$$S_2 = p^{-1}n^{-k} \sum_{\Delta_2(k,r,s)} p(p-1)\cdots(p-r)n(n-1)\cdots(n-s+1)\mathrm{E}(X_{\Delta_2(k,r,s)}) = 0.$$

By Lemma 3.3, for a graph of $\Delta_3(k,r,s)$, we have $r+s<k$. Since the variable $x_{\Delta(k,r,s)}$ is bounded by $(2C/\tilde{\sigma})^{2k}$, we conclude that

$$\begin{aligned} S_3 &= p^{-1}n^{-k} \sum_{\Delta_3(k,r,s)} p(p-1)\cdots(p-r)n(n-1)\cdots(n-s+1)\mathrm{E}(X_{\Delta(k,r,s)}) \\ &= O(n^{-1}). \end{aligned}$$

Now let us evaluate S_1. For a graph in $\Delta_1(k,r)$ (with $s=k-r$), each pair of coincident edges consists of a down edge and an up edge; say, the edge (i_a, j_a) must coincide with the edge (j_a, i_a). This pair of coincident edges corresponds to the expectation $\mathrm{E}(|X_{i_a,j_a}|^2) = 1$. Therefore, $\mathrm{E}(X_{\Delta_1(k,r)}) = 1$. By Lemma 3.4,

$$\begin{aligned} S_1 &= p^{-1}n^{-k} \sum_{\Delta_1(k,r)} p(p-1)\cdots(p-r)n(n-1)\cdots(n-s+1)\mathrm{E}(X_{\Delta_1(k,r)}) \\ &= \sum_{r=0}^{k-1} \frac{y_n^r}{r+1}\binom{k}{r}\binom{k-1}{r} + O(n^{-1}) \\ &= \beta_k + o(1), \end{aligned}$$

where $y_n = p/n \to y \in (0,\infty)$. The proof of (3.1.6) is complete.

The proof of (3.1.7). Recall

$$\begin{aligned} &\mathrm{Var}(\beta_k(\mathbf{S}_n)) \\ &= p^{-2}n^{-2k} \sum_{\mathbf{i},\mathbf{j}} [\mathrm{E}(X_{G_1(\mathbf{i}_1,\mathbf{j}_1)} X_{G_2(\mathbf{i}_2,\mathbf{j}_2)}) - \mathrm{E}(X_{G_1(\mathbf{i}_1,\mathbf{j}_1)})\mathrm{E}(X_{G_2(\mathbf{i}_2,\mathbf{j}_2)})]. \end{aligned}$$

Similar to the proof of Theorem 2.5, if G_1 has no edges coincident with edges of G_2 or $G = G_1 \cup G_2$ has an overall single edge, then

$$\mathrm{E}(X_{G_1(\mathbf{i}_1,\mathbf{j}_1)} X_{G_2(\mathbf{i}_2,\mathbf{j}_2)}) - \mathrm{E}(X_{G_1(\mathbf{i}_1,\mathbf{j}_1)})\mathrm{E}(X_{G_2(\mathbf{i}_2,\mathbf{j}_2)}) = 0$$

by independence between X_{G_1} and X_{G_2}.

Similar to the arguments in Subsection 2.1.3, one may show that the number of noncoincident vertices of G is not more than $2k$. By the fact that the terms are bounded, we conclude that assertion (3.1.7) holds and consequently conclude the proof of Theorem 3.7.

Remark 3.9. The existence of the second moment of the entries is obviously necessary and sufficient for the Marčenko-Pastur law since the limiting distribution involves the parameter σ^2.

3.2 Generalization to the Non-iid Case

Sometimes it is of practical interest to consider the case where the entries of $\mathbf{X}_n$ depend on n and for each n they are independent but not necessarily identically distributed. As in Section 2.2, we shall briefly present a proof of the following theorem.

Theorem 3.10. *Suppose that, for each n, the entries of $\mathbf{X}$ are independent complex variables with a common mean μ and variance σ^2. Assume that $p/n \to y \in (0,\infty)$ and that, for any $\eta > 0$,*

$$\frac{1}{\eta^2 np} \sum_{jk} \mathrm{E}(|x_{jk}^{(n)}|^2 I(|x_{jk}^{(n)}| \geq \eta\sqrt{n})) \to 0. \tag{3.2.1}$$

Then, with probability one, $F^{\mathbf{S}}$ tends to the Marčenko-Pastur law with ratio index y and scale index σ^2.

Proof. We shall only give an outline of the proof of this theorem. The details are left to the reader. Without loss of generality, we assume that $\mu = 0$ and $\sigma^2 = 1$. Similar to what we did in the proof of Theorem 2.9, we may select a sequence $\eta_n \downarrow 0$ such that condition (3.2.1) holds true when η is replaced by η_n. In the following, once condition (3.2.1) is used, we always mean this condition with η replaced by η_n.

Applying Theorem A.44 and the Bernstein inequality, by condition (3.2.1), we may truncate the variables $x_{ij}^{(n)}$ at $\eta_n\sqrt{n}$. Then, applying Corollary A.42, by condition (3.2.1), we may recentralize and rescale the truncated variables. Thus, in the rest of the proof, we shall drop the superscript (n) from the variables for brevity. We further assume that

$$\begin{aligned} &1)\ |x_{ij}| < \eta_n\sqrt{n}, \\ &2)\ \mathrm{E}(x_{ij}) = 0 \ \text{ and } \ \mathrm{Var}(x_{ij}) = 1. \end{aligned} \tag{3.2.2}$$

By arguments to those in the proof of Theorem 2.9, one can show the following two assertions:

$$\mathrm{E}(\beta_k(\mathbf{S}_n)) = \sum_{r=0}^{k-1} \frac{y_n^r}{r+1}\binom{k}{r}\binom{k-1}{r} + o(1) \tag{3.2.3}$$

and

$$\mathrm{E}\,|\beta_k(\mathbf{S}_n) - \mathrm{E}\,(\beta_k(\mathbf{S}_n))|^4 = o(n^{-2}). \tag{3.2.4}$$

The proof of Theorem 3.10 is then complete.

3.3 Proof of Theorem 3.10 by the Stieltjes Transform

As an illustration applying Stieltjes transforms to sample covariance matrices, we give a proof of Theorem 3.10 in this section. Using the same approach of truncating, centralizing, and rescaling as we did in the last section, we may assume the additional conditions given in (3.2.2).

3.3.1 Stieltjes Transform of the M-P Law

Let $z = u + iv$ with $v > 0$ and $s(z)$ be the Stieltjes transform of the M-P law.

Lemma 3.11.

$$s(z) = \frac{\sigma^2(1-y) - z + \sqrt{(z - \sigma^2 - y\sigma^2)^2 - 4y\sigma^4}}{2yz\sigma^2}. \tag{3.3.1}$$

Proof. When $y < 1$, we have

$$s(z) = \int_a^b \frac{1}{x-z} \frac{1}{2\pi x y \sigma^2} \sqrt{(b-x)(x-a)}dx,$$

where $a = \sigma^2(1-\sqrt{y})^2$ and $b = \sigma^2(1+\sqrt{y})^2$.

Letting $x = \sigma^2(1 + y + 2\sqrt{y}\cos w)$ and then setting $\zeta = e^{iw}$, we have

$$\begin{aligned} s(z) &= \int_0^\pi \frac{2}{\pi} \frac{1}{(1+y+2\sqrt{y}\cos w)(\sigma^2(1+y+2\sqrt{y}\cos w) - z)} \sin^2 w dw \\ &= \frac{1}{\pi}\int_0^{2\pi} \frac{((e^{iw} - e^{-iw})/2i)^2}{(1+y+\sqrt{y}(e^{iw}+e^{-iw}))(\sigma^2(1+y+\sqrt{y}(e^{iw}+e^{-iw})) - z)} dw \\ &= -\frac{1}{4i\pi} \oint_{|\zeta|=1} \frac{(\zeta - \zeta^{-1})^2}{\zeta(1+y+\sqrt{y}(\zeta+\zeta^{-1}))(\sigma^2(1+y+\sqrt{y}(\zeta+\zeta^{-1})) - z)} d\zeta \\ &= -\frac{1}{4i\pi} \oint_{|\zeta|=1} \frac{(\zeta^2-1)^2}{\zeta((1+y)\zeta + \sqrt{y}(\zeta^2+1))(\sigma^2(1+y)\zeta + \sqrt{y}\sigma^2(\zeta^2+1) - z\zeta)} d\zeta. \end{aligned} \tag{3.3.2}$$

The integrand function has five simple poles at

$$\begin{aligned} \zeta_0 &= 0, \\ \zeta_1 &= \frac{-(1+y)+(1-y)}{2\sqrt{y}}, \\ \zeta_2 &= \frac{-(1+y)-(1-y)}{2\sqrt{y}}, \end{aligned}$$

$$\begin{aligned} \zeta_3 &= \frac{-\sigma^2(1+y) + z + \sqrt{\sigma^4(1-y)^2 - 2\sigma^2(1+y)z + z^2}}{2\sigma^2\sqrt{y}}, \\ \zeta_4 &= \frac{-\sigma^2(1+y) + z - \sqrt{\sigma^4(1-y)^2 - 2\sigma^2(1+y)z + z^2}}{2\sigma^2\sqrt{y}}. \end{aligned}$$

By elementary calculation, we find that the residues at these five poles are

$$\frac{1}{y\sigma^2}, \quad \mp\frac{1-y}{yz} \quad \text{and} \quad \pm\frac{1}{\sigma^2 yz}\sqrt{\sigma^4(1-y)^2 - 2\sigma^2(1+y)z + z^2}.$$

Noting that $\zeta_3\zeta_4 = 1$ and recalling the definition for the square root of complex numbers, we know that both the real part and imaginary part of $\sqrt{\sigma^2(1-y)^2 - 2\sigma^2(1+y)z + z^2}$ and $-\sigma^2(1+y)+z$ have the same signs and hence $|\zeta_3| > 1$, $|\zeta_4| < 1$. Also, $|\zeta_1| = |-\sqrt{y}| < 1$ and $|\zeta_2| = |-1/\sqrt{y}| > 1$. By Cauchy integration, we obtain

$$\begin{aligned} s(z) &= -\frac{1}{2}\left(\frac{1}{y\sigma^2} - \frac{1}{\sigma^2 yz}\sqrt{\sigma^4(1-y)^2 - 2\sigma^2(1+y)z + z^2} - \frac{1-y}{yz}\right) \\ &= \frac{\sigma^2(1-y) - z + \sqrt{(z-\sigma^2-y\sigma^2)^2 - 4y\sigma^4}}{2yz\sigma^2}. \end{aligned}$$

This proves equation (3.3.1) when $y < 1$.

When $y > 1$, since the M-P law has also a point mass $1 - 1/y$ at zero, $s(z)$ equals the integral above plus $-(y-1)/yz$. In this case, $|\zeta_3| = |-\sqrt{y}| > 1$ and $|\zeta_4| = |-1/\sqrt{y}| < 1$, and thus the residue at ζ_4 should be counted into the integral. Finally, one finds that equation (3.3.1) still holds. When $y = 1$, the equation is still true by continuity in y.

3.3.2 Proof of Theorem 3.10

Let the Stieltjes transform of the ESD of $\mathbf{S}_n$ be denoted by $s_n(z)$. Define

$$s_n(z) = \frac{1}{p}\operatorname{tr}(\mathbf{S}_n - z\mathbf{I}_p)^{-1}.$$

As in Section 2.3, we shall complete the proof by the following three steps:

(i) For any fixed $z \in \mathbb{C}^+$, $s_n(z) - \mathrm{E}s_n(z) \to 0$, a.s.
(ii) For any fixed $z \in \mathbb{C}^+$, $\mathrm{E}s_n(z) \to s(z)$, the Stieltjes transform of the M-P law.
(iii) Except for a null set, $s_n(z) \to s(z)$ for every $z \in \mathbb{C}^+$.

Similar to Section 2.3, the last step is implied by the first two steps and thus its proof is omitted. We now proceed with the first two steps.

Step 1. Almost sure convergence of the random part

$$s_n(z) - \mathrm{E}s_n(z) \to 0, \quad \text{a.s.} \tag{3.3.3}$$

Let $\mathrm{E}_k(\cdot)$ denote the conditional expectation given $\{\mathbf{x}_{k+1}, \cdots, \mathbf{x}_n\}$. Then, by the formula

$$(\mathbf{A} + \boldsymbol{\alpha}\boldsymbol{\beta}^*)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\boldsymbol{\alpha}\boldsymbol{\beta}^*\mathbf{A}^{-1}}{1 + \boldsymbol{\beta}^*\mathbf{A}^{-1}\boldsymbol{\alpha}} \tag{3.3.4}$$

we obtain

$$\begin{aligned} s_n(z) - \mathrm{E}s_n(z) &= \frac{1}{p}\sum_{k=1}^{n}[\mathrm{E}_k\mathrm{tr}(\mathbf{S}_n - z\mathbf{I}_p)^{-1} - \mathrm{E}_{k-1}\mathrm{tr}(\mathbf{S}_n - z\mathbf{I}_p)^{-1}] \\ &= \frac{1}{p}\sum_{k=1}^{n}\gamma_k, \end{aligned}$$

where, by Theorem A.5,

$$\begin{aligned} \gamma_k &= (\mathrm{E}_k - \mathrm{E}_{k-1})[\mathrm{tr}(\mathbf{S}_n - z\mathbf{I}_p)^{-1} - \mathrm{tr}(\mathbf{S}_{nk} - z\mathbf{I}_p)^{-1}] \\ &= -[\mathrm{E}_k - \mathrm{E}_{k-1}]\frac{\mathbf{x}_k^*(\mathbf{S}_{nk} - z\mathbf{I}_p)^{-2}\mathbf{x}_k}{1 + \mathbf{x}_k^*(\mathbf{S}_{nk} - z\mathbf{I}_p)^{-1}\mathbf{x}_k} \end{aligned}$$

and $\mathbf{S}_{nk} = \mathbf{S}_n - \mathbf{x}_k\mathbf{x}_k^*$. Note that

$$\begin{aligned} &\left|\frac{\mathbf{x}_k^*(\mathbf{S}_{nk} - z\mathbf{I}_p)^{-2}\mathbf{x}_k}{1 + \mathbf{x}_k^*(\mathbf{S}_{nk} - z\mathbf{I}_p)^{-1}\mathbf{x}_k}\right| \\ &\le \frac{\mathbf{x}_k^*((\mathbf{S}_{nk} - u\mathbf{I}_p)^2 + v^2\mathbf{I}_p)^{-1}\mathbf{x}_k}{\Im(1 + \mathbf{x}_k^*(\mathbf{S}_{nk} - z\mathbf{I}_p)^{-1}\mathbf{x}_k)} = \frac{1}{v}. \end{aligned}$$

Noticing that $\{\gamma_k\}$ forms a sequence of bounded martingale differences, by Lemma 2.12 with $p = 4$, we obtain

$$\begin{aligned} \mathrm{E}|s_n(z) - \mathrm{E}s_n(z)|^4 &\le \frac{K_4}{p^4}\mathrm{E}\left(\sum_{k=1}^{n}|\gamma_k|^2\right)^2 \\ &\le \frac{4K_4n^2}{v^4p^4} = O(n^{-2}), \end{aligned}$$

which, together with the Borel-Cantelli lemma, implies (3.3.3). The proof is complete.

Step 2. Mean convergence

We will show that

$$\mathrm{E}s_n(z) \to s(z), \tag{3.3.5}$$

where $s(z)$ is defined in (3.3.1) with $\sigma^2 = 1$.

By Theorem A.4, we have

$$s_n(z) = \frac{1}{p}\sum_{k=1}^{p}\frac{1}{\frac{1}{n}\boldsymbol{\alpha}_k'\overline{\boldsymbol{\alpha}}_k - z - \frac{1}{n^2}\boldsymbol{\alpha}_k'\mathbf{X}_k^*(\frac{1}{n}\mathbf{X}_k\mathbf{X}_k^* - z\mathbf{I}_{p-1})^{-1}\mathbf{X}_k\overline{\boldsymbol{\alpha}}_k}, \tag{3.3.6}$$

where $\mathbf{X}_k$ is the matrix obtained from $\mathbf{X}$ with the k-th row removed and $\boldsymbol{\alpha}_k'$ $(n \times 1)$ is the k-th row of $\mathbf{X}$.

Set

$$\varepsilon_k = \frac{1}{n}\boldsymbol{\alpha}_k'\overline{\boldsymbol{\alpha}}_k - 1 - \frac{1}{n^2}\boldsymbol{\alpha}_k'\mathbf{X}_k^*\left(\frac{1}{n}\mathbf{X}_k\mathbf{X}_k^* - z\mathbf{I}_{p-1}\right)^{-1}\mathbf{X}_k\overline{\boldsymbol{\alpha}}_k + y_n + y_nz\mathrm{E}s_n(z), \tag{3.3.7}$$

where $y_n = p/n$. Then, by (3.3.6), we have

$$\mathrm{E}s_n(z) = \frac{1}{1 - z - y_n - y_nz\mathrm{E}s_n(z)} + \delta_n, \tag{3.3.8}$$

where

$$\delta_n = -\frac{1}{p}\sum_{k=1}^{p}\mathrm{E}\left(\frac{\varepsilon_k}{(1 - z - y_n - y_nz\mathrm{E}s_n(z))(1 - z - y_n - y_nz\mathrm{E}s_n(z) + \varepsilon_k)}\right). \tag{3.3.9}$$

Solving $\mathrm{E}s_n(z)$ from equation (3.3.8), we get two solutions:

$$\begin{aligned} s_1(z) &= \frac{1}{2y_nz}(1 - z - y_n + y_nz\delta_n + \sqrt{(1 - z - y_n - y_nz\delta_n)^2 - 4y_nz}), \\ s_2(z) &= \frac{1}{2y_nz}(1 - z - y_n + y_nz\delta_n - \sqrt{(1 - z - y_n - y_nz\delta_n)^2 - 4y_nz}). \end{aligned}$$

Comparing this with (3.3.1), it suffices to show that

$$\mathrm{E}s_n(z) = s_1(z) \tag{3.3.10}$$

and

$$\delta_n \to 0. \tag{3.3.11}$$

We show (3.3.10) first. Making $v \to \infty$, we know that $\mathrm{E}s_n(z) \to 0$ a[…] hence $\delta_n \to 0$ by (3.3.8). This shows that $\mathrm{E}s_n(z) = s_1(z)$ for all z with lar[…] imaginary part. If (3.3.10) is not true for all $z \in \mathbb{C}^+$, then by the continui[…] of s_1 and s_2, there exists a $z_0 \in \mathbb{C}^+$ such that $s_1(z_0) = s_2(z_0)$, which impl[…] that

$$(1 - z_0 - y_n + y_nz\delta_n)^2 - 4y_nz_0(1 + \delta_n(1 - z_0 - y_n)) = 0.$$

Thus,

$$\mathrm{E}s_n(z_0) = s_1(z_0) = \frac{1 - z_0 - y_n + y_n z_0 \delta_n}{2y_n z_0}.$$

Substituting the solution δ_n of equation (3.3.8) into the identity above, we obtain

$$\mathrm{E}s_n(z_0) = \frac{1 - z_0 - y_n}{y_n z_0} + \frac{1}{y_n + z_0 - 1 + y_n z_0 \mathrm{E}s_n(z_0)}. \tag{3.3.12}$$

Noting that for any Stieltjes transform $s(z)$ of probability F defined on $\mathbb{R}^+$ and positive y, we have

$$\begin{aligned}\Im(y + z - 1 + yzs(z)) &= \Im\left(z - 1 + \int_0^\infty \frac{yxdF(x)}{x - z}\right)\\ &= v\left(1 + \int_0^\infty \frac{yxdF(x)}{(x-u)^2 + v^2}\right) > 0. \end{aligned}\tag{3.3.13}$$

In view of this, it follows that the imaginary part of the second term in (3.3.12) is negative. If $y_n \le 1$, it can be easily seen that $\Im(1 - z_0 - y_n)/(y_n z_0) < 0$. Then we conclude that $\Im \mathrm{E}s_n(z_0) < 0$, which is impossible since the imaginary part of the Stieltjes transform should be positive. This contradiction leads to the truth of (3.3.10) for the case $y_n \le 1$.

For the general case, we can prove it in the following way. In view of (3.3.12) and (3.3.13), we should have

$$y_n + z_0 - 1 + y_n z_0 \mathrm{E}s_n(z_0) = \sqrt{y_n z_0}. \tag{3.3.14}$$

Now, let $\underline{s}_n(z)$ be the Stieltjes transform of the matrix $\frac{1}{n}\mathbf{X}^*\mathbf{X}$. Noting that $\frac{1}{n}\mathbf{X}^*\mathbf{X}$ and $\mathbf{S}_n = \frac{1}{n}\mathbf{X}\mathbf{X}^*$ have the same set of nonzero eigenvalues, we have the relation between s_n and $\underline{s}_n$ given by

$$s_n(z) = y_n^{-1}\underline{s}_n(z) - \frac{1 - 1/y_n}{z}.$$

Note that the equation above is true regardless of whether $y_n > 1$ or $y_n \le 1$. From this we have

$$y_n - 1 + y_n z_0 \mathrm{E}s_n(z_0) = z_0 \mathrm{E}\underline{s}_n(z_0).$$

Substituting this into (3.3.14), we obtain

$$1 + \mathrm{E}\underline{s}_n(z_0) = \sqrt{y}/\sqrt{z_0},$$

which leads to a contradiction that the imaginary part of LHS is positive and that of the RHS is negative. Then, (3.3.10) is proved.

Now, let us consider the proof of (3.3.11). Rewrite

$$\begin{aligned}\delta_n &= -\frac{1}{p}\sum_{k=1}^p \left(\frac{\mathrm{E}\varepsilon_k}{(1 - z - y_n - y_n z\mathrm{E}s_n(z))^2}\right)\\ &\quad + \frac{1}{p}\sum_{k=1}^p \mathrm{E}\left(\frac{\varepsilon_k^2}{(1 - z - y_n - y_n z\mathrm{E}s_n(z))^2(1 - z - y_n - y_n z\mathrm{E}s_n(z) + \varepsilon_k)}\right)\\ &= J_1 + J_2.\end{aligned}$$

At first, by assumptions given in (3.2.2), we note that

$$\begin{aligned}|\mathrm{E}\varepsilon_k| &= \left|-\frac{1}{n^2}\mathrm{E}\mathrm{tr}\mathbf{X}_k^*\left(\frac{1}{n}\mathbf{X}_k\mathbf{X}_k^* - z\mathbf{I}_{p-1}\right)^{-1}\mathbf{X}_k + y_n + y_n z\mathrm{E}s_n(z)\right|\\ &= \left|-\frac{1}{n}\mathrm{E}\mathrm{tr}\left(\frac{1}{n}\mathbf{X}_k\mathbf{X}_k^* - z\mathbf{I}_{p-1}\right)^{-1}\frac{1}{n}\mathbf{X}_k\mathbf{X}_k^* + y_n + y_n z\mathrm{E}s_n(z)\right|\\ &\le \frac{1}{n} + \frac{|z|y_n}{n}\mathrm{E}\left|\mathrm{tr}\left(\frac{1}{n}\mathbf{X}_k\mathbf{X}_k^* - z\mathbf{I}_{p-1}\right)^{-1} - s_n(z)\right|\\ &\le \frac{1}{n} + \frac{|z|y_n}{nv} \to 0,\end{aligned}\tag{3.3.15}$$

which implies that $J_1 \to 0$.

Now we prove $J_2 \to 0$. Since

$$\begin{aligned}&\Im(1 - z - y_n - y_n z\mathrm{E}s_n(z) + \varepsilon_k)\\ &= \Im\left(\frac{1}{n}\boldsymbol{\alpha}_k'\overline{\boldsymbol{\alpha}}_k - z - \frac{1}{n^2}\boldsymbol{\alpha}_k'\mathbf{X}_k^*\left(\frac{1}{n}\mathbf{X}_k\mathbf{X}_k^* - z\mathbf{I}_{p-1}\right)^{-1}\mathbf{X}_k\overline{\boldsymbol{\alpha}}_k\right)\\ &= -v\left(1 + \frac{1}{n^2}\boldsymbol{\alpha}_k'\mathbf{X}_k^*\left[\left(\frac{1}{n}\mathbf{X}_k\mathbf{X}_k^* - u\mathbf{I}_{p-1}\right)^2 + v^2\mathbf{I}_{p-1}\right]^{-1}\mathbf{X}_k\overline{\boldsymbol{\alpha}}_k\right) < -v,\end{aligned}$$

combining this with (3.3.13), we obtain

$$\begin{aligned}|J_2| &\le \frac{1}{pv^3}\sum_{k=1}^p \mathrm{E}|\varepsilon_k|^2\\ &= \frac{1}{pv^3}\sum_{k=1}^p [\mathrm{E}|\varepsilon_k - \widetilde{\mathrm{E}}(\varepsilon_k)|^2 + \mathrm{E}|\widetilde{\mathrm{E}}\varepsilon_k - \mathrm{E}(\varepsilon_k)|^2 + (\mathrm{E}(\varepsilon_k))^2],\end{aligned}$$

where $\widetilde{\mathrm{E}}(\cdot)$ denotes the conditional expectation given $\{\boldsymbol{\alpha}_j, j = 1, ..., k-1, k+1, ..., p\}$. In the estimation of J_1, we have proved that

$$|\mathrm{E}(\varepsilon_k)| \le \frac{1}{n} + \frac{|z|y}{nv} \to 0.$$

Write $\mathbf{A} = (a_{ij}) = \mathbf{I}_n - \frac{1}{n}\mathbf{X}_k^*(\frac{1}{n}\mathbf{X}_k\mathbf{X}_k^* - z\mathbf{I}_{p-1})^{-1}\mathbf{X}_k$. Then, we have

$$\varepsilon_k - \tilde{\mathrm{E}}\varepsilon_k = \frac{1}{n}\left(\sum_{i=1}^{n} a_{ii}(|x_{ki}|^2 - 1) + \sum_{i\neq j} a_{ij}x_{ki}\bar{x}_{kj}\right).$$

By elementary calculation, we have

$$\begin{aligned}
&\frac{1}{n^2}\tilde{\mathrm{E}}|\varepsilon_k' - \tilde{\mathrm{E}}\varepsilon_k|^2 \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n}|a_{ii}|^2(\mathrm{E}|x_{ki}|^4 - 1) + \sum_{i\neq j}[|a_{ij}|^2\mathrm{E}|x_{ki}|^2\mathrm{E}|x_{kj}|^2 + a_{ij}^2\mathrm{E}x_{ki}^2\mathrm{E}x_{kj}^2]\right) \\
&\le \frac{1}{n^2}\left(\sum_{i=1}^{n}|a_{ii}|^2(\eta_n^2 n) + 2\sum_{i\neq j}|a_{ij}|^2\right) \\
&\le \frac{\eta_n^2}{v^2} + \frac{2}{nv^2}.
\end{aligned}$$

Here, we have used the fact that $|a_{ii}| \le v^{-1}$.

Using the martingale decomposition method in the proof of (3.3.3), we can show that

$$\begin{aligned}
&\mathrm{E}|\widetilde{\mathrm{E}}\varepsilon_k - \mathrm{E}\varepsilon_k|^2 \\
&= \frac{|z|^2y^2}{n^2}\mathrm{E}\left|\mathrm{tr}\left(\frac{1}{n}\mathbf{X}_k\mathbf{X}_k^* - z\mathbf{I}_{p-1}\right)^{-1} - \mathrm{Etr}\left(\frac{1}{n}\mathbf{X}_k\mathbf{X}_k^* - z\mathbf{I}_{p-1}\right)^{-1}\right|^2 \\
&\le \frac{|z|^2y^2}{nv^2} \to 0.
\end{aligned}$$

Combining the three estimations above , we have completed the proof of the mean convergence of the Stieltjes transform of the ESD of $\mathbf{S}_n$.

Consequently, Theorem 3.10 is proved by the method of Stieltjes transforms.

Chapter 4
Product of Two Random Matrices

In this chapter, we shall consider the LSD of a product of two random m trices, one of them a sample covariance matrix and the other an arbitra Hermitian matrix. This topic is related to two areas: The first is the study the LSD of a multivariate F-matrix that is a product of a sample covarian matrix and the inverse of another sample covariance matrix, independent each other. Multivariate F plays an important role in multivariate data ana ysis, such as two-sample tests, MANOVA (multivariate analysis of varianc and multivariate linear regression. The second is the investigation of the LS of a sample covariance matrix when the population covariance matrix is a bitrary. The sample covariance matrix under a general setup is, as mention in Chapter 3, fundamental in multivariate analysis.

Pioneering work was done by Wachter [290], who considered the limiti distribution of the solutions to the equation

$$\det(\mathbf{X}_{1,n_1}\mathbf{X}_{1,n_1}' - \lambda\mathbf{X}_{2,n_2}\mathbf{X}_{2,n_2}') = 0, \quad (4.0.$$

where $\mathbf{X}_{j,n_j}$ is a $p \times n_j$ matrix whose entries are iid $N(0,1)$ and $\mathbf{X}_{1,}$ is independent of $\mathbf{X}_{2,n_2}$. When $\mathbf{X}_{2,n_2}\mathbf{X}_{2,n_2}'$ is of full rank, the solutio to (4.0.1) are n_2/n_1 times the eigenvalues of the multivariate F-matr $(\frac{1}{n_1}\mathbf{X}_{1,n_1}\mathbf{X}_{1,n_1}')(\frac{1}{n_2}\mathbf{X}_{2,n_2}\mathbf{X}_{2,n_2}')^{-1}$.

Yin and Krishnaiah [304] established the existence of the LSD of the m trix sequence $\{\mathbf{S}_n\mathbf{T}_n\}$, where $\mathbf{S}_n$ is a standard Wishart matrix of dimensi p and degrees of freedom n with $p/n \to y \in (0,\infty)$, $\mathbf{T}_n$ is a positive defini matrix satisfying $\beta_k(\mathbf{T}_n) \to H_k$, and the sequence H_k satisfies the Carl man condition (see (B.1.4)). In Yin [300], this result was generalized to t case where the sample covariance matrix is formed based on iid real rando variables of mean zero and variance one. Using the result of Yin and K ishnaiah [304], Yin, Bai, and Krishnaiah [302] showed the existence of t LSD of the multivariate F-matrix. The explicit form of the LSD of multiva ate F-matrices was derived in Bai, Yin, and Krishnaiah [40] and Silverste

Z. Bai and J.W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices, Second Edition*, Springer Series in Statistics, DOI 10.1007/978-1-4419-0661-8_4,