JOURNAL OF MULTIVARIATE ANALYSIS 15, 295–324 (1984)

Some Limit Theorems on the Eigenvectors of Large Dimensional Sample Covariance Matrices

JACK W. SILVERSTEIN*

Department of Mathematics,
North Carolina State University, Raleigh, North Carolina 27650

Communicated by P. R. Krishnaiah

Let $\{v_{ij}\}$ $i,j=1,2,\ldots,$ be i.i.d. standardized random variables. For each n, let $V_n=(v_{ij})$ $i=1,2,\ldots,n$; $j=1,2,\ldots,s=s(n)$, where $(n/s)\to y>0$ as $n\to\infty$, and let $M_n=(1/s)\,V_nV_n^{\mathrm{T}}$. Previous results [7, 8] have shown the eigenvectors of M_n to display behavior, for n large, similar to those of the corresponding Wishart matrix. A certain stochastic process X_n on $[0, 1]$, constructed from the eigenvectors of M_n, is known to converge weakly, as $n\to\infty$, on $D[0, 1]$ to Brownian bridge when v_{11} is $N(0, 1)$, but it is not known whether this property holds for any other distribution. The present paper provides evidence that this property may hold in the non-Wishart case in the form of limit theorems on the convergence in distribution of random variables constructed from integrating analytic function w.r.t. $X_n(F_n(x))$, where F_n is the empirical distribution function of the eigenvalues of M_n. The theorems assume certain conditions on the moments of v_{11} including $E(v_{11}^4)=3$, the latter being necessary for the theorems to hold.

1. INTRODUCTION

This paper continues the investigation on the asymptotic behavior of eigenvectors of a class of sample covariance matrices defined as follows:

Let $\{v_{ij}\}$, $i, j=1,2,\ldots,$ be i.i.d. random variables such that

$$E(v_{11})=0, \qquad E(v_{11}^2)=1, \tag{1.1}$$

$$E(v_{11}^4)=3, \tag{1.2}$$

* This research was supported by the National Science Foundation under Grant MCS-8101703.

Received March 22, 1982; revised October 30, 1983.

AMS 1980 subject classifications: primary 60B15; secondary 62H99.

Key Words and phrases: Sample covariance matrices, behavior of eigenvectors, Brownian bridge, weak convergence, multidimensional method of moments.

0047-259X/84 $3.00

and either

$$\text{all moments of } v_{11} \text{ exist,} \tag{1.3a}$$

or

$$E(|v_{11}|^m) \leqslant m^{\alpha m} \qquad \text{for all} \quad m \geqslant 2 \text{ and for some } \alpha > 0. \tag{1.3b}$$

For each n let $M_n = (1/s)\, V_n V_n^{\mathrm{T}}$, where $V_n = (v_{ij})$, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, s = s(n)$ and $(n/s) \to y > 0$ as $n \to \infty$.

Previous results have shown similarity of behavior between the eigenvectors of M_n for n large and vectors generated by the Haar measure h_n on $\mathcal{O}_n$, the group of $n \times n$ orthogonal matrices [7, 8]. To be more specific, consider the spectral decomposition $O_n \Lambda_n O_n^{\mathrm{T}}$ of M_n, where the eigenvalues of M_n are arranged along the diagonal of Λ_n in nondecreasing order. Although ambituities do exist in constructing $O_n \in \mathcal{O}_n$ from M_n we may as well assume the distribution of O_n to be uniform whenever ambituities arise. For example, O_n will have the same distribution as $O_n J$ for each diagonal matrix J containing ± 1's along its diagonal. Thus, we may consider a unique Borel probability measure ω_n on $\mathcal{O}_n$ induced from M_n.

The previous results suggest that for n large ω_n and h_n are somehow close. They developed from investigating whether or not the sequence $\{\omega_n\}$ satisfies certain properties or arbitrary sequences $\{\gamma_n\}$, (where for each n, γ_n is a Borel probability measure on $\mathcal{O}_n$), which are known to be satisfied by $\{h_n\}$. One of three such properties on $\{\gamma_n\}$ has, except for v_{11} being $N(0, 1)$, as yet not been shown to be satisfied by $\{\omega_n\}$ under any other assumptions on v_{11}. It is stated below:

The sequence $\{\gamma_n\}$ satisfies property III if for every sequence $\{x_n\}$, $x_n \in \mathbb{R}^n$, of nonrandom unit vectors, if $(y_1, y_2, \ldots, y_n)^{\mathrm{T}} = O_n^{\mathrm{T}} x_n$, where O_n is γ_n distributed, and if $X_n\colon [0, 1] \to \mathbb{R}$ is defined as

$$X_n(t) = \frac{\sqrt{n}}{\sqrt{2}} \left(\sum_{i=1}^{[nt]} y_i^2 - \frac{[nt]}{n} \right) \qquad ([s] \equiv \text{greatest integer} \leqslant s),$$

then $X_n \xrightarrow{\mathscr{D}} W^0$ as $n \to \infty$, where W^0 is Brownian bridge and $\mathscr{D}$ denotes weak convergence of random elements in $D[0, 1]$ [3].

The reason $\{h_n\}$ satisfies III follows from the distribution of $O_n^{\mathrm{T}} x_n$, when O_n is h_n distributed, being the same as the distribution of a normalized vector of i.i.d. mean zero Gaussian components. The exceptions mentioned above follows from the fact that when v_{11} is $N(0, 1)$, M_n is the Wishart matrix, $W(I_n, s)$, in which case $\omega_n = h_n$.

At this point it is necessary to state two results on the limiting behavior of the *eigenvalues* of M_n defined under weaker conditions. First, for M_n defined as above except replacing (1.2) and (1.3) by the existence of a $\delta > 0$ such that $E(|v_{11}|^{2+\delta}) < \infty$, it is known [5, 9] that the empirical distribution function of the eigenvalues of M_n (i.e., $F_n(x) = 1/n$ (# of eigenvalues $\leqslant x$))

converges almost surely for every $x \geqslant 0$ as $n \to \infty$ to a nonrandom distribution function $F_y(x)$, where for $0 \leqslant y \leqslant 1$, $F_y(x)$ has density

$$f_y(x) = \begin{cases} \dfrac{\sqrt{(x-(1-\sqrt{y})^2)((1+\sqrt{y})^2 - x)}}{2\pi y x} \\ \qquad \text{for} \quad (1-\sqrt{y})^2 < x < (1+\sqrt{y})^2, \\ 0 \qquad \text{otherwise,} \end{cases} \tag{1.4}$$

and for $1 < y < \infty$

$$F_y(x) = \begin{cases} 1 - 1/y & \text{for} \quad 0 \leqslant x \leqslant (1-\sqrt{y})^2, \\ 1 - 1/y + \displaystyle\int_{(1-\sqrt{y})^2}^{x} f_y(t)\,dt & \text{for} \quad (1-\sqrt{y})^2 \leqslant x \leqslant (1+\sqrt{y})^2, \\ 1 & \text{for} \quad x \geqslant (1+\sqrt{y})^2. \end{cases} \tag{1.5}$$

The second result concerns the limiting behavior of $\lambda_{\max}^{(n)}$, the largest eigenvalue of M_n. Under the assumptions $\{v_{ij}\}$ i.i.d., (1.1), and (1.3b) it is shown in [4] that

$$\lambda_{\max}^{(n)} \xrightarrow{\text{a.s.}} (1+\sqrt{y})^2 \qquad \text{as} \quad n \to \infty. \tag{1.6}$$

In proving (1.6) it is shown that if $z > (1+\sqrt{y})^2$ and w satisfies

$$w > \max((3/\ln(z/(1+\sqrt{y})^2)), 5), \tag{1.7}$$

then for all n sufficiently large

$$E((\lambda_{\max}^{(n)}/z)^{[w \ln(n)]}) = O(n^{-2}). \tag{1.8}$$

(Note that in [4], $s \to \infty$ while $n = n(s)$, and (1.8) is proven without any change in notation. However, it is not difficult to verify (1.8) for $n \to \infty$ and $s = s(n)$.)

Returning to eigenvectors and assumptions (1.1), (1.2), (1.3b) on v_{11} we make the following observation. Suppose $\{\omega_n\}$ satisfies III. Let $\{f_i\}_{i=1}^{\infty}$ be an arbitrary sequence of absolutely continuous functions defined on $[0, \infty)$, and such that $f_i(0) = 0$ for all i. If $\{\omega_n\}$ satisfies III, then it follows from (1.6) and the theory of weak convergence of measures on function spaces that for any sequence $\{x_n\}$ of unit vectors:

$$\left\{\int_0^\infty f_i(x)\,dX_n(F_n(x))\right\}_{i=1}^{\infty}$$

$$= \left\{\frac{\sqrt{n}}{\sqrt{2}}\left(x_n^{\mathrm{T}} f_i(M_n) x_n - \frac{1}{n}\operatorname{tr} f_i(M_n)\right)\right\}_{i=1}^{\infty} \xrightarrow{\mathscr{D}} (\text{as } n \to \infty) \tag{1.9}$$

$$\left\{\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f_i(x)\,dW_x^y\right\}_{i=1}^{\infty}, \tag{1.10}$$

where $W_x^y \equiv W^0_{F_y(x)}$, $f_i(M_n)$ is the matrix derived from M_n after applying f_i to each of its eigenvalues in the spectral decomposition of M_n, and $\mathscr{D}$ denotes convergence in distribution on $\mathbb{R}^\infty$ [8] (the condition $f_i(0)=0$ is assumed merely for convenience since $\int_0^\infty f(x)\,dX_n(F_n(x))=0$ for $f(x)=\text{const.}$). The limiting random variables (1.10) are well-defined stochastic integrals. They are jointly normal, each having mean zero, and for any i, j it can be shown that

$$E\left(\left(\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f_i(x)\,dW_x^y\right)\left(\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f_j(x)\,dW_x^y\right)\right)$$
$$=\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f_i(x)\,f_j(x)\,f_y(x)\,dx$$
$$-\left(\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f_i(x)\,f_y(x)\,dx\right)\left(\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f_j(x)\,f_y(x)\,dx\right). \quad (1.11)$$

Considering $f(x)=x$ as one of the f_i's, it is shown in [8] that, assuming (1.1) and (1.3b), (1.2) is necessary for III to hold.

The aim of this paper is to verify $(1.9)\to^{\mathscr{D}}(1.10)$ for f_i analytic. In Section 2 we will prove

THEOREM 1. *Let $\{x_n\}$, $x_n\in\mathbb{R}^n$, be an arbitrary sequence of nonrandom unit vectors. Then, under assumptions* (1.1), (1.2), *and* (1.3a), $(1.9)\to^{\mathscr{D}}(1.10)$ *as $n\to\infty$ for $f_i(x)=x^i$. Moreover, if* (1.2) *does not hold, then there exists sequences $\{x_n\}$ for which* (1.9) (*with $f_i(x)=x^i$*) *fails to converge in distribution.*

The proof of Theorem 1 uses a multidimensional method of moments. In particular, it is shown for any positive integers i and j

$$E\left(\int_0^\infty x^i\,dX_n(F_n(x))\int_0^\infty x^j\,dX_n(F_n(x))\right)\to(1.11)$$
$$\text{as } n\to\infty \text{ with } f_i=x^i, f_j=x^j. \quad (1.12)$$

Using Theorem 1 we will prove in Sections 3 and 4.

THEOREM 2. *Let $\{f_i\}_{i=1}^\infty$ be functions defined on $[0,\infty)$, where for each i, $f_i(0)=0$, and f_i is analytic at* 0, *with corresponding radius of convergence greater than $(1+\sqrt{y})^2$. Let $\{x_n\}$, $x_n\in\mathbb{R}^n$, be an arbitrary sequence of nonrandom unit vectors. Then, under assumptions* (1.1), (1.2), *and* (1.3b), $(1.9)\to^{\mathscr{D}}(1.10)$ *as $n\to\infty$.*

The proof of Theorem 2 is mainly contained in Section 3. Essentially, the problem is transformed into showing convergence of functions on an appropriate L^2 space. Two lemmas are required. The second one, deriving a uniform growth condition on $E((\int_0^\infty x^r\, dX_n(F_n(x)))^2)$, has a long proof. The proof will be given in Section 4.

It is believed this lemma may be more fully exploited in proving, at the very least, other consequences of III applied to M_n. It is also believed that Theorem 2 will hold under much weaker conditions on the higher moments of v_{11}. To this end the truncation method introduced in [1] and used in [5] on sample covariance matrices may prove useful.

The result in [8] raises doubts to the intuitive notion that all large dimensional sample covariance matrices of sample vectors having i.i.d. mean zero components should have the distribution of their eigenvectors close to being Haar distributed. It shows the distribution of v_{11} needs to be similar to $N(0, 1)$ in order for III to hold. Theorems 1 and 2, however, show we get some interesting behavior similar to Haar measure once $E(v_{11}^4) = 3$. It may be that this is all that is needed to satisfy III, or it may be that a closer relationship to $N(0, 1)$ is necessary.

The conclusions of Theorems 1 and 2 are not sufficient to verify III. At the very least, the distributions of $\int_0^\infty f_i(x)\, dX_n(F_n(x))$ need to converge *uniformly* over certain classes of f_i [2, 6]. It is believed, though, that an investigation of uniformity will lead to a significantly clearer understanding of precisely what is required to verify III.

2. Proof of Theorem 1

The proof requires two lemmas:

Lemma 2.1. *For any integer $r \geq 1$, $(1/\sqrt{n})\,(\operatorname{tr} M_n^r - E(\operatorname{tr} M_n^r)) \to^{\text{i.p.}} 0$ as $n \to \infty$.*

Proof. We will show for any integers $r_1, r_2 \geq 1$, $(1/n)\operatorname{Cov}(\operatorname{tr} M_n^{r_1}, \operatorname{tr} M_n^{r_2}) \to 0$ as $n \to \infty$. We have

$$
\begin{aligned}
&s^{r_1+r_2} \operatorname{Cov}(\operatorname{tr} M_n^{r_1}, \operatorname{tr} M_n^{r_2}) \\
&= \sum_{\substack{i_1 \cdots i_{r_1}, k_1 \cdots k_{r_1} \\ i'_1 \cdots i'_{r_2}, k'_1 \cdots k'_{r_2}}} E[(v_{i_1k_1} v_{i_2k_1} v_{i_2k_2} \cdots v_{i_{r_1}k_{r_1}} v_{i_1k_{r_1}} \\
&\qquad - E(v_{i_1k_1} v_{i_2k_1} v_{i_2k_2} \cdots v_{i_{r_1}k_{r_1}} v_{i_1k_{r_1}})) \\
&\qquad \times (v_{i'_1k'_1} v_{i'_2k'_1} v_{i'_2k'_2} \cdots v_{i'_{r_2}k'_{r_2}} v_{i'_1k'_{r_2}} \\
&\qquad - E(v_{i'_1k'_1} v_{i'_2k'_1} v_{i'_2k'_2} \cdots v_{i'_{r_2}k'_{r_2}} v_{i'_1k'_{r_2}}))] \\
&= \sum_{\substack{i_1 \cdots i_{r_1}, k_1 \cdots k_{r_1} \\ i'_1 \cdots i'_{r_2}, k'_1 \cdots k'_{r_2}}} E(v_{i_1k_1} \cdots v_{i_1k_{r_1}} v_{i'_1k'_1} \cdots v_{i'_1k'_{r_2}}) \\
&\qquad - E(v_{i_1k_1} \cdots v_{i_1k_{r_1}})\, E(v_{i'_1k'_1} \cdots v_{i'_1k'_{r_2}}).
\end{aligned} \tag{2.1}
$$

Notice a zero term can occur in at least two ways:

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

(2) no v_{ik} equals a $v_{i'k'}$.

Let us divide up the sum into a finite number of sums, independent of n (for n sufficiently large), each one grouping the v_{ik}'s and $v_{i'k'}$'s due to a particular grouping of the indices. Consider one of these sums which avoids (1) and (2). For this sum constraints are placed on some of the indices. For example, i_1 may be constrained to i_3', or to both i_2 and i_5', or it may be free, that is, not constrained to any other i or i'. We have for this sum

$$\sum_{\substack{a_1 \cdots a_{r'} \\ b_1 \cdots b_{r'}}} A^1_{a_1b_1} A^2_{a_2b_2} \cdots A^{r'}_{a_{r'}b_{r'}} - B^1_{a_1b_1} B^2_{a_2b_2} \cdots B^{r'}_{a_{r'}b_{r'}}, \tag{2.2}$$

where the a's correspond to the i's and i''s, the b's to the k's and k''s, and, because 1 is not satisfied, $r' \leqslant r_1 + r_2$. Also, the $A^l_{a_lb_l}$'s and $B^l_{a_lb_l}$'s correspond, respectively, to the first and second terms in the last expression of (2.1). The sum will not include a term where any (a_i, b_i) equals another (a_j, b_j). Constraints still remain on some of the a_i's and some of the b_i's. In fact, because of the weaving pattern of the i's, i''s, k's, and k''s, and because (2) is avoided, if $r' > 1$, then for every (a_i, b_i), either a_i or b_i is constrained.

It is clear that (2.2) is a polynomial $p(n, s)$ in n and s having coefficients bounded for all n, with total degree not exceeding the number of classes in the partition of $\{a_1, ..., a_{r'}, b_1, ..., b_{r'}\}$ imposed by the constraints.

We have three cases:

Case I. $r' = 1$. Then $\deg(p(n, s)) \leqslant 2$ and $(1/ns^{r_1+r_2}) \cdot (2.2) \to 0$ as $n \to \infty$.

Case II. $r' > 1$ and no a_i or b_i is free. Then $\deg(p(n, s)) \leqslant r'$ and again $(1/ns^{r_1+r_2}) \cdot (2, 2) \to 0$ as $n \to \infty$.

Case III. $r' > 1$ and there is at least one a_i or b_i free. Assume without loss of generality b_l is free. Then $A^l_{a_lb_l}$ and $B^l_{a_lb_l}$ must be formed from pairs of adjacent v_{ik}'s and (or) adjacent $v_{i'k'}$'s. For example, if $v_{i_3k_3}$ is involved, then so must $v_{i_4k_3}$. Moreover, i_4 would be constrained to i_3. By summing on b_l we find that (2.2) is asymptotically proportional to

$$n \cdot \sum_{\substack{a_1 \cdots a_{l-1} a_{l+1} \cdots a_{r'} \\ b_1 \cdots b_{l-1} b_{l+1} \cdots b_{r'}}} A^1_{a_1b_1} \cdots A^{r'}_{a_{r'}b_{r'}} - B^1_{a_1b_1} \cdots B^{r'}_{a_{r'}b_{r'}}, \tag{2.3}$$

where the sum is one of those sums arising from (2.1) having a new r_1 and r_2, whose new $r_1 + r_2$ is less than the original $r_1 + r_2$. For this new sum 1 is clearly avoided. If the new sum has no v_{ik} paired with a $v_{i'k'}$, or if all v_{ik}'s or all $v_{i'k'}$'s have been removed (that is, the new r_1 or $r_2 = 0$) then (2.3) is clearly 0.

If the latter does not occur we repeat Case III until it does or until we arrive at Case I or Case II.

LEMMA 2.2. *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. We have

$$s^r \left(E(x_n^{\mathrm{T}} M_n^r x_n) - E\left(\frac{1}{n} \operatorname{tr} M_n^r\right)\right) \tag{2.4}$$
$$= \sum_{\substack{ij \\ i_2 \cdots i_r \\ k_1 \cdots k_r}} x_i x_j E(v_{ik_1} v_{i_2 k_1} \cdots v_{i_r k_r} v_{jk_r}) - \frac{1}{n} \sum_{\substack{i_1 \cdots i_r \\ k_1 \cdots k_r}} E(v_{i_1 k_1} \cdots v_{i_1 k_r}).$$

It is clear that

$$\sum_{\substack{i_2 \cdots i_r \\ k_1 \cdots k_r}} v_{ik_1} \cdots v_{ik_r}, \qquad i = 1, 2, ..., n \tag{2.5}$$

are identically distributed and that

$$\sum_{\substack{i_2 \cdots i_r \\ k_1 \cdots k_r}} v_{ik_1} \cdots v_{jk_r}, \qquad i \neq j \tag{2.6}$$

are identically distributed.

Therefore

$$s^r \left(E(x_n^{\mathrm{T}} M_n^r x_n) - E\left(\frac{1}{n} \operatorname{tr} M_n^r\right)\right) = \sum_{i \neq j} x_i x_j \sum_{\substack{i_2 \cdots i_r \\ k_1 \cdots k_r}} E(v_{ik_1} \cdots v_{jk_r}). \tag{2.7}$$

As in Lemma 1 we get a zero term whenever a v_{ab} is alone. For fixed $i \neq j$ let

$$\sum_{\substack{a_1 \cdots a_{r'} \\ b_1 \cdots b_{r'}}} A^1_{a_1 b_1} \cdots A^{r'}_{a_{r'} b_{r'}} \tag{2.8}$$

be one of the groupings where no v_{ab} is alone, and where the sum is arranged so that a_1 incorporates i and a_2 incorporates j. We claim that $r' < r$. If not then (2.8) involves exact pairings of the v_{ik}'s, so v_{ik_1} must be paired with some $v_{\underline{i}k}$ and the companion to $v_{\underline{i}k}$, say $v_{\underline{\underline{i}}k}$, must be paired to another $v_{\underline{\underline{\underline{i}}}k}$, and so on. We see that all the i_l's must be constrained together, but this cannot happen since j is constrained to an i_l and $i \neq j$.

As before, (2.8) is a polynomial $p(n, s)$ in n and s and for each l a_l and b_l cannot both be free.

We have two cases:

Case I. No a_l for $l = 3, ..., r'$ or b_l is free. Then $\deg(p(n, s)) \leqslant (r' - 2 + r')/2 = r' - 1 \leqslant r - 2$ and since $|\sum_{i \neq j} x_i x_j| \leqslant n - 1$,

$$\frac{\sqrt{n}}{s^r}\left(\sum_{i \neq j} x_i x_j\right) \cdot (2.8) = \frac{\sqrt{n}}{s^r}\, p(n, s)\left(\sum_{i \neq j} x_i x_j\right) \to 0 \qquad \text{as} \quad n \to \infty. \tag{2.9}$$

Case II. At least one a_l for $l = 3, ..., r'$ or b_l is free. By summing on one of the free indices we find, as in Case III of Lemma 1, that (2.8) is asymptotically proportional to n times another sum which arises from (2.7) having a smaller r. This procedure is repeated until Case I is eventually reached.

To finish the proof of Theorem 1 we see that from Lemmas 2.1 and 2.2 it is sufficient to show

$$\left\{\frac{\sqrt{n}}{\sqrt{2}}\,(x_n^{\mathrm{T}} M_n^i x_n - E(x_n^{\mathrm{T}} M_n^i x_n))\right\}_{i=1}^{\infty} \xrightarrow{\mathscr{D}} \left\{\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^i \, dW_x^y\right\}_{i=1}^{\infty} \qquad \text{as} \quad n \to \infty. \tag{2.10}$$

Using a multidimensional version of the method of moments we will show that all mixed moments are bounded and that any asymptotic behavior depends solely on $E(v_{11})$, $E(v_{11}^2)$, and $E(v_{11}^4)$. We know that Theorem 1 is true when v_{11} is $N(0, 1)$ and because of Lemmas 2.1 and 2.2, (2.10) holds also. Bounded mixed moments will imply, when v_{11} is $N(0, 1)$, the mixed moments of (2.10) converge to the proper values. The dependence of the limiting behavior of the mixed moments on $E(v_{11})$, $E(v_{11}^2)$, and $E(v_{11}^4)$ imply the moments is general will converge to the same values. The fact that a multivariate normal distribution is uniquely determined by its moments will then imply (2.10).

Let $m \geqslant 2$ and $r_1, r_2, ..., r_m$ be arbitrary positive integers. Consider

$$\begin{aligned} & s^{r_1 + \cdots + r_m} E[(x_n^{\mathrm{T}} M_n^{r_1} x_n - E(x_n^{\mathrm{T}} M_n^{r_1} x_n)) \cdots (x_n^{\mathrm{T}} M_n^{r_m} x_n - E(x_n^{\mathrm{T}} M_n^{r_m} x_n))] \\ &= s^{r_1 + \cdots + r_m} E \sum_{\substack{i^1 j^1 \\ \vdots \\ i^m j^m}} x_{i^1} x_{j^1} \cdots x_{i^m j^m} ((M_n^{r_1})_{i^1 j^1} - E((M_n^{r_1})_{i^1 j^1})) \\ & \quad \cdots ((M_n^{r_m})_{i^m j^m} - E((M_n^{r_m})_{i^m j^m})). \end{aligned} \tag{2.11}$$

We can divide this sum into a number of terms, independent of n, each one being a particular grouping of the x_i's and x_j's. Consider one such sum.

Let l, $0 \leqslant l \leqslant 2m$, be the number of free x_i's and (or) x_j's. We allow an x_i to be matched up with an x_j. It is easy to see that the sum on the x_i's and x_j's alone in the grouping is bounded in absolute value by $n^{l/2}$. Expanding further, we have for fixed i's and j's

$$E \sum_{\substack{i_2^1 \cdots i_{r_1}^1 k_1^1 \cdots k_{r_1}^1 \\ \vdots \\ i_2^m \cdots i_{r_m}^m, k_1^m \cdots k_{r_m}^m}} (v_{i^1 k_1^1} \cdots v_{j^1 k_{r_1}^1} - E(v_{i^1 k_1^1} \cdots v_{j^1 k_{r_1}^1}))$$
$$\cdots (v_{i^m k_1^m} \cdots v_{j^m k_{r_m}^m} - E(v_{i^m k_1^m} \cdots v_{j^m k_{r_m}^m})). \tag{2.12}$$

Note that before the expected value is taken in (2.12) the random variables, in the particular grouping of the x_i's and x_j's, are identically distributed. As in Lemma 1 a zero term in (2.12) can occur when

(1) a v_{ik} appears alone,

(2) for a given t no $v_{p_t q_t}$ equals a $v_{p_{t'} q_{t'}}$, $t' \neq t$.

Consider one way of grouping the v_{pq}'s so that (1) and (2) are avoided, and consider one of the 2^m terms gotten from expanding (2.12) along the expressions in parentheses. We have for this sum

$$\sum A^1_{a_1 b_1} \cdots A^{r'}_{a_{r'} b_{r'}}, \tag{2.13}$$

where no term is included for which any (a_i, b_i) equals another (a_j, b_j). Let us compute the maximum of indices involved in this summation. From (2.12) we see that this number cannot exceed $2(r_1 + \cdots + r_m) - m$. However, for each free i or j, there must be an i_p^t constrained to it, which reduces this number by one. Since there are l free i's and (or) j's, we find the maximum number of running indices in (2.13) to be

$$2(r_1 + \cdots + r_m) - m - l. \tag{2.14}$$

Again, we have that (2.13) is a polynomial $p(n, s)$ in n and s having coefficients bounded for all n, with total degree not exceeding the number of classes in the partition of the running indices in (2.13) imposed by the constraints. This degree cannot exceed (2.14).

We have three cases:

Case I. None of the running indices are free. Then, because 1 does not hold, it follows that every running index i_t and k_t must be constrained to at least one other running index. From (2.14) we find $\deg(p(n, s)) \leqslant$

$(r_1 + \cdots + r_m) - (m/2) - (l/2)$ and is strictly less if m or l is odd. We therefore have that

$$n^{m/2} \frac{1}{s^{r_1 + \cdots + r_m}} \sum x_{i1} x_{j1} \cdots x_{im} x_{jm} \, p(n, s) \tag{2.15}$$

(the sum being the fixed grouping of the x_i's and x_j's considered) either goes to zero or stays bounded as $n \to \infty$. We will have (2.15) converging to zero if three or more of the k_p^t's are constrained together. This will occur if a fixed v_{pq} appears three times or more than four times in an A_{ab}^t. We conclude that in Case I the asymptotic behavior depends only on $E(v_{11})$, $E(v_{11}^2)$ and $E(v_{11}^4)$.

Case II. $r' = 1$. Then $\deg(p(n, s)) = 1$ and we find that (2.15) is asymptotically proportional to

$$\sum x_i^{2m} \frac{n^{m/2} n}{n^{r_1 + \cdots + r_m}} \tag{2.16}$$

which does not go to zero only when $m = 2$, and $r_1 = r_2 = 1$. In this case (2.16) is bounded and $p(n, s)$ is either $s \cdot E(v_{11}^4)$ or $s \cdot [E(v_{11}^2)]^2$.

Case III. $r' > 1$, and at least one of the running indices is free. As in Case III of Lemma 1 and Case II of Lemma 2 we can sum on a free index and arrive at another sum arising from (2.12) having a reduced $r_1 + \cdots + r_m$ and possibly a smaller m. The latter occurs when a b_t is summed and the corresponding a_t is not constrained to any other $a_{\underline{t}}$. This can only happen if whenever an index from the νth factor in any term from (2.12) is involved in a_t, then $v_{i^\nu k_1^\nu}$, $v_{j^\nu k_{r_\nu}^\nu}$, and all the $v_{l_\alpha^\nu k_\beta^\nu}$ are contained in $A_{a_t b_t}^t$. Because (2) is avoided there must be at least two different ν's in this grouping. Let m' be the number of different ν's and assume without loss of generality the first m' factors in (2.12) are involved in the grouping. After summing on b_t we find that (2.15) is asymptotically proportional to (2.16) (with m replaced by m') times a new (2.15) having a new m equal to $m - m' \geqslant 1$. Therefore, either (2.16) goes to zero or is bounded and $A_{a_t b_t}^t$ is either $E(v_{11}^4)$ or $[E(v_{11}^2)]^2$.

If the free index is summed with its mate constrained to another index then the new sum will have the same m and the ratio of the old (2.15) to the new (2.15) will not go to zero if and only if $A_{a_t b_t}^t = E(v_{11}^2)$.

Case III is repeated until Case I or Case II is reached or until $m = 1$ in which case the sum is zero. We conclude that the mixed moments are bounded and the limiting behavior depends only on $E(v_{11})$, $E(v_{11}^2)$ and $E(v_{11}^4)$.

This completes the proof of the main part of Theorem 1.

To see what happens when $E(v_{11}^4) \neq 3$, consider $m=2$, $r_1=1$, $r_2=2$. After considering all possible groupings of the indices we arrive at

$$E\left(\left(\frac{\sqrt{n}}{\sqrt{2}}\left(x_n^{\mathrm{T}}M_nx_n - E(x_n^{\mathrm{T}}M_nx_n)\right)\right)\left(\frac{\sqrt{n}}{\sqrt{2}}\left(x_n^{\mathrm{T}}M_n^2x_n - E(x_n^{\mathrm{T}}M_n^2x_n)\right)\right)\right)$$

$$\begin{aligned}&\sim\left(\sum_{i\neq j} x_i^2x_j^2\right)(2y+y^2)+\left(\sum_i x_i^4\right)\left((E(v_{11}^4)-1)y\right.\\&\quad\left.+\frac{1}{2}(E(v_{11}^4)-1)y^2\right)=(2y+y^2)+\left(\sum_i x_i^4\right)\left((E(v_{11}^4)-1)y\right.\\&\quad\left.+\frac{1}{2}(E(v_{11}^4)-1)y^2-(2y+y^2)\right).\end{aligned}\tag{2.17}$$

The coefficient of $(\sum_i x_i^4)$ is zero if and only if $E(v_{11}^4)=3$. If $E(v_{11}^4)\neq 3$ then, since $\sum_i x_i^4$ can range between $1/n$ and 1, sequences $\{x_n\}$ can be formed where (2.17) will not converge. For these sequences $\{(\sqrt{n}/\sqrt{2})(x_n^{\mathrm{T}}M_n^rx_n-(1/n)\operatorname{tr} M_n^r)\}_{r=1}^{\infty}$ will not converge in distribution, since the above proof shows the mixed moments are bounded for any value of $E(v_{11}^4)$.

3. Proof of Theorem 2

Before proceeding we make the remark that $X_n(F_n(x))=(\sqrt{n}/\sqrt{2})(G_n(x)-F_n(x))$, where $G_n(x)$ is the (random) probability distribution function placing mass y_i^2 at the ith smallest eigenvalue of M_n, $i=1,2,\ldots,n$.

Lemma 3.1. *Let a and b be such that $(1+\sqrt{y})^2<b<a$. Let w satisfy (1.7) for $z=b$. Then, for all positive integers $i\leqslant 2[w\ln(n)]$ and for any probability distribution function G assigning random or nonrandom mass to the eigenvalues of M_n*

$$E\left(\int_b^{\infty} x^i dG(x)\right)\leqslant\frac{C_0a^i}{n^2}\quad \textit{for all } n \textit{ sufficiently large,}\tag{3.1}$$

where the size of n and the constant C_0 do not depend upon i. (Note that throughout the following $\int_0^b$ and $\int_b^{\infty}$ denote, respectively, integration on $[0,b]$ and (b,∞).)

Proof. Let I_j denote the indicator function on the event $\{b+(a-b)j<\lambda(n)_{\max}\leqslant b+(a-b)(j+1)\}$, $j=0,1,2,\ldots$. Then for $n\geqslant 2$ we have

$$
\begin{aligned}
E\left(\int_b^\infty x^i\,dG(x)\right) &= E\left(\sum_{j=0}^\infty I_j\int_b^\infty x^i\,dG(x)\right) = \sum_{j=0}^\infty E\left(I_j\int_b^\infty x^i\,dG(x)\right)\\
&\leqslant \sum_{j=0}^\infty (b+(a-b)(j+1))^i P(\lambda_{\max}^{(n)} > b+(a-b)j)\\
&\leqslant E(\lambda(n)_{\max}^{[3w\ln(n)]})\sum_{j=0}^\infty \frac{(b+(a-b)(j+1))^i}{(b+(a-b)j)^{[3w\ln(n)]}}\\
&\leqslant E(\lambda(n)_{\max}^{[3w\ln(n)]})(a/b)^i\left(\frac{1}{b^{[3w\ln(n)]-i}}\right.\\
&\quad\left.+\frac{1}{(a-b)}\int_b^\infty \frac{1}{x^{[3w\ln(n)]-i}}\,dx\right)\\
&= E(\lambda(n)_{\max}^{[3w\ln(n)]})(a/b)^i\left(\frac{1}{b^{[3w\ln(n)]-i}}\right.\\
&\quad\left.+\frac{1}{(a-b)}\frac{1}{([3w\ln(n)]-i-1)}\frac{1}{b^{[3w\ln(n)]-i-1}}\right)\\
&= \left(1+\frac{b}{(a-b)}\frac{1}{([3w\ln(n)]-i-1)}\right)\frac{E(\lambda(n)_{\max}^{[3w\ln(n)]})}{b^{[3w\ln(n)]}}\,a^i.
\end{aligned}
\tag{3.2}
$$

Therefore, because of (1.8) we get (3.1).

COROLLARY 3.1. *For any positive integers i and j*

$$
\begin{aligned}
&E\left(\int_0^b x^i\,dX_n(F_n(x))\int_0^b x^j\,dX_n(F_n(x))\right)\\
&\quad\to E\left(\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^i\,dW_x^y\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^j\,dW_x^y\right)\qquad\text{as}\quad n\to\infty.
\end{aligned}
\tag{3.3}
$$

Proof. We have

$$
\begin{aligned}
&\left|E\left(\int_0^b x^i\,dX_n(F_n(x))\int_0^b x^j\,dX_n(F_n(x))\right)\right.\\
&\qquad\left.-E\left(\int_0^\infty x^i\,dX_n(F_n(x))\int_0^\infty x^j\,dX_n(F_n(x))\right)\right|\\
&\leqslant\left(E\left(\left(\int_b^\infty x^i\,dX_n(F_n(x))\right)^2\right)\right)^{1/2}\left(E\left(\left(\int_0^\infty x^j\,dX_n(F_n(x))\right)^2\right)\right)^{1/2}\\
&\quad+\left(E\left(\left(\int_0^\infty x^i\,dX_n(F_n(x))\right)^2\right)\right)^{1/2}\left(E\left(\left(\int_b^\infty x^j\,dX_n(F_n(x))\right)^2\right)\right)^{1/2}\\
&\quad+\left(E\left(\left(\int_b^\infty x^i\,dX_n(F_n(x))\right)^2\right)\right)^{1/2}\left(E\left(\left(\int_b^\infty x^j\,dX_n(F_n(x))\right)^2\right)\right)^{1/2}.
\end{aligned}
\tag{3.4}
$$

By Lemma 3.1 we have for any positive integer r

$$E\left(\left(\int_b^\infty x^r\,dX_n(F_n(x))\right)^2\right)$$

$$\leqslant n\left(E\left(\left(\int_b^\infty x^r\,dG_n(x)\right)^2\right)+E\left(\left(\int_b^\infty x^r\,dF_n(x)\right)^2\right)\right)$$

$$\leqslant n\left(E\left(\int_b^\infty x^{2r}\,dG_n(x)\right)+E\left(\int_b^\infty x^{2r}\,dF_n(x)\right)\right)\leqslant 2C_0\frac{a^{2r}}{n} \quad (3.5)$$

for all n sufficiently large. Therefore, from (1.12) we get (3.3).

LEMMA 3.2. *For any $w>0$ we have for all n sufficiently large,*

$$E\left(\left(\int_0^\infty x^r\,dX_n(F_n(x))\right)^2\right)\leqslant k_1\left(1+\left(\frac{n}{s}\right)^{1/2}\right)^{4r} r^{k_2} \text{ for all positive}$$

integers $r\leqslant[w\ln(n)]$, where k_1 and k_2 do not depend upon r. (3.6)

(Proof given in Section 4.)

COROLLARY 3.2. *Let a, b, and w be as in Lemma 3.1. Then for all n sufficiently large*

$$E\left(\left(\int_0^b x^r\,dX_n(F_n(x))\right)^2\right)\leqslant \underline{k}_1 a^{2r} r^{\underline{k}_2} \text{ for all positive integers}$$

$r\leqslant[w\ln(n)]$, where $\underline{k}_1$ and $\underline{k}_2$ do not depend upon r. (3.7)

Proof. Because of Lemma 3.1 we see that (3.5) holds for all n sufficiently large and for all $r\leqslant[w\ln(n)]$. Using this together with (3.6) we have for all n sufficiently large and for all $r\leqslant[w\ln(n)]$,

$$E\left(\left(\int_0^b x^r\,dX_n(F_n(x))\right)^2\right)$$

$$\leqslant 2\left(E\left(\left(\int_0^\infty x^r\,dX_n(F_n(x))\right)^2\right)+E\left(\left(\int_b^\infty x^r\,dX_n(F_n(x))\right)^2\right)\right)$$

$$\leqslant 2\left(k_1\left(1+\left(\frac{n}{s}\right)^{1/2}\right)^{4r} r^{k_2}+2C_0\frac{a^{2r}}{n}\right). \quad (3.8)$$

Since $\lim_{n\to\infty}(1+(n/s)^{1/2})^2=(1+\sqrt{y})^2<a$, it is clear that constants $\underline{k}_1$ and $\underline{k}_2$ can be found for which (3.7) holds.

We proceed now with the proof of the theorem.

Let $d \geqslant 1$ be an arbitrary integer. Let b be such that $(1+\sqrt{y})^2 < b$ and is less than the smallest radius of convergence at $x=0$ among the first d f_i's. Since (1.6) holds we have

$$\int_0^\infty f_i(x)\,dX_n(F_n(x)) - \int_0^b f_i(x)\,dX_n(F_n(x)) \xrightarrow{\text{a.s.}} 0 \qquad \text{as} \quad n \to \infty \quad (3.9)$$

for all i. Therefore it is sufficient to show for any d

$$\left\{\int_0^b f_i(x)\,dX_n(F_n(x))\right\}_{i=1}^d \xrightarrow{\mathscr{D}} \left\{\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f_i(x)\,dW_x^y\right\}_{i=1}^d \qquad \text{as} \quad n \to \infty. \quad (3.10)$$

Because of Theorem 1 and (3.9) we know that

$$\left\{\int_0^b x^j\,dX_n(F_n(x))\right\}_{j=1}^\infty \xrightarrow{\mathscr{D}} \left\{\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^j\,dW_x^y\right\}_{j=1}^\infty \qquad \text{as} \quad n \to \infty. \quad (3.11)$$

Let L_n^2 be the Hilbert space of square integrable random variables spanned by $\{\int_0^b x^j\,dX_n(F_n(x))\}_{j=1}^\infty$, and, on a probability space where Brownian bridge is defined, let L^2 be the corresponding Hilbert space spanned by $\{\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^j\,dW_x^y\}_{j=1}^\infty$. Let g be a continuous function on $[0,b]$ satisfying $g(0)=0$. Since g can be uniformly approximated by polynomials, and since $X_n(F_n(x))$ is the difference between two bounded random measures we have for every n

$$\int_0^b g(x)\,dX_n(F_n(x)) \in L_n^2. \quad (3.12)$$

Because of (1.11) it also follows that

$$\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} g(x)\,dW_x^y \in L^2. \quad (3.13)$$

By the Skorohod representation theorem there exists a probability space and for each n a sequence $\hat{X}_n$ of random variables such that $\hat{X}_n \sim \{\int_0^b x^j\,dX_n(F_n(x))\}_{j=1}^\infty$, plus a sequence $\hat{X}$ such that $\hat{X} \sim \{\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^j\,dW_x^y\}_{j=1}^\infty$, all defined on the space, satisfying

$$\hat{X}_n \to \hat{X} \text{ (pointwise)} \qquad \text{as} \quad n \to \infty \quad (3.14)$$

everywhere on the space. It is clear that $\{L_n^2\}$ and L^2 have their analogues $\{\underline{L}_n^2\}$ and $\underline{L}^2$ on the new space. In the following we will make the identifications $\hat{X}_n = \{\int_0^b x^j\,dX_n(F_n(x))\}$ and $\hat{X} = \{\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^j\,dW_x^y\}$ which will cause no problem when dealing with distribution and algebraic properties of random variables on the new space.

Since (3.3) and (3.14) hold we have

$$\int_0^b x^j \, dX_n(F_n(x)) \xrightarrow{\mathscr{L}^2} \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^j \, dW_x^y \qquad \text{as} \quad n \to \infty, j = 1, 2,..., \tag{3.15}$$

where $\mathscr{L}^2$ denotes convergence in the Hilbert space of square integrable random variables on the new space.

To prove (3.10) we will show for any function f satisfying the conditions of the theorem

$$\int_0^b f(x) \, dX_n(F_n(x)) \xrightarrow{\mathscr{L}^2} \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f(x) \, dW_x^y \qquad \text{as} \quad n \to \infty \tag{3.16}$$

since this will imply the $\mathscr{L}^2$ convergence of arbitrary linear combinations of $\{\int_0^b f_i(x) \, dX_n(F_n(x))\}_{i=1}^d$ to the appropriate random variables.

For each n let $e_1^n, e_2^n,...,$ be the basis (possibly finite) for $\underline{L}_n^2$ arising from the Gram–Schmidt orthonormalization process on $\{\int_0^b x^i \, dX_n(F_n(x))\}_{j=1}^{\infty}$, where e_m^n is a linear combination of the first m linearly independent variables from this sequence. Similarly, let $e_1, e_2,...,$ be the basis for $\underline{L}^2$ arising from $\{\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^j \, dW_x^y\}_{j=1}^{\infty}$. We will show the latter sequence to be linearly independent. For any numbers $\alpha_1, \alpha_2,..., \alpha_m$ we have $E((\sum_{i=1}^m \alpha_i \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^i \, dW_x^y)^2) = 0$ and (1.11) implies

$$\begin{aligned} 0 &= \sum \alpha_i \alpha_j \left(\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^{i+j} f_y(x) \, dx - \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^i f_y(x) \, dx \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^j f_y(x) \, dx \right) \\ &= \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} \left(\sum_{i=1}^m \alpha_i x^i - \int_0^{(1+\sqrt{y})^2} \sum_{i=1}^m \alpha_i \underline{x}^i f_y(\underline{x}) \, d\underline{x} \right)^2 f_y(x) \, dx, \end{aligned} \tag{3.17}$$

which implies $\sum_{i=1}^m \alpha_i x^i = \text{const.}$ for infinitely many x's which of course implies $\alpha_i = 0$, $i = 1, 2,..., m$.

Therefore, because of (3.3) and (3.15), for any m we have for n sufficiently large: $e_1^n, e_2^n,..., e_m^n$ all defined, are linear combinations of $\{\int_0^b x^i \, dX_n(F_n(x))\}_{i=1}^m$, and

$$e_m^n \xrightarrow{\mathscr{L}^2} e_m \qquad \text{as} \quad n \to \infty. \tag{3.18}$$

Writing

$$\int_0^b f(x) \, dX_n(F_n(x)) = \sum_j E\left(e_j^n \int_0^b f(x) \, dX_n(F_n(x)) \right) e_j^n \tag{3.19}$$

and

$$\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f(x)\,dW_x^y = \sum_{j=1}^{\infty} E\left(e_j \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f(x)\,dW_x^y\right) e_j, \tag{3.20}$$

we see that because of (3.18), in order to prove (3.16) we need to show

$$E\left(\left(\int_0^b f(x)\,dX_n(F_n(x))\right)^2\right) \to E\left(\left(\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f(x)\,dW_x^y\right)^2\right) \qquad \text{as} \quad n\to\infty \tag{3.21}$$

and

$$E\left(e_j^n \int_0^b f(x)\,dX_n(F_n(x))\right) \to E\left(e_j \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f(x)\,dW_x^y\right) \qquad \text{as} \quad n\to\infty \tag{3.22}$$

for every j. We will do this by showing for all functions f and g satisfying the conditions of the theorem

$$E\left(\int_0^b f(x)\,dX_n(F_n(x)) \int_0^b g(x)\,dX_n(F_n(x))\right)$$
$$\to E\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f(x)\,dW_x^y \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} g(x)\,dW_x^y \qquad \text{as} \quad n\to\infty. \tag{3.23}$$

Let $f(x)=\sum_{r=1}^{\infty}\alpha_r x^r$, $g(x)=\sum_{r=1}^{\infty}\beta_r x^r$, and let $a>b$ be less than the smaller of two radii of convergence. Let w satisfy (1.7) for $z=b$ and

$$w > \frac{-1}{2\ln(b/a)}. \tag{3.24}$$

We have

$$\left| E\left(\int_0^b f(x)\,dX_n(F_n(x)) \int_0^b g(x)\,dX_n(F_n(x))\right)\right.$$
$$\left. - E\left(\int_0^b \sum_{r=1}^{[w\ln(n)]} \alpha_r x^r\,dX_n(F_n(x)) \int_0^b \sum_{r=1}^{[w\ln(n)]} \beta_r x^r\,dX_n(F_n(x))\right)\right|$$
$$\leqslant \left(E\left(\left(\int_0^b \sum_{r=1}^{[w\ln(n)]} \alpha_r x^r\,dX_n(F_n(x))\right)^2\right)\right)^{1/2}$$
$$\times \left(E\left(\left(\int_0^b \left(g(x) - \sum_{r=1}^{[w\ln(n)]} \beta_r x^r\right) dX_n(F_n(x))\right)^2\right)\right)^{1/2}$$

$$+\left(E\left(\left(\int_0^b \sum_{r=1}^{[w\ln(n)]} \beta_r x^r\, dX_n(F_n(x))\right)^2\right)\right)^{1/2}$$
$$\times\left(E\left(\left(\int_0^b \left(f(x)-\sum_{r=1}^{[w\ln(n)]} \alpha_r x^r\right) dX_n(F_n(x))\right)^2\right)\right)^{1/2}$$
$$+\left(E\left(\left(\int_0^b \left(f(x)-\sum_{r=1}^{[w\ln(n)]} \alpha_r x^r\right) dX_n(F_n(x))\right)^2\right)\right)^{1/2}$$
$$\times\left(E\left(\left(\int_0^b \left(g(x)-\sum_{r=1}^{[w\ln(n)]} \beta_r x^r\right) dX_n(F_n(x))\right)^2\right)\right)^{1/2}. \tag{3.25}$$

By Corollary 3.2 we have for all n sufficiently large

$$\left| E\left(\int_0^b \sum_{r=1}^{[w\ln(n)]} \alpha_r x^r\, dX_n(F_n(x)) \int_0^b \sum_{r=1}^{[w\ln(n)]} \beta_r x^r\, dX_n(F_n(x))\right)\right|$$
$$=\left|\sum_{r_1,r_2\leqslant[w\ln(n)]} \alpha_{r_1}\beta_{r_2} E\left(\int_0^b x^{r_1}\, dX_n(F_n(x)) \int_0^b x^{r_2}\, dX_n(F_n(x))\right)\right|$$
$$\leqslant \sum_{r_1,r_2\leqslant[w\ln(n)]} |\alpha_{r_1}|\,|\beta_{r_2}| \left(E\left(\left(\int_0^b x^{r_1}\, dX_n(F_n(x))\right)^2\right)\right)^{1/2}$$
$$\times\left(E\left(\left(\int_0^b x^{r_2}\, dX_n(F_n(x))\right)^2\right)\right)^{1/2}$$
$$\leqslant \underline{k}_1 \sum_{r_1,r_2\leqslant[w\ln(n)]} |\alpha_{r_1}|\,|\beta_{r_2}|\, a^{r_1} r_1^{(\underline{k}_2/2)} a^{r_2} r_2^{(\underline{k}_2/2)}$$
$$\leqslant \underline{k}_1 \left(\sum_{r=1}^{\infty} |\alpha_r|\, a^r r^{(\underline{k}_2/2)}\right)\left(\sum_{r=1}^{\infty} |\beta_r|\, a^r r^{(\underline{k}_2/2)}\right) < \infty. \tag{3.26}$$

Also, from (1.11) it follows that

$$E\left(\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} f(x)\, dW_x^y \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} g(x)\, dW_x^y\right)$$
$$=\sum_{r_1,r_2} \alpha_{r_1}\beta_{r_2} E\left(\int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^{r_1}\, dW_x^y \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} x^{r_2}\, dW_x^y\right). \tag{3.27}$$

Therefore, from (3.3), (3.26), (3.27), and the dominated convergence theorem we have

$$E\left(\int_0^b \sum_{r=1}^{[w\ln(n)]} \alpha_r x^r\, dX_n(F_n(x))\right.$$
$$\left.\times\int_0^b \sum_{r=1}^{[w\ln(n)]} \beta_r x^r\, dX_n(F_n(x))\right) \to (3.27) \qquad \text{as} \quad n\to\infty. \tag{3.28}$$

Because of (3.24) we also have

$$\left(E\left(\left(\int_0^b\left(f(x)-\sum_{r=0}^{[w\ln(n)]}a_rx^r\right)dX_n(F_n(x))\right)^2\right)\right)^{1/2}$$

$$\leqslant n^{1/2}\sum_{r=[x\ln(n)]+1}^{\infty}|a_r|\,b^r$$

$$\leqslant n^{1/2}\left(\frac{b}{a}\right)^{w\ln(n)}\sum_{r=[w\ln(n)]+1}^{\infty}|a_r|\,a^r$$

$$=n^{1/2+w\ln(b/a)}\sum_{r=[w\ln(n)]+1}^{\infty}|a_r|\,a^r\to 0 \qquad \text{as} \quad n\to\infty. \tag{3.29}$$

Therefore, from (3.25), (3.28), and (3.29) we get (3.23). This completes the proof of Theorem 2.

4. Proof of Lemma 3.2

Defining

$$v[a,b,i_2,\ldots,i_r;k_1,\ldots,k_r]=v_{ak_1}v_{i_2k_1}v_{i_2k_2}v_{i_3k_2}\cdots v_{i_rk_r}v_{bk_r} \tag{4.1}$$

we have

$$\frac{2s^{2r}}{n}\times(2.6)=E\left(\left(x^{\mathrm{T}}(V_nV_n^{\mathrm{T}})^rx_n-\frac{1}{n}\operatorname{tr}(V_nV_n^{\mathrm{T}})^r\right)^2\right) \tag{4.2}$$

$$=\sum x_ix_jx_{\underline{i}}x_{\underline{j}}\operatorname{Cov}(v[i,j,i_2,\ldots,i_r;k_1,\ldots,k_r],v[\underline{i},\underline{j};\underline{i}_2,\ldots,\underline{i}_r;\underline{k}_1,\ldots,\underline{k}_r]) \tag{4.3}$$

$$-\frac{2}{n}\sum x_ix_j\operatorname{Cov}(v[i,j,k_2,\ldots,i_r;k_1,\ldots,k_r],v[\underline{i},\underline{i};\underline{i}_2,\ldots,\underline{i}_r;\underline{k}_1,\ldots,\underline{k}_r]) \tag{4.4}$$

$$+\frac{1}{n^2}\sum\operatorname{Cov}(v[i,i;i_2,\ldots,i_r;k_1,\ldots,k_r],v[\underline{i},\underline{i};\underline{i}_2,\ldots,\underline{i}_r;\underline{k}_1,\ldots,\underline{k}_r]) \tag{4.5}$$

$$+\Bigg(\sum x_ix_jE(v[i,j;i_2,\ldots,i_r;k_1,\ldots,k_r]$$

$$-\frac{1}{n}\sum E(v[i,i;i_2,\ldots,i_r;k_1,\ldots,k_r])\Bigg)^2. \tag{4.6}$$

Since $(V_nV_n^{\mathrm{T}})^r_{ii}$, $i=1,\ldots,n$ are identically distributed we have

$$(4.6) = \Bigg(\sum_{i \neq j} x_i x_j E(v[i, j; i_2, \ldots, i_r; k_1, \ldots, k_r]) \Bigg)^2 \tag{4.7}$$

$$\leqslant \sum_{\substack{i \neq j \\ \underline{i} \neq \underline{j}}} x_i x_j x_{\underline{i}} x_{\underline{j}} E(v[i, j; i_2, \ldots, i_r; k_1, \ldots, k_r]\, v[\underline{i}, \underline{j}; \underline{i}_2, \ldots, \underline{i}_r; \underline{k}_1, \ldots, \underline{k}_r]). \tag{4.8}$$

Grouping the identically distributed variables together we arrive at

$$(4.2) \tag{4.9}$$

$$\leqslant \left(\sum x_i^4 - \frac{1}{n} \right)$$

$$\times \sum \operatorname{Cov}(v[1, 1; i_2, \ldots, i_r; k_1, \ldots, k_r], v[1, 1; \underline{i}_2, \ldots, \underline{i}_r; \underline{k}_1, \ldots, \underline{k}_r]) \tag{4.10}$$

$$+ \left(\frac{1}{n} - \sum x_i^4 \right)$$

$$\times \sum \operatorname{Cov}(v[1, 1; i_2, \ldots, i_r; k_1, \ldots, k_r], v[2, 2; \underline{i}_2, \ldots, \underline{i}_r; \underline{k}_1, \ldots, \underline{k}_r]) \tag{4.11}$$

$$+ 4 \left(\frac{1}{n} + \left(\sum x_i^3 \right) \left(\sum x_i \right) - \left(\sum x_i^4 + \frac{1}{n} \left(\sum x_i \right)^2 \right) \right)$$

$$\times \sum \operatorname{Cov}(v[1, 1; i_2, \ldots, i_r; k_1, \ldots, k_r], v]1, 2; \underline{i}_2, \ldots, \underline{i}_r; \underline{k}_1, \ldots, \underline{k}_r\}) \tag{4.12}$$

$$+ 4 \left(\sum x_i^4 + \frac{1}{n} \left(\sum x_i \right)^2 - \left(\frac{1}{n} + \left(\sum x_i^3 \right) \left(\sum x_i \right) \right) \right)$$

$$\times \sum \operatorname{Cov}(v[1, 1; i_2, \ldots, i_r; k_1, \ldots, k_r], v[2, 3; \underline{i}_2, \ldots, \underline{i}_r; \underline{k}_1, \ldots, \underline{k}_r]) \tag{4.13}$$

$$+ 2 \left(1 - \sum x_i^4 \right)$$

$$\times \operatorname{Cov}(v[1, 2; i_2, \ldots, i_r; k_1, \ldots, k_r], v[1, 2; \underline{i}_2, \ldots, \underline{i}_r; \underline{k}_1, \ldots, \underline{k}_r]) \tag{4.14}$$

$$+ 4 \Bigg(\sum_{\substack{i, j, k \\ \text{distinct}}} x_i^2 x_j x_k \Bigg)$$

$$\times \sum \operatorname{Cov}(v[1, 2; i_2, \ldots, i_r; k_1, \ldots, k_r], v[1, 3; \underline{i}_2, \ldots, \underline{i}_r; \underline{k}_1, \ldots, \underline{k}_r]) \tag{4.15}$$

$$+ \Bigg(\sum_{\substack{i, j, \underline{i}, \underline{j} \\ \text{distinct}}} x_i x_j x_{\underline{i}} x_{\underline{j}} \Bigg)$$

$$\times \sum \operatorname{Cov}[v[1, 3; i_2, \ldots, i_r; k_1, \ldots, k_r], v[2, 4; \underline{i}_2, \ldots, \underline{i}_r; \underline{k}_1, \ldots, \underline{k}_r] \tag{4.16}$$

$$+ 2 \left(1 - \sum x_i^4 \right)$$

$$\times \sum E(v[1, 2; i_2, \ldots, i_r; k_1, \ldots, k_r], v[1, 2; \underline{i}_2, \ldots, \underline{i}_r; \underline{k}_1, \ldots, \underline{k}_r]) \tag{4.17}$$

$$+4\left(\sum_{\substack{i,j,k\\ \text{distinct}}} x_i^2 x_j x_k\right)$$

$$\times \sum E(v[1,2;i_2,\dots,i_r;k_1,\dots,k_r]\, v[1,3;\underline{i}_2,\dots,\underline{i}_r;\underline{k}_1,\dots,\underline{k}_r]) \tag{4.18}$$

$$+\left(\sum_{\substack{i,j,\underline{i},\underline{j}\\ \text{distinct}}} x_i x_j x_{\underline{i}} x_{\underline{j}}\right)$$

$$\times \sum E(v[1,3;i_2,\dots,i_r;k_1,\dots,k_r]\, v[2,4;\underline{i}_2,\dots,\underline{i}_r;\underline{k}_1,\dots,\underline{k}_r]). \tag{4.19}$$

Notice in (4.10)–(4.16) a zero term will occur whenever

(a) a v_{ik} or $v_{\underline{ik}}$ appears alone

or

(b) no v_{ik} equals a $v_{\underline{ik}}$,

and a zero term in (4.17)–(4.19) will occur whenever (a) occurs.

As in [4] we will define a V-path for any $r_1, r_2 \geqslant 1$ as an ordered sequence of $2(r_1+r_2)$ elements in V_n such that

(i) the first element is arbitrary,

(ii) the second element is in the same column as the first element, the third element is in the same row as the second, etc., until the $2r_1$th element is reached,

(iii) the $2r_1+1$th element is arbitrary,

(iv) the $2r_1+2$th element is in the same column as the $2r_1+1$th element, etc.,

(v) every element appearing in the path appears at least twice.

It is evident that each nonzero term in (4.10)–(4.19) is the result of a V-path with $r_1 = r_2 = r$.

Let r_0 and c_0 be, respectively, the number of rows and columns of V_n entered by a given V-path. Let $l = r_0 + c_0$. It is clear that l is the number of distinct indices in the V-path. We will first determine a bound on l for each type of V-path in (4.10)–(4.19) resulting in a nonzero term. There are essentially six different types:

$v_{ik_1} : \cdots : v_{jk_{r_1}} : v_{\underline{j}\underline{k}_1} : \cdots : v_{\underline{i}\underline{k}_{r_2}}$, $i \neq j$ appearing in (4.14) and (4.17), which can be seen by reversing the order on a set of indices, say, $\underline{i}_2,\dots,\underline{i}_{r_2}; \underline{k}_1,\dots,\underline{k}_{r_2}$, (4.20)

$v_{1k_1} : \cdots : v_{1k_{r_1}} : v_{1\underline{k}_1} : \cdots : v_{1\underline{k}_{r_2}}$ appearing in (4.10) ((4.20) and (4.21) are those V-paths studied in [4]), (4.21)

$$v_{ik_1} : \cdots : v_{jk_{r_1}} : v_{\underline{j}\underline{k}_1} : \cdots : v_{\underline{j}\underline{k}_{r_2}}, \quad i \neq j, \text{ appearing in (4.12), (4.15), and (4.18)}, \tag{4.22}$$

$$v_{ik_1} : \cdots : v_{ik_{r_1}} : v_{j\underline{k}_1} : \cdots : v_{j\underline{k}_{r_2}}, \; i \neq j, \text{ appearing in (4.11)}, \tag{4.23}$$

$$v_{ik_1} : \cdots : v_{ik_{r_1}} : v_{\underline{i}\underline{k}_1} : \cdots : v_{\underline{j}\underline{k}_{r_2}}, \quad i,\ \underline{i},\ \underline{j} \text{ distinct, appearing in (4.13)}, \tag{4.24}$$

and

$$v_{ik_1} : \cdots : v_{jk_{r_1}} : v_{\underline{i}\underline{k}_1} : \cdots : v_{\underline{j}\underline{k}_{r_2}}, \quad i,\ \underline{i}, j, \underline{j} \text{ distinct, appearing in (4.16) and (4.19)}. \tag{4.25}$$

For a V-path of type (4.20) ler $r' \leqslant 2r$ be the number of distinct elements of V_n appearing in the path and let $(a_1, b_1),..., (a_{r'}, b_{r'})$ be a listing of the indices on the distinct elements where the a's (b's) correspond to the rows (columns) of V_n. Each a_m (b_m) is either constrained, that is equal, to at least one other $a_{m'}$ ($b_{m'}$), or is free, that is, not constrained to any other index. Notice that any pair cannot have both indices free.

We have three cases:

(I) $r' = 1$. Then $l = 2 = r' + 1 \leqslant 2r + 1$.

(II) $r' > 1$, and at most one index is free. Let $t = 0, 1$ be the number of free indices. Then $l \leqslant t + ((2r' - t)/2) = r' + t/2$ so that $l \leqslant r' \leqslant 2r$.

(III) $r' > 1$, and at least two indices are free. Note that a free index corresponds to either stationary moves in the V-path, that is, two adjacent elements in the path being the same, or to the first and last elements being the same. It follows that the removal of any pair (a_i, b_i) containing a free index results in the remaining collection of pairs corresponding to a new V-path of type (4.20) or (4.21) having length $\leqslant 2(2r - 1)$ with $r' - 1$ distinct elements of V_n. Then l is one plus the number of distinct indices in the new V-path.

Case (III) is repeated until (I) or (II) is reached. Suppose (III) is entered a total of j times. It is clear that $j \leqslant 2r - 1$. Then (I) or (II) is reached with a V-path of length $\leqslant 2(2r - j)$, $r' = j$ distinct elements of V_n, and $l - j$ distinct indices. If (II) is reached, then $l \leqslant r' - j + j = r' \leqslant 2r$. If I is reached then $l \leqslant j + r' - j + 1 = r' + 1 \leqslant 2r + 1$.

We conclude that for (4.20) $l \leqslant 2r + 1$.

For a V-path of type (4.21) it is necessary to use the fact that (b) will not hold for any nonzero term of (4.10). Let r' and $(a_1, b_1),..., (a_{r'}, b_{r'})$ be as above where we may as well assume the first r'' pairs are those for which each is associated with an element in V_n appearing as a v_{ik} and as a $v_{\underline{i}\underline{k}}$. Let r''' be the number of pairs having row 1 indices.

We have five cases:

(I) $r' = 1$. Then as above $l = 2 = r' + 1 \leqslant 2r$.

(II) $r' > 1$, and at most one index is free. Then as above $l \leqslant r' \leqslant 2r$.

(III) $r' > 1$, $r'''' = 1$, making row 1 a free index, and there is at least one other free index. Then there is at least one element of V_n appearing at least four times in the path. Therefore $r' \leqslant 2r - 1$. Removing one of the pairs not having a row 1 index results in pairs corresponding to a new path of length $\leqslant 2(2r-1)$ being of type (4.21) except (b) may possibly hold. This case is repeated with, say, j entries until one of the above cases is reached. If (I) is reached then $j \leqslant 2r-2$, $r' - j = 1$, $l - j = 2$, $l \leqslant r' + 1$, and $l \leqslant 2r$. If (II) is reached, then $l \leqslant r' \leqslant 2r$.

(IV) $r''' = 2$, with, say $(1, b)$, $(1, b')$ as the corresponding pairs, at least one of b and b' is free, at most one of the indices from the first r'' pairs is free, and all other indices are constrained. It follows that the only way b and b' can both be free is if $r' = 2$ and $r > 1$; otherwise (b) will hold. This is true even for paths in (4.20) so long as (b) does not hold. For this case $l = 3 = r' + 1 < 2r$. If b, say, is only free, then let $t = 0, 1$ be the number of free indices in the first r'' pairs. Then there are at least t distinct elements of V_n each appearing at least four times in the path. We have then $r \leqslant 1 + t + ((4r - 2 - 4t)/2) = 2r - t$ and $l \leqslant 1 + t + ((2r' - (1 + t))/2) = r' + ((1 + t)/2) \leqslant 2r + (1/2) - (t/2)$. Therefore, $l \leqslant r' + 1$ and $l \leqslant 2r$.

(V) $r''' \geqslant 2$ and at least one index is free. Then either (II) or (IV) holds, or it is possible to remove a pair having a free index resulting in a path of type (4.20) or (4.21) of length $\leqslant 2(2r-1)$ for which there are at least two pairs with row 1 indices and for which (b) is still avoided. This case is repeated until eventually (II) or (IV) is reached. For this case we have $l \leqslant r' + 1$ and $l \leqslant 2r$.

We conclude that for paths of type (4.21) $l \leqslant r' + 1$ and $l \leqslant 2r$.

For the remaining four types it can be shown that $l \leqslant 2r$ for (4.22), (4.24), and (4.25), while $l \leqslant 2r + 1$ for (4.23). The reasoning follows in the same way as above where the avoidance of (b) is necessary for (4.23) and (4.24). It turns out that for (4.22), (4.23), and (4.24), $l \leqslant r' + 1$, whereas for (4.25) $l \leqslant r' + 2$. The fact that $l \leqslant r' + 2$ for all types will be used later.

The bounds on l for the sums in (4.10)–(4.19) are summarized in

TABLE I

l	Sums
$2r$	(4.10), (4.12), (4.13), (4.15), (4.16), (4.18), (4.19)
$2r+1$	(4.11), (4.14), (4.17)

We next consider the coefficients involving the x_i's in each of (4.10)–(4.19). It is a simple matter to show the existence of a constant k for which the coefficients are bounded in absolute value by quantities given in

TABLE II

(4.10)–(4.14), (4.17)	k
(4.15), (4.18)	kn
(4.16), (4.19)	kn^2

Our next step is to determine bounds on each of the sums in (4.10)–(4.19). The arguments and notation used are modifications of those found in [4].

Consider one of the sums in (4.10)–(4.19). Let β_l be the number of V-paths having l distinct indices, and let α_l be an upper bound on the absolute value of any term having l distinct indices. We will use as a bound on any sum

$$\sum \alpha_l \beta_l, \tag{4.26}$$

where the limits on l will depend on the particular sum we are considering.

To derive a value for α_l, consider a V-path of length $2\underline{r}$, where $\underline{r} = r$ or $2r$. Suppose $v_{i_j k_j}$, $j = 1, 2, \ldots, r'$ are the distinct elements of V_n with $v_{i_j k_j}$ appearing n_j times. Let $\underline{j}$ be the number of distinct elements of V_n appearing exactly twice in the V-path. Using (1.3) and the relationship between the geometric and arithmetic means we have

$$E\left(\left|\prod v_{i_j k_j}^{n_j}\right|\right) \leqslant \prod_{n_j \geqslant 3} n_j^{\alpha n_j} \leqslant \left(\frac{\sum_{n_j \geqslant 3} n_j^2}{\sum_{n_j \geqslant 3} n_j}\right)^{\alpha \sum_{n_j \geqslant 3} n_j} \leqslant (2\underline{r})^{\alpha \sum_{n_j \geqslant 3} n_j}. \tag{4.27}$$

For terms in (4.17)–(4.19) we may take $(4r)^{\alpha \sum_{n_j \geqslant 3} n_j}$ for α_l. As for terms in (4.10)–(4.16) let $v_{ak_1} : \cdots : v_{bk_r} \ : \ v_{c\underline{k}_1} : \cdots : v_{d\underline{k}_r}$ be an arbitrary V-path appearing in these terms. Let $\underline{j}^{(1)}$, $n_j^{(1)}$, and $\underline{j}^{(2)}$, $n_j^{(2)}$ be the values associated with $v_{ak_1} : \cdots : v_{bk_r}$ and $v_{c\underline{k}_1} : \cdots : v_{d\underline{k}_r}$ respectively. If (a) holds for either of $v_{ak_1} : \cdots : v_{bk_r}$ or $v_{c\underline{k}_1} : \cdots : v_{d\underline{k}_r}$ then

$$E(v_{ak_1}, \ldots, v_{bk_r})\, E(v_{c\underline{k}_1}, \ldots, v_{d\underline{k}_r}) = 0. \tag{4.28}$$

If not, then $\underline{j}^{(1)} + \underline{j}^{(2)} = \underline{j}$, $\sum_{n_j^{(1)} \geqslant 3} n_j^{(1)} + \sum_{n_j^{(2)} \geqslant 3} n_j^{(2)} = \sum_{n_j \geqslant 3} n_j$, and from (4.27) we have

$$E(|v_{ak_1}, \ldots, v_{bk_r}|)\, E(|v_{c\underline{k}_1}, \ldots, v_{d\underline{k}_r}|) \leqslant (2r)^{\alpha \sum_{n_j \geqslant 3} n_j}. \tag{4.29}$$

We have

$$\sum_{n_j \geqslant 3} n_j = 2\underline{r} - 2\underline{j} \geqslant 3(r' - \underline{j}) \tag{4.30}$$

which implies

$$\sum_{n_j \geqslant 3} n_j \leqslant 6(\underline{r} - \underline{r}') \leqslant 6(2r - r'). \tag{4.31}$$

Therefore, using the fact that $l \leqslant r' + 2$, we take for a suitable bound on all terms in (4.10)–(4.19)

$$\alpha_l = 2(4r)^{6\alpha(2r-l)+12\alpha}. \tag{4.32}$$

We will determine two possible bounds for β_l, one valid for all l and r, another valid for $l \geqslant \frac{8}{5}r + \frac{9}{5}$ and $r \geqslant 6$. To do this we need to define a *canonical V-path*. It is a V-path appearing in (4.10)–(4.19) for which the first element is v_{11}, and the $2r+1$th element is either v_{11}, v_{12}, v_{21}, or v_{22}, and continuing in order from the first element, whenever an element enters a new row or column, it is the next available one. For example,

$$v_{11} : v_{31} : v_{31} : v_{11} : v_{21} : v_{41} : v_{41} : v_{21} \tag{4.33}$$

is a canonical V-path appearing in (4.11) for $r = 2$. Given any V-path, there corresponds a unique canonical V-path. For example,

$$v_{18} : v_{68} : v_{68} : v_{18} : v_{28} : v_{98} : v_{98} : v_{28} \tag{4.34}$$

corresponds to (4.33) by making the row associations $1 \to 1$, $6 \to 3$, $9 \to 4$, and the column association $8 \to 1$.

Let m_{r_0,c_0} be the number of canonical V-paths using r_0 rows and c_0 columns of V_n. Since $r \leqslant [w \ln(n)]$ and each term in (4.10)–(4.19) has 1, 2, 3, or 4 indices fixed, for n sufficiently large there are

$$\frac{n!}{(n-(r_0-j))!}\,\frac{s!}{(s-c_0)!} \tag{4.35}$$

V-paths associated with a canonical V-path, where j depends on the sum under consideration (see Table III)

TABLE III

j	Sums
1	(4.10)
2	(4.11), (4.12), (4.14), (4.17)
3	(4.13), (4.15), (4.18)
4	(4.16), (4.19)

We have

$$\beta_l = \sum_{r_0=j}^{l-1} m_{r_0,l-r_0} \frac{n!}{(n-(r_0-j))!} \frac{s!}{(s-(l-r_0))!}$$

$$\leqslant s^{l-j} \sum_{r_0=j}^{l-1} m_{r_0,l-r_0} \left(\frac{n}{s}\right)^{r_0}. \tag{4.36}$$

For a bound on β_l valid for all l and r we note that any canonical V-path with l distinct indices has all its elements contained in the upper left $l \times l$ submatrix of V_n. Therefore, the number of canonical V-paths is bounded by $(l^2)^{4r}$, and we have

$$\beta_l \leqslant s^{l-j} \beta^l l^{8r}, \tag{4.37}$$

where $\beta \geqslant 1$ is any bound on (n/s) for all n.

The other bound on β_l, valid for $l \geqslant \frac{8}{5}r + \frac{9}{5}$ and $r \geqslant 6$ is

$$s^{l-j} r^{\xi_1(2r-l)+\xi_2} \left(1 + \left(\frac{n}{s}\right)^{1/2}\right)^{4r}, \tag{4.38}$$

where ξ_1 and ξ_2 are positive numbers, and j depends on the sum under consideration. Before deriving this bound we will use it to complete the proof of the lemma.

Consider one of the sums in (4.10)–(4.19). Let j be the appropriate value according to Table III. Let $t_1 = 0$ or 1 be such that l in any term is bounded by $2r + t_1$, as given in Table I. Let $t_2 = 0$, 1 or 2 be such that the coefficient involving the x_i's is bounded in absolute value by kn^{t_2}, as given in Table II. Then, using (4.26), (4.32), (4.37), (4.38), and the fact that $l \geqslant 2$ for any term, for $r \geqslant 6$ the sum, multiplied by $n/2s^{2r}$, is bounded in absolute value by

$$\left(\frac{n}{2s^{2r}}\right) \times kn^{t_2} \times \sum_{2 \leqslant l \leqslant (8/5)r + (9/5)} 2(4r)^{12\alpha r} s^{l-j} \beta^l l^{8r} \tag{4.39}$$

$$+ \left(\frac{n}{2s^{2r}}\right) \times kn^{t_2} \times \sum_{(8/5)r + (9/5) \leqslant l \leqslant 2r + t_1}$$

$$2(4r)^{6\alpha(2r-l) + 12\alpha} s^{l-j} r^{\xi_1(2r-l)+\xi_2} \left(1 + \left(\frac{n}{s}\right)^{1/2}\right)^{4r}. \tag{4.40}$$

We have

$$(4.39) \leqslant k \frac{n^{1+t_2}}{s^{2r}} s^{-j} (4r)^{12\alpha r + 8r + 1} s^{(8/5)r + (9/5)} \beta^{2r}$$

$$= kn^{1+t_2} s^{-j} \left(\frac{(4r)^{12\alpha + 8 + (1/r)} \beta^2}{s^{((2/5) - (9/5r))}}\right)^r. \tag{4.41}$$

Upon inspection of Tables II and III we find that for any of (4.10)–(4.19), $1+t_2-j\leqslant 0$. Also, since $r\leqslant[w\ln(n)]$ and $n/s\to y>0$ as $n\to\infty$, we see that the quantity in parentheses is less than 1 for all n sufficiently large, independent of r.

We have

$$(4.40)\leqslant k\left(1+\left(\frac{n}{s}\right)^{1/2}\right)^{4r}\frac{n^{1+t_2}}{s^{2r}}s^{-j}(4r)^{12\alpha r+12\alpha}r^{\xi_1 2r+\xi_2}\times\sum_{(8/5)r+(9/5)\leqslant l\leqslant 2r+t_1}\left(\frac{s}{r^{\xi_1}(4r)^{6\alpha}}\right)^{l}. \tag{4.42}$$

For n sufficiently large, $s/r^{\xi_1}(4r)^{6\alpha}>1$, so that

$$\begin{aligned}(4.40)&\leqslant k\left(1+\left(\frac{n}{s}\right)^{1/2}\right)^{4r}\frac{n^{1+t_2}}{s^{2r}}s^{-j}(4r)^{12\alpha r+12\alpha+1}r^{\xi_1 2r+\xi_2}\\&\quad\times\left(\frac{s}{r^{\xi_1}(4r)^{6\alpha}}\right)^{2r+t_1}\\&=k\left(1+\left(\frac{n}{s}\right)^{1/2}\right)^{4r}n^{1+t_2}s^{-j+t_1}(4r)^{12\alpha+1-t_1}r^{\xi_2-t_1}.\end{aligned} \tag{4.43}$$

Again, upon inspection of Tables I–III, for any of (4.10)–(4.19) $1+t_1+t_2-j\leqslant 0$. Therefore, it is clear that constants k_1 and k_2 can be found so that for all n sufficiently large (3.6) holds for $6\leqslant r\leqslant[w\ln(n)]$.

For values of $r\leqslant 6$ we simply note that for each $r\geqslant 1$ the limiting value of $E((\int_0^\infty x^r\,dX_n(F_n(x)))^2)$, (1.11), is bounded by $(1+\sqrt{y})^{4r}$ so that, with possibly larger values of k_1 and k_2 (3.6) is true for all n sufficiently large.

To derive the bound (4.38) on β_l valid for $l\geqslant\frac{8}{5}r+\frac{9}{5}$ and $r\geqslant 6$, we need to determine bounds on the m_{r_0,c_0}'s. To this end we define an element of a canonical V-path to be a *row* (*column*) *innovation* if it is the first entry into a row (column) (proceeding from the first element). We will consider the first element to be a column innovation. Notice that the $2r+1$th element can be both a row and column innovation at the same time.

We will distinguish the following four types of elements:

Type 1. Row innovations.

Type 2. Column innovations.

Type 3. Elements which are the first to repeat a row or column innovation.

Type 4. All other elements.

Except for the $2r+1$th element (which can be a type 1 and type 2 element) all other elements fit into only one of four types for a given canonical V-path. There are $r_0 - 1$ type 1 elements, $l - r_0$ type 2 elements, $l-2$ or $l-1$ type 3 elements, depending on whether the $2r+1$th element is both a row and column innovation or not, and, for the same reason, $4r - 2(l-2)$ or $4 - 2(l-1)$ type 4 elements.

Let us determine a bound on the number of ways the four types can be distributed among the $4r$ elements. Consider the case when the $2r+1$th element is both a row and column innovation. Excluding the first and $2r+1$th elements, there are $r_0 - 2$ type 1 elements distributed among $2r$ elements, $l - r_0 - 2$ type 2 elements distributed among $2r-2$ elements, and $l-2$ type 3 elements distributed among the remaining $4r - l + 2$ elements.

Therefore for this case we get as a bound

$$\binom{2r}{r_0 - 2}\binom{2r-2}{l-r_0-2}\binom{4r-l+2}{l-2}. \tag{4.44}$$

Note this case can only occur for $2 \leqslant r_0 \leqslant l-2$.

Consider now the case when the $2r+1$th element is not both a row and column innovation. As above, we arrive at the bound

$$\binom{2r+1}{r_0-1}\binom{2r-1}{l-r_0-1}\binom{4r-l+1}{l-1}. \tag{4.45}$$

Therefore, the number of ways the four types can be distributed among the $4r$ elements is bounded by

$$2(2r+1)(2r-1)\,l(l-1)\binom{2r}{r_0-2}\binom{2r-2}{l-r_0-2}\binom{4r-l+2}{l}. \tag{4.46}$$

Our next step is to derive a bound on the number of canonical V-paths associated with a given distribution of the four types. We first notice that the $2r+1$th element is always one of four possible elements of V_n and each element other than the first and $2r+1$th has either its row or column position determined by the element appearing before it. For each element in the path we will determine a bound on the number of possible elements it can be, assuming knowledge of the possibilities for the elements before it, and of the particular element of V_n the $2r+1$th element is.

If the $2r+1$th element is not a row (column) innovation and is in row (column) 2 of V_n, then one of the r $(r-1)$ elements before it which results from a row (column) move is a row (column) innovation moving into row (column) 2. All other row (column) innovations are unambiguous, each haying to move into the next empty row (column) other than 2. Any other possibility for the $2r+1$th element will lead to no other ambiguity for elements of types 1 and 2. Therefore these types contribute a factor bounded by

$$r^2 \tag{4.47}$$

Since the elements in a canonical V-path are contained in the upper left $2r \times 2r$ submatrix of V_n, each of the type 4 elements, other than the $2r+1$th element, introduces a factor of at most $2r$. Therefore, the type 4 elements contribute a factor bounded by

$$(2r)^{2(2r-l)+4}. \tag{4.48}$$

For type 3 elements we need to distinguish between three subtypes:

(a) Those which follow an innovation element. For these types other than the $2r+1$th element there is no ambiguity; they must repeat the previous element, since there is no other element in the row or column to choose from.

(b) Those which follow type 4 elements. As above we get that these types contribute a factor of at most (4.48).

(c) Those which follow type 3 elements. Consider one of these types other than the $2r+1$th element for which there is an ambiguous choice. Suppose it must be chosen from the same column, say column k, as the previous type 3 element. Then before the previous type 3 element there must have been at least 3 unpaired innovation elements in column k. We will determine a bound on the maximum number of times this can happen.

Assume that the $2r+1$th element does not lie between the first innovation in k (which can be the $2r+1$th element) and the given element. The first innovation in k is a column innovation and all subsequent innovations in k must be row innovations made by moving along in column k. After the column innovation in k the next element must be chosen from the same column. It either repeats the previous element, is a type 4 element, or is a row innovarion, in which case there will be two unpaired innovations in k. In the latter, the next move can either repeat the row innovation or move to another column. In any event, there is no way three unpaired innovations can be formed in column k unless the path leaves the column and re-enters with a type 4 element (a type 3 element entering k will reduce the number of unpaired innovations by one). Each subsequent element forming three or more unpaired innovations in k must be preceeded by an entry into k with a type 4 element, as long as the $2r+1$th element is not reached.

Now, the only way the $2r+1$th element can disturb the above scheme is if it is a row 2 innovation directly into column k with $k=1$ or 2. Then it is possible for one ambiguity to occur in column k without a type 4 element preceeding. Note the $2r+1$th element cannot also be a column innovation, thus avoiding the addition of another ambiguity along a row.

The above arguments also apply to rows.

We have then a maximum of $2(2r-l)+4$ (when the $2r+1$th element is both a row and column innovation) type 3 elements of (c) type for which an ambiguity can occur.

With i ranging along the number of ambiguities we therefore have for a bound on the factor contributed by type 3 elements of (c) type the quantity

$$\sum_{i=0}^{2(2r-l)+4} \binom{l-1}{i} (2r)^i. \tag{4.49}$$

Under the assumption $l \geqslant \frac{8}{5}r + \frac{9}{5}$ the largest term in (4.49) is the $2(2r-l)+4$ term $(l \geqslant \frac{8}{5}r + \frac{9}{5} \Leftrightarrow 2(2r-l)+4 \leqslant (l-1)/2)$ and since $l \leqslant 2r+1$ we can bound (4.49) by

$$(r+1)\binom{l-1}{2(2r-l)+4}(2r)^{2(2r-l)+4}. \tag{4.50}$$

Combining the above we get

$$\begin{aligned}
m_{r_0,l-r_0} &\leqslant 2(2r+1)(2r-1)\, l(l-1) \\
&\quad \times \binom{2r}{r_0-2}\binom{2r-2}{l-r_0-2}\binom{4r-l+2}{l} \times 4 \times r^2 \\
&\quad \times ((2r)^{2(2r-l)+4}) \times ((2r)^{2(2r-l)+4}) \times (r+1)\binom{l-1}{2(2r-l)+4} \\
&\quad \times (2r)^{2(2r-l)+4} \\
&= 2(2r+1)(2r-1)(r+1)\binom{2r}{r_0-2}\binom{2r-2}{l-r_0-2} \\
&\quad \times \frac{(l-1)(4r-l+2)!}{(4r-2l+2)!\,(4r-2l+4)!\,(3l-4r-5)!}(2r)^{6(2r-l)+14} \\
&\leqslant 2(2r+1)(r+1)\binom{2r}{r_0}^2 \\
&\quad \times \frac{(r_0!)^2((2r-r_0)!)^2}{(r_0-2)!\,(2r-(r_0-2))!\,(l-r_0-2)!\,(2r-(l-r_0))!} \\
&\quad \times \frac{(4r-l+2)!}{(3l-4r-5)!}(2r)^{6(2r-l)+14} \\
&\leqslant 2(2r+1)(r+1)\binom{2r}{r_0}^2 \frac{(r_0!)\, r_0(r_0-1)(2r-r_0)!}{(l-r_0-2)!\,(2r-(l-r_0))!} \\
&\quad \times (4r)^{4(2r-l)+7}(2r)^{6(2r-l)+14}
\end{aligned}$$

$$\leqslant 2(2r+1)(r+1)\binom{2r}{r_0}^2 (4r)^{4(2r-l)+7}(2r)^{7(2r-l)+19}$$

$$\leqslant \binom{2r}{r_0}^2 r^{22(2r-l)+56} \tag{4.51}$$

valid for $r \geqslant 6$.

Therefore using (4.36) we have for $l \geqslant \frac{8}{5}r + \frac{9}{5}$, $r \geqslant 6$

$$\beta_l \leqslant s^{l-j} r^{22(2r-l)+56} \sum_{r_0=j}^{l-1} \binom{2r}{r_0}^2 \left(\frac{n}{s}\right)^{r_0}$$

$$\leqslant s^{l-j} r^{22(2r-l)+56} \left(1 + \left(\frac{n}{s}\right)^{1/2}\right)^{4r}, \tag{4.52}$$

where the relation between j and the sum under consideration is given in Table II. Thus, (4.38) is established with $\xi_1 = 22$, $\xi_2 = 56$.

References

[1] Arnold, L. (1967). On the asymptotic distribution of the eigenvalues of random matrices. *J. Math. Anal. Appl.* **20** 262–268.

[2] Bartoszynski, R. (1961). A characterization of the weak convergence of measures. *Ann. Math. Statist.* **32** 561–576.

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

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

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

[6] Ranga Rao, R. (1962). Relations between weak and uniform convergence of measures with applications. *Ann. Math. Statist.* **33** 659–680.

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

[8] Silverstein, J. W. (1981). Describing the behavior of eigenvectors of random matrices using sequences of measures on orthogonal groups. *SIAM J. Math. Anal.* **12** 274–281.

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