ALEA, Lat. Am. J. Probab. Math. Stat. 20, 395–423 (2023)
DOI: 10.30757/ALEA.v20-14

Joint convergence of sample cross-covariance matrices

Monika Bhattacharjee, Arup Bose and Apratim Dey

Dept. of Math., Indian Institute of Technolgy Bombay
E-mail address: monaiidexp.gamma@gmail.com

Stat.-Math Unit, Indian Statistical Institute
E-mail address: bosearu@gmail.com

Dept. of Stat., Stanford University
E-mail address: apd1995@stanford.edu

Abstract. Suppose X and Y are $p \times n$ matrices each with entries that have mean 0, variance 1, and which have all moments of any order that 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 which is a polynomial in the above scaled and centered matrices exists and has compact support.

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 empirical spectral distribution (ESD) 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

Received by the editors March 28th, 2021; accepted May 20th, 2022.

2010 *Mathematics Subject Classification.* Primary 60B20, Secondary 46L54.

Key words and phrases. 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.

First Author has been supported by the seed grant Spons/MA/10001870-1/2019 from IIT Bombay and Start-up Research Grant RD/0120-SERB000-031 from SERB, Dept. of Science and Technology, Govt. of India. Second Author has been supported by J.C. Bose National Fellowship, Dept. of Science and Technology, Govt. of India.

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) when the number of rows and columns of X grow at the same rate. The joint convergence of the generalized covariance matrices and the convergence of the ESD of their symmetric matrix polynomials was dealt in Bhattacharjee and Bose (2016b, 2017). In particular, when $y = 0$, the asymptotic freeness of independent $\sqrt{np^{-1}}(S - I_p)$-matrices follows as a consequence.

The semi-circle law is a central probability law in free probability and originally arose from the study of a Wigner matrix W_n. In its simplest form, this is a real symmetric matrix whose entries are i.i.d. with mean 0 and variance 1. The LSD 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 its major and minor diameters being $2(1+\rho)$ and $2(1-\rho)$. 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. Suppose X and Y are two $p \times n$ random matrices where the entries (x_{ij}, y_{ij}), $1 \leq i \leq p$, $1 \leq j \leq n$ are independent bivariate random variables with mean 0, variance 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.

Apparently, there are only a few articles on the behaviour of general cross-covariance matrices. Akemann et al. (2009) and Vinayak and Benet (2014) respectively analysed its characteristic polynomial and its spectral domain, and Akemann et al. (2021) established the weak convergence of the ESD when the entries are complex Gaussian, $0 < y < \infty$, and $\rho \in [-1, 1]$. In particular, joint convergence of independent cross-covariance matrices and LSD of their polynomials do not appear to have been studied earlier for general $\rho \in (-1, 1)$.

We study the joint convergence of independent copies $\{C_l\}$, *appropriately centered and scaled*, 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.

A square matrix of order p whose (i, j)-th entry is a_{ij} for all $1 \leq i, j \leq p$ will be denoted by $A = ((a_{ij}))_{p\times p}$. 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, $\varphi_p(aa^*) \geq 0$, and $\varphi_p(ab) = \varphi_p(ba)$ for all a, b). Elements $\{A_l\}$, $1 \leq i \leq t$ from $\mathcal{M}_p(\mathbb{C})$ are said to converge jointly if for every polynomial $\Pi(A_l, 1 \leq l \leq t)$ in the variables $\{A_l, A_l^*\}$, $\varphi_p(\Pi(A_l, 1 \leq l \leq t)) = p^{-1}\mathbb{E}[\text{Trace}(\Pi(A_l, 1 \leq l \leq t))]$ converges as $p \to \infty$. In that case, let the $*$-algebra generated by indeterminates $\{a_l\}$ (called the limit algebra generated by $\{A_l\}$) be equipped with the tracial and positive state φ defined through the above limits as

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

The matrices $\{A_l\}$ are said to be asymptotically free if the limit variables $\{a_l\}$ are free with respect to the limit state φ in the limit algebra.

Two different cases arise: for each l, either (i) $n_l, p \to \infty, p/n_l \to y_l, 0 < y_l < \infty$, or (ii) $p \to \infty, p/n_l \to 0$. 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 case (i), $\{C_l\}$ are asymptotically free and converge jointly in the sense of (1.2) to, say, $\{c_l\}$. The moments of each c_l depend only on y_l and ρ_l and we give a formula for their free cumulants.

In case (ii), the free cumulants of $\{C_l\}$ vanish in the limit, except those of order one and hence we need appropriate centering and scaling of $\{C_l\}$ to get a non-trivial result. We show that $\{\sqrt{n_l p^{-1}}(C_l - \rho_l I_p)\}$, $1 \leq l \leq t$ converge to free elliptic variables with parameters ρ_l^2. The algebraic convergence results on independent $\sqrt{np^{-1}}(S - I_p))$-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 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 as follows. 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 ESD of A_p. If we take a further expectation with respect to the underlying law of the random variables, then that 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.

Now consider any symmetric matrix polynomial Π in the matrices $\{C_l, C_l^*\}$. Then the moments of the EESD of Π are $p^{-1}\mathbb{E}[\text{Trace}(\Pi^k)]$. Our algebraic convergence results with respect to φ_p imply 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. As a consequence, the EESD of any such 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. It then follows, for example, that the distribution of the singular values of C_l converges almost surely. When the matrix poylnomial is not symmetric, the convergence of the above tracial moments is not sufficient to guarantee the convergence of the EESD or the ESD. However, 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 notions of non-commutative probability that we shall need. For details see Nica and Speicher (2006). Let $(\mathcal{A}, \varphi)$ be a $*$-probability space. Suppose $\{a_i : i \in I\} \subset \mathcal{A}$. Then the numbers $\{\varphi(\Pi(a_i, a_i^* : i \in I)) : \Pi \text{ is a finite degree monomial}\}$ are called their *joint $*$-moments* (we shall refer to them just as moments). If $a \in \mathcal{A}$ is a *self-adjoint* element, that is, $a^* = a$, then μ_a is the *probability law of a* if it is the *unique* law on $\mathbb{R}$ such that

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

If μ_a satisfies (2.1) and is compactly supported, then it is indeed unique.

Let $(\mathcal{A}, \varphi)$ be a $*$-probability space. Define *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 $\varphi(V)[a_1, a_2, \ldots, a_n] := \varphi_s(a_{i_1}, a_{i_2}, \ldots, a_{i_s}) = \varphi(a_{i_1} a_{i_2} \cdots a_{i_s})$ if $V = \{i_1, i_2, \ldots, i_s\}$ with $i_1 < i_2 < \cdots < i_s$. Note that the order of the variables has 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) and the the Möbius function μ, the following relations can be shown:

$$\kappa_\pi[a_1, a_2, ..., a_n] = \sum_{\sigma \in NC(n):\ \sigma \leq \pi} \varphi_\sigma[a_1, a_2, \ldots, a_n]\mu(\sigma, \pi), \pi \in NC(n),\ n \geq 1 \tag{2.6}$$

$$\begin{aligned}\kappa_{\mathbf{1}_n}[a_1, a_2, \ldots, a_n] &= \kappa_n(a_1, a_2, \ldots, a_n), \\ \kappa_2(a_1, a_2) &= \varphi(a_1 a_2) - \varphi(a_1)\varphi(a_2), \\ \varphi(a_1 a_2 \cdots a_n) &= \sum_{\sigma \in NC(n):\ \sigma \leq \mathbf{1}_n} \kappa_\sigma[a_1, a_2, \ldots, a_n], \qquad (2.7) \\ \varphi_\pi[a_1, a_2, \ldots, a_n] &= \sum_{\sigma \in NC(n):\ \sigma \leq \pi} \kappa_\sigma[a_1, a_2, \ldots, a_n].\end{aligned}$$

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 $1 \leq l \leq t$. The (i, j)-th entries of X_l and Y_l are denoted by $X_{ij,(p,n_l)}^{(l)}$ and $Y_{ij,(p,n_l)}^{(l)}$ 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 $p \geq 1$ and $n_l \geq 1$, the pairs of random variables $(X^{(l)}_{ij,(p,n_l)}, Y^{(l)}_{ij,(p,n_l)})$ are independent across $1 \leq i \leq p$, $1 \leq j \leq n_l$, $1 \leq l \leq t$.
(b) For all $1 \leq i \leq p$, $1 \leq j \leq n_l$, $1 \leq l \leq t$, $p, n_l \geq 1$,

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

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

3.1. *Convergence of* $\{C_l : 1 \leq l \leq t\}$ *when* $y_l \neq 0$. 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

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

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

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

where

$$S(\boldsymbol{\eta}_k) := S(\eta_1, \ldots, \eta_k) := \sum_{1 \leq u \leq k, \ \eta_{k+1} = \eta_1} \delta_{\eta_u \eta_{u+1}} \ \text{ and } \ \eta_1, \eta_2, \ldots, \eta_k \in \{1, *\}. \tag{3.1}$$

It is interesting to consider 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 $\rho = 0$, only *alternating free cumulants* of c survive and hence c is a (tracial) *R-diagonal* element. See Nica and Speicher (2006) for the definition and properties of such elements. These free cumulants are given by

$$\kappa_{2k}(c, c^*, c, c^* \ldots, c^*) = \kappa_{2k}(c^*, c, c^*, c, \ldots, c) = y^{2k-1}, \ \forall\ k \geq 1. \tag{3.2}$$

The variable c can be linked 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,

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

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

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

Remark 3.3. 1. Since $\{c_l\}$ are free in Theorem 3.2, their joint free cumulants can, in principle, be written down using the marginal free cumulants.
2. Capitaine and Casalis (2004) proved the asymptotic freeness of the sample covariance matrices for the case $y_l \neq 0, \ \forall\, 1 \le l \le t$. This result follows from Theorem 3.2(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.
3. One can impose different patterns on the correlations and then of course the limit will depend on the patterns. In particular, for the Toeplitz pattern $\mathbb{E}(X_{ij}^{(l)} Y_{ij}^{(l)}) = \rho_{|i-j|}^{(l)}$ suggested by a Referee, it can be anticipated that the convergence would hold. However, substantial amount of calculations would be needed to prove the convergences and identify the limit variables.

We shall use the following notation:

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

For two random variables X and Y, $X \stackrel{\mathcal{D}}{=} Y$ will mean that they have identical probability laws. Henceforth, when we are dealing with only one sequence of matrices, we shall drop the index l and when we are working with any two indices, then without loss, they are taken to be 1 and 2.

Example 3.4. 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 $c_1 c_2^*$ can be calculated using this fact as follows. First note that all its free cumulants, except the even order alternating free cumulants, are 0. These alternating free cumulants are given by

$$\begin{aligned} & \kappa_{2k}(c_1c_2^*, c_2c_1^*, \ldots, c_1c_2^*, c_2c_1^*) = \kappa_{2k}(c_2c_1^*, \ldots, c_1c_2^*, c_2c_1^*, c_1c_2^*) \\ = & \sum_{\substack{\pi, \sigma \in \text{NC}(k) \\ \sigma \le K(\pi)}} \left(\prod_{V \in \pi} y_1^{2(\#V)-1} \right) \left(\prod_{W \in \sigma} y_2^{2(\#W)-1} \right), \\ & \qquad\qquad \text{(by (3.2) and Exercise 15.24(2) of Nica and Speicher (2006))} \\ = & \; y_1^k y_2^k \sum_{\pi \in \text{NC}(k)} y_1^{k-|\pi|} \sum_{\sigma \le K(\pi)} y_2^{k-|\sigma|} \\ = & \sum_{\pi \in \text{NC}(k)} \kappa_\pi[y_1 M_{y_1}, \ldots, y_1 M_{y_1}] \varphi_{K(\pi)}[y_2 M_{y_2}, \ldots, y_2 M_{y_2}] \;\; (M_{y_1} \text{ and } M_{y_2} \text{ are free}) \\ = & \; \varphi((y_1 M_{y_1} y_2 M_{y_2})^k) = \varphi(((\sqrt{y_1 M_{y_1}} y_2 M_{y_2} \sqrt{y_1 M_{y_1}})^{1/2})^{2k}). \end{aligned} \tag{3.3}$$

Thus for $\rho = 0$, the $*$-distributions of $c_1 c_2^*$ and $u\mathrm{P}(1, (\sqrt{y_1 M_{y_1}} y_2 M_{y_2} \sqrt{y_1 M_{y_1}})^{1/2})$ are identical where u is Haar unitary and is free of $\mathrm{P}(1, (\sqrt{y_1 M_{y_1}} y_2 M_{y_2} \sqrt{y_1 M_{y_1}})^{1/2})$. Also note that yM_y is itself a compound free Poisson variable $\mathrm{P}(y^{-1}, \delta_{\{y^2\}})$.

Example 3.5. By Theorem 3.2(b), the ESD of $C + C^*$ converges weakly almost surely to the law whose free cummulants are given by

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

where $S(\boldsymbol{\eta}_k)$ is as defined in (3.1). If $\rho = 0$, then 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 when $\rho = 0$, 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 3.6. By Theorem 3.2(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 (3.2), we have

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

the k-th moment of yM_y. So, the LSD of $n^{-2}XY^*YX^*$ is the law of $\mathrm{P}(1, yM_y)$.

Example 3.7. By Theorem 3.2(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 3.2(a) cannot be further simplified. However for $\rho = 0$, recall that $c_1c_2^*$ is tracial R-diagonal and (3.3) holds. Therefore,

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

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 3.2: (a) 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 in this special case, that the same proof works for the general case. Outline of the modifications needed are given in Section 4.2. 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, X and Y are involved. To keep track of this, let

$$(\epsilon_{2s-1}, \epsilon_{2s}) = \begin{cases} (1,2) & \text{if } \eta_s = 1, \\ (2,1) & \text{if } \eta_s = *. \end{cases} \tag{3.4}$$

Observe that ϵ_{2s-1} can never be equal to ϵ_{2s} and hence

$$\delta_{\epsilon_{2s-1}\epsilon_{2s}} = 0 \text{ for all } s.$$

Also note that, for any $s \geq 1$,

$$\begin{aligned} \eta_s = 1 \;&\Leftrightarrow\; \epsilon_{2s-1} = 1 \;\Leftrightarrow\; \epsilon_{2s} = 2, \\ \eta_s = * \;&\Leftrightarrow\; \epsilon_{2s-1} = 2 \;\Leftrightarrow\; \epsilon_{2s} = 1. \end{aligned}$$

Define

$$A_l^{(\epsilon_{2s-1})} = \begin{cases} X_l & \text{if } \epsilon_{2s-1} = \eta_s = 1, \\ Y_l & \text{if } \epsilon_{2s-1} = 2 \text{ (or } \eta_s = *), \end{cases} \tag{3.5}$$

$$A_l^{(\epsilon_{2s})} = \begin{cases} X^*_l & \text{if } \epsilon_{2s} = 1 \text{ (or } \eta_s = *), \\ Y^*_l & \text{if } \epsilon_{2s} = 2 \text{ (or } \eta_s = 1). \end{cases}$$

Hence,

$$\begin{aligned} \eta_s = 1 \;\Leftrightarrow\; A_l^{(\epsilon_{2s-1})}A_l^{(\epsilon_{2s})} &= A_l^{(1)}A_l^{(2)} = X_lY_l^*, \\ \eta_s = * \;\Leftrightarrow\; A_l^{(\epsilon_{2s-1})}A_l^{(\epsilon_{2s})} &= A_l^{(2)}A_l^{(1)} = Y_lX_l^*. \end{aligned}$$

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 $\eta_s \in \{1, *\}$,

$$L_p := p^{-1}n^{-k}\mathbb{E}\mathrm{Tr}(A_{\beta_1}^{(\epsilon_1)}A_{\beta_2}^{(\epsilon_2)} \cdots A_{\beta_{2k}}^{(\epsilon_{2k})}) \tag{3.6}$$

converges to the appropriate limit. Upon expansion,

$$L_p = \frac{1}{pn^k}\sum_{I_{2k}} \mathbb{E} \prod_{1 \leq s \leq 2k, \; i_{2k+1} = i_1} A_{\beta_s}^{(\epsilon_s)}(i_s, i_{s+1}), \tag{3.7}$$

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

$$I_{2k} = \{(i_1, i_2, \ldots, i_{2k}): \; 1 \leq i_{2s-1} \leq p, \; 1 \leq i_{2s} \leq n, \; 1 \leq s \leq 2k\}. \tag{3.8}$$

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

Since 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, we split up the sum into indices that match. Consider any connected bipartite graph between the *distinct* odd and even indices, $I = \{i_{2s-1} : 1 \leq s \leq k\}$ and $J = \{i_{2s} : 1 \leq s \leq k\}$. Then we need to consider only those cases where each edge appears at least twice. There are at most k such distinct edges and since the graph is connected,

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

By Assumption I, there is a common bound for all expectations involved. Hence, noting also that $\frac{p}{n} \to y > 0$, the total expectation of the terms involved in this graph is of the order

$$O(\frac{p^{\#I}n^{\#J}}{pn^k}) = O(p^{\#I+\#J-(k+1)}). \tag{3.9}$$

Moreover, due to the moment assumption,

$$\frac{1}{pn^k}\mathbb{E}\mathrm{Tr}\left(\prod_{s=1}^{2k} A_{\beta_s}^{(\epsilon_s)}\right) = O(1) \text{ for all } 1 \le \beta_s \le t,\ \epsilon_s \in \{1,2\},\ 1 \le s \le 2k,\ k \ge 1. \tag{3.10}$$

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

$$P_2(2k) = \{\pi : \pi \text{ is a pair-partition of } \{1, 2, \ldots, 2k\}\}. \tag{3.11}$$

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

$$a_{r,s} = \begin{cases} 1 & \text{if } r, s \text{ are both odd or both even,} \\ 0 & \text{otherwise.} \end{cases} \tag{3.12}$$

Then we have

$$\lim_{p\to\infty} L_p = \lim_{p\to\infty} \frac{1}{pn^k} \sum_{I_{2k}} \mathbb{E}\big[\prod_{s=1}^{2k} A_{\beta_s}^{(\epsilon_s)}(i_s, i_{s+1})\big] = \sum_{\pi\in P_2(2k)} \lim_{p\to\infty} \frac{1}{pn^k} \sum_{I_{2k}} \prod_{(r,s)\in\pi} E(r,s) \text{ say,} \tag{3.13}$$

where, suppressing the dependence on other variables,

$$\begin{aligned} E(r,s) &= \mathbb{E}\big[A_{\beta_r}^{(\epsilon_r)}(i_r, i_{r+1}) A_{\beta_s}^{(\epsilon_s)}(i_s, i_{s+1})\big] \\ &= \delta_{\beta_r\beta_s}(\rho(1-\delta_{\epsilon_r\epsilon_s}) + \delta_{\epsilon_r\epsilon_s})(\delta_{i_r i_s}\delta_{i_{r+1}i_{s+1}} a(r,s) + \delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}(1 - a(r,s))) \\ &= \delta_{\beta_r\beta_s}(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\epsilon_r\epsilon_s})(\delta_{i_r i_s}\delta_{i_{r+1}i_{s+1}} a(r,s) + \delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}(1 - a(r,s))). \end{aligned} \tag{3.14}$$

Since $|\rho| \le 1$, each $E(r,s)$ is a sum of two factors—one factor is bounded by $\delta_{i_r i_s}\delta_{i_{r+1}i_{s+1}}$ and the other is bounded by $\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}$. When we expand $\prod_{(r,s)\in\pi} E(r,s)$, each term involves a product of these δ-values. For any of these terms to contribute, all the corresponding δ must equal 1. Now, using arguments similar to those in the proof of Theorem 3.2.6 in Bose (2018), it follows that except for the term described below, for any other term, the number of restrictions on the indices so that all the δ-values are 1 is such that the total number of choices for the indices is of a lower order than the denominator, and hence these terms do not contribute in the limit. Indeed the only term that survives in $\prod_{(r,s)\in\pi} E(r,s)$ in the limit 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

$$\sum_{\pi\in\mathrm{NC}_2(2k)} \lim_{n\to\infty} \frac{1}{pn^k} \sum_{I_{2k}} \prod_{(r,s)\in\pi} \delta_{\beta_r\beta_s}\big(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\epsilon_r\epsilon_s}\big)(1 - a(r,s)) \prod_{r=1}^{2k} \delta_{i_r i_{\gamma\pi(r)}}, \tag{3.15}$$

where γ denotes the cyclic permutation $1 \to 2 \to \cdots \to 2k \to 1$. But if $\pi \in \mathrm{NC}_2(2k)$ and (r,s) is a block of π, then r and s necessarily have opposite parities and hence $a(r,s) = 0$. Therefore (see Section 4.1 for a detailed proof of (3.16)),

$$\lim L_p = \sum_{\pi\in\mathrm{NC}_2(2k)} \lim_{n\to\infty} \frac{1}{pn^k} \sum_{I_{2k}} \prod_{(r,s)\in\pi} \delta_{\beta_r\beta_s}\big(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\epsilon_r\epsilon_s}\big) \prod_{r=1}^{2k} \delta_{i_r i_{\gamma\pi(r)}}. \tag{3.16}$$

The right side of above simplifies to

$$\sum_{\pi\in \mathrm{NC}_2(2k)} \prod_{(r,s)\in\pi} \delta_{\beta_r\beta_s}(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_r\epsilon_s}) \lim_{n\to\infty} \frac{\#\{I_{2k}: i_r = i_{\gamma\pi(r)} \text{ for all } r\}}{pn^k}. \tag{3.17}$$

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

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

For any two sequences of real numbers $\{a_n\}$ and $\{b_n\}$, we use $a_n = \Omega(b_n)$ to indicate $b_n^{-1}a_n \to 1$ as $n\to\infty$. 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

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

and hence using (3.17),

$$\begin{aligned}
\lim L_p &= \sum_{\pi\in \mathrm{NC}_2(2k)} \prod_{(r,s)\in\pi} \delta_{\beta_r\beta_s}(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_r\epsilon_s}) \lim_{n\to\infty} \frac{\#\{I_{2k}: i_r = i_{\gamma\pi(r)}\forall r\}}{pn^k} \\
&= \sum_{m=0}^{k-1} y^m \sum_{\substack{\pi\in \mathrm{NC}_2(2k):\\ S(\gamma\pi)=m+1}} \prod_{(r,s)\in\pi} \delta_{\beta_r\beta_s}(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_r\epsilon_s}) \qquad (3.18)\\
&= \sum_{m=0}^{k-1} y^m \sum_{\substack{\pi\in \mathrm{NC}_2(2k):\\ S(\gamma\pi)=m+1}} \prod_{(r,s)\in\pi} \delta_{\beta_r\beta_s}(\rho^{T(\pi)}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{T(\pi)0}), \qquad (3.19)
\end{aligned}$$

where

$$T(\pi) = \#\{(r,s)\in\pi : \delta_{\epsilon_r\epsilon_s} = 0\} \text{ for any non-crossing pair-partition } \pi. \tag{3.20}$$

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 in the limit NCP, say, $(\mathcal{A},\varphi)$. We still have to identify the limit and prove the asymptotic freeness. For this we need to go from $\mathrm{NC}_2(2k)$ to $\mathrm{NC}(k)$. Define

$$\begin{aligned}
\tilde{J}_i &= \{j\in\{1,2,\ldots,2k\}: \beta_j = i\}, \quad 1\le i\le t, \qquad (3.21)\\
\tilde{B}_k &= \{\pi\in \mathrm{NC}_2(2k): \pi = \cup_{i=1}^t \pi_i,\ \pi_i\in \mathrm{NC}_2(\tilde{J}_i),\ 1\le i\le t\},\\
\tilde{B}_{m,k} &= \{\pi\in\tilde{B}_k: S(\gamma\pi) = m+1\}.
\end{aligned}$$

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

$$\lim L_p = \sum_{m=0}^{k-1} y^m \sum_{\pi\in\tilde{B}_{m,k}} (\rho^{T(\pi)}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{T(\pi)0}). \tag{3.22}$$

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

$$\tilde{T}(S) = \sum_{1 \leq u \leq r,\ j_{r+1}=j_1} \delta_{\eta_{j_u}\eta_{j_{u+1}}}, \quad \mathcal{T}(\pi) = \sum_{i=1}^{m} \tilde{T}(V_i).$$

Consider the function $f : \mathrm{NC}_2(2k) \to \mathrm{NC}(k)$ as follows. Take $\pi \in \mathrm{NC}_2(2k)$. Suppose (r, s) is a block of π. Let $\lceil \cdot \rceil$ be the ceiling function. Then $\lceil r/2 \rceil$ and $\lceil s/2 \rceil$ are put in the same block in $f(\pi) \in \mathrm{NC}(k)$. Let $1 \leq j_1 < j_2 < \cdots < j_r \leq k$ and $\pi \in \mathrm{NC}_2(2k)$. Then

$$(2j_1 - 1, 2j_r), (2j_1, 2j_2 - 1), (2j_2, 2j_3 - 1), \ldots, (2j_{r-1}, 2j_r - 1) \in \pi \Leftrightarrow (j_1, j_2, \ldots, j_r) \in f(\pi). \quad (3.23)$$

Therefore, each $\pi \in \mathrm{NC}_2(2k)$ has a unique $f(\pi) \in \mathrm{NC}(k)$ and each $\sigma \in \mathrm{NC}(k)$ has a unique $\pi \in \mathrm{NC}_2(2k)$, $\sigma = f(\pi)$. Thus f is a bijection between $\mathrm{NC}_2(2k)$ and $\mathrm{NC}(k)$. Further,

$$\begin{aligned} & (2j_1 - 1, 2k), (2j_1, 2j_2 - 1), (2j_2, 2j_3 - 1), \ldots, (2j_{r-1}, 2k - 1) \in \pi \\ \Leftrightarrow\ & (2j_1, 2j_2, \ldots, 2j_{r-1}, 2k) \in \gamma\pi \Leftrightarrow (j_1, j_2, \ldots, j_{r-1}, k) \in f(\pi). \end{aligned} \quad (3.24)$$

Therefore, clearly, for any $\pi \in \mathrm{NC}_2(2k)$, we have

$$\begin{aligned} & \text{Number blocks in } \pi \text{ which starts with odd elements} \\ = \ & \text{Number of blocks in } \gamma\pi \text{ which contains only even elements} \\ = \ & \text{Number of blocks in } f(\pi). \end{aligned} \quad (3.25)$$

As $\gamma\pi$ contains $(k+1)$ blocks and each block has either all odd or all even elements, by (3.25),

$$\begin{aligned} & S(\gamma\pi) = m + 1 \\ \Leftrightarrow\ & \text{Number of blocks in } \gamma\pi \text{ which contains only even elements} = (k+1) - (m+1) = k - m \\ \Leftrightarrow\ & \text{Number of blocks in } f(\pi) = k - m. \end{aligned}$$

This shows that f is also a bijection between $\tilde{B}_{m,k}$ and $B_{k-m,k}$.

By (3.4), we also have

$$\epsilon_r \neq \epsilon_s \Leftrightarrow \begin{cases} (\epsilon_r, \epsilon_{r+1}) = (\epsilon_{s-1}, \epsilon_s) \Leftrightarrow \eta_{\lceil r/2 \rceil} = \eta_{\lceil s/2 \rceil} & \text{if } r \text{ is odd, } s \text{ is even,} \\ (\epsilon_{r-1}, \epsilon_r) = (\epsilon_s, \epsilon_{s+1}) \Leftrightarrow \eta_{\lceil r/2 \rceil} = \eta_{\lceil s/2 \rceil} & \text{if } r \text{ is even, } s \text{ is odd,} \\ (\epsilon_r, \epsilon_{r+1}) \neq (\epsilon_s, \epsilon_{s+1}) \Leftrightarrow \eta_{\lceil r/2 \rceil} \neq \eta_{\lceil s/2 \rceil} & \text{if } r \text{ and } s \text{ are both odd,} \\ (\epsilon_{r-1}, \epsilon_r) \neq (\epsilon_{s-1}, \epsilon_s) \Leftrightarrow \eta_{\lceil r/2 \rceil} \neq \eta_{\lceil s/2 \rceil} & \text{if } r \text{ and } s \text{ are both even.} \end{cases} \quad (3.26)$$

Now, using the first two lines of (3.26), it is immediate that

$$\delta_{\epsilon_{2u-1}, \epsilon_{2v}} = 0 \Leftrightarrow \delta_{\eta_u \eta_v} = 1 \ \forall\ 1 \leq u \leq v < k,$$
$$\delta_{\epsilon_{2u}, \epsilon_{2v+1}} = 0 \Leftrightarrow \delta_{\eta_u \eta_{v+1}} = 1 \ \forall\ 1 \leq u \leq v < k.$$

Therefore, $T(\pi) = \mathcal{T}(f(\pi)) \ \forall\ \pi \in \tilde{B}_{m,k}$ 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, going back to the proof, by (3.22), we have

$$\begin{aligned}
\lim L_p &= \sum_{m=0}^{k-1} y^m \sum_{\pi \in B_{k-m,k}} \big(\rho^{T(f^{-1}(\pi))}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{T(f^{-1}(\pi))0}\big) \\
&= \sum_{m=0}^{k-1} y^m \sum_{\pi \in B_{k-m,k}} \big(\rho^{\mathcal{T}(\pi)}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\mathcal{T}(\pi)0}\big) \\
&= \sum_{m=0}^{k-1} y^{k-m-1} \sum_{\pi \in B_{m+1,k}} \big(\rho^{\mathcal{T}(\pi)}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\mathcal{T}(\pi)0}\big). \qquad (3.27)
\end{aligned}$$

The last equality is obtained by a change of variable $k-m-1 \to m$. Then, (3.27) implies that

$$\begin{aligned}
\varphi(c_{\alpha_1}^{\eta_1} c_{\alpha_2}^{\eta_2} \cdots c_{\alpha_k}^{\eta_k}) &= \sum_{m=0}^{k-1} \sum_{\substack{\pi \in B_{m+1,k} \\ \pi=\{V_1,V_2,\ldots,V_{m+1}\}}} \prod_{l=1}^{m+1} y^{\#V_l - 1}\big(\rho^{\tilde{T}(V_l)}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\tilde{T}(V_l)0}\big) \\
&= \sum_{\pi \in B_k} \prod_{l=1}^{\#\pi} y^{\#V_l - 1}(\rho^{\tilde{T}(V_l)}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\tilde{T}(V_l)0}). \qquad (3.28)
\end{aligned}$$

Hence, by moment-free cumulant relation, we have

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

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

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

This completes the proof of Theorem 3.2 (a) for the state φ_p for the special case when the values of ρ_l and of y_l are same. On a careful scrutiny, it should be clear that the above proof, with some added complexity in notations, continues to hold when the values of ρ_l and of y_l are not necessarily equal. An outline proof is given in Section 4.2.

It remains to show the algebraic convergence with respect to the state $\tilde{\varphi}_p$. Suppose Π is any arbitrary polynomial. We have already shown that $p^{-1}\mathbb{E}[\text{Trace}(\Pi)] \to \beta$ (say). It is now enough to show that $p^{-1}\text{Trace}(\Pi) \to \beta$ almost surely. In Section 4.3 we will show that

$$\mathbb{E}\Big[p^{-1}\text{Trace}(\Pi) - p^{-1}E[\text{Trace}(\Pi)]\Big]^4 = O(p^{-2}). \qquad (3.29)$$

Then an application of the Borel-Cantelli Lemma would finish the proof of Theorem 3.2 (a).
(b) For this part, we can now use Lemma 1.2.4 in Bose (2018) quoted below. Let $\{A_p\}$ be a sequence of real symmetric random matrices of order p and consider the following conditions:
(M1) $p^{-1}\mathbb{E}\text{Tr}(A_p^k) \to \beta_k$ for all $k \ge 1$.

(M4) $\mathbb{E}(p^{-1}\text{Tr}(A_p^k) - \mathbb{E}(p^{-1}\text{Tr}(A_p^k)))^4 = O(p^{-2})$ for all $k \geq 1$.

(C) $\sum_{k=1}^{\infty} \beta_{2k}^{-\frac{1}{2k}} = \infty$. Then we have the following lemma.

Lemma 3.8. *If (M1), (M4) and (C) hold, then the ESD of A_p converges weakly almost surely to a unique probability distribution whose moment sequence is determined by $\{\beta_k \;\; k \geq 1\}$.*

Now suppose that Π is symmetric in $\{C_l, C_l^* : \; 1 \leq l \leq t\}$. Then by Theorem 3.2(a), and (3.29), the conditions (M1) and (M4) are satisfied. Suppose Π is of degree r. Then from (3.28),

$$\lim p^{-1}\mathbb{E}\text{Tr}(\Pi^k) \leq \#\text{NC}_2(2rk) \leq 2^{2rk} \text{ for all } k \geq 1. \tag{3.30}$$

Hence (C) is also satisfied. Thus Lemma 3.8 implies that the LSD of any symmetric polynomial Π in $\{C_l, C_l^* : \; 1 \leq l \leq t\}$ exists (almost surely). Moreover, by (3.30), there is a $C > 0$, depending on the degree of Π, such that the limiting k-th moment is bounded by C^k for all k. Then by Problem 9, Section 4, Chapter 2, Page 95 in Ash and Doléans-Dade (2000), the support of the LSD of Π is contained in $[-C, \; C]$. This completes the proof of Theorem 3.2(b). □

Remark 3.9. Lemma 3.8 is applicable only when Π is a symmetric matrix. To the best of our knowledge, no general results are known on the existence of the LSD of Π when it is non-symmetric. Establishing the existence of the LSD for non-symmetric matrices is invariably challenging. For example, see Akemann et al. (2021) which deals with the LSD of the non-symmetric matrix C_1 when the entries are complex Gaussian and $y > 0$. For more examples of other non-symmetric matrices for which the LSD is known to exist, see Bordenave and Chafaï (2012), O'Rourke et al. (2015), Bose and Hachem (2020) and Bose and Hachem (2023+).

3.2. *Convergence of $\{C_l : 1 \leq l \leq t\}$ when $y = 0$.* Now suppose $pn^{-1} \to y = 0$ as $n, p \to \infty$. Then the limit in Theorem 3.2 is degenerate. For a non-degenerate limit,, we now need a centering as well as a different scaling. Towards this, 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, $\mathbb{E}(X_{11}) = 0$, and $\text{Var}(X_{11}) = 1, \mathbb{E}(X_{11}^4) < \infty$. Then the ESD of $\sqrt{np^{-1}}(S - I_p)$ converges weakly almost surely to the standard semi-circle law. Here I_p is the identity matrix of order p. For a proof, see Bose (2018). This proof actually shows that if Assumption I (c) also holds, 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 extend this algebraic result to several cross-covariance matrices.

Before we state the result, we need to recall the definition of elliptic variables.

Definition 3.10. Suppose $(\mathcal{A}, \varphi)$ is an NCP. An element $e \in \mathcal{A}$ is said to be a (standard) *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 criterion 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 $1 \le i \le k$, $\epsilon_i \in \{1, *\}$ and $\tau_i \in \{1, \ldots, t\}$, the following holds for the joint moments:

$$\varphi(e_{\tau_1}^{\epsilon_1} e_{\tau_2}^{\epsilon_2} \cdots e_{\tau_{2k}}^{\epsilon_{2k}}) = \sum_{\pi \in \mathrm{NC}_2(2k)} \rho_1^{T_1(\pi)} \ldots \rho_m^{T_m(\pi)} \prod_{(r,s) \in \pi} \delta_{\tau_r \tau_s}. \tag{3.31}$$

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 3.11. *Suppose Assumption I holds with $y_l = 0$, for all $1 \le l \le t$.*
(a) Then $E_l = \sqrt{n_l p^{-1}}(C_l - \rho_l I_p), 1 \le l \le t$ as elements of $(\mathcal{M}_p, \varphi_p)$ converge jointly to free elliptic variables $\{e_l\}, 1 \le l \le t$, with parameters $\{\rho_l^2\}, 1 \le l \le t$. The convergence also holds with respect to the state $\tilde{\varphi}_p$ almost surely. The limiting state is tracial.
(b) Let $\Pi := \Pi(\{E_l : 1 \le l \le t\})$ be any finite degree real matrix polynomial in $\{E_l, E_l^* : 1 \le l \le t\}$ and which is a symmetric matrix. Then the ESD of Π converges weakly almost surely to the compactly supported probability law of the self-adjoint variable $\Pi(\{e_l : 1 \le l \le t\})$.

Example 3.12. Theorem 3.11 is useful to find the LSD of any appropriately centered and scaled symmetric polynomial of $\{C_l : 1 \le l \le t\}$. For example, consider

$$\begin{aligned}
\Pi &= \sqrt{\frac{\min(n_1, n_2)}{p}}(C_1 + C_1^* + C_1 C_2^* + C_2 C_1^* - 2\rho_1(1+\rho_2)I_p) \\
&= \sqrt{\frac{\min(n_1, n_2)}{p}}\Big[(C_1 + C_1^* - 2\rho_1 I_p) + (C_1 - \rho_1 I_p)(C_2 - \rho_2 I_p)^* + (C_2 - \rho_2 I_p)(C_1 - \rho_1 I_p)^* \\
&\qquad + \rho_2(C_1 + C_1^* - 2\rho_1 I_p) + \rho_1(C_2 + C_2^* - 2\rho_2 I_p)\Big] \\
&= \sqrt{\min(n_1, n_2)}\Big[n_1^{-1/2}(E_1 + E_1^*) + \sqrt{p} n_1^{-1/2} n_2^{-1/2}(E_1 E_2^* + E_2 E_1^*) \\
&\qquad + n_1^{-1/2}\rho_2(E_1 + E_1^*) + n_2^{-1/2}\rho_1(E_2 + E_2^*)\Big].
\end{aligned} \tag{3.32}$$

Hence it follows that

$$\Pi \xrightarrow{*} \begin{cases} (1+\rho_2)(e_1 + e_1^*) + y_{12}^{1/2}(e_2 + e_2^*) & \text{if} \quad \lim_{p\to\infty} n_1/n_2 = y_{12} \le 1, \\ y_{12}^{-1/2}(1+\rho_2)(e_1 + e_1^*) + (e_2 + e_2^*) & \text{if} \quad \lim_{p\to\infty} n_1/n_2 = y_{12} \ge 1. \end{cases} \tag{3.33}$$

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 (3.33).

Proof of Theorem 3.11: (a) As before, we shall give the detailed proof only for the special case where all the n_l's and the ρ_l's are equal to say n and ρ respectively. For any $k \ge 1$ and $\epsilon_1, ..., \epsilon_k \in \{1, *\}$, we will need to consider the limit of the following as $p, n \to \infty$ with $p/n \to 0$:

$$\frac{1}{p}\mathbb{E}\mathrm{Tr}(E_{l_1}^{\epsilon_1} \cdots E_{l_k}^{\epsilon_k}) = \frac{1}{p(np)^{k/2}} \sum_{\substack{1 \le i_1, \ldots, i_k \le p \\ 1 \le j_1, \ldots, j_k \le n}} \mathbb{E}\Big[\prod_{t=1}^{k} \big(a_{l_t i_t j_t} a_{l_t i_{t+1} j_t} - \rho \delta_{i_t i_{t+1}}\big)\Big], \tag{3.34}$$

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

$$(a_{l_r i_r j_r}, a_{l_r i_{r+1} j_r}) = \begin{cases} (x^{(l_r)}_{i_r j_r}, y^{(l_r)}_{i_{r+1} j_r}) & \text{if } \epsilon_r = 1, \\ \\ (y^{(l_r)}_{i_r j_r}, x^{(l_r)}_{i_{r+1} j_r}) & \text{if } \epsilon_r = *. \end{cases} \tag{3.35}$$

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

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

(i) If there is a pair $(i_r, j_r) \in P$ that appears only once, then $(i_r, j_r) \neq (i_{r+1}, j_r)$ and hence $i_r \neq i_{r+1}$. Hence $a_{l_r i_r j_r}$ is independent of all other variables. This implies that

$$\mathbb{E}\Big[\prod_{t=1}^{k}\big(a_{l_t i_t j_t} a_{l_t i_{t+1} j_t} - \rho\delta_{i_t i_{t+1}}\big)\Big] \;=\; \mathbb{E}[a_{l_r i_r j_r}]\mathbb{E}\Big[a_{l_r i_{r+1} j_r}\prod_{t\neq r}\big(a_{l_t i_t j_t} a_{l_t i_{t+1} j_t} - \rho\delta_{i_t i_{t+1}}\big)\Big] = 0. \tag{3.36}$$

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 for convenience, we continue to denote 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. Then,

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

Hence we can restrict attention to the further reduced subset of P (we continue to denote it by 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 call these 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, $\#E \le k$. Since every j-index was originally matched, $\#V_J \le k/2$. We also know from the connectedness property that

$$\#V_I + \#V_J \le \#E + 1 \le k + 1. \tag{3.37}$$

Hence the contribution to (3.34) is bounded above by

$$O\left(\frac{p^{\#V_I} n^{\#V_J}}{p^{k/2+1} n^{k/2}}\right) = O\left(\frac{p^{k+1-\#V_J} n^{\#V_J}}{p^{k/2+1} n^{k/2}}\right) = O\left(\left(\frac{p}{n}\right)^{k/2-\#V_J}\right). \tag{3.38}$$

Therefore, for all $l_i \in \{1, 2 \ldots, t\}$, $\epsilon_i \in \{1, *\}$ and $1 \le i \le k$, we have

$$p^{-1}\mathbb{E}\mathrm{Tr}(E^{\epsilon_1}_{l_1} \cdots E^{\epsilon_k}_{l_k}) = O(1). \tag{3.39}$$

If $\#V_J < \dfrac{k}{2}$, then the above expression goes to 0. So the only possible non-zero contribution to the limit of (3.34) will come when $\#V_J = k/2$. This immediately implies that if k is odd, then there can be no such contribution, and hence

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

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

$$O\left(\frac{p^{\#V_I} n^{\#V_J}}{p^{k/2+1} n^{k/2}}\right) = O(p^{\#V_I-(k/2+1)}) = O(p^{\#V_I-(m+1)}). \tag{3.40}$$

On the other hand, from (3.37), 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

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

Hence, such a combination cannot contribute to (3.34). 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 denote this reduced subset by P. As a consequence, (3.34) reduces to

$$p^{-1}\mathbb{E}\mathrm{Tr}(E_{l_1}^{\epsilon_1} \cdots E_{l_k}^{\epsilon_k}) \approx p^{-(m+1)} n^{-m} \sum_{P} \mathbb{E}\big(\prod_{r=1}^{2m} a_{l_r i_r j_r} a_{l_r i_{r+1} j_r}\big). \tag{3.41}$$

As a consequence of the above discussions, we have two situations: (i) $(i_r, j_r) = (i_s, j_s)$ 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 $i_{r+1} = i_{s+1}$. (ii) if $(i_r, j_r) = (i_{s+1}, j_s)$ then $i_{r+1} = i_s$. 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 (3.35), the moment structure, and the above developments, it is easy to verify that the possibly contributing part of the above sum (and hence of (3.34)) can be re-expressed as

$$\sum_{\substack{i_1,\ldots,i_{2m}\\ j_1,\ldots,j_{2m}}} \sum_{\pi\in P_2(2m)} \frac{1}{p^{m+1}n^m} \prod_{(r,s)\in\pi} \Big[\delta_{l_r l_s}\delta_{\epsilon_r\epsilon_s}(\rho^2\delta_{i_r i_{s+1}}\delta_{j_r j_s}\delta_{i_{r+1} i_s} + \delta_{i_r i_s}\delta_{j_r j_s}\delta_{i_{r+1} i_{s+1}})$$
$$+\delta_{l_r l_s}(1-\delta_{\epsilon_r\epsilon_s})(\rho^2\delta_{i_r i_s}\delta_{j_r j_s}\delta_{i_{r+1} i_{s+1}} + \delta_{j_r j_s}\delta_{i_r i_{s+1}}\delta_{i_{r+1} i_s})\Big] \tag{3.42}$$

which equals

$$\sum_{\substack{i_1,\ldots,i_{2m}\\ j_1,\ldots,j_{2m}}} \sum_{\pi\in P_2(2m)} \frac{1}{p^{m+1}n^m} \prod_{(r,s)\in\pi} \delta_{l_r l_s}\delta_{j_r j_s}\Big[\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}\big(\rho^2\delta_{\epsilon_r\epsilon_s} + (1-\delta_{\epsilon_r\epsilon_s})\big)$$
$$+\delta_{i_{r+1} i_{s+1}}\delta_{i_s i_r}\big(\delta_{\epsilon_r\epsilon_s} + \rho^2(1-\delta_{\epsilon_r\epsilon_s})\big)\Big]. \tag{3.43}$$

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 (see Theorem 3.3.1 in Bose (2018)). This immediately implies that the limit of (3.34) and hence of (3.42) and (3.43) in this special case equals $\#\mathrm{NC}_2(2m) = C_m$, the m-th Catalan number. This means that

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

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

$$\sum_{i_1,\ldots,i_{2m},j_1,\ldots,j_{2m}} \frac{1}{p^{m+1}n^m} \prod_{(r,s)\in\pi} \delta_{j_r j_s}\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}} = 1. \tag{3.44}$$

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

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

So $i_r=i_{\gamma\pi(r)}$ for each r if the above product is 1. But this means that the i-index is constant on each block of $\gamma\pi$. As $\pi\in\mathrm{NC}_2(2m)$, $|\gamma\pi|=m+1$ and thus there are $\Omega(p^{m+1})$ choices in total for the i's. This establishes (3.44).

These arguments establish that

$$\lim_{n\to\infty}\sum_{\pi\in P_2(2m)}\sum_{\substack{i_1,\dots,i_{2m}\\ j_1,\dots,j_{2m}}}\frac{1}{p^{m+1}n^m}\prod_{(r,s)\in\pi}\delta_{j_r j_s}(\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}+\delta_{i_r i_s}\delta_{i_{r+1}i_{s+1}}) \tag{3.45}$$

$$=\lim\sum_{\pi\in\mathrm{NC}_2(2m)}\frac{1}{p^{m+1}n^m}\sum_{\substack{i_1,\dots,i_{2m}\\ j_1,\dots,j_{2m}}}\prod_{(r,s)\in\pi}\delta_{j_r j_s}\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}. \tag{3.46}$$

This implies that the rest of the terms in (3.45) must go to 0, since these quantities are all non-negative. Therefore,

$$\sum_{\pi\in P_2(2m)}\sum_{\substack{i_1,\dots,i_{2m}\\ j_1,\dots,j_{2m}}}\frac{1}{p^{m+1}n^m}\prod_{(r,s)\in\pi}\delta_{j_r j_s}\delta_{i_{r+1}i_{s+1}}\delta_{i_s i_r}\to 0, \tag{3.47}$$

$$\sum_{\pi\in P_2(2m)\setminus\mathrm{NC}_2(2m)}\sum_{\substack{i_1,\dots,i_{2m}\\ j_1,\dots,j_{2m}}}\frac{1}{p^{m+1}n^m}\prod_{(r,s)\in\pi}\delta_{j_{r+1}j_s}\delta_{i_r i_{s+1}}\delta_{i_s i_r}\to 0. \tag{3.48}$$

Now we consider the general case where $t\ge 1$ but for the moment assume that ρ_l's are equal but the common value is not necessarily equal to 1. Going back to (3.43), we have

$$\begin{aligned}
&\sum_{\pi\in P_2(2m)}\sum_{\substack{i_1,\dots,i_{2m}\\ j_1,\dots,j_{2m}}}\frac{1}{p^{m+1}n^m}\prod_{(r,s)\in\pi}\delta_{l_r l_s}\delta_{j_r j_s}\delta_{i_{r+1}i_{s+1}}\delta_{i_s i_r}\big(\delta_{\epsilon_r\epsilon_s}+\rho^2(1-\delta_{\epsilon_r\epsilon_s})\big)\\
\le\ &\sum_{\pi\in P_2(2m)}\sum_{\substack{i_1,\dots,i_{2m}\\ j_1,\dots,j_{2m}}}\frac{1}{p^{m+1}n^m}\prod_{(r,s)\in\pi}\delta_{j_r j_s}\delta_{i_{r+1}i_{s+1}}\delta_{i_s i_r}\to 0,\ \text{by (3.47)},\\
&\sum_{\pi\in P_2(2m)\setminus\mathrm{NC}_2(2m)}\sum_{\substack{i_1,\dots,i_{2m}\\ j_1,\dots,j_{2m}}}\frac{1}{p^{m+1}n^m}\prod_{(r,s)\in\pi}\delta_{l_r l_s}\delta_{j_r j_s}\delta_{i_r i_{s+1}}\delta_{i_{s+1}i_r}\big(\rho^2\delta_{\epsilon_r\epsilon_s}+(1-\delta_{\epsilon_r\epsilon_s})\big)\\
\le\ &\sum_{\pi\in P_2(2m)\setminus\mathrm{NC}_2(2m)}\sum_{\substack{i_1,\dots,i_{2m}\\ j_1,\dots,j_{2m}}}\frac{1}{p^{m+1}n^m}\prod_{(r,s)\in\pi}\delta_{j_r j_s}\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}\to 0,\ \text{by (3.48)}.
\end{aligned} \tag{3.49}$$

Thus we conclude that the expression in (3.43) converges to

$$\sum_{\pi\in\mathrm{NC}_2(2m)}\prod_{(r,s)\in\pi}\delta_{l_r l_s}(\rho^2\delta_{\epsilon_r\epsilon_s}+(1-\delta_{\epsilon_r\epsilon_s}))\quad\text{as}\quad n\to\infty. \tag{3.50}$$

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

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

On the other hand, if $e_1, ..., e_{2m}$ are freely elliptic elements each with parameter ρ^2 in some NCP $(\mathcal{A}, \varphi)$, then by (3.31), it is easy to check that the above expression is nothing but $\varphi(e_{l_1}^{\epsilon_1} ... e_{l_{2m}}^{\epsilon_{2m}})$. 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 (see Section 4.4 for more details). This completes the proof of the first part of (a).

As discussed in the proof of Theorem 3.2, convergence with respect to the state $\tilde{\varphi}_p$ follows by the Borel-Cantelli Lemma after it is established that (3.29) holds for any polynomial Π in the matrices $\{E_l\}$. Proof of this proceeds along lines similar to that in the proof of the first part of (a). Some more details are given in Section 4.3. See the proof of Theorem 3.5 in Bhattacharjee and Bose (2016b) for similar arguments in a different context. This complete the proof of (a).

(b) This is also similar to the proof of Theorem 3.2(b). Let the finite degree polynomial Π of $\{E_l, E_l^* : 1 \le l \le t\}$ be symmetric. By Theorem 3.11(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. Hence 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. ☐

4. Additional technical details

4.1. *Proof of (3.16).* Take any $\pi \in P_2(2k)$. It contains at least one block with both odd elements if and only if it contains at least one block with both even elements. We shall first show that any such π does not contribute in the limit. Then we shall show that any $\pi \in P_2(2k)$ which is crossing, also does not do so.

Let $P_2^{\mathrm{odd}}(2k)$ be the subsets of $P_2(2k)$ which has at least one block of all odd elements. Then

$$P_2^{\mathrm{odd}}(2k) = \cup_{u=0}^{k-2} \cup_{v=1}^{k-1} P_2^{(u,v,\mathrm{odd})}(2k),$$

where

$$P_2^{(u,v,\mathrm{odd})}(2k) \quad = \quad \{\pi \in P_2(2k) : \ (2u+1, 2v+1) \in \pi\}. \tag{4.1}$$

Recall the notation $E(r, s)$ from (3.14). Note that when $\pi \in P_2^{(u,v,\mathrm{odd})}(2k)$, we have

$$E(2u+1, 2v+1) = \delta_{\beta_{2u+1}\beta_{2v+1}} (\rho^{1-\epsilon_{2u+1}\epsilon_{2v+1}} (1 - \delta_{\rho 0}) + \delta_{\rho 0} \delta_{\epsilon_{2u+1}\epsilon_{2v+1}}).$$

Hence, we have,

$$\begin{aligned}
&\sum_{\pi \in P_2^{(u,v,\text{odd})}(2k)} \frac{1}{pn^k} \sum_{I_{2k}} \prod_{(r,s)\in\pi} E(r,s) \\
= \;& (\delta_{\beta_{2u+1}\beta_{2v+1}}(\rho^{1-\epsilon_{2u+1}\epsilon_{2v+1}}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\epsilon_{2u+1}\epsilon_{2v+1}}))\delta_{i_{2u+1}i_{2v+1}}\delta_{i_{2u+2}i_{2v+2}} \\
& \qquad \times \frac{1}{pn^k} \sum_{I_{2k}} \mathbb{E}\Big(\prod_{\substack{1\le s\le 2u \\ 2u+2\le s\le 2v \\ 2v+2\le s\le 2k \\ i_{2k+1}=i_1}}^{2u} A_{\beta_s}^{\epsilon_s}(i_s, i_{s+1}) \Big) \\
= \;& \left(\frac{1}{n}\right) \frac{1}{pn^{k-1}} \mathbb{E}\mathrm{Tr}\left(\left(\prod_{s=1}^{2u} A_{\beta_s}^{\epsilon_s} \right) \left(\prod_{s=2u+2}^{2v} A_{\beta_s}^{\epsilon_s} \right)^* \left(\prod_{s=2v+2}^{2k} A_{\beta_s}^{\epsilon_s} \right) \right) \\
& \qquad \times (\delta_{\beta_{2u+1}\beta_{2v+1}}(\rho^{1-\epsilon_{2u+1}\epsilon_{2v+1}}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\epsilon_{2u+1}\epsilon_{2v+1}})) \\
= \;& O(n^{-1}), \quad \text{by (3.10)}. \qquad (4.2)
\end{aligned}$$

Hence,

$$\sum_{\pi \in P_2^{\text{odd}}(2k)} \frac{1}{pn^k} \sum_{I_{2k}} \prod_{(r,s)\in\pi} E(r,s) = o(1), \tag{4.3}$$

which immediately leads to

$$\lim_{p\to\infty} L_p = \sum_{\pi \in P_2(2k)\setminus P_2^{\text{odd}}(2k)} \lim_{p\to\infty} \frac{1}{pn^k} \sum_{I_{2k}} \prod_{(r,s)\in\pi} E(r,s). \tag{4.4}$$

To uncover the contribution of the remaining π's that appear above, we first make a general observation. Suppose $\pi \in P_2(2k) \setminus P_2^{\text{odd}}(2k)$. Then there exists an $(r,s) \in \pi$, where r and s have opposite parity and hence

$$a(r,s) = 0 \quad \text{and} \quad E(r,s) = \delta_{\beta_r\beta_s}(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\epsilon_r\epsilon_s})\delta_{i_r i_{s+1}}\delta_{i_s i_{r+1}}.$$

Now we shall show that no crossing partition $\pi \in P_2(2k) \setminus P_2^{\text{odd}}(2k)$ contributes to (4.4). We split the subset of all crossing pair-partitions of $P_2(2k) \setminus P_2^{\text{odd}}(2k)$ in the following manner. Let

$$\begin{aligned}
P_{2,1}^{u_1,u_2,v_1,v_2}(2k) &= \{\pi \in P_2(2k) : \; (2u_1+1, 2v_1), (2u_2+1, 2v_2) \in \pi, 2u_1+1 < 2u_2+1 < 2v_1 < 2v_2\}, \\
P_{2,1}(2k) &= \Big(\cup_{u_1=0}^{k-3} \cup_{u_2=u_1+1}^{k-2} \cup_{v_1=u_2+1}^{k-1} \cup_{v_2=v_1+1}^{k} P_{2,1}^{u_1,u_2,v_1,v_2}(2k) \Big) \setminus P_2^{\text{odd}}(2k), \\
P_{2,2}^{u_1,u_2,v_1,v_2}(2k) &= \{\pi \in P_2(2k) : \; (2u_1, 2v_1+1), (2u_2+1, 2v_2) \in \pi, 2u_1 < 2u_2+1 < 2v_1+1 < 2v_2\}, \\
P_{2,2}(2k) &= \Big(\cup_{u_1=1}^{k-2} \cup_{u_2=u_1}^{k-2} \cup_{v_1=u_1+1}^{k-1} \cup_{v_2=v_1+1}^{k} P_{2,2}^{u_1,u_2,v_1,v_2}(2k) \Big) \setminus P_2^{\text{odd}}(2k), \\
P_{2,3}^{u_1,u_2,v_1,v_2}(2k) &= \{\pi \in P_2(2k) : \; (2u_1+1, 2v_1), (2u_2, 2v_2+1) \in \pi, 2u_1+1 < 2u_2 < 2v_1 < 2v_2+1\}, \\
P_{2,3}(2k) &= \Big(\cup_{u_1=0}^{k-3} \cup_{u_2=u_1+1}^{k-2} \cup_{v_2=u_2+1}^{k-1} \cup_{v_2=v_1}^{k-1} P_{2,3}^{u_1,u_2,v_1,v_2}(2k) \Big) \setminus P_2^{\text{odd}}(2k), \\
P_{2,4}^{u_1,u_2,v_1,v_2}(2k) &= \{\pi \in P_2(2k) : \; (2u_1, 2v_1+1), (2u_2, 2v_2+1) \in \pi, 2u_1 < 2u_2 < 2v_1+1 < 2v_2+1\}, \\
P_{2,4}(2k) &= \Big(\cup_{u_1=1}^{k-3} \cup_{u_2=u_1+1}^{k-2} \cup_{v_2=u_2}^{k-2} \cup_{v_2=v_1+1}^{k-1} P_{2,4}^{u_1,u_2,v_1,v_2}(2k) \Big) \setminus P_2^{\text{odd}}(2k).
\end{aligned}$$

Then $\big(\cup_{i=1}^4 P_{2,i}(2k)\big) \subset (P_2(2k) \setminus P_2^{\text{odd}}(2k))$, containing all crossing pair-partitions. We denote the contributions of these partitions to L_p by $T_i, 1 \le i \le 4$. To investigate them, define

$$\begin{aligned}
\mathcal{S}_1 \;=\; & \{s \in \{1,2,\ldots,2k\} : \; 1 \le s \le 2u_1, 2v_1+2 \le s \le 2v_2-1, 2u_2+3 \le s \le 2v_1-1, \\
& \quad 2u_1+3 \le s \le 2u_2, 2v_2+2 \le s \le i_{2k}, i_{2k+1} = i_1\}.
\end{aligned}$$

Then T_1 equals

$$\mathbb{E}\Big[(\delta_{\beta_{2u_1+1}\beta_{2v_1}}(\rho^{1-\epsilon_{2u_1+1}\epsilon_{2v_1}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_{2u_1+1}\epsilon_{2v_1}}))\delta_{i_{2u_1+1}i_{2v_1+1}}\delta_{i_{2u_1+2}i_{2v_1}}A^{\epsilon_{2v_1}}_{\beta_{2v_1}}(i_{2v_1},i_{2u_1+3})$$
$$(\delta_{\beta_{2u_2+1}\beta_{2v_2}}(\rho^{1-\epsilon_{2u_2+1}\epsilon_{2v_2}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_{2u_2+1}\epsilon_{2v_2}}))\delta_{i_{2u_2+1}i_{2v_2+1}}\delta_{i_{2u_2+2}i_{2v_2}}A^{\epsilon_{2u_2+1}}_{\beta_{2u_2+1}}(i_{2u_2+1},i_{2v_2+2})$$
$$\frac{1}{n^k p}\sum_{I_{2k}}\Big(\prod_{s\in\mathcal{S}_1}A^{\epsilon_s}_{\beta_s}(i_s,i_{s+1})\Big)A^{\epsilon_{2u_1+1}}_{\beta_{2u_1+1}}(i_{2u_1+1},i_{2v_1+2})A^{\epsilon_{2v_2}}_{\beta_{2v_2}}(i_{2v_2},i_{2u_2+3})\Big]$$
$$= \Big(\frac{1}{n^2}\Big)\frac{1}{n^{k-2}p}\mathbb{E}\mathrm{Tr}\Big[\Big(\prod_{s_1=1}^{2u_1+1}A^{\epsilon_{s_1}}_{\beta_{s_1}}\Big)\Big(\prod_{s_2=2v_1+2}^{2v_2}A^{\epsilon_{s_2}}_{\beta_{s_2}}\Big)\Big(\prod_{s_3=2u_2+3}^{2v_1}A^{\epsilon_{s_3}}_{\beta_{s_3}}\Big)\Big(\prod_{s_4=2u_1+3}^{2u_2+1}A^{\epsilon_{s_4}}_{\beta_{s_4}}\Big)$$
$$\Big(\prod_{s_5=2v_2+2}^{2k}A^{\epsilon_{s_5}}_{\beta_{s_5}}\Big)\Big](\delta_{\beta_{2u_1+1}\beta_{2v_1}}(\rho^{1-\epsilon_{2u_1+1}\epsilon_{2v_1}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_{2u_1+1}\epsilon_{2v_1}}))$$
$$(\delta_{\beta_{2u_2+1}\beta_{2v_2}}(\rho^{1-\epsilon_{2u_2+1}\epsilon_{2v_2}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_{2u_2+1}\epsilon_{2v_2}}))$$
$$= O(n^{-2}),\quad \text{by (3.10)}. \tag{4.5}$$

Similarly, it can be seen that T_2 equals

$$= \Big(\frac{1}{n^2}\Big)\frac{1}{n^{k-2}p}\mathbb{E}\mathrm{Tr}\Big[\Big(\prod_{s_1=1}^{2u_1}A^{\epsilon_{s_1}}_{\beta_{s_1}}\Big)\Big(\prod_{s_2=2v_1+3}^{2v_2}A^{\epsilon_{s_2}}_{\beta_{s_2}}\Big)\Big(\prod_{s_3=2u_2+3}^{2v_1+1}A^{\epsilon_{s_3}}_{\beta_{s_3}}\Big)\Big(\prod_{s_4=2u_1+2}^{2u_2+1}A^{\epsilon_{s_4}}_{\beta_{s_4}}\Big)$$
$$\Big(\prod_{s_5=2v_2+2}^{2k}A^{\epsilon_{s_5}}_{\beta_{s_5}}\Big)\Big](\delta_{\beta_{2u_1}\beta_{2v_1+1}}(\rho^{1-\epsilon_{2u_1}\epsilon_{2v_1+1}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_{2u_1}\epsilon_{2v_1+1}}))$$
$$(\delta_{\beta_{2u_2+1}\beta_{2v_2}}(\rho^{1-\epsilon_{2u_2+1}\epsilon_{2v_2}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_{2u_2+1}\epsilon_{2v_2}}))$$
$$= O(n^{-2}),\quad \text{by (3.10)}. \tag{4.6}$$

Further, T_3 equals

$$= \Big(\frac{1}{n^2}\Big)\frac{1}{n^{k-2}p}\mathbb{E}\mathrm{Tr}\Big[\Big(\prod_{s_1=1}^{2u_1+1}A^{\epsilon_{s_1}}_{\beta_{s_1}}\Big)\Big(\prod_{s_2=2v_1+2}^{2v_2+1}A^{\epsilon_{s_2}}_{\beta_{s_2}}\Big)\Big(\prod_{s_3=2u_2+2}^{2v_1}A^{\epsilon_{s_3}}_{\beta_{s_3}}\Big)\Big(\prod_{s_4=2u_1+3}^{2u_2}A^{\epsilon_{s_4}}_{\beta_{s_4}}\Big)$$
$$\Big(\prod_{s_5=2v_2+3}^{2k}A^{\epsilon_{s_5}}_{\beta_{s_5}}\Big)\Big](\delta_{\beta_{2u_1+1}\beta_{2v_1}}(\rho^{1-\epsilon_{2u_1+1}\epsilon_{2v_1}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_{2u_1+1}\epsilon_{2v_1}}))$$
$$(\delta_{\beta_{2u_2}\beta_{2v_2+1}}(\rho^{1-\epsilon_{2u_2}\epsilon_{2v_2+1}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_{2u_2}\epsilon_{2v_2+1}}))$$
$$= O(n^{-2}),\quad \text{by (3.10)}. \tag{4.7}$$

Finally, T_4 equals

$$= \Big(\frac{1}{n^2}\Big)\frac{1}{n^{k-2}p}\mathbb{E}\mathrm{Tr}\Big[\Big(\prod_{s_1=1}^{2u_1}A^{\epsilon_{s_1}}_{\beta_{s_1}}\Big)\Big(\prod_{s_2=2v_1+3}^{2v_2+1}A^{\epsilon_{s_2}}_{\beta_{s_2}}\Big)\Big(\prod_{s_3=2u_2+2}^{2v_1+1}A^{\epsilon_{s_3}}_{\beta_{s_3}}\Big)\Big(\prod_{s_4=2u_1+2}^{2u_2}A^{\epsilon_{s_4}}_{\beta_{s_4}}\Big)$$
$$\Big(\prod_{s_5=2v_2+3}^{2k}A^{\epsilon_{s_5}}_{\beta_{s_5}}\Big)\Big](\delta_{\beta_{2u_1}\beta_{2v_1+1}}(\rho^{1-\epsilon_{2u_1}\epsilon_{2v_1+1}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_{2u_1}\epsilon_{2v_1+1}}))$$
$$(\delta_{\beta_{2u_2}\beta_{2v_2+1}}(\rho^{1-\epsilon_{2u_2}\epsilon_{2v_2+1}}(1-\delta_{\rho 0})+\delta_{\rho 0}\delta_{\epsilon_{2u_2}\epsilon_{2v_2+1}}))$$
$$= O(n^{-2}),\quad \text{by (3.10)}. \tag{4.8}$$

Hence, $T_1 + T_2 + T_3 + T_4 = o(1)$. Since $\mathrm{NC}_2(2k) = P_2(2k) \setminus (P_2^{\mathrm{odd}}(2k) \cup (\cup_{i=1}^4 P_{2,i}(2k)))$, we thus conclude that only non-crossing pair-partitions contribute to the limit. Now recalling (4.4), we can conclude that

$$\begin{aligned}
\lim_{p\to\infty} L_p &= \sum_{\pi\in \mathrm{NC}_2(2k)} \lim_{p\to\infty} \frac{1}{pn^k} \sum_{I_{2k}} \prod_{(r,s)\in\pi} E(r,s) \\
&= \sum_{\pi\in \mathrm{NC}_2(2k)} \lim_{n\to\infty} \frac{1}{pn^k} \sum_{I_{2k}} \prod_{(r,s)\in\pi} \delta_{\beta_r\beta_s}\left(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\epsilon_r\epsilon_s}\right) \prod_{r=1}^{2k} \delta_{i_r i_{\gamma\pi(r)}} \\
&= \sum_{\pi\in \mathrm{NC}_2(2k)} \lim_{n\to\infty} \frac{1}{pn^k} \sum_{I_{2k}} \prod_{(r,s)\in\pi} \delta_{\beta_r\beta_s}\left(\rho^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho 0}) + \delta_{\rho 0}\delta_{\epsilon_r\epsilon_s}\right) \prod_{r=1}^{2k} \delta_{i_r i_{\gamma\pi(r)}}.
\end{aligned}$$

This completes the proof of (3.16).

4.2. *Proof of Theorem 3.2(a) for φ_p when $\{(\rho_l, y_l)\}$ are not necessarily equal.* We outline the changes needed in the proof that we gave for the equal case. Recall $\tilde{J}_i$ defined in (3.21). Let $2\nu_i$ be the cardinality of $\tilde{J}_i$ for all $1 \le i \le t$. Therefore, $\sum_{i=1}^t \nu_i = k$. Now, we replace n^k by $\prod_{j=1}^{2k} \sqrt{n_{\beta_j}} = \prod_{i=1}^t n_i^{\nu_i}$ wherever it appears, in particular, at the following places:

(1) Equations (3.6), (3.7), (3.10), (3.13), (4.4), (3.15) and (3.17);
(2) the denominator of (3.9);
(3) the first three lines of (4.2) and first two lines of (4.5);
(4) the first line of (4.6)-(4.8).

The range of i_{2s} in (3.8) is now between 1 and $n_{\beta_{2s}} = n_{\alpha_s}$ for all $1 \le s \le k$. The term $n^{\#J}$ in the numerator of (3.9) should be replaced by $(\max\{n_l : \ 1 \le l \le t\})^{\#J}$. The last lines of (4.2), (4.5)–(4.8) should be replaced by $O((\min\{n_l : \ 1 \le l \le t\})^{-1})$. The term ρ throughout (3.13)–(3.18) should be replaced by ρ_{β_r}.

For any partition π of $\{1, 2, \dots, 2k\}$, define

$$\tilde{S}(\pi) = \text{ number of blocks in } \pi \text{ whose smallest element is odd.}$$

It is easy to see that $S(\gamma\pi) = k + 1 - \tilde{S}(\pi)$ for all non-crossing pair-partitions π. For $1 \le m_i \le \nu_i$, $1 \le i \le t$ and $k \ge 1$, define

$$\tilde{B}_{m_1,m_2,\dots,m_t,k} = \{\pi = \cup_{i=1}^t \pi_i \in \mathrm{NC}_2(2k) : \ \pi_i \in \mathrm{NC}_2(\tilde{J}_i), \ \tilde{S}(\pi_i) = m_i, \ 1 \le i \le t\}.$$

Note that, for all $0 \le m \le k-1$ and $k \ge 1$, we have

$$\tilde{B}_k = \bigcup_{\substack{1\le m_i\le\nu_i \\ 1\le i\le t}} \tilde{B}_{m_1,m_2,\dots,m_t,k}, \quad \text{and} \quad \tilde{B}_{m,k} = \bigcup_{\substack{1\le m_i\le\nu_i,\ 1\le i\le t, \\ \sum_{i=1}^t m_i = k-m}} \tilde{B}_{m_1,m_2,\dots,m_t,k}. \tag{4.9}$$

Therefore,

$$\begin{aligned}
\lim L_p &= \sum_{\pi\in \mathrm{NC}_2(2k)} \prod_{(r,s)\in\pi} \delta_{\beta_r\beta_s}(\rho_{\beta_r}^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho_{\beta_r}0})+\delta_{\rho_{\beta_r}0}\delta_{\epsilon_r\epsilon_s}) \lim_{p\to\infty} \frac{\#\{I_{2k}: i_r = i_{\gamma\pi(r)} \forall r\}}{p\prod_{i=1}^t n_i^{\nu_i}} \\
&= \sum_{1\le m_i\le \nu_i,\ 1\le i\le t} \left(\lim_{p\to\infty} \frac{p^{(k+1)-\sum_{i=1}^t m_i}\prod_{i=1}^t n_i^{m_i}}{p\prod_{i=1}^t n_i^{\nu_i}}\right) \\
&\qquad \times \sum_{\pi\in \tilde{B}_{m_1,m_2,\ldots,m_t,k}} \prod_{i=1}^t \prod_{(r,s)\in\pi_i} (\rho_{\beta_r}^{1-\delta_{\epsilon_r\epsilon_s}}(1-\delta_{\rho_{\beta_r}0})+\delta_{\rho_{\beta_r}0}\delta_{\epsilon_r\epsilon_s}) \\
&= \sum_{1\le m_i\le \nu_i,\ 1\le i\le t} \left(\prod_{i=1}^t y_i^{\nu_i-m_i}\right) \sum_{\pi\in \tilde{B}_{m_1,m_2,\ldots,m_t,k}} \prod_{i=1}^t (\rho_i^{T(\pi_i)}(1-\delta_{\rho_i 0})+\delta_{\rho_i 0}\delta_{T(\pi_i)\ 0}). \quad (4.10)
\end{aligned}$$

Note that the cardinality of J_i is ν_i. For $1\le m_i\le \nu_i$, $1\le i\le t$ and $k\ge 1$, define

$$B_{m_1,m_2,\ldots,m_t,k} = \{\pi = \cup_{i=1}^t \pi_i \in \mathrm{NC}(k):\ \pi_i\in \mathrm{NC}(J_i),\ \pi_i \text{ has } m_i \text{ blocks},\ 1\le i\le t\}.$$

Note that, for all $0\le m\le k-1$ and $k\ge 1$, we have

$$B_k = \bigcup_{1\le m_i\le\nu_i,\ 1\le i\le t} B_{m_1,m_2,\ldots,m_t,k}, \quad \text{and} \quad B_{k-m,k} = \bigcup_{1\le m_i\le\nu_i,\ 1\le i\le t,\ \sum_{i=1}^t m_i = k-m} B_{m_1,m_2,\ldots,m_t,k}.$$

Moreover, by (3.23)–(3.26), it is clear that f (defined around (3.23)) is a bijection between $\tilde{B}_{m_1,m_2,\ldots,m_t,k}$ and $B_{m_1,m_2,\ldots,m_t,k}$. Also $T(\pi_i) = \mathcal{T}(f(\pi_i))$ for all $\pi = \cup_{i=1}^t \pi_i \in \tilde{B}_{m_1,m_2,\ldots,m_t,k}$ and $\mathcal{T}(\pi_i) = T(f^{-1}(\pi_i))$ for all $\pi = \cup_{i=1}^t\pi_i \in B_{m_1,m_2,\ldots,m_t,k}$.

Hence, by (4.10), we have

$$\begin{aligned}
\lim L_p &= \sum_{1\le m_i\le \nu_i,\ 1\le i\le t} \sum_{\pi\in B_{m_1,m_2,\ldots,m_t,k}} \prod_{i=1}^t y_i^{\nu_i-m_i}(\rho_i^{\mathcal{T}(\pi_i)}(1-\delta_{\rho_i 0})+\delta_{\rho_i 0}\delta_{\mathcal{T}(\pi_i)\ 0}) \\
&= \varphi_p(c_{\alpha_1}^{\eta_1}c_{\alpha_2}^{\eta_2}\cdots c_{\alpha_k}^{\eta_k}).
\end{aligned}$$

Hence, by moment-free cumulant relation, we have

$$\begin{aligned}
\kappa_\pi[c_{\alpha_1}^{\eta_1}, c_{\alpha_2}^{\eta_2},\ldots,c_{\alpha_k}^{\eta_k}] &= 0 \text{ for all } \pi\in \mathrm{NC}(k)-B_k, \\
\kappa_\pi[c_{\alpha_1}^{\eta_1}, c_{\alpha_2}^{\eta_2},\ldots,c_{\alpha_k}^{\eta_k}] &= \prod_{i=1}^t\prod_{j=1}^{m_i} y_i^{\#V_{ij}-1}(\rho_i^{\tilde{T}(V_{ij})}(1-\delta_{\rho_i 0})+\delta_{\rho_i 0}\delta_{\tilde{T}(V_{ij})0}) \text{ for all } \pi\in B_{m_1,m_2,\ldots,m_t,k},
\end{aligned}$$

where $\pi_i = \{V_{ij}:\ 1\le j\le m_i\}$ for all $1\le i\le t$. This implies that

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

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

$$\kappa_k(c_l^{\eta_1}, c_l^{\eta_2},\ldots,c_l^{\eta_k}) = y_l^{k-1}(\rho_l^{S(\boldsymbol{\eta}_k)}(1-\delta_{\rho_l 0})+\delta_{\rho_l 0}\delta_{S(\boldsymbol{\eta}_k)0}).$$

This completes the proof of Theorem 3.2 (a) for the state φ_p.

4.3. *Proof of (3.29).* Equation (3.29) is derived by a counting argument which is similar to what has already been used in the proof of Theorem 3.2(a) for the state φ_p. Similar argument can also be found in Section 2 in the Supplementary material of Bhattacharjee and Bose (2016a) in a different context.

Proof of (3.29) for $y_l > 0$: Let Π be any polynomial in $\{C_l, C_l^* :\ 1 \le l \le t\}$. Write $\Pi = \sum_{i=1}^{r} m_i$ where m_i's are monomials in $\{C_l, C_l^* :\ 1 \le l \le t\}$. Then for some $C > 0$,

$$\begin{aligned}(\mathrm{Tr}(\Pi) - \mathbb{E}(\mathrm{Tr}(\Pi)))^4 &= (\sum_{i=1}^{r}(\mathrm{Tr}(m_i) - \mathbb{E}(\mathrm{Tr}(m_i))))^4 \\ &\le C\sum_{i=1}^{r}(\mathrm{Tr}(m_i) - \mathbb{E}(\mathrm{Tr}(m_i)))^4.\end{aligned}$$

Hence it is enough to prove (3.29) when Π is a monomial. Now

$$\begin{aligned}\mathbb{E}(\mathrm{Tr}(\Pi) - \mathbb{E}(\mathrm{Tr}(\Pi)))^4 \quad &= \quad \mathbb{E}(\mathrm{Tr}(\Pi))^4 - 4\mathbb{E}(\mathrm{Tr}(\Pi))^3\mathbb{E}(\mathrm{Tr}(\Pi)) \\ &\quad\quad +6\mathbb{E}(\mathrm{Tr}(\Pi))^2(\mathbb{E}(\mathrm{Tr}(\Pi)))^2 - 3(\mathbb{E}(\mathrm{Tr}(\Pi)))^4 \\ &= \quad T_1 - 4T_2T_4 + 6T_3T_4^2 - 3T_4^4, \quad \text{say,} \end{aligned} \tag{4.11}$$

where $T_i = \mathbb{E}(\mathrm{Tr}(\Pi))^{5-i},\ 1 \le i \le 4$. Without loss of generality assume that Π is of the form,

$$\Pi = (n_{\alpha_1}^{-1}X_{\alpha_1}Y_{\alpha_1}^*)^{\eta_1}\cdots(n_{\alpha_k}^{-1}X_{\alpha_k}Y_{\alpha_k}^*)^{\eta_k}. \tag{4.12}$$

For any two non-negative integers i and j, the remainder obtained after dividing i by j is denoted by $(i \bmod j)$. Moreover, $\beta_{2rk+i} = \beta_i$ and $\epsilon_{2rk+i} = \epsilon_i$ for all $1 \le i \le 2k$ and $1 \le r \le 3$. Using arguments similar to those used in the proof of (3.13), for $1 \le i \le 4$, we have

$$\begin{aligned}\lim_{p\to\infty} T_i \quad &= \quad \sum_{\pi\in P_2(2(5-i)k)} \lim_{p\to\infty} \frac{1}{\left(\prod_{j=1}^{2(5-i)k}\sqrt{n_{\beta_j}}\right)} \sum_{I_{2(5-i)k}} \left(\prod_{(r,s)\in\pi} E_1(r,s)\right) \\ &\quad\quad \mathbb{E}\left(\prod_{1\le j\le 2k}\prod_{r=0}^{4-i} A_{\beta_{2rk+j}}^{(\epsilon_{2rk+j})}(i_{2rk+j}, i_{2rk+(j \bmod 2k)+1})\right) \\ &= \quad \sum_{\pi\in P_2(2(5-i)k)} \lim_{p\to\infty} G_{i\pi}, \quad \text{say,}\end{aligned} \tag{4.13}$$

where, suppressing the dependence on other variables,

$$E_1(r,s) \quad = \quad \delta_{\beta_r\beta_s}(\delta_{i_ri_s}\delta_{i_{r+1}i_{s+1}}a(r,s) + \delta_{i_ri_{s+1}}\delta_{i_si_{r+1}}(1-a(r,s))).$$

We break up the sums in each T_i into sums over appropriate subsets. For this, let $[a,b]$ denote the set of integers from a to b. Consider the following subsets of $\mathrm{NC}_2(rR)$ for $r = 2,3,4$.

$$\begin{aligned}
\mathcal{F}_{11,R} &= \{\pi = \cup_{i=1}^4 \pi_i \in P_2(4R): \ \pi_i \in P_2([(i-1)R+1, iR]) \ \forall\ 1 \le i \le 4\},\\
\mathcal{F}_{12,R} &= \{\pi = \pi_1 \cup \pi_2 \cup \pi_3 \in P_2(4R) \setminus \mathcal{F}_{11,R}:\\
&\qquad \pi_1 \in P_2([rR+1, (r+1)R]),\ \pi_2 \in P_2([sR+1, (s+1)R]),\\
&\qquad \pi_3 \in P_2(\cup_{\substack{0\le u\le 3\\ u\ne r,s}}[uR+1, (u+1)R]) \text{ for some } r \ne s,\ r,s = 0,1,2,3\},\\
\mathcal{F}_{13,R} &= \{\pi = \pi_1 \cup \pi_2 \in P_2(4R) \setminus (\mathcal{F}_{11,R} \cup \mathcal{F}_{12,R}):\\
&\qquad \pi_1 \in P_2([rR+1, (r+1)R] \cup [sR+1, (s+1)R]),\\
&\qquad \pi_2 \in P_2(\cup_{\substack{0\le u\le 3\\ u\ne r,s}}[uR+1, (u+1)R]) \text{ for some } r \ne s,\ r,s = 0,1,2,3\},\\
\mathcal{F}_{14,R} &= \{\pi = \pi_1 \cup \pi_2 \in P_2(4R) \setminus (\cup_{i=1}^3 \mathcal{F}_{1i,R}):\\
&\qquad \pi_1 \in P_2([rR+1, (r+1)R]),\\
&\qquad \pi_2 \in P_2([1,4R] \setminus [rR+1, (r+1)R]) \text{ for some } r = 0,1,2,3\},\\
\mathcal{F}_{21,R} &= \{\pi = \cup_{i=1}^3 \pi_i \in P_2(3R): \ \pi_i \in P_2([(i-1)R+1, iR]) \ \forall\ 1 \le i \le 3\},\\
\mathcal{F}_{22,R} &= \{\pi = \pi_1 \cup \pi_2 \in P_2(3R) \setminus \mathcal{F}_{21,R}: \ \pi_1 \in P_2([rR+1, (r+1)R]),\\
&\qquad \pi_2 \in P_2([1,3R] \setminus [rR+1, (r+1)R]),\ \text{for some } r = 0,1,2\},\\
\mathcal{F}_{31,R} &= \{\pi = \pi_1 \cup \pi_2 \in P_2(2R): \ \pi_i \in P_2([(i-1)R+1, iR]),\ i = 1,2\},\\
\mathcal{F}_{15,R} &= P_2(4R) \setminus \left(\cup_{i=1}^4 \mathcal{F}_{1i,R}\right),\ \mathcal{F}_{23,R} = P_2(3R) \setminus (\mathcal{F}_{21,R} \cup \mathcal{F}_{22,R}),\ \mathcal{F}_{32,R} = P_2(2R) \setminus \mathcal{F}_{31,R}. \quad (4.14)
\end{aligned}$$

Split T_1, T_2, T_3 as

$$T_{1i} = \sum_{\pi \in \mathcal{F}_{1i,2k}} G_{1\pi}, \quad T_{2i} = \sum_{\pi \in \mathcal{F}_{2i,2k}} G_{2\pi}, \quad T_{3i} = \sum_{\pi \in \mathcal{F}_{3i,2k}} G_{3\pi}. \tag{4.15}$$

Now note that

$$T_{11} = T_4^4,\ T_{21} = T_4^3,\ T_{31} = T_4^2,\ T_{12} = 6T_{32}T_4^2,\ T_{22} = 3T_{32}T_4,\ T_{14} = 4T_{23}T_4. \tag{4.16}$$

Using this and (4.11),

$$\lim \mathbb{E}(\mathrm{Tr}(\Pi) - \mathbb{E}(\mathrm{Tr}(\Pi)))^4 = \lim(T_{13} + T_{15}).$$

Next consider the following decomposition of $\mathcal{F}_{13,R}$.

$$\begin{aligned}
\mathcal{F}_{13,R} &= \cup_{u_1,u_2,u_3,u_4=1}^{2k} \cup_{i=1}^3 \mathcal{F}_{13i,R}^{(u_1,u_2,u_3,u_4)}, \text{ where}\\
\mathcal{F}_{131,R}^{(u_1,u_2,u_3,u_4)} &= \{\pi \in \mathcal{F}_{13,R}: \ (u_1, R+u_2), (2R+u_3, 3R+u_4) \in \pi\},\\
\mathcal{F}_{132,R}^{(u_1,u_2,u_3,u_4)} &= \{\pi \in \mathcal{F}_{13,R}: \ (u_1, 2R+u_2), (R+u_3, 3R+u_4) \in \pi\},\\
\mathcal{F}_{133,R}^{(u_1,u_2,u_3,u_4)} &= \{\pi \in \mathcal{F}_{13,R}: \ (u_1, 3R+u_2), (R+u_3, 2R+u_4) \in \pi\}.
\end{aligned}$$

We split T_{13} accordingly into pieces T_{13i} where

$$T_{13i}^{(u_1,u_2,u_3,u_4)} = \sum_{\pi \in \mathcal{F}_{13i,2k}^{(u_1,u_2,u_3,u_4)}} G_{1\pi},\ 1 \le i \le 3,\ 1 \le u_1,u_2,u_3,u_4 \le 2k. \tag{4.17}$$

Also note that

$$
\begin{aligned}
& T_{131}^{(u_1,u_2,u_3,u_4)} \\
& \quad = \delta_{\beta_{u_1}\beta_{2k+u_2}}(\rho_{\beta_{u_1}}^{1-\delta_{\epsilon_{u_1}\epsilon_{2k+u_2}}}(1-\delta_{\rho_{\beta_{u_1}}0})+\delta_{\rho_{\beta_{u_1}}0}\delta_{\epsilon_{u_1}\epsilon_{2k+u_2}}) \\
& \quad \times\ \delta_{\beta_{4k+u_3}\beta_{6k+u_4}}(\rho_{\beta_{4k+u_3}}^{1-\delta_{\epsilon_{4k+u_3}\epsilon_{6k+u_4}}}(1-\delta_{\rho_{\beta_{2k+u_3}}0})+\delta_{\rho_{\beta_{4k+u_3}}0}\delta_{\epsilon_{4k+u_3}\epsilon_{6k+u_4}}) \\
& \quad \times\ \frac{1}{\left(\prod_{j=1}^{8k}\sqrt{n_{\beta_j}}\right)}\mathbb{E}\mathrm{Tr}\Big(\Big(\prod_{j=u_1+1}^{2k}A_{\beta_j}^{(\epsilon_j)}\Big)\Big(\prod_{j=1}^{u_1-1}A_{\beta_j}^{(\epsilon_j)}\Big)\Big(\prod_{j=2k+u_2+1}^{4k}A_{\beta_j}^{(\epsilon_j)}\Big)\Big(\prod_{j=2k+1}^{2k+u_2-1}A_{\beta_j}^{(\epsilon_j)}\Big)\Big) \\
& \quad \times\ \mathbb{E}\mathrm{Tr}\Big(\Big(\prod_{j=4k+u_3+1}^{6k}A_{\beta_j}^{(\epsilon_j)}\Big)\Big(\prod_{j=4k+1}^{4k+u_3-1}A_{\beta_j}^{(\epsilon_j)}\Big)\Big(\prod_{j=6k+u_4+1}^{8k}A_{\beta_j}^{(\epsilon_j)}\Big)\Big(\prod_{j=6k+1}^{6k+u_4-1}A_{\beta_j}^{(\epsilon_j)}\Big)\Big) \\
& \quad = O(1), \quad \text{by (3.10)}.
\end{aligned}
$$

Similarly, $T_{13i}^{(u_1,u_2,u_3,u_4)} = O(1)$ for $i = 2, 3$. Hence, $T_{13} = O(1)$.

Next consider the following decomposition of $\mathcal{F}_{15,R}$.

$$
\begin{aligned}
\mathcal{F}_{15,R} &= \bigcup_{1\le u_j\le 2k,\ 1\le j\le 6} \mathcal{F}_{15,R}^{(u_1,u_2,u_3,u_4,u_5,u_6)} \quad \text{where} \\
\mathcal{F}_{15,R}^{(u_1,u_2,u_3,u_4,u_5,u_6)} &= \{\pi\in\mathcal{F}_{15,R}:\ (u_1,R+u_2),(R+u_3,2R+u_4),(2R+u_5,3R+u_6)\in\pi\},
\end{aligned}
$$

and split up up T_{15} accordingly into pieces

$$
T_{15}^{(u_1,u_2,u_3,u_4,u_5,u_6)} = \sum_{\pi\in\mathcal{F}_{15,2k}^{(u_1,u_2,u_3,u_4,u_5,u_6)}} G_{1\pi},\ 1\le u_j\le 2k,\ 1\le j\le 6. \tag{4.18}
$$

Also note that, for $u_2 < u_3$ and $u_4 < u_5$, we have

$$
\begin{aligned}
T_{15}^{(u_1,u_2,u_3,u_4,u_5,u_6)} &= \delta_{\beta_{u_1}\beta_{2k+u_2}}(\rho_{\beta_{u_1}}^{1-\delta_{\epsilon_{u_1}\epsilon_{2k+u_2}}}(1-\delta_{\rho_{\beta_{u_1}}0})+\delta_{\rho_{\beta_{u_1}}0}\delta_{\epsilon_{u_1}\epsilon_{2k+u_2}}) \\
&\times\ \delta_{\beta_{2k+u_3}\beta_{4k+u_4}}(\rho_{\beta_{2k+u_3}}^{1-\delta_{\epsilon_{2k+u_3}\epsilon_{4k+u_4}}}(1-\delta_{\rho_{\beta_{2k+u_3}}0})+\delta_{\rho_{\beta_{2k+u_3}}0}\delta_{\epsilon_{2k+u_3}\epsilon_{4k+u_4}}) \\
&\times\ \delta_{\beta_{4k+u_5}\beta_{6k+u_6}}(\rho_{\beta_{4k+u_5}}^{1-\delta_{\epsilon_{4k+u_5}\epsilon_{6k+u_6}}}(1-\delta_{\rho_{\beta_{4k+u_5}}0})+\delta_{\rho_{\beta_{4k+u_5}}0}\delta_{\epsilon_{4k+u_5}\epsilon_{6k+u_6}}) \\
&\times\ \frac{1}{\left(\prod_{j=1}^{8k}\sqrt{n_{\beta_j}}\right)}\mathbb{E}\mathrm{Tr}\Big(\Big(\prod_{j=u_1+1}^{2k}A_{\beta_j}^{(\epsilon_j)}\Big)\Big(\prod_{j=1}^{u_1-1}A_{\beta_j}^{(\epsilon_j)}\Big)\Big(\prod_{j=2k+u_2+1}^{2k+u_3-1}A_{\beta_j}^{(\epsilon_j)}\Big) \\
&\times\ \Big(\prod_{j=4k+u_4+1}^{4k+u_5-1}A_{\beta_j}^{(\epsilon_j)}\Big)\Big(\prod_{j=6k+u_6+1}^{8k}A_{\beta_j}^{(\epsilon_j)}\Big)\Big(\prod_{j=6k+1}^{6k+u_6-1}A_{\beta_j}^{(\epsilon_j)}\Big) \\
&\times\ \Big(\prod_{j=4k+u_5+1}^{6k}A_{\beta_j}^{(\epsilon_j)}\Big)\Big(\prod_{j=4k+1}^{4k+u_4-1}A_{\beta_j}^{(\epsilon_j)}\Big)\Big(\prod_{j=2k+u_3+1}^{4k}A_{\beta_j}^{(\epsilon_j)}\Big) \\
&\times\ \Big(\prod_{j=2k+1}^{2k+u_2-1}A_{\beta_j}^{(\epsilon_j)}\Big)\Big) = O(p^{-2}).
\end{aligned}
$$

Similar argument also works for other choices of $u_1,\ldots,u_6$. Hence, $T_{15} = O(p^{-2})$ and hence

$$
\mathbb{E}(\mathrm{Tr}(\Pi) - \mathbb{E}(\mathrm{Tr}(\Pi)))^4 = O(1).
$$

This completes the proof of (3.29) for the case $y_l > 0$.

Proof of (3.29) for the case $y_l = 0$: This proof is similar to the proof presented for the $y_l \neq 0$ case. The only differences are as follows.

1. Π in (4.12) should be replaced by $\Pi = E_{\alpha_1}^{\eta_1} \cdots E_{\alpha_k}^{\eta_k}$.
2. $G_{i\pi}$ in (4.13) should be replaced by

$$G_{i\pi} = \frac{1}{p^{k/2} \prod_{j=1}^{(5-i)k} \sqrt{n_{\alpha_j}}} \sum_{I_{(5-i)k}} \mathbb{E}\Big(\prod_{r=1}^{k} a_{l_r i_r j_r} a_{l_r i_{(r \bmod k)+1} j_r}\Big),$$

and $\lim_{p\to\infty} T_i = \sum_{\pi \in P_2((5-i)k)} \lim_{p\to\infty} G_{i\pi}$.

3. For $1 \le j \le 3$, the set $\mathcal{F}_{ji,2k}$ in (4.15) should be replaced by $\mathcal{F}_{ji,k}$.
4. The set $\mathcal{F}_{13i,2k}$ in (4.17) should be replaced by $\mathcal{F}_{13i,R}$.
5. The set $\mathcal{F}_{15,2k}$ in (4.18) should be replaced by $\mathcal{F}_{15,k}$.

As before it can be shown easily that $T_{13} = O(1)$ and $T_{15} = O(p^{-2})$. This completes the proof of (3.29) for $\{E_l\}$.

4.4. *Proof of Theorem 3.11(a) for φ_p when $\{\rho_l\}$ are not necessarily equal.* 1. The term $n^{k/2}$ in the denominator of (3.34) and (3.41), should be replaced by $\prod_{j=1}^{k} \sqrt{n_{l_j}}$.

2. In (3.38) and (3.40), n should be replaced by $\min\{n_l : 1 \le l \le t\}$.
3. In (3.42), (3.43) and (3.50), ρ should be replaced by ρ_{l_r}.
4. In (3.42)–(3.49), n^m should be replaced by $\prod_{j=1}^{2m} \sqrt{n_{l_j}}$ wherever it appears.

Finally, (3.50) will be

$$\begin{aligned}\sum_{\pi \in \mathrm{NC}_2(2m)} \prod_{(r,s)\in\pi} \delta_{l_r l_s}(\rho_{l_r}^2 \delta_{\epsilon_r\epsilon_s} + (1-\delta_{\epsilon_r\epsilon_s})) &= \sum_{\pi \in \mathrm{NC}_2(2m)} \prod_{(r,s)\in\pi} \delta_{l_r l_s}(\rho_{l_r}^2)^{\delta_{\epsilon_r\epsilon_s}} \\ &= \sum_{\pi \in \mathrm{NC}_2(2k)} \rho_1^{T_1(\pi)} \ldots \rho_m^{T_m(\pi)} \prod_{(r,s)\in\pi} \delta_{\tau_r\tau_s}\end{aligned}$$

where $T_\tau(\pi)$ is defined after (3.31). This completes the proof of Theorem 3.11(a).

Figure 4.2 reports the simulation results for a few polynomials when p/n is small (0.03) for different values of ρ.

Acknowledgment. We thank the referees for their constructive comments, and these have led to a very significant improvement in the article in terms of both substance and presentation. We also thank the Associate Editor for her/his encouragement to submit a revision.

References

Adhikari, K. and Bose, A. Brown measure and asymptotic freeness of elliptic and related matrices. *Random Matrices Theory Appl.*, **8** (2), 1950007, 23 (2019). MR3943901.

Akemann, G., Byun, S.-S., and Kang, N.-G. A non-Hermitian generalisation of the Marchenko-Pastur distribution: from the circular law to multi-criticality. *Ann. Henri Poincaré*, **22** (4), 1035–1068 (2021). MR4229527.

Akemann, G., Phillips, M. J., and Sommers, H.-J. Characteristic polynomials in real Ginibre ensembles. *J. Phys. A*, **42** (1), 012001, 9 (2009). MR2525254.

Ash, R. B. and Doléans-Dade, C. A. *Probability and measure theory*. Harcourt/Academic Press, Burlington, MA, second edition (2000). ISBN 0-12-065202-1. MR1810041.

Bai, Z. and Silverstein, J. W. *Spectral analysis of large dimensional random matrices*. Springer Series in Statistics. Springer, New York, second edition (2010). ISBN 978-1-4419-0660-1. MR2567175.

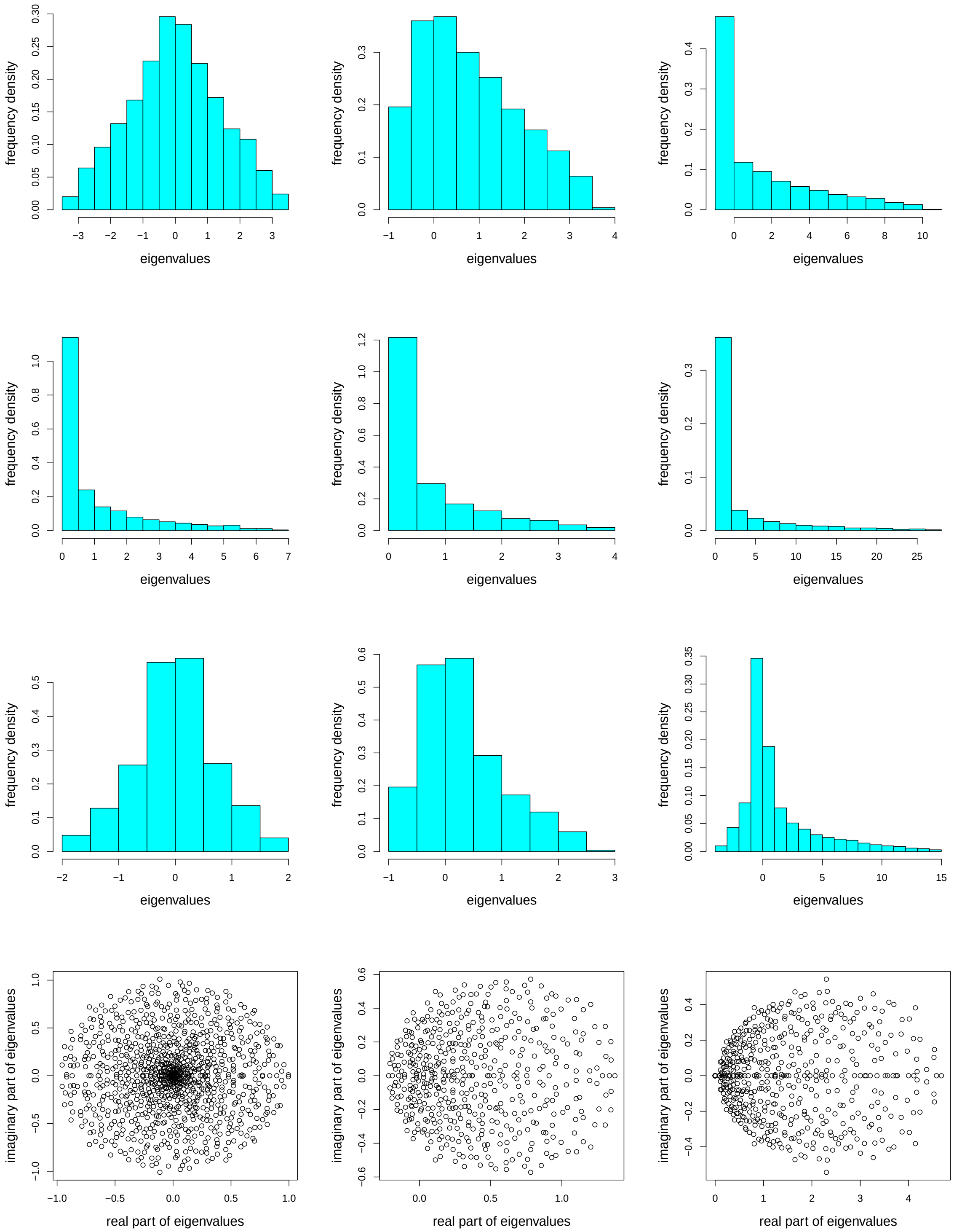

FIGURE 4.1. Histogram for ESD of $(C + C^*)$, CC^* and $C_1C_2^* + C_2C_1^*$ ($n_1 = n$, $n_2 = 2n$, $\rho_1 = \rho_2 = \rho$) in Rows 1-3 respectively. ESD of C in Row 4. Column 1: $n = p = 500$, $\rho = 0$, Column 2: $n = 1000$, $p = 500$, $\rho = 0.4$ Column 3: $n = 500$, $p = 1000$, $\rho = 0.8$.

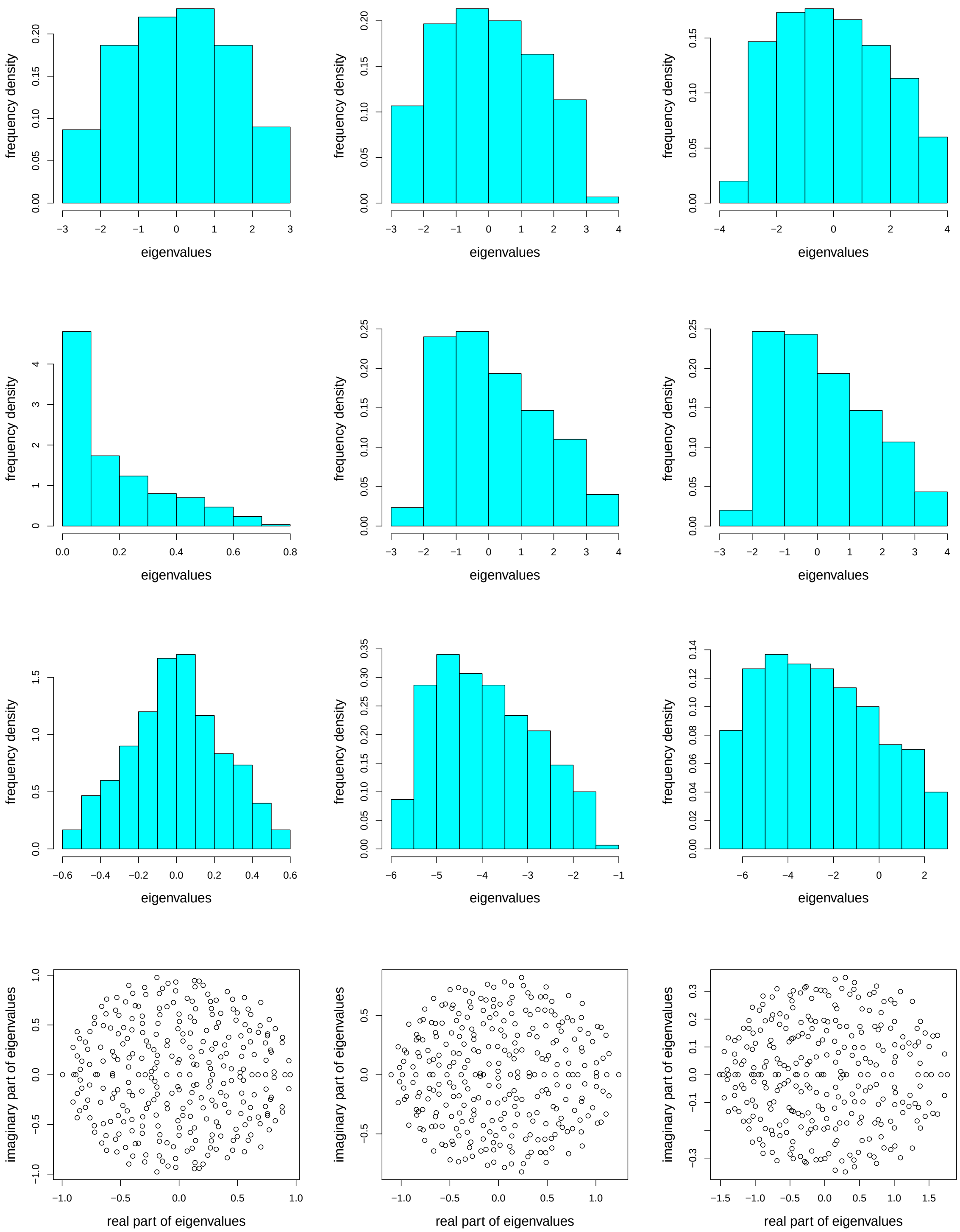

FIGURE 4.2. Histogram for ESD of $\sqrt{np^{-1}}(C + C^* - 2\rho \mathrm{I}_p)$, $\sqrt{np^{-1}}(CC^* - \rho^2 \mathrm{I}_p)$ and $\sqrt{n_2 p^{-1}}(C_1 C_2^* + C_2 C_1^* - 2\rho^2 \mathrm{I}_p)$ $(n_1 = n,\ n_2 = 2n,\ \rho_1 = \rho_2 = \rho)$ in Rows 1-3 respectively. Scatter plot for ESD of $\sqrt{np^{-1}}(C - \rho \mathrm{I}_p)$ 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.

Bhattacharjee, M. and Bose, A. Large sample behaviour of high dimensional autocovariance matrices. *Ann. Statist.*, **44** (2), 598–628 (2016a). MR3476611.

Bhattacharjee, M. and Bose, A. Polynomial generalizations of the sample variance-covariance matrix when $pn^{-1} \to 0$. *Random Matrices Theory Appl.*, **5** (4), 1650014, 41 (2016b). MR3569842.

Bhattacharjee, M. and Bose, A. Matrix polynomial generalizations of the sample variance-covariance matrix when $pn^{-1} \to y \in (0, \infty)$. *Indian J. Pure Appl. Math.*, **48** (4), 575–607 (2017). MR3741695. Erratum **49** (5), 783–788 (2018) MR3882010.

Bordenave, C. and Chafaï, D. Around the circular law. *Probab. Surv.*, **9**, 1–89 (2012). MR2908617.

Bose, A. *Patterned random matrices*. CRC Press, Boca Raton, FL (2018). ISBN 978-1-138-59146-2. MR3823788.

Bose, A. and Hachem, W. Smallest singular value and limit eigenvalue distribution of a class of non-Hermitian random matrices with statistical application. *J. Multivariate Anal.*, **178**, 104623, 24 (2020). MR4082540.

Bose, A. and Hachem, W. Spectral measure of empirical autocovariance matrices of high-dimensional Gaussian stationary processes (2023+). To appear in *Random Matrices Theory Appl.*, online ready. DOI: 10.1142/S2010326322500538.

Capitaine, M. and Casalis, M. Asymptotic freeness by generalized moments for Gaussian and Wishart matrices. Application to beta random matrices. *Indiana Univ. Math. J.*, **53** (2), 397–431 (2004). MR2060040.

Marčenko, V. A. and Pastur, L. A. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, **72 (114)**, 507–536 (1967). MR0208649.

Nguyen, H. H. and O'Rourke, S. The elliptic law. *Int. Math. Res. Not. IMRN*, **2015** (17), 7620–7689 (2015). MR3403996.

Nica, A. and Speicher, R. *Lectures on the combinatorics of free probability*, volume 335 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge (2006). ISBN 978-0-521-85852-6; 0-521-85852-6. MR2266879.

O'Rourke, S., Renfrew, D., Soshnikov, A., and Vu, V. Products of independent elliptic random matrices. *J. Stat. Phys.*, **160** (1), 89–119 (2015). MR3357969.

Vinayak and Benet, L. Spectral domain of large nonsymmetric correlated Wishart matrices. *Phys. Rev. E*, **90**, 042109 (2014). DOI: 10.1103/PhysRevE.90.042109.