Electronic Journal of Probability

Electron. J. Probab. **17** (2012), no. 95, 1–22.
ISSN: 1083-6489 DOI: 10.1214/EJP.v17-2006

Fluctuations of eigenvalues for random Toeplitz and related matrices

Dang-Zheng Liu* Xin Sun† Zheng-Dong Wang‡

Abstract

Consider random symmetric Toeplitz matrices $T_n = (a_{i-j})_{i,j=1}^n$ with matrix entries $a_j, j = 0, 1, 2, \cdots$, being independent real random variables such that

$$\mathbb{E}[a_j] = 0, \quad \mathbb{E}[|a_j|^2] = 1 \;\text{ for }\; j = 0, 1, 2, \cdots,$$

(homogeneity of 4-th moments)

$$\kappa = \mathbb{E}[|a_j|^4],$$

and further (uniform boundedness)

$$\sup_{j\geq 0} \mathbb{E}[|a_j|^k] = C_k < \infty \quad \text{for} \quad k \geq 3.$$

Under the assumption of $a_0 \equiv 0$, we prove a central limit theorem for linear statistics of eigenvalues for a fixed polynomial with degree at least 2. Without this assumption, the CLT can be easily modified to a possibly non-normal limit law. In a special case where a_j's are Gaussian, the result has been obtained by Chatterjee for some test functions. Our derivation is based on a simple trace formula for Toeplitz matrices and fine combinatorial analysis. Our method can apply to other related random matrix models, including Hermitian Toeplitz and symmetric Hankel matrices. Since Toeplitz matrices are quite different from Wigner and Wishart matrices, our results enrich this topic.

Keywords: Toeplitz (band) matrix; Hankel matrix; Random matrices; Linear statistics of eigenvalues; Central limit theorem.

AMS MSC 2010: 60B20; 60F05.

Submitted to EJP on May 4, 2012, final version accepted on October 28, 2012.

Supersedes arXiv:1010.3394v2.

*University of Science and Technology of China, China. E-mail: dzliu@ustc.edu.cn

†Massachusetts Institute of Technology, USA. E-mail: xinsun89@math.mit.edu

‡Peking University, China. E-mail: zdwang@pku.edu.cn

1 Introduction and main results

Toeplitz matrices appear very often in mathematics and physics and also in plenty of applications, see Grenander and Szegő's book [14] for a detailed introduction to deterministic Toeplitz matrices. The study of random Toeplitz matrices with independent entries is proposed by Bai in his review paper [2]. Since then, the literature around the asymptotic distribution of eigenvalues for random Toeplitz and related matrices is very large, including the papers of Basak and Bose [4], Bose et al. [6, 7], Bryc et al. [8], Hammond and Miller [15], Kargin [19], Liu and Wang [21], Massey et al. [22], Sen and Virág [24, 25]. We refer to [4, 24, 25] for recent progress. However, the study of fluctuations of eigenvalues for random Toeplitz matrices is quite little, to the best of our knowledge, the only known result comes from Chatterjee [9] in the special case where the matrix entries are Gaussian distributions. In this paper we will derive a central limit theorem (CLT for short) for linear statistics of eigenvalues of random Toeplitz and related matrices.

In the literature fluctuations of eigenvalues for random matrices have been extensively studied. The investigation of central limit theorems for linear statistics of eigenvalues of random matrices dates back to the work of Jonsson [18] on Gaussian Wishart matrices. Similar work for the Wigner matrices was obtained by Sinai and Soshnikov [26]. For further discussion on Wigner (band) matrices and Wishart matrices and their generalized models, we refer to Bai and Silverstein's book [3], recent papers [1, 9] and the references therein. For another class of invariant random matrix ensembles, Johansson [17] proved a general result which implies CLT for linear statistics of eigenvalues. Recently, Dumitriu and Edelman [13] and Popescu [23] proved that CLT holds for tridiagonal random matrix models.

Another important contribution is the work of Diaconis et al. [12, 11], who proved similar results for random unitary matrices. These results are closely connected to Szegö's limit theorem (see [16]) for the determinant of Toeplitz matrices with (j,k) entry $\widehat{g}(j-k)$, where $\widehat{g}(k)=\frac{1}{2\pi}\int_0^{2\pi} g(e^{i\theta})e^{-ik\theta}d\,\theta$, see [10, 5] and references therein for connections between random matrices and Toeplitz determinants. Here we emphasize that Szegö's limit theorem implies a CLT (see [16]).

Now we turn to our model. The matrix of the form $T_n=(a_{i-j})_{i,j=1}^n$ is called a Toeplitz matrix. If we introduce the Toeplitz or Jordan matrices $B=(\delta_{i+1,j})_{i,j=1}^n$ and $F=(\delta_{i,j+1})_{i,j=1}^n$, respectively called the "backward shift" and "forward shift" because of their effect on the elements of the standard basis $\{e_1,\cdots,e_n\}$ of $\mathbb{R}^n$, then an $n\times n$ matrix T can be written in the form

$$T=\sum_{j=0}^{n-1} a_{-j}B^j+\sum_{j=1}^{n-1} a_jF^j \tag{1.1}$$

if and only if T is a Toeplitz matrix where $a_{-n+1},\cdots,a_0,\cdots,a_{n-1}$ are complex numbers [20]. It is worth emphasizing that this representation of a Toeplitz matrix is of vital importance as the starting point of our method. The "shift" matrices B and F exactly present the information of the traces.

Consider a Toeplitz band matrix as follows. Given a band width $b_n<n$, let

$$\eta_{ij}=\begin{cases}1, & |i-j|\le b_n;\\ 0, & \text{otherwise.}\end{cases} \tag{1.2}$$

Then a Toeplitz band matrix is

$$T_n=(\eta_{ij}\,a_{i-j})_{i,j=1}^n. \tag{1.3}$$

Moreover, the Toeplitz band matrix T_n can also be rewritten in the form

$$T_n = \sum_{j=0}^{b_n} a_{-j} B^j + \sum_{j=1}^{b_n} a_j F^j = a_0 I_n + \sum_{j=1}^{b_n} \left(a_{-j} B^j + a_j F^j \right), \tag{1.4}$$

where I_n is the identity matrix. Obviously, a Toeplitz matrix can be considered as a band matrix with the bandwidth $b_n = n-1$. In this paper, the basic model under consideration consists of $n \times n$ random symmetric Toeplitz band matrices $T_n = (\eta_{ij}\, a_{i-j})_{i,j=1}^n$ in Eq. (1.3). We assume that $a_j = a_{-j}$ for $j = 1, 2, \cdots$, and $\{a_j\}_{j=1}^{\infty}$ is a sequence of independent real random variables such that

$$\mathbb{E}[a_j] = 0, \;\; \mathbb{E}[|a_j|^2] = 1 \;\text{ for }\;\; j = 1, 2, \cdots, \tag{1.5}$$

(homogeneity of 4-th moments)

$$\kappa = \mathbb{E}[|a_j|^4], \tag{1.6}$$

and further (uniform boundedness)

$$\sup_{j \ge 1} \mathbb{E}[|a_j|^k] = C_k < \infty \quad \text{for} \quad k \ge 3. \tag{1.7}$$

In addition, we also assume $a_0 \equiv 0$ (we will explain in Remarks 1.3 and 5.5 below!) and the bandwidth $b_n \to \infty$ but $b_n/n \to b \in [0, 1]$ as $n \to \infty$.

Set $A_n = \frac{T_n}{\sqrt{b_n}}$, a linear statistic of eigenvalues $\lambda_1, \cdots, \lambda_n$ of A_n is a function of the form

$$\frac{1}{n} \sum_{j=1}^{n} f(\lambda_j), \tag{1.8}$$

where f is some fixed function. In particular, when $f(x) = x^p$ we write

$$\omega_p = \frac{\sqrt{b_n}}{n} \sum_{j=1}^{n} \left(\lambda_j^p - \mathbb{E}[\lambda_j^p] \right). \tag{1.9}$$

These ω_p, $p = 2, 3, \ldots$, are our main objects. Note that $\frac{1}{n} \sum_{j=1}^{n} \left(\lambda_j^p - \mathbb{E}[\lambda_j^p] \right)$ converges weakly to zero as $n \longrightarrow \infty$, moreover under the condition

$$\sum_{j=1}^{\infty} \frac{1}{b_n^2} < \infty$$

we have a strong convergence, see [4, 19, 21]. We remark that the fluctuations

$$\frac{1}{n} \sum_{j=1}^{n} \lambda_j^p - \frac{1}{n} \sum_{j=1}^{n} \mathbb{E}[\lambda_j^p]. \tag{1.10}$$

are of order $\frac{1}{\sqrt{b_n}}$ while those for Wigner matrices are of order $\frac{1}{n}$. Thus Toeplitz case is more like the case of classical central limit theorem, especially when $b_n = n$. This shows that the correlations between eigenvalues for Toeplitz matrices are much weaker than those for Wigner matrices (another different phenomenon is that the limiting distribution for random Toeplitz matrices has unbounded support, see [8, 15]). The potential reasons for this phenomenon may come from the fact that the order of the number of independent variables is $O(n)$ for Toeplitz matrices while it is $O(n^2)$ for Wigner matrices. On the other hand, the eigenvalues of random Toeplitz matrices are obviously not independent, so it is different from the case of CLT for independent variables.

In the special case that the matrix entries a_j are Gaussian distributions, by using his notion of “second order Poincaré inequalities” Chatterjee in [9] proved the following theorem:

Theorem ([9], Theorem 4.5) *Consider the Gaussian Toeplitz matrices $T_n = (a_{i-j})_{i,j=1}^n$, i.e. $a_j = a_{-j}$ for $j = 1, 2, \cdots$, and $\{a_j\}_{j=0}^{\infty}$ is a sequence of independent standard Gaussian random variables. Let p_n be a sequence of positive integers such that $p_n = o(\log n/ \log \log n)$. Let $A_n = T_n/\sqrt{n}$, then, as $n \to \infty$,*

$$\frac{\mathrm{tr}(A_n^{p_n}) - \mathbb{E}[\mathrm{tr}(A_n^{p_n})]}{\sqrt{\mathrm{Var}\left(\mathrm{tr}(A_n^{p_n})\right)}} \text{ converges in total variation to } N(0,1).$$

The CLT *also holds for* $\mathrm{tr}(f(A_n))$*, when f is a fixed nonzero polynomial with nonnegative coefficients.*

The author remarked that the theorem above is only for Gaussian Toeplitz matrices based on the obvious fact: considering the function $f(x) = x$, CLT may not hold for linear statistics of non-Gaussian Toeplitz matrices. The author also remarked that the theorem above says nothing about the limiting formula of the variance $\mathrm{Var}\left(\mathrm{tr}(A_n^{p_n})\right)$. However, we assert that CLT holds for a test function $f(x) = x^{2p}$ even for non-Gaussian Toeplitz matrices. When $f(x) = x^{2p+1}$ the fluctuation is Gaussian if and only if the diagonal random variable a_0 is Gaussian. Moreover, if we suppose $a_0 \equiv 0$, we can obtain CLT for any fixed polynomial test functions. On the other hand, we can calculate the variance in terms of integrals associated with pair partitions. Unfortunately, our method fails to deal with the test function $f(x) = x^{p_n}$, where p_n depends on n.

Our study is inspired by the work of Sinai and Soshnikov [26], but new ideas are needed since the structure of Toeplitz matrices is quite different from that of Wigner matrices. More specifically, we first expand $\mathrm{tr}(T_n^p)$ into $\sum_i \sum_J a_J I_J$ where the “balanced" vector $J = (j_1, \ldots, j_p) \in \{-b_n, \ldots, b_n\}^p$ with $\sum_{l=1}^p j_l = 0$, $a_J = \prod_{l=1}^p a_l$ and $I_J = \prod_{k=1}^p \chi_{[1,n]}(i + \sum_{l=1}^k j_l)$. So we can reduce estimates about higher moments of $\mathrm{tr}(T_n^p)$ to the combinatorial analysis of correlated “balanced" vectors. In addition, our method can apply to other related random matrix models, including Hermitian Toeplitz matrices and Hankel matrices.

Now we state the main theorem as follows.

Theorem 1.1. *Let T_n be a real symmetric ((1.5)–(1.7)) random Toeplitz band matrix with bandwidth b_n, where $b_n/n \to b \in [0,1]$ and $b_n \to \infty$ as $n \to \infty$. Set $A_n = T_n/\sqrt{b_n}$ and*

$$\omega_p = \frac{\sqrt{b_n}}{n} \left(\mathrm{tr}(A_n^p) - \mathbb{E}[\mathrm{tr}(A_n^p)]\right). \tag{1.11}$$

For every $p \geq 2$, we have

$$\omega_p \longrightarrow N(0, \sigma_p^2) \tag{1.12}$$

in distribution as $n \to \infty$. Moreover, for a given polynomial

$$Q(x) = \sum_{j=2}^{p} q_j x^j \tag{1.13}$$

with degree $p \geq 2$, set

$$\omega_Q = \frac{\sqrt{b_n}}{n} \left(\mathrm{tr}Q(A_n) - \mathbb{E}[\mathrm{tr}Q(A_n)]\right), \tag{1.14}$$

we also have

$$\omega_Q \longrightarrow N(0, \sigma_Q^2) \tag{1.15}$$

in distribution as $n \to \infty$. Here the variances σ_p^2 and σ_Q^2 will be given in section 4.

From the proof of our main theorem, we can easily derive an interesting result concerning product of independent variables whose subscripts satisfy certain "balance" condition. When $p = 2$, it is a direct result from the classical central limit theorem. Here we state it but omit its proof, see Remark 5.6 for detailed explanation.

Corollary 1.2. *Suppose that $a_j = a_{-j}$ for $j = 1, 2, \cdots$, and $\{a_j\}_{j=1}^{\infty}$ is a sequence of independent random variables satisfying the assumptions (1.5)–(1.7).*

For every $p \geq 2$,

$$\frac{1}{k^{\frac{p-1}{2}}} \sum_{0 \neq j_1, \ldots, j_p = -k}^{k} \left(\prod_{l=1}^{p} a_{j_l} - \mathbb{E}[\prod_{l=1}^{p} a_{j_l}] \right) \delta_{0, \sum_{l=1}^{p} j_l} \tag{1.16}$$

converges in distribution to a Gaussian distribution $N(0, \sigma_p^2)$ as k tends to infinity. Here the variance σ_p^2 corresponds to the case $b = 0$ as in section 4.

Remark 1.3 (on the diagonal entry a_0)**.** *Rewrite $A_n(a_0) = A_n(0) + \frac{a_0}{\sqrt{b_n}} I_n$, where $A_n(0)$ denotes the matrix with $a_0 = 0$. It is easy to see that*

$$\omega_p(a_0) = \omega_p(0) + p\frac{\mathrm{tr}(A_n^{p-1}(0))}{n}(a_0 - \mathbb{E}[a_0]) + O(b_n^{-1/2}), \tag{1.17}$$

which converges in distribution to the distribution of $\sigma_p \mathfrak{n} + pM_{p-1}(a_0 - \mathbb{E}[a_0])$. Here $\mathfrak{n}$ is the standard normal distribution, independent of a_0, and M_{p-1} is the (p-1)-th moment in Theorem 3.1. Since $M_p = 0 \Longleftrightarrow p$ is odd, if a_0 is neither a constant a.s. nor Gaussian then the fluctuation is not Gaussian for odd p.

The remaining part of the paper is arranged as follows. Some integrals associated with pair partitions are defined in section 2. Sections 3, 4 and 5 are devoted to the proof of Theorem 1.1. We will extend our main result to other models closely related to Toeplitz matrices in section 6.

2 Integrals associated with pair partitions

In order to calculate the moments of the limiting distribution and the limiting covariance matrix of random variables ω_p, we first review some basic combinatorical concepts, and then define some integrals associated with pair partitions.

Definition 2.1. *Let the set $[n] = \{1, 2, \cdots, n\}$.*

(1) We call $\pi = \{V_1, \cdots, V_r\}$ a partition of $[n]$ if the blocks V_j $(1 \leq j \leq r)$ are pairwise disjoint, non-empty subsets of $[n]$ such that $[n] = V_1 \cup \cdots \cup V_r$. The number of blocks of π is denoted by $|\pi|$, and the number of elements of V_j is denoted by $|V_j|$.

(2) Without loss of generality, we assume that $V_1, \cdots, V_r$ have been arranged such that $s_1 < s_2 < \cdots < s_r$, where s_j is the smallest number of V_j. Therefore we can define the projection $\pi(i) = j$ if i belongs to the block V_j; furthermore for two elements p, q of $[n]$ we write $p \sim_\pi q$ if $\pi(p) = \pi(q)$.

(3) The set of all partitions of $[n]$ is denoted by $\mathcal{P}(n)$, and the subset consisting of all pair partitions, i.e. all $|V_j| = 2$, $1 \leq j \leq r$, is denoted by $\mathcal{P}_2(n)$. Note that $\mathcal{P}_2(n)$ is an empty set if n is odd.

(4) Suppose p, q are positive integers and $p+q$ is even, we denote a subset of $\mathcal{P}_2(p+q)$ by $\mathcal{P}_2(p, q)$, which consists of such pair partitions π: there exists $1 \leq i \leq p < j \leq p+q$ such that $i \sim_\pi j$ (we say that there is one crossing match in π).

(5) When p and q are both even, we denote a subset of $\mathcal{P}(p+q)$ by $\mathcal{P}_{2,4}(p, q)$, which consists of such partitions $\pi = \{V_1, \cdots, V_r\}$ satisfying

(i) $|V_j| = 2$, $1 \leq j \leq r, j \neq i$ and $|V_i| = 4$ for some i.

(ii) $V_j \subseteq \{1,2,\ldots,p\}$ or $\{p+1,p+2,\ldots,p+q\}$ *for* $1 \le j \le r, j \ne i$.

(iii) two elements of V_i *come from* $\{1,2,\ldots,p\}$ *and the other two come from* $\{p+1,p+2,\ldots,p+q\}$.

For other cases of p *and* q*, we assume* $\mathcal{P}_{2,4}(p,q)$ *is an empty set.*

Now we define several types of definite integrals associated with $\pi \in \mathcal{P}_2(p,q)$ or $\pi \in \mathcal{P}_{2,4}(p,q)$. For reader's convenience, we suggest omitting them for the moment and refer to them when needed in sections 3 and 4. Let the parameter $b \in [0,1]$.

First, for $\pi \in \mathcal{P}_2(p,q)$ we set

$$\epsilon_\pi(i) = \begin{cases} 1, & i \text{ is the smallest number of } \pi^{-1}(\pi(i)); \\ -1, & \text{otherwise.} \end{cases} \tag{2.1}$$

For every pair partition $\pi \in \mathcal{P}_2(p,q)$, we construct a projective relation between two groups of unknowns $y_1,\ldots,y_{p+q}$ and $x_1,\ldots,x_{\frac{p+q}{2}}$ as follows:

$$\epsilon_\pi(i)\, y_i = \epsilon_\pi(j)\, y_j = x_{\pi(i)} \tag{2.2}$$

whenever $i \sim_\pi j$. Thus, we have an identical equation

$$\sum_{j=1}^{p+q} y_j \equiv 0. \tag{2.3}$$

For $x_0, y_0 \in [0,1]$ and $x_1,\ldots,x_{\frac{p+q}{2}} \in [-1,1]$, we define two kinds of integrals with Type I by

$$f_I^-(\pi) = \\ \int_{[0,1]^2\times[-1,1]^{\frac{p+q}{2}}} \delta\left(\sum_{i=1}^p y_i\right) \prod_{j=1}^p \chi_{[0,1]}(x_0 + b\sum_{i=1}^j y_i) \prod_{j'=p+1}^{p+q} \chi_{[0,1]}(y_0 + b\sum_{i=p+1}^{j'} y_i)\, d\, y_0 \prod_{l=0}^{\frac{p+q}{2}} d\, x_l \tag{2.4}$$

and

$$f_I^+(\pi) = \\ \int_{[0,1]^2\times[-1,1]^{\frac{p+q}{2}}} \delta\left(\sum_{i=1}^p y_i\right) \prod_{j=1}^p \chi_{[0,1]}(x_0 + b\sum_{i=1}^j y_i) \prod_{j'=p+1}^{p+q} \chi_{[0,1]}(y_0 - b\sum_{i=p+1}^{j'} y_i)\, d\, y_0 \prod_{l=0}^{\frac{p+q}{2}} d\, x_l. \tag{2.5}$$

Here δ is the Dirac function and χ is the indicator function. Note that $\sum_{j=1}^p y_j \ne 0$ by the definition of $\mathcal{P}_2(p,q)$, therefore the above integrals are multiple integrals in $(\frac{p+q}{2}+1)$ variables.

Next, for $\pi = \{V_1,\ldots,V_{\frac{p+q}{2}-1}\} \in \mathcal{P}_{2,4}(p,q)$ (denoting the block with four elements by V_i), we set for $\pi(k) \ne i$

$$\tau_\pi(k) = \begin{cases} 1, & k \text{ is the smallest number of } \pi^{-1}(\pi(k)); \\ -1, & \text{otherwise} \end{cases} \tag{2.6}$$

while for $\pi(k) = i$

$$\tau_\pi(k) = \begin{cases} 1, & k \text{ is the smallest or largest number of } \pi^{-1}(\pi(k)); \\ -1, & \text{otherwise.} \end{cases} \tag{2.7}$$

To every partition $\pi \in \mathcal{P}_{2,4}(p,q)$, we construct a projective relation between two groups of unknowns $y_1,\ldots,y_{p+q}$ and $x_1,\ldots,x_{\frac{p+q}{2}-1}$ as follows:

$$\tau_\pi(i)\, y_i = \tau_\pi(j)\, y_j = x_{\pi(i)} \tag{2.8}$$

whenever $i \sim_\pi j$. Then two kinds of integrals with Type II are defined respectively by

$$f_{II}^{-}(\pi) = \int_{[0,1]^2\times[-1,1]^{\frac{p+q}{2}-1}} \prod_{j=1}^{p} \chi_{[0,1]}(x_0 + b\sum_{i=1}^{j} y_i) \prod_{j'=p+1}^{p+q} \chi_{[0,1]}(y_0 + b\sum_{i=p+1}^{j'} y_i)\, d\,y_0 \prod_{l=0}^{\frac{p+q}{2}-1} d\,x_l \tag{2.9}$$

and

$$f_{II}^{+}(\pi) = \int_{[0,1]^2\times[-1,1]^{\frac{p+q}{2}-1}} \prod_{j=1}^{p} \chi_{[0,1]}(x_0 + b\sum_{i=1}^{j} y_i) \prod_{j'=p+1}^{p+q} \chi_{[0,1]}(y_0 - b\sum_{i=p+1}^{j'} y_i)\, d\,y_0 \prod_{l=0}^{\frac{p+q}{2}-1} d\,x_l.$$

3 Mathematical expectation

In this section, we will review some results about the moments of the limiting distribution of eigenvalues in [21], for the convenience of the readers and further discussion.

Theorem 3.1. *$\mathbb{E}[\frac{1}{n}\mathrm{tr}(A_n^{2k})] = M_{2k} + o(1)$ and $\mathbb{E}[\frac{1}{n}\mathrm{tr}(A_n^{2k+1})] = o(1)$ as $n \longrightarrow \infty$ where*

$$M_{2k} = \sum_{\pi\in\mathcal{P}_2(2k)} \int_{[0,1]\times[-1,1]^k} \prod_{j=1}^{2k} \chi_{[0,1]}(x_0 + b\sum_{i=1}^{j} \epsilon_\pi(i)\, x_{\pi(i)}) \prod_{l=0}^{k} \mathrm{d}\, x_l. \tag{3.1}$$

Let us first give a lemma about traces of Toeplitz band matrices. Although its proof is simple, it is very useful in treating random matrix models closely related to Toeplitz matrices.

Lemma 3.2. *For Toeplitz band matrices $T_{l,n} = (\eta_{ij}\, a_{l,i-j})_{i,j=1}^n$ with the bandwidth b_n where $a_{l,-n+1},\cdots,a_{l,n-1}$ are complex numbers and $l = 1,\ldots,p$, we have the trace formula*

$$\mathrm{tr}(T_{1,n}\cdots T_{p,n}) = \sum_{i=1}^{n} \sum_{J} a_J\, I_J\, \delta_{0,\sum_{l=1}^{p} j_l}, \qquad p \in \mathbb{N}. \tag{3.2}$$

Here $J = (j_1,\ldots,j_p) \in \{-b_n,\ldots,b_n\}^p$, $a_J = \prod_{l=1}^{p} a_{l,j_l}$, $I_J = \prod_{k=1}^{p} \chi_{[1,n]}(i + \sum_{l=1}^{k} j_l)$ and the summation $\sum_J$ runs over all possibilities that $J \in \{-b_n,\ldots,b_n\}^p$.

Proof. For the standard basis $\{e_1,\cdots,e_n\}$ of the Euclidean space $\mathbb{R}^n$, we have

$$T_{p,n}\, e_i = \sum_{j=0}^{b_n} a_{p,-j} B^j\, e_i + \sum_{j=1}^{b_n} a_{p,j} F^j\, e_i = \sum_{j=-b_n}^{b_n} a_{p,j}\, \chi_{[1,n]}(i+j)\, e_{i+j}.$$

Repeating $T_{l,n}$'s effect on the basis, we have

$$T_{1,n}\cdots T_{p,n}\, e_i = \sum_{j_1,\cdots,j_p=-b_n}^{b_n} \prod_{l=1}^{p} a_{l,j_l} \prod_{k=1}^{p} \chi_{[1,n]}(i + \sum_{l=1}^{k} j_l)\, e_{i+\sum_{l=1}^{p} j_l}.$$

By $\mathrm{tr}(T_{1,n}\cdots T_{p,n}) = \sum_{i=1}^{n} e_i^t\, T_{1,n}\cdots T_{p,n}\, e_i$, we complete the proof. □

We will mainly use the above trace formula in the case where $T_{1,n} = \cdots = T_{p,n} = T_n$. Since $a_0 \equiv 0$, from the Kronecker delta symbol in the trace formula of (3.2), it suffices to consider these $J = (j_1, \ldots, j_p) \in \{\pm 1, \ldots, \pm b_n\}^p$ with the addition of $\sum_{k=1}^{p} j_k = 0$. We remark that $a_0 \equiv 0$ is not necessary to Theorem 3.1. In fact it is sufficient to ensure Theorem 3.1 if all finite moments of random variable a_0 exist and its expectation is zero.

Definition 3.3. *Let $J = (j_1, \ldots, j_p) \in \{\pm 1, \ldots, \pm b_n\}^p$, we say J is balanced if $\sum_{k=1}^{p} j_k = 0$. The component j_u of J is said to be coincident with j_v if $|j_u| = |j_v|$ for $1 \le u \ne v \le p$.*

For $J \in \{\pm 1, \ldots, \pm b_n\}^p$, we construct a set of numbers with multiplicities

$$S_J = \{|j_1|, \ldots, |j_p|\}. \tag{3.3}$$

We call S_J the *projection* of J.

The balanced J's can be classified into three categories.

Category 1 (denoted by $\Gamma_1(p)$): J is said to belong to category 1 if each of its components is coincident with exactly one other component of the opposite sign. It is obvious that $\Gamma_1(p)$ is an empty set when p is odd.

Category 2 ($\Gamma_2(p)$) consists of all those vectors such that S_J has at least one number with multiplicity 1.

Category 3 ($\Gamma_3(p)$) consists of all other balanced vectors in $\{\pm 1, \ldots, \pm b_n\}^p$. For $J \in \Gamma_3(p)$, either S_J has one number with multiplicity at least 3, or each of S_J has multiplicity 2 but at least two of the components are the same, which are denoted respectively by $\Gamma_{31}(p)$ and $\Gamma_{32}(p)$.

We are now ready to prove Theorem 3.1.

Proof of Theorem 3.1. By Lemma 3.2, we have

$$\mathbb{E}[\frac{1}{n}\mathrm{tr}(A_n^{2k})] = \frac{1}{nb_n^k}\sum_{i=1}^{n}\sum_{J}\mathbb{E}[a_J]\, I_J\, \delta_{0,\sum_{l=1}^{2k} j_l} = \sum\nolimits_1 + \sum\nolimits_2 + \sum\nolimits_3, \tag{3.4}$$

where

$$\sum\nolimits_l = \frac{1}{nb_n^k}\sum_{i=1}^{n}\sum_{J\in\Gamma_l(2k)}\mathbb{E}[a_J]\, I_J, \quad l = 1, 2, 3. \tag{3.5}$$

By the definition of the categories and the assumptions on the entries of the random matrices, we obtain

$$\sum\nolimits_2 = 0.$$

Next, we divide $\sum_3$ into two parts

$$\sum\nolimits_3 = \sum\nolimits_{31} + \sum\nolimits_{32},$$

where

$$\sum\nolimits_{3l} = \frac{1}{nb_n^k}\sum_{i=1}^{n}\sum_{J\in\Gamma_{3l}(2k)}\mathbb{E}[a_J]\, I_J, \quad l = 1, 2. \tag{3.6}$$

For $J \in \Gamma_3(2k)$, we denote the number of distinct elements of S_J by t. By the definition of the category, we have $t \le k$. Note that the random variables whose subscripts have different absolute values are independent. Once we have specified the distinct numbers of S_J, the subscripts $j_1, \cdots, j_{2k}$ are determined in at most $2^{2k}k^{2k}$ ways. If $J \in \Gamma_{31}(2k)$, then $t \leq \frac{2k-1}{2}$. Again by independence and the assumptions on the matrix elements (1.7), we find

$$|\sum\nolimits_{31}| \leq \frac{1}{nb_n^k}\sum_{i=1}^{n} O(b_n^{\frac{2k-1}{2}}) = o(1).$$

When $J \in \Gamma_{32}(2k)$, there exist $p_0, q_0 \in [2k]$ such that

$$j_{p_0} = j_{q_0} = \frac{1}{2}(j_{p_0} + j_{q_0} - \sum_{q=1}^{2k} j_q).$$

We can choose the other at most $k-1$ distinct numbers, which determine $j_{p_0} = j_{q_0}$. This shows that there is a loss of at least one degree of freedom, thus the contribution of such terms is $O(b_n^{-1})$, i.e.

$$|\sum\nolimits_{32}| = o(1).$$

Since the main contribution comes from the category 1, each term $\mathbb{E}[a_J] = 1$ for $J \in \Gamma_1(2k)$. So we can rewrite

$$\mathbb{E}[\frac{1}{n}\mathrm{tr}(A_n^{2k})] = o(1) + \frac{1}{nb_n^k}\sum_{\pi \in \mathcal{P}_2(2k)}\sum_{i=1}^{n}\sum_{j_1,\cdots,j_k=-b_n}^{b_n}\prod_{l=1}^{2k}\chi_{[1,n]}(i + \sum_{q=1}^{l}\epsilon_\pi(q)j_{\pi(q)}). \tag{3.7}$$

For fixed $\pi \in \mathcal{P}_2(2k)$,

$$\frac{1}{nb_n^k}\sum_{i=1}^{n}\sum_{j_1,\cdots,j_k=-b_n}^{b_n}\prod_{l=1}^{2k}\chi_{[1,n]}(i + \sum_{q=1}^{l}\epsilon_\pi(q)j_{\pi(q)}),$$

i.e.

$$\frac{1}{nb_n^k}\sum_{i=1}^{n}\sum_{j_1,\cdots,j_k=-b_n}^{b_n}\prod_{l=1}^{2k}\chi_{[\frac{1}{n},1]}(\frac{i}{n} + \sum_{q=1}^{l}\epsilon_\pi(q)\frac{b_n}{n}\frac{j_{\pi(q)}}{b_n})$$

can be considered as a Riemann sum of the definite integral

$$\int_{[0,1]\times[-1,1]^k}\prod_{l=1}^{2k} I_{[0,1]}(x_0 + b\sum_{q=1}^{l}\epsilon_\pi(q)x_{\pi(q)})\prod_{l=0}^{k}\mathrm{d}\,x_l.$$

As in the above arguments, by Lemma 3.2 and the assumptions on the matrix elements (1.7), we have

$$\left|\mathbb{E}[\frac{1}{n}\mathrm{tr}(A_n^{2k+1})]\right| \leq \frac{1}{nb_n^{\frac{2k+1}{2}}}\sum_{i=1}^{n} O(b_n^k) = o(1) \tag{3.8}$$

since $\mathcal{P}_2(2k-1) = \emptyset$.

This completes the proof. □

Remark 3.4. *When $b = 0$, we can easily get $M_{2k} = 2^k(2k-1)!!$. This is just the $2k$-moment of the normal distribution with variance 2, which is also obtained independently by Basak and Bose [4] and Kargin [19]. However, for $b > 0$ it is quite difficult to calculate M_{2k} because the integrals in the sum of (3.1) are not all the same for different partitions π's, some of which are too hard to evaluate.*

4 covariance

In this section we evaluate the covariance of ω_p and ω_q. Recall

$$\omega_p = \frac{\sqrt{b_n}}{n}\left(\mathrm{tr}(A_n^p) - \mathbb{E}[\mathrm{tr}(A_n^p)]\right) = \frac{1}{n} b_n^{-\frac{p-1}{2}} \sum_{i=1}^{n} \sum_{J} I_J \left(a_J - \mathbb{E}[a_J]\right) \delta_{0, \sum_{l=1}^{p} j_l}, \tag{4.1}$$

thus

$$\begin{aligned}\mathbb{E}[\omega_p\, \omega_q] &= \frac{1}{n^2} b_n^{-\frac{p+q}{2}+1} \sum_{i,i'} \sum_{J,J'} I_J\, I_{J'} \mathbb{E}[(a_J - \mathbb{E}[a_J])\,(a_{J'} - \mathbb{E}[a_{J'}])] \\ &= \frac{1}{n^2} b_n^{-\frac{p+q}{2}+1} \sum_{i,i'} \sum_{J,J'} I_J\, I_{J'} \left(\mathbb{E}[a_J a_{J'}] - \mathbb{E}[a_J]\mathbb{E}[a_{J'}]\right), \end{aligned} \tag{4.2}$$

where the summation $\sum\limits_{J,J'}$ runs over all balanced vectors $J \in \{\pm 1, \dots, \pm b_n\}^p$ and $J' \in \{\pm 1, \dots, \pm b_n\}^q$.

The main result of this section can be stated as follows:

Theorem 4.1. *Using the notations in section 2, for fixed $p,\ q \geq 2$, as $n \longrightarrow \infty$ we have*

$$\mathbb{E}[\omega_p\, \omega_q] \longrightarrow \sigma_{p,q} = \sum_{\pi \in \mathcal{P}_2(p,q)} \left(f_I^-(\pi) + f_I^+(\pi)\right) + (\kappa - 1) \sum_{\pi \in \mathcal{P}_{2,4}(p,q)} \left(f_{II}^-(\pi) + f_{II}^+(\pi)\right) \tag{4.3}$$

when $p+q$ is even and

$$\mathbb{E}[\omega_p\, \omega_q] = o(1) \tag{4.4}$$

when $p+q$ is odd.

When $p = q$ we denote $\sigma_{p,q}$ by σ_p^2, where σ_p denotes the standard deviation. From the theorem above, we obtain the variance of ω_Q in Theorem 1.1

$$\sigma_Q^2 = \sum_{i=2}^{p} \sum_{j=2}^{p} q_i q_j \sigma_{i,j}. \tag{4.5}$$

By the independence of matrix entries, the only non-zero terms in the sum of (4.2) come from pairs of balanced vectors $J = (j_1, \dots, j_p)$ and $J' = (j'_1, \dots, j'_q)$ such that

(i) The *projections* S_J and $S_{J'}$ of J and J' have at least one element in common;

(ii) Each number in the union of S_J and $S_{J'}$ occurs at least two times.

Definition 4.2. *Any ordered pair of balanced vectors $J = (j_1, \dots, j_p)$ and $J' = (j'_1, \dots, j'_q)$ satisfying (i) above is called correlated. If $j_u \in J$ (j_u is a component of J) and $|j_u| \in S_J \bigcap S_{J'}$, then j_u is called a joint point of the ordered correlated pair.*

To observe which correlated pairs lead to the main contribution to the covariance, we next construct a balanced vector of dimension $(p+q-2)$ from each correlated pair J of dimension p and J' of dimension q. Although the corresponding map of correlated pairs to such balanced vectors is not one to one, the number of pre-images for a balanced vector of dimension $(p+q-2)$ is finite (only depending on p and q). We will study the resulting balanced vectors in a similar way as in section 3.

Proof of Theorem 4.1. Let us first construct a map from the ordered correlated pair J and J' as follows. Let $j_u \in J$ be the first joint point (whose subscript is the smallest) of the ordered correlated pair J and J', and let j'_v be the first element in J' such that $|j_u| = |j'_v|$. If $j_u = -j'_v$, we construct a vector $L = (l_1, \dots, l_{p+q-2})$ such that

$$l_1 = j_1, \dots, l_{u-1} = j_{u-1}, l_u = j'_1, \dots, l_{u+v-2} = j'_{v-1},$$
$$l_{u+v-1} = j'_{v+1}, \dots, l_{u+q-2} = j'_q, l_{u+q-1} = j_{u+1}, \dots, l_{p+q-2} = j_p.$$

It is obvious that

$$\sum_{k=1}^{p+q-2} l_k = 0,$$

so L is balanced. If $j_u = j'_v$, then from J and $-J' = (-j'_1, \dots, -j'_q)$ we proceed as in the same way above. We call this process of constructing L from J and J' a *reduction* step and denote it by $L = J\bigvee_{|j_u|} J'$.

Remark 4.3. *From the construction above, for any joint point of J and J' a reduction step can be done in the same way. Given $\theta \in S_J \bigcap S_{J'}$, when saying $J\bigvee_\theta J'$, we mean that there exists some joint point of J and J' j_u satisfying $|j_u| = \theta$ and $J\bigvee_\theta J'$ is the vector after this reduction step. In this section, j_u is always the first joint point. While in section 5, j_u may denote other joint points which is clear in the context.*

Notice that the reduction might cause the appearance of one number with multiplicity 1 in S_L, although each number in the union of S_J and $S_{J'}$ occurs at least two times. If so, the resulting number with multiplicity 1 in S_L must be coincident with the joint point j_u. In addition, to estimate which terms lead to main contribution to higher moments of $\mathrm{tr}(A_n^p)$, we will use the *reduction* steps and mark the appearance of the numbers with multiplicity 1 in section 5.

Next, assume we have a balanced vector L of dimension $(p+q-2)$, we shall estimate in how many different ways it can be obtained from correlated pairs of dimensions p and q. First, we have to choose some component l_u in the first half of the vector, $1 \le u \le p$ such that

$$\left|\sum_{i=u}^{u+q-2} l_i\right| \neq |l_j|, \quad j = 1, \dots, u-1. \tag{4.6}$$

Set $J = (j_1, \dots, j_p)$ with

$$j_1 = l_1, \dots, j_{u-1} = l_{u-1}, j_u = \sum_{i=u}^{u+q-2} l_i, j_{u+1} = l_{u+q-1}, \dots, j_p = l_{p+q-2}. \tag{4.7}$$

We also have to choose some component l_{u+v-1}, $1 \le v \le q-1$ such that

$$\left|\sum_{i=u}^{u+q-2} l_i\right| \neq |l_j|, \quad j = u, \dots, u+v-2 \tag{4.8}$$

whenever $v \ge 2$. Set $J' = (j'_1, \dots, j'_q)$ with

$$j'_1 = l_u, \dots, j'_{v-1} = l_{u+v-2}, j'_v = -\sum_{i=u}^{u+q-2} l_i, j'_{v+1} = l_{u+v-1}, \dots, j'_p = l_{u+q-2}. \tag{4.9}$$

If j_u is the joint point of the constructed correlated pair J and J' and j'_v is the corresponding element in J', then the pair $\{J, J'\}$ or $\{J, -J'\}$ is the pre-image of L. Note that since when $u = v = 1$ the conditions (4.6) and (4.8) are satisfied, the pre-image of L always exists. A simple estimation shows that the number of pre-images of L is at most $2pq$, not depending on n (we will see this fact plays an important role in the estimation of higher moments in section 5).

There is at most one element with multiplicity 1 in S_L. If so, this number will be determined by others because of the balance of L. Consequently, the degree of freedom for such terms is at most $\frac{p+q-2-1}{2}$. Therefore, the sum of these terms will be $O(b_n^{-1/2})$, which can be omitted. Now we suppose each number in S_L occurs at least two times. Recall the procedure in section 3, and we know that the main contribution to the covariance (4.2) comes from the $L \in \Gamma_1(p+q-2)$, which implies $\mathbb{E}[\omega_p\,\omega_q] = o(1)$ when $p+q$ is odd. When $p+q$ is even, for $L \in \Gamma_1(p+q-2)$ the weight

$$\mathbb{E}[a_J a_{J'}] - \mathbb{E}[a_J]\mathbb{E}[a_{J'}] = \mathbb{E}[\prod_{s=1}^{p} a_{j_s} \prod_{t=1}^{q} a_{j'_t}] - \mathbb{E}[\prod_{s=1}^{p} a_{j_s}]\mathbb{E}[\prod_{t=1}^{q} a_{j'_t}] \tag{4.10}$$

equals to 1 if j_u is not coincident with any component of L; otherwise the weight is either $\mathbb{E}[|a_{j_u}|^4] = \kappa$ or $\mathbb{E}[|a_{j_u}|^4] - \big(\mathbb{E}[|a_{j_u}|^2]\big)^2 = \kappa - 1$.

So far we have found the terms leading to the main contribution, now we calculate the variance. Based on whether or not the fourth moment appears, we evaluate the covariance. If the fourth moment doesn't appear, then $j_1, \ldots, j_p, j'_1, \ldots, j'_q$ match in pairs. By their subscripts they can be treated as pair partitions of $\{1, 2, \ldots, p, p+1, \ldots, p+q\}$ but with at least one crossing match (i.e., $\mathcal{P}_2(p,q)$ as in section 2). Thus, for every $\pi \in \mathcal{P}_2(p,q)$, the summation can be a Riemann sum and its limit becomes $f_I^-(\pi)$ (it is $f_I^+(\pi)$ when the first coincident components in J and J' have the same sign). On the other hand, if the fourth moment does appear, then $j_1, \ldots, j_p, j'_1, \ldots, j'_q$ match in pairs except that there exists a block with four elements. Therefore, from the balance of $\sum_{k=1}^{p} j_k = 0$ and $\sum_{k=1}^{q} j'_k = 0$, we know that the main contribution must come from such partitions: $j_1, \ldots, j_p$ and $j'_1, \ldots, j'_q$ both form pair partitions; the block with four elements take respectively from a pair of $j_1, \ldots, j_p$ and $j'_1, \ldots, j'_q$. Otherwise, the degree of freedom decreases by at least one. Similarly, for every $\pi \in \mathcal{P}_{2,4}(p,q)$, the corresponding summation can be a Riemann sum and its limit becomes $f_{II}^-(\pi)$ (it is $f_{II}^+(\pi)$ when the first coincident components in J and J' have the same sign).

In a similar way as in section 3, noting that the coincident components in J and J' may have the same or opposite sign, we conclude with the notations in section 2 that

$$\mathbb{E}[\omega_p\,\omega_q] \longrightarrow \sum_{\pi \in \mathcal{P}_2(p,q)} \big(f_I^-(\pi) + f_I^+(\pi)\big) + (\kappa - 1) \sum_{\pi \in \mathcal{P}_{2,4}(p,q)} \big(f_{II}^-(\pi) + f_{II}^+(\pi)\big)$$

as $n \longrightarrow \infty$.

This completes the proof. $\square$

5 Higher Moments of $\mathrm{tr}(A_n^p)$

Let $\mathfrak{B}_{n,p}$ denote the set of all balanced vectors $J = (j_1, \ldots, j_p) \in \{\pm 1, \ldots, \pm b_n\}^p$. Let $\mathfrak{B}_{n,p,i}$ ($1 \le i \le n$) be a subset of $\mathfrak{B}_{n,p}$ such that $J \in \mathfrak{B}_{n,p,i}$ if and only if

$$\forall\, t \in \{1, \ldots, p\}, \quad 1 \le i + \sum_{q=1}^{t} j_q \le n.$$

With these notations, Lemma 3.2 can be rewritten as

$$\mathrm{tr}(T_n^p) = \sum_{i=1}^{n} \sum_{J \in \mathfrak{B}_{n,p,i}} a_J. \tag{5.1}$$

To finish the proof of Theorem 1.1, it is sufficient to show that given $p_1, p_2, \cdots, p_l \geq 2$ and $l \geq 1$, as $n \to \infty$ we have

$$\mathbb{E}[\omega_{p_1}\omega_{p_2}\cdots\omega_{p_l}] \longrightarrow \mathbb{E}[g_{p_1}g_{p_2}\cdots g_{p_l}], \tag{5.2}$$

where $\{g_p\}_{p\geq 2}$ is a centered Gaussian family with covariances $\sigma_{p,q} = \mathbb{E}[g_p g_q]$.

Then a CLT for

$$\omega_Q = \frac{\sqrt{b_n}}{n}\left(\mathrm{tr}Q(A_n) - \mathbb{E}[\mathrm{tr}Q(A_n)]\right) \tag{5.3}$$

follows, with the variance

$$\sigma_Q^2 = \sum_{i=2}^{p}\sum_{j=2}^{p} q_i q_j \sigma_{i,j}. \tag{5.4}$$

The main idea is rather straightforward: in an analogous way to the one used in Eq. (4.2), we will deal with

$$\begin{aligned}&\mathbb{E}[\omega_{p_1}\cdots\omega_{p_l}] = \\ &n^{-l}\cdot b_n^{-\frac{p_1+\cdots+p_l-l}{2}} \sum_{i_1,\ldots,i_l=1}^{n} \sum_{J_1\in\mathfrak{B}_{n,p_1,i_1},\ldots,J_l\in\mathfrak{B}_{n,p_l,i_l}} \mathbb{E}\Big[\prod_{t=1}^{l}(a_{J_t} - \mathbb{E}[a_{J_t}])\Big].\end{aligned} \tag{5.5}$$

Remember that two balanced vectors $J = (j_1, \ldots, j_p)$ and $J' = (j'_1, \ldots, j'_q)$ are called correlated if the corresponding *projections* S_J and $S_{J'}$ of J and J' have at least one element in common.

Definition 5.1. *Given a set of balanced vectors $\{J_1, J_2, \ldots, J_l\}$, a subset of balanced vectors $J_{m_{j_1}}$,$J_{m_{j_2}}$,...,$J_{m_{j_t}}$ is called a cluster if*
1) for any pair J_{m_i},J_{m_j} from the subset one can find a chain of vectors J_{m_s}, also belonging to the subset, which starts with J_{m_i} ends with J_{m_j}, such that any two neighboring vectors are correlated;
2) the subset $\{J_{m_{j_1}}, J_{m_{j_2}}, \ldots, J_{m_{j_t}}\}$ cannot be enlarged with the preservation of 1).

It is clear that the vectors corresponding to different clusters are disjoint. By this reason the mathematical expectation in (5.5) decomposes into the product of mathematical expectations corresponding to different clusters. We will show that the leading contribution to (5.5) comes from products where all clusters consist exactly of two vectors, as is stated in Lemma 5.2 below.

Lemma 5.2. *Provided $l \geq 3$, we have*

$$n^{-l}\cdot b_n^{-\frac{p_1+\cdots+p_l-l}{2}} \sum_{i_1,\ldots,i_l=1}^{n} \sum_{J_1\in\mathfrak{B}_{n,p_1,i_1},\ldots,J_l\in\mathfrak{B}_{n,p_l,i_l}}{}^{\star} \mathbb{E}\Big[\prod_{t=1}^{l}(a_{J_t} - \mathbb{E}[a_{J_t}])\Big] = o(1) \tag{5.6}$$

where the sum $\sum^{\star}$ in (5.6) is taken over l vectors which exactly form a $cluster$.

For fixed p, all the involved moments no higher than p are $O(1)$ because of uniform boundedness of matrix entries. On the other hand, $0 \leq \prod_{t=1}^{l} I_{J_t} \leq 1$. So to prove Lemma 5.2, we just need to count the number of terms in (5.6). As before, to complete the estimation it suffices to replace $\mathfrak{B}_{n,p,i}$ by $\mathfrak{B}_{n,p}$. That is, it suffices to prove

$$b_n^{-\frac{p_1+\cdots+p_l-l}{2}} \sum_{J_1\in\mathfrak{B}_{n,p_1},\ldots,J_l\in\mathfrak{B}_{n,p_l}}{}^{\star} 1 = o(1) \tag{5.7}$$

where the summation $\sum^{\star}$ is taken over l vectors which exactly form a $cluster$.

Instead of Lemma 5.2, we will prove

Lemma 5.3. *Provided $l \geq 3$, $\mathbf{p} = (p_1, \ldots, p_l)$ with positive integers $p_1, \ldots, p_l \geq 2$. Let $\mathfrak{B}_{\mathbf{p}}$ be a subset of the Cartesian product $\mathfrak{B}_{n,p_1} \times \cdots \times \mathfrak{B}_{n,p_l}$ such that $(J_1, J_2, \ldots, J_l) \in \mathfrak{B}_{\mathbf{p}}$ if and only if*
(i) any element in $\bigcup_{i=1}^{l} S_{J_i}$ has at least multiplicity two in the union;
(ii) $J_1, J_2, \ldots, J_l$ make a cluster*;*
(iii) $0 \notin \bigcup_{i=1}^{l} S_{J_i}$.
Then we claim that

$$card(\mathfrak{B}_{\mathbf{p}}) = o(b_n^{\frac{p_1+p_2+\cdots+p_l-l}{2}}). \tag{5.8}$$

Notice that we list the condition (iii) which looks redundant from $J \in \{\pm 1, \ldots, \pm b_n\}^p$ to emphasize the importance of 0 (i.e., the diagonal matrix entry a_0). In fact, if $p_1, p_2, \cdots, p_l$ are all even, even for these $J \in \{0, \pm 1, \ldots, \pm b_n\}^p$ under condition (i) and (ii), the estimation above is true.

Proof of Lemma 5.3. The intuitive idea of the proof is as follows: from condition (i) in Lemma 5.3, regardless of the correlating condition, the cardinality of $\mathfrak{B}_{\mathbf{p}}$(denoted by card$(\mathfrak{B}_{\mathbf{p}})$ for short) is $O(b_n^{\frac{p_1+p_2+\cdots+p_l}{2}})$. We can say that the freedom degree is $\frac{p_1+p_2+\cdots+p_l}{2}$. But each correlation means two vectors share a common element so that it will decrease the freedom degree by one. To form a *cluster* we need $l-1$ correlations. So card$(\mathfrak{B}_{\mathbf{p}}) = O(b_n^{\frac{p_1+p_2+\cdots+p_l}{2}-(l-1)})$. When $l \geq 3$, $l-1 > \frac{l}{2}$. Thus we obtain the desired estimation in Lemma 5.3. However, adding one correlation does not necessarily lead the freedom degree to decrease by one. Some may be redundant. So we have to make use of the correlations more efficiently.

As in section 4, we do the *reduction* steps as long as the structure of the $cluster$ is preserved. To say precisely, we start from checking J_1 and J_2. If j_u is the first joint point of J_1 and J_2 satisfying the condition that $J_1 \bigvee_{|j_u|} J_2$ can still form a *cluster* with the other vectors, then we do this $reduction$ step. If this kind of j_u does not exist, we turn to check J_1 and J_3 in the same way. After each $reduction$ step, we have a new cluster of vectors $\widetilde{J_1}, \widetilde{J_2}, \ldots, \widetilde{J_{\tilde{l}}}$. We continue to check $\widetilde{J_1}$ and $\widetilde{J_2}$ as before. If we cannot do any $reduction$ step, we stop.

Suppose that we did m *reduction* steps in total. Then we have a new $cluster$ of vectors $J'_1, J'_2, \ldots, J'_{l'}$ and the dimension of J'_i is p'_i for $1 \leq i \leq l'$. From the reduction process, for any $\theta \in S_{J'_{\alpha_1}} \bigcap S_{J'_{\alpha_2}}$, $J'_{\alpha_1} \bigvee_\theta J'_{\alpha_2}$ cannot form a *cluster* with the other vectors. The resulting $cluster$ still satisfies condition (ii) in Lemma 5.3. However, the condition (i) may fail because the joint point of a pair of correlated vectors can be a tripartite one, thus after a *reduction* step its multiplicity becomes one.

Note that after a $reduction$ step the number of pre-images of the resulting vector only depends on the dimensions of the involving vectors, not depending on n. Thus we only need to estimate the degree of freedom of the reduced vectors $J'_1, J'_2, \ldots, J'_{l'}$.

Since after one *reduction* step the total dimension of vectors will decrease by two and the number of vectors will decrease by one, we have

$$\sum_{i=1}^{l} p_i = \sum_{i=1}^{l'} p'_i + 2m \tag{5.9}$$

and

$$l = l' + m. \tag{5.10}$$

Denote by l_0 the number of single multiplicity elements in $\bigcup_{i=1}^{l'} S_{J'_i}$. Since one *reduction* step will add at most one element with single multiplicity, therefore

$$l_0 \leq m. \tag{5.11}$$

Below, we will proceed according to two cases: $l' > 1$ and $l' = 1$.

In the case $l' > 1$

To complete the proof of this case, we need some definitions and notations.

Let $\mathcal{U}$ be the set consisting of all elements which belong to at least two of $S_{J'_1}, \ldots, S_{J'_{l'}}$, i.e. $\mathcal{U} = \{\theta | \exists i \neq j \quad \text{s.t.} \quad \theta \in S_{J'_i} \bigcap S_{J'_j}\}$. Since $l' > 1$ and $J'_1, \ldots, J'_{l'}$ form a cluster, $\mathcal{U} \neq \emptyset$. For any $\theta \in \bigcup_{i=1}^{l'} S_{J'_i}$, set $H_\theta =: \{J'_i | \, \theta \text{ or } -\theta \in J'_i, \quad 1 \leq i \leq l'\}$ and denote the number of vectors in H_θ by $h_\theta = \mathrm{card}(H_\theta)$. Obviously, $\{J'_i | i \leq l'\} = \bigcup_{\theta \in \mathcal{U}} H_\theta$.

We notice the following three facts.

Fact 1: For any $\theta \in \mathcal{U}$, $h_\theta \geq 3$. In fact, from the definition of $\mathcal{U}$, $h_\theta \geq 2$. If $h_\theta = 2$, the two vectors in H_θ can still be reduced to one vector and the reduction doesn't affect their connection with the other vectors, which is a contradiction with our assumption that $J'_1, J'_2, \ldots, J'_{l'}$ cannot be reduced.

Fact 2: For any $\theta \in \mathcal{U}$ and $J'_i \in H_\theta$, the multiplicity of θ in $S_{J'_i}$ is one . Otherwise, there are two vectors belonging to H_θ, for example, $J'_1, J'_2 \in H_\theta$ but $S_{J'_1}$ has two θ's, then $J'_1 \bigvee_\theta J'_2$ can be a $reduction$ step and $J'_1 \bigvee_\theta J'_2, J'_3, \ldots, J'_{l'}$ still form a cluster.

Fact 3: For any different elements θ and γ in $\mathcal{U}$, $\mathrm{card}(H_\theta \bigcap H_\gamma) \leq 1$. Otherwise, suppose J'_i and J'_j belong to $H_\theta \bigcap H_\gamma$ and J'_k is an element of H_θ other than J'_i and J'_j. From Fact 1, J'_k must exist. Now $J'_i \bigvee_\theta J'_k$ can form a $reduction$ step since $J'_i \bigvee_\theta J'_k$ can be correlated with J'_j by γ and other correlations won't be broken.

Definition 5.4. $\mathcal{V} \subset \mathcal{U}$ *is called a* dominating set *of* $\{J'_1, \ldots, J'_{l'}\}$ *if* $\{J'_i | 1 \leq i \leq l'\} = \bigcup_{\theta \in \mathcal{V}} H_\theta$.

Choose a minimal *dominating set* denoted by $\mathcal{U}_0$, which means that any proper subset of $\mathcal{U}_0$ is not a *dominating set*. Since $\mathcal{U}$ is a finite set, $\mathcal{U}_0$ must exist. Let $\mathcal{U}_0 = \{\theta_i | 1 \leq i \leq t\}$. For any $1 \leq j \leq t$, since $\mathcal{U}_0 \backslash \{\theta_j\}$ is not a *dominating set* , there exists $J'_{k_j} \in H_{\theta_j} \backslash \bigcup_{i \neq j} H_{\theta_i}$. Once we have already known the elements of $\bigcup_{i=1}^{l'} S_{J'_i} \backslash \mathcal{U}_0$, we know all the elements in $S_{J'_{k_j}}$ other than θ_j, thus θ_j will be determined by the balance of J'_{k_j}.

Set

$$h = \sum_{i=1}^{t} h_{\theta_i},$$

then the different way of choice of $\bigcup_{i=1}^{l'} S_{J'_i} \backslash \mathcal{U}_0$ is $O(b_n^{\frac{\sum_{i=1}^{l'} p'_i - l_0 - h}{2} + l_0})$. From Eqs. (5.9) and (5.10), we have

$$\frac{\sum_{i=1}^{l'} p'_i - l_0 - h}{2} + l_0 = \frac{(\sum_{i=1}^{l} p_i - l) - (m - l_0) - (h - l')}{2}. \tag{5.12}$$

Note that $m - l_0 \geq 0$ from Eq. (5.11). If $h > l'$, we have

$$O(b_n^{\frac{(\sum_{i=1}^{l} p_i - l) - (m - l_0) - (h - l')}{2}}) = o(b_n^{\frac{\sum_{i=1}^{l} p_i - l}{2}}). \tag{5.13}$$

If $h = l'$ and $t = 1$, the analysis is easy but a little complex. We will deal with it later.

Now we focus on the situation that $h = l'$ and $t > 1$. In this case, since $\{J'_i | i \leq l'\} = \bigcup_{\theta \in \mathcal{U}_0} H_\theta$, $l' \leq \sum_{i=1}^{t} \mathrm{card}(H_{\theta_i}) = h$ and the equity is true iff $H_{\theta_i} \bigcap H_{\theta_j} = \emptyset$ for any $1 \leq i \neq j \leq t$. Because $J'_1, J'_2, \ldots, J'_{l'}$ form a cluster, without lost of generality, we can assume that $(\bigcup_{J' \in H_{\theta_1}} S_{J'}) \bigcap (\bigcup_{J' \in H_{\theta_2}} S_{J'}) \neq \emptyset$. Thus there exist $J'_{s_1} \in H_{\theta_1}$,$J'_{s_2} \in H_{\theta_2}$ and $\gamma \in S_{J'_{s_1}} \bigcap S_{J'_{s_2}}$. We know that $\gamma \in \mathcal{U} \backslash \mathcal{U}_0$ since $H_{\theta_i} \bigcap H_{\theta_j} = \emptyset$ for any $1 \leq i \neq j \leq t$. From Fact 3, $\mathrm{card}(H_{\theta_i} \bigcap H_\gamma) = 0$ or 1. Given the elements of $(\bigcup_{i=1}^{l'} S_{J'_i}) \backslash (\mathcal{U}_0 \bigcup \{\gamma\})$, θ_i can be decided by the balance of some $J' \in H_{\theta_i} \backslash H_\gamma$. Then γ can be decided by any vector in

H_γ. Thus we have $O(b_n^{\frac{\sum_{i=1}^{l'} p'_i - l_0 - h - h_\gamma}{2} + l_0})$ ways to decide $J'_1, J'_2, \ldots, J'_{l'}$. From Eq.(5.12), one gets

$$O(b_n^{\frac{\sum_{i=1}^{l'} p'_i - l_0 - h - h_\gamma}{2} + l_0}) = O(b_n^{\frac{(\sum_{i=1}^{l} p_i - l) - (m - l_0) - (h - l') - h_\gamma}{2}}) = o(b_n^{\frac{\sum_{i=1}^{l} p_i - l}{2}}). \tag{5.14}$$

If $h = l'$ and $t = 1$, which means that θ_1 appears exactly one time in each $S_{J_i}(1 \le i \le l')$, we will divide the situation into three subcases.

Case I: $l_0 > 0$.

Without loss of generality, we suppose $\alpha \in J'_1$ is an element with single multiplicity in $\bigcup_{i=1}^{l'} J'_i$. If the elements except for θ_1 and α are known, θ_1 can be determined from the balance of J'_2 and then α can be determined from the balance of J'_1. So we have

$$O(b_n^{\frac{\sum_{i=1}^{l'} p'_i - l_0 - h_{\theta_1}}{2} + l_0 - 1})$$

ways of choice in sum. From $h_{\theta_1} = h = l'$ and Eqs. (5.9) and (5.10), we get

$$O(b_n^{\frac{\sum_{i=1}^{l'} p'_i - l_0 - h_{\theta_1}}{2} + l_0 - 1}) = O(b_n^{\frac{\sum_{i=1}^{l} p_i - l}{2} - 1}) = o(b_n^{\frac{\sum_{i=1}^{l} p_i - l}{2}}).$$

Case II: $l_0 = 0$ **and** $m > 0$.

As in case I above, θ_1 is determined by other elements. To determine all the elements other than θ_1, we have $O(b_n^{\frac{\sum p'_i - l'}{2}})$ ways. From Eqs. (5.9) and (5.10), we have

$$O(b_n^{\frac{\sum p'_i - l'}{2}}) = O(b_n^{\frac{\sum p_i - l}{2} - \frac{m}{2}}) = o(b_n^{\frac{\sum_{i=1}^{l} p_i - l}{2}}).$$

Case III: $l_0 = 0$ **and** $m = 0$.

In this case, we cannot do any *reduction* step.

We claim that there exists $\gamma \in S_{J_1}$ other than θ_1 such that the number of $+\gamma$ and $-\gamma$ in J_1 are not equal (when p_1 is even we can always find some γ other than θ_1 such that γ occurs odd times in S_{J_1}). Otherwise, $\theta_1 = 0$ because of the balance of J_1, which contradicts condition (iii) in Lemma 5.3.

To determine all the elements except for θ_1 and γ, we have

$$O(b_n^{\frac{(p_1 + \ldots + p_l) - (h_{\theta_1} + 1)}{2}})$$

ways. Since $H_{\theta_1} = \{J_1, J_2, \ldots, J_l\}$, we have $H_\gamma = \{J_1\}$ (otherwise a reduction can be done). So we can determine θ_1 from the balance of J_2. Then γ will be determined by the balance of J_1. Since $h_{\theta_1} = l$, we have

$$O(b_n^{\frac{(p_1 + \ldots + p_l) - (h_{\theta_1} + 1)}{2}}) = o(b_n^{\frac{(p_1 + \ldots + p_l) - l}{2}}).$$

Now we complete the proof in the case of $l' > 1$.

In the case $l' = 1$

We also divide this case into two subcases.

Case I′: $l_0 > 0$.

Suppose $\alpha \in J_1'$ is an element with single multiplicity. If the elements other than α are known, α can be determined from the balance of J_1'. So we have totally $O(b_n^{\frac{\sum_{i=1}^{l'} p_i' - l_0}{2} + l_0 - 1})$ ways. From $l' = 1$ and Eqs. (5.9), (5.10) and (5.11), we have

$$O(b_n^{\frac{\sum_{i=1}^{l'} p_i' - l_0}{2} + l_0 - 1}) = O(b_n^{\frac{\sum_{i=1}^{l} p_i - l}{2} - \frac{m - l_0 + 1}{2}}) = o(b_n^{\frac{\sum_{i=1}^{l} p_i - l}{2}}).$$

Case II′: $l_0 = 0$.

In this case, every element in $S_{J_1'}$ appears at least twice. Thus we have totally $O(b_n^{\frac{\sum_{i=1}^{l'} p_i'}{2}})$ ways. From $l' = 1$ and Eqs. (5.9) and (5.10), it follows that

$$O(b_n^{\frac{\sum_{i=1}^{l'} p_i'}{2}}) = O(b_n^{\frac{\sum_{i=1}^{l} p_i - l}{2} - \frac{l-2}{2}}) = o(b_n^{\frac{\sum_{i=1}^{l} p_i - l}{2}})$$

since $l \geq 3$.

Now we have completed the proof in the case of $l' = 1$.

The proof of Lemma 5.3 is then complete. □

We now turn to the proof of the main Theorem 1.1.

Proof of Theorem 1.1. By Theorem 4.1, Lemma 5.3 and Wick formulas for centered Gaussian distributions, it's easy to complete the proof. □

Remark 5.5. *From the analysis of case III above, we can also understand why the technical but necessary condition of $a_0 \equiv 0$ is assumed in the introduction. But when p is even, the assumption of $a_0 \equiv 0$ is not necessary.*

Remark 5.6. *From the calculation of mathematical expectation, variance, and higher moments in sections 3, 4, 5, the factor I_J in the trace formula of Lemma 3.2 can be replaced by a more general non-negative integrable bounded function. For example, if $I_J \equiv 1$ for all J, then the results of asymptotic distribution and fluctuation hold the same as in case where $b = 0$. More precisely, as $n \longrightarrow \infty$*

$$\frac{1}{b_n^{\frac{p}{2}}} \sum_{j_1,\ldots,j_p=-b_n}^{b_n} \mathbb{E}[\prod_{l=1}^{p} a_{j_l}]\delta_{0,\sum_{l=1}^{p} j_l} \longrightarrow \begin{cases} 2^{\frac{p}{2}}(p-1)!!, & p \text{ is even;} \\ 0, & p \text{ is odd} \end{cases} \tag{5.15}$$

and the fluctuation can be stated as in Corollary 1.2 (For conciseness we replace b_n by k there). From this point of view, our method gives a mechanism of CLT.

6 Extensions to other models

In this section, we will make use of our method to deal with some random matrix models closely related to Toeplitz matrices, including Hermitian Toeplitz band matrices and Hankel band matrices. Before we start our extensions, we first review and generalize the key procedures or arguments in calculating mathematical expectation (the distribution of eigenvalues), covariance and higher moments in sections 3, 4, 5 respectively as follows:

1)"Good" trace formulae. See Lemma 3.2. The simple formula both represents the form of the matrix and translates our object of matrix entries to its subscripts $j_1, \ldots, j_p$, which are integers satisfying some homogeneous equation. In the present paper, we mainly encounter such homogeneous equations as

$$\sum_{l=1}^{p} \tau_l j_l = 0, \tag{6.1}$$

where τ_l can take ± 1. We write $J = (j_1, \ldots, j_p)$.

2)"Balanced" vectors via some homogeneous equations. We can generalize the concept of balance: a vector $J = (j_1, \ldots, j_p)$ is said to be balanced if its components satisfy one of finite fixed equations with the form of (6.1). From the balance of a vector, one can determine an element by the others and solve one of the equations.

3) Reduction via a joint point. Eliminate the joint point from two correlated vectors in some definite way, and one gets a new balanced vector.

4) Choose a minimal dominating set by which the freedom degree can be reduced case-by-case. We should particularly be careful about 0.

The CLT is essentially a consequence of the fact that we can omit the terms which have a cluster consisting of more than two vectors. Thus if the argument in section 5 is valid, by Wick formulas the CLT is true. In the following models, we will establish a good trace formula in each case, and then the balanced vector and reduction step can be defined in a natural way. With these equipment the analysis in section 5 is still valid after a small adaption.

The mathematical expectation and covariance vary from case to case. But the way to calculate them is similar. We will only state main results without details, since the proofs would have been overloaded with unnecessary notations and minor differences. When necessary, we will point out the differences.

6.1 Hermitian Toeplitz band matrices

The case is very similar to real symmetric Toeplitz band matrices, except that we now consider n-dimensional complex Hermitian matrices $T_n = (\eta_{ij}\, a_{i-j})_{i,j=1}^n$. We assume that $\mathrm{Re}\, a_j = \mathrm{Re}\, a_{-j}$ and $\mathrm{Im}\, a_j = -\mathrm{Im}\, a_{-j}$ for $j = 1, 2, \cdots$, and $\{\mathrm{Re}\, a_j, \mathrm{Im}\, a_j\}_{j\in\mathbb{N}}$ is a sequence of independent real random variables such that

$$\mathbb{E}[a_j] = 0, \quad \mathbb{E}[|a_j|^2] = 1 \ \text{ and } \mathbb{E}[a_j^2] = 0 \text{ for } \ j \in \mathbb{N}, \tag{6.2}$$

(homogeneity of 4-th moments)

$$\kappa = \mathbb{E}[|a_j|^4], \tag{6.3}$$

and further (uniform boundedness)

$$\sup_{j\in\mathbb{Z}} \mathbb{E}[|a_j|^k] = C_k < \infty \quad \text{for} \quad k \in \mathbb{N}. \tag{6.4}$$

In addition, we also assume $a_0 \equiv 0$ and the bandwidth $b_n \to \infty$ but $b_n/n \to b \in [0, 1]$.

Theorem 6.1. *With above assumptions and notations, Theorem 1.1 also holds for random Hermitian Toeplitz band matrices.*

Remark 6.2. *The distribution of eigenvalues for this case has been proved to be the same as the real case in [21]. The covariance like in (4.2) is slightly different from the real case because $\mathbb{E}[a_j^2] = 0$, which is given by*

$$\sum_{\pi\in\mathcal{P}_2(p,q)} f_I^-(\pi) + (\kappa-1) \sum_{\pi\in\mathcal{P}_{2,4}(p,q)} \left(f_{II}^-(\pi) + f_{II}^+(\pi)\right). \tag{6.5}$$

6.2 Hankel band matrices

A Hankel matrix $H_n = (h_{i+j-1})_{i,j=1}^n$ is closely related to a Toeplitz matrix. Explicitly, let $P_n = (\delta_{i-1,n-j})_{i,j=1}^n$ the “backward identity" permutation, then for a Toeplitz matrix of the form $T_n = (a_{i-j})_{i,j=1}^n$ and a Hankel matrix of the form $H_n = (h_{i+j-1})_{i,j=1}^n$, P_nT_n is a Hankel matrix and P_nH_n is a Toeplitz matrix. In this paper we always write a Hankel band matrix $H_n = P_nT_n$ where $T_n = (\eta_{ij}\, a_{i-j})_{i,j=1}^n$ is a Toeplitz band matrix with bandwidth b_n and the matrix entries $a_{-n+1}, \cdots, a_0, \cdots, a_{n-1}$ are real-valued, thus H_n is a real symmetric matrix.

For Hankel band matrices, as in Toeplitz case we also have a trace formula and its derivation is similar, see [21].

Lemma 6.3. $\mathrm{tr}(H_n^p)$=

$$\begin{cases} \sum\limits_{i=1}^{n} \sum\limits_{j_1,\cdots,j_p=-b_n}^{b_n} \prod\limits_{l=1}^{p} a_{j_l} \prod\limits_{l=1}^{p} \chi_{[1,n]}(i - \sum\limits_{q=1}^{l}(-1)^q j_q)\ \delta_{0, \sum\limits_{q=1}^{p}(-1)^q j_q}, & p \text{ even;} \\ \sum\limits_{i=1}^{n} \sum\limits_{j_1,\cdots,j_p=-b_n}^{b_n} \prod\limits_{l=1}^{p} a_{j_l} \prod\limits_{l=1}^{p} \chi_{[1,n]}(i - \sum\limits_{q=1}^{l}(-1)^q j_q)\ \delta_{2i-1-n, \sum\limits_{q=1}^{p}(-1)^q j_q}, & p \text{ odd.} \end{cases}$$

From this trace formula, our method can apply to the case that p is even. We consider random Hankel matrices satisfying the following assumptions: $\{a_j : j \in \mathbb{Z}\}$ is a sequence of independent real random variables such that

$$\mathbb{E}[a_j] = 0, \quad \mathbb{E}[|a_j|^2] = 1 \ \text{ for } \ j \in \mathbb{Z}, \tag{6.6}$$

(homogeneity of 4-th moments)

$$\kappa = \mathbb{E}[|a_j|^4], j \in \mathbb{Z} \tag{6.7}$$

and further (uniform boundedness)

$$\sup_{j\in\mathbb{Z}} \mathbb{E}[|a_j|^k] = C_k < \infty \quad \text{for} \quad k \geq 3. \tag{6.8}$$

In addition, we also assume the bandwidth $b_n \to \infty$ but $b_n/n \to b \in [0,1]$ as $n \to \infty$.

Theorem 6.4. *Let H_n be a real symmetric ((6.6)–(6.8)) random Hankel band matrix with the bandwidth b_n, where $b_n/n \to b \in [0,1]$ but $b_n \to \infty$ as $n \to \infty$. Set $A_n = H_n/\sqrt{b_n}$ and*

$$\zeta_p = \frac{\sqrt{b_n}}{n}\left(\mathrm{tr}(A_n^{2p}) - \mathbb{E}[\mathrm{tr}(A_n^{2p})]\right). \tag{6.9}$$

Then

$$\zeta_p \longrightarrow N(0, \tilde{\sigma}_p^2) \tag{6.10}$$

in distribution as $n \to \infty$. Moreover, for a given polynomial

$$Q(x) = \sum_{j=0}^{p} q_j x^j \tag{6.11}$$

with degree $p \geq 1$, set

$$\zeta_Q = \frac{\sqrt{b_n}}{n}\left(\mathrm{tr}Q(A_n^2) - \mathbb{E}[\mathrm{tr}Q(A_n^2)]\right), \tag{6.12}$$

we also have

$$\zeta_Q \longrightarrow N(0, \tilde{\sigma}_Q^2) \tag{6.13}$$

in distribution as $n \to \infty$. Here the variances $\tilde{\sigma}_p^2$ and $\tilde{\sigma}_Q^2$ will be given below.

We remark that the limit of $\frac{1}{n}\mathrm{tr}(H_n/\sqrt{b_n})^p$ can be calculated in the same way as in Toeplitz case, see also [21]. Next, we derive briefly the variances $\tilde{\sigma}_p^2$ and $\tilde{\sigma}_Q^2$.

By Lemma 6.3, rewrite

$$\zeta_p = \frac{1}{n b_n^{\frac{2p-1}{2}}} \sum_{i=1}^{n} \sum_J I_J \left(a_J - \mathbb{E}[a_J]\right) \, \delta_{0, \sum_{l=1}^{2p}(-1)^l j_l}. \tag{6.14}$$

Here $J = (j_1, \ldots, j_{2p}) \in \{-b_n, \ldots, b_n\}^{2p}$, $a_J = \prod_{l=1}^{2p} a_{j_l}$, $I_J = \prod_{k=1}^{2p} \chi_{[1,n]}(i - \sum_{l=1}^{k}(-1)^l j_l)$ and the summation $\sum_J$ runs over all possibile $J \in \{-b_n, \ldots, b_n\}^{2p}$. We call a vector $J = (j_1, \ldots, j_{2p}) \in \{-b_n, \ldots, b_n\}^{2p}$ is balanced if

$$\sum_{l=1}^{2p}(-1)^l j_l = 0. \tag{6.15}$$

Since a_j and a_{-j} are independent when $j \neq 0$, S_J, the projection of J, should be J itself (forget the order of its components).

The only difference is the definition of balance and projection. Using the same technique in section 4, we get

$$\mathbb{E}[\zeta_p \, \zeta_q] \longrightarrow \tilde{\sigma}_{p,q} = \sum_{\pi \in \mathcal{P}_2(2p,2q)} g_I(\pi) + (\kappa - 1) \sum_{\pi \in \mathcal{P}_{2,4}(2p,2q)} g_{II}(\pi) \tag{6.16}$$

as $n \to \infty$.

Here for a fixed pair partition π, we construct a projective relation between two groups of unknowns $y_1, \ldots, y_{2p+2q}$ and $x_1, \ldots, x_{p+q}$ as follows:

$$y_i = y_j = x_{\pi(i)} \tag{6.17}$$

whenever $i \sim_\pi j$.

If $\pi \in \mathcal{P}_2(2p, 2q)$, let

$$g_I(\pi) = \int_{[0,1]^2 \times [-1,1]^{p+q}} \delta\left(\sum_{i=1}^{2p}(-1)^i y_i\right) \chi_{\left\{\sum_{i=2p+1}^{2p+2q}(-1)^i y_i = 0\right\}} \tag{6.18}$$
$$\prod_{j=1}^{2p} \chi_{[0,1]}(x_0 - b\sum_{i=1}^{j}(-1)^i y_i) \prod_{j'=2p+1}^{2p+2q} \chi_{[0,1]}(y_0 - b\sum_{i=2p+1}^{j'}(-1)^i y_i) \, d\,y_0 \prod_{l=0}^{p+q} d\,x_l.$$

If $\pi \in \mathcal{P}_{2,4}(2p, 2q)$, let

$$g_{II}(\pi) = \int_{[0,1]^2 \times [-1,1]^{p+q-1}} \chi_{\left\{\sum_{i=1}^{2p}(-1)^i y_i = 0\right\}} \chi_{\left\{\sum_{i=2p+1}^{2p+2q}(-1)^i y_i = 0\right\}} \tag{6.19}$$
$$\prod_{j=1}^{2p} \chi_{[0,1]}(x_0 - b\sum_{i=1}^{j}(-1)^i y_i) \prod_{j'=2p+1}^{2p+2q} \chi_{[0,1]}(y_0 - b\sum_{i=2p+1}^{j'}(-1)^i y_i) \, d\,y_0 \prod_{l=0}^{p+q-1} d\,x_l.$$

Because of the existence of the characteristic function in the integrals of type I and II, we see that $g_I(\pi) \neq 0$ if and only if $\sum_{i=2p+1}^{2p+2q}(-1)^i y_i \equiv 0$ when $\sum_{i=1}^{2p}(-1)^i y_i = 0$. Denote the subset of $\mathcal{P}_2(2p, 2q)$ consisting of this kind of π by $\mathcal{P}_2^I(2p, 2q)$.

$g_{II}(\pi) \neq 0$ if and only if $\sum_{i=2p+1}^{2p+2q}(-1)^i y_i \equiv \sum_{i=1}^{2p}(-1)^i y_i \equiv 0$. Denote the subset of $\mathcal{P}_{2,4}(2p,2q)$ consisting of this kind of π by $\mathcal{P}_{2,4}^{II}(2p,2q)$. From the definition of balance, if we denote $V_i = \{i_1,i_2,i_3,i_4\}$ to be the block with four elements and $1 \le i_1 < i_2 \le 2p < i_3 < i_4 \le 2q$, then i_1+i_2 and i_3+i_4 must be odd. Moreover, $j+k$ is odd provided $j,k \notin V_i$ and $i \sim_\pi j$.

From the discussion above, we get

$$\mathbb{E}[\zeta_p\, \zeta_q] \longrightarrow \tilde{\sigma}_{p,q} = \sum_{\pi \in \mathcal{P}_2^I(2p,2q)} g_I(\pi) + (\kappa-1) \sum_{\pi \in \mathcal{P}_{2,4}^{II}(2p,2q)} g_{II}(\pi), \tag{6.20}$$

$$\tilde{\sigma}_p^2 = \tilde{\sigma}_{p,p} \tag{6.21}$$

and

$$\tilde{\sigma}_Q^2 = \sum_{i=1}^{p}\sum_{j=1}^{p} q_i q_j \tilde{\sigma}_{i,j}. \tag{6.22}$$

Similar to Corollary 1.2, we get another central limit theorem for product of independent random variables.

Corollary 6.5. *Suppose that $\{a_j : j \in \mathbb{Z}\}$ is a sequence of independent random variables satisfying the assumptions (6.6)–(6.8). For every $p \ge 1$,*

$$\frac{1}{k^{\frac{2p-1}{2}}} \sum_{j_1,\ldots,j_{2p}=-k}^{k} \left(\prod_{l=1}^{2p} a_{j_l} - \mathbb{E}[\prod_{l=1}^{2p} a_{j_l}] \right) \delta_{0,\sum_{l=1}^{2p}(-1)^l j_l} \tag{6.23}$$

converges in distribution to a Gaussian distribution $N(0,\tilde{\sigma}_p^2)$ as k tends to infinity. Here the variance $\tilde{\sigma}_p^2$ corresponds to the case $b=0$ above.

Remark 6.6. *The derivation of higher moments can be calculated in the same way as in section 5. Careful examination shows that the argument in case III in section 5 (the existence of γ) is the only part depending on the definition of balance. However, since $2p$ and $2q$ are even, the dimensions of all vectors involved are even. Remark 5.5 is still valid so that one needn't care much about the definition of balance. Moreover, a_0 needn't equal to 0.*

References

[1] Anderson, G. and Zeitouni, O.: A CLT for a band matrix model. *Probab. Theory Related Fields* **134** no. 2, 283–338 (2006). MR-2222385

[2] Bai, Z. : Methodologies in spectral analysis of large-dimensional random matrices, a review, *Statist. Sinica.* **9** (1999), 611–677. MR-1711663

[3] Bai, Z., Silverstein, J. W. : Spectral analysis of large dimensional random matrices. Science Press, Beijing, 2006

[4] Basak, A. and Bose, A.: Limiting spectral distribution of some band matrices, *Periodica Mathematica Hungarica*, Vol. 63(1), 2011, 113–150. MR-2853175

[5] Basor, E. L. : Toeplitz determinants, Fisher–Hartwig symbols and random matrices, *Recent Perspectives in Random Matrix Theory and Number Theory, London Math. Soc. Lecture Note Ser.* **322** (2005), 309–336. MR-2166467

[6] Bose, A., Chatterjee, S. and Gangopadhyay, S. : Limiting spectral distribution of large dimensional random matrices, *J. Indian Statist. Assoc.* **41** (2003), 221–259. MR-2101995

[7] Bose, A. ,Mitra, J. : Limiting spectral distribution of a special circulant, *Stat. Probab. Letters* **60** (2002), no.1, 111–120. MR-1945684

[8] Bryc, W., Dembo, A. and Jiang, T.: Spectral measure of large random Hankel, Markov and Toeplitz matrices, *Ann. Probab.* **34** (2006), no.1, 1–38. MR-2206341

[9] Chatterjee, S.: Fluctuations of eigenvalues and second order Poincar*é* inequalities, *Probab. Theory Related Fields* **143**, 1–40 (2009). MR-2449121

[10] Diaconis, P. : Patterns in eigenvalues: The 70th Josiah Willard Gibbs lecture, *Bull. Amer. Math. Soc.* **40** (2003), 115–178. MR-1962294

[11] Diaconis, P. and Evans, S. N.: Linear functionals of eigenvalues of random matrices. *Trans. Amer. Math. Soc.* **353** 2615–2633(2001). MR-1828463

[12] Diaconis, P. and Shahshahani, M. : On the eigenvalues of random matrices: Studies in applied probability. *J. Appl. Probab.* **31A**, 49–62(1994). MR-1274717

[13] Dumitriu, I. and Edelman, A.: Global spectrum fluctuation for the β-Hermite and β-Laguerre ensembles via matrix models, *Journal of Mathematical Physics* **47**(6), 063302–063336(2006) MR-2239975

[14] Grenander, U., Szegő, G.: *Toeplitz forms and their applications*. University of California Press, Berkeley-Los Angeles, 1958. MR-0094840

[15] Hammond, C. and Miller, S. J.: Distribution of eigenvalues for the ensemble of real symmetric Toeplitz matrices, *J. Theoret. Probab.* **18** (2005), 537–566. MR-2167641

[16] Johansson, K.: On Szegő's asymptotic formula for Toeplitz determinants and generalizations. Bull. Sci. Math. (2) **112**, 257–304 (1988). MR-0975365

[17] Johansson, K.: On the fluctuation of eigenvalues of random Hermitian matrices. *Duke Math. J.* **91**, 151–204 (1998). MR-1487983

[18] Jonsson, D.: Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Mult. Anal.* **12**, 1–38(1982). MR-0650926

[19] Kargin, V.: Spectrum of random Toeplitz matrices with band structures, *Elect. Comm. in Probab.* **14** (2009), 412–421 MR-2551851

[20] Horn, R. A. , Johnson, C. R.: *Matrix Analysis*, Cambridge University Press, New York, 1985. MR-0832183

[21] Liu, D.-Z., Wang, Z.-D.: Limit Distribution of Eigenvalues for Random Hankel and Toeplitz Band Matrices, *J. Theor. Probab.* **24**(2011), 988–1001 MR-2851241

[22] Massey, A., Miller, S. J. and Sinsheimer, J.: Distribution of Eigenvalues of Real Symmetric Palindromic Toeplitz Matrices and Circulant Matrices, *J. Theoret. Probab.* **20** (2007), 637–662. MR-2337145

[23] Popescu, I.: General tridiagonal random matrix models, limiting distributions and fluctuations. *Probab. Theory Related Fields* **114**, 179–220 (2009). MR-2480789

[24] Sen, A. and Virág, B.: Absolute continuity of the limiting eigenvalue distribution of the random Toeplitz matrix, *Elect. Comm. in Probab.* **16** (2011), 706–711 MR-2861434

[25] Sen, A. and Virág, B.: The top eigenvalue of the random Toeplitz matrix and the Sine kernel, arXiv:1109.5494.

[26] Sinai, Y. and Soshnikov, A.: Central Limit Theorem for Traces of Large Random Symmetric Matrices With Independent Matrix Elements, *Bol. Soc. Bras. Mat* **29** (1998), 1–24. MR-1620151

Acknowledgments. The first and third authors were supported by the National Natural Science Foundation of China (Grant # 11171005). The work of D.-Z. L. was partially done during his stay at the School of Mathematical Sciences of Peking University.