Supplement to

When Moving-Average Models Meet High-Frequency Data: Uniform Inference on Volatility

Rui Da*
University of Chicago

Dacheng Xiu†
University of Chicago

This version: August 13, 2017

Abstract

This supplementary appendix contains all mathematical proofs.

*Address: 5807 S Woodlawn Avenue, Chicago, IL 60637, USA. E-mail address: `rui.da@chicagobooth.edu`.

†Address: 5807 S Woodlawn Avenue, Chicago, IL 60637, USA. E-mail address: `dacheng.xiu@chicagobooth.edu`.

Appendix A Proofs of Main Results

A.1 Notation

In this section we define some general notation which will be used throughout the remaining parts of this paper. In some sections below, we will define additional notation, which is effective only in the corresponding sections unless otherwise specified.

Part 1. For any $x = (x_1, x_2, \dots, x_q) \in \mathbb{R}^q$, we set $x_j = 0$ for any $j \geq q+1$ and denote

$$\begin{aligned} \|x\|^2_{(q')} &= \sum_{j=q'+1}^{\infty} x_j^2, \quad \|x\| = \|x\|_{(0)} \\ \|x\|_{1,(q')} &= \sum_{j=q'+1}^{\infty} |x_j|, \quad \|x\|_1 = \|x\|_{1,(0)}, \quad 1^\intercal \cdot x = \sum_{j=1}^{\infty} x_j. \end{aligned} \tag{A.1}$$

For a sequence of real random variables $\{a_n\}$, we use "$a_n = O_{\mathrm{P}}(n^{-\infty})$" to mean "$a_n = o_{\mathrm{P}}(n^{-\alpha})$ for all $\alpha > 0$ fixed". We use $a \vee b$ and $a \wedge b$ to denote $\max\{a, b\}$ and $\min\{a, b\}$, respectively. Let $\mathcal{M}_d$ denote the set of all $d \times d$ matrices. For a generic continuous-time stochastic process $\mathcal{V}$, we set $\mathcal{V}_j = \mathcal{V}_{t_j}$ whenever there is no ambiguity. Moreover, we define $\Delta_j^n \mathcal{V} = \mathcal{V}_j - \mathcal{V}_{j-1}$, regardless of whether $\mathcal{V}$ is cast in continuous time or discrete time.

Part 2. For any m and h, let $O_m, \mathbb{D}_m^h, \mathbb{F}_m^h \in \mathcal{M}_m$ be

$$(O_m)_{ij} = \sqrt{\frac{2}{m+1}} \sin \frac{ij\pi}{m+1}, \quad (\mathbb{D}_m^h)_{ij} = \delta_{i,j}(2 - \delta_{h,0}) \cos \frac{hj\pi}{m+1},$$

$$(\mathbb{F}_m^h)_{ij} = \mathbb{1}_{\{h=|i-j|\}} - \mathbb{1}_{\{h=i+j\}} - \mathbb{1}_{\{h=2m+2-(i+j)\}}.$$

For any m, any $(\sigma^2, \iota^2, \theta, \Delta_n)$ and any (σ^2, γ), define

$$D_m(\iota^2, \theta) = \frac{\iota^2}{2\pi} \sum_{h=0}^{\infty} \mathbb{D}_m^h \int_{-\pi}^{\pi} g(\lambda; \theta) e^{\mathrm{i}\lambda h} d\lambda, \quad V_m(\sigma^2, \iota^2, \theta, \Delta_n) = \sigma^2 \Delta_n \mathbb{I}_m + (2\mathbb{I}_m - \mathbb{D}_m^1) D_m(\iota^2, \theta),$$

$$\Omega_m(\sigma^2, \iota^2, \theta, \Delta_n) = O_m V_m(\sigma^2, \iota^2, \theta, \Delta_n) O_m,$$

$$D_m(\gamma) = \sum_{h=0}^{\infty} \gamma_h \mathbb{D}_m^h, \quad V_m(\sigma^2, \gamma, \Delta_n) = \sigma^2 \Delta_n \mathbb{I}_m + (2\mathbb{I}_m - \mathbb{D}_m^1) D_m(\gamma),$$

$$\Omega_m(\sigma^2, \gamma, \Delta_n) = O_m V_m(\sigma^2, \gamma, \Delta_n) O_m,$$

where $\mathbb{I}_m$ is the m-dimensional identity matrix. In addition, introduce short-hand notation:

$$\Omega_n(\sigma^2, \iota^2, \theta) = \Omega_n(\sigma^2, \iota^2, \theta, \Delta_n), \quad \Omega_n(\sigma^2, \gamma) = \Omega_n(\sigma^2, \gamma, \Delta_n),$$

$$\Omega_{D,n}(\sigma^2, \iota^2, \theta) = (\mathbb{I}_{J_d} \otimes \Omega_{n_d}(\sigma^2, \iota^2, \theta, \Delta_n)) \oplus \Omega_{n'_d}(\sigma^2, \iota^2, \theta, \Delta_n),$$

$$\Omega_{D,n}(\sigma^2,\gamma) = (\mathbb{I}_{J_d}\otimes\Omega_{n_d}(\sigma^2,\gamma,\Delta_n))\oplus\Omega_{n'_d}(\sigma^2,\gamma,\Delta_n),$$

where $n_d = [n^{7/8}]$, $J_d = [n^{1/8}]-1$ and $n'_d = n - n_d J_d$. For any n, let $\Omega_n^U, \Omega_n^B, \Omega_n^J, \Omega_n^Y, \Omega_n^{Y,B} \in \mathcal{M}_n$ be

$$\Omega_n^U = \Omega_{D,n}(0,(\iota^2)^{(n)},\theta^{(n)}), \quad (\Omega_n^B)_{ij} = \delta_{i,j}\int_{t_{i-1}^n}^{t_i^n}\sigma_s^2 ds,$$

$$(\Omega_n^J)_{ij} = \delta_{i,j}\sum_{t_{i-1}^n<s\le t_i^n}(\Delta X_s)^2, \quad \Omega_n^Y = \Omega_n^U+\Omega_n^B+\Omega_n^J, \quad \Omega_n^{Y,B} = \Omega_n^U+\Omega_n^B.$$

Moreover, let

$$\tau(n,i) = \begin{cases} t_{in_d}, & i\le J_d \\ T, & i\ge J_d+1\end{cases}, \quad \sigma_{C,t} = \sigma_{\tau(n,J_d)}\mathbb{1}_{\{\tau(n,J_d)\le t\}} + \sum_{i=0}^{J_d-1}\sigma_{\tau(n,i)}\mathbb{1}_{\{\tau(n,i)\le t<\tau(n,i+1)\}}$$

and $Y_n^C\in\mathbb{R}^n$, $\Omega_n^C, \Omega_n^{Y,C}\in\mathcal{M}_n$ be defined by

$$Y_{n,i}^C = \sigma_{C,t_{i-1}^n}\Delta_i^n W + \Delta_i^n U, \quad (\Omega_n^C)_{ij} = \delta_{i,j}(t_i^n - t_{i-1}^n)\sigma^2_{C,t_i^n}, \quad \Omega^{Y,C} = \Omega_n^U + \Omega_n^C.$$

Part 3. Introduce short-hand notation $\mathcal{L}$:

$$\mathcal{L}(A,B,C) = -\frac{1}{2}\log\det A - \frac{1}{2}\mathrm{tr}(B^{-1}C), \quad \mathcal{L}(A,C) = \mathcal{L}(A,A,C).$$

For any $(\sigma^2,\iota^2,\theta)$, let

$$L_{A,n}(\sigma^2,\iota^2,\theta) = \mathcal{L}(\Omega_n, Y_nY_n^\intercal), \quad L_{D,n}(\sigma^2,\iota^2,\theta) = \mathcal{L}(\Omega_n,\Omega_{D,n},Y_nY_n^\intercal), \quad \bar{L}_n(\sigma^2,\iota^2,\theta) = \mathcal{L}(\Omega_n,\Omega_{D,n},\Omega_n^Y),$$

$$L_{C,n}(\sigma^2,\iota^2,\theta) = \mathcal{L}(\Omega_n,\Omega_{D,n},Y_n^C(Y_n^C)^\intercal), \quad \bar{L}_{C,n}(\sigma^2,\iota^2,\theta) = \mathcal{L}(\Omega_n,\Omega_{D,n},\Omega_n^{Y,C}),$$

where we omit the arguments $(\sigma^2,\iota^2,\theta)$ of Ω_n and $\Omega_{D,n}$ for notational simplicity.

A.2 Proofs of Proposition 1, Theorem 1 and Corollary 1

We prove Proposition 1 by contradiction based on Step 8 of Section A.4. Adopting the notation $\mathcal{R}_f := -\frac{2}{n}(\bar{L}_n^\star(\varsigma^{(n)}(\widehat{q}_n)) - \bar{L}_n^\star(\varsigma_\infty^{(n)}))$ introduced in (A.24), we deduce two implications on $\mathcal{R}_f$ under the assumptions of Proposition 1. First, $(e1)$ in (A.25) yields that $\widehat{q}_{n,2} < q^\star$ leads to $\mathcal{R}_f \ge K^{-1}$ as $\theta^\star_{q^\star}\neq 0$. Second, $(e2)$ in (A.25) implies that $\widehat{q}_{n,2} > q^\star$ leads to $\mathcal{R}_f = 0$ as $\theta^\star\in\Theta(q^\star)$. Combined with (A.26), we conclude neither $\widehat{q}_{n,2} < q^\star$ nor $\widehat{q}_{n,2} > q^\star$ can occur with nonvanishing probability.

Given Proposition 1, the asymptotic result of volatility estimation in Theorem 1 is a special case of either Theorem 3 or Theorem 4. The joint asymptotic results of both volatility and noise estimations follow from extending (A.19) in Step 5 of Section A.4 under fixed q based on the same reasoning underlying Step 4 of Section A.4 and using Lemma D5 and Lemma D6.

Corollary 1 is a direct result of applying the delta method to Theorem 1.

A.3 Proof of Theorem 2

Step 1. (Main Proof) Let

$$\Sigma_n^{(0)} = \Omega_n^B + \Sigma_n(0,\iota^2,\theta), \quad \Sigma_n^{(1)} = \Omega_n^B + \Sigma_n(0,\zeta^2,0), \quad \widetilde{\Omega}_{n+2q}^B = (\Delta_n \mathbb{I}_q) \oplus \Omega_n^B \oplus (\Delta_n \mathbb{I}_q),$$

$$\Sigma_{n+2q}^{(2)} = \widetilde{\Omega}_{n+2q}^B + \Sigma_{n+2q}(0,\iota^2,\theta), \quad \Sigma_{n+2q}^{(3)} = \widetilde{\Omega}_{n+2q}^B + \Sigma_{n+2q}(0,\zeta^2,0).$$

Note that $\Sigma_n^{(0)}$ and $\Sigma_n^{(1)}$ are the covariance matrices of the observations of $\mathcal{E}_n^{(0)}$ and $\mathcal{E}_n^{(1)}$ respectively. We let $\mathcal{E}_n^{(2)}$ and $\mathcal{E}_n^{(3)}$ denote the Gaussian experiments whose covariance matrices are $\Sigma_{n+2q}^{(2)}$ and $\Sigma_{n+2q}^{(3)}$.

The desired result follows from

$$\Delta_{\mathrm{LC}}(\mathcal{E}_n^{(2)},\mathcal{E}_n^{(3)}) \underset{(a1)}{\lesssim} n^{-1/4} + n^{1/4-\alpha}(\log n)^{2\alpha}, \quad \Delta_{\mathrm{LC}}(\mathcal{E}_n^{(1)},\mathcal{E}_n^{(3)}) \underset{(a2)}{\lesssim} n^{-1/4}, \quad \Delta_{\mathrm{LC}}(\mathcal{E}_n^{(0)},\mathcal{E}_n^{(2)}) \underset{(a3)}{\lesssim} n^{-1/4}.$$

To show $(a1)$-$(a3)$, we employ that the Le Cam distance can be bounded by the Hellinger distance for Gaussian measures and further by the Hilber-Schmidt norm (cf. Section A.1 in Reiß (2011)). Concretely, if the probability measures of the observations, possibly with transformations and randomizations, from two statistical experiments are P and Q, the Le Cam distance between the two experiments is bounded by the Hellinger distance between P and Q, which we denote by $H(\mathrm{P},\mathrm{Q})$. Further, for multi-dimensional Gaussian laws $\mathcal{N}(0,\Sigma_A)$and $\mathcal{N}(0,\Sigma_B)$, we have

$$H^2(\mathcal{N}(0,\Sigma_A),\mathcal{N}(0,\Sigma_B)) \le 2\|\Sigma_A^{-1/2}(\Sigma_B-\Sigma_A)\Sigma_A^{-1/2}\|_{\mathrm{HS}}^2,$$

where $\|\cdot\|_{\mathrm{HS}}$ denotes the Hilbert-Schmidt norm.

Based on this method, we show $(a1)$ in Step 2 below, $(a2)$ in Step 4 below, while proving $(a3)$ is similar to Step 4.

Step 2. (Proof of $(a1)$) We start by constructing a transformation of $\mathcal{E}_n^{(3)}$. Let

$$\gamma_h = \iota^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h}, \quad P_{n+2q} = O_{n+2q}\left(\frac{1}{\zeta^2}\sum_{h=0}^{q} \gamma_h \mathbb{D}_{n+2q}^h\right)^{1/2} O_{n+2q}, \quad \widetilde{\Sigma}_{n+2q}^{(3)} = P_{n+2q}\Sigma_{n+2q}^{(3)}P_{n+2q}.$$

We suppress the subscript $n+2q$ of all matrices involved below in this step and Step 2. $(a1)$ follows from

$$\begin{aligned}\Delta_{\mathrm{LC}}(\mathcal{E}_n^{(2)},\mathcal{E}_n^{(3)}) \;&\underset{(b1)}{\le}\; H(\mathcal{N}(0,\Sigma^{(3)}),\mathcal{N}(0,\widetilde{\Sigma}^{(3)})) \underset{(b2)}{\lesssim} \|(\Sigma^{(3)})^{-1/2}(\widetilde{\Sigma}^{(3)}-\Sigma^{(3)})(\Sigma^{(3)})^{-1/2}\|_{\mathrm{HS}} \\ &\underset{(b3)}{\le}\; \|(\Sigma^{(3)})^{-1/2}(P^2-\mathbb{I})\widetilde{\Omega}^B(\Sigma^{(3)})^{-1/2}\|_{\mathrm{HS}} + \|(\Sigma^{(3)})^{-1/2}P(\widetilde{\Omega}^B P - P\widetilde{\Omega}^B)(\Sigma^{(3)})^{-1/2}\|_{\mathrm{HS}} \\ &=:\; \mathcal{R}_a + \mathcal{R}_b \underset{(b4)}{\lesssim} n^{-1/4} + n^{1/4-\alpha}(\log n)^{2\alpha},\end{aligned}$$

where $(b1)$ and $(b2)$ hold by the remark in Step 1, $(b3)$ holds by the triangle inequality, and $(b4)$ holds by Step 3 below.

Step 3. (Bounds of $\mathcal{R}_a$ and $\mathcal{R}_b$) The desired bound of $\mathcal{R}_a$ follows from

$$\begin{aligned}\mathcal{R}_a^2 &\underset{(c1)}{=} \operatorname{tr}((\Sigma^{(3)})^{-1/2}\widetilde{\Omega}^B(P^2-\mathbb{I})(\Sigma^{(3)})^{-1}(P^2-\mathbb{I})\widetilde{\Omega}^B(\Sigma^{(3)})^{-1/2}) \\ &\underset{(c2)}{\lesssim} \operatorname{tr}((\Sigma^{(3)})^{-1/2}\widetilde{\Omega}^B\widetilde{\Omega}^B(\Sigma^{(3)})^{-1/2}) \underset{(c3)}{\lesssim} \Delta_n^2\operatorname{tr}(\Omega(1,1,0,\Delta_n)^{-1}) \underset{(c4)}{\lesssim} n^{-1/2},\end{aligned}$$

where $(c1)$ holds by the definitions, $(c2)$ is proved below, $(c3)$ holds by $\widetilde{\Omega}^B \lesssim \mathbb{I}$ and $\Omega(1,1,0,\Delta_n) \lesssim \Sigma^{(3)}$, and $(c4)$ holds by Lemma D6.

Now we show $(c2)$. Observing $\Omega(0,1,0,\Delta_n) \lesssim \Sigma^{(3)}$, it suffices to derive

$$(P^2-\mathbb{I})\Omega(0,1,0,\Delta_n)^{-1}(P^2-\mathbb{I}) \lesssim \mathbb{I}.$$

Indeed, we have

$$\begin{aligned}O(P^2-\mathbb{I})\Omega(0,1,0,\Delta_n)^{-1}(P^2-\mathbb{I})O &\underset{(c5)}{=} \left(\frac{1}{\zeta^2}\sum_{h=0}^{q}\gamma_h\mathbb{D}^h-\mathbb{I}\right)V(0,1,0,\Delta_n)^{-1}\left(\frac{1}{\zeta^2}\sum_{h=0}^{q}\gamma_h\mathbb{D}^h-\mathbb{I}\right) \\ &\underset{(c6)}{\lesssim} V(0,1,0,\Delta_n) \underset{(c7)}{\lesssim} \mathbb{I}.\end{aligned}$$

Here $(c5)$ is due to definition, $(c6)$ can be verified using that all matrices involved are diagonal, and $(c7)$ is obvious.

The desired bound of $\mathcal{R}_b$ follows from

$$\begin{aligned}\mathcal{R}_b^2 &\underset{(d1)}{=} \operatorname{tr}((\Sigma^{(3)})^{-1/2}(\widetilde{\Omega}^BP-P\widetilde{\Omega}^B)^\intercal P(\Sigma^{(3)})^{-1}P(\widetilde{\Omega}^BP-P\widetilde{\Omega}^B)(\Sigma^{(3)})^{-1/2}) \\ &\underset{(d2)}{\lesssim} \operatorname{tr}(\Omega(1,1,0,\Delta_n)^{-1}(\widetilde{\Omega}^BP-P\widetilde{\Omega}^B)^\intercal\Omega(1,1,0,\Delta_n)^{-1}(\widetilde{\Omega}^BP-P\widetilde{\Omega}^B)) \\ &\underset{(d3)}{=} \sum_{i,j,k,l}\Omega(1,1,0,\Delta_n)^{-1}_{ij}P_{jk}(\widetilde{\Omega}^B_{kk}-\widetilde{\Omega}^B_{jj})\Omega(1,1,0,\Delta_n)^{-1}_{kl}P_{li}(\widetilde{\Omega}^B_{ll}-\widetilde{\Omega}^B_{ii}) \\ &\underset{(d4)}{\lesssim} \Delta_n^2\sum_{i,j,k,l}\Omega(1,1,0,\Delta_n)^{-1}_{ij}P_{jk}(\Delta_n(\log n)^2)^\alpha\Omega(1,1,0,\Delta_n)^{-1}_{kl}P_{li}(\Delta_n(\log n)^2)^\alpha \\ &\underset{(d5)}{\lesssim} \Delta_n^2(\Delta_n(\log n)^2)^{2\alpha}\operatorname{tr}(\Omega(1,1,0,\Delta_n)^{-2}) \underset{(d6)}{\lesssim} n^{1/2-2\alpha}(\log n)^{4\alpha}.\end{aligned}$$

Here $(d1)$ holds by definition, $(d2)$ holds by $\Omega(1,1,0,\Delta_n) \lesssim \Sigma^{(3)}$ and $P \lesssim \mathbb{I}$, $(d3)$ holds by definition, $(d4)$ holds by the α Hölder continuity of σ_s and $P_{ij} \sim n^{-\infty}$ for all $|i-j| \gtrsim (\log n)^2$, $(d5)$ holds by $P \lesssim \mathbb{I}$, and $(d6)$ holds by Lemma D6.

Step 4. (Proof of $(a2)$) We observe that for any observation $\mathcal{Y} = \{Y_j\}_{j=1}^{n+2q}$ whose covariance matrix is $\Sigma^{(3)}_{n+2q}$, the covariance matrix of $\{Y_j\}_{j=q+1}^{n+2q}$ is $\Sigma^{(1)}_n$, which means $\mathcal{E}^{(1)}_n$ is asymptotically less informative than $\mathcal{E}^{(3)}_n$. Hence, we only need to show $\mathcal{E}^{(1)}_n$ is asymptotically more informative than

$\mathcal{E}_n^{(3)}$. To this end, based on (direct) observation $\mathcal{Y} := \{Y_j\}_{j=q+1}^{n+2q}$ from $\mathcal{E}_n^{(3)}$, we explicitly construct the statistic $\widetilde{\mathcal{Y}} := \{\widetilde{Y}_j\}_{j=1}^{n+4q}$, whose covariance matrix is denoted by $\widetilde{\Sigma}_{n+2q}^{(1)}$. Our proof can conclude once we show $H^2(\mathcal{L}(\mathcal{Y}), \mathcal{L}(\widetilde{\mathcal{Y}}))$ tends to zero. We define

$$\widetilde{Y}_j = \begin{cases} Y_j, & q+1 \le j \le n+2q \\ (\mathcal{U}_{j-1} - \mathcal{U}_{j-1})\zeta & 1 \le j \le q, \text{ or } n+q+1 \le j \le n+2q \\ \widetilde{\epsilon}_+ - \mathcal{U}_{j-1}\zeta & j = q \\ \mathcal{U}_j\zeta - \widetilde{\epsilon}_- & j = n+q+1 \end{cases},$$

where $\{\mathcal{U}_j\}$ are i.i.d. standard normal and independent on $\mathcal{Y}$, and

$$\widetilde{\epsilon}_+ := -\sum_{j=1}^{[\sqrt{n}]}\left(1 - \frac{j}{[\sqrt{n}]}\right)Y_{j+q}, \quad \widetilde{\epsilon}_- := \sum_{j=1}^{[\sqrt{n}]}\left(1 - \frac{j}{[\sqrt{n}]}\right)Y_{n+q+1-j}.$$

One can verify, for all $1 \le i \le j \le n+2q$,

$$(\widetilde{\Sigma}_{n+2q}^{(1)} - \Sigma_{n+2q}^{(3)})_{ij} \underset{(e1)}{=} \begin{cases} -\Delta_n & i = j \le q-1, \text{ or }, i = j \ge n+q+1 \\ \zeta^2 \frac{[\sqrt{n}]-1}{[\sqrt{n}]^2} + \sum_{k=1}^{[\sqrt{n}]}(\Omega_n^B)_{kk}\left(1 - \frac{k}{[\sqrt{n}]}\right)^2 - \Delta_n & i = j = q \\ \zeta^2 \frac{[\sqrt{n}]-1}{[\sqrt{n}]^2} + \sum_{k=1}^{[\sqrt{n}]}(\Omega_n^B)_{n+1-k,n+1-k}\left(1 - \frac{k}{[\sqrt{n}]}\right)^2 - \Delta_n & i = j = n+q+1 \\ -(\Omega_n^B)_{j-q,j-q}\left(1 - \frac{j}{[\sqrt{n}]}\right) & i = q, 1 \le j-q \le [\sqrt{n}] \\ (\Omega_n^B)_{i-q,i-q}\left(1 - \frac{i}{[\sqrt{n}]}\right) & j = n+q+1, 1 \le n+q+1-i \le [\sqrt{n}] \\ 0 & \text{otherwise} \end{cases}.$$

Hence, we can bound the Hellinger distance as (we suppress the subscript $n+2q$)

$$\begin{aligned} H^2(\mathcal{L}(\mathcal{Y}), \mathcal{L}(\widetilde{\mathcal{Y}})) &\underset{(e2)}{\le} 2\|(\Sigma^{(3)})^{-1/2}(\widetilde{\Sigma}^{(1)} - \Sigma^{(3)})(\Sigma^{(3)})^{-1/2}\|_{\mathrm{HS}}^2 \\ &\underset{(e3)}{=} 2\sum_{i,j,k,l}(\Sigma^{(3)})_{ij}^{-1}(\widetilde{\Sigma}^{(1)} - \Sigma^{(3)})_{jk}(\Sigma^{(3)})_{kl}^{-1}(\widetilde{\Sigma}^{(1)} - \Sigma^{(3)})_{li} \underset{(e4)}{\lesssim} n^{-1/2}, \end{aligned}$$

where $(e2)$ holds by Step 1, $(e3)$ holds by definition, and $(e4)$ holds by $(e1)$ and Lemma D3.

A.4 Proof of Theorem 3

Step 1. (Notation) This step contains the notation needed in the proof. We will always recall the items introduced here when we use them in the remaining steps.

For any q, let

$$\left.\begin{array}{l}\widehat{\zeta}_n^2(q) = \widehat{\iota}_n^2(q)(1+1^\intercal\cdot\widehat{\theta}_n(q))^2,\\ \theta^{(n)}(q) = \arg\min_{\theta\in\Theta(q)}\int_{-\pi}^{\pi}\frac{g(\lambda;\theta^{(n)})}{g(\lambda;\theta)}d\lambda, \quad \iota^{(n)}(q)^2 = \min_{\theta\in\Theta(q)}\frac{(\iota^{(n)})^2}{2\pi}\int_{-\pi}^{\pi}\frac{g(\lambda;\theta^{(n)})}{g(\lambda;\theta)}d\lambda,\\ \zeta^{(n)}(q)^2 = \iota^{(n)}(q)^2(1+1^\intercal\cdot\theta^{(n)}(q))^2, \quad \sigma^{(n)}(q)^2 = \arg\min_{\sigma^2}\left(\sigma\left(2-\frac{(\zeta^2)^{(n)}}{\zeta^{(n)}(q)^2}\right)+\frac{C_T}{\sigma}\right).\end{array}\right\} \tag{A.2}$$

Moreover, define

$$q_n^\star = \arg\min_q (qn^{-1} + \|\theta^{(n)}\|_{(q)}^2), \tag{A.3}$$

and

$$\left.\begin{array}{l}h_n(\iota^2,\theta) = \log\iota^2 + \frac{(\iota^2)^{(n)}}{\iota^2}\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{g(\lambda;\theta^{(n)})}{g(\lambda;\theta)}d\lambda, \quad \widetilde{h}_n(\sigma^2,\zeta^2) = \frac{\sigma}{\zeta} + \frac{C_T}{2T\zeta\sigma} - \frac{\sigma(\zeta^2)^{(n)}}{2\zeta^3},\\ \bar{L}_n^\star(\sigma^2,\iota^2,\theta) = -\frac{n}{2}h_n(\iota^2,\theta) - \frac{n}{2}\sqrt{\Delta_n}\widetilde{h}_n(\sigma^2,\zeta^2), \text{ with } \zeta^2 = \iota^2(1+1^\intercal\cdot\theta)^2.\end{array}\right\} \tag{A.4}$$

We introduce more notation to facilitate our analysis. Define

$$\Pi_n^\varsigma(q) = \{\varsigma = (\varsigma_1,\varsigma_2,\ldots,\varsigma_{q+2})^\intercal \in \mathbb{R}^{q+2} : (\varsigma_1 n^{1/4},\varsigma_2,\varsigma_3,\ldots,\varsigma_{q+2})^\intercal \in \Pi_n(q)\}. \tag{A.5}$$

For any $\varsigma \in \Pi_n^\varsigma(q)$ and $S_n \in \{L_n, L_{A,n}, L_{D,n}, \bar{L}_n, \bar{L}_n^\star\}$, we let

$$h_n(\varsigma) = h_n(\iota^2,\theta), \quad \widetilde{h}_n(\varsigma) = \widetilde{h}_n(\sigma^2,\zeta^2), \quad S_n(\varsigma) = S_n(\sigma^2,\iota^2,\theta), \tag{A.6}$$

with $\iota^2 = \varsigma_2$, $\theta = (\varsigma_3,\ldots,\varsigma_{q+2})^\intercal$, $\sigma^2 = \varsigma_1 n^{1/4}$, and $\zeta^2 = \iota^2(1+1^\intercal\cdot\theta)^2$. Furthermore, with any $\varsigma \in \Pi_n^\varsigma(q)$ and $s \in \{A, D\}$ we write $\Xi_n(\varsigma), \bar{\Xi}_n(\varsigma), \Xi_{s,n}(\varsigma) \in \mathbb{R}^{q+2}$ and $\partial\Xi_n(\varsigma), \partial\Xi_{s,n}(\varsigma), \partial\bar{\Xi}_n(\varsigma), \partial\bar{\Xi}_n^\star(\varsigma) \in \mathcal{M}_{q+2}$ such that

$$(\Xi_n(\varsigma)_j, \bar{\Xi}_n(\varsigma)_j, \Xi_{s,n}(\varsigma)_j) = -\frac{1}{n}\frac{\partial}{\partial\varsigma_j}(L_n(\varsigma), \bar{L}_n(\varsigma), L_{s,n}(\varsigma)), \tag{A.7}$$

$$(\partial\Xi_n(\varsigma)_{ij}, \partial\Xi_{s,n}(\varsigma)_{ij}, \partial\bar{\Xi}_n(\varsigma)_{ij}, \partial\bar{\Xi}_n^\star(\varsigma)_{ij}) = -\frac{1}{n}\frac{\partial^2}{\partial\varsigma_i\partial\varsigma_j}(L_n(\varsigma), L_{s,n}(\varsigma), \bar{L}_n(\varsigma), \bar{L}_n^\star(\varsigma)). \tag{A.8}$$

Moreover, we set

$$\left.\begin{array}{l}\widehat{\varsigma}_n(q) = (\widehat{\sigma}_n^2(q)n^{-1/4}, \widehat{\iota}_n^2(q), \widehat{\theta}_n(q)), \quad \varsigma^{(n)}(q) = (\sigma^{(n)}(q)^2 n^{-1/4}, \iota^{(n)}(q)^2, \theta^{(n)}(q)),\\ \varsigma_\infty^{(n)} = (C_T n^{-1/4}, (\iota^{(n)})^2, \theta^{(n)}).\end{array}\right\} \tag{A.9}$$

Furthermore, define $\partial\bar{\Xi}_n^{\star\star}(a^2,q) \in \mathcal{M}_{q+2}$ as

$$\partial\bar{\Xi}_n^{\star\star}(a^2,q)_{ij} = \delta_{1,i}\delta_{1,j}\frac{\sqrt{T}}{8a^3\zeta^{(n)}} + \delta_{2,i}\delta_{2,j}\frac{1}{2(\iota^4)^{(n)}} + \mathbb{1}_{\{3\leqslant i,j\leqslant q+2\}}\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{e^{\mathrm{i}(i-j)\lambda}}{g(\lambda;\theta^{(n)})}d\lambda \tag{A.10}$$

and let $\partial\bar{\Xi}_n^{\star\star}(q) = \partial\bar{\Xi}_n^{\star\star}(C_T, q)$. Finally, with any sequence $\{(q_n) : n \geq 1\}$ we define

$$\left.\begin{aligned} \widehat{\mathcal{R}}_n(q_n) &:= (\widehat{\iota}_n^2(q_n) - (\iota^{(n)})^2)^2 + \|\widehat{\theta}_n(q_n) - \theta^{(n)}\|^2 + (\widehat{\sigma}_n^2(q_n) - C_T)^2, \\ \mathcal{R}^{(n)}(q_n) &:= (\iota^{(n)}(q_n)^2 - (\iota^{(n)})^2)^2 + \|\theta^{(n)}(q_n) - \theta^{(n)}\|^2 + (\sigma^{(n)}(q_n)^2 - C_T)^2. \end{aligned}\right\} \tag{A.11}$$

Step 2. (Framework of the proof) We decompose the proof of Theorem 3 into four parts.

First, we show in Step 3 below that under $\widehat{\mathcal{R}}_n(q_n) = o_{\mathrm{P}}(1)$, $\mathcal{R}^{(n)}(q_n) = o_{\mathrm{P}}(1)$, $(\zeta^{(n)})^2 \to b^2 \in (0, \infty)$ and $q_n n^{-1/4} \to 0$, it holds

$$\frac{(\zeta^{(n)})^{-1/2}\Delta_n^{-1/4}(\widehat{\sigma}^2(q_n) - \sigma^{(n)}(q_n)^2)}{\sqrt{\frac{1}{T}(5C(4)_T C_T^{-1/2} + 3C_T^{-3/2})}} \xrightarrow{\mathcal{L}-s} \mathcal{N}(0,1).$$

Recall we define $\widehat{\mathcal{R}}_n(q_n)$ and $\mathcal{R}^{(n)}(q_n)$ in (A.11)

Second, by Step 7 , we have that the first claim still holds without assuming $(\zeta^2)^{(n)} \to b^2 \in (0, \infty)$.

Third, in Step 8 below, we derive that for $j \in \{1, 2\}$,

$$\widehat{\mathcal{R}}_n(\widehat{q}_{n,j}) = o_{\mathrm{P}}(1), \quad \text{and} \quad \mathcal{R}^{(n)}(\widehat{q}_{n,j}) = o_{\mathrm{P}}(1).$$

Fourth, and finally, we prove in Step 9 that for $j \in \{1, 2\}$,

$$\sigma^{(n)}(\widehat{q}_{n,j})^2 = C_T + o_{\mathrm{P}}(n^{-1/4}), \quad \text{and} \quad \widehat{q}_{n,j} = o_{\mathrm{P}}(n^{1/4}).$$

Other steps provide auxiliary results.

Step 3. (Proof of the first claim in Step 2) For notational simplicity, from this step to Step 6 we omit the arguments q_n in $\widehat{\varsigma}_n(q_n)$ $\varsigma^{(n)}(q_n)$, $\partial\bar{\Xi}_n^{\star\star}(a^2, q_n)$ and $\partial\bar{\Xi}_n^{\star\star}(q_n)$ (but not other quantities) whenever possible.

In view of the definitions of $\widehat{\varsigma}_n$ given by (A.9) and of $\Xi_n(\varsigma)$ given by (A.7) and (A.6), we have for any sequence $\{q_n\}$,

$$\Xi_n(\widehat{\varsigma}_n) = 0_{q_n+2}. \tag{A.12}$$

This claim comes from two facts. First, $(\widehat{\sigma}_n^2(q_n), \widehat{\iota}_n^2(q_n), \widehat{\theta}_n(q_n))$ maximizes the quasi-log likelihood over $\Pi(q_n)$. Second, $(\widehat{\sigma}_n^2(q_n), \widehat{\iota}_n^2(q_n), \widehat{\theta}_n(q_n))$ is an interior point (with probability approaching one) from $\widehat{\mathcal{R}}_n(q_n) = o_{\mathrm{P}}(1)$.

Next, we apply the mean value theorem to the FOCs (A.12). Formally, in view of the definition of $\partial\Xi_n(\varsigma)$ given by (A.8), we write for all $1 \leq i \leq q_n + 2$,

$$\Xi_n(\varsigma^{(n)})_i + \sum_{j=1}^{q_n+2} \partial\Xi_n(\bar{\varsigma}_n(i))_{ij}\big(\widehat{\varsigma}_n - \varsigma^{(n)}\big)_j = 0. \tag{A.13}$$

Here $\bar{\varsigma}_n(i) = \alpha_n(i)\widehat{\varsigma}_n + (1 - \alpha_n(i))\varsigma^{(n)}$ with $\alpha_n(i) \in [0, 1]$. We introduce some simplifying notation. Let

$\bar{\varsigma}_n = (\bar{\varsigma}_n(1), \bar{\varsigma}_n(2), \ldots, \bar{\varsigma}_n(q_n+2))$ and let $\partial\Xi_n(\bar{\varsigma}_n) \in \mathcal{M}_{q_n+2}$ be defined by $\partial\Xi_n(\bar{\varsigma}_n)_{ij} = \partial\Xi_n(\bar{\varsigma}_n(i))_{ij}$. Using them, (A.13) can be rewritten as a matrix equation:

$$\widehat{\varsigma}_n - \varsigma^{(n)} = -\partial\Xi_n(\bar{\varsigma}_n)^{-1}\Xi_n(\varsigma^{(n)}). \tag{A.14}$$

Decompose the first row of (A.14)

$$\left(\widehat{\varsigma}_n - \varsigma^{(n)}\right)_1 = \Big(-(\partial\bar{\Xi}_n^{\star\star})^{-1}\Xi_n(\varsigma^{(n)})\Big)_1 + \Big(\underbrace{\Big[(\partial\bar{\Xi}_n^{\star\star})^{-1} - \partial\Xi_n(\bar{\varsigma}_n)^{-1}\Big]\Xi_n(\varsigma^{(n)})}_{=:\mathcal{R}}\Big)_1.$$

Here the subscripts $_1$ denote the first components of the vectors. We define $\partial\bar{\Xi}_n^{\star\star}$ in (A.10). Using $\mathcal{R} = o_{\mathrm{P}}(n^{-1/2})$ proved in Step 5 below, we obtain

$$\begin{aligned} n^{1/4}(\widehat{\sigma}^2(q_n) - \sigma^{(n)}(q_n)^2) \underset{(a1)}{=} n^{1/2}(\widehat{\varsigma}_n - \varsigma^{(n)})_1 \quad &\underset{(a2)}{=} \quad -\frac{8C_T^{3/2}\zeta^{(n)}}{T^2} n^{1/2}\Xi_n(\varsigma^{(n)})_1 + o_{\mathrm{P}}(1) \\ &\underset{(a3)}{\xrightarrow{\mathcal{L}-s}} \quad \mathcal{N}\left(0, \frac{b}{\sqrt{T}}(5C_T^{-1/2}C(4)_T + 3C_T^{3/2})\right). \end{aligned}$$

Here $(a1)$ holds by the definitions of $\widehat{\varsigma}_n$ and $\varsigma^{(n)}$, $(a2)$ holds by the definition of $\partial\bar{\Xi}_n^{\star\star}$, and $(a3)$ holds by Step 4 below.

Step 4. (CLT of the score function) In this step we show

$$n^{1/2}\Xi_n(\varsigma^{(n)})_1 \xrightarrow{\mathcal{L}-s} \mathcal{N}\left(0, \frac{5\sqrt{T}}{64}\frac{C(4)_T}{C_T^{7/2}b} + \frac{3\sqrt{T}}{64}\frac{1}{C_T^{3/2}b}\right). \tag{A.15}$$

This step is challenging in that there is no analytic expression of Σ_n^{-1} that appears in L_n and Ξ_n. To proceed, we need to precisely approximate Σ_n with some quantity that can be easily inverted. Proposition 3 demonstrates the failure of the classic Whittle approximation.[1] Instead, we approximate Σ_n with Ω_n, introduced in Part 2 of Section A.1. The inverse of Ω_n is given by Lemma D3. Moreover, this approximation is sufficiently accurate. Formally, Lemma B1 shows that the asymptotical behavior of $n^{1/2}\Xi_n(\varsigma^{(n)})_1$ is the same as that of $n^{1/2}\Xi_{A,n}(\varsigma^{(n)})_1$:

$$\left|\Xi_n(\varsigma^{(n)})_1 - \Xi_{A,n}(\varsigma^{(n)})_1\right| = o_{\mathrm{P}}(n^{-1/2}). \tag{A.16}$$

Here $\Xi_{A,n}(\varsigma)$ is defined based on Ω_n by (A.7), (A.6) and Part 3 of Section A.1.

The inverse of Ω_n in Lemma D3 reveals that $(\Omega_n^{-1})_{i,j} \neq 0$ for all i, j. Hence showing the central limit theorem for $\Xi_{A,n}(\varsigma^{(n)})_1$ becomes nonstandard. To resolve this, we further approximate Ω_n with a block-diagonal counterpart $\Omega_{D,n}$, introduced in Part 2 of Section A.1. Again, we show this

[1]See the last paragraph in Step 4 of Proof of Lemma B1 for the explanation from the technical perspective.

approximation is precise enough:

$$\left|\Xi_{A,n}(\varsigma^{(n)})_1 - \Xi_{D,n}(\varsigma^{(n)})_1\right| = o_{\mathrm{P}}(n^{-1/2}), \tag{A.17}$$

which holds by Lemma B2. And here $\Xi_{D,n}(\varsigma)$ is defined based on $\Omega_{D,n}$ by (A.7), (A.6) and Part 3 of Section A.1.

Given the block-diagonal structure of $\Omega_{D,n}$, we can make use of central limit theorems for triangular arrays (see, e.g., Theorem 2.2.15 in Jacod and Protter (2011)). Indeed, Lemma B3 shows that under $(\zeta^{(n)})^2 \to b^2 \in (0,\infty)$,

$$n^{1/2}\Xi_{D,n}(\varsigma^{(n)})_1 \xrightarrow{\mathcal{L}-s} \mathcal{N}\left(0, \frac{5\sqrt{T}}{64b}\frac{C(4)_T}{C_T^{7/2}} + \frac{3\sqrt{T}}{64}\frac{1}{C_T^{3/2}b}\right). \tag{A.18}$$

In view of the triangle inequality, we readily deduce (A.15).

Step 5. (Auxiliary: Bound on $\mathcal{R}$) In this step we show $\mathcal{R} = o_{\mathrm{P}}(n^{-1/2})$.

Recall the definition of $\partial\bar{\Xi}_n^{\star\star}$ in (A.10). Using $(\partial\bar{\Xi}_n^{\star\star})^{-1}_{1j} = 0$ for $j \neq 1$ by construction, we write

$$\mathcal{R} = \left((\partial\bar{\Xi}_n^{\star\star})^{-1}_{11} - \partial\Xi_n(\bar{\varsigma}_n)^{-1}_{11}\right) \times \Xi_n(\varsigma^{(n)})_1 + \sum_{j\geq 2}\partial\Xi_n(\bar{\varsigma}_n)^{-1}_{1j}\Xi_n(\varsigma^{(n)})_j =: \mathcal{R}_a \times \Xi_n(\varsigma^{(n)})_1 + \mathcal{R}_b.$$

We bound the first term by (A.15) and $\mathcal{R}_a = o_{\mathrm{P}}(1)$. Meanwhile, using Cauchy-Shwartz inequality, we can deduce $\mathcal{R}_b = o_{\mathrm{P}}(n^{-1/2})$ from

$$\mathcal{R}_c := \sum_{i\geq 2}\left(\partial\Xi_n(\bar{\varsigma}_n)^{-1}_{1i}\right)^2 \underset{(b1)}{=} o_{\mathrm{P}}(n^{-1/4}) \quad \text{and} \quad \mathcal{R}_d := \sum_{j\geq 2}\left(\Xi_n(\varsigma^{(n)})_j\right)^2 \underset{(b2)}{=} o_{\mathrm{P}}(n^{-3/4}).$$

First, we show the bound on $\mathcal{R}_d$. Using Chebyshev's inequality and recalling that we require $n^{1/4}q_n \to 0$, it suffices to show that for all $2 \leq j \leq q_n + 2$,

$$\mathbb{E}\left(\left|\Xi_n(\varsigma^{(n)})_j - \Xi_{A,n}(\varsigma^{(n)})_j\right|^2\right) \underset{(b3)}{=} o_{\mathrm{P}}(n^{-1}), \quad \text{and} \quad \mathbb{E}\left(\left|\Xi_{A,n}(\varsigma^{(n)})_j\right|^2\right) \underset{(b4)}{=} O_{\mathrm{P}}(n^{-1}). \tag{A.19}$$

Indeed, $(b3)$ can be shown following the same reasoning of Lemma B1. And $(b4)$ can be proved based on Lemma D3 and Lemma D6.

Second, we prove $\mathcal{R}_a = o_{\mathrm{P}}(1)$ and $\mathcal{R}_c = o_{\mathrm{P}}(n^{-1/4})$. Let $\mathcal{U}_n := \partial\Xi_n(\bar{\varsigma}_n)_{11}$ and we partition the matrix $\partial\bar{\Xi}_n(\bar{\varsigma}_n)$ into four blocks:

$$\partial\Xi_n(\bar{\varsigma}_n) =: \begin{pmatrix} \mathcal{U}_n & \mathcal{Q}_n^\intercal \\ \mathcal{V}_n & \mathcal{W}_n \end{pmatrix}. \tag{A.20}$$

Next, apply block matrix inversion to obtain

$$\mathcal{R}_a = (\mathcal{U}_n - \mathcal{Q}_n^\intercal \mathcal{W}_n^{-1}\mathcal{V}_n)^{-1} - (\partial\bar{\Xi}_n^{\star\star})_{11}^{-1} \quad \text{and} \quad \mathcal{R}_c = \|(\mathcal{U}_n - \mathcal{Q}_n^\intercal \mathcal{W}_n^{-1}\mathcal{V}_n)^{-1}\mathcal{Q}_n^\intercal \mathcal{W}_n^{-1}\|^2.$$

In view of Caucy-Schwarz, the two bounds follow from

$$\|\mathcal{W}_n^{-1}\| = O_{\mathrm{P}}(1), \quad \|\mathcal{V}_n\|^2 \vee \|\mathcal{Q}_n\|^2 = o_{\mathrm{P}}(n^{-1/4}), \quad \text{and} \quad |\mathcal{U}_n - (\partial\bar{\Xi}_n^{\star\star})_{11}| = o_{\mathrm{P}}(1). \tag{A.21}$$

We prove (A.21) in the next step. The current step ends.

Step 6. (Auxiliary: Bounds on $|\mathcal{U}_n - (\partial\bar{\Xi}_n^{\star\star})_{11}|$, $\|\mathcal{V}_n\|$, $\|\mathcal{Q}_n\|$ and $\|\mathcal{W}_n^{-1}\|$) We only show $|\mathcal{U}_n - (\partial\bar{\Xi}_n^{\star\star})_{11}| = o_{\mathrm{P}}(1)$, $\|\mathcal{W}_n^{-1}\| = O_{\mathrm{P}}(1)$, while $\|\mathcal{V}_n\|^2 \vee \|\mathcal{Q}_n\|^2 = O_{\mathrm{P}}(n^{-1/4})$ is simpler.

Let $x \in \mathbb{R}^{q_n+2}$. In view of (A.10) and (A.20), the definition of matrix norm implies

$$|\mathcal{U}_n - (\partial\bar{\Xi}_n^{\star\star})_{11}| \leq \|\partial\Xi_n(\bar{\varsigma}_n) - \partial\bar{\Xi}_n^{\star\star}\|,$$

$$\text{and} \quad \|\mathcal{W}_n^{-1}\| \leq \Big(\inf_{\|x\|=1} |x^\intercal \partial\Xi_n(\bar{\varsigma}_n)x|\Big)^{-1} \leq \Big(\inf_{\|x\|=1} \Big||x^\intercal\partial\bar{\Xi}_n^{\star\star}x| - |x^\intercal(\partial\Xi_n(\bar{\varsigma}_n) - \partial\bar{\Xi}_n^{\star\star})x|\Big|\Big)^{-1}.$$

Hence, it suffices to show

$$\inf_{\|x\|=1} |x^\intercal\partial\bar{\Xi}_n^{\star\star}x| \geq K^{-1} + o_{\mathrm{P}}(1) \quad \text{and} \quad \|\partial\Xi_n(\bar{\varsigma}_n) - \partial\bar{\Xi}_n^{\star\star}\| = o_{\mathrm{P}}(1). \tag{A.22}$$

Lemma B4 proves the second claim in (A.22). Now we prove the first claim. In view of the definition of $\partial\bar{\Xi}_n^{\star\star}$, one can write

$$x^\intercal\partial\bar{\Xi}_n^{\star\star}x = \frac{\sqrt{T}|x_1|^2}{8C_T^{3/2}\zeta^{(n)}} + \frac{|x_2|^2}{2(\iota^4)^{(n)}} + \frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{\left|\sum_{j=3}^{q_n+2} x_j e^{\mathbb{i}j\lambda}\right|^2}{g(\lambda;\theta^{(n)})} d\lambda \geq \|x\|^2.$$

Here the last inequality holds by the upper bound on $g(\lambda;\theta^{(n)})$ and that $\int_{-\pi}^{\pi} e^{\mathbb{i}j\lambda}d\lambda = 0$ for any nonzero integer j. We readily deduce the first claim in (A.22).

Step 7. (Proof of the second claim in Step 2) The proof is a repetition of Proof of Theorem 5.

Step 8. (Proof of the third claim in Step 2) Throughout the rest of the proof, we do not keep the subscript j of $\widehat{q}_{n,j}$ unless necessary.

First, we derive the bound on $\widehat{\mathcal{R}}_n(\widehat{q}_n)$ for $j \in \{1, 2\}$. In view of the definitions of $\widehat{\varsigma}_n(q)$ and $\varsigma_\infty^{(n)}$ given by (A.9), and (A.1), we rewrite $\widehat{\mathcal{R}}_n(\widehat{q}_n)$ as

$$\widehat{\mathcal{R}}_n(\widehat{q}_n) = \|\widehat{\varsigma}_n(\widehat{q}_n) - \varsigma_\infty^{(n)}\|_{(1)}^2 + |\widehat{\sigma}_n^2(\widehat{q}_n) - C_T|^2.$$

By Step 11, $\|\widehat{\varsigma}_n(\widehat{q}_n) - \varsigma_\infty^{(n)}\|_{(1)}^2 = o_{\mathrm{P}}(1)$ leads to $\widehat{\sigma}_n^2(\widehat{q}_n) - C_T = o_{\mathrm{P}}(1)$, hence showing the former is enough. In view of (A.4) and (A.6), this follows from

$$\|\widehat{\varsigma}_n(\widehat{q}_n) - \varsigma_\infty^{(n)}\|_{(1)}^2 \underset{(d1)}{\lesssim} -\frac{2}{n}(\bar{L}_n^\star(\widehat{\varsigma}_n(\widehat{q}_n)) - \bar{L}_n^\star(\varsigma_\infty^{(n)})) + O_{\mathrm{P}}(n^{-1/2}) \underset{(d2)}{=} o_{\mathrm{P}}(1),$$

where $(d1)$ holds by Lemma D1. To see $(d2)$, using $\widehat{q}_{n,j} = \arg\min_q (q(2\delta_{1,j} + \delta_{2,j}\log n) - 2L_n(\widehat{\varsigma}_n(q)))$ and $L_n(\widehat{\varsigma}_n(\widehat{q}_{n,j})) = \max_{\varsigma\in\Pi_n^\varsigma(\widehat{q}_{n,j})} L_n(\varsigma)$, we conclude

$$-\frac{2}{n}(\bar{L}_n^\star(\widehat{\varsigma}_n(\widehat{q}_{n,j})) - \bar{L}_n^\star(\varsigma_\infty^{(n)})) \le \mathcal{R}_{e1} + \mathcal{R}_{e2} + \mathcal{R}_{e3} + \frac{1}{n}(q_n^\star - \widehat{q}_{n,j})(2\delta_{1,j} + \delta_{2,j}\log n),$$

with

$$\mathcal{R}_{e1} := -\frac{2}{n}(\bar{L}_n^\star(\varsigma^{(n)}(q_n^\star)) - \bar{L}_n^\star(\varsigma_\infty^{(n)})), \quad \mathcal{R}_{e2} := \frac{2}{n}(L_n(\widehat{\varsigma}_n(\widehat{q}_n)) - \bar{L}_n^\star(\widehat{\varsigma}_n(\widehat{q}_n))),$$

$$\mathcal{R}_{e3} := -\frac{2}{n}(L_n(\varsigma^{(n)}(q_n^\star)) - \bar{L}_n^\star(\varsigma^{(n)}(q_n^\star))).$$

Recall $q_n^\star$ is defined by (A.3). In view of $q_n^\star - \widehat{q}_n \le q_n^\star = O_{\mathrm{P}}(n^{1/4}(\log n)^{-1})$, we only need to bound $\mathcal{R}_{e1}$, $\mathcal{R}_{e2}$ and $\mathcal{R}_{e3}$. Bound on $\mathcal{R}_{e1}$ is given by

$$\mathcal{R}_{e1} \underset{(d3)}{\lesssim} \|\theta^{(n)}\|_{(q_n^\star)}^2 \underset{(d4)}{=} O_{\mathrm{P}}(n^{-3/4}),$$

where $(d3)$ holds by Step 10 below and $(d4)$ holds by Condition 1. The other two bounds are special cases of

$$\sup_{\varsigma\in\Pi_n^\varsigma(q), q\lesssim n^{1/2}} |L_n(\varsigma) - \bar{L}_n^\star(\varsigma)| = o_{\mathrm{P}}(n), \tag{A.23}$$

which holds by Lemma B5. The reason is that $\varsigma^{(n)}(q), \widehat{\varsigma}_n(q) \in \Pi_n^\varsigma(q)$ due to their definitions (A.9) and (A.5). We readily deduce the bound of $\widehat{\mathcal{R}}_n(\widehat{q}_n)$.

Second, we show the bound on $\mathcal{R}^{(n)}(\widehat{q}_n)$. In view of the definitions of $\varsigma^{(n)}(q)$ and $\varsigma_\infty^{(n)}$ given by (A.9), and (A.1), we rewrite $\mathcal{R}^{(n)}(\widehat{q}_n)$ as

$$\mathcal{R}^{(n)}(\widehat{q}_n) = \|\varsigma^{(n)}(\widehat{q}_n) - \varsigma_\infty^{(n)}\|_{(1)}^2 + |\sigma^{(n)}(\widehat{q}_n)^2 - C_T|^2.$$

The definition of $\sigma^{(n)}(q)^2$, plus Cauchy-Schwarz and Condition 1, yields that $\|\varsigma^{(n)}(\widehat{q}_n) - \varsigma_\infty^{(n)}\|_{(1)}^2 = o_{\mathrm{P}}(1)$ leads to $\sigma^{(n)}(\widehat{q}_n)^2 - C_T = o_{\mathrm{P}}(1)$. Let

$$\mathcal{R}_f := -\frac{2}{n}(\bar{L}_n^\star(\varsigma^{(n)}(\widehat{q}_n)) - \bar{L}_n^\star(\varsigma_\infty^{(n)})). \tag{A.24}$$

We obtain the bound on $\|\varsigma^{(n)}(\widehat{q}_n) - \varsigma_\infty^{(n)}\|_{(1)}^2$ from

$$\|\varsigma^{(n)}(\widehat{q}_n) - \varsigma_\infty^{(n)}\|_{(1)}^2 \underset{(e1)}{\lesssim} \mathcal{R}_f + o_{\mathrm{P}}(1) \quad \text{and} \quad \mathcal{R}_f \underset{(e2)}{\lesssim} \|\theta^{(n)}\|_{(\widehat{q}_n)}^2 \underset{(e3)}{=} o_{\mathrm{P}}(1). \tag{A.25}$$

Lemma D1, plus Cauchy-Schwarz and Condition 1, imply ($e1$). Step 10 proves ($e2$). Using $\widehat{\mathcal{R}}_n(\widehat{q}_n) = o_{\mathrm{P}}(1)$ proved above and observing $\|\theta^{(n)}\|^2_{(\widehat{q}_n)} \le \widehat{\mathcal{R}}_n(\widehat{q}_n)$, we obtain ($e3$).

Step 9. (Proof of the fourth claim in Step 2) In this step, we prove $\sigma^{(n)}(\widehat{q}_n)^2 - C_T = o_{\mathrm{P}}(n^{-1/4})$ and $\widehat{q}_n = o_{\mathrm{P}}(n^{1/4})$. The definition of $\sigma^{(n)}(q)^2$, plus Cauchy-Schwarz and Condition 1, imply

$$\sigma^{(n)}(\widehat{q}_n)^2 - C_T = o_{\mathrm{P}}(n^{1/4}) + O_{\mathrm{P}}((\widehat{q}_n + 1)^{1/2}\|\varsigma^{(n)}(\widehat{q}_n) - \varsigma^{(n)}_\infty\|_{(1)}).$$

We hence only need to show $\widehat{q}_n = o_{\mathrm{P}}(n^{1/4})$ and $\|\varsigma^{(n)}(\widehat{q}_n) - \varsigma^{(n)}_\infty\|^2_{(1)} = O_{\mathrm{P}}(n^{-3/4})$.

Next, using Lemma D1 and Cauchy-Schwarz, we deduce that $\widehat{q}_n = o_{\mathrm{P}}(n^{1/4})$ and $n^{3/4}\|\varsigma^{(n)}(\widehat{q}_n) - \varsigma^{(n)}_\infty\|^2_{(1)} \xrightarrow{\mathrm{P}} \infty$ together lead to $n^{3/4}\mathcal{R}_f \xrightarrow{\mathrm{P}} \infty$. Thereby, it suffices to show $\widehat{q}_n = o_{\mathrm{P}}(n^{1/4})$ and $\mathcal{R}_f = O_{\mathrm{P}}(n^{-3/4})$.

Lemma B6 proves for $j \in \{1, 2\}$,

$$n\mathcal{R}_f \lesssim (q^\star_n - \widehat{q}_{n,j})(\delta_{1,j} + \delta_{2,j}\log n - \delta_{2,j}) + o_{\mathrm{P}}(|q^\star_n - \widehat{q}_{n,j}|(\delta_{1,j} + \delta_{2,j}\log n) + \log n). \tag{A.26}$$

In view of (A.26), plus $q^\star_n = O_{\mathrm{P}}(n^{1/4}(\log n)^{-1})$ from Condition 1, we deduce that $n^{1/4}\widehat{q}_n \to a^2 \in (0, \infty) \cup \{\infty\}$ yields $|\mathcal{R}_f| \ge \frac{1}{K}n^{-3/4}$. We hence can establish the bound on $\widehat{q}_n$ by showing $|\mathcal{R}_f| \ge \frac{1}{K}n^{-3/4}$ cannot occur with nonvanishing probability. This is obvious from ($e2$) in Step 8 and $\|\theta^{(n)}\|^2_{(\widehat{q}_n)} = o_{\mathrm{P}}(n^{-3/4})$ by $n^{1/4}\widehat{q}_n \to a^2$ and Condition 1. On the other hand, $q^\star_n = O_{\mathrm{P}}(n^{1/4}(\log n)^{-1})$ and (A.26) alone give the bound on $\|\varsigma^{(n)}(\widehat{q}_n) - \varsigma^{(n)}_\infty\|^2_{(1)}$.

Step 10. (Auxiliary: Proof of ($d3$) and ($e2$)) In view of the construction of AIC and BIC, both bounds are special cases of the claim that for any $\{q_n\}$ satisfying $n^{1/2}q_n \to 0$,

$$\mathcal{R}_g \coloneqq -\frac{2}{n}(\bar{L}^\star_n(\varsigma^{(n)}(q_n)) - \bar{L}^\star_n(\varsigma^{(n)}_\infty)) \lesssim \|\theta^{(n)}\|^2_{(q_n)}. \tag{A.27}$$

From the definition of $\bar{L}^\star_n$ and $\varsigma^{(n)}(q_n)$, and writting

$$\mathcal{R}_{g1} \coloneqq \log\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{g(\lambda; \theta^{(n)})}{g(\lambda; \theta^{(n)}(q_n))} d\lambda\right) \quad \text{and} \quad \mathcal{R}_{g2} \coloneqq \Delta_n^{1/2}(\widetilde{h}_n(\varsigma^{(n)}(q_n)) - \widetilde{h}_n(\varsigma^{(n)}_\infty)),$$

we have $\mathcal{R}_g = \mathcal{R}_{g1} + \mathcal{R}_{g2}$, and hence it is sufficient to show $\mathcal{R}_{g1} + \mathcal{R}_{g2} \lesssim \|\theta^{(n)}\|^2_{(q_n)}$.

When $n^{1/2}\|\theta^{(n)}\|^2_{(q_n)} \to \infty$, the inequality directly holds by $\mathcal{R}_{g1} \lesssim \|\theta^{(n)}\|^2_{(q_n)}$ due to Lemma D1. Thus, by the standard subsequence argument, it suffices to show the claim

$$\|\theta^{(n)}\|^2_{(q_n)} \lesssim n^{-1/2} \quad \Longrightarrow \quad \mathcal{R}_{g2} \lesssim \|\theta^{(n)}\|^2_{(q_n)}. \tag{A.28}$$

First, under $\|\theta^{(n)}\|^2_{(q_n)} \lesssim n^{-1/2}$, we have

$$\begin{aligned}|\zeta^{(n)}(q_n)^2 - (\zeta^{(n)})^2| &\underset{(f1)}{\lesssim} \|\varsigma^{(n)}(q_n) - \varsigma^{(n)}_\infty\|_{1,(1)} \underset{(f2)}{=} \sum_{j=2}^{[n^{1/4}]} |\varsigma^{(n)}(q_n)_j - \varsigma^{(n)}_{\infty,j}| + \sum_{j=[n^{1/4}]}^{\infty} |\theta_j^{(n)}| \\ &\underset{(f3)}{\lesssim} (n^{1/8} \vee (q_n+1)^{1/2}) \|\varsigma^{(n)}(q_n) - \varsigma^{(n)}\|_{(1)} + o_{\mathrm{P}}(n^{-1/4}) \\ &\underset{(f4)}{\lesssim} (n^{1/8} \vee (q_n+1)^{1/2}) \|\theta^{(n)}\|_{(q_n)} + o_{\mathrm{P}}(n^{-1/4}) \underset{(f5)}{=} o_{\mathrm{P}}(1),\end{aligned}$$

where $(f1)$ and $(f2)$ hold by definition, $(f3)$ holds by Cauchy-Schwarz and Condition 1, $(f4)$ holds by Lemma D1, and $(f5)$ holds by $q_n = o(n^{1/2})$.

Second, under $|\zeta^{(n)}(q_n)^2 - (\zeta^{(n)})^2| = o_{\mathrm{P}}(1)$, we have

$$\mathcal{R}_{g2} = \left(\left(\frac{2(\zeta^2)^{(n)}}{\widetilde{\zeta}^2_n(q_n)} - \frac{(\zeta^4)^{(n)}}{\widetilde{\zeta}^4_n(q_n)} \right)^{1/2} - 1 \right) \frac{C_T^{1/2} \Delta_n^{1/2}}{\zeta^{(n)}} \leq 0 \lesssim \|\theta^{(n)}\|^2_{(q_n)},$$

which holds by $1 = \max_x (2x - x^2)$.

Hence (A.28) indeed holds and we readily deduce (A.27).

Step 11. (Auxiliary: Bound on $|\widehat{\sigma}^2_n(\widehat{q}_n) - C_T|$) In this step, we demonstrate that $\|\widehat{\varsigma}_n(\widehat{q}_n) - \varsigma^{(n)}_\infty\|^2_{(1)} = o_{\mathrm{P}}(1)$ leads to $\widehat{\sigma}^2_n(\widehat{q}_n) - C_T = o_{\mathrm{P}}(1)$.

In view of the definitions of $\widehat{\varsigma}_n(q)$ and the triangle inequality, this follows from three claims:

$$\frac{\partial L_n(\widehat{\varsigma}_n(\widehat{q}_n))}{\partial \varsigma_1} \underset{(g1)}{=} 0, \qquad \sup_{(\sigma^2, \iota^2, \theta) \in \Pi(\widehat{q}_n)} \left| \frac{\partial L_n(\sigma^2, \iota^2, \theta)}{\partial \sigma^2} - \frac{\partial \bar{L}^\star_n(\sigma^2, \iota^2, \theta)}{\partial \sigma^2} \right| \underset{(g2)}{=} o_{\mathrm{P}}(n^{1/2}),$$

and under $\|\widehat{\varsigma}_n(\widehat{q}_n) - \varsigma^{(n)}_\infty\|^2_{(1)} = o_{\mathrm{P}}(1)$,

$$\mathcal{R}_h \coloneqq \left| \frac{1}{n^{3/4}} \frac{\partial \bar{L}^\star_n(\widehat{\varsigma}_n(\widehat{q}_n))}{\partial \varsigma_1} + \frac{\sqrt{T}(\widehat{\sigma}^2_n(\widehat{q}_n) - C_T)}{8\zeta^{(n)} \widehat{\sigma}^3_n(\widehat{q}_n)} \right| \underset{(g3)}{=} o_{\mathrm{P}}(1).$$

Here $(g1)$ holds as it is one of the FOCs and $(g2)$ holds by generalizing Lemma B5. To see $(g3)$, note $\mathcal{R}_h \lesssim |\widehat{\zeta}^2_n(q_n) - (\zeta^2)^{(n)}|$ and apply Cauchy-Schwarz and Condition 1.

A.5 Proof of Corollary 2

In view of Theorem 3 and Proof of Corollary 1 in Belloni, Chernozhukov, and Hansen (2014), it suffices to show

$$|\widehat{\zeta}^2_n(\widehat{q}_{n,j}) - (\zeta^{(n)})^2| = o_{\mathrm{P}}(1), \quad |\widehat{C}_n(4)_T - C(4)_T| = o_{\mathrm{P}}(1),$$

where the convergences hold uniformly over all DGP sequences $\{\mathrm{P}^{(n)}\}$ satisfying $\mathrm{P}^{(n)} \in \mathbb{P}^{(n)}$ for each n.

First, Theorem 16.4.2 in Jacod and Protter (2011) proves the second convergence.

Second, using the definition of $\widehat{\zeta}_n^2$ and Cauchy-Schwarz inequality, and adopting the notation $\widehat{\mathcal{R}}_n(q_n)$ introduced in Step 1 of Section A.4, we can conclude that for $j \in \{1,2\}$ and all $\alpha_n \to \infty$,

$$\begin{aligned} |\widehat{\zeta}_n^2(\widehat{q}_{n,j}) - (\zeta^{(n)})^2| &\lesssim |\widehat{\iota}_n^2(\widehat{q}_{n,j}) - (\iota^{(n)})^2| + \sqrt{\alpha_n}\|\theta_n(\widehat{q}_{n,j}) - \theta^{(n)}\| + \sum_{k=\alpha_n+1}^{\infty} |\theta_n(\widehat{q}_{n,j})_k - \theta_k^{(n)}| \\ &\lesssim \sqrt{\alpha_n \widehat{\mathcal{R}}_n(\widehat{q}_{n,j})} + o_{\mathrm{P}}(1). \end{aligned}$$

Here the last inequality holds by $\theta_n(\widehat{q}_{n,j}), \theta^{(n)} \in \Theta(\infty)$. Step 10 of Section A.4 proves $\widehat{\mathcal{R}}_n(\widehat{q}_{n,j}) = o_{\mathrm{P}}(1)$. We thus obtain the first convergence, and conclude this proof.

A.6 Proof of Theorem 4

Step 1. (Notation) We introduce all the notation. There are three parts.

Part 1. We introduce notation below which will (mostly) be used to analyze the case where noise magnitude is relatively large. Let

$$\Pi_n^{\varsigma}(q) = \left\{\varsigma \in \mathbb{R}^{q+2} : \left(\varsigma_1\Big(n\exp(\varsigma_2)\Big(1 + \sum_{j=3}^{q+2}\varsigma_j\Big)^2\Big)^{1/4}, \exp(\varsigma_2), \varsigma_3, \ldots, \varsigma_{q+2}\right)^{\intercal} \in \Pi_n(q)\right\}.$$

Let $\theta^{(n)}(q) = (\theta_1^{(n)}, \theta_2^{(n)}, \ldots, \theta_q^{(n)})^{\intercal}$, $\gamma^{(n)}(q) = (\gamma_0^{(n)}, \gamma_1^{(n)}, \ldots, \gamma_q^{(n)})^{\intercal}$, and

$$\varsigma^{(n)}(q) = (C_T n^{-1/4}(\zeta^{(n)})^{-1/2}, \log(\iota^{(n)})^2, \theta^{(n)}(q)).$$

For any $\varsigma \in \Pi_n^{\varsigma}(q)$ and any $A_n \in \{L_n, L_{D,n}, \bar{L}_n\}$, we let

$$A_n(\varsigma) = A_n(\sigma^2, \iota^2, \theta),$$

with $\iota^2 = \exp(\varsigma_2), \theta = (\varsigma_3, \ldots, \varsigma_{q+2})^{\intercal}, \zeta^2 = \iota^2(1 + 1^{\intercal}\cdot\theta)^2$ and $\sigma^2 = \varsigma_1(n\zeta^2)^{1/4}$. Moreover, with any $\varsigma \in \Pi_n^{\varsigma}(q)$ and $s \in \{A, D\}$ we associate $\Xi_n(\varsigma), \bar{\Xi}_n(\varsigma), \Xi_{s,n}(\varsigma) \in \mathbb{R}^{q+2}$ and $\partial\Xi_n(\varsigma), \partial\bar{\Xi}_n(\varsigma) \in \mathcal{M}_{q+2}$ such that

$$\begin{aligned} \Xi_n(\varsigma)_j = -\frac{1}{n}\frac{\partial L_n(\varsigma)}{\partial\varsigma_j}, \quad \bar{\Xi}_n(\varsigma_n)_j = -\frac{1}{n}\frac{\partial \bar{L}_n(\varsigma)}{\partial\varsigma_j}, \quad \Xi_{s,n}(\varsigma)_j = -\frac{1}{n}\frac{\partial L_{s,n}(\varsigma)}{\partial\varsigma_j}, \\ \partial\Xi_n(\varsigma)_{ij} = -\frac{1}{n}\frac{\partial^2\Xi_n(\varsigma)}{\partial\varsigma_i\partial\varsigma_j}, \quad \text{and} \quad \partial\bar{\Xi}_n(\varsigma)_{ij} = -\frac{1}{n}\frac{\partial^2\bar{L}_n(\varsigma)}{\partial\varsigma_i\partial\varsigma_j}. \end{aligned} \tag{A.29}$$

Furthermore, let $\partial\bar{\Xi}^{\star} \in \mathcal{M}_{q+2}$ be defined as

$$(\partial\bar{\Xi}^{\star})_{ij} = \delta_{1,i}\delta_{1,j}\frac{\sqrt{T}}{8C_T^{3/2}} + \delta_{2,i}\delta_{2,j}\frac{1}{2} + \mathbb{1}_{\{3\le i,j\le q+2\}}\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{e^{\mathrm{i}(i-j)\lambda}}{g(\lambda;\theta^{\star})}d\lambda.$$

Part 2. We introduce more notation which will (mostly) be used to analyze the case where noise magnitude is relatively small. For any $(\sigma^2, \gamma) \in \Pi_n^{(\sigma^2,\gamma)}(q)$, let $\beta_n(\sigma^2, \gamma) \in \mathbb{R}^{q+2}$ be defined as

$$\beta_n(\sigma^2, \gamma) = (\sigma^2, \Delta_n^{-1}\gamma_0, \Delta_n^{-1}\gamma_1, \dots, \Delta_n^{-1}\gamma_q)^\intercal. \tag{A.30}$$

Let $\widehat{\beta}_n(q), \beta^{(n)}(q) \in \mathbb{R}^{q+2}$ be defined as

$$\widehat{\beta}_n(q) = \beta_n(\widehat{\sigma}_n^2(q), \widehat{\gamma}_n(q)), \quad \beta^{(n)}(q) = \beta_n(C_T, \gamma^{(n)}(q)), \tag{A.31}$$

where $\gamma^{(n)}(q)$ is the $(q+1)$ vector of up-to-qth-order autocovariances of U at stage n. Let

$$\Pi_n^{\beta}(q) = \left\{\beta = (\beta_1, \beta_2, \dots, \beta_{q+2})^\intercal \in \mathbb{R}^{q+2} : \beta = \beta_n(\sigma^2, \gamma) \text{ with } (\sigma^2, \gamma) \in \Pi_n^{(\sigma^2,\gamma)}(q)\right\}.$$

For any $\beta \in \Pi_n^{\beta}(q)$, and any $A_n \in \{L_n, L_{D,n}, \bar{L}_n\}$, we let

$$A_n(\beta) = A_n(\sigma^2, \gamma),$$

with $\sigma^2 = \beta_1$, $\gamma_j = \Delta_n \beta_{j+2}$, $0 \le j \le q$. Furthermore, with any $\beta \in \Pi_n^{\beta}(q)$ and $s \in \{A, D\}$ we define $\Xi_n(\beta), \bar{\Xi}_n(\beta), \Xi_{s,n}(\beta) \in \mathbb{R}^{q+2}$ and $\partial\Xi_n(\beta) \in \mathcal{M}_{q+2}$ such that

$$\begin{gathered}\Xi_n(\beta)_j = -\frac{1}{n}\frac{\partial L_n(\beta)}{\partial \beta_j}, \quad \bar{\Xi}_n(\beta)_j = -\frac{1}{n}\frac{\partial \bar{L}_n(\beta)}{\partial \beta_j}, \quad \Xi_{s,n}(\beta)_j = -\frac{1}{n}\frac{\partial L_{s,n}(\beta)}{\partial \beta_j}, \\ \partial\Xi_n(\beta)_{ij} = -\frac{1}{n}\frac{\partial^2 L_n(\beta)}{\partial \beta_i \partial \beta_j}, \quad \text{and} \quad \partial\bar{\Xi}_n(\beta)_{ij} = -\frac{1}{n}\frac{\partial^2 \bar{L}_n(\beta)}{\partial \beta_i \partial \beta_j}.\end{gathered} \tag{A.32}$$

Part 3. Finally, we introduce the notation below which will be used to show the consistency.

According to Theorem 4.1.1, Proposition 4.5.3, Proposition 3.2.1 and Theorem 3.1.2 in Brockwell and Davis (1991), for all q and n given and all $(\sigma^2, \gamma) \in \Pi_n^{(\sigma^2,\gamma)}(q)$, $f(\lambda; \sigma^2, \gamma, \Delta_n)$ is the spectral density function of an invertible moving-average process of order at most q. Therefore, for all $(\sigma^2, \gamma) \in \Pi_n^{(\sigma^2,\gamma)}(q)$, there exist unique $\chi^2(\sigma^2, \gamma, \Delta_n) \in \mathbb{R}_+$ and $\phi(\sigma^2, \gamma, \Delta_n) \in \mathbb{R}^{q+1}$ such that

$$\chi^2(\sigma^2, \gamma, \Delta_n) g(\lambda; \phi(\sigma^2, \gamma, \Delta_n)) = f(\lambda; \sigma^2, \gamma, \Delta_n), \quad \inf_{x \in \mathbb{C}, |x| \le 1} \left| 1 + \sum_{j=1}^{\infty} \phi(\sigma^2, \gamma, \Delta_n)_j x^j \right| > 0,$$

with $f(\lambda; \sigma^2, \gamma, \Delta_n)$ defined by (4.16). We set

$$\Pi_n^{(\chi^2,\phi)}(q) = \left\{(\chi^2, \phi) : \chi^2 = \chi^2(\sigma^2, \gamma, \Delta_n), \phi = \phi(\sigma^2, \gamma, \Delta_n), \text{ for some } (\sigma^2, \gamma) \in \Pi_n^{(\sigma^2,\gamma)}(q)\right\}.$$

We also set

$$\widehat{\chi}_n^2(q) = \chi^2(\widehat{\sigma}_n^2(q), \widehat{\gamma}_n(q), \Delta_n), \quad \widehat{\phi}_n(q) = \phi(\widehat{\sigma}_n^2(q), \widehat{\gamma}_n(q), \Delta_n),$$

$$\chi^{(n)}(q)^2 = \chi^2(C_T, \gamma^{(n)}(q), \Delta_n), \quad \phi^{(n)}(q) = \phi(C_T, \gamma^{(n)}(q), \Delta_n).$$

Finally, for all $(\chi^2, \phi) \in \Pi_n^{(\chi^2,\phi)}(q)$, we define $\Sigma_n(\chi^2, \phi) \in \mathcal{M}_n$ by

$$\Sigma_n(\chi^2, \phi)_{ij} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \chi^2 g(\lambda; \phi) e^{\mathbb{i}(i-j)\lambda} d\lambda, \quad \forall 1 \le i, j \le n,$$

and set

$$L_n(\chi^2, \phi) = \mathcal{L}(\Sigma_n(\chi^2, \phi), Y_n Y_n^\intercal), \quad \bar{L}_n^\star(\chi^2, \phi, q) = -\frac{n}{2} \log \chi^2 - \frac{n}{2} \frac{\chi^{(n)}(q)^2}{2\pi\chi^2} \int_{-\pi}^{\pi} \frac{g(\lambda; \phi^{(n)}(q))}{g(\lambda; \phi)} d\lambda.$$

Step 2. (Main) Lemma C5 proves

$$\text{AVAR}\left(C_T, 0, 1, C(4)_T - C_T^2, \text{cum}_4[\varepsilon]\right) = \frac{1}{T}(4q + 6)C(4)_T.$$

Hence, we only need to prove scenarios (i) and (iii) of the theorem. Step 4 shows the latter.

Now we prove scenario (i). In view of the definition of $\varsigma^{(n)}(q)$ in Part 1 of Step 1 above, we write $\widehat{\zeta}_n^2(q) = \sum_{|j|<q} \widehat{\gamma}_n(q)_j$, and

$$n^{1/4}(\zeta^{(n)})^{-1/2}(\widehat{\sigma}_n^2(q) - C_T) = n^{1/2}(\widehat{\sigma}_n^2(q)\underbrace{(n\widehat{\zeta}_n^2(q))^{-1/4} - \varsigma^{(n)}(q)_1)}_{=:\mathcal{R}_a} - n^{1/4}\widehat{\sigma}_n^2(q)\underbrace{((\widehat{\zeta}_n^2(q))^{-1/4} - (\zeta^{(n)})^{-1/2})}_{=:\mathcal{R}_b}.$$

Using $\widehat{\sigma}_n^2(q) = O_{\mathrm{P}}(1)$ and $\|\widehat{\gamma}_n - \gamma^{(n)}\| = (\iota^{(n)})^2 o_{\mathrm{P}}(1)$ from (A.37) below, we obtain $\mathcal{R}_b = o_{\mathrm{P}}(1)$. Meanwhile, Step 3 below proves

$$\mathcal{R}_a \stackrel{\mathcal{L}-s}{\to} \mathcal{MN}\left(0, \frac{5C_T^{-1/2}C(4)_T + 3C_T^{3/2}}{\sqrt{T}}\right), \tag{A.33}$$

which concludes this step.

Step 3. (Proof of (A.33)) We rely on the notation introduced in Part 1 of Step 1. From this step to Step 5 we omit the dependence of $\widehat{\varsigma}_n$ (defined below), $\varsigma^{(n)}$, $\widehat{\sigma}_n^2$, $\widehat{\gamma}_n$, and $\gamma^{(n)}$ on q whenever possible.

First, writing $f(\lambda; \gamma) := \sum_{j=-\infty}^{\infty} \gamma_{|j|} e^{\mathbb{i}j\lambda}$, plus the Cauchy-Schwarz inequality and that q is fixed, we have

$$\begin{aligned}\inf_\lambda f(\lambda; \widehat{\gamma}_n) \ge \inf_\lambda f(\lambda; \gamma^{(n)}) - K\|\widehat{\gamma}_n - \gamma^{(n)}\| &\underset{(a1)}{=} \inf_\lambda f(\lambda; \gamma^{(n)}) + (\iota^{(n)})^2 o_{\mathrm{P}}(1) \\ &\underset{(a2)}{\ge} K^{-1}(\iota^{(n)})^2 + (\iota^{(n)})^2 o_{\mathrm{P}}(1) > 0.\end{aligned} \tag{A.34}$$

Here $(a1)$ holds by $\|\widehat{\gamma}_n - \gamma^{(n)}\| = (\iota^{(n)})^2 o_{\mathrm{P}}(1)$ from (A.37) below. $(a2)$ is obvious. Hence, according to Theorem 4.1.1, Proposition 4.5.3, Proposition 3.2.1 and Theorem 3.1.2 in Brockwell and Davis

(1991), there exists unique $(\widehat{\iota}_n^2(q), \widehat{\theta}_n(q))$ such that

$$\widehat{\iota}_n^2(q) g(\lambda; \widehat{\theta}_n(q)) = f(\lambda; \widehat{\gamma}_n), \quad \forall \lambda \in [-\pi, \pi]; \quad \widehat{\theta}_n(q) \in \Theta_n(q).$$

Moreover, we define

$$\widehat{\varsigma}_n(q) = (\widehat{\sigma}_n^2(q)(n\widehat{\zeta}_n^2(q))^{-1/4}, \log(\widehat{\iota}_n^2(q)), \widehat{\theta}_n(q)) \in \Pi_n^{\varsigma}(q). \tag{A.35}$$

Second, we derive the equations which $\widehat{\varsigma}_n$ satisfies. We follow the same strategy used by Step 3 of Section A.4. In view of the definitions of $\widehat{\varsigma}_n$ and $\Xi_n(\varsigma)$, we have

$$\Xi_n(\widehat{\varsigma}_n) = 0_{q+2}. \tag{A.36}$$

Same as in Section A.4, this claim comes from two facts. First, $(\widehat{\sigma}_n^2(q), \widehat{\iota}_n^2(q), \widehat{\theta}_n(q))$ maximizes the quasi-log likelihood over $\Pi^{(\sigma,\gamma^2)}(q)$ by construction. Second, $(\widehat{\sigma}_n^2(q), \widehat{\iota}_n^2(q), \widehat{\theta}_n(q))$ is an interior point (with probability approaching one). The second fact holds by

$$\|\widehat{\gamma}_n - \gamma^{(n)}\| = (\iota^{(n)})^2 o_{\mathrm{P}}(1) \quad \text{and} \quad |\widehat{\sigma}_n^2 - C_T| = o_{\mathrm{P}}(1), \tag{A.37}$$

proved in Step 6. Applying the mean value theorem to the FOCs (A.36), we conclude for all $1 \le i \le q_n + 2$,

$$\Xi_n(\varsigma^{(n)})_i + \sum_{j=1}^{q+2} \partial \Xi_n(\bar{\varsigma}_n(i))_{ij} (\widehat{\varsigma}_n - \varsigma^{(n)})_j = 0. \tag{A.38}$$

Here $\bar{\varsigma}_n(i) = \alpha_n(i)\widehat{\varsigma}_n + (1 - \alpha_n(i))\varsigma^{(n)}$ with $\alpha_n(i) \in [0, 1]$. We introduce some simplifying notation. Let $\bar{\varsigma}_n = (\bar{\varsigma}_n(1), \bar{\varsigma}_n(2), \ldots, \bar{\varsigma}_n(q+2))$ and let $\partial \Xi_n(\bar{\varsigma}_n) \in \mathcal{M}_{q+2}$ be defined by $\partial \Xi_n(\bar{\varsigma}_n)_{ij} = \partial \Xi_n(\bar{\varsigma}_n(i))_{ij}$. In view of them, (A.38) can be rewritten as a matrix equation:

$$\widehat{\varsigma}_n - \varsigma^{(n)} = -\partial \Xi_n(\bar{\varsigma}_n)^{-1} \Xi_n(\varsigma^{(n)}). \tag{A.39}$$

Third, we prove (A.33) based on (A.35) and (A.39). Note $\mathcal{R}_a = n^{1/2}(\widehat{\varsigma}_n - \varsigma^{(n)})_1$ from (A.35). Moreover, Lemma C1 proves that under (A.37),

$$\|(\partial \bar{\Xi}^\star)^{-1} - \partial \Xi_n(\bar{\varsigma}_n)^{-1}\| = o_{\mathrm{P}}(1). \tag{A.40}$$

Hence, using the definition of $\partial \bar{\Xi}^\star$, plus that q is fixed and Chebyshev's inequality, (A.33) follows from

$$n^{1/2} \Xi_n(\varsigma^{(n)})_1 \xrightarrow{\mathcal{L}-s} \mathcal{MN}\left(0, \frac{5\sqrt{T}}{64} C(4)_T C_T^{-7/2} + \frac{3\sqrt{T}}{64} C_T^{-3/2}\right), \quad \text{and} \quad \mathbb{E}(n\|\Xi_n(\varsigma^{(n)})\|^2) = O_{\mathrm{P}}(1), \tag{A.41}$$

which hold by Step 5.

Step 4. (Proof of scenario (iii)) We conduct the proof in this step using the parameters $\widehat{\beta}_n(q)$ and $\beta^{(n)}(q)$, defined by (A.31) and (A.30) in Part 2 of Step 1. Also in this step we omit their arguments q.

In view of the definition of the function AVAR, the claim in scenario (iii) follows immediately from

$$n^{1/2}(\widehat{\beta}_n - \beta^{(n)}) \xrightarrow{\mathcal{L}-s} \mathcal{MN}\left(0, W(C_T, \gamma^\star, 1)^{-1}\widetilde{W}\left(C_T, \gamma^\star, 1, C_T - C_T^2, \mathrm{cum}_4[\varepsilon]\right) W(C_T, \gamma^\star, 1)^{-1}\right). \tag{A.42}$$

Next, Step 6 below proves under $\Delta_n^{-1}(\iota^{(n)})^2 \to a^2 \in [0, \infty)$,

$$\|\widehat{\beta}_n - \beta^{(n)}\| = o_{\mathrm{P}}(1). \tag{A.43}$$

Hence, in view of the definition of $\Xi_n(\beta)$, we can first obtain the FOCs and then apply the mean value theorem in the same way as we do in Step 3 above. Indeed, we derive

$$\widehat{\beta}_n - \beta^{(n)} = -\partial\Xi_n(\bar{\beta}_n)^{-1}\Xi_n(\beta^{(n)}). \tag{A.44}$$

Here $\partial\Xi_n(\bar{\beta}_n) \in \mathcal{M}_{q+2}$ is defined by $\partial\Xi_n(\bar{\beta}_n)_{ij} = \partial\Xi_n(\bar{\beta}_n(i))_{ij}$, where $\bar{\beta}_n(i) = \alpha_n(i)\widehat{\beta}_n + (1 - \alpha_n(i))\beta^{(n)}$ with $\alpha_n(i) \in [0, 1]$ and $\bar{\beta}_n = (\bar{\beta}_n(1), \bar{\beta}_n(2), \ldots, \bar{\beta}_n(q+2))$.

Moreover, Lemma C2 proves that under (A.43),

$$\|\partial\Xi_n(\bar{\beta}_n)^{-1} - 2W(C_T, \gamma^\star, 1)^{-1}\| = o_{\mathrm{P}}(1), \tag{A.45}$$

while Step 5 below shows that

$$n^{1/2}\Xi_n(\beta^{(n)}) \xrightarrow{\mathcal{L}-s} \mathcal{MN}\left(0, \frac{1}{4}\widetilde{W}\left(C_T, \gamma^\star, 1, C_T - C_T^2, \mathrm{cum}_4[\varepsilon]\right)\right). \tag{A.46}$$

Using (A.44), (A.45) and (A.46), we readily deduce (A.42) and thus scenario (iii) of the theorem.

Step 5. (Auxiliary: Properties of scores) In this step we show (A.41) and (A.46) following the reasoning of Step 4 of Section A.4.

In view of the triangle inequality, the first claim in (A.41) comes from

$$|\Xi_n(\varsigma^{(n)})_1 - \Xi_{A,n}(\varsigma^{(n)})_1| \underset{(b1)}{=} o_{\mathrm{P}}(n^{-1/2}), \quad |\Xi_{A,n}(\varsigma^{(n)})_1 - \Xi_{D,n}(\varsigma^{(n)})_1| \underset{(b2)}{=} o_{\mathrm{P}}(n^{-1/2}),$$

$$\text{and} \quad n^{1/2}\Xi_{D,n}(\varsigma^{(n)})_1 \underset{(b3)}{\xrightarrow{\mathcal{L}-s}} \mathcal{MN}\left(0, \frac{5\sqrt{T}}{64}C(4)_T C_T^{-7/2} + \frac{3\sqrt{T}}{64}C_T^{-3/2}\right).$$

All three results are proved in Lemma C3. Note $\Xi_{A,n}(\varsigma)$ and $\Xi_{D,n}(\varsigma)$ are introduced in Part 1 of Step 1, making use of Part 3 of Section A.1. See Step 4 of Section A.4 for the consideration behind these definitions and $\Xi_{A,n}(\beta)$ and $\Xi_{D,n}(\beta)$) involved below.

The second claim in (A.41) follows from the same analysis.

With $\Xi_{A,n}(\beta)$ and $\Xi_{A,n}(\beta)$ defined in Part 2 of Step 1, Lemma C3 also proves

$$\|\Xi_n(\beta^{(n)}) - \Xi_{A,n}(\beta^{(n)})\| \underset{(b4)}{=} o_{\mathrm{P}}(n^{-1/2}), \quad \|\Xi_{A,n}(\beta^{(n)}) - \Xi_{D,n}(\beta^{(n)})\| \underset{(b5)}{=} o_{\mathrm{P}}(n^{-1/2}),$$

$$\text{and} \quad n^{1/2}\Xi_{D,n}(\beta^{(n)}) \underset{(b6)}{\overset{\mathcal{L}-s}{\to}} \mathcal{MN}\left(0, \frac{1}{4}\widetilde{W}\left(C_T, \gamma^\star, 1, C(4)_T - C_T^2, \mathrm{cum}_4[\varepsilon]\right)\right).$$

Again combined with the triangle inequality, they imply (A.46).

Step 6. (Consistency) The goal of this step is to prove (A.41) under $\Delta_n^{-1}(\iota^{(n)})^2 \to \infty$ and (A.46) under $\Delta_n^{-1}(\iota^{(n)})^2 \to a^2 < \infty$. We only demonstrate the latter, while showing the former is similar.

We rely on the notation introduced in Part 3 of Step 1 above. In view of the relevant definitions, we only need to show under $\Delta_n^{-1}(\iota^{(n)})^2 \to a^2 < \infty$,

$$\Delta_n^{-1}|\widehat{\chi}_n^2(q) - \chi^{(n)}(q)^2| + \|\widehat{\phi}_n(q) - \phi^{(n)}(q)\| = o_{\mathrm{P}}(1). \tag{A.47}$$

Next, we state two facts. Lemma C4 proves

$$\sup_{(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(q)} \left|\frac{1}{n}L_n(\chi^2, \phi) - \frac{1}{n}\bar{L}_n^\star(\chi^2, \phi, q)\right| = o_{\mathrm{P}}(1). \tag{A.48}$$

And by slightly generalizing Lemma D1, one obtains

$$-\frac{1}{n}\bar{L}_n^\star(\chi^2, \phi, q) + \frac{1}{n}\bar{L}_n^\star(\chi^{(n)}(q)^2, \phi^{(n)}(q), q) \underset{(c1)}{=} o_{\mathrm{P}}(1) \quad \Longrightarrow \quad \left|\frac{\chi^2}{\chi^{(n)}(q)^2} - 1\right|^2 + \|\phi - \phi^{(n)}(q)\|^2 \underset{(c2)}{=} o_{\mathrm{P}}(1).$$

Now we prove (A.47) by contradiction. Assume it does not hold, then one deduce $(c2)$ does not hold as $\chi^{(n)}(q)^2 = O_{\mathrm{P}}(\Delta_n)$ by construction. As a result, $(c1)$ does not hold for $(\chi^2, \phi) = (\widehat{\chi}_n^2(q), \widehat{\phi}_n(q))$. One can verify, however, $\bar{L}_n^\star(\cdot, \cdot, q)$ is maximized at $(\chi^{(n)}(q)^2, \phi^{(n)}(q))$ over $\Pi_n^{(\chi^2,\phi)}(q)$, which means the LHS of $(c1)$ can only be non-negative. This indicates that, with a nonvanishing probability, $-\frac{1}{n}\bar{L}_n^\star(\widehat{\chi}_n^2(q), \widehat{\phi}_n(q), q) + \frac{1}{n}\bar{L}_n^\star(\chi^{(n)}(q)^2, \phi^{(n)}(q), q) \geq \frac{1}{K}$. In view of (A.48), we deduce that, with a nonvanishing probability, $-\frac{1}{n}L_n(\widehat{\chi}_n^2(q), \widehat{\phi}_n(q)) + \frac{1}{n}L_n(\chi^{(n)}(q)^2, \phi^{(n)}(q)) \geq \frac{1}{K}$. This is in contradiction with the fact that $(\widehat{\chi}_n^2(q), \widehat{\phi}_n(q))$ maximizes L_n over $\Pi_n^{(\chi^2,\phi)}(q)$ and we obtain (A.47).

A.7 Proof of Proposition 2

Step 1. (Main proof) We only prove the result for $j = 1$. The case of $j = 2$ can be proved in the same way. For each $b \geq 0$ and each n, we can always find $\mathrm{P}_b^{(n)} \in \mathbb{P}^{(n)}$ under which $((\iota^{(n)})^2, \theta^{(n)}) =$

$\left(bC_T\Delta_n n^{-1/2}, \frac{1}{2}\right)$. Hence, it suffices to show

$$\liminf_{n\to\infty} \inf_{\widehat{G}_{n,1}(x)} \sup_{b\geq 0} \mathrm{P}_b^{(n)}\left(|\widehat{G}_{n,1}(x) - G_{n,1}(x)| > \frac{1}{K}\right) > 0. \tag{A.49}$$

In view of Lemma 3.1 in Leeb and Pötscher (2006), (A.49) follows from three claims. First, for each $x \in \mathbb{R}$ and under $\mathrm{P}_b^{(n)}$, $G_{n,1}(x)$ converges to $G_{\infty,1}(x,b)$ as $n \to \infty$. Second, $G_{\infty,1}(x,b)$ as a function of b is nonconstant for all $x \in \mathbb{R}$. Third, the sequence $\mathrm{P}_b^{(n)}$ is contiguous with respect to the sequence $\mathrm{P}_0^{(n)}$ for every $b \geq 0$. We prove the first two in Step 2, while the last one is proved in Step 6.

Step 2. (Proof of the first two claims in Step 1) We explicitly construct the limit of $G_{n,1}(x,b)$. Let $\mathcal{U}$ and $\bar{\mathcal{U}}$ be two mutually independent standard Gaussian random variables. Let

$$\begin{aligned} G_{\infty,1}(x,b) = \mathrm{P}\left(\sqrt{6}\mathcal{U} + \frac{3b}{4} \leq \frac{x}{C_T}\right)\mathrm{P}\left(\left|\overline{\mathcal{U}} - \frac{b}{2}\right| < \sqrt{2 - \frac{b^2}{32}}\right) \\ + \mathrm{P}\left(\sqrt{6}\mathcal{U} + 2\overline{\mathcal{U}} \leq \frac{x}{C_T}\right)\mathrm{P}\left(\left|\mathcal{U} - \frac{b}{2}\right| \geq \sqrt{2 - \frac{b^2}{32}}\right). \end{aligned}$$

One can easily verify that $G_{\infty,1}(x,b)$ as a function of b is nonconstant. Hence, it suffices to show $G_{n,1}(x)$ converges to $G_{\infty,1}(x,b)$ under $\mathrm{P}_b^{(n)}$.

To do this, we need more notation. Let

$$\phi^{(n)}(q) = \arg\min_{\phi\in\Theta(q+1)} \int_{-\pi}^{\pi} \frac{f(\lambda; C_T, \gamma^{(n)}(1), \Delta_n)}{g(\lambda;\phi)} d\lambda, \quad \chi^{(n)}(q)^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f(\lambda; C_T, \gamma^{(n)}(1), \Delta_n)}{g(\lambda; \phi^{(n)}(q))} d\lambda,$$

where the function f is defined by (4.16). Moreover, let $\beta^{(n)}(q), \widehat{\beta}_n(q) \in \mathbb{R}^{q+2}$ be defined by

$$\begin{aligned} \beta^{(n)}(q)_j &= \frac{\chi^{(n)}(q)^2}{2\pi\Delta_n}\int_{-\pi}^{\pi} g(\lambda;\phi^{(n)}(q))e^{\mathbb{i}(j-1)\lambda} d\lambda, \\ \widehat{\beta}_n(q)_j &= \frac{1}{2\pi\Delta_n}\int_{-\pi}^{\pi} f(\lambda; \widehat{\sigma}_n^2(q), \widehat{\gamma}_n(q), \Delta_n)e^{\mathbb{i}(j-1)\lambda} d\lambda, \end{aligned}$$

where $j \in \{1, 2, \ldots, q+2\}$.

First, we prove in Step 3 below

$$\begin{aligned} \widehat{\sigma}_n^2(0) - C_T &= (\widehat{\beta}_n(0)_1 - \beta^{(n)}(0)_1) + 2(\widehat{\beta}_n(0)_2 - \beta^{(n)}(0)_2) + \frac{3b}{4}C_T n^{-1/2} + O_{\mathrm{P}}(n^{-1}), \\ \widehat{\sigma}_n^2(1) - C_T &= (\widehat{\beta}_n(1)_1 - \beta^{(n)}(1)_1) + 2(\widehat{\beta}_n(1)_2 - 2\beta^{(n)}(1)_2) + 2(\widehat{\beta}_n(1)_3 - \beta^{(n)}(1)_3) \\ &\quad + O_{\mathrm{P}}(n^{-1}). \end{aligned} \tag{A.50}$$

Second, Step 4 proves that $\widehat{q}_{n,1} \wedge 1 = 0$ if and only if

$$2 - \frac{1}{32}b^2 - \left(-\frac{b}{2} + \frac{n^{1/2}}{C_T}(\widehat{\beta}_n(1)_3 - \beta^{(n)}(1)_3)\right)^2 \underset{(b4)}{\geq} 0. \tag{A.51}$$

Third, Step 5 shows

$$n^{1/2}\begin{pmatrix} \widehat{\beta}_n(0) - \beta^{(n)}(0) \\ \widehat{\beta}_n(1) - \beta^{(n)}(1) \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, 2C_T^2\left(\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & & \\ & \frac{1}{2} & \\ & & \frac{1}{2} \end{pmatrix}\right)\right). \tag{A.52}$$

That $G_{n,1}(x)$ converges to $G_{\infty,1}(x,b)$ under $\mathrm{P}_b^{(n)}$ readily follows. Indeed, (A.51) indicates that asymptotically the selected order is determined by the realization of $n^{1/2}(\widehat{\beta}_n(1)_3 - \beta^{(n)}(1)_3)$. Meanwhile, from (A.52) we observe that $n^{1/2}(\widehat{\beta}_n(1)_3 - \beta^{(n)}(1)_3)$ is asymptotically independent of $n^{1/2}(\widehat{\beta}_n(q)_j - \beta^{(n)}(q)_j)$ for every $(q,j) \in \{0,1\} \times \{1,2\}$. Moreover, (A.52) implies that $n^{1/2}(\widehat{\beta}_n(0)_j - \beta^{(n)}(0)_j)$ and $n^{1/2}(\widehat{\beta}_n(1)_j - \beta^{(n)}(1)_j)$ are asymptotically perfectly correlated for both $j \in \{1,2\}$. This implication, plus (A.50), means that the first two terms in $\widehat{\sigma}_n^2(0) - C_T$ and $\widehat{\sigma}_n^2(1) - C_T$ are asymptotically equivalent.

Step 3. (Auxiliary: Proof of (A.50)) The two results follow from that $\widehat{\sigma}_n^2(q) = \sum_{j=1}^{\infty}(2-\delta_{j,1})\widehat{\beta}_n(q)$ and that

$$\left\|\beta^{(n)}(0) - C_T\left(1 + \frac{5b}{4\sqrt{n}}, -\frac{b}{4\sqrt{n}}\right)\right\| = O_{\mathrm{P}}(n^{-1}), \tag{A.53}$$

$$\left\|\beta^{(n)}(1) - C_T\left(1 + \frac{3b}{2\sqrt{n}}, -\frac{b}{4\sqrt{n}}, -\frac{b}{2\sqrt{n}}\right)\right\| = O_{\mathrm{P}}(n^{-1}). \tag{A.54}$$

Both can be verified using the definitions.

Step 4. (Auxiliary: Proof of (A.51)) We start by introducing some notation. Let

$$L_n(\beta) = -\frac{1}{2}\log\det(\Sigma_n(\beta)) - \frac{1}{2}\mathrm{tr}(\Sigma_n(\beta)^{-1}Y_nY_n^\intercal),$$

where the $n_T \times n_T$ matrix $\Sigma_n(\beta)$ is defined by $\Sigma_n(\beta)_{i,j} = \beta_{|j-i|+1}\Delta_n$. Let

$$\Pi_n^\beta(q) = \{\beta \in \mathbb{R}^{q+2} : \Sigma_n(\beta) = \Sigma_n(\sigma^2,\gamma); (\sigma^2,\gamma) \in \Pi_n^{(\sigma^2,\gamma)}(q)\}.$$

In view of the the definition of $\mathrm{AIC}_n(q)$, plus $\widehat{\beta}_n(1) = \arg\max_{\beta\in\Pi_n^\beta(1)} L_n(\beta)$ by construction, we use the mean value theorem to conclude

$$\mathrm{AIC}_n(1) - \mathrm{AIC}_n(0) = 2 + \sum_{i,j}\frac{\partial^2 L_n(\widehat{\beta}_n(1))}{\partial\beta_i\partial\beta_j}(\widehat{\beta}_n(0)_i - \widehat{\beta}_n(1)_i)(\widehat{\beta}_n(0)_j - \widehat{\beta}_n(1)_j) + o_{\mathrm{P}}(n\|\widehat{\beta}_n(0) - \widehat{\beta}_n(1)\|^2).$$

On the other hand, using $\widehat{\beta}_n(1)_j = C_T\delta_{j,1} + o_{\mathrm{P}}(1)$ from (A.54) and (A.52), we deduce

$$-\frac{2}{n}\frac{\partial^2 L_n(\widehat{\beta}_n(1))}{\partial\beta_i\partial\beta_j} = \frac{1}{C_T^2}\delta_{i,j}(2-\delta_{j,1}) + o_{\mathrm{P}}(1). \tag{A.55}$$

Finally, (A.52), (A.53) and (A.54) imply

$$(\widehat{\beta}_n(1)_1 - \widehat{\beta}_n(0)_1)^2 = \frac{1}{16}C_T^2 b^2 n^{-1} + o_{\mathrm{P}}(n^{-1}), \quad (\widehat{\beta}_n(1)_2 - \widehat{\beta}_n(0)_2)^2 = o_{\mathrm{P}}(n^{-1}).$$

Using the last row of (A.52), plus the definition of $\widehat{q}_{n,1}$, we readily obtain (A.51).

Step 5. (Proof of (A.52)) One can show (A.52) by following the proofs of scenarios (ii) and (iii) of Theorem 4 in Section A.6.

Step 6. (Proof of the last claim in Step 1) In view of Le Cam's first lemma (see, e.g., Lemma 6.4 in van der Vaart (2000)), and using

$$\log(d\mathrm{P}_0^{(n)}/d\mathrm{P}_b^{(n)}) = L_n\left(C_T, 0, \frac{1}{2}\right) - L_n\left(C_T, bC_T\Delta_n n^{-1/2}, \frac{1}{2}\right) =: \mathcal{U}_n,$$

it suffices to show that $\exp(\mathcal{U}_n)$ converges in distribution under $\mathrm{P}_b^{(n)}$ to a random variable which is almost surely positive.

Introduce short-hand notation $\beta^{(n,b)} = C_T\left(1+\frac{3b}{2\sqrt{n}}, -\frac{b}{4\sqrt{n}}, -\frac{b}{2\sqrt{n}}\right)$. It follows

$$\begin{aligned} L_n\left(C_T, bC_T\Delta_n n^{-1/2}, \frac{1}{2}\right) &= L_n(\widehat{\beta}_n(1)) + \frac{1}{2}(\beta^{(n,b)} - \widehat{\beta}_n(1))^\intercal \frac{\partial^2 L_n(\widehat{\beta}_n(1))}{\partial\beta\partial\beta}(\beta^{(n,b)} - \widehat{\beta}_n(1)) + o_{\mathrm{P}}(1), \\ L_n\left(C_T, 0, \frac{1}{2}\right) &= L_n(\widehat{\beta}_n(1)) + \frac{1}{2}(\beta^{(n,0)} - \widehat{\beta}_n(1))^\intercal \frac{\partial^2 L_n(\widehat{\beta}_n(1))}{\partial\beta\partial\beta}(\beta^{(n,0)} - \widehat{\beta}_n(1)) + o_{\mathrm{P}}(1), \end{aligned}$$

which hold by $L_n\left(C_T, bC_T\Delta_n n^{-1/2}, \frac{1}{2}\right) = L_n(\beta^{(n,b)}) + o_{\mathrm{P}}(1)$ from the construction of $L_n(\beta)$ and (A.54), the mean value theorem, and (A.52).

Making use of (A.55) and (A.52), it follows that under $\mathrm{P}_b^{(n)}$,

$$\mathcal{U}_n \xrightarrow{\mathcal{L}} \mathcal{N}\left(-\frac{23b^2}{32}, \frac{23b^2}{16}\right),$$

which concludes the proof.

A.8 Proof of Theorem 5

We introduce a set of parameter sequences:

$$\Gamma_0(\{n\},\{\alpha_n\}) = \left\{\{(\iota_n^2,\theta_n) : n\geq 1\} : \theta_n \in \Theta(\infty), \Delta_n^{1/2}\sum_{j\geq\bar{q}}|(\theta_n)_j| < \alpha_n\right\},$$

where we recall that $\Theta(q)$ is defined as

$$\Theta(q) = \left\{\theta : \frac{1}{K} < \inf_{\lambda} g(\lambda, \theta) \leq \sup_{\lambda} g(\lambda, \theta) < K; \theta_j = 0, \forall j > q. \right\}.$$

Here we write the argument $\{n\}$ to emphasize how $\{(\iota_n^2, \theta_n)\}$ is indexed and $\{\alpha_n\}$ is any sequence of positive constants. In addition, for any set of parameter sequences Γ, let $\mathbb{P}(\Gamma)$ be the set of all sequences of DGP $\{\mathrm{P}^{(n)}\}$ satisfying that there exists some $\{(\iota_n^2, \theta_n)\} \in \Gamma$ such that, for each n, Assumptions 1 - 3 hold with $\mathrm{P} = \mathrm{P}^{(n)}$ and $((\iota^{(n)})^2, \theta^{(n)}) = (\iota_n^2, \theta_n)$.

We let $F(\cdot)$ be the standard Gaussian CDF. Furthermore, to simplify the notation, we set, for any probability measure P,

$$\mathrm{P}(\Delta_n, \alpha) = \mathrm{P}\left(\frac{\widehat{\sigma}_n^2(q) - \frac{1}{T}C(2)_T}{\Delta_n^{1/2}\sqrt{\mathrm{AVAR}\left(\frac{1}{T}C(2)_T, \gamma^{(n)}, \Delta_n, \frac{1}{T}C(4)_T - \frac{1}{T^2}C(2)_T, \mathrm{cum}_4[\varepsilon]\right)}} \in (-\infty, F^{-1}(\alpha))\right).$$

The claim of the theorem is equivalent with that, for any $\alpha_n \to 0$,

$$\inf_{\{\mathrm{P}^{(n)}\} \in \mathbb{P}(\Gamma_0(\{n\}, \{\alpha_n\}))} \liminf_{\Delta_n \to 0} \mathrm{P}^{(n)}(\Delta_n, \alpha) = \alpha = \sup_{\{\mathrm{P}^{(n)}\} \in \mathbb{P}(\Gamma_0(\{n\}, \{\alpha_n\}))} \limsup_{\Delta_n \to 0} \mathrm{P}^{(n)}(\Delta_n, \alpha). \tag{A.56}$$

To proceed, we introduce two more sets of parameter sequences:

$$\Gamma_1(a^2, \{n\}, \{\alpha_n\}) = \{\{(\iota_n^2, \theta_n) : n \geq 1\} \in \Gamma_0(\{n\}, \{\alpha_n\}) : \Delta_n^{-1}\iota_n^2 \to a^2 \in [0, \infty)\},$$

and

$$\Gamma_2(\{n\}, \{\alpha_n\}) = \{\{(\iota_n^2, \theta_n) : n \geq 1\} \in \Gamma_0(\{n\}, \{\alpha_n\}) : \Delta_n^{-1}\iota_n^2 \to \infty\}.$$

We notice that Theorem 4 and Lemma C5 together indicate that, for any $\{\mathrm{P}^{(n)}\} \in \mathbb{P}(\Gamma_1(a^2, \{n\}, \{\alpha_n\}) \cup \Gamma_2(\{n\}, \{\alpha_n\}))$ with $a^2 \in [0, \infty)$ and $\alpha_n \to 0$,

$$\lim_{\Delta_n \to 0} \mathrm{P}^{(n)}(\Delta_n, \alpha) = \alpha. \tag{A.57}$$

We show the first equality in (A.56), since proving the second one is a simple repetition. Let $\{\mathrm{P}_0^{(n)}\} \in \mathbb{P}(\Gamma_0(\{n\}, \{\alpha_n\}))$ be a sequence of probability measures such that

$$\liminf_{\Delta_n \to 0} \mathrm{P}_0^{(n)}(\Delta_n, \alpha) = \inf_{\{\mathrm{P}^{(n)}\} \in \mathbb{P}(\Gamma_0(\{n\}, \{\alpha_n\}))} \liminf_{\Delta_n \to 0} \mathrm{P}^{(n)}(\Delta_n, \alpha).$$

Such a sequence always exists. Let $\{u_n : n \geq 1\}$ be a subsequence of $\{n\}$ such that

$$\lim_{\Delta_n \to 0} \mathrm{P}_0^{(u_n)}(\Delta_{u_n}, \alpha) = \inf_{\{\mathrm{P}^{(n)}\} \in \mathbb{P}(\Gamma_0(\{n\}, \{\alpha_n\}))} \liminf_{\Delta_n \to 0} \widetilde{\mathrm{P}}_n(\Delta_n, \alpha).$$

Such a sequence always exists, too. The first equality in (A.56) follows once we show that there exists a subsequence $\{v_n\}$ of $\{u_n\}$ such that, for any $\alpha_n \to 0$,

$$\lim_{\Delta_n \to 0} \mathrm{P}_0^{(v_n)}(\Delta_{v_n}, \alpha) = \alpha. \tag{A.58}$$

To this end, we explicitly construct $\{v_n\}$. Let

$$\Gamma_3(\{n\}, \{\alpha_n\}) = \left\{ \{(\iota_n^2, \theta_n) : n \geq 1\} \in \Gamma_0(\{n\}, \{\alpha_n\}) : \limsup_{n \to \infty} \Delta_n^{-1} \iota_n^2 < \infty \right\},$$

and

$$\Gamma_4(\{n\}, \{\alpha_n\}) = \left\{ \{(\iota_n^2, \theta_n) : n \geq 1\} \in \Gamma_0(\{n\}, \{\alpha_n\}) : \limsup_{n \to \infty} \Delta_n^{-1} \iota_n^2 = \infty \right\}.$$

Observe that by construction for any $\{\mathrm{P}_0^{(u_n)}\} \in \mathbb{P}(\Gamma_0(\{u_n\}, \{\alpha_n\}))$, we have

$$\{\mathrm{P}_0^{(u_n)}\} \in \mathbb{P}(\Gamma_3(\{u_n\}, \{\alpha_n\}) \cup \Gamma_4(\{u_n\}, \{\alpha_n\})).$$

If $\{\mathrm{P}_0^{(u_n)}\} \in \mathbb{P}(\Gamma_3(\{u_n\}, \{\alpha_n\}))$, we can always find some subsequence $\{v_n\}$ of $\{u_n\}$ such that $\{\mathrm{P}_0^{(v_n)}\} \in \mathbb{P}(\Gamma_1(a^2, \{v_n\}, \{\alpha_n\}))$ with $a^2 < \infty$. If $\{\mathrm{P}_0^{(u_n)}\} \in \mathbb{P}(\Gamma_4(\{u_n\}, \{\alpha_n\}))$, then one can always find some subsequence $\{v_n\}$ of $\{u_n\}$ such that $\{\mathrm{P}_0^{(v_n)}\} \in \mathbb{P}(\Gamma_2(\{v_n\}, \{\alpha_n\}))$. Thus, to prove the first equality in (A.56), it suffices to demonstrate that (A.58) holds for any $\{\mathrm{P}_0^{(v_n)}\} \in \mathbb{P}(\Gamma_1(a^2, \{v_n\}, \{\alpha_n\}) \cup \Gamma_2(\{v_n\}, \{\alpha_n\}))$ with any $\{v_n\}$, $a^2 \in [0, \infty)$ and $\alpha_n \to 0$. Given (A.57), this can be proved by showing that, for any such $\{\mathrm{P}_0^{(v_n)}\}$, there always exists a sequence $\{\mathrm{P}_1^{(n)}\}$ such that

$$\{\mathrm{P}_1^{(n)}\} \in \mathbb{P}(\Gamma_1(a^2, \{n\}, \{\alpha_n\}) \cup \Gamma_2(\{n\}, \{\alpha_n\})) \text{ and } \mathrm{P}_1^{(v_n)} = \mathrm{P}_0^{(v_n)}, \forall n \geq 1. \tag{A.59}$$

By definition, for any $\{\mathrm{P}_0^{(v_n)}\} \in \mathbb{P}(\Gamma_1(a^2, \{v_n\}, \{\alpha_n\}))$, there is a sequence $\{(\iota_{v_n}^2, \theta_{v_n})\} \in \Gamma_1(a^2, \{v_n\}, \{\alpha_n\})$ such that Assumptions 1 - 3 hold with $\mathrm{P} = \mathrm{P}^{(v_n)}$ and $((\iota^2)^{(v_n)}, \theta^{(v_n)}) = (\iota_{v_n}^2, \theta_{v_n})$ for each v_n with $n \geq 1$. We construct another sequence $\{(\iota_k^{2\#}, \theta_k^{\#}) : k \geq 1\}$ as follows: (i) $\forall k = v_n$, define $(\iota_k^{2\#}, \theta_k^{\#}) = (\iota_{v_n}^2, \theta_{v_n})$, and (ii) $\forall k \in (v_n, v_{n+1})$, define $(\iota_k^{2\#}, \theta_k^{\#}) = (\iota_{v_n}^2 \Delta_{v_n}^{-1} \Delta_k, \theta_{v_n})$. Let $\{\mathrm{P}_1^{(n)}\}$ be a sequence such that $\mathrm{P}_1^{(v_n)} = \mathrm{P}_0^{(v_n)}$ for each v_n with $n \geq 1$ and Assumptions 1 - 3 hold with $\mathrm{P} = \mathrm{P}^{(v_n)}$ and $((\iota^2)^{(n)}, \theta^{(n)}) = (\iota_n^{2\#}, \theta_n^{\#})$ for each $n \geq 1$. Such a sequence always exists. On the other hand, we observe that $\{(\iota_n^{2\#}, \theta_n^{\#})\} \in \Gamma_1(a^2, \{n\}, \{\alpha_n\})$ with $a^2 \in (0, \infty)$. This means $\{\mathrm{P}_1^{(n)}\} \in \mathbb{P}(\Gamma_1(a^2, \{n\}))$. In the same way, we can construct some $\{\mathrm{P}_1^{(n)}\} \in \mathbb{P}(\Gamma_2(\{n\}, \{\alpha_n\}))$ with $\mathrm{P}_1^{(v_n)} = \mathrm{P}_0^{(v_n)}, \forall n \geq 1$, for any $\{\mathrm{P}_0^{(v_n)}\} \in \mathbb{P}(\Gamma_1(a^2, \{v_n\}, \{\alpha_n\}))$. So (A.59) is satisfied and we obtain the first equality in (A.56).

A.9 Proof of Corollary 3

The desired result follows from Theorem 5, Proof of Corollary 1 in Belloni, Chernozhukov, and Hansen (2014), and

$$\frac{\mathrm{AVAR}\left(\widehat{\sigma}_n^2(q), \widehat{\gamma}_n(q), \Delta_n, \widehat{C}_n(4)_T - \widehat{\sigma}_n^4(q), \widehat{\mathrm{cum}}_4[\varepsilon]\right)}{\mathrm{AVAR}\left(C_T, \gamma^{(n)}(q), \Delta_n, C(4)_T - C_T^2, \mathrm{cum}_4[\varepsilon]\right)} = 1 + o_{\mathrm{P}}(1). \tag{A.60}$$

We now show (A.60). We give the consistency of the estimates of all the arguments of the function AVAR, under both drifting sequences. Formally, under $\Delta_n^{-1}(\iota^{(n)})^2 \to \infty$,

$$\begin{aligned} &(\zeta^{(n)})^{-2} \sum_{|j| \le q} \widehat{\gamma}_n(q)_{|j|} \underset{(a1)}{=} 1 + o_{\mathrm{P}}(1), \quad \widehat{\sigma}_n^2(q) \underset{(a2)}{=} C_T + o_{\mathrm{P}}(1), \\ &\widehat{C}_n(4)_T \underset{(a3)}{=} C(4)_T + o_{\mathrm{P}}(1), \quad \widehat{\mathrm{cum}}_4[\varepsilon] \underset{(a4)}{=} O_{\mathrm{P}}(1). \end{aligned} \tag{A.61}$$

Here $(a1)$ holds by the first claim in (A.37). $(a2)$ holds by scenario (i) of Theorem 4. $(a2)$ holds by Theorem 16.4.2 in Jacod and Protter (2011). $(a4)$ is obvious.

Under $\Delta_n^{-1}(\iota^{(n)})^2 \to a^2 < \infty$,

$$\begin{aligned} &\|\widehat{\gamma}_n(q) - \gamma^{(n)}(q)\| \underset{(b1)}{=} o_{\mathrm{P}}(\Delta_n), \quad \widehat{\sigma}_n^2(q) \underset{(b2)}{=} C_T + o_{\mathrm{P}}(1), \\ &\widehat{C}_n(4)_T \underset{(b3)}{=} C(4)_T + o_{\mathrm{P}}(1), \quad a^4 \widehat{\mathrm{cum}}_4[\varepsilon] \underset{(b4)}{=} a^4 \mathrm{cum}_4[\varepsilon] + o_{\mathrm{P}}(1). \end{aligned} \tag{A.62}$$

Here $(b1)$ holds by (A.43), (A.31) and (A.30). $(b2)$ holds by scenario (iii) of Theorem 4. $(a2)$ holds by Theorem 16.4.2 in Jacod and Protter (2011). $(b4)$ holds by $\widehat{\mathrm{cum}}_4[\varepsilon] = \mathrm{cum}_4[\varepsilon] + o_{\mathrm{P}}(1)$ under $a^2 > 0$, which we show now. Observe that, under $a^2 > 0$, the combination of $(b1)$ above and the reasoning underlying (A.34) leads to $\inf_\lambda f(\lambda; \widehat{\gamma}_n(q)) > 0$ with $f(\lambda;\gamma)$ given by (4.13). The construction of $\widehat{\mathrm{cum}}_4[\varepsilon]$ (see the paragraph following (7.24)) implies that $\widehat{\mathrm{cum}}_4[\varepsilon]$ will be given by (7.24). One can verify that $(c4)$ holds in this case by using $(b1)$ -$(b3)$ above and observing

$$\begin{aligned} \frac{1}{n_T}\sum_{j=1}^{n_T}(Y_{n,j})^4 &= \frac{1}{n_T}\sum_{j=1}^{n_T}((U_j - U_{j-1})^4 + (\Delta_j^n X)^4 + 6(\Delta_j^n X)^2(U_j - U_{j-1})^2) + o_{\mathrm{P}}(\Delta_n^2) \\ &= \mathrm{cum}_4[\varepsilon](\iota^{(n)})^4 \sum_{j=0}^{q+1}(\theta_j^{(n)} - \theta_{j-1}^{(n)})^4 + 12(\gamma^{(n)}(q)_0 - \gamma^{(n)}(q)_1)^2 \\ &\quad + C(4)_T\Delta_n^2 + 12C_T\Delta_n(\gamma^{(n)}(q)_0 - \gamma^{(n)}(q)_1) + o_{\mathrm{P}}(\Delta_n^2). \end{aligned}$$

In view of (A.61) and (A.62), we deduce (A.60) from Lemma C5 and the proof concludes.

A.10 Proof of Proposition 3

Step 1. (Main proof) Throughout the proof we use instead the argument (σ^2, ι^2) and introduce $\widehat{\iota}^2_{w,n} = (\widehat{\gamma}_{w,n})_0$, as we take $q = 0$.

We start by providing the uniform asymptotic approximations to $L_{w,n}(\sigma^2, \iota^2)$ and $\frac{\partial}{\partial\sigma^2} L_{w,n}(\sigma^2, \iota^2)$:

$$f_n(\sigma^2, \iota^2) = -\frac{n}{2}\log \iota^2 - \frac{n\iota^{2\star}}{2\iota^2}$$

$$\text{and} \quad g_n(\sigma^2, \iota^2) = \frac{T}{8\sigma^2\iota}(\sigma^2\Delta_n)^{-1/2}\left(C_T + \frac{U_0^2 + U_n^2}{T} - \sigma^2\frac{\iota^{2\star}}{\iota^2}\right) - \frac{T}{4\iota}(\sigma^2\Delta_n)^{-1/2}\left(1 - \frac{\iota^{2\star}}{\iota^2}\right).$$

Formally, Step 2 below shows

$$\sup_{(\sigma^2,\iota^2)\in\Pi(0)} |L_{w,n}(\sigma^2, \iota^2) - f_n(\sigma^2, \iota^2)| = o_{\mathrm{P}}(n), \tag{A.63}$$

$$\sup_{(\sigma^2,\iota^2)\in\Pi(0)} \left|\frac{\partial}{\partial\sigma^2} L_{w,n}(\sigma^2, \iota^2) - g_n(\sigma^2, \iota^2)\right| = o_{\mathrm{P}}(n^{1/2}). \tag{A.64}$$

Hence, in view of that $(\widehat{\sigma}^2_{w,n}, \widehat{\iota}^2_{w,n})$ maximizes $L_{w,n}(\sigma^2, \iota^2)$, we readily obtain the desired result. Indeed, we can first prove by contradiction that (A.63) indicates $\widehat{\iota}^2_{w,n} = \iota^{2\star} + o_{\mathrm{P}}(1)$. Given the consistency of $\widehat{\iota}^2_{w,n}$, it then follows from (A.64) that $\widehat{\sigma}^2_{w,n} = C_T + T^{-1}(U_0^2 + U_n^2) + o_{\mathrm{P}}(1)$.

Step 2. (Uniform asymptotic approximations) In this step we prove (A.63) and (A.64). We need some notation. Let $\mathbb{J}_n, \mathbb{K}_n^+, \mathbb{K}_n^- \in \mathcal{M}_n$ be defined by

$$(\mathbb{J}_n)_{ij} = 2\delta_{ij} - \delta_{i,j+1} - \delta_{i,j-1}, \quad (\mathbb{K}_n^+)_{ij} = \delta_{1,i}\delta_{1,j}, \quad (\mathbb{K}_n^-)_{ij} = \delta_{n,i}\delta_{n,j}.$$

Based on these matrices, we introduce a "pseudo-log likelihood":

$$\bar{L}_{w,n}(\sigma^2, \iota^2) = -\frac{1}{2}\log\det V_{w,n} - \frac{1}{2}\mathrm{tr}(\Omega_{w,n}^{-1}\Sigma_{w,n}),$$

with

$$\Sigma_{w,n} = C_T\Delta_n\mathbb{I}_n + \iota^{2\star}\mathbb{J}_n + (U_0^2 - \iota^{2\star})\mathbb{K}_n^+ + (U_n^2 - \iota^{2\star})\mathbb{K}_n^-.$$

Step 3 below proves

$$\sup_{(\sigma^2,\iota^2)\in\Pi(0)} |L_{w,n}(\sigma^2, \iota^2) - \bar{L}_{w,n}(\sigma^2, \iota^2)| = o_{\mathrm{P}}(n), \tag{A.65}$$

$$\sup_{(\sigma^2,\iota^2)\in\Pi(0)} \left|\frac{\partial}{\partial\sigma^2} L_{w,n}(\sigma^2, \iota^2) - \frac{\partial}{\partial\sigma^2}\bar{L}_{w,n}(\sigma^2, \iota^2)\right| = o_{\mathrm{P}}(n^{1/2}). \tag{A.66}$$

Therefore, in view of the triangle inequality, (A.63) and (A.64) follow from the fact that f_n and g_n are, respectively, uniform asymptotic approximations of $\bar{L}_{w,n}(\sigma^2, \iota^2)$ and $\frac{\partial}{\partial\sigma^2}\bar{L}_{w,n}(\sigma^2, \iota^2)$. Now we show this is true. Let $\mathbb{D}_{w,n} \in \mathcal{M}_n$ be defined by $(\mathbb{D}_{w,n})_{ij} = 2\delta_{i,j}\left(1 - \cos\frac{2\pi j}{n}\right)$. We prove in Step 4 below

$$(O_{w,n}^\dagger \Sigma_{w,n} O_{w,n})_{jj} = \left(\left(C_T\Delta_n + \frac{U_0^2+U_n^2}{n} \right) \mathbb{I}_n + \frac{n-1}{n}\iota^{2\star}\mathbb{D}_{w,n} \right)_{jj}, \quad \forall 1 \le j \le n. \tag{A.67}$$

Using this, plus the definition of $\bar{L}_{w,n}$ and recalling $\Omega_{w,n} = O_{w,n}V_{w,n}O_{w,n}$, we conclude

$$\begin{aligned}
\bar{L}_{w,n}(\sigma^2,\iota^2) &= -\frac{1}{2}\log\det V_{w,n} - \frac{1}{2}\left(C_T\Delta_n + \frac{U_0^2+U_n^2}{n} \right)\mathrm{tr}(V_{w,n}^{-1}) - \frac{n-1}{2n}\iota^{2\star}\mathrm{tr}(V_{w,n}^{-1}\mathbb{D}_{w,n}),\\
\frac{\partial \bar{L}_{w,n}(\sigma^2,\iota^2)}{\partial\sigma^2} &= -\frac{\Delta_n}{2}\left(1-\frac{\iota^{2\star}}{\iota^2}\right)\mathrm{tr}(V_{w,n}^{-1}) + \frac{\Delta_n}{2}\left(C_T\Delta_n + \frac{U_0^2+U_n^2}{n} - \sigma^2\Delta_n\frac{\iota^{2\star}}{\iota^2}\right)\mathrm{tr}(V_{w,n}^{-2}).
\end{aligned}$$

Then the desired result follows from that uniformly over $(\sigma^2,\iota^2)\in\Pi(0)$,

$$\begin{aligned}
&\log\det V_{w,n} = n\log\iota^2 + +O_{\mathrm{P}}(n^{1/2}), \quad \mathrm{tr}(V_{w,n}^{-1}\mathbb{D}_{w,n}) = \frac{n}{\iota^2} + O_{\mathrm{P}}(n^{1/2}),\\
&\mathrm{tr}(V_{w,n}^{-1}) = \frac{n}{2\iota}(\sigma^2\Delta_n)^{-1/2} + O_{\mathrm{P}}(n), \quad \mathrm{tr}(V_{w,n}^{-2}) = \frac{n}{4\iota}(\sigma^2\Delta_n)^{-\frac{3}{2}} + O_{\mathrm{P}}(n^2),
\end{aligned} \tag{A.68}$$

which hold by Step 4 below.

Step 3. (Auxiliary: Uniform convergence) In this step we show (A.65) and (A.66). Given the definition of $\Pi(0)$, a simplified version of the reasoning in Step 5 of Lemma C4 indicates that it suffices to show the pointwise convergence. Namely, for all fixed $(\sigma^2,\iota^2)\in\Pi(0)$,

$$L_{w,n}(\sigma^2,\iota^2) - \bar{L}_{w,n}(\sigma^2,\iota^2) = o_{\mathrm{P}}(n) \quad \text{and} \quad \underbrace{\frac{\partial}{\partial\sigma^2}L_{w,n}(\sigma^2,\iota^2) - \frac{\partial}{\partial\sigma^2}\bar{L}_{w,n}(\sigma^2,\iota^2)}_{=:\mathcal{R}} = o_{\mathrm{P}}(n^{1/2}).$$

We only show the latter, from where the inconsistency of $\widehat{\sigma}_{w,n}^2$ originates. Observing $\mathbb{E}(\mathcal{R})=0$ by construction, the claim follows from

$$\mathrm{Var}(\mathcal{R}_b(n,\sigma^2,\iota^2)) \underset{(a1)}{\lesssim} \Delta_n^2\mathrm{tr}(V_{w,n}^{-2}O_{w,n}^\dagger\widetilde{\Sigma}_{w,n}O_{w,n}V_{w,n}^{-2}O_{w,n}^\dagger\widetilde{\Sigma}_{w,n}O_{w,n}) \underset{(a2)}{=} O_{\mathrm{P}}(n^{1/2}),$$

where $\widetilde{\Sigma}_{w,n} = C_T\Delta_n\mathbb{I}_n + \iota^{2\star}(\mathbb{J}_n - \mathbb{K}_n^+ - \mathbb{K}_n^-)$. Here $(a1)$ holds by the normality of the observations, while $(a2)$ holds by

$$(O_{w,n}^\dagger\widetilde{\Sigma}_{w,n}O_{w,n})_{jj} = \left(C_T\Delta_n\mathbb{I}_n + \frac{n-1}{n}\iota^{2\star}\mathbb{D}_{w,n}\right)_{jj} \quad \text{for all } 1\le j\le n, \tag{A.69}$$

$$\text{and} \quad \mathrm{tr}(V_{w,n}^{-4}\Delta_n^j\mathbb{D}_{w,n}^{2-j}) = O_{\mathrm{P}}(n^{5/2}), \quad \text{for all } j\in\{0,1,2\}. \tag{A.70}$$

Both (A.69) and (A.70) are from Step 4.

Step 4. (Auxiliary: Matrices and traces). In this step we show (A.67), (A.69), (A.68) and (A.70).

The first two follow from $O_{w,n}^{\dagger}O_{w,n} = \mathbb{I}_n$ and for all $1 \le j, l \le n$,

$$
\begin{aligned}
e^{\mathrm{i}\frac{2\pi}{n}j\cdot l}\sum_{m=1}^{n}(\mathbb{J}_n)_{lm}e^{-\mathrm{i}\frac{2\pi}{n}j\cdot m} &= \delta_{1,l}\left(2\left(1-\cos\frac{2\pi j}{n}\right)+1\right)+\delta_{n,l}\left(2\left(1-\cos\frac{2\pi j}{n}\right)+e^{\mathrm{i}\frac{2\pi}{n}j(l-1)}\right) \\
&\quad +\mathbb{1}_{\{2\le l\le n-1\}}\left(2\left(1-\cos\frac{2\pi j}{n}\right)+1\right).
\end{aligned}
$$

To see (A.68) and (A.70), using the same reasoning of Lemma D6, deduce

$$
(V_{w,n}^{-1})_{kk} \underset{(e1)}{=} \sum_{j=0}^{+\infty}(2-\delta_{j,0})\frac{1}{\iota^2}\frac{z^*}{1-(z^*)^2}(z^*)^{|j|}\cos\left(\frac{2\pi k}{n}\cdot j\right), \quad \text{with } z^* = 1-\frac{\sigma\Delta_n^{1/2}}{\iota}+O_{\mathrm{P}}(n^{-1}).
$$

Then (A.68) and (A.70) are results of $\sum_{k=0}^{n-1}\cos\left(\frac{2\pi k}{n}\cdot j\right) = n\delta_{j,0}$ for $0 \le j < n$ and observing for all $j \ge 1$,

$$
\frac{\partial V_{w,n}^{-j}}{\partial\sigma^2} = -j\Delta_n V_{w,n}^{-j-1}, \quad \iota^2\mathbb{D}_{w,n} = V_{w,n}-\sigma^2\Delta_n\mathbb{I}_n.
$$

A.11 Proof of Proposition 4

We only prove the result for the QMLE. Note for all (ι^2,θ) satisfying $\gamma_j = \frac{\iota^2}{2\pi}\int_{-\pi}^{\pi}g(\lambda;\theta)d\lambda$ for all $0 \le j \le q$, it holds that

$$
\mathrm{vec}(\mathcal{W}_n(\sigma^2,\gamma)) = -\frac{1}{n}\frac{d\mathrm{vec}(\Sigma_n(\sigma^2,\iota^2,\theta)^{-1})}{d(\sigma^2,\iota^2,\theta)}W(\sigma^2,\iota^2,\theta,\Delta_n)^{-1}(1,0_{q+1}),
$$

with

$$
W(\sigma^2,\iota^2,\theta,\Delta_n) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\partial\log f(\lambda;\sigma^2,\iota^2,\theta,\Delta_n)}{\partial(\sigma^2,\iota^2,\theta)}\left(\frac{\partial\log f(\lambda;\sigma^2,\iota^2,\theta,\Delta_n)}{\partial(\sigma^2,\iota^2,\theta)}\right)^{\intercal}d\lambda, \tag{A.71}
$$

$$
\text{and} \quad f(\lambda;\sigma^2,\iota^2,\theta,\Delta_n) = \sigma^2\Delta_n+\iota^2|1-e^{\mathrm{i}\lambda}|^2 g(\lambda;\theta).
$$

Below we omit the argument $(C_T,\gamma^\star)$ of Σ_n, Ω_n and $\mathcal{W}_n$ and use W_n to refer to $W(C_T,\iota^{2\star},\theta^\star,\Delta_n)$. Then the desired result follows from

$$
\mathrm{tr}(\mathcal{W}_n\Omega_n^Y) = C_T+o_{\mathrm{P}}(n^{-1/4}) \quad \text{and} \quad \mathrm{tr}(\mathcal{W}_n(Y_nY_n^{\intercal}-\Omega_n^Y)) = \widehat{\sigma}_n^2-C_T+o_{\mathrm{P}}(n^{-1/4}). \tag{A.72}
$$

In view of the observation made above, plus we can replace Σ_n in the definition of $\mathcal{W}_n$ by Ω_n using the arguments underlying Lemma B1, the first claim of (A.72) follows from two groups of technical results. The first group is

$$
\frac{\partial\mathrm{tr}(\Omega_n^{-1}\Omega_n^Y)}{\partial\sigma^2} = -\frac{1}{2}\frac{\Delta_n^{1/2}n}{\zeta^\star C_T^{1/2}}+o_{\mathrm{P}}(n^{1/4}), \quad \frac{\partial\mathrm{tr}(\Omega_n^{-1}\Omega_n^Y)}{\partial\iota^2} = -\frac{n}{\iota^{2\star}}+o_{\mathrm{P}}(n^{3/4}),
$$

$$\text{and}\quad \frac{\partial \mathrm{tr}(\Omega_n^{-1}\Omega_n^Y)}{\partial \theta_j} = o_{\mathrm{P}}(n^{3/4}) \quad \text{for all } 3 \le j \le q+2.$$

They hold by Lemma D3 and Lemma D6. The second group is

$$(W_n^{-1})_{11} = \frac{4C_T^{3/2}\zeta^\star}{\Delta_n^{1/2} n} + o_{\mathrm{P}}(n^{-3/4}), \quad (W_n^{-1})_{12} = -\frac{C_T \iota^{2\star}}{n} + o_{\mathrm{P}}(n^{-5/4}),$$

$$\text{and}\quad (W_n^{-1})_{1j} = O_{\mathrm{P}}(n^{-1}), \quad \text{for all } 3 \le j \le q+2.$$

To obtain them, first calculate $(W_n)_{ij}$ using the reasoning underlying Lemma D3, then apply the formula for inverting block matrices.

Using the second group of results above, the second claim of (A.72) comes from Step 4 of Section A.4 and Step 1 of Lemma B3.

A.12 Proof of Proposition 5

Step 1. (Main proof of claim (i)) In this step we prove claim (i). Claim (ii) will be proved in Step 4. Below we do not keep the arguments $(\sigma^2, \iota^2, \theta)$ unless necessary. We use W_n to refer to $W(\sigma^2, \iota^2, \theta, \Delta_n)$ defined by (A.71).

Define $n \times n$ matrices $\mathcal{R}_a$ and $\mathcal{R}_b(k)$ with $1 \le k \le q+1$ as

$$\mathcal{R}_a = \frac{\partial \Sigma_n^{-1}}{\partial \sigma^2}, \quad \mathcal{R}_b(k) = \delta_{k,1}\frac{\partial \Sigma_n^{-1}}{\partial \iota^2} + \mathbb{1}_{\{k\ge 2\}}\frac{\partial \Sigma_n^{-1}}{\partial \theta_{k-1}}.$$

By the construction of $\mathcal{W}_n$ and the observation made at the beginning of Section A.11, claim (i) follows from that for all $n^{1/2+\alpha} \le i \le n - n^{1/2+\alpha}$ with $0 < a < \frac{1}{2}$ fixed and $|l| \le q+1$,

$$(W_n^{-1})_{1,1}((\mathcal{R}_a)_{i,i+l} - (\mathcal{R}_a)_{i,i}) + \sum_{k=1}^{q+1}(W_n^{-1})_{1,k+1}(\mathcal{R}_b(k)_{i,i+l} - \mathcal{R}_b(k)_{i,i}) = o_{\mathrm{P}}(n^{-1/4}). \tag{A.73}$$

Before we prove (A.73), introduce short-hand notation $\nu = (\iota^2, \theta^\intercal)^\intercal \in \mathbb{R}^{q+1}$ and reparameterize the spectral density for noise process as

$$f(\lambda; \nu) = \nu_1 \Big| 1 + \sum_{j=1}^{q} \nu_{j+1} e^{\mathrm{i} j\lambda}\Big|^2.$$

For all $\nu \in \mathbb{R}^{q+1}$, define $(q+1)\times(q+1)$ matrix $\mathcal{R}_c(\nu)$ as

$$\mathcal{R}_c(\nu) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\partial \log f(\lambda;\nu)}{\partial \nu}\left(\frac{\partial \log f(\lambda;\nu)}{\partial \nu}\right)^\intercal d\lambda.$$

Then (A.73) can be derived from two groups of technical results.

The first group is

$$(W_n^{-1})_{1,1} = 4\sigma^3\zeta\Delta_n^{-1/2} + O_{\mathrm{P}}(1), \tag{A.74}$$

$$\text{and} \quad (W_n^{-1})_{1,k+1} = -\sigma^2 \Bigg(\underbrace{\left(\frac{\partial \log f(0;\nu)}{\partial \nu}\right)^{\intercal} \mathcal{R}_c(\nu)^{-1}}_{=:-\sigma^2\mathcal{R}_d(\nu,k)}\Bigg)_k + O_{\mathrm{P}}(n^{-1/2}), \tag{A.75}$$

where $1 \le k \le q+1$. Here the subscript $_k$ denotes the kth component of the $(q+1)$-dimensional vector. We now show these results are true. By the reasoning underlying Lemma D5, we conclude for $1 \le k, r \le q+1$.

$$(W_n)_{11} = \frac{1}{4}\frac{\Delta_n^{1/2}}{\sigma^3\zeta} + O_{\mathrm{P}}(n^{-1}), \quad (W_n)_{1,1+k} = \frac{1}{4}\frac{\Delta_n^{1/2}}{\sigma\zeta}\frac{\partial \log f(0;\nu)}{\partial \nu_k} + O_{\mathrm{P}}(n^{-1}),$$

$$(W_n)_{1+k,1+r} = \mathcal{R}_c(\nu)_{k,r} + O_{\mathrm{P}}(n^{-1/2}).$$

Applying block matrix inversion immediately yields (A.74) and (A.75).

The second group is that for all $|l| \le K$,

$$(\mathcal{R}_a)_{i,i+l} - (\mathcal{R}_a)_{i,i} \underset{(a4)}{=} \frac{\Delta_n^{1/2}}{8\sigma\zeta^3} \times l^2 + O_{\mathrm{P}}(n^{-1}), \tag{A.76}$$

$$\text{and} \quad \sum_{k=1}^{q+1} \mathcal{R}_d(\nu,k)(\mathcal{R}_b(k)_{i,i+l} - \mathcal{R}_b(k)_{i,i}) \underset{(a5)}{=} \frac{1}{2\zeta^2} \times l^2 + O_{\mathrm{P}}(n^{-1/2}). \tag{A.77}$$

We prove them in Step 2 below.

Step 2. (Auxiliary: Proof of (A.76) and (A.77)) First, (A.76) follows from that for all $|l| \le K$,

$$\begin{aligned}(\mathcal{R}_a)_{i,i+l} - (\mathcal{R}_a)_{i,i} &\underset{(a1)}{=} -\Delta_n(\Omega_n(\sigma^2,\iota^2,\theta)^{-2}_{i,i+l} - \Omega_n(\sigma^2,\iota^2,\theta)^{-2}_{i,i}) + O_{\mathrm{P}}(n^{-\infty}) \\ &\underset{(a2)}{=} -\Delta_n(\Omega_n(\sigma^2,\zeta^2,0)^{-2}_{i,i+l} - \Omega_n(\sigma^2,\zeta^2,0)^{-2}_{i,i}) + O_{\mathrm{P}}(n^{-1}) \underset{(c3)}{=} \frac{\Delta_n^{1/2}}{8\sigma\zeta^3} \times l^2 + O_{\mathrm{P}}(n^{-1}).\end{aligned}$$

Here $(a1)$ holds by the argument underlying Lemma B1 and $n^{1/2+\alpha} \le i, j \le n - n^{1/2+\alpha}$. $(a2)$ and $(a3)$ are easy to verify using Lemma D2 and Lemma D5.

Next, we prove (A.77). For all $|l| \le K$, write

$$\begin{aligned}\mathcal{R}_b(k)_{i,i+l} - \mathcal{R}_b(k)_{i,i} &\underset{(a4)}{=} \left(\frac{\partial \Omega_n(\sigma^2,\iota^2,\theta)^{-1}}{\partial \nu_k}\right)_{i,i+l} - \left(\frac{\partial \Omega_n(\sigma^2,\iota^2,\theta)^{-1}}{\partial \nu_k}\right)_{i,i} + O_{\mathrm{P}}(n^{-\infty}) \\ &\underset{(a5)}{=} \zeta^2\left(\Omega_n(\sigma^2,\zeta^2,0)^{-1}O_n\frac{\partial D_n(\iota^2,\theta)^{-1}}{\partial \nu_k}O_n\right)_{i,i+l} \\ &\quad -\zeta^2\left(\Omega_n(\sigma^2,\zeta^2,0)^{-1}O_n\frac{\partial D_n(\iota^2,\theta)^{-1}}{\partial \nu_k}O_n\right)_{ii} + O_{\mathrm{P}}(n^{-1/2})\end{aligned}$$

$$\underset{(a6)}{=} \quad -\frac{1}{2}\sum_{s=0}^{|l|-1}(2-\delta_{s,0})(|l|-s)\times\left(O_n\frac{\partial D_n(\iota^2,\theta)^{-1}}{\partial\nu_k}O_n\right)_{i,i+s}+O_{\mathrm{P}}(n^{-1/2}).$$

Here $(a4)$ holds by the argument underlying Lemma B1 and $n^{1/2+\alpha}\leq i,j\leq n-n^{1/2+\alpha}$. $(a5)$ and $(a6)$ are easy to verify using Lemma D2 and Lemma D5. In view of this, we obtain (A.77) by writing for all $0\leq s\leq q$,

$$\begin{aligned}\sum_{k=1}^{q+1}\mathcal{R}_d(\nu,k)\left(O_n\frac{\partial D_n(\iota^2,\theta)^{-1}}{\partial\nu_k}O_n\right)_{i,i+s} &\underset{(a7)}{=} \sum_{k=1}^{q+1}\mathcal{R}_d(\nu,k)\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\partial f^{-1}(\lambda;\nu)}{\partial\nu_k}e^{\mathrm{i}\lambda s}d\lambda+O_{\mathrm{P}}(n^{-1/2})\\ &:= \mathcal{R}_e(\nu,s)+O_{\mathrm{P}}(n^{-1/2})\underset{(a8)}{=}-\frac{1}{\zeta^2}+O_{\mathrm{P}}(n^{-1/2}),\end{aligned}$$

Here $(a7)$ holds by the same reasoning as in Step 1 of the proof of Lemma D3. $(a8)$ holds by Step 3 below.

Step 3. (Auxiliary: Proof of $(a8)$) For each $\nu\in\mathbb{R}^{q+1}$, define two $(q+1)$-dimensional vectors $h_+(\nu)$ and $h_-(\nu)$ and a $(q+1)\times(q+1)$ matrix $\mathcal{R}_f(\nu)$ as

$$h_+(\nu)_s=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(\lambda;\nu)e^{\mathrm{i}\lambda(s-1)}d\lambda,\quad h_-(\nu)_s=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{f(\lambda;\nu)}e^{\mathrm{i}\lambda(s-1)}d\lambda,\quad 1\leq s\leq q+1,\tag{A.78}$$

$$\mathcal{R}_f(\nu,\widetilde{\nu})=\frac{\partial h_+(\nu)}{\partial\nu}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\partial f(\lambda;\nu)}{\partial\nu}\left(\frac{\partial f^{-1}(\lambda;\widetilde{\nu})}{\partial\nu}\right)^{\intercal}d\lambda\right)^{-1}\left(\frac{\partial h_-(\nu)}{\partial\nu}\right)^{\intercal}.$$

From the construction of $\mathcal{R}_e(\nu,s)$ and using $\zeta^2=\sum_{s=1}^{q+1}(2-\delta_{s,1})h_+(\nu)_s$, one can verify

$$-\zeta^2\mathcal{R}_e(\nu,s)=\sum_{r=1}^{q+1}(2-\delta_{r,1})\mathcal{R}_f(\nu,\nu)_{r,s+1},\quad\forall\,0\leq s\leq q.$$

Hence, the desired result $(a8)$ follows from $(\mathcal{R}_f(\nu,\nu)^{-1})_{r,s}=(2-\delta_{r,1})\delta_{r,s}$. To see this, observe

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\partial f(\lambda;\nu)}{\partial\nu_r}\frac{\partial f^{-1}(\lambda;\widetilde{\nu})}{\partial\nu_s}d\lambda &\underset{(b1)}{=} \frac{\partial^2}{\partial\nu_r\partial\nu_s}\frac{1}{2\pi}\int_{-\pi}^{\pi}f(\lambda;\nu)f^{-1}(\lambda;\widetilde{\nu})d\lambda\\ &\underset{(b2)}{=} \frac{\partial^2}{\partial\nu_r\partial\nu_s}\sum_{s=1}^{q+1}h_+(\nu)_s h_-(\widetilde{\nu})_s(2-\delta_{s,1}),\end{aligned}$$

where $(b1)$ is obvious and $(b2)$ holds by (A.78) and $\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{\mathrm{i}\lambda s}d\lambda=\delta_{s,0}$. This yields $(a8)$:

$$(\mathcal{R}_f(\nu,\widetilde{\nu})^{-1})_{r,s}=\frac{\partial^2}{\partial h_+(\nu)_r\partial h_-(\nu)_s}\sum_{s=1}^{q+1}h_+(\nu)_s h_-(\widetilde{\nu})_s(2-\delta_{s,1})=(2-\delta_{r,1})\delta_{r,s}.$$

Step 4. (Proof of claim (ii)) Given (A.74) and (A.75) above, claim (ii) follows from

$$(\mathcal{R}_a)_{i,j} = \frac{1}{4\sigma^3\zeta\Delta_n^{1/2}}\left(1-\frac{\sqrt{\sigma^2\Delta_n}}{\zeta}+O(n^{-1})\right)^{|j-i|}(1+O(n^{-1/2})), \quad \sup_k|\mathcal{R}_b(k)_{i,j}| \lesssim (\mathcal{R}_a)_{i,j},$$

where both results hold by Lemma D3.

A.13 Proof of Theorem 6

Step 1. (Notation) Denote $\widetilde{C}_T = \frac{1}{T}[X,X]_T$, $I_T = \frac{1}{T}\int_0^T \iota_s^2\xi_s^{-1}ds$ and $\widetilde{I}_T = \left(\int_0^T \xi_s^{-1}ds\right)^{-1}\int_0^T \iota_s^2\xi_s^{-1}ds$. We adopt all the notation introduced in Step 1 of Section A.4, except that we replace every C_T, $(\iota^{(n)})^2$ and $(\zeta^{(n)})^2$ there with $\widetilde{C}_T$, $\widetilde{I}_T$ and $g(0;\theta^{(n)})\widetilde{I}_T$, respectively. We also introduce some additional notation below.

Part 1. For $q \geq 1$ we consider the successive jump times $\{T(r,m) : m \geq 1\}$ of the Poisson process $\underline{\mu}((0,t]\times\{z : 1/r < \Gamma(z) \leq 1/(r-1)\})$. Let $\{T_m\}_{m\geq 1}$ denotes any reordering of $\{T(r,m) : r,m \geq 1, T(r,m) < T\}$ and $T_0 = 0$. We additionally set $P_r = \{r : \exists r' \geq 1, r' \leq r, T_m = T(r',m')\}$ and define the following processes:

$$X^{J,r} = (\delta\mathbb{1}_{\{z:\Gamma(z)>1/r\}})\star\underline{\mu}, \quad X^{J,r-} = (\delta\mathbb{1}_{\{z:\Gamma(z)\leq 1/r\}})\star\left(\underline{\mu}-\underline{\nu}\right),$$

$$X_t^{B,r} = X_0 + \int_0^t \mu_s^r ds + \int_0^t \sigma_s dW_s, \text{ with } \mu_t^r = \mu_t - \int_{\{z:F(z)>1/r,|\delta(t,z)|\leq 1\}}\delta(t,z)\lambda(dz),$$

$$Y_{n,j}^{r-} = Y_{n,j} - \Delta_j^n X^{J,r}, \quad Y_n^{r-} = (Y_{n,1}^{r-}, Y_{n,2}^{r-}, \ldots, Y_{n,n}^{r-})^\intercal,$$

$$Y_{n,j}^{B,r} = \Delta_j^n X^{B,r} + \Delta_j^n U, \quad Y_n^{B,r} = (Y_{n,1}^{B,r}, Y_{n,2}^{B,r}, \ldots, Y_{n,n}^{B,r})^\intercal.$$

For any n, let $\Omega_n^{J,r}, \Omega_n^{J,r-}, \Omega_n^{Y,r-} \in \mathcal{M}_n$ be defined by

$$(\Omega_n^{J,r})_{ij} = \delta_{i,j}\sum_{m\in P_r:t_{i-1}^n\leq T_m<t_i^n}(\Delta X_{T_m})^2, \quad \Omega_n^{J,r-} = \Omega_n^J - \Omega_n^{J,r}, \quad \Omega_n^{Y,r-} = \Omega_n^{Y,B} + \Omega_n^{J,r-},$$

where we recall Ω_n^J and $\Omega_n^{Y,B}$ are defined in Part 2 of Section A.1.

Part 2. Using the definition of $\mathcal{L}(.,.)$ given by Part 3 of Section A.1, define

$$\begin{aligned} L_n^{B,r}(\sigma^2,\iota^2,\theta) &= \mathcal{L}(\Omega_n, Y_n^{B,r}(Y_n^{B,r})_n^\intercal), \quad \bar{L}_n^B(\sigma^2,\iota^2,\theta) = \mathcal{L}(\Omega_n,\Omega_n^{Y,B}), \\ L_n^{r-}(\sigma^2,\iota^2,\theta) &= \mathcal{L}(\Omega_n, Y_n^{r-}(Y_n^{r-})^\intercal), \quad \bar{L}_n^{r-}(\sigma^2,\iota^2,\theta) = \mathcal{L}(\Omega_n,\Omega_n^{Y,r-}), \end{aligned} \tag{A.79}$$

where we omit the argument $(\sigma^2,\iota^2,\theta)$ of Ω_n for notational simplicity.

Recall $\Pi_n^\varsigma(q)$ from Step 1 of Section A.4. For any $\varsigma \in \Pi_n^\varsigma(q)$ and $A_n \in \{L_n^{r-}, \bar{L}_n^{r-}, L_n^{B,r}, \bar{L}_n^B\}$, we let

$$A_n(\varsigma) = A_n(\sigma^2,\iota^2,\theta).$$

Moreover, with any $\varsigma \in \Pi_n^{\varsigma}(q)$ we define $\Xi_n^{r-}(\varsigma), \bar{\Xi}_n^{r-}(\varsigma), \Xi_n^{B,r}(\varsigma), \bar{\Xi}_n^{B}(\varsigma) \in \mathbb{R}^{q+2}$ such that

$$\begin{aligned} \Xi_n^{r-}(\varsigma)_j &= -\frac{1}{n}\frac{\partial L_n^{r,-}(\varsigma)}{\partial \varsigma_j}, \quad \bar{\Xi}_n^{r-}(\varsigma_n)_j = -\frac{1}{n}\frac{\partial \bar{L}_n^{r,-}(\varsigma)}{\partial \varsigma_j}, \\ \Xi_n^{B,r}(\varsigma)_j &= -\frac{1}{n}\frac{\partial L_n^{B,r}(\varsigma)}{\partial \varsigma_j}, \quad \bar{\Xi}_n^{B}(\varsigma)_j = -\frac{1}{n}\frac{\partial \bar{L}_n^{B}(\varsigma)}{\partial \varsigma_j}. \end{aligned} \tag{A.80}$$

Part 3. We denote by $\Omega(n,r)$ the set of all ω such that for any $m, m' \in P_r$, it holds that $n^{3/4}\Delta_n \leq T_m(\omega) \leq T_n - n^{3/4}\Delta_n$, and $|T_m(\omega) - T_{m'}(\omega)| > n^{3/4}\Delta_n$, and also $T_m(\omega)/\Delta_n$ is not an integer. Since the set $\{T_m : m \in P_r\}$ is finite, we have

$$\Omega(n,r) \longrightarrow \Omega \quad \text{a.s., as } n \to \infty.$$

Furthermore, we set a standard Wiener process $\{B_t\}_{t\geq 0}$ and set four independent sequences of i.i.d. random variables $\{\mathcal{U}_{m+}, \mathcal{U}_{m-}, \overline{\mathcal{U}}_{m+}, \overline{\mathcal{U}}_{m-}\}_{m\geq 1}$, all defined on an extension of the original space, independent of $\mathcal{F}$, and such that for each m, $\mathcal{U}_{m+}$, $\mathcal{U}_{m-}$, $\overline{\mathcal{U}}_{m+}$ and $\overline{\mathcal{U}}_{m-}$ are all standard Gaussian. We also let $\mathcal{Z}_{1,T}$ and $\mathcal{Z}_{2,T}(\delta)$ be two real random variables defined by

$$\begin{aligned} \mathcal{Z}_{1,T} =& \frac{\int_0^T \xi_s^{-1} ds}{8T^{5/4}} \int_0^T \sqrt{\frac{5\sigma_s^4 \xi_s}{\widetilde{C}_T^{7/2}} + \frac{2\sigma_s^2 \iota_s^2}{\widetilde{C}_T^{5/2} I_T} + \frac{\iota_s^4 \xi_s^{-1}}{\widetilde{C}_T^{3/2} I_T^2}} dB_s, \\ \mathcal{Z}_{2,T}(\delta) =& \frac{\int_0^T \xi_s^{-1} ds}{8T^{5/4}\widetilde{C}_T^{7/4}} \sum_{m\geq 1} \Delta X_{T_m} \Big(\sqrt{5}\sigma_{T_m}\xi_{T_m}^{1/2}\mathcal{U}_{m+} + \sqrt{5}\sigma_{T_m-}\xi_{T_m-}^{1/2}\mathcal{U}_{m-} \\ & + \sqrt{\widetilde{C}_T/I_T}(\iota_{T_m}\overline{\mathcal{U}}_{m+} + \iota_{T_m-}\overline{\mathcal{U}}_{m-}) \Big), \end{aligned}$$

where the argument is to emphasize the dependence of $\widetilde{\mathcal{Z}}_{2,T}$, through the jumps of X, on the function δ.

Step 2. (Main) Theorem 6 comes from four claims:

First, under $\widehat{\mathcal{R}}_n(q_n) = o_{\mathrm{P}}(1)$, $\mathcal{R}^{(n)}(q_n) = o_{\mathrm{P}}(1)$, $g(0;\theta^{(n)}) \to b^2 \in (0,\infty)$ and $q_n n^{-1/4} \to 0$, it holds

$$\frac{(Tg(0;\theta^{(n)})I_T)^{-1/4} n^{1/4}(\widehat{\sigma}^2(q_n) - \sigma^{(n)}(q_n)^2)}{\sqrt{\frac{1}{T}(5(C(4,\xi)_T + D(4,\xi)_T)\widetilde{C}_T^{-1/2} + \widetilde{C}_T^{3/2}B(\xi)_T)}} \xrightarrow{\mathcal{L}-s} \mathcal{N}(0,1).$$

Note that the definitions of $\widehat{\mathcal{R}}_n(q_n)$, $\mathcal{R}^{(n)}(q_n)$ and $\sigma^{(n)}(q_n)^2$ in this proof are different from those in Section A.4. Recall the modifications we impose at the beginning of Step 1. These changes are to accommodate the presence of jumps, random sampling and noise heteroskedasticity.

Second, the first claim still holds without assuming $g(0;\theta^{(n)}) \to b^2 \in (0,\infty)$.

Third, for $j \in \{1,2\}$,

$$\widehat{\mathcal{R}}_n(\widehat{q}_{n,j}) = o_{\mathrm{P}}(1), \quad \text{and} \quad \mathcal{R}^{(n)}(\widehat{q}_{n,j}) = o_{\mathrm{P}}(1).$$

Fourth, for $j \in \{1,2\}$,

$$\sigma^{(n)}(\widehat{q}_{n,j})^2 = \widetilde{C}_T + o_{\mathrm{P}}(n^{-1/4}), \quad \text{and} \quad \widehat{q}_{n,j} = o_{\mathrm{P}}(n^{1/4}).$$

We only prove the first claim in Step 3 below to demonstrate how to generalize the results of Section A.4 to the setup under Assumptions 4 - 6, since the other three claims follow from the same reasoning.

Step 3. (Proof of the first claim in Step 2) Following the same analysis of Step 3 of of Section A.4, it suffices to prove, under Assumptions 4 - 6, $\mathcal{R} = o_{\mathrm{P}}(n^{-1/2})$ therein and

$$(g(0;\theta^{(n)})I_T)^{1/4} n^{1/2} \Xi_n(\varsigma^{(n)})_1 \xrightarrow{\mathcal{L}-s} \mathcal{Z}_{1,T} + \mathcal{Z}_{2,T}(\delta). \tag{A.81}$$

Recall $\varsigma^{(n)}$ and $\Xi_n(\varsigma)$ are defined by (A.9) and (A.7). Note (A.81) goes back to (A.15) in the absence of jumps, random sampling and noise heteroskedasticity.

To facilitate the proof of (A.15), we need some notation. In view of (A.80), define

$$\mathcal{R}_a(n,r) := \Xi_n^{r,-}(\varsigma^{(n)})_1 - \bar{\Xi}_n^{r,-}(\varsigma^{(n)})_1 - \Xi_n^{B,r}(\varsigma^{(n)})_1 + \bar{\Xi}_n^{B,r}(\varsigma^{(n)})_1,$$

$$\mathcal{R}_b(n,r) := \Xi_n(\varsigma^{(n)})_1 - \Xi_n^{r,-}(\varsigma^{(n)})_1 + \bar{\Xi}_n^{r,-}(\varsigma^{(n)})_1.$$

Using the same reasoning of Step 4 of Section A.4, plus the facts that $X^{B,r}$ has no jump component by its definition in Step 1 and that μ_t^r is a bounded drift, we deduce for all $r \geq 1$,

$$(g(0;\theta^{(n)})I_T)^{1/4} n^{1/2} (\Xi_n^{B,r}(\varsigma^{(n)}) - \bar{\Xi}_n^{B,r}(\varsigma^{(n)}))_1 \xrightarrow{\mathcal{L}-s} \mathcal{Z}_{1,T}. \tag{A.82}$$

Next, Step 4 below proves

$$\lim_{r\to\infty} \limsup_{n\to\infty} n^{1/2} |\mathcal{R}_a(n,r)| = 0. \tag{A.83}$$

Moreover, Burkholder-Davis-Gundy inequality and Assumption 4 imply

$$\lim_{r\to\infty} \mathbb{E}(\sup_{s\leq T} |\mathcal{Z}_{2,T}(\delta(r)) - \mathcal{Z}_{2,T}(\delta)|^2) \lesssim \lim_{r\to\infty} \mathbb{E}\left(\sum_{s\leq T} |\Delta X_s|^2 \mathbb{1}_{\{z:\Gamma(z)\leq 1/q\}}\right) = 0. \tag{A.84}$$

Finally, in view of (A.80), (A.79) and Part 3 of Step 1 of Section A.1, we have

$$\begin{aligned}(g(0;\theta^{(n)})I_T)^{1/4} n^{1/2} \mathcal{R}_b(n,r) &\overset{a.s.}{=} -\sum_{m\in P_r} \Delta X_{T_m} n^{-\frac{1}{4}} \Delta_n (g(0;\theta^{(n)})I_T)^{1/4} \sum_{j=1}^{n} \Omega_n(\varsigma^{(n)})^{-2}_{I_m^n,j} Y_{n,j}^{r-} \\ &=: -\sum_{m\in P_r} \Delta X_{T_m} \zeta_m^{n,r} \xrightarrow{\mathcal{L}-s} \mathcal{Z}_{2,T}(\delta(r)),\end{aligned} \tag{A.85}$$

where we denote $I_m^n = [T_m/\Delta_n] + 1$, $\delta(r)(\omega,t,z) = \delta(\omega,t,z)\mathbb{1}_{\{z:\Gamma(z)>1/r\}}$. The last convergence holds by the fact that $\{T_m : m \in P_r\}$ is finite and that for each $m \in P_r$,

$$\zeta_m^{n,r} \overset{\mathcal{L}-s}{\to} \frac{\int_0^T \xi_s^{-1} ds}{8T^{5/4}\widetilde{C}_T^{7/4}} \left(\sqrt{5}\sigma_{T_m}\xi_{T_m}^{1/2}\mathcal{U}_{m+} + \sqrt{5}\sigma_{T_m-}\xi_{T_m-}^{1/2}\mathcal{U}_{m-} + \sqrt{\widetilde{C}_T/I_T}(\iota_{T_m}\overline{\mathcal{U}}_{m+} + \iota_{T_m-}\overline{\mathcal{U}}_{m-}) \right). \tag{A.86}$$

In view of the triangle inequality, combination of (A.82), (A.83), (A.84) and (A.85) readily yields (A.81).

Step 4. (Auxiliary: Proof of (A.83)) Let

$$\mathcal{R}_c(n,r) := \sum_{1\le j\le n} \Omega_n(\varsigma^{(n)})_{jj}^{-2}\Delta_j^n X^{J,r-}\Delta_j^n X^{B,r}, \quad \mathcal{R}_d(n,r) := \sum_{1\le i<j\le n} \Omega_n(\varsigma^{(n)})_{ij}^{-2}\Delta_i^n X^{J,r-}\Delta_j^n X^{B,r},$$

$$\mathcal{R}_e(n,r) := \sum_{1\le j<i\le n} \Omega_n(\varsigma^{(n)})_{ij}^{-2}\Delta_i^n X^{J,r-}\Delta_j^n X^{B,r}, \quad \mathcal{R}_f(n,r) := \sum_{1\le i,j\le n} \Omega_n(\varsigma^{(n)})_{ij}^{-2}\Delta_i^n X^{J,r-}\Delta_j^n U.$$

(A.83) follows from

$$\begin{aligned} \mathcal{R}_a(n,r) &\underset{(a1)}{=} \frac{1}{2n}\frac{\partial}{\partial\varsigma_1}\mathrm{tr}(\Omega_n(\varsigma^{(n)})^{-1}(Y_n^{r-}(Y_n^{r-})^\intercal - Y_n^{B,r}(Y_n^{B,r})^\intercal - \Omega_n^{J,r-})) \\ &\underset{(a2)}{=} -\frac{\Delta_n}{n^{3/4}}(\mathcal{R}_c(n,r) + \mathcal{R}_d(n,r) + \mathcal{R}_e(n,r) + \mathcal{R}_f(n,r)) \underset{(a3)}{=} o_{\mathrm{P}}(n^{-5/4}). \end{aligned}$$

Here $(a1)$ and $(a2)$ hold by definition, while $(a3)$ holds by the similar analysis in Lemma B3 and

$$\lim_{r\to\infty}\limsup_{n\to\infty}\mathbb{E}(\Delta_n^{-1}|\Delta_i^n X^{J,r-}|^2) \lesssim \lim_{r\to\infty}\int_{\{z:\Gamma(z)\le 1/r\}}\Gamma(z)^2\lambda(dz) = 0.$$

Step 5. (Auxiliary: Proof of (A.86)) Using $\Omega(n,r) \to \Omega$ a.s., that $|T_m(\omega) - T_{m'}(\omega)| > Tn^{-1/4}$ for all $m, m' \in P_r, \omega \in \Omega(n,r)$, that $\Omega_n(\varsigma_n)_{ij}^{-1} = O_{\mathrm{P}}(n^{-\infty})$ for all $|i-j| \gtrsim n^{3/4}$ by Lemma D3 and that $X^{B,r} = X - X^{J,r} - X^{J,r-}$ by Part 1 of Step 1, one can decompose

$$\zeta_m^{n,r} = \sum_{l=1}^3 \bar{\zeta}(l)_m^{n,q} + \sum_{l=1}^3 \bar{\zeta}'(l)_m^{n,q} + O_{\mathrm{P}}(n^{-\infty}),$$

where

$$\begin{aligned} \bar{\zeta}(1)_m^{n,q} &= n^{-\frac14}\Delta_n(g(0;\theta^{(n)})I_T)^{1/4}\sum_{j=0}^{[n^{3/4}]}(\Omega_n(\varsigma^{(n)})_{I_m^n,I_m^n+j}^{-2} - \Omega_n(\varsigma^{(n)})_{I_m^n,I_m^n+j+1}^{-2})U_{I_m^n+j}, \\ \bar{\zeta}(2)_m^{n,q} &= n^{-\frac14}\Delta_n(g(0;\theta^{(n)})I_T)^{1/4}\sum_{j=1-[n^{3/4}]}^{-1}(\Omega_n(\varsigma^{(n)})_{I_m^n,I_m^n+j}^{-2} - \Omega_n(\varsigma^{(n)})_{I_m^n,I_m^n+j+1}^{-2})U_{I_m^n+j}, \\ \bar{\zeta}(3)_m^{n,q} &= n^{-\frac14}\Delta_n(g(0;\theta^{(n)})I_T)^{1/4}\sum_{j=1}^{[n^{3/4}]}\Omega_n(\varsigma^{(n)})_{I_m^n,I_m^n+j}^{-2}\Delta_{I_m^n+j}^n X^{B,r}, \end{aligned}$$

$$\bar{\zeta}(4)^{n,q}_m = n^{-\frac{1}{4}}\Delta_n(g(0;\theta^{(n)})I_T)^{1/4}\sum_{j=1-[n^{3/4}]}^{-1}\Omega_n(\varsigma^{(n)})^{-2}_{I^n_m,I^n_m+j}\Delta^n_{I^n_m+j}X^{B,r},$$

$$\bar{\zeta}'(1)^{n,q}_m = n^{-\frac{1}{4}}\Delta_n(g(0;\theta^{(n)})I_T)^{1/4}\sum_{j=1}^{[n^{3/4}]}\Omega_n(\varsigma^{(n)})^{-2}_{I^n_m,I^n_m+j}\Delta^n_{I^n_m+j}X^{J,r-},$$

$$\bar{\zeta}'(2)^{n,q}_m = n^{-\frac{1}{4}}\Delta_n(g(0;\theta^{(n)})I_T)^{1/4}\sum_{j=1-[n^{3/4}]}^{-1}\Omega_n(\varsigma^{(n)})^{-2}_{I^n_m,I^n_m+j}\Delta^n_{I^n_m+j}X^{J,r-},$$

$$\bar{\zeta}'(3)^{n,q}_m = n^{-\frac{1}{4}}\Delta_n(g(0;\theta^{(n)})I_T)^{1/4}\Omega_n(\varsigma^{(n)})^{-2}_{I^n_m,I^n_m}(\Delta^{I^n_m}_n X^{J,r-}+\Delta^{I^n_m}_n X^{B,r}).$$

Show $\bar{\zeta}'(l)^{n,q}_m = o_{\mathrm{P}}(1)$ for $l=1,2,3$ is the same as in Step 4. Following the standard steps based on the CLT for triangular array (see Theorem 2.2.15 in Jacod and Protter (2011)) and using Lemma D6 lead to

$$\bar{\zeta}(1)^{n,q}_m \xrightarrow{\mathcal{L}-s} \frac{\iota_{T_m}\int_0^T \xi_s^{-1}ds}{8T^{5/4}\widetilde{C}_T^{5/4}I_T^{1/2}}\overline{\mathcal{U}}_{m+}, \quad \bar{\zeta}(2)^{n,q}_m \xrightarrow{\mathcal{L}-s} \frac{\iota_{T_m-}\overline{\mathcal{U}}_{m-}\int_0^T \xi_s^{-1}ds}{8T^{5/4}\widetilde{C}_T^{5/4}I_T^{1/2}},$$

$$\bar{\zeta}(3)^{n,q}_m \xrightarrow{\mathcal{L}-s} \frac{\sqrt{5}\sigma_{T_m}\xi^{1/2}_{T_m}\mathcal{U}_{m+}\int_0^T \xi_s^{-1}ds}{8T^{5/4}\widetilde{C}_T^{7/4}}, \quad \bar{\zeta}(4)^{n,q}_m \xrightarrow{\mathcal{L}-s} \frac{\sqrt{5}\sigma_{T_m-}\xi^{1/2}_{T_m-}\mathcal{U}_{m-}\int_0^T \xi_s^{-1}ds}{8T^{5/4}\widetilde{C}_T^{7/4}}.$$

A.14 Proof of Corollary 4

The desired result follows from (a) the same reasoning as in Section A.5, (b) Theorem 6 and (c)

$$\widehat{C}_n(4)_T \underset{(a1)}{=} C(4,\xi)_T + o_{\mathrm{P}}(1), \quad \widehat{D}_n(4)_T \underset{(a2)}{=} D(4,\xi)_T + o_{\mathrm{P}}(1), \quad \widehat{B}_n(\widehat{q}) \underset{(a3)}{=} B(\xi)_T + o_{\mathrm{P}}(1).$$

Here $(a1)$ and $(a2)$ hold by generalizing, respectively, Theorem 16.4.2 and Theorem 16.5.4 in Jacod and Protter (2011). Moreover, $(a3)$ holds by

$$\widehat{B}'_n(1) = \frac{2}{T}\int_0^T \iota_s^2\sigma_s^2 ds \times A^{(n)} + o_P(1), \quad \widehat{B}'_n(2) = \frac{1}{T}\sum_{s\le T}(\iota_s^2+\iota_{s-}^2)(\Delta X_s)^2 \times A^{(n)} + o_P(1),$$

$$\widehat{B}'_n(3) = \frac{1}{T}\int_0^T \iota_s^4\xi_s^{-1}ds \times \left(A^{(n)}\right)^2 + o_P(1), \quad \widehat{\gamma}_n(\widehat{q})_0 - \widehat{\gamma}_n(\widehat{q})_1 = \frac{1}{T}\int_0^T \iota_s^2\xi_s^{-1}ds \times A^{(n)} + o_P(1),$$

where $A^{(n)} = \sum_{j=0}^\infty \left(\theta_j^{(n)}\right)^2 - \sum_{j=0}^\infty \theta_j^{(n)}\theta_{j+1}^{(n)}$. These four results follow from generalizing Theorem 16.5.1 and Theorem 16.5.4 in Jacod and Protter (2011) and using the third claim in Step 2 of Section A.13.

A.15 Proof of Theorem 7

In view of the definitions of $\widehat{\gamma}_n(\widehat{q}_{n,j})$ and $\widehat{\rho}_n(\widehat{q}_{n,j})$, (6.23) follows immediately from (6.22) and the fact that $\gamma_0^{(n)} \ge K^{-1}$. We hence only need to prove (6.22). Further, in the interest of space, we restrict ourselves to proving (6.22) under the assumptions of Theorem 3. One can construct the proof which accomodates the more general setup originally assumed in Theorem 7 by following Section A.13. Formally, we aim to show under the assumptions of Theorem 3 and for $j \in \{1, 2\}$,

$$\|\widehat{\theta}_n(\widehat{q}_{n,j}) - \theta^{(n)}\| + |\widehat{\iota}_n^2(\widehat{q}_{n,j}) - (\iota^{(n)})^2| = O_{\mathrm{P}}(n^{-3/8}). \tag{A.87}$$

We adopt all the notation introduced in Step 1 of Section A.4. In view of the definitions of $\widehat{\varsigma}_n(q)$ and $\varsigma_\infty^{(n)}$ given in (A.9), plus using $\|\cdot\|_{(1)}$ introduced in (A.1), we rewrite (A.87) as $\|\widehat{\varsigma}_n(\widehat{q}_{n,j}) - \varsigma_\infty^{(n)}\|_{(1)} = O_{\mathrm{P}}(n^{-3/8})$. This, on the other hand, follows from

$$\|\widehat{\varsigma}_n(\widehat{q}_{n,j}) - \varsigma^{(n)}(\widehat{q}_{n,j})\|_{(1)} = O_{\mathrm{P}}(n^{-3/8}) \quad \text{and} \quad \|\varsigma^{(n)}(\widehat{q}_{n,j}) - \varsigma_\infty^{(n)}\|_{(1)} = O_{\mathrm{P}}(n^{-3/8}). \tag{A.88}$$

Here we use the triangle inequality. Recall $\varsigma^{(n)}(q)$ is also given in (A.9). Step 9 of Section A.4 already proves the second claim in (A.88). We deduce the first claim in (A.88) from (A.14), (A.22), (A.19), (A.15), and the fact that $\widehat{q}_{n,j} = o_{\mathrm{P}}(n^{1/4})$ due to Step 9 of Section A.4.

Appendix B Supplement to Section A.4

In the statements of all lemmas in this section, we use "(xx) holds" to mean "Under the conditions given in the context, (xx) in Section A.4 holds". Moreover, we adopt all the notation introduced in Step 1 of Section A.4.

Lemma B1. *(A.16) holds.*

Proof. Step 1. (Main proof) As in the context in which (A.16) appears, throughout the proof we drop the argument q_n of $\varsigma^{(n)}(q_n)$. For all $\varsigma \in \Pi^\varsigma(q_n)$ we define $\Sigma_n(\varsigma)$ and $\Omega_n(\varsigma)$ by reparameterizing $\Sigma_n(\sigma^2, \iota^2, \theta)$ and $\Omega_n(\sigma^2, \iota^2, \theta)$. Formally, let $\Sigma_n(\varsigma) = \Sigma_n(\sigma^2, \iota^2, \theta)$ and $\Omega_n(\varsigma) = \Omega_n(\sigma^2, \iota^2, \theta)$ with $\iota^2 = \varsigma_2$, $\theta = (\varsigma_3, \ldots, \varsigma_{q_n+2})^\intercal$ and $\sigma^2 = \varsigma_1 n^{1/4}$. Recall $\Omega_n(\sigma^2, \iota^2, \theta)$ is introduced in Part 2 of Section A.1. Using $\Sigma_n(\varsigma)$ and $\Omega_n(\varsigma)$, plus the definitions of $\Xi_n(\varsigma)$ and $\Xi_{A,n}(\varsigma)$ given by (A.7), (A.6) and Part 3 of Section A.1, we rewrite (A.16) as

$$n^{-1/4} \underbrace{\frac{\partial}{\partial \varsigma_1} \log\det(\Sigma_n(\varsigma^{(n)})\Omega_n(\varsigma^{(n)})^{-1})}_{=:\mathcal{R}_a} + n^{-1/4} \underbrace{Y_n^\intercal \frac{\partial}{\partial \varsigma_1}(\Sigma_n(\varsigma^{(n)})^{-1} - \Omega_n(\varsigma^{(n)})^{-1}) Y_n}_{=:\mathcal{R}_b} = o_{\mathrm{P}}(n^{1/4}). \tag{B.1}$$

We hence only need to prove $\mathcal{R}_a = o_{\mathrm{P}}(n^{1/4})$ and $\mathcal{R}_b = o_{\mathrm{P}}(n^{1/4})$. Observe $n^{-1/4}\frac{\partial}{\partial \varsigma_1}\Sigma_n(\varsigma) = n^{-1/4}\frac{\partial}{\partial \varsigma_1}\Omega_n(\varsigma) = \Delta_n$. Then we can write

$$\mathcal{R}_a = \Delta_n \mathrm{tr}(\Sigma_n(\varsigma^{(n)})^{-1} - \Omega_n(\varsigma^{(n)})^{-1}) \quad \text{and} \quad \mathcal{R}_b = -\Delta_n Y_n^\intercal (\Sigma_n(\varsigma^{(n)})^{-2} - \Omega_n(\varsigma^{(n)})^{-2}) Y_n. \tag{B.2}$$

Here we use $\log\det A = \mathrm{tr}\log A$. The challenge we face is that we do not have an analytical expression for Σ_n^{-1}. However, we observe

$$\Sigma_n^{-1} = \Omega_n^{-1} - \Omega_n^{-1} R_n \Omega_n^{-1} + \Omega_n^{-1} R_n \Sigma_n^{-1} R_n \Omega_n^{-1}, \quad \text{with} \quad R_n(\varsigma) := \Sigma_n(\varsigma) - \Omega_n(\varsigma). \tag{B.3}$$

Although in the last term of the RHS Σ_n^{-1} still appears, it turns out we can replace it with Ω_n^{-1} for the purpose of bounding $\mathcal{R}_a$ and $\mathcal{R}_b$. We elaborate the reasoning underlying this in Step 2. Now we apply (B.3) to (B.2). Introduce simplifying notation

$$\left.\begin{aligned} &\mathcal{R}_{a1}(\varsigma) := \operatorname{tr}(\Omega_n^{-1} R_n \Omega_n^{-1}), \quad \mathcal{R}_{a2}(\varsigma) := \operatorname{tr}(\Omega_n^{-1} R_n \Sigma_n^{-1} R_n \Omega_n^{-1}),\\ &\mathcal{R}_{b1}(\varsigma) := Y_n^\intercal \Omega_n^{-1} R_n \Omega_n^{-2} Y_n, \quad \mathcal{R}_{b2}(\varsigma) := Y_n^\intercal \Omega_n^{-1} R_n \Sigma_n^{-1} R_n \Omega_n^{-2} Y_n,\\ &\mathcal{R}_{b3}(\varsigma) := Y_n^\intercal \Omega_n^{-1} R_n \Sigma_n^{-2} R_n \Omega_n^{-1} Y_n. \end{aligned}\right\} \tag{B.4}$$

Here we drop the arguments ς of Ω_n and Σ_n. Then we can rewrite (B.2) as

$$\Delta_n^{-1}\mathcal{R}_a = -\mathcal{R}_{a1}(\varsigma^{(n)}) + \mathcal{R}_{a2}(\varsigma^{(n)}) \quad \text{and} \quad \Delta_n^{-1}\mathcal{R}_b = 2\mathcal{R}_{b1}(\varsigma^{(n)}) - 2\mathcal{R}_{b2}(\varsigma^{(n)}) - \mathcal{R}_{b3}(\varsigma^{(n)}).$$

In view of the triangle inequality, the desired result (B.1) follows from the fact that for all $A \in \{\mathcal{R}_{a1}, \mathcal{R}_{a2}, \mathcal{R}_{b1}, \mathcal{R}_{b2}, \mathcal{R}_{b3}\}$, $A(\varsigma^{(n)}) = o_{\mathrm{P}}(n^{5/4})$. Step 2 below proves $A(\varsigma^{(n)}) = o_{\mathrm{P}}(n^{5/4})$ for $A = \mathcal{R}_{b2}$. The other cases only require simpler analysis.

Step 2. (Bound on $\mathcal{R}_{b2}(\varsigma^{(n)})$) In this step we show $\mathcal{R}_{b2}(\varsigma^{(n)}) = o_{\mathrm{P}}(n^{5/4})$. Note Σ_n is positive definite as it is a covariance matrix. Then the definition of $\mathcal{R}_{b2}(\varsigma)$ given by (B.4) and the Cauchy-Schwarz inequality indicate that $\mathcal{R}_{b2}(\varsigma^{(n)}) = o_{\mathrm{P}}(n^{5/4})$ follows from

$$Y_n^\intercal \Omega_n^{-1} R_n \Sigma_n^{-1} R_n \Omega_n^{-1} Y_n = o_{\mathrm{P}}(n^{1/4}) \quad \text{and} \quad Y_n^\intercal \Omega_n^{-2} R_n \Sigma_n^{-1} R_n \Omega_n^{-2} Y_n = o_{\mathrm{P}}(n^{9/4}). \tag{B.5}$$

Here Ω_n, R_n and Σ_n are evaluated at $\varsigma^{(n)}$. We only show the first claim in (B.5) as a demonstration.

In Step 1 we mention that we can replace Σ_n^{-1} with Ω_n^{-1}. Obviously we can do so if $\Sigma_n(\varsigma)^{-1} \leq K\Omega_n(\varsigma)^{-1}$ for all $\varsigma \in \Pi_n^{\varsigma}(q_n)$. Here $\leq$ stands for Loewner partial order. Indeed, Corollary 7.7.4 in Horn and Johnson (2013) states that for any two Hermitian matrices A and B, if $A < B$, then $A^{-1} < B^{-1}$. Hence, we only need $\Omega_n(\varsigma) \lesssim \Sigma_n(\varsigma)$. Step 3 below proves for all $(\sigma^2, \iota^2, \theta) \in \Pi(q_n)$,

$$\Sigma_n(\sigma^2, \iota^2, 0) \lesssim K\Sigma_n(\sigma^2, \iota^2, \theta). \tag{B.6}$$

On the other hand, in view of the definitions of Ω_n and V_n given in Part 2 of Section A.1, plus $K^{-1} \leq g(\lambda;\theta) \leq K$ for all $\lambda \in [-\pi, \pi]$ and $\theta \in \Theta(\infty)$, we conclude for all $(\sigma^2, \iota^2, \theta) \in \Pi(q_n)$,

$$\Omega_n(\sigma^2, \iota^2, \theta) = O_n V_n(\sigma^2, \iota^2, \theta, \Delta_n) O_n \lesssim O_n V_n(\sigma^2, \iota^2, 0, \Delta_n) O_n = \Omega_n(\sigma^2, \iota^2, 0) = \Sigma_n(\sigma^2, \iota^2, 0). \tag{B.7}$$

Here the last equality can be verified using Lemma D2. Given (B.6) and (B.7) we have $\Omega_n(\sigma^2, \iota^2, \theta) \lesssim \Sigma_n(\sigma^2, \iota^2, \theta)$ for all $(\sigma^2, \iota^2, \theta) \in \Pi(q_n)$ hence $\Omega_n(\varsigma) \lesssim \Sigma_n(\varsigma)$ for all $\varsigma \in \Pi^{\varsigma}(q_n)$. Thereby, the first claim in (B.5) follows from

$$Y_n^\intercal \Omega_n^{-1} R_n \Omega_n^{-1} R_n \Omega_n^{-1} Y_n = o_{\mathrm{P}}(n^{1/4}), \tag{B.8}$$

with Ω_n, R_n and Σ_n evaluated at $\varsigma^{(n)}$. We prove (B.8) in Step 4 below.

Step 3. (Auxiliary: Proof of (B.6)) In this step we bound $\Sigma_n(\sigma^2, \iota^2, \theta)$ from below by $\Sigma_n(\sigma^2, \iota^2, 0)$. For all $x = (x_1, x_2, \ldots, x_n)^\intercal \in \mathbb{R}^n$, define $\widetilde{x} = (\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_{n+1})^\intercal \in \mathbb{R}^{n+1}$ by $\widetilde{x}_j = x_{j-1} - x_j$ with $x_0 = x_{n+1} = 0$. We deduce (B.6) from

$$x^\intercal \Sigma_n(\sigma^2, \iota^2, \theta) x = \sigma^2 \Delta_n \|x\|^2 + \iota^2 \widetilde{x}^\intercal \Gamma(\theta)_{n+1} \widetilde{x} \quad \text{and} \quad \Gamma(\theta)_{n+1} \sim \mathbb{I}_{n+1}. \tag{B.9}$$

Here we define $\Gamma(\theta)_{n+1} \in \mathcal{M}_{n+1}$ by $(\Gamma(\theta)_{n+1})_{ij} = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\lambda; \theta) e^{\mathrm{i}(i-j)\lambda} d\lambda$. $\mathbb{I}_{n+1}$ is the $(n+1)$-dimensional identity matrix. The first claim in (B.9) holds by definition of Σ_n. The second claim in (B.9) comes from Proposition 4.5.3 in Brockwell and Davis (1991) and the fact that $K^{-1} \le g(\lambda; \theta) \le K$ for all $\lambda \in [-\pi, \pi]$ and $\theta \in \Theta(\infty)$.

Step 4. (Proof of (B.8)) First, we derive the expression of R_n. In view of the definition of Ω_n given by Part 2 of Section A.1, plus Lemma D2 and the definition of $\{\mathbb{F}_n^h : 0 \le h \le n\}$ given by (D.1) therein, we write

$$\Omega_n(\sigma^2, \iota^2, \theta) = \sigma^2 \Delta_n \mathbb{I}_n + \sum_{h=0}^{q_n} \gamma_h (2\mathbb{F}_n^h - \mathbb{F}_n^{h+1} - \mathbb{F}_n^{h-1}). \tag{B.10}$$

Recall (3.6) for the connection between γ and (ι^2, θ). Here $\mathbb{F}_n^h = 0_{n\times n}$ for $h = -1$ by convention. On the other hand, we rewrite Σ_n defined by (3.5) as

$$\Sigma_n(\sigma^2, \iota^2, \theta) = \sigma^2 \Delta_n \mathbb{I}_n + \sum_{h=0}^{q_n} \gamma_h (2\mathbb{G}_n^h - \mathbb{G}_n^{h+1} - \mathbb{G}_n^{h-1}). \tag{B.11}$$

To write R_n in a more compact form, define $\mathbb{K}_n^h, \widetilde{\mathbb{K}}_n^h \in \mathcal{M}_n$ by $(\mathbb{K}_n^h)_{ij} = \mathbb{1}_{\{h=i+j\}} - \mathbb{1}_{\{h+1=i+j\}}$ and $(\widetilde{\mathbb{K}}_n^h)_{ij} = (\mathbb{K}_n^h)_{n-i,n-j}$. Obviously $\mathbb{K}_n^h + \widetilde{\mathbb{K}}_n^h = \mathbb{G}_n^h - \mathbb{G}_n^{h+1} - \mathbb{F}_n^h + \mathbb{F}_n^{h+1}$, hence (B.10) and (B.11) lead to

$$R_n(\varsigma^{(n)}) = \Sigma_n(\varsigma^{(n)}) - \Omega_n(\varsigma^{(n)}) = \sum_{h=0}^{q_n-1} (\gamma_h^{(n)} - \gamma_{h+1}^{(n)})(\mathbb{K}_n^h + \widetilde{\mathbb{K}}_n^h). \tag{B.12}$$

Here $\gamma_{q_n+1}^{(n)} = 0$ by convention.

Next, we apply (B.12) to (B.8). In view of the symmetry between $\mathbb{K}_n^h$ and $\widetilde{\mathbb{K}}_n^h$, plus the fact that $\Omega_n(\varsigma)_{i,j}^{-1} = O_{\mathrm{P}}(n^{-\infty})$ for all $|i-j| \sim n^{1/2+\alpha}$ for all $\alpha > 0$ due to Lemma D3, we can replace $R_n(\varsigma^{(n)})$ in (B.8) with $\widetilde{R}_n(\varsigma^{(n)}) := \sum_{h=0}^{q_n} (\gamma_h^{(n)} - \gamma_{h+1}^{(n)}) \mathbb{K}_n^h$ for the purpose of proving it. Then (B.8) becomes

$$\sum_{h=0}^{q_n-1} \sum_{l=0}^{q_n-1} (\gamma_h^{(n)} - \gamma_{h+1}^{(n)})(\gamma_l^{(n)} - \gamma_{l+1}^{(n)}) Y_n^\intercal \Omega_n^{-1} \mathbb{K}_n^h \Omega_n^{-1} \mathbb{K}_n^l \Omega_n^{-1} Y_n = o_{\mathrm{P}}(n^{1/4}). \tag{B.13}$$

In the context in which (A.16) appears we have $q_n n^{-1/4} \to 0$. The definition of $\Theta(\infty)$ states $\sum_{j=1}^{\infty} j|\theta_j| \le K$ for all $\theta \in \Theta(\infty)$, hence $\sum_{h=0}^{\infty} h|\gamma_h^{(n)}| \lesssim 1$. Therefore, applying Chebyshev's inequality, we only need to prove for all $n^{1/4}h_n \to 0$ and $n^{1/4}l_n \to 0$,

$$\mathbb{E}|Y_n^\intercal \Omega_n^{-1} \mathbb{K}_n^{h_n} \Omega_n^{-1} \mathbb{K}_n^{l_n} \Omega_n^{-1} Y_n| \lesssim n^{1/4} h_n l_n. \tag{B.14}$$

This is done in Step 5.

Before we move to Step 5, we briefly explain why the approximation based on Ω_n is more precise than the Whittle approximation. Plugging (D.1) into Lemma D3 yields that the magnitude of $(\Omega_n^{-1})_{i,j}$ shrinks as either i or j approaches 1 or n, keeping $|i-j|$ invariant. Inverse of the Whittle covariance matrix $\Omega_{w,n}$, however, does not possess this property. Meanwhile, $\mathbb{K}_n^h$ only has nonzero entries in the top-left $h \times h$ submatrix. Close scrutiny of either (B.13) or (B.14) reveals that the two facts combined lead to a smaller approximation error.

Step 5. (Auxiliary: Proof of (B.14)) Introduce short-hand notation:

$$A(1)_j \coloneqq (\Omega_n^{-1} Y_n)_j, \quad A(2)_{ij} \coloneqq (\Omega_n^{-1})_{i,j} - (\Omega_n^{-1})_{i,j+1}, \quad \text{and} \quad A(3)_{ij} \coloneqq A(2)_{ij} - A(2)_{i+1,j}.$$

In view of the notation, plus the definition $\mathbb{K}_n^h$, we conclude

$$\begin{aligned} Y_n^\intercal \Omega_n^{-1} \mathbb{K}_n^h \Omega_n^{-1} \mathbb{K}_n^l \Omega_n^{-1} Y_n = \sum_{i=1}^{h-1} \sum_{j=1}^{l-1} A(1)_i A(1)_j A(3)_{h-i,h-j} - A(1)_h \sum_{j=1}^{l-1} A(1)_j A(2)_{1,l-j} \\ - A(1)_l \sum_{i=1}^{h-1} A(1)_i A(2)_{1,h-j} + A(1)_h A(1)_l (\Omega_n^{-1})_{11}. \end{aligned} \tag{B.15}$$

We hence deduce (B.14) by applying the Cauchy-Schwarz inequality and the following four estimates to (B.15):

$$\left. \begin{aligned} &\mathbb{E}(A(1)_{h_n}^2) \lesssim h_n, \quad \mathbb{E}((A(1)_{h_n} - A(1)_{h_n+1})^2) \lesssim 1, \\ &|(\Omega_n^{-1})_{1h_n}| \lesssim 1, \quad |A(3)_{h_n,l_n}| \lesssim n^{-1/2} + \mathbb{1}_{\{|h_n - l_n| \le 2\}}. \end{aligned} \right\}$$

All the estimates are direct results of Lemma D3. Note $A(3)_{i,j}$ is a linear combination of four entries of Ω_n^{-1}. Due to such a combination, for $|h_n - l_n| \ge 3$, the magnitude of $A(3)_{h_n,l_n}$ is reduced by a factor of n, compared to $(\Omega_n^{-1})_{i,j} \sim \sqrt{n}$. This combination originates from the special structure in (B.12). ∎

Lemma B2. *(A.17) holds.*

Proof. All the contents in the Proof of Lemma B1 until (B.2) stay valid if we replace Σ_n with Ω_n, Ω_n with $\Omega_{D,n}$, and (A.16) with (A.17). Unlike the situation there, we do know the analytical expressions of both Ω_n^{-1} and $\Omega_{D,n}^{-1}$ as given by Lemma D3. Note $\Omega_{D,n}$, introduced in Part 2 of Section A.1, is a block-diagonal matrix and we apply Lemma D3 to each block. Instead of (B.3), we use

$$\Omega_n^{-1} = \Omega_{D,n}^{-1} - \Omega_n^{-1} R_n \Omega_{D,n}^{-1} \quad \text{and} \quad R_n := \Omega_n - \Omega_{D,n}.$$

The justification for doing so is the same as the one mentioned in the last paragraph in Step 4 of the Proof of Lemma B1. Indeed, by definition R_n here only has nonzero entries near the top-left or right-bottom corners of the blocks which $\Omega_{D,n}$ consists of. According to Lemma D3, $(\Omega_{D,n}^{-1})_{i,j}$ shrinks when either i or j appoaches the borders of those blocks. Moreover, locally R_n also maintains a structure similar to (B.12). See the comment at the end of the Proof of Lemma B1. One hence can prove (A.17) following the same reasoning of the Proof of Lemma B1. Note we shall skip Steps 2 and 3 there, as we know both $\Omega_{D,n}^{-1}$ and Ω_n^{-1}. ∎

Lemma B3. *(A.18) holds.*

Proof. Step 1. (Main Proof) As in the context in which (A.18) appears, we drop the argument q_n of $\varsigma^{(n)}(q_n)$ whenever there is no ambiguity. In Step 4 of Section A.4, we point out the contruction of $\Xi_{D,n}$ is to allow the application of central limit theorems for triangular arrays. Now we explicitly write $\Xi_{D,n}(\varsigma^{(n)})_1$ as a partial sum of a triangular array.

First, for all $\varsigma \in \Pi^{\varsigma}(q_n)$ we define $\Omega_n(\varsigma)$ and $\Omega_{D,n}(\varsigma)$ by reparameterizing $\Omega_n(\sigma^2, \iota^2, \theta)$ and $\Omega_{D,n}(\sigma^2, \iota^2, \theta)$. Formally, let $\Omega_n(\varsigma) = \Omega_n(\sigma^2, \iota^2, \theta)$ and $\Omega_{D,n}(\varsigma) = \Omega_{D,n}(\sigma^2, \iota^2, \theta)$ with $\iota^2 = \varsigma_2$, $\theta = (\varsigma_3, \ldots, \varsigma_{q_n+2})^\intercal$ and $\sigma^2 = \varsigma_1 n^{1/4}$. Recall $\Omega_n(\sigma^2, \iota^2, \theta)$ is introduced in Part 2 of Section A.1. Note $n^{-1/4}\frac{\partial}{\partial \varsigma_1}\Omega_n(\varsigma) = n^{-1/4}\frac{\partial}{\partial \varsigma_1}\Omega_{D,n}(\varsigma) = \Delta_n$. In view of them, plus the definitions of $\Xi_{D,n}(\varsigma)$ and $\bar{\Xi}_n(\varsigma)$ given by (A.7), (A.6) and Part 3 of Section A.1, we write

$$\Xi_{D,n}(\varsigma^{(n)})_1 - \bar{\Xi}_n(\varsigma^{(n)})_1 = -\frac{\Delta_n}{2n^{3/4}} \mathrm{tr}(\Omega_{D,n}(\varsigma^{(n)})^{-2}(Y_n Y_n^\intercal - \Omega_n^Y)). \tag{B.16}$$

The block-diagonal structure of $\Omega_{D,n}$ indicates we can write the RHS of (B.16) as a sum over those blocks. To do this formally, we introduce for all $1 \leq j \leq J_d + 1$,

$$\begin{cases} \mathcal{U}_n(j) = \mathrm{tr}(\Omega_{n(j)}(\varsigma^{(n)})^{-2} Y_n(j) Y_n(j)^\intercal) \\ \overline{\mathcal{U}}_n(j) = \mathrm{tr}(\Omega_{n(j)}(\varsigma^{(n)})^{-2} \Omega_n^Y(j)) \end{cases} \quad \text{with} \quad n(j) = n_d + \delta_{j, J_d+1}(n_d' - n_d). \tag{B.17}$$

Here $Y_n(j)$ and $\Omega_n^Y(j)$ are defined by (B.31) and (B.32) in Step 4 below. Recall we define J_d, n_d and n_d' in Part 2 of Section A.1 when introducing $\Omega_{D,n}$. In view of (B.17), we can rewrite (B.16) as

$$\Xi_{D,n}(\varsigma^{(n)})_1 - \bar{\Xi}_n(\varsigma^{(n)})_1 = \frac{\Delta_n}{2n^{3/4}} \sum_{j=1}^{J_d+1} (\mathcal{U}_n(j) - \overline{\mathcal{U}}_n(j)). \tag{B.18}$$

Step 2 below proves that the RHS of (B.18), multiplied by $n^{1/2}$, converges stably in law to the RHS of (A.18). Hence (A.18) follows from $\bar{\Xi}_n(\varsigma^{(n)})_1 = o_{\mathrm{P}}(n^{-1/2})$. Indeed, in view of the definitions of $\bar{\Xi}_n(\varsigma)$ and $\bar{\Xi}_n^\star$ given by (A.7), (A.6) and Part 3 of Section A.1, plus $n^{-1/4}\frac{\partial}{\partial\varsigma_1}\Omega_{D,n}(\varsigma) = \Delta_n$, we conclude

$$\bar{\Xi}_n(\varsigma^{(n)})_1 = \frac{\Delta_n}{2n^{3/4}}\mathrm{tr}(\Omega_n^{-1}) - \frac{\Delta_n}{2n^{3/4}}\mathrm{tr}(\Omega_{D,n}^{-1}\Omega_n^Y) = \bar{\Xi}_n^\star(\varsigma^{(n)})_1 + o_{\mathrm{P}}(n^{-1/2}) = o_{\mathrm{P}}(n^{-1/2}). \tag{B.19}$$

Here we drop the arguments $\varsigma^{(n)}$ of Ω_n and $\Omega_{D,n}$. One can deduce the second equality in (B.19) using Lemmas D3 and D6. The last equality in (B.19) comes from the fact that $\bar{\Xi}_n^\star(\varsigma^{(n)}(q_n))_1 \lesssim n^{-1/4}\left|\frac{d}{d\sigma^2}\widetilde{h}_n(\sigma_n^2(q_n), \zeta^{(n)}(q_n)^2)\right| = 0$ due to the definitions of $\bar{\Xi}_n^\star$, $\widetilde{h}_n$, $\varsigma^{(n)}(q_n)$ and $\sigma_n^2(q_n)$. They are all given in Step 1 of Section A.4.

Step 2. (Stable convergence) This step proves under $(\zeta^{(n)})^2 \to b^2 \in (0,\infty)$,

$$\frac{\Delta_n}{2n^{1/4}}\sum_{j=1}^{J_d+1}(\mathcal{U}_n(j) - \overline{\mathcal{U}}_n(j)) \xrightarrow{\mathcal{L}-s} \mathcal{N}\left(0, \frac{5\sqrt{T}}{64b}\frac{C(4)_T}{C_T^{7/2}} + \frac{3\sqrt{T}}{64}\frac{1}{C_T^{3/2}b}\right). \tag{B.20}$$

Stochastic volatility of the price process X_t complicates the analysis. We first show that we are able to replace the stochastic volatility process σ_t with a piecewise constant process, in that the asymptotic behavior of the LHS of (B.20) would not be altered. Then we prove the stable central limit theorem for the piecewise constant volatility process.

Now we formalize the aforementioned procedure. Define for all $1 \le j \le J_d + 1$,

$$\begin{cases} \mathcal{V}_n(j) = \mathrm{tr}(\Omega_{n(j)}(\varsigma^{(n)})^{-2}Y_n^C(j)Y_n^C(j)^\intercal) \\ \overline{\mathcal{V}}_n(j) = \mathrm{tr}(\Omega_{n(j)}(\varsigma^{(n)})^{-2}\Omega_n^{Y,C}(j)) \end{cases} \quad \text{with} \quad n(j) = n_d + \delta_{j,J_d+1}(n_d' - n_d). \tag{B.21}$$

Here $Y_n^C(j)$ and $\Omega_n^{Y,C}(j)$ are defined by (B.31) and (B.32) in Step 4 below based on Y_n^C and $\Omega_n^{Y,C}$. Note Y_n^C and $\Omega_n^{Y,C}$ are specifically constructed to be the counterparts of Y_n and Ω_n^Y under piecewise constant volatility. See Part 2 of Section A.1 for the definitions. Step 3 below proves replacing X_t with the piecewise-constant volatility counterpart has an asymptotically negligible effect:

$$\frac{\Delta_n}{2n^{1/4}}\sum_{j=1}^{J_d+1}\underbrace{(\mathcal{U}_n(j) - \overline{\mathcal{U}}_n(j) - \mathcal{V}_n(j) + \overline{\mathcal{V}}_n(j))}_{=:\mathcal{R}_b(j)} = o_{\mathrm{P}}(1). \tag{B.22}$$

Hence, we only need to show $\frac{\Delta_n}{2n^{1/4}}\sum_{j=1}^{J_d+1}(\mathcal{V}_n(j) - \overline{\mathcal{V}}_n(j))$ converges stably in law to the RHS of (B.20). Let $\mathcal{G}_j = \sigma(U_i : i \le j)$ be the pre-σ field of the noise sequences at time t_j. In view of Theorem 2.2.15 in Jacod and Protter (2011), plus denoting $\mathcal{F}(j) = \mathcal{F}_{\tau(n,j-1)} \otimes \mathcal{G}_{(j-1)n_d}$ and $\check{\delta}_j = 1 + \delta_{j,J_d+1}\left(\frac{n_d}{n_d'} - 1\right)$, the stable convergence of $\frac{\Delta_n}{2n^{1/4}}\sum_{j=1}^{J_d+1}(\mathcal{V}_n(j) - \overline{\mathcal{V}}_n(j))$ follows from that under $(\zeta^{(n)})^2 \to b^2 \in (0,\infty)$ and for all $1 \le j \le J_d + 1$,

$$J_d\frac{\Delta_n}{2n^{1/4}}\mathbb{E}[\mathcal{V}_n(j) - \overline{\mathcal{V}}_n(j)|\mathcal{F}(j)] = o_{\mathrm{P}}(1); \tag{B.23}$$

$$\check{\delta}_j J_d\left(\frac{\Delta_n}{2n^{1/4}}\right)^2\mathbb{E}[(\mathcal{V}_n(j)-\overline{\mathcal{V}}_n(j))^2|\mathcal{F}(j)]\xrightarrow{\mathrm{P}}\underbrace{\frac{\sqrt{T}}{64b}\left(\frac{5\sigma^4_{\tau(n,j-1)}}{C_T^{7/2}}+\frac{2\sigma^2_{\tau(n,j-1)}}{C_T^{5/2}}+\frac{1}{C_T^{3/2}}\right)}_{:=\mathcal{R}_a(j)};\tag{B.24}$$

$$J_d\left(\frac{\Delta_n}{2n^{1/4}}\right)^4\mathbb{E}[(\mathcal{V}_n(j)-\overline{\mathcal{V}}_n(j))^4|\mathcal{F}(j)]=o_{\mathrm{P}}(1);\tag{B.25}$$

and for process $M=W$ and for all processes M in a set $\mathcal{M}$ of bounded martingales which are orthogonal to W and such that the family $(M_\infty : M\in\mathcal{M})$ is total in the space $\mathbb{L}^1(\Omega,\mathcal{F},\mathrm{P})$

$$J_d\frac{\Delta_n}{2n^{1/4}}\mathbb{E}[\mathcal{V}_n(j)(M_{\tau(n,j)}-M_{\tau(n,j-1)})|\mathcal{F}(j)]=o_{\mathrm{P}}(1).\tag{B.26}$$

Recall we define $\tau(n,j)$ in Part 2 of Section A.1. One can verify (B.23) using Itô isometry and Lemma D3. In view of Lemmas D3 and D6, plus (B.21), we conclude (B.24) from:

$$\begin{aligned}&\check{\delta}_j J_d\left(\frac{\Delta_n}{2n^{1/4}}\right)^2\mathbb{E}[(\mathcal{V}_n(j)-\overline{\mathcal{V}}_n(j))^2|\mathcal{F}(j)]\\ =\;& o_{\mathrm{P}}(1)+2J_d\left(\frac{\Delta_n}{2n^{1/4}}\right)^2\mathrm{tr}(\Omega_{n'_d}(\varsigma^{(n)})^{-2}\Omega_n^{Y,C}(j)\Omega_{n'_d}(\varsigma^{(n)})^{-2}\Omega_n^{Y,C}(j))=o_{\mathrm{P}}(1)+\mathcal{R}_a(j).\end{aligned}$$

We derive (B.25) using the same reasoning. A slight modification of Step 4 of the proof of Lemma 5.7 in Jacod, Li, Mykland, Podolskij, and Vetter (2009) yields (B.26).

Step 3. (Proof of (B.22)) Recall we define $\mathcal{R}_b(j)$ in (B.22). In view of the fact that $\mathcal{R}_b(j)$ is $\mathcal{F}(k)$-measurable for all $k\geq j+1$, we deduce (B.22) from that for all $1\leq j\leq J_d+1$,

$$\mathbb{E}[\mathcal{R}_b(j)|\mathcal{F}(j)]=o_{\mathrm{P}}\left(J_d^{-1}\left(\frac{\Delta_n}{2n^{3/4}}\right)^{-1}\right)\quad\text{and}\quad\mathbb{E}[\mathcal{R}_b(j)^2|\mathcal{F}(j)]=o_{\mathrm{P}}\left(J_d^{-1}\left(\frac{\Delta_n}{2n^{3/4}}\right)^{-2}\right).\tag{B.27}$$

The first claim in (B.27) follows from Itô isometry and Lemma D3. Now we prove the second claim in (B.27). We only consider $1\leq j\leq J_d$ to simplify the exposition. Define for all $1\leq j\leq J_d$,

$$\Delta_i^n X(j)=\Delta^n_{(j-1)n_d+i}X,\quad \Delta_i^n W(j)=\Delta^n_{(j-1)n_d+i}W,\quad \sigma_C(j)=\sigma_{C,\tau(n,j-1)},\quad \Delta_i^n U(j)=\Delta^n_{(j-1)n_d+i}U,$$

$$\begin{aligned}\mathcal{R}_{b1}(j)&=\sum_{i=1}^{n_d}\Omega_{n_d}(\varsigma^{(n)})^{-2}_{ii}((\Delta_i^n X(j))^2-\Omega_n^B(j)_{ii}),\\ \mathcal{R}_{b2}(j)&=\sum_{i=1}^{n_d}\Omega_{n_d}(\varsigma^{(n)})^{-2}_{ii}((\sigma_C(j)\Delta_i^n W(j))^2-\Omega_n^C(j)_{ii}),\\ \mathcal{R}_{b3}(j)&=\sum_{k=1}^{n_d}\sum_{i=1}^{n_d}\Omega_{n_d}(\varsigma^{(n)})^{-2}_{ik}(\Delta_i^n X(j)-\sigma_C(j)\Delta_i^n W(j))\Delta_k^n U(j),\\ \mathcal{R}_{b4}(j)&=\sum_{k=1}^{n_d}\sum_{i=k+1}^{n_d}\Omega_{n_d}(\varsigma^{(n)})^{-2}_{ik}(\Delta_i^n X(j)-\sigma_C(j)\Delta_i^n W(j))(\Delta_k^n X(j)-\sigma_C(j)\Delta_k^n W(j)),\end{aligned}$$

$$\mathcal{R}_{b5}(j) = \sum_{k=1}^{n_d} \sum_{i=k+1}^{n_d} \Omega_{n_d}(\varsigma^{(n)})^{-2}_{ik}(\Delta_i^n X(j) - \sigma_C(j)\Delta_i^n W(j))\sigma_C(j)\Delta_k^n W(j), \tag{B.28}$$

$$\mathcal{R}_{b6}(j) = \sum_{k=1}^{n_d} \sum_{i=k+1}^{n_d} \Omega_{n_d}(\varsigma^{(n)})^{-2}_{ik}\sigma_C(j)\Delta_i^n W(j)(\Delta_k^n X(j) - \sigma_C(j)\Delta_k^n W(j)).$$

One can verify

$$\mathcal{R}_b(j) = \mathcal{R}_{b1}(j) + \mathcal{R}_{b2}(j) + 2\mathcal{R}_{b3}(j) + 2\mathcal{R}_{b4}(j) + \mathcal{R}_{b5}(j) + \mathcal{R}_{b6}(j).$$

Hence, by the triangle inequality, it suffices to show for all $l \in \{1, 2, 3, 4, 5, 6\}$,

$$\mathbb{E}[\mathcal{R}_{bl}(j)^2|\mathcal{F}(j)] = o_{\mathrm{P}}\left(J_d^{-1}\left(\frac{\Delta_n}{2n^{3/4}}\right)^{-2}\right). \tag{B.29}$$

We only show the analysis of $\mathcal{R}_{b5}(j)$, while the others only require simpler analysis. Introduce short-hand notation:

$$\left.\begin{aligned} \mathcal{R}_{c1}(j,i) &= 2\int_{t^n_{jn_d+i-1}}^{t^n_{jn_d+i}} (\sigma_s - \sigma_C(j))^2 ds + 2\left(\int_{t^n_{jn_d+i-1}}^{t^n_{jn_d+i}} |\mu_s| ds\right)^2, \\ \mathcal{R}_{c2}(j,i,k) &= \Omega_{n_d}(\varsigma^{(n)})^{-2}_{ik}\sigma_C(j)\Delta_k^n W(j). \end{aligned}\right\} \tag{B.30}$$

In view of (B.28) and (B.30), plus the Cauchy-Schwarz inequality and the continuity of σ_s, we deduce (B.29) for $l = 5$ from

$$\begin{aligned} \mathbb{E}[\mathcal{R}_{b5}(j)^2|\mathcal{F}(j)] &\leq \sum_{i=1}^{n_d} \sqrt{\mathbb{E}[\mathcal{R}_{c1}(j,i)^2|\mathcal{F}(j)]}\sqrt{\mathbb{E}\left[\left(\sum_{k=1}^{i-1}\mathcal{R}_{c2}(j,i,k)\right)^4 \middle| \mathcal{F}(j)\right]} \\ &\lesssim \sum_{i=1}^{n_d} o_{\mathrm{P}}(1) n_d \Delta_n \times \sigma_C^4(j)\left(\sum_{k=1}^{i-1}(\Omega_{n_d}(\varsigma^{(n)})^{-2}_{ik})^2\Delta_n\right)^2 = o_{\mathrm{P}}\left(J_d^{-1}\left(\frac{\Delta_n}{2n^{3/4}}\right)^{-2}\right). \end{aligned}$$

Here the last equality comes from Lemma D3 and the definitions of n_d and J_d given in Part 2 of Section A.1.

Step 4. (Auxiliary notation) This step defines some notation we use in the proof. For all $A \in \{Y_n, Y_n^C\}$, we partition A by defining $A(j) \in \mathbb{R}^{n_d}$ for all $1 \leq j \leq J_d$ and $A(J_d + 1) \in \mathbb{R}^{n'_d}$ as

$$\left.\begin{aligned} A(j)_k &= A_{(j-1)n_d+k}, \forall 1 \leq j \leq J_d, 1 \leq k \leq n_d, \\ A(J_d+1)_k &= A_{n_d J_d + k}, \forall 1 \leq k \leq n'_d. \end{aligned}\right\} \tag{B.31}$$

For all $A \in \{\Omega_n^Y, \Omega_n^{Y,C}\}$, we partition A by defining $A(j) \in \mathcal{M}_{n_d}$ for all $1 \leq j \leq J_d$ and $A(J_d + 1) \in$

$\mathcal{M}_{n_d'}$ as

$$\left.\begin{aligned} &A(j)_{kl} = A_{(j-1)n_d+k,(j-1)n_d+l}, \forall 1 \le j \le J_d, 1 \le k, l \le n_d, \\ &A(J_d+1)_{kl} = A_{n_d J_d+k, n_d J_d+l}, \forall 1 \le k, l \le n_d. \end{aligned}\right\} \tag{B.32}$$

∎

Lemma B4. *The second claim in (A.22) holds.*

Proof. Step 1. (Main proof) Recall the definitions of $\bar{\varsigma}_n$ and $\partial\Xi_n(\bar{\varsigma}_n)$ are given after (A.13). We hence construct a corresponding family of parameter spaces indexed by q and α:

$$\bar{\Pi}_n^{\varsigma}(q,\alpha) = \left\{\varsigma = (\varsigma(1), \ldots, \varsigma(q+2)) : \varsigma(i) \in \Pi_n^{\varsigma}(q), \|\varsigma(i) - \varsigma^{(n)}\|_{(1)}^2 \vee |\varsigma(i)_1 n^{1/4} - C_T| \le \alpha, \forall i\right\}.$$

Namely, an element of $\bar{\Pi}_n^{\varsigma}(q,\alpha)$ is a $(q+2)\times(q+2)$ matrix and each column of the matrix belongs to a subset of $\Pi_n^{\varsigma}(q)$. Indeed, $\bar{\varsigma}_n$ constructed in Step 3 of Section A.4 satisfies $\bar{\varsigma}_n \in \bar{\Pi}_n^{\varsigma}(q_n, \infty)$. Provided $\widehat{\mathcal{R}}_n(q_n) = o_{\mathrm{P}}(1)$ and $\mathcal{R}^{(n)}(q_n) = o_{\mathrm{P}}(1)$, we have $\bar{\varsigma}_n \in \bar{\Pi}_n^{\varsigma}(q_n, \alpha_n)$ for some $\alpha_n = o_{\mathrm{P}}(1)$. Moreover, for all $\varsigma \in \bar{\Pi}_n^{\varsigma}(q,\alpha)$, we somewhat abuse the notation and define $\partial\bar{\Xi}_n^{\star}(\varsigma) \in \mathcal{M}_q$ by $\partial\bar{\Xi}_n^{\star}(\varsigma)_{ij} = \partial\bar{\Xi}_n^{\star}(\varsigma(i))_{ij}$. Recall the RHS is defined by (A.8) as $\varsigma(i) \in \Pi_n^{\varsigma}(q)$. We hence deduce the second claim in (A.22), by using the triangle inequality, from that for all $\alpha_n \to 0$,

$$\sup_{\varsigma \in \bar{\Pi}_n^{\varsigma}(q_n,\alpha_n)} \|\partial\Xi_n(\varsigma) - \partial\bar{\Xi}_n^{\star}(\varsigma)\| = o_{\mathrm{P}}(1) \quad \text{and} \quad \sup_{\varsigma \in \bar{\Pi}_n^{\varsigma}(q_n,\alpha_n)} \|\partial\bar{\Xi}_n^{\star}(\varsigma) - \partial\bar{\Xi}_n^{\star\star}(q_n)\| = o_{\mathrm{P}}(1). \tag{B.33}$$

Step 2 below proves the second claim in (B.33). As for the first claim, we deduce the uniform convergence here from the stochastic equicontinuity $\sup_{\varsigma \in \bar{\Pi}_n^{\varsigma}(q_n,\alpha_n)} \|\partial\Xi_n(\varsigma) - \partial\Xi_n(\varsigma^{(n)})\| = o_{\mathrm{P}}(1)$ and the pointwise convergence $\|\partial\Xi_n(\varsigma^{(n)}) - \partial\bar{\Xi}_n^{\star}(\varsigma^{(n)})\| = o_{\mathrm{P}}(1)$. The challenge of proving the former is that both the dimension of the parameter space $\bar{\Pi}_n^{\varsigma}(q_n,\alpha_n)$ and that of $\partial\Xi_n(\varsigma)$ are q_n+2, which can grow as n increases. The method to solve this issue is the same as the one Step 2 uses. On the other hand, we can derive the pointwise convergence from repeating Step 4 of Section A.4, hence the first claim in (B.33).

Step 2. (Proof of the second claim in (B.33)) Throughout the rest of the proof, all the results hold uniformly over $\varsigma \in \bar{\Pi}_n^{\varsigma}(q_n,\alpha_n)$ for some $\alpha_n \to 0$. We hence do not keep the supremum over ς unless necessary. Moreover, we introduce simplifying notation $\mathcal{R}_a(\varsigma) = \partial\bar{\Xi}_n^{\star}(\varsigma) - \partial\bar{\Xi}_n^{\star\star}(q_n)$. Our goal is to show $\sup_{x\in\mathbb{R}^{q_n+2}:\|x\|=1} |x^{\intercal}\mathcal{R}_a(\varsigma)x| = o_{\mathrm{P}}(1)$. In the interest of space, we only analyze the contribution of the $q_n \times q_n$ bottom-right submatrix of $\mathcal{R}_a(\varsigma)$. Contributions from the other terms only require simpler analysis. We denote the submatrix by $\mathcal{R}_b(\varsigma)$ and we intend to prove

$$\sup_{x\in\mathbb{R}^{q_n}:\|x\|=1} |x^{\intercal}\mathcal{R}_b(\varsigma)x| = o_{\mathrm{P}}(1). \tag{B.34}$$

Writing $\iota^2(i) = \varsigma(i)_2$ and $\theta(i) = (\varsigma(i)_3, \varsigma(i)_4, \ldots, \varsigma(i)_{q_n+2})^{\intercal}$, plus the boundedness of $\iota^{(n)}$, the definition of $\partial\bar{\Xi}_n^\star$ given by (A.8) and Cauchy-Schwarz, we conclude

$$|x^\intercal \mathcal{R}_b(\varsigma)x| \lesssim \underbrace{\left| \sum_{1\le i,j\le q_n} x_i x_j \int_{-\pi}^{\pi} g(\lambda;\theta^{(n)}) \frac{\partial^2}{\partial\theta_i\partial\theta_j}\left(\frac{(\iota^2(i))^{-1}}{g(\lambda;\theta(i))} - \frac{(\iota^{-2})^{(n)}}{g(\lambda;\theta^{(n)})}\right) d\lambda \right|}_{=:|x^\intercal \mathcal{R}_c(\varsigma)x|} + o_{\mathrm{P}}(\|x\|^2). \quad \text{(B.35)}$$

To bound $|x^\intercal \mathcal{R}_c(\varsigma)x|$, we introduce short-hand notation

$$\mathcal{Q}_n(1,j,\lambda) = \frac{g(\lambda;\theta^{(n)})}{\iota^2(j)g^2(\lambda;\theta(j))} - \frac{g(\lambda;\theta^{(n)})}{(\iota^{(n)})^2 g^2(\lambda;\theta^{(n)})}, \quad \alpha_n(1,j,l) = \int_{-\pi}^{\pi} \mathcal{Q}_n(1,j,\lambda)e^{\mathrm{i}l\lambda}d\lambda.$$
$$\mathcal{Q}_n(2,j,\lambda) = \frac{g_-^2(\lambda,\theta(j))g(\lambda;\theta^{(n)})}{\iota^2(j)g^3(\lambda;\theta(j))} - \frac{g_-^2(\lambda,\theta^{(n)})g(\lambda;\theta^{(n)})}{(\iota^{(n)})^2 g^3(\lambda;\theta^{(n)})}, \quad \alpha_n(2,j,l) = \int_{-\pi}^{\pi} \mathcal{Q}_n(2,j,\lambda)e^{\mathrm{i}l\lambda}d\lambda. \quad \text{(B.36)}$$

In view of the definition of $|x^\intercal \mathcal{R}_c(\varsigma)x|$ given in (B.35) and the notation (B.36), plus the triangle inequality and the Cauchy-Schwarz inequality, we conclude

$$\begin{aligned}
|x^\intercal \mathcal{R}_c(\varsigma)x| &\le 2\left|\sum_{1\le j\le q_n}\sum_{1\le j+l\le q_n} x_j x_{j+l}\alpha_n(1,j,l)\right| + 2\left|\sum_{1\le j\le q_n}\sum_{1\le l-j\le q_n} x_j x_{l-j}\alpha_n(2,j,l)\right| \\
&\le 2\|x\|^2 \sup_j \sum_{|l|\le q_n} |\alpha_n(1,j,l)| + 2\|x\|^2 \sup_j \sum_{0\le l\le 2q_n} |\alpha_n(2,j,l)|.
\end{aligned}$$

Step 3 proves both $\sup_j \sum_{|l|\le q_n} |\alpha_n(1,j,l)|$ and $\sup_j \sum_{0\le l\le 2q_n} |\alpha_n(2,j,l)|$ converge in probability to zero, hence the desired result (B.34).

Step 3. (Auxiliary: Bounds on $\alpha_n(1,j,l)$ and $\alpha_n(2,j,l)$) In this step we prove $\sup_j \sum_{|l|\le q_n} |\alpha_n(1,j,l)| = o_{\mathrm{P}}(1)$ and $\sup_j \sum_{0\le l\le 2q_n} |\alpha_n(2,j,l)| = o_{\mathrm{P}}(1)$. We use properties of Fourier transform. Concretely, recalling (B.36), we define for $p \in \{1,2\}$,

$$\widetilde{\mathcal{Q}}(p,j,\lambda,l) = \mathcal{Q}_n\left(p,j,\lambda+\frac{\pi}{l}\right) + \mathcal{Q}_n\left(p,j,\lambda-\frac{\pi}{l}\right) - 2\mathcal{Q}_n(p,j,\lambda).$$

Then the periodicity of $e^{\mathrm{i}l\lambda}$ indicates

$$\alpha_n(p,j,l) = -\frac{1}{4}\int_{-\pi}^{\pi} \widetilde{\mathcal{Q}}(p,j,\lambda,l)e^{\mathrm{i}l\lambda}d\lambda. \quad \text{(B.37)}$$

On the other hand, in view of the definition of $\widetilde{\mathcal{Q}}$, plus recalling $\|\cdot\|_{1,(1)}$defined by (A.1), one can verify

$$\max_{p\in\{1,2\}} |\widetilde{\mathcal{Q}}(p,j,\lambda,l)| \quad \lesssim \quad \|\varsigma(j)-\varsigma^{(n)}\|_{1,(1)} \sum_{k=1}^{q_n}(|\theta(j)_k| + |\theta_k^{(n)}|)\left|e^{\mathrm{i}k\left(\lambda+\frac{\pi}{l}\right)} + e^{\mathrm{i}k\left(\lambda-\frac{\pi}{l}\right)} - 2e^{\mathrm{i}k\lambda}\right|$$

$$+\sum_{k=1}^{q_n}|\theta(j)_k-\theta_k^{(n)}|\left|e^{\mathrm{i}k(\lambda+\frac{\pi}{l})}+e^{\mathrm{i}k(\lambda-\frac{\pi}{l})}-2e^{\mathrm{i}k\lambda}\right| \tag{B.38}$$
$$=: \quad \|\varsigma(j)-\varsigma^{(n)}\|_{1,(1)}\mathcal{R}_d(j,l)+\mathcal{R}_e(j,l).$$

Given (B.37) and (B.38), the bounds on $\alpha_n(1,j,l)$ and $\alpha_n(2,j,l)$ follow from those on $\|\varsigma(j)-\varsigma^{(n)}\|_{1,(1)}$, $\mathcal{R}_d(j,l)$, and $\mathcal{R}_e(j,l)$. First, the combination of the definition of $\Theta(q)$, Condition 1, $\alpha_n\to 0$ and the definition of $\bar{\Pi}_n^{\varsigma}(q_n,\alpha_n)$ indicates $\sup_j\|\varsigma(j)-\varsigma^{(n)}\|_{1,(1)}=o_{\mathrm{P}}(1)$. The same combination, plus the constructions of $\mathcal{R}_d$ and $\mathcal{R}_e$, also lead to

$$\sup_j\sum_{l=1}^{q_n}\mathcal{R}_d(j,l)\lesssim\sup_j\sum_{l=1}^{q_n}\sum_{k=1}^{q_n}(|\theta(j)_k|+|\theta_k^{(n)}|)\left(\frac{k^2}{l^2}\wedge 1\right)\lesssim\sup_j\sum_{k=1}^{q_n}k(|\theta(j)_k|+|\theta_k^{(n)}|)\lesssim 1,$$
$$\text{and}\quad \sup_j\sum_{l=1}^{q_n}\mathcal{R}_e(j,l)\lesssim\sup_j\sum_{k=1}^{q_n}k|\theta(j)_k-\theta_k^{(n)}|=o_{\mathrm{P}}(1).$$

We readily deduce $\sup_j\sum_{|l|\le q_n}|\alpha_n(1,j,l)|=o_{\mathrm{P}}(1)$ and $\sup_j\sum_{0\le l\le 2q_n}|\alpha_n(2,j,l)|=o_{\mathrm{P}}(1)$. ∎

Lemma B5. *(A.23) holds.*

Proof. In view of the triangle inequality, (A.23) follows from that for all $q_n\lesssim n^{1/2}$,

$$\sup_{\varsigma\in\Pi_n^{\varsigma}(q_n)}|L_n(\varsigma)-L_{A,n}(\varsigma)|=o_{\mathrm{P}}(n)\quad\text{and}\quad\sup_{\varsigma\in\Pi_n^{\varsigma}(q_n)}|L_{A,n}(\varsigma)-\bar{L}_n^{\star}(\varsigma)|=o_{\mathrm{P}}(n). \tag{B.39}$$

Recall we define $L_{A,n}(\varsigma)$ in (A.6). We can prove the first claim in (B.39) by repeating the proof of Lemma B1. The proof would be simpler here for two reasons. One is that the bound we require here is less sharp. The other is that only Σ_n^{-1} and Ω_n^{-1} themselves are involved as we do not take derivatives here. We hence only prove the second claim in (B.39).

Define $\Omega_n(\varsigma)$ by reparameterizing $\Omega_n(\sigma^2,\iota^2,\theta)$: we let $\Omega_n(\varsigma)=\Omega_n(\sigma^2,\iota^2,\theta)$ with $\iota^2=\varsigma_2$, $\theta=(\varsigma_3,\dots,\varsigma_{q+2})^{\intercal}$ and $\sigma^2=\varsigma_1 n^{1/4}$. So we can write $L_{A,n}(\varsigma)=\mathcal{L}(\Omega_n(\varsigma),Y_nY_n^{\intercal})$ with $\mathcal{L}(\cdot,\cdot)$ defined in Part 3 of Section A.1. In view of the definitions of $\bar{L}_n^{\star}(\varsigma)$ and $h_n(\varsigma)$ given by (A.4) and (A.6), we conclude $\bar{L}_n^{\star}(\varsigma)=-\frac{n}{2}h_n(\varsigma)+o_{\mathrm{P}}(n)$ holds uniformly over $\varsigma\in\Pi_n^{\varsigma}(q_n)$. On the other hand, using Lemma D2 one can verify $|\log\det\Omega_n(\varsigma)-n\log\varsigma_2|=o(n)$ holds uniformly. Hence, letting $\check{h}_n(\varsigma)=h_n(\varsigma)-n\log\varsigma_2$, the second claim in (B.39) follows from

$$\sup_{\varsigma\in\Pi_n^{\varsigma}(q_n)}|Y_n^{\intercal}\Omega_n(\varsigma)^{-1}Y_n-\check{h}_n(\varsigma)|=o_{\mathrm{P}}(n). \tag{B.40}$$

We decompose the proof of the uniform convergence (B.40) into those of pointwise convergence and stochastic equicontinuity, following the same argument used in the proofs of Theorem 2.1 and Corollary 2.2 in Newey (1991). We first prove the pointwise convergence. In view of the triangle inequality, we conclude for all deterministic sequence $\{\varsigma_n \in \Pi_n^{\varsigma}(q_n), n \geq 1\}$,

$$|Y_n^\intercal \Omega_n(\varsigma)^{-1} Y_n - \check{h}_n(\varsigma)| \leq |Y_n^\intercal \Omega_n(\varsigma)^{-1} Y_n - Y_n^\intercal \Omega_{D,n}(\varsigma)^{-1} Y_n| + |Y_n^\intercal \Omega_{D,n}(\varsigma)^{-1} Y_n - h_n(\varsigma)| = o_{\mathrm{P}}(n).$$

The last equality holds by a repetition of the proof of Lemma B1. It would be simpler here for the two reasons mentioned above. The consideration behind introducing $\Omega_{D,n}$ and $L_{A,n}$ above is the same as that discussed in Step 4 of Section A.4.

As for the stochastic equicontinuity, we need that for $A_n(\varsigma) \in \{Y_n^\intercal \Omega_n(\varsigma)^{-1} Y_n, \check{h}_n(\varsigma)\}$,

$$\sup_{\varsigma \in \Pi_n^{\varsigma}(q_n); j \leq q_n+2} \left| (n^{-1/4}\delta_{j,1} + \mathbb{1}_{\{j \neq 1\}}) \frac{\partial A_n(\varsigma)}{\partial \varsigma_j} \right| = O_{\mathrm{P}}(n). \tag{B.41}$$

The additional factor $n^{-1/4}\delta_{j,1}$ compared to Assumption 3A in Newey (1991) arises from the definition of $\Pi_n^{\varsigma}(q_n)$. Slightly different from the compact space setup of in Newey (1991), here the dimension of $\Pi_n^{\varsigma}(q_n)$, $q_n + 2$, can diverge as n grows. However, $\Theta(\infty)$, which the definition of $\Pi_n^{\varsigma}(q_n)$ hinges on, requires the 1-summability of $j\theta_j$ for all $\theta \in \Theta(\infty)$. Also, note that in (B.41) the supremum is also taken over j. These two facts guarantee the argument of Newey (1991) can go through in our setup.

When $A_n(\varsigma) = \check{h}_n(\varsigma)$, verifying (B.41) is trivial as we have the analytical expression of $\check{h}_n(\varsigma)$. Now we prove (B.41) for $A_n(\varsigma) = Y_n^\intercal \Omega_n(\varsigma)^{-1} Y_n$. Let $V_n(\varsigma)$ be the reparameterization of $V_n(\sigma^2, \iota^2, \theta, \Delta_n)$. Formally, let $V_n(\varsigma) = V_n(\sigma^2, \iota^2, \theta, \Delta_n)$ with $\iota^2 = \varsigma_2$, $\theta = (\varsigma_3, \ldots, \varsigma_{q+2})^\intercal$ and $\sigma^2 = \varsigma_1 n^{1/4}$. Then it follows $V_n(\varsigma) = O_n \Omega_n(\varsigma) O_n$. In view of the diagonality of $V_n(\varsigma)$, we write

$$\frac{\partial Y_n^\intercal \Omega_n(\varsigma)^{-1} Y_n}{\partial \varsigma_j} = -\sum_{k=1}^{n} V_n(\varsigma)_{kk}^{-2} \frac{\partial V_n(\varsigma)_{kk}}{\partial \varsigma_j} (O_n Y_n)_k^2.$$

Let $\varsigma_n^* = (n^{-1/4}, 1, 0_{q_n})$. One can verify, using the definition of $V_n(\sigma^2, \iota^2, \theta, \Delta_n)$ given in Part 2 of Section A.1, that uniformly over $1 \leq k \leq n$,

$$\sup_{\varsigma \in \Pi_n^{\varsigma}(q_n)} |V_n(\varsigma)_{kk}^{-1}| \lesssim V_n(\varsigma_n^*)_{kk}^{-1}, \quad \sup_{\varsigma \in \Pi_n^{\varsigma}(q_n)} \left| \frac{\partial V_n(\varsigma)_{kk}}{\partial \varsigma_1} \right| \lesssim \frac{\partial V_n(\varsigma_n^*)_{kk}}{\partial \varsigma_1},$$
$$\text{and} \quad \sup_{\varsigma \in \Pi_n^{\varsigma}(q_n); 2 \leq j \leq q_n+2} \left| \frac{\partial V_n(\varsigma)_{kk}}{\partial \varsigma_j} \right| \lesssim \frac{\partial V_n(\varsigma_n^*)_{kk}}{\partial \varsigma_2}.$$

Thus we deduce the uniform convergence (B.41) from the following pointwise convergence:

$$\sup_{j \in \{1,2\}} \left| (n^{-1/4}\delta_{j,1} + \delta_{j,2}) \frac{\partial Y_n^\intercal \Omega_n(\varsigma_n^*)^{-1} Y_n}{\partial \varsigma_j} \right| = O_{\mathrm{P}}(n),$$

which can be verified using the expression of Ω_n^{-1} given by Lemma D3. The proof ends. ∎

Lemma B6. *(A.26) holds.*

Proof. Step 1. (Main proof) We introduce some notation:

$$\mathcal{R}_a(n,q) = \bar{L}_n^\star(\varsigma_\infty^{(n)}) - \bar{L}_n^\star(\varsigma^{(n)}(q)) \quad \text{and} \quad \mathcal{R}_b(n,j) = \bar{L}_n^\star(\varsigma^{(n)}(q_n^\star)) - \bar{L}_n^\star(\varsigma^{(n)}(\widehat{q}_{n,j})). \tag{B.42}$$

Recall that the quantity $\mathcal{R}_f$ in (A.26) is defined by (A.24). Note we drop the subscript j of $\widehat{q}_{n,j}$ in (A.24). Using (B.42), we can bound $n\mathcal{R}_f$ by $2\mathcal{R}_b(n,j)$, up to $o_{\mathrm{P}}(n^{1/4})$. Formally, for both $j \in \{1,2\}$,

$$\frac{n}{2}\mathcal{R}_f = \mathcal{R}_a(n,\widehat{q}_{n,j}) = \mathcal{R}_b(n,j) + \mathcal{R}_a(n,q_n^\star) \leq \mathcal{R}_b(n,j) + K\|\theta^{(n)}\|^2_{(q_n^\star)} = \mathcal{R}_b(n,j) + o_{\mathrm{P}}(n^{1/4}).$$

Here the inequality holds by the general result (A.27). The last equality holds by the definition of $q_n^\star$ in (A.3) and Condition 1. Recall the notation $\|\cdot\|^2_{(q)}$ is introduced in (A.1). Hence, (A.26) would follow from an appropriate bound on $\mathcal{R}_b(n,j)$. To formalize this, we need more notation:

$$\mathcal{R}_c(n,q) = L_n(\widehat{\varsigma}_n(q)) - L_n(\varsigma^{(n)}(q)) \quad \text{and} \quad \mathcal{R}_d(n,q) = L_n(\varsigma^{(n)}(q)) - \bar{L}_n^\star(\varsigma^{(n)}(q)). \tag{B.43}$$

Indeed, using (B.43), plus the fact that $\widehat{q}_{n,j} = \arg\min_q(q(2\delta_{1,j} + \delta_{2,j}\log n) - 2L_n(\widehat{\varsigma}_n(q)))$ due to the constructions of AIC and BIC, we conclude for $j \in \{1,2\}$,

$$\begin{aligned}\mathcal{R}_b(n,j) &\leq (q_n^\star - \widehat{q}_{n,j})\left(\delta_{1,j} + \frac{1}{2}\delta_{2,j}\log n\right) - \mathcal{R}_c(n,q_n^\star) + \mathcal{R}_c(n,\widehat{q}_{n,j}) - \mathcal{R}_d(n,q_n^\star) + \mathcal{R}_d(n,\widehat{q}_{n,j}) \\ &\leq \frac{q_n^\star - \widehat{q}_{n,j}}{2}(\delta_{1,j} + \delta_{2,j}\log n - \delta_{2,j}) + o_{\mathrm{P}}(|q_n^\star - \widehat{q}_{n,j}|(\delta_{1,j} + \delta_{2,j}\log n) + \log n).\end{aligned} \tag{B.44}$$

Step 2 below proves the last equality. We already deduce (A.26).

Step 2. (Auxiliary: Proof of the last inequality in (B.44)) Recall the notation $\widehat{\mathcal{R}}_n(q_n)$ and $\mathcal{R}^{(n)}(q_n)$ introduced at the end of Step 1 of Section A.4. We write

$$\widehat{\mathcal{R}}_n(\widehat{q}_n) = o_{\mathrm{P}}(1), \quad \mathcal{R}^{(n)}(\widehat{q}_n) = o_{\mathrm{P}}(1), \quad \widehat{\mathcal{R}}_n(q_n^\star) = o_{\mathrm{P}}(1), \quad \mathcal{R}^{(n)}(q_n^\star) = o_{\mathrm{P}}(1). \tag{B.45}$$

Step 8 of Section A.4 already proves the first two equations. Proving the last two is simpler as we know *a priori* the property of $q_n^\star$. Step 3 below shows for all $\{q_n\}$ satisfying $\widehat{\mathcal{R}}_n(q_n) = o_{\mathrm{P}}(1)$ and $\mathcal{R}^{(n)}(q_n) = o_{\mathrm{P}}(1)$,

$$\mathcal{R}_c(n,q_n) = \frac{1}{2}q_n + o_{\mathrm{P}}(q_n + \log n). \tag{B.46}$$

On the other hand, using Lemmas D3 and D1, one can deduce $\mathcal{R}_d(n,\widehat{q}_{n,j}) - \mathcal{R}_d(n,q_n^\star) = o_{\mathrm{P}}(\mathcal{R}_b(n,j))$. Hence we readily derive the last inequality in (B.44).

Step 3. (Auxiliary: Proof of (B.46)) We introduce short-hand notation:

$$\mathcal{R}_e(n,q_n) := \Xi_n(\varsigma^{(n)}(q_n))^\intercal \partial\bar{\Xi}_n^{\star\star}(q_n)^{-1}\Xi_n(\varsigma^{(n)}(q_n)).$$

First of all, we have for some $\bar{\varsigma}_n(q_n) = \alpha_n\widehat{\varsigma}_n(q_n) + (1-\alpha_n)\varsigma^{(n)}(q_n) \in \mathbb{R}^{q_n+2}$ with $0 \le \alpha_n \le 1$,

$$\mathcal{R}_c(n, q_n) = \frac{1}{2}(\widehat{\varsigma}_n(q_n) - \varsigma^{(n)}(q_n))^\intercal \partial\Xi_n(\bar{\varsigma}_n(q_n))(\widehat{\varsigma}_n(q_n) - \varsigma^{(n)}(q_n)),$$

which holds by noting $\frac{\partial}{\partial\varsigma_j}L_n(\widehat{\varsigma}_n(q_n)) = 0$ for all $1 \le j \le q_n + 2$ and the definition of $\partial\Xi_n$. Second, repeating the reasoning of Step 3 and Step 8, respectively, of Section A.4, we deduce that under (B.45),

$$\mathbb{E}(\|\widehat{\varsigma}_n(q_n) - \varsigma^{(n)}(q_n)\|^2) = O(q_n), \quad \|\partial\Xi_n(\bar{\varsigma}_n(q_n)) - \partial\bar{\Xi}_n^{\star\star}(q_n)\| = o_{\mathrm{P}}(1).$$

$$\mathbb{E}(\mathcal{R}_e(n, q_n)) = q_n + o(q_n + \log n), \quad \mathbb{E}(\mathcal{R}_e(n, q_n)^2) = q_n^2 + o(q_n^2 + (\log n)^2).$$

In view of Chebyshev's inequality and the Cauchy-Schwarz inequality, the above results readily yield (B.46). ∎

Appendix C Supplement to Sections A.6 and A.8

In the statements of all lemmas in this section, we use "(xx) holds" to mean "Under the conditions given in the context, (xx) in Section A.6 holds". Moreover, we adopt all the notation introduced in Step 1 of Section A.6.

Lemma C1. *(A.40) holds.*

Proof. Step 1. (Main proof) First, because q is finite, it suffices to prove (A.40) under the assumption that $\bar{\varsigma}_n(i) = \bar{\varsigma}_n(j)$ for all $1 \le i, j \le q + 2$. Further, we somewhat abuse the notation and use $\bar{\varsigma}_n$ to denote $\bar{\varsigma}_n(i)$, which is invariant over i. Hence $\bar{\varsigma}_n \in \mathbb{R}^{q+2}$ now.

Recalling the definition of $\partial\bar{\Xi}_n$ in (A.29), plus the triangle inequality, we deduce $\|\partial\Xi_n(\bar{\varsigma}_n) - \partial\bar{\Xi}_n^\star(\bar{\varsigma}_n)\| = o_{\mathrm{P}}(1)$ from

$$\|\partial\Xi_n(\bar{\varsigma}_n) - \partial\bar{\Xi}_n(\bar{\varsigma}_n)\| = o_{\mathrm{P}}(1) \quad \text{and} \quad \|\partial\bar{\Xi}_n(\bar{\varsigma}_n) - \partial\bar{\Xi}^\star\| = o_{\mathrm{P}}(1). \tag{C.1}$$

Step 4 below proves the second claim in (C.1). To prove the first claim in (C.1), define $\bar{\Pi}_n^\varsigma(q) = \{\varsigma \in \Pi_n^\varsigma(q) : \varsigma_2 - \varsigma_2^{(n)} \in [K^{-1}, K]\}$. In view of (A.37), we have $\bar{\varsigma}_n \in \bar{\Pi}_n^\varsigma(q)$ and hence only need to show

$$\sup_{\varsigma \in \bar{\Pi}_n^\varsigma(q)} \|\partial\Xi_n(\varsigma) - \partial\bar{\Xi}_n(\varsigma)\| = o_{\mathrm{P}}(1). \tag{C.2}$$

This is done in Step 2. Following the proof of the first claim in (A.22) in Step 6 of Section A.4, we conclude $\partial\bar{\Xi}^\star$ has all its eigenvalue strictly greater than zero. Combined with (C.1), we readily deduce (A.40).

Step 2. (Uniform convergence) In this step we prove (C.2). Again note q is finite, so the scenario here is essentially the same as showing uniform convergences of scalars. In particular, following the argument underlying Theorem 2.1 and Corollary 2.2 in Newey (1991), the uniform convergence here comes from pointwise convergence and stochastic equicontinuity. To formalize the proof, introduce

$$\begin{aligned}
\mathcal{R}_n(\varsigma) &= \frac{1}{2n}\mathrm{tr}\left(\frac{\partial^2\Sigma_n(\varsigma)^{-1}}{\partial\varsigma\partial\varsigma^\intercal}Y_nY_n^\intercal\right), \quad \mathcal{R}_{A,n}(\varsigma) = \frac{1}{2n}\mathrm{tr}\left(\frac{\partial^2\Omega_n(\varsigma)^{-1}}{\partial\varsigma\partial\varsigma^\intercal}Y_nY_n^\intercal\right), \\
\mathcal{R}_{D,n}(\varsigma) &= \frac{1}{2n}\mathrm{tr}\left(\frac{\partial^2\Omega_{D,n}(\varsigma)^{-1}}{\partial\varsigma\partial\varsigma^\intercal}Y_nY_n^\intercal\right), \quad \overline{\mathcal{R}}_n(\varsigma) = \frac{1}{2n}\mathrm{tr}\left(\frac{\partial^2\Omega_{D,n}(\varsigma)^{-1}}{\partial\varsigma\partial\varsigma^\intercal}\Omega_n^Y\right).
\end{aligned} \tag{C.3}$$

Here for $A_n \in \{\Sigma_n, \Omega_n, \Omega_{D,n}\}$, we define $A_n(\varsigma)$ by reparameterizing $A_n(\sigma^2, \iota^2, \theta)$, where $\Omega_n(\sigma^2, \iota^2, \theta)$ and $\Omega_{D,n}(\sigma^2, \iota^2, \theta)$ are introduced in Part 2 of Section A.1. Formally, $A_n(\varsigma) = A_n(\sigma^2, \iota^2, \theta)$ under $\iota^2 = \exp(\varsigma_2), \theta = (\varsigma_3, \ldots, \varsigma_{q+2})^\intercal, \zeta^2 = \iota^2(1 + 1^\intercal \cdot \theta)^2$ and $\sigma^2 = \varsigma_1(n\zeta^2)^{1/4}$.

In view of the triangular inequality, plus the definitions of $\partial\Xi_n(\varsigma)$ and $\partial\bar{\Xi}_n(\varsigma)$ in (A.29), the pointwise consistency comes from that for any deterministic sequence $\{\varsigma_n \in \bar{\Pi}_n^\varsigma(q) : n \geq 1\}$,

$$\|\mathcal{R}_n(\varsigma_n) - \mathcal{R}_{A,n}(\varsigma_n)\| = o_{\mathrm{P}}(1), \quad \|\mathcal{R}_{A,n}(\varsigma_n) - \mathcal{R}_{D,n}(\varsigma_n)\| = o_{\mathrm{P}}(1), \quad \|\mathcal{R}_{D,n}(\varsigma_n) - \overline{\mathcal{R}}_n(\varsigma_n)\| = o_{\mathrm{P}}(1). \tag{C.4}$$

One can prove the three claims by repeating Lemma B1, Lemma B2 and Lemma B3, respectively. The proof would be simpler, due to finite q. For the consideration behind introducing $\mathcal{R}_{A,n}$ and $\mathcal{R}_{D,n}$, see Step 4 of Section A.4.

As for the stochastic equicontinuity, we only analyze $\mathcal{R}_{A,n}$ as a demonstration. Concretely, we show for all $1 \leq i, j, k \leq q + 2$,

$$\sup_{\varsigma \in \bar{\Pi}_n^\varsigma(q)} \left| (\mathbb{1}_{\{k \neq 1\}} + \delta_{k,1} n^{-1/4}(\iota^{(n)})^{-1/2}) \frac{\partial}{\partial\varsigma_k}\mathcal{R}_{A,n}(\varsigma)_{ij} \right| = O_{\mathrm{P}}(1). \tag{C.5}$$

Here the additional factor $\delta_{k,1}n^{-1/4}(\iota^{(n)})^{-1/2}$ compared to Assumption 3A in Newey (1991) arises due to the reparameterization of $(\sigma^2, \iota^2, \theta)$ into ς. We prove (C.5) in Step 3 below.

Step 3. (Auxiliary: Stochastic equicontinuity) In this step we prove (C.5). Due to multiple differentiations in $\frac{\partial}{\partial\varsigma_k}\mathcal{R}_{A,n}(\varsigma)_{ij}$, we need somewhat lengthy notation:

$$\mathcal{S}_{A,n}(1, \varsigma)_{ijk} = -\frac{\partial^3 V_n(\varsigma)}{\partial\varsigma_i\partial\varsigma_j\partial\varsigma_k}, \quad \mathcal{S}_{A,n}(2, \varsigma)_{ijk} = 2\frac{\partial^2 V_n(\varsigma)}{\partial\varsigma_i\partial\varsigma_j}V_n(\varsigma)^{-1}\frac{\partial V_n(\varsigma)}{\partial\varsigma_k},$$

$$\mathcal{S}_{A,n}(3, \varsigma)_{ijk} = -2\frac{\partial V_n(\varsigma)}{\partial\varsigma_i}V_n(\varsigma)^{-1}\frac{\partial V_n(\varsigma)}{\partial\varsigma_j}V_n(\varsigma)^{-1}\frac{\partial V_n(\varsigma)}{\partial\varsigma_k}, \quad \mathcal{S}_{A,n}(4, \varsigma)_{ijk} = \mathcal{S}_{A,n}(2, \varsigma)_{jki},$$

$$\mathcal{S}_{A,n}(5, \varsigma)_{ijk} = \mathcal{S}_{A,n}(2, \varsigma)_{ikj}, \quad \mathcal{S}_{A,n}(6, \varsigma)_{ijk} = \mathcal{S}_{A,n}(3, \varsigma)_{ikj}, \quad \mathcal{S}_{A,n}(7, \varsigma)_{ijk} = \mathcal{S}_{A,n}(3, \varsigma)_{kij}.$$

Here we define $V_n(\varsigma)$ by reparameterizing $V_n(\sigma^2, \iota^2, \theta, \Delta_n)$ introduced in Part 2 of Section A.1. Formally, $V_n(\varsigma) = V_n(\sigma^2, \iota^2, \theta, \Delta_n)$ under $\iota^2 = \exp(\varsigma_2), \theta = (\varsigma_3, \ldots, \varsigma_{q+2})^\intercal, \zeta^2 = \iota^2(1 + 1^\intercal \cdot \theta)^2$ and $\sigma^2 = \varsigma_1(n\zeta^2)^{1/4}$.

Recalling $\Omega_n = O_nV_nO_n$ from Part 2 of Section A.1, plus the rule of matrix differentiation, one can verify

$$\frac{\partial}{\partial \varsigma_k}\mathcal{R}_{A,n}(\varsigma)_{ij} = -\sum_{l=1}^{7}\frac{1}{2n}(V_n(\varsigma)^{-1}O_nY_n)^\intercal \mathcal{S}_{A,n}(l,\varsigma)_{ijk}(V_n(\varsigma)^{-1}O_nY_n). \tag{C.6}$$

We focus on the contribution of the third summand in the RHS of (C.6). We shall prove

$$\sup_{i,j,k;\varsigma\in\bar{\Pi}^\varsigma_n(q)} |(\mathbb{1}_{\{k\neq 1\}} + \delta_{k,1}n^{-1/4}(\iota^{(n)})^{-1/2})(V_n(\varsigma)^{-1}O_nY_n)^\intercal \mathcal{S}_{A,n}(3,\varsigma)_{ijk}(V_n(\varsigma)^{-1}O_nY_n)| = O_{\mathrm{P}}(n). \tag{C.7}$$

Introduce short-hand notation:

$$\mathcal{V}_{ll}(i,j,k) := (\mathbb{1}_{\{k\neq 1\}} + \delta_{k,1}n^{-1/4}(\iota^{(n)})^{-1/2})\left|\frac{\partial V_n(\varsigma)_{ll}}{\partial \varsigma_i}V_n(\varsigma)_{ll}^{-1}\frac{\partial V_n(\varsigma)_{ll}}{\partial \varsigma_j}V_n(\varsigma)_{ll}^{-1}\frac{\partial V_n(\varsigma)_{ll}}{\partial \varsigma_k}\right|.$$

The diagonality of V_n indicates

$$|(O_nY_n)^\intercal \mathcal{S}_{A,n}(3,\varsigma)_{ijk}O_nY_n| \le 2\sum_{1\le l\le n}\mathcal{V}_{ll}(i,j,k)(V_n(\varsigma)^{-1}O_nY_n)_l^2.$$

Given this and recalling again $\Omega_n = O_nV_nO_n$, the desired result (C.7) follows from two facts. The first fact is

$$\delta_{1,i}+\delta_{1,j}+\delta_{1,k} = p \in \{0,1,2,3\} \quad\Longrightarrow\quad \mathcal{V}_{ll}(i,j,k) \le K(n^{1/4}(\iota^{(n)})^{1/2})^{p\wedge 2}\Delta_n^p V_n(\varsigma)_{ll}^{1-p},$$

which holds by the definition of $\bar{\Pi}^\varsigma_n(q)$. The second is

$$p\ge 1 \quad\Longrightarrow\quad \sup_{\varsigma\in\bar{\Pi}^\varsigma_n(q)} Y_n\Omega_n(\varsigma)^{-p}Y_n \le KY_n\Omega_n(1,(\iota^{(n)})^2,0)^{-p}Y_n = O_{\mathrm{P}}((\iota^{(n)})^{-1}n^{p+1/2}\vee n).$$

Here the inequality holds by the definition of $\bar{\Pi}^\varsigma_n(q)$ and the equality holds by Lemma D4.

Step 4. (Proof of the second claim in (C.1)) Let $\bar{\iota}_n^2 = \exp((\bar{\varsigma}_n)_2)$, $\bar{\theta}_n = ((\bar{\varsigma}_n)_3,\ldots,(\bar{\varsigma}_n)_{q+2})$ and $\bar{\sigma}_n^2 = (\bar{\varsigma}_n)_1(n\bar{\iota}_n^2 g(0;\bar{\theta}_n))^{1/4}$. Let $\bar{\zeta}_n^2 = \bar{\iota}_n^2 g(0;\bar{\theta}_n)$. By definition, we can write for all $3\le i,j\le q+2$,

$$\left.\begin{aligned}
&\partial\bar{\Xi}_n(\bar{\varsigma}_n)_{11} \underset{(a1)}{=} -\tfrac{\bar{\zeta}_n}{\sqrt{n}}\tfrac{\partial^2 \bar{L}_n(\bar{\sigma}_n^2,\bar{\iota}_n^2,\bar{\theta}_n)}{\partial\sigma^2\partial\sigma^2}, \quad \partial\bar{\Xi}_n(\bar{\varsigma}_n)_{22} \underset{(a2)}{=} -\tfrac{1}{n}\tfrac{\partial^2 \bar{L}_n(\bar{\sigma}_n^2,\bar{\iota}_n^2,\bar{\theta}_n)}{\partial\log(\iota^2)\partial\log(\iota^2)},\\
&\partial\bar{\Xi}_n(\bar{\varsigma}_n)_{12} = -\tfrac{\bar{\zeta}_n^{1/2}}{n^{3/4}}\tfrac{\partial^2 \bar{L}_n(\bar{\sigma}_n^2,\bar{\iota}_n^2,\bar{\theta}_n)}{\partial\sigma^2\partial\log(\iota^2)}, \quad \partial\bar{\Xi}_n(\bar{\varsigma}_n)_{ij} \underset{(a3)}{=} -\tfrac{1}{n}\tfrac{\partial^2 \bar{L}_n(\bar{\sigma}_n^2,\bar{\iota}_n^2,\bar{\theta}_n)}{\partial\theta_{i-2}\partial\theta_{j-2}},\\
&\partial\bar{\Xi}_n(\bar{\varsigma}_n)_{1j} = -\tfrac{\bar{\zeta}_n^{1/2}}{n^{3/4}}\tfrac{\partial^2 \bar{L}_n(\bar{\sigma}_n^2,\bar{\iota}_n^2,\bar{\theta}_n)}{\partial\sigma^2\partial\theta_{j-2}}, \quad \partial\bar{\Xi}_n(\bar{\varsigma}_n)_{2j} = -\tfrac{1}{n}\tfrac{\partial^2 \bar{L}_n(\bar{\sigma}_n^2,\bar{\iota}_n^2,\bar{\theta}_n)}{\partial\log(\iota^2)\partial\theta_{j-2}}.
\end{aligned}\right\} \tag{C.8}$$

We show $|\partial\bar{\Xi}_n(\bar{\varsigma}_n)_{ij} - \partial\bar{\Xi}^\star_{ij}|$ converges for all i and j based on the above expressions. We only analyze three cases (i) $i=j=2$, (ii) $i=j=1$ and (iii) $i,j\ge 3$, while the other cases are simpler repetitions. Step 5 below proves case (i), Step 6 proves case (ii), and Step 7 proves case (iii).

In those steps, we set for all m and $(\iota^2, \theta, \Delta_n)$,

$$\Lambda_m(\iota^2, \theta, \Delta_n) = \Omega_m(0, \iota^2, \theta, \Delta_n), \quad \Lambda_{D,n}(\iota^2, \theta) = \Omega_{D,n}(0, \iota^2, \theta).$$

Moreover, we drop the arguments $(\bar{\sigma}_n^2, \bar{\iota}_n^2, \bar{\theta}_n)$.

Step 5. (Auxiliary: Case (i) in Step 4) Using the definition of $\bar{L}_n$ and observing $\Lambda_{n,D}$ commutes with $\Omega_{n,D}$, we deduce

$$\begin{aligned} -\frac{1}{n}\frac{\partial^2 \bar{L}_n}{\partial \log(\iota^2)\partial \log(\iota^2)} =& \frac{1}{2n}\mathrm{tr}(\Omega_{D,n}^{-1}\Lambda_{D,n}) - \frac{1}{2n}\mathrm{tr}(\Omega_{D,n}^{-2}\Lambda_{D,n}^2) \\ &- \frac{1}{2n}\mathrm{tr}(\Omega_{D,n}^{-2}\Lambda_{D,n}\Omega_n^Y) + \frac{1}{n}\mathrm{tr}(\Omega_{D,n}^{-3}\Lambda_{D,n}^2\Omega_n^Y). \end{aligned} \tag{C.9}$$

Given $(a2)$ in (C.8) and (C.9), the desired result follows from two facts. The first is

$$j \in \{1,2,3\} \quad \Longrightarrow \quad \mathrm{tr}(\Omega_{D,n}^{-j}\Lambda_{D,n}^j) = n + o_{\mathrm{P}}(n) \quad \text{and} \quad j \in \{1,2\} \quad \Longrightarrow \quad \mathrm{tr}(\Omega_{D,n}^{-j-1}\Lambda_{D,n}^j\Delta_n) = o_{\mathrm{P}}(n),$$

which holds by $\Delta_n^{-1}\bar{\iota}_n^2 \to \infty$ and by generalizing Lemma D6 using Lemma D4.

The second fact is that for all $j \geq 1$,

$$\begin{aligned} \left|\mathrm{tr}(\Omega_{D,n}^{-j-1}\Lambda_{D,n}^j(\Omega_n^Y - \Lambda_{D,n}))\right| =& \left|\mathrm{tr}\left(\Omega_{D,n}^{-\frac{j+1}{2}}\Lambda_{D,n}^{\frac{j}{2}}\Omega_n^B\Lambda_{D,n}^{\frac{j}{2}}\Omega_{D,n}^{-\frac{j+1}{2}}\right)\right| \\ &+ \left|\mathrm{tr}\left(\Omega_{D,n}^{-\frac{j+1}{2}}\Lambda_{D,n}^{\frac{j}{2}}(\Omega_n^U - \Lambda_{D,n})\Lambda_{D,n}^{\frac{j}{2}}\Omega_{D,n}^{-\frac{j+1}{2}}\right)\right| \\ \leq& \left|\mathrm{tr}(\Omega_{D,n}^{-j-1}\Lambda_{D,n}^j\Delta_n)\right| + \left|\mathrm{tr}(\Omega_{D,n}^{-j-1}\Lambda_{D,n}^{j+1})\right| \times o_{\mathrm{P}}(1). \end{aligned} \tag{C.10}$$

Here the equality holds by the triangle inequality and observing $\Lambda_{n,D}$ commutes with $\Omega_{n,D}$. The inequalities holds by observing $\|\Omega_n^B\| \lesssim \Delta_n$ and the fact that for some $\{\alpha_n > 0 : n \geq 1\}$ with $\alpha_n = o_{\mathrm{P}}(1)$, it holds

$$-\alpha_n\Lambda_{D,n} \leq \Omega_n^U - \Lambda_{D,n} \leq \alpha_n\Lambda_{D,n}.$$

This result comes from the constructions of Ω_n^U and $\Lambda_{D,n}$ and observing $\|D_m((\iota^{(n)})^2, \theta^{(n)}) - D_m(\bar{\iota}_n^2, \theta^{(n)})\| \lesssim \|\bar{\varsigma}_n - \varsigma^{(n)}\|$ due to (A.37).

Step 6. (Auxiliary: Case (ii) in Step 4) From the definition of $\bar{L}_n$, it follows

$$-\frac{\bar{\zeta}_n}{\sqrt{n}}\frac{\partial^2 \bar{L}_n}{\partial\sigma^2\partial\sigma^2} = \frac{\Delta_n^2\bar{\zeta}_n}{2\sqrt{n}}\mathrm{tr}(\Omega_{D,n}^{-2}) - \frac{\Delta_n^2\bar{\zeta}_n}{\sqrt{n}}\mathrm{tr}(\Omega_{D,n}^{-3}\Omega_n^U) - \frac{\Delta_n^2\bar{\zeta}_n}{\sqrt{n}}\mathrm{tr}(\Omega_{D,n}^{-3}\Omega_n^B). \tag{C.11}$$

Given $(a1)$ in (C.8) and (C.11), the desired result follows from three facts. The first fact is

$$\mathrm{tr}(\Omega_{D,n}^{-2}) = \frac{n}{4(\bar{\sigma}_n^2\Delta_n)^{3/2}\bar{\zeta}_n} \quad \text{and} \quad \mathrm{tr}(\Omega_{D,n}^{-3}\Lambda_{D,n}) = \frac{n}{16(\bar{\sigma}_n^2\Delta_n)^{3/2}\bar{\zeta}_n},$$

which holds by $\Delta_n^{-1}\bar{\iota}_n^2 \to \infty$ and by generalizing Lemma D6 using Lemma D4.

The second fact is that for all $j \geq 1$,

$$\left|\mathrm{tr}(\Omega_{D,n}^{-3}(\Omega_n^U - \Lambda_{D,n}))\right| = \left|\mathrm{tr}(\Omega_{D,n}^{-3/2}(\Omega_n^U - \Lambda_{D,n})\Omega_{D,n}^{-3/2})\right| = \left|\mathrm{tr}(\Omega_{D,n}^{-3/2}\Lambda_{D,n}\Omega_{D,n}^{-3/2})\right| \times o_{\mathrm{P}}(1),$$

which holds by the same reasoning underlying (C.10).

Finally, Lemma D4 leads to

$$\sum_{j=1}^{n}\left|(\Omega_{D,n}^{-3})_{jj} - n^{-1}\mathrm{tr}(\Omega_{D,n}^{-3})\right| = o_{\mathrm{P}}\left(n^{-1}\mathrm{tr}(\Omega_{D,n}^{-3})\right).$$

Step 7. (Auxiliary: Case (iii) in Step 4) Making use of the definition of $\bar{L}_n$, we conclude for all $1 \leq i, j \leq q$,

$$\begin{aligned}-\frac{1}{n}\frac{\partial^2 \bar{L}_n}{\partial\theta_i\partial\theta_j} =& \frac{1}{2n}\mathrm{tr}\left(\Omega_{D,n}^{-1}\frac{\partial^2\Omega_{D,n}}{\partial\theta_i\partial\theta_j}\right) - \frac{1}{2n}\mathrm{tr}\left(\Omega_{D,n}^{-1}\frac{\partial\Omega_{D,n}}{\partial\theta_i}\Omega_{D,n}^{-1}\frac{\partial\Omega_{D,n}}{\partial\theta_j}\right)\\ &- \frac{1}{2n}\mathrm{tr}\left(\Omega_{D,n}^{-1}\frac{\partial^2\Omega_{D,n}}{\partial\theta_i\partial\theta_j}\Omega_{D,n}^{-1}\Omega_n^Y\right) + \frac{1}{n}\mathrm{tr}\left(\Omega_{D,n}^{-1}\frac{\partial\Omega_{D,n}}{\partial\theta_i}\Omega_{D,n}^{-1}\frac{\partial\Omega_{D,n}}{\partial\theta_j}\Omega_{D,n}^{-1}\Omega_n^Y\right).\end{aligned} \tag{C.12}$$

Given $(a3)$ in (C.8) and (C.12), the desired result follows from (a) for all $1 \leq i, j \leq q$,

$$\mathrm{tr}\left(\Omega_{D,n}^{-2}\frac{\partial\Omega_{D,n}}{\partial\theta_i}\frac{\partial\Omega_{D,n}}{\partial\theta_j}\right) = nW(\bar{\theta}_n)_{ij} + o_{\mathrm{P}}(n) \quad \text{and} \quad \mathrm{tr}(\Omega_{D,n}^{-1}\Delta_n) = o_{\mathrm{P}}(n); \tag{C.13}$$

(b) that $\Omega_{D,n}$ Ω_n^Y can be simultaneously diagonalized by $O_{D,n} \coloneqq (\mathbb{I}_{J_d} \otimes O_{n_d}) \oplus O_{n'_d}$; (c) that $O_{D,n}$ does not depend on the argument $(\sigma^2, \iota^2, \theta)$ and hence commutes with the differentiation operation; and (d) for all $1 \leq k \leq n$ and setting $V_{D,n} \coloneqq O_{D,n}\Omega_{D,n}O_{D,n}$,

$$\begin{aligned}&\left|\frac{\partial^2 (V_{D,n})_{kk}}{\partial\theta_i\partial\theta_j}\right| \leq K(V_{D,n})_{kk}, \quad \left|\frac{\partial(V_{D,n})_{kk}}{\partial\theta_i}\frac{\partial(V_{D,n})_{kk}}{\partial\theta_j}\right| \leq K(V_{D,n}^2)_{kk},\\ &\text{and} \quad |(O_{D,n}\Omega_n^Y O_{D,n} - \Omega_{D,n})_{kk}| \leq K\Delta_n + \alpha_n(V_{D,n})_{kk}.\end{aligned} \tag{C.14}$$

Here (C.13) holds by $\Delta_n^{-1}\bar{\iota}_n^2 \to \infty$ and by generalizing Lemma D6 using Lemma D4. The first two claims in (C.14) hold by construction. The last claim in (C.14) originates from $\|D_m((\iota^{(n)})^2, \theta^{(n)}) - D_m(\bar{\iota}_n^2, \theta^{(n)})\| \lesssim \|\bar{\varsigma}_n - \varsigma^{(n)}\|$ due to (A.37). ∎

Lemma C2. *(A.45) holds.*

Proof. Step 1. (Main proof) This lemma is the counterpart of Lemma C1 under $\Delta_n^{-1}(\iota^{(n)})^2 \to a^2 < \infty$. Close scrutiny of the proof of Lemma C1 reveals that we can follow the same strategy here.

Imitating (C.3), we introduce short-hand notation

$$\begin{aligned}\mathcal{R}_n(\beta) &= \frac{1}{2n}\mathrm{tr}\left(\frac{\partial^2\Sigma_n(\beta)^{-1}}{\partial\beta\partial\beta^\intercal}Y_nY_n^\intercal\right), \quad \mathcal{R}_{A,n}(\beta) = \frac{1}{2n}\mathrm{tr}\left(\frac{\partial^2\Omega_n(\beta)^{-1}}{\partial\beta\partial\beta^\intercal}Y_nY_n^\intercal\right),\\ \mathcal{R}_{D,n}(\beta) &= \frac{1}{2n}\mathrm{tr}\left(\frac{\partial^2\Omega_{D,n}(\beta)^{-1}}{\partial\beta\partial\beta^\intercal}Y_nY_n^\intercal\right), \quad \overline{\mathcal{R}}_n(\beta) = \frac{1}{2n}\mathrm{tr}\left(\frac{\partial^2\Omega_{D,n}(\beta)^{-1}}{\partial\beta\partial\beta^\intercal}\Omega_n^Y\right).\end{aligned} \tag{C.15}$$

Here for $A_n \in \{\Sigma_n, \Omega_n, \Omega_{D,n}\}$, we define $A_n(\beta)$ by reparameterizing $A_n(\sigma^2,\gamma)$, where $\Omega_n(\sigma^2,\gamma)$ and $\Omega_{D,n}(\sigma^2,\gamma)$ are introduced in Part 2 of Section A.1. Formally, $A_n(\beta) = A_n(\sigma^2,\gamma)$ under $\sigma^2 = \beta_1$ and $\gamma_j = \Delta_n\beta_{j+2}$ for all $0 \le j \le q$. Moreover, corresponding to $\bar{\Pi}_n^{\varsigma}(q)$ defined in Step 1 of the proof of Lemma C1, here we define $\bar{\Pi}_n^{\beta}(q) = \{\beta \in \Pi_n^{\beta}(q) : \|\beta - \beta^{(n)}\| \in [K^{-1}, K]\}$.

Indeed, following Steps 1 and 2 of the proof of Lemma C1, plus using (A.43) and recalling the definitions of $\partial\Xi_n(\beta)$ and $\partial\bar{\Xi}_n(\beta)$ in (A.32), we only need to prove three results. The first is the pointwise convergence of $\|\partial\Xi_n(\beta) - \partial\bar{\Xi}_n(\beta)\|$. Namely, for any deterministic sequence $\{\beta_n \in \bar{\Pi}_n^{\beta}(q) : n \ge 1\}$,

$$\|\mathcal{R}_n(\beta_n) - \mathcal{R}_{A,n}(\beta_n)\| = o_{\mathrm{P}}(1), \quad \|\mathcal{R}_{A,n}(\beta_n) - \mathcal{R}_{D,n}(\beta_n)\| = o_{\mathrm{P}}(1), \quad \|\mathcal{R}_{D,n}(\beta_n) - \overline{\mathcal{R}}_n(\beta_n)\| = o_{\mathrm{P}}(1),$$

which is the counterpart of (C.4) under $\Delta_n^{-1}(\iota^{(n)})^2 \to a^2 < \infty$. We omit the detailed proof of it. See the discussion after (C.4).

The second is the stochastic equicontinuity of $\|\partial\Xi_n(\beta) - \partial\bar{\Xi}_n(\beta)\|$. Formally, due to finite q, it suffices to show for all $1 \le i,j,k \le q+2$ and $A(\beta)_{ij} \in \{\mathcal{R}_n(\beta)_{ij}, \mathcal{R}_{A,n}(\beta)_{ij}, \mathcal{R}_{D,n}(\beta)_{ij}, \overline{\mathcal{R}}_n(\beta)_{ij}\}$,

$$\sup_{\varsigma\in\bar{\Pi}_n^{\beta}(q)} \left|\frac{\partial}{\partial\beta_k}A(\beta)_{ij}\right| = O_{\mathrm{P}}(1). \tag{C.16}$$

We prove this in Step 2 below.

Finally, Step 3 below proves the third result we need. We show, for all $\{\beta_n \in \mathbb{R}^{q+2} : n \ge 1\}$ satisfying $\|\beta_n - \beta^{(n)}\| = o_{\mathrm{P}}(1)$,

$$\left\|\partial\bar{\Xi}_n(\beta_n) + \frac{1}{2}W(C_T, \gamma^\star, 1)\right\| = o_{\mathrm{P}}(1). \tag{C.17}$$

Step 2. (Stochastic equicontinuity) In this step we show (C.16).

First, it holds by construction that for all $1 \le i,j \le q+2$ and $\widetilde{A}_n(\beta) \in \{\Sigma_n(\beta), \Omega_n(\beta), \Omega_{D,n}(\beta)\}$,

$$\left\|\frac{\partial}{\partial\beta_i}\frac{\partial}{\partial\beta_j}\widetilde{A}_n(\beta)\right\| = 0 \quad \text{and} \quad \left\|\frac{\partial}{\partial\beta_i}\widetilde{A}_n(\beta)\right\| \le K. \tag{C.18}$$

Second, in view of the definitions of $\Omega_n(\beta)$ and $\Omega_{D,n}(\beta)$ given in Step 1, plus the definition of

$\bar{\Pi}_n^{\beta}(q)$ and Proposition 4.5.3 in Brockwell and Davis (1991), we conclude

$$\sup_{\beta\in\bar{\Pi}_n^{\beta}(q)} \|\Omega_n(\beta)\| \vee \|\Omega_n(\beta)^{-1}\| \vee \|\Omega_{D,n}(\beta)\| \vee \|\Omega_{D,n}(\beta)^{-1}\| \vee \|\Sigma_n(\beta)\| \vee \|\Sigma_n(\beta)^{-1}\| \leq K. \quad \text{(C.19)}$$

Observing $\mathrm{tr}(Y_n Y_n^\intercal) = O_{\mathrm{P}}(1)$ due to $\Delta_n^{-1}(\iota^{(n)})^2 \to a^2 < \infty$, plus (C.15) and the rule of matrix differentiation, (C.18) and (C.19) readily yield (C.16).

Step 3. (Proof of (C.17)) Below we drop the argument β_n of $\Omega_{D,n}$ for simplicity.

From the definitions of $\partial\bar{\Xi}_n(\beta)$ given by (A.32), it follows for all $1 \leq i,j \leq q+2$,

$$\partial\bar{\Xi}_n(\beta_n)_{ij} = -\frac{1}{2n}\mathrm{tr}\left(\Omega_{D,n}^{-1}\frac{\partial\Omega_{D,n}}{\partial\beta_i}\Omega_{D,n}^{-1}\frac{\partial\Omega_{D,n}}{\partial\beta_j}\right) + \frac{1}{n}\mathrm{tr}\left(\Omega_{D,n}^{-1}\frac{\partial\Omega_{D,n}}{\partial\beta_i}\Omega_{D,n}^{-1}\frac{\partial\Omega_{D,n}}{\partial\beta_j}\Omega_{D,n}^{-1}\Omega_n^{Y}\right), \quad \text{(C.20)}$$

where we use $\frac{\partial}{\partial\beta_i}\frac{\partial}{\partial\beta_j}\Omega_{D,n} = 0$ from its definition.

Moreover, using $\Delta_n^{-1}(\iota^{(n)})^2 \to a^2 < \infty$ and by generalizing Lemma D6 with Lemma D4, we conclude for all $1 \leq i,j \leq q+2$,

$$\mathrm{tr}\left(\Omega_{D,n}^{-2}\frac{\partial\Omega_{D,n}}{\partial\beta_i}\frac{\partial\Omega_{D,n}}{\partial\beta_j}\right) = nW(C_T,\gamma^\star,1)_{ij} + o_{\mathrm{P}}(n). \quad \text{(C.21)}$$

Note $\|\Omega_n^Y - \Omega_{D,n}\| \leq K\|\beta_n - \beta^{(n)}\| = o_{\mathrm{P}}(1)$. Hence, in view of the second claim in (C.18) and (C.19), plus $\|AB\| \leq \|A\|\|B\|$ for all $A, B \in \mathcal{M}_n$, (C.20) and (C.21) already indicate (C.17). ∎

Lemma C3. *(b1) and (b4) hold; (b2) and (b5) hold; (b3) and (b6) hold.*

Proof. Using Lemma D4 instead of Lemma D3, (b1) and (b4) follow from the same reasoning as in the proof of Lemma B1; (b2) and (b5) hold by that of Lemma B2; (b3) and (b6) follow from that of Lemma B3. ∎

Lemma C4. *(A.48) holds.*

Proof. Step 1. (Main proof) In view of the triangle inequality, (A.48) follows if there is a function $L_{A,n}(\chi^2,\phi)$ on $\Pi_n^{(\chi^2,\phi)}(q)$ such that

$$\sup_{(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(q)} |L_n(\chi^2,\phi) - L_{A,n}(\chi^2,\phi)| = o_{\mathrm{P}}(n)$$
$$\text{and} \quad \sup_{(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(q)} |L_{A,n}(\chi^2,\phi) - \bar{L}_n^\star(\chi^2,\phi,q)| = o_{\mathrm{P}}(n). \quad \text{(C.22)}$$

We construct $L_{A,n}(\chi^2,\phi)$ explicitly. Let $D_n(\chi^2,\phi) = D_n(\iota^2,\theta)$ with $(\chi^2,\phi) = (\iota^2,\theta)$ and $D_n(\iota^2,\theta)$ given by Part 2 of Section A.1. Let $\Omega_n(\chi^2,\phi) = O_n D_n(\chi^2,\phi) O_n$. Recalling $\mathcal{L}(\cdot,\cdot)$ introduced in Part 3 of Section A.1, we define $L_{A,n}(\chi^2,\phi)$ as

$$L_{A,n}(\chi^2,\phi) = \mathcal{L}(\Omega_n(\chi^2,\phi), Y_n Y_n^\intercal). \quad \text{(C.23)}$$

Using Lemma D2 we observe that for $(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(0)$, it holds $\Omega_n(\chi^2,\phi)=\Sigma_n(\chi^2,\phi)$ and hence $L_{A,n}(\chi^2,\phi)=L_n(\chi^2,\phi)$, an evidence that $L_{A,n}(\chi^2,\phi)$ is a good approximation of $L_n(\chi^2,\phi)$. Indeed, for all fixed q, one can prove the first claim in (C.22) following the argument in the proof of Lemma B1. Although the parameter space is different, the proof would be simpler as the bound we require here is not sharp. We hence only prove the second claim in (C.22).

Using Lemma D2 and the construction of $\Omega_n(\chi^2,\phi)$, one can easily verify that $\frac{1}{n}\log\det\Omega_n(\chi^2,\phi)=\log\chi^2+O(n^{-1}\log n)$ holds uniformly over $\Pi_n^{(\chi^2,\phi)}(q)$. Therefore, recalling the definition of $\bar{L}_n^\star(\chi^2,\phi,q)$ given in Part 3 of Step 1 of Section A.6, plus (C.23), it suffices to show that uniformly over $\Pi_n^{(\chi^2,\phi)}(q)$,

$$\mathcal{R}_a(\chi^2,\phi):=\underbrace{Y_n^\intercal\Omega_n(\chi^2,\phi)^{-1}Y_n}_{=:\mathcal{R}_b(\chi^2,\phi)}-\underbrace{\frac{\chi^{(n)}(q)^2}{2\pi\chi^2}\int_{-\pi}^{\pi}\frac{g(\lambda;\phi^{(n)}(q))}{g(\lambda;\phi)}d\lambda}_{=:\mathcal{R}_c(\chi^2,\phi)}=o_{\mathrm{P}}(n). \tag{C.24}$$

One challenge of showing (C.24) is that the parameter space $\Pi_n^{(\chi^2,\phi)}(q)$ is too large for us to handle $\mathcal{R}_a(\chi^2,\phi)$ in a unified way. To resolve this, we define a family of subsets of $\Pi_n^{(\chi^2,\phi)}(q)$ indexed by α_1 and α_2:

$$\Pi_n^{(\chi^2,\phi)}(q,\alpha_1,\alpha_2)=\{(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(q):\alpha_1\le\chi^2\le\alpha_2\}.$$

Obviously, $\Pi_n^{(\chi^2,\phi)}(q)=\Pi_n^{(\chi^2,\phi)}(q,0,\alpha)\cup\Pi_n^{(\chi^2,\phi)}(q,\alpha,\infty)$ for all α. Hence by the classic subsequence argument, it suffices to show for all $\alpha_n\to\infty$ and all K fixed,

$$\sup_{(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(q,\alpha_n\Delta_n,\infty)}|\mathcal{R}_a(\chi^2,\phi)|=o_{\mathrm{P}}(n)\quad\text{and}\quad\sup_{(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(q,0,K\Delta_n)}|\mathcal{R}_a(\chi^2,\phi)|=o_{\mathrm{P}}(n). \tag{C.25}$$

The first claim is proved in Step 2 below. Step 3 proves the second claim.

Step 2. (Proof of the first claim in (C.25)) In view of the triangle inequality and the definitions of $\mathcal{R}_b(\chi^2,\phi)$ and $\mathcal{R}_c(\chi^2,\phi)$ given in (C.24), the first claim in (C.25) follows from the uniform convergences of both $|\mathcal{R}_b(\chi^2,\phi)|$ and $|\mathcal{R}_c(\chi^2,\phi)|$. The latter is simpler. So we only show the former:

$$\sup_{(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(q,\alpha_n\Delta_n,\infty)}|\mathcal{R}_b(\chi^2,\phi)|=o_{\mathrm{P}}(n). \tag{C.26}$$

Using the definitions of $\mathcal{R}_b$, write

$$\mathcal{R}_b(\chi^2,\phi)=Y_n^\intercal O_nD_n(\chi^2,\phi)^{-1}O_nY_n=Y_n^\intercal O_n\sum_{h=0}^{\infty}\rho_h(\chi^2,\phi)\mathbb{D}_n^hO_nY_n,$$

where $\rho_h(\chi^2,\phi)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{e^{\mathrm{i}h\lambda}}{\chi^2g(\lambda;\phi)}d\lambda$. The last equality holds by the definition of D_n and the discrete Fourier transfrom argument (see, e.g., Step 1 of the proof of Lemma D3). In view of Lemma D2, plus $\Delta_n^{-1}(\iota^{(n)})^2 \to a^2 < \infty$, conclude

$$\mathbb{E}|Y_n^\intercal O_n \mathbb{D}_n^h O_n Y_n| = \mathbb{E}|Y_n^\intercal \mathbb{F}_n^h Y_n| \le K\mathbb{1}_{\{h\le q+1\}} + Kn^{-1/2}.$$

Hence, we only need to show $\rho_h(\chi^2,\phi)$ shrinks fast enough as n increases. Indeed, Step 5 below proves

$$\sup_{(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(q)} |\rho_h(\chi^2,\phi)| \le K\chi^{-2}\left(1-\sqrt{\chi^{-2}\Delta_n}/K\right)^h. \tag{C.27}$$

Noting that $\mathbb{D}_n^h$ is diagonal and positive definite for every h and applying Chebyshev's inequality, we readily deduce (C.26).

Step 3. (Proof of the second claim in (C.25)) As in the proof of Lemma C1, we follow Newey (1991) and decompose the proof of uniform convergence into those of pointwise convergence and stochastic equicontinuity. Concretely, following the reasoning in Step 5 of the proof of Theorem 4, we conclude for all fixed $a^2 \in \mathbb{R}_+$ and $b \in \Theta(q+1)$,

$$\mathcal{R}_a(a^2\Delta_n, b) = o_{\mathrm{P}}(n).$$

On the other hand, Step 5 below proves for all $1 \le j \le q+1$ and $A \in \{\mathcal{R}_b, \mathcal{R}_c\}$,

$$\sup_{(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(q,0,K\Delta_n)} \left|\frac{\Delta_n\partial}{\partial\chi^2}A(\chi^2,\phi)\right| = O_{\mathrm{P}}(n) \quad \text{and} \quad \sup_{(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(q,0,K\Delta_n)} \left|\frac{\partial}{\chi^2\partial\phi_j}A(\chi^2,\phi)\right| = O_{\mathrm{P}}(n), \tag{C.28}$$

from which the stochastic equicontinuity follows. Here the additional factor Δ_n compared to Assumption 3A in Newey (1991) arises due to the definition of $\Pi_n^{(\chi^2,\phi)}(q,0,K\Delta_n)$.

Step 4. (Auxiliary: Proof of (C.27)) Note for all $(\chi^2,\phi) \in \Pi_n^{(\chi^2,\phi)}(q)$, we can define $\{\widetilde{\rho}_j(\phi)\}_{j=1}^\infty$ by

$$1 + \sum_{j=1}^\infty \widetilde{\rho}_j(\phi)e^{\mathrm{i}j\lambda} = \frac{1}{1+\sum_{j=1}^{q+1}\phi_j e^{\mathrm{i}j\lambda}} \quad \text{for all} \quad \lambda \in (-\pi,\pi).$$

With $\widetilde{\rho}_0(\phi) = 1$ by convention, the definition of $\rho_h(\chi^2,\phi)$ indicates

$$\rho_h(\chi^2,\phi) = \chi^{-2}\sum_{j=0}^\infty \widetilde{\rho}_j(\phi)\widetilde{\rho}_{j+h}(\phi). \tag{C.29}$$

Therefore, we only need the behavior of $\chi^{-1}\widetilde{\rho}_j(\phi)$ for large j. To do so, using the definition of $\Pi_n^{(\chi^2,\phi)}(q)$, plus that both q and $\|\phi\|$ are finite, we obtain

$$\inf_{z\in\mathbb{C}:|z|\le 1} \chi^2\left|1+\sum_{j=1}^{q+1}\phi_j z^j\right| \ge K^{-1}\Delta_n \quad \text{and} \quad \sup_{z\in\mathbb{C}:|z|\le 1+\sqrt{\chi^2\Delta_n}/K} \frac{d}{dz}\left|1+\sum_{j=1}^{q+1}\phi_j z^j\right| \le K.$$

A direct result of the two inequalities is that $\inf_{z\in\mathbb{C}:|z|\le 1+\sqrt{\chi^{-2}\Delta_n}/K}\left|1+\sum_{j=1}^{q+1}\phi_j z^j\right|>0$ holds uniformly over $(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(q)$, which, on the other hand, immediately indicates

$$\sup_{(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(q)}|\widetilde{\rho}_j(\phi)|\le\left(1-\sqrt{\chi^{-2}\Delta_n}/K\right)^h.$$

Plugging this back into (C.29), we already deduce (C.27).

Step 5. (Auxiliary: Proof of (C.28)) We only show the second claim in (C.28) for $A=\mathcal{R}_b$. Using that D_n is diagonal, write

$$\frac{\partial}{\partial\phi_j}\mathcal{R}_b(\chi^2,\phi)=-\sum_{k=1}^n D_n(\chi^2,\phi)_{kk}^{-2}\frac{\partial D_n(\chi^2,\phi)_{kk}}{\partial\phi_j}(O_nY_n)_k^2.$$

Let $g_+(\lambda,\phi)=1+\sum_{j=1}^\infty\phi_j e^{\mathbb{i}j\lambda}$. In view of the definition of D_n given in Step 1, we conclude uniformly over $(\chi^2,\phi)\in\Pi_n^{(\chi^2,\phi)}(q,0,K\Delta_n)$ and $1\le k\le n$,

$$\begin{aligned}\left|D_n(\chi^2,\phi)_{kk}^{-2}\frac{\partial D_n(\chi^2,\phi)_{kk}}{\partial\phi_j}\right| =&2\chi^{-2}g^{-2}\left(\frac{k\pi}{n+1};\phi\right)\left|\mathrm{Re}\left(g_+\left(e^{\mathbb{i}\frac{k}{n+1}};\phi\right)e^{\mathbb{i}\frac{jk}{n+1}}\right)\right|\\ \le&2(q+1)\chi^{-2}g^{-2}\left(e^{\mathbb{i}\frac{k}{n+1}};\phi\right)\le K\chi^2\Delta_n^{-2}\le K\Delta_n^{-1}.\end{aligned}$$

Here all three inequalities hold by the definition of $\Pi_n^{(\chi^2,\phi)}(q,0,K\Delta_n)$. The desired result hence follows from $\sum_{k=1}^n(O_nY_n)_k^2=O_{\mathrm{P}}(1)$ due to the orthogonality of O_n and $\Delta_n^{-1}(\iota^{(n)})^2\to a^2<\infty$. ∎

Lemma C5. *Let Assumptions 1 - 3 hold. For $\theta_j^{(n)}\to\theta_j^\star$ and any fixed positive integer q which satisfies $\Delta_n^{1/2}\sum_{j=q}^\infty|\theta_j^{(n)}|\to 0$, we have:*

(i) Under $\Delta_n^{-1}(\iota^{(n)})^2\to\infty$,

$$(\zeta^{(n)})^{-1}\Delta_n^{1/2}AVAR(C_T,\gamma^{(n)},\Delta_n,C(4)_T-C_T^2,cum_4[\varepsilon])=\frac{1}{T}(5C_T^{-1/2}C(4)_T+3C_T^{3/2})+o_{\mathrm{P}}(1).$$

(ii) Under $\Delta_n^{-1}(\iota^{(n)})^2\to a^2\in[0,\infty)$,

$$AVAR(C_T,\gamma^{(n)},\Delta_n,C(4)_T-C_T^2,cum_4[\varepsilon])=AVAR(C_T,\gamma^\star,1,C(4)_T-C_T^2,cum_4[\varepsilon])+o_{\mathrm{P}}(1),$$

with $\gamma^\star\in\mathbb{R}^{q+1}$ defined by $\gamma_j^\star=a^2\sum_{j=0}^{q-j}\theta_l^\star\theta_{l+j}^\star, 0\le j\le q$.

(iii)

$$AVAR(C_T,0,1,C(4)_T-C_T^2,cum_4[\varepsilon])=\frac{1}{T}(4q+6)C(4)_T.$$

Proof. Step 1. (Main proof of (i)) We first analyze the case with $\Delta_n^{-1}(\iota^2)^{(n)}\to\infty$. For notational simplicity, we denote $\Delta_n^{-1}(\iota^{(n)})^2$ by α_n. Let $\rho^{(n)}=(\rho_0^{(n)},\rho_1^{(n)},\rho_2^{(n)},\dots,\rho_{q+1}^{(n)})^\intercal$ be such that for any

$0 \leq k \leq q+1$ and any $\lambda \in (-\pi, \pi)$,

$$\alpha_n \rho_0^{(n)} \prod_{k=1}^{q+1} |1 - \rho_k^{(n)} e^{\mathbb{i}\lambda}|^2 = f(\lambda; C_T, \gamma^{(n)}, \Delta_n)$$
$$\text{and} \quad \rho_0^{(n)} > 0, \quad 1 > |\rho_1^{(n)}| \geq |\rho_2^{(n)}| \geq \ldots \geq |\rho_{q+1}^{(n)}|.$$

Following the reasoning in Step 3 of the proof of Lemma D3, plus q is fixed here and $\alpha_n \to \infty$, one can deduce

$$\begin{aligned} &\rho_0^{(n)} \geq K^{-1}\alpha_n, \quad \Delta_n^{-1/2}(\zeta)^{(n)}(1-\rho_1^{(n)}) = \sqrt{C_T} + o_{\mathrm{P}}(1), \\ &\inf_{j\geq 1}(1 - |\rho_j^{(n)}|) \geq K^{-1}, \quad \text{and} \quad \prod_{k=2}^{q+1}(1-\rho_k^{(n)})^2 = g(0; \theta^{(n)}) + o_{\mathrm{P}}(1). \end{aligned} \tag{C.30}$$

In view of them, the desired result in (i) follows from

$$T \times \mathrm{AVAR}(C_T, \gamma^{(n)}, \Delta_n, Q, c) = 8C_T^2(1-\rho_1^{(n)})^{-1} + \frac{5C_T^2 Q}{\alpha_n^2}(1-\rho_1^{(n)})^{-5} \prod_{k=2}^{q+1}(1-\rho_k^{(n)})^{-4} + O_{\mathrm{P}}(1), \tag{C.31}$$

which holds by Step 2 below.

Step 2. (Proof of (C.31)) Denote the Jacobian matrix $\frac{\partial(C_T, \gamma^{(n)})}{\partial \rho^{(n)}}$ by J_n. Define $Z(\rho^{(n)}), \widetilde{Z}(\rho^{(n)}, Q) \in \mathcal{M}_{q+2}$ as

$$Z(\rho^{(n)}) = J_n^\intercal W(C_T, \gamma^{(n)}, \Delta_n) J_n, \quad \widetilde{Z}(\rho^{(n)}, Q, c) = J_n^\intercal \widetilde{W}(C_T, \gamma^{(n)}, \Delta_n, Q, c) J_n.$$

Let $J_n(1)$ be the first row of J_n. Then the definition of the function AVAR immediately gives

$$\mathrm{AVAR}(C_T, \gamma^{(n)}, \Delta_n, Q, c) = \frac{1}{T} J_n(1) Z(\rho^{(n)})^{-1} \widetilde{Z}(\rho^{(n)}, Q, c) Z(\rho^{(n)})^{-1} J_n(1)^\intercal.$$

Noting $C_T = f(0; C_T, \gamma^{(n)}, \Delta_n)$, plus the definition of $\rho^{(n)}$ and (C.30), we obtain

$$J_n(1)_{k+1} = \frac{\partial C_T}{\partial \rho_k^{(n)}} = \delta_{k,0} O_{\mathrm{P}}(\alpha_n^{-1}) - \delta_{k,1} 2C_T(1-\rho_1^{(n)})^{-1} + \mathbb{1}_{\{k\geq 2\}} O_{\mathrm{P}}(1).$$

Thereby it suffices to calculate $Z(\rho^{(n)})$ and $\widetilde{Z}(\rho^{(n)}, Q, c)$. Indeed, Step 3 below proves for all $i, k \geq 3$ and $j \geq 2$,

$$\begin{aligned} &Z(\rho^{(n)})_{11}^{-1} = O_{\mathrm{P}}(\alpha_n^2), \quad Z(\rho^{(n)})_{1j}^{-1} = 0, \quad Z(\rho^{(n)})_{ij}^{-1} = O_{\mathrm{P}}(1), \\ &Z(\rho^{(n)})_{22}^{-1} = 1 - \rho_1^{(n)} + o_{\mathrm{P}}(\alpha_n^{-1/2}), \quad Z(\rho^{(n)})_{2k}^{-1} = O_{\mathrm{P}}(\alpha_n^{-1/2}). \end{aligned} \tag{C.32}$$

Setting $\check{Z}(\rho^{(n)}) = \frac{1}{2Q}(\widetilde{Z}(\rho^{(n)}, Q, 0) - 2Z(\rho^{(n)}))$, Step 4 below proves for all $i, j \geq 3$,

$$\begin{aligned}
&\check{Z}(\rho^{(n)})_{22} = \frac{5\Delta_n^2}{8(\iota^4)^{(n)}}(1-\rho_1^{(n)})^{-5}\prod_{k=2}^{\bar{q}+1}(1-\rho_k^{(n)})^{-4} + O_{\mathrm{P}}(\alpha_n^{-2}),\\
&\check{Z}(\rho^{(n)})_{11} = O_{\mathrm{P}}(\alpha_n^{-5/2}), \quad \check{Z}(\rho^{(n)})_{12} = O_{\mathrm{P}}(\alpha_n^{-1}), \quad \check{Z}(\rho^{(n)})_{1j} = O_{\mathrm{P}}(\alpha_n^{-3/2}),\\
&\check{Z}(\rho^{(n)})_{2j} = O_{\mathrm{P}}(1), \quad \check{Z}(\rho^{(n)})_{ij} = O_{\mathrm{P}}(\alpha_n^{-1/2}).
\end{aligned} \tag{C.33}$$

We skip analyzing the contribution of the argument c to $\widetilde{Z}(\rho^{(n)}, Q, c)$. Plugging (C.32) and (C.33) back leads to (C.31).

Step 3. (Auxiliary: Entries of $Z(\rho^{(n)})^{-1}$) In this step we prove (C.32).

From the definition of $Z(\rho^{(n)})$, one can calculate, for all $i, j \geq 2$,

$$Z(\rho^{(n)})_{11} = (\rho_0^{(n)})^{-2}, \quad Z(\rho^{(n)})_{1j} = 0, \quad Z(\rho^{(n)})_{ij} = \frac{2}{1-\rho_{i-1}^{(n)}\rho_{j-1}^{(n)}}.$$

Introduce short-hand notation $\mathcal{U}_n \in \mathcal{M}_q, \mathcal{V}_n \in \mathbb{R}^q$ defined by $(\mathcal{U}_n)_{ij} = Z(\rho^{(n)})_{i+2,j+2}$ and $(\mathcal{V}_n)_j = Z(\rho^{(n)})_{2,j+2}$ and let $\mathcal{W}_n = Z(\rho^{(n)})_{22} - \mathcal{V}_n^{\intercal}\mathcal{U}_n^{-1}\mathcal{V}_n$. Applying the block matrix inversion formula, one obtains

$$Z(\rho^{(n)})^{-1} = \begin{pmatrix} (Z(\rho^{(n)})_{11})^{-1} & 0 & 0 \\ 0 & \mathcal{W}_n^{-1} & -\mathcal{W}_n^{-1}\mathcal{V}_n^{\intercal}\mathcal{U}_n \\ 0 & \mathcal{U}_n^{-1}\mathcal{V}_n\mathcal{W}_n & \mathcal{U}_n^{-1} + \mathcal{U}_n^{-1}\mathcal{V}_n\mathcal{W}_n^{-1}\mathcal{V}_n^{\intercal}\mathcal{U}_n^{-1} \end{pmatrix}.$$

Based on this, (C.32) follow from

$$\mathcal{W}_n^{-1} = 1 - \rho_1^{(n)} + o_{\mathrm{P}}((\iota^{-1})^{(n)}\Delta_n^{1/2}),$$

which holds by (a) $\mathbb{I}_n \lesssim \mathcal{U}_n$, (b) $\max_{i\geq 3, j\geq 2}|Z(\rho^{(n)})_{ij}| = O_{\mathrm{P}}(1)$, and (c) $1 - \rho_1^{(n)} \sim \alpha_n^{-1/2}$. All three are from (C.30).

Step 4. (Auxiliary: Entries of $\check{Z}(\rho^{(n)})$) In this step we prove (C.33). We only show the second result therein. In view of (C.30), (C.33) follows from

$$\begin{aligned}
\check{Z}(\rho^{(n)})_{11} &= (\rho_0^{(n)})^{-4}\frac{1}{2\pi \mathbb{i}}\oint_{|z|=1} z\prod_{k=1}^{q+1}(1-\rho_k^{(n)}z)^{-2}(z-\rho_k^{(n)})^{-2}dz\\
&= (\rho_0^{(n)})^{-4}\sum_{l=1}^{q+1}\lim_{z\to\rho_l^{(n)}}\frac{d}{dz}\left(z(z-\rho_l^{(n)})^2\prod_{k=1}^{q+1}(1-\rho_k^{(n)}z)^{-2}(z-\rho_k^{(n)})^{-2}\right).
\end{aligned}$$

Here the first equality comes from the definition of $\check{Z}$. The second comes from Cauchy's residue theorem and $|\rho_k^{(n)}| < 1, \forall k$.

Step 5. (Proof of (ii)) Now we move to the case $\Delta_n^{-1}(\iota^2)^{(n)} \to a^2 \in [0, \infty)$. The desired result follows from

$$\|\Delta_n^{-1}\gamma^{(n)} - \gamma^\star\| = o_{\mathrm{P}}(1) \quad \text{and} \quad \mathrm{AVAR}(\sigma^2, \gamma, \delta, Q, c) = \mathrm{AVAR}(\sigma^2, \delta^{-1}\gamma, 1, Q, c).$$

Both hold by construction.

Step 6. (Proof of (iii)) Define $\mu(\sigma^2, \gamma) = (\mu_0, \mu_1, \ldots, \mu_{q+1}) \in \mathbb{R}^{q+2}$ by

$$f(\lambda; \mu(\sigma^2, \gamma)) = \sum_{j=-q-1}^{q+1} \mu(\sigma^2, \gamma)_{|j|} e^{\mathbb{i} j\lambda} = f(\lambda; \sigma^2, \gamma, 1).$$

The desired result follows from (a)

$$\widetilde{W}(C_T, 0, 1, Q, c) = \left(2 + \frac{2Q}{C_T^2}\right) W(C_T, 0, 1),$$

which holds by the definition of W and $\widetilde{W}$ and $f(\lambda; C_T, 0, 1) = C_T$; (b)

$$(W(\sigma^2, \gamma, 1)^{-1})_{11} = \left(\frac{\partial \sigma^2}{\partial \mu}\right)^\intercal \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \log f(\lambda; \mu(\sigma^2, \gamma))}{\partial \mu} \frac{\partial \log f(\lambda; \mu(\sigma^2, \gamma))}{\partial \mu^\intercal}\right)^{-1} \frac{\partial \sigma^2}{\partial \mu},$$

which is obvious; and (c) observing $\frac{\partial}{\partial \mu_k} f(\lambda; \mu(\sigma^2, \gamma)) = (2 - \delta_{k,0}) \cos k\lambda$ and $\sigma^2 = f(0; \mu(\sigma^2, \gamma))$ and using the the orthogonality of cosine functions. ∎

Appendix D Auxiliary Lemmas

Lemma D1. *Let*

$$h_n^\#(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{g(\lambda; \theta^{(n)})}{g(\lambda; \theta)} d\lambda, \quad h_n(\iota^2, \theta) = \log \iota^2 + \frac{(\iota^2)^{(n)}}{\iota^2} h_n^\#(\theta).$$

For any n and any $\theta, \theta^{(n)} \in \Theta(\infty)$, $\iota^2, (\iota^2)^{(n)} \in [K^{-1}, K]$, the following claims hold,

(i) $\min_{\theta \in \Theta(q_n)} h_n^\#(\theta) - 1 \lesssim \|\theta^{(n)}\|_{(q_n)}^2$, *and (ii)* $(\iota^2 - (\iota^2)^{(n)})^2 + \|\theta - \theta^{(n)}\|^2 \lesssim h_n(\iota^2, \theta) - h_n((\iota^2)^{(n)}, \theta^{(n)})$.

Proof. It suffices to prove

$$\|\theta - \theta^{(n)}\|^2 \underset{(a1)}{\lesssim} h_n^\#(\theta) - 1 \underset{(a2)}{\lesssim} \|\theta - \theta^{(n)}\|^2,$$

because (i) follows from $(a2)$ directly and (ii) follows from $(a1)$ and

$$h_n(\iota^2, \theta) - h_n((\iota^2)^{(n)}, \theta^{(n)}) \geq \log h_n^\#(\theta).$$

$(a1)$ and $(a2)$ follow from

$$h_n^{\#}(\theta) - 1 \underset{(a3)}{=} -1 + \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|1 + \frac{\sum_{j=1}^{\infty}(\theta_j^{(n)} - \theta_j)e^{\mathbb{i}j\lambda}}{1+\sum_{j=0}^{\infty}\theta_j e^{\mathbb{i}j\lambda}}\right|^2 d\lambda$$

$$\underset{(a4)}{=} \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|\sum_{j=1}^{\infty}(\theta_j^{(n)} - \theta_j)e^{\mathbb{i}j\lambda}\right|^2}{\left|1+\sum_{j=0}^{\infty}\theta_j e^{\mathbb{i}j\lambda}\right|^2} d\lambda \underset{(a5)}{\in} \left[\frac{1}{K}\|\theta - \theta^{(n)}\|^2, K\|\theta - \theta^{(n)}\|^2\right].$$

Here $(a3)$ holds by the definition of $h_n^{\#}$. $(a4)$ holds by (a) that there exists $\{\rho_j : j \geq 1\}$ satisfying $\sum_{j=1}^{\infty}|\rho_j| < \infty$ and $\frac{1}{1+\sum_{j=0}^{\infty}\theta_j e^{\mathbb{i}j\lambda}} = 1 + \sum_{j=1}^{\infty}\rho_j e^{\mathbb{i}j\lambda}$ due to the definition of $\Theta(q)$ and (b) $\int_{-\pi}^{\pi} e^{\mathbb{i}j\lambda}d\lambda = 0$ for all nonzero integer j. $(a5)$ holds by (a) the upper and lower boundedness of $\left|1+\sum_{j=0}^{\infty}\theta_j e^{\mathbb{i}j\lambda}\right|$ due to the definition of $\Theta(q)$ and (b) $\int_{-\pi}^{\pi} e^{\mathbb{i}j\lambda}d\lambda = 0$ for any nonzero integer j. ∎

Lemma D2. *For all integers m and h satisfying $0 \leq h \leq m$, it holds*

$$\mathbb{D}_m^h = O_m \mathbb{F}_m^h O_m,$$

with $\mathbb{F}_m^h \in \mathcal{M}_m$ given by

$$(\mathbb{F}_m^h)_{ij} = \mathbb{1}_{\{h=|i-j|\}} - \mathbb{1}_{\{h=i+j\}} - \mathbb{1}_{\{h=2m+2-(i+j)\}}. \tag{D.1}$$

Proof. We drop the subscript m of matrix O_m. The claim follows from

$$(O\mathbb{F}_m^h O)_{pq} \underset{(a1)}{=} \sum_{i=1}^{m} O_{pi}\left(\sum_{j=1}^{m}\mathbb{1}_{\{h=|i-j|\}}O_{jq} - \sum_{j=-m}^{-1}\mathbb{1}_{\{h=i-j\}}O_{-j,q} - \sum_{j=m+2}^{2m+1}\mathbb{1}_{\{h=j-i\}}O_{2m+2-j,q}\right)$$

$$\underset{(a2)}{=} \sum_{i=1}^{m} O_{pi}\left(\sum_{j=1}^{m}\mathbb{1}_{\{h=|i-j|\}}O_{jq} + \sum_{j=-m}^{-1}\mathbb{1}_{\{h=|i-j|\}}O_{jq} + \sum_{j=m+2}^{2m+1}\mathbb{1}_{\{h=|i-j|\}}O_{jq}\right)$$

$$\underset{(a3)}{=} \sum_{i=1}^{m} O_{pi}(O_{i+h,q} + O_{i-h,q}) \underset{(a4)}{=} \delta_{p,q}(2-\delta_{h,0})\cos\frac{hq\pi}{n+1},$$

where $(a1)$ holds by the definition of $\mathbb{F}_m^h$, $(a2)$ holds by the definition of O, $(a3)$ is obvious, and $(a4)$ holds by the orthogonality of O. ∎

Lemma D3. *Suppose there exists a fixed $\alpha > 0$ such that $m\Delta_n^{1/2+\alpha} \to \infty$. Let $q_n n^{-1/2} \to 0$. Define $\mathbb{F}_m^h$ by (D.1). Under $(\sigma^2, \iota^2, \theta) \in \Pi(q_n)$, it holds*

$$V_m^{-1}(\sigma^2, \iota^2, \theta, \Delta_n) = \sum_{h=0}^{m}\rho_h(\sigma^2, \iota^2, \theta, \Delta_n)\mathbb{D}_m^h, \quad \Omega_m^{-1}(\sigma^2, \iota^2, \theta, \Delta_n) = \sum_{h=0}^{m}\rho_h(\sigma^2, \iota^2, \theta, \Delta_n)\mathbb{F}_m^h,$$

with

$$\rho_h(\sigma^2,\iota^2,\theta,\Delta_n)=\frac{1}{2\sigma\zeta\Delta_n^{1/2}}\left(1-\frac{\sqrt{\sigma^2\Delta_n}}{\zeta}+O(\Delta_n q_n)\right)^h\left(1+O\left(\sqrt{\Delta_n}q_n\right)\right),\quad \zeta^2=\iota^2 g(0;\theta) \tag{D.2}$$

Proof. Step 1. (Main proof) Given the expression of V_m^{-1}, that of Ω_m^{-1} directly follows by applying Lemma D2. So it suffices to derive V_m^{-1}.

First, for all $z\in\mathbb{C}$ with $z\neq 0$, let

$$\mathcal{V}(z;\sigma^2,\iota^2,\theta,\Delta_n):=\sigma^2\Delta_n+(2-z-z^{-1})\iota^2\theta(z)\theta(z^{-1}). \tag{D.3}$$

We prove in Step 2 below a key property:

$$\mathcal{V}(z;\sigma^2,\iota^2,\theta,\Delta_n)^{-1}=\sum_{h=-\infty}^{\infty}\rho_{|h|}(\sigma^2,\iota^2,\theta,\Delta_n)z^h, \tag{D.4}$$

where $\rho_h(\sigma^2,\iota^2,\theta,\Delta_n)$ satisfies (D.2).

Next, we connect V_m and $\mathcal{V}$. Indeed, in view of (D.3) and the definitions of V_m and $\mathbb{D}_m^h$, one can verify

$$V_m(\sigma^2,\iota^2,\theta,\Delta_n)_{kk}=\mathcal{V}\left(e^{\mathbb{i}\frac{k\pi}{m+1}};\sigma^2,\iota^2,\theta,\Delta_n\right),\quad \forall 1\le k\le m.$$

Hence, noting $(z^h+z^{-h})\left(1-\frac{1}{2}\delta_{h,0}\right)=(\mathbb{D}_m^h)_{kk}$ for $z=e^{\mathbb{i}\frac{k\pi}{m+1}}$ and that $V_m(\sigma^2,\iota^2,\theta,\Delta_n)$ is diagonal by construction, we readily deduce the desired result.

Step 2. (Proof of (D.4)) Throughout the rest of the proof we drop the arguments $(\sigma^2,\iota^2,\theta)$ whenever possible.

Step 3 below proves that for each n sufficiently large, there exists a unique complex number z_n^* such that

$$1-Kq_n^{-1}\le|z_n^*|\le 1\quad\text{and}\quad\mathcal{V}(z_n^*;\Delta_n)=0. \tag{D.5}$$

In other words, asymptotically z_n^* is the solution of $\mathcal{V}(z;\Delta_n)=0$ which is closest to the unit circle in the complex plane. Further, by noting that $\mathcal{V}(z;\Delta_n)=0$ indicates $\mathcal{V}(z^{-1};\Delta_n)=0$, we can define $\widetilde{\mathcal{V}}(z;\Delta_n)$ as

$$(z-z_n^*)\left(\frac{1}{z}-z_n^*\right)\widetilde{\mathcal{V}}(z;\Delta_n)=\mathcal{V}(z;\Delta_n), \tag{D.6}$$

where z can be any nonzero complex number. Moreover, let $\{\widetilde{\rho}_h(\Delta_n)\}_{h=-\infty}^{\infty}$ be the coefficients of Laurent expansion of $\widetilde{\mathcal{V}}(z;\Delta_n)^{-1}$:

$$\widetilde{\rho}_h(\Delta_n)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{e^{-ih\lambda}}{\widetilde{\mathcal{V}}(e^{\mathbb{i}\lambda};\Delta_n)}d\lambda. \tag{D.7}$$

Note $\rho_h(\Delta_n)=\rho_{-h}(\Delta_n)$ due to $\mathcal{V}(z;\Delta_n)=\mathcal{V}(z^{-1};\Delta_n)$.

Using the constructions of $\widetilde{\mathcal{V}}$ and $\widetilde{\rho}$, one can verify

$$\mathcal{V}(z;\Delta_n)^{-1} = \frac{1}{1-(z_n^*)^2}\sum_{h=-\infty}^{\infty} z^h \sum_{j=-\infty}^{\infty} \widetilde{\rho}_j(\Delta_n)(z_n^*)^{|j-h|}.$$

Recalling (D.4), all we need is to show

$$\rho_h(\sigma^2,\iota^2,\theta,\Delta_n) = \frac{1}{1-(z_n^*)^2}\sum_{j=-\infty}^{\infty} \widetilde{\rho}_j(\Delta_n)(z_n^*)^{|j-h|}$$

indeed satisfies (D.2). To this end, in Step 3 we derive the explict expression of z_n^*, up to $O(\Delta_n q_n)$:

$$z_n^* = \frac{2\zeta^2+\sigma^2\Delta_n}{2\zeta^2} - \frac{\sqrt{\sigma^2\Delta_n(4\zeta^2+\sigma^2\Delta_n)}}{2\zeta^2} + O(\Delta_n q_n). \tag{D.8}$$

Given this, the desired result follows from three facts. The first fact is that $\widetilde{\rho}_h(z_n^*;\Delta_n)$ shrinks faster than $(z_n^*)^h$ as h increases. Formally, for all $h \geq 0$

$$|\widetilde{\rho}_h(\Delta_n)| \leq K q_n (1-K^{-1}q_n^{-1})^h, \tag{D.9}$$

which holds by Step 4. We recall $q_n n^{-1/2} \to 0$. The second fact is that

$$\sum_{h=-\infty}^{+\infty} \widetilde{\rho}_h(\Delta_n) = \widetilde{\mathcal{V}}(1;\Delta_n) = \mathcal{V}(1;\Delta_n)(1-z_n^*)^{-2},$$

which holds by (D.6) and (D.7). The last fact is that $\sum_{j=m}^{\infty}(z_n^*)^j \lesssim n^{-\alpha'}$ for all positive α', which holds by $m\Delta_n^{1/2+\alpha} \to \infty$ and (D.8).

Step 3. (Auxiliary: Behavior of the zeros of $\mathcal{V}$) In this step we show that there is a unique z_n^* satisfying (D.5) and the expression of z_n^* is given by (D.8). To obtain the expression, simply write $z_n^* = 1 + a\sqrt{\Delta_n} + b\Delta_n + \ldots$ and match the coefficients to let $\mathcal{V}(z_n^*;\Delta_n) = 0$. Hence we only need to show the uniqueness.

First, we show there exists a unique real solution within $[1-Kq_n^{-1}, 1]$. Observing $\mathcal{V}(1;\Delta_n) = \sigma^2\Delta_n > 0$, it follows from

$$\inf_{z\in[1-Kq_n^{-1},1]} \frac{1}{(1-z)}\frac{d\mathcal{V}(z;\Delta_n)}{dz} \geq K^{-1},$$

which holds by

$$\sup_{z\in\mathbb{C}:|z|\leq 1+Kq_n^{-1}} \left|\frac{d\theta(z)}{dz}\right| \leq K \tag{D.10}$$

due to the definition of $\Pi(q_n)$.

Second, we show that z_n^* must be real for large n. Suppose $z_n^* = |z_n^*|e^{i\varphi_n^*}$ with $\varphi \in [0, 2\pi)$. We prove $\varphi_n^* = 0$ by contradiction. In view of the fact that for all $|z| = 1$, $\mathcal{V}(z;\Delta_n) \geq \sigma^2\Delta_n > 0$ by construction, it suffices to show that $\varphi_n^* \neq 0$ indicates $|z_n^*| = 1$. To see this, write

$$\begin{aligned}\operatorname{Im}(\mathcal{V}(z_n^*;\Delta_n)) =& -\sin\varphi_n^* \underbrace{\left(|z_n^*| - \frac{1}{|z_n^*|}\right)}_{=:\mathcal{R}_a} \iota^2 \operatorname{Re}\left(\theta(z_n^*)\theta\left(\frac{1}{z_n^*}\right)\right) \\ &+ \cos\varphi_n^* \underbrace{\left(2 - |z_n^*| - \frac{1}{|z_n^*|}\right)}_{=:\mathcal{R}_b} \iota^2 \operatorname{Im}\left(\theta(z_n^*)\theta\left(\frac{1}{z_n^*}\right)\right).\end{aligned}$$

The desired result follows from two facts. The first is that $\operatorname{Im}(\mathcal{V}(z_n^*;\Delta_n)) = 0$ due to $\mathcal{V}(z_n^*;\Delta_n) = 0$. The second is that asymptotically $\mathcal{R}_a$ dominates $\mathcal{R}_b$. We see this from

$$\operatorname{Re}\left(\theta(z_n^*)\theta\left(\frac{1}{z_n^*}\right)\right) \geq K^{-1} \quad \text{and} \quad \left|\operatorname{Im}\left(\theta(z_n^*)\theta\left(\frac{1}{z_n^*}\right)\right)\right| \leq (1 - |z_n^*|) \times |\sin\varphi_n^*|.$$

Recalling $1 - Kq_n^{-1} \leq |z_n^*| \leq 1$, the two inequalities hold by writing $\theta(z_n^*)\theta\left(\frac{1}{z_n^*}\right)$ as an Taylor expansion around $e^{\mathbb{i}\varphi_n^*}$ and (D.10).

Step 4. (Auxiliary: Bound on $\widetilde{\rho}_h$) In this step we prove (D.9).

According to Theorem 4.1.1, Proposition 4.5.3, Proposition 3.2.1. and Theorem 3.1.2. in Brockwell and Davis (1991), for each n and $(\sigma^2, \iota^2, \theta) \in \Pi(q_n)$, $\widetilde{\mathcal{V}}(e^{\mathbb{i}\lambda};\Delta_n)$ is the spectral density function of an invertible moving-average process of order at most q_n. Therefore, there exists unique $\eta(\Delta_n) \in [0,\infty)$ and $\nu(\Delta_n) \in \mathbb{R}^{q_n}$ such that

$$\eta^2(\Delta_n) g(\lambda;\nu(\Delta_n)) = \widetilde{\mathcal{V}}(e^{\mathbb{i}\lambda};\Delta_n) \quad \text{and} \quad \inf_{z\in\mathbb{C}:|z|\leq 1}\left|1 + \sum_{j=1}^{q_n} \nu_j(\Delta_n) z^j\right| > 0.$$

We hence can further define $\{\widetilde{\nu}_j(\Delta_n)\}_{j=1}^{\infty}$ by letting

$$1 + \sum_{j=1}^{\infty} \widetilde{\nu}_j(\Delta_n) e^{\mathbb{i}j\lambda} = \frac{1}{1 + \sum_{j=1}^{q_n} \nu_j(\Delta_n) e^{\mathbb{i}j\lambda}} \quad \text{for all} \quad \lambda \in (-\pi,\pi).$$

In view of (D.5) and (D.6), $\widetilde{\mathcal{V}}(z;\Delta_n)$ is never zero for z satisfying $1 - Kq_n^{-1} \leq |z| \leq 1 + Kq_n^{-1}$. This indicates $\sum_{j=1}^{\infty} \widetilde{\nu}_j z^j$ converges for $1 - Kq_n^{-1} \leq |z| \leq 1 + Kq_n^{-1}$, hence $|\widetilde{\nu}_j(\Delta_n)| \leq K(1 - K^{-1}q_n^{-1})^j$. Moreover, one can also show $\eta^2(\Delta_n) \geq K^{-1}$. The desired result follows by noting $\widetilde{\rho}_h(\Delta_n) = \eta^{-2}(\Delta_n)\sum_{j=0}^{\infty} \nu_j(\Delta_n)\nu_{j+h}(\Delta_n)$, where $\nu_0(\Delta_n) = 1$ by convention. ∎

Lemma D4. *Suppose there exists a fixed $\alpha > 0$ such that $m\Delta_n^{1/2+\alpha} \to \infty$. Let q be fixed. Let (σ^2, θ) and $\{\iota_n^2 : n \geq 1\}$ satisfy $\frac{1}{K} \leq \sigma^2 \leq K, \theta \in \Theta(q)$ and $\iota_n^2 \leq K$ for each n. Define $\mathbb{F}_m^h$ by (D.1). It holds, for some $\{\rho_h(\sigma^2, \iota_n^2, \theta, \Delta_n)\}_{h=0}^{\infty}$,*

$$V_m^{-1}(\sigma^2, \iota_n^2, \theta, \Delta_n) = \sum_{h=0}^{m} \rho_h(\sigma^2, \iota_n^2, \theta, \Delta_n)\mathbb{D}_m^h, \quad \Omega_m^{-1}(\sigma^2, \iota_n^2, \theta, \Delta_n) = \sum_{h=0}^{m} \rho_h(\sigma^2, \iota_n^2, \theta, \Delta_n)\mathbb{F}_m^h.$$

In addition, (i) under $\Delta_n^{-1}\iota_n^2 \to \infty$, it holds

$$\rho_h(\sigma^2, \iota_n^2, \theta, \Delta_n) = \frac{1}{2\sigma\zeta_n\Delta_n^{1/2}}\left(1 - \frac{\sqrt{\sigma^2\Delta_n}}{\zeta_n} + O(\Delta_n\iota_n^{-2})\right)^h \left(1 + O\left(\sqrt{\Delta_n\iota_n^{-2}}\right)\right),$$

with $\zeta_n^2 = \iota_n^2 g(0;\theta)$; and (ii) under $\Delta_n^{-1}\iota_n^2 \to a^2 \in [0,\infty)$, it holds

$$\rho_h(\sigma^2, \iota_n^2, \theta, \Delta_n) = \frac{1}{2\pi\Delta_n}\int_{-\pi}^{\pi} \frac{e^{ih\lambda}}{\sigma^2 + a^2|1-e^{\mathrm{i}\lambda}|^2 g(\lambda;\theta)} d\lambda = O\big(\Delta_n^{-1}(1-\alpha')^h\big),$$

for some fixed $\alpha' > 0$.

Proof. The desired results in both cases follow from the same argument of the proof of Lemma D3. Case (i) is a direct result of using $\Delta_n\iota_n^{-2} \to 0$ and noting that $\mathcal{V}(\sigma^2, \iota^2, \theta, \Delta_n)$ is homogeneous of degree one in ι^2 and Δ_n. Case (ii) is even simpler by observing

$$\Delta_n^{-1}\mathcal{V}(\sigma^2, \iota_n^2, \theta, \Delta_n) = \mathcal{V}(\sigma^2, \Delta_n^{-1}\iota_n^2, \theta, 1) = \mathcal{V}(\sigma^2, a^2, \theta, 1) + o(1).$$

∎

Lemma D5. *Suppose $(\sigma^2, \iota^2, \theta) \in \Pi(q)$ with q fixed. As $n \to \infty$, it holds (i)*

$$V_n(\sigma^2, \iota^2, \theta, \Delta_n) = \frac{1}{\widetilde{\zeta}^2} V_n(\widetilde{\sigma}^2, \widetilde{\zeta}^2, 0, \Delta_n) O_n D_n(\widetilde{\iota}^2, \widetilde{\theta}) O_n, \text{ with } |\widetilde{\iota}^2 - \iota^2| + \|\widetilde{\theta} - \theta\| + |\widetilde{\zeta}^2 - \iota^2 g(0;\theta)| \lesssim n^{-1/2},$$

and (ii)

$$\Omega_n^{-1}(\sigma^2, \zeta^2, 0)_{ij} = b_n(z_n^{|i-j|} - z_n^{i+j} - z_n^{2n+2-i-j}) + O(n^{-\infty}),$$

where b_n and z_n do not depend on i or j, and they satisfy

$$b_n = \frac{1}{2\sigma\zeta\Delta_n^{1/2}} + O(1), \quad z_n = 1 - \frac{\sqrt{\sigma^2\Delta_n}}{\zeta} + O(n^{-1}).$$

Proof. The desired results follow from a simplified version of the proof of Lemma D3. Concretely, the fact that q is finite allows us to solve explicitly the $q+1$ zeros of $\mathcal{V}(z; \sigma^2, \iota^2, \theta, \Delta_n)$, up to $O(n^{-1})$. Hence, we do not need Steps 2 - 4 there and directly obtain the first result. The second result follows by additionally applying Lemma D2. ∎

Lemma D6. *Suppose there exists a fixed $\alpha > 0$ such that $\Delta_n^{1/2+\alpha} m \to \infty$. Under $(\sigma^2, \iota^2, \theta) \in \Pi(q_n)$ and omitting the arguments $(\sigma^2, \iota^2, \theta, \Delta_n)$ of V_m and D_m, it holds, (i)*

$$j \in \{1,2,3,4\} \quad \Longrightarrow \quad tr(V_m^{-j}) = \frac{\lambda_j m}{\zeta(\sigma^2\Delta_n)^{j-1/2}} - \frac{1}{2(\sigma^2\Delta_n)^j} + O(mn^{j-1}q_n),$$

with

$$\lambda_1 = \frac{1}{2}, \quad \lambda_2 = \frac{1}{4}, \quad \lambda_3 = \frac{3}{16}, \quad \lambda_4 = \frac{5}{32};$$

and (ii) for $1 \leq i, j \leq q_n$,

$$tr\left(V_m^{-1}\frac{\partial V_m}{\partial \theta_j}\right) = O(n^{-1}q_n), \quad tr\left(V_m^{-2}\frac{\partial V_m}{\partial \theta_i}\frac{\partial V_m}{\partial \theta_j}\right) = 2W(\theta)_{ij} + O(n^{-1}q_n).$$

Proof. The desired result in (i) follows from Lemma D3 and the fact that

$$\frac{\partial V_m^{-j}}{\partial \sigma^2} = -j\Delta_n V_m^{-j-1}.$$

We only show the second result in (ii) as the other one is simpler. For $l \in \{0, 1, 2\}$, let

$$\mathcal{R}_a(i,j,l) = \mathrm{tr}\left(V_m^{-l}D_m^{-2}\frac{\partial D_m}{\partial \theta_i}\frac{\partial D_m}{\partial \theta_j}\right).$$

It follows from, for all $1 \leq i, j \leq q_n$, (a)

$$\mathrm{tr}\left(V_m^{-2}\frac{\partial V_m}{\partial \theta_i}\frac{\partial V_m}{\partial \theta_j}\right) = \mathcal{R}_a(i,j,0) - 2\sigma^2\Delta_n\mathcal{R}_a(i,j,1) + \sigma^4\Delta_n^2\mathcal{R}_a(i,j,2),$$

which holds by the definition of V_m; (b)

$$l \in \{1, 2\} \quad \Longrightarrow \quad \mathcal{R}_a(i,j,l) = O(n^{j-1/2} + n^{j-1}q_n),$$

which holds by the result in (i) and the fact that $\sup_{i,k}|(D_m^{-1})_{kk}| \vee \left|\frac{\partial (D_m)_{kk}}{\partial \theta_i}\right| \leq K$; and (c)

$$\mathcal{R}_a(i,j,0) = 2W(\theta)_{ij} + O(n^{-1}q_n),$$

which holds by

$$\left(D_m^{-1}\frac{\partial D_m}{\partial \theta_j}\right)_{kk} = \frac{\partial}{\partial \theta_j}\log g\left(e^{\mathbb{i}\frac{k}{m+1}};\theta\right) \quad \text{and} \quad \sum_{k=1}^{m}\left(e^{\mathbb{i}\frac{k}{m+1}}\right)^j = m\delta_{j,0} + O(1) = \frac{m}{2\pi}\int_{-\pi}^{\pi}(e^{\mathbb{i}\lambda})^j d\lambda + O(1).$$

∎

References

Belloni, A., V. Chernozhukov, and C. Hansen (2014): "Inference on Treatment Effects after Selection among High-Dimensional Controls," *Review of Economic Studies*, 81, 608–650.

Brockwell, P. J., and R. A. Davis (1991): *Time Series: Theory and Methods*. Springer-Verlag, New York, second edn.

Horn, R. A., and C. R. Johnson (2013): *Matrix Analysis*. Cambridge University Press, second edn.

Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter (2009): "Microstructure Noise in the Continuous Case: The Pre-Averaging Approach," *Stochastic Processes and Their Applications*, 119, 2249–2276.

Jacod, J., and P. Protter (2011): *Discretization of Processes*. Springer-Verlag.

Leeb, H., and B. M. Pötscher (2006): "Performance limits for estimators of the risk or distribution of shrinkage-type estimators, and some general lower risk-bound results," *Econometric Theory*, 22(1), 69–97.

Newey, W. K. (1991): "Uniform convergence in probability and stochastic equicontinuity," *Econometrica: Journal of the Econometric Society*, pp. 1161–1167.

Reiss, M. (2011): "Asymptotic equivalence for inference on the volatility from noisy observations," *Annals of Statistics*, 39, 772–802.

van der Vaart, A. W. (2000): *Asymptotic statistics*. Cambridge university press.