Bernoulli **24**(1), 2018, 80–114
DOI: 10.3150/16-BEJ867

The logarithmic law of sample covariance matrices near singularity

XUEJUN WANG[1], XIAO HAN[2,*] and GUANGMING PAN[2,**]

[1] *Division of Statistics, School of Mathematical Sciences, Anhui University, Hefei 230601, P.R. China. E-mail: wxjahdx2000@126.com*
[2] *School of Physical and Mathematical Sciences, Nanyang Technological University, 637371, Singapore. E-mail: *xhan011@e.ntu.edu.sg; **gmpan@ntu.edu.sg*

Let $B = (b_{jk})_{p\times n} = (Y_1, Y_2, \ldots, Y_n)$ be a collection of independent real random variables with mean zero and variance one. Suppose that Σ is a p by p population covariance matrix. Let $X_k = \Sigma^{1/2} Y_k$ for $k = 1, 2, \ldots, n$ and $\hat{\Sigma}_1 = \frac{1}{n}\sum_{k=1}^n X_k X_k^T$. Under the moment condition $\sup_{p,n} \max_{1\le j\le p, 1\le k\le n} \mathbb{E} b_{jk}^4 < \infty$, we prove that the log determinant of the sample covariance matrix $\hat{\Sigma}_1$ satisfies

$$\frac{\log\det\hat{\Sigma}_1 - \sum_{k=1}^p \log(1-\frac{k}{n}) - \log\det\Sigma}{\sqrt{-2\log(1-\frac{p}{n})}} \xrightarrow{d} N(0,1),$$

when $p/n \to 1$ and $p < n$. For $p = n$, we prove that

$$\frac{\log\det\hat{\Sigma}_1 + n\log n - \log(n-1)! - \log\det\Sigma}{\sqrt{2\log n}} \xrightarrow{d} N(0,1).$$

Keywords: central limit theorem; covariance matrix; determinant; logarithmic law

1. Introduction and main results

Let $A_n = (a_{ij})_{n,n}$ be a $n \times n$ random matrix where the entries $\{a_{ij}, 1 \le i, j \le n\}$ are independent real random variables. The determinant of the random matrix A_n is a very important function of the matrix and it has been investigated by many authors under different settings. See, for instance, [6–10,12,15–19], and so on. An important topic in the random determinant theory is to establish the central limit theorems (CLT, in short) for the log-determinant $\log|\det A_n|$, such as [10] for random Gaussian matrices, [19] for Wigner matrices, [13] for general real i.i.d. random matrices under an exponential tail condition on the entries, and [4] for general real i.i.d. random matrices based on the existence of 4th moment of the matrix entries, and so forth. The determinant of random matrices has many applications. For instance, computing the volume of random parallelotopes in random geometry involves the determinant. More precisely, we suppose $\mathbf{V} = (\mathbf{V}_1, \ldots, \mathbf{V}_p)$ is a $n \times p$ ($p \le n$) random matrix with independent columns. Then the convex hull of these p points in $\mathbb{R}^n$ determines a p-parallelotope almost surely. Moreover, $\sqrt{\det(\mathbf{V}^T\mathbf{V})}$ is the volume of this p-parallelotope. One can refer to [14] for more details. The determinant of the sample covariance matrices is also commonly used for constructing hypothesis tests such as the likelihood ratio statistics in multivariate statistics (see [1], for instance). Furthermore, the difference of the log determinants of two sample covariance matrices is a necessary information in quadratic discriminant analysis for classification. In view of these applications, it is important to investigate the properties of the log determinant of the sample covariance matrices.

Motivated by the applications mentioned above, Cai *et al.* [5] studied the limiting law of the log determinant of the sample covariance matrices for the high-dimensional Gaussian distributions. Specifically, let $X_1, X_2, \dots, X_{n+1}$ be an independent random sample from the p-dimensional Gaussian distribution $N_p(\mu, \Sigma)$. The sample covariance matrix is

$$\hat{\Sigma} = \frac{1}{n}\sum_{k=1}^{n+1}(X_k - \bar{X})(X_k - \bar{X})^T, \tag{1.1}$$

where $\bar{X} = \frac{1}{n+1}\sum_{k=1}^{n+1} X_k$. By using the Bartlett decomposition and an analysis of the characteristic functions, [5] established a unified central limit theorem for the log determinant of $\hat{\Sigma}$ in the high-dimensional setting where the dimension p may grow with the sample size n with the only restriction that $p \le n$. Let

$$\nu_{n,p} = \sum_{k=1}^{p}\left[\psi\left(\frac{n-k+1}{2}\right) - \log\left(\frac{n}{2}\right)\right], \qquad \sigma_{n,p} = \left(\sum_{k=1}^{p}\frac{2}{n-k+1}\right)^{1/2}, \tag{1.2}$$

where $\psi(x) = \frac{\partial}{\partial z}\log\Gamma(z)|_{z=x}$ is the Digamma function with $\Gamma(z)$ being the gamma function. Cai *et al.* [5] proved that

$$\frac{\log\det\hat{\Sigma} - \nu_{n,p} - \log\det\Sigma}{\sigma_{n,p}} \xrightarrow{d} N(0,1) \tag{1.3}$$

if $X_1, X_2, \dots, X_{n+1} \overset{\text{i.i.d.}}{\sim} N_p(\mu, \Sigma)$ and $n \to \infty$ with $p \le n$.

CLT of the form (1.3) provides a nice unified expression for $p \le n$. Cai *et al.* [5] also worked out some explicit expressions of the mean $\nu_{n,p}$ in the case when $\lim_n p/n < 1$ and in the case when $p/n \to 1$ but with $(n-p) \to \infty$. Particularly, they proved that for these cases

$$\nu_{n,p} = \sum_{k=1}^{p}\log\left(1 - \frac{k}{n}\right) + O\left(\frac{1}{n}\right), \qquad \sigma_{np} = \sqrt{-2\log\left(1 - \frac{p}{n}\right)} + O\left(\frac{1}{n}\right). \tag{1.4}$$

However, it is not easy to work out a simple expression of $\nu_{n,p}$ from (1.2) for the remaining cases of $p/n \to 1$ as pointed out by [5]. Moreover, their results are for multivariate normal distributions.

In view of this, we further study the central limit theorem for the log determinant of the sample covariance matrices under the non-Gaussian samples and in the setting of $p/n \to 1$ with $p \le n$.

Before introducing the main results of the paper, we first list some notations. Let $Y_k = (Y_{1k}, Y_{2k}, \dots, Y_{pk})^T$ for $k = 1, 2, \dots, n$. Suppose that $Y_1, Y_2, \dots, Y_n$ are independent. Moreover assume further that $Y_{1k}, Y_{2k}, \dots, Y_{pk}$ are independent for each $k = 1, 2, \dots, n$. Let Σ be a p by

p population covariance matrix which is positive definite. Let $X_k = \Sigma^{1/2} Y_k$ for $k = 1, 2, \ldots, n$. Thus, $X_1, X_2, \ldots, X_n$ are independent and can be viewed as a sample drawn from the population with mean 0 and covariance matrix Σ. To estimate the log determinant of the covariance matrix Σ, we consider the following sample covariance matrix

$$\hat{\Sigma}_1 = \frac{1}{n} \sum_{k=1}^{n} X_k X_k^T. \tag{1.5}$$

For positive integers n and p define the constant $\tau_{n,p}$ by

$$\tau_{n,p} = \sum_{i=0}^{p-1} \log\left(1 - \frac{i+1}{n}\right) = \sum_{k=1}^{p} \log\left(1 - \frac{k}{n}\right). \tag{1.6}$$

Denote

$$B = (b_{jk})_{p \times n} = (Y_1, Y_2, \ldots, Y_n) = \begin{pmatrix} \mathbf{b}_1^T \\ \mathbf{b}_2^T \\ \vdots \\ \mathbf{b}_p^T \end{pmatrix}. \tag{1.7}$$

It is straightforward from (1.5) to see that

$$\begin{aligned} \log\det \hat{\Sigma}_1 - \log\det \Sigma &= \log\det\big(\Sigma^{-1/2} \hat{\Sigma}_1 \Sigma^{-1/2}\big) \\ &= \log\det\left(\frac{1}{n} \sum_{k=1}^{n} Y_k Y_k^T\right) = \log\det\left(\frac{1}{n} B B^T\right). \end{aligned} \tag{1.8}$$

It follows that

$$\log\det \hat{\Sigma}_1 - \tau_{n,p} - \log\det \Sigma = \log\det\big(B B^T\big) - p \log n - \tau_{n,p}. \tag{1.9}$$

If $p/n \to r \in (0, 1)$, then the central limit theorem for the log determinant of $\hat{\Sigma}_1$ based on the existence of 4th moment of the i.i.d. matrix entries can be obtained from Theorem 1.1 of [2] with $f(x) = \log x$. We aim to consider the case $p/n \to 1$ with $p \le n$.

We are now at a position to state the main results.

Theorem 1.1. *Assume that $\{b_{ij}\}$ are independent random variables with mean zero and variance one. Moreover, suppose that $p = n$ and*

$$\sup_n \max_{1 \le j \le n} \mathbb{E} b_{j1}^4 < \infty. \tag{1.10}$$

Then the log determinant of the sample covariance matrix $\hat{\Sigma}_1$ satisfies

$$\frac{\log\det \hat{\Sigma}_1 + n \log n - \log(n-1)! - \log\det \Sigma}{\sqrt{2 \log n}} \xrightarrow{d} N(0, 1). \tag{1.11}$$

Theorem 1.2. *Assume that* $\{b_{ij}\}$ *are independent random variables with mean zero and variance one. Moreover, suppose that*

$$\sup_p \max_{1\le j\le p} \mathbb{E}b_{j1}^4 < \infty \tag{1.12}$$

and

$$\frac{p}{n} \to 1, \qquad p < n. \tag{1.13}$$

Then the log determinant of the sample covariance matrix $\hat{\Sigma}_1$ *satisfies*

$$\frac{\log\det\hat{\Sigma}_1 - \tau_{n,p} - \log\det\Sigma}{\sqrt{-2\log(1-\frac{p}{n})}} \xrightarrow{d} N(0,1). \tag{1.14}$$

Remark 1.1. Theorem 1.2 is consistent with (1.3) and (1.4) for the case when $p/n \to 1$ with $(n-p) \to \infty$ and extends it to include all cases $p/n \to 1$ as long as $p < n$. Moreover when $p = n-1$ and $p = n-2$ (corresponding to the case of $(n-p)$ is bounded) one may prove from (1.2) that the difference $(\nu_{n,n-1} - \nu_{n,n-2})$ is of order $O(\log n)$. This is consistent with the result that the difference $(\tau_{n,n-1} - \tau_{n,n-2})$ of order $O(\log n)$ from (1.6). The assumption of i.i.d. Gaussian entries in [5] is also weakened by i.i.d. entries with the fourth moment assumption.

Our paper is organized as follows. Section 2 is to introduce the QR decomposition and obtain its martingale expression. Section 3 collects some lemmas while Section 4 presents the details of the proof with the aid of some additional lemmas, whose proofs are given in the Appendix.

Throughout the paper, let C denote a positive constant which is not necessarily the same in each appearance and $\lfloor x\rfloor$ denote the integer part of x. For two vectors $\mathbf{v}$ and $\mathbf{w}$, let $(\mathbf{v},\mathbf{w}) = \mathbf{v}^T\mathbf{w}$ stand for the inner product. We use $\|\cdot\|_2$ and $\|\cdot\|_{\mathrm{op}}$ to represent the Euclidean norm of a vector and the operator norm of a matrix, respectively. For the set A, let A^c denote the complement set of A and $I(A)$ stand for the indicator function of A.

2. QR decomposition

This section is to use the method of QR decomposition to derive an alternative expression of the determinant. To state this method rigorously, we need the following proposition, whose proof will be given in Appendix.

Proposition 2.1. *For the matrix* $B = (b_{jk})_{p\times n}$ *defined in Section 1, we can find a modified matrix* $\bar{B} = (\bar{b}_{jk})_{p\times n}$ *satisfying the assumptions in Theorems 1.1–1.2 such that*

$$\mathbb{P}(\textit{all square submatrices of } \bar{B} \textit{ are invertible}) = 1 \tag{2.1}$$

and

$$\mathbb{P}\bigl(\log\det\bigl(BB^T\bigr) - \log\det\bigl(\bar{B}\bar{B}^T\bigr) = o(1)\bigr) = 1 - o(1). \tag{2.2}$$

We can always work under the following assumption by Proposition 2.1.

Assumption C_0**.** Let $B = (b_{jk})_{p\times n}$ be a p by n random matrix with

$$\sup_p \max_{1\le j\le p} \mathbb{E}b_{j1}^4 < \infty,$$

where $\{b_{jk}, 1 \le j \le p, 1 \le k \le n\}$ is a collection of independent real random variables with common mean 0 and variance 1, and $b_{j1}, b_{j2}, \ldots, b_{jn}$ have the identical distribution for each $1 \le j \le p$. Moreover, we assume that all square submatrices of B are invertible with probability one.

Now let us apply the Gram–Schmidt process to the columns of the full column rank matrix A to get its QR decomposition, where $A = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p]$ is a $n \times p$ matrix with $n \ge p$. Define the projection of a vector $\mathbf{a}$ on a vector $\mathbf{e}$ by

$$\mathbf{proj}_e\mathbf{a} = \frac{(\mathbf{e}, \mathbf{a})}{(\mathbf{e}, \mathbf{e})}\mathbf{e}.$$

Then

$$\begin{aligned}
&\mathbf{u}_1 = \mathbf{a}_1, \qquad \mathbf{e}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|_2},\\
&\mathbf{u}_2 = \mathbf{a}_2 - \mathbf{proj}_{\mathbf{e}_1}\mathbf{a}_2, \qquad \mathbf{e}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|_2},\\
&\mathbf{u}_3 = \mathbf{a}_3 - \mathbf{proj}_{\mathbf{e}_1}\mathbf{a}_3 - \mathbf{proj}_{\mathbf{e}_2}\mathbf{a}_3, \qquad \mathbf{e}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|_2},\\
&\quad\vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots\\
&\mathbf{u}_p = \mathbf{a}_p - \sum_{j=1}^{p-1}\mathbf{proj}_{\mathbf{e}_j}\mathbf{a}_p, \qquad \mathbf{e}_p = \frac{\mathbf{u}_p}{\|\mathbf{u}_p\|_2}.
\end{aligned}$$

Rearrange the above equations we may write

$$A = QR,$$

where

$$Q = [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p]$$

and

$$R = \begin{pmatrix}
(\mathbf{e}_1, \mathbf{a}_1) & (\mathbf{e}_1, \mathbf{a}_2) & (\mathbf{e}_1, \mathbf{a}_3) & \cdots & (\mathbf{e}_1, \mathbf{a}_p)\\
0 & (\mathbf{e}_2, \mathbf{a}_2) & (\mathbf{e}_2, \mathbf{a}_3) & \cdots & (\mathbf{e}_2, \mathbf{a}_p)\\
0 & 0 & (\mathbf{e}_3, \mathbf{a}_3) & \cdots & (\mathbf{e}_3, \mathbf{a}_p)\\
\vdots & \vdots & \vdots & \ddots & (\mathbf{e}_p, \mathbf{a}_p)\\
0 & 0 & 0 & \ddots & (\mathbf{e}_p, \mathbf{a}_p)
\end{pmatrix}_{p\times p} \doteq (r_{ij})_{p\times p}.$$

Recall that $B=(b_{jk})_{p\times n}=[\mathbf{b}_1,\mathbf{b}_2,\dots,\mathbf{b}_p]^T$ is a p by n random matrix with $p\le n$. Applying the QR decomposition with $A=B^T$ and $\mathbf{a}_i=\mathbf{b}_i$ for $1\le i\le p$, we have

$$B^T=QR,\qquad B=R^TQ^T,$$

and

$$\det\bigl(BB^T\bigr)=\det\bigl(R^TQ^TQR\bigr)=\det\bigl(R^TR\bigr)=\prod_{i=1}^{p}r_{ii}^2=\prod_{i=0}^{p-1}r_{i+1,i+1}^2. \tag{2.3}$$

Let $B_{(j)}$ be the $j\times n$ rectangular matrix formed by the first j rows of B. Hence, $B_{(1)}=\mathbf{b}_1^T$ and $B_{(p)}=B$. Let V_i be the subspace generated by the first i rows of B and $P_i=(p_{jk}(i))_{n\times n}$ be the projection matrix onto the space $V_i^{\perp}$. $P_0=I_n$ is the n by n identity matrix. It is easily checked that

$$r_{11}^2=(\mathbf{e}_1,\mathbf{b}_1)^2=\|\mathbf{b}_1\|_2^2=\mathbf{b}_1^TP_0\mathbf{b}_1, \tag{2.4}$$

and

$$\begin{aligned} r_{i+1,i+1}^2&=(\mathbf{e}_{i+1},\mathbf{b}_{i+1})^2=\bigl(\mathbf{e}_{i+1}^T\mathbf{b}_{i+1}\bigr)^2=\left(\frac{\mathbf{u}_{i+1}^T\mathbf{b}_{i+1}}{\|\mathbf{u}_{i+1}\|_2}\right)^2\\ &=\left(\frac{\mathbf{b}_{i+1}^TP_i^T\mathbf{b}_{i+1}}{\|P_i\mathbf{b}_{i+1}\|_2}\right)^2=\mathbf{b}_{i+1}^TP_i\mathbf{b}_{i+1},\qquad 1\le i\le p-1. \end{aligned} \tag{2.5}$$

It follows from the definition of V_i and the Assumption C_0 that $B_{(i)}B_{(i)}^T$ is invertible with probability one that

$$P_i=\bigl(p_{jk}(i)\bigr)=I_n-B_{(i)}^T\bigl(B_{(i)}B_{(i)}^T\bigr)^{-1}B_{(i)}. \tag{2.6}$$

By (2.4), (2.5) and (2.6), we have

$$\mathbb{E}\bigl(r_{i+1,i+1}^2|P_i\bigr)=\operatorname{tr}P_i=n-i,\qquad 0\le i\le p-1. \tag{2.7}$$

The equality (2.3) yields that

$$\begin{aligned} \log\det\bigl(BB^T\bigr)&=\sum_{i=0}^{p-1}\log r_{i+1,i+1}^2\\ &=\sum_{i=0}^{p-1}\log\frac{r_{i+1,i+1}^2}{n-i}+\log\bigl[n(n-1)\cdots(n-p+1)\bigr]. \end{aligned} \tag{2.8}$$

For $0\le i\le p-1$, set

$$X_{i+1}=\frac{r_{i+1,i+1}^2-(n-i)}{n-i} \tag{2.9}$$

and

$$R_{i+1} = \log(1 + X_{i+1}) - \left(X_{i+1} - \frac{X_{i+1}^2}{2}\right). \tag{2.10}$$

Hence, we have

$$\log \frac{r_{i+1,i+1}^2}{n-i} = X_{i+1} - \frac{X_{i+1}^2}{2} + R_{i+1}, \tag{2.11}$$

which, together with (1.6), (1.9), yields that

$$\begin{aligned}
&\frac{\log\det\hat{\Sigma}_1 - \tau_{n,p} - \log\det\Sigma}{\sqrt{-2\log(1-\frac{p}{n})}} \\
&= \frac{\log\det(BB^T) - p\log n - \tau_{n,p}}{\sqrt{-2\log(1-\frac{p}{n})}} \\
&= \frac{\sum_{i=0}^{p-1}\log\frac{r_{i+1,i+1}^2}{n-i} + \log[n(n-1)\cdots(n-p+1)] - p\log n - \tau_{n,p}}{\sqrt{-2\log(1-\frac{p}{n})}} \\
&= \frac{1}{\sqrt{-2\log(1-\frac{p}{n})}}\sum_{i=0}^{p-1}X_{i+1} - \frac{1}{\sqrt{-2\log(1-\frac{p}{n})}}\left[\sum_{i=0}^{p-1}\frac{X_{i+1}^2}{2} + \log\left(1-\frac{p}{n}\right)\right] \\
&\quad + \frac{1}{\sqrt{-2\log(1-\frac{p}{n})}}\sum_{i=0}^{p-1}R_{i+1} \\
&= \frac{1}{\sqrt{-2\log(1-\frac{p}{n})}}\sum_{i=0}^{p-s_1}X_{i+1} - \frac{1}{\sqrt{-2\log(1-\frac{p}{n})}}\left[\sum_{i=0}^{p-s_1}\frac{X_{i+1}^2}{2} + \log\left(1-\frac{p}{n}\right)\right] \\
&\quad + \frac{1}{\sqrt{-2\log(1-\frac{p}{n})}}\sum_{i=0}^{p-s_1}R_{i+1} + \frac{1}{\sqrt{-2\log(1-\frac{p}{n})}}\sum_{i=p-s_1+1}^{p-1}\log\frac{r_{i+1,i+1}^2}{n-i} \\
&\doteq I_1 - I_2 + I_3 + I_4,
\end{aligned} \tag{2.12}$$

where $s_1 = \lfloor(-\log\varepsilon_n)^{1/4}\rfloor$ and $\varepsilon_n = 1 - p/n \to 0$.

Crudely speaking, the main route is to prove that the first term of (2.12) converges weakly to the standard Gaussian distribution and the remaining three terms tend to zero in probability.

Let $\mathcal{F}_0 \doteq \{\Omega, \phi\}$ and $\mathcal{F}_i$ be the σ-algebra generated by the first i rows of B. It follows from (2.7) and (2.9) that

$$\mathbb{E}(X_{i+1}|\mathcal{F}_i) = 0, \qquad 0 \le i \le p-1, \tag{2.13}$$

which yields that $X_1, X_2, \ldots, X_p$ is a martingale difference sequence with respect to the filtration $\mathcal{F}_0 \doteq \{\Omega, \phi\} \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_{p-1}$. This is a very important fact to prove that the first term of (2.12) converges weakly to the standard Gaussian distribution by using the classical CLT for martingales.

3. Some lemmas

This section is to list important lemmas to be used and prove a crucial one. We start with two lemmas from [4].

Lemma 3.1. *Let $A_n = (a_{ij})_{n\times n}$ be a square random matrices, where $\{a_{ij}, 1 \le i, j \le n\}$ is a collection of independent real random variables with common mean* 0 *and variance* 1. *Moreover, we assume*

$$\sup_n \max_{1\le i,j\le n} \mathbb{E}a_{ij}^4 < \infty. \tag{3.1}$$

Then we have the logarithmic law for $\det A_n^2$*: as n tends to infinity,*

$$\frac{\log\det A_n^2 - \log(n-1)!}{\sqrt{2\log n}} \xrightarrow{d} N(0,1). \tag{3.2}$$

Lemma 3.2. *Let $\{x_i, 1 \le i \le n\}$ be independent real random variables with common mean zero and variance* 1. *Moreover, we assume that* $\max_i \mathbb{E}|x_i|^l \le \nu_l$. *Let $M_n = (m_{ij})_{n\times n}$ be a nonnegative definite matrix which is deterministic. Then there exists a positive constant C such that*

$$\mathbb{E}\left|\sum_{i=1}^n m_{ii}x_i^2 - \operatorname{tr} M_n\right|^4 \le C\left[\nu_8 \operatorname{tr} M_n^4 + \nu_4^2\left(\operatorname{tr} M_n^2\right)^2\right] \tag{3.3}$$

and

$$\mathbb{E}\left|\sum_{u\ne v} m_{uv}x_u x_v\right|^4 \le C\nu_4^2\left(\operatorname{tr} M_n^2\right)^2. \tag{3.4}$$

The next one is the classical CLT for martingales, which can be found in the book of [11], for instance.

Lemma 3.3. *Let $\{S_{ni}, F_{ni}, 1 \le i \le k_n, n \ge 1\}$ be a zero-mean, square-integrable martingale array with difference Z_{ni}. Suppose that*

$$\max_{1\le i\le k_n} |Z_{ni}| \xrightarrow{\mathbb{P}} 0, \tag{3.5}$$

and

$$\sum_{i=1}^{k_n} Z_{ni}^2 \xrightarrow{d} 1. \tag{3.6}$$

Moreover, $\mathbb{E}(\max_{1\le i\le k_n} Z_{ni}^2)$ *is bounded in* n. *Then we have*

$$S_{nk_n} \xrightarrow{d} N(0,1). \tag{3.7}$$

The next one is the moment inequality for quadratic forms. One can refer to Lemma B.26 in [3], for instance.

Lemma 3.4. *Let* $\mathbf{A} = (a_{ij})$ *be an* $n \times n$ *nonrandom matrix and* $\mathbf{X} = (x_1, x_2, \dots, x_n)$ *be a random vector of independent entries. Assume that* $\mathbb{E}x_i = 0$, $\mathbb{E}|x_i|^2 = 1$ *and* $\mathbb{E}|x_j|^\lambda \le \nu_\lambda$. *Then for any* $p \ge 1$,

$$\mathbb{E}\big|\mathbf{X}^*\mathbf{A}\mathbf{X} - \operatorname{tr}\mathbf{A}\big|^p \le C_p\big[\big(\nu_4 \operatorname{tr}\big(\mathbf{A}\mathbf{A}^*\big)\big)^{p/2} + \nu_{2p}\operatorname{tr}\big(\mathbf{A}\mathbf{A}^*\big)^{p/2}\big],$$

where C_p *is a constant depending on* p *only.*

In [13] and [4], they estimated the entries of the projection matrix P_i individually. We also obtain a similar estimate. To present it, set

$$s_1 = \lfloor(-\log\varepsilon_n)^{1/4}\rfloor, \qquad s_2 = \lfloor p\log^{-20a} p\rfloor, \tag{3.8}$$

where $\varepsilon_n = 1 - p/n \to 0$ and $a > 1/8$.

Lemma 3.5. *Recall the projection matrix* P_i *defined in* (2.6). *If* $\lim_{n\to\infty}\frac{p}{n} = 1$, *then we have*

$$\mathbb{E}\Big(\max_k p_{kk}(i)\Big) \le C\log^{-8a} p \tag{3.9}$$

and

$$\mathbb{P}\Big(\max_k p_{kk}(i) \ge \log^{-7a} p\Big) = O\big(n^{-1/3}\big) \tag{3.10}$$

for $p - s_2 \le i \le p-1$ *and* $n - p = O(n/\log^{20a} n)$.

Since the proof of Lemma 3.5 is just a slight modification of that in [4] under our setting we delay it until Appendix.

However, in our setting, the inequality (3.9) is not accurate when $n - p \ge n^{1-\delta}$ for any $0 < \delta < 1$. To improve it, our strategy is to treat it globally to get an estimate for their sum (see (3.12) below). The following is a crucial lemma. For $0 \le i \le p - s_1$ and $1 \le k \le n$, let

$$q_{kk}(i) = \frac{p_{kk}(i)}{n-i}.$$

Lemma 3.6. *Under the above notations, if $n/2 \le i \le p - s_1$, then we have*

$$\mathbb{E} p_{kk}^2(i) \le C\left[\frac{1}{1 + \frac{1}{n}\mathbb{E}\operatorname{tr} G_{(i)}(\alpha)}\right]^2 \tag{3.11}$$

for $1 \le k \le n$, where

$$G_{(i)}(\alpha) = \left(\frac{1}{n} B_{(i)} B_{(i)}^T + \alpha I_i\right)^{-1}, \qquad \alpha = n^{-1/6}.$$

In addition, if $n - p \ge n^{19/20}$, then we have

$$\sum_{i=0}^{p-s_1}\sum_{k=1}^{n} \mathbb{E} q_{kk}^2(i) = O(1). \tag{3.12}$$

We now start with the proof of Lemma 3.6. Let $\mathbf{b}_j(i)$ be the jth column of $B_{(i)}$ and $B_{(i,j)}$ denote the matrix obtained from $B_{(i)}$ by deleting the jth column $\mathbf{b}_j(i)$. Denote $\alpha = \alpha_n = n^{-1/6}$ and

$$\begin{aligned} G_{(i,k)}(\alpha) &= \left(\frac{1}{n} B_{(i,k)} B_{(i,k)}^T + \alpha I_i\right)^{-1}, \\ G_{(i)}(\alpha) &= \left(\frac{1}{n} B_{(i)} B_{(i)}^T + \alpha I_i\right)^{-1}. \end{aligned}$$

Write

$$G_{(i)}(\alpha) = \left(\frac{1}{n} B_{(i,k)} B_{(i,k)}^T + \frac{1}{n}\mathbf{b}_k(i)\mathbf{b}_k^T(i) + \alpha I_i\right)^{-1}.$$

It follows from (2.6) and the Sherman–Morrison formula that

$$\begin{aligned} p_{kk}(i) &= 1 - \mathbf{b}_k^T(i)\big(B_{(i)} B_{(i)}^T\big)^{-1}\mathbf{b}_k(i) \\ &= 1 - \mathbf{b}_k^T(i)\big(B_{(i,k)} B_{(i,k)}^T + \mathbf{b}_k(i)\mathbf{b}_k^T(i)\big)^{-1}\mathbf{b}_k(i) \\ &= 1 - \frac{\mathbf{b}_k^T(i)(B_{(i,k)} B_{(i,k)}^T)^{-1}\mathbf{b}_k(i)}{1 + \mathbf{b}_k^T(i)(B_{(i,k)} B_{(i,k)}^T)^{-1}\mathbf{b}_k(i)} \\ &= \frac{1}{1 + \mathbf{b}_k^T(i)(B_{(i,k)} B_{(i,k)}^T)^{-1}\mathbf{b}_k(i)} \\ &\le \left(1 + \frac{1}{n}\mathbf{b}_k^T(i) G_{(i,k)}(\alpha)\mathbf{b}_k(i)\right)^{-1}, \end{aligned} \tag{3.13}$$

which implies that

$$\begin{aligned}
\mathbb{E}p_{kk}^2(i) &\le \mathbb{E}\Bigg[\frac{1}{1+\frac{1}{n}\mathbf{b}_k^T(i)G_{(i,k)}(\alpha)\mathbf{b}_k(i)}\Bigg]^2 \\
&= \mathbb{E}\Bigg[\frac{1}{1+\frac{1}{n}\mathbf{b}_k^T(i)G_{(i,k)}(\alpha)\mathbf{b}_k(i)} - \frac{1}{1+\frac{1}{n}\mathbb{E}\operatorname{tr}G_{(i,k)}(\alpha)} + \frac{1}{1+\frac{1}{n}\mathbb{E}\operatorname{tr}G_{(i,k)}(\alpha)}\Bigg]^2 \\
&\le 2\Bigg[\frac{1}{1+\frac{1}{n}\mathbb{E}\operatorname{tr}G_{(i,k)}(\alpha)}\Bigg]^2 \\
&\quad + 2\mathbb{E}\Bigg[\frac{1}{1+\frac{1}{n}\mathbf{b}_k^T(i)G_{(i,k)}(\alpha)\mathbf{b}_k(i)} - \frac{1}{1+\frac{1}{n}\mathbb{E}\operatorname{tr}G_{(i,k)}(\alpha)}\Bigg]^2 \\
&\le 2\Bigg[\frac{1}{1+\frac{1}{n}\mathbb{E}\operatorname{tr}G_{(i,k)}(\alpha)}\Bigg]^2 + 2\mathbb{E}\Bigg[\frac{\frac{1}{n}\mathbf{b}_k^T(i)G_{(i,k)}(\alpha)\mathbf{b}_k(i) - \frac{1}{n}\mathbb{E}\operatorname{tr}G_{(i,k)}(\alpha)}{1+\frac{1}{n}\mathbb{E}\operatorname{tr}G_{(i,k)}(\alpha)}\Bigg]^2.
\end{aligned} \tag{3.14}$$

Write

$$K \doteq \mathbb{E}\left|\frac{1}{n}\mathbf{b}_k^T(i)G_{(i,k)}(\alpha)\mathbf{b}_k(i) - \frac{1}{n}\mathbb{E}\operatorname{tr}G_{(i,k)}(\alpha)\right|^2. \tag{3.15}$$

We obtain by the C_r-inequality that

$$\begin{aligned}
K &\le \frac{2}{n^2}\mathbb{E}\big|\mathbf{b}_k^T(i)G_{(i,k)}(\alpha)\mathbf{b}_k(i) - \operatorname{tr}G_{(i,k)}(\alpha)\big|^2 + \frac{2}{n^2}\mathbb{E}\big|\operatorname{tr}G_{(i,k)}(\alpha) - \mathbb{E}\operatorname{tr}G_{(i,k)}(\alpha)\big|^2 \\
&\doteq K_1 + K_2.
\end{aligned} \tag{3.16}$$

It follows from Lemma 3.4 that

$$\begin{aligned}
K_1 &\le \frac{C}{n^2}\big[\operatorname{tr}G_{(i,k)}^2(\alpha) + \operatorname{tr}G_{(i,k)}^2(\alpha)\big] \\
&\le \frac{C}{n^2}\big[i\alpha^{-2} + i\alpha^{-2}\big] \le \frac{C}{n\alpha^2}.
\end{aligned} \tag{3.17}$$

Similarly to the proof of (5.10) of [4], we have

$$K_2 \le \frac{C}{n\alpha^4}. \tag{3.18}$$

From (3.16)–(3.18), we can see that $K = o(1)$, which, together with (3.14), yields that

$$\mathbb{E}p_{kk}^2(i) \le C\Bigg[\frac{1}{1+\frac{1}{n}\mathbb{E}\operatorname{tr}G_{(i,k)}(\alpha)}\Bigg]^2. \tag{3.19}$$

Note that

$$\begin{aligned}
&\left[\frac{1}{1+\frac{1}{n}\mathbb{E}\operatorname{tr} G_{(i,k)}(\alpha)}\right]^2 \\
&\quad \le 2\left[\frac{1}{1+\frac{1}{n}\mathbb{E}\operatorname{tr} G_{(i)}(\alpha)}\right]^2 + 2\left[\frac{1}{1+\frac{1}{n}\mathbb{E}\operatorname{tr} G_{(i,k)}(\alpha)} - \frac{1}{1+\frac{1}{n}\mathbb{E}\operatorname{tr} G_{(i)}(\alpha)}\right]^2 \qquad (3.20)\\
&\quad \le 2\left[\frac{1}{1+\frac{1}{n}\mathbb{E}\operatorname{tr} G_{(i)}(\alpha)}\right]^2 + 2\left[\frac{\frac{1}{n}\mathbb{E}\operatorname{tr} G_{(i)}(\alpha) - \frac{1}{n}\mathbb{E}\operatorname{tr} G_{(i,k)}(\alpha)}{1+\frac{1}{n}\mathbb{E}\operatorname{tr} G_{(i)}(\alpha)}\right]^2.
\end{aligned}$$

By the Sherman–Morrison formula again, we have

$$\operatorname{tr} G_{(i)}(\alpha) - \operatorname{tr} G_{(i,k)}(\alpha) = -\frac{\frac{1}{n}\mathbf{b}_k^T(i) G_{(i,k)}^2(\alpha)\mathbf{b}_k(i)}{1+\frac{1}{n}\mathbf{b}_k^T(i) G_{(i,k)}(\alpha)\mathbf{b}_k(i)}, \qquad (3.21)$$

which yields that

$$\begin{aligned}
\left|\operatorname{tr} G_{(i)}(\alpha) - \operatorname{tr} G_{(i,k)}(\alpha)\right| &= \frac{\frac{1}{n}\mathbf{b}_k^T(i) G_{(i,k)}^2(\alpha)\mathbf{b}_k(i)}{1+\frac{1}{n}\mathbf{b}_k^T(i) G_{(i,k)}(\alpha)\mathbf{b}_k(i)} \\
&\le \alpha^{-1},
\end{aligned}$$

and thus,

$$\begin{aligned}
\left|\frac{1}{n}\mathbb{E}\operatorname{tr} G_{(i)}(\alpha) - \frac{1}{n}\mathbb{E}\operatorname{tr} G_{(i,k)}(\alpha)\right|^2 &\le \frac{1}{n^2}\mathbb{E}\left|\operatorname{tr} G_{(i)}(\alpha) - \operatorname{tr} G_{(i,k)}(\alpha)\right|^2 \\
&\le \frac{1}{n^2\alpha^2} = o(1).
\end{aligned} \qquad (3.22)$$

Hence, the desired result (3.11) follows immediately from (3.19)–(3.22).

The next aim is to show that (3.12) holds. Note that $q_{kk}(i) = p_{kk}(i)/(n-i) \le 1/(n-i)$ and $\sum_{k=1}^n q_{kk}(i) = \sum_{k=1}^n p_{kk}(i)/(n-i) = 1$. We have

$$\begin{aligned}
\sum_{i=0}^{p-s_1}\sum_{k=1}^{n}\mathbb{E}q_{kk}^2(i) &= \sum_{i=0}^{\lfloor n/2\rfloor}\sum_{k=1}^{n}\mathbb{E}q_{kk}^2(i) + \sum_{i=\lfloor n/2\rfloor+1}^{p-s_1}\sum_{k=1}^{n}\mathbb{E}q_{kk}^2(i) \\
&\le \sum_{i=0}^{\lfloor n/2\rfloor}\frac{1}{n-i} + \sum_{i=\lfloor n/2\rfloor+1}^{p-s_1}\frac{1}{(n-i)^2}\sum_{k=1}^{n}\mathbb{E}p_{kk}^2(i) \qquad (3.23)\\
&\le O(1) + C\sum_{i=\lfloor n/2\rfloor+1}^{p-s_1}\frac{n}{(n-i)^2}\left[\frac{1}{1+\frac{1}{n}\mathbb{E}\operatorname{tr} G_{(i)}(\alpha)}\right]^2.
\end{aligned}$$

By Lemma A.1, we can see that for $n/2 \le i \le p - s_1$,

$$\mathbb{E}\Big(\frac{1}{i}\operatorname{tr} G_{(i)}(\alpha)\Big) = s_i(\alpha) + O\big(n^{-1/6}\big), \tag{3.24}$$

where

$$s_i(\alpha) = 2\bigg(\alpha + 1 - \frac{i}{n} + \sqrt{\Big(\alpha + 1 - \frac{i}{n}\Big)^2 + \frac{4\alpha i}{n}}\bigg)^{-1}. \tag{3.25}$$

Hence, we have by (3.23)–(3.25) that

$$\begin{aligned}\sum_{i=0}^{p-s_1}\sum_{k=1}^{n}\mathbb{E}q_{kk}^2(i) &\le O(1) + C\sum_{i=\lfloor n/2\rfloor+1}^{p-s_1}\frac{n}{(n-i)^2}\bigg[\frac{1}{1+\frac{i}{n}\cdot\mathbb{E}(\frac{1}{i}\operatorname{tr} G_{(i)}(\alpha))}\bigg]^2 \\ &\le O(1) + C\sum_{i=\lfloor n/2\rfloor+1}^{p-s_1}\frac{n}{(n-i)^2}\Bigg[\frac{1}{1+\frac{i}{n}\cdot\frac{2}{\alpha+1-\frac{i}{n}+\sqrt{(\alpha+1-\frac{i}{n})^2+\frac{4\alpha i}{n}}}}\Bigg]^2.\end{aligned} \tag{3.26}$$

One can see that if

$$\frac{i}{n} \le \frac{2+\alpha-\sqrt{\alpha^2+4\alpha}}{2},$$

namely,

$$i \le n\bigg(1 - \frac{\sqrt{\alpha^2+4\alpha}-\alpha}{2}\bigg),$$

then we have

$$\frac{n-i}{n} \ge \alpha \quad \text{and} \quad \Big(\frac{n-i}{n}\Big)^2 \ge \alpha\cdot\frac{i}{n}. \tag{3.27}$$

Noting that $n - p \ge n^{19/20}$, $i \le p - s_1$ and $\frac{\sqrt{\alpha^2+4\alpha}-\alpha}{2} \sim \sqrt{\alpha} = n^{-1/12}$, we have

$$i < p \le n\big(1 - n^{-1/20}\big) \le n\bigg(1 - \frac{\sqrt{\alpha^2+4\alpha}-\alpha}{2}\bigg).$$

Hence, (3.27) is satisfied, which together with (3.26) yields that

$$\begin{aligned}\sum_{i=0}^{p-s_1}\sum_{k=1}^{n}\mathbb{E}q_{kk}^2(i) &\le O(1) + C\sum_{i=\lfloor n/2\rfloor+1}^{p-s_1}\frac{n}{(n-i)^2}\bigg[\frac{1}{1+\frac{2i}{n}\cdot\frac{n}{(2+2\sqrt{2})(n-i)}}\bigg]^2 \\ &= O(1) + C\sum_{i=\lfloor n/2\rfloor+1}^{p-s_1}\frac{n}{(n-i)^2}\bigg[\frac{(2+2\sqrt{2})(n-i)}{(2+2\sqrt{2})(n-i)+2i}\bigg]^2\end{aligned} \tag{3.28}$$

$$\leq O(1) + C \sum_{i=\lfloor n/2 \rfloor+1}^{p-s_1} \frac{n}{(n-i)^2} \cdot \frac{(n-i)^2}{n^2}$$

$$= O(1) + C \sum_{i=\lfloor n/2 \rfloor+1}^{p-s_1} \frac{1}{n} = O(1).$$

Therefore, (3.12) holds. This completes the proof of Lemma 3.6.

4. Proofs of the main results

4.1. Proof of Theorem 1.1

The proof of Theorem 1.1 follows directly from Lemma 3.1, which was provided by [4]. Here, we give the details as follows.

Note that the square matrix $B = (b_{jk})_{n\times n}$ satisfies the conditions of Lemma 3.1. Applying Lemma 3.1 with $A_n = B$, we have by (1.10) that

$$\frac{\log\det B^2 - \log(n-1)!}{\sqrt{2\log n}} \xrightarrow{d} N(0,1), \tag{4.1}$$

which together with (1.8) yields the desired result (1.11) immediately. The proof is complete.

4.2. Proof of Theorem 1.2

For convenience of the reader, we present the complete proof of Theorem 1.2.

By Proposition 2.1, we assume that all square submatrices of B are invertible. In view of (2.12), to prove (1.14), it suffices to show the following:

(i)

$$I_1 \doteq \frac{1}{\sqrt{-2\log(1-\frac{p}{n})}} \sum_{i=0}^{p-s_1} X_{i+1} \xrightarrow{d} N(0,1); \tag{4.2}$$

(ii)

$$I_2 \doteq \frac{1}{\sqrt{-2\log(1-\frac{p}{n})}} \left[\sum_{i=0}^{p-s_1} \frac{X_{i+1}^2}{2} + \log\left(1-\frac{p}{n}\right)\right] \xrightarrow{\mathbb{P}} 0; \tag{4.3}$$

(iii)

$$I_3 \doteq \frac{1}{\sqrt{-2\log(1-\frac{p}{n})}} \sum_{i=0}^{p-s_1} R_{i+1} \xrightarrow{\mathbb{P}} 0; \tag{4.4}$$

(iv) if the last s_1 rows of B are Gaussian, then

$$I_4 \doteq \frac{1}{\sqrt{-2\log(1-\frac{p}{n})}} \sum_{i=p-s_1+1}^{p-1} \log \frac{r_{i+1,i+1}^2}{n-i} \xrightarrow{\mathbb{P}} 0; \tag{4.5}$$

(v) let $\bar{B}$ be a random matrix satisfying the basic Assumptions C_0 and differ from B only in the last s_1 rows. Then

$$\frac{\log\det(BB^T)}{\sqrt{-2\log(1-\frac{p}{n})}} - \frac{\log\det(\bar{B}\bar{B}^T)}{\sqrt{-2\log(1-\frac{p}{n})}} \xrightarrow{\mathbb{P}} 0, \tag{4.6}$$

and thus,

$$\frac{\log\det(BB^T) - p\log n - \tau_{n,p}}{\sqrt{-2\log(1-\frac{p}{n})}} - \frac{\log\det(\bar{B}\bar{B}^T) - p\log n - \tau_{n,p}}{\sqrt{-2\log(1-\frac{p}{n})}} \xrightarrow{\mathbb{P}} 0. \tag{4.7}$$

4.2.1. *Proofs of* (i) *and* (ii)

Unlike [13] and [4], we have to distinguish two cases to prove (i) and (ii).

Case I. $1 \le n-p \le n^{19/20}$.
In this case, one can easily see that

$$\frac{1}{\sqrt{-2\log(1-\frac{p}{n})}} = O\left(\frac{1}{\sqrt{\log n}}\right).$$

The remaining proofs of (i) and (ii) are similar to the corresponding ones of Theorem 1.1 in [4] by using Lemmas 3.5, 3.2 and 3.3.

Case II. $n^{19/20} \le n-p = n\varepsilon_n$, where $\varepsilon_n = 1-p/n \to 0$.

Applying Lemma 3.3 with $k_n = p - s_1$, $Z_{ni} = X_i/\sqrt{-2\log(1-\frac{p}{n})}$, we only need to show that

$$\frac{1}{\sqrt{-2\log(1-\frac{p}{n})}} \max_{0\le i\le p-s_1} |X_{i+1}| \xrightarrow{\mathbb{P}} 0, \tag{4.8}$$

$$\frac{1}{-2\log(1-\frac{p}{n})} \sum_{i=0}^{p-s_1} X_{i+1}^2 \xrightarrow{\mathbb{P}} 1, \tag{4.9}$$

and

$$\frac{1}{-2\log(1-\frac{p}{n})} \mathbb{E}\Big(\max_{0\le i\le p-s_1} X_{i+1}^2\Big) \le C \tag{4.10}$$

for some positive constant C independent of p and n.

First, we prove (4.8). It suffices to show that for any $\varepsilon > 0$,

$$\sum_{i=0}^{p-s_1} \mathbb{P}\left(\frac{1}{\sqrt{-2\log(1-\frac{p}{n})}}|X_{i+1}| \ge \varepsilon\right) \to 0 \qquad \text{as } n \to \infty. \tag{4.11}$$

For $0 \le i \le p - s_1$, denote

$$q_{jk}(i) = \frac{1}{n-i} p_{jk}(i), \qquad Q_i \doteq \big(q_{jk}(i)\big)_{n\times n} \doteq \frac{1}{n-i} P_i.$$

Noting that $\operatorname{tr} Q_i = 1$ from $\operatorname{tr} P_i = n - i$, we have by (2.5) and (2.9) that

$$X_{i+1} = \sum_{k=1}^{n} q_{kk}(i)\big(b_{i+1,k}^2 - 1\big) + \sum_{u\ne v} q_{uv}(i) b_{i+1,u} b_{i+1,v} \doteq U_{i+1} + V_{i+1}. \tag{4.12}$$

Hence, to prove (4.11), we only need to show

$$\frac{1}{-2\log(1-\frac{p}{n})} \sum_{i=0}^{p-s_1} \mathbb{E}U_{i+1}^2 \to 0 \qquad \text{as } n \to \infty \tag{4.13}$$

and

$$\frac{1}{[-2\log(1-\frac{p}{n})]^2} \sum_{i=0}^{p-s_1} \mathbb{E}V_{i+1}^4 \to 0 \qquad \text{as } n \to \infty. \tag{4.14}$$

Note that

$$\mathbb{E}U_{i+1}^2 = \mathbb{E}\left[\sum_{k=1}^{n} q_{kk}(i)\big(b_{i+1,k}^2 - 1\big)\right]^2 \le C \sum_{k=1}^{n} \mathbb{E}q_{kk}^2(i). \tag{4.15}$$

We have by $n^{19/20} \le n - p$, $p/n \to 1$ and (3.12) that

$$\frac{1}{-2\log(1-\frac{p}{n})} \sum_{i=0}^{p-s_1} \mathbb{E}U_{i+1}^2 \le \frac{C}{-2\log(1-\frac{p}{n})} \sum_{i=0}^{p-s_1} \sum_{k=1}^{n} \mathbb{E}q_{kk}^2(i) = o(1),$$

which implies (4.13). On the other hand, noting that $\varepsilon_n = 1 - p/n \to 0$ and $s_1 = \lfloor(-\log\varepsilon_n)^{1/4}\rfloor \to \infty$, we have by (3.4) that

$$\begin{aligned}\frac{1}{[-2\log(1-\frac{p}{n})]^2} \sum_{i=0}^{p-s_1} \mathbb{E}V_{i+1}^4 &\le \frac{C}{(\log\varepsilon_n)^2} \sum_{i=0}^{p-s_1} \mathbb{E}\big(\operatorname{tr} Q_i^2\big)^2 \\ &\le \frac{C}{(\log\varepsilon_n)^2} \sum_{i=0}^{p-s_1} \frac{1}{(n-i)^2}\end{aligned}$$

$$
\begin{aligned}
&\le \frac{C}{(\log \varepsilon_n)^2} \cdot \frac{1}{s_1} \\
&= o(1),
\end{aligned}
$$

which implies (4.14). Hence, (4.8) is proved.

Next, we prove (4.9). Note that (ii) implies (4.9) directly. So we only need to show (ii), namely,

$$
\frac{1}{\sqrt{-2\log(1-\frac{p}{n})}}\left[\sum_{i=0}^{p-s_1} X_{i+1}^2 + 2\log\left(1-\frac{p}{n}\right)\right] \xrightarrow{\mathbb{P}} 0, \tag{4.16}
$$

which can be implied by

$$
\frac{1}{\sqrt{-2\log(1-\frac{p}{n})}}\sum_{i=0}^{p-s_1}\left[X_{i+1}^2 - \mathbb{E}\left(X_{i+1}^2|\mathcal{F}_i\right)\right] \xrightarrow{\mathbb{P}} 0, \tag{4.17}
$$

and

$$
\frac{1}{\sqrt{-2\log(1-\frac{p}{n})}}\left[\sum_{i=0}^{p-s_1} \mathbb{E}\left(X_{i+1}^2|\mathcal{F}_i\right) + 2\log\left(1-\frac{p}{n}\right)\right] \xrightarrow{\mathbb{P}} 0. \tag{4.18}
$$

It can be checked that

$$
\begin{aligned}
\mathbb{E}\left(X_{i+1}^2|\mathcal{F}_i\right) &= \mathbb{E}\left[\left(\sum_{j,k=1}^{n} q_{jk}(i) b_{i+1,j} b_{i+1,k} - 1\right)^2 \middle| \mathcal{F}_i\right] \\
&= \frac{2}{n-i} + \sum_{k=1}^{n} q_{kk}^2(i)\left(\mathbb{E}b_{i+1,k}^4 - 3\right)
\end{aligned} \tag{4.19}
$$

and

$$
\begin{aligned}
&\sum_{i=0}^{p-s_1} \frac{2}{n-i} + 2\log\left(1-\frac{p}{n}\right) \\
&\quad = 2\left[\log n - \log(n-p+s_1)\right] + O\left(\frac{1}{n-p+s_1}\right) + 2\log\left(1-\frac{p}{n}\right) \\
&\quad = 2\left[\log(n-p) - \log(n-p+s_1)\right] + O\left(\frac{1}{s_1}\right) \\
&\quad = 2\log\left(1 - \frac{s_1}{n-p+s_1}\right) + O\left(\frac{1}{s_1}\right) \\
&\quad = o(1),
\end{aligned} \tag{4.20}
$$

where the last equality above uses the fact that $n^{19/20} \le n - p$, $\varepsilon_n = 1 - p/n \to 0$ and $s_1 = \lfloor(-\log \varepsilon_n)^{1/4}\rfloor \to \infty$. It follows from $n^{19/20} \le n - p$, $1 - p/n \to 0$ and (3.12) again that

$$\frac{1}{\sqrt{-2\log(1-\frac{p}{n})}} \sum_{i=0}^{p-s_1} \sum_{k=1}^{n} \mathbb{E} q_{kk}^2(i) = o(1). \tag{4.21}$$

Hence, (4.18) follows immediately from (4.19)–(4.21).

For (4.17), we conclude from (4.12) and (4.19) that

$$\begin{aligned}
&X_{i+1}^2 - \mathbb{E}\big(X_{i+1}^2 | \mathcal{F}_i\big) \\
&\quad = 2\Bigg[-\sum_{u=1}^{n} q_{uu}(i)\big(b_{i+1,u}^2 - 1\big) + 2\sum_{u \neq v} q_{uv}^2(i)\big(b_{i+1,u}^2 b_{i+1,v}^2 - 1\big) \\
&\qquad + \sum_{u \neq v} q_{uu}(i) q_{vv}(i)\big(b_{i+1,u}^2 b_{i+1,v}^2 - 1\big) \\
&\qquad + \sum_{u_1 \neq v_1, u_2 \neq v_2, \{u_1, v_1\} \neq \{u_2, v_3\}} q_{u_1 v_1}(i) q_{u_2 v_2}(i) b_{i+1,u_1} b_{i+1,v_1} b_{i+1,u_2} b_{i+1,v_2}\Bigg] \\
&\qquad + 2\Bigg[\frac{1}{2}\sum_{u=1}^{n} q_{uu}^2(i)\big(b_{i+1,u}^4 - \mathbb{E} b_{i+1,u}^4\big) \\
&\qquad + \left(\sum_{u=1}^{n} q_{uu}(i)\big(b_{i+1,u}^2 - 1\big)\right) \cdot \left(\sum_{u \neq v} q_{uv}(i) b_{i+1,u} b_{i+1,v}\right)\Bigg] \\
&\quad \doteq 2W_1(i) + 2W_2(i)
\end{aligned}$$

(one may refer to formula below (3.1) in [4]). To prove (4.17), it suffices to show

$$J_1 \doteq \frac{1}{-2\log(1-\frac{p}{n})} \mathbb{E}\left(\sum_{i=0}^{p-s_1} W_1(i)\right)^2 \to 0 \qquad \text{as } n \to \infty \tag{4.22}$$

and

$$J_2 \doteq \frac{1}{\sqrt{-2\log(1-\frac{p}{n})}} \sum_{i=0}^{p-s_1} \mathbb{E}\big|W_2(i)\big| \to 0 \qquad \text{as } n \to \infty. \tag{4.23}$$

First, we deal with (4.23). Noting that $n-p=n\varepsilon_n$, where $\varepsilon_n=1-p/n\to 0$, we obtain by the Cauchy–Schwarz inequality that

$$
\begin{aligned}
J_2 &\le \frac{C}{\sqrt{-\log\varepsilon_n}}\sum_{i=0}^{p-s_1}\sum_{k=1}^{n}\mathbb{E}q_{kk}^2(i)\\
&\quad+\frac{C}{\sqrt{-\log\varepsilon_n}}\sum_{i=0}^{p-s_1}\left[\mathbb{E}\left(\sum_{u=1}^{n}q_{uu}(i)\left(b_{i+1,u}^2-1\right)\right)^2\right]^{1/2}\cdot\left[\mathbb{E}\left(\sum_{u\neq v}q_{uv}(i)b_{i+1,u}b_{i+1,v}\right)^2\right]^{1/2}\\
&\le \frac{C}{\sqrt{-\log\varepsilon_n}}\sum_{i=0}^{p-s_1}\sum_{k=1}^{n}\mathbb{E}q_{kk}^2(i)+\frac{C}{\sqrt{\log\log n}}\sum_{i=0}^{p-s_1}\left(\frac{1}{n-i}\right)^{1/2}\cdot\left(\sum_{k=1}^{n}\mathbb{E}q_{kk}^2(i)\right)^{1/2}\\
&\doteq J_{21}+J_{22}.
\end{aligned}
\tag{4.24}
$$

It follows from (3.12) that $J_{21}=o(1)$. For J_{22}, noting that $q_{kk}(i)=\frac{p_{kk}(i)}{n-i}\le\frac{1}{n-i}$ and $\sum_k q_{kk}(i)=1$, we have by (3.11) and the proofs of (3.26) and (3.28) that

$$
\begin{aligned}
J_{22}&=\frac{C}{\sqrt{-\log\varepsilon_n}}\sum_{i=0}^{p-s_1}\left(\frac{1}{n-i}\right)^{1/2}\cdot\left(\sum_{k=1}^{n}\mathbb{E}q_{kk}^2(i)\right)^{1/2}\\
&\le\frac{C}{\sqrt{-\log\varepsilon_n}}\sum_{i=0}^{\lfloor n/2\rfloor}\frac{1}{n-i}+\frac{C}{\sqrt{-\log\varepsilon_n}}\sum_{i=\lfloor n/2\rfloor+1}^{p-s_1}\left(\frac{1}{n-i}\right)^{1/2}\cdot\left(\sum_{k=1}^{n}\mathbb{E}q_{kk}^2(i)\right)^{1/2}\\
&\le o(1)+\frac{C}{\sqrt{-\log\varepsilon_n}}\sum_{i=\lfloor n/2\rfloor+1}^{p-s_1}\left(\frac{1}{n-i}\right)^{3/2}\cdot\left(\sum_{k=1}^{n}\mathbb{E}p_{kk}^2(i)\right)^{1/2}\\
&\le o(1)+\frac{C}{\sqrt{-\log\varepsilon_n}}\sum_{i=\lfloor n/2\rfloor+1}^{p-s_1}\frac{n^{1/2}}{(n-i)^{3/2}}\cdot\frac{1}{1+\frac{1}{n}\mathbb{E}\operatorname{tr}G_{(i)}(\alpha)}\\
&\le o(1)+\frac{C}{\sqrt{-\log\varepsilon_n}}\sum_{i=\lfloor n/2\rfloor+1}^{p-s_1}\frac{n^{1/2}}{(n-i)^{3/2}}\cdot\frac{n-i}{n}\\
&=o(1)+\frac{C}{\sqrt{-\log\varepsilon_n}}\sum_{i=\lfloor n/2\rfloor+1}^{p-s_1}\frac{1}{(n-i)^{1/2}\cdot n^{1/2}}\\
&=o(1).
\end{aligned}
\tag{4.25}
$$

This finishes the proof of (4.23) in view of (4.24), (4.25) and $J_{21}=o(1)$.

Second, we deal with (4.22). It is easy to check that

$$
\mathbb{E}W_1(i)=0,\qquad \mathbb{E}W_1(i)W_1(j)=0,\qquad i\neq j,
$$

which yields that

$$
\begin{aligned}
J_1 &= \frac{1}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}W_1^2(i)\\
&\le \frac{C}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}\left(\sum_{u=1}^{n}q_{uu}^2(i)\right)+\frac{C}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}\Bigg(\sum_{u\neq v,u\neq w}q_{uv}^2(i)q_{uw}^2(i)\Bigg)\\
&\quad+\frac{C}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}\Bigg(\sum_{u\neq v,u\neq w}q_{uu}^2(i)q_{vv}(i)q_{ww}(i)\Bigg) \qquad (4.26)\\
&\quad+\frac{C}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}\Bigg(\sum_{u_1\neq v_1,u_2\neq v_2}\left|q_{u_1v_1}(i)q_{u_1v_2}(i)q_{u_2v_1}(i)q_{u_2v_2}(i)\right|\Bigg)\\
&\doteq J_{11}+J_{12}+J_{13}+J_{14}.
\end{aligned}
$$

Noting that $n^{19/20}\le n-p$, $1-p/n\to 0$ and $\sum_{u,v}q_{uv}^2(i)=\frac{1}{n-i}$, we conclude from (3.12) and Cauchy–Schwarz's inequality (if needed) again that

$$
J_{11}\doteq\frac{C}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}\left(\sum_{u=1}^{n}q_{uu}^2(i)\right)=o(1), \qquad (4.27)
$$

$$
\begin{aligned}
J_{12}&\doteq\frac{C}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}\Bigg(\sum_{u\neq v,u\neq w}q_{uv}^2(i)q_{uw}^2(i)\Bigg)\\
&\le\frac{C}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}\left(\sum_{u,v}q_{uv}^2(i)\right)^2 \qquad (4.28)\\
&=\frac{C}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\frac{1}{(n-i)^2}=o(1),
\end{aligned}
$$

$$
\begin{aligned}
J_{13}&\doteq\frac{C}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}\Bigg(\sum_{u\neq v,u\neq w}q_{uu}^2(i)q_{vv}(i)q_{ww}(i)\Bigg)\\
&\le\frac{C}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}\left(\sum_{u=1}^{n}q_{uu}^2(i)\right)\left(\sum_{v=1}^{n}q_{vv}(i)\right)^2 \qquad (4.29)\\
&=\frac{C}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}\left(\sum_{u=1}^{n}q_{uu}^2(i)\right)=o(1),
\end{aligned}
$$

and

$$
\begin{aligned}
J_{14} &\doteq \frac{C}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}\Big(\sum_{u_1\neq v_1,u_2\neq v_2}\big|q_{u_1v_1}(i)q_{u_1v_2}(i)q_{u_2v_1}(i)q_{u_2v_2}(i)\big|\Big) \\
&\le \frac{C}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}\Big(\sum_{u,v}q_{uv}^2(i)\Big)^2 = o(1).
\end{aligned}
\tag{4.30}
$$

Hence, (4.22) follows immediately from (4.27)–(4.30), which together with (4.23) yields (4.17). Combining (4.17) with (4.18), we can get (4.16). Thus, (4.9) and (ii) have been proved.

To end the proof of (i), we claim that (4.10) holds. Actually, it follows from (4.19) and (3.12) that

$$
\begin{aligned}
\frac{1}{-2\log(1-\frac{p}{n})}\mathbb{E}\Big(\max_{0\le i\le p-s_1}X_{i+1}^2\Big) &\le \frac{1}{-2\log(1-\frac{p}{n})}\sum_{i=0}^{p-s_1}\mathbb{E}X_{i+1}^2 \\
&\le \frac{C}{-2\log(1-\frac{p}{n})}\Bigg[\sum_{i=0}^{p-s_1}\frac{1}{n-i}+\sum_{i=0}^{p-s_1}\sum_{k=1}^{n}\mathbb{E}q_{kk}^2(i)\Bigg] \\
&\le C\cdot\frac{\log n-\log(n-p+s_1)}{\log n-\log(n-p)}+o(1) \\
&\le C,
\end{aligned}
$$

which implies (4.10). This completes the proofs of (i) and (ii).

4.2.2. *Proof of* (iii) *in* (4.4)

If $1\le n-p\le n^{19/20}$, then the proof of (iii) is still similar to the corresponding one of Theorem 1.1 in [4]. So we only need to consider $n^{19/20}\le n-p=n\varepsilon_n$, where $\varepsilon_n=1-p/n\to 0$.

To prove (iii), we need the following two lemmas, whose proofs will be given in Appendix.

Lemma 4.1. *If* $X_{i+1}\ge -1+(\log\log\varepsilon_n^{-1})^{-1}$, *where* $\varepsilon_n=1-p/n\to 0$, *then one has*

$$
|R_{i+1}|\le C\big(U_{i+1}^2+|V_{i+1}|^{2+\delta}\big)\log\log\varepsilon_n^{-1}
$$

for any $0<\delta\le 1$. *Here* $C\doteq C(\delta)$ *is a positive constant depending on* δ *only.*

Lemma 4.2. *Under the Assumption* C_0, *we have*

$$
\sum_{i=0}^{p-s_1}\mathbb{P}\big(X_{i+1}<-1+\big(\log\log\varepsilon_n^{-1}\big)^{-1}\big)\to 0 \qquad \textit{as } n\to\infty,
$$

where $\varepsilon_n=1-p/n\to 0$.

Applying Lemmas 4.1 and 4.2, we have with probability $1 - o(1)$ that

$$|R_{i+1}| \le C\big(U_{i+1}^2 + |V_{i+1}|^{2+\delta}\big) \log\log \varepsilon_n^{-1}, \qquad 0 \le i \le p - s_1.$$

Hence, to prove (iii), it suffices to show

$$\frac{\log\log \varepsilon_n^{-1}}{\sqrt{-2\log(1-\frac{p}{n})}} \sum_{i=0}^{p-s_1} \big(U_{i+1}^2 + |V_{i+1}|^{2+\delta}\big) \xrightarrow{\mathbb{P}} 0. \tag{4.31}$$

Noting that $n^{19/20} \le n - p = n\varepsilon_n$, where $\varepsilon_n = 1 - p/n \to 0$, we have by (3.12) that

$$\frac{\log\log \varepsilon_n^{-1}}{\sqrt{-2\log(1-\frac{p}{n})}} \sum_{i=0}^{p-s_1} \mathbb{E}U_{i+1}^2 \le \frac{C\log\log \varepsilon_n^{-1}}{\sqrt{\log \varepsilon_n^{-1}}} \sum_{i=0}^{p-s_1} \sum_{k=1}^{n} \mathbb{E}q_{kk}^2(i) = o(1). \tag{4.32}$$

Moreover, we have

$$\begin{aligned} \frac{\log\log \varepsilon_n^{-1}}{\sqrt{-2\log(1-\frac{p}{n})}} \sum_{i=0}^{p-s_1} \mathbb{E}|V_{i+1}|^{2+\delta} &\le \frac{C\log\log \varepsilon_n^{-1}}{\sqrt{\log \varepsilon_n^{-1}}} \sum_{i=0}^{p-s_1} \big(\mathbb{E}V_{i+1}^4\big)^{\frac{2+\delta}{4}} \\ &\le \frac{C\log\log \varepsilon_n^{-1}}{\sqrt{\log \varepsilon_n^{-1}}} \sum_{i=0}^{p-s_1} (n-i)^{-\frac{2+\delta}{2}} \\ &= o(1), \end{aligned} \tag{4.33}$$

which together with (4.32) yields (4.31). This completes the proof of (iii).

4.2.3. *Proofs of* (iv) *and* (v)

For (iv), note that when the last s_1 rows of B are Gaussian, $\{r_{i+1,i+1}^2, i = p - s_1 + 1, \ldots, p-1\}$ are independent random variables and

$$r_{i+1,i+1}^2 = \mathbf{b}_{i+1}^T P_i \mathbf{b}_{i+1} \sim \chi^2_{\mathrm{tr}(P_i)} = \chi^2_{n-i}, \qquad i = p - s_1 + 1, \ldots, p-1.$$

One may refer to Section 7 of [13] for instance.

Similarly to the proof of equality (7.1) of [13], one can get (4.5) immediately.

Finally, we prove (v). The proof is different from that in [4]. We have to distinguish two different cases according as $(n-p) \ge d_n = s_1^{1/16}$ or $(n-p) < d_n$. Moreover, we use induction to handle the case of $(n-p) < d_n$. In addition, our overall strategy is to replace one row at each step, and derive the difference between the distributions of the logarithms of the magnitudes of two adjacent determinants. Hence, it suffices to compare two matrices with only one different row. Without loss of generality, we only need to compare two random matrices $B = (b_{jk})_{p\times n}$ and $\bar{C} = (\bar{b}_{jk})_{p\times n}$ satisfying Assumption C_0 such that they only differ in the last row. Assume that $b_{jk}=\bar{b}_{jk}$, $1\le j\le p-1$, $1\le k\le n$, and $\mathbf{b}_p^T$ and $\bar{\mathbf{b}}_p^T$ are independent. Here we use $\mathbf{b}_p^T$ and $\bar{\mathbf{b}}_p^T$ to denote the pth row of B and $\bar{C}$.

To prove (v), we consider two cases.

(i) $n-p\ge d_n=s_1^{1/16}\to\infty$, where $s_1=\lfloor(-\log\varepsilon_n)^{1/4}\rfloor$ and $\varepsilon_n=1-p/n\to 0$. Set

$$B=(b_{jk})_{p\times n}=\begin{pmatrix}\mathbf{b}_1^T\\ \vdots\\ \mathbf{b}_{p-1}^T\\ \mathbf{b}_p^T\end{pmatrix}=\begin{pmatrix}B_{(p-1)}\\ \mathbf{b}_p^T\end{pmatrix},\qquad \bar{C}=(\bar{b}_{jk})_{p\times n}=\begin{pmatrix}\mathbf{b}_1^T\\ \vdots\\ \mathbf{b}_{p-1}^T\\ \bar{\mathbf{b}}_p^T\end{pmatrix}=\begin{pmatrix}B_{(p-1)}\\ \bar{\mathbf{b}}_p^T\end{pmatrix}.$$

Note that the following basic properties of determinant:

$$\det\begin{pmatrix}A & B\\ C & D\end{pmatrix}=\det\bigl(A-BD^{-1}C\bigr)\det D,$$
$$\det\bigl(A+\mathbf{u}\mathbf{v}^T\bigr)=\bigl(1+\mathbf{v}^TA^{-1}\mathbf{u}\bigr)\det A.$$

These and (2.6) yield

$$\begin{aligned}
&\det\bigl(BB^T\bigr)\\
&=\det\begin{pmatrix}B_{(p-1)}B_{(p-1)}^T & B_{(p-1)}\mathbf{b}_p\\ \mathbf{b}_p^TB_{(p-1)}^T & \mathbf{b}_p^T\mathbf{b}_p\end{pmatrix}\\
&=\det\bigl(B_{(p-1)}B_{(p-1)}^T-B_{(p-1)}\mathbf{b}_p\bigl(\mathbf{b}_p^T\mathbf{b}_p\bigr)^{-1}\mathbf{b}_p^TB_{(p-1)}^T\bigr)\cdot\det\bigl(\mathbf{b}_p^T\mathbf{b}_p\bigr)\\
&=\bigl[1-\mathbf{b}_p^TB_{(p-1)}^T\bigl(B_{(p-1)}B_{(p-1)}^T\bigr)^{-1}B_{(p-1)}\mathbf{b}_p\bigl(\mathbf{b}_p^T\mathbf{b}_p\bigr)^{-1}\bigr]\det\bigl(B_{(p-1)}B_{(p-1)}^T\bigr)\cdot\mathbf{b}_p^T\mathbf{b}_p\\
&=\mathbf{b}_p^T\bigl[I_n-B_{(p-1)}^T\bigl(B_{(p-1)}B_{(p-1)}^T\bigr)^{-1}B_{(p-1)}\bigr]\mathbf{b}_p\cdot\det\bigl(B_{(p-1)}B_{(p-1)}^T\bigr)\\
&=\mathbf{b}_p^TP_{p-1}\mathbf{b}_p\cdot\det\bigl(B_{(p-1)}B_{(p-1)}^T\bigr)\\
&=r_{pp}^2\cdot\det\bigl(B_{(p-1)}B_{(p-1)}^T\bigr),
\end{aligned}\tag{4.34}$$

where

$$\begin{aligned}
r_{pp}^2&=\mathbf{b}_p^TP_{p-1}\mathbf{b}_p=\mathbf{b}_p^T\bigl(p_{jk}(p-1)\bigr)_{n\times n}\mathbf{b}_p\\
&=\sum_{\substack{1\le j,k\le n\\ j\ne k}}p_{jk}(p-1)b_{pj}b_{pk}+\sum_{k=1}^np_{kk}(p-1)b_{pk}^2.
\end{aligned}\tag{4.35}$$

Similarly, we have

$$\begin{aligned}
\det\bigl(\bar{C}\bar{C}^T\bigr)&=\det\begin{pmatrix}B_{(p-1)}B_{(p-1)}^T & B_{(p-1)}\bar{\mathbf{b}}_p\\ \bar{\mathbf{b}}_p^TB_{(p-1)}^T & \bar{\mathbf{b}}_p^T\bar{\mathbf{b}}_p\end{pmatrix}\\
&=\bar{r}_{pp}^2\cdot\det\bigl(B_{(p-1)}B_{(p-1)}^T\bigr),
\end{aligned}\tag{4.36}$$

where

$$\begin{aligned}\bar{r}_{pp}^2 &= \bar{\mathbf{b}}_p^T P_{p-1}\bar{\mathbf{b}}_p = \bar{\mathbf{b}}_p^T \big(p_{jk}(p-1)\big)_{n\times n}\bar{\mathbf{b}}_p \\ &= \sum_{\substack{1\le j,k\le n\\ j\ne k}} p_{jk}(p-1)\bar{b}_{pj}\bar{b}_{pk} + \sum_{k=1}^{n} p_{kk}(p-1)\bar{b}_{pk}^2.\end{aligned} \tag{4.37}$$

From (4.34) and (4.36), we can get that

$$\frac{\log\det(BB^T)}{\sqrt{-2\log(1-\frac{p}{n})}} - \frac{\log\det(\bar{C}\bar{C}^T)}{\sqrt{-2\log(1-\frac{p}{n})}} = \frac{\log r_{pp}^2 - \log\mathbb{E}r_{pp}^2}{\sqrt{-2\log(1-\frac{p}{n})}} - \frac{\log\bar{r}_{pp}^2 - \log\mathbb{E}\bar{r}_{pp}^2}{\sqrt{-2\log(1-\frac{p}{n})}}, \tag{4.38}$$

where $\mathbb{E}r_{pp}^2 = \mathbb{E}\bar{r}_{pp}^2 = n-p+1$ from (2.7).

We next show that

$$\frac{\log r_{pp}^2 - \log\mathbb{E}r_{pp}^2}{\sqrt{-2\log(1-\frac{p}{n})}} = O_{\mathbb{P}}\big(s_1^{-3/2}\big) \tag{4.39}$$

and

$$\frac{\log\bar{r}_{pp}^2 - \log\mathbb{E}\bar{r}_{pp}^2}{\sqrt{-2\log(1-\frac{p}{n})}} = O_{\mathbb{P}}\big(s_1^{-3/2}\big). \tag{4.40}$$

Note that $P_{p-1} = I_n - B_{(p-1)}^T(B_{(p-1)}B_{(p-1)}^T)^{-1}B_{(p-1)}$ is a projection matrix. Hence, the eigenvalues of P_{p-1} are 0 or 1. Since $\mathrm{Rank}(P_{p-1}) = \mathrm{tr}(P_{p-1}) = n-p+1$, we have

$$\|P_{p-1}\|^2 = \sum_{j,k=1}^{n} p_{jk}^2(p-1) = n-p+1. \tag{4.41}$$

It can be easily checked by (4.35) that

$$\mathrm{Var}\big(r_{pp}^2|\mathcal{F}_{p-1}\big) = \big(\mathbb{E}b_{p1}^4 - 3\big)\sum_{k=1}^{n} p_{kk}^2(p-1) + 2\|P_{p-1}\|^2.$$

By (4.41), we then have

$$\mathrm{Var}\big(r_{pp}^2|\mathcal{F}_{p-1}\big) = O(n-p+1). \tag{4.42}$$

Let

$$t_{np} = \sqrt{-2\log\Big(1-\frac{p}{n}\Big)}\cdot s_1^{-3/2} = \sqrt{-2\log\Big(1-\frac{p}{n}\Big)}\cdot\big(\big\lfloor(-\log\varepsilon_n)^{1/4}\big\rfloor\big)^{-3/2} > 0.$$

Noting that $\mathbb{E}r_{pp}^2 = n - p + 1$, we conclude from Markov's inequality that

$$\begin{aligned}
&\mathbb{P}\big(\big|\log r_{pp}^2 - \log \mathbb{E}r_{pp}^2\big| > t_{np}\big) \\
&= \mathbb{P}\left(\left|\log \frac{r_{pp}^2}{\mathbb{E}r_{pp}^2}\right| > t_{np}\right) \\
&= \mathbb{P}\left(\frac{r_{pp}^2}{\mathbb{E}r_{pp}^2} > e^{t_{np}}\right) + \mathbb{P}\left(\frac{r_{pp}^2}{\mathbb{E}r_{pp}^2} < e^{-t_{np}}\right) \\
&= \mathbb{P}\big(r_{pp}^2 - \mathbb{E}r_{pp}^2 > \big(e^{t_{np}} - 1\big)\mathbb{E}r_{pp}^2\big) + \mathbb{P}\big(r_{pp}^2 - \mathbb{E}r_{pp}^2 < \big(e^{-t_{np}} - 1\big)\mathbb{E}r_{pp}^2\big) \\
&\leq \frac{\mathbb{E}|r_{pp}^2 - (n-p+1)|^2}{(e^{t_{np}} - 1)^2 \cdot (n-p+1)^2} + \frac{\mathbb{E}|r_{pp}^2 - (n-p+1)|^2}{(e^{-t_{np}} - 1)^2 \cdot (n-p+1)^2} \\
&\doteq H_1 + H_2.
\end{aligned} \tag{4.43}$$

In the following, we show that $H_1 = o(1)$ and $H_2 = o(1)$.

It follows from (2.5) and (2.6) that $\mathbb{E}(r_{pp}^2|\mathcal{F}_{p-1}) = n - p + 1$, which together with (4.42) yields that

$$\begin{aligned}
\mathbb{E}\big|r_{pp}^2 - (n-p+1)\big|^2 &= \mathbb{E}\big[\mathbb{E}\big(\big|r_{pp}^2 - (n-p+1)\big|^2|\mathcal{F}_{p-1}\big)\big] \\
&= \mathbb{E}\big[\mathrm{Var}\big(r_{pp}^2|\mathcal{F}_{p-1}\big)\big] \\
&\leq C(n-p+1).
\end{aligned} \tag{4.44}$$

Noting that

$$d_n \leq n - p = n\varepsilon_n, \qquad \varepsilon_n = 1 - p/n \to 0,$$

we can see that

$$C_2(-\log \varepsilon_n)^{1/8} \leq t_{np} \leq C_3\sqrt{\log n}, \tag{4.45}$$

where C_2 and C_3 are two positive constants independent of n and p. Hence, we have by (4.44), (4.45) and $e^x - 1 \geq x$ that

$$\begin{aligned}
H_1 &\doteq \frac{\mathbb{E}|r_{pp}^2 - (n-p+1)|^2}{(e^{t_{np}} - 1)^2 \cdot (n-p+1)^2} \\
&\leq \frac{C}{t_{np}^2 \cdot (n-p+1)} \\
&\leq \frac{C}{(-\log \varepsilon_n)^{1/4}} = o(1).
\end{aligned} \tag{4.46}$$

For H_2, noting that $(\frac{e^{t_{np}}}{e^{t_{np}}-1})^2$ is non-increasing for $t_{np} \geq C_2(-\log \varepsilon_n)^{1/8} > 0$, we have by (4.44) and (4.45) again that

$$\begin{aligned} H_2 &\doteq \frac{\mathbb{E}|r_{pp}^2 - (n-p+1)|^2}{(e^{-t_{np}}-1)^2 \cdot (n-p+1)^2} \\ &= \left(\frac{e^{t_{np}}}{e^{t_{np}}-1}\right)^2 \cdot \frac{\mathbb{E}|r_{pp}^2 - (n-p+1)|^2}{(n-p+1)^2} \\ &\leq \frac{C}{n-p+1} \leq \frac{C}{d_n} = o(1). \end{aligned} \tag{4.47}$$

Hence, (4.39) follows immediately from (4.43), (4.46) and (4.47). Similarly, we can also prove (4.40). From (4.38)–(4.40), we have

$$\frac{\log\det(BB^T)}{\sqrt{-2\log(1-\frac{p}{n})}} - \frac{\log\det(\bar{C}\bar{C}^T)}{\sqrt{-2\log(1-\frac{p}{n})}} = O_{\mathbb{P}}\big(s_1^{-3/2}\big). \tag{4.48}$$

Then after $s_1 = \lfloor(-\log \varepsilon_n)^{1/4}\rfloor$ steps replacing, we can get (4.6) immediately, and thus, (4.7) holds.

(ii) $1 \leq n-p < d_n$. Noticing that $|\frac{\log n}{-2\log(1-\frac{p}{n})}|$ is bounded we replace $-2\log(1-\frac{p}{n})$ by $\log n$ in the sequel. We consider the special case $n=p$ first. Since B and $\bar{C}$ are $n \times n$ matrices, it is elementary that

$$\det(B) = \sum_{k=1}^n b_{nk} B_{nk}, \qquad \det(\bar{C}) = \sum_{k=1}^n \bar{b}_{nk} B_{nk},$$

where B_{nk} is the cofactor of b_{nk} and $\bar{b}_{nk}$. Similar to (4.38), it suffices to prove that

$$\frac{\log(\sum_{k=1}^n b_{nk} B_{nk})^2}{\sqrt{\log n}} - \frac{\log(\sum_{k=1}^n \bar{b}_{nk} B_{nk})^2}{\sqrt{\log n}} = O_{\mathbb{P}}\big(s_1^{-3/2}\big),$$

i.e.

$$\begin{aligned} &\frac{\log(\sum_{k=1}^n b_{nk} B_{nk})^2 - \log(\sum_{k=1}^n B_{nk}^2)}{\sqrt{\log n}} - \frac{\log(\sum_{k=1}^n \bar{b}_{nk} B_{nk})^2 - \log(\sum_{k=1}^n B_{nk}^2)}{\sqrt{\log n}} \\ &= O_{\mathbb{P}}\big(s_1^{-3/2}\big). \end{aligned} \tag{4.49}$$

Set

$$\tilde{t}_{np} = \sqrt{\log n} \cdot s_1^{-3/2} = \sqrt{\log n} \cdot \big(\lfloor(-\log \varepsilon_n)^{1/4}\rfloor\big)^{-3/2} \to \infty,$$

and

$$\Delta = \sqrt{\sum_{k=1}^n B_{nk}^2}.$$

Similar to (4.43), we have

$$\begin{aligned}
&\mathbb{P}\left(\left|\log\left(\sum_{k=1}^{n} b_{nk}B_{nk}\right)^2 - \log \Delta^2\right| > \tilde{t}_{np}\right) \\
&= \mathbb{P}\left(\frac{|\sum_{k=1}^{n} b_{nk}B_{nk}|}{\Delta} > e^{\tilde{t}_{np}/2}\right) + \mathbb{P}\left(\frac{|\sum_{k=1}^{n} b_{nk}B_{nk}|}{\Delta} < e^{-\tilde{t}_{np}/2}\right).
\end{aligned} \tag{4.50}$$

Referring to the proof of (v) in [4], we can find that

$$\sup_x \left|\mathbb{P}\left(\frac{\sum_{k=1}^{n} b_{nk}B_{nk}}{\Delta} < x\right) - \Psi(x)\right| = o(1),$$

and

$$\sup_x \left|\mathbb{P}\left(\frac{\sum_{k=1}^{n} \bar{b}_{nk}B_{nk}}{\Delta} < x\right) - \Psi(x)\right| = o(1),$$

where $\Psi(x)$ is the cumulative distribution function of the standard normal distribution. From the above two equations and the fact that $\tilde{t}_{np} \to \infty$, we have (4.50) $= o(1)$, which implies that (4.49) holds.

For the general case when $n - p < d_n$, we use induction on p. The aim is to prove that

$$\frac{\log\det(B_{(p)}B_{(p)}^T)}{\sqrt{\log n}} - \frac{\log\det(\bar{C}_{(p)}\bar{C}_{(p)}^T)}{\sqrt{\log n}} = O_{\mathbb{P}}\big((n-p+1)s_1^{-3/2}\big). \tag{4.51}$$

Note that (4.51) holds for $p = n$ by (4.49). Now suppose that (4.51) is true for $p = p_1$ such that $0 \le n - p_1 \le d_n - 1$. We next prove that (4.51) holds for $p = p_1 - 1$. In order to apply (4.51) for $p = p_1$, we define two new matrices based on $B_{(p)}$ and $\bar{C}_{(p)}$:

$$D = \begin{pmatrix} B_{(p)} \\ \mathbf{b}_p^T \end{pmatrix}, \qquad \bar{E} = \begin{pmatrix} \bar{C}_{(p)} \\ \bar{\mathbf{b}}_p^T \end{pmatrix},$$

where $\bar{\mathbf{b}}_p^T$ is Gaussian, D and $\bar{E}$ only differ in the row next to the last one. According to induction ((4.51) holds for $p = p_1$), we have

$$\frac{\log\det(DD^T)}{\sqrt{\log n}} - \frac{\log\det(EE^T)}{\sqrt{\log n}} = O_{\mathbb{P}}\big((n-p)s_1^{-3/2}\big). \tag{4.52}$$

Similar to (4.34), we expand D and $\bar{E}$ from the last row and obtain

$$\log\det\big(DD^T\big) = \hat{r}_{pp}^2 + \log\det\big(B_{(p)}B_{(p)}^T\big), \qquad \log\det\big(EE^T\big) = \tilde{r}_{pp}^2 + \log\det\big(\bar{C}_{(p)}\bar{C}_{(p)}^T\big),$$

where

$$\hat{r}_{pp}^2 = \bar{\mathbf{b}}_p^T\big(I_n - B_{(p)}^T\big(B_{(p)}B_{(p)}^T\big)^{-1}B_{(p)}\big)\bar{\mathbf{b}}_p, \qquad \tilde{r}_{pp}^2 = \bar{\mathbf{b}}_p^T\big(I_n - \bar{C}_{(p)}^T\big(\bar{C}_{(p)}\bar{C}_{(p)}^T\big)^{-1}\bar{C}_{(p)}\big)\bar{\mathbf{b}}_p.$$

Therefore,

$$\frac{\log\det(DD^T)}{\sqrt{\log n}} - \frac{\log\det(EE^T)}{\sqrt{\log n}} = \frac{\log\det(B_{(p)}B_{(p)}^T)}{\sqrt{\log n}} - \frac{\log\det(\bar{C}_{(p)}\bar{C}_{(p)}^T)}{\sqrt{\log n}} + \frac{\log \hat{r}_{pp}^2}{\sqrt{\log n}} - \frac{\log \tilde{r}_{pp}^2}{\sqrt{\log n}}.$$

Hence, by (4.52), in order to show (4.51), it suffices to prove that

$$\frac{\log \hat{r}_{pp}^2 - \log(n-p)}{\sqrt{\log n}} - \frac{\log \tilde{r}_{pp}^2 - \log(n-p)}{\sqrt{\log n}} = O_{\mathbb{P}}\big(s_1^{-3/2}\big). \tag{4.53}$$

Since $\bar{\mathbf{b}}_p^T$ is Gaussian, $\hat{r}_{pp}^2$ and $\tilde{r}_{pp}^2 \sim \chi_{n-p}^2$. By a simple calculation, it is straightforward to get

$$\mathbb{P}\big(\big|\log \hat{r}_{pp}^2 - \log(n-p)\big| > \tilde{t}_{np}\big) + \mathbb{P}\big(\big|\log \tilde{r}_{pp}^2 - \log(n-p)\big| > \tilde{t}_{np}\big) = o(1),$$

which implies (4.53) directly. Thus, by the induction arguments, we have shown that for any $1 \le n-p \le d_n$,

$$\frac{\log\det(BB^T)}{\sqrt{-2\log(1-\frac{p}{n})}} - \frac{\log\det(\bar{C}\bar{C}^T)}{\sqrt{-2\log(1-\frac{p}{n})}} = O_{\mathbb{P}}\big(d_n s_1^{-3/2}\big) = O_{\mathbb{P}}\big(s_1^{-23/16}\big). \tag{4.54}$$

Then, after s_1 steps replacing, we can show that (4.7) holds. This completes the proof of the theorem.

Remark 4.1. To conclude this section, we present the differences between Theorem 1.2 and the corresponding ones of [13] and [4], and the novelty of the present approach.

First, the present paper establishes the logarithmic law of sample covariance matrices based on a rectangular random matrix $B = (b_{jk})_{p\times n}$, while [13] and [4] established the logarithmic law of sample covariance matrices based on a square random matrix $B = (b_{jk})_{n\times n}$.

Second, to make the log determinant manageable we use the QR decomposition to obtain a martingale decomposition of the determinant, while [13] and [4] directly made use of the Girko's method of perpendiculars.

Third, [13] and [4] estimated $p_{kk}(i)$ (see the definition above (2.4)) individually. In our setting, such an estimate (see the inequality (3.9)) is not enough when $n-p \ge n^{1-\delta}$ for any $0 < \delta < 1$. Our strategy is to treat it globally. Namely, we estimate the summation $\sum_{i=0}^{p-s_1}\sum_{k=1}^{n}\mathbb{E}p_{kk}^2(i)$ instead of $\mathbb{E}(\max_k p_{kk}(i))$. We get that $\sum_{i=0}^{p-s_1}\sum_{k=1}^{n}\mathbb{E}p_{kk}^2(i) = O(1)$ instead of $\sum_{i=0}^{p-s_1}\sum_{k=1}^{n}\mathbb{E}p_{kk}^2(i) = O(\log\log n)$ given in [13] and [4].

Finally, to prove the Gaussian replacement for the last s_1 row of B will not affect CLT, we have to distinguish two different cases according as $(n-p) \ge d_n = s_1^{1/16}$ or $(n-p) < d_n$. Moreover, we use mathematical induction to handle the case of $(n-p) < d_n$ and an appropriate expansion of the log determinant to avoid resorting to the Berry–Essen bound for quadratic forms. However, [13] and [4] used the classical Berry–Esseen bound for the sum of independent random variables.

Appendix

A.1. Proof of Proposition 2.1

Set $A_n = (b_{jk})_{n\times n}$. Note that $B = (b_{jk})_{p\times n}$ is then a submatrix of $A_n = (b_{jk})_{n\times n}$. Applying Cauchy's interlacing law, we can get that

$$s_1(A_n) \geq s_1(B), \qquad s_p(B) \geq s_n(A_n), \tag{A.1}$$

where $s_1(A_n), s_2(A_n), \dots, s_n(A_n)$ are singular values of matrix A_n such that $s_1(A_n) \geq s_2(A_n) \geq \cdots \geq s_n(A_n)$, and $s_1(B), s_2(B), \dots, s_p(B)$ are singular values of matrix B such that $s_1(B) \geq s_2(B) \geq \cdots \geq s_p(B)$.

For $s_1(A_n)$ and $s_n(A_n)$, as in [4] we have for some positive constant L and C,

$$\mathbb{P}\big(s_1(A_n) \geq n\big) \leq Cn^{-1/2}, \qquad \mathbb{P}\big(s_n(A_n) \geq n^{-L}\big) = 1 - O\left(\frac{\sqrt{\log n}}{\sqrt{n}}\right), \tag{A.2}$$

which, together with (A.1), yields that

$$\mathbb{P}\big(s_1(B) \geq n\big) \leq \mathbb{P}\big(s_1(A_n) \geq n\big) \leq Cn^{-1/2}, \tag{A.3}$$

and

$$\mathbb{P}\big(s_p(B) \geq n^{-L}\big) \geq \mathbb{P}\big(s_n(A_n) \geq n^{-L}\big) = 1 - O\left(\frac{\sqrt{\log n}}{\sqrt{n}}\right). \tag{A.4}$$

Now let θ_0 follow the uniform distribution on the interval $[-\sqrt{3}, \sqrt{3}]$ independent of B. Let θ_{jk}, $1 \leq k \leq p$, $1 \leq k \leq n$ be independent copies of θ_0. And we set $\bar{B} = (\bar{b}_{jk})_{p\times n}$, where $\bar{b}_{jk} = (1 - \varepsilon_n^2)^{1/2} b_{jk} + \varepsilon_n \theta_{jk}$. Here we choose $\varepsilon_n = n^{-(100+2L)n}$. Denote $\Theta_n = (\theta_{jk})_{p\times n}$. It follows from Weyl's inequality that

$$\big|s_i(\bar{B}) - \big(1 - \varepsilon_n^2\big)^{1/2} s_i(B)\big| \leq \varepsilon_n \|\Theta_n\|_{\mathrm{op}} \leq Cn^{-(99+2L)n}, \qquad i = 1, 2, \dots, p. \tag{A.5}$$

Combining (A.4) and (A.5), we have with probability $1 - O(\frac{\sqrt{\log n}}{\sqrt{n}})$ that

$$\begin{aligned} \det\big(\bar{B}\bar{B}^T\big) &= \prod_{i=1}^{p} s_i^2(\bar{B}) = \big(1 - \varepsilon_n^2\big)^p \big(1 + O\big(n^{-(99+L)n}\big)\big)^{2p} \prod_{i=1}^{p} s_i^2(B) \\ &= \big(1 + o(1)\big) \det\big(BB^T\big), \end{aligned} \tag{A.6}$$

which implies (2.2). Noting that θ_0 is a continuous random variable and $\bar{b}_{jk} = (1 - \varepsilon_n^2)^{1/2} b_{jk} + \varepsilon_n \theta_{jk}$, we can get (2.1) immediately. The proof is completed.

A.2. Proof of Lemma 3.5

To prove Lemma 3.5, we need the following lemma, whose proof is similar to that of Lemma 5.2 in [4]. So we omit the details.

Lemma A.1. *Let $X = (x_{ij})_{p\times n}$ be a random matrix, where $n/2 \le p \le n$ and $\{x_{ij}, 1 \le i \le p, 1 \le j \le n\}$ is a collection of real independent random variables with mean zero and variance* 1. *Moreover, we assume that* $\sup_n \max_{i,j} \mathbb{E}x_{ij}^4 < \infty$. *Denote $G_{(p)}(\alpha) = (\frac{1}{n}XX^T + \alpha I_p)^{-1}$, where $\alpha = n^{-1/6}$. Then we have*

$$\mathbb{E}\left(\frac{1}{p}\operatorname{tr} G_{(p)}(\alpha)\right) = s_p(\alpha) + O\left(n^{-1/6}\right) \tag{A.7}$$

and

$$\operatorname{Var}\left(\frac{1}{n}\operatorname{tr} G_{(p)}(\alpha)\right) = O\left(n^{-1/3}\right), \tag{A.8}$$

where

$$s_p(\alpha) = 2\left(\alpha + 1 - \frac{p}{n} + \sqrt{\left(\alpha + 1 - \frac{p}{n}\right)^2 + \frac{4\alpha p}{n}}\right)^{-1}.$$

Proof of Lemma 3.5. First, we prove (3.9). In view of (3.13), to prove (3.9), it suffices to show

$$\mathbb{E}\max_k\left(1 + \frac{1}{n}\mathbf{b}_k^T(i)G_{(i,k)}(\alpha)\mathbf{b}_k(i)\right)^{-1} \le C\log^{-8a} p \tag{A.9}$$

for $p - s_2 \le i \le p - 1$.

By the Sherman–Morrison formula again, we have

$$\operatorname{tr} G_{(i)}(\alpha) - \operatorname{tr} G_{(i,k)}(\alpha) = -\frac{\frac{1}{n}\mathbf{b}_k^T(i)G_{(i,k)}^2(\alpha)\mathbf{b}_k(i)}{1 + \frac{1}{n}\mathbf{b}_k^T(i)G_{(i,k)}(\alpha)\mathbf{b}_k(i)}, \tag{A.10}$$

and thus,

$$\operatorname{tr} G_{(i)}(\alpha) \le \operatorname{tr} G_{(i,k)}(\alpha).$$

For some small constant $0 < \varepsilon < 1/2$, for $p - s_2 \le i \le p - 1$ let

$$\chi(i) = \left(\frac{1}{n}\operatorname{tr} G_{(i)}(\alpha) \ge \log^{10a} p\right),$$

$$D_1(i) = \left(\frac{1}{n}\max_k\left|\sum_{1\le u\ne v\le i} G_{(i,k)}(u,v)b_{uk}b_{vk}\right| < \varepsilon\right),$$

$$D_2(i) = \left(\max_k\sum_{j=1}^{i} G_{(i,k)}(j,j)b_{jk}^2 \ge \log^{-a} p \cdot \operatorname{tr} G_{(i,k)}(\alpha)\right).$$

Let $G_{(i,k)}(u,v)$ denote the (u,v)th entry of $G_{(i,k)}(\alpha)$. When there is no confusion, we will omit the parameter α from the notations $G_{(i,k)}(\alpha)$ and $G_{(i)}(\alpha)$. It can be easily checked that

$$
\begin{aligned}
&\mathbb{E}\max_k\left(1+\frac{1}{n}\mathbf{b}_k^T(i)G_{(i,k)}\mathbf{b}_k(i)\right)^{-1}\\
&\le \mathbb{P}\left(\frac{1}{n}\operatorname{tr}G_{(i)}(\alpha)\le \log^{10a}p\right)\\
&\quad+\mathbb{E}I\big(\chi(i)\big)\max_k\left(1+\frac{1}{n}\mathbf{b}_k^T(i)G_{(i,k)}\mathbf{b}_k(i)\right)^{-1}\big[I\big(D_1(i)\big)+I\big(D_1^c(i)\big)\big]\\
&\le \mathbb{P}\left(\frac{1}{n}\operatorname{tr}G_{(i)}(\alpha)\le \log^{10a}p\right)+\mathbb{E}I\big(\chi(i)\big)\max_k\left(1+\frac{1}{n}\sum_{j=1}^{i}G_{(i,k)}(j,j)b_{jk}^2-\varepsilon\right)^{-1}\\
&\quad+\mathbb{P}\left(\frac{1}{n}\max_k\left|\sum_{1\le u\ne v\le i}G_{(i,k)}(u,v)b_{uk}b_{vk}\right|\ge\varepsilon\right)\\
&=\mathbb{P}\left(\frac{1}{n}\operatorname{tr}G_{(i)}(\alpha)\le \log^{10a}p\right)\\
&\quad+\mathbb{E}I\big(\chi(i)\big)\max_k\left(1+\frac{1}{n}\sum_{j=1}^{i}G_{(i,k)}(j,j)b_{jk}^2-\varepsilon\right)^{-1}\big[I\big(D_2(i)\big)+I\big(D_2^c(i)\big)\big]\\
&\quad+\mathbb{P}\left(\frac{1}{n}\max_k\left|\sum_{1\le u\ne v\le i}G_{(i,k)}(u,v)b_{uk}b_{vk}\right|\ge\varepsilon\right) \qquad\qquad (A.11)\\
&\le \mathbb{P}\left(\frac{1}{n}\operatorname{tr}G_{(i)}(\alpha)\le \log^{10a}p\right)+\mathbb{E}I\big(\chi(i)\big)\max_k\left(1+\log^{-a}p\cdot\frac{1}{n}\operatorname{tr}G_{(i,k)}-\varepsilon\right)^{-1}\\
&\quad+C\sum_{k=1}^{n}\mathbb{P}\left(\sum_{j=1}^{i}G_{(i,k)}(j,j)b_{jk}^2<\log^{-a}p\cdot\operatorname{tr}G_{(i,k)},\ \frac{1}{n}\operatorname{tr}G_{(i)}(\alpha)\ge\log^{10a}p\right)\\
&\quad+\mathbb{P}\left(\frac{1}{n}\max_k\left|\sum_{1\le u\ne v\le i}G_{(i,k)}(u,v)b_{uk}b_{vk}\right|\ge\varepsilon\right)\\
&\le \mathbb{P}\left(\frac{1}{n}\operatorname{tr}G_{(i)}(\alpha)\le \log^{10a}p\right)+\mathbb{E}I\big(\chi(i)\big)\left(1+\log^{-a}p\cdot\frac{1}{n}\operatorname{tr}G_{(i)}-\varepsilon\right)^{-1}\\
&\quad+C\sum_{k=1}^{n}\mathbb{P}\left(\sum_{j=1}^{i}G_{(i,k)}(j,j)b_{jk}^2<\log^{-a}p\cdot\operatorname{tr}G_{(i,k)},\ \frac{1}{n}\operatorname{tr}G_{(i)}(\alpha)\ge\log^{10a}p\right)\\
&\quad+\mathbb{P}\left(\frac{1}{n}\max_k\left|\sum_{1\le u\ne v\le i}G_{(i,k)}(u,v)b_{uk}b_{vk}\right|\ge\varepsilon\right)\\
&\doteq W_1+W_2+W_3+W_4.
\end{aligned}
$$

Noting that $n - p = O(n/\log^{20a} n)$ and similarly to the proof of Lemma 3.3 in [4], we have by Lemmas 3.2 and A.1 that

$$\begin{aligned} &W_1 = o\bigl(n^{-1/3}\bigr), \qquad W_2 \le C\log^{-8a} p, \\ &W_3 = o\bigl(n^{-1/3}\bigr), \qquad W_4 = O\bigl(n^{-1/3}\bigr), \end{aligned} \tag{A.12}$$

which together with (A.11) yields (A.9). Hence, (3.9) is proved.

We next prove (3.10). Similarly to the proof of (A.11), we have by (3.13) that

$$\begin{aligned} &\mathbb{P}\Bigl(\max_k p_{kk}(i) \ge \log^{-7a} p\Bigr) \\ &\quad \le \mathbb{P}\biggl(\max_k\biggl(1 + \frac{1}{n}\mathbf{b}_k^T(i)G_{(i,k)}\mathbf{b}_k(i)\biggr)^{-1} \ge \log^{-7a} p\biggr) \\ &\quad \le \mathbb{P}\biggl(\frac{1}{n}\operatorname{tr} G_{(i)}(\alpha) \le \log^{10a} p\biggr) \\ &\qquad + \mathbb{P}\biggl(\max_k\biggl(1 + \frac{1}{n}\mathbf{b}_k^T(i)G_{(i,k)}\mathbf{b}_k(i)\biggr)^{-1} \ge \log^{-7a} p, \frac{1}{n}\operatorname{tr} G_{(i)}(\alpha) \ge \log^{10a} p\biggr) \\ &\quad \le \mathbb{P}\biggl(\frac{1}{n}\operatorname{tr} G_{(i)}(\alpha) \le \log^{10a} p\biggr) \\ &\qquad + \mathbb{P}\Biggl(\max_k\Biggl(1 + \frac{1}{n}\sum_{j=1}^{i} G_{(i,k)}(j,j)b_{jk}^2 - \varepsilon\Biggr)^{-1} \ge \log^{-7a} p, \frac{1}{n}\operatorname{tr} G_{(i)}(\alpha) \ge \log^{10a} p\Biggr) \\ &\qquad + \mathbb{P}\biggl(\frac{1}{n}\max_k\biggl|\sum_{1\le u\ne v\le i} G_{(i,k)}(u,v)b_{uk}b_{vk}\biggr| \ge \varepsilon\biggr) \\ &\quad \le \mathbb{P}\biggl(\frac{1}{n}\operatorname{tr} G_{(i)}(\alpha) \le \log^{10a} p\biggr) \\ &\qquad + \mathbb{P}\biggl(\biggl(1 + \log^{-a} p \cdot \frac{1}{n}\operatorname{tr} G_{(i)} - \varepsilon\biggr)^{-1} \ge \log^{-7a} p, \frac{1}{n}\operatorname{tr} G_{(i)}(\alpha) \ge \log^{10a} p\biggr) \\ &\qquad + C\sum_{k=1}^{n}\mathbb{P}\Biggl(\sum_{j=1}^{i} G_{(i,k)}(j,j)b_{jk}^2 < \log^{-a} p \cdot \operatorname{tr} G_{(i,k)}, \frac{1}{n}\operatorname{tr} G_{(i)}(\alpha) \ge \log^{10a} p\Biggr) \\ &\qquad + \mathbb{P}\biggl(\frac{1}{n}\max_k\biggl|\sum_{1\le u\ne v\le i} G_{(i,k)}(u,v)b_{uk}b_{vk}\biggr| \ge \varepsilon\biggr) \\ &\quad \doteq W_1 + W_5 + W_3 + W_4. \end{aligned} \tag{A.13}$$

We have proved that $W_1 = o(n^{-1/3})$, $W_3 = o(n^{-1/3})$ and $W_4 = O(n^{-1/3})$ by (A.12). Moreover, we claim that

$$\begin{aligned} W_5 &\doteq \mathbb{P}\left(\left(1 + \log^{-a} p \cdot \frac{1}{n}\operatorname{tr} G_{(i)} - \varepsilon\right)^{-1} \geq \log^{-7a} p, \frac{1}{n}\operatorname{tr} G_{(i)}(\alpha) \geq \log^{10a} p\right) \\ &= 0 \end{aligned} \tag{A.14}$$

for sufficiently large p. In fact,

$$W_5 \leq \mathbb{P}\left(C \log^{-8a} p \geq \log^{-7a} p, \frac{1}{n}\operatorname{tr} G_{(i)}(\alpha) \geq \log^{10a} p\right) = 0 \tag{A.15}$$

for sufficiently large p, which yields $W_5 = 0$ for sufficiently large p. Hence, (3.10) is complete. This completes the proof of the lemma. □

A.3. Proofs of Lemmas 4.1 and 4.2

The proof of Lemma 4.1 in the paper is similar to that of Lemma 4.1 in [4]. The only difference is that "$\log n$" in [4] is replaced by "$\log \varepsilon_n^{-1}$", where $\varepsilon_n = 1 - p/n \to 0$. So we omit the details of the proof.

Next, we give the proof of Lemma 4.2. It is easily checked by (4.12) that

$$\begin{aligned} &\mathbb{P}\big(X_{i+1} < -1 + \big(\log\log \varepsilon_n^{-1}\big)^{-1}\big) \\ &\quad \leq \mathbb{P}\left(\sum_{k=1}^{n} q_{kk}(i) b_{i+1,k}^2 < 2\big(\log\log \varepsilon_n^{-1}\big)^{-1}\right) \\ &\qquad + \mathbb{P}\left(\left|\sum_{u\neq v} q_{uv}(i) b_{i+1,u} b_{i+1,v}\right| \geq \frac{1}{2}\big(\log\log \varepsilon_n^{-1}\big)^{-1}\right). \end{aligned} \tag{A.16}$$

Denote

$$\hat{b}_{jk} = b_{jk} I\big(|b_{jk}| \leq \big(\log\log \varepsilon_n^{-1}\big)^2\big), \qquad \tilde{b}_{jk} = \frac{\hat{b}_{jk} - \mathbb{E}\hat{b}_{jk}}{\sqrt{\operatorname{Var}(\hat{b}_{jk})}}.$$

It follows from the assumption $\sup_p \max_{1\leq j\leq p} \mathbb{E} b_{j1}^4 < \infty$ that

$$\mathbb{E}\hat{b}_{jk} = O\big(\big(\log\log \varepsilon_n^{-1}\big)^{-6}\big), \qquad \operatorname{Var}(\hat{b}_{jk}) = 1 + O\big(\big(\log\log \varepsilon_n^{-1}\big)^{-4}\big).$$

Consequently,

$$\tilde{b}_{jk} = \hat{b}_{jk} + O\big(\big(\log\log \varepsilon_n^{-1}\big)^{-2}\big),$$

which implies that

$$\tilde{b}_{jk}^2 \leq 2\hat{b}_{jk}^2 + O\big(\big(\log\log \varepsilon_n^{-1}\big)^{-4}\big) \leq 2\hat{b}_{jk}^2 + \big(\log\log \varepsilon_n^{-1}\big)^{-2}$$

for sufficiently large n. Therefore,

$$\begin{aligned}\mathbb{P}\Big(\sum_{k=1}^{n} q_{kk}(i)b_{i+1,k}^2 < 2\big(\log\log\varepsilon_n^{-1}\big)^{-1}\Big) &\le \mathbb{P}\Big(\sum_{k=1}^{n} q_{kk}(i)\hat{b}_{i+1,k}^2 < 2\big(\log\log\varepsilon_n^{-1}\big)^{-1}\Big)\\ &\le \mathbb{P}\Big(\sum_{k=1}^{n} q_{kk}(i)\tilde{b}_{i+1,k}^2 < C\big(\log\log\varepsilon_n^{-1}\big)^{-1}\Big)\\ &\le \mathbb{P}\Big(\Big|\sum_{k=1}^{n} q_{kk}(i)\tilde{b}_{i+1,k}^2 - 1\Big| \ge \frac{1}{2}\Big),\end{aligned}$$

which together with (3.3) yields that

$$\begin{aligned}&\mathbb{P}\Big(\sum_{k=1}^{n} q_{kk}(i)b_{i+1,k}^2 < 2\big(\log\log\varepsilon_n^{-1}\big)^{-1}\Big)\\ &\quad\le C\Big[\big(\log\log\varepsilon_n^{-1}\big)^8\cdot \operatorname{tr} Q_i^4 + \big(\operatorname{tr} Q_i^2\big)^2\Big] \qquad \text{(A.17)}\\ &\quad= C\Big[\big(\log\log\varepsilon_n^{-1}\big)^8\cdot (n-i)^{-3} + (n-i)^{-2}\Big].\end{aligned}$$

Moreover, it follows from (3.4) that

$$\mathbb{P}\Big(\Big|\sum_{u\neq v} q_{uv}(i)b_{i+1,u}b_{i+1,v}\Big| \ge \frac{1}{2}\big(\log\log\varepsilon_n^{-1}\big)^{-1}\Big) \le C(n-i)^{-2}\cdot\big(\log\log\varepsilon_n^{-1}\big)^4. \quad \text{(A.18)}$$

Noting that $s_1 = \lfloor(-\log\varepsilon_n)^{1/4}\rfloor$, we have by (A.16)–(A.18) that

$$\begin{aligned}&\sum_{i=0}^{p-s_1}\mathbb{P}\big(X_{i+1} < -1 + \big(\log\log\varepsilon_n^{-1}\big)^{-1}\big)\\ &\quad\le C\Big[\big(\log\log\varepsilon_n^{-1}\big)^8\cdot\sum_{i=0}^{p-s_1}\frac{1}{(n-i)^3} + \big(\log\log\varepsilon_n^{-1}\big)^4\cdot\sum_{i=0}^{p-s_1}\frac{1}{(n-i)^2}\Big]\\ &\quad\le C\Big[\frac{(\log\log\varepsilon_n^{-1})^8}{(\log\varepsilon_n^{-1})^{1/2}} + \frac{(\log\log\varepsilon_n^{-1})^4}{(\log\varepsilon_n^{-1})^{1/4}}\Big]\\ &\quad\to 0 \qquad \text{as } n\to\infty.\end{aligned}$$

This completes the proof of Lemma 3.6.

Acknowledgements

Xuejun Wang was partially supported by the Natural Science Foundation of Anhui Province (1508085J06) and Natural Science Foundation of China (11671012). Guangming Pan (corresponding author) were partially supported by Singapore Ministry of Eduction (MOE) AcRF Tier 2 grants of # ARC 14/11 and of 2014-T2-2-060 and by a MOE Tier 1 Grant RG25/14 at the Nanyang Technological University, Singapore.

References

[1] Anderson, T.W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Hoboken, NJ: Wiley. MR1990662
[2] Bai, Z.D. and Silverstein, J.W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** 553–605. MR2040792
[3] Bai, Z. and Silverstein, J.W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. New York: Springer. MR2567175
[4] Bao, Z., Pan, G. and Zhou, W. (2015). The logarithmic law of random determinant. *Bernoulli* **21** 1600–1628. MR3352055
[5] Cai, T.T., Liang, T. and Zhou, H.H. (2015). Law of log determinant of sample covariance matrix and optimal estimation of differential entropy for high-dimensional Gaussian distributions. *J. Multivariate Anal.* **137** 161–172. MR3332804
[6] Costello, K.P. and Vu, V. (2009). Concentration of random determinants and permanent estimators. *SIAM J. Discrete Math.* **23** 1356–1371. MR2556534
[7] Gīrko, V.L. (1980). The central limit theorem for random determinants. *Theory Probab. Appl.* **24** 729–740.
[8] Girko, V.L. (1990). *Theory of Random Determinants*. Dordrecht: Kluwer Academic. MR1080966
[9] Girko, V.L. (1997). A refinement of the central limit theorem for random determinants. *Theory Probab. Appl.* **42** 121–129.
[10] Goodman, N.R. (1963). The distribution of the determinant of a complex Wishart distributed matrix. *Ann. Math. Stat.* **34** 178–180. MR0145619
[11] Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*. New York: Academic Press. MR0624435
[12] Komlós, J. (1968). On the determinant of random matrices. *Studia Sci. Math. Hungar.* **3** 387–399. MR0238371
[13] Nguyen, H.H. and Vu, V. (2014). Random matrices: Law of the determinant. *Ann. Probab.* **42** 146–167. MR3161483
[14] Nielsen, J. (1999). The distribution of volume reductions induced by isotropic random projections. *Adv. in Appl. Probab.* **31** 985–994. MR1747452
[15] Rempała, G. and Wesołowski, J. (2005). Asymptotics for products of independent sums with an application to Wishart determinants. *Statist. Probab. Lett.* **74** 129–138. MR2169371
[16] Rouault, A. (2007). Asymptotic behavior of random determinants in the Laguerre, Gram and Jacobi ensembles. *ALEA Lat. Am. J. Probab. Math. Stat.* **3** 181–230. MR2365642
[17] Tao, T. and Vu, V. (2006). On random ± 1 matrices: Singularity and determinant. *Random Structures Algorithms* **28** 1–23. MR2187480
[18] Tao, T. and Vu, V. (2008). Random matrices: The circular law. *Commun. Contemp. Math.* **10** 261–307. MR2409368
[19] Tao, T. and Vu, V. (2012). A central limit theorem for the determinant of a Wigner matrix. *Adv. Math.* **231** 74–101. MR2935384

Received September 2015 and revised February 2016