Submitted

ALMOST SURE LIMIT OF THE SMALLEST EIGENVALUE OF THE SAMPLE CORRELATION MATRIX

BY HAN XIAO AND WANG ZHOU

National University of Singapore

Let $X^{(n)} = (X_{ij})$ be a $p \times n$ data matrix, where the n columns form a random sample of size n from a certain p-dimensional distribution. Let $R^{(n)} = (\rho_{ij})$ be the $p \times p$ sample correlation coefficient matrix of $X^{(n)}$; and $S^{(n)} = (1/n)X^{(n)}\left(X^{(n)}\right)^* - \bar{X}\bar{X}^*$ be the sample covariance matrix of $X^{(n)}$, where $\bar{X}$ is the mean vector of the n observations. Assuming that X_{ij}'s are independent and identically distributed with finite fourth moment, we show that the smallest eigenvalue of $R^{(n)}$ converges almost surely to the limit $(1-\sqrt{c}\,)^2$ as $n \to \infty$ and $p/n \to c \in (0\,,\,\infty)$. We accomplish this by showing that the smallest eigenvalue of $S^{(n)}$ converges almost surely to $(1-\sqrt{c}\,)^2$.

1. Introduction. Suppose $X^{(n)} = (X_{ij})$ is a $p \times n$ data matrix, where the n columns form a random sample of size n from a certain p-dimensional distribution. Let $R^{(n)} = (\rho_{ij})$ be the $p \times p$ sample correlation coefficient matrix of $X^{(n)}$, where ρ_{ij} is the usual Pearson correlation coefficient between the i-th row and the j-th row of $X^{(n)}$. We are interested in the strong limits of the extreme eigenvalues of this matrix as its dimensions tend to infinity.

There are two random matrices which are closely related with the sample correlation matrix. One is the sample covariance matrix $S^{(n)}$ defined by

$$S^{(n)} = (S_{ij}^{(n)}) = \frac{1}{n}X^{(n)}\left(X^{(n)}\right)^* - \bar{X}\bar{X}^*,$$

where $\bar{X}$ is the mean vector of the n observations. Let

$$D^{(n)} = diag\left\{\sqrt{S_{11}^{(n)}}, \sqrt{S_{22}^{(n)}}, \dots, \sqrt{S_{pp}^{(n)}}\right\},$$

then $R^{(n)}$ could be expressed as

$$R^{(n)} = \left(D^{(n)}\right)^{-1} S^{(n)} \left(D^{(n)}\right)^{-1}. \tag{1.1}$$

AMS 2000 subject classifications: Primary 60H15; secondary 62H99

Keywords and phrases: Random matrix, sample correlation coefficient matrix, sample covariance matrix, smallest eigenvalue

The other one is a simplified version of the sample covariance matrix given by

$$\mathbb{S}^{(n)} = (1/n)X^{(n)}\left(X^{(n)}\right)^{*}.$$

REMARKS. (1) For notational economy, we will omit the super-index (n) from now on when there is no confusion. (2) In the literature, $\mathbb{S}$ is often referred under the name "sample covariance matrix". However, in this paper, we rename it by *simplified sample covariance matrix* to avoid confusion.

Suppose $\lambda_1(\mathbb{S}), \lambda_2(\mathbb{S}), \ldots, \lambda_p(\mathbb{S})$ are the p eigenvalues of $\mathbb{S}$ in increasing order. While the definition of the largest eigenvalue is clear, one needs to examine that of the smallest one.

Since $\operatorname{rank}(\mathbb{S}) \le n$ when $p \ge n$, the $(p-n)$ smallest eigenvalues are all zero. Hence we define the smallest eigenvalue of the matrix $\mathbb{S}$ as

$$\lambda_{\min}(\mathbb{S}) = \begin{cases} \lambda_1(\mathbb{S}), & \text{if } p < n; \\ \lambda_{p-n+1}(\mathbb{S}), & \text{if } p \ge n. \end{cases} \tag{1.2}$$

It is not hard to see that if the *empirical spectral distribution* $F^{\mathbb{S}}$ of $\mathbb{S}$ almost surely converges to the *Marčenko-Pastur law* F_c with the density

$$F_c'(x) = \begin{cases} \frac{1}{2\pi c x}\sqrt{(b-x)(x-a)}, & \text{if } a \le x \le b; \\ 0, & \text{otherwise}; \end{cases} \tag{1.3}$$

and a point mass $1 - 1/c$ at the origin if $c > 1$, where $a = (1-\sqrt{c}\,)^2$ and $b = (1+\sqrt{c}\,)^2$, then

$$\liminf \lambda_{\max}(\mathbb{S}) \ge b = (1+\sqrt{c}\,)^2 \quad \text{a.s.}$$
$$\limsup \lambda_{\min}(\mathbb{S}) \le a = (1-\sqrt{c}\,)^2 \quad \text{a.s.}$$

However, the following converse assertions

$$\limsup \lambda_{\max}(\mathbb{S}) \le b = (1+\sqrt{c}\,)^2 \quad \text{a.s.} \tag{1.4}$$
$$\liminf \lambda_{\min}(\mathbb{S}) \ge a = (1-\sqrt{c}\,)^2 \quad \text{a.s.} \tag{1.5}$$

are not trivial.

Yin, Bai and Krishnaiah (1988) established (1.4). The following modified version is an immediate consequence of their original result.

THEOREM 1.1. *Let X be the up-left $p \times n$ corner of a double array $\{X_{uv} : u, v = 1, 2, \ldots\}$ of independent and identically distributed (i.i.d.) complex random variables (r.v.s) with zero mean and unit variance. If $E|X_{11}|^4 < \infty$, then as $n \to \infty$ and $p/n \to c \in (0\,, \infty)$,*

$$\lim \lambda_{max}(\mathbb{S}) = b = (1+\sqrt{c}\,)^2 \quad \text{a.s.}$$

It is much more difficult to establish (1.5) than (1.4). Bai and Yin (1993) devised a unified approach to prove (1.4) and (1.5) at the same time. As an immediate consequence of their result, we have the following theorem.

THEOREM 1.2. *Under the same conditions of Theorem 1.1,*

$$\lim \lambda_{min}(\mathbb{S}) = a = (1 - \sqrt{c}\,)^2 \quad \text{a.s.}$$

More than ten years later, Jiang (2004) proved that the largest eigenvalue of the sample correlation matrix R converges to the limit $(1 + \sqrt{c}\,)^2$ with probability one as $n \to \infty$ and $p/n \to c \in (0\,,\,\infty)$. Jiang (2004) also conjectured that the smallest eigenvalue of R converges to $(1 - \sqrt{c}\,)^2$ a.s..

Since $\operatorname{rank}(R) \leq n - 1$, as in (1.2), the smallest eigenvalue of the matrix R can be defined as

$$\lambda_{\min}(R) = \begin{cases} \lambda_1(R), & \text{if } p < n; \\ \lambda_{p-n+2}(R), & \text{if } p \geq n. \end{cases} \tag{1.6}$$

In this paper, we prove Jiang's conjecture.

THEOREM 1.3. *Let X be the up-left $p \times n$ corner of a double array $\{X_{uv} : u, v = 1, 2, \ldots\}$ of i.i.d. complex r.v.s with unit variance. If $E|X_{11}|^4 < \infty$, then as $n \to \infty$ and $p/n \to c \in (0\,,\,\infty)$,*

$$\lim \lambda_{min}(R) = a = (1 - \sqrt{c}\,)^2 \quad \text{a.s.}$$

We accomplish the proof of Theorem 1.3 by establishing the following result on the sample covariance matrix S. Note that the definition of the smallest eigenvalue of S is given by replacing R in (1.6) by S.

THEOREM 1.4. *Under the same conditions of Theorem 1.3,*

$$\lim \lambda_{min}(S) = a = (1 - \sqrt{c}\,)^2 \quad \text{a.s.}$$

The paper is organized as follows. In section 2, we show how Theorem 1.4 implies Theorem 1.3. The proof of Theorem 1.4 will be completed in Section 3. The auxiliary lemmas are collected in the last section.

2. From Sample Covariance Matrix to Sample Correlation Matrix. Our task is this section is to prove Theorem 1.3 by Theorem 1.4. Actually the argument here parallels that in Jiang (2004). We repeat it for the completeness of the whole proof.

Since we are interested in sample covariance matrix and sample correlation matrix, we can assume that $EX_{11} = 0$. According Theorem 1.4, it suffices to show that

$$\sqrt{\lambda_{\min}(R)} - \sqrt{\lambda_{\min}(S)} \to 0 \quad \text{a.s.} \tag{2.1}$$

Note that the sample covariance matrix S could be written as $S = (1/n)XPX$, where P is the $n \times n$ projection matrix defined as $I - \frac{1}{n}\mathbb{1}\mathbb{1}^T$, and $\mathbb{1}$ is the $n \times 1$ vector whose entries are all 1's. Since $R = D^{-1}SD^{-1}$ (see (1.1)), by Lemma 4.1

$$\begin{aligned}\left|\sqrt{\lambda_{\min}(R)} - \sqrt{\lambda_{\min}(S)}\right| &\leq \|D^{-1}\frac{1}{\sqrt{n}}XP - \frac{1}{\sqrt{n}}XP\| \\ &\leq \|D^{-1} - I\| \cdot \|\frac{1}{\sqrt{n}}X\|,\end{aligned} \tag{2.2}$$

Since $E|X_{11}|^4 < \infty$, due to Lemma 4.4, we know that

$$\max_{1\leq j\leq p}\left|\frac{\sum_{i=1}^n X_{ij}^2}{n} - 1\right| \to 0 \quad \text{a.s.}, \tag{2.3}$$

and

$$\max_{1\leq j\leq p}\bar{X}_j \to 0 \quad \text{a.s.} \tag{2.4}$$

where $\bar{X}_j$ is the j-th entry of the mean vector $\bar{X}$. Combine (2.3) and (2.4), we have

$$\max_{1\leq j\leq p}\left|\frac{\sum_{i=1}^n (X_{ij} - \bar{X}_j)^2}{n} - 1\right| \leq \max_{1\leq j\leq p}\left|\frac{\sum_{i=1}^n X_{ij}^2}{n} - 1\right| + \max_{1\leq j\leq p}\bar{X}_j^2 \to 0 \quad \text{a.s.},$$

and this implies that

$$\|D^{-1} - I\| = \max_{1\leq j\leq p}\left|\frac{\sqrt{n}}{\sqrt{\sum_{i=1}^n (X_{ij} - \bar{X}_j)^2}} - 1\right| \to 0 \quad \text{a.s.} \tag{2.5}$$

By Theorem 1.1,

$$\|\frac{1}{\sqrt{n}}X\| = \sqrt{\lambda_{\max}(\mathbb{S})} \to 1 + \sqrt{c} \quad \text{a.s.}$$

This together with (2.2) and (2.5) proves (2.1).

3. Proof of Theorem 1.4. We first derive the limiting spectral distribution of S. Since $S = \mathbb{S} - \bar{X}\bar{X}^*$, by Lemma 4.2, we know that

$$\|F^S - F^{\mathbb{S}}\| \le \frac{1}{p}\, rank(\bar{X}\bar{X}^*) = \frac{1}{p}.$$

Since convergence in the sup norm implies the weak convergence of the distribution functions, we know F^S also converges to the Marčenko-Pastur law, and hence

$$\limsup \lambda_{\min}(S) \le a = (1 - \sqrt{c}\,)^2 \quad \text{a.s.}$$

Therefore, in order to prove Theorem 1.4, it suffices to show that

$$\liminf \lambda_{\min}(S) \ge a = (1 - \sqrt{c})^2 \quad \text{a.s.} \tag{3.1}$$

Note that when $c = 1$, the situation is trivial. When $c > 1$, p is larger than n when n is very large. In this case we will consider $\lambda_{\min}(S) = \lambda_{p-n+2}(S)$, which is the $(p - n + 2)$-th smallest eigenvalue of S. According to Corollary 4.3.3 of Horn and Johnson (1985), we have

$$\lambda_{p-n+2}(\Sigma) \ge \lambda_{p-n+1}(S).$$

As an immediate consequence of this fact and Theorem 1.2, we know that (3.1) holds when $c > 1$. Now we shall prove (3.1) when $0 < c < 1$, and the long proof will be divided into several steps.

3.1. *Truncation.* For $C > 0$, let $Y_{ij} = X_{ij} I_{\{|X_{ij}| \le C\}} - EX_{ij} I_{\{|X_{ij}| \le C\}}$, $Y = (Y_{ij})$ and $\tilde{S} = (1/n)YPY^*$. Denote the eigenvalues of S and $\tilde{S}$ by λ_k and $\tilde{\lambda}_k$ (in increasing order). Since these are the squares of the k-th smallest singular values of $(1\sqrt{n})XP$ and $(1\sqrt{n})YP$ (respectively); it follows from Lemma 4.1 that

$$\max_{k \le n} |\lambda_k^{1/2} - \tilde{\lambda}_k^{1/2}| \le \frac{1}{\sqrt{n}}\|X - Y\|$$

Since $X_{ij} - Y_{ij} = X_{ij} I_{\{|X_{ij}| > C\}} - EX_{ij} I_{\{|X_{ij}| > C\}}$, from Theorem 1.1, we have, with probability one,

$$\limsup_{n\to\infty} \max_{k \le n} |\lambda_k^{1/2} - \tilde{\lambda}_k^{1/2}| \le (1 + \sqrt{c}) E^{1/2} |X_{11}|^2 I_{\{|X_{ij}| > C\}}$$

Since $E|X_{11}|^2 = 1$, we can make the above bound arbitrarily small by choosing C sufficiently large. Thus, in the following investigation, we can assume that the underlying variables are uniformly bounded. It is easy to verify that we can rescale the variables such that the assumption $E|X_{11}|^2 = 1$ still holds.

3.2. *An Equivalent Problem.* Suppose that the smallest eigenvalue of S is smaller than $a = (1-\sqrt{c})^2$, than there exists a non-zero vector β, such that:

$$S\beta = \left(\mathbb{S} - \bar{X}\bar{X}^*\right)\beta = \lambda_{\min}(S)\beta.$$

which is equivalent to

$$\left(\mathbb{S} - \lambda_{\min}(S)I\right)\beta = \bar{X}\bar{X}^*\beta \tag{3.2}$$

If the smallest eigenvalue of $\mathbb{S}$ is not smaller than a, then we have that $\bar{X}^*\beta \neq 0$ and the matrix $(\mathbb{S} - \lambda_{\min}(S)I)$ is nonsingular. In this case (3.2) can be inverted to give

$$\beta = \left(\mathbb{S} - \lambda_{\min}(S)I\right)^{-1}\bar{X}\bar{X}^*\beta.$$

If we multiply both sides of the above equation by $\bar{X}^*$, we will get

$$\bar{X}^*\beta = \bar{X}^*\left(\mathbb{S} - \lambda_{\min}(S)I\right)^{-1}\bar{X}\bar{X}^*\beta.$$

Since $\bar{X}^*\beta \neq 0$, we can obtain that

$$\bar{X}^*\left(\mathbb{S} - \lambda_{\min}(S)I\right)^{-1}\bar{X} = 1. \tag{3.3}$$

The above arguments (especially (3.3)) provide the basic idea that we will make use of to state the current problem in an equivalent form which is given by the following lemma.

Lemma 3.1. *If $P\left(\liminf \lambda_{min}(S) < a\right) > 0$, then for some $0 < \lambda < a$,*

$$P\left(\limsup \bar{X}^*\left(\mathbb{S} - \lambda I\right)^{-1}\bar{X} \geq 1\right) > 0.$$

In other words, if

$$\limsup \bar{X}^*\left(\mathbb{S} - \lambda I\right)^{-1}\bar{X} < 1 \quad a.s. \quad \forall\, 0 < \lambda < a \tag{3.4}$$

then we will have the desirable property

$$\liminf \lambda_{min}(S) \geq a \quad a.s.$$

Proof: If $P\left(\liminf \lambda_{\min}(S) < a\right) > 0$, then there exists a small $\epsilon > 0$, such that $P\left(\liminf \lambda_{\min}(S) < a - 3\epsilon\right) > 0$. For simplicity, we will denote the event $\{\liminf \lambda_{\min}(S) < a - 3\epsilon\}$ by E_0. Let B_n denote the event $\{\lambda_{\min}(\mathbb{S}) \leq a - \epsilon\}$, from Lemma 4.5, we know that $P(B_n) = o(n^{-l})$ for any $l > 0$. Hence it is easy to see that for some N large enough, $P\left(E_0 \setminus \cup_{n=N}^{\infty} B_n\right) > 0$. We use E to denote the event $E_0 \setminus \cup_{n=N}^{\infty} B_n$. For each $\omega \in E$, the following two properties hold:

$$\liminf \lambda_{\min}(S(\omega)) < a - 3\epsilon, \quad \lambda_{\min}(\mathbb{S}(\omega)) > a - \epsilon, \forall\, n \geq N.$$

Since $\liminf \lambda_{\min}(S(\omega)) < a - 3\epsilon$, we can find a subsequence n_k, such that

$$\liminf \lambda_{\min}^{(n_k)}(S(\omega)) \to \lambda(\omega) < a - 3\epsilon.$$

When k is large enough, $\lambda_{\min}^{(n_k)}(S(\omega)) < a - 2\epsilon$, and hence we have from (3.3)

$$\sup_{0<\lambda<a-2\epsilon} \left\{ \left[\bar{X}^* \left(\mathbb{S}^{(n_k)} - \lambda I \right)^{-1} \bar{X} \right] \Big|_{\omega} \right\} \geq 1.$$

Note that $\bar{X}^* \left(\mathbb{S} - \lambda I \right)^{-1} \bar{X}$ is an increasing function of λ, we have

$$\left[\bar{X}^* \left(\mathbb{S}^{(n_k)} - (a - 2\epsilon) I \right)^{-1} \bar{X} \right] \Big|_{\omega} \geq 1,$$

and this means that

$$\limsup \left[\bar{X}^* \left(\mathbb{S}^{(n)} - (a - 2\epsilon) I \right)^{-1} \bar{X} \right] \Big|_{\omega} \geq 1.$$

Therefore, we know that for $\lambda = a - 2\epsilon$,

$$P \left(\limsup \bar{X}^* \left(\mathbb{S} - \lambda I \right)^{-1} \bar{X} \geq 1 \right) \geq P(E) > 0,$$

which completes the proof. ■

Now our target is to prove (3.4) when $0 < c < 1$. Suppose $0 < \lambda < a$, let $2\epsilon = a - \lambda$. We expand $\bar{X}^* \left(\mathbb{S} - \lambda I \right)^{-1} \bar{X}$ as

$$\begin{aligned} \bar{X}^* \left(\mathbb{S} - \lambda I \right)^{-1} \bar{X} &= \frac{1}{n} \left(X_1 + \cdots + X_n \right)^* \left(\mathbb{S} - \lambda I \right)^{-1} \frac{1}{n} \left(X_1 + \cdots + X_n \right) \\ &= \frac{1}{n^2} \sum_{i=1}^{n} X_i^* \left(\mathbb{S} - \lambda I \right)^{-1} X_i + \frac{1}{n^2} \sum_{i \neq j} X_i^* \left(\mathbb{S} - \lambda I \right)^{-1} X_j \end{aligned}$$

Let

$$T_1 = \frac{1}{n^2} \sum_{i=1}^{n} X_i^* \left(\mathbb{S} - \lambda I \right)^{-1} X_i \tag{3.5}$$

$$T_2 = \frac{1}{n^2} \sum_{i \neq j} X_i^* \left(\mathbb{S} - \lambda I \right)^{-1} X_j, \tag{3.6}$$

we will consider T_1 and T_2 respectively.

3.3. *Non-negative Terms.* Let $S_i = \mathbb{S} - (1/n)X_iX_i^*$. Using Lemma 4.3, we may write T_1 as

$$\begin{aligned} T_1 &= \frac{1}{n^2}\sum_{i=1}^{n} \frac{X_i^*\left(S_i - \lambda I\right)^{-1} X_i}{1 + \frac{1}{n} X_i^*\left(S_i - \lambda I\right)^{-1} X_i} \\ &= \frac{1}{n}\sum_{i=1}^{n} \frac{\frac{1}{n}X_i^*\left(S_i - \lambda I\right)^{-1} X_i}{1 + \frac{1}{n} X_i^*\left(S_i - \lambda I\right)^{-1} X_i}. \end{aligned}$$

We use E_i to denote the event $\{\lambda_{\min}(S_i) > a - \epsilon\}$, and let $E = \bigcap_{i=1}^n E_i$. Again from Lemma 4.5, we know that $P(E_i^c) = o(n^{-l})$ for any $l > 0$, and hence $P(E^c) = o(n^{-l})$ for any $l > 0$. On the event E, $\left\|\left(S_i - \lambda I\right)^{-1}\right\| \le 1/\epsilon$, and the matrix $\left(S_i - \lambda I\right)^{-1}$ is non-negative definite. Since all the entries of X are bounded by a constant C, we have

$$\left|\frac{1}{n}X_i^*\left(S_i - \lambda I\right)^{-1} X_i\right| \le \frac{1}{n}\left\|X_i^*\right\| \left\|\left(S_i - \lambda I\right)^{-1}\right\| \left\|X_i\right\| \le \frac{C}{\epsilon}.$$

Therefore we know on the event E,

$$T_1 \le \frac{\frac{C}{\epsilon}}{1 + \frac{C}{\epsilon}} < 1.$$

Since $P(E^c) = o(n^{-l})$ for any $l > 0$, by Borel-Cantelli Lemma we know that

$$\limsup T_1 \le \frac{\frac{C}{\epsilon}}{1 + \frac{C}{\epsilon}} < 1 \quad \text{a.s.} \tag{3.7}$$

3.4. *Crossed Terms.* Now we will focus on T_2, and it suffices to show that:

$$\lim T_2 = 0 \quad \text{a.s.} \tag{3.8}$$

Let $S_{ij} = \mathbb{S} - (1/n)X_iX_i^* - (1/n)X_jX_j^*$. By Lemma 4.3, we can write T_2 as

$$T_2 = \frac{1}{n^2}\sum_{i \ne j} \frac{X_i^*\left(S_{ij} - \lambda I\right)^{-1} X_j}{\left[1 + \frac{1}{n}X_i^*\left(S_i - \lambda I\right)^{-1} X_i\right]\left[1 + \frac{1}{n}X_j^*\left(S_{ij} - \lambda I\right)^{-1} X_j\right]}.$$

This expression plays the central role in our investigation.

In the previous parts, we have defined the matrix S_i and S_{ij}. Similarly, we can define such kind of matrix with more sub-indices, such as $S_{i_1 i_2 j_1}$, $S_{i_1 i_2 j_1 j_2}$, *etc.* In general, let $\Lambda \subset \mathbb{N}$ be an finite index set, S_Λ is defined as

$$S_\Lambda = \mathbb{S} - \frac{1}{n}\sum_{i \in \Lambda} X_iX_i^*.$$

For simplicity we use the notation A_Λ to denote the following matrix

$$A_\Lambda = (S_\Lambda - \lambda I)^{-1}.$$

3.4.1. *Change S_i to S_{ij} in the Denominator.* Motivated by the symmetry, we first change S_i in the denominator of T_2 to S_{ij}, and denote the new term by T_3

$$\begin{aligned} T_3 &= \frac{1}{n^2} \sum_{i \neq j} \frac{X_i^* (S_{ij} - \lambda I)^{-1} X_j}{\left[1 + \frac{1}{n} X_i^* (S_{ij} - \lambda I)^{-1} X_i\right] \left[1 + \frac{1}{n} X_j^* (S_{ij} - \lambda I)^{-1} X_j\right]} \\ &= \frac{1}{n^2} \sum_{i \neq j} \frac{X_i^* A_{ij} X_j}{\left(1 + \frac{1}{n} X_i^* A_{ij} X_i\right) \left(1 + \frac{1}{n} X_j^* A_{ij} X_j\right)}. \end{aligned}$$

Our task in this step is to show that

$$D_{23} = T_2 - T_3 \to 0 \quad \text{a.s.} \tag{3.9}$$

According to Lemma 4.3, we can write D_{23} as

$$\begin{aligned} D_{23} &= \frac{1}{n^2} \sum_{i \neq j} \frac{X_i^* A_{ij} X_j \left(\frac{1}{n} X_i^* A_{ij} X_i - \frac{1}{n} X_i^* A_i X_i\right)}{\left(1 + \frac{1}{n} X_i^* A_i X_i\right) \left(1 + \frac{1}{n} X_i^* A_{ij} X_i\right) \left(1 + \frac{1}{n} X_j^* A_{ij} X_j\right)} \\ &= \frac{1}{n^4} \sum_{i \neq j} \frac{X_i^* A_{ij} X_j X_i^* A_{ij} X_j X_j^* A_{ij} X_i}{\left(1 + \frac{1}{n} X_i^* A_i X_i\right) \left(1 + \frac{1}{n} X_i^* A_{ij} X_i\right) \left(1 + \frac{1}{n} X_j^* A_{ij} X_j\right)^2} \end{aligned}$$

In order to control the norm of the matrix A_{ij}, we will confine it on the event $E_{ij} = \{\lambda_{\min}(S_{ij}) > a - \epsilon\}$; and we consider

$$\bar{D}_{23} = \frac{1}{n^4} \sum_{i \neq j} \frac{X_i^* A_{ij} X_j X_i^* A_{ij} X_j X_j^* A_{ij} X_i}{\left(1 + \frac{1}{n} X_i^* A_i X_i\right) \left(1 + \frac{1}{n} X_i^* A_{ij} X_i\right) \left(1 + \frac{1}{n} X_j^* A_{ij} X_j\right)^2} I_{ij}.$$

By Borel-Cantelli Lemma and Lemma 4.5, it is not difficult to see that the difference between D_{23} and $\bar{D}_{23}$ tends to zero with probability one. Hence it suffices to consider $\bar{D}_{23}$ in the following. On the event E_{ij}, A_i and A_{ij} are positive definite, and hence

$$\left(1 + \frac{1}{n} X_i^* A_i X_i\right) \left(1 + \frac{1}{n} X_i^* A_{ij} X_i\right) \left(1 + \frac{1}{n} X_j^* A_{ij} X_j\right)^2 \geq 1;$$

so in order to prove (3.9), it is enough to show that

$$D'_{23} = \frac{1}{n^4} \sum_{i \neq j} |X_i^* A_{ij} X_j X_i^* A_{ij} X_j X_j^* A_{ij} X_i| \to 0 \quad \text{a.s.} \tag{3.10}$$

Since on the event E_{ij}, the norm of A_{ij} is bounded by $1/\epsilon$; due to Lemma 4.6, we know that

$$\begin{aligned} E|X_i^* A_{ij} I_{ij} X_j|^r &= E\left[E\left(|X_i^* A_{ij} I_{ij} X_j|^r \,|\, A_{ij}\right)\right] \\ &\leq K_r n^{r/2} \quad \text{for any } r \geq 2; \end{aligned}$$

and similarly

$$E|X_j^* A_{ij} I_{ij} X_i|^r \leq K_r n^{r/2} \quad \text{for any } r \geq 2;$$

where K_r is a constant only depending on r. Making use of these orders, together with Hölder inequality, we can compute the third moment of D'_{23}, and the result is given by

$$E|D'_{23}|^3 = O(n^{-3/2});$$

Therefore, (3.10) follows the Borel-Cantelli lemma.

Remark. Note that when we transfer form D_{23} to $\bar{D}_{23}$, what we do is to confine the matrix A_{ij} on the event $E_{ij} = \{\lambda_{\min}(S_{ij}) > a - \epsilon\}$ so that its normal could be controlled by $1/\epsilon$. We should further note that $A_{ij} I_{ij}$ is still independent with X_i and X_j. In general, we could confine the matrix A_Λ on the event $E_\Lambda = \{\lambda_{\min}(S_\Lambda) > a - \epsilon\}$ to control its norm. In the subsequent investigation, we should use this kind of confinement again and again. Fortunately, due to Lemma 4.5 and Borel-Cantelli Lemma, none of these confinements will change the strong limit under condiseration; and a straightforward but tedious argument could furnish the justification if necessary. To circumvent such tedium, we can conveniently assume that:

Assumption (i). A_Λ is non-negative definite and $\|A_\Lambda\| \leq (1/\epsilon)$ for any finite index set Λ;
Assumption (ii). A_Λ and $\{X_i,\, i \in \Lambda\}$ are independent.

3.4.2. *Remove X_i and X_j in the Denominator.* We first show that in the denominator $X_i^* A_{ij} X_i$ can be replaced by trA_{ij}. Let

$$T_4 = \frac{1}{n^2} \sum_{i \neq j} \frac{X_i^* A_{ij} X_j}{\left(1 + \frac{1}{n} trA_{ij}\right)\left(1 + \frac{1}{n} X_j^* A_{ij} X_j\right)},$$

our task is to show that:

$$D_{43} = T_4 - T_3 \to 0 \quad \text{a.s.}$$

We write D_{43} as

$$D_{43} = \frac{1}{n^2} \sum_{i \neq j} \frac{X_i^* A_{ij} X_j \left(\frac{1}{n} X_i^* A_{ij} X_i - \frac{1}{n} tr A_{ij} \right)}{\left(1 + \frac{1}{n} tr A_{ij}\right) \left(1 + \frac{1}{n} X_i^* A_{ij} X_i\right) \left(1 + \frac{1}{n} X_j^* A_{ij} X_j\right)}.$$

It is convenient to consider the following

$$\bar{D}_{43} = \frac{1}{n^2} \sum_{i \neq j} \frac{X_i^* A_{ij} X_j \left(\frac{1}{n} X_i^* A_{ij} X_i - \frac{1}{n} tr A_{ij} \right)}{\left(1 + \frac{1}{n} tr A_{ij}\right) \left(1 + \frac{1}{n} tr A_{ij}\right) \left(1 + \frac{1}{n} X_j^* A_{ij} X_j\right)}.$$

instead of D_{43}. The reason is that by computation (again due to Lemma 4.6 and Hölder inequality), we can find that:

$$E|D_{43} - \bar{D}_{43}|^3 = O(n^{-3/2})$$

and hence $D_{43} - \bar{D}_{43} \to 0$ almost surely. For the similar reason, we can consider simply

$$\tilde{D}_{43} = \frac{1}{n^2} \sum_{i \neq j} \frac{X_i^* A_{ij} X_j \left(\frac{1}{n} X_i^* A_{ij} X_i - \frac{1}{n} tr A_{ij} \right)}{\left(1 + \frac{1}{n} tr A_{ij}\right)^3}.$$

For simplicity we use $S(i,j)$ to denote

$$S(i,j) = \frac{X_i^* A_{ij} X_j \left(\frac{1}{n} X_i^* A_{ij} X_i - \frac{1}{n} tr A_{ij} \right)}{\left(1 + \frac{1}{n} tr A_{ij}\right)^3}.$$

Now we will compute the fourth absolute moment of $\tilde{D}_{43}$, and we first expand $E|\tilde{D}_{43}|^4$ as

$$E|\tilde{D}_{43}|^4 = \frac{1}{n^8} \sum_{\substack{i_1 \neq j_1 \\ i_2 \neq j_2 \\ i_3 \neq j_3 \\ i_4 \neq j_4}} E\left[S(i_1, j_1) \bar{S}(i_2, j_2) S(i_3, j_3) \bar{S}(i_4, j_4) \right],$$

where $\bar{S}(i,j)$ is the complex conjugate of $S(i,j)$. Totally we need to use eight sub-indices here, although some of them may have the same value.

According to Assumption (i), Hölder inequality and Lemma 4.6, we know that

(3.11)
$$\begin{aligned} E|S(i,j)|^r &\leq \left(E(X_i^* A_{ij} X_j)^{2r}\right)^{1/2} \left(E\left(\frac{1}{n}X_i^* A_{ij} X_i - \frac{1}{n} trA_{ij}\right)^{2r}\right)^{1/2} \\ &\leq K_r n^{r/2} n^{-r/2} = O(1) \quad \text{for any } r \geq 2. \end{aligned}$$

Now it is easy to verify that the contribution of those terms with less than or equal to six different sub-indices in $E|\tilde{D}_{43}|^4$ is of the order $O(n^{-2})$, which is summable. Therefore, in order to show that

$$\tilde{D}_{43} \to 0 \quad \text{a.s.,} \tag{3.12}$$

we only need to consider the following two cases.

<u>Case 1: seven different indices</u>

When there are seven different indices, the summand has finite different forms depending on which two indices are the same. We only deal with the following kind of summands here,

$$E\left[S(i_1,j_1)\bar{S}(i_2,j_2)S(i_3,j_3)\bar{S}(i_3,j_4)\right] \tag{3.13}$$

The other forms of summand can be treated similarly.

Now for convenience we define an useful operator Δ_i. Let $f(A_\Lambda)$ be a function which involves the matrix A_Λ, and assume $i \notin \Lambda$. Δ_i is defined as

$$\Delta_i(f(A_\Lambda)) = f(A_\Lambda) - f(A_{\Lambda \cup \{i\}}).$$

For the term in (3.13), in the ideal situation, if X_{j_1} is independent with other parts, then the conditional expectation of X_{j_1} given all the other observations is zero, which means the expectation in (3.13) is zero. Unfortunately, this is not the case, because X_{j_1} is involved in matrices $A_{i_2j_2}$, $A_{i_3j_3}$ and $A_{i_3j_4}$. However, motivated by this idea, we can consider the following term

$$E\Bigg(\frac{X_{i_1}^* A_{i_1j_1} X_{j_1}\left(\frac{1}{n}X_{i_1}^* A_{i_1j_1} X_{i_1} - \frac{1}{n}trA_{i_1j_1}\right)}{\left(1+\frac{1}{n}trA_{i_1j_1}\right)^3} \tag{3.14}$$

$$
\times \frac{X_{j_2}^* A_{i_2 j_1 j_2} X_{i_2} \left(\frac{1}{n} X_{i_2}^* A_{i_2 j_1 j_2} X_{i_2} - \frac{1}{n} tr A_{i_2 j_1 j_2}\right)}{\left(1 + \frac{1}{n} tr A_{i_2 j_1 j_2}\right)^3}
\times \frac{X_{i_3}^* A_{i_3 j_1 j_3} X_{j_3} \left(\frac{1}{n} X_{i_3}^* A_{i_3 j_1 j_3} X_{i_3} - \frac{1}{n} tr A_{i_3 j_1 j_3}\right)}{\left(1 + \frac{1}{n} tr A_{i_3 j_1 j_3}\right)^3}
\times \frac{X_{j_4}^* A_{i_3 j_1 j_4} X_{i_3} \left(\frac{1}{n} X_{i_3}^* A_{i_3 j_1 j_4} X_{i_3} - \frac{1}{n} tr A_{i_3 j_1 j_4}\right)}{\left(1 + \frac{1}{n} tr A_{i_3 j_1 j_4}\right)^3} \Bigg).
$$

For simplicity we introduce the notation $S_k(i,j)$

$$
S_k(i,j) = \frac{X_i^* A_{ijk} X_j \left(\frac{1}{n} X_i^* A_{ijk} X_i - \frac{1}{n} tr A_{ijk}\right)}{\left(1 + \frac{1}{n} tr A_{ijk}\right)^3},
$$

and (3.14) can be written as

$$
E\left[S(i_1,j_1)\bar{S}_{j_1}(i_2,j_2)S_{j_1}(i_3,j_3)\bar{S}_{j_1}(i_3,j_4)\right]
$$

Note that now all the matrices involved in (3.14) are independent of X_{j_1}, so the expectation in (3.14) is zero, and hence subtracting (3.14) from (3.13) will not change the expectation in (3.13). This leads us to consider

(3.15)

$$
\begin{aligned}
S(i_1,j_1)\bar{S}(i_2,j_2)S(i_3,j_3)\bar{S}(i_3,j_4) &- S(i_1,j_1)\bar{S}_{j_1}(i_2,j_2)S_{j_1}(i_3,j_3)\bar{S}_{j_1}(i_3,j_4) \\
=& S(i_1,j_1)\left[\Delta_{j_1}\bar{S}(i_2,j_2)\right]S(i_3,j_3)\bar{S}(i_3,j_4) \\
&+ S(i_1,j_1)\bar{S}_{j_1}(i_2,j_2)\left[\Delta_{j_1}S(i_3,j_3)\right]\bar{S}(i_3,j_4) \\
&+ S(i_1,j_1)\bar{S}_{j_1}(i_2,j_2)S_{j_1}(i_3,j_3)\left[\Delta_{j_1}\bar{S}(i_3,j_4)\right].
\end{aligned}
$$

The explicit formula of $[\Delta_{j_1}S(i_2,j_2)]$ is given by

$$
\begin{aligned}
\Delta_{j_1}S(i_2,j_2) &= S(i_2,j_2) - S_{j_1}(i_2,j_2) \\
=& \frac{\left[-\Delta_{j_1}\left(1 + \frac{1}{n} tr A_{i_2 j_2}\right)^3\right] X_{i_2}^* A_{i_2 j_2} X_{j_2} \left(\frac{1}{n} X_{i_2}^* A_{i_2 j_2} X_{i_2} - \frac{1}{n} tr A_{i_2 j_2}\right)}{\left(1 + \frac{1}{n} tr A_{i_2 j_2}\right)^3 \left(1 + \frac{1}{n} tr A_{i_2 j_1 j_2}\right)^3} \\
&+ \frac{\left[\Delta_{j_1}\left(X_{i_2}^* A_{i_2 j_2} X_{j_2}\right)\right]\left(\frac{1}{n} X_{i_2}^* A_{i_2 j_2} X_{i_2} - \frac{1}{n} tr A_{i_2 j_2}\right)}{\left(1 + \frac{1}{n} tr A_{i_2 j_1 j_2}\right)^3} \\
&+ \frac{X_{i_2}^* A_{i_2 j_1 j_2} X_{j_2}\left[\Delta_{j_1}\left(\frac{1}{n} X_{i_2}^* A_{i_2 j_2} X_{i_2} - \frac{1}{n} tr A_{i_2 j_2}\right)\right]}{\left(1 + \frac{1}{n} tr A_{i_2 j_1 j_2}\right)^3},
\end{aligned} \tag{3.16}
$$

where

$$\Delta_{j_1}\left(1+\frac{1}{n}trA_{i_2j_2}\right)^3 = \left[\Delta_{j_1}\left(1+\frac{1}{n}trA_{i_2j_2}\right)\right]\left(1+\frac{1}{n}trA_{i_2j_2}\right)^2 + \left(1+\frac{1}{n}trA_{i_2j_1j_2}\right)\left[\Delta_{j_1}\left(1+\frac{1}{n}trA_{i_2j_2}\right)\right]\left(1+\frac{1}{n}trA_{i_2j_2}\right) + \left(1+\frac{1}{n}trA_{i_2j_1j_2}\right)^2\left[\Delta_{j_1}\left(1+\frac{1}{n}trA_{i_2j_2}\right)\right] \tag{3.17}$$

$$\Delta_{j_1}\left(1+\frac{1}{n}trA_{i_2j_2}\right) = \frac{1}{n}\frac{\frac{1}{n}X_{j_1}^*A_{i_2j_1j_2}^2X_{j_1}}{1+\frac{1}{n}X_{j_1}^*A_{i_2j_1j_2}X_{j_1}} \tag{3.18}$$

$$\Delta_{j_1}\left(X_{i_2}^*A_{i_2j_2}X_{j_2}\right) = \frac{1}{n}\frac{X_{i_2}^*A_{i_2j_1j_2}X_{j_1}X_{j_1}^*A_{i_2j_1j_2}X_{j_2}}{1+\frac{1}{n}X_{j_1}^*A_{i_2j_1j_2}X_{j_1}} \tag{3.19}$$

$$\Delta_{j_1}\left(\frac{1}{n}X_{i_2}^*A_{i_2j_2}X_{i_2} - \frac{1}{n}trA_{i_2j_2}\right) = \frac{1}{n^2}\frac{X_{i_2}^*A_{i_2j_1j_2}X_{j_1}X_{j_1}^*A_{i_2j_1j_2}X_{j_2}}{1+\frac{1}{n}X_{j_1}^*A_{i_2j_1j_2}X_{j_1}} - \frac{1}{n}\frac{\frac{1}{n}X_{j_1}^*A_{i_2j_1j_2}^2X_{j_1}}{1+\frac{1}{n}X_{j_1}^*A_{i_2j_1j_2}X_{j_1}}. \tag{3.20}$$

Combining equations (3.16) to (3.20), again by Lemma 4.6 and Hölder inequality, we find that

$$E\left|\Delta_{j_1}S(i_2,j_2)\right|^r = O(n^{-r/2}) \quad \text{for any } r \geq 2. \tag{3.21}$$

We can verify that $E\left|\Delta_{j_1}S(i_3,j_3)\right|^r$ and $E\left|\Delta_{j_1}S(i_3,j_4)\right|^r$ also have the above order. Therefore, by (3.11) and (3.15), the order of (3.13) is $O(n^{-1/2})$. Furthermore, the same order can be verified for all the other terms with seven different sub-indices. Since the number of the terms with seven different sub-indices is at most $O(n^7)$, we know the contribution of these terms in $E|\tilde{D}_{43}|^4$ is of the order $O(n^{-3/2})$, which is summable.

Remark. Note that if we compute the order of (3.13) directly by Lemma 4.6 and Hölder inequality, the result will be $O(1)$. By taking the difference between (3.13) and (3.14), the order is reduced by $n^{1/2}$. This order reduction method will be used frequently in the subsequent discussions.

Case 2: eight different indices

Now we consider the terms with eight different indices which have the form

$$E\left[S(i_1,j_1)\bar{S}(i_2,j_2)S(i_3,j_3)\bar{S}(i_4,j_4)\right]. \tag{3.22}$$

In order to simplify the long expressions, we introduce another operator Ψ_i. Let $f(A_{\Lambda_1}, A_{\Lambda_2}, \ldots, A_{\Lambda_m})$ be a function which involves the matrix $A_{\Lambda_1}, A_{\Lambda_2}$,

$\ldots, A_{\Lambda_m}$, Ψ_i is defined as

$$\Psi_i(f(A_{\Lambda_1}, A_{\Lambda_2}, \ldots, A_{\Lambda_m})) = f(A_{\Lambda'_1}, A_{\Lambda'_2}, \ldots, A_{\Lambda'_m}).$$

where

$$\Lambda'_k = \Lambda_k \cup \{i\}, \quad , k = 1, 2, \ldots, m.$$

As an example of this operator, note that $\Psi_k S(i,j) = S_k(i,j)$.

We compute the order of (3.22) along similar lines in Case 1. However, the procedure and computation here are more complicated, so we only describe the main steps and omit the details. We know due to (3.11),

$$\left|E\left[S(i_1,j_1)\bar{S}(i_2,j_2)S(i_3,j_3)\bar{S}(i_4,j_4)\right]\right| = O(1).$$

Our purpose is to reduce this order to $O(n^{-3/2})$ so that the contribution of all the terms with eight different indices in $E\left|\tilde{D}_{43}\right|^4$ is of the order $O(n^{-3/2})$, which is summable.

Motivated by the order reduction method in Case 1, we begin by considering the following term

$$E\left[S(i_1,j_1)\bar{S}_{j_1}(i_2,j_2)S_{j_1}(i_3,j_3)\bar{S}_{j_1}(i_4,j_4)\right].$$

Since all the matrices involved in $S(i_1,j_1)\bar{S}_{j_1}(i_2,j_2)S_{j_1}(i_3,j_3)\bar{S}_{j_1}(i_4,j_4)$ are independent of X_{j_1}, we know that

$$E\left[S(i_1,j_1)\bar{S}_{j_1}(i_2,j_2)S_{j_1}(i_3,j_3)\bar{S}_{j_1}(i_4,j_4)\right] = 0;$$

and hence

$$\begin{aligned} E[S(i_1,j_1)\bar{S}(i_2,j_2)S(i_3,j_3)\bar{S}(i_4,j_4)] &= E[S(i_1,j_1)\bar{S}(i_2,j_2)S(i_3,j_3)\bar{S}(i_4,j_4)] \\ &\quad - E[S(i_1,j_1)\bar{S}_{j_1}(i_2,j_2)S_{j_1}(i_3,j_3)\bar{S}_{j_1}(i_4,j_4)]. \end{aligned}$$

Thus, it suffices to consider the difference

$$\begin{aligned} &S(i_1,j_1)\bar{S}(i_2,j_2)S(i_3,j_3)\bar{S}(i_4,j_4) - S(i_1,j_1)\bar{S}_{j_1}(i_2,j_2)S_{j_1}(i_3,j_3)\bar{S}_{j_1}(i_4,j_4) \\ &= S(i_1,j_1)\left[\Delta_{j_1}\bar{S}(i_2,j_2)\right]S(i_3,j_3)\bar{S}(i_4,j_4) \quad (3.23) \\ &+ S(i_1,j_1)\bar{S}_{j_1}(i_2,j_2)\left[\Delta_{j_1}S(i_3,j_3)\right]\bar{S}(i_4,j_4) \quad (3.24) \\ &+ S(i_1,j_1)\bar{S}_{j_1}(i_2,j_2)S_{j_1}(i_3,j_3)\left[\Delta_{j_1}\bar{S}(i_4,j_4)\right]. \quad (3.25) \end{aligned}$$

Because of (3.21), we know the expectations of (3.23)$\sim$(3.25) are of the order $O(n^{-1/2})$, so we need to reduce these orders further. For the term (3.23), since

$$ES_{j_2}(i_1,j_1)\left[\Delta_{j_1}\bar{S}(i_2,j_2)\right]S_{j_2}(i_3,j_3)\bar{S}_{j_2}(i_4,j_4) = 0,$$

it is enough to consider

$$\begin{aligned}
&S(i_1,j_1)\left[\Delta_{j_1}\bar{S}(i_2,j_2)\right] S(i_3,j_3)\bar{S}(i_4,j_4) \\
&\qquad -S_{j_2}(i_1,j_1)\left[\Delta_{j_1}\bar{S}(i_2,j_2)\right] S_{j_2}(i_3,j_3)\bar{S}_{j_2}(i_4,j_4) \\
&\qquad =\left[\Delta_{j_2}S(i_1,j_1)\right]\left[\Delta_{j_1}\bar{S}(i_2,j_2)\right] S(i_3,j_3)\bar{S}(i_4,j_4) \qquad (3.26)\\
&\qquad +S_{j_2}(i_1,j_1)\left[\Delta_{j_1}\bar{S}(i_2,j_2)\right]\left[\Delta_{j_2}S(i_3,j_3)\right]\bar{S}(i_4,j_4) \qquad (3.27)\\
&\qquad +S_{j_2}(i_1,j_1)\left[\Delta_{j_1}\bar{S}(i_2,j_2)\right] S_{j_2}(i_3,j_3)\left[\Delta_{j_2}\bar{S}(i_4,j_4)\right]. \qquad (3.28)
\end{aligned}$$

It is not hard to see that the expectations of (3.26)∼(3.28) are of the order $O(n^{-1})$, so we have to use our order reduction method one more time. For the term (3.26), since

$$E\left[\Psi_{j_3}\Delta_{j_2}S(i_1,j_1)\right]\left[\Psi_{j_3}\Delta_{j_1}\bar{S}(i_2,j_2)\right] S(i_3,j_3)\bar{S}_{j_3}(i_4,j_4)=0,$$

we will consider

$$\begin{aligned}
&\left[\Delta_{j_2}S(i_1,j_1)\right]\left[\Delta_{j_1}\bar{S}(i_2,j_2)\right] S(i_3,j_3)\bar{S}(i_4,j_4) \\
&\qquad -\left[\Psi_{j_3}\Delta_{j_2}S(i_1,j_1)\right]\left[\Psi_{j_3}\Delta_{j_1}\bar{S}(i_2,j_2)\right] S(i_3,j_3)\bar{S}_{j_3}(i_4,j_4) \\
&\qquad =\left[\Delta_{j_3}\Delta_{j_2}S(i_1,j_1)\right]\left[\Delta_{j_1}\bar{S}(i_2,j_2)\right] S(i_3,j_3)\bar{S}(i_4,j_4) \qquad (3.29)\\
&\qquad +\left[\Psi_{j_3}\Delta_{j_2}S(i_1,j_1)\right]\left[\Delta_{j_3}\Delta_{j_1}\bar{S}(i_2,j_2)\right] S(i_3,j_3)\bar{S}(i_4,j_4) \qquad (3.30)\\
&\qquad +\left[\Psi_{j_3}\Delta_{j_2}S(i_1,j_1)\right]\left[\Psi_{j_3}\Delta_{j_1}\bar{S}(i_2,j_2)\right] S(i_3,j_3)\left[\Delta_{j_3}\bar{S}(i_4,j_4)\right]. \qquad (3.31)
\end{aligned}$$

In fact, the explicit expression of $[\Delta_{j_3}\Delta_{j_2}S(i_1,j_1)]$ and $[\Delta_{j_3}\Delta_{j_1}S(i_2,j_2)]$ can be obtained, from which, the orders of these two terms can be computed. However, the computation is very tedious, so we will omit the details; and only write the results here:

$$E\left|\Delta_{j_3}\Delta_{j_2}S(i_1,j_1)\right|^r=O(n^{-r}) \qquad (3.32)$$

$$E\left|\Delta_{j_3}\Delta_{j_1}S(i_2,j_2)\right|^r=O(n^{-r}) \qquad (3.33)$$

Due to (3.11), (3.21), (3.32) and (3.33), we know that the expectations of (3.29) ∼ (3.31) are of the order $O(n^{-3/2})$. All the other terms could be treated similarly.

With the results from the above two cases, we can complete the proof of (3.12), which leads us to consider

$$T_4=\frac{1}{n^2}\sum_{i\neq j}\frac{X_i^*A_{ij}X_j}{\left(1+\frac{1}{n}trA_{ij}\right)\left(1+\frac{1}{n}X_j^*A_{ij}X_j\right)}.$$

Similarly, in the denominator of T_4, $X_j^*A_{ij}X_j$ can also be replaced by trA_{ij}. In the following parts, we will focus on

$$T_5=\frac{1}{n^2}\sum_{i\neq j}\frac{X_i^*A_{ij}X_j}{\left(1+\frac{1}{n}trA_{ij}\right)^2}, \qquad (3.34)$$

and our task is to show that

$$T_5 \to 0 \quad \text{a.s.} \tag{3.35}$$

3.4.3. *Proof of (3.35).* Let

$$T(i,j) = \frac{X_i^* A_{ij} X_j}{\left(1 + \frac{1}{n} tr A_{ij}\right)^2},$$

then we can simplify the expression of T_5

$$T_5 = \frac{1}{n^2} \sum_{i \neq j} T(i,j).$$

Since T_5 is real, in order to prove (3.35), it is enough to show that

$$T_5^2 \to 0 \quad \text{a.s.} \tag{3.36}$$

We expand T_5^2 as

$$T_5^2 = \frac{1}{n^4} \sum_{i \neq j} T(i,j)^2 + \frac{1}{n^4} \sum_{i \neq j} T(i,j) T(j,i) + \frac{1}{n^4} \sum_{\substack{i_1 \neq j_1 \\ i_2 \neq j_2 \\ \{i_1, j_1\} \neq \{i_2, j_2\}}} T(i_1, j_1) T(i_2, j_2)$$

By simple computation (again due to Lemma 4.6 and Hölder inequality), we can find that

$$E \left| \frac{1}{n^4} \sum_{i \neq j} T(i,j)^2 \right|^2 = O(n^{-2}),$$

and therefore

$$\frac{1}{n^4} \sum_{i \neq j} T(i,j)^2 \to 0 \quad \text{a.s.}$$

Similarly,

$$\frac{1}{n^4} \sum_{i \neq j} T(i,j) T(j,i) \to 0 \quad \text{a.s.}$$

Therefore, in order to prove (3.36), it suffices to show that

$$\frac{1}{n^4} \sum_{\substack{i_1 \neq j_1 \\ i_2 \neq j_2 \\ \{i_1, j_1\} \neq \{i_2, j_2\}}} T(i_1, j_1) T(i_2, j_2) \to 0 \quad \text{a.s.} \tag{3.37}$$

In (3.37), we require $i_1 \neq j_1$, $i_2 \neq j_2$ and $\{i_1, j_1\} \neq \{i_2, j_2\}$, so there may be three or four different sub-indices.

When there are three different indices, the summand has finite different forms depending on which two indices are the same. We only consider the following kind of summand here

$$C_3 = \frac{1}{n^4} \sum_{\substack{i_1 \neq j_1 \\ i_1 \neq j_2 \\ j_1 \neq j_2}} T(i_1, j_1) T(i_1, j_2), \tag{3.38}$$

the summands of other forms could be treated similarly. We will compute the absolute second moment of C_3

$$E|C_3|^2 = \frac{1}{n^8} \sum ET(i_1, j_1) T(i_1, j_2) \bar{T}(k_1, l_1) \bar{T}(k_1, l_2) \tag{3.39}$$

Along almost the same lines as in Section 3.4.2, we can find that

$$E|C_3|^2 = O(n^{-3/2}).$$

As an immediate consequence of this order and Borel-Cantelli Lemma, we know

$$C_3 \to 0 \quad \text{a.s.} \tag{3.40}$$

For the terms with four different indices

$$C_4 = \frac{1}{n^4} \sum_{\substack{i_1 \neq j_1 \\ i_1 \neq j_2 \\ \{i_1, j_1\} \cap \{i_2, j_2\} = \emptyset}} T(i_1, j_1) T(i_2, j_2), \tag{3.41}$$

the same order reduction method used before can also be applied. Although the computation will be more complicated and tedious, we can prove that (details are omitted)

$$E|C_4|^2 = O(n^{-3/2})^{\dagger},$$

therefore, by Borel-Cantelli lemma,

$$C_4 \to 0 \quad \text{a.s.} \tag{3.42}$$

By (3.40) and (3.42), (3.37) is proved, and hence (3.36) is proved. As a result, we have (3.35), which is

$$T_5 \to 0 \quad \text{a.s.}$$

---

†In fact, it can be shown that the order is $O(n^{-2})$. However, $O(n^{-3/2})$ is small enough

Now we are in the position to conclude the proof of Theorem 1.4. By the discussion in Sections 3.4.1 and 3.4.2, we know that (3.35) leads to (3.8), which is

$$\lim T_2 = 0 \quad \text{a.s.}$$

In section 3.3, we show that (see (3.7))

$$\limsup T_1 < 1 \quad \text{a.s.}$$

Combining these two results, we have succeeded in proving (3.4)

$$\limsup \bar{X}^* \left(\mathbb{S} - \lambda I\right)^{-1} \bar{X} < 1 \quad \text{a.s.} \quad \forall\, 0 < \lambda < a$$

when $0 < c < 1$. As a result of Lemma 3.1, this means that we have established (3.1)

$$\liminf \lambda_{\min}(S) \geq a = (1 - \sqrt{c})^2 \quad \text{a.s.}$$

when $0 < c < 1$. The proof of Theorem 1.4 is now completed.

4. Some Lemmas. We first introduce a classical result in linear algebra. In fact it is Corollary 7.3.8 of Horn and Johnson (1985).

Lemma 4.1. *Suppose A and B are $m \times n$ complex matrices; and let $q = \min\{m\,, n\}$. If $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_q$ are the singular values of A and $\tau_1 \geq \tau_2 \geq \cdots \tau_q$ are the singular values of B, then*

$$|\sigma_i - \tau_i| \leq \|A - B\|, \quad \text{for all } i = 1, 2, \ldots, q,$$

where $\|A\|$ denotes the spectrum norm of the complex matrix A, which is defined as the largest singular value of A.

The following *rank inequality*, which helps us to measure the difference between two empirical distributions, was proved in Silverstein and Bai (1995).

Lemma 4.2. *For $n \times n$ Hermitian matrices A and B*

$$\|F^A - F^B\| \leq \frac{1}{n}\, rank(A - B),$$

where $\|f\| = \sup_x |f(x)|$.

In the subsequent lemma, we list three equalities which are used frequently in our proof. They could be proved by simple computation.

LEMMA 4.3. *Suppose A is an $n \times n$ complex matrix and $\beta \in \mathbb{C}^n$. If both A and $(A + \beta\beta^*)$ are non-singular and $1 + \beta^* A^{-1}\beta \neq 0$, then:*

$$(A + \beta\beta^*)^{-1}\beta = \frac{A^{-1}\beta}{1 + \beta^* A^{-1}\beta} \tag{4.1}$$

$$\beta^*(A + \beta\beta^*)^{-1} = \frac{\beta^* A^{-1}}{1 + \beta^* A^{-1}\beta} \tag{4.2}$$

$$A^{-1} - (A + \beta\beta^*)^{-1} = \frac{A^{-1}\beta\beta^* A^{-1}}{1 + \beta^* A^{-1}\beta} \tag{4.3}$$

The following lemma, which could be viewed as a generalization of Marcinkiewicz strong law of large numbers (see Loève (1963), pp. 242-243), was proved in Bai and Yin (1993).

LEMMA 4.4. *Let $\{X_{ij},\ i, j = 1, 2, \ldots\}$ be a double array of i.i.d. complex r.v.s. Let $\alpha > 1/2$, $\beta \geq 0$, and $M > 0$ be constants. Then, as $n \to \infty$,*

$$\max_{j \leq Mn^\beta} \left| n^{-\alpha} \sum_{i=1}^{n} (X_{ij} - c) \right| \to 0 \quad a.s.;$$

if and only if the following conditions are true:

(i) $E|X_{11}|^{(1+\beta)/\alpha} < \infty$;

(ii) $c = \begin{cases} EX_{11}, & \text{if } \alpha \leq 1; \\ \text{any value in } \mathbb{C}, & \text{if } \alpha > 1. \end{cases}$

The next result was proved in Bai and Silverstein (2004) (see (1.9b) and the theorem in the appendix).

LEMMA 4.5. *Under the conditions of Theorem 1.1, if the underlying variables are uniformly bounded, then we have when $c \in (0\,,\,1)$*

$$P(\lambda_{min}(\mathbb{S}) \leq \eta) = o(n^{-l})$$

for any $0 < \eta < (1 - \sqrt{c})^2$ and any positive l.

The first two inequalities in the following lemma were originally proved in Bai and Silverstein (1998) (Lemma 2.7 and Lemma A.1) by martingale inequalities. We also state some simple consequences for our purpose.

LEMMA 4.6. *Let $Y = (Y_1, Y_2, \ldots Y_n)^T$ be a random vector containing i.i.d. standardized complex entries, B be an $n \times n$ non-negative definite Hermitian matrix, and C be an $n \times n$ complex matrix, then*

$$E|Y^*BY|^p \le K_p \Big((tr\, B)^p + E|Y_1|^{2p} tr\, B^p \Big) \quad \text{for any } p \ge 1$$

$$E|Y^*CY - tr\, C|^p \le K_p \Big((E|Y_1|^4 tr\, CC^*)^{p/2} + E|Y_1|^{2p} tr\, (CC^*)^{p/2} \Big) \quad \text{for any } p \ge 2$$

If all the entries of Y are bounded by a constant M_1, and the norm of the non-negative definite Hermitian matrix A is bounded by another constant M_2, then we have the following immediate consequences.

$$E|Y^*AY|^p \le K_p n^p \quad \text{for any } p \ge 1; \tag{4.4}$$

$$E|Y^*AY - tr\, A|^p \le K_p n^{p/2} \quad \text{for any } p \ge 2; \tag{4.5}$$

and if Z is i.i.d. with Y, then

$$E|Y^*AZ|^p \le K_p n^{p/2} \quad \text{for any } p \ge 2. \tag{4.6}$$

These K_p's are constants only depending on p, and they do not need to have the same value

Acknowledgements. The author would like to thank Zhidong Bai for his helpful suggestions.

REFERENCES

[1] Bai, Z. D. (1999). Methodologies in spectral analysis of large dimensional random matrices, a review. *Stat. Sin.* **9** 611-677

[2] Bai, Z. D. and Silverstein, Jack W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** 553-605

[3] Bai, Z. D. and Silverstein, Jack W. (1998). No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.* **26** 316-345

[4] Bai, Z. D., Silverstein, Jack W. and Yin, Y. Q. (1988). A note on the largest eigenvalue of a large dimensional sample covariance matrix. *J. Multivariate Anal.* **26** 166-168

[5] Bai, Z. D. and Yin, Y. Q. (1993). Limit of the smallest eigenvalue of large dimensional covariance matrix. *Ann. Probab.* **21** 1275-1294

[6] Horn, R. A. and Johnson, C. R. (1985). *Matrix Analysis*. Cambridge Univ. Press.

[7] Jiang, Tiefeng (2004). The limiting distributions of eigenvalues of sample correlation matrices. *Sankhyā* **66** 35-48

[8] Loève, M. (1963). *Probability Theory*, 3ed. Van Nostrand, Princeton, N.J.

[9] Silverstein, Jack W. and Bai, Z. D. (1995). On the empirical distribution of eigenvalues of a class of large dimensional ramdom matrices. *J. Multivariate Anal.* **54** 175-192

[10] Yin, Y. Q., Bai, Z. D. and Krishnaiah, P. R. (1988). On the limit of the largest eigenvalue of the large dimensional sample covariance matrix. *Probab. Theory Relat. Fields* **78** 509-521

Department of Statistics and Applied Probability
6 Science Drive 2, National University of Singapore
Singapore 117546
e-mail: g0403223@nus.edu.sg; stazw@nus.edu.sg