Convergence rates in curve estimation

Sylvain Delattre & Aurélie Fischer

Laboratoire de Probabilités, Statistique et Modélisation, UMR CNRS 8001
Université Paris Cité, 75013 Paris
sylvain.delattre@lpsm.paris;aurelie.fischer@lpsm.paris

2024

We consider the problem of estimating the image of an unknown function g from the observation of $X_1, \ldots, X_n$ drawn according to a model $X_i = g(U_i) + \xi_i$, $i = 1, \ldots, n$, where the U_i, $i = 1, \ldots, n$, are independent and the errors ξ_i, $i = 1, \ldots, n$, are also independent and in L^{2q}. We propose an estimator based on minimizing an L^q criterion and study its rate of convergence in Hausdorff distance.

Keywords – Curve estimation, additive noise, principal curves, length constraint.

2020 Mathematics Subject Classification: 62G05

1 Introduction

1.1 First insight into the main results

Let $n \geq 2$. We consider the model

$$X_i = g(U_i) + \xi_i, \quad i = 1, \ldots, n, \tag{1.1}$$

where the curve g is unknown, the ξ_i are independent random variables, the U_i are independent random variables with distribution $\mu_i \geq c\lambda$ on $[0, 1]$ for $i = 1, \ldots, n$. Note that we do not require that the noise ξ_i is independent from U_i, so that the context is different from that of a deconvolution problem, even if the distribution of the noise were known. We assume that g is rectifiable and Lipschitz. Moreover, we suppose that g is injective and $\text{reach}(\text{Im}g) \geq r > 0$, where $\text{reach}(\text{Im}g)$ denotes the maximal radius of a ball rolling on it (see Federer, 1959).

The first important result is the construction of a sequence of curves to estimate the image of the unknown function g in Hausdorff distance. Asymptotically, the length of the proposed estimator is equal to the length of g. We provide an upper bound for the rate of convergence of these estimators for the convergence in probability, in Hausdorff distance. Here, convergence is to be understood in the sense of the double asymptotic where the sample size n gets large and the noise gets small.

1.2 Related work

Estimating the image of g recalls the question of filament estimation and is a particular case, in dimension 1, of manifold estimation.

Genovese et al. (2012a) study an additive noise model of the form (1.1), where the curve g is parameterized by arc-length, normalized to $[0,1]$. The U_i, $i=1,\dots,n$, are assumed to have a common density with respect to the Lebesgue measure on $[0,1]$, which is bounded and bounded away from zero. The noise is supported in a ball $B(0,\sigma)$, with $\sigma < \mathrm{reach}(g)$, and has a bounded density with respect to the Lebesgue measure. This density is continuous on $\mathring{B}(0,\sigma)$, nondecreasing and symmetric, with a regularity condition on the boundary of the support. For an open curve (with endpoints), in addition, $|g(1)-g(0)|/2 > \sigma$. In the plane $\mathbb{R}^2$, the authors estimate the support of the distribution of the observations, and its boundary, in order to find its medial axis. Clutter noise is also considered: in this situation, the observations are sampled from a mixture density $(1-\eta)u(x)+\eta h(x)$, where u is the uniform density over some compact set, and h is the density of points on the shape. Considering manifolds without boundary, with dimension lower than the dimension of the ambient space, contained in a compact set, Genovese et al. (2012b) investigates their estimation from an additive model taking the form

$$X_i = G_i + \varepsilon_i, \quad i=1,\dots,n. \tag{1.2}$$

Here the random vectors G_i are drawn uniformly on the shape M, and the noise is drawn uniformly on the normal to the manifold, at distance at most $\sigma < \mathrm{reach}(M)$. The article Genovese et al. (2012c) is also dedicated to manifold estimation, in a noiseless model, in the presence of clutter noise, as well as in an additive noise model, with known Gaussian noise, which is related to density deconvolution. Estimating manifolds without boundary, with low dimension and a lower bound on the reach, is also the purpose of Aamari and Levrard (2018, 2019). The observations have a common density with respect to the d-dimensional Hausdorff measure of the manifold. Is is supposed bounded and bounded away from zero. In Aamari and Levrard (2018), estimation relies on Tangential Delaunay Complexes. It is performed in the noiseless case, with additive noise, bounded by σ, and under clutter noise. Aamari and Levrard (2019) deal with compact manifolds belonging to specific smoothness classes. The authors examine the noiseless situation, as well as centered bounded noise perpendicular to the manifold. Estimators based on local polynomials are proposed.

In Delattre and Fischer (2024), we consider a model of the same general form as (1.1). The goal is also to estimate the image of the unknown function g. We assume a weak condition on the noise: more precisely, we suppose that $\frac{1}{n}\sum_{i=1}^{n}|\xi_i|^q$ converges in probability to 0 as n tends to infinity. In this context, we construct a convergent estimator, but, without any further information on the noise, it is not possible to get a convergence rate. The purpose in Delattre and Fischer (2024) was to check if it is possible to design a procedure leading to a convergent estimator when neither the length nor the shape of the noise is known, which is not completely obvious at first sight.

1.3 Organization of the paper

The manuscript is organized as follows. In Section 2, we set up notation, provide some definitions, and describe the model. In Section 3, we define the estimator and state the main result, which is an upper bound for the Hausdorff distance between the image of the unknown curve g and the estimator.

2 Notation and the model

2.1 Notation and first definitions

For $d \geq 1$, the space $\mathbb{R}^d$ is equipped with the standard Euclidean norm, denoted by $|\cdot|$. For $u \in \mathbb{R}$, $\lfloor u \rfloor$ denotes the floor of x.

For $x \in \mathbb{R}^d$, $A \subset \mathbb{R}^d$, let $d(x, A) = \inf_{y \in A} |x - y|$ denote the distance from point x to set A.

We denote by $d_H(A, B)$ the Hausdorff distance between two sets A and B, given by

$$d_H(A, B) = \sup_{a \in A} d(a, B) \vee \sup_{b \in B} d(b, A).$$

For $\rho > 0$, let $B(x, \rho)$ denote the open ball of $\mathbb{R}^d$ with center x and radius ρ.

Moreover, for $d \geq 1$, λ denotes the Lebesgue measure on $\mathbb{R}^d$. Let α_d stand for the volume of the unit ball in $\mathbb{R}^d$.

The reach of a set $A \subset \mathbb{R}^d$ is the supremum of the radii ρ such that every point at distance at most ρ of A has a unique projection on A. More formally, following Federer (1959), we set for $A \subset \mathbb{R}^d$

$$\text{reach}(A) = \sup \left\{ \rho \geq 0 \mid \forall x \in \mathbb{R}^d \quad d(x, A) \leq \rho \Rightarrow \exists! a \in A \quad d(x, a) = d(x, A) \right\} \in [0, +\infty].$$

For $A \subset \mathbb{R}^d$ and $r \geq 0$, we denote by

$$A \oplus r = \left\{ x \in \mathbb{R}^d \mid d(x, A) \leq r \right\}$$

the r-enlargement of A.

A continuous function from $[0, 1]$ to $\mathbb{R}^d$ will be called a curve. If a curve f is rectifiable, its length will be denoted by $\mathscr{L}(f)$.

2.2 Description of the model

Let $q \in [1, \infty)$ and $n \geq 2$. We consider the model

$$X_i = g(U_i) + \xi_i, \quad i = 1, \ldots, n, \tag{2.1}$$

where the curve g is unknown, the ξ_i are independent random variables, the U_i are independent random variables with distribution $\mu_i \geq c\lambda$ on $[0, 1]$ for some constant $c > 0$, for $i = 1, \ldots, n$. We assume that g is rectifiable and $\mathscr{L}(g)$-Lipschitz and that it is injective, with $\text{reach}(\text{Im}g) \geq r > 0$.

We also suppose that:

- $E[|\xi_i|^q] \leq m_q^q$,
- $\text{Var}\big(|\xi_i|^q\big) \leq \sigma_q^{2q}$.

3 Convergent estimation

3.1 Definition of the estimator and upper bound in Hausdorff distance

The definition of the estimator is based on the notion of principal curves with bounded length, introduced by Kégl et al. (2000). Let

$$\Delta_n(f) = \left(\frac{1}{n}\sum_{i=1}^{n} d(X_i, \mathrm{Im} f)^q\right)^{1/q} = \left(\frac{1}{n}\sum_{i=1}^{n} d(g(U_i)+\xi_i, \mathrm{Im} f)^q\right)^{1/q}.$$

Let $\gamma > 0$. For each $L \geq 0$, we define $\hat{f}_L$ such that

$$\mathscr{L}(\hat{f}_L) \leq L, \quad \Delta_n(\hat{f}_L) \leq \min_{\mathscr{L}(f)\leq L} \Delta_n(f) + \gamma. \tag{3.1}$$

In the next theorem, which is the main result in this article, we define an estimator $\hat{f}_{\hat{L}}$ based on this optimization criterion (3.1), and state an upper bound for the Hausdorff distance between its image and the image of the unknown curve g.

Theorem 3.1. *Let $n \geq 1$, $\varepsilon \in (0, 1/2]$. We set*

$$s = s(q, \varepsilon, n) = \left(m_q^q + n^{-1/2}(1/\varepsilon - 1)^{1/2}\sigma_q^q\right)^{1/q}$$

and define

$$\hat{L} = \min\{L \in \delta\mathbb{N}, \Delta_n(\hat{f}_L) \leq s + \gamma\}.$$

With probability larger than $1 - 3\varepsilon$,

$$d_H(\mathrm{Im}\hat{f}_{\hat{L}}, \mathrm{Im} g) \leq F\left[3 + \frac{\alpha_d d}{\alpha_{d-1}} + \frac{\mathscr{L}(g)(d-1)}{r+F}\right] + \delta,$$

where

$$F = 2s + \gamma + \left((q+1)2^{q+3}\right)^{\frac{1}{q+1}} \left(\frac{\mathscr{L}(g)}{c}\right)^{\frac{1}{q+1}} (2s+\gamma)^{\frac{q}{q+1}} + 16(q+1)\frac{\ln(\frac{n^{2(q+1)}}{\varepsilon})}{n}\frac{\mathscr{L}(g)}{c}.$$

Proposition 3.1. *Assume that m_q, γ and δ tend to 0. Then, the length $\hat{L}$ converges in probability to $\mathscr{L}(g)$ as n tends to infinity.*

3.2 Proof of Theorem 3.1

A first step consists in examining the behavior of the noise of the model, which is the purpose of the next statement.

Lemma 3.1. *The errors in Model 2.1 satisfy*

$$P\left(\frac{1}{n}\sum_{i=1}^{n} |\xi_i|^q \geq s^q\right) \leq \varepsilon.$$

Proof. By the Chebyshev-Cantelli inequality, we have

$$\begin{aligned} P\left(\frac{1}{n}\sum_{i=1}^{n} |\xi_i|^q \geq s^q\right) &= P\left(\frac{1}{n}\sum_{i=1}^{n} |\xi_i|^q - m_q^q \geq \frac{\sigma_q^q}{\sqrt{n}}\left(\frac{1}{\varepsilon} - 1\right)^{1/2}\right) \\ &\leq P\left(\frac{1}{n}\sum_{i=1}^{n}(|\xi_i|^q - E[|\xi_i|^q]) \geq \frac{\sigma_q^q}{\sqrt{n}}\left(\frac{1}{\varepsilon} - 1\right)^{1/2}\right) \\ &\leq \varepsilon. \end{aligned}$$

□

Lemma 3.2. *We have*
$$P\left(\Delta_n(\hat{f}_{n,\mathscr{L}(g)}) \geq s+\gamma\right) \leq \varepsilon.$$

Proof. We have

$$\begin{aligned}\Delta_n(g) &= \Big(\frac{1}{n}\sum_{i=1}^n d\big(g(U_i)+\xi_i, \mathrm{Im}(g)\big)^q\Big)^{1/q}\\ &\leq \Big(\frac{1}{n}\sum_{i=1}^n |\xi_i|^q\Big)^{1/q}.\end{aligned}$$

Consequently, as $\Delta_n(\hat{f}_{\mathscr{L}(g)}) \leq \Delta_n(g)+\gamma$, using Lemma 3.1,

$$P(\Delta_n(\hat{f}_{\mathscr{L}(g)}) \geq s+\gamma) \leq P(\Delta_n(g)+\gamma \geq s+\gamma) \leq P(\Delta_n(g) \geq s) \leq \varepsilon.$$

□

Then, with probability larger than $1-\varepsilon$, $\hat{L} \leq \mathscr{L}(g)+\delta$. Set

$$M := \Big(\frac{1}{n}\sum_{i=1}^n d\big(g(U_i), \mathrm{Im}(\hat{f})\big)^q\Big)^{1/q},$$

and
$$S = \max_{1\leq i\leq n} d\big(g(U_i), \mathrm{Im}(\hat{f})\big).$$

Lemma 3.3. *We have* $P\big(M \geq 2s+\gamma\big) \leq \varepsilon$ *and* $S \leq n^{1/q}M$*.*

Proof. By Minkowski's inequality,

$$\begin{aligned}M &\leq \Delta_n(\hat{f}) + \Big(\frac{1}{n}\sum_{i=1}^n |\xi_i|^q\Big)^{1/q}\\ &\leq s+\gamma+\Big(\frac{1}{n}\sum_{i=1}^n |\xi_i|^q\Big)^{1/q}.\end{aligned}$$

Hence, by Lemma 3.1,
$$P\big(M \geq 2s+\gamma\big) \leq \varepsilon.$$

We have

$$\begin{aligned}S^q &= \max_{1\leq i\leq n} d\big(g(U_i), \mathrm{Im}(\hat{f})\big)^q\\ &\leq \sum_{i=1}^n d\big(g(U_i), \mathrm{Im}(\hat{f})\big)^q,\end{aligned}$$

thus
$$S \leq \Big(\sum_{i=1}^n d\big(g(U_i), \mathrm{Im}(\hat{f})\big)^q\Big)^{1/q} = n^{1/q}M.$$

□

We set

$$H_1 = \sup_{u\in[0,1]} d(g(u), \mathrm{Im}\hat{f})$$

and

$$H_2 = \sup_{u\in[0,1]} d(\hat{f}(u), \mathrm{Im}g).$$

In order to study $d_H(\mathrm{Im}\hat{f}, \mathrm{Im}g) = H_1 \vee H_2$, we will first control H_1 in Proposition 3.2 and then H_2 in Proposition 3.4 below. Putting these two results together leads to Theorem 3.1.

Proposition 3.2. *Let $n \geq 1$, $\varepsilon \in (0, 1/2]$. With probability larger than $1 - 3\varepsilon$,*

$$H_1 \leq 2s + \gamma + \left((q+1)2^{q+3}\right)^{\frac{1}{q+1}} \left(\frac{\mathscr{L}(g)}{c}\right)^{\frac{1}{q+1}} (2s+\gamma)^{\frac{q}{q+1}} + 16(q+1)\frac{\ln(\frac{n^{2(q+1)}}{\varepsilon})}{n}\frac{\mathscr{L}(g)}{c}.$$

The proof relies on Lemma 3.4 and Lemma 3.6 below.

Lemma 3.4. *Let $U_{(1)} \leq U_{(2)} \leq \cdots \leq U_{(n)}$ the order statistics corresponding to $U_1, \ldots, U_n$. Set $U_{(0)} = 0$ and $U_{(n+1)} = 1$. On has for all $t > 0$*

$$P\Big(\max_{1\leq i\leq n+1} U_{(i)} - U_{(i-1)} \geq \frac{\log(n+1)}{n}t\Big) \leq \frac{3}{t(n+1)^{ct/3-1}\log(n+1)}.$$

Proof. Let $s \in (0, 1]$ and set $k = \lfloor 2/s \rfloor + 1$. One has

$$\begin{aligned}\Big\{\max_{1\leq i\leq n+1} U_{(i)} - U_{(i-1)} \geq s\Big\} &= \Big\{\exists a \in [0, 1-s]\ \forall i = 1, \ldots, n\ U_i \notin (a, a+s)\Big\}\\ &\subset \Big\{\exists \ell \in \{1, \ldots, k\}\ \forall i = 1, \ldots, n\ U_i \notin ((\ell-1)/k, \ell/k)\Big\}\end{aligned}$$

since $1/k \leq s/2$. It follows that

$$\begin{aligned}P\Big(\max_{1\leq i\leq n+1} U_{(i)} - U_{(i-1)} \geq s\Big) &\leq \sum_{\ell=1}^{k} P\Big(\bigcap_{i=1}^{n}\{U_i \notin ((\ell-1)/k, \ell/k)\}\Big)\\ &= \sum_{\ell=1}^{k}\prod_{i=1}^{n} P\Big(U_i \notin ((\ell-1)/k, \ell/k)\Big)\\ &\leq k\big(1 - c/k\big)^n\\ &\leq k\exp(-cn/k) \quad \text{since } 1 - c/k \leq \exp(-c/k)\\ &\leq \frac{3}{s}\exp(-cns/3) \quad \text{since } k \leq 3/s\end{aligned}$$

Using that $\max_{1\leq i\leq n+1} U_{(i)} - U_{(i-1)} \leq 1$ a.s., one obtains

$$\forall s > 0 \quad P\Big(\max_{1\leq i\leq n+1} U_{(i)} - U_{(i-1)} \geq s\Big) \leq \frac{3}{s}\exp(-cns/3)$$

□

For $t \in [0, 1]$ and $x \geq 0$, we set (where $z_+ = \max(z, 0)$)

$$Z_n(t, x) = \frac{1}{n}\sum_{i=1}^{n}\Big(x - |U_i - t|\Big)_+^q \tag{3.2}$$

Lemma 3.5. *For all $t \in [0,1]$, $x \in (0,1/2]$, $z \geq 0$ one has*

$$P\Big(Z_n(t,x) \leq \frac{cx^{q+1}}{q+1} - z\Big) \leq \exp\Big(-\frac{n(2q+1)z^2}{2cx^{2q+1}}\Big)$$

Proof. Let $t \in [0,1]$ and $x \in (0,1/2]$. One has $Z_n(t,x) \geq \tilde{Z}_n(t,x)$ where

$$\tilde{Z}_n(t,x) = \begin{cases} \frac{1}{n}\sum_{i=1}^n 1_{\{t \leq U_i \leq t+x\}}\big(t+x-U_i\big)^q & \text{if } t \leq 1/2 \\ \frac{1}{n}\sum_{i=1}^n 1_{\{t-x \leq U_i \leq t\}}\big(U_i-(t-x)\big)^q & \text{if } t > 1/2 \end{cases}$$

thus it suffices to prove the result for $\tilde{Z}_n(t,x)$.

First consider the case where $t \in [0,1/2]$. Set $Y_i = 1_{\{t \leq U_i \leq t+x\}}\big(t+x-U_i\big)^q$, $1 \leq i \leq n$. Remark that, since $t+x \leq 1$ and $\mathscr{L}(U_i) \geq c\lambda$, one has $E(Y_i) \geq cx^{q+1}/(q+1)$. Let $z \geq 0$. Since the Y_i are independent, for all $b > 0$ one has

$$\begin{aligned} P\big(\tilde{Z}_n(t,x) - \frac{cx^{q+1}}{q+1} \leq -z\big) &\leq \exp\big(-bz + b\frac{cx^{q+1}}{q+1}\big)E\Big(\exp(-b\tilde{Z}_n(t,x))\Big) \\ &\leq \exp\big(-bz + b\frac{cx^{q+1}}{q+1}\big)\prod_{i=1}^n E\big(\exp(-\frac{b}{n}Y_i)\big) \end{aligned}$$

and moreover, using that

$$\exp(-\frac{b}{n}Y_i) = 1 - 1_{\{t \leq U_i \leq t+x\}}\big(1-\exp(-\frac{b}{n}(t+x-U_i)^q\big)$$

and $\mathscr{L}(U_i) \geq c\lambda$,

$$\begin{aligned} E\big(\exp(-\frac{b}{n}Y_i)\big) &\leq 1 - c\int_{[t,t+x]}\Big(1-\exp\big(-\frac{b}{n}(t+x-u)^q\big)\Big)du \\ &= 1 - c\int_0^x\Big(1-\exp\big(-\frac{b}{n}y^q\big)\Big)dy \\ &\leq 1 - c\int_0^x\Big(\frac{b}{n}y^q - \frac{b^2}{2n^2}y^{2q}\Big)dy \quad \text{since } 1-e^{-z} \geq z - z^2/2 \text{ for } z \geq 0 \\ &= 1 - c\Big(\frac{b}{n}\frac{x^{q+1}}{q+1} - \frac{b^2}{2n^2}\frac{x^{2q+1}}{2q+1}\Big) \end{aligned}$$

Therefore for all $b > 0$

$$\begin{aligned} P\big(\tilde{Z}_n(t,x) - \frac{cx^{q+1}}{q+1} \leq -z\big) &\leq \exp\big(-bz + b\frac{cx^{q+1}}{q+1}\big)\Big(1 - c\Big(\frac{b}{n}\frac{x^{q+1}}{q+1} - \frac{b^2}{2n^2}\frac{x^{2q+1}}{2q+1}\Big)\Big)^n \\ &\leq \exp\Big(-bz + \frac{cb^2}{2n}\frac{x^{2q+1}}{2q+1}\Big) \end{aligned}$$

using that $1-u \leq e^{-u}$. Minimizing with respect to b, we get

$$P\big(\tilde{Z}_n(t,x) - \frac{cx^{q+1}}{q+1} \leq -z\big) \leq \exp\big(-\frac{n(2q+1)z^2}{2cx^{2q+1}}\big)$$

as desired. The proof in the case where $t \in (1/2,1]$ is similar since

$$Z_n(t,x) = \frac{1}{n}\sum_{i=1}^n\Big(x - |1-U_i-(1-t)|\Big)_+^q$$

and $\mathscr{L}(1-U_i) \geq c\lambda$. $\square$

Lemma 3.6. *For all $\eta \in (0,1]$, for all $p \in \mathbb{N}^*$, for all n such that $\frac{8(q+1)^2\log(1/\eta)}{(2q+1)cn} \le 1/2$, one has*

$$P\Big(\exists t \in [0,1] \quad \exists x \in \Big[\frac{8(q+1)^2\log(1/\eta)}{(2q+1)cn}, \frac{1}{2}\Big] \quad Z_n(t,x) \le \frac{cx^{q+1}}{2(q+1)} - \frac{3}{p}\Big) \le p^2\eta.$$

Proof. Set $a(n,\eta) = \frac{8(q+1)^2\log(1/\eta)}{(2q+1)cn}$. From Lemma 3.5, we have

$$\forall t \in [0,1]\ \forall x \in [0,1/2], \quad P\Big(Z_n(t,x) \le \frac{cx^{q+1}}{q+1} - \Big(\frac{2c\log(1/\eta)}{n(2q+1)}\Big)^{1/2} x^{q+1/2}\Big) \le \eta$$

and, for $x \ge a(n,\eta)$,

$$\Big(\frac{2c\log(1/\eta)}{n(2q+1)}\Big)^{1/2} x^{q+1/2} \le \frac{cx^{q+1}}{2(q+1)}.$$

Thus,

$$\forall t \in [0,1]\ \forall x \in \Big[a(n,\eta), \frac{1}{2}\Big] \quad P\Big(Z_n(t,x) \le \frac{cx^{q+1}}{2(q+1)}\Big) \le \eta.$$

We compute

$$\frac{\partial}{\partial t} Z_n(t,x) = -\frac{1}{n}\sum_{i=1}^n q\ \mathrm{sign}(t-U_i)\big[x - |U_i - t|\big]_+^{q-1},$$

so that

$$\Big|\frac{\partial}{\partial t} Z_n(t,x)\Big| \le \frac{q}{2^{q-1}} \le 2.$$

Set $Y_n(t,x) = Z_n(t,x) - \frac{cx^{q+1}}{2(q+1)}$ for $x, t \in \mathbb{R}$. Then,

$$\frac{\partial}{\partial x} Y_n(t,x) = \frac{1}{n}\sum_{i=1}^n q\ \big[x - |U_i - t|\big]_+^{q-1} - \frac{cx^q}{2}$$

and

$$\Big|\frac{\partial}{\partial t} Y_n(t,x)\Big| \le \frac{q}{2^{q-1}} \vee \frac{c}{2^{q+1}} = \frac{q}{2^{q-1}} \le 2.$$

Since $t \mapsto Y_n(t,x)$ is 2-Lipschitz for each $x \in [0,1/2]$ and $x \mapsto Y(t,x)$ is 2-Lipschitz on $[0,1/2]$ for each t, we have

$$\min_{t\in[0,1]} \min_{x\in[a(n,\eta),1/2]} Y_n(t,x) \ge \min\Big\{Y_n\Big(\frac{i}{p}, \frac{j}{2p}\Big) - \frac{3}{p}; i,j \in \{1,\dots,p\},\ \frac{j}{2p} \ge a(n,\eta)\Big\}$$

$$\begin{aligned} P\Big(\min_{t\in[0,1]} \min_{x\in[a(n,\eta),1/2]} Y_n(t,x) \le -\frac{3}{p}\Big) &\le P\Big(\min\Big\{Y_n\Big(\frac{i}{p}, \frac{j}{2p}\Big); i,j \in \{1,\dots,p\},\ \frac{j}{2p} \ge a(n,\eta)\Big\} \le 0\Big) \\ &\le \sum_{i=1}^p \sum_{j=1}^p 1_{\big\{\frac{j}{2p} \ge a(n,\eta)\big\}} P\Big(Y_n\Big(\frac{i}{p}, \frac{j}{2p}\Big) \le 0\Big) \\ &\le p^2\eta. \end{aligned}$$

□

Proposition 3.3. *For all $n \ge 1$, for all $\varepsilon \in (0,1/2]$,*

$$P\Big(S \le M + \Big(\frac{(q+1)2^{q+3}}{c}\Big)^{\frac{1}{q+1}} \mathscr{L}(g)^{\frac{1}{q+1}} M^{\frac{q}{q+1}} + \frac{16(q+1)^2\log(n^{2(q+1)}/\varepsilon)}{(2q+1)cn}\mathscr{L}(g)\Big) \ge 1 - \varepsilon.$$

Proof. 1. Let I a random variable taking its values in $\{1,\dots,n\}$ such that $d\big(g(U_I),Im(\hat f)\big)=S$. Since the function $t\mapsto d\big(g(t),Im(\hat f)\big)$ is $\mathscr{L}(g)$-Lipschitz, one has for $1\le i\le n$:

$$d\big(g(U_I),Im(\hat f)\big)\le d\big(g(U_i),Im(\hat f)\big)+\mathscr{L}(g)|U_i-U_I|\le d\big(g(U_i),Im(\hat f)\big)+\mathscr{L}(g)$$

therefore, by Minkowski inequality,

$$S\le M+\mathscr{L}(g) \tag{3.3}$$

and

$$\begin{aligned}M^q&\ge\frac{1}{n}\sum_{i=1}^n\Big[d\big(g(U_I),Im\hat f\big)-\mathscr{L}(g)|U_i-U_I|\Big]_+^q\\&=\mathscr{L}(g)^q\frac{1}{n}\sum_{i=1}^n\Big[S/\mathscr{L}(g)-|U_i-U_I|\Big]_+^q\\&=\mathscr{L}(g)^qZ_n\big(U_I,S/\mathscr{L}(g)\big)\end{aligned}$$

where Z_n is given by (3.2). Hence

$$M^q\ge\mathscr{L}(g)^q\min_{t\in[0,1]}Z_n\big(t,S/\mathscr{L}(g)\big) \tag{3.4}$$

2. Let $\varepsilon\in(0,1/2]$ and set

$$a(n,\varepsilon)=\frac{8(q+1)^2\log(n^{2(q+1)}/\varepsilon)}{(2q+1)cn}$$

We have to show that

$$P\bigg(S\le M+\Big(\frac{(q+1)2^{q+3}}{c}\Big)^{\frac{1}{q+1}}\mathscr{L}(g)^{\frac{1}{q+1}}M^{\frac{q}{q+1}}+2a(n,\varepsilon)\mathscr{L}(g)\bigg)\ge1-\varepsilon.$$

If $a(n,\varepsilon)>1/2$ then $S\le M+2a(n,\varepsilon)\mathscr{L}(g)$ since we have (3.3) and the desired result holds.

3. From now on we assume that $a(n,\varepsilon)\le1/2$ and we define the event

$$\Omega_1=\left\{\min_{t\in[0,1]}\min_{x\in[a(n,\varepsilon),1/2]}Z_n(t,x)\ge\frac{cx^{q+1}}{2(q+1)}-\frac{3}{n^{q+1}}\right\}$$

Applying Lemma 3.6 with $p=n^{q+1}$ and $\eta=\varepsilon/n^{2(q+1)}$, we obtain $P(\Omega_1)\ge1-\varepsilon$. Therefore it suffices to prove that on Ω_1:

$$S\le M+\Big(\frac{(q+1)2^{q+3}}{c}\Big)^{\frac{1}{q+1}}\mathscr{L}(g)^{\frac{1}{q+1}}M^{\frac{q}{q+1}}+2a(n,\varepsilon)\mathscr{L}(g) \tag{3.5}$$

4. On $\{\mathscr{L}(g)a(n,\varepsilon)\le S<\mathscr{L}(g)/2\}\cap\Omega_1$, one has

$$\begin{aligned}M^q&\ge\mathscr{L}(g)^q\min_{t\in[0,1]}Z_n\Big(t,\frac{S}{\mathscr{L}(g)}\Big)\\&\ge\mathscr{L}(g)^q\Big(\frac{c}{2(q+1)}\Big(\frac{S}{\mathscr{L}(g)}\Big)^{q+1}-\frac{3}{n^{q+1}}\Big)\end{aligned}$$

therefore

$$S^{q+1} \leq \frac{2(q+1)\mathscr{L}(g)M^q}{c} + \frac{6(q+1)\mathscr{L}(g)^{q+1}}{cn^{q+1}}$$

$$S \leq \Big(\frac{2(q+1)\mathscr{L}(g)}{c}\Big)^{1/(q+1)} M^{q/(q+1)} + \Big(\frac{6(q+1)}{c}\Big)^{1/(q+1)} \frac{\mathscr{L}(g)}{n}$$

Moreover, using that $\varepsilon \leq 1/2$ and $q \geq 1$, one has $a(n,\varepsilon) \geq \left(\frac{6(q+1)}{c}\right)^{1/(q+1)} \frac{1}{n}$. Thus (3.5) holds true on $\{\mathscr{L}(g)a(n,\varepsilon) \leq S < \mathscr{L}(g)/2\} \cap \Omega_1$.

5. On $\{S \geq \mathscr{L}(g)/2\} \cap \Omega_1$, one has

$$\begin{aligned} M^q &\geq \mathscr{L}(g)^q \min_{t\in[0,1]} Z_n\Big(t, \frac{1}{2}\Big) \\ &\geq \mathscr{L}(g)^q \Big(\frac{c}{(q+1)2^{q+2}} - \frac{3}{n^{q+1}}\Big) \\ &\geq \mathscr{L}(g)^q \frac{c}{(q+1)2^{q+3}} \qquad \text{using that } a(n,\varepsilon) \leq 1/2 \end{aligned}$$

that is

$$\mathscr{L}(g) \leq \frac{2^{1+3/q}(q+1)^{1/q}}{c^{1/q}} M$$

One obtains on $\{S \geq \mathscr{L}(g)/2\} \cap \Omega_1$:

$$\begin{aligned} S &\leq M + \mathscr{L}(g) \\ &\leq M + \mathscr{L}(g) \wedge \Big(\frac{2^{1+3/q}(q+1)^{1/q}}{c^{1/q}} M\Big) \\ &\leq M + \mathscr{L}(g)^{1/(q+1)} \frac{2^{(q+3)/(q+1)}(q+1)^{1/(q+1)}}{c^{1/(q+1)}} M^{q/(q+1)} \\ &\leq M + \Big(\frac{2^{q+3}(q+1)\mathscr{L}(g)}{c}\Big)^{1/(q+1)} M^{q/(q+1)} \end{aligned}$$

using that $x \wedge y \leq x^{1/(q+1)} y^{q/(q+1)}$. We see that (3.5) holds true on $\{S \geq \mathscr{L}(g)/2\} \cap \Omega_1$.

6. All it remains is to prove that (3.5) holds true on $\{S < a(n,\varepsilon)\mathscr{L}(g)\}$. This is straightforward.

$\square$

Lemma 3.7. *Recall that*

$$H_1 = \sup_{u\in[0,1]} d(g(u), Im\hat{f}), \quad S = \max_{1\leq i\leq n} d\big(g(U_i), Im(\hat{f})\big).$$

For all $n \geq 1$, for all $\varepsilon \in (0,1]$

$$P\bigg(H_1 \leq S + \mathscr{L}(g)\frac{3\log(n+1)}{cn}\Big(1 + \frac{\log^+\big(\frac{1}{\varepsilon \log(n+1)}\big)}{\log(n+1)}\Big)\bigg) \geq 1 - \varepsilon$$

Proof. Since the function $t \mapsto d\big(g(t), Im(\hat{f})\big)$ is $\mathscr{L}(g)$-Lipschitz, one has

$$H_1 \leq S + \mathscr{L}(g) \max_{1\leq i\leq n+1} \big(U_{(i)} - U_{(i-1)}\big)$$

where $U_{(1)} \leq U_{(2)} \leq \cdots \leq U_{(n)}$ are the order statistics associated to $U_1, \ldots, U_n$ (with $U_{(0)} = 0$ and $U_{(n+1)} = 1$). Therefore, it suffices to prove that

$$P\Big(\max_{1\leq i\leq n+1}\Big(U_{(i)} - U_{(i-1)}\Big) \geq \frac{\log(n+1)}{n} t(n,\varepsilon)\Big) \leq \varepsilon$$

with

$$t(n,\varepsilon) = \frac{3}{c}\Big(1 + \frac{\log^+\big(\frac{1}{\varepsilon\log(n+1)}\big)}{\log(n+1)}\Big)$$

From Lemma 3.4, we have

$$\begin{aligned}
P\Big(\max_{1\leq i\leq n+1} U_{(i)} - U_{(i-1)} \geq \frac{\log(n+1)}{n} t(n,\varepsilon)\Big) &\leq \frac{3}{t(n,\varepsilon)(n+1)^{ct(n,\varepsilon)/3-1}\log(n+1)} \\
&\leq \frac{1}{(n+1)^{ct(n,\varepsilon)/3-1}\log(n+1)} \quad \text{since } t(n,\varepsilon) \geq 3 \\
&\leq \varepsilon
\end{aligned}$$

□

The next proposition allows to control the term $H_2 = \sup_{u\in[0,1]} d(\hat{f}(u), \mathrm{Im}g)$.

Proposition 3.4. *On the event where $\hat{L} \leq \mathscr{L}(g) + \delta$, the quantity $H_2 = \sup_{u\in[0,1]} d(\hat{f}(u), \mathrm{Im}g)$ satisfies*

$$H_2 \leq H_1 \left[2 + \frac{\alpha_d d}{\alpha_{d-1}} + \frac{\mathscr{L}(g)(d-1)}{r + H_1}\right] + \delta.$$

Proof. Recall that, by definition of H_1,

$$\mathrm{Im}g \subset \mathrm{Im}\hat{f} \oplus H_1.$$

Assume that $H_2 \geq 2H_1$. (Otherwise, there is nothing to show, since H_2 is of the order H_1.) Let t^* such that $d(\hat{f}(t^*), \mathrm{Im}g) = H_2$. We set $x^* := \hat{f}(t^*)$. Let

$$\hat{f}_0 : t \in [0, t^*] \mapsto \hat{f}(t^* - t), \quad \hat{f}_1 : t \in [t^*, 1] \mapsto \hat{f}(t).$$

Note that $\mathrm{Im}\hat{f}$ is the union of $\mathrm{Im}\hat{f}_0$ and $\mathrm{Im}\hat{f}_1$. Since $\mathrm{Im}g \subset \mathrm{Im}\hat{f} \oplus H_1$ and $\mathrm{Im}g \cap B(x^*, H_2) = \emptyset$, $\mathrm{Im}\hat{f}$ is not included in the ball $B(x^*, H_2 - H_1)$. Hence, there exists a first time of exit from $B(x^*, H_2 - H_1)$ of at least one of the two branches $\hat{f}_0$ and $\hat{f}_1$. When they exist, the exit times are denoted by t_0^* and t_1^* respectively, and we set

$$\hat{h}_0 : t \in [0, t_0^*] \mapsto \hat{f}(t_0^* - t), \quad \hat{h}_1 : t \in [t_1^*, 1] \mapsto \hat{f}(t).$$

1. We start by observe that $\hat{L} \geq H_2 - H_1$, so that $H_2 \leq H_1 + \hat{L}$. Then, on the event where $\hat{L} \leq \mathscr{L}(g) + \delta$, we have
$$H_2 \leq H_1 + \mathscr{L}(g) + \delta.$$

2. a) Now, let us first address the case where only one branch, say $\hat{f}_1$, exits from $B(x^*, H_2 - H_1)$.

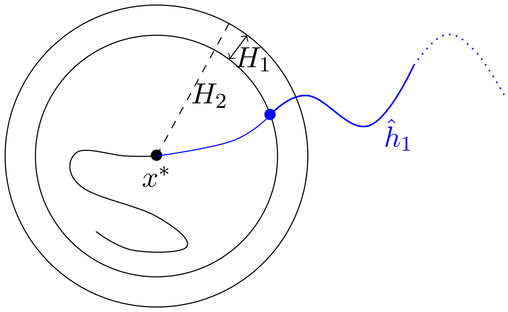

Figure 1: The branch $\hat{f}_1$ exits from the ball $B(x^*, H_2 - H_1)$, the part after first exit is denoted by $\hat{h}_1$.

In this case, $\mathrm{Im}g \subset \mathrm{Im}\hat{h}_1 \oplus H_1$. Thus, $\mathrm{Im}g \oplus r \subset \mathrm{Im}\hat{h}_1 \oplus (r + H_1)$ and consequently

$$\lambda(\mathrm{Im}g \oplus r) \leq \lambda(\mathrm{Im}\hat{h}_1 \oplus (r + H_1))$$

From Federer (1959), since $\mathrm{Im}g$ has a reach larger than r, we have

$$\lambda(\mathrm{Im}g \oplus r) = \mathscr{L}(g)\alpha_{d-1}r^{d-1} + \alpha_d r^d,$$

and from Mosconi and Tilli (2005) (Lemma 4.2), we have

$$\lambda(\mathrm{Im}\hat{h}_1 \oplus (r + H_1)) \leq \mathscr{L}(\hat{h}_1)\alpha_{d-1}(r + H_1)^{d-1} + \alpha_d(r + H_1)^d.$$

Note that $\mathscr{L}(\hat{h}_1) \leq \hat{L} + H_1 - H_2$. On the event where $\hat{L} \leq \mathscr{L}(g) + \delta$, we get

$$\mathscr{L}(g)\left[\alpha_{d-1}r^{d-1} - \alpha_{d-1}(r + H_1)^{d-1}\right] \leq (H_1 - H_2 + \delta)\alpha_{d-1}(r+H_1)^{d-1} - \alpha_d r^d + \alpha_d(r+H_1)^d,$$

that is

$$\begin{aligned}
&H_2\alpha_{d-1}(r + H_1)^{d-1} \\
&\quad\leq (H_1 + \delta)\alpha_{d-1}(r + H_1)^{d-1} + \alpha_d\left[(r + H_1)^d - r^d\right] + \mathscr{L}(g)\alpha_{d-1}\left[(r + H_1)^{d-1} - r^{d-1}\right] \\
&\quad\leq (H_1 + \delta)\alpha_{d-1}(r + H_1)^{d-1} + H_1\alpha_d d(r + H_1)^{d-1} + H_1\mathscr{L}(g)\alpha_{d-1}(d-1)(r + H_1)^{d-2} \\
&\quad\leq H_1\left[(\alpha_{d-1} + \alpha_d d)(r + H_1)^{d-1} + \mathscr{L}(g)\alpha_{d-1}(d-1)(r + H_1)^{d-2}\right] + \delta\alpha_{d-1}(r + H_1)^{d-1}.
\end{aligned}$$

Thus,

$$H_2 \leq H_1\left[1 + \frac{\alpha_d d}{\alpha_{d-1}} + \frac{\mathscr{L}(g)(d-1)}{r + H_1}\right] + \delta.$$

b) Let us turn to the situation where there are two branches, $\hat{h}_0$ and $\hat{h}_1$, exiting from $B(x^*, H_2 - H_1)$. Using again that $\mathrm{Im}g \subset \mathrm{Im}\hat{f} \oplus H_1$ and $\mathrm{Im}g \cap B(x^*, H_2) = \emptyset$, we know that

$$\mathrm{Im}g \subset \left(\mathrm{Im}\hat{h}_0 \oplus H_1\right) \cup \left(\mathrm{Im}\hat{h}_1 \oplus H_1\right)$$

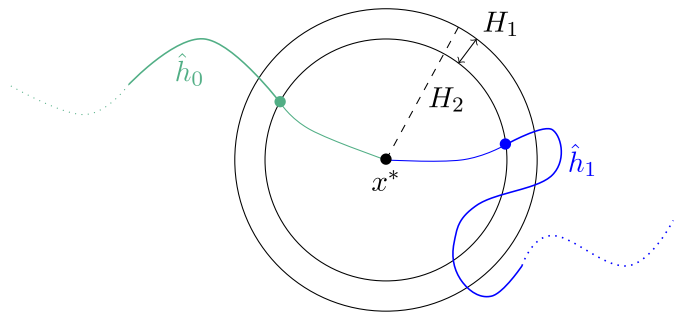

Figure 2: Two branches $\hat{f}_0$ and $\hat{f}_1$ exit from the ball $B(x^*, H_2 - H_1)$, the parts after first exit are denoted by $\hat{h}_0$ and $\hat{h}_1$ respectively.

Now,

$$\mathrm{Im}g \oplus r \subset \big(\mathrm{Im}\hat{h}_0 \oplus (r + H_1)\big) \cup \big(\mathrm{Im}\hat{h}_1 \oplus (r + H_1)\big) \tag{3.6}$$

As already noted above,

$$\lambda(\mathrm{Im}g \oplus r) = \mathscr{L}(g)\alpha_{d-1}r^{d-1} + \alpha_d r^d.$$

Moreover,

$$\begin{aligned}&\lambda\big(\mathrm{Im}\hat{h}_0 \oplus (r + H_1) \cup \mathrm{Im}\hat{h}_1 \oplus (r + H_1)\big)\\ &= \lambda\big(\mathrm{Im}\hat{h}_0 \oplus (r + H_1)) + \lambda\big(\mathrm{Im}\hat{h}_1 \oplus (r + H_1)) - \lambda\big(\mathrm{Im}\hat{h}_0 \oplus (r + H_1) \cap \mathrm{Im}\hat{h}_1 \oplus (r + H_1)\big).\end{aligned}$$

Here,

$$\lambda\big(\mathrm{Im}\hat{h}_0 \oplus (r+H_1)) + \lambda\big(\mathrm{Im}\hat{h}_1 \oplus (r+H_1)) \leq (\mathscr{L}(\hat{h}_0) + \mathscr{L}(\hat{h}_1))\alpha_{d-1}(r+H_1)^{d-1} + 2\alpha_d(r+H_1)^d$$

Since $\mathrm{Im}g \subset \mathrm{Im}\hat{h}_0 \oplus H_1 \cup \mathrm{Im}\hat{h}_1 \oplus H_1$, by connectivity of $\mathrm{Im}g$, there exists $x \in \mathrm{Im}g \cap \mathrm{Im}\hat{h}_0 \oplus H_1 \cap \mathrm{Im}\hat{h}_1 \oplus H_1$. Hence, there exist $x_0 \in \mathrm{Im}\hat{h}_0$ and $x_1 \in \mathrm{Im}\hat{h}_1$ such that $|x - x_0| \leq H_1$ and $|x - x_1| \leq H_1$, so that $\eta := |x_0 - x_1| \leq 2H_1$. Observe that $B(x_0, r + H_1) \subset \mathrm{Im}\hat{h}_0 \oplus (r + H_1)$ and $B(x_1, r + H_1) \subset \mathrm{Im}\hat{h}_1 \oplus (r + H_1)$, so that $B(x_0, r + H_1) \cap B(x_1, r + H_1) \subset \mathrm{Im}\hat{h}_0 \oplus (r + H_1) \cap \mathrm{Im}\hat{h}_1 \oplus (r + H_1)$.

Consequently,

$$\begin{aligned}\lambda\big(\mathrm{Im}\hat{h}_0 \oplus (r + H_1) \cap \mathrm{Im}\hat{h}_1 \oplus (r + H_1)\big) &\geq \lambda\big(B(x_0, r + H_1) \cap B(x_1, r + H_1)\big)\\ &:= V_d(r + H_1, \eta).\end{aligned}$$

Here, $V_d(\rho, c)$ denotes the volume of the intersection of two d-dimensional balls with radius ρ, with their centers c apart. Note that this quantity corresponds to twice the volume of a hyperspherical cap with radius ρ and height $\rho - c/2$.

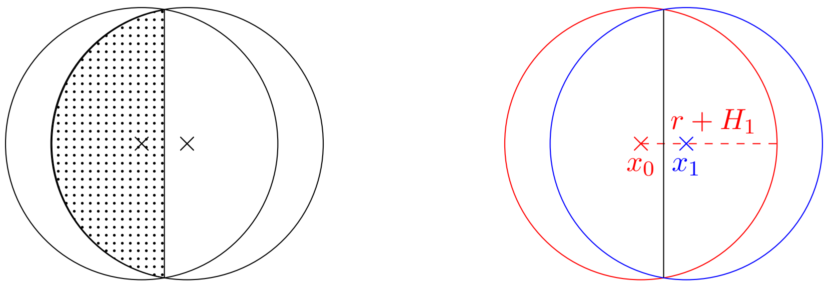

Figure 3: Illustration in dimension 2 of the intersection of two balls: the volume is twice the volume of a spherical cap.

The computation of $V_d(\rho, c)$ may be found for instance in Li (2011):

$$V_d(\rho, c) = 2\alpha_{d-1}\rho^d \int_0^{\arccos(\frac{c}{2\rho})} \sin^d(t)dt.$$

Here, we need a lower bound for

$$V_d(r + H_1, \eta) = 2\alpha_{d-1}(r + H_1)^d \int_0^{\arccos(\frac{\eta}{2(r+H_1)})} \sin^d(t)dt.$$

Let us compute the first and second derivatives of $V_d(r + H_1, \eta)$ with respect to η. We have

$$\begin{aligned}\frac{\partial V_d}{\partial \eta}(r + H_1, \eta) &= \alpha_{d-1}(r + H_1)^{d-1} \arccos'\left(\frac{\eta}{2(r + H_1)}\right) \sin^d\left(\arccos\left(\frac{\eta}{2(r + H_1)}\right)\right) \\ &= -\alpha_{d-1}(r + H_1)^{d-1}\left(1 - \frac{\eta^2}{4(r + H_1)^2}\right)^{\frac{d-1}{2}},\end{aligned}$$

$$\frac{\partial^2 V_d}{\partial \eta^2}(r + H_1, \eta) = \frac{\alpha_{d-1}(r + H_1)^{d-3}(d-1)\eta}{4}\left(1 - \frac{\eta^2}{4(r + H_1)^2}\right)^{\frac{d-3}{2}}.$$

Note that $\frac{\eta^2}{4(r+H_1)^2} < 1$, since $\eta \le 2H_1$. Thus, $\frac{\partial^2 V_d}{\partial \eta^2}(r + H_1, \eta) \ge 0$, and the function $\eta \mapsto \frac{\partial V_d}{\partial \eta}(r+H_1, \eta)$ is nondecreasing. Moreover, by the mean value theorem, there exists $\zeta \in (0, \eta)$ such that

$$\begin{aligned}V_d(r + H_1, \eta) &= V_d(r + H_1, 0) + \eta\frac{\partial V_d}{\partial \eta}(r + H_1, \zeta) \\ &= \lambda(B(0, r + H_1)) + \eta\frac{\partial V_d}{\partial \eta}(r + H_1, \zeta).\end{aligned}$$

Consequently, by monotonicity of $\eta \mapsto \frac{\partial V_d}{\partial \eta}(r + H_1, \eta)$, we obtain

$$\begin{aligned}V_d(r + H_1, \eta) &\ge \lambda(B(0, r + H_1)) + \eta\frac{\partial V_d}{\partial \eta}(r + H_1, 0) \\ &= \alpha_d(r + H_1)^d - \eta\alpha_{d-1}(r + H_1)^{d-1}.\end{aligned}$$

Finally,

$$\lambda\big(\mathrm{Im}\hat{h}_0 \oplus (r+H_1) \cap \mathrm{Im}\hat{h}_1 \oplus (r+H_1)\big) \geq \alpha_d (r+H_1)^d - \eta\alpha_{d-1}(r+H_1)^{d-1}.$$

Hence,

$$\begin{aligned}
&\lambda\big(\mathrm{Im}\hat{h}_0 \oplus (r+H_1) \cup \mathrm{Im}\hat{h}_1 \oplus (r+H_1)\big)\\
&\quad \leq \mathscr{L}(\hat{h}_0)\alpha_{d-1}(r+H_1)^{d-1} + \alpha_d(r+H_1)\\
&\qquad + \mathscr{L}(\hat{h}_1)\alpha_{d-1}(r+H_1)^{d-1} + \alpha_d(r+H_1)\\
&\qquad - \alpha_d(r+H_1)^d + \eta\alpha_{d-1}(r+H_1)^{d-1}\\
&\quad \leq (\mathscr{L}(\hat{h}_0) + \mathscr{L}(\hat{h}_1))\alpha_{d-1}(r+H_1)^{d-1} + \alpha_d(r+H_1)^d\\
&\qquad + \eta\alpha_{d-1}(r+H_1)^{d-1}\\
&\quad \leq (\hat{L} - 2(H_2 - H_1))\alpha_{d-1}(r+H_1)^{d-1} + \alpha_d(r+H_1)^d\\
&\qquad + 2H_1\alpha_{d-1}(r+H_1)^{d-1}.
\end{aligned}$$

On the event where $\hat{L} \leq \mathscr{L}(g) + \delta$, using the inclusion (3.6), we obtain

$$\begin{aligned}
&\mathscr{L}(g)\left[\alpha_{d-1}r^{d-1} - \alpha_{d-1}(r+H_1)^{d-1}\right]\\
&\quad \leq \delta\alpha_{d-1}(r+H_1)^{d-1} - 2H_2\alpha_{d-1}(r+H_1)^{d-1} + 2H_1\alpha_{d-1}(r+H_1)^{d-1}\\
&\qquad\qquad + \alpha_d(r+H_1)^d - \alpha_d r^d + 2H_1\alpha_{d-1}(r+H_1)^{d-1},
\end{aligned}$$

that is

$$\begin{aligned}
&2H_2\alpha_{d-1}(r+H_1)^{d-1}\\
&\quad \leq \delta\alpha_{d-1}(r+H_1)^{d-1} + \mathscr{L}(g)\alpha_{d-1}\left[(r+H_1)^{d-1} - r^{d-1}\right] + \alpha_d\left[(r+H_1)^d - r^d\right]\\
&\qquad + 4H_1\alpha_{d-1}(r+H_1)^{d-1}\\
&\quad \leq \delta\alpha_{d-1}(r+H_1)^{d-1} + \mathscr{L}(g)\alpha_{d-1}H_1(d-1)(r+H_1)^{d-2} + \alpha_d H_1 d(r+H_1)^{d-1}\\
&\qquad + 4H_1\alpha_{d-1}(r+H_1)^{d-1}.
\end{aligned}$$

Hence,

$$H_2 \leq H_1\left[2 + \frac{\alpha_d d}{2\alpha_{d-1}} + \frac{\mathscr{L}(g)(d-1)}{2(r+H_1)}\right] + \frac{\delta}{2}.$$

□

3.3 Proof of Proposition 3.1

The next lemma, based on the computation of the volume of r-enlargements, allows to link $\hat{L}$ and $\mathscr{L}(g)$. We recall the notation $H_1 = \sup_{u\in[0,1]} d(g(u), \mathrm{Im}\hat{f})$.

Lemma 3.8. *We have*

$$(\mathscr{L}(g) - \hat{L})\alpha_{d-1}\eta^{d-1} \leq H_1\left[\hat{L}\alpha_{d-1}(d-1)(H_1+\eta)^{d-2} + \alpha_d d(H_1+\eta)^{d-1}\right].$$

Proof. By definition of H_1,

$$\mathrm{Im} g \subset \mathrm{Im}\hat{f} \oplus H_1.$$

Thus,

$$\mathrm{Im} g \oplus r \subset \mathrm{Im}\hat{f} \oplus (H_1 + r).$$

According to Federer (1959), the volumes of both enlargements satisfy:

$$\lambda(\mathrm{Im} g \oplus r) = \mathscr{L}(g)\alpha_{d-1} r^{d-1} + \alpha_d r^d,$$

$$\lambda(\mathrm{Im}\hat{f} \oplus (H_1 + r)) \leq \hat{L}\alpha_{d-1}(H_1 + r)^{d-1} + \alpha_d (H_1 + r)^d.$$

Hence,

$$\mathscr{L}(g)\alpha_{d-1} r^{d-1} + \alpha_d r^d \leq \hat{L}\alpha_{d-1}(H_1 + r)^{d-1} + \alpha_d (H_1 + r)^d,$$

so that

$$(\mathscr{L}(g) - \hat{L})\alpha_{d-1} r^{d-1} \leq H_1 \left[\hat{L}\alpha_{d-1}(d-1)(H_1 + r)^{d-2} + \alpha_d d (H_1 + r)^{d-1}\right].$$

□

Let us assume that m_q, α and δ tend to 0, as n tends to infinity. Then, H_1 tends to 0 and the difference $\mathscr{L}(g) - \hat{L}$ converges to a limit which is less than or equal to 0. Hence, on the event where $\hat{L} \leq \mathscr{L}(g) + \delta$, the selected length $\hat{L}$ converges in probability to the length of the objective curve $\mathscr{L}(g)$.

References

E. Aamari and C. Levrard. Stability and minimax optimality of tangential delaunay complexes for manifold reconstruction. *Discrete & Computational Geometry*, 59(4):923–971, 2018.

E. Aamari and C. Levrard. Nonasymptotic rates for manifold, tangent space and curvature estimation. *The Annals of Statistics*, 47(1):177–204, 2019.

E. Aamari, C. Aaron, and C. Levrard. Minimax boundary estimation and estimation with boundary. *Bernoulli*, 2023.

S. Delattre and A. Fischer. Estimation via length-constrained curves under small noise. 2024.

H. Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93: 418–491, 1959.

C. R. Genovese, M. Perone-Pacifico, I. Verdinelli, and L. Wasserman. The geometry of nonparametric filament estimation. *Journal of the American Statistical Association*, 107:788–799, 2012a.

C. R. Genovese, M. Perone-Pacifico, I. Verdinelli, and L. Wasserman. Minimax manifold estimation. *Journal of Machine Learning Research*, 13:1263–1291, 2012b.

C. R. Genovese, M. Perone-Pacifico, I. Verdinelli, and L. Wasserman. Manifold estimation and singular deconvolution under Hausdorff loss. *The Annals of Statistics*, 40(2):941–963, 2012c.

B. Kégl, A. Krzyżak, T. Linder, and K. Zeger. Learning and design of principal curves. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22:281–297, 2000.

S. Li. Concise formulas for the area and volume of a hyperspherical cap. *Asian J. Math. Stat.*, 4(1):66–70, 2011. ISSN 1994-5418,2077-2068. doi: 10.3923/ajms.2011.66.70. URL https://doi.org/10.3923/ajms.2011.66.70.

S. J. N. Mosconi and P. Tilli. Γ-convergence for the irrigation problem. *Journal of Convex Analysis*, 12:145–158, 2005.