MULTISCALE MODEL. SIMUL.
Vol. 4, No. 3, pp. 992–1039

BANDELET IMAGE APPROXIMATION AND COMPRESSION*

E. LE PENNEC† AND S. MALLAT‡

Abstract. Finding efficient geometric representations of images is a central issue to improving image compression and noise removal algorithms. We introduce *bandelet* orthogonal bases and frames that are adapted to the geometric regularity of an image. Images are approximated by finding a best bandelet basis or frame that produces a sparse representation. For functions that are uniformly regular outside a set of edge curves that are geometrically regular, the main theorem proves that bandelet approximations satisfy an optimal asymptotic error decay rate. A bandelet image compression scheme is derived. For computational applications, a fast discrete bandelet transform algorithm is introduced, with a fast best basis search which preserves asymptotic approximation and coding error decay rates.

Key words. wavelets, bandelets, geometric representation, nonlinear approximation

AMS subject classifications. 41A25, 42C40, 65T60

DOI. 10.1137/040619454

1. Introduction.

When a function defined over $[0,1]^2$ has singularities that belong to regular curves, one may take advantage of this geometrical regularity to optimize its approximation from M parameters. Most current procedures such as M-term separable wavelet approximations are locally isotropic and thus cannot take advantage of such geometric regularity. This paper introduces a new class of bases, with elongated multiscale *bandelet* vectors that follow the geometry, to optimize the approximation.

For functions f that are Hölderian of order α over $[0,1]^2$, M-term separable wavelet approximations f_M satisfy $\|f - f_M\|^2 \leq C\, M^{-\alpha}$, where $\|\cdot\|$ stands for the L^2 norm. The decay exponent α is optimal in the sense that no other approximation scheme can improve it for all such functions. If f is Hölderian of order $\alpha \geq 1$ over $[0,1]^2 - \{\mathcal{C}_\gamma\}_{1\leq\gamma\leq G}$, where the $\mathcal{C}_\gamma$ are finite length curves along which f is discontinuous, then its M-term wavelet approximation satisfies only $\|f - f_M\|^2 \leq C\, M^{-1}$. The existence of discontinuities drives entirely the decay of the approximation error. If the $\mathcal{C}_\gamma$ are regular curves, several approaches [1, 6, 8, 18] have already been proposed to improve the decay of this wavelet approximation error for $1 \leq \alpha \leq 2$. When $\alpha \geq 1$ is unknown, the issue addressed by this paper is to find an approximation scheme that is asymptotically as efficient as if f was Hölderian of order α over its whole support and to derive an image compression scheme. This is particularly important to approximate and compress images, where the contours of objects create edge transitions along piecewise regular curves.

Section 2 reviews nonlinear approximation results for piecewise regular images including edges. In the neighborhood of an edge, the image gray levels vary regularly in directions parallel to the edge, but they have sharp transitions across the

*Received by the editors November 23, 2004; accepted for publication (in revised form) April 7, 2005; published electronically September 8, 2005.
http://www.siam.org/journals/mms/4-3/61945.html

†Centre de Mathématiques Appliquées, Ecole Polytechnique, 91128 Palaiseau Cedex, France. Current address: Laboratoire de Probabilités et Modèles Aléatoires, Université Paris 7, 75251 Paris Cedex 05, France (lepennec@math.jussieu.fr).

‡Centre de Mathématiques Appliquées, Ecole Polytechnique, 91128 Palaiseau Cedex, France (stephane.mallat@polytechnique.fr).

edge. This anisotropic regularity is specified by a *geometric flow* that is a vector field that indicates local direction of regularity. Section 3 constructs *bandelets*, which are anisotropic wavelets that are warped along this geometric flow, and *bandelet orthonormal bases* in bands around edges. We study the approximation in bandelet bases of functions including edges over such bands. *Bandelet frame* of $\mathbf{L^2}[0,1]^2$ are defined in section 4.1 as a union of bandelet bases in different bands. A dictionary of bandelet frames is constructed in section 4.2 with dyadic square segmentations of $[0,1]^2$ and parameterized geometry flows. The main theorem of section 4.2 proves that a *best bandelet frame* obtained by minimizing an appropriate Lagrangian cost function in the bandelet dictionary yields asymptotically optimal approximations of piecewise regular functions. If f is Hölderian of order $\alpha \geq 1$ over $[0,1]^2 - \{\mathcal{C}_\gamma\}_{1\leq\gamma\leq G}$, where the $\mathcal{C}_\gamma$ are Hölderian of order α, then the main theorem proves that an approximation from M parameters in a best bandelet frame satisfies $\|f - f_M\|^2 \leq C\, M^{-\alpha}$.

To compress images in bits, an image transform code is defined in section 5. It is proved that a best bandelet frame yields a distortion rate that nearly reaches the Kolmogorov asymptotic lower bound up to a logarithmic factor. For numerical implementations over digital images, section 6.1 discretizes bandelet bases and frames. Section 6.2 describes a fast algorithm that finds a best discrete bandelet frame with an approximation error that decays like $M^{-\alpha}$ up to a logarithmic factor.

2. Geometric image model. We begin by establishing a mathematical model for geometrically regular images using the notion of edge. This model incorporates the fact that the image intensity is not necessarily singular at edge locations, which is why edge detection is an ill-posed problem. We then review existing constructive procedures to approximate such geometrically regular functions.

Functions that are regular everywhere outside a set of regular edge curves define a first simple model of geometrically regular functions. Let $\mathbf{C}^\alpha(\Lambda)$ be the space of Hölderian functions of order α over $\Lambda \subset \mathbb{R}^n$ defined for $\alpha > 0$:

(2.1)
$$\mathbf{C}^\alpha(\Lambda) = \left\{ \begin{array}{l} f,\, \mathbb{R}^n \to \mathbb{R} \;:\; \forall |\beta| = \lfloor\!\lfloor \alpha \rfloor\!\rfloor,\, \frac{\partial^{|\beta|}}{\partial x_1^{\beta_1} \cdots \partial x_n^{\beta_n}} f \text{ exists and satisfies} \\ \sup_{(x,y)\in\Lambda^2} \left| \frac{\partial^{|\beta|}}{\partial x_1^{\beta_1} \cdots \partial x_n^{\beta_n}} f(x) - \frac{\partial^{|\beta|}}{\partial x_1^{\beta_1} \cdots \partial x_n^{\beta_n}} f(y) \right| \times \|x-y\|^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor - \alpha} < \infty \end{array} \right\}$$

with

$$\lfloor\!\lfloor \alpha \rfloor\!\rfloor = \begin{cases} \lfloor \alpha \rfloor & \text{if } \alpha \notin \mathbb{N}, \\ \alpha - 1 & \text{if } \alpha \in \mathbb{N} \end{cases}$$

and $\lfloor \alpha \rfloor$ the integer just below α. For α integer, the space $\mathbf{C}^\alpha$ is slightly larger than the space of function having bounded derivatives up to order α. The norm $\|f\|_{\mathbf{C}^\alpha(\Lambda)}$ used through this paper is defined by

(2.2)
$$\|f\|_{\mathbf{C}^\alpha(\Lambda)} = \max \left(\begin{array}{l} \sup\limits_{x\in\Lambda} \max\limits_{|\beta|\leq \lfloor\!\lfloor \alpha \rfloor\!\rfloor} \dfrac{\partial^{|\beta|}}{\partial x_1^{\beta_1} \cdots \partial x_n^{\beta_n}} f(x), \\ \sup\limits_{(x,y)\in\Lambda^2} \max\limits_{|\beta| = \lfloor\!\lfloor \alpha \rfloor\!\rfloor} \left| \dfrac{\partial^{|\beta|}}{\partial x_1^{\beta_1} \cdots \partial x_n^{\beta_n}} f(x) - \dfrac{\partial^{|\beta|}}{\partial x_1^{\beta_1} \cdots \partial x_n^{\beta_n}} f(y) \right| \times \|x-y\|^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor - \alpha} \end{array} \right).$$

We say that an edge curve is Hölderian of order α if the coordinates in $\mathbb{R}^2$ of the points along this curve have a parameterization by arc length which is Hölderian of order α. An image model with geometrically regular edges is obtained by imposing that $f \in \mathbf{C}^\alpha(\Lambda)$ for $\Lambda = [0,1]^2 - \{\mathcal{C}_\gamma\}_{1\leq\gamma\leq G}$, where the $\mathcal{C}_\gamma$ are Hölderian of order α edge curves. For most images, this model is too simplistic because most often the image intensity has a sharp variation but is not singular across an edge. In particular, discontinuities of the image intensity created by occlusions in the visual scene are blurred by optical diffraction effects. These blurring effects along edges can be modeled through a convolution with an unknown kernel of compact support $h(x)$. This means that we can write $f(x) = \tilde{f} \star h(x)$, where $\tilde{f} \in \mathbf{C}^\alpha(\Omega)$ for $\Omega = [0,1]^2 - \{\mathcal{C}_\gamma\}_{1\leq\gamma\leq G}$.

When $f \neq \tilde{f}$, finding from f the exact locations of the edges $\mathcal{C}_\gamma$ is an ill-posed problem, especially since h is unknown. The difficulty in locating blurred edges is well known in image processing [3]. The goal of this paper is to find an approximation f_M from M parameters which satisfies

$$\forall M > 0, \qquad \|f - f_M\|^2 \leq C\, M^{-\alpha}, \tag{2.3}$$

with a constant C that does not depend upon the blurring kernel h.

A wavelet approximation decomposes f in an orthonormal wavelet basis and reconstructs f_M from a partial sum of M wavelets corresponding to the largest amplitude coefficients. Over a class of functions f whose total variation are uniformly bounded then one can prove [5] that

$$\|f - f_M\|^2 \leq C\, M^{-1}. \tag{2.4}$$

This result applies to functions having discontinuities along regular edges. However, wavelet bases are unable to take advantage of existing geometric regularity in order to improve the asymptotic error decay M^{-1}.

Many beautiful ideas have already been studied to find approximation schemes that take into account geometric image regularity. A surprising result by Candès and Donoho [1] shows that it is not necessary to estimate the image geometry to obtain efficient approximations. A curvelet frame is composed of multiscale elongated and rotated wavelet-type functions. Candès and Donoho [2] prove that if $\alpha = 2$, then an approximation f_M with M curvelets satisfies

$$\|f - f_M\|^2 \leq C\, M^{-2}\, (\log_2 M)^3. \tag{2.5}$$

Up to the $(\log_2 M)^3$ factor, this approximation result is thus asymptotically optimal for $\alpha = 2$. However, it is not adaptive in α in the sense that the optimal decay rate $M^{-\alpha}$ is not obtained when $\alpha < 2$ or $\alpha > 2$ [8].

Instead of choosing a priori a basis to approximate f, one can rather adapt the approximation scheme to the image geometry. For example, one can construct an approximation f_M which is piecewise linear over an optimized triangulation including M triangles and satisfies $\|f - f_M\|^2 \leq C\, M^{-2}$. This requires adapting the triangulation to the edge geometry and to the blurring scale [12, 7]. However, there is no known polynomial complexity algorithm which computes such an approximation f_M with an error that always decays like M^{-2}. Incorporating an unknown blurring kernel in the geometric model makes the problem much more difficult. Indeed, smooth edges are more difficult to detect than sharp singularities, and the triangulation must be adapted to the size of the blurring to approximate precisely the image transitions along the edges.

Most adaptive approximation schemes that have been developed so far can efficiently approximate geometrically regular images only if the edges are singularities, meaning that $h = \delta$. In particular, many image processing algorithms have been developed to construct such approximations by detecting edges and constructing regular approximations between the edges where the image is uniformly regular [17]. To obtain fast polynomial time algorithms, Donoho introduced multiscale strategies to approximate the image geometry. Dictionaries constructed with wedgelets are used to compute approximations of functions that are $\mathbf{C}^2$ away from $\mathbf{C}^2$ edges [9]. The approximation bound $\|f - f_M\|^2 \leq C\, M^{-\alpha}$ holds only for $1 \leq \alpha \leq 2$ and if there is no blurring ($h = \delta$). Wakin et al. [18] propose a compression scheme that mixes the wedgelets and the wavelets to obtain better practical approximation results while following a similar geometry optimization scheme. A different strategy developed by Cohen and Matei [6] uses a nonlinear subdivision scheme to construct an approximation f_M that can reach a similar error decay bound if the image has sharp discontinuities along $\mathbf{C}^2$ edges.

The goal of this paper is to find a approximation f_M from M parameters that satisfies $\|f - f_M\|^2 \leq C\, M^{-\alpha}$ for any $\alpha \geq 1$ and any blurring kernel h.

3. Approximation in orthonormal bandelet bases.

3.1. Geometric flow and bandelet bases. An image having geometrically regular edges, as in the model of section 2, has sharp transitions when moving across edges but has regular variations when moving parallel to these edges. This displacement parallel to edges can be characterized by a *geometric flow*, which is a field of parallel vectors that give the local direction in which f has regular variations. Bandelet orthonormal bases are constructed by warping anisotropic wavelets bases with this geometric flow.

A geometric flow is a vector field $\vec{\tau}(x_1, x_2)$ which gives directions in which f has regular variations in the neighborhood of each (x_1, x_2). In the neighborhood of an edge, the flow is typically parallel to the tangents of the edge curve. To construct an orthogonal basis with a geometric flow, we shall impose that the flow is locally either parallel in the vertical direction, and hence constant in this direction, or parallel in the horizontal direction. To simplify the explanations, we shall first consider horizontal or vertical *horizon models* [9] (or *boundary fragments* [13]), which are functions f that include a single edge $\mathcal{C}$ whose tangents have an angle with the horizontal or vertical direction that remains smaller than $\pi/3$, so that $\mathcal{C}$ can be parameterized horizontally or vertically by a function g.

Suppose that f is a horizontal horizon model. We define a vertically parallel flow whose angle with the horizontal direction is smaller than $\pi/3$. Such a flow can be written

$$\vec{\tau}(x_1, x_2) = \vec{\tau}(x_1) = (1, g'(x_1)) \quad \text{with} \quad |g'(x_1)| \leq 2. \tag{3.1}$$

A *flow line* is an integral curve of the flow whose tangent at (x_1, x_2) is collinear to $\vec{\tau}(x_1, x_2)$. Let $g(x)$ be a primitive of $g'(x)$ defined by $g(x) = \int_0^x g'(x)\,\mathrm{d}x$ that we shall call *flow integral*. Flow lines are sets of points $(x_1, x_2) \in \Omega$ which satisfy $x_2 = g(x_1) + \mathrm{cst}$.

A band B parallel to this flow is defined by

$$B = \Big\{(x_1, x_2) \ : \ x_1 \in [a_1, b_1],\, x_2 \in [g(x_1) + a_2, g(x_1) + b_2]\Big\}. \tag{3.2}$$

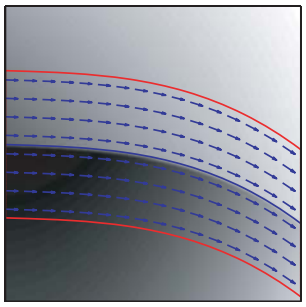

FIG. 3.1. *Horizontal horizon model with a flow induced by the edge and the corresponding band.*

The band height a_2 and b_2 are chosen so that a neighborhood of $\mathcal{C}$ is included in B, as illustrated in Figure 3.1.

If the flow directions are sufficiently parallel to the edge directions, then $f(x)$ has regular variations along each flow line $(x_1, g(x_1) + \text{cst})$. As a consequence, Figure 3.2 shows that the warped image

$$Wf(x_1, x_2) = f(x_1, x_2 + g(x_1)) \tag{3.3}$$

has regular variations along the horizontal lines (x_2 constant) in the rectangle obtained by warping the band B:

$$WB = \{(x_1, x_2) \ : \ (x_1, x_2 + g(x_1)) \in B\} \ = \ \{(x_1, x_2) \ : \ x_1 \in [a_1, b_1],\ x_2 \in [a_2, b_2]\}. \tag{3.4}$$

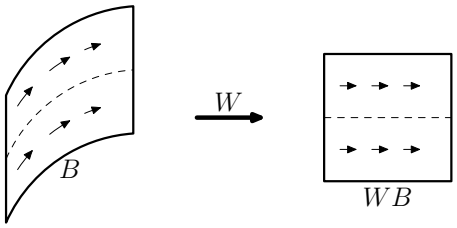

FIG. 3.2. *A band B and its warped band WB.*

If $\Psi(x_1, x_2)$ is a function having vanishing moments along x_1 for x_2 fixed, since $Wf(x_1, x_2)$ is regular along x_1, the resulting inner product $\langle Wf, \Psi \rangle$ will be small. Note that

$$\langle Wf, \Psi \rangle = \langle f, W^* \Psi \rangle, \tag{3.5}$$

where W^* is the adjoint of the operator W defined in (3.3). This suggests decomposing f over the warped image by W^* of a basis having vanishing moments along x_1. Observe then that

$$W^* f(x_1, x_2) = W^{-1} f(x_1, x_2) = f(x_1, x_2 - g(x_1)). \tag{3.6}$$

Since W is an orthogonal operator, an orthonormal family warped with $W^* = W^{-1}$ remains orthonormal. By inverse warping, an orthogonal wavelet basis of the rectangle WB thus yields an orthogonal basis over the band B with basis functions having vanishing moments along the flow lines.

A separable wavelet basis is defined from one-dimensional wavelet $\psi(t)$ and a scaling function $\phi(t)$ that are here chosen compactly supported, which are dilated and translated [15, 16]:

$$\psi_{j,m}(t) = \frac{1}{\sqrt{2^j}}\,\psi\Big(\frac{t-2^j m}{2^j}\Big) \quad \text{and} \quad \phi_{j,m}(t) = \frac{1}{\sqrt{2^j}}\,\phi\Big(\frac{t-2^j m}{2^j}\Big). \tag{3.7}$$

Following [15], the index j goes to $-\infty$ when the wavelet scale 2^j decreases. In this paper, j is thus typically a negative integer. The family of separable wavelets

$$\left\{ \begin{array}{lll} \phi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2) & , & \psi_{j,m_1}(x_1)\,\phi_{j,m_2}(x_2) \\ & , & \psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2) \end{array} \right\}_{(j,m_1,m_2)\in I_{WB}} \tag{3.8}$$

defines an orthonormal basis over the rectangle WB if one modifies appropriately the wavelets whose support intersects the boundary of WB [4]. The index set I_{WB} depends upon the width and length of the rectangle WB. The wavelet construction is slightly modified to cope with a real anisotropic rectangle: it is started with scaling functions of the size of the order of the largest dimension instead of the smallest dimension. This basis could thus already include some anisotropic functions. This ensures that a polynomial on a rectangle could always be reproduced with a fixed number of coefficients.

Since W is orthogonal, applying its inverse to each of the wavelets (3.8) yields an orthonormal basis of $\mathbf{L^2}(B)$ that is called a *warped wavelet basis*:

$$\left\{ \begin{array}{lll} \phi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2-g(x_1)) & , & \psi_{j,m_1}(x_1)\,\phi_{j,m_2}(x_2-g(x_1)) \\ & , & \psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2-g(x_1)) \end{array} \right\}_{(j,m_1,m_2)\in \mathbf{I}_{WB}}. \tag{3.9}$$

Warped wavelets are separable along the x_1 variable and along the $x_2' = x_2 - g(x_1)$ variable, so that for a given x_2' one follows the geometric flow lines within the band B.

The wavelet $\psi(t)$ has $p \geq \alpha$ vanishing moments, but $\phi(t)$ has no vanishing moments. As a consequence, the separable wavelets $\psi_{j,m_1}(x_1)\,\phi_{j,m_2}(x_2)$ and $\psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2)$ have vanishing moments along x_1 but not $\phi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2)$. However, $\{\phi_{j,m_1}(x_1)\}_{m_1}$ is an orthonormal basis of a multiresolution space which also admits an orthonormal basis of wavelets $\{\psi_{l,m_1}(x_1)\}_{l>j,m_1}$ that have vanishing moments, besides a constant number of scaling functions. This suggests replacing the orthogonal family $\{\phi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2)\}_{j,m_1,m_2}$ by the family $\{\psi_{l,m_1}(x_1)\,\psi_{j,m_2}(x_2)\}_{j,l>j,m_1,m_2}$ which generates the same space. This is called a *bandeletization*.

After applying the inverse warping W^{-1} the resulting functions $\psi_{l,m_1}(x_1)\,\psi_{j,m_2}(x_2-g(x_1))$ are called bandelets because their support is parallel to the flow lines and is more elongated ($2^l > 2^j$) in the direction of the geometric flow. Inserting these bandelets in the warped wavelet basis (3.9) yields a bandelet orthonormal basis of $\mathbf{L^2}(B)$:

$$\left\{ \begin{array}{lll} \psi_{l,m_1}(x_1)\,\psi_{j,m_2}(x_2-g(x_1)) & , & \psi_{j,m_1}(x_1)\,\phi_{j,m_2}(x_2-g(x_1)) \\ & , & \psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2-g(x_1)) \end{array} \right\}_{j,l>j,m_1,m_2}. \tag{3.10}$$

Suppose now that f is a vertical horizon model, with an edge along a curve $\mathcal{C}$ whose tangents have an angle smaller than $\pi/3$ with the vertical direction. We then

define a geometric flow $\vec{\tau}(x_1, x_2)$ that is parallel in the horizontal direction and which has an angle smaller than $\pi/3$ with the vertical direction:

$$\vec{\tau}(x_1, x_2) = \vec{\tau}(x_2) = (g'(x_2), 1) \quad \text{with} \quad |g'(x_2)| \leq 2. \tag{3.11}$$

Flow lines are points (x_1, x_2) with $x_1 = g(x_2) + \text{cst}$.

We define a band which is parallel to the geometric flow:

$$B = \Big\{ (x_1, x_2) \ : \ x_1 \in [g(x_2) + a_1, g(x_2) + b_1] \ , \ x_2 \in [a_2, b_2] \Big\}. \tag{3.12}$$

The width $\ell_1 = b_1 - a_1$ of the band is adjusted so that a neighborhood of $\mathcal{C}$ is included in B. Similarly to the previous case, the warped wavelet basis of $\mathbf{L^2}(B)$ is then defined by

$$\left\{ \begin{array}{lll} \phi_{j,m_1}(x_1 - g(x_2))\, \psi_{j,m_2}(x_2) & , & \psi_{j,m_1}(x_1 - g(x_2))\, \phi_{j,m_2}(x_2) \\ & , & \psi_{j,m_1}(x_1 - g(x_2))\, \psi_{j,m_2}(x_2) \end{array} \right\}_{(j,m_1,m_2) \in \mathbf{I}_{WB}}. \tag{3.13}$$

The bandeletization replaces each family of scaling functions $\{\phi_{j,m_2}(x_2)\}_{m_2}$ by a family of orthonormal wavelets that generates the same approximation space. The resulting bandelet orthonormal basis of $\mathbf{L^2}(B)$ is

$$\left\{ \begin{array}{lll} \phi_{j,m_1}(x_1 - g(x_2))\, \psi_{j,m_2}(x_2) & , & \psi_{j,m_1}(x_1 - g(x_2))\, \psi_{l,m_2}(x_2) \\ & , & \psi_{j,m_1}(x_1 - g(x_2))\, \psi_{j,m_2}(x_2) \end{array} \right\}_{j,l>j,m_1,m_2}. \tag{3.14}$$

3.2. Bandelet approximation in a band. The bandelet approximation of geometrically regular horizon models is studied in bands around their edges. The following definition introduces such geometrical regular functions according to the model of section 2.

DEFINITION 3.1. *A function f is a $\mathbf{C}^\alpha$ horizon model over $[0,1]^2$ if*

- *$f = \tilde{f}$ or $f = \tilde{f} \star h$, with $\tilde{f} \in \mathbf{C}^\alpha(\Lambda)$ for $\Lambda = [0,1]^2 - \{\mathcal{C}\}$,*
- *the blurring kernel h has a support included in $[-s, s]^2$ and is $\mathbf{C}^\alpha$ with $\|h\|_{\mathbf{C}^\alpha} \leq s^{-(2+\alpha)}$,*
- *the edge curve $\mathcal{C}$ is Hölderian of order α and its tangents have an angle with the horizontal or vertical direction that remains smaller than $\pi/3$, with a distance larger than s from the horizontal (respectively, vertical) boundary.*

Observe that the kernel h can be rewritten as a dilation by s: $h(x) = s^{-2}\, h_1(s^{-1}x)$, where h_1 has a support in $[-1, 1]$ and $\|h_1\|_1 = \|h\|_1$. Normalizing the amplitude of h by setting $\|h\|_{\mathbf{C}^\alpha} = s^{-(2+\alpha)}$ is equivalent to setting $\|h_1\|_{\mathbf{C}^\alpha} = 1$. The convolution with h diffuses the edge $\mathcal{C}$ over a tube $\mathcal{C}_s$ defined as the set of points within a distance s of $\mathcal{C}$. Outside this tube f is uniformly $\mathbf{C}^\alpha$, but within this tube its regularity depends upon s, which may be an arbitrarily small variable. We study the bandelet approximation of f within a band B that includes the tube $\mathcal{C}_s$.

A bandelet basis is constructed from a geometric flow. To optimize the approximation of f with M parameters in a bandelet basis, it is necessary to specify this geometric flow with as few parameters as possible. A vertically parallel flow is specified in a horizontal band B of length $\ell_1 = b_1 - a_1$ by parameterizing the flow integral $g(x_1)$ over a family of orthogonal scaling functions $\{\theta_{k,m}(t)\}_{1 \leq m \leq \ell_1 2^{-k}}$ at a scale 2^k.

The scaling functions $\{\theta_{k,m}(t)\}_{1\leq m\leq \ell_1 2^{-k}}$ whose support does not intersect the border of $[a_1, b_1]$ can be written $\theta_{k,m}(t) = \theta(2^{-k}t - m)$, and

$$\forall t \in [a_1, b_1] \quad , \quad g(t) = \sum_{m=1}^{\ell_1 2^{-k}} \alpha_m \, \theta_{k,m}(t) \quad \text{with} \quad \alpha_m = \langle g, \theta_{k,m} \rangle \, \|\theta_{k,m}\|^{-2}. \tag{3.15}$$

We suppose that the space $\mathbf{V}_k$ generated by the orthogonal family $\{\theta_{k,n}(t)\}_{1\leq n\leq \ell_1 2^{-k}}$ includes polynomials of degree p over $[a_1, b_1]$ and that $\theta(t)$ is compactly supported and p times differentiable. The decomposition (3.15) defines a parameterized flow that depends upon the scale 2^k and the $(b_1 - a_1)2^{-k}$ coefficients α_n. Since $|g'(t)| \leq 2$ and g is defined up to a constant, one can set $g(a_1) = 0$ so that $|g(t)| \leq 2\,\ell_1$. As a result, one can verify that there exists $C_\theta > 0$ that depend only upon θ such that $|\alpha_m| \leq C_\theta \, (b_1 - a_1)$.

There are many possible approaches to computing a geometric flow for a horizon model. Section 4.2 describes an algorithm that computes the flow by minimizing the approximation error of the resulting bandelet representation. One can also use a simpler and computationally more efficient approach that minimizes a Sobolev norm in the direction of the flow [14]. If the horizon model is horizontal, the edge $\mathcal{C}$ can be parameterized horizontally by $(x_1, c(x_1))$. For each x_1, because of the blurring effect, the position of the edge can be estimated only up to an error bounded by the size s of the blurring kernel. The resulting estimate $\tilde{g}(x_1)$ of the edge position $c(x_1)$ satisfies $\|\tilde{g} - c\|_\infty \leq C_d \, s$, where C_d depends upon precision of the algorithm that is used.

Let $P_{\mathbf{V}_k}$ be the orthogonal projection on the approximation space $\mathbf{V}_k$. To represent the geometry with few coefficients, a regularized edge is considered:

$$g(t) = P_{\mathbf{V}_k} \tilde{g}(t) = \sum_{m=1}^{\ell_1 2^{-k}} \alpha_m \, \theta_{k,m}(t) \quad \text{with} \quad \alpha_m = \langle \tilde{g}, \theta_{k,m} \rangle \, \|\theta_{k,m}\|^{-2}. \tag{3.16}$$

We shall prove that if the scale 2^k is small enough, then the distance between $\mathcal{C}$ and the resulting regularized edge remains of the order of s.

Let us now construct a bandelet orthonormal basis over a band B parallel to the geometric flow defined by the flow integral g, with a one-dimensional wavelet ψ which has $p \geq \alpha$ vanishing moments. To simplify notation, in the following we write $\mathcal{B} = \{b_m\}_m$, the orthonormal bandelet basis defined in (3.10). An approximation of f in $\mathcal{B}$ is obtained by keeping only the bandelet coefficients above a threshold T:

$$f_M = \sum_{\substack{m \\ |\langle f, b_m \rangle| \geq T}} \langle f, b_m \rangle \, b_m. \tag{3.17}$$

The total number of parameters is $M = M_G + M_B$, where M_G is the number of parameters α_n in (3.16) that define the geometric flow in B, and M_B is the number of bandelet coefficients above T. The following theorem computes the resulting approximation error.

THEOREM 3.2. *Let f be a $\mathbf{C}^\alpha$ horizon model with an edge parameterized by c and $1 \leq \alpha < p$. Let $\tilde{g}$ be an edge estimation such that $\|\tilde{g} - c\|_\infty \leq C_d \, s$. There exists C such that for any threshold T, the approximation error in a bandelet basis defined by the flow integral $g = P_{\mathbf{V}_k}(\tilde{g})$ with $2^k = \max(\|c\|_{\mathbf{C}^\alpha}^{-1/\alpha}, 1) \max(s, T^{2\alpha/(\alpha+1)})^{1/\alpha}$ satisfies*

$$\|f - f_M\|^2 \leq C \, C_f^2 \, \ell_1^{\alpha+1} \, M^{-\alpha} \tag{3.18}$$

with

$$C_f = \max(\|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)}\, \max(\|c\|^\alpha_{\mathbf{C}^\alpha}, 1)\, \max(\|c\|^\alpha_{\mathbf{C}^\alpha}, C_d, 1), \min(\|c\|^{(\alpha+1)/(2\alpha)}_{\mathbf{C}^\alpha}, 1), \|\tilde{f}\|^{\alpha+1}_{\mathbf{C}^1(\Lambda)}).$$

The constant C_f is defined as the maximum of three quantities:

- $\|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)} \max(\|c\|^\alpha_{\mathbf{C}^\alpha}, 1) \max(\|c\|^\alpha_{\mathbf{C}^\alpha}, C_d, 1)$ controls the regularity of Wf;
- $\min(\|c\|^{(\alpha+1)/(2\alpha)}_{\mathbf{C}^\alpha}, 1)$ appears in the geometry approximation;
- $\|\tilde{f}\|^{\alpha+1}_{\mathbf{C}^1(\Lambda)}$ controls the error $f - \tilde{f}$ away from the singularities.

This does not correspond to any fine optimization in the relationship between T, s, and the regularity of f and $\mathcal{C}_\gamma$ but on the natural quantities that arise in the proof.

Proof of Theorem 3.2. We first prove the following proposition that considers the case $T \leq s^{(\alpha+1)/(2\alpha)}$ where the geometric approximation scale is $2^k = \max(\|c\|^{-1/\alpha}_{\mathbf{C}^\alpha}, 1)\, s^{1/\alpha}$.

PROPOSITION 3.3. *Under the hypotheses of Theorem* 3.2*, there exists C that depends only upon the edge geometry such that for any threshold $T \leq s^{(\alpha+1)/(2\alpha)}$ the bandelet approximation error satisfies*

$$\|f - f_M\|^2 \leq C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{2\alpha/(\alpha+1)} \tag{3.19}$$

and

$$M \leq \max(C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{-2/(\alpha+1)}, C\, |\log_2 T|). \tag{3.20}$$

Proof of Proposition 3.3. Let g be the projection of $\tilde{g}$ on the space $\mathbf{V}_k$ with $2^k = \max(\|c\|^{-1/\alpha}_{\mathbf{C}^\alpha}, 1)\, s^{1/\alpha}$, B the associated band, and $\mathcal{B}$ the associated bandelet basis.

By construction, the number of geometric parameters M_G satisfies

$$M_G \leq \max((b_1 - a_1)\, \max(\|c\|^{-1/\alpha}, 1)^{-1}\, s^{-1/\alpha}, C) \tag{3.21}$$

and, as $s \geq T^{2\alpha/(\alpha+1)}$,

$$M_G \leq \max(C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{-2/(\alpha+1)}, C\, |\log_2 T|), \tag{3.22}$$

and we should now focus on the number M_B of bandelet coefficients above T.

The bandelet basis $\mathcal{B}$ of B is separated in two families: $\mathcal{B}_1$ the bandelets whose support does not intersect $\mathcal{C}_s$ and $\mathcal{B}_2$ the bandelets whose support does intersect $\mathcal{C}_s$.

Using the orthogonality of the basis $\mathcal{B}$, we verify that

$$\|f - f_M\|^2 = \sum_{\substack{m \\ |\langle f, b_m\rangle| < T}} |\langle f, b_m\rangle|^2, \tag{3.23}$$

$$\|f - f_M\|^2 = \sum_{\substack{b_m \in \mathcal{B}_1 \\ |\langle f, b_m\rangle| < T}} |\langle f, b_m\rangle|^2 + \sum_{\substack{b_m \in \mathcal{B}_2 \\ |\langle f, b_m\rangle| < T}} |\langle f, b_m\rangle|^2, \tag{3.24}$$

and $M_B = M_{B,1} + M_{B,2}$, where $M_{B,1}$ and $M_{B,2}$ are, respectively, the number of bandelet coefficients above the threshold T in $\mathcal{B}_1$ and $\mathcal{B}_2$.

The proof is then divided into two lemmas. Lemma 3.4 takes care of the outer bandelets, $\mathcal{B}_1$, that do not intersect the singularities and relies on the regularity of f.

LEMMA 3.4. *Under the hypotheses of Theorem* 3.2*, there exists a constant* C *such that*

$$\sum_{\substack{b_m \in \mathcal{B}_1 \\ |\langle f, b_m \rangle| < T}} |\langle f, b_m \rangle|^2 \leq C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{2\alpha/(\alpha+1)} \tag{3.25}$$

and

$$M_{B,1} \leq \max(C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{-2/(\alpha+1)}, C). \tag{3.26}$$

Lemma 3.5 considers the inner bandelets of $\mathcal{B}_2$ that capture the sharp transitions of f.

LEMMA 3.5. *Under the hypotheses of Theorem* 3.2*, there exists a constant* C *such that*

$$\sum_{\substack{b_m \in \mathcal{B}_2 \\ |\langle f, b_m \rangle| < T}} |\langle f, b_m \rangle|^2 \leq C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{2\alpha/(\alpha+1)} \tag{3.27}$$

and

$$M_{B,2} \leq \max(C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{-2/(\alpha+1)}, C\, |\log_2 T|). \tag{3.28}$$

Combining (3.22), (3.25), (3.26), (3.27), and (3.28) allows us to conclude.

To prove Lemmas 3.4 and 3.5, we go back to the definition of the bandelets. With a slight abuse of notation, we note $b_{l,j,m}(x_1, x_2)$ are the bandelets defined by $\psi_{l,m_1}(x_1)\, \psi_{j,m_2}(x_2 - g(x_1))$ if $j < l$ and by either $\psi_{j,m_1}(x_1)\, \psi_{j,m_2}(x_2 - g(x_1))$ or $\psi_{j,m_1}(x_1)\, \phi_{j,m_2}(x_2 - g(x_1))$ if $j = l$. In the last case, both $\psi_{j,m_1}(x_1)\, \psi_{j,m_2}(x_2 - g(x_1))$ and $\psi_{j,m_1}(x_1)\, \phi_{j,m_2}(x_2 - g(x_1))$ have vanishing moments along x_1 and, as this is the only direction in which the moments are used, could be handled in the exact same way. From now on, we suppose that $b_{l,j,m}(x_1, x_2) = \psi_{l,m_1}(x_1)\, \psi_{j,m_2}(x_2 - g(x_1))$.

As seen in (3.5) and (3.10),

$$\begin{aligned} \langle f, b_{l,j,m} \rangle &= \langle f(x_1, x_2 + g(x_1)), \psi_{l,m_1}(x_1)\, \psi_{j,m_2}(x_2) \rangle \\ &= \langle Wf(x_1, x_2), \psi_{l,m_1}(x_1)\, \psi_{j,m_2}(x_2) \rangle. \end{aligned} \tag{3.29}$$

The regularity of Wf is the key to controlling the bandelet coefficient.

If f is $\mathbf{C}^k$ in the neighborhood of a point $(x_1, x_2 + g(x_1))$, a straightforward calculation shows that along the x_2-axis

$$\left| \frac{\partial^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor} Wf}{\partial x_2^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}}(x_1, x_2') - \frac{\partial^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor} Wf}{\partial x_2^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}}(x_1, x_2) \right| \leq \|f\|_{\mathbf{C}^\alpha} |x_2' - x_2|^{\alpha - \lfloor\!\lfloor \alpha \rfloor\!\rfloor}.$$

Along the x_1-axis, a control on the regularity of g is required, and after some calculation one derives that if g is $\mathbf{C}^\alpha$, then Wf is also $\mathbf{C}^\alpha$ and that

$$\left| \frac{\partial^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor} Wf}{\partial x_1^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}}(x_1', x_2) - \frac{\partial^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor} Wf}{\partial x_1^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}}(x_1, x_2) \right| \leq \|f\|_{\mathbf{C}^\alpha} \max(\|g\|, \|g\|^\alpha) |x_1' - x_1|^{\alpha - \lfloor\!\lfloor \alpha \rfloor\!\rfloor}.$$

The definition of g implies that the difference between g and the true curve c, $g - c$, is even $\mathbf{C}^\alpha$ as stated by the following lemma proved in Appendix A.3.

LEMMA 3.6. *There exists a constant C such that if c is $\mathbf{C}^\alpha$ over $[a_1, b_1]$ and $\tilde{g}$ satisfies $\|\tilde{g} - c\|_\infty \leq C_d\, s$, then $g = P_{\mathbf{V}_k}\tilde{g}$ with $2^k = \max(\|c\|_{\mathbf{C}^\alpha}^{-1/\alpha}, 1)\, s^{1/\alpha}$ is $\mathbf{C}^\alpha$ and satisfies*

$$\forall \beta \leq \lfloor\!\lfloor \alpha \rfloor\!\rfloor, \quad \|(g-c)^{(\beta)}\|_\infty \leq C\, \max(\|c\|_{\mathbf{C}^\alpha}, C_d, 1)\, s^{1-\beta/\alpha} \tag{3.30}$$

and

$$\forall x \quad |x - x_0| \leq K 2^k, \quad \|(g-c)^{(\lfloor\!\lfloor \alpha \rfloor\!\rfloor)}(x) - (g-c)^{(\lfloor\!\lfloor \alpha \rfloor\!\rfloor)}(x_0)\|_\infty \leq C\, \max(\|c\|_{\mathbf{C}^\alpha}, C_d, 1)\, |x - x_0|^{\alpha - \lfloor\!\lfloor \alpha \rfloor\!\rfloor}. \tag{3.31}$$

An immediate consequence of this lemma is that g is $\mathbf{C}^\alpha$ with a norm bounded by $C \max(\|c\|_{\mathbf{C}^\alpha}, C_d, 1) + \|c\|_{\mathbf{C}^\alpha}$.

The proof of Lemma 3.4 in Appendix A.1 combines this lemma with the regularity of f outside $\mathcal{C}_s$ to obtain a regularity of Wf along x_1 and thus obtain a decay of the bandelet coefficients with their vanishing moments in this direction.

In Appendix A.2 to obtain Lemma 3.5, some regularity along x_1 for Wf in the neighborhood of the smoothed singularity is required. This is given by the following lemma proved in Appendix A.4.

LEMMA 3.7. *Suppose f is a $\mathbf{C}^\alpha$ horizon model with $s > 0$. If g satisfies*

$$\forall \beta \leq \lfloor\!\lfloor \alpha \rfloor\!\rfloor, \quad \|(g-c)^{(\beta)}\|_\infty \leq C\, \max(\|c\|_{\mathbf{C}^\alpha}, C_d, 1)\, s^{1-\beta/\alpha} \tag{3.32}$$

and

$$\forall x \quad |x - x_0| \leq K s^{1/\alpha}, \quad \|(g-c)^{(\lfloor\!\lfloor \alpha \rfloor\!\rfloor)}(x) - (g-c)^{(\lfloor\!\lfloor \alpha \rfloor\!\rfloor)}(x_0)\|_\infty \leq C\, \max(\|c\|_{\mathbf{C}^\alpha}, C_d, 1)\, |x - x_0|^{\alpha - \lfloor\!\lfloor \alpha \rfloor\!\rfloor}, \tag{3.33}$$

then

$$\left| \frac{\partial^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}}{\partial x_1^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}} Wf(x_1', x_2) - \frac{\partial^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}}{\partial x_1^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}} Wf(x_1, x_2) \right| \leq C\, \|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)}\, \max(\|c\|_{\mathbf{C}^\alpha}^\alpha, 1)\, \max(\|c\|_{\mathbf{C}^\alpha}^\alpha, C_d^\alpha, 1)\, s^{-1}\, |x_1' - x_1|^{\alpha - \lfloor\!\lfloor \alpha \rfloor\!\rfloor}. \tag{3.34}$$

This bound and the condition on the threshold $T \leq s^{(\alpha+1)/(2\alpha)}$ are sufficient to prove Lemma 3.5. □

The hypotheses of Proposition 3.3 require a small threshold, $T \leq s^{\alpha+1/(2\alpha)}$, but a similar result also holds for larger T.

PROPOSITION 3.8. *Under the hypotheses of Theorem 3.2, there exists a constant C such that for all thresholds T the resulting bandelet approximation error satisfies*

$$\|f - f_M\|^2 \leq C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{2\alpha/(\alpha+1)} \tag{3.35}$$

and

$$M \leq \max(C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{-2/(\alpha+1)}, C\, |\log_2 T|). \tag{3.36}$$

Proof of Proposition 3.8. If $\|\tilde{g} - c\|_\infty \leq C_d\, s$ and $T^{2\alpha/(\alpha+1)} \leq s$, Proposition 3.3 applies. Otherwise, $T^{2\alpha/(\alpha+1)} > s$, and we prove that the condition on the edge estimation of Theorem 3.2, $\|\tilde{g} - c\|_\infty \leq C_d\, s$, can even be relaxed to $\|\tilde{g} - c\|_\infty \leq C_d \max(s, T^{2\alpha/(\alpha+1)})$.

Indeed, let $f_{\mathrm{mod}} = \tilde{f} \star h_{\mathrm{mod}}$, where

$$h_{\mathrm{mod}} = \left(\frac{s}{T^{2\alpha/(\alpha+1)}}\right)^2 h\left(\frac{s}{T^{2\alpha/(\alpha+1)}}\, x\right) \tag{3.37}$$

is a dilation of the smoothing kernel h supported in $[-T^{2\alpha/(\alpha+1)}, T^{2\alpha/(\alpha+1)}]^2$, which satisfies $\|h_{\mathrm{mod}}\|_{\mathbf{C}^\alpha} \leq (T^{2\alpha/(\alpha+1)})^{(2+\alpha)}$.

The following lemma proved in Appendix A.5 implies

$$\|f - f_{\mathrm{mod}}\|^2 \leq C\, \|f\|^2_{\mathcal{C}^1(\Lambda)}\, \ell_1\, T^{2\alpha/(\alpha+1)}. \tag{3.38}$$

LEMMA 3.9. *If f is a $\mathbf{C}^\alpha$ horizon model, there exists a constant C such that*

$$\left\| \|h\|_1 \tilde{f} - f \right\|^2 \leq C\, \|\tilde{f}\|^2_{\mathbf{C}^1(\Lambda)}\, \ell_1\, s. \tag{3.39}$$

Furthermore, Proposition 3.3 applies to f_{mod}, so

$$\|f_{\mathrm{mod}} - f_{\mathrm{mod},M_{\mathrm{mod}}}\|^2 \leq C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{2\alpha/(\alpha+1)} \tag{3.40}$$

and

$$M_{\mathrm{mod}} \leq \max(C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{-2/(\alpha+1)}, |\log_2 T|). \tag{3.41}$$

If we let J' be the set of bandelets such that $|\langle f_{\mathrm{mod}}, b_m\rangle| \geq T$, as

$$\|f - f_{\mathrm{mod}}\|^2 \leq C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{2\alpha/(\alpha+1)}, \tag{3.42}$$

one can verify that

$$\sum_{m \notin J'} |\langle f, b_m\rangle|^2 \leq 2\left(\sum_{m \notin J'} |\langle f_{\mathrm{mod}}, b_m\rangle|^2 + \sum_{m \notin J'} |\langle f - f_{\mathrm{mod}}, b_m\rangle|^2\right) \tag{3.43}$$

$$\leq 2(\|f_{\mathrm{mod}} - f_{\mathrm{mod},M_{\mathrm{mod}}}\|^2 + \|f - f_{\mathrm{mod}}\|^2) \tag{3.44}$$

$$\sum_{m \notin J'} |\langle f, b_m\rangle|^2 \leq C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{2\alpha/(\alpha+1)}, \tag{3.45}$$

while $\mathrm{Card}\,(J') \leq C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{-2/(\alpha+1)}$. To conclude, we rely on the optimality of the thresholding strategy in the following sense.

LEMMA 3.10. *Let $\mathcal{B} = \{u_m\}_{m \in J}$ be a family of functions and $J_T = \{m : |\langle f, u_m\rangle| > T\}$. If $J' \subset J$, then*

$$\sum_{m \notin J_T} |\langle f, u_m\rangle|^2 + T^2\, \mathrm{Card}\,(J_T) \leq \sum_{m \notin J'} |\langle f, u_m\rangle|^2 + T^2\, \mathrm{Card}\,(J'). \tag{3.46}$$

Proof of Lemma 3.10. For every $J' \subset J$,

$$\sum_{m \notin J'} |\langle f, u_m\rangle|^2 + T^2\, \mathrm{Card}\,(J') = \sum_{m \in J} ((1 - \mathbf{1}_{m \in J'})|\langle f, u_n\rangle|^2 + \mathbf{1}_{m \in J'} T^2). \tag{3.47}$$

This implies that each term of the sum is minimized if $m \in J'$ only when $|\langle f, u_m\rangle|^2 \geq T^2$; i.e., the sum is minimal for $J' = J_T$. □

Indeed, this implies that the bounds obtained for the subset J' remain valid for the subset J_T obtained with the thresholding. As proved in Lemma 3.10,

(3.48)
$$\sum_{m \notin J_T} |\langle f, u_m\rangle|^2 + T^2 \,\mathrm{Card}\,(J_T) \leq \sum_{m \notin J'} |\langle f, u_m\rangle|^2 + T^2 \,\mathrm{Card}\,(J')$$

(3.49)
$$\leq C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{2\alpha/(\alpha+1)} + T^2\, C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{-2/(\alpha+1)},$$

(3.50)
$$\sum_{m \notin J_T} |\langle f, u_m\rangle|^2 + T^2 \,\mathrm{Card}\,(J_T) \leq C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{2\alpha/(\alpha+1)},$$

which immediately implies (3.35) and (3.36). □

This last lemma completes the proof of Theorem 3.2. □

Theorem 3.2 proves that for a horizon model the bandelet approximation error has an optimal asymptotic decay if the flow integral g is computed at an appropriate scale 2^k that depends upon f, from an estimation $\tilde{g}$ that is sufficiently precise. Section 4.2 explains how to compute such a flow and find the approximation scale 2^k, in a more general setting, with a *best basis* strategy that minimizes the approximation error.

4. Bandelet frames and approximations.

4.1. Bandelet frames. To approximate piecewise regular images having several edges over $[0,1]^2$, the image support is partitioned into regions $[0,1]^2 = \cup_i \Omega_i$, inside each of which the restriction of f either is uniformly regular or is a horizon model or has two edges that meet. This means that in each Ω_i f has either no edge or a single edge whose tangents have an angle smaller than $\pi/3$ with the horizontal or with the vertical direction, or f has an edge junction. This is illustrated by Figure 4.1 with square regions. The size of the squares becomes smaller in the neighborhood of junctions. Bandelet bases are defined over bands that include the regions Ω_i, and it is shown that their union defines a frame of $\mathbf{L^2}[0,1]^2$.

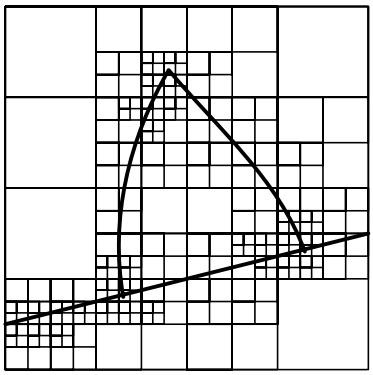

FIG. 4.1. *Partition of an image.*

If f is uniformly regular in a region Ω_i, then there is no need to define a geometric flow. Similarly, if f has an edge junction in Ω_i, then there is no geometric regularity,

and no appropriate geometric flow can be defined. In both cases, f is decomposed in a separable wavelet basis $\mathcal{B}_i = \{b_{i,m}\}_m$ that is constructed over the smallest rectangle B_i that includes Ω_i. If f is a horizon model over Ω_i, then a vertically or horizontally parallel geometric flow is defined over Ω_i. Let B_i be the most narrow band parallel to the flow in Ω_i and that includes Ω_i. Figure 4.2 gives an example. Section 3.1 explains how to construct a bandelet orthonormal basis $\mathcal{B}_i = \{b_{i,m}\}_m$ of $\mathbf{L^2}(B_i)$. The following proposition proves that the union of such orthonormal bases over a segmentation of $[0,1]^2$ defines a frame of $\mathbf{L^2}([0,1]^2)$ that is called a *bandelet frame*. We write P_{Ω_i}, the orthogonal projector defined by

$$P_{\Omega_i} f(x) = \begin{cases} f(x) & \text{if } x \in \Omega_i, \\ 0 & \text{if } x \notin \Omega_i. \end{cases} \tag{4.1}$$

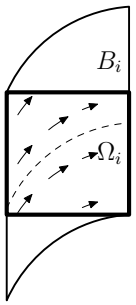

FIG. 4.2. *A square Ω_i with a flow and its associated minimum band B_i.*

PROPOSITION 4.1. *For any segmentation $[0,1]^2 = \cup_i \Omega_i$ and $f \in \mathbf{L^2}[0,1]^2$*

$$f = \sum_{i,m} \langle f, b_{i,m} \rangle \, \tilde{b}_{i,m} \quad \textit{with} \quad \tilde{b}_{i,m} = P_{\Omega_i} b_{i,m} \tag{4.2}$$

and

$$\|f\|^2 \leq \sum_{i,m} |\langle f, b_{i,m} \rangle|^2. \tag{4.3}$$

If the sup over all $x \in [0,1]^2$ of the number of bands B_i that include x is a finite number A, then

$$\|f\|^2 \geq \frac{1}{A} \sum_{i,m} |\langle f, b_{i,m} \rangle|^2, \tag{4.4}$$

and the union of bandelet bases $\mathcal{F} = \cup_i \mathcal{B}_i$ is a frame of $\mathbf{L^2}([0,1]^2)$.

Proof of Proposition 4.1. Since $\mathcal{B}_i = \{b_{i,m}\}_m$ is an orthonormal basis of $\mathbf{L^2}(B_i)$, $P_{B_i} f = \sum_m \langle f, b_{i,m} \rangle \, b_{i,m}$. Moreover, $\Omega_i \subset B_i$ and $[0,1]^2 = \cup_i \Omega_i$, so

$$f = \sum_i P_{\Omega_i} f = \sum_i P_{\Omega_i} P_{B_i} f = \sum_{i,m} \langle f, b_{i,m} \rangle \, P_{\Omega_i} b_{i,m}, \tag{4.5}$$

which proves (4.2).

To prove (4.3) we shall verify a slightly more general property that will be useful later:

$$\text{If} \quad \tilde{f} = \sum_{i,m} \alpha_{i,m}\, \tilde{b}_{i,m}, \quad \text{then} \quad \|\tilde{f}\|^2 \leq \sum_{i,m} |\alpha_{i,m}|^2. \tag{4.6}$$

Indeed, $P_{\Omega_i}\tilde{f} = P_{\Omega_i} \sum_m \alpha_{i,m}\, b_{i,m}$, so $\|P_{\Omega_i}\tilde{f}\|^2 \leq \sum_m |\alpha_{i,m}|^2$, and since $\|\tilde{f}\|^2 = \sum_i \|P_{\Omega_i}\tilde{f}\|^2$ we get (4.6). Applying (4.6) to (4.2) yields (4.3).

Since the sup over all $x \in [0,1]^2$ of the number of regions B_i that includes x is A, one verifies that $\sum_i \|f\|^2_{B_i} \leq A\|f\|^2$ from which we derive (4.4) by inserting (4.5). □

This scheme provides a direct reconstruction of an image from its bandelet coefficients. The discontinuous nature of the border $\tilde{b}_{i,m}$ leads to blocking effects which have to be avoided in image processing but which do not degrade the error decay. To suppress these discontinuities, the reconstruction can be computed with the classical iterative frame algorithm on the full bandelet frame. This would yield a smaller approximation error and may avoid the blocking effect but requires an iterative reconstruction algorithm. Another solution is proposed in [14], where the bandelets themselves are modified in order to cross the boundaries, which removes the blocking artefacts. The bandelet lifting scheme removes the blocking effects, but we then have no proof that the resulting best basis algorithm described in the next section yields an approximation whose error decay is optimal for geometrically regular functions.

4.2. Approximation in a dictionary of bandelet frames. A bandelet frame is defined by geometric parameters that specify the segmentation of the image support into subregions Ω_i and by the geometric flow in each Ω_i. A dictionary of bandelet frames is constructed with segmentations in dyadic square regions and a parameterization of the geometric flow in each region. Within this dictionary, a best bandelet frame is defined by minimizing a Lagrangian cost function. The main theorem computes the error when approximating a geometrically regular function in a best bandelet frame.

As in the wedgelet bases of Donoho [9], the image is segmented in dyadic square regions obtained by successive subdivisions of square regions into four squares of twice smaller width. For a square image support of width L, a square region of width $L\,2^{\lambda}$ is represented by a node at the depth $|\lambda|$ of a quadtree. A square subdivided into four smaller squares corresponds to a node having four children in the quadtree. Figure 4.3 gives an example of a dyadic square image segmentation with the corresponding quadtree. This tree will be coded by the list of the splitting decision at each inner node.

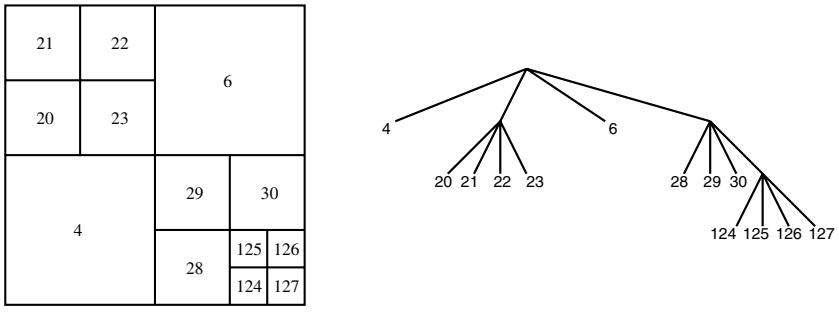

FIG. 4.3. *A dyadic square segmentation and its corresponding quadtree.*

In each dyadic square Ω_i of size 2^{λ_i} a variable indicates if the basis is a wavelet or bandelet basis and in this last case whether the geometric flow is constant vertically or horizontally. If it exists, the flow is characterized by the decomposition (3.15) of an integral curve over a family of scaling functions at a scale 2^{k_i}. Let $\mathcal{F} = \cup_i \mathbf{B}_i$ be the bandelet frame resulting from such a dyadic segmentation of $[0,1]^2$ and such a parameterized flow. An approximation of f in $\mathcal{F}$ is obtained by keeping only the bandelet coefficients above a threshold T:

$$f_M = \sum_{\substack{i,m \\ |\langle f, b_{i,m}\rangle| \geq T}} \langle f, b_{i,m}\rangle \, \tilde{b}_{i,m}, \tag{4.7}$$

with $\tilde{b}_{i,m} = P_{\Omega_i} b_{i,m}$. The total number of parameters is $M = M_S + M_G + M_B$, where M_S is the number of geometric parameters that defines the dyadic image segmentation, M_G is the number of parameters that defines the geometric flows in all the dyadic squares, and M_B is the number of bandelet coefficients above T. Optimizing the frame means finding $\mathcal{F}$, which depends upon f and M, in a dictionary $\mathcal{D}$ of bandelet frame so that

$$\|f - f_M\|^2 \leq C \, M^{-\beta} \tag{4.8}$$

for the largest possible exponent β.

Since $f - f_M = \sum_{|\langle f, b_{i,m}\rangle| < T} \langle f, b_{i,m}\rangle \, \tilde{b}_{i,m}$ we derive from (4.6) that

$$\|f - f_M\|^2 \leq \sum_{\substack{i,m \\ |\langle f, b_{i,m}\rangle| < T}} |\langle f, b_{i,m}\rangle|^2. \tag{4.9}$$

To control $\|f - f_M\|^2$ for a fixed number of parameters $M = M_S + M_G + M_B$, we thus need to minimize $\sum_{|\langle f, b_{i,m}\rangle| < T} |\langle f, b_{i,m}\rangle|^2$. This is done indirectly by minimizing a Lagrangian, similarly to the strategy used by Donoho [9] to optimize the geometry of wedgelet approximations:

$$\mathcal{L}(f, T, \mathcal{F}) = \sum_{\substack{b_{i,m} \in \mathcal{F} \\ |\langle f, b_{i,m}\rangle| < T}} |\langle f, b_{i,m}\rangle|^2 + T^2 \, M \quad \text{with} \quad M = M_S + M_G + M_B. \tag{4.10}$$

The Lagrangian multiplier is T^2 because at the threshold level the squared amplitude T^2 of a bandelet coefficient should increase the Lagrangian by the same amount as an increase by 1 of the number M_B of bandelet coefficients. The *best* bandelet frame $\mathcal{F}$ is the one that minimizes the Lagrangian (4.10) over a dictionary of bandelets.

To reduce the dictionary $\mathcal{D}$ to a finite size, the resolution of the image geometry is limited to T^2, which we shall see is sufficient to approximate f efficiently. This means first that the widths of the square image regions remain larger than T^2, and hence that the maximum depth of the quadtree representing the dyadic square image segmentation is $\lfloor\lfloor |\log_2 T^2| \rfloor\rfloor$. It also means that the decomposition coefficients α_n of the geometry defined in (3.15) are quantized:

$$Q_{T^2}(\alpha_n) = q \, T^2 \quad \text{if} \quad (q - 1/2) \, T^2 \leq \alpha_n < (q + 1/2) \, T^2 \quad \text{with} \quad q \in \mathbb{Z}.$$

In a dyadic square of size 2^λ, since $|\alpha_n| \leq C_\theta \, 2^\lambda$, we necessarily have $|q| \leq C_\theta \, 2^\lambda \, T^{-2} \leq C_\theta \, T^{-2}$. The dictionary $\mathcal{D}_{T^2}$ of bandelet frames constructed over this geometry of

resolution T^2 thus has a finite size, and we can find a best basis that minimizes the Lagrangian (4.10) over this dictionary.

We now concentrate on the approximation capabilities of a best bandelet frame obtained by minimizing $\mathcal{L}(f,T,\mathcal{F})$ over the dictionary of frames $\mathcal{D}_{T^2}$ when f has some geometric regularity. The following definition specifies the geometric regularity conditions on f that will used in the remainder of the paper.

DEFINITION 4.2. *A function f is $\mathbf{C}^\alpha$ geometrically regular over $[0,1]^2$ if*

- *$f = \tilde{f}$ or $f = \tilde{f} \star h$, with $\tilde{f} \in \mathbf{C}^\alpha(\Lambda)$ for $\Lambda = [0,1]^2 - \{\mathcal{C}_\gamma\}_{1\leq\gamma\leq G}$,*
- *the blurring kernel h is $\mathbf{C}^\alpha$, compactly supported in $[-s,s]^2$ and $\|h\|_{\mathbf{C}^\alpha} \leq s^{-(2+\alpha)}$,*
- *the edge curves $\mathcal{C}_\gamma$ are Hölderian of order α and do not intersect tangentially if $\alpha > 1$.*

The following theorem gives an upper bound of the approximation error of f in a best bandelet frame.

THEOREM 4.3. *Let f be a $\mathbf{C}^\alpha$ geometrically regular function and $1 \leq \alpha < p$. There exists C that depends only upon the edge geometry such that for any $T > 0$ the thresholding approximation of f in a best bandelet frame of $\mathcal{D}_{T^2}$ yields an approximation f_M that satisfies*

$$\|f - f_M\|^2 \leq C\, C_f^2\, \max(1, \ell_{\mathcal{C}})^{\alpha+1}\, M^{-\alpha} \tag{4.11}$$

with

$$M \leq C\, \max(1, \ell_{\mathcal{C}})\, C_f^{2/(\alpha+1)} T^{-2/(\alpha+1)}, \tag{4.12}$$

where $C_f = \max(\|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)} \max(\|c\|_{\mathbf{C}^\alpha}^\alpha, 1)^2, \|c\|_{\mathbf{C}^\alpha}^{(\alpha+1)/(2\alpha)}, \|f\|_{\mathbf{C}^1(\Lambda)}^{\alpha+1})$ and $\ell_{\mathcal{C}}$ is the total length of the curves.

The remarks on the nonoptimality of the constant C_f of Theorem 3.2 apply here. In addition, the dependency of C on the edge geometry is not explicitly controlled but involves the number of curves G as well as the geometric configuration (distance, angle of the crossing, etc.).

Proof of Theorem 4.3. The theorem proof is based on the following lemma, which exhibits a bandelet frame in $\mathcal{D}_{T^2}$ having suitable approximation properties when $\alpha > 1$, and thus the edges have tangents. The remaining case, $\alpha = 1$, is obtained with the classical wavelet basis that is included in the dictionary.

LEMMA 4.4. *Under the hypotheses of Theorem* 4.3 *and if $\alpha > 1$, for any $T > 0$ there exists a bandelet frame $\mathcal{F} \in \mathcal{D}_{T^2}$ such that*

$$\mathcal{L}(f,T,\mathcal{F}) \leq C\, C_f^{2/(\alpha+1)}\, \max(1, \ell_{\mathcal{C}})\, T^{2\alpha/(\alpha+1)} \tag{4.13}$$

for a constant C that depends only upon the edge geometry.

We first prove that this lemma implies Theorem 4.3 and will then prove the lemma. Let $\mathcal{F}' = \{b'_{i,m}\}_{i,m}$ be a best bandelet frame that minimizes $\mathcal{L}(f,T,\mathcal{F})$ in $\mathcal{D}_{T^2}$. Clearly, $\mathcal{L}(f,T,\mathcal{F}') \leq \mathcal{L}(f,T,\mathcal{F})$, where $\mathcal{F}$ is the frame of Lemma 4.4, and hence

$$\mathcal{L}(f,T,\mathcal{F}') = \sum_{\substack{b'_{i,m}\in\mathcal{F}' \\ |\langle f, b'_{i,m}\rangle| < T}} |\langle f, b'_{i,m}\rangle|^2 + T^2 M \leq C\, C_f^{2/(\alpha+1)}\, \max(1, \ell_{\mathcal{C}})\, T^{2\alpha/(\alpha+1)}. \tag{4.14}$$

As a result

$$\sum_{\substack{b'_{i,m}\in\mathcal{F}'\\ |\langle f,b'_{i,m}\rangle|<T}} |\langle f,b'_{i,m}\rangle|^2 \le C\, C_f^{2/(\alpha+1)}\, \max(1,\ell_{\mathcal{C}})\, T^{2\alpha/(\alpha+1)} \tag{4.15}$$

and

$$M \le C\, C_f^{2/(\alpha+1)}\, \max(1,\ell_{\mathcal{C}})\, T^{2\alpha/(\alpha+1)} T^{-2}, \tag{4.16}$$

$$M \le C\, C_f^{2/(\alpha+1)}\, \max(1,\ell_{\mathcal{C}})\, T^{-2/(\alpha+1)}. \tag{4.17}$$

Inserting (4.17) in (4.15) yields

$$\sum_{\substack{b'_{i,m}\in\mathcal{F}'\\ |\langle f,b'_{i,m}\rangle|<T}} |\langle f,b'_{i,m}\rangle|^2 \le C\, {C_f}^2\, \max(1,\ell_{\mathcal{C}})\, M^{-\alpha}, \tag{4.18}$$

and since we saw in (4.9) that

$$\|f-f_M\|^2 \le \sum_{\substack{b'_{i,m}\in\mathcal{F}'\\ |\langle f,b'_{i,m}\rangle|<T}} |\langle f,b'_{i,m}\rangle|^2,$$

this proves the theorem result (4.11). □

The core of Theorem 4.3 is thus Lemma 4.4, whose proof is constructive.

Proof of Lemma 4.4. We first design a dyadic square segmentation associated with f and T and a flow in each square so that the resulting bandelet frame $\mathcal{F}$ satisfies (4.13). This segmentation is constructed by separating squares that are close to an edge from the others. *Regular* squares $i \in I_{\mathcal{R}}$ are squares which are distant by more than s from all edges $\mathcal{C}_\gamma$. In such squares, f is uniformly regular. No geometric flow is defined in regular squares, which means that the bandelet basis is a separable wavelet basis.

Let us now consider squares that include or are adjacent to a single edge. We say that a region Ω includes a single horizontal edge component if it is crossed horizontally by a single edge $\mathcal{C}_\gamma$ that remains at distance smaller than s and whose tangents have angles with the horizontal direction smaller than $\pi/3$ and if there is no other edge at a distance smaller than s. It is tempting to use this definition directly on the Ω_i. However, to efficiently approximate f with such bandelets, one must also verify that the larger band B_i includes no other edge component. To enforce this property, if the size of Ω_i is 2^λ, a larger vertical rectangle $\widetilde{\Omega}_i$ of height $7\times 2^\lambda$ and width 2^λ, centered in Ω_i, is considered. We shall verify that $B_i \subset \widetilde{\Omega}_i$.

A square Ω_i is said to be a *horizontal edge* square if Ω_i is at a distance smaller than s from an edge curve and if $\widetilde{\Omega}_i$ includes a single horizontal edge component $\mathcal{C}_\gamma$ and remains at a distance larger than s from the two straight horizontal segments of the boundary of $[0,1]^2$, and we write $i \in I_{\mathcal{H}}$. In this case, a vertically parallel geometric flow integral is defined in Ω_i by approximating c_i at a scale $2^k = \max(\|c\|_{\mathbf{C}^\alpha}^{1/\alpha}, 1)\, \eta^{1/\alpha}$, where $\eta = \max(s, T^{2\alpha/(\alpha+1)})$ plays the role of a geometric resolution, and by quantizing the resulting coefficients:

$$g_i(x) = \sum_{n=1}^{2^{\lambda-k}} Q_{T^2}\left(\frac{\langle c_i,\theta_{k,n}\rangle}{\|\theta_{k,n}\|^2}\right)\theta_{k,n}(x) \tag{4.19}$$

with

$$Q_{T^2}(x) = qT^2 \quad \text{if} \quad (q-1/2)T^2 \leq x < (q+1/2)T^2 \quad \text{with} \quad q \in \mathbb{Z}.$$

Lemma 3.6 proves that choosing $2^k = \max(\|c\|_{\mathbf{C}^\alpha}^{-1/\alpha}, 1)\,\eta^{1/\alpha}$ implies that the error between g_i and c_i satisfies $\|g_i - c_i\|_\infty \leq C\,\max(\|c\|_{\mathbf{C}^\alpha}, 1)\,\eta$.

Vertical edge squares are defined similarly by inverting the roles of the horizontal and vertical directions, and hence of x_1 and x_2. If Ω_i is a square of size 2^λ, then the $\widetilde{\Omega}_i$ is a horizontal rectangle of height 2^λ and width $7 \times 2^\lambda$. A square Ω_i is said to be a vertical edge square if Ω_i is at a distance smaller than s from an edge curve and if $\widetilde{\Omega}_i$ includes a single vertical edge component $\mathcal{C}_\gamma$ and remains at a distance larger than s from the two straight vertical segments of the boundary of $[0,1]^2$. We then write $i \in I_\mathcal{V}$. A horizontally parallel flow is then defined in Ω_i by approximating the edge parameterization c_i as in (4.19). A square that is both a horizontal edge square and a vertical edge square is considered as a horizontal edge square by default.

The following algorithm constructs a dyadic image segmentation by labeling regular, vertical edge, and horizontal edge squares and introduces a third residual class called *junction* squares because we shall see that they are close to junction points. No geometric flow is defined in junction squares because they are close to several edges, and the corresponding bandelet basis is thus a separable wavelet basis.

- **Initialization**: Label the square $[0,1]^2$ a *temporary* square.
- **Step 1:** Split into four every *temporary* square and remove it.
- **Step 2:** Label, in the following order, each new subdivided square Ω_i as a
  - *regular* square $i \in I_\mathcal{R}$ if it is at a distance larger than s from all edges,
  - *junction* square $i \in I_\mathcal{J}$ if its size is smaller than η,
  - *horizontal edge* square $i \in I_\mathcal{H}$ if $\widetilde{\Omega}_i$ includes a single horizontal edge component,
  - *vertical edge* square $i \in I_\mathcal{V}$ if $\widetilde{\Omega}_i$ includes a single vertical edge component,
  - *temporary* square otherwise.
- **Step 3:** Go to Step 1 if there remain *temporary* squares.

Figure 4.4 illustrates such a labeled segmentation while Figure 4.5 displays a close-up on a junction of an intermediate partition.

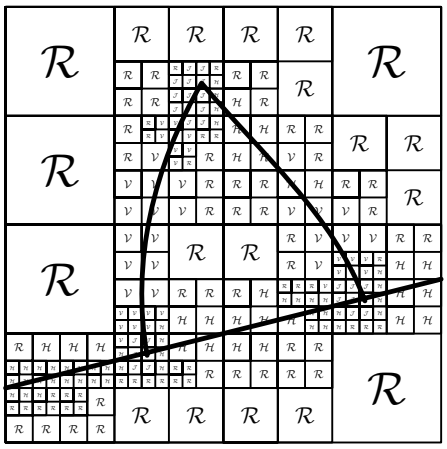

FIG. 4.4. *Partition with dyadic square labeled $\mathcal{J}$, $\mathcal{R}$, $\mathcal{H}$, and $\mathcal{V}$.*

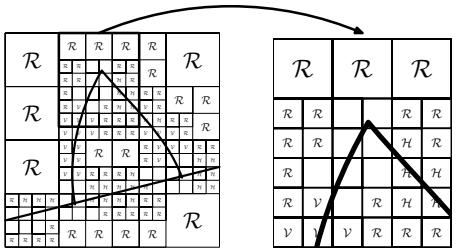

FIG. 4.5. *Close-up on a junction zone with an intermediate partition with dyadic square labeled* $\mathcal{R}$*,* $\mathcal{H}$*,* $\mathcal{V}$*, and temporary square.*

From such a dyadic image partition, we associate a bandelet frame $\mathcal{F}$ as the union of the bandelet bases defined with the geometric flow constructed (or not) in each square depending upon their label. We now prove that the resulting Lagrangian satisfies the property (4.13) of Lemma 4.4. To evaluate the Lagrangian,

$$\mathcal{L}(f,T,\mathcal{F}) = \sum_{\substack{b_{i,m}\in\mathcal{F}\\ |\langle f,b_{i,m}\rangle|<T}} |\langle f,b_{i,m}\rangle|^2 + T^2\,(M_S + M_G + M_G), \tag{4.20}$$

we separate the geometrical cost $M_S + M_G$ and decompose it into

$$\begin{aligned}\mathcal{L}(f,T,\mathcal{F}) = T^2\,(M_S + M_G) + \sum_{i\in I_{\mathcal{R}}} \tilde{\mathcal{L}}(f,T,\mathcal{B}_i) + \sum_{i\in I_{\mathcal{J}}} \tilde{\mathcal{L}}(f,T,\mathcal{B}_i)\\ + \sum_{i\in I_{\mathcal{H}}} \tilde{\mathcal{L}}(f,T,\mathcal{B}_i) + \sum_{i\in I_{\mathcal{V}}} \tilde{\mathcal{L}}(f,T,\mathcal{B}_i),\end{aligned} \tag{4.21}$$

with

$$\tilde{\mathcal{L}}(f,T,\mathcal{B}_i) = \sum_{\substack{b_{i,m}\in\mathcal{B}_i\\ |\langle f,b_{i,m}\rangle|<T}} |\langle f,b_{i,m}\rangle|^2 + T^2\,M_{B,i}, \tag{4.22}$$

where $M_{B,i}$ is the number of inner products $|\langle f,b_{i,m}\rangle| \geq T$ for a fixed i.

To prove that $\mathcal{L}(f,T,\mathcal{F}) \leq C\,C_f^{2/(\alpha+1)}\,\max(1,\ell_{\mathcal{C}})\,T^{2\alpha/(\alpha+1)}$, a similar bound for $T^2\,(M_S + M_G)$ and each of the four partial sums corresponding to different classes of squares in (4.21) shall be proved.

The first lemma characterizes the dyadic segmentation obtained by our splitting algorithm and computes the resulting number of geometric parameters. Its proof can be found in Appendix B.1.

LEMMA 4.5. *There exists a constant* C *that depends upon the edge geometry such that the resulting dyadic image segmentation defined recursively includes at most* $C\,|\log_2 \max(s, T^{2\alpha/(\alpha+1)})|$ *squares with at most* C *junction squares. Furthermore,*

$$T^2\,(M_S + M_G) \leq C\,\ell_{\mathcal{C}}\,\max(\|c\|_{\mathbf{C}^\alpha}^{-1/\alpha}, 1)\,T^{2\alpha/(\alpha+1)}. \tag{4.23}$$

The next lemma computes the Lagrangian value on regular squares by using standard wavelet approximation properties for uniformly regular functions. Its proof is in Appendix B.2.

LEMMA 4.6. *There exists a constant C that depends upon the edge geometry such that the sum of the partial Lagrangian $\tilde{\mathcal{L}}$ over each regular square Ω_i ($i \in I_{\mathcal{R}}$) satisfies*

$$\sum_{i \in I_{\mathcal{R}}} \tilde{\mathcal{L}}_i(f, T, \mathcal{B}_i) \leq C\, C_f^{2/(\alpha+1)}\, T^{2\alpha/(\alpha+1)}. \tag{4.24}$$

For junction squares the Lagrangian value is calculated using the fact that there are few such squares and that they have a small size bounded by η. The lemma's proof is in Appendix B.3.

LEMMA 4.7. *There exists a constant C that depends upon the edge geometry such that the sum of the partial Lagrangian $\tilde{\mathcal{L}}$ over all junction squares Ω_i ($i \in I_{\mathcal{J}}$) satisfies*

$$\sum_{i \in I_{\mathcal{J}}} \tilde{\mathcal{L}}_i(f, T, \mathcal{B}_i) \leq C\, C_f^{2/(\alpha+1)}\, T^{2\alpha/(\alpha+1)}. \tag{4.25}$$

The final lemma gives an upper bound on the Lagrangian of bandelets constructed over vertical edge squares. Its proof is in Appendix B.4. It is at the core of our construction and relies on the precision of the geometric flow used in each edge square.

LEMMA 4.8. *There exists a constant C that depends upon the edge geometry such that the sum of the partial Lagrangian $\tilde{\mathcal{L}}$ over all horizontal edge squares Ω_i ($i \in I_{\mathcal{H}}$) satisfies*

$$\sum_{i \in I_{\mathcal{H}}} \tilde{\mathcal{L}}(f, T, \mathcal{B}_i) \leq C\, C_f^{2/(\alpha+1)}\, \max(1, \ell_{\mathcal{C}})\, T^{2\alpha/(\alpha+1)}. \tag{4.26}$$

Transposing this result to vertical edge squares proves that

$$\sum_{i \in I_{\mathcal{V}}} \tilde{\mathcal{L}}(f, T, \mathcal{B}_i) \leq C\, C_f^{2/(\alpha+1)}\, \max(1, \ell_{\mathcal{C}})\, T^{2\alpha/(\alpha+1)}. \tag{4.27}$$

Inserting (4.23), (4.24), (4.25), (4.26), and (4.27) in (4.21) proves the result (4.13) of Lemma 4.4. □

This theorem provides a constructive approximation scheme with an optimal approximation bound. The decay rate $M^{-\alpha}$ of the error is optimal, as it is the same as the optimal one for uniformly $\mathbf{C}^\alpha$ functions. It is much better than the decay rate M^{-1} for the wavelets and improves the decay rate $M^{-2}\,(\log_2 M)^3$ of the curvelets even if $\alpha = 2$. Furthermore, the Lagrangian minimization does not require any information on the regularity parameter α or on the smoothing kernel h. However, an exhaustive search to minimize the Lagrangian in the dictionary $\mathcal{D}_{T^2}$ requires an exponential number of operations which prohibits its practical use. Section 6.2 introduces a modified dictionary and a fast algorithm that finds a best bandelet basis with a polynomial complexity, at the cost of adding a logarithmic factor in the resulting approximation error.

5. Image compression. For image compression, we must minimize the total number of bits R required to encode the approximation as opposed to the number of parameters M. An image is compressed in a bandelet frame by first coding the segmentation of the image support and a geometric flow in each region of the segmentation. The decomposition coefficients of the image in the resulting bandelet frame are then quantized and stored with a binary code. This very simple algorithm does not provide a scalable scheme but does give an almost optimal distortion rate.

We denote by R the resulting total number of bits to encode a bandelet frame and the bandelet coefficients of f in this frame. It can be decomposed into

$$R = R_S + R_G + R_B, \tag{5.1}$$

where R_S is the number of bits to encode the dyadic square segmentation, R_G is the number of bits to encode the flow in each square region, and R_B is the number of bits to encode the quantized bandelet coefficients.

We saw in section 4.2 that a dyadic square segmentation of $[0,1]^2$ is represented by a quadtree whose leaves are the square regions of the partition. Each interior node of the tree corresponds to a square that is split into four subsquares. This splitting decision is encoded with a bit equal to 1. The leaves of the tree correspond to squares which are not split, which is encoded with a bit equal to 0. With this code, the number of bits R_S that specify the segmentation quadtree is thus equal to the number of nodes of this quadtree.

Over a square of size 2^λ, the geometric flow is parameterized at a scale 2^k in (3.15) by $2^{\lambda-k}$ quantized coefficients $\alpha_m = q\,T^2$ with $|q|T^2 \leq C_\theta$. We thus need $2^{\lambda-k}\log_2(C_\theta T^{-2})$ bits to encode these α_m. The number of bits R_G to encode the flows is the sum of these values over all squares where a flow is defined plus the number of bits required to specify the scale.

In a bandelet frame $\mathcal{F} = \{b_{i,m}\}_{i,m}$, all bandelet coefficients $\langle f, b_{i,m}\rangle$ are uniformly quantized with a uniform quantizer Q_T of step T:

$$Q_T(x) = qT \quad \text{if} \quad (q-1/2)T \leq x < (q+1/2)T \quad \text{with} \quad q \in \mathbb{Z}.$$

The indices i, m of the M_B nonzero quantized coefficients are encoded together with the value $Q_T(\langle f, b_{i,m}\rangle) \neq 0$. The proof of Theorem 4.3 shows that the M_B nonzero quantized bandelet coefficients whose amplitude is larger than T appear at a scale larger than $C_\psi^{-1}\,\|f\|_\infty^{-1}\,T$. Since there are $C_\psi^2\,\|f\|_\infty^2\,T^{-2}$ such coefficients, to encode an index i, m is equivalent to encoding an integer in $[1, C_\psi^2\,\|f\|_\infty^2\,T^{-2}]$ which requires $\log_2(C_\psi^2\,\|f\|_\infty^2\,T^{-2})$ bits. Since $|\langle f, b_{i,m}\rangle| \leq \|f\|$, each nonzero quantized coefficient is encoded with $\lceil |\log_2(\|f\|/T)| \rceil$ bits. The total number of bits to encode the quantized bandelet coefficients thus satisfies

$$R_B \leq M_B\,(\lceil |\log_2(\|f\|/T)| \rceil + \log_2(C_\psi^2\,\|f\|_\infty^2\,T^{-2})). \tag{5.2}$$

The image restored from its bandelet coefficients is

$$f_R = \sum_{i,m} Q_T(\langle f, b_{i,m}\rangle)\,\tilde{b}_{i,m}, \tag{5.3}$$

and the resulting distortion is $D(R) = \|f - f_R\|^2$. Since $f - f_R = \sum_{i,m}(\langle f, b_{i,m}\rangle - Q_T(\langle f, b_{i,m}\rangle))\,\tilde{b}_{i,m}$ we derive from (4.6) that

$$D(R) = \|f - f_R\|^2 \leq \sum_{i,m} |\langle f, b_{i,m}\rangle - Q_T(\langle f, b_{i,m}\rangle)|^2. \tag{5.4}$$

For the M_B nonzero quantized coefficients, we have $|x - Q_T(x)|^2 \leq T^2/4$, so

$$D(R) \leq \sum_{\substack{i,m \\ |\langle f, b_{i,m}\rangle| \leq T/2}} |\langle f, b_{i,m}\rangle|^2 + M_B T^2/4, \tag{5.5}$$

$$D(R) \leq \mathcal{L}(f, T/2, \mathcal{F}). \tag{5.6}$$

This proves that a small distortion rate can be obtained by finding a best bandelet frame in $\mathcal{D}_{T^2}$ that minimizes $\mathcal{L}(f, T/2, \mathcal{F})$. The following theorem computes the resulting decay of $D(R)$ as a function of R for a geometrically regular image in a best bandelet frame.

THEOREM 5.1. *Let f be a $\mathbf{C}^\alpha$ geometrically regular function and $\alpha < p$. There exists C that depends only upon the edge geometry such that for any $T > 0$ coding f in a best bandelet frame that minimizes $\mathcal{L}(f, T/2, \mathcal{F})$ over $\mathcal{D}_{T^2}$ yields a distortion rate that satisfies*

$$D(R) = \|f - f_R\|^2 \leq C\, C_f^2\, \max(1, \ell_{\mathcal{C}})^{\alpha+1} R^{-\alpha}\, |\log_2 R|^{\alpha+1} \tag{5.7}$$

with

$$R \leq C\, C_f^{2/(\alpha+1)}\, \max(1, \ell_{\mathcal{C}})\, T^{-2/(\alpha+1)}\, |\log_2(T)|, \tag{5.8}$$

where $C_f = \max(\|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)} \max(\|c\|_{\mathbf{C}^\alpha}^\alpha, 1)^2, \|c\|_{\mathbf{C}^\alpha}, \|\tilde{f}\|_{\mathbf{C}^1(\Lambda)}^{\alpha+1})$ and $\ell_{\mathcal{C}}$ is the total length of the edge curves.

Proof of Theorem 5.1. The compressed image f_R is given by (5.3) in the bandelet frame that minimizes $\mathcal{L}(f, T/2, \mathcal{F})$.

To bound the distortion $D(R)$, we use that Lemma 4.4 proves that the best bandelet frame in $\mathcal{D}_{T^2}$ satisfies

$$\mathcal{L}(f, T/2, \mathcal{F}) \leq C\, C_f^{2/(\alpha+1)}\, \max(1, \ell_{\mathcal{C}})\, T^{2\alpha/(\alpha+1)}. \tag{5.9}$$

Inserting this bound in (5.6) implies

$$D(R) \leq C\, C_f^{2/(\alpha+1)}\, \max(1, \ell_{\mathcal{C}})\, T^{2\alpha/(\alpha+1)}. \tag{5.10}$$

The corresponding bit rate R is decomposed in (5.1) in three terms: R_S the number of bits to encode the dyadic square segmentation, R_G the number of bits to encode the flow in each square region, and R_B the number of bits to encode the quantized bandelet coefficients.

As the number of splits that specify the best segmentation is of order $\log T$ as shown in Lemma 4.5, R_S is bounded by the same quantity. The specification of the flow requires a flow/no flow bit for each square. For the squares of size 2^λ with a flow, $\log_2 \lambda \leq \log_2 \log_2(C_\psi^2\, \|f\|_\infty^2\, T^{-2})$ bits are needed to specify the scale 2^k and $2^{\lambda-k} \log_2(C_\theta\, T^{-2})$ for the quantized coefficients. As shown in Lemma 4.5, there are at most $\log T$ squares and at most $T^{-2}\, T^{2\alpha/(\alpha+1)}$ quantized coefficients, so R_G satisfies

$$R_G \leq \log_2 T(1 + \log_2 \log_2(C_\psi^2\, \|f\|_\infty^2\, T^{-2})) + T^{-2}\, T^{2\alpha/(\alpha+1)} \log_2(C_\theta\, T^{-2}). \tag{5.11}$$

Lemma 4.4 implies that the number of nonzero bandelet coefficients M_B satisfies

$$M_B \leq C\, C_f^{2/(\alpha+1)}\, \max(1, \ell_{\mathcal{C}})\, T^{-2}\, T^{2\alpha/(\alpha+1)}. \tag{5.12}$$

Inserting this bound in (5.2), one obtains

$$R_B \leq C\, C_f^{2/(\alpha+1)}\, \max(1, \ell_{\mathcal{C}})\, T^{-2}\, T^{2\alpha/(\alpha+1)} (\lceil |\log_2(\|f\|/T)| \rceil + \log_2(C_\psi\, \|f\|_\infty^2\, T^{-2})). \tag{5.13}$$

We thus have

$$R = R_S + R_G + R_B, \tag{5.14}$$

$$R \leq C\, C_f^{2/(\alpha+1)}\, \max(1, \ell_{\mathcal{C}})\, T^{-2/(\alpha+1)}\, |\log_2(T)|. \tag{5.15}$$

To conclude, one can verify that inserting this bound in (5.10) yields

$$D(R) \leq C\, C_f^2\, \max(1, \ell_{\mathcal{C}})^{\alpha+1}\, R^{-\alpha}\, |\log_2(R)|^{\alpha+1}. \qquad \square \tag{5.16}$$

This theorem proves that the asymptotic decay of a bandelet transform code reaches the Kolmogorov lower bound up to the $|\log_2 R|^{\alpha+1}$ term. Indeed, the class of images that we consider includes the class of images that are $\mathbf{C}^\alpha$ over $[0,1]^2$, and we know that for such a class of images the Kolmogorov lower bound of the distortion rate decays like $R^{-\alpha}$ [10].

6. Discretized image and discretized bandelets.

6.1. Discrete orthogonal bandelet bases. A discrete image measured by a CCD camera is an array of pixels obtained by averaging the input analog intensity $f(x_1, x_2)$ over square photoreceptors. We thus do not have access to f, but we explain how to compute discrete bandelet coefficients from these pixels with a fast algorithm.

At a discretization scale ϵ, a receptor covers a square of surface ϵ^2. Let $\mathbf{1}_{n_1,n_2}(x)$ be the indicator function of the square $[n_1,\, n_1+1] \times [n_2,\, n_2+1]$. A pixel value is the average of $f(x_1, x_2)$ over the photoreceptor surface multiplied by a renormalization factor ϵ. The resulting discretized image values are

$$\overline{f}[n_1, n_2] = \epsilon^{-1} \langle f(x)\,,\, \mathbf{1}_{n_1,n_2}(\epsilon^{-1}\, x)\rangle = \epsilon^{-1} \iint f(x_1, x_2)\, \mathbf{1}_{n_1,n_2}(\epsilon^{-1}\, x_1, \epsilon^{-1}\, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2 \tag{6.1}$$

for $0 \leq n_1, n_2 \leq \epsilon^{-1}$. This choice of renormalization for the average ensures $\|\overline{f}\| \simeq \|f\|$. In the following we shall suppose that ϵ^{-1} is an integer.

Discretized orthogonal bandelets are defined with the same approach as in section 3.1. Suppose that the analog image $f(x_1, x_2)$ has a single horizontal edge along a curve $\mathcal{C}$, an edge with tangents having an angle smaller than $\pi/3$ with the horizontal direction. We define a vertically parallel flow that approximates the tangents of this edge $\vec{\tau}(x_1, x_2) = \vec{\tau}(x_1) = (1, g'(x_1))$. We suppose that the flow integral $g(x) = \int_0^x g(t)\mathrm{d}t$ is parameterized in a basis of scaling functions according to (3.15).

The continuous variable warping operator W in (3.3) is approximated by a discretized warping operator that operates over the image sampling grid:

$$\overline{Wf}[n_1, n_2] = f[n_1, n_2 + \lceil \epsilon^{-1} g(\epsilon\, n_1)\rceil], \tag{6.2}$$

where $\lceil x \rceil$ is the integer just above x. As in (3.2), a band B parallel to this flow is defined from this warping over the sampling grid by

$$B = \left\{ (n_1, n_2) \;:\; n_1 \in [a_1, b_1],\, n_2 \in [\lceil \epsilon^{-1} g(\epsilon\, n_1)\rceil + a_2\,,\, \lceil \epsilon^{-1} g(\epsilon\, n_1)\rceil + b_2] \right\} \tag{6.3}$$

so that $\mathcal{C} \subset B$.

If $\overline{f}$ has a support in B, then $\overline{Wf}$ has a support in $\overline{W}B$ which is a rectangle whose sides are horizontal and vertical. The discrete separable orthonormal wavelet basis over the rectangle $\overline{W}B$ can be written

$$\left\{ \begin{array}{lll} \overline{\phi}_{j,m_1}[n_1]\, \overline{\psi}_{j,m_2}[n_2] & , & \overline{\psi}_{j,m_1}[n_1]\, \overline{\phi}_{j,m_2}[n_2] \\ & , & \overline{\psi}_{j,m_1}[n_1]\, \overline{\psi}_{j,m_2}[n_2] \end{array} \right\}_{(j,m_1,m_2) \in I_{\overline{W}B}}. \tag{6.4}$$

The operator $\overline{W}$ is orthogonal, and hence

$$\left\{ \begin{array}{lll} \overline{W}^{-1}\overline{\phi}_{j,m_1}[n_1]\,\overline{\psi}_{j,m_2}[n_2] & , & \overline{W}^{-1}\overline{\psi}_{j,m_1}[n_1]\,\overline{\phi}_{j,m_2}[n_2] \\ & , & \overline{W}^{-1}\overline{\psi}_{j,m_1}[n_1]\,\overline{\psi}_{j,m_2}[n_2] \end{array} \right\}_{(j,m_1,m_2)\in I_{\overline{W}B}} \tag{6.5}$$

is an orthonormal wavelet basis of signals defined in the band B. Using the fact that $\langle \overline{W}\overline{f}, \Psi\rangle = \langle \overline{f}, \overline{W}^{-1}\Psi\rangle$ the discrete warped wavelet transform of $\overline{f}$ is computed with the fast discrete wavelet transform of $\overline{W}\overline{f}$, which requires $O(\#B\epsilon^{-2})$ operations, where $\#B$ is the surface of the band B.

As in section 3.1, the bandeletization replaces the discrete warped wavelets $\{\overline{W}^{-1}\overline{\phi}_{j,m_1}[n_1]\,\overline{\psi}_{j,m_2}[n_2]\}_{m_1,m_2}$ that do not have any vanishing moments along n_1 by a corresponding equivalent family of discrete bandelets $\{\overline{W}^{-1}\overline{\psi}_{l,m_1}[n_1]\,\overline{\psi}_{j,m_2}[n_2]\}_{l>j,m_1,m_2}$ that have vanishing moments along the first coordinate (n_1). The bandelet coefficients of $\overline{f}$ are computed by applying a one-dimensional discrete wavelet transform over the corresponding warped wavelet coefficients of $\overline{f}$, which also requires $O(\#B\,\epsilon^{-2})$ operations. Algorithmic details are given in [14].

If the edge $\mathcal{C}$ has an angle smaller than $\pi/3$ with the vertical direction, then we define a flow that is parallel horizontally. By transposing the procedure previously described and exchanging n_1 and n_2, a similar bandelet orthonormal basis is calculated over a band B parallel to this flow.

In the following, a discrete bandelet basis with a horizontally parallel or vertically parallel flow is written $\mathbf{B} = \{\overline{b}_m\}_m$.

6.2. Approximation in a best discrete bandelet frame. A discrete bandelet frame is defined as a union of discrete orthogonal bandelet bases constructed over the regions of a dyadic square image segmentation $[0,1]^2 = \cup_i \Omega_i$. This section proves that the approximation result of Theorem 4.3 remains valid for discrete images decomposed in a discrete bandelet frame.

In each square Ω_i, we construct either a separable discrete wavelet basis or a bandelet basis over the smallest band B_i including Ω_i and which is parallel to the flow in Ω_i. One can verify that $\#B_i \leq 3\#\Omega_i$. The wavelet or bandelet basis associated with Ω_i is written $\overline{\mathbf{B}}_i = \{\overline{b}_{i,m}\}_m$, and the union of these bases $\overline{\mathcal{F}} = \cup_i \mathbf{B}_i$ is a discrete frame. The corresponding discrete wavelet or bandelet coefficients of $\overline{f}$ are computed with a fast algorithm that requires $O(\#\Omega_i\epsilon^{-2})$ operations. Since $\sum_i \#\Omega_i = 1$ the fast transforms over all dyadic squares of $[0,1]^2$ require $O(\epsilon^{-2})$ operations.

Let P_{Ω_i} be the orthogonal projector over signals having a support in Ω_i. We verify, as in Proposition 4.1, that

$$\overline{f} = \sum_{i,m} \langle \overline{f}, \overline{b}_{i,m}\rangle\, P_{\Omega_i}\overline{b}_{i,m} \tag{6.6}$$

and

$$\|\overline{f}\|^2 \leq \sum_{i,m} |\langle \overline{f}, \overline{b}_{i,m}\rangle|^2. \tag{6.7}$$

An approximation $\overline{f}_M$ is obtained by keeping all coefficients above a threshold T:

$$\overline{f}_M = \sum_{\substack{i,m \\ |\langle \overline{f}, \overline{b}_{i,m}\rangle| \geq T}} \langle \overline{f}, \overline{b}_{i,m}\rangle\, P_{\Omega_i}\overline{b}_{i,m}. \tag{6.8}$$

The total number of parameters is $M = M_S + M_G + M_B$, where M_S is the number of parameters describing the segmentation, M_G is the number of geometric parameters describing the flow, and M_B is the number of bandelet coefficients above T.

To minimize the error $\|\overline{f} - \overline{f}_M\|$, as in section 4.2, we search for a best bandelet frame which minimizes the Lagrangian:

$$\mathcal{L}(\overline{f}, T, \overline{\mathcal{F}}) = \sum_{\substack{\overline{b}_{i,m} \in \overline{\mathcal{F}} \\ |\langle \overline{f}, \overline{b}_{i,m} \rangle| < T}} |\langle \overline{f}, \overline{b}_{i,m} \rangle|^2 + T^2 M \quad \text{with} \quad M = M_S + M_G + M_B. \tag{6.9}$$

The complexity of the best bandelet frame search in section 4.2 is driven by the complexity to find the best geometric flow in a square, which is exponential. This exponential complexity makes it impossible to use such a best bandelet search algorithm in numerical computations. A polynomial complexity algorithm is introduced by choosing geometric flows that are piecewise polynomial.

In the bandelet dictionary of section 4.2, over a square Ω of size 2^λ, the flow is parameterized in a family of scaling functions (3.15) at a scale 2^k. We replace such a flow by a piecewise polynomial flow over the $2^{\lambda-k}$ intervals of sizes 2^k:

$$\forall r \in \mathbb{Z} \quad \text{with} \quad 1 \le r \le 2^{\lambda-k}, \quad \forall t \in [r2^k\,, (r+1)2^k), \quad g(t) = \sum_{n=1}^{p} \alpha_{r,n}\, \theta_n(2^{-k}t - r), \tag{6.10}$$

where $\{\theta_n\}_{1 \le n \le p}$ is an orthogonal basis of the space of polynomials of degree p over $[0,1]$ and which vanish at 0. However, instead of considering this as a single flow in the square Ω of width 2^λ, we shall view it as a subdivision of Ω into $2^{\lambda-k}$ rectangles Ω_r of length 2^λ and width 2^k, inside each of which the flow is a polynomial parameterized by $\{\alpha_{n,r}\}_{1 \le n \le p}$.

To construct an image partition, we first begin with a dyadic square segmentation $[0,1]^2 = \cup_i \Omega_i$. If Ω_i has a geometric flow, then it is subdivided into 2^{k_i} subrectangles $\Omega_{i,r}$ having polynomial geometric flows. An orthogonal bandelet basis $\overline{\mathcal{B}}_{i,r} = \{b_{i,r,m}\}_m$ is defined over each band $B_{i,r}$ associated with a subrectangle $\Omega_{i,r}$. The union of these bandelet orthogonal bases defines a bandelet frame $\overline{\mathcal{F}}_i$ over Ω_i: $\overline{\mathcal{F}}_i = \cup_{r=1}^{2^{\lambda-k}} \overline{\mathcal{B}}_{i,r}$. If Ω_i has no geometric flow, then $\overline{\mathcal{F}}_i$ is the discrete separable wavelet basis defined over Ω_i. The union of these families of bandelets for all Ω_i defines a bandelet frame $\overline{\mathcal{F}} = \cup_i \overline{\mathcal{F}}_i$ over the image support such that

$$\overline{f} = \sum_{b_{i,r,m} \in \overline{\mathcal{F}}} \langle \overline{f}, \overline{b}_{i,r,m} \rangle\, P_{\Omega_{i,r}} \overline{b}_{i,r,m} \tag{6.11}$$

and

$$\|\overline{f}\|^2 \le \sum_{b_{i,r,m} \in \overline{\mathcal{F}}} |\langle \overline{f}, \overline{b}_{i,r,m} \rangle|^2. \tag{6.12}$$

The geometric resolution is limited to T^2 by imposing that the size of each square Ω_i and each rectangle $\Omega_{i,r}$ is larger than T^2 and that the flow parameters $\{\alpha_{r,n}\}_{1 \le n \le p}$ are quantized uniformly with a step equal to T^2. Since $|g'(t)| \le 2$ one can verify that $|\alpha_{r,n}| \le 2^k$, so after quantization it can take $C\, 2^{k+1}\, T^{-2}$ possible values. As a consequence, there are $\mathcal{O}(2^k T^{-2p})$ different polynomial flows over each $\Omega_{i,r}$. We denote by $\mathcal{D}_{T^2}$ the dictionary of all such bandelet frames.

To find the bandelet frame which minimizes the Lagrangian we use a classification and regression tree (CART) algorithm that takes advantage of the additivity of the Lagrangian:

$$\mathcal{L}(\overline{f}, T, \overline{\mathcal{F}}) = \sum_i \mathcal{L}(\overline{f}, T, \overline{\mathcal{F}}_i) \tag{6.13}$$

with

$$\mathcal{L}(\overline{f}, T, \overline{\mathcal{F}}_i) = \sum_{\substack{\overline{b}_{i,m} \in \overline{\mathcal{F}}_i \\ |\langle \overline{f}, \overline{b}_{i,m} \rangle| < T}} |\langle \overline{f}, \overline{b}_{i,m} \rangle|^2 + M_{S,i} + M_{G,i} + M_{B,i}, \tag{6.14}$$

where $M_{S,i}$ is a proportion of the nodes used to specify the segmentation of Ω_i, $M_{G,i}$ is the number of parameters to specify the geometric flow in Ω_i, and $M_{B,i}$ is the number of bandelet coefficients $|\langle \overline{f}, \overline{b}_{i,m} \rangle| \geq T$. As any node is shared by the children of their four subtrees, $M_{S,i} = \sum_{j=0}^{\ell_i - 1} 4^{\ell_i - j}$, and one verifies that $\sum_i M_{S,i} = M_S$.

For each dyadic square Ω_i, the CART algorithm first computes the optimal geometric flow and the resulting best bandelet frame $\overline{\mathcal{F}}_i$ which yields a minimum value $\mathcal{L}(\overline{f}, T, \overline{\mathcal{F}}_i)$. Then a bottom-up CART optimization uses the additivity (6.13) to find the best dyadic partition $\cup_i \Omega_i$ of the image support $[0,1]^2$. When going up the tree, at each node corresponding to a square Ω_i, the minimum Lagrangian value $\mathcal{L}(\overline{f}, T, \overline{\mathcal{F}}_i)$ calculated on Ω_i is compared with an optimal sum of Lagrangian values corresponding to the four subsquares. The minimum of these two is kept and associated with Ω_i, with the corresponding best segmentation configuration. Appendix C proves that the complexity of this algorithm is polynomial in $\mathcal{O}(\epsilon^{-2} T^{-2p})$. The exponent p of this polynomial complexity algorithm depends upon the maximum degree of the polynomial flows. In the following, we shall suppose that p is equal to the number of vanishing moments p of the wavelet ψ that is used to construct the bandelet bases.

Despite the image discretization, the following theorem proves that the best bandelet frame computed with piecewise polynomial flows yields an approximation error that has the same optimal asymptotic decay as in Theorem 4.3 up to a logarithmic factor.

THEOREM 6.1. *Let f be a $\mathbf{C}^\alpha$ geometrically regular function and $\overline{f}$ be its discretization at a scale ϵ. For any $\gamma > 0$, there exists a constant C_γ that depends only upon the edge geometry and γ such that for any $T \geq \gamma \epsilon^{1/2}$ the discrete best bandelet frame in $\mathcal{D}_{T^2}$ yields an approximation $\overline{f}_M$ that satisfies*

$$\|\overline{f} - \overline{f}_M\|^2 \leq C_\gamma \, C_f^2 \, \max(1, \ell_{\mathcal{C}})^{\alpha+1} \, M^{-\alpha} (\log_2 M)^{\alpha+1} \tag{6.15}$$

with

$$M \leq C_\gamma \, C_f^{2/(\alpha+1)} \, \max(1, \ell_{\mathcal{C}}) \, T^{-2/(\alpha+1)} \, |\log_2 T|, \tag{6.16}$$

where $C_f = \max(\|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)} \max(\|c\|_{\mathbf{C}^\alpha}^\alpha, 1)^2, \|c\|_{\mathbf{C}^\alpha}^{(\alpha+1)/(2\alpha)}, \|\tilde{f}\|_{\mathbf{C}^1(\Lambda)}^{\alpha+1})$ and $\ell_{\mathcal{C}}$ is the total length of the curves.

One could note that if $T^{2\alpha/(\alpha+1)} \, |\log_2 T| \leq s$, a stronger result holds as the logarithmic factor disappears.

Proof of Theorem 6.1. To prove (6.15), it is sufficient to prove the existence of a discrete frame $\overline{\mathcal{F}}'$ such that

$$\mathcal{L}(\overline{f}, T, \overline{\mathcal{F}}') \leq C_\gamma \, C_f^{2/(\alpha+1)} \, \max(1, \ell_{\mathcal{C}}) \, T^{2\alpha/(\alpha+1)} \, |\log_2 T|. \tag{6.17}$$

Indeed, as in the proof of Theorem 4.3, the best frame $\overline{\mathcal{F}}$ thus also satisfies

$$\mathcal{L}(\overline{f}, T, \overline{\mathcal{F}}) \leq C_\gamma \, C_f^{2/(\alpha+1)} \, \max(1, \ell_{\mathcal{C}}) \, T^{2\alpha/(\alpha+1)} \, |\log_2 T|. \tag{6.18}$$

This implies

$$\|\overline{f} - \overline{f}_M\|^2 \leq C_\gamma \, C_f^{2/(\alpha+1)} \, \max(1, \ell_{\mathcal{C}}) \, T^{2\alpha/(\alpha+1)} \, |\log_2 T| \tag{6.19}$$

and

$$M \leq C_\gamma \, C_f^{2/(\alpha+1)} \, \max(1, \ell_{\mathcal{C}}) \, T^{-2/(\alpha+1)} \, |\log_2 T|. \tag{6.20}$$

Inserting (6.20) in (6.19) yields

$$\|\overline{f} - \overline{f}_M\|^2 \leq C_\gamma \, C_f^2 \, \max(1, \ell_{\mathcal{C}})^{\alpha+1} \, M^{-\alpha} \, (\log_2 M)^{\alpha+1}. \tag{6.21}$$

The proof of the existence of a bandelet frame $\overline{\mathcal{F}}'$ of discrete images that satisfies (6.17) relies on the existence of a bandelet frame $\mathcal{F}'$ of $\mathbf{L^2}[0,1]^2$ that satisfies a condition similar to (6.17), which is given by the following lemma.

LEMMA 6.2. *Under the hypotheses of Theorem* 6.1*, there exists a constant* C *that depends upon the edges such that, for any* T*, one can construct a bandelet frame* $\mathcal{F}'$ *of* $\mathbf{L^2}[0,1]^2$ *satisfying*

$$\mathcal{L}(f, T, \mathcal{F}') \leq C \, C_f^{2/(\alpha+1)} \, \max(1, \ell_{\mathcal{C}}) \, T^{2\alpha/(\alpha+1)} \, |\log_2 T| \tag{6.22}$$

with a number of edge rectangles bounded by $C \, \ell_{\mathcal{C}} \, T^{-2/(\alpha+1)}$.

The following lemma shows that the corresponding discrete frame $\overline{\mathcal{F}}'$ obtained with the same geometry is closely related to the continuous frame: both frames yield bandelet coefficients that are close.

LEMMA 6.3. *There exists a constant* C *that depends upon the edge geometry such that*

$$\sum_{i,m} |\langle \overline{f}, \overline{b}_{i,m} \rangle - \langle f, b_{i,m} \rangle|^2 \leq C \, (\|\tilde{f}\|_\infty \, \ell_{\mathcal{C}} + \|\tilde{f}\|_{\mathbf{C}^2(\Lambda)} \, \epsilon) \, \epsilon. \tag{6.23}$$

To prove (6.17) let us consider $J' = \{(i,m),\ |\langle f, b_{i,m} \rangle| \geq T\}$ and

$$\begin{aligned} \mathcal{L}(\overline{f}, T, \overline{\mathcal{F}}) &\leq \sum_{(i,m) \notin J'} |\langle \overline{f}, \overline{b}_{i,m} \rangle|^2 + T^2 \, \mathrm{Card}\,(J) + T^2 (M_S + M_G) \qquad (6.24) \\ &\leq 2 \left(\sum_{(i,m) \notin J'} |\langle f, b_{i,m} \rangle|^2 + T^2 \, \mathrm{Card}\,(J) + T^2 (M_S + M_G) \right. \\ &\qquad \left. + \sum_{(i,m) \notin J'} |\langle \overline{f}, \overline{b}_{i,m} \rangle - \langle f, b_{i,m} \rangle|^2 \right). \qquad (6.25) \end{aligned}$$

Using Lemma 3.10 and $J_T = \{(i,m),\, |\langle \overline{f}, \overline{b}_{i,m}\rangle| \geq T\}$, we get

$$\leq 2\left(\sum_{(i,m)\notin J_T} |\langle f, b_{i,m}\rangle|^2 + T^2 \operatorname{Card}(J_T) + T^2(M_S + M_G) + \sum_{(i,m)\notin J'} |\langle \overline{f}, \overline{b}_{i,m}\rangle - \langle f, b_{i,m}\rangle|^2\right), \tag{6.26}$$

$$\mathcal{L}(\overline{f}, T, \overline{\mathcal{F}}) \leq 2\left(\mathcal{L}(f, T, \mathcal{F}') + \sum_i \sum_n |\langle \overline{f}, \overline{b}_{i,m}\rangle - \langle f, b_{i,m}\rangle|^2\right). \tag{6.27}$$

Inserting (6.22) and (6.23) in (6.27) yields (6.17), as by hypothesis ϵ is majored by $\gamma^2 T^2$.

The proof of Lemma 6.2 itself is very similar to that of Lemma 4.4 up to the segmentation. The edge squares are first subdivided into smaller squares of size smaller than $\eta^{1/(2\alpha)}$ in order to avoid some too anisotropic structure. The resulting edge squares are further subdivided into rectangles of width smaller than $\max(\|c\|_{\mathbf{C}^\alpha}^{-1/\alpha}, 1)\,\eta^{1/\alpha}$ to allow the use of a polynomial flow. This subdivision yields the logarithmic term of the lemma: on each such rectangle, the number of required bandelet coefficients is of order $\log_2 T$, while the total number of these rectangles is of order $\ell_{\mathcal{C}}\eta^{-1/\alpha}$. □

The condition $T \geq \gamma\epsilon^{1/2}$ of Theorem 6.1 is a consequence of Lemma 6.3 which controls the differences between the discrete bandelet coefficients of the discrete image and the bandelet coefficients of the original analog image. From the discretization process (6.1) any linear reconstruction of the samples yields a square error with respect to f that can be bounded only by an order of ϵ. Indeed, f can be discontinuous along curves whose locations are unknown. Since we are using a linear warping operator, we cannot reduce this error.

Numerically, in the regular regions of f, the discrete bandelet scheme is improved if we replace the 0-order interpolation of (6.2) with a higher-order interpolation [14]. Unfortunately, this destroys the energy conservation properties of the bandelet basis proposed here, and it does not improve the asymptotic decay of the error because of the presence of discontinuities.

The choice of an L^2 normalized averaging for the discretization can be relaxed: the result holds indeed for any $\overline{f}[n_1, n_2] = f \star \phi_\epsilon(n_1\epsilon, n_2\epsilon)$ with ϕ_ϵ a local averaging function defined from a compactly supported function ϕ such that $\|\phi\|_2 = 1$ and $\|\phi\|_1 = 0$ by $\phi_\epsilon(x_1, x_2) = \epsilon^{-1}\phi(x_1, x_2)$. A possible choice for ϕ is thus a compactly supported scaling function associated with a wavelet [11].

This theorem provides a constructive approximation scheme with a polynomial complexity and a decay rate optimal up to a logarithmic factor. As in section 5, this implies a compression result. The same coding strategy can be used in this context, and the logarithmic factor of Theorem 6.1, which does not appear in Theorem 4.3, just modifies the exponent of the $|\log R|$ factor. Under the hypotheses of Theorem 6.1, we get

$$D(R) \leq C C_f^2 \max(1, \ell_{\mathcal{C}})^{\alpha+1} R^{-\alpha} |\log R|^{2\alpha+1}. \tag{6.28}$$

Hence, the distortion rate of the discrete bandelet coder reaches the Kolmogorov lower bound $R^{-\alpha}$ up to a logarithmic factor $|\log_2(R)|^{2\alpha+1}$.

As in Theorem 4.3, the choice of a different basis in each square region yields a fast algorithm to optimize the geometry, but this can create discontinuities at the boundaries of the region. An implementation of the bandelet transform in [14] overcomes this difficulty with an adapted lifting scheme, for which there is no proof of optimality.

Finally, although the algorithm is polynomial, it is still computationally intensive, and most of the bandelet algorithm implementation [14] replaces the full geometry exploration with a faster geometry exploration obtained from a geometry estimation similar to the one described in section 3.2. As long as the jumps of the discontinuities do not vanish to zero, the estimation remains precise enough to obtain a good geometry. The corresponding optimization algorithm yields an error decay of order $M^{-\alpha}$ with a low-order polynomial complexity.

Appendix A. Proofs of lemmas for Theorem 3.2.

A.1. Lemma 3.4: The outer bandelets.

Proof of Lemma 3.4. We saw in (3.5) and (3.10) that

$$\langle f, b_{l,j,m}\rangle = \langle f(x_1, x_2 + g(x_1)), \psi_{l,m_1}(x_1)\psi_{j,m_2}(x_2)\rangle. \tag{A.1}$$

Furthermore, since $f = \tilde{f} \star h$ and $\tilde{f}$ is uniformly $\mathbf{C}^\alpha$ over the convolution domain, as long as we remain away from the smoothed singularity,

$$\|f\|_{\mathbf{C}^\alpha((x_1,x_2))} \leq \|h\|_1 \, \|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)}, \tag{A.2}$$

where $\|f\|_{\mathbf{C}^\alpha((x_1,x_2))}$ is the Hölder norm of exponent α at the point (x_1, x_2). As $\|h\|_{\mathbf{C}^\alpha} \leq s^{-(2+\alpha)}$ implies $\|h\|_1 \leq 1$,

$$\|f\|_{\mathbf{C}^\alpha((x_1,x_2))} \leq \|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)}. \tag{A.3}$$

Now (A.3) yields along x_1

$$\left|\frac{\partial^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor} Wf}{\partial x_1^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}}(x_1', x_2) - \frac{\partial^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor} Wf}{\partial x_2^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}}(x_1, x_2)\right| \leq \|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)} \, \max(\|c\|^\alpha_{\mathbf{C}^\alpha}, 1)|x_1' - x_1|^{\alpha - \lfloor\!\lfloor \alpha \rfloor\!\rfloor} \tag{A.4}$$

and along x_2

$$\left|\frac{\partial^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor} Wf}{\partial x_2^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}}(x_1, x_2') - \frac{\partial^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor} Wf}{\partial x_2^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}}(x_1, x_2)\right| \leq \|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)} \, |x_2' - x_2|^{\alpha - \lfloor\!\lfloor \alpha \rfloor\!\rfloor}. \tag{A.5}$$

Using the vanishing moments of the wavelets along either x_1 or x_2 in (A.1) and since $l \geq j$,

$$|\langle f, b_{l,j,m}\rangle| \leq C \, \|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)} \, \max(\|c\|^\alpha_{\mathbf{C}^\alpha}, 1) \, 2^{j/2} \, 2^{(\alpha+1/2)l}, \tag{A.6}$$

$$|\langle f, b_{l,j,m}\rangle| \leq C \, C_f \, 2^{j/2} \, 2^{(\alpha+1/2)l}. \tag{A.7}$$

Indeed, assuming $b_{l,j,m}(x_1, x_2) = \psi_{l,n_1}(x_1)\psi_{j,n_2}(x_2 - g(x_1))$ (the case $b_{l,j,m}(x_1, x_2) = \psi_{l,n_1}(x_1)\phi_{j,n_2}(x_2 - g(x_1))$ is similar),

$$\langle f, b_{l,j,m}\rangle = \iint Wf(x_1, x_2)\psi_{l,n_1}(x_1)\psi_{j,n_2}(x_2 - g(x_1))\mathrm{d}x_1\mathrm{d}x_2 \tag{A.8}$$

with a sequence of integration by parts

$$= (-1)^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor} \iint \frac{\partial^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}}{\partial x_1^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}} W f(x_1, x_2) 2^{l\alpha} \psi_{l,n_1}^{[\lfloor\!\lfloor \alpha \rfloor\!\rfloor]}(x_1) \psi_{j,n_2}(x_2 - g(x_1)) \mathrm{d}x_1 \mathrm{d}x_2, \tag{A.9}$$

where $\psi^{[\lfloor\!\lfloor \alpha \rfloor\!\rfloor]}$ is the primitive of ψ of order $\lfloor\!\lfloor \alpha \rfloor\!\rfloor$ which still has a vanishing moment. We thus have for $x_1' = 2^l n_1$, which is in the support of the wavelet ψ_{l,n_1},

$$\begin{aligned} = (-1)^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor} \iint & \left(\frac{\partial^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}}{x_1^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}} W f(x_1, x_2) - \frac{\partial^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}}{\partial x_1^{\lfloor\!\lfloor \alpha \rfloor\!\rfloor}} W f(x_1', x_2) \right) \\ & \times 2^{l\alpha} \psi_{l,n_1}^{[\lfloor\!\lfloor \alpha \rfloor\!\rfloor]}(x_1) \psi_{j,n_2}(x_2 - g(x_1)) \mathrm{d}x_1 \mathrm{d}x_2, \end{aligned} \tag{A.10}$$

which gives the bound (A.6) when combined with (A.4) and $\|\psi_{l,n_1}^{[d]}\|_1 = 2^{\alpha/2} \|\psi^{[d]}\|_1$ for any d.

If we let $J = \{(l, j, m) : b_{l,j,m} \in \mathcal{B}_1\}$,

$$J' = \{(l, j, m) \in J \ : \ C\, C_f\, 2^{j/2}\, 2^{(\alpha+1/2)l} \geq T\}, \tag{A.11}$$

and

$$J_T = \{(l, j, m) \in J \ : \ |\langle f, b_{l,j,m} \rangle| \geq T\}, \tag{A.12}$$

we verify that $J_T \subset J'$.

Since there are at most $\max(\ell_1\, 2^{-l}, K) \times \max(\ell_2\, 2^{-j}, K)$ bandelets in $\mathcal{B}_1$ at the scale of index (l, j), where K is the size of the support of ψ, one can verify with a summation over l and j that

$$\mathrm{Card}\,(J') \leq \sum_{\substack{l,j \\ C\, C_f\, 2^{j/2}\, 2^{(\alpha+1/2)l} \geq T}} \max(\ell_1\, 2^{-l}, K) \times \max(\ell_2\, 2^{-j}, K) \tag{A.13}$$

and

$$\mathrm{Card}\,(J') \leq C \max(\ell_1\, \ell_2\, C_f^{2/(\alpha+1)}\, T^{-2/(\alpha+1)}, K^2), \tag{A.14}$$

so

$$\mathrm{Card}\,(J_T) \leq C \max(\ell_1\, \ell_2\, C_f^{2/(\alpha+1)}\, T^{-2/(\alpha+1)}, K^2). \tag{A.15}$$

Combining the bound on the number of coefficients with (A.7) yields for any scale of index (l, j)

$$\sum_m |\langle f, b_{l,j,m} \rangle|^2 \leq C\, \ell_1\, \ell_2\, C_f^2\, 2^{2\alpha l}. \tag{A.16}$$

Summing over l and j, it results that

$$\sum_{(l,j,m) \notin J'} |\langle f, b_{l,j,m} \rangle|^2 \leq \sum_{\substack{l,j \\ C\, C_f\, 2^{j/2}\, 2^{(\alpha+1/2)l} \leq T}} C\, \ell_1\, \ell_2\, C_f^2\, 2^{2\alpha l}, \tag{A.17}$$

$$\sum_{(l,j,m) \notin J'} |\langle f, b_{l,j,m} \rangle|^2 \leq C\, \ell_1\, \ell_2\, C_f^{2/(\alpha+1)}\, T^{2\alpha/(\alpha+1)}. \tag{A.18}$$

Now

$$\sum_{(l,j,m)\notin J_T} |\langle f, b_{l,j,m}\rangle|^2 = \sum_{(l,j,m)\notin J'} |\langle f, b_{l,j,m}\rangle|^2 + \sum_{(l,j,m)\in J'\setminus J_T} |\langle f, b_{l,j,m}\rangle|^2 \tag{A.19}$$

and, as $(l,j,m) \in J' \setminus J_T$ implies $|\langle f, b_{l,j,m}\rangle| \leq T$,

$$\leq \sum_{(l,j,m)\notin J'} |\langle f, b_{l,j,m}\rangle|^2 + \mathrm{Card}\,(J')\,T^2. \tag{A.20}$$

Inserting (A.13) and (A.17) yields

$$\sum_{(l,j,m)\notin J_T} |\langle f, b_{l,j,m}\rangle|^2 \leq C\,\ell_1\,\ell_2\,C_f^{2/(\alpha+1)}\,T^{2\alpha/(\alpha+1)}. \tag{A.21}$$ □

A.2. Lemma 3.5: The inner bandelets.

Proof of Lemma 3.5. Let $J = \{(l,j,m) : b_{l,j,m} \in \mathcal{B}_2\}$ and $J_T = \{(l,j,m) \in J : |\langle f, b_{l,j,m}\rangle| \geq T\}$.

In this proof, the bandelets $b_{l,j,m} \in \mathcal{B}_2$ are separated according to the scale 2^j in four categories: $2^j \leq 2^{j_\star}$ with

$$2^{j_\star} = C\,C_f^{-1/(\alpha+1)}\,s\,(T^{2\alpha/(\alpha+1)}s^{-1})^{1/(2\alpha)}, \tag{A.22}$$

$2^{j_\star} \leq 2^j \leq s$, $s \leq 2^j \leq s^{1/\alpha}$, and $s^{1/\alpha} < 2^j$.

We first prove that the total energy of the bandelet coefficients $\langle f, b_{l,j,m}\rangle$ with $j \leq j_\star$ is small:

$$\sum_{\substack{(l,j,m)\in J\\ j\leq j_\star}} |\langle f, b_{l,j,m}\rangle|^2 \leq C\,C_f^{2/(\alpha+1)}\,\ell_1\,T^{2\alpha/(\alpha+1)}. \tag{A.23}$$

On one hand, as J_T is a subset of J, (A.23) implies

$$\sum_{\substack{(l,j,m)\notin J_T\\ j\leq j_\star}} |\langle f, b_{l,j,m}\rangle|^2 \leq C\,C_f^{2/(\alpha+1)}\,\ell_1\,T^{2\alpha/(\alpha+1)}. \tag{A.24}$$

On the other hand, as $(l,j,m) \in J_T$ implies $|\langle f, b_{l,j,m}\rangle| \geq T$, (A.23) also implies

$$\mathrm{Card}\,(\{(l,j,m) \in J_T,\, j \leq j_\star\}) \leq C\,C_f^{2/(\alpha+1)}\,\ell_1\,T^{-2/(\alpha+1)}. \tag{A.25}$$

To prove (A.23), we use that the bandelets $b_{l,j,m}$ with $j \leq j_\star$ are obtained with an orthogonal change of bases from warped wavelets of scale $2^j \leq 2^{j_\star}$ as described in section 4.1, so the energy of the corresponding bandelet coefficients is bounded by the one of the wavelet coefficients. Let $\{\Psi^d_{j,m}\}$ be the wavelet basis of B, where $\Psi^d_{j,m}$ stands for $\phi_{j,m_1}(x_1)\psi_{j,m_2}(x_2)$, $\psi_{j,m_1}(x_1)\phi_{j,m_2}(x_2)$, $\psi_{j,m_1}(x_1)\psi_{j,m_2}(x_2)$ for $d \in \{1,2,3\}$, respectively, and define

$$J_\star = \{(j,m)\ :\ j \leq j_\star\}. \tag{A.26}$$

As the space generated by $\{b_{l,j,m} : j \leq j_\star\}$ is included in the space generated by $\{\Psi^d_{j,m} : j \leq j_\star\}$, one verifies that

$$\sum_{(l,j,m),j\leq j_\star} |\langle f, b_{l,j,m}\rangle|^2 \leq \sum_{(j,m)\in J_\star} |\langle f, W^{-1}\Psi^d_{j,m}\rangle|^2. \tag{A.27}$$

To prove (A.23), it is thus sufficient to prove that

$$\sum_{(j,m)\in J_\star} |\langle f, W^{-1}\Psi^d_{j,m}\rangle|^2 \le C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{2\alpha/(\alpha+1)} \tag{A.28}$$

or, equivalently, as $\langle f, W^{-1}\Psi^d_{j,m}\rangle = \langle Wf, \Psi^d_{j,m}\rangle$,

$$\sum_{(j,m)\in J_\star} |\langle Wf, \Psi^d_{j,m}\rangle|^2 \le C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{2\alpha/(\alpha+1)}. \tag{A.29}$$

In B, one can verify that Wf has the same regularity as f. Three different kinds of wavelets are distinguished in (A.29):

- The wavelets that do not intersect the smoothed singularities: there are at most $\ell_1\,\ell_2\,2^{-2j}$ such wavelets at the scale 2^j, and the bounds (A.4) and (A.5) on Wf imply

$$|\langle Wf, \Psi^d_{j,m}\rangle| \le C\, \|\tilde f\|_{\mathbf{C}^\alpha(\Lambda)}\, \max(\|c\|^\alpha_{\mathbf{C}^\alpha}, 1)\, 2^{(\alpha+1)j}, \tag{A.30}$$

$$|\langle Wf, \Psi^d_{j,m}\rangle| \le C\, C_f\, 2^{(\alpha+1)j}. \tag{A.31}$$

- The wavelets that do intersect the singularities with a scale $2^j \ge s$: there are at most $C\,\ell_1\,2^{-j}$ such wavelets at the scale 2^j, and, using $\|Wf\|_\infty \le \|f\|_\infty \le \|\tilde f\|_\infty$, one verifies

$$|\langle Wf, \Psi^d_{j,m}\rangle| \le C\, \|\tilde f\|_\infty\, 2^j, \tag{A.32}$$

$$|\langle Wf, \Psi^d_{j,m}\rangle| \le C\, C_f\, 2^j. \tag{A.33}$$

- The wavelets that do intersect the singularities with a scale $2^j \le s$: there are at most $C\,\ell_1\,s\,2^{-2j}$ such wavelets at the scale 2^j and, as $f = \tilde f \star h$ and h is $\mathbf{C}^\alpha$, we get

$$\|f\|_{\mathbf{C}^\alpha(\Omega_i)} \le \|\tilde f\|_\infty\, s^2\, \|h\|_{\mathbf{C}^\alpha} \le C_d\, \|\tilde f\|_\infty\, s^{-\alpha}, \tag{A.34}$$

  which yields

$$|\langle Wf, \Psi^d_{j,m}\rangle| \le C\, C_f\, s^{-\alpha}\, 2^{(\alpha+1)j}. \tag{A.35}$$

Combining (A.31), (A.33), and (A.35) with the respective bounds on the number of coefficients gives bounds on the energy of the coefficients that eventually yield (A.29).

The remaining bandelets are the ones that intersect the smoothed singularities at a scale $j \ge j_\star$, and the corresponding coefficients are controlled with the regularity of the geometry.

As stated in the proof of Theorem 3.2, Lemma 3.7 implies when $|x_1' - x_1| \le K s^{1/\alpha}$

$$\left|\frac{\partial^{\lfloor\alpha\rfloor} Wf}{\partial x_1^{\lfloor\alpha\rfloor}}(x_1', x_2) - \frac{\partial^{\lfloor\alpha\rfloor} Wf}{\partial x_1^{\lfloor\alpha\rfloor}}(x_1', x_2)\right| \le C\, \|\tilde f\|_{\mathbf{C}^\alpha(\Lambda)}\, \max(\|c\|^\alpha_{\mathbf{C}^\alpha}, 1) \times \max(\|c\|^\alpha_{\mathbf{C}^\alpha}, C_d^\alpha, 1)^2\, s^{-1} |x_1' - x_1|^{\alpha - \lfloor\alpha\rfloor}, \tag{A.36}$$

$$\left|\frac{\partial^{\lfloor\alpha\rfloor} Wf}{\partial x_1^{\lfloor\alpha\rfloor}}(x_1', x_2) - \frac{\partial^{\lfloor\alpha\rfloor} Wf}{\partial x_1^{\lfloor\alpha\rfloor}}(x_1', x_2)\right| \le C\, C_f\, s^{-1} |x_1' - x_1|^{\alpha - \lfloor\alpha\rfloor}. \tag{A.37}$$

Using the vanishing moment of the bandelets and the size of their support, (A.37) implies

$$|\langle f(x_1,x_2), b_{l,j,m}\rangle| \le C\, C_f\, s^{-1}\, 2^{(\alpha+1/2)l}\, 2^{j/2}, \tag{A.38}$$

which is sufficient for $s \ge 2^j \ge 2^{j_\star}$.

Let $J_- = \{(l,j,m) \in J : s \ge 2^j \ge 2^{j_\star},\, C\, C_f\, s^{-1}\, 2^{(\alpha+1/2)l}\, 2^{j/2} \ge T\}$. The number of bandelets is $\max(\ell_1\, 2^{-l}, K)\, \max(s\, 2^{-j}, K)$ at each scale, and summing over l and j yields

$$\mathrm{Card}\,(J_-) \le \max(C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{-2/(\alpha+1)}, |j_\star|). \tag{A.39}$$

Using (A.38), we obtain for each scale of index (l,j)

$$\sum_m |\langle f, b_{l,j,m}\rangle|^2 \le C\, \ell_1\, C_f^2\, s^{-1}\, 2^{2\alpha l} \tag{A.40}$$

and thus by a summation over l and j

$$\sum_{(l,j,m)\notin J_-,\, s\ge 2^j\ge 2^{j_\star}} |\langle f, b_{l,j,m}\rangle|^2 \le C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{2\alpha/(\alpha+1)}. \tag{A.41}$$

As, if $s \ge 2^j \ge 2^{j_\star}$, $(l,j,m) \notin J_T$ implies $(l,j,m) \notin J$ or $(l,j,m) \in J$ and $|\langle f, b_{l,j,m}\rangle| \le T$, combining (A.39) and (A.41) implies

$$\begin{aligned}\sum_{(l,j,m)\notin J_T,\, s\ge 2^j\ge 2^{j_\star}} |\langle f, b_{l,j,m}\rangle|^2 \le & \sum_{(l,j,m)\notin J,\, s\ge 2^j\ge 2^{j_\star}} |\langle f, b_{l,j,m}\rangle|^2 \\ & + \sum_{(l,j,m)\in J\setminus J_T,\, s\ge 2^j\ge 2^{j_\star}} |\langle f, b_{l,j,m}\rangle|^2,\end{aligned} \tag{A.42}$$

$$\sum_{(l,j,m)\notin J_T,\, s\ge 2^j\ge 2^{j_\star}} |\langle f, b_{l,j,m}\rangle|^2 \le C\, C_f^{2/(\alpha+1)}\, \ell_1\, T^{2\alpha/(\alpha+1)}, \tag{A.43}$$

and, as $(l,j,m) \in J_T$ implies $|\langle f, b_{l,j,m}\rangle| \ge T$ and $(i,j,m) \in J_-$,

$$\mathrm{Card}\left(\{(l,j,m) \in J_T,\, s \ge 2^j \ge 2^{j_\star}\}\right) \le C\, \max(C_f^{2/(\alpha+1)}\, \ell_1\, T^{-2/(\alpha+1)}, |j_\star|). \tag{A.44}$$

Away from the smoothed singularity, f is regular, and we still have when $|x_1' - x_1| \le K s^{1/\alpha}$

$$\left|\frac{\partial^{\mathbb{L}\alpha\mathbb{J}} Wf}{\partial x_1^{\mathbb{L}\alpha\mathbb{J}}}(x_1', x_2) - \frac{\partial^{\mathbb{L}\alpha\mathbb{J}} Wf}{\partial x_1^{\mathbb{L}\alpha\mathbb{J}}}(x_1', x_2)\right| \le C\|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)}\, \max(\|c\|_{\mathbf{C}^\alpha}^{\alpha}, 1)|x_1' - x_1|^{\alpha - \mathbb{L}\alpha\mathbb{J}}, \tag{A.45}$$

$$\left|\frac{\partial^{\mathbb{L}\alpha\mathbb{J}} Wf}{\partial x_1^{\mathbb{L}\alpha\mathbb{J}}}(x_1', x_2) - \frac{\partial^{\mathbb{L}\alpha\mathbb{J}} Wf}{\partial x_1^{\mathbb{L}\alpha\mathbb{J}}}(x_1', x_2)\right| \le C\, C_f\, |x_1' - x_1|^{\alpha - \mathbb{L}\alpha\mathbb{J}}. \tag{A.46}$$

Combining this bound and (A.37) with the definition of the bandelets solves the third case.

Indeed,

(A.47)
$$\langle f(x_1,x_2), b_{l,j,m}\rangle = \int_{x_1}\int_{x_2} f(x_1, x_2+g(x_1))\psi_{l,m_1}(x_1)\mathrm{d}x_1\psi_{j,m_2}(x_2)\mathrm{d}x_2,$$

so, splitting the integral depending on the distance along x_2 from the curve,

(A.48)
$$\begin{aligned}|\langle f(x_1,x_2), b_{l,j,m}\rangle| \le &\int_{x_2,|x_2|\le 4\,s}\left|\int_{x_1} f(x_1,x_2+g(x_1))\psi_{l,m_1}(x_1)\mathrm{d}x_1\right||\psi_{j,m_2}(x_2)|\,\mathrm{d}x_2\\ &+\int_{x_2,|x_2|\ge 4\,s}\left|\int_{x_1} f(x_1,x_2+g(x_1))\psi_{l,m_1}(x_1)\mathrm{d}x_1\right||\psi_{j,m_2}(x_2)|\,\mathrm{d}x_2\end{aligned}$$

using, respectively, (A.37) and (A.46) and the vanishing moments of the wavelets

(A.49)
$$\begin{aligned}&\le \int_{x_1,|x_1|\ge 4\,s} C\,C_f 2^{(\alpha+1/2)l}\,|\psi_{j,m_2}(x_2)|\mathrm{d}x_2\\ &\quad+\int_{x_2,|x_2|\le 4\,s} C\,C_f\,s^{-1}2^{\,(\alpha+1/2)l}|\psi_{j,m_2}(x_2)|\mathrm{d}x_2,\end{aligned}$$

(A.50)
$$|\langle f(x_1,x_2), b_{l,j,m}\rangle| \le C\,C_f\,2^{(\alpha+1/2)l}2^{-j/2}.$$

Let $J_+ = \{(l,j,m)\in J : s^{1/\alpha}\ge 2^j\ge s,\ C\,C_f\,2^{(\alpha+1/2)l}\,2^{-j/2}\ge T\}$. As the number of inner bandelets is $\max(\ell_1\,2^{-l}, K)$ at each scale, where K is the size of the support of ψ, a summation over l and j gives

$$\mathrm{Card}\,(J_+) \le \max(C\,C_f^{2/(\alpha+1)}\,\ell_1\,T^{-2/(\alpha+1)}, K\,|\log_2 s|). \tag{A.51}$$

With (A.50), a bound on the energy of the coefficients at a given scale is obtained. Summing over l and j yields

$$\sum_{(l,j,m)\notin J_+,\,2^j\ge s} |\langle f(x_1,x_2), b_{l,j,m}\rangle|^2 \le C\,C_f^{2/(\alpha+1)}\,\ell_1\,T^{-2\alpha/(\alpha+1)}. \tag{A.52}$$

As, if $2^j\ge s$, $(l,j,m)\notin J_T$ implies $(l,j,m)\notin J_+$ or $(l,j,m)\in J_+$ and $|\langle f, b_{l,j,m}\rangle|\le T$, we verify that this implies

$$\sum_{(l,j,m)\notin J_T,\,2^j\ge s} |\langle f, b_{l,j,m}\rangle|^2 \le C\,C_f^{2/(\alpha+1)}\,\ell_1\,T^{2\alpha/(\alpha+1)} \tag{A.53}$$

and, as $(l,j,m)\in J_T$ implies $|\langle f, b_{l,j,m}\rangle|\ge T$,

$$\mathrm{Card}\left(\{(l,j,m)\in J_T, 2^j\ge s\}\right) \le C\,C_f^{2/(\alpha+1)}\,\ell_1\,T^{-2/(\alpha+1)}. \tag{A.54}$$

Finally, there are fewer than $\max(l_1 s^{-1/\alpha}, K)$ bandelets above the scale $s^{1/\alpha}$, so combining (A.24), (A.43), and (A.53) as well as (A.25), (A.44), and (A.54) finishes the proof. □

A.3. Lemma 3.6: Regularity of the approximated curve.

Proof of Lemma 3.6. Given any geometry g that is closer than s from a $\mathbf{C}^\alpha$ curve c, we will show that a suitable projection of g will satisfy all the conditions. Now let π_{x_0} be the Taylor polynomial of order $\lfloor\!\lfloor \alpha \rfloor\!\rfloor$ of c at x_0,

$$c - P_{\mathbf{V}_k} g = (c - P_{\mathbf{V}_k}(\pi_{x_0})) + (P_{\mathbf{V}_k}(\pi_{x_0}) - P_{\mathbf{V}_k}(c)) + (P_{\mathbf{V}_k}(c) - P_{\mathbf{V}_k}(g)), \tag{A.55}$$

so for the derivatives at x_0

$$\begin{aligned}(c - P_{\mathbf{V}_k} g)^{(\beta)}(x_0) = (c - P_{\mathbf{V}_k}(\pi_{x_0}))^{(\beta)}(x_0) + (P_{\mathbf{V}_k}(\pi_{x_0}) - P_{\mathbf{V}_k}(c))^{(\beta)}(x_0) \\ + (P_{\mathbf{V}_k}(c) - P_{\mathbf{V}_k}(g))^{(\beta)}(x_0).\end{aligned} \tag{A.56}$$

By the polynomial reproduction property of the θ, $P_{\mathbf{V}_k}\pi_{x_0} = \pi_{x_0}$, so, using the definition of π_{x_0},

$$\forall \beta \leq \lfloor\!\lfloor \alpha \rfloor\!\rfloor, \quad (c - P_{\mathbf{V}_k}(\pi_{x_0}))^{(\beta)}(x_0) = 0 \tag{A.57}$$

and with the regularity of c

$$|(c - P_{\mathbf{V}_k}(\pi_{x_0}))^{(\lfloor\!\lfloor \alpha \rfloor\!\rfloor)}(x) - (c - P_{\mathbf{V}_k}(\pi_{x_0}))^{(\lfloor\!\lfloor \alpha \rfloor\!\rfloor)}(x_0)| \leq \|c\|_{C^\alpha} |x - x_0|^{\alpha - \lfloor\!\lfloor \alpha \rfloor\!\rfloor}. \tag{A.58}$$

The two remaining terms of (A.56) are derivatives of a projection on $\mathbf{V}_k$ and give similar bounds. Indeed, for any u,

$$(P_{\mathbf{V}_k} u)^{(\beta)}(x_0) = \sum_k \langle u, \theta_{k,n} \rangle \theta_{k,n}^{(\beta)} \tag{A.59}$$

and, as the $\theta_{k,n}$ have a support of size $K2^k$,

$$(P_{\mathbf{V}_k} u)^{(\beta)}(x_0) = \sum_{|n2^k - x_0| \leq K2^k} \langle u, \theta_{k,n} \rangle \, \theta_{k,n}^{(\beta)}, \tag{A.60}$$

and

$$|(P_{\mathbf{V}_k} u)^{(\beta)}(x_0)| \leq 2K \max_{|n2^k - x_0| \leq K2^k} |\langle u, \theta_{k,n} \rangle| \, \|\theta_{k,n}^{(\beta)}\|_\infty, \tag{A.61}$$

so, as $\|\theta_{k,n}^{(\beta)}\|_\infty \leq C\, 2^{-k\beta}$,

$$|(P_{\mathbf{V}_k} u)^{(\beta)}(x_0)| \leq C\, 2^{-k\beta} \max_{|n2^k - x_0| \leq K2^k} |\langle u, \theta_{k,n} \rangle|. \tag{A.62}$$

For $\beta = \lfloor\!\lfloor \alpha \rfloor\!\rfloor$, the $\mathbf{C}^\alpha$ regularity of θ yields, as long as $|x - x_0| \leq K2^k$,

$$|(P_{\mathbf{V}_k} u)^{(\lfloor\!\lfloor \alpha \rfloor\!\rfloor)}(x) - (P_{\mathbf{V}_k} u)^{(\lfloor\!\lfloor \alpha \rfloor\!\rfloor)}(x_0)| \leq C\, 2^{-k\alpha} \max_{|n2^k - x_0| \leq 2K2^k} |\langle u, \theta_{k,n} \rangle| |x - x_0|^{\alpha - \lfloor\!\lfloor \alpha \rfloor\!\rfloor}. \tag{A.63}$$

By definition of the Taylor polynomial, $|\pi_{x_0}(x) - c(x)| \leq \|c\|_{\mathbf{C}^\alpha} |x - x_0|^\alpha$, so

$$\max_{|n2^k - x_0| \leq 2K2^k} |\langle \pi_{x_0} - c, \theta_{k,n} \rangle| \leq C \, \|c\|_{\mathbf{C}^\alpha} \, 2^{k\alpha}. \tag{A.64}$$

Inserting (A.64) in (A.62) with $u = \pi_{x_0} - c$ yields

$$|(P_{\mathbf{V}_k}\pi_{x_0} - P_{\mathbf{V}_k}c)^{(\beta)}(x_0)| \leq C\,\|c\|_{\mathbf{C}^\alpha}\,2^{-k\beta}\,2^{k\alpha} \tag{A.65}$$

and for all x, $|x - x_0| \leq K2^k$,

$$|(P_{\mathbf{V}_k}\pi_{x_0} - P_{\mathbf{V}_k}c)^{(\lfloor\alpha\rfloor)}(x) - (P_{\mathbf{V}_k}\pi_{x_0} - P_{\mathbf{V}_k}c)^{(\lfloor\alpha\rfloor)}(x_0)| \leq C\,\|c\|_{\mathbf{C}^\alpha}\,|x - x_0|^{\alpha-\lfloor\alpha\rfloor}. \tag{A.66}$$

By hypothesis $\|c - g\|_\infty \leq C_d\,s$, so $|\langle c - g, \theta_{k,n}\rangle| \leq C\,C_d\,s$, and inserting this bound in (A.62) with $u = c - g$ yields

$$|(P_{\mathbf{V}_k}c - P_{\mathbf{V}_k}g)^{(\beta)}(x_0)| \leq C\,C_d\,s\,2^{-k\beta} \tag{A.67}$$

and for all x, $|x - x_0| \leq K2^k$,

$$|(P_{\mathbf{V}_k}c - P_{\mathbf{V}_k}g)^{(\lfloor\alpha\rfloor)}(x) - (P_{\mathbf{V}_k}c - P_{\mathbf{V}_k}g)^{(\lfloor\alpha\rfloor)}(x_0)| \leq C\,s\,2^{-k\alpha}\,|x - x_0|^{\alpha-\lfloor\alpha\rfloor}. \tag{A.68}$$

As $2^k = \max(\|c\|_{\mathbf{C}^\alpha}^{-1/\alpha}, 1)\,s^{1/\alpha}$, combining (A.57), (A.65), and (A.67) as well as (A.58), (A.66), and (A.68) concludes the proof. □

A.4. Lemma 3.7: Regularity of the warped function.

Proof of Lemma 3.7. As h is α differentiable, so is $f = \tilde{f} \star h$. Furthermore,

$$f(x_1, x_2 + g(x_1)) = \int_u \tilde{f}(x_1 - u_1, x_2 + g(x_1) - u_2)h(u_1, u_2)\mathrm{d}u. \tag{A.69}$$

Now for each u_1, an acceptable variable change replaces u_2 by $u_2 + g(x_1) - c(x_1 - u_1)$:

$$f(x_1, x_2 + g(x_1)) = \int_u \tilde{f}(x_1 - u_1, x_2 + c(x_1 - u_1) - u_2)h(u_1, u_2 + g(x_1) - c(x_1 - u_1))\mathrm{d}u. \tag{A.70}$$

As $f(x_1 - u_1, x_2 + c(x_1 - u_1) - u_2)$ is by hypothesis α times differentiable along x_1, the inner part of the integral is thus differentiable and

$$\begin{aligned}&\left|\frac{\partial^a}{\partial x_1^a}(\tilde{f}(x_1 - u_1, x_2 + c(x_1 - u_1) - u_2)h(u_1, u_2 + g(x_1) - c(x_1 - u_1)))\right| \\ &\quad= \left|\sum_{d=0}^{a}\binom{d}{a}\frac{\partial^{a-d}}{\partial x_1^{a-d}}(\tilde{f}(x_1 - u_1, x_2 + c(x_1 - u_1) - u_2))\right. \\ &\qquad\left.\times\,\frac{\partial^d}{\partial x_1^d}(h(u_1, u_2 + g(x_1) - c(x_1 - u_1)))\right|.\end{aligned} \tag{A.71}$$

As $\tilde{f}(x_1 - u_1, x_2 + c(x_1 - u_1) - u_2)$ is $\mathbf{C}^\alpha$ with a constant bounded by $\|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)}\max(\|c\|_{\mathbf{C}^\alpha}^\alpha, 1)$, the first factor in each term of the sum can be controlled by

$$\left|\frac{\partial^{a-d}}{\partial x_1^{a-d}}(\tilde{f}(x_1 - u_1, x_2 + c(x_1 - u_1) - u_2))\right| \leq C\,\|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)}\,\max(\|c\|_{\mathbf{C}^\alpha}^\alpha, 1) \tag{A.72}$$

and for any x_1'

(A.73)
$$\left|\frac{\partial^{\lfloor\!\lfloor\alpha\rfloor\!\rfloor}}{\partial x_1^{\lfloor\!\lfloor\alpha\rfloor\!\rfloor}}(\tilde f(x_1'-u_1,x_2+c(x_1'-u_1)-u_2))-\frac{\partial^{\lfloor\!\lfloor\alpha\rfloor\!\rfloor}}{\partial x_1^{\lfloor\!\lfloor\alpha\rfloor\!\rfloor}}(\tilde f(x_1-u_1,x_2+c(x_1-u_1)-u_2))\right| \le C\,\|\tilde f\|_{\mathbf{C}^\alpha(\Lambda)}\,\max(\|c\|_{\mathbf{C}^\alpha}^{\alpha},1)|x_1'-x_1|^{\alpha-\lfloor\!\lfloor\alpha\rfloor\!\rfloor}.$$

The second factor is bounded with the help of the Faa di Bruno formula that gives the derivatives of $h(u_1,u_2+g(x_1)-c(x_1-u_1))$ seen as a composed function:

(A.74)
$$\begin{aligned}&\left|\frac{\partial^d}{\partial x_1^d}(h(u_1,u_2+g(x_1)-c(x_1-u_1)))\right|\\ &\quad=\left|\sum_{k_1+2k_2+\cdots+dk_d=d}\frac{d!}{k_1!\cdots k_d!}\frac{\partial^k}{\partial x_1^k}(h(u_1,u_2+g(x_1)-c(x_1-u_1)))\right.\\ &\qquad\left.\times\left(\frac{g^{(1)}(x_1)-c^{(1)}(x_1-u_1)}{1!}\right)^{k_1}\cdots\left(\frac{g^{(d)}(x_1)-c^{(d)}(x_1-u_1)}{d!}\right)^{k_d}\right|\end{aligned}$$

with $k=k_1+k_2+\cdots+k_d$.

The regularity of both c and $g-c$ and the small support of h implies, as $|u_1|\le s$,

$$\begin{aligned}&\left|g^{(d)}(x_1)-c^{(d)}(x_1-u_1)\right|\\ &\qquad\le\left|g^{(d)}(x_1)-c^{(d)}(x_1)\right|+\left|c^{(d)}(x_1)-c^{(d)}(x_1-u_1)\right| \qquad\text{(A.75)}\\ &\qquad\le\begin{cases}C\,\max(\|c\|_{\mathbf{C}^\alpha},C_d,1)\,s^{1-d/\alpha}+\|c\|_{\mathbf{C}^\alpha}\,s & \text{if } d<\lfloor\!\lfloor\alpha\rfloor\!\rfloor,\\ C\,\max(\|c\|_{\mathbf{C}^\alpha},C_d,1)s^{1-\lfloor\!\lfloor\alpha\rfloor\!\rfloor/\alpha}+2\|c\|_{\mathbf{C}^\alpha}s^{\alpha-\lfloor\!\lfloor\alpha\rfloor\!\rfloor} & \text{if } d=\lfloor\!\lfloor\alpha\rfloor\!\rfloor,\end{cases} \qquad\text{(A.76)}\end{aligned}$$

so

$$\left|g^{(d)}(x_1)-c^{(d)}(x_1-u_1)\right|\le C\,\max(\|c\|_{\mathbf{C}^\alpha},C_d,1)\,s^{1-d/\alpha}. \qquad\text{(A.77)}$$

Futhermore, one derives for any x_1', $|x_1'-x_1|\le Ks^{1/\alpha}$,

$$\begin{aligned}&\left|(g^{(\lfloor\!\lfloor\alpha\rfloor\!\rfloor)}(x_1')-c^{(\lfloor\!\lfloor\alpha\rfloor\!\rfloor)}(x_1'-u_1))-(g^{(\lfloor\!\lfloor\alpha\rfloor\!\rfloor)}(x_1)-c^{(\lfloor\!\lfloor\alpha\rfloor\!\rfloor)}(x_1-u_1))\right|\\ &\qquad\le C\,\max(\|c\|_{\mathbf{C}^\alpha},C_d,1)\,|x_1'-x_1|^{\alpha-\lfloor\!\lfloor\alpha\rfloor\!\rfloor}.\end{aligned} \qquad\text{(A.78)}$$

Now h itself is $\mathbf{C}^\alpha$ with $\|h\|_{\mathcal{C}^\alpha}\le s^{-(2+\alpha)}$, so

$$\left|\frac{\partial^k}{\partial x_1^k}h(u_1,u_2+x_1)\right|\le s^{-2}\,s^{-k} \qquad\text{(A.79)}$$

and

$$\left|\frac{\partial^{\lfloor\!\lfloor\alpha\rfloor\!\rfloor}}{\partial x_1^{\lfloor\!\lfloor\alpha\rfloor\!\rfloor}}h(u_1,u_2+x_1')-\frac{\partial^{\lfloor\!\lfloor\alpha\rfloor\!\rfloor}}{\partial x_1^{\lfloor\!\lfloor\alpha\rfloor\!\rfloor}}h(u_1,u_2+x_1)\right|\le s^{-2}\,s^{-\alpha}|x_1'-x_1|^{\alpha-\lfloor\!\lfloor\alpha\rfloor\!\rfloor}. \qquad\text{(A.80)}$$

Combining the bounds (A.77) and (A.79) as well as the bounds (A.78) and (A.80) with (A.74) yields after some calculations

(A.81)
$$\left|\frac{\partial^d}{\partial x_1^d} h(u_1, u_2 + g(x_1) + c(x_2 - u_2))\right| \leq C\, s^{-2}\, \max(\|c\|^{\alpha}_{\mathbf{C}^\alpha}, C_d^\alpha, 1)\, s^{-d/\alpha}$$

and for any x_1', $|x_1' - x_1| \leq K s^{1/\alpha}$,

(A.82)
$$\left|\frac{\partial^d}{\partial x_1^d} h(u_1, u_2 + g(x_1') + c(x_2 - u_2)) - \frac{\partial^d}{\partial x_1^d} h(u_1, u_2 + g(x_1) + c(x_2 - u_2))\right| \leq C\, s^{-2}\, \max(\|c\|^{\alpha}_{\mathbf{C}^\alpha}, C_d^\alpha, 1)\, s^{-1} |x_1' - x_1|^{\alpha - \lfloor\!\lfloor \alpha \rfloor\!\rfloor}.$$

Inserting (A.72) and (A.81) in (A.71), we obtain

(A.83)
$$\left|\frac{\partial^a}{\partial x_1^a}\left(\tilde{f}(x_1 - u_1, x_2 + c(x_1 - u_1) - u_2) h(u_1, u_2 + g(x_1) - c(x_1 - u_1))\right)\right| \leq C\, \max(\|c\|^{\alpha}_{\mathbf{C}^\alpha}, C_d^\alpha, 1)\, s^{-2}\, s^{-a/\alpha}$$

and with (A.73) and (A.82) for any x_1', $|x_1' - x_1| \leq K s^{1/\alpha}$,

(A.84)
$$\begin{aligned}&\left|\frac{\partial^a}{\partial x_1^a}\left(\tilde{f}(x_1' - u_1, x_2 + c(x_1' - u_1) - u_2) h(u_1, u_2 + g(x_1') - c(x_1 - u_1))\right)\right.\\ &\quad\left. - \frac{\partial^a}{\partial x_1^a}\left(\tilde{f}(x_1 - u_1, x_2 + c(x_1 - u_1) - u_2) h(u_1, u_2 + g(x_1) - c(x_1 - u_1))\right)\right|\\ &\quad\leq C\, \max(\|c\|^{\alpha}_{\mathbf{C}^\alpha}, C_d^\alpha, 1)\, s^{-2}\, s^{-1}\, |x_1' - x_1|^{\alpha - \lfloor\!\lfloor \alpha \rfloor\!\rfloor}.\end{aligned}$$

Now

(A.85)
$$\begin{aligned}\frac{\partial^a}{\partial x_1^a} f(x_1, x_2 + g(x_1)) = \int_u \frac{\partial^a}{\partial x_1^a}\Big(&\tilde{f}(x_1 - u_1, x_2 + c(x_1 - u_1) - u_2)\\ &\times h(u_1, u_2 + g(x_1) - c(x_1 - u_1))\Big)\mathrm{d}u,\end{aligned}$$

so the bounds (A.83) and (A.84) combined with the finite support of h of size s^2 conclude the proof. ∎

A.5. Lemma 3.9: Smoothing effect.

Proof of Lemma 3.9. By construction,

(A.86)
$$\|h\|_1 \tilde{f}(x) - f(x) = \|h\|_1 \tilde{f}(x) - \int \tilde{f}(x - u) h(u) \mathrm{d}u$$

(A.87)
$$= \int (\tilde{f}(x) - \tilde{f}(x - u)) h(u) \mathrm{d}u,$$

so

(A.88)
$$\left|\|h\|_1 \tilde{f}(x) - f(x)\right| \leq \max_{u \in [-s,s]^2} |\tilde{f}(x) - \tilde{f}(x - u)|\, \|h\|_1 \leq \max_{u \in [-s,s]^2} |\tilde{f}(x) - \tilde{f}(x - u)|.$$

Now, if $x \notin \mathcal{C}_s$, the regularity of $\tilde{f}$ yields

$$\forall x \notin \mathcal{C}_s, \quad \left| \|h\|_1 \tilde{f}(x) - f(x) \right| \leq \|\tilde{f}\|_{\mathbf{C}_1} s \tag{A.89}$$

and otherwise

$$\forall x \in \mathcal{C}_s, \quad \left| \|h\|_1 \tilde{f}(x) - f(x) \right| \leq 2 \|\tilde{f}\|_\infty \leq 2 \|\tilde{f}\|_{\mathbf{C}_1}. \tag{A.90}$$

Now the area of $\mathcal{C}_S$ and B can be controlled: $\#\mathcal{C}_s \leq C \ell_1 s$ and $\#B \leq C \ell_1 \ell_2$. So,

$$\left\| \|h\|_1 \tilde{f}(x) - f(x) \right\|^2 = \int_{x \notin \mathcal{C}_s} \left| \|h\|_1 \tilde{f}(x) - f(x) \right|^2 \mathrm{d}x + \int_{x \in \mathcal{C}_s} \left| \|h\|_1 \tilde{f}(x) - f(x) \right|^2 \mathrm{d}x. \tag{A.91}$$

Inserting (A.88), (A.89), (A.90), and the bounds on the areas yields

$$\leq C \ell_1 \ell_2 \|\tilde{f}\|^2_{\mathbf{C}_1} s^2 + C \ell_1 s \|\tilde{f}\|^2_{\mathbf{C}_1}, \tag{A.92}$$

so we conclude that

$$\left\| \|h\|_1 \tilde{f}(x) - f(x) \right\|^2 \leq C \|\tilde{f}\|^2_{\mathbf{C}_1} \ell_1 s. \qquad \square \tag{A.93}$$

Appendix B. Proofs of lemmas for Theorem 4.3.

B.1. Lemma 4.5: Geometry construction.

Proof of Lemma 4.5. We give the main arguments of the proof without the details. We prove that the recursive splitting occurs, after a finite number of steps, only near the junctions. This allows us to control the number of each kind of square as well as the number of parameters required to describe these squares.

The nontangency condition implies that there is a minimum angle $\theta_0 > 0$ between the tangents of the edges at the junctions. The regularity of the curves allows us to define a neighborhood of each junction in which the tangents do not vary by more than $\theta_0/3$. The angle between the tangents of two different curves thus remains larger than $\theta_0/3$, and the geometry around the junction is close to the geometry of the junction of half-lines as illustrated in the close-up of Figure 4.5. Outside this neighborhood, there is a minimal distance $d > 0$ between the curves, and so in a finite number, independent of the geometric precision $\eta = \max(s, T^{2\alpha/(\alpha+1)})$, of steps each dyadic square is either a regular square or an edge square. The corresponding number of square is thus uniformly bounded.

In the neighborhood of the junction, the recursive splitting continues until the size of the squares is of order η, but one can verify that, after a few steps, the number of squares around each junction that can be labeled as *temporary* square is bounded by a constant. As there is only a finite number of such junctions, this implies that the number of the edge squares as well as the number of regular squares is bounded by $C |\log_2 \eta|$ and that the number of junction squares is bounded by C.

The segmentation is specified by the M_S inner nodes of the corresponding dyadic tree. There is a finite number, independent of η, of splits outside the neighborhood of

the junction and at most a constant number of splits at each scale near the junctions. So, as $\eta \geq T^{2\alpha/(\alpha+1)}$,

$$M_S \leq C' + C' \left|\log_2 \eta\right| \leq C \left|\log_2 T\right|. \tag{B.1}$$

As seen in Lemma 3.6, it is enough to keep $\max(2^{\lambda_i} \min(\|c\|_{\mathbf{C}^\alpha}^{1/\alpha}, 1)\, \eta^{-1/\alpha}, K)$ coefficients for each edge square Ω_i of size 2^{λ_i}, where K is the size of the support of ψ. Summing these bounds over all edge squares yields

$$M_G \leq C\, \ell_{\mathcal{C}}\, \min(\|c\|_{\mathbf{C}^\alpha}^{1/\alpha}, 1)\, \eta^{-1/\alpha} + C \left|\log_2 T\right| K, \tag{B.2}$$

$$M_G \leq C\, \ell_{\mathcal{C}}\, \min(\|c\|_{\mathbf{C}^\alpha}^{1/\alpha}, 1)\, T^{-2/(\alpha+1)}, \tag{B.3}$$

$$M_G \leq C\, C_f\, \ell_{\mathcal{C}} T^{-2/(\alpha+1)}, \tag{B.4}$$

where $\ell_{\mathcal{C}}$ is the total length of the edge curves.

Combining (B.1) and (B.4) gives (4.23). □

B.2. Lemma 4.6: Wavelets over regular squares.

Proof of Lemma 4.6. Let Ω_i be a regular square of size 2^λ. By definition, f is $\mathbf{C}^\alpha$ in Ω_i, and this regularity implies

$$|\langle f, \Psi_{j,m}^d \rangle| \leq C \|f\|_{\mathbf{C}^\alpha(\Omega_i)} 2^{(\alpha+1)j}. \tag{B.5}$$

Inserting (A.3) in (B.5) yields

$$|\langle f, \Psi_{j,m}^d \rangle| \leq C \|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)} 2^{(\alpha+1)j}. \tag{B.6}$$

We define now a cutting scale 2^{j_0} as the largest scale such that $C\|\tilde{f}\|_{\mathcal{C}^\alpha}(\Lambda) 2^{(\alpha+1)j}$ is smaller than T, so

$$\frac{1}{2}\left(\frac{T}{C\|\tilde{f}\|_{\mathcal{C}^\alpha(\Lambda)}}\right)^{1/(\alpha+1)} \leq 2^{j_0} \leq \left(\frac{T}{C\|\tilde{f}\|_{\mathcal{C}^\alpha(\Lambda)}}\right)^{1/(\alpha+1)}. \tag{B.7}$$

If we let $J' = \{(j,m);\, 2^j \geq 2^{j_0}\}$, Lemma 3.10 implies

$$\tilde{\mathcal{L}}_i(f, T, \mathcal{B}_i) \leq \sum_{(j,m)\notin J'} |\langle f, \Psi_{j,m}^d \rangle|^2 + T^2 \operatorname{Card}(J'). \tag{B.8}$$

The number of wavelets at the scale 2^j is $\max(2^{2\lambda}2^{-2j}, K^2)$, with K the size of the support of ψ, so with the definition of J

$$\operatorname{Card}(J') \leq C \max(2^{2\lambda}2^{-2j_0}, K^2). \tag{B.9}$$

Combining the bound on the number of wavelets and the inequality (B.6), one obtains that, for any scale 2^j, the energy at this scale satisfies

$$\sum_m |\langle f, \Psi_{j,m}^d \rangle|^2 \leq C \|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)}^2 2^{2\lambda} 2^{2\alpha j} \tag{B.10}$$

and, since $(j,m) \in J$ for $2^j \geq 2^{j_0}$,

$$\sum_{(j,m)\notin J} |\langle f, \Psi_{j,m}^d \rangle|^2 \leq C \|\tilde{f}\|_{\mathbf{C}^\alpha(\Lambda)}^2 2^{2\lambda} 2^{2\alpha j_0}. \tag{B.11}$$

Combining (B.9) and (B.11) with (B.8) and inserting the definition of 2^{j_0} of (B.7) eventually yields

$$\tilde{\mathcal{L}}_i(f,T,\mathcal{B}_i) \leq C \max(2^{2\lambda}\|f\|^{2/(\alpha+1)}_{\mathbf{C}^\alpha(\Lambda)} T^{2\alpha/(\alpha+1)}, K^2T^2). \tag{B.12}$$

As proved in Lemma 4.5, there are at most $C\,|\log T|$ such squares in the chosen partition, and their total area, $\sum_{i\in} 2^{2\lambda_i}$, is bounded by the area of $[0,1]^2$. Summing the inequality (B.12) over all regular squares thus gives

$$\sum_{i\in\mathcal{R}} \tilde{\mathcal{L}}_i(f,T,\mathcal{B}_i) \leq C \max(\|f\|^{2/(\alpha+1)}_{\mathbf{C}^\alpha(\Lambda)} T^{2\alpha/(\alpha+1)}, K^2\,|\log_2 T|\,T^2). \tag{B.13}$$

Thus as $\log|T|K^2T^2 \leq \|f\|^{2/(\alpha+1)}_{\mathbf{C}^\alpha(\Lambda)} T^{2\alpha/(\alpha+1)}$ for T small enough it proves (B.8). □

B.3. Lemma 4.7: Wavelets over junction squares.

Proof of Lemma 4.7. Lemma 4.5 proves that the number of such squares is bounded by a constant C of the same order as the number of curves that is bounded by definition. It is thus sufficient to prove a bound similar to (4.25) for each junction square Ω_i.

By definition, the size of Ω_i is smaller than $\eta = \max(s, T^{2\alpha/(\alpha+1)})$. If $\eta = T^{2\alpha/(\alpha+1)}$, the result holds immediately as $\|f\|^2_{\Omega_i} \leq \|f\|_\infty \eta^2 \leq \|f\|_\infty T^{2\alpha/(\alpha+1)}$ and $\|f\|_\infty \leq \|\tilde{f}\|_\infty \|h\|_1 \leq \|\tilde{f}\|_\infty$. Otherwise $\eta = s \geq T^{2\alpha/(\alpha+1)}$, and we use the regularity of h to obtain the bound.

Indeed, combining (A.34) and (B.5) yields

$$|\langle f, \Psi^d_{j,m}\rangle| \leq C\|\tilde{f}\|_\infty\, s^{-\alpha}\, 2^{(\alpha+1)j}. \tag{B.14}$$

We define now the cutting scale 2^{j_0} as the largest scale such that $C\|\tilde{f}\|_\infty s^{-\alpha} 2^{(\alpha+1)j} \leq T$, so

$$\frac{1}{2}\left(\frac{Ts^\alpha}{C\|\tilde{f}\|_\infty}\right)^{1/(\alpha+1)} \leq 2^{j_0} \leq \left(\frac{Ts^\alpha}{C\|\tilde{f}\|_\infty}\right)^{1/(\alpha+1)}. \tag{B.15}$$

For $J' = \{(j,m);\, 2^j \geq 2^{j_0}\}$, Lemma 3.10 implies

$$\tilde{\mathcal{L}}_i(f,T,\mathcal{B}_i) \leq \sum_{(j,m)\notin J'} |\langle f, \Psi^d_{j,m}\rangle|^2 + T^2\,\mathrm{Card}\,(J'). \tag{B.16}$$

The number of wavelets at the scale 2^j is bounded by $\max(s^2 2^{-2j}, K^2)$, as the size of the square is smaller than s, and thus, with the definition of J',

$$\mathrm{Card}\,(J') \leq C \max(s^2 2^{-2j_0}, K^2). \tag{B.17}$$

Using (B.14), we derive that at any scale 2^j

$$\sum_m |\langle f, \Psi^d_{j,m}\rangle|^2 \leq C \max(s^2\|\tilde{f}\|^2_\infty s^{-2\alpha} 2^{2\alpha j}, \|\tilde{f}\|^2_\infty s^{-2\alpha} 2^{2(\alpha+1)j}), \tag{B.18}$$

so

$$\sum_{(j,m)\notin J} |\langle f, \Psi^d_{j,m}\rangle|^2 \leq C \max(s^2\|\tilde{f}\|^2_\infty 2^{2\alpha j_0}, \|\tilde{f}\|^2_\infty s^{-\alpha} 2^{2(\alpha+1)j_0}). \tag{B.19}$$

Inserting these bounds in (B.16) and using the definition of 2^{j_0} in (B.15) yields

$$\tilde{\mathcal{L}}_i(f,T,\mathcal{B}_i) \leq C s^{2/(\alpha+1)} \|\tilde{f}\|_\infty^{2/(\alpha+1)} T^{2\alpha/(\alpha+1)}, \tag{B.20}$$

so

$$\tilde{\mathcal{L}}_i(f,T,\mathcal{B}_i) \leq C \|\tilde{f}\|_\infty^{2/(\alpha+1)} T^{2\alpha/(\alpha+1)}. \qquad \square \tag{B.21}$$

B.4. Lemma 4.8: Bandelets.

Proof of Lemma 4.8. Theorem 3.2, and more specifically Proposition 3.8, applies to any band B_i associated with a region Ω_i with $\tilde{g} = Q_{T^2}(P_{\mathbf{V}_k}(g))$, $\ell_1 = 2^{\lambda_i}$, $2^k = \eta^{1/\alpha}$, and $C_d = C \max(\|c\|_{\mathbf{C}^\alpha}, \|\theta\|_\infty p)$, so $\|Q_{T^2}(P_{\mathbf{V}_k}(g)) - g\| \leq C_d \eta$. Combining (3.35) and (3.36) thus implies that

$$\tilde{\mathcal{L}}_i(f,T,\mathcal{B}_i) \leq C \max(2^{\lambda_i} C_f^{2/(\alpha+1)} T^{2\alpha/(\alpha+1)}, T^2 |\log_2 T|). \tag{B.22}$$

Summing (B.22) over the horizontal edge squares yields

$$\sum_{i \in I_{\mathcal{H}}} \tilde{\mathcal{L}}_i(f,T,\mathcal{B}_i) \leq \sum_{i \in I_{\mathcal{H}}} C \max(C_f^{2/(\alpha+1)} 2^{\lambda_i} T^{2\alpha/(\alpha+1)}, T^2 |\log_2 T|), \tag{B.23}$$

$$\sum_{i \in I_{\mathcal{H}}} \tilde{\mathcal{L}}_i(f,T,\mathcal{B}_i) \leq \sum_{i \in I_{\mathcal{H}}} C C_f^{2/(\alpha+1)} T^{2\alpha/(\alpha+1)} 2^{\lambda_i} + \sum_{i \in I_{\mathcal{V}}} C T^2 |\log_2 T|. \tag{B.24}$$

Lemma 4.5 proves that there are at most $C |\log T|$ edge squares and that the sum of their size is bounded by a constant $C \ell_{\mathcal{C}}$, so

$$\sum_{i \in I_{\mathcal{H}}} \tilde{\mathcal{L}}_i(f,T,\mathcal{B}_i) \leq C C_f^{2/(\alpha+1)} \ell_{\mathcal{C}} T^{2\alpha/(\alpha+1)}, \tag{B.25}$$

which proves (4.26) and concludes Lemma 4.8. $\square$

Appendix C. Complexity of the CART algorithm. We denote by $\mathcal{C}(2^\lambda)$ the numerical complexity to find this best geometric flow over a square Ω_i of size 2^λ. Since there are $2^{-2\lambda}$ square regions of size 2^λ in $[0,1]^2$, the total numerical complexity to find the best bandelet bases in all dyadic squares is $\sum_{\lambda=\log_2 T^2}^{0} 2^{-2\lambda} \mathcal{C}(2^\lambda)$.

Since the minimum size of squares Ω_i is T^2, the segmentation quadtree has a depth at most equal to $|\log_2 T^2|$. The number of comparisons of the bottom-up optimization algorithm is proportional to the number of nodes of a full quadtree of depth $|\log_2 T^2|$, which is $\mathcal{O}(T^{-4})$. The total number of operations to find the best bandelet frame is therefore

$$\sum_{\lambda=\log_2 T^2}^{0} 2^{-2\lambda} \mathcal{C}(2^\lambda) + \mathcal{O}(T^{-4}). \tag{C.1}$$

Let us now compute the computational complexity $\mathcal{C}(2^\lambda)$ to find the geometric flow and the corresponding bandelet frame $\overline{\mathcal{F}}_i$ which yields a minimum Lagrangian value $\mathcal{L}(\overline{f}, T, \overline{\mathcal{F}}_i)$ over a square Ω_i of size 2^λ. For this purpose all possible geometric flows are explored. If there is no flow, then $\overline{\mathcal{F}}_i$ is a discrete wavelet basis of Ω_i. The wavelet coefficients and the corresponding Lagrangian value are computed with $\mathcal{O}(\#\Omega_i \epsilon^{-2}) = \mathcal{O}(2^{2\lambda} \epsilon^{-2})$ operations. If there is a horizontally parallel flow, then

Ω_i is subdecomposed into $2^{\lambda-k}$ rectangles $\Omega_{i,m}$ of length 2^k and height 2^λ, where 2^k is an adjustable scale variable. There are $\mathcal{O}((2^k\,T^{-2})^p)$ different polynomial flows over each $\Omega_{i,m}$. Computing the corresponding coefficients of $\overline{f}$ for each polynomial flow and its Lagrangian cost with the algorithm of section 4.1 requires $\mathcal{O}(2^{\lambda+k}\,\epsilon^{-2})$ operations. Among the $\mathcal{O}((2^k\,T^{-2})^p)$ polynomial flows, finding the one that minimizes the Lagrangian can thus be done with $\mathcal{O}(2^{\lambda+(p+1)k}\,\epsilon^{-2}\,T^{-2p})$ operations. Hence, the total number of operations to find the best horizontally parallel flow is

$$\sum_{k=\log_2 T^2}^{\lambda} 2^{\lambda-k}\,\mathcal{O}(2^{\lambda+(p+1)k}\,\epsilon^{-2}\,T^{-2p}) = \mathcal{O}(2^{(p+2)\lambda}\,\epsilon^{-2}T^{-2p}). \tag{C.2}$$

The same argument applies to vertically parallel flows. Combining all possibilities (no flow, horizontally or vertically parallel flows), it results that the total number of operations to find the best bandelet frame over a square of size 2^λ is $\mathcal{C}(2^\lambda) = \mathcal{O}(2^{(p+2)\lambda}\,\epsilon^{-2}T^{-2p})$. Inserting this in (C.1) shows that the numerical complexity to find the best bandelet frame over $[0,1]^2$ is

$$\sum_{\lambda=\log_2 T^2}^{0} 2^{-2\lambda}\,\mathcal{C}(2^\lambda) + \mathcal{O}(T^{-4}) = \mathcal{O}(\epsilon^{-2}\,T^{-2p}). \tag{C.3}$$

Appendix D. Proofs of lemmas for Theorem 6.1.

D.1. Lemma 6.2: Existence of a bandelet frame with polynomial geometry.

Proof of Lemma 6.2. To show the existence of a frame satisfying the conditions of the lemma, we use the same strategy as in the proof of Lemma 4.4 and focus on the case $T^{2\alpha/(\alpha+1)} \leq s$ with $\alpha > 1$.

The dyadic squares of this frame are exactly the ones of Lemma 4.5. A wavelet basis is still used in the regular squares and the junction squares. The only difference is in the edge squares which are further subdivided first in an isotropic way and then along the direction of the edge in order to use a bandelet basis with a polynomial geometry.

As the Lagrangian remains the same for the previous regular and junction squares, we will thus study only the modification of the edge squares.

The first subdivision is handled with Lemma 4.6 and yields at most $T^{-2/(\alpha+1)}$ regular squares for which the Lagrangian is bounded by KT^2. The contribution of these new regular squares is thus at most of order $KT^{2\alpha/(\alpha+1)}$, and thus contribution of all the regular squares is still of the right order.

The remaining dyadic edge squares are further subdivided along the edge direction, while their size in this direction is larger than $2^{k_0} = \max(\|c\|_{\mathbf{C}^\alpha}^{-1/\alpha}, 1)\,s^{1/\alpha}$.

On a rectangle of width $2^k \leq 2^{k_0}$ starting at $2^k m$, the geometry g is specified by a polynomial of degree $p \geq \alpha$. It is chosen as the discretized projection of the Taylor polynomial $\pi_{2^k m}(c)$ of the parameterization c of the curve at the point $2^k m$, so

$$\pi_{2^k m}(c)(x) = \sum_{\beta \leq \alpha} \frac{c^{(\beta)}(2^k m)}{\beta!}(x - 2^k m)^\beta \tag{D.1}$$

and

$$g(x) = \sum_{n=1}^{p} Q_{T^2}\left(\frac{\langle \pi_{2^k m}(c), \theta_n(2^{-k}t - m)\rangle}{\|\theta_n\|_1 2^k}\right)\theta_n(2^{-k}t - m). \tag{D.2}$$

This requires p coefficients per rectangle. As $2^k \leq \max(\|c\|_{\mathbf{C}^\alpha}^{-1/\alpha}, 1)\, s^{1/\alpha}$, one can verify then that this implies

$$|(c-g)^{(\beta)}(x)| \leq C\, \max(\|c\|_{\mathbf{C}^\alpha}, C_d)\, s^{1-\beta/\alpha} \quad \forall \beta \leq \lfloor\!\lfloor \alpha \rfloor\!\rfloor \tag{D.3}$$

and

$$|(c-g)^{(\lfloor\!\lfloor \alpha \rfloor\!\rfloor)}(x) - (c-g)^{(\lfloor\!\lfloor \alpha \rfloor\!\rfloor)}(x_0)| \leq C\, \max(\|c\|_{\mathbf{C}^\alpha}, C_d)\, |x-x_0|^{\alpha - \lfloor\!\lfloor \alpha \rfloor\!\rfloor} \tag{D.4}$$

with $C_d = \|\theta\|_{\mathbf{C}^\alpha} p$ that do not depend on f and could be incorporated in the first constant, so these are the same bounds as the ones obtained in Lemma 3.6. The proofs of Lemmas 3.4 and 3.5 can be repeated almost identically, with special care in Lemma 3.4 for the bandelets coming from the anisotropy of the supporting rectangle, to obtain for each rectangle

$$\tilde{\mathcal{L}}(f, T, \mathcal{B}_{i,r}) \leq C \max(C_f^{2/(\alpha+1)}\, 2^k\, T^{2\alpha/(\alpha+1)}, T^2\, K\, |\log_2 T|). \tag{D.5}$$

There are at most $C\, \ell_{\mathcal{C}}\, \min(\|c\|_{\mathbf{C}^\alpha}^{1/\alpha}, 1)\, s^{-1/\alpha}$ rectangles of size $\max(\|c\|_{\mathbf{C}^\alpha}^{-1/\alpha}, 1)\, s^{1/\alpha}$ and at most $C\, |\log_2 T|$ dyadic squares of smaller size, so

$$\begin{aligned}\sum_{i,r} \tilde{\mathcal{L}}(f, T, \mathcal{B}_{i,r}) &\leq C C_f^{2/(\alpha+1)} \ell_{\mathcal{C}}\, T^{2\alpha/(\alpha+1)} + C\, \ell_{\mathcal{C}}\, \min(\|c\|_{\mathbf{C}^\alpha}^{1/\alpha}, 1)\, s^{-1/\alpha}\, T^2\, K\, |\log_2 T| \\ &\quad + C\, |\log_2 T|\, T^2\, K\, |\log_2 T|\end{aligned} \tag{D.6}$$

and for T small enough

$$\sum_{i,r} \tilde{\mathcal{L}}(f, T, \mathcal{B}_{i,r}) \leq C\, C_f^{2/(\alpha+1)}\, \ell_{\mathcal{C}} \max(T^{2\alpha/(\alpha+1)}, s^{-1/\alpha}\, T^2\, |\log_2 T|). \tag{D.7}$$

For each rectangle, there is a constant number p of parameters, so we can obtain

$$\sum_{i,r} \mathcal{L}(f, T, \mathcal{B}_{i,r}) \leq C\, C_f^{2/(\alpha+1)}\, \ell_{\mathcal{C}} \max(T^{2\alpha/(\alpha+1)}, s^{-1/\alpha}\, T^2\, |\log_2 T|). \tag{D.8}$$

As $s \geq T^{2\alpha/(\alpha+1)}$ this implies

$$\sum_{i,r} \mathcal{L}(f, T, \mathcal{B}_{i,r}) \leq C\, C_f^{2/(\alpha+1)}\, \ell_{\mathcal{C}}\, T^{2\alpha/(\alpha+1)}\, |\log_2 T| \tag{D.9}$$

and concludes the proof when $T^{2\alpha/(\alpha+1)} \leq s$. One should note that as long as $T^{2\alpha/(\alpha+1)}\, |\log_2 T|^\alpha \leq s$ the logarithmic factor disappears.

The result for $T^{2\alpha/(\alpha+1)} \geq s$ is then obtained as in the proof of Theorem 3.2 with Lemma 3.9. □

D.2. Lemma 6.3: Discretization.

Proof of Lemma 6.3. As the integral flow g is defined in a continuous way, a continuous bandelet basis $\{b_{l,j,m_1,m_2}\}$ can be constructed over any band B. Over the warped band WB, with $WB = \Omega$ if there is no warping, we define

$$W\tilde{f}_{WB}(x_1, x_2) = \sum \overline{Wf}[n_1, n_2] \phi_{j_0,n_1}(x_1) \phi_{j_0,n_2}(x_2) \tag{D.10}$$

or

$$W\tilde{f}_{WB}(x_1,x_2)=\sum \overline{f}[n_1,n_2+\lceil g(n_1\epsilon)\epsilon^{-1}\rceil]\phi_{j_0,n_1}(x_1)\phi_{j_0,n_2}(x_2), \tag{D.11}$$

with $2^{j_0}=\epsilon$, so

$$\langle \overline{f},\overline{b}_{l,j,m_1,m_2}\rangle=\langle \overline{Wf},\overline{\psi}_{j_1,m_1}[n_1]\overline{\psi}_{j_2,m_2}[n_2]\rangle, \tag{D.12}$$

$$\langle \overline{f},\overline{b}_{l,j,m_1,m_2}\rangle=\langle W\tilde{f}_{W\Omega}(x_1,x_2),\psi_{j,m_1}(x_1)\psi_{l,m_2}(x_2)\rangle. \tag{D.13}$$

Now, over the rectangle WB,

$$\sum_{l\geq j>j_0}|\langle \overline{f},\overline{b}_{l,j,m}\rangle-\langle f,b_{l,j,m}\rangle|^2=\sum_{l\geq j>j_0}\Big|\langle W\tilde{f}_{W\Omega}(x_1,x_2),\psi_{j,m_1}(x_1)\psi_{l,n_2}(x_2)\rangle - \langle Wf(x_1,x_2),\psi_{j,m_1}(x_1)\psi_{l,n_2}(x_2)\rangle\Big|^2 \tag{D.14}$$

using the orthogonality of the wavelet basis

$$=\sum\Big|\langle W\tilde{f}_{W\Omega}(x_1,x_2),\phi_{j_0,n_1}(x_1)\phi_{j_0,n_2}(x_2)\rangle - \langle Wf,\phi_{j_0,n_1}(x_1)\phi_{j_0,n_2}(x_2)\rangle\Big|^2 \tag{D.15}$$

and as

$$\langle W\tilde{f}_{W\Omega}(x_1,x_2),\phi_{j_0,m_1}(x_1)\phi_{j_0,n_2}(x_2)\rangle=\langle Wf(x_1,x_2),W^{-1}(\epsilon^{-1}\mathbf{1}_{m_1\epsilon,(m_2+\lceil g(m_1\epsilon)\epsilon^{-1}\rceil)\epsilon})\rangle,$$

we have

$$=\sum\Big|\langle Wf(x_1,x_2),W^{-1}(\epsilon^{-1}\mathbf{1}_{m_1,m_2+\lceil g(m_1\epsilon)\epsilon^{-1}\rceil})(x_1,x_2)\rangle - \langle Wf(x_1,x_2),\phi_{j_0,m_1}(x_1)\phi_{j_0,m_2}(x_2)\rangle\Big|^2, \tag{D.16}$$

$$\sum_{l\geq j>j_0}|\langle \overline{f},\overline{b}_{l,j,m}\rangle-\langle f,b_{l,j,m}\rangle|^2=\sum|\langle Wf,\widetilde{\Phi}_{j_0,m_1,m_2}\rangle-\langle Wf,\Phi_{j_0,m_1,m_2}\rangle|^2, \tag{D.17}$$

with $\widetilde{\Phi}_{j_0,m_1,m_2}(x_1,x_2)=W^{-1}(\epsilon^{-1}\mathbf{1}_{m_1\epsilon,(m_2+\lceil g(m_1\epsilon)\epsilon^{-1}\rceil)\epsilon})(x_1,x_2)$ and $\Phi_{j_0,m_1,m_2}(x_1,x_2)=\phi_{j_0,m_1}(x_1)\phi_{j_0,m_2}(x_2)$.

Now we will temporarily admit that for any family of scaling functions $\Phi_{j_0,m}$ of scale 2^{j_0} and integral 2^{j_0} of support included in the ball of radius $K\,2^{j_0}$ centered on the point $m2^{j_0}$ included in Ω and such that the total length of curve in Ω is bounded by ℓ, for any set $\{x_{j_0,m}:|x_{j,m}-2^{j_0}m|\leq K2^{j_0}\}$,

$$\sum_m|\langle f,\Phi_{j_0,m}\rangle-2^{j_0}f(x_{j_0,m})|^2\leq C(l+|\Omega|2^{j_0}). \tag{D.18}$$

This can be applied simultaneously to $\{\widetilde{\Phi}_{j_0,m}\}$ and $\{\Phi_{j_0,m}\}$ to obtain

$$\left(\sum_m|\langle Wf,\phi_{j_0,m}\rangle-\langle Wf,\tilde{\phi}_{j_0,m}\rangle|^2\right)^{1/2}\leq 2\left(C2^{j_0}(l+|\Omega|2^{j_0})\right)^{1/2}. \tag{D.19}$$

Summing over all rectangles gives (6.23) and concludes the proof.

We now prove (D.18). As the scaling function Φ_{j_0} is of integral 2^{j_0},

$$\langle f, \Phi_{j_0,m}\rangle - 2^{j_0} f(x_{j_0,m}) = \int (f(x) - f(x_{j_0,m}))\Phi_{j_0,m}(x)\mathrm{d}x \tag{D.20}$$

and

$$|\langle f, \Phi_{j_0,m}\rangle - 2^{j_0} f(x_{j_0,m})| \leq K^2\, 2^{j_0} \max_{|x-x_{j_0,m}|\leq K2^{j_0}} |f(x) - f(x_{j_0,m})|. \tag{D.21}$$

We now bound $\max_{|x-x_{j_0,m}|\leq K2^{j_0}} |f(x) - f(x_{j_0,m})|$ with respect to the position of $x_{j_0,m}$:

- If $x_{j_0,m}$ is such that one remains at a distance greater than s from the discontinuities, the regularity of f is used to obtain

$$\max_{|x-x_{j_0,m}|\leq K2^{j_0}} |f(x) - f(x_{j_0,m})| \leq \|\tilde{f}\|_{\mathbf{C}^1(\Lambda)}\, K\, 2^{j_0}. \tag{D.22}$$

- Otherwise, the regularity of h implies $\|f\|_{\mathbf{C}^1(\Lambda)} \leq \|\tilde{f}\|_\infty s^{-1}$, so

$$\max_{|x-x_{j_0,m}|\leq C2^{j_0}} |f(x) - f(x_{j_0,m})| \leq \|\tilde{f}\|_{\mathbf{C}^1(\Lambda)}\, s^{-1}\, K\, 2^{j_0}. \tag{D.23}$$

- Finally, if s is smaller than 2^{j_0}, one still has

$$\max_{|x-x_{j_0,m}|\leq C2^{j_0}} |f(x) - f(x_{j_0,m})| \leq 2\|f\|_\infty \leq C\|\tilde{f}\|_{\mathbf{C}^1(\Lambda)}. \tag{D.24}$$

Whatever 2^{j_0}, there are at most $\max(\#\Omega\, 2^{-2j_0}, K^2)$ coefficients where the bound (D.22) applies. If $2^{j_0} \geq s$, the smoothing does not affect the coefficients, and the bound (D.24) yields

(D.25)
$$\sum_m |\langle f, \Phi_{j_0,m}\rangle - 2^{j_0} f(x_{j_0,m})|^2 \leq C\, \#\Omega\, 2^{-2j_0}\, \|\tilde{f}\|^2_{\mathbf{C}^1(\Lambda)} (2^{2j_0})^2 + C\ell 2^{-j0} \|\tilde{f}\|^2_{\mathbf{C}^1(\Lambda)} (2^{j_0})^2,$$

(D.26)
$$\sum_m |\langle f, \Phi_{j_0,m}\rangle - 2^{j_0} f(x_{j_0,m})|^2 \leq C\, \|\tilde{f}\|^2_{\mathbf{C}^1(\Lambda)}\, 2^{j_0}(\ell + \#\Omega 2^{j_0}).$$

Otherwise $2^{j_0} \leq s$, the discontinuities are already smoothed at the scale 2^{j_0}, the number of affected coefficients is $\max(C\,\ell\, s\, 2^{-2j_0}, K)$, and the bound (D.23) gives

$$\begin{aligned}\sum_m |\langle f, \Phi_{j_0,m}\rangle - 2^{j_0} f(x_{j_0,m})|^2 &\leq C\, \#\Omega\, 2^{-2j_0}\, \|f\|^2_{\mathbf{C}^1(\Lambda)}\, (2^{2j_0})^2 \\ &\quad + C\,\ell\, s\, 2^{-2j_0}\, \|\tilde{f}\|^2_{\mathbf{C}^1(\Lambda)}\, (s^{-1}2^{2j_0})^2\end{aligned} \tag{D.27}$$

and, as $s^{-1} \leq 2^{-j_0}$,

$$\sum_m |\langle f, \Phi_{j_0,m}\rangle - 2^{j_0} f(x_{j_0,m})|^2 \leq C\, \|\tilde{f}\|^2_{\mathbf{C}^1(\Lambda)}\, 2^{j_0}(\ell + \#\Omega\, 2^{j_0}). \tag{D.28}$$

□

REFERENCES

[1] E. Candès and D. Donoho, *Curvelets: A surprisingly effective nonadaptive representation of objects with edges*, in Curves and Surfaces Fitting, L. L. Schumaker, A. Cohen, and C. Rabut, eds., Vanderbilt University Press, Nashville, TN, 1999.

[2] E. Candes and D. Donoho, *New tight frames of curvelets and optimal representations of objects with piecewise c_2 singularities*, Comm. Pure Appl. Math., 57 (2004), pp. 219–266.

[3] A. Canny, *A computational approach to edge detection*, IEEE Trans. Patt. Anal. Mach. Intell., 8 (1986), pp. 679–698.

[4] A. Cohen, I. Daubechies, and J. C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, Comm. Pure Appl. Math., 45 (1992), pp. 485–560.

[5] A. Cohen, R. deVore, P. Petrushev, and H. Xu, *Non linear approximation and the space $BV(\mathbb{R}^2)$*, Amer. J. Math., 121 (1999), pp. 587–628.

[6] A. Cohen and B. Matei, *Nonlinear subdivisions schemes: Applications to image processing*, in Tutorial on Multiresolution in Geometric Modelling, A. Iske, E. Quack, and M. Floater, eds., Springer, Heidelberg, 2002, pp. 93–97.

[7] L. Demaret, N. Dyn, and A. Iske, *Image compression by linear splines over adaptive triangulations*, Signal Process., to appear.

[8] M. N. Do and M. Vetterli, *Contourlets*, in Beyond Wavelets, J. Stoeckler and G. V. Welland, eds., Academic Press, New York, 2003.

[9] D. Donoho, *Wedgelets: Nearly-minimax estimation of edges*, Ann. Statist., 27 (1999), pp. 353–382.

[10] D. Donoho, *Counting Bits with Kolmogorov and Shannon*, Tech. rep., Stanford University, Stanford, CA, 2000.

[11] C. Dossal, E. Le Pennec, and S. Mallat, *Geometric Image Estimation with Bandelets*, Tech. rep., CMAP, École Polytechnique, Palaiseau, France, 2005.

[12] M. Jansen, H. Choi, S. Lavu, and R. Baraniuk, *Multiscale image processing using normal triangulated meshes*, in Proceedings of the International Conference on Image Processing, Thessaloniki, Greece, 2001.

[13] A. Korostelev and A. Tsybakov, *Minimax Theory of Image Reconstruction*, Lecture Notes in Statist. 82, Springer, New York, 1993.

[14] E. Le Pennec and S. Mallat, *Sparse geometrical image representation with bandelets*, IEEE Trans. Image Process., 14 (2005), pp. 423–438.

[15] S. Mallat, *A theory for multiresolution signal decomposition: The wavelet representation*, PAMI, 11 (1989), pp. 674–693.

[16] Y. Meyer, *Wavelets and Operators*, Cambridge University Press, Cambridge, UK, 1993.

[17] R. Shukla, P. L. Dragotti, M. N. Do, and M. Vetterli, *Rate-distortion optimized tree structured compression algorithms for piecewise smooth images*, IEEE Trans. Image Process., 14 (2005), pp. 343–359.

[18] M. Wakin, J. Romberg, H. Choi, and R. Baraniuk, *Rate-distortion optimized image compression using wedgelets*, in Proceedings of the IEEE International Conference on Image Processing, Rochester, NY, 2002.