Matrix polynomial generalizations of the sample variance-covariance matrix when $pn^{-1} \to y \in (0, \infty)$

Monika Bhattacharjee
Statistics and Mathematics Unit
Indian Statistical Institute
203 B.T. Road, Kolkata 700108
INDIA
monaiidexp.gamma@gmail.com

Arup Bose [*]
Statistics and Mathematics Unit
Indian Statistical Institute
203 B.T. Road, Kolkata 700108
INDIA
bosearu@gmail.com

April 30, 2016

Abstract

Let $\{Z_u = ((\varepsilon_{u,i,j}))_{p \times n}\}$ be random matrices where $\{\varepsilon_{u,i,j}\}$ are independently distributed. Suppose $\{A_i\}$, $\{B_i\}$ are non-random matrices of order $p \times p$ and $n \times n$ respectively. Consider all $p \times p$ random matrix polynomials $\mathbb{P} = \prod_{i=1}^{k_l} \left(n^{-1} A_{t_i} Z_{j_i} B_{s_i} Z_{j_i}^* \right)$ $A_{t_{k_l+1}}$. We show that under appropriate conditions on the above matrices, the elements of the non-commutative $*$- probability space Span$\{\mathbb{P}\}$ with state $p^{-1}E\text{Tr}$ converge. As a by product, we also show that the Limiting Spectral Distribution of any self-adjoint polynomial in Span$\{\mathbb{P}\}$ exists almost surely.

Key words and phrases. Independent matrix, moment method, Stieltjes transformation, limiting spectral distribution, semi-circle law, non-crossing partition, non-commutative probability space, $*$-algebra, free cumulants.

AMS 2010 Subject Classifications. Primary 58C40; Secondary 46L54.

1 Introduction

The *empirical spectral distribution* (ESD) of a $p \times p$ (random) matrix A_p is the (random) probability distribution with mass $1/p$ at each of its eigenvalues. If it converges weakly (almost surely) to a non-degenerate probability distribution, then the latter is called the *limiting spectral*

---

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

distribution (LSD) of A_p. One of the most important LSD result in random matrix theory is the Marčenko-Pastur LSD of the sample variance-covariance matrix. This result and some of its variations are discussed later in this section.

These results motivate us to study the LSD of a comprehensive class of self-adjoint matrix polynomial generalizations of the sample variance-covariance matrix. This class is described later.

Note that such LSD are inherently marginal convergence results and throw no light on the joint convergence when there is more than one matrix. The appropriate joint convergence notion comes via non-commutative probability spaces.

Recall that a *non-commutative probability space* (NCP) $(\mathcal{A}, \varphi)$ consists of a unital $*$-algebra $\mathcal{A}$ and a linear functional $\varphi : \mathcal{A} \to \mathbb{R}$ with $\varphi(1_{\mathcal{A}}) = 1$. φ is called a *state* of $\mathcal{A}$. The elements of $\mathcal{A}$ are called *non-commutative variables*. The numbers $\{\varphi(\Pi(a_i, a_i^* : 1 \leq i \leq k)) : \Pi \text{ is a polynomial}\}$ are *joint moments* or is the *joint distribution* of $\{a_i : 1 \leq i \leq k\}$. We shall say that a sequence of NCP $(\mathcal{A}_p = \text{Span}(\{a_i^{(p)}, a_i^{(p)*} : i \geq 1\}), \varphi_p)$ *converges* to $(\mathcal{A} = \text{Span}(\{a_i, a_i^* : i \geq 1\}, \varphi)$ if for any polynomial Π,

$$\lim_{p\to\infty} \varphi_n(\Pi(a_i^{(p)}, a_i^{(p)*} : i \geq 1)) = \varphi(\Pi(a_i, a_i^* : i \geq 1)). \tag{1.1}$$

Now suppose $\mathbb{M}_p$ is a class of square random matrices of order p. Let $\mathcal{M}_p$ be the *-algebra generated by $\mathbb{M}_p$ and its adjoints. Let $\phi_p = p^{-1}E\text{Tr}$ be the state defined on $\mathcal{M}_p$. Then one can consider the convergence of $(\mathcal{M}_p, \phi_p)$. There is a huge literature on such convergence for different generators $\mathbb{M}_p$. See Nica and Speicher [2006] and Anderson et al. [2009]. Some of these results that are relevant to our work are collected for quick reference in Lemma 2.1.

Now we introduce the class of matrices that is the focus of this article. Suppose we have matrices $Z_u = ((\varepsilon_{u,i,j}))_{p\times n},\ 1 \leq u \leq U$, where $\{\varepsilon_{u,i,j} : u, i, j \geq 0\}$ are independent. Also suppose $\{B_{2i-1}\}$ and $\{B_{2i}\}$ are constant matrices of order $p \times p$ and $n \times n$ respectively. Consider all $p \times p$ matrices

$$\mathbb{P}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} = \prod_{i=1}^{k_l} \left(n^{-1} A_{l,2i-1} Z_{u_{l,i}} A_{l,2i} Z^*_{u_{l,i}}\right) A_{l,2k_l+1}, \tag{1.2}$$

where $\{A_{l,2i-1}\}$, $\{A_{l,2i}\}$ and $Z_{u_{l,i}}$ are matrices from the collections $\{B_{2i-1}, B^*_{2i-1}\}$, $\{B_{2i}, B^*_{2i}\}$ and $\{Z_i\}$ respectively.

Consider the sequence of NCP $(\mathcal{U}_p,\ \phi_p)$, where

$$\mathcal{U}_p = \text{Span}\left(\mathbb{P}_{l,(u_{l,1},\ldots,u_{l,k_l})} :\ l, k_l \geq 1\right). \tag{1.3}$$

Note that the usual sample variance-covariance matrix $n^{-1}Z_1Z_1^*$ is a variable in the above space. We are interested in the convergence of $(\mathcal{U}_p,\ \phi_p)$ as $p, n \to \infty$ and $pn^{-1} \to y \in (0,\ \infty)$. The case $y = 0$ is very different in nature and shall be treated separately elsewhere.

If $\mathbb{M}_p = \{A_p\}$ consists of a single real symmetric square matrix A_p of order p, then there is a connection between the convergence of $(\text{Span}(A_p), \phi_p)$ and the LSD of A_p. This link is apparent if we look at the moment method to establish LSD.

The h-th order moment of the ESD of a $p \times p$ real symmetric matrix A_p equals $\beta_h(A_p) := \frac{1}{p}\,\text{Tr}(A_p^h)$. Consider the following conditions.

(M1) For every $h \geq 1, E(\beta_h(A_p)) \to \beta_h$,

(M4) $\sum_{p=1}^{\infty} E(\beta_h(A_p) - E(\beta_h(A_p)))^4 < \infty,\ \ \forall h \geq 1$, and

(C) The sequence $\{\beta_h\}$ satisfies Carleman's condition, $\sum_{h=1}^{\infty} \beta_{2h}^{-\frac{1}{2h}} = \infty$.

If (M1), (M4) and (C) hold, then the ESD of A_p converges a.s. to the distribution F determined uniquely by the moments $\{\beta_h\}$. The condition (M1) is same as the convergence of $(\text{Span}(A_p), \phi_p)$.

Another widely used method to establish the LSD that is relevant to us is the Stieltjes transformation method. The Stieltjes transformation for any finite measure μ on the real line equals

$$m_\mu(z) = \int \frac{1}{x - z} \mu(dx), \quad z \in \mathbb{C}^+ := \{x + iy : x \in \mathbb{R}, y > 0\}. \tag{1.4}$$

Pointwise convergence of Stieltjes transforms to a Stieltjes transform implies the convergence of the corresponding distributions. In random matrix theory this convergence is often proved by linking the Stieltjes transform of the ESD to the resolvent and showing convergence by martingale convergence methods.

We now list some results on convergence of $(\mathcal{M}_p, \phi_p)$ when the generator $\mathbb{M}_p$ is singleton, and on LSD of square self-adjoint random matrices that are relevant to the present work.

(A) Let $W = ((\omega_{ij}))$ be a Wigner matrix (a square symmetric matrix with independent mean 0 variance 1 entries) of order p. Suppose

$(A0)$ $\sup_{i,j} E|\omega_{ij}|^r < \infty$, $\forall r \geq 1$ **or** for some sequence $\eta_p \downarrow 0$, $|\omega_{ij}| < \eta_p \sqrt{p}$, $\forall i, j$ almost surely.

Let s be a standard semicircle variable. Then $(\text{Span}(p^{-1/2}W), \phi_p)$ converges to $(\text{Span}(s), \varphi)$, where φ is the usual expectation operator. Moreover, the almost sure LSD of $p^{-1/2}W$ is also s.

(B) Let $Z = ((\varepsilon_{ij}))$ be a $p \times n$ matrix where $\{\varepsilon_{ij} : i, j \geq 1\}$ are independently distributed with mean 0 and variance 1 and satisfy Assumption $(A0)$ with $\{\omega_{ij}\}$ replaced by $\{\varepsilon_{ij}\}$. Suppose $p \to \infty$, $p/n \to y \in (0, \infty)$. Let the random variable m_y obey the Marčenko-Pastur law with parameter y. Then $(\text{Span}(n^{-1}ZZ^*), \phi_p)$ converges to $(\text{Span}(m_y), \varphi)$, where φ is the usual expectation operator. Moreover, the a.s. sure LSD of $n^{-1}ZZ^*$ is distributed as m_y.

Incidentally, Assumption $(A0)$ can be relaxed significantly for both (A) and (B). For a detailed account, see Bai and Silverstein [2010], Anderson et al. [2009] and Couillet and Debbah [2011].

Before we describe more LSD results, we need the notion of freeness. Free independence of non-commutative variables is the analogue of the usual independence of commutative (random) variables. Commutative variables (with bounded support) are independent if and only if all mixed cumulants vanish. Similarly, free independence is equivalent to the vanishing of all joint free cumulants. Free cumulants and joint moments are in one-to-one correspondence via the Möbius transformation and its inverse on the POSET of all *non-crossing partitions*. The joint moments of free variables are computable in terms of the moments of individual variables, albeit via complicated algorithms. The notion of freness of variables extends to freeness of sub-algebras in the natural way.

Now suppose we have NCPs $\{(\mathcal{A}_u, \varphi_u)\}_{1 \leq u \leq r}$. Then, analogous to the product space in the commutative case, we can have $(\mathcal{A}, \varphi)$, the *free product* of $\{(\mathcal{A}_u, \varphi_u)\}$ so that the restriction of φ on $\mathcal{A}_u$ is φ_u and $\{\mathcal{A}_u\}$ are considered as free sub-algebras of $\mathcal{A}$. For more details see [Nica and Speicher, 2006].

Now we state further LSD and joint convergence results. As before, for all these LSD results, (A0) can be relaxed considerably.

(C) Let W be as in (A) and A be any $p \times p$ non-negative definite norm bounded matrix whose LSD exists. Then Bai and Zhang [2010] established that the a.s. sure LSD of $p^{-1/2}A^{1/2}WA^{1/2}$ and obtained its Stieltjes transform.

Let s be as in (A) and suppose LSD of A is distributed as the variable a. Further assume s and a are freely independent. Then $(\text{Span}(p^{-1/2}A^{1/2}WA^{1/2}), \phi_p)$ converges to $(\text{Span}(sa), \varphi)$, where φ is the state corresponding to the free product of the $*$-algebras $(\text{Span}(s), \varphi)$ and $(\text{Span}(a), \varphi)$. For more details see Couillet and Debbah [2011].

(D) Suppose Z and A are respectively as in (B) and (C). Suppose $p \to \infty$, $p/n \to y \in (0, \infty)$. Then under appropriate assumptions, Bai and Silverstein [2010] established the Stieltjes transformation of the almost sure LSD of $p^{-1}A^{1/2}ZZ^*A^{1/2}$.

Suppose m_y and a are respectively as in (B) and (C). Further assume m_y and a are freely independent. Then $(\text{Span}\{p^{-1}A^{1/2}ZZ^*A^{1/2}\}, \phi_p)$ converges to $(\text{Span}\{m_y a\}, \varphi)$, where φ is the state corresponding to the free product of the $*$-algebras $(\text{Span}(m_y), \varphi)$ and $(\text{Span}(a), \varphi)$. For more details see Couillet and Debbah [2011].

For (C) and (D), the norm bounded assumption on A can be relaxed by truncating ESD of A. For details see Bai and Zhang [2010] and [Bai and Silverstein, 2010].

(E) Consider the high dimensional vector linear process,

$$X_t = \sum_{j=0}^{q} \psi_j \varepsilon_{t-j} \ \ \forall t, n \geq 1. \tag{1.5}$$

For all t, X_t and $\varepsilon_t = (\varepsilon_{t,1}, \varepsilon_{t,2}, \ldots, \varepsilon_{t,p})^*$ are p-dimensional vectors, ψ_j are $p \times p$ *coefficient matrices* and $\psi_0 = I_p$, the identity matrix of order p. The *sample autocovariance matrix* of order u is defined as

$$\hat{\Gamma}_u = \frac{1}{n}\sum_{t=1}^{n-u} X_t X_{(t+u)}^*, \ \ 0 \leq u \leq n-1. \tag{1.6}$$

Recently researchers have concentrated on the self-adjoint polynomials $\Pi(\hat{\Gamma}_u, \hat{\Gamma}_u^* : u \geq 0)$ in these matrices when $p \to \infty$, $p/n \to y \in (0, \infty)$.

For each $u \geq 1$, define P_u to be an $n \times n$ matrix whose elements are 1 on the upper u-th diagonal and 0 otherwise. $P_0 = I_n$, the identity matrix of order n and $P_{-u} = P_u^*$, $\forall u \geq 1$. Let

$$\Delta_u = \frac{1}{n}\sum_{j,j'=0}^{q} \psi_j Z P_{j-j'+u} Z^* \psi_{j'}^*, \ \forall u \geq 0. \tag{1.7}$$

Suppose $\{\varepsilon_{ij} : i, j \geq 1\}$ are independently distributed with mean 0 and variance 1 and satisfy Assumption ($A0$) with $\{\omega_{ij}\}$ replaced by $\{\varepsilon_{ij}\}$. Bhattacharjee and Bose [2016] proved that, under appropriate assumptions on $\{\psi_j\}$, for any self-adjoint polynomial Π, the almost sure LSD of both $\Pi(\hat{\Gamma}_u, \hat{\Gamma}_u^* : u \geq 0)$ and $\Pi(\Delta_u, \Delta_u^* : u \geq 0)$ exist and they are identical. This LSD can be described as follows.

Let s be a semi-circle variable. Consider two NCP $(\text{Span}(\{a_j, a_j^* : j \geq 0\}), \varphi)$ and $(\text{Span}(\{c_j, c_j^* : j \geq 0\}), \varphi)$ such that for any monomial m,

$$\begin{aligned} \varphi(m(\{a_j, a_j^* : j \geq 0\})) &= y(1+y)^{-1} \lim p^{-1}\text{Tr}(m(\{a_j, a_j^* : j \geq 0\})), \\ \varphi(m(\{c_j, c_j^* : j \geq 0\})) &= y(1+y)^{-1} \lim n^{-1}\text{Tr}(m(\{P_j, P_{-j} : j \geq 0\})). \end{aligned}$$

φ is the state corresponding to the free product of $(\mathrm{Span}(s), \varphi)$, $(\mathrm{Span}(\{a_j, a_j^* : j \geq 0\}), \varphi)$ and $(\mathrm{Span}(\{c_j, c_j^* : j \geq 0\}), \varphi)$. Let $g_u = (1+y)\sum_{j,j'=0}^{q} a_j s c_{j-j'+u} s a_{j'}$, $\forall u \geq 0$. Then the almost sure LSD of $\Pi(\hat{\Gamma}_u, \hat{\Gamma}_u^* : u \geq 0)$ and $\Pi(\Delta_u, \Delta_u^* : u \geq 0)$ is uniquely identified by the moment sequence

$$\begin{aligned}\lim p^{-1} E\mathrm{Tr}(\Pi(\hat{\Gamma}_u, \hat{\Gamma}_u^* : u \geq 0))^k &= \lim p^{-1} E\mathrm{Tr}(\Pi(\Delta_u, \Delta_u^* : u \geq 0))^k \\ &= y^{-1}(1+y)\varphi(\Pi(g_u, g_u^* : u \geq 0))^k, \ \forall k \geq 1. \qquad (1.8)\end{aligned}$$

In particular, the above result implies the results of Pfaffel and Schlemm [2011], Yao [2012], Jin et al. [2014] and Liu et al. [2015].

It is important to note that each of these results involve one random matrix Z and limited choice of the non-random matrices. Collectively, that is a very small sub-collection of variables from $\mathcal{U}_p$. Moreover, almost all of the references above use the Stietljes transform method to derive these results.

It is interesting to ask what happens when we allow several independent random matrices and any $n \times n$ matrices instead of the specific $\{P_u\}$ and, when we also seek the joint limit properties of their polynomials? In other words, when we look at the entire NCP $(\mathcal{U}_p, \phi_p)$ or when we try to find the LSD of any self-adjoint element from it.

Our approach to the solution is borrowed from the literature on the convergence of rectangular matrices. There is a related set up of *rectangular probability spaces* to deal with several rectangular matrices. For example see Benaych-Georges [2009], Benaych-Georges [2010], Benaych-Georges and Nadakuditi [2012], Speicher and Vargas [2012] and Anderson et al. [2009]. In particular, Benaych-Georges [2009] proved that independent, bi-unitarily invariant rectangular matrices, when embedded in an algebra of larger square matrices, are asymptotically free with amalgamation over a commutative finite dimensional sub-algebra.

Taking a cue from Bhattacharjee and Bose [2016] and from Benaych-Georges [2009], we embed the matrices $\{Z_u\}$, $\{B_{2i-1}\}$ and $\{B_{2i}\}$ into appropriate $(n+p)\times(n+p)$ matrices. The $p \times n$ matrix Z is not necessarily bi-unitarily invariant and (unlike Benaych-Georges [2009]) Z and Z^* are embedded into the same Wigner matrix of order $(n+p)$.

Thus, let

$$W_u = \begin{pmatrix} W_{p\times p}^{(1u)} & Z_u \\ Z_u^* & W_{n\times n}^{(2u)} \end{pmatrix}, \qquad (1.9)$$

where $\{W^{(iu)} : i = 1, 2, \ u \geq 1\}$ are independent Wigner matrices and are independent of $\{Z_u\}$.

For any matrices B and D of order p and n respectively, let $\bar{B}$ and $\underline{D}$ of order $(n+p)$ be the matrices

$$\bar{B} = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}, \ \underline{D} = \begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix}. \qquad (1.10)$$

It is easy to see that

$$\bar{\mathbb{P}}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} = \prod_{i=1}^{k_l}\left(n^{-1}\bar{A}_{l,2i-1}W_{u_{l,i}}\underline{A}_{l,2i}W_{u_{l,i}}^*\right)\bar{A}_{l,2k_l+1}. \qquad (1.11)$$

Note that the right side of (1.11) is a polynomial in Wigner and deterministic matrices.

Consider the sequence of NCP $(\bar{\mathcal{U}}_p, \phi_{n+p})$, where

$$\bar{\mathcal{U}}_p = \mathrm{Span}\left(\bar{\mathbb{P}}_{l,(u_{l,1},\ldots,u_{l,k_l})} : \ l, k_l \geq 1\right). \qquad (1.12)$$

Note that $\bar{\mathcal{U}}_p$ also forms a $*$-algebra. Convergence of $(\bar{\mathcal{U}}_p, \phi_{n+p})$ is easy to describe by using asymptotic freeness of several independent Wigner and deterministic matrices (for details see Lemma 2.1). Then we express the limit of $(\mathcal{U}_p, \phi_p)$ in terms of the limit of $(\bar{\mathcal{U}}_p, \phi_{n+p})$.

In Section 2.3, we outline the idea behind the limit. Theorem 3.1 states the convergence of $(\bar{\mathcal{U}}_p, \phi_{n+p})$ and $(\mathcal{U}_p, \phi_p)$. The limiting NCP can be expressed in terms of some free variables.

Theorem 4.1 states that the LSD of any symmetric polynomial in $\{\mathbb{P}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})}\}$ exists and the limit can be expressed in terms of some freely independent variables.

Incidentally, as stated earlier, most of the existing LSD results are obtained using the Stieltjes transformation method. It is not clear how this method could be used for any arbitrary symmetric polynomial. In Theorem 4.1, we fall back on the moment method to derive LSD of such polynomials. To link these two methods, in Theorem 4.2, we derive the Stieltjes transformation of the LSD of a large sub-class of matrices.

All the existing LSD results (A)–(E) mentioned earlier, follow from Theorems 4.1 and 4.2. In Example 4.2, we provide some LSD results, which do not seem to be known in the literature.

Our results can be used to obtain the LSD of polynomials in sample autocovariance matrices for more than one independent moving average processes. That in turn could be used for statistical applications. We shall deal with such applications elsewhere.

2 Preliminaries

2.1 Assumptions

We first list all the assumptions that are required.

Let $Z_u = ((\varepsilon_{u,i,j}))_{p\times n},\ 1 \le u \le U$ be $p \times n$ random matrices, where $\{\varepsilon_{u,i,j} : u, i, j \ge 1\}$ are independently distributed with $E(\varepsilon_{u,i,j}) = 0$, $E|\varepsilon_{u,i,j}|^2 = 1$ and $\sup_{u,i,j} E|\varepsilon_{u,i,j}|^4 < \infty$. We assume that

(A1) For some $\eta > 0,\ \delta \in (0, 2]$,

$$P(|\varepsilon_{u,i,j}| < \eta p^{\frac{1}{2+\delta}}) = 1, 1 \le i \le 2n,\ 1 \le j \le p$$

or, $\sup_{u,i,j} E|\varepsilon_{u,i,j}|^k < c_k < \infty\ \forall k$.

If there is only one u i.e., if $U = 1$, we will write $\varepsilon_{i,j}$ and Z respectively for $\varepsilon_{1,i,j}$ and Z_1. Assumption $(A1)$ will be weakened later for some corollaries and applications by means of truncation techniques.

Now we move to the assumptions on the deterministic matrices $\{B_i\}$. A square matrix A of order p is said to be *norm bounded* if

$$\sup_p \|A\|_2 = \sup_p \sqrt{\text{largest eigenvalue of } A^*A} < \infty.$$

(A2) $\{B_{2i-1} : 1 \le i \le K\}$ are norm bounded $p \times p$ matrices and $(\text{Span}(B_{2i-1}, B^*_{2i-1} : 1 \le i \le K), \phi_p)$ converges.

(A3) $\{B_{2i} : 1 \le i \le L\}$ are norm bounded $n \times n$ matrices and $(\text{Span}(B_{2i}, B^*_{2i} : 1 \le i \le L), \phi_n)$ converges.

Note that we do not assume the joint convergence of $\{B_i : i \ge 1\}$.

2.2 A useful freeness result

The following well known result records the joint convergence of several Wigner and deterministic matrices. This will be needed to prove the convergence of $(\mathcal{U}_p, \phi_p)$ and identify the limits. For (a) and (b) see Anderson et al. [2009]. (c) follows from (a), (b) and Theorem 11.12, Page 180 of Nica and Speicher [2006]. (d) is immediate from (a), (b) and (c).

Lemma 2.1. *Let* $W_p^{(1)}, W_p^{(2)}, \ldots, W_p^{(r)}$ *be* r *independent Wigner matrices of order* p *such that each matrix individually satisfies (A0). Let* $D_p^{(1)}, D_p^{(2)}, \ldots, D_p^{(2q)}$ *be* $2q$ *constant matrices of order* p *with bounded norm such that, for* $\epsilon = 0, 1$, $(\mathit{Span}\{D_p^{(2i-\epsilon)}, D_p^{(2i-\epsilon)*} : 1 \leq i \leq q\}, \phi_p)$ *converges. Then the following statements hold. As* $p \to \infty$,

(a) $p^{-1/2}W_p^{(1)}, p^{-1/2}W_p^{(2)}, \ldots, p^{-1/2}W_p^{(r)}$ *are asymptotically free.*

(b) For $\epsilon = 0$ *or* 1, *the collections* $\{p^{-1/2}W_p^{(i)}\}$ *and* $\{D_p^{(2i-\epsilon)}, D_p^{(2i-\epsilon)*}\}$ *are asymptotically free.*

(c) The collections $\{p^{-1}W_p^{(i)}D_p^{(2j)}W_p^{(i)}, p^{-1}W_p^{(i)}D_p^{(2j)*}W_p^{(i)} : i, j \geq 1\}$ *and* $\{D_p^{(2i-1)}, D_p^{(2i-1)*}\}$ *are asymptotically free.*

(d) Let $\epsilon_i = 1, *, \ \forall 1 \leq i \leq 2k$. *To compute* $\lim p^{-1}E\mathrm{Tr}(p^{-k}\prod_{i=1}^{K} D_p^{(2i-1)\epsilon_{2i-1}}W_p^{(i)}D_p^{(2i)\epsilon_{2i}}W_p^{(i)})$ *one can assume that the collections* $\{W_p^{(i)}\}$, $\{D_p^{(2i-1)}, D_p^{(2i-1)*}\}$ *and* $\{D_p^{(2i)}, D_p^{(2i)*}\}$ *are asymptotically free.*

2.3 Idea behind the limit of $(\mathcal{U}_p, \phi_p)$

To see how freeness comes into the picture and hence how it motivates the limiting NCP of $(\mathcal{U}_p, \phi_p)$, let us focus on a particular element of $\mathcal{U}_p$:

$$P \quad = \quad n^{-1}(B_1Z_1B_2Z_1^*B_3 + B_5Z_1B_4Z_1^*B_7 + B_9Z_1B_6Z_1^*B_{11} + B_{13}Z_1B_8Z_1^*B_{15}).$$

For the purpose of this illustration, assume that P is self-adjoint. For illustration we consider only a self-adjoint polynomial. Similar idea works for non-self-adjoint polynomials also. Our primary goal is to show that for all $r \geq 1$, $\lim \phi_p(P^r) = \lim p^{-1}E\mathrm{Tr}(P^r)$ exists.

Note that for any integer r, if the right/left side limits below exist, then

$$\lim(n+p)^{-1}\mathrm{Tr}(\bar{B}^r) \quad = \quad y(1+y)^{-1}\lim p^{-1}\mathrm{Tr}(B^r), \tag{2.1}$$

$$\lim(n+p)^{-1}\mathrm{Tr}(\underline{D}^r) \quad = \quad (1+y)^{-1}\lim n^{-1}\mathrm{Tr}(D^r), \text{ and} \tag{2.2}$$

$$\lim p^{-1}\mathrm{Tr}(P^r) \quad = \quad y^{-1}(1+y)\lim(n+p)^{-1}\mathrm{Tr}(\bar{P}^r). \tag{2.3}$$

Thus it is enough to deal with $\bar{P}$. Note that

$$n\bar{P} = \bar{B}_1W_1\underline{B}_2W_1\bar{B}_3 + \bar{B}_5W_1\underline{B}_4W_1\bar{B}_7 + \bar{B}_9W_1\underline{B}_6W_1\bar{B}_{11} + \bar{B}_{13}W_1\underline{B}_8W_1\bar{B}_{15}.$$

Thus $\bar{P}^r$ involves polynomials in these enlarged matrices. So it is a question of computing the limiting trace of such polynomials. Note that by (A2) and (A3), as $p, n(p) \to \infty$, $(\mathrm{Span}\{\bar{B}_{2i-1}, \bar{B}_{2i-1}^* : 1 \leq i \leq 8\}, \phi_{n+p})$ and $(\mathrm{Span}\{\underline{B}_{2i}, \underline{B}_{2i}^* : 1 \leq i \leq 4\}, \phi_{n+p})$ converge respectively to $(\mathrm{Span}\{\bar{b}_{2i-1}, \bar{b}_{2i-1}^* : 1 \leq i \leq 8\}, \varphi_1)$ and $(\mathrm{Span}\{\underline{b}_{2i}, \underline{b}_{2i}^* : 1 \leq i \leq 4\}, \varphi_2)$, say. If $\{\varepsilon_{1,i,j}\}$ satisfies (A1), then $(\mathrm{Span}\{(n+p)^{-1/2}W_1\}, \phi_{n+p})$ converges to $(\mathrm{Span}\{s\}, \varphi_3)$, say, where s is a standard semi-circle variable. Finally, by Lemma 2.1 (d) and for the polynomial $\bar{P}$, *in the limit*, the matrices $(n+p)^{-1/2}W_1$, $\{\bar{B}_{2i-1}, \bar{B}^*_{2i-1} : 1 \leq i \leq 8\}$ and $\{\underline{B}_{2i}, \underline{B}^*_{2i} : 1 \leq i \leq 4\}$ are free variables. Therefore, $s, \{\bar{b}_{2i-1}, \bar{b}^*_{2i-1} : \ 1 \leq i \leq 8\}$ and $\{\underline{b}_{2i}, \underline{b}^*_{2i} : \ 1 \leq i \leq 4\}$ are free in some NCP $(\mathcal{A}, \varphi)$.

Thus using the above observations in conjunction with equations (2.1)- (2.3), we can conclude that $\lim \phi_{n+p}(\bar{P}^r)$ and $\lim \phi_p(P^r)$ exist and

$$\begin{aligned}
\lim \phi_{n+p}(\bar{P}^r) &= \lim \Big(\frac{n+p}{n}\Big)^r \Big(\phi_{n+p}(\sum_{i=1}^{4} \bar{B}_{4i-3} \frac{W_1}{\sqrt{n+p}} \underline{B}_{2i} \frac{W_1}{\sqrt{n+p}} \bar{B}_{4i-1})^r\Big) \\
&= \varphi\Big((1+y) \sum_{i=1}^{4} \bar{b}_{4i-3} s \underline{b}_{2i} s \bar{b}_{4i-1}\Big)^r \text{ and} \qquad (2.4) \\
\lim \phi_p(P^r) &= y^{-1}(1+y)\varphi\Big((1+y) \sum_{i=1}^{4} \bar{b}_{4i-3} s \underline{b}_{2i} s \bar{b}_{4i-1}\Big)^r. \qquad (2.5)
\end{aligned}$$

The right side of (2.5), involving free variables, are then the limit moments of P.

This is the idea we implement for the general matrices (1.2). To embed $\{Z_u\}$, we need independent Wigner matrices $\{W_u\}$ and by Lemma 2.1 (a), these yield U many free semi-circle variables. The limits can then be expressed in terms of polynomials in the free semi-circle variables, and the limits of $(\text{Span}\{\bar{B}_{2i-1}, \bar{B}^*_{2i-1} : 1 \leq i \leq 8\}, \phi_{n+p})$ and $(\text{Span}\{\underline{B}_{2i}, \underline{B}^*_{2i}\}, \phi_{n+p})$, where the two limit collections are free.

3 NCP convergence result

Let $(\mathcal{S} = \text{Span}\{s_u : u \geq 1\}, \varphi_s)$ be an NCP of free standard semi-circular variables $\{s_u\}$. Define

$$\begin{aligned}
(\mathcal{A}_{\text{odd}} = \text{Span}\{b_{2i-1}, b^*_{2i-1} : 1 \leq i \leq K\}, \varphi_{\text{odd}}) &= \lim(\text{Span}\{B_{2i-1}, B^*_{2i-1} : 1 \leq i \leq K\}, \phi_p), \qquad (3.1) \\
(\bar{\mathcal{A}}_{\text{odd}} = \text{Span}\{\bar{b}_{2i-1}, \bar{b}^*_{2i-1} : 1 \leq i \leq K\}, \bar{\varphi}_{\text{odd}}) &= \lim(\text{Span}\{\bar{B}_{2i-1}, \bar{B}^*_{2i-1} : 1 \leq i \leq K\}, \phi_{n+p}), \qquad (3.2) \\
(\mathcal{A}_{\text{even}} = \text{Span}\{b_{2i}, b^*_{2i} : 1 \leq i \leq L\}, \varphi_{\text{even}}) &= \lim(\text{Span}\{B_{2i}, B^*_{2i} : 1 \leq i \leq L\}, \phi_n), \qquad (3.3) \\
(\bar{\mathcal{A}}_{\text{even}} = \text{Span}\{\underline{b}_{2i}, \underline{b}^*_{2i} : 1 \leq i \leq L\}, \bar{\varphi}_{\text{even}}) &= (\text{Span}\{\underline{B}_{2i}, \underline{B}^*_{2i} : 1 \leq i \leq L\}, \phi_{n+p}). \qquad (3.4)
\end{aligned}$$

The relation between φ and $\bar{\varphi}$ is as follows. For any polynomial Π,

$$\begin{aligned}
\varphi_{\text{odd}}(\Pi(b_{2i-1}, b^*_{2i-1} : 1 \leq i \leq K)) &= y^{-1}(1+y)\bar{\varphi}_{\text{odd}}(\Pi(\bar{b}_{2i-1}, \bar{b}^*_{2i-1} : 1 \leq i \leq K)), \qquad (3.5) \\
\varphi_{\text{even}}(\Pi(b_{2i}, b^*_{2i} : 1 \leq i \leq L)) &= (1+y)\bar{\varphi}_{\text{even}}(\Pi(\underline{b}_{2i}, \underline{b}^*_{2i} : 1 \leq i \leq L)). \qquad (3.6)
\end{aligned}$$

Let

$$(\mathcal{A}, \bar{\varphi}) = \text{free product of } (\mathcal{S}, \varphi_s), (\bar{\mathcal{A}}_{\text{odd}}, \bar{\varphi}_{\text{odd}}) \text{ and } (\bar{\mathcal{A}}_{\text{even}}, \bar{\varphi}_{\text{even}}). \qquad (3.7)$$

Consider the sub-algebra $\bar{\mathcal{U}}$ of $\mathcal{A}$ as

$$\bar{\mathcal{U}} = \text{Span}\Big(\bar{p}_{l,(u_{l,1}, u_{l,2}, \ldots, u_{l,k_l})} := (1+y)^{k_l} \prod_{i=1}^{k_l} (\bar{a}_{l,2i-1} s_{u_{l,i}} \underline{a}_{l,2i} s_{u_{l,i}}) \bar{a}_{l,2k_l+1} : l \geq 1\Big) \qquad (3.8)$$

where $\bar{a}_{l,2i-1} \in \{\bar{b}_{2i-1}, \bar{b}^*_{2i-1}\}$, $\underline{a}_{l,2i} \in \{\underline{b}_{2i}, \underline{b}^*_{2i}\}$ and $s_{u_{l,i}} \in \{s_u\}$. Note that $\bar{\mathcal{U}}$ forms a $*$-algebra.

Then we have the following Theorem.

Theorem 3.1. *Suppose Assumptions* $(A1)-(A3)$ *hold and* $p, n(p) \to \infty$, $p/n \to y > 0$. *Then*
(a) $(\bar{\mathcal{U}}_p, \phi_{n+p}) \to (\bar{\mathcal{U}}, \bar{\varphi})$, *and*
(b) for any polynomial Π,

$$\lim \phi_p(\Pi(\mathbb{P}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1)) = \frac{1+y}{y}\bar{\varphi}(\Pi(\bar{p}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1)). \tag{3.9}$$

Hence, $(\mathcal{U}_p, \phi_p)$ *converges to* $(\mathcal{U} := \mathrm{Span}(p_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1), \varphi)$, *say, where*

$$\varphi(\Pi(p_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1)) = \frac{1+y}{y}\bar{\varphi}(\Pi(\bar{p}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1)).$$

Proof. (a) Note that, it is enough to show that for any polynomial Π,

$$\lim \phi_{n+p}(\Pi(\bar{\mathbb{P}}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1)) \quad = \quad \bar{\varphi}(\Pi(\bar{p}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1)). \tag{3.10}$$

Now,

$$\Pi(\bar{\mathbb{P}}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1) = \Pi\Big(\prod_{i=1}^{k_l}\Big(n^{-1}\bar{A}_{l,2i-1}W_{u_{l,i}}\underline{A}_{l,2i}W^*_{u_{l,i}}\Big)\bar{A}_{l,2k_l+1} : l \geq 1\Big). \tag{3.11}$$

By using (3.2), (3.4) and Lemma 2.1 (a), (d), the NCP $(\mathrm{Span}(\{\bar{B}_{2i-1}, \bar{B}^*_{2i-1} : i \geq 1\}, \phi_{n+p})$, $(\mathrm{Span}\{\underline{B}_{2i}, \underline{B}^*_{2i} : i \geq 1\}, \phi_{n+p})$ and $(\mathrm{Span}\{(n+p)^{-1/2}W_u : 1 \leq u \leq U\}, \phi_{n+p})$ respectively converge to $(\bar{\mathcal{A}}_{\mathrm{odd}}, \bar{\varphi}_{\mathrm{odd}})$, $(\bar{\mathcal{A}}_{\mathrm{even}}, \bar{\varphi}_{\mathrm{even}})$ and $(\mathcal{S}, \varphi_s)$ and they are asymptotically free. Note that $\{\bar{B}_{2i-1}, \bar{B}^*_{2i-1}\}$ and $\{\underline{B}_{2i}, \underline{B}^*_{2i}\}$ are not in general asymptotically free. They are asymptotically free in polynomials where $\{B_{2i-1}, B^*_{2i-1}\}$ and $\{B_{2i}, B^*_{2i}\}$ are respectively enclosed within (Z^*, Z) and (Z, Z^*). Therefore, observing (3.7), (3.10) holds.

(b) Note that for any polynomial Π,

$$\begin{aligned}
\lim \phi_p(\Pi(\{\mathbb{P}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1\})) &= \lim \frac{n+p}{p}\phi_{n+p}(\Pi(\{\bar{\mathbb{P}}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1\})) \\
&= y^{-1}(1+y)\lim \phi_{n+p}(\Pi(\{\bar{\mathbb{P}}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1\})) \\
&= y^{-1}(1+y)\bar{\varphi}(\Pi(\{\bar{p}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1\})) \\
&= \varphi(\Pi(\{p_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1\})), \text{ (say).}
\end{aligned}$$

This completes the proof of Theorem 3.1 (b). □

4 LSD of symmetric polynomials in $\{\mathbb{P}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})}\}$

The following Theorem guarantees the existence of the LSD of any symmetric polynomial in $\{\mathbb{P}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l}}\}$.

Theorem 4.1. *Suppose (A1)-(A3) hold and* $p, n(p) \to \infty$, $p/n \to y > 0$. *Then the LSD of any symmetric polynomial* $\Pi(\mathbb{P}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1)$ *exists almost surely and it is uniquely determined by the (usual) moment sequence*

$$\begin{aligned}
\lim \phi_p(\Pi(\mathbb{P}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1))^k &= \varphi(\Pi(p_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1))^k \\
&= y^{-1}(1+y)\bar{\varphi}(\Pi(\bar{p}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1))^k, \ \forall k \geq 1.
\end{aligned}$$

Proof. We need to establish (M1), (M4) and (C). The (M1) condition is nothing but (3.9) in Theorem 3.1 (b). Now we shall establish (M4) and (C).

Proof of (M4). We need the following lemma, whose proof is deferred to Section 5.1.

Lemma 4.1. *Suppose (A1)-(A3) hold and $p, n(p) \to \infty$, $p/n \to y > 0$. Let $\mathbb{P}_u \in \operatorname{Span}\{\mathbb{P}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})}\}$, $u \geq 0$. Let for $1 \leq i \leq T$, $m_i(\mathbb{P}_u, \mathbb{P}_u^* : u \geq 0)$ be polynomials. Let*

$$\mathcal{P}_i = \operatorname{Tr}(m_i(\mathbb{P}_u, \mathbb{P}_u^* : u \geq 0)) \text{ and } \mathcal{P}_i^0 = E\mathcal{P}_i.$$

For $d \geq 1$, define

$$\mathcal{S}_d = \text{set of all pair partitions } \{(i_1, i_2), (i_3, i_4), \ldots, (i_{2d-1}, i_{2d})\} \text{ of } \{1, 2, \ldots, 2d\}.$$

Then, for all $d \geq 1$,

$$\lim E\left[\Pi_{i=1}^T \left(\mathcal{P}_i - \mathcal{P}_i^0\right)\right] = \begin{cases} 0 \text{ if } T = 2d-1, \\ \sum_{\mathcal{S}_d} \prod_{k=1}^d \lim E[(\mathcal{P}_{i_{2k-1}} - \mathcal{P}^0_{i_{2k-1}})(\mathcal{P}_{i_{2k}} - \mathcal{P}^0_{i_{2k}})], \text{ if } T = 2d. \end{cases}$$

For any polynomial $\Pi(\mathbb{P}_u, \mathbb{P}_u^* :\ u \geq 0)$, taking $T = 4$ and $\mathcal{P}_i = \operatorname{Tr}\left(\Pi(\mathbb{P}_u, \mathbb{P}_u^* :\ u \geq 0)\right)^h$ in Lemma 4.1, we have,

$$E\left[\frac{1}{p}\operatorname{Tr}\left(\Pi(\mathbb{P}_u, \mathbb{P}_u^* :\ u \geq 0)\right)^h - E\left(\frac{1}{p}\operatorname{Tr}\left(\Pi(\mathbb{P}_u, \mathbb{P}_u^* :\ u \geq 0)\right)^h\right)\right]^4 = O(p^{-4}) = O(n^{-4})$$

and hence (M4) is established.

Proof of (C). We have to show, for any symmetric polynomial Π,

$$\sum_{k=1}^{\infty} \left(y^{-1}(1+y)\bar{\varphi}(\Pi(\bar{p}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1))^{2k}\right)^{-1/2k} = \infty. \tag{4.1}$$

Now note that

$$\sum_{k=1}^{\infty} \left(y^{-1}(1+y)\bar{\varphi}(\Pi(\bar{p}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1))^{2k}\right)^{-1/2k} \geq \frac{y}{1+y}\sum_{k=1}^{\infty} \left(\bar{\varphi}(\Pi(\bar{p}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1))^{2k}\right)^{-1/2k}.$$

Therefore, to prove (4.1), it is enough to show that

$$\sum_{k=1}^{\infty} \left(\bar{\varphi}(\Pi(\bar{p}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1))^{2k}\right)^{-1/2k} = \infty. \tag{4.2}$$

Moreover, to prove (4.2), it is enough to show that for some $C > 0$,

$$\bar{\varphi}(\Pi(\bar{p}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1))^{2k} \leq C^{2k},\ \forall k \geq 1. \tag{4.3}$$

The following lemma is useful in this proof. Its proof is given in Section 5.3.

Lemma 4.2. *Let s be a standard semicircle variable. For all $\{\bar{a}_{2i-1}\} \in \{\bar{b}_{2i-1}, \bar{b}^*_{2i-1}\}$, $\{\bar{a}_{2i}\} \in \{\bar{b}_{2i}, \bar{b}^*_{2i}\}$, $h \geq 1$ and for some $C_1 > 0$, we have $|\bar{\varphi}(\bar{a}_1 s \underline{a}_2 s \bar{a}_3 \ldots \underline{a}_{2h} s)| \leq C_1^{2h}$.*

Now by (3.8), note that

$$\Pi(\bar{p}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1) = \sum_{i=1}^{T} g_i, \quad \text{where } \; g_i = \bar{a}_{1,i} s \underline{a}_{2,i} s \cdots \underline{a}_{2l_i,i} s, \;\; \forall i \geq 1, \tag{4.4}$$

$\bar{a}_{2j-1,i} \in \{\bar{b}_{2i-1}, \bar{b}^*_{2i-1} : i \geq 1\}$ and $\underline{a}_{2j,i} \in \{\underline{b}_{2i}, \underline{b}^*_{2i} : i \geq 1\}$. Now, by Lemma 4.2, there is a $C_1, C_2 > 0$ such that

$$\begin{aligned} \bar{\varphi}(\Pi(\bar{p}_{l,(u_{l,1},u_{l,2},\ldots,u_{l,k_l})} : l \geq 1))^{2k} = \bar{\varphi}\left(\sum_{i=1}^{T} g_i\right)^{2k} &= \sum_{1 \leq i_1, i_2, \ldots, i_{2k} \leq T} \bar{\varphi}(g_{i_1} g_{i_2} \ldots g_{i_{2k}}) \\ &\leq \sum_{1 \leq i_1, i_2, \ldots, i_{2k} \leq T} |\bar{\varphi}(g_{i_1} g_{i_2} \ldots g_{i_{2k}})| \leq C_1^{2\sum_{j=1}^{2k} l_{i_j}} T^{2k} \leq C_2^{2k}. \end{aligned}$$

Hence, (4.3) is proved and (C) follows. This completes the proof of Theorem 4.1. □

4.1 Stieltjes transform

Most of the existing LSD discussed in Section 1 are in terms of the Stieltjes transform. Therefore to show how these results follow from Theorem 4.1, we need to study the Stieltjes transform of our LSD. *If needed, we enlarge the collections $\{B_{2i-1}\}$ and $\{B_{2i}\}$ so that they are closed under $*$ operation. All results proved so far continue to remain valid.* Consider the polynomials $\Delta \in \mathcal{U}_p$ of the form

$$\Delta = \frac{1}{n} \sum_{i=1}^{q} B_{4i-3} Z B_{2i} Z^* B_{4i-1}. \tag{4.5}$$

We assume appropriate conditions on $\{B_i\}$ so that Δ is symmetric.

Note that all the existing LSD results ((A)-(E)) are for random matrices which are special cases of Δ. Moreover, the matrices $\{\Delta_u\}$, which are defined in (1.7) and which will approximate $\{\hat{\Gamma}_u\}$, are also special cases of Δ. By Theorem 4.1, under (A1)-(A3), the almost sure LSD of Δ exists and it is characterized by the moment sequence

$$\lim \phi_p(\Delta)^k = \lim \frac{1}{p} E\mathrm{Tr}(\Delta)^k = \frac{1+y}{y} \bar{\varphi}(\bar{\delta})^k, \; \forall k \geq 1, \tag{4.6}$$

where

$$\bar{\delta} = (1+y) \sum_{i=1}^{q} \bar{b}_{4i-3} s \underline{b}_{2i} s \bar{b}_{4i-1}. \tag{4.7}$$

Recall that $\{\bar{b}_{2i-1}\}$ and $\{\underline{b}_{2i}\}$ are respectively limits of $\{\bar{B}_{2i-1}\}$ and $\{\underline{B}_{2i}\}$. Moreover, s, $\{\bar{b}_{2i-1}\}$ and $\{\underline{b}_{2i}\}$ are free (by Theorem 4.1, as far as computing limits of polynomials of the form (4.5) is concerned).

Let ν be a probability measure on $\mathbb{R}$ with $\nu(-K, K) = 1$ for some $K > 0$. Then we have the following formal power series expansion of the Stieltjes transformation $m(z)$ for $|z| > K$ and $z \in \mathbb{C}^+$,

$$m_\nu(z) = -\frac{1}{z} E_\nu\left(\frac{1}{1 - \frac{X}{z}}\right) = -\frac{1}{z} - \frac{E_\nu(X)}{z^2} - \frac{E_\nu(X^2)}{z^3} - \cdots . \tag{4.8}$$

This relation is crucial in linking the moment and the Stieltjes transform approach. Since $m_\nu(z)$ is analytic for $z \in \mathbb{C}^+$, in principle it suffices to identify it only for large enough $z \in \mathbb{C}^+$.

Note that by (A2), (A3) and Lemma 4.2, there is $C > 0$ such that $|\bar{\phi}(\bar{\delta}^k)| \leq C^k, \ \forall k$. Hence, there is a unique probability measure on $\mathbb{R}$, say $\bar{\mu}$, characterized by the moment sequence $\{\bar{\phi}(\bar{\delta}^k)\}$. Let μ be the probability measure on $\mathbb{R}$ corresponding to the LSD of Δ. Note that by (4.6),

$$\int_{\mathbb{R}} x^k d\mu = \frac{1+y}{y} \int_{\mathbb{R}} x^k d\bar{\mu}, \ \forall k \geq 1. \tag{4.9}$$

Let $m_{\bar{\mu}}(z)$ and $m_\mu(z)$ be respectively the Stieltjes transforms of $\bar{\mu}$ and μ. We first describe $m_{\bar{\mu}}(z)$. Then it is easy to express $m_\mu(z)$ in terms of $m_{\bar{\mu}}(z)$.

To describe $m_{\bar{\mu}}(z)$, we write infinite sums of the form $\sum_{1 \leq i_1, i_2, \ldots, i_k < \infty} a_{i_1} a_{i_2} \ldots a_{i_k}$ in the sense that

$$\varphi\Big(\sum_{1 \leq i_1, i_2, \ldots, i_k < \infty} a_{i_1} a_{i_2} \ldots a_{i_k} \Big) = \sum_{1 \leq i_1, i_2, \ldots, i_k < \infty} \varphi(a_{i_1} a_{i_2} \ldots a_{i_k}),$$

whenever

$$\sum_{1 \leq i_1, i_2, \ldots, i_k < \infty} |\varphi(a_{i_1} a_{i_2} \ldots a_{i_k})| < \infty. \tag{4.10}$$

Moreover, we write $(1-a)^{-1} := \sum_{i=0}^{\infty} a^i$.

Let,

$$\begin{aligned} d &= \{\bar{b}_{4i-3} : 1 \leq i \leq q\}, \ \ e = \{\bar{b}_{4i-1} : 1 \leq i \leq q\}, \ \text{ and} && (4.11) \\ f &= \{\underline{b}_{2i} : 1 \leq i \leq q\}, \ \ h(d,e,f) = (1+y) \sum_{i=1}^{q} \bar{b}_{4i-3} \underline{b}_{2i} \bar{b}_{4i-1}. && (4.12) \end{aligned}$$

Define

$$R_j(f) = \bar{\varphi}\left(h(d,e,f) \bar{\delta}^{j-1} | f \right) := (1+y) \sum_{i=0}^{q} \bar{\varphi}(\bar{b}_{4i-3} \bar{b}_{4i-1} \bar{\delta}^{j-1}) \underline{b}_{2i}. \tag{4.13}$$

Note that $R_j(f) \in \mathcal{A}, \ \forall j \geq 1$. Let,

$$\begin{aligned} K(z,f) &= \sum_{i=1}^{\infty} z^{-i} R_i(f) && (4.14) \\ B(d,e,z) &= \bar{\varphi}\left(h(d,e,f)(1+K(z,f))^{-1} | d, e \right), && (4.15) \\ &:= (1+y) \sum_{i=1}^{q} \bar{\varphi}\left(\underline{b}_{2i} (1+yK(z,f))^{-1} \right) \bar{b}_{4i-3} \bar{b}_{4i-1} && (4.16) \\ &= (1+y) \sum_{j=0}^{\infty} \sum_{i=1}^{q} \bar{\varphi}(\underline{b}_{2i} (-K(z,f))^{i}) \bar{b}_{4i-3} \bar{b}_{4i-1}. && (4.17) \\ G(d,e,z) &= (B(d,e,z)-z)^{-1} = z^{-1} \sum_{i=0}^{\infty} z^{-i} (B(d,e,z))^{-i}. && (4.18) \end{aligned}$$

The following lemma guarantees the existence of $K(z,f), B(d,e,z)$ and $G(d,e,z)$. Its proof is given in Section 5.4.

Lemma 4.3. *$K(z, f)$, $B(d, e, z)$ and $G(d, e, z)$ exist for sufficiently large $|z|$, in the sense of (4.10).*

The following Theorem provides $m_{\bar{\mu}}(z)$ and $m_{\mu}(z)$ and is proved in Section 5.5.

Theorem 4.2. *Suppose (A1)-(A3) hold and $p, n(p) \to \infty$, $p/n \to y > 0$.*

(a) Then for $z \in \mathbb{C}^+$, $|z|$ large, $m_{\bar{\mu}}(z)$ is given by

$$m_{\bar{\mu}}(z) \;=\; \bar{\varphi}(G(d, e, z)), \tag{4.19}$$

where $G(d, e, z)$ satisfies (4.14), (4.15) and (4.18). Moreover, $K(z, f)$ in (4.14) also satisfies

$$\begin{aligned} K(z, f) \;&=\; \bar{\varphi}(h(d, e, f)(B(d, e, z) - z)^{-1}|f) && (4.20) \\ &:=\; (1 + y) \sum_{i=1}^{q} \bar{\varphi}(\bar{b}_{4i-3}\bar{b}_{4i-1}(B(d, e, z) - z)^{-1})\underline{b}_{2i}. && (4.21) \end{aligned}$$

(b) For $z \in \mathbb{C}^+$, $m_{\mu}(z)$ is given by

$$m_{\bar{\mu}}(z) = \frac{y}{1 + y} m_{\mu}(z) - \frac{1}{1 + y}\frac{1}{z}. \tag{4.22}$$

4.2 Examples

It is easy to see that Theorem 4.1 implies Theorem 3.1 of Bhattacharjee and Bose [2016]. In this section, we shall discuss two examples. First one involves existing LSD results which follow from Theorem 3.1 of Bhattacharjee and Bose [2016] and hence also from Theorems 4.1 and 4.2. In the second example, we provide some apparently new LSD results.

Example 4.1. Recall $\{\Delta_u\}$ defined in (1.7) and $\{\psi_j\}$ matrices in (1.5). Suppose (A1) holds and $p, n(p) \to \infty, p/n \to y \in (0, \infty)$.

(a) Suppose $\psi_j = 0,\ \forall j \geq 1$ and hence $\Delta_u = n^{-1} Z P_u Z^*\ \forall u \geq 0$.

(i) (Marčenko and Pastur [1967]) The almost sure LSD of Δ_0 exists and it is the well Marčenko-Pastur law with parameter y.

(ii) (Bai and Silverstein [2010]) Let A be a $p \times p$ symmetric non-negative definite matrix whose LSD exists. Then the almost sure LSD of $A^{1/2}\Delta_0 A^{1/2}$ exists.

(iii) (Bhattacharjee and Bose [2016]) The almost sure LSD of $(n/p)^2 \Delta_u \Delta_u^*$ exists for all $u \geq 1$ and it is the Bessel$(2, y^{-1})$ law.

(b) (Pfaffel and Schlemm [2011] and Bhattacharjee and Bose [2016]) Suppose $\psi_j = \lambda_j I,\ \forall j \geq 1$. Then for all $u \geq 1$, the almost sure LSD of $\Delta_u + \Delta_u^*$ exists. This LSD is the compound free Poisson distribution (see pages 206 − 208 in Nica and Speicher [2006]) with rate y^{-1} and the jump distribution $y\cos(u\theta)h(\lambda, \theta)$, where $\theta \sim U(0, 2\pi)$ and $h(\lambda, \theta) = |\sum_{j=0}^{q} e^{ij\theta}\lambda_j|$.

(c) (Bhattacharjee and Bose [2016]) If $\{\psi_j\}$ are norm bounded and converge jointly, then almost sure LSD of $\Delta_u + \Delta_u^*$ exists. Liu et al. [2015] had proved this existence under stronger assumption on $\{\psi_j\}$.

In all the above cases, (A1) can be replaced by significantly weaker moment assumption. All the above results hold if we replace $\{\Delta_u\}$ by $\{\hat{\Gamma}_u\}$, the sample autocovariance matrices defined in (1.6). For details see Bhattacharjee and Bose [2016].

We now give some examples of LSD that do not seem to be known in the literature.

Example 4.2. Suppose (A1), (A2), (A3), (3.1) hold and $p, n(p) \to \infty, p/n \to y > 0$. Let,

$$Y_1 = n^{-1} Z_1 B_2 Z_1^*,\ Y_2 = n^{-1} Z_1 B_4 Z_1^*,\ Y_3 = n^{-1} Z_2 B_4 Z_2^*. \tag{4.23}$$

Suppose the standard semi-circle variables $s_1, s_2, \{\bar{b}_1, \bar{b}_2, \bar{b}_3\}$ and $\{\underline{b}_4, \underline{b}_5\}$ are free. Define

$$g_1 = (1+y) s_1 \underline{b}_2 s_1,\ g_2 = (1+y) s_1 \underline{b}_4 s_1,\ g_3 = (1+y) s_2 \underline{b}_2 s_2, \tag{4.24}$$

(i) Suppose B_1, B_3 are $p \times p$ real symmetric matrices and B_2 is an $n \times n$ real symmetric matrix. Then the almost sure LSD of $B_1 Y_1 B_3 Y_1^* B_1$ exists and it is uniquely determined by the moment sequence

$$\lim \phi_p (B_1 Y_1 B_3 Y_1^* B_1)^k = y^{-1}(1+y) \varphi(\bar{b}_1 g_1 \bar{b}_3 g_1 \bar{b}_1)^k\ \forall k \geq 1. \tag{4.25}$$

(ii) The almost sure LSD of $B_1 Y_1 B_3 Y_2 B_5 + B_5^* Y_2^* B_3^* Y_1^* B_1^*$ exists and it is uniquely determined by the moment sequence

$$\lim \phi_p (B_1 Y_1 B_3 Y_2 B_5 + B_5^* Y_2^* B_3^* Y_1^* B_1^*)^k = y^{-1}(1+y) \varphi(\bar{b}_1 g_1 \bar{b}_3 g_2 \bar{b}_5 + \bar{b}_5^* g_2^* \bar{b}_3^* g_1^* \bar{b}_1^*)^k\ \forall k \geq 1. \tag{4.26}$$

(iii) The almost sure LSD of $B_1 Y_1 B_3 Y_3 B_5 + B_5^* Y_3^* B_3^* Y_1^* B_1^*$ exists and it is uniquely determined by the moment sequence

$$\lim \phi_p (B_1 Y_1 B_3 Y_3 B_5 + B_5^* Y_3^* B_3^* Y_1^* B_1^*)^k = y^{-1}(1+y) \varphi(\bar{b}_1 g_1 \bar{b}_3 g_3 \bar{b}_5 + \bar{b}_5^* g_3^* \bar{b}_3^* g_1^* \bar{b}_1^*)^k\ \forall k \geq 1. \tag{4.27}$$

Results like 4.1 (iii) can be used to study two independent moving average processes jointly. We shall deal with such applications elsewhere.

5 Proofs

5.1 Proof of Lemma 4.1

We first prove the result for monomials $\{m_i\}$ in $\{\mathbb{P}_u, \mathbb{P}_u^*\}$ involving *one independent matrix Z*. Once the result is proved for monomials, it is easy to see that it continues to hold for polynomials (see Step 1 below). Moreover it will also be clear that the arguments continue to hold when more than one independent matrices $\{Z_u\}$ are involved. The proof is completed in two steps.

Step 1. First we show that it is enough to prove the lemma for monomials. Consider arbitrary $p \times p$ matrices $\{A_{ik}, C_{ik} : 1 \leq k \leq r_i, 1 \leq i \leq T\} \subset \text{Span}\{B_{2i-1}, B_{2i-1}^*\}$ and $n \times n$ matrices $\{B_{ik} : 1 \leq k \leq r_i, 1 \leq i \leq T\} \subset \text{Span}\{B_{2i}, B_{2i}^*\}$. Define

$$\pi_i = n^{-r_i} \text{Tr}\Big(\prod_{k=1}^{r_i} A_{ik} Z B_{ik} Z^* C_{ik}\Big) \text{ and } \pi_i^0 = E\pi_i,\ 1 \leq i \leq T. \tag{5.1}$$

For all $d \geq 1$, consider the equations

$$\lim E\left[\Pi_{i=1}^{T} \left(\pi_i - \pi_i^0\right)\right] = \begin{cases} 0 \text{ if } T = 2d-1, \\ \sum_{S_d} \prod_{k=1}^{d} \lim E[(\pi_{i_{2k-1}} - \pi_{i_{2k-1}}^0)(\pi_{i_{2k}} - \pi_{i_{2k}}^0)], \text{ if } T = 2d. \end{cases} \tag{5.2}$$

We now prove that (5.2) $\implies$ (4.1) i.e. it is enough to prove Lemma 4.1 for monomials only. In the next step, we prove (5.2).

Note that for some matrices $\{A_{iks}, C_{iks}\} \in \text{Span}\{B_{2i-1}, B_{2i-1}^*\}$ and $\{B_{iks}\} \in \text{Span}\{B_{2i}, B_{2i}^*\}$, we can write

$$\begin{aligned}
\mathcal{P}_i &= \sum_{k=1}^{t_i} \text{Tr}\left(n^{-r_k} \prod_{s=1}^{r_k} A_{iks} Z B_{iks} Z^* C_{iks}\right) = \sum_{k=1}^{t_i} S_{ik}, \text{ (say)}. \\
\mathcal{P}_i^0 &= \sum_{i=1}^{t_i} E(S_{ik}) = \sum_{i=1}^{t_i} S_{ik}^0, \text{ (say)}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\lim E\Big(\prod_{i=1}^{T} (\mathcal{P}_i - \mathcal{P}_i^0)\Big) &= \lim E\Big(\prod_{i=1}^{T} \sum_{k=1}^{t_i} (S_{ik} - S_{ik}^0)\Big) = \lim E\Big(\prod_{i=1}^{T} \sum_{k=1}^{t_i} T_{ik}\Big), \text{ (say)} \\
&= \lim E\Big(\sum_{1 \le k_i \le t_i} \prod_{i=1}^{T} T_{ik_i}\Big) = \sum_{1 \le k_i \le t_i} \lim E\Big(\prod_{i=1}^{T} T_{ik_i}\Big) \\
&= \begin{cases} 0, & \text{if } T = 2d-1 \\ \sum_{1 \le k_i \le t_i} \sum_{\mathcal{S}_d} \prod_{s=1}^{d} \lim E(T_{i_{2s-1} k_{i_{2s-1}}} T_{i_{2s} k_{i_{2s}}}), & \text{if } T = 2d. \end{cases}
\end{aligned}$$

The last equality holds by (5.2). Therefore, (4.1) follows from (5.2) when T is odd.

When T is even, we have

$$\sum_{\mathcal{S}_d} \prod_{s=1}^{d} \sum_{k_{2s-1}, k_{2s}} \lim E(T_{i_{2s-1} k_{i_{2s-1}}} T_{i_{2s} k_{i_{2s}}}) = \sum_{\mathcal{S}_d} \prod_{k=1}^{d} \lim E[(\mathcal{P}_{i_{2k-1}} - \mathcal{P}_{i_{2k-1}}^0)(\mathcal{P}_{i_{2k}} - \mathcal{P}_{i_{2k}}^0)].$$

Therefore, (4.1) follows from (5.2) for all $T \geq 1$.

Step 2. **Proof of (5.2).** Let $A(i, j)$ be the (i, j)-th element of the matrix A. Note that, for all $1 \leq i \leq T$,

$$\begin{aligned}
n^{r_i} \pi_i = \text{Tr}\Big(\prod_{k=1}^{r_i} A_{ik} Z B_{ik} Z^* C_{ik}\Big) &= \sum_{\substack{1 \le u_t^{(i)} \le p,\ 1 \le v_s^{(i)} \le p \\ 1 \le t \le 3r_i,\ 1 \le s \le 2r_i}} \prod_{\substack{k=1 \\ u_{3r_i+1}^{(i)} = u_1^{(i)}}}^{r_i} A_{ik}(u_{3k-2}^{(i)}, u_{3k-1}^{(i)}) \varepsilon_{u_{3k-1}^{(i)}, v_{2k-1}^{(i)}} B_{ik}(v_{2k-1}^{(i)}, v_{2k}^{(i)}) \\
&\qquad \varepsilon_{u_{3k}^{(i)}, v_{2k}^{(i)}} C_{ik}(u_{3k}^{(i)}, u_{3k+1}^{(i)}).
\end{aligned}$$

For fix $1 \leq i \leq T$, we define

$$\mathcal{U}_i = \{(u_{3k+\delta}^{(i)}, v_{2k+\delta}^{(i)}) : 1 \leq k \leq r_i, \delta = -1, 0, u_{3r_i+1}^{(i)} = u_1^{(i)}\}. \tag{5.3}$$

Note that $\mathcal{U}_i$ is the set of all indices attached with ε's in the expansion of π_i given in (5.3). An index $(u_{3k+\delta}^{(i)}, v_{2k+\delta}^{(i)})$ is said to be matched if there is at least one $(k', \delta', i') \neq (k, \delta, i)$ with $(u_{3k+\delta}^{(i)}, v_{2k+\delta}^{(i)}) = (u_{3k'+\delta'}^{(i')}, v_{2k'+\delta'}^{(i')})$. Now note that $E\left[\prod_{i=1}^{T} \left(\pi_i - \pi_i^0\right)\right]$ involves all indices in $\cup_{i=1}^{T} \mathcal{U}_i$ (if we expand $\{\pi_i\}$ using the last equality of (5.3)). As $\{\varepsilon_{i,j}\}$ are independent and have mean 0, all indices in $\cup_{i=1}^{T}\mathcal{U}_i$ need to be matched to guarantee a non-zero contribution. For each $1 \leq i \leq T$, consider the following sets of matched indices.

$$\begin{aligned} B_i &= \text{ set of all matchings where for each } (k,\delta) \text{, there is at least one} \\ & \quad (k',\delta') \neq (k,\delta) \text{ with } (u^{(i)}_{3k+\delta}, v^{(i)}_{2k+\delta}) = (u^{(i)}_{3k'+\delta'}, v^{(i)}_{2k'+\delta'}) \text{ and for } i \neq i', \\ & \quad \text{there is no } (k',\delta',i') \text{ such that } (u^{(i)}_{3k+\delta}, v^{(i)}_{2k+\delta}) = (u^{(i')}_{3k'+\delta'}, v^{(i')}_{2k'+\delta'}) \,. \end{aligned}$$

Consider the disjoint decomposition $\cup_{i=1}^{T+1}C_i$ of all possible matchings of indices in $\cup_{i=1}^{T}\mathcal{U}_i$, where

$$C_1 = B_1,\ C_i = (\cap_{j=1}^{i-1}B_j^c) \cap B_i \ \forall 2 \leq i \leq T,\ C_{T+1} = \cap_{i=1}^{T}B_i^c. \tag{5.4}$$

Let for any set A, E_A be the usual expectation restricting on the set A. We shall first show that

$$E\left[\Pi_{i=1}^{T}\left(\pi_i - \pi_i^0\right)\right] \quad = \quad E_{C_{T+1}}(\Pi_{i=1}^{T}\pi_i). \tag{5.5}$$

For this purpose, we need more analysis for the set C_i. Define

$$\begin{aligned} \mathcal{S}_i &= \text{ set of all matchings of indices in } \mathcal{U}_i \text{, and} \\ \mathcal{S}_{-i} &= \text{ set of all matchings of indices in } \cup_{j\neq i}\mathcal{U}_j \text{ such that for each } 1 \leq j < i \\ & \quad \text{ there is at least one index in } \mathcal{U}_j \text{ which matches with some index} \\ & \quad \text{ in } \mathcal{U}_k,\ k \neq j, i. \end{aligned}$$

Note that

$$C_i = (\cap_{j=1}^{i-1}B_j^c) \cap B_i = \{(\sigma_1 \cup \sigma_2) : \sigma_1 \in \mathcal{S}_i,\ \sigma_2 \in \mathcal{S}_{-i}\}. \tag{5.6}$$

Also note that

$$E(\Pi_{i=1}^{T}(\pi_i - \pi_i^0)) = E_{C_i}((\Pi_{j=1}^{i}\pi_j)\Pi_{j=i+1}^{T}(\pi_j - \pi_j^0)) + \text{ other terms.}$$

Then for all $2 \leq i \leq T$, we have

$$\begin{aligned} & E_{C_i}(\Pi_{j=1}^{i}\pi_j\Pi_{j=i+1}^{T}(\pi_j - \pi_j^0)) = \sum_{\sigma\in C_i} E_\sigma(\Pi_{j=1}^{i}\pi_j\Pi_{j=i+1}^{T}(\pi_j - \pi_j^0)) \qquad (5.7) \\ = & \sum_{\sigma_1\in\mathcal{S}_i,\ \sigma_2\in\mathcal{S}_{-i}} E_{\sigma_1}(\pi_i)E_{\sigma_2}(\Pi_{j=1}^{i-1}\pi_j\Pi_{j=i+1}^{T}(\pi_j - \pi_j^0)) \\ & \qquad [\text{as } C_i \subset B_i \text{ and under } B_i, \{\varepsilon_{u,v} : (u,v) \in \mathcal{U}_i\} \text{ are} \\ & \qquad\qquad \text{independent of } \{\varepsilon_{u,v} : (u,v) \in \cup_{j\neq i}\mathcal{U}_j\}] \\ = & \Big(\sum_{\sigma_1\in\mathcal{S}_i} E_{\sigma_1}(\pi_i)\Big)\Big(\sum_{\sigma_2\in\mathcal{S}_{-i}} E_{\sigma_2}(\Pi_{j=1}^{i-1}\pi_j\Pi_{j=i+1}^{T}(\pi_j - \pi_j^0))\Big) \\ = & \pi_i^0\Big(E_{\cap_{j=1}^{i-1}B_j^c}(\Pi_{j=1}^{i-1}\pi_j\Pi_{j=i+1}^{T}(\pi_j - \pi_j^0))\Big). \end{aligned}$$

Similarly,

$$E_{B_1}(\pi_1\Pi_{i=2}^{T}(\pi_i - \pi_i^0)) \quad = \quad \pi_1^0E(\Pi_{i=2}^{T}(\pi_i - \pi_i^0)). \tag{5.8}$$

Now, the left side of (5.5) equals,

$$E\left[\Pi_{i=1}^{T}\left(\pi_i - \pi_i^0\right)\right] \quad = E_{B_1}(\pi_1\Pi_{i=2}^{T}(\pi_i - \pi_i^0)) + E_{B_1^c}(\pi_1\Pi_{i=2}^{T}(\pi_i - \pi_i^0)) - \pi_1^0E(\Pi_{i=2}^{T}(\pi_i - \pi_i^0))$$

$$\begin{aligned}
&= E_{B_1^c}(\pi_1\Pi_{i=2}^T(\pi_i - \pi_i^0)),\\
&= E_{B_1^c\cap B_2}(\pi_1\pi_2\Pi_{i=3}^T(\pi_i - \pi_i^0)) + E_{B_1^c\cap B_2^c}(\pi_1\pi_2\Pi_{i=3}^T(\pi_i - \pi_i^0))\\
&\qquad\qquad\qquad\qquad\qquad\qquad -\pi_2^0 E_{B_1^c}(\pi_1\Pi_{i=3}^T(\pi_i - \pi_i^0))\\
&= E_{B_1^c\cap B_2^c}(\pi_1\pi_2\Pi_{i=3}^T(\pi_i - \pi_i^0)), \quad \text{(by (5.7), for } T = 2)\\
&\qquad\qquad \vdots\\
&= E_{B_1^c\cap B_2^c\cap\ldots\cap B_T^c}(\Pi_{i=1}^T\pi_i), \text{ by repeated application of (5.7) for } 3 \le i \le T.\\
&= E_{C_{T+1}}(\Pi_{i=1}^T\pi_i).
\end{aligned}$$

Therefore, (5.5) is established. Next we shall analyze the set C_{T+1} and identify the set of matchings which contribute in the limit.

Two index sets $\mathcal{U}_i$ and $\mathcal{U}_{i'}$ are said to be connected if there is (k,δ) and (k',δ') with $(u_{3k+\delta}^{(i)}, v_{2k+\delta}^{(i)}) = (u_{3k+\delta}^{(i')}, v_{2k'+\delta'}^{(i')})$. Also a collection of index sets $\{\mathcal{U}_{i_1}, \mathcal{U}_{i_2}, \ldots, \mathcal{U}_{i_s}\}$, $s \ge 2$, is said to form a connected group if for each $1 \le k \le s-1$, $\mathcal{U}_{i_k}$ and $\mathcal{U}_{i_{k+1}}$ is connected. Note that, in a typical matching in C_{T+1}, for each i, $\mathcal{U}_i$ is connected with some other $\mathcal{U}_{i'}$, $i' \ne i$. Therefore, each matching in C_{T+1} corresponds to some disjoint connected groups each of length at least 2. Consider the following disjoint decomposition of C_{T+1}.

$$\begin{aligned}
C_{T+1} &= \bigcup_{\substack{2\le g_1,g_2,\ldots,g_R\le T\\ \sum_{j=1}^R g_j = T,\ R\ge 1}} G(g_1,g_2,\ldots,g_R), \text{ where}\\
G(g_1,g_2,\ldots,g_R) &= \text{set of all such matchings in } C_{T+1} \text{ which form exactly}\\
&\quad\ R \text{ connected groups of length } g_1,g_2,\ldots,g_R.
\end{aligned}$$

Hence, by (5.5), we have

$$E\left[\Pi_{i=1}^T\left(\pi_i - \pi_i^0\right)\right] = E_{C_{T+1}}(\Pi_{i=1}^T\pi_i) = \sum_{\substack{2\le g_1,g_2,\ldots,g_R\le T\\ \sum_{j=1}^R g_j = T,\ R\ge 1}} E_{G(g_1,g_2,\ldots,g_R)}(\Pi_{i=1}^T\pi_i).$$

Consider a typical matching σ in $G(g_1,g_2,\ldots,g_R)$ with connected groups $G_{\sigma 1}, G_{\sigma 2}, \ldots, G_{\sigma R}$ respectively of lengths $g_1,g_2,\ldots,g_R$. Note that, for a fixed σ, $\{G_{\sigma k} : 1 \le k \le R\}$ forms a partition of $\{\mathcal{U}_k : 1 \le k \le T\}$. Also note that, if $i \ne j$, then no index in $G_{\sigma i}$ matches with any index in $G_{\sigma j}$. Hence, by independence of $\{\varepsilon_{i,j}\}$,

$$\begin{aligned}
E_{G(g_1,g_2,\ldots,g_R)}(\Pi_{i=1}^T\pi_i) &= \sum_{\sigma}\prod_{k=1}^{R} E_{G_{\sigma k}}(\pi_1,\pi_2,\ldots,\pi_T), \text{ where} && (5.9)\\
E_{G_{\sigma k}}(\pi_1,\pi_2,\ldots,\pi_T) &= E_{G_{\sigma k}}\Big(\prod_{\substack{i_j:\ \mathcal{U}_{i_j}\in G_{\sigma k}\\ 1\le j\le g_j}} \pi_{i_j}\Big),\ \forall 1 \le k \le R,
\end{aligned}$$

and $E_{G_{\sigma k}}$ is the usual expectation restricting on the matchings in $G_{\sigma k}$. For the time being assume that the following claim is true. We shall prove the claim later.

Claim. $E_{G_{\sigma k}}(\pi_1,\pi_2,\ldots,\pi_T) = O(n^{-g_k+2})$, $\forall \sigma, k$.

Therefore, for all $\sigma \in G(g_1, g_2, \ldots, g_R)$,

$$\prod_{k=1}^{R} E_{G_{\sigma k}}(\pi_1, \pi_2, \ldots, \pi_T) = O(n^{-\sum(g_j - 2)}) = O(n^{-T+2R}) \tag{5.10}$$

As $G(g_1, g_2, \ldots, g_R)$ is a finite set, by (5.9), we have

$$E_{G(g_1, g_2, \ldots, g_R)}(\Pi_{i=1}^{T} \pi_i) = O(n^{-T+2R}). \tag{5.11}$$

Note that as $g_1, g_2, \ldots, g_R \geq 2$, the maximum possible value of R is $[T/2]$, the greatest integer $\leq T/2$.

First suppose T is odd. Then we always have $T - 2R > 0$ and hence, using (5.11), $\lim E_{G(g_1, g_2, \ldots, g_R)}(\Pi_{i=1}^{T} \pi_i) = 0$. As a consequence, using (5.9), we have

$$\lim E\left[\Pi_{i=1}^{T}\left(\pi_i - \pi_i^0\right)\right] = 0, \text{ if } T \text{ is odd}, \tag{5.12}$$

proving (5.2) when T is odd.

Now suppose T is even, say $T = 2d$. Then note that

$$T - 2R \begin{cases} = 0, & \text{for } G(2, 2, \ldots, 2), R = d \\ > 0, & \text{otherwise.} \end{cases} \tag{5.13}$$

Therefore, by (5.11),

$$\lim E_{G(g_1, g_2, \ldots, g_R)}(\Pi_{i=1}^{T} \pi_i) = 0, \text{ if } G(g_1, g_2, \ldots, g_R) \neq G(2, 2, \ldots, 2), \tag{5.14}$$

and hence by (5.9), we have

$$\lim E\left[\Pi_{i=1}^{T}\left(\pi_i - \pi_i^0\right)\right] = \lim E_{G(2,2,\ldots,2)}(\Pi_{i=1}^{T} \pi_i). \tag{5.15}$$

It remains to identify the right side of (5.15) as the right side of (5.2). Note that a typical matching in $G(2, 2, \ldots, 2)$ involves d groups each with length 2. Hence, there is a one-to-one correspondence of $G(2, 2, \ldots, 2)$ and $\mathcal{S}_d$, set of all pair partitions of $\{1, 2, \ldots, 2d\}$. The one-to-one correspondence is as follows. Consider $\sigma = \{(i_1, i_2), (i_3, i_4), \ldots, (i_{2d-1}, i_{2d})\} \in \mathcal{S}_d$, then for every $1 \leq k \leq d$, $\{\mathcal{U}_{i_{2k-1}}, \mathcal{U}_{i_{2k}}\}$ forms a connected group. Therefore, by (5.9), we have

$$E_{G(2,2,\ldots,2)}(\Pi_{i=1}^{T} \pi_i) = \sum_{\sigma \in \mathcal{S}_d} \prod_{k=1}^{d} E_{\{\mathcal{U}_{i_{2k-1}}, \mathcal{U}_{i_{2k}}\}}(\pi_1, \pi_2, \ldots, \pi_T). \tag{5.16}$$

Let D be the set of all such matchings of indices in $\mathcal{U}_{i_{2k-1}} \cup \mathcal{U}_{i_{2k}}$ such that $\{\mathcal{U}_{i_{2k-1}}, \mathcal{U}_{i_{2k}}\}$ are connected. Note that

$$\begin{aligned} E_{\{\mathcal{U}_{i_{2k-1}}, \mathcal{U}_{i_{2k}}\}}(\pi_1, \pi_2, \ldots, \pi_T) &= \sum_{\sigma \in D} E_{\sigma}(\pi_{i_{2k-1}} \pi_{i_{2k}}) \\ &= E\left(\left(\pi_{i_{2k-1}} - \pi^0_{i_{2k-1}}\right)\left(\pi_{i_{2k}} - \pi^0_{i_{2k}}\right)\right), \text{ by (5.5) for } T = 2. \end{aligned}$$

Therefore, by (5.16) and (5.17), we have

$$E_{G(2,2,\ldots,2)}(\Pi_{i=1}^{T} \pi_i) = \sum_{\sigma \in \mathcal{S}_d} \prod_{k=1}^{d} E\left(\left(\pi_{i_{2k-1}} - \pi^0_{i_{2k-1}}\right)\left(\pi_{i_{2k}} - \pi^0_{i_{2k}}\right)\right). \tag{5.17}$$

Now substituting (5.17) in (5.15), we have established (5.2) for $T = 2d$. Therefore, by Steps 1 and 2, proof of (4.1) and hence of Lemma 4.1 is complete when one independent matrix is involved, provided the claim is true.

Proof of claim. As $\{\pi_i\}$ are commutative, it is enough to show that

$$E_C(\pi_1\pi_2\ldots\pi_g) = O(n^{-g+2}), \tag{5.18}$$

where C is the set of all matchings of indices in $\cup_{i=1}^g \mathcal{U}_i$ such that $\{\mathcal{U}_i : 1 \le i \le g\}$ forms a connected group. Recall $\{\pi_i\}$ from (5.3). Consider the following decomposition of C.

$$\begin{aligned} C &= \bigcup_{1\le t_j \le k_j \le r_j} C(k_j, t_j : 1 \le j \le g), \text{ where} \\ C(k_j, t_j : 1 \le j \le g) &= \text{ set of all matchings (may be pair, non-pair,} \\ &\qquad \text{crossing, non-crossing) in } C \text{ such that} \\ &\qquad (u^{(i)}_{3k_i}, v^{(i)}_{2k_i}) = (u^{(i+1)}_{3t_{i+1}-1}, v^{(i+1)}_{2t_{i+1}-1}),\ \forall 1 \le i \le g-1. \end{aligned} \tag{5.19}$$

Therefore,

$$E_C(\pi_1\pi_2\ldots\pi_g) = \sum_{1\le t_j\le k_j\le r_j} E_{C(k_j,t_j:1\le j\le g)}(\pi_1\pi_2\ldots\pi_g). \tag{5.20}$$

Now for convenience of writing, let us denote, for all $1 \le i \le g$,

$$\begin{aligned} D_i &= (\Pi_{k=1}^{t_i-1} A_{ik}(Z/\sqrt{n})B_{ik}(Z^*/\sqrt{n})C_{ik})A_{it_i}, \\ E_i &= B_{it_i}(Z^*/\sqrt{n})C_{it_i}(\Pi_{k=t_i+1}^{k_i-1} A_{ik}(Z/\sqrt{n})B_{ik}(Z^*/\sqrt{n})C_{ik})A_{ik_i}ZB_{ik_i}, \\ F_i &= C_{ik_i}(\Pi_{k=k_i+1}^{r_i} A_{ik}(Z/\sqrt{n})B_{ik}(Z^*/\sqrt{n})C_{ik}), \text{ and} \end{aligned}$$

hence,

$$n\pi_i = \operatorname{Tr}(D_iZE_iZ^*F_i),\ \forall 1 \le i \le g. \tag{5.21}$$

Therefore,

$$\begin{aligned} & n^g E_{C(k_j,t_j:1\le j\le g)}(\pi_1\pi_2\ldots\pi_g) \\ = {} & E_{C(k_j,t_j:1\le j\le g)}(\Pi_{i=1}^g \operatorname{Tr}(D_iZE_iZ^*F_i)),\ \text{(by (5.21))} \\ = {} & \sum_{\substack{\{u_{ij},v_{ik},\ 1\le i\le g\} \\ j=1,2,3,\ k=1,2}} E_{C(k_j,t_j:1\le j\le g)}(\Pi_{i=1}^g D_i(u_{i1},u_{i2})\varepsilon_{u_{i2},v_{i1}}E_i(v_{i1},v_{i2})\varepsilon_{u_{i3},v_{i2}}F_i(u_{i3},u_{i1})) \\ = {} & \sum_{\sigma\in C(k_j,t_j:1\le j\le g)} \sum_{\substack{\{u_{ij},v_{ik},\ 1\le i\le g\} \\ j=1,2,3,\ k=1,2}} E_\sigma(\Pi_{i=1}^g D_i(u_{i1},u_{i2})\varepsilon_{u_{i2},v_{i1}}E_i(v_{i1},v_{i2})\varepsilon_{u_{i3},v_{i2}}F_i(u_{i3},u_{i1})). \end{aligned}$$

Now by (5.19), we have $(u_{i3}, v_{i2}) = (u_{(i+1)2}, v_{(i+1)1})$, $\forall 1 \le i \le g-1$ and therefore,

$$\begin{aligned} & \sum_{\substack{\{u_{ij},v_{ik},\ 1\le i\le g\} \\ j=1,2,3,\ k=1,2}} E_\sigma(\Pi_{i=1}^g D_i(u_{i1},u_{i2})\varepsilon_{u_{i2},v_{i1}}E_i(v_{i1},v_{i2})\varepsilon_{u_{i3},v_{i2}}F_i(u_{i3},u_{i1})) \\ = {} & E_\sigma(\operatorname{Tr}(Z(\Pi_{i=1}^g E_iZ^*Z)Z^*(\Pi_{i=1}^g F_{g+1-i}D_{g+1-i}))). \end{aligned}$$

Hence,

$$n^g E_{C(k_j,t_j:1\le j\le g)}(\pi_1\pi_2\ldots\pi_g) \;=\; \sum_{\sigma\in C(k_j,t_j:1\le j\le g)} E_\sigma(\mathrm{Tr}(Z(\Pi_{i=1}^g E_iZ^*Z)Z^*(\Pi_{i=1}^g F_{g+1-i}D_{g+1-i}))). \quad (5.22)$$

Using the same idea as in the proof of (M1) condition for Theorem 4.1, one can show that

$$\lim n^{-2}E_\sigma\mathrm{Tr}(Z(\Pi_{i=1}^g E_iZ^*Z)Z^*(\Pi_{i=1}^g F_{g+1-i}D_{g+1-i})) \;=\; \begin{cases} O(1), & \text{if } \sigma \text{ is non-crossing pair matching} \\ o(1), & \text{if } \sigma \text{ is not non-crossing or pair matching.}\end{cases}$$

Hence by (5.22), $E_{C(k_j,t_j:1\le j\le g)}(\pi_1\pi_2\ldots\pi_g) = O(n^{-g+2})$. Therefore, by (5.20), (5.18) follows and the claim is established.

This completes the proof of Lemma 4.1 for one independent matrix Z. Note that if we have more than one independent matrix $\{Z_i\}$, then also the above proof will remain unchanged except $\varepsilon_{u,v}$ (the (u,v)-th element of Z) will be replaced by $\varepsilon_{i,u,v}$ (the (u,v)-th element of Z_i). □

5.2 Algorithm to compute moments of free variables

As we have discussed in Section 1, all joint moments of free variables are computable in terms of the moments of the individual variables. The algorithm for computing moments under freeness is different from the product rule under usual independence. For our purpose, a typical term in the moment calculations (see for example (2.4)) is

$$\varphi(d_0sb_1sd_1sb_2sd_2\cdots sb_nsd_n), \quad \text{where } \{b_i\}, \{d_i\} \text{ and } s \text{ are free.} \quad (5.23)$$

In this section, we shall discuss the algorithm for computing (5.23) in terms of the moments of $\{b_i\}$, $\{d_i\}$ and s. Note that, since $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, our φ satisfies $\varphi(ab) = \varphi(ba)$, $\forall a, b$.

Let $NC(n)$ be the set of all non-crossing partitions of $\{1, 2, \ldots, n\}$. Define recursively a family of multiplicative, multilinear functionals $\{\varphi_\pi : n \ge 1, \pi \in NC(n))\}$ by the following formula. If $\pi = \{V_1, V_2, \ldots, V_r\} \in NC(n)$, then

$$\varphi_\pi[a_1, a_2, \ldots, a_n] := \varphi(V_1)[a_1, a_2, \ldots, a_n]\cdots\varphi(V_r)[a_1, a_2, \ldots, a_n], \quad (5.24)$$

where

$$\varphi(V)[a_1, a_2, \ldots, a_n] := \varphi_n(a_{i_1}a_{i_2}\cdots a_{i_s}) \text{ for } V = (i_1 < i_2 < \ldots < i_s). \quad (5.25)$$

Let $NC_2(2n)$ be the the set of all non-crossing pair partitions of $\{1, 2, \ldots, 2n\}$ and $K(\pi) \in NC(n)$ be the Kreweras complement of the partition π (see Definition 9.21 in Nica and Speicher [2006]). Then we have the following lemma.

Lemma 5.1. *(a) Suppose $\varphi(ab) = \varphi(ba)$ $\forall a, b$. Then*

$$\begin{aligned}\varphi(d_0sb_1sd_1sb_2\cdots sd_n) &= \sum_{\pi\in NC_2(2n)} \varphi_{K(\pi)}[b_1, d_1, b_2, d_2, \ldots, b_n, d_nd_0] && (5.26)\\ &= \sum_{\pi\in NC(n)} \varphi_\pi[b_1, b_2, \ldots, b_n]\varphi_{K(\pi)}[d_1, d_2, \ldots, d_nd_0] && (5.27)\\ &= \sum_{\pi\in NC(n)} \varphi_\pi[d_1, d_2, \ldots, d_nd_0]\varphi_{K(\pi)}[b_1, b_2, \ldots, b_n]. && (5.28)\end{aligned}$$

(*b*) *Fix* $1 = k_0 < k_1 < \ldots < k_t \leq n$ *and the following subset of* $NC_2(2n)$ *as*

$$\mathcal{S} = \{\pi \in NC_2(2n) :\ \{2k_i, 2k_{i+1} - 1\} \in \pi,\ 0 \leq i \leq t,\ k_{t+1} = k_0\}.$$

Then

$$\sum_{\pi\in\mathcal{S}} \varphi_{K(\pi)}[b_1, d_1, b_2, d_2, \ldots, b_n, d_n d_0] \quad = \quad \varphi\Big(\prod_{s=0}^{t} b_{k_s}\Big)\prod_{s=1}^{t+1}\varphi(d_{k_{s-1}} s b_{k_{s-1}+1} s d_{k_{s-1}+1} \cdots s d_{k_s-1}), \tag{5.29}$$

where $k_0 = 1, d_{k_{t+1}-1} = d_n d_0$.

Relation (5.26) follows by (22.10) of Nica and Speicher [2006]. By freeness of $\{b_i\}$ and $\{d_i\}$, and by properties of the Kreweras complement (Exercises 9.41 (1), 9.42 (1) and (2) in Nica and Speicher [2006]), (5.27) and (5.28) follow from (5.26). Relation (5.29) follows from the multiplicative property (5.24) and (5.25) of partitions and from certain properties of Kreweras complement. A detailed proof of (5.29) is given in Section 9 of Supplementary file Bhattacharjee and Bose [2016].

5.3 Proof of Lemma 4.2

By Assumptions (A2) and (A3), there exists $C > 0$ such that

$$\sup_{1\leq i\leq K}\sup_{p} \|B_{2i-1}\|_2 \quad = \quad \sup_{1\leq i\leq K}\sup_{p} \|\bar{B}_{2i-1}\|_2 \leq C, \text{ and} \tag{5.30}$$

$$\sup_{1\leq i\leq L}\sup_{n} \|B_{2i}\|_2 \quad = \quad \sup_{1\leq i\leq L}\sup_{n} \|\bar{B}_{2i}\|_2 \leq C. \tag{5.31}$$

Therefore, $\forall h \geq 1$ and $1 \leq i \leq K$

$$\bar{\varphi}(\bar{b}^*_{2i-1}\bar{b}_{2i-1})^h \quad = \quad \lim \frac{1}{n+p}\mathrm{Tr}(\bar{B}^*_{2i-1}\bar{B}_{2i-1})^h \leq \sup_p \|\bar{B}^*_{2i-1}\bar{B}_{2i-1}\|_2^h \leq C^{2h}. \tag{5.32}$$

Similarly,

$$\bar{\varphi}(\underline{b}^*_{2i}\underline{b}_{2i})^h \leq C^{2h}, \ \ \forall h \geq 1,\ 1 \leq i \leq L. \tag{5.33}$$

Also note that, for all $\bar{a}_{2i-1} \in \{\bar{b}_{2i-1}, \bar{b}^*_{2i-1} : 1 \leq i \leq K\}$, $\underline{a}_{2i} \in \{\underline{b}_{2i}, \underline{b}^*_{2i} : 1 \leq i \leq L\}$ and $h \geq 1$, there exists $\{h_i : 1 \leq i \leq 2h\}$ such that

$$|\bar{\varphi}(\bar{a}_1\underline{a}_2\bar{a}_3 \ldots \bar{a}_{2h-1}\underline{a}_{2h})| \leq \prod_{i=1}^{h}\left(\bar{\varphi}(\bar{a}^*_{2i-1}\bar{a}_{2i-1})^{h_{2i-1}}\right)^{1/h_{2i-1}}\prod_{i=1}^{h}\left(\bar{\varphi}(\underline{a}^*_{2i}\underline{a}_{2i})^{h_{2i}}\right)^{1/h_{2i}}.$$

Hence, by (5.32) and (5.33)

$$|\bar{\varphi}(\bar{a}_1\underline{a}_2\bar{a}_3 \ldots \bar{a}_{2h-1}\underline{a}_{2h})| \quad \leq \quad C^{2h}, \ \forall h \geq 1. \tag{5.34}$$

Therefore, applying Lemma 5.1 (a) and using the fact that $\varphi(s^{2h}) = \#NC_2(2h) \leq 2^{2h}, \ \forall h \geq 1$,

$$|\bar{\varphi}(\bar{a}_1 s\underline{a}_2 s\bar{a}_3 \ldots \underline{a}_{2h} s)| \leq C^{2h}(\#NC_2(2h)) \leq (2C)^{2h}.$$

Hence the proof of Lemma 4.2 is complete. □

5.4 Proof of Lemma 4.3

Note that there is a $C > 0$ such that for any $\{\bar{a}_{2i-1}\} \in \{\bar{b}_{2i-1}, \bar{b}^*_{2i-1}\}$, $\{\underline{a}_{2i}\} \in \{\underline{b}_{2i}, \underline{b}^*_{2i}\}$ and $h \geq 1$, we have

$$|\bar{\varphi}(\bar{a}_1\bar{a}_3 \dots \bar{a}_{2h-1})| \leq C^h, \ \ |\bar{\varphi}(\underline{a}_2\underline{a}_4 \dots \underline{a}_{2h})| \leq C^h \text{ and } |\bar{\varphi}(\bar{\delta})^h| \leq C^h. \tag{5.35}$$

Proof of (5.35) is along the same lines as the proof of Condition (C) in Theorem 4.1. Hence we omit it.

We first show that

$$(K(z,f))^r = \sum_{1 \leq j_1, j_2, \dots, j_r < \infty} z^{-\sum_{k=1}^r j_k} \bar{\varphi}(\prod_{k=0}^r R_{j_k}(f)) \tag{5.36}$$

exists. For this, we have to show

$$\sum_{1 \leq j_1, j_2, \dots, j_r < \infty} |z|^{-\sum_{k=1}^r j_k} |\bar{\varphi}(\prod_{k=0}^r R_{j_k}(f))| < \infty, \ \forall r \geq 1. \tag{5.37}$$

Now, note that, by (4.13)

$$R_j(f) = \bar{\varphi}(h(d,e,f)\bar{\delta}^{j-1}|f) = \sum_{i=0}^q \bar{\varphi}(\bar{b}_{4i-3}\bar{b}_{4i-1}\bar{\delta}^{j-1})\underline{b}_{2i}.$$

Therefore,

$$\begin{aligned}
|\bar{\varphi}(\prod_{k=0}^r R_{j_k}(f))| &\leq \sum_{i_1, i_2, \dots, i_r = 0}^q \left(\prod_{k=1}^r |\bar{\varphi}(\bar{b}_{4i_k-3}\bar{b}_{4i_k-1}\bar{\delta}^{j_k-1})| \right) |\bar{\varphi}(\prod_{k=1}^r \underline{b}_{2i_k})| \\
&\leq \sum_{i_1, i_2, \dots, i_r = 0}^q \left(\prod_{k=1}^r \left(\bar{\varphi}(\bar{b}_{4i_k-3}\bar{b}_{4i_k-1}\bar{b}^*_{4i_k-1}\bar{b}^*_{4i_k-3})\bar{\varphi}(\bar{\delta})^{2j_k-2} \right)^{1/2} \right) |\bar{\varphi}(\prod_{k=1}^r \underline{b}_{2i_k})| \\
&\leq (q+1)^r C^{2r+\sum_{k=1}^r j_k}, \ \text{ by (5.35)} \\
&\leq C_1^{\sum_{k=1}^r j_k}, \ \text{ for some } C_1 > C > 0.
\end{aligned} \tag{5.38}$$

Therefore, for $|z| > \eta C_1$, $\eta > 1$, $(K(z,f))^r$ in (5.36) exists and moreover, for all $r \geq 1$, we have

$$\bar{\varphi}(K(z,f))^r = \bar{\varphi}\Big(\sum_{j=1}^\infty \frac{R_j(f)}{z^j}\Big)^r \ = \ \Big(\sum_{j_1, j_2, \dots, j_r = 1}^\infty \frac{\bar{\varphi}(R_{j_1}(f)R_{j_2}(f) \dots R_{j_r}(f))}{z^{j_1}z^{j_2} \dots z^{j_r}} \Big).$$

Therefore,

$$\begin{aligned}
|\bar{\varphi}(K(z,f))^r| &\leq \Big(\sum_{j_1, j_2, \dots, j_r = 1}^\infty \frac{|\bar{\varphi}(R_{j_1}(f)R_{j_2}(f) \dots R_{j_r}(f))|}{|z|^{j_1}|z|^{j_2} \dots |z|^{j_r}} \Big) \\
&\leq \sum_{j_1, j_2, \dots, j_r = 1}^\infty \frac{C_1^{\sum_{k=1}^r j_k}}{|z|^{j_1}|z|^{j_2} \dots |z|^{j_r}}, \ \text{ by (5.38)}
\end{aligned}$$

$$\leq \left(\sum_{j=1}^{\infty} \frac{C_1^j}{(\eta C_1)^j}\right)^r \leq \left(\frac{1}{\eta - 1}\right)^r, \text{ as } \eta > 1. \tag{5.39}$$

Note that by (4.17), (we shall show below that this infinite sum in (5.40) exists)

$$B(d, e, z) = \sum_{i=0}^{\infty} \sum_{j=0}^{q} \bar{\varphi}(\underline{b}_{2j}(-K(z, f))^i)\bar{b}_{4j-3}\bar{b}_{4j-1} = \sum_{i=0}^{\infty} A_i, \text{ say.} \tag{5.40}$$

Now, for all $|z| > \eta C_1, \eta > 1$,

$$\begin{aligned}
|\bar{\varphi}(A_{i_1}A_{i_2}\dots A_{i_r})| &= \sum_{j_1, j_2, \dots, j_r = 0}^{q} \left(\prod_{k=1}^{r} |\bar{\varphi}(\underline{b}_{2j_k}(-K(z, f))^{i_k}|\right) |\bar{\varphi}(\prod_{k=1}^{r} \bar{b}_{4j_k-3}\bar{b}_{4j_k-1})| \\
&\leq \sum_{j_1, j_2, \dots, j_r = 0}^{q} \left(\prod_{k=1}^{r} \left(\bar{\varphi}(\underline{b}_{2j_k}\underline{b}^*_{2j_k})\bar{\varphi}(K(z, f))^{2i_k}\right)^{1/2}\right) |\bar{\varphi}(\prod_{k=1}^{r} \bar{b}_{4j_k-3}\bar{b}_{4j_k-1})| \\
&\leq \sum_{j_1, j_2, \dots, j_r = 0}^{q} \left(\prod_{k=1}^{r} C^2 \left(\frac{1}{\eta - 1}\right)^{2i_k}\right)^{1/2} C^{2r} \\
&\qquad \text{by (5.35) and (5.39)} \\
&\leq (q + 1)^r C^{3r} \left(\prod_{k=1}^{r} \left(\frac{1}{\eta - 1}\right)^{i_k}\right)
\end{aligned} \tag{5.41}$$

Therefore,

$$\begin{aligned}
\sum_{1 \leq i_1, i_2, \dots, i_r < \infty} |\bar{\varphi}(A_{i_1}A_{i_2}\dots A_{i_r})| &\leq (q + 1)^r C^{3r} \left(\sum_{i=0}^{\infty} \left(\frac{1}{\eta - 1}\right)^i\right)^r, \text{ (by (5.41))} \\
&\leq \left(\frac{(q + 1)C(\eta - 1)}{(\eta - 2)}\right)^r, \text{ as } \eta > 2 \\
&\leq C_2^r, \text{ for some } C_2 > C_1 > C > 0.
\end{aligned} \tag{5.42}$$

Therefore, observing (5.40), $(B(d, e, z))^r$ exists for all $|z| > \eta C_1$, $\eta > 2$. Now, using the same arguments as in the existence of $K(z, f)$ above, it is easy to see that $G(d, e, z)$ defined in (4.18) exists for all $|z| > C_2$. This completes the proof of Lemma 4.3. □

5.5 Proof of Theorem 4.2

To prove Theorem 4.2, we need a lemma that provides a recursion formula for the moments of $\bar{\mu}$. For convenience of writing, let us denote $D_i = \bar{B}_{4i-3}$, $E_i = \bar{B}_{4i-1}$, $d_i = \bar{b}_{4i-3}$ and $e_i = \bar{b}_{4i-1}$, $f_i = \underline{b}_{2i}$ for all $1 \leq i \leq q$. For any polynomial $\Pi = \Pi(D_j, D_j^*, E_j, E_j^* : j \geq 0)$, let $\Pi^0 = \Pi(d_j, d_j^*, e_j, e_j^* : j \geq 0)$. Recall $\{R_j(f)\}$ defined in (4.13). For all $j \geq 0$, let

$$S_j(f, \Pi) = \bar{\varphi}\left(\bar{\Pi}^0 h(d, e, f)\bar{\delta}^{j-1}|f\right) := \sum_{i=0}^{q} \bar{\varphi}(\bar{\Pi}^0 \bar{b}_{4i-3}\bar{b}_{4i-1}\bar{\delta}^{j-1})\underline{b}_{2i}. \tag{5.43}$$

Lemma 5.2. *Suppose (A1)-(A3) hold and* $p, n(p) \to \infty$*,* $p/n \to y > 0$*. Then for any polynomial* $\Pi = \Pi(D_j, D_j^*, E_j, E_j^* : j \geq 0)$*, we have*

$$\lim (n+p)^{-1} E\mathrm{Tr}(\Pi \bar{\Delta}^r) = \sum_{t=1}^{r} \bar{\varphi}\Big[\sum_{\substack{1 \leq i_1, i_2, \ldots, i_t \leq r \\ \sum_{j=1}^{t} i_j = r}} S_{i_1}(f, \Pi) \big(\prod_{k=2}^{t} R_{i_k}(f) \big) \Big].$$

Proof. By Theorem 4.1, it is immediate that

$$\lim (n+p)^{-1} E\mathrm{Tr}(\Pi \Delta^r) = \bar{\varphi}(\Pi^0 \bar{\delta}^r) = (1+y)^r \sum_{\substack{j_k=1 \\ 1 \leq k \leq r}}^{q} \bar{\varphi}\left(\Pi^0 \prod_{k=1}^{r} d_{j_k} s f_{j_k} s e_{j_k} \right)$$

$$= \sum_{\sigma \in NC_2(2r)} \tau_\sigma, \quad \text{by (5.26)} \tag{5.44}$$

where

$$\tau_\sigma = (1+y)^r \sum_{\substack{j_k=1 \\ 1 \leq k \leq r}}^{q} \bar{\varphi}_{K(\sigma)}[f_{j_1}, e_{j_1} d_{j_2}, f_{j_2}, e_{j_2} d_{j_3}, \ldots, e_{j_r} d_{j_1} \Pi^0],$$

and $K(\sigma)$ is the Kreweras complement. Now to compute (5.44), we consider the decomposition of $NC_2(2r) = \cup_{t=1}^{r} \mathcal{P}_t^{2r}$, where $\mathcal{P}_1^{2r} = \{\sigma \in NC_2(2r) : \{1, 2\} \in \sigma\}$ and for all $2 \leq t \leq r$,

$$\begin{aligned} \mathcal{P}_t^{2r} = \{ & \sigma \in NC_2(2r) : \{2k_0 - 1, 2k_t\}, \{2k_0, 2k_1 - 1\}, \{2k_1, 2k_2 - 1\}, \ldots, \\ & \{2k_{t-2}, 2k_{t-1} - 1\} \in \sigma, \ 1 = k_0 < k_1 < k_2 < \ldots < k_{t-1} \leq r \}. \end{aligned}$$

Hence, (5.44) is equivalent to

$$\lim (n+p)^{-1} E\mathrm{Tr}(\Pi \bar{\Delta}^r) = \sum_{t=1}^{r} \mathcal{T}_t, \tag{5.45}$$

where for all $1 \leq t \leq r$,

$$\mathcal{T}_t = \sum_{\sigma \in \mathcal{P}_t^{2r}} \tau_\sigma = \sum_{1 = k_0 < k_1 < k_2 < \ldots < k_{t-1} \leq r} g(t, k_1, k_2, \ldots, k_{t-1}), \tag{5.46}$$

and

$$\begin{aligned} & (1+y)^{-t-1} g(t+1, k_1, k_2, \ldots, k_t) \\ = & \sum_{1 \leq j_{k_s} \leq q} \bar{\varphi}\big(\prod_{s=0}^{t} f_{j_{k_s}} \big) \prod_{s=0}^{t} \bar{\varphi}\big(e_{j_{k_s}} \bar{\delta}^{k_{s+1} - k_s - 1} d_{j_{k_{(s+1)}}} \big), \\ & \text{(by Lemma 5.1 (b) and where } k_{t+1} = r+1 \text{ and } d_{j_{k_{t+1}}} = d_{j_{k_0}} \Pi^0) . \end{aligned} \tag{5.47}$$

Therefore, (5.47) is equal to

$$= \sum_{1 \leq j_{k_s} \leq q} \bar{\varphi}\big(\prod_{s=0}^{t} f_{j_{k_s}} \big) \prod_{s=0}^{t} \bar{\varphi}\big(e_{j_{k_s}} \bar{\delta}^{k_{s+1} - k_s - 1} d_{j_{k_{(s+1)}}} \big)$$

$$= \bar{\varphi}\Big(\sum_{j_{k_s}, j_{k_{s+1}}} \prod_{s=0}^{t} f_{j_{k_s}} \bar{\varphi}(e_{j_{k_s}} \bar{\delta}^{k_{s+1}-k_s-1} d_{j_{k_{(s+1)}}}) \Big),$$

where $j_{k_{t+1}} = j_{k_0}$.

Note that

$$\sum_{j'_{k_s}, j_{k_{s+1}}} f_{j_{k_s}} \bar{\varphi}(e_{j_{k_s}} \bar{\delta}^{k_{s+1}-k_s-1} d_{j_{k_{(s+1)}}}) = \begin{cases} (1+y)^{-1} R_{(k_{s+1}-k_s)}(f), \ 1 \le s \le t-1 \\ (1+y)^{-1} S_{(r+1-k_t)}(f, \Pi), \ s = t. \end{cases} .$$

Hence for all $1 = k_0 < k_1 < k_2 < \ldots < k_t \le r$, we have

$$g(t+1, k_1, k_2, \ldots, k_t) \ = \ \bar{\varphi}(S_{(r+1-k_t)}(f, \Pi) \prod_{s=0}^{t-1} R_{(k_{s+1}-k_s)}(f)).$$

Therefore, by (5.46), for all $1 \le t \le r$,

$$\begin{aligned} \mathcal{T}_t &= \sum_{\substack{1=k_0<k_1<\ldots<k_t\le r \\ k_{t+1}=r+1}} \bar{\varphi}[S_{(r+1-k_t)}(f, \Pi) \prod_{s=0}^{t-1} R_{(k_{s+1}-k_s)}(f)] \\ &= \sum_{\substack{1\le i_1, i_2, \ldots, i_t \le r \\ i_1+i_2+\ldots+i_t=r}} \bar{\varphi}(S_{i_1}(f, \Pi) \prod_{s=2}^{t} R_{i_s}(f)). \end{aligned}$$

Hence, by using (5.45) and (5.46), Lemma 5.2 follows. □

Now we are ready to prove Theorem 4.2.

(a) Define

$$D = \sum_{i=1}^{\infty} z^{-i} \bar{\delta}^i. \tag{5.48}$$

Note that by the last part of (5.35), D exists for sufficiently large $|z|$. Moreover, by (4.8), $\varphi(D) = m_{\bar{\mu}}(z)$.

To establish (4.20), note that

$$\begin{aligned} K(z, f) &= \sum_{i=1}^{\infty} z^{-i} \bar{\varphi}(h(d, e, f) \bar{\delta}^{i-1} | f), \text{ by (4.13) and (4.14)} \\ &= -z^{-1} \bar{\varphi}(h(d, e, f) | f) - z^{-1} \bar{\varphi}(h(d, e, f) D | f) \end{aligned} \tag{5.49}$$

where for f' with same property as in f, we have

$$\begin{aligned} &\bar{\varphi}(h(d, e, f) D | f) \\ = &\sum_{r=1}^{\infty} z^{-r} \sum_{t=1}^{r} \bar{\varphi}\Big[\Big(\sum_{\substack{1\le i_1, i_2, \cdots, i_t \le r \\ i_1+i_2+\cdots+i_t=r}} S_{i_1}(f', h(d, e, f)) \prod_{s=2}^{t} R_{i_s}(f') \Big) | f \Big] \end{aligned}$$

(by Lemma 5.2)

$$
\begin{aligned}
&= \sum_{t=1}^{\infty} \bar{\varphi}\Big[\sum_{r=t}^{\infty} z^{-r} \Big(\sum_{\substack{1 \le i_1, i_2, \cdots, i_t \le r \\ i_1 + i_2 + \cdots + i_t = r}} S_{i_1}(f', h(d,e,f)) \prod_{s=2}^{t} R_{i_s}(f') \Big) | f \Big] \\
&= \sum_{t=1}^{\infty} \bar{\varphi}\Big[(K(z,f'))^{t-1} \left(\sum_{r=1}^{\infty} z^{-r} S_r(f', h(d,e,f)) \right) | f \Big] \\
&= \bar{\varphi}\Big[(1 + K(z,f'))^{-1} \left(\sum_{r=1}^{\infty} z^{-r} S_r(f', h(d,e,f)) \right) | f \Big] \\
&= \bar{\varphi}\left(\sum_{r=1}^{\infty} z^{-r} \bar{\delta}^{r-1} h(d,e,f) \bar{\varphi}\Big[h(d,e,f)(1 + yK(z,f'))^{-1} \Big] | f \right) \\
&= z^{-1} \bar{\varphi}(h(d,e,f)B(d,e,z)|f) + z^{-1} \varphi(Dh(d,e,f)B(d,e,z)|f).
\end{aligned}
$$

In a similar fashion, using $h(d,e,f)B(d,e,z)$ instead of $h(d,e,f)$ in the above steps,

$$
\bar{\varphi}(Dh(d,e,f)B(d,e,z)|f) = z^{-1}\bar{\varphi}(h(d,e,f)B^2(d,e,z)|f) + z^{-1}\bar{\varphi}(Dh(d,e,f)B^2(d,e,z)|f).
$$

Finally iterating we have

$$
\bar{\varphi}(h(d,e,f)D|f) = \sum_{r=1}^{\infty} z^{-r} \bar{\varphi}(h(d,e,f)B^r(d,e,z)|f). \tag{5.50}
$$

Hence,

$$
\begin{aligned}
&K(z,f) \\
= \quad &-z^{-1}\bar{\varphi}(\sum_{r=0}^{\infty} h(d,e,f) z^{-r} B^r(d,e,z)|f) = \bar{\varphi}(h(d,e,f)(B(d,e,z) - z)^{-1}|f),
\end{aligned} \tag{5.51}
$$

which is (4.20) in Theorem 4.2.

Note that the above steps from (5.49) leading to (5.51) remain valid if we replace $h(d,e,f)$ by 1 in (5.49). This yields (instead of (5.51)),

$$
\begin{aligned}
m_{\bar{\mu}}(z) &= -z^{-1} \sum_{i=0}^{\infty} \left(\frac{B(d,e,z)}{z} \right)^i && (5.52) \\
&= \bar{\varphi}((B(d,e,z) - z)^{-1}), && (5.53)
\end{aligned}
$$

which is (4.19) in Theorem 4.2. Hence the proof of Theorem 4.2 (a) is complete.

(b) Let δ_0 be the degenerate probability measure at 0. Then (4.22) follows immediately by noting that

$$
\bar{\mu} = \frac{y}{1+y}\mu + \frac{1}{1+y}\delta_0. \tag{5.54}
$$

Hence the proof of Theorem 4.2 is complete.

References

G. Anderson, A. Guionnet, and O. Zeitouni. *An Introduction to Random matrices.* Cambridge University Press, Cambridge, UK, 2009.

Z. D. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices.* Springer, 2010.

Z. D. Bai and L. X. Zhang. The limiting spectral distribution of the product of the Wigner matrix and a nonnegative definite matrix. *J. Mult. Anal.*, 101(9):1927–1949, 2010.

F. Benaych-Georges. Rectangular random matrices, related convolution. *Probab. Theory Related Fields*, 144(3-4):471–515, 2009.

F. Benaych-Georges. On a surprising relation between the Marčenko–Pastur law, rectangular and square free convolutions. *Ann. de l'institut Henri Poincaré (B)*, 46(3):644–652, 2010.

F. Benaych-Georges and R. R. Nadakuditi. The singular values and vectors of low rank perturbations of large rectangular random matrices. *J. Mult. Anal.*, 111:120–135, 2012.

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

R. Couillet and M. Debbah. *Random matrix methods for wireless communications.* Cambridge University Press, Cambridge, UK, 2011.

B. Jin, C. Wang, Z. D. Bai, K. K. Nair, and M. Harding. Limiting spectral distribution of a symmetrized auto-cross covariance matrix. *Ann. Appl. Probab.*, 24(3):1199–1225, 2014.

H. Liu, A. Aue, and D. Paul. On the Marčenko–Pastur law for linear time series. *Ann. Statist.*, 43(2):675–712, 2015.

V. Marčenko and L. Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1:457–483, 1967.

A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*. Cambridge University Press, Cambridge, UK, 2006.

O. Pfaffel and E. Schlemm. Eigenvalue distribution of large sample covariance matrices of linear processes. *Probab. Math. Statist.*, 31(2):313–329, 2011.

R. Speicher and C. Vargas. Free deterministic equivalents, rectangular random matrix models, and operator-valued free probability theory. *Random Matrices: Theory and Applications*, 1 (02):1150008, 2012.

J. Yao. A note on a Marčenko–Pastur type theorem for time series. *Statist. Probab. Lett.*, 82: 22–28, 2012. doi: 10.1006/jmva.1995.1083.