Nonparametric density estimation for intentionally corrupted functional data

Aurore Delaigle and Alexander Meister

University of Melbourne, Australia and Universität Rostock, Germany

Supplementary Material

S1 Proofs

S1.1 Side results

To prove (3.7), note that since $\varphi_\ell \in L_2([0,1])$, we can write $\varphi_\ell = \sum_{j=1}^\infty \varphi_{\ell,j}\psi_j$. Since $\Gamma_X\varphi_\ell = \lambda_\ell\varphi_\ell$, we deduce that $\sum_{j,k=1}^\infty \varphi_{\ell,j}\,\langle\psi_k,\Gamma_X\psi_j\rangle\,\psi_k \;=\; \lambda_\ell\sum_{k=1}^\infty \varphi_{\ell,k}\,\psi_k$. Multiplying both sides of this equality by ψ_k and taking the integral we obtain (3.7).

To prove (3.8), note that, using Fubini's theorem and integration by parts, we have

$$
\begin{aligned}
\langle \psi_k, \Gamma_X \psi_j \rangle &= \int_0^1 \psi_k(t) \big(\Gamma_X \psi_j\big)(t)\, dt = E\Big\{ \int_0^1 \psi_k(t) X'(t)\, dt \int_0^1 \psi_j(s) X'(s)\, ds \Big\} \\
&= \int_0^1 \psi_k'(t) \int_0^1 \big[E\big\{X(t)X(s)\big\}\big] \psi_j'(s)\, ds\, dt \\
&= \int_0^1 \psi_k'(t) \int_0^1 \Big(\big[E\big\{Y(t)Y(s)\big\}\big] - \sigma^2 \min(s,t)\Big) \psi_j'(s)\, ds\, dt \\
&= \int_0^1 \psi_k'(t) \int_0^1 E\big\{Y(t)Y(s)\big\} \psi_j'(s)\, ds\, dt - \sigma^2 \int_0^1 \psi_k(t) \psi_j(t)\, dt \\
&= \mathcal{M}_{j,k} - \sigma^2 \cdot 1\{j = k\}\,,
\end{aligned}
$$

where we used the fact that $\int_0^1 \psi_k'(t) \int_0^1 \min(s,t) \psi_j'(s)\, ds\, dt = \int_0^1 \psi_k(t)\psi_j(t)\, dt$.

In order to provide a more general/ abstract view of a major step (S1.17) in the proof of Theorem 4, we mention that the supremum of a statistical risk $E_\theta \|\hat\theta - \theta\|^2$ over all $\theta \in \Theta$ is estimated from below by a Bayesian risk with respect to some a-priori distribution Q on the parameter space Θ. Therein Θ is a subset of a separable Hilbert space with the norm $\|\cdot\|$. Moreover impose that the data distribution has the density $f(\theta;\cdot)$ with respect to some dominating σ-finite measure μ on the action space Ω. In order to calculate the smallest Bayesian risk, consider the classical argument that

$$\begin{aligned}
E_Q\, E_\theta\, \|\hat{\theta} - \theta\|^2 \;=\; & \iint \|\hat{\theta}(\omega) - \theta\|^2 f(\theta;\omega) d\mu(\omega) dQ(\theta) \\
= & \iint \big\| \big(\hat{\theta}(\omega) - \tilde{\theta}(\omega)\big) + \big(\tilde{\theta}(\omega) - \theta\big) \big\|^2 f(\theta;\omega) d\mu(\omega) dQ(\theta) \\
\geq & \iint \|\tilde{\theta}(\omega) - \theta\|^2 f(\theta;\omega) d\mu(\omega) dQ(\theta) \\
& + 2 \int \Big\langle \hat{\theta}(\omega) - \tilde{\theta}(\omega), \tilde{\theta}(\omega) \int f(\theta;\omega) dQ(\theta) - \int \theta f(\theta;\omega) dQ(\theta) \Big\rangle d\mu(\omega)\,,
\end{aligned} \tag{S1.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product associated with $\| \cdot \|$ and the integrals inside the inner product may be understood as Bochner integrals. Putting

$$\tilde{\theta}(\omega) := \int \theta f(\theta;\omega) dQ(\theta) \,/\, \int f(\theta;\omega) dQ(\theta)\,,$$

the last term in (S1.1) vanishes so that $\tilde{\theta}$ is the Bayes estimator of θ with respect to Q and $\| \cdot \|^2$. Thus the minimal Bayesian risk (Bayesian risk of $\tilde{\theta}$) equals

$$\begin{aligned}
& E_Q\, E_\theta\, \|\tilde{\theta} - \theta\|^2 \\
&= \iint \Big\| \int \theta' f(\theta';\omega) dQ(\theta') \,/\, \int f(\theta'';\omega) dQ(\theta'') - \theta \Big\|^2 f(\theta;\omega) d\mu(\omega) dQ(\theta) \\
&= \int \Big\| \int \theta' f(\theta';\omega) dQ(\theta') \Big\|^2 \Big\{ \int f(\theta'';\omega) dQ(\theta'') \Big\}^{-2} \int f(\theta;\omega) dQ(\theta) d\mu(\omega) \\
&- 2 \int \Big\langle \int \theta' f(\theta';\omega) dQ(\theta'), \int \theta f(\theta;\omega) dQ(\theta) \Big\rangle \Big\{ \int f(\theta'';\omega) dQ(\theta'') \Big\}^{-1} d\mu(\omega) \\
&+ \int \|\theta\|^2 \underbrace{\int f(\theta;\omega) d\mu(\omega)}_{=1} dQ(\theta)\,,
\end{aligned}$$

so that

$$E_Q E_\theta \|\tilde{\theta} - \theta\|^2 = \int \|\theta\|^2 dQ(\theta) - \int \Big\| \int \theta' f(\theta';\omega) dQ(\theta') \Big\|^2 / \Big\{ \int f(\theta;\omega) dQ(\theta) \Big\} d\mu(\omega).$$

This corresponds to the lower bound on the minimax risk which is applied in (S1.17).

S1.2 Proof of Theorem 1

Since the measure P_V of V_1 is known, we can identify the measure P_Y from the Radon-Nikodym derivative $f_Y = dP_Y/dP_V$. Suppose there exist two measures P_X and $\tilde{P}_X$, each of which is a candidate for the true measure of X_1, and both of which lead to the same measure P_Y of $Y_1 = X_1 + V_1$. Consider the functional characteristic functions ψ_X, $\tilde{\psi}_X$ and ψ_Y, defined by

$$\begin{aligned}
\psi_X(t) &= \int \exp\Big\{ i \int_0^1 t(u)x(u)\, du \Big\}\, dP_X(x)\,, \\
\tilde{\psi}_X(t) &= \int \exp\Big\{ i \int_0^1 t(u)x(u)\, du \Big\}\, d\tilde{P}_X(x)\,, \\
\psi_Y(t) &= \int \exp\Big\{ i \int_0^1 t(u)x(u)\, du \Big\}\, dP_Y(x)\,, \\
\psi_V(t) &= \int \exp\Big\{ i \int_0^1 t(u)x(u)\, du \Big\}\, dP_V(x) \\
&= \exp\Big\{ -\frac{1}{2}\sigma^2 \int_0^1 \int_0^1 t(u) \min(u,u') t(u')\, du\, du' \Big\}\,,
\end{aligned}$$

for any $t \in L_2([0,1])$. It follows from the independence of X_1 and V_1 that $\psi_X(t) \cdot \psi_V(t) = \psi_Y(t) = \tilde{\psi}_X(t) \cdot \psi_V(t)$ for all $t \in L_2([0,1])$. Since ψ_V does not vanish anywhere, the above equality implies that $\psi_X = \tilde{\psi}_X$. Now, for $u \in [0,1]$, we put $t(u) = t_h(u) = h^{-1} \sum_{j=1}^{2^m-1} \tau_j \cdot K\big\{(u - j/2^m)/h\big\}$, where $m > 0$ is integer, the τ_j's are real coefficients, and for a bandwidth parameter $h \in (0, 2^{-m}]$ and a kernel function $K : \mathbb{R} \to \mathbb{R}$, which is non-negative, continuous, supported on the interval $[-1,1]$ and integrates to one.

For any fixed m and τ_j, $j = 1, \ldots, 2^m - 1$, we have $\lim_{h\to 0} \int_0^1 t_h(u)x(u)\,du = \sum_{j=1}^{2^m-1} \tau_j \cdot x(j/2^m)$, for any $x \in C_0([0,1])$. By dominated convergence it follows that $\psi_X(t_h) = \tilde{\psi}_X(t_h)$ tend to the characteristic functions of the random vector

$$X_1^{[m]} = \big(X_1(1/2^m), \ldots, X_1((2^m - 1)/2^m)\big)$$

at $\tau = (\tau_1, \ldots, \tau_{2^m-1})$ under the probability measure P_X and $\tilde{P}_X$, respectively, as $h \downarrow 0$. Since τ can be chosen to be any vector in $\mathbb{R}^{2^m-1}$, the above mentioned characteristic functions are equal. It is well known that the characteristic function of any random vector in $\mathbb{R}^{2^m-1}$ determines its distribution uniquely so that the distributions of $X_1^{[m]}$ under the basic measure P_X, on the one hand, and $\tilde{P}_X$, on the other hand, are identical. Thus, for some arbitrary $s \in C_0([0,1])$, we have

$$P_X\big(\{x \in C_{0,0}([0,1]) \,:\, x(j/2^m) \le s(j/2^m),\, \forall j = 1,\ldots,2^m-1\}\big)$$
$$= \tilde{P}_X\big(\{x \in C_{0,0}([0,1]) \,:\, x(j/2^m) \le s(j/2^m),\, \forall j = 1,\ldots,2^m-1\}\big)\,. \tag{S1.2}$$

The countable set $\mathcal{Q} = \bigcup_{m\in\mathbb{N}}\{k/2^m : k = 1,\ldots,2^m-1\}$ is dense in the interval $[0,1]$. Hence the following events coincide:

$$\{x \in C_{0,0}([0,1]) : x(u) \le s(u),\, \forall u \in [0,1]\}$$
$$= \{x \in C_{0,0}([0,1]) : x(u) \le s(u),\, \forall u \in \mathcal{Q}\}$$
$$= \bigcap_{m\in\mathbb{N}} \{x \in C_{0,0}([0,1]) \,:\, x(j/2^m) \le s(j/2^m),\, \forall j = 1,\ldots,2^m-1\}.$$

Therefore we obtain that

$$P_X\big(\{x \in C_{0,0}([0,1]) \,:\, x(u) \le s(u),\, \forall u \in [0,1]\}\big)$$
$$= \lim_{m\to\infty} P_X\big(\{x \in C_{0,0}([0,1]) \,:\, x(j/2^m) \le s(j/2^m),\, \forall j = 1,\ldots,2^m-1\}\big)\,. \tag{S1.3}$$

The corresponding equality holds true for the measure $\tilde{P}_X$.

Combining (S1.2) and (S1.3) we deduce that

$$P_X\big(\{x \in C_{0,0}([0,1]) \,:\, x(u) \le s(u),\, \forall u \in [0,1]\}\big)$$
$$= \tilde{P}_X\big(\{x \in C_{0,0}([0,1]) \,:\, x(u) \le s(u),\, \forall u \in [0,1]\}\big)\,,$$

for any $s \in C_0([0,1])$. The system of the sets

$$\{x \in C_{0,0}([0,1]) \,:\, x(u) \le s(u),\, \forall u \in [0,1]\}, \qquad s \in C_0([0,1])\,,$$

is stable with respect to intersection and generates the Borel σ-field $\mathfrak{B}(C_{0,0}([0,1]))$. Therefore, by the uniqueness theorem for measures, we conclude that $P_X = \tilde{P}_X$.

S1.3 Proof of Proposition 1

Let x and $\tilde{x}$ be two realizations of the functional random variable X. Thanks to Assumptions 1 and 2, we may impose that $x(0) = \tilde{x}(0) = 0$ and that $\max\{\|x'\|_2, \|\tilde{x}'\|_2\} \leq C_{X,1}$. For any $t_1, \ldots, t_n \in [0,1]$ we introduce the vector $F = \big(x(t_j) - \tilde{x}(t_j)\big)^T_{j=1,\ldots,n}$ and the matrix $M = \big\{EW(t_j)W(t_k)\big\}_{j,k=1,\ldots,n}$. According to Proposition 7 in Hall et al. (2013), in order to prove privacy it suffices to show that $\big|M^{-1/2}F\big| \leq \sigma\alpha/c(\beta)$, where we may put $c(\beta) = \sqrt{2\log(2/\beta)}$ according to Proposition 3 in Hall et al. (2013). Without any loss of generality we assume that $t_1 \leq \cdots \leq t_n$ since $\big|(PMP^T)^{-1/2}(PF)\big|^2 = F^T M^{-1} F = \big|M^{-1/2}F\big|^2$, for any $n \times n$-permutation matrix P. Then,

$$M = \begin{pmatrix} t_1 & t_1 & t_1 & \cdots & t_1 \\ t_1 & t_2 & t_2 & \cdots & t_2 \\ t_1 & t_2 & t_3 & \cdots & t_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_1 & t_2 & t_3 & \cdots & t_n \end{pmatrix}.$$

Writing $\Delta_j = (F_j - F_{j-1})/(t_j - t_{j-1})$ if $t_j > t_{j-1}$ and $x'(t_j) - \tilde{x}'(t_j)$ if

$t_j = t_{j-1}$; and $Y_j = \Delta_j - \Delta_{j+1}$, where we set $F_0 = t_0 = 0$ and $\Delta_{n+1} = 0$, we consider that

$$\sum_{l=1}^{k-1} t_l Y_l + \sum_{l=k}^{n} t_k Y_l \;=\; \sum_{l=1}^{k-1} t_l(\Delta_l - \Delta_{l+1}) + t_k \Delta_k \;=\; t_1 \Delta_1 + \sum_{l=2}^{k} (t_l + t_{l-1})\Delta_l \;=\; F_k \,,$$

for all integer $k = 1, \ldots, n$ so that $MY = F$, where $Y = (Y_j)^T_{j=1,\ldots,n}$. We deduce that the left hand side of the above system of equations equals

$$F^T M^{-1} F = F^T Y = \sum_{j=1}^{n} F_j \Delta_j - \sum_{j=1}^{n} F_j \Delta_{j+1} = \sum_{j=1}^{n} \frac{(F_j - F_{j-1})^2}{t_j - t_{j-1}} \,. \quad \text{(S1.4)}$$

As $F_j - F_{j-1} \;=\; \int_{t_{j-1}}^{t_j} \big\{x'(t) - \tilde{x}'(t)\big\} dt$ for $j = 1, \ldots, n$, the Cauchy-Schwarz inequality in $L_2([0,1])$ yields that (S1.4) has the upper bound

$$\sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} \big|x'(t) - \tilde{x}'(t)\big|^2 dt \;\leq\; 4C_{X,1}^2 \,,$$

which completes the proof of the proposition. □

Proof of Lemma 1: (a) Expanding X_1' in the orthonormal basis $\{\varphi_j\}_j$ we get

$$\int_0^1 X_1'(t)\, dV_1(t) \;=\; \sum_{j=1}^{\infty} \langle X_1', \varphi_j \rangle \cdot \beta'_{V_1,j} \,, \quad \|X_1'\|_2^2 \;=\; \sum_{j=1}^{\infty} \big|\langle X_1', \varphi_j \rangle\big|^2 \,,$$

where the infinite sums should be understood as mean squared limits.

Since, for any integer m, $\mathfrak{A}_m$ is a subset of the σ-field generated by V_1, we have that

$$
\begin{aligned}
E\big\{f_Y(V_1)\mid \mathfrak{A}_m\big\} \;&=\; E\Big\{\exp\Big(\frac{1}{\sigma^2}\int_0^1 X_1'(t)\,dV_1(t)-\frac{1}{2\sigma^2}\int_0^1|X_1'(t)|^2\,dt\Big)\Big|\mathfrak{A}_m\Big\}\\
&=\; E\Big\{\exp\Big(\frac{1}{\sigma^2}\sum_{j=1}^{\infty}\langle X_1',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big)\cdot\exp\big(-\|X_1'\|_2^2/(2\sigma^2)\big)\Big|\mathfrak{A}_m\Big\}\\
&=\; E\Big\{\exp\Big(\frac{1}{\sigma^2}\sum_{j=1}^{m}\langle X_1',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big)\cdot\exp\big(-\|X_1'\|_2^2/(2\sigma^2)\big)\\
&\qquad\cdot E\Big\{\exp\Big(\frac{1}{\sigma^2}\sum_{j>m}\langle X_1',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big)\Big|\mathfrak{A}_m,X_1'\Big\}\Big|\mathfrak{A}_m\Big\}\\
&=\; E\Big\{\exp\Big(\frac{1}{\sigma^2}\sum_{j=1}^{m}\langle X_1',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big)\cdot\exp\big(-\|X_1'\|_2^2/(2\sigma^2)\big)\\
&\qquad\cdot\prod_{j>m}E\Big\{\exp\Big(\frac{1}{\sigma^2}\langle X_1',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big)\Big|X_1'\Big\}\Big|\mathfrak{A}_m\Big\}
\end{aligned}
\tag{S1.5}
$$

holds true almost surely.

Applying, to the last term in (S1.5), the fact that

$$
E\big\{\exp\big(t\delta\big)\big\} \;=\; \exp\big(t^2/2\big)\,, \tag{S1.6}
$$

for all $\delta \sim N(0,1)$ and $t \in \mathbb{R}$, we deduce that

$$E\big\{f_Y(V_1) \mid \mathfrak{A}_m\big\} \;=\; E\Big\{\exp\Big(\frac{1}{\sigma^2}\sum_{j=1}^m \langle X_1', \varphi_j\rangle \cdot \beta_{V_1,j}'\Big) \cdot \exp\big(-\|X_1'\|_2^2/(2\sigma^2)\big) \cdot \exp\Big(\frac{1}{2\sigma^2}\sum_{j>m} \big|\langle X_1', \varphi_j\rangle\big|^2\Big)\Big|\mathfrak{A}_m\Big\}$$

$$= E\Big\{\exp\Big(\frac{1}{\sigma^2}\sum_{j=1}^m \langle X_1', \varphi_j\rangle \cdot \beta_{V_1,j}'\Big) \cdot \exp\Big(-\frac{1}{2\sigma^2}\sum_{j=1}^m \big|\langle X_1', \varphi_j\rangle\big|^2\Big)\Big|\mathfrak{A}_m\Big\}$$

$$= \int \exp\Big(\frac{1}{\sigma^2}\sum_{j=1}^m \langle x', \varphi_j\rangle \cdot \beta_{V_1,j}' - \frac{1}{2\sigma^2}\sum_{j=1}^m \big|\langle x', \varphi_j\rangle\big|^2\Big)\, dP_X(x)$$

$$= f_Y^{[m]}(\beta_{V_1,1}', \ldots, \beta_{V_1,m}')$$

almost surely, which completes the proof of part (a).

(b) Using the result of part (a) we have

$$E\big|f_Y^{[m]}(\beta_{V_1,1}', \ldots, \beta_{V_1,m}') - f_Y(V_1)\big|^2$$

$$= E\left[E\big\{\big|f_Y^{[m]}(\beta_{V_1,1}', \ldots, \beta_{V_1,m}') - f_Y(V_1)\big|^2 \mid \mathfrak{A}_m\big\}\right]$$

$$= E\left[\operatorname{var}\big\{f_Y(V_1) \mid \mathfrak{A}_m\big\}\right]. \tag{S1.7}$$

Using Fubini's theorem, we get

$$
\begin{aligned}
&E\left[\mathrm{var}\big\{f_Y(V_1)\mid \mathfrak{A}_m\big\}\right]\\
&\quad= E\left\{\mathrm{var}\Big(\int \exp\Big(\frac{1}{\sigma^2}\sum_{j=1}^{\infty}\langle x',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big)\cdot\exp\big(-\|x'\|_2^2/(2\sigma^2)\big)\,dP_X(x)\Big|\mathfrak{A}_m\Big)\right\}\\
&\quad= \iint \exp\big(-\{\|x_1'\|_2^2+\|x_2'\|_2^2\}/(2\sigma^2)\big)\,E\Big\{\exp\Big(\frac{1}{\sigma^2}\sum_{j=1}^{m}\langle x_1'+x_2',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big)\Big\}\\
&\quad\cdot\mathrm{cov}\Big\{\exp\Big(\frac{1}{\sigma^2}\sum_{j>m}\langle x_1',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big),\exp\Big(\frac{1}{\sigma^2}\sum_{j>m}\langle x_2',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big)\Big\}\\
&\qquad\qquad\qquad dP_X(x_1)\,dP_X(x_2)\,. \qquad (S1.8)
\end{aligned}
$$

Using (S1.6) again we deduce that

$$
E\Big\{\exp\big(\sigma^{-2}\sum_{j=1}^{m}\langle x_1'+x_2',\varphi_j\rangle\cdot\beta'_{V_1,j}\big)\Big\}=\exp\Big\{\sum_{j=1}^{m}\big|\langle x_1'+x_2',\varphi_j\rangle\big|^2/(2\sigma^2)\Big\}\,,
$$

and that

$$
\begin{aligned}
&\mathrm{cov}\Big\{\exp\Big(\frac{1}{\sigma^2}\sum_{j>m}\langle x_1',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big),\exp\Big(\frac{1}{\sigma^2}\sum_{j>m}\langle x_2',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big)\Big\}\\
&= E\Big\{\exp\Big(\frac{1}{\sigma^2}\sum_{j>m}\langle x_1'+x_2',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big)\Big\}\\
&-\Big[E\Big\{\exp\Big(\frac{1}{\sigma^2}\sum_{j>m}\langle x_1',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big)\Big\}\Big]\cdot\Big[E\Big\{\exp\Big(\frac{1}{\sigma^2}\sum_{j>m}\langle x_2',\varphi_j\rangle\cdot\beta'_{V_1,j}\Big)\Big\}\Big]\\
&=\exp\Big(\frac{1}{2\sigma^2}\sum_{j>m}\big|\langle x_1'+x_2',\varphi_j\rangle\big|^2\Big)-\exp\Big(\frac{1}{2\sigma^2}\sum_{j>m}\big\{\big|\langle x_1',\varphi_j\rangle\big|^2+\big|\langle x_2',\varphi_j\rangle\big|^2\big\}\Big)\,.
\end{aligned}
$$

Plugging these equalities into (S1.8) we conclude that

$$E\left[\mathrm{var}\big\{f_Y(V_1)\mid \mathfrak{A}_m\big\}\right] = \iint \Big[\exp\big\{-(\|x_1'\|_2^2+\|x_2'\|_2^2-\|x_1'+x_2'\|_2^2)/(2\sigma^2)\big\}$$

$$-\exp\Big\{-\frac{1}{2\sigma^2}\sum_{j=1}^m\big(\big|\langle x_1',\varphi_j\rangle\big|^2+\big|\langle x_2',\varphi_j\rangle\big|^2-\big|\langle x_1'+x_2',\varphi_j\rangle\big|^2\big)\Big\}\Big]dP_X(x_1)dP_X(x_2)$$

$$= \iint\Big\{\exp\big(\langle x_1',x_2'\rangle/\sigma^2\big)-\exp\Big(\frac{1}{\sigma^2}\sum_{j=1}^m\langle x_1',\varphi_j\rangle\langle x_2',\varphi_j\rangle\Big)\Big\}\,dP_X(x_1)\,dP_X(x_2)\,. \tag{S1.9}$$

Let X_2 denote an independent copy of X_1. Then (S1.9) satisfies

$$E\Big\{\exp\big(\langle X_1',X_2'\rangle/\sigma^2\big)-\exp\Big(\frac{1}{\sigma^2}\sum_{j=1}^m\langle X_1',\varphi_j\rangle\langle X_2',\varphi_j\rangle\Big)\Big\}$$

$$\leq \frac{1}{\sigma^2}E\Big|\sum_{j>m}\langle X_1',\varphi_j\rangle\langle X_2',\varphi_j\rangle\Big|\cdot\exp\big(C_{X,1}^2/\sigma^2\big)$$

$$\leq \frac{1}{\sigma^2}\cdot\exp\big(C_{X,1}^2/\sigma^2\big)\cdot\Big(\sum_{j,j'>m}\big|\langle\varphi_j,\Gamma_X\varphi_{j'}\rangle\big|^2\Big)^{1/2},$$

where we used the mean value theorem and the Cauchy-Schwarz inequality.

S1.4 Proof of Theorem 2

Let $V \sim P_V$ denote a functional random variable which is independent of $X_1,\ldots,X_n$ and $W_1,\ldots,W_n$, and let $\beta_{V,j}' = \int_0^1 \varphi_j(t)\,dV(t)$. Since $\hat f_Y^{[m,K]}(V)$ is measurable in the σ-field generated by $\beta_{V,1}',\ldots,\beta_{V,m}',Y_1,\ldots,Y_n$ and as

$$f_Y^{[m]}(\beta_{V,1}',\ldots,\beta_{V,m}') = E\big\{f_Y(V)\mid \beta_{V,1}',\ldots,\beta_{V,m}'\big\}$$

$$= E\big\{f_Y(V)\mid \beta_{V,1}',\ldots,\beta_{V,m}',Y_1,\ldots,Y_n\big\},$$

a.s., by Lemma 1(a), we have

$$\begin{aligned}
\mathcal{R}\big(\hat f_Y^{[m,K]}, f_Y\big) &= E\big|\hat f_Y^{[m,K]}(V) - f_Y(V)\big|^2 \\
&= E\big[E\big\{\big|\hat f_Y^{[m,K]}(V) - f_Y(V)\big|^2 \mid \beta'_{V,1},\ldots,\beta'_{V,m},Y_1,\ldots,Y_n\big\}\big] \\
&= E\left[\operatorname{var}\big\{f_Y(V) \mid \beta'_{V,1},\ldots,\beta'_{V,m},Y_1,\ldots,Y_n\big\}\right] \\
&\qquad\qquad + E\left|\hat f_Y^{[m,K]}(V) - f_Y^{[m]}(\beta'_{V,1},\ldots,\beta'_{V,m})\right|^2 \\
&= E\left[\operatorname{var}\big\{f_Y(V) \mid \beta'_{V,1},\ldots,\beta'_{V,m}\big\}\right] + E\left|\hat f_Y^{[m,K]}(V) - f_Y^{[m]}(\beta'_{V,1},\ldots,\beta'_{V,m})\right|^2 \\
&\le \mathcal{D} + E\big|\hat f_Y^{[m,K]}(V) - f_Y^{[m]}(\beta'_{V,1},\ldots,\beta'_{V,m})\big|^2, \qquad \text{(S1.10)}
\end{aligned}$$

using also Lemma 1(b). Using the definition (3.6) of the estimator $\hat f_Y^{[m,K]}$ and Parseval's identity with respect to the orthonormal basis of the $H_{k_1,\ldots,k_m}$ in $L_{2,g_1}(\mathbb{R}^m)$, we get

$$\begin{aligned}
E\big|\hat f_Y^{[m,K]}(V) - f_Y^{[m]}(\beta'_{V,1},\ldots,\beta'_{V,m})\big|^2 &= E\Big\|\sum_{k_1,\ldots,k_m\ge 0} 1\{k_1+\cdots+k_m\le K\} \\
&\qquad \cdot \frac{1}{n}\sum_{j=1}^n H_{k_1,\ldots,k_m}(\beta'_{Y_j,1}/\sigma,\ldots,\beta'_{Y_j,m}/\sigma)\cdot H_{k_1,\ldots,k_m} - f_Y^{[m]}(\sigma\cdot)\Big\|_{g_1}^2 \\
= \sum_{k_1,\ldots,k_m\ge 0} 1\{k_1+\cdots+k_m\le K\} & \\
\cdot E\Big|\frac{1}{n}\sum_{j=1}^n H_{k_1,\ldots,k_m}(\beta'_{Y_j,1}/\sigma,\ldots,\beta'_{Y_j,m}/\sigma) &- \big\langle f_Y^{[m]}(\sigma\cdot), H_{k_1,\ldots,k_m}\big\rangle_{g_1}\Big|^2 + \mathcal{B}.
\end{aligned}$$

(S1.11)

Since, from (3.5),

$$E\Big\{\frac{1}{n}\sum_{j=1}^{n} H_{k_1,\ldots,k_m}(\beta'_{Y_j,1}/\sigma,\ldots,\beta'_{Y_j,m}/\sigma)\Big\} \;=\; \big\langle f_Y^{[m]}(\sigma\cdot), H_{k_1,\ldots,k_m}\big\rangle_{g_1}\,,$$

it follows that

$$\begin{aligned} E\Big|\frac{1}{n}\sum_{j=1}^{n} H_{k_1,\ldots,k_m}(\beta'_{Y_j,1}/\sigma,\ldots,\beta'_{Y_j,m}/\sigma) - \big\langle f_Y^{[m]}(\sigma\cdot), H_{k_1,\ldots,k_m}\big\rangle_{g_1}\Big|^2 \\ \;=\; \mathrm{var}\Big(\frac{1}{n}\sum_{j=1}^{n} H_{k_1,\ldots,k_m}(\beta'_{Y_j,1}/\sigma,\ldots,\beta'_{Y_j,m}/\sigma)\Big) \\ \;\le\; \frac{1}{n}\cdot E H^2_{k_1,\ldots,k_m}(\beta'_{Y_1,1}/\sigma,\ldots,\beta'_{Y_1,m}/\sigma)\,. \end{aligned}$$

Using the fact that the Hermite polynomials form an Appell sequence (see e.g. Appell, 1880) we deduce that

$$\begin{aligned} E\big\{ H^2_{k_1,\ldots,k_m}(\beta'_{Y_1,1}/\sigma,\ldots,\beta'_{Y_1,m}/\sigma)\big\} \;=\; E\Big[\prod_{l=1}^{m} E\big\{H^2_{k_l}(\beta'_{X_1,l}/\sigma+\beta'_{V,1,l}/\sigma)\mid X'_1\big\}\Big] \\ = E\Big[\prod_{l=1}^{m}\frac{1}{k_l!}E\Big\{\Big(\sum_{j=0}^{k_l}\sqrt{j!}\binom{k_l}{j}\sigma^{j-k_l}{\beta'_{X_1,l}}^{k_l-j}H_j(\beta'_{V,1,l}/\sigma)\Big)^2\mid X'_1\Big\}\Big] \\ = E\Big[\prod_{l=1}^{m}\Big\{\frac{1}{k_l!}\sum_{j,j'=0}^{k_l}\sqrt{j!j'!}\binom{k_l}{j}\binom{k_l}{j'}\sigma^{j+j'-2k_l}{\beta'_{X_1,l}}^{2k_l-j-j'} \\ \cdot\int H_j(t)H_{j'}(t)\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\,dt\Big\}\Big] \\ = E\Big[\prod_{l=1}^{m}\Big\{\frac{1}{k_l!}\sum_{j=0}^{k_l}j!\binom{k_l}{j}^2\sigma^{2(j-k_l)}{\beta'_{X_1,l}}^{2(k_l-j)}\Big\}\Big] \\ =\; E\Big[\prod_{l=1}^{m}\Big\{\sum_{j=0}^{k_l}\frac{1}{j!}\binom{k_l}{j}\sigma^{-2j}{\beta'_{X_1,l}}^{2j}\Big\}\Big] \\ \le E\Big[\prod_{l=1}^{m}\big\{1+(\beta'_{X_1,l}/\sigma)^2\big\}^{k_l}\Big] \;\le\; \exp\big(KC^2_{X,1}/\sigma^2\big)\,, \qquad \text{(S1.12)} \end{aligned}$$

using the orthonormality of the $H_{k_1,\ldots,k_m}$ with respect to $\langle\cdot,\cdot\rangle_{g_1}$. Using elementary arguments from combinatorics, we also have

$$\#\big\{(k_1,\ldots,k_m)\in\mathbb{N}_0\,:\,k_1+\cdots+k_m\le K\big\}\;=\;\binom{K+m}{K}.$$

Combined with (S1.12), this implies that the first term in (S1.11) is bounded from above by $\mathcal{V}$. Combining this with the other derivations above completes the proof of the theorem.

S1.5 Proof of Theorem 3

The next lemma gives an upper bound on the term $\mathcal{B}$ defined in Theorem 2. It will be used to prove the theorem.

Lemma 1. *Under Assumptions 1 and 2, the term $\mathcal{B}$ in Theorem 2 satisfies $\mathcal{B}=\mathcal{O}\big\{(C_{X,1}/\sigma)^{2K}(2C_{X,1}/\sigma+\sqrt{2})^{2K}/(K+1)!\big\}$, where the constants contained in $\mathcal{O}(\cdots)$ only depend on $C_{X,1}$ and σ.*

Proof of Lemma 1: By Taylor expansion we can write $f_Y^{[m]}=T_{m,K}+R_{m,K}$ where $R_{m,K}$ is a remainder term that will be treated below, and

$$T_{m,K}(s_1,\ldots,s_m)\;=\;E\Big\{\sum_{k=0}^{K}\frac{1}{k!}\sigma^{-2k}\Big(\sum_{j=1}^{m}\beta'_{X_1,j}\cdot s_j\Big)^k\exp\Big(-\frac{1}{2\sigma^2}\sum_{j=1}^{m}{\beta'_{X_1,j}}^2\Big)\Big\}$$

$$=\sum_{k=0}^{K}\frac{1}{k!}\sigma^{-2k}\sum_{j_1,\ldots,j_k=1}^{m}\Big(\prod_{l=1}^{k}s_{j_l}\Big)\cdot E\Big\{\Big(\prod_{l=1}^{k}\beta'_{X_1,j_l}\Big)\exp\Big(-\frac{1}{2\sigma^2}\sum_{j=1}^{m}{\beta'_{X_1,j}}^2\Big)\Big\}.$$

(Assumption 2 guarantees integrability of the above terms). Now $T_{m,K}(\sigma\cdot)$ is an m-variate polynomial of degree $\le K$, so that $T_{m,K}(\sigma\cdot)$ is contained in the linear subspace $\mathcal{H}_{m,K}$ of $L_{2,g_1}(\mathbb{R}^m)$. It follows from there that

$$\mathcal{B} \le \big\|R_{m,K}(\sigma\cdot)\big\|_{g_1}^2 . \tag{S1.13}$$

Next, using the Lagrange representation, the remainder term $R_{m,K}$ has the following upper bound:

$$\begin{aligned}\big|R_{m,K}(s_1,\ldots,s_m)\big| \le \frac{1}{(K+1)!} E\Big[\Big|\frac{1}{\sigma^2}\sum_{j=1}^m \beta'_{X_1,j}\cdot s_j\Big|^{K+1} \\ \cdot \max\Big\{\exp\Big(-\frac{1}{\sigma^2}\sum_{j=1}^m \beta'_{X_1,j}\cdot s_j\Big),1\Big\}\exp\Big(-\frac{1}{2\sigma^2}\sum_{j=1}^m {\beta'_{X_1,j}}^2\Big)\Big],\end{aligned}$$

so that, by Jensen's inequality,

$$\begin{aligned}\big\|R_{m,K}(\sigma\cdot)\big\|_{g_1}^2 \le \frac{1}{[(K+1)!]^2} E\Big[\Big|\frac{1}{\sigma^2}\sum_{j=1}^m \beta'_{X_1,j}\cdot \beta'_{V,j}\Big|^{2(K+1)} \\ \cdot \max\Big\{\exp\Big(-\frac{2}{\sigma^2}\sum_{j=1}^m \beta'_{X_1,j}\cdot \beta'_{V,j}\Big),1\Big\}\exp\Big(-\frac{1}{\sigma^2}\sum_{j=1}^m {\beta'_{X_1,j}}^2\Big)\Big].\end{aligned} \tag{S1.14}$$

Conditionally on X_1', the random variable $\sum_{j=1}^m \beta'_{X_1,j}\cdot\beta'_{V,j}/\sigma^2$ is normally distributed with mean 0 and variance $\kappa_m^2 = \sum_{j=1}^m {\beta'_{X_1,j}}^2/\sigma^2$. Thus, the right hand side of (S1.14) can be expressed as

$$\frac{1}{\{(K+1)!\}^2} E\Big(\kappa_m^{2(K+1)}\exp\big(-\kappa_m^2\big)\, E\big[\delta^{2(K+1)}\cdot\max\big\{\exp\big(2\kappa_m\delta\big),1\big\}\mid X_1'\big]\Big),$$

where $\delta \sim N(0,1)$ and X_1' are independent. Thus (S1.14) has the following upper bound:

$$
\begin{aligned}
&\frac{1}{\{(K+1)!\}^2} E\Big\{\kappa_m^{2(K+1)} \exp\big(-\kappa_m^2\big)\, E\delta^{2(K+1)}\Big\} \\
&\qquad + \frac{1}{\{(K+1)!\}^2} E\Big[\kappa_m^{2(K+1)} \exp\big(-\kappa_m^2\big)\, E\big\{\delta^{2(K+1)} \exp\big(2\kappa_m \delta\big) \mid X_1'\big\}\Big] \\
&= \frac{1}{\{(K+1)!\}^2} E\Big\{\kappa_m^{2(K+1)} \exp\big(-\kappa_m^2\big)\Big\}\, 2^{K+1}\Gamma(K+3/2)/\sqrt{\pi} \\
&\quad + \frac{1}{\{(K+1)!\}^2} E\Big\{\kappa_m^{2(K+1)} \exp\big(-\kappa_m^2\big) \int s^{2(K+1)} \exp\big(2\kappa_m s - s^2/2\big)\, ds\Big\}/\sqrt{2\pi} \\
&\le \mathcal{O}\big\{(2C_{X,1}/\sigma)^{2K}/(K+1)!\big\} \\
&\quad + \frac{1}{\{(K+1)!\}^2} E\Big\{\kappa_m^{2(K+1)} \exp\big(\kappa_m^2\big) \int (s+2\kappa_m)^{2(K+1)} \exp\big(-s^2/2\big)\, ds\Big\}/\sqrt{2\pi} \\
&\le \mathcal{O}\big\{(2C_{X,1}/\sigma)^{2K}/(K+1)!\big\} \\
&+ \frac{1}{\{(K+1)!\}^2} E\Big\{\kappa_m^{2(K+1)} \exp\big(\kappa_m^2\big)\big(\sqrt{2}+2\kappa_m\big)^{2(K+1)}\Big\} \max\Big\{1, \Gamma(K+3/2)/\sqrt{\pi}\Big\} \\
&= \mathcal{O}\big\{(C_{X,1}/\sigma)^{2K}(2C_{X,1}/\sigma+\sqrt{2})^{2K}/(K+1)!\big\},
\end{aligned}
$$

where we have used Assumption 2, which guarantees that $\kappa_m \le C_{X,1}/\sigma$; the fact that $E\delta^{2(K+1)} = \Gamma(K+3/2)2^{K+1}/\sqrt{\pi}$ and Minkowski's inequality.

□

Proof of Theorem 3. Since Assumption 2 holds, we can apply Theorem 2. First we consider the variance term $\mathcal{V}$. Using Stirling's approximation, we have

$$\begin{aligned}
&\exp\big(KC_{X,1}^2/\sigma^2\big)\binom{K+m}{K} \\
&\asymp \frac{1}{\sqrt{2\pi}}\exp\big(KC_{X,1}^2/\sigma^2\big)\sqrt{\frac{1}{m}+\frac{1}{K}}\cdot\big(1+m/K\big)^K\cdot\big(1+K/m\big)^m \\
&\leq \frac{1}{\sqrt{2\pi}}\cdot\exp\big\{K(1+C_{X,1}^2/\sigma^2+\log 2)\big\}\cdot(m/K)^K\,,
\end{aligned}$$

since $m \geq K$ for n sufficiently large. We deduce that

$$\limsup_{n\to\infty}\sup_{P_X\in\mathcal{F}_X}(\log\mathcal{V})/\log n \;=\; \gamma(1/\gamma-1)-1 \;=\; -\gamma\,.$$

Using Lemma 1, an upper bound for $\log\mathcal{B}$ is given by const. $\cdot\, K - K\log K$, where the constant is uniform over all $P_X \in \mathcal{F}_X$, so that

$$\limsup_{n\to\infty}\sup_{P_X\in\mathcal{F}_X}(\log\mathcal{B})/\log n \;=\; -\gamma\,.$$

Finally, under Assumption 3, $\mathcal{D} = \mathcal{O}\big(n^{-C_{X,3}C_M/2}\big)$ uniformly over all $P_X \in \mathcal{F}_X$. The assumption $C_M > 2/C_{X,3}$ guarantees that $\mathcal{D}$ is asymptotically negligible. $\square$

S1.6 Proof of Theorem 4

We define

$$f_{\mathcal{K}}(x) = K^{K/2}(m-K)^{(m-K)/2}\Big\{\prod_{k\in\mathcal{K}}f\big(\sqrt{K}x_k\big)\Big\}\cdot\Big\{\prod_{k\in\{1,\ldots,m\}\backslash\mathcal{K}}\big|f\big(\sqrt{m-K}x_k\big)\big|\Big\}\,,$$

for all $x \in \mathbb{R}^m$, any integers $m > K > 0$, any subset $\mathcal{K} \subseteq \{1, \ldots, m\}$ with $\#\mathcal{K} = K$ and $f = 1_{(0,1/2]} - 1_{[-1/2,0)}$. Then we introduce the functions

$$f_\theta(x) \;=\; \binom{m}{K}^{-1} \sum_{\mathcal{K}} \big|f_{\mathcal{K}}(x)\big| \;+\; \binom{m}{K}^{-1} \sum_{\mathcal{K}} \theta_{\mathcal{K}} f_{\mathcal{K}}(x)\,,$$

for any vector $\theta = \{\theta_{\mathcal{K}}\}_{\mathcal{K}}$ with $\theta_{\mathcal{K}} \in \{-1, 1\}$. All f_θ's are m-variate Lebesgue probability densities. Then we define the probability measure $\tilde{P}_\theta$ on $\mathfrak{B}(\mathbb{R}^m)$ by

$$\tilde{P}_\theta(B) \;=\; (1-\eta)\cdot 1_B(0) + \eta \int_B f_\theta(x)dx\,, \qquad B \in \mathfrak{B}(\mathbb{R}^m)\,,$$

for some $\eta \in (0,1)$ still to be selected. Now let $\tilde{X} = (\tilde{X}_1, \ldots, \tilde{X}_m)$ be some m-dimensional random vector with the measure $\tilde{P}_\theta$. Then $P_{X,\theta}$ denotes the image measure of the functional random variable X_1 on $\mathfrak{B}(C_{0,0}([0,1]))$ with

$$X_1(t) \;=\; \sum_{j=1}^m \tilde{X}_j \cdot \int_0^t \varphi_j(s)ds\,, \qquad t \in [0,1]\,. \tag{S1.15}$$

Now we show that $P_{X,\theta} \in \mathcal{F}_X$ for all vectors θ. As the φ_j's are continuously differentiable, Assumption 1 holds true. Moreover, the support of each $f_{\mathcal{K}}$ is included in the m-dimensional ball around zero with the radius 1. Therefore the measure $\tilde{P}_\theta$ is also supported on a subset of this ball so that $\|X_1'\|_2^2 \;=\; |\tilde{X}|^2 \;\le\; 1\,, \qquad$ a.s.. Hence Assumption 2 is satisfied. We have that

$$\int_0^1 \varphi_j(s)\big(\Gamma_X\varphi_{j'}\big)(s)\,ds \;=\; 1\big\{\max\{j,j'\} \le m\big\}\cdot E\tilde{X}_j\tilde{X}_{j'} \;=\; 1\{j = j' \le m\}\cdot\frac{\eta}{6m}\,,$$

for all $K \geq 3$ where we have used the fact that f is an odd function. Putting

$$\eta \;=\; 6\sqrt{m}\sqrt{C_{X,2}} \cdot \exp\big(- C_{X,3} m^\gamma /2\big)\,, \tag{S1.16}$$

for m sufficiently large, Assumption 3 is satisfied as well.

Following a usual strategy for the proof of lower bounds, we bound the supremum of the statistical risk from below by the Bayesian risk where the a priori distribution of θ is such that all $\theta_{\mathcal{K}}$'s are i.i.d. $\{-1,1\}$-valued random variables with $P(\theta_{\mathcal{K}} = 1) = 1/2$. Applying the standard formula for the minimal Bayesian risk we deduce that

$$\begin{aligned}\sup_{P_X \in \mathcal{F}_X} \mathcal{R}\big(\hat{f}_n, f_Y\big) \geq & E_\theta \int \big|f_{Y,\theta}(v)\big|^2 dP_V(v) \\ & - \iint \big|E_\theta f_{Y,\theta}(u) f_{Y,\theta}^{(n)}(v)\big|^2 dP_V(u) / E_\theta f_{Y,\theta}^{(n)}(v) dP_V^{(n)}(v)\,,\end{aligned} \tag{S1.17}$$

where $f_{Y,\theta}$ denotes the density of Y_1 with respect to P_V when $X_1 \sim P_{X,\theta}$, and $P_V^{(n)}$ and $f_{Y,\theta}^{(n)}$ denote the n-fold product measure and product density of P_V and $f_{Y,\theta}$, respectively. Note that $P_{Y,\theta}^{(n)}$ is the measure of the observed data. For details on the proof of (S1.17), see Section S1.1.

By Lemma 1 and Equation (3.5), the $L_2(P_V)$-inner product of $f_{Y,\theta'}$ and

$f_{Y,\theta''}$ equals

$$\begin{aligned}
\int f_{Y,\theta'}(v) f_{Y,\theta''}(v) dP_V(v) &= E f^{[m]}_{Y,\theta'}(\beta'_{V_1,1},\ldots,\beta'_{V_1,m}) f^{[m]}_{Y,\theta''}(\beta'_{V_1,1},\ldots,\beta'_{V_1,m}) \\
&= \int \Big\{ \int g_\sigma(s-x)\, d\tilde{P}_{\theta'}(x) \Big\} \Big\{ \int g_\sigma(s-x') d\tilde{P}_{\theta''}(x') \Big\} / g_\sigma(s)\, ds \\
&= \iint \Big\{ \int g_\sigma(s-x)\, g_\sigma(s-x')/g_\sigma(s)\, ds \Big\}\, d\tilde{P}_{\theta'}(x)\, d\tilde{P}_{\theta''}(x') \\
&= \iint \exp\big(x^\dagger x'/\sigma^2\big)\, d\tilde{P}_{\theta'}(x)\, d\tilde{P}_{\theta''}(x') \\
&=: \big\langle \tilde{P}_{\theta'}, \tilde{P}_{\theta''} \big\rangle_{\exp}, \qquad \text{(S1.18)}
\end{aligned}$$

for all $\theta', \theta'' \in \{-1,1\}^{\mathcal{K}}$ since $\tilde{X}$ coincides with the vector $(\beta'_{X_1,1},\ldots,\beta'_{X_1,m})$ from (3.1). Note that $\langle\cdot,\cdot\rangle_{\exp}$ represents an inner product on the linear space of all finite signed measures Q on $\mathfrak{B}(\mathbb{R}^m)$ such that the support of the measure $|Q|$ is included in the m-dimensional closed unit ball around 0. By a slight abuse of the notation we write $\langle f_{\mathcal{K}}, f_{\mathcal{K}'}\rangle_{\exp}$ for the corresponding inner product of the signed measures which are induced by the functions $f_{\mathcal{K}}$ and $f_{\mathcal{K}'}$. We show that the $f_{\mathcal{K}}$ form an orthogonal system with respect to this inner product; precisely we have that

$$\begin{aligned}
\langle f_{\mathcal{K}}, f_{\mathcal{K}'}\rangle_{\exp} &= \iint \exp\big(x^\dagger x'/\sigma^2\big)\, f_{\mathcal{K}}(x)\, f_{\mathcal{K}'}(x')\, dx\, dx' \\
&= 1\{\mathcal{K}=\mathcal{K}'\} \cdot \Big[\iint \exp\big\{st/(\sigma^2 K)\big\}\, f(s)\, f(t)\, ds\, dt \Big]^K \\
&\qquad \cdot \Big[\iint \exp\big\{st/\big(\sigma^2(m-K)\big)\big\}\, |f(s)|\, |f(t)|\, ds\, dt \Big]^{m-K} \\
&= 1\{\mathcal{K}=\mathcal{K}'\} \cdot \big(16\,\sigma^2 K\big)^{-K} \cdot \big\{1 \pm o(1)\big\}, \qquad \text{(S1.19)}
\end{aligned}$$

if K and $m-K$ tend to infinity as $n\to\infty$.

Combining (S1.18), (S1.19) and the fact that the $\theta_{\mathcal K}$'s are centered random variables we deduce that the first term in (S1.17) equals

$$E_\theta \int \big|f_{Y,\theta}(v)\big|^2 dP_V(v) \,=\, \|S\|_{\exp}^2 \,+\, \eta^2 \binom{m}{K}^{-2} \sum_{\mathcal K} \big\|f_{\mathcal K}\big\|_{\exp}^2\,, \tag{S1.20}$$

where $\|\cdot\|_{\exp}$ stands for the norm which is induced by $\langle\cdot,\cdot\rangle_{\exp}$ and the measure S on $\mathfrak{B}(\mathbb{R}^m)$ is defined by

$$S(B) \,=\, (1-\eta)\,1_B(0) \,+\, \eta \binom{m}{K}^{-1} \sum_{\mathcal K} \int_B |f_{\mathcal K}(x)|\,dx\,, \qquad B\in\mathfrak{B}(\mathbb{R}^m)\,.$$

The second term in (S1.17) is

$$\begin{aligned} &E_{\theta',\theta''} \int \Big\{\int f_{Y,\theta'}(u) f_{Y,\theta''} dP_V(u)\Big\} f_{Y,\theta'}^{(n)}(v) f_{Y,\theta''}^{(n)}(v) / \big\{E_\theta f_{Y,\theta}^{(n)}(v)\big\} dP_V^{(n)}(v) \\ &\quad = \|S\|_{\exp}^2 + \eta^2 \binom{m}{K}^{-2} \sum_{\mathcal K} \big\|f_{\mathcal K}\big\|_{\exp}^2 \cdot \int \big\{E_\theta \theta_{\mathcal K} f_{Y,\theta}^{(n)}(v)\big\}^2 / E_\theta f_{Y,\theta}^{(n)}(v) dP_V^{(n)}(v)\,, \end{aligned}$$

where, here, θ' and θ'' denote two independent copies of θ. There we have used the fact that

$$E_\theta \int f_{Y,\theta}^{(n)}(v)\,dP_V^{(n)}(v) \,=\, 1\,, \qquad E_\theta \theta_{\mathcal K} f_{Y,\theta}^{(n)} \,=\, \frac12 E_\theta f_{Y,\theta(\mathcal K,+)}^{(n)} - \frac12 E_\theta f_{Y,\theta(\mathcal K,-)}^{(n)}\,, \tag{S1.21}$$

where $\theta(\mathcal K,\pm)$ denotes the vector θ with $\theta_{\mathcal K}$ replaced by ±1; hence,

$$\int E_\theta \theta_{\mathcal K} f_{Y,\theta}^{(n)}(v)\,dP_V^{(n)}(v) \,=\, 0\,.$$

Together with (S1.20) this implies that the right hand side of (S1.17) equals

$$\eta^2 \binom{m}{K}^{-2} \sum_{\mathcal{K}} \|f_{\mathcal{K}}\|_{\exp}^2 \cdot \Big[1 - \int \big\{E_\theta \theta_{\mathcal{K}} f_{Y,\theta}^{(n)}(v)\big\}^2 / E_\theta f_{Y,\theta}^{(n)}(v)\, dP_V^{(n)}(v)\Big]\,. \tag{S1.22}$$

Using (S1.21) and the fact that $E_\theta f_{Y,\theta}^{(n)} = \frac{1}{2} E_\theta f_{Y,\theta(\mathcal{K},+)}^{(n)} + \frac{1}{2} E_\theta f_{Y,\theta(\mathcal{K},-)}^{(n)}$, we establish that

$$\begin{aligned} 1 - \int \big\{E_\theta \theta_{\mathcal{K}} f_{Y,\theta}^{(n)}(v)\big\}^2 / E_\theta f_{Y,\theta}^{(n)}(v) dP_V^{(n)}(v) & \\ \geq 2 \int \sqrt{E_\theta f_{Y,\theta(\mathcal{K},+)}^{(n)}(v)} \sqrt{E_\theta f_{Y,\theta(\mathcal{K},-)}^{(n)}(v)}\, dP_V^{(n)}(v) - 1 & \\ \geq 2 E_\theta \int \sqrt{f_{Y,\theta(\mathcal{K},+)}^{(n)}(v)} \sqrt{f_{Y,\theta(\mathcal{K},-)}^{(n)}(v)}\, dP_V^{(n)}(v) \; - \; 1 & \\ = 2 E_\theta \Big(\int \sqrt{f_{Y,\theta(\mathcal{K},+)}(v)} \sqrt{f_{Y,\theta(\mathcal{K},-)}(v)}\, dP_V(v)\Big)^n \; - \; 1\,, & \end{aligned} \tag{S1.23}$$

by the Cauchy-Schwarz inequality. The Hellinger affinity between the densities $f_{Y,\theta(\mathcal{K},+)}$ and $f_{Y,\theta(\mathcal{K},-)}$ is bounded from below by the corresponding χ^2-distance, i.e.

$$\int \sqrt{f_{Y,\theta(\mathcal{K},+)}(v)} \sqrt{f_{Y,\theta(\mathcal{K},-)}(v)}\, dP_V(v) \;\geq\; 1 - \frac{1}{2} \chi^2 \big\{f_{Y,\theta(\mathcal{K},+)}, f_{Y,\theta(\mathcal{K},-)}\big\}\,,$$

where $\chi^2(f,g) = \int (f-g)^2 / f\, dP_V$. We refer to the book of Tsybakov (2009) for an intensive review on these information distances. We deduce that

$$f_{Y,\theta(\mathcal{K},+)}(V_1) \;=\; f_{Y,\theta}^{[m]}(\beta'_{V_1,1}, \ldots, \beta'_{V_1,m}) \;\geq\; 1 - \eta\,, \text{ a.s.}\,.$$

Equipped with this inequality and (S1.18) we consider that

$$\chi^2\big(f_{Y,\theta(\mathcal{K},+)}, f_{Y,\theta(\mathcal{K},-)}\big) \;\le\; \frac{1}{1-\eta}\,\big\|\tilde{P}_{\theta(\mathcal{K},+)} - \tilde{P}_{\theta(\mathcal{K},-)}\big\|_{\exp}^2 \;\le\; \frac{4\eta^2}{1-\eta}\binom{m}{K}^{-2}\big\|f_{\mathcal{K}}\big\|_{\exp}^2 .$$

Combining this with (S1.19), (S1.22) and (S1.23) we obtain that

$$\sup_{P_X\in\mathcal{F}_X} \mathcal{R}\big(\hat{f}_n, f_Y\big) \ge \eta^2\binom{m}{K}^{-1}(16\sigma^2 K)^{-K} \cdot \big\{1\pm o(1)\big\}\cdot\Big(1-\frac{2\eta^2}{1-\eta}\binom{m}{K}^{-2}(16\sigma^2K)^{-K}\cdot\big\{1\pm o(1)\big\}\Big)^n . \tag{S1.24}$$

Now we take $m = \lfloor (D_M \log n)^{1/\gamma}\rfloor$ and $K = \lfloor D_K(\log n)/\log(\log n)\rfloor$ for some constants $D_M, D_K > 0$. Whenever $-D_M C_{X,3} - 2D_K(1/\gamma - 1) - D_K < -1$, the inequality (S1.24), together with (S1.16), yields that

$$\liminf_{n\to\infty}\sup_{P_X\in\mathcal{F}_X}\big\{\log\mathcal{R}\big(\hat{f}_n, f_Y\big)\big\}/\log n \;\ge\; -D_M C_{X,3} - D_K/\gamma .$$

We may choose $D_K = \gamma/(2-\gamma)$ and $D_M > 0$ arbitrarily close to 0, which completes the proof of the theorem. □

S1.7 Proof of Theorem 5

The proof follows a usual structure of adaptivity proofs for cross-validation techniques, see e.g. Section 2.5.1 in the book of Meister (2009) for a related proof in the field of density deconvolution.

Let (m_n, K_n) be defined as in the statement of Theorem 3 and define the set

$$G' = \left\{(m,K) \in G : \mathcal{R}(\hat{f}_Y^{[m,K]}, f_Y) > 2\,\mathcal{R}(\hat{f}_Y^{[m_n,K_n]}, f_Y)\right\}.$$

Using the notation $\|g\|_{P_V}^2 = \int |g(x)|^2 dP_V(x)$, for any $g \in L_2(P_V)$, we need to prove that $\lim_{n\to\infty} \sup_{P_X \in \mathcal{F}_X} P\big(n^\gamma \|\hat{f}_Y^{[\hat{m},\hat{K}]} - f_Y\|_{P_V}^2 \geq n^d\big) = 0$.

By Markov's inequality we have

$$\begin{aligned} & P\big(n^\gamma \|\hat{f}_Y^{[\hat{m},\hat{K}]} - f_Y\|_{P_V}^2 \geq n^d\big) \\ & \quad \leq \sum_{(m,K)\in G\backslash G'} P\big(\|\hat{f}_Y^{[m,K]} - f_Y\|_{P_V}^2 > n^{-\gamma+d}\big) + P\big[(\hat{m},\hat{K}) \in G'\big] \\ & \quad \leq 2(\#G)\cdot n^{\gamma-d}\cdot \mathcal{R}\big(\hat{f}_Y^{[m_n,K_n]}, f_Y\big) + \sum_{(m,k)\in G'} P\big(\hat{m} = m, \hat{K} = K\big). \end{aligned} \tag{S1.25}$$

By Theorem 3, the first term in (S1.25) converges to 0 as $n \to \infty$ uniformly over $P_X \in \mathcal{F}_X$. It remains to study the second term.

On the event $\{\hat{m} = m, \hat{K} = K\}$, we have $\mathrm{CV}(m,K) \leq \mathrm{CV}(m_n,K_n)$ and, hence also

$$\begin{aligned} & \big\|\hat{f}_Y^{[m,K]}\big\|_{P_V}^2 - E\big\|\hat{f}_Y^{[m,K]}\big\|_{P_V}^2 - 2\Delta_1(m,K) - 4\Delta_2(m,K) + \mathcal{R}\big(\hat{f}_Y^{[m,K]}, f_Y\big) \\ & \leq \big\|\hat{f}_Y^{[m_n,K_n]}\big\|_{P_V}^2 - E\big\|\hat{f}_Y^{[m_n,K_n]}\big\|_{P_V}^2 - 2\Delta_1(m_n,K_n) - 4\Delta_2(m_n,K_n) \\ & \qquad + \mathcal{R}\big(\hat{f}_Y^{[m_n,K_n]}, f_Y\big), \end{aligned} \tag{S1.26}$$

where $\Delta_1(m,K) = \{n(n-1)\}^{-1} \sum_{i\neq i'} \sum_{\mathbf{k}\in\mathcal{K}(m,K)} \overline{\Xi}(i,m,\mathbf{k}) \cdot \overline{\Xi}(i',m,\mathbf{k})$, $\Delta_2(m,K) = n^{-1} \sum_{i=1}^n \sum_{\mathbf{k}\in\mathcal{K}(m,K)} \big\{E\,\Xi(1,m,\mathbf{k})\big\} \cdot \overline{\Xi}(i,m,\mathbf{k})$, $\mathcal{K}(m,K) =$

$\{\mathbf{k} \in \mathbb{N}_0^m \,:\, k_1 + \cdots + k_m \le K\}$, $\Xi(j,m,\mathbf{k}) \,=\, H_{\mathbf{k}}\big(\beta'_{Y_j,1}/\sigma,\ldots,\beta'_{Y_j,m}/\sigma\big)$ and $\overline{\Xi}(j,m,\mathbf{k}) \,=\, \Xi(j,m,\mathbf{k}) - E\,\Xi(j,m,\mathbf{k})$.

The first terms of both sides of the inequality at (S1.26) can be represented as follows, using the orthonormality of the $H_{\mathbf{k}}$'s:

$$
\begin{aligned}
&\|\hat f_Y^{[m,K]}\|_{P_V}^2 - E\|\hat f_Y^{[m,K]}\|_{P_V}^2 \\
&\quad = \sum_{\mathbf{k}\in\mathcal{K}(m,K)} \Big\{ \Big|\frac{1}{n}\sum_{i=1}^n \Xi(i,m,\mathbf{k})\Big|^2 - E\Big|\frac{1}{n}\sum_{i=1}^n \Xi(i,m,\mathbf{k})\Big|^2 \Big\} \\
&\quad = \frac{1}{n^2}\sum_{i,i'}\sum_{\mathbf{k}\in\mathcal{K}(m,K)} \big\{\Xi(i,m,\mathbf{k})\Xi(i',m,\mathbf{k}) - E\,\Xi(i,m,\mathbf{k})\Xi(i',m,\mathbf{k})\big\} \\
&\quad = (1-1/n)\big\{\Delta_1(m,K) + 2\Delta_2(m,K)\big\} + \Delta_3(m,K)\,,
\end{aligned}
$$

where $\Delta_3(m,K) \,= n^{-2}\sum_{i=1}^n \sum_{\mathbf{k}\in\mathcal{K}(m,K)} \big\{\Xi^2(i,m,\mathbf{k}) - E\,\Xi^2(i,m,\mathbf{k})\big\}$. Together with (S1.26) this implies that, for $(m,K) \in G'$, $\mathcal{R}\big(\hat f_Y^{[m,K]}, f_Y\big)/2 \,\le\, \Delta_4(m,K,m_n,K_n)$, where

$$
\begin{aligned}
\Delta_4(m,K,m_n,K_n) &= (1+1/n)\big\{\big|\Delta_1(m,K)\big| + \big|\Delta_1(m_n,K_n)\big| \\
&\quad + 2\big|\Delta_2(m,K) - \Delta_2(m_n,K_n)\big|\big\} + |\Delta_3(m,K)| + |\Delta_3(m_n,K_n)|\,.
\end{aligned}
$$

Hence the second term in (S1.25) has the following upper bound:

$$
2\sum_{(m,K)\in G'} \big\{\mathcal{R}\big(\hat f_Y^{[m,K]}, f_Y\big)\big\}^{-1}\big\{E\Delta_4^2(m,K,m_n,K_n)\big\}^{1/2}\,. \tag{S1.27}
$$

In order to bound (S1.27), we need a lower bound on $\mathcal{R}\big(\hat f_Y^{[m,K]}, f_Y\big)$. Theorem 2 provides only an upper bound to this term but an inspection of the proof of this theorem – in particular (S1.10) to (S1.12) – yields that

$$\mathcal{R}\big(\hat{f}_Y^{[m,K]}, f_Y\big) \;\geq\; \mathcal{B}(m,K) + \mathcal{V}_n^*(m,K) + \mathcal{D}^*(m) - \frac{1}{n}\big\|f_Y^{[m]}(\sigma\cdot)\big\|_{g_1}^2\,, \quad \text{(S1.28)}$$

where $\mathcal{B}(m,K)$ is the term $\mathcal{B}$ from Theorem 2 and

$$\mathcal{D}^*(m) \;=\; E\,\mathrm{var}\{f_Y(V_1)|\mathfrak{A}_m\}\,, \quad \mathcal{V}_n^*(m,K) \;=\; \frac{1}{n}\binom{K+m}{K}\,.$$

Here we have used the fact that $E\big\{H^2_{k_1,\ldots,k_m}(\beta'_{Y_1,1}/\sigma,\ldots,\beta'_{Y_1,m}/\sigma)\big\} \;\geq\; 1$, which comes from the first lines of (S1.12).

In order to bound (S1.27), we also need an upper bound for $E\Delta_4^2(m,K,m_n,K_n)$, which involves Δ_1 to Δ_3. For Δ_1 we have

$$\begin{aligned}
E\big|\Delta_1(m,K)\big|^2 \;&=\; \frac{2}{n(n-1)}\sum_{\mathbf{k},\mathbf{k}'\in\mathcal{K}(m,K)}\big[\mathrm{cov}\big\{\Xi(1,m,\mathbf{k}),\Xi(1,m,\mathbf{k}')\big\}\big]^2 \\
&= \frac{2}{n(n-1)}\sum_{\mathbf{k},\mathbf{k}'\in\mathcal{K}(m,K)}\big[\big\langle H_{\mathbf{k}}, H_{\mathbf{k}'}f_Y^{[m]}(\sigma\cdot)\big\rangle_{g_1} \\
&\qquad\qquad - \big\langle H_{\mathbf{k}}, f_Y^{[m]}(\sigma\cdot)\big\rangle_{g_1}\cdot\big\langle H_{\mathbf{k}'}, f_Y^{[m]}(\sigma\cdot)\big\rangle_{g_1}\big]^2 \\
&\leq \frac{4}{n(n-1)}\sum_{\mathbf{k},\mathbf{k}'\in\mathcal{K}(m,K)}\big\langle H_{\mathbf{k}}, H_{\mathbf{k}'}f_Y^{[m]}(\sigma\cdot)\big\rangle_{g_1}^2 \\
&\qquad\qquad + \frac{4}{n(n-1)}\Big(\sum_{\mathbf{k}\in\mathcal{K}(m,K)}\big\langle H_{\mathbf{k}}, f_Y^{[m]}(\sigma\cdot)\big\rangle_{g_1}^2\Big)^2 \\
&\leq \frac{4}{n(n-1)}\sum_{\mathbf{k}'\in\mathcal{K}(m,K)}\big\|H_{\mathbf{k}'}f_Y^{[m]}(\sigma\cdot)\big\|_{g_1}^2 + \frac{4}{n(n-1)}\big\|f_Y^{[m]}(\sigma\cdot)\big\|_{g_1}^4 \\
&\leq \frac{4}{n-1}\big\|\{f_Y^{[m]}(\sigma\cdot)\}^2\big\|_{g_1}\cdot\sup_{\mathbf{k}\in\mathcal{K}(m,K)}\big\|H_{\mathbf{k}}^2\big\|_{g_1}\mathcal{V}_n^*(m,K) + \frac{4}{n(n-1)}\big\|\{f_Y^{[m]}(\sigma\cdot)\}^2\big\|_{g_1}^2,
\end{aligned}$$

(S1.29)

where we have used Parseval's identity with respect to the orthonormal system $H_{\mathbf{k}}$.

For the term involving Δ_2 we have

$$\begin{aligned}
&E\Big|\Delta_2(m,K)-\Delta_2(m_n,K_n)\Big|^2 \\
&\quad\le \frac{1}{n}\,E\Big[\sum_{\mathbf{k}\in\mathcal{K}(m,K)}\Xi(1,m,\mathbf{k})\{E\,\Xi(1,m,\mathbf{k})\} \\
&\qquad\qquad - \sum_{\mathbf{k}\in\mathcal{K}(m_n,K_n)}\Xi(1,m_n,\mathbf{k})\{E\,\Xi(1,m_n,\mathbf{k})\}\Big]^2 \\
&\quad= \frac{1}{n}\,E\Big[\sum_{\mathbf{k}\in\mathbb{N}_0^{\overline{m}}}\big\{1_{\mathcal{K}^{\overline{m}}(m,K)}(\mathbf{k})-1_{\mathcal{K}^{\overline{m}}(m_n,K_n)}(\mathbf{k})\big\}\Xi(1,\overline{m},\mathbf{k})\{E\,\Xi(1,\overline{m},\mathbf{k})\}\Big]^2 \\
&\quad= \frac{1}{n}\sum_{\mathbf{k},\mathbf{k}'}\big\{1_{\mathcal{K}^{\overline{m}}(m,K)}(\mathbf{k})-1_{\mathcal{K}^{\overline{m}}(m_n,K_n)}(\mathbf{k})\big\}\big\{1_{\mathcal{K}^{\overline{m}}(m,K)}(\mathbf{k}')-1_{\mathcal{K}^{\overline{m}}(m_n,K_n)}(\mathbf{k}')\big\} \\
&\qquad\qquad \cdot\langle H_{\mathbf{k}},f_Y^{[\overline{m}]}(\sigma\cdot)\rangle_{g_1}\langle H_{\mathbf{k}'},f_Y^{[\overline{m}]}(\sigma\cdot)\rangle_{g_1}\big\langle H_{\mathbf{k}},H_{\mathbf{k}'}f_Y^{[\overline{m}]}(\sigma\cdot)\big\rangle_{g_1}\,,
\end{aligned} \tag{S1.30}$$

with $\overline{m}=\max\{m,m_n\}$ and $\mathcal{K}^{\overline{m}}(m,K)=\big\{\mathbf{k}\in\mathcal{K}(\overline{m},K)\,:\,k_l=0,\,\forall l>m\big\}$. Here we have used the fact that $H_0\equiv 1$ and $E\big\{f_Y^{[\overline{m}]}(\beta'_{V_1,1},\ldots,\beta'_{V_1,\overline{m}})\mid \mathfrak{A}_m\big\}=f_Y^{[m]}(\beta'_{V_1,1},\ldots,\beta'_{V_1,m})$ a.s., which follows from Lemma 1(a).

Applying the Cauchy-Schwarz inequality and Parseval's identity, we get that the right hand side of (S1.30) has the following upper bound:

$$\frac{1}{n}\Big(\sum_{\mathbf{k}}\big|1_{\mathcal{K}^{\overline{m}}(m,K)}(\mathbf{k})-1_{\mathcal{K}^{\overline{m}}(m_n,K_n)}(\mathbf{k})\big|\langle H_{\mathbf{k}},f_y^{[\overline{m}]}(\sigma\cdot)\rangle_{g_1}^2\Big) \cdot\Big(\sum_{\mathbf{k}\in\mathcal{K}(\overline{m},\overline{K})}\big\|H_{\mathbf{k}}\cdot f_Y^{[\overline{m}]}(\sigma\cdot)\big\|_{g_1}^2\Big)^{1/2}$$

$$\leq \mathcal{V}_n^*(\overline{m},\overline{K})\cdot\Big\{\big\|\mathcal{P}_{\mathcal{H}(m,K)}f_Y^{[\overline{m}]}(\sigma\cdot)\big\|_{g_1}^2+\big\|\mathcal{P}_{\mathcal{H}(m_n,K_n)}f_Y^{[\overline{m}]}(\sigma\cdot)\big\|_{g_1}^2 -2\big\|\mathcal{P}_{\mathcal{H}(\underline{m},\underline{K})}f_Y^{[\overline{m}]}(\sigma\cdot)\big\|_{g_1}^2\Big\}$$
$$\cdot\binom{\overline{m}+\overline{K}}{\overline{m}}^{-1/2}\cdot\big\|\{f_Y^{[\overline{m}]}(\sigma\cdot)\}^2\big\|_{g_1}^{1/2}\cdot\max\big\{\|H_{\mathbf{k}}^2\|_{g_1}^{1/2}:\mathbf{k}\in\mathcal{K}(\overline{m},\overline{K})\big\}\,, \tag{S1.31}$$

where $\overline{K}=\max\{K,K_n\}$, $\underline{m}=\min\{m,m_n\}$, $\underline{K}=\min\{K,K_n\}$ and $\mathcal{P}_{\mathcal{H}(m,k)}$ denotes the orthogonal projector onto the linear subspace $\mathcal{H}(m,K)$ of $L_{2,g_1}(\mathbb{R}^m)$.

Since

$$\begin{aligned}\|\mathcal{P}_{\mathcal{H}(m,K)}f_Y^{[\overline{m}]}(\sigma\cdot)\|_{g_1}^2 &= \|f_Y^{[m]}(\sigma\cdot)\|_{g_1}^2-\mathcal{B}(m,K)\\ &= E|f_Y(V_1)|^2-\mathcal{D}^*(m)-\mathcal{B}(m,K)\,,\end{aligned}$$

then using Lemma 1(a), the right hand side of (S1.31) has the following upper bound:

$$3\mathcal{V}_n^*(\overline{m},\overline{K})\cdot\big\{\mathcal{D}^*(m)+\mathcal{D}^*(m_n)+\mathcal{B}(m,K)+\mathcal{B}(m_n,K_n)\big\}\cdot\binom{\overline{m}+\overline{K}}{\overline{m}}^{-1/2}$$
$$\cdot\big\|\{f_Y^{[\overline{m}]}(\sigma\cdot)\}^2\big\|_{g_1}^{1/2}\cdot\max\big\{\|H_{\mathbf{k}}^2\|_{g_1}^{1/2}:\mathbf{k}\in\mathcal{K}(\overline{m},\overline{K})\big\}\,, \tag{S1.32}$$

since $\mathcal{B}(m,K)$ decreases as K increases.

Finally, for the term involving Δ_3 we have

$$E\big|\Delta_3(m,K)\big|^2 \,\le\, n^{-3}\, E\Big\{\sum_{\mathbf{k}\in\mathcal{K}(m,K)} \Xi^2(1,m,\mathbf{k})\Big\}^2$$
$$\le\, \frac{1}{n}\,\big\{\mathcal{V}_n^*(m,K)\big\}^2\, \big\|\big\{f_Y^{[m]}(\sigma\cdot)\big\}^2\big\|_{g_1}\cdot \max\big\{\|H_{\mathbf{k}}^4\|_{g_1} : \mathbf{k}\in\mathcal{K}(m,K)\big\}\,. \tag{S1.33}$$

In order to bound the terms (S1.29), (S1.32) and (S1.33), we need some technical results. Using the explicit sum representation of the Hermite polynomials we write

$$\int H_k^\ell(x)\frac{1}{\sqrt{2\pi}}\exp(-x^2/2)dx$$
$$= \sum_{i_1,\ldots,i_l=0}^{\lfloor k/2\rfloor} \frac{(k!)^{l/2}\cdot(-2)^{-i_1-\cdots-i_\ell}}{i_1!\cdots i_\ell!\cdot(k-2i_1)!\cdots(k-2i_\ell)!}\int x^{k\ell-2(i_1+\cdots+i_\ell)}\frac{1}{\sqrt{2\pi}}\exp(-x^2/2)dx$$
$$\le \sum_{i=0}^{\ell\lfloor k/2\rfloor} 2^{-i}(k!)^{\ell/2}\frac{\Gamma(k\ell/2+1/2-i)}{i!(k\ell-2i)!}$$
$$\times \sum_{i_1+\cdots+i_\ell=i}\binom{i}{i_1,\ldots,i_\ell}\binom{k\ell-2i}{k-2i_1,\ldots,k-2i_\ell}\Big/\sqrt{\pi}$$
$$\le \sum_{i=0}^{\ell\lfloor k/2\rfloor} 2^{-i}\binom{k\ell/2}{i}\ell^{k\ell-i} \,\le\, (\ell^2+\ell/2)^{k\ell/2}\,,$$

for any $k\in\mathbb{N}_0$ and any even integer $\ell>0$. Furthermore we have

$$\big\|\{f_Y^{[m]}(\sigma\cdot)\}^2\big\|_{g_1}^2 \,=\, E\big|E\big\{f_Y(V_1)|\mathfrak{A}_m\big\}\big|^4 \,\le\, Ef_Y^4(V_1)$$
$$\le\, E\exp\Big(-\frac{2}{\sigma^2}\int_0^1 |X_1'(t)|^2\,dt\Big)\,E\Big\{\exp\Big(\frac{4}{\sigma^2}\int_0^1 X_1'(t)\,dV_1(t)\Big)\big|X_1\Big\}$$
$$\le\, \exp\big\{(8/\sigma^4)\,C_{X,1}^2\big\}\,,$$

where we used (S1.6) and Assumption 2. Applying these results to (S1.29), (S1.32) and (S1.33) and recalling (S1.28), we deduce that (S1.27) has the upper bound

$$(\log n)^{1+1/\gamma_0}\, D_0^{\overline{K}} \left[\left\{ n\, \mathcal{R}(\hat{f}_Y^{[m,K]}, f_Y) \right\}^{-1/2} + \binom{\overline{m}+\overline{K}}{\overline{m}}^{-1/4} \right],$$

for some global finite constant $D_0 > 0$, so that (S1.27) converges to zero uniformly over $P_X \in \mathcal{F}_X$. This completes the proof of the theorem. □

Bibliography

Appell, P. (1988). Sur une classe de polynômes. *Annales scientifiques de l'École Normale Supérieure 2me série* **9**, 119–144.

Hall, R., Rinaldo, A. and Wasserman, L. (2013). Differential privacy for functions and functional data. *J. Mach. Learn. Research* **14**, 703–727.

Meister, A. (2009). *Deconvolution problems in nonparametric statistics*, Springer.

Tsybakov, A.B. (2009). *Introduction to nonparametric estimation*, Springer.