Supplementary material for Nonlinear Shrinkage Estimation of Large Integrated Covariance Matrices

1 Market Trade Simulations

We simulate intra-day price data and construct minimum variance portfolio for each of the seven methods compared in Section 4·2.

Following Barndorff-Nielsen et al. (2011) and Fan et al. (2012), we simulate $p = 100$ and $p = 200$ stock prices for 210 days, which is 42 weeks of trading days. We use $X_t^{o(i)} = X_t^{(i)} + \epsilon_t^{(i)}$, where $X_t^{(i)}$ is the underlying log-price, and $\epsilon_t^{(i)}$ models the market microstructure noise, with $\epsilon_t^{(i)} \sim N(0, 0.0005^2)$ and are assumed to be independent of each other. The log-price $X_t^{(i)}$ is generated by

$$dX_t^{(i)} = \mu^{(i)}dt + \rho^{(i)}\sigma_t^{(i)}dB_t^{(i)} + \{1 - (\rho^{(i)})^2\}^{1/2}\sigma_t^{(i)}dW_t + \nu^{(i)}dZ_t \quad (i = 1, \ldots, 100),$$

where $\{W_t\}$, $\{Z_t\}$ and the $\{B_t^{(i)}\}$'s are all independent standard Brownian motions. The processes $\{Z_t\}$ and $\{W_t\}$ play the role of factors, when a real market usually has a market factor in the asset returns. The spot volatility $\sigma_t^{(i)} = \exp(\varrho_t^{(i)})$ follows the independent Ornstein-Uhlenbeck process

$$d\varrho_t^{(i)} = \alpha^{(i)}(\beta_0^{(i)} - \varrho_t^{(i)})dt + \beta_1^{(i)}dU_t^{(i)},$$

where the $\{U_t^{(i)}\}$'s are independent standard Brownian motions. We use $(\mu^{(i)}, \beta_0^{(i)}, \beta_1^{(i)}, \alpha^{(i)}, \rho^{(i)}) = (0.03x_1^{(i)}, -x_2^{(i)}, 0.75x_3^{(i)}, -1/40x_4^{(i)}, -0.7)$ and $\nu^{(i)} = exp(\beta_0^{(i)})$, where the $x_j^{(i)}$'s are independent and uniformly distributed on the interval $[0.7, 1.3]$. The initial value of each log-price is set at $X_0^{(i)} = 1$ and the starting spot volatility $\varrho_0^{(i)} = 0$.

We simulate the trading times independently from the price data assuming the transaction times for each stock follow independent Poisson processes with rates $\lambda_1, \ldots, \lambda_{100}$ respectively, where $\lambda_i = 0.01i(23400)$ $(i = 1, \ldots, 100)$, because normal trading time for one day has 23400 seconds.

After simulating the data, we split a trading day into 15-minute intervals, and set the price data for each stock at the end of each interval as the price observed at the trade right before the end of the interval. The data is used to calculate various integrated covariance matrix estimators. At the start, we invest 1 unit of capital using (14) constructed from different estimators of $\mathbf{\Sigma}_p$. We re-evaluate the portfolio weights every week, using either the past 8 weeks or 4 weeks of data as the training set. To gauge performances, we compute

Table 1: Market trades simulation results. Standard errors for weekly maximum exposures are subscripted. Graphical lasso is omitted because of non-convergence issues. RCV, realized covariance; Grand Avg, grand average; NONLIN, nonlinear shrinkage; POET, principal orthogonal complement thresholding; SCAD, adaptive thresholding with the smoothly clipped absolute deviation penalty.

	$p = 100$					$p = 200$		
	Annualized Return	Annualized Std. Dev.	Sharpe Ratio	Maximum Exposure (%)	Actual Risk	Sharpe Ratio	Maximum Exposure (%)	Actual Risk
	Weekly rebalancing with 4-week training windows							
Theoretical	14.6	15.5	0.94	$3.3_{1.2}$	13.1	0.91	$2.9_{1.0}$	13.1
RCV	23.1	15.4	1.50	$14.4_{5.3}$	14.5	0.92	$18.2_{6.8}$	16.6
Proposed	21.0	15.3	1.37	$7.5_{2.6}$	13.4	0.98	$9.1_{3.8}$	13.6
Banding	23.2	15.2	1.53	$8.3_{3.0}$	13.6	1.12	$9.8_{3.8}$	14.2
Grand Avg	21.1	15.2	1.38	$7.5_{2.6}$	13.4	0.98	$9.2_{3.8}$	13.6
NONLIN	21.1	15.2	1.39	$7.4_{2.6}$	13.4	1.03	$8.7_{3.6}$	13.6
POET	21.2	15.3	1.38	$7.6_{2.7}$	13.4	1.03	$7.9_{3.2}$	13.6
SCAD	21.1	15.0	1.41	$8.1_{3.0}$	13.5	1.20	$10.8_{4.1}$	14.3
	Weekly rebalancing with 8-week training windows							
Theoretical	21.2	15.8	1.34	$2.4_{1.2}$	13.1	1.56	$2.2_{1.1}$	13.1
RCV	25.9	15.6	1.66	$10.0_{5.1}$	13.8	1.14	$12.7_{6.9}$	14.6
Proposed	23.1	15.6	1.48	$6.1_{3.1}$	13.2	1.66	$8.6_{4.7}$	13.4
Banding	24.5	15.8	1.55	$5.8_{2.9}$	13.3	1.65	$7.2_{3.9}$	13.5
Grand Avg	22.5	15.6	1.44	$6.2_{3.1}$	13.2	1.65	$8.6_{4.7}$	13.4
NONLIN	22.8	15.7	1.46	$6.0_{3.0}$	13.2	1.65	$8.0_{4.5}$	13.4
POET	22.7	15.6	1.45	$5.6_{2.8}$	13.3	1.61	$6.8_{3.7}$	13.4
SCAD	24.2	15.8	1.53	$5.6_{2.8}$	13.3	1.65	$7.8_{4.3}$	13.5

the annualized return $\widehat{\mu}$ and the annualized out-of-sample standard deviation $\widehat{\sigma}$, defined similarly as those in Section 4·2, together with the Sharpe ratio $\widehat{\mu}/\widehat{\sigma}$ and the actual risk of portfolios.

From Table 1, all methods, except for the realized covariance, are similar in terms of the actual risk. The annualized returns are higher for all methods when using 8-week training windows instead of 4, while the annualized out-of-sample standard deviation or the actual risk do not change much. The maximum exposure level also decreased in general as we change from 4-week training windows to 8-week. Overall, all measures from all methods are not too different from one another. The results in Section 4·2 in the paper are quite different.

2 Proof of Results

Proof of Lemma 1. Since M is finite, if we can show

$$\big\|\mathrm{diag}(\mathbf{P}_1^{(j)\mathrm{T}}\widetilde{\mathbf{\Phi}}_2^{(j)}\mathbf{P}_1^{(j)})\mathrm{diag}^{-1}(\mathbf{P}_1^{(j)\mathrm{T}}\mathbf{\Phi}\mathbf{P}_1^{(j)}) - 1\big\| \to 0$$

almost surely for $j = 1$, then the result of the lemma is established. Dropping the superscript (j) in all notations, we are proving

$$\big\|\mathrm{diag}(\mathbf{P}_1^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{P}_1)\mathrm{diag}^{-1}(\mathbf{P}_1^{\mathrm{T}}\mathbf{\Phi}\mathbf{P}_1) - 1\big\| \to 0$$

almost surely. Write $\mathbf{P}_1 = (\mathbf{p}_{11}, \ldots, \mathbf{p}_{1p})$, it is equivalent to showing

$$\max_{i=1,\ldots,p}\Big|\frac{\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{p}_{1i} - \mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}}{\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}}\Big| \to 0$$

almost surely.

We can write

$$\Delta\mathbf{X}_\ell = \Big(\int_{\tau_{n,\ell-1}}^{\tau_{n,\ell}} \gamma_t^2 dt\Big)^{1/2}\mathbf{\Phi}^{1/2}\mathbf{Z}_\ell, \tag{1}$$

where the $\mathbf{Z}_\ell$'s are independent of each other and each $\mathbf{Z}_\ell$ is a p dimensional vector with independent standard normal entries. Then we can decompose

$$\frac{\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{p}_{1i} - \mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}}{\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}} = I_{i1} + I_{i2} \quad (i = 1, \ldots, p),$$

where

$$I_{i1} = \frac{\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{p}_{1i} - n_2^{-1}\sum_{\ell\in I_2}(\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}^{1/2}\mathbf{Z}_\ell)^2}{\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}}, \quad I_{i2} = \frac{n_2^{-1}\sum_{\ell\in I_2}(\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}^{1/2}\mathbf{Z}_\ell)^2 - \mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}}{\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}}.$$

Since we can write

$$n_2^{-1}\sum_{\ell\in I_2}(\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}^{1/2}\mathbf{Z}_\ell)^2 = \mathbf{p}_{1i}^{\mathrm{T}}\big(n_2^{-1}\sum_{\ell\in I_2}\mathbf{\Phi}^{1/2}\mathbf{Z}_\ell\mathbf{Z}_\ell^{\mathrm{T}}\mathbf{\Phi}^{1/2}\big)\mathbf{p}_{1i} \quad (i = 1, \ldots, p),$$

and $n_2^{-1}\sum_{\ell\in I_2}\mathbf{\Phi}^{1/2}\mathbf{Z}_\ell\mathbf{Z}_\ell^{\mathrm{T}}\mathbf{\Phi}^{1/2}$ is a proper sample covariance matrix for estimating $\mathbf{\Phi}$, Lemma 1 of Lam (2016a) implies almost surely,

$$\max_{i=1,\ldots,p}|I_{i2}| \to 0.$$

Hence it remains to show that $\max_{i=1,\ldots,p}|I_{i1}| \to 0$ almost surely. To this end, consider

$$\begin{aligned}
\max_{i=1,\ldots,p}|I_{i1}| &= \max_{i=1,\ldots,p}\left|\frac{n_2^{-1}\sum_{\ell\in I_2}\frac{(\mathbf{p}_{1i}^{\mathrm{T}}\boldsymbol{\Phi}^{1/2}\mathbf{Z}_\ell)^2}{\mathbf{Z}_\ell^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{Z}_\ell/p} - \mathbf{p}_{1i}^{\mathrm{T}}\big(n_2^{-1}\sum_{\ell\in I_2}\boldsymbol{\Phi}^{1/2}\mathbf{Z}_\ell\mathbf{Z}_\ell^{\mathrm{T}}\boldsymbol{\Phi}^{1/2}\big)\mathbf{p}_{1i}}{\mathbf{p}_{1i}^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{p}_{1i}}\right| \\
&= \max_{i=1,\ldots,p}\left|\frac{n_2^{-1}\sum_{\ell\in I_2}\big(\frac{1}{\mathbf{Z}_\ell^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{Z}_\ell/p} - 1\big)\mathbf{p}_{1i}^{\mathrm{T}}\boldsymbol{\Phi}^{1/2}\mathbf{Z}_\ell\mathbf{Z}_\ell^{\mathrm{T}}\boldsymbol{\Phi}^{1/2}\mathbf{p}_{1i}}{\mathbf{p}_{1i}^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{p}_{1i}}\right| \\
&\le \max_{\ell\in I_2}\left|\frac{1}{\mathbf{Z}_\ell^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{Z}_\ell/p} - 1\right|(1+\max_{i=1,\ldots,p}|I_{i2}|) \to 0
\end{aligned}$$

almost surely, if we can show further that $\max_{\ell\in I_2}\left|(\mathbf{Z}_\ell^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{Z}_\ell/p)^{-1} - 1\right| \to 0$ almost surely.

Using Lemma 2.7 of Bai & Silverstein (1998),

$$E\{\mathbf{Z}_\ell^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{Z}_\ell - \mathrm{tr}(\boldsymbol{\Phi})\}^6 \le K_6\{E^3|z_{\ell,1}|^4\mathrm{tr}^3(\boldsymbol{\Phi}^2) + E|z_{\ell,1}|^{12}\mathrm{tr}(\boldsymbol{\Phi}^6)\},$$

where K_6 is a constant independent of ℓ, n and p. This implies that, since $\mathrm{tr}(\boldsymbol{\Phi}) = p$,

$$\begin{aligned}
E\left(\max_{\ell\in I_2}\left|\frac{\mathbf{Z}_\ell^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{Z}_\ell}{p} - 1\right|^6\right) &\le n_2K_6\Big\{E^3|z_{\ell,1}|^4\frac{\mathrm{tr}^3(\boldsymbol{\Phi}^2)}{p^6} + E|z_{\ell,1}|^{12}\frac{\mathrm{tr}(\boldsymbol{\Phi}^6)}{p^6}\Big\} \\
&= O(n_2p^{-3}).
\end{aligned}$$

The rate in the last line comes from Assumption 2 that $\boldsymbol{\Theta}_t\boldsymbol{\Theta}_t^{\mathrm{T}} = \gamma_t^2\boldsymbol{\Phi}$ has all its eigenvalues uniformly bounded away from 0 and infinity, so that $\mathrm{tr}^3(\boldsymbol{\Phi}^2) = O(p^3)$ and $\mathrm{tr}(\boldsymbol{\Phi}^6) = O(p)$, and the fact that the higher order moments of the $z_{\ell,1}$'s are all finite since they are all normally distributed.

Finally, since $n_2 = O(n^{1/2})$ and p has the same order as n, we have $O(n_2p^{-3}) = O(n^{-5/2})$. Since $\sum_{n\ge1} n^{-5/2} < \infty$, through the Borel-Cantelli lemma, we have proved that $\max_{\ell\in I_2}\left|\mathbf{Z}_\ell^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{Z}_\ell/p - 1\right| \to 0$ almost surely, meaning that $\max_{\ell\in I_2}\left|(\mathbf{Z}_\ell^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{Z}_\ell/p)^{-1} - 1\right| \to 0$ almost surely as well. This completes the proof of the lemma. $\square$

Proof of Theorem 1. Dropping the superscript (j), by Lemma 1, $\widehat{\boldsymbol{\Phi}} = \mathbf{P}_1\mathrm{diag}(\mathbf{P}_1^{\mathrm{T}}\widetilde{\boldsymbol{\Phi}}_2\mathbf{P}_1)\mathbf{P}_1^{\mathrm{T}}$ defined in (7) is almost surely positive definite since all its eigenvalues are almost surely $\mathbf{p}_{1i}^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{p}_{1i}$ $(i = 1,\ldots,p)$ as $n,p\to\infty$ such that $p/n \to c > 0$, and Assumption 2 ensures that these values are uniformly bounded away from 0 and infinity. Hence, the proof of the theorem completes if we can show that $\mathrm{tr}(\boldsymbol{\Sigma}_p^{\mathrm{RCV}})/p > 0$ uniformly almost surely.

To this end, consider

$$\begin{aligned}\left|\frac{\operatorname{tr}(\mathbf{\Sigma}_p^{\mathrm{RCV}})}{p}-\int_0^1\gamma_t^2dt\right| &= \left|\frac{\sum_{\ell=1}^n \Delta\mathbf{X}_\ell^{\mathrm{T}}\Delta\mathbf{X}_\ell}{p}-\int_0^1\gamma_t^2dt\right|\\ &= \left|\frac{\sum_{\ell=1}^n\int_{\tau_{n,\ell-1}}^{\tau_{n,\ell}}\gamma_t^2dt\mathbf{Z}_\ell^{\mathrm{T}}\mathbf{\Phi}\mathbf{Z}_\ell}{p}-\int_0^1\gamma_t^2dt\right|\\ &\leq \max_{\ell=1,\dots,n}\left|\frac{\mathbf{Z}_\ell^{\mathrm{T}}\mathbf{\Phi}\mathbf{Z}_\ell}{p}-1\right|\int_0^1\gamma_t^2dt.\end{aligned}$$

From the proof of $\max_{\ell\in I_2}\left|\mathbf{Z}_\ell^{\mathrm{T}}\mathbf{\Phi}\mathbf{Z}_\ell/p-1\right|\to 0$ almost surely in the last part of the proof of Lemma 1, we can replace n_2 there by n and conclude that

$$\begin{aligned}E\left(\max_{1\leq\ell\leq n}\left|\frac{\mathbf{Z}_\ell^{\mathrm{T}}\mathbf{\Phi}\mathbf{Z}_\ell}{p}-1\right|^6\right) &\leq nK_6\left\{E^3|z_{\ell,1}|^4\frac{\operatorname{tr}^3(\mathbf{\Phi}^2)}{p^6}+E|z_{\ell,1}|^{12}\frac{\operatorname{tr}(\mathbf{\Phi}^6)}{p^6}\right\}\\ &= O(np^{-3})=O(n^{-2}).\end{aligned}$$

Since $\sum_{n\geq 1}n^{-2}<\infty$, we can conclude that $\max_{\ell=1,\dots,n}\left|\mathbf{Z}_\ell^{\mathrm{T}}\mathbf{\Phi}\mathbf{Z}_\ell/p-1\right|\to 0$ almost surely. This shows that $\operatorname{tr}(\mathbf{\Sigma}_p^{\mathrm{RCV}})/p\to\int_0^1\gamma_t^2dt$ almost surely, which is uniformly larger than 0 by Assumption 1. This completes the proof of the theorem. $\square$

To prove Theorem 2, we first present the following lemma and its proof. We drop the superscript (j) in the definition of $\mathbf{P}_1^{(j)},\widetilde{\mathbf{\Phi}}_i^{(j)}(i=1,2;\ j=1,\dots,M)$ defined in (7) through the lemma, which is true for each $j=1,\dots,M$.

Lemma 2 *Let all the assumptions in Lemma* 1 *hold, together with Assumption* 4 *and* 5*. Denote by* $v_1^{(1)}\geq\cdots\geq v_p^{(1)}$ *the eigenvalues of* $\widetilde{\mathbf{\Phi}}_1$ *with corresponding eigenvectors* $\mathbf{p}_{11},\dots,\mathbf{p}_{1p}$*, and* $v_1\geq\cdots\geq v_p$ *the eigenvalues of* $\check{\mathbf{\Phi}}$ *defined in (*5*) with corresponding eigenvectors* $\mathbf{p}_1,\dots,\mathbf{p}_p$*. Then there exist positive functions* $\delta_1(\cdot)=\delta(\cdot)$ *and distribution functions* $F_1=F$ *such that almost surely,*

$$p^{-1}\sum_{j=1}^p\mathbf{1}_{\{x\geq v_j^{(1)}\}}\to F_1(x),\quad p^{-1}\sum_{j=1}^p\mathbf{1}_{\{x\geq v_j\}}\to F(x),$$

$$p^{-1}\sum_{j=1}^p\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}\mathbf{1}_{\{x\geq v_j^{(1)}\}}\to\int_{-\infty}^x\delta_1(\lambda)dF_1(\lambda),\quad p^{-1}\sum_{j=1}^p\mathbf{p}_i^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_i\mathbf{1}_{\{x\geq v_j\}}\to\int_{-\infty}^x\delta(\lambda)dF(\lambda).$$

We do not write down the explicit form of the functions $\delta_1(\cdot)$ and $\delta(\cdot)$ since they are not important for the proof of any subsequent theorems. Interested readers are referred to equation (2.7) and (2.9) of Lam (2016a).

Proof of Lemma 2. Write $\Delta \mathbf{X}_\ell = \big(\int_{\tau_{n,\ell-1}}^{\tau_{n,\ell}} \gamma_t^2 dt\big)^{1/2} \mathbf{\Phi}^{1/2}\mathbf{Z}_\ell$ as in (1). Define

$$\check{\mathbf{\Phi}}_{\text{sam}} = n^{-1}\sum_{\ell=1}^{n} \mathbf{\Phi}^{1/2}\mathbf{Z}_\ell\mathbf{Z}_\ell^{\mathrm{T}}\mathbf{\Phi}^{1/2}, \quad \widetilde{\mathbf{\Phi}}_{i,\text{sam}} = n_i^{-1}\sum_{\ell\in I_i} \mathbf{\Phi}^{1/2}\mathbf{Z}_\ell\mathbf{Z}_\ell^{\mathrm{T}}\mathbf{\Phi}^{1/2} \quad (i=1,2),$$

which are all proper sample covariance matrices. Let $v_{1,\text{sam}}^{(i)} \geq \cdots \geq v_{p,\text{sam}}^{(i)}$ be the eigenvalues of $\widetilde{\mathbf{\Phi}}_{i,\text{sam}}$ with corresponding eigenvectors $\mathbf{p}_{i1,\text{sam}}, \ldots, \mathbf{p}_{ip,\text{sam}}$. Also, let $v_{1,\text{sam}} \geq \cdots \geq v_{p,\text{sam}}$ be the eigenvalues of $\check{\mathbf{\Phi}}_{\text{sam}}$ with corresponding eigenvectors $\mathbf{p}_{1,\text{sam}}, \ldots, \mathbf{p}_{p,\text{sam}}$. Suppose we are able to show the following. Almost surely,

$$\begin{aligned}
p^{-1}\mathrm{tr}\big\{(\widetilde{\mathbf{\Phi}}_{1,\text{sam}} - z\mathbf{I}_p)^{-1}\big\} - p^{-1}\mathrm{tr}\big\{(\widetilde{\mathbf{\Phi}}_1 - z\mathbf{I}_p)^{-1}\big\} &\to 0 \quad (z\in\mathbb{C}^+),\\
p^{-1}\mathrm{tr}\big\{(\check{\mathbf{\Phi}}_{\text{sam}} - z\mathbf{I}_p)^{-1}\big\} - p^{-1}\mathrm{tr}\big\{(\check{\mathbf{\Phi}} - z\mathbf{I}_p)^{-1}\big\} &\to 0 \quad (z\in\mathbb{C}^+),\\
p^{-1}\mathrm{tr}\big\{(\widetilde{\mathbf{\Phi}}_{1,\text{sam}} - z\mathbf{I}_p)^{-1}\mathbf{\Phi}\big\} - p^{-1}\mathrm{tr}\big\{(\widetilde{\mathbf{\Phi}}_1 - z\mathbf{I}_p)^{-1}\mathbf{\Phi}\big\} &\to 0 \quad (z\in\mathbb{C}^+),\\
p^{-1}\mathrm{tr}\big\{(\check{\mathbf{\Phi}}_{\text{sam}} - z\mathbf{I}_p)^{-1}\mathbf{\Phi}\big\} - p^{-1}\mathrm{tr}\big\{(\check{\mathbf{\Phi}} - z\mathbf{I}_p)^{-1}\mathbf{\Phi}\big\} &\to 0 \quad (z\in\mathbb{C}^+).
\end{aligned} \tag{2}$$

The left hand side of the above can in fact be written as differences of Stieltjes transforms of certain nondecreasing functions. The differences in their inverse Stieltjes transforms must then converge to 0 almost surely as well, i.e., at the point of continuity x of these nondecreasing functions, almost surely,

$$\begin{aligned}
p^{-1}\sum_{j=1}^{p}\mathbf{1}_{\{x\geq v_{j,\text{sam}}^{(1)}\}} - p^{-1}\sum_{j=1}^{p}\mathbf{1}_{\{x\geq v_{j}^{(1)}\}} &\to 0,\\
p^{-1}\sum_{j=1}^{p}\mathbf{1}_{\{x\geq v_{j,\text{sam}}\}} - p^{-1}\sum_{j=1}^{p}\mathbf{1}_{\{x\geq v_{j}\}} &\to 0,\\
p^{-1}\sum_{j=1}^{p}\mathbf{p}_{1j,\text{sam}}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1j,\text{sam}}\mathbf{1}_{\{x\geq v_{j,\text{sam}}^{(1)}\}} - p^{-1}\sum_{j=1}^{p}\mathbf{p}_{1j}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1j}\mathbf{1}_{\{x\geq v_{j}^{(1)}\}} &\to 0,\\
p^{-1}\sum_{j=1}^{p}\mathbf{p}_{j,\text{sam}}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{j,\text{sam}}\mathbf{1}_{\{x\geq v_{j,\text{sam}}\}} - p^{-1}\sum_{j=1}^{p}\mathbf{p}_{j}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{j}\mathbf{1}_{\{x\geq v_{j}\}} &\to 0.
\end{aligned} \tag{3}$$

Interested readers are referred to Lam (2016a) or Ledoit & Péché (2011) for the definitions of Stieltjes transform and its inversion. Theorem 4 of Ledoit & Péché (2011) indicates that for all $x \in \mathbb{R}$, there exist functions $\delta_1(\cdot)$ and $\delta(\cdot)$ with corresponding distribution functions F_1 and F such that almost surely,

$$p^{-1}\sum_{j=1}^{p}\mathbf{1}_{\{\lambda\geq v_{j,\mathrm{sam}}^{(1)}\}} \to F_1(\lambda),$$

$$p^{-1}\sum_{j=1}^{p}\mathbf{1}_{\{\lambda\geq v_{j,\mathrm{sam}}\}} \to F(\lambda),$$

$$p^{-1}\sum_{j=1}^{p}\mathbf{p}_{1j,\mathrm{sam}}^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{p}_{1j,\mathrm{sam}}\mathbf{1}_{\{x\geq v_{j,\mathrm{sam}}^{(1)}\}} \to \int_{-\infty}^{x}\delta_1(\lambda)dF(\lambda),$$

$$p^{-1}\sum_{j=1}^{p}\mathbf{p}_{j,\mathrm{sam}}^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{p}_{j,\mathrm{sam}}\mathbf{1}_{\{x\geq v_{j,\mathrm{sam}}\}} \to \int_{-\infty}^{x}\delta(\lambda)dF(\lambda).$$

At the same time, since $p/n_1, p/n \to c > 0$, Theorem 4.1 of Bai & Silverstein (2010) says that the two limits F_1 and F are equal, and since $\delta_1(\cdot)$ and $\delta(\cdot)$ depend on F_1 and F and on the same $c > 0$, we must have $\delta_1(\cdot) = \delta(\cdot)$ also.

With the above, (3) immediately implies the results we need. Hence it remains to show (2).

To prove the first and third results of (2), consider

$$\begin{aligned}
&p^{-1}\mathrm{tr}\big\{(\widetilde{\boldsymbol{\Phi}}_1 - z\mathbf{I}_p)^{-1}\boldsymbol{\Phi}^k\big\}\\
&= p^{-1}\mathrm{tr}\big[\{\mathbf{I}_p - (\widetilde{\boldsymbol{\Phi}}_{1,\mathrm{sam}} - z\mathbf{I}_p)^{-1}(\widetilde{\boldsymbol{\Phi}}_{1,\mathrm{sam}} - \widetilde{\boldsymbol{\Phi}}_1)\}^{-1}(\widetilde{\boldsymbol{\Phi}}_{1,\mathrm{sam}} - z\mathbf{I}_p)^{-1}\boldsymbol{\Phi}^k\big]\\
&= p^{-1}\mathrm{tr}\big\{(\widetilde{\boldsymbol{\Phi}}_{1,\mathrm{sam}} - z\mathbf{I}_p)^{-1}\boldsymbol{\Phi}^k\big\} + R \quad (k = 0, 1),
\end{aligned}$$

where

$$R = \sum_{j\geq 1} p^{-1}\mathrm{tr}\big[\{(\widetilde{\boldsymbol{\Phi}}_{1,\mathrm{sam}} - z\mathbf{I}_p)^{-1}(\widetilde{\boldsymbol{\Phi}}_{1,\mathrm{sam}} - \widetilde{\boldsymbol{\Phi}}_1)\}^j(\widetilde{\boldsymbol{\Phi}}_{1,\mathrm{sam}} - z\mathbf{I}_p)^{-1}\boldsymbol{\Phi}^k\big] \quad (k = 0, 1).$$

The term R comes from a Neumann series expansion, which is valid since almost surely

$$\begin{aligned} r = \left\|(\widetilde{\mathbf{\Phi}}_{1,\mathrm{sam}} - z\mathbf{I}_p)^{-1}(\widetilde{\mathbf{\Phi}}_{1,\mathrm{sam}} - \widetilde{\mathbf{\Phi}}_1)\right\| &\le \left\|(\widetilde{\mathbf{\Phi}}_{1,\mathrm{sam}} - z\mathbf{I}_p)^{-1}\right\|\left\|\widetilde{\mathbf{\Phi}}_{1,\mathrm{sam}} - \widetilde{\mathbf{\Phi}}_1\right\| \\ &\le \frac{1}{|\mathrm{Im}(z)|}\max_{\ell\in I_1}\left|\frac{1}{\mathbf{Z}_\ell^{\mathrm{T}}\mathbf{\Phi}\mathbf{Z}_\ell/p} - 1\right|\left\|\widetilde{\mathbf{\Phi}}_{1,\mathrm{sam}}\right\| \\ &\le \frac{\left\|\mathbf{\Phi}^{1/2}\right\|^2}{|\mathrm{Im}(z)|}\max_{\ell\in I_1}\left|\frac{1}{\mathbf{Z}_\ell^{\mathrm{T}}\mathbf{\Phi}\mathbf{Z}_\ell/p} - 1\right|\left\|n_1^{-1}\sum_{\ell\in I_1}\mathbf{Z}_\ell\mathbf{Z}_\ell^{\mathrm{T}}\right\| \\ &\le \frac{\left\|\mathbf{\Phi}^{1/2}\right\|^2(1+\sqrt{c})^2}{|\mathrm{Im}(z)|}\max_{\ell\in I_1}\left|\frac{1}{\mathbf{Z}_\ell^{\mathrm{T}}\mathbf{\Phi}\mathbf{Z}_\ell/p} - 1\right| \to 0, \end{aligned}$$

where we used Lemma S.2 of Lam (2016b) to conclude $\left\|(\widetilde{\mathbf{\Phi}}_{1,\mathrm{sam}} - z\mathbf{I}_p)^{-1}\right\| \le 1/|\mathrm{Im}(z)|$. We used Theorem 5.11 of Bai & Silverstein (2010) for $\left\|n_1^{-1}\sum_{\ell\in I_1}\mathbf{Z}_\ell\mathbf{Z}_\ell^{\mathrm{T}}\right\| \le (1+\sqrt{c})^2$ almost surely, since $p/n_1 \to c > 0$. Finally, the term $\max_{\ell\in I_1}\left|(\mathbf{Z}_\ell^{\mathrm{T}}\mathbf{\Phi}\mathbf{Z}_\ell/p)^{-1} - 1\right| \to 0$ almost surely by the last part of the proof of Lemma 1. Hence almost surely,

$$|R| \le \sum_{j\ge 1} r^j\left\|\mathbf{\Phi}^k\right\|\left\|(\widetilde{\mathbf{\Phi}}_{1,\mathrm{sam}} - z\mathbf{I}_p)^{-1}\right\| \le \frac{r\left\|\mathbf{\Phi}^k\right\|}{(1-r)|\mathrm{Im}(z)|} \to 0,$$

so that we have proved the first and third results in (2). For the other two results, the proof follows exactly the same lines as before after replacing $\widetilde{\mathbf{\Phi}}_{1,\mathrm{sam}}$ by $\check{\mathbf{\Phi}}_{\mathrm{sam}}$ and $\widetilde{\mathbf{\Phi}}_1$ by $\check{\mathbf{\Phi}}$. This completes the proof of the lemma. □

Proof of Theorem 2. Observe that

$$\widehat{\mathbf{\Sigma}}_{m,M} = \frac{1}{M}\sum_{j=1}^{M}\widehat{\mathbf{\Sigma}}_m^{(j)}, \quad \widehat{\mathbf{\Sigma}}_m^{(j)} = \frac{\mathrm{tr}(\mathbf{\Sigma}_p^{\mathrm{RCV}})}{p}\widehat{\mathbf{\Phi}}^{(j)}.$$

If we can prove that $\mathrm{EffLoss}(\mathbf{\Sigma}_p, \widehat{\mathbf{\Sigma}}_m^{(j)}) \le 0$ almost surely, then we can follow exactly the same lines in the proof of Theorem 6 in Lam (2016a) to complete the proof Theorem 2.

Since M is finite, we drop the superscript (j) in $\widehat{\mathbf{\Sigma}}_m^{(j)}$, and define

$$\widehat{\mathbf{\Sigma}}_m = \frac{\mathrm{tr}(\mathbf{\Sigma}_p^{\mathrm{RCV}})}{p}\widehat{\mathbf{\Phi}} = \frac{\mathrm{tr}(\mathbf{\Sigma}_p^{\mathrm{RCV}})}{p}\mathbf{P}_1\mathrm{diag}(\mathbf{P}_1^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{P}_1)\mathbf{P}_1^{\mathrm{T}}.$$

It remains to show that $\mathrm{EffLoss}(\mathbf{\Sigma}_p, \widehat{\mathbf{\Sigma}}_m) \le 0$. Consider the Frobenius loss first. Define $\theta = \int_0^1 \gamma_t^2 dt$, $\widehat{\theta} = \mathrm{tr}(\mathbf{\Sigma}_p^{\mathrm{RCV}})/p$ and $\mathbf{\Phi}_{\mathrm{Ideal}} = \mathbf{P}\mathrm{diag}(\mathbf{P}^{\mathrm{T}}\mathbf{\Phi}\mathbf{P})\mathbf{P}^{\mathrm{T}}$. Then

$$\mathrm{EffLoss}(\mathbf{\Sigma}_p, \widehat{\mathbf{\Sigma}}_m) \le 1 - \left\{ \frac{p^{-1/2}\left\|(\widehat{\theta}-\theta)\widehat{\mathbf{\Phi}}\right\|_{\mathrm{F}}}{p^{-1/2}\theta\left\|\mathbf{\Phi}_{\mathrm{Ideal}} - \mathbf{\Phi}\right\|_{\mathrm{F}}} + \frac{p^{-1/2}\left\|\widehat{\mathbf{\Phi}} - \mathbf{\Phi}\right\|_{\mathrm{F}}}{p^{-1/2}\left\|\mathbf{\Phi}_{\mathrm{Ideal}} - \mathbf{\Phi}\right\|_{\mathrm{F}}} \right\}^{-2}, \tag{4}$$

where $p^{-1/2}\left\|(\widehat{\theta}-\theta)\widehat{\mathbf{\Phi}}\right\|_F \le |\widehat{\theta}-\theta| \max_{i=1,\dots,p} \mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{p}_{1i} \to 0$ almost surely by Lemma 1 and the proof of Theorem 1. Since we are equivalently assuming that $p^{-1/2}\left\|\mathbf{\Phi}_{\mathrm{Ideal}} - \mathbf{\Phi}\right\|_{\mathrm{F}} \not\to 0$, it remains to show that $p^{-1/2}\left\|\widehat{\mathbf{\Phi}} - \mathbf{\Phi}\right\|_{\mathrm{F}}/(p^{-1/2}\left\|\mathbf{\Phi}_{\mathrm{Ideal}} - \mathbf{\Phi}\right\|_{\mathrm{F}}) \to 1$ almost surely. Observe that

$$\frac{p^{-1}\left\|\widehat{\mathbf{\Phi}} - \mathbf{\Phi}\right\|_{\mathrm{F}}^2}{p^{-1}\left\|\mathbf{\Phi}_{\mathrm{Ideal}} - \mathbf{\Phi}\right\|_{\mathrm{F}}^2} = \frac{p^{-1}\sum_{i=1}^p(\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{p}_{1i} - \mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i})^2}{p^{-1}\left\|\mathbf{P}\mathrm{diag}(\mathbf{P}^{\mathrm{T}}\mathbf{\Phi}\mathbf{P})\mathbf{P}^{\mathrm{T}} - \mathbf{\Phi}\right\|_{\mathrm{F}}^2} + \frac{p^{-1}\left\|\mathbf{P}_1\mathrm{diag}(\mathbf{P}_1^{\mathrm{T}}\mathbf{\Phi}\mathbf{P}_1)\mathbf{P}_1^{\mathrm{T}} - \mathbf{\Phi}\right\|_{\mathrm{F}}^2}{p^{-1}\left\|\mathbf{P}\mathrm{diag}(\mathbf{P}^{\mathrm{T}}\mathbf{\Phi}\mathbf{P})\mathbf{P}^{\mathrm{T}} - \mathbf{\Phi}\right\|_{\mathrm{F}}^2}.$$

By Assumptions 4 and 5, and the results of Lemma 2, almost surely,

$$\begin{aligned} p^{-1}\left\|\mathbf{P}_1\mathrm{diag}(\mathbf{P}_1^{\mathrm{T}}\mathbf{\Phi}\mathbf{P}_1)\mathbf{P}_1^{\mathrm{T}} - \mathbf{\Phi}\right\|_{\mathrm{F}}^2 &= p^{-1}\mathrm{tr}(\mathbf{\Phi}^2) - p^{-1}\sum_{i=1}^p(\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i})^2 \\ &\to \int \tau^2 dH(\tau) - \int \delta_1^2(\lambda)dF_1(\lambda) \\ &= \int \tau^2 dH(\tau) - \int \delta^2(\lambda)dFs(\lambda), \end{aligned}$$

which is non-zero if $\mathbf{\Phi} \ne \mathbf{I}_p$. This is also the almost sure limit of $p^{-1}\left\|\mathbf{\Phi}_{\mathrm{Ideal}} - \mathbf{\Phi}\right\|_{\mathrm{F}}^2$, and hence $p^{-1/2}\left\|\widehat{\mathbf{\Phi}} - \mathbf{\Phi}\right\|_{\mathrm{F}}/(p^{-1/2}\left\|\mathbf{\Phi}_{\mathrm{Ideal}} - \mathbf{\Phi}\right\|_{\mathrm{F}}) \to 1$ almost surely if we can also show that $p^{-1}\sum_{i=1}^p(\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{p}_{1i} - \mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i})^2 \to 0$ almost surely. By Lemma 1, almost surely,

$$p^{-1}\sum_{i=1}^p(\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{p}_{1i} - \mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i})^2 \le \max_{i=1,\dots,p}\left|\frac{\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{p}_{1i} - \mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}}{\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}}\right| \max_{i=1,\dots,p}\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i} \to 0.$$

This completes the proof for the Frobenius loss.

For the inverse Stein loss, by Lemma 2, almost surely,

$$\begin{aligned} p^{-1}L(\mathbf{\Sigma}_p, \mathbf{\Sigma}_{\mathrm{Ideal}}) &= p^{-1}\sum_{i=1}^p \log(\mathbf{p}_i^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_i) - p^{-1}\sum_{i=1}^p \log(v_{n,i}) \\ &\to \int \log(\delta(\lambda))dF(\lambda) - \int \log(\tau)dH(\tau), \end{aligned}$$

where $v_{n,i}$ is the ith largest eigenvalue of $\mathbf{\Phi}$. Now consider the decomposition

$$\begin{aligned}
p^{-1}L(\mathbf{\Sigma}_p, \widehat{\mathbf{\Sigma}}_p) &= I_1 + I_2 + I_3 + I_4 + I_5,\\
I_1 &= \log(\widehat{\theta}/\theta),\\
I_2 &= \Big(\frac{\theta}{\widehat{\theta}} - 1\Big)p^{-1}\sum_{i=1}^{p}\left(\frac{\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}}{\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{p}_{1i}}\right),\\
I_3 &= p^{-1}\sum_{i=1}^{p}\left(\frac{\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}}{\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{p}_{1i}}\right) - 1,\\
I_4 &= p^{-1}\sum_{i=1}^{p}\log\left(\frac{\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}}{\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{p}_{1i}}\right),\\
I_5 &= p^{-1}\sum_{i=1}^{p}\log(\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}) - p^{-1}\sum_{i=1}^{p}\log(v_{n,i}).
\end{aligned}$$

We can prove that $I_1, I_2, I_3, I_4 \to 0$ almost surely by the proof of Theorem 1, and the result of Lemma 1. By Lemma 2, we can show that almost surely,

$$I_5 \to \int \log(\delta(\lambda))dF(\lambda) - \int \log(\tau)dH(\tau),$$

so that $p^{-1}L(\mathbf{\Sigma}_p, \mathbf{\Sigma}_{\mathrm{Ideal}})/\{p^{-1}L(\mathbf{\Sigma}_p, \widehat{\mathbf{\Sigma}}_p)\} \to 1$ almost surely, showing $\mathrm{EffLoss}(\mathbf{\Sigma}_p, \widehat{\mathbf{\Sigma}}_m) \to 0$ almost surely. This completes the proof of the theorem. $\square$

Proof of Theorem 4. For a matrix $A = (a_{ij})$, define $\|A\|_1 = \max_j \sum_{i\geq 1}|a_{ij}|$ and $\|A\|_\infty = \max_i \sum_{j\geq 1}|a_{ij}|$. Then almost surely,

$$\begin{aligned}
p^{1/2}\|\widehat{\mathbf{w}}_{\mathrm{opt}}\|_\infty &\leq \frac{p^{1/2}\|\widehat{\mathbf{\Sigma}}_p^{-1}\|_1}{\mathbf{1}_p^{\mathrm{T}}\widehat{\mathbf{\Sigma}}_p^{-1}\mathbf{1}_p} = \frac{p^{1/2}\|\widehat{\mathbf{\Phi}}^{-1}\|_\infty}{\mathbf{1}_p^{\mathrm{T}}\widehat{\mathbf{\Phi}}^{-1}\mathbf{1}_p} \leq \frac{p^{1/2}\{p^{1/2}\lambda_{\min}^{-1}(\widehat{\mathbf{\Phi}})\}}{p\lambda_{\min}(\widehat{\mathbf{\Phi}}^{-1})}\\
&= \frac{\lambda_{\max}(\widehat{\mathbf{\Phi}})}{\lambda_{\min}(\widehat{\mathbf{\Phi}})} \leq \frac{\max_{i=1,\ldots,p}\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{p}_{1i}}{\min_{i=1,\ldots,p}\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\mathbf{\Phi}}_2\mathbf{p}_{1i}} \to \frac{\max_{i=1,\ldots,p}\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}}{\min_{i=1,\ldots,p}\mathbf{p}_{1i}^{\mathrm{T}}\mathbf{\Phi}\mathbf{p}_{1i}}\\
&\leq \frac{\lambda_{\max}(\mathbf{\Phi})}{\lambda_{\min}(\mathbf{\Phi})} = \mathrm{Cond}(\mathbf{\Phi}),
\end{aligned}$$

where we used the results of Lemma 1 for the almost sure convergence. For the theoretical minimum variance portfolio $\mathbf{w}_{\text{opt}}$,

$$p^{1/2}\left\|\mathbf{w}_{\text{opt}}\right\|_\infty \le \frac{p^{1/2}\left\|\boldsymbol{\Sigma}_p^{-1}\right\|_1}{\mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Sigma}_p^{-1}\mathbf{1}_p} = \frac{p^{1/2}\left\|\boldsymbol{\Phi}^{-1}\right\|_\infty}{\mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Phi}^{-1}\mathbf{1}_p} \le \frac{p^{1/2}\{p^{1/2}\lambda_{\min}^{-1}(\boldsymbol{\Phi})\}}{p\lambda_{\min}(\boldsymbol{\Phi}^{-1})}$$
$$= \frac{\lambda_{\max}(\boldsymbol{\Phi})}{\lambda_{\min}(\boldsymbol{\Phi})} = \mathrm{Cond}(\boldsymbol{\Phi}).$$

For the actual risk bound, almost surely,

$$pR^2(\widehat{\mathbf{w}}_{\text{opt}}) = \frac{p\mathbf{1}_p^{\mathrm{T}}\widehat{\boldsymbol{\Sigma}}_p^{-1}\boldsymbol{\Sigma}_p\widehat{\boldsymbol{\Sigma}}_p^{-1}\mathbf{1}_p}{(\mathbf{1}_p^{\mathrm{T}}\widehat{\boldsymbol{\Sigma}}_p^{-1}\mathbf{1}_p)^2} \le \frac{p\{p\lambda_{\max}^2(\widehat{\boldsymbol{\Sigma}}_p^{-1})\lambda_{\max}(\boldsymbol{\Sigma}_p)\}}{p^2\lambda_{\min}^2(\widehat{\boldsymbol{\Sigma}}_p^{-1})} = \left\{\frac{\lambda_{\max}(\widehat{\boldsymbol{\Phi}})}{\lambda_{\min}(\widehat{\boldsymbol{\Phi}})}\right\}^2\lambda_{\max}(\boldsymbol{\Sigma}_p)$$
$$= \left(\frac{\max_{i=1,\ldots,p}\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\boldsymbol{\Phi}}_2\mathbf{p}_{1i}}{\min_{i=1,\ldots,p}\mathbf{p}_{1i}^{\mathrm{T}}\widetilde{\boldsymbol{\Phi}}_2\mathbf{p}_{1i}}\right)^2\lambda_{\max}(\boldsymbol{\Sigma}_p) \to \left(\frac{\max_{i=1,\ldots,p}\mathbf{p}_{1i}^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{p}_{1i}}{\min_{i=1,\ldots,p}\mathbf{p}_{1i}^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{p}_{1i}}\right)^2\lambda_{\max}(\boldsymbol{\Sigma}_p)$$
$$\le \left\{\frac{\lambda_{\max}(\boldsymbol{\Phi})}{\lambda_{\min}(\boldsymbol{\Phi})}\right\}^2\lambda_{\max}(\boldsymbol{\Sigma}_p) = \mathrm{Cond}^2(\boldsymbol{\Phi})\lambda_{\max}(\boldsymbol{\Sigma}_p).$$

Finally, for the theoretical minimum variance portfolio $\mathbf{w}_{\text{opt}}$,

$$pR^2(\mathbf{w}_{\text{opt}}) = \frac{p\mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Sigma}_p^{-1}\boldsymbol{\Sigma}_p\boldsymbol{\Sigma}_p^{-1}\mathbf{1}_p}{(\mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Sigma}_p^{-1}\mathbf{1}_p)^2} = \frac{p}{\mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Sigma}_p^{-1}\mathbf{1}_p} \le \lambda_{\max}(\boldsymbol{\Sigma}_p).$$

This completes the proof of the theorem. □

References

Bai, Z. & Silverstein, J. (2010). *Spectral Analysis of Large Dimensional Random Matrices.* New York: Springer, 2nd ed.

Bai, Z. D. & Silverstein, J. (1998). No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *The Annals of Probability* **26**, 316–345.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. & Shephard, N. (2011). Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics* **162**, 149 – 169.

Fan, J., Li, Y. & Yu, K. (2012). Vast volatility matrix estimation using high- frequency data for portfolio selection. *Journal of the American Statistical Association* **107**, 412–428.

Lam, C. (2016a). Nonparametric eigenvalue-regularized precision or covariance matrix estimator. *Ann. Statist.* **44**, 928–953.

Lam, C. (2016b). Supplement to "Nonparametric eigenvalue-regularized precision or covariance matrix estimator". *Ann. Statist.* .

Ledoit, O. & Péché, S. (2011). Eigenvectors of some large sample covariance matrix ensembles. *Probability Theory and Related Fields* **151**, 233–264.