Supplementary Materials to "On estimation of nonparametric regression models with autoregressive and moving average errors"

Qi Zheng **Yunwei Cui** **Rongning Wu**

A. Proofs of Propositions 1–3

Proposition 1 *Suppose Conditions (C1) - (C4) hold. There exists some constants δ_1 and δ_2, such that for all $\|\boldsymbol{\beta}-\boldsymbol{\beta}_*\| \le \delta_1$, $\|(\boldsymbol{\phi}^{\mathrm{T}}, \boldsymbol{\theta}^{\mathrm{T}}) - (\boldsymbol{\phi}_*^{\mathrm{T}}, \boldsymbol{\theta}_*^{\mathrm{T}})\| \le \delta_2$,*

(i) $|\zeta_t| \le \eta_t$, $|\zeta_t(\boldsymbol{\xi}_*) - \boldsymbol{\phi}_*(B)\boldsymbol{\theta}_*^{-1}(B)R_t - \zeta_t| \le r^t\eta_0$, $|\zeta(\boldsymbol{\xi})| \le \eta_t + C_2(\Delta+\delta_1)$, *and* $|\zeta_t(\boldsymbol{\xi}) - \zeta_t(\boldsymbol{\xi}_*)| \le C_3\delta_2\eta_t + C_2C_3\delta_2(\delta_1+\Delta) + C_2\delta_1$,

(ii) $\|\mathbf{D}_t(\boldsymbol{\xi})\|_\infty \le \omega_t$, $\mathbf{D}_{t1}(\boldsymbol{\xi}_*) - \mathbf{Q}_{t1} = \mathbf{0}$, *and* $\left\|\left(\mathbf{D}_{t2}^{\mathrm{T}}(\boldsymbol{\xi}_*), \mathbf{D}_{t3}^{\mathrm{T}}(\boldsymbol{\xi}_*)\right) - (\mathbf{Q}_{t2}^{\mathrm{T}}, \mathbf{Q}_{t3}^{\mathrm{T}})\right\|_\infty \le r^t\eta_0 + C_2\Delta$,

(iii) $\|\mathbf{H}_t(\boldsymbol{\xi})\|_{\max} \le \omega_t$, $\mathbf{H}_{t,11}(\boldsymbol{\xi}_*) - \mathbf{V}_{t,11} = \mathbf{0}$, *and* $\|\mathbf{H}_t(\boldsymbol{\xi}_*) - \mathbf{V}_t\|_{\max} \le r^t\eta_0 + C_2\Delta$,

where $\eta_t = C_1\sum_{j=0}^\infty r^j|\epsilon_{t-j}|$, $\omega_t = \max\left\{C_2, r^{-(p+q)}\eta_t + C_2(\Delta+\delta_1)\right\}$, *and C_3 is defined in Lemma 7.*

Proof of Proposition 1: By Condition (C2), the roots of $\boldsymbol{\theta}_*(z)$ lie in the region $|z| > c$ for some $|c| > 1$. The same will be true for the roots of $\theta(z)$ for all $(\boldsymbol{\phi}^{\mathrm{T}}, \boldsymbol{\theta}^{\mathrm{T}})^{\mathrm{T}}$ which are sufficiently close to $(\boldsymbol{\phi}_*^{\mathrm{T}}, \boldsymbol{\theta}_*^{\mathrm{T}})^{\mathrm{T}}$.

(i) By definition, $\theta(B)\zeta_t(\boldsymbol{\xi}) = \phi(B)\epsilon_t(\boldsymbol{\beta})$. Following Section 3.1 in [2] and Proposition A.1. of [3], there exists some $\delta_2 > 0$ such that the coefficients $\pi_j(\boldsymbol{\phi}, \boldsymbol{\theta})$ in the power series expansion
$$\sum_{j=0}^\infty \pi_j(\boldsymbol{\phi}, \boldsymbol{\theta})z^j = \frac{1-\phi_1 z - \cdots - \phi_p z^p}{1+\theta_1 z + \cdots + \theta_q z^q}$$
are bounded by $C_1 r^j$, if $\|(\boldsymbol{\phi}^{\mathrm{T}}, \boldsymbol{\theta}^{\mathrm{T}}) - (\boldsymbol{\phi}_*^{\mathrm{T}}, \boldsymbol{\theta}_*^{\mathrm{T}})\| \le \delta_2$. Thus, $|\zeta_t| = \left|\boldsymbol{\phi}_*(B)\boldsymbol{\theta}_*^{-1}(B)\epsilon_t\right| \le C_1\sum_{j=0}^\infty r^j|\epsilon_{t-j}| = \eta_t$. Moreover,
$$\begin{aligned}
&\left|\zeta_t(\boldsymbol{\xi}_*) - \frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}R_t - \zeta_t\right| \\
=& \left|\sum_{j=0}^{t-1}\pi_j(\boldsymbol{\phi}_*, \boldsymbol{\theta}_*)(\epsilon_{t-j}(\boldsymbol{\beta}_*) - R_{t-j}) - \sum_{j=0}^\infty \pi_j(\boldsymbol{\phi}_*, \boldsymbol{\theta}_*)\epsilon_{t-j}\right| \\
\le& \left|\sum_{j=t}^\infty \pi_j(\boldsymbol{\phi}_*, \boldsymbol{\theta}_*)\epsilon_{t-j}\right| \le C_1\sum_{j=t}^\infty r^j|\epsilon_{t-j}| \le r^t\eta_0.
\end{aligned}$$

In addition,

$$\begin{aligned}
|\zeta_t(\boldsymbol{\xi})| &= \left|\frac{\boldsymbol{\phi}(B)}{\boldsymbol{\theta}(B)}\epsilon_t(\boldsymbol{\beta})\right| = \left|\sum_{j=0}^{t-1}\pi_j(\boldsymbol{\phi},\boldsymbol{\theta})\epsilon_{t-j}(\boldsymbol{\beta})\right| \\
&\le \left|\sum_{j=0}^{t-1}\pi_j(\boldsymbol{\phi},\boldsymbol{\theta})\Big(Y_{t-j}-g_0(X_{t-j})\Big)\right| + \left|\sum_{j=0}^{t-1}\pi_j(\boldsymbol{\phi},\boldsymbol{\theta})R_{t-j}\right| \\
&\qquad + \left|\sum_{j=0}^{t-1}\pi_j(\boldsymbol{\phi},\boldsymbol{\theta})\mathbf{W}_{t-j}^{\mathrm{T}}\Big(\boldsymbol{\beta}_*-\boldsymbol{\beta}\Big)\right| \\
&\le C_1\sum_{j=0}^{t-1} r^j\,|\epsilon_{t-j}| + C_1\sum_{j=0}^{t-1} r^j\Delta + C_1\sum_{j=0}^{t-1} r^j\delta_1 \le \eta_t + C_2(\Delta+\delta_1),
\end{aligned}$$

where the second to the last inequality follows from Theorem 5.4.2 in [5] with l_∞ norm and $\|\boldsymbol{\beta}-\boldsymbol{\beta}_*\|_\infty \le \|\boldsymbol{\beta}-\boldsymbol{\beta}_*\| \le \delta_1$.

By Lemma 7,

$$\begin{aligned}
&|\zeta_t(\boldsymbol{\xi})-\zeta_t(\boldsymbol{\xi}_*)| \le \left|\left(\frac{\boldsymbol{\phi}(B)}{\boldsymbol{\theta}(B)}-\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}\right)\epsilon_t(\boldsymbol{\beta})\right| + \left|\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}\left(\epsilon_t(\boldsymbol{\beta})-\epsilon_t(\boldsymbol{\beta}_*)\right)\right| \\
&\le \left|\left(\frac{\boldsymbol{\phi}(B)}{\boldsymbol{\theta}(B)}-\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}\right)\Big(\epsilon_t - \big(\epsilon_t-\epsilon_t(\boldsymbol{\beta}_*)\big) - \big(\epsilon_t(\boldsymbol{\beta}_*)-\epsilon_t(\boldsymbol{\beta})\big)\Big)\right| + C_2\delta_1 \\
&\le C_3\delta_2\eta_t + C_2C_3\delta_2(\delta_1+\Delta) + C_2\delta_1
\end{aligned}$$

(ii) For $1 \le l \le J$,

$$|[\mathbf{D}_t(\boldsymbol{\xi})]_l| = \left|\frac{\boldsymbol{\phi}(B)}{\boldsymbol{\theta}(B)}[\mathbf{W}_t]_l\right| = \left|\sum_{j=0}^{t-1}\pi_j(\boldsymbol{\phi},\boldsymbol{\theta})\,[\mathbf{W}_{t-j}]_l\right| \le C_1\sum_{j=0}^{\infty} r^j\,\|\mathbf{W}_{t-j}\|_\infty \le C_2,$$

where the last inequality also follows from Theorem 5.4.2 in [5] with l_∞ norm.

For $J+1 \le l \le J+p$, let $l_1 = l-J$. We have $1 \le l_1 \le p$. Then,

$$\begin{aligned}
|[\mathbf{D}_t(\boldsymbol{\xi})]_l| &= \left|\frac{1}{\boldsymbol{\theta}(B)}\epsilon_{t-l_1}(\boldsymbol{\beta})\right| \\
&\le C_1\sum_{j=0}^{t-l_1-1} r^j\Big(|\epsilon_{t-l_1-j}| + |\epsilon_{t-l_1-j}(\boldsymbol{\beta}_*)-\epsilon_{t-l_1-j}| \\
&\qquad + |\epsilon_{t-l_1-j}(\boldsymbol{\beta})-\epsilon_{t-l_1-j}(\boldsymbol{\beta}_*)|\Big) \\
&\le C_1\sum_{j=0}^{t-l_1-1} r^j\,|\epsilon_{t-l_1-j}| + C_2(\Delta+\delta_1) \le r^{-p}\eta_t + C_2(\Delta+\delta_2).
\end{aligned}$$

For $J+p+1 \le l \le J+p+q$, by the same arguments, $|[\mathbf{D}_t(\boldsymbol{\xi})]_l| \le r^{-q}\eta_t + C_2\left(\Delta+\delta_1\right).$

Thus, $\|\mathbf{D}_t(\boldsymbol{\xi})\|_\infty \le \omega_t$.

For $1 \le l \le J$, $[\mathbf{D}_t(\boldsymbol{\xi}_*)]_l = \left[\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}\mathbf{W}_t\right]_l = [\mathbf{Q}_t]_l$. For $J+1 \le l \le J+p$, let $l_1 = l - J$.

$$\begin{aligned}|[\mathbf{D}_t(\boldsymbol{\xi}_*)]_l - [\mathbf{Q}_t]_l| &= \left|\frac{1}{\boldsymbol{\theta}_*(B)}\epsilon_{t-l_1}(\boldsymbol{\beta}_*) - \frac{1}{\boldsymbol{\phi}_*(B)}\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}\epsilon_{t-l_1}\right| \\ &\le C_1 \sum_{j=0}^{t-l_1-1} r^j \left|\epsilon_{t-l_1-j}(\boldsymbol{\beta}_*) - \epsilon_{t-l_1-j}\right| + C_1 \sum_{j=t-l_1}^{\infty} r^j \left|\epsilon_{t-l_1-j}\right| \le r^t\eta_0 + C_2\Delta,\end{aligned}$$

For $J+1 \le l \le J+p$, the argument is similar. Therefore, we have $\|\mathbf{D}_t(\boldsymbol{\xi}_*) - \mathbf{Q}_t\|_\infty \le r^t\eta_0 + C_2\Delta$.

(iii) The proof (iii) follow from the same arguments as used for (ii), We thus omit the details here.

We thus complete the proof of Proposition 1. □

Proposition 2 *Suppose Conditions (C1)–(C4) hold. If $J = n^{1/(2\alpha+1)}$, for any $C > 0$,*

$$\sup_{\mathbf{h}\in\Omega(C)} |T_1(\mathbf{h}) - T(\mathbf{h})| \to_p 0.$$

Proof of Proposition 2: If $J = n^{1/(2\alpha+1)}$, then $J^2 \log n = n^{2/(2\alpha+1)} \log n = o(n^{1/2})$, and $J^{-2\alpha+1/2} = n^{(-2\alpha+1/2)/(2\alpha+1)} = o(n^{-1/2})$, as $\alpha \ge 2$. Thus, the conditions in Lemmas 4–6 are satisfied. The proof directly follows from Lemmas 4–6. □

Proposition 3 *Under the same conditions as in Proposition 2, given any $0 < \varepsilon < 1$, there exists some $C_\varepsilon > 0$, such that*

$$P\left(\inf_{\mathbf{h}\in\bar{\Omega}(C_\varepsilon)\bigcup\Omega^c(C_\varepsilon)} T_1(\mathbf{h}) > 1\right) > 1 - \varepsilon.$$

Proof of Proposition 3: If $J = O(n^{1/(2\alpha+1)})$, then $J^3 \log n = o(n)$ and $J^{-(\alpha+1/2)} = O(n^{-1/2})$ under Condition (C1).

Simple algebra shows that

$$\begin{aligned}T_1(\mathbf{h}) &= -2\sum_{t=1}^{n} \mathbf{h}^{\mathrm{T}}\mathbf{Q}_t\zeta_t - 2\sum_{t=1}^{n} \mathbf{h}^{\mathrm{T}}\mathbf{Q}_t\left(\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}R_t\right) + \mathbf{h}^{\mathrm{T}}\sum_{t=1}^{n}(\zeta_t\mathbf{Q}_t)(\zeta_t\mathbf{Q}_t)^{\mathrm{T}}\mathbf{h} \\ &=: I + II + III.\end{aligned}$$

We first consider I.

$$-2\sum_{t=1}^{n} \mathbf{h}^{\mathrm{T}}\mathbf{Q}_t\zeta_t = -2\sum_{t=1}^{n} \mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\zeta_t - 2\sum_{t=1}^{n}\left(\mathbf{h}_2^{\mathrm{T}}\mathbf{Q}_{t2} + \mathbf{h}_3^{\mathrm{T}}\mathbf{Q}_{t3}\right)\zeta_t =: I_1 + I_2.$$

We have

$$
\begin{aligned}
&E\left[\sup_{\|\mathbf{h}_1\|\le 1}\left(\sum_{t=1}^n \mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\zeta_t\right)^2\right]\\
&\le E\left[\sup_{\|\mathbf{h}_1\|\le 1}\|\mathbf{h}_1\|^2\left\|\sum_{t=1}^n \mathbf{Q}_{t1}\zeta_t\right\|^2\right] = \sigma^2 E\left[\sum_{t=1}^n \mathbf{Q}_{t1}^{\mathrm{T}}\mathbf{Q}_{t1}\right] = n\sigma^2 E\left[\mathbf{Q}_{t1}^{\mathrm{T}}\mathbf{Q}_{t1}\right]\\
&= n\sigma^2 \times \mathrm{trace}(E\left[\mathbf{Q}_{t1}\mathbf{Q}_{t1}^{\mathrm{T}}\right]) \le nJC_2^2\lambda_{\max}J^{-1}\sigma^2 = nC_2^2\lambda_{\max}\sigma^2,
\end{aligned}
$$

where the first equality follows from Cauchy-Schwartz inequality, the third equality follows from the fact that $E\left[\mathbf{Q}_{t1}^{\mathrm{T}}\mathbf{Q}_{t1}\right]$ is the trace of $E\left[\mathbf{Q}_{t1}\mathbf{Q}_{t1}^{\mathrm{T}}\right]$, and the last inequality follows from Proposition 4. Therefore, by Markov inequality,

$$
P\left(\sup_{\|\mathbf{h}_1\|\le 1}\left|\sum_{t=1}^n \mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\zeta_t\right| > aC_2\sqrt{n\lambda_{\max}}\sigma\right) \le \frac{nC_2^2\lambda_{\max}\sigma^2}{a^2nC_2^2\lambda_{\max}\sigma^2} = \frac{1}{a^2}.
$$

Thus, $\sup_{\|\mathbf{h}_1\|\le C,\mathbf{h}_1\neq 0}(nJ^{-1})^{-1/2}\|\mathbf{h}_1\|^{-1}|I_1| = CO_p(1)$. Similar arguments can be applied to show that $\sup_{\|(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}})\|\le C,(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}})\neq 0} n^{-1/2}\|(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}})\|^{-1}|I_2| = CO_p(1)$.

Next, we evaluate II.

$$
\begin{aligned}
II &= -2\sum_{t=1}^n \mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\left(\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}R_t\right) - 2\sum_{t=1}^n(\mathbf{h}_2^{\mathrm{T}}\mathbf{Q}_{t2}+\mathbf{h}_3^{\mathrm{T}}\mathbf{Q}_{t3})\left(\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}R_t\right)\\
&=: II_1 + II_2.
\end{aligned}
$$

We can show that $\sup_{\|\mathbf{h}_1\|\le 1,\mathbf{h}_1\neq 0}\left|\left(E\left[\left|\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right|\right]\right)^{-1}\mathbb{E}_n\left[\left|\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right|\right]-1\right| \to_p 0$, by the same arguments used in the proof of Lemma 2. This and the facts that $|R_t| \le \Delta \le C_0J^{-\alpha}$ and $|\boldsymbol{\phi}_*(B)\boldsymbol{\theta}_*^{-1}(B)R_t| \le C_0C_2J^{-\alpha}$ together yield that

$$
\begin{aligned}
&\sup_{\|\mathbf{h}_1\|\le C,\mathbf{h}_1\neq 0}\|\mathbf{h}_1\|^{-1}|II_1| = 2\sup_{\|\mathbf{h}_1\|\le 1,\mathbf{h}_1\neq 0}\left|\sum_{t=1}^n \mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\left(\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}R_t\right)\right|\\
&\le 2C_0C_2J^{-\alpha}n\sup_{\|\mathbf{h}_1\|\le 1,\mathbf{h}_1\neq 0}\mathbb{E}_n\left[\left|\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right|\right]\\
&\le 2C_0C_2J^{-\alpha}n\sup_{\|\mathbf{h}_1\|\le 1,\mathbf{h}_1\neq 0}E\left[\left|\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right|\right](1+o_p(1))\\
&\le 2C_0C_2J^{-\alpha}n\sup_{\|\mathbf{h}_1\|\le 1,\mathbf{h}_1\neq 0}\sqrt{E\left[\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\mathbf{Q}_{t1}^{\mathrm{T}}\mathbf{h}_1\right]}(1+o_p(1))\\
&\le 2C_0C_2^2\sqrt{\lambda_{\max}}J^{-\alpha-1/2}n(1+o_p(1)).
\end{aligned}
$$

Likewise, $\sup_{\|(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}})\|\le C,(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}})\neq 0}\|(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}})\|^{-1}|II_2| \le 2C_5J^{-\alpha}n(1+o_p(1))$, for some constant C_5 that only depends on Σ and σ.

Now we assess III. Let $\tilde{\lambda}_{\min}$ denote the smallest eigenvalue of $\mathbf{\Sigma}$.

$$
\begin{aligned}
& n\boldsymbol{\xi}^{\mathrm{T}} E\left[\left(\zeta_t \mathbf{Q}_t\right)\left(\zeta_t \mathbf{Q}_t\right)^{\mathrm{T}}\right] \boldsymbol{\xi} \\
& = n\sigma^2 \mathbf{h}_1^{\mathrm{T}} E\left[\mathbf{Q}_{t1}\mathbf{Q}_{t1}^{\mathrm{T}}\right]\mathbf{h}_1 + n\sigma^2\left(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}}\right)\mathbf{\Sigma}\left(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}}\right)^{\mathrm{T}} \\
& \geq n\sigma^2\left(J^{-1}\lambda_{\min}\|\mathbf{h}_1\|^2 + \tilde{\lambda}_{\min}\|(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}})\|^2\right).
\end{aligned}
$$

By Lemmas 2 and 3, uniformly over $\|\mathbf{h}_1\| \leq C$ and $\|\left(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}}\right)\| \leq C$,

$$
\begin{aligned}
III \geq & n\sigma^2\left(J^{-1}\lambda_{\min}\|\mathbf{h}_1\|^2 + \tilde{\lambda}_{\min}\|(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}})\|^2\right) \\
& - 2\|\mathbf{h}_1\|\|(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}})\|)O_p(J^{1/2}n^{1/2}\log n).
\end{aligned}
$$

Combining the results for I, II, and III yields that

$$
\begin{aligned}
& T_1(\mathbf{h}) \\
& \geq n\sigma^2\left(J^{-1}\lambda_{\min}\|\mathbf{h}_1\|^2 + \tilde{\lambda}_{\min}\|(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}})\|^2\right) \\
& \quad - 2\|\mathbf{h}_1\|\|(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}})\|O_p\left(J^{1/2}n^{1/2}\log n\right) - n^{1/2}\left(\|\mathbf{h}_1\| + \|\left(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}}\right)\|\right)O_p(1) \\
& \quad - 2C_0C_2^2\sqrt{\lambda_{\max}}J^{-\alpha-1/2}n\|\mathbf{h}_1\|(1+o_p(1)) - 2C_5J^{-\alpha}n\|(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}})\|,
\end{aligned}
$$

uniformly over $\|\mathbf{h}_1\| \leq C$ and $\|\left(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}}\right)\| \leq C$. Then,

$$
\begin{aligned}
& \inf_{\|\mathbf{h}_1\| = CJn^{-1/2}, \|\left(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}}\right)\| \leq CJ^{1/2}n^{-1/2}} T_1(\mathbf{h}) \\
\geq & n\sigma^2C^2\lambda_{\min}Jn^{-1}(1+o_p(1)) - C^2J^2n^{-1/2}\log nO_p(1) - 2CJO_p(1) \\
& - 2C\left(C_0C_2^2\sqrt{\lambda_{\max}} + C_5\right)J^{-\alpha+1/2}n^{1/2}(1+o_p(1)), \text{ and} \\
& \inf_{\|\mathbf{h}_1\| \leq CJn^{-1/2}, \|\left(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}}\right)\| = CJ^{1/2}n^{-1/2}} T_1(\mathbf{h}) \\
\geq & n\sigma^2C^2\tilde{\lambda}_{\min}Jn^{-1}(1+o_p(1)) - 2C^2J^2n^{-1/2}\log nO_p(1) - 2CJO_p(1) \\
& - 2C\left(C_0C_2^2\sqrt{\lambda_{\max}} + C_5\right)J^{-\alpha+1/2}n^{1/2}(1+o_p(1)).
\end{aligned}
$$

Since $J^{-(\alpha+1/2)} = O(n^{-1/2})$, then given any $0 < \varepsilon < 1$, there exists some sufficiently large C_ε, such that $P\left(\inf_{\mathbf{h}\in\bar{\Omega}(C_\varepsilon)} T_1(\mathbf{h}) \geq 1\right) > 1 - \varepsilon$, This, coupled with the convexity of $T_1(\mathbf{h})$, completes the proof of Proposition 3. □

B. Preliminary proposition and lemmas

Proposition 4 *If Condition (C4) is satisfied,*

$$
\sup_{\|\mathbf{h}_1\|=1, \|(\boldsymbol{\phi}^{\mathrm{T}}, \boldsymbol{\theta}^{\mathrm{T}}) - (\boldsymbol{\phi}_*^{\mathrm{T}}, \boldsymbol{\theta}_*^{\mathrm{T}})\| \leq \delta_2} \mathbf{h}_1^{\mathrm{T}} E\left[\left(\frac{\boldsymbol{\phi}(B)}{\boldsymbol{\theta}(B)}\mathbf{W}_t\right)\left(\frac{\boldsymbol{\phi}(B)}{\boldsymbol{\theta}(B)}\mathbf{W}_t^{\mathrm{T}}\right)\right]\mathbf{h}_1 \leq \lambda_{\max}J^{-1}C_2^2,
$$

where δ_2 *is chosen as in Proposition 1.*

Proof of Proposition 4: Given any $\mathbf{h}_1$, such that $\|\mathbf{h}_1\| = 1$,

$$\mathbf{h}_1^{\mathrm{T}} E\left[\left(\frac{\boldsymbol{\phi}(B)}{\boldsymbol{\theta}(B)}\mathbf{W}_t\right)\left(\frac{\boldsymbol{\phi}(B)}{\boldsymbol{\theta}(B)}\mathbf{W}_t^{\mathrm{T}}\right)\right]\mathbf{h}_1 = E\left[\left(\sum_{i=0}^{\infty}\pi_i(\boldsymbol{\phi},\boldsymbol{\theta})\mathbf{h}_1^{\mathrm{T}}\mathbf{W}_{t-i}\right)\left(\sum_{j=0}^{\infty}\pi_j(\boldsymbol{\phi},\boldsymbol{\theta})\mathbf{h}_1^{\mathrm{T}}\mathbf{W}_{t-j}\right)\right]$$

$$= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\pi_i(\boldsymbol{\phi},\boldsymbol{\theta})\pi_j(\boldsymbol{\phi},\boldsymbol{\theta})E\left[\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{W}_{t-i}\right)\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{W}_{t-j}\right)\right]$$

$$\leq \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\pi_i(\boldsymbol{\phi},\boldsymbol{\theta})\pi_j(\boldsymbol{\phi},\boldsymbol{\theta})\frac{1}{2}\left(\mathbf{h}_1^{\mathrm{T}}E\left[\mathbf{W}_{t-i}\mathbf{W}_{t-i}^{\mathrm{T}}\right]\mathbf{h}_1 + \mathbf{h}_1^{\mathrm{T}}E\left[\mathbf{W}_{t-j}\mathbf{W}_{t-j}^{\mathrm{T}}\right]\mathbf{h}_1\right)$$

$$\leq \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\pi_i(\boldsymbol{\phi},\boldsymbol{\theta})\pi_j(\boldsymbol{\phi},\boldsymbol{\theta})\lambda_{\max}J^{-1} \leq \lambda_{\max}J^{-1}\left(C_1\sum_{i=0}^{\infty}r^i\right)^2 = \lambda_{\max}J^{-1}C_2^2,$$

where the last inequality follows from the argument for part (i) in Proposition 1. This completes the proof of Proposition 4. □

Lemma 1 *Suppose Condition (C3) holds. Then* (i) $P(|\zeta_t| > v) \leq 2\exp\left(\frac{-v^2}{2(C_B^2 + C_B v)}\right)$ *and* (ii) *for any a sequence* $\{a_t, t \geq 0\}$ *and* $k \geq 1$, $E[|\sum_{i=0}^{\infty} a_i\zeta_{t-i}|^k] \leq (\sum_{i=0}^{\infty}|a_i|)^k k! C_B^k/2$.

Proof of Lemma 1: By Bernstein's condition, for any $0 < u < C_B^{-1}$,

$$E\left[\exp(u|\zeta_t|)\right] = 1 + uE\left[|\zeta_t|\right] + \frac{u^2\sigma^2}{2} + \sum_{k=3}^{\infty}u^k\frac{E\left[|\zeta_t|^k\right]}{k!} \leq 1 + \frac{1}{2} + \frac{u^2C_B^2}{2} + \frac{1}{2}\sum_{k=3}^{\infty}(|u|C_B)^k$$

$$< 2\left(1 + \frac{u^2C_B^2}{2 - 2C_B|u|}\right) \leq 2\exp\left(\frac{u^2C_B^2}{2 - 2C_B|u|}\right),$$

where the last inequality follows from $1 + u \leq \exp(u)$. By Markov's inequality,

$$P(|\zeta_t| > v) = P(\exp(u|\zeta_t|) > \exp(uv)) \leq \frac{E\left[\exp(u|\zeta_t|)\right]}{\exp(uv)} < 2\exp\left(\frac{-v^2}{2(C_B^2 + C_B v)}\right),$$

where the equality follows by choosing $u = v/(C_B v + C_B^2)$.

We prove (ii) by induction. Since $\{\zeta_t\}$ is a i.i.d. sequence,

$$E\left[|a_0\zeta_t + a_1\zeta_{t-1}|^k\right] \leq \sum_{j=0}^{k}\binom{k}{j}E\left[|a_0|^j|\zeta_t|^j\right]E\left[|a_1|^{k-j}|\zeta_{t-1}|^{k-j}\right]$$

$$\leq \sum_{j=0}^{k}\binom{k}{j}|a_0|^j|a_1|^{k-j}\frac{j!C_B^j}{2}\frac{(k-j)!C_B^{k-j}}{2} \leq (|a_0| + |a_1|)^k\frac{k!C_B^k}{2}.$$

Suppose (ii) holds for $l-1$. Then,

$$E\left[\left|\sum_{i=0}^{l-1} a_i\zeta_{t-i} + a_l\zeta_{t-l}\right|^k\right] \le \sum_{j=0}^{k}\binom{k}{j}\left(\sum_{i=0}^{l-1}|a_i|\right)^j \frac{j!C_B^j}{2}|a_l|^{k-j}\frac{(k-j)!C_B^{k-j}}{2}$$
$$\le \left(\sum_{i=0}^{l}|a_i|\right)^k \frac{k!C_B^k}{2}.$$

Thus, (ii) holds for l.

This completes the proof of Lemma 1. □

Lemma 2 *Suppose Conditions (C1) – (C4) hold. There exists some constant $C_4 > 0$ that does not depend on n, such that if $J = O(n^{1/(2\alpha+1)})$,*
(i)

$$P\left(\sup_{\|\mathbf{h}_1\|\le 1}\left|\mathbb{G}_n\left[\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\zeta_t^2\right]\right| > 7C_2\sqrt{C_4 J\log n}\right) \le 2\exp(-6J\log n).$$

(ii)

$$\sup_{\|\mathbf{h}_1\|\le 1, \mathbf{h}_1\neq 0}\left|\left(\sigma^2 E\left[\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\right]\right)^{-1}\mathbb{E}_n\left[\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\zeta_t^2\right]\right| = 1 + o_p(1).$$

(iii)

$$P\left(\sup_{\|\mathbf{h}_1\|\le 1, \|\mathbf{h}_2\|\le 1} n^{-1/2}\left|\mathbb{G}_n\left[\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\mathbf{Q}_{t2}^{\mathrm{T}}\mathbf{h}_2\right]\right| > 7p^{1/2}\sqrt{C_4 J n^{-1}\log n}\right) \le 2p\exp(-6J\log n).$$
$$P\left(\sup_{\|\mathbf{h}_1\|\le 1, \|\mathbf{h}_3\|\le 1} n^{-1/2}\left|\mathbb{G}_n\left[\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\mathbf{Q}_{t3}^{\mathrm{T}}\mathbf{h}_3\right]\right| > 7q^{1/2}\sqrt{C_4 J n^{-1}\log n}\right) \le 2q\exp(-6J\log n).$$

Proof of Lemma 2: By Theorem 5.4.2 in [5] again, we have

$$\sup_{\|\mathbf{h}_1\|\le 1}\left|\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right| = \sup_{\|\mathbf{h}_1\|\le 1}\left|\sum_{i=0}^{\infty}\pi_{*i}\mathbf{h}_1^{\mathrm{T}}\mathbf{W}_{t-i}\right| \le \sum_{i=0}^{\infty}|\pi_{*i}| \le C_2.$$

Part (i): Let $\mathcal{C} := \{\mathcal{C}(\mathbf{h}_{1,l})\}$ be a collection of cubes that cover the ball $\|\mathbf{h}_1\| \le 1$, where $\mathcal{C}(\mathbf{h}_{1,l})$ is a cube containing $\mathbf{h}_{1,l}$ with sides of length n^{-2}. Then $|\mathcal{C}| \le (2n^2)^J$. For any $\mathbf{h}_1 \in \mathcal{C}(\mathbf{h}_{1,l})$, $\|\mathbf{h}_1 - \mathbf{h}_{1,l}\|_\infty \le n^{-2}$, and we have

$$\mathbb{E}_n\left[\mathbf{h}_{1,l}^{\mathrm{T}}\mathbf{Q}_{t1}\mathbf{Q}_{t1}^{\mathrm{T}}\mathbf{h}_{1,l}\right] = \mathbb{E}_n\left[\left(\mathbf{h}_{1,l}^{\mathrm{T}}\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}\mathbf{W}_t\right)^2\right] = 2C_2\mathbb{E}_n\left[\frac{1}{2C_2}\left(\sum_{i=0}^{\infty}\pi_{*i}h_l(X_{t-i})\right)^2\right],$$

where $h_l(\cdot) = \mathbf{h}_{1,l}^{\mathrm{T}}\mathbf{B}(\cdot)$.

By Condition (C4) and the fact that $h_l(\cdot)$ is a smooth function, $\{h_l(X_t)\}$ is also β-mixing with coefficients $\beta(k; h_l(X_t)) \leq 2\exp(-d_1 k^{\gamma_1})$ (see, e.g., page 443 in [6]). Define $\psi(x) := x^2/(2C_2)$ and let $\{Z_t(h_l) := \psi\left(\sum_{i=0}^{\infty} \pi_{*i} h_l(X_{t-i})\right)\}$. According to Page 446 in [6], the following bound holds for the τ-mixing coefficients associated to the sequence $\{\sum_{i=0}^{\infty} \pi_{*i} h_l(X_{t-i})\}$:

$$\tau\left(k; \sum_{i=0}^{\infty} \pi_{*i} h_l(X_{t-i})\right) \leq 2C_2 \sum_{j\geq k} |\pi_{*j}| + 4C_2 \sum_{j=0}^{k-1} |\pi_{*j}| \beta_{h_l(X)}^{1/2}(k-j)$$

$$\leq 2C_2 \sum_{j\geq k} C_1 r^j + 8C_2 \sum_{j=0}^{k-1} C_1 r^j \exp(-d_1(k-j)^{\gamma_1}/2) \leq c_2 \exp(-d_2 k^{\gamma_2}), \tag{1}$$

for some constants $c_2, d_2, \gamma_2 > 0$ that do not depend on $\mathbf{h}_{1,l}$ and any $k \geq 1$. Here, we refer the definition of τ-mixing coefficients to Equation (2.3) in [6]. It is easy to check that $\psi(x)$ is 1-Lipschitz function. According to Page 446 in [6] again, the same bound hold for the τ-mixing coefficients associated to the sequence $\{Z_t(h_l)\}$, that is, $\tau(k; Z_t(h_l)) \leq c_2 \exp(-d_2 k^{\gamma_2})$, for any $k \geq 1$.

Since $\{X_t\}$ and $\{\zeta_t\}$ are independent, $\{Z_t(h_l)\zeta_t^2\}$ is also a τ-mixing sequence with $\tau(k; Z_t(h_l)\zeta_t^2) \leq \exp(-d_2 k^{\gamma_2})$ for any $k \geq 1$. By the fact that $h_l(X_t)$ is bounded and Lemma 1, $P\left(|Z_t(h_l)\zeta_t^2| > v\right) \leq \exp(1 - (v/d_3)^{1/2})$ for some constant $d_3 > 0$. Applying Theorem 1 in [6] yields

$$P\left(\left|n^{1/2}\mathbb{G}_n\left[Z_t(h_l)\zeta_t^2\right]\right| > x\right) \leq n\exp\left(-\frac{x^{\gamma_3}}{C_4}\right) + \exp\left(-\frac{x^2}{C_4 n}\right) + \exp\left(-\frac{x^2}{C_4 n}\exp\left(\frac{x^{(1-\gamma_3)\gamma_3}}{C_4(\log x)^{\gamma_3}}\right)\right),$$

for $1/\gamma_3 = 2 + 1/\gamma_2$ and some constant C_4 that does not depend on $\mathbf{h}_{1,l}$. We choose $x = 3\sqrt{C_4 J n \log n}$ and obtain that $P\left(\left|n^{1/2}\mathbb{G}_n\left[Z_t(h_l)\zeta_t^2\right]\right| > 3\sqrt{C_4 J n \log n}\right) \leq 3\exp(-9J\log n)$, when n is sufficiently large. Therefore,

$$\begin{aligned}
&P\left(\max_l n^{-1/2}\left|\mathbb{G}_n\left[\left(\mathbf{h}_{1,l}^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\zeta_t^2\right]\right| > 6C_2\sqrt{C_4 J n^{-1}\log n}\right) \\
&= P\left(\max_l n^{-1/2}\left|\mathbb{G}_n\left[\left(\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}\mathbf{h}_{1,l}^{\mathrm{T}}\mathbf{W}_t\right)^2\zeta_t^2\right]\right| > 6C_2\sqrt{C_4 J n^{-1}\log n}\right) \\
&= P\left(2C_2\max_l \left|n^{1/2}\mathbb{G}_n\left[Z_t(h_l)\zeta_t^2\right]\right| > 6C_2\sqrt{C_4 J n\log n}\right) \leq 3(2n^2)^J\exp(-9J\log n) \\
&\leq \exp(-6J\log n).
\end{aligned} \tag{2}$$

Since, for any $\mathbf{h}_1 \in \mathcal{C}(\mathbf{h}_{1,l})$, $\left|(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1})^2 - (\mathbf{h}_{1,l}^{\mathrm{T}}\mathbf{Q}_{t1})^2\right| \leq 2C_2\sum_{i=0}^{\infty}|\pi_{*i}|\left|(\mathbf{h}_1 - \mathbf{h}_{1,l})^{\mathrm{T}}\mathbf{W}_{t-i}\right| \leq 2C_2\sum_{i=1}^{\infty}|\pi_{*i}|\left\|\mathbf{h}_1 - \mathbf{h}_{1,l}\right\|_\infty \leq 2C_2^2 n^{-2}$, we immediately have

$$\sup_{\mathbf{h}_1\in\mathcal{C}(\mathbf{h}_{1,l})} E\left[\left|(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1})^2 - (\mathbf{h}_{1,\ell}^{\mathrm{T}}\mathbf{Q}_{t1})^2\right]\right| \leq 2C_2^2 n^{-2}. \tag{3}$$

By Lemma 1 and the boundedness of $\mathbf{h}_1\mathbf{Q}_{t1}$, when n is sufficiently large

$$P\left(\sup_{\mathbf{h}_1\in\mathcal{C}(\mathbf{h}_{1,l})}\left|\mathbb{E}_n\left[\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\zeta_t^2-\left(\mathbf{h}_{1,l}^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\zeta_t^2\right]\right|>C_2\sqrt{C_4Jn\log n}/2\right)$$
$$\leq P\left(\max_{1\leq t\leq n}\zeta_t^2\geq C_2^{-1}\sqrt{C_4Jn\log n}n^2/4\right)\leq\exp(-6J\log n). \tag{4}$$

Noting that

$$\sup_{\mathbf{h}_1\in\mathcal{C}(\mathbf{h}_{1,l})}n^{-1/2}\left|\mathbb{G}_n\left[(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1})^2\zeta_t^2\right]\right|\leq\sup_{\mathbf{h}_1\in\mathcal{C}(\mathbf{h}_{1,l})}\left|\mathbb{E}_n\left[(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1})^2\zeta_t^2-(\mathbf{h}_{1,l}^{\mathrm{T}}\mathbf{Q}_{t1})^2\zeta_t^2\right]\right|$$
$$+n^{-1/2}\left|\mathbb{G}_n\left[(\mathbf{h}_{1,l}^{\mathrm{T}}\mathbf{Q}_{t1})^2\zeta_t^2\right]\right|+\sup_{\mathbf{h}_1\in\mathcal{C}(\mathbf{h}_{1,l})}E\left[\left|(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1})^2\zeta_t^2-(\mathbf{h}_{1,l}^{\mathrm{T}}\mathbf{Q}_{t1})^2\zeta_t^2\right|\right],$$

combining (2)–(4) yields

$$P\left(\sup_{\|\mathbf{h}_1\|\leq1}n^{-1/2}\left|\mathbb{G}_n\left[(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1})^2\zeta_t^2\right]\right|>7C_2\sqrt{C_4Jn^{-1}\log n}\right)\leq2\exp(-6J\log n). \tag{5}$$

Part (ii): We use the following simple fact

$$\sup_n|A_n|=\sup_n|A_nB_nB_n^{-1}|\leq\sup_n|A_nB_n|\sup_n|B_n^{-1}|=\sup_n|A_nB_n|\inf_n|B_n|^{-1}$$
$$\Rightarrow\sup_n|A_nB_n|\geq\sup_n|A_n|\inf_n|B_n|$$

By Part (i) and Condition (C4), it is easily seen that

$$2\exp(-6J\log n)$$
$$\geq P\left(\sup_{\|\mathbf{h}_1\|=1}\left|\mathbb{E}_n\left[(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1})^2\zeta_t^2\right]-E\left[\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\zeta_t^2\right]\right|>7C_2\sqrt{C_4Jn^{-1}\log n}\right)$$
$$\geq P\Bigg(\sup_{\|\mathbf{h}_1\|=1}\left|\mathbb{E}_n\left[\left(\sigma^2E\left[\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\right]\right)^{-1}\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\zeta_t^2\right]-1\right|$$
$$\times\inf_{\|\mathbf{h}_1\|=1}\left(\sigma^2E\left[\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\right]\right)>7C_2\sqrt{C_4Jn^{-1}\log n}\Bigg)$$
$$\geq P\Bigg(\sup_{\|\mathbf{h}_1\|=1}\left|\mathbb{E}_n\left[\left(\sigma^2E\left[\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\right]\right)^{-1}\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\zeta_t^2\right]-1\right|(\sigma^2\lambda_{\min}J^{-1})$$
$$>7C_2\sqrt{C_4Jn^{-1}\log n}\Bigg)$$
$$\geq P\Bigg(\sup_{\|\mathbf{h}_1\|=1}\left|\mathbb{E}_n\left[\left(\sigma^2E\left[\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\right]\right)^{-1}\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\zeta_t^2\right]-1\right|>\frac{7C_2\sqrt{C_4Jn^{-1}\log n}}{\sigma^2\lambda_{\min}J^{-1}}\Bigg).$$

Since $J^3 n^{-1}\log n \to 0$, we have

$$\begin{aligned}
&\sup_{\|\mathbf{h}_1\|\le 1, \mathbf{h}_1\neq 0}\left|\left(\sigma^2 E\left[(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1})^2\right]\right)^{-1}\mathbb{E}_n\left[\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\zeta_t^2\right]\right| \\
=&\sup_{\|\mathbf{h}_1\|\le 1, \mathbf{h}_1\neq 0}\left|\left(\sigma^2 E\left[((\mathbf{h}_1/\|\mathbf{h}_1\|)^{\mathrm{T}}\mathbf{Q}_{t1})^2\right]\right)^{-1}\mathbb{E}_n\left[\left((\mathbf{h}_1/\|\mathbf{h}_1\|)^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\zeta_t^2\right]\right| \\
=&\sup_{\|\mathbf{h}_1\|=1}\left|\left(\sigma^2 E\left[(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1})^2\right]\right)^{-1}\mathbb{E}_n\left[\left(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\right)^2\zeta_t^2\right]\right| = 1 + o_p(1)
\end{aligned}$$

Part (iii): Noting that ζ_t's are independent, according to page 446 in [6], we obtain the bound of the τ-mixing coefficient associated to the sequence $\left\{\boldsymbol{\phi}_*^{-1}\zeta_t\right\}$ for any $k \ge 1$:

$$\tau(k;\boldsymbol{\phi}_*^{-1}\zeta_t) \le 2C_B \sum_{l\ge k}|\rho_{*l}| \le 2C_B C_1 \frac{r^k}{1-r} = 2C_B C_1 (1-r)^{-1}\exp(k\log r).$$

Since $\{X_t\}$ and $\{\zeta_t\}$ are independent, then $\mathbf{h}_{1,l}^{\mathrm{T}}\mathbf{Q}_{t1}$ and $\boldsymbol{\phi}_*^{-1}\zeta_{t-j}$ are also independent for any $j = 1,\cdots,p$. By the definition of τ-mixing coefficients, the the τ-mixing coefficient associated to the sequence $\left\{\mathbf{h}_{1,l}^{\mathrm{T}}\mathbf{Q}_{t1}\boldsymbol{\phi}_*^{-1}\zeta_{t-j}\right\}$ is

$$\tau\left(k;\mathbf{h}_{1,l}^{\mathrm{T}}\mathbf{Q}_{t1}\boldsymbol{\phi}_*^{-1}\zeta_{t-j}\right) \le 2C_B C_1(1-r)^{-1}\exp(k\log r) + \exp(-d_2 k^{\gamma_2}) \le c_4\exp(-d_4 k^{\gamma_4}),$$

for some constants $c_4, d_4, \gamma_4 > 0$. Following from the same arguments used for Part (i), we can show that for any $1 \le j \le p$,

$$P\left(\sup_{\|\mathbf{h}_1\|\le 1} n^{-1/2}\left|\mathbb{G}_n\left[\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\boldsymbol{\phi}_*^{-1}\zeta_{t-j}\right]\right| > 7\sqrt{C_4 J n^{-1}\log n}\right) \le 2\exp(-6J\log n).$$

Therefore,

$$\begin{aligned}
&P\left(\sup_{\|\mathbf{h}_1\|\le 1, \|\mathbf{h}_2\|\le 1} n^{-1/2}\left|\mathbb{G}_n\left[\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\mathbf{Q}_{t2}^{\mathrm{T}}\mathbf{h}_2\right]\right| > 7p^{1/2}\sqrt{C_4 J n^{-1}\log n}\right) \\
\le\ & P\left(\sup_{\|\mathbf{h}_2\|\le 1}\|\mathbf{h}_2\|_1 \max_{1\le j\le p}\left\{\sup_{\|\mathbf{h}_1\|\le 1} n^{-1/2}\left|\mathbb{G}_n\left[\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\boldsymbol{\phi}_*^{-1}\zeta_{t-j}\right]\right|\right\} > 7p^{1/2}\sqrt{C_4 J n^{-1}\log n}\right) \\
\le\ & \sum_{j=1}^{p} P\left(\sup_{\|\mathbf{h}_1\|\le 1} n^{-1/2}\left|\mathbb{G}_n\left[\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\boldsymbol{\phi}_*^{-1}\zeta_{t-j}\right]\right| > 7\sqrt{C_4 J n^{-1}\log n}\right) \le 2p\exp(-6J\log n).
\end{aligned}$$

Similarly, we can establish the probability bound for $n^{-1/2}\left|\mathbb{G}_n\left[\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1}\mathbf{Q}_{t3}^{\mathrm{T}}\mathbf{h}_3\right]\right|$. This completes the proof of Lemma 2. □

Lemma 3 *Suppose Conditions (C1) – (C4) hold. Then,*
(i) $\sup_{\|(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}})\|\le 1}\left|\mathbb{E}_n\left[(\mathbf{h}_2^{\mathrm{T}}\mathbf{Q}_{t2} + \mathbf{h}_3^{\mathrm{T}}\mathbf{Q}_{t3})^2\zeta_t^2\right] - \sigma^2\left(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}}\right)\boldsymbol{\Sigma}\left(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}}\right)^{\mathrm{T}}\right| \to_{a.s.} 0,$
(ii) $\mathbb{G}_n\left[(\mathbf{h}_2^{\mathrm{T}}\mathbf{Q}_{t2} + \mathbf{h}_3^{\mathrm{T}}\mathbf{Q}_{t3})\zeta_t\right] \to_d \left(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}}\right) N(0,\sigma^2\boldsymbol{\Sigma})$, *given any* $\left(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}}\right)$ *such that* $\|(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}})\| \le$

C, *for any* $C > 0$.
(iii) $\mathbb{G}_n\left[(\mathbf{h}_2^{\mathrm{T}}\mathbf{Q}_{t2}+\mathbf{h}_3^{\mathrm{T}}\mathbf{Q}_{t3})\zeta_t\right] \to_d \left(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}}\right)N(0,\sigma^2\boldsymbol{\Sigma})$ *on* $\|(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}})\|\le C$, *for any* $C>0$.

Proof of Lemma 3: Since $\{\zeta_t\}$ is an i.i.d. sequence, then $\mathbf{h}_2^{\mathrm{T}}\mathbf{Q}_{t2}$ and ζ_t are independent. Thus,

$$\sup_{\|\mathbf{h}_2\|\le 1} E\left[(\mathbf{h}_2^{\mathrm{T}}\mathbf{Q}_{t2}\zeta_t)^2\right] \le E\left[\|\mathbf{h}_2\|^2\|\mathbf{Q}_{t2}\|^2\right]\sigma^2 \le pE\left[\left(\frac{1}{\boldsymbol{\phi}_*(B)}\zeta_t\right)^2\right]\sigma^2 \le pC_2^2C_B^2\sigma^2.$$

where the first inequality follows from Cauchy-Schwarz inequality and the third inequality follows from Lemma 1. Similarly, $\sup_{\|\mathbf{h}_3\|\le 1} E\left[(\mathbf{h}_3^{\mathrm{T}}\mathbf{Q}_{t3}\zeta_t)^2\right] \le qC_2^2C_B^2\sigma^2$. By Condition (3) and uniform ergodic theorem (e.g., Theorem 2.7 in [8]), we have

$$\sup_{\|(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}})\|\le 1}\left|\mathbb{E}_n\left[(\mathbf{h}_2^{\mathrm{T}}\mathbf{Q}_{t2}+\mathbf{h}_3^{\mathrm{T}}\mathbf{Q}_{t3})^2\zeta_t^2\right]-\sigma^2\left(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}}\right)\boldsymbol{\Sigma}\left(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}}\right)^{\mathrm{T}}\right| \to_{a.s.} 0.$$

Since $(\mathbf{Q}_{t2},\mathbf{Q}_{t3})$ belongs to $\mathcal{F}_{-\infty}^{t-1}$, the σ-field generated by $\{\zeta_k, k\le t-1\}$, $\left\{\zeta_t(\mathbf{h}_2^{\mathrm{T}}\mathbf{Q}_{t2}+\mathbf{h}_3^{\mathrm{T}}\mathbf{Q}_{t3})\right\}$ is a martingale difference sequence that adapts to the filtration $\mathcal{F}_{-\infty}^{t}$. By the martingale central limit theorem,

$$\mathbb{G}_n\left[(\mathbf{h}_2^{\mathrm{T}}\mathbf{Q}_{t2}+\mathbf{h}_3^{\mathrm{T}}\mathbf{Q}_{t3})\zeta_t\right]\to_d N\left(0,\sigma^2\left(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}}\right)\boldsymbol{\Sigma}\left(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}}\right)^{\mathrm{T}}\right).$$

According to the fact that the pointwise convergence of convex functions implies uniform convergence on compact sets (e.g. Theorem 10.8 in [7]) and Theorem 7.1 in [1], we immediately obtain (iii). This completes the proof of Lemma 3. □

Lemmas 4– 6 follow from the steps in [4] and [2].

According to Proposition 1, $|\zeta_t|\le\eta_t$, $\|\mathbf{Q}_t\|_\infty\le\|\mathbf{Q}_t-\mathbf{D}_t(\boldsymbol{\xi}_*)\|_\infty+\|\mathbf{D}_t(\boldsymbol{\xi}_*)\|_\infty\le r^t\eta_0+C_2\Delta+\omega_t =: \chi_t$, and similarly $\|\mathbf{V}_t\|_{\max}\le\chi_t$. Thus,

$$\begin{aligned}
\left|\mathbf{h}^{\mathrm{T}}\mathbf{Q}_t\right| &\le \|\mathbf{h}_1\mathbf{Q}_{t1}\|+\|\mathbf{h}_2^{\mathrm{T}}\mathbf{Q}_{t2}+\mathbf{h}_3^{\mathrm{T}}\mathbf{Q}_{t3}\|\le C_2\|\mathbf{h}_1\|+\chi_t(\sqrt{p}\|\mathbf{h}_2\|+\sqrt{q}\|\mathbf{h}_3\|),\\
\left|\mathbf{h}^{\mathrm{T}}\mathbf{V}_t\mathbf{h}\right| &= \left|2\mathbf{h}_2^{\mathrm{T}}\mathbf{V}_{t,21}\mathbf{h}_1+2\mathbf{h}_3^{\mathrm{T}}\mathbf{V}_{t,31}\mathbf{h}_1+2\mathbf{h}_3^{\mathrm{T}}\mathbf{V}_{t,32}\mathbf{h}_2+\mathbf{h}_3^{\mathrm{T}}\mathbf{V}_{t,33}\mathbf{h}_3\right|\\
&\le 2C_2(\sqrt{p}\|\mathbf{h}_2\|+\sqrt{q}\|\mathbf{h}_3\|)\|\mathbf{h}_1\|+2\sqrt{pq}\chi_t\|\mathbf{h}_2\|\|\mathbf{h}_3\|+q\chi_t\|\mathbf{h}_3\|^2.
\end{aligned}$$

Lemma 4 *Suppose Conditions (C1) – (C4) hold. If* $J^2\log n = o(n^{1/2})$, *then for any* $C>0$, $\sup_{\mathbf{h}\in\Omega(C)}|T_1(\mathbf{h})-T_2(\mathbf{h})|\to_p 0$.

Proof of Lemma 4:: Simple algebra yields that

$$\begin{aligned}
&T_1(\mathbf{h})-T_2(\mathbf{h})\\
&=2\sum_{t=1}^n\zeta_t\mathbf{h}_2^{\mathrm{T}}\mathbf{V}_{t,21}\mathbf{h}_1+2\sum_{t=1}^n\zeta_t\mathbf{h}_3^{\mathrm{T}}\mathbf{V}_{t,31}\mathbf{h}_1+2\sum_{t=1}^n\zeta_t\mathbf{h}_3^{\mathrm{T}}\mathbf{V}_{t,32}\mathbf{h}_2+\sum_{t=1}^n\zeta_t\mathbf{h}_3^{\mathrm{T}}\mathbf{V}_{t,33}\mathbf{h}_3\\
&\quad+\sum_{t=1}^n\left(\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}R_t\right)\mathbf{h}^{\mathrm{T}}\mathbf{V}_t\mathbf{h}-\sum_{t=1}^n\mathbf{h}^{\mathrm{T}}\mathbf{Q}_t\mathbf{h}^{\mathrm{T}}\mathbf{V}_t\mathbf{h}-\frac{1}{4}\sum_{t=1}^n\left(\mathbf{h}^{\mathrm{T}}\mathbf{V}_t\mathbf{h}\right)^2\\
&=: I_1+I_2+I_3+I_4+I_5+I_6+I_7.
\end{aligned}$$

We first consider the item I_1. By the same arguments used to prove (2), we show that

$$\sup_{\mathbf{h}\in\Omega(C),\mathbf{h}\neq 0} n^{1/2}\left|\mathbb{G}_n\left[\zeta_t(\|\mathbf{h}_1\|\|\mathbf{h}_2\|)^{-1}\mathbf{h}_2^{\mathrm{T}}\mathbf{V}_{t,21}\mathbf{h}_1\right]\right| = O_p(\sqrt{Jn\log n}).$$

Since $\{\zeta_t\}$ and $\{X_t\}$ are independent, $E\left[\zeta_t(\|\mathbf{h}_1\|\|\mathbf{h}_2\|)^{-1}\mathbf{h}_2^{\mathrm{T}}\mathbf{V}_{t,21}\mathbf{h}_1\right] = 0$. Thus,

$$\sup_{\mathbf{h}\in\Omega(C)} |I_1| = O_p(\sqrt{Jn\log n})C^2J^{3/2}n^{-1} = o_p(1),$$

as $J^2\log n = o(n^{1/2})$. Likewise, we can prove $\sup_{\mathbf{h}\in\Omega(C)}|I_2| = o_p(1)$.

We then deal with the item I_3. Since $(V_{t,32}, V_{t,33})$ belongs to $\mathcal{F}_{-\infty}^{t-1}$, the σ-field generated by $\{\zeta_k, k\le t-1\}$, $\left\{\zeta_t\left(\|\mathbf{h}_2\|\|\mathbf{h}_3\|\right)^{-1}\mathbf{h}_3^{\mathrm{T}}\mathbf{V}_{t,32}\mathbf{h}_2\right\}$ is a martingale difference sequence that adapts to the filtration $\mathcal{F}_{-\infty}^{t}$. Following the proof used for the item I in Proposition 3 and the fact that p and q are finite, we can show that $\sup_{\|(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}})\|\le CJ^{1/2}n^{-1/2}}|I_3| = O_p(1)n^{1/2}C^2Jn^{-1} = o_p(1)$. Similarly, $\sup_{\|(\mathbf{h}_2^{\mathrm{T}},\mathbf{h}_3^{\mathrm{T}})\|\le CJn^{-1/2}}|I_4| = o_p(1)$.

For the item I_5,

$$\begin{aligned}
&\sup_{\mathbf{h}\in\Omega(C)}|I_5| \le \max_{1\le t\le n}\left|\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}R_t\right|\sup_{\mathbf{h}\in\Omega(C)}\sum_{t=1}^{n}\left|\mathbf{h}^{\mathrm{T}}\mathbf{V}_t\mathbf{h}\right| \\
&\le C_0C_2J^{-\alpha}\sum_{t=1}^{n}\Big(2C_2(\sqrt{p}+\sqrt{q})C^2J^{3/2}n^{-1} + C^2Jn^{-1}(2\sqrt{pq}+q)\chi_t\Big) \\
&= 2C_0C_2^2C^2(\sqrt{p}+\sqrt{q})J^{3/2-\alpha} + C_0C_2C^2(2\sqrt{pq}+q)\mathbb{E}_n\left[\chi_t\right]J^{1-\alpha} = o_p(1),
\end{aligned}$$

where the second inequality follows from (8) and the last equality follows from the ergodic theorem.

Similarly. for the item I_7,

$$\begin{aligned}
&\sup_{\mathbf{h}\in\Omega(C)}|I_7| \le \frac{1}{4}\sum_{t=1}^{n}\Big(2C_2(\sqrt{p}+\sqrt{q})C^2J^{3/2}n^{-1} + C^2Jn^{-1}(2\sqrt{pq}+q)\chi_t\Big)^2 \\
&\le \sum_{t=1}^{n}\left(4C_2^2C^4(\sqrt{p}+\sqrt{q})^2J^3n^{-2} + C^4J^2n^{-2}(2\sqrt{pq}+q)^2\chi_t^2\right) = J^{-1}o_p(1), \qquad (6)
\end{aligned}$$

where the first inequality follows from (8) and the equality follows from the ergodic theorem and the condition $J^2\log n = o(n^{1/2})$.

For the item I_6,

$$\sup_{\mathbf{h}\in\Omega(C)} |I_6| \leq \sup_{\mathbf{h}\in\Omega(C)} \left|\sum_{t=1}^{n}(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1})\mathbf{h}^{\mathrm{T}}\mathbf{V}_t\mathbf{h}\right| + \sup_{\mathbf{h}\in\Omega(C)} \left|\sum_{t=1}^{n}(\mathbf{h}_2^{\mathrm{T}}\mathbf{Q}_{t2}+\mathbf{h}_3^{\mathrm{T}}\mathbf{Q}_{t3})\mathbf{h}^{\mathrm{T}}\mathbf{V}_t\mathbf{h}\right|$$

$$\leq \sup_{\mathbf{h}\in\Omega(C)} \left(\sum_{t=1}^{n}(\mathbf{h}_1^{\mathrm{T}}\mathbf{Q}_{t1})^2\right)^{1/2} \sup_{\mathbf{h}\in\Omega(C)} \left(\sum_{t=1}^{n}\left(\mathbf{h}^{\mathrm{T}}\mathbf{V}_t\mathbf{h}\right)^2\right)^{1/2}$$

$$+ \sum_{t=1}^{n}\left(\chi_t(\sqrt{p}+\sqrt{q})CJ^{1/2}n^{-1/2}\right)\left(2C_2(\sqrt{p}+\sqrt{q})C^2J^{3/2}n^{-1} + C^2Jn^{-1}(2\sqrt{pq}+q)\chi_t\right)$$

$$\leq \left(nC^2J^2n^{-1}(C_2^2\lambda_{\max}J^{-1})(1+o_p(1))\right)^{1/2}J^{-1/2}o_p(1) + 2C^3C_2(\sqrt{p}+\sqrt{q})^2J^2n^{-1/2}\mathbb{E}_n[\chi_t]$$

$$+ C^3(\sqrt{p}+\sqrt{q})(2\sqrt{pq}+q)J^{3/2}n^{-1/2}\mathbb{E}_n[\chi_t^2] = o_p(1),$$

where the second inequality follows from the Cauchy-Schwartz inequality, the third inequality follows from the arguments used for the item II in Proposition 3 and (6), and the last equality follows from the ergodic theorem and the condition $J^2\log n = o(n^{1/2})$.

Combining the results above completes the proof of Lemma 4. □

Lemma 5 *Suppose Conditions (C1) – (C4) hold. If $J^{-2\alpha+1/2} = o(n^{-1/2})$, then for any $C>0$, $\sup_{\mathbf{h}\in\Omega(C)}|T_2(\mathbf{h}) - T_3(\mathbf{h})| \to_p 0$.*

Proof of Lemma 5: Simple algebra yields that

$$\begin{aligned}
&T_2(\mathbf{h}) - T_3(\mathbf{h})\\
&= -2\sum_{t=1}^{n}\zeta_t\Big(\mathbf{h}^{\mathrm{T}}\mathbf{Q}_t - \mathbf{h}^{\mathrm{T}}\mathbf{D}_t(\boldsymbol{\xi}_*)\Big) - \sum_{t=1}^{n}\zeta_t\Big(\mathbf{h}^{\mathrm{T}}\mathbf{V}_t\mathbf{h} - \mathbf{h}^{\mathrm{T}}\mathbf{H}_t(\boldsymbol{\xi}_*)\mathbf{h}\Big)\\
&\quad - 2\sum_{t=1}^{n}\left(\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}R_t\right)\left(\left(\mathbf{h}^{\mathrm{T}}\mathbf{Q}_t + \frac{1}{2}\mathbf{h}^{\mathrm{T}}\mathbf{V}_t\mathbf{h}\right) - \left(\mathbf{h}^{\mathrm{T}}\mathbf{D}_t(\boldsymbol{\xi}_*) + \frac{1}{2}\mathbf{h}^{\mathrm{T}}\mathbf{H}_t(\boldsymbol{\xi}_*)\mathbf{h}\right)\right)\\
&\quad - 2\sum_{t=1}^{n}\left(\zeta_t + \frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}R_t - \zeta_t(\boldsymbol{\xi}_*)\right)\left(\mathbf{h}^{\mathrm{T}}\mathbf{D}_t(\boldsymbol{\xi}_*) + \frac{1}{2}\mathbf{h}^{\mathrm{T}}\mathbf{H}_t(\boldsymbol{\xi}_*)\mathbf{h}\right)\\
&\quad + \left(\sum_{t=1}^{n}\left(\mathbf{h}^{\mathrm{T}}\mathbf{Q}_t\right)^2 - \sum_{t=1}^{n}\left(\mathbf{h}^{\mathrm{T}}\mathbf{D}_t(\boldsymbol{\xi}_*)\right)^2\right) + \frac{1}{4}\left(\sum_{t=1}^{n}(\mathbf{h}^{\mathrm{T}}\mathbf{V}_t\mathbf{h})^2 - \sum_{t=1}^{n}(\mathbf{h}^{\mathrm{T}}\mathbf{H}_t(\boldsymbol{\xi}_*)\mathbf{h})^2\right)\\
&\quad - \sum_{t=1}^{n}\mathbf{h}^{\mathrm{T}}(\mathbf{D}_t(\boldsymbol{\xi}_*) - \mathbf{Q}_t)\mathbf{h}^{\mathrm{T}}\mathbf{V}_t\mathbf{h} - \sum_{t=1}^{n}\mathbf{h}^{\mathrm{T}}\mathbf{Q}_t\mathbf{h}^{\mathrm{T}}(\mathbf{H}_t(\boldsymbol{\xi}_*) - \mathbf{V}_t)\mathbf{h}\\
&\quad - \sum_{t=1}^{n}\mathbf{h}^{\mathrm{T}}(\mathbf{D}_t(\boldsymbol{\xi}_*) - \mathbf{Q}_t)\mathbf{h}^{\mathrm{T}}(\mathbf{H}_t(\boldsymbol{\xi}_*) - \mathbf{V}_t)\mathbf{h}\\
&=: II_1 + II_2 + II_3 + II_4 + II_5 + II_6 + II_7 + II_8 + II_9.
\end{aligned}$$

We first consider II_1. By Condition (C3) and the definitions of $\mathbf{Q}_t$ and $\mathbf{D}_t(\boldsymbol{\xi}_*)$, ζ_t is independent of $(\mathbf{Q}_t, \mathbf{D}_t(\boldsymbol{\xi}_*))$. By Proposition 1,

$$
\begin{aligned}
& E\left[\left(\sup_{\|\mathbf{h}_1\|\le 1, \|(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}})\|\le 1} \sum_{t=1}^{n} \zeta_t \Big(\mathbf{h}^{\mathrm{T}}\mathbf{Q}_t - \mathbf{h}^{\mathrm{T}}\mathbf{D}_t(\boldsymbol{\xi}_*)\Big)\right)^2\right] \\
\le\ & E\left[\sup_{\|(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}})\|\le 1} \|(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}})\|^2 \left\|\sum_{t=1}^{n} \zeta_t \left(\Big(\mathbf{Q}_{t2} - \mathbf{D}_{t2}(\boldsymbol{\xi}_*)\Big)^{\mathrm{T}}, \Big(\mathbf{Q}_{t3} - \mathbf{D}_{t3}(\boldsymbol{\xi}_*)\Big)^{\mathrm{T}}\right)^{\mathrm{T}}\right\|^2\right] \\
\le\ & \sigma^2 E\left[\sum_{t=1}^{n}\left(\Big(\mathbf{Q}_{t2} - \mathbf{D}_{t2}(\boldsymbol{\xi}_*)\Big)^{\mathrm{T}}\Big(\mathbf{Q}_{t2} - \mathbf{D}_{t2}(\boldsymbol{\xi}_*)\Big) + \Big(\mathbf{Q}_{t3} - \mathbf{D}_{t3}(\boldsymbol{\xi}_*)\Big)^{\mathrm{T}}\Big(\mathbf{Q}_{t3} - \mathbf{D}_{t3}(\boldsymbol{\xi}_*)\Big)\right)\right] \\
\le\ & \sigma^2 E\left[\sum_{t=1}^{n}(p+q)(r^t\eta_0 + C_2\Delta)^2\right] \le 2(p+q)\sigma^2 C_2\left(E\left[\eta_0^2\right] + nC_2C_0^2J^{-2\alpha}\right).
\end{aligned}
$$

Thus, by the Markov's inequality, we can show

$$
\begin{aligned}
&\left|\sup_{\|\mathbf{h}_1\|\le 1, \|(\mathbf{h}_2^{\mathrm{T}}, \mathbf{h}_3^{\mathrm{T}})\|\le 1} \sum_{t=1}^{n} \zeta_t \Big(\mathbf{h}^{\mathrm{T}}\mathbf{Q}_t - \mathbf{h}^{\mathrm{T}}\mathbf{D}_t(\boldsymbol{\xi}_*)\Big)\right| \\
&= O_p\left(\sigma\sqrt{2(p+q)\sigma^2 C_2\left(E\left[\eta_0^2\right] + nC_2C_0^2J^{-2\alpha}\right)}\right)
\end{aligned}
$$

Consequently,

$$
\sup_{\mathbf{h}\in\Omega(C)} |II_1| = CJ^{1/2}n^{-1/2}O_p\left(\sigma\sqrt{2(p+q)\sigma^2 C_2\left(E\left[\eta_0^2\right] + nC_2C_0^2J^{-2\alpha}\right)}\right) = o_p(1).
$$

By the arguments for the item I_1 in Lemma 4, it is easy to show $\sup_{\mathbf{h}\in\Omega(C)} |II_2| = o_p(1)$. For the item II_3, by Proposition 1,

$$
\begin{aligned}
& \sup_{\mathbf{h}\in\Omega(C)} |II_3| \\
\le\ & 2\max_{1\le t\le n}\left|\frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}R_t\right| \sup_{\mathbf{h}\in\Omega(C)} \sum_{t=1}^{n}\left(\left|\mathbf{h}^{\mathrm{T}}\mathbf{Q}_t - \mathbf{h}^{\mathrm{T}}\mathbf{D}_t(\boldsymbol{\xi}_*)\right| + \left|\mathbf{h}^{\mathrm{T}}\mathbf{V}_t\mathbf{h} - \mathbf{h}^{\mathrm{T}}\mathbf{H}_t(\boldsymbol{\xi}_*)\mathbf{h}\right|\right) \\
\le\ & C_0C_2J^{-\alpha}\sum_{t=1}^{n}\left(CJ^{1/2}n^{-1/2} + C^2J^{3/2}n^{-1}(p+q)\right)\left(r^t\eta_0 + C_0C_2J^{-\alpha}\right) = o_p(1),
\end{aligned}
$$

as $J^{-2\alpha+1/2} = o(n^{-1/2})$.

Likewise, we obtain $\sup_{\mathbf{h}\in\Omega(C)}(|II_4| + |II_5| + |II_6| + |II_7| + |II_8| + |II_9|) = o_p(1)$. Combining the results above completes the proof of Lemma 5. □

Lemma 6 *Suppose Conditions (C1) – (C4) hold. If $J^2\log n = o(n^{1/2})$, then for any $C > 0$, $\sup_{\mathbf{h}\in\Omega(C)} |T_3(\mathbf{h}) - T(\mathbf{h})| \to_p 0$.*

Proof of Lemma 6: By the mean value theorem,

$$
\begin{aligned}
|T_3(\mathbf{h}) - T(\mathbf{h})| &= \left| \sum_{t=1}^{n} \left(\zeta_t(\boldsymbol{\xi}_*) - \mathbf{h}^{\mathrm{T}}\mathbf{D}_t(\boldsymbol{\xi}_*) - \frac{1}{2}\mathbf{h}^{\mathrm{T}}\mathbf{H}_t(\boldsymbol{\xi}_*)\mathbf{h} \right)^2 - \sum_{t=1}^{n} \zeta_t^2\left(\boldsymbol{\xi}_* + \mathbf{h}\right) \right| \\
&\le \frac{1}{2}\sum_{t=1}^{n} \left| \mathbf{h}^{\mathrm{T}}\Big(\mathbf{H}_t(\boldsymbol{\xi}_*) - \mathbf{H}_t(\boldsymbol{\xi}_{\#,\mathbf{h}})\Big)\mathbf{h} \right| \Bigg[\left|\zeta_t\left(\boldsymbol{\xi}_* + \mathbf{h}\right)\right| + \left|\zeta_t\left(\boldsymbol{\xi}_*\right)\right| \\
&\qquad\qquad + \left|\mathbf{h}^{\mathrm{T}}\mathbf{D}_t(\boldsymbol{\xi}_*)\right| + \frac{1}{2}\left|\mathbf{h}^{\mathrm{T}}\mathbf{H}_t(\boldsymbol{\xi}_*)\mathbf{h}\right| \Bigg]
\end{aligned}
$$

where $\boldsymbol{\xi}_{\#,\mathbf{h}}$ is a point between $\boldsymbol{\xi}_*$ and $\boldsymbol{\xi}_* + \mathbf{h}$ and may depend on t. Let $\mathbf{U}_t = \mathbf{H}_t(\boldsymbol{\xi}_*) - \mathbf{H}_t(\boldsymbol{\xi}_{\#,\mathbf{h}})$. By Proposition 1 and the arguments used in Lemma 7, we can show that $\max\left\{|\mathbf{U}_{t,12}|_{\max}, |\mathbf{U}_{t,13}|_{\max}\right\} \le C_2C_3CJn^{-1/2}$, $\max\left\{|\mathbf{U}_{t,23}|_{\max}, |\mathbf{U}_{t,33}|_{\max}\right\} \le C_3Cn^{-1/2}\eta_t + 2C_2CJn^{-1/2}$. Thus

$$
\begin{aligned}
&\sup_{\mathbf{h}\in\Omega(C)} \left| \mathbf{h}^{\mathrm{T}}\Big(\mathbf{H}_t(\boldsymbol{\xi}_*) - \mathbf{H}_t(\boldsymbol{\xi}_{\#,\mathbf{h}})\Big)\mathbf{h} \right| \\
&\le \sup_{\mathbf{h}\in\Omega(C)} \left| 2\mathbf{h}_2^{\mathrm{T}}\mathbf{U}_{t,21}\mathbf{h}_1 + 2\mathbf{h}_3^{\mathrm{T}}\mathbf{U}_{t,31}\mathbf{h}_1 + 2\mathbf{h}_3^{\mathrm{T}}\mathbf{U}_{t,32}\mathbf{h}_2 + \mathbf{h}_3^{\mathrm{T}}\mathbf{U}_{t,33}\mathbf{h}_3 \right| \\
&\le 2(\sqrt{p} + \sqrt{q})C_2C_3CJ^{1/2}n^{-1/2}\Big(C^2J^{3/2}n^{-1}\Big) \\
&\qquad\qquad + (2\sqrt{pq} + q)C^2Jn^{-1}\left(C_3Cn^{-1/2}\eta_t + 2C_2CJn^{-1/2}\right) \\
&\le C_7(J^2n^{-3/2} + \eta_tJn^{-3/2}),
\end{aligned}
$$

for some $C_7 > 0$. By Proposition 1, the condition $J^2\log n = o(n^{1/2})$, and the ergodic theorem,

$$
\begin{aligned}
\sup_{\mathbf{h}\in\Omega(C)} |T_3(\mathbf{h}) - T(\mathbf{h})| &\le \sum_{t=1}^{n} C_7(J^{3/2}n^{-3/2} + \eta_tn^{-3/2}) \\
&\quad \times \sup_{\mathbf{h}\in\Omega(C)} \left(\left|\zeta_t\left(\boldsymbol{\xi}_* + h\right)\right| + \left|\zeta_t\left(\boldsymbol{\xi}_*\right)\right| + \left|\mathbf{h}^{\mathrm{T}}\mathbf{D}_t(\boldsymbol{\xi}_*)\right| + \left|\mathbf{h}^{\mathrm{T}}\mathbf{H}_t(\boldsymbol{\xi}_*)\mathbf{h}\right| \right) = o_p(1).
\end{aligned}
$$

This completes the proof of Lemma 6. □

Lemma 7 *Under the same conditions as in Proposition 1, for any sequence* $\{a_t\}, t \ge 1$, *there exists some constant* C_3 *such that*

$$
\left| \left(\frac{\boldsymbol{\phi}(B)}{\boldsymbol{\theta}(B)} - \frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)} \right) a_t \right| \le C_3\delta_2 \sum_{i=0}^{\infty} r^i |a_{t-i}|,
$$

where δ_2 *and* r *are defined in Proposition 1*

Proof of Lemma 7: Noting that

$$\frac{\boldsymbol{\phi}(z)}{\boldsymbol{\theta}(z)} - \frac{\boldsymbol{\phi}_*(z)}{\boldsymbol{\theta}_*(z)} = \frac{\boldsymbol{\phi}(z)\left(\boldsymbol{\theta}_*(z) - \boldsymbol{\theta}(z)\right)}{\boldsymbol{\theta}(z)\boldsymbol{\theta}_*(z)} + \frac{\boldsymbol{\phi}(z) - \boldsymbol{\phi}_*(z)}{\boldsymbol{\theta}_*(z)}$$

Let $b_t = \boldsymbol{\theta}_*^{-1}(B)a_t$. We have $|b_t| \leq C_1 \sum_{i=0}^{t} r^i \left|a_{t-i}\right|$. Then,

$$\left|\frac{\boldsymbol{\phi}(B) - \boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)} a_t\right| = \left|\left(\boldsymbol{\phi}(B) - \boldsymbol{\phi}_*(B)\right) b_t\right| = \left|\sum_{j=1}^{p} (\phi - \phi_*) b_{t-j}\right| \leq \|\phi - \phi_*\| \sum_{j=1}^{p} |b_{t-j}|$$

$$\leq \delta_2 \sum_{j=1}^{p} |b_{t-j}| \leq C_1 \delta_2 \sum_{j=1}^{p} \sum_{i=0}^{\infty} r^i \left|a_{t-j-i}\right| \leq C_1 \delta_2 \frac{1 - r^p}{(1-r) r^p} \sum_{i=1}^{\infty} r^i |a_{t-i}|.$$

Similarly, there exists some constant C_1^*, such that

$$\left|\frac{\boldsymbol{\phi}(B)\left(\boldsymbol{\theta}_*(B) - \boldsymbol{\theta}(B)\right)}{\boldsymbol{\theta}(B)\boldsymbol{\theta}_*(B)} a_t\right| \leq C_1^* \delta_2 \frac{1 - r^q}{(1-r) r^q} \sum_{i=1}^{\infty} r^i |a_{t-i}|$$

Thus,

$$\left|\left(\frac{\boldsymbol{\phi}(B)}{\boldsymbol{\theta}(B)} - \frac{\boldsymbol{\phi}_*(B)}{\boldsymbol{\theta}_*(B)}\right) a_t\right| \leq \delta_2 \left(C_1 \frac{1 - r^p}{(1-r) r^p} + C_1^* \frac{1 - r^q}{(1-r) r^q}\right) \sum_{i=1}^{\infty} r^i |a_{t-i}|$$

$$\leq C_3 \delta_2 \sum_{i=0}^{\infty} r^i |a_{t-i}|,$$

where $C_3 > C_1 \frac{1-r^p}{(1-r)r^p} + C_1^* \frac{1-r^q}{(1-r)r^q}$. This completes the proof of Lemma 7. □

References

[1] Billingsley, P. (1999). *Convergence of probability measures.* John Wiley & Sons. New York.

[2] Brockwell, P. J., Davis, R. A. (1991). *Time series: theory and methods, 2nd* edition. Springer Series in Statistics. New York: Springer-Verlag.

[3] Davis, R. A. (1996). Gauss-Newton and M-estimation for ARMA processes with infinite variance. *Stochastic Processes and Their Applications*, *63*(1), 75–95.

[4] Davis, R. A., Dunsmuir, W. T. (1997). Least absolute deviation estimation for regression with ARMA errors. *Journal of Theoretical Probability*, *10*(2), 481–497.

[5] DeVore, R. A. and Lorentz, G. G. (1993). *Constructive approximation.* Springer-Verlag, Berlin Heidelberg.

[6] Merlevède, F., Peligrad, M., Rio, E. (2011). A Bernstein type inequality and moderate deviations for weakly dependent sequences. *Probability Theory and Related Fields*, *151*(3-4), 435–474.

[7] Rockafellar, R. T. (1970). *Convex Analysis*. Princeton University Press, Princeton, NJ.

[8] Straumann, D., Mikosch, T. (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *The Annals of Statistics*, *34*(5), 2449–2495.