JOURNAL OF MULTIVARIATE ANALYSIS 30, 1-16 (1989)

# On the Eigenvectors of Large Dimensional Sample Covariance Matrices

JACK W. SILVERSTEIN*

*North Carolina State University*

*Communicated by C. R. Rao*

Let $\{v_{ij}\}$, $i, j=1, 2, \ldots$, be i.i.d. random variables, and for each $n$ let $M_n=(1/s)V_nV_n^{\mathrm{T}}$, where $V_n=(v_{ij})$, $i=1, 2, \ldots, n$, $j=1, 2, \ldots, s=s(n)$, and $n/s \to y>0$ as $n \to \infty$. Necessary and sufficient conditions are given to establish the convergence in distribution of certain random variables defined by $M_n$. When $E(v_{11}^4)<\infty$ these variables play an important role toward understanding the behavior of the eigenvectors of this class of sample covariance matrices for $n$ large. 

## 1. INTRODUCTION

Let $\{v_{ij}\}$, $i, j=1, 2, \ldots$, be i.i.d. random variables, and for each $n$ let $M_n=(1/s)V_nV_n^{\mathrm{T}}$, where $V_n=(v_{ij})$, $i=1, 2, \ldots, n$, $j=1, 2, \ldots, s=s(n)$, and $n/s \to y>0$ as $n \to \infty$. In [9-11] an investigation into the behavior of the eigenvectors of the class, $M_n$, of sample covariance matrices for large $n$ has led to results suggesting that for standardized $v_{11}$ ($E(v_{11})=0$, $E(v_{11}^2)=1$), the behavior is similar to the Wishart case, that is, when $v_{11}=N(0, 1)$. In this case the orthogonal matrix $O_n$, whose columns contain the eigenvectors of $M_n$ ordered from left to right corresponding to the increasing size of the eigenvalues, induces the Haar (uniform) measure on $\mathcal{O}_n$, the $n \times n$ orthogonal group. Thus, a conjecture has been raised stating that, for general $v_{11}$, $O_n$ is close in some sense to being Haar distributed (Note that

Received March 30, 1987; accepted September 10, 1987.

AMS 1980 subject classifications: Primary 60B15; Secondary 62H99.

Key words and phrases: behavior of eigenvectors, Brownian bridge, convergence in distribution.

* Supported by the National Science Foundation under Grants MCS-8101703-A01 and DMS-8603966.



0047-259X/89 $3.00

$O_n$ is still not well defined because of ambiquities arising from multiple eigenvalues and the choice of direction for each eigenvector. However, it is possible to define uniquely a Borel probability on $\mathcal{O}_n$ naturally induced from the eigenvectors of $M_n$ [10]. We will assume that $O_n$ has this measure for its distribution.) It is suggested in [10] that before defining precisely what is meant by "close," mappings from $\mathcal{O}_n$ to a space $S$, common for all $n$, should be considered. These mappings define a sequence of measures on $S$ which have a known limit when $O_n$ is Haar distributed. It would then be important to determine under what conditions on the distribution of $v_{11}$ this limit is achieved.

Attention has focused on the following maps into $S = D[0, 1]$, the space of rcll functions on $[0, 1]$:

For each $n$ let $x_n \in R^n$, $\|x_n\| = 1$, be nonrandom. With $(y_1, y_2, ..., y_n)^{\mathrm{T}} \equiv O_n^{\mathrm{T}} x_n$ define $X_n\colon [0, 1] \to R$ as

$$X_n(t) = \sqrt{n/2}\left(\sum_{i=1}^{[nt]} (y_i^2 - 1/n)\right)$$

($[a]$ = greatest integer $\leqslant a$).

These mappings have been considered mainly because they carry over to $D[0, 1]$ much of the uniformity of Haar measure. Indeed, when $O_n$ is Haar distributed, $O_n^{\mathrm{T}} x_n$ is uniformly distributed on the unit sphere in $R^n$. As for the limiting behavior, it is straightforward to show that when $O_n$ is Haar distributed $X_n$ converges weakly to Brownian bridge, $W^0$, as $n \to \infty$.

Work on the limiting behavior of $X_n$ for general $v_{11}$ is still in progress. The aim of the present paper is to strengthen a limit theorem in [11] which will be a stepping stone to understanding $\{X_n\}$. It is at this point necessary to review two results on the limiting behavior of the *eigenvalues* of $M_n$.

It is known when $\mathrm{Var}(v_{11}) = 1$ the empirical distribution function $F_n$ of the eigenvalues of $M_n$ ($F_n(x) = (1/n) \times$ (number of eigenvalues of $M_n \leqslant x$)) converges almost surely for every $x \geqslant 0$ to a nonrandom probability distribution function $F_y$ having a density with support on $[(1-\sqrt{y})^2, (1+\sqrt{y})^2]$, and for $y > 1$, $F_y$ places mass $1 - 1/y$ at 0 [7, 8, 12, 13]. Moreover, if $v_{11}$ is standardized, then $\lambda_{\max}(M_n)$, the largest eigenvalue of $M_n$, satisfies

$$\lambda_{\max}(M_n) \to (1+\sqrt{y})^2 \qquad \text{a.s. as} \quad n \to \infty \tag{1.1}$$

if and only if $E(v_{11}^4) < \infty$ [1, 5, 14].

It is now possible to state the purpose of the paper, namely to prove the following

THEOREM. *Let $W_x^y \equiv W^0(F_y(x))$.*

(a) *We have*

$$\begin{aligned}\{\sqrt{n/2}\,(x_n^{\mathrm{T}} M_n^r x_n - (1/n)\operatorname{tr}(M_n^r))\}_{r=1}^{\infty} &= \left\{\int_0^{\infty} x^r\, dX_n(F_n(x))\right\}_{r=1}^{\infty} \\ &\xrightarrow{\mathscr{D}} \left\{\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^r\, dW_x^y\right\}_{r=1}^{\infty} \quad \text{as} \quad n \to \infty \end{aligned} \tag{1.2}$$

*($\mathscr{D}$ denoting weak convergence on $R^{\infty}$) for every sequence $\{x_n\}$, $x_n \in R^n$, $\|X_n\| = 1$ if and only if*

$$E(v_{11}) = 0, \qquad E(v_{11}^2) = 1, \qquad E(v_{11}^4) = 3. \tag{1.3}$$

(b) *If $\int_0^{\infty} x\, dX_n(F_n(x))$ is to converge in distribution to any random variable for each of the $x_n$ sequences $\{(1, 0, ..., 0)^{\mathrm{T}}\}$, $\{(1/\sqrt{n}, 1/\sqrt{n}, ..., 1/\sqrt{n})^{\mathrm{T}}\}$, then necessarily $E(v_{11}^4) < \infty$ and $E(v_{11}) = 0$.*

(c) *If $E(v_{11}^4) < \infty$ but $E[(v_{11} - E(v_{11}))^4]/(\operatorname{Var}(v_{11}))^2 \neq 3$, then there exist sequences $\{x_n\}$ for which*

$$\left(\int_0^{\infty} x\, dX_n(F_n(x)), \int_0^{\infty} x^2\, dX_n(F_n(x))\right)$$

*fails to converge in distribution.*

We note here that the limiting random variables in (1.2) are well-defined stochastic integrals, being jointly normal each with mean 0.

The importance of the theorem lies in its strong relation to the limiting behavior of $X_n$ and to the requirements imposed on the moments of $v_{11}$. If $X_n$ were to converge weakly to $W^0$ for all $\{x_n\}$ and $\operatorname{Var}(v_{11}) = 1$, $E(v_{11}^4) < \infty$, then (1.2) would follow from the above-mentioned behavior of the eigenvalues of $M_n$ and the theory of weak convergence of measures on function space ([10]). The theorem then tells us that $\operatorname{Var}(v_{11}) = 1$, $E(v_{11}^4) < \infty$, and weak convergence of $X_n$ to Brownian bridge for all $\{x_n\}$implies $E(v_{11}) = 0$ and $E(v_{11}^4) = 3$ (we remark that, for $v_{11}$ standardized and possessing moments of all order, the necessary condition on the fourth moment of $v_{11}$ was shown in [11]).

Of greater significance is the fact that, because of the theorem, what essentially remains to be studied is the question of tightness. Indeed, it can be shown that if (1.3) holds then $X_n \to^{\mathscr{D}} W^0$ on $D[0, 1]$ is equivalent to $X_n(F_n(x)) \to^{\mathscr{D}} W_x^y$ on $D[0, \infty)$, $\mathscr{D}$ denoting weak convergence on their respective spaces. We remark here that weak convergence on $D[0, \infty)$ is equivalent to weak convergence on $D[0, b]$ (under the natural projection) for every $b>0$. We can then use the fact that (1.1) and (1.2) would imply the uniqueness of any weak limit $X$ of a subsequence of $\{X_n(F_n(x))\}$ on $[0, b]$ for $b>(1+\sqrt{y})^2$. It would then be a matter of analyzing $X_n(F_n(x))$ for conditions of tightness such as Theorem 15.6 in [2]. At the very least, it is now known that, under the general conditions (1.3), if $X_n$, for some $\{x_n\}$, converges weakly, it must converge to Brownian bridge, and if $E(v_{11}^4)<\infty$ and $X_n$ converges weakly for all $\{x_n\}$, then

$$E(v_{11})=0, \qquad E[(V_{11}-E(v_{11}))^4]/(\mathrm{Var}(v_{11}))^2=3, \tag{1.4}$$

and the limit must be Brownian bridge.

Returning to the eigenvectors of $M_n$, if weak convergence of $X_n$ to Brownian bridge for all $\{x_n\}$ is considered a criterion for $O_n$ to be approximately Haar distributed for $n$ large, then we can say that for $E(v_{11}^4)<\infty$ but (1.4) not holding, the distribution of $O_n$ deviates significantly from Haar measure.

At present nothing is known for the case $E(v_{11}^4)=\infty$. The arguments used in [10] to show the necessary condition (1.2) depend on the boundedness of $\lambda_{\max}(M_n)$. It is possible that $X_n$ could converge weakly to Brownian bridge without (1.2) holding.

The theorem strengthens Theorem 1 of [11] which established (1.2) assuming (1.3) and the existence of all moments of $v_{11}$. The proof of the theorem is given in the next two sections. The proofs of (1.3) $\Rightarrow$ (1.2) and (c), given in Section 3, are similar to those of Theorem 1 in [11] which uses a multidimensional method of moments. In addition to extending those arguments previously used, we will need to perform successive truncations of the elements of $V_n$ using (1.1). The proofs of (1.2) $\Rightarrow$ (1.3) and (b), given in the next section, mainly rely on results in [6] which give conditions for sums of independent random variables to converge in distribution.

## 2. Proof of (1.2) $\Rightarrow$ (1.3), (b)

With $v_{11}=N(0,1)$ and $x_n=(1,0,\ldots,0)^{\mathrm{T}}$, (1.2) for $r=1$, together with the central limit theorem, implies

$$\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x\, dW_x^y = N(0, y).$$

Therefore, with $x_n = (x_1, x_2, ..., x_n)^{\mathrm{T}}$ we have

$$\sqrt{n/2}\,(x_n^{\mathrm{T}} M_n x_n - (1/n)\operatorname{tr}(M_n))$$
$$= \sqrt{n/s}\,1/\sqrt{2s}\sum_{j=1}^{s}\left(\left(\sum_{i=1}^{n} x_i v_{ij}\right)^2 - (1/n)\sum_{i=1}^{n} v_{ij}^2\right)$$
$$\xrightarrow{\mathscr{D}} N(0, y) \qquad \text{as} \quad n \to \infty.$$

We will use this fact on three different types of $x_n$ along with the following:

If $S_s = X_1(s) + X_2(s) + \cdots + X_s(s)$, $X_i(s)$ i.i.d., $i = 1, ..., s$, converges in distribution to $N(0, 1)$ as $s \to \infty$, then $X_1(s) \to^{\text{i.p.}} 0$ [3, p. 191]. Moreover, from [6, Theorem 2, p. 128] we have for every $\varepsilon > 0$

(1) $sP(|X_1(s)| \geqslant \varepsilon) \to 0$ and

(2) $s\operatorname{Var}(X_1(s) I_{(|X_1(s)| < \varepsilon)}) \to 1$

as $s \to \infty$ ($I_A$ denoting the indicator function on the set $A$).

With $x_n = (1, 0, ..., 0)^{\mathrm{T}}$ we can write $\sqrt{n/2}\,(x_n^{\mathrm{T}} M_n x_n - (1/n)\operatorname{tr}(M_n)) = \sqrt{n/s}\,((n-1)/n)(1/\sqrt{2s})\sum_{j=1}^{s}(v_{1j}^2 - (1/(n-1))\sum_{i=2}^{n} v_{ij}^2)$. With $X_1(s) = (1/\sqrt{2s})(v_1^2 - (1/(n-1))\sum_{i=2}^{n} v_i^2)$ ($v_i$ $i = 1, 2, ..., n$ i.i.d. having the same distribution as $v_{11}$) we immediately get $(1/((n-1)\sqrt{s}))\sum_{i=2}^{n} v_i^2 \to^{\text{i.p.}} 0$. We also have for any $\varepsilon > 0$,

$$sP\left(\left|v_1^2 - (1/(n-1))\sum_{i=2}^{n} v_i^2\right| \geqslant \varepsilon\sqrt{s}\right)$$
$$\geqslant sP\left(v_1^2 \geqslant (1/(n-1))\sum_{i=2}^{n} v_i^2 + \varepsilon\sqrt{s}\right)$$
$$\geqslant sP\left(v_1^2 \geqslant (1/(n-1))\sum_{i=2}^{n} v_i^2 + \varepsilon\sqrt{s},\ (1/(n-1))\sum_{i=2}^{n} v_i^2 \leqslant \varepsilon\sqrt{s}\right)$$
$$\geqslant sP(v_1^2 \geqslant 2\varepsilon\sqrt{s})\,P\left((1/(n-1))\sum_{i=2}^{n} v_i^2 \leqslant \varepsilon\sqrt{s}\right).$$

Therefore, from (1) we have $sP(v_1^2 \geqslant 2\varepsilon\sqrt{s}) \to 0$ as $s \to \infty$, which implies $m = E(v_{11}^2) < \infty$. Therefore, $v_1^2 - (1/(n-1))\sum_{i=2}^{n} v_i^2 \to^{\text{i.p.}} v_1^2 - m$ as $s \to \infty$.

Let $A_s = (|v_1^2 - (1/(n-1))\sum_{i=2}^{n} v_i^2| < \varepsilon\sqrt{s})$. From (2) we have

$$\operatorname{Var}\left(\left(v_1^2 - (1/(n-1))\sum_{i=2}^{n} v_i^2\right) I_{A_s}\right) \to 2 \qquad \text{as} \quad s \to \infty. \tag{2.1}$$

Since $E((v_1^2-(1/(n-1))\sum_{i=2}^n v_i^2)I_{A_s})=E(v_1^2 I_{A_s})-E(v_2^2 A_s)$ and $P(A_s)\to 1$ as $s\to\infty$, we can apply the dominated convergence theorem to each term to conclude $E((v_1^2-(1/(n-1))\sum_{i=2}^n v_i^2)I_{A_s})\to 0$ as $s\to\infty$.

Using (2.1) and Fatou's lemma we have $E((v_1^2-m)^2)\leqslant 2$. Thus $E(v_{11}^4)<\infty$.

Writing $1/\sqrt{2s}\sum_{j=1}^s(v_{1j}^2-(1/(n-1))\sum_{i=2}^n v_{ij}^2)$ as

$$1/\sqrt{2s}\sum_{j=1}^s(v_{1j}^2-E(v_{11}^2))-(1/(\sqrt{2s}\,(n-1)))\sum_{j=1}^s\sum_{i=2}^n(v_{ij}^2-E(v_{11}^2)), \tag{2.2}$$

we see that the central limit theorem implies the second sum in (2.2) converges in probability to zero, and in turn, the first sum converges in distribution to $N(0,1)$. We have then

$$E(v_{11}^4)-E^2(v_{11}^2)=2. \tag{2.3}$$

Now let $x_n=(1/\sqrt{n},1/\sqrt{n},...,1/\sqrt{n})^{\mathrm{T}}$. We have

$$\begin{aligned}&\sqrt{n/2}\,(x_n^{\mathrm{T}}M_n x_n-(1/n)\operatorname{tr}(M_n))\\&\quad=\sqrt{n/s}\,(1/n)\,1/\sqrt{2s}\sum_{j=1}^s\left(\left(\sum_{i=1}^n v_{ij}\right)^2-\sum_{i=1}^n v_{ij}^2\right).\end{aligned}$$

With $X_1(s)=(1/n)(1/\sqrt{2s})((\sum_{i=1}^n v_i)^2-\sum_{i=1}^n v_i^2)$ we see that $(1/n^{3/4})\sum_{i=1}^n v_i \xrightarrow{\text{i.p.}} 0$, which implies $E(v_{11})=0$.

Finally, let $x_n=(1/\sqrt{2},1/\sqrt{2},0,...,0)^{\mathrm{T}}$. We have

$$\begin{aligned}&\sqrt{n/2}\,(x_n^{\mathrm{T}}M_n x_n-(1/n)\operatorname{tr}(M_n))\\&\quad=\sqrt{n/s}\,(1/\sqrt{2s})\left(\sum_{j=1}^s((1/2)(v_{1j}+v_{2j})^2-E(v_{11}^2))\right)\\&\qquad-(1/n)\sum_{i,j}(v_{ij}^2-E(v_{11}^2)).\end{aligned}$$

By the central limit theorem we have $(1/(n\sqrt{s}))\sum_{i,j}(v_{ij}^2-E(v_{11}^2))\xrightarrow{\text{i.p.}}0$, and since $E((1/2)(v_{11}+v_{21})^2-E(v_{11}^2))=0$, we have $\operatorname{Var}((1/2)(v_{11}+v_{21})^2)=2$. Therefore, $E(v_{11}^4)+E^2(v_{11}^2)=4$. This, together with (2.3) gives us $E(v_{11}^2)=1$, $E(v_{11}^4)=3$.

To show (b) we use $X_1(s)\xrightarrow{\text{i.p.}}0$ and [6, Theorem 4, p. 124] which implies that if $S_s$ converges in distribution then the quantities in (1) and (2) are, for $\varepsilon>0$ sufficiently small, bounded in $s$. The above arguments will yield $sP(v_1^2\geqslant 2\varepsilon\sqrt{s})$ bounded in $s$ which still gives us $E(v_{11}^2)<\infty$. The statements $E(v_{11}^4)<\infty$ and $E(v_{11})=0$ will then follow as above.

## 3. Proof of (1.3) ⇒ (1.2), (c)

Let $v'_{ij}(n) = v_{ij} I_{(|v_{ij}| \leqslant n^{1/2})}$ and let $M'_n = (1/s) V'_n V'^{\mathrm{T}}_n$, where $V'_n = (v_{ij}(n))(n \times s)$. Then for any measurable function $f_n$ on $n \times n$ matrices and any $\varepsilon > 0$

$$P(|f_n(M_n) - f_n(M'_n)| > \varepsilon) \leqslant Cn^2 P(|v_{11}| > n^{1/2}) \to 0 \quad \text{as} \quad n \to \infty,$$

since $E(v_{11}^4) < \infty$. Therefore,

$$|f_n(M_n) - f_n(M'_n)| \xrightarrow{\text{i.p.}} 0 \quad \text{as} \quad n \to \infty. \tag{3.1}$$

From (1.1) and (3.1) we have

$$\lambda_{\max}(M'_n) \xrightarrow{\text{i.p.}} (1 + \sqrt{y})^2 \quad \text{as} \quad n \to \infty.$$

Let $M''_n = (1/s)(V'_n - E(v'_{11}(n)) 1_n 1_s^{\mathrm{T}})(V'_n - E(v'_{11}(n) 1_n 1_s^{\mathrm{T}})^{\mathrm{T}}$, where $1_m$ denotes the $m$-dimensional vector consisting of 1's. Let $\| \ \|$ denote the spectral norm of any matrix, that is, $\|A\| = \lambda_{\max}^{1/2}(AA^{\mathrm{T}})$.

We have

$$\begin{aligned} |\lambda_{\max}^{1/2}(M'_n) - \lambda_{\max}^{1/2}(M''_n)| &\leqslant \|(1/\sqrt{s})\, E(v'_{11}(n)) 1_n 1_s^{\mathrm{T}}\| \\ &= n^{1/2} |E(v'_{11}(n))| \to 0 \quad \text{as} \quad n \to \infty, \end{aligned}$$

using the fact that $E(v_{11}) = 0$, $E(v_{11}^4) < \infty$ implies

$$(Ev'_{11}(n)) = o(n^{-3/2}).$$

Therefore, $\lambda_{\max}(M''_n) \to^{\text{i.p.}} (1 + \sqrt{y})^2$.

It is a simple matter to show for any $n \times n$ matrices $A$, $B$, and integer $r \geqslant 1$ $\|(A+B)^r - B^r\| \leqslant r\|A\|(\|A\| + \|B\|)^{r-1}$. Therefore,

$$\begin{aligned} &\sqrt{n}\, |x_n(M''_n)^r x_n - x_n^{\mathrm{T}}(M'_r)^r x_n| \\ &\quad \leqslant \sqrt{n}\, \|(M''_n)^r - (M'_n)^r\| \\ &\quad \leqslant \sqrt{n}\, r\|M''_n - M'_n\|(\|M''_n\| + 2\|M'_n\|)^{r-1} \\ &\quad = \sqrt{n}\, r\|M''_n - M'_n\|(\lambda_{\max}(M''_n) + 2\lambda_{\max}(M'_n))^{r-1}. \end{aligned}$$

We have the last factor converging i.p. to $(3(1+\sqrt{y})^2)^{r-1}$.

We also have for matrices $A$, $B$ of the same dimension $\|AA^{\mathrm{T}} - BB^{\mathrm{T}}\| \leqslant \|A - B\|(\|A\| + \|B\|)$. Therefore,

$$\begin{aligned} &\sqrt{n}\, \|M''_n - M'_n\| \\ &\quad \leqslant \sqrt{n}\, \|(1/\sqrt{s})\, E(v'_{11}(n))\, 1_n 1_s^{\mathrm{T}}\|(\lambda_{\max}^{1/2}(M''_n) + \lambda_{\max}^{1/2}(M'_n)) \\ &\quad = n|E(v'_{11}(n))|(\lambda_{\max}^{1/2}(M''_n) + \lambda_{\max}^{1/2}(M'_n)) \xrightarrow{\text{i.p.}} 0 \quad \text{as} \quad n \to \infty. \end{aligned}$$

Therefore,

$$\sqrt{n}\,|x_n^{\mathrm{T}}(M_n'')^r x_n - x_n^{\mathrm{T}}(M_n')^r x_n| \xrightarrow{\text{i.p.}} 0 \qquad \text{as} \quad n \to \infty. \tag{3.2}$$

We also have

$$\sqrt{n}\,|(1/n)\operatorname{tr}((M_n'')^r) - (1/n)\operatorname{tr}((M_n')^r)|$$
$$\leqslant (1/\sqrt{n}) \sum_{i=1}^{n} |\lambda_i''^r - \lambda_i'^r|$$

($\lambda_i''$, $\lambda_i'$ being the respective eigenvalues of $M_n''$, $M_n'$ arranged in nondecreasing order)

$$\leqslant 2r(1/\sqrt{n})\left(\sum_{i=1}^{n} |\lambda_i''^{1/2} - \lambda_i'^{1/2}|\right)$$
$$\times \max((\lambda_{\max}(M_n''))^{(1/2)(2r-1)}, (\lambda_{\max}(M_n'))^{(1/2)(2r-1)}).$$

By Theorem 2 of [4] we have for each $i = 1, 2, ..., n$,

$$|\lambda_i''^{1/2} - \lambda_i'^{1/2}| \leqslant n^{1/2}|E(v_{11}'(n))|.$$

Therefore,

$$\sqrt{n}\,|(1/n)\operatorname{tr}((M_n'')^r) - (1/n)\operatorname{tr}((M_n')^r)|$$
$$= o(n^{(-3/2)+1}) \max((\lambda_{\max}(M_n''))^{(1/2)(2r-1)},$$
$$(\lambda_{\max}(M_n'))^{(1/2)(2r-1)}) \xrightarrow{\text{i.p.}} 0 \qquad \text{as} \quad n \to \infty. \tag{3.3}$$

Therefore, by (3.1), (3.2), and (3.3) we have for each integer $r \geqslant 1$,

$$|\sqrt{n/2}\,(x_n^{\mathrm{T}} M_n'^r x_n - (1/n)\operatorname{tr}(M_n'^r))$$
$$- \sqrt{n/2}\,(x_n^{\mathrm{T}}(M_n'')^r x_n - (1/n)\operatorname{tr}((M_n'')^r))| \xrightarrow{\text{i.p.}} 0 \qquad \text{as} \quad n \to \infty.$$

We see then that it is sufficient to prove (1.2) for $M_n''$.

To simplify notation we will identify $u_{ij}$ with $v_{ij}'(n) - E(v_{ij}'(n))$, suppressing the dependence on $n$. Using the fact that for any random variable $X$ and positive $k$ and $\alpha$,

$$E(|X|^k) < \infty \Rightarrow E(|X|^r I_{(|X| \leqslant n^{\alpha})}) = o(n^{\alpha(r-k)}) \tag{3.4}$$

for any $r > k$, we have $u_{ij}$ i.i.d., $E(u_{11}) = 0$, $E(u_{11}^2) \to 1$, $E(u_{11}^4) \to 3$, as $n \to \infty$, and $E(u_{11}^r) = o(n^{(r/2)-2})$ for $r > 4$.

The remainder of the proof of (1.3) ⇒ (1.2) will appear similar to the one given for Theorem 1 of [11]. However, some of the arguments will be expressed differently to accommodate the truncation of the elements of $M_n$ and to render the proof easier to understand. It will be practically self-contained, making only two references to [11].

We start with

LEMMA 1. *For any integer* $r \geqslant 1$, $(1/\sqrt{n})(\operatorname{tr}((M_n'')^r) - E(\operatorname{tr}((M_n'')^r))) \xrightarrow{\text{i.p.}} 0$ *as* $n \to \infty$.

*Proof.* We have

$$\begin{aligned} s^{2r} &\operatorname{Var}(\operatorname{tr}((M_n'')^r)) \\ &= \sum_{\substack{i_1, \ldots, i_r, k_1, \ldots, k_r \\ i_1', \ldots, i_r', k_1', \ldots, k_r'}} E(u_{i_1 k_1} u_{i_2 k_1} u_{i_2 k_2} \cdots u_{i_r k_r} u_{i_1 k_r} u_{i_1' k_1'} u_{i_2' k_1'} u_{i_2' k_2'} \\ &\qquad \cdots u_{i_r' k_r'} u_{i_1' k_r'}) - E(u_{i_1 k_1} u_{i_2 k_1} u_{i_2 k_2} \cdots u_{i_r k_r} u_{i_1 k_r}) \\ &\qquad \times E(u_{i_1' k_1'} u_{i_2' k_1'} u_{i_2' k_2'} \cdots u_{i_r' k_r'} u_{i_1' k_r'}). \end{aligned} \tag{3.5}$$

A term in (3.5) is zero if

(1) a $u_{ik}$ or $u_{i'k'}$ appears alone, or

(2) no $u_{ik}$ equals a $u_{i'k'}$, or

(3) no $i(k)$ index equals an $i'(k')$ index.

Using the fact that for any random variable $X$ and $a, b \geqslant 0$, $E(|X|^a)\, E(|X|^b) \leqslant E(|X|^{a+b})$, we have

$$s^{2r} \operatorname{Var}(\operatorname{tr}((M_n'')^r)) \leqslant 2 \sum{}' E(|u_{i_1 k_1} \cdots u_{i_1 k_r} u_{i_1' k_1'} \cdots u_{i_1' k_r'}|), \tag{3.6}$$

where $\sum'$ denotes the summation is being taken over those terms of (3.5) for which (1) and (2) (and consequently (3)) are avoided.

Consider one of the (finite) number of ways the sets of indices $\{i_1, \ldots, i_r, i_1', \ldots, i_r'\}$, $\{k_1, \ldots, k_r, k_1', \ldots, k_r'\}$ can each be partitioned. Let $d$ denote the total number of classes making up the two partitions. Associated with the two partitions are the terms in (3.6) (for $n$ large), where indices are equal in value if and only if they belong to the same class. We only consider those partitions resulting in terms for which (1) and (2) are avoided. The number of terms is bounded by $Kn^d$ (since $n/s \to y > 0$). The terms are identical involving, say, $r'$ distinct elements of $V_n$, $1 \leqslant r' \leqslant 2r$.

Choose one of these terms. Let $\mathbf{r}_\alpha$, $\alpha = 2, 3, \ldots, 4r$, denote the number of distinct elements of $U_n \equiv (u_{ij})$ appearing $\alpha$ times in this term. Then $\sum_\alpha \mathbf{r}_\alpha = r'$ and $\sum_\alpha \alpha \mathbf{r}_\alpha = 4r$. Using (3.4), if there is an $\mathbf{r}_\alpha \geqslant 1$ for $\alpha > 4$, then the term is

$$o(n^{((1/2)5 - 2)\mathbf{r}_5 + \cdots + ((1/2)4r - 2)\mathbf{r}_{4r}}) = o(n^{2r - 2r' + \mathbf{r}_2 + (1/2)\mathbf{r}_3}). \tag{3.7}$$

We will show

$$d \leqslant \min(r'+1, 2r). \tag{3.8}$$

In order to do so we will need to verify

$$d \leqslant \min(r'+1, r+\mathbf{r}), \tag{3.9}$$

where $d$ and $r'$ are defined as above for

$$E(|u_{i_1k_1} \cdots u_{i_1k_r} u_{i'_1k'_1} \cdots u_{i'_1k'_{\mathbf{r}}}|) \tag{3.10}$$

for arbitrary $r, \mathbf{r} \geqslant 1$ and where (1) and (2) are avoided. We have $r' \leqslant r + \mathbf{r}$. (3.10) can be written as:

$$A^1_{a_1b_1} A^2_{a_2b_2} \cdots A^{r'}_{a_{r'}b_{r'}}, \tag{3.11}$$

where $A^j_{a_jb_j}$ corresponds to $u_{a_jb_j}$ appearing in (3.10), so that if $u_{a_jb_j}$ appears $t$ times, then $A^j_{a_jb_j} = E(|u^t_{a_jb_j}|)$. The ordered pairs $(a_j, b_j)$ will be distinct, but because of the weaving pattern of the $i$, $i'$, $k$, $k'$ indices and the fact that (3) does not hold, when $r' > 1$, for each $(a_j, b_j)$ either $a_j$ or $b_j$ will be repeated in at least one other ordered pair. Notice that $d$ is equal to the number of distinct $a_j$, $b_j$. We say that $a_j$ or $b_j$ is *free* if it does not appear in any other ordered pair. If $a_j$ or $b_j$ is free then $A^j_{a_jb_j}$ must be formed from adjacent pairs of $u_{ik}$ and (or) $u_{i'k'}$.

We prove (3.9) using induction on $r'$. The case $r' = 1$ being obvious we assume (3.9) is true for all expressions (3.10) with arbitrary $r$, $\mathbf{r}$, where (1) and (3) are avoided and $r' = t-1$. For $r' = t$ consider (3.11). Let $f$ denote the number of free indices. Then $d \leqslant (2r'-f)/2 + f = r' + (f/2)$, so if $f \leqslant 1$ then (3.9) certainly holds. If $f > 1$, then an $A^j_{a_jb_j}$ can be removed from (3.11) resulting in an expression arising from (3.10) where the total number of $i$, $i'$, $k$, $k'$ indices is reduced, having $d-1$ distinct indices and where (1) and (3) are avoided. The inductive hypothesis then yields (3.9).

To complete the proof of Lemma 1 we take a bound on the common value (3.11) of the terms associated with the partitions, multiply by the bound $Kn^d$ on the number of these terms, and divide by $n \times s^{2r}$. If $\mathbf{r}_\alpha = 0$ for all $\alpha > 4$ then (3.11) is bounded and by (3.8) $n^d/n \times s^{2r} = O(1/n)$. Otherwise, we use (3.7) and we have $o(n^{2r-2r'+\mathbf{r}_2+(1/2)\mathbf{r}_3}) n^d/n \times s^{2r} = o(n^{-2r'+\mathbf{r}_2+(1/2)\mathbf{r}_3+d-1})$.

Using (3.8), the exponent of $n$ in the last expression is bounded by $-r' + \mathbf{r}_2 + \frac{1}{2}\mathbf{r}_3 \leqslant 0$. Summing on all appropriate partitions we get $(3.5)/n \times s^{2r} \to 0$ as $n \to \infty$ and we are done.

LEMMA 2. *For any integer* $r \geqslant 1$, $\sqrt{n}\,(E(x_n^{\mathrm{T}}(M''_n)^r x_n - E((1/n)\operatorname{tr}((M''_n)^r))) \to 0$ *as* $n \to \infty$.

*Proof.* Using the fact that the diagonal elements of $(M''_n)^r$ are identically distributed, we have

$$s^r(E(x_n(M''_n)^r x_n) - E((1/n)\,\mathrm{tr}((M''_n)^r)))$$
$$= \sum_{i \neq j} x_i x_j \sum_{\substack{i_2, \ldots, i_r \\ k_1, \ldots, k_r}} E(u_{ik_1} u_{i_2 k_1} \cdots u_{jk_r}).$$

A term in the above summation is zero if

(1) a $u_{ik}$ appears alone.

Notice also that the off-diagonal terms of $(M''_n)^r$ are identically distributed and that $|\sum_{i \neq j} x_i x_j| \leqslant n-1$. Therefore,

$$|s^r(E(x_n^{\mathrm{T}}(M''_n)^r x_n) - E((1/n)\,\mathrm{tr}((M''_n)^r)))|$$
$$< n \sum{}' E(|u_{1k_1} u_{i_2 k_1} \cdots u_{2k_r}|), \tag{3.12}$$

where $\sum'$ denotes the summation avoids (1).

As in Lemma 1 let (3.11) be one of the terms of (3.12) associated with a pair of partitions on $\{i_2, \ldots, i_r\}$, $\{k_1, \ldots, k_r\}$, where (1) is avoided, with $d$, $r'$, and $\mathbf{r}_\alpha$ defined as before. We have $2 \leqslant \alpha \leqslant 2r$ (since (3.11) must involve at least two distinct elements of $U_n$), $\sum \alpha \mathbf{r}_\alpha = 2r$, $\sum \mathbf{r}_\alpha = r'$, and $r' < r$. The last statement follows from the fact that an odd number of the $2r$ $u_{ik}$'s making up (3.11) are taken from the first row of $U_n$. Since 1 and 2 must be among th $a_i$ indices, the number of terms associated with the two partitions is bounded by $Kn^{d-2}$.

As in Lemma 1 it is straightforward to show by induction that $d \leqslant r' + 1$. Since $r' < r$ we must have $d \leqslant r$.

Proceeding as in Lemma 1, if $\mathbf{r}_\alpha = 0$ for all $\alpha > 4$, then (3.11) is bounded and we have

$$n^{3/2} n^{d-2}/s^r = O(n^{-1/2}).$$

Otherwise, we use (3.4) to find that (3.11) is

$$o(n^{r - 2r' + \mathbf{r}_2 + (1/2)\mathbf{r}_3}) \tag{3.13}$$

and we have $(3.13) \times n^{d-(1/2)}/s^r = o(n^{-2r' + \mathbf{r}_2 + (1/2)\mathbf{r}_3 + d - (1/2)})$. The exponent of $n$ is bounded by $-r' + \mathbf{r}_2 + \frac{1}{2}\mathbf{r}_3 + \frac{1}{2} \leqslant 0$ as in Lemma 1, using the additional fact that there is an $\mathbf{r}_\alpha \geqslant 1$ with $\alpha > 4$.

We conclude that $\sqrt{n} \times (3.12)/s^r \to 0$ as $n \to \infty$ which proves Lemma 2.

At this point it is necessary to introduce another truncation of the elements of $V_n$. Let $\bar{v}_{ij} = \bar{v}_{ij}(n) = v_{ij} I_{(|v_{ij}| \leqslant n^{1/4})} - E(v_{ij} I_{(|v_{ij}| \leqslant n^{1/4})})$ and let

$\bar{M}_n=(1/s)\bar{V}_n\bar{V}_n^{\mathrm{T}}$, where $\bar{V}_n=(\bar{v}_{ij})$. We have $E(\bar{v}_{11})=0$, $E(\bar{v}_{11}^2)\to 1$, and $E(\bar{v}_{11}^4)\to 3$ as $n\to\infty$. It will be shown that for every integer $r\geqslant 1$,

$$\sqrt{n}\,(x_n^{\mathrm{T}}(M_n'')^r x_n-E(x_n^{\mathrm{T}}(M_n'')^r x_n)$$
$$-(x_n^{\mathrm{T}}\bar{M}_n^r x_n-E(x_n^{\mathrm{T}}\bar{M}_n^r x_n)))\xrightarrow{\text{i.p.}}0 \qquad \text{as} \quad n\to\infty. \tag{3.14}$$

However, before proving (3.14) we will use it and Lemmas 1 and 2 to complete the proof of the theorem. (1.2) will follow if we can show for any integer $m\geqslant 2$ and positive integers $r_1, r_2, \ldots, r_m$ the asymptotic behavior of

$$n^{m/2}E[(x_n^{\mathrm{T}}\bar{M}_n^{r_1}x_n-E(x_n^{\mathrm{T}}\bar{M}_n^{r_1}x_n))(x_n^{\mathrm{T}}\bar{M}_n^{r_2}x_n-E(x_n^{\mathrm{T}}\bar{M}_n^{r_2}x_n))$$
$$\cdots(x_n^{\mathrm{T}}\bar{M}_n^{r_m}x_n-E(x_n^{\mathrm{T}}\bar{M}_n^{r_m}x_n))] \tag{3.15}$$

depends only on $E(\bar{v}_{11}^2)$ and $E(\bar{v}_{11}^4)$ (see [11, p. 302]).

We have

$$(s^{r_1+\cdots+r_m}/n^{m/2})\times(3.15)$$
$$=\sum_{\substack{i^1,j^1,i_2^1,\ldots,i_{r_1}^1,k_1^1,\ldots,k_{r_1}^1\\ i^m,j^m,i_2^m\ldots,i_{r_m}^m,k_1^m\ldots,k_{r_m}^m}} x_{i^1}x_{j^1}\cdots x_{i^m}x_{j^m}\,E[(\bar{v}_{i^1k_1^1}\cdots\bar{v}_{j^1k_{r_1}^1}$$
$$-E(\bar{v}_{i^1k_1^1}\cdots v_{j^1k_{r_1}^1}))\cdots(\bar{v}_{i^mk_1^m}\cdots\bar{v}_{j^mk_{r_m}^m}-E(\bar{v}_{i^mk_1^m}\cdots\bar{v}_{j^mk_{r_m}^m}))]. \tag{3.16}$$

As before we see that a zero term occurs when

(1) a $\bar{v}_{ik}$ appears alone, or

(2) for a given $t$ no $\bar{v}_{p^tq^t}$ equals a $\bar{v}_{p^{t'}q^{t'}}$, $t'\neq t$.

Consider one of the ways the sets $I\equiv\{i^1, j^1, i_2^1, \ldots, i_{r_1}^1, \ldots, i^m, j^m, i_2^m, \ldots, i_{r_m}^m\}$. $K\equiv\{k_1^1, \ldots, k_{r_1}^1, \ldots, k_1^m, \ldots, k_{r_m}^m\}$ can each be partitioned so that those terms in (3.16) associated with the two partitions avoid (1) and (2). Let $l$ be the number of classes of $I$ indices containing only one element from $J\equiv\{i^1, j^1, i^2, j^2, \ldots, i^m, j^m\}$. Let $d$ denote the number of classes of $I$ indices containing no elements from $J$, plus the number of classes of $K$ indices. Then the contribution to (3.16) of those terms associated with the two partitions is bounded in absolute value by

$$Cn^{(l/2)+d}E(|\bar{v}_{i^1k_1^1}, \ldots, \bar{v}_{j^1k_{r_1}^1}, \ldots, \bar{v}_{i^mk_1^m}, \ldots, \bar{v}_{j^mk_{r_m}^m}|), \tag{3.17}$$

the expected value being one of those associated with the two partitions (it should be mentioned that (3.17) uses the fact that $|\Sigma x_i|\leqslant n^{1/2}$). As in Lemma 1 we let (3.11) denote this expected value, assuming $r'$ distinct elements of $\bar{V}_n$ are involved, with $\mathbf{r}_\alpha$ elements appearing $\alpha$ times, $2\leqslant\alpha\leqslant 2(r_1+\cdots+r_m)$. We have $l\leqslant r'\leqslant r_1+\cdots+r_m$, $\sum_\alpha\mathbf{r}_\alpha=r'$, and $\sum_\alpha\alpha\mathbf{r}_\alpha=2(r_1+\cdots+r_m)$.

Let $d_f$ denote the number of free $b_j$ indices for which $A^j_{a_jb_j}$ involves $\bar{v}_{p^tq^t}$'s for at least two different $t$'s ($1 \leqslant t \leqslant m$). We have

$$d_f \leqslant \mathbf{r}_4 + \mathbf{r}_6 + \cdots + \mathbf{r}_{2(r_1 + \cdots + r_m)}. \tag{3.18}$$

We will show

$$d \leqslant \min(r_1 + \cdots + r_m - (m/2) - (l/2), (r_1 + \cdots + r_m - m - l + r' + d_f)/2). \tag{3.19}$$

Since (1) is avoided, the maximum number of distinct indices from $L \equiv \{i^1_2, ..., i^1_{r_1}, ..., i^m_2, ..., i^m_{r_m}\}$ $(= I - J)$ which contribute to $d$ is bounded by $r_1 + \cdots + r_m - m - l$. Suppose there is an $L$ index which, by itself, forms a class from the $I$ partition, or there is a free $b_j$ index for which $A^j_{a_jb_j}$ involves $\bar{v}_{p^tq^t}$'s for only one value of $t$. Then we have a free $a_i$ or $b_j$ (the other index appearing in another factor of (3.11)) and $A^j_{a_jb_j}$ can be removed from (3.11), resulting in an expression arising from (3.16) avoiding (1) and (2), with $r_1 + \cdots + r_m$ reduced by at least one, but with $m$ and $l$ remaining the same values. The new $d$ value is just one less than the original $d$ value. Then, without loss of generality we may assume:

(3) each $I$ class containing no indices from $J$ has at least two elements, and there are only $d_f$ free $b_j$ indices.

We immediately get $d \leqslant (r_1 + \cdots + r_m - m - l)/2 + (r' - d_f)/2 + d_f = (r_1 + \cdots + r_m - m - l + r' + d_f)/2$.

For the other expression in (3.19) we see that each distinct $b_j$ index is associated with at least four elements from $\bar{V}_n$ so that the number of distinct $b_j$ indices is bounded by $(r_1 + \cdots + r_m)/2$. Thus (3.19) follows.

If there is an $\mathbf{r}_\alpha \geqslant 1$ with $\alpha > 4$ then by (3.4) we have

$$(3.17) = o(n^{(1/2)(r_1 + \cdots + r_m) - r' + (1/2)\mathbf{r}_2 + (1/4)\mathbf{r}_3 + l/2 + d})$$

and

$$(n^{m/2}/s^{r_1 + \cdots + r_m}) \times (3.17)$$
$$= o(n^{-(1/2)(r_1 + \cdots + r_m) - r' + (1/2)\mathbf{r}_2 + (1/4)\mathbf{r}_3 + (m/2) + (l/2) + d}). \tag{3.20}$$

Using (3.18) and (3.19), the exponent of $n$ in the last expression in (3.20) is bounded by

$$-\tfrac{1}{2}r' + \tfrac{1}{2}\mathbf{r}_2 + \tfrac{1}{4}\mathbf{r}_3 + \tfrac{1}{2}(\mathbf{r}_4 + \mathbf{r}_6 + \cdots + \mathbf{r}_{2(r_1 + \cdots r_m)}) \leqslant 0$$

so that $(3.20) \to 0$ as $n \to \infty$.

If $\mathbf{r}_\alpha = 0$ for all $\alpha > 4$, then the expected value in (3.17) is bounded, and by (3.19)

$$(n^{m/2}/s^{r_1 + \cdots + r_m}) \times (3.17) = O(1). \tag{3.21}$$

The only way (3.21) will not converge to 0 is when

$$d = r_1 + \cdots + r_m - (m/2) - (l/2). \tag{3.22}$$

We will be done once we show that (3.22) implies the associated terms contributing to (3.16) depend only on $E(\bar{v}_{11}^2)$ and $E(\bar{v}_{11}^4)$.

In deriving (3.19) we first removed factors from (3.11) involving free indices. Each of these $A^j_{a_j b_j}$ must be $E(|\bar{v}_{11}|^c)$, where $c$ is even. If $c$ is ever four or larger then (3.22) will not hold. We can then assume that (3) holds. From the above we see that, in order for (3.22) to hold, each distinct $b_j$ index must be associated with *exactly* four elements from $\bar{V}_n$, either one element appearing four times or two distinct elements each appearing twice (since (1) is avoided). It follows that any of the associated terms in (3.16) will depend only on $E(\bar{v}_{11}^2)$ and $E(\bar{v}_{11}^4)$.

It remains to verify (3.14). We will show

$$E[(\sqrt{n}\,(x_n^{\mathrm{T}}(M_n'')^r x_n - E(x_n^{\mathrm{T}}(M_n'')^r x_n) - (x_n^{\mathrm{T}}\bar{M}_n^r x_n - E(x_n^{\mathrm{T}}\bar{M}_n^r x_n))))^2] \tag{3.23}$$

converges to zero as $n \to \infty$. We have

$$\begin{aligned}(s^{2r}/n) \times (3.23) = \sum_{\substack{i, j, i_2, \ldots, i_r, k_1, \ldots, k_r \\ i', j', i'_2, \ldots, i'_r, k'_1, \ldots, k'_r}} & x_i x_j x_{i'} x_{j'} E[(u_{ik_1} \cdots u_{jk_r} \\ & - E(u_{ik_1} \cdots u_{jk_r}) - (\bar{v}_{ik_1} \cdots \bar{v}_{jk_r} - E(\bar{v}_{ik_1} \cdots \bar{v}_{jk_r}))) \\ & \times (u_{i'k'_1} \cdots u_{j'k'_r} - E(u_{i'k'_1} \cdots u_{j'k'_r}) - (\bar{v}_{i'k'_1} \cdots \bar{v}_{j'k'_r} \\ & - E(\bar{v}_{i'k'_1} \cdots \bar{v}_{j'k'_r})))].\end{aligned} \tag{3.24}$$

We see that (3.24) is similar to (3.16) with $m = 2$ and $r_1 = r_2 = r$ and that much of the previous arguments carries through. A zero term will occur when

(1) $u_{pq}, \bar{v}_{pq}$ appears alone, or

(2) no $(u_{pq}, \bar{v}_{pq})$ equals a $(u_{p'q'}, \bar{v}_{p'q'})$.

We concentrate on one of the ways $\{i, j, i_2, \ldots, i_r, i', j', i'_2, \ldots, i'_r\}$ and $\{k_1, \ldots, k_r, k'_1, \ldots, k'_r\}$ can each be partitioned so that the associated terms in (3.24) avoid (1) and (2). Define $l$ and $d$ as before. The contribution to (3.24) is bounded in absolute value by

$$Cn^{(l/2)+d}E(|v_{ik_1}I_{(|v_{ik_1}|\leqslant n^{1/2})}\cdots v_{jk_r}I_{(|v_{jk_r}|\leqslant n^{1/2})}\\ \times v_{i'k'_1}I_{(|v_{i'k'_1}|\leqslant n^{1/2})}\cdots v_{j'k'_r}I_{(|v_{j'k'_r}|\leqslant n^{1/2})}|), \tag{3.25}$$

the expected value arising from one of the terms associated with the two partitions. Let (3.11) denote this expected value with $r'$ and $\mathbf{r}_\alpha$, $2\leqslant\alpha\leqslant 4r$, defined as before. We have $l\leqslant r'\leqslant 2r$, $\sum\mathbf{r}_\alpha=r'$, and $\sum\alpha\mathbf{r}_\alpha=4r$.

The bound on $d$ found in (3.19) (with $m=2$, $r_1=r_2=r$) is still valid. If $\mathbf{r}_\alpha=0$ for all $\alpha>4$, then the expected value in (3.25) is bounded. Moreover, the random variable

$$(u_{ik_1}\cdots u_{jk_r}-E(u_{ik_1}\cdots u_{jk_r})-(\bar{v}_{ik_1}\cdots\bar{v}_{jk_r}-E(\bar{v}_{ik_1}\cdots\bar{v}_{jk_r})))\\ \times(u_{i'k'_1}\cdots u_{j'k'_r}-E(u_{i'k'_1}\cdots u_{j'k'_r})-(\bar{v}_{i'k'_1}\cdots\bar{v}_{j'k'_1}-E(\bar{v}_{i'k'_1}\cdots\bar{v}_{j'k'_r})))$$

is bounded in absolute value by an integrable random variable (a constant times the absolute value of a product of $v_{ik}$'s each not appearing more than four times) and converges almost surely to zero. Therefore, by the dominated convergence theorem these terms contributing to (3.23) approach zero as $n\to\infty$.

If there is an $\mathbf{r}_\alpha\geqslant 1$ with $\alpha>4$, then a different bound on $d$ is needed. As in Lemma 1 it is straightforward to show by induction that

$$d\leqslant r'-l/2.$$

once it is shown to be true for $r'=l=4$. In this case $d$ is just the number of distinct $b_j$'s in (3.11). But each $A^i_{a_ib_i}$ must be of the form $E(|v_{11}I_{(|v_{11}|\leqslant n^{1/2})}|^c)$ where $c$ is odd, so that each $b_i$ cannot be free. Therefore, $d\leqslant 2=r'-l/2$.

Using (3.4) and our new bound on $d$ we have

$$(n/s^{2r})\times(3.25)=o(n^{2r-2r'+\mathbf{r}_2+(1/2)\mathbf{r}_3+r'+1-2r})\\ =o(n^{-r'+\mathbf{r}_2+(1/2)\mathbf{r}_3+1})\to 0$$

as $n\to\infty$, using the fact that $\mathbf{r}_\alpha\geqslant 1$ for some $\alpha>4$. Therefore, (3.23) converges to zero as $n\to\infty$ and we are done.

To verify (c) we see that because of (b) we can assume $E(v_{11})=0$ and without loss of generality we can assume $E(v_{11}^2)=1$. As in [11, p. 305] it is straightforward to expand

$$(n/2)\,E[(x_n^{\mathrm{T}}\bar{M}_nx_n-E(x_n^{\mathrm{T}}\bar{M}_nx_n))(x_n^{\mathrm{T}}\bar{M}_n^2x_n-E(x_n^{\mathrm{T}}\bar{M}_n^2x_n))] \tag{3.26}$$

and to find that (3.26) depends asymptotically on $\sum x_i^4$ if and only if $E(v_{11}^4)\neq 3$. This implies that when $E(v_{11}^4)\neq 3$ a sequence $\{x_n\}$ can be formed for which $(\sqrt{n/2}\,(x_n^{\mathrm{T}}\bar{M}_n x_n - E(x_n \bar{M}_n x_n)), \sqrt{n/2}\,(x_n^{\mathrm{T}}\bar{M}_n^2 x_n - E(x_n^{\mathrm{T}}\bar{M}_n^2 x_n)))$ will not converge in distribution (because all mixed moments have been shown above to be bounded), from which (c) follows.

## References


[1] Bai, Z. D., Silverstein, J. 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.

[2] Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.

[3] Breiman, L. (1968). *Probability*, Addison-Wesley, Reading, MA.

[4] Fan, K. (1951). Maximum properties and inequalities for eigenvalues of completely continuous operators. *Proc. Nat. Acad. Sci. USA* **37** 760–766.

[5] Geman, S. (1980). A limit theorem for the norm of random matrices, *Ann. Probab.* **8** 252–261.

[6] Gnedenko, B. V., and Kolmogorov, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, Reading, MA.

[7] Grenander, U., and Silverstein, J. W. (1977). Spectral analysis of networks with random topologies. *SIAM J. Appl. Math.* **32** 499–519.

[8] Jonsson, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** 1–38.

[9] Silverstein, J. W. (1979). On the randomness of eigenvectors generated from networks with random topologies. *SIAM J. Appl. Math.* **37** 235–245.

[10] Silverstein, J. W. (1981). Describing the behavior of eigenvectors of random matrices using sequences of measures on orthogonal groups. *SIAM J. Math. Anal.* **12** 174–281.

[11] Silverstein, J. W. (1984). Some limit theorems on the eigenvectors of large dimensional sample covariance matrices. *J. Multivariate Anal.* **15** 295–324.

[12] Wachter, K. W. (1978) The strong limits of random matrix spectra for sample matrices of independent elements. *Ann. Probab.* **6** 1–18.

[13] Yin, Y. Q. (1986). Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.* **20** 50–68.

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