Supplementary material : Simultaneous off-the-grid learning of mixtures issued from a continuous dictionary

CRISTINA BUTUCEA[1,a], JEAN-FRANÇOIS DELMAS[2,b], ANNE DUTFOY[3,d], CLÉMENT HARDY[2,3,c],

[1]*CREST, ENSAE, IP Paris, France,* [a]*cristina.butucea@ensae.fr*
[2]*CERMICS, École des Ponts, France,* [b]*jean-francois.delmas@enpc.fr,* [c]*clement.hardy@enpc.fr*
[3]*EDF R&D, Palaiseau, France,* [d]*anne.dutfoy@edf.fr*

We give here proofs of the results in Butucea et al. (2023); we keep the same notations and number the equations starting from the last one therein.

A. Auxiliary results in the proof of Theorem 3.1

In the next lemma, we give an upper bound of I_0. Recall the constants C'_N and C_F from Assumption 4.1.

Let $f \in L^2(\nu)$, we define the mapping $v : L^2(\nu) \to L^2(\nu)$ such that for any $z \in \mathcal{Z}$:

$$v(f)(z) = \begin{cases} \operatorname{sign}(f(z)) \dfrac{|f(z)|^{p-1}}{\|f\|_{L^p(\nu)}^{p-1}} & \text{if} \quad \|f\|_{L^p(\nu)} > 0, \\ \nu(\mathcal{Z})^{-\frac{1}{q}} & \text{otherwise,} \end{cases} \tag{48}$$

so that $\|v(f)\|_{L^q(\nu)} = 1$.

Lemma A.1. *Under the assumptions of Theorem 3.1 and with the element $P_1 \in H_T$ from Assumption 4.1 associated to the function $V : \mathcal{Z} \times Q^\star \to \mathbb{R}$ defined by:*

$$V(z, \theta_k^\star) = v(I_{0,k})(z),$$

we get that:

$$I_0 \le C'_N I_2 + (1 - C_F) I_3 + |\hat{\Upsilon}(P_1)|.$$

Proof. We have $\|I_{0,k}\|_{L^p(\nu)} = \|I_{0,k}\|_{L^p(\nu)}^p / \|I_{0,k}\|_{L^p(\nu)}^{p-1}$ and therefore:

$$I_0 := \sum_{k \in S^\star} \|I_{0,k}\|_{L^p(\nu)} = \sum_{k \in S^\star} \int V(z, \theta_k^\star) \left(\left(\sum_{\ell \in \tilde{S}_k(r)} \hat{B}_\ell(z) \right) - B_k^\star(z) \right) \nu(\mathrm{d}z).$$

Let P_1 be an element of L_T from Assumption 4.1 associated to the function V such that properties $(i)-(iv)$ therein hold. By adding and substracting $\sum_{k\in S^\star}\sum_{\ell\in\tilde{S}_k(r)}\left\langle \hat{B}_\ell \phi_T(\hat{\theta}_\ell), P_1\right\rangle_{L_T}$ to I_0 and using the property (ii) satisfied by the element P_1, that is, $\left\langle \phi_T(\theta_k^\star), P_1(z)\right\rangle_T = V(z,\theta_k^\star)$ for all $k\in S^\star$ and ν-almost every $z\in\mathcal{Z}$, we obtain:

$$I_0 = \sum_{k\in S^\star}\sum_{\ell\in\tilde{S}_k(r)}\int \hat{B}_\ell(z)\left(V(z,\theta_k^\star) - \left\langle \phi_T(\hat{\theta}_\ell), P_1(z)\right\rangle_T\right)\nu(\mathrm{d}z) + \hat{\Upsilon}(P_1) - \sum_{\ell\in\tilde{S}(r)^c}\left\langle \hat{B}_\ell\phi_T(\hat{\theta}_\ell), P_1\right\rangle_{L_T}.$$

We deduce, using Hölder's inequality, that:

$$I_0 \le \sum_{k\in S^\star}\sum_{\ell\in\tilde{S}_k(r)}\left\|\hat{B}_\ell\right\|_{L^p(\nu)}\left\|V(\theta_k^\star) - \left\langle \phi_T(\hat{\theta}_\ell), P_1\right\rangle_T\right\|_{L^q(\nu)} + |\hat{\Upsilon}(P_1)| + \sum_{\ell\in\tilde{S}(r)^c}\left\|\hat{B}_\ell\right\|_{L^p(\nu)}\left\|\left\langle \phi_T(\hat{\theta}_\ell), P_1\right\rangle_T\right\|_{L^q(\nu)}.$$

Notice that $\hat{\theta}_\ell \notin \bigcup_{k\in S^\star}\mathcal{B}_T(\theta_k^\star, r)$ for $\ell\in\tilde{S}(r)^c$. Then, by using the properties (ii) and (iii) from Assumption 4.1, we get that Lemma A.1 holds with the constants C_N' and C_F from Assumption 4.1. □

In the next lemma, we give an upper bound of I_1. Recall the constants c_N and c_F from Assumption 4.2. Recall the mapping v defined in (48).

Lemma A.2. *Under the assumptions of Theorem 3.1 and with the element $Q_0\in L_T$ from Assumption 4.2 associated to the function $V:\mathcal{Z}\times\mathcal{Q}^\star\to\mathbb{R}$ defined by:*

$$V(z,\theta_k^\star) = v(I_{1,k})(z),$$

we get that:

$$I_1 \le c_N I_2 + c_F I_3 + |\hat{\Upsilon}(Q_0)|.$$

Proof. We have writing $I_{1,k}(z)$ for $I_{1,k}(r)(z)$:

$$I_1 = \sum_{k\in S^\star}\left\|I_{1,k}\right\|_{L^p(\nu)} = \sum_{k\in S^\star}\int V(z,\theta_k^\star) I_{1,k}(z)\nu(\mathrm{d}z).$$

Let Q_0 be an element of L_T from Assumption 4.2 associated to the function V such that properties $(i)-(iii)$ therein hold. By adding and substracting $\sum_{\ell\in\tilde{S}(r)}\left\langle \hat{B}_\ell\phi_T(\hat{\theta}_\ell), Q_0\right\rangle_{L_T} = \left\langle \hat{B}\Phi_T(\hat{\vartheta}), Q_0\right\rangle_{L_T} - \sum_{\ell\in\tilde{S}(r)^c}\left\langle \hat{B}_\ell\phi_T(\hat{\theta}_\ell), Q_0\right\rangle_{L_T}$ to I_1 and using the triangle inequality, we obtain:

$$I_1 \le \sum_{k\in S^\star}\sum_{\ell\in\tilde{S}_k(r)}\int |\hat{B}_\ell(z)|\left|V(z,\theta_k^\star)\,\mathrm{sign}(\hat{\theta}_\ell-\theta_k^\star)\,\mathfrak{d}_T(\hat{\theta}_\ell,\theta_k^\star) - \left\langle \phi_T(\hat{\theta}_\ell), Q_0(z)\right\rangle_T\right|\nu(\mathrm{d}z) + \sum_{\ell\in\tilde{S}(r)^c}\left|\left\langle \hat{B}_\ell\phi_T(\hat{\theta}_\ell), Q_0\right\rangle_{L_T}\right| + \left|\left\langle \hat{B}\Phi_T(\hat{\vartheta}), Q_0\right\rangle_{L_T}\right|.$$

The property *(i)* of Assumption 4.2 gives that $\left\langle \phi_T(\theta_k^\star), Q_0(z)\right\rangle_T = 0$ for all $k \in S^\star$ and $\nu-$almost every $z \in \mathcal{Z}$. This implies that $\left\langle B^\star \Phi_T(\vartheta^\star), Q_0\right\rangle_{L_T} = 0$. Then, by using the definition of I_2 and I_3 from (42)-(43) and the properties *(i)* and *(ii)* of Assumption 4.2, we obtain:

$$I_1 \leq c_N I_2 + c_F I_3 + \left|\left\langle \hat{B}\Phi_T(\hat{\vartheta}), Q_0\right\rangle_{L_T}\right| = c_N I_2 + c_F I_3 + |\hat{\Upsilon}(Q_0)|,$$

with the constants c_N and c_F from Assumption 4.2. □

At this point, one needs to bound I_2 and I_3. In order to do so, we bound from above and from below the Bregman divergence D_B defined by:

$$D_B = \|\hat{B}\|_{\ell_1, L^p(\nu)} - \|B^\star\|_{\ell_1, L^p(\nu)} - \hat{\Upsilon}(P_0), \tag{49}$$

where $P_0 \in H_T$ is the element given by Assumption 4.1 associated to the function V given by:

$$V(z, \theta_k^\star) = \operatorname{sign}(B_k^\star(z)) \frac{|B_k^\star(z)|^{p-1}}{\|B_k^\star\|_{L_p(\nu)}^{p-1}} \quad \text{for all } k \in S^\star. \tag{50}$$

The next lemma gives a lower bound of the Bregman divergence.

Lemma A.3. *Under the assumptions of Theorem 3.1 and with the constants C_N and C_F of Assumption 4.1, we get that:*

$$D_B \geq C_N I_2 + C_F I_3.$$

Proof. By definition (49) of D_B we have:

$$D_B = \sum_{k \in \hat{S}} \left(\left\|\hat{B}_k\right\|_{L^p(\nu)} - \left\langle \hat{B}_k \phi_T(\hat{\theta}_k), P_0\right\rangle_{L_T} \right) - \sum_{k \in S^\star} \left(\left\|B_k^\star\right\|_{L^p(\nu)} - \left\langle B_k^\star \phi_T(\theta_k^\star), P_0\right\rangle_{L_T} \right).$$

By using the interpolating properties of P_0 from Assumption 4.1 associated to V defined in (50), we have $\sum_{k \in S^\star} \left\|B_k^\star\right\|_{L^p(\nu)} - \left\langle B_k^\star \phi_T(\theta_k^\star), P_0\right\rangle_{L_T} = 0$. Hence, we deduce that:

$$\begin{aligned}
D_B &= \sum_{k \in \hat{S}} \left\|\hat{B}_k\right\|_{L^p(\nu)} - \left\langle \hat{B}_k \phi_T(\hat{\theta}_k), P_0\right\rangle_{L_T} \\
&\geq \sum_{k \in \hat{S}} \left\|\hat{B}_k\right\|_{L^p(\nu)} - \left|\left\langle \hat{B}_k \phi_T(\hat{\theta}_k), P_0\right\rangle_{L_T}\right| \\
&\geq \sum_{k \in \hat{S}} \left\|\hat{B}_k\right\|_{L^p(\nu)} - \left\|\hat{B}_k\right\|_{L^p(\nu)} \left\|\left\langle \phi_T(\hat{\theta}_k), P_0\right\rangle_T\right\|_{L^q(\nu)} \\
&\geq \sum_{\ell \in \tilde{S}(r)} \left\|\hat{B}_\ell\right\|_{L^p(\nu)} \left(1 - \left\|\left\langle \phi_T(\hat{\theta}_\ell), P_0\right\rangle_T\right\|_{L^q(\nu)}\right) + \sum_{k \in \tilde{S}(r)^c} \left\|\hat{B}_k\right\|_{L^p(\nu)} \left(1 - \left\|\left\langle \phi_T(\hat{\theta}_k), P_0\right\rangle_T\right\|_{L^q(\nu)}\right).
\end{aligned}$$

Thanks to properties *(i)* and *(iii)* of Assumption 4.1 and the definitions (42) and (43) of I_2 and I_3, we obtain:

$$D_B \geq \sum_{k \in S^\star} \sum_{\ell \in \tilde{S}_k(r)} C_N \left\|\hat{B}_\ell\right\|_{L^p(\nu)} \mathfrak{d}_T(\hat{\theta}_\ell, \theta_k^\star)^2 + \sum_{k \in \tilde{S}(r)^c} C_F \left\|\hat{B}_k\right\|_{L^p(\nu)} \geq C_N I_2 + C_F I_3,$$

where the constants C_N and C_F are that of Assumption 4.1. □

We now give an upper bound of the Bregman divergence.

Lemma A.4. *Under the assumptions of Theorem 3.1, we have:*

$$\kappa\,\nu(\mathcal{Z})\,D_B \le I_2\left(C'_N M_0 + c_N M_1 + 2^{-1} M_2\right) + I_3\left((2-C_F)M_0 + c_F M_1\right) + |\hat{\Upsilon}(P_1)|M_0 + |\hat{\Upsilon}(Q_0)|M_1 + \kappa\,\nu(\mathcal{Z})\,|\hat{\Upsilon}(P_0)|.$$

Proof. Recall that $Q^\star \subset \Theta_T$. We deduce from (40) that:

$$\kappa(\|\hat{B}\|_{\ell_1, L^p(\nu)} - \|B^\star\|_{\ell_1, L^p(\nu)}) \le \frac{1}{\nu(\mathcal{Z})}\hat{\Upsilon}(W_T) - \frac{1}{2}\hat{R}_T^2 \le \frac{1}{\nu(\mathcal{Z})}\hat{\Upsilon}(W_T).$$

Together with (49), we obtain:

$$\kappa D_B \le \frac{1}{\nu(\mathcal{Z})}|\hat{\Upsilon}(W_T)| + \kappa|\hat{\Upsilon}(P_0)|.$$

Then, use (45) to get the Lemma. □

We recall the events:

$$\mathcal{A}_i = \left\{\frac{1}{\nu(\mathcal{Z})}M_i \le C\,\kappa\right\}, \quad \text{for } i \in \{0,1,2\} \quad \text{and} \quad \mathcal{A} = \mathcal{A}_0 \cap \mathcal{A}_1 \cap \mathcal{A}_2,$$

where:

$$C = \frac{C_F}{2(2 - C_F + c_F)} \wedge \frac{C_N}{2(C'_N + c_N + 2^{-1})}.$$

Lemma A.5. *Under the assumptions of Theorem 3.1, we have on the event $\mathcal{A}$:*

$$C_N I_2 + C_F I_3 \le 2C'\left(|\hat{\Upsilon}(P_1)| + |\hat{\Upsilon}(Q_0)| + |\hat{\Upsilon}(P_0)|\right) \quad \textit{with} \quad C' = C \vee 1.$$

Proof. By combining the upper and lower bounds in Lemmas A.4 and A.3, we deduce that:

$$I_2\left(C_N - \frac{1}{\kappa\,\nu(\mathcal{Z})}\left(C'_N M_0 + c_N M_1 + 2^{-1}M_2\right)\right) + I_3\left(C_F - \frac{1}{\kappa\,\nu(\mathcal{Z})}\left((2-C_F)M_0 + c_F M_1\right)\right) \le \frac{1}{\kappa\,\nu(\mathcal{Z})}|\hat{\Upsilon}(P_1)|M_0 + \frac{1}{\kappa\,\nu(\mathcal{Z})}|\hat{\Upsilon}(Q_0)|M_1 + |\hat{\Upsilon}(P_0)|.$$

We deduce the inequality of the Lemma on the event $\mathcal{A}$. □

B. Proof of Corollary 3.4

This section is dedicated to the proof of Corollary 3.4. We shall apply Theorem 3.1 in the particular case $p = 2$ and $q = 2$. Recall that the measure ν is a sum of n weighted Dirac measures. All the assumptions

of Theorem 3.1 are in force. We shall only give tail bounds for the quantities M_j with $j = 0, 1, 2$ defined in (23).

For $j = 0, 1, 2$ and $\theta \in \Theta_T$, we set $X_j(\theta) = \left\| \left\langle W_T, \phi_T^{[j]}(\theta) \right\rangle_T \right\|_{L^2(\nu)}$. Notice that $M_j = \sup_{\Theta_T} X_j$ and that the process X_j^2 is a χ^2 process.

We first consider $j = 0$. Using (12) and (14), we have that:

$$\|\phi_T(\theta)\|_T^2 = 1 \quad \text{and} \quad \left\|\phi_T^{[1]}(\theta)\right\|_T^2 = \mathcal{K}_T^{[1,1]}(\theta,\theta) = 1.$$

We define two functions f_n and g_n on $\mathbb{R}$ by:

$$f_n(x) = \mathrm{e}^{-x(1-2\sqrt{\frac{n}{x}})} \quad \text{and} \quad g_n(x) = \frac{x^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}\,\mathrm{e}^{-x/2}, \tag{51}$$

where Γ denotes the gamma function. Notice that both functions are decreasing on $[n, +\infty)$.

We set:

$$A = \frac{C^2\,\nu(\mathcal{Z})^2}{\sigma^2 \|a\|_{\ell_\infty} \Delta_T}.$$

Recall Assumption 3.1 on the noise holds. We deduce from Lemma E.2 with $C_1 = C_2 = 1$ and $u = C^2\,\kappa^2\,\nu(\mathcal{Z})^2$, that for $\kappa \geq \sqrt{(n+1)/A}$:

$$\mathbb{P}\left(M_0^2 > C^2\,\kappa^2\,\nu(\mathcal{Z})^2\right) \leq f_n\left(\kappa^2\,A\right) + \frac{4|\Theta_T|_{\mathfrak{d}_T}}{2^{n/2}} g_n\left(\kappa^2 A\right), \tag{52}$$

where $|\Theta_T|_{\mathfrak{d}_T}$ denotes the diameter of the set Θ_T with respect to the metric $\mathfrak{d}_T$.

We consider $j = 1$. We have by (12) and (14) that:

$$\left\|\phi_T^{[1]}(\theta)\right\|_T^2 = 1 \quad \text{and} \quad \left\|\tilde{D}_{1;T}[\phi_T^{[1]}](\theta)\right\|_T^2 = \left\|\phi_T^{[2]}(\theta)\right\|_T^2 = \mathcal{K}_T^{[2,2]}(\theta,\theta).$$

Recall $L_{2,2}$ and $\mathcal{V}_T$ are defined in (17) and (20). Since Assumptions 2.3 and 2.4 hold, we get that for $\theta \in \Theta_T$:

$$\mathcal{K}_T^{[2,2]}(\theta,\theta) \leq L_{2,2} + \mathcal{V}_T \leq 2L_{2,2}.$$

We deduce from Lemma E.2 with $C_1 = 1$, $C_2 = \sqrt{2L_{2,2}}$ and $u = C^2\,\kappa^2\,\nu(\mathcal{Z})^2$, that for $\kappa \geq \sqrt{(n+1)/A}$:

$$\mathbb{P}\left(M_1^2 > C^2\,\kappa^2\,\nu(\mathcal{Z})^2\right) \leq f_n\left(\kappa^2 A\right) + \frac{4\sqrt{2L_{2,2}}\,|\Theta_T|_{\mathfrak{d}_T}}{2^{n/2}} g_n\left(\kappa^2 A\right). \tag{53}$$

We consider $j = 2$. We have by (14) that:

$$\left\|\phi_T^{[2]}(\theta)\right\|_T^2 = \mathcal{K}_T^{[2,2]}(\theta,\theta) \quad \text{and} \quad \left\|\tilde{D}_{1;T}[\phi_T^{[2]}](\theta)\right\|_T^2 = \left\|\phi_T^{[3]}(\theta)\right\|_T^2 = \mathcal{K}_T^{[3,3]}(\theta,\theta).$$

Recall the definition of the function h_∞ from (16) and the constants $L_{2,2}$, L_3, $\mathcal{V}_T$ defined in (17) and (20). Using also Assumption 2.4 so that $\mathcal{V}_T \leq L_{2,2} \wedge L_3$, we get that for all $\theta \in \Theta_T$:

$$\mathcal{K}_T^{[2,2]}(\theta,\theta) \leq L_{2,2} + \mathcal{V}_T \leq 2L_{2,2} \quad \text{and} \quad \mathcal{K}_T^{[3,3]}(\theta,\theta) \leq L_3 + \mathcal{V}_T \leq 2\,L_3.$$

We deduce from Lemma E.2 with $C_1 = \sqrt{2L_{2,2}}$, $C_2 = \sqrt{2L_3}$ and $u = C^2\,\kappa^2\,\nu(\mathcal{Z})^2$, that for

$$\kappa \geq \sqrt{2\,L_{2,2}\,(n+1)/A},$$

we have:

$$\mathbb{P}\left(M_2^2 > C^2\,\kappa^2\,\nu(\mathcal{Z})^2\right) \leq f_n\left(\frac{\kappa^2\,A}{2L_{2,2}}\right) + \frac{4\sqrt{L_3}\,|\Theta_T|_{\mathfrak{d}_T}}{\sqrt{L_{2,2}}\,2^{n/2}}\,g_n\left(\frac{\kappa^2\,A}{2L_{2,2}}\right). \tag{54}$$

Wet set:

$$B = \frac{C_1'^2\,\nu(\mathcal{Z})^2}{\sigma^2\|a\|_{\ell_\infty}\Delta_T} \quad \text{with} \quad C_1' = \sqrt{\frac{C^2}{2L_{2,2}\vee 1}}. \tag{55}$$

We deduce from (52), (53) and (54) that for $\kappa \geq \sqrt{(n+1)/B}$:

$$\sum_{j=0}^{2}\mathbb{P}(M_j > C\,\kappa\,\nu(\mathcal{Z})) \leq 3\left(f_n\left(\kappa^2 B\right) + \frac{C_2'|\Theta_T|_{\mathfrak{d}_T}}{2^{n/2}}\,g_n\left(\kappa^2 B\right)\right), \tag{56}$$

where the constant C_2' is finite positive and defined by:

$$C_2' = 4\left(1 \vee \sqrt{2L_{2,2}} \vee \frac{\sqrt{L_3}}{\sqrt{L_{2,2}}}\right).$$

Recall that the functions f_n and g_n are decreasing on $[n, +\infty)$. We get the following asymptotically-equivalent functions (up to a multiplicative constant) for $f_n(c\,n)$ and $g_n(c\,n)$ and some positive constant c:

$$\begin{aligned} &f_n(c\,n) = \mathrm{e}^{-n(c-2\sqrt{c})} \\ &g_n(c\,n)/2^{\frac{n}{2}} \asymp \mathrm{e}^{-\frac{n}{2}(c-\log(c)-1)+\frac{1}{2}\log(n)} \lesssim \mathrm{e}^{-\frac{n}{2}(c-\log(c)-3/2)}. \end{aligned} \tag{57}$$

Indeed, we use that $\Gamma(n/2) \asymp \mathrm{e}^{\frac{n}{2}\log(\frac{n}{2})-\frac{n}{2}-\frac{1}{2}\log(n)}$. Thus, the constant c determines which of the two terms $f_n(c\,n)$ and $g_n(c\,n)/2^{\frac{n}{2}}$ is dominant.

By solving a second order inequality, we give a lower bound on the tuning parameter κ so that the first right hand term of (56) is bounded by $1/\tau$ for some $\tau > 1$.

Indeed for $\tau > 1$ and $\kappa \geq \sqrt{(n+1)/B}\left(1 + \sqrt{1 + \frac{\log(\tau)}{n}}\right)$ we have:

$$f_n(\kappa^2 B) \leq \frac{1}{\tau}.$$

We also have:

$$g_n(\kappa^2 B) \leq g_n\left(n\left(1 + \sqrt{1 + \frac{\log(\tau)}{n}}\right)^2\right) \leq g_n\,(n)\,\mathrm{e}^{-n/2}/\sqrt{\tau},$$

where we used that g_n is decreasing on $[n, +\infty)$ for the first inequality and that $\log(1+x) \leq x$ for the second. So that, we get:

$$\sum_{j=0}^{2} \mathbb{P}(M_j > C\,\kappa\,\nu(\mathcal{Z})) \leq \frac{3}{\tau} + \frac{3\,C_2'|\Theta_T|_{\mathfrak{d}_T}}{\sqrt{\tau}\,2^{\frac{n}{2}}} g_n(n)\,\mathrm{e}^{-n/2}. \tag{58}$$

Then, by using (57), we deduce an asymptotical equivalence up to a multiplicative constant: $F(n) := g_n(n)\,\mathrm{e}^{-n/2}/2^{n/2} \asymp \mathrm{e}^{-n/2+\log(n)/2}$.

Finally, using the definition of B given in (55), when

$$\kappa \geq C_1 \sigma \sqrt{\frac{\|a\|_{\ell_\infty} \Delta_T\, n}{\nu(\mathcal{Z})^2}} \left(1 + \sqrt{1 + \frac{\log(\tau)}{n}}\right)$$

we get by Theorem 3.1 that the bound (21) stands with probability larger than $1 - C_2\left(\frac{1}{\tau} + \frac{|\Theta_T|_{\mathfrak{d}_T} F(n)}{\sqrt{\tau}}\right)$, where:

$$C_1 = \sqrt{2}/C_1' \quad \text{and} \quad C_2 = 3(1 \vee C_2').$$

This completes the proof of the corollary.

C. Proof of Corollary 3.6

In this section, we prove Corollary 3.6. We shall apply Theorem 3.1 in the particular case $p = 1$ and $q = +\infty$. Recall that the measure ν is a sum of n weighted Dirac measures. All the assumptions of Theorem 3.1 are in force. We shall only give tail bounds for the quantities M_j with $j = 0, 1, 2$ defined by $M_j = \sup_{\Theta_T} X_j$ where $X_j(\theta) = \left\| \left\langle W_T, \phi_T^{[j]}(\theta) \right\rangle_T \right\|_{L^\infty(\nu)}$.

Using Assumption 3.1, we get for any $j = 0, 1, 2$ that:

$$\begin{aligned}\mathbb{P}(M_j > C\,\kappa\,\nu(\mathcal{Z})) &\leq \sum_{z \in \mathcal{Z}} \mathbb{P}\left(\sup_{\Theta_T} \left\langle W_T(z), \phi_T^{[j]}(\theta) \right\rangle_T > C\,\kappa\,\nu(\mathcal{Z})\right) \\ &\leq n\mathbb{P}\left(\sup_{\Theta_T} \left\langle w_T, \phi_T^{[j]}(\theta) \right\rangle_T > C\,\kappa\,\nu(\mathcal{Z})\right).\end{aligned}$$

We use (Butucea et al., 2022, Lemma A.2) that establishes a tail bound for suprema of smooth Gaussian processes and similar arguments as those developed in the proof of (Butucea et al., 2022, Theorem 2.1) to get tail bounds on $\sup_{\Theta_T} \left\langle w_T, \phi_T^{[j]}(\theta) \right\rangle_T$ for $j = 0, 1, 2$. We obtain for any $\tau > 1$ and $\kappa \geq C_3 \sigma \sqrt{\Delta_T \log \tau}/\nu(\mathcal{Z})$ with $C_3 = \frac{2}{C}\left(1 \vee \sqrt{2L_{2,2}}\right)$:

$$\mathbb{P}\left(\sup_{\Theta_T} \left\langle w_T, \phi_T^{[j]}(\theta) \right\rangle_T > C\,\kappa\,\nu(\mathcal{Z})\right) \leq C_4' \left(\frac{|\Theta_T|_{\mathfrak{d}_T}}{\tau\sqrt{\log \tau}} \vee \frac{1}{\tau}\right),$$

where C_4' is a positive constant depending on r and $\mathcal{K}_\infty$ defined in (Butucea et al., 2022, Eq. (84)). We get:

$$\sum_{j=0}^{2} \mathbb{P}(M_j > C\,\kappa\,\nu(\mathcal{Z})) \le 3\,C_4'\,n\left(\frac{|\Theta_T|_{\mathfrak{d}_T}}{\tau\sqrt{\log\tau}} \vee \frac{1}{\tau}\right).$$

Therefore, we obtain by Theorem 3.1 that (21) stands with probability larger than

$$1 - C_4\,n\left(\frac{|\Theta_T|_{\mathfrak{d}_T}}{\tau\sqrt{\log\tau}} \vee \frac{1}{\tau}\right),$$

with $C_4 = 3\,C_4'$ provided the tuning parameter in (3) satisfies $\kappa \ge C_3\sigma\sqrt{\Delta_T\,\log\tau}/\nu(\mathcal{Z})$.

D. Proofs for the construction of certificates

This section is devoted to the proof of Propositions 4.1 and 4.2. We shall first introduce norms that will be useful later in the proof. Then, we shall closely follow the proofs of (Butucea et al., 2022, Propositions 7.4 and 7.5).

Let $p, q \in [1, +\infty]$ such that $p \le q$ and $1/p + 1/q = 1$, let $m, n \in \mathbb{N}$. We define a norm $\|\cdot\|_{*,q}$ on $L^q(\nu, \mathbb{R}^n)$ by:

$$\|f\|_{*,q} = \max_{1\le k\le n} \|f_k\|_{L^q(\nu)}.$$

We shall also define a norm on any matrix $A \in \mathbb{R}^{n\times m}$ by:

$$\|A\|_{\mathrm{op},*,q} = \sup_{\substack{f\in L^q(\nu,\mathbb{R}^m)\\ \|f\|_{*,q}\le 1}} \|Af\|_{*,q}.$$

Recall the definition of the operator norm associated to the ℓ_∞ sup-norm defined for any matrix $A \in \mathbb{R}^{n\times m}$ by:

$$\|A\|_{\mathrm{op},\ell_\infty} = \max_{1\le k\le n} \sum_{1\le \ell\le m} |A_{k,\ell}|.$$

We have the following elementary result.

Lemma D.1. *We have the equality on matrix norms on $\mathbb{R}^{n\times m}$:*

$$\|\cdot\|_{\mathrm{op},\ell_\infty} = \|\cdot\|_{\mathrm{op},*,q}.$$

Proof. Let be $A \in \mathbb{R}^{n\times m}$. We have by definition and the triangle inequality for any $f \in L^q(\nu, \mathbb{R}^m)$:

$$\|Af\|_{*,q} = \max_{1\le k\le n} \left\| \sum_{\ell=1}^{m} A_{k,\ell} f_\ell \right\|_{L^q(\nu)} \le \max_{1\le k\le n} \sum_{\ell=1}^{m} |A_{k,\ell}| \|f_\ell\|_{L^q(\nu)}.$$

Hence for any $f \in L^q(\nu, \mathbb{R}^m)$ such that $\|f\|_{*,q} \le 1$, we have:

$$\|Af\|_{*,q} \le \max_{1\le k\le n} \sum_{1\le \ell\le m} |A_{k,\ell}| = \|A\|_{\mathrm{op},\ell_\infty}.$$

Therefore, we have the bound $\|A\|_{\mathrm{op},*,q} \leq \|A\|_{\mathrm{op},\ell_\infty}$.

Let us show that, in fact, we have an equality between those two norms. We set

$$k^\star = \arg\max_{1\leq k\leq n} \sum_{\ell=1}^{m} |A_{k,\ell}|$$

and we define $f^\star$ so that for almost every $z \in \mathcal{Z}$, $f^\star(z) = \nu(\mathcal{Z})^{-1/q}(\mathrm{sign}(A_{k^\star,1}), \cdots, \mathrm{sign}(A_{k^\star,q}))$. We have $\|f\|_{*,q} = 1$ and $\|Af\|_{*,q} = \|A\|_{\mathrm{op},\ell_\infty}$. Thus, we have $\|A\|_{\mathrm{op},*,q} \geq \|A\|_{\mathrm{op},\ell_\infty}$. Therefore we obtain the equality $\|\cdot\|_{\mathrm{op},\ell_\infty} = \|\cdot\|_{\mathrm{op},*,q}$. □

Since the norm $\|\cdot\|_{\mathrm{op},*,q}$ does not depend on q, we note $\|\cdot\|_{\mathrm{op},*}$ instead of $\|\cdot\|_{\mathrm{op},*,q}$.

Lemma D.2. *Let $x \in \mathbb{R}^m$, $A \in \mathbb{R}^{n\times m}$ and $f \in L^q(\nu, \mathbb{R}^m)$. We have the following inequalities:*

$$\left\|x^\top f\right\|_{L^q(\nu)} \leq \|x\|_{\ell_1}\|f\|_{*,q} \quad \text{and} \quad \|Af\|_{*,q} \leq \|A\|_{\mathrm{op},*}\|f\|_{*,q}.$$

Proof. This is clear since $\left\|x^\top f\right\|_{L^q(\nu)} \leq \sum_{\ell=1}^m |x_\ell|\|f_\ell\|_{L^q(\nu)} \leq \|x\|_{\ell_1}\|f\|_{*,q}$.

□

For a function $f : \mathcal{Z}\times\Theta \to \mathbb{R}$, we note for any $z \in \mathcal{Z}$, $f(z)$ (resp. any $\theta \in \Theta$, $f(\theta)$) the function $f(z,\cdot) : \theta \mapsto f(z,\theta)$ (resp. $f(\cdot,\theta) : z \mapsto f(z,\theta)$). The context in which we shall use this notation will be clear so that there is no confusion.

D.1. Proof of Proposition 4.1(Construction of an interpolating certificate).

Let $T \in \mathbb{N}$ and $s \in \mathbb{N}^*$. Recall Assumptions 2.2 (and thus 2.1 on the regularity of φ_T) and 2.3 on the regularity of the asymptotic kernel $\mathcal{K}_\infty$ are in force. Let $\rho \geq 1$, let $r \in \left(0, 1/\sqrt{2L_{0,2}}\right)$ and $u_\infty \in \left(0, H_\infty^{(2)}(r,\rho)\right)$ such that *(ii)*, *(iii)*, *(iv)* and *(v)* of Proposition 4.1 hold. Recall the definitions (32) and (34) of $\Theta^s_{T,\delta}$ and δ_∞. By assumption $\delta_\infty(u_\infty, s)$ is finite. Let $\vartheta^\star = (\theta_1^\star, \ldots, \theta_s^\star) \in \Theta^s_{T,(2\rho_T\ \delta_\infty(u_\infty,s))\vee(2r)}$. We note $Q^\star = \{\theta_i^\star,\ 1 \leq i \leq s\}$ the set of cardinality s. Let $V : \mathcal{Z}\times Q^\star \to \mathbb{R}$ such that for any $\theta^\star \in Q^\star$, $\left\|V(\theta^\star)\right\|_{L^q(\nu)} = 1$. Let $\alpha, \xi \in L^q(\nu, \mathbb{R}^s)$. We define the function $P_{\alpha,\xi}$ on $\mathcal{Z}$ as:

$$P_{\alpha,\xi}(z) = \sum_{k=1}^{s} \alpha_k(z)\phi_T(\theta_k^\star) + \sum_{k=1}^{s} \xi_k(z)\,\tilde{D}_{1,T}[\phi_T](\theta_k^\star), \tag{59}$$

which belongs to H_T. Recall the definition (13) of the kernel $\mathcal{K}_T$. Using (14), we define the corresponding certificate function on $\mathcal{Z}\times\Theta$ by:

$$\eta_{\alpha,\xi}(z,\theta) = \langle \phi_T(\theta), P_{\alpha,\xi}(z)\rangle_T = \sum_{k=1}^{s} \alpha_k(z)\,\mathcal{K}_T(\theta,\theta_k^\star) + \sum_{k=1}^{s} \xi_k(z)\,\mathcal{K}_T^{[0,1]}(\theta,\theta_k^\star). \tag{60}$$

Notice that the function η is twice continuously differentiable on Θ with respect to its second variable θ due to Assumption 2.1. By Assumption 2.2 on the regularity of φ_T and the positivity of g_T and (14),

we get that for almost every $z \in \mathcal{Z}$ the function $\theta \mapsto \eta_{\alpha,\xi}(z,\theta)$ is of class $\mathcal{C}^3$ on Θ, and that:

$$\tilde{D}_{1;T}\left[\eta_{\alpha,\xi}(z)\right](\theta) = \sum_{k=1}^{s} \alpha_k(z)\,\mathcal{K}_T^{[1,0]}(\theta,\theta_k^\star) + \sum_{k=1}^{s} \xi_k(z)\,\mathcal{K}_T^{[1,1]}(\theta,\theta_k^\star). \tag{61}$$

We give a preliminary technical lemma. Set:

$$\Gamma = \begin{pmatrix} \Gamma^{[0,0]} & \Gamma^{[1,0]\top} \\ \Gamma^{[1,0]} & \Gamma^{[1,1]} \end{pmatrix}, \qquad \text{for } \Gamma^{[i,j]} = \mathcal{K}_T^{[i,j]}(\vartheta^\star). \tag{62}$$

As we have $\mathcal{V}(T) \leq \inf_{\Theta_\infty} g_\infty$, by Lemma 7.3 of Butucea et al. (2022) we have that:

$$\Theta^s_{T,\rho_T\,\delta_\infty(u_\infty,s)} \subseteq \Theta^s_{T,\delta_T(u_T(s),s)} \tag{63}$$

where $u_T(s) = u_\infty + (s-1)\mathcal{V}_1(T)$. Hence we have:

$$(\theta_i^\star, 1 \leq i \leq s) \in \Theta^s_{T,\delta_T(u_T(s),s)}. \tag{64}$$

We deduce from (34), (35), (64) and Lemma D.1 that:

$$\left\|I - \Gamma^{[0,0]}\right\|_{\mathrm{op},*} \leq u_T(s), \quad \left\|I - \Gamma^{[1,1]}\right\|_{\mathrm{op},*} \leq u_T(s), \quad \left\|\Gamma^{[1,0]}\right\|_{\mathrm{op},*} \leq u_T(s) \text{ and } \left\|\Gamma^{[1,0]\top}\right\|_{\mathrm{op},*} \leq u_T(s). \tag{65}$$

We shall write for any $z \in \mathcal{Z}$:

$$\overline{V}(z) = (V(z,\theta_1^\star), \cdots, V(z,\theta_s^\star))^\top. \tag{66}$$

Lemma D.3. *Let be $1 \leq p \leq q \leq +\infty$ such that $1/p + 1/q = 1$. Let $V : \mathcal{Z} \times \mathcal{Q}^\star \to \mathbb{R}$ be a measurable mapping such that for any $\theta^\star \in \mathcal{Q}^\star$, $\left\|V(\cdot,\theta^\star)\right\|_{L^q(\nu)} = 1$. Assume that (65) holds. Assume also that $u_T(s) < 1/2$. Then, there exist $\alpha, \xi \in L^q(\nu, \mathbb{R}^s)$ such that:*

$$\eta_{\alpha,\xi}(z,\theta_k^\star) = V(z,\theta_k^\star) \quad \textit{for} \quad 1 \leq k \leq s, \textit{ for } \nu - \textit{almost every } z, \tag{67}$$

$$\tilde{D}_{1,T}\left[\eta_{\alpha,\xi}(z)\right](\theta_k^\star) = 0 \quad \textit{for} \quad 1 \leq k \leq s, \textit{ for } \nu - \textit{almost every } z, \tag{68}$$

and we have also that:

$$\|\alpha\|_{*,q} \leq \frac{1-u_T(s)}{1-2u_T(s)}, \quad \|\xi\|_{*,q} \leq \frac{u_T(s)}{1-2u_T(s)}, \quad \left\|\alpha - \overline{V}\right\|_{*,q} \leq \frac{u_T(s)}{1-2u_T(s)},$$

and

$$\|\alpha\|_{*,p} \leq \nu(\mathcal{Z})^{1/p-1/q}\frac{1-u_T(s)}{1-2u_T(s)}, \|\xi\|_{*,p} \leq \nu(\mathcal{Z})^{1/p-1/q}\frac{u_T(s)}{1-2u_T(s)},$$

$$\left\|\alpha - \overline{V}\right\|_{*,p} \leq \nu(\mathcal{Z})^{1/p-1/q}\frac{u_T(s)}{1-2u_T(s)}.$$

Remark D.4. The construction of interpolating certificates is different from the one introduced in Golbabaee and Poon (2022) where ν is the counting measure and $q = 2$. Indeed, in Golbabaee and Poon (2022) the mapping ξ is constant and α and ξ solve (67) and $\nabla\left\|\eta_{\alpha,\xi}(\cdot,\theta_k^\star)\right\|^2_{L^2(\nu)} = 0 \quad \text{for} \quad 1 \leq k \leq s$ instead of (68).

Proof. Let $z \in \mathcal{Z}$ such that (67) and (68) are satisfied. By (Butucea et al., 2022, Lemma 10.1), we obtain that:

$$\alpha(z) = \Gamma_{SC}^{-1}\overline{V}(z) \quad \text{and} \quad \xi(z) = -[\Gamma^{[1,1]}]^{-1}\Gamma^{[1,0]}\Gamma_{SC}^{-1}\overline{V}(z).$$

where $\Gamma_{SC} = \Gamma^{[0,0]} - \Gamma^{[1,0]\top}[\Gamma^{[1,1]}]^{-1}\Gamma^{[1,0]}$ and:

$$\|I - \Gamma_{SC}\|_{\mathrm{op},*} = \|I - \Gamma_{SC}\|_{\mathrm{op},\ell_\infty} \leq \frac{u_T(s)}{1-u_T(s)}, \quad \left\|\Gamma_{SC}^{-1}\right\|_{\mathrm{op},*} = \left\|\Gamma_{SC}^{-1}\right\|_{\mathrm{op},\ell_\infty} \leq \frac{1-u_T(s)}{1-2u_T(s)}. \tag{69}$$

We recall that if M is a matrix such that, $\|I - M\|_{\mathrm{op},*} < 1$, then M is non-singular, $M^{-1} = \sum_{i\geq 0}(I-M)^i$ and $\left\|M^{-1}\right\|_{\mathrm{op},*} \leq \left(1 - \|I-M\|_{\mathrm{op},*}\right)^{-1}$. Using (65), (69), the fact that $\left\|\overline{V}\right\|_{*,q} = 1$ and Lemma D.2, we get:

$$\|\alpha\|_{*,q} \leq \left\|\Gamma_{SC}^{-1}\right\|_{\mathrm{op},*}\left\|\overline{V}\right\|_{*,q} \leq \frac{1-u_T(s)}{1-2u_T(s)},$$

$$\|\xi\|_{*,q} \leq \left\|[\Gamma^{[1,1]}]^{-1}\Gamma^{[1,0]}\Gamma_{SC}^{-1}\right\|_{\mathrm{op},*}\left\|\overline{V}\right\|_{*,q} \leq \left\|[\Gamma^{[1,1]}]^{-1}\right\|_{\mathrm{op},*}\left\|\Gamma^{[1,0]}\right\|_{\mathrm{op},*}\left\|\Gamma_{SC}^{-1}\right\|_{\mathrm{op},*} \leq \frac{u_T(s)}{1-2u_T(s)},$$

$$\left\|\alpha - \overline{V}\right\|_{*,q} \leq \left\|(\Gamma_{SC}^{-1} - I)\right\|_{\mathrm{op},*}\left\|\overline{V}\right\|_{*,q} \leq \|\Gamma_{SC} - I\|_{\mathrm{op},*}\left\|\Gamma_{SC}^{-1}\right\|_{\mathrm{op},*} \leq \frac{u_T(s)}{1-2u_T(s)}.$$

Then use that for any $f \in L^q(\nu)$, we have

$$\|f\|_{L^p(\nu)} \leq \nu(\mathcal{Z})^{1/p-1/q}\|f\|_{L^q(\nu)} \tag{70}$$

by Hölder's inequality as $p \leq q$, to obtain the upper bound on the norm $\|\cdot\|_{*,p}$. This finishes the proof. □

We fix $V : \mathcal{Z} \times \mathcal{Q}^\star \to \mathbb{R}$ such that for any $\theta^\star \in \mathcal{Q}^\star$ we have $\left\|V(\theta^\star)\right\|_{L^q(\nu)} = 1$ and we consider $P_{\alpha,\xi}$ and $\eta_{\alpha,\xi}$ with α and ξ characterized by (67) and (68) from Lemma D.3. Let $e_\ell \in \mathbb{R}^s$ be the vector with all the entries equal to zero but the ℓ-th which is equal to 1.

Proof of *(iii)* from Assumption 4.1 with $C_F = \varepsilon_\infty(r/\rho)/10$. Let $\theta \in \Theta_T$ such that $\mathfrak{d}_T(\theta, \mathcal{Q}^\star) > r$ (far region). It is enough to prove that $\left\|\eta_{\alpha,\xi}(\theta)\right\|_{L^q(\nu)} \leq 1 - C_F$. Let $\theta_\ell^\star$ be one of the elements of $\mathcal{Q}^\star$ closest to θ in terms of the metric $\mathfrak{d}_T$. Since $\vartheta^\star \in \Theta^s_{T,2\rho_T\,\delta_\infty(u_\infty,s)}$, we have, by the triangle inequality that for any $k \neq \ell$:

$$2\rho_T\,\delta_\infty(u_\infty, s) < \mathfrak{d}_T(\theta_\ell^\star, \theta_k^\star) \leq \mathfrak{d}_T(\theta_\ell^\star, \theta) + \mathfrak{d}_T(\theta, \theta_k^\star) \leq 2\mathfrak{d}_T(\theta, \theta_k^\star).$$

Hence, we have $\vartheta^\star_{\ell,\theta} \in \Theta^s_{T,\rho_T\,\delta_\infty(u_\infty,s)}$, where $\vartheta^\star_{\ell,\theta}$ denotes the vector $\vartheta^\star$ whose ℓ-th coordinate has been replaced by θ. Then, we obtain from Lemma 7.3 of Butucea et al. (2022) that $\Theta^s_{T,\rho_T\,\delta_\infty(u_\infty,s)} \subseteq \Theta^s_{T,\delta_T(u_T(s),s)}$ and thus:

$$\vartheta^\star_{\ell,\theta} \in \Theta^s_{T,\delta_T(u_T(s),s)}. \tag{71}$$

We denote by $\Gamma_{\ell,\theta}$ (resp. $\Gamma^{[i,j]}_{\ell,\theta}$) the matrix Γ (resp. $\Gamma^{[i,j]}$) in (62) where $\vartheta^\star$ has been replaced by $\vartheta^\star_{\ell,\theta}$. Notice the upper bounds (65) also hold for $\Gamma_{\ell,\theta}$ because of (71). Recall we have that for any

$\theta \in \Theta$, $\mathcal{K}_T(\theta,\theta) = 1$ and $\mathcal{K}_T^{[0,1]}(\theta,\theta) = 0$. Elementary calculations give with $\eta_{\alpha,\xi}$ from Lemma D.3 that:

$$\eta_{\alpha,\xi}(z,\theta) = e_\ell^\top \left(\Gamma_{\ell,\theta}^{[0,0]} - I\right)\alpha(z) + \mathcal{K}_T(\theta,\theta_\ell^\star)\alpha_\ell(z) + e_\ell^\top \Gamma_{\ell,\theta}^{[1,0]\top}\xi(z) + \mathcal{K}_T^{[0,1]}(\theta,\theta_\ell^\star)\xi_\ell(z). \tag{72}$$

By taking the norm $\|\cdot\|_{L^q(\nu)}$ in (72) and using the triangle inequality we get:

$$\begin{aligned}\left\|\eta_{\alpha,\xi}(\theta)\right\|_{L^q(\nu)} \le \left\|\Gamma_{\ell,\theta}^{[0,0]} - I\right\|_{\mathrm{op},*}\|\alpha\|_{*,q} + \|\alpha\|_{*,q}|\mathcal{K}_T(\theta,\theta_\ell^\star)| \\ + \left\|\Gamma_{\ell,\theta}^{[1,0]\top}\right\|_{\mathrm{op},*}\|\xi\|_{*,q} + |\mathcal{K}_T^{[0,1]}(\theta,\theta_\ell^\star)|\|\xi\|_{*,q}.\end{aligned} \tag{73}$$

Since θ belongs to the "far region", we have by definition of $\varepsilon_T(r)$ given in (30) that:

$$|\mathcal{K}_T(\theta,\theta_\ell^\star)| \le 1 - \varepsilon_T(r). \tag{74}$$

The triangle inequality and the definitions (20) of $\mathcal{V}_T$ and (17) of $L_{1,0}$ give:

$$|\mathcal{K}_T^{[0,1]}(\theta,\theta_\ell^\star)| \le L_{1,0} + \mathcal{V}_T. \tag{75}$$

Then, using (65) (which holds for $\Gamma_{\ell,\theta}$ thanks to (71)), we get that:

$$\left\|\eta_{\alpha,\xi}(\theta)\right\|_{L^q(\nu)} \le 1 - \varepsilon_T(r) + \frac{u_T(s)}{1-2u_T(s)}\left(2 + L_{1,0} + \mathcal{V}_T\right).$$

Notice that the function $r \mapsto \varepsilon_\infty(r)$ is increasing. Since $\rho_T \le \rho$, we get by Lemma 7.1 of Butucea et al. (2022) that:

$$\varepsilon_T(r) \ge \varepsilon_\infty(r/\rho_T) - \mathcal{V}_T \ge \varepsilon_\infty(r/\rho) - \mathcal{V}_T. \tag{76}$$

By assumption, we have $u_T(s) \le H_\infty^{(2)}(r,\rho) \le 1/4$. Hence, we have $\frac{1}{1-2u_T(s)} \le 2$. We also have $\mathcal{V}_T \le 1/2$ as $\mathcal{V}_T \le H_\infty^{(1)}(r,\rho)$. Therefore, we get:

$$\left\|\eta_{\alpha,\xi}(\theta)\right\|_{L^q(\nu)} \le 1 - \varepsilon_\infty(r/\rho) + \mathcal{V}_T + u_T(s)\left(5 + 2L_{1,0}\right).$$

The assumption $u_T(s) \le H_\infty^{(2)}(r,\rho)$ gives:

$$u_T(s) \le \frac{8}{10\,(5+2L_{1,0})}\varepsilon_\infty(r/\rho)\cdot \tag{77}$$

The assumption $\mathcal{V}_T \le H_\infty^{(1)}(r,\rho)$ gives $\mathcal{V}_T \le \varepsilon_\infty(r/\rho)/10$. Hence, we have $\left\|\eta_{\alpha,\xi}(\theta)\right\|_{L^q(\nu)} \le 1 - \frac{\varepsilon_\infty(r/\rho)}{10}$. Thus, Property *(iii)* from Assumption 4.1 holds with $C_F = \varepsilon_\infty(r/\rho)/10$.

Proof of *(i)* from Assumption 4.1 with $C_N = \nu_\infty(\rho r)/180$. Let $\theta \in \Theta_T$ such that $\mathfrak{d}_T(\theta, Q^\star) \le r$. Let $\ell \in \{1,\cdots,s\}$ such that $\theta \in \mathcal{B}_T(\theta_\ell^\star, r)$ ("near region"). Thus, it is enough to prove that $\left\|\eta_{\alpha,\xi}(\theta)\right\|_{L^q(\nu)} \le 1 - C_N\,\mathfrak{d}_T(\theta_\ell^\star,\theta)^2$. This will be done by using Lemma E.3 to obtain a quadratic decay on $\eta_{\alpha,\xi}$ from a bound on its second Riemannian derivative.

Recall that the function $\eta_{\alpha,\xi}$ is twice continuously differentiable with respect to its second variable. Differentiating (61) and using that $\mathcal{K}_T^{[2,0]}(\theta,\theta) = -1$ and $\mathcal{K}_T^{[2,1]}(\theta,\theta) = 0$, we deduce that for almost every $z \in \mathcal{Z}$:

$$\tilde{D}_{2;T}[\eta_{\alpha,\xi}(z)](\theta) = e_\ell^\top(I + \Gamma_{\ell,\theta}^{[2,0]})\alpha(z) + \mathcal{K}_T^{[2,0]}(\theta,\theta_\ell^\star)e_\ell^\top\alpha(z) + e_\ell^\top\Gamma_{\ell,\theta}^{[2,1]}\xi(z) + \mathcal{K}_T^{[2,1]}(\theta,\theta_\ell^\star)e_\ell^\top\xi(z). \tag{78}$$

We get:

$$\begin{aligned}\tilde{D}_{2;T}[\eta_{\alpha,\xi}(z)](\theta) - V(z,\theta_\ell^\star)\mathcal{K}_T^{[2,0]}(\theta,\theta_\ell^\star) = e_\ell^\top(I + \Gamma_{\ell,\theta}^{[2,0]})\alpha(z) + \mathcal{K}_T^{[2,0]}(\theta,\theta_\ell^\star)e_\ell^\top(\alpha(z) - \overline{V}(z)) \\ + e_\ell^\top\Gamma_{\ell,\theta}^{[2,1]}\xi(z) + \mathcal{K}_T^{[2,1]}(\theta,\theta_\ell^\star)e_\ell^\top\xi(z). \end{aligned} \tag{79}$$

The triangle inequality and the definition of $\mathcal{V}_T$ give:

$$|\mathcal{K}_T^{[2,0]}(\theta,\theta_\ell^\star)| \le L_{2,0} + \mathcal{V}_T \quad \text{and} \quad |\mathcal{K}_T^{[2,1]}(\theta,\theta_\ell^\star)| \le L_{2,1} + \mathcal{V}_T, \tag{80}$$

where $L_{2,0}$ and $L_{1,2}$ are defined in (17). We deduce from (71), the definition of δ_T in (34) and (35) that:

$$\left\|I + \Gamma_{\ell,\theta}^{[2,0]}\right\|_{\mathrm{op},*} \le u_T(s) \quad \text{and} \quad \left\|\Gamma_{\ell,\theta}^{[2,1]}\right\|_{\mathrm{op},*} \le u_T(s). \tag{81}$$

We deduce from (79) that:

$$\begin{aligned}\left\|\tilde{D}_{2;T}[\eta_{\alpha,\xi}](\theta) - V_\ell(z)\mathcal{K}_T^{[2,0]}(\theta,\theta_\ell^\star)\right\|_{L^q(\nu)} &\le \|\alpha\|_{*,q}\left\|I + \Gamma_{\ell,\theta}^{[2,0]}\right\|_{\mathrm{op},*} + \left\|\alpha - \overline{V}\right\|_{*,q}(L_{2,0} + \mathcal{V}_T) \\ &\quad + \|\xi\|_{*,q}\left(\left\|\Gamma_{\ell,\theta}^{[2,1]}\right\|_{\mathrm{op},*} + L_{2,1} + \mathcal{V}_T\right) \\ &\le \frac{u_T(s)}{1 - 2u_T(s)}(1 + L_{2,0} + L_{2,1} + 2\mathcal{V}_T).\end{aligned}$$

By assumption, we have $u_T(s) \le H_\infty^{(2)}(r,\rho) \le 1/6$. Hence, we have $\frac{1}{1-2u_T(s)} \le 2$. Furthermore, we have by assumption that $\mathcal{V}_T \le H_\infty^{(1)}(r,\rho) \le 1/2$ and $u_T(s) \le H_\infty^{(2)}(r,\rho)$. In particular, we have:

$$u_T(s) \le \frac{8}{9(2L_{2,0} + 2L_{2,1} + 4)}\nu_\infty(\rho r).$$

Therefore, we obtain:

$$\left\|\tilde{D}_{2;T}[\eta_{\alpha,\xi}](\theta) - V(z,\theta_\ell^\star)\mathcal{K}_T^{[2,0]}(\theta,\theta_\ell^\star)\right\|_{L^q(\nu)} \le \frac{8}{9}\nu_\infty(\rho r). \tag{82}$$

We now check that the hypotheses of Lemma E.3-*(ii)* hold in order to obtain a quadratic decay on $\theta \mapsto \left\|\eta_{\alpha,\xi}(\theta)\right\|_{L^q(\nu)}$ from the bound (82). First recall that for almost every $z \in \mathcal{Z}$, $\theta \mapsto \eta_{\alpha,\xi}(z,\theta)$ is twice continuously differentiable and have the interpolation properties (67). By the triangle inequality and since by assumption $\mathcal{V}_T \le L_{2,0}$, we have:

$$\sup_{\Theta_T^2}|\mathcal{K}_T^{[2,0]}| \le L_{2,0} + \mathcal{V}_T \le 2L_{2,0}.$$

Then, Lemma 7.1 of Butucea et al. (2022) ensures that for any θ, θ' in Θ_T such that $\mathfrak{d}_T(\theta, \theta') \le r$ we have:

$$-\mathcal{K}_T^{[2,0]}(\theta, \theta') \ge \nu_\infty(r\rho_T) - \mathcal{V}_T \ge \nu_\infty(\rho r) - \mathcal{V}_T \ge \frac{9}{10}\nu_\infty(\rho r),$$

where we used that the function $r \mapsto \nu_\infty(r)$ is decreasing and $\rho_T \le \rho$ for the second inequality and that $\mathcal{V}_T \le H_\infty^{(1)}(r, \rho) \le \nu_\infty(\rho r)/10$ for the last inequality.

Set $\delta = \frac{8}{9}\nu_\infty(\rho r)$, $\varepsilon = \frac{9}{10}\nu_\infty(\rho r)$, $L = 2L_{2,0}$. As $r < L^{-\frac{1}{2}}$ and $\delta < \varepsilon$, we apply Lemma E.3-*(ii)* and get for $\theta \in \mathcal{B}_T(\theta_\ell^\star, r)$:

$$\left\|\eta_{\alpha,\xi}(\theta)\right\|_{L^q(\nu)} \le 1 - \frac{\nu_\infty(\rho r)}{180}\,\mathfrak{d}_T(\theta, \theta_\ell^\star)^2.$$

Proof of *(ii)* from Assumption 4.1 with $C_N' = (5L_{2,0} + L_{2,1} + 4)/8$. Let $\theta \in \Theta_T$ such that $\mathfrak{d}_T(\theta, Q^\star) \le r$. Let $\ell \in \{1, \cdots, s\}$ such that $\theta \in \mathcal{B}_T(\theta_\ell^\star, r)$ ("near region"). We shall prove that $\left\|\eta_{\alpha,\xi}(\theta) - V(\theta_\ell^\star)\right\|_{L^q(\nu)} \le C_N'\,\mathfrak{d}_T(\theta_\ell^\star, \theta)^2$.

Let us consider the function $f : (z, \theta) \to \eta_{\alpha,\xi}(z, \theta) - V(z, \theta_\ell^\star)$. We will bound $\left\|\tilde{D}_{2;T}[f](\theta)\right\|_{L^q(\nu)}$ on $\mathcal{B}_T(\theta_\ell^\star, r)$ and apply Lemma E.3-*(i)* on f to prove the the inequality of property *(ii)*. Notice that for almost every $z \in \mathcal{Z}$, the map $\theta \mapsto f(z, \theta)$ is twice continuously differentiable. By construction, see (67), we have for almost every $z \in \mathcal{Z}$ that $\tilde{D}_{2;T}[f(z)] = \tilde{D}_{2;T}[\eta_{\alpha,\xi}(z)]$, $f(z, \theta_\ell^\star) = 0$ and $\tilde{D}_{1;T}[f(z)](\theta_\ell^\star) = 0$. We deduce from (78) and the bounds (80) that:

$$\begin{aligned}\left\|\tilde{D}_{2;T}[f](\theta)\right\|_{L^q(\nu)} \le \|\alpha\|_{*,q}\left\|I + \Gamma_{\ell,\theta}^{[2,0]}\right\|_{\mathrm{op},*} &+ \|\alpha\|_{*,q}(L_{2,0} + \mathcal{V}_T) \\ &+ \|\xi\|_{*,q}\left\|\Gamma_{\ell,\theta}^{[2,1]}\right\|_{\mathrm{op},*} + \|\xi\|_{*,q}(L_{2,1} + \mathcal{V}_T).\end{aligned}$$

Using (81), and the bounds on α and ξ from Lemma D.3, we get:

$$\left\|\tilde{D}_{2;T}[f](\theta)\right\|_{L^q(\nu)} \le \frac{1 - u_T(s)}{1 - 2u_T(s)}(L_{2,0} + \mathcal{V}_T + u_T(s)) + \frac{u_T(s)}{1 - 2u_T(s)}(L_{2,1} + \mathcal{V}_T + u_T(s)).$$

Since $u_T(s) \le H_\infty^{(2)}(r, \rho) \le 1/6$ and $\mathcal{V}_T \le H_\infty^{(1)}(r, \rho) \le 1/2$, we get:

$$\left\|\tilde{D}_{2;T}[f](\theta)\right\|_{L^q(\nu)} \le \frac{5}{4}L_{2,0} + \frac{1}{4}L_{2,1} + 1.$$

We get thanks to Lemma E.3-*(i)* on the function f that for any $\theta \in \mathcal{B}_T(\theta_\ell^\star, r)$:

$$\left\|\eta_{\alpha,\xi}(\theta) - V(\theta_\ell^\star)\right\|_{L^q(\nu)} \le \frac{1}{8}\left(5L_{2,0} + L_{1,2} + 4\right)\,\mathfrak{d}_T(\theta, \theta_\ell^\star)^2.$$

Proof of *(iv)* from Assumption 4.1 with $C_B = 2$. Recall the definition of $P_{\alpha,\xi}$ in (59). Elementary calculations give using the definitions of $\Gamma^{[0,0]}$ and $\Gamma^{[1,1]}$ in (62):

$$\left\|P_{\alpha,\xi}\right\|_{L_T}^2 \le 2\left\|\sum_{k=1}^s \alpha_k(z)\phi_T(\theta_k^\star)\right\|_{L_T}^2 + 2\left\|\sum_{k=1}^s \xi_k(z)\,\phi_T^{[1]}(\theta_k^\star)\right\|_{L_T}^2$$

$$
\begin{aligned}
&= 2 \sum_{1\le k,\ell\le s} \mathcal{K}_T(\theta_k^\star,\theta_\ell^\star) \int \alpha_k(z)\alpha_\ell(z)\nu(\mathrm{d}z) + 2 \sum_{1\le k,\ell\le s} \mathcal{K}_T^{[1,1]}(\theta_k^\star,\theta_\ell^\star) \int \xi_k(z)\xi_\ell(z)\nu(\mathrm{d}z) \\
&\le 2\|\alpha\|_{*,q}\|\alpha\|_{*,p} \sum_{1\le k,\ell\le s} |\mathcal{K}_T(\theta_k^\star,\theta_\ell^\star)| + 2\|\xi\|_{*,q}\|\xi\|_{*,p} \sum_{1\le k,\ell\le s} |\mathcal{K}_T^{[1,1]}(\theta_k^\star,\theta_\ell^\star)| \\
&\le 2s\,\|\alpha\|_{*,q}\|\alpha\|_{*,p} \left\|\Gamma^{[0,0]}\right\|_{\mathrm{op},*} + 2s\,\|\xi\|_{*,q}\|\xi\|_{*,p} \left\|\Gamma^{[1,1]}\right\|_{\mathrm{op},*}.
\end{aligned}
$$

Using that $\|I\|_{\mathrm{op},*} = 1$ and (65), we get that:

$$
\left\|\Gamma^{[0,0]}\right\|_{\mathrm{op},*} \le 1 + u_T(s) \quad \text{and} \quad \left\|\Gamma^{[1,1]}\right\|_{\mathrm{op},*} \le 1 + u_T(s).
$$

By assumption we have $u_T(s) \le H_\infty^{(2)}(r,\rho) \le \frac{1}{6}$. Using (70), we deduce that:

$$
\left\|P_{\alpha,\xi}\right\|_{L_T}^2 \le 2(1+u_T(s)) \frac{(1-u_T(s))^2 + u_T(s)^2}{(1-2u_T(s))^2} \nu(\mathcal{Z})^{1/p-1/q} s \le 4\,s\,\nu(\mathcal{Z})^{1/p-1/q}.
$$

This gives:

$$
\left\|P_{\alpha,\xi}\right\|_{L_T} \le 2\sqrt{s}\,\nu(\mathcal{Z})^{1/2p-1/2q}. \tag{83}
$$

We proved that *(i)*-*(iv)* from Assumption 4.1 stand. By assumption we also have that for all $\theta \neq \theta' \in Q^\star : \mathfrak{d}_T(\theta,\theta') > 2r$, therefore Assumption 4.1 holds. This finishes the proof of Proposition 4.1.

D.2. Proof of Proposition 4.2 (Construction of an interpolating derivative certificate)

This section is devoted to the proof of Proposition 4.2. We shall closely follow the proof of (Butucea et al., 2022, Proposition 7.5).

Let $T \in \mathbb{N}$ and $s \in \mathbb{N}^*$. Recall Assumptions 2.2 (and thus 2.1 on the regularity of φ_T) and 2.3 on the regularity of the asymptotic kernel $\mathcal{K}_\infty$ are in force. Let $r > 0$ and $u'_\infty \in (0, 1/6)$ such that *(iii)* and *(iv)* of Proposition 4.2 hold. Recall the definitions (32) and (34) of $\Theta^s_{T,\delta}$ and δ_∞. By assumption $\delta_\infty(u'_\infty, s)$ is finite. Let $\vartheta^\star = (\theta_1^\star,\dots,\theta_s^\star) \in \Theta^s_{T,(2\rho_T\,\delta_\infty(u'_\infty,s))\vee(2r)}$. We note $Q^\star = \{\theta_i^\star,\ 1 \le i \le s\}$ the set of cardinality s.

Let $V : \mathcal{Z} \times Q^\star \to \mathbb{R}$ be such that $\left\|V(\theta^\star)\right\|_{L^q(\nu)} = 1$ for any $\theta^\star \in Q^\star$. Recall the notation $\overline{V}$ defined in (66). Let $\alpha, \xi \in L^q(\nu, \mathbb{R}^s)$. We consider the real-valued function $\eta_{\alpha,\xi}$ defined on $\mathcal{Z} \times \Theta$ by (60).

Recall the definition of $\mathcal{V}_T$ from (20) and define $u'_T(s) = u'_\infty + (s-1)\mathcal{V}_T$. Thanks to (63) and (64), we get that (65) holds with $u_T(s)$ replaced by $u'_T(s)$.

Lemma D.5. *Let be $1 \le p \le q \le +\infty$ such that $1/p + 1/q = 1$. Let $V : \mathcal{Z} \times Q^\star \to \mathbb{R}$ be a measurable mapping such that for any $\theta^\star \in Q^\star$, $\left\|V(\cdot,\theta^\star)\right\|_{L^q(\nu)} = 1$. Assume that we have (65) with $u_T(s)$ replaced by $u'_T(s) < 1/2$. Then, there exist $\alpha, \xi \in L^q(\nu, \mathbb{R}^s)$ such that:*

$$
\eta_{\alpha,\xi}(z,\theta_k^\star) = 0 \quad \textit{for} \quad 1 \le k \le s, \textit{ for } \nu - \textit{almost every } z, \tag{84}
$$

$$
\tilde{D}_{1,T}[\eta_{\alpha,\xi}(z)](\theta_k^\star) = V(z,\theta_k^\star) \quad \textit{for} \quad 1 \le k \le s, \textit{ for } \nu - \textit{almost every } z, \tag{85}
$$

and we also have:

$$\|\alpha\|_{*,q} \leq \frac{u_T'(s)}{1-2u_T'(s)}, \quad \|\xi\|_{*,q} \leq \frac{1-u_T'(s)}{1-2u_T'(s)}, \tag{86}$$

and

$$\|\alpha\|_{*,p} \leq \nu(\mathcal{Z})^{1/p-1/q}\frac{u_T'(s)}{1-2u_T'(s)}, \quad \|\xi\|_{*,p} \leq \nu(\mathcal{Z})^{1/p-1/q}\frac{1-u_T'(s)}{1-2u_T'(s)}. \tag{87}$$

Proof. Let $z \in \mathcal{Z}$ such that (84) and (85) are satisfied. Using the notations from Section D.1, we obtain by (Butucea et al., 2022, Lemma 10.2) that:

$$\alpha(z) = -\Gamma_{SC}^{-1}\Gamma^{[1,0]\top}[\Gamma^{[1,1]}]^{-1}\overline{V}(z) \quad \text{and} \quad \xi(z) = \left(I + [\Gamma^{[1,1]}]^{-1}\Gamma^{[1,0]}\Gamma_{SC}^{-1}\Gamma^{[1,0]\top}\right)[\Gamma^{[1,1]}]^{-1}\overline{V}(z).$$

Using (65), (69) and the fact that $\left\|\overline{V}\right\|_{*,q} = 1$, we readily obtain (86). We then obtain the controls (87) using (70).

□

We fix $V : \mathcal{Z} \times \mathcal{Q}^\star \to \mathbb{R}$ such that for any $\theta^\star \in \mathcal{Q}^\star$ we have $\left\|V(\theta^\star)\right\|_{L^q(\nu)} = 1$ and we consider $P_{\alpha,\xi}$ and $\eta_{\alpha,\xi}$ given by (59) and (60), with α and ξ given by Lemma D.5.

Proof of *(i)* **from Assumption 4.2** with $c_N = (L_{0,2} + L_{2,1} + 7)/8$. We define the function $f : (z,\theta) \mapsto \eta_{\alpha,\xi}(z,\theta) - V(z,\theta_\ell^\star)\,\mathrm{sign}(\theta - \theta_\ell^\star)\mathfrak{d}_T(\theta,\theta_\ell^\star)$ on $\mathcal{Z} \times \Theta$. To prove the Property *(i)*, we will bound $\left\|\tilde{D}_{2;T}[f](\theta)\right\|_{L^q(\nu)}$ on Θ and apply Lemma E.3-*(i)*. Recall $\mathfrak{d}_T(\theta,\theta_\ell^\star) = |G_T(\theta) - G_T(\theta_\ell^\star)|$ with G_T a primitive of $\sqrt{g_T}$, and thus $f(z,\theta) = \eta_{\alpha,\xi}(z,\theta) - V(z,\theta_\ell^\star)(G_T(\theta) - G_T(\theta_\ell^\star))$. We deduce that for ν-almost every $z \in \mathcal{Z}$ the function f is twice continuously differentiable with respect to its second variable on Θ; and elementary calculations give that $\tilde{D}_{2;T}[f(z)](\theta) = \tilde{D}_{2;T}[\eta_{\alpha,\xi}(z)](\theta)$ for any $\theta \in \Theta$ and for ν-almost every $z \in \mathcal{Z}$ as $\tilde{D}_{1;T}[G_T] = 1$ and $\tilde{D}_{2;T}[G_T] = 0$.

Let $\theta \in \Theta_T$ and let $\theta_\ell^\star$ be one of the elements of $\mathcal{Q}^\star$ closest to θ in terms of the metric $\mathfrak{d}_T$. Recall the notations $\Gamma_{\ell,\theta}$ (resp. $\Gamma_{\ell,\theta}^{[i,j]}$) and $\vartheta_{\ell,\theta}^\star$ defined after (71). Since for ν-almost every $z \in \mathcal{Z}$ we have $\tilde{D}_{2;T}[f(z)] = \tilde{D}_{2;T}[\eta_{\alpha,\xi}(z)]$, we deduce from (78) that:

$$\begin{aligned}\left\|\tilde{D}_{2;T}[f](\theta)\right\|_{L^q(\nu)} \leq \left\|I + \Gamma_{\ell,\theta}^{[2,0]}\right\|_{\mathrm{op},*}\|\alpha\|_{*,q} + \|\alpha\|_{*,q}|\mathcal{K}_T^{[2,0]}(\theta,\theta_\ell^\star)| \\ + \|\xi\|_{*,q}\left\|\Gamma_{\ell,\theta}^{[2,1]}\right\|_{\mathrm{op},*} + \|\xi\|_{*,q}|\mathcal{K}_T^{[2,1]}(\theta,\theta_\ell^\star)|.\end{aligned}$$

Notice that (71) holds with $u_T(s)$ replaced by $u_T'(s)$. Using (80) and (81) and the bounds (86) on α and ξ from Lemma D.5, we get:

$$\left\|\tilde{D}_{2;T}[f](\theta)\right\|_{L^q(\nu)} \leq \frac{u_T'(s)}{1-2u_T'(s)}(L_{2,0} + \mathcal{V}_T + u_T'(s)) + \frac{1-u_T'(s)}{1-2u_T'(s)}(L_{2,1} + \mathcal{V}_T + u_T'(s)).$$

By assumption, we have $u_T'(s) \leq 1/6$ and $\mathcal{V}_T \leq 1$. Hence, we obtain:

$$\left\|\tilde{D}_{2;T}[f](\theta)\right\|_{L^q(\nu)} \leq \frac{1}{4}L_{2,0} + \frac{5}{4}L_{2,1} + \frac{7}{4}.$$

Since we have for almost every $z \in \mathcal{Z}$, $f(z, \theta_\ell^\star) = 0$ and $\tilde{D}_{1;T}[f(z)](\theta_\ell^\star) = \tilde{D}_{1;T}[\eta_{\alpha,\xi}(z)](\theta_\ell^\star) - V(z, \theta_\ell^\star) = 0$, using Lemma E.3 *(i)*, we get, with $c_N = (L_{2,0} + 5L_{2,1} + 7)/8$:

$$\left\|\eta_{\alpha,\xi}(\theta) - V(\theta_\ell^\star)\ \mathrm{sign}(\theta - \theta_\ell^\star)\, \mathfrak{d}_T(\theta, \theta_\ell^\star)\right\|_{L^q(\nu)} = \|f(\theta)\|_{L^q(\nu)} \leq c_N\, \mathfrak{d}_T(\theta, \theta_\ell^\star)^2.$$

Proof of *(ii)* **from Assumption 4.2** with $c_F = (5L_{1,0} + 7)/4$. Let $\theta \in \Theta_T$, we shall prove that $\left\|\eta_{\alpha,\xi}(\theta)\right\|_{L^q(\nu)} \leq c_F$. Let $\theta_\ell^\star$ be one of the elements of $Q^\star$ closest to θ in terms of the metric $\mathfrak{d}_T$. We deduce from (73) on the upper bound of $\left\|\eta_{\alpha,\xi}(\theta)\right\|_{L^q(\nu)}$, using (65), the inequality from (15), (75) and the bounds (86) on α and ξ from Lemma D.5 that:

$$\left\|\eta_{\alpha,\xi}(\theta)\right\|_{L^q(\nu)} \leq \frac{u_T'(s)}{1 - 2u_T'(s)}\left(1 + u_T'(s)\right) + \frac{1 - u_T'(s)}{1 - 2u_T'(s)}\left(L_{1,0} + \mathcal{V}_T + u_T'(s)\right).$$

Since $u_T'(s) \leq 1/6$ and $\mathcal{V}_T \leq 1$, we obtain:

$$\left\|\eta_{\alpha,\xi}(\theta)\right\|_{L^q(\nu)} \leq \frac{5}{4}L_{1,0} + \frac{7}{4}.$$

Proof of *(iii)* **from Assumption 4.2** with $c_B = 2$. Using very similar arguments as in the proof of (83) (taking care that the upper bound of the norms $\|\cdot\|_{*,q}$ and $\|\cdot\|_{*,p}$ of α and ξ are given by (86) and (87)) we also get $\left\|P_{\alpha,\xi}\right\|_{L_T} \leq 2\sqrt{s}\,\nu(\mathcal{Z})^{1/2p-1/2q}$.

We proved that *(i)*-*(ii)* from Assumption 4.2 stand for any $\theta \in \Theta_T$. Hence Assumption 4.2 holds for any positive r such that for all $\theta \neq \theta' \in Q^\star : \mathfrak{d}_T(\theta, \theta') > 2r$. This finishes the proof of Proposition 4.2.

E. Auxiliary Lemmas

In this section, we provide the proofs of the intermediate results.

E.1. Proof of Proposition 1.3

We prove the optimization problem (3) is well posed. Denote the objective function of (3) by $F(B, \vartheta)$, that is the penalized risk. Then, we have:

$$\inf_{B \in L^2(\nu, \mathbb{R}^K),\, \vartheta \in \Theta_T^K} F(B, \vartheta) \leq F(0, \vartheta^\star) = \frac{1}{2\nu(\mathcal{Z})}\|Y\|_{L_T}^2.$$

By Minkowski inequality, we have that $\|\cdot\|_{L^p(\nu,\mathbb{R}^K)} \leq \|\cdot\|_{\ell_1, L^p(\nu)}$. Indeed, we have for any $B = (B_1, \ldots, B_K) \in L^2(\nu, \mathbb{R}^K)$, with $B_k \in L^2(\nu)$:

$$\|B\|_{L^p(\nu,\mathbb{R}^K)} := \left\|\left(\sum_{k=1}^K B_k^2\right)^{\frac{1}{2}}\right\|_{L^p(\nu)} \leq \left\|\sum_{k=1}^K |B_k|\right\|_{L^p(\nu)} \leq \|B\|_{\ell_1, L^p(\nu)}.$$

Therefore, the minimization of F over B can be restricted to the centered closed ball $\mathcal{B}_0$ in $L^p(\nu, \mathbb{R}^K)$ of radius $\|Y\|_{L_T}^2/(\kappa 2\nu(\mathcal{Z}))$. We recall that the space $L^p(\nu, \mathbb{R}^K)$ is a reflexive Banach space whose dual is $L^q(\nu, \mathbb{R}^K)$ with $1/p + 1/q = 1$, see (Diestel and Uhl, 1977, Theorem 1 p.98). By Kakutani Theorem, the closed balls of $L^p(\nu, \mathbb{R}^K)$ are therefore compact with respect to the weak topology, see (Brezis, 2011, Theorem 3.17). In particular (3) amounts to minimizing F over the compact set $\mathcal{B}_0 \times \Theta_T^K$.

We show that the objective function F is lower semi-continuous (lsc). Recall that a convex strongly continuous (that is, continuous with respect to the strong topology) real valued function defined on a Banach space is weakly lsc (that is, lsc with respect to the weak topology), see (Brezis, 2011, Corollary 3.9). For any $B \in L^2(\nu, \mathbb{R}^K)$, we have:

$$\begin{aligned} \|B\|_{\ell_1, L^p(\nu)} \leq K^{\frac{1}{q}} \left(\sum_{k=1}^K \|B_k\|_{L^p(\nu)}^p \right)^{\frac{1}{p}} &= K^{\frac{1}{q}} \left(\int \|B(z)\|_{\ell_p}^p \nu(\mathrm{d}z) \right)^{\frac{1}{p}} \\ &\leq K^{\frac{1}{2}} \left(\int \|B(z)\|_{\ell_2}^p \nu(\mathrm{d}z) \right)^{\frac{1}{p}} = K^{\frac{1}{2}} \cdot \|B\|_{L^p(\nu, \mathbb{R}^K)}, \end{aligned} \tag{88}$$

where we used Hölder's inequality. We deduce that the function $B \mapsto \|B\|_{\ell_1, L^p(\nu)}$ is strongly continuous. Since it is also convex, we get it is weakly lsc.

Recall the space $(L_T, \|\cdot\|_{L_T})$ is a Hilbert space, see (Diestel and Uhl, 1977, Section IV). The function $X \mapsto \|Y - X\|_{L_T}$ defined on L_T is weakly lsc as it is strongly continuous and convex. Then, since the function $\vartheta \mapsto \Phi(\vartheta)$ is continuous, we deduce that the function $(B, \vartheta) \mapsto B\Phi(\vartheta)$ is continuous from $L^p(\nu, \mathbb{R}^K) \times \mathbb{R}^K$ to L_T with respect to the product topology of the weak topology on $L^p(\nu, \mathbb{R}^K)$ and the usual topology on $\mathbb{R}^K$. Since the composition of a continuous function by a lsc function is a lsc function, we deduce that the function $(B, \vartheta) \mapsto \|Y - B\Phi(\vartheta)\|_{L_T}$ is lsc (with respect to product topology of the weak topology on $L^p(\nu, \mathbb{R}^K)$ and the usual topology on $\mathbb{R}^K$).

In conclusion, the objective function $(B, \vartheta) \mapsto F(B, \vartheta)$ is lsc (with respect to the product topology of the weak topology on $L^p(\nu, \mathbb{R}^K)$ and the usual topology on $\mathbb{R}^K$). Then, we conclude using that a lsc function on a compact set attains a minimum value, see (Aliprantis and Border, 2006, Theorem 2.43).

E.2. Tail bound for suprema of χ^2 processes

We give a tail bound for suprema of weighted χ^2 processes indexed on an interval $I \subset \mathbb{R}$.

Lemma E.1. *Let $I \subset \mathbb{R}$ be a bounded interval. Assume that $X = (X(\theta), \theta \in I)$ is a real centered Gaussian process with Lipschitz sample paths. Consider the process $Y = \sum_{i=1}^n X_i^2$ where $(X_i, 1 \leq i \leq n)$ are independent copies of X. Then, for an arbitrary $\theta_0 \in I$ and for all $u > n \sup_{\theta \in I} \mathrm{Var}(X(\theta))$, we have:*

$$\mathbb{P}\left(\sup_I Y > u \right) \leq \mathrm{e}^{-\frac{u}{\mathrm{Var}X(\theta_0)}\left(1 - 2\sqrt{\frac{n\mathrm{Var}X(\theta_0)}{u}}\right)} + 4 \int_I \frac{\sqrt{\mathrm{Var}(X'(\theta))}}{2^{n/2}\Gamma(n/2)\sqrt{u}} \left(\frac{u}{\mathrm{Var}(X(\theta))} \right)^{(n+1)/2} \mathrm{e}^{-\frac{u}{2\mathrm{Var}(X(\theta))}} \, \mathrm{d}\theta. \tag{89}$$

Proof. Recall that I is a bounded interval. Hence, the process Y defined on I has Lipschitz sample paths. Then, applying Inequality (122) from Butucea et al. (2022) to the process Y and taking the

expectation, we get, with $M = \sup_I Y$, $a = u > 0$, $b = u + \varepsilon$, $\varepsilon > 0$ and $x_0 = \theta_0$:

$$\int_u^{u+\varepsilon} \mathbb{P}(M \geq t)\,\mathrm{d}t \leq \varepsilon \mathbb{P}(Y(\theta_0) \geq u) + \int_I \mathbb{E}\left[|Y'(\theta)|\mathbf{1}_{\{u<Y(\theta)<u+\varepsilon\}}\right]\,\mathrm{d}\theta. \tag{90}$$

The random variable $Y(\theta_0)$ is a standard χ^2 variable of degree n and therefore we have by (Obozinski, Wainwright and Jordan, 2011, Lemma 11) for $u > n\mathrm{Var}(X(\theta_0))$:

$$\mathbb{P}\left(Y(\theta_0) \geq u\right) \leq \mathrm{e}^{-\frac{u}{\mathrm{Var}(X(\theta_0))}\left(1-2\sqrt{\frac{n\mathrm{Var}(X(\theta_0))}{u}}\right)}. \tag{91}$$

Notice that (91) trivially holds if $\mathrm{Var}(X(\theta_0)) = 0$ as $u > 0$.

We now give a bound of the second term in the right-hand side of (90). Since (X_i', X_i) are independent Gaussian processes for $i = 1, \cdots, n$, we can write for a given $\theta \in I$:

$$X_i'(\theta) = \alpha_\theta X_i(\theta) + \beta_\theta G_i,$$

where $(G_i,\, 1 \leq i \leq n)$ are independent standard Gaussian random variables independent of the variables $(X_i(\theta),\, 1 \leq i \leq n)$ and:

$$\alpha_\theta = \frac{\mathbb{E}[X'(\theta)X(\theta)]}{\mathrm{Var}(X(\theta))} \quad \text{and} \quad \beta_\theta^2 = \mathrm{Var}(X'(\theta)) - \alpha_\theta^2 \mathrm{Var}(X(\theta)),$$

with the convention that $\alpha_\theta = 0$ if $\mathrm{Var}(X(\theta)) = 0$. Since $Y' = 2\sum_{i=1}^n X_i' X_i$ *a.e.*, we get that:

$$\mathbb{E}\left[|Y'(\theta)|\mathbf{1}_{\{u<Y(\theta)<u+\varepsilon\}}\right] \leq 2|\alpha_\theta|\mathbb{E}\left[Y(\theta)\mathbf{1}_{\{u<Y(\theta)<u+\varepsilon\}}\right] + 2|\beta_\theta|\mathbb{E}\left[\left|\sum_{i=1}^n X_i(\theta)G_i\right|\mathbf{1}_{\{u<Y(\theta)<u+\varepsilon\}}\right].$$

Since the variables $(G_i,\, 1 \leq i \leq n)$ and $(X_i(\theta),\, 1 \leq i \leq n)$ are independent, the variable $Z = \sum_{i=1}^n X_i(\theta)G_i$ contidionally to the variables $(X_i(\theta),\, 1 \leq i \leq n)$ is a standard Gaussian random variable of variance $Y(\theta)$. This implies that:

$$\mathbb{E}\left[\left|\sum_{i=1}^n X_i(\theta)G_i\right|\mathbf{1}_{\{u<Y(\theta)<u+\varepsilon\}}\right] = \sqrt{\frac{2}{\pi}}\,\mathbb{E}\left[\sqrt{Y(\theta)}\mathbf{1}_{\{u<Y(\theta)<u+\varepsilon\}}\right].$$

We deduce that:

$$\mathbb{E}\left[|Y'(\theta)|\mathbf{1}_{\{u<Y(\theta)<u+\varepsilon\}}\right] \leq 2\left(|\alpha_\theta|(u+\varepsilon) + \sqrt{\frac{2}{\pi}}\,|\beta_\theta|\sqrt{u+\varepsilon}\right)\mathbb{P}(u < Y(\theta) < u+\varepsilon),$$

The random variable $Y(\theta)$ is distributed as a χ^2 variable and has a density:

$$p_{Y(\theta)}(u) = \frac{u^{n/2-1}}{2^{n/2}\Gamma(n/2)}\left(\frac{1}{\mathrm{Var}(X(\theta))}\right)^{n/2}\mathrm{e}^{-\frac{u}{2\mathrm{Var}(X(\theta))}},$$

where by convention $p_{Y(\theta)}(u)$ is taken equal to 0 if $\mathrm{Var}(X(\theta)) = 0$ and where Γ denotes the gamma function.

Letting ε goes to 0 in (90), using (91), the right continuity of the cdf of M and the monotonicity of the density $p_{Y(\theta)}(u)$ of $Y(\theta)$ on $[n\,\mathrm{Var}X(\theta), +\infty[$, we deduce that for $u > n\,\sup_{\theta\in I}\mathrm{Var}(X(\theta))$:

$$\mathbb{P}(M \geq u) \leq \mathrm{e}^{-\frac{u}{\mathrm{Var}X(\theta_0)}\left(1-2\sqrt{\frac{n\mathrm{Var}X(\theta_0)}{u}}\right)} + 2\int_I \left(|\alpha_\theta| u + \sqrt{\frac{2}{\pi}}\,|\beta_\theta|\sqrt{u}\right) p_{Y(\theta)}(u)\,\mathrm{d}\theta. \tag{92}$$

We now bound the second term of the right-hand side of (92) in two steps. Using that $\beta_\theta^2 \leq \mathrm{Var}(X'(\theta))$, we get that:

$$\sqrt{\frac{2}{\pi}}\,|\beta_\theta|\,\sqrt{u}\,p_{Y(\theta)}(u) \leq \frac{1}{\sqrt{\pi}}\,\frac{\sqrt{\mathrm{Var}(X'(\theta))}}{2^{(n-1)/2}\Gamma(n/2)\sqrt{u}}\left(\frac{u}{\mathrm{Var}(X(\theta))}\right)^{n/2}\mathrm{e}^{-\frac{u}{2\mathrm{Var}(X(\theta))}}. \tag{93}$$

Thanks to the Cauchy-Schwarz inequality, we get $|\alpha_\theta| \leq \sqrt{\mathrm{Var}(X'(\theta))}/\sqrt{\mathrm{Var}(X(\theta))}$. We get that:

$$|\alpha_\theta| u\,p_{Y(\theta)}(u) \leq \frac{\sqrt{\mathrm{Var}(X'(\theta))}}{2^{n/2}\Gamma(n/2)\,\sqrt{u}}\left(\frac{u}{\mathrm{Var}(X(\theta))}\right)^{(n+1)/2}\mathrm{e}^{-\frac{u}{2\mathrm{Var}(X(\theta))}}. \tag{94}$$

Notice that (93) and (94) hold also if $\mathrm{Var}(X(\theta)) = 0$. Using that $\sqrt{\frac{2}{\pi}} + 1 \simeq 1.8 \leq 2$ and that $u \geq \sup_{\theta\in I}\mathrm{Var}(X(\theta))$, we deduce (89) from (92), (93) and (94). □

Recall the functions f_n and g_n defined by (51).

Lemma E.2. *Let $T \in \mathbb{N}$ and $n \in \mathbb{N}^*$ be fixed. Let be $\mathcal{Z} = \{1, \cdots, n\}$. Suppose that Assumptions 2.1 and 2.2 hold. Let h be a function of class C^1 from Θ_T to H_T, with Θ_T a sub-interval of Θ. Assume there exist finite constants C_1 and C_2 such that for all $\theta \in \Theta_T$:*

$$\|h(\theta)\|_T \leq C_1 \quad \textit{and} \quad \left\|\tilde{D}_{1;T}[h](\theta)\right\|_T \leq C_2. \tag{95}$$

Let $(W_T(z), z \in \mathcal{Z})$ be H_T-valued noise processes such that Assumption 3.1 holds. Let $a = (a_1, \cdots, a_n)$ be a sequence of nonnegative real numbers.
Set for any z in the set $\mathcal{Z}$ of cardinality n, $X(z) = (X(z,\theta) = \langle h(\theta), W_T(z)\rangle_T, \theta \in \Theta)$ and $Y = \sum_{z\in\mathcal{Z}} a_z X(z)^2$. Then, we have for $u \geq (n+1)\,\|a\|_{\ell_\infty}\sigma^2\Delta_T C_1^2$:

$$\mathbb{P}\left(\sup_{\theta\in\Theta_T} Y(\theta) > u\right) \leq f_n\left(\frac{u}{\sigma^2\|a\|_{\ell_\infty}\Delta_T C_1^2}\right) + \frac{4\,C_2\,|\Theta_T|_{\mathfrak{d}_T}}{C_1\,2^{n/2}}\,g_n\left(\frac{u}{\sigma^2\|a\|_{\ell_\infty}\Delta_T C_1^2}\right), \tag{96}$$

where $|\Theta_T|_{\mathfrak{d}_T}$ denotes the diameter of the interval Θ_T with respect to the metric $\mathfrak{d}_T$, $\|a\|_{\ell_\infty} = \max_{z\in\mathcal{Z}}|a_z|$ and Γ denotes the classical gamma function.

Proof. First we notice that:

$$\mathbb{P}\left(\sup_{\Theta_T} Y > u\right) \leq \mathbb{P}\left(\sup_{\Theta_T} Z > u/\|a\|_{\ell_\infty}\right), \tag{97}$$

where $Z = \sum_{z\in\mathcal{Z}} X(z)^2$. We shall apply Lemma E.1 to the process Z.

Recall that the Gaussian processes $X(z)$ with $z \in \mathcal{Z}$ are independent with the same distribution as a process denoted $X = (X(\theta), \theta \in \Theta_T)$. The process X has Lipschitz sample paths on Θ_T and $X'(\theta) = \langle \partial_\theta h(\theta), w_T \rangle_T$ for $a.e.\ \theta \in \Theta_T$. By Assumption 3.1, we have for all $\theta \in \Theta_T$ and $z \in \mathcal{Z}$:

$$\mathrm{Var}(X(\theta)) \leq \sigma^2 \Delta_T \|h(\theta)\|_T^2 \leq \sigma^2 \Delta_T C_1^2. \tag{98}$$

We first consider the case where $\Theta_T = [\theta_{\min}, \theta_{\max}]$ is a compact interval with $\theta_{\min} < \theta_{\max}$. Then, according to Lemma E.1, Inequality (89) holds with Y replaced by Z for $u > n\sigma^2 \Delta_T C_1^2$.

Notice that the function $x \mapsto x^{\frac{n+1}{2}} \mathrm{e}^{-x/2}$ is decreasing on $[n+1, +\infty)$ and that the function $x \mapsto \mathrm{e}^{-x\left(1-2\sqrt{\frac{n}{x}}\right)}$ is decreasing on $[n, +\infty)$. Then, plugging (98) in Inequality (89), we obtain for $u > (n+1)\,\sigma^2 \Delta_T C_1^2$:

$$\begin{aligned} \mathbb{P}\left(\sup_{\Theta_T} Z > u\right) &\leq \mathrm{e}^{-\frac{u}{\sigma^2 \Delta_T C_1^2}\left(1-2\sqrt{\frac{n\sigma^2 \Delta_T C_1^2}{u}}\right)} \\ &\quad + \frac{4}{2^{n/2}\Gamma(n/2)\sqrt{u}} \left(\frac{u}{\sigma^2 \Delta_T C_1^2}\right)^{(n+1)/2} \mathrm{e}^{-\frac{u}{2\sigma^2 \Delta_T C_1^2}} \int_{\Theta_T} \sqrt{\mathrm{Var}(X'(\theta))}\, \mathrm{d}\theta. \end{aligned} \tag{99}$$

There exists a geodesic $\gamma : [0,1] \mapsto \Theta_T$ such that $\gamma_0 = \theta_{\min}$, $\gamma_1 = \theta_{\max}$ and $\mathfrak{d}_T(\theta_{\min}, \theta_{\max}) = \int_0^1 |\dot{\gamma}_t| \sqrt{g_T(\gamma_t)}\, \mathrm{d}t$. Hence, a change of variable gives:

$$\int_{\Theta_T} \sqrt{\mathrm{Var}(X'(\theta))}\, \mathrm{d}\theta = \int_0^1 |\dot{\gamma}_t| \sqrt{g_T(\gamma_t) \cdot \frac{\mathrm{Var}(X'(\gamma_t))}{g_T(\gamma_t)}}\, \mathrm{d}t. \tag{100}$$

By Assumption 3.1, we have for all $\theta \in \Theta_T$:

$$\frac{\mathrm{Var}(X'(\theta)))}{g_T(\theta)} \leq \sigma^2 \Delta_T \left\|\tilde{D}_{1;T}[h](\theta)\right\|_T^2 \leq \sigma^2 \Delta_T C_2^2.$$

Using this bound in (100), we get:

$$\int_{\Theta_T} \sqrt{\mathrm{Var}(X'(\theta))}\, \mathrm{d}\theta \leq C_2\, \sigma \sqrt{\Delta_T}\, |\Theta_T|_{\mathfrak{d}_T}, \tag{101}$$

where $|\Theta_T|_{\mathfrak{d}_T}$ is the diameter of the interval Θ_T with respect to the metric $\mathfrak{d}_T$.

Combining (99), (101) and (97), we finally obtain (96) for Θ_T a bounded closed interval. Then, use monotone convergence and the continuity of Z to get (96) for any interval Θ_T. □

E.3. Technical lemma

We consider functions $\eta : \mathcal{Z} \times \Theta \mapsto \mathbb{R}$ and bound the quantities $\|\eta(\theta)\|_{L^q(\nu)}$ on some regions of Θ under some assumptions on the second covariant derivative of η with respect to θ. The following Lemma extends (Poon, Keriven and Peyré, 2021, Lemma 2). The proof is similar, as the latter covers the case where ν is a Dirac measure and $\|\cdot\|_{L^q(\nu)}$ reduces to $|\cdot|$.

Lemma E.3. *Let $q \in [1, +\infty]$. Suppose Assumption 2.2 holds. Consider a function $\eta : \mathcal{Z} \times \Theta$ twice continuously differentiable with respect to its second variable and $\theta_0 \in \Theta_T$.*

(i) Assume that for ν-almost every $z \in \mathcal{Z}$ we have $\eta(z, \theta_0) = 0$ and $\tilde{D}_{1;T}[\eta(z)](\theta_0) = 0$, and that there exist $\delta > 0$ and $r > 0$ such that for any $\theta \in \mathcal{B}_T(\theta_0, r)$ we have:

$$\left\|\tilde{D}_{2;T}[\eta](\theta)\right\|_{L^q(\nu)} \leq \delta. \tag{102}$$

Then, we have $\|\eta(\theta)\|_{L^q(\nu)} < (\delta/2)\, \mathfrak{d}_T(\theta, \theta_0)^2$, for any $\theta \in \mathcal{B}_T(\theta_0, r)$.

(ii) Assume now that for ν-almost every $z \in \mathcal{Z}$, $\eta(z, \theta_0) = V(z)$ and $\tilde{D}_{1;T}[\eta(z)](\theta_0) = 0$ where $V \in L^q(\nu)$ with $\|V\|_{L^q(\nu)} = 1$. Assume there exists a finite positive constant L such that $\sup_{\theta_0, \theta \in \Theta_T} |\mathcal{K}_T^{[0,2]}(\theta_0, \theta)| \leq L$ and there exist $\varepsilon > 0$ and $r \in (0, L^{-\frac{1}{2}})$ such that for any $\theta \in \mathcal{B}_T(\theta_0, r)$, $-\mathcal{K}_T^{[0,2]}(\theta_0, \theta) \geq \varepsilon$. Suppose that for any $\theta \in \mathcal{B}_T(\theta_0, r)$ and $\delta < \varepsilon$:

$$\left\|\tilde{D}_{2;T}[\eta](\theta) - V\mathcal{K}_T^{[0,2]}(\theta_0, \theta)\right\|_{L^q(\nu)} \leq \delta. \tag{103}$$

Then, we have $\|\eta(\theta)\|_{L^q(\nu)} \leq 1 - \frac{(\varepsilon - \delta)}{2}\mathfrak{d}_T(\theta, \theta_0)^2$, for any $\theta \in \mathcal{B}_T(\theta_0, r)$.

References

ALIPRANTIS, C. D. and BORDER, K. C. (2006). *Infinite Dimensional Analysis*, Third ed. Springer, Berlin A hitchhiker's guide.

BREZIS, H. (2011). *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. *Universitext*. Springer, New York.

BUTUCEA, C., DELMAS, J.-F., DUTFOY, A. and HARDY, C. (2022). Off-the-grid learning of sparse mixtures from a continuous dictionary. *arXiv preprint arXiv:2207.00171*.

BUTUCEA, C., DELMAS, J.-F., DUTFOY, A. and HARDY, C. (2023). Simultaneous off-the-grid learning of sparse mixtures from a continuous dictionary.

DIESTEL, J. and UHL, J. J. JR. (1977). *Vector Measures*. *Mathematical Surveys, No. 15*. American Mathematical Society, Providence, R.I. With a foreword by B. J. Pettis.

GOLBABAEE, M. and POON, C. (2022). An off-the-grid approach to multi-compartment magnetic resonance fingerprinting. *Inverse Problems* **38** Paper No. 085002, 31.

OBOZINSKI, G., WAINWRIGHT, M. J. and JORDAN, M. I. (2011). Support union recovery in high-dimensional multivariate regression. *Ann. Statist.* **39** 1–47.

POON, C., KERIVEN, N. and PEYRÉ, G. (2021). The geometry of off-the-grid compressed sensing. *Foundations of Computational Mathematics*.