Appendix D  Coverage of Delta Method AR Inference

Here we compute the under-coverage of delta method AR(1) inference (both lag-augmented and non-augmented) for horizons $h \propto T^{1/2}$. This discussion is merely intended to flesh out the insights of Mikusheva (2012, Section 4.3).

Consider a $\sqrt{T}$-consistent asymptotically normal estimator $\hat{\rho}$ of the AR(1) parameter $\rho$:

$$T^{1/2}(\hat{\rho} - \rho) \xrightarrow{d} N(0, \omega^2).$$

For example, for plain AR(1) inference we would have $\omega^2 = 1 - \rho^2$, and for lag-augmented AR inference we would have $\omega^2 = 1$ (see Appendix B.2.1). Suppose we have available a consistent estimator $\hat{\omega} \xrightarrow{p} \omega > 0$. Let $\hat{\beta}_{AR}(h) \equiv \hat{\rho}^h$ be the plug-in AR estimator of the impulse response $\beta(\rho, h) \equiv \rho^h$ at horizon $h$. The usual delta method standard error for $\hat{\beta}_{AR}(h)$ equals

$$\text{se}_{AR}(h) \equiv T^{-1/2}h|\hat{\rho}|^{h-1}\hat{\omega}.$$

Consider specifically the horizon $h = h_T = \kappa T^{1/2}$, where $\kappa \in (0, \infty)$ is a fixed constant. Let $\Phi(\cdot)$ denote the standard normal CDF, and let $Z$ denote a standard normal random variable. For any $\rho \neq 0$, the coverage probability of the usual $1 - \alpha$ confidence interval equals

$$P_\rho \left( \frac{|\hat{\beta}_{AR}(h_T) - \beta(\rho, h_T)|}{\text{se}_{AR}(h_T)} \leq \Phi^{-1}(1 - \alpha/2) \right).$$
\[ \begin{align*}
&= P_\rho \left( \left| \frac{\hat{\rho}}{\omega \kappa} \left( 1 - e^{-\kappa T^{1/2} (\log \hat{\rho} - \log \rho)} \right) \right| \leq \Phi^{-1}(1 - \alpha/2) \right) \\
&\rightarrow P \left( \left| \frac{\rho}{\omega \kappa} \left( 1 - e^{-\kappa \rho^{-1} \omega Z} \right) \right| \leq \Phi^{-1}(1 - \alpha/2) \right) \\
&= P \left( \left| \frac{\rho}{\omega \kappa} \left( 1 - e^{\kappa |\rho|^{-1} \omega Z} \right) \right| \leq \Phi^{-1}(1 - \alpha/2) \right). \\
\end{align*} \]  

(S1)

The last two steps use the delta method and symmetry of the normal distribution:

\[ T^{1/2} (\log \hat{\rho} - \log \rho) \xrightarrow{d} \frac{1}{\rho} \mathcal{N}(0, \omega^2) \xrightarrow{d} \frac{\omega}{|\rho|} Z. \]

Note that the final expression (S1) for the asymptotic coverage probability is a function only of \( \alpha \) and the quantity \( \zeta \equiv \frac{|\rho|}{\omega \kappa} > 0 \). Denote the asymptotic coverage probability by \( \text{AsyCov}(\alpha, \zeta) \). We can further simplify this as follows:

\[ \text{AsyCov}(\alpha, \zeta) \equiv P \left( \left| \zeta \log \left( 1 - e^{Z/\zeta} \right) \right| \leq \Phi^{-1}(1 - \alpha/2) \right) \\
= P \left( \zeta \log \left( 1 - \frac{\Phi^{-1}(1 - \alpha/2)}{\zeta} \right) \right) \leq Z \leq \zeta \log \left( 1 + \frac{\Phi^{-1}(1 - \alpha/2)}{\zeta} \right) \\
= \Phi \left( \zeta \log \left( 1 + \frac{\Phi^{-1}(1 - \alpha/2)}{\zeta} \right) \right) - \Phi \left( \zeta \log \left( 1 - \frac{\Phi^{-1}(1 - \alpha/2)}{\zeta} \right) \right). \]

Here we use the convention that \( \log(x) = -\infty \) if \( x \leq 0 \). Observe that

\[ \lim_{\zeta \downarrow 0} \text{AsyCov}(\alpha, \zeta) = \Phi \left( \lim_{\zeta \downarrow 0} \zeta \log \left( 1 + \frac{\Phi^{-1}(1 - \alpha/2)}{\zeta} \right) \right) = \Phi(0) = 1/2, \]

\[ \lim_{\zeta \to \infty} \text{AsyCov}(\alpha, \zeta) = \lim_{\zeta \to \infty} \left\{ \Phi \left( \frac{\zeta \Phi^{-1}(1 - \alpha/2)}{\zeta} \right) + \Phi \left( -\zeta \frac{\Phi^{-1}(1 - \alpha/2)}{\zeta} \right) \right\} = 1 - \alpha. \]

Thus, for any significance level \( \alpha < 1/2 \), the AR confidence interval under-covers whenever \( \zeta \) is sufficiently small (i.e., when \( |\rho|/\omega \) is small or \( \kappa = h_T/T^{1/2} \) is large). The coverage is close to the nominal level when \( \zeta \) is large (i.e., \( |\rho|/\omega \) is not too small and \( \kappa \) is small).

Figure S1 plots the coverage shortfall \( (1 - \alpha) - \text{AsyCov}(\alpha, \zeta) \) as a function of \( \zeta = \frac{|\rho|}{\omega \kappa} \), for three conventional significance levels \( \alpha \). For \( \alpha = 0.05 \) and \( 0.10 \), the AR confidence interval under-covers asymptotically regardless of the value of \( \zeta \). For \( \alpha = 0.32 \), the conventional interval is conservative at intermediate values of \( \zeta \) (but under-covers for small \( \zeta \), of course).
Figure S1: Asymptotic coverage shortfall \((1 - \alpha) - \text{AsyCov}(\alpha, \zeta)\) of conventional AR confidence interval at horizon \(h_T = \kappa T^{1/2}\). Horizontal axis: \(\zeta = \frac{|\rho|}{\omega \kappa}\). The three curves correspond to \(\alpha = 0.05, 0.10, 0.32\).

Appendix E Additional Proofs

E.1 Notation

Geometric series of the form \(\sum_{\ell=0}^{h-1}(\rho_t^*(A, \epsilon))^2\) will show up repeatedly in the proofs below. Observe that, for any \(A \in \mathcal{A}(0, C, \epsilon)\) and \(h \in \mathbb{N}\),

\[
1 \leq \sum_{\ell=0}^{h-1} \rho_t^*(A, \epsilon)^{2\ell} \leq \min \left\{ \frac{1}{1 - \rho_t^*(A, \epsilon)^2}, h \right\} \leq \min \left\{ \frac{1}{1 - \rho_t^*(A, \epsilon)}, h \right\} = g(\rho_t^*(A, \epsilon), h)^2.
\]

Recall also the definition of the lag-augmented LP residuals \(\hat{\xi}_{1,t}(h) = y_{1,t+h} - \hat{\beta}_1(h)'y_t - \hat{\gamma}_1(h)'X_t\). We can write

\[
\hat{\xi}_{1,t}(h) - \xi_{1,t}(\rho, h) = (y_{1,t+h} - \hat{\beta}_1(h)'y_t - \hat{\gamma}_1(h)'X_t) - (y_{1,t+h} - \beta_1(A, h)'u_t - \eta_1(A, h)'X_t)
\]
\[ \frac{\sum_{t=1}^{T-hT} \hat{\xi}_{1,t}(hT)^2(\hat{\nu}(hT)'\hat{u}_t(hT)) \leq \sum_{t=1}^{T-hT} \xi_{1,t}(hT)^2(\hat{\nu}'u_t)^2}{(T-hT)v(A_T, h_T, \hat{\nu})^2} \mathcal{P}_A \xrightarrow{p} 0, \]

where we have defined \( \hat{\nu} \equiv \Sigma^{-1}\nu \). Algebra shows that

\[ \left[ \frac{\sum_{t=1}^{T-hT} \hat{\xi}_{1,t}(hT)^2(\hat{\nu}(hT)'\hat{u}_t(hT)) - \xi_{1,t}(A_T, h_T)^2(\hat{\nu}'u_t)^2}{(T-hT)v(A_T, h_T, \hat{\nu})^2} \right] \]

\[ \leq \frac{1}{(T-hT)v(A_T, h_T, \hat{\nu})^2} \sum_{t=1}^{T-hT} \left| \hat{\xi}_{1,t}(hT)(\hat{\nu}(hT)'\hat{u}_t(hT)) \right| - \xi_{1,t}(A_T, h_T)(\hat{\nu}'u_t) \]

\[ \times \left[ \hat{\xi}_{1,t}(hT)(\hat{\nu}(hT)'\hat{u}_t(hT)) \right] - \xi_{1,t}(A_T, h_T)(\hat{\nu}'u_t) + 2\xi_{1,t}(A_T, h_T)(\hat{\nu}'u_t) \]

(as \( (a + b)(a - b) = a^2 - b^2 \))

\[ \leq \left( \frac{\sum_{t=1}^{T-hT} \hat{\xi}_{1,t}(hT)(\hat{\nu}(hT)'\hat{u}_t(hT)) - \xi_{1,t}(A_T, h_T)(\hat{\nu}'u_t)^2}{(T-hT)v(A_T, h_T, \hat{\nu})^2} \right)^{1/2} \]

\[ \times \left( \frac{\sum_{t=1}^{T-hT} \hat{\xi}_{1,t}(hT)(\hat{\nu}(hT)'\hat{u}_t(hT)) - \xi_{1,t}(A_T, h_T)(\hat{\nu}'u_t) + 2\xi_{1,t}(A_T, h_T)(\hat{\nu}'u_t)^2}{(T-hT)v(A_T, h_T, \hat{\nu})^2} \right)^{1/2} \]

Consider the expression in the last line above. By Loève’s inequality (Davidson, 1994, Thm. 9.28), this expression is bounded above by

\[ \left( \frac{\sum_{t=1}^{T-hT} \hat{\xi}_{1,t}(hT)(\hat{\nu}(hT)'\hat{u}_t(hT)) - \xi_{1,t}(A_T, h_T)(\hat{\nu}'u_t)^2}{(T-hT)v(A_T, h_T, \hat{\nu})^2} + 8\sum_{t=1}^{T-hT} \xi_{1,t}(A_T, h_T)^2(\hat{\nu}'u_t)^2 \right)^{1/2} \]
The last fraction above is bounded in probability by Lemma A.6. Thus, it is sufficient to show that
\[
\sum_{t=1}^{T-h_T} \left[ \hat{\xi}_{1,t}(h_T)(\hat{\nu}(h_T)\hat{u}_t(h_T)) - \xi_{1,t}(A_T, h_T)(\hat{\nu}'u_t) \right]^2
\]
\[
\frac{(T - h_T)v(A_T, h_T, \hat{\nu})^2}
\]
converges in probability to zero. To that end, decompose
\[
\hat{\xi}_{1,t}(h_T)(\hat{\nu}(h_T)\hat{u}_t(h_T)) - \xi_{1,t}(A_T, h_T)(\hat{\nu}'u_t)
\]
\[
= (\hat{\xi}_{1,t}(h_T) - \xi_{1,t}(A_T, h_T))(\hat{\nu}'u_t) + (\hat{\nu}(h_T)\hat{u}_t(h_T) - \hat{\nu}'u_T)\xi_{1,t}(A_T, h_T)
\]
\[
+ (\hat{\xi}_{1,t}(h_T) - \xi_{1,t}(A_T, h_T))(\hat{\nu}(h_T)\hat{u}_t(h_T) - \hat{\nu}'u_T).
\]
Hence, by another application of Loève’s inequality,
\[
\sum_{t=1}^{T-h_T} \left[ \hat{\xi}_{1,t}(h_T)(\hat{\nu}(h_T)\hat{u}_t(h_T)) - \xi_{1,t}(A_T, h_T)(\hat{\nu}'u_t) \right]^2
\]
\[
\frac{(T - h_T)v(A_T, h_T, \hat{\nu})^2}
\]
\[
\leq 3 \sum_{t=1}^{T-h_T} \left[ \hat{\xi}_{1,t}(h_T) - \xi_{1,t}(A_T, h_T) \right]^2(\hat{\nu}'u_t)^2
\]
\[
\frac{(T - h_T)v(A_T, h_T, \hat{\nu})^2}
\]
\[
+ 3 \sum_{t=1}^{T-h_T} \left[ \hat{\nu}(h_T)\hat{u}_t(h_T) - \hat{\nu}'u_T \right]^2\xi_{1,t}(A_T, h_T)^2
\]
\[
\frac{(T - h_T)v(A_T, h_T, \hat{\nu})^2}
\]
\[
+ 3 \sum_{t=1}^{T-h_T} \left[ \hat{\xi}_{1,t}(h_T) - \xi_{1,t}(A_T, h_T) \right]^2\hat{\nu}(h_T)\hat{u}_t(h_T) - \hat{\nu}'u_T)^2
\]
\[
\frac{(T - h_T)v(A_T, h_T, \hat{\nu})^2}
\]
\[
\leq 3 \left( \frac{\sum_{t=1}^{T-h_T} \left[ \hat{\xi}_{1,t}(h_T) - \xi_{1,t}(A_T, h_T) \right]^4}{(T - h_T)v(A_T, h_T, \hat{\nu})^4} \right)^{1/2} \times \left( \left\| \hat{\nu} \right\|^4 \sum_{t=1}^{T-h_T} \| u_t \|^4 \right)^{1/2}
\]
\[
+ 3 \left( \frac{\sum_{t=1}^{T-h_T} \left[ \hat{\nu}(h_T)\hat{u}_t(h_T) - \hat{\nu}'u_T \right]^4}{(T - h_T)v(A_T, h_T, \hat{\nu})^4} \right)^{1/2} \times \left( \frac{\sum_{t=1}^{T-h_T} \xi_{1,t}(A_T, h_T)^4}{(T - h_T)v(A_T, h_T, \hat{\nu})^4} \right)^{1/2}
\]
\[
+ 3 \left( \frac{\sum_{t=1}^{T-h_T} \left[ \hat{\xi}_{1,t}(h_T) - \xi_{1,t}(A_T, h_T) \right]^4}{(T - h_T)v(A_T, h_T, \hat{\nu})^4} \right)^{1/2} \times \left( \frac{\sum_{t=1}^{T-h_T} \left[ \hat{\nu}(h_T)\hat{u}_t(h_T) - \hat{\nu}'u_T \right]^4}{(T - h_T)v(A_T, h_T, \hat{\nu})^4} \right)^{1/2}
\]
\[
\equiv 3 \left[ (\hat{R}_1)^{1/2} \times (\hat{R}_2)^{1/2} \right] + 3 \left[ (\hat{R}_3)^{1/2} \times (\hat{R}_4)^{1/2} \right] + 3 \left[ (\hat{R}_1)^{1/2} \times (\hat{R}_3)^{1/2} \right].
\]
It follows from Lemma E.1 below that \( \hat{R}_1 \) tends to zero in probability. \( \hat{R}_2 \) is bounded in probability by Assumption 2(i) and a standard application of Markov’s inequality. We show below that \( \hat{R}_3 \) tends to zero in probability. Another standard application of Markov’s inequality combined with Lemma A.7 implies that \( \hat{R}_4 \) is also uniformly bounded in probability. Hence, the entire expression tends to zero in probability, as needed.
To finish the proof, we prove the claim that \( \hat{R}_3 \) tends to zero in probability. Note that
\[
\hat{R}_3 \leq \|\hat{\nu}(h_T)\|^4 \sum_{t=1}^{T-h_T} \frac{\|\hat{\nu}_t(h_T) - u_t\|^4}{T - h_T} + \|\hat{\nu}(h_T) - \tilde{\nu}\|^4 \sum_{t=1}^{T-h_T} \|u_t\|^4 / (T - h_T).
\]
Since \( \|\hat{\nu}(h_T) - \tilde{\nu}\| \leq \|\hat{\Sigma}(h_T)^{-1} - \Sigma^{-1}\| \times \|\nu\| \), it follows from Lemma A.5(ii), Lemma E.2 below, Assumption 2(i), and an application of Markov’s inequality that the above display tends to zero in probability.

**Lemma E.1** (Negligibility of estimation error in \( \hat{\xi}_{1,t}(h) \)). Let the conditions of Lemma A.2 hold. Let \( w \in \mathbb{R}^n \setminus \{0\} \). Then
\[
\sum_{t=1}^{T-h_T} [\hat{\xi}_{1,t}(h_T) - \xi_{1,t}(A_T, h_T)]^4 \frac{(T-h_T)v(A_T, h_T, w)^4}{P_{A_T}} \rightarrow 0.
\]

**Proof.** Recall equation (S2):
\[
\hat{\xi}_{1,t}(h) - \xi_{1,t}(A, h) = [\beta_1(A, h) - \hat{\beta}_1(h)]'u_t + [\eta_1(A, h) - \hat{\eta}_1(A, h)]'X_t.
\]
By Loève’s inequality (Davidson, 1994, Thm. 9.28),
\[
\sum_{t=1}^{T-h_T} [\hat{\xi}_{1,t}(h) - \xi_{1,t}(\rho, h)]^4 \frac{(T-h_T)v(A_T, h_T, w)^4}{P_{A_T}} \leq 8 \|\hat{\beta}_1(h) - \beta_1(A, h)\|^4 \sum_{t=1}^{T-h_T} \|u_t\|^4 / (T-h_T)
\]
\[
+ 8 \|G(A_T, T - h_T, \epsilon)\| \|\hat{\eta}_1(A, h_T) - \eta_1(A, h_T)\|^4 \sum_{t=1}^{T-h_T} \|G(A_T, T - h_T, \epsilon)^{-1}X_t\|^4 / (T-h_T).
\]
By Assumption 2(i) and Markov’s inequality, we have \((T-h_T)^{-1} \sum_{t=1}^{T-h_T} \|u_t\|^4 = O_{P_{A_T}}(1)\). Lemma A.3(i) then implies that the first term on the right-hand side in the above display tends to zero in probability. Similarly, the second term on the right-hand side of the above display tends to zero in probability by Lemma E.3 below, Lemma A.3(ii), and Markov’s inequality.

**Lemma E.2** (Negligibility of estimation error in \( \hat{u}_t(h) \)). Let the conditions of Lemma A.2 hold. Then
\[
\sum_{t=1}^{T-h_T} \|\hat{u}_t(h_T) - u_t\|^4 \frac{P_{A_T}}{(T-h_T)} \rightarrow 0.
\]
Proof. Since \( \hat{u}_t(h_T) - u_t = [A - \hat{A}(h_T)]X_t \), we have
\[
\frac{\sum_{t=1}^{T-h_T} \| \hat{u}_t(h_T) - u_t \|^4}{T - h_T} \leq \|G(A_T, T - h_T, \epsilon)(\hat{A}(h_T) - A_T)\|^4 \frac{\sum_{t=1}^{T-h_T} \|G(A_T, T - h_T, \epsilon)^{-1}X_t\|^4}{T - h_T}.
\]

Lemma A.3(iii) shows that the first factor after the inequality is \( o_{P_{A_T}}(1) \). Lemma E.3 below and Markov’s inequality show that the second factor is \( O_{P_{A_T}}(1) \).

**Lemma E.3** (Moment bound for \( y_{i,t}^4 \)). Let Assumption 1 and Assumption 2(i) hold. Then, for all \( T \in \mathbb{N}, A \in A(0, C, \epsilon) \), and \( i = 1, \ldots, n \),
\[
\max_{1 \leq t \leq T} E(y_{i,t}^4) \leq \frac{6C_1(E(\|u_t\|^4))^3}{\delta^2 \lambda_{\text{min}}(\Sigma)^2} \times g(\rho_1^*(A, \epsilon), T)^4
\]

where the expectations are taken with respect to the measure \( P_A \), and \( C_1 \) is the constant defined in Lemma E.4 below.

**Proof.** We have defined
\[
\xi_{i,t}(A, h) \equiv \sum_{\ell=1}^{h+1} \beta_i(A, \ell)'u_{t+\ell}.
\]

Since we have set the initial conditions \( y_0 = \ldots = y_{-p+1} = 0 \), we have
\[
y_{i,t} = \sum_{\ell=1}^{t} \beta_i(A, \ell)'u_{t} = \xi_{i,0}(A, t).
\]

Consider any \( w \in \mathbb{R}^n \) such that \( \|w\| = 1 \). Then Lemma A.7 gives
\[
\max_{1 \leq t \leq T} E(y_{i,t}^4) = \max_{1 \leq t \leq T} E[\xi_{i,0}(A, t)^4] \leq \frac{6E(\|u_0\|^4)}{\delta^2 \lambda_{\text{min}}(\Sigma)^2} \times \max_{1 \leq t \leq T} v(A, t, w)^4.
\]

Lemmas E.4 and E.5 below then imply that
\[
\max_{1 \leq t \leq T} E(y_{i,t}^4) \leq \frac{6E(\|u_0\|^4)}{\delta^2 \lambda_{\text{min}}(\Sigma)^2} \times (E(\|u_0\|^4))^2 \|w\|^4 \times \max_{1 \leq t \leq T} \left( \sum_{\ell=0}^{t-1} \|\beta_i(A, \ell)\|^2 \right)^2
\]
\[
= \frac{6(E(\|u_0\|^4))^3}{\delta^2 \lambda_{\text{min}}(\Sigma)^2} \times \left( \sum_{\ell=0}^{T-1} \|\beta_i(A, \ell)\|^2 \right)^2
\]
\[
\leq \frac{6(E(\|u_0\|^4))^3}{\delta^2 \lambda_{\text{min}}(\Sigma)^2} \times \left( \sum_{\ell=0}^{T-1} C_1 \rho_i^*(A, \epsilon)^{2\ell} \right)^2
\]
\[
\frac{6C_1^2(E(\|u_0\|^4))^3}{\delta^2\lambda_{\min}(\Sigma)^2} \times g(\rho_1^*(A, \epsilon), T)^4.
\]

\[\square\]

**Lemma E.4.** Let \(A(L)\) be a lag polynomial such that \(A = (A_1, \ldots, A_p) \in \mathcal{A}(0, C, \epsilon)\) for constants \(C > 0\) and \(0 < \epsilon < 1\). Then, for any \(i = 1, \ldots, n\), the following statements hold.

\begin{enumerate}
\item \(\|\beta_i(A, h)\| \leq C_1 \rho_i^*(A, \epsilon)^h\), where \(C_1 \equiv 1 + 2C \times \frac{1-\epsilon}{\epsilon}\).
\item \(\|\beta_i(A, h + m)\| \leq \rho_i^*(A, \epsilon)^m \times C_2 \sum_{b=0}^{p-1} \|\beta_i(A, h - b)\|\), where \(C_2 \equiv 1 + 4\tilde{C} \left(\frac{1-\epsilon}{\epsilon}\right)\), and \(\tilde{C} \equiv C (1 + C(p - 1))\).
\end{enumerate}

**Proof.** Since \(A\) is in the parameter space \(\mathcal{A}(0, C, \epsilon)\) in Definition 1,

\[\beta_i(A, h) = \rho_i \beta_i(A, h - 1) + \beta_i(B, h).\] (S3)

Thus, applying the equation above recursively,

\[\beta_i(A, h + m) = \rho_i^m \beta_i(A, h) + \sum_{\ell=1}^{m} \rho_i^{m-\ell} \beta_i(B, h + \ell).\]

We now use the above equation to prove each of the two statements of the lemma.

**Part (i).** We have

\[
\|\beta_i(A, h)\| \leq |\rho_i|^h \|\beta_i(A, 0)\| + \sum_{\ell=1}^{h} |\rho_i|^{h-\ell} \|\beta_i(B, \ell)\|
\]

\[
\leq |\rho_i|^h + \sum_{\ell=1}^{h} |\rho_i|^{h-\ell} C(1 - \epsilon)^\ell
\]

(where we have used Lemma E.7 below and \(\beta(A, 0) = I_n\))

\[
\leq \max\{|\rho_i|, 1 - \epsilon/2\}^h + \sum_{\ell=1}^{h} \max\{|\rho_i|, 1 - \epsilon/2\}^{h-\ell} C(1 - \epsilon)^\ell
\]

\[
= \rho_i^*(A, \epsilon)^h \left(1 + C \sum_{\ell=1}^{h} \left(\frac{1 - \epsilon}{\max\{|\rho_i|, 1 - \epsilon/2\}}\right)^\ell\right)
\]

\[
\leq \rho_i^*(A, \epsilon)^h \left(1 + C \sum_{\ell=1}^{\infty} \left(\frac{1 - \epsilon}{1 - \epsilon/2}\right)^\ell\right)
\]

\[
= \rho_i^*(A, \epsilon)^h \left(1 + C \left(\frac{1 - \epsilon}{\epsilon/2}\right)\right).
\]
Part (ii). To establish the remaining inequality, note that
\[
\|\beta_i(A, h + m)\|
\leq |\rho|^m \|\beta_i(A, h)\| + \sum_{\ell=1}^{m} |\rho|^{|m-\ell|} \|\beta_i(B, h + \ell)\|
\leq |\rho|^m \|\beta_i(A, h)\| + \sum_{\ell=1}^{m} |\rho|^{|m-\ell|} \left( \tilde{C} (1 - \epsilon)^{\ell} \sum_{b=0}^{p-2} \|\beta_i(B, h - b)\| \right)
\]
(by Lemma E.7(ii) below)
\[
\leq \max\{|\rho|, 1 - \epsilon/2\}^m \times \left( \|\beta_i(A, h)\| + \tilde{C} \left( \sum_{\ell=1}^{m} \left( \frac{1 - \epsilon}{\max\{|\rho|, 1 - \epsilon/2\}} \right)^{\ell} \right) \left( \sum_{b=0}^{p-2} \|\beta_i(B, h - b)\| \right) \right)
\leq \rho^* \epsilon^m \times \left( \|\beta_i(A, h)\| + 2\tilde{C} \left( \frac{1 - \epsilon}{\epsilon} \right) \left( \sum_{b=0}^{p-2} \|\beta_i(A, h - b)\| + \|\beta_i(A, h - b - 1)\| \right) \right)
\]
where we have used equation (S3))
\[
\leq \rho^* \epsilon^m \times \left( 1 + 4\tilde{C} \left( \frac{1 - \epsilon}{\epsilon} \right) \right) \sum_{b=0}^{p-1} \|\beta_i(A, h - b)\|.
\]

Lemma E.5 (Bounds on \(v(A, h, w)\)). Let Assumption 1 and Assumption 2(i) hold. Then for any \(i = 1, \ldots, n\) and for any matrix of autoregressive parameters \(A\), and any \(h \in \mathbb{N}\)
\[
\delta \times \lambda_{\min}(\Sigma) \leq \frac{1}{\|a\|^2 \sum_{h=0}^{l} \|\beta_i(A, \ell)\|^2} \leq E \left( \|u_t\|^4 \right),
\]
where \(v_i(A, h, w) \equiv E[\xi_{i,t}(A, h)^2(w'u_t)^2]\)

Proof. Algebra shows
\[
v(A, h, w)^2 = E[\xi_{i,t}(A, h)^2(w'u_t)^2]
= E \left[ (\beta_i(A, h - 1)'u_{t+1} + \ldots + \beta_i(A, 0)'u_{t+h}^2)u_t^2 \right]
= E \left[ \left( \sum_{\ell=1}^{h} \sum_{m=1}^{h} \left( \beta_i(A, h - \ell)'u_{t+\ell}u_{t+m}'\beta_i(A, h - m) \right) \right) (w'u_t)^2 \right]
\]

Assumption 1 implies that the last expression above equals
\[
\sum_{\ell=1}^{h} E \left( (\beta_i(A, h - \ell)'u_{t+\ell}^2) (w'u_t)^2 \right). \quad \text{(S4)}
\]
An application of Cauchy-Schwarz gives the upper bound

\[ v(A, h, w)^2 \leq \sum_{\ell=1}^{h} E \left( ((\beta_i(A, h - \ell)'u_{t+\ell})^4 \right)^{1/2} E \left( (w'u_t)^4 \right)^{1/2}. \]

\[ \leq \sum_{\ell=1}^{h} \|\beta_i(A, h - \ell)\|^2 E \left( \|u_{t+\ell}\|^4 \right)^{1/2} \|w\|^2 E \left( \|u_t\|^4 \right)^{1/2} \]

\[ = E \left( \|u_t\|^4 \right) \|w\|^2 \left( \sum_{\ell=0}^{h-1} \|\beta_i(A, \ell)\|^2 \right), \]

where the last line follows from stationarity.

For the lower bound, re-write expression (S4) as

\[ \|w\|^2 \sum_{\ell=1}^{h} \|\beta_i(A, h - \ell)\|^2 E \left( (\omega_1'u_{t+\ell})^2 (\omega_2'u_t)^2 \right). \]

where \(\omega_1, \omega_2\) are vectors of unit norm.

By Assumption 2(i),

\[ E \left( (\omega_1'u_{t+\ell})^2 (\omega_2'u_t)^2 \right) = E \left[ E \left( (\omega_1'u_{t+\ell})^2 \left| \{u_s\}_{s<t+\ell} \right) (\omega_2'u_t)^2 \right) \right] \]

\[ \geq \delta E[(\omega_1'u_t)^2] \]

\[ = \delta \omega_2'E[u_t u_t]' \omega_2 \]

\[ \geq \delta \lambda_{\min}(\Sigma). \]

This gives the lower bound

\[ v(A, h, w)^2 \geq \|w\|^2 \delta \lambda_{\min}(\Sigma) \sum_{\ell=0}^{h-1} \|\beta_i(A, \ell)\|^2, \]

which concludes the proof.

**Lemma E.6.** Partition the identity matrix \(I_{np}\) of dimension \(np \times np\) into \(p\) column blocks of size \(n\):

\[ I_{np} = (J'_1, \ldots, J'_p). \]

Let \(A(L)\) be a lag polynomial of order \(p\) with autoregressive coefficients \(A = (A_1, \ldots, A_p)\). Then, for any \(h, m = 0, 1, \ldots,\)

\[ \beta_i(A, h + m)' = \beta_i(A, h)' (J'_1 A^m J'_1) \]
\[
+ \sum_{j=2}^{p} \left( \sum_{k=0}^{p-j} \beta_i(A, h - 1 - k') A_{j+k} \right) \left( J_{j-1} A^{m-1} J_1' \right),
\]

where we define \( \beta_i(A, \ell) = 0 \) for \( \ell < 0 \).

**Proof.** Define \( \beta(A, \ell) \equiv (\beta_1(A, \ell), \ldots, \beta_n(A, \ell))' \). Then

\[
\beta(A, h + m) \equiv J_1 A^{h+m} J_1' = J_1 A^h A^m J_1' = J_1 A^h I_{np} I_{np} A^m J_1' = \left( \begin{array}{c} J_1 \\ J_2 \\ \vdots \\ J_p \end{array} \right) A^m J_1' = \left( J_1 A^h J_1' \right) \left( J_1 A^m J_1' \right) + \sum_{j=2}^{p} J_1 A^h J_j' A^m J_1' = \beta(A, h) \beta(A, m) + \sum_{j=2}^{p} J_1 A^h J_j' A^m J_1'.
\]

The definition of the companion matrix implies

\[
J_j A = J_{j-1}, \quad j = 2, \ldots, p,
\]

and

\[
A J_j' = J_1' A_j + J_{j+1}', \quad j = 1, \ldots, p - 1, \quad A J_p' = J_1' A_p.
\]

Therefore, for \( j \leq p \),

\[
J_1 A^h J_j' = \sum_{k=0}^{p-j} \beta(A, h - 1 - k) A_{j+k}.
\]

Thus, we have shown that

\[
\beta(A, h + m) = \beta(A, h) \beta(A, m) + \sum_{j=2}^{p} \left( \sum_{k=0}^{p-j} \beta(A, h - 1 - k) A_{j+k} \right) \left( J_{j-1} A^{m-1} J_1' \right).
\]

The lemma follows by selecting the \( i \)-th equation of the above system of equations. \( \square \)
Lemma E.7. Let $B(L)$ be a lag polynomial of order $p - 1$ satisfying $\|B^\ell\| \leq C(1 - \epsilon)^\ell$ for every $\ell = 1, 2, \ldots$. Then the following two statements hold.

i) Define the $n \times n$ matrix $\beta(B, \ell) \equiv (\beta_1(B, \ell), \ldots, \beta_n(B, \ell))'$. Then $\|\beta(B, \ell)\| \leq C(1 - \epsilon)^\ell$ for all $\ell \geq 0$.

ii) $\|\beta_i(B, h + m)\| \leq \tilde{C} \times (1 - \epsilon)^m \times \sum_{\ell=0}^{p-2} \|\beta_i(B, h - \ell)\|$ for all $h, m \geq 0$, where $\tilde{C} \equiv C (1 + C(p - 1))$.

Proof. Let the selector matrix $J_j$ be defined as in Lemma E.6. Part (i) follows immediately from the fact

$$\beta(B, \ell) = J_j B^m J_j'$$

and the assumed bound on $\|B^m\|$.

We now turn to part (ii). Lemma E.6 implies

$$\|\beta_i(B, h + m)\| \leq \|\beta_i(B, h)\| \times \|J_j B^m J_j'\|$$

$$+ \sum_{j=2}^{p-1} \left(\left(\sum_{k=0}^{p-1-j} \|\beta_i(B, h - 1 - k)\| \times \|B_{j+k}\|\right) \|J_{j-1} B^{m-1} J_1'\|\right)$$

$$\leq \|\beta_i(B, h)\| C(1 - \epsilon)^m$$

$$+ \sum_{j=2}^{p-1} \left(\left(\sum_{k=0}^{p-1-j} \|\beta_i(B, h - 1 - k)\| \times \|B_{j+k}\|\right) C(1 - \epsilon)^{m-1}\right)$$

(since $\|J_1 B^m J_1'\| \leq C(1 - \epsilon)^m$ and $\|J_{j-1} B^{m-1} J_1'\| \leq C(1 - \epsilon)^{m-1}$)

$$\leq C(1 - \epsilon)^m \left(\|\beta_i(B, h)\| + \sum_{j=2}^{p-1} \left(\left(\sum_{k=0}^{p-1-j} \|\beta_i(B, h - 1 - k)\| \times C\right)\right)\right)$$

(since $\|B_{j+k}\| = \|J_1 B^m J_{j+k}\| \leq \|B\|$)

$$\leq C(1 - \epsilon)^m \left(\|\beta_i(B, h)\| + C(p - 2) \left(\sum_{k=0}^{p-3} \|\beta_i(B, h - 1 - k)\|\right)\right)$$

$$\leq (1 - \epsilon)^m C (1 + C(p - 2)) \left(\sum_{\ell=0}^{p-2} \|\beta_i(B, h - \ell)\|\right),$$

$$\leq (1 - \epsilon)^m C (1 + C(p - 1)) \left(\sum_{\ell=0}^{p-2} \|\beta_i(B, h - \ell)\|\right).$$

The last step merely ensures that the constant is positive for all $p \geq 1$. Note that, in the case $p = 1$, the sum in the last expression is zero. □
E.3 Proof of Lemma A.3

We first prove the statements (i)–(ii), and then turn to statement (iii). For brevity, denote $G_T \equiv G(A_T, T - h_T, \epsilon)$.

Parts (i)–(ii). Recall the definition $\hat{\eta}(A, h) \equiv A' \hat{\beta}_1(h) + \hat{\gamma}_1(h)$ in equation (S5). Since the OLS coefficients $(\hat{\beta}_1(h)', \hat{\gamma}_1(h)')'$ are a non-singular linear transformation of the OLS coefficients $(\beta_1(h)', \gamma_1(h)')'$, the former vector equals the OLS coefficients in a regression of $y_{1,t+h}$ on $(u_t', X_t')'$, due to the relationship $u_t = y_t - AX_t$. By the representation

$$y_{1,t+h} = \beta_1(A, h)'u_t + \eta_1(A, h)'X_t + \xi_{1,t}(A, h)$$

in equation (19), we can therefore write

$$G_T[\hat{\eta}(A, h) - \eta(A, h)] = \left( \begin{array}{c} \frac{1}{v(A_T, h_T, w)} \left[ \hat{\beta}_1(h_T) - \beta_1(A_T, h_T) \right] \\ \frac{1}{v(A_T, h_T, w)} \sum_{t=1}^{T-h_T} G_T^{-1} u_t X_t' \end{array} \right) \left( \begin{array}{c} \frac{1}{T-h_T} \sum_{t=1}^{T-h_T} G_T^{-1} X_t u_t' \\ \frac{1}{T-h_T} \sum_{t=1}^{T-h_T} G_T^{-1} X_t' G_T^{-1} \end{array} \right)^{-1} \times \left( \begin{array}{c} \frac{1}{T-h_T} \sum_{t=1}^{T-h_T} G_T^{-1} X_t \xi_{1,t}(A_T, h_T) \\ \frac{1}{T-h_T} \sum_{t=1}^{T-h_T} G_T^{-1} \xi_{1,t}(A_T, h_T) \\
\end{array} \right)$$

(S5)

We must prove that the above display tends to zero in probability. $\hat{m}_1$ tends to zero in probability by Lemma A.1 and the fact that Lemma E.5 implies that $v(A_T, h_T, w)/v(A_T, h_T, \tilde{w})$ is uniformly bounded from below and from above for any $\tilde{w} \in \mathbb{R}^n \setminus \{0\}$. $\hat{m}_2$ also tends to zero in probability by Lemma A.4. Hence, it just remains to show that the $n(p + 1) \times n(p + 1)$ symmetric positive semidefinite matrix $\hat{M}^{-1}$ is bounded in probability. It suffices to show that $1/\lambda_{\min}(\hat{M})$ is uniformly asymptotically tight. Consider the $2 \times 2$ block partition of $\hat{M}$ in (S5). The off-diagonal blocks of $\hat{M}$ tend to zero in probability by Lemma E.8 below. Moreover, the upper left block of $\hat{M}$ tends in probability to the positive definite matrix $\Sigma$ by Lemma A.5(i) and Assumption 2. Thus, the tightness of $1/\lambda_{\min}(\hat{M})$ follows from Assumption 3, which pertains to the lower right block of $\hat{M}$. This concludes the proof of the first two statements.
Part (iii). Write
\[(T - h_T)^{1/2}[\hat{A}(h_T) - A_T]G(A_T, T - h_T, \epsilon)\]
\[= \left( \frac{1}{(T - h_T)^{1/2}} \sum_{t=1}^{T-h_T} u_tX'_tG^{-1}_T \right) \times \left( \frac{1}{T - h_T} \sum_{t=1}^{T-h_T} G^{-1}_tX_tX'_tG^{-1}_T \right)^{-1}.\]

The first factor on the right-hand side above is $O_{P_A}(1)$ by Lemma E.8 below, while the second factor is also $O_{P_A}(1)$ by the same argument as in parts (i)–(ii) above.

**Lemma E.8 (OLS denominator).** Let Assumption 1 and Assumption 2(i) hold. Let there be given a sequence $\{A_T\}$ in $A(0, C, \epsilon)$ and a sequence $\{h_T\}$ of nonnegative integers satisfying $T - h_T \to \infty$. Then for any $i, j = 1, \ldots, n$ and $r = 1, \ldots, p$,
\[\frac{\sum_{t=1}^{T-h_T} u_{i,t}y_{j,t-r}}{(T - h_T)^{1/2}g(j^*_T(A, \epsilon), T - h_T)} = O_{P_A}(1).\]

**Proof.** Write $g_{j,T} \equiv g(\rho^*_j(A, \epsilon), T - h_T)$ for brevity. Note that $\{u_{i,t}y_{j,t-r}\}$ is a martingale difference array with respect to the natural filtration $\tilde{F}_t = \sigma(u_t, u_{t-1}, \ldots)$ under Assumption 1. Thus, the sequence is serially uncorrelated, implying that
\[E \left[ \left( \frac{\sum_{t=1}^{T-h_T} u_{i,t}y_{j,t-r}}{(T - h_T)^{1/2}g_{j,T}} \right)^2 \right] = \frac{1}{(T - h_T)g_{j,T}^2} \sum_{t=1}^{T-h_T} E[u_{i,t}^2y_{j,t-r}^2] \leq \frac{1}{g_{j,T}^2} \times [E(u_{i,t}^4)]^{1/2} \times \max_{1 \leq t \leq T-h_T} E(y_{j,t-1}^4)^{1/2} \leq \sqrt{6}C_1(E(||u_t||^4))^2 \leq \frac{\delta\lambda_{\min}(\Sigma)}{\lambda_{\min}(\Sigma)}.
\]

where the last inequality uses Lemma E.3. The lemma follows from Markov’s inequality.

**E.4 Proof of Lemma A.4**

We will show that
\[E \left[ \left( \frac{\sum_{t=1}^{T-h_T} \xi_{t,t}(A_T, h_T)y_{j,t-r}}{(T - h_T)v(A_T, h_T, w)g(\rho^*_j(A_T, \epsilon), T - h_T)} \right)^2 \right] \to 0.
\]
To that end, observe that if \( t \geq s + h_T \), then

\[
E[\xi_{i,t}(A_T, h_T)y_{j,t-r}\xi_{i,s}(A_T, h_T)y_{j,s-r}]
= E[E(\xi_{i,t}(A_T, h_T) \mid u_t, u_{t-1}, \ldots) y_{j,t-r}\xi_{i,s}(A_T, h_T)y_{j,s-r}]
= 0,
\]

by Assumption 1. By symmetry, the far left-hand side above equals 0 also if \( s \geq t + h_T \). Thus,

\[
E \left[ \frac{\sum_{t=1}^{T-h_T} \xi_{i,t}(A_T, h_T)y_{j,t-r}}{(T - h_T)v(A_T, h_T, w)g(\rho'_T(A_T, \epsilon), T - h_T)} \right]^2
\leq \sum_{t=1}^{T-h_T} \sum_{s=1}^{T-h_T} \mathbb{1}(|s - t| < h_T) \frac{|E[\xi_{i,t}(A_T, h_T)y_{j,t-r}\xi_{i,s}(A_T, h_T)y_{j,s-r}]|}{(T - h_T)^2v(A_T, h_T, w)g(\rho'_T(A_T, \epsilon), T - h_T)^2}.
\]

(S6)

We now bound the summands on the right-hand side above. Consider first the case \( s \in (t - h_T, t] \) (we will handle the case \( t \in (s - h_T, s] \) by symmetry). Since the initial conditions for the VAR are zero, we have

\[
y_{j,t-r} = \xi_{j,0}(A_T, t - r).
\]

Thus,

\[
E[\xi_{i,t}(A_T, h_T)y_{j,t-r}\xi_{i,s}(A_T, h_T)y_{j,s-r}]
= E[\xi_{i,t}(A_T, h_T)\xi_{j,0}(A_T, t - r)\xi_{i,s}(A_T, h_T)\xi_{0,j}(A_T, t - r)]
= \sum_{\ell_1=1}^{h_T} \sum_{\ell_2=1}^{h_T} \sum_{m_1=r}^{t-1} \sum_{m_2=r}^{t-1} E \left[ (\beta_i(A_T, h_T - \ell_1) \epsilon_{t+\ell_1}) (\beta_j(A_T, m_1 - r) \epsilon_{t-m_1}) \right.
\times (\beta_i(A_T, h_T - \ell_2) \epsilon_{s+\ell_2}) (\beta_j(A_T, m_2 - r) \epsilon_{s-m_2})]
\]

Consider any summand above defined by its indices \( (\ell_1, \ell_2, m_1, m_2) \). Since \( t + \ell_1 > \max\{t - m_1, s - m_2\} \), Assumption 1 implies that the summand can only be nonzero if \( s + \ell_2 = t + \ell_1 \), which requires \( \ell_1 \leq h_T + s - t \). Moreover, when \( s + \ell_2 = t + \ell_1 \), we also need \( t - m_1 = s - m_2 \) for the summand to be nonzero, which in turn requires \( m_1 \geq t - s + 1 \). Thus,

\[
|E[\xi_{i,t}(A_T, h_T)y_{j,t-r}\xi_{i,s}(A_T, h_T)y_{j,s-r}]|
\]
\[
\begin{align*}
&\leq \sum_{\ell_1=1}^{h_T+s-t} \sum_{m_1=t-s+r}^{t-r} E \left[ \sum_{\ell_1=1}^{h_T+s-t} \sum_{m_1=t-s+r}^{t-r} \beta_i(A_T, h_T - \ell_1) (\beta_i(A_T, h_T - \ell_1 - (t-s)) u_{t+\ell_1}) \times (\beta_j(A_T, m_1 - r) u_{m_1-r}) \right] \\
&\quad \times (\beta_j(A_T, m_1 - r) u_{m_1-r}) (\beta_j(A_T, m_1 - r - (t-s)) u_{m_1-r})] \\
&= \sum_{\ell_1=1}^{h_T+s-t} \sum_{m_1=t-s+r}^{t-r} \|\beta_i(A_T, h_T - \ell_1)\| \times \|\beta_i(A_T, h_T - \ell_1 - (t-s))\| \times \|\beta_j(A_T, m_1 - r)\| \\
&\quad \times \|\beta_j(A_T, m_1 - r - (t-s))\| \times E \left[ \|u_{t+\ell_1}\|^2 \times \|u_{m_1-r}\|^2 \right] \\
&\quad \text{(by Cauchy-Schwarz)} \\
&\leq C_1^2 E(\|u_0\|^4) \sum_{\ell_1=1}^{h_T+s-t} \sum_{m_1=t-s+r}^{t-r} \|\beta_i(A_T, h_T - \ell_1)\| \times \|\beta_i(A_T, h_T - \ell_1 - (t-s))\| \\
&\quad \times \left( \sum_{m_1=t-s+r}^{t-r} \rho_j^*(A, \epsilon)^{2(m_1-r-(t-s))} \right) \\
&\leq E(\|u_0\|^4) \times \rho_j^*(A, \epsilon)^{(t-s)} \left( \sum_{\ell_1=1}^{h_T+s-t} \|\beta_i(A_T, h_T - \ell_1)\| \times \|\beta_i(A_T, h_T - \ell_1 - (t-s))\| \right) \\
&\quad \times \left( \sum_{m_1=t-s+r}^{t-r} \rho_j^*(A, \epsilon)^{2(m_1-r-(t-s))} \right) \\
&\quad \text{(using Lemma E.4 and the definition of } B_i^p(A_T, h_T - \ell_1 - (t-s)) \text{ in Lemma E.9 below)} \\
&= E(\|u_0\|^4) \times \rho_j^*(A, \epsilon)^{(t-s)} \left( \sum_{\ell_0=0}^{h_T-1-(t-s)} B_i^p(A_T, \ell) \right) \left( \sum_{m_0=0}^{s-2r} \rho_j^*(A, \epsilon)^{2m} \right) \\
&\leq E(\|u_0\|^4) \times \rho_j^*(A, \epsilon)^{(t-s)} \left( \sum_{\ell_0=0}^{h_T-1-(t-s)} B_i^p(A_T, \ell) \right) \left( \sum_{m_0=0}^{s-2r} \rho_j^*(A, \epsilon)^{2m} \right) \\
&\leq E(\|u_0\|^4) \times \rho_j^*(A, \epsilon)^{(t-s)} \rho_i^*(A, \epsilon)^{(t-s)} \left( \sum_{\ell_0=0}^{h_T-1-(t-s)} B_i^p(A_T, \ell) \right) g(\rho_j^*(A, \epsilon), T - h_T)^2 \\
&\leq E(\|u_0\|^4) \times \rho_j^*(A, \epsilon)^{(t-s)} \rho_i^*(A, \epsilon)^{(t-s)} \left( \sum_{\ell_0=0}^{h_T-1-(t-s)} B_i^p(A_T, \ell) \right) g(\rho_j^*(A, \epsilon), T - h_T)^2 \\
&\quad \times C_2 p \left( \sum_{\ell_0=0}^{h_T-1} \|\beta_i(A_T, \ell)\|^2 \right) g(\rho_j^*(A, \epsilon), T - h_T)^2 \\
&\quad \text{(by Lemma E.9 below).}
\end{align*}
\]
We have derived the bound in the above display under the assumption $s \in (t - h_T, t]$, but by symmetry, it also applies when $t \in (s - h_T, s]$ if we replace $(t - s)$ with $|t - s|$. Inserting into (S6), we get

$$
E \left[ \left( \sum_{t=1}^{T-h_T} \xi_{i,t}(A_T, h_T) y_{j,t-r} \overline{(T - h_T) v(A_T, h_T, w)} g(\rho_j^*(A_T, \epsilon), T - h_T) \right)^2 \right] 
$$

\[
\leq C_{2p} \times \frac{E(||u_0||^4)}{(T - h_T)^2} \times \frac{\sum_{l=0}^{h_T-1} \beta_l(A_T, \ell)^2}{\sum_{l=0}^{h_T-1} \beta_l(A_T, \ell)^2} \times \sum_{t=1}^{T-h_T} \sum_{s=1}^{T-h_T} \mathbb{1}(|s - t| < h_T) \left( \rho_j^*(A_T, \epsilon) \rho_i^*(A_T, \epsilon) \right)^{|t-s|} 
\]

(\text{where we have used the lower bound of Lemma E.5})

\[
= \frac{C_{2p}}{||w||^2 \times \delta \times \lambda_{\min}(\Sigma)} \times \frac{E(||u_0||^4)}{(T - h_T)^2} \times \sum_{|m| < h_T} \left( 1 - \frac{|m|}{T - h_T} \right) \left( \rho_j^*(A_T, \epsilon) \rho_i^*(A_T, \epsilon) \right)^{|m|} 
\]

\[
\leq \frac{C_{2p}}{||w||^2 \times \delta \times \lambda_{\min}(\Sigma)} \times \frac{E(||u_0||^4)}{(T - h_T)^2} \times \sum_{m=0}^{h_T-1} \left( \rho_j^*(A_T, \epsilon) \rho_i^*(A_T, \epsilon) \right)^m 
\]

\[
\leq \frac{C_{2p}}{||w||^2 \times \delta \times \lambda_{\min}(\Sigma)} \times \frac{E(||u_0||^4)}{(T - h_T)^2} \times \left( \sum_{m=0}^{h_T-1} \rho_j^*(A_T, \epsilon)^{2m} \right)^{1/2} \left( \sum_{m=0}^{h_T-1} \rho_i^*(A_T, \epsilon)^{2m} \right)^{1/2} 
\]

(by Cauchy-Schwarz)

\[
\leq \frac{C_{2p} \times E(||u_0||^4)}{||w||^2 \times \delta \times \lambda_{\min}(\Sigma)} \times \left( \frac{g(\rho_j^*(A_T, \epsilon), T - h_T)}{T - h_T} \right)^{1/2} \left( \frac{g(\rho_i^*(A_T, \epsilon), T - h_T)}{T - h_T} \right)^{1/2} 
\]

\[
\to 0. 
\]

\[\square\]

**Lemma E.9.** Consider any lag polynomial $A(L)$ of order $p$ with autoregressive coefficients $A = (A_1, \ldots, A_p)$. Then for any $h = 1, 2, \ldots,$

$$
\frac{\sum_{\ell=0}^{h-1} B^p_t(A, \ell)}{\sum_{\ell=0}^{h-1} ||\beta_t(A, \ell)||^2} \leq C_{2p}, 
$$

where

$$
B^p_t(A, \ell) \equiv C_2 \sum_{b=0}^{p-1} (||\beta_t(A, \ell)|| \times ||\beta_t(A, \ell - b)||),
$$

and we define $\beta_t(A, \ell) = 0$ whenever $\ell < 0$. Here $C_2$ is the constant defined in Lemma E.4.
Proof. Changing the order of summation, we have
\[
\sum_{\ell=0}^{h-1} \left( \sum_{b=0}^{p-1} \| \beta_i(A, \ell) \| \times \| \beta_i(A, \ell - b) \| \right)
= \sum_{b=0}^{p-1} \left( \sum_{\ell=0}^{h-1} \| \beta_i(A, \ell) \| \times \| \beta_i(A, \ell - b) \| \right)
\leq \sum_{b=0}^{p-1} \left( \sum_{\ell=0}^{h-1} \| \beta_i(A, \ell) \|^2 \right)^{1/2} \times \left( \sum_{\ell=0}^{h-1} \| \beta_i(A, \ell - b) \|^2 \right)^{1/2}
\leq \sum_{b=0}^{p-1} \left( \sum_{\ell=0}^{h-1} \| \beta_i(A, \ell) \|^2 \right)
\quad \text{(since } \| \beta_i(A, \ell - b) \| = 0 \text{ for } \ell - b < 0)
= p \left( \sum_{\ell=0}^{h-1} \| \beta_i(A, \ell) \|^2 \right).
\]

Therefore,
\[
\sum_{\ell=0}^{h-1} B_1(A, \ell) \leq C_2 p \left( \sum_{\ell=0}^{h-1} \| \beta_i(A, \ell) \|^2 \right). \tag*{\square}
\]

E.5 Proof of Lemma A.5

We consider each statement separately.

Part (i). Since \( E(u_t u_t') = \Sigma \) by definition, this statement follows from a standard application of Chebyshev’s inequality, exploiting the summability of the autocovariances of \( \{u_t \otimes u_t\} \), cf. Assumption 2(ii). See for example Davidson (1994, Thm. 19.2).

Part (ii). Using \( \hat{u}_t(h) - u_t = (A - \hat{A}(h))X_t \), we get
\[
\left\| \hat{\Sigma}(h_T) - \frac{1}{T - h_T} \sum_{t=1}^{T-h_T} u_t u_t' \right\|
\leq \frac{1}{T - h_T} \sum_{t=1}^{T-h_T} \| \hat{u}_t(h_T) \hat{u}_t(h_T)' - u_t u_t' \|
\leq \frac{1}{T - h_T} \sum_{t=1}^{T-h_T} \| \hat{u}_t(h_T) - u_t \|^2 + \frac{2}{T - h_T} \sum_{t=1}^{T-h_T} \| (\hat{u}_t(h_T) - u_t) u_t' \|
\leq \| G(A_T, T - h_T, \epsilon)(\hat{A}(h_T) - A_T) \|^2 \times \frac{1}{T - h_T} \sum_{t=1}^{T-h_T} \| G(A_T, T - h_T, \epsilon)^{-1} X_t \|^2
\]
Lemma E.3, Lemma A.3(iii), Lemma E.8, and an application of Markov’s inequality imply that the last expression above is

\[ \frac{o_{P_{\mathcal{A}T}}(1) \times O_{P_{\mathcal{A}T}}(1) + 2 \times o_{P_{\mathcal{A}T}}(1) \times o_{P_{\mathcal{A}T}}(1)}{o_{P_{\mathcal{A}T}}(1)}. \]

### E.6 Proof of Lemma A.6

We would like to show \( \hat{\varsigma} \xrightarrow{P} 1 \), where

\[ \hat{\varsigma} \equiv \frac{1}{T - h_T} \sum_{t=1}^{T-h_T} \frac{\xi_{i,t}(A_T, h_T)^2(w'u_t)^2}{v(A_T, h_T, w)^2}. \]

Note that the summands could be serially correlated under our assumptions. We establish the desired convergence in probability by showing that the variance of \( \hat{\varsigma} \) tends to 0 (since its mean is 1). Observe that

\[
\text{Var}(\hat{\varsigma}) = \frac{1}{(T - h_T)^2 v(A_T, h_T, w)^4} \sum_{t=1}^{T-h_T} \sum_{s=1}^{T-h_T} \text{Cov} \left( \xi_{i,t}(A_T, h_T)^2(w'u_t)^2, \xi_{i,s}(A_T, h_T)^2(w'u_s)^2 \right) \\
= \frac{1}{(T - h_T)^2 v(A_T, h_T, w)^4} \times \sum_{|m|<T-h_T} \left( 1 - \frac{|m|}{T - h_T} \right) \text{Cov} \left( \xi_{i,0}(A_T, h_T)^2(w'u_0)^2, \xi_{i,m}(A_T, h_T)^2(w'u_m)^2 \right) \\
\leq \frac{2}{(T - h_T)^2 v(A_T, h_T, w)^4} \sum_{m=0}^{T-h_T} |\Gamma_T(m)|, \quad (S7)
\]

where we define

\[ \Gamma_T(m) \equiv \text{Cov} \left( \xi_{i,0}(A_T, h_T)^2(w'u_{i,0})^2, \xi_{i,m}(A_T, h_T)^2(w'u_m)^2 \right), \quad m = 0, 1, 2, \ldots \]

By expanding the squares \( \xi_0(\rho, h)^2 \) and \( \xi_m(\rho, h)^2 \), we obtain

\[ \Gamma_T(m) = \sum_{\ell_1=1}^{h_T} \sum_{\ell_2=1}^{h_T} \sum_{\ell_3=1}^{h_T} \sum_{\ell_4=1}^{h_T} \text{Cov} \left( \beta_1(A_T, h_T - \ell_1)'u_{\ell_1} \beta_1(A_T, h_T - \ell_2)'u_{\ell_2}(w'u_0)^2, \right. \\
\left. (\beta_1(A_T, h_T - \ell_3)'u_{m+\ell_3} \beta_1(A_T, h_T - \ell_4)'u_{m+\ell_4}(w'u_m)^2 \right). \]
Consider any summand on the right-hand side above defined by indices \((\ell_1, \ell_2, \ell_3, \ell_4)\). If \(\ell_1 = \ell_2\), then Assumption 1 implies that the covariance in the summand equals zero whenever \(\ell_3 \neq \ell_4\), since in this case at most one of the subscripts \(m + \ell_3\) or \(m + \ell_4\) can equal \(\ell_1\) (= \(\ell_2\)). Thus, if \(\ell_1 = \ell_2\), then the summand can only be nonzero when \(\ell_3 = \ell_4\). If instead \(\ell_1 \neq \ell_2\), then Assumption 1 implies that the summand can only be nonzero when \(\ell_1, \ell_2\} = \{m + \ell_3, m + \ell_4\}\), which in turn requires that \(m < h_T\). Putting these facts together, we obtain

\[
|\Gamma_T(m)| \\
\leq \sum_{\ell_1=1}^{h_T} \sum_{\ell_3=1}^{h_T} \left| \text{Cov} \left( (\beta_i(A_T, h_T - \ell_1)'u_{m+\ell_1}, (\beta_i(A_T, h_T - \ell_3)'u_{\ell_3})^2(w'u_m)^2, (\beta_i(A_T, h_T - \ell_3)'u_{\ell_3})^2(w'u_0)^2 \right) \right| \quad \text{(S8)}
\]

\[
+ \mathbb{1}(m < h_T/2) \sum_{\ell_1=1}^{h_T} \sum_{\ell_3 \neq \ell_1}^{h_T} \left| \text{Cov} \left( (\beta_i(A_T, h_T - \ell_1)'u_{\ell_1}, (\beta_i(A_T, h_T - \ell_2)'u_{\ell_2})^2(w'u_m)^2, (\beta_i(A_T, h_T - (\ell_1 - m))'u_{\ell_1}, (\beta_i(A_T, h_T - (\ell_2 - m))'u_{\ell_2})^2(w'u_0)^2 \right) \right| . \quad \text{(S9)}
\]

Let \(\Gamma_1, T(m)\) and \(\Gamma_2, T(m)\) denote expressions (S8) and (S9), respectively. We will now bound \(\sum_{m=0}^{T-h_T} \Gamma_1, T(m)\) and \(\sum_{m=0}^{T-h_T} \Gamma_2, T(m)\), so that we can ultimately insert these bounds into (S7).

**Bound on \(\sum_{m=0}^{T-h_T} \Gamma_1, T(m)\).** We first bound the term in expression (S8). To do this, we define the unit-norm vectors

\[
\omega_{A_T, h_T, \ell} \equiv \beta_i(A_T, h_T - \ell)/\|\beta_i(A_T, h_T - \ell)\|, \quad \omega_w \equiv w/\|w\|.
\]

By Lemma E.4, the term

\[
\left| \text{Cov} \left( (\beta_i(A_T, h_T - \ell_1)'u_{m+\ell_1})^2(w'u_m)^2, (\beta_i(A_T, h_T - \ell_3)'u_{\ell_3})^2(w'u_0)^2 \right) \right|
\]

is bounded above by

\[
\|w\|^4 C_1^4 \rho_i^*(A_T, \epsilon)^2(h_T - \ell_1) + 2(h_T - \ell_3) |\text{Cov} \left( (\omega_{A_T, h_T, \ell_1} u_{m+\ell_1})^2(w'_m u_m)^2, (\omega_{A_T, h_T, \ell_3} u_{\ell_3})^2(w'_u u_0)^2 \right) | .
\]

Since \(A_T \in \mathcal{A}(0, \epsilon, C)\), we have \(\rho_i^*(A_T, \epsilon) \leq 1\), so

\[
\sum_{m=0}^{T-h_T} \Gamma_1, T(m) \leq \|w\|^4 C_1^4 \sum_{m=0}^{T-h_T} \sum_{\ell_1=1}^{h_T} \sum_{\ell_3=1}^{h_T} \rho_i^*(A_T, \epsilon)^2(h_T - \ell_3)
\]
\[
\begin{aligned}
&\times \left| \text{Cov} \left( (\omega'_{A_T, h_T, \ell_1} u_{m+\ell_1})^2 (\omega'_w u_m)^2, (\omega'_{A_T, h_T, \ell_3} u_{\ell_3})^2 (\omega'_w u_0)^2 \right) \right| \\
\leq &\|w\|^4 C^4_1 \sum_{b_1=1}^{h_T} \rho^*_1(A_T, \epsilon)^{2(h_T-b_1)} \\
&\times \left( \sum_{b_2=-\infty}^{\infty} \sum_{b_3=-\infty}^{\infty} \sup_{\|\omega_j\|=1} \left| \text{Cov} \left( (\omega'_1 u_{b_1})^2 (\omega'_2 u_0)^2, (\omega'_3 u_{b_3+b_2})^2 (\omega'_4 u_{b_3})^2 \right) \right| \right). \quad (S10)
\end{aligned}
\]

Consider the double sum in large parentheses above. If we expand the various squares of the form \((\omega'_i u_t)^2\), then the double sum can be bounded above by at most \(4n^2\) terms of the form
\[
\sum_{b_2=-\infty}^{\infty} \sum_{b_3=-\infty}^{\infty} |\text{Cov} \left( \bar{u}_{j_1, b_1}, \bar{u}_{j_2, b_0}, \bar{u}_{j_3, b_1+b_2}, \bar{u}_{j_4, b_3} \right)|, \quad (S11)
\]

where \(\bar{u}_t = (\bar{u}_{1,t}, \ldots, \bar{u}_{n,t})' \equiv u_t \otimes u_t\), and \(j_1, j_2, j_3, j_4 \in \{1, 2, \ldots, n^2\}\) are summation indices.

By Assumption 2(ii), the process \(\{\bar{u}_t\}\) has absolutely summable cumulants up to order four. We can therefore show there exists a constant \(K \in (0, \infty)\) such that the large parenthesis (S10) is bounded above by \(K\).\(^1\) Consequently,
\[
\sum_{m=0}^{T-h_T} \hat{\Gamma}_{1,T}(m) \leq \|w\|^4 C^4_1 K \sum_{b_1=1}^{h_T} \rho^*_1(A_T, \epsilon)^{2(h_T-b_1)} = \|w\|^4 C^4_1 K \sum_{\ell=0}^{h_T-1} \rho^*_1(A_T, \epsilon)^{2\ell}.
\]

\(^1\)According to Brillinger (2001, Thm. 2.3.2),
\[
\text{Cov} \left( \bar{u}_{j_1, b_1}, \bar{u}_{j_2, b_0}, \bar{u}_{j_3, b_1+b_2}, \bar{u}_{j_4, b_3} \right) = \text{Cov} \left( \bar{u}_{j_2, b_0}, \bar{u}_{j_3, b_2} \right) \text{Cov} \left( \bar{u}_{j_1, b_1}, \bar{u}_{j_4, b_3} \right) + \text{Cov} \left( \bar{u}_{j_2, b_0}, \bar{u}_{j_4, b_3} \right) \text{Cov} \left( \bar{u}_{j_1, b_1}, \bar{u}_{j_3, b_2} \right) + \text{Cum} \left( \bar{u}_{j_2, b_0}, \bar{u}_{j_1, b_1}, \bar{u}_{j_3, b_2}, \bar{u}_{j_4, b_3} \right),
\]

where “Cum” denotes the joint fourth-order cumulant. Thus, the expression (S11) is bounded above by
\[
\left( \sum_{b_2=-\infty}^{\infty} |\text{Cov} \left( \bar{u}_{j_2, b_0}, \bar{u}_{j_3, b_2} \right)| \right) \left( \sum_{b_3=-\infty}^{\infty} |\text{Cov} \left( \bar{u}_{j_1, b_1}, \bar{u}_{j_4, b_3} \right)| \right)
+ \left( \sum_{b_2=-\infty}^{\infty} |\text{Cov} \left( \bar{u}_{j_2, b_0}, \bar{u}_{j_3, b_2} \right)| \right) \left( \sum_{b_3=-\infty}^{\infty} |\text{Cov} \left( \bar{u}_{j_2, b_0}, \bar{u}_{j_4, b_3} \right)| \right)
+ \sum_{b_1=-\infty}^{\infty} \sum_{b_2=-\infty}^{\infty} \sum_{b_3=-\infty}^{\infty} |\text{Cum} \left( \bar{u}_{j_2, b_0}, \bar{u}_{j_1, b_1}, \bar{u}_{j_3, b_2}, \bar{u}_{j_4, b_3} \right)|.
\]

The third term above is finite by Assumption 2(ii), since \(\bar{u}_t \equiv u_t \otimes u_t\) has absolutely summable cumulants up to order 4. Consider the first term above (the second term is handled similarly). The stationarity of \(\bar{u}_t\) implies that this term equals \(\left( \sum_{b_2=-\infty}^{\infty} |\text{Cov} \left( \bar{u}_{j_2, b_0}, \bar{u}_{j_3, b_2} \right)| \right) \left( \sum_{b_3=-\infty}^{\infty} |\text{Cov} \left( \bar{u}_{j_1, b_1}, \bar{u}_{j_4, b_3} \right)| \right)\). By Assumption 2(ii), the autocovariances of \(\{\bar{u}_t\}\) are absolutely summable. This implies the above display is bounded. Thus, we have shown that there exists a constant \(K(J_1, J_2, J_3, J_4)\) (which only depends on the fixed data generating process for \(\{u_t\}\)) that bounds the expression (S11). Picking the largest constant over all summation indices gives the desired result.
Bound on $\sum_{m=0}^{T-h_T} \tilde{\Gamma}_{2,T}(m)$. Expression (S9) can be bounded above by

$$\mathbb{1}(m < h_T) \sum_{\ell_1=1}^{h_T} \sum_{\ell_2 \neq \ell_1} E \left[ |\beta_i(A_T, h_T - \ell_1)'u_{\ell_1}| \times |\beta_i(A_T, h_T - \ell_2)'u_{\ell_2}| \times (w'u_m)^2 \right].$$

Applying Cauchy-Schwarz, we get the upper bound

$$\mathbb{1}(m < h_T) 2 \sum_{\ell_1=1}^{h_T} \sum_{\ell_2 \neq \ell_1} \left( \|w\|^4 \times \|\beta_i(A_T, h_T - \ell_1)\| \times \|\beta_i(A_T, h_T - \ell_2)\| \right. \times \|\beta_i(A_T, h_T - (\ell_1 - m))\| \times \|\beta_i(A_T, h_T - (\ell_2 - m))\| \times E \left[ \|u_{\ell_1}\|^2 \times \|u_{\ell_2}\|^2 \times \|u_m\|^2 \times \|u_0\|^2 \right].$$

Another application of the Cauchy-Schwarz inequality gives

$$E \left[ \|u_{\ell_1}\|^2 \times \|u_{\ell_2}\|^2 \times \|u_m\|^2 \times \|u_0\|^2 \right] \leq E[\|u_i^8\|].$$

Thus,

$$\sum_{m=0}^{T-h_T} \tilde{\Gamma}_{2,T}(m) \leq 2 \times E[\|u_i^8\|] \times \|w\|^4 \times \sum_{m=0}^{h_T-1} \sum_{\ell_1=1}^{h_T} \sum_{\ell_2=1}^{h_T} \left( \|\beta_i(A_T, h_T - \ell_1)\| \times \|\beta_i(A_T, h_T - \ell_2)\| \times \|\beta_i(A_T, h_T - (\ell_1 - m))\| \times \|\beta_i(A_T, h_T - (\ell_2 - m))\| \right).$$

The bound in Lemma E.4 implies that

$$\|\beta_i(A_T, h_T - \ell_1)\| \times \|\beta_i(A_T, h_T - (\ell_1 - m))\|$$

is less than or equal to

$$C_2 \sum_{b=0}^{p-1} \|\beta_i(A_T, h_T - \ell_1)\| \times \|\beta_i(A_T, h_T - \ell_1 - b)\| \times \rho_i^*(A_T, \epsilon)^m, \quad (S12)$$

$$\equiv \beta_i^p(A_T, h_T - \ell_1)$$
for a positive constant $C_2$ that depends on $p$ and $\epsilon$. Thus,

$$
\sum_{m=0}^{T-h_T} \hat{\Gamma}_{2,T}(m)
\leq 2 \times E[ \|u^8_i\| ] \times \|w\|^4 \times \sum_{m=0}^{h_T-1} \sum_{\ell_1=1}^{h_T} \sum_{\ell_2=1}^{h_T} \left( B^p_i(A_T, h_T - \ell_1) \times B^p_i(A_T, h_T - \ell_2) \times \rho_i^*(A_T, \epsilon)^{2m} \right)
= 2 \times E[ \|u^8_i\| ] \times \|w\|^4 \left( \sum_{\ell=0}^{h_T-1} \rho_i^*(A_T, \epsilon)^{2\ell} \right) \left( \sum_{\ell=0}^{h_T-1} B^p_i(A_T, \ell) \right)^2. \quad \text{(S13)}
$$

**Conclusion of Proof.** Putting together (S7), (S8), (S9), and (S13), we get

$$
\text{Var}(\xi) \leq \frac{2\|w\|^4}{(T-h_T)v(A_T, h_T, w)^4} \left\{ C^4_1 K \sum_{\ell=0}^{h_T-1} \rho_i^*(A_T, \epsilon)^{2\ell} \right. \\
+ 2 \times E[ \|u^8_i\| ] \times \left( \sum_{\ell=0}^{h_T-1} \rho_i^*(A_T, \epsilon)^{2\ell} \right) \left( \sum_{\ell=0}^{h_T-1} B^p_i(A_T, \ell) \right)^2 \left\}
\leq \frac{2C^4_1 K \times \sum_{\ell=0}^{h_T-1} \rho_i^*(A_T, \epsilon)^{2\ell}}{(T-h_T) \left( \sum_{\ell=0}^{h_T-1} \|\beta_i(A, \ell)\|^2 \right)^2 \delta^2 \lambda_{\min}(\Sigma)^2} \\
+ \frac{2 \times E[ \|u^8_i\| ] \times \sum_{\ell=0}^{h_T-1} \rho_i^*(A_T, \epsilon)^{2\ell}}{(T-h_T)\delta^2 \lambda_{\min}(\Sigma)^2} \times \frac{\left( \sum_{\ell=0}^{h_T-1} B^p_i(A_T, \ell) \right)^2}{\left( \sum_{\ell=0}^{h_T-1} \|\beta_i(A, \ell)\|^2 \right)^2}
$$

(by the lower bound for $v(A_T, h_T, w)^2$ derived in Lemma E.5)

$$
\leq \frac{2 \left\{ \left( C^4_1 K \times \sum_{\ell=0}^{h_T-1} \rho_i^*(A_T, \epsilon)^{2\ell} \right) + \left( 2 \times E[ \|u^8_i\| ] \times C_2p \times \sum_{\ell=0}^{h_T-1} \rho_i^*(A_T, \epsilon)^{2\ell} \right) \right\}}{(T-h_T)\delta^2 \lambda_{\min}(\Sigma)^2} \\
$$

(where we have used $\sum_{\ell=0}^{\ell_T-1} \|\beta_i(A, \ell)\|^2 \geq \|\beta_i(A, 0)\|^2 = 1$ and Lemma E.9)

$$
= \frac{(2 \times C^4_1 K) + (4 \times E[ \|u^8_i\| ] \times C_2p)}{\delta^2 \lambda_{\min}(\Sigma)^2} \times \frac{\sum_{\ell=0}^{h_T-1} \rho_i^*(A_T, \epsilon)^{2\ell}}{T-h_T}.
$$

The final expression above tends to zero as $T \to \infty$, since

$$
\frac{\sum_{\ell=0}^{h_T-1} \rho_i^*(A_T, \epsilon)^{2\ell}}{T-h_T} \leq \frac{g(\rho_i^*(A_T, \epsilon), h_T)^2}{T-h_T} \to 0.
$$

Thus, $\text{Var}(\xi) \to 0.$
E.7 Proof of Lemma A.7

We prove only the first statement of the lemma, as the proof is completely analogous for the second part. Define the unit-norm vectors

\[ \omega_{A,h,\ell} \equiv \beta_i(A, h - \ell)/\|\beta_i(A, h - \ell)\|, \quad \omega_w \equiv w/\|w\|. \]

In a slight abuse notation, throughout the proof of this lemma we will sometimes write \( \beta_i(h - \ell) \) instead of \( \beta_i(A, h - \ell) \). Expanding the four-fold product \( \xi_{i, \ell}(A, h)^4 \), we obtain

\[
E[\xi_{i, \ell}(A, h)^4(a'u_t)^4] \\
= \sum_{\ell_1=1}^{h} \sum_{\ell_2=1}^{h} \sum_{\ell_3=1}^{h} \sum_{\ell_4=1}^{h} \|\beta_i(h - \ell_1)\| \times \|\beta_i(h - \ell_2)\| \times \|\beta_i(h - \ell_3)\| \times \|\beta_i(h - \ell_4)\| \\
\times E \left[ (\omega_{A,h,\ell_1}^{\prime} u_{t+\ell_1}) \times (\omega_{A,h,\ell_2}^{\prime} u_{t+\ell_2}) \times (\omega_{A,h,\ell_3}^{\prime} u_{t+\ell_3}) \times (\omega_{A,h,\ell_4}^{\prime} u_{t+\ell_4}) \times (w' u_t)^4 \right]. \quad (S14)
\]

By Assumption 1, the summands above equal zero if one of the indices \( \ell_j \) is different from the three other indices. Hence, the only possibly nonzero summands are those for which the four indices appear in two pairs, e.g., \( \ell_1 = \ell_3 \) and \( \ell_2 = \ell_4 \). The typical nonzero summand can thus be written in the form

\[
\|\beta_i(h - \ell)\| \|\beta_i(h - m)\|^2 E \left[ (\omega_{A,h,\ell}^{\prime} u_{t+\ell})^2 \times (\omega_{A,h,m}^{\prime} u_{t+m})^2 \times (w' u_t)^4 \right]
\]

where \( \ell, m \in \{1, \ldots, h\} \). For given \( \ell \) and \( m \), this specific type of summand is obtained precisely when either (i) \( \ell_1 = \ell_2 = \ell \) and \( \ell_3 = \ell_4 = m \), or (ii) \( \ell_1 = \ell_3 = \ell \) and \( \ell_2 = \ell_4 = m \), or (iii) \( \ell_1 = \ell_4 = \ell \) and \( \ell_2 = \ell_3 = m \), or (iv) \( \ell_1 = \ell_2 = m \) and \( \ell_3 = \ell_4 = \ell \), or (v) \( \ell_1 = \ell_3 = m \) and \( \ell_2 = \ell_4 = \ell \), or (vi) \( \ell_1 = \ell_4 = m \) and \( \ell_2 = \ell_3 = \ell \). That is, there are six summands in (S14) of this form. Thus,

\[
E[\xi_{i, \ell}(A, h)^4(w'u_t)^4] = 6 \sum_{\ell=1}^{h} \sum_{m=1}^{h} \left( \|\beta_i(h - \ell)\|^2 \|\beta_i(h - m)\|^2 \times E \left[ (\omega_{A,h,\ell}^{\prime} u_{t+\ell})^2 \times (\omega_{A,h,m}^{\prime} u_{t+m})^2 \times (w' u_t)^4 \right] \right)
\leq 6 \|w\|^4 E(\|u_t\|^8) \sum_{\ell=1}^{h} \sum_{m=1}^{h} \|\beta_i(h - \ell)\|^2 \|\beta_i(h - m)\|^2 \\
(\text{by applying Cauchy-Schwarz twice})
\]
\[ = 6\|w\|^4 E(\|u_t\|^8) \left( \sum_{\ell=0}^{\ell=h-1} \|\beta_i(A, h - \ell)\|^2 \right)^2. \]

It follows from Lemma E.5 that

\[ E \left[ \left( v(A, h, w)^{-1} \xi_t(A, h) u_t \right)^4 \right] \leq \frac{6 E(\|u_t\|^8)}{\delta^2 \lambda_{\min}(\Sigma)^2}. \]

References

