Supplementary Material to "Alternative asymptotics for cointegration tests in large VARs."

Alexei Onatski and Chen Wang
Faculty of Economics, University of Cambridge

August 23, 2016

Abstract

This note contains supplementary material for Onatski and Wang (2016) (OW in what follows).

1 Supplementary Material for Section 2 of OW

1.1 Proof of Corollary OW3

Since the a.s. limits of $PB_{r,p,T}/(pT)$ and $PB_{0,p,T}/(pT)$ are the same as long as $r/p \to 0$, we shall only compute the latter limit. Under the simultaneous asymptotic regime, $p,T \to_c \infty$, we have

$$PB_{0,p,T}/(pT) \xrightarrow{a.s.} \int \lambda \mathrm{d}W(\lambda; c/(1+c), 2/(1+c)).$$

Using the explicit formula for the density of the Wachter distribution (OW9), we obtain,

$$\begin{aligned} \int \lambda \mathrm{d}W(\lambda; c/(1+c), 2/(1+c)) &= \frac{1+c}{2\pi c}\int_{b_-}^{b_+} \frac{\sqrt{(b_+-\lambda)(\lambda-b_-)}}{1-\lambda}\mathrm{d}\lambda \\ &\quad + \max\{0, 2-1/c\}, \end{aligned} \tag{1}$$

where

$$b_\pm = c\left(\sqrt{2} \mp \sqrt{1-c}\right)^{-2}.$$

Denote $\int \lambda \mathrm{d}W(\lambda; c/(1+c), 2/(1+c)) - \max\{0, 2-1/c\}$ as $\mathcal{I}$. Let

$$x = (\lambda - b_-)/(b_+ - b_-)$$

so that $\lambda = b_- + (b_+ - b_-)x$. Then

$$\mathcal{I} = \frac{1+c}{2c\pi}\int_0^1 \frac{(b_+ - b_-)^2\sqrt{(1-x)x}}{1-b_- - (b_+ - b_-)x}\mathrm{d}x.$$

Changing variables to θ where $x = (1-\cos\theta)/2$ so that $\mathrm{d}x = \frac{1}{2}\sin\theta\mathrm{d}\theta$, we obtain

$$\mathcal{I} = -\frac{1+c}{8c\pi}\int_0^\pi \frac{(b_+ - b_-)^2\sin^2\theta}{\frac{2-b_+-b_-}{2} + \frac{b_+-b_-}{2}\cos\theta}\mathrm{d}\theta.$$

Further, letting $z = \cos\theta + \mathrm{i}\sin\theta$ so that

$$\cos\theta = \frac{z+z^{-1}}{2},\ \sin\theta = \frac{z-z^{-1}}{2\mathrm{i}},\ \text{and } \mathrm{d}\theta = \frac{\mathrm{d}z}{\mathrm{i}z},$$

we obtain

$$\mathcal{I} = -\frac{1+c}{16c\pi\mathrm{i}}\oint_{|z|=1} \frac{(b_+ - b_-)^2\left(\frac{z-z^{-1}}{2}\right)^2}{\frac{2-b_+-b_-}{2} + \frac{b_+-b_-}{2}\frac{z+z^{-1}}{2}}\frac{\mathrm{d}z}{z},$$

where the contour integral is taken over the unit circle in the complex plane. Noting that

$$\frac{2-b_+-b_-}{2} + \frac{b_+-b_-}{2}\frac{z+z^{-1}}{2} = (a+bz)\left(a+bz^{-1}\right),$$

where

$$a = \frac{\sqrt{1-b_-}+\sqrt{1-b_+}}{2},\ b = \frac{\sqrt{1-b_-}-\sqrt{1-b_+}}{2},$$

we represent $\mathcal{I}$ in the following form

$$\mathcal{I} = -\frac{1+c}{64c\pi\mathrm{i}}\oint_{|z|=1} \frac{(b_+ - b_-)^2(z^2-1)^2}{a(a+bz)(z+b/a)}\frac{\mathrm{d}z}{z^2}.$$

Since $a > b > 0$, the integrand has poles at 0 and $-b/a$. The corresponding residues are

$$r_0 = \frac{1+c}{2c}\left(a^2+b^2\right),$$

and

$$r_{-b/a} = -\frac{1+c}{2c}\left(a^2-b^2\right),$$

so that

$$\mathcal{I} = \frac{1+c}{2c}\left(a^2+b^2\right) - \frac{1+c}{2c}\left(a^2-b^2\right).$$

Noting that

$$a^2+b^2 = \frac{-c+2c^2+1}{(c+1)^2}$$

and

$$a^2-b^2 = \frac{1-2c}{1+c},$$

we further simplify the above expression for $\mathcal{I}$ to obtain

$$\mathcal{I} = \frac{2c}{c+1}.$$

Therefore,

$$\int \lambda \mathrm{d}W\left(\lambda; c/\left(1+c\right), 2/\left(1+c\right)\right) = \frac{2c}{c+1} + \max\left\{0, 2-1/c\right\},$$

and

$$PB_{0,p,T}/\left(pT\right) \overset{a.s.}{\to} \frac{2c}{c+1} + \max\left\{0, 2-1/c\right\}.$$

1.2 Proof of Corollary OW4

As explained in OW, for $c < 1/2$, we have

$$\frac{1}{Tp} LR_{r,p,T} \geq -\frac{1}{p}\sum_{j=r+1}^{p} \overline{\log}(1-\lambda_j).$$

Therefore,

$$\frac{1}{Tp} LR_{r,p,T} \geq -\int \overline{\log}\left(1-\lambda\right) \mathrm{d}F_{p,T}\left(\lambda\right) + \frac{1}{p}\sum_{j=1}^{r} \overline{\log}(1-\lambda_j).$$

If $r/p \to 0$ as $p, T \to_c \infty$, the second term on the right hand side of the above display converges to zero. Therefore, almost surely,

$$\begin{aligned} \lim_{p,T\to_c\infty}\inf \frac{1}{pT} LR_{r,p,T} &\geq -\int \overline{\log}\left(1-\lambda\right) \mathrm{d}W\left(\lambda; c/\left(1+c\right), 2/\left(1+c\right)\right) \\ &= -\int \log\left(1-\lambda\right) \mathrm{d}W\left(\lambda; c/\left(1+c\right), 2/\left(1+c\right)\right). \end{aligned}$$

Denoting $-\int \log(1-\lambda)\,\mathrm{d}W(\lambda; c/(1+c), 2/(1+c))$ as $\mathcal{I}$, and using the explicit formula for the density of the Wachter distribution (OW9), we obtain

$$\mathcal{I} = -\frac{1+c}{2\pi c}\int_{b_-}^{b_+} \log(1-\lambda)\frac{\sqrt{(b_+-\lambda)(\lambda-b_-)}}{\lambda(1-\lambda)}\mathrm{d}\lambda,$$

where

$$b_\pm = c\left(\sqrt{2} \mp \sqrt{1-c}\right)^{-2}.$$

Let $x = (\lambda - b_-)/(b_+ - b_-)$ so that $\lambda = b_- + (b_+ - b_-)x$. Then

$$\mathcal{I} = -\frac{1+c}{2c\pi}\int_0^1 \frac{\log(1-b_- - (b_+-b_-)x)\sqrt{(1-x)x}\,(b_+-b_-)^2}{((b_+-b_-)x+b_-)(1-b_- - (b_+-b_-)x)}\mathrm{d}x.$$

Changing variables to θ where $x = (1-\cos\theta)/2$ so that $\mathrm{d}x = \frac{1}{2}\sin\theta\mathrm{d}\theta$, we obtain

$$\mathcal{I} = -\frac{1+c}{8c\pi}\int_0^\pi \frac{\log\left(\frac{2-b_+-b_-}{2} + \frac{b_+-b_-}{2}\cos\theta\right)(b_+-b_-)^2\sin^2\theta}{\left(\frac{b_++b_-}{2} - \frac{b_+-b_-}{2}\cos\theta\right)\left(\frac{2-b_+-b_-}{2} + \frac{b_+-b_-}{2}\cos\theta\right)}\mathrm{d}\theta.$$

Further, letting $z = \cos\theta + \mathrm{i}\sin\theta$ so that

$$\cos\theta = \frac{z+z^{-1}}{2},\ \sin\theta = \frac{z-z^{-1}}{2\mathrm{i}},\text{ and } \mathrm{d}\theta = \frac{\mathrm{d}z}{\mathrm{i}z},$$

we obtain

$$\mathcal{I} = \frac{1+c}{16c\pi\mathrm{i}}\oint_{|z|=1} \frac{\log\left(\frac{2-b_+-b_-}{2} + \frac{b_+-b_-}{2}\frac{z+z^{-1}}{2}\right)(b_+-b_-)^2\left(\frac{z-z^{-1}}{2}\right)^2}{\left(\frac{b_++b_-}{2} - \frac{b_+-b_-}{2}\frac{z+z^{-1}}{2}\right)\left(\frac{2-b_+-b_-}{2} + \frac{b_+-b_-}{2}\frac{z+z^{-1}}{2}\right)}\frac{\mathrm{d}z}{z},$$

where the contour integral is taken over the unit circle in the complex plane. Noting that

$$\frac{2-b_+-b_-}{2} + \frac{b_+-b_-}{2}\frac{z+z^{-1}}{2} = (a+bz)\left(a+bz^{-1}\right),$$

where

$$a = \frac{\sqrt{1-b_-}+\sqrt{1-b_+}}{2},\ b = \frac{\sqrt{1-b_-}-\sqrt{1-b_+}}{2},$$

and that

$$\frac{b_++b_-}{2} - \frac{b_+-b_-}{2}\frac{z+z^{-1}}{2} = (e-dz)\left(e-dz^{-1}\right),$$

where

$$e=\frac{\sqrt{b_{+}}+\sqrt{b_{-}}}{2}, d=\frac{\sqrt{b_{+}}-\sqrt{b_{-}}}{2},$$

we represent $\mathcal{I}$ in the following form

$$\begin{aligned}\mathcal{I} &= \frac{1+c}{4c\pi\mathrm{i}}\oint\limits_{|z|=1}\frac{\log\left(\left(a+bz\right)\left(a+bz^{-1}\right)\right)\left(ed\right)^{2}\left(z^{2}-1\right)^{2}}{\left(e-dz\right)\left(ez-d\right)\left(a+bz\right)\left(az+b\right)}\frac{\mathrm{d}z}{z}\\ &= \frac{1+c}{4c\pi\mathrm{i}}\oint\limits_{|z|=1}\frac{\log\left(\left(a+bz\right)\left(a+bz^{-1}\right)\right)ed^{2}\left(z^{2}-1\right)^{2}}{a\left(e-dz\right)\left(z-\frac{d}{e}\right)\left(a+bz\right)\left(z+\frac{b}{a}\right)}\frac{\mathrm{d}z}{z}.\end{aligned}$$

The integral has form $\mathcal{I}=\oint\limits_{|z|=1}\log\left(q(z)q\left(z^{-1}\right)\right)H\left(z\right)z^{-1}\mathrm{d}z$ with $H(z)=H\left(z^{-1}\right)$. Hence, expanding the logarithm yields two identical terms, so that

$$\mathcal{I}=\frac{1+c}{2c\pi\mathrm{i}}\oint\limits_{|z|=1}\frac{\log\left(a+bz\right)ed^{2}\left(z^{2}-1\right)^{2}}{a\left(e-dz\right)\left(z-\frac{d}{e}\right)\left(a+bz\right)\left(z+\frac{b}{a}\right)}\frac{\mathrm{d}z}{z}.$$

Since $a>b>0$ and $e>d>0$, $\log\left(a+bz\right)$ is analytic inside the unit circle and the integrand has three simple poles there: 0, $-b/a$, and d/e. The corresponding residues are

$$r_{0}=-\frac{1+c}{c}\frac{ed}{ab}\log a=-\frac{1+c}{c}\log a,$$

$$\begin{aligned}r_{-b/a} &= \frac{1+c}{c}\frac{\log\left(a-\frac{b^{2}}{a}\right)e^{2}d^{2}\left(a^{2}-b^{2}\right)}{ab\left(ae+db\right)\left(be+ad\right)}\\ &= \frac{1+c}{c}\sqrt{1-b_{-}}\sqrt{1-b_{+}}\log\left(a-\frac{b^{2}}{a}\right)\\ &= \frac{1-2c}{c}\log\left(a-\frac{b^{2}}{a}\right),\end{aligned}$$

and

$$\begin{aligned}r_{d/e} &= \frac{1+c}{c}\frac{\log\left(a+\frac{bd}{e}\right)ed\left(e^{2}-d^{2}\right)}{\left(ae+bd\right)\left(be+ad\right)}\\ &= \frac{1+c}{c}\sqrt{b_{+}}\sqrt{b_{-}}\log\left(a+\frac{bd}{e}\right)\\ &= \log\left(a+\frac{bd}{e}\right).\end{aligned}$$

Summing up, we obtain

$$\mathcal{I} = -\frac{1+c}{c}\log a + \frac{1-2c}{c}\log\left(a - \frac{b^2}{a}\right) + \log\left(a + \frac{bd}{e}\right).$$

Noting that

$$a = \frac{\sqrt{1-c}}{1+c},$$

$$a^2 - b^2 = \frac{1-2c}{1+c},$$

$$e = \frac{\sqrt{2c}}{1+c},$$

and

$$ae + bd = \frac{\sqrt{2c\,(1-c)}}{1+c}$$

we further simplify the above expression for $\mathcal{I}$ to obtain

$$\mathcal{I} = \frac{1+c}{c}\log(1+c) - \frac{1-c}{c}\log(1-c) + \frac{1-2c}{c}\log(1-2c).$$

2 Supplementary Material for Section 3 of OW

2.1 Proof of Theorem OW6

First, let us show that the weak limit $F_0(\lambda)$ of $F_\gamma(\lambda)$ as $\gamma \to 0$ exists and equals the continuous part of the Marchenko-Pastur distribution with density (OW21). By definition and Theorem OW1, $F_\gamma(\lambda)$ is the (scaled) Wachter d.f. $W(\gamma\lambda; \gamma/(1+\gamma), 2\gamma/(1+\gamma))$. Therefore, by (OW9) and (OW10), the density, $f_\gamma(\lambda)$, and the boundaries of the support, $\left[\hat{b}_-, \hat{b}_+\right]$, of the distribution F_γ equal

$$f_\gamma(\lambda) = \frac{1+\gamma}{2\pi}\frac{\sqrt{\left(\hat{b}_+ - \lambda\right)\left(\lambda - \hat{b}_-\right)}}{\lambda(1-\gamma\lambda)}, \text{ and}$$

$$\hat{b}_\pm = \left(\sqrt{2} \mp \sqrt{1-\gamma}\right)^{-2}.$$

As $\gamma \to 0$, $\hat{b}_\pm \to a_\pm$, where $a_\pm = \left(1 \pm \sqrt{2}\right)^2$ as in (OW20), and $f_\gamma(\lambda)$ converges to the density given by (OW21). This implies the weak convergence of $F_\gamma(\lambda)$ to $F_0(\lambda)$ with F_0 supported on $[a_-, a_+]$ and having density (OW21).

To establish the theorem, it remains to show that, as $p \to \infty$, $F_{p,\infty}(\lambda)$ weakly converges to $F_0(\lambda)$, in probability. Recall that the weak convergence is metrized by the Lévy distance $\mathcal{L}(\cdot,\cdot)$. We need to show that for any $\delta > 0$, there exists p_0 such that (s.t.) for all $p > p_0$,

$$\Pr\left(\mathcal{L}\left(F_0, F_{p,\infty}\right) < \delta\right) > 1 - \delta. \tag{2}$$

Let $\gamma > 0$ be so small that

$$\mathcal{L}\left(F_0, F_\gamma\right) < \delta/4. \tag{3}$$

For any p, let T_γ be the smallest even integer satisfying $p/T_\gamma \leq \gamma$. That is,

$$T_\gamma = \min_{T \in 2\mathbb{Z}} \left\{T : p/T \leq \gamma\right\}.$$

For any $T_\infty > T_\gamma$, by the triangle inequality, we have

$$\mathcal{L}\left(F_0, F_{p,\infty}\right) \leq \mathcal{L}\left(F_0, F_\gamma\right) + \mathcal{L}\left(F_\gamma, F_{p,T_\gamma}\right) + \mathcal{L}\left(F_{p,T_\gamma}, F_{p,T_\infty}\right) + \mathcal{L}\left(F_{p,T_\infty}, F_{p,\infty}\right), \tag{4}$$

where F_{p,T_γ} and F_{p,T_∞} denote the empirical distributions of eigenvalues of

$$\frac{T}{p} C D^{-1} C' A^{-1}, \tag{5}$$

with $T = T_\gamma$ and $T = T_\infty$, respectively.

By Theorem OW1, $\mathcal{L}\left(F_\gamma, F_{p,T_\gamma}\right)$ a.s. converges to zero as $p \to \infty$. Therefore, for all sufficiently large p, we have

$$\Pr\left(\mathcal{L}\left(F_\gamma, F_{p,T_\gamma}\right) < \delta/4\right) > 1 - \delta/4. \tag{6}$$

Further, as shown by Johansen (1988, 1991), for any p, as $T_\infty \to \infty$, the eigenvalues of (5) with $T = T_\infty$ jointly converge in distribution to those of

$$\frac{1}{p} \int_0^1 (\mathrm{d}B)\, B' \left(\int_0^1 BB' \mathrm{d}u\right)^{-1} \int_0^1 B\, (\mathrm{d}B)'. \tag{7}$$

Therefore, for any p and all sufficiently large T_∞, we have

$$\Pr\left(\mathcal{L}\left(F_{p,T_\infty}, F_{p,\infty}\right) < \delta/4\right) > 1 - \delta/4. \tag{8}$$

Let us denote the sum of $\mathcal{L}(F_0, F_\gamma)$, $\mathcal{L}\left(F_\gamma, F_{p,T_\gamma}\right)$, and $\mathcal{L}(F_{p,T_\infty}, F_{p,\infty})$ as $\mathcal{L}_{\gamma,p,T_\infty}$. By (4), we have

$$\mathcal{L}(F_0, F_{p,\infty}) \leq \mathcal{L}_{\gamma,p,T_\infty} + \mathcal{L}\left(F_{p,T_\gamma}, F_{p,T_\infty}\right). \tag{9}$$

Inequalities (3), (6), and (8) show that for any $\delta > 0$, there exists $\gamma_\delta > 0$ s.t. for any positive $\gamma < \gamma_\delta$, there is a p_γ s.t. for any $p > p_\gamma$, there is a T_p s.t. for any $T_\infty > T_p$

$$\Pr(\mathcal{L}_{\gamma,p,T_\infty} < 3\delta/4) > 1 - \delta/2. \tag{10}$$

The subscripts in γ_δ, p_γ and T_p signify dependence on the value of the corresponding parameter. Inequalities (10) and (9) would establish (2) as long as we are able to show that for any $\delta > 0$, there exists $\tilde{\gamma}_\delta > 0$ s.t. for any positive $\gamma < \tilde{\gamma}_\delta$, there is a $\tilde{p}_\gamma$ s.t. for any $p > \tilde{p}_\gamma$ *and any* $\tilde{T}_p$, there exists $T_\infty > \tilde{T}_p$ s.t.

$$\Pr\left(\mathcal{L}\left(F_{p,T_\gamma}, F_{p,T_\infty}\right) < \delta/4\right) > 1 - \delta/2. \tag{11}$$

Let us denote $\xi = \sqrt{T}\varepsilon$, where ε is a $p \times T$ matrix with i.i.d. $N(0, 1/T)$ entries, as defined in Section OW2. We shall assume that, as p, T change, ξ represents $p \times T$ sections of a fixed infinite array of i.i.d. standard normal random variables. Consider

$$M_{p,T} = \frac{T}{p}\left(\frac{\xi\xi'}{T}\right)^{-1/2} \frac{\xi\Delta_2'\xi'}{T}\left(\frac{\xi\Delta_1\xi'}{T}\right)^{-1} \frac{\xi\Delta_2\xi'}{T}\left(\frac{\xi\xi'}{T}\right)^{-1/2}.$$

So defined matrix $M_{p,T}$ is identical to the real symmetric matrix $\frac{T}{p}A^{-1/2}CD^{-1}C'A^{-1/2}$. The above definition is formulated in terms of ξ to clarify that $M_{p,T}$ depends on T not only via the term T/p, but also through A, C, and D. Note that F_{p,T_γ} and F_{p,T_∞} are the empirical distributions of eigenvalues of M_{p,T_γ} and M_{p,T_∞}, respectively. The following lemma is established in the next section of this note.

Lemma 1 *For any $\tau > 0$ there exists $\gamma_\tau > 0$ s.t. for any positive $\gamma < \gamma_\tau$, there is a $\tilde{p}_\gamma$ s.t. for any $p > \tilde{p}_\gamma$ and any $\tilde{T}_p$, there exists $T_\infty > \tilde{T}_p$ s.t. with probability larger than $1 - \tau$, $M_{p,T_\gamma} - M_{p,T_\infty}$ can be represented as the sum of two real symmetric matrices S and R,*

$$M_{p,T_\gamma} - M_{p,T_\infty} = S + R,$$

where $\|S\| \leq K\sqrt{\gamma}$, $\operatorname{rank} R \leq \tau p$, *and K is an absolute constant.*

Finally, let F_{SR} be the empirical distribution of eigenvalues of $M_{p,T_\gamma} - S =$

$M_{p,T_\infty}+R$. Then, by Theorem A45 (norm inequality) of Bai and Silverstein (2010),

$$\mathcal{L}\left(F_{p,T_\gamma},F_{SR}\right)\leq\|S\|\leq K\sqrt{\gamma},$$

whereas by their Theorem A43 (rank inequality),

$$\mathcal{L}\left(F_{SR},F_{p,T_\infty}\right)\leq\frac{1}{p}\operatorname{rank}R\leq\tau.$$

Therefore, by Lemma 1 and the triangle inequality, for any $\tau>0$ there exists $\gamma_\tau>0$ s.t. for any positive $\gamma<\gamma_\tau$, there is a $\tilde{p}_\gamma$ s.t. for any $p>\tilde{p}_\gamma$ and any $\tilde{T}_p$, there exists $T_\infty>\tilde{T}_p$ s.t.

$$\Pr\left(\mathcal{L}\left(F_{p,T_\gamma},F_{p,T_\infty}\right)<\tau+K\sqrt{\gamma}\right)>1-\tau.$$

For $\tau=\delta/8$, this inequality implies (11) with $\tilde{\gamma}_\delta=\min\left\{\gamma_\tau,(\delta/8K)^2\right\}$. Combining (11) with (10) yields (2), which completes the proof.

2.2 Proof of Lemma 1

Consider some positive $\gamma<1$. For any p, let

$$T_\gamma=\min_{T\in2\mathbb{Z}}\left\{T:p/T\leq\gamma\right\},$$

and let $T=T_\infty>T_\gamma$. Consider a partition $\xi=\left[\xi_\gamma,\xi_\infty\right]$, where ξ_γ and ξ_∞ are $p\times T_\gamma$ and $p\times(T_\infty-T_\gamma)$, respectively. Further, let $\Delta_{1\gamma}$ and $\Delta_{2\gamma}$ be defined similarly to Δ_1 and Δ_2 with T replaced by T_γ. Define

$$A_\gamma=\frac{\xi_\gamma\xi_\gamma'}{T_\gamma},D_\gamma=\frac{\xi_\gamma\Delta_{1\gamma}\xi_\gamma'}{T_\gamma},\text{ and }C_\gamma=\frac{\xi_\gamma\Delta_{2\gamma}'\xi_\gamma'}{T_\gamma}.$$

Then

$$\begin{aligned}M_{p,T_\gamma}&=\frac{T_\gamma}{p}A_\gamma^{-1/2}C_\gamma D_\gamma^{-1}C_\gamma'A_\gamma^{-1/2},\text{ and}\\M_{p,T_\infty}&=\frac{T}{p}A^{-1/2}CD^{-1}C'A^{-1/2}.\end{aligned}$$

Geman (1980) and Silverstein (1985) established the a.s. convergence (as $p\to\infty$) of the largest and the smallest eigenvalues (which we shall denote as $\lambda_{\max}$ and

$\lambda_{\min}$) of A_γ to $\left(1+\sqrt{\gamma}\right)^2$ and $\left(1-\sqrt{\gamma}\right)^2$, respectively. Hence,

$$\lambda_{\max}\left(A_\gamma^{-1/2}\right) \overset{a.s}{\to} \left(1-\sqrt{\gamma}\right)^{-1}, \text{ and}$$
$$\lambda_{\min}\left(A_\gamma^{-1/2}\right) \overset{a.s.}{\to} \left(1+\sqrt{\gamma}\right)^{-1}.$$

For $\gamma < 1/4$,

$$\left(1-\sqrt{\gamma}\right)^{-1} - 1 < 2\sqrt{\gamma} \text{ and } 1 - \left(1+\sqrt{\gamma}\right)^{-1} < \sqrt{\gamma}.$$

Therefore, for such γ, with probability arbitrarily close to one for all sufficiently large p,

$$\left\| A_\gamma^{-1/2} - I_p \right\| \leq 3\sqrt{\gamma}. \tag{12}$$

Moreover, the inequality

$$\left\| A^{-1/2} - I_p \right\| \leq 3\sqrt{\gamma} \tag{13}$$

holds with probability arbitrarily close to one for all sufficiently large p as well. To see this, consider a $p_\gamma \times T_\infty$ matrix η with $p_\gamma = \lfloor \gamma T_\infty \rfloor \geq p$ (here $\lfloor \cdot \rfloor$ denotes the integer part of a real number), such that the upper $p \times T_\infty$ block of η coincides with ξ, and the remaining part of η consists of i.i.d. $N(0,1)$ variables independent from ξ. Note that A can be viewed as a $p \times p$ principal submatrix of $A_{\gamma\infty} \equiv \eta\eta'/T_\infty$. By Theorem 4.3.15 of Horn and Johnson (1985),

$$\lambda_{\min}\left(A_{\gamma\infty}\right) \leq \lambda_{\min}\left(A\right) \leq \lambda_{\max}\left(A\right) \leq \lambda_{\max}\left(A_{\gamma\infty}\right). \tag{14}$$

Since $A_{\gamma\infty}$ is the sample covariance matrix with $p_\gamma = \lfloor \gamma T_\infty \rfloor$, its largest and smallest eigenvalues a.s. converge to the same limits as those of A_γ, and thus, (14) yields (13).

Inequalities (12) and (13) imply that it is sufficient to establish Lemma 1 with M_{p,T_γ} and M_{p,T_∞} replaced by

$$\tilde{M}_{p,T_\gamma} = C_\gamma \left(\frac{p}{T_\gamma} D_\gamma\right)^{-1} C_\gamma', \text{ and } \tilde{M}_{p,T_\infty} = C\left(\frac{p}{T} D\right)^{-1} C'.$$

Let

$$\alpha_D = \left(\frac{p}{T_\gamma} D_\gamma\right)^{1/2} \left(\frac{p}{T} D\right)^{-1} \left(\frac{p}{T_\gamma} D_\gamma\right)^{1/2} - I_p$$

and

$$\alpha_C = C - C_\gamma.$$

The following lemma is proven in the next section of this note.

Lemma 2 *For any $\tau > 0$ there exists $\gamma_\tau > 0$ s.t. for any positive $\gamma < \gamma_\tau$, there is a $\tilde{p}_\gamma$ s.t. for any $p > \tilde{p}_\gamma$ and any $\tilde{T}_p$, there exists $T_\infty > \tilde{T}_p$ s.t. with probability larger than $1 - \tau$,*

$$\|\alpha_D - \bar{\alpha}_D\| \leq K\gamma \text{ and } \|\alpha_C\| \leq K\sqrt{\gamma},$$

where $\bar{\alpha}_D$ is a matrix with rank no larger than τp, and K is an absolute constant.

Using the identity

$$C\left(\frac{p}{T}D\right)^{-1}C' = (C_\gamma + \alpha_C)\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2}(I_p + \alpha_D)\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2}\left(C'_\gamma + \alpha'_C\right),$$

it is straightforward to verify that

$$\tilde{M}_{p,T_\infty} - \tilde{M}_{p,T_\gamma} = (\beta_1 + \beta'_1) + \beta_2 + (\beta_3 + \beta'_3) + \beta_4, \tag{15}$$

where

$$\beta_1 = \alpha_C\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1}C'_\gamma,$$

$$\beta_2 = C_\gamma\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2}\alpha_D\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2}C'_\gamma,$$

$$\beta_3 = \alpha_C\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2}\alpha_D\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2}C'_\gamma,$$

and

$$\beta_4 = \alpha_C\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2}(I_p + \alpha_D)\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2}\alpha'_C.$$

To bound the norms of the β's, in addition to Lemma 2, we need the following lemma, which we also establish in the next section of this note.

Lemma 3 *As $p \to \infty$, the empirical distribution of eigenvalues of D_γ a.s. converges to a nonrandom distribution with support bounded below by $(17\gamma)^{-1}$ for sufficiently small γ.*

Consider the following decomposition of β_1 in the product of four terms, given in the square brackets,

$$\beta_1 = [\alpha_C] \left[\left(\frac{p}{T_\gamma} D_\gamma \right)^{-1/2} \right] \left[\left(\frac{p}{T_\gamma} D_\gamma \right)^{-1/2} C'_\gamma A_\gamma^{-1/2} \right] \left[A_\gamma^{1/2} \right].$$

By Lemma 2, with high probability,

$$\|\alpha_C\| \leq K\sqrt{\gamma}. \tag{16}$$

Further, since $\lambda_{\max}(A_\gamma) \overset{a.s.}{\to} \left(1+\sqrt{\gamma}\right)^2$, we have, with high probability

$$\|A_\gamma\| \leq K, \tag{17}$$

where K may refer to different constants in different formulae.

Next, by Lemma 3, there exists a decomposition

$$\left(\frac{p}{T_\gamma} D_\gamma \right)^{-1/2} = x_1 + x_2 \tag{18}$$

where $\|x_1\|$ remains bounded in probability by an absolute constant and x_2 has rank no larger than τp for arbitrarily small τ and sufficiently large p. The term x_2 is needed because Lemma 3 does not establish the a.s. convergence of the smallest eigenvalue of D_γ. It only concerns with the a.s. convergence of the empirical distribution of the eigenvalues of D_γ.

Similarly, we can show that

$$\left(\frac{p}{T_\gamma} D_\gamma \right)^{-1/2} C'_\gamma A_\gamma^{-1/2} = x_3 + x_4, \tag{19}$$

where $\|x_3\|$ remains bounded in probability by an absolute constant and x_4 has rank no larger than τp for arbitrarily small τ and sufficiently large p. Indeed, the squared singular values of $\left(\frac{p}{T_\gamma} D_\gamma \right)^{-1/2} C'_\gamma A_\gamma^{-1/2}$ equal the eigenvalues of

$$C_\gamma \left(\frac{p}{T_\gamma} D_\gamma \right)^{-1} C'_\gamma \left(A_\gamma\right)^{-1}. \tag{20}$$

By Theorem OW1, the empirical distribution of these eigenvalues a.s. converges to a distribution with the upper boundary of support $\left(\sqrt{2} - \sqrt{1-\gamma}\right)^{-2}$. Therefore, for any $\tau > 0$, with high probability for sufficiently large p, there are no more than

τp singular values of $\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2} C'_\gamma A_\gamma^{-1/2}$ that are larger than $\left(\sqrt{2}-\sqrt{1-\gamma}\right)^{-1}$, which establishes the existence of the decomposition (19).

Combining (16), (17), (18), and (19), we conclude that with high probability, for all sufficiently large p, β_1 can be decomposed in the sum of a matrix of norm no larger than $K\sqrt{\gamma}$ and a matrix of rank no larger than τp. Similar decompositions are valid for $\beta_2, ..., \beta_4$. The corresponding proofs are very similar to the proof for β_1 and we omit them. Using the decompositions for $\beta_1, ..., \beta_4$ in (15) completes the proof.

2.3 Proofs of Lemmas 2 and 3

2.3.1 Proof of Lemma 2

Let us, first, focus on α_D. Define

$$\alpha \equiv (I_p + \alpha_D)^{-1} - I_p.$$

It is sufficient to prove that, with high probability, for sufficiently large p, α can be decomposed into the sum of a term with norm no larger than $K\sqrt{\gamma}$ and a term with rank no larger than τp. By definition of α_D, we have

$$\alpha = \left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2}\left(\frac{p}{T}D - \frac{p}{T_\gamma}D_\gamma\right)\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2}.$$

Recall that Δ_1 and Δ_2 are $T = T_\infty$-dimensional matrices with $T_\infty > T_\gamma$. Consider partitions $\Delta_1 = \operatorname{diag}\{\Delta_{11}, \Delta_{12}\}$ and $\Delta_2 = \operatorname{diag}\{\Delta_{21}, \Delta_{22}\}$, where Δ_{11} and Δ_{21} are $T_\gamma \times T_\gamma$. Using this notation, represent α in the form

$$\alpha = \left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2}\left(p\frac{\xi_\gamma\Delta_{11}\xi'_\gamma}{T_\infty^2} - \frac{p}{T_\gamma}D_\gamma + p\frac{\xi_\infty\Delta_{12}\xi'_\infty}{T_\infty^2}\right)\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2} \tag{21}$$

By Lemma 3, the proportion of the eigenvalues of $\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2}$ that are larger than 5 converges to zero in probability as $p \to \infty$. In particular, with probability arbitrary close to one, this proportion is smaller than any fixed small positive τ for sufficiently large p. This implies that, with high probability, $\left(\frac{p}{T_\gamma}D_\gamma\right)^{-1/2}$ can be represented as a matrix of rank no larger than τp plus a matrix of norm no larger than 5.

Further, consider

$$p\frac{\xi_\gamma \Delta_{11} \xi_\gamma'}{T_\infty^2} - \frac{p}{T_\gamma} D_\gamma = p\xi_\gamma \left(\frac{\Delta_{11}}{T_\infty^2} - \frac{\Delta_{1\gamma}}{T_\gamma^2} \right) \xi_\gamma'.$$

Recall that the diagonal elements of Δ_{11} (except the first two) have form $\frac{1}{2}\left(1-\cos 2\pi j/T_\infty\right)^{-1}$ with $j \le T_\gamma/2$. The diagonal elements of $\Delta_{1\gamma}$ have a similar form with T_∞ replaced by T_γ. Since

$$\cos x = 1 - \frac{1}{2}x^2 + \frac{1}{4!}x^4 \cos t$$

for some $t \in [0, x]$, we have

$$\frac{1}{2T_\infty^2}\left(1-\cos 2\pi j/T_\infty\right)^{-1} = \frac{1}{(2\pi j)^2}\left(1 - \frac{\cos t}{12}\frac{(2\pi j)^2}{T_\infty^2}\right)^{-1}$$

for some $t \in [0, \pi]$, and hence

$$\frac{1}{2T_\infty^2}\left(1-\cos 2\pi j/T_\infty\right)^{-1} - \frac{1}{(2\pi j)^2} = \frac{\cos t}{12T_\infty^2}\left(1 - \frac{\cos t}{12}\frac{(2\pi j)^2}{T_\infty^2}\right)^{-1}.$$

Since $j \le T_\gamma/2$ and $T_\infty > T_\gamma$, we have

$$1 - \frac{\cos t}{12}\frac{(2\pi j)^2}{T_\infty^2} > 1 - \frac{\pi^2}{12} > \frac{1}{12},$$

and thus

$$\left| \frac{1}{2T_\infty^2}\left(1-\cos 2\pi j/T\right)^{-1} - \frac{1}{(2\pi j)^2} \right| < \frac{1}{T_\infty^2}.$$

A similar inequality holds for the elements of $\Delta_{1\gamma}$:

$$\left| \frac{1}{2T_\gamma^2}\left(1-\cos 2\pi j/T_\gamma\right)^{-1} - \frac{1}{(2\pi j)^2} \right| < \frac{1}{T_\gamma^2}.$$

Therefore,

$$\left| \frac{1}{2T_\infty^2}\left(1-\cos 2\pi j/T_\infty\right)^{-1} - \frac{1}{2T_\gamma^2}\left(1-\cos 2\pi j/T_0\right)^{-1} \right| < \frac{2}{T_\gamma^2}.$$

The absolute value of the difference of the first two diagonal elements of Δ_{11}/T_∞^2

and $\Delta_{1\gamma}/T_\gamma^2$ is obviously smaller than $2/T_\gamma^2$ too. To summarize,

$$\left\| p\frac{\xi_\gamma \Delta_{11} \xi_\gamma'}{T_\infty^2} - \frac{p}{T_\gamma} D_\gamma \right\| < \left\| 2\frac{p}{T_\gamma} \frac{\xi_\gamma \xi_\gamma'}{T_\gamma} \right\| < 4\gamma \tag{22}$$

with high probability for sufficiently small γ. To obtain the last inequality we used the fact that the largest eigenvalue of $\xi_\gamma \xi_\gamma'/T_\gamma$ a.s. converges to $\left(1+\sqrt{\gamma}\right)^2$.

Consider now the component $p\frac{\xi_\infty \Delta_{12} \xi_\infty'}{T_\infty^2}$ of (21). Since, as is straightforward to verify, $1-\cos x > x^2/6$ for $x \in [0,\pi]$, we have

$$2T_\infty^2 \left(1-\cos 2\pi j/T_\infty\right) > \left(2\pi j\right)^2/3. \tag{23}$$

Let us represent Δ_{12} as diag $\left\{\Delta_{12,1}, ..., \Delta_{12,(T_\infty - T_\gamma)/T_\gamma}\right\}$, where each block $\Delta_{12,i}$ is T_γ-dimensional. We can choose T_∞ so that $\left(T_\infty - T_\gamma\right)/T_\gamma$ is an integer, so such a representation is possible. Using the fact that the diagonal elements of $\Delta_{12,i}/T_\infty^2$ have form

$$\frac{1}{2T_\infty^2\left(1-\cos 2\pi j/T_\infty\right)} \text{ with } j = iT_\gamma/2+1, ..., \left(i+1\right)T_\gamma/2-1,$$

we find that the upper bound on the diagonal elements of $\Delta_{12,i}/T_\infty^2$ equals

$$\left[2T_\infty^2\left(1-\cos iT_\gamma \pi/T_\infty\right)\right]^{-1}.$$

By (23), this is no larger than $3/\left(i\pi T_\gamma\right)^2$.

Let us decompose ξ_∞ conformably with Δ_{12} so that $\xi_\infty = \left[\xi_{\infty,1}, ..., \xi_{\infty,(T_\infty - T_\gamma)/T_\gamma}\right]$. Then, from the above, we have

$$\left\| p\frac{\xi_\infty \Delta_{12} \xi_\infty'}{T_\infty^2} \right\| \le \frac{3p}{\pi^2 T_\gamma} \sum_{i=1}^{(T_\infty - T_\gamma)/T_\gamma} \frac{1}{i^2} \left\| \frac{\xi_{\infty,i} \xi_{\infty,i}'}{T_\gamma} \right\|.$$

The Gaussian concentration inequality for the singular values of a rectangular matrix with i.i.d. Gaussian entries (see Theorem II.13 of Davidson and Szarek (2001)) implies that, for any $t>0$,

$$\Pr\left(\left\| \frac{\xi_{\infty,i} \xi_{\infty,i}'}{T_\gamma} \right\| \ge \left(1+\sqrt{\frac{p}{T_\gamma}}+t\right)^2\right) < \exp\left\{-\frac{T_\gamma t^2}{2}\right\}.$$

Take $t = i^{1/4}$. Then,

$$\sum_{i=1}^{(T_\infty - T_\gamma)/T_\gamma} \Pr\left(\left\| \frac{\xi_{\infty,i}\xi'_{\infty,i}}{T_\gamma} \right\| \geq \left(1 + \sqrt{\frac{p}{T_\gamma}} + i^{1/4}\right)^2 \right) < \sum_{i=1}^{\infty} \exp\left\{ -\frac{T_\gamma i^{1/2}}{2} \right\}.$$

Clearly, the right hand side of the above inequality can be made arbitrarily small by choosing sufficiently large T_γ. Therefore, with large probability, for sufficiently large T_γ, all $\left\| \frac{\xi_{\infty,i}\xi'_{\infty,i}}{T_\gamma} \right\|$ are smaller than $\left(1 + \sqrt{\frac{p}{T_\gamma}} + i^{1/4}\right)^2$ and

$$\left\| p \frac{\xi_\infty \Delta_{12} \xi'_\infty}{T_\infty^2} \right\| \leq \frac{3p}{\pi^2 T_\gamma} \sum_{i=1}^{(T_\infty - T_\gamma)/T_\gamma} \frac{\left(1 + \sqrt{\frac{p}{T_\gamma}} + i^{1/4}\right)^2}{i^2} \leq K\gamma \tag{24}$$

for some constant K that does not depend on γ. Using the definition of α, (22), and (24) and recalling that, with high probability, $\left(\frac{p}{T_\gamma} D_\gamma\right)^{-1/2}$ can be represented as a matrix of rank no larger than τp plus a matrix of norm no larger than 5, we obtain that α can be represented as sum of a matrix of rank no larger than $3\tau p$ plus a matrix of norm no larger than $\bar{K}\gamma$, where $\bar{K}$ is some absolute constant.

Let us now focus on α_2. Write α_2 in the following form

$$\alpha_2 = \left(-\frac{\xi\xi'}{2T_\infty} + \frac{\xi_\gamma \left(\Delta'_{21} + I_{T_\gamma}/2\right) \xi'_\gamma}{T_\infty} - C_\gamma \right) + \frac{\xi_\infty \left(\Delta'_{22} + I_{T_\infty - T_\gamma}/2\right) \xi'_\infty}{T_\infty}.$$

Let us denote $\Delta'_{21} + I_{T_\gamma}/2$ as $\hat{\Delta}'_{21}$ and $\Delta'_{2\gamma} + I_{T_\gamma}/2$ as $\hat{\Delta}'_{2\gamma}$. Then

$$-\frac{\xi\xi'}{2T_\infty} + \frac{\xi_\gamma \left(\Delta'_{21} + I_{T_\gamma}/2\right) \xi'_\gamma}{T_\infty} - C_\gamma = -\frac{\xi\xi'}{2T_\infty} + \xi_\gamma \left(\frac{\hat{\Delta}'_{21}}{T_\infty} - \frac{\hat{\Delta}'_{2\gamma}}{T_\gamma} \right) \xi'_\gamma + \frac{\xi_\gamma \xi'_\gamma}{2T_\gamma}.$$

By definition, the block-diagonal elements of $\hat{\Delta}'_{21}$ (except the first block) have form

$$\begin{pmatrix} 0 & -\frac{1}{2}\frac{\sin 2\pi j/T_\infty}{1-\cos 2\pi j/T_\infty} \\ \frac{1}{2}\frac{\sin 2\pi j/T_\infty}{1-\cos 2\pi j/T_\infty} & 0 \end{pmatrix}$$

The block-diagonal elements of $\hat{\Delta}'_{2\gamma}$ have a similar form with T_∞ replaced by T_γ. Now,

$$\sin x = x - \frac{\cos t_1}{3!} x^3 \text{ and } \cos x = 1 - \frac{1}{2} x^2 + \frac{\cos t_2}{4!} x^4$$

for some $t_1, t_2 \in [0,x]$. Therefore, we have

$$\begin{aligned}\frac{1}{2}\frac{\sin 2\pi j/T_\infty}{1-\cos 2\pi j/T_\infty} &= \frac{2\pi j/T_\infty - \frac{\cos t_1}{6}\left(2\pi j/T_\infty\right)^3}{\left(2\pi j/T_\infty\right)^2 - \frac{\cos t_2}{12}\left(2\pi j/T_\infty\right)^4} \\ &= \frac{1}{2\pi j/T_\infty}\frac{1-\frac{\cos t_1}{6}\left(2\pi j/T_\infty\right)^2}{1-\frac{\cos t_2}{12}\left(2\pi j/T_\infty\right)^2},\end{aligned}$$

so that

$$\frac{1}{2T_\infty}\frac{\sin 2\pi j/T_\infty}{1-\cos 2\pi j/T_\infty} - \frac{1}{2\pi j} = \frac{\left(2\pi j/T_\infty\right)^2}{2\pi j}\frac{\frac{\cos t_2}{12}-\frac{\cos t_1}{6}}{1-\frac{\cos t_2}{12}\left(2\pi j/T_\infty\right)^2}$$

and thus,

$$\left|\frac{1}{2T_\infty}\frac{\sin 2\pi j/T_\infty}{1-\cos 2\pi j/T_\infty} - \frac{1}{2\pi j}\right| < \frac{6\pi j}{T_\infty^2}.$$

Similarly,

$$\left|\frac{1}{2T_\gamma}\frac{\sin 2\pi j/T_\gamma}{1-\cos 2\pi j/T_\gamma} - \frac{1}{2\pi j}\right| < \frac{6\pi j}{T_\gamma^2}.$$

Let $\xi_{\gamma 1}$ be a $p \times (T_\gamma - 2)/2$ matrix that consists of the odd columns of ξ_γ (starting from the third one) and let $\xi_{\gamma 2}$ be a $p \times (T_\gamma - 2)/2$ matrix that consists of the even columns (starting from the fourth one) of ξ_γ. Finally, let $\xi_{\gamma 0}$ be the $p \times 2$ matrix of the first two columns of ξ_γ. Then, the latter two inequalities and the fact that $j \leq \frac{T_\gamma}{2}$ imply that

$$\xi_\gamma\left(\frac{\hat{\Delta}_{21}'}{T_\infty} - \frac{\hat{\Delta}_{2\gamma}'}{T_\gamma}\right)\xi_\gamma' = \xi_{\gamma 0}\left(-\frac{1}{2T_\infty} + \frac{1}{2T_\gamma}\right)\xi_{\gamma 0}' + \xi_{\gamma 2}\Gamma\xi_{\gamma 1}' - \xi_{\gamma 1}\Gamma\xi_{\gamma 2}',$$

where Γ is a diagonal matrix with diagonal elements smaller than $3\pi/T_\gamma$ by absolute value. Since with high probability

$$\left\|\xi_{\gamma 0}\left(-\frac{1}{2T_\infty} + \frac{1}{2T_\gamma}\right)\xi_{\gamma 0}'\right\| \leq 2\gamma,$$

we have

$$\left\|\xi_\gamma\left(\frac{\hat{\Delta}_{21}'}{T} - \frac{\hat{\Delta}_{2\gamma}'}{T_\gamma}\right)\xi_\gamma'\right\| \leq 2\gamma + 2\left\|\xi_{\gamma 1}\Gamma\xi_{\gamma 2}'\right\|. \tag{25}$$

On the other hand, $\left\|\xi_{\gamma 1}\Gamma\xi_{\gamma 2}'\right\|$ is the square root of the largest eigenvalue of

$$\xi_{\gamma 1}\Gamma\xi_{\gamma 2}'\xi_{\gamma 2}\Gamma\xi_{\gamma 1}'.$$

Note that the rank of $T_\gamma \Gamma \xi'_{\gamma 2}\xi_{\gamma 2}\Gamma$ is no larger than p, and there exists an orthogonal transformation R such that $RT_\gamma \Gamma \xi'_{\gamma 2}\xi_{\gamma 2}\Gamma R'$ is diagonal with only the first p diagonal elements potentially non-zero. Furthermore, these non-zero diagonal elements will coincide with the eigenvalues of

$$T_\gamma \xi_{\gamma 2}\Gamma^2 \xi'_{\gamma 2}.$$

But

$$T_\gamma \xi_{\gamma 2}\Gamma^2 \xi'_{\gamma 2} \leq \frac{(3\pi)^2}{2}\frac{\xi_{\gamma 2}\xi'_{\gamma 2}}{T_\gamma/2}.$$

Assuming that γ is small, with high probability,

$$\left\| \frac{\xi_{\gamma 2}\xi'_{\gamma 2}}{T_\gamma/2} \right\| < 2.$$

Hence, the only p potentially non-zero diagonal elements of $RT_\gamma \Gamma \xi'_{\gamma 2}\xi_{\gamma 2}\Gamma R'$ are smaller than $(3\pi)^2$ with high probability.

Let $\xi_{\gamma 11}$ be the $p \times p$ matrix that consists of the first p columns of $\xi_{\gamma 1}R'$. Note that the entries of $\xi_{\gamma 11}$ are i.i.d. standard normals. Then, we have

$$\xi_{\gamma 1}\Gamma \xi'_{\gamma 2}\xi_{\gamma 2}\Gamma \xi'_{\gamma 1} \leq \frac{(3\pi)^2}{T_\gamma}\xi_{\gamma 11}\xi'_{\gamma 11}.$$

Since the norm of $\xi_{\gamma 11}\xi'_{\gamma 11}/p$ is smaller than 5 with high probability,

$$\left\| \xi_{\gamma 1}\Gamma \xi'_{\gamma 2}\xi_{\gamma 2}\Gamma \xi'_{\gamma 1} \right\| \leq (9\pi)^2 \gamma$$

with high probability. Combining this with (25), we obtain

$$\left\| \xi_\gamma \left(\frac{\hat{\Delta}'_{21}}{T_\infty} - \frac{\hat{\Delta}'_{2\gamma}}{T_\gamma} \right) \xi'_\gamma \right\| \leq 2\left(\gamma + 9\pi\sqrt{\gamma}\right). \tag{26}$$

Further,

$$\left\| -\frac{\xi\xi'}{2T_\infty} + \frac{\xi_\gamma \xi'_\gamma}{2T_\gamma} \right\| \leq \frac{1}{2}\left\| I_p - \frac{\xi\xi'}{T_\infty} \right\| + \frac{1}{2}\left\| I_p - \frac{\xi_\gamma \xi'_\gamma}{T_\gamma} \right\| \leq 4\sqrt{\gamma}$$

with high probability, for sufficiently large p, T and small γ. Combining this with

(26), we obtain

$$\left\| -\frac{\xi\xi'}{2T_\infty} + \frac{\xi_\gamma\left(\Delta'_{21} + I_{T_\gamma}/2\right)\xi'_1}{T_\infty} - C_\gamma \right\| \leq 2\left(\gamma + 10\pi\sqrt{\gamma}\right). \tag{27}$$

Next, consider $\frac{\xi_\infty\left(\Delta'_{22}+I_{T_\infty-T_\gamma}/2\right)\xi'_\infty}{T_\infty}$ part of α_2. Let $\xi_{\infty 1}$ be a $p\times\left(T_\infty - T_\gamma\right)/2$ matrix that consists of the odd columns of ξ_∞, and let $\xi_{\infty 2}$ be a $p\times\left(T_\infty - T_\gamma\right)/2$ matrix that consists of the even columns of ξ_∞. Then,

$$\frac{\xi_\infty\left(\Delta'_{22} + I_{T_\infty - T_\gamma}/2\right)\xi'_\infty}{T_\infty} = \xi_{\infty 2}\Upsilon\xi'_{\infty 1} - \xi_{\infty 1}\Upsilon\xi'_{\infty 2},$$

where

$$\Upsilon = \operatorname{diag}\left\{\frac{1}{2T_\infty}\frac{\sin 2\pi j/T_\infty}{1-\cos 2\pi j/T_\infty}\right\}$$

with j running from $T_\gamma/2$ to $T_\infty/2-1$. We have

$$\left\|\frac{\xi_\infty\left(\Delta'_{22} + I_{T_\infty - T_\gamma}/2\right)\xi'_\infty}{T_\infty}\right\|^2 \leq 4\left\|\xi_{\infty 2}\Upsilon\xi'_{\infty 1}\right\|^2 = 4\left\|\xi_{\infty 2}\Upsilon\xi'_{\infty 1}\xi_{\infty 1}\Upsilon\xi'_{\infty 2}\right\|.$$

Let R be the orthogonal matrix such that $R\Upsilon\xi'_{\infty 1}\xi_{\infty 1}\Upsilon R'$ is diagonal. Note that the rank of $\Upsilon\xi'_{\infty 1}\xi_{\infty 1}\Upsilon$ is no larger than p. Therefore, there are only p potentially non-zero elements on the diagonal of $R\Upsilon\xi'_{\infty 1}\xi_{\infty 1}\Upsilon R'$. Without loss of generality, these are the first p elements. Let $\xi_{\infty 21}$ be the first p columns of $\xi_{\infty 2}R'$. Then, we have

$$\begin{aligned}\left\|\frac{\xi_\infty\left(\Delta'_{22} + I_{T_\infty - T_\gamma}/2\right)\xi'_\infty}{T_\infty}\right\|^2 &\leq 4\left\|\xi_{\infty 21}\xi'_{\infty 21}\right\|\left\|R\Upsilon\xi'_{\infty 1}\xi_{\infty 1}\Upsilon R'\right\| \\ &= 4\left\|\xi_{\infty 21}\xi'_{\infty 21}\right\|\left\|\xi_{\infty 1}\Upsilon^2\xi'_{\infty 1}\right\|.\end{aligned}$$

Consider the decomposition

$$\xi_{\infty 1} = \left[\xi_{\infty 1,1}, ..., \xi_{\infty 1,(T_\infty - T_\gamma)/T_\gamma}\right],$$

and note that

$$\Upsilon^2 = \operatorname{diag}\left\{\Upsilon_1^2, ..., \Upsilon^2_{(T_\infty - T_\gamma)/T_\gamma}\right\}$$

with

$$\Upsilon_i = \text{diag}\left\{ \frac{1}{2T_\infty} \frac{\sin\left(i\frac{T_\gamma}{2}\frac{2\pi}{T_\infty}\right)}{1-\cos\left(i\frac{T_\gamma}{2}\frac{2\pi}{T_\infty}\right)}, ..., \frac{1}{2T_\infty} \frac{\sin\left(\left(i\frac{T_\gamma}{2}+\frac{T_\gamma}{2}-1\right)\frac{2\pi}{T_\infty}\right)}{1-\cos\left(\left(i\frac{T_\gamma}{2}+\frac{T_\gamma}{2}-1\right)\frac{2\pi}{T_\infty}\right)} \right\}.$$

Note that

$$\begin{aligned}
\left(\frac{1}{2T_\infty}\frac{\sin\left(\left(i\frac{T_\gamma}{2}+k\right)\frac{2\pi}{T_\infty}\right)}{1-\cos\left(\left(i\frac{T_\gamma}{2}+k\right)\frac{2\pi}{T_\infty}\right)}\right)^2 &= \frac{1}{2T_\infty^2}\frac{\cos^2\left(i\frac{T_\gamma}{2}+k\right)\frac{\pi}{T_\infty}}{1-\cos\left(i\frac{T_\gamma}{2}+k\right)\frac{2\pi}{T_\infty}} \\
&\leq \frac{1}{2T_\infty^2}\frac{1}{1-\cos\frac{iT_\gamma\pi}{T_\infty}} \\
&< \frac{3}{T_\gamma^2\pi^2}\frac{1}{i^2}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\left\|\xi_{\infty 1}\Upsilon^2\xi'_{\infty 1}\right\| &\leq \sum_{i=1}^{(T_\infty-T_\gamma)/T_\gamma} \left\|\xi_{\infty 1,i}\Upsilon_i^2\xi'_{\infty 1,i}\right\| \\
&\leq \sum_{i=1}^{(T_\infty-T_\gamma)/T_\gamma} \frac{3}{2T_\gamma^2\pi^2}\frac{1}{i^2}\left\|\frac{\xi_{\infty 1,i}\xi'_{\infty 1,i}}{T_\gamma/2}\right\|.
\end{aligned}$$

Using the large deviation inequality argument as above, we conclude that with high probability,

$$\left\|\xi_{\infty 1}\Upsilon^2\xi'_{\infty 1}\right\| \leq K\frac{1}{T_\gamma},$$

where K is an absolute constant. Therefore,

$$\left\|\frac{\xi_\infty\left(\Delta'_{22}+I_{T_\infty-T_\gamma}/2\right)\xi'_\infty}{T_\infty}\right\|^2 \leq 4K\frac{p}{T_\gamma}\left\|\frac{\xi_{\infty 21}\xi'_{\infty 21}}{p}\right\| \leq K_1\gamma,$$

where K_1 is an absolute constant. This implies that, with high probability,

$$\left\|\frac{\xi_\infty\left(\Delta'_{22}+I_{T_\infty-T_\gamma}/2\right)\xi'_\infty}{T_\infty}\right\| \leq K\sqrt{\gamma}$$

for some absolute constant K. Combining this with (27), we obtain

$$\|\alpha_2\| \leq K\sqrt{\gamma}$$

for some absolute constant K.

2.3.2 Proof of Lemma 3

Let $m \equiv m(z)$ be the Stieltjes transform of the limiting spectral distribution (LSD) of $D_\gamma \equiv \xi_\gamma \Delta_{1\gamma} \xi'_\gamma / T_\gamma$. By Silverstein and Bai (1995), for any $z \in \mathbb{C}^+$, m is the unique solution in $m \in \mathbb{C}^+$ of the equation

$$z = -\frac{1}{m} + \int \frac{t\mathrm{d}H(t)}{1+\gamma m t}, \tag{28}$$

where $H(t)$ is the limit of the empirical distribution of the diagonal elements of $\Delta_{1\gamma}$. Let S_{LSD} be the support of the LSD of D_γ. By continuity, the Stieltjes transform can be defined on S^c_{LSD}, where the superscript c denotes the complementary set. On any open subset of S^c_{LSD}, $m(z)$ still satisfies (28).

Silverstein and Choi (1995) show that S^c_{LSD} can be found as follows. Find $S_m \subset \mathbb{R}$, such that for any $m \in S_m$, $z(m)$ is well defined by (28) and has positive derivative at m. Then S^c_{LSD} coincides with $z(S_m)$.

Since

$$\Delta_{1\gamma} = \mathrm{diag}\left\{ r^{-1}_{\gamma 1} I_2, ..., r^{-1}_{\gamma T_\gamma/2} I_2 \right\}$$

with $r_{\gamma.j+1} = (2 - 2\cos 2\pi j/T_\gamma)$, we have

$$H(t) = \frac{1}{\pi} \arccos\left(1 - \frac{1}{2t}\right).$$

Therefore, $z(m)$ is well defined for $m < -4/\gamma$ and $m > 0$.

Direct computations similar to those in the proof of Lemma OW14, given below, yield

$$z = -\frac{1}{m} + \frac{1}{\sqrt{\gamma^2 m^2 + 4\gamma m}} \text{ for } m > 0, \text{ and} \tag{29}$$

$$z = -\frac{1}{m} - \frac{1}{\sqrt{\gamma^2 m^2 + 4\gamma m}} \text{ for } m < -4/\gamma, \tag{30}$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}m} z = \frac{1}{m^2} \mp \frac{\gamma^2 m + 2\gamma}{(\gamma^2 m^2 + 4\gamma m)^{3/2}}$$

where "−" correspond to $m > 0$ and "+" correspond to $m < -4/\gamma$.

For $m > 0$, $\frac{\mathrm{d}}{\mathrm{d}m} z > 0$ if and only if

$$\gamma \left(\gamma m + 4\right)^3 > m \left(\gamma m + 2\right)^2.$$

Denoting γm as x, we can rewrite the above condition as

$$\gamma^2 \left(x + 4\right)^3 - x \left(x + 2\right)^2 > 0.$$

For sufficiently small $\gamma > 0$, the function on the left hand side of this inequality is strictly decreasing on $x \geq 0$, positive at $x = 0$, and negative as $x \to \infty$. Let $x_0 \equiv \gamma m_0 > 0$ be such that the left hand side becomes zero. Then, a part of S^c_{LSD} consists of the image of $(0, m_0)$ under (29).

For $m < -4/\gamma$, $\frac{\mathrm{d}}{\mathrm{d}m} z > 0$ if and only if

$$\gamma \left(\gamma m + 4\right)^3 < m \left(\gamma m + 2\right)^2.$$

Denoting $\gamma m + 4$ as $x < 0$, we rewrite this condition as

$$\gamma^2 x^3 < \left(x - 4\right) \left(x - 2\right)^2.$$

For sufficiently small γ, this inequality is never satisfied for $x < 0$. Hence S^c_{LSD} consists entirely of the image of $(0, m_0)$ under (29).

Note that as $m \downarrow 0$, $z \downarrow -\infty$. Hence, S^c_{LSD} must have the form $\left(-\infty, d_-\right)$, where d_- is the lower bound of the support of the LSD. Incidentally, we see that the LSD is unbounded from above. Further, for sufficiently small γ,

$$d_- = -\frac{\gamma}{x_0} + \frac{1}{\sqrt{x_0^2 + 4x_0}}.$$

Since x_0 satisfies

$$\gamma^2 \left(x + 4\right)^3 - x \left(x + 2\right)^2 = 0, \tag{31}$$

we must have

$$\sqrt{x_0^2 + 4x_0} = \frac{x_0(x_0 + 2)}{\gamma \left(x_0 + 4\right)}$$

and

$$d_- = \frac{2\gamma}{x_0(x_0 + 2)}$$

Finally, from (31), we have

$$x_0 = 16\gamma^2 + o\left(\gamma^2\right).$$

Therefore,

$$d_- = \frac{1}{16\gamma}\left(1 + o(1)\right) > \frac{1}{17\gamma}$$

for sufficiently small γ.

2.4 Calculation of the integral in equation (OW22)

Using the explicit formula for the density of the continuous part of the Marchenko-Pastur distribution (OW21), we obtain

$$\int \mu \mathrm{d}F_0\left(\mu\right) = \int_{a_-}^{a_+} \frac{\mu}{2\pi} \frac{\sqrt{\left(a_+ - \mu\right)\left(\mu - a_-\right)}}{\mu} \mathrm{d}\mu,$$

where

$$a_{\pm} = \left(1 \pm \sqrt{2}\right)^2.$$

Let $x = \left(\mu - a_-\right) / \left(a_+ - a_-\right)$ so that $\mu = a_- + \left(a_+ - a_-\right) x$. Then

$$\begin{aligned} \int \mu \mathrm{d}F_0\left(\mu\right) &= \frac{\left(a_+ - a_-\right)^2}{2\pi} \int_0^1 \sqrt{\left(1-x\right) x} \mathrm{d}x \\ &= \frac{\left(a_+ - a_-\right)^2}{2\pi} \frac{\pi}{8} = 2. \end{aligned}$$

3 Supplementary Material for Appendix of OW

3.1 Proof of Lemma OW10

Write X_t in the VAR(k) form

$$X_t = \sum_{i=1}^{k} \Pi_i X_{t-i} + \Psi F_t + \varepsilon_t,$$

where Π_i are such that $\Pi = \sum_{i=1}^{k} \Pi_i - I_p$ and $\Gamma_i = -\sum_{j=i+1}^{k} \Pi_j$. Express X_t as a function of the initial values, ε and F (see Theorem 2.1 in Johansen (1995))

$$X_t = \sum_{s=1}^{k} C_{t-s} \sum_{i=1}^{k-s+1} \Pi_{s+i-1} X_{1-i} + \sum_{j=0}^{t-1} C_j \left(\varepsilon_{t-j} + \Psi F_{t-j}\right), \tag{32}$$

where $C_0 = I$ and C_n is defined recursively by

$$C_n = \sum_{j=1}^{k \wedge n} C_{n-j} \Pi_j, \ n = 1, 2, ...$$

Here $k \wedge n$ denotes the minimum of k and n, and $\Pi_j = 0$ for $j > k$. Let us denote $\Pi_1 - I$ as Π_1^* and let $\Pi_j^* = \Pi_j$ for $j \geq 2$. Then, for $n = 1, 2, ...,$

$$\Delta C_n = C_n - C_{n-1} = \sum_{j=1}^{k \wedge n} C_{n-j} \Pi_j^* = \sum_{j=1}^{n-1} \Delta C_{n-j} \sum_{s=1}^{j \wedge k} \Pi_s^* + \sum_{s=1}^{n \wedge k} \Pi_s^*. \tag{33}$$

Clearly the *column* space of ΔC_1 is spanned by the *column* spaces of Π_j^*, $j = 1, ..., k$. Use this as the basis of induction. Suppose that the *column* spaces of each of ΔC_j with $j < n$ are spanned by the *column* spaces of Π_j^*, $j = 1, ..., k$. The identity (33) then implies that the *column* space of ΔC_n is spanned by the *column* spaces of Π_j^*, $j = 1, ..., k$, too.

Now rewrite (32) as

$$X_t = \sum_{s=1}^{k} \sum_{h=0}^{t-s} \Delta C_h \sum_{i=1}^{k} \Pi_{s+i-1} X_{1-i} + \sum_{j=0}^{t-1} \sum_{h=0}^{j} \Delta C_h \left(\varepsilon_{t-j} + \Psi F_{t-j}\right)$$

where $\Delta C_0 = C_0 = I_p$. Represent X_t as a sum $X_t^{(0)} + X_t^{(1)}$, where

$$X_t^{(0)} = \sum_{s=1}^{k} \sum_{i=1}^{k} \Pi_{s+i-1} X_{1-i} + \sum_{j=0}^{t-1} \left(\varepsilon_{t-j} + \Psi F_{t-j}\right) \text{ and} \tag{34}$$

$$X_t^{(1)} = \sum_{s=1}^{k} \sum_{h=1}^{t-s} \Delta C_h \sum_{i=1}^{k} \Pi_{s+i-1} X_{1-i} + \sum_{j=0}^{t-1} \sum_{h=1}^{j} \Delta C_h \left(\varepsilon_{t-j} + \Psi F_{t-j}\right). \tag{35}$$

Since the column spaces of each of ΔC_h with $h \geq 1$ are spanned by those of Π_j^*, $j = 1, ..., k$, the space spanned by $X_t^{(1)}$, $t = 1, ..., T$ is also spanned by the columns spaces of Π_j^*, $j = 1, ..., k$. Since the union of the latter column spaces coincides with the union of the column spaces of Π and Γ, we have

$$\operatorname{rank} X^{(1)} \leq r + \operatorname{rank} \Gamma, \tag{36}$$

where $X^{(1)} = [X_{1-k}^{(1)}, ..., X_T^{(1)}]$ with zero $X_{1-k}^{(1)}, ..., X_0^{(1)}$, and $X_t^{(1)}$ with $t \geq 1$ defined by (35).

Next, represent $X_t^{(0)}$ as a sum $X_t^{(00)}$ and $\tilde{X}_t$, where

$$\begin{aligned} X_t^{(00)} &= \sum_{s=1}^{k} \sum_{i=1}^{k} \Pi_{s+i-1} X_{1-i} + \Psi \sum_{j=0}^{t-1} F_{t-j} \text{ and} & (37) \\ \tilde{X}_t &= \sum_{j=0}^{t-1} \varepsilon_{t-j}, & (38) \end{aligned}$$

and let $X^{(00)} = [X_{1-k}^{(00)}, ..., X_T^{(00)}]$ with $X_t^{(00)} = X_t$ for $t = 1-k, ..., 0$ and $X_t^{(00)}$ with $t \geq 1$ defined by (37). Note that the columns space $X^{(00)}$ is spanned by those of Π_j^*, $j = 1, ..., k$, the column space of the matrix of the initial conditions $[X_{1-k}, ..., X_0]$, and the column space of Ψ. Therefore,

$$\operatorname{rank} X^{(00)} \leq r + \operatorname{rank} \Gamma + k + d_F. \tag{39}$$

Since $X = X^{(1)} + X^{(00)} + \tilde{X}$, inequalities (36) and (39) yield

$$\operatorname{rank} \left(X - \tilde{X} \right) \leq 2 \left(r + \operatorname{rank} \Gamma + k + d_F \right).$$

3.2 Derivation of Equation (OW43)

Applying the Sherman-Morrison-Woodbury formula (OW41) to the right hand side of

$$D^{-1} = \left(D_j + \varepsilon_{(j)} \Delta_{1j} \varepsilon_{(j)}' \right)^{-1},$$

we obtain

$$\begin{aligned} D^{-1} &= D_j^{-1} - D_j^{-1} \varepsilon_{(j)} \left(\Delta_{1j}^{-1} + \varepsilon_{(j)}' D_j^{-1} \varepsilon_{(j)} \right)^{-1} \varepsilon_{(j)}' D_j^{-1} \\ &= D_j^{-1} - D_j^{-1} \varepsilon_{(j)} \left(r_j I_2 + s_j \right)^{-1} \varepsilon_{(j)}' D_j^{-1}. & (40) \end{aligned}$$

Using this and the identity

$$C = C_j + \varepsilon_{(j)} \Delta'_{2j} \varepsilon'_{(j)}, \tag{41}$$

we expand $CD^{-1}C'$ in the following form

$$\begin{aligned}
&C_j D_j^{-1} C'_j + \varepsilon_{(j)} \Delta'_{2j} \varepsilon'_{(j)} D_j^{-1} C'_j - C_j D_j^{-1} \varepsilon_{(j)} \left(r_j I_2 + s_j\right)^{-1} \varepsilon'_{(j)} D_j^{-1} C'_j \\
&+ C_j D_j^{-1} \varepsilon_{(j)} \Delta_{2j} \varepsilon'_{(j)} - \varepsilon_{(j)} \Delta'_{2j} \varepsilon'_{(j)} D_j^{-1} \varepsilon_{(j)} \left(r_j I_2 + s_j\right)^{-1} \varepsilon'_{(j)} D_j^{-1} C'_j \\
&- C_j D_j^{-1} \varepsilon_{(j)} \left(r_j I_2 + s_j\right)^{-1} \varepsilon'_{(j)} D_j^{-1} \varepsilon_{(j)} \Delta_{2j} \varepsilon'_{(j)} + \varepsilon_{(j)} \Delta'_{2j} \varepsilon'_{(j)} D_j^{-1} \varepsilon_{(j)} \Delta_{2j} \varepsilon'_{(j)} \\
&- \varepsilon_{(j)} \Delta'_{2j} \varepsilon'_{(j)} D_j^{-1} \varepsilon_{(j)} \left(r_j I_2 + s_j\right)^{-1} \varepsilon'_{(j)} D_j^{-1} \varepsilon_{(j)} \Delta_{2j} \varepsilon'_{(j)}.
\end{aligned}$$

Recalling that $\varepsilon'_{(j)} D_j^{-1} \varepsilon_{(j)} = s_j$, we further simplify this to obtain

$$\begin{aligned}
CD^{-1}C' \;=\;& C_j D_j^{-1} C'_j - C_j D_j^{-1} \varepsilon_{(j)} \left(r_j I_2 + s_j\right)^{-1} \varepsilon'_{(j)} D_j^{-1} C'_j \\
&+ \varepsilon_{(j)} \Delta'_{2j} r_j \left(r_j I_2 + s_j\right)^{-1} \varepsilon'_{(j)} D_j^{-1} C'_j \\
&+ C_j D_j^{-1} \varepsilon_{(j)} \left(r_j I_2 + s_j\right)^{-1} r_j \Delta_{2j} \varepsilon'_{(j)} \\
&+ \varepsilon_{(j)} \Delta'_{2j} s_j \left(r_j I_2 + s_j\right)^{-1} r_j \Delta_{2j} \varepsilon'_{(j)}.
\end{aligned}$$

Since $M = CD^{-1}C' - zA$, it follows that

$$M^{-1} = \left(M_j + \alpha_j K_j \alpha'_j\right)^{-1}, \tag{42}$$

where

$$\begin{aligned}
M_j &= C_j D_j^{-1} C'_j - zA_j, \\
\alpha_j &= [\varepsilon_{(j)}, C_j D_j^{-1} \varepsilon_{(j)}]
\end{aligned}$$

and

$$K_j = \begin{pmatrix} \Delta'_{2j} s_j \left(r_j I_2 + s_j\right)^{-1} r_j \Delta_{2j} - zI_2 & \Delta'_{2j} r_j \left(r_j I_2 + s_j\right)^{-1} \\ \left(r_j I_2 + s_j\right)^{-1} r_j \Delta_{2j} & -\left(r_j I_2 + s_j\right)^{-1} \end{pmatrix}.$$

Applying the Sherman-Morrison-Woodbury formula to the right hand side of (42), we obtain

$$M^{-1} = M_j^{-1} - M_j^{-1} \alpha_j \left(K_j^{-1} + \alpha'_j M_j^{-1} \alpha_j\right)^{-1} \alpha'_j M_j^{-1}. \tag{43}$$

Since

$$\Delta'_{2j} \Delta_{2j} = \Delta_{1j} = r_j^{-1} I_2, \tag{44}$$

we can write

$$K_j = \begin{pmatrix} \Delta'_{2j} & 0 \\ 0 & I_2 \end{pmatrix} \begin{pmatrix} s_j\left(r_j I_2 + s_j\right)^{-1} r_j - z r_j I_2 & \left(r_j I_2 + s_j\right)^{-1} r_j \\ \left(r_j I_2 + s_j\right)^{-1} r_j & -\left(r_j I_2 + s_j\right)^{-1} \end{pmatrix} \begin{pmatrix} \Delta_{2j} & 0 \\ 0 & I_2 \end{pmatrix},$$

which implies that

$$K_j^{-1} = \frac{1}{1-z} \begin{pmatrix} \Delta_{2j}^{-1} & 0 \\ 0 & I_2 \end{pmatrix} \begin{pmatrix} r_j^{-1} I_2 & I_2 \\ I_2 & z\left(r_j I_2 + s_j\right) - s_j \end{pmatrix} \begin{pmatrix} \Delta_{2j}'^{-1} & 0 \\ 0 & I_2 \end{pmatrix},$$

and therefore, using (44), we obtain

$$K_j^{-1} = \begin{pmatrix} \frac{1}{1-z} I_2 & \frac{1}{1-z} r_j \Delta'_{2j} \\ \frac{1}{1-z} r_j \Delta_{2j} & \frac{z}{1-z} r_j I_2 - s_j \end{pmatrix}. \tag{45}$$

Further, the definitions

$$\begin{aligned} u_j &= \varepsilon'_{(j)} D_j^{-1} C'_j M_j^{-1} \varepsilon_{(j)}, \\ v_j &= \varepsilon'_{(j)} M_j^{-1} \varepsilon_{(j)}, \text{ and} \\ w_j &= \varepsilon'_{(j)} D_j^{-1} C'_j M_j^{-1} C_j D_j^{-1} \varepsilon_{(j)} \end{aligned}$$

yield

$$\alpha'_j M_j^{-1} \alpha_j = \begin{pmatrix} v_j & u'_j \\ u_j & w_j \end{pmatrix}. \tag{46}$$

Using (45) and (46) in (43), we obtain

$$M^{-1} = M_j^{-1} - M_j^{-1} \alpha_j \begin{pmatrix} \frac{1}{1-z} I_2 + v_j & \frac{1}{1-z} r_j \Delta'_{2j} + u'_j \\ \frac{1}{1-z} r_j \Delta_{2j} + u_j & \frac{z}{1-z} r_j I_2 - s_j + w_j \end{pmatrix}^{-1} \alpha'_j M_j^{-1}, \tag{47}$$

which yields

$$\varepsilon'_{(j)} M^{-1} \varepsilon_{(j)} = v_j - [v_j, u'_j] \begin{pmatrix} \frac{1}{1-z} I_2 + v_j & \frac{1}{1-z} r_j \Delta'_{2j} + u'_j \\ \frac{1}{1-z} r_j \Delta_{2j} + u_j & \frac{z}{1-z} r_j I_2 - s_j + w_j \end{pmatrix}^{-1} \begin{bmatrix} v_j \\ u_j \end{bmatrix}$$

3.3 Proof of Lemma OW12

We start from an elementary lemma describing absolute central moments of a quadratic form in i.i.d. normal random variables. Its proof is given in the next section of this note.

Lemma 4 *Let ρ be a positive integer, Ω be a $p \times p$ deterministic complex matrix, and $\xi \sim N_p\left(0, \frac{1}{T} I_p\right)$. Then*

$$\mathbb{E}\left|\xi'\Omega\xi - \frac{1}{T}\operatorname{tr}\Omega\right|^{2\rho} \leq \frac{C_\rho \|\Omega\|^{2\rho} p^\rho}{T^{2\rho}},$$

where

$$C_\rho = 2^{4\rho}(2\rho+1)^{2\rho}\rho^{2\rho+1}.$$

Consider a sequence $\{p, T\} \equiv \{p_T, T\}$ such that $p_T, T \to_c \infty$. We introduce notation p_T to emphasize the fact that the sequence $\{p, T\}$ with $p, T \to_c \infty$ can be indexed by T without loss of generality. Let us use Lemma 4 to prove that, as $p_T, T \to_c \infty$,

$$\max_{j=1,\ldots,T/2} \|s_j - \hat{s}I_2\| \overset{a.s.}{\to} 0. \tag{48}$$

Since the square of the spectral norm is no larger than the sum of the squared elements of the matrix, it is sufficient to prove the element-wise convergences. Take, for example, the element in the second row and the second column of $s_j - \hat{s}I_2$. We need to show that

$$\max_{j=1,\ldots,T/2} \left|\varepsilon_{2j}' D_j^{-1}\varepsilon_{2j} - \hat{s}\right| \overset{a.s.}{\to} 0. \tag{49}$$

For any $\tau > 0$, let $E_{j\tau}$ be the event

$$E_{j\tau} = \left\{\left|\varepsilon_{2j}' D_j^{-1}\varepsilon_{2j} - \hat{s}\right| > 2\tau\right\}.$$

The probability of $E_{j\tau}$ is bounded above as follows

$$\Pr(E_{j\tau}) \leq \Pr(E_{1,j\tau}) + \Pr(E_{2,j\tau}),$$

where

$$\begin{aligned} E_{1,j\tau} &= \left\{\left|\varepsilon_{2j}' D_j^{-1}\varepsilon_{2j} - \frac{1}{T}\operatorname{tr}\left[D_j^{-1}\right]\right| > \tau\right\}, \text{ and} \\ E_{2,j\tau} &= \left\{\left|\frac{1}{T}\operatorname{tr}\left[D_j^{-1} - D\right]\right| > \tau\right\}. \end{aligned}$$

Note that

$$\left\|D_j^{-1}\right\| \leq \left\|D^{-1}\right\| = \frac{1}{\lambda_{\min}(\varepsilon\Delta_1\varepsilon')} \leq \frac{1}{\lambda_{\min}(\Delta_1)\lambda_{\min}(\varepsilon\varepsilon')} = \frac{4}{\lambda_{\min}(\varepsilon\varepsilon')}, \tag{50}$$

where $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue of a matrix. As shown by Bai and Yin (1993),

$$\lambda_{\min}(\varepsilon\varepsilon') \xrightarrow{a.s.} \left(1-\sqrt{c}\right)^2$$

when $p_T, T \to_c \infty$. Fix $\delta > 4\left(1-\sqrt{c}\right)^{-2}$. Then, almost surely, there exists T_δ that does not depend on j, such that

$$\left\|D_j^{-1}\right\| \leq \delta \tag{51}$$

for any $T > T_\delta$. Let $\mathbb{E}_j$ denote the expectation conditional on the event (51). Lemma 4 yields

$$\mathbb{E}_j \left|\varepsilon_{2j}' D_j^{-1} \varepsilon_{2j} - \frac{1}{T} \operatorname{tr}\left[D_j^{-1}\right]\right|^{2\rho} \leq \frac{C_\rho \delta^{2\rho} p_T^\rho}{T^{2\rho}}. \tag{52}$$

We have

$$\begin{aligned} \Pr\left(E_{1,j\tau}\right) &\leq \Pr_j\left(E_{1,j\tau}\right) \Pr\left(\left\|D_j^{-1}\right\| \leq \delta\right) + \Pr\left(\left\|D_j^{-1}\right\| > \delta\right) \\ &\leq \Pr_j\left(E_{1,j\tau}\right) + \Pr\left(\lambda_{\min}(\varepsilon\varepsilon') < 4\delta^{-1}\right), \end{aligned}$$

where $\Pr_j$ is the probability conditional on the event (51). Markov's inequality together with (52) imply that

$$\Pr\left(E_{1,j\tau}\right) \leq \frac{C_\rho \delta^{2\rho} p_T^\rho}{\tau^{2\rho} T^{2\rho}} + \Pr\left(\lambda_{\min}(\varepsilon\varepsilon') < 4\delta^{-1}\right). \tag{53}$$

Further, since the rank of the positive semi-definite matrix $D^{-1} - D_j^{-1}$ is no larger than two, we have by Weyl's theorem (see Theorem 4.3.6 in Horn and Johnson (1985))

$$\frac{1}{T}\left|\operatorname{tr}\left[D_j^{-1} - D\right]\right| \leq \frac{2}{T}\left\|D^{-1}\right\|.$$

Therefore, using (50), we obtain

$$\Pr\left(E_{2,j\tau}\right) \leq \Pr\left(\frac{2}{T}\left\|D^{-1}\right\| > \tau\right) \leq \Pr\left(\lambda_{\min}(\varepsilon\varepsilon') < \frac{8}{T}\tau^{-1}\right). \tag{54}$$

For $\lambda_{\min}(\varepsilon\varepsilon')$, we have the following large deviation inequality (Theorem II.13 in Davidson and Szarek (2001) establishes the inequality given below and a similar inequality for $\lambda_{\max}(\varepsilon\varepsilon')$). For any $\alpha > 0$,

$$\Pr\left(\lambda_{\min}(\varepsilon\varepsilon') < \left(1 - \sqrt{p_T/T} - \alpha\right)^2\right) < \exp\left\{-T\alpha^2/2\right\}. \tag{55}$$

Let

$$\alpha = \left(1 - \sqrt{c}\right)/2 - \delta^{-1/2} > 0,$$

and let $T_{\rho\delta\tau}$ be such that for all $T > T_{\rho\delta\tau}$

$$\left(1 - \sqrt{p_T/T} - \alpha\right)^2 > \max\left\{4\delta^{-1}, \frac{8}{T}\tau^{-1}\right\} \tag{56}$$

and

$$\exp\left\{-T\alpha^2/2\right\} < \frac{C_\rho \delta^{2\rho} p_T^\rho}{\tau^{2\rho}T^{2\rho}}. \tag{57}$$

Let $T_0 = \max\{T_\delta, T_{\rho\delta\tau}\}$. Note that T_0 does not depend on j. Inequalities (53-57) imply that for all $T > T_0$

$$\Pr(E_{j\tau}) \leq \frac{3C_\rho \delta^{2\rho} p_T^\rho}{\tau^{2\rho}T^{2\rho}}.$$

Finally,

$$\Pr\left\{\max_{j=1,\ldots,T/2}\left|\varepsilon_{2j}' D_j^{-1}\varepsilon_{2j} - \hat{s}\right| > 2\tau\right\} \leq \sum_{j=1}^{T/2}\Pr(E_{j\tau}) \leq \frac{3C_\rho \delta^{2\rho} p_T^\rho}{2\tau^{2\rho}T^{2\rho-1}}.$$

For $\rho \geq 2$,

$$\sum_{T=T_0+1}^{\infty}\frac{3C_\rho \delta^{2\rho} p_T^\rho}{2\tau^{2\rho}T^{2\rho-1}} < \infty,$$

which yields (49) by the Borel-Cantelli lemma.

The convergence of the element in the first row and the first column of $s_j - \hat{s}I_2$ can be shown similarly to (49). For the off-diagonal elements, note that

$$\varepsilon_{2j-1}' D_j^{-1}\varepsilon_{2j} = \frac{1}{2}\left(\varepsilon_{2j-1}', \varepsilon_{2j}'\right)\begin{pmatrix} 0 & D_j^{-1} \\ D_j^{-1} & 0 \end{pmatrix}\begin{pmatrix} \varepsilon_{2j-1} \\ \varepsilon_{2j} \end{pmatrix}.$$

Hence, we again can use Lemma 4 and the Borel-Cantelli lemma to obtain desired results.

It remains to prove that

$$\max_{j=1,\ldots,T/2}\|u_j - \hat{u}I_2\| \xrightarrow{a.s} 0,$$

and to establish similar convergences for v_j and w_j. This can be accomplished by closely following the strategy of the above proof of (48), which we leave to the reader. The only two new aspects of the remaining proofs are related to the need for bounds on the spectral norms of $D_j^{-1}C_j'M_j^{-1}$, M_j^{-1}, and $D_j^{-1}C_j'M_j^{-1}C_jD_j^{-1}$, and on the differences between the traces of these matrices and the traces of $D^{-1}C'M^{-1}$, M^{-1}, and $D^{-1}C'M^{-1}CD^{-1}$, respectively. Such bounds can be obtained using the following lemma. Its proof is given in the next section of this note.

Lemma 5 *Let z be as in the definition (OW37) of M, and let $\Im z > 0$ be the imaginary part of z. Then, we have*

$$\left\|M_j^{-1}\right\| \le \frac{1}{(\Im z)\,\lambda_{\min}(\varepsilon\varepsilon')}, \text{ and } \left\|D_j^{-1}C_j'\right\|^2 \le \frac{4\lambda_{\max}(\varepsilon\varepsilon')}{\lambda_{\min}(\varepsilon\varepsilon')}.$$

The same bounds hold for the norms of M^{-1} and $D^{-1}C'$. Further,

$$\left|\operatorname{tr}\left(M_j^{-1} - M^{-1}\right)\right| \le \frac{8}{(\Im z)\,\lambda_{\min}(\varepsilon\varepsilon')},$$

$$\left|\operatorname{tr}\left(D_j^{-1}C_j'M_j^{-1} - D^{-1}C'M^{-1}\right)\right| \le \frac{32\lambda_{\max}^{1/2}(\varepsilon\varepsilon')}{(\Im z)\,\lambda_{\min}^{3/2}(\varepsilon\varepsilon')},$$

and

$$\left|\operatorname{tr}\left(D_j^{-1}C_j'M_j^{-1}C_jD_j^{-1} - D^{-1}C'M^{-1}CD^{-1}\right)\right| \le \frac{96\lambda_{\max}(\varepsilon\varepsilon')}{(\Im z)\,\lambda_{\min}^{2}(\varepsilon\varepsilon')}.$$

3.4 Proofs of Lemmas 4 and 5

3.4.1 Proof of Lemma 4

Without loss of generality, we can assume that $\Omega = A + \mathrm{i}B$, where A and B are real symmetric matrices. Indeed, the expression $\xi'\Omega\xi - \frac{1}{T}\operatorname{tr}\Omega$ does not change if Ω is replaced by $\frac{1}{2}(\Omega + \Omega')$, and the latter matrix obviously has the required form. Using the Hölder inequality, we obtain

$$\left|\xi'\Omega\xi - \frac{1}{T}\operatorname{tr}\Omega\right|^{2\rho} \le 2^{2\rho-1}\left(\left(\xi'A\xi - \frac{1}{T}\operatorname{tr}A\right)^{2\rho} + \left(\xi'B\xi - \frac{1}{T}\operatorname{tr}B\right)^{2\rho}\right). \quad (58)$$

Let UdU' be the spectral decomposition of A, where $d = \operatorname{diag}\{d_1, ..., d_p\}$. The rotational invariance of the distribution $N_p\left(0, \frac{1}{T} I_p\right)$ implies that

$$\mathbb{E}\left(\xi' A\xi - \frac{1}{T}\operatorname{tr} A\right)^{2\rho} = \mathbb{E}\left(\sum_{i=1}^{p} d_i\left(\xi_i^2 - \frac{1}{T}\right)\right)^{2\rho}. \tag{59}$$

Since $\mathbb{E}\left(\xi_i^2 - \frac{1}{T}\right) = 0$, the right hand side of the above equality can be expanded as

$$\sum_{l=1}^{\rho} \sum_{1\le i_1<\cdots<i_l\le p} \sum_{\substack{j_1+\cdots j_l=2\rho \\ j_1\ge 2,\cdots,j_l\ge 2}} (2\rho)! \prod_{t=1}^{l} \frac{(d_{i_t})^{j_t}\,\mathbb{E}\left(\xi_{i_t}^2 - \frac{1}{T}\right)^{j_t}}{j_t!}. \tag{60}$$

Now, let μ_j be the j-th raw moment of a $\chi^2(1)$ random variable. As is well known, $\mu_j = (2j-1)!!$ with a crude upper bound

$$\mu_j \le 2^j j^j.$$

Since the j-th absolute central moment of a $\chi^2(1)$ random variable is no larger than $\mu_j + 1$, we have, for any $j \ge 1$,

$$\mathbb{E}\left|\xi_i^2 - \frac{1}{T}\right|^j \le \frac{\mu_j + 1}{T^j} \le \frac{2^j\,(j+1)^j}{T^j}.$$

Further, clearly

$$|d_i|^j \le \|A\|^j.$$

Using the latter two inequalities in (60), we obtain

$$\begin{aligned}\mathbb{E}\left(\sum_{i=1}^{p} d_i\left(\xi_i^2 - \frac{1}{T}\right)\right)^{2\rho} &\le \frac{\|A\|^{2\rho}(4\rho+2)^{2\rho}}{T^{2\rho}} \sum_{l=1}^{\rho} p^l \sum_{\substack{j_1+\cdots j_l=2\rho \\ j_1\ge 2,\cdots,j_l\ge 2}} \prod_{t=1}^{l} \frac{(2\rho)!}{j_t!} \\ &\le \frac{\|A\|^{2\rho}(4\rho+2)^{2\rho}}{T^{2\rho}} \sum_{l=1}^{\rho} p^l l^{2\rho}.\end{aligned}$$

Using another crude upper bound

$$\sum_{l=1}^{\rho} p^l l^{2\rho} \le \rho^{2\rho+1} p^\rho,$$

we conclude that

$$\mathbb{E}\left(\sum_{i=1}^{p} d_i\left(\xi_i^2-\frac{1}{T}\right)\right)^{2\rho} \leq \frac{\|A\|^{2\rho}(4\rho+2)^{2\rho}\rho^{2\rho+1}p^{\rho}}{T^{2\rho}}. \tag{61}$$

Now note that $\|A\| \leq \|\Omega\|$. This follows, for example, from the equality

$$A^2+B^2=\frac{1}{2}\left(\Omega^*\Omega+\Omega\Omega^*\right),$$

where the superscript '$*$' denotes the operation of transposition and complex conjugation. Therefore, (59) and (61) imply

$$\mathbb{E}\left(\xi' A\xi-\frac{1}{T}\operatorname{tr} A\right)^{2\rho} \leq \frac{\|\Omega\|^{2\rho}(4\rho+2)^{2\rho}\rho^{2\rho+1}p^{\rho}}{T^{2\rho}}.$$

Similarly, we have

$$\mathbb{E}\left(\xi' B\xi-\frac{1}{T}\operatorname{tr} B\right)^{2\rho} \leq \frac{\|\Omega\|^{2\rho}(4\rho+2)^{2\rho}\rho^{2\rho+1}p^{\rho}}{T^{2\rho}}.$$

Using the latter two inequalities in (58), we obtain

$$\mathbb{E}\left|\xi'\Omega\xi-\frac{1}{T}\operatorname{tr}\Omega\right|^{2\rho} \leq \frac{\|\Omega\|^{2\rho}2^{4\rho}(2\rho+1)^{2\rho}\rho^{2\rho+1}p^{\rho}}{T^{2\rho}}.\square$$

3.4.2 Proof of Lemma 5

By definition of M_j, we have

$$\begin{aligned}\left\|M_j^{-1}\right\| &= \left\|A_j^{-1/2}\left(A_j^{-1/2}C_jD_j^{-1}C_j'A_j^{-1/2}-zI_p\right)^{-1}A_j^{-1/2}\right\| \\ &\leq \left\|A_j^{-1}\right\|\left\|\left(A_j^{-1/2}C_jD_j^{-1}C_j'A_j^{-1/2}-zI_p\right)^{-1}\right\|.\end{aligned}$$

On the other hand,

$$\left\|A_j^{-1}\right\| \leq \left\|A^{-1}\right\| = \left\|\left(\varepsilon\varepsilon'\right)^{-1}\right\| = \frac{1}{\lambda_{\min}\left(\varepsilon\varepsilon'\right)}$$

and

$$\left\| \left(A_j^{-1/2} C_j D_j^{-1} C_j' A_j^{-1/2} - z I_p \right)^{-1} \right\| \leq \max_{k=1,\ldots,p} \frac{1}{\left| \lambda_k \left(A_j^{-1/2} C_j D_j^{-1} C_j' A_j^{-1/2} \right) - z \right|}$$

where $\lambda_k(\cdot)$ is the k-th largest eigenvalue of a real symmetric matrix. The above inequality implies that

$$\left\| \left(A_j^{-1/2} C_j D_j^{-1} C_j' A_j^{-1/2} - z I_p \right)^{-1} \right\| \leq \frac{1}{\Im z},$$

and therefore,

$$\left\| M_j^{-1} \right\| \leq \frac{1}{(\Im z)\, \lambda_{\min}(\varepsilon \varepsilon')}. \tag{62}$$

The same bound for $\|M^{-1}\|$ is established similarly.

Further, we have

$$\begin{aligned} \left\| D_j^{-1} C_j' \right\|^2 &= \left\| D_j^{-1} C_j' C_j D_j^{-1} \right\| \\ &= \left\| D_j^{-1} \varepsilon_{-(j)} \Delta_{2(j)}' \varepsilon_{-(j)}' \varepsilon_{-(j)} \Delta_{2(j)} \varepsilon_{-(j)}' D_j^{-1} \right\|, \end{aligned}$$

where $\Delta_{2(j)}$ is the block-diagonal matrix obtained from Δ_2 by removing its j-th 2×2 block, and $\varepsilon_{-(j)}$ is obtained from ε by removing the $2j-1$-th and $2j$-th columns. On the other hand,

$$\begin{aligned} & \left\| D_j^{-1} \varepsilon_{-(j)} \Delta_{2(j)}' \varepsilon_{-(j)}' \varepsilon_{-(j)} \Delta_{2(j)} \varepsilon_{-(j)}' D_j^{-1} \right\| \\ \leq\ & \lambda_{\max}(\varepsilon \varepsilon') \left\| D_j^{-1} \varepsilon_{-(j)} \Delta_{2(j)}' \Delta_{2(j)} \varepsilon_{-(j)}' D_j^{-1} \right\| \\ =\ & \lambda_{\max}(\varepsilon \varepsilon') \left\| D_j^{-1} \varepsilon_{-(j)} \Delta_{1(j)} \varepsilon_{-(j)}' D_j^{-1} \right\| \\ =\ & \lambda_{\max}(\varepsilon \varepsilon') \left\| D_j^{-1} \right\|, \end{aligned}$$

where we used the fact that $\Delta_{2(j)}' \Delta_{2(j)} = \Delta_{1(j)}$ with $\Delta_{1(j)}$ obtained from Δ_1 by removing its j-th 2×2 block. Using (50), we obtain

$$\left\| D_j^{-1} C_j' \right\|^2 \leq \frac{4 \lambda_{\max}(\varepsilon \varepsilon')}{\lambda_{\min}(\varepsilon \varepsilon')}. \tag{63}$$

The same bound for $\|D^{-1} C'\|$ is established similarly.

Now, let us establish the bounds on the differences of traces. As follows from

(47), M_j^{-1} differs from M^{-1} by a matrix of rank no larger than 4. Therefore,

$$\left|\operatorname{tr}\left(M_j^{-1}-M^{-1}\right)\right| \leq 4\left\|M_j^{-1}-M^{-1}\right\| \leq 4\left(\left\|M_j^{-1}\right\|+\left\|M^{-1}\right\|\right).$$

Therefore,

$$\left|\operatorname{tr}\left(M_j^{-1}-M^{-1}\right)\right| \leq \frac{8}{(\Im z)\,\lambda_{\min}(\varepsilon\varepsilon')}.$$

Similarly, $D_j^{-1}C_j'M_j^{-1}$ differs from $D^{-1}C'M^{-1}$ by a matrix with rank no larger than 8. It is because

$$\begin{aligned} D_j^{-1}C_j'M_j^{-1}-D^{-1}C'M^{-1} &= D_j^{-1}C_j'\left(M_j^{-1}-M^{-1}\right)+D_j^{-1}\left(C_j'-C'\right)M^{-1} \\ &\quad+\left(D_j^{-1}-D^{-1}\right)C'M^{-1}, \end{aligned}$$

where the rank of $M_j^{-1}-M^{-1}$ is no larger than 4, and the ranks of $C_j'-C'$ and $D_j^{-1}-D^{-1}$ are no larger than 2 each. Therefore,

$$\begin{aligned} \left|\operatorname{tr}\left(D_j^{-1}C_j'M_j^{-1}-D^{-1}C'M^{-1}\right)\right| &\leq 8\left(\left\|D_j^{-1}C_j'\right\|\left\|M_j^{-1}\right\|+\left\|D^{-1}C'\right\|\left\|M^{-1}\right\|\right) \\ &\leq \frac{32\lambda_{\max}^{1/2}(\varepsilon\varepsilon')}{(\Im z)\,\lambda_{\min}^{3/2}(\varepsilon\varepsilon')}, \end{aligned}$$

where we used (62) and (63). Finally, $D_j^{-1}C_j'M_j^{-1}C_jD_j^{-1}$ differs from $D^{-1}C'M^{-1}CD^{-1}$ by a matrix with rank no larger than 12. Therefore,

$$\left|\operatorname{tr}\left(D_j^{-1}C_j'M_j^{-1}C_jD_j^{-1}-D^{-1}C'M^{-1}CD^{-1}\right)\right| \leq \frac{96\lambda_{\max}(\varepsilon\varepsilon')}{(\Im z)\,\lambda_{\min}^{2}(\varepsilon\varepsilon')}.$$

3.5 Proof of Lemma OW14

First, let us prove that

$$\hat{s} \overset{a.s.}{\to} 4c^2/\left(1-c^2\right). \tag{64}$$

Recall the definition of $\hat{s}$

$$\hat{s}=\frac{1}{T}\operatorname{tr}\left[D^{-1}\right]=\frac{1}{T}\operatorname{tr}\left[\left(\varepsilon\Delta_1\varepsilon'\right)^{-1}\right].$$

Let $F_D(x)$ denote the empirical distribution of the eigenvalues of D, and let

$$\hat{m}_D(z)=\int\frac{1}{x-z}\mathrm{d}F_D(x)$$

be its Stieltjes transform. Then, by Theorem 1.1 of Silverstein and Bai (1995), for any $z \in \mathbb{C}^+$, $\hat{m}_D(z)$ a.s. converges to $m_D(z)$, which satisfies equation

$$z = -\frac{1}{m_D(z)} + \int \frac{\tau \mathrm{d}H(\tau)}{1 + \tau c m_D(z)},$$

where $H(\tau)$ is the limit of the empirical distribution of the diagonal elements of Δ_1, r_j^{-1}, $j = 1, ..., T/2$. Recall that

$$r_{j+1} = 2(1 - \cos\theta j), \ \ j = 1, ..., T/2 - 1.$$

Therefore, $H(\tau)$ is the cumulative distribution function of the random variable $[2(1 - \cos U)]^{-1}$, where U is distributed uniformly on the interval $[0, \pi]$. This fact implies that

$$\begin{aligned} z &= -\frac{1}{m_D(z)} + \frac{1}{\pi}\int_0^\pi \frac{\mathrm{d}u}{2(1-\cos u) + c m_D(z)} \\ &= -\frac{1}{m_D(z)} + \frac{1}{2\pi \mathrm{i}} \oint\limits_{|s|=1} \frac{\mathrm{d}s}{s\left(2\left(1 - \frac{s+s^{-1}}{2}\right) + c m_D(z)\right)} \\ &= -\frac{1}{m_D(z)} - \frac{1}{2\pi \mathrm{i}} \oint\limits_{|s|=1} \frac{\mathrm{d}s}{\left(s^2 - (2 + c m_D(z))\, s + 1\right)}. \end{aligned}$$

The integrand has two poles at

$$s_{1,2} = \frac{c m_D(z) + 2 \pm \sqrt{c^2 m_D^2(z) + 4 c m_D(z)}}{2}.$$

Note that $s_1 s_2 = 1$, which implies that one of them is inside the contour and the other is outside. Therefore, we have

$$\begin{aligned} z &= -\frac{1}{m_D(z)} \pm \frac{1}{s_1 - s_2} \\ &= -\frac{1}{m_D(z)} \pm \frac{1}{\sqrt{c^2 m_D^2(z) + 4 c m_D(z)}} \end{aligned} \tag{65}$$

where the choice of $+$ or $-$ sign depends on which of $s_{1,2}$ is inside the contour. Squaring and rearranging, we obtain

$$c\,(z m_D(z) + 1)^2 \,(c m_D(z) + 4) - m_D(z) = 0. \tag{66}$$

Further, since $\min_{j=1,\ldots,T/2} r_j^{-1} \geq 1/4$, we have

$$\lambda_{\min}(D) = \lambda_{\min}(\varepsilon\Delta_1\varepsilon') \geq \frac{\lambda_{\min}(\varepsilon\varepsilon')}{4} \overset{a.s.}{\to} \frac{\left(1-\sqrt{c}\right)^2}{4}.$$

Therefore, $m_D(z)$ is analytic at $z = 0$, $\hat{m}_D(0) \overset{a.s.}{\to} m_D(0)$, and $m_D(0)$ satisfies equation (66) with $z = 0$. That is,

$$\hat{m}_D(0) \overset{a.s.}{\to} m_D(0) = \frac{4c}{1-c^2}.$$

But $\hat{s} = \frac{p}{T}\hat{m}_D(0)$. Hence, we have (64).

Now, let us turn to the proof of the lemma. Elementary algebra yields the following representation

$$\delta_j = \delta_{1j} + \frac{1}{z}\delta_{2j},$$

where

$$\delta_{1j} = (r_j + \hat{s})(z\hat{v} - 1),$$

and

$$\delta_{2j} = (z\hat{w})(1 + \hat{v} - z\hat{v}) - \hat{s}(z\hat{v}) + r_j(z\hat{u}) - \frac{1-z}{z}(z\hat{u})^2$$

Note that for $z \in \mathbb{C}^+$, $\hat{v} \in \mathbb{C}^+$. Hence, for $z \in \mathbb{C}^+$ such that $\Re z = 0$, we have $\Re(z\hat{v}) < 0$ and

$$|z\hat{v} - 1| > 1. \tag{67}$$

This inequality and (64) imply that, for any $z \in \mathbb{C}^+$ such that $\Re z = 0$,

$$|\delta_{1j}| > \frac{2c^2}{1-c^2}$$

for sufficiently large p, T as $p, T \to_c \infty$, almost surely.

Further, Lemma 5 implies that $|z\hat{u}|, |z\hat{v}|$, and $|z\hat{w}|$ remain bounded for sufficiently large p, T as $p, T \to_c \infty$, almost surely. Moreover, the presence of the imaginary part of z in the denominator of the bound on $\left\|M_j^{-1}\right\|$ in Lemma 5 imply that, for $z \in \mathbb{C}^+$ such that $\Re z = 0$, the value of the bound on $|z\hat{u}|, |z\hat{v}|$, and $|z\hat{w}|$ does not depend on z. In particular, for any such z, $|\delta_{2j}|$ is bounded for sufficiently large p, T as $p, T \to_c \infty$, almost surely, uniformly in j, with the value of the bound independent from z with $\Im z > \zeta$. Hence, by choosing ζ sufficiently large, we can ensure that, for any z with $\Re z = 0$ and $\Im z > \zeta$,

$$\left|\frac{1}{z}\delta_{2j}\right| < \frac{1}{2}\left|\delta_{1j}\right|,$$

and therefore

$$\left|\delta_j\right| > \frac{c^2}{1-c^2},$$

which establishes the lemma.

3.6 Proof of Lemma OW16

The identity $\hat{u} = \tilde{u}$ is established by the following sequence of equalities

$$\begin{aligned} T\hat{u} &= \operatorname{tr} D^{-1}C'M^{-1} = \operatorname{tr} D^{-1}C'\left(CD^{-1}C' - zA\right)^{-1} \\ &= \operatorname{tr}\left(C - zA\left(C'\right)^{-1}D\right)^{-1} = \operatorname{tr}\left(C' - zD\left(C\right)^{-1}A\right)^{-1} \\ &= \operatorname{tr} A^{-1}C\left(C'A^{-1}C - zD\right)^{-1} = \operatorname{tr} A^{-1}C\tilde{M}^{-1} = T\tilde{u}. \end{aligned}$$

The relationship $z\tilde{v} + \hat{s} = \hat{w}$ is obtained as follows

$$\begin{aligned} T\left(z\tilde{v} + \hat{s}\right) &= \operatorname{tr} D^{-1}\left(zI_p\left(C'A^{-1}CD^{-1} - zI_p\right)^{-1} + I_p\right) \\ &= \operatorname{tr} D^{-1}\left(-I_p + C'A^{-1}CD^{-1}\left(C'A^{-1}CD^{-1} - zI_p\right)^{-1} + I_p\right) \\ &= \operatorname{tr} D^{-1}\left(I_p - DC^{-1}A\left(C'\right)^{-1}z\right)^{-1} \\ &= \operatorname{tr} D^{-1}C'\left(CD^{-1}C' - Az\right)^{-1}CD^{-1} = T\hat{w}. \end{aligned}$$

The identity $z\hat{v} + \tilde{s} = \tilde{w}$ is obtained similarly to $z\tilde{v} + \hat{s} = \hat{w}$ by interchanging the roles of D, C and A, C'.

3.7 Derivation of Equation (OW55)

Equations (40), (41), and (47) imply that

$$\begin{aligned} D^{-1}C'M^{-1} &= \left(D_j^{-1} - D_j^{-1}\varepsilon_{(j)}\left(r_jI_2 + s_j\right)^{-1}\varepsilon_{(j)}'D_j^{-1}\right)\left(C_j' + \varepsilon_{(j)}\Delta_{2j}\varepsilon_{(j)}'\right) \\ &\quad \times\left(M_j^{-1} - M_j^{-1}\alpha_j\Omega_j\alpha_j'M_j^{-1}\right), \end{aligned}$$

where $\alpha_j = [\varepsilon_{(j)}, C_j D_j^{-1} \varepsilon_{(j)}]$ and

$$\Omega_j = \begin{pmatrix} \frac{1}{1-z} I_2 + v_j & \frac{1}{1-z} r_j \Delta'_{2j} + u'_j \\ \frac{1}{1-z} r_j \Delta_{2j} + u_j & \frac{z}{1-z} r_j I_2 - s_j + w_j \end{pmatrix}^{-1}.$$

Opening up brackets, we obtain

$$\begin{aligned} D^{-1}C'M^{-1} &= D_j^{-1}C'_jM_j^{-1} - D_j^{-1}\varepsilon_{(j)}\left(r_jI_2+s_j\right)^{-1}\varepsilon'_{(j)}D_j^{-1}C'_jM_j^{-1} \\ &+D_j^{-1}\varepsilon_{(j)}\Delta_{2j}\varepsilon'_{(j)}M_j^{-1} - D_j^{-1}C'_jM_j^{-1}\alpha_j\Omega_j\alpha'_jM_j^{-1} \\ &-D_j^{-1}\varepsilon_{(j)}\left(r_jI_2+s_j\right)^{-1}\varepsilon'_{(j)}D_j^{-1}\varepsilon_{(j)}\Delta_{2j}\varepsilon'_{(j)}M_j^{-1} \\ &+D_j^{-1}\varepsilon_{(j)}\left(r_jI_2+s_j\right)^{-1}\varepsilon'_{(j)}D_j^{-1}C'_jM_j^{-1}\alpha_j\Omega_j\alpha'_jM_j^{-1} \\ &-D_j^{-1}\varepsilon_{(j)}\Delta_{2j}\varepsilon'_{(j)}M_j^{-1}\alpha_j\Omega_j\alpha'_jM_j^{-1} \\ &+D_j^{-1}\varepsilon_{(j)}\left(r_jI_2+s_j\right)^{-1}\varepsilon'_{(j)}D_j^{-1}\varepsilon_{(j)}\Delta_{2j}\varepsilon'_{(j)}M_j^{-1}\alpha_j\Omega_j\alpha'_jM_j^{-1}. \end{aligned}$$

Multiplying from the left by $\varepsilon'_{(j)}$ and from the right by $\varepsilon_{(j)}$, and using the definitions of u_j, v_j, s_j, and w_j, we obtain

$$\begin{aligned} \varepsilon'_{(j)}D^{-1}C'M^{-1}\varepsilon_{(j)} &= u_j - s_j\left(r_jI_2+s_j\right)^{-1}u_j + s_j\Delta_{2j}v_j \\ &-[u_j,w_j]\,\Omega_j\left[v_j,u'_j\right]' - s_j\left(r_jI_2+s_j\right)^{-1}s_j\Delta_{2j}v_j \\ &+s_j\left(r_jI_2+s_j\right)^{-1}[u_j,w_j]\,\Omega_j\left[v_j,u'_j\right]' \\ &-s_j\Delta_{2j}\left[v_j,u'_j\right]\Omega_j\left[v_j,u'_j\right]' \\ &+s_j\left(r_jI_2+s_j\right)^{-1}s_j\Delta_{2j}\left[v_j,u'_j\right]\Omega_j\left[v_j,u'_j\right]'. \end{aligned}$$

Rearranging terms and simplifying gives us

$$\begin{aligned} \varepsilon'_{(j)}D^{-1}C'M^{-1}\varepsilon_{(j)} &= r_j\left(r_jI_2+s_j\right)^{-1}s_j\Delta_{2j}\left(v_j-\left[v_j,u'_j\right]\Omega_j\left[v_j,u'_j\right]'\right) \quad (68) \\ &+r_j\left(r_jI_2+s_j\right)^{-1}\left(u_j-[u_j,w_j]\,\Omega_j\left[v_j,u'_j\right]'\right). \end{aligned}$$

3.8 Proof of Proposition OW17

Multiplying both sides of (68) by Δ'_{2j}, we obtain

$$\begin{aligned} \Delta'_{2j}\varepsilon'_{(j)}D^{-1}C'M^{-1}\varepsilon_{(j)} &= \Delta'_{2j}r_j\left(r_jI_2+s_j\right)^{-1}s_j\Delta_{2j}\left(v_j-\left[v_j,u'_j\right]\Omega_j\left[v_j,u'_j\right]'\right) \\ &+\Delta'_{2j}r_j\left(r_jI_2+s_j\right)^{-1}\left(u_j-[u_j,w_j]\,\Omega_j\left[v_j,u'_j\right]'\right). \quad (69) \end{aligned}$$

Note that, for $j = 1, ..., T/2 - 1$,

$$\Delta_{2,j+1} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sin j\theta}{2-2\cos j\theta} \\ \frac{\sin j\theta}{2-2\cos j\theta} & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -\frac{\cos j\theta/2}{2\sin j\theta/2} \\ \frac{\cos j\theta/2}{2\sin j\theta/2} & -\frac{1}{2} \end{pmatrix}$$

so that

$$r_{j+1}^{1/2}\Delta_{2,j+1} = \begin{pmatrix} -\sin j\theta/2 & -\cos j\theta/2 \\ \cos j\theta/2 & -\sin j\theta/2 \end{pmatrix},$$

and thus, $\left\| r_{j+1}^{1/2}\Delta_{2,j+1} \right\|$ is bounded uniformly in $j = 1, ..., T/2 - 1$. For $r_1^{1/2}\Delta_1$, we have

$$r_1^{1/2}\Delta_1 = -I_2.$$

Hence, $\left\| r_j^{1/2}\Delta_{2,j} \right\|$ is bounded uniformly in $j = 1, ..., T/2$.

Replacing s_j, u_j, v_j, and w_j in (69) by $\hat{s}I_2, \hat{u}I_2, \hat{v}I_2$, and $\hat{w}I_2$, respectively, and noting that

$$\Delta_{2j}'\Delta_{2j} = \Delta_{1j} = r_j^{-1}I_2,$$

we obtain

$$\begin{aligned} & \Delta_{2j}'\varepsilon_{(j)}'D^{-1}C'M^{-1}\varepsilon_{(j)} \\ = \; & \frac{\hat{s}}{r_j + \hat{s}}\left(\hat{v}I_2 - [\hat{v}I_2, \hat{u}I_2]\,\hat{\Omega}_j\,[\hat{v}I_2, \hat{u}I_2]'\right) \\ & + \frac{r_j}{r_j + \hat{s}}\Delta_{2j}'\left(\hat{u}I_2 - [\hat{u}I_2, \hat{w}I_2]\,\hat{\Omega}_j\,[\hat{v}I_2, \hat{u}I_2]'\right) + o(1), \end{aligned} \tag{70}$$

where,

$$\hat{\Omega}_j = \frac{1-z}{\delta_j}\begin{pmatrix} \frac{z}{1-z}r_jI_2 - \hat{s}I_2 + \hat{w}I_2 & -\frac{1}{1-z}r_j\Delta_{2j}' - \hat{u}I_2 \\ -\frac{1}{1-z}r_j\Delta_{2j} - \hat{u}I_2 & \frac{1}{1-z}I_2 + \hat{v}I_2 \end{pmatrix},$$

$$\delta_j = (\hat{w} - \hat{s})(1 + \hat{v} - z\hat{v}) + r_j(\hat{u} + z\hat{v} - 1) - (1-z)\hat{u}^2,$$

and $o(1) \overset{a.s.}{\to} 0$, uniformly in j. The latter convergence follows from Lemma OW12 and the fact that the right hand side of (69) a.s. remains continuous function of s_j, u_j, v_j and w_j with bounded derivatives as $p, T \to_c \infty$. It can be shown using the facts that: (i) $\left\| r_j^{1/2}\Delta_{2,j} \right\|$ is bounded uniformly in $j = 1, ..., T/2$, as show above, (ii) $|r_j + \hat{s}|^{-1}$ remains bounded as $p, T \to_c \infty$, a.s. and uniformly in j, as follows from (64), and (iii) $\left\| \hat{\Omega}_j \right\|$ remains bounded as $p, T \to_c \infty$, a.s. and uniformly in j, as shown in the proof of Proposition OW13.

Consider the first term on the right hand side of (70). Opening brackets and simplifying, we obtain

$$\begin{aligned} & \frac{\hat{s}}{r_j+\hat{s}}\left(\hat{v}I_2-[\hat{v}I_2,\hat{u}I_2]\,\hat{\Omega}_j\,[\hat{v}I_2,\hat{u}I_2]'\right) \\ = \quad & \frac{\hat{s}}{\delta_j\,(r_j+\hat{s})}\left(\hat{v}\,(\hat{w}-\hat{s})\,I_2+r_j\hat{v}\,(\hat{u}-1)\,I_2+r_j\hat{u}\hat{v}\Delta_{2j}+r_j\hat{v}\hat{u}\Delta_{2j}'-\hat{u}^2I_2\right). \end{aligned}$$

Using the identity

$$\Delta_{2j}+\Delta_{2j}'=-I_2,$$

we further simply the above expression to get

$$\frac{\hat{s}}{r_j+\hat{s}}\left(\hat{v}I_2-[\hat{v}I_2,\hat{u}I_2]\,\hat{\Omega}_j\,[\hat{v}I_2,\hat{u}I_2]'\right)=-\frac{\hat{s}\hat{v}}{\delta_j}I_2+\frac{\hat{s}\,(\hat{v}\hat{w}-\hat{u}^2)}{\delta_j\,(r_j+\hat{s})}I_2.$$

Similarly, for the second term on the right hand side of (70), we obtain

$$\frac{r_j}{r_j+\hat{s}}\Delta_{2j}'\left(\hat{u}I_2-[\hat{u}I_2,\hat{w}I_2]\,\hat{\Omega}_j\,[\hat{v}I_2,\hat{u}I_2]'\right)=-\frac{r_j\hat{u}}{\delta_j}\Delta_{2j}'+\frac{r_j\,(\hat{v}\hat{w}-\hat{u}^2)}{\delta_j\,(r_j+\hat{s})}I_2.$$

Summing up the latter two equations, and taking trace, we get

$$\operatorname{tr}\left[\Delta_{2j}'\varepsilon_{(j)}'D^{-1}C'M^{-1}\varepsilon_{(j)}\right]=\frac{r_j\hat{u}-2\hat{s}\hat{v}+2\hat{v}\hat{w}-2\hat{u}^2}{\delta_j}+o(1).$$

Using this equation together with equations (OW53) and (OW54), we obtain

$$1+z\hat{m}=\frac{1}{p}\sum_{j=1}^{T/2}\frac{r_j\hat{u}-2\hat{s}\hat{v}+2\hat{v}\hat{w}-2\hat{u}^2}{\delta_j}+o(1).$$

Recall that $r_{j+1}=2-2\cos j\theta=4\sin^2(j\pi/T)$ and $\delta_{j+1}^{-1}=\hat{h}\,(z,j\pi/T)$. Therefore,

$$1+z\hat{m}=\frac{1}{p}\sum_{j=1}^{T/2-1}\hat{h}\,(z,j\pi/T)\left(4\sin^2(j\pi/T)\,\hat{u}-2\hat{s}\hat{v}+2\hat{v}\hat{w}-2\hat{u}^2\right)+o(1).$$

Since, for z with $\Re z=0$ and $\Im z>\zeta$, the derivative

$$\frac{\mathrm{d}}{\mathrm{d}\varphi}\left[\hat{h}\,(z,\varphi)\left(4\sin^2\varphi\hat{u}-2\hat{s}\hat{v}+2\hat{v}\hat{w}-2\hat{u}^2\right)\right]$$

a.s. remains bounded as $p, T \to_c \infty$, we obtain

$$\begin{aligned} 1 + z\hat{m} &= \frac{1}{\pi c} \int_0^{\pi/2} \hat{h}\,(z, \varphi) \left(4 \sin^2 \varphi \hat{u} - 2\hat{s}\hat{v} + 2\hat{v}\hat{w} - 2\hat{u}^2\right) \mathrm{d}\varphi + o(1) \\ &= \frac{1}{2\pi c} \int_0^{2\pi} \hat{h}\,(z, \varphi) \left(2 \sin^2 \varphi \hat{u} - \hat{s}\hat{v} + \hat{v}\hat{w} - \hat{u}^2\right) \mathrm{d}\varphi + o(1). \end{aligned}$$

Finally, using the identity

$$z\tilde{v} + \hat{s} = \hat{w},$$

we obtain

$$1 + z\hat{m} = \frac{1}{2\pi c} \int_0^{2\pi} \hat{h}\,(z, \varphi) \left(2 \sin^2 \varphi \hat{u} + z\tilde{v}\hat{v} - \hat{u}^2\right) \mathrm{d}\varphi + o(1).$$

3.9 Proof of Proposition OW18

Consider the identity

$$\frac{1}{p} \sum_{j=1}^{T/2} \operatorname{tr}\left[\Delta_{2j} \varepsilon'_{(j)} M^{-1} \varepsilon_{(j)}\right] = \frac{1}{p} \sum_{j=1}^{T/2} \operatorname{tr}\left[\Delta_{1j} \varepsilon'_{(j)} D^{-1} C' M^{-1} \varepsilon_{(j)}\right]. \tag{71}$$

Multiplying both sides of (68) by Δ_{1j}, we obtain

$$\begin{aligned} \Delta_{1j} \varepsilon'_{(j)} D^{-1} C' M^{-1} \varepsilon_{(j)} &= \left(r_j I_2 + s_j\right)^{-1} s_j \Delta_{2j} \left(v_j - \left[v_j, u'_j\right] \Omega_j \left[v_j, u'_j\right]'\right) \\ &\quad + \left(r_j I_2 + s_j\right)^{-1} \left(u_j - \left[u_j, w_j\right] \Omega_j \left[v_j, u'_j\right]'\right). \end{aligned} \tag{72}$$

Further, equation (OW43) gives us

$$\Delta_{2j} \varepsilon'_{(j)} M^{-1} \varepsilon_{(j)} = \Delta_{2j} \left(v_j - [v_j, u'_j] \Omega_j [v_j, u'_j]'\right). \tag{73}$$

Subtracting (73) from (72), we obtain

$$\begin{aligned} &\Delta_{1j} \varepsilon'_{(j)} D^{-1} C' M^{-1} \varepsilon_{(j)} - \Delta_{2j} \varepsilon'_{(j)} M^{-1} \varepsilon_{(j)} \\ =\ &\left(\left(r_j I_2 + s_j\right)^{-1} s_j - I_2\right) \Delta_{2j} \left(v_j - \left[v_j, u'_j\right] \Omega_j \left[v_j, u'_j\right]'\right) \\ &+ \left(r_j I_2 + s_j\right)^{-1} \left(u_j - \left[u_j, w_j\right] \Omega_j \left[v_j, u'_j\right]'\right) \\ =\ &-r_j \left(r_j I_2 + s_j\right)^{-1} \Delta_{2j} \left(v_j - \left[v_j, u'_j\right] \Omega_j \left[v_j, u'_j\right]'\right) \\ &+ \left(r_j I_2 + s_j\right)^{-1} \left(u_j - \left[u_j, w_j\right] \Omega_j \left[v_j, u'_j\right]'\right). \end{aligned}$$

Proceeding as in the above analysis of (69), we obtain

$$\begin{aligned}
& \Delta_{1j}\varepsilon'_{(j)}D^{-1}C'M^{-1}\varepsilon_{(j)} - \Delta_{2j}\varepsilon'_{(j)}M^{-1}\varepsilon_{(j)} \\
= & -\frac{r_j}{\delta_j\left(r_j+\hat{s}\right)}\Delta_{2j}\left(-\hat{v}\hat{s}I_2 + r_j\hat{v}\left(\hat{u}-1\right)I_2 + r_j\hat{u}\hat{v}\Delta_{2j} + r_j\hat{v}\hat{u}\Delta'_{2j}\right) \\
& -\frac{\hat{u}}{\delta_j}I_2 + o(1).
\end{aligned}$$

Taking trace, summing over j, dividing by p, and using (71), we obtain

$$\frac{1}{p}\sum_{j=1}^{T/2}\frac{r_j\hat{v}+2\hat{u}}{\delta_j} = o(1).$$

Replacing the sum by an integral yields

$$\frac{1}{2\pi c}\int_0^{2\pi}\hat{h}\left(z,\varphi\right)\left(2\hat{v}\sin^2\varphi + \hat{u}\right)\mathrm{d}\varphi = o(1).$$

3.10 Derivation of Equation (OW71)

Consider

$$\mathcal{I} = \frac{1}{2\pi}\int_0^{2\pi}\frac{1}{x+2\sin^2\varphi}\mathrm{d}\varphi,$$

where $x \in \mathbb{C}\backslash\left[-2,0\right]$. Changing the variable of integration to $z = \exp\left\{\mathrm{i}\varphi\right\}$, we obtain

$$\begin{aligned}
\mathcal{I} &= \frac{1}{2\pi\mathrm{i}}\oint_{|z|=1}\frac{1}{x-\left(z-z^{-1}\right)^2/2}\frac{\mathrm{d}z}{z} \\
&= -\frac{1}{2\pi\mathrm{i}}\oint_{|z|=1}\frac{2z}{\left(z^2-x_1\right)\left(z^2-x_2\right)}\mathrm{d}z,
\end{aligned}$$

where

$$x_{1,2} = x+1\pm\sqrt{x\left(x+2\right)}.$$

Since $x_1x_2 = 1$, whereas $|x_1| \neq 1$ and $|x_2| \neq 1$, there are only two poles of the integrand that are situated inside the unit circle. They are either $x_1^{1/2}, -x_1^{1/2}$, which we shall call case 1, or $x_2^{1/2}, -x_2^{1/2}$, which we shall call case 2. By Cauchy's residue theorem,

$$\mathcal{I} = \mp\frac{2}{x_1-x_2},$$

with "$-$" corresponding to case 1 and "$+$" corresponding to case 2. Whatever the case, we have

$$\mathcal{I}^2 = \frac{4}{(x_1 - x_2)^2} = \frac{1}{x(x+2)}.$$

References

[1] Bai, Z. D. and J. W. Silverstein, (2010) *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. Springer Verlag, New York.

[2] Bai, Z. D. and Yin, Y. Q. (1993) "Limit of the smallest eigenvalue of large dimensional covariance matrix," *Annals of Probability* 21, 1275–1294.

[3] Davidson, K.R. and S.J. Szarek (2001) Local operator theory, random matrices and Banach spaces, in: W.B. Johnson and J. Lindenstrauss (Eds.), Handbook of the Geometry of Banach Spaces, Vol. I, North-Holland, Amsterdam, 2001, 317–366.

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

[5] Horn, R.A. and C.R. Johnson (1985) *Matrix Analysis*, Cambridge University press.

[6] Johansen, S. (1988) "Statistical Analysis of Cointegrating Vectors," *Journal of Economic Dynamics and Control* 12, 231-254

[7] Johansen, S. (1991) "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica* 59, 1551-1580.

[8] Johansen, S. (1995) *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press.

[9] Onatski, A. and C. Wang (2016) "Alternative asymptotics for cointegration tests in large VARs," manuscript, University of Cambridge.

[10] Silverstein, J.W. (1985) "The smallest eigenvalue of a large dimensional Wishart matrix," *Annals of Probability* 13, 1364-1368.

[11] Silverstein, J.W. and Z.D. Bai (1995) "On the Empirical Distribution of Eigenvalues of a Class of Large Dimensional Random Matrices," *Journal of Multivariate Analysis* 54, 175-192.

[12] Silverstein, J.W. and Choi, S. I. (1995) "Analysis of the limiting spectral distribution of large dimensional random matrices," *Journal of Multivariate Analysis* 54, 295–309.