Supplementary Material

For notational simplicity, the proofs below assume that in the case $p < n$, the support of F is a single compact interval $[a, b] \subset (0, +\infty)$. But they generalize easily to the case where $\mathsf{Supp}(F)$ is the union of a finite number κ of such intervals, as maintained in Assumptions 3.2 and 3.4. On the same grounds, we make a similar assumption on the support of $\underline{F}$ in the case $p > n$; see Section 6.

When there is no ambiguity, the first subscript, n, can be dropped from the notation of the eigenvalues and eigenvectors.

11 Proofs of Mathematical Results in Section 3.2

11.1 Proof of Theorem 3.1

Definition 11.1. *For any integer k, define $\forall x \in \mathbb{R}$, $\Delta_n^{(k)}(x) := p^{-1} \sum_{i=1}^{p} u_i' \Sigma_n^k u_i \times \mathbb{1}_{[\lambda_i, +\infty)}(x)$.*

Lemma 11.1. *Under Assumptions 3.1–3.3, there exists a nonrandom function $\Delta^{(-1)}$ defined on $\mathbb{R}$ such that $\Delta_n^{(-1)}(x)$ converges almost surely to $\Delta^{(-1)}(x)$, for all $x \in \mathbb{R}$. Furthermore, $\Delta^{(-1)}$ is continuously differentiable on $\mathbb{R}$ and satisfies $\forall x \in \mathbb{R}$, $\Delta^{(-1)}(x) = \int_{-\infty}^{x} \delta^{(-1)}(\lambda) dF(\lambda)$, where*

$$\forall \lambda \in \mathbb{R} \qquad \delta^{(-1)}(\lambda) := \begin{cases} 0 & \text{if } \lambda \leq 0, \\ \dfrac{1 - c - 2\,c\,\lambda\,\mathsf{Re}[\breve{m}_F(\lambda)]}{\lambda} & \text{if } \lambda > 0\ . \end{cases}$$

Proof of Lemma 11.1. The proof of Lemma 11.1 follows directly from Ledoit and Péché (2011, Theorem 5) and the corresponding proof, bearing in mind that we are in the case $c < 1$ because of Assumption 3.1. ■

Lemma 11.2. *Under Assumptions 3.1–3.4,*

$$\frac{1}{p}\mathsf{Tr}\big(\Sigma_n^{-1}\widetilde{S}_n\big) \xrightarrow{\text{a.s.}} \int_a^b \widetilde{\varphi}(x)\, d\Delta^{(-1)}(x)\ .$$

Proof of Lemma 11.2. Restrict attention to the set Ω_1 of probability one on which $\Delta_n^{(-1)}(x)$ converges to $\Delta^{(-1)}(x)$, for all x, and on which also the almost sure uniform convergence and the uniform boundedness of Assumption 3.4 hold for all rational, small $\eta > 0$. Wherever necessary, the results in the proof are understood to hold true on this set Ω_1.

Note that

$$\frac{1}{p}\mathsf{Tr}\big(\Sigma_n^{-1}\widetilde{S}_n\big) = \frac{1}{p}\sum_{i=1}^{p}\big(u_i'\Sigma_n^{-1}u_i\big)\widetilde{\varphi}_n(\lambda_i) = \int \widetilde{\varphi}_n(x)\, d\Delta_n^{(-1)}(x)\ . \tag{11.1}$$

Since $\widetilde{\varphi}$ is continuous and $\Delta_n^{(-1)}$ converges weakly to $\Delta^{(-1)}$,

$$\int_a^b \widetilde{\varphi}(x)\, d\Delta_n^{(-1)}(x) \longrightarrow \int_a^b \widetilde{\varphi}(x)\, d\Delta^{(-1)}(x)\ . \tag{11.2}$$

Since $|\widetilde{\varphi}|$ is continuous on $[a,b]$, it is bounded above by a finite constant $\widetilde{K}_1$. Fix $\varepsilon > 0$. Since $\Delta^{(-1)}$ is continuous, there exists a rational $\eta_1 > 0$ such that

$$\left|\Delta^{(-1)}(a+\eta_1) - \Delta^{(-1)}(a)\right| + \left|\Delta^{(-1)}(b) - \Delta^{(-1)}(b-\eta_1)\right| \leq \frac{\varepsilon}{6\,\widetilde{K_1}} \ . \tag{11.3}$$

Since $\Delta_n^{(-1)}(x) \longrightarrow \Delta^{(-1)}(x)$, for all $x \in \mathbb{R}$, there exists $N_1 \in \mathbb{N}$ such that

$$\forall n \geq N_1 \qquad \max_{x \in \{a, a+\eta_1, b-\eta_1, b\}} \left|\Delta_n^{(-1)}(x) - \Delta^{(-1)}(x)\right| \leq \frac{\varepsilon}{24\,\widetilde{K_1}} \ . \tag{11.4}$$

Putting Equations (11.3)–(11.4) together yields

$$\forall n \geq N_1 \qquad \left|\Delta_n^{(-1)}(a+\eta_1) - \Delta_n^{(-1)}(a)\right| + \left|\Delta_n^{(-1)}(b) - \Delta_n^{(-1)}(b-\eta_1)\right| \leq \frac{\varepsilon}{3\,\widetilde{K_1}} \ . \tag{11.5}$$

Therefore, for all $n \geq N_1$,

$$\begin{aligned} &\left|\int_{a+\eta_1}^{b-\eta_1} \widetilde{\varphi}(x)\, d\Delta_n^{(-1)}(x) - \int_a^b \widetilde{\varphi}(x)\, d\Delta_n^{(-1)}(x)\right| \\ &\qquad \leq \widetilde{K}_1 \Big[\left|\Delta_n^{(-1)}(a+\eta_1) - \Delta_n^{(-1)}(a)\right| + \left|\Delta_n^{(-1)}(b) - \Delta_n^{(-1)}(b-\eta_1)\right|\Big] \\ &\qquad \leq \frac{\varepsilon}{3} \ . \end{aligned} \tag{11.6}$$

Since $\widetilde{\varphi}_n(x) \longrightarrow \widetilde{\varphi}(x)$ uniformly over $x \in [a+\eta_1, b-\eta_1]$, there exists $N_2 \in \mathbb{N}$ such that

$$\forall n \geq N_2 \qquad \forall x \in [a+\eta_1, b-\eta_1] \qquad |\widetilde{\varphi}_n(x) - \widetilde{\varphi}(x)| \leq \frac{\varepsilon\, \underline{h}}{3} \ .$$

By Assumption 3.2, there exists $N_3 \in \mathbb{N}$ such that, for all $n \geq N_3$, $\max_{x\in\mathbb{R}} |\Delta_n^{(-1)}(x)| = \mathsf{Tr}(\Sigma_n^{-1})/p$ is bounded by $1/\underline{h}$. Therefore for all $n \geq \max(N_2, N_3)$

$$\left|\int_{a+\eta_1}^{b-\eta_1} \widetilde{\varphi}_n(x)\, d\Delta_n^{(-1)}(x) - \int_{a+\eta_1}^{b-\eta_1} \widetilde{\varphi}(x)\, d\Delta_n^{(-1)}(x)\right| \leq \frac{\varepsilon\, \underline{h}}{3} \times \frac{1}{\underline{h}} = \frac{\varepsilon}{3} \ . \tag{11.7}$$

Arguments analogous to those justifying Equations (11.3)–(11.5) show there exists $N_4 \in \mathbb{N}$ such that

$$\forall n \geq N_4 \qquad \left|\Delta_n^{(-1)}(a+\eta_1) - \Delta_n^{(-1)}(a-\eta_1)\right| + \left|\Delta_n^{(-1)}(b+\eta_1) - \Delta_n^{(-1)}(b-\eta_1)\right| \leq \frac{\varepsilon}{3\,\widetilde{K}} \ ,$$

for the finite constant $\widetilde{K}$ of Assumption 3.4. Therefore, for all $n \geq N_4$,

$$\left|\int_{a-\eta_1}^{b+\eta_1} \widetilde{\varphi}_n(x)\, d\Delta_n^{(-1)}(x) - \int_{a+\eta_1}^{b-\eta_1} \widetilde{\varphi}_n(x)\, d\Delta_n^{(-1)}(x)\right| \leq \frac{\varepsilon}{3} \ . \tag{11.8}$$

Putting together Equations (11.6)–(11.8) implies that, for all $n \geq \max(N_1, N_2, N_3, N_4)$,

$$\left|\int_{a-\eta_1}^{b+\eta_1} \widetilde{\varphi}_n(x)\, d\Delta_n^{(-1)}(x) - \int_a^b \widetilde{\varphi}(x)\, d\Delta_n^{(-1)}(x)\right| \leq \varepsilon \ .$$

Since ε can be chosen arbitrarily small,

$$\int_{a-\eta_1}^{b+\eta_1} \widetilde{\varphi}_n(x)\, d\Delta_n^{(-1)}(x) - \int_a^b \widetilde{\varphi}(x)\, d\Delta_n^{(-1)}(x) \longrightarrow 0 \ .$$

By using Equation (11.2) we get

$$\int_{a-\eta_1}^{b+\eta_1} \widetilde{\varphi}_n(x)\, d\Delta_n^{(-1)}(x) \longrightarrow \int_a^b \widetilde{\varphi}(x)\, d\Delta^{(-1)}(x) \ .$$

Theorem 1.1 of Bai and Silverstein (1998) shows that on a set Ω_2 of probability one, there are no sample eigenvalues outside the interval $[a-\eta_1, a+\eta_1]$, for all n large enough. Therefore, on the set $\Omega := \Omega_1 \cap \Omega_2$ of probability one,

$$\int \widetilde{\varphi}_n(x)\, d\Delta_n^{(-1)}(x) \longrightarrow \int_a^b \widetilde{\varphi}(x)\, d\Delta^{(-1)}(x) \ .$$

Together with Equation (11.1), this proves Lemma 11.2. ■

Lemma 11.3.

$$\frac{1}{p} \log \left[\det\left(\Sigma_n^{-1} \widetilde{S}_n\right)\right] \xrightarrow{\text{a.s.}} \int_a^b \log\left[\widetilde{\varphi}(x)\right] dF(x) - \int \log(t)\, dH(t) \ .$$

Proof of Lemma 11.3.

$$\begin{aligned}
\frac{1}{p} \log \left[\det\left(\Sigma_n^{-1} \widetilde{S}_n\right)\right] &= \frac{1}{p} \log \left[\det\left(\Sigma_n^{-1}\right) \det\left(\widetilde{S}_n\right)\right] \\
&= \frac{1}{p} \log \left[\det\left(\Sigma_n^{-1}\right) \prod_{i=1}^{p} \widetilde{\varphi}_n(\lambda_i)\right] \\
&= \int \log\left[\widetilde{\varphi}_n(x)\right] dF_n(x) - \int \log(t)\, dH_n(t) \ . \qquad (11.9)
\end{aligned}$$

A reasoning analogous to that conducted in the proof of Lemma 11.2 shows that the first term on the right-hand side of Equation (11.9) converges almost surely to $\int_a^b \log\left[\widetilde{\varphi}(x)\right] dF(x)$. Given that H_n converges weakly to H, Lemma 11.3 follows. ■

We are now ready to tackle Theorem 3.1. Lemma 11.1 and Lemma 11.2 imply that

$$\frac{1}{p} \mathsf{Tr}\left(\Sigma_n^{-1} \widetilde{S}_n\right) \xrightarrow{\text{a.s.}} \int_a^b \widetilde{\varphi}(x)\, \frac{1 - c - 2\,c\,x\,\mathsf{Re}[\breve{m}_F(x)]}{x}\, dF(x) \ .$$

Lemma 11.3 implies that

$$-\frac{1}{p} \log \left[\det\left(\Sigma_n^{-1} \widetilde{S}_n\right)\right] - 1 \xrightarrow{\text{a.s.}} \int \log(t)\, dH(t) - \int_a^b \log\left[\widetilde{\varphi}(x)\right] dF(x) - 1 \ .$$

Putting these two results together completes the proof of Theorem 3.1. ■

11.2 Proof of Proposition 3.1

We start with the simpler case where $\forall n \in \mathbb{N}$, $\forall x \in \mathbb{R}$, $\widetilde{\psi}_n(x) \equiv \widetilde{\psi}(x)$. We make implicitly use of Theorem 1.1 of Bai and Silverstein (1998), which states that, for any fixed $\eta > 0$, there are no eigenvalues outside the interval $[a-\eta, b+\eta]$ with probability one, for all n large enough.

For any given estimator $\widetilde{S}_n$ with limiting shrinkage function $\widetilde{\varphi}$, define the univariate function $\forall x, y \in [a, b],\ \widetilde{\psi}(x) := \widetilde{\varphi}(x)/x$ and the bivariate function

$$\forall x, y \in [a, b] \qquad \widetilde{\psi}^\sharp(x, y) := \begin{cases} \dfrac{x\widetilde{\psi}(x) - y\widetilde{\psi}(y)}{x - y} & \text{if} \quad x \neq y \\ x\widetilde{\psi}'(x) + \widetilde{\psi}(x) & \text{if} \quad x = y \ . \end{cases}$$

Since $\widetilde{\psi}$ is continuously differentiable on $[a, b]$, $\widetilde{\psi}^\sharp$ is continuous on $[a, b] \times [a, b]$. Consequently, there exists $K > 0$ such that, $\forall x, y \in [a, b],\ |\widetilde{\psi}^\sharp(x, y)| \leq K$.

Lemma 11.4.

$$\frac{2}{p^2} \sum_{j=1}^{p} \sum_{i>j} \frac{\lambda_j \widetilde{\psi}(\lambda_j) - \lambda_i \widetilde{\psi}(\lambda_i)}{\lambda_j - \lambda_i} \xrightarrow{\text{a.s.}} \int_a^b \int_a^b \widetilde{\psi}^\sharp(x, y)\, dF(x)\, dF(y) \ . \tag{11.10}$$

Proof of Lemma 11.4.

$$\begin{aligned} \frac{2}{p^2} \sum_{j=1}^{p} \sum_{i>j} \frac{\lambda_j \widetilde{\psi}(\lambda_j) - \lambda_i \widetilde{\psi}(\lambda_i)}{\lambda_j - \lambda_i} &= \frac{1}{p^2} \sum_{j=1}^{p} \sum_{i=1}^{p} \widetilde{\psi}^\sharp(\lambda_i, \lambda_j) - \frac{1}{p^2} \sum_{j=1}^{p} \widetilde{\psi}^\sharp(\lambda_j, \lambda_j) \\ &= \int_a^b \int_a^b \widetilde{\psi}^\sharp(x, y)\, dF_n(x)\, dF_n(y) - \frac{1}{p^2} \sum_{j=1}^{p} \widetilde{\psi}^\sharp(\lambda_j, \lambda_j) \ . \end{aligned}$$

Given Equation (3.1), the first term converges almost surely to the right-hand side of Equation (11.10). The absolute value of the second term is bounded by K/p; therefore, it vanishes asymptotically. ■

Lemma 11.5.

$$\int_a^b \int_a^b \widetilde{\psi}^\sharp(x, y)\, dF(x)\, dF(y) = -2 \int_a^b x\widetilde{\psi}(x)\, \mathsf{Re}\left[\breve{m}_F(x)\right]\, dF(x) \ . \tag{11.11}$$

Proof of Lemma 11.5. Fix any $\varepsilon > 0$. Then there exists $\eta_1 > 0$ such that, for all $v \in (0, \eta_1)$,

$$\left| 2 \int_a^b x\widetilde{\psi}(x)\, \mathsf{Re}\left[\breve{m}_F(x)\right]\, dF(x) - 2 \int_a^b x\widetilde{\psi}(x)\, \mathsf{Re}\left[\breve{m}_F(x + iv)\right]\, dF(x) \right| \leq \frac{\varepsilon}{4} \ .$$

The definition of the Stieltjes transform implies

$$-2 \int_a^b x\widetilde{\psi}(x)\, \mathsf{Re}\left[\breve{m}_F(x + iv)\right]\, dF(x) = 2 \int_a^b \int_a^b \frac{x\widetilde{\psi}(x)(x - y)}{(x - y)^2 + v^2}\, dF(x)\, dF(y) \ .$$

There exists $\eta_2 > 0$ such that, for all $v \in (0, \eta_1)$,

$$\left| 2 \int_a^b \int_a^b \frac{x\widetilde{\psi}(x)(x - y)}{(x - y)^2 + v^2} dF(x) dF(y) - 2 \int_a^b \int_a^b \frac{x\widetilde{\psi}(x)(x - y)}{(x - y)^2 + v^2} \mathbb{1}_{\{|x-y| \geq \eta_2\}} dF(x) dF(y) \right| \leq \frac{\varepsilon}{4}$$

$$\text{and} \qquad \left| \int_a^b \int_a^b \widetilde{\psi}^\sharp(x, y)\, dF(x)\, dF(y) - \int_a^b \int_a^b \widetilde{\psi}^\sharp(x, y) \mathbb{1}_{\{|x-y| \geq \eta_2\}}\, dF(x)\, dF(y) \right| \leq \frac{\varepsilon}{4} \ .$$

We have

$$\int_a^b \int_a^b \widetilde{\psi}^\sharp(x,y) \mathbb{1}_{\{|x-y|\geq\eta_2\}}\, dF(x)\, dF(y) = \int_a^b \int_a^b \frac{x\widetilde{\psi}(x) - y\widetilde{\psi}(y)}{x-y} \mathbb{1}_{\{|x-y|\geq\eta_2\}}\, dF(x)\, dF(y)$$
$$= \int_a^b \int_a^b \frac{x\widetilde{\psi}(x)}{x-y} \mathbb{1}_{\{|x-y|\geq\eta_2\}}\, dF(x)\, dF(y)$$
$$+ \int_a^b \int_a^b \frac{y\widetilde{\psi}(y)}{y-x} \mathbb{1}_{\{|y-x|\geq\eta_2\}}\, dF(y)\, dF(x)$$
$$= 2\int_a^b \int_a^b \frac{x\widetilde{\psi}(x)}{x-y} \mathbb{1}_{\{|x-y|\geq\eta_2\}}\, dF(x)\, dF(y)\ .$$

Note that

$$2\int_a^b \int_a^b \frac{x\widetilde{\psi}(x)}{x-y} \mathbb{1}_{\{|x-y|\geq\eta_2\}}\, dF(x)\, dF(y) - 2\int_a^b \int_a^b \frac{x\widetilde{\psi}(x)(x-y)}{(x-y)^2+v^2} \mathbb{1}_{\{|x-y|\geq\eta_2\}}\, dF(x)\, dF(y)$$
$$= 2\int_a^b \int_a^b \frac{x\widetilde{\psi}(x)}{x-y} \frac{v^2}{(x-y)^2+v^2} \mathbb{1}_{\{|x-y|\geq\eta_2\}}\, dF(x)\, dF(y)\ ,$$

and that

$$\forall (x,y) \text{ such that } |x-y| \geq \eta_2 \qquad \frac{v^2}{(x-y)^2+v^2} \leq \frac{v^2}{\eta_2^2+v^2}\ .$$

The quantity on the right-hand side can be made arbitrarily small for fixed η_2 by bringing v sufficiently close to zero. This implies that there exists $\eta_3 \in (0,\eta_1)$ such that, for all $v \in (0,\eta_3)$,

$$\left| 2\int_a^b \int_a^b \frac{x\widetilde{\psi}(x)}{x-y} \mathbb{1}_{\{|x-y|\geq\eta_2\}}\, dF(x)\, dF(y) - 2\int_a^b \int_a^b \frac{x\widetilde{\psi}(x)(x-y)}{(x-y)^2+v^2} \mathbb{1}_{\{|x-y|\geq\eta_2\}}\, dF(x)\, dF(y) \right| \leq \frac{\varepsilon}{4}\ .$$

Putting these results together yields

$$\left| \int_a^b \int_a^b \widetilde{\psi}^\sharp(x,y)\, dF(x)\, dF(y) + 2\int_a^b x\widetilde{\psi}(x)\, \mathsf{Re}\left[\breve{m}_F(x)\right]\, dF(x) \right| \leq \varepsilon\ .$$

Since this holds for any $\varepsilon > 0$, Equation (11.11) follows. ■

Putting together Lemmas 11.4 and 11.5 yields

$$\frac{2}{p^2}\sum_{j=1}^p \sum_{i>j} \frac{\lambda_j\widetilde{\psi}(\lambda_j) - \lambda_i\widetilde{\psi}(\lambda_i)}{\lambda_j - \lambda_i} \xrightarrow{\text{a.s.}} -2\int_a^b x\widetilde{\psi}(x)\, \mathsf{Re}\left[\breve{m}_F(x)\right]\, dF(x)\ .$$

Lemma 11.6. *As n and p go to infinity with their ratio p/n converging to the concentration c,*

$$\log(n) - \frac{1}{p}\sum_{j=1}^p \mathbb{E}[\log(\chi^2_{n-j+1})] \longrightarrow 1 + \frac{1-c}{c}\log(1-c)\ .$$

Proof of Lemma 11.6. It is well known that, for every positive integer ν,

$$\mathbb{E}[\log(\chi^2_\nu)] = \log(2) + \frac{\Gamma'(\nu/2)}{\Gamma(\nu/2)}\ ,$$

where $\Gamma(\cdot)$ denotes the gamma function. Thus,

$$\frac{1}{p}\sum_{j=1}^{p}\mathbb{E}[\log(\chi^2_{n-j+1})] = \log(2) + \frac{1}{p}\sum_{j=1}^{p}\frac{\Gamma'((n-j+1)/2)}{\Gamma((n-j+1)/2)} \ .$$

Formula 6.3.21 of Abramowitz and Stegun (1965) states that

$$\forall x \in (0,+\infty) \qquad \frac{\Gamma'(x)}{\Gamma(x)} = \log(x) - \frac{1}{2x} - 2\int_0^\infty \frac{t\,dt}{(t^2+x^2)(e^{2\pi t}-1)} \ .$$

It implies that

$$\begin{aligned}\log(n) - \frac{1}{p}\sum_{j=1}^{p}\mathbb{E}[\log(\chi^2_{n-j+1})] &= -\frac{1}{p}\sum_{j=1}^{p}\log\left(1-\frac{j-1}{n}\right) + \frac{1}{p}\sum_{k=n-p+1}^{n}\frac{1}{k} \\ &\quad + \frac{1}{p}\sum_{k=n-p+1}^{n}\int_0^\infty \frac{t\,dt}{[t^2+(k/2)^2](e^{2\pi t}-1)} \\ &=: -\frac{1}{p}\sum_{j=1}^{p}\log\left(1-\frac{j-1}{n}\right) + A_n + B_n \ .\end{aligned}$$

It is easy to verify that

$$-\frac{1}{p}\sum_{j=1}^{p}\log\left(1-\frac{j-1}{n}\right) \longrightarrow -\frac{1}{c}\int_0^c \log(1-x)dx = 1 + \frac{1-c}{c}\log(1-c) \ .$$

Therefore, all that remains to be proven is that the two terms A_n and B_n vanish. Using formulas 6.3.2 and 6.3.18 of Abramowitz and Stegun (1965), we see that

$$A_n := \frac{1}{p}\sum_{k=n-p+1}^{n}\frac{1}{k} = \frac{1}{p}\left[\frac{\Gamma'(n)}{\Gamma(n)} - \frac{\Gamma'(n-p+1)}{\Gamma(n-p+1)}\right] = \frac{1}{p}\log\left(\frac{n}{n-p+1}\right) + O\left(\frac{1}{p(n-p+1)}\right) \ ,$$

which vanishes indeed. As for the term B_n, it admits the upper bound

$$B_n := \frac{1}{p}\sum_{k=n-p+1}^{n}\int_0^\infty \frac{t\,dt}{[t^2+(k/2)^2](e^{2\pi t}-1)} \leq \int_0^\infty \frac{t\,dt}{[t^2+((n-p+1)/2)^2](e^{2\pi t}-1)} \ ,$$

which also vanishes. ■

Going back to Equation (2.2), we notice that the term

$$\frac{2}{p}\sum_{j=1}^{p}\lambda_j\widetilde{\psi}'(\lambda_j)$$

remains bounded asymptotically with probability one, since $\widetilde{\psi}'$ is bounded over a compact set.

Putting all these results together shows that the unbiased estimator of risk $\Theta_n(S_n,\widehat{\Sigma})$ converges almost surely to

$$\begin{aligned}&(1-c)\int_a^b \widetilde{\psi}(x)dF(x) - \int_a^b \log[\widetilde{\psi}(x)]dF(x) - 2c\int_a^b x\widetilde{\psi}(x)\mathsf{Re}[\breve{m}_F(x)]dF(x) + \frac{1-c}{c}\log(1-c) \\ &= \int_a^b\left\{\frac{1-c-2\,c\,x\,\mathsf{Re}[\breve{m}_F(x)]}{x}\widetilde{\varphi}(x) - \log[\widetilde{\varphi}(x)]\right\}dF(x) + \int_a^b \log(x)dF(x) + \frac{1-c}{c}\log(1-c) \\ &= \int_a^b\left\{\frac{1-c-2\,c\,x\,\mathsf{Re}[\breve{m}_F(x)]}{x}\widetilde{\varphi}(x) - \log[\widetilde{\varphi}(x)]\right\}dF(x) + \int \log(t)\,dH(t) - 1 \ ,\end{aligned}$$

where the last equality comes from the following lemma.

Lemma 11.7. $\int_a^b \log(x)\,dF(x) + \frac{1-c}{c}\log(1-c) = \int \log(t)\,dH(t) - 1$.

Proof of Lemma 11.7. Setting $\widetilde{\varphi}(x) = x$ for all $x \in \mathsf{Supp}(F)$ in Lemma 11.3 yields

$$\frac{1}{p}\log\left[\det\left(\Sigma_n^{-1}S_n\right)\right] \xrightarrow{\text{a.s.}} \int_a^b \log(x)\,dF(x) - \int \log(t)\,dH(t)\ . \tag{11.12}$$

In addition, note that

$$\begin{aligned}\frac{1}{p}\log\left[\det\left(\Sigma_n^{-1}S_n\right)\right] &= \frac{1}{p}\log\left[\det\left(\Sigma_n^{-1}\frac{1}{n}\sqrt{\Sigma_n}X_n'X_N\sqrt{\Sigma_n}\right)\right]\\ &= \frac{1}{p}\log\left[\det\left(\frac{1}{n}X_n'X_n\right)\right] \xrightarrow{\text{a.s.}} \frac{c-1}{c}\log(1-c) - 1\ ,\end{aligned} \tag{11.13}$$

where the convergence comes from Equation (1.1) of Bai and Silverstein (2004). Comparing Equation (11.12) with Equation (11.13) proves the lemma. ■

It is easy to verify that these results carry through to the more general case where the function $\widetilde{\psi}_n$ can vary across n, as long as it is well behaved asymptotically in the sense of Assumption 3.4. ■

11.3 Proof of Proposition 3.2

We provide a proof by contradiction. Suppose that Proposition 3.2 does not hold. Then there exist $\varepsilon > 0$ and $x_0 \in \mathsf{Supp}(F)$ such that

$$1 - c - 2\,c\,x_0\,\mathsf{Re}[\breve{m}_F(x_0)] \le \frac{a_1}{\overline{h}} - 2\varepsilon\ . \tag{11.14}$$

Since $\breve{m}_F$ is continuous, there exist $x_1, x_2 \in \mathsf{Supp}(F)$ such that $x_1 < x_2$, $[x_1, x_2] \subset \mathsf{Supp}(F)$, and

$$\forall x \in [x_1, x_2] \qquad 1 - c - 2\,c\,x\,\mathsf{Re}[\breve{m}_F(x)] \le \frac{a_1}{\overline{h}} - \varepsilon\ .$$

Define, for all $n \in \mathbb{N}$ and $x \in \mathbb{R}$,

$$\begin{aligned}\overline{\varphi}(x) &:= x\,\mathbb{1}_{[x_1,x_2]}(x)\\ \overline{\varphi}_n(x) &:= \overline{\varphi}(x)\\ \overline{D}_n &:= \mathsf{Diag}\big(\overline{\varphi}_n(\lambda_{n,1}),\ldots,\overline{\varphi}_n(\lambda_{n,p})\big)\\ \overline{S}_n &:= U_n\overline{D}_nU_n'\ .\end{aligned}$$

By Lemmas 11.1–11.2,

$$\frac{1}{p}\mathsf{Tr}\left(\Sigma_n^{-1}\overline{S}_n\right) \xrightarrow{\text{a.s.}} \int \overline{\varphi}(x)\,\frac{1 - c - 2\,c\,x\,\mathsf{Re}[\breve{m}_F(x)]}{x}\,dF(x)\ . \tag{11.15}$$

The left-hand side of Equation (11.15) is asymptotically bounded from below as follows.

$$\begin{aligned}\frac{1}{p}\mathsf{Tr}\left(\Sigma_n^{-1}\overline{S}_n\right) &= \frac{1}{p}\sum_{i=1}^{p} u_{n,i}'\Sigma_n^{-1}u_{n,i} \times \lambda_{n,i}\,\mathbb{1}_{[x_1,x_2]}(\lambda_{n,i})\\ &\ge \frac{\lambda_{n,1}}{\overline{h}}\left[F_n(x_2) - F_n(x_1)\right] \xrightarrow{\text{a.s.}} \frac{a_1}{\overline{h}}\left[F(x_2) - F(x_1)\right]\ .\end{aligned} \tag{11.16}$$

The right-hand side of Equation (11.15) is bounded from above as follows.

$$\int \overline{\varphi}(x)\, \frac{1-c-2\,c\,x\,\mathsf{Re}[\breve{m}_F(x)]}{x}\, dF(x) \leq \left(\frac{a_1}{\overline{h}} - \varepsilon\right) [F(x_2) - F(x_1)] \ . \tag{11.17}$$

Given that $F(x_2) - F(x_1) > 0$, Equations (11.15)–(11.17) form a logical contradiction. Therefore, the initial assumption (11.14) must be false, which proves Proposition 3.2. ■

11.4 Proof of Proposition 3.3

If we compare Equations (3.8) and (3.9), we see that the term $\int \log(t)\, dH(t)$ appears in both, so it can be ignored. The challenge is then to prove that

$$\sum_{k=1}^{\kappa} \int_{a_k}^{b_k} \{c + 2\,c\,x\,\mathsf{Re}[\breve{m}_F(x)] + \log(x)\} dF(x) < \sum_{k=1}^{\kappa} \int_{a_k}^{b_k} \log\left[\frac{x}{1-c-2\,c\,x\,\mathsf{Re}[\breve{m}_F(x)]}\right] dF(x) \ . \tag{11.18}$$

Rearranging terms, we can restate this inequality as

$$\sum_{k=1}^{\kappa} \int_{a_k}^{b_k} \left\{c + 2\,c\,x\,\mathsf{Re}[\breve{m}_F(x)] + \log\left(1-c-2\,c\,x\,\mathsf{Re}[\breve{m}_F(x)]\right)\right\} dF(x) < 0 \ . \tag{11.19}$$

Setting $y := c + 2\,c\,x\,\mathsf{Re}[\breve{m}_F(x)]$ leads us to investigate the function $y \mapsto y + \log(1-y)$. Elementary calculus shows that it is strictly negative over its domain of definition, except at $y = 0$, where it attains its maximum of zero. The condition $y = 0$ is equivalent to $x\,\mathsf{Re}[\breve{m}_F(x)] = -1/2$.

If we set the variable x equal to a_1, the lower bound of the leftmost interval of the support of the limiting sample spectral distribution F, we get

$$a_1\,\mathsf{Re}[\breve{m}_F(a1)] = \mathrm{PV}\int_{-\infty}^{\infty} \frac{a_1}{\lambda - a_1}\, dF(\lambda) \ , \tag{11.20}$$

where PV denotes the Cauchy Principal Value (Henrici, 1988, pp. 259–262). The quantity in Equation (11.20) is nonnegative because $\lambda \geq a_1$ for all $\lambda \in \mathsf{Supp}(F)$. By continuity, there exists some $\beta_1 \in (a_1, b_1]$ such that $x\,\mathsf{Re}[\breve{m}_F(x)] > -1/2$ for all $x \in [a_1, \beta_1]$. This implies that the strict inequality (11.19) is true. ■

11.5 Proof of Proposition 3.4

Subtracting Equation (3.8) from Equation (3.14) shows that the difference between limiting losses $\mathcal{M}_c^S(H, \varphi^M) - \mathcal{M}_c^S\left(H, \varphi^S\right)$ is equal to

$$\int \left\{\frac{1-c-2cx\mathsf{Re}[\breve{m}_F(x)]}{1-c+2cF(x)} - 1 - \log\left[\frac{1-c-2cx\mathsf{Re}[\breve{m}_F(x)]}{1-c+2cF(x)}\right]\right\} dF(x) \ . \tag{11.21}$$

The function $y \longmapsto y - 1 + \log(y)$ is strictly positive over its domain of definition, except at $y = 1$, where it attains its minimum of zero. Therefore

$$\forall x \in \mathsf{Supp}(F) \qquad \frac{1-c-2cx\mathsf{Re}[\breve{m}_F(x)]}{1-c+2cF(x)} - 1 - \log\left[\frac{1-c-2cx\mathsf{Re}[\breve{m}_F(x)]}{1-c+2cF(x)}\right] \geq 0 \ , \tag{11.22}$$

which implies that $\mathcal{M}_c^S(H, \varphi^M) - \mathcal{M}_c^S\left(H, \varphi^S\right) \geq 0$, as we already knew from Corollary 3.1. Elementary calculus shows that the inequality (11.22) is strict if and only if $-x\,\mathsf{Re}[\breve{m}_F(x)] \neq F(x)$. As in the proof of Proposition 3.3, we use a_1, the lower bound of the leftmost interval of the support of the limiting sample spectral distribution F:

$$\forall x \in [0, a_1) \qquad -x\,\mathsf{Re}[\breve{m}_F(x)] = \int \frac{x}{x-\lambda}\,dF(\lambda) = 1 + \int \frac{\lambda}{x-\lambda}\,dF(\lambda)\ , \tag{11.23}$$

which is a strictly decreasing function of x. Therefore, using the fact that $\breve{m}_F$ is continuous (Silverstein and Choi, 1995, Theorem 1.1), $-a_1\,\mathsf{Re}[\breve{m}_F(a_1)]$ is strictly below the value that $-x\,\mathsf{Re}[\breve{m}_F(x)]$ takes at $x = 0$, which is itself zero. It implies $-a_1\,\mathsf{Re}[\breve{m}_F(a_1)] \neq F(a_1)$. By continuity, there exists some $\beta_1' \in (a_1, b_1]$ such that $-x\,\mathsf{Re}[\breve{m}_F(x)] \neq F(x)$ for all $x \in [a_1, \beta_1']$. This in turn implies that the integral in Equation (11.21) is strictly positive. ■

11.6 Proof of Proposition 3.5

The linear shrinkage estimator in Equation (14) of Ledoit and Wolf (2004) is of the form

$$S_n^L := m_n \mathbb{I}_n + \frac{a_n^2}{d_n^2}\,(S_n - m_n \mathbb{I}_n)\ , \tag{11.24}$$

where

$$m_n := \int \lambda\,dF_n(\lambda) \qquad \xrightarrow{\text{a.s.}} \int \lambda\,dF(\lambda) \tag{11.25}$$

$$a_n^2 := \int t^2 dH_n(t) - \left[\int t\,dH_n(t)\right]^2 \qquad \longrightarrow \int t^2 dH(t) - \left[\int t\,dH(t)\right]^2 \tag{11.26}$$

$$d_n^2 := \int \lambda^2 dF_n(\lambda) - \left[\int \lambda\,dF_n(\lambda)\right]^2 \qquad \xrightarrow{\text{a.s.}} \int \lambda^2 dF(\lambda) - \left[\int \lambda\,dF(\lambda)\right]^2 . \tag{11.27}$$

Thus, the linear shrinkage function is $\varphi_n^L : x \longmapsto m_n + \left(a_n^2/d_n^2\right)(x - m_n)$. Under Assumptions 3.1–3.3,

$$\forall x \in \mathsf{Supp}(F) \qquad \varphi_n^L(x) \xrightarrow{\text{a.s.}} \int \lambda\,dF(\lambda) + \frac{\displaystyle\int t^2 dH(t) - \left[\int t\,dH(t)\right]^2}{\displaystyle\int \lambda^2 dF(\lambda) - \left[\int \lambda\,dF(\lambda)\right]^2}\left[x - \int \lambda\,dF(\lambda)\right] . \tag{11.28}$$

Since the support of F is compact, the convergence is uniform. ■

12 Proofs of Theorems in Section 4

12.1 Proof of Theorem 4.1

Lemma 12.1. *Under Assumptions 3.1–3.3, there exists a nonrandom function $\Delta^{(1)}$ defined on $\mathbb{R}$ such that the random function $\Delta_n^{(1)}(x)$ converges almost surely to $\Delta^{(1)}(x)$, for all $x \in \mathbb{R}$. Furthermore, $\Delta^{(1)}$ is continuously differentiable on $\mathbb{R}$ and can be expressed as*

$$\forall x \in \mathbb{R} \qquad \Delta^{(1)}(x) = \begin{cases} 0 & \text{if } x < a, \\ \int_a^x \delta^{(1)}(\lambda) dF(\lambda) & \text{if } x \geq a, \end{cases}$$

where $\forall \lambda \in [a, +\infty)$, $\delta^{(1)}(\lambda) := \lambda / |1 - c - c\,\lambda\,\breve{m}_F(\lambda)|^2$.

Proof of Lemma 12.1. Follows directly from Theorem 4 of Ledoit and Péché (2011). ■

Lemma 12.2. *Under Assumptions 3.1–3.4,*

$$\frac{1}{p}\mathsf{Tr}\big(\Sigma_n \widetilde{S}_n^{-1}\big) \xrightarrow{\text{a.s.}} \int_a^b \frac{1}{\widetilde{\varphi}(x)}\, d\Delta^{(1)}(x)\ .$$

Proof of Lemma 12.2. Note that

$$\frac{1}{p}\mathsf{Tr}\big(\Sigma_n \widetilde{S}_n^{-1}\big) = \frac{1}{p}\sum_{i=1}^{p} \frac{u_i'\Sigma_n u_i}{\widetilde{\varphi}_n(\lambda_i)} = \int \frac{1}{\widetilde{\varphi}_n(x)}\, d\Delta_n^{(1)}(x)\ .$$

The remainder of the proof is similar to the proof of Lemma 11.2 and is thus omitted. ■

Lemma 12.1 and Lemma 12.2 imply that

$$\frac{1}{p}\mathsf{Tr}\big(\Sigma_n \widetilde{S}_n^{-1}\big) \xrightarrow{\text{a.s.}} \int_a^b \frac{x}{\widetilde{\varphi}(x)\ |1 - c - c\,x\,\breve{m}_F(x)|^2}\, dF(x)\ . \tag{12.1}$$

Lemma 11.3 implies that

$$-\frac{1}{p}\log\Big[\det\big(\Sigma_n \widetilde{S}_n^{-1}\big)\Big] - 1 \xrightarrow{\text{a.s.}} \int_a^b \log\left[\widetilde{\varphi}(x)\right] dF(x) - \int \log(t)\, dH(t) - 1\ .$$

Putting these two results together completes the proof of Theorem 4.1. ■

12.2 Proof of Theorem 4.2

Note that

$$\begin{aligned}\frac{1}{p}\mathsf{Tr}\left[\left(\Sigma_n - \widetilde{S}_n\right)^2\right] &= \frac{1}{p}\sum_{i=1}^{p}\left[\tau_{n,i}^2 - 2u_{n,i}'\Sigma_n u_{n,i}\,\widetilde{\varphi}_n(\lambda_{n,i}) + \widetilde{\varphi}_n(\lambda_{n,i})^2\right] \\ &= \int x^2\, dH_n(x) - 2\int \widetilde{\varphi}_n(x)\, d\Delta_n^{(1)}(x) + \int \widetilde{\varphi}_n(x)^2\, dF_n(x)\ .\end{aligned}$$

The remainder of the proof is similar to the proof of Lemma 11.2 and is thus omitted. ■

12.3 Proof of Theorem 4.3

Note that

$$\begin{aligned}\frac{1}{p}\mathsf{Tr}\left[\left(\Sigma_n^{-1} - \widetilde{S}_n^{-1}\right)^2\right] &= \frac{1}{p}\sum_{i=1}^{p}\left[\frac{1}{\tau_{n,i}^2} - 2\frac{u_{n,i}'\Sigma_n^{-1}u_{n,i}}{\widetilde{\varphi}_n(\lambda_{n,i})} + \frac{1}{\widetilde{\varphi}_n(\lambda_{n,i})^2}\right] \\ &= \int \frac{1}{x^2}\, dH_n(x) - 2\int \frac{1}{\widetilde{\varphi}_n(x)}\, d\Delta_n^{(-1)}(x) + \int \frac{1}{\widetilde{\varphi}_n(x)^2}\, dF_n(x)\ .\end{aligned}$$

The remainder of the proof is similar to the proof of Lemma 11.2 and is thus omitted. ■

13 Proof of Theorem 5.2

Define the shrinkage function

$$\forall x \in \mathsf{Supp}\big(F_{n,p}^{\widehat{\boldsymbol{\tau}}_n}\big) \qquad \widehat{\varphi}_n^*(x) := \frac{x}{1 - \dfrac{p}{n} - 2\,\dfrac{p}{n}\,x\,\mathsf{Re}\big[\breve{m}_{n,p}^{\widehat{\boldsymbol{\tau}}_n}(x)\big]}\ .$$

Theorem 2.2 of Ledoit and Wolf (2015) and Proposition 4.3 of Ledoit and Wolf (2012) imply that $\forall x \in \mathsf{Supp}(F)$, $\widehat{\varphi}_n^*(x) \xrightarrow{\text{a.s}} \varphi^*(x)$, and that this convergence is uniform over $x \in \mathsf{Supp}(F)$, apart from arbitrarily small boundary regions of the support. Theorem 5.2 then follows from Corollary 3.1. ■

14 Proofs of Theorems in Section 6

14.1 Proof of Theorem 6.1

Lemma 14.1. *Under Assumptions 3.2, 3.3, and 6.1, there exists a nonrandom function $\Delta^{(-1)}$ defined on $\mathbb{R}$ such that $\Delta_n^{(-1)}(x)$ converges almost surely to $\Delta^{(-1)}(x)$, for all $x \in \mathbb{R} - \{0\}$. Furthermore, $\Delta^{(-1)}$ is continuously differentiable on $\mathbb{R} - \{0\}$ and can be expressed as $\forall x \in \mathbb{R}$, $\Delta^{(-1)}(x) = \int_{-\infty}^{x} \delta^{(-1)}(\lambda) dF(\lambda)$, where*

$$\forall \lambda \in \mathbb{R} \qquad \delta^{(-1)}(\lambda) := \begin{cases} 0 & \text{if } \lambda < 0, \\ \dfrac{c}{c-1} \cdot \breve{m}_H(0) - \breve{m}_{\underline{F}}(0) & \text{if } \lambda = 0, \\ \dfrac{1 - c - 2\,c\,\lambda\,\mathsf{Re}[\breve{m}_F(\lambda)]}{\lambda} & \text{if } \lambda > 0\ . \end{cases}$$

Proof of Lemma 14.1. The proof of Lemma 14.1 follows directly from Ledoit and Péché (2011, Theorem 5) and the corresponding proof, bearing in mind that we are in the case $c > 1$ because of Assumption 6.1. ■

The proof of Theorem 6.1 proceeds as the proof of Theorem 3.1, except that Lemma 14.1 replaces Lemma 11.1. ■

14.2 Proof of Theorem 6.2

Define the shrinkage function

$$\widehat{\varphi}_n^*(0) := \left(\frac{p/n}{p/n - 1} \cdot \widehat{\breve{m}_H}(0) - \widehat{\breve{m}_{\underline{F}}}(0)\right)^{-1},$$

$$\text{and} \quad \forall x \in \mathsf{Supp}\big(\underline{F}_{n,p}^{\widehat{\boldsymbol{\tau}}_n}\big) \qquad \widehat{\varphi}_n^*(x) := \frac{x}{1 - \dfrac{p}{n} - 2\,\dfrac{p}{n}\,x\,\mathsf{Re}\big[\breve{m}_{n,p}^{\widehat{\boldsymbol{\tau}}_n}(x)\big]}\ .$$

First, since both $\widehat{\breve{m}_H}(0)$ and $\widehat{\breve{m}_{\underline{F}}}(0)$ are strongly consistent estimators, $\widehat{\varphi}_n^*(0) \xrightarrow{\text{a.s}} \varphi^*(0)$. Second, Theorem 2.2 of Ledoit and Wolf (2015) and Proposition 4.3 of Ledoit and Wolf (2012) applied to $\underline{F}$ imply that $\forall x \in \mathsf{Supp}(\underline{F})$, $\widehat{\varphi}_n^*(x) \xrightarrow{\text{a.s}} \varphi^*(x)$, and that this convergence is uniform over $x \in \mathsf{Supp}(\underline{F})$, apart from arbitrarily small boundary regions of the support. Theorem 6.2 then follows from Corollary 6.1. ■

15 Proofs of Propositions in Section 7

15.1 Common Notation

Let V_n denote a matrix of eigenvectors of Σ_n arranged to match the ascending order of the eigenvalues vector $\boldsymbol{\tau}_n = (\tau_{n,1}, \ldots, \tau_{n,p})$. Let $v_{n,p}$ denote the pth column vector of the matrix V_n. We can decompose the population covariance matrix Σ_n into its bulk and arrow components according to $\Sigma_n = \Sigma_n^B + \Sigma_n^A$, where

$$\Sigma_n^B := V_n \times \mathsf{Diag}(\tau_{n,1}, \ldots, \tau_{n,p-1}, 0) \times V_n' \tag{15.1}$$

$$\Sigma_n^A := V_n \times \mathsf{Diag}(\underbrace{0, \ldots, 0}_{p-1 \text{ times}}, \tau_{n,p}) \times V_n' \,. \tag{15.2}$$

Note that the $\min(n,p)$ largest eigenvalues of S_n are the same as those of $T_n := n^{-1} X_n \Sigma_n X_n'$, so in many instances we will be able to simply investigate the spectral decomposition of the latter matrix. Equations (15.1)–(15.2) enable us to write $T_n = T_n^B + T_n^A$, where $T_n^B := n^{-1} X_n \Sigma_n^B X_n'$ and $T_n^A := n^{-1} X_n \Sigma_n^A X_n'$.

15.2 Proof of Proposition 7.1

Given that the bulk population eigenvalues are below $\overline{h}$, Theorem 1.1 of Bai and Silverstein (1998) shows that there exists a constant $\overline{B}$ such that the largest eigenvalue of T_n^B is below $\overline{B}$ almost surely for all n sufficiently large. Furthermore, due to the fact that the rank of the matrix T_n^A is one, its second largest eigenvalue is zero. Therefore the Weyl inequalities (e.g., see Theorem 1 in Section 6.7 of Franklin (2000) for a textbook treatment) imply that $\lambda_{n,p-1} \leq \overline{B} + 0 = \overline{B}$ a.s. for sufficiently large n. This establishes the first part of the proposition.

As for the second part, it comes from

$$\frac{\lambda_{n,p}}{\tau_{n,p}} \geq \frac{v_{n,p}' S_n v_{n,p}}{\tau_{n,p}} = \frac{1}{\tau_{n,p}} v_{n,p}' \sqrt{\Sigma_n} \frac{X_n' X_n}{n} \sqrt{\Sigma_n} v_{n,p} = v_{n,p}' \frac{X_n' X_n}{n} v_{n,p} \xrightarrow{\text{a.s.}} 1. \; \blacksquare \tag{15.3}$$

15.3 Proof of Proposition 7.2

Lemma 15.1. *Under Assumptions 3.1, 3.2.a–c, and 3.2.f, there is spectral separation between the arrow and the bulk in the sense that*

$$\sup \left\{ t \in \mathbb{R} : F_{n,p}^{\boldsymbol{\tau}_n}(t) \leq \frac{p-1}{p} \right\} < \inf \left\{ t \in \mathbb{R} : F_{n,p}^{\boldsymbol{\tau}_n}(t) > \frac{p-1}{p} \right\} \tag{15.4}$$

for large enough n.

Proof of Lemma 15.1. From page 5356 of Mestre (2008), a necessary and sufficient condition for spectral separation to occur between the arrow and the bulk is that

$$\exists t \in (\tau_{n,p-1}, \tau_{n,p}) \quad \text{s.t.} \qquad \Theta_n(t) := \frac{1}{p} \sum_{i=1}^{p} \frac{\tau_{n,i}^2}{\left(\tau_{n,i} - t\right)^2} - \frac{1}{c} < 0 \,. \tag{15.5}$$

This is equivalent to the condition that the function $x_F(m)$ defined in Equation (1.6) of Silverstein and Choi (1995) is strictly increasing at $m = -1/t$. Section 4 of Silverstein and Choi

(1995) explains in detail how this enables us to determine the boundaries of the support of $F^{\tau_n}_{n,p}$. Assumption 3.2.f guarantees that

$$\forall i = 1, \ldots, p-1, \quad \forall t \in (\tau_{n,p-1}, \tau_{n,p}), \qquad \frac{\tau_{n,i}^2}{\left(\tau_{n,i} - t\right)^2} \le \frac{\overline{h}^2}{\left(\overline{h} - t\right)^2} , \tag{15.6}$$

therefore a sufficient condition for arrow separation is that

$$\exists t \in (\tau_{n,p-1}, \tau_{n,p}) \quad \text{s.t.} \qquad \theta_n(t) := \frac{p-1}{p} \frac{\overline{h}^2}{\left(\overline{h} - t\right)^2} + \frac{1}{p} \frac{\tau_{n,p}^2}{\left(\tau_{n,p} - t\right)^2} - \frac{1}{c} < 0 . \tag{15.7}$$

The function θ_n is strictly convex on $\left(\overline{h}, \tau_{n,p}\right)$ and goes to infinity as it approaches $\overline{h}$ and $\tau_{n,p}$, therefore it admits a unique minimum on $\left(\overline{h}, \tau_{n,p}\right)$ characterized by the first-order condition

$$\begin{aligned}
\theta_n'(t) = 0 &\Longleftrightarrow 2\,\frac{p-1}{p} \frac{\overline{h}^2}{\left(\overline{h} - t\right)^3} + 2\,\frac{1}{p} \frac{\tau_{n,p}^2}{\left(\tau_{n,p} - t\right)^3} = 0 \\
&\Longleftrightarrow \frac{p-1}{p} \frac{\overline{h}^2}{\left(t - \overline{h}\right)^3} = \frac{1}{p} \frac{\tau_{n,p}^2}{\left(\tau_{n,p} - t\right)^3} \\
&\Longleftrightarrow \left(\frac{p}{p-1}\right)^{1/3} \frac{t - \overline{h}}{\overline{h}^{2/3}} = p^{1/3} \frac{\tau_{n,p} - t}{\tau_{n,p}^{2/3}} \\
&\Longleftrightarrow t = t_n^* := \left(\overline{h}\,\tau_{n,p}\right)^{2/3} \frac{\left(\frac{p}{p-1}\right)^{1/3} \tau_{n,p}^{1/3} + \left(\frac{1}{p}\right)^{1/3} \overline{h}^{1/3}}{\left(\frac{p}{p-1}\right)^{1/3} \overline{h}^{2/3} + \left(\frac{1}{p}\right)^{1/3} \tau_{n,p}^{2/3}} .
\end{aligned}$$

Note that $t_n^* \sim \overline{h}^{2/3} \beta_1^{1/3} p^{2/3}$, therefore

$$\theta_n\left(t_n^*\right) \sim \frac{\overline{h}^2}{\overline{h}^{4/3} \beta_1^{2/3} p^{4/3}} + \frac{\beta_1^2\, p^2}{\beta_1^2\, p^3} - \frac{1}{c} \longrightarrow -\frac{1}{c} < 0 , \tag{15.8}$$

which implies that condition (15.7) is satisfied for large enough n, and the arrow separates from the bulk. ∎

Since the function Θ_n from Equation (15.5) is strictly convex over the interval $\left(\tau_{n,p-1}, t_n^*\right)$, $\lim_{t \searrow \tau_{n,p-1}} \Theta_n(t) = +\infty$ and $\Theta_n(t_n^*) \le \theta_n(t_n^*) < 0$ by Lemma 15.1, Θ_n admits a unique zero in $\left(\tau_{n,p-1}, t_n^*\right)$. Call it $\overline{b}_n$. An asymptotically valid bound for $\overline{b}_n$ is given by the following lemma.

Lemma 15.2. *Under Assumptions 3.1, 3.2.a–c, and 3.2.f,*

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \textit{s.t.} \quad \forall n \ge N \qquad \overline{b}_n \le \left(1 + \sqrt{c + \varepsilon}\right)\overline{h} . \tag{15.9}$$

Proof of Lemma 15.2.

$$\Theta(\overline{b}_n) = 0 \Longleftrightarrow \frac{1}{p} \sum_{i=1}^{p-1} \frac{\tau_{n,i}^2}{\left(\tau_{n,i} - \overline{b}_n\right)^2} + \frac{1}{p} \frac{\tau_{n,p}^2}{\left(\tau_{n,p} - \overline{b}_n\right)^2} = \frac{1}{c} . \tag{15.10}$$

From $\overline{b}_n \le t_n^*$ and $\tau_{n,p} \sim \beta_1 p$ we deduce

$$\frac{1}{p} \frac{\tau_{n,p}^2}{\left(\tau_{n,p} - \overline{b}_n\right)^2} \sim \frac{1}{p} \longrightarrow 0 ; \tag{15.11}$$

therefore,

$$\frac{1}{p}\sum_{i=1}^{p-1}\frac{\tau_{n,i}^2}{\left(\tau_{n,i}-\overline{b}_n\right)^2}\longrightarrow\frac{1}{c}\ . \tag{15.12}$$

This implies that $\forall\varepsilon>0\ \ \exists N\in\mathbb{N}$ s.t. $\forall n\geq N$

$$\begin{aligned}\frac{1}{p}\sum_{i=1}^{p-1}\frac{\tau_{n,i}^2}{\left(\tau_{n,i}-\overline{b}_n\right)^2}&\geq\frac{1}{c+\varepsilon}\\ \frac{p-1}{p}\frac{\overline{h}^2}{\left(\overline{h}-\overline{b}_n\right)^2}&\geq\frac{1}{c+\varepsilon}\\ \frac{\left(\overline{h}-\overline{b}_n\right)^2}{\overline{h}^2}&\leq c+\varepsilon\\ \overline{b}_n&\leq\left(1+\sqrt{c+\varepsilon}\right)\overline{h}.\ \blacksquare\end{aligned} \tag{15.13}$$

Since the function Θ_n from Equation (15.5) is strictly convex over the interval $\left(t_n^*,\tau_{n,p}\right)$, $\lim_{t\nearrow\tau_{n,p}}\Theta_n(t)=+\infty$ and $\Theta_n(t_n^*)\leq\theta_n(t_n^*)<0$ by Lemma 15.1, Θ_n admits a unique zero in $\left(t_n^*,\tau_{n,p}\right)$. Call it $\underline{t}_n$. An asymptotically valid equivalency result for $\underline{t}_n$ is given by the following lemma.

Lemma 15.3. *Under Assumptions 3.1, 3.2.a–c, and 3.2.f,*

$$\tau_{n,p}-\underline{t}_n\sim\frac{\tau_{n,p}}{\sqrt{n}}\ . \tag{15.14}$$

Proof of Lemma 15.3.

$$\Theta(\underline{t}_n)=0\Longleftrightarrow\frac{1}{p}\sum_{i=1}^{p-1}\frac{\tau_{n,i}^2}{\left(\tau_{n,i}-\underline{t}_n\right)^2}+\frac{1}{p}\frac{\tau_{n,p}^2}{\left(\tau_{n,p}-\underline{t}_n\right)^2}=\frac{1}{c}\ . \tag{15.15}$$

From the inequalities $\underline{t}_n\geq t_n^*$ and $\tau_{n,i}\leq\overline{h}$ (for $i=1,\dots,p-1$) we deduce

$$\frac{1}{p}\sum_{i=1}^{p-1}\frac{\tau_{n,i}^2}{\left(\tau_{n,i}-\underline{t}_n\right)^2}\leq\frac{p-1}{p}\frac{\overline{h}^2}{\left(\overline{h}-t_n^*\right)^2}\sim\frac{\overline{h}^{2/3}}{\beta_1^{2/3}p^{4/3}}\longrightarrow 0\ , \tag{15.16}$$

therefore

$$\begin{aligned}\frac{1}{p}\frac{\tau_{n,p}^2}{\left(\tau_{n,p}-\underline{t}_n\right)^2}&\longrightarrow\frac{1}{c}\\ \frac{1}{n}\frac{\tau_{n,p}^2}{\left(\tau_{n,p}-\underline{t}_n\right)^2}&\longrightarrow 1\\ \frac{\tau_{n,p}-\underline{t}_n}{\tau_{n,p}/\sqrt{n}}&\longrightarrow 1.\ \blacksquare\end{aligned}$$

Lemma 15.4. *Define*

$$\underline{\lambda}_n:=\inf\left\{t\in\mathbb{R}:F_{n,p}^{\boldsymbol{\tau}_n}(t)>\frac{p-1}{p}\right\}\ . \tag{15.17}$$

Then under Assumptions 3.1, 3.2.a–c, and 3.2.f,

$$\tau_{n,p}-\underline{\lambda}_n\sim 2\frac{\tau_{n,p}}{\sqrt{n}}\ . \tag{15.18}$$

Proof of Lemma 15.4. Equation (13) of Mestre (2008) gives

$$\underline{\lambda}_n = \underline{t}_n - c\,\underline{t}_n \frac{1}{p} \sum_{i=1}^{p} \frac{\tau_{n,i}}{\tau_{n,i} - \underline{t}_n} \ . \tag{15.19}$$

This is equivalent to plugging $m = -1/\underline{t}_n$ into Equation (1.6) of Silverstein and Choi (1995). These authors' Section 4 explains why method yields the boundary points of $\mathsf{Supp}(F^{\boldsymbol{\tau}_n}_{n,p})$. From Equation (15.19) we deduce

$$1 - \frac{\underline{\lambda}_n}{\underline{t}_n} = c\frac{1}{p}\frac{\tau_{n,p}}{\tau_{n,p} - \underline{t}_n} - c\frac{1}{p}\sum_{i=1}^{p-1} \frac{\tau_{n,i}}{\underline{t}_n - \tau_{n,i}} \ . \tag{15.20}$$

Lemma 15.3 enables us to approximate the first term on the right-hand side by

$$c\frac{1}{p}\frac{\tau_{n,p}}{\tau_{n,p} - \underline{t}_n} \sim \frac{p}{n} \times \frac{1}{p} \times \sqrt{n} = \frac{1}{\sqrt{n}} \ . \tag{15.21}$$

Since $\tau_{n,i} \leq \overline{h} < \underline{t}_n$, the second term is bounded by

$$0 \leq c\frac{1}{p}\sum_{i=1}^{p-1} \frac{\tau_{n,i}}{\underline{t}_n - \tau_{n,i}} \leq c\frac{\overline{h}}{\underline{t}_n - \overline{h}} \sim c\frac{\overline{h}}{\beta_1 p} \ , \tag{15.22}$$

therefore it is negligible with respect to the first term. We conclude by remarking that

$$\begin{aligned}
1 - \frac{\underline{\lambda}_n}{\underline{t}_n} &\sim \frac{1}{\sqrt{n}} \\
\underline{t}_n - \underline{\lambda}_n &\sim \frac{\underline{t}_n}{\sqrt{n}} \sim \frac{\tau_{n,p}}{\sqrt{n}} \\
\tau_{n,p} - \underline{\lambda}_n &= \left(\tau_{n,p} - \underline{t}_n\right) + \left(\underline{t}_n - \underline{\lambda}_n\right) \sim 2\frac{\tau_{n,p}}{\sqrt{n}}. \ \blacksquare
\end{aligned}$$

Lemma 15.5. *Define*

$$\overline{\mu}_n := \sup\left\{ t \in \mathbb{R} : F^{\boldsymbol{\tau}_n}_{n,p}(t) \leq \frac{p-1}{p} \right\} \ . \tag{15.23}$$

Then under Assumptions 3.1, 3.2.a–c, and 3.2.f,

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad s.t. \quad \forall n \geq N \qquad \overline{\mu}_n \leq \left(1 + \sqrt{c + \varepsilon}\right)^2 \overline{h} \ . \tag{15.24}$$

Proof of Lemma 15.5. Equation (13) of Mestre (2008) gives

$$\overline{\mu}_n = \overline{b}_n - c\,\overline{b}_n \frac{1}{p} \sum_{i=1}^{p} \frac{\tau_{n,i}}{\tau_{n,i} - \overline{b}_n} \ . \tag{15.25}$$

This is equivalent to plugging $m = -1/\overline{b}_n$ into Equation (1.6) of Silverstein and Choi (1995). Fix any $i \in \{1, 2, \dots, p-2\}$ and hold $(\tau_{n,j})_{j \neq i}$ constant. Define the function

$$\forall b \in (\tau_{n,p-1}, t^*_n), \quad \forall t \leq \tau_{n,i+1} \qquad F_i(b,t) := b - c\,b\frac{1}{p}\frac{t}{t-b} - c\,b\frac{1}{p}\sum_{\substack{j=1 \\ j \neq i}}^{p} \frac{\tau_{n,j}}{\tau_{n,j} - b} \ . \tag{15.26}$$

Then clearly $\overline{\mu}_n = F_i(\overline{b}_n, \tau_{n,i})$. Viewing $\overline{\mu}_n$ and $\overline{b}_n$ as two univariate functions of $\tau_{n,i}$, we can write:

$$\frac{d\overline{\mu}_n}{d\tau_{n,i}} = \frac{\partial F_i}{\partial b}(\overline{b}_n, \tau_{n,i}) \times \frac{d\overline{b}_n}{d\tau_{n,i}} + \frac{\partial F_i}{\partial t}(\overline{b}_n, \tau_{n,i}) \ . \tag{15.27}$$

But notice that

$$\forall b \in (\tau_{n,p-1}, t_n^*), \quad \forall t \le \tau_{n,i+1} \qquad \frac{\partial F_i}{\partial b}(b,t) = 1 - c\,\frac{1}{p}\frac{t^2}{(t-b)^2} - c\,\frac{1}{p}\sum_{\substack{j=1\\ j\neq i}}^{p} \frac{\tau_{n,j}^2}{(\tau_{n,j}-b)^2} \ ; \tag{15.28}$$

therefore,

$$\frac{\partial F_i}{\partial b}(\overline{b}_n, \tau_{n,i}) = -c\ \Theta_n(\overline{b}_n) \ , \tag{15.29}$$

which is identically equal to zero by Equation (15.5). By the envelope theorem, Equation (15.27) simplifies into

$$\frac{d\overline{\mu}_n}{d\tau_{n,i}} = \frac{\partial F_i}{\partial t}(\overline{b}_n, \tau_{n,i}) = c\,\frac{1}{p}\frac{\overline{b}_n^2}{(\tau_{n,i}-\overline{b}_n)^2} > 0 \ . \tag{15.30}$$

We can thus obtain an upper bound on $\overline{\mu}_n$ by setting $\tau_{n,1}, \dots, \tau_{n,p-2}$ equal to $\tau_{n,p-1}$. In this particular case, $\overline{b}_n$ verifies

$$\frac{p-1}{p}\frac{\tau_{n,p-1}^2}{(\tau_{n,p-1}-\overline{b}_n)^2} + \frac{1}{p}\frac{\tau_{n,p}^2}{(\tau_{n,p}-\overline{b}_n)^2} = \frac{1}{c} \ . \tag{15.31}$$

From Equation (15.13) and $\tau_{n,p} \sim \beta_1 p$ we deduce

$$\frac{p-1}{p}\frac{\tau_{n,p-1}^2}{(\tau_{n,p-1}-\overline{b}_n)^2} \longrightarrow \frac{1}{c} \tag{15.32}$$

$$\frac{\tau_{n,p-1}}{\tau_{n,p-1}-\overline{b}_n} \longrightarrow -\frac{1}{\sqrt{c}} \ . \tag{15.33}$$

Thus, in the particular case where $\tau_{n,1}, \dots, \tau_{n,p-2}$ are all equal to $\tau_{n,p-1}$, $\overline{\mu}_n$ verifies

$$\frac{\overline{\mu}_n}{\overline{b}_n} = 1 - c\frac{p-1}{p}\frac{\tau_{n,p-1}}{\tau_{n,p-1}-\overline{b}_n} - c\frac{1}{p}\frac{\tau_{n,p}}{\tau_{n,p}-\overline{b}_n} \longrightarrow 1+\sqrt{c} \ . \tag{15.34}$$

Remember that, by Equation (15.30), the particular case $\tau_{n,1} = \cdots = \tau_{n,p-2} = \tau_{n,p-1}$ yields an upper bound on $\overline{\mu}_n$ that holds in the general case $\tau_{n,1} \le \cdots \le \tau_{n,p-2} \le \tau_{n,p-1}$, therefore putting together Equations (15.13) and (15.34) yields the conclusion

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \text{s.t.} \quad \forall n \ge N \qquad \overline{\mu}_n \le \left(1+\sqrt{c+\varepsilon}\right)^2 \overline{h}. \blacksquare \tag{15.35}$$

The first part of Proposition 7.2 follows from Lemma 15.5 and from the observation that $q_{n,p}^{p-1}(\boldsymbol{\tau}_n)$ is no greater than $\overline{\mu}_n$ as defined in Equation (15.23). The second part of Proposition 7.2 follows from Lemma 15.4 and from the observation that $q_{n,p}^{p}(\boldsymbol{\tau}_n)$ is no smaller than $\underline{\lambda}_n$ as defined in Equation (15.17). ■

15.4 Proof of Proposition 7.3

The eigenvalues of T_n^B are bounded from below by zero. Given that the bulk population eigenvalues are below $\overline{h}$, Theorem 1.1 of Bai and Silverstein (1998) shows that there exists a constant $\overline{B}$ such that the largest eigenvalue of T_n^B is below $\overline{B}$ almost surely for all n sufficiently large. Therefore the Weyl inequalities imply that

$$\lambda_{n,p}^A \le \lambda_{n,p} \le \lambda_{n,p}^A + \overline{B} \tag{15.36}$$

almost surely for sufficiently large n, where $\lambda_{n,p}^A$ denotes the largest eigenvalue of T_n^A. Furthermore, we have

$$\frac{\lambda_{n,p}^A}{\tau_{n,p}} = v_{n,p}' \frac{X_n' X_n}{n} v_{n,p} \xrightarrow{\text{a.s.}} 1 \; . \tag{15.37}$$

Putting together Equations (15.36) and (15.37) yields $\lambda_{n,p}/\tau_{n,p} \xrightarrow{\text{a.s.}} 1$, as desired. ■

15.5 Proof of Proposition 7.4

Since the function Θ_n from Equation (15.5) is strictly convex over the interval $\big(\tau_{n,p}, +\infty\big)$, $\lim_{t \searrow \tau_{n,p}} \Theta_n(t) = +\infty$ and $\lim_{t \searrow +\infty} \Theta_n(t) = -1/c < 0$, Θ_n admits a unique zero in $\big(\tau_{n,p}, +\infty\big)$. Call it $\bar{t}_n$. An asymptotically valid equivalency result for $\bar{t}_n$ is given by the following lemma.

Lemma 15.6. *Under Assumptions 3.1, 3.2.a–c, and 3.2.f,*

$$\bar{t}_n - \tau_{n,p} \sim \frac{\tau_{n,p}}{\sqrt{n}} \; . \tag{15.38}$$

Proof of Lemma 15.6.

$$\Theta(\bar{t}_n) = 0 \Longleftrightarrow \frac{1}{p} \sum_{i=1}^{p-1} \frac{\tau_{n,i}^2}{\big(\tau_{n,i} - \bar{t}_n\big)^2} + \frac{1}{p} \frac{\tau_{n,p}^2}{\big(\tau_{n,p} - \bar{t}_n\big)^2} = \frac{1}{c} \; . \tag{15.39}$$

From $\bar{t}_n \sim \beta_1 p$ and $\tau_{n,i} \le \overline{h}$ (for $i = 1, \dots, p-1$) we deduce

$$\frac{1}{p} \sum_{i=1}^{p-1} \frac{\tau_{n,i}^2}{\big(\tau_{n,i} - \bar{t}_n\big)^2} \sim \frac{\overline{h}^2}{\beta_1^2 p^2} \longrightarrow 0 \; ; \tag{15.40}$$

therefore,

$$\frac{1}{p} \frac{\tau_{n,p}^2}{\big(\tau_{n,p} - \bar{t}_n\big)^2} \longrightarrow \frac{1}{c}$$

$$\frac{1}{n} \frac{\tau_{n,p}^2}{\big(\tau_{n,p} - \bar{t}_n\big)^2} \longrightarrow 1$$

$$\frac{\bar{t}_n - \tau_{n,p}}{\tau_{n,p}/\sqrt{n}} \longrightarrow 1. \;■$$

Lemma 15.7. *Define*

$$\overline{\lambda}_n := \sup \left\{ t \in \mathbb{R} : F_{n,p}^{\boldsymbol{\tau}_n}(t) < 1 \right\} \; . \tag{15.41}$$

Then under Assumptions 3.1, 3.2.a–c, and 3.2.f,

$$\overline{\lambda}_n - \tau_{n,p} \sim 2 \frac{\tau_{n,p}}{\sqrt{n}} \; . \tag{15.42}$$

Proof of Lemma 15.7. Equation (13) of Mestre (2008) gives

$$\overline{\lambda}_n = \overline{t}_n - c\,\overline{t}_n \frac{1}{p}\sum_{i=1}^{p} \frac{\tau_{n,i}}{\tau_{n,i} - \overline{t}_n} \ . \tag{15.43}$$

This is equivalent to plugging $m = -1/\overline{t}_n$ into Equation (1.6) of Silverstein and Choi (1995). From Equation (15.43) we deduce

$$\frac{\overline{\lambda}_n}{\overline{t}_n} - 1 = c\frac{1}{p}\frac{\tau_{n,p}}{\overline{t}_n - \tau_{n,p}} + c\frac{1}{p}\sum_{i=1}^{p-1} \frac{\tau_{n,i}}{\overline{t}_n - \tau_{n,i}} \ . \tag{15.44}$$

Lemma 15.6 enables us to approximate the first term on the right-hand side by

$$c\frac{1}{p}\frac{\tau_{n,p}}{\underline{t}_n - \tau_{n,p}} \sim \frac{p}{n} \times \frac{1}{p} \times \sqrt{n} = \frac{1}{\sqrt{n}} \ . \tag{15.45}$$

Since $\tau_{n,i} \leq \overline{h} < \overline{t}_n$, the second term is bounded by

$$0 \leq c\frac{1}{p}\sum_{i=1}^{p-1} \frac{\tau_{n,i}}{\overline{t}_n - \tau_{n,i}} \leq c\frac{\overline{h}}{\overline{t}_n - \overline{h}} \sim c\frac{\overline{h}}{\beta_1 p} \ ; \tag{15.46}$$

therefore, it is negligible with respect to the first term. We conclude by remarking that

$$\begin{aligned} \frac{\overline{\lambda}_n}{\overline{t}_n} - 1 &\sim \frac{1}{\sqrt{n}} \\ \overline{\lambda}_n - \overline{t}_n &\sim \frac{\overline{t}_n}{\sqrt{n}} \sim \frac{\tau_{n,p}}{\sqrt{n}} \\ \overline{\lambda}_n - \tau_{n,p} = \left(\overline{\lambda}_n - \overline{t}_n\right) + \left(\overline{t}_n - \tau_{n,p}\right) &\sim 2\frac{\tau_{n,p}}{\sqrt{n}}. \ \blacksquare \end{aligned}$$

The observation that $\underline{\lambda}_n \leq q_{n,p}^{p}(\boldsymbol{\tau}_n) \leq \overline{\lambda}_n$ together with Lemmas 15.4 and 15.7 establishes Proposition 7.4. ■

15.6 Proof of Lemma 7.1

For ease of notation, let us denote $\widetilde{\varphi}_n(\lambda_{n,i})$ by $\widetilde{\varphi}_{n,i}$ in this proof only.

$$\begin{aligned} \mathcal{L}_n^S(\Sigma_n, \widetilde{S}_n) &= \frac{1}{p}\sum_{i=1}^{p} \widetilde{\varphi}_{n,i} \cdot u_{n,i}\Sigma_n^{-1}u_{n,i} + \frac{1}{p}\sum_{i=1}^{p}\log(\tau_{n,i}) - \frac{1}{p}\sum_{i=1}^{p}\log(\widetilde{\varphi}_{n,i}) - 1 \\ \frac{\partial \mathcal{L}_n^S(\Sigma_n, \widetilde{S}_n)}{\partial \widetilde{\varphi}_{n,i}} &= \frac{1}{p}u_{n,i}\Sigma_n^{-1}u_{n,i} - \frac{1}{p}\frac{1}{\widetilde{\varphi}_{n,i}} \end{aligned}$$

The first-order condition is

$$\frac{\partial \mathcal{L}_n^S(\Sigma_n, \widetilde{S}_n)}{\partial \widetilde{\varphi}_{n,i}} = 0 \Longleftrightarrow \widetilde{\varphi}_{n,i} = \frac{1}{u_{n,i}\Sigma_n^{-1}u_{n,i}}. \ \blacksquare$$

15.7 Proof of Proposition 7.5

As in the proof of Proposition 7.3 above, let V_n denote a matrix of eigenvectors of Σ_n arranged to match the nondescending order of the eigenvalues $\boldsymbol{\tau}_n = (\tau_{n,1}, \ldots, \tau_{n,p})$, and let $v_{n,i}$ denote its ith column vector ($i = 1, \ldots, p$). The matrix Σ_n^A defined in Equation (15.2) is a rank-degenerate version of the population covariance matrix where all bulk eigenvalues have been neglected. The sample covariance matrix that corresponds to Σ_n^A is

$$S_n^A := n^{-1}\sqrt{\Sigma_n^A}X_n'X_n\sqrt{\Sigma_n^A} \ . \tag{15.47}$$

It admits a spectral decomposition on the same orthonormal basis as Σ_n and Σ_n^A:

$$S_n^A = V_n \times \mathsf{Diag}\big(\lambda_{n,1}^A, \ldots, \lambda_{n,p}^A\big) \times V_n' \ , \tag{15.48}$$

with all eigenvalues equal to zero except for $\lambda_{n,p}^A = n^{-1}\tau_{n,p} \cdot v_{n,p}'X_n'X_n v_{n,p}$. Viewing S_n as a perturbation of S_n^A, Equation (5.1) of Meyer and Stewart (1988) gives the approximation

$$\forall i = 1, \ldots, p-1 \qquad u_{n,p}'v_{n,i} = \frac{v_{n,i}'\left(S_n - S_n^A\right)v_{n,p}}{\lambda_{n,p}^A - \lambda_{n,i}^A} + O\left(\left(\frac{\tau_{n,i}}{\tau_{n,p}}\right)^2\right) \ , \tag{15.49}$$

from which we deduce

$$\tau_{n,p}\sum_{i=1}^{p-1}\frac{\left(u_{n,p}'v_{n,i}\right)^2}{\tau_{n,i}} = \sum_{i=1}^{p-1}\frac{\tau_{n,p}}{\tau_{n,i}}\left[\frac{v_{n,i}'\left(S_n - S_n^A\right)v_{n,p}}{\lambda_{n,p}^A - \lambda_{n,i}^A}\right]^2 + O\left(\frac{1}{p}\right) \ . \tag{15.50}$$

Note that $\forall i = 1, \ldots, p-1$, $v_{n,i}'S_n^A v_{n,p} = 0$, and $v_{n,i}'S_n v_{n,p} = n^{-1}\sqrt{\tau_{n,i}\,\tau_{n,p}} \cdot v_{n,i}'X_n'X_n v_{n,p}$, therefore this expression simplifies to

$$\tau_{n,p}\sum_{i=1}^{p-1}\frac{\left(u_{n,p}'v_{n,i}\right)^2}{\tau_{n,i}} = \frac{1}{n}\sum_{i=1}^{p-1}\frac{\left(v_{n,i}'X_n'X_n v_{n,p}\right)^2/n}{\left(v_{n,p}'X_n'X_n v_{n,p}/n\right)^2} + O\left(\frac{1}{p}\right) \ . \tag{15.51}$$

By the law of large numbers, $(p-1)^{-1}\sum_{i=1}^{p-1}\left(v_{n,i}'X_n'X_n v_{n,p}\right)^2/n \overset{\text{a.s.}}{\to} 1$ and $v_{n,p}'X_n'X_n v_{n,p}/n \overset{\text{a.s.}}{\to} 1$; therefore,

$$\tau_{n,p}\sum_{i=1}^{p-1}\frac{\left(u_{n,p}'v_{n,i}\right)^2}{\tau_{n,i}} \xrightarrow{\text{a.s.}} c \ . \tag{15.52}$$

Given that $\sum_{i=1}^{p}\left(u_{n,p}'v_{n,i}\right)^2 = 1$, we have $(u_{n,p}'v_{n,p})^2 \xrightarrow{\text{a.s.}} 1$. This enables us to conclude that

$$\tau_{n,p} \cdot u_{n,p}'\Sigma_n^{-1}u_{n,p} = \left(u_{n,p}'v_{n,p}\right)^2 + \tau_{n,p}\sum_{i=1}^{p-1}\frac{\left(u_{n,p}'v_{n,i}\right)^2}{\tau_{n,i}} \xrightarrow{\text{a.s.}} 1 + c \ . \ \blacksquare$$

15.8 Proof of Proposition 7.6

Lemma 15.8. *Let*

$$\boldsymbol{\tau}_n^A := (\underbrace{0, \ldots, 0}_{p-1 \text{ times}}, \tau_{n,p}) \ . \tag{15.53}$$

Then under Assumptions 3.1, 3.2.a–c, and 3.2.f, $1 + \lambda_{n,p}\,\breve{m}_{n,p}^{\boldsymbol{\tau}_n^A}(\lambda_{n,p}) \xrightarrow{\text{a.s.}} 0$.

Proof of Lemma 15.8. By taking the limit of Equation (5.7) as $m \in \mathbb{C}^+$ approaches the real line, we find that for all $\lambda \in (0, +\infty)$, $m := \breve{m}_{n,p}^{\boldsymbol{\tau}_n^A}(\lambda)$ is the unique solution in $\mathbb{C}^+ \cup \mathbb{R}$ to the equation

$$m = -\frac{p-1}{p\lambda} + \frac{1}{p} \frac{1}{\tau_{n,p}\left(1 - \frac{p}{n} - \frac{p}{n}\lambda\, m\right) - \lambda} . \tag{15.54}$$

With the change of variables $\widetilde{m} := pm + (p-1)/\lambda$, Equation (15.54) becomes

$$\widetilde{m} = \frac{1}{\tau_{n,p}\left(1 - \frac{1}{n} - \frac{1}{n}\lambda\, \widetilde{m}\right) - \lambda} \tag{15.55}$$

$$\tau_{n,p}\frac{1}{n}\lambda\widetilde{m}^2 + \left(\tau_{n,p}\frac{1}{n} + \lambda - \tau_{n,p}\right)\widetilde{m} + 1 = 0 . \tag{15.56}$$

(15.56) is a classic quadratic equation whose discriminant is $\Delta = \left(\tau_{n,p}n^{-1} + \lambda - \tau_{n,p}\right)^2 - 4\tau_{n,p}n^{-1}$. In turn, the Equation $\Delta = 0$ is itself a quadratic equation in λ:

$$\lambda^2 - 2\tau_{n,p}\left(\frac{1}{n} + 1\right)\lambda + \tau_{n,p}^2\left(\frac{1}{n} - 1\right)^2 = 0 . \tag{15.57}$$

It admits two distinct real, positive solutions: $\lambda = \tau_{n,p}\left(1 \pm n^{-1/2}\right)^2$. This enables us to factorize the discriminant Δ into

$$\Delta = \left[\lambda - \tau_{n,p}\left(1 + \frac{1}{\sqrt{n}}\right)^2\right] \times \left[\lambda - \tau_{n,p}\left(1 - \frac{1}{\sqrt{n}}\right)^2\right] . \tag{15.58}$$

This factorization shows that Equation (15.56) admits a solution in $\mathbb{C}^+$ if and only if $\lambda \in \left(\tau_{n,p}(1 + n^{-1/2})^2,\ \tau_{n,p}(1 + n^{-1/2})^2\right)$. Over this interval, the solution with positive imaginary part is

$$\widetilde{m} = \frac{\tau_{n,p}(1 - n^{-1}) - \lambda + i \cdot \sqrt{\left[\lambda - \tau_{n,p}(1 - n^{-1/2})^2\right] \times \left[\tau_{n,p}(1 + n^{-1/2})^2 - \lambda\right]}}{2\,\tau_{n,p}\,\lambda/n} . \tag{15.59}$$

This is none other than the Stieltjes transform of the celebrated Marčenko-Pastur (1967) distribution with scale parameter $\tau_{n,p}$ and concentration parameter $1/n$. Changing back to the original variable m, we obtain the following solution for Equation (15.54):

$$m = -\frac{p-1}{p\lambda} + \frac{\tau_{n,p}(1 - n^{-1}) - \lambda + i \cdot \sqrt{\left[\lambda - \tau_{n,p}(1 - n^{-1/2})^2\right] \times \left[\tau_{n,p}(1 + n^{-1/2})^2 - \lambda\right]}}{2\,\tau_{n,p}\,\lambda\, p/n} , \tag{15.60}$$

for all $\lambda \in \left[\tau_{n,p}(1 - n^{-1/2})^2, \tau_{n,p}(1 + n^{-1/2})^2\right]$. Note that this closed interval, along with zero, constitutes the support of $F_{n,p}^{\boldsymbol{\tau}_n^A}$. The general solution for all $\lambda > 0$ can be expressed concisely by introducing the function

$$\forall \lambda \in (0, +\infty) \qquad u(\lambda) := \begin{cases} -1 & \text{if } 0 < \lambda < \tau_{n,p}(1 - n^{-1/2})^2, \\ i & \text{if } \tau_{n,p}(1 - n^{-1/2})^2 \le \lambda \le \tau_{n,p}(1 + n^{-1/2})^2, \\ 1 & \text{if } \tau_{n,p}(1 + n^{-1/2})^2 < \lambda . \end{cases} \tag{15.61}$$

It is

$$\breve{m}_{n,p}^{\boldsymbol{\tau}_n^A}(\lambda) = -\frac{p-1}{p\lambda} + \frac{\tau_{n,p}(1-n^{-1}) - \lambda + u(\lambda)\sqrt{\left|\lambda - \tau_{n,p}(1-n^{-1/2})^2\right| \times \left|\tau_{n,p}(1+n^{-1/2})^2 - \lambda\right|}}{2\,\tau_{n,p}\,\lambda\,p/n}\,,$$

from which we deduce, after simplification,

$$1+\lambda\,\breve{m}_{n,p}^{\boldsymbol{\tau}_n^A}(\lambda) = \frac{1}{2p} + \frac{n}{2p}\left(1-\frac{\lambda}{\tau_{n,p}}\right) + \frac{n\,u(\lambda)}{2p}\sqrt{\left|\frac{\lambda}{\tau_{n,p}} - \left(1-\frac{1}{\sqrt{n}}\right)^2\right| \times \left|\left(1+\frac{1}{\sqrt{n}}\right)^2 - \frac{\lambda}{\tau_{n,p}}\right|}\,. \tag{15.62}$$

Lemma 15.8 then follows by setting $\lambda = \lambda_{n,p}$ in Equation (15.62) and using Proposition 7.3. ■

Lemma 15.9.

$$\forall \lambda \in (0,+\infty) \qquad \left|1+\lambda\,\breve{m}_{n,p}^{\boldsymbol{\tau}_n}(\lambda)\right| \leq \sqrt{\frac{n}{p}}\,. \tag{15.63}$$

Proof of Lemma 15.9. Section 2.2 of Ledoit and Wolf (2012) defines the ancillary function

$$\underline{m}_{n,p}^{\boldsymbol{\tau}_n}(z) := \frac{p-n}{nz} + \frac{p}{n}\,m_{n,p}^{\boldsymbol{\tau}_n}(z), \qquad \forall z \in \mathbb{C}^+\,. \tag{15.64}$$

Call its image $\underline{m}_{n,p}^{\boldsymbol{\tau}_n}(\mathbb{C}^+)$. Equation (1.4) of Silverstein and Choi (1995) states that the function $\underline{m}_{n,p}^{\boldsymbol{\tau}_n}$ has a unique inverse on $\mathbb{C}^+$ given by

$$\underline{z}_{n,p}^{\boldsymbol{\tau}_n}(m) := -\frac{1}{m} + \frac{1}{n}\sum_{i=1}^{p}\frac{\tau_{n,i}}{1+m\,\tau_{n,i}}, \qquad \forall m \in \underline{m}_{n,p}^{\boldsymbol{\tau}_n}\left(\mathbb{C}^+\right)\,. \tag{15.65}$$

The change of variables $m = \underline{m}_{n,p}^{\boldsymbol{\tau}_n}(z) \Longleftrightarrow z = \underline{z}_{n,p}^{\boldsymbol{\tau}_n}(m)$ yields

$$\forall z \in \mathbb{C}^+ \qquad 1+zm_{n,p}^{\boldsymbol{\tau}_n}(z) = \frac{n}{p}\left[1+z\underline{m}_{n,p}^{\boldsymbol{\tau}_n}(z)\right] = \frac{n}{p}\left[1+\underline{z}_{n,p}^{\boldsymbol{\tau}_n}(m)m\right] = \frac{1}{p}\sum_{i=1}^{p}\frac{m\tau_{n,i}}{1+m\tau_{n,i}}\,. \tag{15.66}$$

By Jensen's inequality,

$$\forall m \in \underline{m}_{n,p}^{\boldsymbol{\tau}_n}\left(\mathbb{C}^+\right) \qquad \left(\frac{1}{p}\sum_{i=1}^{p}\mathsf{Re}\left[\frac{m\tau_{n,i}}{1+m\tau_{n,i}}\right]\right)^2 \leq \frac{1}{p}\sum_{i=1}^{p}\left(\mathsf{Re}\left[\frac{m\tau_{n,i}}{1+m\tau_{n,i}}\right]\right)^2 \tag{15.67}$$

$$\left(\frac{1}{p}\sum_{i=1}^{p}\mathsf{Im}\left[\frac{m\tau_{n,i}}{1+m\tau_{n,i}}\right]\right)^2 \leq \frac{1}{p}\sum_{i=1}^{p}\left(\mathsf{Im}\left[\frac{m\tau_{n,i}}{1+m\tau_{n,i}}\right]\right)^2\,. \tag{15.68}$$

Adding these two equations, we obtain

$$\forall m \in \underline{m}_{n,p}^{\boldsymbol{\tau}_n}\left(\mathbb{C}^+\right) \qquad \left|\frac{1}{p}\sum_{i=1}^{p}\frac{m\tau_{n,i}}{1+m\tau_{n,i}}\right|^2 \leq \frac{1}{p}\sum_{i=1}^{p}\left|\frac{m\tau_{n,i}}{1+m\tau_{n,i}}\right|^2\,. \tag{15.69}$$

Since $\underline{z}_{n,p}^{\boldsymbol{\tau}_n}(m) \in \mathbb{C}^+$ for all $m \in \underline{m}_{n,p}^{\boldsymbol{\tau}_n}\left(\mathbb{C}^+\right)$, we have:

$$\forall m \in \underline{m}_{n,p}^{\boldsymbol{\tau}_n}\left(\mathbb{C}^+\right) \qquad \mathsf{Im}\left[\underline{z}_{n,p}^{\boldsymbol{\tau}_n}(m)\right] > 0\,. \tag{15.70}$$

Let $m_1 := \mathsf{Re}[m]$ and $m_2 := \mathsf{Im}[m]$. Equation (1.3) of Silverstein and Choi (1995) implies that $\underline{m}_{n,p}^{\boldsymbol{\tau}_n}(\mathbb{C}^+) \subset \mathbb{C}^+$; therefore, $m_2 > 0$. This enables us to deduce from Equation (15.70) that

$$\forall m \in \underline{m}_{n,p}^{\boldsymbol{\tau}_n}(\mathbb{C}^+) \qquad \mathsf{Im}\left[-\frac{1}{m_1 + i m_2} + \frac{1}{n}\sum_{i=1}^{p} \frac{\tau_{n,i}}{1 + (m_1 + i m_2)\tau_{n,i}}\right] > 0$$

$$\frac{m_2}{m_1^2 + m_2^2} - \frac{1}{n}\sum_{i=1}^{p} \frac{m_2 \tau_{n,i}^2}{(1 + m_1\tau_{n,i})^2 + m_2 \tau_{n,i}^2} > 0$$

$$\frac{1}{p}\sum_{i=1}^{p} \frac{\tau_{n,i}^2 (m_1^2 + m_2^2)}{(1 + m_1\tau_{n,i})^2 + m_2 \tau_{n,i}^2} < \frac{n}{p}$$

$$\frac{1}{p}\sum_{i=1}^{p} \left|\frac{m\tau_{n,i}}{1 + m\tau_{n,i}}\right|^2 < \frac{n}{p} \; . \tag{15.71}$$

Putting together Equations (15.66), (15.69), and (15.71) yields $\forall z \in \mathbb{C}^+ \quad \left|1 + z m_{n,p}^{\boldsymbol{\tau}_n}(z)\right|^2 < n/p$. Lemma 15.9 then follows from taking the limit as $z \in \mathbb{C}^+$ goes to $\lambda \in (0, +\infty)$. ■

By taking the limit of Equation (5.7) as $m \in \mathbb{C}^+$ approaches the real line, we find that for all $\lambda \in (0, +\infty)$, $m := \breve{m}_{n,p}^{\boldsymbol{\tau}_n}(\lambda)$ is the unique solution in $\mathbb{C}^+ \cup \mathbb{R}$ to the equation

$$m = \frac{1}{p}\sum_{i=1}^{p-1} \frac{1}{\tau_{n,i}\left(1 - \dfrac{p}{n} - \dfrac{p}{n}\lambda m\right) - \lambda} + \frac{1}{p}\,\frac{1}{\tau_{n,p}\left(1 - \dfrac{p}{n} - \dfrac{p}{n}\lambda m\right) - \lambda} \; . \tag{15.72}$$

Comparing Equation (15.72) with Equation (15.54) yields

$$\forall \lambda \in (0, +\infty) \qquad \lambda\left[\breve{m}_{n,p}^{\boldsymbol{\tau}_n}(\lambda) - \breve{m}_{n,p}^{\boldsymbol{\tau}_n^A}(\lambda)\right] = \frac{1}{p}\sum_{i=1}^{p-1} \frac{\tau_{n,i}\left[1 - \dfrac{p}{n} - \dfrac{p}{n}\lambda\, \breve{m}_{n,p}^{\boldsymbol{\tau}_n}(\lambda)\right]}{\tau_{n,i}\left[1 - \dfrac{p}{n} - \dfrac{p}{n}\lambda\, \breve{m}_{n,p}^{\boldsymbol{\tau}_n}(\lambda)\right] - \lambda} \; . \tag{15.73}$$

Remember that by Assumption 3.2.f, there exists $\overline{h} > 0$ such that $0 \le \tau_{n,1} \le \cdots \le \tau_{n,p-1} \le \overline{h}$ for all n large enough. Furthermore by Assumption 3.1 there exists $\bar{c}$ such that $p/n \le \bar{c}$ for all n large enough. Lemma 15.9 then yields the following bound for sufficiently large n:

$$\forall \lambda \in \left(\overline{h}\left(1 + \sqrt{\bar{c}}\right), +\infty\right) \qquad \lambda\left|\breve{m}_{n,p}^{\boldsymbol{\tau}_n}(\lambda) - \breve{m}_{n,p}^{\boldsymbol{\tau}_n^A}(\lambda)\right| \le \frac{\overline{h}\left(1 + \sqrt{\bar{c}}\right)}{\lambda - \overline{h}\left(1 + \sqrt{\bar{c}}\right)} \; . \tag{15.74}$$

By Proposition 7.3, this implies

$$\lambda_{n,p}\left[\breve{m}_{n,p}^{\boldsymbol{\tau}_n}(\lambda_{n,p}) - \breve{m}_{n,p}^{\boldsymbol{\tau}_n^A}(\lambda_{n,p})\right] \xrightarrow{\text{a.s.}} 0 \; . \tag{15.75}$$

Using Lemma 15.8, we obtain

$$1 + \lambda_{n,p}\, \breve{m}_{n,p}^{\boldsymbol{\tau}_n}(\lambda_{n,p}) \xrightarrow{\text{a.s.}} 0 \; , \tag{15.76}$$

from which we can finally conclude that

$$\frac{\lambda_{n,p}}{1 - \dfrac{p}{n} - 2\,\dfrac{p}{n}\,\lambda_{n,p}\, \mathsf{Re}\left[\breve{m}_{n,p}^{\boldsymbol{\tau}_n}(\lambda_{n,p})\right]} \overset{\text{a.s.}}{\sim} \frac{\lambda_{n,p}}{1 + \dfrac{p}{n}} \overset{\text{a.s.}}{\sim} \frac{\tau_{n,p}}{1 + c} . \; ■ \tag{15.77}$$

References

Abramowitz, M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, volume 55. Dover publications.

Bai, Z. D. and Silverstein, J. W. (1998). No eigenvalues outside the suppport of the limiting spectral distribution of large-dimensional random matrices. *Annals of Probability*, 26(1):316–345.

Bai, Z. D. and Silverstein, J. W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Annals of Probability*, 32(1A):553–605.

Franklin, J. N. (2000). *Matrix Theory*. Dover Publications, Mineola, N.Y.

Henrici, P. (1988). *Applied and Computational Complex Analysis*, volume 1. Wiley, New York.

Ledoit, O. and Péché, S. (2011). Eigenvectors of some large sample covariance matrix ensembles. *Probability Theory and Related Fields*, 150(1–2):233–264.

Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411.

Ledoit, O. and Wolf, M. (2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices. *Annals of Statistics*, 40(2):1024–1060.

Ledoit, O. and Wolf, M. (2015). Spectrum estimation: a unified framework for covariance matrix estimation and PCA in large dimensions. *Journal of Multivariate Analysis*, 139(2):360–384.

Marčenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Sbornik: Mathematics*, 1(4):457–483.

Mestre, X. (2008). On the asymptotic behavior of the sample estimates of eigenvalues and eigenvectors of covariance matrices. *IEEE Transactions on Signal Processing*, 56(11):5353–5368.

Meyer, C. D. and Stewart, G. W. (1988). Derivatives and perturbations of eigenvectors. *SIAM Journal on Numerical Analysis*, 25(3):679–691.

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.