Central limit theorem for signal-to-interference ratio of reduced rank linear receiver
Pan, G.; Zhou, W.

Published in:
The Annals of Applied Probability

DOI:
10.1214/07-AAP477

Published: 01/01/2008

Document Version
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Let \( s_k = \frac{1}{\sqrt{N}} (v_{1k}, \ldots, v_{Nk})^T \), with \( \{v_{ik}, i, k = 1, \ldots\} \) independent and identically distributed complex random variables. Write \( S_k = (s_1, \ldots, s_{k-1}, s_{k+1}, \ldots, s_K) \), \( P_k = \text{diag}(p_1, \ldots, p_{k-1}, p_{k+1}, \ldots, p_K) \), \( R_k = (S_kP_kS_k^* + \sigma^2 I) \) and \( A_{km} = [s_k, R_k s_k, \ldots, R_k^{m-1} s_k] \). Define \( \beta_{km} = p_k s_k^* A_{km}(A_{km}^* \times R_k A_{km})^{-1} A_{km}^* s_k \), referred to as the signal-to-interference ratio (SIR) of user \( k \) under the multistage Wiener (MSW) receiver in a wireless communication system. It is proved that the output SIR under the MSW and the mutual information statistic under the matched filter (MF) are both asymptotic Gaussian when \( N/K \to c > 0 \). Moreover, we provide a central limit theorem for linear spectral statistics of eigenvalues and eigenvectors of sample covariance matrices, which is a supplement of Theorem 2 in Bai, Miao and Pan [Ann. Probab. 35 (2007) 1532–1572]. And we also improve Theorem 1.1 in Bai and Silverstein [Ann. Probab. 32 (2004) 553–605].

1. Introduction.

1.1. The signal-to-interference ratio (SIR) in engineering. Consider a synchronous direct-sequence code-division multiple-access (CDMA) system. Suppose that there are \( K \) users and that the dimension of the signature sequence \( s_k \) assigned to user \( k \) is \( N \). Let \( x_k \) denote the symbol transmitted by user \( k \), \( p_k \) the power of user \( k \) and \( n \in \mathbb{C}^N \) noise vector with mean zero and covariance matrix \( \sigma^2 I \). Suppose that \( x_k^* s_k \) are independent random variables (r.v.’s) with \( E x_k = 0 \) and \( E x_k^2 = 1 \) and that \( x_k^* s_k \) are independent of \( n \). The discrete time model for the received vector \( r \) is

\[
r = \sum_{k=1}^K \sqrt{p_k} x_k s_k + n.
\]

The goal in wireless communication is to estimate the transmitted \( x_k \) for each user in an appropriate receiver. For simplicity, in the sequel we are only interested
in linear receivers. A linear receiver, represented by a vector \( c_k \), estimates \( x_k \) in a form \( c_k^* r \) (the notation * denotes the complex conjugate transpose of a vector or matrix). The well known linear mean-square error (MMSE) minimizes

\[
E[|x_k - c_k^* r|^2].
\]

To evaluate the linear receivers, a popular performance measure is the output signal-to-interference ratio (SIR),

\[
\frac{p_k(c_k^* s_k)^2}{\sigma^2 c_k^* c_k + \sum_{j \neq k} p_j (c_k^* s_j)^2}
\]

(see Verdú [19] or Tse and Hanly [16]). Ideally, a good receiver should have a higher SIR.

Without loss of generality we focus only user 1. For MMSE receiver, from (1.2) one can solve \( c_1 = R_1^{-1} s_1 \) and then substitute \( c_1 \) into (1.3) to obtain the SIR expression for user 1 as

\[
\hat{\beta}_1 = p_1 s_1^* R_1^{-1} s_1,
\]

where \( R_1 = (S_1 P_1 S_1^* + \sigma^2 I) \), \( S_1 = (s_2, \ldots, s_K) \) and \( P_1 = \text{diag}(p_2, \ldots, p_K) \). It turns out that the choice of \( c_1 \) also maximizes user 1’s SIR. But since MMSE involves a matrix inverse this may be very costly when the spreading factor is high. Based on this reason, some simple and near MMSE performance receivers like reduced-rank linear receiver have been considered.

The basic idea behind a reduced rank is to project the received vector onto a lower dimensional subspace. For the multistage Wiener (MSW), the lower dimensional subspace has been described as a set of recursions by Goldstein, Reed and Scharf [7] and Honig and Xiao [10]. However, we would like to make use of another property of MSW given in Theorem 2 in Honig and Xiao [10] for our purpose, that is, MSW receiver estimates \( x_1 \) through MMSE after producing \( m \)-dimensional project vector \( A_1^* m r \) instead of \( r \), where \( m < n \) and

\[
A_1^* m = [s_1, R_1 s_1, \ldots, R_1^{m-1} s_1].
\]

Similar to (1.4), one can get \( c_1 m = (A_1^* m R_1 A_1 m)^{-1} A_1^* m s_1 \) and the output SIR

\[
\beta_1 m = p_1 s_1^* A_1 m (A_1^* m R_1 A_1 m)^{-1} A_1^* m s_1,
\]

which is the focus of this paper.

The MSW, as a kind of reduced-rank receiver, was first introduced by Goldstein, Reed and Scharf [7]. The receiver is widely employed in practice because the number of stages \( m \) needed to achieve a target SIR, unlike other reduced-rank receivers, does not scale with the system size, that is, dimensionality \( N \) of the system, as remarked by Honig and Xiao [10]. In their subsequent newsletter [11], the authors specially addressed this point. In addition, Honig and Xiao [10] showed
that the SIR of MSW converges to a deterministic limit in a large system. However, as we know, in a finite system, the SIR will fluctuate around the limit. Moreover, such fluctuation will lead to some important performance measures, such as error probability and outage probability. Regarding this promising receiver, we will characterize such fluctuation by providing central limit theorems in this paper.

From now on the signature sequences are modeled as random vectors, that is,

$$s_k = \frac{1}{\sqrt{N}}(v_{1k}, \ldots, v_{Nk})^T,$$

$k = 1, \ldots, K$, where \(\{v_{ik}, i, k = 1, \ldots\}\) are independent and identically distributed (i.i.d.) r.v.’s. Then the SIRs (1.6) may be further analyzed using the random matrices theory when \(K\) and \(N\) go to infinity with their ratio being a positive constant, which is well known as the large system analysis in the wireless communication field.

Tse and Hanly [16] and Verdú and Shamai [20] derived, respectively, the large system SIR and spectral efficiency under MMSE, Matched filter (MF) and decorrelator receiver. Tse and Zeitouni [17] proved that the distribution of SIR under MMSE is asymptotically Gaussian. Later, Bai and Silverstein [4] reported the asymptotic SIR under MMSE for a general model. For more progress in this area, one may see the review paper of Tulino and Verdú [18] and, in addition, refer to the review paper of Bai [2] concerning random matrices theory. Here we would also like to say a few words about our earlier work (Pan, Guo and Zhou [12]). In that paper, the random variables are assumed to be real and we could apply central limit theorems which have appeared in the literature. For example, we made use of main results from Götze and Tikhomirov ([8], page 426: considering real random variables with the sixth moment) and Bai and Silverstein [3] (requiring \(Ev_{11}^4 = 3\) or \(Ev_{11}^4 = 2\)). In the present work we develop a central limit theorem for the statistic of eigenvalues and eigenvectors under the finite fourth moment (see Theorem 1.3), which further gives a central limit theorem for a random quadratic form (see Remark 1.5). And we give a central limit theorem (see Theorem 1.4) for eigenvalues by dropping the assumption \(Ev_{11}^4 = 3\) or \(Ev_{11}^4 = 2\) in Bai and Silverstein [3]. For central limit theorems in other matrix models, we refer to [1].

Our main contribution to engineering is to prove that the distribution of the SIR under MSW, after scaling, is asymptotic Gaussian and that the sum of the SIRs for all users under MF \((m = 1)\), after subtracting a proper value, has a Gaussian limit, which further gives the asymptotic distribution of the sum mutual information under MF.

We introduce some notation before stating our results. Set \(R = (C + \sigma^2I)\), \(C = SPS^*\), \(S = (s_1, \ldots, s_K)\) and \(P = \text{diag}(p_1, \ldots, p_K)\). Suppose that \(F^{c,H}(x)\) and \(H(x)\), respectively, denote the weak limit of the empirical spectral distribution function \(F^{cN}SPS^*\) and \(H_N\) (i.e., \(F^{P}\)), where \(c_N = N/K\). In particular, \(F^{c,H}(x)\) becomes \(F^c(x)\) when \(P\) is the identity matrix, whose probability density was given in
Jonsson [6]. Let \( W^0(t) \) denote a Brownian bridge and \( X \) is independent of \( W^0(t) \), which is \( N(0, Ev^4_{11} - 1) \). Furthermore, let

\[
W_x = W^0(F^c(x)),
\]

\[
\zeta_i = \sum_{u=0}^{i} \binom{i}{u} (\sigma^2)^{i-u} \left( h_u X + \sqrt{2c} \int_{(1-\sqrt{c})^2}^{(1+\sqrt{c})^2} x^u dW^c_x \right),
\]

\( i = 1, \ldots, 2m - 1 \), and \( \zeta_0 = X \), with \( h_u = \int x^u dF^c(cx) \). Define \( a_m = \int (x + \sigma^2)^m dF^{cN,HN}(cx) \) and

\[
\mathbf{b} = (1, a_1, \ldots, a_{m-1})^T, \quad \mathbf{B} = \begin{pmatrix}
a_1 & a_2 & \cdots & a_m \\
a_2 & a_3 & \cdots & a_{m+1} \\
\vdots & \vdots & \ddots & \vdots \\
a_m & a_{m+1} & \cdots & a_{2m-1}
\end{pmatrix},
\]

where \( F^{cN,HN}(x) = F^{c,H}(x) \}_{c=cN,H=HN} \).

In what follows, with a slight abuse of notation, we still use \( a_m \) as a limit, such as (1.8) below, even when \( F^{cN,HN}(x) \) is replaced by \( F^{c,H}(x) \) in the expression of \( a_m \).

**Theorem 1.1.** Suppose that:

(a) \( \{v_{ij}, i, j = 1, \ldots, \} \) are i.i.d. complex r.v.’s with \( Ev_{11} = 0, Ev^2_{11} = 0, E|v_{11}|^2 = 1 \) and \( E|v_{11}|^4 < \infty \).

(b) \( c_N \to c > 0 \) as \( N \to \infty \).

(c) \( p_1 = \cdots = p_K = 1 \). Then, for any finite integer \( m \),

\[
\sqrt{N}(\beta_{1m} - \mathbf{b}^* \mathbf{B}^{-1} \mathbf{b}) \xrightarrow{D} y,
\]

where

\[
y = 2\zeta^* \mathbf{B}^{-1} \mathbf{b} - \mathbf{b}^* \mathbf{B}^{-1} \mathbf{D} \mathbf{B}^{-1} \mathbf{b},
\]

with \( \zeta^* = (\zeta_0, \ldots, \zeta_{m-1}) \) and \( \mathbf{D} = (d_{ij}) = (\zeta_{i+j-1}) \).

**Remark 1.1.** It can be verified that

\[
\text{Cov}\left( \int_{(1-\sqrt{c})^2}^{(1+\sqrt{c})^2} x^i dW^c_x, \int_{(1-\sqrt{c})^2}^{(1+\sqrt{c})^2} x^j dW^c_x \right)
\]

\[
= \int_{(1-\sqrt{c})^2}^{(1+\sqrt{c})^2} x^{i+j} dF^c(x)
\]

\[
- \int_{(1-\sqrt{c})^2}^{(1+\sqrt{c})^2} x^i dF^{c}(x) \int_{(1-\sqrt{c})^2}^{(1+\sqrt{c})^2} x^j dF^c(x).
\]

Moreover, \( X \) is independent of \( \int_{(1-\sqrt{c})^2}^{(1+\sqrt{c})^2} x^i dW^c_x \) and so the variance of \( y \) can be computed, although it is complicated.
The asymptotic distribution of the sum mutual information has been derived for MMSE by Pan, Guo and Zhou [12]. Thus, it is interesting to derive the corresponding asymptotic distribution of the MSW. But, unfortunately, it is rather complicated for the MSW case. At this stage, we can only derive the asymptotic distribution for the sum mutual information for the case \( m = 1 \), which is well known as the MF (see Verdú [19]).

Obviously, when \( m = 1 \), the output SIR for the MSW, \( \beta_{km} \) (the expressions for \( \beta_{km} \) can be derived similarly to \( \beta_{1m} \)), becomes

\[
\beta_k = \frac{p_k (s_k^* s_k)^2}{s_k^* R_k s_k},
\]

with \( R_k = C_k + \sigma^2 I \) and \( C_k = S_k P_k S_k^* \), where \( S_k \) and \( P_k \) are respectively obtained from \( S \) and \( P \) by deleting the \( k \)th column (here we denote \( \beta_{k1} \) by \( \beta_k \)).

**Theorem 1.2.** Suppose that:

(a) \( \{v_{ij}, i, j = 1, \ldots, \} \) are i.i.d. complex r.v.’s with \( E v_{11} = 0, E v_{11}^2 = 0, E|v_{11}| = 1 \) and \( E|v_{11}|^4 < \infty \).

(b) The empirical distribution function of power matrix \( P \) converges weakly to some distribution function \( H(t) \) with all the powers bounded by some constant.

(c) \( c_N \to c > 0 \) as \( N \to \infty \). Then

\[
\sum_{k=1}^{K} \left( \beta_k - \frac{p_k}{\sigma^2 + c} \right) \xrightarrow{D} N(\mu, \tau^2)
\]

with \( p_1 = \cdots = p_K = 1 \),

\[
\mu = \frac{2E|v_{11}|^4 - 3}{c(\sigma^2 + 1/c)^2} + \frac{1}{c^2(\sigma^2 + 1/c)^3},
\]

and \( \tau \) defined in \( (5.34) \).

We would like to point out that the result has been given only for the equal power case \( (p_1 = \cdots = p_K = 1) \) in Theorem 1.2, although the assumptions are concerning different powers. As will be seen, the main difficulty of the different powers case is that matrices \( (SPS^*)^2 \) and \( SP^2S^* \) have different eigenvalues. But, it is worth pointing out that one may establish a central limit theorem for

\[
\sum_{j=1}^{N} (f(\lambda_j) + g(\mu_j))
\]

following a similar line of Bai and Silverstein [3], where \( f, g \) are analytical functions and \( \lambda_j, \mu_j \) denote the eigenvalues of \( P^{1/2}S^*SP^{1/2} \) and \( PS^*SP \), respectively. We do not intend to pursue this direction since the process is lengthy.

Concerning the sum mutual information under the MF, we have the following:
COROLLARY 1.1. Under the conditions of Theorem 1.2,

\begin{equation}
\sum_{k=1}^{K} \left( \log(1 + \beta_k) - \log \left( 1 + \frac{1}{\sigma^2 + c} \right) \right) \xrightarrow{D} N(\mu_1, \tau_1^2)
\end{equation}

with

\[
\mu_1 = \frac{\mu}{1 + (c^{-1} + \sigma^2)^{-1}}
\]

\[
- \frac{2(E|v_{11}|^4 - 2)(c^{-1} + \sigma^2)^2 + 2c^{-1}(1 + c^{-1}) + \sigma^4 + 2\sigma^2c^{-1}}{c(c^{-1} + \sigma^2)^4(1 + (c^{-1} + \sigma^2)^{-1})^2}
\]

and

\[
\tau_1^2 = \frac{\tau^2}{(1 + (c^{-1} + \sigma^2)^{-1})^2}.
\]

1.2. Random matrices. Random matrices have been used in wireless communication since Grant and Alexander’s 1996 conference presentation [9] and it has proved to be a very powerful technique. To prove the preceding theorems, we develop a central limit theorem for the eigenvalues and eigenvectors of the sample covariance matrices, which is a supplement of Theorem 2 in Bai, Miao and Pan [5]. And we also improve Theorem 1.1 in Bai and Silverstein [3]. Obviously, these central limit theorems are interesting themselves.

Let \( c_N T_N^{1/2} S S^* T_N^{1/2} = A_N \) with \( T_N^{1/2} \) being the square root of a nonnegative definite matrix \( T_N \) and \( U_N \Lambda_N U_N^* \) be the spectral decomposition of \( A_N \), where \( \Lambda_N = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N) \), \( U_N = (u_{ij}) \) is a unitary matrix consisting of the orthonormal eigenvectors of \( A_N \). Suppose that \( x_N = (x_{N1}, \ldots, x_{NN})^T \in \mathbb{C}^N, \|x_N\| = 1 \), is nonrandom and \( y = (y_1, y_2, \ldots, y_N)^T = U_N^* x_N \). Let \( F_{A_N}^* \) denote the empirical spectral distribution (ESD) of the matrix \( A_N \) and \( F_{1}^{A_N}(x) \) another ESD of \( A_N \), that is,

\begin{equation}
F_{1}^{A_N}(x) = \sum_{i=1}^{N} |y_i|^2 I(\lambda_i \leq x).
\end{equation}

Let

\[
G_N(x) = \sqrt{N}(F_{A_N}^*(x) - F_{c_N, H}(x)),
\]

and \( m(z) = m_{F_{c_N, H}}(z) \) denote the Stieltjes transform of the limiting empirical distribution function of \( c_N S^* T_N S \). Now it is time to state the following theorem.

THEOREM 1.3. Assume:

(1) \( v_{ij}, i, j = 1, 2, \ldots, \) are i.i.d. with \( E v_{11} = 0, E|v_{11}|^2 = 1 \) and \( E|v_{11}|^4 < \infty \), and \( \lim_{N \to \infty} c_N = c \in (0, \infty) \).
\[ \begin{align*}
(2) & \quad x_N \in \{ x \in \mathbb{C}^N, \| x \|=1 \}. \\
(3) & \quad T_N \text{ is nonrandom Hermitian nonnegative definite such that its spectral norm is bounded in } N, \quad H_N = F^{T_N} \to H, \text{ a proper distribution function and } x_N^*(T - zI)^{-1}x_N \to m_{FH}(z), \text{ where } m_{FH}(z) \text{ denotes the Stieltjes transform of } H(t). \\
(4) & \quad g_1, \ldots, g_k \text{ are defined and analytic on an open region } \mathcal{D} \text{ of the complex plane, which contains the real interval}
\end{align*} \]

\[ \lim_{N} \lambda_{\min}^{T_N}(0,1)(c)(1 - \sqrt{c})^2, \quad \lim_{N} \lambda_{\max}^{T_N}(1 + \sqrt{c})^2, \]

where \( \lambda_{\min}^{T_N} \) and \( \lambda_{\max}^{T_N} \) denote, respectively, the minimum and maximum eigenvalues of \( T_N \).

\[ \sup_{z} \sqrt{N} \left| x_N^* \left( m_{F^{c_N},H_N}(z)T_N + I \right)^{-1}x_N - \int \frac{1}{m_{F^{c_N},H_N}(z)t + 1} dH_N(t) \right| \to 0, \]

as \( n \to \infty \).

\[ \max_{i} \left| e_i^*T_N^{1/2}(zm(z)T_N + zI)^{-1}x_N \right| \to 0, \]

where \( e_i \) is the \( N \times 1 \) column vector with the \( i \)th element being 1 and the rest being 0. Then the following conclusions hold:

(a) If \( v_{11} \) and \( T_N \) are real, the random vector \( (\int g_1(x)dG_N(x), \ldots, \int g_k(x)dG_N(x)) \) converges weakly to a Gaussian vector \( (X_{g_1}, \ldots, X_{g_k}) \), with mean zero and covariance function

\[ \text{Cov}(X_{g_1}, X_{g_2}) = -\frac{1}{2\pi^2} \int_{C_1} \int_{C_2} g_1(z_1)g_2(z_2) \]

\[ \times \frac{(z_2m(z_2) - z_1m(z_1))^2}{c^2z_1z_2(z_2 - z_1)(m(z_2) - m(z_1))} \, dz_1 \, dz_2. \]

The contours \( C_1 \) and \( C_2 \) in the above equality are disjoint, both contained in the analytic region for the functions \( (g_1, \ldots, g_k) \) and both enclosing the support of \( F^{c_n,H_n} \) for all large \( n \).

(b) If \( v_{11} \) is complex, with \( Ev_{11}^2 = 0 \), then the conclusion (a) still holds, but the covariance function reduces to half of the quantity given in (1.15).

Remark 1.2. It is under the assumption \( Ev_{11}^4 = 3 \) in the real case or \( E|v_{11}|^4 = 2 \) in the complex case that Bai, Miao and Pan [5] obtained the above cen-
tral limit theorem. But, when $Ev_{11}^4 \neq 3$ in the real case, there exist sequences $\{x_n\}$ such that
\[
\left( \int x \, dG_N(x), \int x^2 \, dG_N(x) \right)
\]
fails to converge in distribution, as pointed out in Silverstein [13]. Therefore, when $Ev_{11}^4 \neq 3$ in the real case or $E|v_{11}|^4 \neq 2$ in the complex case, to guarantee the central limit theorem, we here impose an additional condition (6), which is implied by
\[
\max_i |x_{Ni}| \to 0,
\]
when $T_N$ becomes a diagonal matrix. Thus, the variance is dependent on the fourth moment of $v_{11}$.

REMARK 1.3. Let $g_1(x) = x, g_2(x) = x^2, \ldots, g_k(x) = x^k$. Then
\[
\sqrt{N} \left( \left( x_N^* A_N x_N - \int x \, dF_{cn,HN}(x) \right), \ldots, \left( x_N^* A_N^k x_N - \int x^k \, dF_{cn,HN}(x) \right) \right)
\]
converges weakly to a Gaussian vector, which is used when proving Theorem 1.1.

To derive Theorem 1.2, we would like to present a central limit theorem for the eigenvalues, which is a little improvement of Theorem 1.1 in Bai and Silverstein [3]. Define
\[
L_N(x) = N \left( F_{A_N}(x) - F_{cn,HN}(x) \right).
\]

THEOREM 1.4. In addition to the assumptions (1), (3) and (4) in Theorem 1.3 [remove the assumption concerning $x_N^*(T_N - zI)^{-2}x_N$ in (3)], suppose that
\[
\frac{1}{N} \sum_{i=1}^{N} e_i^* T_N^{1/2} (m(z_1)T_N + I)^{-1} T_N^{1/2} e_i e_i^* T_N^{1/2} (m(z_2)T_N + I)^{-1} T_N^{1/2} e_i \to h_1(z_1, z_2)
\]
(1.17)
and
\[
\frac{1}{N} \sum_{i=1}^{N} e_i^* T_N^{1/2} (m(z)T_N + I)^{-1}
\]
(1.18)
\[
\times T_N^{1/2} e_i e_i^* T_N^{1/2} (m(z)T_N + I)^{-2} T_N^{1/2} e_i \to h_2(z).
\]
Then the following conclusions hold:
(a) If $v_{11}$ and $T_N$ are real, then $(\int g_1(x) dL_N(x), \ldots, \int g_k(x) dL_N(x))$ converges weakly to a Gaussian vector $(X_{g_1}, \ldots, X_{g_k})$, with mean

$$EX_g = -\frac{1}{2\pi i} \int g(z) \frac{c \int m^3(z) t^2 dH(t)/(1 + tm(z))^3}{(1 - c \int m^2(z) t^2 dH(t)/(1 + tm(z))^2)^2} \, dz$$

and covariance function

$$\text{Cov}(X_{g_1}, X_{g_2}) = -\frac{1}{4\pi^2} \int \int \frac{g_1(z_1) g_2(z_2)}{\left( m(z_1) - m(z_2) \right)^2} \frac{d^2}{dz_1} m(z_1) \frac{d}{dz_2} m(z_2) \, dz_1 \, dz_2$$

(b) If $v_{11}$ is complex with $Ev_{11}^2 = 0$, then (a) holds as well, but the mean is now

$$EX_g = -\frac{E|v_{11}|^4 - 2}{2\pi i} \int g(z) \frac{c m^3(z) h_2(z)}{1 - c \int m^2(z) t^2 dH(t)/(1 + tm(z))^2} \, dz$$

and covariance function

$$\text{Cov}(X_{g_1}, X_{g_2}) = -\frac{c(E|v_{11}|^4 - 2)}{4\pi^2} \int \int \frac{g_1(z_1) g_2(z_2)}{\left( m(z_1) - m(z_2) \right)^2} \frac{d^2}{dz_1} m(z_1) \frac{d}{dz_2} m(z_2) \, dz_1 \, dz_2$$

**Remark 1.4.** When $T_N$ is a diagonal matrix,

$$h_2(z) = \int \frac{t^2 dH(t)}{(m(z)t + 1)^3},$$

$$h_1(z_1, z_2) = \int \frac{t^2 dH(t)}{(m(z_1)t + 1)(m(z_2)t + 1)}.$$
\( g(x) = x^r, \)

\[
\frac{1}{2\pi i} \int g(z) \frac{cm^3(z)h_2(z)}{1 - c \int m^2(z)t^2 dH(t)/(1 + tm(z))} \, dz
\]

\[
= c^{1+r} \sum_{j=0}^{r} \binom{r}{j} \left( \frac{1-c}{c} \right)^j \left( \frac{2r-j}{r-1} \right)
\]

\[
- c^{1+r} \sum_{j=0}^{r} \binom{r}{j} \left( \frac{1-c}{c} \right)^j \left( \frac{2r+1-j}{r-1} \right),
\]

and when \( g_1(x) = x^{r_1} \) and \( g_2(x) = x^{r_2}, \)

\[
- \frac{c}{4\pi^2} \int \int g_1(z_1)g_2(z_2) \frac{d^2}{dz_1 dz_2} \left[ m(z_1)m(z_2)h_1(z_1, z_2) \right] \, dz_1 \, dz_2
\]

\[
= c^{r_1+r_2+1} \sum_{j_1=0}^{r_1} \sum_{j_2=0}^{r_2} \binom{r_1}{j_1} \binom{r_2}{j_2} \left( \frac{1-c}{c} \right)^{j_1+j_2}
\]

\[
\times \left( \frac{2r_1-j_1}{r_1-1} \right) \left( \frac{2r_2-j_2}{r_2-1} \right)
\]

**Remark 1.5.** In applying Theorem 1.4 to Theorem 1.2, we take \( g_1(x) = x + x^2, \) that is, one needs to transform (1.11) into

\[
\sum_{j=1}^{n} (\lambda_j + \lambda_j^2) + u_n,
\]

where the term \( u_n \) will be proved to converge to some constant in probability. Indeed, when using Theorem 1.3 or Theorem 1.4, \( g_1(x) \) is usually taken to be a polynomial function.

The rest of this paper is organized as follows. The proofs of Theorem 1.3 and Theorem 1.1 are given in Sections 2 and 3, respectively. Section 4 includes the argument of Theorem 1.4. Section 5 establishes Theorem 1.2, while the truncation of the underlying r.v.’s. is postponed until the Appendix. Section 6 establishes Corollary 1.1. Throughout this paper, to save notation, \( M \) may denote different constants on different occasions.

**2. Proof of Theorem 1.3.** Let \( A(z) = A_N - z I, A_j(z) = A(z) - s_j s_j^*, \) With a slight abuse of notation, here and in the argument of Theorem 1.4, we use \( s_j \) to denote the \( j \)th column of \( c_N^{1/2} T_N^{1/2} S, \) as in Bai, Miao and Pan [5], but one should note that this \( s_j \) is different from one of other parts. To complete the proof of
where \( E_j = E(\cdot | F_j), F_j = \sigma(s_1, \ldots, s_j) \) and

\[
\mathbf{H}_{nj}(z) = T_N^{1/2} A_j^{-1}(z) \mathbf{x}_n \mathbf{x}_n^* A_j^{-1}(z) T_N^{1/2}.
\]

Define

\[
A_{jk}(z) = A(z) - s_j s_j^* - s_k s_k^*, \quad \varepsilon_k(z) = \beta_{jk}(z) A_j^{-1}(z) s_k s_k^* A_j^{-1}(z),
\]

\[
E \hat{H}_{nj}(z) = T_N^{1/2} E A_j^{-1}(z) \mathbf{x}_n \mathbf{x}_n^* E A_j^{-1}(z) T_N^{1/2}, \quad \beta_{jk}(z) = \frac{1}{1 + s_k^* A_{jk}(z) s_k}.
\]

It is observed that

\[
e_j^* T_N^{1/2} (A_j^{-1}(z_1) - E A_j^{-1}(z_1)) \mathbf{x}_n \mathbf{x}_n^* A_j^{-1}(z_1) T_N^{1/2} e_i
\]

\[
= e_j^* T_N^{1/2} (A_j^{-1}(z_1) - E A_j^{-1}(z_1)) \mathbf{x}_n \mathbf{x}_n^* (A_j^{-1}(z_1) - E A_j^{-1}(z_1)) T_N^{1/2} e_i
\]

\[
+ e_j^* T_N^{1/2} (A_j^{-1}(z_1) - E A_j^{-1}(z_1)) \mathbf{x}_n \mathbf{x}_n^* E A_j^{-1}(z) T_N^{1/2} e_i
\]

\[
= \sum_{k_1, k_2 = 1}^K e_j^* T_N^{1/2} (E_{k_1} A_j^{-1}(z_1) - E_{k_1-1} A_j^{-1}(z_1)) \mathbf{x}_n
\]

\[
\times \mathbf{x}_n^* (E_{k_2} A_j^{-1}(z_1) - E_{k_2-1} A_j^{-1}(z_1)) T_N^{1/2} e_i
\]

\[
+ \sum_{k = 1}^K e_j^* T_N^{1/2} (E_k A_j^{-1}(z_1) - E_{k-1} A_j^{-1}(z_1)) \mathbf{x}_n \mathbf{x}_n^* E A_j^{-1}(z_1) T_N^{1/2} e_i
\]

\[
= \sum_{k_1 \neq j, k_2 \neq j} e_j^* T_N^{1/2} ((E_{k_1} - E_{k_1-1}) \varepsilon_{k_1}(z_1)) \mathbf{x}_n
\]

\[
\times \mathbf{x}_n^* ((E_{k_2} - E_{k_2-1}) \varepsilon_{k_2}(z_1)) T_N^{1/2} e_i
\]

\[
- \sum_{k \neq j} e_j^* T_N^{1/2} ((E_k - E_{k-1}) \varepsilon_k(z_1)) \mathbf{x}_n \mathbf{x}_n^* E A_j^{-1}(z_1) T_N^{1/2} e_i.
\]

This, together with the Burkholder inequality and (4.4) in Bai, Miao and Pan [5],
gives

\[
\begin{align*}
\sum_{i=1}^{N} E_j (H_{nj}(z_1) - T_N^{1/2}(E|A_j^{-1}(z_1))x_n x_n^{*} A_j^{-1}(z_1)T_N^{1/2})_{ii} (E_j H_{nj}(z_2))_{ii} \end{align*}
\]

\[
\leq \sum_{i=1}^{N} E|(H_{nj}(z_1) - T_N^{1/2}(E|A_j^{-1}(z_1))x_n x_n^{*} A_j^{-1}(z_1)T_N^{1/2})_{ii}|^2
\]

\[
\times \sum_{i=1}^{N} E|(H_{nj}(z_2))_{ii}|^2
\]

\[
\leq M \sum_{i=1}^{N} \left[ E \left( \sum_{k \neq j} e_i^{*} T_N^{1/2} ((E_k - E_{k-1}) \epsilon_k(z_1)) x_n \right)^4 \right]^{1/2}
\]

\[
\times \left[ E \left( \sum_{k \neq j} x_n^{*} ((E_k - E_{k-1}) \epsilon_k(z_1)) T_N^{1/2} e_i \right)^4 \right]^{1/2}
\]

\[
+ M \sum_{i=1}^{N} |x_n^{*} E A_j^{-1}(z_1)T_N^{1/2} e_i|^2 E \left( \sum_{k \neq j} e_i^{*} T_N^{1/2} ((E_k - E_{k-1}) \epsilon_k(z_1)) x_n \right)^2
\]

\[
\leq M \varepsilon_N^4 + \frac{M}{N},
\]

which implies

\[
\frac{1}{K} \sum_{j=1}^{K} \sum_{i=1}^{N} E_j (H_{nj}(z_1) - T_N^{1/2}(E|A_j^{-1}(z_1))x_n x_n^{*} A_j^{-1}(z_1)T_N^{1/2})_{ii}
\]

\[
\times (E_j H_{nj}(z_2))_{ii} \xrightarrow{i.p.} 0.
\]

Similarly, one can also prove that

\[
\frac{1}{K} \sum_{j=1}^{K} \sum_{i=1}^{N} (T_N^{1/2}(E|A_j^{-1}(z_1))x_n x_n^{*} E_j (A_j^{-1}(z_1) - E A_j^{-1}(z_1))T_N^{1/2})_{ii}
\]

\[
\times (E_j H_{nj}(z_2))_{ii} \xrightarrow{i.p.} 0
\]

and, therefore,

\[
\frac{1}{K} \sum_{j=1}^{K} \sum_{i=1}^{N} E_j (H_{nj}(z_1) - E \hat{H}_{nj}(z_1))_{ii} (E_j H_{nj}(z_2))_{ii} \xrightarrow{i.p.} 0.
\]
Via an analogous argument,
\[
\frac{1}{K} \sum_{j=1}^{K} \sum_{i=1}^{N} (E \hat{H}_{nj}(z_1))_{ii} E_j(H_{nj}(z_2) - E \hat{H}_{nj}(z_2))_{ii} \overset{i.p.}{\longrightarrow} 0.
\]
Thus, for the proof of (2.1), it is sufficient to show that
\[
\sum_{i=1}^{N} (E \hat{H}_{n1}(z_1))_{ii} (E \hat{H}_{n1}(z_2))_{ii} \overset{i.p.}{\longrightarrow} 0.
\]
(2.3)

To this end, write
\[
A_1(z) = -\left( -\hat{T}_N(z) \right)^{-1} \sum_{k=2}^{K} s_k s_k^* + (-zEm_n(z))T_N,
\]
where \( m_n(z) \) denotes the Stieltjes transform of \( \frac{N}{K} S_1^T T_{NS_1} \) and \( \hat{T}_N(z) = zEm_n(z) \times T_N + zI \). Using equality, similar to (2.2) of Silverstein [15],
\[
m_n(z) = -\frac{1}{z} \sum_{k=2}^{K} \beta_{1k}(z),
\]
(2.4)

we get
\[
EA_1^{-1}(z) - (-\hat{T}_N(z))^{-1} = \left( \hat{T}_N(z) \right)^{-1} E \left[ \left( \sum_{k=2}^{K} s_k s_k^* - (-zEm_n(z))T_N \right) A_1^{-1}(z) \right]
\]
(2.5)

It follows that
\[
e_i^* T_{N}^{1/2} EA_1^{-1}(z)x_n - e_i^* T_{N}^{1/2} (-\hat{T}_N(z))^{-1} x_n
\]
\[
= (K - 1) E \left[ \beta_{12}(z) \left( s_2^* A_{12}^{-1}(z)x_n e_i^* T_{N}^{1/2} (\hat{T}_N(z))^{-1} s_2 \right) \right.
\]
\[
- \frac{1}{K} e_i^* T_{N}^{1/2} (\hat{T}_N(z))^{-1} T_N EA_1^{-1}(z)x_n \left. \right]
\]
(2.6)

where
\[
\rho_1 = (K - 1) E[\beta_{12}(z)b_{12}(z)\xi(z)\alpha(z)],
\]
\[
\rho_2 = \frac{K - 1}{K} E[\beta_{12}(z)e_i^* T_{N}^{1/2} (\hat{T}_N(z))^{-1} T_N (A_{12}^{-1}(z) - A_1^{-1}(z))x_n]?
\]
and
\[ \rho_3 = \frac{K - 1}{K} E[\beta_{12}(z)e_i^*T_N^{1/2}(\hat{T}_N(z))^{-1}T_N(A_1^{-1}(z) - EA_1^{-1}(z))x_n]. \]

Here we also set
\[ \xi(z) = s_2^*A_{12}^{-1}(z)x_n e_i^*T_N^{1/2}(\hat{T}_N(z))^{-1}s_2 \]
\[ - \frac{1}{K} e_i^*T_N^{1/2}(\hat{T}_N(z))^{-1}T_NA_{12}^{-1}(z)x_n \]
and
\[ \alpha(z) = s_2^*A_{12}^{-1}(z)s_2 - \frac{1}{K} \text{Tr}A_{12}^{-1}(z), \quad b_{12}(z) = \frac{1}{1 + (1/K) \text{Tr}A_{12}^{-1}(z)}. \]

According to (4.2) and (4.3) in Bai, Miao and Pan [5], one can conclude that
\[ \max_i |\rho_1| = O(K^{-1/2}), \]
\[ \max_i |\rho_2| = \max_i \frac{K - 1}{K} |E[\beta_{12}^2(z)e_i^*T_N^{1/2}(\hat{T}_N(z))^{-1} \times T_NA_{12}^{-1}(z)s_2s_2^*A_{12}^{-1}(z)x_n]| \]
\[ = O(K^{-1}) \]
and
\[ \max_i |\rho_3| = \max_i \left| \frac{K - 1}{K} E[\beta_{12}(z)b_{12}(z)\alpha(z)e_i^*T_N^{1/2}(\hat{T}_N(z))^{-1} \times T_N(A_1^{-1}(z) - EA_1^{-1}(z))x_n] \right| \]
\[ = O(K^{-1/2}). \]

Hence,
\[ \max_i |e_i^*T_N^{1/2}EA_1^{-1}(z)x_n| \to 0, \]
which, together with the Hölder inequality, guarantees (2.3). Thus, we are done.

3. Proof of Theorem 1.1. It is easy to show that
\[ s_1^*R_1^m s_1 - a_m \overset{i.p.}{\to} 0. \]
It follows that
\[ (3.1) \quad s_1^*A_{1m} - b^* \overset{i.p.}{\to} 0, \quad A_{1m}^*R_1^m A_{1m} - B \overset{i.p.}{\to} 0. \]
It is then observed that
\[
\sqrt{N}(\beta_{1m} - b^*B^{-1}b) = \sqrt{N}(s_1^*A_{1m} - b^*)(A_{1m}^*R_1A_{1m})^{-1}A_{1m}^*s_1 \\
+ \sqrt{N}b^*(A_{1m}^*R_1A_{1m})^{-1}(A_{1m}^*s_1 - b) \\
+ \sqrt{N}b^*((A_{1m}^*R_1A_{1m})^{-1} - B^{-1})b \\
= 2\sqrt{N}(s_1^*A_{1m} - b^*)B^{-1}b \\
- \sqrt{N}b^*B^{-1}(A_{1m}^*R_1A_{1m} - B)B^{-1}b + o_p(1),
\]
(3.2)
where we use (3.1), (3.6) below and an identity
\[
B_1^{-1} - B_2^{-1} = -B_1^{-1}(B_1 - B_2)B_2^{-1},
\]
which holds for any invertible matrices \(B_1\) and \(B_2\). Furthermore, let
\[
b^*B^{-1} = (d_1, \ldots, d_m),
\]
then (3.2) is now equal to
\[
2\sqrt{N} \sum_{i=1}^{m} d_i (s_1^*R_i^{-1}s_1 - a_{i-1}) - \sqrt{N} \sum_{i,j=1}^{m} d_id_j (s_1^*H_u^{-1}s_1 - a_{i-1}) - N \sum_{i=1}^{m} d_i (s_1^*H_u^i s_1 - a_{i-1}).
\]
(3.3)

By the result (1) of Theorem 1.1 of Bai and Silverstein [3], it is easily seen that
\[
\sqrt{N} \left( \frac{1}{N} \text{Tr} R_i^i - a_i \right) \rightarrow i.p. 0.
\]

To derive a central limit theorem for (3.3), it then suffices to develop a multivariate one for \(\{\sqrt{N}(s_1^*R_i^i s_1 - \frac{1}{N} \text{Tr} R_i^i), i = 0, \ldots, 2m - 1\}\).

Set \(H_1 = \mathbf{S}_1\mathbf{S}_1^*\) and \(h_m = \int x^m dF_{\mathbb{C}^N}(cx)\). Note that
\[
\sqrt{N} \left( s_1^*H_u^i s_1 - \frac{1}{N} \text{Tr} H_u^i \right) = \sum_{u=0}^{i} \binom{i}{u} (\sigma^2)^{i-u} \sqrt{N} \left( s_1^*H_u^i s_1 - \frac{1}{N} \text{Tr} H_u^i \right).
\]
(3.4)

Let \(\|s_1\|^2 = \sum_{i=1}^{N} |v_{1i}|^2 / N\). Write
\[
\sqrt{N} \left( s_1^*H_u^i s_1 - \frac{1}{N} \text{Tr} H_u^i \right) \\
= \sqrt{N} \|s_1\|^2 \left( \frac{s_1^*H_u^i s_1}{\|s_1\|^2} - \frac{1}{N} \text{Tr} H_u^i \right) + \sqrt{N} \frac{1}{N} \text{Tr} H_u^i (\|s_1\|^2 - 1).
\]

It is easy to check that
\[
\max_i \left| \frac{v_{1i}/\sqrt{N}}{\|s_1\|} \right| \rightarrow i.p. 0.
\]
Therefore, given $s_1$, it follows from Theorem 1.3 that

\[
\sqrt{N}\left(\frac{s_1^* H_1 s_1}{\|s_1\|^2} - \frac{1}{N} \text{Tr} H_1\right), \ldots, \sqrt{N}\left(\frac{s_1^* H_1 s_1}{\|s_1\|^2} - \frac{1}{N} \text{Tr} H_1^u\right) \xrightarrow{D} \sqrt{2}\left(\int_{(1+\sqrt{c})^2}^{(1-\sqrt{c})^2} x^i \frac{dW^c}{c}, \ldots, \int_{(1-\sqrt{c})^2}^{(1+\sqrt{c})^2} x^u \frac{dW^c}{c}\right)
\]

(regarding the formula, one may refer to Bai, Miao and Pan [5] or Silverstein [13, 14]). However, it is evident that

\[
\sqrt{N}\left(\|s_1\|^2 - 1\right) \xrightarrow{D} X,
\]

where $X \sim N(0, E|v_{11}|^4 - 1)$. Consequently, by the independence of $s_1$ and $H_1$,

\[
\left(\sqrt{N}(s_1^* s_1 - 1), \ldots, \sqrt{N}\left(s_1^* H_1^{2m-1} s_1 - \frac{1}{N} \text{Tr} H_1^{2m-1}\right)\right) \xrightarrow{D} (\xi_0, \ldots, \xi_{2m-1}),
\]

where $\xi_i = h_i X + \frac{\sqrt{N}}{c^i} \int_{(1-\sqrt{c})^2}^{(1+\sqrt{c})^2} x^i dW^c, i = 1, \ldots, 2m - 1$, and $\xi_0 = X$. Then

\[
\left(\sqrt{N}(s_1^* s_1 - 1), \ldots, \sqrt{N}\left(s_1^* R_1^{2m-1} s_1 - \frac{1}{N} \text{Tr} R_1^{2m-1}\right)\right) \xrightarrow{D} (\xi_0, \ldots, \xi_{2m-1}),
\]

where $\xi_i = \sum_{u=0}^{t} \binom{i}{u} (\sigma^2)^{i-u} h_u$.

It follows that

\[
\sqrt{N}(\beta_{1m} - b^* B^{-1} b) \xrightarrow{D} 2 \sum_{i=1}^{m} d_i \xi_i - \sum_{i=1}^{m} \sum_{j=1}^{m} d_i d_j \xi_{i+j-1}.
\]

Thus, we are done.

4. **Proof of Theorem 1.4.** By the argument of Bai and Silverstein [3], it suffices to find the limits of the following sums:

\[
\frac{1}{K^2} \sum_{j=1}^{K} \sum_{i=1}^{N} E_j (T^{1/2}_N A_j^{-1}(z_1) T^{1/2}_N)_{ii} E_j (T^{1/2}_N A_j^{-1}(z_2) T^{1/2}_N)_{ii}
\]

and

\[
\frac{1}{K} \sum_{i=1}^{N} E[(T^{1/2}_N A_j^{-1}(z) T^{1/2}_N)_{ii} (T^{1/2}_N A_j^{-1}(z) (T_N(z))^{-1} T^{1/2}_N)_{ii}]
\]

(see (2.7) and (4.10) in Bai and Silverstein [3]).
Similar to (2.2), it can be verified that
\[
\frac{1}{K^2} \sum_{j=1}^{K} \sum_{i=1}^{N} E_j(T_N^{1/2} A_{ij}^{-1}(z_1)T_N^{1/2} - E(T_N^{1/2} A_{ij}^{-1}(z_1)T_N^{1/2}))_{ii} \\
\times E_j(T_N^{1/2} A_{ij}^{-1}(z_2)T_N^{1/2})_{ii} = O_p(N^{-1/2}).
\]
Consequently, analogous to Theorem 1.3, it remains to find the limit of

\[
\frac{1}{K} \sum_{i=1}^{N} E(T_N^{1/2} A_{1i}^{-1}(z_1)T_N^{1/2})_{ii} E(T_N^{1/2} A_{1i}^{-1}(z_2)T_N^{1/2})_{ii}.
\]

Define
\[
\gamma(z) = s_k^* A_{1k}^{-1}(z)T_N^{1/2} e_i e_i^* T_N^{1/2} (\hat{T}_N(z))^{-1} s_k \\
- \frac{1}{K} e_i^* T_N^{1/2} (\hat{T}_N(z))^{-1} T_N A_{1k}^{-1}(z)T_N^{1/2} e_i.
\]

From (2.5), we have
\[
E(T_N^{1/2} A_{1i}^{-1}(z)T_N^{1/2})_{ii} - e_i^* T_N^{1/2} (-\hat{T}_N(z))^{-1} T_N^{1/2} e_i \\
= \sum_{k=2}^{K} E \left[ \beta_{1k}(z) e_i^* T_N^{1/2} (\hat{T}_N(z))^{-1} s_k s_k^* A_{1k}^{-1}(z)T_N^{1/2} e_i \right. \\
\left. - \beta_{1k}(z) e_i^* T_N^{1/2} \frac{1}{K} (\hat{T}_N(z_1))^{-1} T_N E A_{1}^{-1}(z)T_N^{1/2} e_i \right] \\
= \tau_1(z) + \tau_2(z) + \tau_3(z),
\]
where
\[
\tau_1(z) = (K - 1) E[\beta_{12}(z) b_{12}(z) \gamma(z) \alpha(z)],
\]
\[
\tau_2(z) = \frac{K - 1}{K} E[\beta_{12}(z) e_i^* T_N^{1/2} (\hat{T}_N(z))^{-1} T_N (A_{12}^{-1}(z) - A_1^{-1}(z))T_N^{1/2} e_i]
\]
and
\[
\tau_3(z) = \frac{K - 1}{K} E[\beta_{12}(z) e_i^* T_N^{1/2} (\hat{T}_N(z))^{-1} T_N (A_1^{-1}(z) - E A_1^{-1}(z))T_N^{1/2} e_i].
\]

Therefore, it follows from (4.4) that
\[
\frac{1}{K} \sum_{i=1}^{N} E(T_N^{1/2} A_{1i}^{-1}(z_1)T_N^{1/2})_{ii} E(T_N^{1/2} A_{1i}^{-1}(z_2)T_N^{1/2})_{ii} \\
= \frac{1}{K} \sum_{i=1}^{N} e_i^* T_N^{1/2} (\hat{T}_N(z_1))^{-1} T_N^{1/2} e_i e_i^* T_N^{1/2} (\hat{T}_N(z_2))^{-1} T_N^{1/2} e_i + O \left( \frac{1}{\sqrt{K}} \right),
\]
where the estimate can be obtained as in Theorem 1.3.

Regarding (4.2), due to similar reason, one need only seek the limit of

$$\frac{1}{K} \sum_{i=1}^{N} E[(T_{N}^{1/2} A_{j}^{-1}(z) T_{N}^{1/2})_{ii}] E[(T_{N}^{1/2} A_{j}^{-1}(z) (\hat{T}_{N}(z))^{-1} T_{N}^{1/2})_{ii}].$$

However, as in (4.4), one can conclude that

$$\frac{1}{K} \sum_{i=1}^{N} E[(T_{N}^{1/2} A_{j}^{-1}(z) T_{N}^{1/2})_{ii}] E[(T_{N}^{1/2} A_{j}^{-1}(z) (\hat{T}_{N}(z))^{-1} T_{N}^{1/2})_{ii}]$$

$$= \frac{1}{K} \sum_{i=1}^{N} e_{i}^{*} T_{N}^{1/2} (\hat{T}_{N}(z))^{-1} T_{N}^{1/2} e_{i}^{*} e_{i}^{*} T_{N}^{1/2} (\hat{T}_{N}(z))^{-2} T_{N}^{1/2} e_{i} + O\left(\frac{1}{\sqrt{K}}\right).$$

For later purpose, we now derive (1.23) and (1.24). Note that when $T_{N} = I$, for $z \in C^{+}$,

$$z = -\frac{1}{m(z)} + \frac{c}{1+m(z)} \quad \text{(4.5)}$$

and

$$\frac{d}{dz} m(z) = \frac{m^{2}(z)}{1-cm^{2}(z)/(1+m(z))^{2}} \quad \text{(4.6)}$$

Then for $g(x) = x^{r}$,

$$\frac{1}{2\pi i} \int g(z) \frac{cm^{3}(z) h_{2}(z)}{1-c \int m^{2}(z) t^{2} dH(t)/(1+tm(z))^{2}} dz$$

$$= \frac{c}{2\pi i} \int \frac{(-1/m(z) + c/(1+m(z)))^{r}}{(m(z) + 1)^{3}} m(z) \, dm(z)$$

$$= \frac{c}{2\pi i} \int \frac{(-1/m(z) + c/(1+m(z)))^{r}}{(m(z) + 1)^{2}} \, dm(z)$$

$$- \frac{c}{2\pi i} \int \frac{(-1/m(z) + c/(1+m(z)))^{r}}{(m(z) + 1)^{3}} \, dm(z)$$

$$\triangleq v_{1} - v_{2}.$$

For $v_{1}$, we have

$$v_{1} = c^{r} \frac{c}{2\pi i} \int \frac{((1-c)/c + 1/(1+m(z)))^{r}}{(m(z) + 1)^{2}} \frac{1}{1-(1+m(z))^{-r}} \, dm(z)$$

$$= \frac{c^{1+r}}{2\pi i} \int \sum_{j=0}^{r} \binom{r}{j} \left(\frac{1-c}{c}\right)^{j} \frac{1}{(1+m(z))^{r-j+2}}$$
\[ \times \sum_{k=0}^{\infty} \binom{r+k-1}{k} (1 + m(z))^k \, dm(z) \]

\[ = c^{1+r} \sum_{j=0}^{r} \binom{r}{j} \left( \frac{1-c}{c} \right)^j \left( \frac{2r-j}{r-1} \right). \]

Similarly,

\[ v_2 = c^{1+r} \sum_{j=0}^{r} \binom{r}{j} \left( \frac{1-c}{c} \right)^j \left( \frac{2r+1-j}{r-1} \right). \]

For (1.24), we have

\[ \int z_{1}^{r_1} d \left[ \frac{m(z_1)}{1 + m(z_1)} \right] dz_1 = \int \frac{(-1/m(z_1) + c/(1 + m(z_1)))^r}{(m(z_1) + 1)^2} \, dm(z_1) \]

\[ = 2\pi ic^{r_1} \sum_{j=0}^{r_1} \binom{r_1}{j} \left( \frac{1-c}{c} \right)^j \left( \frac{2r_1-j}{r_1-1} \right). \]

Therefore,

\[ -c^{4} 4\pi^2 \int \int g_1(z_1)g_2(z_2) \frac{d^2}{dz_1 dz_2} [m(z_1)m(z_2)h_1(z_1, z_2)] \, dz_1 dz_2 \]

\[ = c^{r_1+s_2+1} \sum_{j_1=0}^{r_1} \sum_{j_2=0}^{r_2} \binom{r_1}{j_1} \binom{r_2}{j_2} \left( \frac{1-c}{c} \right)^{j_1+j_2} \left( \frac{2r_1-j_1}{r_1-1} \right) \left( \frac{2r_2-j_2}{r_2-1} \right). \]

5. Proof of Theorem 1.2. Since the truncation process is tedious, it is deferred to the Appendix. It may then be assumed that the underlying r.v.'s satisfy

\[ E v_{11} = 0, \quad E |v_{11}|^2 = 1, \quad |v_{11}| \leq \varepsilon_N \sqrt{N}, \]

where \( \varepsilon_N \) is a positive sequence converging to zero.

Define \( \tilde{s}_k = s_k^* R_k s_k - a_1 \). Expand \( (s_k^* R_k s_k)^{-1} \) a little bit as follows:

\[ \frac{1}{s_k^* R_k s_k} = \frac{1}{a_1} - \frac{\tilde{s}_k}{a_1 s_k^* R_k s_k} \]

\[ = \frac{1}{a_1} - \frac{\tilde{s}_k}{a_1} + \frac{(\tilde{s}_k)^2}{a_1^2 s_k^* R_k s_k}. \]

(5.1)

It follows that

\[ \sum_{k=1}^{K} \left( \beta_k - \frac{p_k}{\alpha_1} \right) = G_1 + G_2 + G_3 + G_4, \]

(5.2)
where
\[ G_1 = \frac{1}{a_1} \sum_{k=1}^{K} p_k (s_k^* s_k)^2 - 1), \quad G_2 = -\frac{1}{a_1} \sum_{k=1}^{K} p_k (s_k^* s_k)^2 (\tilde{s}_k) \]
and
\[ G_3 = \frac{1}{a_1^2} \sum_{k=1}^{K} p_k (s_k^* s_k)^2 (\tilde{s}_k)^2, \quad G_4 = -\frac{1}{a_1} \sum_{k=1}^{K} p_k (s_k^* s_k)^2 (\tilde{s}_k)^3 \frac{s_k^* R_k s_k}{s_k^2 R_k s_k}. \]

We will analyze \( G_1, G_2, G_3, G_4 \) one by one and, as will be seen, the contribution from the term \( G_4 \) is negligible.

First consider the term \( G_4 \). Since \( s_k^* R_k s_k \geq \sigma^2 s_k^* s_k \), we have
\[ |G_4| \leq M (G_{41} + \cdots + G_{43}), \]
where
\[ G_{41} = \sum_{k=1}^{K} p_k \left| s_k^* s_k \left( s_k^* R_k s_k - \frac{1}{N} Tr R_k \right)^3 \right| \]
and
\[ G_{42} = \sum_{k=1}^{K} p_k \left| s_k^* s_k \left( \frac{1}{N} Tr R_k - Tr R \right)^3 \right|, \quad G_{43} = \sum_{k=1}^{K} p_k \left| s_k^* s_k \left( \frac{1}{N} Tr R - a_1 \right)^3 \right|. \]

By the Hölder inequality,
\[ EG_{41} \leq M \sum_{k=1}^{K} p_k \left( E (s_k^* s_k - 1)^2 \right)^{1/2} \left( E \left( s_k^* R_k s_k - \frac{1}{N} Tr R_k \right)^6 \right)^{1/2} + M \sum_{k=1}^{K} p_k E \left| s_k^* R_k s_k - \frac{1}{N} Tr R_k \right|^3 \]
\[ = o(1). \]

Indeed, it is easy to verify that
\[ E (s_k^* s_k - 1)^2 = \frac{1}{N} (E |v_{11}|^4 - 1) \]
and that
\[ E \left( s_k^* R_k s_k - \frac{1}{N} Tr R_k \right)^p \leq \frac{M}{N^p} (E |v_{11}|^4 E (Tr R_k^2)^{p/2} + E v_{11}^{2p} E Tr R_k^p) \]
\[ \leq \frac{M}{N^{p/2}} + \frac{M \epsilon_{N}^{2p-4}}{N^2}, \]
(5.4)
where the constant $M$ is independent of $k$. Here we use the fact $R_k \leq MS_k S_k^* + \sigma^2 I$.

Furthermore, it is direct to prove

$$\frac{1}{N^2} \sum_{k=1}^{K} p_k |s_k^* s_k| \overset{i.p.}{\longrightarrow} 0.$$  

This, together with Theorem 1 of Bai and Silverstein [3], leads to

$$G_{43} \overset{i.p.}{\longrightarrow} 0.$$  

In addition, it is also easy to verify that

$$EG_{42} = \frac{1}{N^2} \sum_{k=1}^{K} p_k^4 E(s_k^* s_k)^4 = O\left(\frac{1}{N^2}\right).$$  

Combining the above argument, one can claim that the contribution from $G_4$ can be ignored.

Analyze the term $G_1$ second. Write

$$\sum_{k=1}^{K} p_k (s_k^* s_k)^2 = \sum_{k=1}^{K} p_k (s_k^* s_k - 1)^2 + 2 \sum_{k=1}^{K} p_k s_k^* s_k - \sum_{k=1}^{K} p_k$$

(5.5)

$$= \sum_{k=1}^{K} p_k (s_k^* s_k - 1)^2 + 2 \text{Tr} C - \sum_{k=1}^{K} p_k.$$  

Moreover,

$$E\left(\sum_{k=1}^{K} [p_k (s_k^* s_k - 1)^2 - E(s_k^* s_k - 1)^2]\right)^2$$

$$= \sum_{k=1}^{K} p_k^2 E((s_k^* s_k - 1)^2 - E(s_k^* s_k - 1)^2)^2 = o(1),$$

using

$$E(s_k^* s_k - 1)^4 = o\left(\frac{1}{N}\right).$$

So

$$\sum_{k=1}^{K} p_k (s_k^* s_k - 1)^2 \overset{i.p.}{\longrightarrow} \frac{1}{c} (E|v_{11}|^4 - 1) \int x \, dH(x),$$

and then

$$G_1 = \frac{1}{a_1} \left(\frac{(E|v_{11}|^4 - 1) \int x \, dH(x)}{c} + 2 \text{Tr} C - 2 \sum_{k=1}^{K} p_k\right) + o_p(1).$$
Third, for the term $G_2$, similar to $G_1$,

\begin{equation}
-a_1^2 G_2 = \sum_{k=1}^{K} p_k(\tilde{s}_k)(s_k^* s_k - 1)^2
\end{equation}

\begin{equation}
+ 2 \sum_{k=1}^{K} p_k(\tilde{s}_k)(s_k^* s_k - 1) + \sum_{k=1}^{K} p_k(\tilde{s}_k).
\end{equation}

For the sum in (5.8), we have

\[
E \left| \sum_{k=1}^{K} p_k(\tilde{s}_k)(s_k^* s_k - 1)^2 \right| \leq M \sum_{k=1}^{K} (E(\tilde{s}_k)^2)^{1/2}(E(s_k^* s_k - 1)^4)^{1/2}
= o(1),
\]

where we use (5.6) and

\[
E(\tilde{s}_k)^2 \leq E\left(s_k^* R_k s_k - \frac{1}{N} \text{Tr} R_k\right)^2 + E\left(\frac{1}{N} \text{Tr} R_k - a_1\right)^2 = O\left(\frac{1}{N}\right),
\]

which is accomplished by (5.4) and Theorem 1 of Bai and Silverstein [3]. Similarly to (5.5), we deduce that

\begin{equation}
\sum_{k=1}^{K} p_k^2(s_k^* s_k)^2 = \frac{1}{c} (E|v_{11}|^4 - 1) \int x^2 dH(x)
\end{equation}

\begin{equation}
+ 2 \text{Tr} S P^2 S^* - \sum_{k=1}^{K} p_k^2 + o_p(1).
\end{equation}

Applying $C - p_k s_k s_k^* = C_k$, the second sum of (5.9) is then equal to

\begin{equation}
\sigma^2 \text{Tr} C + \text{Tr} C^2 - a_1 \sum_{k=1}^{K} p_k - \sum_{k=1}^{K} p_k^2(s_k^* s_k)^2
\end{equation}

\begin{equation}
= \sigma^2 \text{Tr} C + \text{Tr} C^2 - a_1 \sum_{k=1}^{K} p_k - \frac{1}{c} (E|v_{11}|^4 - 1) \int x^2 dH(x)
\end{equation}

\begin{equation}
- 2 \text{Tr} S P^2 S^* + \sum_{k=1}^{K} p_k^2 + o_p(1).
\end{equation}

With regard to the first sum of (5.9), its variance will be proved to converge to zero. Now let us provide more details to the reader:

\begin{equation}
\text{Var}\left(\sum_{k=1}^{K} p_k(\tilde{s}_k)(s_k^* s_k - 1)\right) = G_{21} + G_{22},
\end{equation}
where

\[ G_{21} = \sum_{k=1}^{K} p_k^2 E[(\tilde{s}_k)(s_k^* s_k - 1) - E(\tilde{s}_k)(s_k^* s_k - 1)]^2 \]

and

\[ G_{22} = \sum_{k_1 \neq k_2}^{K} p_{k_1} p_{k_2} E\left[ \left( (\tilde{s}_{k_1})(s_{k_1}^* s_{k_1} - 1) - E(\tilde{s}_{k_1})(s_{k_1}^* s_{k_1} - 1) \right) \times \left( (\tilde{s}_{k_2})(s_{k_2}^* s_{k_2} - 1) - E(\tilde{s}_{k_2})(s_{k_2}^* s_{k_2} - 1) \right) \right]. \]

Evidently,

\[ G_{21} \leq \sum_{k=1}^{K} M E[(\tilde{s}_k)(s_k^* s_k - 1)]^2 \]

\[ \leq M \sum_{k=1}^{K} E \left[ \left( s_k^* R_k s_k - \frac{1}{N} \text{Tr} R_k \right)(s_k^* s_k - 1) \right]^2 \]

\[ + M \sum_{k=1}^{K} E \left[ \left( \frac{1}{N} \text{Tr} R_k - a_1 \right)(s_k^* s_k - 1) \right]^2 \]

(5.13)

\[ \leq M \sum_{k=1}^{K} \left[ E \left( s_k^* R_k s_k - \frac{1}{N} \text{Tr} R_k \right) \right]^{4/2} [E(s_k^* s_k - 1)^4]^{1/2} \]

\[ + M \sum_{k=1}^{K} E \left( \frac{1}{N} \text{Tr} R_k - a_1 \right)^2 E(s_k^* s_k - 1)^2 \]

\[ = o(1). \]

Let \( S_{k_1 k_2} \) denote the matrix obtained from \( S_{k_1} \) by deleting the \( k_2 \)th column and, furthermore, \( R_{k_1 k_2} \) and \( C_{k_1 k_2} \) have the same meaning. Split \( R_{k_1} = R_{k_1 k_2} + p_{k_2} s_{k_2} s_{k_2}^* \) and \( R_{k_2} = R_{k_1 k_2} + p_{k_1} s_{k_1} s_{k_1}^* \). Also, for convenience, set

\[ \alpha_{kj} = s_{kj}^* R_{k_1 k_2} s_{kj} - a_1, \quad \gamma_j = s_{kj}^* R_{k_1 k_2} s_{kj} - \frac{1}{N} \text{Tr} R_{k_1 k_2} \]

and

\[ \Upsilon_{kj} = s_{kj}^* s_{kj} - 1, \quad j = 1, 2. \]

\( G_{22} \) is then decomposed as

\[ G_{22} = G_{221} + \cdots + G_{224}, \]
where

\[ G_{221} = \sum_{k_1 \neq k_2} p_{k_1} p_{k_2} \text{Cov}(\alpha_{k_1} \gamma_{k_1}, \alpha_{k_2} \gamma_{k_2}), \]

\[ G_{222} = \sum_{k_1 \neq k_2} p_{k_1}^2 p_{k_2} \text{Cov}(\alpha_{k_1} \gamma_{k_1}, |s_{k_1}^* s_{k_2}|^2 \gamma_{k_2}), \]

\[ G_{223} = \sum_{k_1 \neq k_2} p_{k_1} p_{k_2}^2 \text{Cov}(|s_{k_1}^* s_{k_2}|^2 \gamma_{k_1}, \alpha_{k_2} \gamma_{k_2}) \]

and

\[ G_{224} = \sum_{k_1 \neq k_2} p_{k_1}^2 p_{k_2} \text{Cov}(|s_{k_1}^* s_{k_2}|^2 \gamma_{k_1}, |s_{k_1}^* s_{k_2}|^2 \gamma_{k_2}). \]

The basic idea behind this decomposition is to produce some independent terms when \( R_{k_1 k_2} \) is given, which is very important when estimating the order of some terms.

It is easy to check that

\[ (5.14) \quad E \left( s_k^* R_k s_k - \frac{1}{N} \text{Tr} R_k \right) \left( s_k^* s_k - 1 \right) = \frac{E|v_{11}|^4 - 1}{N^2} E \text{Tr} R_k, \]

and that

\[ (5.15) \quad E \left| s_k^* D s_k - \frac{1}{N} \text{Tr} D \right|^2 = \frac{1}{N^2} (E|v_{11}|^4 - 2) \sum_i [(D)_{ii}]^2 + \frac{1}{N^2} \text{Tr} D D^*, \]

where \( D \) is any constant Hermite matrix.

This gives that \( G_{221} \) is equal to

\[ \sum_{k_1 \neq k_2} p_{k_1} p_{k_2} E \left[ E(\alpha_{k_1} \gamma_{k_1} | R_{k_1 k_2}) E(\alpha_{k_2} \gamma_{k_2} | R_{k_1 k_2}) \right] \]

\[ = \sum_{k_1 \neq k_2} p_{k_1}^2 p_{k_2} E \left[ E(\gamma_{k_1} \gamma_{k_1} | R_{k_1 k_2}) E(\gamma_{k_2} \gamma_{k_2} | R_{k_1 k_2}) \right] \]

\[ = \frac{E|v_{11}|^4 - 1}{N} \sum_{k_1 \neq k_2} p_{k_1} p_{k_2} E \left( \frac{1}{N} \text{Tr} R_{k_1 k_2} - E \frac{1}{N} \text{Tr} R_{k_1 k_2} \right)^2 = O \left( \frac{1}{N} \right), \]

where \( \alpha_k \gamma_k = \alpha_k \gamma_k - E \alpha_k \gamma_k, \gamma_k \gamma_k = \gamma_k \gamma_k - E \gamma_k \gamma_k, \) and we use the independence of \( s_{k_1} \) and \( s_{k_2} \), and

\[ (5.16) \quad E \left( \frac{1}{N} \text{Tr} R_{k_1 k_2} - E \frac{1}{N} \text{Tr} R_{k_1 k_2} \right)^2 = \frac{1}{N^2} E \left| \sum_{j \neq k_1, k_2} p_j (s_j^* s_j - 1) \right|^2 \leq \frac{M}{N^2}, \]
where $M$ is independent of $k_1, k_2$.

After some simple computations, we get

$$E|s^*_k s_k|^2 \gamma_k = \frac{E|v_{11}|^4 - 1}{N^2},$$

(5.17)

$$E(|s^*_k s_k|^2 \gamma_k | s_k) = \frac{E|v_{11}|^4 - 1}{N^2} s^*_k s_k,$$

and so

$$G_{222} = \sum_{k_1 \neq k_2} p^2_{k_1} p^2_{k_2} E((\alpha_{k_1} \bar{\gamma}_{k_1}) E[|s^*_k s_k|^2 \gamma_k - E(|s^*_k s_k|^2 \gamma_k) | s_k])$$

$$= \frac{E|v_{11}|^4 - 1}{N^2} \sum_{k_1 \neq k_2} p^2_{k_1} p^2_{k_2} (E(\alpha_{k_1} \bar{\gamma}_{k_1}) s^*_k s_k)$$

(5.18)

$$= \frac{E|v_{11}|^4 - 1}{N^2} \sum_{k_1 \neq k_2} p^2_{k_1} p^2_{k_2} (E(\alpha_{k_1} \bar{\gamma}_{k_1}) \gamma_{k_1})$$

$$\leq M \sum_{k_1 \neq k_2} \left( E(\alpha^2_{k_1}) \right)^{1/2} (E \gamma_{k_1}^4)^{1/2}$$

$$= O\left( \frac{1}{N} \right).$$

Similarly, one can conclude that

$$G_{223} \to 0.$$

Write

$$G_{224} = \sum_{k_1 \neq k_2} p^2_{k_1} p^2_{k_2} E(|s^*_k s_k|^4 \gamma_k \gamma_k) - \sum_{k_1 \neq k_2} p^2_{k_1} p^2_{k_2} E(|s^*_k s_k|^2 \gamma_k)^2.$$

The second sum converges to zero because of (5.17). For its first sum we have

$$\sum_{k_1 \neq k_2} p^2_{k_1} p^2_{k_2} E(|s^*_k s_k|^4 \gamma_k \gamma_k) = \sum_{k_1 \neq k_2} p^2_{k_1} p^2_{k_2} E(\gamma_k E(|s^*_k s_k|^4 \gamma_k | s_k)),$$

which is less than or equal to

$$M \sum_{k_1 \neq k_2} E(\gamma_k | E\left(\left(s^*_k s_k s^*_s s_{k_1} - \frac{1}{N} \text{Tr} s^*_k s_{k_2}\right)^2 | \gamma_{k_1} | s_k\right))$$

(5.20)

$$+ M \sum_{k_1 \neq k_2} E(\gamma_k | E\left(\left(\frac{1}{N} s^*_k s_{k_2}\right)^2 | \gamma_{k_1} | s_k\right)) = o(1).$$
as

\[
E \left\{ |\gamma_{k_2}| E \left[ \left( \mathbf{s}_{k_1}^* \mathbf{s}_{k_2}^* \mathbf{s}_{k_2} \mathbf{s}_{k_1} - \frac{1}{N} \text{Tr} \mathbf{s}_{k_2}^* \mathbf{s}_{k_2}^* \right)^2 | \gamma_{k_1} \right| \ | \mathbf{s}_{k_2} \right] \right\}
\]

\[
\leq E \left\{ |\gamma_{k_2}| \left[ E \left( \left( \mathbf{s}_{k_1}^* \mathbf{s}_{k_2}^* \mathbf{s}_{k_2} \mathbf{s}_{k_1} - \frac{1}{N} \text{Tr} \mathbf{s}_{k_2}^* \mathbf{s}_{k_2}^* \right)^4 \right) \left| \mathbf{s}_{k_2} \right| \right]^{1/2} \left[ E \left( |\gamma_{k_1}| \right) \right]^{1/2} \right\}
\]

\[
\leq \frac{M \varepsilon^2}{N^{3/2}} E [ |\gamma_{k_2}| (\mathbf{s}_{k_2}^* \mathbf{s}_{k_2})^2 ]
\]

\[
\leq \frac{M \varepsilon^2}{N^{3/2}} E [ |\gamma_{k_2}|^3 ] + \frac{M \varepsilon^2}{N^{3/2}} (E |\gamma_{k_2}|)^{1/2}
\]

\[
= o \left( \frac{1}{N^2} \right)
\]

and

\[
E \left\{ |\gamma_{k_2}| E \left[ \left( \frac{1}{N} \mathbf{s}_{k_2}^* \mathbf{s}_{k_2} \right)^2 | \gamma_{k_1} \right| \ | \mathbf{s}_{k_2} \right] \right\}
\]

\[
\leq \frac{1}{N^2} E [ |\gamma_{k_2}| (\mathbf{s}_{k_2}^* \mathbf{s}_{k_2})^2 ] (E |\gamma_{k_1}|)^{2/2} = O \left( \frac{1}{N^3} \right).
\]

Consequently, \( G_{224} \) converges to zero and then \( G_{22} \) converges to zero. Therefore, via (5.14),

\[
(5.21) \quad \sum_{k=1}^{K} p_k (\mathbf{s}_k^* \mathbf{s}_k - 1) \overset{i.p.}{\longrightarrow} \frac{E |v_{11}|^4 - 1}{c} a_1 \int x \, dH(x).
\]

Combining (5.9)–(5.12) with (5.21), one can conclude that

\[
G_2 = -\frac{1}{a_1^2} \left[ 2 a_1 \frac{E |v_{11}|^4 - 1}{c} \int x \, dH(x) \right.
\]

\[
+ \sigma^2 \text{Tr} \mathbf{C} + \text{Tr} \mathbf{C}^2 - a_1 \sum_{k=1}^{K} p_k
\]

\[
- \frac{1}{c} (E |v_{11}|^4 - 1)
\]

\[
\times \int x^2 \, dH(x) - 2 \text{Tr} \mathbf{S} \mathbf{P}^2 \mathbf{S}^* + \sum_{k=1}^{K} p_k^2 \right] + o_p(1).
\]

Fourth, turn to the term \( G_3 \). It is decomposed as

\[
(5.23) \quad a_1^3 G_3 = G_{31} + G_{32} + G_{33},
\]

where

\[
G_{31} = \sum_{k=1}^{K} p_k (\mathbf{s}_k^* \mathbf{s}_k - 1)^2 (\mathbf{s}_k^* \mathbf{s}_k)^2
\]
(recall $\tilde{s}_k = s_k^* R_k s_k - a_1$) and

$$G_{32} = 2 \sum_{k=1}^{K} p_k (s_k^* s_k - 1) (\tilde{s}_k)^2, \quad G_{33} = \sum_{k=1}^{K} p_k (\tilde{s}_k)^2.$$ 

Applying the Hölder inequality,

$$E|G_{31}| \leq M \sum_{k=1}^{K} \left[ E(s_k^* s_k - 1)^2 \left(s_k^* R_k s_k - \frac{1}{N} \text{Tr} R_k\right)^2 + E(s_k^* s_k - 1)^2 \left(\frac{1}{N} \text{Tr} R_k - a_1\right)^2 \right] \leq M \sum_{k=1}^{K} \left( E(s_k^* s_k - 1)^4 \right)^{1/2} \left( E\left(s_k^* R_k s_k - \frac{1}{N} \text{Tr} R_k\right)^4 \right)^{1/2} + M \sum_{k=1}^{K} E(s_k^* s_k - 1)^2 E\left(\frac{1}{N} \text{Tr} R_k - a_1\right)^2 = o(1).$$

Analogously, one can also obtain

$$E|G_{32}| = o(1).$$

To derive the limit of $G_{33}$, we need to evaluate its variance:

$$E\left(\sum_{k=1}^{K} p_k^2 [(\tilde{s}_k)^2 - E(\tilde{s}_k)^2]\right)^2 = G_{331} + G_{332},$$

where

$$G_{331} = \sum_{k=1}^{K} p_k^2 E[(\tilde{s}_k)^2 - E(\tilde{s}_k)^2]^2$$

and

$$G_{332} = \sum_{k_1 \neq k_2} p_{k_1} p_{k_2} E[(\tilde{s}_{k_1})^2 - E(\tilde{s}_{k_1})^2](\tilde{s}_{k_2})^2 - E(\tilde{s}_{k_2})^2].$$

For $G_{331}$, we have

$$G_{331} \leq M \sum_{k=1}^{K} E[(\tilde{s}_k)^4] \leq M \sum_{k=1}^{K} E\left(s_k^* R_k s_k - \frac{1}{N} \text{Tr} R_k\right)^4 + M \sum_{k=1}^{K} E\left(\frac{1}{N} \text{Tr} R_k - \frac{1}{N} \text{Tr} R\right)^4 + M \sum_{k=1}^{K} E\left(\frac{1}{N} \text{Tr} R - a_1\right)^4 = o(1).$$
In fact, note that \( a_1 = \sigma^2 + \frac{1}{cN} \),

\[
(5.25) \quad E \left( \frac{1}{N} \text{Tr} \mathbf{R} - a_1 \right)^4 = E \left( \frac{1}{N} \sum_{k=1}^{K} p_k (s_k^* s_k - 1) \right)^4 = o \left( \frac{1}{N^2} \right).
\]

Since the treatment of \( G_{332} \) is basically similar to that of \( G_{22} \), we give only an outline. To this end, we expand it as

\[
(5.26) \quad G_{332} = G_{332}^{(1)} + \cdots + G_{332}^{(9)},
\]

where

\[
G_{332}^{(1)} = \sum_{k_1 \neq k_2}^K p_{k_1} p_{k_2} \text{Cov}(\alpha_{k_1}^2, \alpha_{k_2}^2),
\]

\[
G_{332}^{(2)} = \sum_{k_1 \neq k_2}^K p_{k_1}^3 p_{k_2} \text{Cov}(\alpha_{k_1}^2, |s_{k_1}^* s_{k_2}|^4),
\]

\[
G_{332}^{(3)} = 2 \sum_{k_1 \neq k_2}^K p_{k_1}^2 p_{k_2} \text{Cov}(\alpha_{k_1}^2, \alpha_{k_2} |s_{k_1}^* s_{k_2}|^2),
\]

\[
G_{332}^{(4)} = 2 \sum_{k_1 \neq k_2}^K p_{k_1} p_{k_2}^2 \text{Cov}(\alpha_{k_1} |s_{k_1}^* s_{k_2}|^2, \alpha_{k_2}^2),
\]

\[
G_{332}^{(5)} = 2 \sum_{k_1 \neq k_2}^K p_{k_1}^3 p_{k_2} \text{Cov}(\alpha_{k_1}^2 |s_{k_1}^* s_{k_2}|^2, |s_{k_1}^* s_{k_2}|^4),
\]

\[
G_{332}^{(6)} = 4 \sum_{k_1 \neq k_2}^K p_{k_1}^2 p_{k_2}^2 \text{Cov}(\alpha_{k_1} |s_{k_1}^* s_{k_2}|^2, \alpha_{k_2}^2 |s_{k_1}^* s_{k_2}|^2),
\]

\[
G_{332}^{(7)} = \sum_{k_1 \neq k_2}^K p_{k_1} p_{k_2}^3 \text{Cov}(|s_{k_1}^* s_{k_2}|^4, \alpha_{k_2}^2),
\]

\[
G_{332}^{(8)} = 2 \sum_{k_1 \neq k_2}^K p_{k_1}^3 p_{k_2}^3 \text{Var}(|s_{k_1}^* s_{k_2}|^4)
\]

and

\[
G_{332}^{(9)} = 2 \sum_{k_1 \neq k_2}^K p_{k_1}^2 p_{k_2}^3 \text{Cov}(|s_{k_1}^* s_{k_2}|^4, \alpha_{k_2}^2 |s_{k_1}^* s_{k_2}|^2).
\]

We claim that

\[
G_{332} = o(1).
\]
But, in the sequel, as an illustration, only terms $G_{332}^{(1)}$ and $G_{332}^{(8)}$ will be estimated, the argument for all the remaining ones are analogous and then omitted.

Using the Burkholder inequality,

$$E |s_{k_1}^* s_{k_2}|^8 \leq \frac{M}{N^8} \left[ \left( \sum_{i=1}^{N} E |v_{ik_1} v_{ik_2}|^2 \right)^4 + \sum_{i=1}^{N} (E |v_{11}|^8)^2 \right]$$

(5.27)

$$= O \left( \frac{1}{N^3} \right),$$

which leads to

$$|G_{332}^{(8)}| \leq M \sum_{k_1 \neq k_2} p_{k_1}^2 p_{k_2}^3 E |s_{k_1}^* s_{k_2}|^8 = O \left( \frac{1}{N} \right).$$

From (5.15),

$$G_{332}^{(1)} = \sum_{k_1 \neq k_2}^K p_{k_1} p_{k_2} E (E (\alpha_{k_1}^2 - E \alpha_{k_1}^2 | R_{k_1 k_2}) E (\alpha_{k_2}^2 - E \alpha_{k_2}^2 | R_{k_1 k_2}))$$

$$\leq \frac{M}{N^4} \sum_{k_1 \neq k_2}^K \left[ E (\text{Tr} R_{k_1 k_2}^2 - E \text{Tr} R_{k_1 k_2}^2)^2 + E (\text{Tr} R_{k_1 k_2} - Na_1)^4 \right.$$

$$\left. + E \left( \sum_{i=1}^{N} [(R_{k_1 k_2})_{ii}]^2 - E [(R_{k_1 k_2})_{ii}]^2 \right)^2 \right]$$

(5.28)

$$+ E (\text{Tr} R_{k_1 k_2} - Na_1)^2)^2 \right]$$

$$= o(1).$$

In order to get (5.28), we need to analyze the above four terms on the right-hand side of the inequality. First, applying $R_{k_1 k_2} = (R_{k_1 k_2} - R_{k_1}) + R_{k_1}$ twice, we have

$$E (\text{Tr} R_{k_1 k_2}^2 - E \text{Tr} R_{k_1 k_2}^2)^2$$

$$\leq M E (\text{Tr} (R_{k_1 k_2} - R_{k_1}) R_{k_1 k_2} - E \text{Tr} (R_{k_1 k_2} - R_{k_1}) R_{k_1 k_2})^2$$

(5.29)

$$+ M E (\text{Tr} R_{k_1} (R_{k_1 k_2} - R_{k_1}) - E \text{Tr} R_{k_1} (R_{k_1 k_2} - R_{k_1}))^2$$

$$+ M E (\text{Tr} R_{k_1}^2 - E \text{Tr} R_{k_1}^2)^2$$

$$\leq M E (\hat{y}_{k_2}^2)^2 + M E ((s_{k_2}^* s_{k_2})^2 - E (s_{k_2}^* s_{k_2})^2)^2$$

$$+ M E (\text{Tr} R_{k_1}^2 - E \text{Tr} R_{k_1}^2)^2,$$
where \( \hat{y}_k = s_{k_2}^* \mathbf{R}_{k_1 k_2} s_{k_2} - \frac{1}{N} E \text{Tr} \mathbf{R}_{k_1 k_2} \). However, observe that

\[
(5.30) \quad E((s_{k_2}^* s_{k_2})^2 - E(s_{k_2}^* s_{k_2})^2)^2 = O\left(\frac{1}{N}\right)
\]

and that

\[
E(\gamma_{k_2}^2) \leq \frac{M}{N},
\]

using (5.4) and (5.16). Therefore,

\[
E(\text{Tr} \mathbf{R}_{k_1 k_2}^2 - E \text{Tr} \mathbf{R}_{k_1 k_2}^2)^2 \leq \frac{M}{N} + ME(\text{Tr} \mathbf{R}_{k_1}^2 - E \text{Tr} \mathbf{R}_{k_1}^2)^2 \leq \frac{M}{N} + ME(\text{Tr}^2 - E \text{Tr}^2)^2
\]

again, repeating a process analogous to (5.29) in the last step. But this implies

\[
M \sum_{k_1 \neq k_2}^{K} E(\text{Tr} \mathbf{R}_{k_1 k_2}^2 - E \text{Tr} \mathbf{R}_{k_1 k_2}^2)^2 \to 0.
\]

Second,

\[
E(\text{Tr} \mathbf{R}_{k_1 k_2} - Na_1)^4 \leq ME \left| \sum_{k=1}^{K} p_k (s_k^* s_k - 1) \right|^4 + ME(s_k^* s_{k_1})^4 + ME(s_k^* s_{k_2})^4 \leq M,
\]

which shows that the second sum in (5.28) converges to zero. Similarly, one can also prove that the fourth sum in (5.28) converges to zero by a similar argument, as expected. Finally, in order to show that the third sum in (5.28) converges to zero, it is enough to show that

\[
E||[(\mathbf{C})_{ii}]^2 - E[(\mathbf{C})_{ii}]^2||^2 \to 0.
\]

To this end, it suffices to verify that

\[
E||[(\mathbf{C})_{ii}] - E(\mathbf{C})_{ii}]^4 \to 0,
\]

but, as in Theorem 1.3, through martingale difference decomposition, one can get it. Thus, (5.28) holds.

Hence, by (5.15) and an argument similar to Theorem 1.4, we have so far proved that

\[
G_{33} = \frac{a_2}{c} \int_{x} x \, dH(x) + \frac{E|v_{11}|^4}{c} - 2a_1^2 \int_{x} x \, dH(x) + o_p(1)
\]
and then

\[ G_3 = \frac{1}{a_1^3 c} \left( a_2 + (E|v_{11}|^4 - 2)a_1^2 \right) \int x \, dH(x) + o_p(1). \tag{5.31} \]

Summarizing (5.7), (5.22) and (5.31), we conclude that

\[
\sum_{k=1}^{K} \left( \frac{\beta_k - p_k}{a_1} \right) = \left( \frac{2}{a_1} - \frac{\sigma^2}{a_1^2} \right) \text{Tr} C - \frac{1}{a_1^2} \text{Tr} C^2 + \frac{2}{a_1^2} \text{Tr} SP^2 S^* \\
- \frac{1}{a_1} \sum_{k=1}^{K} p_k - \frac{1}{a_1^2} \sum_{k=1}^{K} p_k^2 + \left( \frac{a_2}{ca_1^2} - \frac{1}{ca_1} \right) \int x \, dH(x) \\
+ \frac{E|v_{11}|^4 - 1}{ca_1^2} \int x^2 \, dH(x) + o_p(1), \tag{5.32} \]

(recall that \( a_1 = \sigma^2 + 1/c \)).

Now we let \( p_k = 1, k = 1, \ldots, K \), so that Theorem 1.4 can be applied. In this case (5.32) becomes

\[
\sum_{k=1}^{K} \left( \frac{\beta_k - p_k}{a_1} \right) = \frac{2 + 2/c + \sigma^2}{a_1^2} (\text{Tr} C - K) - \frac{1}{a_1^2} (\text{Tr} C^2 - (1 + 1/c)K) \\
+ \frac{1}{c^2 a_1^4} + \frac{E|v_{11}|^4 - 1}{ca_1^2} + o_p(1), \tag{5.33} \]

[\( a_2 = \frac{(1+1/c)}{c} + \sigma^4 + 2\sigma^2/c \)]. Then Theorem 1.3 follows from Theorem 1.4. The expectation and variance of the limiting normal distribution are also from (1.23) and (1.24) of Bai and Silverstein [3] and our (1.23) and (1.24). Therefore, the variance \( \tau^2 \) is equal to

\[
\frac{(2 + 2/c + \sigma^2)^2 (E|v_{11}|^4 - 1)}{c(\sigma^2 + 1/c)^4} - 2 \frac{(2 + 2/c + \sigma^2)(2c + 2)}{c^2(\sigma^2 + 1/c)^4} (E|v_{11}|^4 - 1) \\
+ \frac{1}{c^4(\sigma^2 + 1/c)^4} (4c^3 + 10c^2 + 4c \\
+ (4c^3 + 8c^2 + 4c)(E|v_{11}|^4 - 2)). \tag{5.34} \]
6. Proof of Corollary 1.1. By the Taylor expansion,
\[
\sum_{k=1}^{K} \left( \log(1 + \beta_k) - \log \left( 1 + \frac{1}{a_1} \right) \right) = \sum_{k=1}^{K} \frac{\left( \beta_k - 1/a_1 \right)}{1 + 1/a_1} - \sum_{k=1}^{K} \frac{(\beta_k - 1/a_1)^2}{2(1 + 1/a_1)^2} + \sum_{k=1}^{K} \frac{(\beta_k - 1/a_1)^3}{3(1 + \psi_k)^3},
\]

with \( \psi_k \) being located in the interval \([1/a_1, \beta_k]\). Similar to Theorem 1.3, it can be shown that
\[
E \left( \sum_{k=1}^{K} \left( \left( \beta_k - \frac{1}{a_1} \right)^2 - E \left( \beta_k - \frac{1}{a_1} \right)^2 \right) \right)^2 = o(1)
\]
and
\[
\sum_{k=1}^{K} \left| \frac{(\beta_k - 1/a_1)^3}{3(1 + \psi_k)^3} \right| \leq M \sum_{k=1}^{K} \left| \left( \beta_k - \frac{1}{a_1} \right)^3 \right| \overset{i.p.}{\longrightarrow} 0.
\]

Now compute \( \sum_{k=1}^{K} E \left( \beta_k - \frac{1}{a_1} \right)^2 \). Applying (5.1),
\[
(6.1) \quad \sum_{k=1}^{K} E \left( \beta_k - \frac{1}{a_1} \right)^2 = \sum_{k=1}^{K} E \left( \frac{(s_k \bar{s}_k)^2 - 1}{a_1} - \frac{(s_k \bar{s}_k)^2 (\bar{s}_k)}{a_1^2} + \frac{(s_k \bar{s}_k)^2 (\bar{s}_k)}{a_1^2} \frac{R_k s_k}{a_1^2} \right)^2.
\]

Again, via an argument analogous to Theorem 1.3, it is easily seen that the contribution from the terms involving \((s_k \bar{s}_k)^2 (\bar{s}_k)/a_1^2 \bar{s}_k R_k s_k\) can be ignored. So (6.1) is equal to
\[
\sum_{k=1}^{K} E \left( \frac{(s_k \bar{s}_k)^2 - 1}{a_1} \right)^2 + \sum_{k=1}^{K} E \left( \frac{(s_k \bar{s}_k)^2 (\bar{s}_k)}{a_1^2} \right)^2 - 2 \sum_{k=1}^{K} E \frac{(s_k \bar{s}_k)^2 - 1}{a_1} \frac{(s_k \bar{s}_k)^2 (\bar{s}_k)}{a_1^2} + o(1).
\]

Combining the argument of Theorem 1.3 with (5.15) and (5.14), the above three terms are equal to, respectively, \( 4(E|v_{11}|^4 - 1)/(ca_1^2) \), \( a_2 + E|v_{11}|^4 - 2)/(ca_1^4) \) and \(-4(E|v_{11}|^4 - 1)/(ca_1^2) \). Thus, we finish the proof.

APPENDIX: TRUNCATION OF UNDERLYING RANDOM VARIABLES IN THEOREM 1.3

Let \( \hat{v}_{ij} = v_{ij} I(|v_{ij}| \leq \epsilon_N \sqrt{N}) \) and \( \bar{v}_{ij} = \hat{v}_{ij} - E \hat{v}_{ij}, \) \( i = 1, \ldots, N, \) \( j = 1, \ldots, K, \) where \( \epsilon_N \) is a positive sequence converging to zero. We use \( \hat{s}_k, \bar{s}_k, \hat{S}_k, \bar{S}_k, \hat{R}_k, \bar{R}_k \) and \( \hat{\beta}_k, \bar{\beta}_k, k = 1, \ldots, K, \) to denote the analogues of \( s_k, S_k, R_k \) and \( \beta_k \) with the elements replaced by \( \hat{v}_{ij} \) or \( \bar{v}_{ij} \).
As in the proof for Theorem 1.3 in Pan, Guo and Zhou [12], one can select the above $\varepsilon_N$ so that

$$\varepsilon_N^{-4} E v_{11}^4 I(|v_{11}| > \varepsilon_N \sqrt{N}) \to 0,$$

and show that

$$\sum_{k=1}^{K} \beta_k - \sum_{k=1}^{K} \hat{\beta}_k \overset{i.p.}{\to} 0.$$

Now consider the re-centralization of random variables. Applying (5.1),

\begin{equation}
\sum_{k=1}^{K} \frac{p_k (\hat{s}_k^* \hat{s}_k)^2}{\hat{s}_k^* \hat{R}_k \hat{s}_k} = U_1 + U_2 + U_3 + U_4,
\end{equation}

where

$$U_1 = \frac{1}{a_1} \sum_{k=1}^{K} p_k (\hat{s}_k^* \hat{s}_k)^2, \quad U_4 = - \frac{1}{a_1^3} \sum_{k=1}^{K} p_k (\hat{s}_k^* \hat{s}_k)^2 (\hat{s}_k^* \hat{R}_k \hat{s}_k - a_1)^3$$

and

$$U_2 = \frac{1}{a_1^2} \sum_{k=1}^{K} p_k (\hat{s}_k^* \hat{s}_k)^2 (\hat{s}_k^* \hat{R}_k \hat{s}_k - a_1),$$

$$U_3 = \frac{1}{a_1^3} \sum_{k=1}^{K} p_k (\hat{s}_k^* \hat{s}_k)^2 (\hat{s}_k^* \hat{R}_k \hat{s}_k - a_1)^2.$$

In the sequel we shall show that $U_4$ converges to zero in probability. Note that

$$\frac{\hat{s}_k^* \hat{s}_k}{\hat{s}_k^* \hat{R}_k \hat{s}_k} \leq \frac{1}{\sigma^2}, \quad \hat{s}_k^* = \bar{s}_k^* + E \hat{s}_k^*.$$

This gives

$$|U_4| \leq M \sum_{k=1}^{K} \hat{s}_k^* \bar{s}_k |\hat{s}_k^* \hat{R}_k \hat{s}_k - a_1|^3$$

\begin{equation}
\leq M \sum_{k=1}^{K} (\bar{s}_k^* \bar{s}_k + \bar{s}_k^* E \hat{s}_k + (E \hat{s}_k^*) \bar{s}_k + E \hat{s}_k^* E \hat{s}_k)
\times |\hat{s}_k^* \hat{R}_k \hat{s}_k - a_1 + \bar{s}_k^* \hat{R}_k \hat{s}_k + (E \hat{s}_k^*) \hat{R}_k \hat{s}_k + (E \hat{s}_k^*) \hat{R}_k E \hat{s}_k|^3.
\end{equation}

Then we need to compute each term of the above expansion.
It is observed that

(A.3) \[
\left( \sum_{k=1}^{K} \bar{s}_k^* \bar{s}_k \Big| \tilde{s}_k^* \bar{R}_k \tilde{s}_k - \frac{1}{N} \text{Tr} \bar{R}_k \right)^3
\]

\[
\leq M \left( \sum_{k=1}^{K} \bar{s}_k^* \bar{s}_k \left| \tilde{s}_k^* \bar{R}_k \tilde{s}_k - \frac{1}{N} \text{Tr} \bar{R}_k \right|^3 + \left| \tilde{s}_k^* (E \hat{\bar{S}}_k) \tilde{s}_k \right|^3 + \left| \tilde{s}_k^* (E \hat{\bar{S}}_k) \tilde{s}_k \right|^3 \right)
\]

(A.4) \[
+ \left| \tilde{s}_k^* \bar{R}_k \left( E \hat{\bar{S}}_k \right) \tilde{s}_k \right|^3 + \left| \tilde{s}_k^* (E \hat{\bar{S}}_k) \tilde{s}_k \right|^3 \right).
\]

It is a simple matter to prove that

\[
\lim_{N \to \infty} E(\tilde{v}_{11})^2 = 1.
\]

Then, appealing (5.4), we have

(A.5) \[
E \left( \sum_{k=1}^{K} \bar{s}_k^* \bar{s}_k \left| \tilde{s}_k^* \bar{R}_k \tilde{s}_k - \frac{1}{N} \text{Tr} \bar{R}_k \right|^3 \right)
\]

\[
\leq M \sum_{k=1}^{K} \left( E \left| \tilde{s}_k^* \bar{s}_k - 1 \right|^2 \right)^{1/2} \left( E \left| \tilde{s}_k^* \bar{R}_k \tilde{s}_k - \frac{1}{N} \text{Tr} \bar{R}_k \right|^6 \right)^{1/2} = o \left( \frac{1}{\sqrt{N}} \right).
\]

For the first term in (A.4), with the notation \( e = \frac{1}{\sqrt{N}}(1, \ldots, 1)^* \) and \( G = \bar{S}_k \bar{P}_k (E \hat{\bar{S}}_k) (E \hat{\bar{S}}_k) \bar{P}_k \bar{S}_k^* \), analogously,

(A.5) \[
E \left( \sum_{k=1}^{K} \bar{s}_k^* \bar{s}_k \left| \tilde{s}_k^* \bar{S}_k \bar{P}_k (E \hat{\bar{S}}_k) \tilde{s}_k \right|^3 \right)
\]

\[
\leq M E \left( \sum_{k=1}^{K} \bar{s}_k^* \bar{s}_k \left| \tilde{s}_k^* \bar{S}_k \bar{P}_k (E \hat{\bar{S}}_k) \tilde{s}_k - \frac{1}{N} \text{Tr} \bar{S}_k \bar{P}_k E \hat{\bar{S}}_k \right|^3 \right)
\]

\[
+ M E \left( \sum_{k=1}^{K} \bar{s}_k^* \bar{s}_k \left| \frac{1}{N} \text{Tr} \bar{S}_k \bar{P}_k E \hat{\bar{S}}_k \right|^3 \right)
\]

\[
\leq \frac{M}{N^{3/2}} \sum_{k=1}^{K} \left[ \left( E \left( \frac{1}{N} \text{Tr} G \right)^3 + E \frac{1}{N} \text{Tr}(G)^3 \right)^{1/2}
\right.
\]

\[
\left. + E \left( \frac{1}{N} \text{Tr} G \right)^{3/2} + E \frac{1}{N} \text{Tr}(G)^{3/2} \right]
\]

\[
+ M \left( E \hat{v}_{11} \right)^3 \left( \sum_{k=1}^{K} E \left| \sum_{j \neq k} p_j \tilde{s}_j^* e \right|^3 \right) = o(1).
\]
Indeed,

\[
(E \hat{v}_{11})^3 \left( \sum_{k=1}^{K} E \left| \sum_{j \neq k} p_j \tilde{s}_j^* e \right|^3 \right) \leq M (E \hat{v}_{11})^3 K N^{3/2} = o\left( \frac{1}{N^2} \right),
\]

and since

\[
G \leq MN (E \hat{v}_{11})^2 \bar{s}_k e e^* \bar{s}_k^* \tilde{s}_k^* \tilde{s}_k \frac{1}{n},
\]

we have

\[
E \left( \frac{1}{N} \text{Tr} G \right)^3 \leq MN^3 (E \hat{v}_{11})^6 E \left( \frac{1}{N} \text{Tr} \bar{s}_1 \tilde{s}_1^* \right)^3.
\]

The other terms can be estimated similarly. Regarding the last term in (A.4), we have

\[
\sum_{k=1}^{K} \bar{s}_k \tilde{s}_k |\text{Tr} \hat{R}_k - a_1|^3 \leq M \sum_{k=1}^{K} (\bar{s}_k \tilde{s}_k)^3 \| \hat{S}_k \|^6 = o\left( \frac{1}{n} \right),
\]

where \( \| \cdot \| \) denotes the spectral norm of a matrix. Therefore, (A.3) converges to zero in probability.

Now

\[
\sum_{k=1}^{K} \bar{s}_k \tilde{s}_k \left| \frac{1}{N} \text{Tr} \hat{R}_k - a_1 \right|^3 \leq M \sum_{k=1}^{K} \bar{s}_k \tilde{s}_k \left( \left| \frac{1}{N} \text{Tr} \hat{R}_k - a_1 \right|^3 + \left| \frac{1}{N} \text{Tr} \bar{s}_k \tilde{s}_k \right|^3 \left| \hat{S}_k^* \right|^3 \right)
\]

(A.6)

\[
+ \left| \frac{1}{N} \text{Tr}(E \hat{S}_k) \tilde{s}_k \tilde{s}_k^* \right|^3 + \left| \frac{1}{N} \text{Tr} E \hat{S}_k \tilde{s}_k \tilde{s}_k^* \right|^3.
\]

For its first term one can get

\[
E \sum_{k=1}^{K} \bar{s}_k \tilde{s}_k \left| \frac{1}{N} \text{Tr} \hat{R}_k - a_1 \right|^3 \leq M \sum_{k=1}^{K} (E(\bar{s}_k \tilde{s}_k - 1)^2)^{1/2} \left( E \left| \frac{1}{N} \text{Tr} \hat{R}_k - a_1 \right|^6 \right)^{1/2}
\]

\[
+ M \sum_{k=1}^{K} E \left| \frac{1}{N} \text{Tr} \hat{R}_k - a_1 \right|^3 = o(1),
\]

as

\[
E \left| \frac{1}{N} \text{Tr} \hat{R}_k - a_1 \right|^p = E \left| \frac{1}{N} \sum_{j \neq k} p_j \bar{s}_j \tilde{s}_j - c \right|^p = O\left( \frac{1}{N^{p/2}} \right).
\]
The argument that the remaining terms of (A.6) converge to zero in probability is similar to above, even simpler and then omitted. Hence, (A.6) converges to zero in probability. This, together with (A.3), leads to

\[(A.7) \sum_{k=1}^{K} \bar{s}_k^* \bar{s}_k |\bar{s}_k^* \hat{R}_k \bar{s}_k - a_1|^3 \xrightarrow{i.p.} 0,\]

which is one term in (A.2). All remaining items of (A.2) can be computed similarly, so we omit it here. Consequently,

\[U_4 \xrightarrow{i.p.} 0\]

and

\[(A.8) \sum_{k=1}^{K} \frac{p_k (\bar{s}_k^* \bar{s}_k)^2}{\bar{s}_k^* R_k \bar{s}_k} = U_1 + U_2 + U_3 + o_p(1).\]

Analogously, one can also show that

\[(A.9) \sum_{k=1}^{K} \frac{p_k (\bar{s}_k^* \bar{s}_k)^2}{\bar{s}_k^* R_k \bar{s}_k} = V_1 + V_2 + V_3 + o_p(1),\]

where

\[V_1 = \frac{1}{a_1} \sum_{k=1}^{K} p_k (\bar{s}_k^* \bar{s}_k)^2, \quad V_2 = -\frac{1}{a_1^2} \sum_{k=1}^{K} p_k (\bar{s}_k^* \bar{s}_k)^2 (\bar{s}_k^* \hat{R}_k \bar{s}_k - \bar{s}_k^* \bar{s}_k),\]

and

\[V_3 = \frac{1}{a_1^3} \sum_{k=1}^{K} p_k (\bar{s}_k^* \bar{s}_k)^2 (\bar{s}_k^* \hat{R}_k \bar{s}_k - \bar{s}_k^* \bar{R}_k \bar{s}_k - a_1)^2.\]

In the following, we show that \(U_i - V_i, i = 1, 2, 3,\) converge to zero in probability. Since all the calculations for \(U_i - V_i\) are similar, as an illustration, we consider \(U_2 - V_2\) only.

Write

\[(A.10) -a_1^2 (U_2 - V_2) = U_{21} + U_{22},\]

where

\[U_{21} = \sum_{k=1}^{K} p_k (\bar{s}_k^* \hat{s}_k)^2 - (\bar{s}_k^* \bar{s}_k)^2 (\bar{s}_k^* \hat{R}_k \bar{s}_k - \bar{s}_k^* \bar{s}_k - a_1)\]

and

\[U_{22} = \sum_{k=1}^{K} p_k (\bar{s}_k^* \hat{s}_k)^2 (\bar{s}_k^* \hat{R}_k \bar{s}_k - \bar{s}_k^* \bar{R}_k \bar{s}_k).\]
Furthermore, expand $U_{21}$ as follows:

\[
U_{21} = \sum_{k=1}^{K} p_k (\tilde{s}_k^* E \hat{s}_k)^2 + ((E \hat{s}_k^*) \tilde{s}_k)^2 + (E \hat{s}_k^* E \hat{s}_k)^2 + 2\tilde{s}_k^* \tilde{s}_k E \hat{s}_k \\
+ 2\tilde{s}_k^* \tilde{s}_k (E \hat{s}_k^*) \tilde{s}_k + 2\tilde{s}_k^* (E \hat{s}_k^*) (E \hat{s}_k^*) \tilde{s}_k + 2\tilde{s}_k^* \tilde{s}_k E \hat{s}_k^* E \hat{s}_k \\
+ 2(E \hat{s}_k^*) \tilde{s}_k (E \hat{s}_k^*) E \hat{s}_k + 2\tilde{s}_k^* E \hat{s}_k E \hat{s}_k^* E \hat{s}_k \\
\times (\tilde{s}_k^* \hat{R}_k \tilde{s}_k - a_1),
\]

which can be easily proved to tend to zero in probability. For example, for one of the terms,

\[
E \sum_{k=1}^{K} p_k (\tilde{s}_k^* \tilde{s}_k) (E \hat{s}_k) (\tilde{s}_k^* \hat{R}_k \tilde{s}_k - a_1) \\
\leq \sqrt{N} \varepsilon N E \hat{v}_{11} \sum_{k=1}^{K} (E (\tilde{s}_k^* \tilde{s}_k)^2)^{1/2} (E (\tilde{s}_k^* \hat{R}_k \tilde{s}_k - a_1)^2)^{1/2} = o\left(\frac{1}{\sqrt{N}}\right).
\]

For $U_{22}$, we have

\[
U_{22} \leq \sum_{k=1}^{K} p_k (\tilde{s}_k^* \hat{s}_k)^2 (\tilde{s}_k^* \hat{R}_k \tilde{s}_k - \tilde{s}_k^* \hat{R}_k \tilde{s}_k) + (E \hat{s}_k^*) \hat{R}_k \tilde{s}_k + (E \hat{s}_k^*) \hat{R}_k (E \hat{s}_k).
\]

Moreover, it is observed that

\[
\sum_{k=1}^{K} p_k (\tilde{s}_k^* \hat{s}_k)^2 (\tilde{s}_k^* \hat{R}_k \tilde{s}_k - \tilde{s}_k^* \hat{R}_k \tilde{s}_k) \\
= \sum_{k=1}^{K} p_k (\tilde{s}_k^* \hat{s}_k)^2 \\
\times (\tilde{s}_k^* \hat{S}_k P_k (E \hat{s}_k^*) \tilde{s}_k + \tilde{s}_k^* (E \hat{S}_k) P_k \tilde{s}_k + \tilde{s}_k^* (E \hat{S}_k) P_k (E \hat{s}_k^*) \tilde{s}_k).
\]

As for the first sum of the above expansion,

\[
E \left| \sum_{k=1}^{K} p_k (\tilde{s}_k^* \hat{s}_k)^2 \tilde{s}_k^* \tilde{s}_k P_k (E \hat{s}_k^*) \tilde{s}_k \right| \\
\leq M \sum_{k=1}^{K} (E (\tilde{s}_k^* \tilde{s}_k)^4)^{1/2} \left( E \left( \tilde{s}_k^* \tilde{s}_k P_k (E \hat{s}_k^*) \tilde{s}_k - \frac{1}{N} \text{Tr} \tilde{s}_k P_k E \hat{s}_k^* \right)^2 \right)^{1/2} \\
+ E \left| \sum_{k=1}^{K} p_k (\tilde{s}_k^* \hat{s}_k)^2 \frac{1}{N} \text{Tr} \tilde{s}_k P_k E \hat{s}_k^* \right| = O\left(\frac{1}{N^2}\right),
\]
where we make use of
\[ E \left( \hat{s}_k^* \hat{s}_k \right)^4 \leq M E \left( \hat{s}_k^* \hat{s}_k - E(\hat{s}_k^* \hat{s}_k) \right)^4 + M (E(\hat{s}_k^* \hat{s}_k))^4 \leq \frac{M}{N} + M \]

and
\[ E \left( \hat{s}_k^* \hat{s}_k P_k (E \hat{S}_k^*) \hat{s}_k - \frac{1}{N} \text{Tr} \hat{S}_k P_k E \hat{S}_k^* \right)^2 \leq \frac{M}{N^2} \text{Tr} E \hat{S}_k^* E \hat{S}_k \leq \frac{M \|E \hat{S}_k\|^2}{N^2} \text{Tr} E \hat{S}_k \hat{S}_k = o \left( \frac{1}{N^3} \right) \]

and
\[ E \left| \sum_{k=1}^{K} p_k (\hat{s}_k^* \hat{s}_k) \frac{1}{N} \text{Tr} \hat{S}_k P_k E \hat{S}_k^* \right| = o \left( \frac{1}{N^{1/2}} \right) \].

Similarly, one can also verify that the other two terms of (A.12) converge to zero in probability, and all the other items of (A.11) converge to zero in probability. So \( U_{22} \xrightarrow{i.p.} 0 \), and then \( U_2 - V_2 \xrightarrow{i.p.} 0 \), as expected. Finally, we get
\[ \sum_{k=1}^{K} \hat{\beta}_k - \sum_{k=1}^{K} \beta_k \xrightarrow{i.p.} 0. \]

Similarly, one can perform the re-normalization step, but it is omitted here.

**Acknowledgments.** The authors would like to thank a referee for his valuable comments, which greatly improved the exposition of this work.

**REFERENCES**


