that the above integral is positive. Let $p_i = y_{i-1} + y_i$ and $q_i = y_{i-1} - y_i$. Now we have the following claim.

Claim: For any $1 \leq i \leq 2k$,

$$y_i = \begin{cases} y_0 + \sum_{j=1}^{i} \alpha_{ij} p_{i_j} & \text{if } i \text{ is even,} \\ -y_0 + \sum_{j=1}^{i} \alpha_{ij} p_{i_j} & \text{if } i \text{ is odd.} \end{cases}$$

We prove this by induction. We know that $\pi(1) \in S$. Clearly, $y_1 = p_1 - y_0$. Now either $\pi(2) \in S$ or $2 \in S'$. If $\pi(2) \in S$, then $y_2 = p_2 - y_1$. Therefore $y_2 = p_2 - p_1 + y_0$. If $2 \in S'$, then $p_2 = p_1$ and $y_2 = p_1 - y_1 = y_0$. So the claim is true for $i = 2$.

Now we assume that the claim is true for all $j < i$, and try to prove it for i. Then either $\pi(i) \in S$ or $i \in S'$.

If $\pi(i) \in S$, then $y_i = p_i - y_{i-1}$. If i is even, then $i - 1$ is odd and hence $y_{i-1} = -y_0 + \sum_{j=1}^{i-1} \alpha_{(i-1)j} p_{i_j}$ by induction hypothesis. Therefore $y_i = y_0 + \sum_{j=1}^{i} \alpha_{ij} p_{i_j}$ where $\alpha_{ii} = 1$. The case where i is odd can be tackled similarly.

If $i \in S'$, then there exists m such that $i_m < i$ and $p_i = p_{i_m}$. Then $y_i = p_{i_m} - y_{i-1}$. Now if i is even, then $i-1$ is odd and $y_{i-1} = -y_0 + \sum_{j=1}^{i-1} \alpha_{(i-1)j} p_{i_j}$ by induction hypothesis. Therefore $y_i = y_0 + \sum_{j=1}^{i-1} \alpha_{ij} p_{i_j}$ where $\alpha_{im} = \alpha_{(i-1)m} + 1$. The case where i is odd can be tackled similarly.

Thus the claim is proved.

Now we perform the following change of variables in (4.3.24):

$$(y_0, y_1, y_2, y_3, \ldots, y_{2k}) \longrightarrow (y_0, -y_1, y_2, -y_3, \ldots, y_{2k}) = (z_0, z_1, z_2, z_3, \ldots, z_{2k}) \text{ (say).}$$

Under this transformation,

$$(p_1, p_2, p_3, \ldots, p_{2k}) \longrightarrow (q_1, -q_2, q_3, \ldots, -q_{2k}).$$

Then from the claim it follows that

$$z_i = z_0 + \sum_{j=1}^{i} \beta_{ij} q_{i_j}, \tag{4.3.25}$$

where $\beta_{ij} = \pm\alpha_{ij}$ according as i_j is odd or even. We shall use the notation $z_i = l_{i,q}^H(z_S)$ to denote the linear relation (4.3.25).

Also note that choosing $y_{i_j}, 0 \le j \le b$ is equivalent to choosing $p_{i_j}, 0 \le j \le b$ (where $p_{i_0} = y_0$). Therefore we can write (4.3.24) as

$$\lim_{n\to\infty} \frac{1}{n^{b+1}} \Pi(\boldsymbol{\omega}) = \int_{-1/2}^{1/2} \int_{-1}^{1} \int_{-1}^{1} \cdots \int_{-1}^{1} \mathbf{1}(-1/2 \le LC_{i,q}^H(z_S) \le 1/2,\ \forall\ i \in S')\ dq_S, \tag{4.3.26}$$

where $dq_S = \prod_{j=0}^{b} dq_{i_j}$ is the $(b+1)$-dimensional Lebesgue measure on $[-\frac{1}{2}, \frac{1}{2}] \times [-1,1]^b$.

Let

$$C = \max\{|\alpha_{ij}| : 1 \le j \le b \text{ and } i \in S'\}.$$

Next we choose ϵ such that $Cb\epsilon < 1/4$. Now, let $|q_{i_j}| < \epsilon$ for $1 \le j \le b$ and $Cb\epsilon - \frac{1}{2} \le z_0 \le \frac{1}{2} - Cb\epsilon$. Then, for all $i \in S'$, $-\frac{1}{2} \le LC_{i,q}^T(z_S) \le \frac{1}{2}$. Also the circuit condition is automatically satisfied. So the integrand in the rhs of (4.3.26) is 1.

Hence

$$\lim_{n\to\infty} \frac{1}{n^{b+1}} |\Pi(\boldsymbol{\omega})| = \alpha(\boldsymbol{\omega}) > 0,$$

where $\alpha(\boldsymbol{\omega})$ is the value of the integral in (4.3.24). Also as $\alpha(\boldsymbol{\omega})$ actually gives the Lebesgue measure of a set in $[-1/2, 1/2]^{b+1}$ as seen in (4.3.24), we have $\alpha(\boldsymbol{\omega}) \le 1$.

This completes the proof of the lemma. □

Now we are ready to prove Theorems 4.2.2-4.2.4.

4.3.2 Proof of Theorem 4.2.2 (Reverse circulant)

Proof. We separate the proof of the theorem into four steps.

Step 1 (Reduction to the case where the entries of Z_n have mean 0): Consider the matrix $\widetilde{Z}_n$ whose entries are $(y_i - \mathbb{E}y_i)$. The entries of $\widetilde{Z}_n$ have mean 0. Now

$$n\ \mathbb{E}[(y_i - \mathbb{E}y_i)^{2k}] = n\mathbb{E}[y_i^{2k}] + n\sum_{j=0}^{2k-1}\binom{2k}{j}\mathbb{E}[y_i^j](\mathbb{E}y_i)^{2k-j}. \tag{4.3.27}$$

The first term of the r.h.s. is equal to $\tilde{g}_{2k,n}(i/n)$ by (4.2.1). For the second term we argue as follows:

$$\begin{aligned}\text{For } j \neq 2k-1,\quad n\ \mathbb{E}[y_i^j](\mathbb{E}y_i)^{2k-j} &= (n^{\frac{1}{2k-j}}\mathbb{E}y_i)^{2k-j}\mathbb{E}[y_i^j]\\ &\overset{n\to\infty}{\longrightarrow} 0, \qquad \text{by condition (4.2.2)}.\end{aligned}$$

$$\begin{aligned}\text{For } j = 2k-1,\quad n\ \mathbb{E}[y_i^{2k-1}]\mathbb{E}y_i &= (\sqrt{n}\mathbb{E}[y_i^{2k-1}])(\sqrt{n}\mathbb{E}y_i)\\ &\overset{n\to\infty}{\longrightarrow} 0, \qquad \text{by condition (4.2.2)}.\end{aligned}$$

Hence from (4.3.27), we see that Condition (4.2.1) is true for the matrix $\widetilde{Z}_n$ with a modified sequence $\tilde{\tilde{g}}_{2k,n}$ that still converges uniformly to $\tilde{g}_{2k}$. However, for ease of notation, we will continue to call this sequence of functions as $\tilde{g}_{2k,n}$. Similarly we can show that (4.2.2) is true for $\widetilde{Z}_n$. Hence Assumption A holds for the matrix $\widetilde{Z}_n$.

Now observe that using Lemma 2.3.1,

$$\begin{aligned}d_2^2(\mathbb{E}F^{Z_n}, \mathbb{E}F^{\widetilde{Z}_n}) &\leq \frac{1}{n}\sum_i n(\mathbb{E}y_i)^2\\ &\leq n\ (\sup_i \mathbb{E}y_i)^2\\ &= (\sup_i \sqrt{n}\ \mathbb{E}y_i)^2 \overset{n\to\infty}{\longrightarrow} 0, \qquad \text{by condition (4.2.2)}.\end{aligned}$$

Hence the limit of EESD of Z_n and $\widetilde{Z}_n$ are same. Hence we can assume the entries of Z_n have mean zero.

Now we prove Part (a) of the theorem by verifying the first moment condition and Carleman's condition of Lemma 2.1.3.

Step 2 (Verification of the first moment condition): From (2.4.4), using the fact that $\mathbb{E}(y_i) = 0$, we have

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathrm{Tr}(Z_n)^k] = \lim_{n\to\infty} \frac{1}{n} \sum_{\pi:\ell(\pi)=k} \mathbb{E}[Y_\pi] = \lim_{n\to\infty} \sum_{b=1}^{k} \sum_{\substack{\omega \text{ matched} \\ \text{with b distinct letters}}} \frac{1}{n} \sum_{\pi\in\Pi(\boldsymbol{\omega})} \mathbb{E}(Y_\pi). \tag{4.3.28}$$

Suppose $\boldsymbol{\omega}$ is a word with b distinct letters where each letter appears $k_1, k_2, \ldots, k_b$ times. Let the jth distinct letter appear for the first time at $i_j, 1 \le j \le b$. Write $(\pi(i_j-1), \pi(i_j))$ as (m_j, l_j). Clearly, $m_1 = \pi(0)$ and $l_1 = \pi(1)$. Let S be the set of distinct generating vertices of $\boldsymbol{\omega}$. Recall the set $E_b(k)$ of all even words with b distinct letter, see Section 2.5.3. Suppose $\boldsymbol{\omega} \in E_b(k)$. Then the contribution of this $\boldsymbol{\omega}$ in (4.3.28) is

$$\frac{1}{n^{b+1}} \sum_{S} \prod_{j=1}^{b} \tilde{g}_{k_j,n}\left(\frac{m_j + l_j - 2 \pmod n}{n}\right). \tag{4.3.29}$$

Now from Step 4 of Lemma 4.3.1, observe that for $j \neq 1$, m_j can be written as a linear combination of $\{l_i; 1 \le i \le j-1\}$ and m_1. By abuse of notation, let m_1 and $l_j, 1 \le j \le b$ denote the indices of the generating vertices. Then, as $n \to \infty$, the above sum goes to

$$\int_{[0,1]^{b+1}} \prod_{j=1}^{b} \Big[\tilde{g}_{k_j}(x_{m_j} + x_{l_j})\mathbf{1}(0 \le x_{m_j} + x_{l_j} \le 1) + \tilde{g}_{k_j}(x_{m_j} + x_{l_j} - 1)\mathbf{1}(x_{m_j} + x_{l_j} > 1)\Big]\, dx_S, \tag{4.3.30}$$

where $dx_S = dx_{m_1} dx_{l_1} \cdots dx_{l_b}$ is the $(b+1)$-dimensional Lebesgue measure on $[0,1]^{(b+1)}$.

By Lemma 4.3.1, it follows that the above integral is over all of $[0,1]^{(b+1)}$ if and only if $\boldsymbol{\omega} \in S_b(k)$. That is, if $\boldsymbol{\omega} \notin S_b(k)$, then the indicator functions in (4.3.30) is non-zero only on a surface whose dimension is at most b. This is because in that case, x_{m_1} and the x_{l_j}'s satisfy a linear equation (see proof of Lemma 4.3.1).

As a result, the contribution of $\boldsymbol{\omega}$ as described in (4.3.30) is equal to 0 if $\boldsymbol{\omega} \in E_b(k) \setminus S_b(k)$. If $\boldsymbol{\omega} \in S_b(k)$, then the contribution of $\boldsymbol{\omega}$ is

$$\int_0^1 \int_0^1 \cdots \int_0^1 \prod_{j=1}^{b} h_{k_j}(x_{m_j}, x_{l_j})\, dx_S, \tag{4.3.31}$$

where $h_{k_j}(x_{m_j}, x_{l_j}) = \tilde{g}_{k_j}(x_{m_j} + x_{l_j})\mathbf{1}(0 \le x_{m_j} + x_{l_j} \le 1) + \tilde{g}_{k_j}(x_{m_j} + x_{l_j} - 1)\mathbf{1}(x_{m_j} + x_{l_j} > 1)$.

Now for any $m \in \mathbb{N}$, let

$$f_{2m}(x) = \int_0^1 h_{2m}(x,y)\,dy = \int_0^1 \tilde{g}_{2m}(x{+}y)\mathbf{1}(0 \le x{+}y \le 1) + \tilde{g}_{2m}(x{+}y{-}1)\mathbf{1}(x{+}y > 1)\,dy.$$

Note that

$$f_{2m}(x) = \int_x^1 \tilde{g}_{2m}(t)\ dt\ + \int_0^x \tilde{g}_{2m}(t)\ dt = \int_0^1 \tilde{g}_{2m}(t)\ dt = c_{2m}, \tag{4.3.32}$$

is independent of x. Then, using (4.3.32), (4.3.31) can be written as

$$\int_0^1\int_0^1 \cdots \int_0^1 \prod_{j=1}^{b-1} h_{k_j}(x_{m_j}, x_{l_j}) c_{s_b}\ dx_{m_1} dx_{l_1} \cdots dx_{l_{b-1}}. \tag{4.3.33}$$

Proceeding in this manner the contribution of $\boldsymbol{\omega}$ from (4.3.31) can be obtained as follows:

$$\int_0^1 \prod_{j=1}^{b} c_{k_j}\ dx_0 = \prod_{j=1}^{b} c_{k_j}. \tag{4.3.34}$$

Now suppose $\boldsymbol{\omega} \notin E(2k)$. Suppose $\boldsymbol{\omega}$ contains b_1 distinct letters that appear an even number of times, and b_2 distinct letters that appear an odd number of times, and $b = b_1 + b_2$. So we assume that for each $\pi \in \Pi(\boldsymbol{\omega})$, k_{j_p}, $1 \le p \le b_1$ are even and k_{j_q}, $b_1 + 1 \le q \le b_1 + b_2$ are odd. Hence the contribution of this $\boldsymbol{\omega}$ to (4.3.28) is as follows:

$$\begin{aligned}
&\frac{1}{n} n^{-b_1} n^{-(b_2 - \frac{1}{2})} \sum_S \prod_{p=1}^{b_1} h_{k_{j_p}}(x_{m_{j_p}}, x_{l_{j_p}}) \prod_{q=b_1+1}^{b_1+b_2} n^{\frac{b_2-1/2}{b_2}} \mathbb{E}\Big[y^{k_{j_q}}_{(t_{i_q}-2)(\mathrm{mod}\ n)}\Big] \\
&= \frac{1}{n^{b_1+b_2+\frac{1}{2}}} \sum_S \prod_{p=1}^{b_1} h_{k_{j_p}}(x_{m_{j_p}}, x_{l_{j_p}}) \prod_{q=b_1+1}^{b_1+b_2} n^{\frac{b_2-1/2}{b_2}} \mathbb{E}\Big[y^{k_{j_q}}_{(t_{i_q}-2)\ (\mathrm{mod}\ n)}\Big].
\end{aligned} \tag{4.3.35}$$

For n large, $n^{\frac{b_2-1/2}{b_2}} \mathbb{E}[y^{k_{j_q}}_{(t_{i_q}-2)\ (\mathrm{mod}\ n)}] < 1$ for any $b_1 + 1 \le q \le b_1 + b_2$ and $\prod_{p=1}^{b_1} h_{k_{j_p}}(x_{m_{j_p}}, x_{l_{j_p}}) \le M$ (independent of n). Now as $\boldsymbol{\omega} \notin S_b(k)$, from Lemma 4.3.1 we have, $|S| \le b$. Thus any word that is not even, contributes 0 as $n \to \infty$.

For any partition $\sigma \in S_b(2k)$, let $\{V_1, \ldots, V_b\}$ be its partition blocks. Then from (4.3.28) and (4.3.34), we have

$$\lim_{n\to\infty} \frac{1}{n} \mathbb{E}[\mathrm{Tr}(Z_n)^{2k}] = \sum_{b=1}^{k} \sum_{\sigma \in S_b(2k)} \prod_{i=1}^{b} c_{|V_i|} = \sum_{\sigma \in S(2k)} c_\sigma. \tag{4.3.36}$$

We also note that $\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathrm{Tr}(Z_n)^{2k+1}] = 0$ for any $k \geq 0$. This proves the first moment condition.

Step 3 (Uniqueness of the LSD): Here we show that $\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathrm{Tr}(Z_n)^{2k}] = \gamma_{2k}$ determines a unique distribution. Note that $\{\gamma_{2k}\}_{k\geq 1}$, being the limit of a moment sequence, is a moment sequence. Moreover,

$$\gamma_{2k} = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathrm{Tr}(Z_n)^{2k}] \leq \sum_{\sigma \in S(2k)} M_\sigma \leq \sum_{\sigma \in \mathcal{P}(2k)} M_\sigma = \alpha_{2k}.$$

As $\{\alpha_{2k}\}$ satisfies Carleman's condition, $\{\gamma_{2k}\}$ also does so. Hence the sequence of moments $\{\gamma_{2k}\}$ determines a unique distribution.

Therefore, there exists a measure ν_R with moment sequence $\{\gamma_{2k}\}$ such that EESD of Z_n converges weakly to ν_R. Further, we see from (4.3.36), (2.5.17) and (2.5.18) that the half cumulants (see Section 2.5.3) of ν_R are $\{c_{2n}\}_{n\geq 1}$ which are defined in (4.3.32).

This completes the proof of Part (a).

Step 4 (Proof of Part (b)): Observe that from Lemma 2.3.3,

$$d_2^2(\mathbb{E}F^{R_n^{(s)}}, \mathbb{E}F^{Z_n}) \leq \frac{1}{n}\mathbb{E}[\mathrm{Tr}(R_n^{(s)} - Z_n)^2] = \frac{1}{n}\sum_i n\mathbb{E}[x_i^2[\mathbf{1}_{[|x_i|>r_n]}], \tag{4.3.37}$$

where the last equality follows as each x_i occurs n times in $R_n^{(s)}$.

Now if $\{r_n\}$ also satisfies condition (3.3.4), then using (4.3.37) and Part (a) we can say that the EESD of $R_n^{(s)}$ converges to ν_R. This proves Part (b). □

4.3.3 Proof of Theorem 4.2.3 (Symmetric Circulant)

Proof. **Step 1. (Reduction to mean zero)**: First of all note that the entries of the matrix Z_n can be assumed to have mean zero. This reduction follows from Step 1 in the proof of Theorem 4.2.2.

Step 2. (Verfication of the first moment condition: From (2.4.4), using the fact that $\mathbb{E}(y_i) = 0$, we have

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}\big[\operatorname{Tr}(Z_n)^k\big] = \lim_{n\to\infty} \frac{1}{n} \sum_{\pi:\ell(\pi)=k} \mathbb{E}[Y_\pi] = \lim_{n\to\infty} \sum_{b=1}^{k} \sum_{\substack{\omega \text{ matched} \\ \text{with b distinct letters}}} \frac{1}{n} \sum_{\pi\in\Pi(\boldsymbol{\omega})} \mathbb{E}[Y_\pi]. \tag{4.3.38}$$

Suppose $\boldsymbol{\omega}$ is a word with b distinct letters where each letter appears $k_1, k_2, \ldots, k_b$ times. Suppose the jth distinct letter appears at the $(\pi(i_j-1), \pi(i_j))$th position for the first time. Denote $(\pi(i_j-1), \pi(i_j))$ as (m_j, l_j). Let $u_i = s_i/n$ be as defined in Lemma 4.3.2, and $U_n = \{0, 1/n, 2/n, \ldots, (n-1)/n\}$.

Let $\boldsymbol{\omega} \in E_b(k)$ where the distinct letters appear $k_1, k_2, \ldots, k_b$ times. Clearly, as observed in Lemma 4.3.2, there are $\prod_{i=1}^{b} \binom{k_i-1}{\frac{k_i}{2}}$ equations for determining the non-generating vertices, once the generating vertices are chosen. For each of the combination of equations, we get the same contribution to the limit due to the structure of the link function. Then, with S as the set of distinct generating vertices of $\boldsymbol{\omega}$, the contribution of each such combination of equations for the word $\boldsymbol{\omega}$ in (4.3.28) is

$$\frac{1}{n^{b+1}} \sum_{S} \prod_{j=1}^{b} \tilde{g}_{k_j,n}\Big(\frac{1}{2} - |\frac{1}{2} - |u_{i_j}||\Big). \tag{4.3.39}$$

Now from Step 4 of Lemma 4.3.2, observe that for each set of linear combinations, whenever $j \neq 1$, m_j can be written as a linear combination of $\{l_i; 1 \leq i \leq j-1\}$ and m_1. By abuse of notation, let m_1 and $l_j, 1 \leq j \leq b$ denote the indices of the generating vertices. Then, as $n \to \infty$, the above sum goes to

$$\int_0^1 \int_0^1 \cdots \int_0^1 \prod_{j=1}^{b} \tilde{g}_{k_j}\Big(\frac{1}{2} - |\frac{1}{2} - |x_{m_j} - x_{l_j}||\Big)\, dx_S, \tag{4.3.40}$$

where $dx_S = dx_{m_1} dx_{l_1} \cdots dx_{l_b}$ is the $(b+1)$-dimensional Lebesgue measure.

Now suppose $\boldsymbol{\omega} \in E_b(k)$. Then for any $m \in \mathbb{N}$,

$$\begin{aligned} f_{2m}(x) &:= \int_0^1 \tilde{g}_{2m}\Big(\frac{1}{2} - |\frac{1}{2} - |x-y||\Big)\, dy \\ &= \int_0^{\frac{1}{2}} \tilde{g}_{2m}(t)\, dt + \int_0^{\frac{1}{2}} \tilde{g}_{2m}(t)\, dt = 2\int_0^{\frac{1}{2}} \tilde{g}_{2m}(t)\, dt = c_{2m}. \end{aligned} \tag{4.3.41}$$

So $f_{2m}(x)$ is independent of x. Then, using (4.3.41), (4.3.40) can be written as

$$\int_0^1 \int_0^1 \cdots \int_0^1 \prod_{j=1}^{b-1} \tilde{g}_{k_j}\Big(\frac{1}{2} - |\frac{1}{2} - |x_{m_j} - x_{l_j}||\Big) c_{s_b}\, dx_{m_1} dx_{l_1} \cdots dx_{l_{b-1}}. \tag{4.3.42}$$

Proceeding in this manner, for $\boldsymbol{\omega} \in E_b(k)$, (4.3.40) can be written as

$$\int_0^1 \prod_{j=1}^{b} c_{k_j}\, dx_0 = \prod_{j=1}^{b} c_{k_j}.$$

As there are $\prod_{i=1}^{b} \binom{k_i - 1}{\frac{k_i}{2}}$ sets of equations that contribute identically, the total contribution for the word $\boldsymbol{\omega} \in E_b(k)$ is

$$\prod_{j=1}^{b} \binom{k_j - 1}{\frac{k_j}{2}} c_{k_j}. \tag{4.3.43}$$

Now suppose $\boldsymbol{\omega} \notin E(2k)$. Suppose $\boldsymbol{\omega}$ contains b_1 distinct letters that appear an even number of times, and b_2 distinct letters that appear an odd number of times and $b = b_1 + b_2$. Using a similar argument as in Step 2 of the proof of Theorem 4.2.2, it is easy to see that the contribution of this $\boldsymbol{\omega}$ is 0 (see (4.3.35)).

Therefore

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathrm{Tr}(Z_n)^{2k+1}] = 0 \quad \text{for any} \quad k \geq 0. \tag{4.3.44}$$

For any partition $\sigma \in E_b(2k)$, let $\{V_1, \ldots, V_b\}$ be its blocks. Then from (4.3.28) and (4.3.43), we have

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathrm{Tr}(Z_n)^{2k}] = \sum_{b=1}^{k} \sum_{\sigma\in E_b(2k)} \prod_{j=1}^{b} \frac{1}{2}\binom{|V_j|}{\frac{|V_j|}{2}} c_{|V_j|} = \sum_{\sigma\in E(2k)} a_\sigma c_\sigma, \tag{4.3.45}$$

where $a_{2n} = \frac{1}{2}\binom{2n}{n}$, and a_σ and c_σ are the multiplicative extensions of the sequence $\{a_{2n}\}$ and $\{c_{2n}\}$ respectively. (4.3.44) and (4.3.45) establishes the first moment condition.

Step 3. (Uniqueness of the LSD): Note that $\{\gamma_{2k}\}_{k\geq 1}$, being the limit of a moment sequence, is a moment sequence. As $a_{2k} \leq 2^{2k}$, we have

$$\gamma_{2k} = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathrm{Tr}(Z_n)^{2k}] \leq \sum_{\sigma\in E(2k)} a_\sigma M_\sigma \leq \sum_{\sigma\in \mathcal{P}(2k)} 2^{2k} M_\sigma = 2^{2k}\alpha_{2k}.$$

Since $\{\alpha_{2k}\}$ satisfies Carleman's condition, $\{\gamma_{2k}\}$ also does so. Hence there is a unique symmetric distribution ν_C with even moments $\{\gamma_{2k}\}_{k\geq 1}$, and the EESD of Z_n converges weakly to ν_C. Also, the odd cumulants of ν_C are 0, and the even cumulants are $\{a_{2n}c_{2n}\}$ as in (4.3.41), see Section 2.5.1. This completes the proof of Part (a).

Step 4. (Proof of Part (b)): Observe that by Lemma 2.3.3,

$$\begin{aligned}
&d_2^2(\mathbb{E}F^{C_n^{(s)}}, \mathbb{E}F^{Z_n})\\
&\leq \frac{1}{n}\mathbb{E}[\mathrm{Tr}(C_n^{(s)} - Z_n)^2]\\
&= \begin{cases} \frac{1}{n}\sum_{i\neq 0} 2n\mathbb{E}[x_i^2 \mathbf{1}_{[|x_i|>r_n]}] + \frac{1}{n} n\mathbb{E}[x_0^2\mathbf{1}_{[|x_i|>r_n]}] & \text{for } n \text{ odd},\\ \frac{1}{n}\sum_{i\neq 0,\frac{n}{2}} 2n\mathbb{E}[x_i^2 \mathbf{1}_{[|x_i|>r_n]}] + \frac{1}{n} n\mathbb{E}[x_0^2\mathbf{1}_{[|x_i|>r_n]}] + \frac{1}{n} n\mathbb{E}[x_{\frac{n}{2}}^2\mathbf{1}_{[|x_i|>r_n]}] & \text{for } n \text{ even}.\end{cases}
\end{aligned} \tag{4.3.46}$$

If $\{r_n\}$ also satisfies Condition (4.2.4), then we see that the rhs of (4.3.46) goes to 0. Hence from (a) we have that the EESD of $C_n^{(s)}$ converges to ν_C. This proves Part (b).

This completes the proof of Theorem 4.2.3. □

Remark 4.3.5. *The sequence $\{a_{2n}\}_{n\geq 1}$ defined in Theorem 4.2.3 is the (even) moment sequence of a unique probability distribution. So, there exists a random variable Z such that for every $n \geq 1$, $\mathbb{E}[Z^{2n}] = a_{2n}$ and $\mathbb{E}[Z^{2n-1}] = 0$.*

Let X be a random variable with density $f(x) = \frac{1}{\pi\sqrt{4-x^2}}$ on $[-2, 2]$. Then $\mathbb{E}[X^k] = \binom{k}{k/2}$. Now let Y be the random variable which takes values 1 and 0 with probability $\frac{1}{2}$ each. Suppose Y is independent of X. It is easy to see that if $Z = XY$, then $\mathbb{E}[Z^{2k}] = a_{2k}$.

4.3.4 Proof of Theorem 4.2.4 (Toeplitz and Hankel)

Proof. We first prove the theorem for the Toeplitz matrix and then for the Hankel matrix.

Toeplitz matrix:

Step 1. (Reduction to mean zero): First of all note that the entries of the Toeplitz matrix Z_n can be assumed to have mean zero. This reduction follows from Step 1 in the proof of Theorem 4.2.2.

Step 2. (Verification of the first moment condition): From (2.4.4) and using the fact that $\mathbb{E}(y_i) = 0$, we have

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathrm{Tr}(Z_n)^k] = \lim_{n\to\infty} \frac{1}{n} \sum_{\pi:\ell(\pi)=k} \mathbb{E}[Y_\pi] = \lim_{n\to\infty} \sum_{b=1}^{k} \sum_{\substack{\omega \text{ matched} \\ \text{with b distinct letters}}} \frac{1}{n} \sum_{\pi\in\Pi(\boldsymbol{\omega})} \mathbb{E}(Y_\pi). \tag{4.3.47}$$

Now suppose $\boldsymbol{\omega}$ is a word with b distinct letters, where each letter appears $k_1, k_2, \ldots, k_b$ times. Suppose the jth distinct letter appears at $(\pi(i_j-1), \pi(i_j))$th position for the first time. Denote $(\pi(i_j-1), \pi(i_j))$ as (m_j, l_j). Let $v_i = \pi(i)/n$ be as defined in Lemma 4.3.3, and $U_n = \{0, 1/n, 2/n, \ldots, (n-1)/n\}$.

Let $\boldsymbol{\omega} \in E_b(k)$. Clearly, as observed in Lemma 4.3.3, there are $\prod_{i=1}^{b}\binom{k_i-1}{\frac{k_i}{2}}$ combinations of equations for the s_j's (and hence v_j's) for determining the non-generating vertices, once the generating vertices have been chosen. Let us denote a generic combination of the v_j's by $LC_i^T(v_S)$ (see (4.3.19)) and the collection of all such sets of linear combinations coreesponding to the word $\boldsymbol{\omega}$, $LC_{\boldsymbol{\omega}}^T$. For each of the combinations of equations, we get a positive (possibly different) contribution (see Lemma 4.3.3). The contribution of each combination $LC_i^T(v_S) \in LC_{\boldsymbol{\omega}}^T$ corresponding to the word $\boldsymbol{\omega}$ in (4.3.47) is

$$\frac{1}{n^{b+1}} \sum_{S} \prod_{j=1}^{b} \tilde{g}_{k_j,n}\big(|v_{m_j} - v_{l_j}|\big)\mathbf{1}(0 \le v_0 + l_i^T(v_S) \le 1,\ \forall\, i \in S'), \tag{4.3.48}$$

where S is the set of distinct generating vertices and S' is the set of indices of the non-generating vertices of $\boldsymbol{\omega}$. By abuse of notation, let m_1 and $l_j, 1 \le j \le b$ denote the indices of the generating vertices. Therefore, as $n \to \infty$, the contribution of $\boldsymbol{\omega}$ in (4.3.47) is given by

$$\sum_{LC_{\boldsymbol{\omega}}^T} \int_{[0,1]^{b+1}} \prod_{j=1}^{b} \tilde{g}_{k_j}\big(|x_{m_j} - x_{l_j}|\big)\mathbf{1}(0 \le x_0 + LC_i^T(x_S) \le 1,\ \forall\, i \in S')\, dx_S, \tag{4.3.49}$$

where $dx_S = dx_{m_1}dx_{l_1}\cdots dx_{l_b}$ denotes the $(b+1)$-dimensional Lebesgue measure on $[0,1]^{b+1}$. As $\boldsymbol{\omega}$ is an even word, from Lemma 4.3.3, it follows that for every set of linear combination $LC_i^T \in LC_{\boldsymbol{\omega}}^T$, there is a set of positive Lebesgue measure in $[0,1]^{b+1}$ where the indicator function in the above integral takes the value 1.

If $\boldsymbol{\omega}$ is not an even word, then following a similar argument as given in Step 2 of the proof of Theorem 4.2.2, it can be shown that its contribution is zero in the limit. As an even word $\boldsymbol{\omega}$ with b distinct letters can be identified with an even partition, say $\sigma \in E_b(k)$, for $k \geq 0$,

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathrm{Tr}(Z_n)^{2k+1}] = 0,$$

and

$$\begin{aligned}\gamma_{2k} &= \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\mathrm{Tr}(Z_n)^{2k}] \\ &= \sum_{b=1}^{k} \sum_{\sigma\in E_b(2k)} \sum_{LC_i^T\in LC_\sigma^T} \int_{[0,1]^{b+1}} \prod_{j=1}^{b} \tilde{g}_{k_j}\left(|x_{m_j}-x_{l_j}|\right)\mathbf{1}(0\leq x_0 + LC_i^T(x_S)\leq 1,\ \forall\ i\in S')\ dx_S.\end{aligned} \tag{4.3.50}$$

This establishes the first moment condition.

Step 3. (Uniqueness of the LSD): Note that $\{\gamma_k\}_{k\geq 1}$, being a limit of a moment sequence is a moment sequence. This limiting moment sequence is dominated by the moment sequence of the Symmetric Circulant case (see (4.3.45)). As the latter satisfies the Carleman's condition, so does γ_k.

Therefore there exists a symmetric distribution ν_T with moment sequence $\{\gamma_{2k}\}$ such that EESD of Z_n converges weakly to ν_T. This completes the proof of Part (a).

Proof of Part (b) is trivial, so we skip it. This ends the proof for the Toeplitz matrix.

Hankel matrix:

Step 1. (Reduction to mean zero): First of all note that the entries of the matrix can be assumed to have mean zero due to Step 1 in the proof of Theorem 4.2.2.

Step 2. (Verification of the first moment condition): From (2.4.4) and using the fact that $\mathbb{E}(y_i) = 0$, we have

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}\big[\operatorname{Tr}(Z_n)^k\big] = \lim_{n\to\infty} \frac{1}{n} \sum_{\pi:\ell(\pi)=k} \mathbb{E}[Y_\pi] = \lim_{n\to\infty} \sum_{b=1}^{k} \sum_{\substack{\omega \text{ matched} \\ \text{with b distinct letters}}} \frac{1}{n} \sum_{\pi\in\Pi(\boldsymbol{\omega})} \mathbb{E}(Y_\pi). \tag{4.3.51}$$

Now suppose $\boldsymbol{\omega}$ is a symmetric word with b distinct letters. Let $v_i = \pi(i)/n$ and $U_n = \{0, 1/n, 2/n, \ldots, (n-1)/n\}$ as defined in Lemma 4.3.4. We know that the v_i's satisfy a linear relation given in (4.3.23). Then the contribution of $\boldsymbol{\omega}$ in (4.3.51) is

$$\frac{1}{n^{b+1}} \sum_{S} \prod_{j=1}^{b} \tilde{g}_{k_j,n}(v_{m_j} + v_{l_j}) \mathbf{1}(0 \le LC_i^H(v_S) \le 1,\ \forall\ i \in S'), \tag{4.3.52}$$

where S is the set of distinct generating vertices, and S' is the set of indices of the non-generating vertices of $\boldsymbol{\omega}$.

By abuse of notation, let m_1 and $l_j, 1 \le j \le b$ denote the indices of the generating vertices. Therefore as $n \to \infty$, for each symmetric word $\boldsymbol{\omega}$, the contribution to the limit in (4.3.51) is given by (see Lemma 4.3.4)

$$\int_{[0,1]^{b+1}} \prod_{j=1}^{b} \tilde{g}_{k_j}(x_{m_j} + x_{l_j}) \mathbf{1}(0 \le LC_i^H(x_S) \le 1,\ \forall\ i \in S')\ dx_S, \tag{4.3.53}$$

where $dx_S = dx_{m_1} dx_{l_1} \cdots dx_{l_b}$ is the $(b+1)$-dimensional Lebesgue measure. As $\boldsymbol{\omega}$ is a symmetric word, from Lemma 4.3.4, it follows that, there is a set of positive Lebesgue measure in $[0,1]^{b+1}$ where the indicator function in the above integral takes the value 1.

If $\boldsymbol{\omega}$ is not a symmetric word, then following a very similar argument as given in the proof of Theorem 4.2.2, it can be shown that its contribution is zero in the limit. Therefore for $k \ge 0$, $\lim\limits_{n\to\infty} \frac{1}{n}\mathbb{E}[\operatorname{Tr}(Z_n)^{2k+1}] = 0$ and

$$\begin{aligned}
\gamma_{2k} &= \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\operatorname{Tr}(Z_n)^{2k}] \\
&= \sum_{b=1}^{k} \sum_{\sigma\in S_b(2k)} \int_{[0,1]^{b+1}} \prod_{j=1}^{b} \tilde{g}_{k_j}(x_{m_j} + x_{l_j}) \mathbf{1}(0 \le LC_i^H(x_S) \le 1,\ \forall\ i \in S')\ dx_S.
\end{aligned} \tag{4.3.54}$$

This establishes the first moment condition in this case.

Step 3. (Uniqueness of the LSD): Note that $\{\gamma_k\}_{k\geq 1}$, being the limit of a moment sequence is a moment sequence and γ_k is dominated by the moment sequence in the Reverse Circulant case (see (4.3.36)). As the latter satisfies the Carleman's condition, so does γ_k.

Hence there exists a symmetric distribution ν_H with (even) moment sequence $\{\gamma_k\}$ such that EESD of Z_n converges weakly to ν_H. This completes the proof of Part (a).

Proof of Part (b) is trivial, so we skip it. □

4.4 Some Corollaries

In this section, we present a few corollaries that follow from Theorems 4.2.2- 4.2.4. In particular we deduce Results 4.1.1- 4.1.8 from Theorems 4.2.2-4.2.4. We also discuss some other models that can be handled using these theorems.

Almost sure convergence in some special cases: Note that Theorems 4.2.2-4.2.4 conclude the convergence of the EESD of the respective matrices. In general, the almost sure convergences do not hold in Theorems 4.2.2–4.2.4 as the matrices are more structured and are constituted of $\mathcal{O}(n)$ independent random variables. As a result, unlike the Wigner case, (3.4.5) is not true in general.

However, many of the previous results, viz, Results 4.1.1-4.1.3, 4.1.5-4.1.8, conclude the almost sure convergence of the ESD. Here we will present a unified lemma that will help us tackle the almost sure convergence in these special cases. As in Chapter 3 we shall use Lemma 2.1.3 to verify the almost sure convergence of the ESD wherever it holds. As seen in Chapter 3, to verify Condition (ii) in Lemma 2.1.3 for the patterned matrices, we shall need an upper bound of the following set (note that in this case the circuits π_i correspond to the respective link functions of $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$).

$$\begin{aligned}\tilde{Q}^b_{k,4} = \{(\pi_1,\pi_2,\pi_3,\pi_4) : \ell(\pi_i) = k;\ \pi_i, 1\leq i\leq 4 \text{ jointly- and cross-matched with}\\ b \text{ distinct letters}\}.\end{aligned}$$

Lemma 4.4.1. *There is a constant $M > 0$ such that*

$$|\tilde{Q}_{k,4}^{b}| \leq Mn^{2k+2} \quad \text{for any } 1 \leq b \leq 2k. \tag{4.4.1}$$

Proof. (Borrowed from the proof of Lemma 1.4.3 (a) in Bose [2018]) Note that the total length of the circuits is $4k$. As the circuits are cross-matched, the number of distinct letters across the four circuits can be at most $2k$. We divide the proof into three cases—(i) when $b \leq 2k-2$, (ii) when $b = 2k-1$, and (iii) when $b = 2k$.

Case 1. $b \leq 2k-2$.

First observe that for any circuit π, if we set aside the first vertex $\pi(0)$, then the number of choices for the generating vertices is at most n^b, where b is the number of distinct letters in π. Moreover once all the generating vertices have been chosen, the number of choices for the non-generating vertices is at most finitely many. This observation will be used repeatedly.

Consider all circuits $(\pi_1, \pi_2, \pi_3, \pi_4)$ of length k, which are jointly-matched and cross-matched, with b distinct letters. Let the number of new distinct letters appearing in π_i be k_i, $i \in \{1,2,3,4\}$. So clearly, $k_1 + k_2 + k_3 + k_4 = b$. Then the generating vertices of each $\pi_i, i = 1,2,3,4$ can be chosen freely in $\mathcal{O}(n^{k_i+1})$ ways. Hence,

$$|\tilde{Q}_{k,4}^{b}| \leq Mn^{b+4} \leq Mn^{2k+2}.$$

Case 2. $b = 2k-1$.

In this case there are two possibilities: (a) one of the letters is repeated four times across $\pi_1, \pi_2, \pi_3, \pi_4$ and all other letters appear exactly two times across the four circuits; (b) two of the letters appear thrice each and all other $(2k-3)$ letters appear exactly twice across $\pi_1, \pi_2, \pi_3, \pi_4$.

Observe that in any case there is one circuit, say π_1 and one letter say, x such that x does not re-occur in π_1. Suppose that x appears in π_1 at the ith position. That is, the L-value corresponding to x is $L(\pi_1(i-1), \pi_1(i))$. We leave aside this letter (L-value) and count the number of ways the rest of the generating vertices can be chosen. Now since there are $(2k-2)$ distinct letters (L-values) just as in Case 1, the number of ways $\pi_i(0), 1 \leq i \leq 4$ and these generating vertices can be chosen is $\mathcal{O}(n^{4+2k-2}) = \mathcal{O}(n^{2k+2})$.

It is to be noted here that the generating vertex $\pi_1(i)$ for the letter x has not been chosen yet. We will show that $\pi_1(i)$ has finitely many choices. To see this observe that while choosing $\pi_i(0), 1 \leq i \leq 4$ and the generating vertices of the $(2k-2)$ letters (except x), the value of $L(\pi_1(i)), \pi_1(i+1)$ as well as $\pi(i+1)$ have been chosen. That leaves only finitely many choices for $\pi_1(i)$ (for the particular link functions that we have considered). Hence in this case too, we have,

$$|\tilde{Q}_{k,4}^{2k-1}| \leq Mn^{2k+2}.$$

Case 3. $b = 2k$.

Note that each letter appears exactly twice across $\pi_1, \pi_2, \pi_3, \pi_4$. So there is a letter x that appears first in π_1 at the ith position and does not re-occur in π_1. Here our approach is to choose another letter across the four circuits with certain properties which we will leave aside while counting the rest of the generating vertices as we did in Case 2. Then, just as in Case 2, we will show that the generating vertices of both these letters have finitely many choices. Towards that we have the following subcases:

(a) π_1 shares a letter with only one of the other circuits, say π_2. Then π_3 has at least one letter, say y that appears first in π_3 at $(\pi_3(j-1), \pi_3(j))$ and does not appear in π_1. As the circuits are cross matched, y also does not re-occur in π_3. Now we rearrange the circuits as $\pi_1, \pi_3, \pi_2, \pi_4$. Next, we set aside the L-values, $L(\pi_1(i-1)), \pi_1(i))$ and $L(\pi_3(j-1)), \pi_3(j))$. Now the rest of the $(2k-2)$ generating vertices and $\pi_i(0), 1 \leq i \leq 4$ can be chosen in $\mathcal{O}(n^{4+2k-2}) = \mathcal{O}(n^{2k+2})$ ways.

Then using the same arguments as in Case 2 for π_1 and π_3, we see that $\pi_1(i)$ and $\pi_3(j)$ have only finitely many choices.

(b) π_1 shares a letter with exactly two other circuits, say π_2 and π_3. Then π_4 has at least one letter, say y that appears first in π_4 at $(\pi_4(j-1), \pi_4(j))$ and does not appear in π_1. As the circuits are cross matched, y also does not re-occur in π_4. Now we rearrange the circuits as $\pi_1, \pi_4, \pi_2, \pi_3$. Then we repeat the argument in (a) for $\pi_1(i)$ and $\pi_4(j)$, leaving them with only finitely many choices.

(c) π_1 shares a letter with all the other three circuits. Here there are two possibilities:

(i) one of the three circuits (π_2, π_3, π_4), say π_2, shares a letter with another circuit, say π_3. So there is a letter in π_2 which appears first in π_2, and re-occurs in π_3. Then, we can repeat the argument in (a) for π_1 and π_2.

(ii) none of the three circuits π_2, π_3, π_4, share any letter with each other. In this case, we identify one letter from π_2 that has appeared in π_1 (and not in π_3) and another letter from π_3 that has appeared in π_1 (and not in π_2). Then we rearrange the circuits as $\pi_2, \pi_3, \pi_1, \pi_4$ and repeat the argument given in (a) for π_2 and π_3.

Thus we get,

$$|\tilde{Q}_{k,4}^{2k}| \leq Mn^{2k+2}.$$

This completes the proof of the lemma. $\square$

Remark 4.4.2. *Note that for the Wigner matrix Lemma 3.4.7 states that $|Q_{k,4}^b| \leq Cn^{b+2}$. However that is not true for the patterned matrices of this chapter. For example consider the jointly- and cross matched circuits $\pi_1, \pi_2, \pi_3, \pi_4$ of length 2, $b = 2$, for the reverse circulant link function, such that the word formed across the four circuits is aabbaabb. Then it is easy to see that $\pi_1(0), \pi_1(1), \pi_2(0), \pi_2(1), \pi_3(0)$ and $\pi_4(0)$ can be freely chosen. Thus $|\tilde{Q}_{k,4}^b|$ cannot be bounded by n^4.*

4.4.1 Fully i.i.d. entries

Corollary 4.4.3. *Results 4.1.1, 4.1.2, 4.1.3 follow from Theorems 4.2.2, 4.2.3 and 4.2.4, respectively.*

Proof. Let A_n be any one of the four patterned matrices, $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$. Let the input sequence of A_n be $\{\frac{1}{\sqrt{n}}x_i : i \geq 0\}$, where x_i are independent and identically distributed with mean 0 and variance 1. In this case we first show that $\{\frac{1}{\sqrt{n}}x_i; i \geq 0\}$ satisfy Assumption A with $\tilde{g}_2 \equiv 1$ and $\tilde{g}_{2k} \equiv 0$ for all $k \geq 2$. The proof is similar to the arguments given in the proof of Corollary 3.5.1.

Let $r_n = n^{-1/3}$. Then $r_n\sqrt{n} \to \infty$ as $n \to \infty$ and

$$\lim_{n\to\infty} n\mathbb{E}\Big[\Big(x_i/\sqrt{n}\Big)^2 \mathbf{1}_{[|x_i/\sqrt{n}|\leq r_n]}\Big] = \lim_{n\to\infty} \mathbb{E}\Big[(x_0)^2 \mathbf{1}_{[|x_0|\leq \sqrt{n}r_n]}\Big] = 1.$$

Also, for any $k > 2$,

$$\begin{aligned}
n\mathbb{E}\Big[\Big(\frac{x_i}{\sqrt{n}}\Big)^k \mathbf{1}_{[|\frac{x_i}{\sqrt{n}}|\leq r_n]}\Big] &= n\ \mathbb{E}\Big[\Big(\frac{x_i}{\sqrt{n}}\Big)^{k-2}\Big(\frac{x_i}{\sqrt{n}}\Big)^2 \mathbf{1}_{[|x_i|\leq r_n\sqrt{n}]}\Big] \\
&\leq r_n^{k-2}\ \mathbb{E}\big[{x_i}^2 \mathbf{1}_{[|y_i|\leq r_n\sqrt{n}]}\big] \\
&\leq r_n^{k-2} \\
&= (n^{-\frac{1}{3}})^{k-2}\ \to\ 0 \quad \text{as } n\to\infty.
\end{aligned}$$

Now for any $t > 0$,

$$\begin{aligned}
\sum_{i=0}^{n-1} \mathbb{E}\big[(x_i/\sqrt{n})^2 \mathbf{1}_{[|x_i/\sqrt{n}|>r_n]}\big] &= \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[x_i^2 \mathbf{1}_{[|x_i|>r_n\sqrt{n}]}] \\
&\leq \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[x_i^2 \mathbf{1}_{[|x_i|>t]}]] \ \ \text{for all large } n, \\
&\xrightarrow{almost surely} \mathbb{E}\big[x_0^2[\mathbf{1}_{[|x_0|>t]}]\big], \ \ \text{as} \ \ n\to\infty.
\end{aligned}$$

As $\mathbb{E}(x_0^2) = 1$, the last term in the above expression tends to zero as $t \to \infty$.

Therefore, from Theorems 4.2.2- 4.2.4, the EESD of A_n converges weakly to a symmetric probability measure ν_A, say. Now, as A_n (with entries $\frac{1}{\sqrt{n}}x_i$) satisfies (4.4.1) (see Step 2 of the proof of Theorem 3.3.1), we have,

$$\begin{aligned}
&\frac{1}{n^4}\mathbb{E}\big[\operatorname{Tr}(A_n^k) - \mathbb{E}(\operatorname{Tr}(A_n^k))\big]^4 = \mathcal{O}(n^{-2}), \quad \text{and therefore,} \\
&\sum_{n=1}^{\infty} \frac{1}{n^4}\mathbb{E}\big[\operatorname{Tr}(A_n^k) - \mathbb{E}(\operatorname{Tr}(A_n^k))\big]^4 < \infty \quad \text{for every } k \geq 1. \qquad (4.4.2)
\end{aligned}$$

Then using Lemma 2.1.3, we can conclude that the ESD of A_n also converges weakly almost surely to ν_A.

We now identify ν_A in each case:

(i) **Reverse Circulant**: Suppose A_n is $R_n^{(s)}$. Then by Theorem 4.2.2,

$$\beta_{2k}(\nu_R) = \sum_{\sigma \in S(2k)} c_\sigma = \sum_{\sigma \in S_k(2k)} 1 = k!.$$

The second last equality holds since $c_{2k} = 0$ for all $k \geq 2$ as $\tilde{g}_{2k} \equiv 0$ for all $k \geq 2$ and $c_2 = \tilde{g}_2 = 1$. Hence ν_R is the *symmetrised Rayleigh distribution* as in Result 4.1.1.

(ii) **Symmetric Circulant**: Now suppose A_n is $C_n^{(s)}$. Then by Theorem 4.2.3,

$$\beta_{2k}(\mu) = \sum_{\sigma \in E(2k)} a_\sigma c_\sigma = \sum_{\sigma \in E_k(2k)} 1 = \frac{(2k)!}{2^k k!}.$$

The second last equality in the above expression holds since $c_{2k} = 0$ for all $k \geq 2$ as $\tilde{g}_{2k} \equiv 0$ for all $k \geq 2$, $c_2 = \tilde{g}_2 = 1$ and $a_2 = 1$. Thus ν_R is the standard normal distribution as in Result 4.1.2.

(iii) **Toeplitz and Hankel**: Similarly, choosing A_n as the Toeplitz and Hankel matrices, we get the almost sure convergence of the ESDs, which were originally showed in Hammond and Miller [2005] and Bryc et al. [2006]. See also Bose and Sen [2008]. $\square$

4.4.2 General triangular i.i.d.

The next corollary states LSD results for general triangular i.i.d. matrices. Here the condition of full i.i.d. entries is relaxed, and the distribution of the entries can vary with n, the size of the matrix.

Corollary 4.4.4. *Let A_n be one of the four patterned matrices, $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$. Suppose for each fixed n, the input sequence $\{x_{i,n} : i \geq 0\}$ are i.i.d. with all moments finite. Assume that for all $k \geq 1$,*

$$n\mathbb{E}[x_{0,n}^k] \to c_k \quad as \quad n \to \infty. \tag{4.4.3}$$

Also assume that the moments of the random variable whose cumulants are $\{0, c_2, 0, c_4, \ldots\}$ satisfy Carleman's condition. Then the EESD of $R_n^{(s)}$, $C_n^{(s)}$, $T_n^{(s)}$ and $H_n^{(s)}$ converge weakly to symmetric probability measures, whose moments are determined by $\{c_{2k}\}_{k \geq 1}$.

Proof. Observe that Assumption B (i), (ii) and (iii) are satisfied with $r_n = \infty$ and $\tilde{g}_{2k} \equiv c_{2k}$ for $k \geq 1$. Hence from Theorems 4.2.2-4.2.4 the EESD of $R_n^{(s)}$, $C_n^{(s)}$, $T_n^{(s)}$ and $H_n^{(s)}$ converge weakly to symmetric probability measures, whose moments are as given in the respective theorems. □

Remark 4.4.5. *Let $\{x_{i,n} : i \geq 1\}$ be i.i.d. with all moments finite, for every fixed n. Assume that $\sum_{i=1}^{n} x_{i,n}$ converges in distribution to a limit distribution F whose cumulants are $\{c_k\}_{k\geq 1}$. This assumption is equivalent to Condition (4.4.3).*

In particular, if we start with an infinitely divisible distribution F with all moments finite, we can definitely find such i.i.d. random variables $\{x_{i,n} : i \geq 1\}$ (Characterization 1 in Bose et al. [2002]).

4.4.3 Sparse entries

Corollary 4.4.6. *Result 4.1.4 follows from Corollary 4.4.4.*

Proof. In this case the input sequence $\{x_{i,n} : i \geq 0\}$ of the patterned matrices are $\text{Ber}(p_n)$ where $np_n \to \lambda > 0$. Observe that, (4.4.3) is satisfied with $c_k = \lambda$ for all $k \geq 1$. Therefore, from Corollary 4.4.4, we obtain that the EESD of A_n converges weakly.

Let us now look at the particular cases.

(i) If A_n is the sparse reverse circulant matrix, then

$$\beta_{2k}(\nu_R) = \sum_{\sigma \in S(2k)} \lambda^{|\sigma|}. \tag{4.4.4}$$

Therefore, the *half cumulants* $(r_{2n})_{n\geq 1}$ of ν_R are $r_{2n} = \lambda, n \geq 1$. For more details on half cumulants, see Section 2.5.3.

(ii) If A_n is the sparse symmetric circulant, then

$$\beta_{2k}(\nu_C) = \sum_{\sigma \in E(2k)} a_\sigma \lambda^{|\sigma|}. \tag{4.4.5}$$

All odd cumulants of ν_C vanish, and its even cumulants are $\{a_{2n}c_{2n}\}_{n\geq 1}$, where a_{2k} is the $2k$th moment of a random variable Z defined in Remark 4.3.5. Now as $c_{2n} = \lambda$

for all n, we get that ν_C is the *compound Poisson distribution* with rate λ and jump distribution F_Z, where F_Z is the distribution of Z (see Remark 4.3.5).

(iii) Now suppose A_n is the sparse Toeplitz matrix. Then its EESD converges weakly to ν_T and (see (4.3.49))

$$\beta_{2k}(\nu_T) = \sum_{\sigma \in E(2k)} \alpha(\sigma)\lambda^{|\sigma|} \tag{4.4.6}$$

where $\alpha(\sigma)$ is obtained from the different linear combinations of s_j's corresponding to σ.

(iv) Finally, suppose A_n is the sparse Hankel matrix. Then its EESD converges weakly to ν_H and

$$\beta_{2k}(\nu_H) = \sum_{\sigma \in S(2k)} \alpha(\sigma)\lambda^{|\sigma|} \tag{4.4.7}$$

where $\alpha(\sigma)$ is obtained as the value of an integral corresponding to σ (see (4.3.53)). $\square$

Remark 4.4.7. *(Sums of sparse matrices) From this discussion on sparse matrices we can also conclude that the EESD of finite sums of i.i.d. copies of sparse matrices with independent entries converge weakly to a symmetric probability measure. This can be observed in the following manner.*

Suppose $A_{n,1}, A_{n,2}, \ldots, A_{n,m}$ are m independent $n \times n$ matrices with any of the four patterns and whose entries are independent Ber(p_n) where $np_n \to \lambda > 0$. Then the entries $\{x_{i,n} : i \geq 0\}$ of $A_n := \sum_{k=1}^{m} A_{n,k}$ are independent Bin(m, p_n). Then for $i \geq 0$,

$$\mathbb{E}[x_{i,n}^k] = mp_n(1-p_n)^{m-1} + \sum_{j=2}^{m} j^k \binom{k}{j} p_n^j (1-p_n)^{m-j} = mp_n + o(p_n).$$

Clearly (4.4.3) is satisfied with $c_k = m\lambda$ for all $k \geq 1$. Hence, from the above discussion, the EESD of A_n converges to a symmetric probability measure in all of the four cases. In each case the $2k$th moments of the limiting distribution are obtained accordingly, replacing λ by $m\lambda$ in (4.4.4), (4.4.5), (4.4.6) and (4.4.7).

4.4.4 Matrices with variance profile

In the next two corollaries, we describe results about the four *patterned matrices with variance profile* $(A_n, \cdot)$, where A_n is any one of the four patterned matrices. Like the

Wigner matrices (see Definitions 3.5.11), we discuss two kinds of variance profiles- *discrete variance profile* and *continuous variance profile.*

Definition 4.4.8. *(a) Discrete variance profile: Suppose $\{x_{i,n}; 0 \leq i \leq 2n\}$ are i.i.d. random variables with mean zero and variance 1, and let $\sigma_d = \{\sigma_i\}_{0 \leq i \leq 2n}$ be uniformly bounded real numbers. Then the patterned matrix (A_n, σ_d), where A_n is one of the four patterned matrices, with discrete variance profile σ_d, is given by (L being the link function for $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$)*

$$(A_n, \sigma_d) = \big((y_{L(i,j),n} = \sigma_{L(i,j)} x_{L(i,j),n})\big)_{1 \leq i,j \leq n}. \tag{4.4.8}$$

(b) Continuous variance profile: Suppose $\{x_{i,n}; i \leq j\}$ are i.i.d. random variables for every fixed n and let σ be a symmetric bounded piecewise continuous function on $[0, 1]$. Then the patterned matrix (A_n, σ_c) (where A_n is one of the four patterned matrices) with continuous variance profile σ_c, is given by (L being the link function for $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$)

$$(A_n, \sigma_c) = \big((y_{L(i,j),n} = \sigma(L(i,j)/n) x_{L(i,j),n})\big)_{1 \leq i,j \leq n}. \tag{4.4.9}$$

Corollary 4.4.9. ***(Discrete variance profile)*** *Consider the matrix (A_n, σ_d) where A_n is either $R_n^{(s)}$ or $C_n^{(s)}$ with input sequence $\{y_{i,n}; 0 \leq i \leq n\}$ as described in (4.4.8). Assume that the random variables $\{x_{i,n}; 0 \leq i \leq n\}$ are i.i.d. for every n, and satisfy the conditions of Corollary 4.4.4.*

Further let $\sigma_d = \{\sigma_i : i \geq 0\}$ satisfy the following:

(i) $\sup_i |\sigma_i| \leq c < \infty$.

(ii) For any $k \geq 1$,

$$\frac{1}{n} \sum_{i=0}^{n-1} \sigma_i^{2k} \to \alpha_{2k}, \text{ say.} \tag{4.4.10}$$

Under these conditions, the EESD of (A_n, σ_d) converges weakly to some symmetric probability distribution $\tilde{\nu}_A$ whose moments are determined by $\{\alpha_{2n}\}_{n \geq 1}$ and $\{c_{2n}\}_{n \geq 1}$.

Proof. We first prove the corollary for the reverse circulant matrix, and then for the symmetric circulant matrix.

Reverse Circulant: Let $c_{2k,n} = n\mathbb{E}[x_{0,n}^{2k}]$. Observe that the entries of $A_n = R_n^{(s)}$ satisfy (4.2.1) with $r_n = \infty$ and $\tilde{g}_{2k,n}(\frac{i}{n}) = \sigma_i^{2k} c_{2k,n}$ for $0 \leq i \leq n-1$. Also because of (4.4.3), we have (4.2.2) with $r_n = \infty$.

Now from Step 2 of the proof of Theorem 4.2.2, observe that the contribution of the non-symmetric words is 0. Suppose $\boldsymbol{\omega}$ is a symmetric word of length $2k$ with b distinct letters which appear $k_1, k_2, \ldots, k_b$ times. Then using (4.3.29), the contribution of this word to the limiting moment is the following:

$$\lim_{n\to\infty} \frac{1}{n^{b+1}} \sum_S \prod_{j=1}^{b} \sigma_{L(m_j,l_j)}^{k_j} \prod_{j=1}^{b} c_{k_j,n}, \tag{4.4.11}$$

where L is the link function of the reverse circulant matrix, S is the set of distinct generating vertices, and (m_j, l_j), $1 \leq j \leq b$ are as in the proof of Theorem 4.2.2.

Observe that for every $k \geq 1$, $c_{k,n}$ does not depend on the values of the generating vertices as $\{x_{i,n}\}$ are i.i.d. Also from (4.4.3) we have $c_{k,n} \to c_k$ as $n \to \infty$. Therefore to obtain the contribution of $\boldsymbol{\omega}$ to the limiting moment, it is enough to compute $\lim\limits_{n\to\infty} \frac{1}{n^{b+1}} \sum\limits_S \prod\limits_{j=1}^{b} \sigma_{L(m_j,l_j)}^{k_j}$.

Now for any $j \in \{1, 2, \ldots, b\}$, $m_j = \pi(i_j - 1)$ is either a generating vertex, or can be written as a linear combination of the *previous* l_q's. Therefore, in any case, $m_j = LC_j(\{l_q : 0 \leq q \leq j-1\})$, where LC_j denotes the linear representation. As a result, m_j does not depend on the value of l_j, and hence does not change with l_j. With this observation we obtain that for any $s \geq 1$,

$$\begin{aligned}
\frac{1}{n}\sum_{l_j=1}^{n} \sigma_{m_j,l_j}^{2s} &= \frac{1}{n} \sum_{l_j:\ m_j+l_j-2<n} \sigma_{m_j+l_j-2}^{2s} + \frac{1}{n} \sum_{l_j:\ m_j+l_j-2\geq n} \sigma_{(m_j+l_j-2)-n}^{2s} \\
&= \frac{1}{n} \sum_{t=m_j-2}^{n-1} \sigma_t^{2s} + \frac{1}{n} \sum_{t=0}^{m_j-3} \sigma_t^{2s} \\
&= \frac{1}{n} \sum_{t=0}^{n-1} \sigma_t^{2s} \to \alpha_{2s} \quad \text{as} \quad n \to \infty.
\end{aligned}$$

As a consequence, we get

$$\lim_{n\to\infty} \frac{1}{n^{b+1}} \sum_S \prod_{j=1}^{b} \sigma_{L(m_j,l_j)}^{k_j} = \prod_{j=1}^{b} \alpha_{k_j}.$$

Hence the contribution of any symmetric word of length $2k$, with b distinct letters that appears $k_1, k_2, \ldots, k_b$ times, to the limiting moment is, $\prod_{j=1}^{b} \alpha_{k_j} c_{k_j}$. Hence we have

$$\lim_{n\to\infty} \mathbb{E}[\mathrm{Tr}(R_n^{(s)})^k] = \begin{cases} \sum_{\pi\in S(k)} \alpha_\pi c_\pi & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd.} \end{cases} \tag{4.4.12}$$

Thus we have verified the first moment condition for the reverse circulant matrix with a variance profile.

As $\sup_i |\sigma_i| \le c$,

$$\sum_{\pi\in S(2k)} \alpha_\pi c_\pi \le \max\{c, 1\}^k \sum_{\pi\in E(2k)} c_\pi.$$

Since the right side of the above inequality satisfies the Carleman's condition, there exists a unique probability measure $\tilde{\nu}_R$ such that its moments $\{\beta_k(\tilde{\nu}_R)\}_{k\ge1}$ are as in (4.4.12). Hence the EESD of $R_n^{(s)}$ converges weakly to $\tilde{\nu}_R$.

Symmetric Circulant: The arguments in this case are similar to the reverse circulant case. Let $c_{2k,n} = n\mathbb{E}[x_{0,n}^{2k}]$. Observe that the entries of $A_n = C_n^{(s)}$ satisfies (4.2.1) and (4.2.2) with $r_n = \infty$ and $\tilde{g}_{2k,n}(\frac{i}{n}) = \sigma_i^{2k} c_{2k,n}$ for $0 \le i \le n-1$.

From the proof of Theorem 4.2.3 observe that, the words that are not even, contribute 0. Now, let $\boldsymbol{\omega}$ be an even word of length $2k$ with b distinct letters which appear $k_1, k_2, \ldots, k_b$ times. From (4.3.43), observe that there are $\prod_{i=1}^{b} \binom{k_i-1}{\frac{k_i}{2}}$ sets of equations that contribute identically. Thus, from (4.3.39) the contribution for each such set of equations is now as follows:

$$\lim_{n\to\infty} \frac{1}{n^{b+1}} \sum_{S} \prod_{j=1}^{b} \sigma_{L(m_j,l_j)}^{k_j} c_{k_j,n}, \tag{4.4.13}$$

where L is the symmetric circulant link function, S is the set of distinct generating vertices and (m_j, l_j), $1 \le j \le b$, are as in the proof of Theorem 4.2.3.

As in the reverse circulant case, it is enough to compute $\lim_{n\to\infty} \frac{1}{n^{b+1}} \sum_{S} \prod_{j=1}^{b} \sigma_{L(m_j,l_j)}^{k_j}$ in order to obtain the contribution of the word $\boldsymbol{\omega}$ to the limiting moment. Note that $m_j = LC_j(\{l_q : 0 \le q \le j-1\})$ where LC_j denotes the linear representation. As a result, m_j does not depend on l_j. Hence for any $s \geq 1$,

$$\frac{1}{n}\sum_{l_j}\sigma^{2s}_{L(m_j,l_j)} = \frac{1}{n}\sum_{t=0}^{m_j-1}\sigma_t^{2s} + \frac{1}{n}\sum_{t=m_j}^{\frac{n}{2}-1}\sigma_t^{2s} + \frac{2}{n}\sum_{t=0}^{\frac{n}{2}-1}\sigma_t^{2s} = 2\frac{2}{n}\sum_{t=0}^{n/2-1}\sigma_t^{2s} \to 2\alpha_{2s}. \quad (4.4.14)$$

As a consequence, we have

$$\lim_{n\to\infty}\frac{1}{n^{b+1}}\sum_S\prod_{j=1}^{b}\sigma^{k_j}_{L(m_j,l_j)} = \prod_{j=1}^{b}2^{k_j}\alpha_{k_j}.$$

Now recall from Lemma 4.3.2, that for $\boldsymbol{\omega}$, there are $\prod_{j=1}^{b} a_{k_j}$ (where $a_{2n} = \frac{1}{2}\binom{2n}{n}$) equations for determining the non-generating vertices once the generating vertices are chosen. Therefore, the contribution of any even word of length $2k$ with b distinct letters to the limiting moment is $\prod\limits_{j=1}^{b} 2^{k_j} a_{k_j}\alpha_{k_j}c_{k_j}$.

Let $\delta_{2k} = 2\alpha_{2k}$. Then we have that

$$\lim_{n\to\infty}\mathbb{E}[\mathrm{Tr}(C_n^{(s)})^k] = \begin{cases} \sum\limits_{\pi\in E(k)} a_\pi\delta_\pi c_\pi & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd.} \end{cases} \quad (4.4.15)$$

So, there is a unique probability measure $\tilde{\nu}_C$ such that its moments are as in (4.4.15). □

Remark 4.4.10. *Note that in Corollary 4.4.9, if the input sequence is $\{\frac{y_{i,n}}{\sqrt{n}}\}$ (as described in (4.4.8)) with $\{x_{i,n}\}$ fully i.i.d. that follows a distribution F for all i and n, then $c_{2k} = 0$ for all $k \geq 2$. So from Corollary 4.4.9, the EESD of $(R_n^{(s)}, \sigma_d)$ and $(C_n^{(s)}, \sigma_d)$ converge weakly to symmetric probability measures that depend on α_2. Again as σ_i is bounded uniformly, (4.4.1) and hence (4.4.2) hold true. Hence in this case the almost sure convergence of the ESD in Corollary 4.4.9 holds.*

(i) The almost sure LSD of $(R_n^{(s)}, \sigma_d)$ with input sequence $\{\frac{y_{i,n}}{\sqrt{n}}\}$, is the distribution of $Y = \sqrt{\alpha_2}\mathcal{R}$, where $\mathcal{R}$ is a random variable with the symmetrised Rayleigh distribution.

(ii) The almost sure LSD of $(C_n^{(s)}, \sigma_d)$ with input sequence $\{\frac{y_{i,n}}{\sqrt{n}}\}$) is the distribution of $\sqrt{2\alpha_2}N$ where N is a standard Gaussian variable.

Corollary 4.4.11. *(**Continuous variance profile**) Consider the matrix (A_n, σ_c) where A_n is any of the four patterned matrices with input sequence $\{y_{i,n}; 0 \leq i \leq 2n\}$*

as described in (4.4.9). Assume that the random variables $\{x_{i,n}; 0 \leq i \leq n\}$ are i.i.d. for every n, and satisfy the conditions of Corollary 4.4.4. Then the EESD of (A_n, σ_c) converges weakly almost surely to a symmetric probability measure $\tilde{\nu}_A$ whose moments are determined by σ and $\{c_{2k}\}_{k\geq 1}$.

Proof. Let $c_{2k,n} = n\mathbb{E}[x_{0,n}^{2k}]$. First observe that the entries of A_n satisfy Assumption B (i) and (ii) with $r_n = \infty$, $\tilde{g}_{2k,n} = \sigma^{2k} c_{2k,n}$ and $\tilde{g}_{2k} = \sigma^{2k} c_{2k}$. Since σ is a continuous function on a compact set $[0, 1]$, it is bounded. Therefore, Assumption B (iii) is also true. Hence from Theorems 4.2.2- 4.2.4, we can conclude that the EESD of A_n converges weakly to a symmetric probability distribution, say $\tilde{\nu}_A$.

In addition, the limiting moments of $(R_n^{(s)}, \sigma_c)$ and $(C_n^{(s)}, \sigma_c)$ can be expressed in a simplified form as follows:

Reverse Circulant, $(R_n^{(s)}, \sigma_c)$: From Theorem 4.2.2, all odd moments of the LSD are 0, and the $2k$th moment is given by

$$\beta_{2k} = \sum_{\pi \in S(2k)} \alpha_\pi c_\pi, \quad \text{where } \alpha_{2m} = \int_0^1 \sigma^{2m}(t)\, dt. \tag{4.4.16}$$

Symmetric Circulant, $(C_n^{(s)}, \sigma_c)$: From Theorem 4.2.3, all odd moments of the LSD are 0, and the $2k$th moment is given by

$$\beta_{2k} = \sum_{\pi \in S(2k)} a_\pi \alpha_\pi c_\pi, \quad \text{where } \alpha_{2m} = 2\int_0^1 \sigma^{2m}(t)\, dt. \tag{4.4.17}$$

The limiting moments for the Toeplitz and Hankel matrices are more involved, and can be calculated using (4.3.50) and (4.3.54) in Theorem 4.2.4. Simply replace the function in the integrand by $\prod_{j=1}^{b} c_{k_j} \sigma^{k_j}(|x_{m_j} - x_{l_j}|)$ and $\prod_{j=1}^{b} c_{k_j} \sigma^{k_j}(x_{m_j} + x_{l_j})$ respectively.

$\square$

Remark 4.4.12. *Just as in Remark 4.4.10, suppose the input sequence is $\{\frac{y_{i,n}}{\sqrt{n}}\}$ (as described in (4.4.9)) in Corollary 4.4.11 with $\{x_{i,n}\}$ fully i.i.d. with a distribution F for all i and n. Then $c_{2k} = 0$ for all $k \geq 2$. So from Corollary 4.4.11, the EESD of $(A_n^{(s)}, \sigma_c)$ converges weakly to symmetric probability measures that depend on σ^2. Again, as σ is bounded, (4.4.1), and hence (4.4.2) hold true. Hence in this case, the almost sure* *convergence of the ESD of* $(A_n^{(s)}, \sigma_c)$ *holds. In particular, for the reverse circulant and symmetric circulant matrices we have the following:*

(i) The almost sure LSD of $(R_n^{(s)}, \sigma_c)$ *(where the input sequence is* $\{\frac{y_{i,n}}{\sqrt{n}}\}$*) is the distribution of* $Y = \sqrt{\alpha_2}\mathcal{R}$*, where* $\mathcal{R}$ *is a random variable with the symmetrised Rayleigh distribution, see* (4.4.16).

(ii) The almost sure LSD of $(C_n^{(s)}, \sigma_c)$ *(where the input sequence is* $\{\frac{y_{i,n}}{\sqrt{n}}\}$*) is the distribution of* $\sqrt{2\alpha_2}N$ *where* N *is a standard Gaussian variable, see* (4.4.17) .

4.4.5 Band and Triangular matrices

We have seen band Wigner matrices in Chapter 3. In a similar spirit, here we discuss the banded versions of the patterned matrices, $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$.

Definition 4.4.13. *(**Banded versions of*** $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$*) Let* m_n *be a sequence of positive integers such that* $m_n \to \infty$ *and* $m_n/n \to \alpha > 0$ *as* $n \to \infty$*. we discuss two banding models- Type I banding* (A_n^b) *and Type II banding* (A_n^B)*.*

Type I banding*:* A_n^b *is the symmetric matrix with entries* $\{y_{L(i,j),n}\}$*, where the input sequence* $y_{i,n}$ *is as follows*

$$y_{i,n} = \begin{cases} x_{i,n} & \text{if } i \le m_n, \\ 0 & \text{otherwise.} \end{cases} \tag{4.4.18}$$

Type II banding*: The Type II band versions* $R_n^{(s)B}$ *of* $R_n^{(s)}$*, and* $T_n^{(s)B}$ *of* $T_n^{(s)}$*, are defined with input sequence* $\hat{y}_{i,n}$ *where*

$$\hat{y}_{i,n} = \begin{cases} x_{i,n} & \text{if } i \le m_n \text{ or } i \ge n - m_n, \\ 0 & \text{otherwise.} \end{cases} \tag{4.4.19}$$

The Type II band versions $H_n^{(s)B}$ *of* $H_n^{(s)}$ *is defined with input sequence* $\tilde{y}_{i,n}$ *where*

$$\tilde{y}_{i,n} = \begin{cases} x_{i,n} & \text{if } n - m_n \le i \le n + m_n, \\ 0 & \text{otherwise.} \end{cases} \tag{4.4.20}$$

For example suppose $m_n \sim \left[\frac{n}{3}\right]$ where $[\cdot]$ is the greatest integer function. At $n = 5$,

$$R_5^{(s)b} = \begin{bmatrix} x_0 & x_1 & 0 & 0 & 0 \\ x_1 & 0 & 0 & 0 & x_0 \\ 0 & 0 & 0 & x_0 & x_1 \\ 0 & 0 & x_0 & x_1 & 0 \\ 0 & x_0 & x_1 & 0 & 0 \end{bmatrix}, \quad C_5^{(s)b} = \begin{bmatrix} x_0 & x_1 & 0 & 0 & x_1 \\ x_1 & x_0 & x_1 & 0 & 0 \\ 0 & x_1 & x_0 & x_1 & 0 \\ 0 & 0 & x_1 & x_0 & x_1 \\ x_1 & 0 & 0 & x_1 & x_0 \end{bmatrix},$$

$$R_5^{(s)B} = \begin{bmatrix} x_0 & x_1 & 0 & 0 & x_4 \\ x_1 & 0 & 0 & x_4 & x_0 \\ 0 & 0 & x_4 & x_0 & x_1 \\ 0 & x_4 & x_0 & x_1 & 0 \\ x_4 & x_0 & x_1 & 0 & 0 \end{bmatrix}, \quad T_5^{(s)B} = \begin{bmatrix} x_0 & x_1 & 0 & 0 & x_4 \\ x_1 & x_0 & x_1 & 0 & 0 \\ 0 & x_1 & x_0 & x_1 & 0 \\ 0 & 0 & x_1 & x_0 & x_1 \\ x_4 & 0 & 0 & x_1 & x_0 \end{bmatrix}.$$

Corollary 4.4.14. *Suppose that the random variables $\{x_{i,n}; i \geq 0\}$ associated with the matrices A_n^b as in (4.4.18), are i.i.d. for every n, and satisfy the conditions of Corollary 4.4.4. Then, the EESD of A_n^b converge weakly to some symmetric probability measures μ_α that depend on $\{c_{2k}\}_{k\geq 1}$, and $\alpha = \lim\limits_{n\to\infty} \dfrac{m_n}{n} > 0$.*

Proof. For every n, define the function σ_n on the interval $[0, 1]$ as

$$\sigma_n(x) = \begin{cases} 1 & \text{if } x \leq m_n/n, \\ 0 & \text{otherwise.} \end{cases}$$

Now observe that the entries $y_{i,n}$ of the matrix A_n^b can be written as $\sigma_n\big(i/n\big)x_{i,n}$.

Observe that, for any $k \geq 1$, $\int \sigma_n^k(x)\, dx \to \int \sigma^k(x)\, dx$ as $n \to \infty$, where σ equals

$$\sigma(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq \alpha, \\ 0 & \text{otherwise.} \end{cases} \tag{4.4.21}$$

Following the proofs in Theorems 4.2.2—4.2.4 and the above convergence of $\int \sigma_n^k(x)\, dx$, it is easy to argue that the first moment condition holds for A_n^b. Hence the EESD of A_n^b converges weakly to a symmetric probability distribution.

The formulae for the limiting moments are as follows:

Reverse Circulant, $R_n^{(s)b}$: Clearly from Theorem 4.2.2, the odd moments of the LSD are all zero, and the $2k$th moment of the limiting distribution is given by

$$\beta_{2k} = \sum_{\pi \in S(2k)} \alpha^{|\pi|} c_\pi, \text{ as } \int_0^1 \sigma^{2m}(t)\ dt = \int_0^\alpha \sigma^{2m}(t)\ dt = \alpha.$$

Symmetric Circulant, $C_n^{(s)b}$: Clearly from Theorem 4.2.3, the odd moments of the LSD are all zero, and the $2k$th moment of the limiting distribution is given by

$$\beta_{2k} = \sum_{\pi \in S(2k)} (2\alpha)^{|\pi|} a_\pi c_\pi, \text{ as } 2\int_0^1 \sigma^{2m}(t)\ dt = 2\int_0^\alpha \sigma^{2m}(t)\ dt = 2\alpha.$$

Similarly, the moments of the Toeplitz and Hankel matrices can be calculated from (4.3.50) and (4.3.54) in Theorem 4.2.4 where the range of the integral is $[0,\alpha]$, and the function in the integrand is replaced by $\prod_{j=1}^b c_{k_j}\sigma^{k_j}(|x_{m_j} - x_{l_j}|)$ and $\prod_{j=1}^b c_{k_j}\sigma^{k_j}(x_{m_j} + x_{l_j})$ respectively. $\square$

Corollary 4.4.15. *Result 4.1.5 follows from Corollary 4.4.14.*

Proof. In this case the entries of A_n^b are $\frac{y_{i,n}}{\sqrt{m_n}}$ (see (4.4.18)), where $\{x_{i,n}; i \geq 0\}_{n\geq 1}$ are fully i.i.d. random variables with mean 0 and variance 1. Now using the truncation argument in the proof of Corollary 4.4.3, it can be seen that, we can assume the variables $\{x_i\}$ to be uniformly bounded. Therefore conditions of Corollary 4.4.14 hold with $c_2 = \frac{1}{\alpha}$ and $c_{2k} = 0$ for $k \geq 2$. Hence, we obtain the convergence of the EESD from Corollary 4.4.14. Again, it can be verified that (4.4.1) and (4.4.2) hold true. Thus we obtain the almost sure convergence of the ESD. This yields Result 4.1.5. $\square$

Corollary 4.4.16. *Result 4.1.7 (a) follows from Theorem 4.2.4.*

Proof. In this case, the entries of the band Toeplitz matrix are $y_{i,n} = \frac{y_i}{\sqrt{(2-\alpha)\alpha n}}$ where $\alpha = \lim_{n\to\infty} \frac{m_n}{n} > 0$, $\{x_i; i \geq 0\}$ are independent variables with mean 0, variance 1 and $\sup_i \mathbb{E}[|x_i|^k] = M_k < \infty$ for $k \geq 3$. Then Assumption B(i), (ii) and (iii) are satisfied with $r_n = \infty$ and $\tilde{g}_2 \equiv 1$, $\tilde{g}_{2k} \equiv 0$ for $k \geq 2$. Thus from Theorem 4.2.4, the EESD of $T_n^{(s)^b}$ converges weakly almost surely to a symmetric probability distribution, say $\gamma_T(\alpha)$ whose even moments depend on α. Further, just like the proof of Corollary 4.4.15, it can be verified here that (4.4.1) and (4.4.2) hold true. Hence we obtain the almost sure convergence of the ESD. □

The next corollaries deals with the Type II banding.

Corollary 4.4.17. *Suppose that the random variables $\{x_{i,n}; i \geq 0\}$ associated with the matrices A_n^B as in (4.4.19), (4.4.20) are i.i.d. for every n, and satisfy the conditions of Corollary 4.4.4. Then, the EESD of A_n^B converge weakly to some symmetric probability measure μ_α that depend on $\{c_{2k}\}_{k\geq 1}$ and $\alpha = \lim_{n\to\infty} \frac{m_n}{n} > 0$.*

Proof. For Type II band versions $R_n^{(s)B}$ of $R_n^{(s)}$ and $T_n^{(s)B}$ of $T_n^{(s)}$, define $\sigma_{n,1}$ on $[0,1)$ as

$$\sigma_{n,1}(x) = \begin{cases} 1 & \text{if } x \leq m_n/n \text{ or } x \geq 1 - m_n/n, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, $\int \sigma_{n,1}^k(x)\, dx \to \int \sigma_1^k(x)\, dx$ as $n \to \infty$ for each $k \geq 1$ where $\sigma_1 = \mathbf{1}_{[0,\alpha]\cup[1-\alpha,1]}$.

For Type II band versions $H_n^{(s)B}$ of $H_n^{(s)}$, we define a function $\sigma_{n,2}$ on $[0,2)$ as

$$\sigma_{n,2}(x) = \begin{cases} 1 & \text{if } 1 - m_n/n \leq x \leq 1 + m_n/n, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, $\int \sigma_{n,2}^k(x)\, dx \to \int \sigma_2^k(x)\, dx$ as $n \to \infty$ for each $k \geq 1$ where $\sigma_2 = \mathbf{1}_{[1-\alpha,1+\alpha]}$.

Following the arguments as in the proof of Corollary 4.4.14, the convergence of the EESD of A_n^B follows. Now we compute the moments of the LSD of these matrices.

Reverse Circulant: Clearly all odd moments of the LSD are 0. From Theorem 4.2.2, the $2k$th moment of the LSD is given by

$$\beta_{2k} = \sum_{\pi \in S(2k)} (2\alpha)^{|\pi|} c_\pi, \text{ as } \int_0^\alpha \sigma_1^{2m}(t)\, dt + \int_{1-\alpha}^1 \sigma_1^{2m}(t)\, dt = 2\alpha.$$

Similarly, the moments of the LSD for $T_n^{(s)B}$ and $H_n^{(s)B}$ can be calculated by using (4.3.50) and (4.3.54) in Theorem 4.2.4, where the function in the integrand is replaced by $\prod_{j=1}^b c_{k_j} \sigma_1^{k_j}(|x_{m_j} - x_{l_j}|)$ and $\prod_{j=1}^b c_{k_j} \sigma_2^{k_j}(x_{m_j} + x_{l_j})$ respectively. As $\int_0^1 \sigma_1^{2m}(t)\, dt$ and $\int_0^1 \sigma_2^{2m}(t)\, dt$ are dependent on α, the limiting distribution also depends on α. □

Corollary 4.4.18. *Result 4.1.6 follows from Corollary 4.4.17.*

The proof of this corollary is similar to that of Corollary 4.4.15, and is omitted.

Corollary 4.4.19. *Result 4.1.7 (b) follows from Theorem 4.2.4.*

The proof of this corollary is similar to that of Corollary 4.4.16, and is omitted.

We have already seen LSD results for triangular Wigner matrices in Chapter 3. Now we shall see some analogous results for triangular versions of $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$.

Definition 4.4.20. *(**Triangular versions of** $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$) Let A_n be one of the $n \times n$ matrices $R_n^{(s)}, C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$. Then the triangular version of A_n, denoted by A_n^u is the matrix whose entries $y_{L(i,j),n}$ are as follows:*

$$y_{L(i,j),n} = \begin{cases} x_{L(i,j),n} & \text{if } (i+j) \leq n+1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.4.22}$$

Note that the triangular reverse circulant and the triangular Hankel matrices are same.

Then we have the following result.

Corollary 4.4.21. *Consider the triangular matrix A_n^u where A_n is one of the matrices $C_n^{(s)}, T_n^{(s)}, H_n^{(s)}$. Assume that the variables $\{x_{i,n}; i \geq 0\}$ associated with the matrices A_n^u (as in (4.4.22)) are i.i.d. random variables with all moments finite for every fixed n, and satisfy conditions of Corollary 4.4.4. For all these three matrices, the EESD of A_n^u converge weakly to some symmetric probability measures μ_A that depend on $\{c_{2k}\}_{k\geq 1}$.*

Proof. The proof of this follows from the same argument given in Corollary 4.4.14, with σ being replaced by $\eta : [0,1]^2 \to \mathbb{R}$ where $\eta(x,y) = \mathbf{1}_{[x+y\leq 1]}$. We skip the details. □

Corollary 4.4.22. *Result 4.1.8 follows from Corollary 4.4.21.*

Proof. In this case the entries of A_n^u are $\frac{y_{i,n}}{\sqrt{n}}$, where $\{y_{i,n}; i \geq 0\}$ are as in (4.4.22), and $\{x_{i,n}; i \geq 0\}_{n\geq 1}$ are i.i.d. random variables with mean 0 and variance 1. Now using the truncation argument in the proof of Corollary 4.4.3, it can be seen that we can assume the variables $\{x_i\}$ to be uniformly bounded. Therefore conditions of Corollary 4.4.21

hold with $c_2 = \frac{1}{\alpha}$ and $c_{2k} = 0$ for $k \geq 2$. Hence from Corollary 4.4.21, we obtain the convergence of the EESD. Again it can be verified that (4.4.1) and (4.4.2) hold true. Thus we obtain the almost sure convergence of the ESD. This yields Result 4.1.8. $\square$

4.5 Simulation

Below we give a few simulations which demonstrate that when the entries are i.i.d. $N(0,1)/\sqrt{n}$ (top rows of Figures 4.1 and 4.2), we have almost sure convergence. However the same is not true when the entries are i.i.d. $Ber(3/n)$ (bottom rows of Figure 4.1 and 4.2). See Remark 4.2.1. Figure 4.3 demonstrates the almost sure convergence of symmetric triangular (top row) Toeplitz and Hankel matrices and banded (bottom row) Toeplitz and Hankel matrices when the non-zero entries are i.i.d. $N(0,1)/\sqrt{n}$.

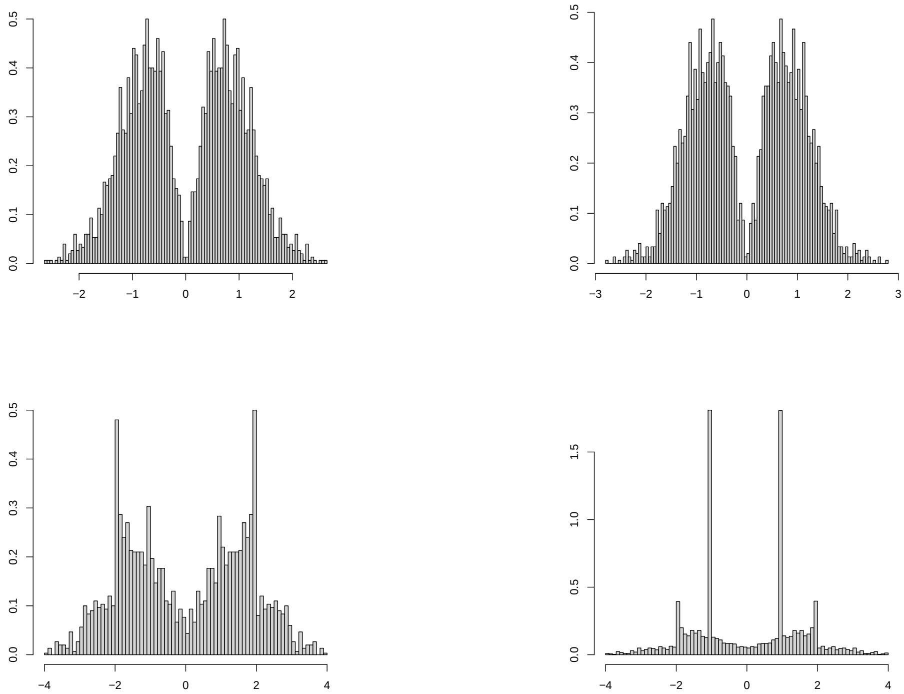


FIGURE 4.1: Two replications of the histograms of the eigenvalues of $R_n^{(s)}$ for $n = 1000$ where the entries are i.i.d. $N(0,1)/\sqrt{n}$ (top row) and i.i.d. Ber$(3/n)$ (bottom row).

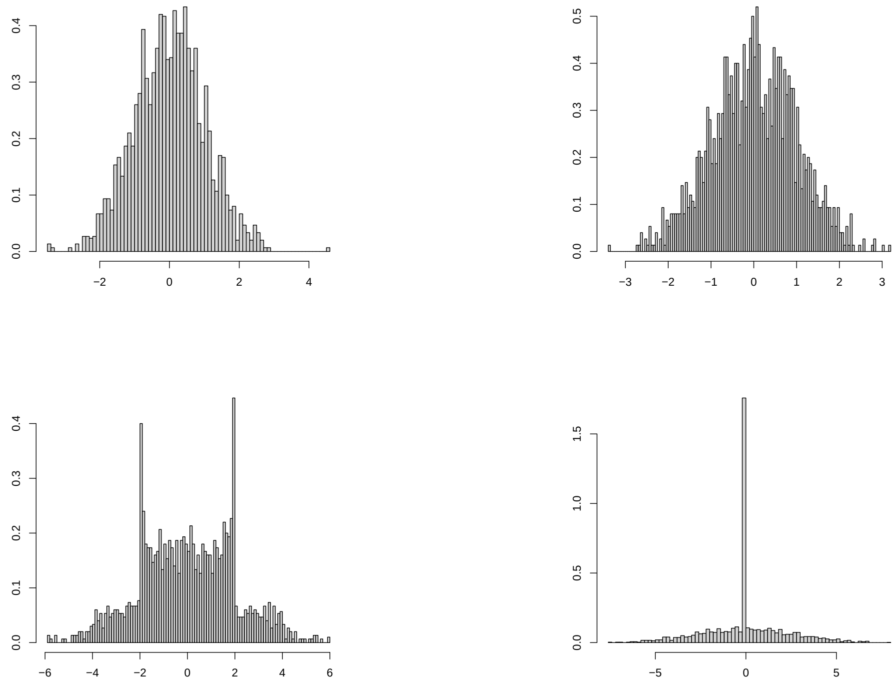


FIGURE 4.2: Two replicated histograms of the eigenvalues of $C_n^{(s)}$ for $n = 1000$. The entries are i.i.d. $N(0,1)/\sqrt{n}$ (top row) and i.i.d. $\mathrm{Ber}(3/n)$ (bottom row).

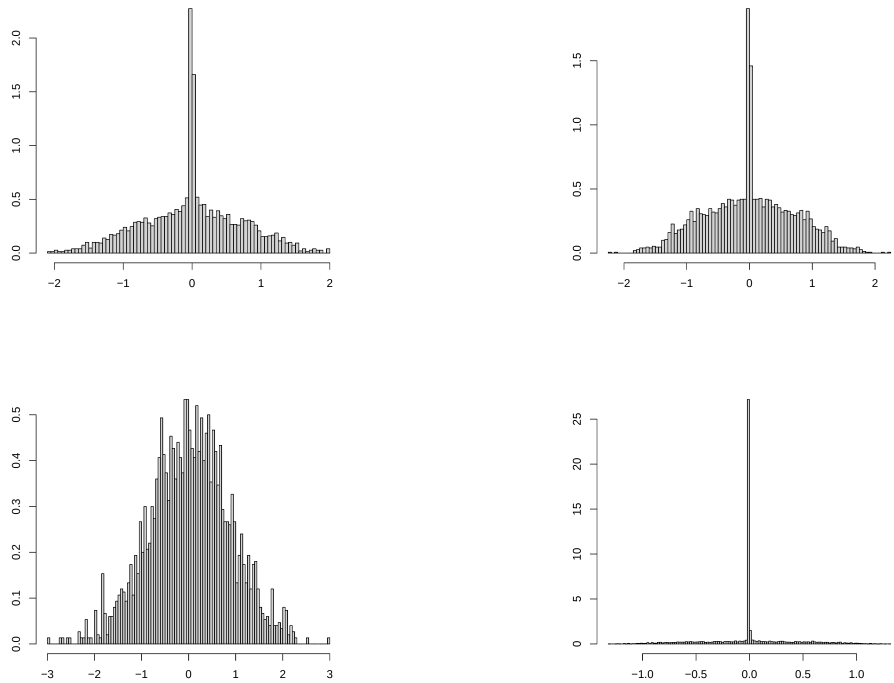


FIGURE 4.3: Two replicated histograms of the eigenvalues for $n = 1000$ of $T_n^{(s)u}$ and $H_n^{(s)u}$ (top row) and $T_n^{(s)b}$ and $H_n^{(s)b}$ with $\alpha = 1/2$ (bottom row). The entries are i.i.d. $N(0,1)/\sqrt{n}$.

Chapter 5

Sample Covariance (S) matrix

Let X_p be a $p \times n$ matrix with real independent entries $\{x_{ij,n} : 1 \leq i \leq p, 1 \leq j \leq n\}$, where $p = p(n), p/n \to y \in (0, \infty)$. The matrix $S = X_p X_p^T$ will be called the *Sample covariance matrix* (without scaling) or the S matrix. In this chapter, we will investigate the empirical spectral distribution of this matrix.

In Section 5.1, we describe a few LSD results that already exist in the literature. These results are closely related to the main result of this chapter, namely Theorem 5.2.1, that is described in Section 5.2. Next, in Section 5.4, we give a proof of Theorem 5.2.1. In Section 5.5, we discuss how the results in Section 5.1 can be concluded from Theorem 5.2.1. We conclude the chapter with some simulations that show the various distributions that can appear as the LSD. *This chapter is based on Bose and Sen [2023].*

5.1 Review of existing results

The S matrix is arguably one of the most important matrices in random matrix theory, with varied applications in physics, statistics and other areas. There have been several works regarding its LSD. When the entries of X_p are i.i.d. with mean zero and finite fourth moment, Marčenko and Pastur [1967] first established the LSD of $\frac{1}{n}S$ and this LSD has been named the Marčenko-Pastur (MP) law. Subsequent works by Grenander and Silverstein [1977], Wachter [1978], Yin [1986], Jonsson [1982], Bai [1999] investigated the existence and properties of the LSD under varied assumptions on the entries. We state the most widely known result in the fully i.i.d. regime. First let us introduce the

MP_y law, where $p/n \to y \in (0, \infty)$. The MP_y law has the following density when $y \leq 1$

$$f_y(x) = \begin{cases} \frac{1}{2\pi x y}\sqrt{(b-x)(x-a)} & \text{if } a \leq x \leq b, \\ 0 & \text{otherwise} \end{cases}$$

where $a = (1-\sqrt{y})^2$ and $b = (1+\sqrt{y})^2$. When $y > 1$, the MP_y law is a mixture of a point mass at 0 and the density $f_{1/y}$ with weights $(1-\frac{1}{y})$ and $\frac{1}{y}$, respectively.

Result 5.1.1. *(**Standardized fully i.i.d. entries**) Suppose the entries $\{x_{ij} : 1 \leq i \leq p, 1 \leq j \leq n\}$ of X_p are i.i.d. with mean 0 and variance 1. Then as $p, n \to \infty$ with $p/n \to y \in (0, \infty)$, the ESD of $\frac{1}{n}S$ converges weakly almost surely to the MP_y law.*

Recall that the even moments of the standard semicircle distribution μ_s are given by $\beta_{2k}(\mu_s) = |NC_2(2k)| = |NC(k)|$. The moments of the MP_y law are related to the set $NC_2(2k)$ and $NC(k)$ (see Bose [2021]). In particular when $y = 1$, it is known that $\beta_k(MP_y) = \beta_{2k}(\mu_s), k \geq 1$.

Heavy-tailed entries: Just like the Wigner case, a natural extension of the above model is made by considering the case where the entries follow a *heavy-tailed* distribution.

Result 5.1.2. *(Belinschi et al. [2009]) Suppose the entries $\{x_{ij}\ ; 1 \leq i \leq p, 1 \leq j \leq n\}$ of X_p are i.i.d. and satisfy $P\{|x_{ij}| > u\} = u^{-\alpha}L(u)$ as $u \to \infty$ where $L(\cdot)$ is slowly varying and $\alpha \in (0, 2)$. Let $a_n = \inf\{u : P[|x_{ij}| > u] \leq 1/n\}$. Then the ESD of $\frac{S}{a_n^2}$ converges to a probability measure μ_α in probability.*

In these works the distribution of the entries of X_p remain unaltered for every n. It is natural to ask what happens when the distribution of the entries depend on n and/or the entries are not identically distributed. We already saw in Chapter 3 that for the Wigner case, an appropriate truncation of the variables at levels that depend on n was crucial in dealing with the heavy-tailed entries. The same strategy is used in the arguments in Belinschi et al. [2009] for dealing with the S matrix with heavy-tailed entries. Thus it becomes relevant to probe the case where the distribution of the entries of X_p is allowed to depend on n, not just due to a scaling constant that depends on n but where a genuine triangular sequence of entries is used.

Size-dependent entries: Recall that such a model referred to as *matrices with exploding moments* was already considered by Zakharevich [2006] for the (symmetric) Wigner matrix. For the S matrix, the matrices X_p with exploding moments have been considered by several authors (Benaych-Georges and Cabanal-Duvillard [2012], Male [2017], Noiry [2018]). We state one of the results below. All the others are similar in spirit, and will be discussed later in Section 5.5.

Result 5.1.3. *(Theorem 3.2 in Benaych-Georges and Cabanal-Duvillard [2012]) Suppose that the entries $\{x_{ij,n}\ ; 1 \leq i \leq p, 1 \leq j \leq n\}$ of X_p are centered i.i.d. for every fixed n and $p = p(n)$. Assume that $p/n \to y > 0$, and there exists a sequence $(d_k)_{k\geq 2}$ with $d_k^{1/k}$ bounded, such that for every $k \geq 2$,*

$$\lim_{n\to\infty} \frac{\mathbb{E}[x_{11,n}^k]}{n^{k/2-1}} = d_k.$$

Then, as $n \to \infty$, the ESD of $\frac{1}{n}S$ converges weakly almost surely to a non-random distribution $\mu_{y,d}$ say, whose moments are determined by $(d_k)_{k\geq 2}$ and y.

When $d_k = 0$ for all $k \geq 3$, $\mu_{y,d}$ is the MP_y law dilated by a coefficient $\sqrt{d_2}$. The formulae for the moments of the LSD have been provided using graphs and *hypergraphs*. Similar results have been studied in Proposition 3.1 of Noiry [2018], and the moments of the LSD have been provided using free probability theory and certain equivalence classes.

Matrices with variance profile: The S matrix, where the entries of X_p are independent but not necessarily identically distributed, have been considered in Yin [1986], Lytova and Pastur [2009], and Bai and Silverstein [2010]. A common theme has been to assume that the entries have equal variances. However, when the matrix X_p has a non-trivial variance profile (Definition 3.5.15 and 3.5.16), Hachem et al. [2006], Zhu [2020], Jin and Xie [2020] have previously studied the distribution of the eigenvalues of the S matrix. We state one of the results below, and discuss the others later in the chapter.

Result 5.1.4. *(Hachem et al. [2006]) Suppose $\sigma : [0,1]^2 \to \mathbb{R}$ is a function such that σ^2 is continuous and bounded. Suppose the entries $\{y_{ij} = \sigma(i/p, j/n)x_{ij}, 1 \leq i \leq p, 1 \leq j \leq n\}$ of X_p are such that $\{x_{ij}, 1 \leq i \leq p, 1 \leq j \leq n\}$ are i.i.d. with mean zero, variance 1 and satisfy*

$$\mathbb{E}[x_{ij}^{4+\epsilon}] < \infty \quad \textit{for some} \quad \epsilon > 0. \tag{5.1.1}$$

Then the ESD of $\frac{1}{n}S$ converges weakly almost surely to a non-random probability distribution.

Triangular matrices: Other available variations of LSD results for the S matrix include the cases where X_p is triangular (Dykema and Haagerup [2004]) or sparse. Dykema and Haagerup [2004] considered upper triangular matrices X_n^u with i.i.d. complex Gaussian entries and studied the LSD of $X_n^{u*}X_n^u$. Let

$$X_n^u = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\ 0 & x_{22} & x_{23} & \cdots & x_{2n} \\ & & & \vdots & \\ 0 & 0 & 0 & \cdots & x_{nn} \end{bmatrix}. \tag{5.1.2}$$

Result 5.1.5. *(Dykema and Haagerup [2004]) Suppose the non-zero entries $\{x_{ij,n}, i \leq j, 1 \leq i, j \leq n\}$ of X_n^u are i.i.d. Gaussian with mean zero and variance 1. Then the ESD of $\frac{1}{n}S^u = \frac{1}{n}X_n^u X_n^{uT}$ converges weakly almost surely to a non-random probability distribution.*

5.2 Main Results

In Theorem 5.2.1, we establish a general LSD result for the S matrix, where the distribution of any entry is allowed to be dependent not only on n, but also on its position in the matrix. We describe a formula for the moments of the LSD using special symmetric partitions. We relate these moments to the moment formulae that have appeared in the results mentioned above, and also to the limiting moments in the Wigner case. Under our assumptions, only the class of special symmetric partitions defined in Section 3.2.1 contribute to the moments.

Consider the matrix $S = X_pX_p^T$, where the entries of X_p are given by the bi-sequence $\{x_{ij,n}\}$. We drop the suffix n and p for convenience wherever there is no scope for confusion. For any real-valued function g on $[0,\ 1]$, $\|g\| := \sup_{0\leq x\leq 1}|g(x)|$ will denote its sup norm. We introduce the following assumptions on the entries $\{x_{ij}\}$.

Assumption A1. There exists a sequence $\{r_n\}$ with $r_n \in [0, \infty]$ such that

(i) For each $k \in \mathbb{N}$,

$$n\, \mathbb{E}\left[x_{ij}^{2k}\mathbf{1}_{\{|x_{ij}|\le r_n\}}\right] = g_{2k,n}\left(\frac{i}{p},\frac{j}{n}\right) \quad \text{for} \quad 1 \le i \le p, 1 \le j \le n, \tag{5.2.1}$$

$$\lim_{n\to\infty} n^{\alpha} \sup_{1\le i\le p, 1\le j\le n} \mathbb{E}\left[x_{ij}^{2k-1}\mathbf{1}_{\{|x_{ij}|\le r_n\}}\right] = 0 \quad \text{for all } \alpha < 1, \tag{5.2.2}$$

where $\{g_{2k,n}\}$ is a sequence of bounded Riemann integrable functions on $[0,1]^2$.

(ii) The functions $g_{2k,n}(\cdot), n \ge 1$ converge uniformly to $g_{2k}(\cdot)$ for each $k \ge 1$.

(iii) With $M_{2k} = \|g_{2k}\|$, $M_{2k-1} = 0$ for all $k \ge 1$, the sequence $\alpha_k = \sum_{\sigma\in\mathcal{P}(2k)} M_\sigma$ satisfies *Carleman's condition*,

$$\sum_{k=1}^{\infty} \alpha_{2k}^{-\frac{1}{2k}} = \infty.$$

All of these conditions are naturally satisfied by well-known models such as, the *standard i.i.d.* case, where the entries of X are $\frac{x_{ij}}{\sqrt{n}}$ with $\{x_{ij}\}$ being i.i.d. with zero mean and finite variance, and the *sparse* case, where entries of X are i.i.d, $Ber(p_n)$ with $np_n \to \lambda > 0$, for every n, etc. (more details in Section 5.5). Now we state our theorem.

Theorem 5.2.1. *Let X be a $p \times n$ real matrix with independent entries $\{x_{ij,n}; 1 \le i \le p, 1 \le j \le n\}$ that satisfy Assumption A1, and $p/n \to y \in (0,\infty)$ as $n \to \infty$. Suppose Z is a $p \times n$ real matrix whose entries are $y_{ij} = x_{ij}\mathbf{1}_{[|x_{ij}|\le r_n]}$. Then the following hold.*

(a) The ESD of $S_Z = ZZ^T$ converges weakly almost surely to a probability measure μ, whose moments are determined by integrals of $\{g_{2k}\}_{k\ge1}$ as described in (5.4.8).

(b) Moreover, if

$$\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{n} x_{ij}^2\mathbf{1}_{\{|x_{ij}|>r_n\}} \to 0, \ \textit{almost surely (or in probability)}, \tag{5.2.3}$$

then the ESD of $S = XX^T$ converges almost surely (or in probability) to the probability measure μ given in (a).

Note that as mentioned Remark 3.3.4, the convergence in (5.2.3) can occur in probability and not almost surely in certain scenarios.

Remark 5.2.2. *In Result 5.1.1, the MP_y law appears as the LSD of $\frac{1}{n}S$ with fully i.i.d.entries and it has bounded support. However, that is not necessarily the case for μ in Theorem 5.2.1. Suppose the entries of X_p satisfy Assumption A1. Let for every $m \geq 1$, $f_{2m}(x) = \int_{[0,1]} g_{2m}(x,y)\, dy$. Now if there exist an $m > 1$ such that $\inf\limits_{t\geq 1} \int_{[0,1]} \left(\frac{f_{2m}(x)}{m!}\right)^t dx = c > 0$, then the LSD μ in Theorem 5.2.1 has unbounded support.*

This has implications on the partition description of moments. As is known, the moments of MP_y can be described via the set of non-crossing pair partitions. In the present case, these partitions are not enough to describe μ, and we need special symmetric partitions described in Definition 3.2.1.

Remark 5.2.3. *It is known that if Y follows the MP_1 law and Y' follows the semi-circle law, then $Y \stackrel{\mathcal{D}}{=} Y'^2$. A similar result holds for μ. Suppose $p/n \to 1$, and the entries of X satisfy Assumption A1 and (5.2.3). Then the ESD of S converges almost surely to a probability distribution μ as given in Theorem 5.2.1. At the same time, consider the Wigner matrix (i.e., a symmetric matrix) with independent entries $\{x_{ij,n}; 1 \leq i \leq j \leq n\}$ that satisfy the conditions of Assumption A1 (see Section 3.3). Then by Theorem 3.3.1, the ESD of W_n converges almost surely to a symmetric probability measure μ'. The two measures μ and μ' are connected. Suppose Y and Y' are two random variables such that $Y \sim \mu$ and $Y' \sim \mu'$. If $\{g_{2k}\}_{k\geq 1}$ are symmetric functions, then $Y \stackrel{\mathcal{D}}{=} Y'^2$.*

Remark 5.2.4. *Consider the matrix*

$$\begin{bmatrix} 0 & X \\ X^T & 0 \end{bmatrix} \tag{5.2.4}$$

whose eigenvalues are $\pm\sqrt{\lambda_i}$, where λ_i are the eigenvalues of $S = XX^T$. As this matrix is symmetric, choosing $g_{2k,n} = 0$ on two appropriate rectangles, the matrix falls under the regime of Chapter 3. Hence with adequate adjustments the LSD of S can be found using Theorem 3.3.1.

Using (5.2.4) we can also conclude the convergence of the EESDs of $S_A = AA^T$ matrix where A is one of the matrices $T^{(s)}, H^{(s)}, R^{(s)}, C^{(s)}, T, H, R$ and C. However, some information like, which words actually contribute positively to the limiting moments and which don't, is lost in this process (see also Remark 6.1.3). In this case using the methods described in Chapter 2 and some extensions of these methods, we are able to *provide more explicit expressions of the limiting moments and hence better understand the limit.*

Here, we have given an independent proof for the LSD of S. We have extended some of the notions like generating vertices, their free choice and the set $\Pi(\boldsymbol{\omega})$ to even and odd generating vertices, their free choices and $\Pi_S(\boldsymbol{\omega})$ in Section 5.3. With the help of these notions, we prove the convergence of the ESD of S and describe its limiting moments. These proofs (Section 5.4) help us understand the proofs of the theorems in Chapter 6 in a better manner.

Further, the equivalence of $SS(2k)$ with hypergraphs and Noiry words as described in Section 5.5.7 bear significance in emphasizing the various combinatorial structures that arise from this set of partitions.

5.3 Some preliminaries

To prove Theorem 5.2.1, we need to extend the notion of circuits and words that were given in Section 2.4 for single symmetric matrices.

First recall that the link function for S is given by a pair of functions as follows.

$$L_1(i,j) = (i,j) \quad \text{and} \quad L_2(i,j) = (j,i).$$

Circuits and Words: In case of the S matrix, a **circuit** π is a function $\pi : \{0, 1, 2, \ldots, 2m\} \to \{1, 2, 3, \ldots, \max(p, n)\}$ with $\pi(0) = \pi(2m)$ and $1 \le \pi(2i) \le p, 1 \le \pi(2i-1) \le n$ for $1 \le i \le m$. We say that the *length* of π is $2m$ and denote it by $\ell(\pi)$. Next, let

$$\xi_\pi(2i-1) = L_1(\pi(2i-2), \pi(2i-1)), 1 \le i \le k$$
$$\xi_\pi(2i) = L_2(\pi(2i-1), \pi(2i)), 1 \le i \le k.$$

Then with $Y_\pi = \prod_{i=1}^{k} y_{\xi_{\pi(2i-1)}} y_{\xi_{\pi(2i)}}$,

$$\begin{aligned}
\mathbb{E}\big[\operatorname{Tr}(S^k)\big] &= \mathbb{E}\big[\operatorname{Tr}(XX^*)^k\big] \\
&= \sum_{\pi:\ell(\pi)=2k} x_{L_1(\pi(0),\pi(1))} x_{L_2(\pi(1),\pi(2))} \cdots x_{L_2(\pi(2k-1),\pi(2k))} \\
&= \sum_{\pi:\ell(\pi)=2k} \mathbb{E}[Y_\pi]. \qquad (5.3.1)
\end{aligned}$$

From (5.3.1), note that the kth moment of an entry of S involves the $2k$th moment of the entries of X. Hence the kth moment of the ESD of S involves circuits of length $2k$.

For any π, the values $L_t(\pi(i-1), \pi(i)), t = 1, 2$ will be called *edges* or L-values. When an edge appears more than once in a circuit π, then it is called *matched*. Any m circuits $\pi_1, \pi_2, \ldots, \pi_m$ are said to be *jointly-matched* if each edge occurs at least twice across all circuits. They are said to be *cross-matched* if each circuit has an edge which occurs in at least one of the other circuits. Circuits π_1 and π_2 are said to be *equivalent* if

$$L_t(\pi_1(i-1), \pi_1(i)) = L_t(\pi_1(j-1), \pi_1(j)) \iff L_t(\pi_2(i-1), \pi_2(i)) = L_t(\pi_2(j-1), \pi_2(j)), t = 1, 2.$$

The above is an equivalence relation on $\{\pi : \ell(\pi) = 2k\}$. Any equivalence class of circuits can be indexed by an element of $\mathcal{P}(2k)$. The positions where the edges match are identified by each block of a partition of $[2k]$. Also, an element of $\mathcal{P}(2k)$ can be identified with a *word* of length k of letters (see Section 2.4).

The class $\Pi_S(\boldsymbol{\omega})$: For a given word $\boldsymbol{\omega}$, this is the set of all circuits which correspond to $\boldsymbol{\omega}$. For any word $\boldsymbol{\omega}$, $\boldsymbol{\omega}[i] = \boldsymbol{\omega}[j] \;\Leftrightarrow \xi_\pi(i) = \xi_\pi(j)$. This implies that,

$$\begin{aligned}
&L_t(\pi(i-1), \pi(i)) = L_t(\pi(j-1), \pi(j)) \quad \text{if } i \text{ and } j \text{ are of same parity}, \; t = 1, 2, \\
&L_t(\pi(i-1), \pi(i)) = L_t'(\pi(j-1), \pi(j)) \quad \text{if } i \text{ and } j \text{ are of different parity}, \; t, t' \in \{1, 2\}, t \neq t'.
\end{aligned}$$

Therefore the class $\Pi_S(\boldsymbol{\omega})$ is given as follows:

$$\begin{aligned}\Pi_S(\boldsymbol{\omega}) &= \{\pi; \boldsymbol{\omega}[i] = \boldsymbol{\omega}[j] \ \Leftrightarrow \xi_\pi(i) = \xi_\pi(j)\} \\ &= \Big\{\pi : \boldsymbol{\omega}[i] = \boldsymbol{\omega}[j] \Leftrightarrow (\pi(i-1), \pi(i)) = (\pi(j-1), \pi(j)) \\ &\qquad \text{or } (\pi(i-1), \pi(i)) = (\pi(j), \pi(j-1))\Big\}. \end{aligned} \tag{5.3.2}$$

From (5.3.1) observe that,

$$\lim_{p\to\infty} \frac{1}{n}\mathbb{E}[\mathrm{Tr}(S^k] = \lim_{n\to\infty} \frac{1}{p} \sum_{\pi:\ell(\pi)=2k} \mathbb{E}[Y_\pi] = \lim_{p\to\infty} \sum_{b=1}^{k} \sum_{\substack{\omega \text{ matched of length 2k} \\ \text{with b distinct letters}}} \frac{1}{p} \sum_{\pi\in\Pi_S(\boldsymbol{\omega})} \mathbb{E}(Y_\pi). \tag{5.3.3}$$

Note that all words that appear above are of length $2k$. For every $k \geq 1$, the words of length $2k$, corresponding to the circuits of S and the Wigner matrix W, are related (see Observation 1 below). We will later find a connection between the kth moment of the LSD of S and the $2k$th moment of the LSD of the Wigner matrix.

Recall the link function $L_W(i,j) = (\min(i,j), \max(i,j))$ of the Wigner matrix. For words corresponding to L_W, the class $\Pi_W(\boldsymbol{\omega})$ is given by

$$\begin{aligned}\Pi_W(\boldsymbol{\omega}) &= \Big\{\pi : \boldsymbol{\omega}[i] = \boldsymbol{\omega}[j] \Leftrightarrow L_W(\pi(i-1), \pi(i)) = L_W(\pi(j-1), \pi(j))\Big\} \\ &= \Big\{\pi : \boldsymbol{\omega}[i] = \boldsymbol{\omega}[j] \Leftrightarrow (\pi(i-1), \pi(i)) = (\pi(j-1), \pi(j)) \\ &\qquad \text{or } (\pi(i-1), \pi(i)) = (\pi(j), \pi(j-1))\Big\}. \end{aligned} \tag{5.3.4}$$

Next, we make a key observation about $\Pi_S(\boldsymbol{\omega})$ and $\Pi_W(\boldsymbol{\omega})$.

Observation 1: Let $\tilde{\Pi}_W(\boldsymbol{\omega})$ be the possibly larger class of circuits for the Wigner Link function with range $1 \leq \pi(i) \leq \max(p,n), 0 \leq i \leq 2k$. Then for a word $\boldsymbol{\omega}$ of length $2k$,

$$\Pi_S(\boldsymbol{\omega}) \subset \tilde{\Pi}_W(\boldsymbol{\omega}). \tag{5.3.5}$$

The definition of *generating and non-generating vertices* in this case remains same as those in Section 2.4. However, since we have $1 \leq \pi(2i) \leq p$ and $1 \leq \pi(2i-1) \leq n$ for every $1 \leq i \leq k$, we now need the notion of *even and odd generating vertices* to distinguish between the two types of vertices.

Even and odd generating vertices: A generating vertex $\pi(i)$ is called even (odd) if i is even (odd). Any word has at least one of each, namely $\pi(0)$ and $\pi(1)$. So for a matched word with b ($\leq k/2$) distinct letters, there can be $(r+1)$ even generating vertices where $0 \leq r \leq b-1$. Observe that

$$\begin{aligned}\big|\Pi_S(\boldsymbol{\omega})\big| = \big|\big\{\big(\pi(0), \pi(1), \ldots, \pi(2k)\big) : 1 \leq \pi(2i) \leq p, 1 \leq \pi(2i-1) \leq n \text{ for } i = 0, 1, \ldots, k,\\ \pi(0) = \pi(2k), \quad \xi_\pi(i) = \xi_\pi(j) \text{ if and only if } \boldsymbol{\omega}[i] = \boldsymbol{\omega}[j]\big\}\big|. \qquad (5.3.6)\end{aligned}$$

Circuits corresponding to a word $\boldsymbol{\omega}$ are completely determined by the generating vertices. The vertex $\pi(0)$ is always generating, and there is one generating vertex for each new letter in $\boldsymbol{\omega}$. So, if $\boldsymbol{\omega}$ has b distinct letters, then it has $(b+1)$ generating vertices. Hence the growth of $|\Pi_S(\boldsymbol{\omega})|$ is determined by the number of generating vertices that can be chosen freely. As $p/n \to y > 0$,

$$|\Pi_S(\boldsymbol{\omega})| = \mathcal{O}(p^{r+1}n^{b-r}) \text{ whenever } \omega \text{ has } b \text{ distinct letters, and } (r+1) \text{ even generating vertices.} \qquad (5.3.7)$$

As in Chapters 3 and 4, the existence of

$$\lim_{p,n\to\infty} \frac{|\Pi_S(\boldsymbol{\omega})|}{p^{r+1}n^{b-r}} \qquad (5.3.8)$$

for every word $\boldsymbol{\omega}$ is tied very intimately to the LSD of S.

5.4 Proof of Theorem 5.2.1

We present a few lemmas that lead to the proof of Theorem 5.2.1. In Lemmas 5.4.1 and 5.4.2, we identify the words that possibly contribute to the limiting moments. Lemma 5.4.3 helps us tackle the almost sure convergence of the ESD. Lemma 5.4.4 helps us in the truncation as well as in the reduction to the case where all entries are centered. Finally we prove Theorem 5.2.1.

Identification of words that may contribute: As discussed in Section 5.3, the existence of $\lim_{n\to\infty} \frac{|\Pi_S(\boldsymbol{\omega})|}{p^{r+1}n^{b-r}}$ is crucial in identifying the words that contribute to the moments of the LSD. So we look into it first.

Lemma 5.4.1. *Suppose $\boldsymbol{\omega} \in SS_b(2k)$ with $(r+1)$ even generating vertices $(0 \le r \le k)$. Then, $\left|\Pi_S(\boldsymbol{\omega})\right| \approx p^{r+1}n^{b-r}$, where $a \approx b$ means that $a/b \to 1$.*

Proof. This result is already known for $b = k$, i.e., for pair matched words (see Bose [2018]). This lemma can be proved using the arguments from their result and the proof of Lemma 3.4.2. We give a sketch of the proof here.

We argue by induction on b, the number of distinct letters. If $b = 1$, then $r = 0$ and $\boldsymbol{\omega} = aa\cdots aa$. Therefore $\pi(0)$ and $\pi(1)$ are the generating vertices, and both can be chosen freely. Thus, $\left|\Pi_S(\boldsymbol{\omega})\right| \approx pn$. Now assume that the result is true upto $b-1$. Then it is enough to prove that if $\boldsymbol{\omega}$ has b distinct letters with $(r+1)$ $(0 \le r \le b-1)$ even generating vertices, then $\left|\Pi_S(\boldsymbol{\omega})\right| \approx p^{r+1}n^{b-r}$.

First let $0 \le r \le b-2$. Suppose the last distinct letter of $\boldsymbol{\omega}$, say, z appears for the first time at the ith position, that is at $(\pi(i-1), \pi(i))$ or $(\pi(i), \pi(i-1))$ (depending on whether i is odd or even). Then z appears in *pure even blocks* (see Lemma 3.2.3). Let m (m even) be the length of the first pure block. Then it can be shown that if we drop the zs from $\boldsymbol{\omega}$, we get a special symmetric word $\boldsymbol{\omega}'$ be the word with $(b-1)$ distinct letters and $(r+1)$ even generating vertices. Therefore, by induction hypothesis, $\left|\Pi_S(\boldsymbol{\omega}')\right| \approx p^{r+1}n^{b-(r+1)}$. Now as $\pi(i)$ is another odd vertex that can be chosen freely, we have $\left|\Pi_S(\boldsymbol{\omega})\right| \approx p^{r+1}n^{b-(r+1)}n = p^{r+1}n^{b-r}$.

Now let $r = b-1$. Then there are $r+1 = b$ even generating vertices (one of them being $\pi(0)$) and b distinct letters in $\boldsymbol{\omega}$. Therefore all letters, except the first, appear for the first time at even positions in $\boldsymbol{\omega}$. So, if z is the last distinct letter of $\boldsymbol{\omega}$, then z appears for the first time at $(\pi(i-1), \pi(i))$ where i is even. If we drop all zs as before from $\boldsymbol{\omega}$, then we get a special symmetric word $\boldsymbol{\omega}'$ with $(b-1)$ distinct letters and $(b-2)$ even generating vertices. Therefore, $\left|\Pi_S(\boldsymbol{\omega}')\right| \approx p^{b-1}n^{b-(b-1)}$. As $\pi(i)$ is another even vertex that can be chosen freely, we have $\left|\Pi_S(\boldsymbol{\omega})\right| \approx p^{b-1}np = p^b n = p^{r+1}n^{b-r}$, $r = b-1$.

This completes the proof of the lemma. $\square$

Lemma 5.4.2. *Let $\boldsymbol{\omega}$ be a word with b distinct letters and $(r+1)$ even generating vertices, $0 \leq r \leq b-1$. Then*

$$\lim_{n\to\infty} \frac{|\Pi_S(\boldsymbol{\omega})|}{n^{b+1}} = \begin{cases} y^{r+1} & \text{if } \omega \in SS_b(2k), \\ 0 & \text{if } \omega \notin SS_b(2k). \end{cases} \tag{5.4.1}$$

Thus, $\lim_{n\to\infty} \frac{|\Pi_S(\boldsymbol{\omega})|}{p^{r+1}n^{b-r}} = 1$ if and only if $\boldsymbol{\omega}$ is a special symmetric word.

Proof. First suppose $\boldsymbol{\omega} \in \mathcal{P}(2k) \setminus SS_b(2k)$. Then from (5.3.5), Lemma 3.4.2 and Lemma 3.4.6, it is easy to see that

$$\lim_{n\to\infty} \frac{|\Pi_S(\boldsymbol{\omega})|}{n^{b+1}} = 0.$$

If $\boldsymbol{\omega} \in SS_b(2k)$, then from Lemma 5.4.1, it immediately follows that $\lim_{n\to\infty} \frac{|\Pi_S(\boldsymbol{\omega})|}{n^{b+1}} = y^{r+1}$. This completes the proof of the lemma. □

Handling almost sure convergence: As was done in Chapter 3, we shall use the moment method to prove our theorem, and hence shall take help of Lemma 2.1.3. To verify the fourth moment condition, consider four circuits $\pi_1, \pi_2, \pi_3, \pi_4$ of length $2k$ each. Just as in Chapter 3, we put a new letter wherever a new edge (or L-value) appears across all the circuits, and define

$$\begin{aligned} Q^b_{k,4} = |\{(\pi_1, \pi_2, \pi_3, \pi_4) : \ell(\pi_i) = 2k, 1 \leq i \leq 4 \text{ jointly- and cross-matched with} \\ b \text{ distinct edges or } b \text{ distinct letters across all } (\pi_i)_{1\leq i\leq 4}\}|. \end{aligned} \tag{5.4.2}$$

Lemma 5.4.3. *There exists a constant C, such that,*

$$Q^b_{k,4} \leq Cn^{b+2} \,. \tag{5.4.3}$$

This was proved for the Wigner link function in Lemma 3.4.7. The arguments in that proof can be used for the S link function here as $1 \leq \pi(2i) \leq p$ and $1 \leq \pi(2i-1) \leq n$, and p and n are comparable for large n. We omit the details.

Lemma 5.4.4. *Suppose $\{x_{ij}; 1 \leq i \leq p, 1 \leq j \leq j \leq n\}$ are independent variables that satisfy Assumption A1 and $y_{ij} = x_{ij}\mathbf{1}_{[|x_{ij}|\leq r_n]}$. Then* *(i)* $\displaystyle\frac{1}{p}\sum_{i,j}(y_{ij}^2 - \mathbb{E}[y_{ij}^2]) \to 0$ *almost surely as* $p \to \infty$.

(ii) Additionally suppose $\frac{1}{p}\displaystyle\sum_{i,j} x_{ij}^2 \mathbf{1}_{[|x_{ij}|>r_n]} \to 0$ *almost surely (or in probability). Then* $\limsup_p \displaystyle\frac{1}{p}\sum_{i,j} x_{ij}^2 < \infty$ *almost surely (or in probability).*

Proof. (i) Let $\epsilon > 0$ be fixed. Then

$$\mathbb{P}\Big[\Big|\frac{1}{p}\sum_{i,j}(y_{ij}^2 - \mathbb{E}[y_{ij}]^2)\Big| > \epsilon\Big] \le \frac{1}{\epsilon^4 p^4}\mathbb{E}\Big[\big(\sum_{i,j} y_{ij}^2 - \mathbb{E}[y_{ij}^2]\big)\big)^4\Big]$$
$$= \frac{1}{\epsilon^4 p^4}\mathbb{E}\Big[\sum_{\substack{i_1,i_2,i_3,i_4\\ j_1,j_2,j_3,j_4}}\prod_{l=1}^{4}(y_{i_l j_l}^2 - \mathbb{E}[y_{i_l j_l}^2])\big)^4\Big].$$

As $\{y_{ij}\}$ are independent, the above inequality becomes

$$\mathbb{P}\Big[\Big|\frac{1}{p}\sum_{i,j}(y_{ij}^2 - \mathbb{E}[y_{ij}]^2)\Big| > \epsilon\Big] \le \frac{1}{\epsilon^4 p^4}\sum_{ij}\mathbb{E}\big[(y_{ij}^2 - \mathbb{E}[y_{ij}^2])^4\big] +$$
$$6\frac{1}{\epsilon^4 p^4}\sum_{\substack{i_1,i_2\\ j_1,j_2}}\mathbb{E}\big[(y_{i_1 j_1}^2 - \mathbb{E}[y_{i_1 j_1}^2])^2(y_{i_2 j_2}^2 - \mathbb{E}[y_{i_2 j_2}^2])^2\big].$$

Now from (5.2.1), as $\{g_{2k,n}\}$ are bounded integrable, the first term in the rhs of the above inequality is $\mathcal{O}(\frac{1}{p^3})$ and the second term is $\mathcal{O}(\frac{1}{p^2})$. Therefore,

$$\sum_p \mathbb{P}\Big[\Big|\frac{1}{p}\sum_{i,j}(y_{ij}^2 - \mathbb{E}[y_{ij}]^2)\Big| > \epsilon\Big] < \infty.$$

Hence by Borel-Cantelli lemma, $\displaystyle\frac{1}{p}\sum_{i,j}(y_{ij}^2 - \mathbb{E}[y_{ij}^2]) \to 0$ almost surely as $p \to \infty$.

(ii) Observe that $\displaystyle\sum_{i,j} x_{ij}^2 = \sum_{i,j}\big(y_{ij}^2 + x_{ij}^2\mathbf{1}_{[|x_{ij}|>r_n]}\big)$. Also note that $\frac{1}{p}\displaystyle\sum_{i,j}\mathbb{E}[y_{ij}]^2 \to \int g_2(x,y)\,dx\,dy$ as $n, p \to \infty$. Then by the condition $\frac{1}{p}\displaystyle\sum_{i,j} x_{ij}^2\mathbf{1}_{[|x_{ij}|>r_n]} \to 0$ almost surely (or in probability) and (i), (ii) hold true. □

Proof of Theorem 5.2.1. We break the proof into five steps.

Step 1 (Reduction to mean zero): Consider the zero mean matrix $\widetilde{Z} = ((y_{ij} - \mathbb{E}y_{ij}))$. Using the same arguments as in Step 1 of the proof of Theorem 3.3.1, it follows that Conditions (5.2.1) and (5.2.2) hold for $\widetilde{Z}$. Thus, Assumption A1 holds for the matrix $\widetilde{Z}$.

Now from Lemma 2.3.1,

$$\begin{aligned} L^4(F^{S_Z}, F^{S_{\tilde{Z}}}) &\leq \frac{2}{p^2}(\mathrm{Tr}(ZZ^T + \widetilde{Z}\widetilde{Z}^T))(\mathrm{Tr}[(Z-\widetilde{Z})(Z-\widetilde{Z})^T]) \\ &\leq \frac{2}{p}\Big(\sum_{i,j}(2y_{ij}^2 + (\mathbb{E}y_{ij})^2 - 2y_{ij}\mathbb{E}y_{ij})\Big)\Big(\frac{1}{p}\sum_{i,j}(\mathbb{E}y_{ij})^2\Big). \end{aligned} \tag{5.4.4}$$

The second factor of the rhs in (5.4.4) is bounded by

$$n(\sup_{i,j}\mathbb{E}y_{ij})^2 = (\sup_{i,j}\sqrt{n}\mathbb{E}y_{ij})^2 \to 0 \quad \text{as } n \to \infty \quad \text{by (5.2.2)}.$$

By Lemma 5.4.4, $\frac{1}{p}\sum_{i,j}(y_{ij}^2 - \mathbb{E}[y_{ij}^2]) \to 0$ almost surely as $p \to \infty$. Also $\mathbb{E}\big[\frac{1}{p}\sum_{ij} y_{ij}^2\big] \to \int_{[0,1]^2} g_2(x,y)\, dx\, dy$. Hence,

$$\mathbb{P}\big[\{\omega : \limsup_p \frac{1}{p}\sum_{i,j} y_{ij}^2(\omega) = \infty\}\big] = 0.$$

Therefore the first term of the rhs in (5.4.4) also tends to zero almost surely. Hence, the LSD of S_Z and $S_{\widetilde{Z}}$ are same almost surely. Thus we can assume that the entries of Z have mean 0.

We will now verify the Conditions (i), (ii) and (iii) of Lemma 2.1.3.

Step 2 (Verification of the first moment condition): First note that, we can write (5.3.1) as

$$\begin{aligned} \lim_{n\to\infty}\frac{1}{p}\mathbb{E}[\mathrm{Tr}(ZZ^T)^k] &= \lim_{n\to\infty}\sum_{b=1}^{k}\Big[\frac{1}{p}\sum_{\omega \in SS_b(2k)}\sum_{\pi\in\Pi(\boldsymbol{\omega})}\mathbb{E}(Y_\pi) + \frac{1}{n}\sum_{\substack{\omega\notin SS(2k) \\ \boldsymbol{\omega} \text{ with b letters}}}\sum_{\pi\in\Pi(\boldsymbol{\omega})}\mathbb{E}(Y_\pi)\Big]. \\ &= T_1 + T_2. \end{aligned} \tag{5.4.5}$$

Suppose that $\boldsymbol{\omega}$ has b distinct letters and let $\pi \in \Pi_S(\boldsymbol{\omega})$. Suppose the jth new letter appears at the $(\pi(i_j-1), \pi(i_j))$th position for the first time, $1 \leq j \leq b$. Let $\mathcal{D}$ denote the set of all distinct generating vertices. Thus $|\mathcal{D}| \leq (b+1)$. (Everywhere else in the thesis, the set of generating vertices is denoted by S. In this chapter we denote it by $\mathcal{D}$ to avoid confusion with the S matrix notation.)

Suppose $\boldsymbol{\omega}$ has b distinct letters but does not belong to $SS(2k)$. Then from Lemma 5.4.2, $|\mathcal{D}| \leq b$. Hence $\boldsymbol{\omega}$, and as a consequence, T_2 has no contribution to (5.4.5).

Now suppose $\boldsymbol{\omega} \in SS_b(2k)$ with $(r+1)$ even generating vertices. By Lemma 5.4.2, all generating vertices of $\boldsymbol{\omega}$ are distinct. For each $j \in \{1, 2, \ldots, b\}$ denote $(\pi(i_j - 1), \pi(i_j))$ as (t_j, l_j). Then $t_1 = \pi(0)$ and $l_1 = \pi(1)$. It is easy to see that any distinct (t_j, l_j) corresponds to a distinct letter in $\boldsymbol{\omega}$. Suppose the jth new letter appears s_j times in $\boldsymbol{\omega}$. Clearly all the s_j are even. So the total contribution of this $\boldsymbol{\omega}$ to T_1 in (5.4.5) is:

$$\frac{1}{pn^b} \sum_{\mathcal{D}} \prod_{j=1}^{b} g_{s_j,n}(t_j/p, l_j/n). \tag{5.4.6}$$

Recall that there are $(r+1)$ even generating vertices in $\mathcal{D}$ (the set of all distinct generating vertices) with range between 1 and p, and $(b-r)$ odd generating vertices with range between 1 and n. So as $n \to \infty$, (5.4.6) converges to

$$y^r \int_{[0,1]^{b+1}} \prod_{j=1}^{b} g_{k_j}(x_{t_j}, x_{l_j}) \prod_{i \in \mathcal{D}} dx_i. \tag{5.4.7}$$

Hence we obtain

$$\lim_{p\to\infty} \frac{1}{p}\mathbb{E}[\operatorname{Tr} S^k] = \sum_{b=1}^{k} \sum_{r=0}^{b-1} \sum_{\substack{\pi \in SS_b(2k) \\ \text{with } (r+1) \\ \text{even generating vertices}}} y^r \int_{[0,1]^{b+1}} \prod_{j=1}^{b} g_{k_j}(x_{t_j}, x_{l_j}) \prod_{i \in \mathcal{D}} dx_i. \tag{5.4.8}$$

This completes the verification of the first moment condition.

Step 3 (Uniqueness of the measure and convergence of the EESD): We have obtained

$$\gamma_k = \lim_{p\to\infty} \frac{1}{p}\mathbb{E}[\operatorname{Tr}(S)^k] \leq \sum_{b=1}^{k} \sum_{r=0}^{b-1} \sum_{\substack{\sigma \in SS_b(2k) \\ \text{with } (r+1) \\ \text{even generating vertices}}} y^r M_\sigma.$$

Let $c = \max(y, 1)$. Then

$$\gamma_k \leq \sum_{\sigma \in SS(2k)} c^k M_\sigma \leq \sum_{\sigma \in \mathcal{P}(2k)} c^k M_\sigma = c^k \alpha_k.$$

As $\{\alpha_k\}$ satisfies Carleman's condition, $\{\gamma_k\}$ also does so. By Lemma 2.1.3, there exists a measure μ with moments $\{\gamma_k\}_{k\geq 1}$ such that the EESD of S_Z converges weakly to μ.

Step 4 (Verification of the fourth moment condition for S_Z): Observe that

$$\frac{1}{p^4}\mathbb{E}[\mathrm{Tr}(ZZ^T)^k \ - \ \mathbb{E}(\mathrm{Tr}(ZZ^T)^k)]^4 \ = \frac{1}{p^4} \sum_{\pi_1,\pi_2,\pi_3,\pi_4} \mathbb{E}[\Pi_{i=1}^4 (Y_{\pi_i} \ - \ \mathbb{E}Y_{\pi_i})]. \tag{5.4.9}$$

By (5.3.5) and Lemma 5.4.3, using the same argument as in Step 2 of the proof of Theorem 3.3.1, we find that for some constant $B > 0$,

$$\frac{1}{p^4}\mathbb{E}[\mathrm{Tr}(ZZ^T)^k \ - \ \mathbb{E}(\mathrm{Tr}(ZZ^T)^k)]^4 \ \leq BM_0' \sum_{b=1}^{4k} \frac{1}{p^{b+3\frac{1}{2}}} p^{b+2} = \mathcal{O}(p^{-\frac{3}{2}}).$$

Thus the fourth moment condition is established for S_Z.

Hence by Lemma 2.1.3, we conclude that the ESD of S_Z converges weakly almost surely to μ. This completes the proof of Part (a).

Step 5 (Proof of Part (b)): From Lemma 2.3.1, we have

$$\begin{aligned} L^4(F^S, F^{S_Z}) &\leq \frac{2}{p^2}(\mathrm{Tr}(XX^T + ZZ^T))(\mathrm{Tr}[(X-Z)(X-Z)^T]) \\ &= \frac{2}{p}\Big(2\sum_{i,j} y_{ij}^2 + \sum_{ij} x_{ij}^2 \mathbf{1}_{[|x_{ij}|>r_n]}\Big)\Big(\frac{1}{p}\sum_{ij} x_{ij}^2 \mathbf{1}_{[|x_{ij}|>r_n]}\Big). \end{aligned} \tag{5.4.10}$$

The second factor in the above equation tends to zero almost surely (or in probability) as $n \to \infty$ due to condition (5.2.3). Now $\dfrac{1}{p}\displaystyle\sum_{i,j}(y_{ij}^2 - \mathbb{E}[y_{ij}^2]) \to 0$ almost surely (see Lemma 5.4.4) and $\mathbb{E}\big[\frac{1}{p}\sum_{ij} y_{ij}^2\big] \to \int_{[0,1]^2} g_2(x,y)\, dx\, dy$, and hence remains bounded. This implies that $\frac{1}{p}\sum_{i,j} y_{ij}^2$ is bounded almost surely. Therefore the first factor in (5.4.10) is bounded almost surely, and thus the rhs of (5.4.10) tends to 0 as $p, n \to \infty$.

From the discussion above, we infer that the ESD of $S = XX^T$ converges to μ almost surely (or in probability) according as the convergence in (5.2.3) holds almost surely (or in probability). This completes the proof of the theorem. □

Now we will prove Remark 5.2.3 that establishes the relation between the LSDs of the S matrix and the Wigner matrix.

Proof of Remark 5.2.3. Suppose Y follows the MP_1 law and Y' follows the semi-circle law. In the case $p = n$, if $\{g_{2k,n}\}$ are symmetric functions, then the assumption on the entries of X_n are no different from that on the entries of W_n in Theorem 3.3.1, and from (3.4.14), we see that $\mathbb{E}[Y^k] = \mathbb{E}[Y'^{2k}], k \geq 1$. Therefore, by the uniqueness criterion of a probability distribution via moments, we have $Y \stackrel{\mathcal{D}}{=} Y'^2$. From (5.4.8), the limiting moments of the LSD depend on $\{g_{2k}\}$ and $\lim p/n$. So, if $p/n \to 1$, and $\{g_{2k}\}$ are symmetric, we have $Y \stackrel{\mathcal{D}}{=} Y'^2$.

However, observe that even though $\{g_{2k,n}\}$ need not be symmetric for every n, the functions $\{g_{2k}\}$ are symmetric, and hence $\mathbb{E}[Y^k] = \mathbb{E}[Y'^{2k}], k \geq 1$ still holds (see (5.4.8)), as the limiting moments depend only on $\{g_{2k}\}$. Thus the rest of the argument holds as above, and we have $Y \stackrel{\mathcal{D}}{=} Y'^2$. □

The fact that μ has unbounded support follows in the same way as Remark 3.3.3. So we omit the details of the proof of Remark 5.2.2.

5.5 Some Corollaries

In this section we present a few corollaries that deal with S matrices where the entries of the matrix X are— (a) fully i.i.d. with finite mean and variance (Corollary 5.5.1), (b) fully i.i.d. with heavy tails (Corollary 5.5.2), (c) triangular i.i.d. (Corollary 5.5.3), (d) sparse i.e., i.i.d. $Ber(p_n)$ (Corollary 5.5.5) and (e) have non-trivial variance structure (Corollaries 5.5.8 and 5.5.10).

5.5.1 Fully i.i.d. entries

Corollary 5.5.1. *Result 5.1.1 follows from Theorem 5.2.1.*

Proof. Suppose $X = ((x_{ij}/\sqrt{n}))$ where $\{x_{ij}\}$ are i.i.d. with distribution F which has mean zero and variance 1.

First, let us verify that the conditions of Assumption A1 are satisfied in this case. Towards that, let $r_n = n^{-1/3}$. Using the same line of reasoning as in Section 3.5.1, it follows that $g_2 \equiv 1$ and $g_{2k} \equiv 0, k > 1$. Thus $M_2 = 1$, $M_{2k} = 0, k \geq 2$ (see (iii) in Assumption A1) and $\alpha_k = \sum_{\sigma \in \mathcal{P}(2k)} 1$ clearly satisfies Carleman's condition. Now for any $t > 0$,

$$\begin{aligned}
\frac{1}{p}\sum_{i,j} (x_{ij}/\sqrt{n})^2[\mathbf{1}_{[|x_{ij}/\sqrt{n}|>r_n]}] &= \frac{1}{np}\sum_{i,j} x_{ij}^2[\mathbf{1}_{[|x_{ij}|>r_n\sqrt{n}]}] \\
&\leq \frac{1}{np}\sum_{i,j} x_{ij}^2[\mathbf{1}_{[|x_{ij}|>t]}] \quad \text{for all large } n, \\
&\xrightarrow{a.s.} \mathbb{E}\big[x_{11}^2[\mathbf{1}_{[|x_{11}|>t]}]\big].
\end{aligned}$$

As $\mathbb{E}[x_{11}^2] = 1$, taking t to infinity, the above limit is 0 almost surely. Hence by Theorem 5.2.1, the ESD of S converges almost surely to μ whose kth moment is given by

$$\begin{aligned}
\beta_k(\mu) &= \sum_{r=0}^{k-1} \sum_{\substack{\pi \in SS_k(2k) \\ \text{with } (r+1) \\ \text{even generating vertices}}} y^r \\
&= \sum_{r=0}^{k-1} \sum_{\substack{\pi \in NC_2(2k) \\ \text{with } (r+1) \\ \text{even generating vertices}}} y^r. \qquad (5.5.1)
\end{aligned}$$

But this is the kth moment of the MP_y law (see Lemma 8.2.1 and 8.2.2 of Bose [2021]). Hence the ESD of $\frac{1}{n}S$ converges to the MP_y law almost surely. Thus Theorem 5.2.1 yields Result 5.1.1. □

5.5.2 Heavy-tailed entries

Corollary 5.5.2. *Result 5.5.2 follows from Theorem 5.2.1.*

Proof. Suppose $\{x_{ij}, 1 \leq i \leq p, 1 \leq j \leq n\}$ are i.i.d. with an α-stable distribution $(0 < \alpha < 2)$ and $a_p = \inf\Big\{u : \mathbb{P}[|x_{ij}| \geq u] \leq \frac{1}{p}\Big\}$. Suppose $n/p \to \gamma \in (0, 1]$ and $X_p = ((x_{ij}/a_p))$. The existence of LSD of $S = XX^T$ using Stieltjes transform, has been proved in Belinschi et al. [2009]. Theorem 5.2.1 may be used to give the following alternative proof.

Recall that for the Wigner matrix with heavy tailed entries, in the proof of Corollary 3.5.2, we used truncation and moment method to prove the convergence of the ESD. That proof can be easily adapted here. For a fixed constant B, let $X_p^B = ((\frac{x_{ij}}{a_p}\mathbf{1}_{[|x_{ij}|\leq Ba_p]}))$. Then X_p^B satisfies Assumption A1. From Theorem 5.2.1, the ESD of $S^B = X^B(X^B)^T$, almost surely converges to say μ_B. The rest of the arguments are as in the proof of Corollary 3.5.2. Thus μ_S converges to $\tilde{\mu}$ in probability, and we have Result 5.1.2. $\square$

5.5.3 General triangular i.i.d.

The next corollary states an LSD result about general triangular i.i.d matrices. The assumptions are very similar to those of Corollary 3.5.3.

Corollary 5.5.3. *Suppose the entries $\{x_{ij,n}; 1 \leq i \leq p, 1 \leq j \leq n\}$ of X_p is a sequence of i.i.d. random variables with distribution F_n that has finite moments of all orders, for every n. Also assume that*

(i) for every $k \geq 1$,

$$n\beta_k(F_n) \to c_k < \infty, \tag{5.5.2}$$

(ii) $\gamma_k = \sum_{\pi \in SS(2k)} c_\pi$ satisfies Carleman's condition.

Then the ESD of $S = XX^T$ converges weakly almost surely to a non-random probability distribution.

Proof. Let $X = ((x_{ij,n}))$ and $p/n \to y > 0$. Condition (5.5.2) implies that Assumption A1 holds with $r_n = \infty$ and $g_{2k} \equiv c_{2k}, k \geq 1$. Therefore by Theorem 5.2.1, the ESD of S converges weakly almost surely to μ with moments

$$\beta_k(\mu) = \sum_{r=0}^{k-1} \sum_{\substack{\pi \in SS(2k) \\ \text{with } (r+1) \\ \text{even generating vertices}}} y^r c_\pi. \tag{5.5.3}$$

$\square$

Corollary 5.5.4. *Result 5.1.3 follows from Corollary 5.5.3.*

Proof. Suppose the entries of X are $\{\frac{x_{ij,n}}{\sqrt{n\mu_n(2)}}\}$ where $x_{ij,n}$ are i.i.d. with distribution μ_n that has mean zero and all moments finite, and

$$\lim_{n\to\infty} \frac{\beta_k(\mu_n)}{n^{k/2-1}\beta_2(\mu_n)^{k/2}} = d_k \quad \text{say, exists for all} \quad k \geq 1.$$

Clearly Conditions (i) and (ii) of Corollary 5.5.3 are satisfied with $r_n = \infty$, $c_2 \equiv 1$ and $c_k \equiv d_k, k \geq 2$. Hence Corollary 5.5.3 can be applied and the resulting LSD μ, has moments as in (5.5.3). Thus we have Result 5.1.3 from Corollary 5.5.3. □

5.5.4 Sparse S

As we have seen in Corollary 3.5.7 of Section 3.5, a special case of the triangular i.i.d. matrix model is where the entries of X_p are i.i.d. *Bernoulli* distribution with parameter p_n such that $np_n \to \lambda > 0$. The next corollary deals with this case for S.

Corollary 5.5.5. *Suppose the entries of X_p are i.i.d. $Ber(p_n)$ for each n, with $np_n \to \lambda > 0$. Then the ESD of $S = XX^T$ converges weakly almost surely to a non-random probability distribution, say, μ_{ber}, whose moments are as follows:*

$$\beta_k(\mu_{ber}) = \sum_{r=0}^{k-1} \sum_{\substack{\pi\in SS(2k) \\ \text{with } (r+1) \\ \text{even generating vertices}}} y^r \lambda^{|\pi|}. \tag{5.5.4}$$

Proof. Observe that (5.5.2) holds with $c_k \equiv \lambda$ for all $k \geq 1$. Thus X_p in this case satisfies conditions (i) and (ii) of Corollary 5.5.3. Then the result follows immediately from Corollary 5.5.3. □

Remark 5.5.6. (*μ_{ber}* ***and the free Poisson distribution***) *Explicit description of μ_{ber} is not available. However, we can say the following. Recall from Section 2.5, $E(2k)$ and $NCE(2k)$, whose blocks are all of even sizes. It is easily seen that $NCE(2k) \subset SS(2k) \subset E(2k)$. Therefore we have the following:*

Case 1: $y \le 1$. *Then from* (5.5.4)

$$\sum_{\pi \in NCE(2k)} (\lambda y)^{|\pi|} < \beta_k(\mu_{ber}) < \sum_{\pi \in E(2k)} \lambda^{|\pi|}. \tag{5.5.5}$$

Case 2: $y > 1$. *Then from* (5.5.4)

$$\sum_{\pi \in NCE(2k)} y^{|\pi|} < \beta_k(\mu_{ber}) < \sum_{\pi \in E(2k)} (\lambda y)^{|\pi|}. \tag{5.5.6}$$

Now suppose $P_1(\gamma)$ *is a random variable which has the free Poisson distribution with mean* γ*, and* $P_2(\gamma)$ *is a random variable which has the Poisson distribution with mean* γ*. Let* Y *be a random variable which takes value* 1 *and* -1 *with probability* $\frac{1}{2}$ *each. Suppose* Y *is independent of* $P_1(\gamma)$ *and* $P_2(\gamma)$*. Consider* $Q_1(\gamma) = P_1(\gamma)Y$ *and* $Q_2(\gamma) = P_2(\gamma)Y$*. Then the moments of* $Q_1(\gamma)$ *and* $Q_2(\gamma)$ *are given as follows:*

$$\mathbb{E}[Q_1^k(\gamma)] = \begin{cases} 0 & \text{if } k \text{ is odd,} \\ \displaystyle\sum_{\pi \in NCE(k)} \gamma^{|\pi|} & \text{if } k \text{ is even.} \end{cases} \tag{5.5.7}$$

$$\mathbb{E}[Q_2^k(\gamma)] = \begin{cases} 0 & \text{if } k \text{ is odd,} \\ \displaystyle\sum_{\pi \in E(k)} \gamma^{|\pi|} & \text{if } k \text{ is even.} \end{cases} \tag{5.5.8}$$

Hence (5.5.5) *and* (5.5.6)*, can be rewritten as*

$$\mathbb{E}[(Q_1(\lambda y))^{2k}] < \beta_k(\mu_{ber}) < \mathbb{E}[(Q_2(\lambda))^{2k}] \quad \text{for every } k \ge 1,\ y \le 1, \tag{5.5.9}$$

$$\mathbb{E}[(Q_1(\lambda))^{2k}] < \beta_k(\mu_{ber}) < \mathbb{E}[(Q_2(\lambda y))^{2k}] \quad \text{for every } k \ge 1,\ y > 1. \tag{5.5.10}$$

Thus μ_{ber} *lies between the square of a compound free Poisson and the square of a compound Poisson distribution in the above sense.*

5.5.5 Matrices with variance profile

In the next two corollaries, we consider X with a variance profile. Recall Wigner matrices with variance profile, $(W_n, \cdot)$ from Definition 3.5.11. Here the matrices $(X_p, \cdot)$ with *discrete variance profile* and *continuous variance profile* are defined similarly.

Definition 5.5.7. *(a) Discrete variance profile: Suppose* $\{x_{ij,n}; 1 \leq i \leq p, 1 \leq j \leq n\}$ *are i.i.d. random variables with mean zero and variance 1, and let* $\{\sigma_{ij}\}_{1\leq i\leq p, 1\leq j\leq n}$ *be uniformly bounded real numbers. Then the matrix* X_p*, with discrete variance profile* σ_d*, is given by*

$$(X_p, \sigma_d) = \big((y_{ij,n} = \sigma_{ij} x_{ij,n})\big)_{1\leq i\leq p, 1\leq j\leq n}. \tag{5.5.11}$$

(b) Continuous variance profile: Let $\{x_{ij,n}; 1 \leq i \leq p, 1 \leq j \leq n\}$ *be i.i.d. random variables for every fixed* n, p*, and* σ *be a bounded piecewise continuous function on* $[0, 1]^2$*. The matrix* X_p*, with continuous variance profile* σ_c*, is given by*

$$(X_p, \sigma_c) = \big((y_{ij,n} = \sigma(i/p, j/n) x_{ij,n})\big)_{1\leq i\leq p, 1\leq j\leq n}. \tag{5.5.12}$$

First we deal with the discrete variance profile case. Recall Corollary 3.5.12 where we described an LSD result for a Wigner matrix with discrete variance profile. We state a similar result for S. Its proof uses arguments similar to the proof of Corollary 3.5.12 and we omit the details.

Corollary 5.5.8. *(**Discrete variance profile**) Consider the matrix* (X_p, σ_d) *with entries* $\{\frac{y_{ij}}{\sqrt{n}} = \frac{\sigma_{ij} x_{ij}}{\sqrt{n}} : 1 \leq i \leq p, 1 \leq j \leq n\}$ *that are independent and satisfy the following conditions:*

(i) $\mathbb{E}x_{ij} = 0$ *and* $\mathbb{E}[x_{ij}^2] = 1$.

(ii) σ_{ij} *satisfy the following:*

$$\sup_{1\leq i\leq p} \left| \frac{1}{n} \sum_{j=1}^{n} \sigma_{ij}^2 - 1 \right| \to 0 \quad \text{as } n \to \infty. \tag{5.5.13}$$

(iii) $\lim_{n\to\infty} \frac{1}{n^2} \sum_{i,j} \mathbb{E}\big[x_{ij}^2\big] \mathbf{1}_{[|x_{ij}|>\eta\sqrt{n}]}\big] = 0$ *for every* $\eta > 0$.

Then the ESD of $(S, \sigma_d) = (X_p, \sigma_d)(X_p, \sigma_d)^T$ *converges weakly almost surely to the* MP_y *law, where* $0 < y = \lim p/n$.

Remark 5.5.9. *Theorem 1.2 in Jin and Xie [2020] states a similar result where (5.5.13) is replaced by* $\frac{1}{n} \sum_i \left| \frac{1}{n} \sum_{j=1}^{n} {\sigma^2}_{ij} - 1 \right| \to 0$*. However, the proof equation of (2.6) there is not very clear.*

Corollary 5.5.10. ***(Continuous variance profile)*** *Consider the matrix (X_p, σ_c) with entries $\{y_{ij,n}, 1 \le i \le p, 1 \le j \le n\}$ as described in (5.5.12). Assume that the variables $\{x_{ij,n}, 1 \le i \le p, 1 \le j \le n\}$ satisfy Conditions (i) and (ii) of Corollary 5.5.3. Then the ESD of $(S, \sigma_c) = (X_p, \sigma_c)(X_p, \sigma_c)^T$ converges weakly almost surely to a non-random probability measure ν whose kth moment is determined by σ and $\{c_{2m}\}_{1 \le m \le 2k}$.*

Proof. To see this, note that $\{y_{ij,n}\}$ satisfy Assumption A1 with $g_{2k} \equiv \sigma^{2k} c_{2k}$. By Theorem 5.2.1, the ESD of (S, σ_c) converges weakly almost surely to a probability measure ν. From Step 3 in the proof of Theorem 5.2.1, for each word in $SS_b(2k)$ with $(r+1)$ even generating vertices and where the distinct letters appear $s_1, s_2, \ldots, s_b$ times, its contribution to the limiting moments is (see (5.4.7))

$$y^r \int_{[0,1]^{b+1}} \prod_{j=1}^{b} \sigma^{s_j}(x_{t_j}, x_{l_j}) \prod_{i \in S} dx_i \prod_{j=1}^{b} c_{s_j}.$$

Here (t_j, l_j) denotes the position of the first appearance of the jth distinct letter in the word. Hence the kth moment of ν is

$$\beta_k(\nu) = \sum_{b=1}^{k} \sum_{r=0}^{b-1} \sum_{\substack{\pi \in SS_b(2k) \\ \text{with } (r+1) \\ \text{even generating vertices}}} y^r \int_{[0,1]^{b+1}} \prod_{j=1}^{b} \sigma^{s_j}(x_{t_j}, x_{l_j}) dx_i \prod_{j=1}^{b} c_{s_j}.$$

This completes the proof. □

Corollary 5.5.11. *Result 5.1.4 follows from Theorem 5.2.1.*

Proof. In this case, the entries of X_p are $\{\frac{\sigma(i/p,j/n)x_{ij}}{\sqrt{n}}, 1 \le i \le p, 1 \le j \le n\}$ with $\{x_{ij}\}$ being centered i.i.d. variables that have variance 1 and $\mathbb{E}[x_{ij}^{4+\epsilon}] < \infty$ for some $\epsilon > 0$. As σ^2 is a continuous function on $[0,1]^2$, we have $||\sigma|| \le c$, where c is a constant. Now using this fact and the same arguments as in the proof of Corollary 5.5.1, the variables $\{\frac{\sigma(i/p,j/n)x_{ij}}{\sqrt{n}}\}$ satisfy Assumption A1 with $t_n = n^{-1/3}$, $g_2 \equiv \sigma^2$ and $g_{2k} \equiv 0, k \ge 2$. Similarly, (5.2.3) is also satisfied. Hence from Theorem 5.2.1, the ESD of YY^T converges weakly almost surely to a non-random probability measure μ whose moments are determined by σ and y. □

5.5.6 Triangular matrix

Now we look into LSD results for S when X is a triangular matrix.

Corollary 5.5.12. *Suppose the variables $\{x_{ij,n}, 1 \leq i \leq p, 1 \leq j \leq n\}$ associated with the matrices X_n^u as described in (5.1.2) are i.i.d. with all moments finite for every fixed n, and satisfy Conditions (i) and (ii) of Corollary 5.5.3. Then the ESD of $S = X_n^u X_n^{uT}$ converges weakly almost surely to a non-random probability distribution whose moments depend on $\{c_{2k}\}_{k\geq 1}$ and y.*

The proof of Corollay 5.5.12 follows in the same manner as Corollary 5.5.10 by considering $\sigma(x,y) : [0,1]^2 \longrightarrow [0,1]$,

$$\sigma(x,y) = \begin{cases} 1 & \text{if } x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

Corollary 5.5.13. *Result 5.1.5 follows from Corollary 5.5.12.*

Proof. Observe that the non-zero entries of X_n^u in this case are $\{\frac{x_{ij}}{\sqrt{n}}\}$. Also using the truncation argument as in Corollary 5.5.1, we can assume the entries to be bounded. Hence the conditions of Corollary 5.5.12 are satisfied with $c_2 = 1$ and $c_{2k} = 0$ for $k \geq 2$. Hence Corollary 5.5.12 implies that the ESD of S converges weakly almost surely to a non-random probability distribution. □

5.5.7 Hypergraphs, Noiry words and $SS(2k)$

It is undoubtedly clear by now that special symmetric partitions play an indispensable role in the LSD of the S matrix. We have already seen the description of $SS(2k)$ in terms of coloured rooted ordered trees in Section 3.2. In this section, we shall that see a few more structures and combinatorial objects which have previously appeared in the literature (Benaych-Georges and Cabanal-Duvillard [2012], Noiry [2018]), can be described via $SS(2k)$. These descriptions bring out the fact that $SS(2k)$ serve as a central combinatorial object in the LSD of the S matrix and ties in the rest of the results effectively.

The LSD of the S matrix with triangular i.i.d. entries, was studied in Benaych-Georges and Cabanal-Duvillard [2012] where the authors used the concepts of *Hypergraphs*.

Definition 5.5.14. *(**Hypergraphs**) Let G be a graph with vertex set V. Let π and τ be partitions, respectively, of V and the edge set. Then the hypergraph $H(\pi,\tau)$ is a graph with vertex set G_π (i.e. π) and edges $\{E_W; W \in \tau\}$, where each edge E_W is the set of blocks $J \in \pi$ such that at least one edge of G_π starting or ending at J belongs to W. Further, if no two of the edges are allowed to have more than one common vertex, then $H(\pi,\tau)$ is said to be a hypergraph with no cycles.*

Details on Hypergraphs is available in Sections 5.3 and 12.3.2 in Benaych-Georges and Cabanal-Duvillard [2012] and Berge [1989], respectively. In Benaych-Georges and Cabanal-Duvillard [2012], their Equation (22) describes the moments of the LSD of S via a sum on *Hypergraphs with no cycles*. Here we shall show that these hypergraphs are in one-one correspondence with the special symmetric words.

Lemma 5.5.15. *For every word $\boldsymbol{\omega} \in SS_b(2k)$, there exists a unique hypergraph $H(\sigma,\tau)$ which has no cycles where $\sigma,\tau \in \mathcal{P}(k)$ with $|\sigma|+|\tau| = b+1$. The converse is also true.*

Proof. Let $\boldsymbol{\omega} \in SS_b(2k)$ with $(r+1)$ and $(b-r)$ even and odd generating vertices respectively. Suppose the even and the odd generating vertices are respectively $\pi(i_{t_0}) = \pi(0), \pi(i_{t_1}), \ldots, \pi(i_{t_r})$ and $\pi(i_{m_1}) = \pi(1), \pi(i_{m_2}), \ldots, \pi(i_{m_{b-r}})$. Let

$$
\begin{aligned}
V_j =& \{\pi(2i) : \pi(2i) = \pi(i_{t_j}), 1 \le i \le k\}, 0 \le j \le r,\\
W_j =& \{\pi(2i-1) : \pi(2i-1) = \pi(i_{m_j}), 1 \le i \le k\}, 1 \le j \le (b-r).
\end{aligned}
$$

Clearly, $\sigma = \{V_j; 0 \le j \le r\}$ and $\tau = \{W_j; 1 \le j \le (b-r)\}$ are two partitions of $\{1,2,\ldots,k\}$. Therefore, we can construct a hypergraph $H(\sigma,\tau)$ where σ is the vertex set, and $\{E_W; W \in \tau\}$ is the edge set (see (5.5.15)).

Now suppose if possible, $H(\sigma,\tau)$ has a cycle. That means by construction, there exists α,β $(\alpha \ne \beta) \in \{1,2,\ldots,(b-r)\}$ and q,l $(q \ne l) \in \{0,1,\ldots,r\}$ such that $V_q, V_l \in W_\alpha \cap W_\beta$. That is, there are edges $(\pi(k_1-1),\pi(k_1)), (\pi(k_2-1),\pi(k_2)), (\pi(k_3-1),\pi(k_3)), (\pi(k_4-1),\pi(k_4))$ with $k_i, 1 \le i \le 4$ odd such that $\pi(k_1-1) \in V_q, \pi(k_1) \in W_\alpha$, $\pi(k_2-1) \in V_q, \pi(k_2) \in W_\beta$, $\pi(k_3-1) \in V_l, \pi(k_3) \in W_\alpha$ and $\pi(k_4-1) \in V_l, \pi(k_4) \in W_\beta$.

As the positions $(\pi(k_i - 1), \pi(k_i), i = 1, 2, 3, 4$ are all distinct, there are four distinct letters that appear at these four positions of $\boldsymbol{\omega}$.

Without loss of generality suppose, from left to right $(\pi(k_4-1), \pi(k_4))$ is the rightmost (among the four positions mentioned above) in $\boldsymbol{\omega}$. Since $\pi(k_4 - 1) \in V_l$, and $\pi(t_l)$ comes before $\pi(k_4 - 1)$, it cannot be chosen freely. Using a similar argument, $\pi(k_4)$ also cannot be chosen freely. Also they have been chosen as generating vertices of three different letters that have appeared in the positions $(\pi(k_i - 1), \pi(k_i)), 1 \leq i \leq 3$. Using Lemma 5.4.2, this is not possible as the letter at $(\pi(k_4 - 1), \pi(k_4))$ is different from the previous three letters. Thus $H(\sigma, \tau)$ does not have a cycle.

Moreover, it is evident by construction that for every special symmetric word, we get a unique $H(\sigma, \tau)$ without any cycles.

Conversely, suppose $H(\sigma, \tau)$ is a hypergraph with no cycles and, $\sigma, \tau \in \mathcal{P}(k)$ with $|\sigma| + |\tau| = b + 1$. We form a word of length $2k$ from it in the following manner.

Let $\sigma = \{V_0, V_1, \ldots, V_r\}$ and $\tau = \{W_1, \ldots, W_{b-r}\}$ (as $|\sigma| + |\tau| = b + 1$). Then we choose the even vertices $\pi(2i), 0 \leq i \leq k-1$ from σ, and odd vertices $\pi(2i-1), 1 \leq i \leq k$ from τ and $\pi(i) = \pi(j)$ if i and j belong to the same block of σ or τ (depending on i and j both being even or odd respectively).

Thus we get a word $\boldsymbol{\omega}$ of length $2k$ whose even and odd generating vertices are $\{\pi(\min\{V_s\})\}_{0 \leq s \leq r}$ and $\{\pi(\min\{W_t\})\}_{1 \leq t \leq (b-r)}$ respectively. Hence there are b distinct letters in $\boldsymbol{\omega}$.

Now as $H(\sigma, \tau)$ does not have a cycle, using the same arguments as above, it can be shown that all the generating vertices have free choice. This can happen only if the word is special symmetric. Thus we obtain $\boldsymbol{\omega} \in SS_b(2k)$ with $(r + 1)$ even generating vertices.

It is easy to see that two different hypergaphs with no cycles cannot give rise to the same special symmetric word.

Hence there is a one-one correspondence between special symmetric words and hypergraphs with no cycles. This completes the proof of the lemma. $\square$

In Proposition 3.1 in Noiry [2018], the author described the limiting moments via equivalence class of words. His notion of words is different from ours described in Section 2.4, and so we call the former *Noiry words*.

Noiry words: Suppose $G = (V, E)$ is a graph with labelled vertices. A word of length $k \geq 1$ on G is a sequence of labels $i_1, i_2, \ldots, i_k$ such that for each $j \in \{1, 2, \ldots, k-1\}$, $\{i_j, i_{j+1}\}$ is a pair of adjacent labels, i.e., the associated vertices are neighbours in G. A word of length k is closed if $i_1 = i_k$. See Section 3 in Noiry [2018] for more details. Such closed words will be called Noiry words.

Equivalence of Noiry words: Let $\boldsymbol{i} = i_1, i_2, \ldots, i_k$ and $\boldsymbol{i'} = i'_1, i'_2, \ldots, i'_k$ be two Noiry words on two labeled graphs G and G' with vertex set V. These words are said to be equivalent if there is a bijection σ of $\{1, 2, \ldots, |V|\}$ such that $\sigma(i_j) = i'_j, 1 \leq j \leq k$. This defines an equivalence relation on the set of all Noiry words, thereby giving rise to equivalence classes of Noiry words.

$\mathbf{W}_k(a, a+1, l, \boldsymbol{b}), \boldsymbol{b} = (b_1, b_2, \ldots, b_a) \in \mathbb{N}^a, b_i \geq 2, \sum_{i=1}^{a} b_i = 2k$, (see Section 3 and equation (3.2) of Noiry [2018]) denotes an equivalence class of Noiry words on a labeled rooted planar tree with a edges, of which l are odd, and each edge is traversed b_i times, $1 \leq i \leq a$. Then the kth moment of the LSD, as given in equation (3.2) of Noiry [2018] is

$$\beta_k(\mu) = \sum_{a=1}^{k} \sum_{l=1}^{a} \alpha^l \sum_{\substack{\boldsymbol{b}=(b_1, b_2, \ldots, b_a) \\ b_i \geq 2, b_1 + \cdots + b_a = 2k}} |\mathbf{W}_k(a, a+1, l, \boldsymbol{b})| \prod_{i=1}^{a} C_{b_i}.$$

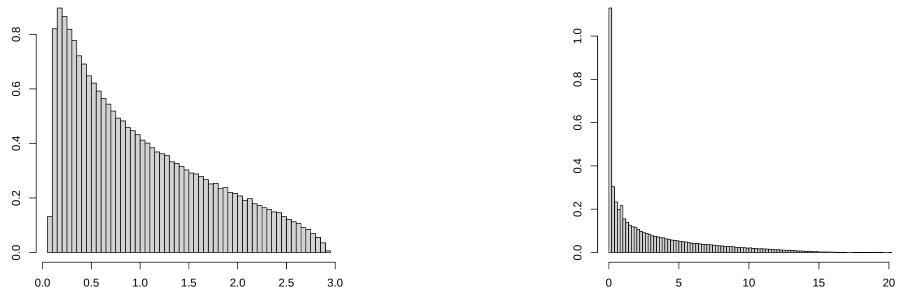
In the next lemma we show how that these equivalence classes of words correspond to special symmetric words.

Lemma 5.5.16. *Each equivalence class* $\mathbf{W}_k(a, a+1, l, \boldsymbol{b}), \boldsymbol{b} = (b_1, b_2, \ldots, b_a) \in \mathbb{N}^a, b_i \geq 2, \sum_{i=1}^{a} b_i = 2k$ *is a word* $\boldsymbol{\omega} \in SS_a(2k)$ *with* l *odd generating vertices, and where each letter appears* $b_i, 1 \leq i \leq a$ *times in* $\boldsymbol{\omega}$.

Proof. Recall from Section 5.3 that we have defined words to be equivalence classes of circuits with the equivalence relation arising from the link functions (see (5.3.2)). Now Noiry words are not equivalence classes to begin with, they form equivalence classes if they are relabeled in a certain way as described above. From this, and how we have defined equivalence of circuits, observe that an equivalence class of Noiry words is nothing but a word in our case. Now the only words with a distinct letters for which $a+1$ generating vertices can be chosen freely, are the special symmetric words with a distinct letters (see Lemma 5.4.2). Thus $\mathbf{W}_k(a, a+1, l, \boldsymbol{b}), \boldsymbol{b} = (b_1, b_2, \ldots, b_a) \in \mathbb{N}^a, b_i \geq 2, \sum_{i=1}^{a} b_i = 2k$ is a word $\boldsymbol{\omega} \in SS_a(2k)$ with l odd generating vertices where each letter appears $b_i, 1 \leq i \leq a$ times in $\boldsymbol{\omega}$. $\square$

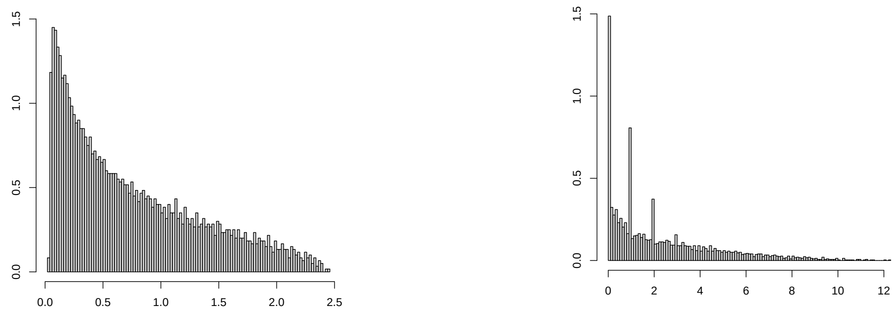
5.6 Simulations

In this section we present a few simulations that demonstrate the different distributions we get as LSDs by considering different kinds of input for the S matrix.



(A) Input is i.i.d $x_{ij} \sim N(0,1)/\sqrt{n}$ for every n. (B) Input is i.i.d $x_{ij} \sim Ber(3/n)$ for every n.

FIGURE 5.1: Histogram of the eigenvalues of S for $p = 1000, n = 2000$, 30 replications.



(A) Input is i.i.d $x_{ij} \sim \frac{N(0,1)}{\sqrt{n}}, i \leq j$. (B) Input is i.i.d $x_{ij} \sim Ber(3/n), i \leq j$ and 0 otherwise.

FIGURE 5.2: Histogram of the eigenvalues of S when X is triangular, for $p = 1000, n = 2000$, 30 replications.

Chapter 6

Other patterned XX^T matrices

In this chapter we look at some real rectangular $p \times n$ random matrices A_p and study the empirical distribution of $S_A = A_p A_p^T$ when $n, p(n) \to \infty$ and p and n are comparable, i.e., $p/n \to y \in (0, \infty)$. We will let A_p to be the symmetric as well as asymmetric versions of the four matrices that have been previously discussed in Chapter 4, namely reverse circulant, circulant, Toeplitz and Hankel, with entries that are real and independent. Hence A_p is any one of the matrices given below.

$$A_p = T^{(s)}, T, H^{(s)}, H, R^{(s)}, R, C^{(s)}, C.$$

$$T^{(s)} = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_{n-1} \\ x_1 & x_0 & x_1 & \cdots & x_{n-2} \\ x_2 & x_1 & x_0 & \cdots & x_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{|p-1|} & x_{|p-2|} & x_{|p-3|} & \cdots & x_{|p-n|} \end{bmatrix}, \quad T = \begin{bmatrix} x_0 & x_{-1} & x_{-2} & \cdots & x_{1-n} \\ x_1 & x_0 & x_{-1} & \cdots & x_{2-n} \\ x_2 & x_1 & x_0 & \cdots & x_{3-n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{p-1} & x_{p-2} & x_{n-3} & \cdots & x_{p-n} \end{bmatrix},$$

$$H^{(s)} = \begin{bmatrix} x_2 & x_3 & x_4 & \cdots & x_{n+1} \\ x_3 & x_4 & x_5 & \cdots & x_{n+2} \\ x_4 & x_5 & x_6 & \cdots & x_{n+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{p+1} & x_{p+2} & x_{n+3} & \cdots & x_{p+n} \end{bmatrix}, \quad H = \begin{bmatrix} x_2 & x_{-3} & x_{-4} & \cdots & x_{-(n+1)} \\ x_3 & x_4 & x_{-5} & \cdots & x_{-(n+2)} \\ x_4 & x_5 & x_6 & \cdots & x_{-(n+3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{p+1} & x_{p+2} & x_{p+3} & \cdots & x_{-(p+n)} \end{bmatrix},$$

$$R^{(s)} = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_{n-1} \\ x_1 & x_2 & x_3 & \cdots & x_0 \\ x_2 & x_3 & x_4 & \cdots & x_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{(p-1)\text{mod } n} & \cdots & & \cdots & x_{(p-2)\text{mod } n} \end{bmatrix},$$

$$R = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_{n-1} \\ x_{-1} & x_2 & x_3 & \cdots & x_0 \\ x_{-2} & x_{-3} & x_4 & \cdots & x_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{-(p-1)\text{mod } n} & \cdots & & \cdots & x_{(p-2)\text{mod } n} \end{bmatrix},$$

$$C^{(s)} = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_1 \\ x_1 & x_0 & x_1 & \cdots & x_2 \\ x_2 & x_1 & x_0 & \cdots & x_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n/2-|n/2-|p-1||} & \cdots & \cdots & \cdots & x_{n/2-|n/2-|p-n||} \end{bmatrix},$$

$$C = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_{n-1} \\ x_1 & x_0 & x_1 & \cdots & x_{n-2} \\ x_2 & x_1 & x_0 & \cdots & x_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{(1-p)(\text{mod } n)} & \cdots & \cdots & \cdots & x_{(n-p)(\text{mod } n)} \end{bmatrix}.$$

We have dropped the suffix p here for ease of reading. The matrices $T^{(s)}, H^{(s)}, R^{(s)}$ and $C^{(s)}$ are the rectangular versions of the symmetric Toeplitz, Hankel, reverse circulant and circulant matrices so that the (i,j)th entry is equal to the (j,i)th entry whenever $1 \leq i, j \leq \min(p,n)$. The matrices T, H, R and C are the *asymmetric* versions of these matrices. These matrices can also be described via link functions, see Section 2.4.

We explore the existence of the LSD of S_A under suitable conditions on the entries of A_p. A brief discussion to relate our results with the models and results that already exist in the literature, are given below.

When the entries of A_p come from a single i.i.d. sequence, the following result is known.

Result 6.0.1. *(Theorems 1(i), 2(i), 3(i), 4(i) in Bose et al. [2010]) Suppose the input sequence $\{x_i : -(n+p) \leq i \leq n+p\}$ of A_p are i.i.d. with mean 0 and variance 1. Then as $p, n \to \infty$ with $p/n \to y \in (0,\infty)$, the ESD of $\frac{1}{n}S_A$ converges weakly almost surely to some non-random probability distribution μ_A, for each matrix A_p.*

We generalise these results by allowing the distribution of the entries to vary with n, as well as with their positions in the matrix (see Theorem 6.1.1). Like Theorems 4.2.2—4.2.4, Theorem 6.1.1 also claims the convergence of the EESD only. The almost sure or in probability convergence of the ESD is not true in general. However, in some special cases as in Chapter 4, the almost sure convergence holds.

We also find some relationships between the LSDs of $S_{R^{(s)}}, S_C, S_T$ and $S_{H^{(s)}}$. For instance, when the entries are i.i.d. for every n, and have exploding moments, the LSDs of S_T and $S_{H^{(s)}}$ are identical; so are the LSDs of S_C and $S_{R^{(s)}}$.

In Section 6.1 we describe our main result, namely Theorem 6.1.1. In Section 6.3.2, we state and prove a few lemmas that lead to the proof of Theorem 6.1.1. In Section 6.4, we present a few well-known models for A like sparse matrices, matrices with variance profile, triangular and band matrices, and conclude the convergence of the ESD of S_A as special cases of Theorem 6.1.1. *This chapter is based on Bose and Sen [2023].*

6.1 Main results

Consider the matrix $S_A = A_p A_p^T$ where the entries of A_p are constructed from the sequence of random variables $\{x_{i,n}; -(n+p) \le i \le (n+p)\}$. We will denote A_p by A and write x_i for $x_{i,n}$. Recall that as $p, n \to \infty$, $p/n \to y \in (0, \infty)$. We introduce the following assumptions on x_i. These assumptions are very similar to Assumption B in Chapter 4.

Assumption B1. Suppose there exists a sequence $\{r_n\}$ with $r_n \in [0, \infty]$ such that

(i) for each $k \in \mathbb{N}$,

$$n\, \mathbb{E}\left[x_i^{2k} \mathbf{1}_{\{|x_i| \le r_n\}}\right] = f_{2k,n}\left(\frac{i}{n}\right) \quad \text{for} \quad -(n+p) \le i \le n+p, \tag{6.1.1}$$

$$\lim_{n\to\infty} n^{\alpha} \sup_{0 \le i \le n-1} \mathbb{E}\left[x_i^{2k-1} \mathbf{1}_{\{|x_i| \le r_n\}}\right] = 0 \ \text{ for all } \alpha < 1, \tag{6.1.2}$$

where $\{f_{k,n}; 0 \le k \le n\}$ is a sequence of bounded and integrable functions on $[-(1+y), 1+y]$.

(ii) For each $k \ge 1$, $f_{2k,n}, n \ge 1$ converge uniformly to a function f_{2k} .

(iii) Let $M_{2k} = \|f_{2k}\|$ (where $\|\cdot\|$ denotes the sup norm) and $M_{2k-1} = 0$ for all $k \ge 1$. Then, $\alpha_k = \sum_{\sigma \in \mathcal{P}(2k)} M_\sigma$ satisfy *Carleman's condition*,

$$\sum_{k=1}^{\infty} \alpha_{2k}^{-\frac{1}{2k}} = \infty.$$

As we will see in Section 6.4, these assumptions are naturally satisfied by various well-known models. Now we state the main theorem of this chapter.

Theorem 6.1.1. *Suppose A is one of the eight rectangular matrices $T^{(s)}, T, H^{(s)}, H, R^{(s)}$, $R, C^{(s)}, C$, with entries $\{x_i\}$ which are independent, and satisfy Assumption B1. Let Z_A be the corresponding $p \times n$ matrix with entries $y_l = x_l \mathbf{1}_{\{|x_l| \le r_n\}}$. Then the EESD of*

$S_{Z_A} = Z_A Z_A^T$ converges weakly to a probability measure μ_A say, whose moment sequence is determined by the functions f_{2k}, $k \geq 1$, in each of the eight cases. Further if

$$\sum_l \mathbb{E}[x_l^2 \mathbf{1}_{\{|x_l| > r_n\}}] \to 0, \tag{6.1.3}$$

then the EESD of $S_A = AA^T$ converges weakly to μ_A.

Remark 6.1.2. *As mentioned earlier, the almost sure or in probability convergence of the ESD to the limit μ_A does not hold in general, unlike that of the S matrix (see Chapter 5). The reason for this is the same as given in Remark 4.2.1 of Chapter 4. This lack of convergence is also clear from the simulations given in Figures 6.1 and 6.2. In particular there is no almost sure convergence in the sparse case. Of course, almost sure convergence can hold in special cases, for example in the fully i.i.d. case.*

Remark 6.1.3. *Note that as mentioned in Remark 5.2.4, we can conclude the convergence of the EESD of AA^T, using matrices of the form*

$$\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \tag{6.1.4}$$

from Chapter 4.

However, for many of the cases, like the Toeplitz and the Circulant matrices, the symmetric and asymmetric versions of the matrices demonstrate a major difference: in the asymmetric Toeplitz and ciculant matrices, the moments are given via the set of all symmetric partitions (all other partitions contribute 0 to the limiting moments) although for the symmetric Toeplitz and circulant matrices, the moments are given via the set of all even partitions. This phenomenon is difficult to unearth if we proceed with the matrix in (6.1.4). Also in the cases of the asymmetric Hankel and reverse circulant matrices, some of the symmetric partitions do not contribute. This is observed in Lemmas 6.3.3 and 6.3.4. These facts are also hard to discover if we take the approach using (6.1.4). Hence the limiting moments cannot be expressed as precisely. Thus, we take help of the machinery developed in Chapters 2 and 5 and derive the LSD results for these matrices independently.

6.2 Some preliminaries

The notion of circuits and words for S_A remain identical as for the S matrix given in Section 5.3. Like the S matrix, notice that circuits π with $\ell(\pi) = 2k$ are required to deal with the kth moment of S_A. For any choice of the link L (see Section 2.4), set

$$\xi_\pi(2i-1) = L(\pi(2i-2), \pi(2i-1)), 1 \le i \le k,$$
$$\xi_\pi(2i) = L(\pi(2i), \pi(2i-1)), 1 \le i \le k.$$

Then, with $Y_\pi = \prod_{i=1}^{k} x_{\xi_{\pi(2i-1)}} x_{\xi_{\pi(2i)}}$,

$$\mathbb{E}\big[\operatorname{Tr}(S_A^k)\big] = \mathbb{E}\big[\operatorname{Tr}(AA^T)^k\big] = \sum_{\pi:\ell(\pi)=2k} \mathbb{E}[Y_\pi]. \tag{6.2.1}$$

The class $\Pi_{S_A}(\boldsymbol{\omega})$: For $\boldsymbol{\omega}$,

$$\Pi_{S_A}(\boldsymbol{\omega}) = \{\pi : \ \boldsymbol{\omega}[i] = \boldsymbol{\omega}[j] \ \Leftrightarrow \xi_\pi(i) = \xi_\pi(j) \text{ for all } i, j\}. \tag{6.2.2}$$

Now,

$$\begin{aligned}\lim_{p\to\infty} \frac{1}{n}\mathbb{E}[\operatorname{Tr}(S_A)^k] &= \lim_{p\to\infty} \frac{1}{p} \sum_{\pi:\ell(\pi)=2k} \mathbb{E}[Y_\pi] \\ &= \lim_{p\to\infty} \sum_{b=1}^{k} \sum_{\substack{\omega \text{ matched of length } 2k \\ \text{with b distinct letters}}} \frac{1}{p} \sum_{\pi\in\Pi_{S_A}(\boldsymbol{\omega})} \mathbb{E}(Y_\pi). \end{aligned}\tag{6.2.3}$$

Note that all words that appear above are of length $2k$. For every $k \ge 1$, the words of length $2k$ corresponding to the circuits of A and S_A, are related. Here we make a key observation in that regard.

Observation (i): Let $A^{(s)}$ stand for any of the symmetric matrices $R^{(s)}, H^{(s)}, C^{(s)}$ or $T^{(s)}$ and let $\Pi_{A^{(s)}}(\boldsymbol{\omega})$ be the possibly larger class of circuits for $A^{(s)}$ with range $1 \le \pi(i) \le \max(p,n), 0 \le i \le 2k$. Let $\Pi_{S_{A^{(s)}}}(\boldsymbol{\omega})$ and $\Pi_{A^{(s)}}(\boldsymbol{\omega})$ denote the set of all circuits corresponding to a word $\boldsymbol{\omega}$ arising from the circuits corresponding to $A^{(s)}$ and

$S_{A^{(s)}}$, respectively. Then, for every $k \geq 1$ and any word $\boldsymbol{\omega}$ of length $2k$,

$$\Pi_{S_A}(\boldsymbol{\omega}) \subset \Pi_{S_{A^{(s)}}}(\boldsymbol{\omega}) \subset \Pi_{A^{(s)}}(\boldsymbol{\omega}). \tag{6.2.4}$$

Even and odd generating vertices are defined exactly as in Section 5.3. Here,

$$\begin{aligned}\big|\Pi_{S_A}(\boldsymbol{\omega})\big| = \Big|\big\{\big(\pi(0), \pi(1), \ldots, \pi(2k)\big) : 1 \leq \pi(2i) \leq p, 1 \leq \pi(2i-1) \leq n \text{ for } i = 0, 1, \ldots, k,\\ \pi(0) = \pi(2k), \quad \xi_\pi(i) = \xi_\pi(j) \text{ if and only if } \boldsymbol{\omega}[i] = \boldsymbol{\omega}[j], 1 \leq i, j \leq 2k\big\}\Big|.\end{aligned} \tag{6.2.5}$$

As $p/n \to y > 0$,

$$|\Pi_{S_A}(\boldsymbol{\omega})| = \mathcal{O}(p^{r+1}n^{b-r}) \text{ if } \omega \text{ has } b \text{ distinct letters and } (r+1) \text{ even generating vertices.} \tag{6.2.6}$$

6.3 Proof of Theorem 6.1.1

As we have been doing in the previous chapters, we first present a few lemmas that are needed for the proof. In Lemmas 6.3.1–6.3.5, we identify the words that possibly contribute positively to the limiting moments. Next, in Lemma 6.3.6, we prove that the entries of A_p (where A_p is any one of the eight matrices $T^{(s)}, T, H^{(s)}, H, R^{(s)}, R, C^{(s)}, C$) can be assumed to have mean zero. Lemma 6.3.7, shows how it suffices to prove the convergence of the EESD for the truncated matrices S_{Z_A} in Theorem 6.1.1. Finally, we finish the proof of Theorem 6.1.1.

6.3.1 Identification of words that contribute

As observed in Chapter 5 for the S matrix, in this case too, $\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}}|\Pi_{S_A}(\boldsymbol{\omega})|$ helps us determine which words contribute to the limiting moments. Here, we look into the existence and value of $\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}}|\Pi_{S_A}(\boldsymbol{\omega})|$ for each of the matrices A_p.

Lemma 6.3.1. ***(Symmetric Toeplitz matrix,** $T^{(s)}$**)*** *Suppose $\boldsymbol{\omega}$ is a word with b distinct letters and $(r+1)$ even generating vertices. Then $\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}}|\Pi_{S_{T^{(s)}}}(\boldsymbol{\omega})| = \alpha_{T^{(s)}}(\boldsymbol{\omega}) > 0$ if $\boldsymbol{\omega}$ is an even word. Else this limit is 0.*

Proof. This proof will be similar to the proof of Lemma 4.3.3, except the fact that the ratio $y_n = p/n$ occurs in certain places because the even generating vertices range from 1 to p. We shall also borrow all the notations from Lemma 4.3.3.

First suppose $\boldsymbol{\omega} \in \mathcal{P}(2k) \setminus E_b(2k)$. Then from (6.2.4) and Lemma 4.3.3, and keeping in mind that $p/n \to y > 0$ as $n \to \infty$, it is easy to see that

$$\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_{T^{(s)}}}(\boldsymbol{\omega})| = 0.$$

Let $\boldsymbol{\omega}$ be an even word, with b distinct letter and $(r+1)$ even generating vertices. As the word is even, as seen in Lemma 4.3.3, the circuit condition is automatically satisfied. So there is no additional restriction while choosing the generating vertices.

Then s_i satisfies (4.3.15) and $\pi(i)$ satisfies (4.3.18).

In this case, we choose

$$v_{2i} = \frac{\pi(2i)}{p} \ \text{ for } 0 \le i \le k, \ v_{2i-1} = \frac{\pi(2i-1)}{n} \quad \text{for} \ \ 1 \le i \le k,$$

$$\text{and } u_i = \frac{s_i}{n} \text{ for } 1 \le i \le 2k.$$

Therefore $v_i = \frac{1}{y_n}(v_{i-1} \pm u_{i_j})$ when i is even, and $v_i = y_n v_{i-1} \pm u_{i_j}$ when i is odd.

Similarly as in proof of Lemma 4.3.3, we can show that For any $1 \le i \le 2k$,

$$v_i = \begin{cases} v_0 + \frac{1}{y_n}\sum_{j=1}^{i} \alpha_{ij} u_{i_j} & \text{if } i \ \text{ is even,} \\ y_n v_0 + \sum_{j=1}^{i} \alpha_{ij} u_{i_j} & \text{if } i \ \text{ is odd,} \end{cases} \tag{6.3.1}$$

where α_{ij} depends on the choice of sign in (4.3.18).

Thus we have, for $1 \le i \le 2k$,

$$v_i = \begin{cases} v_0 + \frac{1}{y_n} LC^T_{i,u,n}(u_S), & \text{if } i \ \text{ is even,} \\ y_n v_0 + LC^T_{i,u}(u_S) & \text{if } i \ \text{ is odd,} \end{cases}$$

where $LC^T_{i,u,n}(u_S)$ denotes a linear combination of $\{u_i : \pi(i) \in S\}$.

Just as in the proof of Lemma 4.3.3, due to (4.3.15) there are $\prod_{i=1}^{b} \binom{k_i - 1}{\frac{k_i}{2}}$ different sets of linear combinations corresponding to each word $\boldsymbol{\omega}$, where $k_1, \ldots, k_b$ are the block sizes of $\boldsymbol{\omega}$, that determine the non-generating vertices.

Hence, we have (see (4.3.20))

$$\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_{T(s)}}(\boldsymbol{\omega})| = \sum_{LC^T_{i,u} \in LC^T_{\boldsymbol{\omega}}} \int_0^1 \int_{-1}^{y} \cdots \int_{-y}^{1} \mathbf{1}(0 \leq x_0 + \frac{1}{y} LC^T_{i,u}(u_S) \leq 1, \ \forall\ 2i \in S')$$
$$\mathbf{1}(0 \leq y x_0 + LC^T_{i,u}(u_S) \leq 1, \ \forall\ (2i-1) \in S')\ dx_0 du_S, \tag{6.3.2}$$

where $du_S = \prod_{j=1}^{b} du_{i_j}$ denotes the $(b+1)$-dimensional Lebesgue measure on $[-1, y]^r \times [-y, 1]^{b-r}$ and $LC^T_{i,u}$ is the limit of $LC^T_{i,u,n}$ as $p/n \to y$.

As $y > 0$, the integrand in (6.3.2) can be shown to be positive on a certain region of $[0, 1] \times [-1, y]^r \times [-y, 1]^{b-r}$ in a similar manner as in the proof of Lemma 4.3.3.

Thus

$$\lim_{n\to\infty} \frac{1}{n^{b+1}} |\Pi_{S_{T(s)}}(\boldsymbol{\omega})| = \alpha(\boldsymbol{\omega}) > 0 \ \text{ for any even word } \boldsymbol{\omega},$$

where $\alpha_{T(s)}(\boldsymbol{\omega})$ is the sum of the integrals defined in (6.3.2).

This completes the proof of the lemma. $\square$

Lemma 6.3.2. *(**(Asymmetric) Toeplitz matrix,** T) Suppose $\boldsymbol{\omega}$ is a word with b distinct letters and $(r+1)$ even generating vertices. Then $\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_T}(\boldsymbol{\omega})| = \alpha_T(\boldsymbol{\omega}) > 0$ if and only if $\boldsymbol{\omega}$ is symmetric.*

Proof. Let

$$s_i = \pi(i) - \pi(i-1), \quad \text{for } 1 \leq i \leq 2k.$$

From (6.2.2), we know that $\boldsymbol{\omega}[i] = \boldsymbol{\omega}[j]$ if and only if $\xi_\pi(i) = \xi_\pi(j)$. This implies

$$\begin{aligned} s_i &= s_j \quad \text{when } i \text{ and } j \text{ are of same parity,} \\ s_i &= -s_j \quad \text{when } i \text{ and } j \text{ are of opposite parity.} \end{aligned} \tag{6.3.3}$$

Now we fix an $\boldsymbol{\omega}$ with b distinct letters which appear at $i_1, i_2, \ldots, i_b$ positions for the first time. Also let $\boldsymbol{\omega}$ have $(r+1)$ even generating vertices. Using the same arguments as in Lemma 4.3.3, choosing $\pi(i_j), 0 \leq j \leq b$ is equivalent to choosing $\pi(0), s_{i_j}, 1 \leq j \leq b$. Next we show that if the word is not symmetric, then $\pi(0)$ and $s_{i_j}, 1 \leq j \leq b$ satisfy a non-trivial linear relation.

Observe that the circuit condition needs to be satisfied automatically. Therefore,

$$\sum_{i=1}^{2k} s_i = \pi(0) - \pi(2k) = 0. \tag{6.3.4}$$

Using (6.3.3), we see that there exists $\alpha_j, 1 \leq j \leq b$ such that

$$\sum_{i=1}^{b} \alpha_j s_{i_j} = 0.$$

Since $\{\pi(0), s_{i_j}, 1 \leq j \leq b\}$ does not satisfy any non-trivial relation, we must have $\alpha_j = 0$ for all $j \in \{1, 2, \ldots, b\}$. Therefore for each j,

$$\big|\{l : s_l = s_{i_j}\}\big| = \big|\{l : s_l = -s_{i_j}\}\big|. \tag{6.3.5}$$

Now from the definition of ξ_π, $\xi_\pi(2i) = s_{2i}, 0 \leq i \leq k$ and $\xi_\pi(2i-1) = -s_{2i-1}, 0 \leq i \leq k$. Therefore for each $j \in \{1, 2, \ldots, b\}$, to satisfy (6.3.5), we must have

$$\Big|\Big\{l : l \text{ even and } \xi_\pi(l) = \xi_\pi(i_j)\Big\}\Big| = \Big|\Big\{l : l \text{ odd and } \xi_\pi(l) = \xi_\pi(i_j)\Big\}\Big|. \tag{6.3.6}$$

That is, each letter appears equal number of times at odd and even places. Hence the word is symmetric.

Therefore, if $\boldsymbol{\omega}$ is not symmetric, the circuit condition gives rise to a linear relation between $\pi(i_j), 0 \leq j \leq b$. So, at least one of the generating vertices (even or odd) is a linear combination of the others, and hence,

$$\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_T}(\boldsymbol{\omega})| = 0 \quad \text{if } \boldsymbol{\omega} \text{ is not symmetric.} \tag{6.3.7}$$

Next, suppose $\boldsymbol{\omega}$ is a symmetric word with b distinct letters and $(r+1)$ even generating vertices. We shall show that $\lim_{n\to\infty} \dfrac{1}{p^{r+1}n^{b-r}} |\Pi_{S_T}(\boldsymbol{\omega})| = \alpha_T(\boldsymbol{\omega}) > 0$.

First observe that (6.3.6) is true for any symmetric word. So from the above discussion it is easy to see that the circuit condition is automatically satisfied.

Suppose the letters make their first appearances at $i_1, i_2, \ldots, i_b$ positions in $\boldsymbol{\omega}$. First we fix the generating vertices $\pi(i_j), 0 \leq j \leq b$. Suppose $S = \{\pi(i_j) : 0 \leq j \leq b\}$ and $S' = \{i : \pi(i) \notin S\}$. For $i \in S'$, $\xi_\pi(i) = \xi_\pi(i_j)$ for some $j \in \{1, 2, \ldots b\}$. Then

$$\begin{aligned} \pi(i) =& s_{i_j} + \pi(i-1) \quad \text{if } i \text{ and } i_j \text{ are of same parity}, \\ \pi(i) =& - s_{i_j} + \pi(i-1) \quad \text{if } i \text{ and } i_j \text{ are of opposite parity}. \end{aligned} \tag{6.3.8}$$

Thus, (6.3.8) is nothing but (4.3.15) where the sign has been determined depending on the parity of i and i_j. So, for $1 \leq i \leq 2k$, $v_i = LC^T_{i,n}(v_S)$, (the notations are the same as in the proof of Lemma 6.3.1) where $LC^T_{i,n}(\cdot)$ is a particular set of linear combinations that has been determined by (6.3.8). Consequently, the rest of the proof is same as that of Lemma 6.3.1. Therefore, with $dx_S = \prod_{j=0}^{b} dx_{i_j}$ as the $(b+1)$-dimensional Lebesgue measure $(x_{i_0} = x_0)$,

$$\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_T}(\boldsymbol{\omega})| = \int_0^1 \int_0^1 \int_0^1 \cdots \int_0^1 \mathbf{1}(0 \leq LC^T_i(x_S) \leq 1,\ \forall\, i \in S')\, dx_S. \tag{6.3.9}$$

The integral is positive now follows from the proof of the same fact in Lemma 6.3.1.

$$\therefore \lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_T}(\boldsymbol{\omega})| = \alpha_T(\boldsymbol{\omega}) > 0 \quad \text{for every symmetric word } \boldsymbol{\omega}, \tag{6.3.10}$$

where α_T is the value of an individual integral in the rhs of (6.3.2).

(6.3.7) and (6.3.10) completes the proof. □

Lemma 6.3.3. *(**Hankel matrices,** $H^{(s)}$ **and** H) Suppose $\boldsymbol{\omega}$ is a word with b distinct letters and $(r+1)$ even generating vertices. Then*

(i) $\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_{H^{(s)}}}(\boldsymbol{\omega})| = \alpha_{H^{(s)}}(\boldsymbol{\omega}) > 0$ *if and only if $\boldsymbol{\omega}$ is a symmetric word.*

(ii) $\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_H}(\boldsymbol{\omega})| = \alpha_H(\boldsymbol{\omega})$ *can only be positive if $\boldsymbol{\omega}$ is a symmetric word.*

Proof. This proof is very similar to that of Lemma 4.3.4, except the fact that the ratio $y_n = p/n$ occurs in certain places due to the fact that the even generating vertices range from 1 to p. We shall also borrow all the notations from Lemma 4.3.4, unless otherwise mentioned and only provide a sketch of the proof. First suppose $\boldsymbol{\omega} \in \mathcal{P}(2k) \setminus S_b(2k)$. Then from (6.2.4) and Lemma 4.3.4, it is easy to see that

$$\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_{H^{(s)}}}(\boldsymbol{\omega})| = \lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_H}(\boldsymbol{\omega})| = 0.$$

Now suppose $\boldsymbol{\omega}$ is a symmetric word, with b distinct letters and $(r+1)$ even generating vertices. Since the word is symmetric, as in the proof of Lemma 4.3.4, the circuit condition is automatically satisfied. Also $t_i(= \pi(i) + \pi(i-1))$ satisfies (4.3.21).

In this case, let

$$v_{2i} = \frac{\pi(i)}{p},\ \ v_{2i-1} = \frac{\pi(i)}{n} \text{ for } 0 \leq i \leq k, S = \{\pi(i_j) : 0 \leq j \leq b\} \text{ and } S' = \{i : \pi(i) \notin S\}.$$

For $1 \leq i \leq 2k$, from the link function and the formula for t_i we have (see (4.3.23))

$$v_i = LC^H_{i,n}(v_S), \tag{6.3.11}$$

where $LC^H_{i,n}(v_S)$ denotes a linear combination of $\{v_i : \pi(i) \in S\}$.

Hence similarly as (4.3.24), we get

$$\begin{aligned}&\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} \Pi_{S_{H^{(s)}}}(\boldsymbol{\omega}) \\ &= \int_{[-1/2,1/2]^{b+1}} \mathbf{1}(-1/2 \leq LC^H_i(x_S) \leq 1/2,\ \forall\, i \in S')\, dx_S,\end{aligned} \tag{6.3.12}$$

where $dx_S = \prod_{j=0}^{b} dx_{i_j}$ denotes the $(b+1)$-dimensional Lebesgue measure on $[-\frac{1}{2}, \frac{1}{2}]^{b+1}$.

Let $y_n = p/n$ and for $1 \leq i \leq k$,

$$p_{2i} = x_{2i-1} + y_n x_{2i},\ \ p_{2i-1} = y_n x_{2i-2} + x_{2i-1}, \tag{6.3.13}$$

$$q_{2i} = x_{2i-1} - y_n x_{2i},\ \ q_{2i-1} = y_n x_{2i-2} - x_{2i-1}. \tag{6.3.14}$$

Then it can be shown that for any $1 \leq i \leq 2k$, (see (4.3.24))

$$x_i = \begin{cases} x_0 + \frac{1}{y_n} \sum_{j=1}^{i} \alpha_{ij} p_{i_j} & \text{if } i \text{ is even,} \\ -y_n x_0 + \sum_{j=1}^{i} \alpha_{ij} p_{i_j} & \text{if } i \text{ is odd.} \end{cases} \tag{6.3.15}$$

Now performing the following change of variables in (6.3.12) (see (4.3.25)) we get :

$$z_i = \begin{cases} z_0 + \frac{1}{y_n} \sum\limits_{j=1}^{i} \beta_{ij} q_{i_j} & \text{if } i \text{ is even,} \\ y_n z_0 + \sum\limits_{j=1}^{i} \beta_{ij} q_{i_j} & \text{if } i \text{ is odd,} \end{cases}$$

where $\beta_{ij} = \pm\alpha_{ij}$ according as i_j is odd or even. We shall use the notation $z_i = LC^H_{i,q,n}(z_S)$ to denote this linear relation.

Therefore we can write (6.3.12) as

$$\lim_{n\to\infty} \frac{1}{p^{r+1} n^{b-r}} \Pi_{S_{H(s)}}(\boldsymbol{\omega}) = \int_{-1/2}^{1/2} \int_{[-\frac{y+1}{2}, \frac{y+1}{2}]^b} \mathbf{1}(-1/2 \leq LC^H_{i,q}(z_S) \leq 1/2,\ \forall\, i \in S')\, dq_S,$$

where $dq_S = \prod\limits_{j=0}^{b} dq_{i_j}$ denotes the $(b+1)$-dimensional Lebesgue measure on $[-\frac{1}{2}, \frac{1}{2}] \times [-\frac{y+1}{2}, \frac{y+1}{2}]^b$ and $LC^H_{i,q}$ denotes the limit of the linear combination $LC^H_{i,q,n}$ as $y_n \to y > 0$.

Now it can be proved that the above integrand is positive on a region of positive measure on $[-\frac{1}{2}, \frac{1}{2}] \times [-\frac{y+1}{2}, \frac{y+1}{2}]^b$. This proof is similar to the proof that the integral in the rhs of (6.3.2) is positive. So we omit the details.

This completes the proof of Part (i).

To prove Part (ii), note that for the asymmetric Hankel link function,

$$\begin{aligned} &\xi_\pi(i) = \xi_\pi(j) \quad \text{if and only if} \quad t_i = t_j \quad \text{and,} \\ &\operatorname{sgn}(\pi(i) - \pi(i-1)) = \operatorname{sgn}(\pi(j) - \pi(j-1)) \quad \text{if } i \text{ and } j \text{ are of same parity, or} \\ &\operatorname{sgn}(\pi(i) - \pi(i-1)) = \operatorname{sgn}(\pi(j-1) - \pi(j)) \quad \text{if } i \text{ and } j \text{ are of opposite parity.} \end{aligned}$$

Let

$$\mathcal{E}_{\boldsymbol{\omega}} = \{0, i_j; i_j \text{ is even}, 1 \le j \le b\} \tag{6.3.16}$$

$$\mathcal{O}_{\boldsymbol{\omega}} = \{i_j; i_j \text{ is odd}, 1 \le j \le b\}. \tag{6.3.17}$$

For every $j \in \{1, 2, \ldots, b\}$, let

$$C^o_{i_j} = \{i; \xi_\pi(i) = \xi_\pi(i_j), i, i_j \text{ are of opposite parity}, 0 \le i \le 2k\} \tag{6.3.18}$$

$$C^e_{i_j} = \{i; \xi_\pi(i) = \xi_\pi(i_j), i, i_j \text{ are of same parity}, 0 \le i \le 2k\} \tag{6.3.19}$$

Using the notations as in the proof of Part (i), we now have that

$$\begin{aligned}
&\left|\Pi_{S_H}(\boldsymbol{\omega})\right| \\
&= \Big|\Big\{(v_0, v_1, \ldots, v_{2k}) : v_{2i} \in U_p, v_{2i-1} \in U_n \text{ for } 0 \le i \le k, v_0 = v_{2k}, v_i = LC^H_{i,n}(v_S), \\
&\operatorname{sgn}(y_n LC^H_{i,n}(v_S) - LC^H_{i-1,n}(v_S)) = \operatorname{sgn}(y_n v_{i_j} - LC^H_{i-1,n}(v_S)) \text{ when } i_j \in \mathcal{E}_{\boldsymbol{\omega}} \text{ and } i \in C^e_{i_j} \\
&\text{ or } \operatorname{sgn}(y_n LC^H_{i-1,n}(v_S) - LC^H_{i,n}(v_S)) = \operatorname{sgn}(y_n v_{i_j} - LC^H_{i_j-1,n}(v_S)) \text{ when } i_j \in \mathcal{E}_{\boldsymbol{\omega}} \text{ and } i \in C^o_{i_j}, \\
&\operatorname{sgn}(y_n LC^H_{i,n}(v_S) - LC^H_{i-1,n}(v_S)) = \operatorname{sgn}(y_n LC^H_{i_j-1,n}(v_S) - v_{i_j-1}) \text{ when } i_j \in \mathcal{O}_{\boldsymbol{\omega}} \text{ and } i \in C^e_{i_j} \\
&\text{ or } \operatorname{sgn}(y_n LC^H_{i-1,n}(v_S) - LC^H_{i,n}(v_S)) = \operatorname{sgn}(y_n v_{i_j} - LC^H_{i-1,n}(v_S)) \text{ when } i_j \in \mathcal{O}_{\boldsymbol{\omega}} \text{ and } i \in C^o_{i_j}\Big\}\Big|.
\end{aligned}$$

Now $\Pi_{S_H}(\boldsymbol{\omega}) \subset \Pi_{S_{H(s)}}(\boldsymbol{\omega})$. Therefore, if $\boldsymbol{\omega}$ is a word with b distinct letters but is not symmetric, by Part (i), $\frac{1}{p^{r+1}b^{n-r}}\left|\Pi_{S_H}(\boldsymbol{\omega})\right| \to 0$ as $n \to \infty$.

Next let $\boldsymbol{\omega} \in S_b(2k)$ with $(r+1)$ even generating vertices. Clearly for $\boldsymbol{\omega}$, $|\mathcal{E}_{\boldsymbol{\omega}}| = r+1$ and $|\mathcal{O}_{\boldsymbol{\omega}}| = b - r$. Now suppose,

$$\begin{aligned}
&f^H_n(v_S) && (6.3.20)\\
&= \prod_{j=1}^{b}\Big[\prod_{i_j \in \mathcal{E}_{\boldsymbol{\omega}}}\Big(\prod_{i \in C^e_{i_j}} \mathbf{1}(\operatorname{sgn}(y_n LC^H_{i,n}(v_S) - LC^H_{i-1,n}(v_S)) = \operatorname{sgn}(y_n v_{i_j} - LC^H_{i-1,n}(v_S))) \\
&\quad \prod_{i \in C^o_{i_j}} \mathbf{1}(\operatorname{sgn}(y_n LC^H_{i-1,n}(v_S) - LC^H_{i,n}(v_S)) = \operatorname{sgn}(y_n v_{i_j} - LC^H_{i_j-1,n}(v_S)))\Big) \\
&\quad \prod_{i_j \in \mathcal{O}_{\boldsymbol{\omega}}}\Big(\prod_{i \in C^e_{i_j}} \mathbf{1}(\operatorname{sgn}(y_n LC^H_{i,n}(v_S) - LC^H_{i-1,n}(v_S)) = \operatorname{sgn}(y_n LC^H_{i_j-1,n}(v_S) - v_{i_j-1})) \\
&\quad \prod_{i \in C^o_{i_j}} \operatorname{sgn}(y_n LC^H_{i-1,n}(v_S) - LC^H_{i,n}(v_S)) = \operatorname{sgn}(y_n v_{i_j} - LC^H_{i-1,n}(v_S)))\Big]. && (6.3.21)
\end{aligned}$$

Let f^H be the limit of f_n^H as $y_n \to y > 0$. Then

$$\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} \Pi_{S_H}(\boldsymbol{\omega}) = \int_0^1 \int_0^1 \cdots \int_0^1 \mathbf{1}(0 \leq LC_i^H(v_S) \leq 1) f^H(v_S)\, dv_S, \tag{6.3.22}$$

where $dv_S = \prod_{j=0}^b dv_{i_j}$ is the $(b+1)-$dimensional Lebesgue integral on $[0,1]^{b+1}$ and LC_i^H is the limit of the linear combination $LC_{i,n}^H$ as $y_n \to y$.

This completes the proof of Part (ii). □

Let $\lfloor \cdot \rfloor$ denotes the greatest integer function.

Lemma 6.3.4. *(**Reverse circulant matrices,** $R^{(s)}$ **and** R) Suppose $\boldsymbol{\omega}$ is a word of length $2k$ with b distinct letters, and $(r+1)$ even generating vertices. Then*

(i) $\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_{R^{(s)}}}(\boldsymbol{\omega})| = \lfloor y \rfloor^{k-(r+1)} + \alpha_{R^{(s)}}(\boldsymbol{\omega}) > 0$ *if and only if $\boldsymbol{\omega}$ is a symmetric word.*

(ii) $\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_R}(\boldsymbol{\omega})| = \lfloor y \rfloor^{k-(r+1)} + \alpha_R(\boldsymbol{\omega})$ *can only be positive if $\boldsymbol{\omega}$ is a symmetric word.*

Proof. First suppose $\boldsymbol{\omega} \in \mathcal{P}(2k) \setminus S_b(2k)$. Then from (6.2.4) and Lemma 4.3.1, and using the fact that $p/n \to y > 0$ as $n \to \infty$, it is easy to see that

$$\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_{R^{(s)}}}(\boldsymbol{\omega})| = \lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_R}(\boldsymbol{\omega})| = 0.$$

This remains true for $R^{(s)}$ and R.

Let us consider the symmetric reverse circulant link function, $L_{R^{(s)}}$ (see Section 2.4).

Now suppose $\boldsymbol{\omega}$ is a symmetric word with b distinct letters and $(r+1)$ even generating vertices. We shall borrow the notations from Lemma 4.3.1 here.

Since the word is symmetric, as in the proof of Lemma 4.3.1, the circuit condition is automatically satisfied. Also t_i (see (4.3.2)) satisfies (4.3.3) and $\pi(i)$ satisfies (4.3.6).

Recall the generating vertices $\pi(i_j)$, $j = 0, 1, 2, \ldots, b$. Also recall $S = \{\pi(i_j) : 0 \leq j \leq b\}$ and $S' = \{i : \pi(i) \notin S\}$. For every $i \in S'$, (see (4.3.6))

$$\pi(i) = \sum_{j<i} \alpha_{ij} \pi(i_j) \pmod{n} \quad \text{for some } \alpha_{ij} \in \mathbb{Z} \tag{6.3.23}$$

Thus for every $i \in S' \setminus \{2k\}$, there exists unique integer $m_{i,n}$ such that

$$1 \leq \sum_{j<i} \alpha_{ij}\pi(i_j) + m_{i,n} \leq n. \tag{6.3.24}$$

As we have already fixed the generating vertices, from (6.3.23) and (6.3.24) it follows that there is a unique choice for $\pi(2i-1), 1 \leq i \leq k$ such that $2i-1 \in S'$. For all $2i \in S', 1 \leq i < k$, we can have $\lfloor y_n \rfloor$ choices as $1 \leq \pi(2i) \leq p$ and $y_n = p/n$. Moreover, there is an additional choice if

$$\sum_{j<2i} \alpha_{2ij}\pi(i_j) + m_{2i,n} \leq p - \lfloor y_n \rfloor n. \tag{6.3.25}$$

Next let

$$\begin{aligned} v_{2i} =& \frac{\pi(i)}{p}, \quad v_{2i-1} = \frac{\pi(i)}{n} \text{ for } 0 \leq i \leq k, \\ L(a) =& \max\{m \in \mathbb{Z}\}, \ F(a) = a - L(a). \end{aligned} \tag{6.3.26}$$

Also let

$$S^- = \{2i : 2i \notin S' \text{ and (6.3.25) holds true}\}. \tag{6.3.27}$$

Now observe that from (6.3.24) and (6.3.25) it follows that for every $i \in S^-$,

$$F(y_n LC^H_{2i,n}(v_S)) \leq y_n - \lfloor y_n \rfloor, \tag{6.3.28}$$

where $LC^H_{2i,n}$ is the set of linear combinations defined in (6.3.11), and F is the function described in (6.3.26).

From (6.2.5) and the discussion above, it is easy to see that for $\boldsymbol{\omega}$ of length $2k$,

$$\frac{1}{p^{r+1}n^{b-r}}\left|\Pi_{S_{R(s)}}(\boldsymbol{\omega})\right| = \lfloor y_n \rfloor^{k-(r+1)} + \sum_{\phi \neq S_0 \subset S^-} \lfloor y_n \rfloor^{|S^- - S_0|} \Big| \\ \Big\{v_S : F(y_n LC^H_{2i,n}(v_S)) \leq y_n - \lfloor y_n \rfloor, \mathbf{1}(2i \in S_0)\Big\}\Big|.$$

Therefore as $n \to \infty$,

$$\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} \left|\Pi_{S_{R^{(s)}}}(\boldsymbol{\omega})\right| = \lfloor y \rfloor^{k-(r+1)} + \sum_{\phi \neq S_0 \subset S^-} \lfloor y \rfloor^{|S^- - S_0|} \int_0^1 \int_0^1 \cdots \int_0^1 \mathbf{1}\big(F(yLC_{2i}^H(v_S)) \le y - \lfloor y \rfloor) \ \forall 2i \in S_0\big) \, dv_S \tag{6.3.29}$$

where $dv_S = \prod_{j=0}^{b} dv_{i_j}$ is the $(b+1)-$dimensional Lebesgue integral on $[0,1]^{b+1}$.

When $y \ge 1$, the rhs of (6.3.29) is positive. We next show that when $y < 1$, the value of the integral $\int_0^1 \int_0^1 \cdots \int_0^1 \mathbf{1}\big(F(yLC_{2i}^H(v_S)) \le y - \lfloor y \rfloor) \ \forall 2i \in S^-\big) \, dv_S$ is positive.

First note that as $y < 1$, $\lfloor y \rfloor = 0$. Now note that we had previously established in the proof of Part (i) of Lemma 4.3.4 that for certain values of $v_S \in [0,1]^{b+1}$, $\mathbf{1}(0 \le LC_i^H(v_S) \le 1, \forall i \in S') = 1$. As, $\{v_S : 0 \le LC_{2i}^H(v_S) \le 1, \forall i \in S^-\} \subset \{v_S : 0 \le LC_i^H(v_S) \le 1, \forall i \in S'\}$, for these chosen values of v_S, we have $\mathbf{1}(0 \le LC_{2i}^H(v_S) \le 1, \forall 2i \in S^-) = 1$. Therefore, with this choice of $v_S \in [0,1]^{b+1}$,

$$yLC_{2i}^H(v_S) \le y < 1 \implies F(yLC_{2i}^H(v_S)) \le y.$$

That is, the integral $\int_0^1 \int_0^1 \cdots \int_0^1 \mathbf{1}\big(F(yLC_{2i}^H(v_S)) \le y - \lfloor y \rfloor) \ \forall 2i \in S^-\big) \, dv_S$ is positive.

Hence the proof of Part (i) is complete.

To prove Part (ii), observe that

$$\begin{aligned} &\xi_\pi(i) = \xi_\pi(j) \quad \text{if and only if} \quad t_i = t_j \pmod n \quad \text{and} \\ &\operatorname{sgn}(\pi(i) - \pi(i-1)) = \operatorname{sgn}(\pi(j) - \pi(j-1)) \quad \text{if } i \text{ and } j \text{ are of same parity, or} \\ &\operatorname{sgn}(\pi(i) - \pi(i-1)) = \operatorname{sgn}(\pi(j-1) - \pi(j)) \quad \text{if } i \text{ and } j \text{ are of opposite parity.} \end{aligned}$$

Now, $\Pi_{S_R}(\boldsymbol{\omega}) \subset \Pi_{S_{R^{(s)}}}(\boldsymbol{\omega})$. Therefore, if $\boldsymbol{\omega}$ is a word with b distinct letters but not symmetric, by Part (i), $\frac{1}{p^{r+1}b^{n-r}} \left|\Pi_{S_R}(\boldsymbol{\omega})\right| \to 0$ as $n \to \infty$.

Let $\boldsymbol{\omega} \in S_b(2k)$ with $(r+1)$ even generating vertices. Then $|\mathcal{E}_{\boldsymbol{\omega}}| = r+1$ and $|\mathcal{O}_{\boldsymbol{\omega}}| = b-r$.

Recall the sets $\mathcal{E}_{\boldsymbol{\omega}}, \mathcal{O}_{\boldsymbol{\omega}}, C^e_{i_j}, C^o_{i_j}$ from (6.3.16), (6.3.17), (6.3.18) and (6.3.19). Similarly we can define the functions f^H_n and f^H. Thus we can conclude

$$\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}}\left|\Pi_{S_R}(\boldsymbol{\omega})\right| = \lfloor y \rfloor^{k-(r+1)} + \sum_{\phi \neq S_0 \subset S^-} \lfloor y \rfloor^{|S^- - S_0|} \int_0^1 \int_0^1 \cdots \int_0^1 \mathbf{1}\big(F(yLC^H_{2i}(v_S)) \leq y - \lfloor y \rfloor) \; \forall 2i \in S_0\big) f^H(v_S)\, dv_S, \tag{6.3.30}$$

where $dv_S = \prod_{j=0}^{b} dv_{i_j}$ is the $(b+1)-$dimensional Lebesgue integral on $[0,1]^{b+1}$.

This completes the proof of Part (ii). $\square$

Recall the sequence $a_{2n} = \frac{1}{2}\binom{2n}{n}, n \geq 1$ from Lemma 4.3.2.

Lemma 6.3.5. *(**Circulant matrices,** $C^{(s)}$ **and** C) Suppose $\boldsymbol{\omega}$ is a word of length $2k$ with b distinct letters and $(r+1)$ even generating vertices. Then*

(i) $\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}}|\Pi_{S_{C^{(s)}}}(\boldsymbol{\omega})| = a_{\boldsymbol{\omega}}\big[\lfloor y \rfloor^{k-(r+1)} + \alpha_{C^{(s)}}(\boldsymbol{\omega})\big] > 0$ *if and only if $\boldsymbol{\omega}$ is an even word. Here $a_{\boldsymbol{\omega}}$ is the multiplicative extension of the sequence a_{2n} when $\boldsymbol{\omega}$ is considered as a partition in $\{1, 2, \ldots, 2k\}$.*

(ii) $\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}}|\Pi_{S_C}(\boldsymbol{\omega})| = \lfloor y \rfloor^{k-(r+1)} + \alpha_C(\boldsymbol{\omega}) > 0$ *if and only if $\boldsymbol{\omega}$ is a symmetric word.*

Proof. (i) The proof is very similar to that of Lemmas 4.3.2, 6.3.1 and 6.3.4. So we skip the proof here.

As $n \to \infty$, we have

$$\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}}\left|\Pi_{S_{C^{(s)}}}(\boldsymbol{\omega})\right| = a_{\boldsymbol{\omega}}\Bigg[\lfloor y \rfloor^{k-(r+1)} + \sum_{\phi \neq S_0 \subset S^-} \lfloor y \rfloor^{|S^- - S_0|} \int_0^1 \int_0^1 \cdots \int_0^1 \mathbf{1}\big(F(yLC^T_{2i}(v_S)) \leq y - \lfloor y \rfloor) \; \forall 2i \in S_0\big)\, dv_S\Bigg], \tag{6.3.31}$$

where $\prod_{i=1}^{b} \binom{k_i - 1}{\frac{k_i}{2}} = a_{\boldsymbol{\omega}}$.

The positivity of the integral follows as in the proof of Lemma 6.3.4.

(ii) Using the same arguments as in the proof of Lemma 6.3.2, it follows that

$$\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} |\Pi_{S_C}(\boldsymbol{\omega})| = 0 \ \text{ if } \boldsymbol{\omega} \text{ is not symmetric.}$$

For a symmetric word $\boldsymbol{\omega}$, the computation of its contribution to the limiting moments is similar to that of Lemma 6.3.4 and hence we omit it. Finally we have,

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{p^{r+1}n^{b-r}} \big|\Pi_{S_C}(\boldsymbol{\omega})\big| =& [\lfloor y\rfloor^{k-(r+1)} + \sum_{\phi\neq S_0\subset S^-} \lfloor y\rfloor^{|S^- - S_0|} \\ & \int_0^1\int_0^1\cdots\int_0^1 \mathbf{1}\big(F(yLC_{2i}^T(v_S)) \le y-\lfloor y\rfloor\big)\ \forall 2i\in S_0\big)\ dv_S.\end{aligned} \tag{6.3.32}$$

As in the proof of Part(i) of Lemma 6.3.4, we can show that the integral above is positive.

This completes the proof of the lemma. □

6.3.2 Proof of Theorem 6.1.1

Lemma 6.3.6. ***(Reduction to mean zero)*** *Recall the matrix Z_A from Theorem 6.1.1. Under the assumptions of Theorem 6.1.1, suppose, $\widetilde{Z_A}$ is the $p\times n$ matrix whose entries are $(y_i - \mathbb{E}y_i)$ and thus have mean 0. Then the EESD of S_{Z_A} and $S_{\widetilde{Z_A}}$ are same in the limit.*

Proof. Consider the matrix $\widetilde{Z_A} = \big((y_i-\mathbb{E}y_i)\big)$. That Conditions (6.1.1) and (6.1.2) are true for $\widetilde{Z_A}$ follows from Step 1 of the proofs of Theorem 4.2.2–4.2.4. Similarly, we can show that (6.1.2) is true for $\widetilde{Z_A}$. Thus, Assumption B1 holds for $\widetilde{Z_A}$.

Recall the Lévy metric from Section 2.3. From Lemma 2.3.7,

$$\begin{aligned}& L^4\big(\mathbb{E}F^{S_{Z_A}}, \mathbb{E}F^{S_{\widetilde{Z_A}}}\big) \\ & \le \frac{2}{p^2}(\mathbb{E}\operatorname{Tr}(Z_AZ_A^T + \widetilde{Z_A}\widetilde{Z_A}^T))(\mathbb{E}\operatorname{Tr}[(Z_A-\widetilde{Z_A})(Z_A-\widetilde{Z_A})^T]) \\ & \le \frac{2}{p}\Big(\sum_{i=-(n+p)}^{n+p} cn\mathbb{E}\big(2y_i^2 + (\mathbb{E}y_i)^2 - 2y_{ij}\mathbb{E}y_{ij}\big)\Big)\frac{1}{p}\Big(\sum_{i=-(n+p)}^{n+p} cn(\mathbb{E}y_i)^2\Big),\end{aligned} \tag{6.3.33}$$

where c is a constant depending on the link function of the matrix A. Observe that for all the eight matrices, the second inequality is true due to the structure of the link functions. The second factor of the rhs in (6.3.33) is bounded by

$$\frac{2}{y_n}(n+p)(\sup_i \mathbb{E}y_i)^2 = \frac{2}{y_n}(\sup_i \sqrt{n}\mathbb{E}y_i)^2 + \frac{2}{y_n}(\sup_i \sqrt{p}\mathbb{E}y_i)^2 \to 0,$$

as $n \to \infty$, $p/n \to y > 0$ by (6.1.2). Again, $\mathbb{E}\big[\frac{1}{p}\sum_i y_i^2\big] \to \int_{[0,1]^2} f_2(x,y)\,dx\,dy$. Therefore, the first term of the rhs in (6.3.33) is bounded uniformly. Hence $L^4\big(\mathbb{E}F^{S_{Z_A}}, \mathbb{E}F^{S_{\widetilde{Z_A}}}\big) \to 0$ as $p \to \infty$. This completes the proof of the lemma. □

Lemma 6.3.7. *(**Truncation**) Under the conditions of Theorem 6.1.1, if the EESD of the matrix S_{Z_A} converges weakly to μ_A, then, with the Assumption (6.1.3), the EESD of S_A (where A is the non-truncated version of Z) converges weakly to μ_A.*

Proof. Observe that from Lemma 2.3.7, we have

$$\begin{aligned} L^4(\mathbb{E}F^{S}, \mathbb{E}F^{S_{Z_A}}) &\le \frac{2}{p^2}(\mathbb{E}\,\mathrm{Tr}(AA^T + Z_AZ_A^T))(\mathbb{E}\,\mathrm{Tr}[(A-Z_A)(A-Z_A)^T]) \\ &\le \frac{2}{p}\Bigg(2cn\sum_{i=-(n+p)}^{n+p}\mathbb{E}[y_i^2] + cn\sum_{i=-(n+p)}^{n+p}\mathbb{E}[x_i^2\mathbf{1}_{[|x_i|>r_n]}]\Bigg) \\ &\quad\Bigg(\frac{1}{p}\sum_{i=-(n+p)}^{n+p} cn\mathbb{E}[x_i^2\mathbf{1}_{[|x_i|>r_n]}]\Bigg). \end{aligned} \tag{6.3.34}$$

The second factor in the above inequality tends to zero as $n \to \infty$ from (6.1.3). Again, the first factor is uniformly bounded as in the proof of Lemma 6.3.6. Thus $L^4(\mathbb{E}F^{S_A}, \mathbb{E}F^{S_{Z_A}}) \to 0$ as $p \to \infty$.

This completes the proof of the lemma. □

Now we are ready to prove Theorem 6.1.1.

Proof of Theorem 6.1.1. Note that eight different matrices are involved here. The arguments in the proof for the different matrices are often repetitive. So we prove the theorem for $T^{(s)}$ in details, and omit the elaborate arguments for the other matrices.

(i) Let $A = T^{(s)}$. First observe that from Lemma 6.3.7, it is enough to prove that the EESD of S_{Z_A} converges to some probability distribution $\mu_{T^{(s)}}$. Further, by Lemma 6.3.6 we may assume that $\mathbb{E}(y_i) = 0$. Therefore, it suffices to verify the first moment condition and the Carleman's condition for S_{Z_A}.

As $\mathbb{E}(y_i) = 0$, from (6.2.1), if $\lim_{p\to\infty} \frac{1}{p} \sum_{\pi \in \Pi_{S_{T(s)}}(\boldsymbol{\omega})} \mathbb{E}(Y_\pi)$ exists for every matched word $\boldsymbol{\omega}$ of length $2k$ with b distinct letters and $(r+1)$ even generating vertices ($k \geq 1, 1 \leq b \leq k, 0 \leq r \leq (b-1)$), then the first moment condition would follow.

Suppose $\boldsymbol{\omega}$ is a word, with b distinct letters, $(r+1)$ even generating vertices, and the distinct letters appear $k_1, k_2, \ldots, k_b$ times. Let the jth distinct letter appear at $(\pi(i_j - 1), \pi(i_j))$th position for the first time. Denote $(\pi(i_j - 1), \pi(i_j))$ as (m_j, l_j). Let us now recall $v_i, 1 \leq i \leq 2k$ and $s_i, u_i, 1 \leq i \leq 2k$ as defined in the proof of Lemma 6.3.1.

First, let $\boldsymbol{\omega} \notin E(2k)$. Suppose $\boldsymbol{\omega}$ contains b_1 distinct letters that appear even number of times and b_2 distinct letters that appear odd number of times where $b = b_1 + b_2$. So we assume that for each $\pi \in \Pi(\boldsymbol{\omega})$, k_{j_p}, $1 \leq p \leq b_1$ are even, and k_{j_q}, $b_1 + 1 \leq q \leq b_1 + b_2$ are odd. Hence the contribution of this $\boldsymbol{\omega}$ to the limiting kth moment is as follows:

$$\frac{1}{pn^{b_1+b_2-\frac{1}{2}}} \sum_{S} \prod_{p=1}^{b_1} f_{k_{j_p}}(|s_{j_p}|) \prod_{q=b_1+1}^{b_1+b_2} n^{\frac{b_2-1/2}{b_2}} \mathbb{E}\Big[y_{|s_{j_q}|}^{k_{j_q}}\Big], \tag{6.3.35}$$

where S is the set of distinct generating vertices for $\boldsymbol{\omega}$.
For n large, $n^{\frac{b_2-1/2}{b_2}} \mathbb{E}[y_{(s_{j_q}-2) \pmod n}^{k_{j_q}}] < 1$ for any $b_1+1 \leq q \leq b_1+b_2$ and $\prod_{p=1}^{b_1} f_{k_{j_p}}(|s_{j_p}|) \leq M$ (independent of n). Now as $\boldsymbol{\omega} \notin E_b(2k)$ and $p/n \to y > 0$, from Lemma 6.3.1 we have, $|S| \leq b$. Hence, as $p, n \to \infty$ and $p/n \to y > 0$, (6.3.35) goes to 0. Thus any word that is not even, contributes 0 to the limiting moments.

Now let $\boldsymbol{\omega} \in E_b(2k)$. Let $|\mathcal{E}_{\boldsymbol{\omega}}|$ and $|\mathcal{O}_{\boldsymbol{\omega}}|$ be as in (6.3.16) and (6.3.17). Clearly, as observed in Lemma 6.3.1, there are $\prod_{i=1}^{b} \binom{k_i - 1}{\frac{k_i}{2}}$ combination of equations for the s_j's (and hence v_j's) for determining the non-generating vertices, once the generating vertices are chosen. Let us denote a generic combination of the v_j's by $LC_i^T(v_S) \in LC_{\boldsymbol{\omega}}^T$ (see (4.3.19)). For each of the combination of equations, we get positive (possibly different) contribution (see Lemma 4.3.3 and Theorem 4.2.4). Then the contribution of each combination LC_i^T corresponding to the word $\boldsymbol{\omega}$ is

$$y_n^r \frac{1}{p^{r+1} n^{b-r}} \sum_{S} \prod_{j=1}^{b} \Big(\prod_{i_j \in \mathcal{E}_{\boldsymbol{\omega}}} f_{k_j,n}(|v_{m_j} - y_n v_{l_j}|) \prod_{i_j \in \mathcal{O}_{\boldsymbol{\omega}}} f_{k_j,n}(|y_n v_{m_j} - v_{l_j}|) \Big)$$
$$\mathbf{1}(0 \leq LC_i^T(v_S) \leq 1, \ \forall\, i \in S'), \tag{6.3.36}$$

where S is the set of distinct generating vertices and S' is the set of indices of the non-generating vertices of $\boldsymbol{\omega}$. By abuse of notation let m_1 and $l_j, 1 \le j \le b$ denote the indices of the generating vertices. Therefore as $p \to \infty$, the contribution of $\boldsymbol{\omega}$ is

$$y^r \sum_{LC_i^T \in LC_{\boldsymbol{\omega}}^T} \int_0^1 \int_0^1 \int_0^1 \cdots \int_0^1 \prod_{j=1}^{b} \Big(\prod_{i_j \in \mathcal{E}_{\boldsymbol{\omega}}} f_{k_j}\left(|x_{m_j} - y x_{l_j}|\right) \prod_{i_j \in \mathcal{O}_{\boldsymbol{\omega}}} f_{k_j}\left(|y x_{m_j} - x_{l_j}|\right) \Big)$$
$$\mathbf{1}(0 \le LC_i^T(x_S) \le 1, \ \forall\, i \in S') \, dx_S, \tag{6.3.37}$$

where $dx_S = dx_{m_1} dx_{l_1} \cdots dx_{l_b}$ denotes the $(b+1)$-dimensional Lebesgue measure on $[0,1]^{b+1}$ and $0 < y = \lim p/n$. As for each $k \ge 1$, there are finitely many even words, the first moment condition is established.

Hence we have,

$$\lim_{n\to\infty} \frac{1}{p} \mathbb{E}[\mathrm{Tr}(S_{Z_A})^k] = \sum_{b=1}^{k} \sum_{r=0}^{b-1} \sum_{\substack{\sigma \in E_b(2k) \text{ with} \\ (r+1) \text{ even generating vertics}}} y^r \sum_{LC_i^T \in LC_\sigma^T} \int_0^1 \int_0^1 \int_0^1 \cdots \int_0^1$$
$$\prod_{j=1}^{b} \Big(\prod_{i_j \in \mathcal{E}_{\boldsymbol{\omega}}} f_{k_j}\left(|x_{m_j} - y x_{l_j}|\right) \prod_{i_j \in \mathcal{O}_{\boldsymbol{\omega}}} f_{k_j}\left(|y x_{m_j} - x_{l_j}|\right) \Big) \mathbf{1}(0 \le LC_i^T(x_S) \le 1, \ \forall\, i \in S') \, dx_S. \tag{6.3.38}$$

Now we show that $\gamma_k = \lim_{n\to\infty} \frac{1}{p} \mathbb{E}[\mathrm{Tr}(S_{Z_A})^k], k \ge 1$ determines a unique distribution.

If $y \le 1$,

$$\gamma_k = \lim_{n\to\infty} \frac{1}{p} \mathbb{E}[\mathrm{Tr}(S_{Z_A})^k] \le \sum_{\sigma \in E(2k)} M_\sigma \le \sum_{\sigma \in \mathcal{P}(2k)} M_\sigma = \alpha_k.$$

As $\{\alpha_k\}$ satisfies Carleman's condition (Assumption B1), $\{\gamma_k\}$ does so. Hence the sequence of moments $\{\gamma_k\}$ determines a unique distribution.

If $y > 1$, $y^r \le y^b, 0 \le r \le b, 1 \le b \le k$ and hence

$$\gamma_k = \lim_{n\to\infty} \frac{1}{p} \mathbb{E}[\mathrm{Tr}(S_Z)^k] \le \sum_{\sigma \in E(2k)} y_\sigma M_\sigma \le \sum_{\sigma \in \mathcal{P}(2k)} y^{2k} M_\sigma = y^{2k} \alpha_k.$$

As, $y \in (1, \infty)$ and α_k satisfies Carleman's condition, $\{\gamma_k\}$ does so. Hence the sequence of moments $\{\gamma_k\}$ determines a unique distribution.

Therefore, there exists a measure $\mu_{T^{(s)}}$ with moment sequence $\{\gamma_k\}$ such that EESD of S_{Z_A} converges weakly to $\mu_{T^{(s)}}$, whose moments are given as in (6.3.38).

This completes the proof of Part (i).

(ii) Let $A = T$. Just as in Part (i), it suffices to verify the first moment condition and the Carleman's condition for S_{Z_A}. The proof is very similar to Part (i). So we omit the details and give the limiting moment formula below, with the notations from Part (i).

For $k \geq 1$,

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}[\mathrm{Tr}(S_{Z_A})^k] = \sum_{b=1}^{k}\sum_{r=0}^{b-1}\sum_{\substack{\sigma\in S_b(2k)\text{ with}\\(r+1)\text{ even generating vertics}}} y^r \int_{[0,1]^{b+1}}\prod_{j=1}^{b}\Big(\prod_{i_j\in\mathcal{E}_\omega} f_{k_j}(yx_{l_j}-x_{m_j})\prod_{i_j\in\mathcal{O}_\omega} f_{k_j}(yx_{m_j}-x_{l_j})\Big) \mathbf{1}(0\leq LC_i^T(x_S)\leq 1,\ \forall\ i\in S')\ dx_S. \tag{6.3.39}$$

As $S(2k)\subset E(2k)$ and the integrand in (6.3.39) is bounded, the same arguments as in Part (i) are applicable. Thus the Carleman's condition for S_{Z_A} is satisfied.

Therefore, there exists a measure μ_T with moment sequence $\{\gamma_k\}$ such that EESD of S_{Z_A} converges weakly to μ_T. This proves the theorem for (asymmetric) Toeplitz matrix.

Since the arguments of the other matrices are similar to the previous parts, we omit the proofs and describe only the limiting moments.

(iii) Let $A = H^{(s)}$. The moments of the LSD $\mu_{H^{(s)}}$ are given by

$$\beta_k(\mu_{H^{(s)}}) = \sum_{b=1}^{k}\sum_{r=0}^{b-1}\sum_{\substack{\sigma\in S_b(2k)\text{ with}\\(r+1)\text{ even generating vertics}}} y^r \int_{[0,1]^{b+1}}\prod_{j=1}^{b}\Big(\prod_{i_j\in\mathcal{E}_\omega} f_{k_j}(x_{m_j}+yx_{l_j})\prod_{i_j\in\mathcal{O}_\omega} f_{k_j}(yx_{m_j}+x_{l_j})\Big) \mathbf{1}(0\leq LC_i^H(x_S)\leq 1,\ \forall\ i\in S')\ dx_S. \tag{6.3.40}$$

(iv) Let $A = H$. The moments of the LSD μ_H are given by

$$\beta_k(\mu_H) = \sum_{b=1}^{k}\sum_{r=0}^{b-1} \sum_{\substack{\sigma\in S_b(2k) \text{ with} \\ (r+1) \text{ even generating vertics}}} y^r$$
$$\int_0^1\int_0^1\int_0^1 \cdots \int_0^1 \prod_{j=1}^{b}\Big(\prod_{i_j\in\mathcal{E}_{\boldsymbol{\omega}}} f_{k_j}\big(\operatorname{sgn}(yx_{l_j}-x_{m_j})(x_{m_j}+yx_{l_j})\big)$$
$$\prod_{i_j\in\mathcal{O}_{\boldsymbol{\omega}}} f_{k_j}\big(\operatorname{sgn}(yx_{m_j}-x_{l_j})(yx_{m_j}+x_{l_j})\big)\Big)$$
$$\mathbf{1}(0\le LC_i^H(x_S)\le 1,\ \forall\, i\in S')f^H(x_S)\ dx_S. \tag{6.3.41}$$

(v) Let $A = R^{(s)}$.

For any $m \geq 1$, let

$$h_{2m,n}(x_1,x_2) = f_{2m,n}(x_1+x_2)\mathbf{1}(0\le x_1+x_2\le 1) + f_{2m,n}(x_1+x_2-1)\mathbf{1}(x_1+x_2>1),$$
$$h_{2m}(x_1,x_2) = f_{2m}(x_1+x_2)\mathbf{1}(0\le x_1+x_2\le 1) + f_{2m}(x_1+x_2-1)\mathbf{1}(x_1+x_2>1). \tag{6.3.42}$$

Then the moments of the LSD $\mu_{R^{(s)}}$ are given by (see Lemma 6.3.4),

$$\beta_k(\mu_{R^{(s)}}) = \sum_{b=1}^{k}\sum_{r=0}^{b-1} \sum_{\substack{\sigma\in S_b(2k) \text{ with} \\ (r+1) \text{ even generating vertics}}} y^r$$
$$\Bigg[\lfloor y\rfloor^{k-(r+1)}\int_{[0,1]^{b+1}}\prod_{j=1}^{b}\Big(\prod_{i_j\in\mathcal{E}_{\boldsymbol{\omega}}} h_{k_j}(x_{m_j},yx_{l_j})\prod_{i_j\in\mathcal{O}_{\boldsymbol{\omega}}} h_{k_j}(yx_{m_j},x_{l_j})\Big)\ dx_S$$
$$+\sum_{\phi\ne S_0\subset S^-}\lfloor y\rfloor^{|S^- - S_0|}\int_{[0,1]^{b+1}}\prod_{j=1}^{b}\Big(\prod_{i_j\in\mathcal{E}_{\boldsymbol{\omega}}} h_{k_j}(x_{m_j},yx_{l_j})$$
$$\prod_{i_j\in\mathcal{O}_{\boldsymbol{\omega}}} h_{k_j}(yx_{m_j},x_{l_j})\Big)\mathbf{1}(F(yLC_{2i}^H(x_S))\le y-\lfloor y\rfloor,\ \forall\, 2i\in S_0)\ dx_S\Bigg]. \tag{6.3.43}$$

(vi) Let $A = R$. Suppose

$$\tilde{h}_{2m}(x,x_2) = f_{2m}(\operatorname{sgn}(x_2-x_1)(x_1+x_2))\mathbf{1}(0\le x_1+x_2\le 1)+$$
$$f_{2m}(\operatorname{sgn}(x_2-x_1)(x_1+x_2-1))\mathbf{1}(x_1+x_2>1). \tag{6.3.44}$$

Then the moments of the LSD μ_R are as follows (see (6.3.30)):

$$\begin{aligned}
\beta_k(\mu_R) = &\sum_{b=1}^{k}\sum_{r=0}^{b-1} \sum_{\substack{\sigma\in S_b(2k) \text{ with} \\ (r+1) \text{ even generating vertics}}} y^r \\
&\Bigg[\lfloor y\rfloor^{k-(r+1))} \int_{[0,1]^{b+1}} \prod_{j=1}^{b} \Bigg(\prod_{i_j\in\mathcal{E}_{\boldsymbol{\omega}}} \tilde{h}_{k_j}(x_{m_j}, yx_{l_j}) \\
&\prod_{i_j\in\mathcal{O}_{\boldsymbol{\omega}}} \tilde{h}_{k_j}(yx_{m_j}, x_{l_j})\Bigg) f^H(x_S)\, dx_S + \sum_{\phi\neq S_0\subset S^-} \lfloor y\rfloor^{|S^- - S_0|} \\
&\int_{[0,1]^{b+1}} \prod_{j=1}^{b} \Bigg(\prod_{i_j\in\mathcal{E}_{\boldsymbol{\omega}}} \tilde{h}_{k_j}(x_{m_j}, yx_{l_j}) \prod_{i_j\in\mathcal{O}_{\boldsymbol{\omega}}} \tilde{h}_{k_j}(yx_{m_j}, x_{l_j})\Bigg) \\
&\mathbf{1}(F(yLC_{2i}^H(x_S)) \le y - \lfloor y\rfloor,\ \forall\ 2i\in S_0) f^H(x_S)\, dx_S\Bigg]. \qquad (6.3.45)
\end{aligned}$$

(vii) Let $A = C^{(s)}$. The moments of the LSD $\mu_{C^{(s)}}$ are given by

$$\begin{aligned}
\beta_k(\mu_{C^{(s)}}) = &\sum_{b=1}^{k}\sum_{r=0}^{b-1} \sum_{\substack{\sigma\in E_b(2k) \text{ with} \\ (r+1) \text{ even generating vertics}}} y^r a_\sigma \\
&\Bigg[\lfloor y\rfloor^{k-(r+1)} \int_{[0,1]^{b+1}} \prod_{j=1}^{b} \Bigg(\prod_{i_j\in\mathcal{E}_{\boldsymbol{\omega}}} f_{k_j}\big(|1/2 - |1/2 - |x_{m_j} - yx_{l_j}|||\big) \\
&\prod_{i_j\in\mathcal{O}_{\boldsymbol{\omega}}} f_{k_j}\big(|1/2 - |1/2 - |yx_{m_j} - x_{l_j}|||\big)\Bigg) + \sum_{\phi\neq S_0\subset S^-} \lfloor y\rfloor^{|S^- - S_0|} \\
&\int_{[0,1]^{b+1}} \prod_{j=1}^{b} \Bigg(\prod_{i_j\in\mathcal{E}_{\boldsymbol{\omega}}} f_{k_j}\big(|1/2 - |1/2 - |x_{m_j} - yx_{l_j}|||\big) \\
&\qquad \prod_{i_j\in\mathcal{O}_{\boldsymbol{\omega}}} f_{k_j}\big(|1/2 - |1/2 - |yx_{m_j} - x_{l_j}|||\big)\Bigg) \\
&\qquad \mathbf{1}\big(F(yLC_{2i}^T(x_S)) \le y - \lfloor y\rfloor, \forall 2i\in S_0\big)\Bigg]. \qquad (6.3.46)
\end{aligned}$$

(viii) Let $A = C$. Suppose

$$\eta_{2m}(yx_1, x_2) = f_{2m}(x_2 - yx_1)\mathbf{1}(0\le x_2 - yx_1 \le 1) + f_{2m}(1 - x_2 + yx_1)\mathbf{1}(x_2 - yx_1 < 0), \qquad (6.3.47)$$

where $(x_1, x_2) \in [0,1]^2$. The moments of the LSD μ_C are given by

$$\beta_k(\mu_C) = \sum_{b=1}^{k} \sum_{r=0}^{b-1} \sum_{\substack{\sigma \in S_b(2k) \text{ with} \\ (r+1) \text{ even generating vertics}}} y^r a_\sigma$$
$$\Bigg[\lfloor y \rfloor^{k-(r+1)} \int_0^1 \cdots \int_0^1 \prod_{j=1}^{b} \Big(\prod_{i_j \in \mathcal{E}_{\boldsymbol{\omega}}} \eta_{k_j}(x_{m_j} - y x_{l_j})$$
$$\prod_{i_j \in \mathcal{O}_{\boldsymbol{\omega}}} \eta_{k_j}(x_{l_j} - y x_{m_j}) \Big) + \sum_{\phi \neq S_0 \subset S^-} \lfloor y \rfloor^{|S^- - S_0|}$$
$$\int_{[0,1]^{b+1}} \prod_{j=1}^{b} \Big(\prod_{i_j \in \mathcal{E}_{\boldsymbol{\omega}}} \eta_{k_j}(x_{m_j} - y x_{l_j}) \prod_{i_j \in \mathcal{O}_{\boldsymbol{\omega}}} \eta_{k_j}(x_{l_j} - y x_{m_j}) \Big)$$
$$\mathbf{1}\big(F(yLC_{2i}^T(x_S)) \le y - \lfloor y \rfloor, \forall 2i \in S_0\big) \Bigg]. \tag{6.3.48}$$

□

6.4 Some Corollaries

As the entries are dependent on i, j, n, the formula for the limiting moments, as derived in (6.3.38), (6.3.39), (6.3.40), (6.3.41), (6.3.43), (6.3.45), (6.3.46) and (6.3.48) can be very complicated. Here we discuss a few special cases where the limiting moment formulae are relatively simple. These special cases would be when the entries of the matrix A are– (a) triangular i.i.d. (Corollary 6.4.1), (b) sparse triangular i.i.d. (Corollary 6.4.4), (c) fully i.i.d. with finite mean and variance (Corollary 6.4.5), (d) have a non-trivial variance structure (Corollary 6.4.6), (e) triangular, i.e., only lower triangular entries are non-zero (Corollary 6.4.8) and (f) have a band structure (Corollary 6.4.10).

6.4.1 General triangular i.i.d. entries

Corollary 6.4.1. *Let A be one of the eight $p \times n$ patterned matrices $R^{(s)}, R, C^{(s)}, C, T^{(s)}, T, H^{(s)}, H$. Suppose the input sequence $\{x_{i,n}\}$ are i.i.d. for every fixed n, with all moments finite. Also assume that*

(i) for all $k \ge 1$,

$$n\mathbb{E}[x_{0,n}^k] \to c_k \quad \text{as} \quad n \to \infty, \tag{6.4.1}$$

(ii) $\gamma_k = \sum_{\pi \in E(2k)} c_\pi$ *satisfies Carleman's condition.*

Then the EESD of $S_A = AA^T$ *converges weakly to a non-random probability distribution, say,* μ_A *for each of the matrices* A.

Proof. Observe that Assumption B1 (i), (ii) and (iii) are satisfied with $r_n = \infty$ and $f_{2k} \equiv c_{2k}$ for $k \geq 1$. Thus Theorem 6.1.1 can be applied to conclude that the EESD of S_A converges to a probability distribution, μ_A. A brief description of the limiting moments is given below.

(i) Suppose $A = T^{(s)}$ whose entries satisfy (6.4.1). By Part (i) of Theorem 6.1.1, the EESD of $S_{T_p^{(s)}}$ converges to $\mu_{T^{(s)}}$ with moment sequence as follows (see (6.3.38)):

$$\beta_k(\mu_{T^{(s)}}) = \sum_{b=1}^{k} \sum_{r=0}^{b} y^r \sum_{\substack{\pi \in E_b(2k) \text{ with} \\ (r+1) \text{ even generating vertices}}} \sum_{LC_\pi^T} c_\pi \int_{[0,1]^{b+1}} \mathbf{1}(0 \leq LC_i^T(x_S) \leq 1, \forall i \in S')\, dx_S. \tag{6.4.2}$$

Note that every word with b distinct letters can be identified as a partition with b blocks, see Section 2.4. Therefore, for every $\pi \in E_b(2k)$, $LC_\pi^T = LC_{\boldsymbol{\omega}}^T$ for the corresponding even word with b distinct letters.

(ii) $A = T$. By Part (ii) of Theorem 6.1.1 (see (6.3.39)):

$$\beta_k(\mu_T) = \sum_{b=1}^{k} \sum_{r=0}^{b} y^r \sum_{\substack{\pi \in S_b(2k) \text{ with} \\ (r+1) \text{ even generating vertices}}} c_\pi \int_{[0,1]^{b+1}} \mathbf{1}(0 \leq LC_i^T(x_S) \leq 1, \forall i \in S')\, dx_S. \tag{6.4.3}$$

(iii) $A = H^{(s)}$. By Part (iii) of Theorem 6.1.1, the EESD of $S_{H^{(s)}}$ converges to $\mu_{H^{(s)}}$ whose moment sequence is as in (6.4.3), where the integrand is replaced by $\mathbf{1}(0 \leq LC_i^H(x_S) \leq 1, \forall i \in S')$ (see (6.3.40)).

(iv) $A = H$. By Part (iv) of Theorem 6.1.1, the EESD of S_{H_p} converges to μ_H whose moment sequence is as in (6.4.3), where the integrand is replaced by $\mathbf{1}(0 \leq LC_i^H(x_S) \leq 1, \forall i \in S')f^H(s)$ (see (6.3.41)).

(v) $A = R^{(s)}$. By Part (v) of Theorem 6.1.1, the EESD of $S_{R_p^{(s)}}$ converges to $\mu_{R^{(s)}}$ with moment sequence as in (6.4.3), and the function inside the square brackets (see (6.3.43))

$$c_\pi\big[\lfloor y\rfloor^{k-(r+1)} + \sum_{\phi\neq S_0\subset S^-}\lfloor y\rfloor^{|S^-\setminus S_0|}\int_{[0,1]^{b+1}}\mathbf{1}(F(yLC_{2i}^H(x_S))\leq y-\lfloor y\rfloor, \forall 2i\in S_0)\ dx_S\big].$$

(vi) $A = R$. By Part (vi) of Theorem 6.1.1, $\beta_k(\mu_R)$ is same as $\beta_k(\mu_{R^{(s)}})$, with an extra factor $f^H(s)$ in the integrand.

(vii) $A = C^{(s)}$. By Part (vii) of Theorem 6.1.1, the limit in this case is $\mu_{C^{(s)}}$ with moment sequence as in (6.4.2), and the function inside the square brackets (see (6.3.46))

$$c_\pi\big[\lfloor y\rfloor^{k-(r+1)}+$$
$$\sum_{\phi\neq S_0\subset S^-}\lfloor y\rfloor^{|S^-\setminus S_0|}\int_{[0,1]^{b+1}}\mathbf{1}(F(yLC_{2i}^T(x_S))\leq y-\lfloor y\rfloor, \forall 2i\in S_0)f^H(x_S)\ dx_S\big].$$

(viii) $A = C$. By Part (viii) of Theorem 6.1.1, $\beta_k(\mu_C)$ is as in (6.4.3), where the function inside the square brackets is (see (6.3.48))

$$c_\pi\big[\lfloor y\rfloor^{k-(r+1)}+$$
$$\sum_{\phi\neq S_0\subset S^-}\lfloor y\rfloor^{|S^-\setminus S_0|}\int_{[0,1]^{b+1}}\mathbf{1}(F(yLC_{2i}^T(x_S))\leq y-\lfloor y\rfloor, \forall 2i\in S_0)f^H(x_S)\ dx_S\big].$$

□

Remark 6.4.2. *(a) The linear combinations LC_i^T and LC_i^H from Lemma 6.3.1, 6.3.2 and 6.3.3 play crucial role in the moments of the LSD of S_A. For S_T and $S_{H^{(s)}}$, only symmetric words can contribute positively to the limiting moments (see Lemmas 6.3.2 and 6.3.3). Now from the proof of Part (i) of Lemma 6.3.2 and Lemma 6.3.3, it follows that for every $i\in S'$, $LC_i^T(v_S) = LC_{i,q}^H(z_S)$. As the z_is are derived by elementary transformations, we have $LC_i^T(v_S) = LC_i^H(v_S)$, for any symmetric word. This will be useful in finding relations between $\mu_T, \mu_{H^{(s)}}, \mu_{R^{(s)}}, \mu_C$. We discuss this next.*

(b) Suppose $y \in \mathbb{N}$. Then, the integrals in Parts (v) and (viii) above are zero and

$$\beta_k(\mu_{R^{(s)}}) = \beta_k(\mu_C) = \sum_{\pi \in S(2k)} y^{k-1} c_\pi.$$

We can say more about these limits even when $y \notin \mathbb{N}$, using Part (a). Recall that the words contributing to the limiting moments for $S_T, S_{H^{(s)}}, S_{R^{(s)}}$ ans S_C are symmetric. Now, from (a) and uniqueness of the limit, it is easy to see that when the variables are triangular i.i.d. and satisfy (6.4.1), we have $\mu_T = \mu_{H^{(s)}}$ and $\mu_{R^{(s)}} = \mu_C$.

Remark 6.4.3. *However, in general the LSDs of S_A for symmetric and the asymmetric cases are not identical. For instance, for the Toeplitz and the circulant matrices, this is evident from the moment formula, as the set of partitions that contribute positively to the limiting moments are different in the two cases. For the Hankel and the reverse circulant, there is an extra factor in the integrand for the asymmetric versions, and this gives rise to the difference in the limit. We illustrate this for $S_{H^{(s)}}$ and S_H below.*

For all words that are special symmetric, the contributions to the symmetric and asymmetric Hankel are same as there are no further restrictions for the signs arising from (6.3.20). However, if $\boldsymbol{\omega} \in S_b(2k) \setminus SS_b(2k)$, some additional conditions do appear in case of asymmetric Hankel.

For instance, let us consider the word abcabc $\in S_3(6) \setminus SS_3(6)$. In case of symmetric Hankel, its contribution to $\mu_{H^{(s)}}$ is

$$C_2^3 \int_0^1 \int_0^1 \int_0^1 \int_0^1 \mathbf{1}(0 \le x_0 + x_1 - x_3, x_2 - x_0 + x_3 \le 1)\ dx_0 dx_1 dx_2 dx_3. \tag{6.4.4}$$

On the other hand, the contribution of abcabc (in case of asymmetric Hankel) to μ_H is

$$\begin{aligned} C_2^3 \int_0^1 \int_0^1 \int_0^1 \int_0^1 \mathbf{1}(0 \le x_0 + x_1 - x_3, x_2 - x_0 + x_3 \le 1)\mathbf{1}\big(sgn(x_1 - x_0) = \\ sgn(2x_3 - x_0 - x_1), sgn(x_1 - x_2) = sgn(x_2 - 2x_0 - x_1 + 2x_3), \\ sgn(x_3 - x_2) = sgn(x_2 - 2x_0 + x_3)\big)\ dx_0 dx_1 dx_2 dx_3. \end{aligned} \tag{6.4.5}$$

The integrand in (6.4.5) is less than that in (6.4.4) due to the extra restrictions arising from the sign functions. Thus, the kth moment of μ_H is in general smaller than that of $\mu_{H^{(s)}}$. A very similar thing occurs in case of $\mu_{R^{(s)}}$ and μ_R.

The next corollary deals with sparse matrices A. This serves as a special case of the general triangular i.i.d. case.

Corollary 6.4.4. *(**Sparse entries**) Suppose the input sequence for the matrix A (where A is any one of the eight matrices), $\{x_{i,n}\}$ are i.i.d. Ber(p_n) where $np_n \to \lambda > 0$. Then the EESD of $S_A = AA^T$ converges weakly to a probability distribution whose moments are determined by λ.*

Proof. Observe that (6.4.1) is satisfied with $c_k = \lambda$ for all $k \geq 1$. Also condition (ii) of Corollary 6.4.1 is satisfied. Therefore from Corollary 6.4.1, the EESD of $S_A = AA^T$ converges weakly to say μ_A whose moments are as in (i)-(viii) of Corollary 6.4.1, where $c_\pi = \lambda^{|\pi|}$ for all $\pi \in \mathcal{P}(2k)$. □

6.4.2 Fully i.i.d. entries

Theorem 6.1.1 concludes the convergence of the EESD of S_A. However, as we will see in the upcoming corollaries, almost sure convergence of the ESD can be obtained only in some cases. To establish the almost sure convergence of the ESD, we will use Lemma 2.1.3, just as we did in case of the S matrix. Recall the set $\tilde{Q}^b_{k,4}$ from (5.4.2) that was used to establish the fourth moment condition for S. Analogous version of Lemma 5.4.3 is not true for S_A. However, it can be shown that

$$|\tilde{Q}^b_{k,4}| \leq n^{2k+2} \quad \text{for any } 1 \leq b \leq 2k. \tag{6.4.6}$$

This was proved for the single symmetric matrices $R^{(s)}, C^{(s)}, T^{(s)}$ and $H^{(s)}$ in Lemma 4.4.1. The arguments in that proof can be used for the S_A matrices here as $1 \leq \pi(2i) \leq p$ and $1 \leq \pi(2i-1) \leq n$, and p and n are comparable for large n. We omit the details.

Corollary 6.4.5. *Result 6.0.1 follows from Theorem 6.1.1.*

Proof. Suppose $A = \left(\left(\frac{x_{L(i,j)}}{\sqrt{n}}\right)\right)$ (for the corresponding link function L for each of the matrices) where $\{x_i\}$ are i.i.d. with distribution F which has mean zero and variance 1. First, let us verify that the conditions of Assumption B1 are satisfied in this case. Towards that, let $r_n = n^{-1/3}$. Using the same line of reasoning as in Corollary 5.5.1, it follows that $g_2 \equiv 1$ and $g_{2k} \equiv 0, k > 1$. Thus $M_2 = 1$, $M_{2k} = 0, k \geq 2$ (see (iii) in Assumption B1) and $\alpha_k = \sum_{\sigma \in \mathcal{P}(2k)} 1$ clearly satisfies Carleman's condition. Also (6.3.7) can be verified similarly as Corollary 5.5.1. Then by Theorem 6.1.1, the EESD of S_A converges weakly to a probability distribution μ_A for each of the matrices A.

The moment formulae are given as in (i)-(viii) in Corollary 6.4.1, where $c_2 = 1$ and $c_{2k} = 0$ for all $k \geq 2$. The words that contribute to the limiting moments are now pair matched.

Now, as S_A satisfies (6.4.6), we have

$$\begin{aligned} \frac{1}{p^4}\mathbb{E}\big[\operatorname{Tr}(S_A^k) - \mathbb{E}(\operatorname{Tr}(S_A^k))\big]^4 &= \mathcal{O}(p^{-2}) \quad \text{and therefore,} \\ \sum_{p=1}^{\infty} \frac{1}{p^4}\mathbb{E}\big[\operatorname{Tr}(S_A^k) - \mathbb{E}(\operatorname{Tr}(S_A^k))\big]^4 &< \infty \quad \text{for every } k \geq 1. \end{aligned} \tag{6.4.7}$$

Then using Lemma 2.1.3, we can conclude that μ_{S_A} converges almost surely. This yields Result 6.0.1 as a special case of Theorem 6.1.1. $\square$

6.4.3 Matrices with variance profile

The next corollary deals with the case of a variance profile, see (4.4.9).

Corollary 6.4.6. ***(Matrices with variance profile)*** *Suppose the input sequence of A (where A is one of the eight matrices) is $\{y_{i,n}\} = \{\sigma(i/n)x_{i,n}; i \geq 0\}$, where $\sigma : [0,1] \to \mathbb{R}$ is a bounded and Riemann integrable function and $\{x_{i,n}; i \in \mathbb{Z}\}$ are i.i.d. random variables with mean zero and all moments finite, and satisfy Conditions (i) and (ii) of Corollary 6.4.1. Then the EESD of S_A for each of the eight patterns of A, converges to a probability distribution μ_A whose moments are determined by σ, y and $\{c_{2k}, k \geq 1\}$.*

Proof. First observe that the entries of A satisfy Assumption B1 (i) and (ii) with $r_n = \infty$, $f_{2k} = \sigma^{2k} c_{2k},\ k \geq 1$. Since σ is bounded, Assumption B1 (iii) is also true. Hence from Theorem 6.1.1, we can conclude that the EESD of S_A converges weakly to μ_A.

From (6.3.38), (6.3.39), (6.3.40), (6.3.41), (6.3.43), (6.3.45), (6.3.46), (6.3.48) and $f_{2k} = \sigma^{2k} c_{2k},\ k \geq 1$, we see that the moments of the limiting distribution are indeed determined by σ and $\{c_{2k}, k \geq 1\}$. $\square$

Remark 6.4.7. *Note that in Corollary 6.4.6, if each $x_{i,n}$ has the same distribution F for all i and n, then, $c_{2k} = 0$ for all $k \geq 2$. Hence the EESD of $\frac{1}{n}S_A$ converges. As σ is* *bounded,* (6.4.6) *and hence* (6.4.7) *hold true. Thus we can conclude that μ_{S_A} converges almost surely to the respective limits μ_A.*

6.4.4 Triangular matrices

Triangular Matrices: As discussed in Chapter 5, the LSD of XX^T has been studied in Dykema and Haagerup [2004], where the entries of the upper triangular matrix X are i.i.d. Gaussian. Later LSD results were proved in Basu et al. [2012] for symmetric triangular matrices with other patterns such as Hankel, Toeplitz and symmetric circulant, and with i.i.d. input. The matrices in Basu et al. [2012] had the entries $y_{L(i,j),n} = x_{L(i,j),n}\mathbf{1}(i+j \leq n+1)$. However, the matrix considered in Dykema and Haagerup [2004] is upper triangular, as in (5.1.2) and not symmetric. It is natural to ask what happens to such matrices when there are other patterns involved.

Let A be any of the eight matrices that are being discussed in this chapter. Let A^U be the matrix whose entries $y_{L(i,j),n}$ are as follows:

$$y_{L(i,j),n} = \begin{cases} x_{L(i,j),n} & \text{if } i \leq j, \\ 0 & \text{otherwise.} \end{cases} \tag{6.4.8}$$

Then we have the following result.

Corollary 6.4.8. *Consider the matrices A^U as defined in* (6.4.8)*. Assume that the variables $\{x_{i,n}; i \geq 0\}$ in* (6.4.8) *are i.i.d. random variables with all moments finite, for every fixed n, and satisfy Conditions (i) and (ii) of Corollary 6.4.1. Then, for each of the eight matrices mentioned above, the EESD of S_{A^U} converges to some probability measure μ_{A^U} that depends on $\{c_{2k}\}_{k\geq 1}$.*

The proof is very similar to that of Corollary 4.4.21, so we skip it.

Remark 6.4.9. *If the entries of A^U are $\frac{y_{i,n}}{\sqrt{n}}$ where $\{y_{i,n}; i \geq 0\}$ are as in* (6.4.8) *and $\{x_{i,n}; i \geq 0\}_{n\geq 1}$ are i.i.d. random variables with mean 0 and variance 1, then using familiar truncation arguments (as in Corollary 6.4.5), the variables $\{y_{i,n}; i \geq 0\}$ can be assumed to be uniformly bounded and hence satisfy Conditions (i) and (ii) of Corollary 6.4.1 with $c_2 = 1$ and $c_{2k} = 0$ for $k \geq 2$. Hence from Corollary 6.4.8, we obtain the* *convergence of the EESD. Again it can be verified that (6.4.7) is true in this case. Thus the ESD of S_{A^U} converges weakly almost surely to a non-random probability measure.*

6.4.5 Band matrices

In Corollaries 4.4.14 and 4.4.17, the LSD of symmetric band matrices where the non-zero entries satisfy (4.4.3) have been discussed. So it was natural to ask what happens to the LSD of the $A^b A^{bT}, A^B A^{BT}$, where A^b and A^B are matrices with entries $y_{L(i,j)} = x_{L(i,j)}\mathbf{1}(L(i,j) \leq m_n)$ and $y_{L(i,j)} = x_{L(i,j)}[\mathbf{1}(L(i,j) \leq m_n) + \mathbf{1}(L(i,j) \geq n - m_n)$ (see Definitions 4.4.18, 4.4.19 and 4.4.20). Here we provide an answer to that question.

Corollary 6.4.10. *Consider the matrices A^b and A^B. Assume that the variables $\{x_{i,n}; i \geq 0\}$ associated with the matrices A^b and A^B (as in (4.4.18), (4.4.19) and (4.4.20)) are i.i.d. random variables with all moments finite, for every fixed n, and satisfy (6.4.1). Suppose $\alpha = \lim_{n\to\infty} \frac{m_n}{n} > 0$. Then, for each of the eight matrices, the EESD of S_{A^b} and S_{A^B} converge to some probability measures μ_α^b and μ_α^B that depend on $\{c_{2k}\}_{k\geq 1}$.*

The proof is very similar to those of Corollaries 4.4.14 and 4.4.17, so we skip it.

Remark 6.4.11. *In Corollary 6.4.10 if $\{x_i\}$ are all fully i.i.d.with mean zero and finite variance, and the entries of the matrices are $\{\frac{y_i}{\sqrt{m_n}}\}$, then additionally (6.4.6) and thereby (6.4.7) holds. Thus the almost sure convergence of the ESDs can be concluded.*

6.5 Simulations

The LSDs of course cannot be universal. A variety of limit distributions are possible and are influenced by the moments of the entries and y(and nothing else). Moreover, even though ESD of S converges almost surely to μ, as noted in Remark 6.1.2, ESD of S_A does not converge almost surely to μ_A in general. Figure 6.1 and 6.2 illustrates this point. Matrices with variance profile serve as natural examples in demonstrating the diversity of the limit distributions. In Figure 6.3 we give some simulated ESD of S_A when A has a variance profile.

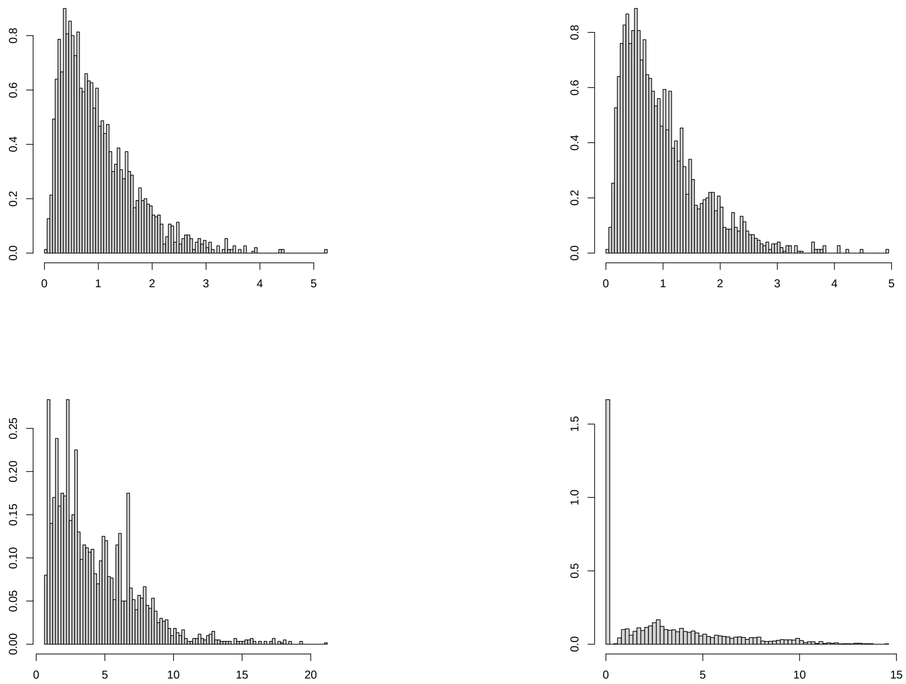


FIGURE 6.1: Histogram of the eigenvalues of $S_{R^{(s)}}$ with entries i.i.d. $N(0,1)/\sqrt{n}$ (top row) and i.i.d. Ber$(3/n)$ (bottom row), $p = 1000, n = 2000$, 2 replications.

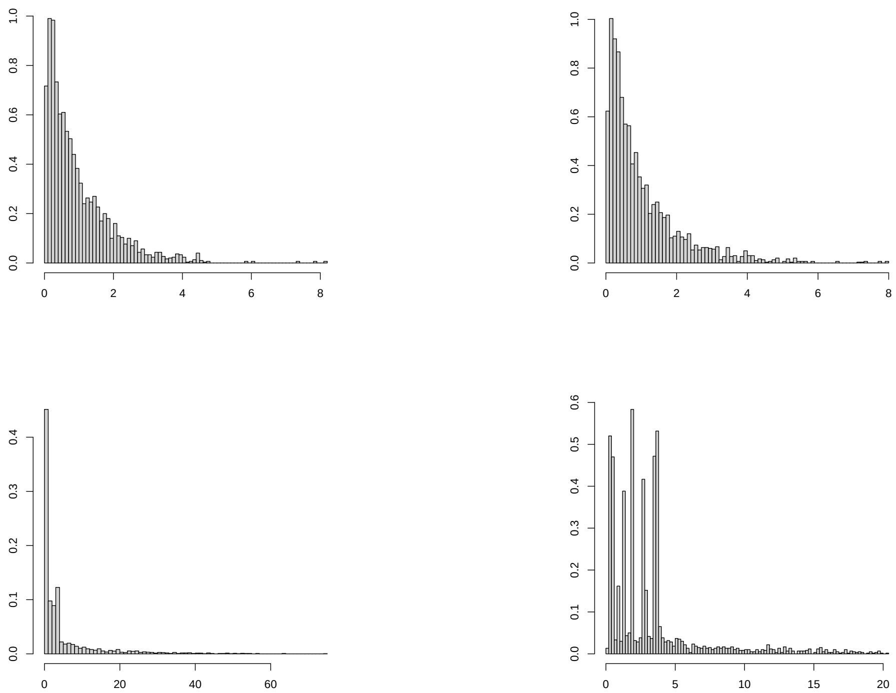
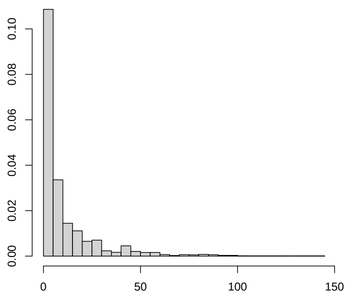
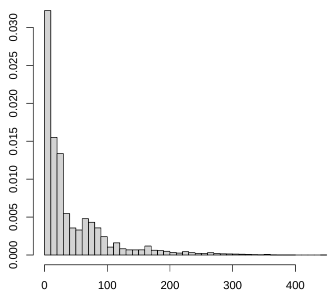


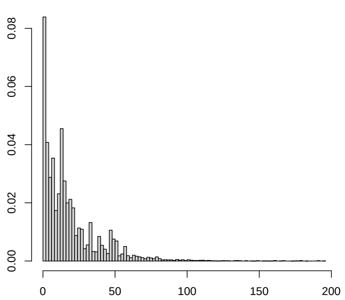
FIGURE 6.2: Histogram of the eigenvalues of $S_{C^{(s)}}$ with entries i.i.d. $N(0,1)/\sqrt{n}$(top row) and i.i.d. Ber$(3/n)$ (bottom row), $p = 1000, n = 2000$, 2 replications.



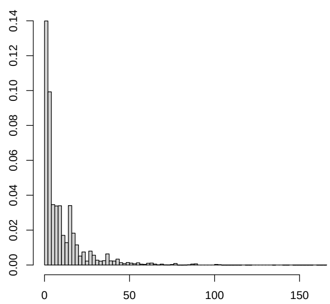
(A) $S_{T(s)}$



(B) $S_{H(s)}$



(C) $S_{R(s)}$



(D) $S_{C(s)}$

FIGURE 6.3: Histogram of the eigenvalues of S_A for $p = 1000, n = 2000$, 30 replications with variance profile $\sigma(x) = x^2 + 4x$ and $x_{ij,n} \sim Ber(3/n)$.

Bibliography

O. H. Ajanki, L. Erdős, and T. Krüger. Universality for general Wigner-type matrices. *Probab. Theory Related Fields*, 169(3-4):667–727, 2017. URL https://doi.org/10.1007/s00440-016-0740-2.

G. Akemann, J. Baik, and P. Di Francesco, editors. *The Oxford Handbook of Random Matrix Theory*. Oxford University Press, Oxford, 2011.

G. W. Anderson and O. Zeitouni. A CLT for a band matrix model. *Probab. Theory Related Fields*, 134(2):283–338, 2006. URL https://doi.org/10.1007/s00440-004-0422-3.

G. W. Anderson, A. Guionnet, and O. Zeitouni. *An Iintroduction to Random Matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

T. W. Anderson. Asymptotic theory for principal component analysis. *Ann. Math. Statist.*, 34:122–148, 1963. URL https://doi.org/10.1214/aoms/1177704248.

L. Arnold. On the asymptotic distribution of the eigenvalues of random matrices. *J. Math. Anal. Appl.*, 20:262–268, 1967. URL https://doi.org/10.1016/0022-247X(67)90089-3.

L. Arnold. On Wigner's semicircle law for the eigenvalues of random matrices. *Z. Wahrsch. Verw. Gebiete*, 19:191–198, 1971. URL https://doi.org/10.1007/BF00534107.

Z. Bai and J. W. Silverstein. *Spectral Analysis of Large Dimensional Random Matrices*. Springer Series in Statistics. Springer, New York, second edition, 2010. URL https://doi.org/10.1007/978-1-4419-0661-8.

Z. D. Bai. Circular law. *Ann. Probab.*, 25(1):494–529, 1997. URL https://doi.org/10.1214/aop/1024404298.

Z. D. Bai. Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica*, 9(3):611–677, 1999. With comments by G. J. Rodgers and Jack W. Silverstein; and a rejoinder by the author.

D. Banerjee and A. Bose. Patterned sparse random matrices: a moment approach. *Random Matrices Theory Appl.*, 6(3):1750011 (40 pages), 2017. URL https://doi.org/10.1142/S2010326317500113.

T. Banica, S. Curran, and R. Speicher. De Finetti theorems for easy quantum groups. *Ann. Probab.*, 40(1):401–435, 2012. URL https://doi.org/10.1214/10-AOP619.

O. Barrett, F. W. K. Firk, S. J. Miller, and C. Turnage-Butterbaugh. From Quantum Systems to L-Functions: Pair Correlation Statistics and Beyond. In *Open Problems in Mathematics*, pages 123–171. Springer, Cham, 2016.

A. Basak and A. Bose. Limiting spectral distributions of some band matrices. *Period. Math. Hungar.*, 63(1):113–150, 2011. URL https://doi.org/10.1007/s10998-011-7113-5.

R. Basu, A. Bose, S. Ganguly, and R. Subhra Hazra. Spectral properties of random triangular matrices. *Random Matrices Theory Appl.*, 1(3):1250003 (22 pages), 2012. URL https://doi.org/10.1142/S2010326312500037.

M. Bauer and O. Golinelli. Random incidence matrices: moments of the spectral density. *J. Statist. Phys.*, 103(1-2):301–337, 2001. URL https://doi.org/10.1023/A:1004879905284.

S. Belinschi, A. Dembo, and A. Guionnet. Spectral measure of heavy tailed band and covariance random matrices. *Comm. Math. Phys.*, 289(3):1023–1055, 2009. URL https://doi.org/10.1007/s00220-009-0822-4.

G. Ben Arous and A. Guionnet. The spectrum of heavy tailed random matrices. *Comm. Math. Phys.*, 278(3):715–751, 2008. URL https://doi.org/10.1007/s00220-007-0389-x.

F. Benaych-Georges and T. Cabanal-Duvillard. Marchenko-Pastur Theorem and Bercovici-Pata bijections for heavy-tailed or localized vectors. *ALEA Lat. Am. J.*

Probab. Math. Stat., 9(2):685–715, 2012. URL https://hal.archives-ouvertes.fr/hal-00690425.

F. Benaych-Georges, A. Guionnet, and C. Male. Central limit theorems for linear statistics of heavy tailed random matrices. *Comm. Math. Phys.*, 329(2):641–686, 2014. URL https://doi.org/10.1007/s00220-014-1975-3.

F. Benaych-Georges, C. Bordenave, and A. Knowles. Largest eigenvalues of sparse inhomogeneous Erdős-Rényi graphs. *Ann. Probab.*, 47(3):1653–1676, 2019. URL https://doi.org/10.1214/18-AOP1293.

C. Berge. *Hypergraphs-Combinatorics of Finite Sets*, volume 45 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1989.

P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, Inc., New York-London-Sydney, 1968.

K. Blackwell, N. Borade, A. Bose, C. D. VI, N. Luntzlara, R. Ma, S. J. Miller, S. S. Mukherjee, M. Wang, and W. Xu. Distribution of eigenvalues of matrix ensembles arising from wigner and palindromic toeplitz blocks. *arXiv preprint arXiv:2102.05839*, 2021.

M. Bolla. Distribution of the eigenvalues of random block-matrices. *Linear Algebra Appl.*, 377:219–240, 2004. URL https://doi.org/10.1016/j.laa.2003.08.015.

B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures Algorithms*, 31(1):3–122, 2007. URL https://doi.org/10.1002/rsa.20168.

C. Bordenave and M. Lelarge. The rank of diluted random graphs. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1389–1402. SIAM, Philadelphia, PA, 2010a.

C. Bordenave and M. Lelarge. Resolvent of large random graphs. *Random Structures Algorithms*, 37(3):332–352, 2010b. URL https://doi.org/10.1002/rsa.20313.

C. Bordenave, M. Lelarge, and J. Salez. The rank of diluted random graphs. *Ann. Probab.*, 39(3):1097–1121, 2011. URL https://doi.org/10.1214/10-AOP567.

C. Bordenave, A. Sen, and B. Virág. Mean quantum percolation. *J. Eur. Math. Soc.*, 19(12):3679–3707, 2017. URL https://doi.org/10.4171/JEMS/750.

A. Bose. *Patterned Random Matrices*. Chapman and Hall, CRC Press, Boca Raton, FL, 2018. URL https://doi.org/10.1201/9780429488436.

A. Bose. *Random Matrices and Non-Commutative Probability*. Chapman and Hall, CRC Press, Boca Raton, FL, 2021. URL https://doi.org/10.1201/9781003144496.

A. Bose and J. Mitra. Limiting spectral distribution of a special circulant. *Statist. Probab. Lett.*, 60(1):111–120, 2002. URL https://doi.org/10.1016/S0167-7152(02)00289-4.

A. Bose and A. Sen. Another look at the moment method for large dimensional random matrices. *Electron. J. Probab.*, 13(21):588–628, 2008. URL https://doi.org/10.1214/EJP.v13-501.

A. Bose and P. Sen. XX^T matrices with independent entries. *ALEA Lat. Am. J. Probab. Math. Stat.*, 20(1):75–125, 2023. URL https://doi.org/10.30757/alea.v20-05.

A. Bose, A. Dasgupta, and H. Rubin. A contemporary review and bibliography of infinitely divisible distributions and processes. *Sankhyā Ser. A*, 64(3, part 2):763–819, 2002. Special issue in memory of D. Basu.

A. Bose, S. Gangopadhyay, and A. Sen. Limiting spectral distribution of XX' matrices. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(3):677–707, 2010. URL https://doi.org/10.1214/09-AIHP329.

A. Bose, S. Guha, R. S. Hazra, and K. Saha. Circulant type matrices with heavy tailed entries. *Statist. Probab. Lett.*, 81(11):1706–1716, 2011a. URL https://doi.org/10.1016/j.spl.2011.07.001.

A. Bose, R. Subhra Hazra, and K. Saha. Half independence and half cumulants. *Electron. Commun. Probab.*, 16:405–422, 2011b. URL https://doi.org/10.1214/ECP.v16-1651.

A. Bose, K. Saha, and P. Sen. Some patterned matrices with independent entries. *Random Matrices Theory Appl.*, 10(3):2150030 (46 pages), 2021. URL https://doi.org/10.1142/S2010326321500301.

W. Bryc, A. Dembo, and T. Jiang. Spectral measure of large random Hankel, Markov and Toeplitz matrices. *Ann. Probab.*, 34(1):1–38, 2006. URL https://doi.org/10.1214/009117905000000495.

T. Carleman. *Les Fonctions quasi analytiques*. Gauthier-Villars, Paris, 1926.

G. Casati and V. Girko. Generalized Wigner law for band random matrices. *Random Oper. Stochastic Equations*, 1(3):279–286, 1993a. URL https://doi.org/10.1515/rose.1993.1.3.279.

G. Casati and V. Girko. Wigner's semicircle law for band random matrices. *Random Oper. Stochastic Equations*, 1(1):15–21, 1993b. URL https://doi.org/10.1515/rose.1993.1.1.15.

G. Casati, B. V. Chirikov, F. M. Izraelev, and J. Ford. Stochastic behavior of a quantum pendulum under a periodic perturbation. In *Stochastic behavior in classical and quantum Hamiltonian systems (Volta Memorial Conf., Como, 1977)*, volume 93 of *Lecture Notes in Phys.*, pages 334–352. Springer, Berlin-New York, 1979.

A. Chakrabarty, R. S. Hazra, F. d. Hollander, and M. Sfragara. Large Deviation Principle for the Maximal Eigenvalue of Inhomogeneous Erdős-Rényi Random Graphs. *J. Theoret. Probab.*, 35(4):2413–2441, 2022. URL https://doi.org/10.1007/s10959-021-01138-w.

F. Chen, Y. Lin, S. J. Miller, and J. Yu. The limiting spectral measure for an ensemble of generalized checkerboard matrices. *PUMP J. Undergrad. Res.*, 4:202–221, 2021. URL https://doi.org/10.1007/s42543-021-00034-w.

B. V. Chirikov. An example of chaotic eigenstates in a complex atom. *Physics Letters A*, 108(2):68–70, 1985. URL https://www.sciencedirect.com/science/article/pii/0375960185905183.

J. B. Conrey. *L-Functions and Random Matrices*. Springer, Berlin, 2001.

X. Ding. Spectral analysis of large block random matrices with rectangular blocks. *Lith. Math. J.*, 54(2):115–126, 2014. URL https://doi.org/10.1007/s10986-014-9231-2.

X. Ding and T. Jiang. Spectral distributions of adjacency and Laplacian matrices of random graphs. *Ann. Appl. Probab.*, 20(6):2086–2117, 2010. URL https://doi.org/10.1214/10-AAP677.

K. Dykema and U. Haagerup. DT-operators and decomposability of Voiculescu's circular operator. *Amer. J. Math.*, 126(1):121–189, 2004. URL http://muse.jhu.edu/journals/american_journal_of_mathematics/v126/126.1dykema.pdf.

A. Edelman. The probability that a random real gaussian matrix haskreal eigenvalues, related distributions, and the circular law. *J. Multivariate Anal.*, 60(2):203–232, 1997. URL https://www.sciencedirect.com/science/article/pii/S0047259X9691653X.

L. Erdős, S. Péché, J. A. Ramírez, B. Schlein, and H.-T. Yau. Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.*, 63(7):895–925, 2010. URL https://doi.org/10.1002/cpa.20317.

L. Erdős, H.-T. Yau, and J. Yin. Bulk universality for generalized Wigner matrices. *Probab. Theory Related Fields*, 154(1-2):341–407, 2012. URL https://doi.org/10.1007/s00440-011-0390-3.

P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

F. W. K. Firk and S. J. Miller. Nuclei, primes and the random matrix connection. *Symmetry*, 1(1):64–105, 2009. URL https://doi.org/10.3390/sym1010064.

P. J. Forrester. *Log-Gases and Random Matrices*, volume 34 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2010. URL https://doi.org/10.1515/9781400835416.

J. S. Geronimo and T. P. Hill. Necessary and sufficient condition that the limit of Stieltjes transforms is a Stieltjes transform. *J. Approx. Theory*, 121(1):54–60, 2003. URL https://doi.org/10.1016/S0021-9045(02)00042-4.

E. N. Gilbert. Random graphs. *Ann. Math. Statist.*, 30:1141–1144, 1959. URL https://doi.org/10.1214/aoms/1177706098.

J. Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *J. Mathematical Phys.*, 6:440–449, 1965. URL https://doi.org/10.1063/1.1704292.

V. Girko, W. Kirsch, and A. Kutzelnigg. A necessary and sufficient condition for the semicircle law. *Random Oper. Stochastic Equations*, 2(2):195–202, 1994. URL https://doi.org/10.1515/rose.1994.2.2.195.

V. L. Girko. The circular law. *Teor. Veroyatnost. i Primenen.*, 29(4):669–679, 1984.

V. L. Girko. The strong circular law. Twenty years later. II. *Random Oper. Stochastic Equations*, 12(3):255–312, 2004. URL https://doi.org/10.1163/1569397042222477.

R. M. Gray. Toeplitz and Circulant Matrices: A review. *Foundations and Trends® in Communications and Information Theory*, 2(3):155–239, 2006.

U. Grenander. *Probabilities on Algebraic Structures*. Almqvist & Wiksell, Stockholm; John Wiley & Sons Inc., New York-London, second edition, 1968.

U. Grenander and J. W. Silverstein. Spectral analysis of networks with random topologies. *SIAM J. Appl. Math.*, 32(2):499–519, 1977. URL https://doi.org/10.1137/0132041.

U. Grenander and G. Szegő. *Toeplitz Forms and their Applications*. Chelsea Publishing Co., New York, second edition, 1984.

A. Guionnet. Bernoulli random matrices. *arXiv preprint arXiv:2112.05506*, 2021.

W. Hachem, P. Loubaton, and J. Najim. The empirical distribution of the eigenvalues of a Gram matrix with a given variance profile. *Ann. Inst. H. Poincaré Probab. Statist.*, 42(6):649–670, 2006. URL https://doi.org/10.1016/j.anihpb.2005.10.001.

A. Hald. The early history of the cumulants and the Gram-Charlier series. *International Statistical Review*, 68(2):137–153, 2000. URL http://www.jstor.org/stable/1403665.

C. Hammond and S. J. Miller. Distribution of eigenvalues for the ensemble of real symmetric Toeplitz matrices. *J. Theoret. Probab.*, 18(3):537–566, 2005. URL https://doi.org/10.1007/s10959-005-3518-5.

A. J. Hoffman and H. W. Wielandt. The variation of the spectrum of a normal matrix. *Duke Math. J.*, 20:37–39, 1953. URL http://projecteuclid.org/euclid.dmj/1077465062.

S. Jackson, S. J. Miller, and T. Pham. Distribution of eigenvalues of highly palindromic Toeplitz matrices. *J. Theoret. Probab.*, 25(2):464–495, 2012. URL https://doi.org/10.1007/s10959-010-0311-x.

S. Jin and J. Xie. A result on the limiting spectral distribution of random matrices with unequal variance entries. *J. Inequal. Appl.*, 174:(13 pages), 2020. URL https://doi.org/10.1186/s13660-020-02440-7.

I. M. Johnstone and D. M. Titterington. Statistical challenges of high-dimensional data. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 367(1906):4237–4253, 2009. URL https://doi.org/10.1098/rsta.2009.0159.

D. Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.*, 12(1):1–38, 1982. URL https://doi.org/10.1016/0047-259X(82)90080-X.

P. Jung. Lévy-Khintchine random matrices and the Poisson weighted infinite skeleton tree. *Trans. Amer. Math. Soc.*, 370(1):641–668, 2018. URL https://doi.org/10.1090/tran/6977.

V. Kargin. Spectrum of random Toeplitz matrices with band structure. *Electron. Commun. Probab.*, 14:412–421, 2009. URL https://doi.org/10.1214/ECP.v14-1492.

N. M. Katz and P. Sarnak. Zeroes of zeta functions and symmetry. *Bull. Amer. Math. Soc. (N.S.)*, 36(1):1–26, 1999. URL https://doi.org/10.1090/S0273-0979-99-00766-1.

J. P. Keating and N. C. Snaith. Random matrices and L-functions. *J. Phys. A*, 36(12):2859–2881, 2003. URL https://doi.org/10.1088/0305-4470/36/12/301.

O. Khorunzhy, M. Shcherbina, and V. Vengerovsky. Eigenvalue distribution of large weighted random graphs. *J. Math. Phys.*, 45(4):1648–1672, 2004. URL https://doi.org/10.1063/1.1667610.

M. Koloğlu, G. S. Kopp, and S. J. Miller. The limiting spectral measure for ensembles of symmetric block circulant matrices. *J. Theoret. Probab.*, 26(4):1020–1060, 2013. URL https://doi.org/10.1007/s10959-011-0391-2.

S. Liang, N. Obata, and S. Takahashi. Asymptotic spectral analysis of generalized Erdős-Rényi random graphs. In *Noncommutative Harmonic Analysis with Applications to*

Probability, volume 78 of *Banach Center Publ.*, pages 211–229. Polish Acad. Sci. Inst. Math., Warsaw, 2007. URL https://doi.org/10.4064/bc78-0-16.

D. Z. Liu and Z. D. Wang. Limit distribution of eigenvalues for random Hankel and Toeplitz band matrices. *J. Theoret. Probab.*, 24(4):988–1001, 2011. URL https://doi.org/10.1007/s10959-009-0260-4.

J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math.*, 169(3):903–991, 2009. URL https://doi.org/10.4007/annals.2009.169.903.

L. Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012. URL https://doi.org/10.1090/coll/060.

A. Lytova and L. Pastur. Central limit theorem for linear eigenvalue statistics of random matrices with independent entries. *Ann. Probab.*, 37(5):1778–1840, 2009. URL https://doi.org/10.1214/09-AOP452.

C. Male. The limiting distributions of large heavy Wigner and arbitrary random matrices. *J. Funct. Anal.*, 272(1):1–46, 2017. URL https://doi.org/10.1016/j.jfa.2016.10.001.

V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457, 1967. URL https://dx.doi.org/10.1070/SM1967v001n04ABEH001994.

A. Massey, S. J. Miller, and J. Sinsheimer. Distribution of eigenvalues of real symmetric palindromic Toeplitz matrices and circulant matrices. *J. Theoret. Probab.*, 20(3):637–662, 2007. URL https://doi.org/10.1007/s10959-007-0078-x.

M. L. Mehta. *Random Matrices*, volume 142 of *Pure and Applied Mathematics*. Elsevier/Academic Press, Amsterdam, third edition, 2004.

M. L. Mehta and M. Gaudin. On the density of eigenvalues of a random matrix. *Nuclear Phys.*, 18:420–427, 1960.

J. A. Mingo and R. Speicher. *Free Probability and Random Matrices*, volume 35 of *Fields Institute Monographs*. Springer, New York, 2017. URL https://doi.org/10.1007/978-1-4939-6942-5.

S. A. Molchanov, L. A. Pastur, and A. Khorunzhii. Limiting eigenvalue distribution for band random matrices. *Theoretical and Mathematical Physics*, 90(2):108–118, 1992.

A. Nica and R. Speicher. *Lectures on the Combinatorics of Free Probability*, volume 335 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2006. URL https://doi.org/10.1017/CBO9780511735127.

N. Noiry. Spectral asymptotic expansion of Wishart matrices with exploding moments. *ALEA Lat. Am. J. Probab. Math. Stat.*, 15(2):897–911, 2018. URL https://doi.org/10.30757/alea.v15-34.

J. Novak. Three lectures on free probability. *Random matrix theory, interacting particle systems, and integrable systems*, 65(309-383):13, 2014.

G. Pan and W. Zhou. Circular law, extreme singular values and potential theory. *J. Multivariate Anal.*, 101(3):645–656, 2010. ISSN 0047-259X. doi: 10.1016/j.jmva.2009.08.005. URL https://doi.org/10.1016/j.jmva.2009.08.005.

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

L. A. Pastur. Spectra of random self adjoint operators. *Russian Mathematical Surveys*, 28(1):1, 1973.

I. Popescu. General tridiagonal random matrix models, limiting distributions and fluctuations. *Probab. Theory Related Fields*, 144(1-2):179–220, 2009. URL https://doi.org/10.1007/s00440-008-0145-y.

R. Rashidi Far, T. Oraby, W. Bryc, and R. Speicher. On slow-fading MIMO systems with nonseparable correlation. *IEEE Trans. Inform. Theory*, 54(2):544–553, 2008. URL https://doi.org/10.1109/TIT.2007.913239.

M. Riesz. Sur le probleme des moments, troisieme note. *Ark. Mat. Fys*, 16:1–52, 1923.

D. Shlyakhtenko. Random Gaussian band matrices and freeness with amalgamation. *Internat. Math. Res. Notices*, (20):1013–1025, 1996. URL https://doi.org/10.1155/S1073792896000633.

R. P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.

K.-T. Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006. URL https://doi.org/10.1007/s11511-006-0002-8.

T. Tao. *Topics in Random Matrix Theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012. URL https://doi.org/10.1090/gsm/132.

T. Tao and V. Vu. Random matrices: the circular law. *Commun. Contemp. Math.*, 10 (2):261–307, 2008. URL https://doi.org/10.1142/S0219199708002788.

T. Tao and V. Vu. Random matrices: universality of ESDs and the circular law. *Ann. Probab.*, 38(5):2023–2065, 2010. ISSN 0091-1798. doi: 10.1214/10-AOP534. URL https://doi.org/10.1214/10-AOP534. With an appendix by Manjunath Krishnapur.

T. Tao and V. Vu. Random matrices: universality of local eigenvalue statistics. *Acta Math.*, 206(1):127–204, 2011. URL https://doi.org/10.1007/s11511-011-0061-3.

C. Villani. Optimal transport, old and new. Notes for the 2005 Saint-Flour summer school. *Fundamental Principles of Mathematical Sciences, Springer*, 3, 2008.

D. Voiculescu. Limit laws for random matrices and free products. *Invent. Math.*, 104 (1):201–220, 1991. URL https://doi.org/10.1007/BF01245072.

D. V. Voiculescu, K. J. Dykema, and A. Nica. *Free Random Variables*, volume 1 of *CRM Monograph Series*. American Mathematical Society, Providence, RI, 1992. URL https://doi.org/10.1090/crmm/001.

K. W. Wachter. The strong limits of random matrix spectra for sample matrices of independent elements. *Ann. Probability*, 6(1):1–18, 1978. URL https://doi.org/10.1214/aop/1176995607.

E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math.*, 62:548–564, 1955. URL https://doi.org/10.2307/1970079.

E. P. Wigner. On the distribution of the roots of certain symmetric matrices. *Ann. of Math.*, 67:325–327, 1958. URL https://doi.org/10.2307/1970008.

Y. Q. Yin. Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.*, 20(1):50–68, 1986. URL https://doi.org/10.1016/0047-259X(86)90019-9.

I. Zakharevich. A generalization of Wigner's law. *Comm. Math. Phys.*, 268(2):403–414, 2006. URL https://doi.org/10.1007/s00220-006-0074-5.

Y. Zhu. A graphon approach to limiting spectral distributions of Wigner-type matrices. *Random Structures Algorithms*, 56(1):251–279, 2020. URL https://doi.org/10.1002/rsa.20894.

G. L. Zitelli. Amalgamated Free Lévy Processes as Limits of Sample Covariance Matrices. *J. Theoret. Probab.*, 35(4):2176–2193, 2022. URL https://doi.org/10.1007/s10959-021-01153-x.

List of Publications

1. Arup Bose and Priyanka Sen:
   XX^T matrices with independent entries, 2022,
   ALEA Lat. Am. J. Probab. Math. Stat.,
   https://doi.org/10.30757/alea.v20-05, (51 pages).

2. Arup Bose, Koushik Saha, Arusharka Sen and Priyanka Sen:
   Random matrices with independent entries: beyond non-crossing partitions.
   Random Matrices Theory Appl., 11(2), 2022, 2250021 (42 pages),
   https://doi.org/10.1142/S2010326322500216.

3. Arup Bose, Koushik Saha and Priyanka Sen:
   Some patterned matrices with independent entries.
   Random Matrices Theory Appl., 10(3), 2021, 2150030 (46 pages),
   https://doi.org/10.1142/S2010326321500301.