Online Appendix: Local Projection Inference is Simpler and More Robust Than You Think

José Luis Montiel Olea Mikkel Plagborg-Møller

December 4, 2020

Appendix D Further Simulation Results

Bivariate VAR(4) model. We first consider the bivariate VAR(p_0) model

$$y_{1,t} = \rho y_{1,t-1} + u_{1,t}, \quad (1 - \tfrac{1}{2}L)^{p_0} y_{2,t} = \tfrac{1}{2} y_{1,t-1} + u_{2,t}, \quad (u_{1,t}, u_{2,t})' \overset{i.i.d.}{\sim} N\left(0, \left(\begin{smallmatrix} 1 & 0.3 \\ 0.3 & 1 \end{smallmatrix}\right)\right),$$

where L is the lag operator, and the parameter ρ indexes the persistence. For $p_0 = 1$, this model reduces to the one considered by Kilian and Kim (2011, section III); we instead set $p_0 = 4$ to generate richer dynamics. The parameters of interest are the reduced-form impulse responses of $y_{2,t}$ with respect to the innovation $u_{1,t}$.

Table S1 shows that the qualitative conclusions from the AR(1) simulation study in Section 2.2 carry over to the present bivariate DGP with $p_0 = 4$. We employ four different inference methods that use the correct estimation lag length $p = p_0$: non-augmented VAR, delta method confidence interval ("AR"); lag-augmented VAR (Inoue and Kilian, 2020), Efron bootstrap interval ("AR-LA$_b$"); local projection with HAR standard errors as in Section 2.2, percentile-t bootstrap interval ("LP$_b$"); and our preferred method, lag-augmented local projection with heteroskedasticity-robust standard errors, percentile-t bootstrap interval ("LP-LA$_b$"). As a fifth method, we consider our preferred procedure with a larger estimation lag length $p = 8$ ("LP-LA$_b^8$"). The bootstrap is a wild recursive residual VAR bootstrap. We set $T = 240$. The nominal confidence level is 90%.

Consistent with the theory in Section 4, lag-augmented local projection achieves good coverage in all cases, except at long horizons $h \geq 36$ when there is a unit root ($\rho = 1$). Over-specifying the lag length to be 8 instead of 4 barely affects the coverage of lag-augmented local projection confidence intervals and only widens them by 3–5% (see columns 2 and 7). Non-augmented delta method VAR inference suffers from poor coverage at long horizons when $\rho \geq 0.95$, while lag-augmented VAR confidence intervals can be very wide.

Table S1: Monte Carlo results: bivariate VAR(4) model

	Coverage					Median length				
h	LP-LA$_b$	LP-LA$_b^8$	LP$_b$	AR-LA$_b$	AR	LP-LA$_b$	LP-LA$_b^8$	LP$_b$	AR-LA$_b$	AR
					$\rho = 0.00$					
1	0.910	0.906	0.906	0.901	0.902	0.234	0.241	0.245	0.229	0.226
6	0.892	0.892	0.899	0.894	0.895	1.481	1.518	1.517	1.310	1.278
12	0.895	0.889	0.895	0.903	0.901	1.605	1.661	1.627	3.813	0.660
36	0.906	0.901	0.905	0.924	1.000	1.694	1.754	1.709	30.081	0.015
60	0.913	0.912	0.911	0.927	1.000	1.825	1.901	1.832	301.439	0.000
					$\rho = 0.50$					
1	0.908	0.906	0.907	0.900	0.900	0.235	0.240	0.244	0.228	0.226
6	0.896	0.890	0.894	0.892	0.889	1.731	1.774	1.776	1.706	1.624
12	0.891	0.880	0.889	0.889	0.897	2.006	2.065	2.037	7.186	1.264
36	0.902	0.897	0.902	0.922	1.000	2.079	2.148	2.101	89.302	0.066
60	0.913	0.909	0.906	0.922	1.000	2.239	2.322	2.262	1517.269	0.001
					$\rho = 0.95$					
1	0.904	0.902	0.907	0.895	0.893	0.235	0.241	0.245	0.230	0.227
6	0.891	0.890	0.888	0.887	0.889	2.296	2.361	2.362	2.407	2.136
12	0.890	0.884	0.891	0.902	0.881	4.542	4.665	4.641	16.838	4.014
36	0.830	0.809	0.832	0.933	0.841	6.295	6.421	6.407	1113.555	5.413
60	0.876	0.859	0.872	0.931	0.763	6.146	6.297	6.343	73988.007	3.253
					$\rho = 1.00$					
1	0.904	0.897	0.900	0.893	0.890	0.236	0.242	0.245	0.230	0.227
6	0.894	0.892	0.890	0.859	0.874	2.381	2.445	2.472	2.450	2.181
12	0.877	0.873	0.872	0.879	0.828	5.278	5.407	5.364	17.862	4.491
36	0.767	0.760	0.769	0.965	0.775	11.346	11.558	11.509	1311.475	8.200
60	0.659	0.654	0.677	0.961	0.751	12.436	12.355	12.750	95033.410	11.423

Coverage probability and median length of nominal 90% confidence intervals at different horizons. Bivariate VAR(4) model with $\rho \in \{0, .5, .95, 1\}$, $T = 240$. 5,000 Monte Carlo repetitions; 2,000 bootstrap iterations.

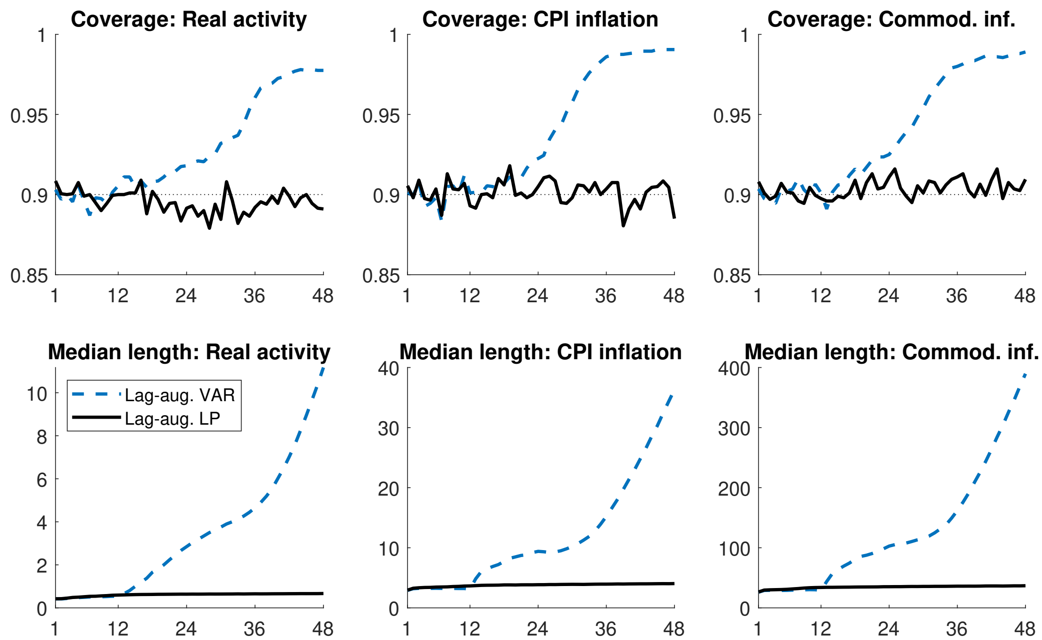
Empirically calibrated VAR(12) models. We additionally consider two empirically calibrated VAR(12) models with four or five observables. The first DGP broadly follows Kilian and Kim (2011, section IV) and is given by the empirical least-squares estimate of a workhorse monetary VAR model estimated on monthly U.S. data for 1984–2018 ($T = 419$). The four variables in the empirical VAR are the Federal Funds Rate, the Chicago Fed National Activity Index, CPI inflation, and real commodity price inflation (CRB Raw Industrials deflated by CPI).[1] The second DGP is based on the main specification in Gertler and Karadi (2015) estimated on their monthly data set for 1990–2012 ($T = 270$).[2] The five variables are industrial production (log levels), CPI (log levels), the 1-year Treasury rate, the Excess Bond Premium, and a monetary shock series given by high-frequency changes in 3-month Federal Funds Futures prices around FOMC announcements. For both DGPs, we simulate data from a Gaussian VAR(12) model with true parameters given by the empirically estimated coefficients and innovation covariance matrix (but no intercept). The sample sizes are the same as in the real data, mentioned earlier.

Figure S1 shows that lag-augmented local projection achieves acceptable coverage in these empirically calibrated DGPs. The figure shows the coverage and median length of 90% confidence intervals for reduced-form impulse responses of selected response variables with respect to an innovation in the Federal Funds Rate (first DGP) or the monetary shock series (second DGP). Our preferred lag-augmented local projection procedure (solid black line) exhibits coverage distortions below 5 percentage points at all horizons for four of the six impulse response functions shown. The distortions only approach 10 percentage points for two response variables at long horizons in the second DGP. This second DGP is very challenging: Four of the eigenvalues of the VAR companion matrix exceed 0.98 in magnitude, while the sample size (270) is small relative to the number of covariates in each equation (60 plus the intercept). The Inoue and Kilian (2020) procedure (dashed blue line) exhibits near-uniform coverage in both DGPs, but this comes at the expense of extremely large confidence interval length at medium and long horizons.

[1] St. Louis FRED codes: CFNAI, CPIAUCSL, FEDFUNDS. Global Financial Data code: CMCRBIND.

[2] The data was downloaded from: `https://www.aeaweb.org/articles?id=10.1257/mac.20130329`

Monte Carlo results: Kilian-Kim VAR(12) specification



Monte Carlo results: Gertler-Karadi VAR(12) specification

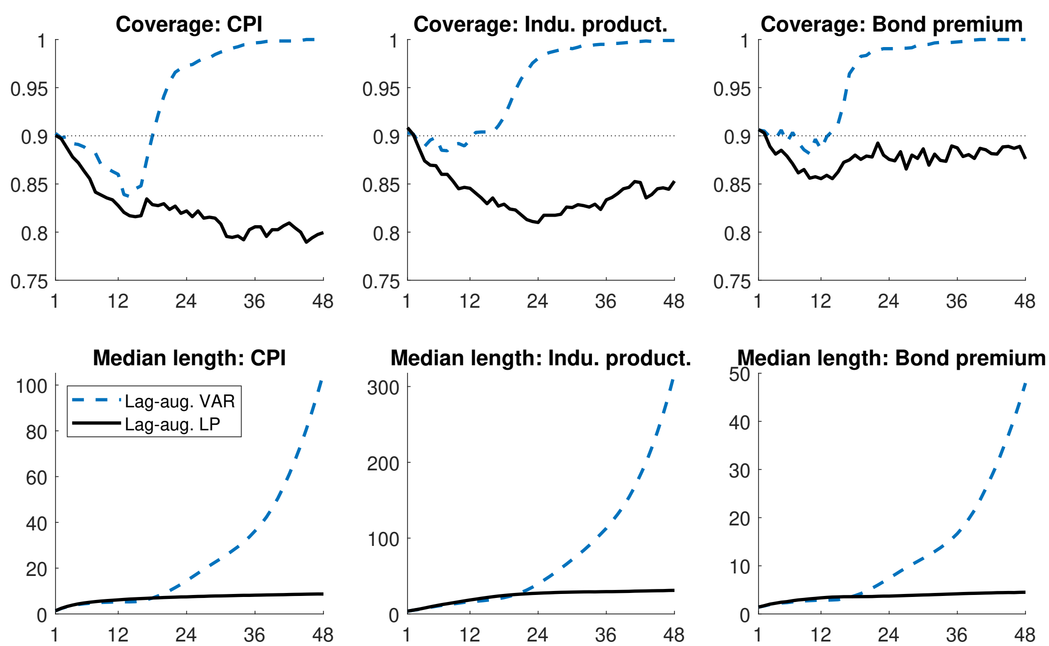


Figure S1: Coverage rate and median length of 90% confidence intervals for reduced-form impulse responses at horizons up to 48 (horizontal axis). Black solid line: lag-augmented local projection, percentile-t bootstrap interval. Blue dashed line: Inoue and Kilian (2020) Efron bootstrap interval. 2,000 Monte Carlo repetitions; 2,000 bootstrap iterations.

Appendix E Additional Proofs

E.1 Notation

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

$$1 \le \sum_{\ell=0}^{h-1} \rho_i^*(A,\epsilon)^{2\ell} \le \min\left\{\frac{1}{1-\rho_i^*(A,\epsilon)^2}, h\right\} \le \min\left\{\frac{1}{1-\rho_i^*(A,\epsilon)}, h\right\} = g(\rho_i^*(A,\epsilon),h)^2.$$

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

$$\begin{aligned}
\hat{\xi}_{1,t}(h) - \xi_{1,t}(\rho,h) &= (y_{1,t+h} - \hat{\beta}_1(h)'y_t - \hat{\gamma}_1(h)'X_t) - (y_{1,t+h} - \beta_1(A,h)'u_t - \eta_1(A,h)'X_t) \\
&= -\hat{\beta}_1(h)'\underbrace{(y_t - AX_t)}_{=u_t} - (\underbrace{\hat{\beta}_1(h)'A + \hat{\gamma}_1(h)'}_{\equiv\hat{\eta}_1(A,h)'})X_t + \beta_1(A,h)u_t + \eta_1(A,h)X_t \\
&= [\beta_1(A,h) - \hat{\beta}_1(h)]'u_t + [\eta_1(A,h) - \hat{\eta}_1(A,h)]'X_t. \qquad \text{(S1)}
\end{aligned}$$

E.2 Proof of Lemma A.2

Define $\hat{\nu}(h_T) \equiv \hat{\Sigma}(h_T)^{-1}\nu$, where $\nu \in \mathbb{R}\backslash\{0\}$ is a user-specified vector. The result follows from Lemma A.6 if we can show that

$$\frac{\sum_{t=1}^{T-h_T} \hat{\xi}_{1,t}(h_T)^2(\hat{\nu}(h_T)'\hat{u}_t(h_T))^2 - \sum_{t=1}^{T-h_T} \xi_{1,t}(h_T)^2(\tilde{\nu}'u_t)^2}{(T-h_T)v(A_T,h_T,\tilde{\nu})^2} \underset{P_{A_T}}{\overset{p}{\to}} 0,$$

where we have defined $\tilde{\nu} \equiv \Sigma^{-1}\nu$. Algebra shows that

$$\begin{aligned}
&\frac{\left|\sum_{t=1}^{T-h_T}\left[\hat{\xi}_{1,t}(h_T)^2(\hat{\nu}(h_T)'\hat{u}_t(h_T))^2 - \xi_{1,t}(A_T,h_T)^2(\tilde{\nu}'u_t)^2\right]\right|}{(T-h_T)v(A_T,h_T,\tilde{\nu})^2} \\
&\le \frac{\sum_{t=1}^{T-h_T}\left|\hat{\xi}_{1,t}(h_T)^2(\hat{\nu}(h_T)'\hat{u}_t(h_T))^2 - \xi_{1,t}(A_T,h_T)^2(\tilde{\nu}'u_t)^2\right|}{(T-h_T)v(A_T,h_T,\tilde{\nu})^2} \\
&= \frac{1}{(T-h_T)v(A_T,h_T,\tilde{\nu})^2}\sum_{t=1}^{T-h_T}\left|\hat{\xi}_{1,t}(h_T)(\hat{\nu}(h_T)'\hat{u}_t(h_T)) - \xi_{1,t}(A_T,h_T)(\tilde{\nu}'u_t)\right| \\
&\qquad \times \left|\hat{\xi}_{1,t}(h_T)(\hat{\nu}(h_T)'\hat{u}_t(h_T)) - \xi_{1,t}(A_T,h_T)(\tilde{\nu}'u_t) + 2\xi_{1,t}(A_T,h_T)(\tilde{\nu}'u_t)\right|
\end{aligned}$$

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

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

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

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

$$\left(2\frac{\sum_{t=1}^{T-h_T} \left[\hat{\xi}_{1,t}(h_T)(\hat{\nu}(h_T)'\hat{u}_t(h_T)) - \xi_{1,t}(A_T, h_T)(\tilde{\nu}'u_t) \right]^2}{(T-h_T)v(A_T, h_T, \tilde{\nu})^2} + 8\frac{\sum_{t=1}^{T-h_T} \xi_{1,t}(A_T, h_T)^2(\tilde{\nu}'u_t)^2}{(T-h_T)v(A_T, h_T, \tilde{\nu})^2} \right)^{1/2}.$$

The last fraction above is bounded in probability by Lemma A.6. Thus, it is sufficient to show that

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

converges in probability to zero. To that end, decompose

$$\begin{aligned}
&\hat{\xi}_{1,t}(h_T)(\hat{\nu}(h_T)'\hat{u}_t(h_T)) - \xi_{1,t}(A_T, h_T)(\tilde{\nu}'u_t) \\
&= (\hat{\xi}_{1,t}(h_T) - \xi_{1,t}(A_T, h_T))(\tilde{\nu}'u_t) + (\hat{\nu}(h_T)'\hat{u}_t(h_T) - \tilde{\nu}'u_T)\xi_{1,t}(A_T, h_T) \\
&\quad + (\hat{\xi}_{1,t}(h_T) - \xi_{1,t}(A_T, h_T))(\hat{\nu}(h_T)'\hat{u}_t(h_T) - \tilde{\nu}'u_T).
\end{aligned}$$

Hence, by another application of Loève's inequality,

$$\begin{aligned}
&\frac{\sum_{t=1}^{T-h_T} \left[\hat{\xi}_{1,t}(h_T)(\hat{\nu}(h_T)'\hat{u}_t(h_T)) - \xi_{1,t}(A_T, h_T)(\tilde{\nu}'u_t) \right]^2}{(T-h_T)v(A_T, h_T, \tilde{\nu})^2} \\
&\leq 3\frac{\sum_{t=1}^{T-h_T} [\hat{\xi}_{1,t}(h_T) - \xi_{1,t}(A_T, h_T)]^2(\tilde{\nu}'u_t)^2}{(T-h_T)v(A_T, h_T, \tilde{\nu})^2} \\
&\quad + 3\frac{\sum_{t=1}^{T-h_T} [\hat{\nu}(h_T)'\hat{u}_t(h_T) - \tilde{\nu}'u_t]^2 \xi_{1,t}(A_T, h_T)^2}{(T-h_T)v(A_T, h_T, \tilde{\nu})^2} \\
&\quad + 3\frac{\sum_{t=1}^{T-h_T} [\hat{\xi}_{1,t}(h_T) - \xi_{1,t}(A_T, h_T)]^2[\hat{\nu}(h_T)'\hat{u}_t(h_T) - \tilde{\nu}'u_t]^2}{(T-h_T)v(A_T, h_T, \tilde{\nu})^2} \\
&\leq 3\left(\frac{\sum_{t=1}^{T-h_T} [\hat{\xi}_{1,t}(h_T) - \xi_{1,t}(A_T, h_T)]^4}{(T-h_T)v(A_T, h_T, \tilde{\nu})^4} \right)^{1/2} \times \left(\|\tilde{\nu}\|^4 \frac{\sum_{t=1}^{T-h_T} \|u_t\|^4}{T-h_T} \right)^{1/2}
\end{aligned}$$

$$+ 3\left(\frac{\sum_{t=1}^{T-h_T}[\hat{\nu}(h_T)'\hat{u}_t(h_T) - \tilde{\nu}'u_t]^4}{T-h_T}\right)^{1/2} \times \left(\frac{\sum_{t=1}^{T-h_T}\xi_{1,t}(A_T,h_T)^4}{(T-h_T)v(A_T,h_T,\tilde{\nu})^4}\right)^{1/2}$$
$$+ 3\left(\frac{\sum_{t=1}^{T-h_T}[\hat{\xi}_{1,t}(h_T) - \xi_t(A_T,h_T)]^4}{(T-h_T)v(A_T,h_T,\tilde{\nu})^4}\right)^{1/2} \times \left(\frac{\sum_{t=1}^{T-h_T}[\hat{\nu}(h_T)'\hat{u}_t(h_T) - \tilde{\nu}'u_t]^4}{T-h_T}\right)^{1/2}$$

(by Cauchy-Schwarz)

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

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

$$\hat{R}_3 \leq \|\hat{\nu}(h_T)\|^4 \frac{\sum_{t=1}^{T-h_T}\|\hat{u}_t(h_T) - u_t\|^4}{T-h_T} + \|\hat{\nu}(h_T) - \tilde{\nu}\|^4 \frac{\sum_{t=1}^{T-h_T}\|u_t\|^4}{T-h_T}.$$

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 $w \in \mathbb{R}^n \backslash \{0\}$. Then*

$$\frac{\sum_{t=1}^{T-h_T}[\hat{\xi}_{1,t}(h_T) - \xi_{1,t}(A_T,h_T)]^4}{(T-h_T)v(A_T,h_T,w)^4} \underset{P_{A_T}}{\overset{p}{\to}} 0.$$

Proof. Recall equation (S1):

$$\hat{\xi}_{1,t}(h) - \xi_{1,t}(A,h) = [\beta_1(A,h) - \hat{\beta}_1(h)]'u_t + [\eta_1(A,h) - \hat{\eta}_1(A,h)]'X_t.$$

By Loève's inequality (Davidson, 1994, Thm. 9.28),

$$\frac{\sum_{t=1}^{T-h_T}[\hat{\xi}_{1,t}(h) - \xi_{1,t}(\rho,h)]^4}{(T-h_T)v(A_T,h_T,w)^4}$$
$$\leq 8\frac{\|\hat{\beta}_1(h) - \beta_1(A_T,h_T)\|^4}{v(A_T,h_T,w)^4}\frac{\sum_{t=1}^{T-h_T}\|u_t\|^4}{T-h_T}$$

$$+ 8\frac{\|G(A_T, T-h_T, \epsilon)[\hat{\eta}_1(A_T, h_T) - \eta_1(A_T, h_T)]\|^4}{v(A_T, h_T, w)^4} \frac{\sum_{t=1}^{T-h_T} \|G(A_T, T-h_T, \epsilon)^{-1} X_t\|^4}{T-h_T}.$$

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*

$$\frac{\sum_{t=1}^{T-h_T} \|\hat{u}_t(h_T) - u_t\|^4}{T-h_T} \underset{P_{A_T}}{\xrightarrow{p}} 0.$$

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 $T \in \mathbb{N}$, $A \in \mathcal{A}(0, C, \epsilon)$, and $i = 1, \dots, n$,*

$$\max_{1 \leq t \leq T} E(y_{i,t}^4) \leq \frac{6C_1(E(\|u_t\|^4))^3}{\delta^2 \lambda_{\min}(\Sigma)^2} \times g(\rho_i^*(A, \epsilon), T)^4$$

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

$$\xi_{i,t}(A, h) \equiv \sum_{\ell=1}^{h} \beta_i(A, \ell)' u_{t+\ell}.$$

Since we have set the initial conditions $y_0 = \ldots = y_{-p+1} = 0$, we have

$$y_{i,t} = \sum_{\ell=1}^{t} \beta_i(A, \ell)' u_\ell = \xi_{i,0}(A, t).$$

Consider any $w \in \mathbb{R}^n$ such that $\|w\| = 1$. Then Lemma A.7 gives

$$\max_{1\le t\le T} E(y_{i,t}^4) = \max_{1\le t\le T} E[\xi_{i,0}(A,t)^4] \le \frac{6E(\|u_0\|^4)}{\delta^2\lambda_{\min}(\Sigma)^2} \times \max_{1\le t\le T} v(A,t,w)^4.$$

Lemmas E.4 and E.5 below then imply that

$$\begin{aligned}\max_{1\le t\le T} E(y_{i,t}^4) &\le \frac{6E(\|u_0\|^4)}{\delta^2\lambda_{\min}(\Sigma)^2} \times (E[\|u_0\|^4])^2\|w\|^4 \times \max_{1\le t\le T}\left(\sum_{\ell=0}^{t-1}\|\beta_i(A,\ell)\|^2\right)^2 \\ &= \frac{6(E(\|u_0\|^4))^3}{\delta^2\lambda_{\min}(\Sigma)^2} \times \left(\sum_{\ell=0}^{T-1}\|\beta_i(A,\ell)\|^2\right)^2 \\ &\le \frac{6(E(\|u_0\|^4))^3}{\delta^2\lambda_{\min}(\Sigma)^2} \times \left(\sum_{\ell=0}^{T-1} C_1\rho_i^*(A,\epsilon)^{2\ell}\right)^2 \\ &\le \frac{6C_1^2(E(\|u_0\|^4))^3}{\delta^2\lambda_{\min}(\Sigma)^2} \times g(\rho_i^*(A,\epsilon),T)^4.\end{aligned}$$

□

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

i) $\|\beta_i(A,h)\| \le C_1\rho_i^*(A,\epsilon)^h$, *where* $C_1 \equiv 1 + 2C \times \frac{1-\epsilon}{\epsilon}$.

ii) $\|\beta_i(A,h+m)\| \le \rho_i^*(A,\epsilon)^m \times C_2\sum_{b=0}^{p-1}\|\beta_i(A,h-b)\|$, *where* $C_2 \equiv 1 + 4\tilde{C}\left(\frac{1-\epsilon}{\epsilon}\right)$, *and* $\tilde{C} \equiv C\left(1 + C(p-1)\right)$.

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

$$\beta_i(A,h) = \rho_i\beta_i(A,h-1) + \beta_i(B,h). \tag{S2}$$

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

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

$$
\begin{aligned}
&\leq \; |\rho_i|^h + \sum_{\ell=1}^{h} |\rho_i|^{h-\ell} C(1-\epsilon)^\ell \\
&\quad\; \text{(where we have used Lemma E.7 below and } \beta(A,0) = I_n\text{)} \\
&\leq \; \max\{|\rho_i|, 1-\epsilon/2\}^h + \sum_{\ell=1}^{h} \max\{|\rho_i|, 1-\epsilon/2\}^{h-\ell} C(1-\epsilon)^\ell \\
&= \; \rho_i^*(A,\epsilon)^h \left(1 + C \sum_{\ell=1}^{h} \left(\frac{1-\epsilon}{\max\{|\rho_i|, 1-\epsilon/2\}}\right)^\ell\right) \\
&\leq \; \rho_i^*(A,\epsilon)^h \left(1 + C \sum_{\ell=1}^{\infty} \left(\frac{1-\epsilon}{1-\epsilon/2}\right)^\ell\right) \\
&= \; \rho_i^*(A,\epsilon)^h \left(1 + C \left(\frac{1-\epsilon}{\epsilon/2}\right)\right).
\end{aligned}
$$

Part (ii). To establish the remaining inequality, note that

$$
\begin{aligned}
&\|\beta_i(A, h+m)\| \\
&\leq |\rho_i|^m \|\beta_i(A,h)\| + \sum_{\ell=1}^{m} |\rho_i|^{m-\ell} \|\beta_i(B, h+\ell)\| \\
&\leq |\rho_i|^m \|\beta_i(A,h)\| + \sum_{\ell=1}^{m} |\rho_i|^{m-\ell} \left(\tilde{C}(1-\epsilon)^\ell \sum_{b=0}^{p-2} \|\beta_i(B, h-b)\|\right) \\
&\text{(by Lemma E.7(ii) below)} \\
&\leq \max\{|\rho_i|, 1-\epsilon/2\}^m \\
&\quad \times \left(\|\beta_i(A,h)\| + \tilde{C}\left(\sum_{\ell=1}^{m}\left(\frac{1-\epsilon}{\max\{|\rho_i|, 1-\epsilon/2\}}\right)^\ell\right)\left(\sum_{b=0}^{p-2} \|\beta_i(B, h-b)\|\right)\right) \\
&\leq \rho_i^*(A,\epsilon)^m \times \left(\|\beta_i(A,h)\| + 2\tilde{C}\left(\frac{1-\epsilon}{\epsilon}\right)\left(\sum_{b=0}^{p-2} \|\beta_i(A, h-b)\| + \|\beta_i(A, h-b-1)\|\right)\right) \\
&\text{(where we have used equation (S2))} \\
&\leq \rho_i^*(A,\epsilon)^m \times \left(1 + 4\tilde{C}\left(\frac{1-\epsilon}{\epsilon}\right)\right) \sum_{b=0}^{p-1} \|\beta_i(A, h-b)\|. \qquad \square
\end{aligned}
$$

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 $h \in \mathbb{N}$*

$$
\delta \times \lambda_{\min}(\Sigma) \leq \frac{1}{\|a\|^2} \frac{v_i(A,h,w)^2}{\sum_{\ell=0}^{h-1} \|\beta_i(A,\ell)\|^2} \leq E\left(\|u_t\|^4\right),
$$

where $v_i(A,h,w) \equiv E[\xi_{i,t}(A,h)^2(w'u_t)^2]$

Proof. Algebra shows

$$\begin{aligned} v(A,h,w)^2 &= E[\xi_{i,t}(A,h)^2(w'u_t)^2] \\ &= E\left[(\beta_i(A,h-1)'u_{t+1} + \ldots + \beta_i(A,0)'u_{t+h})^2 u_t^2\right] \\ &= E\left[\left(\sum_{\ell=1}^{h}\sum_{m=1}^{h}\Big(\beta_i(A,h-\ell)'u_{t+\ell}u_{t+m}'\beta_i(A,h-m)\Big)\right)(w'u_t)^2\right]. \end{aligned}$$

Assumption 1 implies that the last expression above equals

$$\sum_{\ell=1}^{h} E\left((\beta_i(A,h-\ell)'u_{t+\ell})^2\,(w'u_t)^2\right). \tag{S3}$$

An application of Cauchy-Schwarz gives the upper bound

$$\begin{aligned} v(A,h,w)^2 &\leq \sum_{\ell=1}^{h} E\left((\beta_i(A,h-\ell)'u_{t+\ell})^4\right)^{1/2} E\left((w'u_t)^4\right)^{1/2}. \\ &\leq \sum_{\ell=1}^{h} \|\beta_i(A,h-\ell)\|^2\, E\left(\|u_{t+\ell}\|^4\right)^{1/2} \|w\|^2\, E\left(\|u_t\|^4\right)^{1/2} \\ &= E\left(\|u_t\|^4\right)\|w\|^2\left(\sum_{\ell=0}^{h-1}\|\beta_i(A,\ell)\|^2\right), \end{aligned}$$

where the last line follows from stationarity.

For the lower bound, re-write expression (S3) 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 ω_1, ω_2 are vectors of unit norm.

By Assumption 2(i),

$$\begin{aligned} E\left((\omega_1'u_{t+\ell})^2\,(\omega_2'u_t)^2\right) &= E\left[E\Big(\,(\omega_1'u_{t+\ell})^2\;\Big|\;\{u_s\}_{s<t+\ell}\Big)(\omega_2'u_t)^2\right] \\ &\geq \delta E[(\omega_2'u_t)^2] \\ &= \delta\omega_2'E[u_tu_t']\omega_2 \\ &\geq \delta\lambda_{\min}(\Sigma). \end{aligned}$$

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 $h, m = 0, 1, \ldots,$

$$\begin{aligned}\beta_i(A,h+m)' = {}& \beta_i(A,h)'\left(J_1\mathbf{A}^m J_1'\right) \\ & + \sum_{j=2}^{p}\left(\sum_{k=0}^{p-j}\beta_i(A,h-1-k)'A_{j+k}\right)\left(J_{j-1}\mathbf{A}^{m-1}J_1'\right),\end{aligned}$$

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

$$\begin{aligned}\beta(A,h+m) &\equiv J_1\mathbf{A}^{h+m}J_1' \\ &= J_1\mathbf{A}^h\mathbf{A}^m J_1' \\ &= J_1\mathbf{A}^h I_{np}I_{np}'\mathbf{A}^m J_1' \\ &= J_1\mathbf{A}^h[J_1', \ldots, J_p']\begin{bmatrix} J_1 \\ \vdots \\ J_p\end{bmatrix}\mathbf{A}^m J_1' \\ &= \left(J_1\mathbf{A}^h J_1'\right)\left(J_1\mathbf{A}^m J_1'\right) + \sum_{j=2}^{p} J_1\mathbf{A}^h J_j' J_j\mathbf{A}^m J_1' \\ &= \beta(A,h)\beta(A,m) + \sum_{j=2}^{p} J_1\mathbf{A}^h J_j' J_j\mathbf{A}^m J_1'.\end{aligned}$$

The definition of the companion matrix implies

$$J_j\mathbf{A} = J_{j-1}, \quad j = 2, \ldots, p,$$

and

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

Therefore, for $j \leq p$,

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

Thus, we have shown that

$$\begin{aligned}\beta(A, h+m) &= \beta(A,h)\beta(A,m) \\ &\quad + \sum_{j=2}^{p}\left(\left(\sum_{k=0}^{p-j}\beta(A,h-1-k)A_{j+k}\right)\left(J_{j-1}\mathbf{A}^{m-1}J_1'\right)\right).\end{aligned}$$

The lemma follows by selecting the i-th equation of the above system of equations. □

Lemma E.7. *Let $B(L)$ be a lag polynomial of order $p-1$ satisfying $\|\mathbf{B}^\ell\| \leq C(1-\epsilon)^\ell$ 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 $\tilde{C} \equiv C\left(1 + C(p-1)\right)$.

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

$$\beta(B,\ell) = J_1\mathbf{B}^m J_1'$$

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

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

$$\begin{aligned}\|\beta_i(B,h+m)\| &\leq \|\beta_i(B,h)\| \times \|J_1\mathbf{B}^m J_1'\| \\ &\quad + \sum_{j=2}^{p-1}\left(\left(\sum_{k=0}^{p-1-j}\|\beta_i(B,h-1-k)\| \times \|B_{j+k}\|\right)\|J_{j-1}\mathbf{B}^{m-1}J_1'\|\right) \\ &\leq \|\beta_i(B,h)\|\, C(1-\epsilon)^m \\ &\quad + \sum_{j=2}^{p-1}\left(\left(\sum_{k=0}^{p-1-j}\|\beta_i(B,h-1-k)\| \times \|B_{j+k}\|\right) C(1-\epsilon)^{m-1}\right)\end{aligned}$$

$$
\begin{aligned}
&(\text{since } \|J_1 \mathbf{B}^m J_1'\| \le C(1-\epsilon)^m \text{ and } \|J_{j-1}\mathbf{B}^{m-1}J_1'\| \le C(1-\epsilon)^{m-1}\ )\\
&\le C(1-\epsilon)^m \left(\|\beta_i(B,h)\| + \sum_{j=2}^{p-1} \left(\left(\sum_{k=0}^{p-1-j} \|\beta_i(B,h-1-k)\| \times C \right) \right) \right)\\
&(\text{since } \|B_{j+k}\| = \|J_1 \mathbf{B} J_{j+k}'\| \le \|\mathbf{B}\|)\\
&\le C(1-\epsilon)^m \left(\|\beta_i(B,h)\| + C(p-2) \left(\sum_{k=0}^{p-3} \|\beta_i(B,h-1-k)\| \right) \right)\\
&\le (1-\epsilon)^m C\,(1+C(p-2)) \left(\sum_{\ell=0}^{p-2} \|\beta_i(B,h-\ell)\| \right),\\
&\le (1-\epsilon)^m C\,(1+C(p-1)) \left(\sum_{\ell=0}^{p-2} \|\beta_i(B,h-\ell)\| \right).
\end{aligned}
$$

The last step merely ensures that the constant is positive for all $p \ge 1$. Note that, in the case $p = 1$, the sum in the last expression is zero. $\square$

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 $\hat{\eta}_1(A,h) \equiv A'\hat{\beta}_1(h) + \hat{\gamma}_1(h)$ in equation (S1). Since the OLS coefficients $(\hat{\beta}_1(h)', \hat{\eta}_1(A,h)')'$ are a non-singular linear transformation of the OLS coefficients $(\hat{\beta}_1(h)', \hat{\gamma}_1(h)')'$, 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

$$
\begin{aligned}
&\begin{pmatrix} \frac{1}{v(A_T,h_T,w)}[\hat{\beta}_1(h_T) - \beta_1(A_T,h_T)] \\ \frac{1}{v(A_T,h_T,w)} G_T[\hat{\eta}(A_T,h_T) - \eta(A_T,h_T)] \end{pmatrix}\\
&= \begin{pmatrix} \frac{1}{T-h_T}\sum_{t=1}^{T-h_T} u_t u_t' & \frac{1}{T-h}\sum_{t=1}^{T-h_T} u_t X_t' G_T^{-1} \\ \frac{1}{T-h_T}\sum_{t=1}^{T-h_T} G_T^{-1} X_t u_t' & \frac{1}{T-h_T}\sum_{t=1}^{T-h_T} G_T^{-1} X_t X_t' G_T^{-1} \end{pmatrix}^{-1}\\
&\quad \times \begin{pmatrix} \frac{1}{(T-h_T)v(A_T,h_T,w)} \sum_{t=1}^{T-h_T} u_t \xi_{1,t}(A_T,h_T) \\ \frac{1}{(T-h_T)v(A_T,h_T,w)} \sum_{t=1}^{T-h_T} G_T^{-1} X_t \xi_{1,t}(A_T,h_T) \end{pmatrix}
\end{aligned}
\tag{S4}
$$

$$\equiv \hat{M}^{-1} \begin{pmatrix} \hat{m}_1 \\ \hat{m}_2 \end{pmatrix}.$$

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 $\tilde{w} \in \mathbb{R}^n \backslash \{0\}$. $\hat{m}_2$ 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 2×2 block partition of $\hat{M}$ in (S4). 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 Σ 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

$$\begin{aligned}
&(T-h_T)^{1/2}[\hat{A}(h_T) - A_T]G(A_T, T-h_T, \epsilon) \\
&= \left(\frac{1}{(T-h_T)^{1/2}} \sum_{t=1}^{T-h_T} u_t X_t' G_T^{-1} \right) \times \left(\frac{1}{T-h_T} \sum_{t=1}^{T-h_T} G_T^{-1} X_t X_t' G_T^{-1} \right)^{-1}.
\end{aligned}$$

The first factor on the right-hand side above is $O_{P_{A_T}}(1)$ by Lemma E.8 below, while the second factor is also $O_{P_{A_T}}(1)$ 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 $\mathcal{A}(0, C, \epsilon)$ and a sequence $\{h_T\}$ of nonnegative integers satisfying $T - h_T \to \infty$. Then for any $i, j = 1, \dots, n$ and $r = 1, \dots, p$,*

$$\frac{\sum_{t=1}^{T-h_T} u_{i,t} y_{j,t-r}}{(T-h_T)^{1/2} g(\rho_j^*(A, \epsilon), T-h_T)} = O_{P_{A_T}}(1).$$

Proof. Write $g_{j,T} \equiv g(\rho_j^*(A, \epsilon), T-h_T)$ for brevity. Note that $\{u_{i,t} y_{j,t-r}\}_t$ is a martingale difference array with respect to the natural filtration $\tilde{\mathcal{F}}_t = \sigma(u_t, u_{t-1}, \dots)$ under Assumption 1. Thus, the sequence is serially uncorrelated, implying that

$$E\left[\left(\frac{\sum_{t=1}^{T-h_T} u_{i,t} y_{j,t-r}}{(T-h_T)^{1/2} g_{j,T}} \right)^2 \right] = \frac{1}{(T-h_T) g_{j,T}^2} \sum_{t=1}^{T-h_T} E[u_{i,t}^2 y_{j,t-r}^2]$$

$$
\leq \frac{1}{g_{j,T}^2} \times [E(u_{i,t}^4)]^{1/2} \times \max_{1 \leq t \leq T-h_T} E(y_{j,t-1}^4)^{1/2}
$$

$$
= [E(u_{i,t}^4)]^{1/2} \times \left(\frac{\max_{1 \leq t \leq T-h_T} E(y_{j,t-1}^4)}{g_{j,T}^4} \right)^{1/2}
$$

$$
\leq \frac{\sqrt{6} C_1 (E(\|u_t\|^4))^2}{\delta \lambda_{\min}(\Sigma)},
$$

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

$$
E \left[\left(\frac{\sum_{t=1}^{T-h_T} \xi_{i,t}(A_T, h_T) y_{j,t-r}}{(T-h_T) v(A_T, h_T, w) g(\rho_j^*(A_T, \epsilon), T-h_T)} \right)^2 \right] \to 0.
$$

To that end, observe that if $t \geq s + h_T$, then

$$
\begin{aligned}
& E[\xi_{i,t}(A_T, h_T) y_{j,t-r} \xi_{i,s}(\rho_T, h_T) y_{j,s-r}] \\
&= E\left[E(\xi_{i,t}(A_T, h_T) \mid u_t, u_{t-1}, \dots) y_{j,t-r} \xi_s(A_T, h_T) y_{j,s-r} \right] \\
&= 0,
\end{aligned}
$$

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

$$
\begin{aligned}
& E \left[\left(\frac{\sum_{t=1}^{T-h_T} \xi_{i,t}(A_T, h_T) y_{j,t-r}}{(T-h_T) v(A_T, h_T, w) g(\rho_j^*(A_T, \epsilon), T-h_T)} \right)^2 \right] \\
&\leq \sum_{t=1}^{T-h_T} \sum_{s=1}^{T-h_T} \mathbb{1}(|s-t| < h_T) \frac{|E[\xi_{i,t}(A_T, h_T) y_{j,t-r} \xi_{i,s}(A_T, h_T) y_{j,s-r}]|}{(T-h_T)^2 v(A_T, h_T, w)^2 g(\rho_j^*(A_T, \epsilon), T-h_T)^2}.
\end{aligned} \tag{S5}
$$

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,

$$
\begin{aligned}
&E[\xi_{i,t}(A_T,h_T)y_{j,t-r}\xi_{i,s}(A_T,h_T)y_{j,s-r}] \\
&= E[\xi_{i,t}(A_T,h_T)\xi_{j,0}(A_T,t-r)\xi_{i,s}(A_T,h_T)\xi_{0,j}(A_T,t-r)] \\
&= \sum_{\ell_1=1}^{h_T}\sum_{\ell_2=1}^{h_T}\sum_{m_1=r}^{t-1}\sum_{m_2=r}^{s-1} E\Big[(\beta_i(A_T,h_T-\ell_1)'u_{t+\ell_1})\,(\beta_j(A_T,m_1-r)'u_{t-m_1}) \\
&\qquad\qquad \times (\beta_i(A_T,h_T-\ell_2)'u_{s+\ell_2})(\beta_j(A_T,m_2-r)'u_{s-m_2})\Big].
\end{aligned}
$$

Consider any summand above defined by its indices (ℓ_1,ℓ_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 $\ell_1\le h_T+s-t$. 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 $m_1\ge t-s+1$. Thus,

$$
\begin{aligned}
&|E[\xi_{i,t}(A_T,h_T)y_{j,t-r}\xi_{i,s}(A_T,h_T)y_{j,s-r}]| \\
&\le \sum_{\ell_1=1}^{h_T+s-t}\sum_{m_1=t-s+r}^{t-r} E\left[|(\beta_i(A_T,h_T-\ell_1)'u_{t+\ell_1})(\beta_i(A_T,h_T-\ell_1-(t-s))'u_{t+\ell_1})\right. \\
&\qquad\qquad \left.\times (\beta_j(A_T,m_1-r)'u_{m_1-r})(\beta_j(A_T,m_1-r-(t-s))'u_{m_1-r})|\right] \\
&= \sum_{\ell_1=1}^{h_T+s-t}\sum_{m_1=t-s+r}^{t-r} \|\beta_i(A_T,h_T-\ell_1)\|\times\|\beta_i(A_T,h_T-\ell_1-(t-s))\|\times\|\beta_j(A_T,m_1-r)\| \\
&\qquad\qquad \times\|\beta_j(A_T,m_1-r-(t-s))\|\times E\left[\|u_{t+\ell_1}\|^2\times\|u_{m_1-r}\|^2\right]
\end{aligned}
$$

(by Cauchy-Schwarz)

$$
\begin{aligned}
&\le C_1^2E(\|u_0\|^4)\sum_{\ell_1=1}^{h_T+s-t}\sum_{m_1=t-s+r}^{t-r}\|\beta_i(A_T,h_T-\ell_1)\|\times\|\beta_i(A_T,h_T-\ell_1-(t-s))\| \\
&\qquad\qquad \times\rho_j^*(A_T,\epsilon)^{2(m_1-r)-(t-s)}
\end{aligned}
$$

(since $\|\beta_j(A_T,h)\|\le C_1\rho_j^*(A_T,\epsilon)^h$ for any j,h by Lemma E.4)

$$
\begin{aligned}
&\le C_1^2\times E(\|u_0\|^4)\times\rho_j^*(A_T,\epsilon)^{(t-s)}\left(\sum_{\ell_1=1}^{h_T+s-t}\|\beta_i(A_T,h_T-\ell_1)\|\times\|\beta_i(A_T,h_T-\ell_1-(t-s))\|\right) \\
&\qquad \times\left(\sum_{m_1=t-s+r}^{t-r}\rho_j^*(A_T,\epsilon)^{2[m_1-r-(t-s)]}\right) \\
&\le E(\|u_0\|^4)\times\rho_j^*(A,\epsilon)^{(t-s)}\left(\sum_{\ell_1=1}^{h_T+s-t}B_i^p(A_T,h_T-\ell_1-(t-s))\rho_i^*(A,\epsilon)^{(t-s)}\right)
\end{aligned}
$$

$$\times \left(\sum_{m_1=t-s+r}^{t-r} \rho_j^*(A_T,\epsilon)^{2[m_1-r-(t-s)]} \right)$$

(using Lemma E.4 and the definition of $B_i^p(A_T, h_T - \ell_1 - (t-s))$ in Lemma E.9 below)

$$= E(\|u_0\|^4) \times \rho_j^*(A_T,\epsilon)^{(t-s)} \rho_i^*(A_T,\epsilon)^{(t-s)} \left(\sum_{\ell=0}^{h_T-1-(t-s)} B_i^p(A_T,\ell) \right) \left(\sum_{m=0}^{s-2r} \rho_j^*(A_T,\epsilon)^{2m} \right)$$

$$\leq E(\|u_0\|^4) \times \rho_j^*(A_T,\epsilon)^{(t-s)} \rho_i^*(A_T,\epsilon)^{(t-s)} \left(\sum_{\ell=0}^{h_T-1} B_i^p(A_T,\ell) \right) \left(\sum_{m=0}^{T-h_T} \rho_j^*(A,\epsilon)^{2m} \right)$$

$$\leq E(\|u_0\|^4) \times \rho_j^*(A_T,\epsilon)^{(t-s)} \rho_i^*(A_T,\epsilon)^{(t-s)} \left(\sum_{\ell=0}^{h_T-1} B_i^p(A_T,\ell) \right) g(\rho_j^*(A_T,\epsilon), T-h_T)^2$$

$$\leq E(\|u_0\|^4) \times \rho_j^*(A_T,\epsilon)^{(t-s)} \rho_i^*(A_T,\epsilon)^{(t-s)}$$

$$\times C_2 p \left(\sum_{\ell=0}^{h_T-1} \|\beta_i(A_T,\ell)\|^2 \right) g(\rho_j^*(A_T,\epsilon), T-h_T)^2$$

(by Lemma E.9 below).

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 (S5), we get

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

$$\leq C_2 p \times \frac{E(\|u_0\|^4)}{(T-h_T)^2}$$

$$\times \frac{\sum_{\ell=0}^{h_T-1} \|\beta_i(A_T,\ell)\|^2}{v(A_T,h_T,w)^2} \sum_{t=1}^{T-h_T} \sum_{s=1}^{T-h_T} \mathbb{1}(|s-t| < h_T) \left(\rho_j^*(A_T,\epsilon) \rho_i^*(A_T,\epsilon) \right)^{|t-s|}$$

$$\leq \frac{C_2 p}{\|w\|^2 \times \delta \times \lambda_{\min}(\Sigma)} \times \frac{E(\|u_0\|^4)}{(T-h_T)^2} \times \sum_{t=1}^{T-h_T} \sum_{s=1}^{T-h_T} \mathbb{1}(|s-t| < h_T) \left(\rho_j^*(A_T,\epsilon) \rho_i^*(A_T,\epsilon) \right)^{|t-s|}$$

(where we have used the lower bound of Lemma E.5)

$$= \frac{C_2 p}{\|w\|^2 \times \delta \times \lambda_{\min}(\Sigma)} \times \frac{E(\|u_0\|^4)}{(T-h_T)} \times \sum_{|m|<h_T} \left(1 - \frac{|m|}{T-h_T} \right) \left(\rho_j^*(A_T,\epsilon) \rho_i^*(A_T,\epsilon) \right)^{|m|}$$

$$\leq \frac{C_2 p}{\|w\|^2 \times \delta \times \lambda_{\min}(\Sigma)} \times \frac{E(\|u_0\|^4)}{(T-h_T)} \times \sum_{m=0}^{h_T-1} \left(\rho_j^*(A_T) \rho_i^*(A_T,\epsilon) \right)^{m}$$

$$\leq \frac{C_2 p}{\|w\|^2 \times \delta \times \lambda_{\min}(\Sigma)} \times \frac{E(\|u_0\|^4)}{(T-h_T)} \times \left(\sum_{m=0}^{h_T-1} \rho_j^*(A_T,\epsilon)^{2m}\right)^{1/2} \left(\sum_{m=0}^{h_T-1} \rho_i^*(A_T,\epsilon)^{2m}\right)^{1/2}$$

(by Cauchy-Schwarz)

$$\leq \frac{C_2 p \times E(\|u_0\|^4)}{\|w\|^2 \times \delta \times \lambda_{\min}(\Sigma)} \times \left(\frac{g(\rho_i^*(A_T,\epsilon), T-h_T)}{T-h_T}\right)^{1/2} \left(\frac{g(\rho_j^*(A_T,\epsilon), T-h_T)}{T-h_T}\right)^{1/2}$$

$$\to 0.$$

□

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

$$\frac{\sum_{\ell=0}^{h-1} B_i^p(A,\ell)}{\sum_{\ell=0}^{h-1} \|\beta_i(A,\ell)\|^2} \leq C_2 p,$$

where

$$B_i^p(A,\ell) \equiv C_2 \sum_{b=0}^{p-1} \left(\|\beta_i(A,\ell)\| \times \|\beta_i(A,\ell-b)\|\right),$$

and we define $\beta_i(A,\ell) = 0$ whenever $\ell < 0$. Here C_2 is the constant defined in Lemma E.4.

Proof. Changing the order of summation, we have

$$\sum_{\ell=0}^{h-1}\left(\sum_{b=0}^{p-1} \|\beta_i(A,\ell)\| \times \|\beta_i(A,\ell-b)\|\right)$$
$$= \sum_{b=0}^{p-1}\left(\sum_{\ell=0}^{h-1} \|\beta_i(A,\ell)\| \times \|\beta_i(A,\ell-b)\|\right)$$
$$\leq \sum_{b=0}^{p-1}\left(\sum_{\ell=0}^{h-1} \|\beta_i(A,\ell)\|^2\right)^{1/2} \times \left(\sum_{\ell=0}^{h-1} \|\beta_i(A,\ell-b)\|^2\right)^{1/2}$$
$$\leq \sum_{b=0}^{p-1}\left(\sum_{\ell=0}^{h-1} \|\beta_i(A,\ell)\|^2\right)$$

(since $\|\beta_i(A,\ell-b)\| = 0$ for $\ell - b < 0$)

$$= p\left(\sum_{\ell=0}^{h-1} \|\beta_i(A,\ell)\|^2\right).$$

Therefore,

$$\sum_{\ell=0}^{h-1} B_i^p(A,\ell) \leq C_2 p\left(\sum_{\ell=0}^{h-1} \|\beta_i(A,\ell)\|^2\right).$$

□

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

$$\begin{aligned}
&\left\| \hat{\Sigma}(h_T) - \frac{1}{T-h_T} \sum_{t=1}^{T-h_T} u_t u_t' \right\| \\
&\leq \frac{1}{T-h_T} \sum_{t=1}^{T-h_T} \|\hat{u}_t(h_T)\hat{u}_t(h_T)' - u_t u_t'\| \\
&\leq \frac{1}{T-h_T} \sum_{t=1}^{T-h_T} \|\hat{u}_t(h_T) - u_t\|^2 + \frac{2}{T-h_T} \sum_{t=1}^{T-h_T} \|(\hat{u}_t(h_T) - u_t)u_t'\| \\
&\leq \|G(A_T, T-h_T, \epsilon)(\hat{A}(h_T) - A_T)\|^2 \times \frac{1}{T-h_T} \sum_{t=1}^{T-h_T} \|G(A_T, T-h_T, \epsilon)^{-1} X_t\|^2 \\
&\qquad + 2 \times \|G(A_T, T-h_T, \epsilon)(\hat{A}(h_T) - A_T)\| \times \frac{1}{T-h_T} \sum_{t=1}^{T-h_T} \|G(A_T, T-h_T, \epsilon)^{-1} X_t u_t'\|.
\end{aligned}$$

Lemma E.3, Lemma A.3(iii), Lemma E.8, and an application of Markov's inequality imply that the last expression above is

$$o_{P_{A_T}}(1) \times O_{P_{A_T}}(1) + 2 \times o_{P_{A_T}}(1) \times o_{P_{A_T}}(1) = o_{P_{A_T}}(1). \qquad \square$$

E.6 Proof of Lemma A.6

We would like to show $\hat{\varsigma} \xrightarrow[P_{A_T}]{p} 1$, 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

$$
\begin{aligned}
\operatorname{Var}(\hat{\varsigma}) &= \frac{1}{(T-h_T)^2 v(A_T,h_T,w)^4} \sum_{t=1}^{T-h_T} \sum_{s=1}^{T-h_T} \operatorname{Cov}\left(\xi_{i,t}(A_T,h_T)^2(w'u_t)^2, \xi_{i,s}(A_T,h_T)^2(w'u_s)^2\right) \\
&= \frac{1}{(T-h_T)v(A_T,h_T,w)^4} \\
&\quad \times \sum_{|m|<T-h_T} \left(1-\frac{|m|}{T-h_T}\right) \operatorname{Cov}\left(\xi_{i,0}(A_T,h_T)^2(w'u_0)^2, \xi_{i,m}(A_T,h_T)^2(w'u_m)^2\right) \\
&\le \frac{2}{(T-h_T)v(A_T,h_T,w)^4} \sum_{m=0}^{T-h_T} |\Gamma_T(m)|, \qquad \text{(S6)}
\end{aligned}
$$

where we define

$$
\Gamma_T(m) \equiv \operatorname{Cov}\left(\xi_{i,0}(A_T,h_T)^2(w'u_{i,0})^2, \xi_{i,m}(A_T,h_T)^2(w'u_m)^2\right), \quad m=0,1,2,\dots
$$

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

$$
\begin{aligned}
\Gamma_T(m) = \sum_{\ell_1=1}^{h_T}\sum_{\ell_2=1}^{h_T}\sum_{\ell_3=1}^{h_T}\sum_{\ell_4=1}^{h_T} \operatorname{Cov}\Big(&(\beta_i(A_T,h_T-\ell_1)'u_{\ell_1})(\beta_i(A_T,h_T-\ell_2)'u_{\ell_2})(w'u_0)^2, \\
&(\beta_i(A_T,h_T-\ell_3)'u_{m+\ell_3})(\beta_i(A_T,h_T-\ell_4)'u_{m+\ell_4})(w'u_m)^2\Big).
\end{aligned}
$$

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 ℓ_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

$$
\begin{aligned}
&|\Gamma_T(m)| \\
&\le \sum_{\ell_1=1}^{h_T}\sum_{\ell_3=1}^{h_T} \left|\operatorname{Cov}\left((\beta_i(A_T,h_T-\ell_1)'u_{m+\ell_1})^2(w'u_m)^2, (\beta_i(A_T,h_T-\ell_3)'u_{\ell_3})^2(w'u_0)^2\right)\right| \qquad \text{(S7)} \\
&\quad + \mathbb{1}(m<h_T)2\sum_{\ell_1=1}^{h_T}\sum_{\ell_2\neq\ell_1} \Big|\operatorname{Cov}\Big((\beta_i(A_T,h_T-\ell_1)'u_{\ell_1})(\beta_i(A_T,h_T-\ell_2)'u_{\ell_2})(w'u_m)^2, \\
&\qquad (\beta_i(A_T,h_T-(\ell_1-m))'u_{\ell_1})(\beta_i(A_T,h_T-(\ell_2-m))'u_{\ell_2})(w'u_0)^2\Big)\Big|. \qquad \text{(S8)}
\end{aligned}
$$

Let $\tilde{\Gamma}_{1,T}(m)$ and $\tilde{\Gamma}_{2,T}(m)$ denote expressions (S7) and (S8), respectively. We will now bound $\sum_{m=0}^{T-h_T} \tilde{\Gamma}_{1,T}(m)$ and $\sum_{m=0}^{T-h_T} \tilde{\Gamma}_{2,T}(m)$, so that we can ultimately insert these bounds into (S6).

Bound on $\sum_{m=0}^{T-h_T} \tilde{\Gamma}_{1,T}(m)$. We first bound the term in expression (S7). 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

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

is bounded above by

$$\|w\|^4 C_1^4 \rho_i^*(A_T,\epsilon)^{2(h_T-\ell_1)+2(h_T-\ell_3)} \left|\mathrm{Cov}\left((\omega_{A_T,h_T,\ell_1}' u_{m+\ell_1})^2 (\omega_w' u_m)^2, (\omega_{A_T,h_T,\ell_3}' u_{\ell_3})^2 (\omega_w' u_0)^2\right)\right|.$$

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

$$\begin{aligned}
&\sum_{m=0}^{T-h_T} \tilde{\Gamma}_{1,T}(m) \\
&\leq \|w\|^4 C_1^4 \sum_{m=0}^{T-h_T} \sum_{\ell_1=1}^{h_T} \sum_{\ell_3=1}^{h_T} \rho_i^*(A_T,\epsilon)^{2(h_T-\ell_3)} \\
&\qquad\qquad \times \left|\mathrm{Cov}\left((\omega_{A_T,h_T,\ell_1}' u_{m+\ell_1})^2 (\omega_w' u_m)^2, (\omega_{A_T,h_T,\ell_3}' u_{\ell_3})^2 (\omega_w' u_0)^2\right)\right| \\
&\leq \|w\|^4 C_1^4 \sum_{b_1=1}^{h_T} \rho_i^*(A_T,\epsilon)^{2(h_T-b_1)} \\
&\qquad \times \left(\sum_{b_2=-\infty}^{\infty} \sum_{b_3=-\infty}^{\infty} \sup_{\|\omega_j\|=1} \left|\mathrm{Cov}\left((\omega_1{}' u_{b_1})^2 (\omega_2' u_0)^2, (\omega_3' u_{b_3+b_2})^2 (\omega_4' u_{b_3})^2\right)\right|\right).
\end{aligned} \tag{S9}$$

Consider the double sum in large parentheses above. If we expand the various squares of the form $(\omega_j' u_t)^2$, then the double sum can be bounded above by at most $4n^2$ terms of the form

$$\sum_{b_2=-\infty}^{\infty} \sum_{b_3=-\infty}^{\infty} \left|\mathrm{Cov}\left(\tilde{u}_{j_1,b_1} \tilde{u}_{j_2,0}, \tilde{u}_{j_3,b_3+b_2} \tilde{u}_{j_4,b_3}\right)\right|, \tag{S10}$$

where $\tilde{u}_t = (\tilde{u}_{1,t}, \dots, \tilde{u}_{n^2,t})' \equiv u_t \otimes u_t$, and $j_1, j_2, j_3, j_4 \in \{1, 2, \dots, n^2\}$ are summation indices. By Assumption 2(ii), the process $\{\tilde{u}_t\}$ 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 (S9) is bounded above by K.[3] Consequently,

$$\sum_{m=0}^{T-h_T} \tilde{\Gamma}_{1,T}(m) \leq \|w\|^4 C_1^4 K \sum_{b_1=1}^{h_T} \rho_i^*(A_T,\epsilon)^{2(h_T-b_1)} = \|w\|^4 C_1^4 K \sum_{\ell=0}^{h_T-1} \rho_i^*(A_T,\epsilon)^{2\ell}.$$

Bound on $\sum_{m=0}^{T-h_T} \tilde{\Gamma}_{2,T}(m)$. Expression (S8) can be bounded above by

$$\begin{aligned}\mathbb{1}(m<h_T)2\sum_{\ell_1=1}^{h_T}\sum_{\ell_2\neq\ell_1} E\Big[\, &|\beta_i(A_T,h_T-\ell_1)'u_{\ell_1}| \times |\beta_i(A_T,h_T-\ell_2)'u_{\ell_2}| \times (w'u_m)^2 \\ &|\beta_i(A_T,h_T-(\ell_1-m))'u_{\ell_1}| \times |\beta_i(A_T,h_T-(\ell_2-m))'u_{\ell_2}| \times (w'u_0)^2\Big].\end{aligned}$$

Applying Cauchy-Schwarz, we get the upper bound

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

---

[3]According to Brillinger (2001, Thm. 2.3.2),

$$\begin{aligned}\mathrm{Cov}\left(\tilde{u}_{j_1,b_1}\tilde{u}_{j_2,0},\tilde{u}_{j_3,b_2}\tilde{u}_{j_4,b_3}\right) = {}&\mathrm{Cov}\left(\tilde{u}_{j_2,0},\tilde{u}_{j_3,b_2}\right)\mathrm{Cov}\left(\tilde{u}_{j_1,b_1},\tilde{u}_{j_4,b_3}\right)+\mathrm{Cov}\left(\tilde{u}_{j_2,0},\tilde{u}_{j_4,b_3}\right)\mathrm{Cov}\left(\tilde{u}_{j_1,b_1},\tilde{u}_{j_3,b_2}\right)\\ &+\mathrm{Cum}\left(\tilde{u}_{j_2,0},\tilde{u}_{j_1,b_1},\tilde{u}_{j_3,b_2},\tilde{u}_{j_4,b_3}\right),\end{aligned}$$

where "Cum" denotes the joint fourth-order cumulant. Thus, the expression (S10) is bounded above by

$$\begin{aligned}&\left(\sum_{b_2=-\infty}^{\infty}\left|\mathrm{Cov}\left(\tilde{u}_{j_2,0},\tilde{u}_{j_3,b_2}\right)\right|\right)\left(\sum_{b_3=-\infty}^{\infty}\left|\mathrm{Cov}\left(\tilde{u}_{j_1,b_1},\tilde{u}_{j_4,b_3}\right)\right|\right)\\ &+\left(\sum_{b_2=-\infty}^{\infty}\left|\mathrm{Cov}\left(\tilde{u}_{j_1,b_1},\tilde{u}_{j_3,b_2}\right)\right|\right)\left(\sum_{b_3=-\infty}^{\infty}\left|\mathrm{Cov}\left(\tilde{u}_{j_2,0},\tilde{u}_{j_4,b_3}\right)\right|\right)\\ &+\sum_{b_1=-\infty}^{\infty}\sum_{b_2=-\infty}^{\infty}\sum_{b_3=-\infty}^{\infty}\left|\mathrm{Cum}\left(\tilde{u}_{j_2,0},\tilde{u}_{j_1,b_1},\tilde{u}_{j_3,b_2},\tilde{u}_{j_4,b_3}\right)\right|.\end{aligned}$$

The third term above is finite by Assumption 2(ii), since $\tilde{u}_t \equiv u_t \otimes u_t$ has absolutely summable cumulants up to order 4. Consider the first term above (the second term is handled similarly). The stationarity of $\tilde{u}_t$ implies that this term equals $\left(\sum_{b_2=-\infty}^{\infty}|\mathrm{Cov}(\tilde{u}_{j_2,0},\tilde{u}_{j_3,b_2})|\right)\left(\sum_{\ell=-\infty}^{\infty}|\mathrm{Cov}(\tilde{u}_{j_1,0},\tilde{u}_{j_4,\ell})|\right)$. By Assumption 2(ii), the autocovariances of $\{\tilde{u}_t\}$ 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 (S10). Picking the largest constant over all summation indices gives the desired result.

Another application of the Cauchy-Schwarz inequality gives

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

Thus,

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

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

$$\underbrace{C_2 \sum_{b=0}^{p-1} \|\beta_i(A_T, h_T - \ell_1)\| \times \|\beta_i(A_T, h_T - \ell_1 - b)\|}_{\equiv B_i^p(A_T, h_T - \ell_1)} \times \rho_i^*(A_T, \epsilon)^m, \tag{S11}$$

for a positive constant C_2 that depends on p and ϵ. Thus,

$$\begin{aligned}
&\sum_{m=0}^{T-h_T} \tilde{\Gamma}_{2,T}(m) \\
&\leq 2 \times E[\, \|u_t^8\|\,] \times \|w\|^4 \times \sum_{m=0}^{h_T-1} \sum_{\ell_1=1}^{h_T} \sum_{\ell_2=1}^{h_T} \left(B_i^p(A_T, h_T - \ell_1) \times B_i^p(A_T, h_T - \ell_2) \times \rho_i^*(A_T, \epsilon)^{2m}\right) \\
&= 2 \times E[\, \|u_t^8\|\,] \times \|w\|^4 \left(\sum_{\ell=0}^{h_T-1} \rho_i^*(A_T, \epsilon)^{2\ell}\right) \left(\sum_{\ell=0}^{h_T-1} B_i^p(A_T, \ell)\right)^2. \qquad \text{(S12)}
\end{aligned}$$

Conclusion of proof. Putting together (S6), (S7), (S8), and (S12), we get

$$\mathrm{Var}(\hat{\varsigma}) \leq \frac{2\|w\|^4}{(T - h_T) v(A_T, h_T, w)^4} \left\{ C_1^4 K \sum_{\ell=0}^{h_T-1} \rho_i^*(A_T, \epsilon)^{2\ell} \right.$$

$$+ 2 \times E[\,\|u_t^8\|\,] \times \left(\sum_{\ell=0}^{h_T-1} \rho_i^*(A_T,\epsilon)^{2\ell} \right) \left(\sum_{\ell=0}^{h_T-1} B_i^p(A_T,\ell) \right)^2 \Bigg\}$$

$$\leq \left\{ \frac{2C_1^4 K \times \sum_{\ell=0}^{h_T-1} \rho_i^*(A_T,\epsilon)^{2\ell}}{(T-h_T)\left(\sum_{\ell=0}^{h_T-1} \|\beta_i(A,\ell)\|^2\right)^2 \delta^2 \lambda_{\min}(\Sigma)^2} \right.$$

$$\left. + \frac{2 \times E[\,\|u_t^8\|\,] \times \sum_{\ell=0}^{h_T-1} \rho_i^*(A_T,\epsilon)^{2\ell}}{(T-h_T)\delta^2 \lambda_{\min}(\Sigma)^2} \times \frac{\left(\sum_{\ell=0}^{h_T-1} B_i^p(A_T,\ell)\right)^2}{\left(\sum_{\ell=0}^{h_T-1} \|\beta_i(A,\ell)\|^2\right)^2} \right\}$$

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

$$\leq \frac{2\left\{\left(C_1^4 K \times \sum_{\ell=0}^{h_T-1} \rho_i^*(A_T,\epsilon)^{2\ell}\right) + \left(2 \times E[\,\|u_t^8\|\,] \times C_2 p \times \sum_{\ell=0}^{h_T-1} \rho_i^*(A_T,\epsilon)^{2\ell}\right)\right\}}{(T-h_T)\delta^2 \lambda_{\min}(\Sigma)^2}$$

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

$$= \frac{(2 \times C_1^4 K) + (4 \times E[\,\|u_t^8\|\,] \times C_2 p)}{\delta^2 \lambda_{\min}(\Sigma)^2} \times \frac{\sum_{\ell=0}^{h_T-1} \rho_i^*(A_T,\epsilon)^{2\ell}}{T-h_T}.$$

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, $\mathrm{Var}(\hat{\varsigma}) \to 0$. □

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 $\xi_{i,t}(A,h)^4$, we obtain

$$\begin{aligned} & E[\xi_{i,t}(A,h)^4 (a'u_t)^4] \\ &= \sum_{\ell_1=1}^{h}\sum_{\ell_2=1}^{h}\sum_{\ell_3=1}^{h}\sum_{\ell_4=1}^{h} \|\beta_i(h-\ell_1)\| \times \|\beta_i(h-\ell_2)\| \times \|\beta_i(h-\ell_3)\| \times \|\beta_i(h-\ell_4)\| \\ &\qquad \times E\left[(\omega'_{A,h,\ell_1}u_{t+\ell_1}) \times (\omega'_{A,h,\ell_2}u_{t+\ell_2}) \times (\omega'_{A,h,\ell_3}u_{t+\ell_3}) \times (\omega'_{A,h,\ell_4}u_{t+\ell_4}) \times (w'u_t)^4\right]. \end{aligned} \quad \text{(S13)}$$

By Assumption 1, the summands above equal zero if one of the indices ℓ_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

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

where $\ell, m \in \{1, \dots, h\}$. For given ℓ 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 (S13) of this form. Thus,

$$\begin{aligned}
E[\xi_{i,t}(A,h)^4 (w'u_t)^4] &= 6 \sum_{\ell=1}^{h} \sum_{m=1}^{h} \Big(\|\beta_i(h-\ell)\|^2 \|\beta_i(h-m)\|^2 \\
&\qquad \times E\left[(\omega'_{A,h,\ell} u_{t+\ell})^2 \times (\omega'_{A,h,m} u_{t+m})^2 \times (w' u_t)^4\right] \Big) \\
&\leq 6 \|w\|^4 E(\|u_t\|^8) \sum_{\ell=1}^{h} \sum_{m=1}^{h} \|\beta_i(h-\ell)\|^2 \|\beta_i(h-m)\|^2 \\
&\quad \text{(by applying Cauchy-Schwarz twice)} \\
&= 6 \|w\|^4 E(\|u_t\|^8) \left(\sum_{\ell=0}^{h-1} \|\beta_i(A, h-\ell)\|^2 \right)^2 .
\end{aligned}$$

It follows from Lemma E.5 that

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

References

Brillinger, D. R. (2001): *Time Series: Data Analysis and Theory*, Classics in Applied Mathematics, SIAM.

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

Gertler, M. and P. Karadi (2015): “Monetary Policy Surprises, Credit Costs, and Economic Activity,” *American Economic Journal: Macroeconomics*, 7, 44–76.

Inoue, A. and L. Kilian (2020): “The uniform validity of impulse response inference in autoregressions,” *Journal of Econometrics*, 215, 450–472.

Kilian, L. and Y. J. Kim (2011): “How Reliable Are Local Projection Estimators of Impulse Responses?” *Review of Economics and Statistics*, 93, 1460–1466.