Econometrica Supplementary Material

SUPPLEMENT TO "MONTE CARLO CONFIDENCE SETS FOR IDENTIFIED SETS"

(*Econometrica*, Vol. 86, No. 6, November 2018, 1965–2018)

XIAOHONG CHEN
Cowles Foundation for Research in Economics, Yale University

TIMOTHY M. CHRISTENSEN
Department of Economics, New York University

ELIE TAMER
Department of Economics, Harvard University

THIS SUPPLEMENTAL MATERIAL CONSISTS of the following sections:

E. Verification of main conditions for uniformity in examples
F. Proofs of all the results in the main text and additional results

APPENDIX E: VERIFICATION OF MAIN CONDITIONS FOR UNIFORMITY IN EXAMPLES

E.1. *Example 1: Uniform Validity for Missing Data*

Here, we apply Proposition D.1 to establish uniform validity of our procedures. To make the missing data example fit the preceding notation, let $p_\theta = (\tilde\gamma_{11}(\theta), \tilde\gamma_{00}(\theta), 1-\tilde\gamma_{00}(\theta)-\tilde\gamma_{11}(\theta))'$ and let $p = (\tilde\gamma_{11}, \tilde\gamma_{00}, 1-\tilde\gamma_{00}-\tilde\gamma_{11})'$ denote the true probabilities under $\mathbb{P}$. The only requirement on $\mathbf{P}$ is that (36) holds. Therefore, the conclusion of Proposition D.1 holds uniformly over a set of DGPs under which the probability of missing data can drift to zero at rate up to n^{-1}. As $\{p_\theta : \theta \in \Theta, p_\theta > 0\} = \operatorname{int}(\Delta^2)$, Lemma D.6 implies that $\{\sqrt{n}\gamma(\theta) : \theta \in \Theta_{\mathrm{osn}}(\mathbb{P})\}$ covers a ball of radius ρ_n (independently of $\mathbb{P}$) with $\rho_n \to \infty$ as $n \to \infty$. This verifies Assumption D.2.

By concavity, the infimum in the definition of the profile likelihood $PL_n(M(\theta))$ is attained at either the lower or upper bound of $M_I(\theta) = [\tilde\gamma_{11}(\theta), \tilde\gamma_{11}(\theta)+\tilde\gamma_{00}(\theta)]$. Moreover, at both $\mu = \tilde\gamma_{11}(\theta)$ and $\mu = \tilde\gamma_{11}(\theta)+\tilde\gamma_{00}(\theta)$, the profile likelihood is

$$\sup_{\substack{0\le g_{11}\le\mu \\ \mu\le g_{11}+g_{00}\le 1}} \big(n\mathbb{P}_n\mathbb{1}\{yd=1\}\log g_{11} + n\mathbb{P}_n\mathbb{1}\{1-d=1\}\log g_{00} + n\mathbb{P}_n\mathbb{1}\{d-yd=1\}\log(1-g_{11}-g_{00})\big).$$

The constraint $g_{11} \le \mu$ will be the binding constraint at the lower bound and the constraint $\mu \le g_{11}+g_{00}$ will be the binding constraint at the upper bound (wpa1, uniformly in $\mathbb{P}$). These constraints are equivalent to $a_1'\mathbb{I}_0^{1/2}(g-\tilde\gamma(\theta)) \le 0$ and $a_2'\mathbb{I}_0^{1/2}(g-\tilde\gamma(\theta)) \le 0$ for some $a_1 = a_1(\mathbb{P}) \in \mathbb{R}^2$ and $a_2 = a_2(\mathbb{P}) \in \mathbb{R}^2$ with $g = (g_{11}, g_{00})'$ and $\mathbb{I}_0 = \mathbb{I}_0(\mathbb{P})$. It now follows from Proposition D.1 and Lemmas D.6 and D.7 that

$$\left| nPL_n(M_I) - \min_{j\in\{1,2\}} \sup_{\gamma: a_j'\gamma\le 0} \left(\ell_n - \frac{1}{2}\|\sqrt{n}\gamma\|^2 + (\sqrt{n}\gamma)'\mathbb{V}_n\right)\right| = o_{\mathbb{P}}(1)$$

Xiaohong Chen: xiaohong.chen@yale.edu
Timothy M. Christensen: timothy.christensen@nyu.edu
Elie Tamer: elietamer@fas.harvard.edu

 https://doi.org/10.3982/ECTA14525

and

$$\sup_{\theta\in\Theta'_{\mathrm{osn}}(\mathbb{P})}\left|nPL_n\big(M(\theta)\big)-\min_{j\in\{1,2\}}\sup_{\gamma:a'_j(\gamma-\gamma(\theta))\le 0}\left(\ell_n-\frac{1}{2}\|\sqrt{n}\gamma\|^2+(\sqrt{n}\gamma)'\mathbb{V}_n\right)\right|=o_{\mathbb{P}}(1)$$

uniformly in $\mathbb{P}$. Let T_j denote the regular half-space in $\mathbb{R}^2$ defined by the inequality $a'_j\gamma\le 0$ for $j=1,2$. We may write the above as

$$\left|nPL_n(M_I)-\left(\ell_n+\frac{1}{2}\|\mathbb{V}_n\|^2-\max_{j\in\{1,2\}}\inf_{t\in T_j}\|\mathbb{V}_n-t\|^2\right)\right|=o_{\mathbb{P}}(1),$$

$$\sup_{\theta\in\Theta'_{\mathrm{osn}}(\mathbb{P})}\left|nPL_n\big(M(\theta)\big)-\left(\ell_n+\frac{1}{2}\|\mathbb{V}_n\|^2-\max_{j\in\{1,2\}}\inf_{t\in T_j}\left\|\big(\mathbb{V}_n-\sqrt{n}\gamma(\theta)\big)-t\right\|^2\right)\right|=o_{\mathbb{P}}(1)$$

uniformly in $\mathbb{P}$. This verifies the uniform expansion of the profile criterion.

E.2. *Example 3: Uniform Validity of Procedure 2 versus the Bootstrap*

We return to Example 3 considered in Section 5.3.3 and show that our MC CSs (based on the posterior distribution of the profile QLR) are uniformly valid under very mild conditions while bootstrap-based CSs (based on the bootstrap distribution of the profile QLR) can under-cover along certain sequences of DGPs. This reinforces the fact that our MC CSs and bootstrap-based CSs have different asymptotic properties.

Recall that $X_1,\dots,X_n$ are i.i.d. with unknown mean $\mu^*\in\mathbb{R}_+$ and $\mu\in\mathbb{R}_+$ is identified by the moment inequality $\mathbb{E}[\mu-X_i]\le 0$. The identified set for μ is $M_I=[0,\mu^*]$. We consider coverage of the CS for $M_I=[0,\mu^*]$. We introduce a slackness parameter $\eta\in\mathbb{R}_+$ to write this model as a moment equality model $\mathbb{E}[\mu+\eta-X_i]=0$. The parameter space for $\theta=(\mu,\eta)$ is $\Theta=\mathbb{R}^2_+$. The GMM objective function and profile QLR are

$$\begin{aligned}L_n(\mu,\eta)&=-\frac{1}{2}(\mu+\eta-\bar{X}_n)^2,\\ PQ_n(M_I)&=(\mathbb{V}_n\wedge 0)^2-\big((\mathbb{V}_n+\sqrt{n}\mu^*)\wedge 0\big)^2,\\ PQ_n\big(M(\theta)\big)&=\big((\mathbb{V}_n-\sqrt{n}\gamma(\theta))\wedge 0\big)^2-\big((\mathbb{V}_n+\sqrt{n}\mu^*)\wedge 0\big)^2,\end{aligned}\tag{39}$$

where $\gamma(\theta)=\mu+\eta-\mu^*\in[-\mu^*,\infty)$ and $\mathbb{V}_n=\mathbb{V}_n(\mathbb{P})=\sqrt{n}(\bar{X}_n-\mu^*)$.

E.2.1. *Uniform Validity of Procedures 2 and 3*

Let $\mathbf{P}$ be the family of distributions under which the X_i are i.i.d. with mean $\mu^*=\mu^*(\mathbb{P})\in\mathbb{R}_+$ and unit variance and for which

$$\lim_{n\to\infty}\sup_{\mathbb{P}\in\mathbf{P}}\sup_{z\in\mathbb{R}}\big|\mathbb{P}(\mathbb{V}_n\le z)-\Phi(z)\big|=0\tag{40}$$

holds. We first consider uniform coverage of our MC CSs $\widehat{M}_\alpha$ for the identified set $M_I=M_I(\mathbb{P})=[0,\mu^*(\mathbb{P})]$.

To focus solely on the essential ideas, assume the prior on θ induces a uniform prior on γ (the posterior is still proper); this could be relaxed at the cost of more cumbersome notation without changing the results that follow. Letting $z\ge 0$, $\kappa=\sqrt{n}\gamma$, and

$v_n = v_n(\mathbb{P}) = \mathbb{V}_n + \sqrt{n}\mu^*$, we have

$$\Pi_n\big(\{\theta : PQ_n\big(M(\theta)\big) \le z\}|\mathbf{X}_n\big) = \frac{\displaystyle\int_{-\sqrt{n}\mu^*}^{\infty} \mathbb{1}\big\{\big((\mathbb{V}_n - \kappa) \wedge 0\big)^2 - (v_n \wedge 0)^2 \le z\big\} e^{-\frac{1}{2}(\mathbb{V}_n - \kappa)^2}\,\mathrm{d}\kappa}{\displaystyle\int_{-\sqrt{n}\mu^*}^{\infty} e^{-\frac{1}{2}(\mathbb{V}_n - \kappa)^2}\,\mathrm{d}\kappa}.$$

A change of variables with $x = \mathbb{V}_n - \kappa$ yields

$$\begin{aligned}\Pi_n\big(\{\theta : PQ_n\big(M(\theta)\big) \le z\}|\mathbf{X}_n\big) &= \frac{\displaystyle\int_{-\infty}^{v_n} \mathbb{1}\big\{(x \wedge 0)^2 \le z + (v_n \wedge 0)^2\big\} e^{-\frac{1}{2}x^2}\,\mathrm{d}x}{\displaystyle\int_{-\infty}^{v_n} e^{-\frac{1}{2}x^2}\,\mathrm{d}x} \\ &= \mathbb{P}_{Z|\mathbf{X}_n}\big(-\sqrt{z + (v_n \wedge 0)^2} \le Z | Z \le v_n\big) = G(v_n; z).\end{aligned}$$

As we have an explicit form for the posterior distribution of the profile QLR, we can compute the posterior critical value directly rather than resorting to MC sampling. Therefore, Assumption D.6 is not required here (as we can trivially set $\xi_{n,\alpha}^{\mathrm{post},p} = \xi_{n,\alpha}^{\mathrm{mc},p}$). If MC sampling were to be used, we would require that Assumption D.6 holds.

Fix any $\alpha \in (\frac{1}{2}, 1)$. For $v_n \ge 0$, we have

$$G(v_n; z) = \mathbb{P}_{Z|\mathbf{X}_n}(-\sqrt{z} \le Z | Z \le v_n)$$

and so the posterior α-critical value $\xi_{n,\alpha}^{\mathrm{post},p} = \Phi^{-1}((1-\alpha)\Phi(v_n))^2$. Therefore,

$$\begin{aligned}\mathbb{P}\big(PQ_n(M_I) \le \xi_{n,\alpha}^{\mathrm{post},p} | v_n \ge 0\big) &= \mathbb{P}\big((\mathbb{V}_n \wedge 0)^2 \le \Phi^{-1}\big((1-\alpha)\Phi(v_n)\big)^2 | v_n \ge 0\big) \\ &= \mathbb{P}\big(\Phi^{-1}\big((1-\alpha)\Phi(v_n)\big) \le \mathbb{V}_n | v_n \ge 0\big).\end{aligned} \tag{41}$$

Now suppose that $v_n < 0$. Here,we have

$$G(v_n; z) = \mathbb{P}_{Z|\mathbf{X}_n}\Big(-\sqrt{z + v_n^2} \le Z \big| Z \le v_n\Big) = \frac{\Phi(v_n) - \Phi\Big(-\sqrt{z + v_n^2}\Big)}{\Phi(v_n)},$$

from which it follows that $\xi_{n,\alpha}^{\mathrm{post},p} = \Phi^{-1}((1-\alpha)\Phi(v_n))^2 - v_n^2$ and hence

$$\begin{aligned}\mathbb{P}\big(PQ_n(M_I) \le \xi_{n,\alpha}^{\mathrm{post},p} | v_n < 0\big) &= \mathbb{P}\big((\mathbb{V}_n \wedge 0)^2 \le \Phi^{-1}\big((1-\alpha)\Phi(v_n)\big)^2 | v_n < 0\big) \\ &= \mathbb{P}\big(\Phi^{-1}\big((1-\alpha)\Phi(v_n)\big) \le \mathbb{V}_n | v_n < 0\big).\end{aligned} \tag{42}$$

Combining (41) and (42), we obtain

$$\mathbb{P}\big(PQ_n(M_I) \le \xi_{n,\alpha}^{\mathrm{post},p}\big) = \mathbb{P}\big((1-\alpha)\Phi(v_n) \le \Phi(\mathbb{V}_n)\big) \ge \mathbb{P}\big((1-\alpha) \le \Phi(\mathbb{V}_n)\big),$$

which, together with (40), delivers the uniform coverage result for Procedure 2:

$$\liminf_{n\to\infty} \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(\mathbb{M}_I(\mathbb{P}) \subseteq \widehat{M}_\alpha\big) \ge \alpha.$$

For uniform validity of Procedure 3, first note that (39) implies that the inequality

$$\mathbb{P}\big(PQ_n(M_I) \le \chi^2_{1,\alpha}\big) \ge \mathbb{P}\big((\mathbb{V}_n \wedge 0)^2 \le \chi^2_{1,\alpha}\big)$$

holds uniformly in $\mathbb{P}$. It follows by (40) that

$$\liminf_{n\to\infty} \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(\mathbb{M}_I(\mathbb{P}) \subseteq \widehat{M}^\chi_\alpha\big) > \alpha.$$

E.2.2. *Lack of Uniformity of the Bootstrap*

We now show that bootstrap-based CSs for M_I are not uniformly valid when the standard (i.e., nonparametric) bootstrap is used. The bootstrap criterion function $L^\star_n(\mu, \eta)$ is

$$L^\star_n(\mu, \eta) = -\frac{1}{2}\big(\mu + \eta - \bar{X}^\star_n\big)^2,$$

where $\bar{X}^\star_n$ is the bootstrap sample mean. Let $\widehat{M}_I = [0, (\bar{X}_n \vee 0)]$ and $\mathbb{V}^\star_n = \sqrt{n}(\bar{X}^\star_n - \bar{X}_n)$. Consider a subsequence $(\mathrm{P}_n)_{n\in\mathbb{N}} \subset \mathbf{P}$ with $\mu^*(\mathrm{P}_n) = c/\sqrt{n}$ for some $c > 0$ (chosen below). By similar calculations to Section 5.3.3, along this sequence of DGPs, the bootstrapped profile QLR statistic for M_I is

$$\begin{aligned} PQ^\star_n(M_I) &= 2nL^\star_n\big(\hat{\mu}^\star, \hat{\eta}^\star\big) - \inf_{\mu\in\widehat{M}_I} \sup_{\eta\in H_\mu} 2nL^\star_n(\mu, \eta) \\ &= \Big(\big(\mathbb{V}^\star_n + \big((\mathbb{V}_n + c) \wedge 0\big)\big) \wedge 0\Big)^2 - \big(\big(\mathbb{V}^\star_n + \mathbb{V}_n + c\big) \wedge 0\big)^2. \end{aligned}$$

Let $\xi^{\mathrm{boot},p}_{n,\alpha}$ denote the α quantile of the distribution of $PQ^\star_n(M_I)$. Consider

$$\widehat{M}^{\mathrm{boot}}_\alpha = \Big\{\mu : \sup_{\eta\in H_\mu} Q_n(\mu, \eta) \le \xi^{\mathrm{boot},p}_{n,\alpha}\Big\}.$$

We now show that for any $\alpha \in (\frac{1}{2}, 1)$, we may choose $c > 0$ in the definition of $(\mathrm{P}_n)_{n\in\mathbb{N}}$ such that the asymptotic coverage of $\widehat{M}^{\mathrm{boot}}_\alpha$ is strictly less than α along this sequence of DGPs. As

$$PQ^\star_n(M_I) = \Big(\big(\mathbb{V}^\star_n \wedge 0\big)^2 - \big(\big(\mathbb{V}^\star_n + \mathbb{V}_n + c\big) \wedge 0\big)^2\Big)\mathbb{1}\{\mathbb{V}_n + c \ge 0\},$$

it follows that whenever $\mathbb{V}_n + c < 0$, the bootstrap distribution of the profile QLR for M_I is point mass at the origin, and the α quantile of the bootstrap distribution is $\xi^{\mathrm{boot},p}_{n,\alpha} = 0$. However, the QLR statistic for M_I is $PQ_n(M_I) = (\mathbb{V}_n \wedge 0)^2 - ((\mathbb{V}_n + c) \wedge 0)^2$. So whenever $\mathbb{V}_n + c < 0$, we also have that $PQ_n(M_I) = \mathbb{V}^2_n - (\mathbb{V}_n + c)^2 > 0$. Therefore,

$$\mathrm{P}_n\big(M_I(\mathrm{P}_n) \subseteq \widehat{M}^{\mathrm{boot}}_\alpha | \mathbb{V}_n + c < 0\big) = 0.$$

It follows by (40) that for any c for which $\Phi(c) < \alpha$, we have

$$\limsup_{n\to\infty} \mathrm{P}_n\big(M_I(\mathrm{P}_n) \subseteq \widehat{M}^{\mathrm{boot}}_\alpha\big) \le \lim_{n\to\infty} \mathrm{P}_n(\mathbb{V}_n + c \ge 0) < \alpha.$$

E.2.3. *An Alternative Recentering*

An alternative is to recenter the criterion function at $(\bar{X}_n \vee 0)$, that is, one could use instead

$$L_n(\mu, \eta) = -\frac{1}{2}\big(\mu + \eta - (\bar{X}_n \vee 0)\big)^2,$$

similar to the idea of a sandwich (quasi-)likelihood with $(\bar{X}_n \vee 0) = \hat{\gamma}_n$. This maps into the setup described in Appendix D, where

$$nL_n(\theta) = \ell_n - \frac{1}{2}\big(\sqrt{n}\gamma(\theta)\big)^2 + \sqrt{n}\big(\gamma(\theta)\big)\big(\sqrt{n}(\hat{\gamma}_n - \tau)\big),$$

where $\ell_n = -\frac{1}{2}(\sqrt{n}(\hat{\gamma}_n - \tau))^2$, $\theta = (\mu, \eta)$, and

$$\gamma(\theta) = \mu + \eta - \mu^*, \qquad \tau = \mu^*, \qquad \hat{\gamma}_n = (\bar{X}_n \vee 0), \qquad \sqrt{n}(\hat{\gamma}_n - \tau) = \big(\mathbb{V}_n \vee -\sqrt{n}\mu^*\big),$$

where $\mathbb{V}_n = \sqrt{n}(\bar{X}_n - \mu^*)$, $\gamma(\theta) \in [-\mu^*, \infty)$, and $\mu^* \in \mathbb{R}_+$.

Assumptions D.1 and D.2(i)–(iii) hold with $\Theta_{\text{osn}} = \Theta$, $k_n = +\infty$, $T = \mathbb{R}_+$, and $\mathbf{T}v = (v \vee 0)$ (none of the models is singular). We again take a prior on θ that induces a flat prior on γ to concentrate on the essential ideas, verifying Assumption D.3.

For inference on $M_I = [0, \mu^*(\mathbb{P})]$, observe that

$$PQ_n\big(M(\theta)\big) = f\big(\sqrt{n}(\hat{\gamma}_n - \tau) - \sqrt{n}\gamma(\theta)\big), \qquad PQ_n(M_I) = f\big(\sqrt{n}(\hat{\gamma}_n - \tau)\big),$$

where $f(v) = (v \wedge 0)^2$ for each $\mathbb{P}$, verifying Assumption D.5(i). Assumption D.5(ii) also holds for this f. Finally, for Assumption D.5(iii), for any $z, v \geq 0$, we have

$$\mathbb{P}_Z\big(f(Z) \leq z | Z \in v - T\big) = \frac{\Phi(v) - \Phi(-\sqrt{z})}{\Phi(v)} \leq 1 - \Phi(-\sqrt{z}) = \mathbb{P}_Z\big(f(Z) \leq z\big).$$

Theorem D.2, together with (40), delivers uniform coverage for Procedure 2.

Similarly, for uniform validity of Procedure 3, we have

$$\mathbb{P}\big(PQ_n(M_I) \leq \chi^2_{1,\alpha}\big) \geq \mathbb{P}\big((\mathbb{V}_n \wedge 0)^2 \leq \chi^2_{1,\alpha}\big),$$

which, together with (40), delivers uniform coverage for Procedure 3.

Now consider bootstrap-based inference. As before, let $\widehat{M}_I = [0, (\bar{X}_n \vee 0)]$ and consider a subsequence $(\mathrm{P}_n)_{n\in\mathbb{N}} \subset \mathbf{P}$ with $\mu^*(\mathrm{P}_n) = c/\sqrt{n}$ for some $c > 0$. Under P_n, we then have

$$L_n^\star(\mu, \eta) = -\frac{1}{2}\big(\mu + \eta - \big(\bar{X}_n^\star \vee 0\big)\big)^2,$$

$$PQ_n^\star(M_I) = \big(\big[\big(\big(\mathbb{V}_n^\star + \mathbb{V}_n\big) \vee -c\big) - (\mathbb{V}_n \vee -c)\big] \wedge 0\big)^2,$$

and the true QLR statistic is $PQ_n(M_I) = ((\mathbb{V}_n \vee -c) \wedge 0)^2$. We again show that for any $\alpha \in (\frac{1}{2}, 1)$, we may choose $c > 0$ in the definition of $(\mathrm{P}_n)_{n\in\mathbb{N}}$ such that the asymptotic coverage of $\widehat{M}_\alpha^{\text{boot}}$ is strictly less than α along this sequence of DGPs. Observe that when $\mathbb{V}_n < -c$, we have $PQ_n(M_I) = c^2 > 0$ and $PQ_n^\star(M_I) = 0$. Therefore,

$$\mathrm{P}_n\big(M_I(\mathrm{P}_n) \subseteq \widehat{M}_\alpha^{\text{boot}} | \mathbb{V}_n + c < 0\big) = 0.$$

It follows by (40) that for any c for which $\Phi(c) < \alpha$, we again have

$$\limsup_{n\to\infty} \mathrm{P}_n\big(M_I(\mathrm{P}_n) \subseteq \widehat{M}_\alpha^{\mathrm{boot}}\big) \leq \lim_{n\to\infty} \mathrm{P}_n(\mathbb{V}_n + c \geq 0) < \alpha.$$

APPENDIX F: PROOFS AND ADDITIONAL RESULTS

F.1. *Proofs and Additional Lemmas for Sections 2 and 4*

PROOF OF LEMMA 2.1: By (ii), there is a positive sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ with $\varepsilon_n = o(1)$ such that $w_{n,\alpha} \geq w_\alpha - \varepsilon_n$ holds wpa1. Therefore,

$$\begin{aligned}\mathbb{P}(\Theta_I \subseteq \widehat{\Theta}_\alpha) &= \mathbb{P}\Big(\sup_{\theta\in\Theta_I} Q_n(\theta) \leq w_{n,\alpha}\Big)\\ &\geq \mathbb{P}\Big(\sup_{\theta\in\Theta_I} Q_n(\theta) \leq w_\alpha - \varepsilon_n\Big) + o(1),\end{aligned}$$

and the result follows by part (i). If $w_{n,\alpha} = w_\alpha + o_{\mathbb{P}}(1)$, then the proof follows similarly, noting that $|w_{n,\alpha} - w_\alpha| \leq \varepsilon_n$ holds wpa1. *Q.E.D.*

PROOF OF LEMMA 2.2: Follows by similar arguments to the proof of Lemma 2.1. *Q.E.D.*

LEMMA F.1: *Let Assumptions* 4.1(i) *and* 4.2 *hold*. *Then*,

$$\sup_{\theta\in\Theta_{\mathrm{osn}}} \big|Q_n(\theta) - \big\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\big\|^2\big| = o_{\mathbb{P}}(1). \tag{43}$$

And hence $\sup_{\theta\in\Theta_I} Q_n(\theta) = \|\mathbf{T}\mathbb{V}_n\|^2 + o_{\mathbb{P}}(1)$.

PROOF OF LEMMA F.1: By Assumptions 4.1(i) and 4.2, we obtain

$$\begin{aligned}2nL_n(\hat{\theta}) &= \sup_{\theta\in\Theta_{\mathrm{osn}}} 2nL_n(\theta) + o_{\mathbb{P}}(1)\\ &= 2\ell_n + \|\sqrt{n}\hat{\gamma}_n\|^2 - \inf_{\theta\in\Theta_{\mathrm{osn}}} \big\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\big\|^2 + o_{\mathbb{P}}(1) \qquad (44)\\ &= 2\ell_n + \|\mathbf{T}\mathbb{V}_n\|^2 - \inf_{t\in T} \|t - \mathbf{T}\mathbb{V}_n\|^2 + o_{\mathbb{P}}(1),\end{aligned}$$

where $\inf_{t\in T} \|t - \mathbf{T}\mathbb{V}_n\|^2 = 0$ because $\mathbf{T}\mathbb{V}_n \in T$. Now, by Assumption 4.2, for $\theta \in \Theta_{\mathrm{osn}}$,

$$\begin{aligned}Q_n(\theta) &= \big(2\ell_n + \|\mathbf{T}\mathbb{V}_n\|^2 + o_{\mathbb{P}}(1)\big) - \big(2\ell_n + \|\mathbf{T}\mathbb{V}_n\|^2 - \big\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\big\|^2 + o_{\mathbb{P}}(1)\big)\\ &= \big\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\big\|^2 + o_{\mathbb{P}}(1),\end{aligned}$$

where the $o_{\mathbb{P}}(1)$ term holds uniformly over Θ_{osn}. This proves expression (43). Finally, as $\gamma(\theta) = 0$ for $\theta \in \Theta_I$, we have $\sup_{\theta\in\Theta_I} Q_n(\theta) = \|\mathbf{T}\mathbb{V}_n\|^2 + o_{\mathbb{P}}(1)$. *Q.E.D.*

PROOF OF LEMMA 4.1: We first prove equation (22). As $|\Pr(A) - \Pr(A\cap B)| \leq \Pr(B^c)$, we have

$$\begin{aligned}&\sup_z \big|\Pi_n\big(\{\theta : Q_n(\theta) \leq z\}|\mathbf{X}_n\big) - \Pi_n\big(\{\theta : Q_n(\theta) \leq z\} \cap \Theta_{\mathrm{osn}}|\mathbf{X}_n\big)\big|\\ &\quad\leq \Pi_n\big(\Theta_{\mathrm{osn}}^c|\mathbf{X}_n\big) = o_{\mathbb{P}}(1)\end{aligned} \tag{45}$$

by Assumption 4.1(ii). Moreover, by Assumptions 4.1(ii) and 4.3(i),

$$\left|\frac{\int_{\Theta_{\mathrm{osn}}} e^{nL_n(\theta)}\,\mathrm{d}\Pi(\theta)}{\int_{\Theta} e^{nL_n(\theta)}\,\mathrm{d}\Pi(\theta)} - 1\right| = \Pi_n\big(\Theta_{\mathrm{osn}}^c|\mathbf{X}_n\big) = o_{\mathbb{P}}(1),$$

and hence

$$\sup_z \left|\Pi_n\big(\{\theta : Q_n(\theta) \le z\} \cap \Theta_{\mathrm{osn}}|\mathbf{X}_n\big) - \frac{\int_{\{\theta:Q_n(\theta)\le z\}\cap\Theta_{\mathrm{osn}}} e^{nL_n(\theta)}\,\mathrm{d}\Pi(\theta)}{\int_{\Theta_{\mathrm{osn}}} e^{nL_n(\theta)}\,\mathrm{d}\Pi(\theta)}\right| = o_{\mathbb{P}}(1). \tag{46}$$

In view of (45) and (46), it suffices to characterize the large-sample behavior of

$$R_n(z) := \frac{\int_{\{\theta:Q_n(\theta)\le z\}\cap\Theta_{\mathrm{osn}}} e^{nL_n(\theta)-\ell_n-\frac{1}{2}\|\mathbf{T}\mathbb{V}_n\|^2}\,\mathrm{d}\Pi(\theta)}{\int_{\Theta_{\mathrm{osn}}} e^{nL_n(\theta)-\ell_n-\frac{1}{2}\|\mathbf{T}\mathbb{V}_n\|^2}\,\mathrm{d}\Pi(\theta)}. \tag{47}$$

Lemma F.1 and Assumption 4.2 imply that there exists a positive sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ independent of z with $\varepsilon_n = o(1)$ such that the inequalities

$$\sup_{\theta\in\Theta_{\mathrm{osn}}} \left|Q_n(\theta) - \left\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\right\|^2\right| \le \varepsilon_n,$$

$$\sup_{\theta\in\Theta_{\mathrm{osn}}} \left|nL_n(\theta) - \ell_n - \frac{1}{2}\|\mathbf{T}\mathbb{V}_n\|^2 + \frac{1}{2}\left\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\right\|^2\right| \le \varepsilon_n$$

both hold wpa1. Therefore, wpa1, we have

$$\begin{aligned} &e^{-2\varepsilon_n}\frac{\int_{\{\theta:\|\sqrt{n}\gamma(\theta)-\mathbf{T}\mathbb{V}_n\|^2\le z-\varepsilon_n\}\cap\Theta_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma(\theta)-\mathbf{T}\mathbb{V}_n\|^2}\,\mathrm{d}\Pi(\theta)}{\int_{\Theta_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma(\theta)-\mathbf{T}\mathbb{V}_n\|^2}\,\mathrm{d}\Pi(\theta)} \\ &\le R_n(z) \le e^{2\varepsilon_n}\frac{\int_{\{\theta:\|\sqrt{n}\gamma(\theta)-\mathbf{T}\mathbb{V}_n\|^2\le z+\varepsilon_n\}\cap\Theta_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma(\theta)-\mathbf{T}\mathbb{V}_n\|^2}\,\mathrm{d}\Pi(\theta)}{\int_{\Theta_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma(\theta)-\mathbf{T}\mathbb{V}_n\|^2}\,\mathrm{d}\Pi(\theta)} \end{aligned}$$

uniformly in z. Let $\varGamma_{\mathrm{osn}} = \{\gamma(\theta) : \theta \in \varTheta_{\mathrm{osn}}\}$. A change of variables yields

$$\begin{aligned}
& e^{-2\varepsilon_n} \frac{\displaystyle\int_{\{\gamma:\|\sqrt{n}\gamma - \mathbf{T}\mathbb{V}_n\|^2 \le z - \varepsilon_n\} \cap \varGamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma - \mathbf{T}\mathbb{V}_n\|^2} \, \mathrm{d}\varPi_{\varGamma}(\gamma)}{\displaystyle\int_{\varGamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma - \mathbf{T}\mathbb{V}_n\|^2} \, \mathrm{d}\varPi_{\varGamma}(\gamma)} \\
& \quad \le R_n(z) \le e^{2\varepsilon_n} \frac{\displaystyle\int_{\{\gamma:\|\sqrt{n}\gamma - \mathbf{T}\mathbb{V}_n\|^2 \le z + \varepsilon_n\} \cap \varGamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma - \mathbf{T}\mathbb{V}_n\|^2} \, \mathrm{d}\varPi_{\varGamma}(\gamma)}{\displaystyle\int_{\varGamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma - \mathbf{T}\mathbb{V}_n\|^2} \, \mathrm{d}\varPi_{\varGamma}(\gamma)}.
\end{aligned} \tag{48}$$

Recall B_δ from Assumption 4.3(ii). The inclusion $\varGamma_{\mathrm{osn}} \subset B_\delta \cap \varGamma$ holds for all n sufficiently large by Assumption 4.2. Taking n sufficiently large and using Assumption 4.3(ii), we may deduce that there exists a positive sequence $(\bar{\varepsilon}_n)_{n\in\mathbb{N}}$ with $\bar{\varepsilon}_n = o(1)$ such that

$$\left| \frac{\sup_{\gamma \in \varGamma_{\mathrm{osn}}} \pi_{\varGamma}(\gamma)}{\inf_{\gamma \in \varGamma_{\mathrm{osn}}} \pi_{\varGamma}(\gamma)} - 1 \right| \le \bar{\varepsilon}_n$$

for each n. Substituting into (48):

$$\begin{aligned}
& (1 - \bar{\varepsilon}_n) e^{-2\varepsilon_n} \frac{\displaystyle\int_{\{\gamma:\|\sqrt{n}\gamma - \mathbf{T}\mathbb{V}_n\|^2 \le z - \varepsilon_n\} \cap \varGamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma - \mathbf{T}\mathbb{V}_n\|^2} \, \mathrm{d}\gamma}{\displaystyle\int_{\varGamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma - \mathbf{T}\mathbb{V}_n\|^2} \, \mathrm{d}\gamma} \\
& \quad \le R_n(z) \le (1 + \bar{\varepsilon}_n) e^{2\varepsilon_n} \frac{\displaystyle\int_{\{\gamma:\|\sqrt{n}\gamma - \mathbf{T}\mathbb{V}_n\|^2 \le z + \varepsilon_n\} \cap \varGamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma - \mathbf{T}\mathbb{V}_n\|^2} \, \mathrm{d}\gamma}{\displaystyle\int_{\varGamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma - \mathbf{T}\mathbb{V}_n\|^2} \, \mathrm{d}\gamma}
\end{aligned}$$

uniformly in z, where "$\mathrm{d}\gamma$" denotes integration with respect to Lebesgue measure on $\mathbb{R}^{d^*}$.

Let $T_{\mathrm{osn}} = \{\sqrt{n}\gamma : \gamma \in \varGamma_{\mathrm{osn}}\}$ and let B_z denote a ball of radius z in $\mathbb{R}^{d^*}$ centered at the origin. Using the change of variables $\sqrt{n}\gamma - \mathbf{T}\mathbb{V}_n \mapsto \kappa$, we can rewrite the preceding inequalities as

$$(1 - \bar{\varepsilon}_n) e^{-2\varepsilon_n} \frac{\displaystyle\int_{B_{\sqrt{z-\varepsilon_n}} \cap (T_{\mathrm{osn}} - \mathbf{T}\mathbb{V}_n)} e^{-\frac{1}{2}\|\kappa\|^2} \, \mathrm{d}\kappa}{\displaystyle\int_{(T_{\mathrm{osn}} - \mathbf{T}\mathbb{V}_n)} e^{-\frac{1}{2}\|\kappa\|^2} \, \mathrm{d}\kappa} \le R_n(z) \le (1 + \bar{\varepsilon}_n) e^{2\varepsilon_n} \frac{\displaystyle\int_{B_{\sqrt{z+\varepsilon_n}} \cap (T_{\mathrm{osn}} - \mathbf{T}\mathbb{V}_n)} e^{-\frac{1}{2}\|\kappa\|^2} \, \mathrm{d}\kappa}{\displaystyle\int_{(T_{\mathrm{osn}} - \mathbf{T}\mathbb{V}_n)} e^{-\frac{1}{2}\|\kappa\|^2} \, \mathrm{d}\kappa}$$

with the understanding that $B_{\sqrt{z-\varepsilon_n}}$ is empty if $\varepsilon_n > z$.

Let $\nu_{d^*}(A) = (2\pi)^{-d^*/2} \int_A e^{-\frac{1}{2}\|\kappa\|^2} \, \mathrm{d}\kappa$ denote Gaussian measure. We now show that

$$\sup_z \left| \frac{\nu_{d^*}\big(B_{\sqrt{z \pm \varepsilon_n}} \cap (T_{\mathrm{osn}} - \mathbf{T}\mathbb{V}_n)\big)}{\nu_{d^*}(T_{\mathrm{osn}} - \mathbf{T}\mathbb{V}_n)} - \frac{\nu_{d^*}\big(B_{\sqrt{z \pm \varepsilon_n}} \cap (T - \mathbf{T}\mathbb{V}_n)\big)}{\nu_{d^*}(T - \mathbf{T}\mathbb{V}_n)} \right| = o_{\mathbb{P}}(1), \tag{49}$$

$$\sup_z \left| \frac{\nu_{d^*}\big(B_{\sqrt{z\pm\varepsilon_n}} \cap (T - \mathbf{T}\mathbb{V}_n)\big)}{\nu_{d^*}(T - \mathbf{T}\mathbb{V}_n)} - \frac{\nu_{d^*}\big(B_{\sqrt{z}} \cap (T - \mathbf{T}\mathbb{V}_n)\big)}{\nu_{d^*}(T - \mathbf{T}\mathbb{V}_n)} \right| = o_{\mathbb{P}}(1). \tag{50}$$

Consider (49). To simplify presentation, we assume wlog that $T_{\text{osn}} \subseteq T$ (otherwise, we may truncate T and T_{osn} to B_{k_n} as in the proof of Lemma D.3). As

$$\left| \frac{\Pr(A \cap B)}{\Pr(B)} - \frac{\Pr(A \cap C)}{\Pr(C)} \right| \le 2 \frac{\Pr(C \setminus B)}{\Pr(B)} \tag{51}$$

holds for events A, B, C with $B \subseteq C$, we have

$$\sup_z \left| \frac{\nu_{d^*}\big(B_{\sqrt{z\pm\varepsilon_n}} \cap (T_{\text{osn}} - \mathbf{T}\mathbb{V}_n)\big)}{\nu_{d^*}(T_{\text{osn}} - \mathbf{T}\mathbb{V}_n)} - \frac{\nu_{d^*}\big(B_{\sqrt{z\pm\varepsilon_n}} \cap (T - \mathbf{T}\mathbb{V}_n)\big)}{\nu_{d^*}(T - \mathbf{T}\mathbb{V}_n)} \right| \le 2 \frac{\nu_{d^*}\big((T \setminus T_{\text{osn}}) - \mathbf{T}\mathbb{V}_n\big)}{\nu_{d^*}(T_{\text{osn}} - \mathbf{T}\mathbb{V}_n)}.$$

As $\mathbb{V}_n$ is tight and $T \subseteq \mathbb{R}^{d^*}$ has positive volume and T_{osn} covers T, we may deduce that

$$1/\nu_{d^*}(T - \mathbf{T}\mathbb{V}_n) = O_{\mathbb{P}}(1) \quad \text{and} \quad 1/\nu_{d^*}(T_{\text{osn}} - \mathbf{T}\mathbb{V}_n) = O_{\mathbb{P}}(1). \tag{52}$$

It also follows by tightness of $\mathbb{V}_n$ and Assumption 4.2 that $\nu_{d^*}((T \setminus T_{\text{osn}}) - \mathbf{T}\mathbb{V}_n) = o_{\mathbb{P}}(1)$, which proves (49). Result (50) now follows by (52) and the fact that

$$\begin{aligned} &\sup_z \left| \nu_{d^*}\big(B_{\sqrt{z\pm\varepsilon_n}} \cap (T_{\text{osn}} - \mathbf{T}\mathbb{V}_n)\big) - \nu_{d^*}\big(B_{\sqrt{z}} \cap (T_{\text{osn}} - \mathbf{T}\mathbb{V}_n)\big) \right| \\ &\quad \le \sup_z \left| F_{\chi^2_{d^*}}(z \pm \varepsilon_n) - F_{\chi^2_{d^*}}(z) \right| = o(1) \end{aligned}$$

because $\nu_{d^*}(B_{\sqrt{z}}) = F_{\chi^2_{d^*}}(z)$. This completes the proof of result (22).

Part (i) follows by combining (22) and the inequality

$$\sup_z \big(\mathbb{P}_Z\big(\|Z\|^2 \le z | Z \in T - \mathbf{T}v\big) - \mathbb{P}_Z\big(\|\mathbf{T}Z\|^2 \le z\big)\big) \le 0 \quad \text{for all } v \in \mathbb{R}^{d^*} \tag{53}$$

(see Theorem 2 in Chen and Gao (2017)). Part (ii) also follows from (22) by observing that if $T = \mathbb{R}^{d^*}$, then $T - \mathbb{V}_n = \mathbb{R}^{d^*}$. *Q.E.D.*

PROOF OF THEOREM 4.1: We verify the conditions of Lemma 2.1. By Assumption 4.1(i), we have that $L_n(\hat{\theta}) = \sup_{\theta \in \Theta_{\text{osn}}} L_n(\theta) + o_{\mathbb{P}}(n^{-1})$. By Lemma F.1, we have

$$\sup_{\theta \in \Theta_I} Q_n(\theta) = \|\mathbf{T}\mathbb{V}_n\|^2 + o_{\mathbb{P}}(1) \rightsquigarrow \|\mathbf{T}Z\|^2$$

with $Z \sim N(0, I_{d^*})$ when $\Sigma = I_{d^*}$. Let z_α denote the α quantile of the distribution of $\|\mathbf{T}Z\|^2$.

For part (i), Lemma 4.1(i) shows that the posterior distribution of the QLR asymptotically stochastically dominates the distribution of $\|\mathbf{T}Z\|^2$, which implies that $\xi^{\text{post}}_{n,\alpha} \ge z_\alpha + o_{\mathbb{P}}(1)$. Therefore,

$$\xi^{\text{mc}}_{n,\alpha} = z_\alpha + \big(\xi^{\text{post}}_{n,\alpha} - z_\alpha\big) + \big(\xi^{\text{mc}}_{n,\alpha} - \xi^{\text{post}}_{n,\alpha}\big) \ge z_\alpha + \big(\xi^{\text{mc}}_{n,\alpha} - \xi^{\text{post}}_{n,\alpha}\big) + o_{\mathbb{P}}(1) = z_\alpha + o_{\mathbb{P}}(1),$$

where the final equality is by Assumption 4.4.

For part (ii), when $T = \mathbb{R}^{d^*}$ and $\Sigma = I_{d^*}$, we have

$$\sup_{\theta\in\Theta_I} Q_n(\theta) = \|\mathbb{V}_n\|^2 + o_{\mathbb{P}}(1) \rightsquigarrow \chi^2_{d^*}, \quad \text{and hence} \quad z_\alpha = \chi^2_{d^*,\alpha}.$$

Further,

$$\xi^{\mathrm{mc}}_{n,\alpha} = \chi^2_{d^*,\alpha} + \left(\xi^{\mathrm{post}}_{n,\alpha} - \chi^2_{d^*,\alpha}\right) + \left(\xi^{\mathrm{mc}}_{n,\alpha} - \xi^{\mathrm{post}}_{n,\alpha}\right) = \chi^2_{d^*,\alpha} + o_{\mathbb{P}}(1)$$

by Lemma 4.1(ii) and Assumption 4.4. *Q.E.D.*

LEMMA F.2: *Let Assumptions* 4.1(i) *and* 4.2′ *hold*. *Then*:

$$\sup_{\theta\in\Theta_{\mathrm{osn}}} \left|Q_n(\theta) - \left(\left\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\right\|^2 + 2f_{n,\perp}\big(\gamma_\perp(\theta)\big)\right)\right| = o_{\mathbb{P}}(1). \tag{54}$$

And hence $\sup_{\theta\in\Theta_I} Q_n(\theta) = \|\mathbf{T}\mathbb{V}_n\|^2 + o_{\mathbb{P}}(1)$.

PROOF OF LEMMA F.2: Using Assumptions 4.1(i) and 4.2′, we obtain

$$\begin{aligned}
2nL_n(\hat\theta) &= \sup_{\theta\in\Theta_{\mathrm{osn}}} \left(2\ell_n + \|\mathbf{T}\mathbb{V}_n\|^2 - \left\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\right\|^2 - 2f_{n,\perp}\big(\gamma_\perp(\theta)\big)\right) + o_{\mathbb{P}}(1) \\
&= 2\ell_n + \|\mathbf{T}\mathbb{V}_n\|^2 - \inf_{t\in T_{\mathrm{osn}}} \|t - \mathbf{T}\mathbb{V}_n\|^2 - \inf_{\theta\in\Theta_{\mathrm{osn}}} 2f_{n,\perp}\big(\gamma_\perp(\theta)\big) + o_{\mathbb{P}}(1) \qquad (55)\\
&= 2\ell_n + \|\mathbf{T}\mathbb{V}_n\|^2 + o_{\mathbb{P}}(1),
\end{aligned}$$

because $\mathbf{T}\mathbb{V}_n \in T$ and $f_{n,\perp}(\cdot) \geq 0$ with $f_{n,\perp}(0) = 0$ (a.s.), $\gamma_\perp(\theta) = 0$ for all $\theta \in \Theta_I$; thus

$$0 \leq \inf_{\theta\in\Theta_{\mathrm{osn}}} f_{n,\perp}\big(\gamma_\perp(\theta)\big) \leq f_{n,\perp}\big(\gamma_\perp(\bar\theta)\big) = 0 \quad \text{(a.s.) for any } \bar\theta \in \Theta_I.$$

Then by Assumption 4.2′(i) and definition of Q_n, we obtain

$$\begin{aligned}
Q_n(\theta) &= 2\ell_n + \|\mathbf{T}\mathbb{V}_n\|^2 + o_{\mathbb{P}}(1) \\
&\quad - \left(2\ell_n + \|\mathbf{T}\mathbb{V}_n\|^2 - \left\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\right\|^2 - 2f_{n,\perp}\big(\gamma_\perp(\theta)\big) + o_{\mathbb{P}}(1)\right) \\
&= \left\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\right\|^2 + 2f_{n,\perp}\big(\gamma_\perp(\theta)\big) + o_{\mathbb{P}}(1),
\end{aligned}$$

where the $o_{\mathbb{P}}(1)$ term holds uniformly over Θ_{osn}. This proves expression (54).

As $\gamma(\theta) = 0$ and $\gamma_\perp(\theta) = 0$ for $\theta \in \Theta_I$, and $f_{n,\perp}(0) = 0$ (almost surely), we therefore have $\sup_{\theta\in\Theta_I} Q_n(\theta) = \|\mathbf{T}\mathbb{V}_n\|^2 + o_{\mathbb{P}}(1)$. *Q.E.D.*

PROOF OF LEMMA 4.2: We first show that inequality (24) holds. By identical arguments to the proof of Lemma 4.1, it is enough to characterize the large-sample behavior of $R_n(z)$ defined in (47). By Lemma F.2 and Assumption 4.2′, there exists a positive sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ independent of z with $\varepsilon_n = o(1)$ such that

$$\sup_{\theta\in\Theta_{\mathrm{osn}}} \left|Q_n(\theta) - \left(\left\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\right\|^2 + 2f_{n,\perp}\big(\gamma_\perp(\theta)\big)\right)\right| \leq \varepsilon_n,$$

$$\sup_{\theta\in\Theta_{\mathrm{osn}}} \left|nL_n(\theta) - \ell_n - \frac{1}{2}\|\mathbf{T}\mathbb{V}_n\|^2 + \frac{1}{2}\left\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\right\|^2 + f_{n,\perp}\big(\gamma_\perp(\theta)\big)\right| \leq \varepsilon_n$$

both hold wpa1. Also note that for any z, we have

$$\begin{aligned}&\left\{\theta \in \Theta_{\mathrm{osn}} : \left\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\right\|^2 + 2f_{n,\perp}\big(\gamma_\perp(\theta)\big) \pm \varepsilon_n \le z\right\}\\ &\quad\subseteq \left\{\theta \in \Theta_{\mathrm{osn}} : \left\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\right\|^2 \pm \varepsilon_n \le z\right\}\end{aligned}$$

because $f_{n,\perp}(\cdot) \ge 0$. Therefore, wpa1, we have

$$R_n(z) \le e^{2\varepsilon_n} \frac{\displaystyle\int_{\{\theta:\|\sqrt{n}\gamma(\theta)-\mathbf{T}\mathbb{V}_n\|^2\le z+\varepsilon_n\}\cap\Theta_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma(\theta)-\mathbf{T}\mathbb{V}_n\|^2 - f_{n,\perp}(\gamma_\perp(\theta))}\,\mathrm{d}\Pi(\theta)}{\displaystyle\int_{\Theta_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma(\theta)-\mathbf{T}\mathbb{V}_n\|^2 - f_{n,\perp}(\gamma_\perp(\theta))}\,\mathrm{d}\Pi(\theta)}$$

uniformly in z. Define $\Gamma_{\mathrm{osn}} = \{\gamma(\theta) : \theta \in \Theta_{\mathrm{osn}}\}$ and $\Gamma_{\perp,\mathrm{osn}} = \{\gamma_\perp(\theta) : \theta \in \Theta_{\mathrm{osn}}\}$. By similar arguments to the proof of Lemma 4.1, Assumption 4.3′(ii) and a change of variables yield

$$R_n(z) \le e^{2\varepsilon_n}(1+\bar{\varepsilon}_n) \frac{\displaystyle\int_{(\{\gamma:\|\sqrt{n}\gamma-\mathbf{T}\mathbb{V}_n\|^2\le z+\varepsilon_n\}\cap\Gamma_{\mathrm{osn}})\times\Gamma_{\perp,\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma-\mathbf{T}\mathbb{V}_n\|^2 - f_{n,\perp}(\gamma_\perp)}\,\mathrm{d}(\gamma,\gamma_\perp)}{\displaystyle\int_{\Gamma_{\mathrm{osn}}\times\Gamma_{\perp,\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma-\mathbf{T}\mathbb{V}_n\|^2 - f_{n,\perp}(\gamma_\perp)}\,\mathrm{d}(\gamma,\gamma_\perp)},$$

which holds uniformly in z (wpa1) for some $\bar{\varepsilon}_n = o(1)$. By Tonelli's theorem and Assumption 4.2′(ii), the preceding inequality becomes

$$R_n(z) \le e^{2\varepsilon_n}(1+\bar{\varepsilon}_n) \frac{\displaystyle\int_{(\{\gamma:\|\sqrt{n}\gamma-\mathbf{T}\mathbb{V}_n\|^2\le z+\varepsilon_n\}\cap\Gamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma-\mathbf{T}\mathbb{V}_n\|^2}\,\mathrm{d}\gamma}{\displaystyle\int_{\Gamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma-\mathbf{T}\mathbb{V}_n\|^2}\,\mathrm{d}\gamma}.$$

The rest of the proof of inequality (24) follows by similar arguments to the proof of Lemma 4.1. The conclusion now follows by combining inequalities (24) and (53). *Q.E.D.*

PROOF OF THEOREM 4.2: We verify the conditions of Lemma 2.1. Again, we have that $L_n(\hat{\theta}) = \sup_{\theta\in\Theta_{\mathrm{osn}}} L_n(\theta) + o_{\mathbb{P}}(n^{-1})$. By Lemma F.2, when $\Sigma = I_{d^*}$, we have

$$\sup_{\theta\in\Theta_I} Q_n(\theta) = \|\mathbf{T}\mathbb{V}_n\|^2 + o_{\mathbb{P}}(1) \rightsquigarrow \|\mathbf{T}Z\|^2, \tag{56}$$

where $Z \sim N(0, I_{d^*})$. Lemma 4.2 shows that the posterior distribution of the QLR asymptotically stochastically dominates the F_T distribution. The result follows by the same arguments as the proof of Theorem 4.1(i). *Q.E.D.*

LEMMA F.3: *Let Assumptions* 4.1(i) *and* 4.2 *or* 4.2′ *and* 4.5 *hold*. *Then*:

$$\sup_{\theta\in\Theta_{\mathrm{osn}}} \left|PQ_n\big(M(\theta)\big) - f\big(\mathbf{T}\mathbb{V}_n - \sqrt{n}\gamma(\theta)\big)\right| = o_{\mathbb{P}}(1).$$

PROOF OF LEMMA F.3: By display (44) in the proof of Lemma F.1 or display (55) in the proof of Lemma F.2 and Assumption 4.5, we obtain

$$\begin{aligned} PQ_n\big(M(\theta)\big) &= 2nL_n(\hat{\theta}) - 2nPL_n\big(M(\theta)\big) \\ &= 2\ell_n + \|\mathbf{T}\mathbb{V}_n\|^2 - \big(2\ell_n + \|\mathbf{T}\mathbb{V}_n\|^2 - f\big(\mathbf{T}\mathbb{V}_n - \sqrt{n}\gamma(\theta)\big)\big) + o_{\mathbb{P}}(1), \end{aligned}$$

where the $o_{\mathbb{P}}(1)$ term holds uniformly over Θ_{osn}. *Q.E.D.*

PROOF OF LEMMA 4.3: We first prove equation (26) under Assumptions 4.1, 4.2′, 4.3′, and 4.5. The proof under Assumptions 4.1, 4.2, 4.3, and 4.5 follows similarly. By the same arguments as the proof of Lemma 4.1, it suffices to characterize the large-sample behavior of

$$R_n(z) := \frac{\displaystyle\int_{\{\theta: PQ_n(M(\theta)) \le z\} \cap \Theta_{\mathrm{osn}}} e^{nL_n(\theta)}\, \mathrm{d}\Pi(\theta)}{\displaystyle\int_{\Theta_{\mathrm{osn}}} e^{nL_n(\theta)}\, \mathrm{d}\Pi(\theta)}. \tag{57}$$

By Lemma F.3 and Assumption 4.2′, there exists a positive sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ independent of z with $\varepsilon_n = o(1)$ such that the inequalities

$$\sup_{\theta\in\Theta_{\mathrm{osn}}} \big| PQ_n\big(M(\theta)\big) - f\big(\mathbf{T}\mathbb{V}_n - \sqrt{n}\gamma(\theta)\big)\big| \le \varepsilon_n,$$
$$\sup_{\theta\in\Theta_{\mathrm{osn}}} \left| nL_n(\theta) - \ell_n - \frac{1}{2}\|\mathbf{T}\mathbb{V}_n\|^2 - \left(-\frac{1}{2}\big\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\big\|^2 - f_{n,\perp}\big(\gamma_\perp(\theta)\big)\right)\right| \le \varepsilon_n$$

both hold wpa1. Therefore, wpa1, we have

$$\begin{aligned} &e^{-2\varepsilon_n} \frac{\displaystyle\int_{\{\theta: f(\mathbf{T}\mathbb{V}_n - \sqrt{n}\gamma(\theta)) \le z - \varepsilon_n\} \cap \Theta_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\|^2 - f_{n,\perp}(\gamma_\perp(\theta))}\, \mathrm{d}\Pi(\theta)}{\displaystyle\int_{\Theta_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\|^2 - f_{n,\perp}(\gamma_\perp(\theta))}\, \mathrm{d}\Pi(\theta)} \\ &\le R_n(z) \le e^{2\varepsilon_n} \frac{\displaystyle\int_{\{\theta: f(\mathbf{T}\mathbb{V}_n - \sqrt{n}\gamma(\theta)) \le z + \varepsilon_n\} \cap \Theta_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\|^2 - f_{n,\perp}(\gamma_\perp(\theta))}\, \mathrm{d}\Pi(\theta)}{\displaystyle\int_{\Theta_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma(\theta) - \mathbf{T}\mathbb{V}_n\|^2 - f_{n,\perp}(\gamma_\perp(\theta))}\, \mathrm{d}\Pi(\theta)} \end{aligned}$$

uniformly in z. By similar arguments to the proof of Lemma 4.2, we may use the change of variables $\theta \mapsto (\gamma(\theta), \gamma_\perp(\theta))$, continuity of π_{Γ^*} (Assumption 4.3′(ii)), and Tonelli's theorem to restate the preceding inequalities as

$$(1-\bar{\varepsilon}_n)e^{-2\varepsilon_n}\frac{\int_{\{\gamma:f(\mathbf{T}\mathbb{V}_n-\sqrt{n}\gamma)\le z-\varepsilon_n\}\cap\Gamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma-\mathbf{T}\mathbb{V}_n\|^2}\,\mathrm{d}\gamma}{\int_{\Gamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma-\mathbf{T}\mathbb{V}_n\|^2}\,\mathrm{d}\gamma}$$

$$\le R_n(z)\le(1+\bar{\varepsilon}_n)e^{2\varepsilon_n}\frac{\int_{\{\gamma:f(\mathbf{T}\mathbb{V}_n-\sqrt{n}\gamma)\le z+\varepsilon_n\}\cap\Gamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma-\mathbf{T}\mathbb{V}_n\|^2}\,\mathrm{d}\gamma}{\int_{\Gamma_{\mathrm{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma-\mathbf{T}\mathbb{V}_n\|^2}\,\mathrm{d}\gamma},$$

which holds (wpa1) for some $\bar{\varepsilon}_n=o(1)$. Let $f^{-1}(z)=\{\kappa\in\mathbb{R}^{d^*}:f(\kappa)\le z\}$. A second change of variables $\mathbf{T}\mathbb{V}_n-\sqrt{n}\gamma\mapsto\kappa$ yields

$$(1-\bar{\varepsilon}_n)e^{-2\varepsilon_n}\frac{\nu_{d^*}\big(\big(f^{-1}(z-\varepsilon_n)\big)\cap(\mathbf{T}\mathbb{V}_n-T_{\mathrm{osn}})\big)}{\nu_{d^*}(\mathbf{T}\mathbb{V}_n-T_{\mathrm{osn}})}$$

$$\le R_n(z)\le(1+\bar{\varepsilon}_n)e^{2\varepsilon_n}\frac{\nu_{d^*}\big(\big(f^{-1}(z+\varepsilon_n)\big)\cap(\mathbf{T}\mathbb{V}_n-T_{\mathrm{osn}})\big)}{\nu_{d^*}(\mathbf{T}\mathbb{V}_n-T_{\mathrm{osn}})}$$

uniformly in z, where it should be understood that $\mathbf{T}\mathbb{V}_n-T_{\mathrm{osn}}$ is the Minkowski sum $\mathbf{T}\mathbb{V}_n+(-T_{\mathrm{osn}})$ with $-T_{\mathrm{osn}}=\{-\kappa:\kappa\in T_{\mathrm{osn}}\}$.

The result now follows by similar arguments to the proof of Lemma 4.1, noting that

$$\sup_{z\in I}\left|\frac{\nu_{d^*}\big(\big(f^{-1}(z\pm\varepsilon_n)\big)\cap(\mathbf{T}\mathbb{V}_n-T)\big)}{\nu_{d^*}(\mathbf{T}\mathbb{V}_n-T)}-\mathbb{P}_{Z|\mathbf{X}_n}\big(f(Z)\le z|Z\in\mathbf{T}\mathbb{V}_n-T\big)\right|$$

$$\le\sup_{z\in I}\left|\frac{\nu_{d^*}\big(f^{-1}(z\pm\varepsilon_n)\big)-\nu_{d^*}\big(f^{-1}(z)\big)}{\nu_{d^*}(\mathbf{T}\mathbb{V}_n-T)}\right|=o_{\mathbb{P}}(1),$$

where the final equality is by uniform continuity of bounded, monotone continuous functions and display (52).

Part (i) follows by combining equation (26) and the following inequality:

$$\mathbb{P}_Z\big(f(Z)\le z|Z\in\mathbf{T}v-T\big)\le\mathbb{P}_Z\big(f(\mathbf{T}Z)\le z\big)\tag{58}$$

holds for all $z\in I$ and for any given $v\in\mathbb{R}^{d^*}$. To prove this, it suffices to show that

$$\nu_{d^*}\big(f^{-1}(z)\cap(\mathbf{T}v-T)\big)\le\nu_{d^*}(\mathbf{T}v-T)\times\nu_{d^*}\big(\big\{\kappa\in\mathbb{R}^{d^*}:f(\mathbf{T}\kappa)\le z\big\}\big)\tag{59}$$

holds for all $z\in I$ and any given $v\in\mathbb{R}^{d^*}$. Since f is quasiconvex, we have

$$\nu_{d^*}\big(f^{-1}(z)\cap(\mathbf{T}v-T)\big)\le\nu_{d^*}\big(\big(f^{-1}(z)-T^o\big)\cap(\mathbf{T}v-T)\big)\le\nu_{d^*}\big(f^{-1}(z)-T^o\big)\times\nu_{d^*}(\mathbf{T}v-T),$$

where the first inequality is because $f^{-1}(z)\subseteq f^{-1}(z)-T^o=\{\kappa_1+\kappa_2:\kappa_1\in f^{-1}(z),-\kappa_2\in T^o\}$ as $0\in T^o$ and the second inequality is by Theorem 1 of Chen and Gao (2017) (taking $A=\{\mathbf{T}v\}$, $B=f^{-1}(z)$, $C=-T$, and $D=-T^o$ in their notation). Hence (59) holds whenever

$$\nu_{d^*}\big(f^{-1}(z) - T^o\big) \le \nu_{d^*}\big(\{\kappa \in \mathbb{R}^{d^*} : f(\mathbf{T}\kappa) \le z\}\big) \tag{60}$$

holds, which does hold when f is subconvex.

Part (ii) also follows from equation (26) by observing that if $T = \mathbb{R}^{d^*}$, then $T - \mathbb{V}_n = \mathbb{R}^{d^*}$. *Q.E.D.*

PROOF OF THEOREM 4.3: We verify the conditions of Lemma 2.2. Again, we have that $L_n(\hat\theta) = \sup_{\theta\in\Theta_{\text{osn}}} L_n(\theta) + o_{\mathbb{P}}(n^{-1})$.

To prove Theorem 4.3(i), let ξ_α denote the α quantile of $f(\mathbf{T}Z)$. By Lemma 4.3(i), we have

$$\xi^{\text{mc},p}_{n,\alpha} = \xi_\alpha + \big(\xi^{\text{post},p}_{n,\alpha} - \xi_\alpha\big) + \big(\xi^{\text{mc},p}_{n,\alpha} - \xi^{\text{post},p}_{n,\alpha}\big) \ge \xi_\alpha + \big(\xi^{\text{mc},p}_{n,\alpha} - \xi^{\text{post},p}_{n,\alpha}\big) + o_{\mathbb{P}}(1) = \xi_\alpha + o_{\mathbb{P}}(1),$$

where the final equality is by Assumption 4.6. Since G_T is continuous at its α quantile ξ_α, from the proof of Lemma 4.3(i), it is clear that Theorem 4.3(i) remains valid under the weaker condition that (i) f is quasiconvex and (ii) $\mathbb{P}_Z(Z \in (f^{-1}(\xi_\alpha) - T^o)) \le G_T(\xi_\alpha)$.

To prove Theorem 4.3(ii), when $T = \mathbb{R}^{d^*}$ we have $PQ_n(M_I) \rightsquigarrow f(Z)$. Let ξ_α denote the α quantile of $f(Z)$. Then:

$$\xi^{\text{mc}}_{n,\alpha} = \xi_\alpha + \big(\xi^{\text{post}}_{n,\alpha} - \xi_\alpha\big) + \big(\xi^{\text{mc}}_{n,\alpha} - \xi^{\text{post}}_{n,\alpha}\big) = \xi_\alpha + o_{\mathbb{P}}(1)$$

by Lemma 4.3(ii) and Assumption 4.6. *Q.E.D.*

PROOF OF THEOREM 4.4: By Lemma 2.2, it is enough to show that $\Pr(W^* \le w) \ge F_{\chi^2_1}(w)$ holds for $w \ge 0$, where $W^* = \max_{i\in\{1,2\}} \inf_{t\in T_i} \|Z - t\|^2$.

Case 1: $d^* = 1$. Wlog let $T_1 = [0,\infty)$ and $T_1^o = (-\infty, 0]$. If $T_2 = T_1$, then $\mathbf{T}_1^o Z = \mathbf{T}_2^o Z = (Z \wedge 0)$, so $W^* = (Z\wedge 0)^2 \le Z^2 \sim \chi^2_1$. If $T_2 = T_1^o$, then $\mathbf{T}_1^o Z = (Z \wedge 0)$ and $\mathbf{T}_2^o Z = (Z \vee 0)$, so $W^* = Z^2 \sim \chi^2_1$. In either case, we have $\Pr(W^* \le w) \ge F_{\chi^2_1}(w)$ for any $w \ge 0$.

Case 2: $d^* = 2$. Wlog let $T_1 = \{(x,y) : y \le 0\}$; then T_1^o is the positive y-axis. Let $Z = (X,Y)'$. If $T_1 = T_2$, then $\mathbf{T}_1^o Z = \mathbf{T}_2^o Z = (Y \vee 0)$, so $W^* = (Y \vee 0)^2 \le Y^2 \sim \chi^2_1$. If $T_2 = \{(x,y) : y \ge 0\}$, then T_2^o is the negative y-axis. So, in this case, $\mathbf{T}_1^o Z = (Y \vee 0)$, $\mathbf{T}_2^o = (Y \wedge 0)$, and so $W^* = Y^2 \sim \chi^2_1$.

Now let T_2 be the rotation of T_1 by $\varphi \in (0,\pi)$ radians. This is plotted in Figure 6 for $\varphi \in (0, \pi/2)$ (left panel) and $\varphi \in (\frac{\pi}{2}, \pi)$ (right panel). The axis of symmetry is the line $y = -x\cot(\frac{\varphi}{2})$, which bisects the angle between T_1^o and T_2^o.

Suppose $Z = (X,Y)'$ lies in the half-space $Y \ge -X\cot(\frac{\varphi}{2})$. There are three options:

- $Z \in (T_1 \cap T_2)$ (purple region): $\mathbf{T}_1^o Z = 0$, $\mathbf{T}_2^o Z = 0$, so $W^* = 0$.
- $Z \in (T_1^c \cap T_2)$ (red region): $\mathbf{T}_1^o Z = (0,Y)'$, $\mathbf{T}_2^o Z = 0$, so $W^* = Y^2$.
- $Z \in (T_1^c \cap T_2^c)$ (white region): $\mathbf{T}_1^o Z = (0,Y)'$. To calculate $\mathbf{T}_2^o Z$, observe that if we rotate about the origin by $-\varphi$, then the polar cone $\mathbf{T}_2^o$ becomes the positive y-axis. Under the rotation, $\mathbf{T}_2^o Z = (0, Y^*)$ where Y^* is the y-value of the rotation of (X,Y) by negative φ. The point (X,Y) rotates to $(X\cos\varphi + Y\sin\varphi, Y\cos\varphi - X\sin\varphi)$, so we get $\|\mathbf{T}_2^o Z\|^2 = (Y\cos\varphi - X\sin\varphi)^2$. We assumed $Y \ge -X\cot(\frac{\varphi}{2})$. By the half-angle formula $\cot(\frac{\varphi}{2}) = \frac{\sin\varphi}{1-\cos\varphi}$, this means that $Y \ge Y\cos\varphi - X\sin\varphi$. But $Y\cos\varphi - X\sin\varphi \ge 0$ as $Y \ge X\tan\varphi$. Therefore, $(Y\cos\varphi - X\sin\varphi)^2 \le Y^2$ and so $W^* = Y^2$.

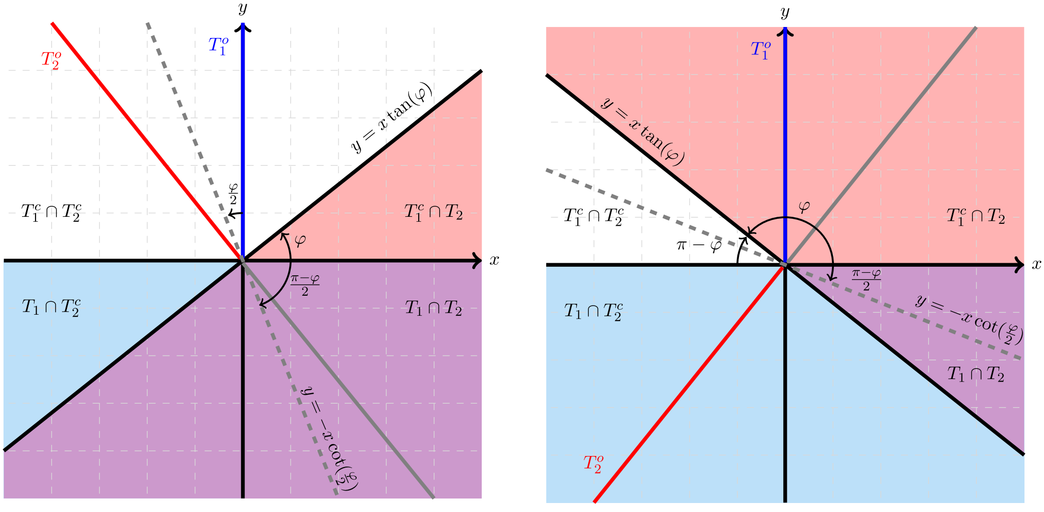

FIGURE 6.—Cones and polar cones for the proof of Theorem 4.4.

We have shown that $W^* \leq Y^2$ whenever $Y \geq -X\cot(\frac{\varphi}{2})$. Now, for any $w \geq 0$,

$$\begin{aligned}\Pr\left(W^* \leq w \middle| Y \geq -X\cot\left(\frac{\varphi}{2}\right)\right) &\geq \Pr\left(Y^2 \leq w \middle| Y \geq -X\cot\left(\frac{\varphi}{2}\right)\right) \\ &= \Pr\big(Y^2 \leq w | V \geq 0\big),\end{aligned} \tag{61}$$

where $V = Y\sin(\frac{\varphi}{2}) + X\cos(\frac{\varphi}{2})$. Note that Y and V are jointly normal with mean 0, unit variance, and correlation $\rho = \sin(\frac{\varphi}{2})$. The pdf of Y given $V \geq 0$ is

$$f(y|V \geq 0) = \frac{\int_0^\infty f_{Y|V}(y|v) f_V(v)\,\mathrm{d}v}{\int_0^\infty f_V(v)\,\mathrm{d}v} = 2f_Y(y)\big(1 - F_{V|Y}(0|y)\big).$$

As $V|Y = y \sim N(\rho y, (1-\rho^2))$, we have

$$F_{V|Y}(0|y) = \Phi\left(\frac{-\rho y}{\sqrt{1-\rho^2}}\right) = 1 - \Phi\left(\frac{\rho}{\sqrt{1-\rho^2}}y\right)$$

and so

$$f(y|V \geq 0) = 2\phi(y)\Phi\left(\frac{\rho}{\sqrt{1-\rho^2}}y\right).$$

Therefore,

$$\Pr\big(Y^2 \leq w | V \geq 0\big) = \Pr(-\sqrt{w} \leq y \leq \sqrt{w} | V \geq 0) = \int_{-\sqrt{w}}^{\sqrt{w}} 2\phi(y)\Phi\left(\frac{\rho}{\sqrt{1-\rho^2}}y\right)\mathrm{d}y. \tag{62}$$

But differentiating the right-hand side of (62) with respect to ρ gives

$$\frac{\mathrm{d}}{\mathrm{d}\rho}\int_{-\sqrt{w}}^{\sqrt{w}} 2\phi(y)\Phi\left(\frac{\rho}{\sqrt{1-\rho^2}}y\right)\mathrm{d}y = \frac{1}{\left(1-\rho^2\right)^{3/2}}\int_{-\sqrt{w}}^{\sqrt{w}} 2y\phi(y)\phi\left(\frac{\rho}{\sqrt{1-\rho^2}}y\right)\mathrm{d}y = 0$$

for any $\rho \in (-1,1)$, because $y\phi(y)\phi(\rho y/\sqrt{1-\rho^2})$ is an odd function. Therefore, the probability in display (62) does not depend on the value of ρ. Setting $\rho = 0$, we obtain

$$\Pr\left(Y^2 \le w | V \ge 0\right) = \int_{-\sqrt{w}}^{\sqrt{w}} 2\phi(y)\Phi(0)\,\mathrm{d}y = \Phi(\sqrt{w}) - \Phi(-\sqrt{w}) = F_{\chi_1^2}(w).$$

Therefore, by inequality (61), we have

$$\Pr\left(W^* \le w | Y \ge -X\cot\left(\frac{\varphi}{2}\right)\right) \ge F_{\chi_1^2}(w).$$

By symmetry, we also have $\Pr(W^* \le w | Y < -X\cot(\frac{\varphi}{2})) \ge F_{\chi_1^2}(w)$. Therefore, we have shown that $\Pr(W^* \le w) \ge F_{\chi_1^2}(w)$ holds for each $w \ge 0$. A similar argument applies when T_2 is the rotation of T_1 by $\varphi \in (-\pi, 0)$ radians. This completes the proof of the case $d^* = 2$.

Case 3: $d^* \ge 3$. As T_1 and T_2 are closed half-spaces, we have $T_1 = \{z \in \mathbb{R}^{d^*} : a'z \le 0\}$ and $T_2 = \{z \in \mathbb{R}^{d^*} : b'z \le 0\}$ for some $a, b \in \mathbb{R}^{d^*} \setminus \{0\}$. The polar cones are the rays $T_1^o = \{sa : s \ge 0\}$ and $T_2^o = \{sb : s \ge 0\}$. There are three sub-cases to consider.

Case 3a: $a = sb$ for some $s > 0$. Let $u_a = \frac{a}{\|a\|}$. Here $T_1 = T_2$, $\mathbf{T}_1^o Z = \mathbf{T}_2^o Z = 0$ if $Z \in T_1$, and

$$\mathbf{T}_1^o Z = \mathbf{T}_2^o Z = u_a\left(Z'u_a\right) \quad \text{if } Z \notin T_1 \text{ (i.e., if } Z'u_a > 0).$$

Therefore, $W^* = (Z'u_a \vee 0)^2 \le (Z'u_a)^2 \sim \chi_1^2$.

Case 3b: $a = sb$ for some $s < 0$. Here $T_1 = -T_2$ and $T_1^o = -T_2^o$, so $\mathbf{T}_1^o Z = 0$ and $\mathbf{T}_2^o Z = u_a(Z'u_a)$ if $Z \in T_1$ (i.e., if $Z'u_a \le 0$) and $\mathbf{T}_1^o Z = u_a(Z'u_a)$ and $\mathbf{T}_2^o Z = 0$ if $Z \notin T_1$ (i.e., if $Z'u_a > 0$). Therefore, $W^* = (Z'u_a)^2 \sim \chi_1^2$.

Case 3c: a and b are linearly independent. Without loss of generality,[22] we can take T_1^o to be the positive y-axis (i.e., $a = (0, a_2, 0, \dots, 0)'$ for some $a_2 > 0$) and take T_2^o to lie in the (x, y)-plane (i.e., $b = (b_1, b_2, 0, \dots, 0)'$ for some $b_1 \ne 0$).

Now write $Z = (X, Y, U)$ where $U \in \mathbb{R}^{d^*-2}$. Note that $a'Z = a_2 Y$ and $b'Z = b_1 X + b_2 Y$. So only the values of X and Y matter in determining whether or not Z belongs to T_1 and T_2.

Without loss of generality, we may assume that $(b_1, b_2)'$ is, up to scale, a rotation of $(0, a_2)'$ by $\varphi \in (0, \pi)$ (the case $(-\pi, 0)$ can be handled by similar arguments, as in Case 2).

Suppose that $Y \ge -X\cot(\frac{\varphi}{2})$. As in Case 2, there are three options:

- $Z \in (T_1 \cap T_2)$: $\mathbf{T}_1^o Z = 0$, $\mathbf{T}_2^o Z = 0$, so $W^* = 0$.
- $Z \in (T_1^c \cap T_2)$: $\mathbf{T}_1^o Z = (0, Y, 0, \dots, 0)'$, $\mathbf{T}_2^o Z = 0$, so $W^* = Y^2$.
- $Z \in (T_1^c \cap T_2^c)$: $\|\mathbf{T}_1^o Z\|^2 = Y^2$ and $\|\mathbf{T}_2^o Z\|^2 = (Y\cos\varphi - X\sin\varphi)^2 \le Y^2$, so $W^* = Y^2$.

Arguing as in Case 2, we obtain $\Pr(W^* \le w | Y \ge -X\cot(\frac{\varphi}{2})) \ge F_{\chi_1^2}(w)$. By symmetry, we also have $\Pr(W^* \le w | Y < -X\cot(\frac{\varphi}{2})) \ge F_{\chi_1^2}(w)$. Therefore, $\Pr(W^* \le w) \ge F_{\chi_1^2}(w)$. *Q.E.D.*

[22] By Gram-Schmidt, we can always define a new set of coordinate vectors $e_1, e_2, \dots, e_{d^*}$ for $\mathbb{R}^{d^*}$ with $e_2 = u_a$ and such that b is in the span of e_1 and e_2.

PROOF OF PROPOSITION 4.1: It follows from condition (i) and display (44) or display (55) that

$$2nL_n(\hat{\theta}) = 2\ell_n + \|\mathbb{V}_n\|^2 + o_{\mathbb{P}}(1).$$

Moreover, applying conditions (ii) and (iii), we obtain

$$\begin{aligned}\inf_{\mu\in M_I}\sup_{\eta\in H_\mu} 2nL_n(\mu,\eta) &= \min_{\mu\in\{\underline{\mu},\overline{\mu}\}}\sup_{\eta\in H_\mu} 2nL_n(\mu,\eta) + o_{\mathbb{P}}(1)\\ &= \min_{\mu\in\{\underline{\mu},\overline{\mu}\}}\Big(2\ell_n + \|\mathbb{V}_n\|^2 - \inf_{t\in T_\mu}\|\mathbb{V}_n - t\|^2\Big) + o_{\mathbb{P}}(1).\end{aligned}$$

Therefore,

$$\sup_{\mu\in M_I}\inf_{\eta\in H_\mu} Q_n(\mu,\eta) = \max_{\mu\in\{\underline{\mu},\overline{\mu}\}}\inf_{t\in T_\mu}\|\mathbb{V}_n - t\|^2 + o_{\mathbb{P}}(1).$$

The result now follows from $\Sigma = I_{d^*}$. *Q.E.D.*

F.2. *Proofs and Additional Lemmas for Section 5*

PROOF OF PROPOSITION 5.1: Wlog we can take $\tilde{\gamma}_0 = 0$. Also take n large enough that $\{\tilde{\gamma} : \|\tilde{\gamma}\| \le n^{-1/4}\} \subseteq U$. Then, by condition (b), for any such $\tilde{\gamma}$, we have

$$nL_n(\tilde{\gamma}) = nL_n(\tilde{\gamma}_0) + (\sqrt{n}\tilde{\gamma})'(\sqrt{n}\mathbb{P}_n\dot{\ell}_{\tilde{\gamma}_0}) + \frac{1}{2}(\sqrt{n}\tilde{\gamma})'(\mathbb{P}_n\ddot{\ell}_{\tilde{\gamma}^*})(\sqrt{n}\tilde{\gamma}),$$

where $\tilde{\gamma}^*$ is in the segment between $\tilde{\gamma}$ and $\tilde{\gamma}_0$ for each element of $\mathbb{P}_n\ddot{\ell}_{\tilde{\gamma}^*}$. We may deduce from Lemma 2.4 of Newey and McFadden (1994) that $\sup_{\tilde{\gamma}:\|\tilde{\gamma}\|\le n^{-1/4}}\|(\mathbb{P}_n\ddot{\ell}_{\tilde{\gamma}^*}) - P_0(\ddot{\ell}_{\tilde{\gamma}_0})\| = o_{\mathbb{P}}(1)$ holds under conditions (a) and (b). As this term is $o_{\mathbb{P}}(1)$, we can choose a positive sequence $(r_n)_{n\in\mathbb{N}}$ with $r_n\to\infty$, $r_n = o(n^{1/4})$ such that $r_n^2\sup_{\tilde{\gamma}:\|\tilde{\gamma}\|\le n^{1/4}}\|(\mathbb{P}_n\ddot{\ell}_{\tilde{\gamma}^*}) - P_0(\ddot{\ell}_{\tilde{\gamma}_0})\| = o_{\mathbb{P}}(1)$. Assumption 4.2 then holds over $\Theta_{\text{osn}} = \{\theta\in\Theta : \|\tilde{\gamma}(\theta)\| \le r_n/\sqrt{n}\}$ with $\ell_n = nL_n(\tilde{\gamma}_0)$, $\gamma(\theta) = \mathbb{I}_0^{1/2}\tilde{\gamma}(\theta)$, $\sqrt{n}\hat{\gamma}_n = \mathbb{V}_n = \mathbb{I}_0^{-1/2}\mathbb{G}_n(\dot{\ell}_{\tilde{\gamma}_0})$, and $\Sigma = I_{d^*}$ because $\mathbb{I}_0 = P_0(\dot{\ell}_{\tilde{\gamma}_0}\dot{\ell}'_{\tilde{\gamma}_0})$.

It remains to show that the posterior concentrates on Θ_{osn}. Choose ε sufficiently small that $U_\varepsilon = \{\tilde{\gamma} : \|\tilde{\gamma}\| < \varepsilon\} \subseteq U$. By a similar expansion to the above and condition (c), we have $D_{\text{KL}}(p_0\|q_{\tilde{\gamma}}) = -\frac{1}{2}\tilde{\gamma}'P_0(\ddot{\ell}_{\tilde{\gamma}^*})\tilde{\gamma}$, where $\tilde{\gamma}^*$ is in the segment between $\tilde{\gamma}^*$ and $\tilde{\gamma}_0$. As $\|P_0(\ddot{\ell}_{\tilde{\gamma}^*}) + \mathbb{I}_0\| \to 0$ as $\|\tilde{\gamma}\|\to 0$, we may reduce ε so that $\inf_{\tilde{\gamma}\in U_\varepsilon}\|P_0(\ddot{\ell}_{\tilde{\gamma}^*}) + \mathbb{I}_0\| \le \frac{1}{2}\lambda_{\min}(\mathbb{I}_0)$. On U_ε we then have that there exist finite positive constants $\underline{c}$ and $\overline{c}$ such that $\underline{c}\|\tilde{\gamma}\|^2 \le D_{\text{KL}}(p_0\|q_{\tilde{\gamma}}) \le \overline{c}\|\tilde{\gamma}\|^2$. Also note that $\inf_{\tilde{\gamma}\in\tilde{\Gamma}\setminus U_\varepsilon} D_{\text{KL}}(p_0\|q_{\tilde{\gamma}}) =: \delta$ with $\delta > 0$ by identifiability of $\tilde{\gamma}_0$, continuity of the map $\tilde{\gamma}\mapsto P_0\ell_{\tilde{\gamma}}$, and compactness of $\tilde{\Gamma}$. Standard consistency arguments (e.g., the Corollary to Theorem 6.1 in Schwartz (1965)) then imply that $\Pi_n(U_\varepsilon|\mathbf{X}_n) \to_{a.s.} 1$. As the posterior concentrates on U_ε and $\Theta_{\text{osn}} \subset U_\varepsilon$ for all n sufficiently large, it is enough to confine attention to U_ε. We have shown that $\underline{c}\|\tilde{\gamma}\|^2 \le D_{\text{KL}}(p_0\|q_{\tilde{\gamma}}) \le \overline{c}\|\tilde{\gamma}\|^2$ holds on U_ε. It now follows by the parametric Bernstein–von Mises theorem (e.g., Theorem 10.1 in van der Vaart (2000)) that the posterior contracts at a $\sqrt{n}$-rate, verifying Assumption 4.1(ii). *Q.E.D.*

For the following lemma, let $(r_n)_{n\in\mathbb{N}}$ be a positive sequence with $r_n\to\infty$ and $r_n = o(n^{1/2})$, $\mathcal{P}_{\text{osn}} = \{p\in\mathcal{P} : h(p,p_0) \le r_n/\sqrt{n}\}$ and $\Theta_{\text{osn}} = \{\theta\in\Theta : h(p_\theta,p_0) \le r_n/\sqrt{n}\}$. For each $p\in\mathcal{P}$ with $p\neq p_0$, define $S_p = \sqrt{p/p_0} - 1$ and $s_p = S_p/h(p,p_0)$. Recall the definitions of $\overline{\mathcal{D}}_\varepsilon$, the tangent cone $\mathcal{T}$, and the projection $\mathbb{T}$ from Section 5.1.2. We say $\mathcal{P}$ is

r_n-DQM if each p is absolutely continuous with respect to p_0 and, for each $p \in \mathcal{P}$, there are $g_p \in \mathcal{T}$ and $R_p \in L^2(\lambda)$ such that

$$\sqrt{p} - \sqrt{p_0} = g_p \sqrt{p_0} + h(p, p_0) R_p$$

with $\sup\{r_n \|R_p\|_{L^2(\lambda)} : h(p, p_0) \le r_n/\sqrt{n}\} \to 0$ as $n \to \infty$. Let $\overline{\mathcal{D}}_\varepsilon^2 = \{d^2 : d \in \overline{\mathcal{D}}_\varepsilon\}$.

LEMMA F.4: *Let the following conditions hold*:
(i) $\mathcal{P}$ *is* r_n*-DQM*;
(ii) *there exists* $\varepsilon > 0$ *such that* $\overline{\mathcal{D}}_\varepsilon^2$ *is* P_0*-Glivenko–Cantelli and* $\overline{\mathcal{D}}_\varepsilon$ *has envelope* $D \in L^2(P_0)$ *with* $\max_{i \le i \le n} D(X_i) = o_{\mathbb{P}}(\sqrt{n}/r_n^3)$;
(iii) $\sup_{p \in \mathcal{P}_{\text{osn}}} |\mathbb{G}_n(S_p - \mathbb{T}S_p)| = o_{\mathbb{P}}(n^{-1/2})$;
(iv) $\sup_{p \in \mathcal{P}_{\text{osn}}} |(\mathbb{P}_n - P_0) S_p^2| = o_{\mathbb{P}}(n^{-1})$.
Then:

$$\sup_{\theta \in \Theta_{\text{osn}}} \left| nL_n(\theta) - \left(n\mathbb{P}_n \log p_0 - \frac{1}{2} n P_0\big((2\mathbb{T}S_{p_\theta})^2\big) + n\mathbb{P}_n(2\mathbb{T}S_{p_\theta}) \right) \right| = o_{\mathbb{P}}(1).$$

If, in addition, $\mathcal{T}$ *is linear with finite dimension* $d^* \ge 1$, *then Assumption* 4.2 *holds over* Θ_{osn} *with* $\ell_n = n\mathbb{P}_n \log p_0$, $\sqrt{n}\hat{\gamma}_n = \mathbb{V}_n = \mathbb{G}_n(\psi)$, $\Sigma = I_{d^*}$, *and* $\gamma(\theta)$ *defined in* (27).

PROOF OF LEMMA F.4: We first prove

$$\sup_{p \in \mathcal{P}_{\text{osn}}} \left| n\mathbb{P}_n \log(p/p_0) - 2n\mathbb{P}_n\big(S_p - P_0(S_p)\big) + n\big(\mathbb{P}_n S_p^2 + h^2(p, p_0)\big) \right| = o_{\mathbb{P}}(1) \tag{63}$$

by adapting arguments used in Theorem 1 of Azaïs, Gassiat, and Mercadier (2009), Theorem 3.1 in Gassiat (2002), and Theorem 2.1 in Liu and Shao (2003).

Take n large enough that $r_n/\sqrt{n} \le \varepsilon$. Then, for each $p \in \mathcal{P}_{\text{osn}} \setminus \{p_0\}$,

$$n\mathbb{P}_n \log(p/p_0) = 2n\mathbb{P}_n S_p - n\mathbb{P}_n S_p^2 + 2n\mathbb{P}_n S_p^2 r(S_p), \tag{64}$$

where $r(u) = (\log(1+u) - u - \frac{1}{2}u^2)/u^2$ and $\lim_{u \to 0} |r(u)/(\frac{1}{3}u) - 1| = 0$. By condition (ii), $\max_{1 \le i \le n} |S_p(X_i)| \le r_n/\sqrt{n} \times \max_{1 \le i \le n} D(X_i) = o_{\mathbb{P}}(r_n^{-2})$ uniformly for $p \in \mathcal{P}_{\text{osn}}$. This implies that $\sup_{p \in \mathcal{P}_{\text{osn}}} \max_{1 \le i \le n} |r(S_p(X_i))| = o_{\mathbb{P}}(r_n^{-2})$. Therefore, by the Glivenko–Cantelli condition in (ii),

$$\sup_{p \in \mathcal{P}_{\text{osn}}} \left| 2n\mathbb{P}_n S_p^2 r(S_p) \right| \le 2r_n^2 \times o_{\mathbb{P}}\big(r_n^{-2}\big) \times \sup_{p \in \mathcal{P}_{\text{osn}}} \mathbb{P}_n s_p^2 = o_{\mathbb{P}}(1) \times \big(1 + o_{\mathbb{P}}(1)\big) = o_{\mathbb{P}}(1).$$

Display (63) now follows by adding and subtracting $2nP_0(S_p) = -nh^2(p, p_0)$ to (64).

Each element of $\mathcal{T}$ has mean zero and so $P_0(\mathbb{T}S_p) = 0$ for each p. By condition (iii):

$$\sup_{p \in \mathcal{P}_{\text{osn}}} \left| \mathbb{P}_n\big(S_p - P_0(S_p) - \mathbb{T}S_p\big) \right| = n^{-1/2} \times \sup_{p \in \mathcal{P}_{\text{osn}}} \left| \mathbb{G}_n(S_p - \mathbb{T}S_p) \right| = o_{\mathbb{P}}\big(n^{-1}\big).$$

It remains to show

$$\sup_{p \in \mathcal{P}_{\text{osn}}} \left| \mathbb{P}_n\big(S_p^2\big) + h^2(p, p_0) - 2P_0\big((\mathbb{T}S_p)^2\big) \right| = o_{\mathbb{P}}\big(n^{-1}\big). \tag{65}$$

By condition (iv) and $P_0(S_p^2) = h^2(p, p_0)$, to establish (65) it is enough to show

$$\sup_{p\in\mathcal{P}_{\mathrm{osn}}} \left|P_0\big(S_p^2\big) - P_0\big((\mathbb{T}S_p)^2\big)\right| = o\big(n^{-1}\big).$$

Observe by definition of $\mathbb{T}$ and condition (i), for each $p \in \mathcal{P}$, there is a $g_p \in \mathcal{T}$ and remainder $R_p^* = R_p/\sqrt{p_0}$ such that $S_p = g_p + h(p, p_0)R_p^*$, and so

$$\|S_p - \mathbb{T}S_p\|_{L^2(P_0)} \le \|S_p - g_p\|_{L^2(P_0)} = h(p, p_0)\big\|R_p^*\big\|_{L^2(P_0)} = h(p, p_0)\|R_p\|_{L^2(\lambda)}. \tag{66}$$

By Moreau's decomposition theorem and inequality (66), we may deduce

$$\sup_{p\in\mathcal{P}_{\mathrm{osn}}} \left|P_0\big(S_p^2\big) - P_0\big((\mathbb{T}S_p)^2\big)\right| = \sup_{p\in\mathcal{P}_{\mathrm{osn}}} \|S_p - \mathbb{T}S_p\|^2_{L^2(P_0)} \le \sup_{p\in\mathcal{P}_{\mathrm{osn}}} h(p, p_0)^2\|R_p\|^2_{L^2(\lambda)},$$

which is $o(n^{-1})$ by condition (i) and definition of $\mathcal{P}_{\mathrm{osn}}$. This proves the first result.

The second result is immediate by defining $\mathbb{V}_n = \mathbb{G}_n(\psi)$ with $\psi = (\psi_1, \dots, \psi_{d^*})'$ where $\psi_1, \dots, \psi_{d^*}$ is an orthonormal basis for Span$(\mathcal{T})$, and $\gamma(\theta)$ as in (27), then noting that $P_0((\mathbb{T}(2S_{p_\theta}))^2) = \gamma(\theta)'P_0(\psi\psi')\gamma(\theta) = \|\gamma(\theta)\|^2$. *Q.E.D.*

PROOF OF PROPOSITION 5.2: We verify the conditions of Lemma F.4. By DQM (condition (b)), we have $\sup\{\|R_p\|_{L^2(\lambda)} : h(p, p_0) \le n^{-1/4}\} \to 0$ as $n \to \infty$. Therefore, we may choose a sequence $(a_n)_{n\in\mathbb{N}}$ with $a_n \le n^{1/4}$ but $a_n \to \infty$ slowly enough that

$$\sup\big\{a_n\|R_p\|_{L^2(\lambda)} : h(p, p_0) \le a_n/\sqrt{n}\big\} \to 0 \quad \text{as } n \to \infty$$

and hence $\sup\{r_n\|R_p\|_{L^2(\lambda)} : h(p, p_0) \le r_n/\sqrt{n}\} \to 0$ as $n \to \infty$ for any slowly diverging positive sequence $(r_n)_{n\in\mathbb{N}}$ with $r_n \le a_n$. This verifies condition (i) of Lemma F.4.

For condition (ii), $\overline{\mathcal{D}}_\varepsilon^2$ is Glivenko–Cantelli by condition (c) and Lemma 2.10.14 of van der Vaart and Wellner (1996). Moreover, it follows from the envelope condition (in condition (c)) that $\max_{1\le i\le n} D(X_i) = o_{\mathbb{P}}(n^{1/2})$. We can therefore choose a positive sequence $(c_n)_{n\in\mathbb{N}}$ with $c_n \to \infty$ such that $c_n^3 \max_{1\le i\le n} D(X_i) = o_{\mathbb{P}}(n^{1/2})$ and so $\max_{1\le i\le n} D(X_i) = o_{\mathbb{P}}(n^{1/2}/r_n^3)$ for any $0 < r_n \le c_n$.

For condition (iv), as $\overline{\mathcal{D}}_\varepsilon^2$ is Glivenko–Cantelli, we may choose a positive sequence $(b_n)_{n\in\mathbb{N}}$ with $b_n \to \infty$ such that $b_n^2 \sup_{s_p\in\overline{\mathcal{D}}_\varepsilon} |(\mathbb{P}_n - P_0)s_p^2| = o_{\mathbb{P}}(1)$. Therefore, for any $0 < r_n \le b_n$, we have

$$\sup_{p:h(p,p_0)\le r_n/\sqrt{n}} \left|(\mathbb{P}_n - P_0)S_p^2\right| \le \sup_{p:h(p,p_0)\le r_n/\sqrt{n}} r_n^2\left|(\mathbb{P}_n - P_0)s_p^2\right|/n = o_{\mathbb{P}}\big(n^{-1}\big).$$

Finally, for condition (iii), note that conditions (b) and (c) imply that $\overline{\mathcal{D}}_\varepsilon^o := \{s_p - \mathbb{T}s_p : s_p \in \overline{\mathcal{D}}_\varepsilon\}$ is Donsker. Also note that the singleton $\{0\}$ is the only limit point of $\overline{\mathcal{D}}_\varepsilon^o$ as $\varepsilon \searrow 0$ because

$$\sup\big\{\|s_p - \mathbb{T}s_p\|_{L^2(P_0)} : h(p, p_0) \le \varepsilon\big\} \le \sup\big\{\|R_p\|_{L^2(\lambda)} : h(p, p_0) \le \varepsilon\big\} \to 0 \quad (\text{as } \varepsilon \to 0)$$

by DQM (condition (b)). Asymptotic equicontinuity of $\mathbb{G}_n$ on $\overline{\mathcal{D}}_\varepsilon^o$ then implies that

$$\sup_{p:h(p,p_0)\le n^{-1/4}} \left|\mathbb{G}_n(s_p - \mathbb{T}s_p)\right| = o_{\mathbb{P}}(1).$$

We can therefore choose a positive sequence $(d_n)_{n\in\mathbb{N}}$ with $d_n \le n^{1/4}$ but $d_n \to \infty$ slowly enough that $d_n \sup_{p:h(p,p_0)\le n^{-1/4}} |\mathbb{G}_n(s_p - \mathbb{T}s_p)| = o_{\mathbb{P}}(1)$ and so, for any $0 < r_n \le d_n$,

$$\sup_{p:h(p,p_0)\le r_n/\sqrt{n}} \left|\mathbb{G}_n(S_p - \mathbb{T}S_p)\right| \le \frac{r_n}{\sqrt{n}} \sup_{p:h(p,p_0)\le n^{-1/4}} \mathbb{G}_n(s_p - \mathbb{T}s_p) = o_{\mathbb{P}}\big(n^{-1/2}\big).$$

The result follows by taking $r_n = (a_n \wedge b_n \wedge c_n \wedge d_n)$. *Q.E.D.*

PROOF OF PROPOSITION 5.3: We first show that

$$\sup_{\theta:\|g(\theta)\|\le r_n/\sqrt{n}} \left| nL_n(\theta) - \left(-\frac{1}{2}\big(\mathbb{T}\big(\sqrt{n}g(\theta)\big) + Z_n\big)'\Omega^{-1}\big(\mathbb{T}\big(\sqrt{n}g(\theta)\big) + Z_n\big)\right)\right| = o_{\mathbb{P}}(1) \tag{67}$$

holds for a positive sequence $(r_n)_{n\in\mathbb{N}}$ with $r_n \to \infty$, $r_n = o(n^{1/4})$ and $Z_n = \mathbb{G}_n(\rho_{\theta^*})$. Take n large enough that $n^{-1/4} \le \varepsilon_0$. By conditions (a)–(c) and Lemma 2.10.14 of van der Vaart and Wellner (1996), we have that $\sup_{\theta:\|g(\theta)\|\le n^{-1/4}} \|\mathbb{P}_n(\rho_\theta\rho_\theta') - \Omega\| = o_{\mathbb{P}}(1)$. Therefore, we may choose a positive sequence $(a_n)_{n\in\mathbb{N}}$ with $a_n \to \infty$, $a_n = o(n^{1/4})$ such that $\sup_{\theta:\|g(\theta)\|\le n^{-1/4}} a_n^2\|\mathbb{P}_n(\rho_\theta\rho_\theta') - \Omega\| = o_{\mathbb{P}}(1)$ and hence

$$\sup_{\theta:\|g(\theta)\|\le r_n/\sqrt{n}} \big\|\mathbb{P}_n\big(\rho_\theta\rho_\theta'\big) - \Omega\big\| = o_{\mathbb{P}}\big(r_n^{-2}\big) \tag{68}$$

for any $0 < r_n \le a_n$.

Notice that $Z_n \rightsquigarrow N(0, \Omega)$ by condition (a) and that the covariance of each element of $\rho_\theta(X_i) - \rho_{\theta^*}(X_i)$ vanishes uniformly over Θ_I^ε as $\varepsilon \to 0$ by condition (c). Asymptotic equicontinuity of $\mathbb{G}_n$ (which holds under (a)) then implies that $\sup_{\theta:\|g(\theta)\|\le n^{-1/4}} \|\mathbb{G}_n(\rho_\theta) - Z_n\| = o_{\mathbb{P}}(1)$. We can therefore choose a positive sequence $(b_n)_{n\in\mathbb{N}}$ with $b_n \to \infty$, $b_n = o(n^{1/4})$ as $n \to \infty$ such that $b_n \sup_{\theta:\|g(\theta)\|\le b_n/\sqrt{n}} \|\mathbb{G}_n(\rho_\theta) - Z_n\| = o_{\mathbb{P}}(1)$ and hence

$$\sup_{\theta:\|g(\theta)\|\le r_n/\sqrt{n}} \big\|\sqrt{n}\mathbb{P}_n\rho_\theta - \big(\sqrt{n}g(\theta) + Z_n\big)\big\| = o_{\mathbb{P}}\big(r_n^{-1}\big) \tag{69}$$

for any $0 < r_n \le b_n$.

Condition (d) implies that we may choose a sequence $(c_n)_{n\in\mathbb{N}}$ with $c_n \to \infty$, $c_n = o(n^{1/4})$ such that $\sup_{\theta:\|g(\theta)\|\le c_n/\sqrt{n}} \sqrt{n}\|g(\theta) - \mathbb{T}g(\theta)\| = o(c_n^{-1})$ and so

$$\sup_{\theta:\|g(\theta)\|\le r_n/\sqrt{n}} \big\|\sqrt{n}g(\theta) - \mathbb{T}\big(\sqrt{n}g(\theta)\big)\big\| = o\big(r_n^{-1}\big) \tag{70}$$

for any $0 < r_n \le c_n$.

Result (67) now follows by taking $r_n = (a_n \wedge b_n \wedge c_n)$ and using (68), (69), and (70). To complete the proof, expanding the quadratic in (67), we obtain

$$\begin{aligned}
-\frac{1}{2}\big(\mathbb{T}\big(\sqrt{n}g(\theta)\big) + Z_n\big)'\Omega^{-1}\big(\mathbb{T}\big(\sqrt{n}g(\theta)\big) + Z_n\big) &= -\frac{1}{2}Z_n'\Omega^{-1}Z_n - \frac{1}{2}\big\|\big[\Omega^{-1/2}\mathbb{T}\big(\sqrt{n}g(\theta)\big)\big]_1\big\|^2 \\
&\quad - \big[\Omega^{-1/2}Z_n\big]_1'\big[\Omega^{-1/2}\mathbb{T}\big(\sqrt{n}g(\theta)\big)\big]_1,
\end{aligned}$$

and the result follows with $\ell_n = Z_n'\Omega^{-1}Z_n$, $\gamma(\theta) = [\Omega^{-1/2}\mathbb{T}g(\theta)]_1$, and $\mathbb{V}_n = -[\Omega^{-1/2}Z_n]_1$. *Q.E.D.*

PROOF OF PROPOSITION 5.4: Follows by similar arguments to the proof of Proposition 5.3, noting that, by condition (e), we may choose a positive sequence $(a_n)_{n\in\mathbb{N}}$ with $a_n \to \infty$ slowly such that $a_n^2\|\widehat{W} - \Omega^{-1}\| = o_{\mathbb{P}}(1)$. Therefore, $\|\widehat{W} - \Omega^{-1}\| = o_{\mathbb{P}}(r_n^{-2})$ holds for any $0 < r_n \le a_n$. *Q.E.D.*

LEMMA F.5: *Consider the missing data model with a flat prior on Θ. Suppose that the model is point-identified* (*i.e., the true* $\eta_2 = 1$). *Then Assumption* 4.1(ii) *holds for*

$$\Theta_{\text{osn}} = \left\{\theta : \left|\tilde{\gamma}_{11}(\theta) - \tilde{\gamma}_{11}\right| \le r_n/\sqrt{n},\, \tilde{\gamma}_{00}(\theta) \le r_n/n\right\}$$

for any positive sequence $(r_n)_{n\in\mathbb{N}}$ *with* $r_n \to \infty$, $r_n/\sqrt{n} = o(1)$.

PROOF OF LEMMA F.5: The flat prior on Θ induces a flat prior on $\{(a, b) \in [0, 1] : 0 \le a \le 1 - b\}$ under the map $\theta \mapsto (\tilde{\gamma}_{11}(\theta), \tilde{\gamma}_{00}(\theta))$. Take n large enough that $[\tilde{\gamma}_{11} - r_n/\sqrt{n}, \tilde{\gamma}_{11} + r_n/\sqrt{n}] \subseteq [0, 1]$ and $r_n/n < 1$. Then, with $S_n := \sum_{i=1}^n Y_i$, we have

$$\begin{aligned}
\Pi_n\big(\Theta_{\text{osn}}^c|\mathbf{X}_n\big) &= \frac{\int_{[0,\tilde{\gamma}_{11}-r_n/\sqrt{n}]\cup[\tilde{\gamma}_{11}+r_n/\sqrt{n},1]}\int_0^{1-a}(a)^{S_n}(1-a-b)^{n-S_n}\,\mathrm{d}b\,\mathrm{d}a}{\int_0^1\int_0^{1-a}(a)^{S_n}(1-a-b)^{n-S_n}\,\mathrm{d}b\,\mathrm{d}a} \\
&\quad + \frac{\int_{\tilde{\gamma}_{11}-r_n/\sqrt{n}}^{\tilde{\gamma}_{11}+r_n/\sqrt{n}}\int_{r_n/n}^{1-a}(a)^{S_n}(1-a-b)^{n-S_n}\,\mathrm{d}b\,\mathrm{d}a}{\int_0^1\int_0^{1-a}(a)^{S_n}(1-a-b)^{n-S_n}\,\mathrm{d}b\,\mathrm{d}a} \\
&=: I_1 + I_2.
\end{aligned}$$

Integrating I_1 first with respect to b yields

$$I_1 = \frac{\int_{[0,\tilde{\gamma}_{11}-r_n/\sqrt{n}]\cup[\tilde{\gamma}_{11}+r_n/\sqrt{n},1]}(a)^{S_n}(1-a)^{n-S_n+1}\,\mathrm{d}a}{\int_0^1(a)^{S_n}(1-a)^{n-S_n+1}\,\mathrm{d}a} = \mathbb{P}_{U|S_n}\big(|U - \tilde{\gamma}_{11}| > r_n/\sqrt{n}\big),$$

where $U|S_n \sim \text{Beta}(S_n + 1, n - S_n + 2)$. Note that this implies

$$\mathbb{E}[U|S_n] = \frac{S_n + 1}{n + 3}, \qquad \text{Var}[U|S_n] = \frac{(S_n + 1)(n - S_n + 2)}{(n + 3)^2(n + 4)}.$$

By the triangle inequality, the fact that $\mathbb{E}[U|S_n] = \tilde{\gamma}_{11} + O_{\mathbb{P}}(n^{-1/2})$, and Chebyshev's inequality,

$$\begin{aligned}
I_1 &\le \mathbb{P}_{U|S_n}\big(\big|U - \mathbb{E}[U|S_n]\big| > r_n/(2\sqrt{n})\big) + \mathbb{1}\big\{\big|\mathbb{E}[U|S_n] - \tilde{\gamma}_{11}\big| > r_n/(2\sqrt{n})\big\} \\
&= \mathbb{P}_{U|S_n}\big(\big|U - \mathbb{E}[U|S_n]\big| > r_n/(2\sqrt{n})\big) + o_{\mathbb{P}}(1) \\
&\le \frac{4}{r_n^2}\frac{\left(\frac{S_n}{n} + \frac{1}{n}\right)\left(1 - \frac{S_n}{n} + \frac{2}{n}\right)}{\left(1 + \frac{3}{n}\right)^2\left(1 + \frac{4}{n}\right)} + o_{\mathbb{P}}(1) = o_{\mathbb{P}}(1).
\end{aligned}$$

Similarly, for n sufficiently large,

$$I_2 = \frac{\int_{\tilde{\gamma}_{11}-r_n/\sqrt{n}}^{\tilde{\gamma}_{11}+r_n/\sqrt{n}} (a)^{S_n}\big(1-a-(r_n/n)\big)^{n-S_n+1}\,\mathrm{d}a}{\int_0^1 (a)^{S_n}(1-a)^{n-S_n+1}\,\mathrm{d}a} \le \frac{\int_0^{1-r_n/n} (a)^{S_n}\big(1-a-(r_n/n)\big)^{n-S_n+1}\,\mathrm{d}a}{\int_0^1 (a)^{S_n}(1-a)^{n-S_n+1}\,\mathrm{d}a}.$$

Using the change of variables $a \mapsto c(a) := \frac{1-a-r_n/n}{1-r_n/n}$ in the numerator yields

$$I_2 \le \big(1-(r_n/n)\big)^{n+2} \frac{\int_0^1 (1-c)^{S_n}(c)^{n-S_n+1}\,\mathrm{d}c}{\int_0^1 (a)^{S_n}(1-a)^{n-S_n+1}\,\mathrm{d}a} = \big(1-(r_n/n)\big)^{n+2} \to 0.$$

Therefore, $\Pi_n(\Theta^c_{\mathrm{osn}}|\mathbf{X}_n) = o_{\mathbb{P}}(1)$, as required. *Q.E.D.*

F.3. *Proofs for Appendix B*

PROOF OF THEOREM B.1: We first derive the asymptotic distribution of $\sup_{\theta\in\Theta_I} Q_n(\theta)$ under $\mathrm{P}_{n,a}$. By similar arguments to the proof of Theorem 4.1, we have

$$\sup_{\theta\in\Theta_I} Q_n(\theta) = \|\mathbb{V}_n\|^2 + o_{\mathrm{P}_{n,a}}(1) \overset{\mathrm{P}_{n,a}}{\rightsquigarrow} \chi^2_{d^*}\big(a'a\big).$$

Identical arguments to the proof of Lemma 4.1 yield

$$\sup_z \big|\Pi_n\big(\{\theta : Q_n(\theta) \le z\}|\mathbf{X}_n\big) - F_{\chi^2_{d^*}}(z)\big| = o_{\mathrm{P}_{n,a}}(1).$$

Therefore, $\xi^{\mathrm{mc}}_{n,\alpha} = \chi^2_{d^*,\alpha} + o_{\mathrm{P}_{n,a}}(1)$ and we obtain

$$\mathrm{P}_{n,a}(\Theta_I \subseteq \widehat{\Theta}_\alpha) = \Pr\big(\chi^2_{d^*}\big(a'a\big) \le \chi^2_{d^*,\alpha}\big) + o(1),$$

as required. *Q.E.D.*

PROOF OF THEOREM B.2: By similar arguments to the proof of Theorem 4.3, we have

$$PQ_n(M_I) = f(\mathbb{V}_n) + o_{\mathrm{P}_{n,a}}(1) \overset{\mathrm{P}_{n,a}}{\rightsquigarrow} f(Z+a),$$

where $Z \sim N(0, I_{d^*})$. Identical arguments to the proof of Lemma 4.3 yield

$$\sup_{z\in I} \big|\Pi_n\big(\{\theta : PQ_n\big(M(\theta)\big) \le z\}|\mathbf{X}_n\big) - \mathbb{P}_{Z|\mathbf{X}_n}\big(f(Z) \le z\big)\big| = o_{\mathrm{P}_{n,a}}(1)$$

for a neighborhood I of z_α. Therefore, $\xi^{\mathrm{mc},p}_{n,\alpha} = z_\alpha + o_{\mathrm{P}_{n,a}}(1)$ and we obtain

$$\mathrm{P}_{n,a}(M_I \subseteq \widehat{M}_\alpha) = \mathbb{P}_Z\big(f(Z+a) \le z_\alpha\big) + o(1),$$

as required. *Q.E.D.*

F.4. *Proofs for Appendix C*

PROOF OF LEMMA C.1: By equations (45) and (46) in the proof of Lemma 4.1, it suffices to characterize the large-sample behavior of

$$R_n(z) := \frac{\int_{\{\theta : Q_n(\theta) \le z\} \cap \Theta_{\mathrm{osn}}} e^{-\frac{1}{2}Q_n(\theta)} \, \mathrm{d}\Pi(\theta)}{\int_{\Theta_{\mathrm{osn}}} e^{-\frac{1}{2}Q_n(\theta)} \, \mathrm{d}\Pi(\theta)}.$$

By Assumption C.2(i), there exists a positive sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ with $\varepsilon_n = o(1)$ such that $(1-\varepsilon_n)h(\gamma(\theta)-\hat{\gamma}_n) \le \frac{a_n}{2}Q_n(\theta) \le (1+\varepsilon_n)h(\gamma(\theta)-\hat{\gamma}_n)$ holds uniformly over Θ_{osn} wpa1. Therefore,

$$\begin{aligned}
&\frac{\int_{\{\theta : 2a_n^{-1}(1+\varepsilon_n)h(\gamma(\theta)-\hat{\gamma}_n) \le z\} \cap \Theta_{\mathrm{osn}}} e^{-a_n^{-1}(1+\varepsilon_n)h(\gamma(\theta)-\hat{\gamma}_n)} \, \mathrm{d}\Pi(\theta)}{\int_{\Theta_{\mathrm{osn}}} e^{-a_n^{-1}(1-\varepsilon_n)h(\gamma(\theta)-\hat{\gamma}_n)} \, \mathrm{d}\Pi(\theta)} \\
&\quad \le R_n(z) \le \frac{\int_{\{\theta : 2a_n^{-1}(1-\varepsilon_n)h(\gamma(\theta)-\hat{\gamma}_n) \le z\} \cap \Theta_{\mathrm{osn}}} e^{-a_n^{-1}(1-\varepsilon_n)h(\gamma(\theta)-\hat{\gamma}_n)} \, \mathrm{d}\Pi(\theta)}{\int_{\Theta_{\mathrm{osn}}} e^{-a_n^{-1}(1+\varepsilon_n)h(\gamma(\theta)-\hat{\gamma}_n)} \, \mathrm{d}\Pi(\theta)}.
\end{aligned}$$

By similar arguments to the proof of Lemma 4.1, under Assumption 4.3 there exists a positive sequence $(\bar{\varepsilon}_n)_{n\in\mathbb{N}}$ with $\bar{\varepsilon}_n = o(1)$ such that, for all n sufficiently large, we have

$$\begin{aligned}
&(1-\bar{\varepsilon}_n)\frac{\int_{\{\gamma : 2a_n^{-1}(1+\varepsilon_n)h(\gamma-\hat{\gamma}_n) \le z\} \cap \Gamma_{\mathrm{osn}}} e^{-a_n^{-1}(1+\varepsilon_n)h(\gamma-\hat{\gamma}_n)} \, \mathrm{d}\gamma}{\int_{\Gamma_{\mathrm{osn}}} e^{-a_n^{-1}(1-\varepsilon_n)h(\gamma-\hat{\gamma}_n)} \, \mathrm{d}\gamma} \\
&\quad \le R_n(z) \le (1+\bar{\varepsilon}_n)\frac{\int_{\{\gamma : 2a_n^{-1}(1-\varepsilon_n)h(\gamma-\hat{\gamma}_n) \le z\} \cap \Gamma_{\mathrm{osn}}} e^{-a_n^{-1}(1-\varepsilon_n)h(\gamma-\hat{\gamma}_n)} \, \mathrm{d}\gamma}{\int_{\Gamma_{\mathrm{osn}}} e^{-a_n^{-1}(1+\varepsilon_n)h(\gamma-\hat{\gamma}_n)} \, \mathrm{d}\gamma}
\end{aligned}$$

under the change of variables $\theta \mapsto \gamma(\theta)$, where $\Gamma_{\mathrm{osn}} = \{\gamma(\theta) : \theta \in \Theta_{\mathrm{osn}}\}$.

Assumption C.2(ii) implies that

$$a_n^{-1}(1\pm\varepsilon_n)h(\gamma-\hat{\gamma}_n) = h\big(a_n^{-r_1}(1\pm\varepsilon_n)^{r_1}(\gamma_1-\hat{\gamma}_{n,1}), \dots, a_n^{-r_{d^*}}(1\pm\varepsilon_n)^{r_{d^*}}(\gamma_{d^*}-\hat{\gamma}_{n,d^*})\big).$$

Using a change of variables:

$$\gamma \mapsto \kappa_{\pm}(\gamma) = \big(a_n^{-r_1}(1\pm\varepsilon_n)^{r_1}(\gamma_1-\hat{\gamma}_{n,1}), \dots, a_n^{-r_{d^*}}(1\pm\varepsilon_n)^{r_{d^*}}(\gamma_{d^*}-\hat{\gamma}_{n,d^*})\big)$$

(with choice of sign as appropriate) and setting $r^* = r_1 + \cdots + r_{d^*}$, we obtain

$$\begin{aligned}
&(1-\bar{\varepsilon}_n)\frac{(1-\varepsilon_n)^{r^*}}{(1+\varepsilon_n)^{r^*}}\frac{\int_{\{\kappa:2h(\kappa)\leq z\}\cap K_{\mathrm{osn}}^+} e^{-h(\kappa)}\,\mathrm{d}\kappa}{\int e^{-h(\kappa)}\,\mathrm{d}\kappa} \\
&\qquad \leq R_n(z) \leq (1+\bar{\varepsilon}_n)\frac{(1+\varepsilon_n)^{r^*}}{(1-\varepsilon_n)^{r^*}}\frac{\int_{\{\kappa:2h(\kappa)\leq z\}} e^{-h(\kappa)}\,\mathrm{d}\kappa}{\int_{K_{\mathrm{osn}}^+} e^{-h(\kappa)}\,\mathrm{d}\kappa}
\end{aligned} \tag{71}$$

uniformly in z, where $K_{\mathrm{osn}}^+ = \{\kappa_+(\gamma) : \gamma \in \Gamma_{\mathrm{osn}}\}$.

We can use a change of variables for $\kappa \mapsto t = 2h(\kappa)$ to obtain

$$\begin{aligned}
\int_{\{\kappa:h(\kappa)\leq z/2\}} e^{-h(\kappa)}\,\mathrm{d}\kappa &= 2^{-r^*}\mathrm{V}(S)\int_0^z e^{-t/2}t^{r^*-1}\,dt, \\
\int e^{-h(\kappa)}\,\mathrm{d}\kappa &= 2^{-r^*}\mathrm{V}(S)\int_0^\infty e^{-t/2}t^{r^*-1}\,dt,
\end{aligned} \tag{72}$$

where $\mathrm{V}(S)$ denotes the volume of the set $S = \{\kappa : h(\kappa) = 1\}$.

For the remaining integrals over K_{osn}^+, we first fix any $\omega \in \Omega$ so that $K_{\mathrm{osn}}^+(\omega)$ becomes a deterministic sequence of sets. Let $C_n(\omega) = K_{\mathrm{osn}}^+(\omega) \cap B_{k_n}$. Assumption C.2(iii) gives $\mathbb{R}_+^{d^*} = \overline{\bigcup_{n\geq 1} C_n(\omega)}$ for almost every ω. Now, clearly,

$$\int e^{-h(\kappa)}\,\mathrm{d}\kappa \geq \int_{K_{\mathrm{osn}}^+(\omega)} e^{-h(\kappa)}\,\mathrm{d}\kappa \geq \int \mathbb{1}\big\{\kappa \in C_n(\omega)\big\}e^{-h(\kappa)}\,\mathrm{d}\kappa \to \int e^{-h(\kappa)}\,\mathrm{d}\kappa$$

(by dominated convergence) for almost every ω. Therefore,

$$\int_{K_{\mathrm{osn}}^+} e^{-h(\kappa)}\,\mathrm{d}\kappa \to_p 2^{-r^*}\mathrm{V}(S)\int_0^\infty e^{-t/2}t^{r^*-1}\,dt. \tag{73}$$

We may similarly deduce that

$$\sup_z\left|\int_{\{\kappa:h(\kappa)\leq 2z\}\cap K_{\mathrm{osn}}^+} e^{-h(\kappa)}\,\mathrm{d}\kappa - 2^{-r^*}\mathrm{V}(S)\int_0^z e^{-t/2}t^{r^*-1}\,dt\right| = o_{\mathbb{P}}(1). \tag{74}$$

The result follows by substituting (72), (73), and (74) into (71). *Q.E.D.*

PROOF OF THEOREM C.1: We verify the conditions of Lemma 2.1. Lemma C.1 shows that the posterior distribution of the QLR is asymptotically $F_\Gamma = \Gamma(r^*, 1/2)$, and hence $\xi_{n,\alpha}^{\mathrm{post}} = z_\alpha + o_{\mathbb{P}}(1)$, where z_α denotes the α quantile of F_Γ. By assumption, $\sup_{\theta\in\Theta_I} Q_n(\theta) \rightsquigarrow F_\Gamma$. Then,

$$\xi_{n,\alpha}^{\mathrm{mc}} = z_\alpha + \big(\xi_{n,\alpha}^{\mathrm{post}} - z_\alpha\big) + \big(\xi_{n,\alpha}^{\mathrm{mc}} - \xi_{n,\alpha}^{\mathrm{post}}\big) = z_\alpha + o_{\mathbb{P}}(1),$$

where the final equality is by Assumption 4.4. *Q.E.D.*

F.5. *Proofs and Additional Lemmas for Appendix D*

PROOF OF LEMMA D.1: By condition (i), there exists a positive sequence $(\varepsilon_n)_{n\in\mathbb{N}}$, $\varepsilon_n = o(1)$ such that $\sup_{\mathbb{P}\in\mathbf{P}} \mathbb{P}(\sup_{\theta\in\Theta_I(\mathbb{P})} Q_n(\theta) - W_n > \varepsilon_n) = o(1)$. Let $\mathcal{A}_{n,\mathbb{P}}$ denote the event on which $\sup_{\theta\in\Theta_I(\mathbb{P})} Q_n(\theta) - W_n \le \varepsilon_n$. Then,

$$\begin{aligned}\inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(\Theta_I(\mathbb{P}) \subseteq \widehat{\Theta}_\alpha\big) &\ge \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(\{\Theta_I(\mathbb{P}) \subseteq \widehat{\Theta}_\alpha\} \cap \mathcal{A}_{n,\mathbb{P}}\big)\\ &= \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\Big(\Big\{\sup_{\theta\in\Theta_I(\mathbb{P})} Q_n(\theta) \le v_{\alpha,n}\Big\} \cap \mathcal{A}_{n,\mathbb{P}}\Big)\\ &\ge \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(\{W_n \le v_{\alpha,n} - \varepsilon_n\} \cap \mathcal{A}_{n,\mathbb{P}}\big),\end{aligned}$$

where the second equality is by the definition of $\widehat{\Theta}_\alpha$. As $\mathbb{P}(A \cap B) \ge 1 - \mathbb{P}(A^c) - \mathbb{P}(B^c)$, we therefore have

$$\begin{aligned}\inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(\Theta_I(\mathbb{P}) \subseteq \widehat{\Theta}_\alpha\big) &\ge 1 - \sup_{\mathbb{P}\in\mathbf{P}} \mathbb{P}(W_n > v_{\alpha,n} - \varepsilon_n) - \sup_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(\mathcal{A}_{n,\mathbb{P}}^c\big)\\ &= 1 - \Big(1 - \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}(W_n \le v_{\alpha,n} - \varepsilon_n)\Big) - o(1)\\ &= \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}(W_n \le v_{\alpha,n} - \varepsilon_n) - o(1)\\ &\ge \alpha - o(1),\end{aligned}$$

where the final line is by condition (ii) and definition of $\mathcal{A}_{n,\mathbb{P}}$. *Q.E.D.*

PROOF OF LEMMA D.2: Follows by similar arguments to the proof of Lemma D.1. *Q.E.D.*

We use the next lemma several times in the following proofs.

LEMMA F.6: *Let $T \subseteq \mathbb{R}^d$ be a closed convex cone and let $\mathbf{T}$ denote the projection onto T. Then,*

$$\big\|\mathbf{T}(x + t) - t\big\| \le \|x\|$$

for any $x \in \mathbb{R}^d$ and $t \in T$.

PROOF OF LEMMA F.6: Let $\mathbf{T}^o$ denote the projection onto the polar cone T^o of T. As $u't \le 0$ holds for any $u \in T^o$ and $\|\mathbf{T}v\| \le \|v\|$ holds for any $v \in \mathbb{R}^d$, we obtain

$$\big\|\mathbf{T}(x + t)\big\|^2 + 2\big(\mathbf{T}^o(x + t)\big)'t \le \big\|\mathbf{T}(x + t)\big\|^2 \le \|x + t\|^2.$$

Subtracting $2(x + t)'t$ from both sides and using the fact that $v = \mathbf{T}v + \mathbf{T}^o v$ yields:

$$\big\|\mathbf{T}(x + t)\big\|^2 - 2\big(\mathbf{T}(x + t)\big)'t \le \|x + t\|^2 - 2(x + t)'t.$$

Adding $\|t\|^2$ to both sides and completing the square gives $\|\mathbf{T}(x+t) - t\|^2 \le \|x + t - t\|^2 = \|x\|^2$. *Q.E.D.*

In view of Lemma F.6 and Assumption D.2(i), for each $\mathbb{P} \in \mathbf{P}$, we have

$$\left\| \sqrt{n}(\hat{\gamma}_n - \tau) \right\| \le \|\mathbb{V}_n\|. \tag{75}$$

LEMMA F.7: *Let Assumptions D.*1(i) *and D.*2 *hold*. *Then*,

$$\sup_{\theta \in \Theta_{\text{osn}}} \left| Q_n(\theta) - \left\| \sqrt{n}\gamma(\theta) - \sqrt{n}(\hat{\gamma}_n - \tau) \right\|^2 - 2f_{n,\perp}\big(\gamma_\perp(\theta)\big) \right| = o_{\mathbb{P}}(1) \tag{76}$$

uniformly in $\mathbb{P}$.

If, *in addition*, *Assumption D.*5(i) *holds*, *then*

$$\sup_{\theta \in \Theta_{\text{osn}}} \left| PQ_n\big(M(\theta)\big) - f\big(\sqrt{n}(\hat{\gamma}_n - \tau) - \sqrt{n}\gamma(\theta)\big) \right| = o_{\mathbb{P}}(1) \tag{77}$$

uniformly in $\mathbb{P}$.

PROOF OF LEMMA F.7: To show (76), by Assumptions D.1(i) and D.2(i), (iii),

$$\begin{aligned} nL_n(\hat{\theta}) &= \sup_{\theta \in \Theta_{\text{osn}}} \left(\ell_n + \frac{n}{2}\|\hat{\gamma}_n - \tau\|^2 - \frac{1}{2}\left\| \sqrt{n}\gamma(\theta) - \sqrt{n}(\hat{\gamma}_n - \tau) \right\|^2 - f_{n,\perp}\big(\gamma_\perp(\theta)\big) \right) + o_{\mathbb{P}}(1) \\ &= \ell_n + \frac{n}{2}\|\hat{\gamma}_n - \tau\|^2 - \inf_{\theta \in \Theta_{\text{osn}}} \frac{1}{2}\left\| \sqrt{n}\gamma(\theta) - \sqrt{n}(\hat{\gamma}_n - \tau) \right\|^2 + o_{\mathbb{P}}(1) \end{aligned} \tag{78}$$

uniformly in $\mathbb{P}$. But observe that by Assumption D.2(i), (ii), for any $\epsilon > 0$,

$$\begin{aligned} &\sup_{\mathbb{P} \in \mathbf{P}} \mathbb{P}\Big(\inf_{\theta \in \Theta_{\text{osn}}} \left\| \sqrt{n}\gamma(\theta) - \sqrt{n}(\hat{\gamma}_n - \tau) \right\|^2 > \epsilon \Big) \\ &\quad \le \sup_{\mathbb{P} \in \mathbf{P}} \mathbb{P}\left(\left\{ \inf_{t \in (T - \sqrt{n}\tau) \cap B_{k_n}} \left\| t - \sqrt{n}(\hat{\gamma}_n - \tau) \right\|^2 > \epsilon \right\} \cap \left\{ \|\hat{\gamma}_n - \tau\| < \frac{k_n}{\sqrt{n}} \right\} \right) \\ &\quad\quad + \sup_{\mathbb{P} \in \mathbf{P}} \mathbb{P}\left(\|\hat{\gamma}_n - \tau\| \ge \frac{k_n}{\sqrt{n}} \right), \end{aligned}$$

where $\inf_{t \in (T - \sqrt{n}\tau) \cap B_{k_n}} \|t - \sqrt{n}(\hat{\gamma}_n - \tau)\|^2 = 0$ whenever $\|\sqrt{n}(\hat{\gamma}_n - \tau)\| < k_n$ (because $\sqrt{n}\hat{\gamma}_n \in T$). Notice $\|\sqrt{n}(\hat{\gamma}_n - \tau)\| = o_{\mathbb{P}}(k_n)$ uniformly in $\mathbb{P}$ by (75) and the condition $\|\mathbb{V}_n\| = O_{\mathbb{P}}(1)$ (uniformly in $\mathbb{P}$). This proves (76). Result (77) follows by Assumption D.5(i). *Q.E.D.*

PROOF OF LEMMA D.3: We only prove the case with singularity; the case without singularity follows similarly. By identical arguments to the proof of Lemma 4.2, it is enough to characterize the large-sample behavior of $R_n(z)$ defined in equation (47) uniformly in $\mathbb{P}$. By Lemma F.7 and Assumption D.2(i)–(iii), there exist a positive sequence $(\varepsilon_n)_{n \in \mathbb{N}}$ independent of z with $\varepsilon_n = o(1)$ and a sequence of events $(\mathcal{A}_n)_{n \in \mathbb{N}}$ (possibly depending on $\mathbb{P}$) with $\inf_{\mathbb{P} \in \mathbf{P}} \mathbb{P}(\mathcal{A}_n) = 1 - o(1)$ such that

$$\sup_{\theta \in \Theta_{\text{osn}}} \left| Q_n(\theta) - \left(\left\| \sqrt{n}\gamma(\theta) - \sqrt{n}(\hat{\gamma}_n - \tau) \right\|^2 + 2f_{n,\perp}\big(\gamma_\perp(\theta)\big) \right) \right| \le \varepsilon_n,$$

$$\sup_{\theta \in \Theta_{\text{osn}}} \left| nL_n(\theta) - \ell_n - \frac{n}{2}\|\hat{\gamma}_n - \tau\|^2 + \frac{1}{2}\left\| \sqrt{n}\gamma(\theta) - \sqrt{n}(\hat{\gamma}_n - \tau) \right\|^2 + f_{n,\perp}\big(\gamma_\perp(\theta)\big) \right| \le \varepsilon_n$$

both hold on $\mathcal{A}_n$ for all $\mathbb{P} \in \mathbf{P}$. Also note that for any $z \in \mathbb{R}$ and any singular $\mathbb{P} \in \mathbf{P}$, we have

$$\begin{aligned}&\left\{\theta \in \Theta_{\text{osn}} : \left\|\sqrt{n}\gamma(\theta) - \sqrt{n}(\hat{\gamma}_n - \tau)\right\|^2 + 2f_{n,\perp}\big(\gamma_\perp(\theta)\big) \le z + \varepsilon_n\right\} \\ &\quad \subseteq \left\{\theta \in \Theta_{\text{osn}} : \left\|\sqrt{n}\gamma(\theta) - \sqrt{n}(\hat{\gamma}_n - \tau)\right\|^2 \le z + \varepsilon_n\right\}\end{aligned}$$

because $f_{n,\perp} \ge 0$. Therefore, on $\mathcal{A}_n$, we have

$$R_n(z) \le e^{2\varepsilon_n} \frac{\int_{\{\theta : \|\sqrt{n}\gamma(\theta) - \sqrt{n}(\hat{\gamma}_n - \tau)\|^2 \le z + \varepsilon_n\} \cap \Theta_{\text{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma(\theta) - \sqrt{n}(\hat{\gamma}_n - \tau)\|^2 - f_{n,\perp}(\gamma_\perp(\theta))} \, \mathrm{d}\Pi(\theta)}{\int_{\Theta_{\text{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma(\theta) - \sqrt{n}(\hat{\gamma}_n - \tau)\|^2 - f_{n,\perp}(\gamma_\perp(\theta))} \, \mathrm{d}\Pi(\theta)}$$

uniformly in z, for all $\mathbb{P} \in \mathbf{P}$.

Define $\Gamma_{\text{osn}} = \{\gamma(\theta) : \theta \in \Theta_{\text{osn}}\}$ and $\Gamma_{\perp,\text{osn}} = \{\gamma_\perp(\theta) : \theta \in \Theta_{\text{osn}}\}$ (if $\mathbb{P}$ is singular). The condition $\sup_{\mathbb{P} \in \mathbf{P}} \sup_{\theta \in \Theta_{\text{osn}}} \|(\gamma(\theta), \gamma_\perp(\theta))\| \to 0$ in Assumption D.2(i) implies that, for all n sufficiently large, we have $\Gamma_{\text{osn}} \times \Gamma_{\perp,\text{osn}} \subset B_\delta^*$ for all $\mathbb{P} \in \mathbf{P}$. By similar arguments to the proof of Lemma 4.2, we use Assumption D.3(ii), a change of variables, and Tonelli's theorem to obtain

$$R_n(z) \le e^{2\varepsilon_n}(1 + \bar{\varepsilon}_n) \frac{\int_{(\{\gamma : \|\sqrt{n}\gamma - \sqrt{n}(\hat{\gamma}_n - \tau)\|^2 \le z + \varepsilon_n) \cap \Gamma_{\text{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma - \sqrt{n}(\hat{\gamma}_n - \tau)\|^2} \, \mathrm{d}\gamma}{\int_{\Gamma_{\text{osn}}} e^{-\frac{1}{2}\|\sqrt{n}\gamma - \sqrt{n}(\hat{\gamma}_n - \tau)\|^2} \, \mathrm{d}\gamma},$$

which holds uniformly in z for all $\mathbb{P} \in \mathbf{P}$ (on $\mathcal{A}_n$ with n sufficiently large) for some sequence $(\bar{\varepsilon}_n)_{n \in \mathbb{N}}$ with $\bar{\varepsilon}_n = o(1)$. A second change of variables with $\sqrt{n}\gamma - \sqrt{n}(\hat{\gamma}_n - \tau) \mapsto \kappa$ yields

$$R_n(z) \le e^{2\varepsilon_n}(1 + \bar{\varepsilon}_n) \frac{\nu_{d^*}\big(\{\kappa : \|\kappa\|^2 \le z + \varepsilon_n\} \cap \big(T_{\text{osn}} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)\big)}{\nu_{d^*}\big(T_{\text{osn}} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)},$$

where $T_{\text{osn}} = \{\sqrt{n}\gamma : \gamma \in \Gamma_{\text{osn}}\} = \{\sqrt{n}\gamma(\theta) : \theta \in \Theta_{\text{osn}}\}$.

Recall that $B_\delta \subset \mathbb{R}^{d^*}$ denotes a ball of radius δ centered at zero. To complete the proof, it is enough to show that

$$\sup_z \left| \frac{\nu_{d^*}\big(B_{\sqrt{z+\varepsilon_n}} \cap \big(T_{\text{osn}} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)\big)}{\nu_{d^*}\big(T_{\text{osn}} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)} - \frac{\nu_{d^*}\big(B_{\sqrt{z}} \cap (T - \sqrt{n}\hat{\gamma}_n)\big)}{\nu_{d^*}(T - \sqrt{n}\hat{\gamma}_n)} \right| = o_{\mathbb{P}}(1) \tag{79}$$

uniformly in $\mathbb{P}$. We split this into three parts. First note that

$$\begin{aligned}&\sup_z \left| \frac{\nu_{d^*}\big(B_{\sqrt{z+\varepsilon_n}} \cap \big(T_{\text{osn}} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)\big)}{\nu_{d^*}\big(T_{\text{osn}} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)} - \frac{\nu_{d^*}\big(B_{\sqrt{z+\varepsilon_n}} \cap \big(T_{\text{osn}} \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)\big)}{\nu_{d^*}\big(T_{\text{osn}} \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)} \right| \\ &\quad \le 2 \frac{\nu_{d^*}\big(\big((T_{\text{osn}} \setminus B_{k_n}) - \sqrt{n}(\hat{\gamma}_n - \tau)\big)\big)}{\nu_{d^*}\big(T_{\text{osn}} \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)} \\ &\quad \le 2 \frac{\nu_{d^*}\big(B_{k_n}^c - \sqrt{n}(\hat{\gamma}_n - \tau)\big)}{\nu_{d^*}\big(T_{\text{osn}} \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)},\end{aligned} \tag{80}$$

where the first inequality is by (51) and the second is by the inclusion $(T_{\text{osn}} \setminus B_{k_n}) \subseteq B_{k_n}^c$. As $\|\sqrt{n}(\hat{\gamma}_n - \tau)\| \le \|\mathbb{V}_n\|$ (by 75) where $\|\mathbb{V}_n\| = O_{\mathbb{P}}(1)$ uniformly in $\mathbb{P}$ and $\inf_{\mathbb{P}\in\mathbf{P}} k_n(\mathbb{P}) \to \infty$ and $d^* = d^*(\mathbb{P}) \le \overline{d} < \infty$, we have

$$\nu_{d^*}\big(B_{k_n}^c - \sqrt{n}(\hat{\gamma}_n - \tau)\big) = o_{\mathbb{P}}(1)$$

uniformly in $\mathbb{P}$. Also notice that, by Assumption D.2(ii),

$$\begin{aligned}
&\frac{\nu_{d^*}\big(B_{\sqrt{z+\varepsilon_n}} \cap \big(T_{\text{osn}} \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)\big)}{\nu_{d^*}\big(T_{\text{osn}} \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)} \\
&\quad = \frac{\nu_{d^*}\big(B_{\sqrt{z+\varepsilon_n}} \cap \big((T - \sqrt{n}\tau) \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)\big)}{\nu_{d^*}\big((T - \sqrt{n}\tau) \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)},
\end{aligned}$$

where, by similar arguments to (80),

$$\begin{aligned}
&\sup_z \left| \frac{\nu_{d^*}\big(B_{\sqrt{z+\varepsilon_n}} \cap \big((T - \sqrt{n}\tau) \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)\big)}{\nu_{d^*}\big((T - \sqrt{n}\tau) \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)} - \frac{\nu_{d^*}(B_{\sqrt{z+\varepsilon_n}} \cap (T - \sqrt{n}\hat{\gamma}_n)}{\nu_{d^*}(T - \sqrt{n}\hat{\gamma}_n)} \right| \\
&\quad \le 2\frac{\nu_{d^*}\big(\big((T - \sqrt{n}\tau) \setminus B_{k_n}\big) - \sqrt{n}(\hat{\gamma}_n - \tau)\big)}{\nu_{d^*}\big((T - \sqrt{n}\tau) \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)} \qquad (81) \\
&\quad \le 2\frac{\nu_{d^*}\big(B_{k_n}^c - \sqrt{n}(\hat{\gamma}_n - \tau)\big)}{\nu_{d^*}\big((T - \sqrt{n}\tau) \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)}. \qquad (82)
\end{aligned}$$

A sufficient condition for the right-hand side of display (82) to be $o_{\mathbb{P}}(1)$ (uniformly in $\mathbb{P}$) is that

$$1/\nu_{d^*}\big((T - \sqrt{n}\tau) \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big) = O_{\mathbb{P}}(1) \quad \text{(uniformly in } \mathbb{P}). \qquad (83)$$

But notice that $\sqrt{n}(\hat{\gamma}_n - \tau)$ is uniformly tight (by (75) and the condition $\|\mathbb{V}_n\| = O_{\mathbb{P}}(1)$ uniformly in $\mathbb{P}$) and $T - \sqrt{n}\tau \supseteq T$. We may therefore deduce by the condition $\inf_{\mathbb{P}\in\mathbf{P}} \nu_{d^*}(T) > 0$ in Assumption D.2(ii) that (83) holds, and so

$$\sup_z \left| \frac{\nu_{d^*}\big(B_{\sqrt{z+\varepsilon_n}} \cap \big((T - \sqrt{n}\tau) \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)\big)}{\nu_{d^*}\big((T - \sqrt{n}\tau) \cap B_{k_n} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)} - \frac{\nu_{d^*}(B_{\sqrt{z+\varepsilon_n}} \cap (T - \sqrt{n}\hat{\gamma}_n)}{\nu_{d^*}(T - \sqrt{n}\hat{\gamma}_n)} \right| = o_{\mathbb{P}}(1)$$

uniformly in $\mathbb{P}$. It also follows that the right-hand side of (80) is $o_{\mathbb{P}}(1)$ (uniformly in $\mathbb{P}$). Hence,

$$\sup_z \left| \frac{\nu_{d^*}\big(B_{\sqrt{z+\varepsilon_n}} \cap \big(T_{\text{osn}} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)\big)}{\nu_{d^*}\big(T_{\text{osn}} - \sqrt{n}(\hat{\gamma}_n - \tau)\big)} - \frac{\nu_{d^*}\big(B_{\sqrt{z+\varepsilon_n}} \cap (T - \sqrt{n}\hat{\gamma}_n)\big)}{\nu_{d^*}(T - \sqrt{n}\hat{\gamma}_n)} \right| = o_{\mathbb{P}}(1)$$

(uniformly in $\mathbb{P}$). To complete the proof of (79), it remains to show that

$$\sup_z \big| \nu_{d^*}\big(B_{\sqrt{z+\varepsilon_n}} \cap (T - \sqrt{n}\hat{\gamma}_n)\big) - \nu_{d^*}\big(B_{\sqrt{z}} \cap (T - \sqrt{n}\hat{\gamma}_n)\big) \big| = o_{\mathbb{P}}(1)$$

holds uniformly in $\mathbb{P}$. But here we have

$$\begin{aligned}
&\sup_z \left|\nu_{d^*}\left(B_{\sqrt{z+\varepsilon_n}} \cap (T - \sqrt{n}\hat{\gamma}_n)\right) - \nu_{d^*}\left(B_{\sqrt{z}} \cap (T - \sqrt{n}\hat{\gamma}_n)\right)\right| \\
&\quad \leq \sup_z \left|\nu_{d^*}(B_{\sqrt{z+\varepsilon_n}} \setminus B_{\sqrt{z}})\right| \\
&\quad = \sup_z \left|F_{\chi^2_{d^*}}(z+\varepsilon_n) - F_{\chi^2_{d^*}}(z)\right| \to 0
\end{aligned}$$

by uniform equicontinuity of $\{F_{\chi^2_d} : d \leq \overline{d}\}$. *Q.E.D.*

PROOF OF THEOREM D.1: We first prove part (i) by verifying the conditions of Lemma D.1. We have $L_n(\hat{\theta}) = \sup_{\theta\in\Theta_{\mathrm{osn}}} L_n(\theta) + o_{\mathbb{P}}(n^{-1})$ uniformly in $\mathbb{P}$. By display (76) in Lemma F.7 and Assumption D.2, we have $\sup_{\theta\in\Theta_I(\mathbb{P})} Q_n(\theta) = \|\mathbf{T}(\mathbb{V}_n + \sqrt{n}\tau) - \sqrt{n}\tau\|^2 + o_{\mathbb{P}}(1)$ uniformly in $\mathbb{P}$. This verifies condition (i) with $W_n = \|\mathbf{T}(\mathbb{V}_n + \sqrt{n}\tau) - \sqrt{n}\tau\|^2$.

For condition (ii), let $\xi_{\alpha,\mathbb{P}}$ denote the α quantile of F_T under $\mathbb{P}$ and let $(\varepsilon_n)_{n\in\mathbb{N}}$ be a positive sequence with $\varepsilon_n = o(1)$. We require that $\|\mathbf{T}(\mathbb{V}_n + \sqrt{n}\tau) - \sqrt{n}\tau\|^2 \leq \|\mathbf{T}\mathbb{V}_n\|^2$ (almost surely) for each $\mathbb{P} \in \mathbf{P}$, which holds trivially when $T = \mathbb{R}^{d^*}$ for each $\mathbb{P}$. By the conditions $\|\mathbf{T}(\mathbb{V}_n + \sqrt{n}\tau) - \sqrt{n}\tau\|^2 \leq \|\mathbf{T}\mathbb{V}_n\|^2$ (almost surely) for each $\mathbb{P} \in \mathbf{P}$, $\sup_{\mathbb{P}\in\mathbf{P}} \sup_z |\mathbb{P}(\|\mathbf{T}\mathbb{V}_n\|^2 \leq z) - \mathbb{P}_Z(\|\mathbf{T}Z\|^2 \leq z)| = o(1)$, and equicontinuity of $\{F_T : \mathbb{P} \in \mathbf{P}\}$ at their α quantiles,

$$\begin{aligned}
\liminf_{n\to\infty} \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}(W_n \leq \xi_{\alpha,\mathbb{P}} - \varepsilon_n) &\geq \liminf_{n\to\infty} \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\left(\|\mathbf{T}\mathbb{V}_n\|^2 \leq \xi_{\alpha,\mathbb{P}} - \varepsilon_n\right) \\
&\geq \liminf_{n\to\infty} \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}_Z\left(\|\mathbf{T}Z\|^2 \leq \xi_{\alpha,\mathbb{P}} - \varepsilon_n\right) \\
&= \alpha.
\end{aligned}$$

By condition D.4, it suffices to show that, for each $\epsilon > 0$,

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\left(\xi_{\alpha,\mathbb{P}} - \xi^{\mathrm{post}}_{n,\alpha} > \epsilon\right) = 0.$$

A sufficient condition is that, for each $\epsilon > 0$,

$$\lim_{n\to\infty} \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\left(\Pi_n\left(\left\{\theta : Q_n(\theta) \leq \xi_{\alpha,\mathbb{P}} - \epsilon\right\} | \mathbf{X}_n\right) < \alpha\right) = 1.$$

By Lemma D.3, there exists a sequence of positive constants $(u_n)_{n\in\mathbb{N}}$ with $u_n = o(1)$ and a sequence of events $(\mathcal{A}_n)_{n\in\mathbb{N}}$ (possibly depending on $\mathbb{P}$) with $\inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}(\mathcal{A}_n) = 1 - o(1)$ such that

$$\Pi_n\left(\left\{\theta : Q_n(\theta) \leq \xi_{\alpha,\mathbb{P}} - \epsilon\right\} | \mathbf{X}_n\right) \leq \mathbb{P}_{Z|\mathbf{X}_n}\left(\|Z\|^2 \leq \xi_{\alpha,\mathbb{P}} - \epsilon | Z \in T - \sqrt{n}\hat{\gamma}_n\right) + u_n$$

holds on $\mathcal{A}_n$ for each $\mathbb{P}$. But by Theorem 2 of Chen and Gao (2017), we also have

$$\mathbb{P}_{Z|\mathbf{X}_n}\left(\|Z\|^2 \leq \xi_{\alpha,\mathbb{P}} - \epsilon | Z \in T - \sqrt{n}\hat{\gamma}_n\right) \leq F_T(\xi_{\alpha,\mathbb{P}} - \epsilon)$$

and hence

$$\Pi_n\left(\left\{\theta : Q_n(\theta) \leq \xi_{\alpha,\mathbb{P}} - \epsilon\right\} | \mathbf{X}_n\right) \leq F_T(\xi_{\alpha,\mathbb{P}} - \epsilon) + u_n$$

holds on $\mathcal{A}_n$ for each $\mathbb{P}$. Also note that by the equicontinuity of $\{F_T : \mathbb{P} \in \mathbf{P}\}$ at their α quantiles,

$$\limsup_{n\to\infty} \sup_{\mathbb{P}\in\mathbf{P}} F_T(\xi_{\alpha,\mathbb{P}} - \epsilon) + u_n < \alpha - \delta \tag{84}$$

for some $\delta > 0$.

We therefore have

$$\begin{aligned}
&\liminf_{n\to\infty} \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(\Pi_n\big(\{\theta : Q_n(\theta) \le \xi_{\alpha,\mathbb{P}} - \epsilon\}|\mathbf{X}_n\big) < \alpha\big) \\
&\ge \liminf_{n\to\infty} \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(\big\{\Pi_n\big(\{\theta : Q_n(\theta) \le \xi_{\alpha,\mathbb{P}} - \epsilon\}|\mathbf{X}_n\big) < \alpha\big\} \cap \mathcal{A}_n\big) \\
&\ge \liminf_{n\to\infty} \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(\big\{F_T(\xi_{\alpha,\mathbb{P}} - \epsilon) + u_n < \alpha\big\} \cap \mathcal{A}_n\big) \\
&\ge 1 - \limsup_{n\to\infty} \sup_{\mathbb{P}\in\mathbf{P}} \mathbb{1}\big\{F_T(\xi_{\alpha,\mathbb{P}} - \epsilon) + u_n \ge \alpha\big\} - \limsup_{n\to\infty} \sup_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(\mathcal{A}_n^c\big) \\
&= 1,
\end{aligned}$$

where the final line is by (84) and definition of $\mathcal{A}_n$.

The proof of part (ii) is similar. *Q.E.D.*

PROOF OF LEMMA D.4: It suffices to characterize the large-sample behavior of $R_n(z)$ defined in (57) uniformly in $\mathbb{P}$. By Lemma F.7 and Assumption D.2(i)–(iii), there exist a positive sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ independent of z with $\varepsilon_n = o(1)$ and a sequence of events $(\mathcal{A}_n)_{n\in\mathbb{N}}$ (possibly depending on $\mathbb{P}$) with $\inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}(\mathcal{A}_n) = 1 - o(1)$ such that

$$\sup_{\theta\in\Theta_{\text{osn}}} \big|PQ_n\big(M(\theta)\big) - f\big(\sqrt{n}(\hat{\gamma}_n - \tau) - \sqrt{n}\gamma(\theta)\big)\big| \le \varepsilon_n,$$

$$\sup_{\theta\in\Theta_{\text{osn}}} \left| nL_n(\theta) - \ell_n - \frac{n}{2}\|\hat{\gamma}_n - \tau\|^2 + \frac{1}{2}\big\|\sqrt{n}\gamma(\theta) - \sqrt{n}(\hat{\gamma}_n - \tau)\big\|^2 + f_{n,\perp}\big(\gamma_\perp(\theta)\big) \right| \le \varepsilon_n$$

both hold on $\mathcal{A}_n$ for all $\mathbb{P} \in \mathbf{P}$. By similar arguments to the proof of Lemma 4.3, wpa1 we obtain

$$\begin{aligned}
(1 - \bar{\varepsilon}_n)e^{-2\varepsilon_n} &\frac{\nu_{d^*}\big(\big(f^{-1}(z - \varepsilon_n)\big) \cap \big(\sqrt{n}(\hat{\gamma}_n - \tau) - T_{\text{osn}}\big)\big)}{\nu_{d^*}\big(\sqrt{n}(\hat{\gamma}_n - \tau) - T_{\text{osn}}\big)} \\
&\le R_n(z) \le (1 + \bar{\varepsilon}_n)e^{2\varepsilon_n} \frac{\nu_{d^*}\big(\big(f^{-1}(z + \varepsilon_n)\big) \cap \big(\sqrt{n}(\hat{\gamma}_n - \tau) - T_{\text{osn}}\big)\big)}{\nu_{d^*}\big(\sqrt{n}(\hat{\gamma}_n - \tau) - T_{\text{osn}}\big)}
\end{aligned}$$

uniformly in z for all $\mathbb{P} \in \mathbf{P}$, for some positive sequence $(\bar{\varepsilon}_n)_{n\in\mathbb{N}}$ with $\bar{\varepsilon}_n = o(1)$. To complete the proof, it remains to show that

$$\sup_{z\in I} \left| \frac{\nu_{d^*}\big(\big(f^{-1}(z + \varepsilon_n)\big) \cap \big(\sqrt{n}(\hat{\gamma}_n - \tau) - T_{\text{osn}}\big)\big)}{\nu_{d^*}\big(\sqrt{n}(\hat{\gamma}_n - \tau) - T_{\text{osn}}\big)} - \frac{\nu_{d^*}\big(f^{-1}(z) \cap (\sqrt{n}\hat{\gamma}_n - T)\big)}{\nu_{d^*}(\sqrt{n}\hat{\gamma}_n - T)} \right| = o_{\mathbb{P}}(1)$$

uniformly in $\mathbb{P}$. This follows by the uniform continuity condition on I in the statement of the lemma, using similar arguments to the proofs of Lemmas 4.3 and D.3. *Q.E.D.*

PROOF OF THEOREM D.2: We verify the conditions of Lemma D.2. We have that $L_n(\hat{\theta}) = \sup_{\theta\in\Theta_{\text{osn}}} L_n(\theta) + o_{\mathbb{P}}(n^{-1})$ uniformly in $\mathbb{P}$. By display (77) in Lemma F.7, we have $PQ_n(M_I) = f(\sqrt{n}(\hat{\gamma}_n - \tau)) + o_{\mathbb{P}}(1)$ uniformly in $\mathbb{P}$. This verifies condition (i) with $W_n = f(\sqrt{n}(\hat{\gamma}_n - \tau)) = f(\mathbf{T}(\mathbb{V}_n + \sqrt{n}\tau) - \sqrt{n}\tau)$.

For condition (ii), let $\xi_{\alpha,\mathbb{P}}$ denote the α quantile of $f(Z)$ under $\mathbb{P}$ and let $(\varepsilon_n)_{n\in\mathbb{N}}$ be a positive sequence with $\varepsilon_n = o(1)$. By Assumption D.5(ii), the condition $\sup_{\mathbb{P}\in\mathbf{P}} \sup_z |\mathbb{P}(f(\mathbb{V}_n) \le z) - \mathbb{P}_Z(f(Z) \le z)| = o(1)$, and equicontinuity of the distributions $\{\mathbb{P}_Z(f(Z) \le z) : \mathbb{P} \in \mathbf{P}\}$ at thier α quantiles, we have

$$\begin{aligned}\liminf_{n\to\infty} \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}(W_n \le \xi_{\alpha,\mathbb{P}} - \varepsilon_n) &\ge \liminf_{n\to\infty} \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(f(\mathbb{V}_n) \le \xi_{\alpha,\mathbb{P}} - \varepsilon_n\big) \\ &\ge \liminf_{n\to\infty} \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}_Z\big(f(Z) \le \xi_{\alpha,\mathbb{P}} - \varepsilon_n\big) \\ &= \alpha.\end{aligned}$$

By condition D.6, it suffices to show that, for each $\epsilon > 0$,

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(\xi_{\alpha,\mathbb{P}} - \xi_{n,\alpha}^{\text{post},p} > \epsilon\big) = 0.$$

A sufficient condition is that

$$\lim_{n\to\infty} \inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\big(\Pi_n\big(\big\{\theta : PQ_n\big(M(\theta)\big) \le \xi_{\alpha,\mathbb{P}} - \epsilon\big\}|\mathbf{X}_n\big) < \alpha\big) = 1.$$

By Lemma D.4, there exists a sequence of positive constants $(u_n)_{n\in\mathbb{N}}$ with $u_n = o(1)$ and a sequence of events $(\mathcal{A}_n)_{n\in\mathbb{N}}$ (possibly depending on $\mathbb{P}$) with $\inf_{\mathbb{P}\in\mathbf{P}} \mathbb{P}(\mathcal{A}_n) = 1 - o(1)$ such that

$$\Pi_n\big(\big\{\theta : PQ_n\big(M(\theta)\big) \le \xi_{\alpha,\mathbb{P}} - \epsilon\big\}|\mathbf{X}_n\big) \le \mathbb{P}_{Z|\mathbf{X}_n}\big(f(Z) \le \xi_{\alpha,\mathbb{P}} - \epsilon | Z \in \sqrt{n}\hat{\gamma}_n - T\big) + u_n$$

holds on $\mathcal{A}_n$ for each $\mathbb{P}$. But by Assumption D.5(iii), we may deduce that

$$\Pi_n\big(\big\{\theta : PQ_n\big(M(\theta)\big) \le \xi_{\alpha,\mathbb{P}} - \epsilon\big\}|\mathbf{X}_n\big) \le \mathbb{P}_Z\big(f(Z) \le \xi_{\alpha,\mathbb{P}} - \epsilon\big) + u_n$$

holds on $\mathcal{A}_n$ for each $\mathbb{P}$. By equicontinuity of the distributions of $\{f(Z) : \mathbb{P} \in \mathbf{P}\}$, we have

$$\limsup_{n\to\infty} \sup_{\mathbb{P}\in\mathbf{P}} \mathbb{P}_Z\big(f(Z) \le \xi_{\alpha,\mathbb{P}} - \epsilon\big) + u_n < \alpha - \delta$$

for some $\delta > 0$. The result now follows by the same arguments as the proof of Theorem D.1. *Q.E.D.*

PROOF OF LEMMA D.5: To simplify notation, let $D_{\theta;p} = \sqrt{\chi^2(p_\theta; p)}$. Define the *generalized score* of $\mathbb{P}_\theta$ with respect to $\mathbb{P}$ as $S_{\theta;p}(x) = g'_{\theta;p} e_x$, where

$$g_{\theta;p} = \frac{1}{D_{\theta;p}} \begin{bmatrix} \frac{p_\theta(1)-p(1)}{p(1)} \\ \vdots \\ \frac{p_\theta(k)-p(k)}{p(k)} \end{bmatrix}.$$

Note that $PS_{\theta;p} = 0$ and $P(S^2_{\theta;p}) = 1$. Also define $u_{\theta;p} = \mathbb{J}_p^{-1} g_{\theta;p}$ and notice that $u_{\theta;p}$ is a unit vector (i.e., $\|u_{\theta;p}\| = 1$). Therefore,

$$\left|S_{\theta;p}(x)\right| \le 1/\Big(\min_{1\le j\le k} \sqrt{p(j)}\Big) \tag{85}$$

for each θ and $\mathbb{P} \in \mathbf{P}$.

For any $p_\theta > 0$, a Taylor series expansion of $\log(u+1)$ about $u = 0$ yields

$$\begin{aligned} nL_n(p_\theta) - nL_n(p) &= n\mathbb{P}_n \log(D_{\theta;p}S_{\theta;p} + 1) \\ &= nD_{\theta;p}\mathbb{P}_n S_{\theta;p} - \frac{nD^2_{\theta;p}}{2}\mathbb{P}_n S^2_{\theta;p} + nD^2_{\theta;p}\mathbb{P}_n\big(S^2_{\theta;p}R(D_{\theta;p}S_{\theta;p})\big), \end{aligned} \tag{86}$$

where $R(u) \to 0$ as $u \to 0$.

By (85), we may choose $(a_n)_{n\in\mathbb{N}}$ to be a positive sequence with $a_n \to \infty$ as $n \to \infty$ such that $a_n \sup_{\theta:p_\theta>0} \max_{1\le i\le n} |S_{\theta;p}(X_i)| = o_{\mathbb{P}}(\sqrt{n})$ (uniformly in $\mathbb{P}$). Then, for any $r_n \le a_n$,

$$\sup_{\theta\in\Theta_{\mathrm{osn}}(\mathbb{P})} \max_{1\le i\le n} \left|D_{\theta;p}S_{\theta;p}(X_i)\right| = o_{\mathbb{P}}(1) \quad \text{(uniformly in } \mathbb{P}). \tag{87}$$

By the two-sided Chernoff bound, for any $\delta \in (0,1)$,

$$\sup_{\mathbb{P}\in\mathbf{P}} \mathbb{P}\left(\max_{1\le j\le k}\left|\frac{\mathbb{P}_n \mathbb{1}\{x=j\}}{p(j)} - 1\right| > \delta\right) \le 2ke^{-n(\inf_{\mathbb{P}\in\mathbf{P}}\min_{1\le j\le k} p(j))\frac{\delta^2}{3}} \to 0 \tag{88}$$

because $\sup_{\mathbb{P}\in\mathbf{P}} \max_{1\le j\le k}(1/p(j)) = o(n)$. It follows that $\mathbb{P}_n(\mathbb{J}_p e_x e_x' \mathbb{J}_p) = I + o_{\mathbb{P}}(1)$ uniformly in $\mathbb{P}$. Also notice that $S^2_{\theta;p}(x) = u'_{\theta;p}\mathbb{J}_p e_x e_x' \mathbb{J}_p u_{\theta;p}$ where each $u_{\theta;p}$ is a unit vector. Therefore,

$$\sup_{\theta:p_\theta>0} \left|\mathbb{P}_n S^2_{\theta;p} - 1\right| = o_{\mathbb{P}}(1) \quad \text{(uniformly in } \mathbb{P}). \tag{89}$$

Substituting (87) and (89) into (86) yields

$$nL_n(p_\theta) - nL_n(p) = nD_{\theta;p}\mathbb{P}_n S_{\theta;p} - \frac{nD^2_{\theta;p}}{2} + nD^2_{\theta;p} \times o_{\mathbb{P}}(1),$$

where the $o_{\mathbb{P}}(1)$ term holds uniformly for all θ with $p_\theta > 0$, uniformly for all $\mathbb{P} \in \mathbf{P}$. We may therefore choose a positive sequence $(b_n)_{n\in\mathbb{N}}$ with $b_n \to \infty$ slowly such that b_n^2 times the $o_{\mathbb{P}}(1)$ term is still $o_{\mathbb{P}}(1)$ uniformly in $\mathbb{P}$. Letting $r_n = (a_n \wedge b_n)$, we obtain

$$\sup_{\theta\in\Theta_{\mathrm{osn}}(\mathbb{P})} \left|nL_n(p_\theta) - nL_n(p) - nD_{\theta;p}\mathbb{P}_n S_{\theta;p} + \frac{nD^2_{\theta;p}}{2}\right| = o_{\mathbb{P}}(1) \quad \text{(uniformly in } \mathbb{P}),$$

where $nD_{\theta;p}\mathbb{P}_n S_{\theta;p} = \sqrt{n}D_{\theta;p}\mathbb{G}_n(S_{\theta;p}) = \sqrt{n}\tilde{\gamma}_{\theta;p}\mathbb{G}_n(\mathbb{J}_p e_x)$ and $D^2_{\theta;p} = \|\tilde{\gamma}_{\theta;p}\|^2$. *Q.E.D.*

PROOF OF PROPOSITION D.1: The quadratic expansion follows from Lemma D.5 and (37) and (38), which give $\|\tilde{\gamma}_{\theta;p}\|^2 = \tilde{\gamma}'_{\theta;p}\tilde{\gamma}_{\theta;p} = \tilde{\gamma}'_{\theta;p}V_p'V_p\tilde{\gamma}_{\theta;p} = \gamma(\theta)'\gamma(\theta)$ and $\tilde{\gamma}'_{\theta;p}\tilde{\mathbb{V}}_{n,p} = \tilde{\gamma}'_{\theta;p}V_p'V_p\tilde{\mathbb{V}}_{n,p} = \gamma(\theta)'\mathbb{V}_n$.

Uniform convergence in distribution is by Proposition A.5.2 of van der Vaart and Wellner (1996), because $\sup_{\mathbb{P}\in\mathbf{P}} \max_{1\le j\le k}(1/p(j)) = o(n)$ implies $\sup_{\mathbb{P}\in\mathbf{P}} |v'_{j,p}\mathbb{J}_p e_x| \le 1/(\min_{1\le j\le k}\sqrt{p(j)}) = o(n^{1/2})$. *Q.E.D.*

PROOF OF PROPOSITION D.2: The condition $\sup_{\mathbb{P}\in\mathbf{P}}\max_{1\leq j\leq k}(1/p(j)) = o(n/\log k)$ ensures that display (88) holds with $k = k(n)\to\infty$. The rest of the proof follows that of Proposition D.1. *Q.E.D.*

PROOF OF LEMMA D.6: Recall that the upper $k-1$ elements of $V_p\tilde{\gamma}_{\theta;p}$ is the vector $\gamma(\theta)=\gamma(\theta;\mathbb{P})$ and the remaining kth element is zero. For each $\mathbb{P}\in\mathbf{P}$, the mapping $\operatorname{int}(\Delta^{k-1})\ni p_\theta\mapsto\gamma(\theta)$ is a homeomorphism. As $\{p_\theta:\theta\in\Theta, p_\theta>0\}=\operatorname{int}(\Delta^{k-1})$ and $p\in\operatorname{int}(\Delta^{k-1})$ for each $\mathbb{P}\in\mathbf{P}$, it follows that $\{\gamma(\theta):\theta\in\Theta, p_\theta>0\}$ contains a ball of radius $\epsilon=\epsilon(\mathbb{P})>0$ for each $\mathbb{P}\in\mathbf{P}$ (because homeomorphisms map interior points to interior points).

Recall that $\theta\in\Theta_{\mathrm{osn}}(\mathbb{P})$ if and only if $\|\gamma(\theta)\|\leq r_n/\sqrt{n}$ (because $\|\gamma(\theta)\|^2=\|\tilde{\gamma}_{\theta;p}\|^2=\chi^2(p_\theta;p)$). Let $\epsilon(\mathbb{P})=\sup\{\epsilon>0:B_\epsilon\subseteq\{\gamma(\theta):\theta\in\Theta, p_\theta>0\}\}$. It suffices to show that $\inf_{\mathbb{P}\in\mathbf{P}}\sqrt{n}\epsilon(\mathbb{P})\to\infty$ as $n\to\infty$. We can map back from any $\gamma\in\mathbb{R}^{k-1}$ by the inverse mapping $q_{\gamma;p}$ given by

$$q_{\gamma;p}(j)=p(j)+\sqrt{p(j)}\big[V_p^{-1}\big((\gamma'\ 0)'\big)\big]_j$$

for $1\leq j\leq k$, where $[V_p^{-1}((\gamma'\ 0)')]_j$ denotes the jth element of $V_p^{-1}((\gamma'\ 0)')$. An equivalent definition of $\epsilon(\mathbb{P})$ is $\inf\{\epsilon>0:q_p(\gamma)\notin\operatorname{int}(\Delta^{k-1})$ for some $\gamma\in B_\epsilon\}$. As $p>0$ and $\sum_{j=1}^k q_{\gamma;p}(j)=1$ for each γ by construction, we therefore need to find the smallest $\epsilon>0$ for which $q_{\gamma;p}(j)=0$ for some j, for some $\gamma\in B_\epsilon$. This is equivalent to finding the smallest $\epsilon>0$ for which

$$\sqrt{p(j)}=-\big[V_p^{-1}\big((\gamma'\ 0)'\big)\big]_j \tag{90}$$

for some j, for some $\gamma\in B_\epsilon$. Also notice that, because the ℓ^2 norm dominates the maximum norm and V_p is an orthogonal matrix, we have

$$\big[V_p^{-1}\big((\gamma'\ 0)'\big)\big]_j\leq\big\|V_p^{-1}\big((\gamma'\ 0)'\big)\big\|=\|\gamma\|\leq\epsilon. \tag{91}$$

It follows from (90) and (91) and the condition $\sup_{\mathbb{P}\in\mathbf{P}}\max_{1\leq j\leq k}(1/p(j))=o(n)$ that $\sqrt{n}\inf_{\mathbb{P}\in\mathbf{P}}\epsilon(\mathbb{P})\geq\frac{\sqrt{n}}{o(\sqrt{n})}\to\infty$ as $n\to\infty$, as required. *Q.E.D.*

PROOF OF LEMMA D.7: First note that

$$\begin{aligned}&\sup_{\theta\in\Theta_{\mathrm{osn}}(\mathbb{P})}\sup_{\mu\in M(\theta)}\Big|\sup_{\eta\in H_\mu}nL_n(p_{\mu,\eta})-\sup_{\eta\in H_\mu:(\mu,\eta)\in\Theta_{\mathrm{osn}}(\mathbb{P})}nL_n(p_{\mu,\eta})\Big|\\&\quad=\sup_{\theta\in\Theta_{\mathrm{osn}}(\mathbb{P})}\sup_{\mu\in M(\theta)}\Big(\inf_{\eta\in H_\mu:(\mu,\eta)\in\Theta_{\mathrm{osn}}(\mathbb{P})}nD_{\mathrm{KL}}(\hat{p}\,\|\,p_{\mu,\eta})-\inf_{\eta\in H_\mu}nD_{\mathrm{KL}}(\hat{p}\,\|\,p_{\mu,\eta})\Big),\end{aligned}$$

where $D_{\mathrm{KL}}(p\,\|\,p_\theta)=\sum_{j=1}^k p(j)\log(p(j)/p_\theta(j))$ and $\hat{p}(j)=\mathbb{P}_n\mathbb{1}\{x=j\}$. By similar arguments to Lemma 3.1 in Liu and Shao (2003), we may deduce that

$$\frac{1}{\chi^2(p_\theta;p)}\big|4h^2(p_\theta,p)-\chi^2(p_\theta;p)\big|\leq\frac{3}{\sqrt{\chi^2(p_\theta;p)}}\max_x\big|S_{\theta,p}(x)\big|h^2(p_\theta,p).$$

Moreover, the proof of Lemma D.5 also shows that $|S_{\theta;p}|\leq 1/(\min_{1\leq j\leq k}\sqrt{p(j)})$ holds for each θ and each $\mathbb{P}\in\mathbf{P}$ so $\max_x|S_{\theta,p}(x)|=o(\sqrt{n})$ uniformly in $\mathbb{P}$. This, together with the fact that $h(p_\theta, p) \leq \sqrt{D_{\mathrm{KL}}(p \parallel p_\theta)} \leq \sqrt{\chi^2(p_\theta; p)}$, yields

$$\frac{1}{\chi^2(p_\theta; p)}\left|4h^2(p_\theta, p) - \chi^2(p_\theta; p)\right| \leq o(\sqrt{n}) \times \sqrt{\chi^2(p_\theta; p)},$$

where the $o(\sqrt{n})$ term holds uniformly for $\theta \in \Theta$ and $\mathbb{P} \in \mathbf{P}$. Let $(a_n)_{n\in\mathbb{N}}$ be a positive sequence with $a_n \leq r_n$ and $a_n \to \infty$ sufficiently slowly that a_n times the $o(\sqrt{n})$ term in the above display is still $o(\sqrt{n})$ (uniformly in θ and $\mathbb{P}$). We then have

$$\sup_{\mathbb{P}\in\mathbf{P}} \sup_{\theta:\chi^2(p_\theta;p)\leq\frac{a_n^2}{n}} \frac{1}{\chi^2(p_\theta; p)}\left|4h^2(p_\theta, p) - \chi^2(p_\theta; p)\right| = o(1). \tag{92}$$

We also want to show that an equivalence holds uniformly over shrinking KL-divergence neighborhoods (rather than χ^2-divergence neighborhoods). By a Taylor expansion of $-\log(u+1)$ about $u = 0$, it is straightforward to deduce that

$$\sup_{\mathbb{P}\in\mathbf{P}} \sup_{\theta\in\Theta:\chi^2(p_\theta;p)\leq\frac{a_n^2}{n}} \left|\frac{D_{\mathrm{KL}}(p \parallel p_\theta)}{\frac{1}{2}\chi^2(p_\theta; p)} - 1\right| = o(1). \tag{93}$$

Condition (88) implies $h^2(p, \hat{p}) \leq \chi^2(\hat{p}; p) = O_{\mathbb{P}}(n^{-1})$ uniformly in $\mathbb{P}$. By the triangle inequality, we have that $\sup_{\theta:\chi^2(p_\theta;p)\leq a_n^2/n} h^2(p_\theta, \hat{p}) = O_{\mathbb{P}}(n^{-1}) + O(a_n^2/n)$ uniformly in $\mathbb{P}$ and hence by (92) that $\sup_{\theta:\chi^2(p_\theta;p)\leq a_n^2/n} \chi^2(p_\theta; \hat{p}) = O_{\mathbb{P}}(n^{-1}) + O(a_n^2/n)$ uniformly in $\mathbb{P}$.

It now follows by (92) and (93) that we may choose positive sequences $(b_n)_{n\in\mathbb{N}}$ and $(\varepsilon_n)_{n\in\mathbb{N}}$ with $b_n \to \infty$ as $n \to \infty$, $b_n = o(a_n)$, $\varepsilon_n b_n^2 = o(1)$, such that (uniformly in $\mathbb{P}$)

$$\begin{aligned}
&n \sup_{\theta:\chi^2(p_\theta;p)\leq\frac{b_n^2}{n}} \sup_{\mu\in M(\theta)} \Big(\inf_{\eta\in H_\mu:\chi^2(p_{\mu,\eta};p)\leq\frac{b_n^2}{n}} D_{\mathrm{KL}}(\hat{p} \parallel p_{\mu,\eta}) - \inf_{\eta\in H_\mu} D_{\mathrm{KL}}(\hat{p} \parallel p_{\mu,\eta})\Big) \\
&\quad = \frac{n}{2} \sup_{\theta:\chi^2(p_\theta;p)\leq\frac{b_n^2}{n}} \sup_{\mu\in M(\theta)} \Big(\inf_{\eta\in H_\mu:\chi^2(p_{\mu,\eta};p)\leq\frac{b_n^2}{n}} \chi^2(p_{\mu,\eta}; \hat{p}) - \inf_{\eta\in H_\mu} \chi^2(p_{\mu,\eta}; \hat{p})\Big) + o_{\mathbb{P}}(1),
\end{aligned}$$

and $\{\theta : \chi^2(p_\theta; \hat{p}) \leq \frac{b_n^2}{n}(1-\varepsilon_n)\} \subseteq \{\theta : \chi^2(p_\theta; p) \leq \frac{b_n^2}{n}\} \subseteq \{\theta : \chi^2(p_\theta; \hat{p}) \leq \frac{b_n^2}{n}(1+\varepsilon_n)\}$ holds wpa1 uniformly in $\mathbb{P}$.

For any $\mu \in M(\theta)$ with θ such that $\chi^2(p_\theta; p) \leq \frac{b_n^2}{n}$, the difference in parentheses in the above display is positive when

$$\inf_{\eta\in H_\mu:\chi^2(p_{\mu,\eta};p)>\frac{b_n^2}{n}} \chi^2(p_{\mu,\eta}; \hat{p}) < \inf_{\eta\in H_\mu:\chi^2(p_{\mu,\eta};p)\leq\frac{b_n^2}{n}} \chi^2(p_{\mu,\eta}; \hat{p})$$

(if the infimum on the left-hand side is over an empty set then the difference in parentheses is zero). As $\{\theta : \chi^2(p_\theta; p) \leq \frac{b_n^2}{n}\} \subseteq \{\theta : \chi^2(p_\theta; \hat{p}) \leq \frac{b_n^2}{n}(1 + \varepsilon_n)\}$ wpa1 uniformly in $\mathbb{P}$, the inequality $\inf_{\eta\in H_\mu:\chi^2(p_{\mu,\eta};p)\leq b_n^2/n} \chi^2(p_{\mu,\eta}; \hat{p}) \leq \frac{b_n^2}{n}(1 + \varepsilon_n)$ holds wpa1 uniformly in $\mathbb{P}$. For the reverse inequality, suppose $\{\eta \in H_\mu : \chi^2(p_{\mu,\eta}; p) > \frac{b_n^2}{n}\}$ is nonempty. As $\{\theta : \chi^2(p_\theta; \hat{p}) > \frac{b_n^2}{n}(1 - \varepsilon_n)\} \supseteq \{\theta : \chi^2(p_\theta; p) > \frac{b_n^2}{n}\}$ holds wpa1 uniformly in $\mathbb{P}$, we have $\inf_{\eta\in H_\mu:\chi^2(p_{\mu,\eta};p)>b_n^2/n} \chi^2(p_{\mu,\eta}; \hat{p}) \geq \frac{b_n^2}{n}(1 - \varepsilon_n)$ wpa1 uniformly in $\mathbb{P}$. The result follows by combining the preceding three inequalities and taking $r_n' = b_n$. *Q.E.D.*

REFERENCES

AZAÏS, J.-M., E. GASSIAT, AND C. MERCADIER (2009): "The Likelihood Ratio Test for General Mixture Models With or Without Structural Parameter," *ESAIM: Probability and Statistics*, 13, 301–327.[18]

CHEN, X., AND F. GAO (2017): "A Reverse Gaussian Correlation Inequality by Adding Cones," *Statistics & Probability Letters*, 123, 84–87.[9,13,29]

GASSIAT, E. (2002): "Likelihood Ratio Inequalities With Application to Various Mixtures," *Annales de L'Institut Henri Poincaré. Section B. Probabilités et Statistiques*, 38 (6), 897–906.[18]

LIU, X., AND Y. SHAO (2003): "Asymptotics for Likelihood Ratio Tests Under Loss of Identifiability," *The Annals of Statistics*, 31 (3), 807–832.[18,33]

NEWEY, W. K., AND D. MCFADDEN (1994): "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics*, Vol. 4. Elsevier, Chapter 36, 2111–2245.[17]

SCHWARTZ, L. (1965): "On Bayes Procedures," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 4 (1), 10–26.[17]

VAN DER VAART, A. W. (2000): *Asymptotic Statistics*. Cambridge University Press.[17]

VAN DER VAART, A. W., AND J. A. WELLNER (1996): *Weak Convergence and Empirical Processes*. Springer-Verlag.[19,20,32]

Guest Editor Keisuke Hirano handled this manuscript.

Manuscript received 5 July, 2016; final version accepted 22 June, 2018; available online 26 July, 2018.