Joint convergence of sample cross-covariance matrices

Monika Bhattacharjee *
Dept. of Math.
Indian Institute of Technolgy
monaiidexp.gamma@gmail.com

Arup Bose †
Stat.-Math Unit
Indian Statistical Institute
bosearu@gmail.com

Apratim Dey
Dept. of Stat.
Stanford University
apd1995@stanford.edu

March 23, 2021

Abstract

Suppose X and Y are $p \times n$ matrices each with mean 0, variance 1 and where all moments of any order are uniformly bounded as $p, n \to \infty$. Moreover, the entries (X_{ij}, Y_{ij}) are independent across i, j with a common correlation ρ. Let $C = n^{-1}XY^*$ be the sample cross-covariance matrix. We show that if $n, p \to \infty, p/n \to y \neq 0$, then C converges in the algebraic sense and the limit moments depend only on ρ. Independent copies of such matrices with same p but different n, say $\{n_l\}$, different correlations $\{\rho_l\}$, and different non-zero y's, say $\{y_l\}$ also converge jointly and are asymptotically free. When $y = 0$, the matrix $\sqrt{np^{-1}}(C - \rho I_p)$ converges to an elliptic variable with parameter ρ^2. In particular, this elliptic variable is circular when $\rho = 0$ and is semi-circular when $\rho = 1$. If we take independent C_l, then the matrices $\{\sqrt{n_l p^{-1}}(C_l - \rho_l I_p)\}$ converge jointly and are also asymptotically free. As a consequence, the limiting spectral distribution of any symmetric matrix polynomial exists and has compact support.

Key words. Sample cross-covariance matrices, joint convergence, free independence, free cumulants, semi-circular variable, circular variable, elliptic variable, cross-covariance variable, Marčenko-Pastur law, compound free Poisson law.

AMS 2000 Subject classification: Primary 60B20, Secondary 46L54.

1 Introduction

The large sample behaviour of the high dimensional sample covariance matrix $S = n^{-1}XX^*$ where X is a $p \times n$ matrix with i.i.d. entries has been extensively studied. Under suitable moment assumptions on the entries, the convergence of its spectral distribution when $n, p \to \infty$ and $p/n \to y \neq 0$ was originally shown in Marčenko and Pastur [1967] and the limit law is now known as the Marčenko-Pastur law. When $y = 0$, the limit spectral distribution (LSD) of $\sqrt{np^{-1}}(S - I_p)$ where I_p is the $p \times p$ identity matrix, is known to be the (standard) semi-circle law. See Bai and Silverstein [2010] and Bose [2018] for book-level expositions of these results. The joint algebraic convergence and asymptotic freeness of independent S-matrices was established in Capitaine and Casalis [2004]. The joint convergence of the generalized covariance

*Research supported by seed grant from IIT Bombay and Start-up Research Grant from SERB, Dept. of Science and Technology, Govt. of India.

†Research supported by J.C. Bose National Fellowship, Dept. of Science and Technology, Govt. of India.

matrices and the convergence of the spectral distribution of their symmetric matrix polynomials was dealt in Bhattacharjee and Bose [2016a] and Bhattacharjee and Bose [2017].

The semi-circle law is a central probability law in free probability and orginally arose from the study of a Wigner matrix W_n which is a real symmetric matrix whose entries are i.i.d. with mean 0 and variance 1. The limit spectral distribution of $n^{-1/2}W_n$ is the (standard) semi-circle law. Moreover, independent copies of these matrices, when all moments are finite, converge jointly in the algebraic sense and are asymptotically free.

If in W_n the (i,j)-th and the (j,i)-th entries have a common correlation ρ, then it is called an elliptic matrix, and in that case the LSD of $n^{-1/2}W_n$ is the uniform law in the interior of an ellipse centered at the origin with lengths of the major and the minor axes being $\sqrt{2(1+\rho)}$ and $\sqrt{2(1-\rho)}$. See Nguyen and O$'$Rourke [2015]. In particular if $\rho = 1$ we recover the semi-circle law result for the Wigner matrix, and if $\rho = 0$ then the LSD is uniformly distributed over the unit disc. It is also known that independent copies of elliptic matrices with possibly different ρ converge jointly to elliptic elements and are asymptotically free. See Adhikari and Bose [2019].

Motivated by the above results, we consider the following high dimensional model: X and Y are two $p \times n$ random matrices where the entries (x_{ij}, y_{ij}), $1 \le i \le p$, $1 \le j \le n$ are independent bivariate random variables with means 0, variances 1 and correlation ρ. Then the matrix $C = n^{-1}XY^*$ is called the sample *cross-covariance matrix*. Note that if $\rho = 1$, then $X = Y$ and we recover the S matrix.

We study the joint convergence of independent copies $\{C_l\}$ of these matrices with possibly different $\{\rho_l\}$ and $\{n_l\}$. This convergence is taken to be convergence as elements of an appropriate $*$-probability space as described below.

Consider the $*$-probability space $(\mathcal{M}_p, \varphi_p)$ where $\mathcal{M}_p$ is the set of all $p \times p$ random matrices:

$$\mathcal{M}_p(\mathbb{C}) = \{A : A = ((a_{ij}))_{p\times p} \text{ and } \quad \mathbb{E}|a_{ij}|^k < \infty \quad \text{for all} \quad i, j, k\}, \tag{1.1}$$

and the state φ_p is defined as

$$\varphi_p(A) = p^{-1}\mathbb{E}[\text{Trace}(A)].$$

Note that φ_p is positive and tracial (that is, $\varphi_p(aa^*) \ge 0$, and $\varphi_p(ab) = \varphi_p(ba)$ for all a, b).

Elements $\{A_l\}$, $1 \le i \le t$ from $\mathcal{M}_p(\mathbb{C})$ are said to converge jointly if for every polynomial $\Pi(A_l, 1 \le l \le t)$ in the variables $\{A_l, A_l^*\}$, $\varphi_p(\Pi(A_l, 1 \le l \le t)) = p^{-1}\mathbb{E}[\text{Trace}(\Pi(A_l, 1 \le l \le t))]$ converges as $p \to \infty$. The limit algebra is taken to be a $*$-algebra generated by indeterminates $\{a_l\}$ and its state is defined through the above limits as

$$\varphi(\Pi(a_l, 1 \le l \le t)) = \lim_{p\to\infty} p^{-1}\mathbb{E}[\text{Trace}(\Pi(A_l, 1 \le l \le t))]. \tag{1.2}$$

We shall refer to this convergence as algebraic convergence. Note that φ is also tracial and positive. The matrices $\{A_l\}$ are said to be free if the limit variables $\{a_l\}$ are free.

Two different cases arise: for each l, either $n_l, p \to \infty, p/n_l \to y_l, 0 < y_l < \infty$ or $p \to \infty, p/n_l \to 0$. Suppose the entries of $\{X_l, Y_l\}$ have mean 0, variance 1, and moments of any order are uniformly bounded. Moreover, across l the matrices are independent. Then in the first case $\{C_l\}$ converge jointly in the sense of (1.2) to say $\{c_l\}$ which are asymptotically free. The moments of each c_l depend only on y_l and ρ_l and we give a formula for their free cumulants. In the second case $\{\sqrt{n_l p^{-1}}(C_l - \rho_l I_p)\}$, $1 \le l \le t$ converge to free elliptic variables with parameters ρ_l^2. The algebraic convergence results on independent S-matrices mentioned earlier are obtained as a special case by using $\rho = 1$.

There is another natural state on $\mathcal{M}_p(\mathbb{C})$ given by

$$\tilde{\varphi}_p(A) = p^{-1}\text{Trace}(A).$$

Convergence with respect to this state is defined as the almost sure convergence of $p^{-1}\text{Trace}(\Pi$ $(A_l, 1 \leq l \leq t))$ for all polynomials Π, and for our purposes the limits are non-random. In the present paper, it turns out that all the convergence results described above that hold for $\{C_l\}$ with respect to φ_p also hold with respect to $\tilde{\varphi}_p$.

A related notion of convergence for a single sequence of random matrices is the convergence of the spectral distribution. Suppose A_p is any $p \times p$ matrix with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_p$. The random probability law which puts mass p^{-1} on each eigenvalue is called the empirical spectral distribution (ESD) of A_p. If we take a further expectation with respect to the underlying law of the random variables, then it defines another probability law, which we shall call the expected empirical spectral distribution (EESD). If the ESD converges weakly (in probability or almost surely), the limit is called the limiting spectral distribution (LSD). If this limit is non-random, then it is also the limit of the EESD. Usually the convergence of the EESD is easier to establish and then the convergence of the ESD to the same limit often follows by a Borel-Cantelli type argument on the moments of the ESD.

Akemann et al. [2008] and Vinayak and Benet [2014] respectively analysed the characteristic polynomial and spectral domain, and Akemann et al. [2020] established the weak convergence of the ESD of a cross-covariance matrix with complex Gaussian entries when $0 < y_l < \infty$. Now consider any symmetric matrix polynomial Π in the matrices $\{C_l, C_l^*\}$. Then the moments of the EESD of Π are $p^{-1}\mathbb{E}[\text{Trace}(\Pi^k)]$. Algebraic convergence with respect to φ_p implies that $p^{-1}\mathbb{E}[\text{Trace}(\Pi^k)]$ converges for all integers $k \geq 1$. It is easily checked that these limiting moments define a unique probability law and hence the EESD of Π converges weakly to this law. This convergence can be upgraded to almost sure convergence of the ESD of Π by an estimation of the fourth moment of $p^{-1}\text{Trace}(\Pi^k) - p^{-1}\mathbb{E}[\text{Trace}(\Pi^k)]$ and applying the Borel-Cantelli Lemma. As an example, the distribution of the singular values of C_l converges almost surely. Simulations suggest that the LSD exists even for non-symmetric polynomials. It is difficult to settle this rigorously for any non-symmetric polynomial, and we do not pursue this issue in this article.

2 Necessary notions from non-commutative probability

We briefly mention some notions of non-commutative probability that we shall need. For further details see Nica and Speicher [2006]. Let $(\mathcal{A}, \varphi)$ be a $*$-probability space. Suppose $\{a_i : i \in I\}$ is a collection of variables from $\mathcal{A}$. Then the numbers $\{\varphi(\Pi(a_i, a_i^* : i \in I)) :$ Π is a finite degree monomial$\}$ are called their *joint $*$-moments*. For simplicity, we shall often write moments instead of $*-$moments.

Suppose $a \in \mathcal{A}$ is a *self-adjoint* element, that is, $a^* = a$. Suppose there a *unique* probability law μ_a on $\mathbb{R}$ such that

$$\varphi(a^n) = \int_{\mathbb{R}} x^n \mu_a(dx), n = 1, 2, \ldots. \tag{2.1}$$

Then μ_a is called the *probability law of a*. If μ_a satisfies (2.1) and is compactly supported, then it is indeed unique.

Let $(\mathcal{A}, \varphi)$ be a $*$-probability space. Define a sequence of *multi-linear functionals* $(\varphi_n)_{n\in\mathbb{N}}$ on $\mathcal{A}^n$ via

$$\varphi_n(a_1, a_2, \ldots, a_n) := \varphi(a_1 a_2 \cdots a_n). \tag{2.2}$$

Extend $\{\varphi_n, n \geq 1\}$ to $\{\varphi_\pi, \pi \in NC(n), n \geq 1\}$ *multiplicatively* in a recursive way by the following formula. If $\pi = \{V_1, V_2, \ldots, V_r\} \in NC(n)$, then

$$\varphi_\pi[a_1, a_2, \ldots, a_n] := \varphi(V_1)[a_1, a_2, \ldots, a_n] \cdots \varphi(V_r)[a_1, a_2, \ldots, a_n], \tag{2.3}$$

where $\varphi(V)[a_1, a_2, \ldots, a_n] := \varphi_s(a_{i_1}, a_{i_2}, \ldots, a_{i_s}) = \varphi(a_{i_1} a_{i_2} \cdots a_{i_s})$ for $V = \{i_1, i_2, \ldots, i_s\}$ with $i_1 < i_2 < \cdots < i_s$. Note that the order of the variables have been preserved. Also note that the two types of braces ( ) and [ ] in (2.2) and (2.3) have different uses. In particular, if $\mathbf{1}_n$ denotes the 1-block partition of $\{1, \ldots, n\}$ then

$$\varphi_{\mathbf{1}_n}[a_1, a_2, \ldots, a_n] = \varphi_n(a_1, a_2, \ldots, a_n) = \varphi(a_1 a_2 \cdots a_n). \tag{2.4}$$

The *joint free cumulant* of order n of $(a_1, a_2, \ldots, a_n)$ is

$$\kappa_n(a_1, a_2, \ldots, a_n) = \sum_{\sigma \in NC(n)} \varphi_\sigma[a_1, a_2, ..., a_n] \mu(\sigma, \mathbf{1}_n), \tag{2.5}$$

where μ is the Möbius function of $NC(n)$. It is called a *mixed free cumulant* if at least *one* pair a_i, a_j are different and $a_i \neq a_j^*$ for some $i \neq j$. For any $\epsilon_i \in \{1, *\}, 1 \leq i \leq n$, $\kappa_n(a^{\epsilon_1}, a^{\epsilon_2}, \ldots, a^{\epsilon_n})$ is called a *marginal free cumulant* of order n of $\{a, a^*\}$. For a self-adjoint element a,

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

is called the n-th free cumulant of a. The free cumulants κ_n in (2.5) are also multi-linear. In particular, for any variables $\{a_i, b_i\}$ and constants $\{c_i\}$,

$$\begin{aligned} \kappa_n(a_1 + b_1, \ldots, a_n + b_n) &= \kappa_n(a_1, \ldots, a_n) + \kappa_n(a_1, b_2, a_3, \ldots, a_n) + \cdots + \kappa_n(b_1, \ldots, b_n), \quad \text{and} \\ \kappa_n(c_1 a, c_2 a, \ldots, c_n a) &= c_1 c_2 \cdots c_n \kappa_n(a, a, \ldots, a) = c_1 c_2 \cdots c_n \kappa_n(a). \end{aligned}$$

Let $\{\kappa_\pi : \pi \in NC(n), n \geq 1\}$ be the multiplicative extension of $\{\kappa_n, n \geq 1\}$. By using (2.5), for any $\pi \in NC(n), \ \ n \geq 1$,

$$\kappa_\pi[a_1, a_2, ..., a_n] := \sum_{\sigma \in NC(n):\ \sigma \leq \pi} \varphi_\sigma[a_1, a_2, \ldots, a_n] \mu(\sigma, \pi). \tag{2.6}$$

Note that

$$\begin{aligned} \kappa_{\mathbf{1}_n}[a_1, a_2, \ldots, a_n] &= \kappa_n(a_1, a_2, \ldots, a_n), \\ \kappa_2(a_1, a_2) &= \varphi(a_1 a_2) - \varphi(a_1)\varphi(a_2). \end{aligned}$$

Using the Möbius function μ, it can be shown that

$$\begin{aligned} \varphi(a_1 a_2 \cdots a_n) &= \sum_{\sigma \in NC(n):\ \sigma \leq \mathbf{1}_n} \kappa_\sigma[a_1, a_2, \ldots, a_n], \\ \varphi_\pi[a_1, a_2, \ldots, a_n] &= \sum_{\sigma \in NC(n):\ \sigma \leq \pi} \kappa_\sigma[a_1, a_2, \ldots, a_n]. \end{aligned} \tag{2.7}$$

In particular, (2.5) and (2.7) establish a one-to-one correspondence between free cumulants and moments. These relations will be referred to as *moment-free cumulant relations.*

Variables $\{a_i : i \in I\}$ are said to be free if their mixed free cumulants are all zero. Similarly, variables $\{a_i^{(p)} :\ i \in I\}$ are said to be asymptotically free if, as $p \to \infty$, they converge jointly to $\{a_i :\ i \in I\}$ which are free.

3 Main results

Let X_l and Y_l be $p \times n_l$ random matrices for all $1 \le l \le t$. The (i,j)-th entries of X_l and Y_l are denoted by $X^{(l)}_{ij,(n_l,p)}$ and $Y^{(l)}_{ij,(n_l,p)}$ respectively. Note that all entries of these matrices may change with p and n_l. We shall often suppress this dependence by dropping the subscripts p and n_l. We make the following assumption on these entries.

Assumption I

(a) For every $p \ge 1$ and $n_l \ge 1$, the pairs of random variables $(X^{(l)}_{ij,(n_l,p)}, Y^{(l)}_{ij,(n_l,p)})$ are independent across $1 \le i \le p$, $1 \le j \le n_l$, $1 \le l \le t$.

(b) For all $1 \le i \le p$, $1 \le j \le n_l$, $1 \le l \le t$, $p, n_l \ge 1$,

$$\mathbb{E}(X^{(l)}_{ij}) = \mathbb{E}(Y^{(l)}_{ij}) = 0, \quad \mathbb{E}[(X^{(l)}_{ij})^2] = \mathbb{E}[(Y^{(l)}_{ij})^2] = 1, \quad \mathbb{E}(X^{(l)}_{ij} Y^{(l)}_{ij}) = \rho_l.$$

(c) $\sup\limits_{p,n_l \ge 1} \sup\limits_{1 \le i \le p} \sup\limits_{1 \le j \le n_l} \mathbb{E}\big(|X^{(l)}_{ij}|^k + |X^{(l)}_{ij}|^k\big) < B_k < \infty$ for all $1 \le l \le t$ and $k \ge 1$.

(d) $n_l = n_l(p) \to \infty$ as $p \to \infty$ such that $n_l^{-1} p \to y_l < \infty$.

Define the sample cross-covariance matrices $C_l = \frac{1}{n_l} X_l Y_l^*$ for all $1 \le l \le t$.

3.1 Convergence of $\{C_l : 1 \le l \le t\}$ when $y_l \ne 0$

The symbol δ will be used in two senses: δ_x will denote the probability measure which puts all mass at x; on the other hand, δ_{xy} is defined as

$$\delta_{xy} = \begin{cases} 1 & \text{if} \quad x = y \\ 0 & \text{if} \quad x \ne y. \end{cases}$$

A compound free Poisson variable with rate λ and jump distribution μ will be denoted by $\mathrm{P}(\lambda, \mu)$ or by $\mathrm{P}(\lambda, X)$ where X is a random variable with probability law μ.

The following variable shall appear in the limit:

Definition 3.1. An element c of a $*$-probability space will be called a *cross-covariance variable* with parameters ρ and $y, 0 < y < \infty$, if its free cumulants are given by

$$\kappa_k(c^{\eta_1}, c^{\eta_2}, \ldots, c^{\eta_k}) = \begin{cases} y^{k-1} \rho^{S(\boldsymbol{\eta}_k)} & \text{if} \quad \rho \ne 0 \\ y^{k-1} \delta_{S(\boldsymbol{\eta}_k)0} & \text{if} \quad \rho = 0. \end{cases}$$

where

$$S(\boldsymbol{\eta}_k) := S(\eta_1, \ldots, \eta_k) := \sum_{\substack{1 \le u \le k \\ \eta_{k+1} = \eta_1}} \delta_{\eta_u \eta_{u+1}}.$$

It is interesting to note what happens for the two special case $\rho = 1$ and $\rho = 0$.

(i) When $\rho = 1$, c is self-adjoint and its free cumulants are given by:

$$\kappa_k(c) = y^{k-1}.$$

Thus c is a compound free Poisson variable $\mathrm{P}(y^{-1}, \delta_{\{y\}})$ with rate $1/y$ and jump distribution $\delta_{\{y\}}$. The moments of c determine the Marčenko-Pastur probability law with parameter y.

(ii) When $\rho = 0$, all odd order free cumulants of c vanish. Moreover, only alternating free cumulants of even order survive, and are given by

$$\kappa_{2k}(c, c^*, c, c^* \ldots, c^*) = \kappa_{2k}(c^*, c, c^*, c, \ldots, c) = y^{2k-1}, \ \forall\ k \geq 1. \tag{3.1}$$

Hence c is a *tracial R-diagonal* element. See Nica and Speicher [2006] for the definition and properties of R-diagonal elements. It is related to a Marčenko-Pastur variable M_y in the following way. Let $\tilde{M}_y$ be a *symmetrized Marčenko-Pastur* variable with parameter y. That is,

$$\kappa_k(\tilde{M}_y) = \begin{cases} \kappa_k(\tilde{M}_y) = y^{k-1} & \text{if } k \text{ is even}, \\ 0 \ \text{ if } k \text{ is odd}. \end{cases}$$

Suppose u is Haar unitary and free of M_y and $\tilde{M}_y$, then it is easy to see that the $*$-distributions of the three variables c (where $\rho = 0$), $u\tilde{M}_y$, and uM_y are identical.

The following theorem states the joint convergence and asymptotic freeness of independent cross-covariance matrices when $y_l \neq 0, \ \forall\ 1 \leq l \leq t$.

Theorem 3.1. *Suppose (X_l, Y_l), $1 \leq l \leq t$ are pairs of $p \times n_l$ random matrices with correlation parameters $\{\rho_l\}$ and whose entries satisfy Assumption I. Suppose $n_l, p \to \infty$ and $p/n_l \to y_l, 0 < y_l < \infty$ for all l. Then the following statements hold for the $p \times p$ cross-covariance matrices $\{C_l := n^{-1} X_l Y_l^*\}$.*

(a) As elements of the C^* probability space $(\mathcal{M}_p(\mathbb{C}), \varphi_p)$, $\{C_l\}$ converge jointly to free variables $\{c_l\}$ where each c_l is a cross-covariance variable with parameters (ρ_l, y_l). The convergence also holds with respect to the state $\tilde{\varphi}_p$ almost surely. The limiting state is tracial.

(b) Let $\Pi := \Pi(\{C_l : \ 1 \leq l \leq t\})$ be any finite degree real matrix polynomial in $\{C_l, C_l^* : \ 1 \leq l \leq t\}$ and which is symmetric. Then the ESD of Π converges weakly almost surely to the compactly supported probability law of the self-adjoint variable $\Pi(\{c_l : \ 1 \leq l \leq t\})$.

Before we prove the theorem, let us make some remarks and give a few examples.

Remark 3.1. *1. Since $\{c_l\}$ are free in Theorem 3.1, their joint free cumulants can be written in principle using the marginal free cumulants.*

2. Capitaine and Casalis [2004] proved the asymptotic freeness of the sample covariance matrices for the case $y_l \neq 0, \ \forall\ 1 \leq l \leq t$. This result follows from Theorem 3.1(a) if we take $\rho = 1$, since in that case, each $X_l = Y_l$ almost surely and each C_l is a sample covariance matrix.

We shall use the following notation:

$$\begin{aligned} \#A &= \text{Number of elements of the set } A, \\ \mathrm{NC}(k) &= \text{Set of non-crossing partitions of } \{1, \ldots, k\}, \\ K(\pi) &= \text{Kreweras complement of a non-crossing partition } pi, \\ |\pi| &= \text{Number of blocks of the partition } \pi. \end{aligned}$$

For two random variable X and Y, $X \stackrel{\mathcal{D}}{=} Y$ will mean that they have identical probability laws. In the following discussion, if we are dealing with only one sequence of matrices, we shall drop the index l. Similarly, if we are working with any two indices, then without loss of generality, we shall take them to be 1 and 2.

Example 3.1. Suppose $\rho_l = 0$, $l = 1, 2$. Since c_1 and c_2 are then free and tracial R-diagonal, by Theorem 15.17 of Nica and Speicher [2006], $c_1 c_2^*$ is also tracial R-diagonal. The free cumulants of $c_1c_2^*$ can be calculated using this fact as follows. First note that all free cumulants except the even order alternating free cumulants are 0. These alternating free cumulants are given by

$$
\begin{aligned}
& \kappa_{2k}(c_1c_2^*, c_2c_1^*, \ldots, c_1c_2^*, c_2c_1^*) = \kappa_{2k}(c_2c_1^*, \ldots, c_1c_2^*, c_2c_1^*, c_1c_2^*) \\
= & \sum_{\substack{\pi,\sigma\in \mathrm{NC}(k)\\ \sigma\leq K(\pi)}} \left(\prod_{V\in\pi} y_1^{2(\#V)-1}\right)\left(\prod_{W\in\sigma} y_2^{2(\#W)-1}\right), \\
& \qquad\qquad \text{(by (3.1) and Exercise 15.24(2) of Nica and Speicher [2006])} \\
= & \; y_1^k y_2^k \sum_{\pi\in \mathrm{NC}(k)} y_1^{k-|\pi|} \sum_{\sigma\leq K(\pi)} y_2^{k-|\sigma|} \\
= & \sum_{\pi\in \mathrm{NC}(k)} \kappa_\pi[y_1M_{y_1}, \ldots, y_1M_{y_1}]\varphi_{K(\pi)}[y_2M_{y_2}, \ldots, y_2M_{y_2}] \quad (M_{y_1} \text{ and } M_{y_2} \text{ are free}) \\
= & \; \varphi((y_1M_{y_1}y_2M_{y_2})^k) = \varphi(((\sqrt{y_1M_{y_1}}y_2M_{y_2}\sqrt{y_1M_{y_1}})^{1/2})^{2k}) \qquad (3.2)
\end{aligned}
$$

Thus for $\rho = 0$, the $*$-distributions of $c_1c_2^*$ and $u\mathrm{P}(1, (\sqrt{y_1M_{y_1}}y_2M_{y_2}\sqrt{y_1M_{y_1}})^{1/2})$ are identical where u is Haar unitary and is free of $\mathrm{P}(1, (\sqrt{y_1M_{y_1}}y_2M_{y_2}\sqrt{y_1M_{y_1}})^{1/2})$. Also note that yM_y is itself a compound free Poisson variable $\mathrm{P}(y^{-1}, \delta_{\{y^2\}})$.

Example 3.2. By Theorem 3.1(b), the ESD of $C + C^*$ converges weakly almost surely to the law whose free cummulants are given by

$$\kappa_k(c + c^*) = y^{k-1} \sum_{\boldsymbol{\eta}_k\in\{1,*\}^k} (\rho^{S(\boldsymbol{\eta}_k)}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{S(\boldsymbol{\eta}_k)0}), \ \forall\ k \geq 1.$$

For $\rho = 0$, the above formula can be simplified to:

$$\kappa_k(c + c^*) = \begin{cases} 2y^{k-1} & \text{if } k \text{ is even} \\ 0, & \text{if } k \text{ is odd} \end{cases}, \ \forall\ k \geq 1.$$

Thus the LSD of $C + C^*$ is the free additive convolution $\mu \boxplus \mu$ where μ is the symmetrized Marčenko-Pastur law with parameter y.

Example 3.3. By Theorem 3.1(b), the ESD of CC^* converges almost surely. The limit law is the law of the self-adjoint variable cc^*. For $\rho \neq 0$, neither the moments nor the free cumulants of cc^* seem to have a simple expression. However, we know that when $\rho = 0$, c is tracial R-diagonal. Hence using Proposition 15.6(2) of Nica and Speicher [2006] and (3.1), we have

$$
\begin{aligned}
\kappa_k(cc^*) &= \sum_{\pi\in \mathrm{NC}(k)} \prod_{V\in\pi} y^{2\#V-1} \\
&= \sum_{r=1}^{k} \#\{\pi \in \mathrm{NC}(k) : \ \pi \text{ has } r \text{ blocks}\} y^{2k-r} \\
&= \sum_{r=1}^{k} \frac{1}{r}\binom{k-1}{r-1}\binom{k}{r-1} y^{2k-r} \\
&= \sum_{r=0}^{k-1} \frac{1}{k-r}\binom{k-1}{k-r-1}\binom{k}{k-r-1} y^{k+r} \\
&= \sum_{r=0}^{k-1} \frac{1}{k-r}\binom{k-1}{r}\binom{k}{r+1} y^{k+r} = \sum_{r=0}^{k-1} \frac{1}{r+1}\binom{k-1}{r}\binom{k}{r} y^{k+r}
\end{aligned}
$$

which is the k-th moment of yM_y. That is, the LSD of $n^{-2}XY^*YX^*$ is the law of the compound free Poisson variable $\mathrm{P}(1, yM_y)$.

Example 3.4. By Theorem 3.1(b), the LSD of $n^{-2}(X_1Y_1^*Y_2X_2^* + X_2Y_2^*Y_1X_1^*)$ exists almost surely. For $\rho \neq 0$, the moment or free cumulant sequence of $(c_1c_2^* + c_2c_1^*)$ obtained from Theorem 3.1(a) cannot be further simplified. However for $\rho = 0$, recall that $c_1c_2^*$ is tracial R-diagonal elementand (3.2) holds. Therefore,

$$\kappa_k(c_1c_2^* + c_2c_1^*) = \begin{cases} 2\varphi(((\sqrt{y_1M_{y_1}}y_2M_{y_2}\sqrt{y_1M_{y_1}})^{1/2})^k) & \text{if } k \text{ is even,} \\ 0 \quad \text{if } k \text{ is odd.} \end{cases}$$

Let $\tilde{\mathrm{P}}(\lambda, X)$ denote a symmetrized compound free Poisson variable—its odd free cumulants are 0 and the even order free cumulants are the same as those of $\mathrm{P}(\lambda, X)$. Then clearly the LSD of $n^{-2}(X_1Y_1^*Y_2X_2^* + X_2Y_2^*Y_1X_1^*)$ is the free additive convolution $\nu \boxplus \nu$ where ν is the probability law of the self-adjoint variable $\tilde{\mathrm{P}}(1, \sqrt{y_1M_{y_1}}y_2M_{y_2}\sqrt{y_1M_{y_1}})^{1/2})$.

Proof of Theorem 3.1. (a) For simplicity, we will prove the result only for the special case where n_l, y_l and ρ_l do not depend on l. It will be clear from the arguments that the same proof works for the general case. Consider a typical monomial

$$(X_{\alpha_1}Y_{\alpha_1}^*)^{\eta_1} \cdots (X_{\alpha_k}Y_{\alpha_k}^*)^{\eta_k}.$$

This product has $2k$ factors. We shall write this monomial in a specific way to facilitate computation. Note that

$$(X_{\alpha_s}Y_{\alpha_s}^*)^{\eta_s} = \begin{cases} X_{\alpha_s}Y_{\alpha_s}^* & \text{if } \eta_s = 1, \\ Y_{\alpha_s}X_{\alpha_s}^* & \text{if } \eta_s = *. \end{cases}$$

For every index l, two types of matrices, namely X and Y are involved. To keep track of this, define

$$(\epsilon_{2s-1}, \epsilon_{2s}) = \begin{cases} (1,2) & \text{if } \eta_s = 1, \\ (2,1) & \text{if } \eta_s = *. \end{cases}$$

Let $\lceil \cdot \rceil$ be the ceiling function. Note that

$$\epsilon_r \neq \epsilon_s \Leftrightarrow \begin{cases} (\epsilon_r, \epsilon_{r+1}) = (\epsilon_{s-1}, \epsilon_s) \Leftrightarrow \eta_{\lceil r/2 \rceil} = \eta_{\lceil s/2 \rceil} & \text{if } r \text{ is odd, } s \text{ is even,} \\ (\epsilon_{r-1}, \epsilon_r) = (\epsilon_s, \epsilon_{s+1}) \Leftrightarrow \eta_{;\lceil r/2 \rceil} = \eta_{\lceil s/2 \rceil} & \text{if } r \text{ is even, } s \text{ is odd.} \end{cases} \tag{3.3}$$

Observe that $\delta_{\epsilon_{2s-1}\epsilon_{2s}} = 0$ for all s. Define

$$A_l^{(\epsilon_{2s-1})} = \begin{cases} X_l & \text{if } \epsilon_{2s-1} = \eta_s = 1 \\ Y_l & \text{if } \epsilon_{2s-1} = 2 \text{ (or } \eta_s = *), \end{cases} \tag{3.4}$$

$$A_l^{(\epsilon_{2s})} = \begin{cases} X_l^* & \text{if } \epsilon_{2s} = 1 \text{ (or } \eta_s = *) \\ Y_l^* & \text{if } \epsilon_{2s} = 2 \text{ (or } \eta_s = 1). \end{cases}$$

Extend the vector $(\alpha_1, \ldots, \alpha_k)$ of length k to the vector of length $2k$ as

$$(\beta_1, \ldots, \beta_{2k}) := (\alpha_1, \alpha_1, \ldots \alpha_k, \alpha_k).$$

We need to show that for all choices of $\alpha_s \in \{1, 2, ..., t\}$ and $\eta_s \in \{1, *\}$,

$$L_p := \frac{1}{pn^k}\mathbb{E}\mathrm{Tr}(A_{\beta_1}^{(\epsilon_1)}A_{\beta_2}^{(\epsilon_2)} \cdots A_{\beta_{2k}}^{(\epsilon_{2k})})$$

converges to the appropriate limit. Upon expansion,

$$L_p = \frac{1}{pn^k} \sum_{I_{2k}} \mathbb{E} \prod_{\substack{1 \le s \le 2k \\ i_{2k+1}=i_1}} A_{\beta_s}^{(\epsilon_s)}(i_s, i_{s+1})$$

where $A_\beta^{(\epsilon)}(i,j)$ denotes the (i,j)th element of $A_\beta^{(\epsilon)}$ for all choices of β, ϵ, i and j, and

$$I_{2k} = \{(i_1, i_2, \dots, i_{2k}):\ 1 \le i_{2s-1} \le p,\ 1 \le i_{2s} \le n,\ 1 \le s \le 2k\}.$$

Observe that the values of these i_j have different ranges p and n, depending on whether j is odd or even.

Note that the expectation of any summand is zero if there is at least one (i_s, i_{s+1}) whose value is not repeated elsewhere in the product. So, as usual, to split up the sum into indices that match, consider any connected bipartite graph between the *distinct* odd and even indices, $I = \{i_{2s-1} : 1 \le s \le k\}$ and $J = \{i_{2s} : 1 \le s \le k\}$. Then we need to consider only those cases where each edge appears at least twice. Hence there can be at most k distinct edges and since the graph is connected,

$$\#I + \#J \le \#E + 1 \le k + 1.$$

By Assumption I, there is a common bound for all expectations involved. Hence, the total expectation of the terms involved in this graph is of the order

$$O(\frac{p^{\#I} n^{\#J}}{pn^k}) = O(p^{\#I + \#J - (k+1)})$$

since $\frac{p}{n} \to y > 0$. As a consequence, only those terms can potentially contribute to the limit for which $\#I + \#J = k+1$. This implies that $\#E = k$. So each edge is repeated exactly twice. Let

$$P_2(2k) = \{\pi : \pi \text{ pair partition of } \{1, 2, \dots, 2k\}\}. \tag{3.5}$$

Then each edges in E corresponds to some $\pi = \{(r,s) : r < s\} \in P_2(2k)$. Let

$$a_{r,s} = \begin{cases} 1 & \text{if } \ r, s \text{ are both odd or both even,} \\ 0 & \text{otherwise.} \end{cases} \tag{3.6}$$

Then we have

$$\lim_{p\to\infty} L_p \ = \ \lim_{n\to\infty} \frac{1}{pn^k} \sum_{I_{2k}} \mathbb{E}[\prod_{s=1}^{2k} A_{\beta_s}^{(\epsilon_s)}(i_s, i_{s+1})] = \sum_{\pi \in P_2(2k)} \lim_{n\to\infty} \frac{1}{pn^k} \sum_{I_{2k}} \prod_{(r,s)\in\pi} E(r,s) \ \text{ say,}$$

where, suppressing the dependence on other variables,

$$\begin{aligned} E(r,s) &= \mathbb{E}\big[A_{\beta_r}^{(\epsilon_r)}(i_r, i_{r+1}) A_{\beta_s}^{(\epsilon_s)}(i_s, i_{s+1})\big] \\ &= \delta_{\beta_r\beta_s}(\rho(1-\delta_{\epsilon_r\epsilon_s}) + \delta_{\epsilon_r\epsilon_s})(\delta_{i_r i_s}\delta_{i_{r+1}i_{s+1}} a(r,s) + \delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}(1-a(r,s))) \\ &= \delta_{\beta_r\beta_s}(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\epsilon_r\epsilon_s})(\delta_{i_r i_s}\delta_{i_{r+1}i_{s+1}} a(r,s) + \delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}(1-a(r,s))). \end{aligned}$$

Recall that $|\rho| \le 1$. Hence each $E(r,s)$ is a sum of two factors—one of them is bounded by $\delta_{i_r i_s}\delta_{i_{r+1}i_{s+1}}$ and the other is bounded by $\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}$. Hence when we expand $\prod_{(r,s)\in\pi} E(r,s)$, each term involves a product of these δ-values. Using arguments similar to those used in the proof of Theorem 3.2.6 in Bose [2018], it is easy to see that the only term that will survive in

$\prod_{(r,s)\in\pi} E(r,s)$ is

$$\prod_{(r,s)\in\pi} \delta_{\beta_r\beta_s}\big(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_r\epsilon_s}\big)\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}(1-a(r,s)).$$

Hence $\lim L_p$ is equal to

$$\sum_{\pi\in\mathrm{NC}_2(2k)} \lim_{n\to\infty}\frac{1}{pn^k}\sum_{I_{2k}}\prod_{(r,s)\in\pi}\delta_{\beta_r\beta_s}\big(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_r\epsilon_s}\big)(1-a(r,s))\prod_{r=1}^{2k}\delta_{i_r i_{\gamma\pi(r)}}. \tag{3.7}$$

But for any $\pi\in\mathrm{NC}_2(2k)$ if $(r,s)\in\pi$ then r and s have different parity and hence $a(r,s)=0$. Let γ denote the cyclic permutation $1\to 2\to\cdots\to 2k\to 1$. Then (3.7) simplifies to

$$\sum_{\pi\in\mathrm{NC}_2(2k)}\prod_{(r,s)\in\pi}\delta_{\beta_r\beta_s}(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_r\epsilon_s})\lim_{n\to\infty}\frac{\#\{I_{2k}: i_r=i_{\gamma\pi(r)}\text{ for all } r\}}{pn^k}. \tag{3.8}$$

Now note that as $\pi\in\mathrm{NC}_2(2k)$, $\gamma\pi$ contains $k+1$ blocks. Moreover, each block of $\gamma\pi$ contains only odd or only even elements. Let

$$S(\gamma\pi) = \text{ Number of blocks in } \ \gamma\pi \ \text{ with only odd elements.}$$

Then the number of blocks of $\gamma\pi$ with only even elements is $k+1-S(\gamma\pi)$. Suppose $\pi\in\mathrm{NC}_2(2k)$ such that $S(\gamma\pi)=m+1$. Then it is clear that

$$\#\{I_{2k}: i_r=i_{\gamma\pi(r)}\text{ for all } r\}=p^{m+1}n^{k+1-(m+1)}$$

and hence using (3.8)

$$\begin{aligned}
\lim L_p &= \sum_{\pi\in\mathrm{NC}_2(2k)}\prod_{(r,s)\in\pi}\delta_{\beta_r\beta_s}(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_r\epsilon_s})\lim_{n\to\infty}\frac{\#\{I_{2k}: i_r=i_{\gamma\pi(r)}\forall r\}}{pn^k}\\
&= \sum_{m=0}^{k-1} y^m \sum_{\substack{\pi\in\mathrm{NC}_2(2k):\\ S(\gamma\pi)=m+1}}\prod_{(r,s)\in\pi}\delta_{\beta_r\beta_s}(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_r\epsilon_s})y^m\\
&= \sum_{m=0}^{k-1} y^m \sum_{\substack{\pi\in\mathrm{NC}_2(2k):\\ S(\gamma\pi)=m+1}}\prod_{(r,s)\in\pi}\delta_{\beta_r\beta_s}(\rho^{T(\pi)}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{T(\pi)0})
\end{aligned}$$

where

$$T(\pi)=\#\{(r,s)\in\pi:\delta_{\epsilon_r\epsilon_s}=0\} \quad\text{for}\quad \pi\in\mathrm{NC}_2(2k).$$

Hence we have proved that $\{C_l: \ 1\le l\le t\}$ converge jointly in $*$-distribution to say $\{c_{l,y,\rho}: 1\le l\le t\}$ which are the limit NCP $(\mathcal{A},\varphi)$. We still have to identify the limit and prove the freeness. For this we need to go from $\mathrm{NC}_2(2k)$ to $\mathrm{NC}(k)$. Define

$$\begin{aligned}
\tilde{J}_i &= \{j\in\{1,2,\ldots,2k\}: \ \beta_j=i\}, \quad 1\le i\le t,\\
\tilde{B}_k &= \{\pi\in\mathrm{NC}_2(2k): \ \pi=\cup_{i=1}^t\pi_i, \ \pi_i\in\mathrm{NC}_2(\tilde{J}_i), \ 1\le i\le t\},\\
\tilde{B}_{m,k} &= \{\pi\in\tilde{B}_k: \ S(\gamma\pi)=m+1\}.
\end{aligned}$$

Note that $\cup_{m=0}^{k-1}\tilde{B}_{m,k} = \tilde{B}_k$ and hence

$$\lim L_p = \sum_{m=0}^{k-1} y^m \sum_{\pi \in \tilde{B}_{m,k}} (\rho^{T(\pi)}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{T(\pi)0}). \tag{3.9}$$

Also define

$$\begin{aligned}
J_i &= \{j \in \{1,2,\ldots,k\} : \ \alpha_j = i\}, \quad 1 \le i \le t, \\
B_k &= \{\pi \in \mathrm{NC}(k) : \ \pi = \cup_{i=1}^t \pi_i, \ \pi_i \in \mathrm{NC}(J_i), \ 1 \le i \le t\}, \\
B_{m,k} &= \{\pi \in B_k : \ \pi \text{ has } m \text{ blocks}\}.
\end{aligned}$$

Note that $\cup_{m=0}^{k-1} B_{m+1,k} = B_k$. For any finite subset $S = \{j_1, j_2, \ldots, j_r\}$ of positive integers, define

$$\tilde{T}(S) = \sum_{\substack{1 \le u \le r \\ j_{r+1}=j_1}} \delta_{\eta_{j_u}\eta_{j_{u+1}}}$$

and for $\pi = \{V_1, V_2, \ldots, V_m\} \in \mathrm{NC}(k)$, define

$$\mathcal{T}(\pi) = \sum_{i=1}^m \tilde{T}(V_i).$$

Consider the bijection $f : \mathrm{NC}_2(2k) \to \mathrm{NC}(k)$ as follows. Take $\pi \in \mathrm{NC}_2(2k)$. Suppose (r,s) is a block of π. Then $\lceil r/2 \rceil$ and $\lceil s/2 \rceil$ are put in the same block in $f(\pi) \in \mathrm{NC}(k)$. Using arguments similar to used in the proof of Lemma 3.2 in Bhattacharjee and Bose [2021], it is easy to see that f is indeed a bijection and is also a bijection between $\tilde{B}_{m,k}$ and $B_{k-m,k}$. Moreover, using (3.3), it is immediate that $T(\pi) = \mathcal{T}(f(\pi)) \ \forall \ \pi \in \tilde{B}_{m,k}$ i.e. $T(f^{-1}(\pi)) = \mathcal{T}(\pi) \ \forall \ \pi \in B_{k-m,k}$.

As an example, let $\pi = \{(1,8),(2,5),(6,7),(3,4),(9,10)\}$. Then $\pi \in \tilde{B}_{2,5}$ and is mapped to $f(\pi) = \{(1,3,4),(3),(5)\} \in B_{3,5}$. Let $(\eta_1,\eta_2,\eta_3,\eta_4,\eta_5) = (1,*,1,*,1)$. Further, $(\epsilon_1,\epsilon_2,\ldots,\epsilon_{10}) = (1,2,2,1,1,2,2,1,1,2)$ and $T(\pi) = \mathcal{T}(f(\pi)) = 3$.

Now (3.9), we have

$$\begin{aligned}
\lim L_p &= \sum_{m=0}^{k-1} y^m \sum_{\pi \in B_{k-m,k}} \left(\rho^{T(f^{-1}(\pi))}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{T(f^{-1}(\pi))0}\right) \\
&= \sum_{m=0}^{k-1} y^m \sum_{\pi \in B_{k-m,k}} \left(\rho^{\mathcal{T}(\pi)}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\mathcal{T}(\pi)0}\right) \\
&= \sum_{m=0}^{k-1} y^{k-m-1} \sum_{\pi \in B_{m+1,k}} \left(\rho^{\mathcal{T}(\pi)}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\mathcal{T}(\pi)0}\right). \qquad (3.10)
\end{aligned}$$

Then, (3.10) implies

$$\begin{aligned}
\varphi(c_{\alpha_1}^{\eta_1} c_{\alpha_2}^{\eta_2} \cdots c_{\alpha_k}^{\eta_k}) &= \sum_{m=0}^{k-1} \sum_{\substack{\pi \in B_{m+1,k} \\ \pi = \{V_1,V_2,\ldots,V_{m+1}\}}} \prod_{l=1}^{m+1} y^{\#V_l - 1}\left(\rho^{\tilde{T}(V_l)}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\tilde{T}(V_l)0}\right) \\
&= \sum_{\pi \in B_k} \prod_{l=1}^{\#\pi} y^{\#V_l - 1}(\rho^{\tilde{T}(V_l)}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\tilde{T}(V_l)0}).
\end{aligned}$$

Hence, by moment-free cumulant relation, we have

$$\begin{aligned}\kappa_\pi[c_{\alpha_1}^{\eta_1}, c_{\alpha_2}^{\eta_2}, \ldots, c_{\alpha_k}^{\eta_k}] &= 0 \text{ for all } \pi \in \mathrm{NC}(k) - B_k, \\ \kappa_\pi[c_{\alpha_1}^{\eta_1}, c_{\alpha_2}^{\eta_2}, \ldots, c_{\alpha_k}^{\eta_k}] &= \prod_{l=1}^{\#\pi} y^{\#V_l - 1}(\rho^{\tilde{T}(V_l)}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\tilde{T}(V_l)0}) \text{ for all } \pi \in B_k.\end{aligned}$$

This implies that

$$\kappa_k(c_{\alpha_1}^{\eta_1}, c_{\alpha_2}^{\eta_2}, \ldots, c_{\alpha_k}^{\eta_k}) = \begin{cases} y^{k-1}\big(\rho^{\mathcal{T}(\mathbf{1}_k)}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\mathcal{T}(\mathbf{1}_k)0}\big), & \text{if } \alpha_1 = \alpha_2 = \cdots = \alpha_k, \\ 0 \ \text{ otherwise.} \end{cases}$$

Therefore $\{c_l : \ 1 \le l \le t\}$ are free across l, and the marginal free cumulant of order k is

$$\begin{aligned}\kappa_k(c_l^{\eta_1}, c_l^{\eta_2}, \ldots, c_l^{\eta_k}) &= y^{k-1}(\rho^{S(\boldsymbol{\eta}_k)}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{S(\boldsymbol{\eta}_k)0}) \\ \text{where } S(\boldsymbol{\eta}_k) &= \mathcal{T}(\mathbf{1}_k) = \sum_{\substack{1\le u\le k \\ u_{k+1}=u_1}} \delta_{\eta_u \eta_{u+1}}.\end{aligned}$$

This completes the proof of Theorem 3.1 (a) for the state φ_p for the special case when the values of ρ_l and of y_l are same. It is easy to see that the above proof continues to hold for the general case, except for notational complexity. We omit the details.

Now we argue convergence with respect to the state $\tilde{\varphi}_p$. Consider any polynomial Π. We have already shown convergence of $p^{-1}\mathbb{E}[\mathrm{Trace}(\Pi^k)]$ to say β_k for all k. We need to show that $p^{-1}\mathrm{Trace}(\Pi^k)$ converges to β_k almost surely. For this it is enough to show that for every k,

$$\mathbb{E}\Big[p^{-1}\mathrm{Trace}(\Pi^k) - p^{-1}E[\mathrm{Trace}(\Pi^k)]\Big]^4 = O(p^{-2}).$$

Then an application of the Borel-Cantelli Lemma would finish the proof. Now, a further simplification is that it is enough to prove this for any monomial. Then the above estimate is obtained by a counting argument, which is similar to, but simpler than what has already been used in the proof so far. We omit the details. See Section 2 in the Supplementary material of Bhattacharjee and Bose [2016b] for similar arguments. This complete the proof of (a).

(b) Now suppose that Π is symmetric. Then by the above argument, all moments of Π converge and there is a $C > 0$, depending on Π such that the limiting kth moment is bounded by C^k for all k. This implies that these moments define a unique probability law say μ with support contained in $[-C, \ C]$, and hence the EESD of Π converges weakly to μ. Again, the almost sure convergence of the ESD can be established by the arguments discussed above. We omit the details. Note that this argument works only if Π is a symmetric matrix. □

3.2 Convergence of $\{C_l : 1 \le l \le t\}$ when $y = 0$

Now suppose $pn^{-1} \to 0$ as $n, p \to \infty$. Then Theorem 3.1 leads to a degenerate distribution when $y = 0$. In this case, we need a centering as well as a different scaling for a non-degenerate limit to exist. Let us quickly recall a known result. Consider the sample covariance matrix $S = n^{-1}XX^*$ where the entries X_{ij} of X are independent and $\mathbb{E}(X_{11}) = 0, \ \mathrm{Var}(X_{11}) = 1, \mathbb{E}(X_{11}^4) < \infty$. Then the empirical spectral distribution (ESD) of $\sqrt{np^{-1}}(S - I_p)$ converges weakly almost surely to the standard semi-circle law where I_p is the identity matrix of order p. For proof, one can see Bose [2018]. This proof actually shows that when all moments are finite, then the moments of the ESD converge to the moments of the semi-circle law and hence there is convergence as elements of $(\mathcal{M}_p(\mathbb{C}), \varphi_p)$ to a semi-circular variable.

We provide a generalization of this result involving cross-covariance matrices. Before we state the result, we need to recall the definition of elliptic variables.

Definition 3.2. Suppose $(\mathcal{A}, \varphi)$ is an NCP. An element $e \in \mathcal{A}$ is said to be an *elliptic* variable with parameter ρ, $-1 \leq \rho \leq 1$ if its free cumulants of order one and of order greater than 2 are zero and its second order free cumulants are given by

$$\kappa_2(e, e) = \kappa_2(e^*, e^*) = \rho, \quad \kappa_2(e, e^*) = \kappa_2(e^*, e) = 1.$$

An elliptic variable has the following representation. Suppose s_1 and s_2 are two free standard semi-circular variables. Define

$$e = \sqrt{\frac{1+\rho}{2}} s_1 + \sqrt{-1}\sqrt{\frac{1-\rho}{2}} s_2.$$

Then e is an elliptic variable with parameter ρ. Note that $\rho = 1$ and $\rho = 0$ yield respectively the standard semi-circular and the standard circular variable.

We use the following crucial fact for variables to be elliptic *and* free: Variables $\{e_i, 1 \leq i \leq t\}$ are elliptic with parameters $\{\rho_i, 1 \leq i \leq t\}$ on an NCP $(\mathcal{A}, \varphi)$ and are free if and only if, for all $k \geq 1$, for all $1 \leq i \leq k$, $\epsilon_i \in \{1, *\}$ and $\tau_i \in \{1, \ldots, t\}$, the following holds for the joint moments:

$$\varphi(e_{\tau_1}^{\epsilon_1} e_{\tau_2}^{\epsilon_2} \cdots e_{\tau_{2k}}^{\epsilon_{2k}}) = \sum_{\pi \in \mathrm{NC}_2(2k)} \rho_1^{T_1(\pi)} \ldots \rho_m^{T_m(\pi)} \prod_{(r,s) \in \pi} \delta_{\tau_r \tau_s}. \tag{3.11}$$

where

$$T_\tau(\pi) := \#\{(r, s) \in \pi \ : \ \delta_{\epsilon_r \epsilon_s} = 1, \tau_r = \tau_s = \tau\}.$$

Note that, it is understood that all odd order moments are 0.

Now we can state our theorem.

Theorem 3.2. *Suppose Assumption I holds with $y_l = 0$, $\forall\ 1 \leq l \leq t$.*

(a) Then $E_l = \sqrt{n_l p^{-1}}(C_l - \rho_l I_p), 1 \leq l \leq t$ as elements of $(\mathcal{M}_p, \varphi_p)$ converge jointly to free elliptic variables $e_1, ..., e_t$, with parameters ρ_l^2. The convergence also holds with respect to the state $\tilde{\varphi}_p$ almost surely. The limiting state is tracial.

(b) Let $\Pi := \Pi(\{E_l : \ 1 \leq l \leq t\})$ be any finite degree real matrix polynomial in $\{E_l, E_l^* : \ 1 \leq l \leq t\}$ and which is symmetric. Then the ESD of Π converges weakly almost surely to the compactly supported probability law of the self-adjoint variable $\Pi(\{e_l : \ 1 \leq l \leq t\})$.

Example 3.5. Theorem 3.2 is useful to find the LSD of any appropriately centered and scaled symmetric polynomial of $\{C_l : \ 1 \leq l \leq t\}$. For example, consider

$$\begin{aligned}
\Pi &= \sqrt{\frac{\min(n_1, n_2)}{p}}(C_1 + C_1^* + C_1 C_2^* + C_2 C_1^* - 2\rho_1(1 + \rho_2) I_p) \\
&= \sqrt{\frac{\min(n_1, n_2)}{p}}\Big[(C_1 + C_1^* - 2\rho_1 I_p) + (C_1 - \rho_1 I_p)(C_2 - \rho_2 I_P)^* + (C_2 - \rho_2 I_p)(C_1 - \rho_1 I_p)^* \\
&\qquad + \rho_2(C_1 + C_1^* - 2\rho_1 I_p) + \rho_1(C_2 + C_2^* - 2\rho_2 I_p)\Big] \\
&= \sqrt{\min(n_1, n_2)}\Big[n_1^{-1/2}(E_1 + E_1^*) + \sqrt{p} n_1^{-1/2} n_2^{-1/2}(E_1 E_2^* + E_2 E_1^*) \\
&\qquad + n_1^{-1/2}\rho_2(E_1 + E_1^*) + n_2^{-1/2}\rho_1(E_2 + E_2^*)\Big].
\end{aligned} \tag{3.12}$$

Hence it follows that

$$\Pi \xrightarrow{*} \begin{cases} (1+\rho_2)(e_1+e_1^*) + y_{12}^{1/2}(e_2+e_2^*) & \text{if} \quad \lim_{p\to\infty} n_1/n_2 = y_{12} \leq 1, \\ \\ y_{12}^{-1/2}(1+\rho_2)(e_1+e_1^*) + (e_2+e_2^*) & \text{if} \quad \lim_{p\to\infty} n_1/n_2 = y_{12} \geq 1. \end{cases} \tag{3.13}$$

Note that depending on whether $y_{12} = 0$ or ∞, the second or the first term respectively drop out from the limit sums. We can conclude that the LSD of Π also exists almost surely and equals the probability law of the self-adjoint variable in (3.13)

Proof of Theorem 3.2. (a) We shall give the detailed proof only for the special case where all the n_l's ρ_l's are equal to say n and ρ respectively. For any $k \geq 1$ and $\epsilon_1, ..., \epsilon_k \in \{1, *\}$, we will consider the limit of the following as $p, n \to \infty$ with $p/n \to 0$:

$$\frac{1}{p}\mathbb{E}\mathrm{Tr}(E_{l_1}^{\epsilon_1} \cdots E_{l_k}^{\epsilon_k}) = \frac{1}{p(np)^{k/2}} \sum_{\substack{1\leq i_1,...,i_k\leq p \\ 1\leq j_1,...,j_k\leq n}} \mathbb{E}\Big[\prod_{t=1}^{k} \big(a_{l_t i_t j_t} a_{l_t i_{t+1} j_t} - \rho\delta_{i_t i_{t+1}}\big)\Big] \tag{3.14}$$

with the understanding that $i_{k+1} = i_1$, and as ordered pairs, for all $1 \leq r \leq k$,

$$(a_{l_r i_r j_r}, a_{l_r i_{r+1} j_r}) = \begin{cases} (x_{i_r j_r}^{(l_r)}, y_{i_{r+1} j_r}^{(l_r)}) & \text{if} \quad \epsilon_r = 1, \\ \\ (y_{i_r j_r}^{(l_r)}, x_{i_{r+1} j_r}^{(l_r)}) & \text{if} \quad \epsilon_r = *. \end{cases}$$

Consider the following collection of all *ordered* pairs of indices that appear in the above formula:

$$P = \{(i_r, j_r), (i_{r+1}, j_r), 1 \leq r \leq k\}.$$

(i) Suppose there is a pair say, $(i_r, j_r) \in P$ that appears only once. Then $(i_r, j_r) \neq (i_{r+1}, j_r)$ and hence $i_r \neq i_{r+1}$. As a consequence, the variable $a_{l_r i_r j_r}$ is independent of all other variables and we get

$$\mathbb{E}\Big[\prod_{t=1}^{k} \big(a_{l_t i_t j_t} a_{l_t i_{t+1} j_t} - \rho\delta_{i_t i_{t+1}}\big)\Big] \quad = \mathbb{E}[a_{l_r i_r j_r}]\mathbb{E}\Big[a_{l_r i_{r+1} j_r} \prod_{t\neq r} \big(a_{l_t i_t j_t} a_{l_t i_{t+1} j_t} - \rho\delta_{i_t i_{t+1}}\big)\Big] = 0. \tag{3.15}$$

The same conclusion holds if a pair (i_{r+1}, j_r) occurs only once in P. So we can restrict attention to the subset of P where each pair is repeated at least twice, and we continue to call this reduced subset by P.

(ii) Suppose in P there is a j_r such $j_r \neq j_s$ for all $s \neq r$, then the pair $(a_{l_r i_r j_r}, a_{l_r i_{r+1} j_r})$ is independent of all other factors in the product. Hence,

$$\mathbb{E}\Big[\prod_{t=1}^{k} \big(a_{l_t i_t j_t} a_{l_t i_{t+1} j_t} - \rho\delta_{i_t i_{t+1}}\big)\Big] = \underbrace{\mathbb{E}[a_{l_r i_r j_r} a_{l_r i_{r+1} j_r} - \rho\delta_{i_r i_{r+1}}]}_{=0} \mathbb{E}\Big[\prod_{t\neq r} \big(a_{l_t i_t j_t} a_{l_t i_{t+1} j_t} - \rho\delta_{i_t i_{t+1}}\big)\Big] = 0.$$

Hence we restrict attention to the subset of P where each j_r occurs in at least four pairs i.e. in $(i_r, j_r), (i_{r+1}, j_r)$ and also in $(i_s, j_s), (i_{s+1}, j_s)$ for some $s \neq r$. We continue to call this reduced subset by P, and the corresponding pairs, edges. If $j_r = j_s$ then they are said to be matched and likewise for the i-vertices.

Define the set of vertices V_I and V_J which are the distinct indices from $\{i_1, ..., i_k\}$ and $\{j_1, ..., j_k\}$ respectively. Note that there are at most $2k$ edges in P but each edge appears at least twice. Let E be the set of distinct edges between the vertices in V_I and V_J. This defines a simple connected bi-partite graph. Then clearly, $\#E \leq k$. Since every j-index was originally matched, $\#V_J \leq k/2$. We also know from the connectedness property that

$$\#V_I + \#V_J \leq \#E + 1 \leq k + 1. \tag{3.16}$$

Hence the contribution to (3.14) is bounded above by

$$O\left(\frac{p^{\#V_I} n^{\#V_J}}{p^{k/2+1} n^{k/2}}\right) = O\left(\frac{p^{k+1-\#V_J} n^{\#V_J}}{p^{k/2+1} n^{k/2}}\right) = O\left(\left(\frac{p}{n}\right)^{k/2-\#V_J}\right). \tag{3.17}$$

If $\#V_J < \dfrac{k}{2}$ then the above expression goes to 0. So the only possible non-zero contribution to the limit of (3.14) will come when $\#V_J = k/2$. This immediately shows that if k is odd, then we do not get such a contribution and hence,

$$\frac{1}{p}\mathbb{E}\mathrm{Tr}(E_{l_1}^{\epsilon_1} \cdots E_{l_k}^{\epsilon_k}) \to 0.$$

So now consider the (contributing) case when k is even and $\#V_J = k/2$, that is, let $k = 2m$ and $\#V_J = m$. Then

$$O\left(\frac{p^{\#V_I} n^{\#V_J}}{p^{k/2+1} n^{k/2}}\right) = O(p^{\#V_I-(k/2+1)}) = O(p^{\#V_I-(m+1)}).$$

On the other hand, from (3.16), when $k = 2m$ and $\#V_J = m$, we get $\#V_I \leq m + 1$. So for a possible non-zero contribution, we must have $\#V_I = m + 1$. This implies that

$$m + 1 + m = \#V_I + \#V_J \leq \#E + 1 \leq 2m + 1$$

and hence $\#E = 2m$. In other words, each edge must appear exactly 2 times.

Suppose $(i_r, j_r) = (i_{r+1}, j_r)$ for some r. Since each edge appears exactly twice, this pair will be independent of all others and therefore

$$\mathbb{E}[\prod_{t=1}^{k}(a_{l_t i_t j_t} a_{l_t i_{t+1} j_t} - \rho\delta_{i_t i_{t+1}})] = \underbrace{\mathbb{E}[a_{l_r i_r j_r} a_{l_r i_{r+1} j_r} - \rho\delta_{i_r i_{r+1}}]}_{=0} \mathbb{E}[\prod_{t \neq r}(a_{l_t i_t j_t} a_{l_t i_{t+1} j_t} - \rho\delta_{i_t i_{t+1}})] = 0.$$

Hence, such a combination cannot contribute to (3.14). So we may assume from now on that for every r, $i_r \neq i_{r+1}$ and hence $\delta_{i_r i_{r+1}} = 0$ always. We continue to call this reduced subset by P. As a consequence, (3.14) reduces to

$$\frac{1}{p}\mathbb{E}\mathrm{Tr}(E_{l_1}^{\epsilon_1} \cdots E_{l_k}^{\epsilon_k}) = \frac{1}{p^{m+1} n^m} \sum_{P} \mathbb{E}(\prod_{r=1}^{2m} a_{l_r i_r j_r} a_{l_r i_{r+1} j_r}).$$

Due to the preceding discussion, we have two situations. Suppose $(i_r, j_r) = (i_s, j_s)$ for some $s \neq r$. Note that j_r and j_s are also adjacent to i_{r+1} and i_{s+1} respectively. Due to the nature of the edge set P, this forces $i_{r+1} = i_{s+1}$. Similarly if $(i_r, j_r) = (i_{s+1}, j_s)$ then it would force $i_{r+1} = i_s$. Let $\mathcal{P}(2m)$ be the set of all possible pair partitions of the set $\{1, ..., 2m\}$. We will think of each block in the partition to represent the equal pairs of edges in the graph. Now recalling the definition (3.15), the moment structure, and the above developments, it is easy to verify that the possibly contributing part of the above sum (and hence of (3.14)) can be re-expressed as

$$\sum_{\substack{i_1,\dots,i_{2m}\\ j_1,\dots,j_{2m}}}\sum_{\pi\in\mathcal{P}_2(2m)}\frac{1}{p^{m+1}n^m}\prod_{(r,s)\in\pi}\Big[\delta_{l_r l_s}\delta_{\epsilon_r\epsilon_s}(\rho^2\delta_{i_r i_{s+1}}\delta_{j_r j_s}\delta_{i_{r+1}i_s}+\delta_{i_r i_s}\delta_{j_r j_s}\delta_{i_{r+1}i_{s+1}})$$
$$+\delta_{l_r l_s}(1-\delta_{\epsilon_r\epsilon_s})(\rho^2\delta_{i_r i_s}\delta_{j_r j_s}\delta_{i_{r+1}i_{s+1}}+\delta_{j_r j_s}\delta_{i_r i_{s+1}}\delta_{i_{r+1}i_s})\Big] \tag{3.18}$$

which equals

$$\sum_{\substack{i_1,\dots,i_{2m}\\ j_1,\dots,j_{2m}}}\sum_{\pi\in P_2(2m)}\frac{1}{p^{m+1}n^m}\prod_{(r,s)\in\pi}\delta_{l_r l_s}\delta_{j_r j_s}\Big[\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}\big(\rho^2\delta_{\epsilon_r\epsilon_s}+(1-\delta_{\epsilon_r\epsilon_s})\big)+\text{other terms}\Big]. \tag{3.19}$$

We first consider a special case and extract some crucial information that will be useful for the general case. Suppose $t=1,\rho=1$. Then $Y=X$, and we can take ϵ_r to be same for all r,s. We know that $\sqrt{np^{-1}}(XX^*-I_p)$ converges to a semi-circular variable. This immediately implies that the limit of (3.14) and hence of (3.18) and (3.19) in this special case equals $\#\mathrm{NC}_2(2m)=C_m$, ;the mth Catalan number. This means that

$$\sum_{i_1,i_2,\dots,i_m=1}^{p}\sum_{j_1,j_2,\dots,j_m=1}^{n}\sum_{\pi\in P_2(2m)}\frac{1}{p^{m+1}n^m}\prod_{(r,s)\in\pi}\delta_{j_r j_s}(\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}+\delta_{i_r i_s}\delta_{i_{r+1}i_{s+1}})\to\sum_{\pi\in\mathrm{NC}_2(2m)}1.$$

Now we note that for $\pi\in\mathrm{NC}_2(2m)$,

$$\sum_{i_1,\dots,i_{2m},j_1,\dots,j_{2m}}\frac{1}{p^{m+1}n^m}\prod_{(r,s)\in\pi}\delta_{j_r j_s}\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}=1.$$

The reason is as follows. If $(r,s)\in\pi$ then $j_r=j_s$, and there are n^m ways of choosing the $j-$indices. Now let $\gamma=(1,2,\dots,2m)$ be the cyclic permutation. Then we note

$$\prod_{(r,s)\in\pi}\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}=\prod_{(r,s)\in\pi}\delta_{i_r i_{\gamma\pi(r)}}\delta_{i_s i_{\gamma\pi(s)}}=\prod_{r=1}^{2m}\delta_{i_r i_{\gamma\pi(r)}}.$$

So $i_r=i_{\gamma\pi(r)}$ for each r if the above product is 1. But this means i is constant on each block of $\gamma\pi$. As $\pi\in\mathrm{NC}_2(2m)$, $|\gamma\pi|=m+1$ and thus there are p^{m+1} choices in total for the i's.

These arguments establish that

$$\lim_{n\to\infty}\sum_{\pi\in P_2(2m)}\sum_{\substack{i_1,\dots,i_{2m}\\ j_1,\dots,j_{2m}}}\frac{1}{p^{m+1}n^m}\prod_{(r,s)\in\pi}\delta_{j_r j_s}(\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}+\delta_{i_r i_s}\delta_{i_{r+1}i_{s+1}}) \tag{3.20}$$
$$=\lim\sum_{\pi\in\mathrm{NC}_2(2m)}\frac{1}{p^{m+1}n^m}\prod_{(r,s)\in\pi}\delta_{j_r j_s}\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}. \tag{3.21}$$

This implies that the rest of the terms in (3.20) must go to 0, since these quantities are all non-negative.

Now we consider the general case where $t\geq 1$ but for the moment assume that ρ_l's are equal but the common value is not necessarily equal to 1. Going back to (3.19) and noting that for each $\pi\in P_2(2m)$ and for each $(r,s)\in\pi$,

$$|\delta_{l_r l_s}\delta_{j_r j_s}(\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}(\rho^2\delta_{\epsilon_r\epsilon_s}+(1-\delta_{\epsilon_r\epsilon_s}))+\text{other terms})|\leq 1,$$

we conclude that the expression in (3.19) converges to

$$\sum_{\pi \in \mathrm{NC}_2(2m)} \prod_{(r,s)\in\pi} \delta_{l_r l_s}(\rho^2 \delta_{\epsilon_r \epsilon_s} + (1-\delta_{\epsilon_r \epsilon_s})) \quad \text{as} \quad n \to \infty.$$

Now note that $\rho^2 \delta_{\epsilon_r \epsilon_s} + (1-\delta_{\epsilon_r \epsilon_s}) = (\rho^2)^{\delta_{\epsilon_r \epsilon_s}}$. Let $T(\pi) = \#\{(r,s) \in \pi : \epsilon_r = \epsilon_s\}$. Then the limit can be re-expressed as

$$\sum_{\pi \in \mathrm{NC}_2(2m)} \prod_{(r,s)\in\pi} (\rho^2)^{\delta_{\epsilon_r \epsilon_s}} \delta_{l_r l_s} = \sum_{\pi \in \mathrm{NC}_2(2m)} \prod_{(r,s)\in\pi} \delta_{l_r l_s} (\rho^2)^{T(\pi)}.$$

If $e_1, \ldots, e_{2m}$ are freely independent elliptic elements each with parameter ρ^2 in some NCP $(\mathcal{A}, \varphi)$, then the above expression is nothing but $\varphi(e_{l_1}^{\epsilon_1} \ldots e_{l_{2m}}^{\epsilon_{2m}})$. This proves the $*$-convergence (for the special case where all ρ_l's and all n_l's are identical.

In particular, if $\rho_l = 1$ for all l, then $\{E^{(l)}, 1 \le i \le t\}$ are asymptotically free semi-circular variables and if $\rho_l = 0$ for all l, then they are asymptotically free circular variables.

If we follow the above proof carefully, then it is clear that when we have possibly different ρ_l's, and n_l's, the argument for negligibility of the terms remains valid. Once we make the allowance for different ρ, the rest of the proof carries through and the product of $\{\rho_l^{T_l(\pi)}\}$ emerges in the limit. This completes the proof of the first part of (a).

As discussed in the proof of Theorem 3.1, convergence with respect to the state $\tilde{\varphi}_p$ follows by the Borel-Cantelli Lemma after it is established that

$$\mathbb{E}\Big[p^{-1}\mathrm{Trace}(\Pi^k) - p^{-1}E[\mathrm{Trace}(\Pi^k)]\Big]^4 = O(p^{-2}). \tag{3.22}$$

Proof of (3.22) proceeds along lines similar to that in the proof of the first part of (a). We omit the details. See the proof of Theorem 3.5 in Bhattacharjee and Bose [2016a] for similar arguments. This complete the proof of (a).

(b) This is also similar to the proof of Theorem 3.1(b). Let the finite degree polynomial Π of $\{E_l, E_l^* : 1 \le l \le t\}$ be symmetric. By Theorem 3.2(a), all moments of Π converge almost surely. Also there is a $C > 0$, depending on Π such that the limiting kth moment is bounded by C^k for all k. This implies that these moments define a unique probability law whose support is a subset of the interval $[-C,\ C]$, and hence the ESD of Π converges weakly almost surely to this law. □

Figure 2 reports the simulation results for a few polynomials when p/n is small (0.03) for different values of ρ.

References

K. Adhikari and A. Bose. Brown measure and asymptotic freeness of elliptic and related matrices. *Random Matrices: Theory and Applications*, 8(2):1950007, 2019. doi: 10.1142/S2010326319500072.

G. Akemann, M. Phillips, and H. Sommers. Characteristic polynomials in real Ginibre ensembles. *Journal of Physics A: Mathematical and Theoretical*, 42(1):012001, 2008.

G. Akemann, S.-S. Byun, and N.-G. Kang. A non-hermitian generalisation of the Marchenko–Pastur distribution: from the circular law to multi-criticality. *Annales Henri Poincaré*, pages 1–34, 2020. Online: DOI: 10.1007/s00023-020-00973-7.

Z. Bai and J. W. Silverstein. *Spectral Analysis of Large Dimensional Random Matrices*. Springer, 2010.

M. Bhattacharjee and A. Bose. Polynomial generalizations of the sample variance-covariance matrix when $pn^{-1} \to 0$. *Random Matrices: Theory and Applications*, 5(04):1650014, 2016a.

M. Bhattacharjee and A. Bose. Large sample behaviour of high dimensional autocovariance matrices. *Annals of Statistics*, 44(2):598–628, 2016b.

M. Bhattacharjee and A. Bose. Matrix polynomial generalizations of the sample variance-covariance matrix when $pn^{-1} \to y \in (0, \infty)$. *Indian Journal of Pure and Applied Mathematics*, 48(4):575–607, 2017. Erratum 49: 783-788, 2018.

M. Bhattacharjee and A. Bose. Asymptotic freeness of sample covariance matrices via embedding. *arXiv preprint arXiv:2101.06481*, 2021.

A. Bose. *Patterned Random Matrices*. CRC Press, 2018.

M. Capitaine and M. Casalis. Asymptotic freeness by generalized moments for Gaussian and Wishart matrices. application to beta random matrices. *Indiana University Mathematics Journal*, pages 397–431, 2004.

V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114):507–536, 1967.

H. H. Nguyen and S. O′Rourke. The elliptic law. *International Mathematics Research Notices*, pages 7620–7689, 2015.

A. Nica and R. Speicher. *Lectures on the Combinatorics of Free Probability*. Cambridge University Press, Cambridge, UK, 2006.

Vinayak and L. Benet. Spectral domain of large nonsymmetric correlated Wishart matrices. *Physical Review E*, 90(4):042109, 2014.

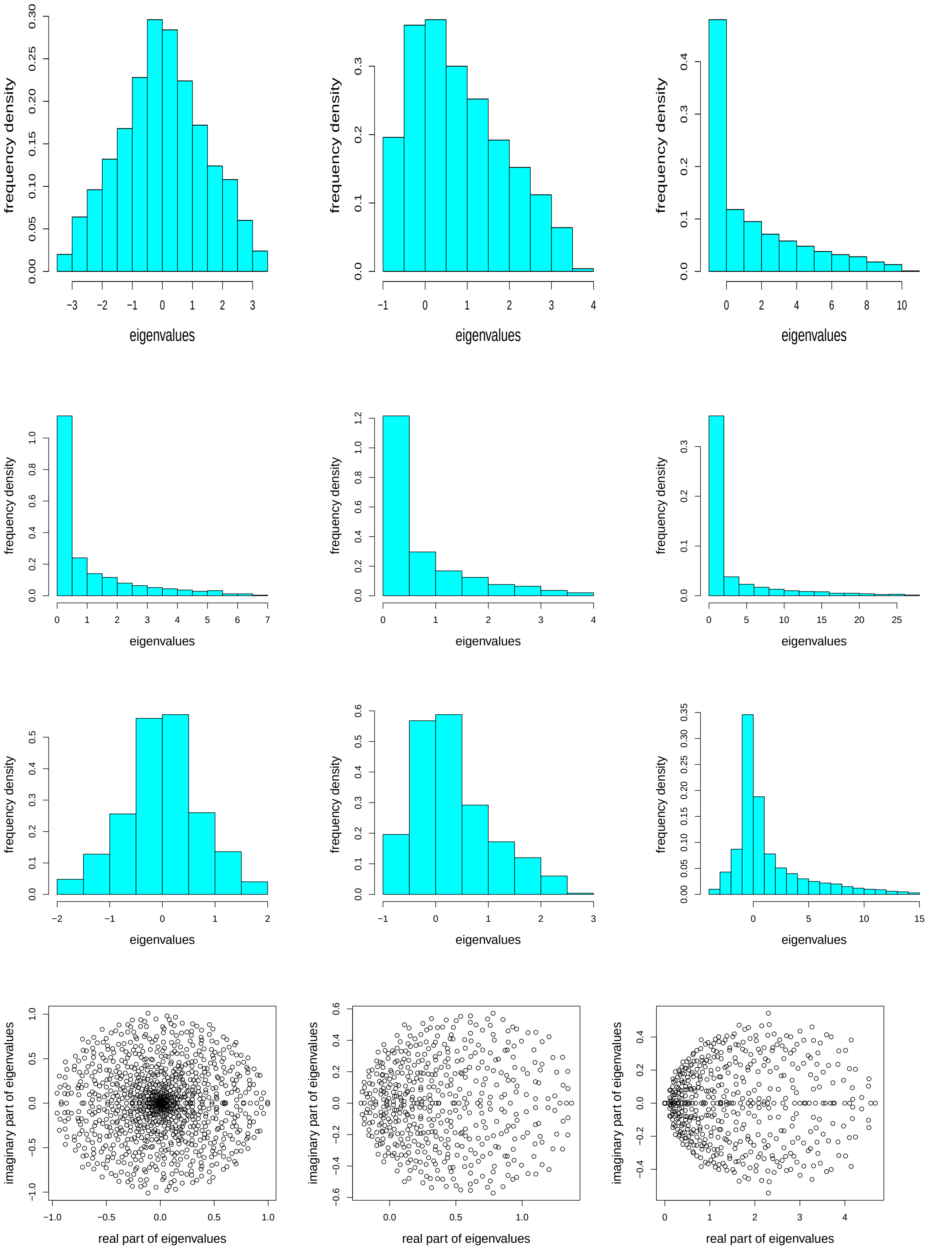

Figure 1: Histogram for ESD of $(C+C^*)$, CC^* and $C_1C_2^*+C_2C_1^*$ $(n_1 = n,\ n_2 = 2n,\ \rho_1 = \rho_2 = \rho)$ in Rows 1-3 respectively. ESD of C in Row 4. Column 1: $n = p = 500$, $\rho = 0$, Column 2: $n = 1000$, $p = 500$, $\rho = 0.4$ Column 3: $n = 500$, $p = 1000$, $\rho = 0.8$.

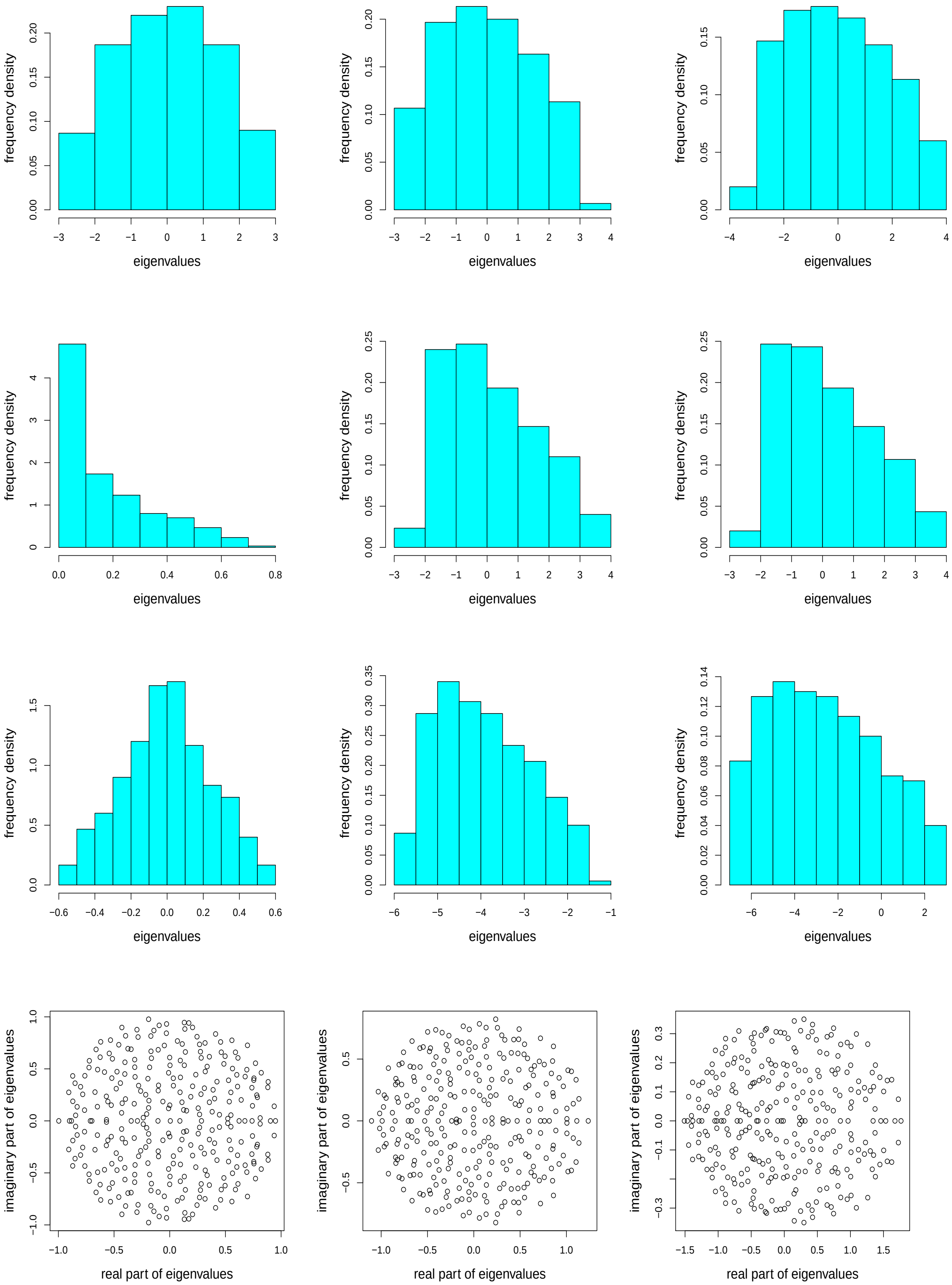

Figure 2: Histogram for ESD of $\sqrt{np^{-1}}(C+C^*-2\rho \mathrm{I}_p)$, $\sqrt{np^{-1}}(CC^*-\rho^2\mathrm{I}_p)$ and $\sqrt{n_2p^{-1}}(C_1C_2^*+C_2C_1^*-2\rho^2\mathrm{I}_p)$ $(n_1=n,\ n_2=2n,\ \rho_1=\rho_2=\rho)$ in Rows 1-3 respectively. Scatter plot for ESD of $\sqrt{np^{-1}}(C-\rho\mathrm{I}_p)$ in Row 4. All figures are for $n=10000$ and $p=300$. The values of ρ are $0, 0.4, 0.8$ respectively in Columns 1-3.