Supplement to
“Accurate and (Almost) Tuning Parameter Free Inference in Cointegrating Regressions”

Karsten Reichold and Carsten Jentsch

May 7, 2021

Equation numbers not preceded by “D.” refer to the main article.

C Asymptotic Critical Values in the Presence of Deterministic Regressors

Table C.1: Asymptotic critical values for $\tau_{\mathrm{IM}}(\hat{\eta}_T)$ in the presence of deterministic regressors d_t in (2.1)

	$m=1$	$m=2$		$m=3$			$m=4$			
%	$s=1$	$s=1$	$s=2$	$s=1$	$s=2$	$s=3$	$s=1$	$s=2$	$s=3$	$s=4$
intercept ($d_t = 1$)										
90.0	64.13	94.15	168.58	126.65	221.45	305.36	162.08	278.05	382.30	481.15
95.0	95.81	140.55	233.15	187.03	297.11	396.56	236.54	372.79	487.71	596.15
97.5	136.10	190.23	292.64	245.93	375.55	488.35	325.56	458.37	587.85	720.31
99.0	187.13	263.92	381.78	338.59	474.31	602.27	421.68	582.89	719.98	872.07
intercept and linear time trend ($d_t = [1, t]'$)										
90.0	90.44	122.19	209.54	152.66	261.47	363.17	180.25	311.22	434.29	545.37
95.0	134.19	171.46	283.33	219.51	354.08	460.37	258.75	423.39	546.31	688.21
97.5	183.51	231.09	357.66	294.26	433.33	569.84	342.56	524.25	686.44	810.09
99.0	243.72	304.08	460.98	409.03	556.42	713.24	478.05	680.76	821.12	977.09
intercept, linear and square time trend ($d_t = [1, t, t^2]'$)										
90.0	115.13	138.49	245.91	175.40	302.95	418.77	205.72	352.96	479.57	608.35
95.0	166.35	200.65	331.26	255.74	402.90	530.94	303.58	465.28	621.70	764.20
97.5	217.42	268.86	401.63	348.51	509.51	637.29	390.59	589.05	762.57	902.89
99.0	290.63	357.58	513.85	472.49	646.48	800.81	527.20	754.26	923.59	1070.32
intercept, linear, square and cubic time trend ($d_t = [1, t, t^2, t^3]'$)										
90.0	137.70	166.87	292.13	197.84	340.61	465.58	229.38	392.80	533.60	680.84
95.0	198.48	237.82	379.15	288.65	446.27	590.05	334.55	509.11	684.33	858.04
97.5	263.30	308.64	467.71	391.70	565.65	720.19	438.56	645.41	853.07	1004.50
99.0	352.56	406.48	587.03	539.71	726.07	903.53	592.44	846.82	1052.20	1222.78

Notes: s is the number of linearly independent restrictions under the null hypothesis on the coefficients corresponding to the $m \geq s$ integrated regressors.

D Proofs of Auxiliary Results

Proof of Lemma 1. The solution of the sample Yule-Walker equations in the regression of w_t on $w_{t-1}, \dots, w_{t-q}$, $t = q+1, \dots, T$, can be written in compact form as a $((m+1) \times q(m+1))$-dimensional matrix

$$\tilde{\mathbf{\Phi}}(q) := [\tilde{\Phi}_1(q), \dots, \tilde{\Phi}_q(q)] = \tilde{\Gamma}\tilde{G}^{-1},$$

with the $((m+1) \times q(m+1))$-dimensional matrix $\tilde{\Gamma} := \left[\tilde{\Gamma}(1), \dots, \tilde{\Gamma}(q)\right]$, the $(q(m+1) \times q(m+1))$-dimensional matrix $\tilde{G} := \left(\tilde{\Gamma}(s-r)\right)_{r,s=1,\dots,q}$ and the $((m+1) \times (m+1))$-dimensional empirical autocovariance matrix of $w_1, \dots, w_T$ at lag $-q+1 \leq h \leq q$, given by

$$\tilde{\Gamma}(h) := T^{-1} \sum_{t=\max\{1,1-h\}}^{\min\{T,T-h\}} (w_{t+h} - \bar{w}_T)(w_t - \bar{w}_T)',$$

where $\bar{w}_T := T^{-1}\sum_{t=1}^{T} w_t$. Analogously, the solution of the sample Yule-Walker equations in the regression of $\hat{w}_t$ on $\hat{w}_{t-1}, \dots, \hat{w}_{t-q}$, $t = q+1, \dots, T$, can be written in compact form as

$$\hat{\mathbf{\Phi}}(q) := [\hat{\Phi}_1(q), \dots, \hat{\Phi}_q(q)] = \hat{\Gamma}\hat{G}^{-1},$$

with $\hat{\Gamma} := \left[\hat{\Gamma}(1), \dots, \hat{\Gamma}(q)\right]$, $\hat{G} := \left(\hat{\Gamma}(s-r)\right)_{r,s=1,\dots,q}$ and $\hat{\Gamma}(h)$ the empirical autocovariance matrix of $\hat{w}_1, \dots, \hat{w}_T$ at lag $-q+1 \leq h \leq q$. Taking the difference of $\tilde{\mathbf{\Phi}}(q)$ and $\hat{\mathbf{\Phi}}(q)$, adding and subtracting $\hat{\Gamma}\tilde{G}^{-1}$ and using

$$\tilde{G}^{-1} - \hat{G}^{-1} = \tilde{G}^{-1}(\hat{G} - \tilde{G})\hat{G}^{-1}$$

leads to

$$\hat{\mathbf{\Phi}}(q) - \tilde{\mathbf{\Phi}}(q) = \hat{\Gamma}\tilde{G}^{-1}(\tilde{G} - \hat{G})\hat{G}^{-1} - (\tilde{\Gamma} - \hat{\Gamma})\tilde{G}^{-1}$$

Hence, we have to consider $\tilde{G} - \hat{G}$ in more detail ($\tilde{\Gamma} - \hat{\Gamma}$ works similarly). A typical block element of $\tilde{G} - \hat{G}$ is

$$\begin{aligned}\tilde{\Gamma}(h) - \hat{\Gamma}(h) = T^{-1} &\sum_{t=\max\{1,1-h\}}^{\min\{T,T-h\}} (w_{t+h} - \bar{w}_T)(w_t - \bar{w}_T)' \\ &- T^{-1} \sum_{t=\max\{1,1-h\}}^{\min\{T,T-h\}} (\hat{w}_{t+h} - \bar{\hat{w}}_T)(\hat{w}_t - \bar{\hat{w}}_T)' \\ &= T^{-1} \sum_{t=\max\{1,1-h\}}^{\min\{T,T-h\}} (w_{t+h} - \bar{w}_T)(w_t - \hat{w}_t - (\bar{w}_T - \bar{\hat{w}}_T))' \\ &- T^{-1} \sum_{t=\max\{1,1-h\}}^{\min\{T,T-h\}} (\hat{w}_{t+h} - w_{t+h} - (\bar{\hat{w}}_T - \bar{w}_T))(\hat{w}_t - \bar{\hat{w}}_T)' \\ &= A_1(h) - A_2(h),\end{aligned}$$

with an obvious definition for $A_1(h)$ and $A_2(h)$. Let us consider $A_1(h)$ in more detail ($A_2(h)$ works similarly). Using $\hat{w}_t = [\hat{u}_t, v_t']'$ and $\hat{u}_t = y_t - x_t'\hat{\beta}_{\mathrm{IM}}$ together with the model equations (2.1) and (2.2), we get

$$w_t - \hat{w}_t = \begin{bmatrix} u_t - \hat{u}_t \\ 0_{m\times 1} \end{bmatrix} = \begin{bmatrix} y_t - x_t'\beta - (y_t - x_t'\hat{\beta}_{\mathrm{IM}}) \\ 0_{m\times 1} \end{bmatrix} = \begin{bmatrix} x_t'(\hat{\beta}_{\mathrm{IM}} - \beta) \\ 0_{m\times 1} \end{bmatrix}. \qquad \text{(D.1)}$$

Before we continue, note that the last equality implies

$$\max_{1\le t\le T} |\hat{w}_t - w_t|_F \le T^{-1/2} \max_{1\le t\le T} |T^{-1/2} x_t|_F |T\left(\hat{\beta}_{\mathrm{IM}} - \beta\right)|_F = O_{\mathbb{P}}(T^{-1/2}),$$

since $\max_{1\le t\le T} |T^{-1/2}x_t|_F = \sup_{0\le r\le 1} |T^{-1/2}x_{\lfloor rT\rfloor}|_F$ converges by Assumption 3 and the continuous mapping theorem to $\sup_{0\le r\le 1} |B_v(r)|_F = O_{\mathbb{P}}(1)$ and $\hat{\beta}_{\mathrm{IM}}$ is rate-T consistent. This proves (B.1).
We now proceed with the proof of (B.2). From (D.1) we obtain

$$\begin{aligned}A_1(h) = T^{-1} &\sum_{t=\max\{1,1-h\}}^{\min\{T,T-h\}} (w_{t+h} - \bar{w}_T) \begin{bmatrix} (x_t - \bar{x}_T)'(\hat{\beta}_{\mathrm{IM}} - \beta) \\ 0_{m\times 1} \end{bmatrix}' \\ &= T^{-1} \sum_{t=\max\{1,1-h\}}^{\min\{T,T-h\}} (w_{t+h} - \bar{w}_T)[(x_t - \bar{x}_T)'(\hat{\beta}_{\mathrm{IM}} - \beta), 0_{1\times m}] \\ &= \tilde{\Gamma}_{w,x}(h)(\hat{\beta}_{\mathrm{IM}} - \beta)e_1',\end{aligned}$$

where $\tilde{\Gamma}_{w,x}(h) := T^{-1}\sum_{t=\max\{1,1-h\}}^{\min\{T,T-h\}}(w_{t+h} - \bar{w}_T)(x_t - \bar{x}_T)'$ is $((m+1)\times m)$-dimensional and $e_1 \coloneqq (1, 0_{1\times m})'$ is the first $(m+1)$-dimensional unit vector. Denoting the part of $\tilde{G} - \hat{G}$ that consists of block-entries $A_1(h)$ by $\tilde{G}_1 - \hat{G}_1$, we get

$$\tilde{G}_1 - \hat{G}_1 = \left[\tilde{\Gamma}_{w,x}(s-r)(\hat{\beta}_{\mathrm{IM}} - \beta)e_1'\right]_{r,s=1,\dots,q} = \tilde{\Gamma}_{w,x}(I_q \otimes ((\hat{\beta}_{\mathrm{IM}} - \beta)e_1')),$$

where $\tilde{\Gamma}_{w,x} \coloneqq (\tilde{\Gamma}_{w,x}(s-r))_{r,s=1,\dots,q}$ is $(q(m+1) \times qm)$-dimensional. For the second factor we have

$$\begin{aligned}
|I_q \otimes ((\hat{\beta}_{\mathrm{IM}} - \beta)e_1')|_F &= \sqrt{\operatorname{tr}\left((I_q \otimes ((\hat{\beta}_{\mathrm{IM}} - \beta)e_1'))' I_q \otimes ((\hat{\beta}_{\mathrm{IM}} - \beta)e_1')\right)} \\
&= \sqrt{\operatorname{tr}\left(I_q \otimes (e_1(\hat{\beta}_{\mathrm{IM}} - \beta)'(\hat{\beta}_{\mathrm{IM}} - \beta)e_1'\right))} \\
&= \sqrt{\operatorname{tr}\left(I_q \otimes \operatorname{diag}\left(\sum_{i=1}^m (\hat{\beta}_{\mathrm{IM},i} - \beta_i)^2, 0, \dots, 0\right)\right)} \\
&= \sqrt{q \sum_{i=1}^m (\hat{\beta}_{\mathrm{IM},i} - \beta_i)^2} \\
&= q^{1/2} T^{-1} |T(\hat{\beta}_{\mathrm{IM}} - \beta)|_F \\
&= O_P(q^{1/2} T^{-1}).
\end{aligned}$$

Next, let us consider $\tilde{\Gamma}_{w,x}$ in more detail. Recall that $w_t = (u_t, v_t')'$ and $x_t = \sum_{k=1}^t v_k$. To avoid lengthy index notation, w.l.o.g. we can assume that $m = 1$ and consider the second element of w_t only (the first element works similarly). We thus consider the scalar quantity

$$\tilde{\Gamma}_{w,x}(h) = T^{-1} \sum_{t=\max\{1,1-h\}}^{\min\{T,T-h\}} (v_{t+h} - \bar{v}_T)(\sum_{k=1}^t v_k - T^{-1} \sum_{i=1}^T \sum_{j=1}^i v_j).$$

Taking the expectation of the squared Frobenius norm of the corresponding $(q \times q)$-

dimensional matrix $\tilde{\Gamma}_{w,x}$ and combining the sums over r and s, leads to

$$\begin{aligned}
&\mathbb{E}(|\tilde{\Gamma}_{w,x}|_F^2) \\
&= \sum_{r,s=1}^{q} T^{-2} \sum_{t_1,t_2=\max\{1,1-(s-r)\}}^{\min\{T,T-(s-r)\}} \mathbb{E}\left[(v_{t_1+s-r} - \bar{v}_T)(\sum_{k_1=1}^{t_1} v_{k_1} - T^{-1}\sum_{i_1=1}^{T}\sum_{j_1=1}^{i_1} v_{j_1})\right. \\
&\qquad\qquad\qquad \left. \times\ (v_{t_2+s-r} - \bar{v}_T)(\sum_{k_2=1}^{t_2} v_{k_2} - T^{-1}\sum_{i_2=1}^{T}\sum_{j_2=1}^{i_2} v_{j_2})\right] \\
&= T^{-2} \sum_{h=-q+1}^{q-1} (q-|h|) \sum_{t_1,t_2=\max\{1,1-h\}}^{\min\{T,T-h\}} \mathbb{E}\left[(v_{t_1+h} - \bar{v}_T)(\sum_{k_1=1}^{t_1} v_{k_1} - T^{-1}\sum_{i_1=1}^{T}\sum_{j_1=1}^{i_1} v_{j_1})\right. \\
&\qquad\qquad\qquad \left. \times\ (v_{t_2+h} - \bar{v}_T)(\sum_{k_2=1}^{t_2} v_{k_2} - T^{-1}\sum_{i_2=1}^{T}\sum_{j_2=1}^{i_2} v_{j_2})\right].
\end{aligned}$$

Note that the last expectation is of the form $\mathbb{E}(ABCD)$ with $\mathbb{E}(A) = \mathbb{E}(B) = \mathbb{E}(C) = \mathbb{E}(D) = 0$. Hence, by using common rules for joint cumulants of centered random variables (see, e. g., Brillinger, 1981), we get

$$\begin{aligned}
&\mathbb{E}(ABCD) \\
&= \mathrm{cum}(A,B,C,D) + \mathbb{E}(AB)\mathbb{E}(CD) + \mathbb{E}(AC)\mathbb{E}(BD) + \mathbb{E}(AD)\mathbb{E}(BC) \\
&= \mathrm{cum}(A,B,C,D) + \mathrm{Cov}(A,B)\mathrm{Cov}(C,D) + \mathrm{Cov}(A,C)\mathrm{Cov}(B,D) \\
&\quad + \mathrm{Cov}(A,D)\mathrm{Cov}(B,C),
\end{aligned}$$

where $\mathrm{Cov}(\cdot,\cdot)$ denotes the covariance of two random variables. Hence, the first term corresponding to the fourth-order cumulant becomes

$$\begin{aligned}
T^{-2} \sum_{h=-q+1}^{q-1} (q-|h|) \sum_{t_1,t_2=\max\{1,1-h\}}^{\min\{T,T-h\}} \mathrm{cum}\Bigg(& v_{t_1+h} - \bar{v}_T, \sum_{k_1=1}^{t_1} v_{k_1} - T^{-1}\sum_{i_1=1}^{T}\sum_{j_1=1}^{i_1} v_{j_1}, \\
& v_{t_2+h} - \bar{v}_T, \sum_{k_2=1}^{t_2} v_{k_2} - T^{-1}\sum_{i_2=1}^{T}\sum_{j_2=1}^{i_2} v_{j_2}\Bigg)
\end{aligned}$$

leading to $2^4 = 16$ terms when exapnding the cumulant. Exemplarily, for the absolute value of the first one (the others work similarly), we get from the common calculation rules for cumulants

$$
\begin{aligned}
&|T^{-2} \sum_{h=-q+1}^{q-1} (q-|h|) \sum_{t_1,t_2=\max\{1,1-h\}}^{\min\{T,T-h\}} \operatorname{cum}\left(v_{t_1+h}, \sum_{k_1=1}^{t_1} v_{k_1}, v_{t_2+h}, \sum_{k_2=1}^{t_2} v_{k_2} \right)| \\
&\leq T^{-2} \sum_{h=-q+1}^{q-1} |q-|h|| \sum_{t_1,t_2=\max\{1,1-h\}}^{\min\{T,T-h\}} \sum_{k_1=1}^{t_1} \sum_{k_2=1}^{t_2} |\operatorname{cum}\left(v_{t_1+h}, v_{k_1}, v_{t_2+h}, v_{k_2}\right)| \\
&\leq \frac{q}{T^2} \sum_{h=-q+1}^{q-1} \sum_{t_1,t_2=1}^{T} \sum_{k_1,k_2=1}^{T} |\operatorname{cum}\left(v_{t_1+h}, v_{k_1}, v_{t_2+h}, v_{k_2}\right)|.
\end{aligned}
$$

By combining the sums over t_1 and t_2 and those over k_1 and k_2, respectively, the above term becomes

$$
\begin{aligned}
&\frac{q}{T^2} \sum_{h=-q+1}^{q-1} \sum_{l=-(T-1)}^{T-1} \sum_{s=\max\{1,1-l\}}^{\min\{T,T-l\}} \sum_{i=-(T-1)}^{T-1} \sum_{j=\max\{1,1-i\}}^{\min\{T,T-i\}} |\operatorname{cum}\left(v_{s+l+h}, v_{j+i}, v_{s+h}, v_j\right)| \\
&\leq \frac{q}{T^2} \sum_{h=-q+1}^{q-1} \sum_{i,l=-(T-1)}^{T-1} \sum_{s,j=1}^{T} |\operatorname{cum}\left(v_{s+l+h}, v_{j+i}, v_{s+h}, v_j\right)| \\
&\leq \frac{q}{T^2} \sum_{h=-q+1}^{q-1} \sum_{i,l=-(T-1)}^{T-1} \sum_{k=-(T-1)}^{T-1} \sum_{r=\max\{1,1-k\}}^{\min\{T,T-k\}} |\operatorname{cum}\left(v_{r+k+l+h}, v_{r+i}, v_{r+k+h}, v_r\right)| \\
&\leq \frac{q}{T} \sum_{h=-q+1}^{q-1} \sum_{i,l=-(T-1)}^{T-1} \sum_{k=-(T-1)}^{T-1} |\operatorname{cum}\left(v_{k+l+h}, v_i, v_{k+h}, v_0\right)|,
\end{aligned}
$$

where we also combined the sums over s and j and made use of the (strict) stationarity of $\{v_t\}_{t\in\mathbb{Z}}$. Finally, combining the sums over h and k, we get the bound

$$
\begin{aligned}
&\frac{q}{T} \sum_{r=-(T+q-2)}^{T+q-2} (2q-1) \sum_{i,l=-(T-1)}^{T-1} |\operatorname{cum}\left(v_{r+l}, v_i, v_r, v_0\right)| \\
&\leq 2\frac{q^2}{T} \sum_{j=-(2T+q-3)}^{2T+q-3} \sum_{r=-(T+q-2)}^{T+q-2} \sum_{i=-(T-1)}^{T-1} |\operatorname{cum}\left(v_j, v_i, v_r, v_0\right)| \\
&\leq 2\frac{q^2}{T} \sum_{j,i,r=-\infty}^{\infty} |\operatorname{cum}\left(v_j, v_i, v_r, v_0\right)| = O(q^2 T^{-1})
\end{aligned}
$$

due to the summability condition imposed on the fourth order cumulants in Assumption 2 and, hence, vanishes for $T \to \infty$. However, the leading term is the term corresponding to $\mathrm{Cov}(A,B)\mathrm{Cov}(C,D)$. That is, we have to consider

$$T^{-2}\sum_{h=-q+1}^{q-1}(q-|h|)\sum_{t_1,t_2=\max\{1,1-h\}}^{\min\{T,T-h\}}\mathrm{Cov}(v_{t_1+h}-\bar{v}_T,\sum_{k_1=1}^{t_1}v_{k_1}-T^{-1}\sum_{i_1=1}^{T}\sum_{j_1=1}^{i_1}v_{j_1})$$
$$\times\,\mathrm{Cov}(v_{t_2+h}-\bar{v}_T,\sum_{k_2=1}^{t_2}v_{k_2}-T^{-1}\sum_{i_2=1}^{T}\sum_{j_2=1}^{i_2}v_{j_2})$$
$$=\sum_{h=-q+1}^{q-1}(q-|h|)\left(T^{-1}\sum_{t=\max\{1,1-h\}}^{\min\{T,T-h\}}\mathrm{Cov}(v_{t+h}-\bar{v}_T,\sum_{k=1}^{t}v_k-T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{i}v_j)\right)^2.$$

Hence, we have to compute

$$T^{-1}\sum_{t=\max\{1,1-h\}}^{\min\{T,T-h\}}\mathrm{Cov}(v_{t+h}-\bar{v}_T,\sum_{k=1}^{t}v_k-T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{i}v_j).$$

This leads to four terms to consider, which can be treated with similar arguments. For the first term we get

$$T^{-1}\sum_{t=\max\{1,1-h\}}^{\min\{T,T-h\}}\mathrm{Cov}(v_{t+h},\sum_{k=1}^{t}v_k)=T^{-1}\sum_{t=\max\{1,1-h\}}^{\min\{T,T-h\}}\sum_{k=1}^{t}\gamma_v(t+h-k),$$

where $\gamma_v(h)$ is the covariance of the one-dimensional process (still assumed for notational brevity) $\{v_t\}_{t\in\mathbb{Z}}$ at lag h. E. g., for $h\geq 0$, this can be exactly computed to equal

$$T^{-1}\sum_{t=1}^{T-h}\sum_{k=1}^{t}\gamma_v(t+h-k)=T^{-1}\sum_{j=h}^{T-1}(T-j)\gamma_v(j)\leq\sum_{j=h}^{T-1}\gamma_v(j)$$

and its absolute vale can be bounded by $\sum_{j=-\infty}^{\infty}|\gamma_v(j)|<\infty$ due to the second-order cumulant condition imposed in Assumption 2. Similar arguments yield the same bound for $h<0$ and for the other three terms. Hence, the term of $\mathbb{E}(|\tilde{\Gamma}_{w,x}|_F^2)$ that corresponds to $\mathrm{Cov}(A,B)\mathrm{Cov}(C,D)$ is of order $O(q^2)$.

In total we thus have $\mathbb{E}(|\tilde{\Gamma}_{w,x}|_F^2)=O(q^2)$. Note that we have proven this result for $m=1$. However, as m is fixed, the result also holds for $m>1$. Therefore, we obtain for the $(q(m+1)\times qm)$-dimensional matrix $\tilde{\Gamma}_{w,x}$ that $|\tilde{\Gamma}_{w,x}|_F=O_{\mathbb{P}}(q)$. It follows that $|\tilde{G}_1-\hat{G}_1|_F=O_{\mathbb{P}}(q)O_P(q^{1/2}T^{-1})=O_{\mathbb{P}}(q^{3/2}T^{-1})$ and similarly also $|\tilde{G}-\hat{G}|_F=O_{\mathbb{P}}(q)O_{\mathbb{P}}(q^{1/2}T^{-1})=O_{\mathbb{P}}(q^{3/2}T^{-1})$.

Further, we have to consider $\tilde{G}^{-1}$. In the following, let $\mu_{\min}(A)$ and $\mu_{\max}(A)$

denote the smallest and largest eigenvalue of a matrix A, respectively and define $G := (\Gamma(s-r))_{r,s=1,\ldots,q} \in \mathbb{R}^{q(m+1)\times q(m+1)}$. Similar to the above, using the fourth-order cumulant condition from Assumption 2, we can show that $|\tilde{G}-G|_F = O_{\mathbb{P}}(qT^{-1/2}) = o_{\mathbb{P}}(1)$. Then, to show boundedness in probability of $\tilde{G}^{-1}$ (similar for $\hat{G}^{-1}$), for all $\epsilon > 0$, we have to find a $K < \infty$ and a $T_0 < \infty$ both large enough such that for all $T > T_0$, it holds that

$$\mathbb{P}(|\tilde{G}^{-1}|_2 > K) < \epsilon,$$

where $|A|_2$ denotes the spectral norm of a matrix A. Let $\epsilon > 0$. Then, due to positive semi-definiteness of $\tilde{G}$ by construction and invertibility, see Meyer and Kreiss (2015, Lemma 3.4 and Remark 3.2), we have positive definiteness of $\tilde{G}$ and, consequently, of $\tilde{G}^{-1}$. Hence, we get

$$\begin{aligned}\mathbb{P}(|\tilde{G}^{-1}|_2 > K) =& \mathbb{P}(\mu_{\max}(\tilde{G}^{-1}) > K) = \mathbb{P}(\mu_{\min}^{-1}(\tilde{G}) > K)\\ =& \mathbb{P}(\mu_{\min}(\tilde{G}) < \frac{1}{K}, |\tilde{G}-G|_2 \geq \delta)\\ &+ \mathbb{P}(\mu_{\min}(\tilde{G}) < \frac{1}{K}, |\tilde{G}-G|_2 < \delta).\end{aligned}$$

Further, as $|\tilde{G}-G|_2 \leq |\tilde{G}-G|_F = o_{\mathbb{P}}(1)$, for any $\delta > 0$, we can choose T_0 large enough to have $\mathbb{P}(|\tilde{G}-G|_2 \geq \delta) \leq \epsilon$. Then the first term on the last right-hand side can be bounded by $\mathbb{P}(|\tilde{G}-G|_2 \geq \delta) \leq \epsilon$. For the second term, as $\tilde{G}$, G and hence also $\tilde{G}-G$ are symmetric with real-valued entries, these matrices are Hermitian such that Weyl's theorem (see, e. g., Theorem 4.3.1 in Horn and Johnson, 2012) applies, leading to the inequality $\mu_{\min}(G) + \mu_{\min}(\tilde{G}-G) \leq \mu_{\min}(\tilde{G})$. It follows that

$$\begin{aligned}&\mathbb{P}(\mu_{\min}(\tilde{G}) < \frac{1}{K}, |\tilde{G}-G|_2 < \delta)\\ &\leq \mathbb{P}(\mu_{\min}(G) + \mu_{\min}(\tilde{G}-G) < \frac{1}{K}, |\tilde{G}-G|_2 < \delta)\\ &= \mathbb{P}(\mu_{\min}(G) < \frac{1}{K} - \mu_{\min}(\tilde{G}-G), |\tilde{G}-G|_2 < \delta).\end{aligned}$$

From symmetry of $\tilde{G}-G$ we get that the eigenvalues of $(\tilde{G}-G)'(\tilde{G}-G)$ are exactly the squared eigenvalues of $\tilde{G} - G$. Hence, the bound

$$|\tilde{G} - G|_2 = \sqrt{\mu_{\max}((\tilde{G} - G)'(\tilde{G} - G))} < \delta$$

implies also $\mu_{\min}(\tilde{G} - G) \geq -\delta$ such that the last right-hand side can be bounded by

$$\mathbb{P}(\mu_{\min}(G) < \frac{1}{K} + \delta, |\tilde{G} - G|_2 < \delta) \leq \mathbb{P}(\mu_{\min}(G) < \frac{1}{K} + \delta). \tag{D.2}$$

Next, note that $\mu_{\min}(G) \geq \tilde{c}$ for some constant $\tilde{c} > 0$ by Assumption 1. Therefore, the right-hand side in (D.2) becomes zero if we choose $\delta < \tilde{c}/2$ small enough and $K > 2/\tilde{c}$ large enough, such that $\frac{1}{K} + \delta < \tilde{c}$. This completes the proof of $|\tilde{G}^{-1}|_2 = O_{\mathbb{P}}(1)$. Furthermore, from Assumption 1 we also get $|\tilde{G}|_2 = O_{\mathbb{P}}(1)$ and similarly $|\hat{G}|_2 = O_{\mathbb{P}}(1)$ and $|\hat{G}|_2 = O_{\mathbb{P}}(1)$. Altogether, we get

$$\begin{aligned}
&|\tilde{\mathbf{\Phi}}(q) - \hat{\mathbf{\Phi}}(q)|_2 \leq |\tilde{\Gamma} - \hat{\Gamma}|_2 |\tilde{G}^{-1}|_2 + |\hat{\Gamma}|_2 |\tilde{G}^{-1}|_2 |\tilde{G} - \hat{G}|_2 |\hat{G}^{-1}|_2 \\
&\leq |\tilde{G} - \hat{G}|_2 \left(|\tilde{G}^{-1}|_2 + |\hat{G}|_2 |\tilde{G}^{-1}|_2 |\hat{G}^{-1}|_2 \right) \\
&\leq |\tilde{G} - \hat{G}|_F \left(|\tilde{G}^{-1}|_2 + |\hat{G}|_2 |\tilde{G}^{-1}|_2 |\hat{G}^{-1}|_2 \right) \\
&= O_{\mathbb{P}}(q^{3/2} T^{-1}) \left(O_{\mathbb{P}}(1) + O_{\mathbb{P}}(1) O_{\mathbb{P}}(1) O_{\mathbb{P}}(1) \right) \\
&= O_{\mathbb{P}}(q^{3/2} T^{-1}).
\end{aligned}$$

Since $\tilde{\mathbf{\Phi}}(q) - \hat{\mathbf{\Phi}}(q)$ is $((m+1) \times q(m+1))$-dimensional and m is fixed, it holds that (see, e. g. Gentle, 2007)

$$|\tilde{\mathbf{\Phi}}(q) - \hat{\mathbf{\Phi}}(q)|_F \leq \sqrt{m+1}\ |\tilde{\mathbf{\Phi}}(q) - \hat{\mathbf{\Phi}}(q)|_2 = O_{\mathbb{P}}(q^{3/2} T^{-1}).$$

This implies that

$$q^{1/2} \sum_{j=1}^{q} |\hat{\Phi}_j(q) - \tilde{\Phi}_j(q)|_F \leq q^{3/2} |\tilde{\mathbf{\Phi}}(q) - \hat{\mathbf{\Phi}}(q)|_F = O_{\mathbb{P}}(q^3 T^{-1}),$$

which is $o_{\mathbb{P}}(1)$ since $q^3 T^{-1} = o(1)$ by Assumption 4. □

In the proofs of Lemma 2 and 3 we repeatedly use the fact that by convexity, $|\sum_{i=1}^{k} z_i|^a \leq k^{a-1} \sum_{i=1}^{k} |z_i|^a$, for all $a, k \geq 1$.

Proof of Lemma 2. Let $\tilde{\varepsilon}_t(q) := w_t - \sum_{j=1}^{q} \tilde{\Phi}_j(q) w_{t-j}$, $t = q+1, \ldots, T$, denote the Yule-Walker residuals in the regression of w_t on $w_{t-1}, \dots, w_{t-q}$, $t = q+1, \dots, T$ and define $\bar{\tilde{\varepsilon}}_T(q) \coloneqq (T-q)^{-1} \sum_{t=q+1}^{T} \tilde{\varepsilon}_t(q)$.[23] For $q+1 \leq t \leq T$ we have

$$
\begin{aligned}
&|\hat{\varepsilon}_t(q) - \bar{\tilde{\varepsilon}}_T(q)|_F^a \\
&\leq \left(|\hat{\varepsilon}_t(q) - \tilde{\varepsilon}_t(q)|_F + |\tilde{\varepsilon}_t(q) - \bar{\tilde{\varepsilon}}_T(q)|_F + |\bar{\hat{\varepsilon}}_T(q) - \bar{\tilde{\varepsilon}}_T(q)|_F \right)^a \\
&\leq \left(|\hat{\varepsilon}_t(q) - \tilde{\varepsilon}_t(q)|_F + |\tilde{\varepsilon}_t(q) - \bar{\tilde{\varepsilon}}_T(q)|_F + (T-q)^{-1} \sum_{t=q+1}^{T} |\hat{\varepsilon}_t(q) - \tilde{\varepsilon}_t(q)|_F \right)^a \\
&\leq 3^{a-1} \left(|\hat{\varepsilon}_t(q) - \tilde{\varepsilon}_t(q)|_F^a + |\tilde{\varepsilon}_t(q) - \bar{\tilde{\varepsilon}}_T(q)|_F^a + \left((T-q)^{-1} \sum_{t=q+1}^{T} |\hat{\varepsilon}_t(q) - \tilde{\varepsilon}_t(q)|_F \right)^a \right).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&(T-q)^{-1} \sum_{t=q+1}^{T} |\hat{\varepsilon}_t(q) - \bar{\tilde{\varepsilon}}_T(q)|_F^a \\
&\leq 3^{a-1} \Bigg((T-q)^{-1} \sum_{t=q+1}^{T} |\hat{\varepsilon}_t(q) - \tilde{\varepsilon}_t(q)|_F^a + (T-q)^{-1} \sum_{t=q+1}^{T} |\tilde{\varepsilon}_t(q) - \bar{\tilde{\varepsilon}}_T(q)|_F^a \\
&\qquad + \left((T-q)^{-1} \sum_{t=q+1}^{T} |\hat{\varepsilon}_t(q) - \tilde{\varepsilon}_t(q)|_F \right)^a \Bigg) \\
&= 3^{a-1} \left(F_{T,a} + (T-q)^{-1} \sum_{t=q+1}^{T} |\tilde{\varepsilon}_t(q) - \bar{\tilde{\varepsilon}}_T(q)|_F^a + (F_{T,1})^a \right),
\end{aligned}
$$

where $F_{T,a} \coloneqq (T-q)^{-1} \sum_{t=q+1}^{T} |\hat{\varepsilon}_t(q) - \tilde{\varepsilon}_t(q)|_F^a$.

[23] In the following, a denotes the fixed $a > 2$ from Assumption 1. However, the results also hold for $1 \leq \tilde{a} < a$.

We now consider $F_{T,a}$ in more detail. For $q+1 \leq t \leq T$ we have

$$
\begin{aligned}
&|\hat{\varepsilon}_t(q) - \tilde{\varepsilon}_t(q)|_F \\
&= |\hat{w}_t - \sum_{j=1}^{q} \hat{\Phi}_j(q)\hat{w}_{t-j} - (w_t - \sum_{j=1}^{q} \tilde{\Phi}_j(q) w_{t-j})|_F \\
&= |\hat{w}_t - w_t - \sum_{j=1}^{q} \left(\hat{\Phi}_j(q) - \tilde{\Phi}_j(q) + \tilde{\Phi}_j(q)\right)(\hat{w}_{t-j} - w_{t-j} + w_{t-j}) + \sum_{j=1}^{q} \tilde{\Phi}_j(q) w_{t-j}|_F \\
&= |\hat{w}_t - w_t - \sum_{j=1}^{q} \left(\hat{\Phi}_j(q) - \tilde{\Phi}_j(q)\right)(\hat{w}_{t-j} - w_{t-j} + w_{t-j}) - \sum_{j=1}^{q} \tilde{\Phi}_j(q)\left(\hat{w}_{t-j} - w_{t-j}\right)|_F \\
&\leq \max_{1\leq t\leq T} |\hat{w}_t - w_t|_F + \sum_{j=1}^{q} |\hat{\Phi}_j(q) - \tilde{\Phi}_j(q)|_F |\hat{w}_{t-j} - w_{t-j}|_F \\
&\quad + \sum_{j=1}^{q} |\tilde{\Phi}_j(q)|_F |\hat{w}_{t-j} - w_{t-j}|_F + \sum_{j=1}^{q} |\hat{\Phi}_j(q) - \tilde{\Phi}_j(q)|_F |w_{t-j}|_F \\
&\leq \max_{1\leq t\leq T} |\hat{w}_t - w_t|_F + \max_{1\leq t\leq T} |\hat{w}_t - w_t|_F \sum_{j=1}^{q} |\hat{\Phi}_j(q) - \tilde{\Phi}_j(q)|_F \\
&\quad + \max_{1\leq t\leq T} |\hat{w}_t - w_t|_F \sum_{j=1}^{q} |\tilde{\Phi}_j(q)|_F + \sqrt{m+1}\, q |\hat{\boldsymbol{\Phi}}(q) - \tilde{\boldsymbol{\Phi}}(q)|_F \; q^{-1} \sum_{j=1}^{q} |w_{t-j}|_F,
\end{aligned}
$$

where we have used that

$$
\begin{aligned}
|\hat{\Phi}_j(q) - \tilde{\Phi}_j(q)|_F &\leq |\hat{\boldsymbol{\Phi}}(q) - \tilde{\boldsymbol{\Phi}}(q)|_F |[0, I_{m+1}, 0]'|_F \\
&= \sqrt{m+1}|\hat{\boldsymbol{\Phi}}(q) - \tilde{\boldsymbol{\Phi}}(q)|_F.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&|\hat{\varepsilon}_t(q) - \tilde{\varepsilon}_t(q)|_F^a \\
&\leq 4^{a-1} \left(\left(\max_{1\leq t\leq T} |\hat{w}_t - w_t|_F \right)^a + \left(\max_{1\leq t\leq T} |\hat{w}_t - w_t|_F \right)^a \left(\sum_{j=1}^{q} |\hat{\Phi}_j(q) - \tilde{\Phi}_j(q)|_F \right)^a \right. \\
&\quad + \left(\max_{1\leq t\leq T} |\hat{w}_t - w_t|_F \right)^a \left(\sum_{j=1}^{q} |\tilde{\Phi}_j(q)|_F \right)^a \\
&\quad \left. + \left(\sqrt{m+1}\, q |\hat{\boldsymbol{\Phi}}(q) - \tilde{\boldsymbol{\Phi}}(q)|_F \right)^a \left(q^{-1} \sum_{j=1}^{q} |w_{t-j}|_F \right)^a \right).
\end{aligned}
$$

It follows that

$$F_{T,a} \leq 4^{a-1} \left(\left(\max_{1\leq t\leq T} |\hat{w}_t - w_t|_F \right)^a + \left(\max_{1\leq t\leq T} |\hat{w}_t - w_t|_F \right)^a \left(\sum_{j=1}^{q} |\hat{\Phi}_j(q) - \tilde{\Phi}_j(q)|_F \right)^a \right.$$
$$+ \left(\max_{1\leq t\leq T} |\hat{w}_t - w_t|_F \right)^a \left(\sum_{j=1}^{q} |\tilde{\Phi}_j(q)|_F \right)^a$$
$$\left. + \left(\sqrt{m+1} q |\hat{\mathbf{\Phi}}(q) - \tilde{\mathbf{\Phi}}(q)|_F \right)^a (T-q)^{-1} \sum_{t=q+1}^{T} \left(q^{-1} \sum_{j=1}^{q} |w_{t-j}|_F \right)^a \right).$$

From Lemma 1 we have $\max_{1\leq t\leq T} |\hat{w}_t - w_t|_F = O_{\mathbb{P}}(T^{-1/2})$ and $\sum_{j=1}^{q} |\hat{\Phi}_j(q) - \tilde{\Phi}_j(q)|_F = O_{\mathbb{P}}(q^{5/2}T^{-1})$. Moreover, the proof of Lemma 1 shows that $q|\hat{\mathbf{\Phi}}(q) - \tilde{\mathbf{\Phi}}(q)|_F = O_{\mathbb{P}}(q^{5/2}T^{-1})$. Further, by (B.3), (B.4) and Assumption 1 we have

$$\begin{aligned} \sum_{j=1}^{q} |\tilde{\Phi}_j(q)|_F &\leq \sum_{j=1}^{q} |\tilde{\Phi}_j(q) - \Phi_j(q)|_F + \sum_{j=1}^{q} |\Phi_j(q) - \Phi_j|_F + \sum_{j=1}^{q} |\Phi_j|_F \\ &\leq q \sup_{1\leq j\leq q} |\tilde{\Phi}_j(q) - \Phi_j(q)|_F + c \sum_{j=q+1}^{\infty} |\Phi_j|_F + \sum_{j=1}^{\infty} |\Phi_j|_F \\ &= O_{\mathbb{P}}(1). \end{aligned}$$

Finally, note that

$$\begin{aligned} (T-q)^{-1} \sum_{t=q+1}^{T} \left(q^{-1} \sum_{j=1}^{q} |w_{t-j}|_F \right)^a &\leq (T-q)^{-1} \sum_{t=q+1}^{T} q^{-a} q^{a-1} \sum_{j=1}^{q} |w_{t-j}|_F^a \\ &= (T-q)^{-1} q^{-1} \sum_{t=q+1}^{T} \sum_{j=1}^{q} |w_{t-j}|_F^a \\ &\leq (T-q)^{-1} \sum_{t=1}^{T-1} |w_t|_F^a, \end{aligned}$$

where the last inequality follows from the fact that each element in the double sum occurs at most q times, i. e., $\sum_{t=q+1}^{T} \sum_{j=1}^{q} |w_{t-j}|_F^a \leq q \sum_{t=1}^{T-1} |w_t|_F^a$. From $\sup_{t\in\mathbb{Z}} \mathbb{E}\left(|w_t|_F^a\right) < \infty$ and Markov's inequality, it follows that $(T-q)^{-1} \sum_{t=1}^{T-1} |w_t|_F^a = O_{\mathbb{P}}(1)$. In total, we thus have $F_{T,a} = O_{\mathbb{P}}((q^{5/2}T^{-1})^a) = o_{\mathbb{P}}(1)$. Therefore,

$$(T-q)^{-1} \sum_{t=q+1}^{T} |\hat{\varepsilon}_t(q) - \bar{\hat{\varepsilon}}_T(q)|_F^a \leq 3^{a-1} \left((T-q)^{-1} \sum_{t=q+1}^{T} |\tilde{\varepsilon}_t(q) - \bar{\tilde{\varepsilon}}_T(q)|_F^a + o_{\mathbb{P}}(1) \right).$$

It thus remains to show that $(T-q)^{-1}\sum_{t=q+1}^{T}|\tilde{\varepsilon}_t(q)-\bar{\tilde{\varepsilon}}_T(q)|_F^a = O_{\mathbb{P}}(1)$. We now follow Park (2002, Proof of Lemma 3.2) and Palm *et al.* (2010, Proof of Lemma 2). Define $\varepsilon_t(q) \coloneqq \varepsilon_t + \sum_{j=q+1}^{\infty}\Phi_j w_{t-j}$ and note that

$$\begin{aligned}
&(T-q)^{-1}\sum_{t=q+1}^{T}|\tilde{\varepsilon}_t(q)-\bar{\tilde{\varepsilon}}_T(q)|_F^a \\
&= (T-q)^{-1}\sum_{t=q+1}^{T}|\tilde{\varepsilon}_t(q)-\varepsilon_t(q)+\varepsilon_t(q)-\varepsilon_t+\varepsilon_t-\bar{\tilde{\varepsilon}}_T(q)|_F^a \\
&\leq 4^{a-1}\left(A_{T,a}+B_{T,a}+C_{T,a}+D_{T,a}\right),
\end{aligned}$$

where

$$\begin{aligned}
A_{T,a} &\coloneqq (T-q)^{-1}\sum_{t=q+1}^{T}|\varepsilon_t|_F^a, \\
B_{T,a} &\coloneqq (T-q)^{-1}\sum_{t=q+1}^{T}|\varepsilon_t(q)-\varepsilon_t|_F^a = (T-q)^{-1}\sum_{t=q+1}^{T}|\sum_{j=q+1}^{\infty}\Phi_j w_{t-j}|_F^a, \\
C_{T,a} &\coloneqq (T-q)^{-1}\sum_{t=q+1}^{T}|\tilde{\varepsilon}_t(q)-\varepsilon_t(q)|_F^a, \\
D_{T,a} &\coloneqq (T-q)^{-1}\sum_{t=q+1}^{T}|\bar{\tilde{\varepsilon}}_T(q)|_F^a = |\bar{\tilde{\varepsilon}}_T(q)|_F^a = |(T-q)^{-1}\sum_{t=q+1}^{T}\tilde{\varepsilon}_t(q)|_F^a.
\end{aligned}$$

We first consider $B_{T,a}$. Note that $\mathbb{E}\left(|B_{T,a}|_F\right) \leq \sup_{t\in\mathbb{Z}}\mathbb{E}\left(|\varepsilon_t(q)-\varepsilon_t|_F^a\right)$. Using Minkowski's inequality, we have

$$\begin{aligned}
\mathbb{E}\left(|\varepsilon_t(q)-\varepsilon_t|_F^a\right) &= \mathbb{E}\left(|\sum_{j=q+1}^{\infty}\Phi_j w_{t-j}|_F^a\right) = \left(\left[\mathbb{E}\left(|\sum_{j=q+1}^{\infty}\Phi_j w_{t-j}|_F^a\right)\right]^{1/a}\right)^a \\
&\leq \left(\sum_{j=q+1}^{\infty}\left[\mathbb{E}\left(|\Phi_j w_{t-j}|_F^a\right)\right]^{1/a}\right)^a \leq \left(\sum_{j=q+1}^{\infty}\left[\mathbb{E}\left(|\Phi_j|_F^a|w_{t-j}|_F^a\right)\right]^{1/a}\right)^a \\
&\leq \left(\sum_{j=q+1}^{\infty}|\Phi_j|_F\left[\mathbb{E}\left(|w_{t-j}|_F^a\right)\right]^{1/a}\right)^a \leq \left(\sum_{j=q+1}^{\infty}|\Phi_j|_F\left[\sup_{t\in\mathbb{Z}}\mathbb{E}\left(|w_t|_F^a\right)\right]^{1/a}\right)^a \\
&\leq \sup_{t\in\mathbb{Z}}\mathbb{E}\left(|w_t|_F^a\right)\left(\sum_{j=q+1}^{\infty}|\Phi_j|_F\right)^a.
\end{aligned}$$

From Assumption 1 we have $\sup_{t\in\mathbb{Z}}\mathbb{E}\left(|w_t|_F^a\right) < \infty$ and $\sum_{j=q+1}^{\infty}|\Phi_j|_F = o(1)$. Markov's inequality thus yields $B_{T,a} = o_{\mathbb{P}}(1)$. Analogously, $\mathbb{E}\left(|A_{T,a}|_F\right) \leq \sup_{t\in\mathbb{Z}}\mathbb{E}\left(|\varepsilon_t|_F^a\right)$.

Using Minkowski's inequality, we have as above

$$
\begin{aligned}
\mathbb{E}\left(|\varepsilon_t|_F^a\right) = \mathbb{E}\left(|w_t - \sum_{j=1}^{\infty} \Phi_j w_{t-j}|_F^a\right) &\le 2^{a-1}\left(\mathbb{E}\left(|w_t|_F^a\right) + \mathbb{E}\left(|\sum_{j=1}^{\infty} \Phi_j w_{t-j}|_F^a\right)\right) \\
&\le 2^{a-1}\left(\sup_{t\in\mathbb{Z}} \mathbb{E}\left(|w_t|_F^a\right) + \mathbb{E}\left(|\sum_{j=1}^{\infty} \Phi_j w_{t-j}|_F^a\right)\right) \\
&\le 2^{a-1} \sup_{t\in\mathbb{Z}} \mathbb{E}\left(|w_t|_F^a\right)\left(1 + \left(\sum_{j=1}^{\infty} |\Phi_j|_F\right)^a\right) < \infty.
\end{aligned}
$$

Using Markov's inequality we conclude that $A_{T,a} = O_{\mathbb{P}}(1)$. We now turn to $C_{T,a}$. By definition,

$$
\begin{aligned}
\tilde{\varepsilon}_t(q) = w_t - \sum_{j=1}^{q} \tilde{\Phi}_j(q) w_{t-j} &= \varepsilon_t(q) - \sum_{j=1}^{q} \left(\tilde{\Phi}_j(q) - \Phi_j\right) w_{t-j} \\
&= \varepsilon_t(q) - \sum_{j=1}^{q} \left(\tilde{\Phi}_j(q) - \Phi_j(q)\right) w_{t-j} - \sum_{j=1}^{q} \left(\Phi_j(q) - \Phi_j\right) w_{t-j}.
\end{aligned}
$$

Hence,

$$
|\tilde{\varepsilon}_t(q) - \varepsilon_t(q)|_F^a \le 2^{a-1}\left(|\sum_{j=1}^{q} \left(\tilde{\Phi}_j(q) - \Phi_j(q)\right) w_{t-j}|_F^a + |\sum_{j=1}^{q} \left(\Phi_j(q) - \Phi_j\right) w_{t-j}|_F^a\right).
$$

It follows that $C_{T,a} = 2^{a-1}\left(C_{1T,a} + C_{2T,a}\right)$, where

$$
\begin{aligned}
C_{1T,a} &\coloneqq (T-q)^{-1} \sum_{t=q+1}^{T} |\sum_{j=1}^{q} \left(\tilde{\Phi}_j(q) - \Phi_j(q)\right) w_{t-j}|_F^a, \\
C_{2T,a} &\coloneqq (T-q)^{-1} \sum_{t=q+1}^{T} |\sum_{j=1}^{q} \left(\Phi_j(q) - \Phi_j\right) w_{t-j}|_F^a.
\end{aligned}
$$

We consider both terms separately. First note that

$$
\begin{aligned}
C_{1T,a} &\le q^{a-1}(T-q)^{-1} \sum_{t=q+1}^{T} \sum_{j=1}^{q} |\tilde{\Phi}_j(q) - \Phi_j(q)|_F^a |w_{t-j}|_F^a \\
&\le q^{a-1}\left(\sup_{1\le j\le q} |\tilde{\Phi}_j(q) - \Phi_j(q)|_F\right)^a (T-q)^{-1} \sum_{t=q+1}^{T} \sum_{j=1}^{q} |w_{t-j}|_F^a \\
&\le \left(q \sup_{1\le j\le q} |\tilde{\Phi}_j(q) - \Phi_j(q)|_F\right)^a (T-q)^{-1} \sum_{t=1}^{T-1} |w_t|_F^a,
\end{aligned}
$$

where the third inequality follows again from the fact that $\sum_{t=q+1}^{T}\sum_{j=1}^{q}|w_{t-j}|_F^a \leq q\sum_{t=1}^{T-1}|w_t|_F^a$. As $(T-q)^{-1}\sum_{t=1}^{T-1}|w_t|_F^a = O_{\mathbb{P}}(1)$ it follows from (B.3) that $C_{1T,a} = o_{\mathbb{P}}(1)$. Moreover, using Minkowski's inequality, we obtain

$$
\begin{aligned}
\mathbb{E}\left(|C_{2T,a}|_F\right) &= (T-q)^{-1}\sum_{t=q+1}^{T}\mathbb{E}\left(|\sum_{j=1}^{q}\left(\Phi_j(q)-\Phi_j\right)w_{t-j}|_F^a\right) \\
&= (T-q)^{-1}\sum_{t=q+1}^{T}\left(\left[\mathbb{E}\left(|\sum_{j=1}^{q}\left(\Phi_j(q)-\Phi_j\right)w_{t-j}|_F^a\right)\right]^{1/a}\right)^a \\
&\leq (T-q)^{-1}\sum_{t=q+1}^{T}\left(\sum_{j=1}^{q}\left[\mathbb{E}\left(|\left(\Phi_j(q)-\Phi_j\right)w_{t-j}|_F^a\right)\right]^{1/a}\right)^a \\
&\leq (T-q)^{-1}\sum_{t=q+1}^{T}\sup_{t\in\mathbb{Z}}\mathbb{E}\left(|w_t|_F^a\right)\left(\sum_{j=1}^{q}|\Phi_j(q)-\Phi_j|_F\right)^a \\
&= \sup_{t\in\mathbb{Z}}\mathbb{E}\left(|w_t|_F^a\right)\left(\sum_{j=1}^{q}|\Phi_j(q)-\Phi_j|_F\right)^a.
\end{aligned}
$$

From (B.4) and Markov's inequality it follows that $C_{2T,a} = o_{\mathbb{P}}(1)$. In total we thus have $C_{T,a} = o_{\mathbb{P}}(1)$. Finally, we consider $D_{T,a}$. It holds that $(T-q)^{-1}\sum_{t=q+1}^{T}\tilde{\varepsilon}_t(q) = D_{1T}+D_{2T}+D_{3T}$, where

$$
\begin{aligned}
D_{1T} &:= (T-q)^{-1}\sum_{t=q+1}^{T}\varepsilon_t, \\
D_{2T} &:= (T-q)^{-1}\sum_{t=q+1}^{T}\left(\varepsilon_t(q)-\varepsilon_t\right), \\
D_{3T} &:= (T-q)^{-1}\sum_{t=q+1}^{T}\left(\tilde{\varepsilon}_t(q)-\varepsilon_t(q)\right).
\end{aligned}
$$

By Chebyshev's weak law of large numbers (White, 2001, p. 25), $D_{1T} \xrightarrow{p} \mathbb{E}\left(\varepsilon_t\right) = 0$, i. e., $D_{1T} = o_{\mathbb{P}}(1)$. Moreover, $|D_{2T}|_F \leq B_{T,1} = o_{\mathbb{P}}(1)$ and $|D_{3T}|_F \leq C_{T,1} = o_{\mathbb{P}}(1)$. By the continuous mapping theorem we thus have $D_T = o_{\mathbb{P}}(1)$. This completes the proof. $\square$

Proof of Lemma 3. It follows from Assumption 2 that

$$
|(T-q)^{-1}\sum_{t=q+1}^{T}\varepsilon_t\varepsilon_t' - \Sigma|_F = o_{\mathbb{P}}(1).
$$

Therefore,

$$|\mathbb{E}^*\left(\varepsilon_t^*\varepsilon_t^{*\prime}\right)-\Sigma|_F \leq |\mathbb{E}^*\left(\varepsilon_t^*\varepsilon_t^{*\prime}\right)-(T-q)^{-1}\sum_{t=q+1}^{T}\varepsilon_t\varepsilon_t'|_F + o_{\mathbb{P}}(1).$$

Moreover,

$$\begin{aligned}
&|\mathbb{E}^*\left(\varepsilon_t^*\varepsilon_t^{*\prime}\right)-(T-q)^{-1}\sum_{t=q+1}^{T}\varepsilon_t\varepsilon_t'|_F \\
&= |(T-q)^{-1}\sum_{t=q+1}^{T}\left(\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)\right)\left(\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)\right)'-\varepsilon_t\varepsilon_t'|_F \\
&= |(T-q)^{-1}\sum_{t=q+1}^{T}\left(\left[\left(\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)\right)-\varepsilon_t\right]\left(\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)\right)'\right. \\
&\qquad\qquad\qquad\left.+\varepsilon_t\left[\left(\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)\right)-\varepsilon_t\right]'\right)|_F \\
&\leq E_{1T}+E_{2T},
\end{aligned}$$

where

$$\begin{aligned}
E_{1T} &= (T-q)^{-1}\sum_{t=q+1}^{T}|\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)-\varepsilon_t|_F|\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)|_F, \\
E_{2T} &= (T-q)^{-1}\sum_{t=q+1}^{T}|\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)-\varepsilon_t|_F|\varepsilon_t|_F.
\end{aligned}$$

The Cauchy-Schwarz inequality yields

$$\begin{aligned}
E_{1T} &\leq \left((T-q)^{-1}\sum_{t=q+1}^{T}|\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)-\varepsilon_t|_F^2(T-q)^{-1}\sum_{t=q+1}^{T}|\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)|_F^2\right)^{1/2}, \\
E_{2T} &\leq \left((T-q)^{-1}\sum_{t=q+1}^{T}|\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)-\varepsilon_t|_F^2(T-q)^{-1}\sum_{t=q+1}^{T}|\varepsilon_t|_F^2\right)^{1/2}.
\end{aligned}$$

From the proof of Lemma 2 we have $(T-q)^{-1}\sum_{t=q+1}^{T}|\varepsilon_t|_F^2 = A_{T,2} = O_{\mathbb{P}}(1)$ and $(T-q)^{-1}\sum_{t=q+1}^{T}|\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)|_F^2 = O_{\mathbb{P}}(1)$. It thus remains to show that

$(T-q)^{-1}\sum_{t=q+1}^{T}|\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)-\varepsilon_t|_F^2=o_{\mathbb{P}}(1)$. To this end note that

$$
\begin{aligned}
&|\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)-\varepsilon_t|_F \\
&\leq |\hat{\varepsilon}_t(q)-\tilde{\varepsilon}_t(q)|_F+|\tilde{\varepsilon}_t(q)-(w_t-\sum_{j=1}^{q}\Phi_j(q)w_{t-j})|_F \\
&\quad +|w_t-\sum_{j=1}^{q}\Phi_j(q)w_{t-j}-\varepsilon_t|_F+|\bar{\hat{\varepsilon}}_T(q)|_F \\
&=|\hat{\varepsilon}_t(q)-\tilde{\varepsilon}_t(q)|_F+|\sum_{j=1}^{q}\left(\tilde{\Phi}_j(q)-\Phi_j(q)\right)w_{t-j}|_F \\
&\quad +|\sum_{j=1}^{q}\left(\Phi_j(q)-\Phi_j\right)w_{t-j}|_F+|\sum_{j=q+1}^{\infty}\Phi_j w_{t-j}|_F+|\bar{\hat{\varepsilon}}_T(q)|_F.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&|\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)-\varepsilon_t|_F^2 \\
&\leq 5\Bigg(|\hat{\varepsilon}_t(q)-\tilde{\varepsilon}_t(q)|_F^2+|\sum_{j=1}^{q}\left(\tilde{\Phi}_j(q)-\Phi_j(q)\right)w_{t-j}|_F^2 \\
&\qquad +|\sum_{j=1}^{q}\left(\Phi_j(q)-\Phi_j\right)w_{t-j}|_F^2+|\sum_{j=q+1}^{\infty}\Phi_j w_{t-j}|_F^2+|\bar{\hat{\varepsilon}}_T(q)|_F^2\Bigg).
\end{aligned}
$$

In the notation of the proof of Lemma 2, we obtain

$$
\begin{aligned}
(T-q)^{-1}\sum_{t=q+1}^{T}|\hat{\varepsilon}_t(q)-\bar{\hat{\varepsilon}}_T(q)-\varepsilon_t|_F^2 &\leq 5\left(F_{T,2}+C_{1T,2}+C_{2T,2}+B_{T,2}+|\bar{\hat{\varepsilon}}_T(q)|_F^2\right) \\
&=5|\bar{\hat{\varepsilon}}_T(q)|_F^2+o_{\mathbb{P}}(1).
\end{aligned}
$$

From $\hat{\varepsilon}_t(q)=\hat{\varepsilon}_t(q)-\tilde{\varepsilon}_t(q)+\tilde{\varepsilon}_t(q)-\varepsilon_t(q)+\varepsilon_t(q)-\varepsilon_t+\varepsilon_t$, with $\varepsilon_t(q)$ as defined in the proof of Lemma 2, it follows that

$$
\begin{aligned}
|\bar{\hat{\varepsilon}}_T(q)|_F &\leq |(T-q)^{-1}\sum_{t=q+1}^{T}\hat{\varepsilon}_t(q)-\tilde{\varepsilon}_t(q)|_F+|(T-q)^{-1}\sum_{t=q+1}^{T}\tilde{\varepsilon}_t(q)-\varepsilon_t(q)|_F \\
&\quad +|(T-q)^{-1}\sum_{t=q+1}^{T}\varepsilon_t(q)-\varepsilon_t|_F+|(T-q)^{-1}\sum_{t=q+1}^{T}\varepsilon_t|_F \\
&\leq F_{T,1}+C_{T,1}+B_{T,1}+|D_{1T}|_F=o_{\mathbb{P}}(1).
\end{aligned}
$$

This completes the proof. □

References

Brillinger, D.R. (1981). Time Series: Data Analysis and Theory. Holden-Day, Inc., San Francisco.

Gentle, J.E. (2007). Matrix Algebra. Springer, New York.

Horn, R.A., Johnson, C.R. (2012). Matrix Analysis (2nd Ed.). Cambridge University Press, New York.

Meyer, M., Kreiss, J.-P. (2015). On the Vector Autoregressive Sieve Bootstrap. *Journal of Time Series Analysis* **36**, 377 – 397.

Palm, F.C., Smeekes, S., Urbain, J.P. (2010). A Sieve Bootstrap Test for Cointegration in a Conditional Error Correction Model. *Econometric Theory* **26**, 647 – 681.

Park, J.Y. (2002). An Invariance Principle for Sieve Bootstrap in Time Series. *Econometric Theory* **18**, 469–490.

White, H. (2001). Asymptotic Theory for Econometricians. Academic Press, San Diego.