Online Supplementary Material to "The Local Projection Residual Bootstrap for AR(1) Models"

Amilcar Velez
Department of Economics
Cornell University
amilcare@cornell.edu

August 27, 2025

C Proof of Auxiliary Results: Uniform Inference

C.1 Proof of the Lemma B.1

Proof. **Notation:** We say a sequence of random variables Z_n is uniformly $O_p(1)$ if $\forall \epsilon > 0$ there exists $M > 0$ and $n_0 \in \mathbf{N}$ such that $P_\rho(|Z_n| > M) < \epsilon$ for any $\rho \in [-1, 1]$ and $n \geq n_0$. Similarly, Z_n is uniformly $o_p(1)$ if $\forall \epsilon, \delta > 0$ there exists $n_0 \in \mathbf{N}$ such that $P_\rho(|Z_n| > \delta) < \epsilon$ for any $\rho \in [-1, 1]$ and $n \geq n_0$.

Item 1: Consider the following derivation:

$$g(\rho, n)\ n^{1/2}\ \left(\hat{\rho}_n - \rho\right) = \left(\frac{g(\rho, n)^{-2} \sum_{t=1}^{n} y_{t-1}^2}{n}\right)^{-1} \left(\frac{\sum_{t=1}^{n} u_t y_{t-1}}{g(\rho, n) n^{1/2}}\right) ,$$

where the first term is uniformly $O_p(1)$ due to Assumption 4.2. The second term is also uniformly $O_p(1)$ due to the following derivation:

$$E\left[\left(\frac{\sum_{t=1}^{n} u_t y_{t-1}}{g(\rho, n) n^{1/2}}\right)^2\right] = \frac{1}{g(\rho, n)^2 n} \sum_{t=1}^{n} E[u_t^2 y_{t-1}^2] \leq E[u_t^4]^{1/2} \max_{1 \leq i \leq n} \left(\frac{E[y_{t-1}^4]}{g(\rho, n)^4}\right)^{1/2} \leq C_8^{1/(4\zeta)} C_{y4}^{1/2} ,$$

where the first inequality follows by Cauchy's and algebra manipulation and the second inequality follows by Assumption 4.1 and part(i) of Lemma MOMT-Y in Xu (2023). The constant C_{y4} depends on the distribution of the sequence $\{u_t : t \geq 1\}$ but does not depend on ρ. Therefore, we conclude $g(\rho, n)\ n^{1/2}\ (\hat{\rho}_n - \rho)$ is uniformly $O_p(1)$ for any $\rho \in [-1, 1]$, which conclude the proof of the lemma.

Item 2: Recall that $\tilde{u}_t = \hat{u}_t - n^{-1} \sum_{t=1}^n \hat{u}_t$, where $\hat{u}_t = y_t - \hat{\rho}_n y_{t-1}$ and $\hat{\rho}_n$ is defined in (12). By Bonferroni's inequality, it is sufficient to prove that there exists $N_0 = N_0(\eta)$ such that

$$P_\rho\left(\left|n^{-1}\sum_{t=1}^n \hat{u}_t^2 - \sigma^2\right| > \sigma^2/4\right) < \eta/2 \tag{C.1}$$

and

$$P_\rho\left(\left(n^{-1}\sum_{t=1}^n \hat{u}_t\right)^2 > \sigma^2/4\right) < \eta/2 \tag{C.2}$$

for any $n \geq N_0$ and any $\rho \in [-1, 1]$. Lemma SIG in Xu (2023) adapted for the case of the AR(1) model implies (C.1). To prove (C.2), we derive the following inequality

$$\begin{aligned}\left(n^{-1}\sum_{t=1}^n \hat{u}_t\right)^2 &= \left(n^{-1}\sum_{t=1}^n u_t + (\rho - \hat{\rho}_n) n^{-1} \sum_{t=1}^n y_{t-1}\right)^2 \\ &\leq 2\left(n^{-1}\sum_{t=1}^n u_t\right)^2 + 2g(\rho, n)^2 (\hat{\rho}_n - \rho)^2 \left(n^{-1}\sum_{t=1}^n \frac{y_{t-1}^2}{g(\rho, n)^2}\right) ,\end{aligned}$$

where we used Loeve's inequality (see Theorem 9.28 in Davidson (1994)) in the inequality above. Note that the first term is uniformly $o_p(1)$ due to the law of large numbers for α-mixing sequences (see Corollary 3.48 in White (2000)) and Assumption 4.1. Since $g(\rho, n)^2 (\hat{\rho}_n - \rho)^2$ is uniformly $o_p(1)$ due to Part 1, it is sufficient to prove that $n^{-1}\sum_{t=1}^n \frac{y_{t-1}^2}{g(\rho,n)^2}$ is uniformly $O_p(1)$. The last claim follows by the next inequality

$$E\left[n^{-1}\sum_{t=1}^n \frac{y_{t-1}^2}{g(\rho, n)^2}\right] \leq n^{-1}\sum_{t=1}^n \left(E\left[\frac{y_{t-1}^4}{g(\rho, n)^4}\right]\right)^{1/2} \leq C_{y4}^{1/2} ,$$

where the last inequality follows by part(i) of Lemma MOMT-Y in Xu (2023). The constant C_{y4} depends on the distribution of the sequence $\{u_t : t \geq 1\}$ but does not depend on ρ. Therefore, $n^{-1}\sum_{t=1}^n \frac{y_{t-1}^2}{g(\rho,n)^2}$ is uniformly $O_p(1)$, which concludes the proof of the lemma.

Item 3: Recall that $\tilde{u}_t = \hat{u}_t - n^{-1} \sum_{t=1}^n \hat{u}_t$, where $\hat{u}_t = y_t - \hat{\rho}_n y_{t-1}$ and $\hat{\rho}_n$ is defined in

(12). By Loeve's inequality (see Theorem 9.28 in Davidson (1994)), we obtain

$$(\hat{u}_t - n^{-1}\sum_{t=1}^{n}\hat{u}_t)^4 = ((\hat{u}_t - u_t) + n^{-1}\sum_{t=1}^{n}\hat{u}_t + u_t)^4 \leq 3^3 \left((\hat{u}_t - u_t)^4 + \left(n^{-1}\sum_{t=1}^{n}\hat{u}_t\right)^4 + u_t^4 \right) .$$

Therefore, it is sufficient to prove that there exists $N_0 = N_0(\eta)$ and $\tilde{K}_4$ such that

$$P\left(n^{-1}\sum_{t=1}^{n}(\hat{u}_t - u_t)^4 > \tilde{K}_4/81\right) < \eta/3 \ , \tag{C.3}$$

$$P\left(\left(n^{-1}\sum_{t=1}^{n}\hat{u}_t\right)^4 > \tilde{K}_4/81\right) < \eta/3 \ , \tag{C.4}$$

$$P\left(n^{-1}\sum_{t=1}^{n}u_t^4 > \tilde{K}_4/81\right) < \eta/3 \ . \tag{C.5}$$

To prove (C.3), we use $\hat{u}_t - u_t = (\rho - \hat{\rho}_n)y_{t-1}$, the following equality

$$n^{-1}\sum_{t=1}^{n}(\hat{u}_t - u_t)^4 = (\hat{\rho}_n - \rho)^4 \ g(\rho, n)^4 \ n^{-1}\sum_{t=1}^{n}\frac{y_{t-1}^4}{g(\rho, n)^4} \ ,$$

Markov inequality, and part (i) of Lemma MOMT-Y in Xu (2023). To verify (C.4), we use (C.2) from the proof of Part 2. Finally, Markov's inequality and Assumption 4.1 implies (C.5). ■

C.2 Proof of the Lemma B.2

Proof. We first prove that there exists $\tilde{M} = \tilde{M}(M) > 0$ and $N_0 = N_0(M) > 0$ such that

$$\rho + \frac{M}{n^{1/2}g(\rho, n)} \leq 1 + \frac{\tilde{M}}{n} \ , \tag{C.6}$$

for all $n \geq N_0$. Note that (C.6) is sufficient to conclude that $\hat{\rho}_n \leq 1 + \tilde{M}/n$ since $\hat{\rho}_n \leq \rho + Mn^{-1/2}/g(\rho_n, n)$.

Let us prove (C.6) by contradiction. That is: suppose that there exist sequences ρ_k, $\tilde{M}_k \to \infty$, and $n_k \to \infty$ such that $\rho_k + M/(n_k^{1/2}g(\rho_k, n_k)) > 1 + \tilde{M}_k/n_k$, for all k. The previous expression is equivalent to

$$M > n_k^{1/2} g(\rho_k, n_k)(1-\rho_k) + \tilde{M}_k \frac{g(\rho_k, n_k)}{n_k^{1/2}} \ . \tag{C.7}$$

Define $a_k = n_k(1-|\rho_{n_k}|)$. Consider the derivation to get a lower bound for $g(\rho_k, n_k)$:

$$\begin{aligned} g(\rho_k, n_k)^2 &= 1 + (1-a_k/n_k)^2 + \ldots + (1-a_k/n_k)^{2(n_k-1)} \\ &= \frac{\{1-(1-a_k/n_k)^{n_k}\}\{1+(1-a_k/n_k)^{n_k}\}}{a_k/n_k\{2-a_k/n_k\}} \\ &\geq \frac{n_k}{a_k} \times \frac{1-e^{-a_k}}{2} \ , \end{aligned}$$

where the last inequality use that $(1-a_k/n_k)^{n_k} = \exp(n_k \log(1-a_k/n_k)) \leq \exp(-a_k)$. Without loss of generality, suppose that $a_k \to a \in [0, +\infty]$; otherwise, we can use a subsequence. We now consider two cases. For the first case, suppose $a_k \to +\infty$. This implies that

$$n_k^{1/2} g(\rho_k, n_k)(1-\rho_k) \geq \left(\frac{1-e^{-a_k}}{2}\right)^{1/2} a_k^{1/2} \to \infty \ ,$$

which contradicts (C.7). For the second case, suppose $a_k \to a$. This implies that

$$\frac{\tilde{M}_k}{\sqrt{n_k}} g(\rho_k, n_k) \geq \left(\frac{1-e^{-a_k}}{2a_k}\right)^{1/2} \tilde{M}_k \to \infty \ ,$$

which contradicts (C.7). Therefore, there exists $N_0 = N_0(M)$ and $\tilde{M} = \tilde{M}(M)$ such that (C.6) holds for $n \geq N_0$. We can adapt the proof to conclude that $\hat{\rho}_n > -1 - \tilde{M}/n$ for all $n \geq N_0$. ■

C.3 Proof of the Lemma B.3

Proof. We prove only item 1 since the proof of item 2 is analogous. The proof of item 1 has three steps. First, we can write $P_\rho\left(|R^*_{n,b}(h)| \leq x \mid Y^{(n)}\right) - (2\Phi(x) - 1) = I_1 + I_2 + I_3$, where $I_1 = J_n(x, h, \hat{P}_n, \hat{\rho}_n) - \Phi(x)$, $I_2 = \Phi(-x) - J_n(-x, h, \hat{P}_n, \hat{\rho}_n)$, and

$$I_3 = P_\rho(R^*_{n,b}(h) = -x \mid Y^{(n)}) \leq P_\rho(R^*_{n,b}(h) \in (-x-\epsilon/2, -x+\epsilon/2] \mid Y^{(n)}) \ .$$

Second, conditional on the event E_n defined in the proof of Theorem 4.1, the inequality (A.1) in the proof of Theorem 4.1 implies that $|I_1| < \epsilon/2$ and $|I_2| < \epsilon/2$ for any $n \geq N_2 = N_2(\epsilon, \eta)$ and any $h \leq h_n$ such that $h_n \leq n$ and $h_n = o(n)$. Also, the inequality (A.1) and algebra manipulation implies $|I_3| < 2\epsilon$. Therefore, we conclude

$$\sup_{h \leq h_n} \sup_{x \in \mathbf{R}} \left| P_\rho \left(|R^*_{n,b}(h)| \leq x \mid Y^{(n)} \right) - (2\Phi(x) - 1) \right| < 3\epsilon \ , \tag{C.8}$$

for any $n \geq N_2$ and any $h_n \leq n$ such that $h_n = o(n)$. Third, taking $x = z_{1-\alpha/2-3\epsilon/2}$ in (C.8), it follows that $\left| P_\rho \left(|R^*_{n,b}(h)| \leq z_{1-\alpha/2-3\epsilon/2} \mid Y^{(n)} \right) - (1 - \alpha - 3\epsilon) \right| < 3\epsilon$, which implies

$$P_\rho \left(|R^*_{n,b}(h)| \leq z_{1-\alpha/2-3\epsilon/2} \mid Y^{(n)} \right) \leq 1 - \alpha \ .$$

By definition of $c^*_n(h, 1-\alpha)$ as in (15), it follows that $c^*_n(h, 1-\alpha) \geq z_{1-\alpha/2-3\epsilon/2}$ holds conditional on the event E_n. We similarly obtain that $c^*_n(h, 1-\alpha) \leq z_{1-\alpha/2+3\epsilon/2}$ holds conditional on E_n. ■

C.4 Proof of Proposition B.1

Proof of Proposition B.1. We use the general subsequence approach of Andrews et al. (2020) to show that the uniform result in the proposition holds. We prove that for any sequence $\{\rho_n : n \geq 1\}$ such that $|\rho_n| \leq 1 + M/n$ and any sequence $\{\sigma^2_n : n \geq 1\} \subset [\underline{\sigma}, \overline{\sigma}]$, there exists subsequences $\{\rho_{n_k} : k \geq 1\}$ and $\{\sigma^2_{n_k} : k \geq 1\}$ such that

$$\lim_{K \to \infty} \lim_{k \to \infty} P \left(g(\rho_{n_k}, n_k)^{-2} n_k^{-1} \sum_{i=1}^{n_k} y^2_{n_k, i-1} \geq 1/K \right) = 1 \ . \tag{C.9}$$

We consider two cases to prove (C.9). The first case is $n_k(1 - |\rho_{n_k}|) \to \infty$ for some subsequence $\{n_k : k \geq 1\}$, which considers the subsequence of ρ_n that stay on the stationary region or go to the boundary at slower rates. The second case is $n_k(1 - |\rho_{n_k}|) \to c \in [-M, +\infty)$ for some subsequence $\{n_k : k \geq 1\}$, which considers the subsequence of ρ_n that goes to the boundary (local-unit-model) or are on it (unit-root model). For both cases, we assume $\sigma^2_{n_k} \to \sigma^2_0$ since any sequence $\{\sigma^2_n : n \geq 1\} \subset [\underline{\sigma}, \overline{\sigma}]$ has always a convergent subsequence. To avoid complicated sub-index notation, we present the algebra derivation using the original sequence.

Case 1: Suppose $n(1 - |\rho_n|) \to \infty$ and $\sigma^2_n \to \sigma^2_0$ as $n \to \infty$. This condition implies that there exists N_0 such that $|\rho_n| \leq 1$ for all $n \geq N_0$, otherwise there is a subsequence $|\rho_{n_k}|$ in $(1, 1 + M/n_k]$ but this cannot occur since $n_k(1 - |\rho_{n_k}|) \in [-M, 0]$. As a result, we have $g(\rho_n, n)^2 = \sum_{\ell=0}^{n-1} \rho_n^{2\ell} \leq (1 - \rho_n^2)^{-1}$ that implies

$$P\left(g(\rho_n, n)^{-2}\, n^{-1} \sum_{t=1}^{n} y_{n,t-1}^2 \geq 1/K \right) \geq P\left(\frac{(1 - \rho_n^2)}{n} \sum_{t=1}^{n} y_{n,t-1}^2 \geq 1/K \right) .$$

Therefore, to verify (C.9) is sufficient to prove that

$$\lim_{K \to \infty} \lim_{n \to \infty} P\left(\frac{(1 - \rho_n^2)}{n} \sum_{t=1}^{n} y_{n,t-1}^2 \geq 1/K \right) = 1 ,$$

which follows if we prove

$$\frac{(1 - \rho_n^2)}{n} \sum_{t=1}^{n} y_{n,t-1}^2 \xrightarrow{p} \sigma_0^2 . \tag{C.10}$$

We prove (C.10) in two steps.

Step 1: Using Assumption B.1 and $y_{n,t-1} = \sum_{\ell=1}^{i-1} \rho_n^{i-1-\ell} u_{n,\ell}$, we derive the following:

$$\begin{aligned} E\left[\frac{(1 - \rho_n^2)}{n} \sum_{t=1}^{n} y_{n,t-1}^2 \right] &= \frac{(1 - \rho_n^2)}{n} \sum_{t=1}^{n} \sum_{\ell=1}^{n-1} E[u_{n,\ell}^2]\ \rho_n^{2(t-1-\ell)}\ I\{1 \leq \ell \leq i - 1\} \\ &= \frac{\sigma_n^2 (n-1)}{n} - \frac{\sigma_n^2}{n} \sum_{\ell=1}^{n-1} \rho_n^{2(n-\ell)} . \end{aligned}$$

We conclude the right-hand side of the previous display converges to σ_0^2 since $\sigma_n^2 \to \sigma_0^2$, $n^{-1} \sum_{\ell=1}^{n} \rho_n^{2(n-\ell)} \leq \{n(1 - |\rho_n|)(1 + |\rho_n|)\}^{-1}$, and $n(1 - |\rho_n|) \to \infty$.

Step 2: We use $E[y_{n,t-1}^2] = g(\rho_n, t-1)^2 \sigma_n^2$ to derive the following decomposition

$$\sum_{i=1}^{n} y_{n,t-1}^2 - E[y_{n,t-1}^2] = \sum_{\ell=1}^{n-1} \left(u_{n,\ell}^2 - \sigma_n^2 \right) b_{n,\ell} + 2 \sum_{\ell=1}^{n-1} u_{n,\ell}\, d_{n,\ell} ,$$

where $b_{n,\ell} = \sum_{i=1+\ell}^{n} \rho_n^{2(i-1-\ell)}$ and $d_{n,\ell} = b_{n,\ell} \sum_{\ell_2=1}^{\ell-1} u_{n,\ell_2} \rho_n^{\ell-\ell_2}$. Note $d_{n,\ell}$ is measurable with respect to the σ-algebra defined by $\{u_{n,k} : 1 \leq k \leq \ell - 1\}$.

The decomposition above, Loeve's inequality (see Theorem 9.28 in Davidson (1994)), and

Assumption B.1 imply that the variance of $(1-\rho_n^2)n^{-1}\sum_{t=1}^n y_{n,t-1}^2$ is lower than

$$\frac{2(1-\rho_n^2)^2}{n^2}\left(\sum_{\ell=1}^{n-1} E\left[\left(u_{n,\ell}^2-\sigma_n^2\right)^2\right] b_{n,\ell}^2 + 4\sum_{\ell=1}^{n-1} E\left[u_{n,\ell}^2\right] E\left[d_{n,\ell}^2\right]\right) .$$

Since $b_{n,\ell}^2 \leq (1-\rho_n^2)^{-2}$ and $E[d_{n,\ell}^2] = b_{n,\ell}^2\sigma_n^2\sum_{\ell_2=1}^{\ell-1}\rho_n^{2(\ell-\ell_2)} \leq \sigma_n^2(1-\rho_n^2)^{-3}$ for $\ell = 1,\ldots,n-1$, and $E[(u_{n,\ell}^2-\sigma_n^2)^2] \leq E[u_{n,\ell}^4] \leq K_4$ by Assumption B.1, the previous display is lower than

$$\frac{2(n-1)K_4}{n^2} + \frac{8(n-1)\sigma_n^4}{n^2(1-\rho_n^2)} ,$$

which goes to 0 since $\sigma_n^4 = E[u_{n,\ell}^2]^2 \leq E[u_{n,\ell}^4] \leq K_4$ and $n(1-\rho_n^2) = n(1-|\rho_n|)(1+|\rho_n|) \to \infty$ as $n \to \infty$. This proves that the variance of the left-hand side on (C.10) goes to zero, which proves (C.10) due to step 1.

Case 2: Suppose $n(1-|\rho_n|) \to c \in [-M,+\infty)$ and $\sigma_n^2 \to \sigma_0^2$ as $n \to \infty$. We first observe that $g(\rho_n,n)^2 \leq n\exp(2M)$ due to $|\rho_n| \leq 1 + M/n$ and the following derivation:

$$g(\rho_n,n)^2 = \sum_{\ell=0}^{n-1}\rho_n^{2\ell} \leq n(1+M/n)^{2n} = n\exp\left(2n\log\left(1+M/n\right)\right) \leq n\exp\left(2M\right) ,$$

where we used that $\log(1+x) \leq x$ for all $x > -1$. By the previous observation

$$P\left(g(\rho_n,n)^{-2}n^{-1}\sum_{t=1}^n y_{n,t-1}^2 \geq 1/K \right) \geq P\left(\frac{\exp(-2M)}{n^2}\sum_{t=1}^n y_{n,t-1}^2 \geq 1/K \right) ,$$

where $\exp(-2M)$ is a constant that does not change as $n \to \infty$ and $K \to \infty$. Therefore, it is sufficient to prove that $\lim_{K\to\infty} \lim_{n\to\infty} P\left(\frac{1}{n^2}\sum_{t=1}^n y_{n,t-1}^2 \geq 1/K \right) = 1$ to verify (C.9), which follows if we prove

$$\frac{1}{n^2}\sum_{t=1}^n y_{n,t-1}^2 \xrightarrow{d} \sigma_0^2\int_0^1 J_{-c}(r)^2 dr , \tag{C.11}$$

where $J_{-c}(r) = \int_0^r e^{-(r-s)c}dW(s)$ and $W(s)$ is a standard Brownian motion.

To prove (C.11), we rely on the results and techniques presented in Phillips (1987). Specifically, we adapt his Lemma 1 part (c) for the sequence of models and the drifting parameter that we consider in this paper. We proceed in two steps. First, we construct a triangular array $\{\tilde{y}_{n,t} : 1 \leq t \leq n, n \geq 1\}$ that verify (C.11). Then, we prove that the constructed sequence verifies that $n^{-2}\sum_{t=1}^{n} y_{n,t-1}^2 - n^{-2}\sum_{t=1}^{n}\tilde{y}_{n,t-1}^2 = o_p(1)$.

Step 1: Define $\tilde{u}_{n,t} = u_{n,t}I\{\rho_n \geq 0\} + (-1)^t u_{n,t}I\{\rho_n < 0\}$ for all $t = 1,\ldots,n$. Note that the sequence $\{\tilde{u}_{n,t} : 1 \leq t \leq n\}$ defines a martingale difference array with the same variance $E[\tilde{u}_{n,t}^2] = \sigma_n^2$ and satisfies that $E[\tilde{u}_{n,t}^2] \in [\underline{\sigma}, \overline{\sigma}]$, and $E[\tilde{u}_{n,t}^4] \leq K_4$. Using this notation, we construct the following triangular array:

$$\tilde{y}_{n,t} = e^{-c/n}\tilde{y}_{n,t-1} + \tilde{u}_{n,t} \ , \quad \tilde{y}_{n,0} = 0 \ ,$$

where $c = \lim_{n\to\infty} n(1-|\rho_n|)$. Denote the sequence of partial sums by $S_{n,j} = \sum_{t=1}^{j} u_{n,t}$ for any $j = 1,\ldots,n$ and $S_{n,0} = 0$. Let us define the following random process

$$X_n(r) = \frac{1}{\sqrt{n}}\frac{1}{\sigma_n}S_{n,[nr]} = \frac{1}{\sqrt{n}}\frac{1}{\sigma_0}S_{n,j-1} \ \text{ if } (j-1)/n \leq r < j/n \ ,$$

and $X_n(1) = \frac{1}{\sqrt{n}}\frac{1}{\sigma_n}S_{n,n}$. By a functional central limit theorem for martingale difference arrays (see Theorem 27.14 in Davidson (1994)), we claim that $\{X_n(r) : r \in [0,1]\}$ converges to the standard Brownian motion process $\{W(r) : r \in [0,1]\}$. To use this result, we prove

$$\text{(a)} \quad \sum_{t=1}^{n}\frac{\tilde{u}_{n,t}^2}{n\sigma_n^2} \xrightarrow{p} 1 \ , \quad \text{(b)} \quad \max_{t\leq n}\left|\frac{\tilde{u}_{n,t}}{\sqrt{n}\sigma_n}\right| \xrightarrow{p} 0 \ , \quad \text{(c)} \quad \lim_{n}\sum_{t=1}^{[nr]}\frac{E[\tilde{u}_{n,t}^2]}{n\sigma_n^2} = r \ .$$

We can verify condition (a) using $\tilde{u}_{n,t}^2 = u_{n,t}^2$, Chebyshev's inequality and Assumption B.1:

$$P\left(\left|\sum_{t=1}^{n}\frac{\tilde{u}_{n,t}^2}{n\sigma_n^2} - 1\right| > \epsilon\right) \leq \frac{n^{-2}}{\epsilon^2}\sum_{i=1}^{n}E[u_{n,t}^4] \leq \frac{K_4}{\epsilon^2 n} \to 0 \quad \text{as } n\to\infty \ ,$$

for any $\epsilon > 0$. To verify condition (b) holds is sufficient to show that

$$nE\left[\frac{\tilde{u}_{n,t}^2}{n\sigma_n^2}I\left\{\left|\frac{\tilde{u}_{n,t}}{\sqrt{n}\sigma_n}\right| > c\right\}\right] \to 0 \ ,$$

for any $c > 0$, where $I\{\cdot\}$ is the indicator function. If the previous display holds, then condition (b) follows by theorem 23.16 in Davidson (1994). To verify the previous condition, note that

$$nE\left[\frac{\tilde{u}_{n,t}^2}{n\sigma_n^2}I\left\{\left|\frac{\tilde{u}_{n,t}}{\sqrt{n}\sigma_n}\right| > c\right\}\right] \leq nE\left[\frac{\tilde{u}_{n,t}^4}{n^2\sigma_n^4c^2}I\left\{\left|\frac{\tilde{u}_{n,t}}{\sqrt{n}\sigma_n}\right| > c\right\}\right] \leq \frac{E[\tilde{u}_{n,t}^4]}{n\sigma_n^4c^2} \ \leq \ \frac{K_4}{n\underline{\sigma}^2c^2} \ ,$$

where the last inequality uses $\tilde{u}_{n,t}^4 = u_{n,t}^4$ and Assumption B.1. Finally, condition (c) holds since $E[\tilde{u}_{n,t}^2] = \sigma_n^2$.

Using the functional central limit theorem, the continuous mapping theorem, and $\sigma_n \to \sigma_0$, we can repeat the arguments presented in the proof of Lemma 1 in Phillips (1987) to conclude that $n^{-2}\sum_{t=1}^n \tilde{y}_{n,t-1}^2 \xrightarrow{d} \sigma_0^2 \int_0^1 J_{-c}(r)^2 dr$.

Step 2: Define $a_n = |\rho_n| e^{c/n}$. We know $y_{n,t} = \sum_{\ell=1}^i \rho_n^{i-\ell} u_{n,t}$ and $\tilde{y}_{n,t} = \sum_{\ell=1}^i e^{-c(i-\ell)/n}\tilde{u}_{n,t}$; therefore, $\tilde{y}_{n,t} = y_{n,t} - R_{n,t}$ if $\rho_n \geq 0$, and $\tilde{y}_{n,t} = (-1)^t y_{n,t} - R_{n,t}$ if $\rho_n < 0$, where

$$R_{n,t} = \sum_{\ell=1}^{t} \left(a_n^{t-\ell} - 1\right) e^{-c(t-\ell)/n} \tilde{u}_{n,\ell} \ .$$

Therefore, we conclude that $y_{n,t}^2 = \tilde{y}_{n,t}^2 + 2\tilde{y}_{n,t}R_{n,t} + R_{n,t}^2$. This implies that

$$\left| \frac{1}{n^2}\sum_{t=1}^n y_{n,t-1}^2 - \frac{1}{n^2}\sum_{t=1}^n \tilde{y}_{n,t-1}^2 \right| \leq \left| \frac{2}{n^2}\sum_{t=1}^n \tilde{y}_{n,t-1}R_{n,t} \right| + \left| \frac{1}{n^2}\sum_{t=1}^n R_{n,t}^2 \right| \ .$$

By Cauchy–Schwartz's inequality, the right-hand side of the previous expression is lower than or equal to

$$2\left(\frac{1}{n^2}\sum_{i=1}^n \tilde{y}_{n,t-1}^2\right)^{1/2} \left(\frac{1}{n^2}\sum_{i=1}^n R_{n,t}^2\right)^{1/2} + \left|\frac{1}{n^2}\sum_{i=1}^n R_{n,t}^2\right| \ . \tag{C.12}$$

By the result at the end of Step 1, we have $n^{-2}\sum_{i=1}^n \tilde{y}_{n,t-1}^2$ is $O_p(1)$. Therefore, it is sufficient to show $n^{-2}\sum_{i=1}^n R_{n,t}^2 \xrightarrow{p} 0$ to conclude that (C.12) converges to zero in probability.

To verify the claim, we first observe that $a_n^j e^{-(i-\ell)c/n} = |\rho_n|^j e^{-(i-\ell-j)c/n} \leq |\rho_n|^j$ for all $j = 0, \ldots, i-\ell-1$, which implies that

$$\begin{aligned} |R_{n,t}| &= \left| \sum_{\ell=1}^t (a_n - 1)(1 + a_n + \ldots + a_n^{i-\ell-1}) e^{-(t-\ell)c/n} \tilde{u}_{n,\ell} \right| \\ &\leq |a_n - 1| \sum_{\ell=1}^t (1 + |\rho_n| + \ldots + |\rho_n|^{t-\ell-1}) |\tilde{u}_{n,\ell}| \ . \end{aligned}$$

Using the previous inequality and $|\rho_n|^j \leq (1 + M/n)^j \leq (1 + M/n)^n \leq e^M$, we obtain

$$|R_{n,t}| \leq e^M |a_n - 1| \sum_{\ell=1}^t (t-\ell)|\tilde{u}_{n,\ell}| \leq e^M |n(a_n - 1)| \sum_{\ell=1}^n |u_{n,\ell}| \ ,$$

for all $t = 1, \ldots, n$, where we used that $|\tilde{u}_{n,\ell}| = |u_{n,\ell}|$ in the last inequality. Then, we derive

$$n^{-2} \sum_{t=1}^{n} R_{n,t}^2 \leq e^{2M} |n(a_n - 1)|^2 \left(n^{-1} \sum_{\ell=1}^{n} |u_{n,\ell}| \right)^2 .$$

By Markovs's inequality and Assumption B.1, we obtain that $n^{-1} \sum_{\ell=1}^{n} |u_{n,\ell}|$ is $O_p(1)$. Analyzing $a_n - 1 = e^{c/n} \left(|\rho_n| - e^{-c/n} \right)$, we can conclude that $n(a_n - 1) = o(1)$, which implies that the right-hand side of the previous display converges to zero in probability. As a result, we conclude that (C.12) converges to zero in probability, which implies that $n^{-2} \sum_{i=1}^{n} y_{n,t-1}^2 - n^{-2} \sum_{i=1}^{n} \tilde{y}_{n,t-1}^2 = o_p(1)$, and by the result at the end of Step 1, we conclude (C.11). ■

C.5 Proof of Theorem B.1

Additional Notation: Define $\xi_{n,t}(\rho_n, h_n) = y_{n,t+h_n} - \beta(\rho_n, h_n) y_{n,t}$ and recall $\beta(\rho, h) \equiv \rho^h$. Algebra shows

$$\xi_{n,t}(\rho_n, h_n) = \sum_{\ell=1}^{h_n} \rho_n^{h_n - \ell} u_{n,t+\ell} . \tag{C.13}$$

Proof. The derivations presented on pages 1811 and 1812 in Montiel Olea and Plagborg-Møller (2021a) imply

$$R_n(h_n) = \frac{\sum_{t=1}^{n-h_n} \xi_{n,t}(\rho_n, h_n) \hat{u}_{n,t}(h_n)}{\left(\sum_{t=1}^{n-h_n} \hat{\xi}_{n,t}(h_n)^2 \hat{u}_{n,t}(h_n)^2 \right)^{1/2}} ,$$

which is equal to

$$\left(\frac{\sum_{t=1}^{n-h_n} \xi_{n,t}(\rho_n, h_n) u_{n,t}}{(n-h_n)^{1/2} V(\rho_n, h_n)^{1/2}} + \frac{\sum_{t=1}^{n-h} \xi_{n,t}(\rho_n, h_n)(\hat{u}_{n,t}(h_n) - u_{n,t})}{(n-h_n)^{1/2} V(\rho_n, h_n)^{1/2}} \right) \times \left(\frac{(n-h_n) V(\rho_n, h_n)}{\sum_{t=1}^{n-h_n} \hat{\xi}_{n,t}(h_n)^2 \hat{u}_{n,t}(h_n)^2} \right)^{1/2} ,$$

where $V(\rho, h) = E[\xi_{n,t}(h)^2 u_{n,t}^2]$. We then follow their approach and prove that under Assumption B.1: for any sequences $\{\rho_n : n \geq 1\} \subset [-1 - M/n, 1 + M/n]$, $\{\sigma_n^2 : n \geq 1\} \subset [\underline{\sigma}, \overline{\sigma}]$, and $\{h_n : n \geq 1\}$ satisfying $h_n = o(n)$ and $h_n \leq n$, we have

$$(i) \quad \frac{\sum_{t=1}^{n-h_n} \xi_{n,t}(\rho_n, h_n) u_{n,t}}{(n-h_n)^{1/2} V(\rho_n, h_n)^{1/2}} \xrightarrow{d} N(0,1) \ ,$$

$$(ii) \quad \frac{\sum_{t=1}^{n-h_n} \xi_{n,t}(\rho_n, h_n)(\hat{u}_{n,t}(h_n) - u_{n,t})}{(n-h_n)^{1/2} V(\rho_n, h_n)^{1/2}} \xrightarrow{p} 0 \ ,$$

$$(iii) \quad \frac{\sum_{t=1}^{n-h_n} \hat{\xi}_{n,t}(h_n)^2 \hat{u}_{n,t}(h_n)^2}{(n-h_n) V(\rho_n, h_n)} \xrightarrow{p} 1 \ .$$

Finally, Lemmas C.4, C.5, and C.7 imply (i), (ii), and (iii), respectively. ∎

Lemma C.1. *Suppose Assumption B.1 holds. Then, for any (ρ_n, σ_n, h_n) such that $|\rho_n| \leq 1 + M/n$, $\sigma_n^2 \in [\underline{\sigma}, \overline{\sigma}]$, and $h_n \leq n$, we have*

$$E\left[\xi_{n,t}(\rho_n, h_n)^4\right] \leq 4 g(\rho_n, h_n)^4 K_4 \ ,$$

where $g(\rho, k) = \left(\sum_{\ell=0}^{k-1} \rho^{2\ell}\right)^{1/2}$ and $\xi_{n,t}(\rho_n, h_n)$ is in (C.13).

Proof. It follows from the proof of Lemma A.7 in Montiel Olea and Plagborg-Møller (2021a). ∎

Lemma C.2. *Suppose Assumption B.1 holds. Then, for any (ρ_n, σ_n, h_n) such that $|\rho_n| \leq 1 + M/n$, $\sigma_n^2 \in [\underline{\sigma}, \overline{\sigma}]$, and $h_n \leq n$, we have*

$$E\left[\left(\frac{\sum_{t=1}^{n-h_n} \xi_{n,t}(\rho_n, h_n) y_{n,t-1}}{(n-h_n) g(\rho_n, n-h_n) g(\rho_n, h_n) \sigma_n^2}\right)^2\right] \leq \frac{n}{n-h_n} \times \frac{h_n}{n-h_n} \times \frac{\sqrt{4K_4}}{\underline{\sigma}} \ ,$$

where $g(\rho, k) = \left(\sum_{\ell=0}^{k-1} \rho^{2\ell}\right)^{1/2}$ and $\xi_{n,t}(\rho_n, h_n)$ is in (C.13).

Proof. The definition of $\xi_{n,t}(\rho_n, h_n)$ in (C.13) implies $\sum_{t=1}^{n-h_n} \xi_{n,t}(\rho_n, h_n) y_{n,t-1} = \sum_{j=1}^{n} u_{n,j} b_{n,j}$, where $b_{n,j} = \sum_{t=j-h_n}^{j-1} \rho_n^{t+h_n-j} \, y_{n,t-1} I\{1 \leq t \leq n-h\}$. Note that $b_{n,j}$ is measurable with respect to the σ-algebra defined by $\{u_{n,k} : 1 \leq k \leq j-2\}$. Using Assumption B.1, we obtain

$$E\left[\left(\sum_{j=1}^{n} u_{n,j} b_{n,j}\right)^2\right] = \sum_{j=1}^{n} E[u_{n,j}^2 b_{n,j}^2] = \sigma_n^2 \sum_{j=1}^{n} E[b_{n,j}^2] \ .$$

Therefore, the derivation above implies $E\left[\left(\sum_{t=1}^{n-h_n} \xi_{n,t}(\rho_n, h_n) y_{n,t-1}\right)^2\right] = \sigma_n^2 \sum_{j=1}^{n} E[b_{n,j}^2]$.

We claim that

$$E[b_{n,j}^2] \leq h_n g(\rho_n, h_n)^2 g(\rho_n, n-h_n)^2 \sqrt{4K_4}\ , \tag{C.14}$$

for any $j = 1, \ldots, n$. The previous claim and Assumption B.1 imply

$$E\left[\left(\frac{\sum_{t=1}^{n-h_n} \xi_{n,t}(\rho_n, h_n) y_{n,t-1}}{(n-h_n) g(\rho_n, n-h_n) g(\rho_n, h_n) \sigma_n^2}\right)^2\right] \leq \frac{nh_n}{(n-h_n)^2 \sigma_n^2} \times \sqrt{4K_4} \leq \frac{nh_n}{(n-h_n)^2 \underline{\sigma}} \times \sqrt{4K_4}\ .$$

To verify (C.14), we consider three cases. The first case is $j \leq h_n$, in which we derive

$$\begin{aligned} E[b_j^2] = E[(\sum_{t=1}^{j-1} \rho^{t+h-j} y_{n,t-1})^2] &\leq (j-1) E[\sum_{t=1}^{j-1} \rho^{2(t+h-j)} y_{n,t-1}^2] \\ &\leq h_n \left(\sum_{t=1}^{j-1} \rho^{2(t+h-j)}\right) g(\rho_n, n-h_n)^2 \sqrt{4K_4} \\ &\leq h_n g(\rho_n, h_n)^2 g(\rho_n, n-h_n)^2 \sqrt{4K_4}\ , \end{aligned}$$

where we use Loeve's inequality (see Theorem 9.28 in Davidson (1994)) in the first inequality above. In the second inequality, we use $E[y_{n,t-1}^2] \leq E[y_{n,t-1}^4]^{1/2}$, $y_{n,t-1} = \xi_{n,0}(\rho_n, t-1)$, Lemma C.1, and $g(\rho_n, t-1)^2 \leq g(\rho_n, n-h_n)^2$ for all $t = 1, \ldots, n-h_n$. Note that we also use $j \leq h_n$, which also implies the last inequality above. The second case is $h_n + 1 \leq j \leq n - h_n + 1$. We follow the same approach as before and conclude $E[b_j^2] \leq h_n g(\rho_n, h_n)^2 g(\rho_n, n-h_n)^2 \sqrt{4K_4}$. In the final case, we have $j \geq n - h_n + 2$. Similarly, we obtain $E[b_j^2] \leq h_n g(\rho_n, h_n)^2 g(\rho_n, n-h_n)^2 \sqrt{4K_4}$. ■

Lemma C.3. *Suppose Assumption B.1 holds. Then, for any (ρ_n, σ_n, h_n) such that $|\rho_n| \leq 1 + M/n$, $\sigma_n^2 \in [\underline{\sigma}, \overline{\sigma}]$, and $h_n \leq n$, we have*

$$E\left[\left(\frac{\sum_{t=1}^{n-h_n} u_{n,t} y_{n,t-1}}{(n-h_n)^{1/2} g(\rho_n, n-h_n)}\right)^2\right] \leq 2K_4\ ,$$

where $g(\rho, k) = \left(\sum_{\ell=0}^{k-1} \rho^{2\ell}\right)^{1/2}$.

Proof. It follows from the proof of Lemma E.8 in Montiel Olea and Plagborg-Møller (2021a). ■

Lemma C.4. *Suppose Assumptions B.1 hold. Then, for any sequences $\{\rho_n : n \geq 1\}$ such that $|\rho_n| \leq 1 + M/n$, $\{\sigma_n^2 : n \geq 1\} \subset [\underline{\sigma}, \overline{\sigma}]$, and $\{h_n : n \geq 1\}$ satisfying $h_n = o(n)$ and*

$h_n \leq n$, we have

$$\frac{\sum_{t=1}^{n-h_n} \xi_{n,t}(\rho_n, h_n) u_{n,t}}{(n-h_n)^{1/2} g(\rho_n, h_n)\sigma_n^2} \xrightarrow{d} N(0,1) \ , \tag{C.15}$$

where $\xi_{n,t}(\rho_n, h_n)$ is in (C.13) and $g(\rho, h)^2 = \sum_{\ell=1}^{h} \rho^{2h}$.

Proof. We adapt the proof of Lemma A1 in Montiel Olea and Plagborg-Møller (2021a). We start by writing the left-hand side term in (C.15) as follows

$$\sum_{t=1}^{n-h_n} \chi_{n,t} \ ,$$

where

$$\chi_{n,t} = \frac{\xi_{n,n-h_n+1-t}(\rho_n, h_n) u_{n,n-h_n+1-t}}{(n-h_n)^{1/2} g(\rho_n, h_n)\sigma_n^2} \ ,$$

for $t = 1, \ldots, n-h_n$. Define the σ-algebra $\mathcal{F}_{n,t} = \sigma\left(u_{n-h_n+j-t} : j \geq 1\right)$. Note that for any $t = 1, \ldots, n-h_n$, $\chi_{n,t}$ is measurable with respect to $\mathcal{F}_{n,t}$. Therefore, the sequence $\{\chi_{n,t} : 1 \leq t \leq n-h_n\}$ is adapted to the filtration $\{\mathcal{F}_{n,t} : 1 \leq t \leq n-h_n\}$. Moreover, $\xi_{n,n-h_n+1-t}(\rho_n, h_n)$ is measurable with respect to $\mathcal{F}_{n,t-1}$ since it is a function of $\{u_{n,n-h_n+j-(t-1)} : 1 \leq j \leq h_n\}$. This implies that $E[\chi_{n,t}|\mathcal{F}_{n,t-1}] = 0$ since

$$E[\chi_{n,t}|\mathcal{F}_{n,t-1}] = \frac{\xi_{n,n-h_n+1-t}(\rho_n, h_n)}{(n-h_n)^{1/2} g(\rho_n, h_n)\sigma_n^2} E[u_{n,n-h_n+1-t}|\mathcal{F}_{n,t-1}]$$

and by Assumptions B.1 we conclude $E[u_{n,n-h_n+1-t}|\mathcal{F}_{n,t-1}] = E[u_{n,n-h_n+1-t}] = 0$.

The derivation presented above proves that the sequence $\{\chi_{n,t} : 1 \leq t \leq n-h_n\}$ is a martingale difference array with respect to the filtration $\{\mathcal{F}_{n,t} : 1 \leq t \leq n-h_n\}$. The result in (C.15) then follows by a martingale central limit theorem (Theorem 24.3 in Davidson (1994)), which requires

$$(i) \quad \sum_{t=1}^{n-h_n} E[\chi_{n,t}^2] = 1 \ , \quad (ii) \quad \sum_{t=1}^{n-h_n} \chi_{n,t}^2 \xrightarrow{p} 1 \ , \quad (iii) \quad \max_{1 \leq t \leq n-h_n} |\chi_{n,t}| \xrightarrow{p} 0 \ .$$

The condition (i) follows by using that $E[\xi_{n,t}(\rho_n, h_n)^2 u_{n,t}^2] = g(\rho_n, h_n)^2 \sigma_n^4$. To prove the condition (ii) is sufficient to show

$$\text{Var}\left(\sum_{t=1}^{n-h_n} \chi_{n,t}^2\right) \to 0 \ . \tag{C.16}$$

To prove (C.16), we first recall that

$$\sum_{t=1}^{n-h_n} \chi_{n,t}^2 = \frac{1}{(n-h_n)g(\rho_n,h_n)^2\sigma_n^4} \times \sum_{t=1}^{n-h_n} \xi_{n,t}(\rho_n,h_n)^2 u_{n,t}^2 \ ,$$

where the second term of the right-hand side of the previous display can be decomposed into the sum of its expected value and another three zero mean terms:

$$(n-h_n)g(\rho_n,h_n)^2\sigma_n^4 + \sum_{j=1}^{n}(u_{n,j}^2-\sigma_n^2)b_{n,j} + \sum_{j=1}^{n} u_{n,j}d_{n,j} + \sum_{j=1}^{n}(u_{n,j}^2-\sigma_n^2)g(\rho_n,h_n)^2\sigma_n^2 \ ,$$

where $b_{n,j} = \sum_{t=j-h_n}^{j-1} \rho_n^{2(h_n+t-j)} u_{n,t}^2 I\{1 \le t \le n-h_n\}$ and

$$d_{n,j} = \sum_{t=j-h_n}^{j-1} \sum_{\ell_2=1}^{j-t-1} u_{n,t+\ell_2}\rho_n^{2h-j+t-\ell_2} u_{n,t}^2 I\{1 \le t \le n-h_n\} \ .$$

Note that $b_{n,j}$ and $d_{n,j}$ are measurable with respect to the $\sigma-$algebra $\sigma(u_{n,k} : 1 \le k \le j-1\}$. By Assumptions B.1 and Loeve's inequality (Theorem 9.28 in Davidson (1994)), we conclude

$$E[b_{n,j}^2] \le h_n \sum_{t=j-h_n}^{j-1} E[\rho_n^{4(h_n+t-j)} u_{n,t}^4] \le h_n g(\rho_n,h_n)^4 K_4$$

and

$$E[d_{n,j}^2] \le h_n E\left[\left(\sum_{\ell_2=1}^{j-i-1} u_{n,i+\ell_2}\rho_n^{2h-j+i-\ell_2} u_{n,t}^2\right)^2\right] \le h_n g(\rho_n,h_n)^4 \sigma_n^2 K_4$$

for all $j = 1, \dots, n$.

We use the decomposition presented above, Assumptions B.1, and Loeve's inequality (see Theorem 9.28 in Davidson (1994)) imply that the left-hand side of (C.16) is lower than or equal to

$$\frac{3\sum_{j=1}^{n} E[(u_{n,j}^2-\sigma_n^2)^2 b_{n,j}^2] + 3\sum_{j=1}^{n} E[u_{n,j}^2 d_{n,j}^2] + 3g(\rho_n,h_n)^4\sigma_n^4 \sum_{j=1}^{n} E[(u_{n,j}^2-\sigma_n^2)^2]}{(n-h_n)^2 g(\rho_n,h_n)^4 \sigma_n^8} \ .$$

By Assumptions B.1 and the upper bounds that we found for $E[b_{n,j}^2]$ and $E[d_{n,j}^2]$, the previous expression is lower than or equal to

$$\frac{3nh_nK_4^2}{(n-h_n)^2\underline{\sigma}^4}+\frac{3nh_nK_4}{(n-h_n)^2\underline{\sigma}^2}+\frac{3nK_4}{(n-h_n)^2\underline{\sigma}^2}\;.$$

The previous expression is $o\,(1)$ since $h_n = o\,(n)$ as $n \to \infty$. This implies that (C.16) holds.

Finally, to verify that condition (iii) holds is sufficient to show

$$(n-h_n)E\left[\chi_{n,t}^2 I\{|\chi_{n,t}^2|>c\}\right]\to 0\;, \tag{C.17}$$

for any $c > 0$, where $I\{\cdot\}$ is the indicator function. If the condition in (C.20) holds, then condition (iii) follows by Theorem 23.16 in Davidson (1994). To verify (C.17), note that

$$\begin{aligned}(n-h_n)E\left[\chi_{n,t}^2 I\{|\chi_{n,t}^2|>c\}\right] &\leq\; (n-h_n)E\left[\frac{\chi_{n,t}^4}{c^2}I\{|\chi_{n,t}^2|>c\}\right]\\ &\leq\; (n-h_n)\frac{E\left[\chi_{n,t}^4\right]}{c^2}\\ &= \frac{E\left[\xi_{n,n-h_n+1-t}(\rho_n,h_n)^4\right]E\left[u_{n,n-h_n+1-t}^4\right]}{(n-h_n)\sigma_n^8 g(\rho_n,h_n)^4c^2}\\ &\leq \frac{4K_4E\left[u_{n,n-h_n+1-t}^4\right]}{(n-h_n)\sigma_n^8c^2}\;,\end{aligned}$$

where the equality above uses Assumption B.1, and the last inequality follows by Lemma C.1. By Assumption B.1 we obtain $(n-h_n)E\left[\chi_{n,t}^2 I\{|\chi_{n,t}^2|>c\}\right] \leq 4K_4^2/((n-h_n)\underline{\sigma}^4c^2)$, which is sufficient to conclude (C.17). ∎

Lemma C.5. *Suppose Assumption B.1 holds. Then, for any sequences $\{\rho_n : n \geq 1\}$ such that $|\rho_n| \leq 1 + M/n$, $\{\sigma_n^2 : n \geq 1\} \subset [\underline{\sigma}, \overline{\sigma}]$, and $\{h_n : n \geq 1\}$ satisfying $h_n = o(n)$ and $h_n \leq n$, we have*

$$\frac{\sum_{t=1}^{n-h}\xi_{n,t}(\rho_n,h_n)(\hat{u}_{n,t}(h_n)-u_{n,t})}{(n-h_n)^{1/2}g(\rho_n,h_n)\sigma_n^2}\overset{p}{\to}0\;, \tag{C.18}$$

where $\hat{u}_{n,t}(h_n) = y_{n,t} - \hat{\rho}_n(h_n)y_{n,t-1}$, $\hat{\rho}_n(h_n)$ is defined in (B.2), and $g(\rho,h)^2 = \sum_{\ell=1}^{h}\rho^{2h}$.

Proof. A proof can be adapted from the proof of Lemma A.4 and Lemma E.8 in Montiel Olea and Plagborg-Møller (2021a). Importantly, their Assumption 3 (relevant for the proof) holds due to Proposition B.1 in Appendix B. ∎

Lemma C.6. *Suppose Assumption B.1 holds. Then, for any sequences $\{\rho_n : n \geq 1\}$ such that $|\rho_n| \leq 1 + M/n$, $\{\sigma_n^2 : n \geq 1\} \subset [\underline{\sigma}, \overline{\sigma}]$, and $\{h_n : n \geq 1\}$ satisfying $h_n = o(n)$ and*

$h_n \le n$, we have

$$(i) \quad \frac{\hat{\beta}_n(h_n) - \beta(\rho_n, h_n)}{g(\rho_n, h_n)} \xrightarrow{p} 0 \ ,$$

$$(ii) \quad \frac{g(\rho_n, n-h_n)\left(\hat{\eta}(\rho_n, h_n) - \eta(\rho_n, h_n)\right)}{g(\rho_n, h_n)} \xrightarrow{p} 0 \ ,$$

$$(iii) \quad (n-h_n)^{1/2} \times g(\rho_n, n-h_n) \times (\hat{\rho}_n(h_n) - \rho_n) = O_p(1) \ ,$$

where $\hat{\eta}(\rho_n, h_n) = \rho_n \hat{\beta}_n(h_n) + \hat{\gamma}_n(h_n)$, $\eta(\rho_n, h_n) = \rho_n \beta(\rho_n, h_n) = \rho_n^{h_n+1}$, $\hat{\beta}_n(h_n)$ and $\hat{\gamma}_n(h_n)$ are defined in (B.1), $\hat{\rho}_n(h_n)$ is in (B.2), and $g(\rho, h)^2 = \sum_{\ell=1}^{h} \rho^{2h}$.

Proof. A proof can be adapted from the proof of Lemma A.3 in Montiel Olea and Plagborg-Møller (2021a). Importantly, their Assumption 3 (relevant for the proof) holds due to Proposition B.1 in Appendix B. ■

Lemma C.7. *Suppose Assumptions B.1 holds. Then, for any sequences $\{\rho_n : n \ge 1\}$ such that $|\rho_n| \le 1 + M/n$, $\{\sigma_n^2 : n \ge 1\} \subset [\underline{\sigma}, \overline{\sigma}]$, and $\{h_n : n \ge 1\}$ satisfying $h_n = o(n)$ and $h_n \le n$, we have*

$$\frac{\sum_{t=1}^{n-h_n} \hat{\xi}_{n,t}(h_n)^2 \hat{u}_{n,t}(h_n)^2}{(n-h_n) g(\rho_n, h_n)^2 \sigma_n^4} \xrightarrow{p} 1 \ ,$$

where $\hat{\xi}_{n,t}(h_n) = y_{n,t+h_n} - \hat{\beta}_n(h_n) y_{n,t} - \hat{\gamma}_n(h_n) y_{n,t-1}$, $\hat{u}_{n,t}(h_n) = y_{n,t} - \hat{\rho}_n(h_n) y_{n,t-1}$, $\hat{\beta}_n(h_n)$ and $\hat{\gamma}_n(h_n)$ are defined in (B.1), $\hat{\rho}_n(h_n)$ is defined in (B.2), and $g(\rho, h)^2 = \sum_{\ell=1}^{h} \rho^{2h}$.

Proof. We adapt the proof of Lemma A.2 in Montiel Olea and Plagborg-Møller (2021a) presented in their Supplemental Appendix E.2. They claim that is sufficient to prove

$$\frac{\sum_{t=1}^{n-h_n} \hat{\xi}_{n,t}(h_n)^2 \hat{u}_{n,t}(h_n)^2}{(n-h_n) g(\rho_n, h_n)^2 \sigma_n^4} - \frac{\sum_{t=1}^{n-h_n} \xi_{n,t}(\rho_n, h_n)^2 u_{n,t}^2}{(n-h_n) g(\rho_n, h_n)^2 \sigma_n^4} \xrightarrow{p} 0 \ , \tag{C.19}$$

since they then can conclude using their Lemma A6, which implies

$$\frac{\sum_{t=1}^{n-h_n} \xi_{n,t}(\rho_n, h_n)^2 u_{n,t}^2}{(n-h_n) g(\rho_n, h_n)^2 \sigma_n^4} \xrightarrow{p} 1 \ .$$

We avoid using their Lemma A6 since its proof requires that the shocks have a finite 8th moment. Instead, we observe that (C.16) presented in the proof of Lemma C.4 implies the previous claim.

To verify (C.19), Montiel Olea and Plagborg-Møller prove that is sufficient to show that

$$\frac{\sum_{t=1}^{n-h_n}\left[\hat{\xi}_{n,t}(h_n)^2\hat{u}_{n,t}(h_n)^2-\xi_{n,t}(\rho_n,h_n)^2u_{n,t}^2\right]^2}{(n-h_n)g(\rho_n,h_n)^2\sigma_n^4} \tag{C.20}$$

converges in probability to zero. To prove that, they derive the following upper bound for (C.20):

$$3[(\hat{R}_1)^{1/2}\times(\hat{R}_2)^{1/2}]+3[(\hat{R}_3)^{1/2}\times(\hat{R}_4)^{1/2}]+3[(\hat{R}_1)^{1/2}\times(\hat{R}_3)^{1/2}]\ ,$$

where

$$\hat{R}_1=\frac{\sum_{t=1}^{n-h_n}\left[\xi_{n,t}(\rho_n,h_n)-\hat{\xi}_{n,t}(h_n)\right]^4}{(n-h_n)g(\rho_n,h_n)^4\sigma_n^8}\ ,\qquad \hat{R}_2=\frac{\sum_{t=1}^{n-h_n}u_{n,t}^4}{n-h_n}$$

$$\hat{R}_3=\frac{\sum_{t=1}^{n-h_n}\left[\hat{u}_{n,t}(h_n)-u_{n,t}\right]^4}{n-h_n}\ ,\qquad \hat{R}_4=\frac{\sum_{t=1}^{n-h_n}\xi_{n,t}(\rho_n,h_n)^4}{(n-h_n)g(\rho_n,h_n)^4\sigma_n^8}\ .$$

In what follows, we use Assumptions B.1 to prove that (i) $\hat{R}_1$ and $\hat{R}_3$ are $o_p(1)$ and (ii) $\hat{R}_2$ and $\hat{R}_4$ are $O_p(1)$, which are sufficient to conclude that (C.20) converges to zero in probability.

To verify $\hat{R}_1$ is $o_p(1)$, let us first observe that

$$\xi_{n,t}(\rho_n,h_n)-\hat{\xi}_{n,t}(h_n)=[\hat{\beta}_n(h_n)-\beta(\rho_n,h_n)]u_{n,t}+[\hat{\eta}_n(\rho_n,h_n)-\eta(\rho_n,h_n)]y_{n,t-1}\ ,$$

where $\hat{\eta}(\rho_n,h_n)=\rho_n\hat{\beta}_n(h_n)+\hat{\gamma}_n(h_n)$ and $\eta(\rho_n,h_n)=\rho_n\beta(\rho_n,h_n)$. Then, using Loeve's inequality (see Theorem 9.28 in Davidson (1994)), we obtain

$$\begin{aligned}\hat{R}_1\leq\ & 8\left(\frac{[\hat{\beta}_n(h_n)-\beta(\rho_n,h_n)]}{g(\rho_n,h_n)}\right)^4\left(\frac{\sum_{t=1}^{n-h_n}u_{n,t}^4}{(n-h_n)\sigma_n^8}\right)\\ &+8\left(\frac{g(\rho_n,n-h_n)[\hat{\eta}_n(\rho_n,h_n)-\eta(\rho_n,h_n)]}{g(\rho_n,h_n)}\right)^4\left(\frac{\sum_{t=1}^{n-h_n}y_{n,t-1}^4}{(n-h_n)g(\rho_n,n-h_n)^4\sigma_n^8}\right)\ .\end{aligned}$$

Note that the first term on the right-hand side in the previous expression goes to zero in probability due to part (i) in Lemma C.6, Markov's inequality, and Assumptions B.1. The second term on the right-hand side in the previous expression goes to zero in probability due to part (ii) in Lemma C.6, Markov's inequality, and using that

$$E[y_{n,t-1}^4]=E[\xi_{n,0}(\rho_n,t-1)^4]\leq g(\rho_n,n-h_n)^44K_4\ , \tag{C.21}$$

where the inequality holds due to Lemma C.1 and $g(\rho_n, t-1)^4 \leq g(\rho_n, n-h_n)^4$ for all $i \leq n - h_n$. This completes the proof of $\hat{R}_1$ is $o_p(1)$.

To prove that $\hat{R}_3$ is $o_p(1)$, note that we can write:

$$\hat{R}_1 = (g(\rho_n, h_n)[\hat{\rho}_n(h_n) - \rho])^4 \frac{\sum_{t=1}^{n-h_n} y_{t-1}^4}{(n-h_n)g(\rho_n, h_n)^4}$$

since $\hat{u}_{n,t}(h_n) - u_{n,t} = (\hat{\rho}_n(h_n) - \rho)y_{t-1}$. Note that the right-hand side in the previous expression goes to zero in probability due to part (iii) in Lemma C.6, Markov's inequality, and using (C.21). This completes the proof of $\hat{R}_3$ is $o_p(1)$. Finally, Markov's inequality and Assumptions B.1 implies that $\hat{R}_2$ is $O_p(1)$. While Markov's inequality and Lemma C.1 implies $\hat{R}_4$ is $O_p(1)$. ■

C.6 Proof of Proposition B.2

Proof. For any $\epsilon > 0$, define the event $E_{n,\epsilon} = \{|R_n(1)| \leq c_n^*(1, 1-\alpha) - \psi(\epsilon)\}$, where $\psi(\epsilon) = z_{1-\alpha/2-3\epsilon/2} - z_{1-\alpha/2-2\epsilon}$ and z_α is the α-quantile of the standard normal distribution.

We will prove that for any (small) $\epsilon > 0$ the following claims hold,

$$\lim_{n\to\infty} P_n\left([1/L, L] \subseteq C^*_{la-ar}(h_n, 1-\alpha) \mid E_{n,\epsilon}\right) = 1 \ , \tag{C.22}$$

$$\liminf_{n\to\infty} P_n(E_{n,\epsilon}) \geq 1 - \alpha - 4\epsilon \ . \tag{C.23}$$

These two claims are sufficient to conclude that

$$\liminf_{n\to\infty} P_n\left([1/L, L] \subseteq C^*_{la-ar}(h_n, 1-\alpha)\right) \geq 1 - \alpha - 4\epsilon \ .$$

Since this holds for any (small) $\epsilon >$), it implies the claim of the proposition.

Claim 1: (C.22) holds. To verify this, we first rewrite the lower and upper bounds of $C^*_{la-ar}(h, 1-\alpha)$ using the definition of $R_n(1)$ as in (6):

$$\left(\hat{\beta}_n(1) - \hat{s}_n(1)c_n^*(1, 1-\alpha)\right)^{h_n} = (1 - c_1/n - \zeta_{n,1})^{h_n} \tag{C.24}$$

$$\left(\hat{\beta}_n(1) + \hat{s}_n(1)c_n^*(1, 1-\alpha)\right)^{h_n} = (1 - c_1/n + \zeta_{n,2})^{h_n} \tag{C.25}$$

where $\zeta_{n,1} = \hat{s}_n(1)\{-R_n(1) + c_n^*(1, 1-\alpha)\}$ and $\zeta_{n,2} = \hat{s}_n(1)\{R_n(1) + c_n^*(1, 1-\alpha)\}$.

Note that we have $\zeta_{n,1} \geq \hat{s}_n(1)\psi(\epsilon)$ and $\zeta_{n,2} \geq \hat{s}_n(1)\psi(\epsilon)$ conditional $E_{n,\epsilon}$. Additionally, we can obtain that

$$\hat{s}_n(1) = c_2^{1/2} n^{-1/4}(1 + o_p(1)) \ , \tag{C.26}$$

which follows by the formula in (4), Lemma C.7 and part 2 of Lemma B.1(adapted for the i.i.d. case that we consider in this proposition), and because $h_n/n^{1/2} \to c_2$ as $n \to \infty$.

Since $h_n/\sqrt{n} \to c_2 > 0$ as $n \to \infty$, it holds that $\lim_{n\to\infty}(1 - Cn^{-1/4})^{h_n} \to 0$ for any positive constant C. This implies that conditional on $E_{n,\epsilon}$, we have that the lower bound of $C^*_{la-ar}(h, 1-\alpha)$ goes to zero. To see this, consider the following derivations using (C.24),

$$(1 - c_1/n - \zeta_{n,1})^{h_n} \overset{(1)}{\leq} \Big(1 - c_2^{1/2} n^{-1/4}(1 + o_p(1))\psi(\epsilon)\Big)^{h_n}$$

where (1) holds by definition of $\zeta_{n,1}$ conditional on E_n, (C.26), and because $\psi(\epsilon)$ is positive by definition. Since $1 + o_p(1)$ is larger than $1 - \delta$ for any small $\delta > 0$ with a high probability for any n sufficiently larger, we can conclude that the right-hand side of the previous display goes to zero with a high probability conditional on E_n.

The previous derivation concludes that the lower bound of $C^*_{la-ar}(h, 1-\alpha)$ goes to zero conditional on E_n, which implies that the lower bound is asymptotically lower than $1/L$ conditional on E_n. Now we will show that the upper bound of $C^*_{la-ar}(h, 1-\alpha)$ is asymptotically larger than L conditional on E_n. To see this, consider the following derivation using (C.25),

$$\begin{aligned}(1 - c_1/n + \zeta_{n,2})^{h_n} &\overset{(1)}{\geq} \Big(1 - c_1/n + c_2^{1/2} n^{-1/4}(1 + o_p(1))\psi(\epsilon)\Big)^{h_n} \\ &\overset{(2)}{\geq} 1 + (-c_1/n + c_2^{1/2} n^{-1/4}(1 + o_p(1))\psi(\epsilon))h_n \\ &= 1 + O(n^{-1/2}) + c_2^{1/2} n^{1/4}(1 + o_p(1))(h_n/\sqrt{n})\end{aligned}$$

where (1) holds by definition of $\zeta_{n,2}$ conditional on E_n, (C.26), and because $\psi(\epsilon)$ is positive by definition, and (2) holds by Bernoulli's inequality. Since $1 + o_p(1)$ is larger than $1 - \delta$ for any small $\delta > 0$ with a high probability for any n sufficiently larger, we can conclude that the right-hand side of the previous display goes to infinity with a high probability conditional on E_n. In particular, the upper bound of $C^*_{la-ar}(h, 1-\alpha)$ is asymptotically larger than L conditional on E_n. This completes the proof of claim 1.

Claim 2: (C.23) holds. We first note that the following inclusion

$$\{|R_n(1)| \leq z_{1-\alpha/2-2\epsilon}\} \subseteq \{|R_n(1)| \leq c_n^*(1, 1-\alpha) - \psi(\epsilon)\} = E_{n,\epsilon}$$

holds with a high probability due to part 1 of Lemma B.3 (adapted for the i.i.d. case that we consider in this proposition). We then observe that the probability of the left-hand side of the previous expression goes to $1-\alpha-4\epsilon$ due to Theorem B.1. This completes the proof of claim 2. ■

D Proof of Auxiliary Results: Asymptotic Refinements

D.1 Proof of the Lemma B.4

Proof. By Lemma D.1, there exists a random variable $\tilde{R}_n(h)$ such that

$$P_\rho\left(\left|R_n(h)-\tilde{R}_n(h)\right|>n^{-1-\epsilon}\right)\leq Cn^{-1-\epsilon}\left(E[|u_t|^k]+E[u_t^{2k}]+E[u_t^{4k}]\right) \ ,$$

where

$$\tilde{R}_n(h)\equiv(n-h)^{1/2}\mathcal{T}\left(\frac{1}{n-h}\sum_{t=1}^{n-h}X_t;\sigma^2,\psi_4^4,\rho\right)$$

and the sequence $\{X_t:1\leq t\leq n-h\}$ is defined in (B.4). Due to Lemma D.1, we know $\mathcal{T}$ is a polynomial. Define $\tilde{J}_n(x,h,P,\rho)=P_\rho\left(\tilde{R}_n(h)\leq x\right)$. Using Bonferroni's inequality, we conclude

$$|P_\rho\left(R_n(h)\leq x\right)-P_\rho\left(\tilde{R}_n(h)\leq x\right)|\leq D_n+P_\rho\left(\left|R_n(h)-\tilde{R}_n(h)\right|>n^{-1-\epsilon}\right) \ .$$

Therefore, $\sup_{x\in\mathbf{R}}|J_n(x,h,P,\rho)-\tilde{J}_n(x,h,P,\rho)|\leq D_n+Cn^{-1-\epsilon}\left(E[|u_t|^k]+E[u_t^{2k}]+E[u_t^{4k}]\right)$, which completes the proof of the Lemma. Note that the constant C is defined in Lemma D.1 and only depends on a, h, k, C_σ, and c_u. ■

Lemma D.1. *Suppose Assumption 5.1 holds. For any fixed $h\in\mathbf{N}$ and $a\in(0,1)$. Then, for any $\rho\in[-1+a,1-a]$ and $\epsilon\in(0,1/2)$, there exist a constant $C=C(a,h,k,C_\sigma,c_u)>0$, where $k\geq 8(1+\epsilon)/(1-2\epsilon)$, and a real-valued function $\mathcal{T}(\cdot\ ;\sigma^2,\psi_4^4,\rho):\mathbf{R}^8\to\mathbf{R}$, such that*

1. $\mathcal{T}(\mathbf{0};\sigma^2,\psi_4^4,\rho)=0$,
2. *$\mathcal{T}(x;\sigma^2,\psi_4^4,\rho)$ is a polynomial of degree 3 in $x\in\mathbf{R}^8$ with coefficients depending continuously differentiable on σ^2, ψ_4^4, and ρ,*
3. $P_\rho\left(\left|R_n(h)-\tilde{R}_n(h)\right|>n^{-1-\epsilon}\right)\leq Cn^{-1-\epsilon}\left(E[|u_t|^k]+E[u_t^{2k}]+E[u_t^{4k}]\right)$,

where $\sigma^2 = E_P[u_1^2]$, $\psi_4^4 = E_P[u_1^4]$,

$$\tilde{R}_n(h) \equiv (n-h)^{1/2}\mathcal{T}\left(\frac{1}{n-h}\sum_{t=1}^{n-h} X_t; \sigma^2, \psi_4^4, \rho\right)$$

and the sequence $\{X_t : 1 \leq t \leq n-h\}$ *is defined in* (B.4). *Furthermore, the asymptotic variance of* $\tilde{R}_n(h)$ *equals one.*

Proof. The proof has two main parts. We first use Lemmas D.2 and D.3 to approximate $R_n(h)$ using functions based on ξ_t, u_t, and y_{t-1}. We then replace y_{t-1} by z_{t-1}. We specifically define the polynomial $\mathcal{T}$.

Part 1: The derivations presented on page 1811 in Montiel Olea and Plagborg-Møller (2021a) implies

$$R_n(h) = \frac{\sum_{t=1}^{n-h}\xi_t(h)\hat{u}_t(h)}{\left(\sum_{t=1}^{n-h}\hat{\xi}_t(h)^2\hat{u}_t(h)^2\right)^{1/2}} ,$$

where $\xi_t(\rho,h) = \sum_{\ell=1}^{h}\rho^{h-\ell}u_{t+\ell}$, $\hat{u}_t(h) = y_t - \hat{\rho}_n(h)y_{t-1}$, $\hat{\xi}_t(h) = y_{t+h} - \left(\hat{\beta}_n(h)y_t + \hat{\gamma}_n(h)y_{t-1}\right)$, and the coefficients $(\hat{\beta}_n(h), \hat{\gamma}_n(h))$ is as in (3) and $\hat{\rho}_n(h)$ is defined in (5). Define

$$f_n = \frac{\sum_{t=1}^{n-h}\xi_t(\rho,h)\hat{u}_t(h)}{n-h} \quad \text{and} \quad g_n = \frac{\sum_{t=1}^{n-h}\hat{\xi}_t(h)^2\hat{u}_t(h)^2}{V(n-h)} - 1 ,$$

where $V = E[\xi_t(\rho,h)^2u_t^2] = \sigma^4\sum_{\ell=1}^{h}\rho^{2(h-\ell)}$. It follows that

$$R_n(h) = (n-h)^{1/2}V^{-1/2}f_n\left(1+g_n\right)^{-1/2} .$$

Lemmas D.2, D.3, and D.4 imply

$$P((n-h)^{1/2}|V^{-1/2}f_n| > \delta) \leq C\delta^{-k}\left(E[|u_t|^k] + E[u_t^{2k}] + E[u_t^{4k}]\right)$$

and

$$P((n-h)^{1/2}|g_n| > \delta) \leq C\delta^{-k}\left(E[|u_t|^k] + E[u_t^{2k}] + E[u_t^{4k}]\right) .$$

Step 1: Define $\tilde{R}_{g,n} = (n-h)^{1/2}f_nV^{-1/2}\left(1 - \frac{1}{2}g_n + \frac{3}{8}g_n^2\right)$. Due to Lemma D.4, we have

$$P\left((n-h)^{3/2}\left|R_n(h) - \tilde{R}_{g,n}\right| > \delta^4\right) \leq C\delta^{-k}\left(E[|u_t|^k] + E[u_t^{2k}] + E[u_t^{4k}]\right)$$

since $P\left(n^{3/2}\left|(1+g_n)^{-1/2}-\left(1-\frac{1}{2}g_n+\frac{3}{8}g_n^2\right)\right|>\delta^3\right)\leq C\delta^{-k}\left(E[|u_t|^k]+E[u_t^{2k}]+E[u_t^{4k}]\right)$.

Step 2: Define

$$\tilde{R}_{f,n}=(n-h)^{1/2}V^{-1/2}\left(f_{1,n}+f_{2,n}+f_{3,n}\right)\left(1-\frac{1}{2}g_n+\frac{3}{8}g_n^2\right) \ ,$$

where $f_n=\sum_{j=1}^4 f_{j,n}$ as in Lemma D.2. We conclude

$$\begin{aligned}P\left((n-h)^{3/2}\left|\tilde{R}_{g,n}-\tilde{R}_{f,n}\right|>\delta^4\right)&=P\left((n-h)^{4/2}\left|V^{-1/2}f_{4,n}\left(1-\frac{1}{2}g_n+\frac{3}{8}g_n^2\right)\right|>\delta^4\right)\\&\leq C\delta^{-k}\left(E[|u_t|^k]+E[u_t^{2k}]+E[u_t^{4k}]\right) \ ,\end{aligned}$$

where the last inequality follows by Lemmas D.4 and D.2.

Step 3: Define

$$\tilde{R}_{fg,n}=(n-h)^{1/2}V^{-1/2}\left(f_{1,n}+f_{2,n}+f_{3,n}-\frac{1}{2}f_{1,n}g_n-\frac{1}{2}f_{2,n}g_n+\frac{3}{8}f_{1,n}g_n^2\right) \ .$$

Lemmas D.2, D.3, and D.4 imply

$$P\left((n-h)^{3/2}\left|\tilde{R}_{f,n}-\tilde{R}_{fg,n}\right|>\delta^4\right)\leq C\delta^{-k}\left(E[|u_t|^k]+E[u_t^{2k}]+E[u_t^{4k}]\right) \ .$$

Step 4: Define

$$\tilde{R}_{y,n}(h)=(n-h)^{1/2}V^{-1/2}\left(\sum_{j=1}^3 f_{j,n}-\frac{1}{2}f_{1,n}g_{1,n}-\frac{1}{2}f_{1,n}g_{2,n}-\frac{1}{2}f_{2,n}g_{1,n}+\frac{3}{8}f_{1,n}g_{1,n}^2\right) \ ,$$

where $g_n=\sum_{j=1}^3 g_{j,n}$ as in Lemma D.3. We use Lemmas D.2, D.3, and D.4 to conclude

$$P\left((n-h)^{3/2}\left|\tilde{R}_{fg,n}-\tilde{R}_{y,n}(h)\right|>\delta^4\right)\leq C\delta^{-k}\left(E[|u_t|^k]+E[u_t^{2k}]+E[u_t^{4k}]\right) \ .$$

Step 5: By Bonferroni's:

$$P\left((n-h)^{3/2}\left|R_n(h)-\tilde{R}_{y,n}(h)\right|>\delta^4\right)\leq C\delta^{-k}\left(E[|u_t|^k]+E[u_t^{2k}]+E[u_t^{4k}]\right) \ .$$

Part 2: We consider $V^{1/2}\mathcal{T}(x;\sigma^2,\psi_4^4,\rho)$ is equal to the following polynomial

$$f_1(x)+f_2(x)+f_3(x)-\frac{1}{2}\left(f_1(x)g_1(x)+f_1(x)g_2(x)+f_2(x)g_1(x)\right)+\frac{3}{8}f_1(x)g_1(x)^2\ , \quad \text{(D.1)}$$

where $f_1(x)=x_1,\ f_2(x)=-\sigma^{-2}(1-\rho^2)x_5x_6,\ f_3(x)=\sigma^{-4}(1-\rho^2)x_5x_6(2\rho x_5+x_2),\ g_1(x)=V^{-1}x_3$, and

$$g_2(x)=V^{-1}\left(\frac{\psi_4^4}{\sigma^4}x_1^2-\frac{2}{\sigma^2}x_1x_4+(1-\rho^2)\left(g(\rho,h)^2x_5^2-\frac{2}{\sigma^2}x_5x_7+x_6^2-\frac{2}{\sigma^2}x_6x_8\right)\right)\ .$$

Note that $\tilde{R}_{y,n}(h)=V^{1/2}\mathcal{T}\left((n-h)^{-1}\sum_{t=1}^{n-h}\tilde{X}_t;\sigma^2,\psi_4^4,\rho\right)$, where

$$\tilde{X}_t=(\xi_tu_t,u_t^2-\sigma^2,(\xi_tu_t)^2-V,\xi_tu_t^3,u_ty_{t-1},\xi_ty_{t-1},\xi_t^2u_ty_{t-1},\xi_tu_t^2y_{t-1})\ .$$

Since $z_t=y_t+\rho^tz_0$, it follows that

$$P\left((n-h)^{1/2}\left|\frac{\sum_{t=1}^{n-h}f_ty_{t-1}}{n-h}-\frac{\sum_{t=1}^{n-h}f_tz_{t-1}}{n-h}\right|>\delta\right)\leq C\delta^{-k}\left(E[u_t^{2k}]+E[u_t^{4k}]\right)$$

for $f_t=u_t,\xi_t,\xi_t^2u_t,\xi_tu_t^2$. Then, Lemma D.4 and step 5 in part 1 implies

$$P\left((n-h)^{3/2}\left|R_n(h)-\tilde{R}_n(h)\right|>\delta^4\right)\leq C\delta^{-k}\left(E[|u_t|^k]+E[u_t^{2k}]+E[u_t^{4k}]\right)\ ,$$

where $\tilde{R}_n(h)=V^{1/2}\mathcal{T}\left((n-h)^{-1}\sum_{t=1}^{n-h}X_t;\sigma^2,\psi_4^4,\rho\right)$ and the sequence $\{X_t:1\leq t\leq n-h\}$ is defined in (C.13). As we mentioned before, the constant C includes the constants C's that appear in Lemmas D.2, D.3, D.5 and D.6 that only depends on a, h, k, C_σ, and c_u. Finally, we take $\delta=n^{(1/2-\epsilon)/4}$ ■

Lemma D.2. *Suppose Assumption 5.1 holds. For any fixed $h,k\in\mathbf{N}$ and $a\in(0,1)$. Define*

$$f_n=\frac{\sum_{t=1}^{n-h}\xi_t(\rho,h)\hat{u}_t(h)}{n-h}$$

where $\xi_t(\rho,h)=\sum_{\ell=1}^{h}\rho^{h-\ell}u_{t+\ell}$, $\hat{u}_t(h)=y_t-\hat{\rho}_n(h)y_{t-1}$, and $\hat{\rho}_n(h)$ is as in (5). Then, for any $\rho\in[-1+a,1-a]$, there exists a constant $C=C(h,k,a,C_\sigma)$ such that we can write

$$f_n=f_{1,n}+f_{2,n}+f_{3,n}+f_{4,n}\ ,$$

where

$$P\left((n-h)^{j/2}|f_{j,n}|\geq\delta^j\right)\leq C\delta^{-k}\left(E[|u_t|^k]+E[u_t^{2k}]+E[u_t^{4k}]\right)$$

for any $\delta<(n-h)^{1/2}$ *and* $j\in\{1,2,3,4\}$.

Proof. In what follows we use $\xi_t=\xi_t(\rho,h)$. Using the definition of $\hat{\rho}_n(h)$, we obtain

$$f_n=\frac{\sum_{t=1}^{n-h}\xi_t u_t}{n-h}-\left(\frac{\sum_{t=1}^{n-h}u_t y_{t-1}}{n-h}\right)\left(\frac{\sum_{t=1}^{n-h}\xi_t y_{t-1}}{n-h}\right)\left(\frac{\sum_{t=1}^{n-h}y_{t-1}^2}{n-h}\right)^{-1}.$$

Let us define the components of f_n as follows:

$$\begin{aligned}
f_{1,n}&=\left(\frac{\sum_{t=1}^{n-h}\xi_t u_t}{n-h}\right)\\
f_{2,n}&=-\frac{(1-\rho^2)}{\sigma^2}\left(\frac{\sum_{t=1}^{n-h}u_t y_{t-1}}{n-h}\right)\left(\frac{\sum_{t=1}^{n-h}\xi_t y_{t-1}}{n-h}\right)\\
f_{3,n}&=\frac{(1-\rho^2)}{\sigma^4}\left(\frac{\sum_{t=1}^{n-h}u_t y_{t-1}}{n-h}\right)\left(\frac{\sum_{t=1}^{n-h}\xi_t y_{t-1}}{n-h}\right)\left(2\rho\frac{\sum_{t=1}^{n-h}u_t y_{t-1}}{n-h}+\frac{\sum_{t=1}^{n-h}u_t^2-\sigma^2}{n-h}\right)\\
f_{4,n}&=f_n-(f_{1,n}+f_{2,n}+f_{3,n})\ .
\end{aligned}$$

Note that by construction $f_n=\sum_{j=1}^4 f_{j,n}$. Lemma D.6 guarantees that each sample average in $f_{1,n}$, $f_{2,n}$, and $f_{3,n}$ verify the conditions to use Lemma D.4, which imply that $P\left((n-h)^{j/2}|f_{j,n}|\geq\delta^j\right)\leq C\delta^{-k}E[u_t^{2k}]$ for $j=1,2,3$, where the constant C includes C_σ and a.

To prove $P\left((n-h)^{4/2}|f_{j,n}|\geq\delta^4\right)\leq C\delta^{-k}\left(E[|u_t|^k]+E[u_t^{2k}]+E[u_t^{4k}]\right)$, we proceed in two steps.

Step 1: Define

$$W_n=\frac{\sum_{t=1}^{n-h}(1-\rho^2)\sigma^{-2}y_{t-1}^2}{n-h}-1\ .$$

We can use (1) and algebra to derive the following identity

$$W_n=-\frac{\sigma^{-2}y_{n-h}^2}{n-h}+2\rho\sigma^{-2}\frac{\sum_{t=1}^{n-h}u_t y_{t-1}}{n-h}+\sigma^{-2}\frac{\sum_{t=1}^{n-h}u_t^2-\sigma^2}{n-h}\ .$$

This implies that

$$f_{3,n} = \frac{(1-\rho^2)}{\sigma^2}\left(\frac{\sum_{t=1}^{n-h} u_t y_{t-1}}{n-h}\right)\left(\frac{\sum_{t=1}^{n-h} \xi_t y_{t-1}}{n-h}\right)\left(W_n + \frac{\sigma^{-2} y_{n-h}^2}{n-h}\right) .$$

Therefore, $f_{4,n} = f_n - f_{1,n} - f_{2,n} - f_{3,n}$ is equal to

$$-\frac{(1-\rho^2)}{\sigma^2}\left(\frac{\sum_{t=1}^{n-h} u_t y_{t-1}}{n-h}\right)\left(\frac{\sum_{t=1}^{n-h} \xi_t y_{t-1}}{n-h}\right)\left((1+W_n)^{-1} - (1-W_n) + \frac{\sigma^{-2} y_{n-h}^2}{n-h}\right) .$$

Step 2: Due to Lemma D.4, it is sufficient to show that

$$P\left((n-h)^{2/2}|(1+W_n)^{-1} - (1-W_n)| > \delta^2\right) \leq C\delta^{-k} E[u_t^{2k}] \tag{D.2}$$

and

$$P\left((n-h)^{2/2}\left|\frac{\sigma^{-2} y_{n-h}^2}{n-h}\right| > \delta^2\right) \leq C\delta^{-k} E[|u_t|^k] . \tag{D.3}$$

Note that (D.2) follows by part 5 in Lemma D.4 since $P(n^{1/2}|W_n| > \delta) \leq C\delta^{-k}E[u_t^{2k}]$ due to Lemma D.6 and $\delta < n^{1/2}$. Finally, (D.3) follows by Markov's inequality and Lemma D.5. As we mentioned before, the constant C includes the constants C's that appear in Lemmas D.5 and D.6 that only depends on a, h, k, and C_σ. ∎

Lemma D.3. *Suppose Assumption 5.1 holds. For any fixed $h, k \in \mathbf{N}$ and $a \in (0,1)$. Define*

$$g_n = \frac{\sum_{t=1}^{n-h} \hat{\xi}_t(h)^2 \hat{u}_t(h)^2}{V(n-h)} - 1$$

where $V = \sigma^2 \sum_{\ell=1}^{h} \rho^{2(h-\ell)}$, $\hat{\xi}_t(h)$ is as in (3) $\hat{u}_t(h) = y_t - \hat{\rho}_n(h) y_{t-1}$, and $\hat{\rho}_n(h)$ is as in (5). Then, for any $\rho \in [-1+a, 1-a]$, there exists a constant $C = C(h,k,a,C_\sigma,c_u)$ such that we can write

$$g_n = g_{1,n} + g_{2,n} + g_{3,n} \ ,$$

where

$$P\left((n-h)^{j/2}|g_{j,n}| \geq \delta^j\right) \leq C\delta^{-k}\left(E[|u_t|^k] + E[u_t^{2k}] + E[u_t^{4k}]\right)$$

for any $\delta < (n-h)^{1/2}$ and $j \in \{1,2,3\}$.

Proof. In what follows we use $\xi_t = \xi_t(\rho, h) = \sum_{\ell=1}^{h} \rho^{h-\ell} u_{t+\ell}$. As we did for the case of f_n in Lemma D.2, we utilize the linear regression formulas to define the components of g_n as functions of the sample average of functions of ξ_t, u_t, and y_{t-1}:

$$
\begin{aligned}
Vg_{1,n} &= \frac{\sum_{t=1}^{n-h}(\xi_t^2 u_t^2 - V)}{n-h} \\
Vg_{2,n} &= \frac{\psi_4^4}{\sigma^4}\left(\frac{\sum_{t=1}^{n-h}\xi_t u_t}{n-h}\right)^2 - \frac{2}{\sigma^2}\left(\frac{\sum_{t=1}^{n-h}\xi_t u_t}{n-h}\right)\left(\frac{\sum_{t=1}^{n-h}\xi_t u_t^3}{n-h}\right) \\
&\quad + (1-\rho^2)g(\rho,h)^2\left(\frac{\sum_{t=1}^{n-h}u_t y_{t-1}}{n-h}\right)^2 - \frac{2(1-\rho^2)}{\sigma^2}\left(\frac{\sum_{t=1}^{n-h}u_t y_{t-1}}{n-h}\right)\left(\frac{\sum_{t=1}^{n-h}\xi_t^2 u_t y_{t-1}}{n-h}\right) \\
&\quad + (1-\rho^2)\left(\frac{\sum_{t=1}^{n-h}\xi_t y_{t-1}}{n-h}\right)^2 - \frac{2(1-\rho^2)}{\sigma^2}\left(\frac{\sum_{t=1}^{n-h}\xi_t y_{t-1}}{n-h}\right)\left(\frac{\sum_{t=1}^{n-h}\xi_t u_t^2 y_{t-1}}{n-h}\right) \\
Vg_{3,n} &= Vg_n - V\left(g_{1,n}+g_{2,n}\right) \ ,
\end{aligned}
$$

where $\psi_4^4 = E[u_t^4]$ and $g(\rho,h)^2 = \sum_{\ell=1}^{h}\rho^{2(h-\ell)}$. By construction $g_n = \sum_{j=1}^{3} g_{j,n}$. Lemma D.6 guarantees that each sample average in $g_{1,n}$ and $g_{2,n}$ verify the conditions to use Lemma D.4, which imply

$$
P\left((n-h)^{j/2}|g_{j,n}| \geq \delta^j\right) \leq C\delta^{-k}\left(E[u_t^{2k}] + E[u_t^{4k}]\right)
$$

for $\delta < (n-h)^{1/2}$ and $j \in \{1,2\}$, where the constant C includes C_σ, a, and c_u (since $\psi_4^4 \leq 24e^{c_u^2}$). In what follows we prove $P\left((n-h)^{3/2}|g_{3,n}| \geq \delta^3\right) \leq C\delta^{-k}\left(E[u_t^{2k}] + E[u_t^{4k}]\right)$.

First, we write $Vg_{3,n} = R_{g,1} + R_{g,2}$, where $R_{g,1}$ and $R_{g,2}$ are specified below. We will prove that $P((n-h)^{3/2}|V^{-1}R_{g,j}| > \delta^3) \leq C\delta^{-k}E[u_t^{4k}]$ for $j = 1,2$. To compute $g_{3,n}$ we use equation (3) and the following equality

$$
\hat{\xi}_t(h) = \xi_t - (\hat{\beta}_n(h) - \beta(\rho,h))u_t - \hat{\eta}_n(\rho,h)y_{t-1} \ ,
$$

where $\hat{\eta}_n(\rho,h) = \rho\hat{\beta}_n(h) + \hat{\gamma}_n(h)$. We also use that $\hat{u}_t(h) = y_t - \hat{\rho}_n(h)y_{t-1} = u_t - (\hat{\rho}_n(h) - \rho)y_{t-1}$. In what follows, we denote $\hat{\beta} = \hat{\beta}_n(h)$, $\hat{\rho} = \hat{\rho}_n(h)$ and $\hat{\eta} = \hat{\eta}_n(\rho,h)$, and $\sum = \sum_{t=1}^{n-h}$ to simplify the heavy notation. We obtain $R_{g,1}$ is equal to

$$
\begin{aligned}
&2\left[(\eta-\hat{\eta}) + \frac{(1-\rho^2)}{\sigma^2}\left(\frac{\sum\xi_t y_{t-1}}{n-h}\right)\right]\left(\frac{\sum\xi_t u_t^2 y_{t-1}}{n-h}\right) + \frac{\sigma^4}{1-\rho^2}\left\{(\eta-\hat{\eta})^2 - \frac{(1-\rho^2)^2}{\sigma^4}\left(\frac{\sum\xi_t y_{t-1}}{n-h}\right)^2\right\} \\
&+ 2\left[(\rho-\hat{\rho}) + \frac{(1-\rho^2)}{\sigma^2}\left(\frac{\sum u_t y_{t-1}}{n-h}\right)\right]\left(\frac{\sum\xi_t^2 u_t y_{t-1}}{n-h}\right) + \psi_4^4\left\{(\beta-\hat{\beta})^2 - \frac{1}{\sigma^4}\left(\frac{\sum\xi_t u_t}{n}\right)^2\right\} \\
&+ 2\left[(\beta-\hat{\beta}) + \frac{1}{\sigma^2}\left(\frac{\sum\xi_t u_t}{n-h}\right)\right]\left(\frac{\sum\xi_t u_t^3}{n-h}\right) + \frac{\sigma^4 g_2^2}{1-\rho^2}\left\{(\rho-\hat{\rho})^2 - \frac{(1-\rho^2)^2}{\sigma^4}\left(\frac{\sum u_t y_{t-1}}{n-h}\right)^2\right\} \ ,
\end{aligned}
$$

and we obtain that $R_{g,2}$ is equal to

$$
\begin{aligned}
&(\beta-\hat{\beta})^2\left(\frac{\sum u_t^4}{n-h}-\psi_4^4\right)+(\rho-\hat{\rho})^2\left(\frac{\sum \xi_t^2 y_{t-1}^2}{n-h}-\frac{\sigma^4 g(\rho,h)^2}{1-\rho^2}\right)+(\eta-\hat{\eta})^2\left(\frac{\sum u_t^2 y_{t-1}^2}{n-h}-\frac{\sigma^4}{1-\rho^2}\right)\\
&2(\beta-\hat{\beta})^2(\rho-\hat{\rho})\left(\frac{\sum u_t^3 y_{t-1}}{n-h}\right)+2(\eta-\hat{\eta})^2(\rho-\hat{\rho})\left(\frac{\sum u_t y_{t-1}^3}{n-h}\right)+2(\beta-\hat{\beta})(\rho-\hat{\rho})^2\left(\frac{\sum \xi_t u_t y_{t-1}^2}{n-h}\right)\\
&+4(\beta-\hat{\beta})(\rho-\hat{\rho})\left(\frac{\sum \xi_t u_t^2 y_{t-1}}{n-h}\right)+2(\eta-\hat{\eta})(\rho-\hat{\rho})^2\left(\frac{\sum \xi_t y_{t-1}^3}{n-h}\right)+4(\eta-\hat{\eta})(\rho-\hat{\rho})\left(\frac{\sum \xi_t u_t y_{t-1}^2}{n-h}\right)\\
&+2(\beta-\hat{\beta})(\eta-\hat{\eta})\left(\frac{\sum u_t^3 y_{t-1}}{n-h}\right)+2(\beta-\hat{\beta})(\eta-\hat{\eta})(\rho-\hat{\rho})^2\left(\frac{\sum u_t y_{t-1}^3}{n-h}\right)\\
&+(\beta-\hat{\beta})^2(\rho-\hat{\rho})^2\left(\frac{\sum u_t^2 y_{t-1}^2}{n-h}-\frac{\sigma^4}{1-\rho^2}\right)+4(\beta-\hat{\beta})(\eta-\hat{\eta})(\rho-\hat{\rho})\left(\frac{\sum u_t^2 y_{t-1}^2}{n-h}-\frac{\sigma^4}{1-\rho^2}\right)\\
&+(\eta-\hat{\eta})^2(\rho-\hat{\rho})^2\left(\frac{\sum y_{t-1}^4}{n-h}-\frac{E[u_t^4]}{1-\rho^4}-6\frac{\rho^2\sigma^4}{(1-\rho^2)(1-\rho^4)}\right)+(\eta-\hat{\eta})^2(\rho-\hat{\rho})^2\frac{E[u_t^4]}{1-\rho^4}\\
&+\frac{\sigma^4}{1-\rho^2}(\beta-\hat{\beta})^2(\rho-\hat{\rho})^2+4(\beta-\hat{\beta})(\eta-\hat{\eta})(\rho-\hat{\rho})\frac{\sigma^4}{1-\rho^2}+(\eta-\hat{\eta})^2(\rho-\hat{\rho})^2\frac{6\rho^2\sigma^4}{(1-\rho^2)(1-\rho^4)}\ .
\end{aligned}
$$

Note that $P((n-h)^{3/2}|V^{-1}R_{g,1}|>\delta^3)\leq C\delta^{-k}\left(E[u_t^{2k}]+E[u_t^{4k}]\right)$ follows by Lemmas D.4, D.6, and D.7, since each term between parenthesis in the definition of $R_{g,2}$ appears in Lemma D.6, the terms in brackets appears in items 1-4 of Lemma D.7, and the terms between curly brackets can be written as the product of terms like parenthesis and brackets terms. Similarly, $P((n-h)^{3/2}|V^{-1}R_{g,2}|>\delta^3)\leq C\delta^{-k}\left(E[u_t^{2k}]+E[u_t^{4k}]\right)$ follows by Lemmas D.4, D.6, and D.7, since each term between parenthesis in the definition of $R_{g,1}$ appears in Lemma D.6 or in items 5-8 of Lemma D.7. ■

Lemma D.4. *Let $\{W_{n,j}:1\leq j\leq r\}$ be a sequence of random variables. Suppose that there exist constants c_j and C such that*

$$P(n^{1/2}|W_{n,j}|>c_j\delta)\leq C\delta^{-k}\ ,$$

for $j=1,..,r$ and some $k\in\mathbf{N}$. Then, for any $r\geq 2$ and $\delta<n^{1/2}$, we have

1. $P(n^{1/2}|\sum_{j=1}^r W_{n,j}|>(\sum_{j=1}^r c_j)\delta)\leq rC\delta^{-k}$.
2. $P(n^{r/2}|\prod_{j=1}^r W_{n,j}|>(\prod_{j=1}^r c_j)\delta^r)\leq rC\delta^{-k}$.
3. $P(n^{1/2}|W_{n,1}+\prod_{j=2}^r W_{n,j}|>(c_1+\prod_{j=2}^r c_j)\delta)\leq 2C\delta^{-k}$.
4. *If $c_1 n^{-1/2}\delta<1$. Then, $P(|W_{n,1}|>1-b)\leq C\delta^{-k}$ for any $b\in(0,1-c_1 n^{-1/2}\delta)$.*

5. *If* $c_1 n^{-1/2}\delta < 1$*. Then, for any* $b \in (0, 1 - c_1 n^{-1/2}\delta)$*, we have*

$$P\left(n^{3/2}|(1+W_{n,1})^{-1/2} - (1 - \frac{1}{2}W_{n,1} + \frac{3}{8}W_{n,1}^2)| > \frac{5}{16}b^{-7/2}c_1^3\delta^3\right) \leq 4C\delta^{-k}$$

and

$$P\left(n^{2/2}|(1+W_{n,1})^{-1} - (1 - W_{n,1})| > \frac{2}{1+(1-b)^3}\delta^2\right) \leq 3C\delta^{-k} \ .$$

Proof. Bonferroni's inequality and $\{|W_{n,1}| > 1-b\} \subseteq \{n^{1/2}|W_{n,1}| > c_1\delta\}$ for $b \leq 1 - c_1 n^{-1/2}\delta$ imply the proof of items 1–4. To prove the first part of item 5, we use Bonferroni's inequality to conclude that the left-hand side in item 5 is lower than or equal to the sum of $P(|W_{n,1}| > 1-b)$ and

$$P\left(n^{3/2}|(1+W_{n,1})^{-1/2} - (1 - \frac{1}{2}W_{n,1} + \frac{3}{8}W_{n,1}^2)| > \frac{5}{16}b^{-7/2}c_1^3\delta^3, |W_{n,1}| \leq 1-b\right) \ .$$

Item 4 implies that the former term is bounded by $C\delta^{-k}$, while the latter term is lower than or equal to

$$P\left(n^{3/2}\frac{5}{16}b^{-7/2}|W_{n,1}|^3 > \frac{5}{16}b^{-7/2}c_1^3\delta^3, |W_{n,1}| \leq 1-b\right) \ ,$$

where the left-hand side term inside the previous probability used the Taylor Polynomial error and $|W_{n,1}| \leq 1-b$. By item 2, the above probability is lower than or equal to $3C\delta^{-k}$. Finally, adding the upper and lower bounds concludes the first part of item 5. The second part is analogous. ■

Lemma D.5. *Suppose Assumption 5.1 holds. For fixed* $a \in (0,1)$ *and* $k \geq 1$*. Then, for any* $|\rho| \leq 1-a$*, there exists a constant* $C = C(a,k) > 0$ *such that*

$$E[|y_n|^k] \leq C\, E[|u_n|^k] \ , \ \forall n \geq 1 \ .$$

and $P(n^{-1/2}|y_n^s| > \delta) \leq C\,\delta^{-k}E[|u_n|^{sk}] \ , \ \forall n \geq 1 \ ,$

Proof. The proof goes by induction. For $k=1$, we have $|y_n| \leq |\rho||y_{n-1}| + |u_n|$, which implies that

$$E[|y_n|] \leq E[|u_n|]\left(1 + |\rho| + ... + |\rho|^{n-1}\right) \leq E[|u_n|]\ a^{-1} \ .$$

Therefore, the constant $C = a^{-1}$. We can also derive $y_n^2 = \rho^2 y_{n-1}^2 + 2\rho y_{n-1}u_n + u_n^2$, which implies $E[y_n^2] \leq \rho^2 E[y_{n-1}^2] + 2|\rho|E[|y_{n-1}u_n|] + E[u_n^2]$. Using that $E[|y_{n-1}u_n|] = E[|y_{n-1}|]E[|u_n|] \leq a^{-1}E[u_n^2]$, we conclude $E[y_n^2] \leq \rho^2 E[y_{n-1}^2] + (2|\rho|a^{-1}+1)E[u_n^2]$, which implies $E[y_n^2] \leq$

$(2|\rho|a^{-1}+1)E[u_n^2]\left(1+\rho^2+...+\rho^{2(n-1)}\right) \leq C_1(a)E[u_n^2]$, where $C_1(a) = (2(1-a)/a+1)/a$. In this case, the constant $C = C_1(a)$.

Let us use $C_1(a)$ to construct $C_2(a)$ and so on. Suppose we have already computed $C_k(a)$. Now, let us compute $C_{k+1}(a)$. We have

$$y_n^{2k+1} = \rho^{2k+1}y_{n-1}^{2k+1} + \sum_{\ell=1}^{2k} C_\ell^{2k+1}\rho^{2k+1-\ell}y_{n-1}^{2k+1-\ell}u_n^\ell + u_n^{2k+1} \ .$$

By triangular inequality and similar arguments as before, we obtain

$$E[y_n^{2k+1}] \leq |\rho|^{2k+1}E[|y_{n-1}|^{2k+1}] + \sum_{\ell=1}^{2k} C_\ell^{2k+1}|\rho|^{2k+1-\ell}E[|y_{n-1}|^{2k+1-\ell}]E[|u_n^\ell|] + E[|u_n|^{2k+1}] \ ,$$

and by the inductive hypothesis, we know $E[|y_{n-1}|^{2k+1-\ell}]E[|u_n^\ell|] \leq C_{k-\ell/2}E[|u_n|^{2k+1}]$, where we used that $E[|X|^a]E[|X|^b] \leq E[|X|^{a+b}]$. Thus, the inductive hypothesis implies that

$$E[y_n^{2k+1}] \leq |\rho|^{2k+1}E[|y_{n-1}|^{2k+1}] + E[|u_n|^{2k+1}]\left(\sum_{\ell=1}^{2k} C_\ell^{2k+1}|\rho|^{2k+1-\ell}C_{k-\ell/2}+1\right) \ ,$$

in a similar way as in the initial case, we conclude. Note that the final constant C only involves a and k. The other case is analogous. ∎

Lemma D.6. *Suppose Assumption 5.1 holds. For a given $h, k \in \mathbf{N}$ and $a \in (0,1)$. Then, for any $|\rho| \leq 1-a$ and $h \in \mathbf{N}$, there exist a constant $C = C(a,k,r,s,h,C_\sigma) > 0$ such that*

1. $P(\ (n-h)^{1/2}|(n-h)^{-1}\sum_{t=1}^{n-h} u_t^r y_{t-1}^s - m_{r,s}| > \delta) \leq C\delta^{-k}E\left[|u_t|^{(r+s)k}\right]$
2. $P(\ (n-h)^{1/2}|(n-h)^{-1}\sum_{t=1}^{n-h} \xi_t u_t^r y_{t-1}^s| > \delta) \leq C\delta^{-k}E\left[|u_t|^{(1+r+s)k}\right]$
3. $P(\ (n-h)^{1/2}|(n-h)^{-1}\sum_{t=1}^{n-h} \xi_t^2 u_t^2 - V| > 3\delta) \leq C\delta^{-k}E[u_t^{4k}]$
4. $P(\ (n-h)^{1/2}|(n-h)^{-1}\sum_{t=1}^{n-h} \xi_t^2 u_t y_{t-1}| > 3\delta) \leq C\delta^{-k}E[u_t^{4k}]$
5. $P(\ (n-h)^{1/2}|(n-h)^{-1}\sum_{t=1}^{n-h} \xi_t^2 y_{t-1}^2 - \sigma^4 g(\rho,h)^2(1-\rho^2)^{-1}| > 5\delta) \leq C\delta^{-k}E[u_t^{4k}]$

for any $\delta > 0$ and $n > h$, where $\xi_t = \xi_t(\rho,h) = \sum_{\ell=1}^h \rho^{h-\ell}u_{t+\ell}$, $g(\rho,h) = \left(\sum_{\ell=1}^h \rho^{2(h-\ell)}\right)^{1/2}$, and $m_{r,s} = E\left[u_t^r\left(\sum_{j\geq 1}\rho^{j-1}u_{t-j}\right)^s\right]$.

Proof. Define the filtration $\mathcal{F}_j = \sigma(u_t : t \leq j)$. In what follows, we use Markov's inequality, Lemmas D.4, D.8, and D.5. The constant C will replace other constants and will only depend on a, k, r, s, h, C_σ, and the constants that appear in Lemmas D.8 and D.5.

Item 1: We prove this item by induction on s. First, consider $s = 0$ and $r \geq 0$. Note that $\{(u_t^r - E[u_t^r], \mathcal{F}_t) : 1 \leq t \leq n-h\}$ define a martingale difference sequence. Therefore, Markov's inequality and Lemma D.8 imply $P(\ (n-h)^{1/2}|(n-h)^{-1}\sum_{t=1}^{n-h} u_t^r - m_{r,0}| > \delta) \leq C\delta^{-k}E\left[|u_t|^{rk}\right]$, since $m_{r,0} = E[u_t^r]$. Let us suppose that item 1 holds for any (r,s) such that $r \geq 0$ and $s \leq s_0$ (this is a strong inductive hypothesis). Next, let us prove item 1 for (r, s_0+1). We write

$$(n-h)^{-1}\sum_{t=1}^{n-h} u_t^r y_{t-1}^{s_0+1} - m_{r,s_0+1} = I_1 + I_2\ ,$$

where $I_1 = (n-h)^{-1}\sum_{t=1}^{n-h}\left(u_t^r - m_{r,0}\right) y_{t-1}^{s_0+1}$ and $I_2 = (n-h)^{-1}\sum_{t=1}^{n-h} m_{r,0}\left(y_{t-1}^{s_0+1} - m_{0,s_0+1}\right)$. Note that $\{(\left(u_t^r - m_{r,0}\right) y_{t-1}^{s_0+1}, \mathcal{F}_t) : 1 \leq t \leq n-h\}$ define a martingale difference sequence; therefore, we conclude that $P((n-h)^{1/2}|I_1| > \delta) \leq C\delta^{-k}E[|u_t|^{k(r+s_0+1)}]$ using Markov's inequality and Lemmas D.8 and D.5. Now, let us write

$$y_t^{s_0+1} = (\rho y_{t-1} + u_t)^{s_0+1} = \rho^{s_0+1} y_{t-1}^{s_0+1} + \sum_{j=1}^{s_0+1}\binom{s_0+1}{j}\rho^{s_0+1-j}u_t^j y_{t-1}^{s_0+1-j}\ ,$$

which implies the following identity

$$(1-\rho^{s_0+1})(n-h)^{-1}\sum_{t=1}^{n-h} y_{t-1}^{s_0+1} = \frac{-y_t^{s_0+1}}{n-h} + \sum_{j=1}^{s_0+1}\binom{s_0+1}{j}\rho^{s_0+1-j}\left((n-h)^{-1}\sum_{t=1}^{n-h} u_t^j y_{t-1}^{s_0+1-j}\right)\ .$$

In a similar way, using $z_t = \sum_{j\geq 1}\rho^{j-1}u_{t-j}$ instead of y_t, we can derive the following identity

$$(1-\rho^{s_0+1})m_{0,s_0+1} = \sum_{j=1}^{s_0+1}\binom{s_0+1}{j}\rho^{s_0+1-j}m_{j,s_0+1-j}\ .$$

Using that $|m_{r,0}|^k = |E[u_t^r]|^k \leq E[|u_t|^{rk}]$, the previous two identities, the inductive hypothesis to $(n-h)^{-1}\sum_{t=1}^{n-h} u_t^j y_{t-1}^{s_0+1-j} - m_{j,s_0+1-j}$ for $j = 1, \ldots, s_0+1$, and Lemmas D.5 and D.4, we conclude that $P((n-h)^{1/2}|I_2| > \delta) \leq C\delta^{-k}E[|u_t|^{k(r+s_0+1)}]$, which completes the proof due to Lemma D.4.

Item 2: Consider the following derivation for a sequence of random variable $f_t \in \mathcal{F}_t$:

$$\sum_{t=1}^{n-h} \xi_t f_t = \sum_{t=1}^{n-h}\sum_{\ell=1}^{h} \rho^{h-\ell} u_{t+\ell}\ f_t = \sum_{t=1}^{n-h}\sum_{j=t+1}^{t+h} \rho^{t+h-j} u_j\ f_t = \sum_{j=1}^{n} u_j b_{n,j}\ ,$$

where $b_{n,j} = \sum_{t=j-h}^{j-1} \rho^{t+h-j}\ f_t\ I\{1 \le t \le n-h\}$. Note that $\{u_j b_{n,j} \in \mathcal{F}_j : 1 \le j \le n-h\}$ defines a martingale difference sequence and

$$\begin{aligned} E[|u_j b_{n,j}|^k] &\le h^{k-1} \sum_{t=j-h}^{j-1} |\rho|^{(t+h-j)k}\ E[|u_j|^k] E[|f_t|^k] I\{1 \le t \le n-h\} \\ &\le C_{k,h,a} E[|u_t|^k] \max_{1 \le t \le n-h} E[|f_t|^k]\ , \end{aligned}$$

where $C_{k,h,a} = h^{k-1}\sum_{\ell=1}^{h}(1-a)^{(h-\ell)k}$. Now we take $f_t = u_t^r y_{t-1}^s$ and use $E[|u_t^r y_{t-1}^s|^k] \le CE[|u_t|^{rk}]E[|u_t|^{sk}]$ due to Lemma D.5. Finally, Jensen's inequality implies $E[|u_t|^k]E[|f_t|^k] \le E[|u_t|^{(1+s+r)k}]$. We conclude due to Lemma D.8.

Item 3–4: We can write $(n-h)^{-1}\sum_{t=1}^{n-h} \xi_t^2 u_t^2 - V$ and $(n-h)^{-1}\sum_{t=1}^{n-h} \xi_t^2 u_t y_{t-1}$ as the sum three martingale difference sequences. For item 3, we use the same decomposition used to prove (C.16) in the proof of Lemma C.4. For item 4, a similar decomposition is possible. For both items, we can conclude as in the proof of item 2 by using Bonferroni's inequality and Lemma D.8; therefore, the details are omitted.

Item 5: Let us write $\sum_{t=1}^{n-h} \xi_t^2 y_{t-1}^2 - \sigma^4 g(\rho,h)^2(1-\rho^2)^{-1}$ as the sum of three terms:

$$\sum_{j=1}^{n}(u_j^2-\sigma^2)b_{n,j} + \sum_{j=1}^{n} u_j d_{n,j} + g(\rho,h)^2\sigma^4(1-\rho^2)^{-1}\sum_{j=1}^{n-h}(\sigma^{-2}(1-\rho^2)y_{t-1}^2 - 1)\ ,$$

where $b_{n,j} = \sum_{t=j-h}^{j-1} \rho^{2(h+t-j)} y_{t-1}^2 I\{1 \le t \le n-h\}$ and

$$d_{n,j} = \sum_{t=j-h}^{j-1}\sum_{\ell_2=1}^{j-t-1} u_{t+\ell_2}\rho^{2h-j+t-\ell_2} y_{t-1}^2 I\{1 \le t \le n-h\}\ .$$

Note that $(u_j^2-\sigma^2)b_{n,j}$ and $u_j d_{n,j}$ define two martingale difference sequences with respect to $\mathcal{F}_{j-1}$. By Lemma D.8 and similar derivations as in the proof of item 2, we obtain

$$I_1 = P(\ |(n-h)^{-1/2}\sum_{j=1}^{n}(u_j^2-\sigma^2)b_{n,j}| > \delta) \le \delta^{-k} C E[u_t^{4k}]$$

and

$$I_2 = P(\ |(n-h)^{-1/2} \sum_{j=1}^{n} u_j d_{n,j}| > \delta) \leq \delta^{-k} C E[u_t^{4k}]\ ,$$

for some constant C that depends on h, k, and a.

By item 1 ($r = 0$ and $s = 2$) and $m_{0,2} = \sigma^2(1-\rho^2)^{-1}$, it follows that

$$I_3 = P(\ |(n-h)^{-1/2} \frac{g(\rho,h)^2 \sigma^4}{1-\rho^2} \sum_{j=1}^{n-h} (\sigma^{-2}(1-\rho^2) y_{t-1}^2 - 1)| > \delta) \leq \delta^{-k} C E[u_t^{4k}]\ ,$$

for some constant C that depends on h, k, a, and C_σ.

Finally, Bonferroni's inequality and the previous inequalities imply

$$P(\ (n-h)^{1/2} |(n-h)^{-1} \sum_{t=1}^{n-h} \xi_t^2 y_{t-1}^2 - \sigma^4 g(\rho,h)^2 (1-\rho^2)^{-1}| > 5\delta) \leq \delta^{-k} C E[u_t^{4k}]\ ,$$

where the constant C absorbs all the previous constants. ■

Lemma D.7. *Suppose Assumption 5.1 holds. For fixed $h, k \in \mathbf{N}$ and $a \in (0,1)$. Then, for any $|\rho| \leq 1-a$, there exist a constant $C = C(a,k,h,C_\sigma) > 0$ such that*

1. $P\left((n-h)^{2/2} \left|\hat{\rho}_n(h) - \rho - (1-\rho^2)\sigma^{-2} \frac{\sum_{t=1}^{n-h} u_t y_{t-1}}{n-h}\right| > \delta^2\right) \leq C\delta^{-k} E[u_t^{2k}]$
2. $P\left((n-h)^{2/2} \left|\hat{\beta}_n(h) - \beta(\rho,h) - \sigma^{-2} \frac{\sum_{t=1}^{n-h} \xi_t(\rho,h) u_t}{n-h}\right| > \delta^2\right) \leq C\delta^{-k} E[u_t^{2k}]$
3. $P\left((n-h)^{2/2} \left|\hat{\gamma}_n(h) - \frac{(1-\rho^2)\sum_{t=1}^{n-h} \xi_t(\rho,h) y_{t-1}}{\sigma^2(n-h)} + \frac{\rho \sum_{t=1}^{n-h} \xi_t(\rho,h) u_t}{\sigma^2(n-h)}\right| > \delta^2\right) \leq C\delta^{-k} E[u_t^{2k}]$
4. $P\left((n-h)^{2/2} \left|\hat{\eta}_n(\rho,h) - \eta(\rho,h) - \frac{(1-\rho^2)\sum_{t=1}^{n-h} \xi_t y_{t-1}}{\sigma^2(n-h)}\right| > \delta^2\right) \leq C\delta^{-k} E[u_t^{2k}]$
5. $P(n^{1/2}|\hat{\rho}_n - \rho| > \delta) \leq C\delta^{-k} E[u_t^{2k}]$
6. $P(n^{1/2}|\hat{\beta}_n(h) - \beta(\rho,h)| > \delta) \leq C\delta^{-k} E[u_t^{2k}]$
7. $P(n^{1/2}|\hat{\gamma}_n(h)| > \delta) \leq C\delta^{-k} E[u_t^{2k}]$
8. $P(n^{1/2}|\hat{\eta}_n - \eta| > \delta) \leq C\delta^{-k} E[u_t^{2k}]$

for any $\delta < n^{1/2}$, where $\hat{\rho}_n(h)$ is as in (5), $(\hat{\beta}_n(h), \hat{\gamma}_n(h))$ is as in (3), and $\xi_t(\rho,h) = \sum_{\ell=1}^{h} \rho^{h-\ell} u_{t+\ell}$. $\hat{\eta}_n(\rho,h) = \rho\hat{\beta}_n(h) + \hat{\gamma}_n(h)$, $\eta(\rho,h) = \rho\beta(\rho,h)$.

Proof. To prove item 1, we first use the definition of $\hat{\rho}_n(h)$,

$$\hat{\rho}_n(h) - \rho = \frac{(n-h)^{-1}\sum_{t=1}^{n-h} u_t y_{t-1}}{(n-h)^{-1}\sum_{t=1}^{n-h} y_{t-1}^2} = \frac{(1-\rho^2)\sum u_t y_{t-1}}{\sigma^2(n-h)}(1+W_n)^{-1} ,$$

where $W_n = (1-\rho^2)\sigma^{-2}(n-h)^{-1}\sum_{t=1}^{n-h} y_{t-1}^2 - 1$. Using this notation, we have

$$\hat{\rho}_n(h) - \rho - \frac{(1-\rho^2)\sum u_t y_{t-1}}{\sigma^2(n-h)} = \frac{(1-\rho^2)\sum u_t y_{t-1}}{\sigma^2(n-h)}\left((1+W_n)^{-1} - 1\right) .$$

Since $P(n^{1/2}|(n-h)^{-1}\sum_{t=1}^{n-h} u_t y_{t-1}| > \delta) \le C\delta^{-k}E[u_t^{2k}]$ holds by Lemma D.6, it is sufficient to show that $P(n^{1/2}|(1+W_n)^{-1} - 1| > \delta) \le C\delta^{-k}E[u_t^{2k}]$ due to Lemma D.4. To prove the last inequality we use $P(n^{1/2}|W_n| > \delta) \le C\delta^{-k}E[u_t^{2k}]$ (which holds by Lemma D.6) and part 5 in Lemma D.4.

The proof of items 2–3 follows from the same arguments as before; therefore, the details are omitted. Finally, the proof of item 4 follows by the results of items 2 and 3, the definition of $\hat{\eta}_n(\rho, h)$ and $\eta(\rho, h)$, and Bonferroni's inequality. Items 5-8 are implied by items 1-4, Bonferroni's inequality, and Lemma D.6. ∎

Lemma D.8. *Let $\{Z_t : 1 \le t \le n\}$ be a martingale difference sequence. Then, for any $k \ge 2$, we have*

$$E\left[\left|n^{-1/2}\sum_{t=1}^{n} Z_t\right|^k\right] \le d_k \beta_{n,k} ,$$

where $\beta_{n,k} = n^{-1}\sum_{t=1}^{n} E[|Z_t|^k]$ and $d_k = (8(k-1)\max\{1, 2^{k-3}\})^k$.

Proof. See Dharmadhikari et al. (1968), where this lemma is the main theorem. ∎

D.2 Proof of the Lemma B.5

Proof. For item 1, for any fixed $\epsilon > 0$, there exist $N_0 = N_0(\epsilon)$ such that the next inclusion

$$\{|\hat{\rho}_n| > 1 - a/2\} \subseteq \{n^{1/2}|\hat{\rho}_n - \rho| > n^{1/2-\epsilon}\} \cup \{|\rho| > 1 - a\} ,$$

holds for any $n \ge N_0$. Since $|\rho| \le 1 - a$, we conclude

$$P\left(|\hat{\rho}_n| > 1 - a/2\right) \le P\left(n^{1/2}|\hat{\rho}_n - \rho| > n^{1/2-\epsilon}\right) \le C_1 n^{-1-\epsilon}E[u_t^{2k_1}] ,$$

for $k_1 \geq 2(1+\epsilon)/(1-2\epsilon)$, where the last inequality follows from Lemma D.7 in Appendix D.1 by taking $\delta = n^{1/2-\epsilon}$. This proves item 1 since $C_1 E[u_t^{2k_1}] \leq C = C(c_u, k_1, C_1)$.

For item 2, we use the definition of $\tilde{u}_t$ in (13), $\hat{u}_t = y_t - \hat{\rho}_n y_{t-1}$, where $\hat{\rho}_n$ is as in (12), and the model (1) to obtain

$$n^{-1}\sum_{t=1}^{n} \tilde{u}_t^r = n^{-1}\sum_{t=1}^{n} \left(\hat{u}_t - \bar{\hat{u}}\right)^r = n^{-1}\sum_{t=1}^{n} \left(u_t + (\rho - \hat{\rho}_n) y_{t-1} - \bar{\hat{u}}\right)^r ,$$

where $\bar{\hat{u}} = n^{-1}\sum_{t=1}^{n} u_t + (\rho - \hat{\rho}_n) n^{-1}\sum_{t=1}^{n} y_{t-1}$. Using the multinomial formula and the previous expression, we have that $n^{-1}\sum_{t=1}^{n} \tilde{u}_t^r$ is equal to

$$n^{-1}\sum_{t=1}^{n} \sum_{r_1+r_2+r_3=r} \binom{r}{r_1, r_2, r_3} u_t^{r_1} \left((\rho-\hat{\rho}_n) y_{t-1}\right)^{r_2} (-\bar{\hat{u}})^{r_3} = I_1 + I_2 + E[u_t^r] ,$$

where

$$\begin{aligned} I_1 &= \sum_{r_1+r_2+r_3=r} \binom{r}{r_1, r_2, r_3} (\rho-\hat{\rho}_n)^{r_2} (-\bar{\hat{u}})^{r_3} \left\{ n^{-1}\sum_{t=1}^{n} u_t^{r_1} y_{t-1}^{r_2} - m_{r_1,r_2} \right\} \\ I_2 &= \sum_{r_1+r_2+r_3=r} \binom{r}{r_1, r_2, r_3} (\rho-\hat{\rho}_n)^{r_2} (-\bar{\hat{u}})^{r_3} m_{r_1,r_2} - E[u_t^r] I\{r_1 = r\} \\ m_{r_1,r_2} &= E\left[u_t^{r_1} \left(\sum_{j\geq 1} \rho^{t-1-j} u_j \right)^{r_2} \right] . \end{aligned}$$

Note that Lemmas D.4, D.6, and D.7 in Appendix D.1 imply that $P(|\rho - \hat{\rho}_n| > n^{-\epsilon}) \leq Cn^{-1-\epsilon}$, $P(|\bar{\hat{u}}| > n^{-\epsilon}) \leq Cn^{-1-\epsilon}$, and $P(|n^{-1}\sum_{t=1}^{n} u_t^{r_1} y_{t-1}^{r_2} - m_{r_1,r_2}| > n^{-\epsilon}) \leq Cn^{-1-\epsilon}$ for some constant C. Therefore, Lemma D.4 implies $P(|I_j| > n^{-\epsilon}) \leq Cn^{-1-\epsilon}$ for $j = 1, 2$. This implies that, for a fixed r, we have $P(|n^{-1}\sum_{t=1}^{n} \tilde{u}_t^r - E[u_t^r]| > n^{-\epsilon}) \leq Cn^{-1-\epsilon}$.

For item 3, we note that item 2 implies $P(|n^{-1}\sum_{t=1}^{n} \tilde{u}_t^2 - E[u_t^2]| > E[u_t^2]/2) \leq Cn^{-1-\epsilon}$. Therefore, we conclude item 3 by taking $\tilde{C}_\sigma = E[u_t^2]/2$. For item 4, we note that item 2 implies $P(|n^{-1}\sum_{t=1}^{n} \tilde{u}_t^{4k} - E[u_t^{4k}]| > 1) \leq Cn^{-1-\epsilon}$. Then, we conclude item 4 by taking $M = E[u_t^{4k}] + 1$. ■

D.3 Proof of Theorem B.2

Proof. The proof of this theorem has two steps.

Step 1: The sample average $(n-h)^{-1/2}\sum_{t=1}^{n-h} X_t$ has a valid Edgeworth expansion up to an error $o\left(n^{-3/2}\right)$ due to the results in Götze and Hipp (1983). Assumption B.2 and the definition of X_t in (B.4) guarantees that we can use Theorem 1.2 in Götze and Hipp (1994), and this in turn implies that we can use the results in Götze and Hipp (1983) (Theorem 2.8 and Remark 2.12). We obtain an approximation error of $o\left(n^{-3/2}\right)$ since $E[|X_t|^6] < +\infty$, which holds due to Assumption B.2.(iii).

Step 2: The proof of Theorem 2 in Bhattacharya and Ghosh (1978) and the Edgeworth expansion for the sample average $(n-h)^{-1/2}\sum_{t=1}^{n-h} X_t$ guarantee the existence of Edgeworth expansion for the distribution $\tilde{J}_n$ defined in (B.6). Furthermore, the function $q_j(x, h, P, \rho)$ for $j = 1, 2, 3$ is a polynomial in x with coefficients that are polynomials of the moments of X_t (up to order $j+2$) since the sequence X_t is strictly stationary ($|\rho| < 1$). In particular, the coefficients of the polynomial $q_j(x, h, P, \rho)$ for $j = 1, 2$ are polynomials of moments of P (up to order 12) and ρ since the moments of X_t can be computed using the moments of u_t and ρ. Moreover, $q_j(x, h, P, \rho) = (-1)^j q_j(-x, h, P, \rho)$ since the sequence X_t is strictly stationary. ■

D.4 Proof of Theorem B.3

Proof. The proof has two steps.

Step 1: Define the events $E_{n,1} = \{|\hat{\rho}_n| \leq 1 - a/2\}$, $E_{n,2} = \{n^{-1}\sum_{t=1}^{n} \tilde{u}_t^2 \geq \tilde{C}_\sigma\}$, and $E_{n,3} = \{n^{-1}\sum_{t=1}^{n} \tilde{u}_t^{4k} < M\}$, where $\tilde{C}_\sigma$ and M are as in Lemma B.5. Define $E_n = E_{n,1} \cap E_{n,2} \cap E_{n,3}$. By Lemma B.5 and Assumption 5.1 it follows that $P(E_n^c) \leq C_2 n^{-1-\epsilon}$ for some constant C_2 that depends on the moments of u_t. Since $k > 8$, it follows that, conditional on E_n, the empirical distribution $\hat{P}_n$ verifies part (i) and (iii) of Assumption B.2. It is important to mention that Götze and Hipp (1994) use part (ii) of Assumption B.2 to guarantee the dependent-data version of the Cramer condition that appears in Götze and Hipp (1983); see Lemma 2.3 in Götze and Hipp (1994).

Step 2: Condition (iii) in Lemma 2.3 in Götze and Hipp (1994) holds for the bootstrap sequence $X_{b,t}^*$ since it holds for the original sequence X_t, otherwise the function F in (B.3) verifies equation (8) in Götze and Hipp (1994). Therefore, the dependent-data version of the Cramer condition holds for the bootstrap sequence $X_{b,t}^*$. The results in Götze and Hipp (1994) implied that Edgeworth expansion exists for the sample average. Then, conditional on the event E_n we can repeat the arguments presented in the proof of Theorem B.2. ■

E Additional Tables

This appendix presents the additional results of the simulations.

ρ	h	RB	RB_{per-t}	RB_{hc3}	WB	WB_{per-t}	GB_{LR}	AA	AA_{hc2}	AA_{hc3}
Design 1: Gaussian i.i.d. shocks										
0.95	1	0.35	0.35	0.35	0.35	0.35	0.13	0.33	0.34	0.35
	6	0.83	0.81	0.83	0.86	0.84	0.54	0.71	0.73	0.74
	12	1.07	1.03	1.07	1.12	1.09	0.81	0.89	0.91	0.93
	18	1.15	1.11	1.15	1.21	1.17	0.97	0.98	1.00	1.03
1.00	1	0.35	0.35	0.35	0.35	0.35	0.07	0.33	0.34	0.35
	6	0.97	0.93	0.97	1.00	0.96	0.42	0.80	0.82	0.84
	12	1.51	1.41	1.51	1.57	1.48	0.77	1.12	1.15	1.17
	18	2.01	1.83	2.01	2.09	1.92	1.07	1.36	1.39	1.42
Design 2: Gaussian GARCH shocks										
0.95	1	0.44	0.43	0.44	0.46	0.45	0.13	0.41	0.43	0.44
	6	0.93	0.91	0.94	1.00	0.98	0.54	0.80	0.82	0.84
	12	1.10	1.06	1.11	1.19	1.15	0.79	0.91	0.94	0.97
	18	1.13	1.09	1.13	1.22	1.18	0.94	0.95	0.98	1.01
1.00	1	0.44	0.43	0.44	0.45	0.45	0.07	0.41	0.42	0.44
	6	1.10	1.06	1.11	1.17	1.13	0.42	0.91	0.93	0.96
	12	1.60	1.50	1.61	1.73	1.63	0.77	1.18	1.22	1.25
	18	2.04	1.86	2.05	2.21	2.04	1.06	1.37	1.41	1.45
Design 3: t-student i.i.d. shocks										
0.95	1	0.33	0.33	0.34	0.33	0.33	0.13	0.31	0.32	0.33
	6	0.81	0.79	0.82	0.84	0.82	0.54	0.68	0.71	0.73
	12	1.05	1.02	1.06	1.10	1.07	0.80	0.86	0.89	0.93
	18	1.14	1.10	1.15	1.19	1.16	0.97	0.94	0.98	1.02
1.00	1	0.33	0.33	0.34	0.34	0.33	0.07	0.31	0.32	0.33
	6	0.94	0.91	0.95	0.97	0.94	0.42	0.77	0.79	0.82
	12	1.49	1.39	1.50	1.54	1.45	0.77	1.07	1.11	1.16
	18	1.96	1.79	1.98	2.03	1.87	1.08	1.30	1.35	1.41
Design 4: mix-gaussian GARCH shocks										
0.95	1	0.46	0.45	0.46	0.46	0.45	0.13	0.42	0.43	0.44
	6	0.89	0.87	0.90	0.96	0.94	0.55	0.77	0.79	0.82
	12	1.01	0.98	1.02	1.11	1.07	0.78	0.86	0.88	0.91
	18	1.02	1.00	1.03	1.12	1.08	0.88	0.89	0.91	0.94
1.00	1	0.46	0.45	0.46	0.46	0.46	0.08	0.42	0.43	0.44
	6	1.06	1.01	1.07	1.12	1.09	0.45	0.87	0.90	0.92
	12	1.50	1.40	1.51	1.62	1.53	0.79	1.11	1.14	1.18
	18	1.88	1.71	1.89	2.06	1.90	1.08	1.28	1.31	1.35

Table E.1: Median length of confidence intervals for $\beta(\rho, h)$ with a nominal level of 90% and $n = 95$. 5,000 simulations and 1,000 bootstrap iterations.

ρ	h	RB	RB_{per-t}	RB_{hc3}	WB	WB_{per-t}	GB_{LR}	AA	AA_{hc2}	AA_{hc3}
Design 1: Gaussian i.i.d. shocks										
0.95	1	80.32	77.86	80.36	79.80	77.42	33.00	80.14	79.98	79.98
	6	92.02	88.44	92.00	91.92	89.02	85.68	91.88	91.92	91.84
	12	91.82	90.14	91.74	91.94	90.56	88.80	91.24	91.28	91.24
	18	91.90	90.36	91.98	91.72	90.22	89.30	91.38	91.40	91.42
1.00	1	79.04	76.14	79.12	79.50	76.40	15.28	79.32	79.34	79.24
	6	92.06	87.66	92.22	91.82	88.64	75.14	91.88	92.02	91.94
	12	92.56	89.86	92.70	92.78	90.12	84.96	93.06	93.00	93.06
	18	92.06	90.72	92.08	92.04	90.50	87.54	92.08	92.14	92.18
Design 2: Gaussian GARCH shocks										
0.95	1	84.48	82.16	84.62	84.64	82.88	41.62	84.26	84.46	84.62
	6	92.30	89.74	92.36	92.44	89.18	87.80	92.22	92.20	92.20
	12	91.94	90.44	91.90	92.14	90.28	89.90	91.72	91.70	91.64
	18	91.76	90.70	91.76	91.66	90.52	90.30	91.50	91.56	91.58
1.00	1	84.30	81.48	84.46	84.36	82.22	16.94	84.34	84.42	84.38
	6	92.52	89.44	92.64	92.70	89.50	75.92	92.62	92.66	92.72
	12	92.78	90.28	92.86	92.40	90.06	85.76	93.02	92.86	92.94
	18	92.14	90.78	92.12	92.20	90.32	88.18	92.28	92.32	92.28
Design 3: t-student i.i.d. shocks										
0.95	1	78.36	74.78	78.54	77.46	74.68	32.14	78.34	78.02	78.32
	6	91.80	88.08	91.80	92.08	88.34	86.58	91.72	91.60	91.54
	12	91.84	90.40	92.00	91.74	90.08	88.90	92.08	92.06	92.06
	18	91.74	89.86	91.50	91.70	89.82	89.52	91.32	91.32	91.44
1.00	1	77.48	73.44	77.98	76.18	73.88	14.62	77.30	77.50	77.68
	6	92.00	88.28	91.78	92.06	88.08	73.00	92.02	91.78	91.72
	12	92.90	89.62	92.80	92.80	89.62	83.60	92.82	92.82	92.66
	18	92.16	90.26	92.20	92.36	90.60	86.60	92.30	92.42	92.40
Design 4: mix-gaussian GARCH shocks										
0.95	1	92.92	86.04	93.40	93.76	91.70	47.00	92.64	92.84	93.18
	6	93.68	91.62	93.72	93.64	92.00	90.54	93.72	93.78	93.80
	12	92.48	91.64	92.38	92.54	91.70	90.92	92.60	92.60	92.52
	18	92.00	91.26	91.82	91.78	90.76	90.86	91.70	91.62	91.74
1.00	1	93.14	84.72	93.72	93.54	91.42	20.76	92.78	93.18	93.44
	6	94.20	91.86	94.20	93.96	92.14	80.00	94.52	94.58	94.50
	12	92.30	91.48	92.28	92.40	91.58	87.54	92.88	92.88	92.80
	18	92.26	92.12	92.26	92.16	91.64	89.22	92.26	92.30	92.20

Table E.2: Coverage probability (in %) of (size-adjusted) confidence intervals for $\beta(\rho, h) \times 0.9$ with a nominal level of 90% and $n = 95$. 5,000 simulations and 1,000 bootstrap iterations.

ρ	h	RB	RB_{per-t}	RB_{hc3}	WB	WB_{per-t}	GB_{LR}	AA	AA_{hc2}	AA_{hc3}
Design 1: Gaussian i.i.d. shocks										
0.95	18	89.44	87.52	89.46	89.66	88.74	90.02	87.20	87.40	87.68
	40	90.30	87.92	90.28	91.06	88.98	90.02	88.74	89.08	89.48
1.00	18	88.62	88.78	88.56	89.14	89.58	90.66	82.16	82.70	83.02
	40	86.56	84.80	86.52	86.64	85.78	90.66	78.56	78.88	79.16
Design 2: Gaussian GARCH shocks										
0.95	18	86.64	86.44	86.70	88.58	88.20	81.98	84.28	84.62	85.22
	40	89.18	86.84	89.24	90.36	87.90	81.98	87.46	87.88	88.32
1.00	18	87.10	87.50	87.22	89.26	89.72	87.56	80.82	81.28	81.94
	40	84.10	82.56	84.10	85.46	85.16	87.56	76.54	76.88	77.30
Design 3: t-student i.i.d. shocks										
0.95	18	89.06	88.04	89.06	89.82	88.78	89.96	86.04	86.68	87.38
	40	89.84	87.26	89.94	90.66	88.70	89.96	88.04	88.72	89.46
1.00	18	89.36	88.96	89.54	89.98	89.42	90.92	82.36	83.00	83.64
	40	85.96	84.90	85.82	86.52	86.24	90.92	77.90	78.54	79.28
Design 4: mix-gaussian GARCH shocks										
0.95	18	83.00	86.32	83.10	84.50	87.50	83.20	81.22	81.54	82.00
	40	87.20	86.18	87.24	88.64	87.56	83.20	86.00	86.50	87.02
1.00	18	84.06	88.68	84.22	86.04	90.42	86.84	76.98	77.34	77.80
	40	81.24	84.06	81.20	82.76	85.90	86.84	73.44	73.98	74.36

Table E.3: Coverage probability (in %) of confidence intervals for $\beta(\rho, h)$ with a nominal level of 90% and $n = 240$. 5,000 simulations and 1,000 bootstrap iterations.

E.1 Monte-Carlo Simulations for a VAR model

We consider a Bivariate VAR(4) model as in Montiel Olea and Plagborg-Møller (2021b):

$$y_{1,t} = \rho y_{1,t-1} + u_{1,t}, \quad \left(1 - \frac{1}{2}L\right)^4 y_{2,t} = \frac{1}{2} y_{1,t-1} + u_{2,t}, \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix} \overset{i.i.d.}{\sim} N\left(0, \begin{pmatrix} 1 & 0.3 \\ 0.3 & 1 \end{pmatrix}\right) .$$

We construct confidence intervals for the reduced-form impulse response of $y_{2,t}$ with respect to the shock $u_{1,t}$, that is, we set $\nu = (1, 0)$ and $i = 2$ according to the notation of Section 7. The confidence intervals that we use are listed below:

1. **RB:** confidence interval as in (26) based on the LP-residual bootstrap in Section 7.
2. $\mathbf{RB}_{per-t}$**:** equal-tailed percentile-t confidence interval based on the LP-residual bootstrap described in Section 7 and similar to the one described in Remark 4.1.
3. **WB:** confidence interval as in (26) but using $c_n^{wb,*}(h, 1-\alpha)$ instead of $c_n^*(h, 1-\alpha)$, where

$c_n^{wb,*}(h, 1-\alpha)$ is based on the LP-wild bootstrap, that is, the LP-residual bootstrap described in Section 7 with a different step 2 as in Remark 3.2.

4. **WB$_{per-t}$:** equal-tailed percentile-t confidence interval based on the LP-wild bootstrap.
5. **AA:** standard confidence interval as in Montiel Olea and Plagborg-Møller (2021a).

Table E.4 presents the results of the simulation in terms of coverage probability and median length. It reports qualitatively similar results as the ones presented in Section 6. The confidence interval **RB** performs better than **AA** and is similar to **WB** for all horizons.

	Coverage					Median length				
h	RB	RB$_{per-t}$	WB	WB$_{per-t}$	AA	RB	RB$_{per-t}$	WB	WB$_{per-t}$	AA
					$\rho = 0.00$					
1	89.700	89.660	90.160	90.040	89.280	0.232	0.232	0.235	0.235	0.228
6	90.440	89.180	91.200	90.060	88.640	1.450	1.440	1.487	1.476	1.376
12	89.320	88.780	90.320	89.860	87.480	1.572	1.565	1.614	1.605	1.490
36	89.940	89.500	90.760	90.560	88.160	1.657	1.653	1.705	1.701	1.577
60	89.300	89.220	90.220	90.200	87.200	1.771	1.767	1.826	1.824	1.671
					$\rho = 0.50$					
1	89.820	89.800	90.160	90.060	89.380	0.232	0.232	0.235	0.235	0.228
6	90.340	88.920	91.040	89.860	88.040	1.705	1.680	1.744	1.720	1.590
12	89.080	87.820	90.040	88.640	86.660	1.968	1.942	2.017	1.997	1.830
36	89.880	89.460	90.840	90.320	88.160	2.036	2.023	2.095	2.082	1.934
60	89.060	89.060	90.120	90.220	86.500	2.178	2.169	2.251	2.238	2.048
					$\rho = 0.95$					
1	89.740	89.740	90.120	90.080	89.320	0.232	0.232	0.236	0.236	0.229
6	89.640	88.880	90.360	89.500	85.960	2.346	2.243	2.393	2.289	2.084
12	87.960	87.860	88.500	88.660	81.500	4.824	4.374	4.935	4.482	3.786
36	82.740	80.660	83.740	82.020	77.380	6.207	5.640	6.421	5.836	5.040
60	88.300	87.300	89.380	88.500	82.460	6.003	5.644	6.197	5.826	5.185
					$\rho = 1.00$					
1	89.900	89.780	90.200	90.100	89.420	0.233	0.233	0.236	0.236	0.229
6	87.420	88.520	87.940	89.180	82.060	2.483	2.333	2.534	2.381	2.152
12	82.920	87.020	83.960	87.860	71.540	5.950	5.142	6.067	5.266	4.306
36	70.060	73.460	70.800	74.580	46.320	14.159	10.863	14.556	11.188	7.528
60	53.060	57.300	54.340	58.560	28.860	13.932	10.433	14.365	10.767	7.904

Table E.4: Coverage probability (in %) and median length of confidence intervals for $\beta(\rho, h)$ with a nominal level of 90% and $n = 240$. 5,000 simulations and 2,000 bootstrap iterations.

References

Andrews, D. W., X. Cheng, and P. Guggenberger (2020): "Generic results for establishing the asymptotic size of confidence sets and tests," *Journal of Econometrics*, 218, 496–531.

Bhattacharya, R. N. and J. K. Ghosh (1978): "On the validity of the formal Edgeworth expansion," *Annals of Statistics*, 6, 434–451.

Davidson, J. (1994): *Stochastic Limit Theory: An Introduction for Econometricians*, Advanced Texts in Econometrics, Oxford University Press.

Dharmadhikari, S. W., V. Fabian, and K. Jogdeo (1968): "Bounds on the Moments of Martingales," *The Annals of Mathematical Statistics*, 39, 1719 – 1723.

Götze, F. and C. Hipp (1983): " Asymptotic expansions for sums of weakly dependent random vectors," *Z. Wahrscheinlichkeitstheorie verw Gebiete*, 64, 211–239.

——— (1994): "Asymptotic distribution of statistics in time series," *Annals of Statistics*, 2062–2088.

Montiel Olea, J. L. and M. Plagborg-Møller (2021a): "Local Projection Inference is Simpler and More Robust Than You Think," *Econometrica*, 89, 1789–1823.

——— (2021b): "Supplement to 'Local Projection Inference is Simpler and More Robust Than You Think'," *Econometrica*, 89, 1789–1823.

Phillips, P. C. B. (1987): "Towards a unified asymptotic theory for autoregression," *Biometrika*, 74, 535–547.

White, H. (2000): *Asymptotic Theory for Econometricians*, Academic Press.

Xu, K.-L. (2023): "Local Projection Based Inference under General Conditions," Tech. rep., No. 2023-001, Center for Applied Economics and Policy Research, Department of Economics, Indiana University Bloomington.