ТЕОРИЯ ВЕРОЯТНОСТЕЙ
Том 69 | И ЕЕ ПРИМЕНЕНИЯ | Выпуск 3
2024

BANERJEE C.*, **SAKHANENKO L. A.***, **ZHU D. C.****

GLOBAL RATE OPTIMALITY OF INTEGRAL CURVE ESTIMATORS IN HIGH ORDER TENSOR MODELS: SUPPLEMENTAL MATERIAL[1)]

В статье приведены доказательства результатов, сформулированных в работе C. Banerjee, L. Sakhanenko, D. C. Zhu “Global rate optimality of integral curve estimators in high order tensor models”, опубликованной в журнале “Теория вероятностей и ее применения”, 2023, **68**(2), с. 95–115.

Ключевые слова и фразы: глобальная минимаксная нижняя граница, полупараметрическая оценка, диффузионная тензорная МРТ с высоким угловым разрешением.

DOI: https://doi.org/10.4213/tvp5566

1. Restatement of the main results.

Theorem 1. *Assume conditions* (A1)–(A6) *and* (A5′) *hold and* $1 \leqslant R \leqslant (M+2)/2$. *Define for* $c, k > 0$,

$$\mathcal{D}_{c,k} = \{D \in \mathcal{D}^2(a_*, G, T)\colon \|D - D_0\|_\infty \leqslant cn^{-2/(d+3)} \text{ and } \|D'' - D_0''\|_\infty \leqslant k\}.$$

Then for any numbers $c > 0$, $T > 0$, *any point* $a_* \in G$, *any* $D \in \mathcal{D}_{c,k}$, *any function* $w \in \mathcal{W}$, *any* $1 \leqslant p \leqslant \infty$, *and some* $k > 0$, *we have*

$$\liminf_{n\to\infty} \inf_{\widehat{X}_n^{(1)},\dots,\widehat{X}_n^{(R)} \in \mathcal{E}_n(a_*,T)} \sup_{D\in\mathcal{D}_{c,k}} \mathbf{E} w\bigl(n^{2/(d+3)}\bigl(\bigl\|\widehat{X}_n^{(1)} - x^{(1)}\bigr\|_{p,T}, \dots, \bigl\|\widehat{X}_n^{(R)} - x^{(R)}\bigr\|_{p,T}\bigr)\bigr) > 0.$$

*Department of Statistics and Probability, Michigan State University, East Lansing, MI, USA; e-mail: banerj25@msu.edu; sakhanen@msu.edu

**Department of Radiology and Gruss Magnetic Resonance Research Center, Albert Einstein College of Medicine, NY, USA; e-mail: zhuda@msu.edu

[1)] This research was partially supported by the National Science Foundation grants DMS–1612867, DMS–2111251 and partially supported by Michigan State University High Performance Computing Center through computational resources provided by the Institute for Cyber–Enabled Research.

Corollary 1. *Assume conditions* (A1)–(A6) *and* (A5′) *hold and* $1 \leqslant R \leqslant (M+2)/2$. *Then for any number* $T > 0$, *any point* $a_* \in G$, *any function* $w \in \mathcal{W}$, *we have*

$$\liminf_{n\to\infty} \inf_{\widehat{X}_n^{(1)},\dots,\widehat{X}_n^{(R)} \in \mathcal{E}_n(a_*,T)} \sup_{D\in\mathcal{D}^2(a_*,G,T)} \mathbf{E}w\big(n^{2/(d+3)}\big(\big\|\widehat{X}_n^{(1)} - x^{(1)}\big\|_{p,T}, \dots, \big\|\widehat{X}_n^{(R)} - x^{(R)}\big\|_{p,T}\big)\big) > 0.$$

Theorem 2. *Assume conditions* (A1)–(A8) *and* (A5′) *hold. Let* $\{\widehat{X}_n^{(r)} : r = 1,\dots,R\}$ *be the CS estimators.*

(a) *If we additionally assume* $\mathbf{E}\|\xi_i\|_q^q < \infty$, *for some* $q \geqslant 4$, *then for each* $r = 1,\dots,R$, *for any number* $T > 0$, *for any point* $a_* \in G$, *and any* $2 \leqslant p \leqslant q$, *we have*

$$\sup_n \sup_{D\in\mathcal{D}^2(a_*,G,T)} \mathbf{E}\big\|n^{2/(d+3)}\big(\widehat{X}_n^{(r)} - x^{(r)}\big)\big\|_{p,T} < \infty.$$

(b) *For each* $r = 1,\dots,R$, *for any number* $T > 0$, *for any point* $a_* \in G$, *we have*

$$\sup_n \sup_{D\in\mathcal{D}^2(a_*,G,T)} \mathbf{E}\big\|n^{2/(d+3)}\big(\widehat{X}_n^{(r)} - x^{(r)}\big)\big\|_{\infty,T} < \infty.$$

2. Proofs. The proof relies on several intermediate results. It is easy to see $\mathcal{D}_{c,k} \subset \mathcal{D}^2(a_*,G,T)$, therefore, Corollary 1 immediately follows from Theorem 1.

2.1. Proof of Theorem 1.

Proof. The proof of Theorem 1 will be based on the following lemmas and are similar to the construction given in [6] and [4]. We additionally provide a result that extends Fano's lemma in the multidimensional parameter space. We use that result in Lemma 3 tailored to fit our problem. The ideas are similar to Theorem 5.7 in [3]. Lemma 2 provides a bound for the function g from section 2 in [1]. We establish the smoothness condition on the perturbed class of tensor fields $\mathcal{D}_{c,k}$ in Lemma 4. Finally, Lemmas 5 and 6 provide a construction for the tensor fields and integral curves which we will utilize in the proof of Theorem 1.

Lemma 1. *For* $r = 1,\dots,R$, *for all* $x \in G_{\epsilon,T}^{(r)}$, *the tensors* $D_b^{(r)}$ *satisfy* $\ker(\mathcal{T}(v_{b,n}^{(r)}(x), D_b^{(r)}(x)) - \lambda_r I) = 0$.

The technical proof is in the last section.

Lemma 2. *Let* f *satisfy condition* (A5′). *Then there exists a* $\delta_1 > 0$ *such that for any vector* u *satisfying* $|u| \leqslant \delta_1$ *we have*

$$g(u) \leqslant C(f,\delta_1)|u|^2$$

with a positive constant $C(f,\delta_1)$.

The proof of Lemma 2 will follow along the same lines of the proof of Lemma 1 from [11]. Now we consider two important propositions. Proposition 1 is *Fano's* lemma described in a multidimensional parameters setting and Proposition 2 provides a bound for the KL-divergence of the densities in our parametric subclass. The proofs of these propositions will be deferred to the appendix sections 4 and 5. Let us define $\|v\|_G^2 = \int_G |v(x)|^2 dx$ for a vector field v.

Proposition 1. *Let X be a random variable whose density f_θ depends on a parameter $\theta = (\theta^{(1)}, \dots, \theta^{(R)})$ from a multidimensional discrete parameter space $\{1 \leqslant \theta^{(r)} \leqslant l_r + 1,\ r = 1, \dots, R,\ l_r \in \mathbf{N}\}$. Suppose for $\eta > 0$, the Kullback–Leibler divergence between f_θ and $f_{\theta'}$ is $K(f_\theta, f_{\theta'}) \leqslant \eta$. Then for an estimator $\Psi(X) = \big(\Psi^{(1)}(X), \dots, \Psi^{(R)}(X)\big)$ of the unknown parameter θ we have*

$$\sup_{\underline{i}} \mathbf{P}_{\underline{i}}\big(\Psi(X) \neq \underline{i}\big) \geqslant 1 - \frac{\eta + \ln 2}{\ln(\mathscr{L} - 1)},$$

where $\underline{i} = (i_1, \dots, i_R)$, $1 \leqslant i_r \leqslant l_r + 1$, $r = 1, \dots, R$, $\mathscr{L} = \prod_{r=1}^R (1 + l_r)$ and $\mathbf{P}_{\underline{i}}$ is the probability induced by $f_{\underline{i}}$.

Proposition 2. *Let $f_{\underline{i}}$ be the density of $(X, B\underline{D}_{\underline{i}}(X) + \Sigma^{1/2}(X)\xi)$, $\underline{i} = (i_1, \dots, i_R)$. Then for any two indices $\underline{i} \neq \underline{j}$, $1 \leqslant i_r, j_r \leqslant l_r + 1$, $r = 1, \dots, R$, we have*

$$\|\underline{D}_{\underline{i}} - \underline{D}_{\underline{j}}\|_G^2 \leqslant \eta \quad \Longrightarrow \quad K(f_{\underline{i}}, f_{\underline{j}}) \leqslant C(f, \delta_1) C_{\Sigma,B} \eta,$$

where $C(f, \delta_1)$ is a constant introduced in Lemma 2 and $C_{\Sigma,B}$ is a constant depending on Σ and B only.

Using these two propositions we will describe and prove Lemma 3, which is a pivotal result in proving Theorem 1 and its Corollary 1.

Lemma 3. *Let $\mathcal{D}_{\mathcal{L}}$ be a class of $\mathscr{L}$ tensor fields D_θ, $\theta = (\theta^{(1)}, \dots, \theta^{(R)})$ inside the parametric subclass of $\mathcal{D}_{c,k}$, such that for any $\theta \neq \widetilde{\theta}$ and positive constants η and $\delta_2^{(r)}$, $r = 1, \dots, R$,*

$$\|\underline{D}_\theta - \underline{D}_{\widetilde{\theta}}\|_G^2 \leqslant \delta_1, \quad \|\underline{D}_\theta - \underline{D}_{\widetilde{\theta}}\|_G^2 \leqslant \eta, \quad \textit{and} \quad \big\|x_\theta^{(r)} - x_{\widetilde{\theta}}^{(r)}\big\|_{1,T} \geqslant 2\delta_2^{(r)},$$

where for each $r = 1, \dots, R$ and θ, we denote by $x_\theta^{(r)} \in \mathbf{R}^d$ the integral curve corresponding to $v_\theta^{(r)}$ starting at a_*. In addition, let f satisfy condition (A5$'$). *Then for any $w \in \mathcal{W}$, we have*

$$\sup_{D \in \mathcal{D}_{\mathcal{L}}} \mathbf{E} w\big(\big\|\widehat{X}_n^{(1)} - x^{(1)}\big\|_{1,T}, \dots, \big\|\widehat{X}_n^{(R)} - x^{(R)}\big\|_{1,T}\big)$$
$$\geqslant \inf_{|x^{(r)}| \geqslant \delta_2^{(r)}:\ r=1,\dots,R} w\big(|x^{(1)}|, \dots, |x^{(R)}|\big) \left(1 - \frac{n C_{\Sigma,B} C(f, \delta_1)\eta + \ln 2}{\ln(\mathscr{L} - 1)}\right).$$

Proof. We will follow the structure of the modified version of Fano's lemma that we described in Proposition 1. Let $\theta = (\theta^{(1)}, \dots, \theta^{(R)})$ is uniformly distributed over $\{1 \leqslant \theta^{(r)} \leqslant l_r + 1,\, r = 1, \dots, R,\, l_r \in \mathbf{N}\}$ such that

$$\mathbf{P}(\theta^{(1)} = i_1,\, \dots,\, \theta^{(R)} = i_R) = \frac{1}{(l_1+1)\cdots(l_R+1)},$$
$$1 \leqslant i_r \leqslant l_r + 1, \qquad r = 1, \dots, R.$$

Let $\mathbf{P}_{\underline{i}}$ be the probability induced by $(X, B\underline{D}_{\underline{i}}(X) + \Sigma^{1/2}(X)\xi)$, $\underline{i} = (i_1, \dots, i_R)$. For each $r = 1, \dots, R$, define the function Ψ as follows:

$$\Psi(X_1, \dots, X_n, \xi_1, \dots, \xi_n) = \big(\Psi^{(1)}(X), \dots, \Psi^{(R)}(X)\big),$$

where $\Psi^{(R)}(X) = \theta^{(r)}$ if $\|\widehat{X}_n^{(r)} - x^{(r)}\|_{1,T} \geqslant \delta_2^{(r)}$. Then by Proposition 1 (a modified version of Fano's lemma) we have

$$\begin{aligned}
&\sup_{D \in \mathcal{D}^2(a_*, G, T)} \mathbf{E} w\big(\big\|\widehat{X}_n^{(1)} - x^{(1)}\big\|_{1,T}, \dots, \big\|\widehat{X}_n^{(R)} - x^{(R)}\big\|_{1,T}\big) \\
&\quad \geqslant \sup_{D \in \mathcal{D}_{\mathcal{L}}} \mathbf{E} w\big(\big\|\widehat{X}_n^{(1)} - x^{(1)}\big\|_{1,T}, \dots, \big\|\widehat{X}_n^{(R)} - x^{(R)}\big\|_{1,T}\big) \\
&\quad \geqslant \inf_{|x^{(r)}| \geqslant \delta_2^{(r)}:\, r=1,\dots,R} w(|x^{(1)}|, \dots, |x^{(R)}|) \\
&\qquad \times \sup_{D \in \mathcal{D}_{\mathcal{L}}} \mathbf{P}\big(\big\|\widehat{X}_n^{(r)} - x_\theta^{(r)}\big\|_{1,T} \geqslant \delta_2^{(r)};\, r = 1, \dots, R\big) \\
&\quad \geqslant \inf_{|x^{(r)}| \geqslant \delta_2^{(r)}:\, r=1,\dots,R} w(|x^{(1)}|, \dots, |x^{(R)}|) \\
&\qquad \times \max_{\underline{i}} \mathbf{P}_{\underline{i}}\big(\psi^{(r)}(X) \neq \theta^{(r)},\, r = 1, \dots, R\big) \\
&\quad \geqslant \inf_{|x^{(r)}| \geqslant \delta_2^{(r)}:\, r=1,\dots,R} w(|x^{(1)}|, \dots, |x^{(R)}|) \left(1 - \frac{nI((X_1, \xi_1), \theta) + \ln 2}{\ln(\mathcal{L} - 1)}\right),
\end{aligned}$$

where $I((X_1, \xi_1), \theta)$ is the Shannon information, which can be bounded from above as in [4]:

$$\begin{aligned}
&I((X_1, \xi_1), \Theta) \\
&= \mathcal{L}^{-1} \sum_{i_1=1}^{l_1+1} \cdots \sum_{i_R=1}^{l_R+1} \int_G \int_{\mathbf{R}^N} \Bigg(\ln \frac{f(\Sigma^{-1/2}(x)(y - B\underline{D}_{\underline{i}}(x)))}{\mathcal{L}^{-1} \sum_{\underline{j}} f(\Sigma^{-1/2}(x)(y - B\underline{D}_{\underline{j}}(x)))} \\
&\qquad\qquad \times f(\Sigma^{-1/2}(x)(y - B\underline{D}_{\underline{i}}(x)))\Bigg)\, dx\, dy \\
&= \mathcal{L}^{-1} \sum_{i_1=1}^{l_1+1} \cdots \sum_{i_R=1}^{l_R+1} \int_G \int_{\mathbf{R}^N} \Bigg(\ln \frac{f(\widetilde{y} - \Sigma^{-1/2}(x) B\underline{D}_{\underline{i}}(x))}{\mathcal{L}^{-1} \sum_{\underline{j}} f(\widetilde{y} - \Sigma^{-1/2}(x) B\underline{D}_{\underline{j}}(x))} \\
&\qquad\qquad \times f\big(\widetilde{y} - \Sigma^{-1/2}(x) B\underline{D}_{\underline{i}}(x)\big) \det(\Sigma^{1/2}(x))\Bigg)\, dx\, d\widetilde{y}.
\end{aligned}$$

Using the change of variable $\widetilde{y} = \Sigma^{-1/2} y$ on y only and using the concavity of log we can further bound $I((X_1, \xi_1), \Theta)$ from above by

$$\mathcal{L}^{-1} \sum_{i_1=1}^{l_1+1} \cdots \sum_{i_R=1}^{l_R+1} \int_G \int_{\mathbf{R}^N} \biggl(\ln \frac{f(\widetilde{y} - \Sigma^{-1/2}(x) B \underline{D}_{\underline{i}}(x))}{f(\widetilde{y} - \Sigma^{-1/2}(x) B \underline{D}(x))} \\ \times f\bigl(\widetilde{y} - \Sigma^{-1/2}(x) B \underline{D}_{\underline{i}}(x)\bigr) \det(\Sigma^{1/2}(x)) \biggr) \, dx \, d\widetilde{y}.$$

Now we can bound the integral above similarly to the proof of Proposition 2. Hence, we get

$$\begin{aligned} I((X_1, \xi_1), \Theta) &\leqslant \mathcal{L}^{-1} \sum_{i_1=1}^{l_1+1} \cdots \sum_{i_R=1}^{l_R+1} C(f, \delta_1) \\ &\quad \times \int_G |\Sigma^{-1/2}(x) B \underline{D}_{\underline{i}}(x) - \Sigma^{-1/2}(x) B \underline{D}(x)|^2 \det(\Sigma^{1/2}(x)) \, dx \\ &\leqslant \mathcal{L}^{-1} \sum_{i_1=1}^{l_1+1} \cdots \sum_{i_R=1}^{l_R+1} C(f, \delta_1) \\ &\quad \times \int_G \|\Sigma^{-1/2}(x)\|_F^2 \|B\|_F^2 \|\underline{D}_{\underline{i}}(x) - \underline{D}(x)\|_F^2 \det(\Sigma^{1/2}(x)) \, dx \\ &\leqslant \mathcal{L}^{-1} \sum_{i_1=1}^{l_1+1} \cdots \sum_{i_R=1}^{l_R+1} C(f, \delta_1) C_{\Sigma,B} \|D_\theta - D_{\widetilde{\theta}}\|_G^2 \leqslant C(f, \delta_1) C_{\Sigma,B} \eta. \end{aligned}$$

This argument along with Propositions 1 and 2 completes the proof. Lemma 3 is proved.

Further, we define a $(P/4)$-separated net in L_1-norm $\mathcal{B}$ by using $b \in \{0,1\}^P$ similar to [10]. The cardinality of such a set $\mathcal{B}$ is larger than $\exp(P/2)$, see [7], [8]. Recall

$$v_{b,n}^{(r)}(x) = v_0^{(r)}(x) \bigl(1 + n^{-\alpha} \varphi_b^{\prime(r)}(t_b^{(r)}(x)) \psi^{(r)}(n^\gamma |a_b(x) - a_*|)\bigr).$$

Note that by construction, the starting point $a_b(x) \in A_\varepsilon$ does not depend on b, therefore, $a_b(x) = a_0(x)$, for all $b \in \{0,1\}^P$. Moreover,

$$t_0^{(r)}(x) = t_b^{(r)}(x) + n^{-\alpha} \varphi_b^{(r)}(t_b^{(r)}(x)) \psi^{(r)}(n^\gamma |a_0(x) - a_*|). \tag{1}$$

Following the construction of [11], for each $r = 1, \ldots, R$, we can construct $\varphi^{(r)}(t)\colon \mathbf{R} \to [-1,1]$ as a twice continuously differentiable function with support $[0,1]$ such that

$$\varphi^{(r)}(0) = \varphi^{(r)}(1) = 0, \qquad -1 < \varphi^{\prime(r)}(t) \leqslant 1,$$

and

$$\varphi_b^{(r)}(t) = \sum_{i=1}^{P} b_i h \varphi^{(r)}\Big(\frac{t-(i-1)h}{h}\Big).$$

Here $P = P_1 n^\delta$, $h = T/P$, $P_1 > 0$ is a constant, and $\delta \geqslant 0$. We will eventually show that the best choice would be $\delta = 0$. Moreover, suppose the components of $\varphi_b^{(r)}$ are supported on equal length subintervals $[(i-1)h, ih]$, $i = 1, \ldots, P$. It is very important for us now to show that the tensor subclass is indeed in $\mathcal{D}^2(a_*, G, T)$, in particular, its members are twice differentiable. The smoothness condition is a requirement as stated in (A1), which is pivotal to our estimation process of the tensor model of the fiber tracts. In what follows we would introduce Lemma 4 which will establish that the tensors in our parametric subclass are twice continuously differentiable.

Lemma 4. *For all $b \in \mathcal{B}$, we have $D_b \in \mathcal{D}_{c,k}$.*

The proof requires careful term by term differentiation of the tensor D_b in its components and elements, therefore, the proof of this lemma is deferred to last subsection. Now we move on to our next lemma which shows that the difference in the tensors for any two $b \in \mathcal{B}$ is bounded in the integrated L_2-norm.

Lemma 5. *For any $b, \widetilde{b} \in \mathcal{B}$, such that $b \neq \widetilde{b}$ and, for any large enough n, we have*

$$\|\underline{D}_b(x) - \underline{D}_{\widetilde{b}}(x)\|_G^2 \leqslant C n^{-(d-1)\gamma + 2\alpha},$$

where $C > 0$ is a constant.

Proof. We have $D_b(x) = \sum_{r=1}^R \lambda_r v_{b,n}^{(r)}(x)^{\otimes M}$. For fixed $b, \widetilde{b} \in \mathcal{B}$, we have

$$\begin{aligned}
&\|D_b(x) - D_{\widetilde{b}}(x)\|_G^2 \\
&= \int_G \bigg\| \sum_{r=1}^R \Big(\lambda_r v_0^{(r)}(x)^{\otimes M} \big(1 + n^{-\alpha} \varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big) \psi^{(r)}(n^\gamma |a_0(x) - a_*|)\big)^M \\
&\qquad - \lambda_r v_0^{(r)}(x)^{\otimes M} \big(1 + n^{-\alpha} \varphi_{\widetilde{b}}^{\prime(r)}\big(t_{\widetilde{b}}^{(r)}(x)\big) \psi^{(r)}(n^\gamma |a_0(x) - a_*|)\big)^M \Big) \bigg\|_F^2 dx \\
= &\int_G \bigg\| \sum_{r=1}^R \lambda_r v_0^{(r)}(x)^{\otimes M} \Big(\big(1 + n^{-\alpha} \varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big) \psi^{(r)}(n^\gamma |a_0(x) - a_*|)\big)^M \\
&\qquad - \big(1 + n^{-\alpha} \varphi_{\widetilde{b}}^{\prime(r)}\big(t_{\widetilde{b}}^{(r)}(x)\big) \psi^{(r)}(n^\gamma |a_0(x) - a_*|)\big)^M \Big) \bigg\|_F^2 dx \\
&\leqslant \int_G \sum_{r=1}^R \lambda_r^2 \|v_0^{(r)}(x)^{\otimes M}\|_F^2 \Big(\big(1 + n^{-\alpha} \varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big) \psi^{(r)}(n^\gamma |a_0(x) - a_*|)\big)^M \\
&\qquad - \big(1 + n^{-\alpha} \varphi_{\widetilde{b}}^{\prime(r)}\big(t_{\widetilde{b}}^{(r)}(x)\big) \psi^{(r)}(n^\gamma |a_0(x) - a_*|)\big)^M \Big)^2 dx
\end{aligned}$$

$$= \int_G \sum_{r=1}^{R} \lambda_r^2 \big\| v_0^{(r)}(x)^{\otimes M} \big\|_F^2$$
$$\times \big[\big(Mn^{-\alpha}\psi^{(r)}(n^\gamma |a_0(x)-a_*|)\big)\big(\varphi_b^{'(r)}\big(t_b^{(r)}(x)\big) - \varphi_{\widetilde b}^{'(r)}\big(t_{\widetilde b}^{(r)}(x)\big)\big)\big) C_{\varphi,\psi}\big]^2\, dx$$
$$\big(\text{where the sequence } C_{\varphi,\psi}^{(n)} = 1 + O(n^{-2\alpha})\big)$$
$$= Cn^{-2\alpha} M^2 \sum_{r=1}^{R} \Big(\lambda_r^2 \sup_{x\in G} \big\| v_0^{(r)}(x)^{\otimes M} \big\|_F^2\Big)$$
$$\times \int_G \big(\psi^{(r)}(n^\gamma |a_0(x)-a_*|)\big)^2 \big(\varphi_b^{'(r)}\big(t_b^{(r)}(x)\big) - \varphi_{\widetilde b}^{'(r)}\big(t_{\widetilde b}^{(r)}(x)\big)\big)^2\, dx. \qquad (2)$$

In the above expression we use the simple identity

$$(1+u)^M - (1+v)^M = (u-v)\left(M + \binom{M}{2}(u+v) + \binom{M}{3}(u^2+uv+v^2) + \cdots\right)$$

and the fact that functions φ, ψ are bounded. Hence, we can find a bounded sequence $C_{\varphi,\psi}^{(n)}$ with which we can bound the $o(n^{-\alpha})$ terms in the expression (2). Now for each $r = 1, \dots, R$, integrals

$$\int_G \big(\psi^{(r)}(n^\gamma |a_0(x)-a_*|)\big)^2 \big(\varphi_b^{'(r)}\big(t_b^{(r)}(x)\big) - \varphi_{\widetilde b}^{'(r)}\big(t_{\widetilde b}^{(r)}(x)\big)\big)^2\, dx$$

can be bounded by $C_r n^{-(d-1)\gamma}$ using Lemma 4 in [11], where $C_r > 0$ is some generic constant depending on r. Hence, we conclude that for some generic $C > 0$

$$\|\underline{D}_b(x) - \underline{D}_{\widetilde b}(x)\|_G^2 \leqslant Cn^{-2\alpha-(d-1)\gamma}.$$

Lemma 5 is proved.

The following lemma shows that the curves driven by the pseudo-eigenvectors are separated in L_1-norm inside the class indexed by $\mathcal{B}$.

Lemma 6. *For each $r = 1, \dots, R$, for any $b, \widetilde b \in \mathcal{B}$ such that $b \neq \widetilde b$, and any large enough n, we have*

$$\big\| x_b^{(r)} - x_{\widetilde b}^{(r)} \big\|_{1,T} \geqslant Cn^{-\alpha-\delta}$$

with some $C > 0$ depending on r.

Proof. Consider the following relations:

$$\big\| x_b^{(r)}(\cdot, a_*) - x_{\widetilde b}^{(r)}(\cdot, a_*) \big\|_{1,T}$$
$$= \int_0^T \big\| x_0^{(r)}\big(t + n^{-\alpha}\varphi_b^{(r)}(t)\psi(0), a_*\big) - x_0^{(r)}\big(t + n^{-\alpha}\varphi_{\widetilde b}^{(r)}(t)\psi(0), a_*\big) \big\|_1\, dt$$
$$= \int_0^T \Bigg(\bigg\| n^{-\alpha}\big(\varphi_b^{(r)}(t) - \varphi_{\widetilde b}^{(r)}(t)\big)\psi(0)$$
$$\times \int_0^1 v_0^{(r)}\big(x\big(t + n^{-\alpha}\psi(0)\big(\pi\varphi_b^{(r)}(t) + (1-\pi)\varphi_{\widetilde b}^{(r)}(t)\big), a_*\big)\big)\, d\pi \bigg\|_1 \Bigg)\, dt$$

$$
\begin{aligned}
&= n^{-\alpha}\psi(0)\int_0^T\Bigl(\bigl|\varphi_b^{(r)}(t)-\varphi_{\widetilde b}^{(r)}(t)\bigr| \\
&\quad\times\Bigl\|\int_0^1 v_0^{(r)}\bigl(x\bigl(t+n^{-\alpha}\psi(0)\bigl(\pi\varphi_b^{(r)}(t)+(1-\pi)\varphi_{\widetilde b}^{(r)}(t)\bigr),a_*\bigr)\bigr)\,d\pi\Bigr\|_1\Bigr)\,dt,
\end{aligned}
$$

for any $b,\widetilde b\in\mathcal{B}, b\neq\widetilde b$, which is bounded from below by

$$
\begin{aligned}
&n^{-\alpha}\psi(0)\Bigl[\Bigl(\inf_{t\in[0,T]}\Bigl\|\int_0^1 v_0^{(r)}\bigl(x\bigl(t+n^{-\alpha}\psi(0)\bigl(\pi\varphi_b^{(r)}(t) \\
&\qquad\qquad +(1-\pi)\varphi_{\widetilde b}^{(r)}(t)\bigr),a_*\bigr)\bigr)\,d\pi\Bigr\|_1\Bigr)\int_0^T|\varphi_b^{(r)}(t)-\varphi_{\widetilde b}^{(r)}(t)|\,dt\Bigr] \\
&\quad= n^{-\alpha}\psi(0)C_r\int_0^T\bigl|\varphi_b^{(r)}(t)-\varphi_{\widetilde b}^{(r)}(t)\bigr|\,dt.
\end{aligned}
$$

In order to complete the proof of this lemma we need to show that

$$
\inf_{t\in[0,T]}\Bigl\|\int_0^1 v_0^{(r)}\bigl(x\bigl(t+n^{-\alpha}\psi(0)\bigl(\pi\varphi_b^{(r)}(t)+(1-\pi)\varphi_{\widetilde b}^{(r)}(t)\bigr),a_*\bigr)\bigr)\,d\pi\Bigr\|_1>0. \quad (3)
$$

Let us suppose there exists an n_0 such that, for all $n\geqslant n_0$,

$$
\inf_{t\in[0,T]}\Bigl\|\int_0^1 v_0^{(r)}\bigl(x\bigl(t+n^{-\alpha}\psi(0)\bigl(\pi\varphi_b^{(r)}(t)+(1-\pi)\varphi_{\widetilde b}^{(r)}(t)\bigr),a_*\bigr)\bigr)\,d\pi\Bigr\|_1=0.
$$

Then there exists $t_0\in[0,T]$ such that, for $n\geqslant n_0$,

$$
\int_0^1 v_0^{(r)}\bigl(x\bigl(t_0+n^{-\alpha}\psi(0)\bigl(\pi\varphi_b^{(r)}(t_0)+(1-\pi)\varphi_{\widetilde b}^{(r)}(t_0)\bigr),a_*\bigr)\bigr)\,d\pi=0. \quad (4)
$$

Using the change of variable $t_0+n^{-\alpha}\psi(0)(\pi\varphi_b^{(r)}(t_0)+(1-\pi)\varphi_{\widetilde b}^{(r)}(t_0))=u$, we can write (4) as

$$
\begin{aligned}
&\int_{t_0+n^{-\alpha}\psi(0)\varphi_{\widetilde b}^{(r)}(t_0)}^{t_0+n^{-\alpha}\psi(0)\varphi_b^{(r)}(t_0)} v_0^{(r)}(x(u,a_*))\frac{du}{n^{-\alpha}\psi(0)(\varphi_b^{(r)}(t_0)\varphi_{\widetilde b}^{(r)}(t_0))}=0 \\
&\implies\quad v_0^{(r)}\bigl(x\bigl(t_{0,n}+n^{-\alpha}\psi(0)\bigl(\pi_n\varphi_b^{(r)}(t_{0,n})+(1-\pi_n)\varphi_{\widetilde b}^{(r)}(t_{0,n})\bigr)\bigr)\bigr)=0,
\end{aligned}
$$

where $\{t_{0,n}\}_{n\geqslant1}$, $t_{0,n}\to t_0$ and $\pi_n\in[0,1]$. This would imply that there exists $x\in\mathcal{X}_0\subset G$ such that $v_0^{(r)}(x)=0$, $x\in\mathcal{X}_0$, which violates the assumption that we have imposed in (A1) and (A3). Therefore, (3) holds true. Now by following the proof of Lemma 5 in [11] the integral $\int_0^T|\varphi_b^{(r)}(t)-\varphi_{\widetilde b}^{(r)}(t)|\,dt$ can be shown to be bounded from below by $Cn^{-\delta}$, where C is a generic constant depending on r. Hence, we can conclude $\|x_b^{(r)}-x_{\widetilde b}^{(r)}\|_{1,T}\geqslant Cn^{-\alpha-\delta}$. The lemma is proved.

Going back to the expression in Theorem 1 using Lemma 3 we get

$$\begin{aligned}
&\inf_{\widehat{X}_n^{(1)},\dots,\widehat{X}_n^{(R)}\in\mathcal{E}_n(a_*,T)}\ \sup_{D\in\mathcal{D}_{c,k}} \mathbf{E}w\big(n^{2/(d+3)}\big(\big\|\widehat{X}_n^{(1)}-x^{(1)}\big\|_{p,T},\\
&\qquad\qquad\qquad\qquad\dots,\big\|\widehat{X}_n^{(R)}-x^{(R)}\big\|_{p,T}\big)\big)\\
&\geqslant \inf_{\widehat{X}_n^{(1)},\dots,\widehat{X}_n^{(R)}\in\mathcal{E}_n(a_*,T)}\ \sup_{D\in\mathcal{D}_{\mathcal{L}}} \mathbf{E}w\big(n^{2/(d+3)}\big(\big\|\widehat{X}_n^{(1)}-x^{(1)}\big\|_{p,T},\\
&\qquad\qquad\qquad\qquad\dots,\big\|\widehat{X}_n^{(R)}-x^{(R)}\big\|_{p,T}\big)\big)\\
&\geqslant \inf_{|x^{(r)}|\geqslant\delta_2^{(r)}:\ r=1,\dots,R} w_n(|x^{(1)}|,\dots,|x^{(R)}|)\left(1-\frac{nC_{\Sigma,B}C(f,\delta_1)\eta+\ln 2}{\ln(\mathcal{L}-1)}\right). \qquad (5)
\end{aligned}$$

In the above expression we substitute $C_{\Sigma,B}C(f,\delta_1)=C>0$ as a generic constant and choose $\eta=Cn^{1-2\alpha-(d-1)\gamma}$, $\delta_2^{(r)}=C/2n^{-\alpha-\delta}$. Also it can be shown that $\mathcal{L}\geqslant\exp(P/2)$, see [7], [8], hence by an algebraic manipulation we get $\ln(\mathcal{L}-1)\geqslant P/2-1$. Therefore, we can rewrite (5) as

$$\begin{aligned}
&\inf_{\widehat{X}_n^{(1)},\dots,\widehat{X}_n^{(R)}\in\mathcal{E}_n(a_*,T)}\ \sup_{D\in\mathcal{D}_{c,k}} \mathbf{E}w\big(n^{2/(d+3)}\big(\big\|\widehat{X}_n^{(1)}-x^{(1)}\big\|_{p,T},\\
&\qquad\qquad\qquad\qquad\dots,\big\|\widehat{X}_n^{(R)}-x^{(R)}\big\|_{p,T}\big)\big)\\
&= \inf_{|x^{(r)}|\geqslant C/2n^{-\alpha-\delta}:\ r=1,\dots,R} w_n(|x^{(1)}|,\dots,|x^{(R)}|)\left(1-\frac{n^{1-2\alpha-(d-1)\gamma}C^2+\ln 2}{P/2-1}\right)\\
&> \frac{1}{2}\inf_{|x^{(r)}|\geqslant(1/2)Ch:\ r=1,\dots,R} w(|x^{(1)}|,\dots,|x^{(R)}|)>0, \qquad (6)
\end{aligned}$$

where $w_n(|x^{(1)}|,\dots,|x^{(R)}|)=w(n^{2/(d+3)}(|x^{(1)}|,\dots,|x^{(R)}|))$. Note that while obtaining (6) we have chosen $\alpha=2/(d+3)$, $\gamma=\alpha/2$ and $\delta=0$. Also we can choose P_1 sufficiently large, where $P=P_1n^\delta$, introduced before such that $C^2<(P_1-2)/4-\ln 2$, which completes the proof of Theorem 1.

2.2. Proof of Theorem 2. To prove part (a), first fix r in $1,\dots,R$. Then using Jensen's inequality, we can write

$$\begin{aligned}
\mathbf{E}\big\|\widehat{X}_n^{(r)}-x^{(r)}\big\|_{p,T} &= \mathbf{E}\left(\int_0^T \big\|\widehat{X}_n^{(r)}(t)-x^{(r)}(t)\big\|_p^p\,dt\right)^{1/p}\\
&\leqslant \left(\int_0^T \mathbf{E}\big\|\widehat{X}_n^{(r)}(t)-x^{(r)}(t)\big\|_p^p\,dt\right)^{1/p}. \qquad (7)
\end{aligned}$$

Recall that $\underline{D}$ denotes the vectorized representation of the super-symmetric tensor D. Using this notation and the definition of an integral curve estimator we can write

$$\widehat{X}_n^{(r)}(t) - x^{(r)}(t) = \int_0^t \big(v^{(r)}\big(\underline{\widehat{D}}_n\big(\widehat{X}_n^{(r)}(s)\big)\big) - v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)\big)\,ds$$
$$= \int_0^t \big(v^{(r)}\big(\underline{D}\big(\widehat{X}_n^{(r)}(s)\big)\big) - v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)\big)\,ds$$
$$+ \int_0^t \big(v^{(r)}\big(\underline{D}\big(\widehat{X}_n^{(r)}(s)\big)\big) - v^{(r)}\big(\underline{D}\big(\widehat{X}_n^{(r)}(s)\big)\big)\big)\,ds$$
$$= \int_0^t \nabla v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)\nabla \underline{D}(x^{(r)}(s))\big(\widehat{X}_n^{(r)}(s) - x^{(r)}(s)\big)\,ds$$
$$+ \int_0^t \nabla v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)\big(\underline{\widehat{D}}_n(x^{(r)}(s)) - \underline{D}(x^{(r)}(s))\big)\,ds + R_n^{(r)}(t),$$

where the remainder term $R_n^{(r)}(t)$ is given by

$$R_n^{(r)}(t) = \int_0^t \big[v^{(r)}\big(\underline{D}\big(\widehat{X}_n^{(r)}(s)\big)\big) - v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)$$
$$- \nabla v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)\nabla \underline{D}(x^{(r)}(s))\big(\widehat{X}_n^{(r)}(s) - x^{(r)}(s)\big)\big]\,ds$$
$$+ \int_0^t \big[v^{(r)}\big(\underline{D}\big(\widehat{X}_n^{(r)}(s)\big)\big) - v^{(r)}\big(\underline{D}\big(\widehat{X}_n^{(r)}(s)\big)\big) - \nabla v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)$$
$$\times \big(\underline{\widehat{D}}_n(x^{(r)}(s)) - \underline{D}(x^{(r)}(s))\big)\big]\,ds.$$

Now similar to [9] we can decompose

$$\widehat{X}_n^{(r)}(s) - x^{(r)}(s) = Z_n^{(r)}(s) + \delta_n^{(r)}(s), \quad \text{for each} \ \ r = 1, \dots, R.$$

Then we can write

$$\widehat{X}_n^{(r)}(t) - x^{(r)}(t) = \int_0^t \nabla v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)\nabla \underline{D}(x^{(r)}(s))\delta_n^{(r)}(s)\,ds + R_n^{(r)}(t)$$
$$+ \int_0^t \nabla v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)\big(\underline{\widehat{D}}_n - \underline{D}\big)(x^{(r)}(s))\,ds$$
$$+ \int_0^t \nabla v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)\nabla \underline{D}(x^{(r)}(s))Z_n^{(r)}(s)\,ds.$$

Here the stochastic processes $\delta_n^{(r)}(t), Z_n^{(r)}(t)$, for any $t \in [0, T]$, are represented by

$$\delta_n^{(r)}(t) = \int_0^t \nabla v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)\nabla \underline{D}(x^{(r)}(s))\delta_n^{(r)}(s)\,ds + R_n^{(r)}(t),$$
$$Z_n^{(r)}(t) = \int_0^t \nabla v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)\big(\underline{\widehat{D}}_n - \underline{D}\big)(x^{(r)}(s))\,ds$$
$$+ \int_0^t \nabla v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)\nabla \underline{D}(x^{(r)}(s))Z_n^{(r)}(s)\,ds.$$

Let us denote $\nabla v^{(r)}(\underline{D}(x^{(r)}(s)))\nabla\underline{D}(x^{(r)}(s)) = v'^{(r)}(x^{(r)}(s))$. Then $Z_n^{(r)}(t)$ satisfies the equation

$$\frac{dZ_n^{(r)}(t)}{dt} = \nabla v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)\big(\widehat{\underline{D}}_n - \underline{D}\big)(x^{(r)}(s)) + Z_n^{(r)}(t), \qquad Z_n^{(r)}(0) = 0. \tag{8}$$

The solution to the stochastic differential equation in (7) can be represented by the d-dimensional random process

$$Z_n^{(r)}(s) = \int_0^t U^{(r)}(t,s)\nabla v^{(r)}\big(\underline{D}(x^{(r)}(s))\big)\big(\widehat{\underline{D}}_n - \underline{D}\big)(x^{(r)}(s))\,ds, \qquad t\in[0,T],$$

where $U^{(r)}\colon \mathbf{R}^2 \to \mathbf{R}^{d^2}$ is the Green function defined as the solution of the PDE given by

$$\frac{\partial U^{(r)}(t,s)}{\partial t} = \nabla v^{(r)}\big(\underline{D}(x^{(r)}(t))\big)\nabla\underline{D}(x^{(r)}(t))U^{(r)}(t,s), \qquad U^{(r)}(s,s) = I_{d\times d}.$$

Now from (7) we have

$$\mathbf{E}\big\|\widehat{X}_n^{(r)}(t) - x^{(r)}(t)\big\|_p^p \leqslant \mathbf{E}\big(\|Z_n^{(r)}(t) - \mathbf{E}Z_n^{(r)}(t)\|_p + \|\mathbf{E}Z_n^{(r)}(t)\|_p + \|\delta_n^{(r)}(t)\|_p\big)^p. \tag{9}$$

Since

$$\widehat{\underline{D}}_n(x) = \frac{1}{nh_n^d}\sum_{j=1}^n K\left(\frac{x - X_j}{h_n}\right)\widetilde{\underline{D}}(X_j),$$

where $\widetilde{\underline{D}}(X_j)$ is the LSE estimator of $\underline{D}(X_j)$, we can write

$$\begin{aligned}\widetilde{\underline{D}}(X_j) &= (B^\top B)^{-1}B^\top Y(X_j) = (B^\top B)^{-1}B^\top(B\underline{D}(X_j) + \Sigma^{1/2}(X_j)\Xi_j)\\ &= \underline{D}(X_j) + \Gamma_j, \qquad \Gamma_j = (B^\top B)^{-1}B^\top\Sigma^{1/2}(X_j)\Xi_j.\end{aligned}$$

Now let us denote $Z_n^{(r)}(t) = \sum_{i=1}^n \chi_i^{(r)}(t)$, where

$$\begin{aligned}\chi_i^{(r)}(t) &= \int f_t^{(r)}(s)K\left(\frac{x^{(r)}(s) - X_j}{h_n}\right)ds\,(\underline{D}(X_i) + \Gamma_j),\\ f_t^{(r)}(s) &= I_{[0,t]}(s)U^{(r)}(t,s)\nabla v^{(r)}\big(\underline{D}(x^{(r)}(s))\big).\end{aligned}$$

Then using similar arguments from [2] we can derive

$$\begin{aligned}\mathbf{E}Z_n^{(r)}(t) = {}&-h_n\int f_t^{(r)}(s)\nabla\underline{D}(x^{(r)}(s))\int uK(u)\,du\,ds\\ &+\frac{1}{2}h_n^2\int f_t^{(r)}(s)\int K(u)\langle\nabla^2\underline{D}(x^{(r)}(s))u,u\rangle\,du\,ds + o(h_n^2).\end{aligned}$$

Since we select the bandwidth $h_n = n^{-1/(d+3)}$, we get

$$\mathbf{E}Z_n^{(r)}(t) = \frac{\mu_\beta^{(r)}(t) + o_P(1)}{\sqrt{nh_n^{d-1}}},$$

where $\mu_\beta^{(r)}(t)$ is the mean of the limiting Gaussian process $\mathcal{G}^{(r)}(t)$ for integral curve estimator $\widehat{X}_n^{(r)}(t)$ and β is the constant that has been introduced in (A8). Hence,

$$\left\|\frac{\mu_\beta^{(r)}(t)}{\sqrt{nh_n^{d-1}}}\right\|_p = n^{-2/(d+3)}\left\|\mu_\beta^{(r)}(t)\right\|_p.$$

Also notice that using similar arguments from [2] we get $\sup_{0\leqslant t\leqslant T}|\delta_n^{(r)}(t)| = o_P((nh_n^{d-1})^{-1/2})$. Now applying Rosenthal's inequality for independent mean zero random variables from [5] to $(\chi_i^{(r)}(t) - \mathbf{E}\chi_i^{(r)}(t))/nh_n^d$ with finite pth moments, we obtain

$$\mathbf{E}\left\|\sum_{i=1}^n \frac{\chi_i^{(r)}(t) - \mathbf{E}\chi_i^{(r)}(t)}{nh_n^d}\right\|_p^p \leqslant \sum_{k=1}^d C^{(r)}(p)\max\left\{\sum_{i=1}^n \mathbf{E}\left|\frac{\chi_{i,k}^{(r)}(t) - \mathbf{E}\chi_{i,k}^{(r)}(t)}{nh_n^d}\right|^p,\right.$$
$$\left.\left(\sum_{i=1}^n \mathbf{E}\left(\frac{\chi_{i,k}^{(r)}(t) - \mathbf{E}\chi_{i,k}^{(r)}(t)}{nh_n^d}\right)^2\right)^{p/2}\right\},$$

where $C^{(r)}(p) > 0$ depends on p and r only. Now we can bound the pth moment of $\chi_{i,k}^{(r)}$ using the similar arguments from [9]. To this goal we have

$$\begin{aligned}
&\mathbf{E}\left|\frac{\chi_{i,k}^{(r)}(t) - \mathbf{E}\chi_{i,k}^{(r)}(t)}{nh_n^d}\right|^p \\
&\leqslant C(nh_n^d)^{-p}\mathbf{E}\bigg|\int_{\mathbf{R}} f_t^{(r)}(s_1)K\left(\frac{x^{(r)}(s_1) - X_i}{h_n}\right)ds_1\,(\underline{D}_k(X_i) + \Gamma_{j,k}) \\
&\times\cdots\times\int_{\mathbf{R}} f_t^{(r)}(s_p)K\left(\frac{x^{(r)}(s_p) - X_i}{h_n}\right)ds_p\,(\underline{D}_k(X_i) + \Gamma_{j,k})\bigg| \\
&\leqslant C'(nh_n^d)^{-p}\int_{\mathbf{R}^{d+p}}\bigg(|f_t^{(r)}(s_1)|\cdots|f_t^{(r)}(s_p)| \\
&\times K\left(\frac{x^{(r)}(s_1) - y}{h_n}\right)\cdots K\left(\frac{x^{(r)}(s_p) - y}{h_n}\right)\bigg)\,ds_1\cdots ds_p\,dy \\
&= C'h_n^d(nh_n^d)^{-p}\int_{\mathbf{R}^{d+p}}\bigg(|f_t^{(r)}(s_1)|\cdots|f_t^{(r)}(s_p)|K(z) \\
&\times K\left(z + \frac{x^{(r)}(s_2) - x^{(r)}(s_1)}{h_n}\right)\cdots K\left(z + \frac{x^{(r)}(s_p) - x^{(r)}(s_1)}{h_n}\right)\bigg)\,ds_1\cdots ds_p\,dz,
\end{aligned}$$

which is

$$
\begin{aligned}
&C' h_n^{d+p-1} (nh_n^d)^{-p} \int_{\mathbf{R}^{d+p}} \biggl(|f_t^{(r)}(s_1)| |f_t^{(r)}(s_1+\tau_2 h)| \cdots |f_t^{(r)}(s_1+\tau_p h)| K(z) \\
&\qquad \times K\biggl(z + \tau_2 \frac{x^{(r)}(s_1+\tau_2 h) - x^{(r)}(s_1)}{\tau_2 h_n} \biggr) \\
&\qquad \times \cdots \times K\biggl(z + \tau_p \frac{x^{(r)}(s_1+\tau_p h) - x^{(r)}(s_1)}{\tau_p h_n} \biggr) \biggr) \, dz \, ds_1 \, d\tau_2 \cdots d\tau_p \\
&\leqslant C' h_n^{d+p-1} (nh_n^d)^{-p} \int_{\mathbf{R}^{p-1}} \widetilde{\Lambda}(\tau_2, \dots, \tau_p) \biggl(\int |f_t^{(r)}(s_1)|^p \, ds_1 \biggr)^{1/p} \\
&\qquad \times \cdots \times \biggl(\int |f_t^{(r)}(s_1+\tau_2 h_n)|^p \, ds_1 \biggr)^{1/p} \\
&\qquad \times \biggl(\int |f_t^{(r)}(s_1+\tau_p h_n)|^p \, ds_1 \biggr)^{1/p} d\tau_2 \cdots d\tau_p \\
&\leqslant C'' h_n^{d+p-1} (nh_n^d)^{-p} \int_{\mathbf{R}^{p-1}} \widetilde{\Lambda}(\tau_2, \dots, \tau_p) \biggl(\int |f_t^{(r)}(s_1)|^p \, ds_1 \biggr) \\
&\leqslant C_1 h_n^{d+p-1} (nh_n^d)^{-p},
\end{aligned}
$$

where

$$
\begin{aligned}
\widetilde{\Lambda}(\tau_2, \dots, \tau_p) = \sup_{s_1, \dots, s_p \in [0,\tau]} &\int_{\mathbf{R}^d} K(z) K\biggl(z + \tau_2 \frac{x^{(r)}(s_2) - x^{(r)}(s_1)}{s_2 - s_1} \biggr) \\
&\times \cdots \times K\biggl(z + \tau_p \frac{x^{(r)}(s_p) - x^{(r)}(s_1)}{s_p - s_1} \biggr) \, dz
\end{aligned}
$$

is analogous to $\Lambda(\tau_1, \tau_2, \tau_3)$ in [2]. Recall that $\underline{D}_k$ is the kth component of vectorized D. Also C_1, C, C', and C'' are positive constants depending on B, D, $U^{(r)}$. Thus, from (9), we obtain, by replacing $h_n = n^{-1/(d+3)}$,

$$
\begin{aligned}
\mathbf{E} \biggl\| \sum_{i=1}^n \frac{\chi_i^{(r)}(t) - \mathbf{E}\chi_i^{(r)}(t)}{nh_n^d} \biggr\|_p^p &\leqslant dC^{(r)}(p) \max \biggl\{ \frac{C_1 n h_n^{d+p-1}}{(nh_n^d)^p}, \biggl(\frac{C_2 n h_n^{d+1}}{(nh_n^d)^2} \biggr)^{p/2} \biggr\} \\
&= dC^{(r)}(p) n^{-3p/(d+3)} \max\bigl\{ C_1' n^{(4-p)/(d+3)}, C_2' n^{p/(d+3)} \bigr\} \\
&= dC^{(r)}(p) C_2' n^{-2p/(d+3)}, \qquad (10)
\end{aligned}
$$

where C_1', C_2' are positive constants. As a result, we have

$$
\begin{aligned}
\mathbf{E} \bigl\| \widehat{X}_n^{(r)} - x^{(r)} \bigr\|_p^p &\leqslant n^{-2p/(d+3)} b\bigl(\bigl\| \mu_\beta^{(r)}(t) \bigr\|_p \bigr), \\
\mathbf{E} \bigl\| \widehat{X}_n^{(r)} - x^{(r)} \bigr\|_{p,T} &\leqslant n^{-2/(d+3)} \biggl(\int_0^T b\bigl(\bigl\| \mu_\beta^{(r)}(t) \bigr\|_p \bigr) \, dt \biggr)^{1/p},
\end{aligned}
$$

for some bounded and continuous function $b\colon [0,\infty) \to [0,\infty)$. This concludes the proof of part (a) in Theorem 2.

To prove part (b), we first notice that we can write

$$\begin{aligned}
&\mathbf{E} \sup_{t\in[0,T]} \max_{1\leqslant k\leqslant d} \bigl|\widehat{X}^{(r)}_{n,k}(t) - x^{(r)}_k(t)\bigr| \\
&\leqslant \|U^{(r)}\|_\infty \mathbf{E} \sup_{t\in[0,T]} \max_{1\leqslant k\leqslant d} \left|\sum_{i=1}^n \frac{\chi^{(r)}_i(t) - \mathbf{E}\chi^{(r)}_i(t)}{nh_n^d}\right| + \sup_{t\in[0,T]} \max_{1\leqslant k\leqslant d} \frac{|\mu^{(r)}_{\beta,k}(t)|}{\sqrt{nh_n^d}} \\
&\quad + \mathbf{E} \sup_{t\in[0,T]} \max_{1\leqslant k\leqslant d} \bigl|\delta^{(r)}_k(t)\bigr|.
\end{aligned}$$

We already know that $\sup_{t\in[0,T]} |\delta^{(r)}(t)| = o_P((nh_n^{d-1})^{-1/2})$. Now using a maximal inequality for $\mathcal{L}_2$-norm we can bound

$$\begin{aligned}
&\mathbf{E} \sup_{t\in[0,T]} \max_{1\leqslant k\leqslant d} \left|\sum_{i=1}^n \frac{\chi^{(r)}_i(t) - \mathbf{E}\chi^{(r)}_i(t)}{nh_n^d}\right| \\
&\qquad \leqslant \sqrt{d} \max_{1\leqslant k\leqslant d} \left(\mathbf{E} \sup_{t\in[0,T]} \left|\sum_{i=1}^n \frac{\chi^{(r)}_i(t) - \mathbf{E}\chi^{(r)}_i(t)}{nh_n^d}\right|^2\right)^{1/2}.
\end{aligned}$$

Let us denote $W^{(r)}(t) = \sum_{i=1}^n \bigl(\chi^{(r)}_i(t) - \mathbf{E}\chi^{(r)}_i(t)\bigr)/(nh_n^d)$, $t \in [0,T]$. Then repeating the first few steps of part (a) with $p = 2$ and

$$f^{(r)}_t(s) = I_{[0,t]}(s) U^{(r)}(t,s) \nabla v^{(r)}\bigl(\underline{D}(x^{(r)}(s))\bigr), \qquad 0 \leqslant t_1 < t_2 \leqslant T,$$

we get

$$\bigl(\mathbf{E}|W^{(r)}(t_1) - W^{(r)}(t_2)|^2\bigr)^{1/2} \leqslant C^{(r)} n^{-2/(d+3)} |t_2 - t_1|. \tag{11}$$

Observe that the diameter of $[0,T]$ in the metric

$$m(t_1,t_2) = C^{(r)} n^{-2/(d+3)} |t_2 - t_1|$$

is $\operatorname{diam}([0,T],m) = TC^{(r)} n^{-2/(d+3)}$ and the maximal number of ε-separated points in $[0,T]$ with respect to metric m is $\operatorname{Dist}(\varepsilon,m) = TC^{(r)} n^{-2/(d+3)}/\varepsilon$. Now using the inequality on p. 100 in [12] yields

$$\begin{aligned}
&\Bigl(\mathbf{E} \sup_{t\in[0,T]} |W^{(r)}(t)|^2\Bigr)^{1/2} \\
&\qquad \leqslant \bigl(\mathbf{E}|W^{(r)}(t_0)|^2\bigr)^{1/2} + K^{(r)} \int_0^{\operatorname{diam}([0,T],m)} \operatorname{Dist}^{1/2}(\varepsilon,m)\, d\varepsilon
\end{aligned}$$

and we can write

$$\begin{aligned}
\int_0^{\operatorname{diam}([0,T],m)} \operatorname{Dist}^{1/2}(\varepsilon,m)\, d\varepsilon &= \int_0^{TC^{(r)} n^{-2/(d+3)}} \left(\frac{TC^{(r)} n^{-2/(d+3)}}{\varepsilon}\right)^{1/2} d\varepsilon \\
&= 2TC^{(r)} n^{-2/(d+3)}.
\end{aligned}$$

On the other hand, from (10) we have $(\mathbf{E}|W^{(r)}(t_0)|^2)^{1/2} \leqslant C^{(r)} n^{-2/(d+3)}$ for some generic constant $C^{(r)} > 0$. Then we have the bound

$$\Big(\mathbf{E} \sup_{t \in [0,T]} |W^{(r)}(t)|^2\Big)^{1/2} \leqslant C^{(r)} n^{-2/(d+3)}.$$

As a result finally we get

$$\mathbf{E} \sup_{t \in [0,T]} \max_{1 \leqslant k \leqslant d} |\widehat{X}_{n,k}^{(r)}(t) - x_k^{(r)}(t)| \leqslant C^{(r)} n^{-2/(d+3)},$$

for some generic $C^{(r)} > 0$, which concludes the proof of Theorem 2.

3. Proof of Lemma 1.

Proof. In order to prove Lemma 1, let us first assume n be sufficiently large so that all the terms with $o(n^{-\alpha})$ can be ignored and all assertions hold up to terms of order $o(n^{-\alpha})$. Then note that, for $r = R$,

$$\begin{aligned}
&\mathcal{T}\big(v_{b,n}^{(R)}(x), D_b^{(R)}(x)\big)_{km} = \mathcal{T}\big(v_0^{(R)}(x), D_0^{(R)}(x)\big)_{km} \\
&= \lambda_R(M-1) \sum_{i_3=1}^{d} \cdots \sum_{i_M=1}^{d} v_{0k}^{(R)}(x) v_{0m}^{(R)}(x) \big(v_{0i_3}^{(R)}(x)\big)^2 \cdots \big(v_{0i_M}^{(R)}(x)\big)^2 - \lambda_R \delta_{km} \\
&= \lambda_R[(M-1) v_{0k} v_{0m} - \delta_{km}].
\end{aligned}$$

To show $\ker(\mathcal{T}(v_0^{(R)}(x), D_0^{(R)}(x)) - \lambda_R I) = 0$, we consider any arbitrary $u \in \mathbf{R}^d$ such that

$$\begin{aligned}
&\big\langle \lambda_R\big[(M-1) v_{0k}^{(R)} v_{0m}^{(R)} - \delta_{km}\big], u\big\rangle = 0 \\
&\Longrightarrow \quad \sum_{m=1}^{d} (M-1) v_{0k}^{(R)} v_{0m}^{(R)} u_m = \sum_{m=1}^{d} \delta_{km} u_m \\
&\Longrightarrow \quad \sum_{k=1}^{d} \sum_{m=1}^{d} (M-1) (v_{0k}^{(R)})^2 v_{0m}^{(R)} u_m = \sum_{k=1}^{d} v_{0k}^{(R)} u_k \\
&\Longrightarrow \quad (M-1) \sum_{m=1}^{d} v_{0m}^{(R)} u_m = \sum_{k=1}^{d} v_{0k}^{(R)} u_k \quad \Longrightarrow \quad \sum_{m=1}^{d} v_{0m}^{(R)} u_m = 0.
\end{aligned}$$

Therefore, it implies that, if $u_k = 0, k = 1, \ldots, d$, then

$$\ker\big(\mathcal{T}\big(v_0^{(R)}(x), D_0^{(R)}(x)\big) - \lambda_R I\big) = 0.$$

Similarly,

$$
\begin{aligned}
&\big(\mathcal{T}\big(v_0^{(p)}(x), D_0^{(p)}(x)\big) - \lambda_p I\big)_{km} \\
&= (M-1)\sum_{r=p}^{R}\lambda_r \sum_{i_3=1}^{d}\cdots\sum_{i_M=1}^{d} v_{0k}^{(R)}(x)v_{0m}^{(R)}(x)v_{0i_3}^{(R)}(x)\cdots v_{0i_M}^{(R)}(x) \\
&\qquad \times v_{0i_3}^{(p)}(x)\cdots v_{0i_M}^{(p)}(x) - \delta_{km}\lambda_p \\
&= \big[(M-1)\lambda_p v_{0k}^{(p)}(x)v_{0m}^{(p)}(x) - \delta_{km}\lambda_p\big] + (M-1)\sum_{r=p+1}^{R}\lambda_r v_{0k}^{(r)}(x)v_{0m}^{(r)}(x)q_{r,p}^{M-2},
\end{aligned}
\tag{12}
$$

where $q_{r,p} = \langle v_0^{(r)}, v_0^{(p)}\rangle$, $p \neq r = 1,\dots,R$. Now consider a generic $u \in \mathbf{R}^d$, which belongs to $\ker(\mathcal{T}(v_0^{(p)}(x), D_0^{(p)}(x)) - \lambda_p I)$, then

$$
\begin{aligned}
&(M-1)\sum_{m=1}^{d}\lambda_p v_{0k}^{(p)}(x)v_{0m}^{(p)}(x)u_m \\
&\qquad + (M-1)\sum_{m=1}^{d}\sum_{r=p+1}^{R}\lambda_r v_{0k}^{(r)}(x)v_{0m}^{(r)}(x)u_m q_{r,p}^{M-2} = \lambda_p u_k.
\end{aligned}
\tag{13}
$$

Then taking a sum over $\sum_{k=1}^{d} v_{0k}^{(p)}(x)$ we get

$$
\begin{aligned}
&(M-1)\lambda_p\sum_{m=1}^{d} v_{0m}^{(p)}(x)u_m + (M-1)\sum_{m=1}^{d}\sum_{r=p+1}^{R}\lambda_r v_{0m}^{(r)}(x)u_m q_{r,p}^{M-1} \\
&\qquad = \lambda_p\sum_{k=1}^{d} v_{0k}^{(p)}(x)u_k, \\
&(M-2)\lambda_p\sum_{m=1}^{d} v_{0m}^{(p)}(x)u_m + (M-1)\sum_{r=p+1}^{R}\lambda_r q_{r,p}^{M-1}\sum_{m=1}^{d} v_{0m}^{(r)}(x)u_m = 0.
\end{aligned}
$$

Similarly, taking $\sum_{k=1}^{d} v_{0k}^{(l)}(x)$ on (13), we get, for $l = p+1,\dots,R$,

$$
\begin{aligned}
&(M-1)\lambda_p q_{l,p}\sum_{m=1}^{d} v_{0m}^{(p)}(x)u_m \\
&\qquad + (M-1)\sum_{r=p+1}^{R}\lambda_r q_{l,r} q_{r,p}^{M-1}\sum_{m=1}^{d} v_{0m}^{(r)}(x)u_m = \lambda_p\sum_{m=1}^{d} v_{0m}^{(l)}(x)u_m.
\end{aligned}
\tag{14}
$$

Suppose $\sum_{m=1}^{d} v_{0m}^{(l)}(x)u_m = A_l$, then together with (13) and (14) we get a linear system of equations, for $l = p+1, \dots, R$. It is of the form

$$\begin{aligned} \lambda_p(M-2)A_p + (M-1)\sum_{r=p+1}^{R} \lambda_r q_{r,p}^{M-1} A_r = 0, \\ \lambda_p(M-1)q_{l,p}A_p + (M-1)\sum_{r=p+1}^{R} \lambda_r q_{l,r} q_{r,p}^{M-2} A_r - \lambda_p A_l = 0. \end{aligned} \tag{15}$$

We can write these $R-p+1$ equations in $R-p+1$ variables A_l, $l = p+1, \dots, R$, given by (15), in matrix notations as $Q_{R,p}A = 0$, where $A = (A_p, A_{p+1}, \dots, A_R)$ and matrix $Q_{R,p}$ is given by

$$\begin{bmatrix} (M-2)\lambda_p & (M-1)\lambda_{p+1}q_{p+1,p}^{M-1} & \dots & (M-1)\lambda_R q_{R,p}^{M-1} \\ (M-1)\lambda_p q_{p+1,p} & (M-1)\lambda_{p+1}q_{p+1,p}^{M-2} - \lambda_p & \dots & (M-1)\lambda_R q_{p+1,R} q_{R,p}^{M-1} \\ \vdots & & & \\ (M-1)\lambda_p q_{R,p} & \dots & \dots & (M-1)\lambda_R q_{R,p}^{M-2} - \lambda_p \end{bmatrix}.$$

Now suppose $\det(Q_{R,p}) \neq 0$, then

$$A_l = \sum_{m=1}^{d} v_{0m}^{(l)} u_m = 0, \qquad l = p, p+1, \dots, R, \tag{16}$$

or in other words

$$\begin{bmatrix} v_{01}^{(p)} & \dots & v_{0d}^{(p)} \\ \vdots & & \\ v_{01}^{(R)} & \dots & v_{0d}^{(R)} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_d \end{bmatrix} = 0.$$

Now putting equation (16) back in (14), we get $u_k = 0$, $k = 1, \dots, d$. Now we need to find under what conditions the $((R-p+1)\times(R-p+1))$-matrix $Q_{R,p}$ is of the full rank for $p = 1, \dots, R$. Thus, in order to ensure $\det(Q_{R,p}) \neq 0$, $p = 1, \dots, R$, consider the following cases.

Case I. If $R \leqslant d$, then after choosing $R-p$ pseudo-eigenvectors we can choose the rest p of pseudo-eigenvectors $v_0^{(k)}$, $k = R-p+1, \dots, R$, that are mutually orthogonal, and as a result we will have $q_{k,m} = 0$, $k, m = R-p+1, \dots, R$. Therefore, as long as $\lambda_p \neq 0$, we will have $\det(Q_{R,p}) \neq 0$.

Case II. If $R > d$, then first let us consider $R = d+1$. In that case, we can choose the first d pseudo-eigenvectors $v_0^{(k)}$ to be orthogonal, hence $q_{i,j} = \delta_{ij}$, $i, j = 1, \dots, d$. Now we need to check $\det(Q_{R,p}) \neq 0$, $p = 1, \dots, R$.

1. When $p = R$, we have $\dim(Q_{R,p}) = 1 \times 1$ and

$$\det(Q_{R,p}) \neq 0 \quad \Longrightarrow \quad \lambda_R(M-2) \neq 0.$$

2. When $p = R-1$, we have $\dim(Q_{R,p}) = 2\times 2$ and

$$\begin{aligned} &\det(Q_{R,p}) \neq 0 \\ &\implies \quad \det\left(\begin{bmatrix} (M-2)\lambda_{R-1} & (M-1)\lambda_R q_{R,R-1}^{M-1} \\ (M-1)\lambda_{R-1} q_{R,R-1} & (M-1)\lambda_R q_{R,R-1}^{M-2} - \lambda_R \end{bmatrix}\right) \neq 0 \\ &\implies \quad (M-1)(M-2)q_{R,R-1}^{M-2} - (M-2) - (M-1)^2 q_{R,R-1}^{M} \neq 0. \quad (17) \end{aligned}$$

Now (17) is a polynomial in $q_{R,R-1}$, that can yield at most M roots, provided $\lambda_{R-1} \neq 0$. Hence, we can choose $v_0^{(R)}$ such that $q_{R,R-1}$ satisfies (17).

3. When $p = R-2$, we have $\dim(Q_{R,p}) = 3\times 3$. Also we use the fact that $q_{R-1,R-2} = 0$ as the first d pseudo-eigenvectors are orthogonal. Here in this case, the $Q_{R,p}$ matrix is given by

$$\begin{bmatrix} (M-2)\lambda_{R-2} & (M-1)\lambda_{R-1} q_{R-1,R-2} & (M-1)\lambda_{R-1} q_{R,R-2}^{M-1} \\ (M-1)\lambda_{R-2} q_{R-1,R-2} & (M-1)\lambda_{R-1} q_{R-1,R-2}^{M-2} - \lambda_{R-1} & (M-1)\lambda_R q_{R-1,R} q_{R,R-2}^{M-2} \\ (M-1)\lambda_{R-2} q_{R,R-2} & (M-1)\lambda_{R-1} q_{R-1,R} q_{R-1,R-2}^{M-2} & (M-1)\lambda_R q_{R,R-2}^{M-2} - \lambda_R \end{bmatrix},$$

and, therefore,

$$\begin{aligned} &\det(Q_{R,p}) \neq 0 \\ &\implies \det\left(\begin{bmatrix} (M-2)\lambda_{R-2} & 0 & (M-1)\lambda_{R-1} q_{R,R-2}^{M-1} \\ 0 & -\lambda_{R-1} & (M-1)\lambda_R q_{R-1,R} q_{R,R-2}^{M-2} \\ (M-1)\lambda_{R-2} q_{R,R-2} & 0 & (M-1)\lambda_R q_{R,R-2}^{M-2} - \lambda_R \end{bmatrix}\right) \neq 0 \\ &\implies \quad -\big((M-1)(M-2)q_{R,R-2}^{M-2} - (M-2) - (M-1)^2 q_{R,R-2}^{M}\big) \neq 0. \end{aligned}$$

Note that, the last assertion is similar to (17) provided $\lambda_R, \lambda_{R-1}, \lambda_{R-2} \neq 0$.

4. Similarly, for $p = R-3$, we have $\dim(Q_{R,p}) = 4\times 4$ and here we use the fact that $q_{R-2,R-3}, q_{R-1,R-3} = 0$ using similar arguments. Hence,

$$\begin{aligned} &\det(Q_{R,p}) \neq 0 \\ &\implies \det\left(\begin{bmatrix} (M-2)\lambda_{R-2} & 0 & 0 & (M-1)\lambda_R q_{R,R-3}^{M-1} \\ 0 & -\lambda_{R-2} & 0 & (M-1)\lambda_R q_{R-2,R} q_{R,R-3}^{M-2} \\ 0 & 0 & -\lambda_{R-1} & (M-1)\lambda_R q_{R-1,R} q_{R,R-3}^{M-2} \\ 0 & 0 & 0 & (M-1)\lambda_R q_{R,R-3}^{M-2} - \lambda_R \end{bmatrix}\right) \neq 0 \\ &\implies (-1)^2\big((M-1)(M-2)q_{R,R-3}^{M-2} - (M-2) - (M-1)^2 q_{R,R-3}^{M}\big) \neq 0. \end{aligned}$$

In this way we can continue and at each step we need to choose $v_0^{(R)}$, in such a way that, for $p = 1,\dots,R-1$,

$$(-1)^{(R-p-1)}\big((M-1)(M-2)q_{R,p}^{M-2} - (M-2) - (M-1)^2 q_{R,p}^{M}\big) \neq 0.$$

Finally, for any $R = d+k$, where $k > 1$, by induction we will get a polynomial expression in

$$\begin{bmatrix} q_{R,R-k+1} & q_{R,R-k} & q_{R,R-k-1} & \cdots & q_{R,p} \\ q_{R-1,R-k+1} & \cdots & & & q_{R,p} \\ \vdots & & & & \\ q_{R-k+1,R-k} & \cdots & & & q_{R-k+1,p} \end{bmatrix},$$

similar to (17), where each of the $q_{R_1,R_2} \in [-1,1]$ and the polynomial will be of a fixed integer degree. Therefore, we can choose $\lambda_1, \dots, \lambda_R \neq 0$ and $\{v_0^{(1)}, \dots, v_0^{(R)}\}$ such that $\det(Q_{R,p}) \neq 0$, $p = 1, \dots, R$. Lemma 1 is proved.

4. Proof of Proposition 1.

Proof. Let $\theta = (\theta^{(1)}, \dots, \theta^{(R)})$ be a random vector such that

$$\mathbf{P}\big(\theta^{(1)} = i_1, \dots, \theta^{(R)} = i_R\big) = \frac{1}{\mathcal{L}}, \qquad 1 \leqslant i_r \leqslant l_r + 1, \quad r = 1, \dots, R.$$

For ease of notation assume $\Psi(X) = (\Psi^{(1)}(X), \dots, \Psi^{(R)}(X))$. Then,

$$\begin{aligned} \sum_{i_1,\dots,i_R} &\mathbf{P}\big(\theta^{(1)} = i_1, \dots, \theta^{(R)} = i_R \bigm| X\big) \ln \mathbf{P}\big(\theta^{(1)} = i_1, \dots, \theta^{(R)} = i_R \bigm| X\big) \\ &= \mathbf{P}(\theta = \Psi(X) \,|\, X) \ln \mathbf{P}(\theta = \Psi(X) \,|\, X) \\ &\quad + \mathbf{P}(\theta \neq \Psi(X) \,|\, X) \ln \mathbf{P}(\theta \neq \Psi(X) \,|\, X) \\ &\quad + \mathbf{P}(\theta \neq \Psi(X) \,|\, X) \sum_{\underline{i} \neq \Psi(X)} \frac{\mathbf{P}(\theta = \underline{i} \,|\, X)}{\mathbf{P}(\theta \neq \Psi(X) \,|\, X)} \ln \frac{\mathbf{P}(\theta = \underline{i} \,|\, X)}{\mathbf{P}(\theta \neq \Psi(X) \,|\, X)} \\ &\geqslant -\ln 2 - \mathbf{P}(\theta \neq \Psi(X) \,|\, X) \ln(\mathcal{L} - 1), \end{aligned}$$

where we applied Lemma (5.1) from the [3] twice. The quantity on the left-hand side of this chain of inequalities will now be bounded from above. Indeed,

$$\mathbf{P}(\theta = \underline{i} \,|\, X) = \mathbf{P}\big(\theta^{(1)} = i_1, \dots, \theta^{(R)} = i_R \bigm| X\big) = \frac{f_{i_1,\dots,i_R}(X)}{\sum_{j_1,\dots,j_R} f_{j_1,\dots,j_R}(X)}.$$

Thus,

$$\begin{aligned} \mathbf{E}&\Big(\sum_{\underline{i}} \mathbf{P}(\theta = \underline{i} \,|\, X) \ln\big(\mathbf{P}(\theta = \underline{i} \,|\, X)\big)\Big) \\ &= \int \Big(\sum_{\underline{i}} \frac{f_{\underline{i}}(x)}{\sum_{\underline{j}} f_{\underline{j}}(x)} \ln\Big(\frac{f_{\underline{i}}(x)}{\sum_{\underline{j}} f_{\underline{j}}(x)}\Big)\Big) \frac{1}{\mathcal{L}} \sum_{\underline{j}} f_{\underline{j}}(x)\, dx \\ &= \frac{1}{\mathcal{L}} \sum_{\underline{i}} \int \ln\Big(\frac{f_{\underline{i}}(x)}{\sum_{\underline{j}} f_{\underline{j}}(x)}\Big) f_{\underline{i}}(x)\, dx, \end{aligned}$$

since,

$$\ln\left(\frac{1}{\mathcal{L}}\sum_{\underline{j}} f_{\underline{j}}(x)\right) \geqslant \frac{1}{\mathcal{L}}\sum_{\underline{j}} \ln f_{\underline{j}}(x) = \frac{1}{\mathcal{L}^2}\sum_{\underline{i},\underline{j}} \int \ln\left(\frac{f_{\underline{i}}}{f_{\underline{j}}}\right) f_{\underline{i}}\, dx - \ln\mathcal{L}$$
$$= \frac{1}{\mathcal{L}^2}\sum_{\underline{i},\underline{j}} K(f_{\underline{i}}, f_{\underline{j}}) - \ln\mathcal{L} \leqslant \beta - \ln\mathcal{L}.$$

We conclude that

$$\beta - \ln\mathcal{L} \geqslant -\ln 2 - \mathbf{P}(\theta \neq \Psi(X))\ln(\mathcal{L}-1).$$

Hence,

$$\sup_{\underline{i}} \mathbf{P}_{\underline{i}}(\Psi(X) \neq \underline{i}) \geqslant \mathbf{P}(\Psi(X) \neq \theta) \geqslant \frac{\ln\mathcal{L} - \beta - \ln 2}{\ln(\mathcal{L}-1)}.$$

Finally, $\sup_{\underline{i}} \mathbf{P}_{\underline{i}}(\Psi(X) \neq \underline{i}) \geqslant 1 - (\beta + \ln 2)/\ln(\mathcal{L}-1)$. Proposition 1 is proved.

5. Proof of Proposition 2.

Proof. Consider the following relations:

$$\begin{aligned}
K(f_{\underline{i}}, f_{\underline{j}}) &= \int_G \int_{\mathbf{R}^N} \ln \frac{f(\Sigma^{-1/2}(x)(y - B\underline{D}_{\underline{i}}(x)))}{f(\Sigma^{-1/2}(x)(y - B\underline{D}_{\underline{j}}(x)))} \\
&\qquad \times f(\Sigma^{-1/2}(x)(y - B\underline{D}_{\underline{i}}(x)))\, dx\, dy \\
&= \int_G \int_{\mathbf{R}^N} \ln \frac{f(\widetilde{y} - \Sigma^{-1/2}(x)B\underline{D}_{\underline{i}}(x))}{f(\widetilde{y} - \Sigma^{-1/2}(x)B\underline{D}_{\underline{j}}(x))} \\
&\qquad \times f(\widetilde{y} - \Sigma^{-1/2}(x)B\underline{D}_{\underline{i}}(x)) \det(\Sigma^{1/2}(x))\, dx\, d\widetilde{y} \\
&\overset{(1)}{=} -\int_G \int_{\mathbf{R}^N} \ln \frac{f(\widetilde{y} - \Sigma^{-1/2}(x)B\underline{D}_{\underline{j}}(x))}{f(\widetilde{y} - \Sigma^{-1/2}(x)B\underline{D}_{\underline{i}}(x))} \\
&\qquad \times f(\widetilde{y} - \Sigma^{-1/2}(x)B\underline{D}_{\underline{i}}(x)) \det(\Sigma^{1/2}(x))\, dx\, d\widetilde{y} \\
&= -\int_G \int_{\mathbf{R}^N} \ln \frac{f\big(z + \Sigma^{-1/2}(x)B(\underline{D}_{\underline{i}}(x) - \underline{D}_{\underline{j}}(x))\big)}{f(z)} f(z) \det(\Sigma^{1/2}(x))\, dx\, d\widetilde{y} \\
&\overset{(2)}{=} -\int_G \int_{\mathbf{R}^N} \ln\left(1 + \frac{f\big(z + \Sigma^{-1/2}(x)B(\underline{D}_{\underline{i}}(x) - \underline{D}_{\underline{j}}(x))\big) - f(z)}{f(z)}\right) \\
&\qquad \times f(z) \det(\Sigma^{1/2}(x))\, dx\, dz.
\end{aligned}$$

Here in equality (1) we used the change of variable $\widetilde{y} = \Sigma^{-1/2}y$ on y and in equality (2) we used the change of variable $z = \widetilde{y} - \Sigma^{-1/2}(x)B\underline{D}_{\underline{i}}(x)$.

Finally, applying Lemma 2 we obtain

$$
\begin{aligned}
K(f_{\underline{i}}, f_{\underline{j}}) &\leqslant C(f,\delta_1)\int_G |\Sigma^{-1/2}(x)B(\underline{D}_{\underline{i}}(x)-\underline{D}_{\underline{j}}(x))|^2 \det(\Sigma^{1/2}(x))\,dx \\
&\leqslant C(f,\delta_1)\int_G \|\Sigma^{-1/2}(x)\|_F^2\|B\|_F^2\|\underline{D}_{\underline{i}}(x)-\underline{D}_{\underline{j}}(x)\|_F^2 \det(\Sigma^{1/2}(x))\,dx \\
&\leqslant C(f,\delta_1)C_{\Sigma,B}\|D_\theta - D_{\widetilde{\theta}}\|_G^2 \leqslant C(f,\delta_1)C_{\Sigma,B}\beta.
\end{aligned}
$$

Proposition 2 is proved.

6. Proof of Lemma 4.

Proof. Following the arguments in Lemma 3 of [11] we note that the quantity $t_b^{(r)}$, $r=1,\dots,R$, is twice continuously differentiable. Therefore, the implicit differentiation of the equation (1) would imply the gradient and the Hessian of $t_b^{(r)}$, $r=1,\dots,R$, are bounded for all sufficiently large n. In other words, for each $r=1,\dots,R$,

$$
\|t_b^{'(r)}(x)\|_2 \leqslant \frac{|t_0^{'(r)}(x)|+\varepsilon_1^{(r)}}{1-\varepsilon_1^{(r)}} \quad \text{and} \quad \|t_b^{''(r)}(x)\|_F \leqslant \varepsilon_2^{(r)},
$$

for some positive constants $\varepsilon_1^{(r)}, \varepsilon_2^{(r)}$. Also,

$$
\varphi_b^{'(r)}(t) = \sum_{i=1}^P b_i \varphi^{'(r)} \frac{t-(i-1)h}{h}, \quad \varphi_b^{''(r)}(t) = \sum_{i=1}^P b_i \varphi^{''(r)} \frac{t-(i-1)h}{h}\,\frac{n^\delta P_1}{T},
$$

$$
\varphi_b^{'''(r)}(t) = \sum_{i=1}^P b_i \varphi^{'''(r)} \frac{t-(i-1)h}{h}\,\frac{n^{2\delta} P_1^2}{T^2}.
$$

Recall $D_0(x)=\sum_{r=1}^R \lambda_r v_0^{(r)}(x)^{\otimes M}$. Then we can rewrite the tensor elements

$$
D_{0,i_1,\dots,i_M}(x) = \sum_{r=1}^R \lambda_r \Bigl(\prod_{j=1}^M v_{0,i_j}^{(r)}(x)\Bigr).
$$

Taking the first and the second derivatives of $D_{0,i_1,\dots,i_M}(x)$ in x_u and x_u, x_v, respectively, we get

$$
\begin{aligned}
\frac{\partial D_{0,i_1,\dots,i_M}(x)}{\partial x_u} &= \sum_{r=1}^R \lambda_r \Bigl(\sum_{j=1}^M \frac{\partial v_{0,i_j}^{(r)}(x)}{\partial x_u} \prod_{l=1,\,l\neq j}^M v_{0,i_l}^{(r)}(x)\Bigr), \\
\frac{\partial^2 D_{0,i_1,\dots,i_M}(x)}{\partial x_u\,\partial x_v} &= \sum_{r=1}^R \lambda_r \Bigl(\sum_{j=1}^M \frac{\partial^2 v_{0,i_j}^{(r)}(x)}{\partial x_u\,\partial x_v}\Bigl(\prod_{l=1,\,l\neq j}^M v_{0,i_j}^{(r)}(x)\Bigr) \\
&\quad + \sum_{\substack{j_1,j_2=1\\ j_1\neq j_2}}^M \frac{\partial v_{0,i_{j_1}}^{(r)}(x)}{\partial x_u}\,\frac{\partial v_{0,i_{j_2}}^{(r)}(x)}{\partial x_v}\Bigl(\prod_{\substack{l=1\\ l\neq j_1,j_2}}^M v_{0,i_l}^{(r)}(x)\Bigr)\Bigr).
\end{aligned} \tag{18}
$$

Similarly,

$$D_b(x) = \sum_{r=1}^{R} \lambda_r v_0^{(r)}(x)^{\otimes M} \big(1 + n^{-\alpha} \varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{(r)}(n^\gamma |a_b(x) - a_*|)\big)^M.$$

Hence, we can rewrite the above tensor elementwise as follows:

$$\begin{aligned} &D_{b,i_1,\ldots,i_M}(x) \\ &= \sum_{r=1}^{R} \lambda_r v_{0,i_1}^{(r)}(x) \ldots v_{0,i_M}^{(r)}(x)\big(1 + n^{-\alpha} \varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{(r)}(n^\gamma |a_0(x) - a_*|)\big)^M \\ &= \sum_{r=1}^{R} \lambda_r \Bigg(\prod_{j=1}^{M} v_{0,i_j}^{(r)}(x)\Bigg)\big(1 + n^{-\alpha} \varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{(r)}(n^\gamma |a_0(x) - a_*|)\big)^M. \end{aligned}$$

Now taking the partial derivative of the tensor component $D_{b,i_1,\ldots,i_M}(x)$ with respect to x_u, we get

$$\begin{aligned} \frac{\partial D_{b,i_1,\ldots,i_M}(x)}{\partial x_u} &= \sum_{r=1}^{R} \lambda_r \big(1 + n^{-\alpha} \varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{(r)}(n^\gamma |a_0(x) - a_*|)\big)^M \\ &\times \Bigg(\sum_{j=1}^{M} \frac{\partial v_{0,i_j}^{(r)}(x)}{\partial x_u} \prod_{l=1, l\neq j}^{M} v_{0,i_l}^{(r)}(x)\Bigg) + \sum_{r=1}^{R} \lambda_r n^{-\alpha} \Bigg(\prod_{j=1}^{M} v_{0,i_j}^{(r)}(x)\Bigg) \\ &\quad \times \Bigg(M\big(1 + n^{-\alpha} \varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{(r)}(n^\gamma |a_0(x) - a_*|)\big)^{M-1} \\ &\quad \times \Bigg(\varphi_b^{\prime\prime(r)}\big(t_b^{(r)}(x)\big) \frac{\partial t_b^{(r)}(x)}{\partial x_u} \psi^{(r)}(n^\gamma |a_0(x) - a_*|) \\ &\quad\quad + n^\gamma \varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{\prime(r)}(n^\gamma |a_0(x) - a_*|) \frac{(a_0(x) - a_*)^T}{|a_0(x) - a_*|} \frac{\partial a_0(x)}{\partial x_u}\Bigg)\Bigg). \end{aligned}$$

Further, the second order derivative of the tensor component $D_{b,i_1,\ldots,i_M}(x)$ with respect to x_u and x_v, where $u \neq v$, is given by

$$\begin{aligned} \frac{\partial^2 D_{b,i_1,\ldots,i_M}(x)}{\partial x_u \partial x_v} &= \sum_{r=1}^{R} \lambda_r \big(1 + n^{-\alpha} \varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{(r)}(n^\gamma |a_0(x) - a_*|)\big)^M \\ &\times \Bigg(\sum_{j=1}^{M} \frac{\partial^2 v_{0,i_j}^{(r)}(x)}{\partial x_u \, \partial x_v} \Bigg(\prod_{l\neq j} v_{0,i_l}^{(r)}(x)\Bigg) \\ &\quad + \sum_{\substack{j_1,j_2=1 \\ j_1\neq j_2}}^{M} \frac{\partial v_{0,i_{j_1}}^{(r)}(x)}{\partial x_u} \frac{\partial v_{0,i_{j_2}}^{(r)}(x)}{\partial x_v} \Bigg(\prod_{\substack{l=1 \\ l\neq j_1,j_2}}^{M} v_{0,i_l}^{(r)}(x)\Bigg)\Bigg) \end{aligned}$$

$$
\begin{aligned}
&+ 2Mn^{-\alpha}\sum_{r=1}^{R}\lambda_r\big(1+n^{-\alpha}\varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{(r)}(n^\gamma|a_0(x)-a_*|)\big)^{M-1}\\
&\times\Bigg(\varphi_b^{\prime\prime(r)}\big(t_b^{(r)}(x)\big)\,\frac{\partial t_b^{(r)}(x)}{\partial x_u}\,\psi^{(r)}(n^\gamma|a_0(x)-a_*|)\\
&\quad + n^\gamma\varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{\prime(r)}(n^\gamma|a_0(x)-a_*|)\frac{(a_0(x)-a_*)^T}{|a_0(x)-a_*|}\,\frac{\partial a_0(x)}{\partial x_u}\Bigg)\\
&\times\Bigg(\sum_{j=1}^{M}\frac{\partial v_{0,i_j}^{(r)}(x)}{\partial x_u}\prod_{l=1,\,l\neq j}^{M}v_{0,i_l}^{(r)}(x)\Bigg)\\
&+ M(M-1)n^{-2\alpha}\Bigg(\varphi_b^{\prime\prime(r)}\big(t_b^{(r)}(x)\big)\,\frac{\partial t_b^{(r)}(x)}{\partial x_u}\psi^{(r)}(n^\gamma|a_0(x)-a_*|)\\
&\quad + n^\gamma\varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{\prime(r)}(n^\gamma|a_0(x)-a_*|)\frac{(a_0(x)-a_*)^T}{|a_0(x)-a_*|}\,\frac{\partial a_0(x)}{\partial x_u}\Bigg)^2\\
&+ Mn^{-\alpha}\sum_{r=1}^{R}\lambda_r\Bigg(\prod_{j=1}^{M}v_{0,i_j}^{(r)}(x)\Bigg)\big(1+n^{-\alpha}\varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{(r)}(n^\gamma|a_0(x)-a_*|)\big)^{M-1}\\
&\times\Bigg(\varphi_b^{\prime\prime\prime(r)}\big(t_b^{(r)}(x)\big)\Bigg(\frac{\partial t_b^{(r)}(x)}{\partial x_u}\Bigg)^T\frac{\partial t_b^{(r)}(x)}{\partial x_v}\,\psi^{(r)}(n^\gamma|a_0(x)-a_*|)\\
&\quad + \varphi_b^{\prime\prime(r)}\big(t_b^{(r)}(x)\big)\frac{\partial^2 t_b^{(r)}(x)}{\partial x_u\partial x_v}\psi^{(r)}(n^\gamma|a_0(x)-a_*|)\\
&\quad + n^\gamma\varphi_b^{\prime\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{\prime(r)}(n^\gamma|a_0(x)-a_*|)\\
&\quad\times\Bigg(\frac{\partial t_b^{(r)}(x)}{\partial x_u}\Bigg(\frac{(a_0(x)-a_*)^T}{|a_0(x)-a_*|}\,\frac{\partial a_0(x)}{\partial x_v}\Bigg)+\frac{\partial t_b^{(r)}(x)}{\partial x_v}\Bigg(\frac{(a_0(x)-a_*)^T}{|a_0(x)-a_*|}\frac{\partial a_0(x)}{\partial x_u}\Bigg)\Bigg)\\
&\quad + n^{2\gamma}\varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{\prime\prime(r)}(n^\gamma|a_0(x)-a_*|)\\
&\quad\times\Bigg(\frac{(a_0(x)-a_*)^T}{|a_0(x)-a_*|}\,\frac{\partial a_0(x)}{\partial x_v}\Bigg)\Bigg(\frac{(a_0(x)-a_*)^T}{|a_0(x)-a_*|}\frac{\partial a_0(x)}{\partial x_u}\Bigg)\\
&\quad + n^\gamma\varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{\prime(r)}(n^\gamma|a_0(x)-a_*|)\\
&\quad\times\Bigg(\Bigg(\frac{\partial a_0(x)}{\partial x_v}\Bigg)^T\frac{\partial a_0(x)}{\partial x_u}\,\frac{1}{|a_0(x)-a_*|}+\frac{(a_0(x)-a_*)^T}{|a_0(x)-a_*|}\,\frac{\partial^2 a_0(x)}{\partial x_u\partial x_v}\\
&\qquad - \frac{(a_0(x)-a_*)^T}{|a_0(x)-a_*|^3}\,\frac{\partial a_0(x)}{\partial x_u}(a_0(x)-a_*)^T\,\frac{\partial a_0(x)}{\partial x_v}\Bigg)\Bigg). \qquad (19)
\end{aligned}
$$

The leading term in (19) is

$$
\begin{aligned}
&n^{2\gamma-\alpha}\varphi_b^{\prime(r)}\big(t_b^{(r)}(x)\big)\psi^{\prime\prime(r)}(n^\gamma|a_0(x)-a_*|)\\
&\qquad\times\Bigg(\frac{(a_0(x)-a_*)^T}{|a_0(x)-a_*|}\,\frac{\partial a_0(x)}{\partial x_v}\Bigg)\Bigg(\frac{(a_0(x)-a_*)^T}{|a_0(x)-a_*|}\frac{\partial a_0(x)}{\partial x_u}\Bigg).
\end{aligned}
$$

Now by construction the vector fields $v_0^{(r)}(x)$, $r = 1, \dots, R$, are bounded and twice continuously differentiable, and the functions $\varphi^{(r)}$, $\psi^{(r)}$, $r = 1, \dots, R$, are bounded thrice and twice continuously differentiable, respectively. Therefore, in the expression of (19), if $\alpha > 0$ and $\gamma = \alpha/2$, then D_b satisfies $\|D_b - D_0\|_\infty \leqslant cn^{-\alpha}$ and $\|D_b'' - D_0''\|_\infty \leqslant k$ for some constant $k > 0$. Lemma 4 is proved.

REFERENCES

1. C. Banerjee, L. A. Sakhanenko, D. C. Zhu, "Global rate optimality of integral curve estimators in high order tensor models", *Теория вероятн. и ее примен.*, **68**:2 (2023), 301–321; C. Banerjee, L. A. Sakhanenko, D. C. Zhu, "Global rate optimality of integral curve estimators in high order tensor models", *Theory Probab. Appl.*, **68**:2 (2023), 250–266.
2. O. Carmichael, L. Sakhanenko, "Estimation of integral curves from high angular resolution diffusion imaging (HARDI) data", *Linear Algebra Appl.*, **473** (2015), 377–403.
3. L. Devroye, *A course in density estimation*, Progr. Probab. Statist., **14**, Birkhäuser Boston, Inc., Boston, MA, 1987, xx+183 pp.
4. R. Hasminskii, I. Ibragimov, "On density estimation in the view of Kolmogorov's ideas in approximation theory", *Ann. Statist.*, **18**:3 (1990), 999–1010.
5. M. Ibragimov, R. Ibragimov, "Optimal constants in the Rosenthal inequality for random variables with zero odd moments", *Statist. Probab. Lett.*, **78**:2 (2008), 186–189.
6. И. А. Ибрагимов, Р. З. Хасьминский, *Асимптотическая теория оценивания*, Наука, М., 1979, 527 с.; англ. пер.: I. A. Ibragimov, R. Z. Has'minskii, *Statistical estimation. Asymptotic theory*, Appl. Math., **16**, Springer-Verlag, New York–Berlin, 1981, vii+403 pp.
7. И. А. Ибрагимов, Р. З. Хасьминский, "Об оценке плотности распределения", *Исследования по математической статистике*. IV, Зап. науч. сем. ЛОМИ, **98**, Изд-во "Наука", Ленинград. отд., Л., 1980, 61–85; англ. пер.: I. A. Ibragimov, R. Z. Khas'minskii, "Estimation of distribution density", *J. Soviet Math.*, **21**:1 (1983), 40–57.
8. И. А. Ибрагимов, Р. З. Хасьминский, "Еще об оценке плотности распределения", *Исследования по математической статистике*. V, Зап. науч. сем. ЛОМИ, **108**, Изд-во «Наука», Ленинград. отд., Л., 1981, 72–88; англ. пер.: I. A. Ibragimov, R. Z. Khas'minskii, "More on the estimation of distribution densities", *J. Soviet Math.*, **25**:3 (1984), 1155–1165.
9. V. Koltchinskii, L. Sakhanenko, Songhe Cai, "Integral curves of noisy vector fields and statistical problems in diffusion tensor imaging: nonparametric kernel estimation and hypotheses testing", *Ann. Statist.*, **35**:4 (2007), 1576–1607.
10. L. Sakhanenko, "How to choose the number of gradient directions for estimation problems from noisy diffusion tensor data", *Contemporary developments in statistical theory*, A festschrift for Hira Lal Koul, Springer Proc. Math. Stat., **68**, Springer, Cham, 2014, 305–310.
11. Л. А. Саханенко, "Глобальная оптимальная скорость сходимости в модели для исследования тензоров диффузии", *Теория вероятн. и ее примен.*, **55**:1

(2010), 19–35; англ. пер.: L. Sakhanenko, "A global optimal convergence rate in a model for the diffusion tensor imaging", *Theory Probab. Appl.*, **55**:1 (2011), 77–90.

12. A. W. van der Vaart, J. A. Wellner, *Weak convergence and empirical processes. With applications to statistics*, Springer Ser. Statist., Springer-Verlag, New-York, 1996, xvi+508 pp.

Поступила в редакцию
1.XI.2021