Online Supplementary Appendix

Siddhartha Chib*

Edward Greenberg†

Anna Simoni‡

January 5, 2022

D.1 Notation

The notation $\lesssim$ (resp. $\asymp$) will be used to denote inequality (resp. equality) up to a fixed constant. For a vector v, denote by $\|v\|_\infty \triangleq \max_i |v_i|$. The indicator function of an event $\mathscr{A}$ is denoted by both $I[\mathscr{A}]$ and $\mathbb{1}_{\mathscr{A}}$.

The true data distribution, conditional on z, of the sample is denoted by $P_*^{(n)}$ in both designs and is as specified in Assumptions 1 and 5-7. The true data distribution, conditional on z_i, for one observation is denoted by P_*^i in both designs for short. The Lebesgue density of $P_*^{(n)}$ (resp. P_*^i) is denoted by $p_*^{(n)}$ (resp. $p_*(y_i|z_i)$ in the sharp RD design and $p_*(y_i, x_i|z_i)$ in the fuzzy RD design). In the proof we exploit the fact that we can write the student-t distribution as a mixture of normal distributions: in the sharp RD design $p_*(y_i|z_i) = t_\nu(y_i|g_j^*(z_i), \sigma_{j*}^2) = \int \mathscr{N}(y_i|z_i; g_j^*, \sigma_{j*}^2/\xi_i)\mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i$ for $j = 0, 1$ (and similarly in the fuzzy RD design). The expectation taken with respect to $P_*^{(n)}$ is denoted by $\mathbf{E}_*$. For a probability measure P, the notation Ph will abbreviate $\int h dP$. Define the Kullback-Leibler divergence between two probability measures P and Q as $K(P, Q) \triangleq P \log(p/q)$, where p and q are the Lebesgue densities of P and Q respectively, and define the discrepancy measure

$$V(P, Q) \triangleq P \left|\log(p/q) - K(P, Q)\right|^2 .$$

*Olin Business School, Washington University in St. Louis, Campus Box 1133, 1 Bookings Drive, St. Louis, MO 63130. e-mail: chib@wustl.edu.

†formerly, Department of Economics, Washington University in St. Louis, passed away in September 2016, co-author in remembrance of contributions to this research.

‡CREST, CNRS, ENSAE, École Polytechnique, Université Paris-Saclay, 5, Avenue Henry Le Chatelier, 91120 Palaiseau - France, e-mail: anna.simoni@ensae.fr

For a set $\mathscr{F}$ and a metric d on it let $N(\epsilon, \mathscr{F}, d)$ be the ϵ-covering number which is defined as the minimum number of balls of radius ϵ needed to cover $\mathscr{F}$ see *e.g.* Van der Vaart [2000].

Notation specific to the Sharp RD design. Let $\mathscr{G}_0 \triangleq \mathscr{C}^\delta[a,\tau)$, $\mathscr{G}_1 \triangleq \mathscr{C}^\delta[\tau,b]$ and

$$B_*^c(\epsilon_{n_j}, n^{-1/2}) \triangleq \{(g_j,\sigma_j^2) \in \mathscr{G}_j \times \mathbb{R}_+; \|g_j - g_j^*\|_{n_j} \geq M_n\epsilon_{n_j},\, |1-\sigma_j/\sigma_{j*}| \geq \widetilde{M}_n/\sqrt{n_j}\}$$

for $j=0,1$ and any sequences $M_n, \widetilde{M}_n \to \infty$. For $j=0,1$, let $P^{(n_j)}_{g_j,\sigma_j^2}$ (resp. $P^i_{g_j,\sigma_j^2}$) denote the conditional distribution of the sample $y^{(n_j)}$ given $(z^{(n_j)}, g_j, \sigma_j^2)$ (resp. of y_i given (z_i, g_j, σ_j^2)), and $p^{(n_j)}_{g_j,\sigma_j^2}$ (resp. $p_{g_j,\sigma_j^2}(y_i|z_i)$) denote the Lebesgue density of $P^{(n_j)}_{g_j,\sigma_j^2}$ (resp. of $P^i_{g_j,\sigma_j^2}$). Denote the subsamples $I_0 \triangleq \{i; z_i < \tau\}$ and $I_1 \triangleq \{i; z_i \geq \tau\}$. For $j=0,1$, the Kullback-Leibler neighborhood of (g_j^*, σ_{j*}^2) is defined as, $\forall \epsilon_{n_j} > 0$

$$B_n^{KL}((g_j^*,\sigma_{j*}^2),\epsilon_{n_j}) \triangleq \left\{(g_j,\sigma_j^2) \in \mathscr{G}_j \times \mathbb{R}_+;\ K\left(P_*^{(n_j)}, P^{(n_j)}_{g_j,\sigma_j^2}\right) \leq n_j\epsilon_{n_j}^2,\, V\left(P_*^{(n_j)}, P^{(n_j)}_{g_j,\sigma_j^2}\right) \leq 2n_j\epsilon_{n_j}^2\right\}. \quad \text{(D.1)}$$

Finally, in the proofs we use the following sequence of measurable sets for $M_k = 4\gamma_k n_k \epsilon_n^2 \asymp n_k\epsilon_n^2 \asymp n\epsilon_n^2$, where $\gamma_k \triangleq \max_{1\leq j\leq m_k^*} |(\boldsymbol{D}^{-1}_{\boldsymbol{\alpha_k}} \boldsymbol{T}_{\boldsymbol{\alpha_k}} \boldsymbol{D}^{-1'}_{\boldsymbol{\alpha_k}})_{jj}|$, $\boldsymbol{\alpha}_k \triangleq \boldsymbol{\alpha} I[k=0] + \boldsymbol{\beta} I[k=1]$ and $m_k^* \asymp \left(\frac{n_k}{\log(n_k)}\right)^{1/(2\delta+1)}$, $k=0,1$:

$$\mathscr{C}_{n,0} \triangleq \bigcup_{m_0=1}^{m_0^*} \{g_{m_0}(z) \in \mathscr{S}_{m_0};\ \boldsymbol{\alpha} \in [-M_0, M_0]^{m_0}\},$$

$$\mathscr{C}_{n,1} \triangleq \bigcup_{m_1=1}^{m_1^*} \{g_{m_1}(z) \in \mathscr{S}_{m_1};\ \boldsymbol{\beta} \in [-M_1, M_1]^{m_1}\}, \quad \text{(D.2)}$$

where $g_{m_0}(z) \triangleq \boldsymbol{B}_0(z)'\boldsymbol{\alpha}$ and $g_{m_1}(z) \triangleq \boldsymbol{B}_1(z)'\boldsymbol{\beta}$.

Notation specific to the Fuzzy RD design. For the Fuzzy RD design we in addition use the following notation. Let $\mathscr{G}_{0n} = \mathscr{G}_{1a} \triangleq \mathscr{C}^\delta[a,b]$, and

$$\mathscr{F} \triangleq \Big\{ \Big(g_{0,0n}(s,z) I[s\in\{c,n\}] + g_{1a}(z) I[s=a]\Big) I[z<\tau] + \Big(g_{1,1a}(s,z) I[s\in\{c,a\}] + g_{0n}(z) I[s=n]\Big) I[z\geq\tau], \quad \text{with } g_{0,0n}(s,z) \triangleq g_0(z)I[s=c] + g_{0n}(z)I[s=n],\ g_{1,1a}(s,z) \triangleq g_1(z)I[s=c] + g_{1a}(z)I[s=a],\ g_j \in \mathscr{G}_j, j=0,1,0n,1a \Big\}.$$

We use the notation $\bar{n}_{10} \triangleq n_{00} + n_{10}$, $\bar{n}_{01} \triangleq n_{00} + n_{10} + n_{01}$. For a given sequence $\boldsymbol{s} \triangleq \{\boldsymbol{s_0}, \boldsymbol{s_1}\}$

where $\boldsymbol{s_k} \triangleq \{s_i\}_{i\in I_{kk}}$ with values in $\{c, n\}$ if $k=0$ and in $\{c, a\}$ if $k=1$ denote:

$$\text{for } f \in \mathscr{F}, \quad \boldsymbol{f_s} \triangleq (\boldsymbol{f_{s_0}}', f(n, z_{n_{00}+1}), \ldots, f(n, z_{\bar{n}_{10}}), f(a, z_{\bar{n}_{10}+1}), \ldots, f(n, z_{\bar{n}_{01}}), \boldsymbol{f_{s_1}}')',$$
$$\boldsymbol{f_{s_0}} \triangleq (f(s_1, z_1), \ldots, f(s_{n_{00}}, z_{n_{00}}))', \qquad \boldsymbol{f_{s_1}} \triangleq (f(s_{\bar{n}_{01}+1}, z_{\bar{n}_{01}+1}), \ldots, f(s_n, z_n))',$$
$$\boldsymbol{\sigma^2_{0s_0}} \triangleq (\sigma^2_{0s_1}, \ldots, \sigma^2_{0s_{n_{00}}})', \qquad \boldsymbol{\sigma^2_{1s_1}} \triangleq (\sigma^2_{1s_{\bar{n}_{01}+1}}, \ldots, \sigma^2_{1s_n})',$$
$$\boldsymbol{g_{0s_0}} \triangleq (g_{0s_1}(z_1), \ldots, g_{0s_{n_{00}}}(z_{n_{00}}))', \qquad \boldsymbol{g_{1s_1}} \triangleq (g_{1s_{\bar{n}_{01}+1}}(z_{\bar{n}_{01}+1}), \ldots, g_{1s_n}(z_n))'. \quad \text{(D.3)}$$

Moreover, define $B_*(\epsilon_n, n^{-1/2}) \triangleq \{(f, \sigma_j^2) \in \mathscr{F} \times \mathbb{R}_+$ for $j = 0, 1, 0n, 1a; |1 - \sigma_j/\sigma_{j*}| \geq \widetilde{M}_n/\sqrt{n},$ and $\forall \boldsymbol{s}, \|\boldsymbol{f_s} - \boldsymbol{f^*_s}\|_n \geq M_n\epsilon_n\}$ for any sequences $M_n, \widetilde{M}_n \to \infty$.

Let $\boldsymbol{g} \triangleq (g_0, g_1, g_{0n}, g_{1a})$, $\boldsymbol{\theta} \triangleq (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\alpha}_n, \boldsymbol{\beta}_a)$, $\boldsymbol{\sigma}^2 \triangleq (\sigma_0^2, \sigma_1^2, \sigma_{0n}^2, \sigma_{1a}^2)$, $\boldsymbol{q} \triangleq (q_c, q_n, q_a)$. Let $P^{(n)}_{\boldsymbol{g},\boldsymbol{\sigma^2}}$ (resp. $P^i_{\boldsymbol{g},\boldsymbol{\sigma^2}}$) denote the conditional distribution of the sample $(y^{(n)}, x^{(n)})$ given $(z^{(n)}, \boldsymbol{g}, \boldsymbol{\sigma^2}, \boldsymbol{q})$ (resp. of (y_i, x_i) given $(z_i, \boldsymbol{g}, \boldsymbol{\sigma^2}, \boldsymbol{q})$), and $p^{(n)}_{\boldsymbol{g},\boldsymbol{\sigma^2}}$ (resp. $p_{\boldsymbol{g},\boldsymbol{\sigma^2}}(y_i, x_i|z_i)$) denote the Lebesgue density of $P^{(n)}_{\boldsymbol{g},\boldsymbol{\sigma^2}}$ (resp. of $P^i_{\boldsymbol{g},\boldsymbol{\sigma^2}}$). The Kullback-Leibler neighborhood of $(\boldsymbol{g}^*, \boldsymbol{\sigma^2_*}, \boldsymbol{q}^*)$ is defined as, $\forall \varepsilon > 0$

$$B_n^{KL}((\boldsymbol{g}^*, \boldsymbol{\sigma^2_*}, \boldsymbol{q}^*), \varepsilon) \triangleq \Big\{(g_j, \sigma_j^2) \in \mathscr{G}_j \times \mathbb{R}_+ \text{ for } j = 0, 1, 0n, 1a, \boldsymbol{q} \in [0,1]^3;$$
$$K\left(P_*^{(n)}, P^{(n)}_{\boldsymbol{g},\boldsymbol{\sigma^2}}\right) \leq n\varepsilon^2,\ V\left(P_*^{(n)}, P^{(n)}_{\boldsymbol{g},\boldsymbol{\sigma^2}}\right) \leq 2n\varepsilon^2\Big\}. \quad \text{(D.4)}$$

Finally, in the proofs we use the following sequence of measurable sets for $M_k = 4\gamma n_k\epsilon_n^2 \asymp n_k\epsilon_n^2 \asymp n\epsilon_n^2$, where $\gamma \triangleq \max_{1\leq j\leq m_k^*} |(\boldsymbol{D}^{-1}_{\boldsymbol{\alpha_k}}\boldsymbol{T}_{\boldsymbol{\alpha_k}}\boldsymbol{D}^{-1'}_{\boldsymbol{\alpha_k}})_{jj}|$, $\boldsymbol{\alpha}_k \triangleq \boldsymbol{\alpha}I[k=0] + \boldsymbol{\beta}I[k=1] + \boldsymbol{\alpha_n}I[k = 0n] + \boldsymbol{\beta_a}I[k=1a]$ and $m_k^* \asymp \left(\frac{n_k}{\log(n_k)}\right)^{1/(2\delta+1)}$, $k = 0, 1, 0n, 1a$:

$$\mathscr{C}_{n,0} \triangleq \bigcup_{m_0=1}^{m_0^*} \{g_{m_0}(z) \in \mathcal{S}_{m_0};\ \boldsymbol{\alpha} \in [-M_0, M_0]^{m_0}\},$$
$$\mathscr{C}_{n,1} \triangleq \bigcup_{m_1=1}^{m_1^*} \{g_{m_1}(z) \in \mathcal{S}_{m_1};\ \boldsymbol{\beta} \in [-M_1, M_1]^{m_1}\},$$
$$\mathscr{C}_{n,0n} \triangleq \bigcup_{m_n=1}^{m_n^*} \{g_{m_n}(z) \in \mathcal{S}_{m_n};\ \boldsymbol{\alpha}_n \in [-M_n, M_n]^{m_n}\},$$
$$\mathscr{C}_{n,1a} \triangleq \bigcup_{m_a=1}^{m_a^*} \{g_{m_a}(z) \in \mathcal{S}_{m_a};\ \boldsymbol{\beta}_a \in [-M_a, M_a]^{m_a}\}. \quad \text{(D.5)}$$

where $g_{m_0}(z) \triangleq \boldsymbol{B}_{00}(z)'\boldsymbol{\alpha}$, $g_{m_1}(z) \triangleq \boldsymbol{B}_{11}(z)'\boldsymbol{\beta}$, $g_{m_n}(z) \triangleq \boldsymbol{B}_{0,n}(z)'\boldsymbol{\alpha}_n$ and $g_{m_a}(z) \triangleq \boldsymbol{B}_{1,a}(z)'\boldsymbol{\beta}_a$, and

$$
\begin{aligned}
\mathscr{F}_n \triangleq \Big\{ & \Big(g_{0,0n}(s,z)I[s\in\{c,n\}] + g_{1a}(z)I[s=a]\Big)I[z<\tau] + \Big(g_{1,1a}(s,z)I[s\in\{c,a\}] \\
& + g_{0n}(z)I[s=n]\Big)I[z\geq\tau],\ g_{0,0n}(s,z) \triangleq g_0(z)I[s=c] + g_{0n}(z)I[s=n], \\
& g_{1,1a}(s,z) \triangleq g_1(z)I[s=c] + g_{1a}(z)I[s=a],\ g_j \in \mathscr{C}_{n,j},\ j=0,1,0n,1a\Big\}.
\end{aligned}
$$

Prior specification for both the Sharp and the Fuzzy RD designs. The prior specification for $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\alpha}_n, \boldsymbol{\beta}_a, \boldsymbol{\lambda}, \sigma_0^2, \sigma_1^2, \sigma_n^2, \sigma_a^2, \boldsymbol{q})$, where $\boldsymbol{\lambda} \triangleq (\lambda_0, \lambda_1, \lambda_n, \lambda_a)$ and $\boldsymbol{q} \triangleq (q_c, q_n, q_a)$ is an independent prior as follows:

$$
\begin{aligned}
\boldsymbol{\alpha}|\lambda_0 &\sim \mathscr{N}_{m_0}(\boldsymbol{D}_\alpha^{-1}\boldsymbol{\alpha}_0, \lambda_0^{-1}\boldsymbol{D}_\alpha^{-1}\boldsymbol{T}_\alpha\boldsymbol{D}_\alpha^{-1'}), \qquad \boldsymbol{\beta}|\lambda_1 \sim \mathscr{N}_{m_1}(\boldsymbol{D}_\beta^{-1}\boldsymbol{\beta}_0, \lambda_1^{-1}\boldsymbol{D}_\beta^{-1}\boldsymbol{T}_\beta\boldsymbol{D}_\beta^{-1'}), \\
\boldsymbol{\alpha}_n|\lambda_n &\sim \mathscr{N}_{m_n}(\boldsymbol{D}_n^{-1}\boldsymbol{\alpha}_{0n}, \lambda_n^{-1}\boldsymbol{D}_n^{-1}\boldsymbol{T}_n\boldsymbol{D}_n^{-1'}), \qquad \boldsymbol{\beta}_a|\lambda_a \sim \mathscr{N}_{m_a}(\boldsymbol{D}_a^{-1}\boldsymbol{\beta}_{0a}, \lambda_a^{-1}\boldsymbol{D}_a^{-1}\boldsymbol{T}_a\boldsymbol{D}_a^{-1'}), \\
\lambda_j &\sim \mathscr{G}a\left(\frac{a_{j0}}{2}, \frac{b_{j0}}{2}\right), \qquad j=0,1,n,a \\
\sigma_j^2 &\sim \mathscr{IG}\left(\frac{\nu_{00}}{2}, \frac{\delta_{00}}{2}\right), \text{ for } j=0,1, \quad \sigma_n^2 \sim \mathscr{IG}\left(\frac{\nu_{0n}}{2}, \frac{\delta_{0n}}{2}\right), \quad \sigma_a^2 \sim \mathscr{IG}\left(\frac{\nu_{0a}}{2}, \frac{\delta_{0a}}{2}\right), \\
\boldsymbol{q} &\sim Dir(n_{0c}, n_{0n}, n_{0a}),
\end{aligned} \tag{D.6}
$$

where $\nu_{00} > 2$.

E Proofs for Section 2

E.1 Proof of Theorem 2.1

For every n, denote $X^{(n)} \triangleq (y^{(n)}, z^{(n)})$. The potential outcome y_j, $j=0,1$ is given by

$$y_0 = g_0(z) + \sigma_0\varepsilon_0, \qquad z<\tau, \tag{E.1}$$

$$y_1 = g_1(z) + \sigma_1\varepsilon_1, \qquad z\geq\tau, \tag{E.2}$$

where for $j=0,1$, $\varepsilon_j \sim t_\nu(0,1)$ are error terms independent of z and independent between them. Let $I_0 \triangleq \{i; z_i<\tau\} = \{i=1,\dots n_0\}$ and $I_1 \triangleq \{i; z_i\geq\tau\} = \{i=n_0+1,\dots n_1\}$. Moreover, for $k=0,1$ denote by D_{n_k} the denominator of the posterior distribution of (g_k, σ_k^2):

$$D_{n_k} \triangleq \int \prod_{\{i\in I_k\}} p_{g_k,\sigma_k^2}(y_i|z_i)\pi(g_k|\lambda_k)\pi(\lambda_k)\pi(\sigma_k^2)d\lambda_k d\sigma_k^2 dg_k,$$

where $p_{g_k,\sigma_k^2}(y_i|z_i) \triangleq t_\nu(y_i|g_k(z_i), \sigma_k^2)$. In the following we use the notation $\boldsymbol{\alpha}_k \triangleq \boldsymbol{\alpha} I[k=0] + \boldsymbol{\beta} I[k=1]$ for $k = 0, 1$ (and similarly $\boldsymbol{\alpha}_{k0}$ for the prior mean). Therefore, for $k = 0, 1$, the marginal posterior of (g_k, σ_k^2) is

$$\pi((g_k,\sigma_k^2)|X^{(n_k)}) =$$
$$\frac{1}{D_{n_k}} \int \prod_{\{i\in I_k\}} t_\nu(y_i|\boldsymbol{B}_k(z_i)'\boldsymbol{\alpha}_k + g_k^\perp, \sigma_k^2)\mathscr{N}_{m_k}(\boldsymbol{D}_{\boldsymbol{\alpha}_k}^{-1}\boldsymbol{\alpha}_{k0}, \lambda_k^{-1}\boldsymbol{D}_{\boldsymbol{\alpha}_k}^{-1}\boldsymbol{T}_{\boldsymbol{\alpha}_k}\boldsymbol{D}_{\boldsymbol{\alpha}_k}^{-1'}) \otimes \delta_0(dg_k^\perp)d\pi(\lambda_k)\pi(\sigma_k^2),$$

where δ_0 denotes the Dirac mass at zero and $g_k^\perp \triangleq g_k^\perp(z) \triangleq g_k(z) - \boldsymbol{B}_k(z)'\boldsymbol{\alpha}_k$. For $k = 0, 1$, for some $C > 0$ and $\epsilon_{n_k} > 0$ denote by $\mathscr{A}_k^c$ the following event:

$$\mathscr{A}_k^c \triangleq \left\{ \int \frac{\prod_{i\in I_k} p_{g_k,\sigma_k^2}(y_i|z_i)}{p_*^{(n_k)}} d\overline{\pi}(\boldsymbol{\alpha}_k, g_k^\perp, \lambda_k, \sigma_k^2) \leq e^{-(1+C)n_k\epsilon_{n_k}^2} \right\}$$

where $d\overline{\pi}(\boldsymbol{\alpha}_k, g_k^\perp, \lambda_k, \sigma_k^2)$ denotes the prior supported on $B_n^{KL}((g_j^*, \sigma_{j*}^2), \epsilon_{n_j})$. By Ghosal and van der Vaart [2007, Lemma 10], $P_*^{(n_k)}[\mathscr{A}_k^c] \leq \frac{1}{C^2 n_k \epsilon_{n_k}^2}$ for every $C > 0$.

To prove the result of the theorem we use the test approach which relies on the uniformly exponential consistent test ϕ_{n_k} constructed in Lemma E.4. By this lemma, for $k = 0, 1$ there exists a test ϕ_{n_k} that satisfies:

$$Q_{g_k^*,\sigma_{k*}^2}^{(n_k)}\phi_{n_k} \leq e^{n_k\epsilon_{n_k}^2}(1 - e^{-KM^2 n_k\epsilon_{n_k}^2})^{-1}e^{-KM^2 n_k \epsilon_{n_k}^2} \qquad \text{and}$$
$$Q_{g_k,\sigma_k^2}^{(n_k)}(1-\phi_{n_k}) \leq e^{-KM^2 n_k\epsilon_{n_k}^2 j^2}, \qquad \forall (g_k,\sigma_k^2) \in \mathscr{C}_{n,k} \times \left[\frac{1}{2n_k}, e^{3n_k\epsilon_{n_k}^2}\right] \text{ such that } |1-\sigma_k/\sigma_{k*}| > j\varepsilon_\sigma$$
$$\text{and } \|\Xi_k^{1/2}(g_k - g_k^*)\|_{n_k} > M\epsilon_{n_k} j, \ \forall j \in \mathbb{N}, \quad \text{(E.3)}$$

for some $\varepsilon_\sigma > 1/(2n_k^2)$, where $Q_{g_k^*,\sigma_{k*}^2}^{(n_k)} \triangleq \prod_{\{i\in I_k\}} \mathscr{N}(y_i|z_i; g_k^*, \sigma_{k*}^2/\xi_i)$ which, conditional on $\{\xi_i\}_{i=1}^n$, is the true model, $Q_{g_k,\sigma_k^2}^{(n_k)} \triangleq \prod_{\{i\in I_k\}} \mathscr{N}(y_i|z_i; g_k, \sigma_k^2/\xi_i)$ and $\Xi_0 \triangleq \mathrm{diag}(\xi_1, \ldots, \xi_{n_0})$, $\Xi_1 \triangleq \mathrm{diag}(\xi_{n_0+1}, \ldots, \xi_{n_1})$.

We now use the tests ϕ_{n_0} and ϕ_{n_1} to upper bound the posterior of $B_*^c(\epsilon_{n_k}, n^{-1/2})$. For this we use the fact that, since $t_\nu(y_i|g_j^*(z_i), \sigma_{j*}^2) = \int \mathscr{N}(y_i|z_i; g_j^*, \sigma_{j*}^2/\xi_i)\mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i$ for $j = 0, 1$, the true Lebesgue density can be written as a mixture of normal distributions:

$$p_*^{(n)} = \int Q_{g_0^*,\sigma_{0*}^2}^{(n_0)} Q_{g_1^*,\sigma_{1*}^2}^{(n_1)} \prod_{i=1}^n \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i.$$

For $k = 0, 1$:

$$\begin{aligned}
\mathbf{E}_*[\pi(B_*^c(\epsilon_{n_k}, n^{-1/2})|X^{(n_k)})] &\leq \mathbf{E}_*[\overbrace{\pi(B_*^c(\epsilon_{n_k}, n^{-1/2})|X^{(n_k)})}^{\leq 1}\phi_{n_k}] \\
&+ \mathbf{E}_*[\overbrace{\pi(B_*^c(\epsilon_{n_k}, n^{-1/2})|X^{(n_k)})}^{\leq 1}\mathbb{1}_{\mathscr{A}_k^c}] + \mathbf{E}_*[\pi(B_*^c(\epsilon_{n_k}, n^{-1/2})|X^{(n_k)})(1-\phi_{n_k})\mathbb{1}_{\mathscr{A}_k}] \\
&\leq \int Q^{(n_k)}_{g_k^*,\sigma^2_{k*}}\phi_{n_k} \prod_{\{i\in I_k\}} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i + \frac{1}{C^2 n_k \epsilon^2_{n_k}} \\
&+ \mathbf{E}_*\left[\frac{\int\int_{B_*^c(\epsilon_{n_k}, n^{-1/2})} \frac{\prod_{\{i\in I_k\}} p_{g_k,\sigma_k^2}(y_i|z_i)}{p_*^{(n_k)}} d\pi(\boldsymbol{\alpha}_k, g_k^\perp, \lambda_k, \sigma_k^2)}{D_{n_k}/p_*^{(n_k)}}(1-\phi_{n_k})\mathbb{1}_{\mathscr{A}_k}\right] \\
&\leq \int \frac{e^{-(KM^2-1)n_k\epsilon^2_{n_k}}}{(1-e^{-Kn_kM^2\epsilon^2_{n_k}})} \prod_{i\in I_k} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i + \frac{1}{C^2 n_k \epsilon^2_{n_k}} \\
&+ \mathbf{E}_*\left[\frac{\int\int_{B_*^c(\epsilon_{n_k}, n^{-1/2})} \prod_{\{i\in I_k\}} \frac{p_{g_k,\sigma_k^2}(y_i|z_i)}{p_*(y_i|z_i)} d\pi(\boldsymbol{\alpha}_k, g_k^\perp, \lambda_k, \sigma_k^2)}{D_{n_k}/p_*^{(n_k)}}(1-\phi_{n_k})\mathbb{1}_{\mathscr{A}_k}\right]
\end{aligned}$$

where to get the third inequality we have used the first line in (E.3) with $M = \xi_{\min}^{1/2} M_n/J$ for $J \in \mathbb{N}$ and $\xi_{\min} \triangleq \min_i(\Xi_k)_{i,i}$. If M is sufficiently large to ensure that $KM^2 - 1 > KM^2/2$, $\frac{e^{-(KM^2-1)n_k\epsilon^2_{n_k}}}{(1-e^{-Kn_kM^2\epsilon^2_{n_k}})} \leq e^{-KM^2 n\epsilon^2_{n_k}/2}$ and by using the definition of $\mathscr{A}_k$ we obtain

$$\begin{aligned}
\mathbf{E}_*[\pi(B_*^c(\epsilon_{n_k}, n^{-1/2})|X^{(n_k)})] &\leq \int e^{-K\xi_{\min}M_n^2 n_k \epsilon^2_{n_k}/(2J^2)} \prod_{i\in I_k} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i + \frac{1}{C^2 n_k\epsilon^2_{n_k}} \\
+ \mathbf{E}_*\Big[\int\int_{B_*^c(\epsilon_{n_k}, n^{-1/2})} \prod_{\{i\in I_k\}} \frac{p_{g_k,\sigma_k^2}(y_i|z_i)}{p_*(y_i|z_i)} d\pi(\boldsymbol{\alpha}, g_k^\perp, \lambda_k, \sigma_k^2)(1-\phi_{n_k})\Big] &\frac{e^{(1+C)n_k\epsilon^2_{n_k}}}{\pi(B_n^{KL}((g_j^*, \sigma^2_{j*}), \epsilon_{n_j}))}.
\end{aligned} \tag{E.4}$$

Next, remark that the density function of the minimum $\xi_{\min}$ is

$$p(\xi_{\min}) = n_k \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)} (\lambda_{\min})^{\nu/2-1} e^{-\xi_{\min}(\nu/2)} \left(\frac{\Gamma(\nu/2, \nu/2\xi_{\min})}{\Gamma(\nu/2)}\right)^{n_k-1},$$

where $\Gamma(\nu/2, \nu/2\xi_{\min})$ is the upper incomplete gamma function. Therefore,

$$\int e^{-K\xi_{\min}M_n^2 n_k \epsilon_{n_k}^2/(2J^2)} \prod_{i \in I_k} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i$$

$$\leq \int e^{-K\xi_{\min}M_n^2 n_k \epsilon_{n_k}^2/(2J^2)} n_k \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)} (\lambda_{\min})^{\nu/2-1} e^{-\xi_{\min}(\nu/2)} d\xi_{\min}$$

$$\leq 2e^{-Kc_q M_n^2 n_k \epsilon_{n_k}^2/(2J^2)} n_k \int_{c_q}^{+\infty} \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)} (\lambda_{\min})^{\nu/2-1} e^{-\xi_{\min}(\nu/2)} d\xi_{\min}$$

$$\leq 2e^{-n_k \epsilon_{n_k}^2 (Kc_q M_n^2/(2J^2) - \log(n_k)/(n_k \epsilon_{n_k}^2))}, \quad \text{(E.5)}$$

where, for $q \in (0, 1/2)$, c_q is the q-quantile of a $\mathscr{G}a\left(\frac{\nu}{2}, \frac{\nu}{2} + \frac{KM_n^2 n_k \epsilon_{n_k}^2}{2J^2}\right)$. Moreover, by using the result of Lemma E.1: $\pi(B_n^{KL}(g_k^*, \epsilon_{n_k})) \gtrsim e^{-n_k \epsilon_{n_k}^2}$ for $k = 0, 1$, and by Fubini's theorem

$$\mathbf{E}_*[\pi(B_*^c(\epsilon_{n_k}, n^{-1/2})|X^{(n_k)})] \lesssim e^{-n_k \epsilon_{n_k}^2 (Kc_q M_n/(2J) - \epsilon_{n_k})} + \frac{1}{C^2 n_k \epsilon_{n_k}^2} + e^{(1+C+1)n_k \epsilon_{n_k}^2} \times$$
$$\int \int_{B_*^c(\epsilon_{n_k}, n^{-1/2})} \int \underbrace{\prod_{\{i \in I_k\}} \mathscr{N}(y_i | z_i; \boldsymbol{\alpha}, g_0^{\perp}, \sigma_k^2/\xi_i)}_{=Q_{g_k, \sigma_k^2}^{(n_k)}} (1 - \phi_{n_k}) dy_i d\pi(\boldsymbol{\alpha}, g_k^{\perp}, \lambda_k, \sigma_k^2) \prod_{i \in I_k} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i \quad \text{(E.6)}$$

where we have used again the fact that $p_{g_k, \sigma_k^2}(y_i|z_i) = \int \mathscr{N}(y_i|z_i; \boldsymbol{\alpha}_k, g_k^{\perp}, \sigma_k^2/\xi_i) \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i$ for $k = 0, 1$. Set for $k = 0, 1$ and for $j \in \mathbb{N}$:

$$B_{k,j} \triangleq \{(g_k, \sigma_k^2) \in \mathscr{C}_{n,k} \times \left[\frac{1}{2n_k}, e^{3n_k \epsilon_{n_k}^2}\right]; M\epsilon_{n_k} j < \|\Xi_k^{1/2}(g_k - g_{k*})\|_n \leq 2jM\epsilon_{n_k},$$
$$j\varepsilon_\sigma < |1 - \sigma_k/\sigma_{k*}| < 2j\varepsilon_\sigma\}$$

and $\mathscr{G}_{k,n} \triangleq \{(g_k, \sigma_k^2) \in \mathscr{C}_{n,k} \times \left[\frac{1}{2n_k}, e^{3n_k \epsilon_{n_k}^2}\right]; \|g_k - g_{k*}\|_n \geq M_n \epsilon_{n_k}, |1 - \sigma_k/\sigma_{k*}| > \widetilde{M}_n/\sqrt{n}\}$

$$\subseteq \left\{(g_k, \sigma_k^2) \in \mathscr{C}_{n,k} \times \left[\frac{1}{2n_k}, e^{3n_k \epsilon_{n_k}^2}\right]; \|\Xi_k^{1/2}(g_j - g_j^*)\|_n \geq \xi_{\min}^{1/2} M_n \epsilon_{n_k}, \ |1 - \sigma_k/\sigma_{k*}| > \widetilde{M}_n/\sqrt{n}\right\},$$

then $B_*^c(\epsilon_{n_k}, n^{-1/2}) \subseteq (\mathscr{G}_k \times \mathbb{R}_+) \backslash (\mathscr{C}_{n,k} \times \left[\frac{1}{2n_k}, e^{3n_k \epsilon_{n_k}^2}\right]) \cup \mathscr{G}_{k,n}$ and $\mathscr{G}_{k,n} \subseteq \bigcup_{j \geq J} B_{k,j}$ for $\xi_{\min}^{1/2} M_n = JM$ and $\widetilde{M}/\sqrt{n_k} = J\varepsilon_\sigma$. Therefore, by decomposing the integral over $B_*^c(\epsilon_{n_k}, n^{-1/2})$ in the sum of two integrals over the ranges $(\mathscr{G}_k \times \mathbb{R}_+) \backslash (\mathscr{C}_{n,k} \times \left[\frac{1}{2n_k}, e^{3n_k \epsilon_{n_k}^2}\right])$ and $\mathscr{G}_{k,n}$ and by upper bounding $(1-\phi_{n_k})$ by 1 over $\mathscr{G}_k\backslash\mathscr{C}_{n,k}$, for $k=0,1$, we get:

$$
\begin{aligned}
&\int\int_{B_*^c(\epsilon_{n_k},n^{-1/2})}\int Q^{(n_k)}_{g_k,\sigma_k^2}(1-\phi_{n_k})\,d\pi(\boldsymbol{\alpha},g_k^{\perp},\lambda_k,\sigma_k^2)\prod_{i\in I_k}\mathscr{G}a\left(\xi_i;\frac{\nu}{2},\frac{\nu}{2}\right)d\xi_i\\
&\quad\leq \pi(\mathscr{G}_k\backslash\mathscr{C}_{n,k})\int_{(2n_k)^{-1}}^{e^{3n_k\epsilon_{n_k}^2}}\pi(\sigma_k^2)d\sigma_k^2+\underbrace{\pi(\mathscr{G}_k)}_{\leq 1}\left(\int_0^{(2n_k)^{-1}}\pi(\sigma_k^2)d\sigma_k^2+\int_{e^{3n_k\epsilon_{n_k}^2}}^{+\infty}\pi(\sigma_k^2)d\sigma_k^2\right)\\
&\qquad+\int\int_{\mathscr{G}_{k,n}}\int Q^{(n_k)}_{g_k,\sigma_k^2}(1-\phi_{n_k})\,d\pi(\boldsymbol{\alpha},g_k^{\perp},\lambda_k,\sigma_k^2)\prod_{i\in I_k}\mathscr{G}a\left(\xi_i;\frac{\nu}{2},\frac{\nu}{2}\right)d\xi_i\\
&\quad\leq e^{-n_k\epsilon_{n_k}^2\eta/(2\delta+1)}+e^{-n_k\epsilon_{n_k}^2\delta_{00}}+\frac{e^{-3n_k\epsilon_{n_k}^2}\delta_{00}}{\nu_{00}-2}+\int\sum_{j\geq J}e^{-Kj^2M^2n_k\epsilon_{n_k}^2}\prod_{i\in I_k}\mathscr{G}a\left(\xi_i;\frac{\nu}{2},\frac{\nu}{2}\right)d\xi_i
\end{aligned}
\tag{E.7}
$$

where we have used Lemma E.2 to upper bound $\pi(\mathscr{G}_k\backslash\mathscr{C}_{n,k})$, the concentration inequality for sub-Gamma random variables to upper bound the integral over $\left[0,\frac{1}{2n_k}\right]$ (since $(\sigma_k^{-2}-\mathbf{E}[\sigma_{\mathbf{k}}^{-\mathbf{2}}])$ is sub-Gamma $\left(2\frac{\nu_{00}}{\delta_{00}^2},\frac{2}{\delta_{00}}\right)$), and the second inequality in (E.3) with $JM=\xi_{\min}^{1/2}M_n$ to control the term $Q^{(n_k)}_{g_k,\sigma_k^2}(1-\phi_{n_k})$. By using the same argument to get (E.5) we obtain $\mathbf{E}[e^{-KJ^2M^2n_k\epsilon_{n_k}^2}]\leq 2\exp\{-n_k\epsilon_{n_k}^2(Kc_qM_n^2-\epsilon_{n_k})\}$. By putting this, (E.6) and (E.7) together we get that for $k=0,1$:

$$
\begin{aligned}
\mathbf{E}_*[\pi(B_*^c(\epsilon_{n_k},n^{-1/2})|X^{(n_k)})]&\lesssim e^{-n_k\epsilon_{n_k}^2(Kc_qM_n^2/(2J^2)-\epsilon_{n_k})}+\frac{1}{C^2n_k\epsilon_{n_k}^2}\\
&\quad+e^{(1+C+1)n_k\epsilon_{n_k}^2}\left(e^{-n_k\epsilon_{n_k}^2\eta/(2\delta+1)}+e^{-n_k\epsilon_{n_k}^2\delta_{00}}+e^{-3n_k\epsilon_{n_k}^2}+e^{-n_k\epsilon_{n_k}^2(Kc_qM_n^2-\epsilon_{n_k})}\right)
\end{aligned}
$$

which converges to zero for every M_n,K sufficiently large and every $0<C<\min\{1,\delta_{00}-2\}$ such that $1+C+1<\frac{\eta}{(2\delta+1)}$ which implies $\eta>\min\{3,\delta_{00}\}(2\delta+1)$. This establishes the statement of the theorem.

E.2 Proof of Theorem 2.2

For every n, denote $X^{(n)} \triangleq (y^{(n)}, z^{(n)})$. For given $m_0, m_1 \in \mathbb{N}$, define the functional spaces $\mathscr{G}_{m_0}$ and $\mathscr{G}_{m_1}$ as

$$\mathscr{G}_{m_0} \triangleq \Big\{g_{m_0}(z);\, g_{m_0}(z) = \boldsymbol{B}_0(z)'\boldsymbol{\alpha},\, z \in \{z_1, \dots, z_{n_0}\}, \boldsymbol{\alpha} \in \mathbb{R}^{m_0}\Big\} \subset L_2(P_{n_0})$$
$$\mathscr{G}_{m_1} \triangleq \Big\{g_{m_1}(z);\, g_{m_1}(z) = \boldsymbol{B}_1(z)'\boldsymbol{\beta},\, z \in \{z_{n_0+1}, \dots, z_{n_1}\}, \boldsymbol{\beta} \in \mathbb{R}^{m_1}\Big\} \subset L_2(P_{n_1}),$$

where P_{n_j} has support $z^{(n_j)}$ for $j = 0, 1$. Associated to each space (and then to each measure P_{n_j}), define the linear functionals

$$L_{z^{(n_0)}} : \mathscr{G}_{m_0} \to \mathbb{R}$$
$$g \mapsto \boldsymbol{e}'_{m_0,m_0}\mathbb{B}_0^{-1}\frac{1}{n_0}\sum_{i=1}^{n_0} \boldsymbol{B}_0(z_i)g(z_i)I[z_i < \tau]$$
$$L_{z^{(n_1)}} : \mathscr{G}_{m_1} \to \mathbb{R}$$
$$g \mapsto \boldsymbol{e}'_{m_1,1}\mathbb{B}_1^{-1}\frac{1}{n_1}\sum_{i=n_0+1}^{n_1} \boldsymbol{B}_1(z_i)g(z_i)I[z_i \geq \tau]$$

where $\boldsymbol{e}_{i,j}$ denotes the $(i \times 1)$ canonical vector with all components equal to zero but the j-th one, $\mathbb{B}_0 \triangleq \frac{1}{n_0}\sum_{i=1}^{n_0} \boldsymbol{B}_0(z_i)\boldsymbol{B}_0(z_i)'I[z_i < \tau]$ and $\mathbb{B}_1 \triangleq \frac{1}{n_1}\sum_{i=n_0+1}^{n_1} \boldsymbol{B}_1(z_i)\boldsymbol{B}_1(z_i)'I[z_i \geq \tau]$. Therefore, for every $g_{m_0} \in \mathscr{G}_{m_0}$, $g_{m_1} \in \mathscr{G}_{m_1}$, $\boldsymbol{\alpha}_{[m_0]} = L_{z^{(n_0)}}g_{m_0}$ and $\boldsymbol{\beta}_{[1]} = L_{z^{(n_1)}}g_{m_1}$. The linear functionals $L_{z^{(n_0)}}$ and $L_{z^{(n_1)}}$ can be written: $\forall \varphi_0 \in L_2(P_{n_0})$ and $\forall \varphi_1 \in L_2(P_{n_1})$,

$$L_{z^{(n_0)}}\varphi_0 = \langle \varphi_0, \ell_{z^{(n_0)}} \rangle_{P_{n_0}}, \qquad L_{z^{(n_1)}}\varphi_1 = \langle \varphi_1, \ell_{z^{(n_1)}} \rangle_{P_{n_1}}$$

where for $j = 0, 1$, $\langle \cdot, \cdot \rangle_{P_{n_j}}$ (resp. $\|\cdot\|_{n_j}$) denotes the scalar product (resp. the induced norm) in $L_2(P_{n_j})$ and

$$\ell_{z^{(n_0)}}(s) \triangleq \boldsymbol{e}'_{m_0,m_0}\mathbb{B}_0^{-1}\boldsymbol{B}_0(z)I[z < \tau] \in L_2(P_{n_0}), \quad \ell_{z^{(n_1)}}(s) \triangleq \boldsymbol{e}'_{m_1,1}\mathbb{B}_1^{-1}\boldsymbol{B}_1(z)I[z \geq \tau] \in L_2(P_{n_1}).$$

Therefore, by the Riesz theorem, $\|L_{z^{(n_0)}}\|^2_{n_0} = \|\ell_{z^{(n_0)}}\|^2_{n_0} = e'_{m_0,m_0}\mathbb{B}_0^{-1}e_{m_0,m_0}$ and $\|L_{z^{(n_1)}}\|^2_{n_1} = \|\ell_{z^{(n_1)}}\|^2_{n_1} = e'_{m_1,1}\mathbb{B}_0^{-1}e_{m_1,1}$. By the definition of $\boldsymbol{B}_j(z)$, $j = 0, 1$, there exists a constant $c_j > 0$ such that $\|\ell_{z^{(n_j)}}\|_{n_j} \leq c_j < \infty$ for every n_j.

Since $|ATE - ATE^*| \triangleq \Big|(\boldsymbol{\beta}_{[1]} - \boldsymbol{\alpha}_{[m_0]}) - (\boldsymbol{\beta}^*_{[1]} - \boldsymbol{\alpha}^*_{[m_0]})\Big| \leq |\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| + |\boldsymbol{\alpha}_{[m_0]} - \boldsymbol{\alpha}^*_{[m_0]}|$ it holds that (by denoting with Σ the event $\{|1-\sigma_j/\sigma_{j*}| \geq \widetilde{M}_n/\sqrt{n_j}$ for $j=0,1\}$)

$$\pi\left(|ATE - ATE^*| \geq M_{1,n}\epsilon_n,\, \underbrace{|1-\sigma_j/\sigma_{j*}| \geq \widetilde{M}_n/\sqrt{n_j} \text{ for } j=0,1}_{\triangleq \Sigma} \middle| X^{(n)}\right)$$
$$\leq \pi\left(|\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| + |\boldsymbol{\alpha}_{[m_0]} - \boldsymbol{\alpha}^*_{[m_0]}| \geq M_{1,n}\epsilon_n,\, \Sigma, |\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| \geq \frac{M_{1,n}\epsilon_n}{2} \middle| X^{(n)}\right)$$
$$+ \pi\left(|\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| + |\boldsymbol{\alpha}_{[m_0]} - \boldsymbol{\alpha}^*_{[m_0]}| \geq M_{1,n}\epsilon_n,\, \Sigma, |\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| < \frac{M_{1,n}\epsilon_n}{2} \middle| X^{(n)}\right). \quad \text{(E.8)}$$

We analyse the two terms in the right hand side of (E.8) separately by starting from the first one. Since for two events A and B, $\pi(A \cap B|X^{(n)}) \leq \pi(B|X^{(n)})$ and by using the independence of the posterior of (σ_0^2) and $(\boldsymbol{\beta}_{[1]}, \sigma_1^2)$ we get:

$$\pi\left(|\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| + |\boldsymbol{\alpha}_{[m_0]} - \boldsymbol{\alpha}^*_{[m_0]}| \geq M_{1,n}\epsilon_n,\, \Sigma, |\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| \geq \frac{M_{1,n}\epsilon_n}{2} \middle| X^{(n)}\right)$$
$$\leq \pi\left(\Sigma, |\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| \geq \frac{M_{1,n}\epsilon_n}{2} \middle| X^{(n)}\right)$$
$$\leq \pi\left(|1-\sigma_1^2/\sigma_{1*}^2| \geq \widetilde{M}_n\epsilon_n, |\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| \geq \frac{M_{1,n}\epsilon_n}{2} \middle| X^{(n_1)}\right) \underbrace{\pi\left(|1-\sigma_0/\sigma_{0*}| \geq \frac{\widetilde{M}_n}{\sqrt{n_j}} \middle| X^{(n_0)}\right)}_{\leq 1}$$
$$\leq \pi\left(\|L_{z^{(n_1)}}\|_{n_1} \|(g_{m_1} - g^*_{m_1})\|_n \geq \frac{M_{1,n}\epsilon_{n_1}}{2},\, \left|1 - \frac{\sigma_1}{\sigma_{1*}}\right| \geq \frac{\widetilde{M}_n}{\sqrt{n_j}} \middle| X^{(n_1)}\right) \quad \text{(E.9)}$$

which converges to zero in $P_*^{(n_1)}$-probability by the result of Theorem 2.1 with $M_n = M_{1,n}/(2c_1)$ and where we have used $\|L_{z^{(n_1)}}\|_{n_1} \leq c_1 < \infty$. We now analyse the second term on the right hand side of (E.8). By using the independence of the posterior of $(\boldsymbol{\alpha}_{[m_0]}, \sigma_0^2)$ and $(\boldsymbol{\beta}_{[1]}, \sigma_1^2)$ we get:

$$
\begin{aligned}
\pi & \left(|\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| + |\boldsymbol{\alpha}_{[m_0]} - \boldsymbol{\alpha}^*_{[m_0]}| \geq M_{1,n}\epsilon_n,\, \Sigma, |\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| < \frac{M_{1,n}\epsilon_n}{2} \right| X^{(n)} \right) \\
& \leq \pi \left(\frac{M_{1,n}\epsilon_n}{2} + |\boldsymbol{\alpha}_{[m_0]} - \boldsymbol{\alpha}^*_{[m_0]}| \geq M_{1,n}\epsilon_n,\, \Sigma, |\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| < \frac{M_{1,n}\epsilon_n}{2} \right| X^{(n)} \right) \\
& \leq \pi \left(|\boldsymbol{\alpha}_{[m_0]} - \boldsymbol{\alpha}^*_{[m_0]}| \geq \frac{M_{1,n}\epsilon_n}{2},\, \left|1 - \frac{\sigma_0}{\sigma_{0*}}\right| \geq \widetilde{M}_n/\sqrt{n_j} \right| X^{(n_0)} \right) \\
& \qquad \times \underbrace{\pi \left(|\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| < \frac{M_{1,n}\epsilon_n}{2},\, \left|1 - \frac{\sigma_1}{\sigma_{1*}}\right| \geq \widetilde{M}_n/\sqrt{n_j} \right| X^{(n_1)} \right)}_{\leq 1} \\
& \leq \pi \left(|L_{z^{(n_0)}}(g_{m_0} - g^*_{m_0})| \geq \frac{M_{1,n}\epsilon_{n_0}}{2},\, \left|1 - \frac{\sigma_0}{\sigma_{0*}}\right| \geq \widetilde{M}_n/\sqrt{n_j} \right| X^{(n_0)} \right) \\
& \leq \pi \left(\|L_{z^{(n_0)}}\|_n \|(g_{m_0} - g^*_{m_0})\|_n \geq \frac{M_{1,n}\epsilon_{n_0}}{2},\, \left|1 - \frac{\sigma_0}{\sigma_{0*}}\right| \geq \widetilde{M}_n/\sqrt{n_j} \right| X^{(n_0)} \right), \quad \text{(E.10)}
\end{aligned}
$$

where we have used the independence of the posterior of $(\boldsymbol{\alpha}_{[m_0]}, \sigma_0^2)$ and $(\boldsymbol{\beta}_{[1]}, \sigma_1^2)$ to get the second inequality. By the result of Theorem 2.1 with $M_n = M_{1,n}/(2c_0)$, (E.10) converges to zero in $P_*^{(n_0)}$-probability. By putting together (E.8), (E.9) and (E.10) we get the statement of the theorem.

E.3 Technical Lemmas

Lemma E.1. *Assume the conditions of Theorem 2.1 hold. Then, for $k = 0, 1$ there exists a $N > 0$ such that $\forall n_k \geq N$ and $\forall \epsilon_{n_k} > 0$*

$$
\pi(B_n^{KL}((g_k^*, \sigma_{k*}^2), \epsilon_{n_k})) \quad \gtrsim \quad \exp\left\{-n_k \epsilon_{n_k}^2\right\}. \quad \text{(E.11)}
$$

Proof. All along the proof k takes values in $\{0, 1\}$. By Lemma E.3 there exists a $N > 0$ such that $\forall n_k > N$ and $\forall \epsilon_{n_k} > 0$:

$$
\begin{aligned}
\pi(B_n^{KL}((g_k^*, \sigma_{k*}^2), \epsilon_{n_k})) & \gtrsim e^{\log \epsilon_{n_k}} \pi(g_k \in \mathscr{G}_k; \|g_k - g_k^*\|_\infty^2 \leq \sigma_{k*}^2 \epsilon_{n_k}^2 / 2) \\
& = e^{\log \epsilon_{n_k}} \pi \left\{ g_k \in \mathscr{G}_k; \|g_k - g_k^*\|_\infty \leq \frac{\sigma_{k*}\epsilon_{n_k}}{\sqrt{2}} \right\}. \quad \text{(E.12)}
\end{aligned}
$$

Because $\|g_k - g_k^*\|_\infty \leq \|g_k - g_{m_k}\|_\infty + \|g_{m_k} - g^*_{m_k}\| + \|g^*_{m_k} - g_k^*\|_\infty$, $\|g^*_{m_k} - g^*_k\|_\infty \leq C_2 m_k^{-\delta}$, and $\|g_{m_k} - g^*_{m_k}\|_\infty \leq C_3 m_k^{1/2} \|\boldsymbol{\alpha}_k - \boldsymbol{\alpha}_k^*\|$ for some constants $C_2, C_3 > 0$ (by boundedness of natural cubic splines), we have,

$$\pi\left\{g_k \in \mathscr{G}_k; \|g_k - g_k^*\|_\infty \le \frac{\sigma_{k*}\epsilon_{n_k}}{\sqrt{2}}\right\} \ge \pi\left\{\boldsymbol{\beta} \in \mathbb{R}^{m_k}; C_2 m_k^{-\delta} + C_3 m_k^{1/2}\|\boldsymbol{\alpha}_k - \boldsymbol{\alpha}_k^*\| \le \frac{\sigma_{k*}\epsilon_{n_k}}{\sqrt{2}}\right\}$$
$$= \sum_{m_k=1}^{\infty} \pi\left(\boldsymbol{\alpha}_k \in \mathbb{R}^{m_k}; C_2 m_k^{-\delta} + C_3 m_k^{1/2}\|\boldsymbol{\alpha_k} - \boldsymbol{\alpha_k}^*\|) \le \frac{\sigma_{k*}\epsilon_{n_k}}{\sqrt{2}} \middle| m_k\right)\pi(m_k) \quad \text{(E.13)}$$

where we have used the fact that the prior on $(g_k - g_{m_k})$ is degenerate on zero. Therefore, by setting $\epsilon_{n_k} = \frac{2C_2}{\sigma_{k*}}\left(\frac{\log n}{n}\right)^{\delta/(2\delta+1)}$, by replacing it and $m_k = m_k^* \asymp \left(\frac{n}{\log n}\right)^{1/(2\delta+1)}$ and since $\epsilon_{n_k} \asymp (m_k^*)^{-\delta}$ we obtain

$$\pi\left\{g_k \in \mathscr{G}_k; \|g_k - g_k^*\|_\infty \le \frac{\sigma_{k*}\epsilon_{n_k}}{\sqrt{2}}\right\} \quad \ge \quad \pi\left(\boldsymbol{\beta} \in \mathbb{R}^{m_k^*}; \|\boldsymbol{\beta} - \boldsymbol{\beta}^*\| \le (m_k^*)^{-\delta-1/2}\frac{C_2}{C_3}\right) \quad \text{(E.14)}$$

and denote $\widetilde{C} \triangleq C_2/C_3$. By using the result of Lemma G.1 with $U = \boldsymbol{\beta}$, which is valid under Assumption 4 and for every $(m_k^*)^{-\delta-1/2}\widetilde{C} \in (0, \sqrt{\rho_{\max} m_k^*/(4\lambda_k)})$, where $\rho_{\max}$ denotes the maximum eigenvalue of $D_\beta^{-1} T_\beta D_\beta^{-1'}$, we get:

$$\pi\left(\boldsymbol{\alpha_k} \in \mathbb{R}^{m_k};\ \|\boldsymbol{\alpha_k} - \boldsymbol{\alpha_k}^*\| \le (m_k^*)^{-\delta-1/2}\widetilde{C}\right) \ge \widetilde{C}^{m_k^*}\,(m_k^*)^{-(\delta+1/2)m_k^* - m_k^*/2}$$
$$\times \int_{\{\lambda_k \le \rho_{\max} m_k^*/(4\widetilde{C}^2)\}} \lambda_k^{m_k^*/2} e^{-\lambda_k\|w\|(m_k^*)^{-\delta-1/2}\widetilde{C}/(\rho_{\max}^{1/2})} e^{-\lambda_k\|w\|^2/2}\lambda_k^{a_{10}/2-1} e^{-b_{k0}\lambda_k/2}\frac{(b_{k0}/2)^{a_{k0}/2}}{\Gamma(a_{k0}/2)} d\lambda_k$$
$$= \widetilde{C}^{m_k^*}\,(m_k^*)^{-(\delta+1)m_k^*}\,\frac{(\widetilde{b}_{k0}/2)^{-\widetilde{a}_{k0}/2}(b_{k0}/2)^{a_{k0}/2}\Gamma(\widetilde{a}_{k0}/2)}{\Gamma(a_{k0}/2)}\widetilde{\pi}\left(\lambda_k \le \rho_{\max} m_k^*/(4\widetilde{C}^2)\right) \quad \text{(E.15)}$$

where $\widetilde{\pi}$ is taken with respect to a $Ga\left(\frac{\widetilde{a}_{k0}}{2}, \frac{\widetilde{b}_{k0}}{2}\right)$ distribution with $\widetilde{a}_{k0} \triangleq (a_{k0} + m_k^*)$ and $\widetilde{b}_{k0} \triangleq (b_{k0} + \|w\|^2) + 2\frac{\|w\|(m_k^*)^{-\delta-1/2}\widetilde{C}}{\sqrt{\rho_{\max}}}$, and where we have used the following inclusion of events: $\{(m_k^*)^{-(\delta+1/2)} \le \sqrt{\rho_{\max} m_k^*/(4\lambda_k\widetilde{C}^2)}\} \supseteq \{1 \le \sqrt{\rho_{\max} m_k^*/(4\lambda_k\widetilde{C}^2)}\}$ since $(m_k^*)^{-(\delta+1/2)} \le 1$.

We start with the computation of $\widetilde{\pi}\left(\lambda_k \le \rho_{\max} m_k^*/(4\widetilde{C}^2)\right)$. For this we can use the estimate of the tail of sub-gamma random variables[1] (see *e.g.* Boucheron et al. [2013]) to get

$$\widetilde{\pi}\left(\lambda_k > \frac{\rho_{\max} m_k^*}{4\widetilde{C}^2}\right) \leq \exp\left\{-\frac{k}{2}\left(\widetilde{b}_{k0}\frac{\rho_{\max} m_k^*}{4\widetilde{C}^2} + \sqrt{\widetilde{a}_{k0}\left(2\widetilde{b}_{k0}\frac{\rho_{\max} m_k^*}{4\widetilde{C}^2} - \widetilde{a}_{k0}\right)}\right)\right\} \leq \exp\left\{-\frac{\widetilde{b}_{k0}}{2}\frac{\rho_{\max} m_k^*}{4\widetilde{C}^2}\right\}.$$

Hence,

$$\widetilde{\pi}\left(\lambda_k \leq \rho_{\max} m_k^*/(4\widetilde{C}^2)\right) = 1 - \widetilde{\pi}\left(\lambda_k > \frac{\rho_{\max} m_k^*}{4\widetilde{C}^2}\right) \geq 1 - \exp\left\{-\frac{\widetilde{b}_{k0}}{2}\frac{\rho_{\max} m_k^*}{4\widetilde{C}^2}\right\} \geq \frac{1}{2} \quad \text{(E.16)}$$

for every n large enough.

We now analyse term $\frac{(\widetilde{b}_{k0}/2)^{-\widetilde{a}_{k0}/2}(b_{k0}/2)^{a_{k0}/2}\Gamma(\widetilde{a}_{k0}/2)}{\Gamma(a_{k0}/2)}$. For this, first remark that a version of the Stirling's formula gives the bound $\sqrt{2\pi}n^{n+1/2}e^{-n} \leq n! \leq e^{-n+1}n^{n+1/2}$. By using these bounds and $a_{k0} = 2\eta m_k^*$ we obtain: $\forall m_k^* \in \mathbb{N}$,

$$\begin{aligned}\frac{\Gamma(\widetilde{a}_{k0}/2)}{\Gamma(a_{k0}/2)} = \frac{(\widetilde{a}_{k0}-1)!}{(\frac{a_{k0}}{2}-1)!} &\geq \sqrt{2\pi}\left(\frac{\widetilde{a}_{k0}}{2}-1\right)^{(\widetilde{a}_{k0}-1)/2}\left(\frac{a_{k0}}{2}-1\right)^{-(a_{k0}-1)/2} e^{-(\widetilde{a}_{k0}/2-1)}e^{(a_{k0}-1)/2} \\ &= \sqrt{2\pi}e^{-(m_k^*-1)/2}\left(m_k^*\left(\eta+\frac{1}{2}-\frac{1}{m_k^*}\right)\right)^{(m_k^*(2\eta+1)-1)/2}\left(m_k^*\left(\eta-\frac{1}{m_k^*}\right)\right)^{-(2\eta m_k^*-1)/2} \\ &\geq \sqrt{2\pi}e^{-(m_k^*-1)/2}\left(m_k^*\left(\eta-\frac{1}{m_k^*}\right)\right)^{(m_k^*2\eta-1)/2}\left(m_k^*\left(\eta-\frac{1}{m_k^*}\right)\right)^{-(2\eta m_k^*-1)/2} = \sqrt{2\pi}e^{-m_k^*/2}\end{aligned} \quad \text{(E.17)}$$

where we have used the fact that $\eta + \frac{1}{2} - \frac{1}{m_k^*} > \eta - \frac{1}{m_k^*}$ to get the penultimate inequality and

[1] If λ_k is distributed as a $Ga\left(\frac{\widetilde{a}_{k0}}{2}, \frac{\widetilde{b}_{k0}}{2}\right)$ according to $\widetilde{\pi}$, then its centered version $\lambda_k - \mathbf{E}[\lambda_\mathbf{k}]$ its sub-Gamma $\left(2\frac{\widetilde{a}_{k0}}{\widetilde{b}_{k0}^2}, \frac{2}{\widetilde{b}_{k0}}\right)$, where $\mathbf{E}[\lambda_\mathbf{k}] = \widetilde{\mathbf{a}}_{\mathbf{k0}}/\widetilde{\mathbf{b}}_{\mathbf{k0}}$.

$(m_k^* - 1) < m_k^*$ in the last inequality. By putting (E.15)-(E.17) together we get:

$$
\begin{aligned}
\pi\left(\boldsymbol{\alpha_k} \in \mathbb{R}^{m_k};\ \|\boldsymbol{\alpha_k} - \boldsymbol{\alpha_k^*}\| \le (m_k^*)^{-\delta-1/2}\widetilde{C}\right) &\ge \exp\Big\{m_k^*\log(\widetilde{C}) - (\delta+1)m_k^*\log(m_k^*) + \frac{1}{2}\log(2\pi) \\
&\quad - \frac{1}{2}m_k^* - \frac{\widetilde{a}_{k0}}{2}\log(\widetilde{b}_{k0}/2) + \frac{a_{k0}}{2}\log(b_{k0}/2) - \log(2)\Big\} \\
&\ge \exp\Big\{ - m_k^*\log(n)\Big[\frac{\delta+1}{2\delta+1} - \frac{(\delta+1)\log\log(n)}{(2\delta+1)\log(n)} - \frac{\log(\widetilde{C})}{\log(n)} - \frac{\log(2\pi)}{2m_k^*\log(n)} + \frac{1}{\log(n)} \\
&\quad + \frac{2\eta+1}{2\log(n)}\log(\widetilde{b}_{k0}/2) - \frac{\eta}{\log(n)}\log(b_{k0}/2) + \frac{\log(2)}{m_k^*\log(n)}\Big]\Big\} \\
&\ge \exp\Big\{ - m_k^*\log(n)\Big[\frac{\delta+1}{2\delta+1} - 1 - 1 + 1 + 1 - 1 + 1\Big]\Big\} \gtrsim \exp\{-m_k^*\log(n)(\delta+1)/(2\delta+1)\}
\end{aligned}
$$

where we have used $b_{k0} > 2$ and n large to get the first inequality in the last line. Therefore, if $m_k^* \asymp \left(\frac{n}{\log n}\right)^{1/(2\delta+1)}$

$$
\begin{aligned}
\pi(B_n^{KL}((g_k^*, \sigma_{k*}^2), \epsilon_{n_k})) &\gtrsim e^{\log\epsilon_{n_k}}\exp\left\{-m_j^*\log(n)\frac{\delta+1}{2\delta+1}\right\} \\
&= \exp\left\{-n_k\epsilon_{n_k}^2\left(\frac{\delta+1}{2\delta+1} - \frac{\log(\epsilon_{n_k})}{n\epsilon_{n_k}^2}\right)\right\} \ge \exp\left\{-n_k\epsilon_{n_k}^2\frac{2\delta+1}{2\delta+1}\right\} \quad \text{(E.18)}
\end{aligned}
$$

since for the values of m_k^* and ϵ_{n_k} given in the theorem $m_k^*\log n \asymp n_k\epsilon_{n_k}^2$, and since $\log(\epsilon_{n_k})/(n_k\epsilon_{n_k}^2) \ge -\delta/(2\delta+1)$. This establishes (E.11).

□

Lemma E.2. *Assume the conditions of Theorem 2.1 hold. Then, for $k = 0, 1$ the sequence of measurable sets $\mathscr{C}_{n,k}$ defined in* (D.2) *satisfies*

$$
\pi(\mathscr{G}_k \backslash \mathscr{C}_{n,k}) \lesssim \exp\left\{-n_k\epsilon_{n_k}^2\frac{\eta}{2\delta+1}\right\}, \quad \text{(E.19)}
$$

where η is defined in Theorem 2.1.

Proof. For $k = 0, 1$, take as $\mathscr{C}_{n,k}$ the set defined in (D.2) and recall the notation $\boldsymbol{\alpha}_k \triangleq \boldsymbol{\alpha}I[k = 0] + \boldsymbol{\beta}I[k = 1]$. The complement $\mathscr{G}_k \backslash \mathscr{C}_{n,k}$ is

$$
\mathscr{G}_k \backslash \mathscr{C}_{n,k} = \mathscr{G}_k \backslash \mathscr{S}_{m_k^*} \cup \bigcup_{m_k=1}^{m_k^*}\Big\{g_{m_k} \in \mathscr{S}_{m_k},\ \|\boldsymbol{\alpha}_k\|_\infty > M_k,\Big\},
$$

where $g_{m_k}(z) \triangleq \boldsymbol{B}_k(z)'\boldsymbol{\alpha}_k$. By using the fact that the prior has nonzero mass only on the subspace of natural cubic splines on the set of $m_k = m_k^*$ knots for $k = 0, 1$, the prior of this event is

$$\pi(\mathscr{G}_k \backslash \mathscr{C}_{n,k}) \leq \pi\Big(g_{m_k} \in \mathcal{S}_{m_k};\ \|\boldsymbol{\alpha}_k\|_\infty > M_k\Big). \tag{E.20}$$

From Lemma G.2 applied with $U = \boldsymbol{\alpha_k}$, $m = m_k^*$ and $M = M_k$, which is valid under Assumption 4 (ii), we know that $\pi(\|\boldsymbol{\alpha}_k\|_\infty > M_k|\lambda_k) \leq \exp\left\{\log(m_k^*) - \frac{\lambda_k(M_k/2 - \|\boldsymbol{D_{\alpha_k}}^{-1}\boldsymbol{\alpha}_{k0}\|_\infty^2)}{2\gamma}\right\}$ for $M_k \geq 1$ and $\gamma \triangleq \max_{1\leq j\leq m_k^*} |(\boldsymbol{D_{\alpha_k}}^{-1}\boldsymbol{T_{\alpha_k}}\boldsymbol{D_{\alpha_k}}^{-1'})_{jj}|$. Hence, by taking M_k sufficiently large to guarantee that $M_k/2 - \|\boldsymbol{D_{\alpha_k}}^{-1}\boldsymbol{\alpha}_{k0}\|_\infty^2 + 2\gamma b_{k0} > 0$, we have

$$\begin{aligned}\pi(\|\boldsymbol{\alpha_k}\|_\infty > M_k) &\leq e^{\log(m_k^*)}\bigg(\int_0^\infty e^{-\left\{\lambda_k(M_k/2 - \|D_{\alpha_k}^{-1}\boldsymbol{\alpha}_{k0}\|_\infty^2)/(2\gamma)\right\}}\mathscr{G}a\left(\lambda_k; \frac{a_{k0}}{2}, \frac{b_{k0}}{2}\right) d\lambda_k \\ &= \exp\left\{\log(m_k^*) - \frac{a_{k0}}{2}\log\left(\frac{b_{k0}}{2} + \frac{M_k/2 - \|\boldsymbol{D_{\alpha_k}}^{-1}\boldsymbol{\alpha}_{k0}\|_\infty^2}{2\gamma}\right) + \frac{a_{k0}}{2}\log\left(\frac{b_{k0}}{2}\right)\right\}.\end{aligned}$$

By replacing the values $a_{k0} = 2\eta m_k^*$, $M_k = 4\gamma n_k\epsilon_{n_k}^2 \asymp n_k\epsilon_{n_k}^2$ and the value for ϵ_{n_k} we obtain:

$$\begin{aligned}&\pi(\|\boldsymbol{\alpha_k}\|_\infty > M_k) \leq \\ &\exp\left\{\log(m_k^*) - \eta m_k^*\log\left(n_k\epsilon_{n_k}^2\left(\frac{b_{k0}}{2n_k\epsilon_{n_k}^2} + 1 - \frac{\|\boldsymbol{D_{\alpha_k}}^{-1}\boldsymbol{\alpha}_{k0}\|_\infty^2}{2\gamma n_k\epsilon_{n_k}^2}\right)\right) + \eta m_k^*\log\left(\frac{b_{k0}}{2}\right)\right\} \\ &= \exp\left\{-m_k^*\log(n_k)\left(\frac{\eta}{2\delta+1} + o(1)\right)\right\} \asymp \exp\left\{-n_k\epsilon_{n_k}^2\frac{\eta}{2\delta+1}\right\}.\end{aligned} \tag{E.21}$$

□

Lemma E.3. *Let us consider model* (E.1)-(E.2)*. For $j = 0, 1$, suppose that $\varepsilon_{ji} \sim t_\nu(0, 1)$, $\forall i \in I_j$, where 0 and 1 denote the location and scale parameters respectively of a Student distribution, and suppose that ε_{0i} and ε_{1i} are independent between them and across i. Let $B_n^{KL}((f_j^*, \sigma_{j*}^2), \epsilon)$ be as defined in* (D.1) *and π be an independent prior on g_j and σ_j^2. Then, for $j = 0, 1$ there exists a $N > 0$ such that $\forall n_j \geq N$ and $\forall \epsilon_{n_j} > 0$*

$$\pi(B_n^{KL}((g_j^*, \sigma_{j*}^2), \epsilon_{n_j})) \geq \pi\left(|\sigma_{j*}^2 - \sigma_j^2| \leq \epsilon_{n_j}\sigma_{j*}^2/\sqrt{2\nu},\ \sigma_j^2 \geq \sigma_{j*}^2\right) \times \pi\left(\|g_j - g_j^*\|_\infty^2 \leq \sigma_{j*}^2\epsilon_{n_j}^2/2\right). \tag{E.22}$$

Moreover, let the prior on σ_j^2 be specified as in (D.6)*. Then, there exists a $N > 0$ such that $\forall n \geq N$ and $\forall \epsilon_{n_j} > 0$*

$$\pi(B_n^{KL}((g_j^*, \sigma_{j*}^2), \epsilon_{n_j})) \gtrsim e^{\log\epsilon_{n_j}}\pi\left(\|g_j - g_j^*\|_\infty^2 \leq \sigma_{j*}^2\epsilon_{n_j}^2/2\right). \tag{E.23}$$

Proof. All along the proof j take values in $\{0,1\}$. From the distribution of ε_{ji} it follows that $y_i|z_i; g_j, \sigma_j^2 \sim t_\nu(y_i|g_j(z_i), \sigma_j)$ where σ_j is the scale parameter and $g_j(z_i)$ is the location parameter, for every $i \in I_j$. Simple algebra shows that $K\left(P_*^{(n_j)}, P_{g_j,\sigma_j^2}^{(n_j)}\right) = \sum_{i\in I_j} K\left(P_*^i, P_{g_j,\sigma_j^2}^i\right)$ and that

$$V\left(P_*^{(n_j)}, P_{g_j,\sigma_j^2}^{(n_j)}\right) = \int \left|\sum_{i\in I_j}\left(\log\frac{dP_*^i}{dP_{g_j,\sigma_j^2}^i} - K\left(P_*^i, P_{g_j,\sigma_j^2}^i\right)\right)\right|^2 dP_*^{(n_j)}$$

$$= \int \sum_{i\in I_j}\left(\log\frac{dP_*^i}{dP_{g_j,\sigma_j^2}^i} - K\left(P_*^i, P_{g_j,\sigma_j^2}^i\right)\right)^2 dP_*^{(n_j)} + 2\sum_{i>l} Cov_*\left(\log\frac{dP_*^i}{dP_{g_j,\sigma_j^2}^i}, \log\frac{dP_*^l}{dP_{g_j,\sigma_j^2}^l}\right)$$

$$= \sum_{i\in I_j} V(P_*^i, P_{g_j,\sigma_j^2}^i)$$

where Cov_* denotes the covariance with respect to $P_*^{(n_j)}$. Therefore,

$$B_n^{KL}((g_j^*, \sigma_{j*}^2), \epsilon_{n_j}) =$$
$$\left\{(g_j, \sigma_j^2) \in \mathscr{G}_j \times \mathbb{R}_+; \sum_{i\in I_j} K(P_*^i, P_{g_j,\sigma_j^2}^i) \le n_j\epsilon_{n_j}^2, \sum_{i\in I_j} V(P_*^i, P_{g_j,\sigma_j^2}^i) \le 2n_j\epsilon_{n_j}^2\right\}$$

for some constant $0 < C_2 \le 2$. In the following of the proof we leave implicit the fact that $(g_j, \sigma_j^2) \in \mathscr{G}_j \times \mathbb{R}_+$ in all the events that we define. Denote $\mathscr{A}_1 \triangleq \{\frac{1}{n_j}\sum_{i\in I_j}^n K(P_*^i, P_{g_j,\sigma_j^2}^i) \le \epsilon_{n_j}^2\}$. We first upper bound $K(P_*^i, P_{g_j,\sigma_j^2}^i)$ in order to find an event included in $\mathscr{A}_1$. Denote $\widetilde{\alpha} = (\nu+1)/2$, $g_{j,i}^* = g_j^*(z_i)$ and $g_{j,i} = g_j(z_i)$. Then, for $j = 0,1$:

$$K(P_*^i, P_{g_j,\sigma_j^2}^i) = \int \ln\left[\frac{t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)}{t_\nu(y_i|g_{j,i}, \sigma_j^2)}\right] t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)dy_i$$

$$= \int \ln\left(\frac{\sigma_j}{\sigma_{j*}}\frac{\left(1+\frac{(y_i-g_{j,i}^*)^2}{\nu\sigma_{j*}^2}\right)^{-\widetilde{\alpha}}}{\left(1+\frac{(y_i-g_{j,i})^2}{\nu\sigma_j^2}\right)^{-\widetilde{\alpha}}}\right) t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)dy_i$$

$$= \int \ln\left(\frac{\sigma_j^{-\nu}}{\sigma_{j*}^{-\nu}}\left(\frac{\nu\sigma_j^2 + (y_i-g_{j,i})^2}{\nu\sigma_{j*}^2 + (y_i-g_{j,i}^*)^2}\right)^{\widetilde{\alpha}}\right) t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)dy_i.$$

Since for $j = 0,1$, $\sigma_{j*}/\sigma_j > 0$ and $(\nu\sigma_j^2 + (y_i - g_{j,i})^2)/(\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2) > 0$, we use the inequality $\ln(x) \leq x - 1, \forall x > 0$, and obtain:

$$
\begin{aligned}
K(P_*^i, P_{g_j,\sigma_j^2}^i) &\leq \frac{\nu}{2} \int \frac{\sigma_{j*}^2 - \sigma_j^2}{\sigma_j^2} t_\nu(y_i | g_{j,i}^*, \sigma_{j*}) dy_i \\
&\quad + \widetilde{\alpha} \int \left(\frac{\nu\sigma_j^2 + (y_i - g_{j,i})^2 - \nu\sigma_{j*}^2 - (y_i - g_{j,i}^*)^2}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2} \right) t_\nu(y_i | g_{j,i}^*, \sigma_{j*}^2) dy_i \\
&= \frac{\nu(\sigma_{j*}^2 - \sigma_j^2)}{2\sigma_j^2} + \widetilde{\alpha} \int \left(\frac{\nu\sigma_j^2 + (y_i - g_{j,i})^2 - \nu\sigma_{j*}^2 - (y_i - g_{j,i}^*)^2}{\nu\sigma_{j*}^2} \right) \frac{\nu\sigma_{j*}^2 t_\nu(y_i | g_{j,i}^*, \sigma_{j*}^2)}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2} dy_i.
\end{aligned}
\tag{E.24}
$$

Remark that $\frac{\nu\sigma_{j*}^2}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2} t_\nu(y_i | g_{j,i}^*, \sigma_{j*}^2) = t_{\nu+2}(y_i | g_{j,i}^*, \breve{\sigma}_{j*}^2) \frac{\nu}{\nu+1}$ with $\breve{\sigma}_{j*}^2 \triangleq \sigma_{j*}^2 \nu / (\nu + 2)$. In fact,

$$
\begin{aligned}
\frac{\nu\sigma_{j*}^2}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2} t_\nu(y_i | g_{j,i}^*, \sigma_{j*}^2) &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}\sigma_*} \left(1 + \frac{(y_i - g_{j,i}^*)^2(\nu+2)}{\nu\sigma_{j*}^2(\nu+2)}\right)^{-(\nu+1)/2-1} \\
&= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}\sigma_{j*}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)\sqrt{\pi(\nu+2)}\breve{\sigma}_{j*}}{\Gamma\left(\frac{\nu+2+1}{2}\right)} t_{\nu+2}(y_i | g_{j,i}^*, \breve{\sigma}_{j*}^2) \\
&= \frac{\nu}{\nu+1} t_{\nu+2}(y_i | g_{j,i}^*, \breve{\sigma}_{j*}^2),
\end{aligned}
\tag{E.25}
$$

where we have used $\Gamma(z+1) = z\Gamma(z)$. By plugging this result in (E.24) we obtain:

$$
\begin{aligned}
K(P_*^i, P_{g_j,\sigma_j^2}^i) &\leq \frac{\nu(\sigma_{j*}^2 - \sigma_j^2)}{2\sigma_j^2} + \frac{\nu}{2} \int \left(\frac{\nu\sigma_j^2 + (y_i - g_{j,i})^2 - \nu\sigma_{j*}^2 - (y_i - g_{j,i}^*)^2}{\nu\sigma_{j*}^2} \right) t_{\nu+2}(y_i | g_{j,i}^*, \breve{\sigma}_{j*}^2) dy_i \\
&= \frac{\nu(\sigma_{j*}^2 - \sigma_j^2)^2}{2\sigma_{j*}^2\sigma_j^2} + \frac{1}{2} \int \left(\frac{2(y_i - g_{j,i}^*)(g_{j,i}^* - g_{j,i}) + (g_{j,i}^* - g_{j,i})^2}{\sigma_{j*}^2} \right) t_{\nu+2}(y_i | g_{j,i}^*, \breve{\sigma}_{j*}^2) dy_i \\
&= \frac{1}{2\sigma_{j*}^2} \left(\frac{\nu(\sigma_{j*}^2 - \sigma_j^2)^2}{\sigma_j^2} + (g_{j,i}^* - g_{j,i})^2 \right)
\end{aligned}
\tag{E.26}
$$

since $(y_i - g_{j,i}^*)$ has mean equal to zero. It follows that

$$
\sum_{i \in I_j} K(P_*^i, P_{g_j,\sigma_j^2}^i) = \frac{n_j}{2\sigma_{j*}^2} \left(\frac{\nu(\sigma_{j*}^2 - \sigma_j^2)^2}{\sigma_j^2} + \|g_j^* - g_j\|_n^2 \right). \tag{E.27}
$$

Therefore,

$$\mathscr{A}_1 \supseteq \left\{ \|g_j^* - g_j\|_{n_j}^2 \leq \frac{\sigma_{j*}^2 \epsilon_{n_j}^2}{2} \right\} \cap \left\{ \frac{(\sigma_{j*}^2 - \sigma_j^2)^2}{\sigma_j^2} \leq \frac{\sigma_{j*}^2 \epsilon_{n_j}^2}{2\nu} \right\} \triangleq \mathscr{A}_2.$$

In addition, denote $\mathscr{A}_3 \triangleq \{|\sigma_{j*}^2 - \sigma_j^2| \leq \epsilon_{n_j}\sigma_{j*}^2/\sqrt{2\nu},\ \sigma_j^2 \geq \sigma_{j*}^2\}$ and $\mathscr{A}_4 \triangleq \{\|g_j^* - g_j\|_n^2 \leq \sigma_{j*}^2\epsilon_{n_j}^2/2\}$ so that $\mathscr{A}_2 \supseteq \mathscr{A}_3 \cap \mathscr{A}_4$ which we will use later on. Moreover, define the event $\mathscr{A}_5 \triangleq \{\|g_j - g_j^*\|_\infty^2 \leq \sigma_{j*}^2\epsilon_{n_j}^2/2\}$.

Next, we show that $\{\sum_{i\in I_j} V(P_*^i, P_{g_j,\sigma_j^2}^i) \leq 2n_j\epsilon_{n_j}^2\} \supseteq \mathscr{A}_3\cap\mathscr{A}_4\cap\mathscr{A}_5$. On $\mathscr{A}_3\cap\mathscr{A}_4\cap\mathscr{A}_5$, $(g_{j,i}, \sigma_j^2)$ is close to $(g_{j,i}^*, \sigma_{j*}^2)$ for every $i \in I_j = 1, \dots, n$. Therefore, on $\mathscr{A}_3 \cap \mathscr{A}_4 \cap \mathscr{A}_5$, $P_{g_j,\sigma_j^2}^i$ is close to P_*^i and we can use the Taylor expansion of $\ln(u)$ around $u = 1$: $\ln(u) = (u-1) + o(u-1)$ with $u = \frac{\nu\sigma_j^2 + (y_i - g_{j,i})^2}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2}$. In particular, on $\mathscr{A}_3\cap\mathscr{A}_4\cap\mathscr{A}_5$, $u-1 \leq \epsilon_{n_j}(1/\sqrt{\nu} + \epsilon_{n_j}/\nu - 2(y_i - g_{j,i}^*)/(\nu\sigma_{j*}))$ which converges to zero as $\epsilon_{n_j} \to 0$. We use this result to deal with $V(P_*^i, P_{g_j,\sigma_j^2}^i)$ which we first develop as:

$$\begin{aligned}
V(P_*^i, P_{g_j,\sigma_j^2}^i) &= \int \left| \ln\left[\frac{t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)}{t_\nu(y_i|g_{j,i}, \sigma_j^2)}\right] - K(P_*^i, P_{g_j,\sigma_j^2}^i) \right|^2 t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)dy_i \\
&= \int \Bigg| \ln\left[\frac{\sigma_j}{\sigma_{j*}}\left(\frac{\nu\sigma_j^2 + (y_i - g_{j,i})^2}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2}\right)^{(\nu+1)/2}\left(\frac{\sigma_{j*}}{\sigma_j}\right)^{(\nu+1)}\right] \\
&\quad - \int \ln\left[\left(\frac{\sigma_{j*}}{\sigma_j}\right)^\nu \left(\frac{\nu\sigma_j^2 + (y_i - g_{j,i})^2}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2}\right)^{(\nu+1)/2}\right] t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)dy_i \Bigg|^2 t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)dy_i \\
&= \frac{(\nu+1)^2}{4} \int \left| \ln\left(\frac{\nu\sigma_j^2 + (y_i - g_{j,i})^2}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2}\right) - \int \ln\left(\frac{\nu\sigma_j^2 + (y_i - g_{j,i})^2}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2}\right) t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)dy_i \right|^2 \\
&\qquad \times t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)dy_i.
\end{aligned}$$

Then, by applying the Taylor expansion of $\ln(u)$ as indicated above we get:

$$\begin{aligned}
V(P_*^i, P_{g_j,\sigma_j^2}^i) &= \frac{(\nu+1)^2}{4} \int \Bigg| \frac{\nu(\sigma_j^2 - \sigma_{j*}^2) + (g_{j,i} - g_{j,i}^*)^2 + 2(y_i - g_{j,i}^*)(g_{j,i}^* - g_{j,i})}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2} \\
&- \int \frac{\nu(\sigma_j^2 - \sigma_{j*}^2) + (g_{j,i} - g_{j,i}^*)^2 + 2(y_i - g_{j,i}^*)(g_{j,i}^* - g_{j,i})}{\nu\sigma_{j*}^2} \frac{\nu\sigma_{j*}^2}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2} t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)dy_i \Bigg|^2 \\
&\qquad \times t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)dy_i
\end{aligned}$$

By using the fact that $\frac{\nu\sigma_{j*}^2}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2} t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2) = t_{\nu+2}(y_i|g_{j,i}^*, \check{\sigma}_{j*}^2)\frac{\nu}{\nu+1}$ with $\check{\sigma}_{j*}^2 \triangleq \sigma_{j*}^2\nu/(\nu+2)$

the inner integral is equal to $[\nu(\sigma_j^2 - \sigma_{j*}^2) + (g_{j,i} - g_{j,i}^*)^2]/(\sigma_{j*}^2(\nu+1))$. Moreover, by using the fact that

$$\left(\frac{\nu\sigma_{j*}^2}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2}\right)^2 t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}\sigma_{j*}}\left(1 + \frac{(y_i - g_{j,i}^*)^2(\nu+4)}{\nu\sigma_{j*}^2(\nu+4)}\right)^{-(\nu+1)/2-2}$$
$$= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}\sigma}\frac{\Gamma\left(\frac{\nu+4}{2}\right)\sqrt{\pi(\nu+4)}\overline{\sigma}_{j*}}{\Gamma\left(\frac{\nu+4+1}{2}\right)} t_{\nu+4}(y_i|g_{j,i}^*, \overline{\sigma}_{j*}^2) = \frac{(\nu+2)\nu}{(\nu+3)(\nu+1)} t_{\nu+4}(y_i|g_{j,i}^*, \overline{\sigma}_{j*}^2),$$

with $\overline{\sigma}_{j*}^2 = \nu\sigma_{j*}^2/(\nu+4)$ we obtain:

$$V(P_*^i, P_{g_j,\sigma_j^2}^i) = \frac{(\nu+1)^2}{4}\times$$
$$\Big(\int \frac{[\nu(\sigma_j^2 - \sigma_{j*}^2) + (g_{j,i} - g_{j,i}^*)^2 + 2(y_i - g_{j,i}^*)(g_{j,i}^* - g_{j,i}^*)]^2}{\nu^2\sigma_{j*}^4}\left(\frac{\nu\sigma_{j*}^2}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2}\right)^2 t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)dy_i$$
$$+ \frac{\left[\nu(\sigma_j^2 - \sigma_{j*}^2) + (g_{j,i} - g_{j,i}^*)^2\right]^2}{\sigma_{j*}^4(\nu+1)^2} - 2\frac{[\nu(\sigma_j^2 - \sigma_{j*}^2) + (g_{j,i} - g_{j,i}^*)^2]}{\sigma_{j*}^2(\nu+1)}\times$$
$$\int \frac{[\nu(\sigma_j^2 - \sigma_{j*}^2) + (g_{j,i} - g_{j,i}^*)^2 + 2(y_i - g_{j,i}^*)(g_{j,i}^* - g_{j,i}^*)]}{\nu\sigma_{j*}^2}\frac{\nu\sigma_{j*}^2}{\nu\sigma_{j*}^2 + (y_i - g_{j,i}^*)^2} t_\nu(y_i|g_{j,i}^*, \sigma_{j*}^2)dy_i\Big)$$

By computing the integral this gives:

$$V(P_*^i, P_{g_j,\sigma_j^2}^i) = \frac{(\nu+1)^2}{4}\Big(\frac{\nu(\sigma_j^2 - \sigma_{j*}^2)^2(\nu+2)}{\sigma_{j*}^4(\nu+3)(\nu+1)} + \frac{(g_{j,i}^* - g_{j,i})^4\nu(\nu+2)}{\nu^2\sigma_{j*}^4(\nu+3)(\nu+1)} + \frac{4(g_{j,i}^* - g_{j,i})^2\sigma_{j*}^2}{\sigma_{j*}^4(\nu+3)(\nu+1)}$$
$$+ \frac{2(\sigma_j^2 - \sigma_{j*}^2)(g_{j,i}^* - g_{j,i})^2}{\nu\sigma_{j*}^4} - \frac{[\nu(\sigma_j^2 - \sigma_{j*}^2) + (g_{j,i}^* - g_{j,i})^2]^2}{\sigma_{j*}^4(\nu+1)^2}\Big)$$
$$= \Big(\frac{(\sigma_j^2 - \sigma_{j*}^2)^2}{\sigma_{j*}^4}\left[\frac{\nu(\nu+2)}{(\nu+3)(\nu+1)} - \frac{\nu^2}{(\nu+1)^2}\right] + \frac{(g_{j,i} - g_{j,i}^*)^4}{\sigma_{j*}^4}\left[\frac{\nu+2}{\nu(\nu+3)(\nu+1)} - \frac{1}{(\nu+1)^2}\right]$$
$$+ 2\frac{(g_{j,i} - g_{j,i}^*)^2}{\sigma_{j*}^2}\left(\frac{2}{(\nu+3)(\nu+1)} + \frac{\sigma_j^2 - \sigma_{j*}^2}{\nu\sigma_{j*}^2} - \frac{\nu(\sigma_j^2 - \sigma_{j*}^2)}{(\nu+1)^2\sigma_{j*}^2}\right)\Big)\frac{(\nu+1)^2}{4}$$
$$\leq \frac{\nu(\sigma_j^2 - \sigma_{j*}^2)^2}{2\sigma_{j*}^4} + \frac{(g_{j,i} - g_{j,i}^*)^4}{2\sigma_{j*}^4\nu} + \frac{(g_{j,i} - g_{j,i}^*)^2}{\sigma_{j*}^2} + \frac{(g_{j,i} - g_{j,i}^*)^2(\sigma_j^2 - \sigma_{j*}^2)}{\sigma_{j*}^4}\frac{(1+2\nu)}{2\nu}.$$

Therefore, by using the Holder inequality which implies $\frac{1}{n_j}\sum_{i=1}^{n_j}(g_{j,i} - g_{j,i}^*)^4 \leq \|g_j - g_j^*\|_\infty^2 \|g_j -$

$g_j^*\|_{n_j}^2$ and because we are on $\mathscr{A}_3 \cap \mathscr{A}_4$, we get

$$\sum_{i\in I_j} V(P_*^i, P_{g_j,\sigma_j^2}^i) \leq n_j\Big(\frac{\nu(\sigma_j^2-\sigma_{j*}^2)^2}{2\sigma_{j*}^4} + \frac{\|g_j-g_j^*\|_\infty^2\|g_j-g_j^*\|_{n_j}^2}{2\sigma_{j*}^4\nu}$$
$$+\frac{\|g_j-g_j^*\|_{n_j}^2}{\sigma_{j*}^2} + \frac{\|g_j-g_j^*\|_{n_j}^2(\sigma_j^2-\sigma_{j*}^2)}{\sigma_{j*}^4}\frac{(1+2\nu)}{2\nu}\Big)$$
$$\leq n_j\left(\frac{\epsilon_{n_j}^2}{2} + \frac{\epsilon_{n_j}^2}{4\nu\sigma_{j*}^2}\|g_j-g_j^*\|_\infty^2 + \frac{\epsilon_{n_j}^2}{2} + \epsilon_{n_j}^3\frac{(1+2\nu)}{4\nu\sqrt{2\nu}}\right).$$

Moreover, on $\mathscr{A}_5$:

$$\sum_{i\in I_j} V(P_*^i, P_{g_j,\sigma_j^2}^i) \leq n\epsilon_{n_j}^2\left(\frac{1}{2} + \frac{\epsilon_{n_j}^2}{8\nu} + \frac{1}{2} + \epsilon_{n_j}\frac{(1+2\nu)}{4\nu\sqrt{2\nu}}\right) \lesssim n\epsilon_{n_j}^2.$$

This shows that there exists an $N>0$ such that $\forall n \geq N$ and $\forall \epsilon_{n_j} > 0$, $\{\sum_{i\in I_j}^n V(P_*^i, P_{g_j,\sigma_j^2}^i) \leq 2n\epsilon_{n_j}^2\} \supseteq \mathscr{A}_3 \cap \mathscr{A}_4 \cap \mathscr{A}_5$. Then, by remarking that

$$\pi(B_n^{KL}((g_j^*,\sigma_{j*}^2),\epsilon_{n_j})) \geq \int_{\mathscr{A}_3\cap\mathscr{A}_4\cap\{\sum_{i\in I_j} V(P_*^i,P_{g_j,\sigma_j^2}^i)\leq 2n\epsilon_{n_j}^2\}} d\pi(g_j,\sigma_j^2)$$
$$= \int_{\mathscr{A}_3\cap\mathscr{A}_4\cap\{\sum_{i\in I_j}^n V(P_*^i,P_{g_j,\sigma_j^2}^i)\leq 2n_j\epsilon_{n_j}^2\}\cap\mathscr{A}_5} d\pi(g_j,\sigma_j^2) + \int_{\mathscr{A}_3\cap\mathscr{A}_4\cap\{\sum_{i=1}^n V(P_*^i,P_{g_j,\sigma_j^2}^i)\leq 2n\epsilon_{n_j}^2\}\cap\mathscr{A}_5^c} d\pi(g_j,\sigma_j^2)$$
$$\geq \int_{\mathscr{A}_3\cap\mathscr{A}_4\cap\{\sum_{i\in I_j}^n V(P_*^i,P_{g_j}^i)\leq 2n\epsilon_{n_j}^2\}\cap\mathscr{A}_5} d\pi(g_j,\sigma_j^2). \quad \text{(E.28)}$$

we get that for $j=0,1$:

$$\pi(B_n^{KL}((g_j^*,\sigma_{j*}^2),\epsilon_{n_j})) \geq \pi\left(\mathscr{A}_3\cap\mathscr{A}_4\cap\mathscr{A}_5\right)$$
$$\geq \pi\left(|\sigma_{j*}^2-\sigma_j^2| \leq \epsilon_{n_j}\sigma_{j*}^2/\sqrt{2\nu},\ \sigma_j^2 \geq \sigma_{j*}^2\right)\pi\left(\|g_j-g_j^*\|_\infty^2 \leq \sigma_*^2\epsilon_{n_j}^2\right),$$

where in the last line we have used the independence of the prior for σ_j^2 and g_j and the inequality $\|g_j-g_j^*\|_{n_j}^2 \leq \|g_j-g_j^*\|_\infty^2$ for $j\in\{0,1\}$ which implies $\mathscr{A}_5 \subseteq \mathscr{A}_4$. This establishes result (E.22) in the lemma. To show result (E.23):

$$\pi\left(|\sigma_{j*}^2-\sigma_j^2|\le \epsilon_{n_j}\sigma_{j*}^2/\sqrt{\nu},\sigma_j^2\ge\sigma_{j*}^2\right)=\int_{(\sigma_{j*}^2-\frac{\epsilon_{n_j}\sigma_{j*}^2}{\sqrt{\nu}})\vee\sigma_{j*}^2}^{(\sigma_{j*}^2+\frac{\epsilon_{n_j}\sigma_{j*}^2}{\sqrt{\nu}})}\frac{(\delta_{00}/2)^{\nu_{00}/2}}{\Gamma(\nu_{00}/2)}(\sigma_j^2)^{-\nu_{00}/2-1}e^{-\delta_{00}/(2\sigma_j^2)}d\sigma_j^2$$

$$\ge(\sigma_{j*}^2+\frac{\epsilon_{n_j}\sigma_{j*}^2}{\sqrt{\nu}})^{-(\nu_{00}/2+1)}e^{-\delta_{00}/(2\sigma_{j*}^2)}\int_{\sigma_{j*}^2}^{(\sigma_{j*}^2+\frac{\epsilon_{n_j}\sigma_{j*}^2}{\sqrt{\nu}})}\frac{(\delta_{00}/2)^{\nu_{00}/2}}{\Gamma(\nu_{00}/2)}d\sigma_j^2$$

$$=(\sigma_{j*}^2+\frac{\epsilon_{n_j}\sigma_{j*}^2}{\sqrt{\nu}})^{-(\nu_{00}/2+1)}e^{-\delta_{00}/(2\sigma_{j*}^2)}\frac{(\delta_{00}/2)^{\nu_{00}/2}}{\Gamma(\nu_{00}/2)}\left[\left(\sigma_{j*}^2+\frac{\epsilon_{n_j}\sigma_{j*}^2}{\sqrt{\nu}}\right)-\sigma_{j*}^2\right]\asymp e^{\log\epsilon_{n_j}}$$

since $(\sigma_{j*}^2-\frac{\epsilon_{n_j}\sigma_{j*}^2}{\sqrt{\nu}})\vee\sigma_{j*}^2=\sigma_{j*}^2$.

E.4 Testing

Lemma E.4 (Testing). *For each $k=0,1$, let $q_{k*}\triangleq(g_k^*,\sigma_{k*}^2)$ and $\mathbf{E}_{q_{k*}}$ (resp. $\mathbf{E}_{q_k}$) denote the expectation taken with respect to the distribution $\mathcal{N}_{n_k}(g_k^*,\sigma_{k*}^2)$ (resp. $\mathcal{N}_{n_k}(g_k,\sigma_k^2)$). For each $k=0,1$, there exists a test ϕ_{n_k} such that for some $K,M>0$ and $\forall\ell\in\mathbb{N}$:*

$$\mathbf{E}_{q_{k*}}\phi_{n_k}\le\frac{e^{-(M^2K-1)n_k\epsilon_{n_k}^2}}{1-e^{-M^2Kn_k\epsilon_{n_k}^2}},\qquad \sup_{q_k=(g_k,\sigma_k^2)\in\mathscr{A}_{n,k,\ell}}\mathbf{E}_{q_k}(1-\phi_{n_k})\le e^{-M^2K\ell n_k\epsilon_{n_k}^2}\tag{E.29}$$

where $\mathscr{A}_{n,k,\ell}\triangleq\{g_k\in\mathscr{C}_{n,k};\|\Xi_k^{1/2}(g_k-g_k^*)\|_{n_k}>\ell M\epsilon_{n_k}\}\times\{\sigma_k^2\in\left[\frac{1}{2n_k},e^{n_k\epsilon_{n_k}^2}\right];|1-\sigma_k/\sigma_{k*}|>\ell\varepsilon_\sigma\}$ for some $\varepsilon_\sigma>1/(2n_k^2)$, $\Xi_0\triangleq\mathrm{diag}(\xi_1,\dots,\xi_{n_0})$, $\Xi_1\triangleq\mathrm{diag}(\xi_{n_0+1},\dots,\xi_{n_1})$ and $\{\xi_i\}_i$ are the latent variables in the mixture representation of the student-t distribution.

Proof. All along the proof k can take values in $\{0,1\}$. To construct the test ϕ_{n_k}, we first consider the set $\mathcal{Q}_j:=\{(g_k,\sigma_k^2)\in\mathscr{C}_{n,k}\times\left[\frac{1}{2n_k},e^{n_k\epsilon_{n_k}^2}\right];j\varepsilon<\|\Xi_k^{1/2}(g_k-g_k^*)\|_{n_k}<2j\varepsilon,j\varepsilon_\sigma<|1-\sigma_k/\sigma_{k*}|<2j\varepsilon_\sigma\}$ for a given $j\in\mathbb{N}$, $k\in\{0,1\}$, $\varepsilon>0$ and $\varepsilon_\sigma>0$. Consider a maximal set of points $q_{j,1},\dots,q_{j,N_j}\in\mathcal{Q}_j$, with $q_{j,l}=(g_k^l,\sigma_{k,l}^2)$ that satisfy $\|\Xi_k^{1/2}(g_k^l-g_k^{l'})\|_{n_k}\ge\varepsilon/3$ and $|\sigma_{k,l}^2-\sigma_{k,l'}^2|>\varepsilon_\sigma\sigma_{k*}^2/2$, for every $l,l'=1,\dots,N_j$. Then, consider N_j balls $B_{j,l}$ around each of these $q_{j,l}$ constructed as

$$B_{j,l}:=\left\{q\in\mathcal{Q}_j;\|\Xi_k^{1/2}(g_k^l-g_k)\|_{n_k}\le\varepsilon/3,\ |\sigma_{k,l}^2-\sigma_k^2|\le\varepsilon_\sigma\sigma_{k*}^2/2\right\}\tag{E.30}$$

and denote by η the radius of these balls. These balls cover $\mathcal{Q}_j$ as otherwise the set $q_{j,1},\dots,q_{j,N_j}$ would not be maximal. Moreover, $N_j=N(\eta,\mathcal{Q}_j,\|\Xi_k^{1/2}\cdot\|_{n_k}+|\cdot|)$. For every ball $B_{j,l}$ let us construct a test $\phi_{j,l}$ as: $\phi_{j,l} = \mathbb{1}_{\{dP_{q_{k,l}}/dP_* \geq 1\}}$, where $P_{q_{k,l}} = \mathcal{N}_{n_k}(g_k^l, \sigma_{k,l}^2)$, $q_{k,l} \in B_{j,l}$ and $P_{q_{k*}} = \mathcal{N}_{n_k}(g_k^*, \sigma_{k*}^2)$. Denote $q_{k*} \triangleq (g_k^*, \sigma_{k*}^2)$, and $\mathbf{E}_q$ is the expectation taken with respect to the distribution P_q. Since $q_{k,l} \in B_{j,l} \subset \mathcal{Q}_j$, this test satisfies:

$$\begin{aligned}
\mathbf{E}_{q_{k*}}\phi_{j,l} &= \int \mathbb{1}_{\{\sqrt{dP_{q_{k,l}}/dP_{k*}} \geq 1\}} dP_{q_{k*}} \leq \int \sqrt{dP_{q_{k,l}} dP_{q_{k*}}} \\
&\leq \left(\frac{\sigma_{k,l}}{\sigma_{k*}} + \frac{\sigma_{k*}}{\sigma_{k,l}}\right)^{-n_k/2} 2^{n_k/2} \exp\{-n_k \|\Xi_k^{1/2}(g_k^l - g_k^*)\|_{n_k}^2 / (4(\sigma_{k,l}^2 + \sigma_{k*}^2))\} \\
&< e^{-n_k j^2 \varepsilon^2/(8\sigma_{k*}^2)}, \qquad \text{(E.31)} \\
\mathbf{E}_{q_k}(1-\phi_{j,l}) &\leq \sqrt{\mathbf{E}_{q_{k,l}}(1-\phi_{j,l})}\sqrt{\mathbf{E}_{q_{k,l}}(dP_{q_k}/dP_{q_{k,l}})^2} \\
&< e^{-n_k j^2 \varepsilon^2/(16\sigma_{k*}^2)} e^{n_k j^2 \varepsilon^2/(18\sigma_{k*}^2)}, \qquad \forall q_k \in B_{j,l} \qquad \text{(E.32)}
\end{aligned}$$

where to get (E.31) we have used the inequality $|1-\sigma_k/\sigma_*| > j\varepsilon_\sigma$ which implies that $\left(\frac{\sigma_{k,l}}{\sigma_{0*}} + \frac{\sigma_{0*}}{\sigma_{k,l}}\right) > 1 + j\varepsilon_\sigma + (1-j\varepsilon_\sigma)^{-1} > 2$ and that $\sigma_{k,l}^2 + \sigma_{k*}^2 = \sigma_{k*}^2(\frac{\sigma_{k,l}^2}{\sigma_{k*}^2} + 1) < \sigma_{k*}^2(1 - j\varepsilon_\sigma + 1) \leq 2\sigma_{k*}^2$. To get (E.32) we have used the Cauchy-Schwartz inequality, the result in Lemma E.5 and the fact that for $q_k \in B_{j,l}$:

$$\mathbf{E}_{q_{k,l}}(1-\phi_{j,l}) = \int \mathbb{1}_{\{\sqrt{dP_{q_*}/dP_{q_{k,l}}} > 1\}} dP_{q_{k,l}} < \int \sqrt{dP_{q_{k,l}} dP_{q_*}} < e^{-n_k j^2 \varepsilon^2/(16\sigma_{k*}^2)}.$$

This implies that $\sup_{q_k \in B_{j,l}} \mathbf{E}_{q_k}(1-\phi_{j,l}) < e^{-n_k j \varepsilon^2/(144\sigma_{0*}^2)}$, and by choosing $\varepsilon = 12\sigma_{k*}\sqrt{\widetilde{M}_1}\epsilon_{n_k}$ it follows that

$$\mathbf{E}_{q_{k*}}\phi_{j,l} \leq e^{-\widetilde{M}_1 j n_k \epsilon_{n_k}^2} \qquad \text{and} \qquad \sup_{q_k \in B_{j,l}} \mathbf{E}_{q_k}(1-\phi_{j,l}) \leq e^{-\widetilde{M}_1 j n_k \epsilon_{n_k}^2}.$$

Let $\phi_{n_k} \triangleq \max_{j \in \mathbb{N}, l \in \{1,\ldots,N_j\}} \phi_{j,l}$. Then, if $N_j \lesssim \exp(n_k \epsilon_{n_k}^2)$, $\forall j \in \mathbb{N}$

$$\begin{aligned}
\mathbf{E}_{q_{k*}}\phi_{n_k} \leq \sum_{j=1}^{\infty}\sum_{l=1}^{N_j} \mathbf{E}_{q_*}\phi_{j,l} \leq \sum_{j=1}^{\infty}\sum_{l=1}^{N_j} e^{-\widetilde{M}_1 j n_k \epsilon_{n_k}^2} &\leq \sum_{j=1}^{\infty} N_j e^{-\widetilde{M}_1 j n_k \epsilon_{n_k}^2} \\
&\leq e^{n_k \epsilon_{n_k}^2}\left(\frac{1}{1 - e^{-\widetilde{M}_1 n_k \epsilon_{n_k}^2}} - 1\right) = \frac{e^{-(\widetilde{M}_1 - 1) n_k \epsilon_{n_k}^2}}{1 - e^{-\widetilde{M}_1 n_k \epsilon_{n_k}^2}}
\end{aligned}$$

and, since $\mathcal{Q}_j \subseteq \bigcup_{k=1,\dots,N_j} B_{j,k}$, we have for every $\ell \in \mathbb{N}$

$$\sup_{\{g_k \in \mathscr{C}_{n,k}; \|\Xi_k^{1/2}(g_k - g_k^*)\|_{n_k} > \ell\varepsilon\}} \sup_{\{\sigma_k^2 \in [1/(2n_k), e^{3n_k\epsilon_{n_k}^2}]; |1-\sigma_k/\sigma_*| > \ell\varepsilon_\sigma\}} \mathbf{E}_{q_k}(1-\phi_{n_k}) \leq \sup_{q_k \in \bigcup_{j>\ell} \mathcal{Q}_j} \mathbf{E}_{q_k}(1-\phi_{n_k})$$
$$\leq \sup_{j>\ell} \sup_{1\leq l \leq N_j} \sup_{q_k \in B_{j,l}} \mathbf{E}_{q_k}(1-\phi_{j,l}) \leq \sup_{j>\ell} \sup_{1\leq l\leq N_j} e^{-\widetilde{M}_1 j n_k \epsilon_{n_k}^2} = e^{-\widetilde{M}_1 \ell n_k \epsilon_{n_k}^2}.$$

Finally, remark that $\ell\varepsilon = \ell 12\sigma_{k*}\sqrt{\widetilde{M}_1}\epsilon_{n_k} =: \ell M \epsilon_{n_k}$ and we can write $\widetilde{M}_1 = M^2 K$ with $K = 1/(144\sigma_{k*}^2)$.

To conclude the proof we have to show that $N_j \lesssim \exp(n_k \epsilon_{n_k}^2)$. Since for every $j \in \mathbb{N}$, $\mathcal{Q}_j \subseteq \mathscr{C}_{n,k} \times \left[\frac{1}{2n_k}, e^{3n_k\epsilon_{n_k}^2}\right]$, then

$$N_j = N(\eta, \mathcal{Q}_j, \|\Xi_k^{1/2}\cdot\|_{n_k} + |\cdot|) \leq N(\eta, \mathscr{C}_{n,k} \times [(2n_k)^{-1}, e^{3n_k\epsilon_{n_k}^2}], \|\Xi_k^{1/2}\cdot\|_{n_k} + |\cdot|)$$
$$\leq N\left(\varepsilon/3, \mathscr{C}_{n,k}, \|\Xi_k^{1/2}\cdot\|_{n_k}\right) \times N\left(\varepsilon_\sigma \sigma_{k*}^2/2, \left\{\sigma_k^2 \in \mathbb{R}_+;\ \frac{1}{2n_k} \leq \sigma_k^2 < e^{3n_k\epsilon_{n_k}^2}\right\}, |\cdot|\right). \quad \text{(E.33)}$$

We start by considering the first factor. Remark that (by denoting $\boldsymbol{\alpha}_k \triangleq \boldsymbol{\alpha} I[k=0] + \boldsymbol{\beta} I[k=1]$) and by using the value of ε

$$\log N\left(\varepsilon/3, \mathscr{C}_{n,k}, \|\Xi_k^{1/2}\cdot\|_{n_k}\right) \leq \log N\Big(\varepsilon/3, \bigcup_{m_k=1}^{m_k^*} \{\boldsymbol{B}_k(z)'\boldsymbol{\alpha}_k; \boldsymbol{\alpha}_k \in \mathbb{R}^{m_k}, \|\boldsymbol{\alpha}_k\|_\infty \leq M_k\}, \|\Xi_k^{1/2}\cdot\|_{n_k}\Big)$$
$$\leq \log N\Big(4\underline{\sigma}_k \epsilon_{n_k}, \bigcup_{m_k=1}^{m_k^*} \{\boldsymbol{B}_k(z)'\boldsymbol{\alpha}_k;\ \boldsymbol{\alpha}_k \in \mathbb{R}^{m_k}, \|\boldsymbol{\alpha}_k\|_\infty \leq M_k\}, \|\Xi_k^{1/2}\cdot\|_{n_k}\Big).$$

By Van der Vaart [2000, Example 19.7], since $4\underline{\sigma}_k\epsilon_{n_k} > \underline{\sigma}_k\epsilon_{n_k}$, there exist a constant K_k depending on $\boldsymbol{\alpha}_k$ only such that (by denoting with $\|\cdot\|$ the Euclidean norm) the penultimate inequality below holds:

$$\log N\left(\varepsilon/3, \mathscr{C}_{n,k}, \|\Xi_k^{1/2}\cdot\|_{n_k}\right) \leq \log \sum_{m_k=1}^{m_k^*} N\Big(\underline{\sigma}_k\epsilon_{n_k}/\sqrt{m_k^*}, \{\boldsymbol{\alpha}_k \in \mathbb{R}^{m_k}, \|\boldsymbol{\alpha}_k\|_\infty \leq M_k\}, \|\cdot\|\Big)$$
$$\leq \log m_k^* N\Big(\underline{\sigma}_k\epsilon_{n_k}/\sqrt{m_k^*}, \{\boldsymbol{\alpha}_k \in \mathbb{R}^{m_k^*}, \|\boldsymbol{\alpha}_k\|_\infty \leq M_k\}, \|\cdot\|\Big)$$
$$\leq \log\left(m_k^* K_k \left(\frac{\sqrt{m_k^*}2M_k}{\underline{\sigma}_k\epsilon_{n_k}}\sqrt{m_k^*}\right)^{m_k^*}\right) \leq \log\left(m_k^* K_k \left(\frac{\sqrt{m_k^*}2M_k}{\underline{\sigma}_k\epsilon_{n_k}}\sqrt{m_k^*}\right)^{m_k^*}\right)$$

where we have used the fact that $\underline{\sigma}_k\epsilon_{n_k} \geq \underline{\sigma}_k\epsilon_{n_k}/\sqrt{m_k^*}$ for every $m_k^* \geq 1$ to get the first inequality and the fact that the diameter of the hypercube $[-M_k, M_k]^{m_k^*}$ is $\sqrt{m_k^*}2M_k$ to get the penultimate inequality. By replacing the optimal values $m_k^* \asymp \left(\frac{n_k}{\log n_k}\right)^{1/(2\delta+1)}$, $\epsilon_{n_k} \asymp \left(\frac{\log n_k}{n_k}\right)^{\delta/(2\delta+1)}$ and $M_k \asymp n_k \epsilon_{n_k}^2 = n_k^{1/(2\delta+1)} (\log n_k)^{2\delta/(2\delta+1)}$ we obtain

$$\sup_{\varepsilon > 3\underline{\sigma}_k \epsilon_{n_k}} \log N\left(\varepsilon/3, \mathscr{C}_{n,k}, \|\Xi_k^{1/2} \cdot \|_{n_k}\right) \lesssim m_k^* \log(n_k) \left(\frac{\delta+1}{2\delta+1} + 3\right) \lesssim n_k \epsilon_{n_k}^2$$

since $m_k^* \log(n) \asymp n\epsilon_{n_k}^2$. We now analyse the second factor in (E.33). Since $\varepsilon_\sigma \sigma_{k*}^2/2 > \underline{\sigma}_k^2/(2n_k^2)$ we obtain

$$\log N\left(\varepsilon_\sigma \sigma_{k*}^2/2, \left\{\sigma_k^2 \in \mathbb{R}_+;\ (2n_k)^{-1} \leq \sigma^2 < e^{3n_k \epsilon_{n_k}^2}\right\}, |\cdot|\right) \\ \leq \log N\left(\underline{\sigma}_k^2/(2n_k^2), \{\sigma_k^2 \in \mathbb{R}_+;\ (2n_k)^{-1} \leq \sigma^2 < e^{3n_k \epsilon_{n_k}^2}\}, |\cdot|\right) \leq 4n_k \epsilon_{n_k}^2$$

where the last inequality follows from the definition of ϵ_{n_k} and $\log(n_k)/n_k \leq \epsilon_{n_k}^2$. □

Lemma E.5. *For $k \in \{0,1\}$ and given $j, l \in \mathbb{N}$ let $P_{q_{k,l}} = \mathscr{N}_{n_k}(g_k^l, \sigma_{k,l}^2)$ and consider the balls $B_{j,l}$ defined in* (E.30)*. For every $q_k \in B_{j,l}$ it holds*

$$\sqrt{\mathbf{E}_{q_k^l}(dP_{q_k}/dP_{q_{k,l}})^2} \leq \exp\{j^2 \varepsilon^2 n_k/(18\sigma_{k*}^2)\}.$$

Proof. First, remark that simple algebra shows that

$$\mathbf{E}_{q_{k,l}}\left(\frac{dP_{q_k}}{dP_{q_{k,l}}}\right)^2 = \mathbf{E}_{q_k}\left(\frac{dP_{q_k}}{dP_{q_{k,l}}}\right) = \left(\frac{\sigma_{k,l}^2}{\sigma_k^2}\right)^{n_k/2} \left(2 - \frac{\sigma_k^2}{\sigma_{k,l}^2}\right)^{-n_k/2} \exp\left\{\sum_{i \in I_{00}} \frac{[\xi_i(g_k^l - g_k)]^2}{2\sigma_{k,l}^2 - \sigma_k^2}\right\}. \tag{E.34}$$

We start by bounding the first two factors:

$$\left(\frac{\sigma_{k,l}^2}{\sigma_k^2}\right)^{n_k/2} \left(2 - \frac{\sigma_k^2}{\sigma_{k,l}^2}\right)^{-n_k/2} = \left[\frac{\sigma_{k,l}^2}{\sigma_k^2} \frac{\sigma_{k,l}^2}{2\sigma_{k,l}^2 - \sigma_k^2}\right]^{n_k/2} = \left[\frac{\sigma_{k,l}^2}{\sigma_k^2} \frac{1}{2 - \sigma_k^2/\sigma_{k,l}^2}\right]^{n_k/2} \\ \leq \left[\frac{\sigma_{k,l}^2}{\sigma_k^2} \frac{1}{1 - \varepsilon_\sigma \sigma_{k*}^2/(2\sigma_{k,l}^2)}\right]^{n_k/2} \leq \left[\frac{(1-\varepsilon_\sigma)}{(1+\varepsilon_\sigma)} \frac{1}{1 - \varepsilon_\sigma/2(1+\varepsilon_\sigma)}\right]^{n_k/2} \\ = \left[\frac{2(1-\varepsilon_\sigma)}{2+\varepsilon_\sigma}\right]^{n_k/2} \leq \left[\frac{2(1-\varepsilon_\sigma)}{2}\right]^{n_k/2} \leq 1, \tag{E.35}$$

where we have used the facts that, since $q_k \in B_{j,l}$: $\sigma_k^2/\sigma_{k,l}^2 \leq 1 + \varepsilon_\sigma \sigma_{k*}^2/(2\sigma_{k,l}^2)$, and since both $q_k, q_{k,l} \in \mathcal{Q}_j$: $\sigma_{k,l}^2/\sigma_k^2 = (\sigma_{k,l}^2/\sigma_{k*}^2)/(\sigma_k^2/\sigma_{k*}^2) \leq (1-\varepsilon_\sigma)/(1+\varepsilon_\sigma)$. By using again the fact that

$q_k \in B_{j,l}$ and $q_{k,l} \in \mathscr{Q}_j$ the following inequalities hold: $\sigma_{k*}^2 < (1+\varepsilon/2)\sigma_{k*}^2 = (1+\varepsilon_\sigma - \varepsilon_\sigma/2)\sigma_{k*}^2 \leq (1-\varepsilon_\sigma\sigma_{k*}^2/(2\sigma_{k,l}^2))\sigma_{k,l}^2 \leq (2-\frac{\sigma_k^2}{\sigma_{k,l}^2})\sigma_{k,l}^2 = 2\sigma_{k,l}^2 - \sigma_k^2$ and then $(2\sigma_{k,l}^2 - \sigma_k^2)^{-1} < \sigma_{k*}^{-2}$, so that we get

$$\exp\left\{\sum_{i\in I_{00}} \xi_i (g_k^l - g_k)^2 (2\sigma_{k,l}^2 - \sigma_k^2)^{-1}\right\} \leq \exp\left\{\varepsilon^2 n_k/(9\sigma_{k*}^2)\right\} \leq \exp\left\{j^2\varepsilon^2 n_k/(9\sigma_{k*}^2)\right\}$$

for every $j \in \mathbb{N}$. This together with (E.35) allows to conclude that $\mathbf{E}_{q_{k,l}}\left(\frac{dP_{q_k}}{dP_{q_{k,l}}}\right)^2 \leq \exp\{j^2\varepsilon^2 n_k/(9\sigma_{k*}^2)\}$.
□

F Proofs for Section 4

We introduce new notation which we use in this section for convenience together with the already introduced notation. For every n, denote $X^{(n)} \triangleq (y^{(n)}, x^{(n)}, z^{(n)})$. Let $g_{0c} \triangleq g_0$, $g_{1c} \triangleq g_1$, $\sigma_{0c}^2 \triangleq \sigma_0^2$ and $\sigma_{1c}^2 \triangleq \sigma_1^2$. For the sample size of the cells I_{10} and I_{01} we use both the notations $n_{0n} \triangleq n_{10}$ and $n_{1a} \triangleq n_{01}$; moreover $I_{0n} \triangleq I_{10}$ and $I_{1a} \triangleq I_{01}$. We also use the notation $\boldsymbol{B}_0 \triangleq \boldsymbol{B}_{00}$, $\boldsymbol{B}_1 \triangleq \boldsymbol{B}_{11}$, $\boldsymbol{B}_{0n} \triangleq \boldsymbol{B}_{0,n}$, $\boldsymbol{B}_{1a} \triangleq \boldsymbol{B}_{1,a}$.

F.1 Posterior consistency conditional on the type

Theorem F.1. *Assume that for some $-\infty < a < b < \infty$ there exists a $\delta > 0$ and a constant $C_2 > 0$ such that: $g_0^* \in \mathscr{G}_0 \triangleq \mathscr{C}^\delta[a,\tau)$, $g_1^* \in \mathscr{G}_1 \triangleq \mathscr{C}^\delta[\tau,b]$ and $g_{0n}^*, g_{1a}^* \in \mathscr{C}^\delta[a,b]$, and $\|g_j^* - g_{m_k}^*\|_\infty \leq C_2 m_k^{-\delta}$ for $j = 0,1,0n,1a$ and $k = 0,1,n,a$. Moreover, $0 < \underline{\sigma}_j^2 < \sigma_{j*}^2 < \overline{\sigma}_j^2 < \infty$, for $j = 0,1,0n,1a$ and two constants $\underline{\sigma}_j^2, \overline{\sigma}_j^2$. Let π be the prior on $(\boldsymbol{g}, \boldsymbol{\sigma}^2, \boldsymbol{q}, \{\lambda_j, j = 0,1,n,a\})'$ specified in Section 4.7 with $m_j = m_j^* \asymp \left(\frac{n}{\log n}\right)^{1/(2\delta+1)}$, $\nu_{00} > 2$, $\delta_{00} > 2 + 6\delta/(2\delta+1)$, and $a_{j0} = 2\eta m_j^*$ for $j = 0,1,n,a$, and $\eta > \min\{3, \delta_{00}\}(2\delta+1)/4 + 3\delta/(2(2\delta+1))$. Moreover, suppose that Assumptions 1-3 and 5-9 hold and let $n_{lj} \asymp n$ for $l,j = 0,1$. Then, for all $M_n \to \infty$,*

$$P_*^{(n)}\pi\Big((f,\sigma_j^2) \in \mathscr{F} \times \mathbb{R}_+ \textit{ for } j = 0,1,0n,1a; |1-\sigma_j/\sigma_{j*}| \geq \widetilde{M}_n/\sqrt{n},$$
$$\textit{and } \forall \boldsymbol{s}, \|\boldsymbol{f_s} - \boldsymbol{f_s^*}\|_n \geq M_n\epsilon_n \Big| y^{(n)}, x^{(n)}, z^{(n)}\Big) \to 0$$

where $\epsilon_n \asymp \left(\frac{\log n}{n}\right)^{\delta/(2\delta+1)}$.

Proof. For every n, denote $X^{(n)} \triangleq (y^{(n)}, x^{(n)}, z^{(n)})$. We use the notation $\boldsymbol{g} \triangleq (g_0, g_1, g_{0n}, g_{1a})$,

$\boldsymbol{\theta} \triangleq (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\alpha}_n, \boldsymbol{\beta}_a)$, $\boldsymbol{\sigma}^2 \triangleq (\sigma_0^2, \sigma_1^2, \sigma_{0n}^2, \sigma_{1a}^2)$, $\boldsymbol{\lambda} \triangleq (\lambda_0, \lambda_1, \lambda_n, \lambda_a)$, $\boldsymbol{q} \triangleq (q_c, q_n, q_a)$, and

$$\begin{aligned}\prod_{i=1}^{n} p_{\boldsymbol{g},\boldsymbol{\sigma}^2}(y_i, x_i|z_i) \triangleq &\prod_{\{i\in I_{00}\}} \sum_{k\in\{c,n,a\}} t_\nu(y_i|z_i, g_{0s_i}, \sigma_{0s_i}^2, s_i = k) Pr(x_i = 0|z_i < \tau, \theta, s_i = k) q_k \\ &\times \prod_{\{i\in I_{10}\}} \sum_{k\in\{c,n,a\}} t_\nu(y_i|z_i, g_{0s_i}, \sigma_{0s_i}^2, s_i = k) Pr(x_i = 0|z_i < \tau, \theta, s_i = k) q_k \\ &\times \prod_{\{i\in I_{11}\}} \sum_{k\in\{c,n,a\}} t_\nu(y_i|z_i, g_{1s_i}, \sigma_{1s_i}^2, s_i = k) Pr(x_i = 0|z_i < \tau, \theta, s_i = k) q_k \\ &\times \prod_{\{i\in I_{01}\}} \sum_{k\in\{c,n,a\}} t_\nu(y_i|z_i, g_{1s_i}, \sigma_{1s_i}^2, s_i = k) Pr(x_i = 0|z_i < \tau, \theta, s_i = k) q_k\end{aligned}$$

to denote the likelihood. The marginal posterior for $(\boldsymbol{g}, \boldsymbol{\sigma}^2)$ writes

$$\begin{aligned}\pi(\boldsymbol{g}, \boldsymbol{\sigma}^2|D^{(n)}) &\propto \int \prod_{i=1}^{n} p_{\boldsymbol{g},\boldsymbol{\sigma}^2}(y_i, x_i|z_i)\pi(g_0)\pi(g_1)\pi(g_{0n})\pi(g_{1a})\pi(\sigma_0^2)\pi(\sigma_1^2)\pi(\sigma_{0n}^2)\pi(\sigma_{1a}^2)\pi(d\boldsymbol{q}) \\ &= \int \prod_{\{i\in I_{00}\}} \sum_{s_i\in\{c,n\}} t_\nu(y_i|z_i, g_{0s_i}, \sigma_{0s_i}^2) q_{s_i} \prod_{\{i\in I_{10}\}} t_\nu(y_i|z_i, g_{0n}, \sigma_n^2) q_n \prod_{\{i\in I_{11}\}} \sum_{s_i\in\{c,a\}} t_\nu(y_i|z_i, g_{1s_i}, \sigma_{1s_i}^2) q_{s_i} \\ &\times \prod_{\{i\in I_{01}\}} t_\nu(y_i|z_i, g_{1a}, \sigma_a^2) q_a \times \pi(g_0)\pi(g_1)\pi(g_{0n})\pi(g_{1a})\pi(\sigma_0^2)\pi(\sigma_1^2)\pi(\sigma_{0n}^2)\pi(\sigma_{1a}^2) d\pi(q_c, q_a, q_n) \\ &= \int \sum_{s_1\in\{c,n\}} \times \ldots \times \sum_{s_{n_{00}}\in\{c,n\}} \sum_{s_{n_{01}+1}\in\{c,n\}} \times \ldots \times \sum_{s_{n_{11}}\in\{c,n\}} \prod_{\{i\in I_{00}\}} t_\nu(y_i|z_i, g_{0s_i}, \sigma_{0s_i}^2) q_{s_i} \\ &\times \prod_{\{i\in I_{11}\}} t_\nu(y_i|z_i, g_{1s_i}, \sigma_{1s_i}^2) q_{s_i} \prod_{\{i\in I_{10}\}} t_\nu(y_i|z_i, g_{0n}, \sigma_n^2) q_n \prod_{\{i\in I_{01}\}} t_\nu(y_i|z_i, g_{1a}, \sigma_a^2) q_a \\ &\times \pi(g_0)\pi(g_1)\pi(g_{0n})\pi(g_{1a})\pi(\sigma_0^2)\pi(\sigma_1^2)\pi(\sigma_{0n}^2)\pi(\sigma_{1a}^2) d\pi(q_c, q_a, q_n). \quad \text{(F.1)}\end{aligned}$$

Remark that given our specification of the prior, we have $\pi(g_0) = \int \mathscr{N}_{m_0}(\boldsymbol{D}_\alpha^{-1}\boldsymbol{\alpha}_0, \lambda_0^{-1}\boldsymbol{D}_\alpha^{-1}\boldsymbol{T}_\alpha\boldsymbol{D}_\alpha^{-1'}) \otimes \delta_0(dg_0^\perp)d\pi(\lambda_0)$ where δ_0 denotes the Dirac mass at zero and $g_0^\perp \triangleq g_0^\perp(z) \triangleq g_0(z) - \boldsymbol{B}_{00}(z)'\boldsymbol{\alpha}$ (and similarly for $\pi(g_1)$, $\pi(g_{0n})$, $\pi(g_{1a})$ and $g_1^\perp$, $g_{0n}^\perp$ and $g_{1a}^\perp$).

Moreover, for $k \in \{0, 1\}$ and a given sequence $\{s_i\}_{i=1}^{n_{kk}}$:

$$\prod_{\{i\in I_{kk}\}} t_\nu(y_i|z_i, g_{ks_i}, \sigma_{ks_i}^2) = \int \underbrace{\prod_{\{i\in I_{kk}\}} \mathscr{N}(y_i|z_i, g_{ks_i}, \sigma_{ks_i}^2/\xi_i)}_{\triangleq Q^{(n_{kk})}_{\boldsymbol{g_{ks_k}}, \boldsymbol{\sigma^2_{ks_k}}}} \prod_{\{i\in I_{kk}\}} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i$$

and similarly for cells I_{10} and I_{01}: $\prod_{\{i\in I_{10}\}} t_\nu(y_i|z_i, g_{0n}, \sigma_n^2) = \int Q^{(n_{10})}_{g_{0n},\sigma_n^2} \prod_{\{i\in I_{10}\}} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i$ and $\prod_{\{i\in I_{01}\}} t_\nu(y_i|z_i, g_{1a}, \sigma_a^2) = \int Q^{(n_{01})}_{g_{1a},\sigma_a^2} \prod_{\{i\in I_{01}\}} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i$. The true Lebesgue density can be written in a similar way. For instance, the true Lebesgue density of $(y_i, x_i = 0)|\{z_i < \tau\}, \boldsymbol{g^*_{0s_0}}, \boldsymbol{\sigma^2_{0s_0*}}$ can be written:

$$
\begin{aligned}
p_*^{(n_{00})} &= \sum_{s_1\in\{c,n\}} q^*_{s_1} \times \ldots \times \sum_{s_{n_{00}}\in\{c,n\}} q^*_{s_{n_{00}}} \prod_{\{i\in I_{00}\}} t_\nu(y_i|z_i, g^*_{0s_i}, \sigma^2_{0s_i*}) \\
&= \int \sum_{s_1\in\{c,n\}} q^*_{s_1} \times \ldots \times \sum_{s_{n_{00}}\in\{c,n\}} q^*_{s_{n_{00}}} \underbrace{\prod_{\{i\in I_{00}\}} \mathscr{N}(y_i|z_i, g^*_{0s_i}, \sigma^2_{0s_i*}/\xi_i)}_{\triangleq Q^{(n_{00})}_{g^*_{0\boldsymbol{s_0}},\sigma^2_{0\boldsymbol{s_0}*}}} \prod_{\{i\in I_{00}\}} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i,
\end{aligned}
$$

where $\boldsymbol{s_0} \triangleq (s_1, \ldots, s_{n_{00}})$.

For some $C > 0$ and $\epsilon_n > 0$ denote by $\mathscr{A}$ the following event:

$$
\mathscr{A} \triangleq \left\{ \int \frac{\prod_{i=1}^n p_{\boldsymbol{g},\boldsymbol{\sigma}^2}(y_i, x_i|z_i)}{p_*^{(n)}} d\pi(\boldsymbol{\theta}, \boldsymbol{\sigma}^2, \boldsymbol{g}^\perp, \boldsymbol{\lambda}, \boldsymbol{q}) \geq e^{-(1+C)n\epsilon_n^2} \right\},
$$

where $\boldsymbol{g}^\perp \triangleq (g_0^\perp, g_1^\perp, g_{0n}^\perp, g_{1a}^\perp)$, $d\pi(\boldsymbol{\theta}_k, g_k^\perp, \lambda_k, \sigma_k^2)$ denotes the prior supported on $B_n^{KL}((\boldsymbol{g^*}, \boldsymbol{\sigma^2_*}, \boldsymbol{q^*}), \epsilon_n)$. By Ghosal and van der Vaart [2007, Lemma 10], $P_*^{(n)}[\mathscr{A}^c] \leq \frac{1}{C^2 n\epsilon_n^2}$ for every $C > 0$ where $\mathscr{A}^c$ denotes the complement of $\mathscr{A}$.

To prove the result of the theorem we use the test approach which relies on the uniformly exponential consistent test ϕ_n constructed in Lemma F.4. By this lemma, for every given sequence $\boldsymbol{s_k} = \{s_i\}_{i\in I_{kk}}$ with values in $\{c, n\}$ if $k = 0$ and in $\{c, a\}$ if $k = 1$, there exists a test ϕ_n that satisfies:

$$
\begin{aligned}
&Q^{(n)}_{\boldsymbol{f^*_s},\boldsymbol{\sigma^2_{s*}}} \phi_n \leq e^{n\epsilon_n^2}(1 - e^{-KM^2 n\epsilon_n^2})^{-1} e^{-KM^2 n\epsilon_n^2} \qquad \text{and} \\
&Q^{(n)}_{\boldsymbol{f_s},\boldsymbol{\sigma^2_s}}(1 - \phi_{n_k}) \leq e^{-KM^2 n\epsilon_n^2 j^2}, \qquad \forall(\boldsymbol{f_s}, \boldsymbol{\sigma^2_s});\ (f, \sigma_l^2) \in \mathscr{F}_n \times \left[\frac{1}{2n}, e^{6n\epsilon_n^2}\right] \text{ for } l = 0, 1, 1a, 0n; \\
&\qquad \|\Xi^{1/2}(\boldsymbol{f_s} - \boldsymbol{f^*_s})\|_n > \ell\varepsilon, |1 - \sigma_l/\sigma_{l*}| > \ell\varepsilon_\sigma \qquad \text{(F.2)}
\end{aligned}
$$

for $\varepsilon = 12\max(\sigma_{0*}, \sigma_{1*}, \sigma_{0n*}, \sigma_{1a*})M\sqrt{K}\epsilon_n$ for some constants $K, M > 0$, some $\varepsilon_\sigma > 1/(2n^2)$. Here, we have used the notation $\boldsymbol{\sigma^2_s} \triangleq (\boldsymbol{\sigma^2_{0s_0}}, \boldsymbol{\sigma^2_{1s_1}}, \sigma^2_{0n}, \sigma^2_{1a})$, $Q^{(n)}_{\boldsymbol{f^*_s},\boldsymbol{\sigma^2_{s*}}} \triangleq Q^{(n_{00})}_{\boldsymbol{g^*_{0s_0}},\boldsymbol{\sigma^2_{0s_0*}}} Q^{(n_{11})}_{\boldsymbol{g^*_{1s_1}},\boldsymbol{\sigma^2_{1s_1*}}} Q^{(n_{10})}_{g^*_{0n},\sigma^2_{0n*}} Q^{(n_{01})}_{g^*_{1a},\sigma^2_{1a*}}$ and $Q^{(n)}_{\boldsymbol{f_s},\boldsymbol{\sigma^2_s}} \triangleq Q^{(n_{00})}_{\boldsymbol{g_{0s_0}},\boldsymbol{\sigma^2_{0s_0}}} Q^{(n_{11})}_{\boldsymbol{g_{1s_1}},\boldsymbol{\sigma^2_{1s_1}}} Q^{(n_{10})}_{g_{0n},\sigma^2_{0n}} Q^{(n_{01})}_{g_{1a},\sigma^2_{1a}}$. Remark that, conditional on $\{\xi_i\}_{i=1}^n$ and on a sequence $\boldsymbol{s}$, $Q^{(n)}_{\boldsymbol{f^*_s},\boldsymbol{\sigma^2_{s*}}}$ is the true model. Moreover, we have used the notation $\Xi \triangleq \mathrm{diag}(\xi_i, i \in$

$I_{00}, \xi_i \in I_{10}, \xi_i \in I_{01}, \xi_i \in I_{11})$.

Recall the definition of $B_*(\epsilon_n, n^{-1/2})$. We now use the test ϕ_n to upper bound the posterior of $B_*(\epsilon_n, n^{-1/2})$. So,

$$\mathbf{E}_*[\pi(B_*(\epsilon_n, n^{-1/2})|X^{(n)})] = \mathbf{E}_*[\pi(B_*(\epsilon_n, n^{-1/2})|X^{(n)})\phi_n] \\ + \mathbf{E}_*[\overbrace{\pi(B_*(\epsilon_n, n^{-1/2})|X^{(n)})}^{\leq 1} \mathbb{1}_{\mathscr{A}^c}] + \mathbf{E}_*[\pi(B_*(\epsilon_n, n^{-1/2})|X^{(n)})(1-\phi_n)\mathbb{1}_{\mathscr{A}}] \quad \text{(F.3)}$$

where $\mathbf{E}_*[\pi(B_*(\epsilon_n, n^{-1/2})|X^{(n)})\mathbb{1}_{\mathscr{A}^c}] \leq \frac{1}{C^2 n\epsilon_n^2}$. We analyse the first and third terms separately. Let us start from the first term:

$$\mathbf{E}_*[\overbrace{\pi(B_*(\epsilon_n, n^{-1/2})|X^{(n)})}^{\leq 1}\phi_n] \leq \mathbf{E}_*[\phi_n] \leq \int \sum_{s_1 \in \{c,n\}} q^*_{s_1} \times \ldots \times \sum_{s_{n_{00}} \in \{c,n\}} q^*_{s_{n_{00}}} \\ \times \sum_{s_{n_{01}+1} \in \{c,n\}} q^*_{s_{n_{01}+1}} \times \ldots \times \sum_{s_{n_{11}} \in \{c,n\}} q^*_{s_{n_{11}}} Q^{(n)}_{\boldsymbol{f^*_s}, \boldsymbol{\sigma^2_{s*}}} \phi_n \prod_{\{i=1\}}^{n} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i \\ \leq \int \underbrace{\sum_{s_1 \in \{c,n\}} q^*_{s_1} \times \ldots \times \sum_{s_{n_{00}} \in \{c,n\}} q^*_{s_{n_{00}}} \sum_{s_{n_{01}+1} \in \{c,n\}} q^*_{s_{n_{01}+1}} \times \ldots \times \sum_{s_{n_{11}} \in \{c,n\}} q^*_{s_{n_{11}}}}_{\leq 1} \\ \times \frac{e^{-(KM^2-1)n_k\epsilon_n^2}}{(1-e^{-Kn_kM^2\epsilon_n^2})} \prod_{i=1}^{n} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i \\ \leq \int \frac{e^{-(KM^2-1)n_k\epsilon_n^2}}{(1-e^{-Kn_kM^2\epsilon_n^2})} \prod_{i=1}^{n} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i$$

by using the first line in (F.2) with $M = \xi_{\min}^{1/2} M_n/J$ for $J \in \mathbb{N}$ and $\xi_{\min} \triangleq \min_{1\leq i\leq n} \xi_i$. If M is sufficiently large to ensure that $KM^2 - 1 > KM^2/2$, $\frac{e^{-(KM^2-1)n\epsilon_n^2}}{(1-e^{-KnM^2\epsilon_n^2})} \leq e^{-KM^2 n\epsilon_n^2/2}$. Next, remark that the density function of the minimum $\xi_{\min}$ is

$$p(\xi_{\min}) = n\frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)}(\lambda_{\min})^{\nu/2-1}e^{-\xi_{\min}(\nu/2)}\left(\frac{\Gamma(\nu/2, \nu/2\xi_{\min})}{\Gamma(\nu/2)}\right)^{n-1},$$

where $\Gamma(\nu/2, \nu/2\xi_{\min})$ is the upper incomplete gamma function. Therefore, since $\frac{\Gamma(\nu/2,\nu/2\xi_{\min})}{\Gamma(\nu/2)} \leq$

1,

$$\int e^{-K\xi_{\min}M_n^2 n\epsilon_n^2/(2J^2)} \prod_{i=1}^{n} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i$$

$$\leq \int e^{-K\xi_{\min}M_n^2 n\epsilon_n^2/(2J^2)} n \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)} (\lambda_{\min})^{\nu/2-1} e^{-\xi_{\min}(\nu/2)} d\xi_{\min}$$

$$\leq 2e^{-Kc_q M_n^2 n\epsilon_n^2/(2J^2)} n \int_{c_q}^{+\infty} \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)} (\lambda_{\min})^{\nu/2-1} e^{-\xi_{\min}(\nu/2)} d\xi_{\min}$$

$$\leq 2e^{-n\epsilon_n^2(Kc_\tau M_n^2/(2J^2)-\log(n)/(n\epsilon_n^2))}, \quad \text{(F.4)}$$

where, for $\tau \in (0, 1/2)$, c_τ is the τ-quantile of a $\mathscr{G}a\left(\frac{\nu}{2}, \frac{\nu}{2} + \frac{KM_n^2 n_k \epsilon_n^2}{2J^2}\right)$. Therefore,

$$\mathbf{E}_*[\overbrace{\pi(B_*(\epsilon_n, n^{-1/2})|X^{(n)})}^{\leq 1} \phi_n] \leq 2e^{-n\epsilon_n^2(Kc_\tau M_n^2/(2J^2)-\log(n)/(n\epsilon_n^2))}. \quad \text{(F.5)}$$

Next, we analyse the third term in (F.3). By using the definition of $\mathscr{A}$ we obtain

$$\mathbf{E}_*[\pi(B_*(\epsilon_n, n^{-1/2})|X^{(n)})(1-\phi_n)\mathbb{1}_{\mathscr{A}}] \leq$$

$$\mathbf{E}_*\Big[\int\int_{B_*(\epsilon_n, n^{-1/2})} \prod_{\{i=1\}}^{n} \frac{p_{\boldsymbol{g},\boldsymbol{\sigma^2}}(y_i, x_i|z_i)}{p_*(y_i, x_i|z_i)} d\pi(\boldsymbol{\theta}, \boldsymbol{\sigma}^2, \boldsymbol{g}^\perp, \boldsymbol{\lambda}, \boldsymbol{q})(1-\phi_n)\mathbb{1}_{\mathscr{A}}\Big]$$

$$\times \left[\pi(B_n^{KL}((\boldsymbol{g^*}, \boldsymbol{\sigma^2_*}, \boldsymbol{q^*}), \epsilon_n))\right]^{-1} e^{(1+C)n\epsilon_n^2}$$

$$\leq e^{(1+C+4)n\epsilon_n^2}\mathbf{E}_*\Big[\int\int_{B_*(\epsilon_n, n^{-1/2})} \prod_{\{i=1\}}^{n} \frac{p_{\boldsymbol{g},\boldsymbol{\sigma^2}}(y_i, x_i|z_i)}{p_*(y_i, x_i|z_i)} d\pi(\boldsymbol{\theta}, \boldsymbol{\sigma}^2, \boldsymbol{g}^\perp, \boldsymbol{\lambda}, \boldsymbol{q})(1-\phi_n)\mathbb{1}_{\mathscr{A}}\Big],$$

where we have used the result of Lemma F.1 to get the second inequality: $\pi(B_n^{KL}((\boldsymbol{g^*}, \boldsymbol{\sigma^2_*}, \boldsymbol{q^*}), \epsilon_n)) \gtrsim$

$e^{-4n\epsilon_n^2}$. By Fubini's theorem and by using the mixture representation of the model we get:

$$\mathbf{E}_*[\pi(B_*(\epsilon_n, n^{-1/2})|X^{(n)})(1-\phi_n)\mathbb{1}_{\mathscr{A}}] \lesssim e^{(1+C+4)n\epsilon_n^2} \times$$
$$\int\int \underbrace{\sum_{s_1 \in \{c,n\}} q_{s_1}}_{\leq 1} \times \ldots \times \sum_{s_{n_{00}} \in \{c,n\}} q_{s_{n_{00}}} \sum_{s_{n_{01}+1} \in \{c,n\}} q_{s_{n_{01}+1}} \times \ldots \times \sum_{s_{n_{11}} \in \{c,n\}} q_{s_{n_{11}}}$$
$$\int_{B_*(\epsilon_n, n^{-1/2})} Q^{(n)}_{\boldsymbol{f_s}, \boldsymbol{\sigma_s^2}} (1-\phi_n) d\pi(\boldsymbol{\theta}, \boldsymbol{\sigma^2}, \boldsymbol{g}^{\perp}, \boldsymbol{\lambda}, \boldsymbol{q}) \prod_{i=1}^{n} \mathscr{G}a\left(\xi_i; \frac{\nu}{2}, \frac{\nu}{2}\right) d\xi_i. \quad \text{(F.6)}$$

For $k = 0, 1$ set

$$B_{\boldsymbol{s},j} \triangleq \Big\{(f, \sigma_l^2) \in \mathscr{F}_n \times \left[\frac{1}{2n}, e^{6n\epsilon_n^2}\right] \text{ for } l = 0, 1, 0n, 1a; j\varepsilon_\sigma < |1 - \sigma_l/\sigma_{l*}| < 2j\varepsilon_\sigma,$$
$$\text{and for given } \boldsymbol{s}, jM\epsilon_n < \|\Xi^{1/2}(\boldsymbol{f_s} - \boldsymbol{f_s}^*)\|_n < 2jM\epsilon_n\Big\}$$

and

$$\widetilde{\mathscr{G}_n} \triangleq \Big\{(f, \sigma_j^2) \in \mathscr{F}_n \times \left[\frac{1}{2n}, e^{6n\epsilon_n^2}\right] \text{ for } j = 0, 1, 0n, 1a; |1 - \sigma_j/\sigma_{j*}| \geq \widetilde{M}_n/\sqrt{n},$$
$$\text{and } \forall \boldsymbol{s}, \|\boldsymbol{f_s} - \boldsymbol{f_s^*}\|_n \geq M_n\epsilon_n\Big\}$$
$$\subseteq \Big\{(f, \sigma_j^2) \in \mathscr{F}_n \times \left[\frac{1}{2n}, e^{6n\epsilon_n^2}\right] \text{ for } j = 0, 1, 0n, 1a; |1 - \sigma_j/\sigma_{j*}| \geq \widetilde{M}_n/\sqrt{n},$$
$$\text{and for given } \boldsymbol{s}, \|\boldsymbol{f_s} - \boldsymbol{f_s^*}\|_n \geq M_n\epsilon_n\Big\}$$
$$\subseteq \Big\{(f, \sigma_j^2) \in \mathscr{F}_n \times \left[\frac{1}{2n}, e^{6n\epsilon_n^2}\right] \text{ for } j = 0, 1, 0n, 1a; |1 - \sigma_j/\sigma_{j*}| \geq \widetilde{M}_n/\sqrt{n},$$
$$\text{and for given } \boldsymbol{s}, \|\Xi^{1/2}(\boldsymbol{f_s} - \boldsymbol{f_s^*})\|_n \geq \xi_{\min}^{1/2} M_n\epsilon_n\Big\},$$

for $\xi_{\min} \triangleq \min_{1 \leq i \leq n} \xi_i$. Then $B_*(\epsilon_n, n^{-1/2}) \subseteq (\mathscr{F} \times \mathbb{R}_+^4) \backslash \left(\mathscr{F}_n \times \left[\frac{1}{2n}, e^{6n\epsilon_n^2}\right]^4\right) \cup \widetilde{\mathscr{G}_n}$ and $\widetilde{\mathscr{G}_n} \subseteq \bigcup_{j \geq J} B_{\boldsymbol{s},j}$ for $\xi_{\min}^{1/2} M_n = JM$ and $\widetilde{M}/\sqrt{n} = J\varepsilon_\sigma$. Therefore, by decomposing the integral over $B_*(\epsilon_n, n^{-1/2})$ in the sum of two integrals over the ranges $(\mathscr{F} \times \mathbb{R}_+^4) \backslash (\mathscr{F}_n \times \left[\frac{1}{2n}, e^{6n\epsilon_n^2}\right]^4)$ and $\widetilde{\mathscr{G}_n}$ and by upper bounding $(1-\phi_n)$ by 1 over $(\mathscr{F} \times \mathbb{R}_+^4) \backslash \left(\mathscr{F}_n \times \left[\frac{1}{2n}, e^{6n\epsilon_n^2}\right]^4\right)$, we can upper bound the last factor in (F.6) as:

$$
\begin{aligned}
\int\int_{B_*(\epsilon_n,n^{-1/2})} & Q^{(n)}_{\boldsymbol{f_s},\boldsymbol{\sigma_s^2}}(1-\phi_n)d\pi(\boldsymbol{\theta},\boldsymbol{\sigma^2},\boldsymbol{g}^\perp,\boldsymbol{\lambda})\prod_{i=1}^n \mathscr{G}a\left(\xi_i;\frac{\nu}{2},\frac{\nu}{2}\right)d\xi_i \\
&\leq \int_{\mathscr{F}\setminus\mathscr{F}_n}\prod_{k\in\{0,1,0n,1a\}}\underbrace{\int_{(2n)^{-1}}^{e^{6n\epsilon_n^2}} d\pi(\sigma_k^2)}_{\leq 1}\,d\pi(\boldsymbol{\theta},\boldsymbol{g}^\perp,\boldsymbol{\lambda}) \\
&\quad + \frac{(3^4-1)}{2}\int_{\mathscr{F}}\left(\int_0^{(2n)^{-1}}\pi(\sigma_0^2)d\sigma_0^2+\int_{e^{6n\epsilon_n^2}}^{+\infty}\pi(\sigma_0^2)d\sigma_0^2\right)d\pi(\boldsymbol{\theta},\boldsymbol{g}^\perp) \\
&\quad + \int\int_{\widetilde{\mathscr{G}_n}} Q^{(n)}_{\boldsymbol{f_s},\boldsymbol{\sigma_s^2}}(1-\phi_n)\,d\pi(\boldsymbol{\theta},\boldsymbol{\sigma^2},\boldsymbol{g}^\perp,\boldsymbol{\lambda})\prod_{i=1}^n \mathscr{G}a\left(\xi_i;\frac{\nu}{2},\frac{\nu}{2}\right)d\xi_i \\
&\leq \prod_{k\in\{0,1,0n,1a\}}\pi(\mathscr{G}_k\setminus\mathscr{C}_{n,k})+\frac{(3^4-1)}{2}\underbrace{\pi(\mathscr{F})}_{\leq 1}\left(\int_0^{(2n)^{-1}}\pi(\sigma_0^2)d\sigma_0^2+\int_{e^{6n\epsilon_n^2}}^{+\infty}\pi(\sigma_0^2)d\sigma_0^2\right) \\
&\quad + \int\int_{\widetilde{\mathscr{G}_n}} Q^{(n)}_{\boldsymbol{f_s},\boldsymbol{\sigma_s^2}}(1-\phi_n)\,d\pi(\boldsymbol{\theta},\boldsymbol{\sigma^2},\boldsymbol{g}^\perp,\boldsymbol{\lambda},\boldsymbol{q})\prod_{i=1}^n \mathscr{G}a\left(\xi_i;\frac{\nu}{2},\frac{\nu}{2}\right)d\xi_i \\
&\leq e^{-n\epsilon_n^2 4\eta/(2\delta+1)}+\frac{3^4}{2}\left(e^{-n\epsilon_n^2\delta_{00}}+\frac{e^{-6n\epsilon_n^2}\delta_{00}}{\nu_{00}-2}\right)+\int\sum_{j\geq J}e^{-Kj^2M^2n\epsilon_n^2}\prod_{i=1}^n \mathscr{G}a\left(\xi_i;\frac{\nu}{2},\frac{\nu}{2}\right)d\xi_i
\end{aligned}
\tag{F.7}
$$

where we have used Lemma F.2 to upper bound $\pi(\mathscr{G}_k\setminus\mathscr{C}_{n,k})$, the concentration inequality for sub-Gamma random variables to upper bound the integral over $\left[0,\frac{1}{2n_k}\right]$ (since $(\sigma_k^{-2}-\mathbf{E}[\sigma_{\mathbf{k}}^{-\mathbf{2}}])$ is sub-Gamma $\left(2\frac{\nu_{00}}{\delta_{00}^2},\frac{2}{\delta_{00}}\right)$), and the second inequality in (F.2) with $JM=\xi_{\min}^{1/2}M_n$ to control the term $Q^{(n)}_{\boldsymbol{f_s},\boldsymbol{\sigma_s^2}}(1-\phi_n)$. By using the same argument to get (F.4) we obtain $\mathbf{E}[e^{-KJ^2M^2n\epsilon_n^2}]\leq 2\exp\{-n\epsilon_n^2(Kc_qM_n^2-\epsilon_n)\}$. Hence,

$$
\begin{aligned}
\mathbf{E}_*[\pi(B_*(\epsilon_n,n^{-1/2})|X^{(n)})(1-\phi_n)\mathbb{1}_{\mathscr{A}}] \lesssim e^{(1+C+4)n\epsilon_n^2}\Big(& e^{-n\epsilon_n^2 4\eta/(2\delta+1)}+e^{-n\epsilon_n^2\delta_{00}} \\
&+e^{-6n_k\epsilon_{n_k}^2}+2e^{-n\epsilon_n^2(Kc_qM_n^2-\epsilon_n)}\Big).
\end{aligned}
$$

By putting (F.3) and (F.5) together we get :

$$\mathbf{E}_*[\pi(B_*(\epsilon_n, n^{-1/2})|X^{(n)})] \lesssim e^{-n\epsilon_n^2(Kc_\tau M_n^2/(2J^2)-\epsilon_n)} + \frac{1}{C^2 n\epsilon_n^2} + e^{(5+C)n\epsilon_n^2}\left(e^{-n\epsilon_n^2 4\eta/(2\delta+1)} + e^{-n\epsilon_n^2\delta_{00}} + e^{-6n\epsilon_n^2} + e^{-n\epsilon_n^2(Kc_q M_n^2-\epsilon_n)}\right)$$

which converges to zero for every M_n, K sufficiently large and every $0 < C < \min\{2, \delta_{00} - 5\}$ such that $5+C < \frac{4\eta}{(2\delta+1)}$ which implies $\eta > \min\{7, \delta_{00}\}(2\delta+1)/4$. This establishes the statement of the theorem.

□

F.2 Proof of Theorem 4.2

For given $m_0, m_n \in \mathbb{N}$ and for every $s \in \{c, n\}$, let $\boldsymbol{\alpha}_s \triangleq \boldsymbol{\alpha} I[s = c] + \boldsymbol{\alpha}_n I[s = n]$, $\boldsymbol{\alpha} \in \mathbb{R}^{m_0}$, $\boldsymbol{\alpha}_n \in \mathbb{R}^{m_n}$, and $\forall z$: $\boldsymbol{B}_{0s}(z) \triangleq \boldsymbol{B}_{00}(z)I[s = c] + \boldsymbol{B}_{00,n}(z)I[s = n]$, with $\boldsymbol{B}_{00}(z) \in \mathbb{R}^{m_0}$ and $\boldsymbol{B}_{00,n}(z) \in \mathbb{R}^{m_n}$. For given $m_0, m_n, m_a, m_1 \in \mathbb{N}$, define the following functional spaces

$$\mathscr{G}_{m_0,m_n} \triangleq \Big\{g_{m_0,m_n}(s, z; \boldsymbol{\alpha}, \boldsymbol{\alpha}_n) = \boldsymbol{B}_{0s}(z)'\boldsymbol{\alpha}_s : \{c, n\} \times \boldsymbol{z_{00}} \to \mathbb{R},\ \boldsymbol{\alpha} \in \mathbb{R}^{m_0}, \boldsymbol{\alpha}_n \in \mathbb{R}^{m_n}\Big\},$$

$$\mathscr{G}_{m_n} \triangleq \Big\{g_{m_n}(z; \boldsymbol{\alpha}_n) = \boldsymbol{B}_{10,n}(z)'\boldsymbol{\alpha}_n : \boldsymbol{z_{10}} \to \mathbb{R},\ \boldsymbol{\alpha}_n \in \mathbb{R}^{m_n}\Big\},$$

$$\mathscr{G}_{m_a} \triangleq \Big\{g_{m_a}(z; \boldsymbol{\beta}_a) = \boldsymbol{B}_{01,a}(z)'\boldsymbol{\beta}_a : \boldsymbol{z_{01}} \to \mathbb{R},\ \boldsymbol{\beta}_n \in \mathbb{R}^{m_a}\Big\},$$

$$\mathscr{G}_{m_1,m_a} \triangleq \Big\{g_{m_1,m_a}(s, z; \boldsymbol{\beta}, \boldsymbol{\beta}_a) = \boldsymbol{B}_{1s}(z)'\boldsymbol{\beta}_s : \{c, a\} \times \boldsymbol{z_{11}} \to \mathbb{R},\ \boldsymbol{\beta} \in \mathbb{R}^{m_1}, \boldsymbol{\beta}_a \in \mathbb{R}^{m_a}\Big\},$$

where for every $s \in \{c, a\}$, $\boldsymbol{\beta}_s \triangleq \boldsymbol{\beta} I[s = c] + \boldsymbol{\beta}_a I[s = a]$, $\boldsymbol{\beta} \in \mathbb{R}^{m_1}$, $\boldsymbol{\beta}_a \in \mathbb{R}^{m_a}$, and $\boldsymbol{B}_{1s}(z) \triangleq \boldsymbol{B}_{11}(z)I[s = c] + \boldsymbol{B}_{11,a}(z)I[s = a]$, with $\boldsymbol{B}_{11}(z) \in \mathbb{R}^{m_1}$, $\forall z$, and $\boldsymbol{B}_{11,a}(z) \in \mathbb{R}^{m_a}$, $\forall z$. Moreover, for $\boldsymbol{m} \triangleq (m_0, m_n, m_a, m_1)$ define

$$\mathscr{F}_{\boldsymbol{m}} \triangleq \Big\{f(s, z) = \Big(g_{m_0,m_n}(s, z)I[s \in \{c, n\}] + g_{m_a}(z)I[s = a]\Big)I[z < \tau] + \Big(g_{m_1,m_a}(s, z)I[s \in \{c, a\}] + g_{m_n}(z)I[s = n]\Big)I[z \geq \tau],\ g_{m_0,m_n}(s, z) \in \mathscr{G}_{m_0,m_n}, g_{m_a}(z) \in \mathscr{G}_{m_a}, g_{m_1,m_a}(s, z) \in \mathscr{G}_{m_1,m_a}, g_{m_n} \in \mathscr{G}_{m_n}\Big\}$$

and for every sequence $\boldsymbol{s}$, $\mathscr{F}_{\boldsymbol{m}} \subset L_2(P_{n,\boldsymbol{s}})$ where $P_{n,\boldsymbol{s}}$ has support $(s_1, z_1), \ldots, (s_{n_{00}}, z_{n_{00}})$, $(n, z_{n_{00}+1}), \ldots, (n, z_{\bar{n}_{10}})$, $(a, z_{\bar{n}_{10}+1}), \ldots, (a, z_{\bar{n}_{01}})$, $(s_{\bar{n}_{01}+1}, z_{\bar{n}_{01}+1}), \ldots, (s_n, z_n)$. Associated to every sequence $\boldsymbol{s}$ we also define two linear functionals:

$$
\begin{aligned}
& L_{0\boldsymbol{s}} : \mathscr{F}_{\boldsymbol{m}} \to \mathbb{R} \\
& \quad f(s,z) \mapsto \boldsymbol{e}'_{m_0,m_0}\mathbb{B}_{00}^{-1}\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{B}_{00}(z_i)f(s_i,z_i)I[s_i=c]I[z_i<\tau] \\
& \qquad\qquad = \boldsymbol{e}'_{m_0,m_0}\frac{\sharp\{i; s_i=c, z_i<\tau\}}{n}\mathbb{B}_{00}^{-1}\frac{\sum_{i=1}^{n}\boldsymbol{B}_{00}(z_i)g_{m_0,m_n}(s_i,z_i)I[s_i=c]I[z_i<\tau]}{\sharp\{i; s_i=c, z_i<\tau\}} \\
& L_{1\boldsymbol{s}} : \mathscr{F}_{\boldsymbol{m}} \to \mathbb{R} \\
& \quad f(s,z) \mapsto \boldsymbol{e}'_{m_1,1}\mathbb{B}_{11}^{-1}\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{B}_{11}(z_i)f(s_i,z_i)I[s_i=c]I[z_i\geq\tau] \\
& \qquad\qquad = \boldsymbol{e}'_{m_1,1}\frac{\sharp\{i; s_i=c, z_i\geq\tau\}}{n}\mathbb{B}_{11}^{-1}\frac{\sum_{i=1}^{n}\boldsymbol{B}_{11}(z_i)g_{m_1,m_a}(s_i,z_i)I[s_i=c]I[z_i\geq\tau]}{\sharp\{i; s_i=c, z_i\geq\tau\}}
\end{aligned}
$$

where $\boldsymbol{e}_{i,j}$ denotes the $(i \times 1)$ canonical vector with all components equal to zero but the j-th one, $\mathbb{B}_{00} \triangleq \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{B}_{00}(z_i)\boldsymbol{B}_{00}(z_i)'I[s_i=c]I[z_i<\tau] = P_{n,\boldsymbol{s}}\boldsymbol{B}_{00}(z_i)\boldsymbol{B}_{00}(z_i)'I[s_i=c]I[z_i<\tau]$ and $\mathbb{B}_{11} \triangleq \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{B}_{11}(z_i)\boldsymbol{B}_{11}(z_i)'I[z_i=c]I[z_i\geq\tau]$. Therefore, for every $f \in \mathscr{F}_{\boldsymbol{m}}$ and every sequence $\boldsymbol{s}$, $\boldsymbol{\alpha}_{[m_0]} = L_{0\boldsymbol{s}}f$ and $\boldsymbol{\beta}_{[1]} = L_{1\boldsymbol{s}}f$. The linear functionals $L_{0\boldsymbol{s}}$ and $L_{1\boldsymbol{s}}$ can be written: $\forall\varphi \in L_2(P_{n,\boldsymbol{s}})$,

$$
L_{0\boldsymbol{s}}\varphi = \langle\varphi, \ell_{0\boldsymbol{s}}\rangle_{P_{n,\boldsymbol{s}}}, \qquad L_{1\boldsymbol{s}}\varphi = \langle\varphi, \ell_{1\boldsymbol{s}}\rangle_{P_{n,\boldsymbol{s}}}
$$

where for every sequence $\boldsymbol{s}$, $\langle\cdot,\cdot\rangle_{P_{n,\boldsymbol{s}}}$ (resp. $\|\cdot\|_{n,\boldsymbol{s}}$) denotes the scalar product (resp. the induced norm) in $L_2(P_{n,\boldsymbol{s}})$ and

$$
\begin{aligned}
\ell_{0\boldsymbol{s}}(s,z) & \triangleq \boldsymbol{e}'_{m_0,m_0}\mathbb{B}_{00}^{-1}\boldsymbol{B}_{00}(z)I[s=c]I[z<\tau], \\
\ell_{1\boldsymbol{s}}(s,z) & \triangleq \boldsymbol{e}'_{m_1,1}\mathbb{B}_{11}^{-1}\boldsymbol{B}_{11}(z)f(s,z)I[s=c]I[z\geq\tau].
\end{aligned}
$$

Therefore, by the Riesz theorem, for $j=0,1$, $\|L_{j\boldsymbol{s}}\|_{n,\boldsymbol{s}}^2 = \|\ell_{j\boldsymbol{s}}\|_{n,\boldsymbol{s}}^2$ and $\|\ell_{0\boldsymbol{s}}\|_{n,\boldsymbol{s}}^2 = e'_{m_0,m_0}\mathbb{B}_{00}^{-1}e_{m_0,m_0}I[s=c]I[z<\tau]$, $\|\ell_{1\boldsymbol{s}}\|_{n,\boldsymbol{s}}^2 = e'_{m_1,1}\mathbb{B}_{11}^{-1}e_{m_1,1}I[s=c]I[z\geq\tau]$. By the definition of $\boldsymbol{B}_{jj}(z)$, $j=0,1$, there exists a constant $c_j > 0$ such that $\|\ell_{j\boldsymbol{s}}\|_{n,\boldsymbol{s}} \leq c_j < \infty$ for every n and for every $\boldsymbol{s}$.

Since $|CATE - CATE^*| \triangleq \left|(\boldsymbol{\beta}_{[1]} - \boldsymbol{\alpha}_{[m_0]}) - (\boldsymbol{\beta}^*_{[1]} - \boldsymbol{\alpha}^*_{[m_0]})\right| \leq |\boldsymbol{\beta}_{[1]} - \boldsymbol{\beta}^*_{[1]}| + |\boldsymbol{\alpha}_{[m_0]} - \boldsymbol{\alpha}^*_{[m_0]}|$

it holds that (by denoting with Σ the event $\{|1-\sigma_j/\sigma_{j*}| \geq \widetilde{M}_n\epsilon_n \text{ for } j=0,1,0n,1a\}$)

$$\pi\left(|CATE-CATE^*| \geq M_{1,n}\epsilon_n,\ \underbrace{|1-\sigma_j^2/\sigma_{j*}^2| \geq \widetilde{M}_n\epsilon_n \text{ for } j=0,1,0n,1a}_{\triangleq\Sigma}\middle| X^{(n)}\right)$$
$$\leq \pi\left(|\boldsymbol{\beta}_{[1]}-\boldsymbol{\beta}^*_{[1]}| + |\boldsymbol{\alpha}_{[m0]}-\boldsymbol{\alpha}^*_{[m0]}| \geq M_{1,n}\epsilon_n,\ \Sigma\middle| X^{(n)}\right)$$
$$\leq \pi\left(\forall \boldsymbol{s}, \forall f \in \mathscr{F}_{\boldsymbol{m}};\ |L_{1\boldsymbol{s}}(f-f^*)| + |L_{0\boldsymbol{s}}(f-f^*)| \geq M_{1,n}\epsilon_n,\ \Sigma| X^{(n)}\right)$$
$$\leq \pi\left(\forall \boldsymbol{s}, \forall f \in \mathscr{F}_{\boldsymbol{m}};\ (\|L_{1\boldsymbol{s}}\|_{n,\boldsymbol{s}} + \|L_{0\boldsymbol{s}}\|_{n,\boldsymbol{s}})\ \|f-f^*\|_{n,\boldsymbol{s}} \geq M_{1,n}\epsilon_n,\ \Sigma| X^{(n)}\right). \quad \text{(F.8)}$$

Remark that $\forall f \in \mathscr{F}_{\boldsymbol{m}}$, for $f^* \in \mathscr{F}_{\boldsymbol{m}}$ and for every given $\boldsymbol{s} = (\boldsymbol{s_0}', \boldsymbol{s_1})'$ with $\boldsymbol{s_0} \triangleq (s_1, \ldots, s_{n_{00}})$, $\boldsymbol{s_1} \triangleq (s_{\bar{n}_{01}+1}, \ldots, s_n)$:

$$\|f-f^*\|^2_{n,\boldsymbol{s}} \leq$$
$$\frac{1}{n}\sum_{i\in I_{00}}\left[2(g_{0s_i}(z_i)-g_{0s_i}(z_i)^*)^2 + 4(g_{m_0,m_n}(s_i,z_i)-g_{0s_i}(z_i))^2 + 4(g_{0s_i}(z_i)^* - g_{m_0,m_n}(s_i,z_i)^*)^2\right]$$
$$+\sum_{j\in\{0n,1a\}}\frac{1}{n}\sum_{i\in I_j}\left[2(g_j(z_i)-g_j(z_i)^*)^2 + 4(g_{m_j}(z_i)-g_j(z_i))^2 + 4(g_j(z_i)^* - g_{m_j}(z_i)^*)^2\right]$$
$$+\frac{1}{n}\sum_{i\in I_{11}}\left[2(g_{1s_i}(z_i)-g_{1s_i}(z_i)^*)^2 + 4(g_{m_1,m_a}(s_i,z_i)-g_{1s_i}(z_i))^2 + 4(g_{1s_i}(z_i)^* - g_{m_1,m_a}(s_i,z_i)^*)^2\right]$$
$$= 2\|\boldsymbol{f_s}-\boldsymbol{f^*_s}\|^2_n + 4\|f-\boldsymbol{f_s}\|^2_{n,\boldsymbol{s}} + 4\|\boldsymbol{f^*_s}-f^*\|^2_{n,\boldsymbol{s}}, \quad \text{(F.9)}$$

where $\boldsymbol{f_s}$ (resp. $\boldsymbol{f^*_s}$) is defined in (D.3) (resp. for the true value of the parameters) and remark that $f^* \in \mathscr{F}_{\boldsymbol{m}}$. Remark that the second term in the right hand side of (F.9) has zero prior mass and, under the assumption of the theorem $\|\boldsymbol{f^*_s}-f^*\|_{n,\boldsymbol{s}} \leq C_2(\min\{m_0,m_1,m_{0n},m_{1a}\})^{-\delta}$. By plugging this in (F.8) we get:

$$\pi\left(|CATE-CATE^*| \geq M_{1,n}\epsilon_n,\ \Sigma| X^{(n)}\right)$$
$$\leq \pi\left(\forall \boldsymbol{s}, \forall f \in \mathscr{F};\ \|\boldsymbol{f_s}-\boldsymbol{f^*_s}\|^2_n \geq \frac{1}{2}\left(\frac{M^2_{1,n}\epsilon^2_n}{(c_0+c_1)^2} - 4C_2^2(\min\{m_0,m_1,m_{0n},m_{1a}\})^{-2\delta}\right),\ \Sigma\middle| X^{(n)}\right)$$
$$= \pi\left(\forall \boldsymbol{s}, \forall f \in \mathscr{F};\ \|\boldsymbol{f_s}-\boldsymbol{f^*_s}\|^2_n \geq \frac{\epsilon^2_n}{2}\left(\frac{M^2_{1,n}}{(c_0+c_1)^2} - 4C_2^2\right),\ \Sigma\middle| X^{(n)}\right) \quad \text{(F.10)}$$

where we have set $m_0 \asymp m_1 \asymp m_{0n} \asymp m_{1a} \asymp (n/\log(n))^{1/(2\delta+1)}$ to get the last line. If $M_{1,n}$ is large enough such that $\frac{M^2_{1,n}}{2(c_0+c_1)^2} - 2C_2^2 > 0$, then (F.10) converges to zero by Theorem F.1 with $M_n \leq \frac{M^2_{1,n}}{2(c_0+c_1)^2} - 2C_2^2$.

F.3 Technical Lemmas

Lemma F.1. *Assume the conditions of Theorem F.1 hold. Then, there exists an* $N > 0$ *such that* $\forall n \geq N$ *and* $\forall \epsilon_n > 0$

$$\pi(B_n^{KL}((\boldsymbol{g}^*, \boldsymbol{\sigma}^{\boldsymbol{2}}_{\boldsymbol{*}}, \boldsymbol{q}^*), \epsilon_n)) \quad \gtrsim \quad \exp\left\{-4n\epsilon_n^2\right\}. \tag{F.11}$$

Proof. By Lemma F.3 there exists an $N > 0$ such that $\forall n \geq N$ and $\forall \epsilon_n > 0$:

$$\pi(B_n^{KL}((\boldsymbol{g}^*, \boldsymbol{\sigma}^{\boldsymbol{2}}_{\boldsymbol{*}}, \boldsymbol{q}^*), \epsilon_n)) \gtrsim e^{6\log\epsilon_n} \prod_{j\in\{0,1,0n,1a\}} \pi\left(g_j \in \mathscr{G}_j;\ \|g_j - g_j^*\|_\infty^2 \leq \frac{\epsilon_n^2 \underline{\sigma}_j^2}{6}\right). \tag{F.12}$$

So, we have to analyze the four factors in the product. We start by considering $\pi\left\{g_1 \in \mathscr{G}_1; \|g_1 - g_1^*\|_\infty \leq \frac{\underline{\sigma}_j\epsilon_n}{\sqrt{6}}\right\}$ and we lower bound it. Because $\|g_1 - g_1^*\|_\infty \leq \|g_1 - g_{m_1}\|_\infty + \|g_{m_1} - g^*_{m_1}\|_\infty + \|g^*_{m_1} - g_1^*\|_\infty$, $\|g^*_{m_1} - g_1^*\|_\infty \leq C_2 m_1^{-\delta}$, and $\|g_{m_1} - g^*_{m_1}\|_\infty \leq C_3 m_1^{1/2}\|\boldsymbol{\beta} - \boldsymbol{\beta}^*\|$ (by boundedness of natural cubic splines), we have,

$$\pi\left\{g_1 \in \mathscr{G}_1;\ \|g_1 - g_1^*\|_\infty \leq \frac{\epsilon_n \underline{\sigma}_j}{\sqrt{6}}\right\} \ \geq \ \pi\left\{\boldsymbol{\beta} \in \mathbb{R}^{m_1}; C_2 m_1^{-\delta} + C_3 m_1^{1/2}\|\boldsymbol{\beta} - \boldsymbol{\beta}^*\| \leq \frac{\epsilon_n \underline{\sigma}_j}{\sqrt{6}}\right\}, \tag{F.13}$$

where we have used the fact that the prior on $(g_1 - g_{m_1})$ is degenerate on zero. Therefore, by setting $\epsilon_n = \frac{2\sqrt{6}C_2}{\underline{\sigma}_j}\left(\frac{\log n}{n}\right)^{\delta/(2\delta+1)}$, by replacing it and $m_1 = m_1^* \asymp \left(\frac{n}{\log n}\right)^{1/(2\delta+1)}$ and since $\epsilon_n \asymp (m_1^*)^{-\delta}$ we obtain

$$\pi\left\{g_1 \in \mathscr{G}_1;\ \|g_1 - g_1^*\|_\infty \leq \frac{\underline{\sigma}_j \epsilon_n}{\sqrt{6}}\right\} \geq \pi\left(\boldsymbol{\beta} \in \mathbb{R}^{m_1^*}; \|\boldsymbol{\beta} - \boldsymbol{\beta}^*\| \leq (m_1^*)^{-\delta-1/2}\frac{C_2}{C_3}\right) \tag{F.14}$$

and denote $\widetilde{C} \triangleq C_2/C_3$. From the proof of Lemma E.1 it follows that

$$\pi\left(\boldsymbol{\beta} \in \mathbb{R}^{m_1};\ \|\boldsymbol{\beta} - \boldsymbol{\beta}^*\| \leq (m_1^*)^{-\delta-1/2}\widetilde{C}\right) \gtrsim \exp\{-m_1^*\log(n)(\delta+1)/(2\delta+1)\}.$$

In a similar way we can lower bound $\pi\left\{g_j \in \mathscr{G}_j;\ \|g_j - g_j^*\|_\infty \leq \frac{\epsilon_n \underline{\sigma}_j}{\sqrt{6}}\right\}$ for $j = 0, 0n, 1a$. Therefore, since $\log(\epsilon_n)/(n\epsilon_n^2) \geq -\epsilon_n^{1/\delta}\delta/(2\delta+1)$

$$\pi(B_n^{KL}((\boldsymbol{g}^*, \boldsymbol{\sigma_*^2}, \boldsymbol{q}^*), \epsilon_n)) \gtrsim e^{6\log\epsilon_n} \exp\left\{-\sum_{j\in\{0,1,n,a\}} m_j^* \log(n)\frac{\delta+1}{2\delta+1}\right\}$$
$$= \exp\left\{-n\epsilon_n^2\left(\frac{4(\delta+1)}{2\delta+1} - \frac{6\log\epsilon_n}{n\epsilon_n^2}\right)\right\}$$
$$\geq \exp\left\{-n\epsilon_n^2\left(\frac{4(\delta+1)}{2\delta+1} + \frac{6\delta\epsilon_n^{1/\delta}}{2\delta+1}\right)\right\} \asymp \exp\left\{-4n\epsilon_n^2\right\} \quad \text{(F.15)}$$

if $m_j^* \asymp \left(\frac{n}{\log n}\right)^{1/(2\delta+1)}$ for $j = 0, 1, n, a$ so that $m_j^* \log n \asymp n\epsilon_n^2$ and for every N such that $\epsilon_n^{1/\delta} \leq 2/3$ for $n > N$. This establishes (F.11).

□

Lemma F.2. *Assume the conditions of Theorem F.1 hold. Then, for $k = 0, 1, 0n, 1a$ the sequence of measurable sets $\mathscr{C}_{n,k}$ defined in* (D.5) *satisfies*

$$\pi(\mathscr{G}_k \backslash \mathscr{C}_{n,k}) \lesssim \exp\left\{-n_k\epsilon_n^2\frac{\eta}{2\delta+1}\right\}. \quad \text{(F.16)}$$

Proof. The proof is the same as the proof of Lemma E.2 and so we omit it.

□

Lemma F.3. *Let us consider the model specified in Assumption 7. Let $B_n^{KL}((\boldsymbol{g}^*, \boldsymbol{\sigma_*^2}, \boldsymbol{q}^*), \epsilon_n)$ be as defined in* (D.4) *and π be an independent prior on $(\boldsymbol{g}, \boldsymbol{\sigma^2}, \boldsymbol{\lambda}, \boldsymbol{q})$. Then, there exists an $N > 0$ such that $\forall n \geq N$ and $\forall \epsilon_n > 0$*

$$\pi(B_n^{KL}((\boldsymbol{g}^*, \boldsymbol{\sigma_*^2}, \boldsymbol{q}^*), \epsilon_n)) \geq \prod_{j\in\{0,1,0n,1a\}} \pi\left(|\sigma_{j*}^2 - \sigma_j^2| \leq \epsilon_n\sigma_{j*}^2/(2\sqrt{6\nu}), \sigma_j^2 \geq \sigma_{j*}^2\right)$$
$$\times \prod_{j\in\{0,1,0n,1a\}} \pi\left(\|g_j - g_j^*\|_\infty^2 \leq \frac{\epsilon_n^2\sigma_{j*}^2}{6}\right)\pi\left(\ln\left(\frac{q_\ell^*}{q_\ell}\right) \leq \epsilon_n^2/18,\ \ell = c, n, a\right). \quad \text{(F.17)}$$

Moreover, let the prior on $(\boldsymbol{\sigma^2}, \boldsymbol{q})$ be specified as in (D.6)*. Then, there exists an $N > 0$ such that $\forall n \geq N$ and $\forall \epsilon_n > 0$*

$$\pi(B_n^{KL}((\boldsymbol{g}^*, \boldsymbol{\sigma_*^2}, \boldsymbol{q}^*), \epsilon_n)) \gtrsim e^{6\log\epsilon_n} \prod_{j\in\{0,1,0n,1a\}} \pi\left(\|g_j - g_j^*\|_\infty^2 \leq \frac{\epsilon_n^2\sigma_{j*}^2}{6}\right). \quad \text{(F.18)}$$

Proof. By using similar steps as in the proof of Lemma (E.3), it is easy to show that

$$B_n^{KL}((\boldsymbol{g}^*, \boldsymbol{\sigma^2_*}, \boldsymbol{q}^*), \epsilon_n) \supseteq \Big\{(g_j, \sigma_j^2) \in \mathscr{G}_j \times \mathbb{R}_+ \text{ for } j = 0, 1, 0n, 1a, \boldsymbol{q} \in [0,1]^3; \frac{1}{n}\sum_{i=1}^n K(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq \epsilon_n^2, \frac{1}{n}\sum_{i=1}^n V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq C_2\epsilon_n^2\Big\}$$

for some constant $C_2 > 0$. In the following of the proof we leave implicit the fact that $(g_j, \sigma_j^2) \in \mathscr{G}_j \times \mathbb{R}_+$ for $j = 0, 1, 0n, 1a, \boldsymbol{q} \in [0,1]^3$ in all the events that we define.

Denote $\mathscr{A}_1 \triangleq \{\frac{1}{n}\sum_{i=1}^n K(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq \epsilon_n^2\}$. We first upper bound $K(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i)$ in order to find an event included in $\mathscr{A}_1$. We simplify notation by defining: $g_{ji} \triangleq g_j(z_i)$ and $g_{ji}^* \triangleq g_j^*(z_i)$ for $j = 0, 1, 0n, 1a$. Consider the observations in I_{00}: $\forall i \in I_{00}$:

$$\begin{aligned} K(P_*^i, P_{g_0,\sigma_0^2}^i) &= \\ \int \ln\left[\frac{t_\nu(y_i|g_{0i}^*, \sigma_*^2)q_c^* + t_\nu(y_i|g_{0ni}^*, \sigma_{0n*}^2)q_n^*}{t_\nu(y_i|g_{0i}, \sigma^2)q_c + t_\nu(y_i|g_{0ni}, \sigma_{0n}^2)q_n}\right] &\left(t_\nu(y_i|g_{0i}^*, \sigma_*^2)q_c^* + t_\nu(y_i|g_{0ni}^*, \sigma_{0n*}^2)q_n^*\right) dy_i \\ &\leq \int \ln\left[\frac{t_\nu(y_i|g_{0i}^*, \sigma_*^2)q_c^*}{t_\nu(y_i|g_{0i}, \sigma^2)q_c}\right] t_\nu(y_i|g_{0i}^*, \sigma_*^2)dy_i \\ &\quad + \int \ln\left[\frac{t_\nu(y_i|g_{0ni}^*, \sigma_{0n*}^2)q_n^*}{t_\nu(y_i|g_{0ni}, \sigma_{0n}^2)q_n}\right] t_\nu(y_i|g_{0ni}^*, \sigma_{0n*}^2)dy_i \end{aligned} \quad \text{(F.19)}$$

where to get the last inequality we have used the inequality: $a\log(a/b) \leq \sum_i a_i \log(a_i/b_i)$ for $a_i, b_i \geq 0$ and $a = \sum_i a_i$, $b = \sum_i b_i$, and the fact that $q_c, q_n \leq 1$. Hence, to get the explicit form for each integral, we can use (E.26) with $g_{j,i}$, $g_{j,i}^*$, σ_j^2 and σ_{j*}^2 replaced by g_{0i}, g_{0i}^*, σ_0^2 and σ_{0*}^2, respectively (and by $g_{0n,i}$, $g_{0n,i}^*$, σ_{0n}^2 and σ_{0n*}^2, respectively) to get:

$$\begin{aligned} \forall i \in I_{00}, \qquad K(P_*^i, P_{g_0,\sigma_0^2}^i) \leq &\frac{1}{2\sigma_{0*}^2}\left(\frac{\nu(\sigma_{0*}^2 - \sigma_0^2)^2}{\sigma_0^2} + (g_{0i}^* - g_{0i})^2\right) + \ln\left(\frac{q_c^*}{q_c}\right) \\ &+ \frac{1}{2\sigma_{0n*}^2}\left(\frac{\nu(\sigma_{0n*}^2 - \sigma_{0n}^2)^2}{\sigma_{0n}^2} + (g_{0ni}^* - g_{0ni})^2\right) + \ln\left(\frac{q_n^*}{q_n}\right). \end{aligned} \quad \text{(F.20)}$$

In a similar way, the following holds $\forall i \in I_{11}$:

$$K(P_*^i, P_{g_1,\sigma_1^2}^i) \leq \frac{1}{2\sigma_{1*}^2}\left(\frac{\nu(\sigma_{1*}^2-\sigma_1^2)^2}{\sigma_1^2} + (g_{1i}^*-g_{1i})^2\right) + \ln\left(\frac{q_c^*}{q_c}\right) + \frac{1}{2\sigma_{1a*}^2}\left(\frac{\nu(\sigma_{1a*}^2-\sigma_{1a}^2)^2}{\sigma_{1a}^2} + (g_{1ai}^*-g_{1ai})^2\right) + \ln\left(\frac{q_a^*}{q_a}\right), \quad \text{(F.21)}$$

while for the observations in I_{01} and I_{10} result (E.26) directly applies. By putting these results together we get:

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^n K(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq{}& \frac{1}{n}\sum_{i\in I_{00}} \frac{1}{2\sigma_{0*}^2}\left(\frac{\nu(\sigma_{0*}^2-\sigma_0^2)^2}{\sigma_0^2} + (g_{0i}^*-g_{0i})^2\right)\\
&+ \frac{1}{n}\sum_{i\in I_{00}} \frac{1}{2\sigma_{0n*}^2}\left(\frac{\nu(\sigma_{0n*}^2-\sigma_{0n}^2)^2}{\sigma_{0n}^2} + (g_{0ni}^*-g_{0ni})^2\right) + \ln\left(\frac{q_n^*}{q_n}\right) + \ln\left(\frac{q_c^*}{q_c}\right)\\
&+ \frac{1}{n}\sum_{i\in I_{10}} \frac{1}{2\sigma_{0n*}^2}\left(\frac{\nu(\sigma_{0n*}^2-\sigma_{0n}^2)^2}{\sigma_{0n}^2} + (g_{0ni}^*-g_{0ni})^2\right) + \ln\left(\frac{q_n^*}{q_n}\right)\\
&+ \frac{1}{n}\sum_{i\in I_{01}} \frac{1}{2\sigma_{1a*}^2}\left(\frac{\nu(\sigma_{1a*}^2-\sigma_{1a}^2)^2}{\sigma_{1a}^2} + (g_{1ai}^*-g_{1ai})^2\right) + \ln\left(\frac{q_a^*}{q_a}\right)\\
&+ \frac{1}{n}\sum_{i\in I_{11}} \frac{1}{2\sigma_{1*}^2}\left(\frac{\nu(\sigma_{1*}^2-\sigma_1^2)^2}{\sigma_1^2} + (g_{1i}^*-g_{1i})^2\right)\\
&+ \frac{1}{n}\sum_{i\in I_{11}} \frac{1}{2\sigma_{1a*}^2}\left(\frac{\nu(\sigma_{1a*}^2-\sigma_{1a}^2)^2}{\sigma_{1a}^2} + (g_{1ai}^*-g_{1ai})^2\right) + \ln\left(\frac{q_a^*}{q_a}\right) + \ln\left(\frac{q_c^*}{q_c}\right). \quad \text{(F.22)}
\end{aligned}$$

Denote

$$\begin{aligned}
\mathscr{A}_2 \triangleq \Big\{\frac{1}{n}\Big(&\sum_{i\in I_{00}}\left[\frac{(g_{0i}-g_{0i}^*)^2}{\sigma_{0*}^2} + \frac{(g_{0ni}-g_{0ni}^*)^2}{\sigma_{0n*}^2}\right] + \sum_{i\in I_{10}}\frac{(g_{0ni}-g_{0ni}^*)^2}{\sigma_{0n*}^2} + \sum_{i\in I_{01}}\frac{(g_{1ai}-g_{1ai}^*)^2}{\sigma_{1a*}^2}\\
&+ \sum_{i\in I_{11}}\left[\frac{(g_{1i}-g_{1i}^*)^2}{\sigma_{1*}^2} + \frac{(g_{1ai}-g_{1ai}^*)^2}{\sigma_{1a*}^2}\right]\Big) \leq \frac{2\epsilon_n^2}{3}\Big\},
\end{aligned}$$

and

$$\mathscr{A}_3 \triangleq \{|\sigma_{j*}^2-\sigma_j^2|^2 \leq \epsilon_n^2\sigma_{j*}^4/(24\nu), \sigma_j^2 \geq \sigma_{j*}^2, \text{ for } j=0,1,0n,1a\}$$

$$\mathscr{A}_4 \triangleq \{\ln\left(\frac{q_c^*}{q_c}\right) \leq \epsilon_n^2/18,\ \ln\left(\frac{q_n^*}{q_n}\right) \leq \epsilon_n^2/18,\ \ln\left(\frac{q_a^*}{q_a}\right) \leq \epsilon_n^2/18\}.$$

Therefore, $\mathscr{A}_1 \supseteq \mathscr{A}_2 \cap \mathscr{A}_3 \cap \mathscr{A}_4$.

Define the event $\mathscr{A}_5 \triangleq \{\|g_j - g_j^*\|_\infty^2 \leq \epsilon_n^2 \sigma_{j*}^2/6, \text{ for } j = 0, 1, 0n, 1a\}$ and remark that

$$\pi(B_n^{KL}((\boldsymbol{g}^*, \boldsymbol{\sigma_*^2}, \boldsymbol{q}^*), \epsilon_n)) \geq \int_{\mathscr{A}_2 \cap \mathscr{A}_3 \cap \mathscr{A}_4 \cap \{\frac{1}{n}\sum_{i=1}^n V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq C_2 \epsilon_n^2\}} d\pi(\boldsymbol{g}, \boldsymbol{\sigma^2}, \boldsymbol{q})$$

$$= \int_{\mathscr{A}_2 \cap \mathscr{A}_3 \cap \mathscr{A}_4 \cap \{\frac{1}{n}\sum_{i=1}^n V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq C_2 \epsilon_n^2\} \cap \mathscr{A}_5} d\pi(\boldsymbol{g}, \boldsymbol{\sigma^2}, \boldsymbol{q}) + \int_{\mathscr{A}_2 \cap \mathscr{A}_3 \cap \mathscr{A}_4 \cap \{\frac{1}{n}\sum_{i=1}^n V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq C_2 \epsilon_n^2\} \cap \mathscr{A}_5^c} d\pi(\boldsymbol{g}, \boldsymbol{\sigma^2}, \boldsymbol{q})$$

$$\geq \int_{\mathscr{A}_2 \cap \mathscr{A}_3 \cap \mathscr{A}_4 \cap \{\frac{1}{n}\sum_{i=1}^n V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq C_2 \epsilon_n^2\} \cap \mathscr{A}_5} d\pi(\boldsymbol{g}, \boldsymbol{\sigma^2}, \boldsymbol{q}).$$

Therefore, we have to analyze the event $\mathscr{A}_2 \cap \mathscr{A}_3 \cap \mathscr{A}_4 \cap \{\frac{1}{n}\sum_{i=1}^n V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq C_2 \epsilon_n^2\} \cap \mathscr{A}_5$. On $\mathscr{A}_2 \cap \mathscr{A}_3 \cap \mathscr{A}_4 \cap \mathscr{A}_5$, $(\boldsymbol{g}, \boldsymbol{\sigma^2}, \boldsymbol{q})$ is close to $(\boldsymbol{g}, \boldsymbol{\sigma^2}, \boldsymbol{q})$. Therefore, on $\mathscr{A}_2 \cap \mathscr{A}_3 \cap \mathscr{A}_4 \cap \mathscr{A}_5$, $P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i$ is close to P_*^i and $u \triangleq \frac{p_*(y_i, x_i|z_i, \boldsymbol{g}^*, \boldsymbol{\sigma_*^2}, \boldsymbol{q}^*)}{p_{\boldsymbol{g},\boldsymbol{\sigma^2}}(y_i, x_i|z_i, \boldsymbol{g}, \boldsymbol{\sigma^2}, \boldsymbol{q})}$ is close to 1 for n sufficiently large and for every (y_i, x_i, z_i). Hence, since the function $f(u) \triangleq u(\log u)^2$ is convex for every $u > e^{-1}$ we can use the inequality: $a\left[\log(a/b)\right]^2 \leq \sum_i a_i \left[\log(a_i/b_i)\right]^2$ for $a_i, b_i \geq 0$ and $a = \sum_i a_i$, $b = \sum_i b_i$, which for n sufficiently large is valid with $a = p_*(y_i, x_i|z_i, \boldsymbol{g}^*, \boldsymbol{\sigma_*^2}, \boldsymbol{q}^*)$ and $b = p_{\boldsymbol{g},\boldsymbol{\sigma^2}}(y_i, x_i|z_i, \boldsymbol{g}, \boldsymbol{\sigma^2}, \boldsymbol{q})$. By using this result, on $\mathscr{A}_2 \cap \mathscr{A}_3 \cap \mathscr{A}_4 \cap \mathscr{A}_5$ we have for n sufficiently large: $\forall i \in I_{00}$,

$$\begin{aligned}
V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) &= \int \left|\ln(u) - K(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i)\right|^2 \left(t_\nu(y_i|g_{0i}^*, \sigma_{0*}^2)q_c^* + t_\nu(y_i|g_{ni}^*, \sigma_{0n*}^2)q_n^*\right) dy_i \\
&\leq \int |\ln(u)|^2 \left(t_\nu(y_i|g_{0i}^*, \sigma_*^2)q_c^* + t_\nu(y_i|g_{ni}^*, \sigma_{0n*}^2)q_n^*\right) dy_i \\
&\leq \int \left|\ln\left(\frac{t_\nu(y_i|g_{0i}^*, \sigma_{0*}^2)q_c^*}{t_\nu(y_i|g_{0i}, \sigma_0^2)q_c}\right)\right|^2 t_\nu(y_i|g_{0i}^*, \sigma_{0*}^2)q_c^* dy_i \\
&\quad + \int \left|\ln\left(\frac{t_\nu(y_i|g_{0ni}^*, \sigma_{0n*}^2)q_n^*}{t_\nu(y_i|g_{0ni}, \sigma_{0n}^2)q_n}\right)\right|^2 t_\nu(y_i|g_{0ni}^*, \sigma_{0n*}^2)q_n^* dy_i. \quad \text{(F.23)}
\end{aligned}$$

We start by analyzing the first term in the right hand side of (F.23):

$$
\begin{aligned}
\int & \left|\ln\left(\frac{t_\nu(y_i|g_{0i}^*,\sigma_{0*}^2)q_c^*}{t_\nu(y_i|g_{0i},\sigma_0^2)q_c}\right)\right|^2 t_\nu(y_i|g_{0i}^*,\sigma_{0*}^2)q_c^* dy_i \\
&= \int \left|\ln\left(\left(\frac{\sigma_{0*}}{\sigma_0}\right)^\nu \left(\frac{\nu\sigma_0^2+(y_i-g_{0i})^2}{\nu\sigma_{0*}^2+(y_i-g_{0i}^*)^2}\right)^{(\nu+1)/2} \frac{q_c^*}{q_c}\right)\right|^2 t_\nu(y_i|g_{0i}^*,\sigma_{0*}^2)q_c^* dy_i \\
\leq 4\frac{\nu^2}{4}&\left(\ln\left(\frac{\sigma_{0*}^2}{\sigma_0^2}\right)\right)^2 + 2\int\left|\ln\left(\frac{\nu\sigma_0^2+(y_i-g_{0i})^2}{\nu\sigma_{0*}^2+(y_i-g_{0i}^*)^2}\right)^{(\nu+1)/2}\right|^2 t_\nu(y_i|g_{0i}^*,\sigma_{0*}^2)q_c^* dy_i + 4\ln^2\left(\frac{q_c^*}{q_c}\right).
\end{aligned}
\tag{F.24}
$$

Since we are on $\mathscr{A}_3$ and by using the inequality $\ln(x) \leq x-1, \forall x > 0$ we have the upper bound:

$$
4\frac{\nu^2}{4}\left(\ln\left(\frac{\sigma_{0*}^2}{\sigma_0^2}\right)\right)^2 = 4\frac{\nu^2}{4}\left(\ln\left(\frac{\sigma_0^2}{\sigma_{0*}^2}\right)\right)^2 \leq \nu^2\frac{(\sigma_0^2-\sigma_{0*}^2)^2}{\sigma_{0*}^4} \leq \frac{\nu\epsilon_n^2}{9}. \tag{F.25}
$$

By using again the inequality $\ln(x) \leq x-1, \forall x > 0$, and the fact that $\left(\frac{\nu\sigma_{0*}^2}{\nu\sigma_{0*}^2+(y_i-g_{0i}^*)^2}\right)^2 t_\nu(y_i|g_{0i}^*,\sigma_{0*}^2) = \frac{(\nu+2)\nu}{(\nu+3)(\nu+1)}t_{\nu+4}(y_i|g_{0i}^*,\bar\sigma_{0*}^2)$ with $\bar\sigma_{0*}^2 \triangleq \frac{\nu\sigma_{0*}^2}{\nu+4}$ we have the upper bound:

$$
\begin{aligned}
2\int & \left|\ln\left(\frac{\nu\sigma_0^2+(y_i-g_{0i})^2}{\nu\sigma_{0*}^2+(y_i-g_{0i}^*)^2}\right)^{(\nu+1)/2}\right|^2 t_\nu(y_i|g_{0i}^*,\sigma_{0*}^2)q_c^* dy_i \\
&\leq \frac{(\nu+1)^2}{2}\int\left|\frac{\nu(\sigma_0^2-\sigma_{0*}^2)+(g_{0i}-g_{0i}^*)^2+2(y_i-g_{0i}^*)(g_{0i}^*-g_{0i})}{\nu\sigma_{0*}^2+(y_i-g_{0i}^*)^2}\right|^2 t_\nu(y_i|g_{0i}^*,\sigma_{0*}^2)q_c^* dy_i \\
= \frac{\nu(\nu+1)(\nu+2)}{2(\nu+3)}&\left[\frac{(\sigma_0^2-\sigma_{0*}^2)^2}{\sigma_{0*}^4} + \frac{(g_{0i}-g_{0i}^*)^4}{\nu^2\sigma_{0*}^4} + \frac{4(g_{0i}^*-g_{0i})^2\bar\sigma_{0*}^2(\nu+4)}{\nu^2\sigma_{0*}^4(\nu+2)} + 2\frac{(\sigma_0^2-\sigma_{0*}^2)}{\nu\sigma_{0*}^4}(g_{0i}-g_{0i}^*)^2\right] \\
&= \frac{\nu(\nu+1)(\nu+2)}{2(\nu+3)\sigma_{0*}^4}(\sigma_0^2-\sigma_{0*}^2)^2 + \frac{(\nu+1)(\nu+2)}{2\nu\sigma_{0*}^4(\nu+3)}(g_{0i}-g_{0i}^*)^4 + \frac{2(\nu+1)}{\sigma_{0*}^2(\nu+3)}(g_{0i}^*-g_{0i})^2 \\
&\qquad + \frac{(\nu+1)(\nu+2)}{(\nu+3)\sigma_{0*}^4}(g_{0i}-g_{0i}^*)^2(\sigma_0^2-\sigma_{0*}^2).
\end{aligned}
\tag{F.26}
$$

By substituting (F.25) and (F.26) in (F.24) and using the fact that we are on $\mathscr{A}_3$ to further upper bound (F.26), we obtain:

$$\int \left| \ln \left(\frac{t_\nu(y_i|g_{0i}^*, \sigma_{0*}^2) q_c^*}{t_\nu(y_i|g_{0i}, \sigma_0^2) q_c} \right) \right|^2 t_\nu(y_i|g_{0i}^*, \sigma_{0*}^2) q_c^* dy_i$$
$$\leq \frac{\nu\epsilon_n^2}{9} + \frac{\epsilon_n^2(\nu+1)(\nu+2)}{48(\nu+3)} + \frac{(\nu+1)(\nu+2)}{2\nu\sigma_{0*}^4(\nu+3)}(g_{0i} - g_{0i}^*)^4 + \frac{2(\nu+1)}{\sigma_{0*}^2(\nu+3)}(g_{0i}^* - g_{0i})^2$$
$$+ \frac{\epsilon_n(\nu+1)(\nu+2)}{2\sqrt{6\nu}(\nu+3)\sigma_{0*}^2}(g_{0i} - g_{0i}^*)^2 + 4\ln^2\left(\frac{q_c^*}{q_c}\right). \quad \text{(F.27)}$$

In a similar way, we obtain the upper bound for the second term in the right hand side of (F.23):

$$\int \left| \ln \left(\frac{t_\nu(y_i|g_{0ni}^*, \sigma_{0n*}^2) q_n^*}{t_\nu(y_i|g_{0ni}, \sigma_{0n}^2) q_n} \right) \right|^2 t_\nu(y_i|g_{0ni}^*, \sigma_{0n*}^2) q_n^* dy_i$$
$$\leq \frac{\nu\epsilon_n^2}{9} + \frac{\epsilon_n^2(\nu+1)(\nu+2)}{48(\nu+3)} + \frac{(\nu+1)(\nu+2)}{2\nu\sigma_{0n*}^4(\nu+3)}(g_{0ni} - g_{0ni}^*)^4 + \frac{2(\nu+1)}{\sigma_{0n*}^2(\nu+3)}(g_{0ni}^* - g_{0ni})^2$$
$$+ \frac{\epsilon_n(\nu+1)(\nu+2)}{2\sqrt{6\nu}(\nu+3)\sigma_{0n*}^2}(g_{0ni} - g_{0ni}^*)^2 + 4\ln^2\left(\frac{q_n^*}{q_n}\right) \quad \text{(F.28)}$$

and therefore,

$$\frac{1}{n}\sum_{i\in I_{00}} V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq \frac{2\nu\epsilon_n^2}{9} + \frac{\epsilon_n^2(\nu+1)(\nu+2)}{24(\nu+3)} + \frac{(\nu+1)(\nu+2)}{2\nu(\nu+3)}\frac{1}{n}\sum_{i\in I_{00}}\left[\frac{(g_{0i} - g_{0i}^*)^4}{\sigma_{0*}^4} + \frac{(g_{0ni} - g_{0ni}^*)^4}{\sigma_{0n*}^4}\right]$$
$$+ \frac{2(\nu+1)}{(\nu+3)}\frac{1}{n}\sum_{i\in I_{00}}\left[\frac{(g_{0i}^* - g_{0i})^2}{\sigma_{0*}^2} + \frac{(g_{0ni}^* - g_{0ni})^2}{\sigma_{0n*}^2}\right]$$
$$+ \frac{\epsilon_n(\nu+1)(\nu+2)}{2\sqrt{6\nu}(\nu+3)}\frac{1}{n}\sum_{i\in I_{00}}\left[\frac{(g_{0i} - g_{0i}^*)^2}{\sigma_{0*}^2} + \frac{(g_{0ni} - g_{0ni}^*)^2}{\sigma_{0n*}^2}\right] + 4\left(\ln^2\left(\frac{q_c^*}{q_c}\right) + \ln^2\left(\frac{q_n^*}{q_n}\right)\right). \quad \text{(F.29)}$$

By using similar arguments we obtain the following upper bounds for the other cells I_{10}, I_{01} and I_{11}:

$$\frac{1}{n}\sum_{i\in I_{10}} V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq \frac{\nu\epsilon_n^2}{9} + \frac{\epsilon_n^2(\nu+1)(\nu+2)}{48(\nu+3)} + \frac{(\nu+1)(\nu+2)}{2\nu(\nu+3)}\frac{1}{n}\sum_{i\in I_{10}}\frac{(g_{0ni} - g_{0ni}^*)^4}{\sigma_{0n*}^4}$$
$$+ \frac{2(\nu+1)}{(\nu+3)}\frac{1}{n}\sum_{i\in I_{10}}\frac{(g_{0ni}^* - g_{0ni})^2}{\sigma_{0n*}^2} + \frac{\epsilon_n(\nu+1)(\nu+2)}{2\sqrt{6\nu}(\nu+3)}\frac{1}{n}\sum_{i\in I_{10}}\frac{(g_{0ni} - g_{0ni}^*)^2}{\sigma_{0n*}^2} + 4\ln^2\left(\frac{q_n^*}{q_n}\right), \quad \text{(F.30)}$$

$$\frac{1}{n}\sum_{i\in I_{01}} V(P_*^i, P^i_{\boldsymbol{g},\boldsymbol{\sigma^2}}) \le \frac{\nu\epsilon_n^2}{9} + \frac{\epsilon_n^2(\nu+1)(\nu+2)}{48(\nu+3)} + \frac{(\nu+1)(\nu+2)}{2\nu(\nu+3)}\frac{1}{n}\sum_{i\in I_{01}}\frac{(g_{1ai}-g^*_{1ai})^4}{\sigma^4_{1a*}}$$
$$+\frac{2(\nu+1)}{(\nu+3)}\frac{1}{n}\sum_{i\in I_{01}}\frac{(g^*_{1ai}-g_{1ai})^2}{\sigma^2_{1a*}} + \frac{\epsilon_n(\nu+1)(\nu+2)}{2\sqrt{6\nu}(\nu+3)}\frac{1}{n}\sum_{i\in I_{01}}\frac{(g_{1ai}-g^*_{1ai})^2}{\sigma^2_{1a*}} + 4\ln^2\left(\frac{q^*_a}{q_a}\right) \quad \text{(F.31)}$$

and

$$\frac{1}{n}\sum_{i\in I_{11}} V(P_*^i, P^i_{\boldsymbol{g},\boldsymbol{\sigma^2}}) \le \frac{2\nu\epsilon_n^2}{9} + \frac{\epsilon_n^2(\nu+1)(\nu+2)}{24(\nu+3)} + \frac{(\nu+1)(\nu+2)}{2\nu(\nu+3)}\frac{1}{n}\sum_{i\in I_{11}}\left[\frac{(g_{1i}-g^*_{1i})^4}{\sigma^4_{1*}} + \frac{(g_{1ai}-g^*_{1ai})^4}{\sigma^4_{1a*}}\right]$$
$$+\frac{2(\nu+1)}{(\nu+3)}\frac{1}{n}\sum_{i\in I_{11}}\left[\frac{(g^*_{1i}-g_{1i})^2}{\sigma^2_{1*}} + \frac{(g^*_{1ai}-g_{1ai})^2}{\sigma^2_{1a*}}\right]$$
$$+\frac{\epsilon_n(\nu+1)(\nu+2)}{2\sqrt{6\nu}(\nu+3)}\frac{1}{n}\sum_{i\in I_{11}}\left[\frac{(g_{1i}-g^*_{1i})^2}{\sigma^2_{1*}} + \frac{(g_{1ai}-g^*_{1ai})^2}{\sigma^2_{1a*}}\right] + 4\left(\ln^2\left(\frac{q^*_c}{q_c}\right) + \ln^2\left(\frac{q^*_a}{q_a}\right)\right). \quad \text{(F.32)}$$

By putting together these results we obtain:

$$\frac{1}{n}\sum_{i=1}^{n} V(P_*^i, P^i_{\boldsymbol{g},\boldsymbol{\sigma^2}}) \le \frac{6\nu\epsilon_n^2}{9} + \frac{\epsilon_n^2(\nu+1)(\nu+2)}{8(\nu+3)} + \frac{(\nu+1)(\nu+2)}{2\nu(\nu+3)}\frac{1}{n}\Big(\sum_{i\in I_{00}}\left[\frac{(g_{0i}-g^*_{0i})^4}{\sigma^4_{0*}} + \frac{(g_{0ni}-g^*_{0ni})^4}{\sigma^4_{0n*}}\right]$$
$$+\sum_{i\in I_{10}}\frac{(g_{0ni}-g^*_{0ni})^4}{\sigma^4_{0n*}} + \sum_{i\in I_{01}}\frac{(g_{1ai}-g^*_{1ai})^4}{\sigma^4_{1a*}} + \sum_{i\in I_{11}}\left[\frac{(g_{1i}-g^*_{1i})^4}{\sigma^4_{1*}} + \frac{(g_{1ai}-g^*_{1ai})^4}{\sigma^4_{1a*}}\right]\Big)$$
$$+\frac{2(\nu+1)}{(\nu+3)}\frac{1}{n}\Big(\sum_{i\in I_{00}}\left[\frac{(g^*_{0i}-g_{0i})^2}{\sigma^2_{0*}} + \frac{(g^*_{0ni}-g_{0ni})^2}{\sigma^2_{0n*}}\right] + \sum_{i\in I_{10}}\frac{(g^*_{0ni}-g_{0ni})^2}{\sigma^2_{0n*}} + \sum_{i\in I_{01}}\frac{(g^*_{1ai}-g_{1ai})^2}{\sigma^2_{1a*}}$$
$$+\sum_{i\in I_{11}}\left[\frac{(g^*_{1i}-g_{1i})^2}{\sigma^2_{1*}} + \frac{(g^*_{1ai}-g_{1ai})^2}{\sigma^2_{1a*}}\right]\Big)$$
$$+\frac{\epsilon_n(\nu+1)(\nu+2)}{2\sqrt{6\nu}(\nu+3)}\frac{1}{n}\Big(\sum_{i\in I_{00}}\left[\frac{(g_{0i}-g^*_{0i})^2}{\sigma^2_{0*}} + \frac{(g_{0ni}-g^*_{0ni})^2}{\sigma^2_{0n*}}\right] + \sum_{i\in I_{10}}\frac{(g_{0ni}-g^*_{0ni})^2}{\sigma^2_{0n*}} + \sum_{i\in I_{01}}\frac{(g_{1ai}-g^*_{1ai})^2}{\sigma^2_{1a*}}$$
$$+\sum_{i\in I_{11}}\left[\frac{(g_{1i}-g^*_{1i})^2}{\sigma^2_{1*}} + \frac{(g_{1ai}-g^*_{1ai})^2}{\sigma^2_{1a*}}\right]\Big) + 4\left(2\ln^2\left(\frac{q^*_c}{q_c}\right) + 2\ln^2\left(\frac{q^*_n}{q_n}\right) + 2\ln^2\left(\frac{q^*_a}{q_a}\right)\right). \quad \text{(F.33)}$$

Since we are on $\mathscr{A}_2$, then we can further upper bound the previous quantity as

$$\frac{1}{n}\sum_{i=1}^{n} V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq \frac{6\nu\epsilon_n^2}{9} + \frac{\epsilon_n^2(\nu+1)(\nu+2)}{8(\nu+3)} + \frac{(\nu+1)(\nu+2)}{2\nu(\nu+3)}\frac{1}{n}\Big(\sum_{i\in I_{00}}\left[\frac{(g_{0i}-g_{0i}^*)^4}{\sigma_{0*}^4} + \frac{(g_{0ni}-g_{0ni}^*)^4}{\sigma_{0n*}^4}\right] + \sum_{i\in I_{10}}\frac{(g_{0ni}-g_{0ni}^*)^4}{\sigma_{0n*}^4} + \sum_{i\in I_{01}}\frac{(g_{1ai}-g_{1ai}^*)^4}{\sigma_{1a*}^4} + \sum_{i\in I_{11}}\left[\frac{(g_{1i}-g_{1i}^*)^4}{\sigma_{1*}^4} + \frac{(g_{1ai}-g_{1ai}^*)^4}{\sigma_{1a*}^4}\right]\Big) + \frac{2(\nu+1)}{(\nu+3)}\frac{2\epsilon_n^2}{3} + \frac{\epsilon_n(\nu+1)(\nu+2)}{2\sqrt{6\nu}(\nu+3)}\frac{2\epsilon_n^2}{3} + 4\left(2\ln^2\left(\frac{q_c^*}{q_c}\right) + 2\ln^2\left(\frac{q_n^*}{q_n}\right) + 2\ln^2\left(\frac{q_a^*}{q_a}\right)\right). \tag{F.34}$$

Moreover, since we are on $\mathscr{A}_4$, then we can further upper bound the previous quantity as

$$\frac{1}{n}\sum_{i=1}^{n} V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq \frac{6\nu\epsilon_n^2}{9} + \frac{\epsilon_n^2(\nu+1)(\nu+2)}{8(\nu+3)} + \frac{(\nu+1)(\nu+2)}{2\nu(\nu+3)}\frac{1}{n}\Big(\|g_0-g_0^*\|_\infty^2\sum_{i\in I_{00}}\frac{(g_{0i}-g_{0i}^*)^2}{\sigma_{0*}^4} + \|g_{0n}-g_{0n}^*\|_\infty^2\sum_{i\in I_{00}}\frac{(g_{0ni}-g_{0ni}^*)^2}{\sigma_{0n*}^4} + \|g_{0n}-g_{0n}^*\|_\infty^2\sum_{i\in I_{10}}\frac{(g_{0ni}-g_{0ni}^*)^2}{\sigma_{0n*}^4} + \|g_{1a}-g_{1a}^*\|_\infty^2\sum_{i\in I_{01}}\frac{(g_{1ai}-g_{1ai}^*)^2}{\sigma_{1a*}^4} + \|g_1-g_1^*\|_\infty^2\sum_{i\in I_{11}}\frac{(g_{1i}-g_{1i}^*)^2}{\sigma_{1*}^4} + \|g_{1a}-g_{1a}^*\|_\infty^2\sum_{i\in I_{11}}\frac{(g_{1ai}-g_{1ai}^*)^2}{\sigma_{1a*}^4}\Big) + \frac{2(\nu+1)}{(\nu+3)}\frac{2\epsilon_n^2}{3} + \frac{\epsilon_n(\nu+1)(\nu+2)}{2\sqrt{6\nu}(\nu+3)}\frac{2\epsilon_n^2}{3} + \frac{2\epsilon_n^4}{27}. \tag{F.35}$$

Finally, because we are on $\mathscr{A}_5$ we obtain:

$$\frac{1}{n}\sum_{i=1}^{n} V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq \frac{6\nu\epsilon_n^2}{9} + \frac{\epsilon_n^2(\nu+1)(\nu+2)}{8(\nu+3)} + \frac{\epsilon_n^4(\nu+1)(\nu+2)}{18\nu(\nu+3)} + \frac{2(\nu+1)}{(\nu+3)}\frac{2\epsilon_n^2}{3} + \frac{\epsilon_n(\nu+1)(\nu+2)}{2\sqrt{6\nu}(\nu+3)}\frac{2\epsilon_n^2}{3} + \frac{2\epsilon_n^4}{27} = \epsilon_n^2[C + \mathscr{O}(\epsilon_n) + \mathscr{O}(\epsilon_n^2)]. \tag{F.36}$$

This shows that on $\mathscr{A}_2 \cap \mathscr{A}_3 \cap \mathscr{A}_4 \cap \mathscr{A}_5$ there exists an $N > 0$ such that $\forall n \geq N$ it holds $\frac{1}{n}\sum_{i=1}^{n} V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq C_2\epsilon_n^2$ for $C_2 = C + \mathscr{O}(\epsilon_n)$, then $\{\frac{1}{n}\sum_{i=1}^{n} V(P_*^i, P_{\boldsymbol{g},\boldsymbol{\sigma^2}}^i) \leq C_2\epsilon_n^2\} \supseteq$

$\mathscr{A}_2 \cap \mathscr{A}_3 \cap \mathscr{A}_4 \cap \mathscr{A}_5$ and

$$\pi(B_n^{KL}((\boldsymbol{g}^*, \boldsymbol{\sigma_*^2}, \boldsymbol{q^*}), \epsilon_n)) \geq \pi\left(\mathscr{A}_2 \cap \mathscr{A}_3 \cap \mathscr{A}_4 \cap \mathscr{A}_5\right)$$
$$\geq \prod_{j\in\{0,1,0n,1a\}} \pi\left(|\sigma_{j*}^2 - \sigma_j^2| \leq \epsilon_n \sigma_{j*}^2/(2\sqrt{6\nu}), \sigma_j^2 \geq \sigma_{j*}^2\right) \pi\left(\ln\left(\frac{q_j^*}{q_j}\right) \leq \epsilon_n^2/18,\ j = c, n, a\right) \pi(\mathscr{A}_2 \cap \mathscr{A}_5).$$

By noticing that $\pi(\mathscr{A}_2 \cap \mathscr{A}_5) \supseteq \{\|g_j - g_j^*\|_\infty^2 \leq \frac{\epsilon_n^2 \sigma_{j*}^2}{6},\ \text{for } j = 0, 1, 0n, 1a\}$ since $\frac{1}{n_{00}} \sum_{i\in I_{00}} \frac{(g_{0i} - g_{0i*})^2}{\sigma_*^2} \leq \sigma_*^{-2} \|g_0 - g_0^*\|_\infty^2$ (and similarly for the other terms in $\mathscr{A}_2$), we obtain:

$$\pi(B_n^{KL}((\boldsymbol{g}^*, \boldsymbol{\sigma_*^2}, \boldsymbol{q^*}), \epsilon_n)) \geq \prod_{j\in\{0,1,0n,1a\}} \pi\left(|\sigma_{j*}^2 - \sigma_j^2| \leq \epsilon_n \sigma_{j*}^2/(2\sqrt{6\nu}), \sigma_j^2 \geq \sigma_{j*}^2\right)$$
$$\times \pi\left(\ln\left(\frac{q_j^*}{q_j}\right) \leq \epsilon_n^2/18,\ j = c, n, a\right) \prod_{j\in\{0,1,0n,1a\}} \pi\left(\|g_j - g_j^*\|_\infty^2 \leq \frac{\epsilon_n^2 \sigma_{j*}^2}{6}\right).$$

This establishes result (F.17) in the lemma. Finally we have to show result (F.18). For $j = 0, 1, 0n, 1a$,

$$\pi\left(|\sigma_{j*}^2 - \sigma_j^2| \leq \epsilon_n \sigma_{j*}^2/(2\sqrt{6\nu}), \sigma_j^2 \geq \sigma_{j*}^2\right) = \int_{(\sigma_{j*}^2 - \frac{\epsilon_n \sigma_{j*}^2}{2\sqrt{6\nu}}) \vee \sigma_{j*}^2}^{(\sigma_{j*}^2 + \frac{\epsilon_n \sigma_{j*}^2}{2\sqrt{6\nu}})} \frac{(\delta_{00}/2)^{\nu_{00}/2}}{\Gamma(\nu_{00}/2)} (\sigma_j^2)^{-\nu_{00}/2-1} e^{-\delta_{00}/(2\sigma_j^2)} d\sigma_j^2$$
$$\geq (\sigma_{j*}^2 + \frac{\epsilon_n \sigma_{j*}^2}{2\sqrt{6\nu}})^{-(\nu_{00}/2+1)} e^{-\delta_{00}/(2\sigma_{j*}^2)} \int_{\sigma_{j*}^2}^{(\sigma_{j*}^2 + \frac{\epsilon_n \sigma_{j*}^2}{2\sqrt{6\nu}})} \frac{(\delta_{00}/2)^{\nu_{00}/2}}{\Gamma(\nu_{00}/2)} d\sigma_j^2$$
$$= (\sigma_{j*}^2 + \frac{\epsilon_n \sigma_{j*}^2}{2\sqrt{6\nu}})^{-(\nu_{00}/2+1)} e^{-\delta_{00}/(2\sigma_{j*}^2)} \frac{(\delta_{00}/2)^{\nu_{00}/2}}{\Gamma(\nu_{00}/2)} \left[\left(\sigma_{j*}^2 + \frac{\epsilon_n \sigma_{j*}^2}{2\sqrt{6\nu}}\right) - \sigma_{j*}^2\right] \asymp e^{\log \epsilon_n}$$

since $(\sigma_{j*}^2 - \frac{\epsilon_n \sigma_{j*}^2}{\sqrt{\nu}}) \vee \sigma_{j*}^2 = \sigma_{j*}^2$. We now analyse the probability $\pi\left(\ln\left(\frac{q_j^*}{q_j}\right) \leq \epsilon_n^2/18,\ j = c, n, a\right)$. The event $\{\ln\left(\frac{q_j^*}{q_j}\right) \leq \epsilon_n^2/18,\ j = c, n, a\}$ and the restriction $q_a = 1 - q_n - q_c$ together are equivalent to the restrictions:

$$q_c^* e^{-\epsilon_n^2/18} \leq q_c, \quad q_n^* e^{-\epsilon_n^2/18} \leq q_n, \quad q_a^* e^{-\epsilon_n^2/18} \leq 1 - q_c - q_n,$$

where the last restriction $q_a^* e^{-\epsilon_n^2/18} \leq 1 - q_c - q_n$ is equivalent to $q_n \leq 1 - q_c - q_a^* e^{-\epsilon_n^2/18}$. By putting together these three restrictions and the restrictions $q_j < 1$, $j = c, n, a$ we get:

$$q_c^* e^{-\epsilon_n^2/18} \leq q_c \leq e^{-\epsilon_n^2/18} \quad \text{and} \quad q_n^* e^{-\epsilon_n^2/18} \leq q_n \leq \min\{1 - q_c - q_a^* e^{-\epsilon_n^2/18}, 1\}.$$

Denote by $\mathscr{A}_c$ and $\mathscr{A}_n$ these two events. The restrictions for q_n in $\mathscr{A}_n$ are feasible if and only if $\min\{1-q_c-q_a^* e^{-\epsilon_n^2/18}, 1\} > q_n^* e^{-\epsilon_n^2/18}$ which in turn is equivalent to $q_c < 1-q_n^* e^{-\epsilon_n^2/18}-q_a^* e^{-\epsilon_n^2/18}$. Therefore, since $\{q_c^* e^{-\epsilon_n^2/18} \leq q_c < 1-q_n^* e^{-\epsilon_n^2/18}-q_a^* e^{-\epsilon_n^2/18}\} = \{q_c^* e^{-\epsilon_n^2/18} \leq q_c < 1+q_c^* e^{-\epsilon_n^2/18}-e^{-\epsilon_n^2/18}\} \triangleq \mathscr{B}_c$, we have that

$$\pi\left(\ln\left(\frac{q_j^*}{q_j}\right) \leq \epsilon_n^2/18,\ j=c,n,a\right) \geq \pi(\mathscr{A}_c \cap \mathscr{A}_n | \mathscr{B}_c)\pi(\mathscr{B}_c)$$

where conditional on $\mathscr{B}_c$ the event $\mathscr{A}_c \cap \mathscr{A}_n$ has a strictly positive probability. Moreover, since $q_c \sim \mathscr{B}e(n_{0c}, n_{0n}+n_{0a})$,

$$\begin{aligned}\pi(\mathscr{B}_c) = \int_{q_c^* e^{-\epsilon_n^2/18}}^{1+q_c^* e^{-\epsilon_n^2/18}-e^{-\epsilon_n^2/18}} & q_c^{n_{0c}}(1-q_c)^{n_{0n}+n_{0a}} dq_c \\ & > (q_c^* e^{-\epsilon_n^2/18})^{n_{0c}}(1-(1+q_c^* e^{-\epsilon_n^2/18}-e^{-\epsilon_n^2/18}))^{n_{0n}+n_{0a}}(1-e^{-\epsilon_n^2/18}) \\ = (q_c^* e^{-\epsilon_n^2/18})^{n_{0c}}(e^{-\epsilon_n^2/18}-q_c^* e^{-\epsilon_n^2/18})^{n_{0n}+n_{0a}} & \frac{\epsilon_n^2}{18}\left(1-\frac{\epsilon_n^2}{18(2!)}+\frac{\epsilon_n^4}{(18)^2(3!)}-\dots\right) \gtrsim \epsilon_n^2 \end{aligned} \quad \text{(F.37)}$$

where we have used a Taylor expansion of the exponential function in the last line. So, we conclude that $\pi\left(\ln\left(\frac{q_j^*}{q_j}\right) \leq \epsilon_n^2/18,\ j=c,n,a\right) \gtrsim e^{2\log(\epsilon_n)}$.

□

F.4 Testing

Lemma F.4 (Testing). *For a given sequence* $\boldsymbol{s} \triangleq (\boldsymbol{s_0}, \boldsymbol{s_1})$ *where* $\boldsymbol{s_k} \triangleq \{s_i\}_{i\in I_{kk}}$ *with values in* $\{c,n\}$ *if* $k=0$ *and in* $\{c,a\}$ *if* $k=1$ *recall the definition of* $\boldsymbol{f_s}$ *given in section D.1. Moreover,* $\Xi \triangleq \mathrm{diag}(\xi_i, i \in I_{00}, \xi_i \in I_{10}, \xi_i \in I_{01}, \xi_i \in I_{11})$.

For a given $\boldsymbol{s}$, $\ell \in \mathbb{N}$, $\varepsilon \triangleq 12\max(\sigma_{0*}, \sigma_{1*}, \sigma_{0n*}, \sigma_{1a*})M\sqrt{K}\epsilon_n$ *for some constants* $K, M > 0$ *and* $\varepsilon_\sigma > 1/(2n^2)$, *let* $\mathscr{A}_{\ell,\boldsymbol{s}} \triangleq \{(f,\sigma_k^2) \in \mathscr{F}_n \times \left[\frac{1}{2n}, e^{6n\epsilon_n^2}\right]$ *for* $k = 0,1,1a,0n; \|\Xi^{1/2}(\boldsymbol{f_s} - \boldsymbol{f_s^*})\|_n > \ell\varepsilon, |1-\sigma_k/\sigma_{k*}| > \ell\varepsilon_\sigma\}$. *Denote* $b \triangleq (\boldsymbol{f_s}, \boldsymbol{\sigma^2_{0s_0}}, \boldsymbol{\sigma^2_{1s_1}}, \sigma_{0n}^2, \sigma_{1a}^2)$ *and by* b_* *its true value. For each* b, *let* $\mathbf{E}_b$ *denote the expectation taken with respect to the distribution* P_b*:*

$$P_b = \prod_{i\in I_{00}} \mathscr{N}(g_{0s_i}, \sigma_{0s_i}^2/\xi_i) \prod_{i\in I_{11}} \mathscr{N}(g_{1s_i}, \sigma_{1s_i}^2/\xi_i) \prod_{i\in I_{10}} \mathscr{N}(g_{0n}, \sigma_{0n}^2/\xi_i) \prod_{i\in I_{01}} \mathscr{N}(g_{1a}, \sigma_{1a}^2/\xi_i). \quad \text{(F.38)}$$

Then, for every given sequence $\boldsymbol{s} \triangleq (\boldsymbol{s_0}, \boldsymbol{s_1})$ *where* $\boldsymbol{s_k} = \{s_i\}_{i \in I_{kk}}$, $k = 0, 1$, *there exists a test* ϕ_n *such that for some* $K, M > 0$ *and* $\forall \ell \in \mathbb{N}$*:*

$$\mathbf{E}_{b_*}\phi_n \leq \frac{e^{-(M^2K-1)n\epsilon_n^2}}{1 - e^{-M^2Kn\epsilon_n^2}}, \qquad \sup_{b \in \mathscr{A}_{\ell,\boldsymbol{s}}} \mathbf{E}_b(1-\phi_n) \leq e^{-M^2K\ell n\epsilon_n^2}. \tag{F.39}$$

Proof. To construct the test ϕ_n for a given sequence $\boldsymbol{s}$ we first consider the set $\mathscr{Q}_j := \{(f, \sigma_k^2) \in \mathscr{F}_n \times \left[\frac{1}{2n}, e^{6n\epsilon_n^2}\right]$ for $k = 0, 1, 1a, 0n; j\varepsilon < \|\Xi^{1/2}(\boldsymbol{f_s} - \boldsymbol{f_s^*})\|_n < 2j\varepsilon, j\varepsilon_\sigma < |1 - \sigma_k/\sigma_{k*}| < 2j\varepsilon_\sigma\}$ for a given $j \in \mathbb{N}$, $\varepsilon > 0$ and $\varepsilon_\sigma > 0$. Consider a maximal set of points $b_{j,1}, \ldots, b_{j,N_j} \in \mathscr{Q}_j$, with $b_{j,l} = (\boldsymbol{f_s^{jl}}, \boldsymbol{\sigma^2_{0s_0,jl}}, \boldsymbol{\sigma^2_{1s_1,jl}}, \sigma^2_{0n,jl}, \sigma^2_{1a,jl})$ such that for every $l, l' = 1, \ldots, N_j$ satisfy $\|\Xi_k(\boldsymbol{f_s^{jl}} - \boldsymbol{f_s^{jl'}})\|_n \geq \varepsilon\sigma_{k\min*}/(3\sigma_{k\max*})$ and $|\sigma^2_{k,jl} - \sigma^2_{k,jl'}| > \varepsilon_\sigma\sigma^2_{k*}/2$ for $k = 0, 1, 0n, 1a$ and where we have denoted $\sigma_{0\max*} \triangleq \max\{\sigma_{0*}, \sigma_{0n*}\}$, $\sigma_{1\max*} \triangleq \max\{\sigma_{1*}, \sigma_{1a*}\}$, $\sigma_{0\min*} \triangleq \min\{\sigma_{0*}, \sigma_{0n*}\}$ and $\sigma_{1\min*} \triangleq \min\{\sigma_{1*}, \sigma_{1a*}\}$ for $k = 0, 1$.

Then, consider N_j balls $B_{j,l}$ around each of these $b_{j,l}$ constructed as

$$B_{j,l} := \Big\{b \in \mathscr{Q}_j; \|\Xi_k^{1/2}(\boldsymbol{f_s} - \boldsymbol{f_s^{jl}})\|_n \leq \varepsilon\sigma_{k\min*}/(3\sigma_{k\max*}),\ |\sigma^2_{k,jl} - \sigma^2_k| \leq \varepsilon_\sigma\sigma^2_{k*}/2, \text{ for } k = 0, 1, 0n, 1a\Big\} \tag{F.40}$$

where $\boldsymbol{s}$ is the same sequence that characterises $b_{j,l}$, and denote by η the radius of these balls. These balls cover $\mathscr{Q}_j$ as otherwise the set $b_{j,1}, \ldots, b_{j,N_j}$ would not be maximal. Moreover, $N_j \leq N(\eta, \mathscr{Q}_j, \|\Xi^{1/2} \cdot \|_n + | \cdot |)$. For every ball $B_{j,l}$ let us construct a test $\phi_{j,l}$ as: $\phi_{j,l} = \mathbb{1}_{\{dP_{b_l}/dP_{b_*} \geq 1\}}$, where

$$P_{b_l} = \prod_{i \in I_{00}} \mathscr{N}(g^l_{0s_i}, \sigma^2_{0s_i,l}/\xi_i) \prod_{i \in I_{11}} \mathscr{N}(g^l_{1s_i}, \sigma^2_{1s_i,l}/\xi_i) \prod_{i \in I_{10}} \mathscr{N}(g^l_{0n}, \sigma^2_{0n,l}/\xi_i) \prod_{i \in I_{01}} \mathscr{N}(g^l_{1a}, \sigma^2_{1a,l}/\xi_i),$$

$b_l \in B_{j,l}$ and

$$P_{b_*} = \prod_{i \in I_{00}} \mathscr{N}(g^*_{0s_i}, \sigma^2_{0s_i,*}/\xi_i) \prod_{i \in I_{11}} \mathscr{N}(g^*_{1s_i}, \sigma^2_{1s_i,*}/\xi_i) \prod_{i \in I_{10}} \mathscr{N}(g^*_{0n}, \sigma^2_{0n,*}/\xi_i) \prod_{i \in I_{01}} \mathscr{N}(g^*_{1a}, \sigma^2_{1a,*}/\xi_i).$$

Denote $b_* \triangleq (\boldsymbol{f_s^*}, \boldsymbol{\sigma^2_{0s_0*}}, \boldsymbol{\sigma^2_{1s_1*}}, \sigma^2_{0n,*}, \sigma^2_{1a,*})$, and $\mathbf{E}_b$ (resp. $\mathbf{E}_{b_*}$) is the expectation taken with respect to the distribution P_b (resp. P_{b_*}). Since $b_l \in B_{j,l} \subset \mathcal{Q}_j$, this test satisfies:

$$\begin{aligned}
\mathbf{E}_{b_*}\phi_{j,l} &= \int \mathbb{1}_{\{\sqrt{dP_{b_l}/dP_{b_*}}\geq 1\}} dP_{b_*} \leq \int \sqrt{dP_{b_l} dP_{b_*}} \\
&\leq \prod_{k\in\{0,1\}} \prod_{i\in I_{kk}} \left(\frac{\sigma_{ks_i,l}}{2\sigma_{ks_i*}} + \frac{\sigma_{ks_i*}}{2\sigma_{ks_i,l}}\right)^{-1/2} \exp\left\{-\sum_{i\in I_{kk}} \frac{\xi_i\left(g^l_{ks_i}(z_i) - g^*_{ks_i}(z_i)\right)^2}{4(\sigma^2_{ks_i,l} + \sigma^2_{ks_i*})}\right\} \\
&\quad \times \prod_{k\in\{0n,1a\}} \prod_{i\in I_k} \left(\frac{\sigma_{k,l}}{2\sigma_{k*}} + \frac{\sigma_{k*}}{2\sigma_{k,l}}\right)^{-1/2} \exp\left\{-\sum_{i\in I_k} \frac{\xi_i\left(g^l_k(z_i) - g^*_k(z_i)\right)^2}{4(\sigma^2_{k,l} + \sigma^2_{k*})}\right\} \\
&\qquad < \exp\left\{-\left(\frac{n}{\sigma^2_{0\max *}} + \frac{n}{\sigma^2_{1\max *}} + \frac{n}{\sigma^2_{0n*}} + \frac{n}{\sigma^2_{1a*}}\right)\frac{j^2\varepsilon^2}{8}\right\}, \quad \text{(F.41)}
\end{aligned}$$

and

$$\begin{aligned}
\forall b \in B_{j,l}, \qquad \mathbf{E}_b(1-\phi_{j,l}) &\leq \sqrt{\mathbf{E}_{b_l}(1-\phi_{j,l})}\sqrt{\mathbf{E}_{b_l}(dP_b/dP_{b_l})^2} < \exp\left\{-\frac{nj^2\varepsilon^2}{144\sigma^2_{0\max *}}\right\} \\
&\times \exp\left\{-\frac{nj^2\varepsilon^2}{144\sigma^2_{1\max *}}\right\} \exp\left\{-\frac{nj^2\varepsilon^2}{144\sigma^2_{0n*}}\right\} \exp\left\{-\frac{nj^2\varepsilon^2}{144\sigma^2_{1a\max *}}\right\}, \quad \text{(F.42)}
\end{aligned}$$

where $\sigma^2_{0\max *} \triangleq \max\{\sigma^2_{0*}, \sigma^2_{0n*}\}$ and $\sigma^2_{1\max *} \triangleq \max\{\sigma^2_{1*}\sigma^2_{1a*}\}$. To get (F.41) we have used the inequality $|1-\sigma_k/\sigma_{k*}| > j\varepsilon_\sigma$ which implies that $\left(\frac{\sigma_{ks_i,l}}{\sigma_{ks_i*}} + \frac{\sigma_{ks_i*}}{\sigma_{ks_i,l}}\right) > 1 + j\varepsilon_\sigma + (1-j\varepsilon_\sigma)^{-1} > 2$ since $j\varepsilon_\sigma > j\varepsilon_\sigma/(1+j\varepsilon_\sigma)$ and $1 - j\varepsilon_\sigma < 1 + j\varepsilon_\sigma$, and the inequality $\sigma^2_{ks_i,l} + \sigma^2_{ks_i*} = \sigma^2_{ks_i*}(\frac{\sigma^2_{ks_i,l}}{\sigma^2_{ks_i*}} + 1) < \sigma^2_{ks_i*}((1-j\varepsilon_\sigma)^2 + 1) \leq 2\sigma^2_{ks_i*}$. To get (F.42) we have used the Cauchy-Schwartz inequality, the result in Lemma F.5 and the fact that for $b \in B_{j,l}$:

$$\mathbf{E}_{b_l}(1-\phi_{j,l}) = \int \mathbb{1}_{\{\sqrt{dP_{b_*}/dP_{b_l}}>1\}} dP_{b_l} < \int \sqrt{dP_{b_l} dP_{b_*}} < e^{-\left(\frac{n}{\sigma^2_{0\max *}} + \frac{n}{\sigma^2_{1\max *}} + \frac{n}{\sigma^2_{0n*}} + \frac{n}{\sigma^2_{1a*}}\right)\frac{j^2\varepsilon^2}{8}}.$$

This implies that $\sup_{b\in B_{j,l}} \mathbf{E}_b(1-\phi_{j,l}) < e^{-\frac{j^2\varepsilon^2}{144}\left(\frac{n}{\sigma^2_{0\max *}} + \frac{n}{\sigma^2_{1\max *}} + \frac{n}{\sigma^2_{0n*}} + \frac{n}{\sigma^2_{1a*}}\right)}$, and by choosing $\varepsilon = 12\max(\sigma_{0*}, \sigma_{1*}, \sigma_{0n*}, \sigma_{1a*})\sqrt{\widetilde{M}_1}\epsilon_n$ for some constant $\widetilde{M}_1$ it follows that

$$\mathbf{E}_{b_*}\phi_{j,l} \;\leq\; e^{-\widetilde{M}_1 jn\epsilon^2_n} \qquad \text{and} \qquad \sup_{b\in B_{j,l}} \mathbf{E}_b(1-\phi_{j,l}) \leq e^{-\widetilde{M}_1 jn\epsilon^2_n}.$$

Let $\phi_n \triangleq \max_{j\in\mathbb{N},l\in\{1,\dots,N_j\}} \phi_{j,l}$. Then, if $N_j \le \exp(n\epsilon_n^2)$, $\forall j \in \mathbb{N}$

$$\mathbf{E}_{b_*}\phi_n \le \sum_{j=1}^{\infty}\sum_{l=1}^{N_j}\mathbf{E}_{b_*}\phi_{j,l} \le \sum_{j=1}^{\infty}\sum_{l=1}^{N_j} e^{-\widetilde{M}_1 jn\epsilon_n^2} \le \sum_{j=1}^{\infty} N_j e^{-\widetilde{M}_1 jn\epsilon_n^2}$$
$$\le e^{n\epsilon_n^2}\left(\frac{1}{1-e^{-\widetilde{M}_1 n\epsilon_n^2}} - 1\right) = \frac{e^{-(\widetilde{M}_1-1)n\epsilon_n^2}}{1-e^{-\widetilde{M}_1 n\epsilon_n^2}}.$$

Moreover, for a given $\ell \in \mathbb{N}$, $\varepsilon > 0$ and $\varepsilon_\sigma > 0$, let $\mathscr{A}_\ell \triangleq \{(f,\sigma_k^2) \in \mathscr{F}_n \times \left[\frac{1}{2n}, e^{6n\epsilon_n^2}\right]$ for $k = 0, 1, 1a, 0n$; $\|\Xi^{1/2}(\boldsymbol{f_s} - \boldsymbol{f_s^*})\|_n > \ell\varepsilon, |1-\sigma_k/\sigma_{k*}| > \ell\varepsilon_\sigma$ for $k = 0, 1, 0n, 1a\}$ and, since $\mathscr{Q}_j \subseteq \bigcup_{k=1,\dots,N_j} B_{j,k}$, we have for every $\ell \in \mathbb{N}$

$$\sup_{b\in\mathscr{A}_\ell}\mathbf{E}_b(1-\phi_n) \le \sup_{b\in\bigcup_{j>\ell}\mathscr{Q}_j}\mathbf{E}_b(1-\phi_n)$$
$$\le \sup_{j>\ell}\sup_{1\le l\le N_j}\sup_{b\in B_{j,l}}\mathbf{E}_b(1-\phi_{j,l}) \le \sup_{j>\ell}\sup_{1\le l\le N_j} e^{-\widetilde{M}_1 jn\epsilon_n^2} = e^{-\widetilde{M}_1 \ell n\epsilon_n^2}.$$

Finally, remark that $\ell\varepsilon = \ell 12\max(\sigma_{0*},\sigma_{1*},\sigma_{0n*},\sigma_{1a*})\sqrt{\widetilde{M}_1}\epsilon_n =: \ell M\epsilon_n$ and we can write $\widetilde{M}_1 = M^2K$ with $K = 1/(144\max(\sigma_{0*}^2,\sigma_{1*}^2,\sigma_{0n*}^2,\sigma_{1a*}^2))$.

To conclude the proof we have to show that $N_j \lesssim \exp(n\epsilon_n^2)$. Let $\Xi_k \triangleq \mathrm{diag}(\xi_i, i \in I_{kk})$ for $k \in \{0,1\}$ and $\Xi_k \triangleq \mathrm{diag}(\xi_i, i \in I_k)$ for $k \in \{0n, 1a\}$, and remark that $\|\Xi^{1/2}\cdot\|_n \le \|\Xi_0^{1/2}\cdot\|_{n_{00}} + \|\Xi_1^{1/2}\cdot\|_{n_{11}} + \|\Xi_{0n}^{1/2}\cdot\|_{n_{10}} + \|\Xi_{1a}^{1/2}\cdot\|_{n_{01}}$. By this and since for every $j \in \mathbb{N}$, $\mathscr{Q}_j \subseteq \mathscr{C}_{n,0}\times\mathscr{C}_{n,1}\times\mathscr{C}_{n,0n}\times\mathscr{C}_{n,1a}\times\left[\frac{1}{2n}, e^{6n\epsilon_n^2}\right]^4$, then

$$N_j = N(\eta, \mathscr{Q}_j, \|\Xi_0^{1/2}\cdot\|_{n_{00}} + \|\Xi_1^{1/2}\cdot\|_{n_{11}} + \|\Xi_{0n}^{1/2}\cdot\|_{n_{10}} + \|\Xi_{1a}^{1/2}\cdot\|_{n_{01}} + |\cdot|)$$
$$\le N(\eta, \mathscr{C}_{n,0}\times\mathscr{C}_{n,1}\times\mathscr{C}_{n,0n}\times\mathscr{C}_{n,1a}\times\left[\frac{1}{2n}, e^{6n\epsilon_n^2}\right]^4, \|\Xi_0^{1/2}\cdot\|_{n_{00}} + \|\Xi_1^{1/2}\cdot\|_{n_{11}} + \|\Xi_{0n}^{1/2}\cdot\|_{n_{10}} + \|\Xi_{1a}^{1/2}\cdot\|_{n_{01}} + |\cdot|)$$
$$\le \prod_{k=0,1,0n,1a}\left[N\left(\frac{\varepsilon\sigma_{k\min *}}{3\sigma_{k\max *}}, \mathscr{C}_{n,k}, \|\Xi_k^{1/2}\cdot\|_{n_k}\right)\times N\left(\varepsilon_\sigma\sigma_{k*}^2/2, \{\sigma_k^2\in\mathbb{R}_+;\ (2n)^{-1}\le\sigma_k^2 < e^{6n\epsilon_n^2}\}, |\cdot|\right)\right]. \tag{F.43}$$

We start by considering the factor $N\left(\frac{\varepsilon\sigma_{k\min *}}{3\sigma_{k\max *}}, \mathscr{C}_{n,k}, \|\Xi_{n_k}^{1/2}\cdot\|_{n_{kk}}\right)$. Remark that (by using the notation $\boldsymbol{\alpha}_k \triangleq \boldsymbol{\alpha} I[k=0] + \boldsymbol{\beta} I[k=1] + \boldsymbol{\alpha}_n I[k=0n] + \boldsymbol{\beta}_a I[k=1a]$)

$$\begin{aligned}
&\log N\left(\frac{\varepsilon \sigma_{k\min *}}{3\sigma_{k\max *}}, \mathscr{C}_{n,k}, \|\Xi_k^{1/2}\cdot\|_{n_k}\right)\\
&\qquad \leq \log N\Big(\frac{\varepsilon \sigma_{k\min *}}{3\sigma_{k\max *}}, \bigcup_{m_k=1}^{m_k^*}\{\boldsymbol{B}_k(z)'\boldsymbol{\alpha}_k; \boldsymbol{\alpha}_k \in \mathbb{R}^{m_k}, \|\boldsymbol{\alpha}_k\|_\infty \leq M_k\}, \|\Xi_k^{1/2}\cdot\|_{n_k}\Big)\\
&\qquad\qquad \leq \log N\Big(\epsilon_n \underline{\sigma}, \bigcup_{m_k=1}^{m_k^*}\{\boldsymbol{B}_k(z)'\boldsymbol{\alpha}_k;\ \boldsymbol{\alpha}_k \in \mathbb{R}^{m_k}, \|\boldsymbol{\alpha}_k\|_\infty \leq M_k\}, \|\Xi_k^{1/2}\cdot\|_{n_k}\Big)
\end{aligned}$$

where we have used the value of ε and the inequality $\frac{\varepsilon\sigma_{k\min *}}{3\sigma_{k\max *}} \geq 4\epsilon_n\sqrt{\widetilde{M}_1}\underline{\sigma} \geq \epsilon_n\underline{\sigma}$ with $\underline{\sigma} \triangleq \min\{\underline{\sigma}_0, \underline{\sigma}_1, \underline{\sigma}_{0n}, \underline{\sigma}_{1a}\}$. By Van der Vaart [2000, Example 19.7] there exists a constant K_k depending on $\boldsymbol{\alpha}_k$ only such that (by denoting with $\|\cdot\|$ the Euclidean norm) the penultimate inequality below holds:

$$\begin{aligned}
\log N\left(\frac{\varepsilon \sigma_{k\min *}}{3\sigma_{k\max *}}, \mathscr{C}_{n,k}, \|\Xi_k^{1/2}\cdot\|_{n_k}\right) &\leq \log\sum_{m_k=1}^{m_k^*} N\Big(\epsilon_n\underline{\sigma}/\sqrt{m_k^*}, \{\boldsymbol{\alpha}_k \in \mathbb{R}^{m_k}, \|\boldsymbol{\alpha}_k\|_\infty \leq M_k\}, \|\cdot\|\Big)\\
&\leq \log m_k^* N\Big(\epsilon_n\underline{\sigma}/\sqrt{m_k^*}, \{\boldsymbol{\alpha}_k \in \mathbb{R}^{m_k^*}, \|\boldsymbol{\alpha}_k\|_\infty \leq M_k\}, \|\cdot\|\Big)\\
&\leq \log\left(m_k^* K_k\left(\frac{\sqrt{m_k^*}2M_k}{\epsilon_n\underline{\sigma}}\sqrt{m_k^*}\right)^{m_k^*}\right) \leq \log\left(m_k^* K_k\left(\frac{\sqrt{m_k^*}2M_k}{\epsilon_n\underline{\sigma}}\sqrt{m_k^*}\right)^{m_k^*}\right)
\end{aligned}$$

where we have used the fact that $\epsilon_n\underline{\sigma} \geq \epsilon_n\underline{\sigma}/\sqrt{m_k^*}$ for every $m_k^* \geq 1$ to get the first inequality and the fact that the diameter of the hypercube $[-M_k, M_k]^{m_k^*}$ is $\sqrt{m_k^*}2M_k$ to get the penultimate inequality. By replacing the optimal values $m_k^* \asymp \left(\frac{n_k}{\log n_k}\right)^{1/(2\delta+1)}$, $\epsilon_n \asymp \left(\frac{\log n}{n}\right)^{\delta/(2\delta+1)}$, $n_k \asymp n$ and $M_k \asymp n\epsilon_n^2 = n^{1/(2\delta+1)}(\log n)^{2\delta/(2\delta+1)}$ we obtain

$$\sup_{\varepsilon > 3\epsilon_n\underline{\sigma}\sigma_{k\max *}/\sigma_{k\min *}} \log N\left(\frac{\varepsilon \sigma_{k\min *}}{3\sigma_{k\max *}}, \mathscr{C}_{n,k}, \|\Xi_k^{1/2}\cdot\|_{n_k}\right) \lesssim m_k^*\log(n_k)\left(\frac{\delta+1}{2\delta+1}+3\right) \lesssim n_k\epsilon_n^2$$

since $m_k^*\log(n) \asymp n\epsilon_n^2$. We now analyse the second factor in (F.43). Since $\varepsilon_\sigma\sigma_{k*}^2/2 > \underline{\sigma}_k^2/(2n^2)$ we obtain

$$\begin{aligned}
&\log N\left(\varepsilon_\sigma\sigma_{k*}^2/2, \left\{\sigma_k^2 \in \mathbb{R}_+;\ (2n_k)^{-1} \leq \sigma^2 < e^{6n_k\epsilon_{n_k}^2}\right\}, |\cdot|\right)\\
&\qquad\qquad \leq \log N\left(\underline{\sigma}_k^2/(2n^2), \{\sigma_k^2 \in \mathbb{R}_+;\ (2n_k)^{-1} \leq \sigma^2 < e^{6n_k\epsilon_{n_k}^2}\}, |\cdot|\right) \leq 4n_k\epsilon_{n_k}^2
\end{aligned}$$

where the last inequality follows from the definition of ϵ_n and $\log(n)/n \leq \epsilon_n^2 \leq 1$. $\square$

Lemma F.5. *For given $j, l \in \mathbb{N}$ consider the balls $B_{j,l}$ defined in* (F.40) *and $b_l \in B_{j,l}$, and let $P_{b_l} = \prod_{i \in I_{00}} \mathcal{N}(g^l_{0s_i}, \sigma^2_{0s_i,l}/\xi_i) \prod_{i \in I_{11}} \mathcal{N}(g^l_{1s_i}, \sigma^2_{1s_i,l}/\xi_i) \prod_{i \in I_{10}} \mathcal{N}(g^l_{0n}, \sigma^2_{0n,l}/\xi_i) \prod_{i \in I_{01}} \mathcal{N}(g^l_{1a}, \sigma^2_{1a,l}/\xi_i)$. For every $b \in B_{j,l}$ it holds*

$$\sqrt{\mathbf{E}_{b_l}(dP_b/dP_{b_l})^2} \leq \exp\left\{\frac{j^2\varepsilon^2}{18}\left(\frac{n}{\sigma^2_{0\max *}} + \frac{n}{\sigma^2_{1\max *}} + \frac{n}{\sigma^2_{0n*}} + \frac{n}{\sigma^2_{1a*}}\right)\right\}$$

where P_b is equal to P_{b_l} with b_l replaced by b.

Proof. First, remark that simple algebra shows that

$$\begin{aligned}
&\mathbf{E}_{b_l}\left(\frac{dP_b}{dP_{b_l}}\right)^2 = \mathbf{E}_b\left(\frac{dP_b}{dP_{b_l}}\right) \\
&= \prod_{k\in\{0,1\}}\prod_{i\in I_{kk}}\left(\frac{\sigma^2_{ks_i,l}}{\sigma^2_{ks_i}}\right)^{n_{kk}/2}\left(2-\frac{\sigma^2_{ks_i}}{\sigma^2_{ks_i,l}}\right)^{-n_{kk}/2}\exp\left\{\sum_{i\in I_{kk}}\frac{[\xi_i(g^l_{ks_i}(z_i)-g_{ks_i}(z_i))]^2}{2\sigma^2_{ks_i,l}-\sigma^2_{ks_i}}\right\} \\
&\qquad \prod_{k\in\{0n,1a\}}\prod_{i\in I_k}\left(\frac{\sigma^2_{k,l}}{\sigma^2_k}\right)^{n_k/2}\left(2-\frac{\sigma^2_k}{\sigma^2_{k,l}}\right)^{-n_k/2}\exp\left\{\sum_{i\in I_{kk}}\frac{[\xi_i(g^l_k(z_i)-g_k(z_i))]^2}{2\sigma^2_{k,l}-\sigma^2_k}\right\}. \quad \text{(F.44)}
\end{aligned}$$

We bound the three factors in the product in the second line, the other terms can be bounded in the same way by only changing the indices. We start by bounding the first two factors:

$$\begin{aligned}
\left(\frac{\sigma^2_{ks_i,l}}{\sigma^2_{ks_i}}\right)^{n_{kk}/2}\left(2-\frac{\sigma^2_{ks_i}}{\sigma^2_{ks_i,l}}\right)^{-n_{kk}/2} &= \left[\frac{\sigma^2_{ks_i,l}}{\sigma^2_{ks_i}}\frac{\sigma^2_{ks_i,l}}{2\sigma^2_{ks_i,l}-\sigma^2_{ks_i}}\right]^{n_{kk}/2} = \left[\frac{\sigma^2_{ks_i,l}}{\sigma^2_{ks_i}}\frac{1}{2-\sigma^2_{ks_i}/\sigma^2_{ks_i,l}}\right]^{n_{kk}/2} \\
\leq \left[\frac{\sigma^2_{ks_i,l}}{\sigma_{ks_i^2}}\frac{1}{1-\varepsilon_\sigma\sigma^2_{ks_i*}/(2\sigma^2_{ks_i,l})}\right]^{n_{kk}/2} &\leq \left[\frac{(1-\varepsilon_\sigma)}{(1+\varepsilon_\sigma)}\frac{1}{1-\varepsilon_\sigma/2(1+\varepsilon_\sigma)}\right]^{n_{kk}/2} \\
&= \left[\frac{2(1-\varepsilon_\sigma)}{2+\varepsilon_\sigma}\right]^{n_{kk}/2} \leq \left[\frac{2(1-\varepsilon_\sigma)}{2}\right]^{n_{kk}/2} \leq 1 \quad \text{(F.45)}
\end{aligned}$$

where we have used the facts that, since $b \in B_{j,l}$: $\sigma^2_{ks_i}/\sigma^2_{ks_i,l} \leq 1 + \varepsilon_\sigma\sigma^2_{ks_i*}/(2\sigma^2_{ks_i,l})$, and since both $b, b_l \in \mathcal{Q}_j$: $\sigma^2_{ks_i,l}/\sigma^2_{ks_i} = (\sigma^2_{ks_i,l}/\sigma^2_{ks_i*})/(\sigma^2_{ks_i}/\sigma^2_{ks_i*}) \leq (1-\varepsilon_\sigma)/(1+\varepsilon_\sigma)$. By using again the fact that $b \in B_{j,l}$ and $b_l \in \mathcal{Q}_j$ the following inequalities hold: $\sigma^2_{ks_i*} < (1+\varepsilon/2)\sigma^2_{ks_i*} = (1+\varepsilon_\sigma-\varepsilon_\sigma/2)\sigma^2_{ks_i*} \leq (1-\varepsilon_\sigma\sigma^2_{ks_i*}/(2\sigma^2_{ks_i,l}))\sigma^2_{ks_i,l} \leq (2-\frac{\sigma^2_{ks_i}}{\sigma^2_{ks_i,l}})\sigma^2_{ks_i,l} = 2\sigma^2_{ks_i,l} - \sigma^2_{ks_i}$ and then

$(2\sigma^2_{ks_i,l} - \sigma^2_{ks_i})^{-1} < \sigma^{-2}_{ks_i*}$, so that we get for $k = 0, 1$:

$$\exp\left\{\sum_{i\in I_{kk}} \frac{\xi_i(g^l_{ks_i}(z_i) - g_{ks_i}(z_i))^2}{2\sigma^2_{ks_i,l} - \sigma^2_{ks_i}}\right\} \leq \exp\left\{\sum_{i\in I_{kk}} \frac{\xi_i(g^l_{ks_i}(z_i) - g_{ks_i}(z_i))^2}{\sigma^2_{ks_i*}}\right\}$$
$$\leq \exp\left\{\frac{\varepsilon^2}{9\sigma^2_{k\max *}}\right\}$$

since we are on $B_{j,l}$ and where $\sigma^2_{k\max *} \triangleq \max\{\sigma^2_{0*}, \sigma^2_{0n*}\}$ for $k = 0$, and $\sigma^2_{k\max *} \triangleq \max\{\sigma^2_{1*}, \sigma^2_{1a*}\}$ for $k = 1$. This together with (F.45),

$$\exp\left\{\sum_{i\in I_{10}} \frac{\xi_i(g^l_{0n}(z_i) - g_{0n}(z_i))^2}{2\sigma^2_{0n,l} - \sigma^2_{0n}}\right\} \leq \exp\left\{\frac{\varepsilon^2}{9\sigma^2_{0n*}}\right\}$$

and a similar expression for observations in I_{01} allows to conclude that for every $j \in \mathbb{N}$:

$$\mathbf{E}_{b_l}\left(\frac{dP_b}{dP_{b_l}}\right)^2 \leq \exp\left\{\frac{j^2\varepsilon^2}{9}\left(\frac{n}{\sigma^2_{0\max *}} + \frac{n}{\sigma^2_{1\max *}} + \frac{n}{\sigma^2_{0n*}} + \frac{n}{\sigma^2_{1a*}}\right)\right\}.$$

□

G Auxiliary results

Lemma G.1. *Let $U \sim \mathcal{N}_m(U_0, A/\lambda)$, $m \in \mathbb{N}_+$, $\lambda \in \mathbb{R}_+$, $U_0 \in \mathbb{R}^m$, A be a covariance matrix, and let $\rho_{\max}$ denote the maximum eigenvalue of A. Assume that:* (i) *there exist finite $w \in \mathbb{R}^m$ and $U^* \in \mathbb{R}^m$ such that $U^* - U_0 = A^{1/2}w$ where $A^{1/2}$ is such that $A^{1/2}(A^{1/2})' = A$, and* (ii) $0 < \rho_{\max} < \infty$. *Then for all $\varepsilon \in (0, \sqrt{\rho_{\max} m/(4\lambda)})$ it holds:*

$$\pi\left(\boldsymbol{\beta} \in \mathbb{R}^m;\ \|U - U^*\| \leq \varepsilon\right) \gtrsim e^{-\lambda\|w\|\varepsilon/\rho^{1/2}_{\max}(A)} e^{-\|w\|^2\lambda/2} \lambda^{m/2}\varepsilon^m m^{-m/2}.$$

Proof. By making the change of variable $\lambda^{1/2}A^{-1/2}(U - U^*) = z$ and by using assumption *(i)*, the probability $\pi\left(U \in \mathbb{R}^m;\ \|U - U^*\| \leq \varepsilon\right)$ can be bounded from below as follows: for $\widetilde{\varepsilon} \triangleq$

$\sqrt{\lambda}\rho_{\max}^{-1/2}(A)\varepsilon,$

$$
\begin{aligned}
\pi\left(U \in \mathbb{R}^m;\ \|U - U^*\| \leq \varepsilon\right) &= \int_{\{\|U-U^*\|\leq\varepsilon\}} (2\pi)^{-m/2}\left|\frac{A}{\lambda}\right|^{-1/2} e^{-\lambda(U-U_0)'A^{-1}(U-U_0)/2}dU \\
&\geq \int_{\{\|A^{1/2'}z\|\leq\sqrt{\lambda}\varepsilon\}} \frac{1}{(2\pi)^{1/2}} e^{-z'z/2} e^{-\lambda^{1/2}\|w\|\sqrt{z'z}} dz\, e^{-\lambda\|w\|^2/2} \\
&\geq \int_{\{\|z\|\leq\widetilde{\varepsilon}\}} \frac{1}{(2\pi)^{m/2}} e^{-\frac{z'z}{2}} dz\, e^{-\lambda^{1/2}\|w\|\widetilde{\varepsilon}} e^{-\lambda\|w\|^2/2}, \quad \text{(G.1)}
\end{aligned}
$$

where we have used the Cauchy-Schwartz inequality to get the first inequality: $w'z \leq \sqrt{w'wz'z}$, and to get the second inequality we have used the fact that $\{\|A^{1/2}z\| \leq \sqrt{\lambda}\varepsilon\} \supset \{\|z\| \leq \sqrt{\lambda}\rho_{\max}^{-1/2}(A)\varepsilon\}$. Remark that $\widetilde{\varepsilon}$ is well defined under assumption *(ii)* because there exists a constant $c > 0$ such that $\rho_{\max}(A) \geq c$. Next, we have to compute $\int_{\{\|z\|\leq\widetilde{\varepsilon}\}} \frac{1}{(2\pi)^{m/2}} e^{-\frac{z'z}{2}} dz = P(\|z\| \leq \widetilde{\varepsilon})$ where P is taken with respect to a $\mathcal{N}_m(0, I_m)$. Remark that $P(|z_1| \leq \eta) \geq 2\eta\phi(\eta_0)$ for any $\eta \in (0, \eta_0)$, where $\phi(\cdot)$ denotes the Lebesgue density of a $\mathcal{N}(0, 1)$ and $z_1 \sim \mathcal{N}(0, 1)$.

$$
\begin{aligned}
P(\|z\| \leq \widetilde{\varepsilon}) &\geq P\left(\max_{1\leq j\leq m}|z_j| \leq m^{-1/2}\widetilde{\varepsilon}\right) = \left[P(|z_1| \leq m^{-1/2}\widetilde{\varepsilon})\right]^m \\
&\geq \left[2\widetilde{\varepsilon}m^{-1/2}\phi(\widetilde{\varepsilon}_0)\right]^m, \qquad \forall m^{-1/2}\widetilde{\varepsilon} \in (0, \widetilde{\varepsilon}_0) \\
&= \left[2\sqrt{\lambda/\rho_{\max}(A)}\varepsilon m^{-1/2}\phi(1/2)\right]^m, \qquad \text{for } \widetilde{\varepsilon}_0 = \frac{1}{2} \\
&\geq \left[\sqrt{\frac{\lambda}{4\rho_{\max}(A)m}}\varepsilon\right]^m \gtrsim \lambda^{m/2}\varepsilon^m m^{-m/2} \qquad \text{(G.2)}
\end{aligned}
$$

where we have used the fact that $2\phi(1/2) \geq 1/2$ to get the penultimate inequality, and assumption *(ii)* to get the last inequality. By putting together (G.1) and (G.2) we establish the result of the lemma.

□

Lemma G.2. *Let $U \sim \mathcal{N}_m(U_0, A/\lambda)$, $m \in \mathbb{N}_+$, $\lambda \in \mathbb{R}_+$, $U_0 \in \mathbb{R}^m$, A be a covariance matrix. Assume that $0 < \rho_{\max} < \infty$ where $\rho_{\max}$ denote the maximum eigenvalue of A. Then, for every $M \geq 1$:*

$$
\pi(\|U\|_\infty > M) \leq \exp\left\{\log(m^*) - \frac{\lambda(M/2 - \|U_0\|_\infty^2)}{2\gamma}\right\}
$$

where $\gamma \triangleq \max_{1\leq j\leq m^*}|A_{jj}|$ and $0 < \gamma < \infty$.

Proof. We upper bound $\pi(\|U\|_\infty > M)$ in the following way:

$$\begin{aligned}
\pi(\|U\|_\infty > M) &\leq \pi(\max_{1\leq j\leq m} |U_j - (U_0)_j| > M - \|U_0\|_\infty) \\
&\leq \pi\left(\bigcup_{j=1}^{m}\{|U_j - (U_0)_j| > M - \|U_0\|_\infty\}\right) \\
&\leq \sum_{j=1}^{m} \pi\left(|U_j - (U_0)_j| > M - \|U_0\|_\infty\right) \\
&\leq \sum_{j=1}^{m} \exp\left\{-\frac{(M - \|U_0\|_\infty)^2}{2Var(U_j)}\right\}
\end{aligned}$$

where to get the last inequality we have used the well known concentration properties for Gaussian measures. We can upper bound $Var(U_j)$ as follows: $Var(U_j) = A_{jj}/\lambda \leq \max_{1\leq j\leq m}|A_{jj}|/\lambda = \gamma/\lambda$ where γ is as defined in the statement of the Lemma. Remark that such a constant exists under the assumption of the lemma since $\max_{1\leq j\leq m}|A_{jj}| \leq \|A\|_\infty \leq \rho_{\max}^{1/2}(A)$. Therefore, by using the inequalities $(c-b)^2 \geq c^2/2 - b^2$ and $M^2 \geq M$ we obtain

$$\begin{aligned}
\pi(\|U\|_\infty > M) &\leq m\exp\left\{-\frac{\lambda(M - \|U_0\|_\infty)^2}{2\gamma}\right\} \\
&\leq m\exp\left\{-\frac{\lambda(M^2/2 - \|U_0\|_\infty^2)}{2\gamma}\right\} \\
&\leq \exp\left\{\log(m) - \frac{\lambda(M/2 - \|U_0\|_\infty^2)}{2\gamma}\right\}.
\end{aligned}$$

□

References

S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford, 2013.

S. Ghosal and A. van der Vaart. Convergence rates of posterior distributions for non iid observations. *The Annals of Statistics*, 35(1):192–223, 02 2007.

A. W. Van der Vaart. *Asymptotic Statistics*. Lectures on Probability Theory, École d'Été de Probailités de St. Flour XX, 2000.