Noname manuscript No.
(will be inserted by the editor)

Supplementary Material for “Bayesian High-dimensional Semi-parametric Inference beyond sub-Gaussian Errors”

Kyoungjae Lee · Minwoo Chae · Lizhen Lin

Received: date / Accepted: date

1 Notation for Proofs

For given real-valued functions l and u, we define the bracket $[l,u]$ as the set of all functions f such that $l \le f \le u$. We call a bracket $[l,u]$ an ϵ-bracket if $d(l,u) < \epsilon$ for a given constant $\epsilon > 0$ and a (semi-)metric d. For a given class of real-valued functions $\mathcal{F}$, the bracketing number $N_{[]}(\epsilon, \mathcal{F}, d)$ is the minimal number of ϵ-brackets which is needed to cover $\mathcal{F}$. The covering number $N(\epsilon, \mathcal{F}, d)$ is the minimal number of ϵ-balls, $\{g : d(f,g) < \epsilon\}$, which is needed to cover $\mathcal{F}$.

For given constant $\epsilon > 0$, the class of real-valued functions $\mathcal{F}$ on $\mathbb{R}^p \times \mathbb{R}$ and the data $D_n = \{(Y_1, x_1), \ldots, (Y_n, x_n)\}$, we denote $N^n_{[]}(\epsilon, \mathcal{F})$ as the minimal number of partition $\{\mathcal{F}_1, \ldots, \mathcal{F}_N\}$ of $\mathcal{F}$ such that

$$\sup_{1\le j\le N} \frac{1}{n}\sum_{i=1}^n \mathbb{E}_{\theta_0,\eta_0}\left[\sup_{f,g\in\mathcal{F}_j} |f(x_i,Y_i) - g(x_i,Y_i)|^2\right] \le \epsilon^2.$$

We define the set of density functions

$$\mathcal{H}_{\rm mix} := \left\{\eta(\cdot) = \int \phi_\sigma(\cdot - z) d\overline{F}(z) : \sigma > 0,\ F \in \mathcal{M}[-C'n, C'n]\right\} \quad (1.1)$$

for the constant $C' > 0$ used in (**??**). Recall that $\overline{F} = (F + F^-)/2$, $dF^-(z) = dF(-z)$ and $\phi_\sigma(z) = (\sqrt{2\pi}\sigma)^{-1}\exp\{-z^2/(2\sigma^2)\}$, for any $z \in \mathbb{R}$.

Kyoungjae Lee
Department of Statistics, Inha University
E-mail: leekjstat@gmail.com

Minwoo Chae
Department of Industrial and Management Engineering, Pohang University of Science and Technology

Lizhen Lin
Department of Applied and Computational Mathematics and Statistics, The University of Notre Dame

2 Proofs for Posterior Convergence Rates

Lemma 1 *Assume that the prior conditions (2)-(8) hold and* η_0 *satisfies (D1)-(D4). If* $\log p \le n^2$*, then there exists a constant* $C_{\rm lower} > 0$ *not depending on* (n,p) *such that the* $\mathbb{P}_{\theta_0,\eta_0}$*-probability of the event*

$$\begin{aligned}\int_{\Theta\times\mathcal{H}_{\rm mix}} & R_n(\theta,\eta)d\Pi(\theta,\eta) \\ &\ge \exp\left[C_{\rm lower}\{\log \pi_p(s_0) - s_0\log p - \lambda\|\theta_0\|_1 - n\tilde{\epsilon}_n^2\}\right] \end{aligned} \tag{2.2}$$

converges to 1 as $n\to\infty$*, where* $\tilde\epsilon_n = n^{-\beta/(2\beta+\kappa^*)}(\log n)^{t_0}$ *and* $t_0 = \{\kappa^*(1+\tau^{-1}+\beta^{-1})+1\}/(2+\kappa^*\beta^{-1})$.

Proof Let $\tilde\sigma_{0n}^{\beta} = \tilde\epsilon_n(\log(1/\tilde\epsilon_n))^{-1}$, and define

$$\begin{aligned}\widetilde{\mathcal{H}}_n := \Big\{\eta\in\mathcal{H}_{\rm mix} : \mathbb{E}_{\eta_0}\left(\log \eta_0/\eta\right) \le A\tilde\epsilon_n^2,\ \mathbb{E}_{\eta_0}\left(\log\eta_0/\eta\right)^2 \le A\tilde\epsilon_n^2, \\ \sigma^{-2} \le \tilde\sigma_{0n}^{-2}(1+\tilde\sigma_{0n}^{2\beta})\Big\},\end{aligned} \tag{2.3}$$

for some constant $A>0$, and

$$\widetilde\Theta_n := \left\{\theta\in\Theta : \|\theta-\theta_0\|_1 \le n^{-5},\ S_\theta = S_0\right\}.$$

Note that

$$\begin{aligned}\int_{\Theta\times\mathcal{H}_{\rm mix}} R_n(\theta,\eta)d\Pi(\theta,\eta) &\ge \int_{\widetilde\Theta_n\times\widetilde{\mathcal{H}}_n} R_n(\theta,\eta)d\Pi(\theta,\eta) \\ &= \int_{\widetilde\Theta_n\times\widetilde{\mathcal{H}}_n} R_n(\theta,\eta)d\widetilde\Pi(\theta,\eta)\cdot \Pi(\widetilde\Theta_n\times\widetilde{\mathcal{H}}_n),\end{aligned}$$

where $\widetilde\Pi = \Pi|_{\widetilde\Theta_n\times\widetilde{\mathcal{H}}_n}$ is the restricted and renormalized prior on $\widetilde\Theta_n\times\widetilde{\mathcal{H}}_n$, that is, $\widetilde\Pi(\cdot) = \Pi(\cdot\cap\widetilde\Theta_n\times\widetilde{\mathcal{H}}_n)/\Pi(\widetilde\Theta_n\times\widetilde{\mathcal{H}}_n)$. We will show that

$$\Pi(\widetilde\Theta_n\times\widetilde{\mathcal{H}}_n) \ge \exp\left[\tilde C_1\left(\log\pi_p(s_0) - s_0\log p - \lambda\|\theta_0\|_1 - n\tilde\epsilon_n^2\right)\right] \tag{2.4}$$

for some constant $\tilde C_1>0$ and all sufficiently large n, and

$$\begin{aligned}\mathbb{P}_{\theta_0,\eta_0}&\left(\int_{\widetilde\Theta_n\times\widetilde{\mathcal{H}}_n} R_n(\theta,\eta)d\widetilde\Pi(\theta,\eta) \le \exp(-\tilde C_2 n\tilde\epsilon_n^2)\right) \\ &\le \frac{2(A+M^2)}{(\tilde C_2 - A - 2M)^2 n\tilde\epsilon_n^2}\end{aligned} \tag{2.5}$$

for some constant $\tilde C_2 > A+2M$. Then, (2.4) and (2.5) complete the proof by taking $C_{\rm lower} = (\tilde C_1\vee\tilde C_2)$.

To obtain inequality (2.4), because $\Pi(\widetilde\Theta_n\times\widetilde{\mathcal{H}}_n) = \Pi_\Theta(\widetilde\Theta_n)\,\Pi_{\mathcal{H}}(\widetilde{\mathcal{H}}_n)$, we derive lower bounds for $\Pi_\Theta(\widetilde\Theta_n)$ and $\Pi_{\mathcal{H}}(\widetilde{\mathcal{H}}_n)$ separately. By Lemma 2, we have

$$\Pi_{\mathcal{H}}(\widetilde{\mathcal{H}}_n) \ge \exp(-C_{\mathcal{H}} n\tilde\epsilon_n^2) \tag{2.6}$$

for all sufficiently large n and some constant $C_{\mathcal{H}} > 0$ not depending on (n, p). By the definition of Π_Θ, we have

$$\Pi_\Theta(\widetilde{\Theta}_n) = \int_{\widetilde{\Theta}_n} d\Pi_\Theta(\theta) = \pi_p(s_0)\binom{p}{s_0}^{-1}\int_{\widetilde{\Theta}_n} g_{S_0}(\theta_{S_0})d\theta_{S_0}$$

and

$$\begin{aligned}
&\int_{\widetilde{\Theta}_n} g_{S_0}(\theta_{S_0})d\theta_{S_0}\\
&\geq e^{-\lambda\|\theta_0\|_1}\int_{\widetilde{\Theta}_n} g_{S_0}(\theta_{S_0}-\theta_{0,S_0})d\theta_{S_0}\\
&= e^{-\lambda\|\theta_0\|_1}\int_{\widetilde{\Theta}_n}\left(\frac{\lambda}{2}\right)^{s_0} e^{-\lambda\|\theta_{S_0}-\theta_{0,S_0}\|_1}d\theta_{S_0}\\
&\geq e^{-\lambda\|\theta_0\|_1}\left(\frac{\lambda}{2}\right)^{s_0} e^{-\lambda n^{-5}}\int_{\{\theta_{S_0}\in\mathbb{R}^{s_0}:\|\theta_{S_0}-\theta_{0,S_0}\|_2\leq(s_0 n^{10})^{-1/2}\}} d\theta_{S_0}\\
&\geq e^{-\lambda\|\theta_0\|_1}\left(\frac{\lambda}{2}\right)^{s_0} e^{-\lambda n^{-1/2}}\frac{\pi^{s_0/2}}{\Gamma(s_0/2+1)}(s_0 n^{10})^{-s_0/2}.
\end{aligned}$$

Thus, the lower bound for $\Pi_\Theta(\widetilde{\Theta}_n)$ is given by

$$\begin{aligned}
&\Pi_\Theta(\widetilde{\Theta}_n)\\
&\geq \pi_p(s_0)\binom{p}{s_0}^{-1} e^{-\lambda\|\theta_0\|_1-\lambda n^{-1/2}}\left(\frac{\lambda\sqrt{\pi}}{2\sqrt{s_0}n^5}\right)^{s_0}\frac{1}{\Gamma(s_0/2+1)}\\
&\geq \pi_p(s_0)p^{-s_0}\Gamma(s_0+1)e^{-\lambda\|\theta_0\|_1-\sqrt{\log p}}\left(\frac{\sqrt{\pi}\sqrt{n}/p}{2\sqrt{s_0}n^5}\right)^{s_0}\frac{1}{\Gamma(s_0/2+1)}\\
&\geq \exp\left\{\log\pi_p(s_0) - s_0\log p - \lambda\|\theta_0\|_1 - \sqrt{\log p}\right\}\left(\frac{1}{\sqrt{s_0}n^5 p}\right)^{s_0}\\
&\geq \exp\left\{\log\pi_p(s_0) - s_0\log p - \lambda\|\theta_0\|_1 - \frac{1}{2}s_0\log p - s_0\log(\sqrt{s_0}n^5 p)\right\}\\
&\geq \exp\left[8\big\{\log\pi_p(s_0) - s_0\log p - \lambda\|\theta_0\|_1\big\}\right]
\end{aligned}$$

for all sufficiently large n because we assume $p \geq n$. Thus,

$$\begin{aligned}
&\Pi(\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n)\\
&= \Pi_\Theta(\widetilde{\Theta}_n)\Pi_{\mathcal{H}}(\widetilde{\mathcal{H}}_n)\\
&\geq \exp\left[8\big\{\log\pi_p(s_0) - s_0\log p - \lambda\|\theta_0\|_1\big\}\right]\exp(-C_{\mathcal{H}}n\tilde{\epsilon}_n^2)\\
&\geq \exp\left[(8\vee C_{\mathcal{H}})\big\{\log\pi_p(s_0) - s_0\log p - \lambda\|\theta_0\|_1 - n\tilde{\epsilon}_n^2\big\}\right],
\end{aligned}$$

which implies (2.4) by taking $\tilde{C}_1 = (8\vee C_{\mathcal{H}})$.

By the Jensen's inequality,

$$\begin{aligned}
&\mathbb{P}_{\theta_0,\eta_0}\left(\int_{\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n} R_n(\theta,\eta)d\widetilde{\Pi}(\theta,\eta) \le \exp(-\tilde{C}_2 n\tilde{\epsilon}_n^2)\right) \\
&\le \mathbb{P}_{\theta_0,\eta_0}\left(\int_{\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n} \sum_{i=1}^n \Big\{\log\frac{\eta(Y_i-x_i^T\theta)}{\eta_0(Y_i-x_i^T\theta_0)}\Big\} d\widetilde{\Pi}(\theta,\eta) \le -\tilde{C}_2 n\tilde{\epsilon}_n^2\right) \\
&= \mathbb{P}_{\theta_0,\eta_0}\left(\sqrt{n}(\widetilde{\mathbb{P}}_n - P_0) \le -\tilde{C}_2\sqrt{n}\tilde{\epsilon}_n^2 - \sqrt{n}P_0\right), \qquad (2.7)
\end{aligned}$$

where $\widetilde{\mathbb{P}}_n := n^{-1}\sum_{i=1}^n \int_{\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n} \log[\eta(Y_i-x_i^T\theta)/\eta_0(Y_i-x_i^T\theta_0)]d\widetilde{\Pi}(\theta,\eta)$ and $P_0 := \mathbb{E}_{\theta_0,\eta_0}[\widetilde{\mathbb{P}}_n]$. Note that

$$\begin{aligned}
-P_0 &\le \max_i\ \mathbb{E}_{\theta_0,\eta_0}\left[\int_{\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n} \log\frac{\eta_0(Y_i-x_i^T\theta_0)}{\eta(Y_i-x_i^T\theta)} d\widetilde{\Pi}(\theta,\eta)\right] \\
&= \max_i \int_{\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n} \mathbb{E}_{\theta_0,\eta_0}\left(\log\frac{\eta_0(Y_i-x_i^T\theta_0)}{\eta(Y_i-x_i^T\theta)}\right) d\widetilde{\Pi}(\theta,\eta) \\
&= \max_i \int_{\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n} \mathbb{E}_{\theta_0,\eta_0}\left(\log\frac{\eta_0(Y_i-x_i^T\theta_0)}{\eta(Y_i-x_i^T\theta_0)} + \log\frac{\eta(Y_i-x_i^T\theta_0)}{\eta(Y_i-x_i^T\theta)}\right) d\widetilde{\Pi}(\theta,\eta) \\
&\le A\tilde{\epsilon}_n^2 + \max_i \int_{\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n} \int \log\frac{\eta(y_i-x_i^T\theta_0)}{\eta(y_i-x_i^T\theta)} \eta_0(y_i-x_i^T\theta_0)\, dy_i\ d\widetilde{\Pi}(\theta,\eta)
\end{aligned}$$

and

$$\begin{aligned}
&\int \log\frac{\eta(y-x^T\theta_0)}{\eta(y-x^T\theta)}\eta_0(y-x^T\theta_0)dy \\
&\le |x^T(\theta-\theta_0)| \int |\dot{\ell}_\eta(y-x^T\theta_0+tx^T(\theta_0-\theta))|\eta_0(y-x^T\theta_0)dy \qquad (2.8)
\end{aligned}$$

for some $t\in[0,1]$ by the mean value theorem. Note that for any $y\in\mathbb{R}$,

$$\begin{aligned}
\sup_{\eta\in\widetilde{\mathcal{H}}_n} |\dot{\ell}_\eta(y)| &\le \sup_{\eta\in\widetilde{\mathcal{H}}_n} \frac{\frac{1}{\sigma^2}\int |y-z|\phi_\sigma(y-z)d\bar{F}(z)}{\int \phi_\sigma(y-z)d\bar{F}(z)} \\
&\le \sup_{\eta\in\widetilde{\mathcal{H}}_n} \frac{1}{\sigma^2}(|y|+C'n) \\
&\le \tilde{\sigma}_{0n}^{-2}(1+\tilde{\sigma}_{0n}^{2\beta})(|y|+C'n) \\
&\le n^2(|y|+n)
\end{aligned}$$

for all sufficiently large n. The above supremum is essentially taken over (F,σ) satisfying (2.3) because of definitions of (1.1) and (2.3). Thus, the right hand side of (2.8) is bounded above by

$$
\begin{aligned}
& M\sqrt{\log p}\,\|\theta-\theta_0\|_1 \int n^2\left(|y-x^T\theta_0|+|x^T(\theta-\theta_0)|+n\right)\eta_0(y-x^T\theta_0)dy \\
\leq\ & M\sqrt{\log p}\,\|\theta-\theta_0\|_1 n^2 \\
& \times\left\{\int |y-x^T\theta_0|\eta_0(y-x^T\theta_0)dy + M\sqrt{\log p}\|\theta-\theta_0\|_1+n\right\} \\
\leq\ & 2M\sqrt{\log p}\; n^{-2} \;\leq\; 2Mn^{-1}
\end{aligned}
$$

for all sufficiently large n on $\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n$, because we assume condition (D2) and $\log p\leq n^2$. Therefore, (2.7) is bounded above by

$$
\begin{aligned}
& \mathbb{P}_{\theta_0,\eta_0}\left(\sqrt{n}(\widetilde{\mathbb{P}}_n-P_0)\leq -\tilde{C}_2\sqrt{n}\tilde{\epsilon}_n^2+\sqrt{n}(A\tilde{\epsilon}_n^2+2Mn^{-1})\right) \\
\leq\ & \mathbb{P}_{\theta_0,\eta_0}\left(\sqrt{n}(\widetilde{\mathbb{P}}_n-P_0)\leq -(\tilde{C}_2-A-2M)\sqrt{n}\tilde{\epsilon}_n^2\right) \\
\leq\ & \frac{1}{(\tilde{C}_2-A-2M)^2 n\tilde{\epsilon}_n^4} \\
& \times\max_i \mathrm{Var}_{\theta_0,\eta_0}\left[\int_{\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n}\log\eta(Y_i-x_i^T\theta)-\log\eta_0(Y_i-x_i^T\theta_0)d\widetilde{\Pi}(\theta,\eta)\right] \\
\leq\ & \frac{1}{(\tilde{C}_2-A-2M)^2 n\tilde{\epsilon}_n^4} \\
& \times\max_i \mathbb{E}_{\theta_0,\eta_0}\left[\int_{\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n}\left(\log\eta(Y_i-x_i^T\theta)-\log\eta_0(Y_i-x_i^T\theta_0)\right)d\widetilde{\Pi}(\theta,\eta)\right]^2 \\
\leq\ & \frac{1}{(\tilde{C}_2-A-2M)^2 n\tilde{\epsilon}_n^4} \\
& \times\max_i \mathbb{E}_{\theta_0,\eta_0}\left[\int_{\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n}\left(\log\eta(Y_i-x_i^T\theta)-\log\eta_0(Y_i-x_i^T\theta_0)\right)^2 d\widetilde{\Pi}(\theta,\eta)\right] \\
=\ & \frac{1}{(\tilde{C}_2-A-2M)^2 n\tilde{\epsilon}_n^4}\max_i\int_{\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n}\mathbb{E}_{\theta_0,\eta_0}\left(\log\frac{\eta_0(Y_i-x_i^T\theta_0)}{\eta(Y_i-x_i^T\theta)}\right)^2 d\widetilde{\Pi}(\theta,\eta)
\end{aligned}
$$

for all sufficiently large n and any constant $\tilde{C}_2>A+2M$. The second and fourth inequalities follow from the Chebyshev's inequality and Jensen's inequality, respectively. Note that

$$
\begin{aligned}
& \mathbb{E}_{\theta_0,\eta_0}\left(\log\frac{\eta_0(Y_i-x_i^T\theta_0)}{\eta(Y_i-x_i^T\theta)}\right)^2 \\
\leq\ & 2\mathbb{E}_{\theta_0,\eta_0}\left(\log\frac{\eta_0(Y_i-x_i^T\theta_0)}{\eta(Y_i-x_i^T\theta_0)}\right)^2+2\mathbb{E}_{\theta_0,\eta_0}\left(\log\frac{\eta(Y_i-x_i^T\theta_0)}{\eta(Y_i-x_i^T\theta)}\right)^2 \\
\leq\ & 2A\tilde{\epsilon}_n^2+2\mathbb{E}_{\theta_0,\eta_0}\left(\log\frac{\eta(Y_i-x_i^T\theta_0)}{\eta(Y_i-x_i^T\theta)}\right)^2
\end{aligned}
$$

and

$$\begin{aligned}&\int\left(\log\frac{\eta(y-x^T\theta_0)}{\eta(y-x^T\theta)}\right)^2\eta_0(y-x^T\theta_0)dy\\ &\le\{x^T(\theta-\theta_0)\}^2\int\left|\dot\ell_\eta(y-x^T\theta_0+tx^T(\theta_0-\theta))\right|^2\eta_0(y-x^T\theta_0)dy\\ &\le M^2\log p\|\theta-\theta_0\|_1^2n^4\Big\{\int 2y^2\eta_0(y)dy+4M^2\log p\|\theta-\theta_0\|_1^2+4n^2\Big\}\\ &\le M^2n^{-1}\end{aligned}$$

for all sufficiently large n on $\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n$. Thus, we have

$$\mathbb{P}_{\theta_0,\eta_0}\left(\int_{\widetilde{\Theta}_n\times\widetilde{\mathcal{H}}_n}R_n(\theta,\eta)d\widetilde{\Pi}(\theta,\eta)\le\exp(-\tilde C_2n\tilde\epsilon_n^2)\right)\le\frac{2(A+M^2)}{(\tilde C_2-A-2M)^2n\tilde\epsilon_n^2}$$

for all sufficiently large n, which completes the proof. ■

Lemma 2 *Under the conditions in Lemma 1,*

$$\Pi_{\mathcal{H}}(\widetilde{\mathcal{H}}_n)\ge\exp(-C_{\mathcal{H}}n\tilde\epsilon_n^2),$$

for some constant $C_{\mathcal{H}}>0$ not depending on (n,p), where $\widetilde{\mathcal{H}}_n$ and $\tilde\epsilon_n$ are defined at (2.3) and Lemma 1, respectively.

Proof We closely follow the steps in the proof of Theorem 4 in **?**. We consider the univariate density case while the original proof in **?** considers d-dimensional case.

By Proposition 1 in **?**, there exist constants δ, s_0, a_0, B_0 and K_0 not depending on (n,p) such that

$$d_H(\eta_0,K_\sigma\tilde h_\sigma)\le K_0\sigma^\beta \tag{2.9}$$

and

$$\mathbb{P}_{\theta_0,\eta_0}(E_\sigma^c)\le B_0\sigma^{4\beta+2\nu+8}$$

for any $\sigma\in(0,s_0)$, where $K_\sigma\tilde h_\sigma=\int\phi_\sigma(x-z)\tilde h_\sigma(z)dz$, $\tilde h_\sigma$ is a probability density function with support inside $(-a_\sigma,a_\sigma)$, $a_\sigma=a_0\{\log(1/\sigma)\}^\tau$ and $E_\sigma:=\{x\in\mathbb{R}:\eta_0(x)\ge\sigma^{(4\beta+2\nu+8)/\delta}\}\subset\{x\in\mathbb{R}:|x|\le a_\sigma\}$. Fix $b_1>\{1\vee1/(2\beta)\}$ such that $\tilde\epsilon_n^{b_1}\{\log(1/\tilde\epsilon_n)\}^{5/4}\le\tilde\epsilon_n$. Let $S_{\tilde\sigma_{0n}}=\{\sigma>0:\sigma^{-2}\in[\tilde\sigma_{0n}^{-2},\tilde\sigma_{0n}^{-2}(1+\tilde\sigma_{0n}^{2\beta})]\}$, where $\tilde\sigma_{0n}=\tilde\epsilon_n^{1/\beta}\{\log(1/\tilde\epsilon_n)\}^{-1/\beta}$. Suppose that $\sigma\in S_{\tilde\sigma_{0n}}$.

By Corollary B1 in **?**, there exists a probability measure $F_\sigma=\sum_{j=1}^N p_j\delta_{z_j}$ satisfying

$$d_H(K_\sigma\tilde h_\sigma,\eta_{F_\sigma,\sigma})\le\tilde A_1\tilde\epsilon_n^{b_1}\{\log(1/\tilde\epsilon_n)\}^{1/4}, \tag{2.10}$$

where $N\le D_0\sigma^{-1}\{\log(1/\sigma)\}^{1/\tau}\log(1/\tilde\epsilon_n)$, $z_i\in[-a_\sigma,a_\sigma]$ $(i=1,\dots,n)$ and $\min_{i\ne j}|z_i-z_j|\ge\sigma\tilde\epsilon_n^{2b_1}$, for some universal constants $\tilde A_1$ and $D_0>0$. Note that $N \le D_0\sigma^{-1}\{\log(1/\sigma)\}^{1/\tau}\log(1/\tilde{\epsilon}_n) \le D_1\sigma^{-1}\{\log(1/\tilde{\epsilon}_n)\}^{1+1/\tau}$ for some universal constant $D_1 > 0$.

Let $U_j = \{x \in \mathbb{R} : |x - z_j| \le \sigma\tilde{\epsilon}_n^{2b_1}/4\}$ for all $j = 1, \dots, N$. Then, one can choose $U_{N+1}, \dots, U_K$ such that (i) $\{U_1, \dots, U_K\}$ is a partition of $[-a_\sigma, a_\sigma]$, (ii) each U_j $(j = N+1, \dots, K)$ has a diameter at most σ and (iii) $K \le D_2\sigma^{-1}\{\log(1/\tilde{\epsilon}_n)\}^{1+1/\tau}$ for some universal constant $D_2 > 0$. Furthermore, one can extend this to a partition $\{U_1, \dots, U_M\}$ of $[-C'n, C'n]$ such that $M \le D_2'\sigma^{-1}\{\log(1/\tilde{\epsilon}_n)\}^{1+1/\tau} \le D_2'\tilde{\epsilon}_n^{-1/\beta}\{\log(1/\tilde{\epsilon}_n)\}^{1+1/\tau+1/\beta}$ and $D_3\sigma\tilde{\epsilon}_n^{2b_1} \le \alpha(U_j) \le 1$ for all $j = 1, \dots, M$ and for some universal constants D_2' and $D_3 > 0$ because of the continuity and positivity of α.

Let $p_j = 0$ for all $j = N+1, \dots, M$. Define $\mathcal{P}_{\tilde{\sigma}_{0n}}$ as the set of probability measures F on $[-C'n, C'n]$ such that

$$\sum_{j=1}^{M} |F(U_j) - p_j| \le 2\tilde{\epsilon}_n^{2b_1} \quad \text{and} \quad \min_{1\le j\le M} F(U_j) \ge \frac{1}{2}\tilde{\epsilon}_n^{4b_1}.$$

Then, we have $\tilde{\epsilon}_n^{2b_1}M \le D_2'\tilde{\epsilon}_n^{2b_1-1/\beta}\{\log(1/\tilde{\epsilon}_n)\}^{1+1/\tau+1/\beta} \le 1$ and $\min_{1\le j\le M}\alpha(U_j) \ge D_3\sigma\tilde{\epsilon}_n^{2b_1} \ge D_3\tilde{\epsilon}_n^{4b_1}$ for all large n. By Lemma 10 in **?**,

$$\begin{aligned}\pi(\mathcal{P}_{\tilde{\sigma}_{0n}}) &\ge C_1 \exp\big\{ - c_1 M \log(1/\tilde{\epsilon}_n)\big\} \\ &\ge C_1 \exp\big[- c_1 D_2'\tilde{\epsilon}_n^{-1/\beta}\{\log(1/\tilde{\epsilon}_n)\}^{2+1/\tau+1/\beta}\big]\end{aligned}$$

for some universal constants C_1 and $c_1 > 0$. In fact, $C_1 = \Gamma(\alpha([-C'n, C'n]))$, but it can be replaced with a universal constant not depending on n by considering $\Gamma(\alpha([-C'n, C'n])) \ge \Gamma(\alpha([-C', C'])) =: C_1$. Also note that, by (**??**),

$$\begin{aligned}\pi(S_{\tilde{\sigma}_{0n}}) &\ge a_6\tilde{\sigma}_{0n}^{-2a_4}\tilde{\sigma}_{0n}^{2\beta a_5}\exp(-C''\tilde{\sigma}_{0n}^{2\beta}) \\ &\ge D_4 \exp\big[- D_5\tilde{\epsilon}_n^{-\kappa/\beta}\{\log(1/\tilde{\epsilon}_n)\}^{\kappa/\beta}\big]\end{aligned}$$

for some universal constant $D_4 > 0$ and some constant $D_5 > 0$ depending only on $C'' > 0$ in (**??**). Therefore, by Lemma B1 in **?** with $V_j = U_j$ for $j = 1, \dots, N$ and $V_0 = \cup_{j=N+1}^{M} U_j$, we have

$$d_H(\eta_{F_\sigma,\sigma}, \eta_{F,\sigma}) \le \tilde{A}_2\tilde{\epsilon}_n^{b_1} \tag{2.11}$$

for any $F \in \mathcal{P}_{\tilde{\sigma}_{0n}}$, $\sigma \in S_{\tilde{\sigma}_{0n}}$ and some constant $\tilde{A}_2 > 0$ not depending on (n, p). Thus, by (2.9)–(2.11),

$$d_H(\eta_0, \eta_{F,\sigma}) \le \tilde{A}_3\tilde{\sigma}_{0n}^{\beta}$$

for any $F \in \mathcal{P}_{\tilde\sigma_{0n}}$, $\sigma \in S_{\tilde\sigma_{0n}}$ and some constant $\tilde A_3 > 0$ not depending on (n,p). Note that $d_H^2(\eta_0, \eta_{F,\sigma}) = d_H^2(\eta_0, \eta_{F^-,\sigma})$ due to condition (D4) and

$$
\begin{aligned}
d_H^2(\eta_0, \eta_{\bar F,\sigma}) &= \int \big(\sqrt{\eta_0} - \sqrt{\eta_{\bar F,\sigma}}\big)^2 d\mu \\
&= \int \big(\sqrt{\eta_0} - \sqrt{(\eta_{F,\sigma} + \eta_{F^-,\sigma})/2}\big)^2 d\mu \\
&= \int_{|\sqrt{\eta_0} - \sqrt{\eta_{F,\sigma}}| > |\sqrt{\eta_0} - \sqrt{\eta_{F^-,\sigma}}|} \big(\sqrt{\eta_0} - \sqrt{(\eta_{F,\sigma} + \eta_{F^-,\sigma})/2}\big)^2 d\mu \\
&+ \int_{|\sqrt{\eta_0} - \sqrt{\eta_{F,\sigma}}| \le |\sqrt{\eta_0} - \sqrt{\eta_{F^-,\sigma}}|} \big(\sqrt{\eta_0} - \sqrt{(\eta_{F,\sigma} + \eta_{F^-,\sigma})/2}\big)^2 d\mu \\
&\le \int_{|\sqrt{\eta_0} - \sqrt{\eta_{F,\sigma}}| > |\sqrt{\eta_0} - \sqrt{\eta_{F^-,\sigma}}|} \big(\sqrt{\eta_0} - \sqrt{\eta_{F,\sigma}}\big)^2 d\mu \\
&+ \int_{|\sqrt{\eta_0} - \sqrt{\eta_{F,\sigma}}| \le |\sqrt{\eta_0} - \sqrt{\eta_{F^-,\sigma}}|} \big(\sqrt{\eta_0} - \sqrt{\eta_{F^-,\sigma}}\big)^2 d\mu \\
&\le d_H^2(\eta_0, \eta_{F,\sigma}) + d_H^2(\eta_0, \eta_{F^-,\sigma}) = 2 d_H^2(\eta_0, \eta_{F,\sigma}).
\end{aligned}
$$

Therefore, we have

$$
d_H(\eta_0, \eta_{\bar F,\sigma}) \le \sqrt{2} \tilde A_3 \tilde\sigma_{0n}^{\beta}
$$

for any $F \in \mathcal{P}_{\tilde\sigma_{0n}}$ and $\sigma \in S_{\tilde\sigma_{0n}}$.

Note that for any $F \in \mathcal{P}_{\tilde\sigma_{0n}}$, $\sigma \in S_{\tilde\sigma_{0n}}$ and $x \in [-a_\sigma, a_\sigma]$,

$$
\begin{aligned}
\frac{\eta_{\bar F,\sigma}(x)}{\eta_0(x)} &\ge \{\sup_{t\in\mathbb{R}} \eta_0(t)\}^{-1} (2\pi\tilde\sigma_{0n}^2)^{-1/2} \int \exp\Big\{-\frac{(x-z)^2}{2\tilde\sigma_{0n}^2}\Big\} d\bar F(z) \\
&\ge K_1 \tilde\sigma_{0n}^{-1} \{F(U_{J(x)}) \wedge F(U_{J(-x)})\} \ \ge \ \frac{K_1}{2} \tilde\sigma_{0n}^{-1} \tilde\epsilon^{4b_1}
\end{aligned}
$$

for some universal constant $K_1 > 0$, where $J(x)$ is the index $j \in \{1, \dots, M\}$ for which $x \in U_j$. On the other hand, for any $F \in \mathcal{P}_{\tilde\sigma_{0n}}$, $\sigma \in S_{\tilde\sigma_{0n}}$ and $x \notin [-a_\sigma, a_\sigma]$,

$$
\begin{aligned}
\frac{\eta_{\bar F,\sigma}(x)}{\eta_0(x)} &\ge K_1 \tilde\sigma_{0n}^{-1} \int_{|z| \le a_\sigma} \exp\Big\{-\frac{(x-z)^2}{2\tilde\sigma_{0n}^2}\Big\} d\bar F(z) \\
&\ge K_1 \tilde\sigma_{0n}^{-1} \exp\Big(-\frac{2x^2}{\tilde\sigma_{0n}^2}\Big) F\big(Z : |Z| \le a_\sigma\big) \\
&\ge K_1 \tilde\sigma_{0n}^{-1} \exp\Big(-\frac{2x^2}{\tilde\sigma_{0n}^2}\Big) (1 - 2\tilde\epsilon_n^{2b_1}) \\
&\ge \frac{K_1}{2} \tilde\sigma_{0n}^{-1} \exp\Big(-\frac{2x^2}{\tilde\sigma_{0n}^2}\Big)
\end{aligned}
$$

for all large n. The third inequality holds because $F \in \mathcal{P}_{\tilde{\sigma}_{0n}}$. Define $\vartheta = \tilde{\sigma}_{0n}^{-1}\tilde{\epsilon}_n^{4b_1}K_1/2$, then $\log(1/\vartheta) \le K_2\log(1/\tilde{\epsilon}_n)$ for some constant $K_2 > 0$ depending only on b_1. Then, for any $F \in \mathcal{P}_{\tilde{\sigma}_{0n}}$ and $\sigma \in S_{\tilde{\sigma}_{0n}}$,

$$\begin{aligned}
&\mathbb{E}_{\eta_0}\Big[\Big\{\log\big(\frac{\eta_0}{\eta_{\bar{F},\sigma}}\big)\Big\}^2 I\Big(\frac{\eta_{\bar{F},\sigma}}{\eta_0} \le \vartheta\Big)\Big] \\
&\le \int_{|x|>a_{\tilde{\sigma}_{0n}}} \Big\{\log\big(\frac{\eta_0(x)}{\eta_{\bar{F},\sigma}(x)}\big)\Big\}^2 \eta_0(x)dx \\
&\le \int_{|x|>a_{\tilde{\sigma}_{0n}}} \Big[\log\Big\{\frac{2\tilde{\sigma}_{0n}}{K_1}\exp\Big(\frac{2x^2}{\tilde{\sigma}_{0n}^2}\Big)\Big\}\Big]^2 \eta_0(x)dx \\
&\le \frac{K_3}{\tilde{\sigma}_{0n}^4}\int_{|x|>a_{\tilde{\sigma}_{0n}}} x^4\eta_0(x)dx \\
&\le \frac{K_3}{\tilde{\sigma}_{0n}^4}\Big(\mathbb{E}_{\eta_0}X^8\Big)^{1/2}\mathbb{P}_{\eta_0}(E_{\tilde{\sigma}_{0n}}^c) \\
&\le K_4\tilde{\sigma}_{0n}^{2\beta+\nu}
\end{aligned}$$

for some constants K_3 and $K_4 > 0$ not depending on (n,p) by construction of $E_{\tilde{\sigma}_{0n}}$. Since $\vartheta < e^{-1}$, it implies that

$$\mathbb{E}_{\eta_0}\Big\{\log\big(\frac{\eta_0}{\eta_{\bar{F},\sigma}}\big)I\Big(\frac{\eta_{\bar{F},\sigma}}{\eta_0} \le \vartheta\Big)\Big\} \le K_4\tilde{\sigma}_{0n}^{2\beta+\nu}.$$

Therefore, by Lemma B2 in **?**, for any $F \in \mathcal{P}_{\tilde{\sigma}_{0n}}$ and $\sigma \in S_{\tilde{\sigma}_{0n}}$,

$$\begin{aligned}
&\mathbb{E}_{\eta_0}\Big\{\log\Big(\frac{\eta_0}{\eta_{\bar{F},\sigma}}\Big)\Big\} \\
&\le d_H^2(\eta_0,\eta_{\bar{F},\sigma})\big\{1+2\log(1/\vartheta)\big\} + 2\mathbb{E}_{\eta_0}\Big\{\log\big(\frac{\eta_0}{\eta_{\bar{F},\sigma}}\big)I\Big(\frac{\eta_{\bar{F},\sigma}}{\eta_0} \le \vartheta\Big)\Big\} \\
&\le 2\tilde{A}_3^2\tilde{\sigma}_{0n}^{2\beta}\{1+2K_2\log(1/\tilde{\epsilon}_n)\} + 2K_4\tilde{\sigma}_{0n}^{2\beta+\nu} \\
&\le 2\tilde{A}_3^2(12+2K_2^2)\tilde{\epsilon}_n^2
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}_{\eta_0}\Big\{\log\Big(\frac{\eta_0}{\eta_{\bar{F},\sigma}}\Big)\Big\}^2 \\
&\le d_H^2(\eta_0,\eta_{\bar{F},\sigma})\big[12+2\big\{\log(1/\vartheta)\big\}^2\big] + 8\mathbb{E}_{\eta_0}\Big[\Big\{\log\big(\frac{\eta_0}{\eta_{\bar{F},\sigma}}\big)\Big\}^2 I\Big(\frac{\eta_{\bar{F},\sigma}}{\eta_0} \le \vartheta\Big)\Big] \\
&\le 2\tilde{A}_3^2\tilde{\sigma}_{0n}^{2\beta}\big[12+2K_2^2\{\log(1/\tilde{\epsilon}_n)\}\big] + 8K_4\tilde{\sigma}_{0n}^{2\beta+\nu} \\
&\le 2\tilde{A}_3^2(12+K_2^2)\tilde{\epsilon}_n^2.
\end{aligned}$$

Thus, by taking $A = 2\tilde{A}_3^2(12 + 2K_2^2)$ in $\widetilde{\mathcal{H}}_n$ defined at (2.3), we have

$$\begin{aligned}
&\Pi_{\mathcal{H}}(\widetilde{\mathcal{H}}_n)\\
&\geq \Pi_{\mathcal{H}}\big((F,\sigma) : F \in \mathcal{P}_{\tilde{\sigma}_{0n}}, \sigma \in S_{\tilde{\sigma}_{0n}}\big)\\
&\geq C_1 D_4 \exp\Big[- c_1 D_2' \tilde{\epsilon}_n^{-1/\beta} \{\log(1/\tilde{\epsilon}_n)\}^{2+1/\tau+1/\beta} - D_5 \tilde{\epsilon}_n^{-\kappa/\beta}\{\log(1/\tilde{\epsilon}_n)\}^{\kappa/\beta}\Big]\\
&\geq C_1 D_4 \exp\Big[- (c_1 D_2' \vee D_5)\tilde{\epsilon}_n^{\kappa^*/\beta}\{\log(1/\tilde{\epsilon}_n)\}^{2+1/\tau+\kappa^*/\beta}\Big]\\
&\geq \exp\big\{ - (c_1 D_2' \vee D_5) n\tilde{\epsilon}_n^2\big\}
\end{aligned}$$

for all large n and some constants c_1, D_2' and $D_5 > 0$ not depending on (n,p). By taking $C_{\mathcal{H}} = (c_1 D_2' \vee D_5)$, it completes the proof. ■

*Proof (Proof of Theorem **??**)* Suppose $\lambda\|\theta_0\|_1 \leq C_\lambda s_0 \log p$ for some constant $C_\lambda > 0$. Let $B := \{(\theta,\eta) : s_\theta \geq R\}$ for some $R > s_0$ and E_n be the event (2.2), then we have

$$\begin{aligned}
&\mathbb{E}_{\theta_0,\eta_0}\Pi(B \mid D_n)\\
&\leq \mathbb{E}_{\theta_0,\eta_0}\left[\Pi(B \mid D_n) I_{E_n}\right] + \mathbb{P}_{\theta_0,\eta_0}(E_n^c)\\
&\leq \mathbb{E}_{\theta_0,\eta_0}\left[\frac{\int_B R_n(\theta,\eta) d\Pi(\theta,\eta)}{\int R_n(\theta,\eta) d\Pi(\theta,\eta)} I_{E_n}\right] + o(1)\\
&\leq \exp\left[C_{\mathrm{lower}}\big\{ - \log \pi_p(s_0) + s_0 \log p + \lambda\|\theta_0\|_1 + n\tilde{\epsilon}_n^2\big\}\right] \cdot \Pi(B) + o(1)\\
&\leq \exp\Big[C_{\mathrm{lower}}\big\{(A_3+1)s_0\log p + s_0 \log p + C_\lambda s_0 \log p + n^{\frac{\kappa^*}{2\beta+\kappa^*}}(\log n)^{2t_0}\big\}\Big]\\
&\quad \times \Pi(B) + o(1)\\
&\leq \exp\Big[C_{\mathrm{lower}}(A_3 + 2 + C_\lambda)\{s_0 \vee n^{\frac{\kappa^*}{2\beta+\kappa^*}}(\log n)^{2t_0-1}\}\log p\Big] \cdot \Pi(B) + o(1)
\end{aligned}$$

by Lemma 1 and condition (**??**). Note that

$$\begin{aligned}
\Pi(B) &\leq \sum_{s=R}^{p} \pi_p(s_0)\left(\frac{A_2}{p^{A_4}}\right)^{s-s_0}\\
&\leq 2\pi_p(s_0)\left(\frac{A_2}{p^{A_4}}\right)^{R-s_0}\\
&\leq \exp\Big\{ - (R - s_0)\frac{A_4}{2}\log p\Big\}
\end{aligned}$$

by condition (**??**). Thus, we have

$$\begin{aligned}
&\mathbb{E}_{\theta_0,\eta_0}\Pi(B \mid D_n)\\
&\leq \exp\Big[- \Big\{(K_{\mathrm{dim}} - 1)\frac{A_4}{2} - C_{\mathrm{lower}}(A_3 + 2 + C_\lambda)\Big\}\\
&\quad \times \{s_0 \vee n^{\frac{\kappa^*}{2\beta+\kappa^*}}(\log n)^{2t_0-1}\}\log p\Big] + o(1)\\
&= o(1)
\end{aligned}$$

by taking $R = K_{\text{dim}}\{s_0 \vee n^{\frac{\kappa^*}{2\beta+\kappa^*}}(\log n)^{2t_0-1}\}$ for some large constant $K_{\text{dim}} > 1 + 2A_4^{-1}C_{\text{lower}}(A_3 + 2 + C_\lambda)$, which completes the proof. ■

Proof (Proof of Theorem **??***)* Define

$$\Theta_n := \left\{\theta \in \Theta : \|\theta - \theta_0\|_1 \le p^2(p + \sqrt{n}) + \|\theta_0\|_1,\ s_\theta \le s_n/2\right\}$$

and for positive constants C_1 and C_2, which will be described below, define

$$\mathcal{H}_n := \left\{\eta(\cdot) = \int \phi_\sigma(\cdot - z)d\bar{F}(z) \text{ with } F = \sum_{h=1}^{\infty}\pi_h\delta_{z_h} : \right.$$
$$\left. z_h \in [-a_n, a_n], h \le H_n; \sum_{h>H_n}\pi_h < \epsilon_n; \sigma^2 \in [\sigma_{0n}^2, \sigma_{0n}^2(1+\epsilon_n^2)^{M_n})\right\}, \tag{2.12}$$

where $a_n^{a_1} = \sigma_{0n}^{-2a_2} = M_n = n, \epsilon_n^2 = C_1 s_n \log p/n$ and $H_n = \lfloor C_2 s_n \log p/\log n\rfloor$. We first prove that

$$\mathbb{E}_{\theta_0,\eta_0}\Pi(\theta \in \Theta_n^c \mid D_n) = o(1) \quad \text{and} \tag{2.13}$$
$$\mathbb{E}_{\theta_0,\eta_0}\Pi(\eta \in \mathcal{H}_n^c \mid D_n) = o(1). \tag{2.14}$$

Suppose $\lambda\|\theta_0\|_1 \le C_\lambda s_0 \log p$ for some constant $C_\lambda > 0$. By Lemma 1 and Theorem **??**,

$$\begin{aligned}
&\mathbb{E}_{\theta_0,\eta_0}\Pi(\theta \in \Theta_n^c \mid D_n)\\
&\le \mathbb{E}_{\theta_0,\eta_0}\Pi\big(\|\theta - \theta_0\|_1 > p^2(p+\sqrt{n}) + \|\theta_0\|_1 \mid D_n\big) + \mathbb{E}_{\theta_0,\eta_0}\Pi(s_\theta > s_n/2 \mid D_n)\\
&\le \mathbb{E}_{\theta_0,\eta_0}\left[\Pi\left(\|\theta - \theta_0\|_1 > p^2(p+\sqrt{n}) + \|\theta_0\|_1 \mid D_n\right) I_{E_n}\right] + o(1)\\
&\le \Pi_\Theta\left(\|\theta - \theta_0\|_1 > p^2(p+\sqrt{n}) + \|\theta_0\|_1\right)\cdot \exp\left\{\frac{C_{\text{lower}}(A_3 + 2 + C_\lambda)}{2K_{\text{dim}}}s_n \log p\right\}\\
&\quad + o(1),
\end{aligned}$$

where E_n is the event (2.2). Note that

$$\begin{aligned}
&\Pi_\Theta\left(\|\theta - \theta_0\|_1 > p^2(p+\sqrt{n}) + \|\theta_0\|_1\right)\\
&\le \Pi_\Theta\left(\|\theta\|_1 > p^2(p+\sqrt{n})\right)\\
&= \sum_{s=1}^{p}\Pi_\Theta\left(\|\theta\|_1 > p^2(p+\sqrt{n}) \mid s_\theta = s\right)\pi_p(s)\\
&\le \sum_{s=1}^{p} s\cdot \max_{1\le h\le s}\Pi_\Theta\left(|\theta_h| > p(p+\sqrt{n})\right)\cdot p^{-A_4 s}A_2^s\\
&\le p\cdot \exp\left(-\lambda p(p+\sqrt{n})\right)\\
&\le \exp\left\{-\frac{1}{2}(n+p)\right\}
\end{aligned}$$

because $\lambda p \geq \sqrt{n}$. Thus, (2.13) holds due to condition $s_n \log p = o(n)$. On the other hand, by Proposition 2 of **?**,

$$\begin{aligned}
&\Pi_{\mathcal{H}}(\mathcal{H}_n^c) \\
&\lesssim H_n \exp(-C'' a_n^{a_1}) + \left\{ \frac{e\alpha(\mathbb{R})}{H_n} \log \frac{1}{\epsilon_n} \right\}^{H_n} + \exp(-C'' \sigma_{0n}^{-2a_2}) \\
&\quad + \sigma_{0n}^{-2a_3} \left(1 + \epsilon_n^2\right)^{-2M_n a_3} \\
&\leq C_2 \frac{s_n \log p}{\log n} \exp(-C'' n) + \exp(-C_2 s_n \log p) + \exp(-C'' n) \\
&\quad + \exp(-C_1 a_3 s_n \log p) \\
&\leq \exp\Big\{ - \frac{(C_1 a_3 \wedge C_2)}{2} s_n \log p \Big\}.
\end{aligned}$$

Then, we have

$$\begin{aligned}
&\mathbb{E}_{\theta_0,\eta_0} \Pi(\eta \in \mathcal{H}_n^c \mid D_n) \\
&\leq \mathbb{E}_{\theta_0,\eta_0} \left[\Pi\left(\eta \in \mathcal{H}_n^c \mid D_n\right) I_{E_n} \right] + o(1) \\
&\lesssim \Pi_{\mathcal{H}}(\mathcal{H}_n^c) \cdot \exp\left\{ \frac{C_{\text{lower}}(A_3 + 2 + C_\lambda)}{2K_{\text{dim}}} s_n \log p \right\} + o(1) \\
&\leq \exp\left[-\frac{1}{2} \Big\{ (C_1 a_3 \wedge C_2) - \frac{C_{\text{lower}}(A_3 + 2 + C_\lambda)}{2K_{\text{dim}}} \Big\} s_n \log p \right] \;=\; o(1)
\end{aligned}$$

for some large constants C_1 and $C_2 > 0$. Thus, we have proved (2.13) and (2.14).

By Lemma 2 and Lemma 9 of **?**, if for some nonincreasing function $\epsilon \mapsto N(\epsilon)$ and some $\epsilon_n' \geq 0$,

$$N\left(\frac{\epsilon}{36}, \Theta_n \times \mathcal{H}_n, d_n\right) \leq N(\epsilon),$$

for all $\epsilon > \epsilon_n'$, then there exists test functions ϕ_n such that

$$\begin{aligned}
\mathbb{P}_{\theta_0,\eta_0} \phi_n &\lesssim \exp\left(-\frac{n}{2}\epsilon_n^2 + \log N(\epsilon_n)\right) \quad \text{and} \\
\sup_{\substack{(\theta,\eta) \in \Theta_n \times \mathcal{H}_n \\ d_n((\theta,\eta),(\theta_0,\eta_0)) > \epsilon_n}} \mathbb{P}_{\theta,\eta}(1 - \phi_n) &\lesssim \exp\left(-\frac{n}{2}\epsilon_n^2\right)
\end{aligned} \tag{2.15}$$

for all $\epsilon_n > \epsilon'_n$. For any $(\theta^i, \eta_i) \in \Theta_n \times \mathcal{H}_n$, $i = 1, 2$,

$$\begin{aligned}
& d_H^2(\eta_1(\cdot - x^T\theta^1), \eta_2(\cdot - x^T\theta^2)) \\
&= \int \left(\sqrt{\eta_1(y - x^T\theta^1)} - \sqrt{\eta_2(y - x^T\theta^2)}\right)^2 dy \\
&\le 2\int \left\{\sqrt{\eta_1(y - x^T\theta^1)} - \sqrt{\eta_1(y - x^T\theta^2)}\right\}^2 dy \\
&+ 2\int \left\{\sqrt{\eta_1(y - x^T\theta^2)} - \sqrt{\eta_2(y - x^T\theta^2)}\right\}^2 dy \\
&\le 2\left\{|x^T(\theta^1 - \theta^2)|^2 \int \left(\int_0^1 \frac{\dot\eta_1(y + td_{12})}{\sqrt{\eta_1(y + td_{12})}} dt\right)^2 dy + d_H^2(\eta_1, \eta_2)\right\} \\
&\le 2\left\{M^2 \log p\, \|\theta^1 - \theta^2\|_1^2 \int_0^1 \int \left(\frac{\dot\eta_1(y + td_{12})}{\eta_1(y + td_{12})}\right)^2 \eta_1(y + td_{12}) dy dt \right. \\
&\qquad \left. + d_H^2(\eta_1, \eta_2)\right\} \\
&\le 2\Big\{M^2 \log p\, \|\theta^1 - \theta^2\|_1^2 \cdot n^{1/a_2} + d_H^2(\eta_1, \eta_2)\Big\},
\end{aligned}$$

where $d_{12} := x^T(\theta^1 - \theta^2)$. The last inequality holds because

$$\begin{aligned}
\left(\frac{\dot\eta(y)}{\eta(y)}\right)^2 \eta(y) &= \frac{\{\dot\eta(y)\}^2}{\eta(y)} \\
&\le \frac{\left\{\int \frac{|y-z|}{\sigma^2} \phi_\sigma(y - z) d\overline{F}(z)\right\}^2}{\eta(y)} \\
&\le \int \left(\frac{y - z}{\sigma^2}\right)^2 \phi_\sigma(y - z) d\overline{F}(z)
\end{aligned}$$

by Hölder's inequality and

$$\begin{aligned}
\int \left(\frac{\dot\eta(y)}{\eta(y)}\right)^2 \eta(y) dy &\le \int\int \left(\frac{y - z}{\sigma^2}\right)^2 \phi_\sigma(y - z) d\overline{F}(z) dy \\
&= \frac{1}{\sigma^2} \int\int \left(\frac{y - z}{\sigma}\right)^2 \phi_\sigma(y - z) dy d\overline{F}(z) \\
&= \frac{1}{\sigma^2} \;\le\; \sigma_{0n}^{-2} = \; n^{1/a_2}.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&\log N\left(\frac{\epsilon}{36}, \Theta_n \times \mathcal{H}_n, d_n\right) \\
&\lesssim \log N\left(\frac{\epsilon}{72Mn^{1/(2a_2)}\sqrt{\log p}}, \Theta_n, \|\cdot\|_1\right) + \log N\left(\frac{\epsilon}{72}, \mathcal{H}_n, d_H\right) \\
&\le \log\left(\sum_{j=0}^{s_n/2}\binom{p}{j}\Big[\frac{p^2(p+\sqrt{n}) + \|\theta_0\|_1}{\epsilon}72M\sqrt{n^{1/a_2}\log p}\Big]^j\right) \\
&\quad + K\Big\{H_n \log\left(\frac{a_n}{\sigma_{0n}\epsilon}\right) + H_n \log\left(\frac{1}{\epsilon}\right) + \log M_n\Big\} \\
&\le \log\left(\sum_{j=0}^{s_n/2}\Big[\frac{p\{p^2(p+\sqrt{n}) + \|\theta_0\|_1\}}{\epsilon}72M\sqrt{n^{1/a_2}\log p}\Big]^j\right) \\
&\quad + K\Big\{H_n \log\left(\frac{a_n}{\sigma_{0n}\epsilon}\right) + H_n \log\left(\frac{1}{\epsilon}\right) + \log M_n\Big\} \\
&\le s_n \log\left(\frac{p^4}{\epsilon}\right) + K\Big\{\frac{C_2 s_n \log p}{\log n}\log\left(\frac{n^{1/a_1+1/(2a_2)}}{\epsilon}\right) + \frac{C_2 s_n\log p}{\log n}\log\left(\frac{1}{\epsilon}\right) \\
&\qquad + \log n\Big\} \\
&=: \log N(\epsilon)
\end{aligned}$$

for some universal constant $K > 0$ by Proposition 2 of **?**. Note that in the last term, we do not have the term $M_n\epsilon_n^2$ while Proposition 2 in **?** includes this term, because they considered d-dimensional densities. It is easy to see that from their proof, the term $M_n\epsilon_n^2$ can be omitted if we focus on univariate ($d = 1$) densities. Note that

$$\begin{aligned}
\log N(\epsilon_n) &\le 5s_n\log p + KC_2\big\{2 + a_1^{-1} + (2a_2)^{-1}\big\}s_n\log p \\
&= \big[5 + KC_2\big\{2 + a_1^{-1} + (2a_2)^{-1}\big\}\big]s_n\log p
\end{aligned}$$

Thus, by (2.15), there exist test functions ϕ_n such that

$$\mathbb{P}_{\theta_0,\eta_0}\phi_n \lesssim \exp\left(-\frac{C_1}{2}s_n\log p + \big[5 + KC_2\big\{2 + a_1^{-1} + (2a_2)^{-1}\big\}\big]s_n\log p\right)$$

and

$$\sup_{\substack{(\theta,\eta)\in\Theta_n\times\mathcal{H}_n \\ d_n((\theta,\eta),(\theta_0,\eta_0))>\epsilon_n}} \mathbb{P}_{\theta,\eta}(1-\phi_n) \lesssim \exp\left(-\frac{C_1}{2}s_n\log p\right).$$

Therefore, by Lemma 1, for a large constant $C_1 > 0$ such that $C_1 > 10 + 2KC_2\{2 + a_1^{-1} + (2a_2)^{-1}\}$ and $C_1 > C_{\text{lower}}(A_3 + 2 + C_\lambda)/K_{\text{dim}}$,

$$\begin{aligned}
&\mathbb{E}_{\theta_0,\eta_0}\Pi\left(d_n((\theta,\eta),(\theta_0,\eta_0)) > \epsilon_n \;\middle|\; D_n\right)\\
&\le \mathbb{E}_{\theta_0,\eta_0}\Pi\left((\theta,\eta)\in\Theta_n\times\mathcal{H}_n : d_n((\theta,\eta),(\theta_0,\eta_0)) > \epsilon_n \;\middle|\; D_n\right) + o(1)\\
&\le \mathbb{E}_{\theta_0,\eta_0}\left[\Pi\left((\theta,\eta)\in\Theta_n\times\mathcal{H}_n : d_n((\theta,\eta),(\theta_0,\eta_0)) > \epsilon_n \;\middle|\; D_n\right)(1-\phi_n)\right]\\
&\quad + o(1)\\
&\lesssim \sup_{\substack{(\theta,\eta)\in\Theta_n\times\mathcal{H}_n\\ d_n((\theta,\eta),(\theta_0,\eta_0))>\epsilon_n}} \mathbb{P}_{\theta,\eta}(1-\phi_n)\cdot\exp\left\{\frac{C_{\text{lower}}(A_3+2+C_\lambda)}{2K_{\text{dim}}} s_n\log p\right\} + o(1)\\
&= o(1).
\end{aligned}$$

It completes the proof by taking $K_{\text{Hel}} = \sqrt{C_1} > \sqrt{C_{\text{lower}}(A_3 + 2 + C_\lambda)/K_{\text{dim}}} \vee \sqrt{10 + 2KC_2\{2 + a_1^{-1} + (2a_2)^{-1}\}}$. ■

Proof (Proof of Corollary **??***)* Let $(T_z(\eta))(x) = \eta(x+z)$. Note that for any η_0 satisfying (D1)-(D4) and $\eta \in \mathcal{H}_{\text{mix}}$,

$$\begin{aligned}
\inf_{z\in\mathbb{R}} d_H(\eta, T_z(\eta_0)) &\le d_H(\eta, T_{x^T(\theta-\theta_0)}(\eta_0))\\
&= \left[\int \left(\sqrt{\eta(y)} - \sqrt{\eta_0(y + x^T(\theta-\theta_0))}\,\right)^2 dy\right]^{1/2}\\
&= \left[\int \left(\sqrt{\eta(y - x^T\theta)} - \sqrt{\eta_0(y - x^T\theta_0)}\,\right)^2 dy\right]^{1/2}\\
&= d_H(\eta(\cdot - x^T\theta), \eta_0(\cdot - x^T\theta_0)),
\end{aligned}$$

thus

$$\begin{aligned}
\inf_{z\in\mathbb{R}} d_H(\eta, T_z(\eta_0)) &\le \left[\frac{1}{n}\sum_{i=1}^n d_H^2(\eta(\cdot - x_i^T\theta), \eta_0(\cdot - x_i^T\theta_0))\right]^{1/2}\\
&= d_n((\theta,\eta),(\theta_0,\eta_0)).
\end{aligned}$$

For any $z \in \mathbb{R}$,

$$\begin{aligned}
d_H^2(\eta_0, T_z(\eta_0)) &= \int (\sqrt{\eta_0(y+z)} - \sqrt{\eta_0(y)})^2 dy \\
&\le z^2 \int \left(\int_0^1 \frac{\dot{\eta}_0(y+tz)}{\sqrt{\eta_0(y+tz)}} dt \right)^2 dy \\
&\le z^2 \int \int_0^1 \left(\frac{\dot{\eta}_0(y+tz)}{\eta_0(y+tz)} \right)^2 \eta_0(y+tz) dt dy \\
&= z^2 \int_0^1 \int \left(\frac{\dot{\eta}_0(y+tz)}{\eta_0(y+tz)} \right)^2 \eta_0(y+tz) dy dt \\
&= z^2 \mathbb{E}_{\eta_0} \left(\frac{\dot{\eta}_0}{\eta_0} \right)^2 \\
&\le z^2 \mathbb{E}_{\eta_0} \left(\frac{|\dot{\eta}_0|}{\eta_0} \right)^{2\beta+\nu} \ \le \ z^2 C_{2\beta+\nu}
\end{aligned}$$

for some constant $C_{2\beta+\nu} > 0$ depending only on (β, ν) because of condition (D3) on η_0 and $2\beta + \upsilon \ge 2$.

If $|z| \le d_H(\eta, \eta_0)/(2\sqrt{C_{2\beta+\nu}})$, then

$$\begin{aligned}
d_H(\eta, T_z(\eta_0)) &\ge d_H(\eta, \eta_0) - d_H(\eta_0, T_z(\eta_0)) \\
&\ge d_H(\eta, \eta_0) - \sqrt{C_{2\beta+\nu}}|z| \\
&\ge \frac{1}{2} d_H(\eta, \eta_0),
\end{aligned}$$

and otherwise, if $|z| > d_H(\eta, \eta_0)/(2\sqrt{C_{2\beta+\nu}})$

$$\begin{aligned}
d_H(\eta, T_z(\eta_0)) &\ge \frac{1}{2} d_V(\eta, T_z(\eta_0)) \\
&= \sup_B |\eta(B) - T_z(\eta_0)(B)| \\
&\ge \left| \int_0^\infty \eta(y) dy - \int_0^\infty \eta_0(y+z) dy \right| \\
&= \left| \int_0^\infty \eta(y) dy - \int_{-z}^\infty \eta_0(y+z) dy - \int_0^{-z} \eta_0(y+z) dy \right| \\
&= \int_0^{|z|} \eta_0(y) dy \qquad (2.16) \\
&\ge \{ \int_0^1 \eta_0(y) dy \} \wedge \{ (2\sqrt{C_{2\beta+\nu}})^{-1} d_H(\eta, \eta_0) \inf_{0 \le y \le 1} \eta_0(y) \},
\end{aligned}$$

where (2.16) holds due to the symmetric assumption (D4) and $\eta \in \mathcal{H}_{\text{mix}}$.

Thus, we have

$$\begin{aligned} K_{\mathrm{Hel}}\sqrt{\frac{s_n\log p}{n}} &\geq d_n\big((\theta,\eta),(\theta_0,\eta_0)\big) \\ &\geq \inf_{z\in\mathbb{R}} d_H(\eta, T_z(\eta_0)) \\ &\geq \left[\frac{1}{2}\wedge\left\{\frac{1}{2\sqrt{C_{2\beta+\nu}}}\inf_{0\leq y\leq 1}\eta_0(y)\right\}\right] d_H(\eta,\eta_0) \end{aligned}$$

because $s_n\log p = o(n)$, which completes the proof by taking $K_{\mathrm{eta}} = K_{\mathrm{Hel}}\Big[\frac{1}{2}\wedge \Big\{\frac{1}{2\sqrt{C_{2\beta+\nu}}}\inf_{0\leq y\leq 1}\eta_0(y)\Big\}\Big]^{-1}$. ■

3 Proofs for Bernstein von-Mises Theorem

We first present three lemmas (Lemma 3, Lemma 4 and Lemma 5), which directly appear in the proof of Theorem ??. Other auxiliary results used to prove these lemmas will be provided in Section 5.

Lemma 3 *Assume that the prior conditions* (??), (??) *and* (??)-(??) *hold. Let*

$$\mathcal{H}_n' = \Bigg\{\eta(\cdot) = \int \phi_\sigma(\cdot - z)d\bar{F}(z) \text{ with } F = \sum_{h=1}^\infty \pi_h\delta_{z_h}: \\ z_h\in[-a_n,a_n], h\leq H_n; \sum_{h>H_n}\pi_h < \epsilon_n; \sigma^2\in[\sigma_{0n}^2, \log n\wedge\{\sigma_{0n}^2(1+\epsilon_n^2)^{M_n}\})\Bigg\},$$

where $a_n = (\log n)^{\frac{2}{\tau}}$, $\epsilon_n^2 = C_1 s_n\log p/n$, $H_n = \lfloor C_2 s_n\log p/\log n\rfloor$, $\sigma_{0n}^{-2a_2} = s_n\log p$, $M_n = n$ *for some positive constants* C_1 *and* C_2, *and define*

$$\mathcal{H}_n^* := \left\{\eta\in\mathcal{H}_n' : d_H(\eta,\eta_0)\leq K_{\mathrm{eta}}\sqrt{s_n\log p/n}\right\}. \tag{3.17}$$

Then,

$$\mathbb{E}_{\theta_0,\eta_0}\mathit{\Pi}\left(\eta\in(\mathcal{H}_n^*)^c \mid D_n\right) = o(1)$$

for any η_0 *satisfying (D1)-(D5).*

Proof We have

$$\begin{aligned} &\mathbb{E}_{\theta_0,\eta_0}\mathit{\Pi}\left(\eta\in(\mathcal{H}_n^*)^c \mid D_n\right) \\ &\leq \mathbb{E}_{\theta_0,\eta_0}\mathit{\Pi}\left(\eta\in(\mathcal{H}_n')^c \mid D_n\right) \\ &+ \mathbb{E}_{\theta_0,\eta_0}\mathit{\Pi}\left(d_H(\eta,\eta_0) > K_{\mathrm{eta}}\sqrt{\frac{s_n\log p}{n}}\,\Bigg|\, D_n\right). \end{aligned} \tag{3.18}$$

Note that Lemma 1 still holds for the prior $\Pi_{\mathcal{H}}$ with the support conditions (??) and (??) because the proof of Theorem 4 of ? can be easily modified for the priors with the restricted support with (??) and (??). Thus,

$$\begin{aligned}
&\mathbb{E}_{\theta_0,\eta_0}\Pi\left(\eta \in (\mathcal{H}_n')^c \mid Dn\right) \\
&\le \mathbb{E}_{\theta_0,\eta_0}\left[\Pi\left(\eta \in (\mathcal{H}_n')^c \mid D_n\right) I_{E_n}\right] + o(1) \\
&\le \Pi_{\mathcal{H}}\big((\mathcal{H}_n')^c\big)\exp\Big\{\frac{C_{\text{lower}}(A_3+2+C_\lambda)}{2K_{\text{dim}}}\tilde{s}_n \log p\Big\} + o(1),
\end{aligned}$$

where E_n is the event (2.2), $\tilde{s}_n = 2K_{\text{dim}}\{s_0 \vee n^{\frac{\kappa^*}{2\beta+\kappa^*}}(\log n)^{2t_0-1}\}$ and $t_0 = \{\kappa^*(1+\tau^{-1}+\beta^{-1})+1\}/(2+\kappa^*\beta^{-1})$. With a slight modification of the proof of Proposition 2 in ?,

$$\begin{aligned}
\Pi_{\mathcal{H}}\big((\mathcal{H}_n')^c\big) &\lesssim H_n \exp\Big\{-C'' a_n^{a_1}\Big\} + \Big\{\frac{e\alpha(\mathbb{R})}{H_n}\log\Big(\frac{1}{\epsilon_n}\Big)\Big\}^{H_n} \\
&\quad + \exp(-C''\sigma_{0n}^{-2a_2}) + \sigma_{0n}^{-2a_3}(1+\epsilon_n^2)^{-2M_n a_3} \\
&\le \exp\Big\{-\frac{1}{2}(C_1 a_3 \wedge C_2 \wedge C'')s_n \log p\Big\}.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\mathbb{E}_{\theta_0,\eta_0}\Pi\left(\eta \in (\mathcal{H}_n')^c \mid Dn\right) \\
&\lesssim \exp\Big\{-\frac{1}{2}(C_1 a_3 \wedge C_2 \wedge C'')s_n \log p + \frac{C_{\text{lower}}(A_3+2+C_\lambda)}{2K_{\text{dim}}}\tilde{s}_n \log p\Big\} + o(1) \\
&= o(1)
\end{aligned}$$

for some large constant $K_{\text{dim}} > 1$. Furthermore, it is easy to see that Corollary ?? also holds for for the prior $\Pi_{\mathcal{H}}$ with (??) and (??), which implies that (3.18) is of order $o(1)$. ■

Lemma 4 *Suppose that* $(s_n \log p)^{1+\frac{8}{a_2}} = o(n^{1-\zeta})$ *holds for some constant* $\zeta > 0$. *Further assume that* $\psi(s_n)$ *is bounded away from zero. Let* $A_S := \{h \in \mathbb{R}^{|S|} : \|h\|_1 > M_n s_n\sqrt{\log p}\}$ *for some sequence* M_n *such that* $\sqrt{\log p} = o(M_n)$. *Then*

$$\sup_{S\in\mathcal{S}_n}\sup_{\eta\in\mathcal{H}_n^*}\frac{\int_{A_S}\exp\left(h^T G_{n,\eta,S} - \frac{1}{2}h^T V_{n,\eta,S}h\right)dh}{\int_{\mathbb{R}^{|S|}}\exp\left(h^T G_{n,\eta,S} - \frac{1}{2}h^T V_{n,\eta,S}h\right)dh} = o_{P_0}(1), \qquad (3.19)$$

where $\mathcal{H}_n^*$ *defined at* (3.17) *and*

$$\mathcal{S}_n := \left\{S : |S| \le \frac{s_n}{2},\ \|\theta_{0,S^c}\|_2 \le \frac{K_{\text{theta}}}{\psi(s_n)}\sqrt{\frac{s_n \log p}{n}}\right\}.$$

Proof Note that

$$\mathbb{E}_{\theta_0,\eta_0}\left(\sup_{S\in\mathcal{S}_n}\sup_{\eta\in\mathcal{H}_n^*}\|G_{n,\eta,S}\|_\infty\right) \lesssim \log p$$

by Lemma 12 and $|h^T G_{n,\eta,S}| \le \|h\|_1 \cdot \|G_{n,\eta,S}\|_\infty$. Also note that

$$\begin{aligned} h^T V_{n,\eta,S} h &= \nu_\eta \cdot h^T \Sigma_S h \\ &= \frac{\nu_\eta}{n} \cdot \|X_S h\|_2^2 \\ &\ge \nu_\eta \cdot \phi^2(s_n)\|h\|_1^2 \cdot \frac{1}{s_n} \;\ge\; \nu_\eta \cdot \psi^2(s_n)\|h\|_1^2 \cdot \frac{1}{s_n}. \end{aligned}$$

Thus, we have

$$\begin{aligned} \sup_{S\in\mathcal{S}_n} \sup_{h\in A_S} \sup_{\eta\in\mathcal{H}_n^*} \frac{|h^T G_{n,\eta,S}|}{h^T V_{n,\eta,S} h} &\lesssim \sup_{S\in\mathcal{S}_n} \sup_{h\in A_S} \sup_{\eta\in\mathcal{H}_n^*} \frac{\|h\|_1 \cdot \|G_{n,\eta,S}\|_\infty \cdot s_n}{\nu_\eta \psi^2(s_n) \cdot \|h\|_1^2} \\ &\le o_{P_0}(1), \end{aligned}$$

because $\sqrt{\log p} = o(M_n)$ and $\nu_{\eta_0} \gtrsim 1$ holds by Lemma 7 and assumptions on η_0. It implies that

$$\begin{aligned} &\sup_{S\in\mathcal{S}_n} \sup_{\eta\in\mathcal{H}_n^*} \int_{A_S} \exp\left(h^T G_{n,\eta,S} - \frac{1}{2} h^T V_{n,\eta,S} h\right) dh \\ &\le \sup_{S\in\mathcal{S}_n} \sup_{\eta\in\mathcal{H}_n^*} \int_{A_S} \exp\left(-C h^T V_{n,\eta,S} h\right) dh \\ &\le \int_{A_S} \exp\left(-\tilde{C}\|h\|_2^2\right) dh \\ &\le \left(\sqrt{\pi} M_n^2 s_n \log p\right)^{\frac{s_n}{2}} \exp\left(-\frac{1}{3}\tilde{C}' M_n^2 s_n \log p\right) \end{aligned}$$

for some positive constants $C, \tilde{C}$ and $\tilde{C}'$, and all sufficiently large n with $\mathbb{P}_{\theta_0,\eta_0}$-probability tending to 1. It is easy to show that

$$\begin{aligned} &\int \exp\left(h^T G_{n,\eta,S} - \frac{1}{2} h^T V_{n,\eta,S} h\right) dh \\ &= (2\pi)^{\frac{|S|}{2}} |V_{n,\eta,S}|^{-\frac{1}{2}} \exp\left(\frac{1}{2\nu_\eta}\|H_S \dot{L}_{n,\eta}\|_2^2\right), \end{aligned}$$

where $H_S = X_S(X_S^T X_S)^{-1} X_S^T$ and $\dot{L}_{n,\eta} = \left(\dot{\ell}_\eta(y_i - x_i^T\theta_0)\right)_{i=1}^n \in \mathbb{R}^n$. Therefore, the log of the left hand side of (3.19) is bounded above by

$$\begin{aligned} &\frac{s_n}{2}\log\left(\sqrt{\pi} M_n^2 s_n \log p\right) - \frac{1}{3}\tilde{C}' M_n^2 s_n \log p - \frac{|S|}{2}\log(2\pi) + \frac{1}{2}\log|V_{n,\eta,S}| \\ &-\frac{1}{2\nu_\eta}\|H_S \dot{L}_{n,\eta}\|_2^2 \\ &\le \frac{s_n}{2}\log\left(\sqrt{\pi} M_n^2 s_n \log p\right) - \frac{1}{3}\tilde{C}' M_n^2 s_n \log p + \frac{s_n}{4}\log\left(M_n^2 \nu_\eta\right) \end{aligned}$$

with $\mathbb{P}_{\theta_0,\eta_0}$-probability tending to 1. The last term tends to $-\infty$ as $n \to \infty$, thus we get the desired result. ■

Define

$$\Theta_n^* := \Bigg\{\theta \in \Theta : S_\theta \in \mathcal{S}_n,\ \|\theta-\theta_0\|_1 \le K_{\rm theta}\frac{s_n}{\phi(s_n)}\sqrt{\frac{\log p}{n}}, \qquad (3.20)$$
$$\|\theta-\theta_0\|_2 \le K_{\rm theta}\frac{1}{\psi(s_n)}\sqrt{\frac{s_n\log p}{n}},\ \|X(\theta-\theta_0)\|_2 \le K_{\rm theta}\sqrt{s_n\log p}\Bigg\},$$

and let $M_n\Theta_n^*$ be the variant of Θ_n^* with $M_n K_{\rm theta}$ instead of $K_{\rm theta}$.

Lemma 5 (Misspecified LAN: version 1) *Suppose that* $s_n^6(\log p)^{11} = o(n^{1-\zeta})$, $(s_n\log p)^{1+\frac{15}{a_2}} = o(n^{1-\zeta})$ *and* $(s_n\log p)^{6+\frac{5}{4a_2}}(\log p)^{\frac{5}{2}} = o(n^{1-\zeta})$ *hold for some constant* $\zeta>0$. *Further assume that* $\psi(s_n)$ *is bounded away from zero. Define* Θ_n^* *and* $\mathcal{H}_n^*$ *as* (3.20) *and* (3.17), *respectively, and let*

$$r_n(\theta,\eta) := L_n(\theta,\eta) - L_n(\theta_0,\eta) - \sqrt{n}(\theta-\theta_0)^T\mathbb{G}_n\dot{\ell}_{\theta_0,\eta_0} + \frac{n}{2}(\theta-\theta_0)^T V_{n,\eta_0}(\theta-\theta_0).$$

Then, we have

$$\mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\theta\in M_n\Theta_n^*}\sup_{\eta\in\mathcal{H}_n^*}|r_n(\theta,\eta)|\right) = o(1)$$

for any η_0 *satisfying (D1)-(D5) and some sequence* M_n *such that* $\sqrt{\log p} = o(M_n)$.

Proof Define $\tilde{r}_n(\theta,\eta)$ as in Lemma 11. Note that

$$\begin{aligned}
&\mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\theta\in M_n\Theta_n^*}\sup_{\eta\in\mathcal{H}_n^*}|r_n(\theta,\eta)|\right)\\
&\le \mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\theta\in M_n\Theta_n^*}\sup_{\eta\in\mathcal{H}_n^*}|r_n(\theta,\eta)-\tilde{r}_n(\theta,\eta)|\right)\\
&\quad+\mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\theta\in M_n\Theta_n^*}\sup_{\eta\in\mathcal{H}_n^*}|\tilde{r}_n(\theta,\eta)|\right),
\end{aligned}$$

and, by Lemma 11,

$$\begin{aligned}
&\mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\theta\in M_n\Theta_n^*}\sup_{\eta\in\mathcal{H}_n^*}|\tilde{r}_n(\theta,\eta)|\right)\\
&\lesssim \frac{M_n^2 s_n^2}{\phi^2(s_n)}\log p\cdot\sqrt{\frac{s_n(\log p)^3+(s_n\log p)^{\frac{3}{a_2}}(\log p)^4}{n}}(s_n\log p)^{\zeta'}\\
&+\frac{M_n^3 s_n}{\phi(s_n)}\sqrt{\frac{\log p}{n}}\cdot s_n(\log p)^{\frac{3}{2}}\\
&= o(1)
\end{aligned}$$

for some small constant $\zeta' > 0$ and some sequence M_n when $(s_n \log p)^{1+\frac{15}{a_2}} = o(n^{1-\zeta})$ and $(s_n \log p)^{6+\frac{5}{4a_2}} (\log p)^{\frac{5}{2}} = o(n^{1-\zeta})$. Thus, it suffices to show that

$$\mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\theta\in M_n\Theta_n^*}\sup_{\eta\in\mathcal{H}_n^*}|r_n(\theta,\eta)-\tilde{r}_n(\theta,\eta)|\right)=o(1).$$

By the definition of $r_n(\theta,\eta)$ and $\tilde{r}_n(\theta,\eta)$,

$$|r_n(\theta,\eta)-\tilde{r}_n(\theta,\eta)| \le \sqrt{n}\left|(\theta-\theta_0)^T\mathbb{G}_n\left(\dot{\ell}_{\theta_0,\eta}-\dot{\ell}_{\theta_0,\eta_0}\right)\right| \tag{3.21}$$

$$+\frac{n}{2}\left|(\theta-\theta_0)^T(V_{n,\eta}-V_{n,\eta_0})(\theta-\theta_0)\right|. \tag{3.22}$$

The supremum of (3.22) is easily bounded above by

$$\begin{aligned}
&\sup_{\theta\in M_n\Theta_n^*}\sup_{\eta\in\mathcal{H}_n^*} n\left|(\theta-\theta_0)^T(V_{n,\eta}-V_{n,\eta_0})(\theta-\theta_0)\right|\\
=&\sup_{\theta\in M_n\Theta_n^*}\sup_{\eta\in\mathcal{H}_n^*}|\nu_\eta-\nu_{\eta_0}|\cdot\|X(\theta-\theta_0)\|_2^2\\
\lesssim&\sup_{\theta\in M_n\Theta_n^*}\sup_{\eta\in\mathcal{H}_n^*}\epsilon_n^{\frac{2}{5}-\frac{\zeta}{2}}M_n^2 s_n\log p
\end{aligned}$$

by Lemma 8, where $\epsilon_n = K_{\mathrm{eta}}\sqrt{s_n\log p/n}$, which is of order $o(1)$ under the assumption $s_n^6(\log p)^{11}=o(n^{1-\zeta})$. Note that

$$\begin{aligned}
\sqrt{n}\left|(\theta-\theta_0)^T\mathbb{G}_n\left(\dot{\ell}_{\theta_0,\eta}-\dot{\ell}_{\theta_0,\eta_0}\right)\right| &\le \sqrt{n}\|\theta-\theta_0\|_1\cdot\|\mathbb{G}_n(\dot{\ell}_{\theta_0,\eta}-\dot{\ell}_{\theta_0,\eta_0})\|_\infty\\
&\lesssim \frac{M_n s_n}{\phi(s_n)}\sqrt{\log p}\cdot\sup_{\eta\in\mathcal{H}_n^*}\|\mathbb{G}_n(\dot{\ell}_{\theta_0,\eta}-\dot{\ell}_{\theta_0,\eta_0})\|_\infty.
\end{aligned}$$

Define

$$\mathcal{L}_{n,j}:=\left\{M_n s_n\sqrt{\log p}\cdot e_j^T\left(\dot{\ell}_{\theta_0,\eta}-\dot{\ell}_{\theta_0,\eta_0}\right):\eta\in\mathcal{H}_n^*\right\}$$

and $\mathcal{L}_n := \cup_{j=1}^p\mathcal{L}_{n,j}$, where e_j is the jth unit vector in $\mathbb{R}^p$. Then $L_n(x,y) := M\sqrt{\log p}\cdot M_n s_n\sqrt{\log p}\cdot\sup_{\eta\in\mathcal{H}_n^*}|\dot{\ell}_\eta(y)-\dot{\ell}_{\eta_0}(y)|$ is an envelop function of $\mathcal{L}_n$, and

$$\begin{aligned}
\|L_n\|_n &\lesssim M_n s_n\log p\cdot\left\{\mathbb{E}_{\theta_0,\eta_0}\left[\sup_{\eta\in\mathcal{H}_n^*}\left(\dot{\ell}_\eta(Y)-\dot{\ell}_{\eta_0}(Y)\right)^2\right]\right\}^{\frac{1}{2}}\\
&\lesssim M_n s_n\log p\cdot\left(\frac{s_n\log p}{n}\right)^{\frac{1}{5}-\zeta}
\end{aligned}$$

by Lemma 8. We will use Corollary A.1 in **?**, which implies

$$\begin{aligned}
&M_n s_n\sqrt{\log p}\cdot\mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\eta\in\mathcal{H}_n^*}\|\mathbb{G}_n(\dot{\ell}_{\theta_0,\eta}-\dot{\ell}_{\theta_0,\eta_0})\|_\infty\right)\\
\lesssim&\int_0^{\|L_n\|_n}\sqrt{\log N_{[]}^n(\epsilon,\mathcal{L}_n)}d\epsilon.
\end{aligned}$$

Note that

$$N^n_{[]}(\epsilon, \mathcal{L}_{n,j}) \le N_{[]}\left(\frac{\epsilon}{MM_n s_n \log p}, \mathcal{G}_n, L_2(P_{\eta_0})\right),$$

where $\mathcal{G}_n := \{\dot{\ell}_\eta : \eta \in \mathcal{H}^*_n\}$, and

$$\begin{aligned} \log N_{[]}\left(\epsilon, \mathcal{G}_n, L_2(P_{\eta_0})\right) &\le \log N_{[]}\left(\epsilon^\gamma, \mathcal{H}^*_n, d_H\right) \\ &\le \log N_{[]}\left(\epsilon^\gamma, \mathcal{H}_n, d_H\right). \end{aligned} \tag{3.23}$$

Let $a_n = (\log n)^{\frac{2}{\tau}}, b_{1n} = (s_n \log p)^{-\frac{1}{2a_2}}$ and $b_{2n} = \sqrt{\log n}$. By Lemma 3 of **?**,

$$\log N\left(\epsilon, \mathcal{H}_n, \|\cdot\|_\infty\right) \lesssim \frac{a_n}{b_{1n}} \cdot \log\frac{1}{\epsilon} \cdot \left(\log\frac{1}{\epsilon} + \log\frac{a_n}{b_{1n}}\right).$$

Now we use the similar argument to the proof of Theorem 6 of **?**. Define

$$H(x) = b^{-1}_{1n}\phi\left(\frac{x}{2b_{2n}}\right) I(|x| > 2a_n) + b^{-1}_{1n}\phi(0) I(|x| \le 2a_n),$$

where ϕ is the density function of the standard normal distribution. H is an envelop function for $\mathcal{H}_n$. For some $\varrho > 0$, let $g_1, \dots, g_T$ be a ϱ-net for $\|\cdot\|_\infty$, $l_i := (g_i - \varrho) \vee 0$ and $u_i := (g_i + \varrho) \wedge H$. Then, the brackets $[l_i, u_i]$ cover $\mathcal{H}_n$. Let $\varrho = C\epsilon^2(a_n b_{2n})^{-1}[\log(1/\epsilon)]^{-\frac{1}{2}}$ for some constant $C > 0$, then for $D_n = 2a_n b_{2n}[\log(1/\epsilon)]^{\frac{1}{2}} > 2a_n$,

$$\begin{aligned} \int (u_i - l_i)\, d\mu &\lesssim \|u_i - l_i\|_\infty \cdot D_n + \int_{|x|>D_n} \frac{1}{b_{1n}}\phi\left(\frac{x}{2b_{2n}}\right) dx \\ &\lesssim \varrho \cdot D_n + \frac{b_{2n}}{b_{1n}} \exp\left(-\frac{D^2_n}{8b^2_{2n}}\right) \\ &\lesssim \epsilon^2 + \frac{b_{2n}}{b_{1n}} \cdot \epsilon^{ca^2_n} \\ &\lesssim \epsilon^2 \end{aligned}$$

for some constant $c > 0$ and any $\epsilon < 1$. The second inequality follows from the Chernoff's inequality. Thus,

$$\begin{aligned} \log N_{[]}\left(\epsilon, \mathcal{H}_n, d_H\right) &\le \log N_{[]}\left(\epsilon^2, \mathcal{H}_n, \|\cdot\|_1\right) \\ &\le \log N\left(C \cdot \frac{\epsilon^2}{a_n b_{2n}}\left[\log\frac{1}{\epsilon}\right]^{-\frac{1}{2}}, \mathcal{H}_n, \|\cdot\|_\infty\right) \\ &\lesssim \frac{a_n}{b_{1n}} \cdot \left[\left(\log\frac{1}{\epsilon}\right)^2 + (\log n)^2\right], \end{aligned}$$

and by (3.23),

$$\begin{aligned} \log N^n_{[]}\left(\epsilon, \mathcal{L}_n\right) &\le \log p + \log N_{[]}\left(\frac{\epsilon}{MM_n s_n \log p}, \mathcal{G}_n, L_2(P_{\eta_0})\right) \\ &\lesssim \log p + (s_n \log p)^{\frac{1}{2a_2}} \left[\log n\right]^{\frac{2}{\tau}} \cdot \left[\left(\log\frac{1}{\epsilon}\right)^2 + (\log n)^2\right]. \end{aligned}$$

Then by Corollary A.1 in ?, we have

$$\begin{aligned}
&\mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\eta\in\mathcal{H}_n^*}\|\mathbb{G}_n(\dot{\ell}_{\theta_0,\eta}-\dot{\ell}_{\theta_0,\eta_0})\|_\infty\right)\cdot\frac{M_n s_n}{\phi(s_n)}\sqrt{\log p}\\
&\lesssim \int_0^{\|L_n\|_n}\sqrt{\log N_{[]}^n(\epsilon,\mathcal{L}_n)}d\epsilon\\
&\lesssim \int_0^{\|L_n\|_n}\sqrt{\log p}+(s_n\log p)^{\frac{1}{4a_2}}\left[\log n\right]^{\frac{1}{\tau}}\cdot\left(\log\frac{1}{\epsilon}+\log n\right)d\epsilon\\
&\lesssim \|L_n\|_n\sqrt{\log p}+(s_n\log p)^{\frac{1}{4a_2}}\left[\log n\right]^{\frac{1}{\tau}+1}\cdot\int_0^{\|L_n\|_n}\log\frac{1}{\epsilon}\,d\epsilon\\
&\lesssim M_n s_n \log p\cdot\left(\frac{s_n\log p}{n}\right)^{\frac{1}{5}-\zeta'}\left\{\sqrt{\log p}+(s_n\log p)^{\frac{1}{4a_2}}\left[\log n\right]^{\frac{1}{\tau}+1}\right\} \qquad (3.24)
\end{aligned}$$

because $\int_0^u \log(1/\epsilon)d\epsilon \le \int_0^u \epsilon^{-1+\zeta''}d\epsilon \lesssim u^{1-\zeta''}$ for any small $\zeta''>0$ and $0<u<1$. (3.24) converges to zero as $n\to\infty$ under the assumptions $(s_n\log p)^{6+\frac{5}{4a_2}}(\log p)^{\frac{5}{2}} = o(n^{1-\zeta})$ and $s_n^6(\log p)^{11} = o(n^{1-\zeta})$ for some constant $\zeta>0$. Thus, we have shown (3.21), and this completes the proof. ■

Now, we prove Theorem ?? using the above results (Lemma 3, Lemma 4 and Lemma 5) and posterior convergence rate results (Theorem ??, Corollary ?? and Corollary ??).

Proof (Proof of Theorem ??) Let Θ_n^* and $\mathcal{H}_n^*$ be defined as (3.20) and (3.17), respectively. Define $\breve{\Pi}_\Theta := \Pi_\Theta|_{M_n\Theta_n^*}$ and $\breve{\Pi}_\mathcal{H} := \Pi_\mathcal{H}|_{\mathcal{H}_n^*}$ as the restricted and renormalized priors on $M_n\Theta_n^*$ and $\mathcal{H}_n^*$, respectively. Let $\breve{\Pi}(\cdot|D_n)$ be the posterior distribution corresponding to the prior $\breve{\Pi}=\breve{\Pi}_\Theta\times\breve{\Pi}_\mathcal{H}$. We first prove that

$$d_V\left(\breve{\Pi}(\cdot|D_n),\Pi(\cdot|D_n)\right)=o_{P_0}(1) \quad \text{and} \qquad (3.25)$$

$$d_V\left(\breve{\Pi}^\infty(\cdot|D_n),\Pi^\infty(\cdot|D_n)\right)=o_{P_0}(1), \qquad (3.26)$$

where $\breve{\Pi}^\infty(\cdot|D_n) := \Pi^\infty(\cdot|D_n)|_{M_n\Theta_n^*}$. Note that for any measurable set $A\in\Theta\times\mathcal{H}$,

$$\begin{aligned}
\breve{\Pi}(A\mid D_n) &= \frac{\Pi(A\cap[M_n\Theta_n^*\times\mathcal{H}_n^*]\mid D_n)}{\Pi(M_n\Theta_n^*\times\mathcal{H}_n^*\mid D_n)}\\
&= \frac{\Pi(A\mid D_n)-\Pi(A\cap[M_n\Theta_n^*\times\mathcal{H}_n^*]^c\mid D_n)}{\Pi(\Theta\times\mathcal{H}_{\text{mix}}\mid D_n)-\Pi([M_n\Theta_n^*\times\mathcal{H}_n^*]^c\mid D_n)}\\
&= \Pi(A\mid D_n)+o_{P_0}(1)
\end{aligned}$$

by Corollaries ??, ?? and Lemma 3, which implies (3.25). Define

$$\mathcal{S}_n := \left\{S:|S|\le\frac{s_n}{2},\quad \|\theta_{0,S^c}\|_2\le\frac{K_{\text{theta}}}{\psi(s_n)}\sqrt{\frac{s_n\log p}{n}}\right\}, \qquad (3.27)$$

$\Theta_S^* := \{\theta_S \in \mathbb{R}^{|S|} : \widetilde{\theta}_S \in M_n \Theta_n^*\}$ and $H_S := \sqrt{n}(\Theta_S^* - \theta_{0,S})$ for some sequence M_n such that $\sqrt{\log p} = o(M_n)$ and

$$\sup_{\theta \in M_n \Theta_n^*} \sup_{\eta \in \mathcal{H}_n^*} |r_n(\theta, \eta)| = o_{P_0}(1),$$

where $r_n(\theta, \eta)$ is defined in Lemma 5. Then,

$$d\breve{\Pi}(\theta \mid D_n) = \sum_{S \in \mathcal{S}_n} \widetilde{w}_S \cdot d\widetilde{Q}_S(\theta_S) d\delta_0(\theta_{S^c}),$$
$$d\breve{\Pi}^\infty(\theta \mid D_n) = \sum_{S \in \mathcal{S}_n} \widetilde{w}_S^\infty \cdot n^{-\frac{|S|}{2}} d\widetilde{N}_{n,S}(h_S) d\delta_0(\theta_{S^c}),$$

where $\widetilde{Q}_S = Q_S|_{\Theta_S^*}$ and $\widetilde{N}_{n,S} := N_{n,S}|_{H_S}$ are the restricted and renormalized distributions,

$$\widetilde{w}_S := \frac{Q_S(\Theta_S^*)}{\sum_{S' \in \mathcal{S}_n} w_{S'} Q_{S'}(\Theta_{S'}^*)} \cdot w_S,$$
$$\widetilde{w}_S^\infty := \frac{N_{n,S}(H_S)}{\sum_{S' \in \mathcal{S}_n} w_{S'} N_{n,S'}(H_{S'})} \cdot w_S,$$

and $h_S = \sqrt{n}(\theta_S - \theta_{0,S}) \in H_S$. It is easy to show that

$$\sup_{S \in \mathcal{S}_n} \left| 1 - \frac{w_S}{\widetilde{w}_S^\infty} \right| = o_{P_0}(1) \quad \text{and} \tag{3.28}$$
$$\sup_{S \in \mathcal{S}_n} d_V\left(N_{n,S}, \widetilde{N}_{n,S}\right) = o_{P_0}(1) \tag{3.29}$$

hold by Theorem **??** and Lemma 4. Then, by Lemma 4.5 in **?**,

$$\begin{aligned}
& d_V\left(\breve{\Pi}^\infty(\cdot|D_n), \Pi^\infty(\cdot|D_n)\right) \\
& \le 2 d_V(\widetilde{w}^\infty, w) + \sum_{S \in \mathcal{S}} w_S d_V(\widetilde{N}_{n,S}, N_{n,S}) \\
& \le 2 \sum_{S \in \mathcal{S}_n} \widetilde{w}_S^\infty \left| 1 - \frac{w_S}{\widetilde{w}_S^\infty} \right| + \sum_{S \in \mathcal{S}_n} w_S \cdot \sup_{S \in \mathcal{S}_n} d_V(\widetilde{N}_{n,S}, N_{n,S}) \\
& \quad + 4 \sum_{S \in \mathcal{S}_n^c} w_S,
\end{aligned}$$

where $w = (w_S)_{S \in \mathcal{S}}$ and $\widetilde{w}^\infty = (\widetilde{w}_S^\infty)_{S \in \mathcal{S}_n}$. It implies that (3.26) holds by (3.28), (3.29) and Theorem **??**.

Now we have (3.25) and (3.26), so it suffices to prove that

$$d_V\left(\breve{\Pi}(\cdot|D_n), \breve{\Pi}^\infty(\cdot|D_n)\right) = o_{P_0}(1). \tag{3.30}$$

Again by Lemma 4.5 in **?**, if we show that

$$d_V(\widetilde{w}, \widetilde{w}^{\infty}) = o_{P_0}(1) \quad \text{and} \tag{3.31}$$

$$\sup_{S\in\mathcal{S}_n} d_V(\widetilde{Q}_S, \widetilde{N}_{n,S}) = o_{P_0}(1), \tag{3.32}$$

where $\widetilde{w} = (\widetilde{w}_S)_{S\in\mathcal{S}_n}$, it implies the desired result, (3.30). Note that

$$\begin{aligned}
d_V(\widetilde{w}, \widetilde{w}^{\infty}) &= \sum_{S\in\mathcal{S}_n} |\widetilde{w}_S - \widetilde{w}_S^{\infty}| \\
&= \sum_{S\in\mathcal{S}_n} \left|1 - \frac{\widetilde{w}_S}{\widetilde{w}_S^{\infty}}\right| \cdot \widetilde{w}_S^{\infty} \\
&= \sum_{S\in\mathcal{S}_n} \left|1 - Q_S(\Theta_S^*)\frac{w_S}{\widetilde{w}_S^{\infty}}(1 + o_{P_0}(1))\right| \cdot \widetilde{w}_S^{\infty} \\
&= \sum_{S\in\mathcal{S}_n} |1 - Q_S(\Theta_S^*)(1 + o_{P_0}(1))| \cdot \widetilde{w}_S^{\infty} \\
&\le \sup_{S\in\mathcal{S}_n} (1 - Q_S(\Theta_S^*)) + o_{P_0}(1) \;\; = \;\; o_{P_0}(1).
\end{aligned}$$

The third and fourth equality hold by Theorem **??**, Corollary **??** and (3.28), respectively. Thus, we have proved (3.31). For any measurable set B,

$$\begin{aligned}
&\breve{\Pi}(\theta_S \in B \mid D_n, \eta, S_\theta = S) \\
&= \frac{\int_{B\cap\Theta_S^*} \exp\Big(L_n(\widetilde{\theta}_S, \eta) - L_n(\theta_0, \eta)\Big) \cdot g_S(\theta_S)/g_S(\theta_{0,S})\, d\theta_S}{\int_{\Theta_S^*} \exp\Big(L_n(\widetilde{\theta}_S, \eta) - L_n(\theta_0, \eta)\Big) \cdot g_S(\theta_S)/g_S(\theta_{0,S})\, d\theta_S} \\
&= \frac{\int_{B\cap\Theta_S^*} \exp\big(\sqrt{n}(\theta_S - \theta_{0,S})^T G_{n,\eta_0,S} - \frac{n}{2}(\theta_S - \theta_{0,S})^T V_{n,\eta_0,S}(\theta_S - \theta_{0,S})\big)\, d\theta_S}{\int_{\Theta_S^*} \exp\big(\sqrt{n}(\theta_S - \theta_{0,S})^T G_{n,\eta_0,S} - \frac{n}{2}(\theta_S - \theta_{0,S})^T V_{n,\eta_0,S}(\theta_S - \theta_{0,S})\big)\, d\theta_S} \\
&\quad + o_{P_0}(1)
\end{aligned}$$

by Lemma 5 and

$$\begin{aligned}
\sup_{S\in\mathcal{S}_n}\sup_{\theta_S\in\Theta_S^*}\left|\log\frac{g_S(\theta_S)}{g_S(\theta_{0,S})}\right| &= \sup_{S\in\mathcal{S}_n}\sup_{\theta_S\in\Theta_S^*}\Big|\log\exp\left(\lambda\|\theta_{0,S} - \theta_S\|_1\right)\Big| \\
&\lesssim \sup_{S\in\mathcal{S}_n} \lambda\cdot\frac{M_n s_n}{\phi(s_n)}\sqrt{\frac{\log p}{n}} \;\; = \;\; o(1)
\end{aligned}$$

for some sequence M_n such that $\sqrt{\log p} = o(M_n)$ because we assume $\lambda s_n \log p = o(\sqrt{n})$. Then,

$$\begin{aligned}
\widetilde{Q}_S(h_S \in B) &= \int_{\mathcal{H}_n^*} \breve{\Pi}(h_S \in B \mid D_n, \eta, S_\theta = S)\, d\breve{\Pi}(\eta \mid D_n, S_\theta = S) \\
&= \int_{\mathcal{H}_n^*} \widetilde{N}_{n,S}(B)\, d\breve{\Pi}(\eta \mid D_n, S_\theta = S) + o_{P_0}(1) \\
&= \widetilde{N}_{n,S}(B) + o_{P_0}(1),
\end{aligned}$$

which implies $\sup_{S\in\mathcal{S}_n} d_V(\widetilde{Q}_S, \widetilde{N}_{n,S}) = o_{P_0}(1)$. ■

4 Proof for Strong Model Selection Consistency

*Proof (Proof of Theorem **??**)* Define $\mathcal{S}_n$ and $\breve{\Pi}$ as in the proof of Theorem 3.5. Define the set $\mathcal{S}_n' = \{S \in \mathcal{S}_n : S \supsetneq S_0\}$, then it suffices to show that $\breve{\Pi}(S_\theta \in \mathcal{S}_n' \mid D_n) \longrightarrow 0$ by (3.25). Note that

$$\begin{aligned}&\breve{\Pi}(S_\theta = S \mid D_n, \eta)\\ &= \frac{\pi_p(|S|)\binom{p}{|S|}^{-1} \int_{\Theta_S^*} \exp\left(L_n(\widetilde{\theta}_S, \eta) - L_n(\theta_0, \eta)\right) g_S(\theta_S) d\theta_S}{\sum_{S \in \mathcal{S}_n} \pi_p(|S|)\binom{p}{|S|}^{-1} \int_{\Theta_S^*} \exp\left(L_n(\widetilde{\theta}_S, \eta) - L_n(\theta_0, \eta)\right) g_S(\theta_S) d\theta_S}.\end{aligned}$$

Then, by Lemma 5,

$$\begin{aligned}&\breve{\Pi}(S_\theta \in \mathcal{S}_n' \mid D_n, \eta)\\ &= \frac{\sum_{S \in \mathcal{S}_n'} \pi_p(|S|)\binom{p}{|S|}^{-1} \int_{\Theta_S^*} \exp\left(L_n(\widetilde{\theta}_S, \eta) - L_n(\theta_0, \eta)\right) g_S(\theta_S) d\theta_S}{\sum_{S \in \mathcal{S}_n} \pi_p(|S|)\binom{p}{|S|}^{-1} \int_{\Theta_S^*} \exp\left(L_n(\widetilde{\theta}_S, \eta) - L_n(\theta_0, \eta)\right) g_S(\theta_S) d\theta_S}\\ &\le \sum_{S \in \mathcal{S}_n'} \frac{\widehat{w}_S}{\widehat{w}_{S_0}} e^{2\xi_n}\\ &\lesssim \sum_{s=s_0+1}^{s_n/2} \frac{\pi_p(s)}{\pi_p(s_0)} \binom{s}{s_0} \left(\frac{\lambda\sqrt{\pi}}{\sqrt{2\nu_{\eta_0}}}\right)^{s-s_0}\\ &\quad \times \max_{|S|=s} \left[\frac{|X_{S_0}^T X_{S_0}|^{1/2}}{|X_S^T X_S|^{1/2}} \exp\left\{\frac{1}{2\nu_{\eta_0}} \|(H_S - H_{S_0})\dot{L}_{n,\eta_0}\|_2^2\right\}\right]\end{aligned}$$

for any η and some sequence $\xi_n \to 0$, where

$$\begin{aligned}&\widehat{w}_S\\ &\propto \pi_p(|S|)\binom{p}{|S|}^{-1} \times\\ &\int_{\Theta_S^*} \exp\left(\sqrt{n}(\theta_S - \theta_{0,S})^T G_{n,S} - \frac{n}{2}(\theta_S - \theta_{0,S})^T V_{n,S}(\theta_S - \theta_{0,S})\right) g_S(\theta_S) d\theta_S.\end{aligned}$$

Note that, by the condition on π_p and the definition of $\psi^2(s)$, $\pi_p(s)/\pi_p(s_0) \le A_2^{s-s_0} p^{-A_4(s-s_0)}$ and $|X_{S_0}^T X_{S_0}|/|X_S^T X_S| \le (n\psi^2(s_n))^{|S|-s_0}$ for any $S \in \mathcal{S}_n'$. Thus, it suffices to prove that

$$\begin{aligned}&\mathbb{P}_{\eta_0}\left(\frac{1}{2\nu_{\eta_0}} \|(H_S - H_{S_0})\dot{L}_{n,\eta_0}\|_2^2 > K_{\rm sel}(s - s_0)\log p, \text{ for some } S \in \mathcal{S}_n'\right)\\ &= o(1) \qquad (4.33)\end{aligned}$$

for some positive constant $K_{\rm sel}$ depending only on η_0 such that $A_4 > K_{\rm sel}$.

The left hand side of (4.33) is bounded above by

$$\sum_{s=s_0+1}^{s_n/2} \binom{p-s_0}{s-s_0} \mathbb{P}_{\eta_0}\left(\|(H_S - H_{S_0})\dot{L}_{n,\eta_0}\|_2^2 > 2\nu_{\eta_0} K_{\rm sel}(s-s_0)\log p \right)$$
$$\leq \sum_{s=s_0+1}^{s_n/2} \binom{p-s_0}{s-s_0} e^{-t\cdot 2\nu_{\eta_0} K_{\rm sel}(s-s_0)(\log p - \nu_{\eta_0} K_{\rm sel}^{-1})} \times \mathbb{E}_{\theta_0,\eta_0} e^{t\|(H_S-H_{S_0})\dot{L}_{n,\eta_0}\|_2^2}$$

for any $t > 0$, where $H_S = X_S(X_S^T X_S)^{-1} X_S^T$. Note that $\dot{\ell}_{\eta_0}(y_i - x_i^T\theta_0)$ is a sub-Gaussian by assumption. By Lemma B.2 in **?** (Hanson-Wright inequality),

$$\mathbb{E}_{\theta_0,\eta_0} e^{t_0\|(H_S-H_{S_0})\dot{L}_{n,\eta_0}\|_2^2} \lesssim e^{C(|S|-s_0)}$$

for some positive constants C and t_0 depending only on η_0. Thus, if we choose $K_{\rm sel} = (\nu_{\eta_0} t_0)^{-1}$, the left hand side of (4.33) tends to zero as $n \to \infty$. ■

5 Auxiliary Lemmas

We first introduce Lemma 6, which is used to prove lemmas 7, 8 and 9.

Lemma 6 *Let B be a subset of $\mathbb{R}$ and for given $\epsilon > 0$, p and q be probability densities on $\mathbb{R}$ such that $d_H^2(p,q) \leq \epsilon^2$. Suppose $M_\delta^2 := \int_B p(p/q)^\delta < \infty$ for some $\delta \in (0,1)$. Then,*

$$\int_B p\left(\log\frac{p}{q}\right)^2 \leq 20\epsilon^2 \left[\frac{1}{\delta}\left(1 \vee \log\frac{M_\delta}{\epsilon}\right)\right]^2.$$

Proof The main strategy for the proof is similar to the proof of Theorem 5 in **?**. Note that

$$\int_B p\left(\log\frac{p}{q}\right)^2 \leq \int_{0<p/q\leq K^2} p\left(\log\frac{p}{q}\right)^2 + \int_{B\cap(p/q>K^2)} p\left(\log\frac{p}{q}\right)^2$$

for any $K > 0$. Let $K^\delta = e \vee (M_\delta/\epsilon) > 1$ and $r = \sqrt{p/q} - 1$. Then,

$$\begin{aligned}
\int_{0<p/q\leq K^2} p\left(\log\frac{p}{q}\right)^2 &= \int_{-1<r\leq K-1} q(r+1)^2(2\log(r+1))^2 \\
&= \int_{-1<r\leq K-1,\, r\neq 0} qr^2\left(\frac{r+1}{r}\right)^2 (2\log(r+1))^2 \\
&\leq 16\int_{-1<r\leq K-1,\, r\neq 0} qr^2(\log K)^2 \;\;\leq\;\; 16\epsilon^2(\log K)^2
\end{aligned}$$

because $(x+1)/x\log(x+1)$ is increasing for $x>-1, x\neq 0$ and $\int qr^2 = d_H^2(p,q)\leq \epsilon^2$ by assumption. On the other hand,

$$\begin{aligned}\int_{B\cap(p/q>K^2)} p\left(\log\frac{p}{q}\right)^2 &= \int_{B\cap(p/q>K^2)} p\left(\frac{p}{q}\right)^\delta \frac{(\log\frac{p}{q})^2}{(\frac{p}{q})^\delta}\\ &\leq \int_{B\cap(p/q>K^2)} p\left(\frac{p}{q}\right)^\delta \frac{(2\log K)^2}{K^{2\delta}}\\ &\leq 4M_\delta^2\frac{(\log K)^2}{K^{2\delta}},\end{aligned}$$

because $\log x/x^\delta$ is decreasing for $x\geq e^{1/\delta}$. Thus, we have

$$\begin{aligned}\int_B p\left(\log\frac{p}{q}\right)^2 &\leq 16\epsilon^2(\log K)^2 + 4M_\delta^2\frac{(\log K)^2}{K^{2\delta}}\\ &\leq 20\epsilon^2\left[\frac{1}{\delta}\left(1\vee\log\frac{M_\delta}{\epsilon}\right)\right]^2\end{aligned}$$

by the definition of K. ■

The following lemma gives a (uniform) convergence rate for the score function, which plays an important role in proving the BvM theorem. This lemma is used to prove lemmas 4 and 12.

Lemma 7 *Let $\epsilon_n = K_{\text{eta}}\sqrt{s_n\log p/n}$ and assume that $(s_n\log p)^2 = o(n)$. For any constant $\zeta>0$, there exists a constant $K_\zeta>0$ not depending on (n,p) such that*

$$\int \sup_{\eta\in\mathcal{H}_n^*}\left(\dot\ell_\eta(y)-\dot\ell_{\eta_0}(y)\right)^2 dP_{\eta_0}(y)\leq K_\zeta\,(\epsilon_n)^{\frac{4}{5}-\zeta}\,(s_n\log p)^{\frac{16}{5a_2}}$$

for any η_0 satisfying (D1)-(D5) and all sufficiently large n, where $\mathcal{H}_n^*$ defined at (3.17).

Proof We first state some inequalities that we frequently use in the proof. For any $\eta \in \mathcal{H}_n^*$ and any $y \in \mathbb{R}$,

$$
\begin{aligned}
|\ell_\eta(y)| &= \Big|\log\Big\{\int (2\pi\sigma^2)^{-1/2}\exp\big(-(y-z)^2/(2\sigma^2)\big)d\overline{F}(z)\Big\}\Big| \\
&\le \Big|\log\Big\{(s_n\log p)^{\frac{1}{2a_2}}\exp\Big(-(y^2+(\log n)^{\frac{4}{\tau}})(s_n\log p)^{\frac{1}{a_2}}\Big)\Big\}\Big| \\
&\le \frac{1}{a_2}\log(s_n\log p)+\big\{y^2+(\log n)^{\frac{4}{\tau}}\big\}(s_n\log p)^{\frac{1}{a_2}} \\
&\le 2\big\{y^2+(\log n)^{\frac{4}{\tau}}\big\}(s_n\log p)^{\frac{1}{a_2}}, \\
|\dot{\ell}_\eta(y)| &= \Big|\frac{\int -(\frac{y-z}{\sigma^2})\phi_\sigma(y-z)d\overline{F}(z)}{\int \phi_\sigma(y-z)d\overline{F}(z)}\Big| \\
&\le \frac{1}{\sigma^2}\big\{|y|+(\log n)^{\frac{2}{\tau}}\big\} \\
&\le \big\{|y|+(\log n)^{\frac{2}{\tau}}\big\}(s_n\log p)^{\frac{1}{a_2}}, \\
|\ddot{\ell}_\eta(y)| &= \Big|\frac{\ddot{\eta}(y)}{\eta(y)}-\Big\{\frac{\dot{\eta}(y)}{\eta(y)}\Big\}^2\Big| \\
&\le \frac{|\ddot{\eta}(y)|}{\eta(y)}+|\dot{\ell}_\eta(y)|^2 \\
&\le \frac{1}{\eta(y)}\Big|\int\frac{1}{\sigma^2}\phi_\sigma(y-z)d\overline{F}(z)+\int\frac{(y-z)^2}{\sigma^4}\phi_\sigma(y-z)d\overline{F}(z)\Big| \\
&\quad +2\big\{y^2+(\log n)^{\frac{4}{\tau}}\big\}(s_n\log p)^{\frac{2}{a_2}} \\
&\le \frac{1}{\sigma^2}+\frac{2}{\sigma^4}\big\{y^2+(\log n)^{\frac{4}{\tau}}\big\}+2\big\{y^2+(\log n)^{\frac{4}{\tau}}\big\}(s_n\log p)^{\frac{2}{a_2}} \\
&\le 5\big\{y^2+(\log n)^{\frac{4}{\tau}}\big\}(s_n\log p)^{\frac{2}{a_2}}
\end{aligned}
$$

and

$$
\begin{aligned}
|\dddot{\ell}_\eta(y)| &= \Big|\frac{\dddot{\eta}(y)}{\eta(y)}-\frac{\dot{\eta}(y)\ddot{\eta}(y)}{\{\eta(y)\}^2}-2\dot{\ell}_\eta(y)\ddot{\ell}_\eta(y)\Big| \\
&\le \frac{1}{\eta(y)}\Big\{\int\frac{(y-z)}{\sigma^4}\phi_\sigma(y-z)d\overline{F}(z)+\int\frac{2|y-z|}{\sigma^4}\phi_\sigma(y-z)d\overline{F}(z) \\
&\quad +\int\frac{|y-z|^3}{\sigma^6}\phi_\sigma(y-z)d\overline{F}(z)\Big\} \\
&\quad +\big\{|y|+(\log n)^{\frac{2}{\tau}}\big\}(s_n\log p)^{\frac{1}{a_2}}\,3\big\{y^2+(\log n)^{\frac{4}{\tau}}\big\}(s_n\log p)^{\frac{2}{a_2}} \\
&\quad +2\big\{|y|+(\log n)^{\frac{2}{\tau}}\big\}(s_n\log p)^{\frac{1}{a_2}}\,5\big\{y^2+(\log n)^{\frac{4}{\tau}}\big\}(s_n\log p)^{\frac{2}{a_2}} \\
&\le 43\big\{|y|^3+(\log n)^{\frac{6}{\tau}}\big\}(s_n\log p)^{\frac{3}{a_2}}.
\end{aligned}
$$

Assume that a small $\zeta > 0$ is given. Let $A = \{y \in \mathbb{R} : |y| \le C_1 \left(\log(1/\epsilon_n)\right)^{\frac{1}{\tau}}\}$ for some large constant $C_1 > 0$. Note that

$$\begin{aligned}&\int_{A^c} \sup_{\eta \in \mathcal{H}_n^*} \left(\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y)\\ &\lesssim \int_{A^c} \sup_{\eta \in \mathcal{H}_n^*} \left(\dot{\ell}_\eta(y)\right)^2 dP_{\eta_0}(y) + \int_{A^c} \left(\dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y).\end{aligned}$$

It is easy to show that

$$\begin{aligned}&\int_{A^c} \sup_{\eta \in \mathcal{H}_n^*} \left(\dot{\ell}_\eta(y)\right)^2 dP_{\eta_0}(y)\\ &\lesssim \int_{y > C_1 (\log \frac{1}{\epsilon_n})^{\frac{1}{\tau}}} \left(y^2 + [\log n]^{\frac{4}{\tau}}\right) e^{-by^\tau} dy \cdot (s_n \log p)^{\frac{2}{a_2}}\\ &\lesssim (\epsilon_n)^{\frac{b}{2} C_1^\tau} \cdot (s_n \log p)^{\frac{2}{a_2}} (\log n)^{\frac{4}{\tau}} \ \lesssim \ \epsilon_n\end{aligned}$$

for some constant large $C_1 > 0$ by the assumption $(s_n \log p)^2 = o(n)$. Since

$$\begin{aligned}\int_{A^c} \left(\dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y) &\lesssim \int_{y > C_1 (\log \frac{1}{\epsilon_n})^{\frac{1}{\tau}}} (|y|^{\gamma_1} + 1) e^{-by^\tau} dy\\ &\lesssim \epsilon_n\end{aligned}$$

for some large constant $C_1 > 0$, we have

$$\int_{A^c} \sup_{\eta \in \mathcal{H}_n^*} \left(\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y) \lesssim \epsilon_n.$$

Thus, it suffices to prove

$$\int_{A} \sup_{\eta \in \mathcal{H}_n^*} \left(\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y) \le K_\zeta \, (\epsilon_n)^{\frac{4}{5} - \zeta} \, (s_n \log p)^{\frac{16}{5a_2}}$$

for some positive constants ζ and K_ζ not depending on (n, p).

Define for any x and $y \in \mathbb{R}$,

$$d_\eta(x, y) := \frac{\ell_\eta(y + x) - \ell_\eta(y)}{x} - \frac{\ell_{\eta_0}(y + x) - \ell_{\eta_0}(y)}{x},$$

then we have that

$$\begin{aligned}&\int_{A} \sup_{\eta \in \mathcal{H}_n^*} \left(\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y)\\ &\lesssim \int_{A} \sup_{\eta \in \mathcal{H}_n^*} \left(\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y) - d_\eta(x, y)\right)^2 dP_{\eta_0}(y) \qquad (5.34)\\ &+ \frac{1}{x^2} \int_{A} \sup_{\eta \in \mathcal{H}_n^*} \left(x \, d_\eta(x, y)\right)^2 dP_{\eta_0}(y). \qquad (5.35)\end{aligned}$$

One can obtain the upper bound for (5.34) using

$$\begin{aligned}|\dot\ell_\eta(y)-\dot\ell_{\eta_0}(y)-d_\eta(x,y)| &\le \left|\dot\ell_\eta(y)-\frac{\ell_\eta(y+x)-\ell_\eta(y)}{x}\right|\\ &\quad+\left|\dot\ell_{\eta_0}(y)-\frac{\ell_{\eta_0}(y+x)-\ell_{\eta_0}(y)}{x}\right|\\ &\le |x|\cdot\left\{|\ddot\ell_\eta(y_1)|+|\ddot\ell_{\eta_0}(y_2)|\right\}\\ &\lesssim |x|\cdot\left\{y^2+(\log n)^{\frac{4}{\tau}}\right\}(s_n\log p)^{\frac{2}{\alpha_2}}\\ &\lesssim |x|(s_n\log p)^{\frac{2}{\alpha_2}}(\log n)^{\frac{4}{\tau}}\end{aligned}$$

for any $\eta\in\mathcal{H}_n^*$, $y\in A$, small $|x|$ and some $|y-y_1|\vee|y-y_2|\le|x|$ by the Taylor expansion. Thus,

$$\int_A\sup_{\eta\in\mathcal{H}_n^*}\Big(\dot\ell_\eta(y)-\dot\ell_{\eta_0}(y)-d_\eta(x,y)\Big)^2dP_{\eta_0}(y)\lesssim x^2\cdot(s_n\log p)^{\frac{4}{\alpha_2}}\,[\log n]^{\frac{8}{\tau}}. \tag{5.36}$$

Note that $|x\,d_\eta(x,y)|\le|\ell_\eta(y+x)-\ell_{\eta_0}(y+x)|+|\ell_\eta(y)-\ell_{\eta_0}(y)|$ and

$$\begin{aligned}&\int_A\sup_{\eta\in\mathcal{H}_n^*}\left(\ell_\eta(y+x)-\ell_{\eta_0}(y+x)\right)^2dP_{\eta_0}(y)\\ &=\int_A\sup_{\eta\in\mathcal{H}_n^*}\left(\ell_\eta(y+x)-\ell_{\eta_0}(y+x)\right)^2\eta_0(y+x)\cdot\frac{\eta_0(y)}{\eta_0(y+x)}dy\\ &\lesssim\int_A\sup_{\eta\in\mathcal{H}_n^*}\left(\ell_\eta(y+x)-\ell_{\eta_0}(y+x)\right)^2\eta_0(y+x)\cdot e^{b'|y|^{\tau'}}dy\end{aligned}$$

provided that $|x|$ is small, by condition (D5). To calculate the upper bound for (5.35), we first find an upper bound for $f_\eta(y):=(\ell_\eta(y)-\ell_{\eta_0}(y))^2\eta_0(y)$ on $y\in A$ and $\eta\in\mathcal{H}_n^*$. Let $\delta_n:=\epsilon_n\log(1/\epsilon_n)$ and $B:=\left\{y\in\mathbb{R}:|y|\le 2C_1(\log(1/\delta_n))^{\frac{1}{\tau}}\right\}$, so that $A\subset B$ for all sufficiently large n. By the triangle inequality and the definition of $\mathcal{H}_n^*$,

$$\begin{aligned}|\dot f_\eta(y)|&=\left|2(\ell_\eta(y)-\ell_{\eta_0}(y))(\dot\ell_\eta(y)-\dot\ell_{\eta_0}(y))\eta_0(y)+(\ell_\eta(y)-\ell_{\eta_0}(y))^2\dot\eta_0(y)\right|\\ &\lesssim\sqrt{f_\eta(y)}\sqrt{\eta_0(y)}\Big(|\dot\ell_\eta(y)-\dot\ell_{\eta_0}(y)|+|\ell_\eta(y)-\ell_{\eta_0}(y)|\cdot|\dot\ell_{\eta_0}(y)|\Big)\\ &\lesssim\sqrt{f_\eta(y)}\,(s_n\log p)^{\frac{1}{\alpha_2}}\,(\log n)^{\frac{4}{\tau}},\end{aligned}\tag{5.37}$$

and

$$\begin{aligned}&|\ddot f_\eta(y)|\\ &\lesssim\eta_0(y)\bigg\{\Big(\dot\ell_\eta(y)-\dot\ell_{\eta_0}(y)\Big)^2+|\ddot\ell_\eta(y)-\ddot\ell_{\eta_0}(y)|\cdot|\ell_\eta(y)-\ell_{\eta_0}(y)|\\ &+|\ell_\eta(y)-\ell_{\eta_0}(y)|\cdot|\dot\ell_\eta(y)-\dot\ell_{\eta_0}(y)|\cdot|\dot\ell_\eta(y)|+(\ell_\eta(y)-\ell_{\eta_0}(y))^2\,|\ddot\ell_{\eta_0}(y)|\bigg\}\\ &\lesssim(s_n\log p)^{\frac{3}{\alpha_2}}\,(\log n)^{\frac{8}{\tau}}\end{aligned}\tag{5.38}$$

for any $\eta \in \mathcal{H}_n^*$ and $y \in \mathbb{R}$. By the Taylor expansion,

$$\begin{aligned}
&|f_\eta(y+x) - f_\eta(y)| \\
&\lesssim |x|\sqrt{f_\eta(y)}\,(s_n \log p)^{\frac{1}{a_2}}\,(\log n)^{\frac{4}{\tau}} + x^2\,(s_n \log p)^{\frac{3}{a_2}}\,(\log n)^{\frac{8}{\tau}} \\
&\lesssim (s_n \log p)^{\frac{1}{a_2}}\,[\log n]^{\frac{4}{\tau}} \left\{ |x|\sqrt{f_\eta(y)} + x^2\,(s_n \log p)^{\frac{2}{a_2}}\,(\log n)^{\frac{4}{\tau}} \right\}
\end{aligned}$$

for any $y \in \mathbb{R}$ and small $|x|$. If we take $|x| \le C\,(s_n \log p)^{-\frac{3}{2a_2}}\,(\log n)^{-\frac{4}{\tau}}\sqrt{f_\eta(y)}$ for some small constant $C > 0$, it implies $|f_\eta(y+x) - f_\eta(y)| \le f_\eta(y)/2$ for any $y \in \mathbb{R}$ and small $|x|$. Therefore, for any fixed $y_0 \in A$, we have $f_\eta(y_0 + x) > f_\eta(y_0)/2$ for any $|x| \le C\,(s_n \log p)^{-\frac{3}{2a_2}}\,(\log n)^{-\frac{4}{\tau}}\sqrt{f_\eta(y_0)}$ for some small constant $C > 0$. Then,

$$\begin{aligned}
\int_B f_\eta(y)dy &\ge \int_{|y-y_0| \le C(s_n \log p)^{-\frac{3}{2a_2}}[\log n]^{-\frac{4}{\tau}}\sqrt{f_\eta(y_0)}} f_\eta(y)dy \\
&\gtrsim (s_n \log p)^{-\frac{3}{2a_2}}\,(\log n)^{-\frac{4}{\tau}}\,(f_\eta(y_0))^{\frac{3}{2}} \qquad (5.39)
\end{aligned}$$

for any $y_0 \in A$ and $\eta \in \mathcal{H}_n^*$. On the other hand,

$$1/\eta(y) \lesssim (\log n)^{\frac{1}{2}} \exp\{2(s_n \log p)^{\frac{1}{a_2}}(\log n)^{\frac{4}{\tau}}\}$$

for any $y \in B$ and $\eta \in \mathcal{H}_n^*$, which implies

$$\begin{aligned}
\int_B \left\{\frac{\eta_0(y)}{\eta(y)}\right\}^{\delta} \eta_0(y)dy &\lesssim \int_B \eta_0(y)^{1+\delta}(\log n)^{\frac{\delta}{2}} \exp\{2\delta(s_n \log p)^{\frac{1}{a_2}}(\log n)^{\frac{4}{\tau}}\}dy \\
&\lesssim 1
\end{aligned}$$

by taking $\delta = (s_n \log p)^{-\frac{1}{a_2}}(\log n)^{-\frac{4}{\tau}}$. Thus, by Lemma 6, we have

$$\int_B f_\eta(y)dy \lesssim \delta_n^2\,(s_n \log p)^{\frac{2}{a_2}}\,[\log n]^{\frac{12}{\tau}} \qquad (5.40)$$

for any $\eta \in \mathcal{H}_n^*$. By combining (5.39) and (5.40), it implies that

$$f_\eta(y_0) \lesssim \delta_n^{\frac{4}{3}}\,(s_n \log p)^{\frac{7}{3a_2}}\,[\log n]^{\frac{32}{3\tau}} \qquad (5.41)$$

for any $y_0 \in A$ and $\eta \in \mathcal{H}_n^*$.

Next, we claim that if $f_\eta(y) \lesssim \delta_n^{d_1}\,(s_n \log p)^{d_2}\,[\log n]^{d_3}$ for some d_1, d_2 and $d_3 > 0$, then we have $f_\eta(y) \lesssim \delta_n^{1+\frac{3}{8}d_1 - \zeta}\,(s_n \log p)^{\frac{3}{8}d_2 + \frac{3}{2a_2}}\,[\log n]^{\frac{3}{8}d_3 + \frac{7}{\tau}}$ for any $y \in A$ and $\eta \in \mathcal{H}_n^*$. Suppose that $f_\eta(y) \lesssim \delta_n^{d_1}\,(s_n \log p)^{d_2}\,[\log n]^{d_3}$ on $y \in A$ and $\eta \in \mathcal{H}_n^*$ for some positive constants d_1, d_2 and d_3. Due to (5.41), there exist constants $d_1 = 4/3, d_2 = 7/(3a_2)$ and $d_3 = 32/(3\tau)$ satisfying $f_\eta(y) \lesssim \delta_n^{d_1} (s_n \log p)^{d_2} [\log n]^{d_3}$. Note that for any small constant $\zeta > 0$,

$$\begin{aligned} &|\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)|\sqrt{\eta_0(y)} \\ &\lesssim |x| \left(|\ddot{\ell}_\eta(y_1)| + |\ddot{\ell}_{\eta_0}(y_2)|\right) \sqrt{\eta_0(y)} \\ &+ \frac{|\ell_\eta(y+x) - \ell_{\eta_0}(y+x)| + |\ell_\eta(y) - \ell_{\eta_0}(y)|}{|x|} \sqrt{\eta_0(y)} \\ &\lesssim |x| (s_n \log p)^{\frac{2}{a_2}} [\log n]^{\frac{4}{\tau}} + \frac{e^{\frac{b'}{2}|y|^{\tau'}}}{|x|} \cdot \delta_n^{\frac{d_1}{2}} (s_n \log p)^{\frac{d_2}{2}} [\log n]^{\frac{d_3}{2}} \\ &\lesssim |x| (s_n \log p)^{\frac{2}{a_2}} [\log n]^{\frac{4}{\tau}} + \frac{1}{|x|} \delta_n^{\frac{d_1}{2} - 4\zeta} (s_n \log p)^{\frac{d_2}{2}} [\log n]^{\frac{d_3}{2}} \end{aligned} \tag{5.42}$$

for some $|y - y_1| \vee |y - y_2| \le |x|$, thus

$$|\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)|\sqrt{\eta_0(y)} \lesssim \delta_n^{\frac{d_1}{4} - 2\zeta} (s_n \log p)^{\frac{d_2}{4} + \frac{1}{a_2}} [\log n]^{\frac{d_3}{4} + \frac{2}{\tau}} \tag{5.43}$$

on $y \in A$ and $\eta \in \mathcal{H}_n^*$, by taking $|x| = \delta_n^{\frac{d_1}{4} - 2\zeta} (s_n \log p)^{\frac{d_2}{4} - \frac{1}{a_2}} [\log n]^{\frac{d_3}{4} - \frac{2}{\tau}}$. Then, by (5.37),

$$|\dot{f}_\eta(y)| \lesssim \delta_n^{\frac{3}{4}d_1 - 2\zeta} (s_n \log p)^{\frac{3}{4}d_2 + \frac{1}{a_2}} [\log n]^{\frac{3}{4}d_3 + \frac{2}{\tau}}$$

for any $y \in A$ and $\eta \in \mathcal{H}_n^*$, which implies that

$$f_\eta(y + x) \ge \frac{1}{2} f_\eta(y)$$

for any $y \in A, \eta \in \mathcal{H}_n^*$, $|x| \le C_3 \delta_n^{-\frac{3}{4}d_1 + 2\zeta} (s_n \log p)^{-\frac{3}{4}d_2 - \frac{1}{a_2}} [\log n]^{-\frac{3}{4}d_3 - \frac{2}{\tau}} f_\eta(y)$ and for some small constant $C_3 > 0$, by the first-order Taylor expansion. Thus, similar to (5.39),

$$\int_B f_\eta(y)dy \gtrsim (f_\eta(y_0))^2 \delta_n^{-\frac{3}{4}d_1 + 2\zeta} (s_n \log p)^{-\frac{3}{4}d_2 - \frac{1}{a_2}} [\log n]^{-\frac{3}{4}d_3 - \frac{2}{\tau}},$$

for any $y_0 \in A, \eta \in \mathcal{H}_n^*$ and small $\zeta > 0$. Again by (5.40),

$$f_\eta(y) \lesssim \delta_n^{1 + \frac{3}{8}d_1 - \zeta} (s_n \log p)^{\frac{3}{8}d_2 + \frac{3}{2a_2}} [\log n]^{\frac{3}{8}d_3 + \frac{7}{\tau}}, \tag{5.44}$$

for any $y \in A, \eta \in \mathcal{H}_n^*$ and small $\zeta > 0$.

Note that the upper bound (5.44) is obtained from the assumption $\sup_{\eta \in \mathcal{H}_n^*} f_\eta(y) \lesssim \delta_n^{d_1} (s_n \log p)^{d_2} [\log n]^{d_3}$. Thus, by applying the claim repeatedly, one can check that $\sup_{\eta \in \mathcal{H}_n^*} f_\eta(y) \lesssim \delta_n^{\frac{8}{5} - 2\zeta} (s_n \log p)^{\frac{12}{5a_2}} [\log n]^{\frac{56}{5\tau}}$ for any $y \in A$ and a given small constant $\zeta > 0$.

Therefore, we finally obtain the following upper bound

$$\begin{aligned}
&\int_A \sup_{\eta\in\mathcal{H}_n^*} \left(\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y) \\
&\lesssim x^2 \cdot (s_n \log p)^{\frac{4}{a_2}} [\log n]^{\frac{8}{\tau}} \\
&\quad + \frac{1}{x^2} \int_A \sup_{\eta\in\mathcal{H}_n^*} \left(\ell_\eta(y+x) - \ell_{\eta_0}(y+x)\right)^2 \eta_0(y+x) \cdot e^{b'|y|^{\tau'}} dy \\
&\lesssim x^2 \cdot (s_n \log p)^{\frac{4}{a_2}} [\log n]^{\frac{8}{\tau}} + \frac{\delta_n^{\frac{8}{5}-2\zeta} (s_n \log p)^{\frac{12}{5a_2}} [\log n]^{\frac{56}{5\tau}}}{x^2}
\end{aligned}$$

by (5.36). By taking $|x| = \delta_n^{\frac{2}{5}-\frac{\zeta}{2}} (s_n \log p)^{-\frac{2}{5a_2}} [\log n]^{\frac{4}{5\tau}}$,

$$\begin{aligned}
\int_A \sup_{\eta\in\mathcal{H}_n^*} \left(\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y) &\leq K_\zeta \delta_n^{\frac{4}{5}-\zeta} (s_n \log p)^{\frac{16}{5a_2}} [\log n]^{\frac{48}{5\tau}} \\
&\leq K_\zeta \epsilon_n^{\frac{4}{5}-2\zeta} (s_n \log p)^{\frac{16}{5a_2}}
\end{aligned}$$

for some constant $K_\zeta > 0$ not depending on (n, p). ■

This lemma gives slightly faster convergence rate, under stronger condition, compared with Lemma 7, and is used to prove the misspecified LAN (Lemma 5). Although Lemma 8 seems similar to Lemma 7, we stated them separately to avoid assuming redundant conditions for Lemma 7.

Lemma 8 *Let $\epsilon_n = K_{\text{eta}}\sqrt{s_n \log p/n}$. For any constant $\zeta > 0$, there exists a constant $K_\zeta > 0$ not depending on (n, p) such that*

$$\int \sup_{\eta\in\mathcal{H}_n^*} \left(\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y) \leq K_\zeta (\epsilon_n)^{\frac{4}{5}-\zeta}$$

for any η_0 satisfying (D1)-(D5) and all sufficiently large n, provided that $(s_n \log p)^{1+\frac{15}{a_2}} = o(n^{1-\zeta})$, where $\mathcal{H}_n^*$ defined at (3.17).

Proof Assume that a small $\zeta > 0$ is given. Let $\varphi_n := \epsilon_n^{\frac{4}{5}-\zeta} (s_n \log p)^{\frac{6}{5a_2}} [\log n]^{\frac{4}{\tau}}$, $A' := \{y \in A : \eta_0(y) \gtrsim \varphi_n^2\}$ and $B' := \{y \in B : \eta_0(y) \gtrsim \varphi_n^2\}$, where A and B are defined in Lemma 7. Note that

$$\begin{aligned}
\int_{(A')^c} \sup_{\eta\in\mathcal{H}_n^*} \left(\dot{\ell}_\eta(y)\right)^2 dP_{\eta_0}(y) &\lesssim \int_{A^c} \sup_{\eta\in\mathcal{H}_n^*} \left(\dot{\ell}_\eta(y)\right)^2 dP_{\eta_0}(y) \\
&\quad + \int_{A\cap\{y:\,\eta_0(y)\lesssim\varphi_n^2\}} \sup_{\eta\in\mathcal{H}_n^*} \left(\dot{\ell}_\eta(y)\right)^2 dP_{\eta_0}(y) \\
&\lesssim \epsilon_n + \varphi_n^2 (s_n \log p)^{\frac{2}{a_2}} [\log n]^{\frac{4}{\tau}} \cdot \int_A (y^2+1) dy \\
&\lesssim \epsilon_n + \varphi_n^2 (s_n \log p)^{\frac{2}{a_2}} [\log n]^{\frac{7}{\tau}} \\
&\lesssim \epsilon_n^{\frac{4}{5}-\zeta},
\end{aligned}$$

provided that $(s_n \log p)^{1+\frac{11}{a_2}} = o(n)$. Similarly, it is easy to check that

$$\int_{(A')^c} \sup_{\eta \in \mathcal{H}_n^*} \left(\dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y) \lesssim \epsilon_n^{4/5-\zeta}.$$

Hence, it suffices to show that

$$\int_{A'} \sup_{\eta \in \mathcal{H}_n^*} \left(\dot{\ell}_{\eta}(y) - \dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y) \leq K_\zeta \left(\epsilon_n\right)^{\frac{4}{5}-\zeta}$$

for some positive constants ζ and K_ζ.

Note that similar to (5.42),

$$\begin{aligned} &|\ddot{\ell}_\eta(y) - \ddot{\ell}_{\eta_0}(y)|\sqrt{\eta_0(y)} \\ &\lesssim |x| \{|\dddot{\ell}_\eta(y_1)| + |\dddot{\ell}_{\eta_0}(y_2)|\} \sqrt{\eta_0(y)} + \frac{e^{\frac{b'}{2}|y|^{\tau'}}}{|x|} |\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)|\sqrt{\eta_0(y)} \quad (5.45) \\ &\lesssim |x| \left(s_n \log p\right)^{\frac{3}{a_2}} [\log n]^{\frac{6}{\tau}} + \frac{1}{|x|} \delta_n^{\frac{2}{5}-\zeta} \left(s_n \log p\right)^{\frac{8}{5a_2}} [\log n]^{\frac{4}{\tau}} \end{aligned}$$

for some $|y - y_1| \vee |y - y_2| \leq |x|$ on $y \in A'$ and $\eta \in \mathcal{H}_n^*$ by (5.43). Then, by taking appropriate $|x|$, we have

$$\begin{aligned} |\ddot{\ell}_\eta(y) - \ddot{\ell}_{\eta_0}(y)|\sqrt{\eta_0(y)} &\lesssim \delta_n^{\frac{1}{5}-\zeta} \left(s_n \log p\right)^{\frac{23}{10a_2}} [\log n]^{\frac{5}{\tau}} \\ &\lesssim \left(s_n \log p\right)^{\frac{4}{5a_2}} \end{aligned} \quad (5.46)$$

on $y \in A'$ and $\eta \in \mathcal{H}_n^*$, because we assume that $(s_n \log p)^{1+\frac{15}{a_2}} = o(n^{1-\zeta})$. Suppose that $\sup_{\eta \in \mathcal{H}_n^*} |\ddot{\ell}_\eta(y) - \ddot{\ell}_{0\eta}(y)|\sqrt{\eta_0(y)} \lesssim (s_n \log p)^K$ and $\sup_{\eta \in \mathcal{H}_n^*} f_\eta(y) \lesssim \delta_n^{d_1} \left(s_n \log p\right)^{d_2} [\log n]^{d_3}$ on $y \in B'$ for some positive constants K, d_1, d_2 and d_3. Note that from the proof of Lemma 7 and the definition of B',

$$\frac{\eta_0(y)}{\eta(y)} \lesssim \exp\left(\frac{\varphi_n}{\sqrt{\eta_0(y)}}\right) \lesssim 1$$

for any $y \in B'$ and $\eta \in \mathcal{H}_n^*$, then, similar to (5.40), it is easy to show that

$$\int_{B'} f_\eta(y) dy \lesssim \delta_n^2, \quad (5.47)$$

by Lemma 6. Applying (5.42),

$$|\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)|\sqrt{\eta_0(y)} \lesssim \delta_n^{\frac{d_1}{4}-\zeta} \left(s_n \log p\right)^{\frac{d_2}{4}+\frac{K}{2}} [\log n]^{\frac{d_3}{4}} \quad (5.48)$$

for any $y \in A'$ and $\eta \in \mathcal{H}_n^*$. Then by (5.47) and the similar arguments to the proof of Lemma 7, we have

$$f_\eta(y) \lesssim \delta_n^{1+\frac{3}{8}d_1-\zeta} \left(s_n \log p\right)^{\frac{3}{8}d_2+\frac{K}{4}} [\log n]^{\frac{3}{8}d_3}$$

for any $y \in A'$ and $\eta \in \mathcal{H}_n^*$. By a recursion, one can check that d_1, d_2 and d_3 converge to $8/5 - \zeta, 2K/5$ and 0, respectively. Thus, by (5.48), we have

$$|\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)|\sqrt{\eta_0(y)} \lesssim \delta_n^{\frac{2}{5}-\zeta} (s_n \log p)^{\frac{3}{5}K} \tag{5.49}$$

for any $y \in A'$ and $\eta \in \mathcal{H}_n^*$, and it implies that

$$\begin{aligned}|\ddot{\ell}_\eta(y) - \ddot{\ell}_{0\eta}(y)|\sqrt{\eta_0(y)} &\lesssim \delta_n^{\frac{1}{5}-\zeta} (s_n \log p)^{\frac{3}{2a_2}+\frac{3}{10}K} [\log n]^{\frac{5}{\tau}} \\ &\lesssim (s_n \log p)^{\frac{3}{10}K}\end{aligned}$$

for any $y \in A'$ and $\eta \in \mathcal{H}_n^*$ by (5.45). Thus, we obtain $\sup_{\eta\in\mathcal{H}_n^*} |\ddot{\ell}_\eta(y) - \ddot{\ell}_{0\eta}(y)|\sqrt{\eta_0(y)} \lesssim (s_n \log p)^{\frac{3}{10}K}$ from the assumption $\sup_{\eta\in\mathcal{H}_n^*} |\ddot{\ell}_\eta(y) - \ddot{\ell}_{0\eta}(y)| \times\sqrt{\eta_0(y)} \lesssim (s_n \log p)^K$ on $y \in A'$. Suppose that a small constant $\zeta' > 0$ is given, then we have $\sup_{\eta\in\mathcal{H}_n^*} |\ddot{\ell}_\eta(y)| \lesssim (s_n \log p)^{\zeta'}$ on $y \in B'$ by repeatedly applying the above arguments. Finally, by (5.49),

$$\left(\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)\right)^2 \eta_0(y) \lesssim \delta_n^{\frac{4}{5}-\zeta}$$

for some given constant $\zeta > 0$, any $y \in A'$ and $\eta \in \mathcal{H}_n^*$. Therefore,

$$\int_{A'} \sup_{\eta\in\mathcal{H}_n^*} \left(\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y) \le K_\zeta (\epsilon_n)^{\frac{4}{5}-\zeta}$$

for some positive constants ζ and K_ζ not depending on (n, p). ∎

Lemma 9 *If* $(s_n \log p)^{1+\frac{11}{2a_2}} = o(n^{1-\zeta})$ *for some constant* $\zeta > 0$*, we have*

$$\sup_{\eta\in\mathcal{H}_n^*} \int \left(\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y) = o(1)$$

for any η_0 *satisfying (D1)-(D5), where* $\mathcal{H}_n^*$ *defined at* (3.17).

Proof Note that

$$\begin{aligned}\int \left(\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)\right)^2 dP_{\eta_0}(y) = &-\int (\ell_\eta(y) - \ell_{\eta_0}(y))(\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y))\dot{\eta}_0(y)dy \\ &-\int (\ell_\eta(y) - \ell_{\eta_0}(y))(\ddot{\ell}_\eta(y) - \ddot{\ell}_{\eta_0}(y))\eta_0(y)dy\end{aligned}$$

follows from the integration by parts. By Lemma 6, (5.43) and (5.46), one can show that the absolute value of the above equality is bounded above by $\epsilon_n^{\frac{6}{5}-\zeta}(s_n \log p)^{\frac{33}{10a_2}}$ for some constant $\zeta > 0$, up to some constant not depending on η, which implies the desired result. ∎

The following lemma is used to prove Lemma 10.

Lemma 10 *Let s_n be a sequence of positive integers. Define*

$$\Theta_{n,1} := \{\theta \in \mathbb{R}^p : s_\theta \le s_n, \|\theta - \theta_0\|_1 \le 1\}$$

and $f_{\theta,\bar\theta,\eta} := (\theta-\theta_0)^T \ddot\ell_{\bar\theta,\eta}(\theta-\theta_0)$. If we assume $(s_n \log p)^{1+\frac{15}{\alpha_2}} = o(n^{1-\zeta})$ for some constant $\zeta > 0$, then for any small constant $\zeta' > 0$,

$$\mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\theta,\bar\theta\in\Theta_{n,1}} \sup_{\eta\in\mathcal{H}_n^*} \frac{1}{\sqrt{n}}\left|\mathbb{G}_n f_{\theta,\bar\theta,\eta}\right|\right) \lesssim \left(\frac{s_n(\log p)^3 + (s_n\log p)^{\frac{3}{\alpha_2}}(\log p)^4}{n}\right)^{\frac{1}{2}} (s_n \log p)^{\zeta'} \tag{5.50}$$

for any η_0 satisfying (D1)-(D5) and all sufficiently large n, where $\mathcal{H}_n^*$ defined at (3.17).

Proof Without loss of generality, we assume that $\theta_0 = 0$. For a given $\zeta' > 0$, define

$$\widetilde{\mathcal{F}}_n := \left\{\widetilde f_{\theta,\bar\theta,\eta} = (s_n\log p)^{-\zeta'}(\log p)^{-1}\cdot f_{\theta,\bar\theta,\eta} : \theta,\bar\theta\in\Theta_{n,1}, \eta\in\mathcal{H}_n^*\right\}. \tag{5.51}$$

Then for any $\widetilde f_{\theta,\bar\theta,\eta} \in \widetilde{\mathcal{F}}_n$,

$$\begin{aligned}&|\widetilde f_{\theta,\bar\theta,\eta}(x,y)|\\ &\le \sup_{\theta,\bar\theta\in\Theta_{n,1}}\sup_{\eta\in\mathcal{H}_n^*}(x^T\theta)^2|\ddot\ell_\eta(y-x^T\bar\theta)|(s_n\log p)^{-\zeta'}(\log p)^{-1} =: \widetilde F_n(x,y).\end{aligned}$$

$\widetilde F_n$ is an envelop function of $\widetilde{\mathcal{F}}_n$ such that $\mathbb{E}_{\theta_0,\eta_0}\widetilde F_n^2(x_i,Y_i) \lesssim 1$ for any $i = 1,\dots,n$ because

$$\begin{aligned}&\mathbb{E}_{\theta_0,\eta_0}\widetilde F_n^2(x,Y)\\ &= \int \sup_{\theta,\bar\theta\in\Theta_{n,1}}\sup_{\eta\in\mathcal{H}_n^*}(x^T\theta)^4|\ddot\ell_\eta(y-x^T\bar\theta)|^2\eta_0(y)dy\cdot(s_n\log p)^{-2\zeta'}(\log p)^{-2}\\ &\lesssim \int_{A'}\sup_{\bar\theta\in\Theta_{n,1}}\sup_{\eta\in\mathcal{H}_n^*}|\ddot\ell_\eta(y-x^T\bar\theta)|^2\eta_0(y)dy\cdot(s_n\log p)^{-2\zeta'}\\ &+ \int_{(A')^c}\sup_{\bar\theta\in\Theta_{n,1}}\sup_{\eta\in\mathcal{H}_n^*}|\ddot\ell_\eta(y-x^T\bar\theta)|^2\eta_0(y)dy\cdot(s_n\log p)^{-2\zeta'}\\ &\lesssim (s_n\log p)^{-2\zeta'} + \int_{A^c}\sup_{\bar\theta\in\Theta_{n,1}}\sup_{\eta\in\mathcal{H}_n^*}|\ddot\ell_\eta(y-x^T\bar\theta)|^2\eta_0(y)dy\cdot(s_n\log p)^{-2\zeta'}\\ &+ \int_{A\cap\{y:\eta_0(y)\lesssim\varphi_n^2\}}\sup_{\bar\theta\in\Theta_{n,1}}\sup_{\eta\in\mathcal{H}_n^*}|\ddot\ell_\eta(y-x^T\bar\theta)|^2\eta_0(y)dy\cdot(s_n\log p)^{-2\zeta'}\\ &\lesssim (s_n\log p)^{-2\zeta'} + (s_n\log p)^{\frac{4}{\alpha_2}}\varphi_n^2(s_n\log p)^{-2\zeta'} \lesssim (s_n\log p)^{-2\zeta'}\end{aligned}$$

provided that $(s_n \log p)^{1+\frac{15}{a_2}} = o(n)$, where A, A' and φ_n are defined in the proof of Lemma 8. Thus, $\|\widetilde{F}_n\|_n^2 = n^{-1}\sum_{i=1}^n \mathbb{E}_{\theta_0,\eta_0}\widetilde{F}_n^2(x_i, Y_i) \lesssim (s_n \log p)^{-2\zeta'}$. We will use Corollary A.1 in **?**, which implies

$$\begin{aligned}
&\mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\theta,\bar{\theta}\in\Theta_{n,1}}\sup_{\eta\in\mathcal{H}_n^*}\frac{1}{\sqrt{n}}\left|\mathbb{G}_n f_{\theta,\bar{\theta},\eta}\right|\right)\\
&\lesssim \int_0^{\|\widetilde{F}_n\|_n}\sqrt{\log N_{[]}^n(\epsilon,\widetilde{\mathcal{F}}_n)}d\epsilon\cdot\frac{(s_n\log p)^{\zeta'}}{\sqrt{n}}\log p. \qquad (5.52)
\end{aligned}$$

Now, we calculate $N_{[]}^n(\epsilon,\widetilde{\mathcal{F}}_n)$ defined at (5.52). For $\theta^j,\bar{\theta}^j\in\Theta_{n,1}$ and $\eta_j\in\mathcal{H}_n^*, j=1,2$, write

$$\widetilde{f}_{\theta^1,\bar{\theta}^1,\eta_1} - \widetilde{f}_{\theta^2,\bar{\theta}^2,\eta_2} \equiv \widetilde{f}_1 + \widetilde{f}_2 + \widetilde{f}_3,$$

where $\widetilde{f}_1 := \widetilde{f}_{\theta^1,\bar{\theta}^1,\eta_1} - \widetilde{f}_{\theta^2,\bar{\theta}^1,\eta_1}$, $\widetilde{f}_2 := \widetilde{f}_{\theta^2,\bar{\theta}^1,\eta_1} - \widetilde{f}_{\theta^2,\bar{\theta}^2,\eta_1}$ and $\widetilde{f}_3 := \widetilde{f}_{\theta^2,\bar{\theta}^2,\eta_1} - \widetilde{f}_{\theta^2,\bar{\theta}^2,\eta_2}$. It is easy to show $|\widetilde{f}_1(x,y)| \lesssim \|\theta^1-\theta^2\|_1\cdot\left(y^2+1\right)(s_n\log p)^{\frac{2}{a_2}}[\log n]^{\frac{4}{\tau}}$ and $|\widetilde{f}_2(x,y)| \lesssim \|\bar{\theta}^1-\bar{\theta}^2\|_1\cdot\left(|y|^3+1\right)(s_n\log p)^{\frac{3}{a_2}}[\log n]^{\frac{6}{\tau}}\sqrt{\log p}$. Then, we have

$$\begin{aligned}
&\mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\theta^1,\theta^2}\sup_{\eta_1,\eta_2}|\widetilde{f}_{\theta^1,\bar{\theta}^1,\eta_1}(x,Y) - \widetilde{f}_{\theta^2,\bar{\theta}^2,\eta_2}(x,Y)|^2\right)\\
&\lesssim \sup_{\theta^1,\theta^2}\|\theta^1-\theta^2\|_1^2(s_n\log p)^{\frac{6}{a_2}}[\log n]^{\frac{12}{\tau}}\log p + \mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\theta^1,\theta^2}\sup_{\eta_1,\eta_2}|\widetilde{f}_3(x,Y)|^2\right).
\end{aligned}$$

To deal with $\widetilde{f}_3$, define

$$\widetilde{\mathcal{G}}_{K_n} := \left\{\ddot{\ell}_\eta\cdot I_{[-K_n,K_n]} : \eta\in\mathcal{H}_n^*\right\}$$

and $\widetilde{H}_{K_n} := \sup_{\eta\in\mathcal{H}_n^*}\max_{k=0,1}\sup_{|y|\le K_n}|\ddot{\ell}_\eta^{(k)}(y)|$ for some $K_n > 0$. Then, Theorem 2.7.1 of **?**, which implies for every $\epsilon > 0$,

$$\begin{aligned}
\log N(\epsilon) &:= \log N(\epsilon,\widetilde{\mathcal{G}}_{K_n},\|\cdot\|_\infty)\\
&\lesssim K_n\cdot\widetilde{H}_{K_n}\cdot\frac{1}{\epsilon}\\
&\lesssim K_n\cdot K_n^3(s_n\log p)^{\frac{3}{a_2}}(\log n)^{\frac{6}{\tau}}\frac{1}{\epsilon}.
\end{aligned}$$

By the definition of the covering number, there is a partition $\{\mathcal{H}^l : 1\le l\le N(\epsilon)\}$ of $\mathcal{H}_n^*$ such that

$$\begin{aligned}
&\int_{|y|\le K_n - M\sqrt{\log p}}\sup_{\theta\in\Theta_{n,1}}\sup_{\eta_1,\eta_2\in\mathcal{H}^l}|\ddot{\ell}_{\eta_1}(y-x^T\theta) - \ddot{\ell}_{\eta_2}(y-x^T\theta)|^2 dP_{\eta_0}(y)\\
&\lesssim \int_{|y|\le K_n - M\sqrt{\log p}}\epsilon^2 dP_{\eta_0}(y) \;\le\; \epsilon^2.
\end{aligned}$$

Let $K_n = C(\log(1/\epsilon))^{1/\tau} + C(\log n)^{1/\tau} + M\sqrt{\log p}$ for some constant $C > 0$, then

$$\begin{aligned}
&\int_{|y|>K_n - M\sqrt{\log p}} \sup_{\theta\in\Theta_{n,1}} \sup_{\eta\in\mathcal{H}_n^*} |\ddot{\ell}_\eta(y - x^T\theta)|^2 dP_{\eta_0}(y) \\
&\lesssim \int_{|y|>K_n - M\sqrt{\log p}} y^4 e^{-b|y|^\tau} dy \cdot (s_n \log p)^{\frac{4}{a_2}} [\log n]^{\frac{8}{\tau}} \\
&\lesssim e^{-\frac{b}{4}K_n^\tau} \cdot (s_n \log p)^{\frac{4}{a_2}} [\log n]^{\frac{8}{\tau}} \ \le \ \epsilon^2.
\end{aligned}$$

Thus, we have

$$\int \sup_{\theta^2,\bar{\theta}^2\in\Theta_{n,1}} \sup_{\eta_1,\eta_2\in\mathcal{H}^l} |\widetilde{f}_3(x,y)|^2 dP_{\eta_0}(y) \lesssim \epsilon^2,$$

for some constant $C > 0$ and any $1 \le l \le N(\epsilon)$.

By the above arguments,

$$\begin{aligned}
\log N^n_{[\,]}(\epsilon, \widetilde{\mathcal{F}}_n) &\lesssim \log N(\epsilon) + \log N\left(\epsilon\, (s_n \log p)^{-\frac{3}{a_2}} [\log n]^{-\frac{6}{\tau}} [\log p]^{-\frac{1}{2}}, \Theta_n, \|\cdot\|_1\right) \\
&\lesssim K_n^4 \,(s_n \log p)^{\frac{3}{a_2}} (\log n)^{\frac{6}{\tau}} \cdot \frac{1}{\epsilon} + s_n \log p + s_n \log \frac{1}{\epsilon} \\
&\lesssim \epsilon^{-\frac{3}{2}} \cdot (s_n \log p)^{\frac{3}{a_2}} (\log n)^{\frac{6}{\tau}} (\log p)^2 + s_n \log p + s_n \log \frac{1}{\epsilon}.
\end{aligned}$$

Hence, by (5.52), we get the inequality (5.50). ■

The following lemma is used to prove Lemma 5.

Lemma 11 (Misspecified LAN: version 2) *Let s_n be a positive integer sequence and ϵ_n be a sequence such that $\epsilon_n \to 0$. Define $\Theta_{n,\epsilon_n} := \{\theta \in \Theta : s_\theta \le s_n, \|\theta-\theta_0\|_1 \le \epsilon_n\}$ and $\tilde{r}_n(\theta,\eta) := L_n(\theta,\eta) - L_n(\theta_0,\eta_0) - \sqrt{n}(\theta-\theta_0)^T \mathbb{G}_n \dot{\ell}_{\theta_0,\eta} + n(\theta-\theta_0)^T V_{n,\eta}(\theta-\theta_0)/2$. If we assume that $(s_n \log p)^{1+\frac{15}{a_2}} = o(n^{1-\zeta})$ for some constant $\zeta > 0$, then*

$$\begin{aligned}
&\mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\theta\in\Theta_{n,\epsilon_n}} \sup_{\eta\in\mathcal{H}_n^*} |\tilde{r}_n(\theta,\eta)|\right) \\
&\lesssim n\epsilon_n^2 \cdot \rho_n + \epsilon_n \sqrt{\log p} \cdot \sup_{\theta\in\Theta_{n,\epsilon_n}} \|X(\theta-\theta_0)\|_2^2, \qquad (5.53)
\end{aligned}$$

for any η_0 satisfying (D1)-(D5) and all sufficiently large n, where $\mathcal{H}_n^*$ defined at (3.17) and

$$\rho_n := \left(\frac{s_n (\log p)^3 + (s_n \log p)^{\frac{3}{a_2}} (\log p)^4}{n}\right)^{\frac{1}{2}} (s_n \log p)^{\zeta'}$$

for a given constant $\zeta' > 0$.

Proof By the Taylor expansion, where $\theta(t) := \theta_0 + t(\theta - \theta_0)$,

$$\begin{aligned} L_n(\theta, \eta) &= L_n(\theta(1), \eta) \\ &= L_n(\theta_0, \eta) + \frac{\partial}{\partial t} L_n(\theta(t), \eta)\big|_{t=0} + \int_0^1 \frac{\partial^2}{\partial t^2} L_n(\theta(t), \eta)(1-t)dt. \end{aligned}$$

Since $\mathbb{E}_{\theta_0,\eta_0} \dot{\ell}_{\theta_0,\eta} = 0$ for every η by (D4), we have that

$$\frac{\partial}{\partial t} L_n(\theta(t), \eta)\big|_{t=0} = \sqrt{n}(\theta - \theta_0)^T \mathbb{G}_n \dot{\ell}_{\theta_0,\eta}$$

and

$$\frac{\partial^2}{\partial t^2} L_n(\theta(t), \eta) = n(\theta - \theta_0)^T \mathbb{P}_n \ddot{\ell}_{\theta(t),\eta}(\theta - \theta_0).$$

Define

$$\begin{aligned} A_{n1}(\theta, \eta) &:= n \int_0^1 (1-t) \frac{1}{\sqrt{n}} \mathbb{G}_n (\theta - \theta_0)^T \ddot{\ell}_{\theta(t),\eta}(\theta - \theta_0) dt, \\ A_{n2}(\theta, \eta) &:= \int_0^1 (1-t) \\ &\qquad \sum_{i=1}^n \Big[(\theta - \theta_0)^T \mathbb{E}_{\theta_0,\eta_0} \Big\{ \ddot{\ell}_{\theta(t),\eta}(x_i, Y_i) - \ddot{\ell}_{\theta_0,\eta}(x_i, Y_i) \Big\} (\theta - \theta_0) \Big] dt, \\ A_{n3}(\theta, \eta) &:= \frac{1}{2} \sum_{i=1}^n (\theta - \theta_0)^T \mathbb{E}_{\theta_0,\eta_0} \ddot{\ell}_{\theta_0,\eta}(x_i, Y_i)(\theta - \theta_0), \end{aligned}$$

then, it is easy to show that

$$\int_0^1 \frac{\partial^2}{\partial t^2} L_n(\theta(t), \eta)(1-t)dt = A_{n1}(\theta, \eta) + A_{n2}(\theta, \eta) + A_{n3}(\theta, \eta).$$

Since

$$\frac{1}{\sqrt{n}} \mathbb{G}_n (\theta - \theta_0)^T \ddot{\ell}_{\theta(t),\eta}(\theta - \theta_0) = \frac{\|\theta - \theta_0\|_1^2}{\sqrt{n}} \mathbb{G}_n \frac{(\theta - \theta_0)^T}{\|\theta - \theta_0\|_1} \ddot{\ell}_{\theta(t),\eta} \frac{(\theta - \theta_0)}{\|\theta - \theta_0\|_1},$$

we have

$$\mathbb{E}_{\theta_0,\eta_0} \left(\sup_{\theta \in \Theta_{n,\epsilon_n}} \sup_{\eta \in \mathcal{H}_n^*} \Big| A_{n1}(\theta, \eta) \Big| \right) \lesssim n\epsilon_n^2 \cdot \rho_n$$

by (5.50) in Lemma 10, provided that $(s_n \log p)^{1+\frac{15}{a_2}} = o(n^{1-\zeta})$ for some $\zeta > 0$. Since $A_{n3}(\theta, \eta) = -n/2 \cdot (\theta - \theta_0)^T V_{n,\eta}(\theta - \theta_0)$, if we only need to show that

$$\sup_{\theta \in \Theta_{n,\epsilon_n}} \sup_{\eta \in \mathcal{H}_n^*} \Big| A_{n2}(\theta, \eta) \Big| \lesssim \epsilon_n \sqrt{\log p} \cdot \sup_{\theta \in \Theta_{n,\epsilon_n}} \|X(\theta - \theta_0)\|_2^2,$$

where $\theta(t) := \theta_0 + t(\theta - \theta_0)$ for $0 \le t \le 1$. To show the above inequality, it suffices to prove that

$$(\theta - \theta_0)^T \left\{ \mathbb{E}_{\theta_0,\eta_0} \ddot{\ell}_{\theta(t),\eta}(x_i, Y_i) - \mathbb{E}_{\theta_0,\eta_0} \ddot{\ell}_{\theta_0,\eta}(x_i, Y_i) \right\} (\theta - \theta_0) \quad (5.54)$$
$$\lesssim |x_i^T(\theta - \theta_0)^T|^2 \sqrt{\log p} \|\theta - \theta_0\|_1$$

for any $i = 1, \ldots, n$. Note that (5.54) is bounded above by

$$|x_i^T(\theta - \theta_0)|^2 \left| \mathbb{E}_{\theta_0,\eta_0} \left(\ddot{\ell}_\eta(Y_i - x_i^T\theta(t)) - \ddot{\ell}_\eta(Y_i - x_i^T\theta_0) \right) \right|$$
$$\lesssim |x_i^T(\theta - \theta_0)|^2 \sqrt{\log p} \|\theta - \theta_0\|_1 \cdot \left| \mathbb{E}_{\theta_0,\eta_0} \dddot{\ell}_\eta(Y_i - x_i^T\theta(t_1)) \right|,$$

for some constant $0 \le t_1 \le t$. Also note that

$$\left| \mathbb{E}_{\theta_0,\eta_0} \left(\dddot{\ell}_\eta(Y - x^T\theta(t_1)) - \dddot{\ell}_{\eta_0}(Y - x^T\theta(t_1)) \right) \right|$$
$$= \left| \int \left(\dddot{\ell}_\eta(y - x^T\theta(t_1)) - \dddot{\ell}_{\eta_0}(y - x^T\theta(t_1)) \right) \eta_0(y - x^T\theta_0) dy \right|$$
$$= \left| \int \left(\dot{\ell}_\eta(y - x^T\theta(t_1)) - \dot{\ell}_{\eta_0}(y - x^T\theta(t_1) \right) \ddot{\eta}_0(y - x^T\theta_0) dy \right|$$
$$\le \left[\int (\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y))^2 \eta_0(y) dy \right]^{\frac{1}{2}}$$
$$\times \left[\int \left(\frac{\ddot{\eta}_0(y - x^T\theta_0)}{\eta_0(y - x^T\theta_0)} \right)^2 \frac{\eta_0(y - x^T\theta_0)}{\eta_0(y - x^T\theta(t_1))} \eta_0(y - x^T\theta_0) dy \right]^{\frac{1}{2}}.$$

The above equality follows from the integration by parts, and the last inequality follows from the Hölder's inequality. The last term is of order $O(1)$ by Lemma 9. Since $\left| \mathbb{E}_{\theta_0,\eta_0} \dddot{\ell}_{\eta_0}(Y - x^T\theta(t_1)) \right| \lesssim 1$, it completes the proof for (5.53). ∎

Finally, the following lemma is used to prove Lemma 4.

Lemma 12 *Suppose that* $(s_n \log p)^{1+\frac{8}{a_2}} = o(n^{1-\zeta})$ *holds for some constant* $\zeta > 0$*, then*

$$\mathbb{E}_{\theta_0,\eta_0} \left(\sup_{\eta \in \mathcal{H}_n^*} \|\mathbb{G}_n \dot{\ell}_{\theta_0,\eta}\|_\infty \right) \lesssim \log p$$

for any η_0 *satisfying (D1)-(D5), where* $\mathcal{H}_n^*$ *defined at* (3.17).

Proof Without loss of generality, we assume that $\theta_0 = 0$. Define

$$\mathcal{F}_n := \left\{ e_j^T \dot{\ell}_{\theta_0,\eta} (\log p)^{-\frac{1}{2}} : 1 \le j \le p, \ \eta \in \mathcal{H}_n^* \right\},$$

where e_j is the jth unit vector in $\mathbb{R}^p$. Then,

$$\sup_{\eta \in \mathcal{H}_n^*} \|\mathbb{G}_n \dot{\ell}_{\theta_0,\eta}\|_\infty = \sup_{f \in \mathcal{F}_n} |\mathbb{G}_n f| \sqrt{\log p}.$$

We first show that $F_n(x,y) := \sup_{\eta \in \mathcal{H}_n^*} |\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)| + |\dot{\ell}_{\eta_0}(y)|$ is an envelop function of $\mathcal{F}_n$ and $\mathbb{E}_{\theta_0,\eta_0} F_n^2(x_i, Y_i) \lesssim 1$ for any $i = 1, \ldots, n$. Note that for any $f \in \mathcal{F}_n$ and $x = (x_1, \ldots, x_p)^T$,

$$\begin{aligned}
|f(x,y)| &= \left| e_j^T \dot{\ell}_{\theta_0,\eta}(x,y) \right| (\log p)^{-\frac{1}{2}} \\
&= \left| x_j \cdot \dot{\ell}_\eta(y) \right| (\log p)^{-\frac{1}{2}} \\
&\lesssim \sup_{\eta \in \mathcal{H}_n^*} |\dot{\ell}_\eta(y) - \dot{\ell}_{\eta_0}(y)| + |\dot{\ell}_{\eta_0}(y)|.
\end{aligned}$$

By Lemma 7, we have $\mathbb{E}_{\theta_0,\eta_0} F_n^2(x_i, Y_i) \lesssim 1$ if $(s_n \log p)^{1+\frac{8}{a_2}} = O(n^{1-\zeta})$ for some $\zeta > 0$. Then, we have

$$\begin{aligned}
\mathbb{E}_{\theta_0,\eta_0}\left(\sup_{\eta \in \mathcal{H}_n^*} \|\mathbb{G}_n \dot{\ell}_{\theta_0,\eta}\|_\infty \right) &\lesssim \int_0^{\|F_n\|_n} \sqrt{\log N_{[]}^n(\epsilon, \mathcal{F}_n)}\, d\epsilon \sqrt{\log p} \\
&\lesssim \int_0^{\|F_n\|_n} \sqrt{\epsilon^{-1} + \log p}\, d\epsilon \sqrt{\log p} \ \lesssim \ \log p,
\end{aligned}$$

where the second inequality follows from Corollary 2.7.4 of **?**. ■