Quantification of model bias underlying the phenomenon of "Einstein from noise"

Shao-Hsuan Wang, Yi-Ching Yao, Wei-Hau Chang and I-Ping Tu

Academia Sinica

Supplementary Material

The online Supplementary Material contains the proofs of Lemmas A6-A8.

S1 Notations

- Vectors and matrices are denoted by boldface letters, e.g. $\boldsymbol{X}, \boldsymbol{V}$.
- $\mathcal{S}^{p-1} = \{(x_1, \ldots, x_p) : x_1^2 + \cdots + x_p^2 = 1\}$ is the $(p-1)$-dimensional unit sphere.
- $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{ip})^\top$, $i = 1, \ldots, n$ are iid uniformly distributed on $\mathcal{S}^{p-1}$.
- $\boldsymbol{r} = (1, 0, \ldots, 0)^\top \in \mathcal{S}^{p-1}$.
- $(\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(n)})$ is a permutation of $(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ such that $\boldsymbol{r}^\top \boldsymbol{X}^{(1)} \geq \cdots \geq \boldsymbol{r}^\top \boldsymbol{X}^{(n)}$ (or equivalently, $X_1^{(1)} \geq \cdots \geq X_1^{(n)}$).

- Θ_i denotes the angle between $\boldsymbol{X}_i$ and $\boldsymbol{r}$.

- Θ_i, $i = 1, \ldots, n$ are iid with the common cdf

$$\begin{aligned} F_p(\theta) &= \int_0^\theta \frac{1}{\sqrt{\pi}} \frac{\Gamma(p/2)}{\Gamma((p-1)/2)} (\sin x)^{p-2} dx \\ &= \int_{\cos\theta}^1 \frac{1}{\sqrt{\pi}} \frac{\Gamma(p/2)}{\Gamma((p-1)/2)} (1-u^2)^{\frac{p-3}{2}} du,\ \theta \in [0, \pi]. \end{aligned}$$

- $\Theta_{1:n} \leq \Theta_{2:n} \leq \cdots \leq \Theta_{n:n}$ denotes the order statistics of $\Theta_1, \ldots, \Theta_n$.

- $\overline{\boldsymbol{X}}_m = \frac{1}{m} \sum_{i=1}^m \boldsymbol{X}^{(i)}$.

- $\rho_{n,p,m} = \boldsymbol{r}^\top \frac{\overline{\boldsymbol{X}}_m}{\|\overline{\boldsymbol{X}}_m\|}$.

- $\boldsymbol{V}_i$ ($\boldsymbol{V}^{(i)}$, respectively) $\in \mathcal{S}^{p-2}$ denotes the normalized subvector of $\boldsymbol{X}_i$ ($\boldsymbol{X}^{(i)}$, respectively) with the first component deleted.

- $\boldsymbol{X}_i = (X_{i1}, (1 - X_{i1}^2)^{1/2} \boldsymbol{V}_i^\top)^\top$.

- $\boldsymbol{X}^{(i)} = (X_1^{(i)}, \nu_i \boldsymbol{V}^{(i)\top})^\top$, where $\nu_i = (1 - X_1^{(i)2})^{1/2}$.

- $\boldsymbol{V}'_i$, $i = 1, \ldots, n$ are iid uniformly distributed on $\mathcal{S}^{p-2}$ independent of $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$.

- $A_{n,p,m} = \left(\frac{1}{m} \sum_{i=1}^m X_1^{(i)} \right)^2$.

- $V_{n,p,m} = \left\| \frac{1}{m} \sum_{i=1}^m \nu_i \boldsymbol{V}'_i \right\|^2$, where $\nu_i = (1 - X_1^{(i)2})^{1/2}$.

- $U_1, U_2, \ldots$ are iid uniform (0,1) random variables.

- $U_{1:n} \le \cdots \le U_{n:n}$ denotes the order statistics of $U_1, \ldots, U_n$.
- $\xi_1, \xi_2, \ldots$ are iid exponential random variables with mean 1.
- $S_0 = 0,\ S_i = \xi_1 + \cdots + \xi_i,\ i = 1, 2, \ldots$

S2 Proof of Lemma A6

Proof of Lemma A6(i). For any fixed (large) $C > 0$ and for $i = 1, \ldots, m$, let $\underline{t}_{n,i} \in [0, 1)$ be such that

$$\begin{aligned}\frac{1}{2} p \ln(1 - \underline{t}_{n,i}^2) &= \min\left\{-\ln \frac{n}{i} + \frac{1}{2}\ln\ln\frac{n}{i} + \frac{1}{2}\ln(4\pi) - C, 0\right\} \\ &= -\ln\frac{n}{i} + \frac{1}{2}\ln\ln\frac{n}{i} + \frac{1}{2}\ln(4\pi) - C \text{ (for large } n). \end{aligned} \tag{S.1}$$

Since $m/n = o(1)$ and $(\ln n)^2/p = O(1)$, we have $\underline{t}_{n,i} \to 0^+$ (uniformly in $1 \le i \le m$) as $n \to \infty$. Thus

$$\ln(1 - \underline{t}_{n,i}^2) = -\frac{2\ln(n/i)}{p}\left(1 + O\left(\frac{\ln\ln(n/m)}{\ln(n/m)}\right)\right), \tag{S.2}$$

$$\ln(1 - \underline{t}_{n,i}^2) = -\underline{t}_{n,i}^2(1 + O(\underline{t}_{n,i}^2)) = -\underline{t}_{n,i}^2\left(1 + O\left(\frac{\ln n}{p}\right)\right), \tag{S.3}$$

and

$$\frac{2\ln(n/i)}{p\underline{t}_{n,i}^2} = 1 + O\left(\frac{\ln\ln(n/m)}{\ln(n/m)}\right) + O\left(\frac{\ln n}{p}\right) = 1 + O\left(\frac{\ln\ln(n/m)}{\ln(n/m)}\right), \tag{S.4}$$

where the $O\left(\frac{\ln\ln(n/m)}{\ln(n/m)}\right)$ term is uniform in $1 \le i \le m$. (Note that $\frac{\ln\ln x}{\ln x}$ is decreasing in $x \in (e^e, \infty)$.)

Note that $S_n = \xi_1 + \cdots + \xi_n$ has the gamma distribution with the scale and shape parameters equal to 1 and n, respectively. For large n,

$$n^{1/2} \ln\left(\frac{S_n}{n}\right) = n^{1/2}\left(\frac{S_n}{n} - 1\right) + O_p\left(n^{-1/2}\right) \xrightarrow{d} N(0,1).$$

Furthermore, by making use of the gamma tail probabilities, it can be readily shown that there exists an $M > 0$ such that

$$\max\left\{\mathrm{E}\left[\ln\frac{S_i}{i}\right]^4, \mathrm{E}\left[\ln\frac{S_{i+1}}{i}\right]^4\right\} \le \frac{M}{i^2},\ i = 1, 2, \ldots$$

Let $h_{n,i}(x) = \ln(nx/i)$, $i = 1, \ldots, m$. We have by Lemma A3

$$\begin{aligned}\mathrm{E}\left[h_{n,i}(U_{i:n})\right]^4 &= \mathrm{E}\left[\ln\left(\frac{nS_i}{iS_{n+1}}\right)\right]^4 = \mathrm{E}\left[\ln\frac{S_i}{i} - \ln\frac{S_{n+1}}{n}\right]^4 \\ &\le 8\left\{\mathrm{E}\left[\ln\frac{S_i}{i}\right]^4 + \mathrm{E}\left[\ln\frac{S_{n+1}}{n}\right]^4\right\} \le \frac{16M}{i^2},\ i = 1, \ldots, m.\end{aligned} \tag{S.5}$$

By the definition of $\overline{F}_p(\cdot)$ given in (A.2), we have

$$\begin{aligned}
\underline{h}_{n,i} := h_{n,i}\left(\overline{F}_p(\cos^{-1}\underline{t}_{n,i})\right) &= \ln\left\{\frac{n}{i\sqrt{\pi}}\frac{\Gamma(p/2)}{\Gamma((p-1)/2)}\left(\frac{(1-\underline{t}_{n,i}^2)^{p-1}}{(p-1)^2\underline{t}_{n,i}^2}\right)^{1/2}\right\} \\
&= \ln\left(\frac{n}{i}\sqrt{\frac{(1-\underline{t}_{n,i}^2)^{p-1}}{2\pi p\underline{t}_{n,i}^2}}\right) + O\left(\frac{1}{p}\right) \quad \text{(by (A.15))} \\
&= \ln\left(\frac{n}{i}\sqrt{\frac{(1-\underline{t}_{n,i}^2)^{p}}{2\pi p\underline{t}_{n,i}^2}}\right) + O\left(\frac{\ln n}{p}\right) \quad \text{(by (S.2))} \\
&= \ln\left(\sqrt{\frac{2\ln(n/i)}{p\underline{t}_{n,i}^2}}e^{-C}\right) + O\left(\frac{\ln n}{p}\right) \quad \text{(by (S.1))} \\
&= -C + O\left(\frac{\ln\ln(n/m)}{\ln(n/m)}\right) + O\left(\frac{\ln n}{p}\right) \quad \text{(by (S.4))} \\
&= -C + O\left(\frac{\ln\ln(n/m)}{\ln(n/m)}\right). \qquad \text{(S.6)}
\end{aligned}$$

Similarly to $\underline{t}_{n,i}$ as defined in (S.1), let $\overline{t}_{n,i} \in [0,1)$ be such that

$$\frac{1}{2}p\ln(1-\overline{t}_{n,i}^2) = \min\left\{-\ln\frac{n}{i} + \frac{1}{2}\ln\ln\frac{n}{i} + \frac{1}{2}\ln(4\pi) + C,\ 0\right\}. \qquad \text{(S.7)}$$

It can be shown that

$$\overline{h}_{n,i} := h_{n,i}\left(\overline{F}_p(\cos^{-1}\overline{t}_{n,i})\left(1+\frac{1}{(p-3)\overline{t}_{n,i}^2}\right)^{-1}\right) = C + O\left(\frac{\ln\ln(n/m)}{\ln(n/m)}\right). \qquad \text{(S.8)}$$

By (S.6) and (S.8), we have

$$\underline{h}_{n,i} < -\frac{C}{2} \text{ and } \overline{h}_{n,i} > \frac{C}{2} \text{ for } i = 1, \ldots, m \text{ for large } n. \qquad \text{(S.9)}$$

Thus, we have for large n

$$
\begin{aligned}
& \mathrm{P}\left(\max_{1\le i\le m}\left|p\ln\left(\sin\,\Theta_{i:n}\right)+\ln\frac{n}{i}-\frac{1}{2}\ln\ln\frac{n}{i}-\frac{1}{2}\ln(4\pi)\right|>C\right)\\
\le\ & \sum_{i=1}^{m}\mathrm{P}\left(\left|p\ln(\sin\,\Theta_{k:n})+\ln\frac{n}{i}-\frac{1}{2}\ln\ln\frac{n}{i}-\frac{1}{2}\ln(4\pi)\right|>C\right)\\
=\ & \sum_{i=1}^{m}\left\{\mathrm{P}\left(p\ln(\sin\,\Theta_{i:n})<\frac{p}{2}\ln(1-\underline{t}_{n,i}^{2})\right)+\mathrm{P}\left(p\ln(\sin\,\Theta_{i:n})>\frac{p}{2}\ln(1-\bar{t}_{n,i}^{2})\right)\right\}\\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\text{(by (S.1) and (S.7))}\\
=\ & \sum_{i=1}^{m}\left\{\mathrm{P}(\cos^2\,\Theta_{i:n}>\underline{t}_{n,i}^{2})+\mathrm{P}(\cos^2\,\Theta_{i:n}<\bar{t}_{n,i}^{2})\right\}\\
\le\ & \sum_{i=1}^{m}\left\{\mathrm{P}(\Theta_{i:n}<\cos^{-1}\underline{t}_{n,i})+\mathrm{P}\left(\cos^{-1}\,\bar{t}_{n,i}<\Theta_{i:n}<\frac{\pi}{2}\right)+2\mathrm{P}\left(\Theta_{i:n}>\frac{\pi}{2}\right)\right\}\\
\le\ & \sum_{i=1}^{m}\left\{\mathrm{P}\left(F_p(\Theta_{i:n})<\overline{F}_p(\cos^{-1}\,\underline{t}_{n,i})\right)+\mathrm{P}\left(F_p(\Theta_{i:n})>\overline{F}_p\left(\cos^{-1}\,\bar{t}_{n,i}\right)\left(1+\frac{1}{(p-3)\bar{t}_{t,n}^{2}}\right)^{-1}\right)\right.\\
& \qquad\qquad\qquad\qquad\left.+2\mathrm{P}\left(F_p(\Theta_{i:n})>\frac{1}{2}\right)\right\}\quad\text{(by Lemma A2)}\\
=\ & \sum_{i=1}^{m}\left\{\mathrm{P}(h_{n,i}(U_{i:n})<\underline{h}_{n,i})+\mathrm{P}(h_{n,i}(U_{i:n})>\overline{h}_{n,i})+2\mathrm{P}\left(U_{i:n}>\frac{1}{2}\right)\right\}\quad\left(\text{since } F_p(\Theta_{i:n})\stackrel{d}{=}U_{i:n}\right)\\
\le\ & \left(\frac{C}{2}\right)^{-4}\sum_{i=1}^{m}\left\{\mathrm{E}[h_{n,i}(U_{i:n})]^4+\mathrm{E}[h_{n,i}(U_{i:n})]^4\right\}+2m\mathrm{P}\left(U_{m:n}>\frac{1}{2}\right)\quad\text{(by (S.9))}\\
\le\ & \frac{512M}{C^4}\sum_{i=1}^{m}\frac{1}{i^2}+2m\mathrm{P}\left(U_{m:n}>\frac{1}{2}\right)\quad\text{(by (S.5))}\\
\le\ & \frac{512M}{C^4}\frac{\pi^2}{6}+2m\mathrm{P}\left(U_{m:n}>\frac{1}{2}\right).
\end{aligned}
\tag{S.10}
$$

Since $\lim_{n\to\infty} m\,\mathrm{P}\left(U_{m:n}>\frac{1}{2}\right)=0$ and since $\frac{512M}{C^4}\frac{\pi^2}{6}$ can be made arbitrarily small by choosing a sufficiently large C, Lemma A6(i) follows from (S.10).

□

Proof of Lemma A6(ii). By Lemma A6(i), for any given $0 < \eta < 1$, there exist $C > 0$ and $N > 0$ such that $\mathrm{P}(A_n) > \eta$ for all $n > N$, where the event A_n is defined by

$$A_n = \left\{ \max_{1\le i\le m} \left| p \ln(\sin\, \Theta_{i:n}) + \ln\, \frac{n}{i} - \frac{1}{2} \ln\ln\, \frac{n}{i} \right| \le C \right\}.$$

Let

$$u_{n,i} = 1 - \exp\left(\frac{-2\ln(n/i)}{p} + \frac{\ln\ln(n/i)}{p} + \frac{2C}{p} \right)$$

and

$$v_{n,i} = 1 - \exp\left(\frac{-2\ln(n/i)}{p} + \frac{\ln\ln(n/i)}{p} - \frac{2C}{p} \right),$$

so that $A_n = \{u_{n,i} \le \cos^2\Theta_{i:n} \le v_{n,i}, i = 1, \dots, m\}$.

Since $(\ln n)^2/p = O(1)$, there exist $C' > 0$ and $N' > 0$ such that for $n > N'$

$$\max_{1\le i\le m} \left| u_{n,i} + \left\{ \frac{-2\ln(n/i)}{p} + \frac{\ln\ln(n/i)}{p} + \frac{2C}{p} \right\} \right| \le C' \frac{(\ln\, n)^2}{p^2},$$
$$\max_{1\le i\le m} \left| v_{n,i} + \left\{ \frac{-2\ln(n/i)}{p} + \frac{\ln\ln(n/i)}{p} - \frac{2C}{p} \right\} \right| \le C' \frac{(\ln\, n)^2}{p^2}.$$

Let $C'' = \sup_{n>N'} (\ln\, n)^2/p$, so that for $n > N'$ and $i = 1, \dots, m$,

$$u_{n,i} \ge \frac{2\ln(n/i)}{p} - \frac{\ln\ln(n/i)}{p} - \frac{2C + C'C''}{p}$$

and

$$v_{n,i} \le \frac{2\ln(n/i)}{p} - \frac{\ln\ln(n/i)}{p} + \frac{2C + C'C''}{p}.$$

We have for $n > \max\{N, N'\}$

$$\begin{aligned}
\eta < \mathrm{P}(A_n) &= \mathrm{P}(u_{n,i} \le \cos^2 \Theta_{i:n} \le v_{n,i}, i = 1, \dots, m) \\
&\le \mathrm{P}\left(\max_{1\le i\le m} \left| \cos^2 \Theta_{i:n} - \frac{2\ln(n/i)}{p} + \frac{\ln\ln(n/i)}{p} \right| < \frac{2C + C'C''}{p} \right) \\
&= \mathrm{P}\left(\max_{1\le i\le m} \left| \frac{-p}{2} \cos^2 \Theta_{i:n} + \ln\, \frac{n}{i} - \frac{1}{2}\ln\ln\, \frac{n}{i} \right| < C + \frac{C'C''}{2} \right).
\end{aligned}$$

This proves Lemma A6(ii). □

S3 Proof of Lemma A7

Proof of Lemma A7(i). For $\boldsymbol{w}_m = (w_1, \dots, w_m)$ and $u_i > 0$, let

$$g(u_1, \dots, u_m; \boldsymbol{w}_m) := \frac{1}{m} \sum_{i=1}^{m} w_i \left\{ \ln \frac{n}{i} + \ln u_i \right\}.$$

Let $\boldsymbol{w}'_m = (1, 1, \dots, 1)$ and $\boldsymbol{w}''_m = (1/m, 2/m, \dots, m/m)$. To derive the asymptotic behavior of $A_{n,p,m}$, we first establish

$$g\left(\overline{F}_p(\Theta_{1:n}), \dots, \overline{F}_p(\Theta_{m:n}); \boldsymbol{w}_m\right) = o_p(1) \tag{S.11}$$

for $\boldsymbol{w}_m = \boldsymbol{w}'_m,\ \boldsymbol{w}''_m$. To show (S.11), it suffices to prove that for $\boldsymbol{w}_m = \boldsymbol{w}'_m,\ \boldsymbol{w}''_m$,

$$g\left(F_p(\Theta_{1:n}), \dots, F_p(\Theta_{m:n}); \boldsymbol{w}_m\right) \stackrel{d}{=} g\left(U_{1:n}, \dots, U_{m:n}; \boldsymbol{w}_m\right) = O_p(m^{-1/2}), \tag{S.12}$$

$$g\left(F_p(\Theta_{1:n}), \dots, F_p(\Theta_{m:n}); \boldsymbol{w}_m\right) = g\left(\overline{F}_p(\Theta_{1:n}), \dots, \overline{F}_p(\Theta_{m:n}); \boldsymbol{w}_m\right) + o_p(1). \tag{S.13}$$

To prove (S.12), we have by Lemma A3

$$
\begin{aligned}
g(U_{1:n},\ldots,U_{m:n};\boldsymbol{w}_m) &\overset{d}{=} g\left(\frac{S_1}{S_{n+1}},\ldots,\frac{S_m}{S_{n+1}};\boldsymbol{w}_m\right)\\
&= \frac{1}{m}\sum_{i=1}^{m} w_i\left\{\ln\frac{n}{m}+\sum_{\ell=i}^{m}\frac{1}{\ell}+\ln\frac{S_i}{S_{n+1}}\right\}+\frac{1}{m}\sum_{i=1}^{m} w_i\left\{-\sum_{\ell=i}^{m}\frac{1}{\ell}+\ln\frac{m}{i}\right\}\\
&:= R_1(\boldsymbol{w}_m)+R_2(\boldsymbol{w}_m). \qquad \text{(S.14)}
\end{aligned}
$$

Here,

$$
\begin{aligned}
R_1(\boldsymbol{w}_m) &= \frac{1}{m}\sum_{i=1}^{m} w_i\left(\ln\frac{n}{m}+\sum_{\ell=i}^{m}\frac{1}{\ell}+\ln\frac{S_i}{S_{n+1}}\right)\\
&= \frac{1}{m}\sum_{i=1}^{m} w_i\left(\ln\frac{S_i}{S_{m+1}}+\sum_{\ell=i}^{m}\frac{1}{\ell}\right)+\left(\ln\frac{S_{m+1}}{m}-\ln\frac{S_{n+1}}{n}\right)\left(\frac{1}{m}\sum_{i=1}^{m} w_i\right)\\
&:= R_1'(\boldsymbol{w}_m)+R_1''(\boldsymbol{w}_m), \qquad \text{(S.15)}
\end{aligned}
$$

where

$$
\begin{aligned}
R_1'(\boldsymbol{w}_m) &= \frac{1}{m}\sum_{i=1}^{m} w_i\left(\ln\frac{S_i}{S_{m+1}}+\sum_{\ell=i}^{m}\frac{1}{\ell}\right)\\
&\overset{d}{=} \frac{1}{m}\sum_{i=1}^{m} w_i\left(\ln U_{i:m}+\sum_{\ell=i}^{m}\frac{1}{\ell}\right)\\
&\overset{d}{=} -\frac{1}{m}\sum_{i=1}^{m} w_i\left(e_{(m+1-i):m}-\sum_{\ell=i}^{m}\frac{1}{\ell}\right) = -\frac{1}{m}\sum_{i=1}^{m} w_i\sum_{\ell=i}^{m}\left(Y_\ell-\frac{1}{\ell}\right), \qquad \text{(S.16)}
\end{aligned}
$$

where $\xi_{1:m}\le\xi_{2:m}\le\cdots\le\xi_{m:m}$ are the order statistics of iid exponential random variables $\xi_1,\ldots,\xi_m$ with mean 1 and $Y_i := \xi_{(m-i+1):m}-\xi_{(m-i):m}$, $i = 1,\ldots,m$ ($\xi_{0:m}:=0$). Note that the Y_i's are independent and exponentially distributed with respective means $1/i, i = 1, \ldots, m$.

For $\boldsymbol{w}_m = \boldsymbol{w}'_m = (1, 1, \ldots, 1)$, by Chebyshev's inequality,

$$R'_1(\boldsymbol{w}'_m) \stackrel{d}{=} \frac{-1}{m} \sum_{i=1}^{m} \sum_{\ell=i}^{m} \left(Y_\ell - \frac{1}{\ell} \right) = \frac{-1}{m} \sum_{i=1}^{m} i \left(Y_i - \frac{1}{i} \right) = O_p \left(m^{-1/2} \right). \tag{S.17}$$

Also,

$$R''_1(\boldsymbol{w}'_m) := \left(\ln \frac{S_{m+1}}{m} - \ln \frac{S_{n+1}}{n} \right) \left(\frac{1}{m} \sum_{i=1}^{m} w'_i \right) = \ln \frac{S_{m+1}}{m} - \ln \frac{S_{n+1}}{n} = O_p(m^{-1/2}),$$

which together with (S.17) implies

$$R_1(\boldsymbol{w}'_m) = R'_1(\boldsymbol{w}'_m) + R''_1(\boldsymbol{w}'_m) = O_p(m^{-1/2}). \tag{S.18}$$

Since $m^{-1} \sum_{i=1}^{m} \ln (m/i) = 1 + O(\ln m/m)$, we have

$$R_2(\boldsymbol{w}'_m) := \frac{1}{m} \sum_{i=1}^{m} \left(- \sum_{\ell=i}^{m} \frac{1}{\ell} + \ln \frac{m}{i} \right) = -1 + \frac{1}{m} \sum_{i=1}^{m} \ln \frac{m}{i} = O \left(\frac{\ln m}{m} \right),$$

which together with (S.14) and (S.18) yields (S.12) for $\boldsymbol{w}_m = \boldsymbol{w}'_m$.

For $\boldsymbol{w}_m = \boldsymbol{w}''_m$, by (S.16) and Chebyshev's inequality,

$$R'_1(\boldsymbol{w}''_m) \stackrel{d}{=} \frac{-1}{m} \sum_{i=1}^{m} \frac{i}{m} \sum_{\ell=i}^{m} \left(Y_\ell - \frac{1}{\ell} \right) = \frac{-1}{2m^2} \sum_{i=1}^{m} i(i+1) \left(Y_i - \frac{1}{i} \right) = O_p(m^{-1/2}),$$

$$R''_1(\boldsymbol{w}''_m) = \left(\ln \frac{S_{m+1}}{m} - \ln \frac{S_{n+1}}{n} \right) \left(\frac{1}{m} \sum_{i=1}^{m} \frac{i}{m} \right) = O_p(m^{-1/2}),$$

$$R_2(\boldsymbol{w}''_m) = \frac{1}{m} \sum_{i=1}^{m} \frac{i}{m} \left\{ \sum_{\ell=i}^{m} -\frac{1}{\ell} + \ln \frac{m}{i} \right\} = -\frac{1}{4} + O \left(\frac{1}{m} \right) + \frac{1}{m} \sum_{i=1}^{m} \frac{i}{m} \ln \frac{m}{i} = O \left(\frac{\ln m}{m} \right),$$

where we have used the fact that $\frac{1}{m} \sum_{i=1}^{m} \frac{i}{m} \ln \frac{m}{i} = \frac{1}{4} + O \left(\frac{\ln m}{m} \right)$. Thus, by (S.14) and (S.15), $g(U_{1:n},\ldots,U_{m:n};\boldsymbol{w}''_m)=O_p(m^{-1/2})$. This establishes (S.12) for $\boldsymbol{w}_m=\boldsymbol{w}''_m$.

To prove (S.13), note that $\mathrm{P}(\Theta_{m:n}<\pi/2)\to 1$. For $\boldsymbol{w}_m=(w_1,\ldots,w_m)$ with $0\le w_i\le 1,\ i=1,\ldots,m$, on event $\{\Theta_{m:n}<\pi/2\}$, we have by Lemma A2

$$\begin{aligned}
g(F_p(\Theta_{1:n}),\ldots,F_p(\Theta_{m:n});\boldsymbol{w}_m) &\le g(\overline{F}_p(\Theta_{1:n}),\ldots,\overline{F}_p(\Theta_{m:n});\boldsymbol{w}_m)\\
&=\frac{1}{m}\sum_{i=1}^{m} w_i\left\{\ln\frac{n}{i}+\ln\overline{F}_p(\Theta_{i:n})\right\}\\
&\le\frac{1}{m}\sum_{i=1}^{m} w_i\left\{\ln\frac{n}{i}+\ln F_p(\Theta_{i:n})\right\}+\max_{1\le i\le m}\ln\frac{\overline{F}_p(\Theta_{i:n})}{F_p(\Theta_{i:n})}\\
&\le g(F_p(\Theta_{1:n}),\ldots,F_p(\Theta_{m:n});\boldsymbol{w}_m)+\ln\left(1+\frac{1}{(p-3)\cos^2\Theta_{m:n}}\right).
\end{aligned}\tag{S.19}$$

By Lemma A6(ii), $p\cos^2\Theta_{m:n}=2\ln\frac{n}{m}(1+o_p(1))$, implying that

$$\ln\left(1+\frac{1}{(p-3)\cos^2\Theta_{m:n}}\right)=O_p\left(\frac{1}{\ln\,(n/m)}\right),$$

which together with (S.19) establishes (S.13). So we have proved (S.11).

We now prove Lemma A7(i). We have by the definition of $\overline{F}_p(\theta)$ given in (A.2),

$$\begin{aligned}
&g(\overline{F}_p(\Theta_{1:n}),\ldots,\overline{F}_p(\Theta_{m:n});\boldsymbol{w}_m)=\frac{1}{m}\sum_{i=1}^{m} w_i\left\{\ln\frac{n}{i}+\ln\overline{F}_p(\Theta_{i:n})\right\}\\
&=\frac{1}{m}\sum_{i=1}^{m} w_i\left[\ln\frac{n}{i}+\ln c_p+\left\{\frac{p-1}{2}\ln\left(1-\cos^2\Theta_{i:n}\right)-\frac{1}{2}\ln\left(2\pi p\cos^2\Theta_{i:n}\right)\right\}\right],
\end{aligned}\tag{S.20}$$

where $c_p := \frac{\Gamma(p/2)}{\Gamma((p-1)/2)} \frac{\sqrt{2p}}{p-1} = 1 + O(p^{-1})$ by (A.15). By Lemma A6(ii)

$$\max_{1\le i\le m}\left|\cos^4\Theta_{i:n} - \frac{4(\ln(n/m))^2}{p^2}\right| \le \max_{1\le i\le m}\left|\cos^4\Theta_{i:n} - \frac{4(\ln(n/i))^2}{p^2}\right| + \frac{4}{p^2}\left[(\ln n)^2 - \left(\ln\frac{n}{m}\right)^2\right]$$
$$= O_p\left(\frac{\ln n \ln\ln n}{p^2}\right) + O_p\left(\frac{\ln n \ln m}{p^2}\right), \tag{S.21}$$

and $\max_{1\le i\le m}\cos^6\Theta_{i:n} = O_p\left(\frac{(\ln n)^3}{p^3}\right)$. (S.22)

By (S.22) and $\ln(1-x) = -x - x^2/2 + O(x^3)$,

$$\max_{1\le i\le m}\left|\ln(1-\cos^2\Theta_{i:n}) + \cos^2\Theta_{i:n} + \frac{1}{2}\cos^4\Theta_{i:n}\right| = O_p\left(\frac{(\ln n)^3}{p^3}\right). \tag{S.23}$$

Letting $R_{n,i} := p\cos^2\Theta_{i:n} - 2\ln\frac{n}{i} + \ln\ln\frac{n}{i}$, we have $\max_{1\le i\le m}|R_{n,i}| = O_p(1)$ by Lemma A6(ii), so that

$$\max_{1\le i\le m}\left|\ln(p\cos^2\Theta_{i:n}) - \ln\left(2\ln\frac{n}{i}\right)\right| = \max_{1\le i\le m}\left|\ln\left(1 - \frac{\ln\ln(n/i) - R_{n,i}}{2\ln(n/i)}\right)\right|$$
$$= O_p\left(\frac{\ln\ln(n/m)}{\ln(n/m)}\right). \tag{S.24}$$

For $0 \le w_i \le 1$, $i = 1, \dots, m$, let

$$A_n := \frac{1}{m}\sum_{i=1}^m w_i\left\{\frac{p-1}{2}\ln(1-\cos^2\Theta_{i:n}) - \frac{1}{2}\ln(2\pi p\cos^2\Theta_{i:n})\right\}, \tag{S.25}$$
$$B_n := \frac{1}{m}\sum_{i=1}^m w_i\left\{\frac{p}{2}\cos^2\Theta_{i:n} + \frac{1}{2}\ln(4\pi) + \frac{1}{2}\ln\ln\frac{n}{i} + \frac{(\ln(n/m))^2}{p}\right\}. \tag{S.26}$$

It follows from Lemma A6(ii) and (S.21)-(S.24) that

$$\begin{aligned}|A_n+B_n| &\le \max_{1\le i\le m}\left\{\frac{1}{2}\left|\ln(1-\cos^2\Theta_{i:n})\right|+\frac{p}{2}\left|\ln(1-\cos^2\Theta_{i:n})+\cos^2\Theta_{i:n}+\frac{1}{2}\cos^4\Theta_{i:n}\right|\right.\\ &\quad \left.+\frac{p}{4}\left|\cos^4\Theta_{i:n}-\frac{4(\ln(n/m))^2}{p^2}\right|+\frac{1}{2}\left|\ln(p\cos^2\Theta_{i:n})-\ln\left(2\ln\frac{n}{i}\right)\right|\right\}\\ &= O_p\left(\frac{\ln n}{p}\right)+O_p\left(\frac{(\ln n)^3}{p^2}\right)+O_p\left(\frac{\ln n\ln\ln n}{p}\right)\\ &\quad +O_p\left(\frac{\ln n\ln m}{p}\right)+O_p\left(\frac{\ln\ln(n/m)}{\ln(n/m)}\right).\end{aligned} \tag{S.27}$$

We have

$$A_n+B_n=O_p\left(\frac{(\ln\ n)^2}{p}\right)+O_p\left(\frac{\ln\ln(n/m)}{\ln(n/m)}\right)=O_p(1). \tag{S.28}$$

So, by (S.11), (S.20), (S.25), (S.26) and (S.28), we have for $\boldsymbol{w}_m=\boldsymbol{w}'_m,\ \boldsymbol{w}''_m$,

$$\begin{aligned}&\frac{1}{m}\sum_{i=1}^m w_i\left\{\ln\frac{n}{i}-\left(\frac{p}{2}\cos^2\Theta_{i:n}+\frac{1}{2}\ln(4\pi)+\frac{1}{2}\ln\ln\frac{n}{i}+\frac{(\ln(n/m))^2}{p}\right)\right\}\\ &=\frac{1}{m}\sum_{i=1}^m w_i\ln\frac{n}{i}-B_n\\ &=g\left(\overline{F}_p(\Theta_{1:n}),\ldots,\overline{F}_p(\Theta_{m:n});\boldsymbol{w}_m\right)-\frac{1}{m}\sum_{i=1}^m w_i\ln c_p-A_n-B_n\\ &=O_p(1).\end{aligned} \tag{S.29}$$

Noting that $\ln(1+x)\le x$ for $x>-1$, $\frac{1}{m}\sum_{i=1}^m\ln\frac{m}{i}=1+O\left(\frac{\ln m}{m}\right)$ and $\frac{1}{m}\sum_{i=1}^m\frac{i}{m}\ln\frac{m}{i}=\frac{1}{4}+O\left(\frac{\ln m}{m}\right)$, we have

$$\frac{1}{m}\sum_{i=1}^m\ln\frac{n}{i}=\ln\frac{n}{m}+\frac{1}{m}\sum_{i=1}^m\ln\frac{m}{i}=\ln\frac{n}{m}+1+O\left(\frac{\ln m}{m}\right), \tag{S.30}$$

$$0 \leq \frac{1}{m}\sum_{i=1}^{m} \ln\ln\frac{n}{i} - \ln\ln\frac{n}{m} = \frac{1}{m}\sum_{i=1}^{m} \ln\left\{\frac{\ln(m/i)}{\ln(n/m)} + 1\right\} \leq \frac{1}{m}\sum_{i=1}^{m}\frac{\ln(m/i)}{\ln(n/m)} = O\left(\frac{1}{\ln(n/m)}\right) = o(1), \tag{S.31}$$

$$\frac{1}{m}\sum_{i=1}^{m}\frac{i}{m}\ln\frac{n}{i} = \left(\frac{1}{m}\sum_{i=1}^{m}\frac{i}{m}\ln\frac{n}{m}\right) + \left(\frac{1}{m}\sum_{i=1}^{m}\frac{i}{m}\ln\frac{m}{i}\right) = \frac{m+1}{2m}\ln\frac{n}{m} + \frac{1}{4} + O\left(\frac{\ln m}{m}\right), \tag{S.32}$$

$$0 \leq \frac{1}{m}\sum_{i=1}^{m}\frac{i}{m}\left(\ln\ln\frac{n}{i} - \ln\ln\frac{n}{m}\right) = \frac{1}{m}\sum_{i=1}^{m}\frac{i}{m}\ln\left\{\frac{\ln(m/i)}{\ln(n/m)} + 1\right\} \leq \frac{1}{m}\sum_{i=1}^{m}\frac{i}{m}\frac{\ln(m/i)}{\ln(n/m)} = O\left(\frac{1}{\ln(n/m)}\right) = o(1). \tag{S.33}$$

By (S.29) with $\boldsymbol{w}_m = \boldsymbol{w}'_m$, (S.30) and (S.31) ,

$$\frac{p}{2}\left(\frac{1}{m}\sum_{i=1}^{m}\cos^2\Theta_{i:n}\right) = \ln\frac{n}{m} + 1 - \frac{1}{2}\ln(4\pi) - \frac{1}{2}\ln\ln\frac{n}{m} - \frac{1}{p}\left(\ln\frac{n}{m}\right)^2 + O_p(1). \tag{S.34}$$

By (S.29) with $\boldsymbol{w}_m = \boldsymbol{w}''_m$, (S.32) and (S.33) ,

$$\frac{p}{2}\left(\frac{1}{m}\sum_{i=1}^{m}\frac{i}{m}\cos^2\Theta_{i:n}\right) = \frac{m+1}{2m}\left\{\ln\frac{n}{m} + \frac{1}{2} - \frac{1}{2}\ln(4\pi) - \frac{1}{2}\ln\ln\frac{n}{m} - \frac{1}{p}\left(\ln\frac{n}{m}\right)^2\right\} + O_p(1). \tag{S.35}$$

On the event $\{\Theta_{m:n} < \pi/2\}$, we have $\cos\Theta_{1:n} \geq \cos\Theta_{2:n} \geq \cdots \geq$

$\cos\Theta_{m:n} > 0$, so that

$$\left(\frac{1}{m}\sum_{i=1}^{m}\cos\Theta_{i:n}\right)^2 \geq \frac{1}{m^2}\sum_{i=1}^{m}(2i-1)\cos^2\Theta_{i:n}. \tag{S.36}$$

By Schwarz' Inequality, $(m^{-1}\sum_{i=1}^{m}\cos\Theta_{i:n})^2 \leq m^{-1}\sum_{i=1}^{m}\cos^2\Theta_{i:n}$, which together with (S.34)-(S.36) yields

$$\begin{aligned}
0 &\leq \frac{p}{2}\left(\frac{1}{m}\sum_{i=1}^{m}\cos^2\Theta_{i:n}\right) - \frac{p}{2}\left(\frac{1}{m}\sum_{i=1}^{m}\cos\Theta_{i:n}\right)^2 \\
&\leq \frac{p}{2}\left(\frac{1}{m}\sum_{i=1}^{m}\cos^2\Theta_{i:n}\right) - \frac{p}{2}\left(\frac{1}{m^2}\sum_{i=1}^{m}(2i-1)\cos^2\Theta_{i:n}\right) + o_p(1) \\
&= \frac{p}{2}\left(\frac{m+1}{m}\right)\left(\frac{1}{m}\sum_{i=1}^{m}\cos^2\Theta_{i:n}\right) - \frac{p}{2}\left(\frac{2}{m^2}\sum_{i=1}^{m}i\cos^2\Theta_{i:n}\right) + o_p(1) \\
&= \frac{m+1}{2m} + O_p(1) = O_p(1).
\end{aligned} \tag{S.37}$$

It follows from (S.34) and (S.37) that

$$\frac{p}{2}A_{n,p,m} = \frac{p}{2}\left(\frac{1}{m}\sum_{i=1}^{m}\cos\Theta_{i:n}\right)^2 = \frac{p}{2}\left(\frac{1}{m}\sum_{i=1}^{m}\cos^2\Theta_{i:n}\right) + O_p(1) = \ln\frac{n}{m} - \frac{1}{2}\ln\ln\frac{n}{m} + O_p(1),$$

completing the proof of Lemma A7(i). □

Proof of Lemma A7(ii). Define for $u_1,\ldots,u_m > 0$,

$$f(u_1,\ldots,u_m) = \ln\frac{n}{m} + 1 - \left(\frac{1}{m}\sum_{i=1}^{m}\sqrt{-\ln\left(\min\{u_i,1\}\right)}\right)^2. \tag{S.38}$$

We claim that

$$\text{(a) } f(F_p(\Theta_{1:n}),\ldots,F_p(\Theta_{m:n})) \overset{d}{=} f\left(\frac{S_1}{S_{n+1}},\ldots,\frac{S_m}{S_{n+1}}\right)$$
$$= \frac{1}{m}\sum_{i=1}^{m}\ln\frac{S_i}{S_{m+1}} + 1 + \ln\frac{S_{m+1}}{m} + O_p\left(\frac{1}{\ln n}\right) + O_p\left(\frac{(\ln m)^3}{(\ln n)^2}\right)$$
$$= o_p(1),$$

$$\text{(b) } f(\overline{F}_p(\Theta_{1:n}),\ldots,\overline{F}_p(\Theta_{m:n})) = f(F_p(\Theta_{1:n}),\ldots,F_p(\Theta_{m:n})) + O_p\left(\frac{1}{\ln n}\right),$$

$$\text{(c) } 2f(\overline{F}_p(\Theta_{1:n}),\ldots,\overline{F}_p(\Theta_{m:n})) = -pA_{n,p,m} + 2\ln\frac{n}{m} - \ln\ln\frac{n}{m} - \ln(4\pi) + 2 - \frac{2}{p}\left(\ln\frac{n}{m}\right)^2$$
$$+ O_p\left(\frac{(\ln\ln\, n)^2}{\ln(n/m)}\right) + O_p\left(\frac{\ln m\ln\ln(n/m)}{\ln(n/m)}\right),$$

from which Lemma A7(ii) follows.

We first prove (a). By Lemma A3,

$$f\big(F_p(\Theta_{1:n}),\ldots,F_p(\Theta_{m:n})\big) \overset{d}{=} f(U_{1:n},\ldots,U_{m:n}) \overset{d}{=} f\left(\frac{S_1}{S_{n+1}},\ldots,\frac{S_m}{S_{n+1}}\right). \tag{S.39}$$

A simple application of the triangle inequality yields for $i = 1,\ldots,m$,

$$|\ln n - \ln S_i|^{1/2} - \left|\ln\frac{S_{n+1}}{n}\right|^{1/2} \le (\ln S_{n+1} - \ln S_i)^{1/2} \le |\ln n - \ln S_i|^{1/2} + \left|\ln\frac{S_{n+1}}{n}\right|^{1/2},$$

implying that

$$\max_{1\le i\le m}\left|(\ln S_{n+1} - \ln S_i)^{1/2} - |\ln n - \ln S_i|^{1/2}\right| \le \left|\ln\frac{S_{n+1}}{n}\right|^{1/2} = O_p(n^{-1/4}). \tag{S.40}$$

Since $m^{-1}\sum_{i=1}^{m}|\ln n - \ln S_i|^{1/2} \le \max\{|\ln n - \ln S_1|^{1/2}, |\ln n - \ln S_m|^{1/2}\} =$

$O_p((\ln n)^{1/2})$, we have by (S.40)

$$f\left(\frac{S_1}{S_{n+1}},\cdots,\frac{S_m}{S_{n+1}}\right)=\ln\frac{n}{m}+1-\left(\frac{1}{m}\sum_{i=1}^{m}|\ln n-\ln S_i|^{1/2}+O_p(n^{-1/4})\right)^2$$
$$=\ln\frac{n}{m}+1-\left(\frac{1}{m}\sum_{i=1}^{m}|\ln n-\ln S_i|^{1/2}\right)^2+O_p\left(\frac{(\ln n)^{1/2}}{n^{1/4}}\right)$$
$$=\ln\frac{n}{m}+1-\left(\frac{1}{m}\sum_{i=1}^{m}a_i\right)^2+O_p\left(\frac{(\ln n)^{1/2}}{n^{1/4}}\right), \tag{S.41}$$

where $a_i:=|\ln n-\ln S_i|^{1/2}$. Thus, we have by $\sqrt{1-x}=1-x/2+O(x^2)$,

$$\frac{1}{m}\sum_{i=1}^{m}\left(a_i-\frac{1}{m}\sum_{\ell=1}^{m}a_\ell\right)^2=\frac{\ln n}{m}\sum_{i=1}^{m}\left(\left|1-\frac{\ln S_i}{\ln n}\right|^{1/2}-\frac{1}{m}\sum_{\ell=1}^{m}\left|1-\frac{\ln S_\ell}{\ln n}\right|^{1/2}\right)^2$$
$$=\frac{\ln n}{m}\sum_{i=1}^{m}\left(\frac{\ln S_i}{2\ln n}-\frac{1}{m}\sum_{\ell=1}^{m}\frac{\ln S_\ell}{2\ln n}+O_p\left(\frac{(\ln m)^2}{(\ln n)^2}\right)\right)^2$$
$$=\frac{1}{4m\ln n}\sum_{i=1}^{m}\left(\ln S_i-\frac{1}{m}\sum_{\ell=1}^{m}\ln S_\ell\right)^2+O_p\left(\frac{(\ln m)^3}{(\ln n)^2}\right)$$
$$=O_p\left(\frac{1}{\ln n}\right)+O_p\left(\frac{(\ln m)^3}{(\ln n)^2}\right), \tag{S.42}$$

where the third equality follows from the fact that $\max_{1\le i\le m}|-\ln S_i+m^{-1}\sum_{\ell=1}^{m}\ln S_\ell|=O_p(\ln m)$ and the last equality is due to the fact that

$$\frac{1}{m}\sum_{i=1}^{m}\left(\ln S_i-\frac{1}{m}\sum_{\ell=1}^{m}\ln S_\ell\right)^2=\frac{1}{m}\sum_{i=1}^{m}\left(\ln\frac{S_i}{S_{m+1}}-\frac{1}{m}\sum_{\ell=1}^{m}\ln\frac{S_\ell}{S_{m+1}}\right)^2$$
$$\stackrel{d}{=}\frac{1}{m}\sum_{i=1}^{m}\left(\ln U_{i:m}-\frac{1}{m}\sum_{\ell=1}^{m}\ln U_{\ell:m}\right)^2$$
$$=\frac{1}{m}\sum_{i=1}^{m}\left(\ln U_i-\frac{1}{m}\sum_{\ell=1}^{m}\ln U_\ell\right)^2\to 1 \text{ a.s.}$$

Note also that

$$\frac{1}{m}\sum_{i=1}^{m}\ln\frac{S_i}{S_{m+1}} \overset{d}{=} \frac{1}{m}\sum_{i=1}^{m}\ln U_{i:m} = \frac{1}{m}\sum_{i=1}^{m}\ln U_i \to -1 \text{ a.s.} \tag{S.43}$$

By (S.41), (S.42), and (S.43),

$$\begin{aligned}
f\left(\frac{S_1}{S_{n+1}},\cdots,\frac{S_m}{S_{n+1}}\right) &= \ln\frac{n}{m} + 1 - \frac{1}{m}\sum_{i=1}^{m}a_i^2 + \frac{1}{m}\sum_{i=1}^{m}\left(a_i - \frac{1}{m}\sum_{\ell=1}^{m}a_\ell\right)^2 + O_p\left(\frac{(\ln n)^{1/2}}{n^{1/4}}\right)\\
&= \ln\frac{n}{m} + 1 - \frac{1}{m}\sum_{i=1}^{m}a_i^2 + O_p\left(\frac{1}{\ln n}\right) + O_p\left(\frac{(\ln m)^3}{(\ln n)^2}\right)\\
&= \frac{1}{m}\sum_{i=1}^{m}\ln S_i + 1 - \ln m + O_p\left(\frac{1}{\ln n}\right) + O_p\left(\frac{(\ln m)^3}{(\ln n)^2}\right)\\
&= \frac{1}{m}\sum_{i=1}^{m}\ln\frac{S_i}{S_{m+1}} + 1 + \ln\frac{S_{m+1}}{m} + O_p\left(\frac{1}{\ln n}\right) + O_p\left(\frac{(\ln m)^3}{(\ln n)^2}\right)\\
&= o_p(1),
\end{aligned}$$

proving (a).

Next we prove (b). We first show that if $x_i \geq y_i \geq 0$, $i = 1,\ldots,n$, then

$$\left(\frac{1}{n}\sum_{1\leq i\leq n}\sqrt{x_i}\right)^2 - \left(\frac{1}{n}\sum_{1\leq i\leq n}\sqrt{x_i - y_i}\right)^2 \leq 2\left(\max_{1\leq i\leq n}\sqrt{x_i}\right)\left(\max_{1\leq i\leq n}\frac{1}{\sqrt{x_i}}\right)\left(\max_{1\leq i\leq n}y_i\right). \tag{S.44}$$

Since $\sqrt{x_i - y_i} \geq \sqrt{x_i} - y_i/\sqrt{x_i} \geq 0$, we have

$$\begin{aligned}
\left(\frac{1}{n}\sum_{1\leq i\leq n}\sqrt{x_i - y_i}\right)^2 &\geq \left(\frac{1}{n}\sum_{1\leq i\leq n}\sqrt{x_i} - \frac{1}{n}\sum_{1\leq i\leq n}\frac{y_i}{\sqrt{x_i}}\right)^2\\
&\geq \left(\frac{1}{n}\sum_{1\leq i\leq n}\sqrt{x_i}\right)^2 - 2\left(\max_{1\leq i\leq n}\sqrt{x_i}\right)\left(\max_{1\leq i\leq n}\frac{1}{\sqrt{x_i}}\right)\left(\max_{1\leq i\leq n}y_i\right),
\end{aligned}$$

from which (S.44) follows.

By Lemma A6(ii)

$$\cos^2 \Theta_{m:n} = \frac{1}{p}\left(2\ln\frac{n}{m} - \ln\ln\frac{n}{m}\right) + O_p\left(\frac{1}{p}\right). \tag{S.45}$$

Let $\mathcal{E}_n$ denote the event that $\Theta_{m:n} < \pi/2$ and $\overline{F}_p(\Theta_{i:n}) < 1$ for $i = 1, \ldots, m$. Note that $\mathrm{P}(\Theta_{m:n} < \pi/2) \to 1$. On $\{\Theta_{m:n} < \pi/2\}$, we have by Lemma A2

$$\max_{1\le i\le m} \overline{F}_p(\Theta_{i:n}) = \overline{F}_p(\Theta_{m:n}) \le \left(1 + \{(p-3)\cos^2\Theta_{m:n}\}^{-1}\right) F_p(\Theta_{m:n}) \to 0 \text{ in probability}$$

(since $p\cos^2\Theta_{m:n} \to \infty$ in probability by (S.45) and $F_p(\Theta_{m:n}) \overset{d}{=} U_{m:n} \to 0$ in probability). So $\mathrm{P}(\mathcal{E}_n) \to 1$. The following calculations are carried out on the event $\mathcal{E}_n$, so that $0 < \overline{F}_p(\Theta_{i:n}) < 1$ and $\sqrt{-\ln\overline{F}_p(\Theta_{i:n})}$ is well defined for $i = 1, \ldots, m$. We have

$$\begin{aligned} f(\overline{F}_p(\Theta_{1:n}), \ldots, \overline{F}_p(\Theta_{m:n})) &= \ln\frac{n}{m} + 1 - \left(\frac{1}{m}\sum_{i=1}^m \sqrt{-\ln\overline{F}_p(\Theta_{i:n})}\right)^2 \\ &= \ln\frac{n}{m} + 1 - \left(\frac{1}{m}\sum_{i=1}^m \sqrt{-\ln F_p(\Theta_{i:n}) - \ln\frac{\overline{F}_p(\Theta_{i:n})}{F_p(\Theta_{i:n})}}\right)^2 \\ &= \ln\frac{n}{m} + 1 - \left(\frac{1}{m}\sum_{i=1}^m \sqrt{-\ln F_p(\Theta_{i:n})}\right)^2 + R_n \\ &= f(F_p(\Theta_{1:n}), \ldots, F_p(\Theta_{m:n})) + R_n, \end{aligned} \tag{S.46}$$

where R_n satisfies (on the event $\mathcal{E}_n$)

$$\begin{aligned}
0 < R_n &= \left(\frac{1}{m}\sum_{i=1}^{m}\sqrt{-\ln\, F_p(\Theta_{i:n})}\right)^2 - \left(\frac{1}{m}\sum_{i=1}^{m}\sqrt{-\ln F_p(\Theta_{i:n}) - \ln\frac{\overline{F}_p(\Theta_{i:n})}{F_p(\Theta_{i:n})}}\right)^2 \\
&\leq 2\left\{\max_{1\leq i\leq m}\sqrt{-\ln F_p(\Theta_{i:n})}\right\}\left\{\max_{1\leq i\leq m}\frac{1}{\sqrt{-\ln F_p(\Theta_{i:n})}}\right\}\left\{\max_{1\leq i\leq m}\ln\frac{\overline{F}_p(\Theta_{i:n})}{F_p(\Theta_{i:n})}\right\} \text{ (by (S.44))} \\
&\leq 2\left\{\sqrt{-\ln F_p(\Theta_{1:n})}\right\}\left\{\frac{1}{\sqrt{-\ln F_p(\Theta_{m:n})}}\right\}\ln\left(1+\frac{1}{(p-3)\cos^2\Theta_{m:n}}\right) \text{ (by Lemma A2)} \\
&= 2\sqrt{\ln n}(1+o_p(1))\frac{1}{\sqrt{\ln(n/m)}}(1+o_p(1))\left\{\frac{1}{(p-3)\cos^2\Theta_{m:n}}(1+o_p(1))\right\} \\
&= \frac{(\ln n)^{1/2}}{(\ln(n/m))^{3/2}}\frac{2\ln(n/m)}{(p-3)\cos^2\Theta_{m:n}}(1+o_p(1)) \\
&= \frac{1}{\ln n}(1+o_p(1)) \text{ (by (S.45) and } (\ln m)^3/(\ln n)^2 \to 0),
\end{aligned}$$

where the second equality is due to the fact that $-\ln F_p(\Theta_{i:n}) \overset{d}{=} -\ln U_{i:n} \overset{d}{=} \ln(S_{n+1}/S_i) = (1+o_p(1))\ln(n/i)$ for $i = 1$ and $i = m$. Thus, (b) follows from (S.46).

It remains to prove (c). Recall that by Sterling's formula, we have

$$\frac{\Gamma(p/2)}{\Gamma((p-1)/2)} = \sqrt{\frac{p}{2}}\left(1+O\left(\frac{1}{p}\right)\right) \text{ as } p\to\infty. \tag{S.47}$$

Again, the calculations below are done on the event $\mathcal{E}_n$. We have by (A.2) and (S.38)

$$f\left(\overline{F}_p(\Theta_{1:n}),\cdots\overline{F}_p(\Theta_{m:n})\right) = \ln\frac{n}{m}+1-\left(\frac{1}{m}\sum_{i=1}^{m}L_{i,n}^{1/2}\right)^2, \tag{S.48}$$

where for $i = 1, \ldots, m$,

$$\begin{aligned} L_{i,n} &:= -\ln \overline{F}_p(\Theta_{i:n}) \\ &= -\ln c_p - \frac{p-1}{2} \ln \left(1 - \cos^2 \Theta_{i:n}\right) + \frac{1}{2} \ln \left(2\pi p \cos^2 \Theta_{i:n}\right), \end{aligned} \tag{S.49}$$

where $c_p = \frac{\Gamma(p/2)}{\Gamma((p-1)/2)} \frac{\sqrt{2p}}{p-1} = 1 + O(p^{-1})$ (by (S.47)). Let

$$T_{i,n} := \frac{p}{2} \cos^2 \Theta_{i:n} \text{ and } t_n := \frac{1}{2} \ln \ln \frac{n}{m} + \frac{1}{2} \ln(4\pi) + \frac{1}{p} \left(\ln \frac{n}{m}\right)^2.$$

Then

$$\begin{aligned} L_{i,n} - (T_{i,n} + t_n) &= -\ln c_p + \frac{1}{2} \ln \left(1 - \cos^2 \Theta_{i:n}\right) \\ &- \frac{p}{2} \left(\ln (1 - \cos^2 \Theta_{i:n}) + \cos^2 \Theta_{i:n} + \frac{1}{2} \cos^4 \Theta_{i:n} \right) \\ &+ \frac{1}{2} \left(\ln (p \cos^2 \Theta_{i:n}) - \ln \left(2 \ln \frac{n}{i}\right) \right) \\ &+ \frac{1}{2} \left(\ln \left(2 \ln \frac{n}{i}\right) - \ln \left(2 \ln \frac{n}{m}\right) \right) \\ &+ \frac{1}{p} \left(\left(\frac{p}{2} \cos^2 \Theta_{i:n}\right)^2 - \left(\ln \frac{n}{i}\right)^2 \right) \\ &+ \frac{1}{p} \left(\left(\ln \frac{n}{i}\right)^2 - \left(\ln \frac{n}{m}\right)^2 \right). \end{aligned}$$

By Lemma A6(ii), $\frac{(\ln\ m)^3}{(\ln\ n)^2} = o(1)$ and $\frac{(\ln\ n)^2}{p} = O(1)$, we have

$$\max_{1\le i\le m}\left|\frac{T_{i,n}}{\ln(n/m)} - 1\right| = O_p\left(\frac{\ln\ m + \ln\ln\ n}{\ln(n/m)}\right) = o_p(1), \tag{S.50}$$

$$\max_{1\le i\le m}\left|\ln(1-\cos^2\ \Theta_{i:n})\right| = O_p\left(\frac{\ln\ n}{p}\right),$$

$$\max_{1\le i\le m}\frac{p}{2}\left|\ln(1-\cos^2\ \Theta_{i:n}) + \cos^2\ \Theta_{i:n} + \frac{1}{2}\cos^4\ \Theta_{i:n}\right| = O_p\left(\frac{(\ln\ n)^3}{p^2}\right) = O_p\left(\frac{1}{\ln\ n}\right),\quad \text{(by (S.23))}$$

$$\max_{1\le i\le m}\left|\ln(p\cos^2\ \Theta_{i:n}) - \ln\ \left(2\ln\ \frac{n}{i}\right)\right| = O_p\left(\frac{\ln\ln(n/m)}{\ln(n/m)}\right),$$

$$\max_{1\le i\le m}\left|\ln\ \left(2\ln\ \frac{n}{i}\right) - \ln\ \left(2\ln\ \frac{n}{m}\right)\right| = O\left(\frac{\ln\ m}{\ln\ n}\right),$$

$$\max_{1\le i\le m}\frac{1}{p}\left|\left(\frac{p}{2}\cos^2\ \Theta_{i:n}\right)^2 - \left(\ln\ \frac{n}{i}\right)^2\right| = O_p\left(\frac{\ln\ n\ln\ln\ n}{p}\right) = O_p\left(\frac{\ln\ln\ n}{\ln\ n}\right) = O_p\left(\frac{\ln\ln(n/m)}{\ln(n/m)}\right),$$

$$\max_{1\le i\le m}\frac{1}{p}\left|\left(\ln\ \frac{n}{i}\right)^2 - \left(\ln\ \frac{n}{m}\right)^2\right| = O\left(\frac{\ln\ n\ln\ m}{p}\right) = O\left(\frac{\ln\ m}{\ln\ n}\right), \tag{S.51}$$

which together imply that

$$\max_{1\le i\le m}|L_{i,n} - (T_{i,n} + t_n)| = O_p\left(\frac{\ln\ m}{\ln\ n}\right) + O_p\left(\frac{\ln\ln(n/m)}{\ln(n/m)}\right). \tag{S.52}$$

Thus, we have by (S.48)

$$\begin{aligned}
& f(\overline{F}_p(\Theta_{1:n}), \ldots, \overline{F}_p(\Theta_{m:n})) - \ln \frac{n}{m} - 1 \\
= & -\left(\frac{1}{m}\sum_{i=1}^m L_{i,n}^{1/2}\right)^2 \\
= & -\left(\frac{1}{m}\sum_{i=1}^m \{(T_{i,n}+t_n) + (L_{i,n} - (T_{i,n}+t_n))\}^{1/2}\right)^2 \\
= & -\left(\frac{1}{m}\sum_{i=1}^m T_{i,n}^{1/2}\left\{1 + T_{i,n}^{-1}t_n + O_p\left(\frac{\ln m}{(\ln(n/m))^2}\right) + O_p\left(\frac{\ln\ln(n/m)}{(\ln(n/m))^2}\right)\right\}^{1/2}\right)^2 \quad \text{(by (S.50) and (S.52))} \\
= & -\left(\frac{1}{m}\sum_{i=1}^m T_{i,n}^{1/2}\left\{1 + \frac{1}{2}T_{i,n}^{-1}t_n + O_p\left(\frac{\ln m}{(\ln(n/m))^2})\right) + O_p\left(\frac{\ln\ln(n/m)}{(\ln(n/m))^2}\right)\right\}\right)^2 \\
& \text{(by } \sqrt{1+x} = 1 + \tfrac{1}{2}x + O(x^2)) \\
= & -\left(\frac{1}{m}\sum_{i=1}^m T_{i,n}^{1/2} + \frac{t_n}{2m}\sum_{i=1}^m T_{i,n}^{-1/2} + O_p\left(\frac{\ln m}{(\ln(n/m))^{3/2}}\right) + O_p\left(\frac{\ln\ln(n/m)}{(\ln(n/m))^{3/2}}\right)\right)^2 \\
= & -\left(\frac{1}{m}\sum_{i=1}^m T_{i,n}^{1/2}\right)^2 - \frac{t_n}{m^2}\sum_{i=1}^m\sum_{\ell=1}^m T_{i,n}^{1/2}T_{\ell,n}^{-1/2} + O_p\left(\frac{\ln m}{\ln(n/m)}\right) + O_p\left(\frac{(\ln\ln(n/m))^2}{\ln(n/m)}\right) \\
= & -\frac{p}{2}A_{n,p,m} - t_n\left(1 + O_p\left(\frac{\ln\, m + \ln\ln\, n}{\ln(n/m)}\right)\right) + O_p\left(\frac{\ln m}{\ln\,(n/m)}\right) + O_p\left(\frac{(\ln\ln(n/m))^2}{\ln(n/m)}\right) \quad \text{(by (S.50))} \\
= & \frac{1}{2}\left\{-pA_{n,p,m} - \ln\ln\frac{n}{m} - \ln(4\pi) - \frac{2}{p}\left(\ln\frac{n}{m}\right)^2\right\} + O_p\left(\frac{(\ln\ln\, n)^2}{\ln(n/m)}\right) + O_p\left(\frac{\ln\, m\ln\ln(n/m)}{\ln(n/m)}\right)
\end{aligned}$$

establishing (c). The proof is complete. □

Proof of Lemma A7(iii). In view of (a)-(c) in the proof of Lemma A7(ii) together with $m(\ln\ln n)^4/(\ln n)^2 = o(1)$ and $(\ln n)^2/p = O(1)$, it sufficies to prove

$$m^{1/2}\left\{\frac{1}{m}\sum_{i=1}^m \ln\frac{S_i}{S_{m+1}} + 1 + \ln\frac{S_{m+1}}{m}\right\} \xrightarrow{d} N(0,2). \tag{S.53}$$

By Lemma A3, $\sum_{i=1}^m \frac{S_i}{S_{m+1}}$ is independent of S_{m+1}. Moreover,

$$m^{1/2}\ln\frac{S_{m+1}}{m} = m^{1/2}\left(\frac{S_{m+1}}{m} - 1\right) + o_p(1) \xrightarrow{d} N(0,1)$$

and

$$
\begin{aligned}
m^{-1/2}\sum_{i=1}^{m}\left(\ln\frac{S_i}{S_{m+1}}+1\right) &\overset{d}{=} m^{-1/2}\sum_{i=1}^{m}(\ln U_{i:m}+1)\\
&= m^{-1/2}\sum_{i=1}^{m}(\ln U_i+1) \xrightarrow{d} N(0,1),
\end{aligned}
$$

establishing (S.53). The proof is complete. □

S4 Proof of Lemma A8

Proof of Lemma A8. Since $\boldsymbol{W}_1,\ldots,\boldsymbol{W}_n$ are iid uniformly distributed on $\mathcal{S}^{p-1}$, we may write $\boldsymbol{W}_i=(W_{i1},\ldots,W_{ip})$ with $W_{ij}=\frac{Z_{ij}}{\sqrt{\sum_{k=1}^p Z_{ik}^2}}$ where Z_{ij}, $i=1,\ldots,n,\ j=1,\ldots,p$ are iid standard normal. We have

$$
\begin{aligned}
\sum_{1\le i\ne\ell\le n}\langle\boldsymbol{W}_i,\boldsymbol{W}_\ell\rangle &= \|\sum_{i=1}^{n}\boldsymbol{W}_i\|^2-n\\
&= \sum_{j=1}^{p}\left(\sum_{i=1}^{n}\frac{Z_{ij}}{\sqrt{\sum_{k=1}^p Z_{ik}^2}}\right)^2-n\\
&= \sum_{j=1}^{p}\left\{\frac{1}{\sqrt{p}}\sum_{i=1}^{n}Z_{ij}+\frac{1}{\sqrt{p}}\sum_{i=1}^{n}Z_{ij}\left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}}-1\right)\right\}^2-n\\
&= \frac{1}{p}\|\boldsymbol{A}_{n,p}+\boldsymbol{B}_{n,p}\|^2-n\\
&= \frac{1}{p}\left(\|\boldsymbol{A}_{n,p}\|^2-np+\|\boldsymbol{B}_{n,p}\|^2+2\langle\boldsymbol{A}_{n,p},\boldsymbol{B}_{n,p}\rangle\right), \qquad \text{(S.54)}
\end{aligned}
$$

where $\boldsymbol{A}_{n,p}$ and $\boldsymbol{B}_{n,p}$ are p-dimensional random vectors defined by

$$\boldsymbol{A}_{n,p} = \left(\sum_{i=1}^{n} Z_{ij}\right)_{j=1,\ldots,p},$$
$$\boldsymbol{B}_{n,p} = \left(\sum_{i=1}^{n} Z_{ij}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{ik}^2}} - 1\right)\right)_{j=1,\ldots,p}.$$

Then

$$\|\boldsymbol{A}_{n,p}\|^2 - np = \sum_{j=1}^{p}\left(\sum_{i=1}^{n} Z_{ij}\right)^2 - np \stackrel{d}{=} n\sum_{j=1}^{p}(Z_j^2 - 1),$$

where $Z_1, \ldots, Z_p$ are iid standard normal. So

$$\frac{1}{n\sqrt{2p}}\left(\|\boldsymbol{A}_{n,p}\|^2 - np\right) \stackrel{d}{=} \frac{1}{\sqrt{2p}}\sum_{j=1}^{p}(Z_j^2 - 1) \xrightarrow{d} N(0,1) \text{ as } p \to \infty. \quad \text{(S.55)}$$

Next, we show $\frac{1}{n\sqrt{2p}}\|\boldsymbol{B}_{n,p}\|^2 = o_p(1)$. We have

$$
\begin{aligned}
\mathrm{E}\|\boldsymbol{B}_{n,p}\|^2 &= \sum_{j=1}^{p}\sum_{i=1}^{n}\sum_{i'=1}^{n}\mathrm{E}\left[Z_{ij}\left(\sqrt{\frac{p}{\sum_{k=1}^{p}Z_{ik}^2}}-1\right)Z_{i'j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p}Z_{i'k}^2}}-1\right)\right]\\
&= \sum_{j=1}^{p}\sum_{i=1}^{n}\mathrm{E}\left[Z_{ij}^2\left(\sqrt{\frac{p}{\sum_{k=1}^{p}Z_{ik}^2}}-1\right)^2\right]\\
&= \sum_{i=1}^{n}\sum_{j=1}^{p}\mathrm{E}\left[Z_{ij}^2\frac{p}{\sum_{k=1}^{p}Z_{ik}^2}\right]+\sum_{i=1}^{n}\sum_{j=1}^{p}\mathrm{E}\left[Z_{ij}^2\right]-2\sum_{i=1}^{n}\sum_{j=1}^{p}\mathrm{E}\left[Z_{ij}^2\sqrt{\frac{p}{\sum_{k=1}^{p}Z_{ik}^2}}\right]\\
&= \sum_{i=1}^{n}\mathrm{E}\left[\sum_{j=1}^{p}Z_{ij}^2\frac{p}{\sum_{k=1}^{p}Z_{ik}^2}\right]+np-2\sum_{i=1}^{n}\mathrm{E}\left[\sum_{j=1}^{p}Z_{ij}^2\sqrt{\frac{p}{\sum_{k=1}^{p}Z_{ik}^2}}\right]\\
&= np+np-2\sqrt{p}\sum_{i=1}^{n}\mathrm{E}\sqrt{\sum_{j=1}^{p}Z_{ij}^2}\\
&= 2np-2\sqrt{p}\sum_{i=1}^{n}\mathrm{E}\sqrt{\mathcal{X}_p^2}\\
&= 2np-2\sqrt{p}n2^{1/2}\frac{\Gamma(\frac{p+1}{2})}{\Gamma(\frac{p}{2})}\\
&= 2np-2\sqrt{p}n2^{1/2}\sqrt{\frac{p}{2}}(1+O(p^{-1}))=O(n)\text{ as }n\wedge p\to\infty,
\end{aligned}
$$

where $\mathcal{X}_p^2$ denotes a chi-squared random variable with p degrees of freedom.

So

$$\frac{1}{n\sqrt{2p}}\mathrm{E}\|\boldsymbol{B}_{n,p}\|^2=O(p^{-1/2})\text{ as }n\wedge p\to\infty,$$

which by Chebyshev's inequality implies that

$$\frac{1}{n\sqrt{2p}}\|\boldsymbol{B}_{n,p}\|^2=o_p(1)\text{ as }n\wedge p\to\infty. \tag{S.56}$$

In view of (S.54)-(S.56), it remains to show

$$\frac{1}{n\sqrt{2p}}\langle \boldsymbol{A}_{n,p}, \boldsymbol{B}_{n,p}\rangle = o_p(1) \text{ as } n \wedge p \to \infty. \tag{S.57}$$

By Chebyshev's inequality, it suffices to establish

$$\mathrm{E}\langle \boldsymbol{A}_{n,p}, \boldsymbol{B}_{n,p}\rangle^2 = o(n^2 p). \tag{S.58}$$

We have

$$\begin{aligned}
&\mathrm{E}\langle \boldsymbol{A}_{n,p}, \boldsymbol{B}_{n,p}\rangle^2 = \mathrm{E}\left[\sum_{j=1}^{p}\left(\sum_{i=1}^{n} Z_{ij}\right)\left(\sum_{i'=1}^{n} Z_{i'j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{i'k}^2}}-1\right)\right)\right]^2 \\
&= \mathrm{E}\left\{\left[\sum_{j=1}^{p}\left(\sum_{i=1}^{n} Z_{ij}\right)\left(\sum_{i'=1}^{n} Z_{i'j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{i'k}^2}}-1\right)\right)\right]\left[\sum_{j'=1}^{p}\left(\sum_{\ell=1}^{n} Z_{\ell j'}\right)\left(\sum_{\ell'=1}^{n} Z_{\ell' j'}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{\ell'k}^2}}-1\right)\right)\right]\right\} \\
&= \sum_{j=1}^{p}\sum_{i=1}^{n}\sum_{i'=1}^{n}\sum_{j'=1}^{p}\sum_{\ell=1}^{n}\sum_{\ell'=1}^{n} \mathrm{E}\left\{Z_{ij}Z_{i'j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{i'k}^2}}-1\right)Z_{\ell j'}Z_{\ell' j'}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{\ell'k}^2}}-1\right)\right\} \\
&= C + D,
\end{aligned} \tag{S.59}$$

where

$$C = \sum_{j=1}^{p}\sum_{i=1}^{n}\sum_{i'=1}^{n}\sum_{\ell=1}^{n}\sum_{\ell'=1}^{n} \mathrm{E}\left\{Z_{ij}Z_{i'j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{i'k}^2}}-1\right)Z_{\ell j}Z_{\ell' j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{\ell'k}^2}}-1\right)\right\}$$

and

$$D = \sum_{1\le j\ne j'\le p}\sum_{i=1}^{n}\sum_{i'=1}^{n}\sum_{\ell=1}^{n}\sum_{\ell'=1}^{n} \mathrm{E}\left\{Z_{ij}Z_{i'j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{i'k}^2}}-1\right)Z_{\ell j'}Z_{\ell' j'}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{\ell'k}^2}}-1\right)\right\}.$$

Then

$$\begin{aligned}C &= \sum_{j=1}^{p}\sum_{i=i'=\ell=\ell'} \mathrm{E}\left\{Z_{ij}Z_{i'j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{i'k}^2}}-1\right)Z_{\ell j}Z_{\ell' j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{\ell' k}^2}}-1\right)\right\} \\ &+ \sum_{j=1}^{p}\sum_{i=i'\neq\ell=\ell'} \mathrm{E}\left\{Z_{ij}Z_{i'j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{i'k}^2}}-1\right)Z_{\ell j}Z_{\ell' j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{\ell' k}^2}}-1\right)\right\} \\ &+ \sum_{j=1}^{p}\sum_{i=\ell\neq i'=\ell'} \mathrm{E}\left\{Z_{ij}Z_{i'j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{i'k}^2}}-1\right)Z_{\ell j}Z_{\ell' j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{\ell' k}^2}}-1\right)\right\} \\ &+ \sum_{j=1}^{p}\sum_{i=\ell'\neq i'=\ell} \mathrm{E}\left\{Z_{ij}Z_{i'j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{i'k}^2}}-1\right)Z_{\ell j}Z_{\ell' j}\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{\ell' k}^2}}-1\right)\right\} \\ &:= C_1 + C_2 + C_3 + C_4. \qquad (S.60)\end{aligned}$$

We have

$$C_1 = np\mathrm{E}\left[Z_{ij}^4\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{ik}^2}}-1\right)^2\right] = O(np^{1/2}) \text{ as } n \wedge p \to \infty \qquad (S.61)$$

since by Schwarz' inequality

$$\begin{aligned}\left\{\mathrm{E}\left[Z_{ij}^4\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{ik}^2}}-1\right)^2\right]\right\}^2 &= \left\{\mathrm{E}\left[Z_{ij}^3 Z_{ij}^1\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{ik}^2}}-1\right)^2\right]\right\}^2 \\ &\leq \mathrm{E}Z_{ij}^6\mathrm{E}\left[Z_{ij}^2\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{ik}^2}}-1\right)^4\right] \\ &= 15\mathrm{E}\left[Z_{ij}^2\left(\sqrt{\frac{p}{\sum_{k=1}^{p} Z_{ik}^2}}-1\right)^4\right], \qquad (S.62)\end{aligned}$$

and

$$\begin{aligned}
& \mathrm{E}\left[Z_{ij}^2\left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}}-1\right)^4\right] \\
= & \frac{1}{p}\mathrm{E}\left[\sum_{j=1}^p Z_{ij}^2\left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}}-1\right)^4\right] \\
= & \frac{1}{p}\mathrm{E}\left[\mathcal{X}_p^2\left(\sqrt{\frac{p}{\mathcal{X}_p^2}}-1\right)^4\right], \\
= & p\mathrm{E}[(\mathcal{X}_p^2)^{-1}] - 4p^{1/2}\mathrm{E}[(\mathcal{X}_p^2)^{-1/2}] + 6 - 4p^{-1/2}\mathrm{E}[(\mathcal{X}_p^2)^{1/2}] + 1 \\
= & \frac{p}{p-2} - 4p^{1/2}\frac{1}{\sqrt{2}}\frac{\Gamma((p-1)/2)}{\Gamma(p/2)} + 6 - 4p^{-1/2}\sqrt{2}\frac{\Gamma((p+1)/2)}{\Gamma(p/2)} + 1 \\
= & O(p^{-1}). \qquad \text{(S.63)}
\end{aligned}$$

Next, we have

$$\begin{aligned}
C_2 = C_4 &= pn(n-1)\left\{\mathrm{E}\left[Z_{ij}^2\left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}}-1\right)\right]\right\}^2 \\
&= pn(n-1)\left\{\sqrt{p}\mathrm{E}\frac{Z_{ij}^2}{\sqrt{\sum_{k=1}^p Z_{ik}^2}}-1\right\}^2 \\
&= pn(n-1)\left\{\frac{1}{\sqrt{p}}\mathrm{E}\sqrt{\mathcal{X}_p^2}-1\right\}^2 \\
&= pn(n-1)\left\{\frac{1}{\sqrt{p}}\sqrt{2}\frac{\Gamma((p+1)/2)}{\Gamma(p/2)}-1\right\}^2 \\
&= pn(n-1)\left\{\frac{1}{\sqrt{p}}\sqrt{2}\sqrt{\frac{p}{2}}(1+O(p^{-1}))-1\right\}^2 \\
&= O\left(\frac{n^2}{p}\right) \text{ as } n \wedge p \to \infty, \qquad \text{(S.64)}
\end{aligned}$$

and

$$
\begin{aligned}
C_3 &= pn(n-1)\mathrm{E}Z_{i'j}^2\,\mathrm{E}\left[Z_{ij}^2\left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}}-1\right)^2\right] = pn(n-1)\mathrm{E}\left[Z_{ij}^2\left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}}-1\right)^2\right] \\
&= pn(n-1)\left\{\mathrm{E}p\,\frac{Z_{ij}^2}{\sum_{k=1}^p Z_{ik}^2}+\mathrm{E}Z_{ij}^2-2\sqrt{p}\,\mathrm{E}\frac{Z_{ij}^2}{\sqrt{\sum_{k=1}^p Z_{ik}^2}}\right\} \\
&= pn(n-1)\left\{1+1-\frac{2}{\sqrt{p}}\,\mathrm{E}\sqrt{\mathcal{X}_p^2}\right\} \\
&= pn(n-1)\left\{2-\frac{2}{\sqrt{p}}\,\sqrt{2}\frac{\Gamma((p+1)/2)}{\Gamma(p/2)}\right\} \\
&= pn(n-1)\left\{2-\frac{2}{\sqrt{p}}\,\sqrt{2}\sqrt{\frac{p}{2}}(1+O(p^{-1}))\right\} \\
&= pn(n-1)O(p^{-1}) = O\left(n^2\right) \text{ as } n\wedge p\to\infty. \qquad \text{(S.65)}
\end{aligned}
$$

By (S.60)-(S.65), we have

$$
C = C_1+C_2+C_3+C_4 = O(np^{1/2})+O\left(\frac{n^2}{p}\right)+O(n^2) = o(n^2p) \text{ as } n\wedge p\to\infty. \qquad \text{(S.66)}
$$

We now deal with D. For $j\neq j'$,

$$
\mathrm{E}\left[Z_{ij}Z_{i'j}\left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{i'k}^2}}-1\right)Z_{\ell j'}Z_{\ell' j'}\left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{\ell'k}^2}}-1\right)\right]=0 \quad \text{(S.67)}
$$

if $i\neq i'$ or $\ell\neq\ell'$. (Observe that (S.67) holds for the special case that $j\neq j'$ and $i=\ell'\neq\ell$.) So

$$
\begin{aligned}
D &= p(p-1)n\mathrm{E}\left[Z_{i1}^2\left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}}-1\right)Z_{i2}^2\left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}}-1\right)\right] \\
&+ p(p-1)n(n-1)\left\{\mathrm{E}\left[Z_{ij}^2\left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}}-1\right)\right]\right\}^2 \\
&:= D_1+D_2.
\end{aligned}
$$

We have

$$
\begin{aligned}
D_1 &= np(p-1)\mathrm{E}\left[Z_{i1}^2 Z_{i2}^2 \left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}} - 1\right)^2\right] \\
&= n\mathrm{E}\left[\sum_{1\le j\ne j'\le p} Z_{ij}^2 Z_{ij'}^2 \left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}} - 1\right)^2\right] \\
&= n\mathrm{E}\left[\sum_{j=1}^p\sum_{j'=1}^p Z_{ij}^2 Z_{ij'}^2 \left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}} - 1\right)^2 - \sum_{j=1}^p Z_{ij}^4 \left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}} - 1\right)^2\right] \\
&= n\mathrm{E}\left[\left(\sum_{j=1}^p Z_{ij}^2\right)^2 \left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}} - 1\right)^2\right] - np\mathrm{E}\left[Z_{ij}^4 \left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}} - 1\right)^2\right] \\
&= n\mathrm{E}\left[p\mathcal{X}_p^2 + (\mathcal{X}_p^2)^2 - 2\sqrt{p}(\mathcal{X}_p^2)^{3/2}\right] - npO(p^{-1/2}) \quad \text{(by (S.62) and (S.63))} \\
&= n\left\{p^2 + 3p + p(p-1) - 2\sqrt{p}2^{3/2}\frac{\Gamma((p+3)/2)}{\Gamma(p/2)}\right\} + O(np^{1/2}) \\
&= n\left\{2p^2 + 2p - 2\sqrt{p}2^{3/2}\frac{p+1}{2}\sqrt{\frac{p}{2}}(1+O(p^{-1}))\right\} + O(np^{1/2}) \\
&= O(np) + O(np^{1/2}) \\
&= O(np), \\
D_2 &= p(p-1)n(n-1)\left\{\mathrm{E}\left[Z_{i1}^2\left(\sqrt{\frac{p}{\sum_{k=1}^p Z_{ik}^2}} - 1\right)\right]\right\}^2 = (p-1)\,C_2 = O(n^2),
\end{aligned}
$$

so that $D = D_1 + D_2 = O(np) + O(n^2) = o(n^2 p)$ as $n \wedge p \to \infty$, which together with (S.59) and (S.66) establishes (S.58). The proof is complete.

□