



HAL
open science

Non-local estimators: a new class of multigrid convergent length estimators

Loïc Mazo, Etienne Baudrier

► **To cite this version:**

Loïc Mazo, Etienne Baudrier. Non-local estimators: a new class of multigrid convergent length estimators. 2014. hal-01058445v1

HAL Id: hal-01058445

<https://hal.science/hal-01058445v1>

Preprint submitted on 26 Aug 2014 (v1), last revised 11 Aug 2015 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Non-local estimators: a new class of multigrid convergent length estimators

Loïc Mazo, Étienne Baudrier
ICube, University of Strasbourg, CNRS

August 26, 2014

Abstract

An interesting property for curve length digital estimators is the convergence toward the continuous length and the associate convergence speed when the resolution tends to infinity. On the one hand, DSS based estimators have been proved to converge but only under some convexity and smoothness or polygonal assumptions. On the other hand, we have introduced in a previous paper the sparse estimators and we proved their convergence for Lipschitz functions without convexity assumption. Here, we introduce a wider class of estimators, the *non-local estimators*, that intends to gather sparse estimators and DSS based estimators. We prove their convergence and give an error upper bound for a large class of functions.

1 Introduction

We focus in this paper on one classical digital problem: the length estimation. The problem is to estimate the length of a continuous curve S knowing a digitization of S . As information is lost during the digitization step, there is no reliable estimation without *a priori* knowledge. From a theoretical point of view, a classical criterion to evaluate the quality of a geometric feature estimator is the possession, or not, of the *(multigrid) convergence* property, that is the estimation convergence toward the continuous curve feature when the resolution tends to infinity. The local estimators based on a segmentation of the digital curve in patterns whose size is a constant that does not depends upon the resolution do not satisfy the convergence property even for straight line segments [1]. The adaptive estimators based on a segmentation in Maximal Digital Straight Segments (MDSS) or based on a Minimum Length Polygon (MLP) satisfy the convergence property for smooth, or polygonal, closed simple curves under assumption of convexity [2]. The semi-local estimators [3], and the sparse estimators [4], both based on a segmentation of the curve in patterns

whose size only depends upon the resolution, verifies the convergence property¹ without convexity hypothesis, for smooth curves of class C^2 with the former and for Lipschitz curves with the latter. We present here a new class of length estimators, the *non-local estimators*, that aims to encompass the sparse estimators and the MDSS based estimators.

The paper is organized as follows. In Section 2, some necessary notations and conventions are recalled, then the existing estimators and their convergence properties are detailed. In Section 3, the non-local estimators are defined and the multigrid convergence property is proved for Lipschitz functions under some assumptions satisfied by sparse estimators and MDSS based estimators. Furthermore, an upper bound on the error of the estimator is exhibited for a wide subclass of the Lipschitz functions. Section 4 provides some illustrations of the obtained convergence speed lower bounds and a comparison of the estimations for different kind of non-local estimators. Section 5 concludes the article and gives directions for future works.

2 Background

2.1 Discretization models

In this work, we have restricted ourselves to the digitization of function graphs. So, let us consider a continuous function $g : [a, b] \rightarrow \mathbb{R}$ ($a < b$), its graph $\mathcal{C}(g) = \{(x, g(x)) \mid x \in [a, b]\}$ and a positive real number r , the *resolution*. We assume to have an orthogonal grid in the Euclidean space \mathbb{R}^2 whose set of grid points is $h\mathbb{Z}^2$ where $h = 1/r$ is the *grid spacing*.

The common methods to model the digitization of the graph $\mathcal{C}(g)$ at the resolution r are closely related to each others. In this paper, we assume an *object boundary quantization* (OBQ). This method associates to the graph $\mathcal{C}(g)$ the *h-digitization set*

$$\mathcal{D}^O(g, h) = \{(kh, \lfloor \frac{g(kh)}{h} \rfloor h) \mid k \in \mathbb{Z} \text{ and } kh \in [a, b]\}$$

where $\lfloor \cdot \rfloor$ denotes the floor function. The set $\mathcal{D}^O(g, h)$ contains the uppermost grid points which lie in the hypograph of g , hence it can be understood as a part of the boundary of a solid object. Provided the slope of g is limited by 1 in modulus, $\mathcal{D}^O(g, h)$ is an 8-connected digital curve. Observe that if g is a function of class C^1 such that the set $\{x \in [a, b] \mid |g'(x)| = 1\}$ is finite, then by symmetries on the graph $\mathcal{C}(g)$, it is possible to come down to the case where $|g'| \leq 1$. Nevertheless, in this article, we make no assumption on the slope of the function g .

In the sequel of the article, for any function $f : [a, b] \rightarrow \mathbb{R}$, $L(f)$ denotes the

¹Actually, the convergence of a semi-local estimator depends upon some choice made in its definition (see Theorem 1).

length of the graph $\mathcal{C}(f)$ according to Jordan's definition of length:

$$L(f) = \sup_{a=x_0 < x_1 < \dots < x_n = b} \sum_{i=1}^n \sqrt{(x_i - x_{i-1})^2 + (f(x_i) - f(x_{i-1}))^2}$$

where the supremum is taken over all possible partitions of $[a, b]$ and n is unbounded.

2.2 Local estimators

Local length estimators (see [5] for a short review) are based on parallel computations of the length of fixed size segments of a digital curve. For instance, an 8-connected curve can be split into 1-step segments. For each segment, the computation returns 1 whenever the segment is parallel to the axes and $\sqrt{2}$ when the segment is diagonal. Then all the results are added to give the curve length estimation.

This kind of local computation is the oldest way to estimate the length of a curve and has been widely used in image analysis. Nevertheless, it has not the convergence property. In [1], Kulkarni *et al.* introduce a general definition of local length estimation with sliding segments and prove that such computations cannot give a convergent estimator for straight lines whose slope is small (less than the inverse of the size of the sliding segment). In [6], a similar definition of local length estimation is given with disjoint segments. Again, it is shown that the estimator failed to converge for straight lines (with irrational slopes). This behavior is experimentally confirmed in [2] on a test set of five closed curves. Moreover, the non-convergence is established in [7, 8] for almost all parabolas.

2.3 Adaptive estimators: DSS and MLP

Adaptive length estimators gather estimators relying on a segmentation of the discrete curve that depends on each point of the curve: a move on a point can change the whole segmentation. Unlike local estimators, it is possible to prove the convergence property of adaptive length estimators under some assumptions. Adaptive length estimators include two families of length estimators, namely the Maximal Digital Straight Segment (MDSS) based length estimators and the Minimal Length Polygon (MLP) based length estimators.

Definition and properties of MDSS can be found in [9, 10, 2]. Efficient algorithms have been developed for segmenting curves or function graphs into MDSS and to compute their characteristics in a linear time [9, 11, 10]. The decomposition in MDSS is not unique and depends on the start-point of the segmentation and on the curve travel direction. The convergence property of MDSS estimators has been proved for convex polygons whose MDSS polygonal approximation² is also convex [12, Th. 13 and the proof³]: given a convex

²Though the digitization of a convex set is digitally convex, it does not mean that a polygonal curve related to a convex polygonal curve via a MDSS segmentation process is also convex.

³The hypothesis on the convexity of the MDSS polygone is not assumed in the statement

polygon \mathcal{C} and a grid spacing h (below some threshold), the error between the estimated length $L_{\text{est}}(\mathcal{C}, h)$ and the true length of the polygon $L(\mathcal{C})$ is such that

$$|L(\mathcal{C}) - L_{\text{est}}(\mathcal{C}, h)| \leq (2 + \sqrt{2})\pi h. \quad (1)$$

It must be noticed that there exists a wrong version of the above equation in the literature ⁴. Empirical MDSS multigrid convergence has also been tested in [2, 15] on smooth nonconvex planar curves. The obtained convergence has order 1 as in the convex polygonal case. Nevertheless MDSS multigrid convergence has not been proved under these assumptions. Another way to obtain an estimation of the length of a curve using MDSS is to take the slopes of the MDSSs to estimate the tangent directions and then to compute the length by numerical integration [16, 2, 17]. The estimation is unique and has been proved to be multigrid convergent for smooth curves (of class C^2 with bounded curvature in [2], of class C^3 with strictly positive curvature in [17]). The convergence order is a $\mathcal{O}(h^{\frac{1}{3}})$ $\frac{1}{3}$ in [17] and thus, worse than (1).

Let \mathcal{C} be a simple closed curve lying in-between two polygonal curves γ_1 and γ_2 . Then, there is a unique polygon, the MLP, whose length is minimal between γ_1 and γ_2 . The length of the MLP can be used to estimate the length of the curve \mathcal{C} . At least two MLP based length estimators have been described and proved to be multigrid convergent for convex, smooth or polygonal, simple closed curves, the grid-continua MLP algorithm (GC-MLP) proposed in [18] and the Approximation Saussage MLP (AS-MLP) introduced in [19]. For both of them, and for a given grid spacing h , the error between the estimated length $L_{\text{est}}(\mathcal{C}, h)$ and the true length of the curve $L(\mathcal{C})$ is a $O(h)$:

$$|L(\mathcal{C}) - L_{\text{est}}(\mathcal{C}, h)| \leq Ah$$

where $A = 8$ for GC-MLP and $A \approx 5.844$ for AS-MLP.

On the one hand, as estimators described in this section are adaptive, the convergence theorems rely on strong hypotheses, the proofs are difficult to establish and often incomplete. On the other hand, the study of the MDSS in [15] shows that the MDSS size tends to 0 and their discrete length tends toward infinity as the grid step tends to 0. Thereby, one could ask whether combining a local estimation with an increasing window size as the resolution grows would give a convergent estimator under more general assumptions and/or with simpler proofs of convergence. The following sections explore this question.

of the theorem but it appears in the proof.

⁴The formulation of the right hand side of (1) in the literature is $(2\varepsilon_{DSS} + \sqrt{2})\pi h$ where ε_{DSS} is a bound for the Hausdorff distance between a real straight segment and its discretization, expressed in the unit given by the grid-spacing and it is related to the chord property of the straight lines established by Rosenfeld [13]. Accordingly, in [12], it is said that "the 'classical' value of ε_{DSS} is 1". A few month later, in [14], the authors chose to express the same constant ε_{DSS} in absolute length unit, so they claim that "its 'classical' value is $1/r$ " but they forgot to update the right hand side of the majorization. Afterward, the mistake propagated to [2] and led to an erroneous conclusion.

2.4 Semi-local length estimators

The notion of semi-local estimator appears in [3]. Here, we give a slightly different presentation of this notion. At a given resolution, a semi-local estimator resembles a local estimator: it can be implemented via a parallel computation, each processor handling a fixed size segment of the curve. Nevertheless, in the framework of semi-local estimation, the processors must be aware of the resolution on which the size of the segments depends.

More formally, a pattern of size N is a discrete function from $[0, N] \cap \mathbb{Z}$ to \mathbb{Z} . The diameter of a pattern ω is the real $\|\omega\| = \sqrt{N^2 + \omega(N)^2}$. We denote by \mathcal{P}_N the set of patterns of size N and by \mathcal{P} the set of all patterns: $\mathcal{P} = \bigcup_{N>0} \mathbb{Z}^{[0, N] \cap \mathbb{Z}}$.

A *semi-local estimator* is a pair (H, p) where

- $H: (0, \infty) \rightarrow \mathbb{N}^*$ gives the size of the patterns given a grid spacing h and

$$\lim_{h \rightarrow 0} H(h) = +\infty, \quad (2)$$

$$\lim_{h \rightarrow 0} hH(h) = 0, \quad (3)$$

- $p: \mathcal{P} \rightarrow [0, \infty)$ gives the estimated feature (here, the length) associated to a pattern.

Let $g: [a, b] \rightarrow \mathbb{R}$ be a rectifiable function. At a given grid spacing h corresponding to a resolution r , the digitization $\mathcal{D}^O(g, h)$ of the curve $\mathcal{C}(g)$ is segmented in N_r patterns $\omega_1, \dots, \omega_{N_r}$ of size $H(h)$ and, possibly, a rest ω_* whose size is less than $H(h)$. Then, the length of the curve $\mathcal{C}(g)$ is estimated by

$$L^{\text{SL}}(g, h) = h \sum_{i=1}^{N_r} p(\omega_i).$$

We can see that the patterns of a semi-local length estimator have a discrete size that tends to infinity as the resolution grows to infinity (Equation 2) while their true size tends to zero (Equation 3). Under these constraints, and a constraint on the function p , it is proved in [3] that the semi-local length estimators are multigrid convergent for functions of class C^2 .

Theorem 1 ([3, Prop. 1]). *Let (H, p) be a semi-local estimator such that*

$$\max \{p(\omega) - \|\omega\| \mid \omega \in \mathcal{P}_k\} = o(k) \text{ as } k \rightarrow +\infty.$$

Then, for any function $g \in C^2([a, b])$, the estimation $L^{\text{SL}}(g, h)$ converges toward the length of the curve $\mathcal{C}(g)$. Furthermore, if the term $o(k)$ in the third hypothesis is a constant and $H(h) = \Theta(h^{-\frac{1}{2}})$, then $L(g) - L^{\text{SL}}(g, h) = O(h^{\frac{1}{2}})$.

In the above theorem, the hypothesis states that the function p applied to a pattern ω must return a value close to its diameter. This has led us to define a subfamily of the semi-local length estimators that we present in the next section.

2.5 Sparse length estimators

The notion of sparse estimator is introduced in [4]. They are a subclass of the semi-local estimators for which the estimation of a pattern length is exactly its diameter. Hence, the information given by the points inside a pattern, but its extremities, is discarded. This justifies the name given to this class of estimators.

As semi-local estimators, sparse estimators can be defined by a *sparcity function* $H:]0, +\infty[\rightarrow \mathbb{N}^*$ that verifies the two assertions

$$\lim_{h \rightarrow 0} H(h) = +\infty \text{ and } \lim_{h \rightarrow 0} hH(h) = 0.$$

For any function $g: [a, b] \rightarrow \mathbb{R}$ and any grid spacing $h = 1/r$, the digitization $\mathcal{D}^O(g, h)$ of the curve $\mathcal{C}(g)$ is segmented in N_r patterns $\omega_0, \dots, \omega_{N_r-1}$ of size $H(h)$ and a last pattern ω_{N_r} whose size is less than, or equal to, $H(h)$. The *sparse estimator with parameter H* of the length of the curve $\mathcal{C}(g)$ is defined by

$$L^{Sp}(g, h) = h \sum_{i=0}^{N_h} \|\omega_i\|^2.$$

Sparse estimators have the following convergence properties.

Theorem 2 ([4]). *Let H be a sparsity function and $g: [a, b] \rightarrow \mathbb{R}$ a Lipschitz function. Then, the estimator L^{Sp} converges toward the length of the curve $\mathcal{C}(g)$. Furthermore,*

- if g is of class C^2

$$L(g) - L^{Sp}(g, h) = O(hH(h)) + O\left(\frac{1}{H(h)}\right),$$

- if g is C^2 and concave,

$$L(g) - L^{Sp}(g, h) = O(h^2H(h)^2) + O\left(\frac{1}{H(h)^2}\right),$$

- if g is C^2 with everywhere strictly negative curvature and $h^{\frac{1}{2}} = o(hH(h))$,

$$L(g) - L^{Sp}(g, h) = O(h^2H(h)^2) + O(h).$$

On the one hand, with semi-local and sparse estimators, the pattern size is constant for a given resolution. This is important for an algorithmic point of view. Nevertheless, it does not really matter to prove the convergence and Equations (2), (3) could be expressed in terms of means. On the other hand, it has been proved in [15], under some hypotheses, that the average size of the maximal digital straight segments of a contour verify eqs. (2) and (3). This led us to define a new family of length estimators that is presented in the next section.

3 Non-local estimators

We introduce a new class of length estimators that aims to gather sparse estimators and adaptive estimators. To do so, we need to relax the hypothesis on the length of the codes ω_i by allowing variable lengths. We need also to allow very large patterns when the curve is close to straight segments.

3.1 Definition

We recall that for any non-zero real number α , the generalized mean of parameter α , or α -mean, of a finite sequence of positive numbers $(x_i)_{i=0}^n$ is defined by

$$M_\alpha((x_i)_{i=0}^n) = \left(\frac{1}{n} \sum_{i=0}^n x_i^\alpha \right)^{\frac{1}{\alpha}}.$$

Furthermore, $M_{+\infty}((x_i)_{i=0}^n) = \sup((x_i)_{i=0}^n)$ and $M_{-\infty}((x_i)_{i=0}^n) = \inf((x_i)_{i=0}^n)$. For any $\alpha, \beta \in \overline{\mathbb{R}}$, one has⁵

$$\alpha < \beta \implies M_\alpha((x_i)_{i=0}^n) \leq M_\beta((x_i)_{i=0}^n).$$

When σ is a partition of some interval I , we write $M_\alpha(\sigma)$ for the α -mean of the σ subinterval length sequence. We also write $c_v(\sigma)$ for the coefficient of variation of the σ subinterval length sequence. Recall that the coefficient of variation is the ratio of the standard deviation to the arithmetic mean and that $1 + c_v(\sigma)^2 = M_2(\sigma)^2 / M_1(\sigma)$.

We are now able to give the definition of the non-local estimators. An illustration of the definition is given Figure 1.

Definition 1 (Pattern function). *Let I be a closed real interval. Let $\alpha, \beta \in \overline{\mathbb{R}}$.*

- *A pattern function on I is a function that maps a resolution r to a partition of $rI \cap \mathbb{Z}$.*
- *An α -pattern function is a pattern function such that*

$$\lim_{r \rightarrow +\infty} M_\alpha(\mathcal{A}(r)) = +\infty, \tag{4}$$

- *An (α, β) -pattern function is an α -pattern function such that*

$$M_\beta(\mathcal{A}(r)) = o(r) \quad \text{as } r \rightarrow 0. \tag{5}$$

Definition 2 (Non-local estimator). *Let $g: I \rightarrow \mathbb{R}$ be a rectifiable function and \mathcal{A} be an α -pattern function on I ($\alpha \in \overline{\mathbb{R}}$). The non-local length estimation with parameter \mathcal{A} of the length of the curve $\mathcal{C}(g)$ at the resolution r , denoted $L^{\text{NL}}(g, \mathcal{A}, r)$, is given by the length of the polyline $(P_a^r)_{a \in \mathcal{A}(r)}$ defined by $P_a^r = (a, \lfloor rg(a) \rfloor / r)$.*

⁵ $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

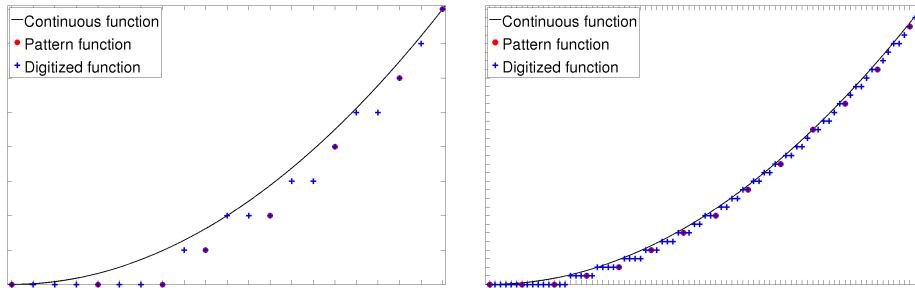


Figure 1: Non-local estimation at two resolutions. The red points are the vertices of the polyline whose length yields the non local estimation of the continuous black curve length at the given resolutions.

A non-local multigrid length estimator is a function that maps a rectifiable function g together with a resolution r to a non-local length estimation $L^{\text{NL}}(g, \mathcal{A}_g, r)$ where \mathcal{A}_g is some α -pattern function. If, besides, \mathcal{A}_g is an (α, β) -pattern function for any g ($\beta \in \overline{\mathbb{R}}$), we say that \mathcal{E} is an M-sparse multigrid length estimator.

In other words, at a given resolution r , a non-local estimator (NLE) splits a discrete curve in patterns in such a way that the pattern discrete sizes tend to infinity, in mean, as r tends to infinity. A M-sparse estimator is a non-local estimator whose patterns have absolute size tending to 0, in mean, as r tends to infinity.

Here we detail the inclusion relations between the non-local estimators and the estimators described in Section 2.

Local estimators The patterns associated to a local estimator have a constant size that does not depend on the resolution. Thereby, the local estimators are not non-local estimators (this justifies the name given to the class of estimators introduced in this paper).

Semi-local and sparse estimators The subinterval-length sequence for sparse estimators is $(H(h), \dots, H(h), rem)$ where $rem \leq H(h)$, $H(h)$ tends to infinity and $hH(h)$ tends to 0 as the grid spacing h tends to 0 (*i.e.*, $r \rightarrow +\infty$). Thus, it is plain that they are M-sparse length estimators. As a matter of fact, M-sparse estimators are derived from sparse estimators by replacing the fixed size patterns and their limit properties by variable ones with in mean limit properties (the 'M' in M-sparse stands for mean).

In the case of semi-local estimators, it is also true that the discrete size of each pattern tends to infinity and that their absolute size tends to 0. Nevertheless, the semi-local length estimators as defined in [3] do not exactly comply to the rule given for the computation of the length in Definition 2 for two reasons.

Firstly, they withdraw the last pattern if it has not the expected size, $H(h)$. Secondly, the result of the estimation is not necessarily based on a polyline because of the length function p (even if it is asymptotically the case when the hypothesis of Theorem 1 is satisfied). Then, strictly speaking, the semi-local estimators are not non-local estimators though they are very close to them (and to M-sparse estimators) when Theorem 1 applies.

MDSS based estimators Though it has not been proved in the same framework nor in the general case, it is likely that MDSS based length estimators are non-local estimators, under the assumption that the length of a MDSS is estimated with its diameter (as it is the case the DGtal library [20]). Indeed, in [15] it is shown that, in the framework of Gauss digitization and for convex shapes with C^3 boundaries and everywhere strictly positive curvature, the average discrete size of all 4-connected maximal segments defined on a discrete boundary is between a $\Theta(r^{\frac{1}{3}})$ and a $\Theta(r^{\frac{1}{3}} \ln(r))$. Our experiments, even with the damped sinusoids s_1 , s_2 and the fractal function f (see Section 4), suggests that this bound could be valid for an MDSS segmentation of a discrete function graph. This hypothesis must now be theoretically confirmed. On the other side, it is plain that a MDSS based estimator is not a M-sparse estimator. For instance, with an affine function it use a single pattern at any resolution.

3.2 Convergence

In this section, we give some sufficient conditions under which the non-local length estimators are convergent for Lipschitz functions. Moreover, Theorem 3 gives a bound on the error at resolution r for Lipschitz functions whose derivative is k -Lipschitz on any interval included in their domain ($k > 0$).

Notations In the remainder of the article, we use the following notations. $I = [a, b]$ is an interval of \mathbb{R} with a non-empty interior and, for any $r > 0$, A^r , resp. B^r , is the smallest, resp. largest, integer in rI . $g: I \rightarrow \mathbb{R}$ is a Lipschitz function whose derivative is denoted g' (from Rademacher's theorem, g is differentiable almost everywhere). The functions g^a , g^r , g^b are resp. the restrictions of the function g to the intervals $[a, A^r h]$, $[A^r h, B^r h]$, $[B^r h, b]$ where h is the grid spacing ($h = \frac{1}{r}$). We also define the function $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ by $\varphi(x) = \sqrt{1+x^2}$. Thus, one has $L(g) = \int_{[a,b]} \varphi \circ g'$.

For any pattern function \mathcal{A} on I and any $r > 0$, we write M_α , resp. c_v , instead of $M_\alpha(\mathcal{A}(r))$, resp. $c_v(\mathcal{A}(r))$, when there is no ambiguity, the number of subintervals in the partition $\mathcal{A}(r)$ is denoted $N_{\mathcal{A}}^r$ and we define two piecewise affine functions, $g_{\mathcal{A}}^r$ and $\lfloor g_{\mathcal{A}}^r \rfloor$, whose graphs are respectively

- the polyline linking the points $(a_i h, g(a_i h))_{i=0}^{N_{\mathcal{A}}^r}$
- the polyline linking the points $(a_i h, \lfloor \frac{g(a_i h)}{h} \rfloor h)_{i=0}^{N_{\mathcal{A}}^r}$

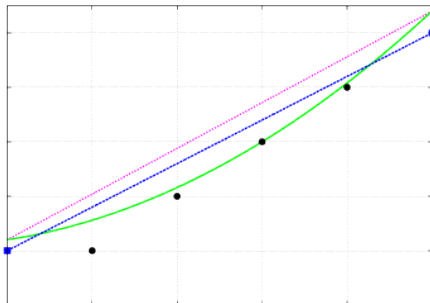


Figure 2: The two main parts of the estimation error: the curve g (in green, solid) to its chord $g_{\mathcal{A}}^r$ (in magenta, dotted-dashed) then the curve chord to the chord $\lfloor g_{\mathcal{A}}^r \rfloor$ (in blue, dashed) of the digitized curve $\mathcal{D}^{\text{O}}(g, h)$ (black points).

where $(a_i)_{i=0}^{N_{\mathcal{A}}^r} = \mathcal{A}(r)$.

Figure 2 shows the three functions g , $g_{\mathcal{A}}^r$, $\lfloor g_{\mathcal{A}}^r \rfloor$ on a subinterval of \mathcal{A} .

The proof of Theorem 3 can be split in three parts. Proposition 1 gives a bound on the error due to the ignorance of the exact abscissas of the curve extremities. Proposition 2 and Proposition 3 evaluate the difference between the length of the curves $\mathcal{C}(g^r)$ and $\mathcal{C}(g_{\mathcal{A}}^r)$ for a given pattern function \mathcal{A} under different assumptions. Given two pattern functions \mathcal{A} and \mathcal{B} , Proposition 4 evaluates the difference between the length of the curves $\mathcal{C}(g_{\mathcal{A}}^r)$ and $\mathcal{C}(\lfloor g_{\mathcal{B}}^r \rfloor)$. The reason to use two distinct α -pattern functions \mathcal{A} and \mathcal{B} comes from the need to deal with non-local estimators as MDSS that produce patterns whose absolute size does not tends to zero unlike M-sparse estimator patterns.

From the Lipschitz hypothesis on g , we derive immediately a bound on the errors due to the loss of the true left and right extremities of the curve $\mathcal{C}(g)$.

Proposition 1. *For any k -Lipschitz function g , we have*

$$L(g^a) + L(g^b) \leq 2\varphi(k)h.$$

Proof. Since g is k -Lipschitz, the slope of any chord of $\mathcal{C}(g)$ is less than k in modulus. It follows that the length of any polyline fitting $\mathcal{C}(g)$ on a subinterval $[c, d]$ of $[a, b]$ is bounded by $\varphi(k)(d - c)$. Then, according to Jordan's definition of arc length, we get

$$L(g|_{[c,d]}) \leq \varphi(k)(d - c) \tag{6}$$

where $g|_{[c,d]}$ denotes the restriction of the function g to the interval $[c, d]$. In particular, $L(g^a) \leq \varphi(k)(A^r h - a) \leq \varphi(k)h$ and $L(g^b) \leq \varphi(k)(b - B^r h) \leq \varphi(k)h$. \square

We now look at the difference between the length of $\mathcal{C}(g^r)$ and the length of the polyline $\mathcal{C}(g_{\mathcal{A}}^r)$.

Proposition 2. For any pattern function \mathcal{A} such that $M_{+\infty}(\mathcal{A}(r)) = o(r)$ as $r \rightarrow +\infty$ and any Lipschitz function g , we have

$$\lim_{r \rightarrow +\infty} L(g_{\mathcal{A}}^r) - L(g^r) = 0.$$

Proof. For any function f defined on an interval J and any partition σ of J , we note $L_{\sigma}(f)$ the length of the polyline interpolating f related to the partition σ . Remember that, from the Jordan's definition of arc length and the triangle inequality, if σ and σ' are two partitions of the interval J and f is a function defined on J , then

$$\sigma \subseteq \sigma' \implies L_{\sigma}(f) \leq L_{\sigma'}(f) \leq L(f).$$

Let $k > 0$ be a Lipschitz constant for g on I . Let $\varepsilon > 0$ and σ a partition of I such that $L(g) - L_{\sigma}(g) < \varepsilon/2$. We denote by n the number of subintervals of the partition σ . Since $M_{+\infty} = o(r)$ as $r \rightarrow +\infty$, there exists a positive real r_0 such that

$$\forall r > r_0, \frac{M_{+\infty}}{r} < \frac{\varepsilon}{2(n-1)(\varphi(k)-1)}.$$

Let $r > r_0$ and $h = \frac{1}{r}$. We set $\sigma_0 = (\sigma \cup \{A^r h, B^r h\}) \cap [A^r h, B^r h]$. It can easily be seen that

$$L(g^r) - L_{\sigma_0}(g^r) < L(g) - L_{\sigma}(g) < \varepsilon/2.$$

We set $\sigma_{0,r} = \sigma_0 \cup \mathcal{A}(r)$. Firstly, we observe that

$$L(g^r) - L_{\sigma_{0,r}}(g^r) \leq L(g^r) - L_{\sigma_0}(g^r) \leq \frac{\varepsilon}{2}$$

for the partition $\sigma_{0,r}$ is finer than the partition σ_0 .

Then we give an upper bound for $L_{\sigma_{0,r}}(g^r) - L(g_{\mathcal{A}}^r)$. Let $\mathcal{A}(r) = (a_i^r)_{i=0}^{N_{\mathcal{A}}^r}$. For any $i \in [1, N_{\mathcal{A}}^r] \cap \mathbb{Z}$, let R_i^j , $0 \leq j \leq n_i$ be the points of $\mathcal{C}(g)$ with abscissas in $\sigma_{0,r} \cap [a_{i-1}^r, a_i^r]$ (thereby, the abscissa of R_i^0 is a_{i-1}^r and the abscissa of $R_i^{n_i}$ is a_i^r). Observe that $n_i \neq 1$ for at most $(n-1)$ values of i (there is at most $n-1$ points in σ_0 that are not in $\mathcal{A}(r)$).

$$\begin{aligned} L_{\sigma_{0,r}}(g^r) - L(g_{\mathcal{A}}^r) &= \sum_{i=1}^{N_{\mathcal{A}}^r} \left(\left(\sum_{j=1}^{n_i} d(R_i^{j-1}, R_i^j) \right) - d(R_i^0, R_i^{n_i}) \right) \\ &\leq \sum_{n_i \neq 1} (\varphi(k) - 1) h (a_i^r - a_{i-1}^r) \quad (\text{for } g \text{ is } k\text{-Lipschitz}) \\ &\leq (n-1)(\varphi(k) - 1) h M_{+\infty} \\ &\leq \frac{\varepsilon}{2}. \end{aligned}$$

Then,

$$\begin{aligned} L(g^r) - L(g_{\mathcal{A}}^r) &\leq (L(g^r) - L_{\sigma_{0,r}}(g^r)) + (L_{\sigma_{0,r}}(g^r) - L(g_{\mathcal{A}}^r)) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &\leq \varepsilon. \end{aligned}$$

We conclude the proof straightforwardly. \square

When the derivative of g is Lipschitz continuous, the next lemma gives us a bound on the difference between the length of the curve $\mathcal{C}(g^r)$ and the length of the polyline $\mathcal{C}(g_{\mathcal{A}}^r)$.

Proposition 3. *If g is k_1 -Lipschitz continuous and g' is k_2 -Lipschitz on each interval included in its domain, we have for any pattern function \mathcal{A} and any resolution $r > 0$*

$$L(g^r) - L(g_{\mathcal{A}}^r) \leq \frac{k_2(b-a)}{2} h M_1(\mathcal{A}(r)) \left(1 + c_v(\mathcal{A}(r))^2\right) + (\varphi(k_1) - 1) \mathcal{H}_\delta(g) \quad (7)$$

with $\delta = h M_{+\infty}(\mathcal{A}(r))$ and $\mathcal{H}_\delta(g) = \max \sum_i w_i$ where the maximum is over all minimal coverings⁶ of the set of points in $[a, b]$ at which g is not differentiable by closed subintervals of $[a, b]$ with diameters $w_i \leq \delta$ and pairwise disjoint interiors.

Proof. Let $\mathcal{A}(r) = (a_i^r)_{i=0}^{N_{\mathcal{A}}^r}$. Let $D \subset [1, N_{\mathcal{A}}^r] \cap \mathbb{Z}$ be the set of subscripts i such that g is differentiable on $(a_{i-1}^r h, a_i^r h)$. Note that the function φ is 1-Lipschitz and thereby $\varphi \circ g'$ is k_2 -Lipschitz.

From the mean value theorem, on each interval $(a_{i-1}^r h, a_i^r h)$, $i \in D$, there exists a real t_i^r such that

$$\frac{g(a_i^r h) - g(a_{i-1}^r h)}{(a_i^r - a_{i-1}^r)h} = g'(t_i^r).$$

Thus, the length of the restriction of $\mathcal{C}(g_{\mathcal{A}}^r)$ to the interval $[a_{i-1}^r h, a_i^r h]$ is

$$\begin{aligned} \sqrt{(g(a_i^r h) - g(a_{i-1}^r h))^2 + h^2(a_i^r - a_{i-1}^r)^2} &= h(a_i^r - a_{i-1}^r) \varphi(g'(t_i^r)) \text{ if } i \in D \text{ and} \\ \sqrt{(g(a_i^r h) - g(a_{i-1}^r h))^2 + h^2(a_i^r - a_{i-1}^r)^2} &\geq h(a_i^r - a_{i-1}^r) \text{ if } i \notin D. \end{aligned}$$

Moreover, for any $i \notin D$, since g is k_1 -Lipschitz, one has by Equation (6)

$$L(g|_{[a_{i-1}^r h, a_i^r h]}) \leq \varphi(k_1) h(a_i^r - a_{i-1}^r)$$

⁶A covering by a family E of sets is minimal if no subfamily of E is itself a covering.

Then,

$$\begin{aligned}
L(g^r) - L(g_{\mathcal{A}}^r) &\leq \sum_{i \in D} \int_{a_{i-1}^r h}^{a_i^r h} \varphi \circ g'(t) - \varphi \circ g'(t_i^r) dt + \\
&\quad (\varphi(k_1) - 1) \sum_{i \notin D} h(a_i^r - a_{i-1}^r) \\
&\leq \sum_{i \in D} \int_{a_{i-1}^r h}^{a_i^r h} k_2 |t - t_i^r| dt + (\varphi(k_1) - 1) \mathcal{H}_\delta(g) \\
&\leq \sum_{i=1}^{N_{\mathcal{A}}^r} k_2 h^2 \frac{(a_i^r - a_{i-1}^r)^2}{2} + (\varphi(k_1) - 1) \mathcal{H}_\delta(g) \\
&\leq \frac{1}{2} k_2 h^2 N_{\mathcal{A}}^r M_2^2 + (\varphi(k_1) - 1) \mathcal{H}_\delta(g) \\
&\leq \frac{1}{2} k_2 h (b-a) \frac{M_1^2 + V}{M_1} + (\varphi(k_1) - 1) \mathcal{H}_\delta(g) \\
&\quad \text{where } V \text{ is the variance of } \mathcal{A}(r) \\
&\leq \frac{1}{2} k_2 (b-a) h M_1 (1 + c_v^2) + (\varphi(k_1) - 1) \mathcal{H}_\delta(g)
\end{aligned}$$

where the penultimate inequality comes from the well known formula $V = M_2^2 - M_1^2$ and from the relation

$$M_1 = \frac{1}{N_{\mathcal{A}}^r} \sum_{i=1}^{N_{\mathcal{A}}^r} (a_i - a_{i-1}) = \frac{1}{N_{\mathcal{A}}^r} (a_{N_{\mathcal{A}}^r}^r - a_0^r) \leq \frac{1}{h N_{\mathcal{A}}^r} (b-a).$$

□

In the above proposition, obviously, the second term of the sum in the right hand side of the inequality vanishes if g is differentiable. In particular, if g is C^2 on I , then g and g' are Lipschitz continuous on I and $\mathcal{H}_\delta(g) = 0$. Then, the difference $L(g^r) - L(g_{\mathcal{A}}^r)$ is a $O(hM_1(1 + c_v^2))$.

When g is differentiable everywhere but on finitely many points, $\mathcal{H}_\delta(g)$ is equal to $n\delta$ (for small enough δ) where n is the number of points where g is not differentiable. Thus the difference $L(g^r) - L(g_{\mathcal{A}}^r)$ is a $O(hM_{+\infty})$. Indeed, for any sequence of positive numbers (x_i) ,

$$M_1(1 + c_v^2) = \frac{M_2^2}{M_1} = \frac{\sum x_i^2}{\sum x_i} \leq \frac{\sum x_i M_{+\infty}}{\sum x_i} \leq M_{+\infty}$$

In A, we detail the calculus of $\mathcal{H}_\delta(g)$ for a fractal function with infinitely many points where g is not differentiable (infinitely many isolated points and infinitely many limit points).

Given a Lipschitz function g on an interval I and \mathcal{A}, \mathcal{B} two pattern functions on I , we now look at the difference between $L(g_{\mathcal{A}}^r)$ and $L(\lfloor g_{\mathcal{B}}^r \rfloor)$ that is between the length of two piecewise affine functions.

Proposition 4. *Let f_1 and f_2 be two piecewise affine functions defined on $[c, d] \subset \mathbb{R}$ ($d > c$) with a common partition having p steps. Suppose that $e_1 \leq f_1 - f_2 \leq e_2$ for some reals e_1, e_2 . Then*

$$|L(f_1) - L(f_2)| \leq \varphi'(k) p (e_2 - e_1)$$

where k is the arithmetic mean of the set $\{\max(|s_{1,i}|, |s_{2,i}|)\}_{i=1}^p$, the reals $s_{1,i}, s_{2,i}$, $1 \leq i \leq p$, being the slopes of f_1 and f_2 on each subinterval of the common partition.

Proof. Let $\sigma = (x_i)_{i=0}^p$ be a common partition for f_1 and f_2 . We write m_i for $x_i - x_{i-1}$ and $s_{1,i}$, resp. $s_{2,i}$, for the slope of f_1 , resp. f_2 , on the interval $[x_{i-1}, x_i]$.

Then

$$L(f_1) - L(f_2) = \sum_{i=1}^p (\varphi(s_{1,i}) - \varphi(s_{2,i})) m_i.$$

From the mean value theorem, we derive that

$$L(f_1) - L(f_2) = \sum_{i=1}^p \varphi'(s_{0,i}) m_i (s_{1,i} - s_{2,i}).$$

where, for any i , $s_{0,i}$ lies between $s_{1,i}$ and $s_{2,i}$.

Note that, for any $i \leq p$,

$$m_i (s_{1,i} - s_{2,i}) = f_1(x_i) - f_2(x_i) - (f_1(x_{i-1}) - f_2(x_{i-1})).$$

Thus, as, by hypothesis, $e_1 \leq f_1 - f_2 \leq e_2$, we get

$$-(e_2 - e_1) \leq m_i (s_{1,i} - s_{2,i}) \leq e_2 - e_1.$$

Therefore,

$$\begin{aligned} |L(f_1) - L(f_2)| &\leq \sum_{i=1}^p |\varphi'(s_{0,i})| |m_i (s_{1,i} - s_{2,i})| \\ &\leq p(e_2 - e_1) M_1((\varphi'(|s_{0,i}|))_{i=1}^p) \\ &\leq p(e_2 - e_1) \varphi'(M_1(|s_{0,i}|)_{i=1}^p) \\ &\leq p(e_2 - e_1) \varphi'(k) \end{aligned}$$

where the penultimate inequality is derived from the concavity of the function φ' on $[0, +\infty)$. \square

Thanks to the four previous propositions, we can state our theorem on the convergence of non-local length estimators. Figure 3 illustrates the theorem hypotheses.

Theorem 3. *Let $g: [a, b] \rightarrow \mathbb{R}$ be a k_1 -Lipschitz function and \mathcal{A} be a 1-pattern function.*

If there exist a $(1, \beta)$ -pattern function \mathcal{B} , $\beta \in [1, +\infty]$, and a real ω such that for any resolution r ,

$$r \|[g_{\mathcal{A}}^r] - [g_{\mathcal{B}}^r]\|_{+\infty} \leq \omega \quad (8)$$

then

- if $\beta = +\infty$, the non-local estimation $L^{\text{NL}}(g, \mathcal{A}, r)$ converges toward the length of the curve $\mathcal{C}(g)$ as r tends to $+\infty$;
- if g' is k_2 -Lipschitz on each interval included in its domain, we have

$$\begin{aligned} L(g) - L^{\text{NL}}(g, \mathcal{A}, r) &\leq \\ &Sh \quad (\text{side error}) \\ &+ ThM_1(1 + c_v^2) + U\mathcal{H}_\delta(g) \quad (\text{discretization error}) \\ &+ V \left(\frac{1}{M_1} + \frac{1}{M'_1} \right) \quad (\text{quantification error}) \end{aligned} \quad (9)$$

where

$$S = 2\varphi(k_1),$$

$$T = \frac{k_2(b-a)}{2}, \quad M_1 = M_1(\mathcal{B}(r)), \quad c_v = c_v(\mathcal{B}(r)),$$

$$U = \varphi(k_1) - 1, \quad \delta = hM_{+\infty}(\mathcal{B}(r)) \quad \text{and} \quad \mathcal{H}_\delta(g) = \max_i \sum_i w_i \quad \text{where}$$

the maximum is over all minimal coverings of the set of points in $[a, b]$ at which g is not differentiable, by closed intervals with diameters $w_i \leq \delta$ and pairwise disjoint interiors,

$$V = (1 + 2\omega)\varphi' \left(k_1 + \frac{1}{M_{-1}(\mathcal{A}(r))} \right) (b-a) \quad \text{and} \quad M'_1 = M_1(\mathcal{A}(r)).$$

Furthermore, if $\mathcal{B}(r) \subseteq \mathcal{A}(r)$, the term $\frac{1}{M_1} + \frac{1}{M'_1}$ in the right hand side of Equation 9 can be replaced by $\frac{1}{M_1}$.

Proof. We write the difference between $L(g)$ and $L([g_{\mathcal{A}}^r])$ as the sum of three terms.

$$L(g) - L([g_{\mathcal{A}}^r]) = (L(g^a) + L(g^b)) + (L(g^r) - L(g_{\mathcal{B}}^r)) + (L(g_{\mathcal{B}}^r) - L([g_{\mathcal{A}}^r])) \quad (10)$$

Each term of the right hand side of (10) is majorized as follows.

- From Proposition 1, we have

$$L(g^a) + L(g^b) \leq 2\varphi(k_1)h \quad (11)$$

Thus, $L(g^a) + L(g^b)$ converges to 0 as $r \rightarrow +\infty$.

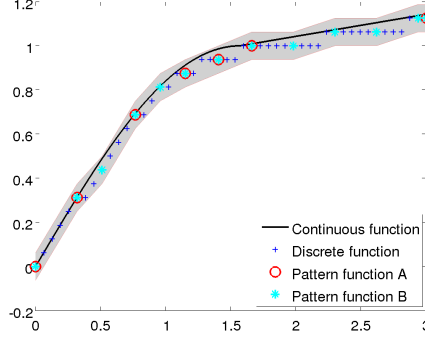


Figure 3: Hypotheses of Theorem 3. The partition related to $\mathcal{A}(r)$ split the discrete curve in patterns whose average size in pixels tends toward infinity as the resolution tends itself toward infinity. Another partition of the curve, related to $\mathcal{B}(r)$, has patterns whose average absolute size tends toward 0 as the resolution tends toward infinity. A tube (in gray) with constant height contains the discrete points related to the partitions $\mathcal{A}(r)$ and $\mathcal{B}(r)$.

- From Proposition 2, we know that $L(g_{\mathcal{B}}^r)$ converges towards $L(g^r)$ as $r \rightarrow +\infty$ provided $\beta = +\infty$.
When g' is k_2 -Lipschitz continuous on the intervals included in its domain, from Proposition 3 we have

$$L(g^r) - L(g_{\mathcal{B}}^r) \leq \frac{k_2(b-a)}{2} h M_1 (1 + c_v^2) + (\varphi(k_1) - 1) \mathcal{H}_\delta(g). \quad (12)$$

- From Hypothesis (8),

$$\| \lfloor g_{\mathcal{A}}^r \rfloor - \lfloor g_{\mathcal{B}}^r \rfloor \|_{+\infty} \leq \omega h.$$

Thus, we have

$$-\omega h \leq g_{\mathcal{B}}^r - \lfloor g_{\mathcal{A}}^r \rfloor \leq (\omega + 1)h.$$

Since g is k_1 -Lipschitz, the piecewise affine function $g_{\mathcal{B}}^r$ is clearly k_1 -Lipschitz. Let $\mathcal{A}(r) = (a_i)_{i=0}^{N_{\mathcal{A}}^r}$. For any $i \in [1, N_{\mathcal{A}}^r]$, the absolute slope of $\lfloor g_{\mathcal{A}}^r \rfloor$ on $[a_{i-1}h, a_ih]$ is bounded by $k_1 + 1/(a_i - a_{i-1})$. Hence, the arithmetic mean of the absolute slopes of $\lfloor g_{\mathcal{A}}^r \rfloor$ is bounded by $k = k_1 + 1/M_{-1}(\mathcal{A}(r))$ (thus $k \leq k_1 + 1$). Then, from Proposition 4, we derive that

$$|L(g_{\mathcal{B}}^r) - L(\lfloor g_{\mathcal{A}}^r \rfloor)| \leq \varphi'(k) N (1 + 2\omega) h$$

where N is the size of the partition $\mathcal{A}(r) \cup \mathcal{B}(r)$. So, we have $N \leq N_{\mathcal{A}}^r + N_{\mathcal{B}}^r$ and we observe that, for any $\mathcal{I} \in \{\mathcal{A}, \mathcal{B}\}$,

$$M_1(\mathcal{I}(r)) = \frac{1}{N_{\mathcal{I}}^r} (B^r - A^r) \leq \frac{1}{h N_{\mathcal{I}}^r} (b - a).$$

Thus, we get

$$|L(g_{\mathcal{B}}^r) - L(\lfloor g_{\mathcal{A}}^r \rfloor)| \leq (1 + 2\omega)\varphi'(k)(b - a) \left(\frac{1}{M_1} + \frac{1}{M_1'} \right). \quad (13)$$

It follows from Equation (13) and the hypotheses that $L(g_{\mathcal{B}}^r) - L(\lfloor g_{\mathcal{A}}^r \rfloor)$ converges to 0 as $r \rightarrow +\infty$.

Thereafter, the convergence is established once $\beta = +\infty$. Moreover, Equation (9) derives obviously from eqs. (11) to (13). \square

Basically, Theorem 3 states that a non-local estimator is convergent for Lipschitz functions provided it is not too far from a M-sparse estimator. We can also infer from this theorem that a non-local length estimation $L^{\text{NL}}(g, \mathcal{A}, r)$ by a M-sparse estimator converges toward $L(g)$ when \mathcal{A} is a $(1, +\infty)$ -pattern function or when, for instance, \mathcal{A} is a $(1, 1)$ -pattern function, g is differentiable, its derivative is Lipschitz and the coefficient of variation of $\mathcal{A}(r)$ is bounded as $r \rightarrow \infty$. Let us now look at the possible options for the pattern function \mathcal{B} .

Choice of the pattern function \mathcal{B} As mentioned at the beginning of Section 3.2 the second pattern function \mathcal{B} is introduced so as to include MDSS estimators for which pattern absolute size needs not tend to zero. For example, it is plain that for an affine function g defined on $[a, b]$, an MDSS segmentation will provide a single pattern of size $B^r - A^r \approx (b - a)r$. Nevertheless, after an MDSS segmentation, it is always possible to artificially subdivide the obtained segments in order to define the pattern function \mathcal{B} . Then, the definition of a straight line segment ensures that Hypothesis (8) of Theorem 3 is satisfied (and the constant ω can be set to 1 in our discretization model). Thereby, together with the constant ω , the piecewise affine function $g_{\mathcal{B}}^r$ provides a tube around the data points attached to the partition $\mathcal{B}(r)$ in which must stay the function $\lfloor g_{\mathcal{A}}^r \rfloor$ to reliably estimate $L(g)$. Thus this tube can be seen as a data fitting. Figure 3 gives an illustration of such a situation. Also, it should be noticed that if we assume the same hypotheses as in [15], C^3 functions with strictly positive curvature, we can take $\mathcal{B} = \mathcal{A}$ with the MDSS segmentation since the existence of a strictly positive minimal curvature on $[a, b]$ ensures that $M_{+\infty}(\mathcal{A}(r)) = o(r)$. Nevertheless, even with finitely many inflexion points, we could not directly applied Theorem 3 with $\mathcal{B} = \mathcal{A}$ to C^3 curves. Indeed, when a MDSS is partitioned in subsegments, the length that comes with the MDSS (its diameter) is not the sum of the diameters of the subsegments and it is the role of (8) in the theorem hypotheses to bound the difference (via Proposition 4). Eventually, since there are infinitely many valid choices of \mathcal{B} for the MDSS estimator, Equation 9 leads to the following bound on the estimation error (for the MDSS estimator):

$$L(g) - L^{\text{NL}}(g, r) \leq \inf_i \left(Sh + ThM_{1,i} (1 + c_{v,i}^2) + U\mathcal{H}_{\delta_i}(g) + V \left(\frac{1}{M_{1,i}} + \frac{1}{M_1'} \right) \right)$$

where the infimum is over all 1-pattern functions \mathcal{B}_i such that $M_{1,i} = M_1(\mathcal{B}_i(r)) = o(r)$ as $r \rightarrow +\infty$ and where $c_{v,i} = c_v(\mathcal{B}_i(r))$ and $\delta_i = hM_{+\infty}(\mathcal{B}_i(r))$.

With sparse estimators, the partition subinterval length sequence provided by the pattern function is $(H(h), \dots, H(h), rem)$ where $\lim_{h \rightarrow 0} H(h) = +\infty$, $\lim_{h \rightarrow 0} hH(h) = 0$ and $rem \leq H(h)$. Thus, taking $\mathcal{B} = \mathcal{A}$, $\omega = 0$ and $\beta = 1$ all the hypotheses of Theorem 3 are satisfied.

Theorem 3 is tested on real functions in the next section.

4 Tests

The tests have been made using the MPFR and GMP libraries through the SAGE [21] software. The results are presented in Table 1 and Table 3. Hereafter, we present the tested non local length estimators (NLE) and the tested functions.

4.1 Protocol

The tested NLE are:

- The sparse estimators $\mathcal{E}^{\text{Sp}2}$ and $\mathcal{E}^{\text{Sp}3}$ with
 1. $\mathcal{E}^{\text{Sp}2}$: the discrete pattern size of $r^{\frac{1}{2}}$ which is generally optimal for such an estimator (see [4]),
 2. $\mathcal{E}^{\text{Sp}3}$: the discrete pattern size $r^{\frac{1}{3}}$ that approximatively corresponds to the asymptotic average discrete size of the MDSS on a smooth curve with positive minimal curvature [15].
- The M-sparse estimator $\mathcal{E}^{\text{rand}}$ with random pattern sizes equidistributed between 1 and $2r^{\frac{1}{2}}$. Hence, for large enough resolutions, the average pattern size should be the same as $\mathcal{E}^{\text{Sp}2}$. Note that for the three previous NLE, we can apply Theorem 3 taking $\mathcal{A} = \mathcal{B}$.
- The MDSS based estimator \mathcal{E}^{DSS} (taking the diameter as the pattern length – see the comment at the end of Section 3.1).

The computed errors are the discretization and the quantization errors. The bounds and the grid spacing h are chosen such that the error on the bounds is null. Then the total error comes down to the sum of both discretization and quantization errors.

There are five test functions that we describe now. The first three satisfy the stronger assumptions on the function g in Theorem 3 (the function and its derivative are Lipschitz continuous), the two last are Lipschitz but their derivatives are not.

- The natural logarithm on the interval $[1, 2]$ where it is 1-Lipschitz, concave and of class C^∞ .

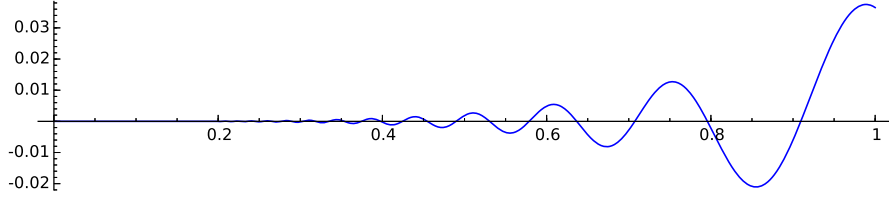


Figure 4: Graph of the function s_1 .

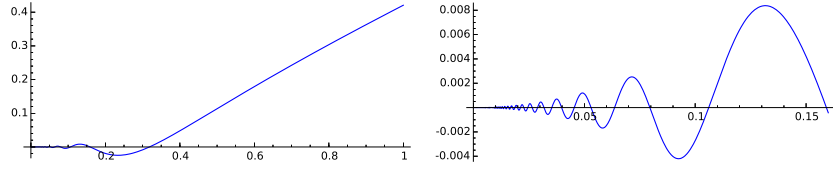


Figure 5: Graph of the function s_2 .

- The function

$$s_1 : \begin{cases} x \in (0, 1] \mapsto \frac{x^4}{25} \sin \frac{20}{x}, \\ 0 \mapsto 0 \end{cases}$$

which has infinitely many inflexion points (but only finitely many inflexion points on any interval $[a, 1]$, $a > 0$), is 1-Lipschitz of class C^1 (C^∞ on $(0, 1]$) with a Lipschitz continuous derivative.

The graph of s_1 is shown Figure 4.

- The function

$$s_2 : \begin{cases} x \in (0, 1] \mapsto \frac{x^2}{2} \sin \frac{1}{x}, \\ 0 \mapsto 0 \end{cases}$$

which has infinitely many inflexion points, is 1-Lipschitz of class C^1 (C^∞ on $(0, 1]$). The convergence is given by Theorem 3 but, unlike to the function s_1 , the derivative of s_2 is not Lipschitz on $(0, 1]$. Then the bound obtained in Proposition 3 is not valid for this function. Nevertheless, only one computation of pattern length (the first one) is concerned by this restriction. The graph of s_2 is shown Figure 5.

- The fractal functions f_1 and f_2 defined on $[0, 1]$ as follows.

We denote by $\mathbf{1}_J$ the characteristic function of the interval $J = [\frac{1}{3}, \frac{2}{3}]$ and by $\{\cdot\}$ the fractional part.

Then, for any $i \in \{1, 2\}$, $f_i = \lim_{n \rightarrow \infty} f_{i,n}$ where

$$f_{1,0} : x \mapsto \frac{1}{2} - \left| x - \frac{1}{2} \right|, \quad f_{2,0} : x \mapsto \frac{1}{2\pi} (1 - \cos(2\pi x)),$$

$$f_{i,1} : x \mapsto \frac{1}{3} f_{i,0}(\{3x\}) \mathbf{1}_J(x)$$

and for $n > 1$,

$$f_{i,n} : \begin{cases} x \notin J \mapsto \frac{1}{3} f_{i,n-1}(\{3x\}), \\ x \in J \mapsto f_{i,n-1}(x). \end{cases}$$

An illustration of the graphs $\mathcal{C}(f_1)$ and $\mathcal{C}(f_2)$ is given Figures 6 and 7. The length of $\mathcal{C}(f_i)$, $i \in \{1, 2\}$ is the length of $\mathcal{C}(f_{i,0})$.

The functions f_1 and f_2 are 1-Lipschitz. The function f_1 is not differentiable on $[0, 1]$ but it is plain that the derivative of f_1 is constant on any interval included in its domain. On the contrary, f_2 is of class C^1 on $[0, 1]$, C^∞ almost everywhere, but its second derivative is not bounded. Thus Equation (9) of Theorem 3 does not apply to f_2 . Furthermore, unlike s_2 , the number of patterns where the second derivative is unbounded tends toward $+\infty$ as the resolution grows.

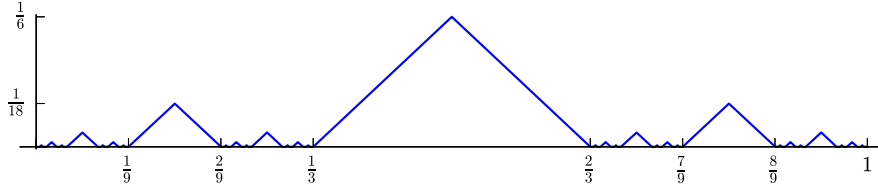


Figure 6: Graph of the fractal function f_1 .

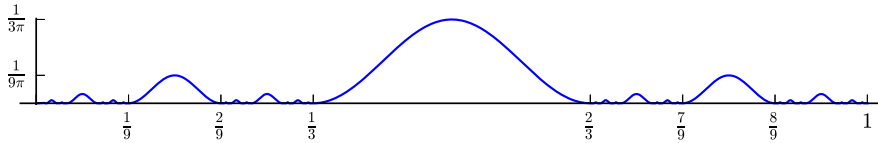


Figure 7: Graph of the fractal function f_2 .

Remark 1. *The pattern functions of \mathcal{E}^{DSS} for the tested functions on the tested resolutions have not contradicted the assumption that MDSS are non-local estimators. Furthermore, the bounds $r^{1/3}$ and $r^{1/3} \log(r)$ for the asymptotic average discrete size of the MDSS on a smooth solid contour with positive minimal curvature fit our experiment (see Figure 8).*

4.2 On the theoretical bounds

The tests gathered in Table 1 aim at comparing the theoretical bounds obtained in Theorem 3 with the experimental values of the discretization error and the

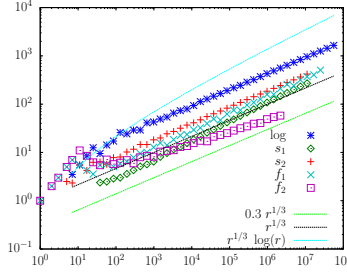


Figure 8: Average discrete size of the MDSS produced by \mathcal{E}^{DSS} on the tested functions according to the resolution.

quantization error. In the table, the blue '+' represents the theoretical upper bound and the green '*' represents the observed error for the resolutions $[1.5^n]$, $n \in [1, 50] \cap \mathbb{N}$.

Discretization error The discretization error has two parts. The second one only concerns the function f_1 for it is the only one that is not differentiable.

The first part of the bound is

$$\frac{k_2(b-a)}{2} h M_1 (1 + c_v^2)$$

where k_2 is a Lipschitz constant for the derivative of the tested function. We took:

	log	s_1	s_2	f_1	f_2
k_2	1	17	∞	0	∞

Since $k_2 = 0$ for f_1 , this function is not concerned by the first part of the bound. When the derivative is not Lipschitz continuous ($k_2 = \infty$), the blue '+' on the plot represent the values of $h M_1 (1 + c_v^2)$.

Since $\mathcal{E}^{\text{rand}}$ has pattern sizes equidistributed in the interval $[1, 2\sqrt{r}]$, for large enough resolutions we have

$$h M_1 (1 + c_v^2) \approx \frac{1}{r} \times \sqrt{r} \times \left(1 + \frac{1}{3}\right) = \frac{4}{3} r^{-\frac{1}{2}}$$

and thereby

$$\frac{k_2(b-a)}{2} h M_1 (1 + c_v^2) \approx \frac{2}{3} k_2 r^{-\frac{1}{2}}.$$

The cyan line on the plots of the discretization error of the functions \log , s_1 , s_2 and f_2 is the graph of $r \mapsto \Gamma r^{-\frac{1}{2}}$ where $\Gamma = \frac{2}{3} k_2$ ($\Gamma = \frac{4}{3}$ for s_2 and f_2).

It can be seen in Table 1 that the observed errors for the functions \log and s_1 are less than the upper bounds given by Theorem 3. Specifically, the power in the measured errors are almost twice the bounds, as far as the resolution is big enough. We know from our previous work [4], that convexity allow to double the

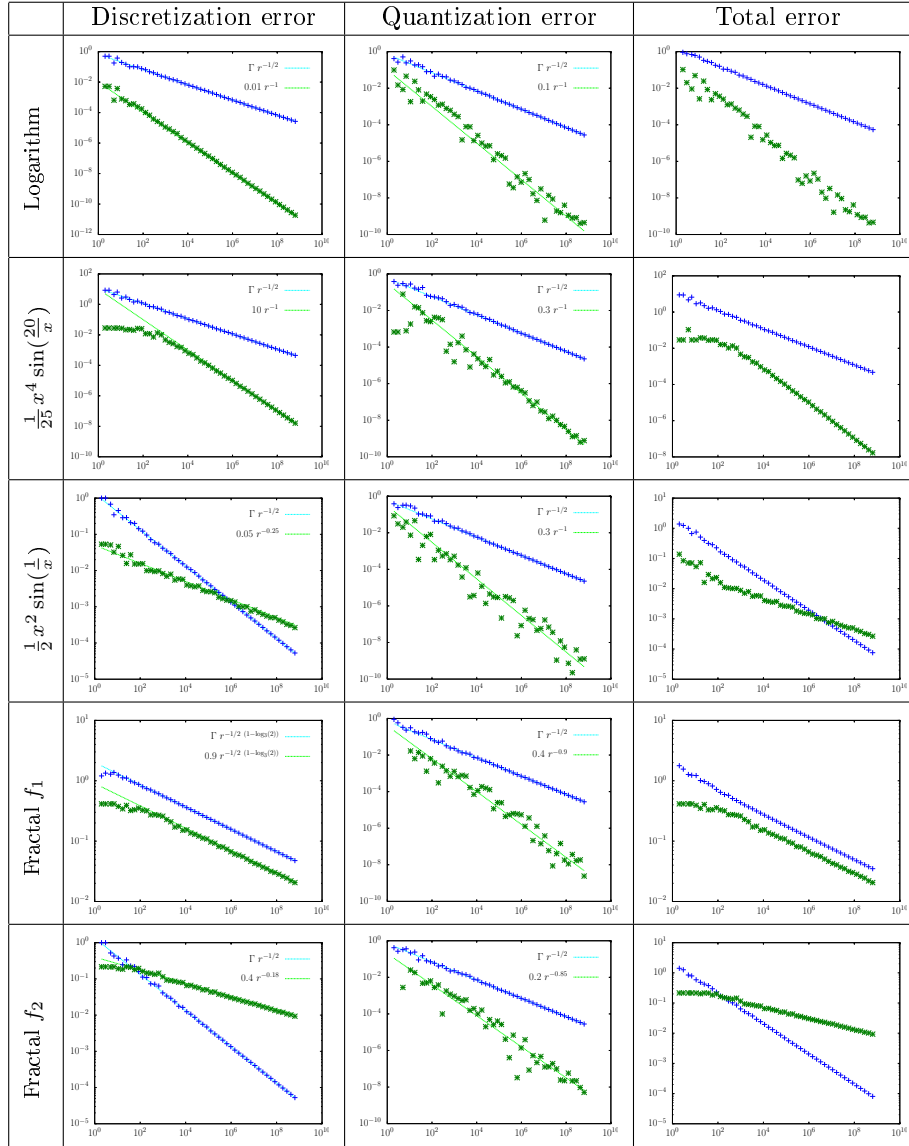


Table 1: Observed length estimation errors and theoretical upper bounds in function of the resolution r for the NLE $\mathcal{E}^{\text{rand}}$. Blue crosses: the theoretical bounds given by Theorem 3. Green stars: the observed errors. Γ : a constant that is computed from the bounds in Theorem 3 and from the expected generalized means of the random pattern sizes (see text).

convergence rate of the sparse estimators. This explains the observed difference for the logarithm (provided we can extend our previous result to the M-sparse estimators). On the contrary, the function s_1 has infinitely many inflexion points. Nevertheless, once the first grid step is passed, it remains only finitely many of them. This could explain the good convergence rate. In a work in preparation, we study this issue by bringing a measure of the set of inflexion points.

On the plots related to the functions s_2 and f_2 , we can see that having an unbounded second derivative drastically decreases the convergence rate, even for s_2 which has a bounded second derivative on any closed subinterval of its domain that does not contain 0.

Regarding the function f_1 , which is 1-Lipschitz, the bound for the discretization error is

$$(\varphi(1) - 1)\mathcal{H}_\delta(g)$$

where $\delta = hM_{+\infty}$. An upper bound for $\mathcal{H}_\delta(g)$ is calculated in A:

$$\mathcal{H}_\delta(g) \leq \frac{15}{4} \delta^{1-\log_3(2)}$$

With $\mathcal{E}^{\text{rand}}$, $\delta = hM_{+\infty} \approx \frac{1}{r} \times 2\sqrt{r} = 2r^{-\frac{1}{2}}$ (for enough large resolutions). Thus, the bound for the function f_1 is approximatively

$$\frac{15(\sqrt{2}-1)}{2^{1+\log_3(2)}} r^{-\frac{1}{2}(1-\log_3(2))}.$$

The cyan line on the plot of the error for the function f_1 is the graph of $r \mapsto \Gamma r^{-\frac{1}{2}(1-\log_3(2))}$ where $\Gamma = 15(\sqrt{2}-1)/2^{1+\log_3(2)}$.

The plot relative to f_1 shows the good accuracy of the upper bound for this non differentiable function. Moreover, if instead of taking $\delta = hM_{+\infty}$ in the bound, which corresponds to the worst case, we take $\delta = hM_1$, that is we cover the points of non differentiability by patterns of average size, both point clouds overlap almost.

Quantization error The upper bound for the quantization error is derived from Proposition 4 and the function just needs be Lipschitz continuous. Thus, this bound is valid for our five test functions.

With the estimator $\mathcal{E}^{\text{rand}}$ we can take just one pattern function ($\mathcal{A} = \mathcal{B}$ and $\omega = 0$). So, the quantization error is bounded by

$$\varphi' \left(k_1 + \frac{1}{M-1} \right) (b-a) \frac{1}{M_1}$$

where k_1 is a Lipschitz constant for the tested function. Thus, for enough large resolutions, the bound on the quantization error for the estimator $\mathcal{E}^{\text{rand}}$ is approximatively

$$\varphi'(k_1) r^{-\frac{1}{2}}.$$

For k_1 , we took

	log	s_1	s_2	f_1	f_2
k_1	1	0.7	0.7	1	1

The cyan line on the plots of the quantization errors (for the five functions) is the graph of $r \mapsto \Gamma r^{-\frac{1}{2}}$ where $\Gamma = \varphi'(k_1)$.

With the quantization error, the results for all tested functions are less than the upper bounds given by Theorem 3. Again the powers in the measured errors are almost twice the bounds, as far as the resolution is big enough. As with the discretization error, this corresponds to a property of concave functions shown in [4] for the sparse estimators and this property could also be valid for the NLEs and for a larger class of functions than the concave ones. Nevertheless this is not true in the worst case. Hereafter, we show a very simple example that proves that the bound cannot be improved for any resolution.

Let us choose a resolution $r = 4k^2$, $k \in \mathbb{N}$ and we assume that, at this resolution, the pattern size is constant, equals to $\sqrt{r} = 2k$. Now, we define the periodic piecewise affine function

$$v_{k,a} = \left(1 + \frac{a}{k}\right) \frac{1}{k} v_0(kx)$$

where $0 < a \ll 1$ and

$$v_0(x) = -\frac{1}{2\pi} \arccos(\cos(2\pi x)).$$

Figure 9 shows the plot of the function $v_{4,0.1}$ on $[0, 1]$. Each pattern is a digital straight segment and the quantization error is maximal on each segment for the difference between the ordinates of the discrete function and the continuous function at the lower end of the segments is $(1 - 2a)h \approx h$ (where h is the grid step). On Table 2, we show the relative difference between the quantization error (QE) and the upper bound (UB) in Theorem 3 for some functions $v_{k,a}$ where $a = 10^{-3}$.

k	1	2^2	2^4	2^6	2^8	2^{10}	2^{12}
r	4	2^6	2^{10}	2^{14}	2^{18}	2^{22}	2^{24}
QE	3e-1	6e-2	1e-2	3e-3	9e-4	2e-4	5e-5
$\frac{\text{QE}-\text{UB}}{\text{QE}}$	2e-1	8e-2	2e-2	6e-3	1e-3	4e-4	1e-4

Table 2: Relative difference between the quantization error for the function $v_{k,a}$ with $a = 10^{-3}$ and the upper bound provided by Theorem 3

4.3 Comparison between the tested estimators

A comparison of the different estimators on the five test function graphs is shown Table 3. The MDSS, Sparse and M-sparse estimators are tested on

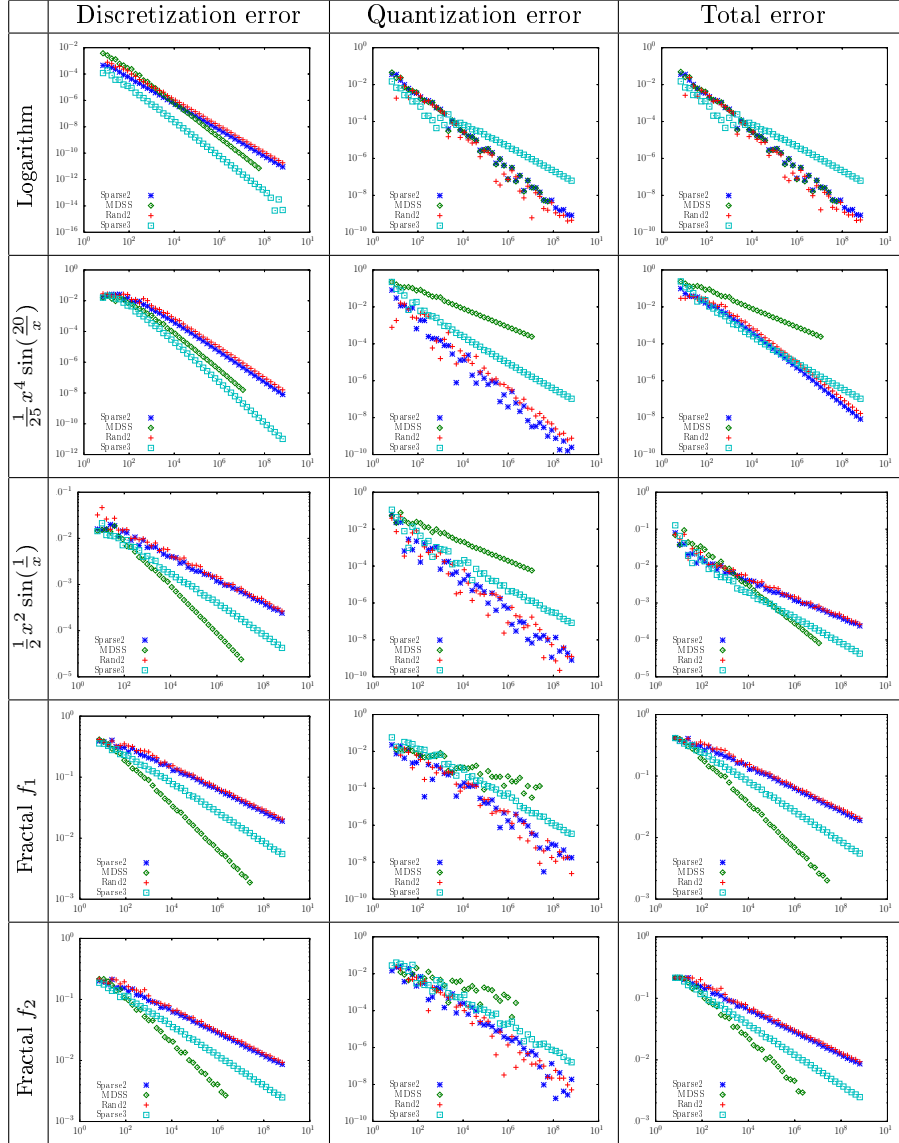


Table 3: Estimation errors in function of the resolution r for the estimators $\mathcal{E}^{\text{Sp}2}$, \mathcal{E}^{DSS} , $\mathcal{E}^{\text{rand}}$ and $\mathcal{E}^{\text{Sp}3}$. The tested functions are \log , s_1 , s_2 , f_1 and f_2 (see text).

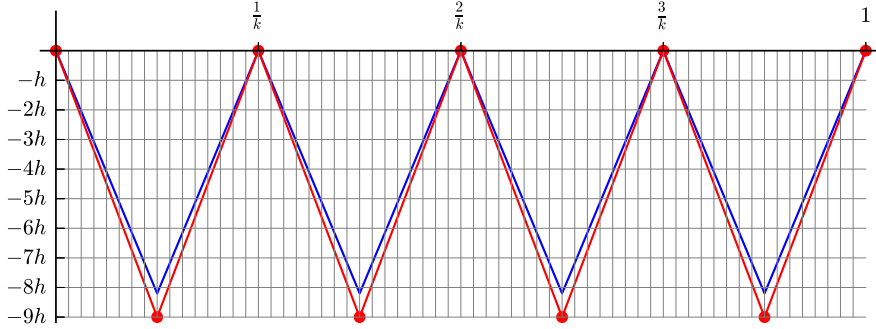


Figure 9: Blue (upper) line: the graph of the function $v_{4,0.1}$. Red points: MDSS end points of the $v_{4,0.1}$ OBQ discretization. Red line: polyline whose length is the non local length estimation of $L(g)$.

the Logarithm, the damped sinusoids s_1 , s_2 and the fractal functions f_1 , f_2 . Three errors are shown for the test: the discretization error (left column), the quantization error (middle column) and the total error (right column).

Discretization error The Discretization Error (DE) is due to the approximation of the curve by chords whose width is given by the pattern function. Thus, generally, the shorter the width, the smaller the DE. This is indeed what we observe in Table 3 where $\mathcal{E}^{\text{MDSS}}$ and \mathcal{E}^{Sp3} , which have a pattern size average of about $r^{\frac{1}{3}}$ while \mathcal{E}^{Sp2} and $\mathcal{E}^{\text{rand}}$ have a pattern size average of $r^{\frac{1}{2}}$. The good performance of $\mathcal{E}^{\text{MDSS}}$ on the f_1 graph which is composed of straight line segments is natural. Contrariwise, as MDSS are adaptive, we could expect that $\mathcal{E}^{\text{MDSS}}$ outperforms \mathcal{E}^{Sp3} on s_1 , s_2 and f_2 . Surprisingly this is only the case on the graph of f_2 .

Quantization error The Quantization Error (QE) comes from the vertical alignment of the chords on the grid. So, generally, a pattern function that produce a small number of patterns yield to a small QE on the contrary to the DE. This is what we observe on the experiment. In particular, the price of the adaptive nature of $\mathcal{E}^{\text{MDSS}}$ is a relatively larger QE.

Total error The influence of the QE and the DE on the total error depends on the studied graph. On the one hand, the order of magnitude of the QE is pretty much the same for the five functions and each NLE (with an exception for $\mathcal{E}^{\text{MDSS}}$ on the logarithm). On the other hand, the order of magnitude of the DE increases significantly from the smooth concave function \log to the fractal functions for the four tested NLE. Surprisingly, the adaptive estimator has relatively bad performances on the two damped sinusoids while it has the best results on the fractal f_2 which is built from a sine. We note also that the sparse estimator

$\mathcal{E}^{\text{Sp}^2}$ and its randomized equivalent, the M-sparse estimator $\mathcal{E}^{\text{rand}}$ have almost the same results for the five functions. Eventually, no estimator gives the best results over all the graph.

5 Conclusion

In this article, we have introduced a new class of length estimators, the non-local estimators (NLE). The NLE gather all the multigrid estimators whose discrete graph length estimation is made with a polyline whose vertices belong to the discrete graph under the assumption that the average of the line segment discrete sizes tends toward infinity as the resolution grows to infinity. The NLE class should encompass the MDSS based length estimators. Nevertheless we still need to prove it formally. We have also defined a subclass of the NLE, the M-sparse estimators, for which the average of the line segment absolute sizes tends toward 0 as the resolution tends toward infinity. We proved that any NLE has the multigrid convergence property for the Lipschitz functions as soon as at any resolution its polyline is close to the one of some M-sparse polyline. The bound on the convergence rate is the same as the one of semi-local and sparse estimators in the general case, under weaker assumptions on the function. The convergence rate can be improved when the function is concave or convex as it has been shown for sparse estimators in [4]. This point will be developed in a work in preparation [22]. We have also to study how the material presented in this article behaves with Jordan curves obtained as boundaries of solid objects through various discretization schemes (and more generally, with any graph of a one-to-one vector-valued function).

A Calculus of $\mathcal{H}_\delta(f)$

We consider the fractal function f defined in Section 4.1. Let E be the set of points in $[0, 1]$ at which g is not differentiable. E contains isolated points whose triadic development is composed of finitely many digits 0 or 2 followed by an infinite sequence of digits 1 and E contains limit points whose triadic development is finite, made of digits 0 or 2 but the last which is any digit.

We set $F(\delta) = \mathcal{H}_\delta(f)$. Recall that

$$F(\delta) = \max \sum_i w_i$$

where the maximum is over the set \mathcal{E}_δ of all minimal coverings of the set E by closed subintervals of $[0, 1]$ with diameters $w_i \leq \delta$ and pairwise disjoint interiors.

For $\delta \geq \frac{1}{9}$, the set $\{I_i\}_{i=1}^9$ where $I_i = [\frac{i-1}{9}, \frac{i}{9}]$ is a minimal covering of E with intervals whose interiors are pairwise disjoint.

Hence, for any $\delta \geq \frac{1}{9}$, we have

$$F(\delta) = 1. \tag{14}$$

When $\delta < \frac{1}{9}$, any minimal covering in \mathcal{E}_δ can be split in three parts: a minimal covering of $E \cap [0, x_1]$ where $x_1 \in [\frac{1}{3}, \frac{4}{9})$, an interval that contains the point $\frac{1}{2}$ and a minimal covering of $E \cap [x_2, 1]$ where $x_2 \in (\frac{5}{9}, \frac{2}{3}]$.

From the recursive definition of f , one can see that

$$\mathcal{H}_\delta(f|_{[0, \frac{1}{3}]}) = \mathcal{H}_\delta(f|_{[\frac{2}{3}, 1]}) = \frac{1}{3}\mathcal{H}_{3\delta}(f)$$

where $f|_J$ denotes the restriction of the function f to an interval J .

Thus, for any $\delta < \frac{1}{9}$,

$$F(\delta) = 2 \times (\frac{1}{3}(F(3\delta) + \delta) + \delta = \frac{2}{3}F(3\delta) + 3\delta \quad (15)$$

By solving the recurrence relation (15) with the initial equality (14), we get

$$\begin{aligned} F(x) &= \left(\frac{2}{3}\right)^n + 3\delta \sum_{i=0}^{n-1} 2^i \quad \text{with } \frac{1}{9} \leq 3^n \delta < \frac{1}{3} \\ &= \left(\frac{2}{3}\right)^n + 3\delta(2^n - 1) \quad \text{with } n = -\lfloor \log_3(\delta) \rfloor + 2 \end{aligned}$$

Therefore,

$$\begin{aligned} F(\delta) &\leq \frac{9}{4} \left(\frac{2}{3}\right)^{-\log_3(\delta)} + 3\delta \left(\frac{1}{2} 2^{-\log_3(\delta)}\right) \\ &\leq \frac{15}{4} \delta^{1-\log_3(2)}. \end{aligned}$$

Note that $\log_3(2)$ is the Hausdorff dimension of the set E .

References

- [1] S. R. Kulkarni, S. K. Mitter, T. J. Richardson, J. N. Tsitsiklis, Local versus nonlocal computation of length of digitized curves, *IEEE Trans. Pattern Anal. Mach. Intell.* 16 (7) (1994) 711–718.
- [2] D. Coeurjolly, R. Klette, A comparative evaluation of length estimators of digital curves, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (2) (2004) 252–257.
- [3] A. Daurat, M. Tajine, M. Zouaoui, Les estimateurs semi-locaux de périmètre, Tech. rep., LSIIT CNRS, UMR 7005, Université de Strasbourg, <http://hal.inria.fr/hal-00576881> (2011).
- [4] L. Mazo, E. Baudrier, About multigrid convergence of some length estimators, in: E. B. et al. (Ed.), *DGCI*, Vol. 8668 of LNCS, Springer, 2014, pp. 214–225.
- [5] R. Klette, A. Rosenfeld, *Digital Geometry*, Morgan Kaufman, 2004.

- [6] M. Tajine, A. Daurat, On local definitions of length of digital curves, in: I. Nyström, G. S. di Baja, S. Svensson (Eds.), DGCI, Vol. 2886 of LNCS, Springer, 2003, pp. 114–123.
- [7] A. Daurat, M. Tajine, M. Zouaoui, Patterns in discretized parabolas and length estimation, in: S. Brlek, C. Reutenauer, X. Provençal (Eds.), DGCI, Vol. 5810 of LNCS, Springer, 2009, pp. 373–384.
- [8] M. Tajine, A. Daurat, Patterns for multigrid equidistributed functions: Application to general parabolas and length estimation, *Theoretical Computer Science* 412 (36) (2011) 4824 – 4840.
- [9] V. Kovalevsky, New definition and fast recognition of digital straight segments and arcs, in: Proc. 10th Int. Conf. on Pattern Recognition, Vol. ii, 1990, pp. 31–34.
- [10] I. Debled-Renesson, J.-P. Reveillès, A linear algorithm for segmentation of digital curves, *Int J. Pattern Recognition and Artificial Intelligence* 09 (04) (1995) 635–662.
- [11] L. Dorst, A. W. Smeulders, Discrete straight line segments: parameters, primitives and properties, *Vision geometry: proceedings of an AMS special session held October 20-21, 1989* 119 (1991) 45–62.
- [12] R. Klette, , J. Žunić, Multigrid convergence of calculated features in image analysis, *J. Mathematical Imaging and Vision* 13 (3) (2000) 173–191.
- [13] A. Rosenfeld, Digital straight line segments, *Computers, IEEE Transactions on C-23* (12) (1974) 1264–1269.
- [14] R. Klette, B. Yip, Evaluation of curve length measurements, *International Conference on Pattern Recognition* 1 (2000) 1610.
- [15] F. De Vieilleville, J.-O. Lachaud, F. Feschet, Convex digital polygons, maximal digital straight segments and convergence of discrete geometric estimators, *Journal of Mathematical Imaging and Vision* 27 (2) (2007) 139–156.
- [16] D. Coeurjolly, Algorithmique et géométrie discrète pour la caractérisation des courbes et des surfaces, Ph.D. thesis, Université Lyon 2 (2002).
- [17] J.-O. Lachaud, Espaces non-euclidiens et analyse d’image : modèles déformables riemanniens et discrets, topologie et géométrie discrète, Habilitation à diriger des recherches, Université de Bordeaux 1 (2006).
- [18] F. Sloboda, B. Zlatko, J. Stoer, On approximation of planar one-dimensional continua, *Advances in Digital and Computational Geometry* (1998) 113–160.
- [19] T. Asano, Y. Kawamura, R. Klette, K. Obokata, Minimum-length polygons in approximation sausages, in: C. Arcelli, L. P. Cordella, G. S. Baja (Eds.), *Visual Form 2001*, Vol. 2059 of LNCS, Springer, 2001, pp. 103–112.

- [20] DGtal: Digital geometry tools and algorithms library, <http://libdgtal.org>.
- [21] W. Stein, et al., Sage Mathematics Software (Version 6.2.0), The Sage Development Team, <http://www.sagemath.org> (2014).
- [22] E. Baudrier, L. Mazo, Results on multigrid convergence speed of non-local estimators for concave functions, in preparation.