Supplement to "A consistent specification test for dynamic quantile models"

(*Quantitative Economics*, Vol. 13, No. 1, January 2022, 125–151)

Peter Horvath
Department of Economics, Duke University

Jia Li
Department of Economics, Duke University

Zhipeng Liao
Department of Economics, UCLA

Andrew J. Patton
Department of Economics, Duke University

This Appendix contains the technical lemmas used in the proofs of the main theorems.

SA. Proofs

Throughout the proofs, we use K to denote a generic finite constant that may change from line to line. For $p \geq 1$, let $\|\cdot\|_p$ denote the L_p norm for random variables. For notational simplicity, we write $\sum_t$ in place of $\sum_{t=1}^n$.

SA.1 *Technical lemmas for Theorem 1*

Lemma S1. *Under Assumptions* 1 *and* 2, *we have for any* $R > 0$,

$$\sup_{\theta \in B_n(R)} n^{-1} \sum_t \big\| P\big(X_t(\theta)\big) - P\big(X_t^\star\big) \big\|^2 = O_p\big(\zeta_{1,n}^2 m_n n^{-1}\big),$$
$$n^{-1} \sum_t \big\| P(\widehat{X}_t) - P\big(X_t^\star\big) \big\|^2 = O_p\big(\zeta_{1,n}^2 m_n n^{-1}\big).$$

Proof. Fix some constant $\eta > 0$. Since the variables $(L_{X,t})_{1 \leq t \leq n}$ are L_p-bounded, we can use a maximal inequality to deduce $\mathbb{E}[\max_{1 \leq t \leq n} |L_{X,t}|] \leq n^{1/p} \max_{1 \leq t \leq n} \|L_{X,t}\|_p \leq Kn^{1/p}$. Note that

$$\sup_{\theta \in B_n(R)} \max_{1 \leq t \leq n} \big\| X_t(\theta) - X_t^\star \big\| \leq K n^{-1/2} \max_{1 \leq t \leq n} |L_{X,t}| = O_p\big(n^{1/p - 1/2}\big).$$

Peter Horvath: peter.horvath@duke.edu
Jia Li: jl410@duke.edu
Zhipeng Liao: zhipeng.liao@econ.ecla.edu
Andrew J. Patton: andrew.patton@duke.edu

Available at http://qeconomics.org. https://doi.org/10.3982/QE1727

Hence, there exists some constant $C > 0$ such that

$$\mathbb{P}\big(\Omega_n^c\big) < \eta/2, \quad \text{where } \Omega_n \equiv \Big\{ \sup_{\theta \in B_n(R)} \max_{1 \le t \le n} \big\| X_t(\theta) - X_t^\star \big\| \le C n^{1/p - 1/2} \Big\}.$$

In particular, $X_t(\theta) \in \mathcal{X} \oplus C n^{1/p-1/2}$ for all t and $\theta \in B_n(R)$ in restriction to Ω_n, and hence,

$$\begin{aligned}
&\sup_{\theta \in B_n(R)} n^{-1} \sum_t \big\| P\big(X_t(\theta)\big) - P\big(X_t^\star\big) \big\|^2 \\
&= \sup_{\theta \in B_n(R)} n^{-1} \sum_{l=1}^{m_n} \sum_t \big| \partial p_l(\bar{X}_{\theta,t})^\top \big(X_t(\theta) - X_t^\star\big) \big|^2 \\
&\le \sup_{\theta \in B_n(R)} n^{-1} \sum_{l=1}^{m_n} \sum_t \big\| \partial p_l(\bar{X}_{\theta,t}) \big\|^2 \big\| X_t(\theta) - X_t^\star \big\|^2 \\
&\le K \zeta_{1,n}^2 m_n n^{-1} \Big(n^{-1} \sum_t L_{X,t}^2 \Big) = O_p\big(\zeta_{1,n}^2 m_n n^{-1}\big),
\end{aligned}$$

where $\bar{X}_{\theta,t}$ is some mean value between $X_t(\theta)$ and $X_t^\star$, the first inequality is by the Cauchy–Schwarz inequality, and the second inequality is by Assumptions 1 and 2. Therefore, there exists some constant $K > 0$ such that

$$\mathbb{P}\bigg(\Big\{ \sup_{\theta \in B_n(R)} n^{-1} \sum_t \big\| P\big(X_t(\theta)\big) - P\big(X_t^\star\big) \big\|^2 > K \zeta_{1,n}^2 m_n n^{-1} \Big\} \cap \Omega_n \bigg) < \eta/2.$$

Hence,

$$\begin{aligned}
&\mathbb{P}\bigg(\sup_{\theta \in B_n(R)} n^{-1} \sum_t \big\| P\big(X_t(\theta)\big) - P\big(X_t^\star\big) \big\|^2 > K \zeta_{1,n}^2 m_n n^{-1} \bigg) \\
&\le \mathbb{P}\bigg(\Big\{ \sup_{\theta \in B_n(R)} n^{-1} \sum_t \big\| P\big(X_t(\theta)\big) - P\big(X_t^\star\big) \big\|^2 > K \zeta_{1,n}^2 m_n n^{-1} \Big\} \cap \Omega_n \bigg) + \mathbb{P}\big(\Omega_n^c\big) < \eta.
\end{aligned}$$

This proves the first assertion of the lemma. Since $\hat{\theta}_n = \theta^\star + O_p(n^{-1/2})$, the second assertion readily follows from the first. □

Below, we denote

$$\widehat{Q}_n(\theta) \equiv n^{-1} \sum_t P\big(X_t(\theta)\big) P\big(X_t(\theta)\big)^\top.$$

LEMMA S2. *Suppose that Assumptions* 1 *and* 2 *hold. Then we have for any* $R > 0$,

$$\sup_{\theta \in B_n(R)} \big\| \widehat{Q}_n(\theta) - Q_n \big\|_S = O_p\big(\delta_{Q,n} + \zeta_{1,n} m_n^{1/2} n^{-1/2}\big) = o_p(1),$$

$$\|\widehat{Q}_n - Q_n\|_S = O_p\big(\delta_{Q,n} + \zeta_{1,n} m_n^{1/2} n^{-1/2}\big) = o_p(1).$$

PROOF. Note that $\widehat{Q}_n^\star = \widehat{Q}_n(\theta^\star)$. By the triangle inequality,

$$\|\widehat{Q}_n(\theta) - \widehat{Q}_n^\star\| \le \Big\| n^{-1}\sum_t \big(P(X_t(\theta)) - P(X_t^\star)\big)\big(P(X_t(\theta)) - P(X_t^\star)\big)^\top \Big\| + 2\Big\| n^{-1}\sum_t P(X_t^\star)\big(P(X_t(\theta)) - P(X_t^\star)\big)^\top \Big\|. \tag{S.1}$$

By the Cauchy–Schwarz inequality and Lemma S1, we have uniformly in $\theta \in B_n(R)$,

$$\Big\| n^{-1}\sum_t \big(P(X_t(\theta)) - P(X_t^\star)\big)\big(P(X_t(\theta)) - P(X_t^\star)\big)^\top \Big\|^2 \le \left(n^{-1}\sum_{l=1}^{m_n}\sum_t \big(p_l(X_t(\theta)) - p_l(X_t^\star)\big)^2 \right)^2 = O_p\big(\zeta_{1,n}^4 m_n^2 n^{-2}\big). \tag{S.2}$$

In addition, we have uniformly in $\theta \in B_n(R)$,

$$\begin{aligned} &\Big\| n^{-1}\sum_t P(X_t^\star)\big(P(X_t(\theta)) - P(X_t^\star)\big)^\top \Big\|^2 \\ &= \sum_{l=1}^{m_n}\Big\| n^{-1}\sum_t P(X_t^\star)\big(p_l(X_t(\theta)) - p_l(X_t^\star)\big)\Big\|^2 \\ &\le \lambda_{\max}\big(\widehat{Q}_n^\star\big)\sum_{l=1}^{m_n} n^{-1}\sum_t \big(p_l(X_t(\theta)) - p_l(X_t^\star)\big)^2 = O_p\big(\zeta_{1,n}^2 m_n n^{-1}\big), \end{aligned} \tag{S.3}$$

where the bound on the last line follows from Lemma S1 and $\lambda_{\max}(\widehat{Q}_n^\star) = O_p(1)$, with the latter implied by Assumption 2. By (S.1), (S.2), and (S.3), we have uniformly in $\theta \in B_n(R)$,

$$\begin{aligned} \|\widehat{Q}_n(\theta) - \widehat{Q}_n^\star\| &= O_p\big(\zeta_{1,n}^2 m_n n^{-1} + \zeta_{1,n} m_n^{1/2} n^{-1/2}\big) \\ &= O_p\big(\zeta_{1,n} m_n^{1/2} n^{-1/2}\big) = o_p(1), \end{aligned}$$

where the second line follows from $\zeta_{1,n} m_n^{1/2} n^{-1/2} = o(1)$ which is implied by Assumption 2(ii). The first assertion of the lemma thus follows from the estimate above and the assumption that $\|\widehat{Q}_n^\star - Q_n\|_S = O_p(\delta_{Q,n})$. The second assertion then follows from $\hat{\theta}_n = \theta^\star + O_p(n^{-1/2})$. □

LEMMA S3. *Under Assumption* 1, *we have for each* $R > 0$,

$$\sup_{\theta \in B_n(R)} n^{-1}\sum_t \big(Z_{t+1}(\theta) - Z_{t+1}^\star\big)^2 = O_p\big(n^{-1/2}\big).$$

PROOF. Fix any $R > 0$. Note that for each $\theta \in B_n(R)$, $f_t(\theta^\star) - Rn^{-1/2}L_t \le f_t(\theta) \le f_t(\theta^\star) + Rn^{-1/2}L_t$, and hence,

$$\big|Z_{t+1}(\theta) - Z_{t+1}(\theta^\star)\big|^2 = |1_{\{Y_{t+1} \le f_t(\theta)\}} - 1_{\{Y_{t+1} \le f_t(\theta^\star)\}}| \le U_t^+ - U_t^-, \tag{S.4}$$

where we set $U_t^{\pm} \equiv 1_{\{Y_{t+1} \le f_t(\theta^\star) \pm R n^{-1/2} L_t\}}$. Therefore,

$$\sup_{\theta \in B_n(R)} n^{-1} \sum_t \big(Z_{t+1}(\theta) - Z_{t+1}(\theta^\star)\big)^2 \le n^{-1} \sum_t \big(U_t^+ - U_t^-\big). \tag{S.5}$$

Recall that $F_{t+1|t}(\cdot)$ is the $\mathcal{F}_t$-conditional distribution function of Y_{t+1}. By Assumption 1,

$$\begin{aligned}\mathbb{E}\big[U_t^+ - U_t^- | \mathcal{F}_t\big] &= F_{t+1|t}\big(f_t(\theta^\star) + R n^{-1/2} L_t\big) - F_{t+1|t}\big(f_t(\theta^\star) - R n^{-1/2} L_t\big) \\ &\le 2 R n^{-1/2} L_t^2. \end{aligned} \tag{S.6}$$

Since L_t is L_2-bounded, $\mathbb{E}[U_t^+ - U_t^-] \le K n^{-1/2}$. The assertion of the lemma then readily follows from this estimate and (S.5). □

LEMMA S4. *Suppose that Assumptions* 1 *and* 2 *hold. Then we have for any* $R > 0$,

$$\sup_{\theta \in B_n(R)} \left\| n^{-1} \sum_t P\big(X_t(\theta)\big)\big(Z_{t+1}(\theta) - Z_{t+1}^\star\big) \right\| = O_p\big(n^{-1/2} + m_n^{1/2}\big(\zeta_{1,n} n^{-5/4} + \zeta_{0,n} n^{-3/4}\big)\big).$$

PROOF. Step 1. By Lemma S2, $\sup_{\theta \in B_n(R)} \|\widehat{Q}_n(\theta) - Q_n\| = o_p(1)$. Since the eigenvalues of Q_n are bounded from above and away from zero, we further deduce

$$\lambda_{\min}^{-1}\big(\widehat{Q}_n(\theta)\big) + \lambda_{\max}\big(\widehat{Q}_n(\theta)\big) = O_p(1), \quad \text{uniformly in } \theta \in B_n(R). \tag{S.7}$$

Recall that $\bar{Z}_{t+1}(\theta) = F_{t+1|t}(f_t(\theta)) - q$ is the $\mathcal{F}_t$-conditional mean of $Z_{t+1}(\theta)$. Let $\widetilde{Z}_{t+1}(\theta) \equiv Z_{t+1}(\theta) - \bar{Z}_{t+1}(\theta)$, which forms a martingale difference sequence with respect to $\mathcal{F}_t$ by construction. By the triangle inequality,

$$\begin{aligned} &\sup_{\theta \in B_n(R)} \left\| n^{-1} \sum_t P\big(X_t(\theta)\big)\big(Z_{t+1}(\theta) - Z_{t+1}^\star\big) \right\| \\ &\le \sup_{\theta \in B_n(R)} \left\| n^{-1} \sum_t P\big(X_t(\theta)\big)\big(\bar{Z}_{t+1}(\theta) - \bar{Z}_{t+1}(\theta^\star)\big) \right\| \\ &\quad + \sup_{\theta \in B_n(R)} \left\| n^{-1} \sum_t P\big(X_t(\theta)\big)\big(\widetilde{Z}_{t+1}(\theta) - \widetilde{Z}_{t+1}(\theta^\star)\big) \right\|. \end{aligned} \tag{S.8}$$

The first term on the majorant side of (S.8) can be bounded as follows:

$$\begin{aligned} &\sup_{\theta \in B_n(R)} \left\| n^{-1} \sum_t P\big(X_t(\theta)\big)\big(\bar{Z}_{t+1}(\theta) - \bar{Z}_{t+1}(\theta^\star)\big) \right\|^2 \\ &\le \sup_{\theta \in B_n(R)} \lambda_{\max}\big(\widehat{Q}_n(\theta)\big) \sup_{\theta \in B_n(R)} n^{-1} \sum_t \big(\bar{Z}_{t+1}(\theta) - \bar{Z}_{t+1}(\theta^\star)\big)^2 \\ &\le O_p(1) \cdot \sup_{\theta \in B_n(R)} \left\{ n^{-1} \sum_t \big(\|\partial_\theta \bar{Z}_{t+1}(\theta^\star)\| \|\theta - \theta^\star\| + \bar{L}_{t+1} \|\theta - \theta^\star\|^2\big)^2 \right\} \\ &= O_p(n^{-1}), \end{aligned}$$

where the first inequality is obtained by using the contraction property of least-square projections, the second inequality is due to Assumption 1, and the last line follows from the L_2-boundedness of $\|\partial_\theta \bar{Z}_t(\theta^\star)\|$ and $\bar{L}_t$. This estimate further implies that

$$\sup_{\theta\in B_n(R)}\left\|n^{-1}\sum_t P\big(X_t(\theta)\big)\big(\bar{Z}_{t+1}(\theta)-\bar{Z}_{t+1}(\theta^\star)\big)\right\| = O_p\big(n^{-1/2}\big).$$

Hence, to prove the assertion of the lemma, it remains to show that the second term on the majorant side of (S.8) satisfies

$$\sup_{\theta\in B_n(R)}\left\|n^{-1}\sum_t P\big(X_t(\theta)\big)\big(\widetilde{Z}_{t+1}(\theta)-\widetilde{Z}_{t+1}(\theta^\star)\big)\right\| \\ = O_p\big(m_n^{1/2}\big(\zeta_{1,n}n^{-5/4}+\zeta_{0,n}n^{-3/4}\big)\big). \tag{S.9}$$

Below, we prove (S.9) in two steps.

Step 2. For ease of notation, we set for each $l\in\{1,\ldots,m_n\}$,

$$\pi_{l,n}(\theta)\equiv n^{-1/2}\sum_t p_l\big(X_t(\theta)\big)\big(\widetilde{Z}_{t+1}(\theta)-\widetilde{Z}_{t+1}(\theta^\star)\big), \quad \theta\in\Theta.$$

In this step, we establish the following technical estimate:

$$\big\|\pi_{l,n}(\theta_1)-\pi_{l,n}(\theta_2)\big\|_p \\ \le K\big(\zeta_{1,n}n^{-1/2}+\zeta_{0,n}\big)\|\theta_1-\theta_2\|^{1/2}, \quad \text{for } \theta_1,\theta_2\in B_n(R). \tag{S.10}$$

Recall that $\widetilde{Z}_{t+1}(\theta)=Z_{t+1}(\theta)-\mathbb{E}[Z_{t+1}(\theta)|\mathcal{F}_t]$. It is then easy to see that

$$\mathbb{E}\big[\big(\widetilde{Z}_{t+1}(\theta)-\widetilde{Z}_{t+1}(\theta^\star)\big)^2|\mathcal{F}_t\big]\le\mathbb{E}\big[\big(Z_{t+1}(\theta)-Z_{t+1}(\theta^\star)\big)^2|\mathcal{F}_t\big].$$

By (S.4) and (S.6), the majorant side of the above inequality can be further bounded by $Kn^{-1/2}L_t^2$ uniformly in $\theta\in B_n(R)$. Hence,

$$\mathbb{E}\big[\big(\widetilde{Z}_{t+1}(\theta)-\widetilde{Z}_{t+1}(\theta^\star)\big)^2|\mathcal{F}_t\big]\le Kn^{-1/2}L_t^2. \tag{S.11}$$

For $\theta_1,\theta_2\in B_n(R)$, we can decompose

$$\pi_{l,n}(\theta_1)-\pi_{l,n}(\theta_2)=n^{-1/2}\sum_t\big[p_l\big(X_t(\theta_1)\big)-p_l\big(X_t(\theta_2)\big)\big]\big(\widetilde{Z}_{t+1}(\theta_1)-\widetilde{Z}_{t+1}(\theta^\star)\big) \\ +n^{-1/2}\sum_t p_l\big(X_t(\theta_2)\big)\big(\widetilde{Z}_{t+1}(\theta_1)-\widetilde{Z}_{t+1}(\theta_2)\big). \tag{S.12}$$

We now derive L_p-bounds for the two terms on the right-hand side of (S.12). By (S.11), Burkholder's inequality, and Hölder's inequality, we have

$$\mathbb{E}\left[\left|n^{-1/2}\sum_t\big[\big(p_l\big(X_t(\theta_1)\big)-p_l\big(X_t(\theta_2)\big)\big]\big(\widetilde{Z}_{t+1}(\theta_1)-\widetilde{Z}_{t+1}(\theta^\star)\big)\right|^p\right]$$

$$\leq K\mathbb{E}\bigg[\bigg|n^{-1}\sum_t\big[(p_l(X_t(\theta_1)) - p_l(X_t(\theta_2))\big]^2 n^{-1/2}L_t^2\bigg|^{p/2}\bigg]$$

$$\leq Kn^{-p/4}\mathbb{E}\bigg[\bigg|\zeta_{1,n}^2\|\theta_1-\theta_2\|^2\Big(n^{-1}\sum_t L_{X,t}^2 L_t^2\Big)\bigg|^{p/2}\bigg]$$

$$\leq Kn^{-p/4}\zeta_{1,n}^p\mathbb{E}\bigg[n^{-1}\sum_t L_{X,t}^p L_t^p\bigg]\|\theta_1-\theta_2\|^p.$$

Since $\mathbb{E}[L_{X,t}^p L_t^p] \leq K$ by Assumption 1, we can bound the L_p-norm of the first term in the decomposition (S.12) as follows:

$$\bigg\|n^{-1/2}\sum_t\big[(p_l(X_t(\theta_1)) - p_l(X_t(\theta_2))\big]\big(\widetilde{Z}_{t+1}(\theta_1) - \widetilde{Z}_{t+1}(\theta^\star)\big)\bigg\|_p$$
$$\leq Kn^{-1/4}\zeta_{1,n}\|\theta_1-\theta_2\|. \tag{S.13}$$

Turning to the second term in the decomposition (S.12), we note that

$$\big(Z_{t+1}(\theta_1) - Z_{t+1}(\theta_2)\big)^2 \leq 1_{\{Y_{t+1}\leq f_t(\theta_1)+|f_t(\theta_1)-f_t(\theta_2)|\}} - 1_{\{Y_{t+1}\leq f_t(\theta_1)-|f_t(\theta_1)-f_t(\theta_2)|\}}.$$

Hence,

$$\mathbb{E}\big[\big(\widetilde{Z}_{t+1}(\theta_1) - \widetilde{Z}_{t+1}(\theta_2)\big)^2|\mathcal{F}_t\big] \leq 2L_t\big|f_t(\theta_1) - f_t(\theta_2)\big| \leq 2L_t^2\|\theta_1-\theta_2\|. \tag{S.14}$$

By (S.14), Burkholder's inequality, and Hölder's inequality,

$$\mathbb{E}\bigg[\bigg|n^{-1/2}\sum_t p_l\big(X_t(\theta_2)\big)\big(\widetilde{Z}_{t+1}(\theta_1) - \widetilde{Z}_{t+1}(\theta_2)\big)\bigg|^p\bigg]$$
$$\leq K\mathbb{E}\bigg[\bigg|n^{-1}\sum_t p_l\big(X_t(\theta_2)\big)^2\mathbb{E}\big[\big(\widetilde{Z}_{t+1}(\theta_1) - \widetilde{Z}_{t+1}(\theta_2)\big)^2|\mathcal{F}_t\big]\bigg|^{p/2}\bigg]$$
$$\leq K\|\theta_1-\theta_2\|^{p/2}\mathbb{E}\bigg[\bigg|n^{-1}\sum_t p_l\big(X_t(\theta_2)\big)^2 L_t^2\bigg|^{p/2}\bigg]$$
$$\leq K\zeta_{0,n}^p\|\theta_1-\theta_2\|^{p/2}n^{-1}\sum_t\mathbb{E}\big[L_t^p\big] \leq K\zeta_{0,n}^p\|\theta_1-\theta_2\|^{p/2}.$$

Hence,

$$\bigg\|n^{-1/2}\sum_t p_l\big(X_t(\theta_2)\big)\big(\widetilde{Z}_{t+1}(\theta_1) - \widetilde{Z}_{t+1}(\theta_2)\big)\bigg\|_p \leq K\zeta_{0,n}\|\theta_1-\theta_2\|^{1/2}. \tag{S.15}$$

Since $\|\theta_1-\theta_2\| \leq Rn^{-1/2}$, the assertion in (S.10) readily follows from (S.13) and (S.15).

Step 3. We shall use a chaining argument to establish (S.9). Construct nested sets $\Theta_{0,n} \subset \Theta_{1,n} \subset \cdots \subset B_n(R)$ such that $\Theta_{0,n} = \{\theta^\star\}$ and for each $j \geq 1$, $\Theta_{j,n}$ is a maximal set of points such that each pair of distinct elements in $\Theta_{j,n}$ has distance greater than $Rn^{-1/2}2^{-j}$. Note that the number of points in Θ_j is less than $C(2^j)^{d_\theta}$ for some constant

$C > 0$ that does not depend on j. Link every point $\theta_{j+1} \in \Theta_{j+1}$ to a unique $\theta_j \in \Theta_j$ such that $\|\theta_{j+1} - \theta_j\| \le Rn^{-1/2}2^{-j}$. Then for any $J \ge 0$ and $\theta_{J+1} \in \Theta_{J+1}$, we can construct a chain $\theta_{J+1}, \ldots, \theta_0$ to $\theta_0 = \theta^\star$, and hence, by the triangle inequality

$$\left|\pi_{l,n}(\theta_{J+1})\right| = \left|\sum_{j=0}^{J}\left[\pi_{l,n}(\theta_{j+1}) - \pi_{l,n}(\theta_j)\right]\right| \le \sum_{j=0}^{J}\max\left|\pi_{l,n}(\theta_{j+1}) - \pi_{l,n}(\theta_j)\right| \tag{S.16}$$

where, for each j, the maximum is taken over all links (θ_{j+1}, θ_j) from Θ_{j+1} to Θ_j (with the total number less than $C(2^{j+1})^{d_\theta}$). We then observe

$$\begin{aligned}
\left\|\max_{\theta\in\Theta_{J+1}}\left|\pi_{l,n}(\theta)\right|\right\|_p &\le \sum_{j=0}^{J}\left\|\max\left|\pi_{l,n}(\theta_{j+1}) - \pi_{l,n}(\theta_j)\right|\right\|_p \\
&\le K\sum_{j=0}^{J}\left(2^j\right)^{d_\theta/p}\max\left\|\pi_{l,n}(\theta_{j+1}) - \pi_{l,n}(\theta_j)\right\|_p \\
&\le K\sum_{j=0}^{J}\left(2^j\right)^{d_\theta/p}\left(\zeta_{1,n}n^{-1/2} + \zeta_{0,n}\right)\left(Rn^{-1/2}2^{-j}\right)^{1/2} \\
&\le K\left(\zeta_{1,n}n^{-3/4} + \zeta_{0,n}n^{-1/4}\right)
\end{aligned}$$

where the first inequality is by (S.16); the second inequality is by a maximal inequality under the L_p-norm; the third inequality follows from (S.10); and the last inequality holds because $\sum_j(2^j)^{d_\theta/p}(2^{-j})^{1/2} < \infty$ as implied by $p > 2d_\theta$. Since the stochastic process $\pi_{l,n}(\theta)$ indexed by θ is separable, by letting $J \to \infty$, we further have

$$\left\|\sup_{\theta\in B_n(R)}\left|\pi_{l,n}(\theta)\right|\right\|_p \le K\left(\zeta_{1,n}n^{-3/4} + \zeta_{0,n}n^{-1/4}\right). \tag{S.17}$$

Finally, note that $\|n^{-1/2}\sum_t P(X_t(\theta))(\widetilde{Z}_{t+1}(\theta) - \widetilde{Z}_{t+1}(\theta^\star))\|^2 = \sum_{l=1}^{m_n}\pi_{l,n}(\theta)^2$. Therefore,

$$\mathbb{E}\left[\sup_{\theta\in B_n(R)}\left\|n^{-1/2}\sum_t P\left(X_t(\theta)\right)\left(\widetilde{Z}_{t+1}(\theta) - \widetilde{Z}_{t+1}\left(\theta^\star\right)\right)\right\|^2\right] \le \sum_{l=1}^{m_n}\left\|\sup_{\theta\in B_n(R)}\left|\pi_{l,n}(\theta)\right|\right\|_p^2.$$

The assertion in (S.9) then readily follows from this estimate and (S.17). □

Lemma S5. *Suppose that Assumptions* 1 *and* 2 *hold. Under the null hypothesis, we have for any* $R > 0$,

$$\sup_{\theta\in B_n(R)}\left\|n^{-1}\sum_t\left(P\left(X_t(\theta)\right) - P\left(X_t\left(\theta^\star\right)\right)\right)u_t^\star\right\| = O_p\left(\zeta_{1,n}m_n^{1/2}n^{-1}\right).$$

Proof. We set $\pi_{l,n}(\theta) \equiv n^{-1/2}\sum_t(p_l(X_t(\theta)) - p_l(X_t(\theta^\star)))u_t^\star$. Note that under the null hypothesis, $(p_l(X_t(\theta)) - p_l(X_t(\theta^\star)))u_t^\star$ forms a martingale difference sequence. For any

$\theta_1, \theta_2 \in \Theta$, we observe

$$\mathbb{E}\big[\big|\pi_{l,n}(\theta_1) - \pi_{l,n}(\theta_2)\big|^p\big] \le K\mathbb{E}\Bigg[\Bigg|n^{-1}\sum_t \big(p_l\big(X_t(\theta_1)\big) - p_l\big(X_t(\theta_2)\big)\big)^2\Bigg|^{p/2}\Bigg]$$
$$\le K\zeta_{1,n}^p\|\theta_1 - \theta_2\|^p n^{-1}\sum_t \mathbb{E}\big[L_{X,t}^p\big] \le K\zeta_{1,n}^p\|\theta_1 - \theta_2\|^p,$$

where the first inequality is by Burkholder's inequality and the boundedness of $u_t^\star$, and the second line follows from Assumptions 1 and 2. Hence,

$$\|\pi_{l,n}(\theta_1) - \pi_{l,n}(\theta_2)\|_p \le K\zeta_{1,n}\|\theta_1 - \theta_2\|. \tag{S.18}$$

Construct $\Theta_{0,n} \subset \Theta_{1,n} \subset \cdots \subset B_n(R)$ as in step 3 in the proof of Lemma S4. Using the same chaining argument but with (S.10) replaced by (S.18), we deduce that

$$\Big\|\max_{\theta\in\Theta_{J+1}}\big|\pi_{l,n}(\theta)\big|\Big\|_p \le \sum_{j=0}^{J}\Big\|\max\big|\pi_{l,n}(\theta_{j+1}) - \pi_{l,n}(\theta_j)\big|\Big\|_p$$
$$\le K\sum_{j=0}^{J}\big(2^j\big)^{d_\theta/p}\max\big\|\pi_{l,n}(\theta_{j+1}) - \pi_{l,n}(\theta_j)\big\|_p$$
$$\le K\zeta_{1,n}n^{-1/2}\sum_{j=0}^{J}\big(2^{-j}\big)^{1-d_\theta/p} \le K\zeta_{1,n}n^{-1/2}.$$

Sending $J \to \infty$, we further deduce

$$\Big\|\sup_{\theta\in B_n(R)}\big|\pi_{l,n}(\theta)\big|\Big\|_p \le K\zeta_{1,n}n^{-1/2}. \tag{S.19}$$

Finally, note that $\|n^{-1}\sum_t(P(X_t(\theta)) - P(X_t(\theta^\star)))u_t^\star\|^2 = \sum_{l=1}^{m_n}\pi_{l,n}(\theta)^2$. Therefore,

$$\mathbb{E}\Bigg[\sup_{\theta\in B_n(R)}\Bigg\|n^{-1/2}\sum_t\big(P\big(X_t(\theta)\big) - P\big(X_t\big(\theta^\star\big)\big)\big)u_t^\star\Bigg\|^2\Bigg] \le \sum_{l=1}^{m_n}\Big\|\sup_{\theta\in B_n(R)}\big|\pi_{l,n}(\theta)\big|\Big\|_p^2.$$

The assertion of the lemma then readily follows from (S.19). □

Lemma S6. *Suppose that Assumptions* 1 *and* 2 *hold. Then we have* $\|\widehat{b}_n - b_n^\star\| = O_p(\delta_{b,n})$, *where*

$$\delta_{b,n} = \begin{cases} llm_n^{1/2}n^{-1/2} & \text{under the null,} \\ \zeta_{1,n}m_n^{1/2}n^{-1/2} + \zeta_{0,n}m_n^{1/2}n^{-3/4} & \text{in general.} \end{cases}$$

Proof. By Lemma S2, $\|\widehat{Q}_n - Q_n\|_S = o_p(1)$. Since the eigenvalues of Q_n are bounded from above and away from zero, we further have

$$\lambda_{\min}^{-1}(\widehat{Q}_n) + \lambda_{\max}(\widehat{Q}_n) = O_p(1). \tag{S.20}$$

Recall that $u_t^\star = Z_{t+1}^\star - \mathbb{E}[Z_{t+1}^\star | X_t^\star]$. By the definition of $\widehat{b}_n$, we can decompose

$$\begin{aligned}\widehat{b}_n - b_n^\star = \widehat{Q}_n^{-1}\Big(n^{-1}\sum_t P(X_t^\star)u_t^\star\Big) &+ \widehat{Q}_n^{-1}\Big(n^{-1}\sum_t \big(P(\widehat{X}_t) - P(X_t^\star)\big)u_t^\star\Big)\\ &+ \widehat{Q}_n^{-1} n^{-1}\sum_t P(\widehat{X}_t)\big(h(X_t^\star) - P(\widehat{X}_t)^\top b_n^\star\big)\\ &+ \widehat{Q}_n^{-1}\Big(n^{-1}\sum_t P(\widehat{X}_t)\big(\widehat{Z}_{t+1} - Z_{t+1}^\star\big)\Big). \end{aligned} \tag{S.21}$$

It remains to bound the four terms on the right-hand side of this decomposition.

First, recall that $\overline{A}_n = \mathrm{Var}(n^{-1/2}\sum_t P(X_t^\star)u_t^\star)$ has bounded eigenvalues (Assumption 2). Hence, $\mathbb{E}[\|n^{-1/2}\sum_t P(X_t^\star)u_t^\star\|^2] = \mathrm{Trace}(\overline{A}_n) \le K m_n$, which, combined with (S.20), implies that

$$\widehat{Q}_n^{-1}\Big(n^{-1}\sum_t P(X_t^\star)u_t^\star\Big) = O_p\big(m_n^{1/2} n^{-1/2}\big). \tag{S.22}$$

Second, by the Cauchy–Schwarz inequality, the boundedness of $u_t^\star$, and Lemma S1,

$$\Big\| n^{-1}\sum_t \big(P(\widehat{X}_t) - P(X_t^\star)\big)u_t^\star\Big\|^2 \le K n^{-1}\sum_t \big\|P(\widehat{X}_t) - P(X_t^\star)\big\|^2 = O_p\big(\zeta_{1,n}^2 m_n n^{-1}\big).$$

In addition, under the null hypothesis, we can apply Lemma S5 to get

$$n^{-1}\sum_t \big(P(\widehat{X}_t) - P(X_t^\star)\big)u_t^\star = O_p\big(\zeta_{1,n} m_n^{1/2} n^{-1}\big). \tag{S.23}$$

Hence,

$$\begin{aligned}&\widehat{Q}_n^{-1}\Big(n^{-1}\sum_t \big(P(\widehat{X}_t) - P(X_t^\star)\big)u_t^\star\Big)\\ &\quad = \begin{cases} O_p\big(\zeta_{1,n} m_n^{1/2} n^{-1}\big) & \text{under the null,}\\ O_p\big(\zeta_{1,n} m_n^{1/2} n^{-1/2}\big) & \text{in general.}\end{cases}\end{aligned} \tag{S.24}$$

Third, we note that

$$\begin{aligned}&\Big\|\widehat{Q}_n^{-1} n^{-1}\sum_t P(\widehat{X}_t)\big(h(X_t^\star) - P(\widehat{X}_t)^\top b_n^\star\big)\Big\|^2\\ &\quad \le \lambda_{\min}^{-1}(\widehat{Q}_n) n^{-1}\sum_t \big(h(X_t^\star) - P(\widehat{X}_t)^\top b_n^\star\big)^2\\ &\quad \le 2\lambda_{\min}^{-1}(\widehat{Q}_n) n^{-1}\sum_t \big(h(\widehat{X}_t) - P(\widehat{X}_t)^\top b_n^\star\big)^2\\ &\qquad + 2\lambda_{\min}^{-1}(\widehat{Q}_n) n^{-1}\sum_t \big(h(\widehat{X}_t) - h(X_t^\star)\big)^2.\end{aligned}$$

Define Ω_n as in the proof of Lemma S1, so that $\widehat{X}_t \in \mathcal{X} \oplus \varepsilon_n$ for some $\varepsilon_n \asymp n^{1/p-1/2}$ in restriction to Ω_n. By (S.20) and Assumption 2, $\lambda_{\min}^{-1}(\widehat{Q}_n) n^{-1} \sum_t (h(\widehat{X}_t) - P(\widehat{X}_t)^\top b_n^\star)^2 = O_p(n^{-1})$. Moreover, since $h(\cdot)$ is continuously differentiable, it is Lipschitz on the compact set $\mathcal{X} \oplus \varepsilon_n$. Therefore,

$$\begin{aligned}\lambda_{\min}^{-1}(\widehat{Q}_n) n^{-1} \sum_t \big(h(\widehat{X}_t) - h\big(X_t^\star\big)\big)^2 &\leq O_p(1) n^{-1} \sum_t \big(\widehat{X}_t - X_t^\star\big)^2 \\ &\leq O_p(1)\bigg(n^{-1} \sum_t L_{X,t}^2\bigg) \big\| \hat{\theta}_n - \theta^\star \big\|^2 \\ &= O_p\big(n^{-1}\big).\end{aligned}$$

Combining the three estimates above yields

$$\bigg\| \widehat{Q}_n^{-1} n^{-1} \sum_t P(\widehat{X}_t)\big(h\big(X_t^\star\big) - P(\widehat{X}_t)^\top b_n^\star\big) \bigg\| = O_p\big(n^{-1/2}\big). \tag{S.25}$$

On the other hand, under the null hypothesis, we have $h(\cdot) = 0$ and $b_n^\star = 0$. We thus have

$$\begin{aligned}&\widehat{Q}_n^{-1} n^{-1} \sum_t P(\widehat{X}_t)\big(h\big(X_t^\star\big) - P(\widehat{X}_t)^\top b_n^\star\big) \\ &\quad = \begin{cases} 0 & \text{under the null,} \\ O_p\big(n^{-1/2}\big) & \text{in general.}\end{cases}\end{aligned} \tag{S.26}$$

Finally, by Lemma S4,

$$\begin{aligned}&\widehat{Q}_n^{-1}\bigg(n^{-1} \sum_t P(\widehat{X}_t)\big(\widehat{Z}_{t+1} - Z_{t+1}^\star\big)\bigg) \\ &\quad = O_p\big(n^{-1/2} + m_n^{1/2}\big(\zeta_{1,n} n^{-5/4} + \zeta_{0,n} n^{-3/4}\big)\big).\end{aligned} \tag{S.27}$$

Combining (S.22), (S.24), (S.26), and (S.27), we deduce that

$$\|\widehat{b}_n - b_n^\star\| = \begin{cases} O_p\big(m_n^{1/2} n^{-1/2} + \zeta_{1,n} m_n^{1/2} n^{-1} + \zeta_{0,n} m_n^{1/2} n^{-3/4}\big) & \text{under the null,} \\ O_p\big(\zeta_{1,n} m_n^{1/2} n^{-1/2} + \zeta_{0,n} m_n^{1/2} n^{-3/4}\big) & \text{in general.}\end{cases}$$

Under the maintained rate condition on $\zeta_{0,n}$ and $\zeta_{1,n}$ (see Assumption 2), we can further reduce the rates displayed above into those asserted in the lemma. □

LEMMA S7. *Suppose that Assumptions* 1 *and* 2 *hold. Then*

$$\|\widehat{A}_n - A_n\|_S = \begin{cases} O_p\big(\delta_{A,n} + \zeta_{1,n} m_n^{1/2} n^{-1/2} + \zeta_{0,n} m_n^{1/2} n^{-1/4}\big) & \textit{under the null,} \\ O_p\big(\delta_{A,n} + \zeta_{0,n} m_n^{1/2} n^{-1/4} + \zeta_{0,n} \zeta_{1,n} m_n n^{-1/2}\big) & \textit{in general.}\end{cases}$$

In particular, $\|\widehat{A}_n - A_n\|_S = o_p(1)$.

PROOF. Step 1. We outline the proof in this step. Recall that $\widehat{A}_n^\star \equiv n^{-1}\sum_t u_t^{\star 2} P(X_t^\star) \times P(X_t^\star)^\top$. By the triangle inequality,

$$\begin{aligned}
\|\widehat{A}_n - \widehat{A}_n^\star\| &= \left\| n^{-1}\sum_t \widehat{u}_t^2 P(\widehat{X}_t)P(\widehat{X}_t)^\top - n^{-1}\sum_t u_t^{\star 2} P(X_t^\star)P(X_t^\star)^\top \right\| \\
&\leq \left\| n^{-1}\sum_t u_t^{\star 2}\big(P(\widehat{X}_t) - P(X_t^\star)\big)\big(P(\widehat{X}_t) - P(X_t^\star)\big)^\top \right\| \\
&\quad + 2\left\| n^{-1}\sum_t u_t^{\star 2} P(X_t^\star)\big(P(\widehat{X}_t) - P(X_t^\star)\big)^\top \right\| \\
&\quad + \left\| n^{-1}\sum_t \big(\widehat{u}_t^2 - u_t^{\star 2}\big)P(\widehat{X}_t)P(\widehat{X}_t)^\top \right\|. \qquad \text{(S.28)}
\end{aligned}$$

Below, we bound the three terms on the majorant side of (S.28) in turn.

We start with the first two terms. Note that $u_t^\star$ is bounded. By the triangle inequality, the Cauchy–Schwarz inequality, and Lemma S1,

$$\begin{aligned}
&\left\| n^{-1}\sum_t u_t^{\star 2}\big(P(\widehat{X}_t) - P(X_t^\star)\big)\big(P(\widehat{X}_t) - P(X_t^\star)\big)^\top \right\| \\
&\quad \leq K n^{-1}\sum_t \big\|P(\widehat{X}_t) - P(X_t^\star)\big\|^2 = O_p\big(\zeta_{1,n}^2 m_n n^{-1}\big). \qquad \text{(S.29)}
\end{aligned}$$

Using Lemma S1, we can also deduce that

$$\begin{aligned}
&\left\| n^{-1}\sum_t u_t^{\star 2} P(X_t^\star)\big(P(\widehat{X}_t) - P(X_t^\star)\big)^\top \right\| \\
&= \left(\sum_{l=1}^{m_n} \left\| n^{-1}\sum_t u_t^{\star 2} P(X_t^\star)\big(p_l(\widehat{X}_t) - p_l(X_t^\star)\big) \right\|^2 \right)^{1/2} \\
&\leq \left(\sum_{l=1}^{m_n} \lambda_{\max}\big(\widehat{Q}_n^\star\big) n^{-1}\sum_t \big\| u_t^{\star 2}\big(p_l(\widehat{X}_t) - p_l(X_t^\star)\big)\big\|^2 \right)^{1/2} \\
&\leq O_p(1)\cdot\left(n^{-1}\sum_t \big\|P(\widehat{X}_t) - P(X_t^\star)\big\|^2 \right)^{1/2} \\
&= O_p\big(\zeta_{1,n} m_n^{1/2} n^{-1/2}\big). \qquad \text{(S.30)}
\end{aligned}$$

Turning to the third term on the majorant side of (S.28), we note that

$$\begin{aligned}
&\left\| n^{-1}\sum_t \big(\widehat{u}_t^2 - u_t^{\star 2}\big)P(\widehat{X}_t)P(\widehat{X}_t)^\top \right\|^2 \\
&\quad = \sum_{l=1}^{m_n} \left\| n^{-1}\sum_t P(\widehat{X}_t)p_l(\widehat{X}_t)\big(\widehat{u}_t^2 - u_t^{\star 2}\big) \right\|^2
\end{aligned}$$

$$\leq \sum_{l=1}^{m_n} \lambda_{\max}(\widehat{Q}_n) n^{-1} \sum_t p_l(\widehat{X}_t)^2 \big(\widehat{u}_t^2 - u_t^{\star 2}\big)^2$$

$$\leq \zeta_{0,n}^2 m_n \lambda_{\max}(\widehat{Q}_n) n^{-1} \sum_{t=1}^{n} \big(\widehat{u}_t^2 - u_t^{\star 2}\big)^2. \tag{S.31}$$

In steps 2–4, below, we shall show that

$$n^{-1} \sum_t \big(\widehat{u}_t^2 - u_t^{\star 2}\big)^2 = O_p\big(n^{-1/2} + \delta_{b,n}^2 + \zeta_{0,n}^2 m_n \delta_{b,n}^4\big). \tag{S.32}$$

Then (S.31) and (S.32) imply

$$\left\| n^{-1} \sum_t \big(\widehat{u}_t^2 - u_t^{\star 2}\big) P(\widehat{X}_t) P(\widehat{X}_t)^\top \right\| = O_p\big(\zeta_{0,n} m_n^{1/2} n^{-1/4} + \zeta_{0,n} m_n^{1/2} \delta_{b,n} + \zeta_{0,n}^2 m_n \delta_{b,n}^2\big)$$

$$= O_p\big(\zeta_{0,n} m_n^{1/2} n^{-1/4} + \zeta_{0,n} m_n^{1/2} \delta_{b,n}\big), \tag{S.33}$$

where $\delta_{b,n}$ is defined in Lemma S6, and the second line follows from $\zeta_{0,n} m_n^{1/2} \delta_{b,n} = o(1)$, which is implied by Assumption 2. Therefore, by (S.29), (S.30), and (S.33),

$$\|\widehat{A}_n - \widehat{A}_n^\star\| = O_p\big(\zeta_{1,n} m_n^{1/2} n^{-1/2} + \zeta_{0,n} m_n^{1/2} n^{-1/4} + \zeta_{0,n} m_n^{1/2} \delta_{b,n}\big). \tag{S.34}$$

Recalling the definition of $\delta_{b,n}$ from Lemma S6, we can further simplify the rate in (S.34) via some elementary calculations, yielding

$$\|\widehat{A}_n - \widehat{A}_n^\star\| = \begin{cases} llO_p\big(\zeta_{0,n} m_n^{1/2} n^{-1/4} + \zeta_{1,n} m_n^{1/2} n^{-1/2}\big) & \text{under the null,} \\ O_p\big(\zeta_{0,n} m_n^{1/2} n^{-1/4} + \zeta_{0,n} \zeta_{1,n} m_n n^{-1/2}\big) & \text{in general.} \end{cases}$$

The assertion of the lemma then follows from this estimate and the assumption that $\|\widehat{A}_n^\star - A_n\|_S = O_p(\delta_{A,n})$. The remaining steps below are devoted to proving (S.32).

Step 2. We collect some technical estimates in this step. Since $\|\hat{\theta}_n - \theta^\star\| = O_p(n^{-1/2})$, we can apply Lemma S3 to get

$$n^{-1} \sum_t \big|\widehat{Z}_{t+1} - Z_{t+1}^\star\big|^2 = O_p\big(n^{-1/2}\big). \tag{S.35}$$

Since $\widehat{Z}_{t+1}$ and $Z_{t+1}^\star$ are bounded, this estimate further implies

$$n^{-1} \sum_t \big|\widehat{Z}_{t+1} - Z_{t+1}^\star\big|^4 \leq n^{-1} \sum_t \big|\widehat{Z}_{t+1} - Z_{t+1}^\star\big|^2 = O_p\big(n^{-1/2}\big). \tag{S.36}$$

Denote $h_n(\cdot) = P(\cdot)^\top b_n^\star$. By Lemma S6 and (S.20),

$$n^{-1} \sum_t \big|\widehat{h}_n(\widehat{X}_t) - h_n(\widehat{X}_t)\big|^2 \leq \lambda_{\max}(\widehat{Q}_n) \big\|\widehat{b}_n - b_n^\star\big\|^2 = O_p\big(\delta_{b,n}^2\big). \tag{S.37}$$

Note that for some $\varepsilon_n \asymp n^{1/p-1/2}$, $\{\widehat{X}_t : 1 \leq t \leq n\} \in \mathcal{X} \oplus \varepsilon_n$ with probability arbitrarily close to one and, in restriction to this event, we have $\|P(\widehat{X}_t)\|^2 \leq \zeta_{0,n}^2 m_n$ by the definition of $\zeta_{0,n}$. Then, by Lemma S6 and (S.37), we have

$$n^{-1}\sum_t \left|\widehat{h}_n(\widehat{X}_t)-h_n(\widehat{X}_t)\right|^4$$
$$\leq \zeta_{0,n}^2 m_n \left\|\widehat{b}_n-b_n^\star\right\|^2 n^{-1}\sum_{t=1}^n \left|\widehat{h}_n(\widehat{X}_t)-h_n(\widehat{X}_t)\right|^2 = O_p\left(\zeta_{0,n}^2 m_n \delta_{b,n}^4\right). \tag{S.38}$$

Finally, we note that (recalling $\widehat{u}_t=\widehat{Z}_{t+1}-\widehat{h}_n(\widehat{X}_t)$ and $u_t^\star = Z_{t+1}^\star - h(X_t^\star)$)

$$\left|\widehat{u}_t-u_t^\star\right| \leq \left|\widehat{Z}_{t+1}-Z_{t+1}^\star\right| + \left|\widehat{h}_n(\widehat{X}_t)-h_n(\widehat{X}_t)\right|$$
$$+\left|h_n(\widehat{X}_t)-h(\widehat{X}_t)\right| + \left|h(\widehat{X}_t)-h\left(X_t^\star\right)\right|. \tag{S.39}$$

Step 3. In this step, we prove (S.32) under the null hypothesis. In this case, $h(\cdot)=0$ and $h_n(\cdot)=0$. Hence, by (S.39), $|\widehat{u}_t-u_t^\star| \leq |\widehat{Z}_{t+1}-Z_{t+1}^\star| + |\widehat{h}_n(\widehat{X}_t)-h_n(\widehat{X}_t)|$. By (S.35) and (S.37), we have

$$n^{-1}\sum_t \left|\widehat{u}_t-u_t^\star\right|^2 \leq Kn^{-1}\sum_t \left|\widehat{Z}_{t+1}-Z_{t+1}^\star\right|^2 + Kn^{-1}\sum_t \left|\widehat{h}_n(\widehat{X}_t)-h_n(\widehat{X}_t)\right|^2$$
$$= O_p\left(n^{-1/2}\right)+O_p\left(\delta_{b,n}^2\right). \tag{S.40}$$

Similarly, by (S.36) and (S.38), we have

$$n^{-1}\sum_t \left|\widehat{u}_t-u_t^\star\right|^4 = O_p\left(n^{-1/2}\right)+O_p\left(\zeta_{0,n}^2 m_n \delta_{b,n}^4\right). \tag{S.41}$$

Consider the following elementary inequality: for any $|x|\leq 1$ and $y\in\mathbb{R}$,

$$\left((x+y)^2-x^2\right)^2 = \left(2xy+y^2\right)^2 \leq 8x^2y^2+2y^4 \leq K\left(y^2+y^4\right). \tag{S.42}$$

Applying this inequality with $x=u_t^\star$ and $y=\widehat{u}_t-u_t^\star$, we deduce

$$n^{-1}\sum_t \left(\widehat{u}_t^2-u_t^{\star 2}\right)^2 \leq Kn^{-1}\sum_t \left|\widehat{u}_t-u_t^\star\right|^2 + Kn^{-1}\sum_t \left|\widehat{u}_t-u_t^\star\right|^4$$
$$= O_p\left(n^{-1/2}+\delta_{b,n}^2+\zeta_{0,n}^2 m_n \delta_{b,n}^4\right). \tag{S.43}$$

This completes the proof of (S.32) under the null hypothesis.

Step 4. In this step, we prove (S.32) in the general case without imposing the null hypothesis. We first observe that

$$n^{-1}\sum_t \left|h(\widehat{X}_t)-h\left(X_t^\star\right)\right|^4 \leq n^{-1}\sum_t \left|h(\widehat{X}_t)-h\left(X_t^\star\right)\right|^2$$
$$\leq K\left\|\hat{\theta}_n-\theta^\star\right\|^2 n^{-1}\sum_t L_{X,t}^2 = O_p\left(n^{-1}\right), \tag{S.44}$$

where the first inequality holds because $h(\cdot)$ is bounded, and the second inequality follows from the Lipschitz continuity of $h(\cdot)$ and $\|X_t(\theta)-X_t^\star\| \leq L_{X,t}\|\theta-\theta^\star\|$. By (S.39), we have for $q = 2$ or 4,

$$n^{-1}\sum_t \left|\widehat{u}_t - u_t^\star\right|^q \le Kn^{-1}\sum_t \left|\widehat{Z}_{t+1} - Z_{t+1}^\star\right|^q + Kn^{-1}\sum_t \left|\widehat{h}_n(\widehat{X}_t) - h_n(\widehat{X}_t)\right|^q$$
$$+ Kn^{-1}\sum_t \left|h_n(\widehat{X}_t) - h(\widehat{X}_t)\right|^q + Kn^{-1}\sum_t \left|h(\widehat{X}_t) - h\big(X_t^\star\big)\right|^q.$$

Then, by (S.35), (S.37), (S.44), and Assumption 2(ii), we deduce

$$n^{-1}\sum_t \left|\widehat{u}_t - u_t^\star\right|^2 = O_p\big(n^{-1/2} + \delta_{b,n}^2\big). \tag{S.45}$$

Similarly, by (S.36), (S.38), (S.44), and Assumption 2(ii), we deduce

$$n^{-1}\sum_t \left|\widehat{u}_t - u_t^\star\right|^4 = O_p\big(n^{-1/2} + \zeta_{0,n}^2 m_n \delta_{b,n}^4\big). \tag{S.46}$$

Using (S.42), (S.45), and (S.46), we derive

$$n^{-1}\sum_t \big(\widehat{u}_t^2 - u_t^{\star 2}\big)^2 \le Kn^{-1}\sum_t \left|\widehat{u}_t - u_t^\star\right|^2 + Kn^{-1}\sum_t \left|\widehat{u}_t - u_t^\star\right|^4$$
$$= O_p\big(n^{-1/2} + \delta_{b,n}^2 + \zeta_{0,n}^2 m_n \delta_{b,n}^4\big).$$

This completes the proof of (S.32), and hence, the assertion of the lemma. □

LEMMA S8. *Suppose that Assumptions* 1 *and* 2 *hold. Then*

$$\sup_{x\in\mathcal{X}} \frac{\left|P(x)^\top \widehat{Q}_n^{-1} n^{-1/2}\sum_t P(\widehat{X}_t)\big(\widehat{Z}_{t+1} - Z_{t+1}^\star\big)\right|}{\sigma_n(x)} = o_p\big((\log n)^{-1/2}\big).$$

PROOF. Step 1**.** We outline the proof in this step. Recall $\bar{Z}_{t+1}(\theta) = \mathbb{E}[Z_{t+1}(\theta)|\mathcal{F}_t]$ and set $\widetilde{Z}_{t+1}(\theta) = Z_{t+1}(\theta) - \bar{Z}_{t+1}(\theta)$. For ease of notation, we denote $\bar{Z}'_{t+1} = \partial_\theta \bar{Z}_{t+1}(\theta^\star)$. Our proof relies on the following decomposition:

$$\frac{P(x)^\top \widehat{Q}_n^{-1} n^{-1/2}\sum_t P(\widehat{X}_t)\big(\widehat{Z}_{t+1} - Z_{t+1}^\star\big)}{\sigma_n(x)} = \sum_{j=1}^{3} R_{j,n}(x), \tag{S.47}$$

where

$$R_{1,n}(x) \equiv \frac{P(x)^\top \widehat{Q}_n^{-1} n^{-1/2}\sum_t P(\widehat{X}_t)\big(\widetilde{Z}_{t+1}(\hat{\theta}_n) - \widetilde{Z}_{t+1}\big(\theta^\star\big)\big)}{\sigma_n(x)},$$

$$R_{2,n}(x) \equiv \frac{P(x)^\top \widehat{Q}_n^{-1} n^{-1/2}\sum_t \big\{P(\widehat{X}_t)\big(\bar{Z}_{t+1}(\hat{\theta}_n) - \bar{Z}_{t+1}\big(\theta^\star\big)\big) - P\big(X_t^\star\big)\bar{Z}_{t+1}^{\prime\top}\big(\hat{\theta}_n - \theta^\star\big)\big\}}{\sigma_n(x)},$$

$$R_{3,n}(x) \equiv \frac{P(x)^\top \widehat{Q}_n^{-1} n^{-1/2} \sum_t P\big(X_t^\star\big) \bar{Z}_{t+1}'^\top \big(\hat{\theta}_n - \theta^\star\big)}{\sigma_n(x)}.$$

By (S.9), we have

$$\left\| n^{-1/2} \sum_t P(\widehat{X}_t)\big(\widetilde{Z}_{t+1}(\hat{\theta}_n) - \widetilde{Z}_{t+1}(\theta^\star)\big) \right\| = O_p\big(\zeta_{1,n} m_n^{1/2} n^{-3/4} + \zeta_{0,n} m_n^{1/2} n^{-1/4}\big),$$

which further implies

$$\begin{aligned} \sup_{x\in\mathcal{X}} \big|R_{1,n}(x)\big| &\le O_p\big(\zeta_{1,n} m_n^{1/2} n^{-3/4} + \zeta_{0,n} m_n^{1/2} n^{-1/4}\big) \\ &= o_p\big((\log n)^{-1/2}\big). \end{aligned} \tag{S.48}$$

In steps 2 and 3 below, we show that

$$\sup_{x\in\mathcal{X}} \big|R_{j,n}(x)\big| = o_p\big((\log n)^{-1/2}\big), \quad \text{for } j = 2, 3. \tag{S.49}$$

The assertion of the lemma then follows from (S.47), (S.48), and (S.49).

Step 2. In this step, we prove (S.49) for the $j = 2$ case. We first observe that, for any $R > 0$,

$$\begin{aligned} &\sup_{\theta\in B_n(R)} \sup_{x\in\mathcal{X}} \frac{\left| P(x)^\top \widehat{Q}_n^{-1} n^{-1/2} \sum_t P(\widehat{X}_t)\big(\bar{Z}_{t+1}(\theta) - \bar{Z}_{t+1}(\theta^\star) - \bar{Z}_{t+1}'^\top(\theta - \theta^\star)\big)\right|}{\sigma_n(x)} \\ &\le O_p(1) \cdot \sup_{\theta\in B_n(R)} \left\| (n\widehat{Q}_n)^{-1/2} \sum_t P(\widehat{X}_t)\big(\bar{Z}_{t+1}(\theta) - \bar{Z}_{t+1}(\theta^\star) - \bar{Z}_{t+1}'^\top(\theta - \theta^\star)\big)\right\| \\ &\le O_p(1) \cdot \sup_{\theta\in B_n(R)} \left(\sum_t \big(\bar{Z}_{t+1}(\theta) - \bar{Z}_{t+1}(\theta^\star) - \bar{Z}_{t+1}'^\top(\theta - \theta^\star)\big)^2 \right)^{1/2} \\ &\le O_p(1) \cdot \left(n^{-2} \sum_t \bar{L}_t^2 \right)^{1/2} = O_p\big(n^{-1/2}\big) = o_p\big((\log n)^{-1/2}\big), \end{aligned}$$

where the first inequality follows from the fact that $\lambda_{\min}(\Sigma_n)$ is bounded away from zero; the second inequality follows from the contraction property of the least-square projection; and the last line follows from the definition of $B_n(R)$ and the L_2-boundedness of $\bar{L}_t$. Since $\hat{\theta}_n - \theta^\star = O_p(n^{-1/2})$, this estimate further implies

$$\begin{aligned} &\sup_{x\in\mathcal{X}} \frac{\left| P(x)^\top \widehat{Q}_n^{-1} n^{-1/2} \sum_t P(\widehat{X}_t)\big(\bar{Z}_{t+1}(\hat{\theta}_n) - \bar{Z}_{t+1}(\theta^\star) - \bar{Z}_{t+1}'^\top(\hat{\theta}_n - \theta^\star)\big)\right|}{\sigma_n(x)} \\ &= o_p\big((\log n)^{-1/2}\big). \end{aligned} \tag{S.50}$$

We further observe that

$$\begin{aligned}
&\sup_{x\in\mathcal{X}} \frac{\left\| P(x)^\top \widehat{Q}_n^{-1} n^{-1} \sum_t \big(P(\widehat{X}_t) - P(X_t^\star)\big) \bar{Z}_{t+1}'^\top \right\|}{\sigma_n(x)} \\
&\leq O_p(1) \left\| n^{-1} \sum_t \big(P(\widehat{X}_t) - P(X_t^\star)\big) \bar{Z}_{t+1}'^\top \right\| \\
&\leq O_p(1) \left(n^{-1} \sum_t \| P(\widehat{X}_t) - P(X_t^\star) \|^2 \right)^{1/2} \left(n^{-1} \sum_t \| \bar{Z}_{t+1}' \|^2 \right)^{1/2} \\
&= O_p\big(\zeta_{1,n} m_n^{1/2} n^{-1/2}\big) = o_p\big((\log n)^{-1/2}\big),
\end{aligned}$$

where the first inequality follows from the fact that $\lambda_{\min}(\widehat{Q}_n)$ and $\lambda_{\min}(\Sigma_n)$ are bounded away from zero with probability approaching 1; the second inequality is by the Cauchy–Schwarz inequality; and the last line follows from Lemma S1. Then, by the Cauchy–Schwarz inequality, we further deduce

$$\begin{aligned}
&\sup_{x\in\mathcal{X}} \frac{\left| P(x)^\top \widehat{Q}_n^{-1} n^{-1/2} \sum_t \big(P(\widehat{X}_t) - P(X_t^\star)\big) \partial_\theta \bar{Z}_{t+1}(\theta^\star)^\top (\hat{\theta}_n - \theta^\star) \right|}{\sigma_n(x)} \\
&= o_p\big((\log n)^{-1/2}\big). \qquad \text{(S.51)}
\end{aligned}$$

The claim in (S.49) for the case $j = 2$ readily follows from (S.50) and (S.51).

Step 3. In this step, we prove (S.49) for the $j = 3$ case. Recall the definition $\eta_t \equiv \bar{Z}_{t+1}' - \mathbb{E}[\bar{Z}_{t+1}' | X_t^\star]$. We then observe, for each $j \in \{1, \ldots, d_\theta\}$,

$$\begin{aligned}
\mathbb{E}\left[\left\| n^{-1} \sum_t P(X_t^\star) \eta_{j,t} \right\|^2 \right] &= n^{-1} \operatorname{Trace}\left(\operatorname{Var}\left(n^{-1/2} \sum_t P(X_t^\star) \eta_{j,t} \right) \right) \\
&\leq K m_n n^{-1}, \qquad \text{(S.52)}
\end{aligned}$$

which further implies that $\|n^{-1} \sum_t P(X_t^\star) \eta_t\| = O_p(m_n^{1/2} n^{-1/2})$. Hence,

$$\begin{aligned}
&\sup_{x\in\mathcal{X}} \frac{\left| P(x)^\top \widehat{Q}_n^{-1} n^{-1/2} \sum_t P(X_t^\star) \eta_t^\top (\hat{\theta}_n - \theta^\star) \right|}{\sigma_n(x)} \\
&\leq O_p(1) \left\| n^{-1} \sum_t P(X_t^\star) \eta_t^\top \right\| = O_p\big(m_n^{1/2} n^{-1/2}\big) = o_p\big((\log n)^{-1/2}\big). \qquad \text{(S.53)}
\end{aligned}$$

Recall that $g(X_t^\star) = \mathbb{E}[\bar{Z}_{t+1}' | X_t^\star]$. Hence, $\mathbb{E}[\|g(X_t^\star)\|^2] \leq \mathbb{E}[\|\bar{Z}_{t+1}'\|^2] \leq K$ and

$$n^{-1} \sum_t \| g(X_t^\star) \|^2 = O_p(1). \qquad \text{(S.54)}$$

By (S.54) and Lemma S2,

$$\begin{aligned}
&\sup_{x\in\mathcal{X}} \frac{\left|P(x)^\top(\widehat{Q}_n^{-1}-Q_n^{-1})n^{-1/2}\sum_t P(X_t^\star)g(X_t^\star)^\top(\hat{\theta}_n-\theta^\star)\right|}{\sigma_n(x)} \\
&\le O_p(1)\|\widehat{Q}_n - Q_n\|_S \left\| n^{-1}\sum_t P(X_t^\star)g(X_t^\star)^\top \right\| \\
&\le O_p(1)\|\widehat{Q}_n - Q_n\|_S \left(n^{-1}\sum_t \|g(X_t^\star)\|^2 \right)^{1/2} \\
&= O_p\big(\delta_{Q,n} + \zeta_{1,n} m_n^{1/2} n^{-1/2}\big) = o_p\big((\log n)^{-1/2}\big). \qquad \text{(S.55)}
\end{aligned}$$

Using (S.53), (S.55), and the triangle inequality, we further deduce that

$$\begin{aligned}
&\sup_{x\in\mathcal{X}}\left| \frac{P(x)^\top \widehat{Q}_n^{-1} n^{-1/2}\sum_t P(X_t^\star)\bar{Z}_{t+1}'^\top(\hat{\theta}_n-\theta^\star)}{\sigma_n(x)} \right. \\
&\quad \left. - \frac{P(x)^\top Q_n^{-1} n^{-1/2}\sum_t P(X_t^\star)g(X_t^\star)^\top(\hat{\theta}_n-\theta^\star)}{\sigma_n(x)} \right| = o_p\big((\log n)^{-1/2}\big). \qquad \text{(S.56)}
\end{aligned}$$

Next, for each $j \in \{1,\ldots,d_\theta\}$, let $g_{j,n}(\cdot) = P(\cdot)^\top \gamma_{j,n}$ and observe that

$$\begin{aligned}
&\sup_{x\in\mathcal{X}} \frac{\left|P(x)^\top Q_n^{-1} n^{-1}\sum_t P(X_t^\star) g_j(X_t^\star) - g_j(x)\right|}{\sigma_n(x)} \\
&\le \sup_{x\in\mathcal{X}} \frac{\left|P(x)^\top Q_n^{-1} n^{-1}\sum_t P(X_t^\star)\big(g_j(X_t^\star) - g_{j,n}(X_t^\star)\big)\right|}{\sigma_n(x)} + \sup_{x\in\mathcal{X}} \frac{|g_j(x)-g_{j,n}(x)|}{\sigma_n(x)} \\
&\le O_p(1)\left(n^{-1}\sum_t \big(g_j(X_t^\star) - g_{j,n}(X_t^\star)\big)^2\right)^{1/2} + K\sup_{x\in\mathcal{X}}|g_j(x)-g_{j,n}(x)| \\
&= o_p\big((\log n)^{-1/2}\big),
\end{aligned}$$

where the last line follows Assumption 1(v). This estimate further implies

$$\begin{aligned}
&\sup_{x\in\mathcal{X}} \frac{\left|P(x)^\top Q_n^{-1} n^{-1/2}\sum_t P(X_t^\star)g(X_t^\star)^\top(\hat{\theta}_n-\theta^\star) - n^{1/2}g(x)^\top(\hat{\theta}_n-\theta^\star)\right|}{\sigma_n(x)} \\
&\le n^{1/2}\|\hat{\theta}_n-\theta^\star\| \sup_{x\in\mathcal{X}} \frac{\left\|P(x)^\top Q_n^{-1} n^{-1}\sum_{t=1}^n P(X_t^\star)g(X_t^\star) - g(x)\right\|}{\sigma_n(x)}
\end{aligned}$$

$$= o_p\big((\log n)^{-1/2}\big). \tag{S.57}$$

Finally, since $\sup_{x\in\mathcal{X}} \|P(x)\|^{-1} = o((\log n)^{-1/2})$ under Assumption 1(vii),

$$\sup_{x\in\mathcal{X}} \frac{\big|n^{1/2} g(x)^\top (\hat{\theta}_n - \theta^\star)\big|}{\sigma_n(x)} \leq O_p(1) \sup_{x\in\mathcal{X}} \frac{\|g(x)\|}{\|P(x)\|} = o_p\big((\log n)^{-1/2}\big). \tag{S.58}$$

Combing (S.56), (S.57), and (S.58), we derive (S.49) for $j=3$ as claimed. This completes the proof of the lemma. □

SA.2 *Technical lemmas for Theorem 2*

In this subsection, we prove Theorem 2 in the main text. We first explicitly introduce some notation for various bootstrap quantities:

$$\widehat{Q}_n^* = n^{-1}\sum_t P(\widehat{X}_t^*)P(\widehat{X}_t^*)^\top, \qquad \widehat{b}_n^* = (\widehat{Q}_n^*)^{-1}\bigg(n^{-1}\sum_t P(\widehat{X}_t^*)\widehat{Z}_{t+1}^*\bigg),$$

$$\widehat{u}_t^* = \widehat{Z}_{t+1}^* - P(\widehat{X}_t^*)^\top \widehat{b}_n^*, \qquad \widehat{A}_n^* = n^{-1}\sum_t P(\widehat{X}_t^*)P(\widehat{X}_t^*)^\top (\widehat{u}_t^*)^2,$$

$$\widehat{\Sigma}_n^* = (\widehat{Q}_n^*)^{-1}\widehat{A}_n^*(\widehat{Q}_n^*)^{-1}.$$

We need some technical lemmas before proving Theorem 2. Below, we use $\mathcal{D}_n$ to denote the σ-field generated by data and use $\mathbb{P}^*$ (resp., $\mathbb{E}^*[\cdot]$) to denote the conditional probability (resp., expectation) given data.

Lemma S9. *Under Assumptions* 1 *and* 2,

$$\|\widehat{Q}_n^* - \widehat{Q}_n\|_S = O_p\big(\zeta_{0,n}\big(\log(m_n)m_n n^{-1}\big)^{1/2}\big) = o_p(1).$$

Proof. Denote $D_t^* \equiv P(\widehat{X}_t^*)P(\widehat{X}_t^*)^\top - \mathbb{E}^*[P(\widehat{X}_t^*)P(\widehat{X}_t^*)^\top]$. We further set

$$R_{D,n} \equiv 2\max_{1\leq t\leq n}\|P(\widehat{X}_t)\|^2, \qquad \sigma_{D,n}^2 \equiv n\lambda_{\max}(\widehat{Q}_n)\max_{1\leq t\leq n}\|P(\widehat{X}_t)\|^2.$$

It is easy to see that

$$\max_{1\leq t\leq n}\|D_t^*\|_S \leq R_{D,n}, \qquad \bigg\|\sum_t \mathbb{E}^*\big[D_t^* D_t^*\big]\bigg\|_S \leq \sigma_{D,n}^2. \tag{S.59}$$

Since the matrix-valued variables $(D_t^*)_{t\geq 1}$ are i.i.d. with zero mean conditional on data, by (S.59), we can invoke the matrix Bernstein inequality (see, e.g., Theorem 1.4 in Tropp (2012)) to deduce that, for any finite constant $C\geq 1$,

$$\begin{aligned}
&\mathbb{P}^*\big(\|\widehat{Q}_n^* - \widehat{Q}_n\|_S \geq C\sqrt{\log(m_n)R_{D,n}n^{-1}}\big)\\
&\quad = \mathbb{P}^*\bigg(\bigg\|\sum_t D_t^*\bigg\|_S \geq C\sqrt{\log(m_n)R_{D,n}n}\bigg)
\end{aligned}$$

$$\leq m_n \exp\left(\frac{-C^2 \log(m_n) R_{D,n} n/2}{\sigma_{D,n}^2 + C\big(\log(m_n)\big)^{1/2} R_{D,n}^{3/2} n^{1/2}/3}\right)$$

$$\leq m_n \exp\left(\frac{-C \log(m_n)/2}{\lambda_{\max}(\widehat{Q}_n) + \big(\log(m_n) R_{D,n} n^{-1}\big)^{1/2}}\right), \tag{S.60}$$

where the second inequality is by $\sigma_{D,n}^2 \leq n\lambda_{\max}(\widehat{Q}_n) R_{D,n}$. Under Assumptions 1 and 2, $R_{D,n} = O_p(\zeta_{0,n}^2 m_n)$ and

$$\log(m_n) R_{D,n} n^{-1} = O_p\big(\log(m_n)\zeta_{0,n}^2 m_n n^{-1}\big) = o_p(1).$$

By Lemma S2, $\lambda_{\max}(\widehat{Q}_n) = O_p(1)$. We can then deduce from (S.60) that

$$\|\widehat{Q}_n - \widehat{Q}_n^*\|_S = O_p\left(\sqrt{\log(m_n)\zeta_{0,n}^2 m_n n^{-1}}\right),$$

which completes the proof. □

Lemma S10. *Under Assumptions* 1 *and* 2, *we have* $\|\widehat{b}_n^* - \widehat{b}_n\| = O_p(\delta_{b,n}^*)$ *where*

$$\delta_{b,n}^* = \begin{cases} m_n^{1/2} n^{-1/2} & \textit{under the null,} \\ \zeta_{0,n} \log(m_n)^{1/2} m_n^{1/2} n^{-1/2} & \textit{in general.} \end{cases}$$

Proof. Since $\mathbb{E}^*[P(\widehat{X}_t^*)\widehat{Z}_{t+1}^*] = n^{-1}\sum_t P(\widehat{X}_t)\widehat{Z}_{t+1}$, we have

$$\widehat{b}_n^* - \widehat{b}_n = (\widehat{Q}_n^*)^{-1}\left(n^{-1}\sum_t P\big(\widehat{X}_t^*\big)\widehat{Z}_{t+1}^*\right) - (\widehat{Q}_n)^{-1}\left(n^{-1}\sum_t P(\widehat{X}_t)\widehat{Z}_{t+1}\right)$$

$$= (\widehat{Q}_n^*)^{-1}\left(n^{-1}\sum_t \big(P\big(\widehat{X}_t^*\big)\widehat{Z}_{t+1}^* - \mathbb{E}^*\big[P\big(\widehat{X}_t^*\big)\widehat{Z}_{t+1}^*\big]\big)\right)$$

$$+ \big((\widehat{Q}_n^*)^{-1} - (\widehat{Q}_n)^{-1}\big)\left(n^{-1}\sum_t P(\widehat{X}_t)\widehat{Z}_{t+1}\right). \tag{S.61}$$

By (S.20) and Lemma S9,

$$\lambda_{\min}^{-1}\big(\widehat{Q}_n^*\big) + \lambda_{\max}\big(\widehat{Q}_n^*\big) = O_p(1). \tag{S.62}$$

Note that $n^{-1}\sum_t (P(\widehat{X}_t^*)\widehat{Z}_{t+1}^* - \mathbb{E}^*[P(\widehat{X}_t^*)\widehat{Z}_{t+1}^*])$ is an average of $\mathcal{D}_n$-conditionally i.i.d. zero-mean elements. Therefore,

$$\mathbb{E}^*\left[\left\|n^{-1}\sum_t \big(P\big(\widehat{X}_t^*\big)\widehat{Z}_{t+1}^* - \mathbb{E}^*\big[P\big(\widehat{X}_t^*\big)\widehat{Z}_{t+1}^*\big]\big)\right\|^2\right]$$

$$= n^{-1}\mathbb{E}^*\big[\big\|P\big(\widehat{X}_t^*\big)\widehat{Z}_{t+1}^* - \mathbb{E}^*\big[P\big(\widehat{X}_t^*\big)\widehat{Z}_{t+1}^*\big]\big\|^2\big]$$

$$\leq n^{-1}\mathbb{E}^*\big[\big\|P\big(\widehat{X}_t^*\big)\widehat{Z}_{t+1}^*\big\|^2\big]$$

$$\leq n^{-2}\sum_t \|P(\widehat{X}_t)\|^2 = n^{-1}\,\mathrm{Trace}(\widehat{Q}_n) = O_p\big(m_n n^{-1}\big), \tag{S.63}$$

where the second inequality is by $|\widehat{Z}^*_{t+1}| \leq 1$ for any t and $\mathbb{E}^*[\|P(\widehat{X}^*_t)\|^2] = n^{-1}\sum_t \|P(\widehat{X}_t)\|^2$, and the last equality is by (S.20). Combining the results in (S.62) and (S.63), we get

$$(\widehat{Q}^*_n)^{-1}\Big(n^{-1}\sum_t\big(P(\widehat{X}^*_t)\widehat{Z}^*_{t+1} - \mathbb{E}^*\big[P(\widehat{X}^*_t)\widehat{Z}^*_{t+1}\big]\big)\Big) = O_p\big(m_n^{1/2}n^{-1/2}\big). \tag{S.64}$$

Since $|\widehat{Z}^*_{t+1}| \leq 1$ for any t, by (S.20) we obtain

$$\Big\|n^{-1}\sum_t P(\widehat{X}_t)\widehat{Z}_{t+1}\Big\|^2 \leq \lambda_{\max}(\widehat{Q}_n)n^{-1}\sum_t \widehat{Z}^2_{t+1} = O_p(1).$$

Hence, by Lemma S9,

$$\big((\widehat{Q}^*_n)^{-1} - (\widehat{Q}_n)^{-1}\big)\Big(n^{-1}\sum_t P(\widehat{X}_t)\widehat{Z}_{t+1}\Big) = O_p\big(\zeta_{0,n}\big(\log(m_n)m_n n^{-1}\big)^{1/2}\big). \tag{S.65}$$

Combining (S.64) and (S.65), we prove the assertion of the lemma in the general case.

If we further impose the null hypothesis, we see that

$$\big((\widehat{Q}^*_n)^{-1} - (\widehat{Q}_n)^{-1}\big)\Big(n^{-1}\sum_t P(\widehat{X}_t)\widehat{Z}_{t+1}\Big) = \big((\widehat{Q}^*_n)^{-1} - (\widehat{Q}_n)^{-1}\big)\widehat{Q}_n\widehat{b}_n = o_p\big(m_n^{1/2}n^{-1/2}\big),$$

where the second equality follows from Lemma S6. This estimate and (S.64) imply the assertion of the lemma under the null. □

LEMMA S11. *Under Assumptions* 1 *and* 2, *we have* $\|\widehat{A}^*_n - \widehat{A}_n\|_S = o_p(1)$. *In addition, under the null hypothesis,* $\|\widehat{A}^*_n - \widehat{A}_n\|_S = O_p(\zeta_{0,n}m_n n^{-1/2})$.

PROOF. We decompose $\widehat{u}^*_t = \widehat{Z}^*_{t+1} - P(\widehat{X}^*_t)^\top\widehat{b}^*_n$ as

$$\widehat{u}^*_t = \widehat{Z}^*_{t+1} - P\big(\widehat{X}^*_t\big)^\top\widehat{b}_n - P\big(\widehat{X}^*_t\big)^\top\big(\widehat{b}^*_n - \widehat{b}_n\big).$$

Therefore,

$$\begin{aligned}
\widehat{A}^*_n &= n^{-1}\sum_t P\big(\widehat{X}^*_t\big)P\big(\widehat{X}^*_t\big)^\top\big(\widehat{u}^*_t\big)^2\\
&= n^{-1}\sum_t P\big(\widehat{X}^*_t\big)P\big(\widehat{X}^*_t\big)^\top\big(\widehat{Z}^*_{t+1} - P\big(\widehat{X}^*_t\big)^\top\widehat{b}_n\big)^2\\
&\quad - 2n^{-1}\sum_t P\big(\widehat{X}^*_t\big)P\big(\widehat{X}^*_t\big)^\top\big(\widehat{Z}^*_{t+1} - P\big(\widehat{X}^*_t\big)^\top\widehat{b}_n\big)P\big(\widehat{X}^*_t\big)^\top\big(\widehat{b}^*_n - \widehat{b}_n\big)\\
&\quad + n^{-1}\sum_t P\big(\widehat{X}^*_t\big)P\big(\widehat{X}^*_t\big)^\top\big(P\big(\widehat{X}^*_t\big)^\top\big(\widehat{b}^*_n - \widehat{b}_n\big)\big)^2\\
&\equiv R_{1,n} - 2R_{2,n} + R_{3,n}.
\end{aligned} \tag{S.66}$$

We analyze these terms in turn, starting with the (leading) term $R_{1,n}$ defined as

$$R_{1,n} \equiv n^{-1} \sum_t P(\widehat{X}_t^*) P(\widehat{X}_t^*)^\top \big(\widehat{Z}_{t+1}^* - P(\widehat{X}_t^*)^\top \widehat{b}_n\big)^2.$$

Note that $\widehat{b}_n$ is $\mathcal{D}_n$-measurable (where $\mathcal{D}_n$ is the σ-field generated by data) and $R_{1,n}$ is the average of conditionally i.i.d. variables. The conditional mean of each summand term is

$$\begin{aligned}
&\mathbb{E}^*\big[P(\widehat{X}_t^*) P(\widehat{X}_t^*)^\top \big(\widehat{Z}_{t+1}^* - P(\widehat{X}_t^*)^\top \widehat{b}_n\big)^2\big] \\
&\quad = n^{-1} \sum_t P(\widehat{X}_t) P(\widehat{X}_t)^\top \big(\widehat{Z}_{t+1} - P(\widehat{X}_t)^\top \widehat{b}_n\big)^2 \\
&\quad = n^{-1} \sum_t P(\widehat{X}_t) P(\widehat{X}_t)^\top \widehat{u}_t^2 = \widehat{A}_n, \qquad \text{(S.67)}
\end{aligned}$$

and the conditional second moment of each centered summand term satisfies

$$\begin{aligned}
&\mathbb{E}^*\Bigg[\bigg\| n^{-1} \sum_t P(\widehat{X}_t^*) P(\widehat{X}_t^*)^\top \big(\widehat{Z}_{t+1}^* - P(\widehat{X}_t^*)^\top \widehat{b}_n\big)^2 - \widehat{A}_n \bigg\|^2\Bigg] \\
&\quad \le n^{-1} \sum_{l_1, l_2 = 1}^{m_n} \mathbb{E}^*\Big[\Big(p_{l_1}(\widehat{X}_t^*) p_{l_2}(\widehat{X}_t^*) \big(\widehat{Z}_{t+1}^* - P(\widehat{X}_t^*)^\top \widehat{b}_n\big)^2\Big)^2\Big] \\
&\quad = n^{-2} \sum_t \sum_{l_1, l_2 = 1}^{m_n} \Big(p_{l_1}(\widehat{X}_t) p_{l_2}(\widehat{X}_t) \big(\widehat{Z}_{t+1} - P(\widehat{X}_t)^\top \widehat{b}_n\big)^2\Big)^2 \\
&\quad = n^{-2} \sum_t \big\| P(\widehat{X}_t) \big\|^4 \widehat{u}_t^4. \qquad \text{(S.68)}
\end{aligned}$$

Next, we observe that

$$\begin{aligned}
&n^{-2} \sum_t \big\| P(\widehat{X}_t) \big\|^4 \widehat{u}_t^4 \\
&\quad \le K n^{-2} \sum_t \big\| P(\widehat{X}_t) \big\|^4 (u_t^\star)^4 + K n^{-2} \sum_t \big\| P(\widehat{X}_t) \big\|^4 (\widehat{u}_t - u_t^\star)^4 \\
&\quad \le K \zeta_{0,n}^2 m_n n^{-2} \sum_t \big\| P(\widehat{X}_t) \big\|^2 + \zeta_{0,n}^4 m_n^2 n^{-2} \sum_t (\widehat{u}_t - u_t^\star)^4 \\
&\quad = O_p\big(\zeta_{0,n}^2 m_n^2 n^{-1}\big) + O_p\big(\zeta_{0,n}^4 m_n^2 n^{-1}\big) O_p\big(n^{-1/2} + \zeta_{0,n}^2 m_n \delta_{b,n}^4\big) \\
&\quad = O_p\big(\zeta_{0,n}^2 m_n^2 n^{-1}\big), \qquad \text{(S.69)}
\end{aligned}$$

where the second inequality follows from $|u_t^\star| \le K$ and $\|P(X_t)\| \le \zeta_{0,n} m_n^{1/2}$, the first equality follows from (S.46), and the last line is implied by the maintained rate conditions in Assumption 2. Therefore,

$$\|R_{1,n} - \widehat{A}_n\| = O_p\big(\zeta_{0,n} m_n n^{-1/2}\big). \qquad \text{(S.70)}$$

Next, we analyze

$$R_{2,n} \equiv n^{-1} \sum_t P(\widehat{X}_t^*) P(\widehat{X}_t^*)^\top (\widehat{Z}_{t+1}^* - P(\widehat{X}_t^*)^\top \widehat{b}_n)(P(\widehat{X}_t^*)^\top (\widehat{b}_n^* - \widehat{b}_n)).$$

Consider any $a_n \in \mathbb{R}^{m_n}$ with $\|a_n\| = 1$. Observe that

$$\begin{aligned}
&a_n^\top R_{2,n}^2 a_n \\
&= \left\| n^{-1} \sum_t P(\widehat{X}_t^*) P(\widehat{X}_t^*)^\top a_n (\widehat{Z}_{t+1}^* - P(\widehat{X}_t^*)^\top \widehat{b}_n)(P(\widehat{X}_t^*)^\top (\widehat{b}_n^* - \widehat{b}_n)) \right\|^2 \\
&\leq \lambda_{\max}(\widehat{Q}_n^*) n^{-1} \sum_t (P(\widehat{X}_t^*)^\top a_n)^2 (\widehat{Z}_{t+1}^* - P(\widehat{X}_t^*)^\top \widehat{b}_n)^2 (P(\widehat{X}_t^*)^\top (\widehat{b}_n^* - \widehat{b}_n))^2 \\
&\leq \lambda_{\max}(\widehat{Q}_n^*) \max_{1 \leq t \leq n} \|P(\widehat{X}_t^*)\|^2 \|\widehat{b}_n^* - \widehat{b}_n\|^2 n^{-1} \sum_t (P(\widehat{X}_t^*)^\top a_n)^2 (\widehat{Z}_{t+1}^* - P(\widehat{X}_t^*)^\top \widehat{b}_n)^2 \\
&\leq O_p(\zeta_{0,n}^2 m_n (\delta_{b,n}^*)^2) \max_{1 \leq t \leq n} (\widehat{Z}_{t+1}^* - P(\widehat{X}_t^*)^\top \widehat{b}_n)^2, \qquad \text{(S.71)}
\end{aligned}$$

where the last line follows from Lemma S10. In addition, since $\widehat{Z}_{t+1}^* - h(\widehat{X}_t^*)$ is bounded and $\max_{1\leq t\leq n} |h(\widehat{X}_t^*) - P(\widehat{X}_t^*)^\top b_n^\star| = o_p(1)$,

$$\begin{aligned}
\max_{1 \leq t \leq n} \left|\widehat{Z}_{t+1}^* - P(\widehat{X}_t^*)^\top \widehat{b}_n\right| &\leq \max_{1 \leq t \leq n} \left|\widehat{Z}_{t+1}^* - h(\widehat{X}_t^*)\right| + \max_{1 \leq t \leq n} \left|h(\widehat{X}_t^*) - P(\widehat{X}_t^*)^\top b_n^\star\right| \\
&\quad + \max_{1 \leq t \leq n} \left|P(\widehat{X}_t^*)^\top (\widehat{b}_n - b_n^\star)\right| \\
&= O_p(1 + \zeta_{0,n} m_n^{1/2} \delta_{b,n}) = O_p(1). \qquad \text{(S.72)}
\end{aligned}$$

Hence, (S.71) further implies that

$$\|R_{2,n}\|_S = O_p(\zeta_{0,n} m_n^{1/2} \delta_{b,n}^*). \qquad \text{(S.73)}$$

It remains to study the term

$$R_{3,n} \equiv n^{-1} \sum_t P(\widehat{X}_t^*) P(\widehat{X}_t^*)^\top (P(\widehat{X}_t^*)^\top (\widehat{b}_n^* - \widehat{b}_n))^2.$$

Consider any $a_n \in \mathbb{R}^{m_n}$ with $\|a_n\| = 1$. Then by the Cauchy–Schwarz inequality

$$\begin{aligned}
a_n^\top R_{3,n}^2 a_n &= \left\| n^{-1} \sum_t P(\widehat{X}_t^*) P(\widehat{X}_t^*)^\top a_n (P(\widehat{X}_t^*)^\top (\widehat{b}_n^* - \widehat{b}_n))^2 \right\|^2 \\
&\leq \lambda_{\max}(\widehat{Q}_n^*) n^{-1} \sum_t (P(\widehat{X}_t^*)^\top a_n)^2 (P(\widehat{X}_t^*)^\top (\widehat{b}_n^* - \widehat{b}_n))^4 \\
&\leq (\lambda_{\max}(\widehat{Q}_n^*))^2 \max_{1 \leq t \leq n} \|P(\widehat{X}_t^*)\|^4 \|\widehat{b}_n^* - \widehat{b}_n\|^4 \\
&\leq O_p(\zeta_{0,n}^4 m_n^2 (\delta_{b,n}^*)^4),
\end{aligned}$$

where the last line follows from (S.62) and Lemma S10. Hence,

$$\|R_{3,n}\|_S = O_p\big(\zeta_{0,n}^2 m_n \delta_{b,n}^{*2}\big). \tag{S.74}$$

Finally, collecting the estimates in (S.70), (S.73), and (S.74), we get

$$\begin{aligned}\big\|\widehat{A}_n^* - \widehat{A}_n\big\|_S &\le \|R_{1,n} - \widehat{A}_n\|_S + 2\|R_{2,n}\|_S + \|R_{3,n}\|_S \\ &= O_p\big(\zeta_{0,n} m_n n^{-1/2} + \zeta_{0,n} m_n^{1/2} \delta_{b,n}^* + \zeta_{0,n}^2 m_n \delta_{b,n}^{*2}\big) \\ &= O_p\big(\zeta_{0,n} m_n^{1/2} \delta_{b,n}^*\big).\end{aligned}$$

The assertions of the lemma then follows from the definition of $\delta_{b,n}^*$ in Lemma S10, and the fact that $\zeta_{0,n}^2 \log(m_n)^{1/2} m_n n^{-1/2} = o(1)$, which is implied by the maintained rate conditions in Assumption 2. □

Lemma S12. *Suppose that Assumptions* 1 *and* 2 *hold. Then under the null hypothesis,*

$$\big\|\widehat{H}_n^* - \widehat{A}_n\big\|_S = O_p\big(\zeta_{0,n} m_n n^{-1/2}\big),$$

where $\widehat{H}_n^* \equiv \mathbb{E}^*[(\widehat{Z}_{t+1}^*)^2 P(\widehat{X}_t^*) P(\widehat{X}_t^*)^\top] - \mathbb{E}^*[\widehat{Z}_{t+1}^* P(\widehat{X}_t^*)]\mathbb{E}^*[P(\widehat{X}_t^*)^\top \widehat{Z}_{t+1}^*]$.

Proof. Note that

$$\mathbb{E}^*\big[P\big(\widehat{X}_t^*\big)\widehat{Z}_{t+1}^*\big] = n^{-1}\sum_{t=1}^n P(\widehat{X}_t)\widehat{Z}_{t+1} = \widehat{Q}_n \widehat{b}_n. \tag{S.75}$$

Under the null hypothesis, $h(x) = 0$ and $b_n^\star = 0$. By (S.20) and Lemma S6,

$$\widehat{Q}_n \widehat{b}_n = O_p\big(m_n^{1/2} n^{-1/2}\big), \tag{S.76}$$

which together with (S.75) implies that

$$\big\|\mathbb{E}^*\big[\widehat{Z}_{t+1}^* P\big(\widehat{X}_t^*\big)\big]\mathbb{E}^*\big[P\big(\widehat{X}_t^*\big)^\top \widehat{Z}_{t+1}^*\big]\big\| = O_p\big(m_n n^{-1}\big). \tag{S.77}$$

Since $\widehat{u}_t = \widehat{Z}_{t+1} - P(\widehat{X}_t)^\top \widehat{b}_n$,

$$\begin{aligned}&\mathbb{E}^*\big[\big(\widehat{Z}_{t+1}^*\big)^2 P\big(\widehat{X}_t^*\big)P\big(\widehat{X}_t^*\big)^\top\big] - \widehat{A}_n \\ &\quad = n^{-1}\sum_t (\widehat{Z}_{t+1})^2 P(\widehat{X}_t)P(\widehat{X}_t)^\top - \widehat{A}_n \\ &\quad = n^{-1}\sum_t \big(P(\widehat{X}_t)^\top \widehat{b}_n + \widehat{u}_t\big)^2 P(\widehat{X}_t)P(\widehat{X}_t)^\top - \widehat{A}_n \\ &\quad = n^{-1}\sum_t \big(P(\widehat{X}_t)^\top \widehat{b}_n\big)^2 P(\widehat{X}_t)P(\widehat{X}_t)^\top \\ &\qquad + 2n^{-1}\sum_t \widehat{u}_t\big(P(\widehat{X}_t)^\top \widehat{b}_n\big)P(\widehat{X}_t)P(\widehat{X}_t)^\top.\end{aligned} \tag{S.78}$$

Let $a_n \in \mathbb{R}^{m_n}$ be such that $\|a_n\| = 1$. By Assumption 2, (S.20), and (S.76),

$$\begin{aligned}
&a_n^\top \Big(n^{-1} \sum_t \big(P(\widehat{X}_t)^\top \widehat{b}_n\big)^2 P(\widehat{X}_t) P(\widehat{X}_t)^\top \Big)^2 a_n \\
&= \Big\| n^{-1} \sum_t P(\widehat{X}_t) \big(P(\widehat{X}_t)^\top \widehat{b}_n\big)^2 \big(P(\widehat{X}_t)^\top a_n\big) \Big\|^2 \\
&\leq \lambda_{\max}(\widehat{Q}_n) n^{-1} \sum_t \big(P(\widehat{X}_t)^\top \widehat{b}_n\big)^4 \big(P(\widehat{X}_t)^\top a_n\big)^2 \\
&\leq \big(\lambda_{\max}(\widehat{Q}_n)\big)^2 \max_{1\leq t\leq n} \big(P(\widehat{X}_t)^\top \widehat{b}_n\big)^4 \\
&\leq \big(\lambda_{\max}(\widehat{Q}_n)\big)^2 \max_{1\leq t\leq n} \big\|P(\widehat{X}_t)\big\|^4 \|\widehat{b}_n\|^4 = O_p\big(\zeta_{0,n}^4 m_n^4 n^{-2}\big),
\end{aligned}$$

which implies that

$$\Big\| n^{-1} \sum_t \big(P(\widehat{X}_t)^\top \widehat{b}_n\big)^2 P(\widehat{X}_t) P(\widehat{X}_t)^\top \Big\|_S = O_p\big(\zeta_{0,n}^2 m_n^2 n^{-1}\big). \tag{S.79}$$

Turning to the second term in (S.78), we note that by (S.20), Lemma S7, and (S.76), we have

$$\begin{aligned}
&a_n^\top \Big(n^{-1} \sum_t \widehat{u}_t \big(P(\widehat{X}_t)^\top \widehat{b}_n\big) P(\widehat{X}_t) P(\widehat{X}_t)^\top \Big)^2 a_n \\
&= \Big\| n^{-1} \sum_t P(\widehat{X}_t) \widehat{u}_t \big(P(\widehat{X}_t)^\top \widehat{b}_n\big) \big(P(\widehat{X}_t)^\top a_n\big) \Big\|^2 \\
&\leq \lambda_{\max}(\widehat{Q}_n) n^{-1} \sum_t \big(P(\widehat{X}_t)^\top \widehat{b}_n\big)^2 \big(\widehat{u}_t P(\widehat{X}_t)^\top a_n\big)^2 \\
&\leq \lambda_{\max}(\widehat{Q}_n) \lambda_{\max}(\widehat{A}_n) \max_{1\leq t\leq n} \big\|P(\widehat{X}_t)\big\|^2 \|\widehat{b}_n\|^2 = O_p\big(\zeta_{0,n}^2 m_n^2 n^{-1}\big),
\end{aligned} \tag{S.80}$$

which implies that

$$\Big\| n^{-1} \sum_t \widehat{u}_t P(\widehat{X}_t)^\top \widehat{b}_n P(\widehat{X}_t) P(\widehat{X}_t)^\top \Big\|_S = O_p\big(\zeta_{0,n} m_n n^{-1/2}\big). \tag{S.81}$$

The assertion of the lemma then follows from (S.77), (S.78), (S.79), and (S.81). □

References

Tropp, J. A. (2012), "User-friendly tail bounds for sums of random matrices." *Foundations of Computational Mathematics*, 12, 389–434. [18]

Co-editor Tao Zha handled this manuscript.

Manuscript received 17 September, 2020; final version accepted 29 May, 2021; available online 9 June, 2021.