Supplementary Materials for "On the estimation of high-dimensional integrated covariance matrix based on high-frequency data with multiple transactions"

Moming Wang[1], Jianhua Hu[1], Ningning Xia[1], Yong Zhou[2]

[1] *Shanghai University of Finance and Economics and* [2] *East China Normal University*

S1. Simulation results for PA-ATVA matrices

In this section, we demonstrate the finite sample performances of Theorem 1 and Theorem 2 by showing that the matrices $\mathcal{B}_M$ and $\mathcal{B}_M^*$ have almost the same empirical spectral distributions as their corresponding sample covariance matrices based on i.i.d. samples drawn from the ICV matrix.

We adopt scenarios for the diffusion process $(\mathbf{X}_t)$ from Xia and Zheng (2018) and Lam et al. (2017). We take the following U-shaped stochastic process (γ_t) as

$$d\gamma_t = -\rho(\gamma_t - \mu_t)dt + \sigma d\widetilde{W}_t, \quad \text{for} \quad t \in [0,1], \tag{S1.1}$$

where $\rho = 10, \sigma = 0.05, \mu_t = \sqrt{0.0009 + 0.0008\cos(2\pi t)}$, and the process

$$\widetilde{W}_t = \sum_{i=1}^{p} W_t^{(i)}/\sqrt{p}$$

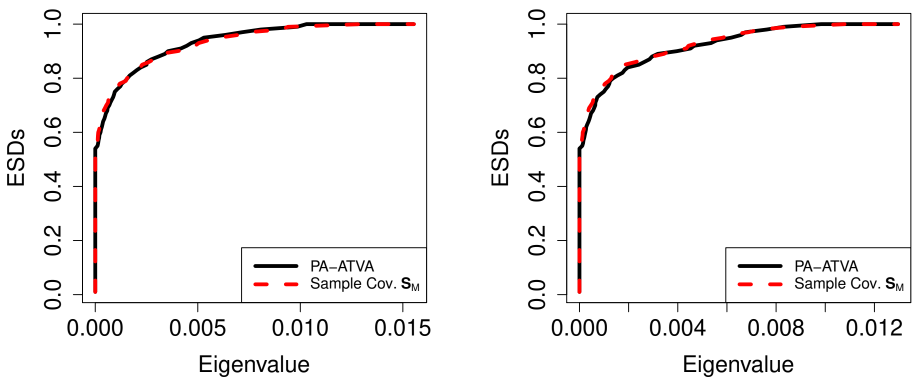
with $W_t^{(i)}$ being the ith component of the Brownian motion $(\mathbf{W}_t)$ that drives the price process. (A.iii) is violated since (γ_t) depends on all components of the Brownian motion. However, our estimate still works, as demonstrated by the simulation studies. We assume that $\mathbf{\Lambda} = (0.5^{|i-j|})_{i,j=1,\dots,p}$ and further rescale it to satisfy the condition $\mathrm{tr}(\mathbf{\Lambda}\mathbf{\Lambda}^{\mathrm{T}}) = p$ when spiked eigenvalues or factor models are considered. The latent log price process $(\mathbf{X}_t)$ follows

$$d\mathbf{X}_t = \gamma_t \mathbf{\Lambda} d\mathbf{W}_t. \tag{S1.2}$$

We investigate the finite sample performance of the PA-ATVA matrices $\mathcal{B}_M$ and $\mathcal{B}_M^*$ in the presence of microstructure noise. It is reasonable to conjecture that the ESDs of the PA-ATVA matrices $\mathcal{B}_M$ and $\mathcal{B}_M^*$ would have similar behavior to that of the sample covariance matrix $\mathbf{S}_M = 1/M \sum_{i=1}^{M} (\mathbf{\Sigma})^{1/2} \mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}} (\mathbf{\Sigma})^{1/2}$, where $\mathbf{Z}_i \overset{i.i.d.}{\sim} N(0, \mathbf{I}_p)$, as the LSDs of $\mathcal{B}_M$, $\mathcal{B}_M^*$ and $\mathbf{S}_M$ are related through the Marčenko-Pastur equation in Theorem 1, Theorem 2 and Theorem 1.1 of Silverstein (1995). Hence, the ESDs of the PA-ATVA matrices $\mathcal{B}_M$, $\mathcal{B}_M^*$ and sample covariance matrix $\mathbf{S}_M$ are compared here under various simulation designs. We set $p = 100$ and $n = 23400$, which represents the case where transactions are recorded per second within one trading day. We simulate the observations from the following additive model $\mathbf{Y}_{t_i} = \mathbf{X}_{t_i} + \boldsymbol{\varepsilon}_{t_i}$, in which the log price $(\mathbf{X}_t)$ follows from the continuous-time process as in (S1.2) and the noise values $(\boldsymbol{\varepsilon}_{t_i})$ are drawn independently from $N(0, 0.0002\mathbf{I}_p)$. The pre-averaging window length h is taken to be $\lfloor n^{0.55} \rfloor = 252$. We use $L_i^{(q)}$ to denote the number of transactions for stock q within time stamp $(t_{i-1}, t_i]$, for $q = 1, \ldots, p$ and $i = 1, \ldots, n$. We designed two transaction schemes as follows.

Design I: For simplicity, $L_i^{(q)} = L_i$ for each stock q within time interval $(t_{i-1}, t_i]$, where the t_i values are arranged as an equally spaced grid in $[0, 1]$. The L_i's are generated independently from a Poisson distribution with parameter 20 for the first and the last hours within 6.5 hours of a trading day, and from a Poisson distribution with parameter 5 for the remaining trading hours. According to the simulation results, shown in the left subfigure of Figure 1, the two ESDs of matrices $\mathcal{B}_M$ and $\mathbf{S}_M$ were very closely matched.

Design II: In order to generate high-frequency data such as that commonly used in practice, we further simulated observations in a highly asynchronous setting. Based on Design I, we allowed variation of $L_i^{(q)}$, which is generated independently from a discrete uniform distribution within the interval $[1, L_i]$, for each $q = 1, \ldots, p$ and $i = 1, \ldots, n$. Right subfigure of Figure 1 displays the ESDs of matrices $\mathcal{B}_M^*$ and $\mathbf{S}_M$ under Design II.



(a) (b)

Figure 1: ESDs of PA-ATVA matrix and sample covariance matrix $\mathbf{S}_M$ for dimension $p = 100$ and observation frequency $n = 23400$. The pre-averaging window length h was taken to be $\lfloor n^{0.55} \rfloor = 252$, with an effective sample size $M = \lfloor n/(2h) \rfloor = 46$. In the left panel, for each stock $q = 1, \ldots, p$, $L_i^{(q)} = L_i$ were generated independently from a Poisson distribution with parameter 20 for the first and the last hours within 6.5 hours of a trading day, and with parameter 5 for the rest of trading hours. In the rightpanel, $L_i^{(q)}$ values were generated independently from discrete uniform distribution $U[1, L_i]$ where L_is are generated with the same method in the left panel.

S2. Figure: asynchronous trading

Figure 2 shows a simplified version of true transactions and recording mechanism under asynchronous trading.

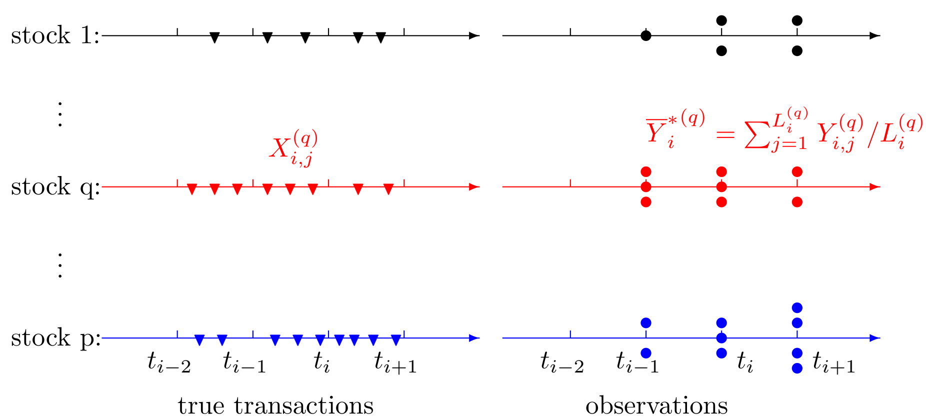


Figure 2: The true transactions vs. observations. Theoretically, the transactions occur consecutively for each stock during each time interval as shown in the left panel. However, in practice, the order of arrival is missing and the number of transaction varies according to the stock and recording interval as shown in the right panel.

S3. Proof of Theorem 1

Theorem 1 is a direct consequence of the following two propositions.

Proposition 1. *Under the assumptions of Theorem 1, the ESD of* $\widetilde{\boldsymbol{\Xi}}$ *converges almost surely, and the limit* $F^{\widetilde{\boldsymbol{\Xi}}}$ *is determined by* $\breve{H}$ *in that its Stieltjes transform* $m_{\widetilde{\boldsymbol{\Xi}}}(z)$ *satisfies the following equation:*

$$m_{\widetilde{\boldsymbol{\Xi}}}(z) = \int_{\tau \in \mathbb{R}} \frac{1}{\tau(1 - c(1 + z m_{\widetilde{\boldsymbol{\Xi}}}(z))) - z} d\breve{H}(\tau), \quad \textit{for } z \in \mathbb{C}^+. \qquad \text{(S3.3)}$$

Proposition 2. *Under the assumptions of Theorem 1, the ESD of the ICV matrix converges almost surely in distribution to a probability distribution* H *as* $p \to \infty$ *defined by*

$$H(x) = \breve{H}(x/\theta), \qquad \text{(S3.4)}$$

where $\theta = \int_0^1 (\gamma_t^*)^2 dt$. *Moreover,*

$$\lim_{p \to \infty} 3 \frac{\sum_{i=1}^M |\Delta \widetilde{\overline{\mathbf{Y}}}_{2i}|^2}{p} = \theta, \qquad \textit{almost surely.} \qquad \text{(S3.5)}$$

Proof of Proposition 1. To prove the convergence of $F^{\widetilde{\boldsymbol{\Xi}}}$, we fistly show that

$$\widetilde{\boldsymbol{\Xi}} = \frac{p}{M} \sum_{i=1}^M \frac{\Delta \widetilde{\overline{\mathbf{Y}}}_{2i} (\Delta \widetilde{\overline{\mathbf{Y}}}_{2i})^{\mathrm{T}}}{|\Delta \widetilde{\overline{\mathbf{Y}}}_{2i}|^2} \quad \text{and} \quad \widetilde{\widetilde{\boldsymbol{\Xi}}} := \frac{p}{M} \sum_{i=1}^M \frac{\Delta \widetilde{\overline{\mathbf{X}}}_{2i} \Delta \widetilde{\overline{\mathbf{X}}}_{2i}^{\mathrm{T}}}{|\Delta \widetilde{\overline{\mathbf{X}}}_{2i}|^2}$$

have the same LSD. Following the same arguments as in the proof of Proposition C.1 of Xia and Zheng (2018), it suffices to show that

$$\max_{1 \le i \le M, 1 \le j \le p} \frac{\sqrt{p} |\Delta \widetilde{\overline{\boldsymbol{\varepsilon}}}_{2i}^{(j)}|}{|\Delta \widetilde{\overline{\mathbf{X}}}_{2i}|} \to 0, \qquad \text{almost surely.} \qquad \text{(S3.6)}$$

We start by showing that there exists a constant $\widetilde{C} > 0$ for large M, such that

$$\min_{1\le i\le M} |\Delta\widetilde{\overline{\mathbf{X}}}_{2i}|^2 \ge \widetilde{C}. \tag{S3.7}$$

Recall $\overline{\mathbf{V}}_i = (1/L_i)\sum_{j=1}^{L_i} \mathbf{V}_{T_{i-1}+j}$, the average of the multiple observations at recording time t_i given in (2.6). We decompose its increments as follows:

$$\begin{aligned}
\Delta\overline{\mathbf{V}}_i :&= \overline{\mathbf{V}}_i - \overline{\mathbf{V}}_{i-1} \\
&= \frac{1}{L_i}\sum_{j=1}^{L_i} \mathbf{V}_{T_{i-1}+j} - \frac{1}{L_{i-1}}\sum_{j=1}^{L_{i-1}} \mathbf{V}_{T_{i-2}+j} \\
&= \frac{1}{L_i}\sum_{j=1}^{L_i} (\mathbf{V}_{T_{i-1}+j} - \mathbf{V}_{T_{i-2}}) - \frac{1}{L_{i-1}}\sum_{j=1}^{L_{i-1}} (\mathbf{V}_{T_{i-2}+j} - \mathbf{V}_{T_{i-2}}) \\
&= \sum_{j=1}^{L_i} a_{i,j}\Delta_{i,j}\mathbf{V} + \sum_{j=1}^{L_{i-1}} b_{i-1,j}\Delta_{i-1,j}\mathbf{V},
\end{aligned} \tag{S3.8}$$

where $a_{i,j} = 1 - \frac{j-1}{L_i}$, $b_{i,j} = \frac{j-1}{L_i}$, and

$$\Delta_{i,j}\mathbf{V} := \mathbf{V}_{T_{i-1}+j} - \mathbf{V}_{T_{i-1}+j-1}, \text{ for } \ j = 1, \dots, L_i, \ i = 2, \dots, n,$$

which is an asymmetric triangular form of $\Delta_{i,j}\mathbf{V}$. Following (S3.8), the log return based on the averaged log prices becomes

$$\Delta\overline{\mathbf{X}}_{2i} = \sum_{j=1}^{L_{2i}} a_{2i,j}\Delta_{2i,j}\mathbf{X} + \sum_{j=1}^{L_{2i-1}} b_{2i-1,j}\Delta_{2i-1,j}\mathbf{X},$$

where $\Delta_{i,j}\mathbf{X} = \mathbf{X}_{T_{i-1}+j} - \mathbf{X}_{T_{i-1}+j-1}$. Thus, the latent pre-averaged return

---

$\Delta\widetilde{\widetilde{\overline{\mathbf{X}}}}_{2i}$ is written as

$$\begin{aligned}
\Delta\widetilde{\widetilde{\overline{\mathbf{X}}}}_{2i} =& \frac{1}{h}\sum_{j=1}^{h}\left(\overline{\mathbf{X}}_{(2i-1)h+j} - \overline{\mathbf{X}}_{(2i-2)h}\right) - \frac{1}{h}\sum_{j=1}^{h}\left(\overline{\mathbf{X}}_{(2i-2)h+j} - \overline{\mathbf{X}}_{(2i-2)h}\right) \\
=& \sum_{l=1}^{2h}\left(1 - \frac{|h-l+1|}{h}\right)\Delta\overline{\mathbf{X}}_{(2i-2)h+l} \\
=& \sum_{l=1}^{2h}\left(1 - \frac{|h-l+1|}{h}\right)\sum_{j=1}^{L_{(2i-2)h+l}} a_{(2i-2)h+l,j}\Delta_{(2i-2)h+l,j}\mathbf{X} \\
&+ \sum_{l=1}^{2h-1}\left(1 - \frac{|h-l|}{h}\right)\sum_{j=1}^{L_{(2i-2)h+l}} b_{(2i-2)h+l,j}\Delta_{(2i-2)h+l,j}\mathbf{X}.
\end{aligned}$$

From the fact that $a_{i,j} + b_{i,j} = 1$, we can further write $\Delta\widetilde{\widetilde{\overline{\mathbf{X}}}}_{2i}$ as

$$\begin{aligned}
\Delta\widetilde{\widetilde{\overline{\mathbf{X}}}}_{2i} =& \frac{1}{h}\sum_{j=1}^{L_{2ih}} a_{2ih,j}\Delta_{2ih,j}\mathbf{X} \\
&+ \sum_{l=1}^{2h-1}\sum_{j=1}^{L_{(2i-2)h+l}}\left(1 - \frac{|h-l+1|}{h}\right)\Delta_{(2i-2)h+l,j}\mathbf{X} \\
&+ \sum_{l=1}^{2h-1}\sum_{j=1}^{L_{(2i-2)h+l}}\left[\frac{|h-l+1| - |h-l|}{h} b_{(2i-2)h+l,j}\right]\Delta_{(2i-2)h+l,j}\mathbf{X} \\
:=& \widetilde{\mathbf{V}}_i + \sqrt{\psi_i}\boldsymbol{\Lambda}\widetilde{\mathbf{Z}}_i, \qquad \text{(S3.9)}
\end{aligned}$$

where

$$\widetilde{\mathbf{V}}_i = \frac{1}{h}\sum_{j=1}^{L_{2ih}} a_{2ih,j} \int_{s_{T_{2ih-1}+j-1}}^{s_{T_{2ih-1}+j}} \boldsymbol{\mu}_t dt$$
$$+ \sum_{l=1}^{2h-1} \sum_{j=1}^{L_{(2i-2)h+l}} \left[\left(1 - \frac{|h-l+1|}{h}\right) + \frac{|h-l+1|-|h-l|}{h} b_{(2i-2)h+l,j}\right]$$
$$\cdot \int_{s_{T_{(2i-2)h+l-1}+j-1}}^{s_{T_{(2i-2)h+l-1}+j}} \boldsymbol{\mu}_t dt,$$

$$\psi_i = \frac{1}{h^2}\sum_{j=1}^{L_{2ih}} a_{2ih,j}^2 \int_{s_{T_{2ih-1}+j-1}}^{s_{T_{2ih-1}+j}} \gamma_t^2 dt$$
$$+ \sum_{l=1}^{2h-1} \sum_{j=1}^{L_{(2i-2)h+l}} \left[\left(1 - \frac{|h-l+1|}{h}\right) + \frac{|h-l+1|-|h-l|}{h} b_{(2i-2)h+l,j}\right]^2$$
$$\cdot \int_{s_{T_{(2i-2)h+l-1}+j-1}}^{s_{T_{(2i-2)h+l-1}+j}} \gamma_t^2 dt,$$

and

$$\widetilde{\mathbf{Z}}_i := \frac{1}{\sqrt{\psi_i}} \frac{1}{h}\sum_{j=1}^{L_{2ih}} a_{2ih,j} \int_{s_{T_{2ih-1}+j-1}}^{s_{T_{2ih-1}+j}} \gamma_t d\mathbf{W}_t$$
$$+ \sum_{l=1}^{2h-1} \sum_{j=1}^{L_{(2i-2)h+l}} \left[\left(1 - \frac{|h-l+1|}{h}\right) + \frac{|h-l+1|-|h-l|}{h} b_{(2i-2)h+l,j}\right]$$
$$\cdot \int_{s_{T_{(2i-2)h+l-1}+j-1}}^{s_{T_{(2i-2)h+l-1}+j}} \gamma_t d\mathbf{W}_t.$$

Without loss of generality, we may assume that γ_t and $\mathbf{W}_t$ are independent, which leads to the fact that each entry of $\widetilde{\mathbf{Z}}_i$ is i.i.d. standard normal. Otherwise, by using a similar trick as in the proof of (3.34) of Zheng and

Li (2011), we have

$$\max_{1\le i\le M} |\frac{1}{p}|\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i|^2 - 1| \to 0, \quad \text{almost surely.} \tag{S3.10}$$

Combining this with the fact that all the entries of $\widetilde{\mathbf{V}}_i$ are of order $O(h/n) = o(1/\sqrt{p})$, we have

$$\frac{\sum_{i=1}^M |\Delta\widetilde{\widetilde{\mathbf{X}}}_{2i}|^2}{p} = \frac{\sum_{i=1}^M |\widetilde{\mathbf{V}}_i + \sqrt{\psi_i}\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i|^2}{p} = \sum_{i=1}^M \psi_i + o_{a.s.}(1). \tag{S3.11}$$

On the other hand, from equation (S3.9), we get

$$|\Delta\widetilde{\widetilde{\mathbf{X}}}_{2i}|^2 = |\widetilde{\mathbf{V}}_i + \sqrt{\psi_i}\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i|^2 \ge |\widetilde{\mathbf{V}}_i|^2 + |\psi_i|\cdot|\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i|^2 - 2|\widetilde{\mathbf{V}}_i||\sqrt{\psi_i}\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i|.$$

Assumption (A.iii) implies that for all i, there exists $C' > 0$ such that $|\psi_i| \ge C'h/n$. Taking this together with Assumption (A.ix) and Equation (S3.10), there exists $C^* > 0$ such that for all n large enough,

$$\min_{1\le i\le M} |\psi_i|\cdot|\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i|^2 \ge C^*.$$

Moreover, $\max_i |\widetilde{\mathbf{V}}_i| = O(\sqrt{p}\times h/n) = o(1)$ follows from Assumption (A.ii). Therefore (S3.7) follows.

Next, we will show that

$$\max_{1\le i\le M, 1\le q\le p} \sqrt{p}|\Delta\widetilde{\widetilde{\boldsymbol{\varepsilon}}}_{2i}^{(q)}| \to 0, \quad \text{almost surely.} \tag{S3.12}$$

By the boundedness of L_i from Assumption (A.xi), $(\bar{\varepsilon}_i^{(q)}) = (1/L_i \sum_{k=1}^{L_i} \varepsilon_{T_{i-1}+k}^{(q)})$ is also a ρ-mixing sequence, and the ρ-mixing coefficients based on $(\bar{\varepsilon}_i^{(q)})$

have the same order as $\rho^q(r)$. Thus, (S3.12) follows by the same proof process used in (C.7) of Xia and Zheng (2018). Togethering with (S3.7), (S3.6) holds.

Finally, following a similar argument to the last part of the proof of Proposition 8 in Zheng and Li (2011), we have that the LSD of $\widetilde{\widetilde{\boldsymbol{\Xi}}}$ is determined by (S3.3). ■

Proof of Proposition 2. The convergence of $F^{\boldsymbol{\Sigma}}$ follows from Assumption (A.vii) and the fact that

$$F^{\boldsymbol{\Sigma}}(x) = F^{\breve{\boldsymbol{\Sigma}}}\left(\frac{x}{\int_0^1 \gamma_t^2 dt}\right) \qquad \text{for all } x \geq 0.$$

Note that

$$\sum_{i=1}^{M} |\Delta\widetilde{\overline{\mathbf{Y}}}_{2i}|^2 = \sum_{i=1}^{M} |\Delta\widetilde{\overline{\mathbf{X}}}_{2i}|^2 + 2\sum_{i=1}^{M} \Delta\widetilde{\overline{\mathbf{X}}}_{2i}^{\mathrm{T}} \Delta\widetilde{\overline{\boldsymbol{\varepsilon}}}_{2i} + \sum_{i=1}^{M} |\Delta\widetilde{\overline{\boldsymbol{\varepsilon}}}_{2i}|^2.$$

The convergence of (S3.6) and inequlity (S3.7) imply that $\sum_{i=1}^{M} |\Delta\widetilde{\overline{\boldsymbol{\varepsilon}}}_{2i}|^2/p \to 0$ almost surely. It remains to prove that

$$\lim_{p\to\infty} 3\frac{\sum_{i=1}^{M} |\Delta\widetilde{\overline{\mathbf{X}}}_{2i}|^2}{p} = \theta, \qquad \text{almost surely,} \tag{S3.13}$$

where $\theta = \int_0^1 (\gamma_t^*)^2 dt$, and

$$\frac{\sum_{i=1}^{M} \Delta\widetilde{\overline{\mathbf{X}}}_{2i}^{\mathrm{T}} \Delta\widetilde{\overline{\boldsymbol{\varepsilon}}}_{2i}}{p} \to 0, \qquad \text{almost surely.} \tag{S3.14}$$

To show (S3.13), by (S3.11), it suffices to show that

$$\lim_{n\to\infty}\sum_{i=1}^{M}\int_{s_{T_{(2i-2)h}}}^{s_{T_{2ih}}}|M\psi_i-\frac{1}{3}(\gamma_s^*)^2|ds=0,$$

almost surely. Suppose that γ_t^* has J jumps at $\{\tau_1,\ldots,\tau_J\}$; then

$$\begin{aligned}
&\sum_{i=1}^{M}\int_{s_{T_{(2i-2)h}}}^{s_{T_{2ih}}}|M\psi_i-\frac{1}{3}(\gamma_s^*)^2|ds\\
=&\sum_{i\in\{\tau_1,\ldots,\tau_J\}}\int_{s_{T_{(2i-2)h}}}^{s_{T_{2ih}}}|M\psi_i-\frac{1}{3}(\gamma_s^*)^2|ds\\
&+\sum_{i\notin\{\tau_1,\ldots,\tau_J\}}\int_{s_{T_{(2i-2)h}}}^{s_{T_{2ih}}}|M\psi_i-\frac{1}{3}(\gamma_s^*)^2|ds\\
:=&\Delta_1+\Delta_2.
\end{aligned}$$

For any $\epsilon>0$ and for sufficiently large n, $|\Delta_1|\le\epsilon$ follows from the boundedness of $|M\psi_i|$ and γ_t^*. For the second term, Δ_2, by defining ψ_i^* by replacing γ_t with γ_t^* from the definition of ψ_i, we have

$$|\Delta_2|\le\Delta_{21}+\Delta_{22}+\Delta_{23}+\Delta_{24},$$

where

$$\Delta_{21} := \sum_{i \notin \{\tau_1, \dots, \tau_J\}} \int_{s_{T_{(2i-2)h}}}^{s_{T_{2ih}}} |M\psi_i - M\psi_i^*| ds,$$

$$\Delta_{22} := \sum_{i \notin \{\tau_1, \dots, \tau_J\}} \int_{s_{T_{(2i-2)h}}}^{s_{T_{2ih}}} |M\psi_i^* - M(\gamma^*_{(2i-2)/h})^2 A_i| ds,$$

$$\Delta_{23} := \sum_{i \notin \{\tau_1, \dots, \tau_J\}} \int_{s_{T_{(2i-2)h}}}^{s_{T_{2ih}}} |M(\gamma^*_{(2i-2)/h})^2 A_i - M(\gamma^*_s)^2 A_i| ds,$$

$$\Delta_{24} := \sum_{i \notin \{\tau_1, \dots, \tau_J\}} \int_{s_{T_{(2i-2)h}}}^{s_{T_{2ih}}} |M(\gamma^*_s)^2 A_i - \frac{1}{3}(\gamma^*_s)^2| ds,$$

$$A_i := \frac{1}{h^2} \sum_{j=1}^{L_{2ih}} a^2_{2ih,j} \Delta s_{2ih,j} + \sum_{l=1}^{2h-1} \sum_{j=1}^{L_{(2i-2)h+l}} \Big[1 - \frac{|h-l+1|}{h} + \\ + \frac{|h-l+1| - |h-l|}{h} b_{(2i-2)h+l,j}\Big]^2 \Delta s_{(2i-2)h+l,j},$$

and $\Delta s_{i,j} := s_{T_{i-1}+j} - s_{T_{i-1}+j-1}$.

We further decompose A_i as $A_i = A_{i1} + A_{i2} + A_{i3} + A_{i4}$, where

$$A_{i1} := \frac{1}{h^2} \sum_{j=1}^{L_{2ih}} a^2_{2ih,j} \Delta s_{2ih,j} = O(\frac{1}{nh^2}),$$

$$A_{i2} := \sum_{l=1}^{2h-1} \left(1 - \frac{|h-l+1|}{h}\right)^2 \sum_{j=1}^{L_{(2i-2)h+l}} \Delta s_{(2i-2)h+l,j} = \frac{2h}{3n} + o(\frac{h}{n}),$$

$$A_{i3} := \sum_{l=1}^{2h-1} \sum_{j=1}^{L_{(2i-2)h+l}} \frac{2}{h} \left(1 - \frac{|h-l+1|}{h}\right) (|h-l+1| - |h-l|) \\ \cdot b_{(2i-2)h+l,j} \Delta s_{(2i-2)h+l,j} = O(\frac{1}{n}),$$

$$A_{i4} := \sum_{l=1}^{2h-1} \sum_{j=1}^{L_{(2i-2)h+l}} \frac{1}{h^2} (|h-l+1| - |h-l|)^2 b^2_{(2i-2)h+l,j} \Delta s_{(2i-2)h+l,j} \\ = O(\frac{1}{nh}),$$

follow from the boundedness of L_i both below and above, the fact that $\Delta s_{i,j} = O(1/n)$ and Assumption (A.x). As (γ_t^*) is continuous in $[s_{T_{(2i-2)h}}, s_{T_{2ih}}]$ when $i \notin \{\tau_1, \ldots, \tau_J\}$, (γ_t) uniformly converges to (γ_t^*) by Assumption (A.iv), and $\psi_i^* = O(h/n)$ by Assumption (A.iii) and (A.xi), for any $\epsilon > 0$ and sufficiently large n, p, it is easy to show that

$$|MA_i - 1/3| \leq \epsilon \quad \text{and} \quad \max\{\Delta_{21}, \Delta_{22}, \Delta_{23}, \Delta_{24}\} < C\epsilon.$$

This completes the proof of (S3.13). Finally, (S3.14) follows from (S3.13) and (S3.6) and (S3.7). ■

S4. Proof of Theorem 2

Suppose that we have $L_i^{(q)} (\geq 1)$ observations for each stock q at recording time $t_i = i/n$, for $q = 1, 2, \ldots, p$ and $i = 1, 2, \ldots, n$. Take $T_i^{(q)} = \sum_{k=1}^{i} L_k^{(q)}$ for $q = 1, \ldots, p$ and $i = 1, \ldots, n$. Recall that for any process $(\mathbf{V}_t)$, $V_{i,j}^{(q)}$ denote the observation of the jth transaction for stock q during time interval $(t_{i-1}, t_i]$, and the true transaction time of $V_{i,j}^{(q)}$ is denoted as $s^{(q)}_{T^{(q)}_{i-1}+j}$, for $j = 1, \ldots, L_i^{(q)}$ satisfying $t_{i-1} \leq s^{(q)}_{T^{(q)}_{i-1}+1} < \cdots < s^{(q)}_{T^{(q)}_{i-1}+L_i^{(q)}} = s^{(q)}_{T_i^{(q)}} \leq t_i$. Thus, $V_{i,j}^{(q)} = V^{(q)}_{s^{(q)}_{T^{(q)}_{i-1}+j}}$. Under asynchronous trading conditions, the average of multiple observations at each recording time t_i is given by

$$\overline{\mathbf{V}}_i^* := \left(\sum_{j=1}^{L_i^{(1)}} \frac{1}{L_i^{(1)}} V_{i,j}^{(1)}, \ldots, \sum_{j=1}^{L_i^{(p)}} \frac{1}{L_i^{(p)}} V_{i,j}^{(p)} \right)^{\mathrm{T}};$$

thus, the increment at trading time t_i becomes

$$\Delta \overline{\mathbf{V}}_i^* := \begin{pmatrix} \sum_{j=1}^{L_i^{(1)}} a_{i,j}^{(1)} \Delta_{i,j}^{(1)} V + \sum_{j=1}^{L_{i-1}^{(1)}} b_{i-1,j}^{(1)} \Delta_{i-1,j}^{(1)} V \\ \vdots \\ \sum_{j=1}^{L_i^{(p)}} a_{i,j}^{(p)} \Delta_{i,j}^{(p)} V + \sum_{j=1}^{L_{i-1}^{(p)}} b_{i-1,j}^{(p)} \Delta_{i-1,j}^{(p)} V \end{pmatrix},$$

where $a_{i,j}^{(q)} = 1 - \frac{j-1}{L_i^{(q)}}$, $b_{i,j}^{(q)} = \frac{j-1}{L_i^{(q)}}$ and $\Delta_{i,j}^{(q)} X = X_{i,j}^{(q)} - X_{i,j-1}^{(q)}$, for $q = 1, 2, \ldots, p$. Similar to the proof of Lemma 1, we first show that there exists a constant $C > 0$ such that for large M

$$\min_{1 \le i \le M} |\Delta \widetilde{\overline{\mathbf{X}}}_{2i}^*|^2 \ge C, \tag{S4.15}$$

and

$$\max_{1 \le i \le M, 1 \le q \le p} \sqrt{p} |\Delta(\widetilde{\overline{\boldsymbol{\varepsilon}}}_{2i}^*)^{(q)}| \to 0, \quad \text{almost surely,} \tag{S4.16}$$

which leads to the result that matrices

$$\frac{p}{M} \sum_{i=1}^{M} \frac{\Delta \widetilde{\overline{\mathbf{Y}}}_{2i}^* (\Delta \widetilde{\overline{\mathbf{Y}}}_{2i}^*)^{\mathrm{T}}}{|\Delta \widetilde{\overline{\mathbf{Y}}}_{2i}^*|^2} \quad \text{and} \quad \frac{p}{M} \sum_{i=1}^{M} \frac{\Delta \widetilde{\overline{\mathbf{X}}}_{2i}^* (\Delta \widetilde{\overline{\mathbf{X}}}_{2i}^*)^{\mathrm{T}}}{|\Delta \widetilde{\overline{\mathbf{X}}}_{2i}^*|^2}$$

have the same LSD. We only need to prove (S4.15), as (S4.16) holds straightforwardly from the proof of (S3.12) given earlier. By a similar decomposition to that used in (S3.9), the return based on the pre-averaged (latent) price can be decomposed as

$$\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i} = \frac{1}{h}\sum_{j=1}^{h}\overline{\mathbf{X}}^*_{(2i-1)h+j} - \frac{1}{h}\sum_{j=1}^{h}\overline{\mathbf{X}}^*_{(2i-2)h+j} = \mathbf{R}_{i1} + \mathbf{M}_i + \mathbf{R}_{i2},$$

where $\mathbf{M}_i = (M_i^{(1)}, \ldots, M_i^{(q)})^{\mathrm{T}}$, $\mathbf{R}_{i\ell} = (R_{i\ell}^{(1)}, \ldots, R_{i\ell}^{(q)})^{\mathrm{T}}$ for $\ell = 1, 2$, and their qth components have the form

$$R_{i1}^{(q)} = \frac{1}{h}\sum_{j=1}^{L_{2ih}^{(q)}} a_{2ih,j}^{(q)}\Delta_{2ih,j}^{(q)}X,$$

$$M_i^{(q)} = \sum_{l=1}^{2h-1}\sum_{j=1}^{L_{(2i-2)h+l}^{(q)}} c_{l,h}\Delta_{(2i-2)h+l,j}^{(q)}X,$$

$$R_{i2}^{(q)} = \sum_{l=1}^{2h-1}\sum_{j=1}^{L_{(2i-2)h+l}^{(q)}} \left[\frac{|h-l+1|-|h-l|}{h} b_{(2i-2)h+l,j}^{(q)}\right]\Delta_{(2i-2)h+l,j}^{(q)}X,$$

where $c_{l,h} = 1 - \frac{|h-l+1|}{h}$. Note that $M_i^{(q)}$ can be reduced to

$$\sum_{l=1}^{2h-1} c_{l,h}\cdot(X^{(q)}_{s^{(q)}_{T_{(2i-2)h+l}}} - X^{(q)}_{s^{(q)}_{T_{(2i-2)h+l-1}}}).$$

We further decompose $M_i^{(q)}$ as $\Phi_i^{(q)} + R_{i3}^{(q)}$, where

$$\Phi_i^{(q)} = \sum_{l=1}^{2h-1} c_{l,h}(X^{(q)}_{t_{(2i-2)h+l}} - X^{(q)}_{t_{(2i-2)h+l-1}}),$$

$$R_{i3}^{(q)} = \sum_{l=1}^{2h-1} c_{l,h}\left(X^{(q)}_{s^{(q)}_{T_{(2i-2)h+l}}} - X^{(q)}_{t_{(2i-2)h+l}} + X^{(q)}_{t_{(2i-2)h+l-1}} - X^{(q)}_{s^{(q)}_{T_{(2i-2)h+l-1}}}\right),$$

and $X_{t_i}^{(q)}$ denotes the log price for stock q at recording time t_i. Let $\mathbf{\Phi}_i = (\Phi_i^{(1)}, \ldots, \Phi_i^{(q)})^{\mathrm{T}}$ and $\mathbf{R}_{i3} = (R_{i3}^{(1)}, \ldots, R_{i3}^{(q)})^{\mathrm{T}}$. Thus,

$$\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i} = \mathbf{\Phi}_i + \mathbf{R}_{i1} + \mathbf{R}_{i2} + \mathbf{R}_{i3}, \tag{S4.17}$$

where

$$\mathbf{\Phi}_i = \sum_{l=1}^{2h-1} \left(1 - \frac{|h-l+1|}{h}\right) \Delta \mathbf{X}_{(2i-2)h+l},$$

and

$$\Delta \mathbf{X}_{(2i-2)h+l} = \begin{pmatrix} X^{(1)}_{t_{(2i-2)h+l}} - X^{(1)}_{t_{(2i-2)h+l-1}} \\ \vdots \\ X^{(p)}_{t_{(2i-2)h+l}} - X^{(p)}_{t_{(2i-2)h+l-1}} \end{pmatrix} = \mathbf{X}_{t_{(2i-2)h+l}} - \mathbf{X}_{t_{(2i-2)h+l-1}},$$

which reduces to the synchronous setting. Using a similar argument to that in (S3.13), we have

$$\lim_{p\to\infty} \frac{3\sum_{i=1}^{M} |\mathbf{\Phi}_i|^2}{p} = \theta. \tag{S4.18}$$

Thus, (S4.15) follows if we can show that there exists a constant $C > 0$ such that for large M

$$\min_{1\le i\le M} |\mathbf{\Phi}_i|^2 \ge C, \tag{S4.19}$$

and

$$\max_{1\le i\le M, 1\le q\le p} \sqrt{p} |R^{(q)}_{i1} + R^{(q)}_{i2} + R^{(q)}_{i3}| \to 0, \quad \text{almost surely.} \tag{S4.20}$$

Notice that (S4.19) follows naturally from the proof of (S3.7). We only need to show the proof of (S4.20). To prove (S4.20), let $R^{(q)}_i = R^{(q)}_{i1} + R^{(q)}_{i2} + R^{(q)}_{i3}$. By C_p inequality and the Burkholder-Davis-Gundy inequality, we have for any $\kappa \ge 1$,

$$E(R^{(q)}_i) < Cn^{-1} \ \text{ and } \ E|R^{(q)}_i - E(R^{(q)}_i)|^{2\kappa} \le Cn^{-\kappa}, \tag{S4.21}$$

for all i, q. Hence, it follows that for any $\varepsilon > 0$ and $\kappa \geq 1$,

$$
\begin{aligned}
&P\left(\max_{i,q} \sqrt{p}|R_i^{(q)} - E(R_i^{(q)})| \geq \varepsilon\right) \\
&\leq \sum_{i,q} \frac{p^{\kappa} E|R_i^{(q)} - E(R_i^{(q)})|^{2\kappa}}{\varepsilon^{2\kappa}} \\
&\leq C \cdot \frac{Mp \cdot p^{\kappa} \cdot n^{-\kappa}}{\varepsilon^{2\kappa}} = O(p^{2-\kappa\beta/(1-\beta)}).
\end{aligned}
$$

We choose κ large enough such that $\kappa\beta/(1-\beta) - 2 > 1$. This proves (S4.20) from the Borel–Cantelli lemma. Thus, (S4.15) holds.

From (S4.15) and (S4.16), we know that

$$
\frac{p}{M} \sum_{i=1}^{M} \frac{\Delta\widetilde{\overline{\mathbf{Y}}}_{2i}^{*}(\Delta\widetilde{\overline{\mathbf{Y}}}_{2i}^{*})^{\mathrm{T}}}{|\Delta\widetilde{\overline{\mathbf{Y}}}_{2i}^{*}|^{2}} \quad \text{and} \quad \frac{p}{M} \sum_{i=1}^{M} \frac{\Delta\widetilde{\overline{\mathbf{X}}}_{2i}^{*}(\Delta\widetilde{\overline{\mathbf{X}}}_{2i}^{*})^{\mathrm{T}}}{|\Delta\widetilde{\overline{\mathbf{X}}}_{2i}^{*}|^{2}}
$$

have the same LSD. Further, (S4.17), (S4.19), (S4.20), and Lemma 2 imply that

$$
\frac{p}{M} \sum_{i=1}^{M} \frac{\Delta\widetilde{\overline{\mathbf{X}}}_{2i}^{*}(\Delta\widetilde{\overline{\mathbf{X}}}_{2i}^{*})^{\mathrm{T}}}{|\Delta\widetilde{\overline{\mathbf{X}}}_{2i}^{*}|^{2}} \quad \text{and} \quad \frac{p}{M} \sum_{i=1}^{M} \frac{\mathbf{\Phi}_i \mathbf{\Phi}_i^{\mathrm{T}}}{|\mathbf{\Phi}_i|^{2}}
$$

have the same LSD. Moreover, from the proof of Theorem 2.3 in Xia and Zheng (2018), we have already known that the LSD of $\frac{p}{M} \sum_{i=1}^{M} \frac{\mathbf{\Phi}_i \mathbf{\Phi}_i^{\mathrm{T}}}{|\mathbf{\Phi}_i|^{2}}$ relates to $\breve{H}$ through the Marčenko-Pastur equation. Therefore, at last, we only need to prove that

$$
\lim_{p\to\infty} \frac{3\sum_{i=1}^{M} |\Delta\widetilde{\overline{\mathbf{Y}}}_{2i}^{*}|^{2}}{p} = \theta, \qquad \text{almost surely.} \tag{S4.22}
$$

Followed by the proof of Proposition 2, (S4.22) holds if we can show that

$$
\lim_{p\to\infty} \frac{3\sum_{i=1}^{M} |\Delta\widetilde{\overline{\mathbf{X}}}_{2i}^{*}|^{2}}{p} = \theta, \tag{S4.23}
$$

and

$$\frac{\sum_{i=1}^{M}(\Delta\widetilde{\overline{\mathbf{X}}}_{2i}^{*})^{\mathrm{T}}\Delta\widetilde{\overline{\boldsymbol{\varepsilon}}}_{2i}^{*}}{p} \to 0, \tag{S4.24}$$

almost surely. Notice that (S4.23) follows from (S4.18) and (S4.20), and (S4.24) follows from (S4.23) and (S4.16). Therefore, the proof of Theorem 2 is complete. ■

S5. Proofs of Theorem 3, 4, 5.

S5.1 Proof of Theorem 3

Under the assumptions of Theorem 3, if we can show that

$$\max_{1\le\ell\le p}\left|\frac{\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_2^{*}\mathbf{u}_{1\ell}-\mathbf{u}_{1\ell}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1\ell}}{\mathbf{u}_{1\ell}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1\ell}}\right| \to 0, \quad \text{almost surely.} \tag{S5.25}$$

Then Theorem 3 is a direct consequence of (S5.25), (S4.22) and Assumptions (A.iv) and (C.i).

To prove (S5.25), we decompose (S5.25) into two parts

$$\frac{\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_2^{*}\mathbf{u}_{1\ell}-\mathbf{u}_{1\ell}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1\ell}}{\mathbf{u}_{1\ell}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1\ell}} = I_{\ell 1}+I_{\ell 2}, \tag{S5.26}$$

where

$$I_{\ell 1}=\frac{\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_2^{*}\mathbf{u}_{1\ell}-\frac{1}{M_2}\sum_{i\in J_2}(\mathbf{u}_{1\ell}^{\mathrm{T}}\boldsymbol{\Lambda}\widetilde{\mathbf{Z}}_i^{*})^2}{\mathbf{u}_{1\ell}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1\ell}}, I_{\ell 2}=\frac{\frac{1}{M_2}\sum_{i\in J_2}(\mathbf{u}_{1\ell}^{\mathrm{T}}\boldsymbol{\Lambda}\widetilde{\mathbf{Z}}_i^{*})^2-\mathbf{u}_{1\ell}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1\ell}}{\mathbf{u}_{1\ell}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1\ell}}.$$

By Assumption (C.iv) and Lemma 1 in Lam (2016), we have $\max_{1\le\ell\le p}|I_{\ell 2}|\to 0$ almost surely.

Now we consider the convergence of $I_{\ell 1}$. we further decompose $I_{\ell 1}$ as follows,

$$\begin{aligned}
&\max_{1\le \ell\le p}|I_{\ell 1}| \\
&\le \max_{1\le \ell\le p}\Bigg\{\left|\frac{\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_2^*\mathbf{u}_{1\ell}-\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_{yx,2}\mathbf{u}_{1\ell}}{\mathbf{u}_{1\ell}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1\ell}}\right|+\left|\frac{\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_{yx,2}\mathbf{u}_{1\ell}-\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_{xx,2}\mathbf{u}_{1\ell}}{\mathbf{u}_{1\ell}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1\ell}}\right| \\
&\quad+\left|\frac{\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_{xx,2}\mathbf{u}_{1\ell}-\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_{x\phi,2}\mathbf{u}_{1\ell}}{\mathbf{u}_{1\ell}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1\ell}}\right|+\left|\frac{\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_{x\phi,2}\mathbf{u}_{1\ell}-\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_{\phi\phi,2}\mathbf{u}_{1\ell}}{\mathbf{u}_{1\ell}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1\ell}}\right| \\
&\quad+\left|\frac{\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_{\phi\phi,2}\mathbf{u}_{1\ell}-M_2^{-1}\sum_{i\in J_2}(\mathbf{u}_{1\ell}^{\mathrm{T}}\boldsymbol{\Lambda}\widetilde{\mathbf{Z}}_i^*)^2}{\mathbf{u}_{1\ell}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1\ell}}\right|\Bigg\},
\end{aligned}$$

where

$$\widetilde{\boldsymbol{\Xi}}_{yx,j}=\frac{p}{M_j}\sum_{i\in J_j}\frac{\Delta\widetilde{\overline{\mathbf{Y}}}^*_{2i}(\Delta\widetilde{\overline{\mathbf{Y}}}^*_{2i})^{\mathrm{T}}}{|\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}|^2},\quad \widetilde{\boldsymbol{\Xi}}_{xx,j}=\frac{p}{M_j}\sum_{i\in J_j}\frac{\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}(\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i})^{\mathrm{T}}}{|\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}|^2},$$

$$\widetilde{\boldsymbol{\Xi}}_{x\phi,j}=\frac{p}{M_j}\sum_{i\in J_j}\frac{\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}(\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i})^{\mathrm{T}}}{|\boldsymbol{\Phi}_i|^2}\ \text{ and }\ \widetilde{\boldsymbol{\Xi}}_{\phi\phi,j}=\frac{p}{M_j}\sum_{i\in J_j}\frac{\boldsymbol{\Phi}_i\boldsymbol{\Phi}_i^{\mathrm{T}}}{|\boldsymbol{\Phi}_i|^2}.$$

The convergence of $|I_{\ell 1}|$ follows directly if we can show the following results,

$$\max_{1\le\ell\le p}\mathbf{u}_{1\ell}^{\mathrm{T}}\left(\frac{p}{M_2}\sum_{i\in J_2}\frac{\Delta\widetilde{\overline{\mathbf{Y}}}^*_{2i}(\Delta\widetilde{\overline{\mathbf{Y}}}^*_{2i})^{\mathrm{T}}}{|\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}|^2}-\frac{p}{M_2}\sum_{i\in J_2}\frac{\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}(\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i})^{\mathrm{T}}}{|\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}|^2}\right)\mathbf{u}_{1\ell}\to 0, \tag{S5.27}$$

$$\max_{1\le\ell\le p}\mathbf{u}_{1\ell}^{\mathrm{T}}\left(\frac{p}{M_2}\sum_{i\in J_2}\frac{\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}(\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i})^{\mathrm{T}}}{|\mathbf{\Phi}_i|^2}-\frac{p}{M_2}\sum_{i\in J_2}\frac{\mathbf{\Phi}_i\mathbf{\Phi}_i^{\mathrm{T}}}{|\mathbf{\Phi}_i|^2}\right)\mathbf{u}_{1\ell}\to 0, \tag{S5.28}$$

$$\max_{1\le i\le M}|\Delta\widetilde{\overline{\mathbf{Y}}}^*_{2i}|\cdot|\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}|^{-1}\to 1,\qquad \max_{1\le i\le M}|\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}|\cdot|\mathbf{\Phi}_i|^{-1}\to 1 \tag{S5.29}$$

$$\max_{1\le\ell\le p}[\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\mathbf{\Xi}}_{\phi\phi,2}\mathbf{u}_{1\ell}-M_2^{-1}\sum_{i\in J_2}(\mathbf{u}_{1\ell}^{\mathrm{T}}\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*)^2]\to 0, \tag{S5.30}$$

$$\limsup_{p\to\infty}\mathbf{u}_{1\ell}^{\mathrm{T}}\frac{p}{M_2}\sum_{i\in J_2}\frac{\Delta\widetilde{\overline{\mathbf{Y}}}^*_{2i}(\Delta\widetilde{\overline{\mathbf{Y}}}^*_{2i})^{\mathrm{T}}}{|\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}|^2}\mathbf{u}_{1\ell}\le C, \tag{S5.31}$$

$$\limsup_{p\to\infty}\mathbf{u}_{1\ell}^{\mathrm{T}}\frac{p}{M_2}\sum_{i\in J_2}\frac{\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}(\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i})^{\mathrm{T}}}{|\mathbf{\Phi}_i|^2}\mathbf{u}_{1\ell}\le C, \tag{S5.32}$$

almost surely, where C is a finite constant. (S5.32) follows from the boundedness of $\|\breve{\mathbf{\Sigma}}\|$, the convergence of $|I_{\ell 2}|$, (S5.30), (S5.29) and (S5.28). (S5.31) follows from (S5.32), (S5.29) and (S5.27). Moreover, (S5.29) has already been shown in the proof of Theorem 2. Thus it only suffices to prove (S5.27), (S5.28) and (S5.30) to finish the proof of Theorem 3. We begin with (S5.30) and then show (S5.28), (S5.27).

We first consider (S5.30). By the definition of $\mathbf{\Phi}_i$, $\mathbf{\Phi}_i$ can be further decomposed into two parts as $\mathbf{\Phi}_i = \phi_i^{1/2}(\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^* + \widetilde{\mathbf{V}}_i^*)$, where

$$\phi_i = \sum_{l=1}^{2h-1}\int_{t_{(2i-2)h+l-1}}^{t_{(2i-2)h+l}} c_{l,h}^2\gamma_t^2 dt, \quad \widetilde{\mathbf{V}}_i^* = \phi_i^{-1/2}\sum_{l=1}^{2h-1}\int_{t_{(2i-2)h+l-1}}^{t_{(2i-2)h+l}} c_{l,h}\boldsymbol{\mu}_t dt, \tag{S5.33}$$

and $\widetilde{\mathbf{Z}}_i^*$s are i.i.d standard normal. Note that

$$\begin{aligned}
&\max_{1\le\ell\le p}|\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\mathbf{\Xi}}_{\phi\phi,2}\mathbf{u}_{1\ell} - M_2^{-1}\sum_{i\in J_2}(\mathbf{u}_{1\ell}^{\mathrm{T}}\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*)^2| \\
\le &\max_{i\in J_2}\left(\frac{1}{|\widetilde{\mathbf{V}}_i^* + \mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*|^2/p} - 1\right)\cdot\max_{1\le\ell\le p}\frac{1}{M_2}\sum_{i\in J_2}\mathbf{u}_{1\ell}^{\mathrm{T}}(\widetilde{\mathbf{V}}_i^* + \mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*)(\widetilde{\mathbf{V}}_i^* + \mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*)^{\mathrm{T}}\mathbf{u}_{1\ell} \\
&+\max_{1\le\ell\le p}\mathbf{u}_{1\ell}^{\mathrm{T}}\big[\frac{1}{M_2}\sum_{i\in J_2}(\widetilde{\mathbf{V}}_i^* + \mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*)(\widetilde{\mathbf{V}}_i^* + \mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*)^{\mathrm{T}} - \frac{1}{M_2}\sum_{i\in J_2}\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*(\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*)^{\mathrm{T}}\big]\mathbf{u}_{1\ell}.
\end{aligned}$$

By Assumption (C.ii), we have $\max_{1\le i\le M}|\widetilde{\mathbf{V}}_i^*|^2 \to 0$. Thus

$$\max_{1\le i\le M}\left|\frac{|\mathbf{\Phi}_i|}{|\phi_i^{1/2}\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*|} - 1\right| \to 0, \quad \|\frac{1}{M_2}\sum_{i\in J_2}\widetilde{\mathbf{V}}_i^*(\widetilde{\mathbf{V}}_i^*)^{\mathrm{T}}\| \le \max_{i\in J_2}|\widetilde{\mathbf{V}}_i^*|^2 \to 0, \tag{S5.34}$$

almost surely. Further, by Cauchy-Schwartz inequality and the convergence of $\max\limits_{1\le\ell\le p}|I_{\ell 2}|$,

$$\begin{aligned}
&\max_{1\le\ell\le p}\mathbf{u}_{1\ell}^{\mathrm{T}}\frac{1}{M_2}\sum_{i\in J_2}\widetilde{\mathbf{V}}_i^*(\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*)^{\mathrm{T}}\mathbf{u}_{1\ell} \\
\le &\max_{1\le\ell\le p}\big(\frac{1}{M_2}\sum_{i\in J_2}\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\mathbf{V}}_i^*(\widetilde{\mathbf{V}}_i^*)^{\mathrm{T}}\mathbf{u}_{1\ell}\big)^{1/2}\cdot\max_{1\le\ell\le p}\big(\frac{1}{M_2}\mathbf{u}_{1\ell}^{\mathrm{T}}\sum_{i\in J_2}\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*(\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*)^{\mathrm{T}}\mathbf{u}_{1\ell}\big)^{1/2}
\end{aligned} \tag{S5.35}$$

$$\to 0, \quad \text{almost surely.}$$

It follows that

$$\max_{1\le\ell\le p}\mathbf{u}_{1\ell}^{\mathrm{T}}\big[\frac{1}{M_2}\sum_{i\in J_2}(\widetilde{\mathbf{V}}_i^*+\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*)(\widetilde{\mathbf{V}}_i^*+\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*)^{\mathrm{T}}-\frac{1}{M_2}\sum_{i\in J_2}\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*(\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*)^{\mathrm{T}}\big]\mathbf{u}_{1\ell}$$

$$\le\max_{1\le\ell\le p}\frac{1}{M_2}\sum_{i\in J_2}\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\mathbf{V}}_i^*(\widetilde{\mathbf{V}}_i^*)^{\mathrm{T}}\mathbf{u}_{1\ell}+2\max_{1\le\ell\le p}\frac{1}{M_2}\sum_{i\in J_2}\mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\mathbf{V}}_i^*(\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*)^{\mathrm{T}}\mathbf{u}_{1\ell}\to 0,$$

almost surely. By (S5.34) and (C.5) of Xia and Zheng (2018), almost surely,

$$\max_{1\le i\le M}|\widetilde{\mathbf{V}}^*+\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^*|^2/p\to 1.$$

Thus (S5.30) follows.

Next we show (S5.28). Let $\mathbf{R}_i=\mathbf{R}_{i1}+\mathbf{R}_{i2}+\mathbf{R}_{i3}=(R_i^{(1)},\ldots,R_i^{(p)})^{\mathrm{T}}$. By (S4.19) and (S4.21), for any $\kappa\ge 1$

$$\max_{1\le i\le M,1\le q\le p}E|R_i^{(q)}|^{2\kappa}\le Cn^{-\kappa},$$

which indicates that

$$\max_{1\le i\le M}E|\mathbf{R}_i|^{2\kappa}=E(\sum_{j=1}^{p}|R_i^{(q)}|^2)^{\kappa}\le Cp^{\kappa}n^{-\kappa}. \tag{S5.36}$$

It follows that for any $\kappa\ge 1$ and $\varepsilon>0$

$$P\left(\max_{1\le i\le M}p|\mathbf{R}_i|^2\ge\varepsilon\right)\le\varepsilon^{-\kappa}\frac{pp^{\kappa}p^{\kappa}}{n^{\kappa}}\to 0.$$

Thus $\max_{1\le i\le M}p|\mathbf{R}_i|^2\to 0$ almost surely by Borel-Cantelli Lemma and Assumption (C.v). Togethering with (S4.19), we can show that

$$\|\frac{p}{M_2}\sum_{i\in J_2}\frac{\mathbf{R}_i\mathbf{R}_i^{\mathrm{T}}}{|\mathbf{\Phi}_i|^2}\|\le\max_{1\le i\le M}C\cdot p|\mathbf{R}_i|^2\to 0,\quad\text{almost surely.}$$

Using the above result, (S4.19) and a similar argument in (S5.35), we have

$$\max_{1\le\ell\le p}\mathbf{u}_{1\ell}^{\mathrm{T}}\left(\frac{p}{M_2}\sum_{i\in J_2}\frac{\Delta\widetilde{\overline{\mathbf{X}}}_{2i}^{*}(\Delta\widetilde{\overline{\mathbf{X}}}_{2i}^{*})^{\mathrm{T}}}{|\mathbf{\Phi}_i|^2}-\frac{p}{M_2}\sum_{i\in J_2}\frac{\mathbf{\Phi}_i\mathbf{\Phi}_i^{\mathrm{T}}}{|\mathbf{\Phi}_i|^2}\right)\mathbf{u}_{1\ell} \tag{S5.37}$$
$$\le\max_{1\le\ell\le p}\frac{p}{M_2}\sum_{i\in J_2}\mathbf{u}_{1\ell}^{\mathrm{T}}\frac{\mathbf{R}_i\mathbf{R}_i^{\mathrm{T}}}{|\mathbf{\Phi}_i|^2}\mathbf{u}_{1\ell}+2\max_{1\le\ell\le p}\frac{p}{M_2}\sum_{i\in J_2}\mathbf{u}_{1\ell}^{\mathrm{T}}\frac{\mathbf{R}_i\mathbf{\Phi}_i^{\mathrm{T}}}{|\mathbf{\Phi}_i|^2}\mathbf{u}_{1\ell}\to 0,$$

almost surely, which indicates (S5.28).

Finally, we show (S5.27). By Assumptions (A.xi) and (A.viii), Lyapunov's inequality and the same argument in (C.8) of Xia and Zheng (2018), we have

$$\max_{1\le i\le M,1\le q\le p}E|(\Delta\widetilde{\overline{\boldsymbol{\varepsilon}}}_{2i}^{*})^{(q)}|^{2\nu}<\frac{C}{h^{\nu}}, \tag{S5.38}$$

where ν is the integer in Assumption (A.viii). For any $\varepsilon>0$, by Markov's inequality,

$$P\left(\max_{1\le i\le M}p|\Delta\widetilde{\overline{\boldsymbol{\varepsilon}}}_{2i}^{*}|^2\ge\varepsilon\right)\le\varepsilon^{-\nu}\frac{pp^{\nu}p^{\nu}}{h^{\nu}}=O(\frac{1}{p^{\beta\nu/(1-\beta)-2\nu-1}}),$$

which means that $p|\Delta\widetilde{\overline{\boldsymbol{\varepsilon}}}_{2i}^{*}|^2\to 0$ almost surely, by Borel-Cantelli Lemma and Assumption (C.v). Following a similar argument as (S5.37), we have (S5.27). Thus the proof of Theorem 3 completes.

Now, we consider data from the factor model

$$d\mathbf{X}_t=\boldsymbol{\mu}_t dt+\gamma_t\mathbf{\Lambda}_f d\mathbf{W}_t^{*}+\gamma_t\mathbf{\Lambda}_B d\mathbf{W}_t.$$

WLOG, we may assume that $\boldsymbol{\mu}_t\equiv 0$. If $p^{-1}\mathrm{tr}(\mathbf{\Lambda}_B\mathbf{\Lambda}_B^{\mathrm{T}})=1$ and $p^{-1}\mathrm{tr}(\mathbf{\Lambda}_f\mathbf{\Lambda}_f^{\mathrm{T}})\to\beta^{*}$ as $p\to\infty$, the increment $\Delta\mathbf{X}_i=\sqrt{\omega_i}(\mathbf{\Lambda}_f\mathbf{f}_i+\mathbf{\Lambda}_B\mathbf{z}_i)$, where $\omega_i=$

$\int_{t_{i-1}}^{t_i} \gamma_t^2 dt$, $\mathbf{f}_i$ are i.i.d. $N(0, \mathbf{I}_r)$, $\mathbf{z}_i$ are i.i.d. $N(0, \mathbf{I}_p)$, and $\mathbf{f}_i$ and $\mathbf{z}_i$ are independent. Then (S5.25) holds if we can further show that

$$\lim_{p\to\infty} \frac{3\sum_{i=1}^{M} |\Delta\widetilde{\widetilde{\mathbf{X}}}_{2i}^{*}|^2}{p(1+\beta^*)} = \theta, \qquad \text{a.s.,} \tag{S5.39}$$

and for any given unit vector $\mathbf{u}_{1\ell}$,

$$\begin{aligned}
&\max_i |\mathbf{u}_{1\ell}^{\mathrm{T}} \boldsymbol{\Lambda}_f \mathbf{f}_i \mathbf{f}_i^{\mathrm{T}} \boldsymbol{\Lambda}_f^{\mathrm{T}} \mathbf{u}_{1\ell} - \mathbf{u}_{1\ell}^{\mathrm{T}} \boldsymbol{\Lambda}_f \boldsymbol{\Lambda}_f^{\mathrm{T}} \mathbf{u}_{1\ell}| \ \to\ 0, \qquad \text{a.s.} \\
&\max_i |\mathbf{u}_{1\ell}^{\mathrm{T}} \boldsymbol{\Lambda}_B \mathbf{z}_i \mathbf{z}_i^{\mathrm{T}} \boldsymbol{\Lambda}_B^{\mathrm{T}} \mathbf{u}_{1\ell} - \mathbf{u}_{1\ell}^{\mathrm{T}} \boldsymbol{\Lambda}_B \boldsymbol{\Lambda}_B^{\mathrm{T}} \mathbf{u}_{1\ell}| \ \to\ 0, \qquad \text{a.s.} \qquad \text{(S5.40)} \\
\text{and } &\max_i |\mathbf{u}_{1\ell}^{\mathrm{T}} \boldsymbol{\Lambda}_f \mathbf{f}_i \mathbf{z}_i^{\mathrm{T}} \boldsymbol{\Lambda}_B^{\mathrm{T}} \mathbf{u}_{1\ell}| \ \to\ 0 \qquad \text{a.s..}
\end{aligned}$$

By the proof of (S4.23), (S5.39) follows from the fact that $\mathrm{E}(|\boldsymbol{\Lambda}_f \mathbf{f}_i + \boldsymbol{\Lambda}_B \mathbf{z}_i|^2/p) \to 1+\beta^*$ as $p\to\infty$ and (S5.40) follows directly from the proof of Lemma 1 in Lam (2016). ■

S5.2 Proof of Theorem 4

To prove Theorem 4, we first provide the following lemma.

Lemma 1. *Let the assumptions in Theorem 4 hold and denote* $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_p)$. *Then the ESDs of* $\widetilde{\widetilde{\boldsymbol{\Xi}}}^*$ *and* $\widetilde{\widetilde{\boldsymbol{\Xi}}}_1^*$ *converge to the same probability* $\widetilde{F}^*$ *almost surely, where* $\widetilde{F}^*$ *is the LSD of matrices* $\frac{p}{M}\sum_{i=1}^{M} \frac{\Delta\widetilde{\widetilde{\mathbf{Y}}}_{2i}^{*}(\Delta\widetilde{\widetilde{\mathbf{Y}}}_{2i}^{*})^{\mathrm{T}}}{|\Delta\widetilde{\widetilde{\mathbf{Y}}}_{2i}^{*}|^2}$ *as* $p/M \to c > 0$. *Meanwhile, there exist positive functions* $\delta(\cdot)$ *such that almost surely,*

$$\begin{aligned}
p^{-1}\sum_{i=1}^{p} \mathbf{u}_{1i}^{\mathrm{T}} \breve{\boldsymbol{\Sigma}} \mathbf{u}_{1i} \mathbf{1}_{\{v_{1i}\le x\}} &\to \int_{-\infty}^{x} \delta(\lambda) d\widetilde{F}^*(\lambda) \quad \textit{and} \\
p^{-1}\sum_{i=1}^{p} \mathbf{u}_{i}^{\mathrm{T}} \breve{\boldsymbol{\Sigma}} \mathbf{u}_{i} \mathbf{1}_{\{v_{i}\le x\}} &\to \int_{-\infty}^{x} \delta(\lambda) d\widetilde{F}^*(\lambda),
\end{aligned}$$

where v_{1i} is the eigenvalues of $\widetilde{\breve{\boldsymbol{\Xi}}}_1^*$ with corresponding eigenvectors $\mathbf{u}_{1i}$ and v_i is the eigenvalues of $\widetilde{\breve{\boldsymbol{\Xi}}}^*$ with corresponding eigenvectors $\mathbf{u}_i$.

We do not write down the explicit form of $\delta(\cdot)$ because it is not important in the proof of any subsequent theorems. Interested readers may refer to (2.7) and (2.9) of Lam (2016).

Proof of Lemma 1. We just show the proof of second part in Lemma 1 because the first part can be obtained using the same argument as in the proof of Theorem 2.

Define $\widetilde{\breve{\boldsymbol{\Xi}}}_{1,\mathrm{spl}}$ and $\widetilde{\breve{\boldsymbol{\Xi}}}_{\mathrm{spl}}$ as

$$\widetilde{\breve{\boldsymbol{\Xi}}}_{1,\mathrm{spl}} = M_1^{-1} \sum_{i \in J_1} \boldsymbol{\Lambda} \widetilde{\mathbf{Z}}_i^* (\widetilde{\mathbf{Z}}_i^*)^{\mathrm{T}} \boldsymbol{\Lambda}^{\mathrm{T}}, \quad \widetilde{\breve{\boldsymbol{\Xi}}}_{\mathrm{spl}} = M^{-1} \sum_{i=1}^{M} \boldsymbol{\Lambda} \widetilde{\mathbf{Z}}_i^* (\widetilde{\mathbf{Z}}_i^*)^{\mathrm{T}} \boldsymbol{\Lambda}^{\mathrm{T}},$$

where $\widetilde{\mathbf{Z}}_i^*$ consists of i.i.d. standard normals. Let $v_{1j,\mathrm{spl}}$, $v_{j,\mathrm{spl}}$ be the eigenvalues of $\widetilde{\breve{\boldsymbol{\Xi}}}_{1,\mathrm{spl}}$ and $\widetilde{\breve{\boldsymbol{\Xi}}}_{\mathrm{spl}}$ with corresponding eigenvectors $\mathbf{u}_{1j,\mathrm{spl}}$, $\mathbf{u}_{j,\mathrm{spl}}$, respectively. By the Theorem 4 of Ledoit and Péché (2011), there exists a positive function $\delta(\cdot)$ such that almost surely,

$$p^{-1} \sum_{i=1}^{p} \mathbf{u}_{1i,\mathrm{spl}}^{\mathrm{T}} \breve{\boldsymbol{\Sigma}} \mathbf{u}_{1i,\mathrm{spl}} \mathbf{1}_{\{x \geq v_{1i,\mathrm{spl}}\}} \to \int_{-\infty}^{x} \delta(\lambda) d\widetilde{F}^*(\lambda),$$
$$p^{-1} \sum_{i=1}^{p} \mathbf{u}_{i,\mathrm{spl}}^{\mathrm{T}} \breve{\boldsymbol{\Sigma}} \mathbf{u}_{i,\mathrm{spl}} \mathbf{1}_{\{x \geq v_{i,\mathrm{spl}}\}} \to \int_{-\infty}^{x} \delta(\lambda) d\widetilde{F}^*(\lambda),$$

To prove the convergence of $p^{-1} \sum_{i=1}^{p} \mathbf{u}_{1i}^{\mathrm{T}} \breve{\boldsymbol{\Sigma}} \mathbf{u}_{1i} \mathbf{1}_{\{v_{1i} \leq x\}}$ and $p^{-1} \sum_{i=1}^{p} \mathbf{u}_{i}^{\mathrm{T}} \breve{\boldsymbol{\Sigma}} \mathbf{u}_{i} \mathbf{1}_{\{v_{i} \leq x\}}$, it suffices to show the convergence of their Stieltjes transforms, that is, almost surely,

$$p^{-1}\mathrm{tr}\big((\widetilde{\mathbf{\Xi}}_{1,\mathrm{spl}} - z\mathbf{I}_p)^{-1}\breve{\mathbf{\Sigma}}\big) - p^{-1}\mathrm{tr}\big((\widetilde{\mathbf{\Xi}}_1^{*} - z\mathbf{I}_p)^{-1}\breve{\mathbf{\Sigma}}\big) \to 0,$$
$$p^{-1}\mathrm{tr}\big((\widetilde{\mathbf{\Xi}}_{\mathrm{spl}} - z\mathbf{I}_p)^{-1}\breve{\mathbf{\Sigma}}\big) - p^{-1}\mathrm{tr}\big((\widetilde{\mathbf{\Xi}}^{*} - z\mathbf{I}_p)^{-1}\breve{\mathbf{\Sigma}}\big) \to 0. \tag{S5.41}$$

Next, we only show the first result in (S5.41) because the second result can be proved similarly. Observe that

$$\begin{aligned}&p^{-1}\mathrm{tr}\big((\widetilde{\mathbf{\Xi}}_1^{*} - z\mathbf{I}_p)^{-1}\breve{\mathbf{\Sigma}}\big) - p^{-1}\mathrm{tr}\big((\widetilde{\mathbf{\Xi}}_{1,\mathrm{spl}} - z\mathbf{I}_p)^{-1}\breve{\mathbf{\Sigma}}\big)\\ =&p^{-1}\mathrm{tr}\big((\widetilde{\mathbf{\Xi}}_{1,\mathrm{spl}} - \widetilde{\mathbf{\Xi}}_1^{*})(\widetilde{\mathbf{\Xi}}_1^{*} - z\mathbf{I}_p)^{-1}(\widetilde{\mathbf{\Xi}}_{1,\mathrm{spl}} - z\mathbf{I}_p)^{-1}\breve{\mathbf{\Sigma}}\big),\end{aligned}$$

By Theorem 5.11 of Bai and Silverstein (2010), there exists a constant C such that

$$\limsup_{p\to\infty} \|\frac{1}{M_1}\sum_{i\in J_1}\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^{*}(\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^{*})^{\mathrm{T}}\| \le C.$$

Combining with (S5.34) and following a similar argument as (S5.30), we have almost surely,

$$\|\frac{p}{M_1}\sum_{i\in J_1}\frac{\mathbf{\Phi}_i\mathbf{\Phi}_i^{\mathrm{T}}}{|\mathbf{\Phi}_i|^2} - \frac{1}{M_1}\sum_{i\in J_1}\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^{*}(\mathbf{\Lambda}\widetilde{\mathbf{Z}}_i^{*})^{\mathrm{T}}\| \to 0. \tag{S5.42}$$

Togethering with (S5.42), (S5.29) and a similar argument as (S5.27), (S5.28), (S5.31), (S5.32), the following statements hold, almost surely,

$$\|\frac{p}{M_1}\sum_{i\in J_1}\frac{\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}(\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i})^{\mathrm{T}}}{|\mathbf{\Phi}_i|^2}-\frac{p}{M_1}\sum_{i\in J_1}\frac{\mathbf{\Phi}_i\mathbf{\Phi}_i^{\mathrm{T}}}{|\mathbf{\Phi}_i|^2}\|\to 0,$$

$$\limsup_{p\to\infty}\|\frac{p}{M_1}\sum_{i\in J_1}\frac{\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}(\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i})^{\mathrm{T}}}{|\mathbf{\Phi}_i|^2}\|\leq C,$$

$$\|\frac{p}{M_1}\sum_{i\in J_1}\frac{\Delta\widetilde{\overline{\mathbf{Y}}}^*_{2i}(\Delta\widetilde{\overline{\mathbf{Y}}}^*_{2i})^{\mathrm{T}}}{|\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}|^2}-\frac{p}{M_1}\sum_{i\in J_1}\frac{\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}(\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i})^{\mathrm{T}}}{|\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}|^2}\|\to 0,$$

$$\limsup_{p\to\infty}\|\frac{p}{M_1}\sum_{i\in J_1}\frac{\Delta\widetilde{\overline{\mathbf{Y}}}^*_{2i}(\Delta\widetilde{\overline{\mathbf{Y}}}^*_{2i})^{\mathrm{T}}}{|\Delta\widetilde{\overline{\mathbf{X}}}^*_{2i}|^2}\|\leq C,$$

where C is a finite constant. Thus $\|\breve{\mathbf{\Xi}}_{1,\mathrm{spl}}-\breve{\mathbf{\Xi}}_1^*\|\to 0$ almost surely. Combining with the facts that $\|(\breve{\mathbf{\Xi}}_1^*-z\mathbf{I}_p)^{-1}\|\leq 1/\Im(z)$ and $\|(\breve{\mathbf{\Xi}}_{1,\mathrm{spl}}-z\mathbf{I}_p)^{-1}\|\leq 1/\Im(z)$, we conclud that

$$p^{-1}\mathrm{tr}\Big((\breve{\mathbf{\Xi}}_{1,\mathrm{spl}}-\breve{\mathbf{\Xi}}_1^*)(\breve{\mathbf{\Xi}}_1^*-z\mathbf{I}_p)^{-1}(\breve{\mathbf{\Xi}}_{1,\mathrm{spl}}-z\mathbf{I}_p)^{-1}\breve{\mathbf{\Sigma}}\Big)\leq\frac{\|\breve{\mathbf{\Sigma}}\|}{\Im^2(z)}\|\breve{\mathbf{\Xi}}_{1,\mathrm{spl}}-\breve{\mathbf{\Xi}}_1^*\|,$$

which converges to 0 almost surely. This completes the proof of Lemma 1.

■

Now, we begin the proof of Theorem 4. Recall that

$$\mathbf{\Sigma}_{\mathrm{oracle}}=\theta_n\cdot\mathbf{U}\mathrm{diag}(\mathbf{U}^{\mathrm{T}}\breve{\mathbf{\Sigma}}\mathbf{U})\mathbf{U}^{\mathrm{T}},\quad \mathbf{\Sigma}=\theta_n\breve{\mathbf{\Sigma}},\quad\text{and}\quad\widehat{\mathbf{\Sigma}}=\widehat{\theta}\,\widehat{\mathbf{\Xi}}^*,$$

where $\theta_n=\int_0^1\gamma_t^2dt$ and $\widehat{\theta}=(3/p)\sum_{i=1}^M|\Delta\widetilde{\overline{\mathbf{Y}}}^*_{2i}|^2$. Write $EL((\widehat{\mathbf{\Sigma}},\mathbf{\Sigma}_{\mathrm{oracle}}))$ as follows,

$$EL(\widehat{\mathbf{\Sigma}},\mathbf{\Sigma}_{\mathrm{oracle}})\leq 1-\left(\frac{p^{-1/2}\|(\hat{\theta}-\theta_n)\widehat{\mathbf{\Xi}}^*\|_F}{p^{-1/2}\|\mathbf{\Sigma}_{\mathrm{oracle}}-\mathbf{\Sigma}\|_F}+\frac{p^{-1/2}\theta_n\|\widehat{\mathbf{\Xi}}^*-\breve{\mathbf{\Sigma}}\|_F}{p^{-1/2}\|\mathbf{\Sigma}_{\mathrm{oracle}}-\mathbf{\Sigma}\|_F}\right)^{-2}.$$

By Assumption (A.iv) and (S4.22), $p^{-1/2}\|(\hat{\theta}-\theta_n)\widehat{\boldsymbol{\Xi}}^*\|_F \leq |\hat{\theta}-\theta_n| \max_{1\leq i\leq p} \mathbf{u}_{1\ell}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_2^*\mathbf{u}_{1\ell} \to 0$, almost surely. Further, from assumption that $p^{-1/2}||\boldsymbol{\Sigma}_{\text{oracle}} - \boldsymbol{\Sigma}||_F \not\to 0$ almost surely, it suffices to show that $p^{-1/2}\theta_n\|\widehat{\boldsymbol{\Xi}}^* - \breve{\boldsymbol{\Sigma}}\|_F$ and $p^{-1/2}\|\boldsymbol{\Sigma}_{\text{oracle}} - \boldsymbol{\Sigma}\|_F$ share the same limit.

By the triangle formula, we write

$$\begin{aligned}\frac{p^{-1}\theta_n^2\|\widehat{\boldsymbol{\Xi}}^* - \breve{\boldsymbol{\Sigma}}\|_F^2}{p^{-1}\|\boldsymbol{\Sigma}_{\text{oracle}} - \boldsymbol{\Sigma}\|_F^2} \leq & \frac{p^{-1}\sum_{i=1}^p(\mathbf{u}_{1i}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_2^*\mathbf{u}_{1i} - \mathbf{u}_{1i}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1i})^2}{p^{-1}\|\mathbf{U}\text{diag}(\mathbf{U}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{U})\mathbf{U}^{\mathrm{T}} - \breve{\boldsymbol{\Sigma}}\|_F^2} \\ & + \frac{p^{-1}\|\mathbf{U}_1\text{diag}(\mathbf{U}_1^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{U}_1)\mathbf{U}_1^{\mathrm{T}} - \breve{\boldsymbol{\Sigma}}\|_F^2}{p^{-1}\|\mathbf{U}\text{diag}(\mathbf{U}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{U})\mathbf{U}^{\mathrm{T}} - \breve{\boldsymbol{\Sigma}}\|_F^2}.\end{aligned}$$

By Assumption (A.vii) and Lemma 1, almost surely,

$$\begin{aligned}p^{-1}||\mathbf{U}_1\text{diag}(\mathbf{U}_1^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{U}_1)\mathbf{U}_1^{\mathrm{T}} - \breve{\boldsymbol{\Sigma}}||_F^2 &= p^{-1}\text{tr}(\breve{\boldsymbol{\Sigma}}^2) - p^{-1}\sum_{i=1}^p(\mathbf{u}_{1i}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1i})^2 \\ &\to \int \tau^2 d\breve{H}(\tau) - \int \delta^2(\lambda)d\widetilde{F}^*(\lambda),\end{aligned}$$

which is non-zero if $\breve{\boldsymbol{\Sigma}} \neq \mathbf{I}_p$ and also it is the limit of $p^{-1}||\mathbf{U}\text{diag}(\mathbf{U}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{U})\mathbf{U}^{\mathrm{T}} - \breve{\boldsymbol{\Sigma}}||_F^2$. By (S5.25), almost surely,

$$\frac{1}{p}\sum_{i=1}^p(\mathbf{u}_{1i}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_2^*\mathbf{u}_{1i} - \mathbf{u}_{1i}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1i})^2 \leq \max_{1\leq i\leq p}\left|\frac{\mathbf{u}_{1i}^{\mathrm{T}}\widetilde{\boldsymbol{\Xi}}_2^*\mathbf{u}_{1i} - \mathbf{u}_{1i}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1i}}{\mathbf{u}_{1i}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1i}}\right|^2 \cdot \max_{1\leq i\leq p}(\mathbf{u}_{1i}^{\mathrm{T}}\breve{\boldsymbol{\Sigma}}\mathbf{u}_{1i})^2 \to 0.$$

This completes the proof of Theorem 4. ■

S5.3 Proof of Theorem 5

The result of asymptotically positive-definite propery follows directly from Theorem 3 and Assumption (C.i). We only need tox show that $EL(\boldsymbol{\Sigma}, \widehat{\boldsymbol{\Sigma}}_B) \leq$

0 almost surely.

Define

$$\widehat{\mathbf{\Sigma}}^{(k)} = 3\frac{\sum_{i=1}^{M} |\Delta \widetilde{\overline{\mathbf{Y}}}^*_{2i}|^2}{p} (\widehat{\mathbf{\Xi}}^*)^{(k)}.$$

Then $\widehat{\mathbf{\Sigma}}_B = B^{-1}\sum_{k=1}^{B} \widehat{\mathbf{\Sigma}}^{(k)}$ for a finite number B. By the property of matrix norm, we have

$$\begin{aligned} EL(\mathbf{\Sigma}, \widehat{\mathbf{\Sigma}}_B) =& 1 - \left(\frac{\|\frac{1}{B}\sum_{k=1}^{B} \widehat{\mathbf{\Sigma}}^{(k)} - \mathbf{\Sigma}\|_F}{\|\mathbf{\Sigma}_{\text{ideal}} - \mathbf{\Sigma}\|_F}\right)^{-2} = 1 - \left(\frac{\|\frac{1}{B}\sum_{k=1}^{B} (\widehat{\mathbf{\Sigma}}^{(k)} - \mathbf{\Sigma})\|_F}{\|\mathbf{\Sigma}_{\text{ideal}} - \mathbf{\Sigma}\|_F}\right)^{-2} \\ \leq& 1 - \left(\frac{1}{B}\sum_{k=1}^{B} \cdot \frac{\|\widehat{\mathbf{\Sigma}}^{(k)} - \mathbf{\Sigma}\|_F}{\|\mathbf{\Sigma}_{\text{ideal}} - \mathbf{\Sigma}\|_F}\right)^{-2}. \end{aligned}$$

Following the same argument as the proof of Theorem 4, we know that $\|\widehat{\mathbf{\Sigma}}^{(k)} - \mathbf{\Sigma}\|_F / \|\mathbf{\Sigma}_{\text{ideal}} - \mathbf{\Sigma}\|_F \to 1$ almost surely. Therefore, $EL(\mathbf{\Sigma}, \widehat{\mathbf{\Sigma}}_B) \leq 0$ almost surely, which completes the proof of Theorem 5. ■

S6. Useful lemmas

Lemma 2. *(Lemma 1 in Zheng and Li (2011)). Suppose that for each* p, $\mathbf{v}_l = (v_l^1, \ldots, v_l^p)^{\mathrm{T}}$ *and* $\boldsymbol{w}_l = (w_l^1, \ldots, w_l^p)^{\mathrm{T}}, l = 1, \ldots, m$, *are all* p-*dimensional vectors. Define*

$$\widetilde{\boldsymbol{S}}_m = \sum_{l=1}^{m} (\boldsymbol{v}_l + \boldsymbol{w}_l)(\boldsymbol{v}_l + \boldsymbol{w}_l)^{\mathrm{T}} \text{ and } \boldsymbol{S}_m = \sum_{l=1}^{m} w_l (w_l)^{\mathrm{T}}.$$

Suppose the following conditions are satisfied:

- $m = m(p)$ *with* $\lim_{p\to\infty} p/m = c > 0$;

* *there exists a sequence* $\epsilon_p = o(1/\sqrt{p})$ *such that for all* p *and all* l*, all the entries of* $\boldsymbol{v}_l$ *are bounded by* ϵ_p *in absolute value;*
* $\limsup_{p\to\infty} \operatorname{tr}(\boldsymbol{S}_m)/p < \infty$ *almost surely.*

Then $L(F^{\widetilde{\boldsymbol{S}}_m}, F^{\boldsymbol{S}_m}) \to 0$ *almost surely, where for any two probability distribution functions* F *and* G*,* $L(F, G)$ *denotes the Levy distance between them.*

References

Bai, Z. and J. W. Silverstein (2010). *Spectral Analysis of Large Dimensional Random Matrices* (Second ed.). New York: Springer.

Jing, B. Y., Z. Liu, and X. B. Kong (2017). Estimating the volatility functionals with multiple transactions. *Econometric Theory* ***33***(**1**), 331–365.

Lam, C. (2016). Nonparametric eigenvalue-regularized precision or covariance matrix estimator. *The Annals of Statistics* ***44***(**3**), 928–953.

Lam, C., P. Feng, and C. Hu (2017). Nonlinear shrinkage estimation of large integrated covariance matrix. *Biometrika* ***104***(**2**), 481–488.

Ledoit, O. and S. Péché (2011). Eigenvectors of some large sample covariance matrix ensembles. *Probability Theory and Related Fields* ***151***(**1**), 233–264.

Liu, Z. (2017). Jump-robust estimation of volatility with simultaneous presence of microstructure noise and multiple observations. *Finance and Stochastics* ***21***(**1**), 427–469.

Liu, Z., X. B. Kong, and B. Y. Jing (2018). Estimating the integrated volatility using high-frequency data with zero durations. *Journal of Econometrics* ***204***(**1**), 18–32.

Silverstein, J. W. (1995). Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *Journal of Multivariate Analysis* ***55***(**2**), 331–339.

Xia, N. and X. Zheng (2018). On the inference about the spectra of high-dimensional covariance matrix based on noisy observations-with applications to integrated covolatility matrix inference in the presence of microstructure noise. *The Annals of Statistics* ***46***(**2**), 500–525.

Zheng, X. and Y. Li (2011). On the estimation of integrated covariance matrices of high dimensional diffusion processes. *Annals of Statistics* ***39***(**6**), 3121–3151.