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

Loïc Mazo, Etienne Baudrier

Non-local estimators: a new class of multigrid convergent length estimators[1]

Loïc Mazo, Étienne Baudrier

ICube, University of Strasbourg, CNRS
300 Bd Sébastien Brant - CS 10413 - 67412 ILLKIRCH, FRANCE

Abstract

An interesting property for curve length digital estimators is the convergence toward the continuous length and the associate convergence speed when the grid spacing tends to zero. 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.

Keywords: discrete geometry, length estimation, multigrid convergence

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 grid spacing tends to zero. The local estimators based on a segmentation of the digital curve in patterns whose size is a constant that does not depends upon the grid spacing do not satisfy the convergence property even for straight line segments [15]. 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 [5]. The semi-local estimators [7], and the sparse estimators [18], both based on a segmentation of the curve in patterns whose size only depends upon the grid spacing, verifies the convergence

property[2] without convexity hypothesis, for smooth functional 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 about the convergence speed 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 $g : [a, b] \to \mathbb{R}$ $(a < b)$, its graph $\mathcal{C}(g) = \{(x, g(x)) \mid x \in [a, b]\}$ and a positive real number h, the *grid spacing*. We assume to have an orthogonal grid in the Euclidean space $\mathbb{R}^2$ whose set of grid points is $h\mathbb{Z}^2$.

The common methods to model the digitization of the graph $\mathcal{C}(g)$ with a grid spacing h 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

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

where $\lfloor \cdot \rfloor$ denotes the floor function. The set $\mathcal{D}(g, h)$ 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, $\mathcal{D}(g, h)$ 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 and by *discrete curve* we mean the graph of a function $\gamma \colon I \to \mathbb{Z}$ where I is an interval of $\mathbb{Z}$.

In the sequel of the article, for any function
$f \colon [a, b] \to \mathbb{R}$, $L(f)$ denotes the length of the graph $\mathcal{C}(f)$ according to Jordan's definition of length:

[2] Actually, the convergence of a semi-local estimator depends upon some choice made in its definition (see Section 2.4).

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

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

2.2. Local estimators

Local length estimators (see [12] 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 [15], 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 [22], 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 [5] on a test set of five closed curves. Moreover, the non-convergence is established in [6, 23] for almost all parabolas.

2.3. Adaptative 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 [14, 9, 5]. Efficient algorithms have been developed for segmenting curves or function graphs into MDSS and to compute their characteristics in a linear time [14, 10, 9]. 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

approximation[3] is also convex [11, Th. 13 and the proof [4]]: given a convex 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

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

It must be noticed that there exists a wrong version of the above equation in the literature [5]. Empirical MDSS multigrid convergence has also been tested in [5, 8] 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 [4, 5, 16]. The estimation is unique and has been proved to be multigrid convergent for smooth curves (of class C^3 with strictly positive curvature in [16], compact boundaries with positive reach in [17],). The convergence order is a $\mathcal{O}(h^{\frac{1}{3}})$ in [16] 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 [20] and the Approximation Saussage MLP (AS-MLP) introduced in [2]. 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 [8] shows that the MDSS size tends to 0 and their discrete length tends toward

[3] Though 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.

[4] The hypothesis on the convexity of the MDSS polygon is not assumed in the statement of the theorem but it appears in the proof.

[5] The 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 [19]. Accordingly, in [11], it is said that "the 'classical' value of ε_{DSS} is 1". A few month later, in [13], 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 [5] and led to an erroneous conclusion.

infinity as the grid spacing tends to zero. Thereby, one could ask whether combining a local estimation with an increasing window size as the grid spacing decreases would give a convergent estimator under more general assumptions and/or with simpler proofs of convergence. The following sections explore this question.

2.4. Semi-local length estimators

The notion of semi-local estimator appears in [7]. At a given grid spacing , 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 grid spacing on which the size of the segments depends.

Given a grid spacing h, a semi-local estimator segments a discrete curve in "patterns" of equal discrete size $H(h)$ and, possibly, a rest whose size is less than $H(h)$. By *pattern*, we mean a finite sequence of discrete points with consecutive abscissae. The function H which controls the discrete size of the patterns, that is the number of pixels in each pattern, is called *pattern function*.

The patterns of a semi-local length estimator have

- a discrete size $H(h)$ that tends to infinity as the grid spacing decreases toward zero (Property P_∞) while
- their true size $hH(h)$ tends to zero (Property P_0).

It is proved in [7] that the semi-local length estimators are multigrid convergent for functions of class C^2 with an error in $O(h^{\frac{1}{2}})$ for the best pattern size choice $H(h) = \Theta(h^{-\frac{1}{2}})$.

For a SL estimator, a pattern length should be close to the Euclidean distance between its extremities. 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 [18]. 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.

Under the hypothesis that the pattern function H has the properties P_∞ and P_0, it is proved in [18] that sparse length estimators are multigrid convergent for Lipschitz functions. Furthermore, for function of class C^2, the error rate is the same as for SL estimators. If, besides, the function is of class C^2 and concave, then the error is in $O(h)$ for the best pattern size choice $H(h) = \Theta(h^{-\frac{1}{2}})$.

On the one hand, with semi-local and sparse estimators, the pattern size is constant for a given grid spacing. This is important for an algorithmic point of view. Nevertheless, it does not really matter to prove the convergence and the properties P_∞ and P_0 could be expressed in terms of means. On the other hand, it has been proved in [8], under some hypotheses, that the average size

of the maximal digital straight segments of a contour verifies the properties P_∞ and P_0. 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 patterns by allowing variable lengths. We need also to allow very large patterns when the curve is close to straight segments.

3.1. Definition

The definition of non-local estimators involves some generalized means. 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, $M_{+\infty}((x_i)_{i=0}^n) = \max((x_i)_{i=0}^n)$ and $M_{-\infty}((x_i)_{i=0}^n) = \min((x_i)_{i=0}^n)$. For any $\alpha, \beta \in \overline{\mathbb{R}}$, one has[6]

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

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 $C(\sigma)$ 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 $1 + C(\sigma)^2 = \big(M_2(\sigma)/M_1(\sigma)\big)^2$.

Let Γ be a discretization with a grid spacing h of a continuous curve. The *pattern function* that we define thereafter produces a segmentation of the discrete curve Γ. The points put forward by this segmentation form a polyline joining the two extremities of Γ. Then the length of this polyline is a *non local length estimation* of the curve. Note that all the points of Γ are ignored in the estimation except the points used as vertices of the polyline. An Illustration of the definition is given Figure 1.

Definition 1 (Pattern function). Let $\alpha, \beta \in \overline{\mathbb{R}}$.

- A *pattern function* is a function that maps a discrete curve Γ and a grid spacing h to a partition of the domain of Γ.

Let C be a set of rectifiable functions.

[6] $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

- An *α-pattern function $\mathcal{A}$ on C* is a pattern function such that, for any rectifiable function $g \in C$,

$$\lim_{h \to 0} M_\alpha(\mathcal{A}(\mathcal{D}(g,h),h)) = +\infty, \tag{2}$$

- An *(α,β)-pattern function $\mathcal{A}$ on C* is an α-pattern function such that, for any rectifiable function $g \in C$,

$$\lim_{h \to 0} M_\beta(\mathcal{A}(\mathcal{D}(g,h),h)) \times h = 0. \tag{3}$$

An α-pattern function, resp. (α,β)-pattern function, is an α-pattern function, resp. (α,β)-pattern function, on the set of all rectifiable functions.

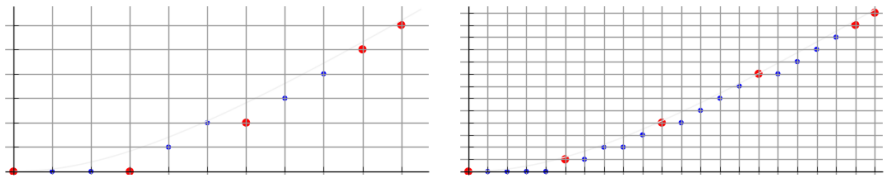
Assuming $\Gamma = \mathcal{D}(g,h)$ where g is a rectifiable curve, Equation (2) means that the pattern size in pixels tends in α-mean toward infinity as h tends toward 0. Equation (3) means that the absolute pattern size tends in β-mean toward 0 as h tends toward 0. For instance, looking at the continuous curve depicted in Figure 1, we could make the following thought experiment. We zoom in the figure to see the patterns evolve as the grid spacing h tends toward 0. We can imagine two ways to do so. Firstly, the size of the windows remains constant while the grid spacing decreases. Then, from Equation (2) we should see smaller and smaller patterns. Secondly, the grid remains unchanged while the window size increases. Then, from Equation (3) we should see larger and larger patterns. Figure 1 (b–d) shows two snapshots of this experiment (at $h = 0.2$ and $h = 0.1$ for two $(1,1)$-pattern functions, one with a (almost) regular partition of the domain and one with a somewhat adaptive partition (smaller curvature, larger patterns).

A *non-local multigrid length estimator* is the choice of an α-pattern function for some $\alpha \in \overline{\mathbb{R}}$. When, furthermore, we use an (α,β)-pattern function to construct our estimator, we say that we have an *M-sparse multigrid length estimator* (the 'M' in M-sparse stands for mean). The α parameter ensures that we do not pick too much points while the β parameter ensures that we pick enough points on the discrete curve to get the convergence. It is worthy to note that in the contrary of the conventional framework of continuous curves rectification from samples, rectification from discretized curves enjoins to use only few points among the available ones as shown the non-convergence of the local estimators.

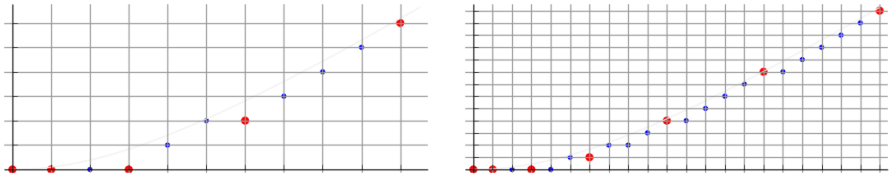
In the sequel, given a rectifiable function g and a grid spacing h, we denote by $L^{\mathrm{NL}}(\mathcal{A},g,h)$ the length of $\mathcal{D}(g,h)$ returned by the non-local estimator related to the pattern function $\mathcal{A}$. By abuse of notation, we also write $\mathcal{A}(g,h)$ instead of $\mathcal{A}(\mathcal{D}(g,h),h)$.

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 grid spacing. Thereby, the local estimators are not non-local estimators (this justifies the name given to the class of estimators introduced in this paper).



(a) $h = 0.2$ (b) $h = 0.1$



(c) $h = 0.2$ (d) $h = 0.1$

Figure 1: Two examples of $(1,1)$-pattern functions. (a–d) Blue points: the discrete curve which is the discretization of $x \mapsto \sqrt{1+x^2}-1$ on the interval $[0, 2.1]$. Red points: extremities of the subintervals provided by the pattern function for the considered grid spacings. (a–b) Patterns with a constant size (but the last) proportional to $\lfloor h^{-\frac{1}{2}} \rfloor$. (c–d) Patterns with increasing sizes $1, 2, 3, \ldots$ (positive integers).

Semi-local and sparse estimators. The subinterval-length sequence for sparse estimators is $(H(h), \ldots, H(h), rem)$ where $rem \leq H(h)$, $H(h)$ tends to infinity and $hH(h)$ tends to 0 as the grid spacing h tends to 0. Thus, it is plain that they are M-sparse length estimators related to $(1,1)$-pattern functions. 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.

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 [7] do not exactly comply to the definition of non-local estimators for two reasons. Firstly, they withdraw the last pattern if it has not the expected size, $H(h)$. Secondly, they make use of a weight function to get the length of the induced polyline. 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 the hypotheses that ensure their convergence apply that is when the weight function is closed to the diameter function.

MDSS based estimators. Though it has not been proved in the general case, it is likely that, in the framework of Gauss digitization, 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 [1]).

Indeed, in [8] it is shown that, 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 $\Theta(h^{-\frac{1}{3}})$ and a $\Theta(h^{\frac{1}{3}}\ln(h))$. 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 also valid for an MDSS segmentation of a discrete function graph. 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 grid spacing.

3.2. Convergence

Our main result (Theorem 8) is stated and proved in Section 3.3. As the price to pay to get a general, and nevertheless precise, theorem is a very cumbersome statement, we present in this section several corollaries which are obtained by restricting the hypotheses of Theorem 8. Thereby, the significance and the implications of each corollary can be more easily understood.

Since the results presented here are rather straightforward consequences of Theorem 8 whose proof is extensively detailed in the following section, no proofs of the corollaries are provided.

The first corollary states the convergence property of M-sparse estimators under the weaker assumptions for the Euclidean function. Nevertheless, this convergence requires that the maximum of the pattern absolute widths tends toward 0. This is much more restrictive than the hypothesis of convergence in mean used in Theorem 8.

Corollary 1. *Let $g\colon [a,b] \to \mathbb{R}$ be a Lipschitz function and $\mathcal{A}$ be a $(1,+\infty)$-pattern function. Then,*

$$\lim_{h\to 0} L^{\mathrm{NL}}(\mathcal{A}, g, h) = L(g).$$

In Corollary 1, due to the weak hypotheses, we cannot give any bound on the convergence speed.

Nevertheless it allows us to obtain the multigrid convergence for the class of Lipschitz functions, including fractal functions such as the function f_2 introduced in Subsection 4.1. The function f_2 is of class C^∞ almost everywhere, but its second derivative is not bounded. As a consequence, the function f_2 does not meet the hypotheses of Corollaries 2 and 3. Exhibiting a fractal function may be surprising, but having properties on function class including fractal ones has two motivations:

1. fractal curves can be found in natural scenes and then as segmented digital object boundaries in the multigrid framework,
2. fractal functions emphasize some length estimator properties that can be hidden otherwise.

Assuming a smooth enough function, we can retrieve for M-sparse estimators the convergence speed obtained in [18] for sparse estimators under somewhat weaker hypotheses.

Corollary 2. *Let $g\colon [a,b] \to \mathbb{R}$ be a differentiable function whose derivative is Lipschitz and $\mathcal{A}$ be a $(1,1)$-pattern function such that $c_v(\mathcal{A}(g,h))$ is upper-bounded as $h \to 0$. Then,*

$$L(g) - L^{\mathrm{NL}}(\mathcal{A},g,h) = O(hM(h)) + O\Big(\frac{1}{M(h)}\Big) \tag{4}$$

where $M(h) = M_1(\mathcal{A}(g,h))$.

In Equation (4) the constants in the big O only depend on $b-a$, $\max g'$ and $\max g''$ and can be explicitly computed.The optimal convergence speed rate in $h^{\frac{1}{2}}$ is then obtained by choosing the pattern function such that $M(h) = O(h^{-\frac{1}{2}})$. Note that in Corollary 2 the pattern function is assumed to be $(1,1)$, which is weaker than the $(1,\infty)$-pattern function asked in Corollary 1. Then the convergence comes mainly from the smoothness hypothesis.

To handle the case of functions that are not differentiable on their domain, we define $\mathcal{H}(\mathcal{A}(g,h))$ as the (Lebesgue) measure of the union of the $\mathcal{A}(g,h)$ open subintervals on which g is not differentiable. For instance, if the function g is differentiable on (a,b) but finitely many points, then $\mathcal{H}(\mathcal{A}(g,h))$ is upper bounded by $nhM_\infty(\mathcal{A}(g,h))$ where n is the number of points in (a,b) where g is not differentiable. In Appendix A, we compute an upper bound of $\mathcal{H}(\mathcal{A}(g,h))$ for a fractal function with infinitely many points where g is not differentiable (infinitely many isolated points and infinitely many limit points).

Corollary 3. *Let $g\colon [a,b] \to \mathbb{R}$ be a Lipschitz function whose derivative is k-Lipschitz on each interval included in its domain of definition for some $k > 0$. Let $\mathcal{A}$ be a $(1,+\infty)$-pattern function. Then,*

$$\begin{aligned} L(g) - L^{\mathrm{NL}}(\mathcal{A},g,h) = \\ O(hM'(h)) + O\Big(\frac{1}{M(h)}\Big) + O(\mathcal{H}(h)) \end{aligned} \tag{5}$$

where $M' = M_\infty(\mathcal{A}(g,h))$, $M = M_1(\mathcal{A}(g,h))$ and $\mathcal{H}(h) = \mathcal{H}(\mathcal{A}(g,h))$.

The upper bound convergence rate depends on $\mathcal{H}(\mathcal{A}(g,h))$ and is not known in general.

3.3. Proof of the main theorem

In this section, we give some sufficient conditions under which the non-local length estimators are convergent for Lipschitz functions. Moreover, Theorem 8 gives a bound on the error at grid spacing h for Lipschitz functions whose derivative is k-Lipschitz ($k > 0$) on any interval included in their domain (of definition).

Notations. In the remainder of the article, we use some notations that we now present. The first ones concern euclidean objects. Thereby, they do not depend upon the grid spacing. The others are related to the grid spacing h and should

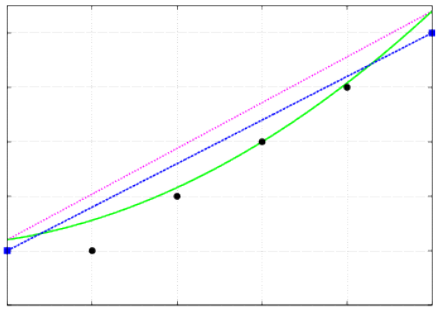


Figure 2: The two main parts of the estimation error: the curve g (in green, solid) to its chord $g_{\mathrm{c}}^{\mathcal{A}}$ (in magenta, dotted-dashed) then the curve chord to the chord $\lfloor g_{\mathrm{c}}^{\mathcal{A}} \rfloor$ (in blue, dashed) of the digitized curve $\mathcal{D}(g,h)$ (black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

be indexed by h. Nevertheless, as we never have to work with two different grid spacings, we prefer to omit the h index to lighten the notations.

$I = [a,b]$ is an interval of $\mathbb{R}$ with a non-empty interior and $g\colon I \to \mathbb{R}$ is a Lipschitz function whose derivative is denoted g' (from Rademacher's theorem, g is differentiable almost everywhere). We also define the function $\varphi\colon \mathbb{R} \to \mathbb{R}$ by $\varphi(x) = \sqrt{1+x^2}$. Thus, one has $L(g) = \int_{[a,b]} \varphi \circ g'$.

Given some grid spacing $h > 0$, A, resp. B, is the smallest, resp. largest, integer such that $Ah \in I$, resp. $Bh \in I$. The functions g_{l}, g_{c}, g_{r} are resp. the restrictions of the function g to the intervals $[a, Ah]$, $[Ah, Bh]$, $[Bh, b]$. For any pattern function $\mathcal{A}$, we write $M_\alpha^{\mathcal{A}}$, resp. $C^{\mathcal{A}}$, instead of $M_\alpha(\mathcal{A}(g,h))$, resp. $C(\mathcal{A}(g,h))$, when there is no ambiguity. The number of subintervals in the partition $\mathcal{A}(g,h)$ is denoted $N^{\mathcal{A}}$, or just N when possible and the integers defining the partition $\mathcal{A}(g,h)$ are $A = a_0 < a_1 < \cdots < a_N = B$. Finally, we define two piecewise affine functions, $g_{\mathrm{c}}^{\mathcal{A}}$ and $\lfloor g_{\mathrm{c}}^{\mathcal{A}} \rfloor$, interpoling the continuous function g_{c} and its discretization (which is equal to $\mathcal{D}(g,h)$) according to the pattern function $\mathcal{A}$. The graph of $g_{\mathrm{c}}^{\mathcal{A}}$, resp $\lfloor g_{\mathrm{c}}^{\mathcal{A}} \rfloor$, is the polyline linking the points $\big(a_i h, g(a_i h)\big)_{i=0}^{N}$ which are on $\mathcal{C}(g)$, resp. the grid points $\big(a_i h, \lfloor \frac{g(a_i h)}{h} \rfloor h\big)_{i=0}^{N}$.

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

The proof of Theorem 8 can be split in three parts. Proposition 4 gives a bound on the error due to the ignorance of the exact abscissas of the curve extremities. Proposition 5 and Proposition 6 evaluate the difference between the length of the curves $\mathcal{C}(g_{\mathrm{c}})$ and $\mathcal{C}(g_{\mathrm{c}}^{\mathcal{A}})$ for a given pattern function $\mathcal{A}$ under different assumptions. Given two pattern functions $\mathcal{A}$ and $\mathcal{B}$, Proposition 7 evaluates the difference between the length of the curves $\mathcal{C}(g_{\mathrm{c}}^{\mathcal{A}})$ and $\mathcal{C}(\lfloor g_{\mathrm{c}}^{\mathcal{B}} \rfloor)$. 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 4. *For any k-Lipschitz function g, we have*

$$L(g_{\mathrm{l}}) + L(g_{\mathrm{r}}) \leq 2\varphi(k)h.$$

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

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

where $g_{|[c,d]}$ denotes the restriction of the function g to the interval $[c,d]$. In particular, $L(g_{\mathrm{l}}) \leq \varphi(k)(Ah-a) \leq \varphi(k)h$ and $L(g_{\mathrm{r}}) \leq \varphi(k)(b-Bh) \leq \varphi(k)h$.

We now look at the difference between the length of $\mathcal{C}(g_{\mathrm{c}})$ and the length of the polyline $\mathcal{C}(g_{\mathrm{c}}^{\mathcal{A}})$. These two curves share the same extremities but these depend on the grid spacing h.

Proposition 5. *For any pattern function $\mathcal{A}$ such that $M_{+\infty}^{\mathcal{A}} \times h \to 0$ as $h \to 0$ and any Lipschitz function g, we have*

$$\lim_{h\to 0} L(g_{\mathrm{c}}^{\mathcal{A}}) - L(g_{\mathrm{c}}) = 0.$$

PROOF. Let $h > 0$. From the discrete partition $\mathcal{A}(g,h) = (a_i)_{i=0}^N$ we define a partition σ of the interval $[a,b]$ as $a \leq a_0 h < a_1 h < \cdots < a_N h \leq b$. Let L_σ be the length of the polyline interpolating g according to the partition σ. On the one hand, since $M_{+\infty}^{\mathcal{A}} \times h \to 0$ as $h \to 0$, $Ah - a < h$ and $b - Bh < h$, the norm of the partition σ tends to 0 as $h \to 0$. Then the Lipschitz-continuity of g ensures that $L(g) - L_\sigma \to 0$ as $h \to 0$. On the other hand, it is plain that $L(g_{\mathrm{c}}) - L(g_{\mathrm{c}}^{\mathcal{A}}) \leq L(g) - L_\sigma(g)$ for any h. We conclude straightforwardly.

When the derivative of g is Lipschitz continuous, the next Proposition gives us a bound on the difference between the length of the curve $\mathcal{C}(g_{\mathrm{c}})$ and the length of the polyline $\mathcal{C}(g_{\mathrm{c}}^{\mathcal{A}})$. In order to deal with the set D of points where g is not differentiable, we need first to explain how we calculate the measure $\mathcal{H}(\mathcal{A}(g,h))$ which is defined by $\mathcal{H}(\mathcal{A}(g,h)) = h\sum_{i\in\mathcal{I}} a_i - a_{i-1}$ where $\{a_i\}_{i=0}^N = \mathcal{A}(g,h)$ and $\mathcal{I} = \{i \in [0,N] \mid (a_{i-1}, a_i) \cap D \neq \emptyset\}$. When there is no ambiguity, we will write $\mathcal{H}^{\mathcal{A}}$ instead of $\mathcal{H}(\mathcal{A}(g,h))$. We also define a less specific measure, $\mathcal{H}(g,\delta) = \sup \sum_i w_i$ where the supremum is over all the minimal coverings [7] of $D \setminus \{a,b\}$ by disjoint open subintervals of (a,b) with diameters $w_i \leq \delta$. The real $\mathcal{H}^{\mathcal{A}}$ is upper bounded by $\mathcal{H}(g, hM_{+\infty}^{\mathcal{A}})$. In Appendix A, we detail the calculus of $\mathcal{H}(g,\delta)$ for a fractal function with infinitely many points where g is not differentiable (infinitely many isolated points and infinitely many limit points).

[7] A covering by a family E of sets is minimal if no subfamily of E is itself a covering.

Proposition 6. *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 grid spacing $h > 0$*

$$L(g_{\mathrm{c}}) - L(g_{\mathrm{c}}^{\mathcal{A}}) \leq ThM_1^{\mathcal{A}} \left(1 + (C^{\mathcal{A}})^2\right) + U\mathcal{H}^{\mathcal{A}} \tag{7}$$

where $T = k_2(b-a)/2$ and $U = (\varphi(k_1) - 1)$.

Proof. Let $h > 0$. We consider the partition $\mathcal{A}(g,h) = (a_i)_{i=0}^N$. Let $\mathcal{J} \subset [1, N] \cap \mathbb{Z}$ be the set of subscripts i such that g is differentiable on $(a_{i-1}h, a_i h)$. 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
$(a_{i-1}h, a_i h)$, $i \in \mathcal{J}$, there exists a real t_i such that

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

Then, if $i \in \mathcal{J}$, the length ℓ_i of the restriction of $\mathcal{C}(g_{\mathrm{c}}^{\mathcal{A}})$ to the interval $[a_{i-1}h, a_i h]$ is such that

$$\ell_i = h(a_i^r - a_{i-1}^r)\, \varphi\big(g'(t_i)\big).$$

Otherwise, we have the obvious lower bound

$$\ell_i \geq h(a_i - a_{i-1}).$$

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

$$L\big(g_{|[a_{i-1}h, a_i h]}\big) \leq \varphi(k_1) h (a_i - a_{i-1})$$

Then,

$$L(g_{\mathrm{c}}) - L(g_{\mathrm{c}}^{\mathcal{A}}) \leq \sum_{i \in \mathcal{J}} \int_{a_{i-1}h}^{a_i h} \varphi \circ g'(t) - \varphi \circ g'(t_i)\, \mathrm{d}t \quad + (\varphi(k_1) - 1) \sum_{i \notin \mathcal{J}} h(a_i - a_{i-1})$$

So,

$$\begin{aligned}
L(g_{\mathrm{c}}) - L(g_{\mathrm{c}}^{\mathcal{A}}) \leq & \\
& \sum_{i \in \mathcal{J}} \int_{a_{i-1}h}^{a_i h} k_2 \left| t - t_i \right| \, \mathrm{d}t + (\varphi(k_1) - 1)\mathcal{H}^{\mathcal{A}} \\
& \leq \sum_{i=1}^{N} k_2 h^2 \frac{(a_i - a_{i-1})^2}{2} + (\varphi(k_1) - 1)\mathcal{H}^{\mathcal{A}} \\
& \leq \frac{1}{2} k_2 h^2 N (M_2^{\mathcal{A}})^2 + (\varphi(k_1) - 1)\mathcal{H}^{\mathcal{A}} \\
& \leq \frac{1}{2} k_2 (b-a) h \frac{(M_1^{\mathcal{A}})^2 + V}{M_1^{\mathcal{A}}} + (\varphi(k_1) - 1)\mathcal{H}^{\mathcal{A}} \\
& \qquad \text{where } V \text{ is the variance of } \mathcal{A}(g,h) \\
& \leq \frac{1}{2} k_2 (b-a) h M_1^{\mathcal{A}} \left(1 + (C^{\mathcal{A}})^2\right) + (\varphi(k_1) - 1)\mathcal{H}^{\mathcal{A}}
\end{aligned}$$

where the penultimate inequality comes from the well known formula $V = (M_2^{\mathcal{A}})^2 - (M_1^{\mathcal{A}})^2$ and from the relation

$$\begin{aligned} M_1^{\mathcal{A}} &= \frac{1}{N}\sum_{i=1}^{N}(a_i - a_{i-1}) \\ &= \frac{1}{N}(a_N - a_0) \\ &\leq \frac{1}{Nh}(b-a). \end{aligned}$$

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 $\mathcal{H}^{\mathcal{A}} = 0$. Then, the difference $L(g_{\mathrm{c}}) - L(g_{\mathrm{c}}^{\mathcal{A}})$ is a $O\big(hM_1^{\mathcal{A}}(1 + (C^{\mathcal{A}})^2)\big)$.

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

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

Given a Lipschitz function g on an interval I and a pattern function $\mathcal{A}$, we now look at the difference between $L(g_{\mathrm{c}}^{\mathcal{A}})$ and $L(\lfloor g_{\mathrm{c}}^{\mathcal{A}} \rfloor)$ that is between the length of two piecewise affine functions. Moreover, to handle the MDSS length estimator, we shall have to use Proposition 7 with two different pattern functions $\mathcal{A}$ and B, that is to compare $L(g_{\mathrm{c}}^{\mathcal{A}})$ and $L(\lfloor g_{\mathrm{c}}^{\mathcal{B}} \rfloor)$.

Proposition 7. *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,

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

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

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 8. *Let $g\colon [a,b] \to \mathbb{R}$ 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 grid spacing h,

$$\| \lfloor g_{\mathrm{c}}^{\mathcal{A}} \rfloor - \lfloor g_{\mathrm{c}}^{\mathcal{B}} \rfloor \|_\infty \leq \omega h \tag{8}$$

then

- *if $\beta = +\infty$, the non-local estimation $L^{\mathrm{NL}}(g, \mathcal{A}, h)$ converges toward the length of the curve $\mathcal{C}(g)$ as h tends to 0;*
- *if g' is k_2-Lipschitz on each interval included in its domain, we have*

$$\begin{aligned} &L(g) - L^{\mathrm{NL}}(g, \mathcal{A}, h) \leq \\ &\quad Sh && \text{(side error)} \\ &+\ ThM_1^{\mathcal{B}}(1 + (C^{\mathcal{B}})^2) + U\mathcal{H}^{\mathcal{B}}) && \text{(discretization error)} \\ &+\ V\left(\tfrac{1}{M_1^{\mathcal{A}}} + \tfrac{1}{M_1^{\mathcal{B}}}\right) && \text{(quantization error)} \end{aligned} \tag{9}$$

where $S = 2\varphi(k_1)$, $T = k_2(b-a)/2$, $U = \varphi(k_1) - 1$ and $V = (1 + 2\omega)\varphi'\Big(k_1 + 1/M_{-1}^{\mathcal{A}}\Big)(b-a)$.

Furthermore, if $\mathcal{B}(g,h) \subseteq \mathcal{A}(g,h)$*, the term* $1/{M_1^{\mathcal{A}}} + 1/{M_1}^{\mathcal{B}}$ *in the right hand side of Equation 9 can be replaced by* $1/M_1^{\mathcal{B}}$*.*

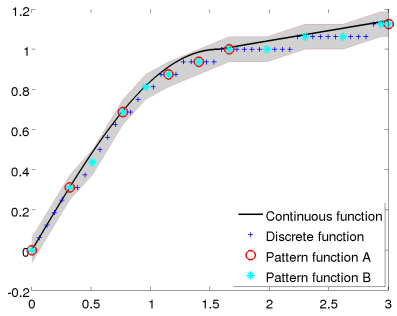


Figure 3: Hypotheses of Theorem 8. The partition related to $\mathcal{A}(g,h)$ splits the discrete curve in patterns whose average size in pixels tends toward infinity as the grid spacing tends toward zero. Another partition of the curve, related to $\mathcal{B}(g,h)$, has patterns whose average absolute size tends toward zero as the grid spacing tends toward zero. A tube (in gray) with constant height contains the discrete points related to the partitions $\mathcal{A}(g,h)$ and $\mathcal{B}(g,h)$.

Proof. We write the difference between $L(g)$ and $L(\lfloor g_{\mathrm{c}}^{\mathcal{A}} \rfloor)$ as the sum of three terms that correspond to the three kinds of errors.

$$L(g) - L(\lfloor g_{\mathrm{c}}^{\mathcal{A}} \rfloor) = (L(g_{\mathrm{l}}) + L(g_{\mathrm{r}})) + \big(L(g_{\mathrm{c}}) - L(g_{\mathrm{c}}^{\mathcal{B}})\big) + \big(L(g_{\mathrm{c}}^{\mathcal{B}}) - L(\lfloor g_{\mathrm{c}}^{\mathcal{A}} \rfloor)\big) \quad (10)$$

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

- From Proposition 4, we have

$$L(g_{\mathrm{l}}) + L(g_{\mathrm{r}}) \leq 2\varphi(k_1)h \quad (11)$$

  Thus, $L(g_{\mathrm{l}}) + L(g_{\mathrm{r}})$ converges to 0 as $h \to 0$.

- From Proposition 5, we know that $L(g_{\mathrm{c}}^{\mathcal{B}})$ converges toward $L(g_{\mathrm{c}})$ as $h \to 0$ provided $\beta = +\infty$.
  Moreover, when g' is k_2-Lipschitz continuous on the intervals included in its domain and whatever is the value of β, from Proposition 6 we have

$$L(g_{\mathrm{c}}) - L(g_{\mathrm{c}}^{\mathcal{B}}) \leq \frac{k_2(b-a)}{2} h M_1^{\mathcal{B}} \left(1 + (C^{\mathcal{B}})^2\right) + (\varphi(k_1) - 1)\mathcal{H}^{\mathcal{B}}. \quad (12)$$

- From Hypothesis (8),

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

Thus, we have

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

Since g is k_1-Lipschitz, the piecewise affine function $g_{\mathrm{c}}^{\mathcal{B}}$ is clearly k_1-Lipschitz. Let $\mathcal{A}(g,h) = (a_i)_{i=0}^{N^{\mathcal{A}}}$. For any $i \in [1, N^{\mathcal{A}}]$, the absolute slope of $\lfloor g_{\mathrm{c}}^{\mathcal{A}} \rfloor$ on $[a_{i-1}h, a_i h]$ is bounded by $k_1 + 1/(a_i - a_{i-1})$. Hence, the arithmetic mean of the absolute slopes of $\lfloor g_{\mathrm{c}}^{\mathcal{A}} \rfloor$ is bounded by $k = k_1 + 1/M_{-1}^{\mathcal{A}}$ (thus $k \leq k_1 + 1$). Then, from Proposition 7, we derive that

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

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

$$M_1\big(\mathcal{I}(g,h)\big) = \frac{1}{N^{\mathcal{I}}}(B - A) \leq \frac{1}{h N^{\mathcal{I}}}(b - a).$$

Thus, we get

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

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

Thereafter, the convergence is established once $\beta = +\infty$. Moreover, Equation (9) derives obviously from (11), (12) and (13).

Basically, Theorem 8 states that a non-local estimator (which rely on an α-pattern function) is convergent for Lipschitz functions provided it is not too far from a M-sparse estimator (which rely on an (α, β)-pattern function). 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 $B - A \approx (b-a)/h$. 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 8 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 $g_{\mathrm{c}}^{\mathcal{B}}$ provides a tube around the

data points attached to the partition $\mathcal{B}(g,h)$ in which must stay the function $\lfloor g_c^{\mathcal{A}} \rfloor$ to reliably estimate $L(g)$. That way 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 [8], 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 $M_{+\infty}^{\mathcal{A}} \times h \to 0$ as $h \to 0$. On the other hand, for C^3 curves with almost linear parts, even with $\mathcal{B} \supset \mathcal{A}$, the real ω would not be null. Besides, 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 7). 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):

$$L(g) - L^{\mathrm{NL}}(g,\mathcal{A},h) \leq Sh + \inf_i \left(ThM_{1,i}\left(1+(C_i)^2\right) + U\mathcal{H}^{\mathcal{A}} + V\left(\frac{1}{M_{1,i}} + \frac{1}{M_1^{\mathcal{A}}}\right)\right)$$

where the infimum is over all 1-pattern functions $\mathcal{B}_i$ such that $M_{1,i} = M_1^{\mathcal{B}_i} = o(1/h)$ as $h \to 0$ and where $C_i = C(\mathcal{B}_i(g,h))$.

With sparse estimators, the partition subinterval length sequence provided by the pattern function is

$$(H(h), \ldots, H(h), rem)$$

where

$$\lim_{h\to 0} H(h) = +\infty, \ \lim_{h\to 0} hH(h) = 0 \ \text{ and } \ rem \leq H(h).$$

Thus, taking $\mathcal{B} = \mathcal{A}$, $\omega = 0$ and $\beta = 1$ all the hypotheses of Theorem 8 are satisfied.

Theorem 8 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 2 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 $\mathcal{E}^{\mathrm{Sp2}}$ and $\mathcal{E}^{\mathrm{Sp3}}$ with
  1. $\mathcal{E}^{\mathrm{Sp2}}$: the discrete pattern size equal to $\lfloor h^{-\frac{1}{2}} \rfloor$ which is generally optimal for such an estimator (see [18]),

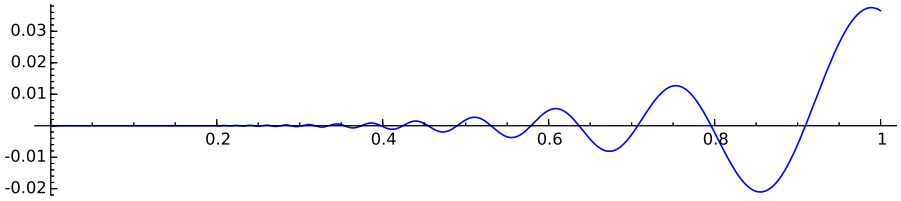


Figure 4: Graph of the function s_1.

2. $\mathcal{E}^{\mathrm{Sp3}}$: the discrete pattern size equal to $\lfloor h^{-\frac{1}{3}} \rfloor$ that approximatively corresponds to the asymptotic average discrete size of the MDSS on a smooth curve with positive minimal curvature [8].

- The M-sparse estimator $\mathcal{E}^{\mathrm{rand}}$ with random pattern sizes equidistributed between 1 and $\lfloor 2h^{-\frac{1}{2}} \rfloor$. Hence, for small enough grid spacings, the average pattern size should be the same as $\mathcal{E}^{\mathrm{Sp2}}$. Note that for the three previous NLE, we can apply Theorem 8 taking $\mathcal{A} = \mathcal{B}$.
- The MDSS based estimator $\mathcal{E}^{\mathrm{DSS}}$ (taking the diameter as the pattern length – see the comment at the end of Section 3.1).

The computed errors are the discretization, the quantization and the total error. The bounds and the grid spacing h are chosen such that the error on the bounds is null.

There are five test functions that we describe now. The first three satisfy the stronger assumptions on the function g in Theorem 8 (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^∞.
- 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 8 but, unlike the function

s_1, the derivative of s_2 is not Lipschitz on $(0, 1]$. Then the bound obtained in Proposition 6 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.

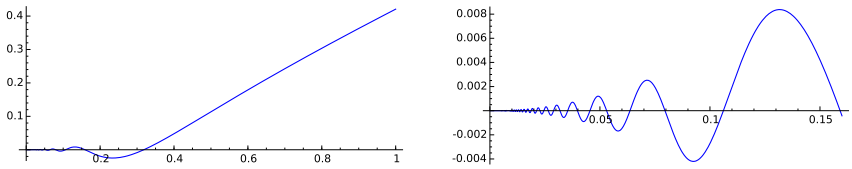


Figure 5: Graph of the function s_2.

- 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 $i \in \{1, 2\}$, $f_i = \lim_{n\to\infty} f_{i,n}$ where

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

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

  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 8 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 grid spacing tends to zero.

Remark 1. The pattern functions of $\mathcal{E}^{\mathrm{DSS}}$ for the tested functions on the tested grid spacings have not contradict the assumption that, in our framework, MDSS are non-local estimators. Furthermore, the bounds $h^{-1/3}$ and $h^{-\frac{1}{3}} \log(1/h)$ for the asymptotic average discrete size of the MDSS on a smooth solid contour with positive minimal curvature [8] fit our experiment on function graphs (see Figure 8).

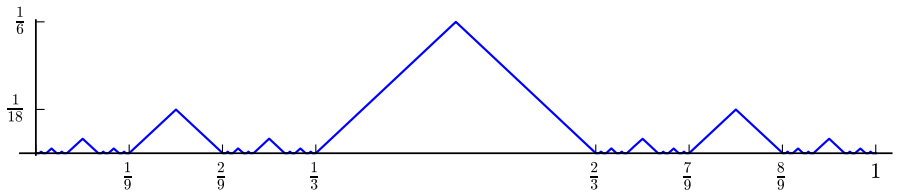


Figure 6: Graph of the fractal function f_1.

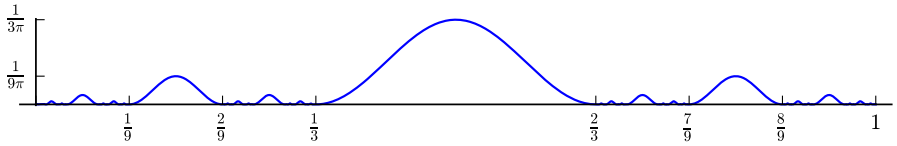


Figure 7: Graph of the fractal function f_2.

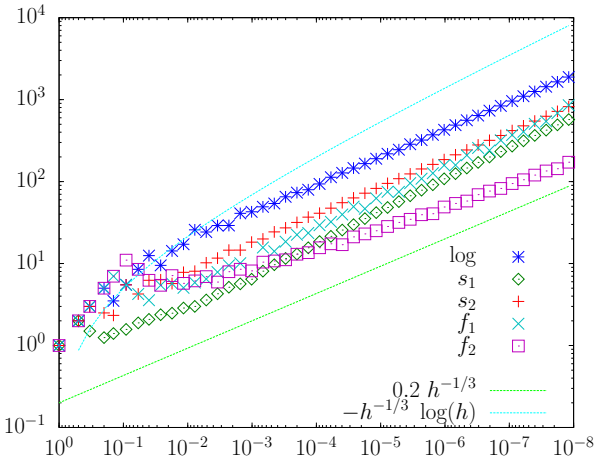


Figure 8: Average discrete size of the MDSS produced by $\mathcal{E}^{\mathrm{DSS}}$ on the tested functions according to the grid spacing.

4.2. On the theoretical bounds

The tests gathered in Table 2 aim at comparing the theoretical bounds obtained in Theorem 8 with the experimental values of the discretization error and the quantization error. To do so, we use the NLE $\mathcal{E}^{\mathrm{rand}}$. In the table, the blue '+' represents the theoretical upper bound and the green '*' represents the observed error for the grid spacings $\lfloor (2/3)^n \rfloor$, $n \in [1, 50] \cap \mathbb{N}$.

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

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

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 Lispchitz continuous ($k_2 = \infty$), the blue '+' on the plot represent the values of $hM_1^{\mathcal{A}}(1+(C^{\mathcal{A}})^2)$.

Since $\mathcal{E}^{\text{rand}}$ has pattern sizes equidistributed in the interval $[1, 2h^{-1/2}]$, for small enough grid spacings we have

$$hM_1^{\mathcal{A}}(1+(C^{\mathcal{A}})^2) \approx \frac{4}{3} h^{\frac{1}{2}}$$

and thereby

$$\frac{k_2(b-a)}{2} hM_1^{\mathcal{A}}(1+(C^{\mathcal{A}})^2) \approx \frac{2}{3} k_2 h^{\frac{1}{2}}.$$

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 $h \mapsto \Gamma h^{\frac{1}{2}}$ where $\Gamma = \frac{2}{3}k_2$ ($\Gamma = \frac{4}{3}$ for s_2 and f_2).

It can be seen in Table 2 that the observed errors for the functions log and s_1 are less than the upper bounds given by Theorem 8. Specifically, the power in the measured errors are almost twice the bounds, as far as the grid spacing is small enough. We know from our previous work [18], that convexity allow to double the 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

$$(\varphi(1)-1)\mathcal{H}^{\mathcal{A}}(f_1, h).$$

An upper bound for $\mathcal{H}^{\mathcal{A}}(f_1, h)$ is calculated in Appendix A:

$$\mathcal{H}^{\mathcal{A}}(f_1, h) \leq \mathcal{H}(f_1, \delta) \leq \frac{15}{4} \delta^{1-\log_3(2)}$$

where $\delta = hM_{+\infty}^{\mathcal{A}}$.

With $\mathcal{E}^{\text{rand}}$, $\delta = hM_{+\infty}^{\mathcal{A}} \approx 2h^{\frac{1}{2}}$ (for small enough grid spacings). Thus, the bound for the function f_1 is approximatively

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

The cyan line on the plot of the error for the function f_1 is the graph of $h \mapsto \Gamma h^{\frac{1}{2}(1-\log_3(2))}$ 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 $\delta = hM_{+\infty}^{\mathcal{A}}$ in the bound, which corresponds to the worst case, we take $\delta = hM_1^{\mathcal{A}}$, that is we cover the points of non differentiability by patterns of average size, both point clouds almost overlap.

Quantization error. The upper bound for the quantization error is derived from Proposition 7 and the function just needs to 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

$$\varphi'\Big(k_1 + \tfrac{1}{M_{-1}^{\mathcal{A}}}\Big)(b-a)\frac{1}{M_1^{\mathcal{A}}}$$

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

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

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 $h \mapsto \Gamma h^{\frac{1}{2}}$ where $\Gamma = \varphi'(k_1)$.

With the quantization error, the results for all tested functions are less than the upper bounds given by Theorem 8. Again the powers in the measured errors are almost twice the bounds, as far as the grid spacing is small enough. As with the discretization error, this corresponds to a property of concave functions shown in [18] 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 grid spacing.

Let us choose a grid spacing $h = k^{-2}/4$, $k \in \mathbb{N}$ and we assume that, at this grid spacing, the pattern size is constant, equals to $h^{-1/2} = 2k$. Now, we define the periodic piecewise affine function

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

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 $(1-2a)h \approx h$. On Table 1, we show the relative difference between the quantization error (QE) and its upper bound (UB) in Theorem 8 for some functions $v_{k,a}$ where $a = 10^{-3}$.

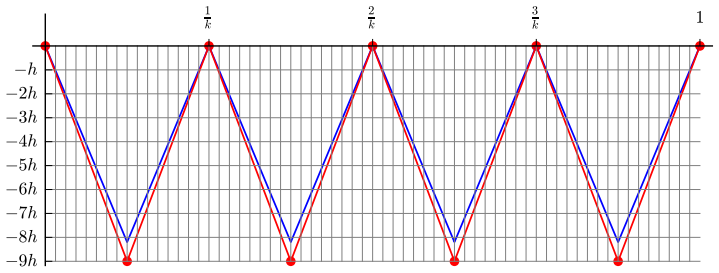


Figure 9: Blue (upper) line: the graph of the function $v_{4,0.1}$. Red points: MDSS end points of the $v_{4,0.1}$ OBQ discretization. Red line: polyline whose length is the non local length estimation of $L(g)$.

k	1	2^2	2^4	2^6	2^8	2^{10}	2^{12}
h	2^{-2}	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 1: Relative difference between the quantization error for the function $v_{k,a}$ with $a = 10^{-3}$ and the upper bound provided by Theorem 8

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 the Logarithm, the dumped 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 $\mathcal{E}^{\mathrm{MDSS}}$ and $\mathcal{E}^{\mathrm{Sp3}}$ which have a pattern size average of about $h^{-\frac{1}{3}}$ converge more quickly than $\mathcal{E}^{\mathrm{Sp2}}$ and $\mathcal{E}^{\mathrm{rand}}$ which have a pattern size average of $h^{-\frac{1}{2}}$. The good performance of $\mathcal{E}^{\mathrm{MDSS}}$ on the f_1 graph which is composed of straight line segments is natural. Contrariwise, as MDSS are adaptive, we could expect that $\mathcal{E}^{\mathrm{MDSS}}$ outperforms $\mathcal{E}^{\mathrm{Sp3}}$ on s_1, s_2 and f_2. Surprisingly this is not the case on the graph of s_1.

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 $\mathcal{E}^{\mathrm{MDSS}}$ 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 $\mathcal{E}^{\mathrm{MDSS}}$ 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. On the dumped sinusoid s_2, the total error for the $\mathcal{E}^{\mathrm{MDSS}}$ estimator benefits from the contrary signs, with same order of magnitude, of the QE and the DE. We note also that the sparse estimator $\mathcal{E}^{\mathrm{Sp2}}$ and its randomized equivalent, the M-sparse estimator $\mathcal{E}^{\mathrm{rand}}$ 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 nonlocal 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 grid spacing tends toward zero. 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 lengths tends toward 0 as the grid spacing itself tends toward zero. We proved that any NLE has the multigrid convergence property for the Lipschitz functions as soon as at any grid spacing 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 [18]. This point will be developed in a work in preparation [3]. 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).

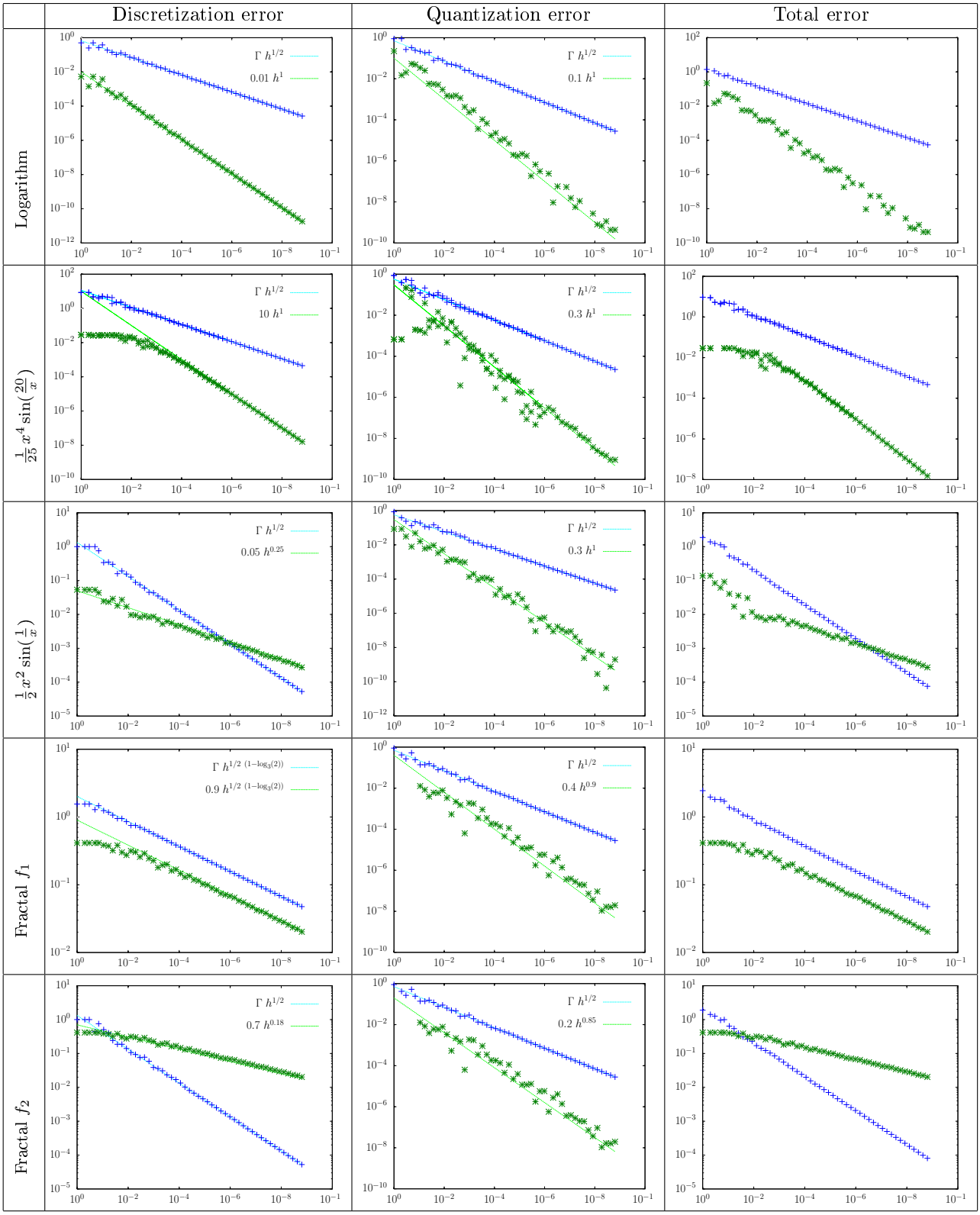


Table 2: Errors in multigrid length estimations with the NLE $\mathcal{E}^{\text{rand}}$. Green stars: the observed errors. Blue crosses: the theoretical upper bounds given by Theorem 8 (Γ is a constant computed from the expected means of the random pattern sizes; see text).

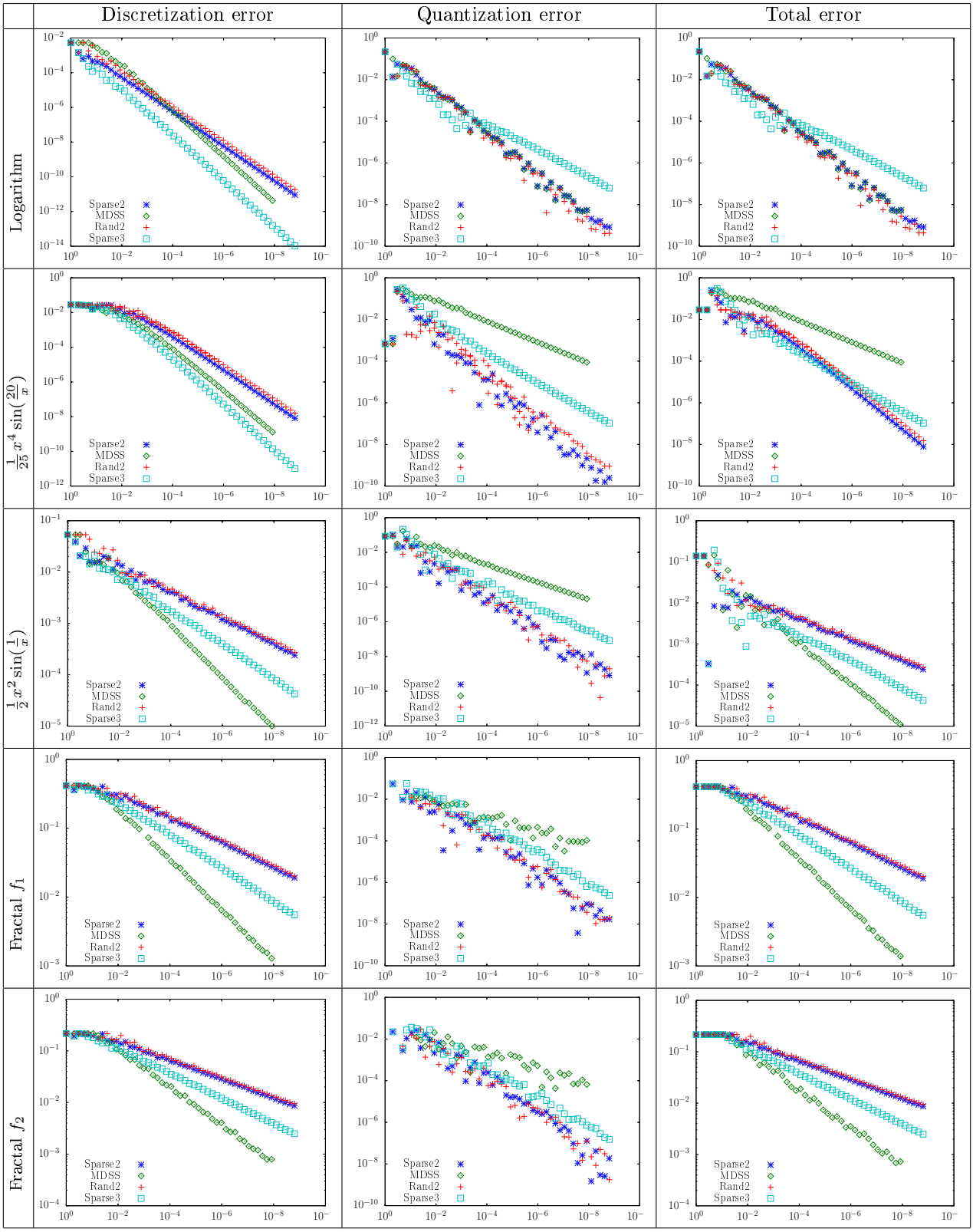


Table 3: Estimation errors in function of the grid spacing h for the estimators $\mathcal{E}^{\mathrm{Sp2}}$, $\mathcal{E}^{\mathrm{DSS}}$, $\mathcal{E}^{\mathrm{rand}}$ and $\mathcal{E}^{\mathrm{Sp3}}$. The tested functions are log, s_1, s_2, f_1 and f_2 (see text).

Appendix A. Calculus of $\mathcal{H}(f,\delta)$

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 $F(\delta) = \mathcal{H}(f,\delta)$. Recall that

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

where the supremum is over the set $\mathcal{E}_\delta$ of all minimal coverings of the set E by disjoint open subintervals of $(0,1)$ with diameters $w_i \leq \delta$.

For $\delta \geq \frac{1}{6}$, the set $\{I_i\}_{i=0}^{6}$ where $I_i = (\frac{2i-1}{12}, \frac{2i+1}{12}) \cap [0,1]$ is a minimal covering of E with open disjoint intervals.
Hence, for any $\delta \geq \frac{1}{6}$, we have

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

When $\delta < \frac{1}{6}$, any minimal covering in $\mathcal{E}_\delta$ can be split in three parts: a minimal covering of $E \cap [0, x_1]$ where $x_1 \in (\frac{1}{3}, \frac{1}{2})$, an interval that contains the point $\frac{1}{2}$ and a minimal covering of $E \cap [x_2, 1]$ where $x_2 \in (\frac{1}{2}, \frac{2}{3})$.

From the recursive definition of f, one can see that

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

where $f_{|J}$ denotes the restriction of the function f to an interval J.
Thus, for any $\delta < \frac{1}{6}$,

$$F(\delta) = 2 \times (\tfrac{1}{3}(F(3\delta) + \delta) \ + \ \delta = \tfrac{2}{3}F(3\delta) + 3\delta \tag{A.2}$$

By solving the recurrence relation (A.2) with the initial equality (A.1), we get

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

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] DGtal: Digital geometry tools and algorithms library. http://libdgtal.org

[2] Asano, T., Kawamura, Y., Klette, R., Obokata, K.: Minimum-length polygons in approximation sausages. In: C. Arcelli, L.P. Cordella, G.S. Baja (eds.) Visual Form 2001, *LNCS*, vol. 2059, pp. 103–112. Springer (2001)

[3] Baudrier, E., Mazo, L.: Results on multigrid convergence speed of non-local estimators for concave functions. In preparation

[4] Coeurjolly, D.: Algorithmique et géométrie discrète pour la caractérisation des courbes et des surfaces. Ph.D. thesis, Université Lyon 2 (2002)

[5] Coeurjolly, D., Klette, R.: A comparative evaluation of length estimators of digital curves. IEEE Trans. Pattern Anal. Mach. Intell. **26**(2), 252–257 (2004)

[6] Daurat, A., Tajine, M., Zouaoui, M.: Patterns in discretized parabolas and length estimation. In: S. Brlek, C. Reutenauer, X. Provençal (eds.) DGCI, *LNCS*, vol. 5810, pp. 373–384. Springer (2009)

[7] Daurat, A., Tajine, M., Zouaoui, M.: Les estimateurs semi-locaux de périmètre. Tech. rep., LSIIT CNRS, UMR 7005, Université de Strasbourg (2011). http://hal.inria.fr/hal-00576881

[8] De Vieilleville, F., Lachaud, J.O., Feschet, F.: Convex digital polygons, maximal digital straight segments and convergence of discrete geometric estimators. Journal of Mathematical Imaging and Vision **27**(2), 139–156 (2007)

[9] Debled-Rennesson, I., Reveillès, J.P.: A linear algorithm for segmentation of digital curves. Int J. Pattern Recognition and Artificial Intelligence **09**(04), 635–662 (1995)

[10] Dorst, L., Smeulders, A.W.: Discrete straight line segments: parameters, primitives and properties. Vision geometry: proceedings of an AMS special session held October 20-21, 1989 **119**, 45–62 (1991)

[11] Klette, R., , Žunić, J.: Multigrid convergence of calculated features in image analysis. J. Mathematical Imaging and Vision **13**(3), 173–191 (2000)

[12] Klette, R., Rosenfeld, A.: Digital Geometry. Morgan Kaufman (2004)

[13] Klette, R., Yip, B.: Evaluation of curve length measurements. International Conference on Pattern Recognition **1**, 1610 (2000)

[14] Kovalevsky, V.: New definition and fast recognition of digital straight segments and arcs. In: Proc. 10th Int. Conf. on Pattern Recognition, vol. ii, pp. 31–34 (1990)

[15] Kulkarni, S.R., Mitter, S.K., Richardson, T.J., Tsitsiklis, J.N.: Local versus nonlocal computation of length of digitized curves. IEEE Trans. Pattern Anal. Mach. Intell. **16**(7), 711–718 (1994)

[16] Lachaud, J.O.: 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)

[17] Lachaud, J.O., Thibert, B.: Properties of Gauss digitized sets and digital surface integration. Tech. rep. (2014). URL https://hal.archives-ouvertes.fr/hal-01070289

[18] Mazo, L., Baudrier, E.: About multigrid convergence of some length estimators. In: E.B. et al. (ed.) DGCI, *LNCS*, vol. 8668, pp. 214–225. Springer (2014)

[19] Rosenfeld, A.: Digital straight line segments. Computers, IEEE Transactions on **C-23**(12), 1264–1269 (1974)

[20] Sloboda, F., Zatko, B., Stoer, J.: On approximation of planar one-dimensional continua. Advances in Digital and Computational Geometry pp. 113–160 (1998)

[21] Stein, W., et al.: Sage Mathematics Software (Version 6.7). The Sage Development Team (2015). http://www.sagemath.org

[22] Tajine, M., Daurat, A.: On local definitions of length of digital curves. In: I. Nyström, G.S. di Baja, S. Svensson (eds.) DGCI, *LNCS*, vol. 2886, pp. 114–123. Springer (2003)

[23] Tajine, M., Daurat, A.: Patterns for multigrid equidistributed functions: Application to general parabolas and length estimation. Theoretical Computer Science **412**(36), 4824 – 4840 (2011)