Compensated Convex Based Transforms
for Image Processing and Shape Interrogation

Antonio Orlando∗, Elaine Crooks† and Kewei Zhang‡

Abstract

This paper reviews some recent applications of the theory of the compensated convex transforms or of the proximity hull as developed by the authors to image processing and shape interrogation with special attention given to the Hausdorff stability and multiscale properties. The paper contains also numerical experiments that demonstrate the performance of our methods compared to the state-of-art ones.

Keywords: Compensated convex transform, Moreau envelope, Proximity hull, Mathematical morphology, Hausdorff-Lipschitz continuity, Image processing, Shape interrogation, Scattered data

2000 Mathematics Subjects Classification number: 90C25, 90C26, 49J52, 52A41, 65K10, 62H35, 14J17, 58K25, 53-XX, 65D17, 53A05, 26B25, 52B55, 65D18

Email: aorlando@herrera.unt.edu.ar, e.c.m.crooks@swansea.ac.uk, kewei.zhang@nottingham.ac.uk

1 Introduction

The compensated convex transforms were introduced in [116, 117] for the purpose of tight approximation of functions defined in $$\mathbb{R}^n$$ and their definitions were originally motivated by the translation method [76, 50, 54, 105] in the study of the quasiconvex envelope in the vectorial calculus of variations (see [45, 115] and references therein) and in the variational approach of material microstructure [19, 20, 21]. Thanks to their smoothness and tight approximation property, these transforms provide geometric convexity-based techniques for general functions that yield novel methods for identifying singularities in functions [122, 123, 118, 120] and new tools for function and image interpolation and approximation [119, 121]. In this paper we present some of the applications that have been tackled by this theory up to date. These range from the detection of features in images or data [122, 123, 120], to multi-scale medial-axis extraction [118], to surface reconstruction from level sets, to approximation of scattered data and noise removal from images, to image inpainting [119, 121].

Suppose $$f : \mathbb{R}^n \to \mathbb{R}$$ satisfies the following growth condition

$$f(x) \geq -A_1|x|^2 - A_2 \quad \text{for any } x \in \mathbb{R}^n,$$

for some constants $$A_1, A_2 \geq 0$$, then the quadratic lower compensated convex transform (lower transform for short) for a given $$\lambda > A_1$$ is defined in [116] by

$$C_\lambda^l(f)(x) = \text{co} \left[ \lambda |\cdot|^2 + f \right] (x) - \lambda |x|^2 \quad x \in \mathbb{R}^n,$$
where $|x|$ is the Euclidean norm of $x \in \mathbb{R}^n$ and $\text{co}[g]$ the convex envelope \cite{58, 88} of a function $g : \mathbb{R}^n \to \mathbb{R}$ bounded below. Similarly, given $f : \mathbb{R}^n \to \mathbb{R}$ satisfying the growth condition

$$f(x) \leq A_1|x|^2 + A_2 \quad \text{for any } x \in \mathbb{R}^n,$$

for some constants $A_1, A_2 \geq 0$, the quadratic upper compensated convex transform (upper transform for short) for a given $\lambda > A_1$ is defined \cite{116} by

$$C^u_\lambda(f)(x) = -C^l_\lambda(-f)(x) = \lambda|x|^2 - \text{co} [\lambda \cdot |x|^2 - f](x) \quad x \in \mathbb{R}^n. \quad (1.4)$$

It is not difficult to verify that if $f$ meets both (1.1) and (1.3), for instance if $f$ is bounded, there holds

$$C^l_\lambda(f)(x) \leq f(x) \leq C^u_\lambda(f)(x) \quad x \in \mathbb{R}^n,$$

thus, the lower and upper compensated convex transforms are $\lambda$-parametrised families of transforms that approximate $f$ from below and above respectively. Furthermore, they have smoothing effects and are tight approximations of $f$ in the sense that if $f$ is $C^{1,1}$ in a neighbourhood of $x_0$, there is a finite $\Lambda > 0$, such that $f(x_0) = C^l_\lambda(f)(x_0)$ (respectively, $f(x_0) = C^u_\lambda(f)(x_0)$ whenever $\lambda \geq \Lambda$. This approximation property, which we refer to as tight approximation, is pivotal in the developments of the theory, because it allows the transforms to be used for detecting singularities of functions by exploiting the fact that it is only when a point $x$ is close to a singularity point of $f$ we might find that the values of $C^l_\lambda(f)(x)$ and $C^u_\lambda(f)(x)$ might be different from that of $f(x)$ \cite{122}. Figure 1 visualizes the smoothing and tight approximation of the mixed transform $C^u_\lambda(C^l_\lambda(f))$ of the squared distance function $f$ to a four-point set. Given the type of singularity of $f$, we apply the lower transform to $f$ which smooth the ‘concave’–like singularity followed by the upper transform that smooths the ‘convex’–like singularity of $C^l_\lambda(f)$ which are unaltered with respect to the original function $f$. This can be appreciated by the graph of the pointwise error $e(x) = |f(x) - C^u_\lambda(C^l_\lambda(f))(x)|$ for $x \in \Omega$ which is zero everywhere but in a neighborhood of the singularities of $f$.

The transforms additionally satisfy the locality property that the values of $C^l_\lambda(f)$, $C^u_\lambda(f)$ at $x \in \mathbb{R}^n$ depend only on the values of $f$ in a neighbourhood of $x$, and are translation invariant in the sense that $C^l_\lambda(f)$, $C^u_\lambda(f)$ are unchanged if the ‘weight’ $|\cdot|^2$ in the formula (1.2) and (1.4) is replaced by
$| \cdot - x_0 |^2$ for any $x_0 \in \mathbb{R}^n$. These last two properties make the explicit calculation of transforms tractable for specific prototype functions $f$, which facilitates the creation of dedicated extractors for a variety of different types of singularity using customised combinations of the transforms.

These new geometric approaches enjoy key advantages over previous image and data processing techniques. The curvature parameter $\lambda$ provides scales for features that allow users to select which size of feature they wish to detect, and the techniques are blind and global, in the sense that images/data are treated as a global object with no a priori knowledge required of, e.g., feature location. Figure 2 displays the $\lambda$–scale dependence in the case of the medial axis where $\lambda$ is associated with the scale of the different branches whereas Figure 3 shows the multiscale feature for given $\lambda$ associated with the height of the different branches of the multiscale medial axis map.

Many of the methods can also be shown to be stable under perturbation and different sampling.
techniques. Most significantly, Hausdorff stability results can be rigorously proved for many of the methods. For example, the Hausdorff-Lipschitz continuity estimate \[ |C^u_\lambda (\chi_E)(x) - C^u_\lambda (\chi_F)(x)| \leq 2\sqrt{\lambda} \text{dist}_H(E,F), \quad x \in \mathbb{R}^n, \]
shows that the upper transform \( C^u_\lambda \) is Hausdorff stable against sampling of geometric shapes defined by their characteristic functions. Such stability is particularly important for the extraction of information when ‘point clouds’ represent sampled domains. If a geometric shape is densely sampled, then from a human vision point of view, one can typically still identify geometric features of the sample and sketch its boundary. From the mathematical/computer science perspective, however, feature identification from sampled domains is challenging and usually methods are justified only by either ad hoc arguments or numerical experiments. Figure 4 displays an instance of this property where we show the edges of the continuous nonnegative function \( f(x,y) = \text{dist}^2((x,y), \partial \Omega) \), with \((x,y) \in \Omega = ([1.5] \times [-1.5, 1.5]) \setminus ([1.5, 0.5] \times [-1.5, -0.5])\), and of its sparse sampling \( f \cdot \chi_A \) where \( A \subset \Omega \) is a sparse set (see Figure 4(a), (b) respectively). Due to the Hausdorff stability of the stable ridge transform, we are able to recover an approximation of the ridges from the sampled image (compare Figure 4(c), (d)).

![Figure 4](image)

Figure 4 (a) Image of \( f(x,y) \); (b) Sampled image of \( f(x,y) \) by random salt and pepper noise; (c) Stable ridges of \( f(x,y) \); (d) Stable ridges from sampled image.

Via fast and robust numerical implementations of the transforms [124], this theory also gives rise to a highly-effective computational toolbox for applications. The efficiency of the numerical computations benefits greatly from the locality property, which holds despite the global nature of the convex envelope itself.

Before we describe the applications of this theory, we provide next alternative characterizations of the compensated convex transforms.

1.1 Related areas: Semiconvex envelope

Given the definitions (1.1) and (1.3), lower and upper compensated convex transforms can be considered as parameterized semiconvex and semiconcave envelopes, respectively, for a given function. The notions of semiconvex and semiconcave functions go back at least to Reshetnyak [87] and have since been studied by many authors in different contexts (see, for example, [5, 6, 18, 91, 92]). Let \( \Omega \subseteq \mathbb{R}^n \) be an open set, we recall that a function \( f : \Omega \rightarrow \mathbb{R} \cup \{+\infty\} \) is semiconvex if there is a constant \( C \geq 0 \) such that \( f(x) = g(x) - C|x|^2 \) with \( g \) a convex function. More general weight functions, such as \( |x|\sigma(|x|) \), for example, are also used in the literature for defining more general semiconvex functions [5, 6, 18, 91, 92]. Since general DC-functions (difference of convex functions)
and semiconvex/semiconcave functions are locally Lipschitz functions in their essential domains (Theorem 2.1.7), Rademacher’s theorem implies that they are differentiable almost everywhere. Fine properties for the singular sets of convex/concave and semiconvex/semiconcave functions have been studied extensively showing that the singular set of a semiconvex/semiconcave function is rectifiable. By applying results and tools of the theory of compensated convex transforms, it is possible therefore to study how such functions can be effectively approximated by smooth functions, whether all singular points are of the same type, that is, for semiconcave (semiconvex) functions, whether all singular points are geometric ‘ridge’ (‘valley’) points, how singular sets can be effectively extracted beyond the definition of differentiability and how the information concerning ‘strengths’ of different singular points can be effectively measured. These are all questions relevant to applications in image processing and computer-aided geometric design. An instance of this study, for example, has been carried out in to study the singular set of the Euclidean squared-distance function \( \text{dist}^2(\cdot, \Omega^c) \) to the complement of a bounded open domain \( \Omega \subset \mathbb{R}^n \) (called the medial axis of the domain \( \Omega \)) and of the weighted squared distance function.

1.2 Related areas: Proximity hull

Another characterization of the compensated convex transforms is in terms of the critical mixed Moreau envelopes, given that

\[
C^l_\lambda(f)(x) = M^\lambda(M_\lambda(f))(x), \quad C^u_\lambda(f)(x) = M_\lambda(M^\lambda(f))(x),
\]

where the Moreau lower and upper envelopes are defined, in our notation, respectively, by

\[
M_\lambda(f)(x) = \inf \{ f(y) + \lambda|y - x|^2, \ y \in \mathbb{R}^n \}, \\
M^\lambda(f)(x) = \sup \{ f(y) - \lambda|y - x|^2, \ y \in \mathbb{R}^n \},
\]

with \( f \) satisfying the growth condition (1.1) and (1.3), respectively. Moreau envelopes play important roles in optimization, nonlinear analysis, optimal control and Hamilton-Jacobi equations, both theoretically and computationally. The mixed Moreau envelopes \( M^\tau(M_\lambda(f)) \) and \( M_\tau(M^\lambda(f)) \) coincide with the Lasry-Lions double envelopes \( (f_\lambda)^\tau \) and \( (f^\lambda)^\tau \) defined in by (2.5) and (2.6), respectively, in the case of \( \lambda = \tau \) and are also referred to in as proximal hull and upper proximal hull, respectively. They have been extensively studied and used as approximation and smoothing methods of not necessarily convex functions. In particular, in the partial differential equation literature, the focus of the study of the mixed Moreau envelopes \( M^\tau(M_\lambda(f)) \) and \( M_\tau(M^\lambda(f)) \) for the case \( \tau > \lambda \) are known, under suitable growth conditions, as the Lasry-Lions regularizations of \( f \) of parameter \( \lambda \) and \( \tau \). In this case, the mixed Moreau envelopes are both \( C^{1,1} \) functions. However, crucially they are not ‘tight approximations’ of \( f \), in contrast with our lower and upper transforms \( C^l_\lambda(f)(x) \) and \( C^u_\lambda(f)(x) \). Generalised inf and sup convolutions have also been considered, for instance in. However, due to the way these regularization operators are defined, proof of mathematical and geometrical results to describe how such approximations work has usually been challenging, making their analysis and applications very difficult. As a result, the study of the proximal hull using the characterization in terms of the compensated convex transform would make them much more accessible and feasible for real world applications.
1.3 Related areas: Mathematical morphology

Moreau lower and upper envelopes have also been employed in mathematical morphology in the 1990’s [60 [108], to define greyscale erosion and dilation morphological operators, whereas the critical mixed Moreau envelopes $M^\lambda(M_\lambda(f))$ and $M_\lambda(M^\lambda(f))$ are greyscale opening and closing morphological operators [96 [102 [98]. In convex analysis, the infimal convolution of $f$ with $g$ is denoted as $f □ g$ and is defined as [88 [23 [41 [90]

$$(f □ g)(x) = \inf_y \{ f(y) + g(x - y) \}.$$  

This is closely related to the erosion of $f$ by $g$, given that

$$(f □ g)(x) = f(x) ⊖ (-g(-x)).$$

Thus if we denote by $b_\lambda(x) = -\lambda|x|^2$ the quadratic structuring function, introduced for the first time in [60 [107 [63 [01 [62], then with the notation of [96 [102 [98 [8], we have

$$M_\lambda(f)(x) = \inf_{y \in \mathbb{R}^n} \{ f(y) - b_\lambda(y - x) \} =: f \ominus b_\lambda,$$

$$M^\lambda(f)(x) = \sup_{y \in \mathbb{R}^n} \{ f(y) + b_\lambda(y - x) \} =: f \oplus b_\lambda$$  

(1.7)

so that (1.5) can be written alternatively as

$$C^\lambda_\lambda(f) = (f \ominus b_\lambda) \oplus b_\lambda \quad \text{and} \quad C^\lambda_\lambda(f) = (f \oplus b_\lambda) \ominus b_\lambda.$$  

(1.8)

The application of $M^\lambda(M_\lambda(f))$ and $M_\lambda(M^\lambda(f))$ in mathematical morphology [96 [102 [98], however, has not met with corresponding success, nor have its properties been fully explored. This is in contrast with the rôle, recognized since its introduction, that is played by paraboloid structuring functions in defining morphological scale-spaces in image analysis [60 [107 [94 [68 [69 [74 [74 [113 [112]. For this and related topics concerning the morphological scale-space representation produced by quadratic structuring functions, we refer to the pioneering works [60 [107]. Here, we would like only to observe that through identity (1.5), we have a direct characterization of the quadratic structuring based opening and closing morphological operators, either in terms of the convex envelope (see (1.2) and (1.4)) or in terms of envelope from below/above with parabolas (see (1.9) and (1.10)). Such characterizations will allow us to derive various new geometric and stability properties for the opening and closing morphological operators. Furthermore, when we apply compensated convex transforms to extract singularities from characteristic functions of compact geometric sets, our operations can be viewed as the application of morphological operations devised for ‘greyscale images’ to ‘binary images’. As a result, it might look not efficient to apply more involved operations for processing binary images, when in the current literature [96 [102 [98] there are ‘binary’ set theoretic morphological operations that have been specifically designed for the tasks under examination. Nevertheless, an advantage of adopting our approach is that the compensated convex transforms of characteristic functions are (Lipschitz) continuous, therefore applying a combination of transforms will produce a landscape of various levels (heights) that can be designed to highlight a specific type of singularity. We can then extract multiscale singularities by taking thresholds at different levels. In fact, the graphs of functions obtained by combinations of compensated convex transforms contain much more geometric information than binary operations that produce simply a yes or no answer. Also, for ‘thin’ geometric structures, such as curves and surfaces, it is difficult to design ‘binary’ morphological operations to be Hausdorff stable.
1.4 Related areas: Quadratic envelopes

From definition (1.2), it also follows that $C^{l}_{\lambda}(f)(x)$ is the envelope of all the quadratic functions with fixed quadratic term $\lambda|x|^2$ that are less than or equal to $f$, that is,

$$C^{l}_{\lambda}(f)(x) = \sup \{ -\lambda|x|^2 + \ell(x) : -\lambda|y|^2 + \ell(y) \leq f(y) \text{ for all } y \in \mathbb{R}^n \text{ and } \ell \text{ affine} \},$$

(1.9)

whereas from (1.4) it follows that $C^{u}_{\lambda}(f)(x)$ is the envelope of all the quadratic functions with fixed quadratic term $\lambda|x|^2$ that are greater than or equal to $f$, that is,

$$C^{u}_{\lambda}(f)(x) = \inf \{ \lambda|x|^2 + \ell(x) : f(y) \leq \lambda|y|^2 + \ell(y) \text{ for all } y \in \mathbb{R}^n \text{ and } \ell \text{ affine} \}.$$

(1.10)

This characterization was first given in [122, Eq. (1.4)] and can be derived by noting that since the convex envelope of a function $g$ can be characterized as the pointwise supremum of the family of all the affine functions which are majorized by $g$, we have then

$$C^{l}_{\lambda}(f)(x) = \text{co}[f + \lambda|\cdot||/(x) - \lambda|x|^2$$

$$= \sup \ell(x) : \ell(y) \leq f(y) + \lambda|y|^2 \text{ for any } y \in \mathbb{R}^n$$

(1.11)

$$= \sup \ell(x) - \lambda|x|^2 : \ell(y) - \lambda|y|^2 \leq f(y) \text{ for any } y \in \mathbb{R}^n,$$

which is (1.9). As stated before, (1.11) can be in turn related directly to the Moreau’s mixed envelope. The characterization (1.9) has been recently also reproposed by [32] for the study of low-rank approximation and compressed sensing.

It is instructive to compare this characterization with (4.7) below about the Moreau envelopes.

1.5 Outline of the Chapter

The plan of the paper is as follows. After this general introduction, we will introduce relevant notation and recall basic results in convex analysis and compensated convex transforms in the next section. In Section 3, we introduce the different compensated convex based transforms that we have been developing. Their definition can be either motivated by a mere application of key properties of the basic transforms, namely the lower and upper transform, or by an ad–hoc designed combinations of the basic transforms so to create a singularity at the location of the feature of interest. Section 4 introduces some of the numerical schemes that can be used for the numerical realization of the compensated convex based transforms, namely of the basic transform given by the lower compensated convex transform. We will therefore describe the convex based and Moreau based algorithms, which can be both used according to whether we refer to the definition (1.2) or the characterization (1.5) of the lower compensated convex transform. Section 5 contains some representative applications of the transformations introduced in this paper. More specifically, we will consider an application to shape interrogation by considering the problem of identifying the location of intersections of manifolds represented by point clouds, and applications of our approximation compensated convex transform to the reconstruction of surfaces using level lines and isolated points, image inpainting and salt & pepper noise removal.

2 Notation and Preliminaries

Throughout the paper $\mathbb{R}^n$ denotes the $n$-dimensional Euclidean space, whereas $|x|$ and $x \cdot y$ are the standard Euclidean norm and inner product respectively, for $x, y \in \mathbb{R}^n$. Given a non-empty
subset $K$ of $\mathbb{R}^n$, $K^c$ denotes the complement of $K$ in $\mathbb{R}^n$, i.e. $K^c = \mathbb{R}^n \setminus K$, $\overline{K}$ its closure, $\text{co}[K]$ the convex hull of $K$, that is, the smallest (with respect to inclusion) convex set that contains the set $K$ and $\chi_K$ its characteristic function, that is, $\chi_K(x) = 1$ if $x \in K$ and $\chi_K(x) = 0$ if $x \in K^c$. The Euclidean distance transform of a non-empty set $K \subset \mathbb{R}^n$ is the function that, at any point $x \in \mathbb{R}^n$, associates the Euclidean distance of $x$ to $K$, which is defined as $\inf \{|x - y|, y \in K\}$ and is denoted as $\text{dist}(x; K)$. Let $\delta > 0$, the open $\delta$-neighbourhood $K^\delta$ of $K$ is then defined by $K^\delta = \{x \in \mathbb{R}^n, \text{dist}(x, K) < \delta\}$ and is an open set. For $x \in \mathbb{R}^n$ and $r > 0$, $B(x; r)$ indicates the open ball with center $x$ and radius $r$ whereas $S(x; r)$ denotes the sphere with center $x$ and radius $r$, that is, $S(x; r) = \partial B(x; r)$ is the boundary of $B(x; r)$. The suplevel set of a function $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ of level $\alpha$ is the set

$$S_\alpha f = \{x \in \Omega : f(x) \geq \alpha\},$$

(2.1)

whereas the level set of $f$ with level $\alpha$ is also defined by (2.1) with the inequality sign replaced by the equality sign. Finally, we use the notation $Df$ to denote the derivative of $f$.

Next we next list some basic properties of compensated convex transforms. Without loss of generality, these properties are stated mainly for the lower compensated convex transform given that it is then not difficult to derive the corresponding results for the upper compensated convex transform using (1.4). Only in the case $f$ is the characteristic function of a set $K$, i.e. $f = \chi_K$, we will refer explicitly to $C^\lambda_\chi(\chi_K)$ given that $C^\lambda_\chi(\chi_K)(x) = 0$ for any $x \in \mathbb{R}^n$ if $K$ is, e.g., a finite set. For details and proofs we refer to [116, 122] and references therein, whereas for the relevant notions of convex analysis we refer to [58, 88, 23].

**Definition 2.1.** Given a function $f : \mathbb{R}^n \to \mathbb{R}$ bounded below, the convex envelope $\text{co}[f]$ is the largest convex function not greater than $f$.

This is a global notion. By Carathéodory’s Theorem [58, 88], we have

$$\text{co}[f](x_0) = \inf_{x_i \in \mathbb{R}^n} \left\{ \sum_{i=1}^{n+1} \lambda_i f(x_i) : \sum_{i=1}^{n+1} \lambda_i = 1, \sum_{i=1}^{n+1} \lambda_i x_i = x_0, \lambda_i \geq 0 \; i = 1, \ldots, n + 1 \right\},$$

(2.2)

that is, the convex envelope of $f$ at a point $x_0 \in \mathbb{R}^n$ depends on the values of $f$ on its whole domain of definition, namely $\mathbb{R}^n$ in this case. We will however introduce also a local version of this concept which will be used to formulate the locality property of the compensated convex transform and is fundamental for our applications.

**Definition 2.2.** Let $r > 0$, $x_0 \in \mathbb{R}^n$. Assume $f : B(x_0; r) \to \mathbb{R}$ to be bounded from below. Then the value of the local convex envelope of $f$ at $x_0$ in $B(x_0; r)$ is defined by

$$\text{co}_{B(x_0; r)}[f](x_0) = \inf_{x_i \in B(x_0; r)} \left\{ \sum_{i=1}^{n+1} \lambda_i f(x_i) : \sum_{i=1}^{n+1} \lambda_i = 1, \sum_{i=1}^{n+1} \lambda_i x_i = x_0, \lambda_i \geq 0 \; i = 1, \ldots, n + 1 \right\}.$$

(2.3)

Unlike the global definition, the infimum in (2.3) is taken only over convex combinations in $B(x_0; r)$ rather than in $\mathbb{R}^n$. 

8
As part of the convex analysis reminder, we also recall the definition of the Legendre-Fenchel transform.

**Definition 2.3.** Let \( f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} \), \( f \neq +\infty \) and there is an affine function minorizing \( f \) on \( \mathbb{R}^n \). The conjugate (or Legendre-Fenchel transform) of \( f \) is

\[
f^* : s \in \mathbb{R}^n \to f^*(s) = \sup_{x \in \mathbb{R}^n} \{ x \cdot s - f(x) \}, \tag{2.4}
\]

and the biconjugate of \( f \) is \((f^*)^*\).

We have then the following results.

**Proposition 2.4.** For \( f \) satisfying the conditions of Definition 2.3, the conjugate \( f^* \) is a lowersemi-continuous convex function and \((f^*)^*\) is equal to the lowersemicontinuous convex envelope of \( f \).

Before stating the properties of interest of the compensated convex transforms, we describe the relationship between the compensated convex transforms and other infimal convolutions.

Let \( f : \mathbb{R}^n \to \mathbb{R} \) satisfy (1.1) and (1.3). As we have mentioned in the introduction, concepts closely related to the compensated convex transforms are the Lasry-Lions regularisations for parameters \( \lambda \) and \( \tau \) with \( 0 < \tau < \lambda \), which are defined in [66] as follows

\[
(f^\lambda)^\tau(x) = \sup_{y \in \mathbb{R}^n} \inf_{u \in \mathbb{R}^n} \{ f(u) + \lambda |u - y|^2 - \tau |y - x|^2 \}
= M^\tau (M_\lambda(f))(x), \tag{2.5}
\]

and

\[
(f^\lambda)_\tau(x) = \inf_{y \in \mathbb{R}^n} \sup_{u \in \mathbb{R}^n} \{ f(u) - \lambda |u - y|^2 + \tau |y - x|^2 \}
= M_\tau (M^\lambda(f))(x). \tag{2.6}
\]

Both \((f^\lambda)^\tau\) and \((f^\lambda)_\tau\) approach \( f \) from below and above respectively, as the parameters \( \lambda \) and \( \tau \) go to \(+\infty\). If \( \lambda = \tau \), then \((f^\lambda)^\lambda = M^\lambda (M_\lambda(f))\) is called proximal hull of \( f \) whereas \((f^\lambda)_\lambda = M_\lambda (M^\lambda(f))\) is refered to as the upper proximal hull of \( f \). It is not difficult to verify that whenever \( \tau > \lambda > 0 \) the following relation holds between the compensated convex transforms, the Moreau envelopes and the Lary–Lions regularizations of \( f \) [116],

\[
M_\lambda(f)(x) \leq M^\lambda (M_\tau(f))(x) \leq C^\lambda_l(f)(x) \leq f(x) \leq C^\tau_u(f)(x) \leq M_\lambda (M^\tau(f))(x) \leq M^\tau(f)(x)
\tag{2.7}
\]

for \( x \in \mathbb{R}^n \).

Given \( f : \mathbb{R}^n \to \mathbb{R} \), we recall also that the lower semicontinuous envelope of \( f \) is defined in [58] [88] by

\[
\underline{f} : x \in \mathbb{R}^n \mapsto \underline{f}(x) = \liminf_{y \to x} f(y), \tag{2.8}
\]

and since there holds

\[
C^\lambda_l(f)(x) = C^\lambda_l(\underline{f})(x) \quad \text{for} \quad x \in \mathbb{R}^n, \tag{2.9}
\]

without loss of generality, in the following we can assume that the functions are lower semicontinuous.

The monotonicity and approximation properties of \( C^\lambda_l(f) \) with respect to \( \lambda \) is described by the following results.
Proposition 2.5. Given \( f : \mathbb{R}^n \to \mathbb{R} \) that satisfies (1.1), then for all \( A_1 < \lambda < \tau < \infty \), we have

\[
C^d_{\lambda}(f)(x) \leq C^l_{\lambda}(f)(x) \leq f(x) \quad \text{for } x \in \mathbb{R}^n, \tag{2.10}
\]

and, for \( \lambda > A_1 \)

\[
\lim_{\lambda \to \infty} C^l_{\lambda}(f)(x) = f(x) \quad \text{for } x \in \mathbb{R}^n. \tag{2.11}
\]

The approximation of \( f \) from below by \( C^l_{\lambda}(f) \) given by (2.11) can be better specified, given that \( C^l_{\lambda}(f) \) realizes a ‘tight’ approximation of the function \( f \) in the following sense (see [116, Theorem 2.3(iv)]).

Proposition 2.6. Let \( f \in C^{1,1}(\bar{B}(x_0; r)) \), \( x_0 \in \mathbb{R}^n, r > 0 \). Then for sufficiently large \( \lambda > 0 \), we have that \( f(x_0) = C^l_{\lambda}(f)(x_0) \). If the gradient of \( f \) is Lipschitz in \( \mathbb{R}^n \) with Lipschitz constant \( L \), then \( C^l_{\lambda}(f)(x) = f(x) \) for all \( x \in \mathbb{R}^n \) whenever \( \lambda \geq L \).

The property of ‘tight’ approximation plays an important role in the definition of the transforms introduced in Section 3. Related to this property is the density property of the lower compensated transform established in [122] that can be viewed as a tight approximation for general bounded functions.

Theorem 2.7. Suppose \( f : \mathbb{R}^n \to \mathbb{R} \) is bounded, satisfying \( |f(x)| \leq M \) for some \( M > 0 \) and for all \( x \in \mathbb{R}^n \). Let \( \lambda > 0 \), \( x_0 \in \mathbb{R}^n \) and define \( R_{\lambda,M} = (2 + \sqrt{2})\sqrt{M/\lambda} \). Then there are \( x_i \in B(x_0; R_{\lambda,M}) \), with \( x_i \neq x_0 \), and \( \lambda_i \geq 0 \) for \( i = 1, \ldots, n + 1 \), satisfying \( \sum_{i=1}^{n+1} \lambda_i = 1 \) and \( \sum_{i=1}^{n+1} \lambda_i x_i = x_0 \), such that

\[
C^l_{\lambda}(f)(x_i) = f(x_i) \quad \text{for } i = 1, \ldots, n + 1.
\]

Since the lower transform satisfies

\[
C^l_{\lambda}(f) \leq f \leq f,
\]

if we consider the following set

\[
T_l(f, \lambda) = \{ x \in \mathbb{R}^n : C^l_{\lambda}(f)(x) = f(x) \},
\]

as a result of Theorem 2.7, the set of points at which the lower compensated convex transform equal the original function satisfies a density property, that is, the closed \( R_{\lambda,M} \)-neighbourhoods of \( T_l(f, \lambda) \) covers \( \mathbb{R}^n \). For any point \( x_0 \in \mathbb{R}^n \), the point \( x_0 \) is contained in the local convex hull \( \text{co} \left[ T_l(f, \lambda) \cap B(x_0; R_{\lambda,M}) \right] \). Furthermore, if \( f \) is bounded and continuous, \( T_l(f, \lambda) \) is exactly the set of points at which \( f \) is \( \lambda \)-semiconvex [31], i.e., points \( x_0 \) where

\[
f(x) \geq f(x_0) + \ell(x) - \lambda |x - x_0|^2 \quad \text{for all } x \in \mathbb{R}^n
\]

with \( \ell \) an affine function satisfying \( \ell(x_0) = 0 \) and Condition (1.1) holds for \( f \).

A fundamental property for the applications is the locality of the compensated convex transforms. For a lowersemi-continuous function that is in addition bounded on any bounded set, the locality property was established for this general case in [116]. We next report its version for a bounded function which is relevant for the applications to image processing and shape interrogation [122].

Theorem 2.8. Suppose \( f : \mathbb{R}^n \mapsto \mathbb{R} \) is bounded, satisfying \( |f(x)| \leq M \) for some \( M > 0 \) and for all
Let $\lambda > 0$ and $x_0 \in \mathbb{R}^n$, then the following locality properties hold,

$$C^l_\lambda(f)(x_0) = \inf \left\{ \sum_{i=1}^{n+1} \lambda_i (f(x_i) + \lambda |x_i - x_0|^2), \quad \lambda_i \geq 0, \sum_{i=1}^{n+1} \lambda_i = 1, \sum_{i=1}^{n+1} \lambda_i x_i = x_0, |x_i - x_0| \leq R_{\lambda,M} \right\},$$

(2.12)

where $R_{\lambda,M}$ is the same as in Theorem 2.7.

Since the convex envelope is affine invariant, it is not difficult to realize that there holds

$$C^l_\lambda(f)(x_0) = \text{co} \left\{ \lambda |\cdot - x_0|^2 + f \right\}(x_0) \quad \text{for} \quad x_0 \in \mathbb{R}^n$$

(2.13)

thus condition (2.12) can be equivalently written as

$$C^l_\lambda(f)(x_0) = \text{co}_{\text{H}(x_0; R_{\lambda,M})} \left[ \lambda |\cdot - x_0|^2 + f \right](x_0).$$

(2.14)

Despite the definition of $C^l_\lambda(f)$ involves the convex envelope of $f + \lambda |\cdot|^2$, the value of the lower transform for a bounded function at a point depends on the values of the function in its $R_{\lambda,M}$-neighborhood. Therefore when $\lambda$ is large, the neighborhood will be very small. If $f$ is globally Lipschitz, this result is a special case of Lemma 3.5.7 at p. 72 of [31].

The following property shows that the mapping $f \mapsto C^l_\lambda(f)$ is nondecreasing, that is we have

**Proposition 2.9.** If $f \leq g$ in $\mathbb{R}^n$ and satisfy (1.1), then

$$C^l_\lambda(f)(x) \leq C^l_\lambda(g)(x) \quad \text{for} \quad x \in \mathbb{R}^n \quad \text{and} \quad \lambda \geq \max \{A_{1,f}, A_{1,g}\}.$$

We conclude this section by stating some results on the Hausdorff stability of the compensated convex transforms. This is the relevant concept of stability we use to assess the change of the transformations with respect to perturbations of the set, thus it refers to the behaviour of the compensated convex transform of the characteristic functions of subsets $K$ of $\mathbb{R}^n$. We first state a result that highlights the geometric structure of the upper transform of $\chi_K$.

**Theorem 2.10.** (Expansion Theorem) Let $E \subset \mathbb{R}^n$ be a non-empty set and let $\lambda > 0$ be fixed, then

$$C^u_\lambda(\chi_E)(x) = \begin{cases} 1, & \text{if} \quad x \in \bar{E}, \\ 0, & \text{if} \quad x \in (\bar{E}^{1/\sqrt{\lambda}})^c, \\ \in (0, 1), & \text{if} \quad x \in E^{1/\sqrt{\lambda}} \setminus \bar{E}. \end{cases}$$

Next, we recall the definition of Hausdorff distance from [10].

**Definition 2.11.** Let $E, F$ be non-empty subsets of $\mathbb{R}^n$. The Hausdorff distance between $E$ and $F$ is defined by

$$\text{dist}_H(E, F) = \inf \left\{ \delta > 0 : F \subset E^\delta \text{ and } E \subset F^\delta \right\}.$$

This definition is also equivalent to saying that

$$\text{dist}_H(E, F) = \max \left\{ \sup_{x \in E} \text{dist}(x; F), \sup_{x \in F} \text{dist}(x; E) \right\}.$$

It is well-known and easy to prove that the Euclidean distance function $\text{dist}(x, K)$ is Hausdorff-
Lipschitz continuous in the sense that for given $K$ and $S \subset \mathbb{R}^n$ non-empty compact sets, we have

$$|\text{dist}(x, K) - \text{dist}(x, S)| \leq \text{dist}_H(K, S).$$

In order to study the Hausdorff-Lipschitz continuity of the upper compensated convex transform of characteristic functions of compact sets, we introduce the distance based function $D^2_\lambda(x, K)$ defined by

$$D^2_\lambda(x, K) = \left( \max \left\{ 0, 1 - \sqrt{\lambda} \text{dist}(x, K) \right\} \right)^2, \quad x \in \mathbb{R}^n. \quad (2.15)$$

Clearly, we have $0 \leq D^2_\lambda(x, K) \leq 1$ in $\mathbb{R}^n$. More precisely, we have

$$D^2_\lambda(x, K) = \begin{cases} 1, & \text{if } x \in K, \\ 0, & \text{if } \text{dist}(x, K) \geq \frac{1}{\sqrt{\lambda}}, \\ \in (0, 1), & \text{if } 0 < \text{dist}(x, K) < \frac{1}{\sqrt{\lambda}}. \end{cases} \quad (2.16)$$

Suppose $E, F \subset \mathbb{R}^n$ are two non-empty closed sets. It is, then, easy to see that

(i) if $E \subset F$,

$$D^2_\lambda(x, E) \leq D^2_\lambda(x, F), \quad x \in \mathbb{R}^n; \quad (2.17)$$

(ii) for $x \in \mathbb{R}^n$, if $E \cap \bar{B}(x, 1/\sqrt{\lambda}) \neq \emptyset$, then

$$D^2_\lambda(x, E) = D^2_\lambda(x, E \cap \bar{B}(x, 1/\sqrt{\lambda})). \quad (2.18)$$

For a given non-empty closed set $K$, by definition of the function $D^2_\lambda(x, K)$, we have

$$0 \leq \chi_K(x) \leq D^2_\lambda(x, K) \leq 1, \quad x \in \mathbb{R}^n.$$ 

The following result establishes the relationship between the upper transform of $\chi_K(x)$ and $D^2_\lambda(x, K)$ and it was established in [122].

**Proposition 2.12.** Let $K \subset \mathbb{R}^n$ be a non-empty closed set and assume $\lambda > 0$. Then, there holds

$$C^\alpha_\lambda(\chi_K)(x) = C^\alpha_\lambda(D^2_\lambda(\cdot, K))(x), \quad x \in \mathbb{R}^n. \quad (2.19)$$

The Hausdorff-Lipschitz continuity of $C^\alpha_\lambda(\chi_K)(x)$ and $C^\alpha_\lambda(D^2_\lambda(\cdot, K))(x)$ were also established in [122].

**Theorem 2.13.** Let $E, F \subset \mathbb{R}^n$ be non-empty compact sets and let $\lambda > 0$ be fixed, then for all $x \in \mathbb{R}^n$,

$$|D^2_\lambda(x, E) - D^2_\lambda(x, F)| \leq 2\sqrt{\lambda} \text{dist}_H(E, F), \quad (2.20)$$

$$|C^\alpha_\lambda(D^2_\lambda(\cdot, E))(x) - C^\alpha_\lambda(D^2_\lambda(\cdot, F))(x)| \leq 2\sqrt{\lambda} \text{dist}_H(E, F). \quad (2.21)$$

Consequently,

$$|C^\alpha_\lambda(\chi_E)(x) - C^\alpha_\lambda(\chi_F)(x)| \leq 2\sqrt{\lambda} \text{dist}_H(E, F). \quad (2.22)$$
3 Compensated convexity based transforms

The lower compensated convex transform \[ (1.2) \] and the upper compensated convex transform \[ (1.4) \] represent building blocks for defining novel transformations to smooth functions, to identify singularities in functions, and to interpolate and approximate data. For the creation of these transformations we follow mainly two approaches. One approach makes a direct use of the basic transforms to single out singularities of the function or to smooth and/or approximate the function. By contrast, the other approach realises a suitably designed combination of the basic transforms that creates the singularity at the location of the feature of interest.

3.1 Smoothing Transform

Let \( f : \mathbb{R}^n \to \mathbb{R} \) satisfy a growth condition of the form
\[
|f(x)| \leq C_1 |x|^2 + C_2
\] (3.1)
for some \( C_1, C_2 > 0 \), then given \( \lambda, \tau > C_1 \), we can define two (quadratic) mixed compensated convex transform as follows
\[
C_{\tau,\lambda}^{\text{u},l}(f)(x) := C_{\lambda}^{\text{u}}(C_{\tau}^{l}(f))(x) \quad \text{and} \quad C_{\lambda,\tau}^{\text{d},u}(f)(x) := C_{\lambda}^{d}(C_{\tau}^{u}(f))(x), \quad x \in \mathbb{R}^n.
\] (3.2)

From \[ (1.4) \], we have that for every \( \lambda, \tau > C_1 \)
\[
C_{\tau,\lambda}^{\text{u},l}(f)(x) = -C_{\tau,\lambda}^{\text{d},u}(-f),
\] (3.3)
hence properties of \( C_{\lambda,\tau}^{\text{d},u}(f) \) follow from those for \( C_{\tau,\lambda}^{\text{u},l}(f) \) and we can thus state appropriate results only for \( C_{\tau,\lambda}^{\text{u},l}(f) \). In this case, then, whenever \( \tau, \lambda > C_1 \) we have that \( C_{\tau,\lambda}^{\text{u},l}(f) \in C^{1,1}(\mathbb{R}^n) \). As a result, if \( f \) is bounded, then \( C_{\tau,\lambda}^{\text{u},l}(f) \in C^{1,1}(\mathbb{R}^n) \) and \( C_{\lambda,\tau}^{\text{d},u}(f) \in C^{1,1}(\mathbb{R}^n) \) for all \( \lambda > 0 \) and \( \tau > 0 \).

This is important in applications of the mixed transforms to image processing, because there the function representing the image takes a value from a fixed range at each pixel point and so is always bounded. The regularizing effect of the mixed transform is visualized in Figure 5 where we display \( C_{\tau,\lambda}^{\text{u},l}(f) \) of the no-differentiable function \( f(x,y) = |x| - |y|, (x,y) \in [-1, 1] \times [-1, 1] \) and of \( f(x,y) + n(x,y) \) with \( n(x,y) \) a bivariate normal distribution with mean value equal to 0.05. The level lines of \( C_{\lambda,\tau}^{\text{d},u}(f) \) and \( C_{\lambda,\tau}^{\text{d},u}(f + n) \) displayed in Figure 5(b) and Figure ??(d), respectively, are smooth curves.
Figure 5  (a) Input function $f(x, y) = |x| - |y|$; (b) Graph of $C_{\lambda, \tau}^{d}(f)$ for $\lambda = 5$ and $\tau = 5$; (c) Input function $f(x, y) + n(x, y)$ with $n(x, y)$ a bivariate normal distribution with mean value equal to 0.05; (d) Graph of $C_{\lambda, \tau}^{u}(f + n)$ for $\lambda = 5$ and $\tau = 5$.

Finally, as a consequence of the approximation result (2.11) and likewise result for $C_{\tau}^{u}(f)$ (see Proposition 2.5) it is then to difficult to establish a similar approximation result also for the mixed transforms and verify that verify there are $\tau_j, \lambda_j \to \infty$ as $j \to \infty$ such that on every compact subset of $\mathbb{R}^n$, there holds

$$C_{\tau_j}^{u}(C_{\lambda_j}^{d})(f) \to f \quad \text{uniformly as } j \to \infty.$$  \hfill (3.4)

3.2 Stable Ridge/Edge Transform

The ridge, valley and edge transforms introduced in [122] are basic operations for extracting geometric singularities. The key property is the tight approximation of the compensated convex transforms (see Proposition 2.6) and the approximation to $f$ from below by $C_{\lambda}^{d}(f)$ and above by $C_{\lambda}^{u}(f)$, respectively.
3.2.1 Basic transforms

Let \( f : \mathbb{R}^n \to \mathbb{R} \) satisfy the growth condition (3.1). The ridge \( R_\lambda(f) \), the valley \( V_\lambda(f) \) and the edge transforms \( E_\lambda(f) \) of scale \( \lambda > C_1 \) are defined respectively by

\[
R_\lambda(f) = f - C^R_\lambda(f); \quad V_\lambda(f) = f - C^V_\lambda(f);
\]

\[
E_\lambda(f) = R_\lambda(f) - V_\lambda(f) = C^R_\lambda(f) - C^V_\lambda(f).
\]  

If \( f \) is of sub-quadratic growth, that is, \(|f(x)| \leq A(1 + |x|^{\alpha})\) with \( 0 \leq \alpha < 2 \), in particular \( f \) can be a bounded function, the requirement for \( \lambda \) in (3.5) is simply \( \lambda > 0 \).

The ridge transform \( R_\lambda(f) = f - C^R_\lambda(f) \) and the valley transform \( V_\lambda(f) = f - C^V_\lambda(f) \) are non-negative and non-positive, respectively, because of the ordering property of the compensated convex transforms and their support set is disjoint to each other. In the applications, we usually consider \(-V_\lambda(f)\) to make the resulting function non-negative. Figure 6 displays the suplevel set of \( R_\lambda(f) \) and \(-V_\lambda(f)\) of the same level for a gray scale image \( f \) compared to the Canny edge filter whereas Figure 7 demonstrates on the test image used in [101] the ability of \( R_\lambda(f) \) to detect edges between different gray levels.

**Figure 6** (a) Input image; (b) Suplevel set of the ridge and valley transform with \( \lambda = 2.5 \) and for the level equal to 0.005 \( \cdot \) max \([R_\lambda(f)]\) and 0.005 \( \cdot \) max \([-V_\lambda(f)]\), respectively; (c) Canny edges.
Figure 7 (a) Input test image from [101]; (b) Suplevel set of the ridge transform with $\lambda = 0.1$ and for the level equal to $0.004 \cdot \max \left[ R_{\lambda}(f) \right]$; (c) Canny edges.

The transforms $R_{\lambda}(f)$ and $V_{\lambda}(f)$ satisfy the following properties

(i) The transforms $R_{\lambda}(f)$ and $V_{\lambda}(f)$ are invariant with respect to translation, in the sense that

\[ R_{\lambda}(f + \ell) = R_{\lambda}(f) \quad \text{and} \quad V_{\lambda}(f + \ell) = V_{\lambda}(f) \]  \hspace{1cm} (3.6)

for all affine functions $\ell \in \text{Aff}(\mathbb{R}^n)$. Consequently, the edge transform $E_{\lambda}(f)$ is also invariant with respect to translation.

(ii) The transforms $R_{\lambda}(f)$ and $V_{\lambda}(f)$ are scale covariant in the sense that

\[ R_{\lambda}(\alpha f) = \alpha R_{\lambda/\alpha}(f) \quad \text{and} \quad V_{\lambda}(\alpha f) = \alpha V_{\lambda/\alpha}(f) \]  \hspace{1cm} (3.7)

for all $\alpha > 0$. Consequently, the edge transform $E_{\lambda}(f)$ is also scale covariant.

(iii) The transforms $R_{\lambda}(f)$, $V_{\lambda}(f)$ and $E_{\lambda}(f)$ are all stable under curvature perturbations in the sense that for any $g \in C^{1,1}(\mathbb{R}^n)$ satisfying $|Dg(x) - Dg(y)| \leq \epsilon |x - y|$, if $\lambda > \epsilon$ then

\[ R_{\lambda+\epsilon}(f) \leq R_{\lambda}(f + g) \leq R_{\lambda-\epsilon}(f); \quad V_{\lambda-\epsilon}(f) \leq V_{\lambda}(f + g) \leq V_{\lambda+\epsilon}(f); \]
\[ E_{\lambda+\epsilon}(f) \leq E_{\lambda}(f + g) \leq E_{\lambda-\epsilon}(f). \]  \hspace{1cm} (3.8)

The numerical experiments depicted in Figure 8 illustrate the affine invariance of the edge transform expressed by (3.6) whereas Figure 9 shows implications of the stability of the the edge transform under curvature perturbations according to (3.8).
Figure 8 (a) A binary image $\chi$ of a Chinese character; (b) Image $255\chi + \ell$ with $\ell = 70(i - j)$ for $1 \leq i \leq 546, 1 \leq j \leq 571$, i.e. the scaled characteristic function of the character plus an affine function; (c) Edges extracted by Canny edge detector; (d) Edges extracted by the edge transform $E_\lambda(f)$ with $\lambda = 0.1$ after thresholding.

Figure 9 (a) A scaled binary image of a Chinese character perturbed by a smooth image; (b) Edges extracted by Canny edge detector; (c) Edges extracted by the edge transform $E_\lambda(f)$ after thresholding.

To get an insight on the geometric structure of the edge transform, it is informative to consider the case where $f$ is the characteristic function of a set. Let $\Omega \subset \mathbb{R}^n$ be a non-empty open regular set such that $\overline{\Omega} \neq \mathbb{R}^n$ and $\Gamma \subset \partial \Omega$, then for $\lambda > 0$, we have that

$$E_\lambda(\chi_{\Omega \cup \Gamma})(x) = \begin{cases} 0 & x \in (\Omega^{1/\sqrt{\lambda}})^c \cup \Omega \setminus (\Omega^c)^{1/\sqrt{\lambda}} \\ \in (0, 1) & x \in \Omega^{1/\sqrt{\lambda}} \setminus \overline{\Omega} \cup (\Omega^c)^{1/\sqrt{\lambda}} \setminus \Omega^c \\ 1 & x \in \partial \Omega. \end{cases}$$  \quad (3.9)

Furthermore, $E_\lambda(\chi_{\Omega \cup \Gamma})$ is continuous in $\mathbb{R}^n$ and, for $x \in \mathbb{R}^n$ there holds

$$\lim_{\lambda \to +\infty} E_\lambda(\chi_{\Omega \cup \Gamma})(x) = \chi_{\partial \Omega}(x),$$  \quad (3.10)

that is, $\lambda$ controls the width of the neighborhood of $\chi_{\partial \Omega}$. As $\lambda \to \infty$, the support of $E_\lambda(\chi_{\Omega \cup \Gamma})$ shrinks to the support of $\chi_{\partial \Omega}$. 
Figure 10 illustrates the behaviour of $E_\lambda(\chi_\Omega)$ by displaying the support of $E_\lambda(\chi_\Omega)$ for different values of $\lambda$.

![Figure 10](image)

Figure 10  Scale effect associated with $\lambda$ on the support of the edge transform of the (a) image $f = 255 \cdot \chi$ of a Chinese character for different values of $\lambda$: (b) $\lambda = 1$; (c) $\lambda = 10$; (d) $\lambda = 100$.

Since the original function $f$ is directly involved in the definitions of the ridge, valley and edge transforms, the transforms (3.5) are not Hausdorff stable if we consider a dense sampling of the original function. It is possible nevertheless to establish a stable versions of ridge and valley transforms in the case that $f$ is the characteristic function $\chi_E$ of a non-empty closed set $E \subset \mathbb{R}^n$. For this result, it is fundamental the observation on the Hausdorff stability of the upper transform of the characteristic function $\chi_E$ of closed sets which motivates the definition of stable ridge transform of $E$ as

$$SR_{\tau,\lambda}(\chi_E) = C^\tau_{\lambda}(\chi_E) - C^4_{\lambda}(C^n_{\lambda}(\chi_E)).$$

(3.11)

For the ridge defined by (3.11) we have that if $E, F$ are non-empty compact subsets of $\mathbb{R}^n$, for $\lambda > 0$ and $\tau > 0$, then there holds

$$|SR_{\lambda,\tau}(\chi_E)(x) - SR_{\lambda,\tau}(\chi_F)(x)| \leq 4\sqrt{\lambda}\text{dist}_{\mathcal{H}}(E,F) \quad (\text{for } x \in \mathbb{R}^n).$$

(3.12)

Figure 11 illustrates the meaning of (3.12). Figure 11(a) displays a domain $E$ represented by a binary image of a cat, (c) shows a domain $F$ obtained by randomly sampling $E$, whereas (b) and (d) picture a suplevel set of the stable ridge transforms of the respective characteristic functions.

![Figure 11](image)

Figure 11  (a) Domain $E$ given by the image of an elephant displayed here as $1 - \chi_E$; (b) Boundary extraction using the stable ridge transform, $SR_{\lambda,\tau}(\chi_E)$, for $\lambda = 0.1$ and $\tau = \lambda/8$; (c) Domain $F$ obtained by randomly sampling $E$; (d) Boundary extraction of the data sample after thresholding the stable ridge transform, $SR_{\lambda,\tau}(\chi_F)$, computed for $\lambda = 0.1$ and $\tau = \lambda/8$.

Similarly to the Stable Ridge Transform of a non-empty compact subset $E$ of $\mathbb{R}^n$, we can then...
define the Stable Valley Transform of $E$ for $\lambda > \tau$ as
\[
SV_{\lambda, \tau}(\chi_E)(x) = V_{\tau}(C^u_\lambda(\chi_E))(x) \quad x \in \mathbb{R}^n, \quad \lambda > \tau > 0,
\]
and the Stable Edge Transform of $E$ for $\lambda > \tau$ as
\[
SE_{\lambda, \tau}(\chi_E)(x) = E_{\tau}(C^u_\lambda(\chi_E))(x) \quad x \in \mathbb{R}^n, \quad \lambda > \tau > 0.
\]
The condition $\lambda > \tau$ is invoked because it is not difficult to see that
\[
C^u_\tau(C^u_\lambda(f)) = \begin{cases} 
C^u_\lambda(f), & \text{for } \lambda \leq \tau \\
C^u_\tau(f), & \text{for } \lambda \geq \tau.
\end{cases}
\]
Hence, if $\lambda \leq \tau$, we would get $SV_{\lambda, \tau}(\chi_E)(x) = 0$ and $SE_{\lambda, \tau}(\chi_E)(x)$ would simply equal to $SR_{\lambda, \tau}(\chi_E)(x)$.

### 3.2.2 Extractable corner points

Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with $|\partial \Omega| = 0$ (i.e. $\partial \Omega$ has zero $n$–dimensional measure) and $x \in \partial \Omega$. We say that the point $x \in \partial \Omega$ is a $\delta$–regular point of $\partial \Omega$ if there is an open ball $B(x_0; \delta) \subset \Omega^c$, $x_0 \in \Omega^c$, $\delta > 0$, such that $x \in \partial B(x_0; \delta)$ and if there is an open ball $B(x_0; \delta) \subset \Omega$, $x_0 \in \Omega$, $\delta > 0$, such that $x \in \partial B(x_0; \delta)$. If the point $x \in \partial \Omega$ meets only the first condition, we refer to it as exterior $\delta$–regular point whereas if it meets only the second condition is called interior $\delta$–regular point. Figure 12 displays the different type of points of $\partial \Omega$.

![Figure 12](image)

The stable ridge transform allows the characterization of such points given that if $x \in \partial \Omega$ is a $\delta$–regular point of $\Omega$ with $\delta > 0$ sufficiently small, in [122] it is shown that there holds
\[
SR_{\lambda, \tau}(\chi_\Omega)(x_0) \leq \frac{(\sqrt{\lambda + \tau} - \sqrt{\tau})^2}{\lambda}. \tag{3.13}
\]
As a result, we define an extractable corner point of $\Omega$ if for at least sufficiently large $\lambda > 0$ and $\tau > 0$,
\[
SR_{\lambda, \tau}(\chi_\Omega)(x_0) > \mu_1(\lambda, \tau), \tag{3.14}
\]
where
\[
\mu_1(\lambda, \tau) := \frac{(\sqrt{\lambda + \tau} - \sqrt{\tau})^2}{\lambda}. \tag{3.15}
\]
is called the standard height for codimension-1 regular boundary points. The analysis of the
behaviour of $SR_{\lambda,\tau}(\chi_{K_a})$ in the case of the prototype exterior corner defined by the set $K_a = \{(x,y) \in \mathbb{R}^2 : |y| \leq ax, a,x \geq 0\}$, with angle $\theta$ satisfying $a = \tan(\theta/2)$ shows that the value of $SR_{\lambda,\tau}(\chi_{K_a})$ at the corner tip $(0,0)$ of $K_a$ is given by

$$SR_{\lambda,\tau}(\chi_{K_a})(0,0) := \mu_2(a,\lambda,\tau) = \begin{cases} \frac{\lambda}{\lambda + (1 + a^2)\tau} & \text{if } a^2 \leq \sqrt{\frac{\lambda + \tau}{\tau}} \\ \frac{1}{2} \left(\sqrt{\lambda + \tau} - \sqrt{\tau}\right)^2 & \text{otherwise} \end{cases}$$

(3.16)

One can then verify that for $a > 0$, and for any $\lambda, \tau > 0$,

$$\mu_2(a,\lambda,\sigma) > \mu_1(\lambda,\tau) \quad \text{and} \quad \lim_{a \to \infty} \mu_2(a,\lambda,\sigma) = \mu_1(\lambda,\sigma).$$

This result means that when the angle $\theta$ approaches $\pi$, the singularity at $(0,0)$ disappears. Figure 13 illustrates the behaviour of $SR_{\lambda,\tau}(\chi_{K_a})$ for different values of the opening angle $\theta$ and for $\tau = \sigma \lambda$ with $\sigma = 1/8$, for which the value of $a$ that separates the two conditions in (3.16) corresponds to $\theta = 2\pi/3$.

**Figure 13** Graph of $SR_{\lambda,\tau}(\chi_{K_a})$ for different pairs of opening angle $\theta$. (a) $\pi/2 - -\pi/2$; (b) $5\pi/12 - -7\pi/12$; (c) $\pi/12 - -11\pi/12$.

Based on this prototype example [122, Example 6.11], one can therefore conclude that $R_\tau(C_1^\infty(\chi_\Omega))$.
can actually detect exterior corners, whereas it might happen that at some $\delta$-singular points of $\partial \Omega$, $R_\tau(C^m_\lambda(\chi_{\Omega}))$ takes on values lower than at the regular points of $\partial \Omega$. As a result, a different Hausdorff stable method will be therefore needed to detect interior corners and boundary intersections of domains.

### 3.2.3 Interior corners

Since a prototype interior corner is defined as the complement of an exterior corner, one could think of detecting interior corners of $\Omega$ by looking at the stable ridge transform of the complement of $\Omega$ in $\mathbb{R}^n$. But this would not provide useful information for geometric objects subject to finite sampling. On the other hand, traditional methods, such as Harris and Susan, as well as other local mask based corner detection methods, would also not apply directly to such a situation. In this case therefore we adopt an indirect approach. This consists of constructing an ad-hoc geometric designed based function that is robust under sampling and is such that its singularities can be identified with the geometric singularities we want to extract: (i) interior corners of a domain, and (ii) intersections of smooth manifolds. By applying one of the transforms introduced in Section 3.2.1 according to the type of singularity, we can detect the singularity of interest. Given a non–empty closed set $K \subset \mathbb{R}^2$ with $K \neq \mathbb{R}^n$, an instance of function whose singularities capture the type of geometric feature of $K$ which we are interested of, is the distance–based function (2.15) for $\lambda > 0$, which we re-write next for ease of reference

$$D^2_\lambda(x, K) = \left(\max\{0, 1 - \sqrt{\lambda} \text{dist}(x, K)\}\right)^2, \quad x \in \mathbb{R}^n. \quad (3.17)$$

Figure 14(a) displays the graph of $D^2_\lambda(x, K)$ for a prototype of interior corner in an $L$–shape domain, and shows that such singularity is of the valley type.

![Figure 14](a) Prototype of internal corner. L–shape domain. (a) Graph of $D^2_\lambda(\cdot, K)$ for $\lambda = 0.0001$; (b) Graph of $V^d_\lambda(\cdot, K)$.

By applying then to $D^2_\lambda(\cdot, K)$ the valley transform (3.5) with the same parameter $\lambda$ as used to compute $D^2_\lambda(\cdot, K)$ itself, we obtain

$$V^d_\lambda(x, K) = -V_\lambda(D^2_\lambda(\cdot, K))(x) = C^m_\lambda(D^2_\lambda(\cdot, K))(x) - D^2_\lambda(x, K), \quad x \in \mathbb{R}^n, \quad (3.18)$$

whose graph is displayed in Figure 14(b). We observe therefore that this transform allows the definition of the set of interior corner points and intersection points of scale $1/\sqrt{\lambda}$ as the support of $V^d_\lambda(\cdot, K)$, that is

$$I_\lambda(K) = \{x \in \mathbb{R}^n, V^d_\lambda(x, K) > 0\}. \quad (3.19)$$
In this manner we obtain a marker which is localized in the neighborhood of the feature. Figure 15 displays, for $\lambda = 0.0001$, the behaviour of $D^2_\lambda(\cdot, K)$, of $V^d_\lambda(\cdot, K)$, and of the suplevel set of $V^d_\lambda(\cdot, K)$ for a level equal to $0.8 \max_{x \in \mathbb{R}^2} \{V^d_\lambda(x, K)\}$ as approximation of $I_\lambda(K)$, considering different opening angles of the interior corner prototype $K$. As for the exterior corner, we observe that the marker reduces and the maximum of $V^d_\lambda(x, K)$ depends on the opening angle of the corner. The larger is the angle, the smaller is the value of $\max_{x \in \mathbb{R}^2} V^d_\lambda(x, K)$ which agrees with what we expect given that the interior angle disappears and the marker vanishes.
Finally, since $D^2_\lambda(\cdot, K)$ is Hausdorff-Lipschitz continuous, it is easy to see that so is $V^d_\lambda(x, K)$.

### 3.3 Stable Multiscale Intersection Transform of Smooth Manifolds

Rather than devising an ad-hoc function that embeds the geometric features as its singularities, one can suitably modify the landscape of the characteristic function of the object and generate singularities which are localised in a neighborhood of the geometric feature of interest. This is for...
instance the rationale behind the transformation introduced in [122]. The objective is to obtain a Hausdorff stable multiscale method that is robust with respect to sampling, so that it can be applied to geometric objects represented by point clouds, and that is able to describe possible hierarchy of features as defined in terms of some characteristic geometric property. If we denote by \( K \subset \mathbb{R}^n \) the union of finitely many smooth compact manifolds \( M_k \), for \( k = 1, \ldots, m \), in this section we are interested to extract two types of types of geometric singularities:

(i) Transversal surface-to-surface intersections.

(ii) Boundary points shared by two smooth manifolds.

These problems are studied extensively in computer-aided geometric design under the general terminology of shape interrogation [85]. The traditional approach to surface-to-surface intersection problems is to consider parameterized polynomial surfaces and to solve systems of algebraic equations numerically based on real algebraic geometry [85]. The application of these methods typically requires some topological information such as triangle mesh connectivity or a parameterization of the geometrical objects, hence they are difficult to implement in the cases of free-form surfaces and of manifolds represented, for instance, by point clouds. For the latter case, other types of approaches are usually used which aim at identifying, according to some criteria, the points that are likely to belong to a neighborhood of the sharp feature. State-of-art methods currently in use are mostly justified by numerical experiments, and their stability properties, under dense sampling of the set \( M \), are not known. Let \( K \subset \mathbb{R}^n \) be a non-empty compact set. By using compensated convex transforms we introduced the intersection extraction transform of scale \( \lambda > 0 \) [122] by

\[
I_\lambda(x; K) = \left| C_{4\lambda}^u(\chi_K)(x) - 2\left( C_{4\lambda}^u(\chi_K)(x) - C_{4\lambda}^l(\chi_K)(x) \right) \right|, \quad x \in \mathbb{R}^n. \tag{3.20}
\]

By recalling the definition of the stable ridge transform \([3.12]\) of scale \( \lambda \) and \( \tau \) for the characteristic function \( \chi_K \), \( I_\lambda(x; K) \) can be expressed in terms of \( \text{SR}_{\lambda, \tau}(\chi_K)(x) \) for \( \tau = \lambda \) as

\[
I_\lambda(x; K) = \left| C_{4\lambda}^u(\chi_K)(x) - 2\text{SR}_{\lambda, \lambda}(\chi_K)(x) \right|, \quad x \in \mathbb{R}^n. \tag{3.21}
\]

As instance of how \( I_\lambda(\cdot; K) \) is used to remove or filter regular points, Figure [16] illustrates the graphs of \( C_{\lambda}^u(\chi_{K_{a=1}})(x) \), \( C_{\lambda}^u(C_{\lambda}^u(\chi_{K_{a=1}}))(x) \) and of the filter \( I_\lambda(\cdot; K_{a=1}) \) in the case of the intersection of two lines perpendicular to each other. This example can be generalized to ‘regular directions’ and ‘regular points’ on manifolds \( K \) and verify that \( I_\lambda(x; K) = 0 \) at these points. Let \( K \subset \mathbb{R}^n \) be a non-empty compact set and \( e \) is a \( \delta \)-regular direction of \( x \in K \), then \( I_\lambda(y; K) = 0 \) for \( y \in [x - \delta e, x + \delta e] := \{ x + t\delta e, \ -1 \leq t \leq 1 \} \) when \( \lambda \geq 1/\delta^2 \). In particular, we have that at the point \( x \),

\[
C_{\lambda}^u(\chi_K)(x) = 1/2. \tag{3.22}
\]

If \( K \) is a \( C^1 \) manifold in a neighbourhood of \( x \in K \) and \( x \) is a \( \delta \)-regular point of \( K \), then \( I_\lambda(y; K) = 0 \) if \( y - x \in N_x \) and \( |y - x| \leq \delta \). Since \( C_{\lambda}^u(\chi_K)(x) = 1 \) for \( x \in K \), by using \( I_\lambda(\cdot; K) \), we have that the regular points will be removed by the transform itself, leaving only points near the singular ones. In this context, for compact \( C^2 \) \( m \)-dimensional manifolds with \( 1 \leq m \leq n - 1 \), since \( I_\lambda(y; K) = 0 \) for all \( \delta \)-regular points \( y \in K \) when \( \lambda > 0 \) is sufficiently large, the condition \( I_\lambda(y; K) = 0 \) can thus be used to define singular points which can be extracted by \( I_\lambda(\cdot; K) \) if there exists a constant \( c_x > 0 \), depending at most only on \( x \), such that \( I_\lambda(x; K) \geq c_x > 0 \) for sufficiently large \( \lambda > 0 \).

From the definition (3.21) of \( I_\lambda(\cdot; K) \) in terms of the stable ridge transform and of the upper transform of the characteristic function of the manifold \( K \), since such transforms are Hausdorff stable, it follows that \( I_\lambda(\cdot; K) \) is also Hausdorff stable, that is, for \( E, F \) non-empty compact subsets
of $\mathbb{R}^n$ and $\lambda > 0$, then there holds
\[ |I_\lambda(x; E) - I_\lambda(x; F)| \leq 12\sqrt{\lambda} \operatorname{dist}_E(E, F), \quad x \in \mathbb{R}^n. \] (3.23)

### 3.4 Stable Multiscale Medial Axis Map

The medial axis of an object is a geometric structure introduced by Blum [24] as a means of providing a compact representation of a shape. Initially defined as the set of the shock points of a grass fire lit on the boundary and required to propagate uniformly inside the object. Closely related definitions of skeleton [29] and cut-locus [114] have since been proposed, and have served for the study of its topological properties [2, 4, 30, 73, 75, 97], its stability [40, 38] and for the development of fast and efficient algorithms for its computation [1, 15, 14, 64, 81]. Hereafter we refer to the definition given in [67]. For a given non-empty closed set $K \subset \mathbb{R}^n$, with $K \neq \mathbb{R}^n$, we define the medial axis $M_K$ of $K$ as the set of points $x \in \mathbb{R}^n \setminus K$ such that $x \in M_K$ if and only if there are at least two different points $y_1, y_2 \in K$, satisfying $\operatorname{dist}(x; K) = |x - y_1| = |x - y_2|$, whereas for a non-empty bounded open set $\Omega \subset \mathbb{R}^n$, the medial axis of $\Omega$ is defined by $M_{\Omega} := \Omega \cap M_{\partial \Omega}$.

The application of the lower transform to study the medial axis $M_K$ of a set $K$ is motivated by the identification of the medial axis with the singularity set of the Euclidean distance function and by the geometric structure of this set [4, 30, 73]. However, for our setting, it is more convenient to consider the squared distance function and to use the identification of the singular set of the squared distance function with the set of points where the squared distance function fails to be locally $C^1$. Since the lower compensated convex transform to the Euclidean squared-distance function gives a smooth ($C^1$) tight approximation outside a neighbourhood of the closure of the medial axis, in [118] the quadratic multiscale medial axis map with scale $\lambda > 0$ is defined as a scaled difference between the squared-distance function and its lower transform, that is,
\[
M_\lambda(x; K) := (1 + \lambda)R_\lambda(\operatorname{dist}^2(\cdot; K))(x) = (1 + \lambda)\left(\operatorname{dist}^2(x; K) - C_\lambda^1(\operatorname{dist}^2(\cdot; K))(x)\right), \tag{3.24}
\]
whereas for a bounded open set $\Omega \subset \mathbb{R}^n$ with boundary $\partial \Omega$, the quadratic multiscale medial axis
map of $\Omega$ with scale $\lambda > 0$ is defined by

$$M_\lambda(x; \Omega) := M_\lambda(x; \partial \Omega) \quad x \in \Omega.$$  

A direct consequence of the definition of $M_\lambda(x, K)$ is that for $x \in \mathbb{R}^n \setminus M_K$ we have

$$\lim_{\lambda \to \infty} M_\lambda(x, K) = 0,$$  

(3.25)

and the limit map $M_\infty(x, K)$ presents well separated values, in the sense that they are zero outside the medial axis and remain strictly positive on it. To gain an insight of the geometric structure of $M$ and the limit map between the two nonzero vectors $y_1 - x$ and $y_2 - x$ for $y_1, y_2 \in K(x)$, then

$$\theta_x = \max\{\angle[y_1 - x, y_2 - x], \; y_1, y_2 \in K(x)\}. \quad (3.26)$$

By means of this geometric parameter, it was shown in [118] that for every $\lambda > 0$ and $x \in M_K$ that

$$\sin^2(\theta_x/2) \text{dist}^2(x, K) \leq M_\lambda(x, K) \leq \text{dist}^2(x, K). \quad (3.27)$$

This result along with the examination of prototype examples ensures that the multiscale medial axis map of scale $\lambda$ keeps a constant height along the part of the medial axis generated by a two-point subset, with the value of the height depending on the distance between the two generating points. Such values can, therefore, be used to define a hierarchy between different parts of the medial axis and one can thus select the relevant parts through simple thresholding, that is, by taking suplevel sets of the multiscale medial axis map, justifying the the word "multiscale" in its definition. For each branch of the medial axis, the multiscale medial axis map automatically defines a scale associated with it. In other words, a given branch has a strength which depends on some geometric features of the part of the set that generates that branch.

An inherent drawback of the medial axis $M_K$ is in fact its sensitivity to boundary details, in the sense that small perturbations of the object (with respect to the Hausdorff distance) can produce huge variations of the corresponding medial axis. This does not occur in the case of the quadratic multiscale medial axis map, given that [118] shifts somehow the focus from the support of $M_\lambda(\cdot, K)$ to the whole map. Let $K, L \subset \mathbb{R}^n$ denote non-empty compact sets and $\mu := \text{dist}_{\mathcal{H}}(K, L)$, it was shown in [118] that for $x \in \mathbb{R}^n$, we have

$$\left|M_\lambda(x; K) - M_\lambda(x; L)\right| \leq \mu(1 + \lambda)(\text{dist}(x; K) + \mu)^2 + 2\text{dist}(x; K) + 2\mu + 1. \quad (3.28)$$

While the medial axis of $K$ is not a stable structure with respect to the Hausdorff distance, its medial axis map $M_\lambda(x; K)$ is by contrast a stable structure. This result complies with (3.28) which shows that as $\lambda$ becomes large, the bound in (3.28) becomes large.

With the aim of giving insights into the implications of the Hausdorff stability of $M_\lambda(x; \partial \Omega)$, we display in Figure [17] the graph of the multiscale medial axis map of a non-convex domain $\Omega$ and of an $\epsilon$-sample $K_\epsilon$ of its boundary. An inspection of the graph of $M_\lambda(x; \partial \Omega)$ and $M_\lambda(x; K_\epsilon)$, displayed in Figure (a) and Figure (b), reveals that both functions take comparable values along the main branches of $M_\Omega$. Also, $M_\lambda(x; K_\epsilon)$ takes small values along the secondary branches, generated by the sampling of the boundary of $\Omega$. These values can therefore be filtered out by a simple thresholding so that a stable approximation of the medial axis of $\Omega$ can be computed. This can be appreciated by looking at Figure (d), which displays a suplevel set of $M_\lambda(x; K_\epsilon)$ that
appears to be a reasonable approximation of the support of \( M_\lambda(x; \partial \Omega) \) shown in Figure 17(c).

**Figure 17** Multiscale Medial Axis Map of a nonconvex domain \( \Omega \) and of an \( \epsilon \)-sample \( K_\epsilon \) of its boundary. (a) Nonconvex domain \( \Omega \); (b) Graph of \( M_\lambda(\cdot; \partial \Omega) \) for \( \lambda = 2.5 \); (c) Support of \( M_\lambda(\cdot; \partial \Omega) \); (d) Graph of \( M_\lambda(\cdot; K_\epsilon) \); (e) Support of \( M_\lambda(\cdot; \Omega) \); (f) Suplevel set of \( M_\lambda(x; K_\epsilon) \) for a threshold equal to 0.15 \( \max_{x \in \mathbb{R}^2} \{ M_\lambda(x; K_\epsilon) \} \).

A relevant implication of (3.28) concerns with the continuous approximation of the medial axis of a shape starting from subsets of the Voronoi diagram of a sample of the shape boundary which is pertinent for shape reconstruction from point clouds. Let us consider an \( \epsilon \)-sample \( K_\epsilon \) of \( \partial \Omega \), that is, a discrete set of points such that \( \text{dist}_H(\partial \Omega, K_\epsilon) \leq \epsilon \). Since the medial axis of \( K_\epsilon \) is the Voronoi diagram of \( K_\epsilon \), if we denote by \( V_\epsilon \) the set of all the vertices of the Voronoi diagram \( \text{Vor}(K_\epsilon) \) of \( K_\epsilon \), and denote by \( P_\epsilon \) the subset of \( V_\epsilon \) formed by the ‘poles’ of \( \text{Vor}(K_\epsilon) \) introduced in [11], (i.e. those vertices of \( \text{Vor}(K_\epsilon) \) that converge to the medial axis of \( \Omega \) as the sample density approaches infinity), then, for \( \lambda > 0 \), it was established in [118] that

\[
\lim_{\epsilon \to 0^+} M_\lambda(x_\epsilon; K_\epsilon) = 0 \quad \text{for} \quad x_\epsilon \in V_\epsilon \setminus P_\epsilon .
\]

Since as \( \epsilon \to 0^+ \), \( K_\epsilon \to \partial \Omega \), and knowing that \( P_\epsilon \to M_\Omega \) [12, 25], then on the vertices of \( \text{Vor}(K_\epsilon) \) that do not tend to \( M_\Omega \), \( M_\lambda(x_\epsilon; K_\epsilon) \) must approach zero in the limit because of (3.25). As a result, in the context of the methods of approximating the medial axis starting from the Voronoi diagram of a sample set (such as those described in [12, 47, 48, 100]), the use of the multiscale medial axis map offers an alternative and much easier tool to construct continuous approximations to the medial axis with guaranteed convergence as \( \epsilon \to 0^+ \).

We conclude this topic by showing how compensated convex transform is used to obtain a fine
result of geometric measure theory. Let us introduce the set \( V_{\lambda,K} \) defined as
\[
V_{\lambda,K} = \{ x \in \mathbb{R}^n : \lambda \text{dist}(x; M_K) \leq \text{dist}(x; K) \},
\]
which represents a neighborhood of \( M_K \). From the property of the tight approximation of the lower transform of the squared-distance function, it was shown in \[118\] that
\[
\text{dist}^2(\cdot; K) \in C^{1,1}(\mathbb{R}^n \setminus V_{\lambda,K}),
\]
and a sharp estimate for the Lipschitz constant of \( D\text{dist}^2(\cdot, K) \) was also obtained. This result can be viewed as a weak Lusin type theorem for the squared-distance function which extends regularity results of the squared-distance function to any closed non-empty subset of \( \mathbb{R}^n \).

3.5 Approximation Transform

The theory of compensated convex transforms can also be applied to define Lipschitz continuous and smooth geometric approximations and interpolations for bounded real-valued functions sampled from either a compact set \( K \) in \( \mathbb{R}^n \) or the complement of a bounded open set \( \Omega \), i.e. \( K = \mathbb{R}^n \setminus \Omega \). The former is motivated by approximating or interpolating sparse data and/or contour lines whereas the latter by the so-called inpainting problem in image processing \[37\], where some parts of the image content are missing. The aim of ‘inpainting’ is to use other information from parts of the image to repair or reconstruct the missing parts.

Let \( f : \mathbb{R}^n \to \mathbb{R} \) denote the underlying function to be approximated, \( f_K : K \subset \mathbb{R}^n \to \mathbb{R} \) the sampled function defined by \( f_K(x) = f(x) \) for \( x \in K \), and \( \Gamma_{f_K} := \{(x,f_K(x)), x \in K\} \) its graph, the setting for the application of the compensated convex transforms to obtain an approximation transform is the following. Given \( M > 0 \), we define first two functions extending \( f_K \) to \( \mathbb{R}^n \setminus K \), namely
\[
\begin{align*}
f_K^{-M}(x) &= f(x)\chi_K(x) - M\chi_{\mathbb{R}^n \setminus K} = \begin{cases} f_K(x), & x \in K, \\ -M, & x \in \mathbb{R}^n \setminus K; \end{cases} \\
f_K^M(x) &= f(x)\chi_K(x) + M\chi_{\mathbb{R}^n \setminus K} = \begin{cases} f_K(x), & x \in K, \\ M, & x \in \mathbb{R}^n \setminus K, \end{cases}
\end{align*}
\]

where \( \chi_G \) denotes the characteristic function of a set \( G \). We then compute the arithmetic average of the proximal hull of \( f_K^M(x) \) and the upper proximal hull of \( f_K^{-M} \) as follows,
\[
A^M_\lambda(f_K)(x) = \frac{1}{2} \left( C^M_\lambda(f_K^M)(x) + C^M_\lambda(f_K^{-M})(x) \right), \quad x \in \mathbb{R}^n,
\]
which we refer to as the average compensated convex approximation transform of \( f_K \) of scale \( \lambda \) and level \( M \) \[119\].

In the case that \( K \subset \mathbb{R}^n \) is a compact set and \( f : \mathbb{R}^n \to \mathbb{R} \) is bounded and uniformly continuous, error estimates are available for \( M \to \infty \) and for \( x \in \text{co}[K] \). If for \( x \in \text{co}[K] \setminus K \) we denote by \( r_c(x) \) the convex density radius as the smallest radius of a closed ball \( B(x; r_c(x)) \) such that \( x \) is in the convex hull of \( K \cap B(x; r_c(x)) \), then for \( \lambda > 0 \) and all \( x \in \text{co}[K] \) there holds
\[
|A^\infty_\lambda(f_K)(x) - f(x)| \leq \omega \left( r_c(x) + \frac{a}{\lambda} + \sqrt{\frac{2b}{\lambda}} \right), \tag{3.33}
\]
where \( \omega = \omega(t) \) is the least concave majorant of the modulus of continuity \( \omega_f \) of \( f \) and \( a \geq 0, b \geq 0 \).
are such that $\omega(t) \leq at + b$ for $t \geq 0$. Error estimates are also available for a finite $M > 0$ under the extra restriction that $f(x) = c_0$ for $|x| \geq r$ where $c_0 \in \mathbb{R}$ and $r > 0$ are constants. In this case, for $R > r$, we extend $f_K$ to be equal to $c_0$ outside a large ball $B(0; R)$ containing $K$ and define $K_R = K \cup B^c(0; R)$. Thus we obtain similar error estimates to (3.33) for $A^M_\lambda(f_{K_R})(x)$. Furthermore, we have that when $M > 0$ is sufficiently large, $A^M_\lambda(f_K)$ approaches $f_K$ in $K$ as $\lambda \to \infty$, whereas if $f$ is a $C^{1,1}$ function and $\lambda > 0$ is large enough, $A^M_\lambda(f_K)$ is an interpolation of $f$ in the convex hull $\co[K]$ of $K$. In the special case of a finite set $K$, the average approximation $A^M_\lambda(f_K)$ defines an approximation for the scattered data $\Gamma_{f_K} = \{(x, f_K(x)), x \in K\}$.

If the closed set $K$ is the complement of a non-empty bounded open set $\Omega \subset \mathbb{R}^n$, we can also obtain estimates that are similar to (3.33). Clearly, $\co[K] = \mathbb{R}^n$ for such a $K$, thus if $f : \mathbb{R}^n \to \mathbb{R}$ is bounded and uniformly continuous, satisfying $|f(x)| \leq A_0$ for some constant $A_0 > 0$ and for all $x \in \mathbb{R}^n$ and $d_\Omega$ denotes the diameter of $\Omega$, then for $\lambda > 0$, $M > A_0 + \lambda d_\Omega^2$ and all $x \in \mathbb{R}^n$, we have

$$|A^M_\lambda(f_K)(x) - f(x)| \leq \omega \left( r_e(x) + \frac{a}{\lambda} + \sqrt{\frac{2b}{\lambda}} \right),$$

where, as for (3.33), the constants $a \geq 0$ and $b \geq 0$ are such that $\omega(t) \leq at + b$ for $t \geq 0$ with $\omega = \omega(t)$ the least concave majorant of the modulus of continuity $\omega_f$ of $f$.

Both the estimates (3.33) and (3.34) can be improved for Lipschitz functions and for $C^{1,1}$ functions.

Another natural and practical question in data approximation and interpolation is the stability of a given method. For approximations and interpolations of sampled functions, we would like to know, for two sample sets which are ‘close’ to each other, say, under the Hausdorff distance [10], whether the corresponding approximations are close to each other. It is easy to see that differentiation and integration based approximation methods are not Hausdorff stable because continuous functions can be sampled over a finite dense set. One of the advantages of the compensated convex approximation is that for a bounded uniformly continuous function $f$, and for fixed $M > 0$ and $\lambda > 0$, the mapping $K \mapsto A^M_\lambda(f_K)(x)$ is continuous with respect to the Hausdorff distance for compact sets $K$, and the continuity is uniform with respect to $x \in \mathbb{R}^n$. This means that if another sampled subset $E \subset \mathbb{R}^n$ (finite or compact) is close to $K$, then the output $A^M_\lambda(f_E)(x)$ is close to $A^M_\lambda(f_K)(x)$ uniformly with respect to $x \in \mathbb{R}^n$. As far as we know, not many known interpolation/approximation methods share such a property.

By using the mixed compensated convex transforms [116], it is possible to define a mixed average compensated convex approximation with scales $\lambda > 0$ and $\tau > 0$ for the sampled function $f_K : K \to \mathbb{R}$ by

$$(SA)^M_{\tau, \lambda}(f_K)(x) = \frac{1}{2}(C^u_{\tau}(C^M_{\lambda}(f_K))(x) + C^d_{\tau}(C^u_{\lambda}(f_K^{-M}))(x)), \quad x \in \mathbb{R}^n. \quad (3.35)$$

Since the mixed compensated convex transforms are $C^{1,1}$ functions [116 Theorem 2.1(iv) and Theorem 4.1(ii)], the mixed average approximation $(SA)^M_{\tau, \lambda}$ is a smooth version of our average approximation. Also, for a bounded function $f : \mathbb{R}^n \to \mathbb{R}$, satisfying $|f(x)| \leq M$, $x \in \mathbb{R}^n$ for some constant $M > 0$, we have the following estimates [122 Theorem 3.13]

$$0 \leq C^u_{\tau}(C^d_{\lambda}(f))(x) - C^d_{\lambda}(f)(x) \leq \frac{16M\lambda}{\tau}, \quad 0 \leq C^u_{\lambda}(f)(x) - C^d_{\tau}(C^u_{\lambda}(f))(x) \leq \frac{16M\lambda}{\tau}$$

29
for all $x \in \mathbb{R}^n$, $\lambda > 0$ and $\tau > 0$, and hence can easily show that for any closed set $K \subset \mathbb{R}^n$,

$$
| (SA)^M_{r,\lambda}(f_K)(x) - A^M_{\lambda}(f_K)(x) | \leq \frac{16M\lambda}{\tau}, \quad x \in \mathbb{R}^n.
$$

This implies that for given $\lambda > 0$ and $M > 0$, the mixed approximation $(SA)^M_{r,\lambda}(f_K)$ converges to the basic average approximation $A^M_{\lambda}(f_K)$ uniformly in $\mathbb{R}^n$ as $\tau \to \infty$, with rate of convergence $16M\lambda/\tau$.

### 4 Numerical Algorithms

The numerical realisation of the convex transforms introduced in Section 3 relies on the availability of numerical schemes for computing the upper and lower transforms of a given function. Because of the relation (1.4) between the upper and lower transform, the computation of the above transforms ultimately boils down to the evaluation of the lower compensated convex transform. As a result, without loss of generality, in the following we refer just to the actual implementation of $C^l_{\lambda}(f)$. With this respect, we can proceed in two different ways according to whether we use definition (1.2) in terms of the convex envelope or the characterization (1.5) as proximity hull of the function and use its definition in terms of the Moreau envelopes. In the following, we describe some algorithms that can be used successfully for the computation of $C^l_{\lambda}(f)$ and discuss their relative merits. Figure 18 summarizes the different approaches considered in this paper.

**Figure 18** Different approaches for computing the Lower Compensated Convex transform $C^l_{\lambda}(f)$.

#### 4.1 Convex based algorithms

Algorithms to compute convex hull such as the ones given in [35, 22] are more suitable for discrete set of points and their complexity is related to the cardinality of the set. An adaptation of these methods to our case, with the set to convexify given by the epigraph of $f + \lambda | \cdot |^2$, does not appear to be very effective, especially for functions defined in subsets of $\mathbb{R}^n$ for $n \geq 2$, compared to the methods that (directly) compute the convex envelope of a function [109, 27, 80, 42].

Of particular interest for applications to image processing, where functions involved are defined on grid of pixels, is the characterization of the convex envelope as the viscosity solution of a nonlinear obstacle problem [80]. An approximated solution is then obtained by using centered finite differences along directions defined by an associated stencil to approximate the first eigenvalue of the Hessian matrix at the grid point. A generalization of the scheme introduced in [80] in terms of the number of convex combinations of the function values at the grid points of the stencil, is briefly summarized in Algorithm 1 and described below. Given a uniform grid of points $x_k \in \mathbb{R}^n$, equally spaced with grid size $h$, denote by $S_{x_k}$ the $d$–point stencil of $\mathbb{R}^n$ with center at $x_k$ defined as $S_{x_k} = \{ x_k + hr, |r|_\infty \leq 1, r \in \mathbb{Z}^n \}$ with $| \cdot |_\infty$ the $\ell^\infty$-norm of $r \in \mathbb{Z}^n$ and $d = \#(S)$, cardinality
of the finite set $S$. At each grid point $x_k$ we compute an approximation of the convex envelope of $f$ at $x_k$ by an iterative scheme where each iteration step $m$ is given by

$$(\text{co } f)_m(x_k) = \min \{ f(x_k), \sum \lambda_i (\text{co } f)_{m-1}(x_i) : \sum \lambda_i = 1, \lambda_i \geq 0, x_i \in S_{x_k} \}$$

with the minimum taken between $f(x_k)$ and only some convex combinations of $(\text{co } f)_{m-1}$ at the stencil grid points $x_i$ of $S_{x_k}$. It is then not difficult to show that the scheme is monotone, thus convergent. However, there is no estimate of the rate of convergence which, in actual applications, appears to be quite slow. Furthermore, results are biased by the type of underlying stencil.

**Algorithm 1** Computation of the convex envelope of $f$ according to [80]

1: Set $m = 1$, $(\text{co } f)_0 = f$, $tol$
2: $\epsilon = \|f\|_{L^2}$
3: while $\epsilon > tol$ do
4: $\forall x_k$, $(\text{co } f)_m(x_k) = \min \{ f(x_k), \sum \lambda_i (\text{co } f)_{m-1}(x_i) : \sum \lambda_i = 1, \lambda_i \geq 0, x_i \in S_{x_k} \}$
5: $\epsilon = \|(\text{co } f)_m - (\text{co } f)_{m-1}\|_{L^2}$
6: $m \leftarrow m + 1$
7: end while

Based on the characterization of the convex envelope of $f$ in terms of the biconjugate $(f^*)^*$ of $f$ [23, 58, 88], where $f^*$ is the Legendre-Fenchel transform of $f$, we can approximate the convex envelope by computing twice the discrete Legendre-Fenchel transform. We can thus improve speed efficiency with respect to a brute force algorithm, which computes $(f^*)^*$ with complexity $O(N^2)$ with $N$ the number of grid points, if we have an efficient scheme to compute the discrete Legendre-Fenchel transform of a function. For functions $f : X \rightarrow \mathbb{R}$ defined on cartesian sets of the type $X = \prod_{i=1}^n X_i$ with $X_i$ intervals of $\mathbb{R}$, $i = 1, \ldots, n$, the Legendre-Fenchel transform of $f$ can be reduced to the iterate evaluation of the Legendre-Fenchel transform of functions dependent only on one variable as follows

$$(\xi_1, \xi_2, \ldots, \xi_n) \in \mathbb{R}^n \rightarrow f^*(\xi_1, \ldots, \xi_n) = \sup_{x \in X} \{ \xi \cdot x - f(x) \}$$

$$= \sup_{x_1, \ldots, x_n - 1 \in \prod_{i=1}^{n-1} X_i} \left\{ x_1 \xi_1 + \ldots + x_{n-1} \xi_{n-1} - \sup_{x_n \in X_n} \{ x_n \xi_n - f(x_1, \ldots, x_{n-1}, x_n) \} \right\}.$$  \hspace{1cm} (4.1)

As a result, one can improve the complexity of the computation of $f^*$ if one has an efficient scheme to compute the Legendre-Fenchel transform of functions of only one variable. For instance, the algorithm described in [70, 57], which exploits an idea of [26] and improves the implementation of [43], computes the discrete Legendre-Fenchel transform in linear time, that is, with complexity $O(N)$. If $g_h$ denote the grid values of a function of one variable, the key idea of [26, 43] is to compute $(g_h)^*$ as approximation of $g^*$ using the following result

$$(g_h)^*(\xi) = (\text{co} [\Pi f_h])^*(\xi), \quad \xi \in \mathbb{R}$$  \hspace{1cm} (4.2)

where $\Pi g_h$ denotes the continuous piecewise affine interpolation of the grid values $g_h$. Therefore by applying an algorithm with linear complexity, for instance the beneath-beyond algorithm [86], to compute the convex envelope $\text{co} [\Pi g_h]$, followed by the use of analytical expressions for the Legendre-Fenchel transform of a convex piecewise affine function yields an efficient method to compute $(g_h)^*$
For functions defined in a bounded domain, in [70] it was recommended to increment the size of the domain for a better precision of the computation of the Legendre-Fenchel transform. The work [57] avoids this by elaborating the exact expression of the Legendre-Fenchel transform of a convex piecewise affine function defined in a bounded domain which is equal to infinity in $\mathbb{R} \setminus X$ or it has an affine variation. In this manner, they can avoid boundary effects. For ease of reference, we report next the analytical expression of $g^*$ in the case where $g : \mathbb{R} \to \mathbb{R}$ is convex piecewise affine.

Without loss of generality, let $x_1 < \ldots < x_N$ be a grid of points of $\mathbb{R}$, $c_1 < \ldots < c_N$ and assume $g : \mathbb{R} \to \mathbb{R}$ to be defined as follows:

$$
g : x \in \mathbb{R} \rightarrow
g_1 + c_1(x_1 - x) \quad \text{if} \quad x_1 \leq x \leq x_i, \quad i = 1, \ldots, N - 1
g_i + c_i(x_i - x) \quad \text{if} \quad x_i \leq x \leq x_{i+1}, \quad i = 1, \ldots, N - 1
g_N + c_N(x_N - x) \quad \text{if} \quad x \geq x_N$$

(4.3)

where $g_i = g(x_i)$ and $c_i$, for $i = 1, \ldots, N$, represent the slopes of each affine piece of $g$. It is not difficult to verify that the analytical expression of $g^*$ is given by [57]

$$
g^* : \xi \in \mathbb{R} \rightarrow
g^*_1 \xi - g_1 \quad \text{if} \quad \xi \leq c_1
g_{i+1} \xi - g_{i+1} \quad \text{if} \quad c_i \leq \xi \leq c_{i+1}, \quad i = 1, \ldots, N - 2
g_N + c_N \quad \text{if} \quad \xi \geq c_N.$$  

(4.4)

Once we know $g^*$, using the decomposition (4.1), we can compute $f^*$ and thus the biconjugate $f^{**}$.

### 4.2 A Moreau envelope based algorithm

The computation of the Moreau envelope is an established task in the field of computational convex analysis [72] that has been tackled by various different approaches aimed at reducing the quadratic complexity of a direct brute force implementation of the transform. Such reduction is achieved, one way or another, by a dimensional reduction. The fundamental idea of the scheme presented in [124], for instance, is the generalization of the Euclidean distance transform of binary images, by replacing the binary image by an arbitrary function on a grid. The decomposition of the structuring element which yields the exact Euclidean distance transform [99] into basic ones, leads to a simple and fast algorithm where the discrete lower Moreau envelope can be computed by a sequence of local operations, using one-dimensional neighborhoods. Unless otherwise stated, in the following, $i, j, k, r, s, p, q \in \mathbb{Z}$ denote integers whereas $m, n \in \mathbb{N}$ are non-negative integers. Given $n \geq 1$, we introduce grid of points of the space $\mathbb{R}^n$ with regular spacing $h > 0$ denoted by $x_k \in \mathbb{R}^n$, $k \in \mathbb{Z}$ and define the discrete lower Moreau envelope at $x_k \in \mathbb{R}^n$ as

$$
M^h \lambda(f)(x_k) = \inf \{ f(x_k + rh) + \lambda h^2 \| r \|^2, \; r \in \mathbb{Z}^n \}.
$$

(4.5)

By taking the infimum in (4.5) over a finite number $m \geq 1$ of directions, we obtain the $m$–th approximation of the discrete Moreau lower envelope $M^h \lambda(f)(x_k)$ which can be evaluated by taking the values $f_m(x_k)$ given by Algorithm 2. For the convergence analysis and convergence rate we refer to [124] where it is shown that the scheme has a linear convergence rate with respect to $h$. 

32
Algorithm 2  Computation of $f_m(x_k)$ at the points $x_k$ of the grid of $\mathbb{R}^n$ of size $h$ for given $m \geq 1$.

1: Set $i = 1$, $m \in \mathbb{N}$
2: $\forall x_k$, $f_0(x_k) = f(x_k)$
3: while $i < m$ do
4: $\tau_i = 2i - 1$
5: $f_i(x_k) = \min\{f_{i-1}(x_k + rh) + \lambda h^2 r^2 \tau_i : r \in \mathbb{Z}^n, |r|_\infty \leq 1 \}$
6: $i \leftarrow i + 1$
7: end while

Likewise the computation of the Legendre-Fenchel transform, in the scheme proposed by [49], the authors apply the dimensional reduction directly to the computation of the Moreau envelope which is factored by $n$ one dimensional Moreau envelope. For instance, in the case of $n = 2$, let $\Omega = X \times Y$, with $X, Y \subset \mathbb{R}$, and $(\xi_1, \xi_2) \in \Omega = X \times Y$, for any $x = (x_1, x_2) \in \mathbb{R}^2$, we have

$$M_\lambda(f)(x_1, x_2) = \inf_{(\xi_1, \xi_2) \in \Omega} \{\lambda|(x_1, x_2) - (\xi_1, \xi_2)|^2 + f(\xi_1, \xi_2)\}$$

$$= \inf_{\xi_1 \in X} \{\lambda|x_1 - \xi_1|^2 + \inf_{\xi_2 \in Y}\{\lambda|x_2 - \xi_2|^2 + f(\xi_1, \xi_2)\}\}.$$  \hspace{1cm} (4.6)

For the computation of $M_\lambda(f)$ with $f$ function of one variable, if we denote by $\mathcal{F}$ the family of parabolas with given curvature $\lambda$ of the following type

$$p_q : x \in \mathbb{R} \to p_q(x) = \lambda|x - q|^2 + f(q),$$

parameterized by $q \in \Omega \subset \mathbb{R}$, we have that

$$M_\lambda(f)(x) = \inf \{p_q(x) : p_q \in \mathcal{F}\},$$ \hspace{1cm} (4.7)

that is, the Moreau envelope of a function of one variable is reduced to the computation of the envelope of parabolas of given curvature $\lambda$. The computation of such envelope is realised by [49] in two steps. In the first one, they compute the envelope by adding the parabolas one at time which is done in linear time, and comparing each parabola to the parabolas that realise the envelope, which is done in constant time, whereas in the second step they compute the value of the envelope at the given point $x \in \mathbb{R}$. The key points of the scheme result from two observations. The first one is that given any two parabolas of $\mathcal{F}$ parameterized by $q, r \in \Omega$, their intersection occurs only at one point with coordinate

$$x_s = \frac{(f(q) - f(r)) + \lambda(q^2 - r^2)}{2\lambda(q - r)},$$

whereas the second one regards the relation between the parabolas so that if $q < r$, then $p_q(x) \leq p_r(x)$ for $x < x_s$ and $p_q(x) \geq p_r(x)$ for $x > x_s$. This scheme allows the evaluation of $M_\lambda(f)(x)$ for any $x \in \mathbb{R}^n$ even if $f$ is defined only on a bounded open set $\Omega$, without any consideration on how to extend $f$ on $\mathbb{R}^n \setminus \Omega$. We will refer next to this scheme as the parabola envelope scheme.

By using the link between the Moreau envelope and the Legendre-Fenchel transform given by [90] [71]

$$M_\lambda(f)(x) = \lambda|x|^2 - 2\lambda \left(\frac{f}{2\lambda} + \frac{|.|^2}{2}\right)^*(x),$$  \hspace{1cm} (4.8)

it is possible to design another scheme to calculate the Moreau envelope by computing the Legendre-Fenchel transform of the augmented function that appears in [48] [71]. In this case, however, special
considerations must be taken about the primary domain, where the Moreau envelope is defined, and the dual domain, which is the one where the Legendre-Fenchel transform is defined.

5 Numerical Examples

In this section we present some illustrative numerical examples of implementation of the transforms introduced in Section 3. We proceed this discussion by the computation of a two-dimensional prototype example with analytical expression of $C^u_\lambda(\chi_K)$ which we use to select the most suitable numerical scheme out of those described in Section 4 for the computation of the compensated convex transforms.

5.1 Prototype Example: Upper transform of a singleton set of $\mathbb{R}^2$

Given the singleton set $K = \{0\} \subset \mathbb{R}^2$, the analytical expression of $C^u_\lambda(\chi_K)$ established in [123, Example 1.2] is given by

$$C^u_\lambda(\chi_K)(x) = \begin{cases} 0, & \text{if } |x| > 1/\sqrt{\lambda}, \\ \lambda(1/\sqrt{\lambda} - |x|)^2, & \text{if } |x| \leq 1/\sqrt{\lambda}. \end{cases} \quad (5.1)$$

We compute then $C^u_\lambda(\chi_K)$ by applying the convex based algorithms, i.e. Algorithm 1 [80] and the biconjugate based scheme (shorted as $BS$ hereafter) [70,57], and the Moreau based algorithms, i.e. Algorithm 2 and the parabola envelope scheme (shorted as $PES$ hereafter) [49]. To compare the accuracy of the schemes, we will consider: (i) the Hausdorff distance between the support of the exact and the computed upper transform,

$$e_H = \text{dist}_H \left( \mathcal{B}(0; 1/\sqrt{\lambda}), \text{sprt} \left( C^{u,h}_\lambda(\chi_K) \right) \right)$$

with $C^{u,h}_\lambda(\chi_K)$ the computed upper compensated transform; (ii) the relative $L^\infty$ error norm given by

$$e_{L^\infty} = \frac{\max_{x \in \mathbb{R}^2} |C^{u,h}_\lambda(\chi_K)(x) - C^u_\lambda(\chi_K)(x)|}{\max_{x \in \mathbb{R}^2} |C^u_\lambda(\chi_K)(x)|}$$

and (iii) the execution time $t_c$ in seconds by a PC with processor Intel® Core™ i7-4510U CPU@2.00 GHz and 8GB of memory RAM.
Figure 19  Supports of the exact and computed upper compensated transform of the characteristic function of a singleton set of $\mathbb{R}^2$ by the different numerical schemes. (a) Exact support given by $B(0; 1/\sqrt{\lambda})$ for $\lambda = 0.01$; (b) Support of $C^{u,h}_\lambda(\chi_K)$ computed by Algorithm 1 [80]; (c) Support of $C^{u,h}_\lambda(\chi_K)$ computed by the biconjugate based scheme [70, 57] for $h_d = 0.001$; (d) Support of $C^{u,h}_\lambda(\chi_K)$ computed by Algorithm 2 [124] which coincides with the one computed using the parabola envelope scheme [49].

Figure 19 displays the support of $C^{u}_\lambda(\chi_K)$ given by $B(0; 1/\sqrt{\lambda})$ and of $C^{u,h}_\lambda(\chi_K)$ computed by the numerical schemes mentioned above. Algorithm 2 and the parabola envelope algorithm yield the same results, thus Figure 19 displays the support as computed by only one of the two schemes. In this case we observe that the support coincides with the exact one. This does not happen for the support computed by the other two schemes. The application of Algorithm 1 evidences the bias of the scheme with the underlying stencil whereas by applying the biconjugate based scheme we note some small error all over the domain. The spread of this error depends on the dual mesh grid size $h_d$. Table 1 reports the values of $t_c$, $e_{L,\infty}$ and $d_{H}$ for the different schemes. For the biconjugate based scheme, we have different results according to the parameter $h_d$ that controls the uniform discretization of the dual mesh. The value $h_d = 1$ means that we are considering the same grid size as the grid of the input function $\chi_K$ whereas lower values for $h_d$ means that we are computing on a finer dual mesh compared to the primal one. The results given in Table 1 show that in terms of the values of $C^{u}_\lambda(\chi_K)$ the biconjugate based scheme is the one that produces the best results (compare the values of $e_{L,\infty}$), but this occurs at the fraction of cost of reducing $h_d$ which means to increase the number of the dual grid nodes and consequently the computational time. The issue of the choice of the dual grid on the accuracy of the computation of the convex envelope by the conjugate has been also tackled and recognized in [42]. However, as already pointed out in the analysis of Figure 19, the support of $C^{u,h}_\lambda(\chi_K)$ computed by the biconjugate scheme is the one to yield the worst value for $e_{H}$.

5.2 Intersection of Sampled Smooth Manifolds

In the following numerical experiments we verify the effectiveness of the filter $I_\lambda(\cdot; K)$ introduced in Section 3.3 and its Hausdorff stability property. We will consider both 2d— and 3d—geometries. The geometry is digitized and input as an image, but also other computer representations of the geometry can clearly be handled. This depends finally on the representation of the input geometry for the numerical scheme that is used to compute the compensated transforms. Figure 20 displays a road network extract from a map of the city of London, and represents a set of 2d curves which intersect to each other in different manner. The Figure shows the position of the local maxima of $I_\lambda(\cdot; K)$ which are seen to coincide with all the crossing and turning points of the given curves. We also have some false positive due to the digitization of the road network.
Algorithm 1

| Scheme                | $t_c$ | $e_{L,\infty}$ | $e_{H}$ |
|-----------------------|-------|-----------------|---------|
| Biconjugate scheme    | 1.9791| 0.0390          | 1.7321  |
| $h_d = 1$             | 0.1575| 48              | 9.4999  |
| $h_d = 0.1$           | 0.2157| 0.2400          | 9       |
| $h_d = 0.01$          | 0.5935| 0.0142          | 7.6158  |
| $h_d = 0.001$         | 16.6603| 0.0032         | 7.5498  |

Algorithm 2

| Scheme                | $t_c$ | $e_{L,\infty}$ | $e_{H}$ |
|-----------------------|-------|-----------------|---------|
| PE scheme             | 0.1246| 0.0249          | 0       |
| PE scheme             | 0.2553| 0.0249          | 0       |

Table 1 Comparison between the different numerical schemes for the computation of $C^\frac{d}{3}(\chi_K)$ for $\lambda = 0.01$. The symbol $h_d$ refers to the dual mesh size of the scheme that computes the convex envelope via the biconjugate.

Figure 20 (a) Medial axis of the road network; (b) Location of the intersection points; (c) Map of the road network and location of the intersection points shown in (b).

Figure 21 displays the results of the application of the filter $I_\lambda(\cdot; K)$ to 3d geometries represented by point clouds. Figure 21(a) displays the Plücker’s conoid of parametric equation

$$x = v \cos u, \quad y = v \sin u, \quad z = \sin 4u \quad \text{for} \ u \in [0, 2\pi], \ v \in [-1, 1],$$

with the location of its singular lines and the parts of surface with higher curvature. Figure 21(b) depicts the intersections between manifolds of different dimensions, namely, in the Figure, we have the Whitney umbrella of the implicit equation $x^2 = y^2 z$, a cylinder and an helix, with the location of their mutual intersections and also of where the Whitney surface intersects itself; finally, Figure 21(c) displays the intersection between a cylinder, planes and an helix.
Figure 21 (a) Plücker surface with identification of its singular lines and surface parts of higher curvatures; (b) Intersections of the Whitney surface of equation $x^2 = y^2z$ with an helix and a cylinder; (c) Intersections of planes with a cylinder and an helix.

The intersection of the line with the plane for the geometry shown in Figure 21 is weaker than the geometric singularities of the surfaces. With this meaning, the values of the local maxima of $I_\lambda(\cdot; K)$ determine a scale between the different type of intersections present in the manifold $K$ and represents the multiscale nature of the filter $I_\lambda(\cdot; K)$.

Finally, the numerical experiments displayed in Figure 22 refer to critical conditions that are not directly covered by the theoretical results we have obtained. Figure 22(a) shows the result of the application of $I_\lambda(\cdot; K)$ to a sphere and a cylinder that are ‘almost’ tangentially intersecting each other, whereas Figure 22(b) illustrates the results of the application of the filter to detect the intersection between loosely sampled piecewise affine functions, a plane and a line.

Figure 22 (a) Tangential intersection of a sampled sphere and cylinder which are ‘almost’ tangentially intersected, and indication of the intersection marker. (b) Intersection markers for the intersection among loosely sampled piecewise affine surfaces of equation $||10x - 75| - |10y - 75| + |10z - 75| - 45||=0$, the circle of equation $(10x - 75)^2 + (10z - 75)^2 \leq 45^2$ on the plane of equation $y = 75$ and the line of equation $x = 75, z = 75$. 
5.3 Approximation Transform

We report here on applications of the average approximation compensated convex transform developed in [119, 121] to three class of problems. These include (i) surface reconstruction from real world data using level lines and single points; (ii) Salt & Pepper noise restoration and (iii) image inpainting.

5.3.1 Level set reconstruction

We consider here the problem of producing a Digital Elevation Map from a sample of the the NASA SRTM global digital elevation model of Earth land. The data provided by the National Elevation Dataset [51] contain geographical coordinates (latitude, longitude and elevation) of points sampled at one arc-second intervals in latitude and longitude. For our experiments, we choose the region defined by the coordinates [N 40°23′25″, N 40°27′37″] × [E 14°47′25″, E 14°51′37″] extracted from the SRTM1 cell N40E014.hgt [103]. Such region consists of an area with extension 7.413 km × 5.844 km and height varying between 115 m and 1282 m, with variegated topography features. In the digitization by the US Geological Survey, each pixel represents a 30 m × 30 m patch. Figure 23(a) displays the elevation model from the SRTM1 data which we refer in the following to as the ground truth model. We will take a sample $f_K$ of such data, make the reconstruction using the $A^N(f_K)$ computed with Algorithm 1 and the AMLE interpolant [7, 33] using the MatLab® code described in [83], and compare them with the ground truth model.

Figure 23 Reconstruction of real-world digital elevation maps. (a) Ground truth model from USGS-STRM1 data relative to the area with geographical coordinates; [N 40°23′25″, N 40°27′37″] × [E 14°47′25″, E 14°51′37″]. (b) Sample set $K_1$ formed by only level lines at regular height interval of 58.35 m. The set $K_1$ contains 14% of the ground truth points. (c) Sample set $K_2$ formed by taking randomly 30% of the points belonging to the level lines of the set $K_1$ and scattered points corresponding to 5% density. The sample set $K_2$ contains 7% of the ground truth points.

In the numerical experiments, we consider two sample data, characterized by different data density and typo of information. The first, which we refer to as sample set $K_1$, consists only of level lines at regular height interval of 658.35 m and contains the 14% of the ground truth real digital data. The second sample set, denoted by $K_2$, has been formed by taking randomly the 30% of the points belonging to the level lines of the set $K_1$ and scattered points corresponding to 5% density so that the sample set $K_2$ amounts to about 7% of the ground truth points. The two sample sets $K_1$ and $K_2$ are shown in Figure 23(b) and Figure 23(c), respectively.
Figure 24 Reconstruction of real-world digital elevation maps. (a) Graph of $A^M_\lambda(f_K)$ for sample set $K_1$. Relative $L^2$-Errors: $\epsilon = 0.0118, \epsilon_K = 0$. Parameters: $\lambda = 2 \cdot 10^3, M = 1 \cdot 10^6$. Total number of iterations: 3818. (b) Graph of $A^M_\lambda(f_K)$ for sample set $K_2$. Relative $L^2$-Errors: $\epsilon = 0.0109, \epsilon_K = 0$. Parameters: $\lambda = 2 \cdot 10^3, M = 1 \cdot 10^6$. Total number of iterations: 1662. (c) Isolines of $A^M_\lambda(f_K)$ from sample set $K_1$ at regular heights of 58.35 m. (d) Isolines of $A^M_\lambda(f_K)$ from sample set $K_2$ at regular heights of 58.35 m.
Figure 25  Reconstruction of real-world digital elevation maps. (a) Graph of the AMLE Interpolant from set $K_1$. Relative $L^2$-Error: $\epsilon = 0.0410$, $\epsilon_K = 0.0110$. Total number of iterations: 11542. (b) Graph of the AMLE Interpolant from set $K_2$. Relative $L^2$-Error: $\epsilon = 0.02863$, $\epsilon_K = 0.0109$. Total number of iterations: 12457. (c) Isolines of the AMLE Interpolant from sample set $K_1$ at regular heights of 58.35 m. (d) Isolines of the AMLE Interpolant from sample set $K_2$ at regular heights of 58.35 m.

The graph of the $A^M_\lambda(f_K)$ interpolant and of the AMLE interpolant for the two sample sets along with the respective isolines at equally spaced heights equal to 58.35 m, are displayed in Figure 24 and Figure 25 respectively, whereas Table 2 contains the values of the relative $L^2$-error $\epsilon$ on $\Omega$ and $\epsilon_K$ on the sample set $K$ between such interpolants and the ground truth model, given by,

$$
\epsilon = \frac{\|f - A^M_\lambda(f_K)\|_{L^2(\Omega)}}{\|f\|_{L^2(\Omega)}} \quad \text{and} \quad \epsilon_K = \frac{\|f_K - A^M_\lambda(f_K)\|_{L^2(K)}}{\|f_K\|_{L^2(K)}},
$$

(5.2)

where $f$ is the ground truth model and $A^M_\lambda(f_K)$ is the average approximation of the sample $f_K$ of $f$ over $K$. We observe that while $A^M_\lambda(f_K)$ yields an exact interpolation of $f_K$ over $\Omega$, this is not the case for the AMLE approximation.

| Sample set | $A^M_\lambda(f_K)$ | AMLE | $A^M_\lambda(f_K)$ | AMLE |
|------------|-------------------|-----|-------------------|-----|
| $K_1$      | 0.0118            | 0.0410 | 0                | 0.0110 |
| $K_2$      | 0.0109            | 0.0286 | 0                | 0.0109 |

Table 2  Relative $L^2$-error for the DEM Reconstruction from the two sample sets using the $A^M_\lambda(f_K)$ and the AMLE interpolant. The realization of $\epsilon_K = 0$ for $A^M_\lambda(f_K)$ says that $A^M_\lambda(f_K)$ yields an exact interpolation of $f_K$ over $\Omega$, unlike the AMLE approximation.

Though both reconstructions are comparable visually to the ground truth model, a closer inspection of the pictures show that the reconstruction from the synthetic data, the AMLE interpolant
does not reconstruct correctly the mountains peaks, which appear to be smoothed, and introduce artificial ridges along the slopes of the mountains. In contrast, the $A^m_N(f_K)$ interpolant appears to better for capturing features of the ground truth model. Finally, we also note that though the sample set $K_1$ contains a number of ground truth points higher than the sample set $K_2$, the reconstruction from $K_2$ appears to be better than the one obtained from $K_1$. This behaviour was found for both interpolations, though it is more notable in the case of the $A^m_N(f_K)$ interpolant. By taking scattered data, we are able to get a better characterization of irregular surfaces, compared to the one obtained from a structured representation such as provided by the level lines.

5.3.2 Salt & Pepper Noise Removal

As an application of scattered data approximation to image processing, we consider here the restoration of an image corrupted by salt & pepper noise. This is an impulse type noise that is caused, for instance, by malfunctioning pixels in camera sensors or faulty memory locations in hardware, so that information is lost at the faulty pixels and the corrupted pixels are set alternatively to the minimum or to the maximum value of the range of the image values. When the noise density is low, about less than 40%, the median filter [13] or its improved adaptive median filter [59], is quite effective for restoring the image. However, this filter loses its denoising power for higher noise density given that details and features of the original image are smeared out. In those cases, other techniques must be applied; one possibility is the two-stage TV-based method proposed in [34] which consists of applying first an adaptive median filter to identify the pixels that are likely to contain noise and construct thus a starting guess which is used in the second stage for the minimization of a functional of the form

$$F(u, y) = \Psi(u, y) + \alpha \Phi(u)$$

where $y$ denotes the noisy image, $\Psi$ is a data-fidelty term and $\Phi$ is a regularization term, with $\alpha > 0$ a parameter. In the following numerical experiments, we consider the image displayed in Figure 26(a) with size 512 $\times$ 512 pixels, damaged by 70% salt & pepper noise. The resulting corrupted image is displayed in Figure 26(b) where on average only 78643 pixels out of the total 262144 pixels carry true information. The true image values represent our sample function $f_K$ whereas the set of the true pixels forms our sample set $K$. To assess the restoration performance we use the peak signal-to-noise ratio (PSNR) which is expressed in the units of dB and, for an 8−bit image, i.e. with values in the range [0, 255], is defined by

$$\text{PSNR} = 10 \log_{10} \frac{255^2}{m \times n \sum_{i,j} |f_{i,j} - r_{i,j}|^2}$$

where $f_{i,j}$ and $r_{i,j}$ denote the pixels values of the original and restored image, respectively, and $m$, $n$ denote the size of the image $f$. In our numerical experiments, we have considered the following cases. The first one assumes the set $K$ to be given by the noise-free interior pixels of the corrupted image together with the boundary pixels of the original image. In the second case, $K$ is just the set of the noise-free pixels of the corrupted image, without any special consideration on the image boundary pixels. In analysing this second case, to reduce the boundary effects produced by the application of Algorithm [1] and Algorithm [2], we have applied our method to an enlarged image and then restricted the resulting restored image to the original domain. The enlarged image has been obtained by padding a fixed number of pixels before the first image element and after the last image element along each dimension, making mirror reflections with respect to the boundary. The values used for padding are all from the corrupted image. In our examples, we have considered
two versions of enlarged images, obtained by padding the corrupted image with 2 pixels and 10 pixels, respectively. Table 3, Table 4 and Table 5 compare the values of the PSNR of the restored images by our method and the TV-based method applied to the corrupted image with noise-free boundary and to the two versions of the enlarged images with the boundary values of the enlarged images given by the padded noisy image data. We observe that there are no important variations in the denoising result between the different methods of treating the image boundary. This is also reflected by the close value of the PSNR of the resulting restored images. For 70% salt & pepper noise, Figure 26(c) and Figure 26(d) display the restored image \( A_M^\lambda (f_K) \) by Algorithm 1 and Algorithm 2, respectively, with \( K \) equal to the true set that has been enlarged by two pixels, whereas Figure 26(e) and Figure 26(f) show the restored image by the Adaptive median Filter and the TV-based method [28, 34] using the same set \( K \). Although the visual quality of the images restored from 70% noise corruption is comparable between our method and the TV-based method, the PSNR using our method with Algorithm 1 is higher than that for the TV-based method in all of the experiments reported in Table 3, Table 4 and Table 5. An additional advantage of our method is its speed. Our method does not require initialisation which is in contrast with the two-stage TV-based method, for which the initialisation, for instance, is given by the restored image using an adaptive median filter.

| Noise Density | Algorithm 1 | Algorithm 2 | TV       |
|--------------|-------------|-------------|----------|
| 70% (6.426 dB)| 26.674 dB   | 26.634 dB   | 26.506 dB|
| 90% (5.371 dB)| 23.117 dB   | 22.968 dB   | 22.521 dB|
| 99% (4.938 dB)| 18.424 dB   | 18.357 dB   | 17.420 dB|

**Table 3** Comparison of PSNR of the restored images by the compensated convexity based method \( A_M^\lambda (f_K) \) by applying the Moreau based scheme (Algorithm 1) and the convex based scheme (Algorithm 2), and by the two-stage TV-based method (TV), with the set \( K \) with noise–free boundary.

| Noise Density | Algorithm 1 | Algorithm 2 | TV       |
|--------------|-------------|-------------|----------|
| 70% (6.426 dB)| 26.642 dB   | 26.020 dB   | 26.475 dB|
| 90% (5.371 dB)| 23.078 dB   | 22.654 dB   | 22.459 dB|
| 99% (4.938 dB)| 18.240 dB   | 18.026 dB   | 17.314 dB|

**Table 4** Comparison of PSNR of the restored images by the compensated convexity based method \( A_M^\lambda (f_K) \) by applying the Moreau based scheme (Algorithm 1) and the convex based scheme (Algorithm 2), and by the two-stage TV-based method (TV), with the set \( K \) padded by two pixels.
Finally, to demonstrate the performance of our method in some extreme cases of very sparse data, we consider cases of noise density equal to 90% and 99%. Figure 27 displays the restored image by the compensated convexity based method \( A_{\lambda}^{M}(f_K) \) by applying the Moreau based scheme (Algorithm 1) and the convex based scheme (Algorithm 2), and by the two-stage TV-based method (TV), with the set \( K \) padded by ten pixels. As far as the visual quality of the restored images is concerned, and to the extent that such judgement can make sense given the high level of noise density, the inspection of Figure 27 seems to indicate that \( A_{\lambda}^{M}(f_K) \) gives a better approximation of details than the TV-based restored image. This is also reflected by the values of the PSNR index in the Table 3, Table 4 and Table 5.

| Noise Density | \( A_{\lambda}^{M}(f_K) \) (Algorithm 1) | \( A_{\lambda}^{M}(f_K) \) (Algorithm 2) | TV |
|---------------|-----------------------------------|-----------------------------------|-----|
| 70% (6.426 dB)| 26.640 dB                         | 26.020 dB                         | 26.468 dB |
| 90% (5.371 dB)| 23.068 dB                         | 22.654 dB                         | 22.446 dB |
| 99% (4.938 dB)| 18.342 dB                         | 18.026 dB                         | 17.330 dB |

Table 5 Comparison of PSNR of the restored images by the compensated convexity based method \( A_{\lambda}^{M}(f_K) \) by applying the Moreau based scheme (Algorithm 1) and the convex based scheme (Algorithm 2), and by the two-stage TV-based method (TV), with the set \( K \) padded by ten pixels.
Figure 26  (a) Original image; (b) Original image covered by a salt & pepper noise density of 70%. PSNR = 6.426 dB; (c) Restored image $A^M(f_K)$ by Moreau based scheme (Algorithm 2) with the set $K$ padded by two pixels. PSNR = 26.020 dB. $\lambda = 20$, $M = 1E13$. Total number of iterations: 21. (d) Restored image $A^M(f_K)$ by Convex based scheme (Algorithm 1) with the set $K$ padded by two pixels. PSNR = 26.642 dB. $\lambda = 20$, $M = 1E13$. Total number of iterations: 1865. (e) Restored image by the Adaptive Median filter [59] used as starting guess for the two-stage TV-based method described in [28, 33]. Window size $w = 33$ pixels. PSNR = 22.519 dB. (f) Restored image by the two-stage TV-based method described in [28, 33] with the set $K$ padded by two pixels. PSNR = 26.475 dB. Total number of iterations: 3853.
5.3.3 Inpainting

Inpainting is the problem where we are given an image that is damaged in some parts and we want to reconstruct the values in the damaged part on the basis of the known values of the image. This topic has attracted lot of interest especially as an application of TV related models [37, 95]. The main motivation is that functions of bounded variations provide the appropriate functional setting given that such functions are allowed to have jump discontinuities [9]. These authors usually argue that continuous functions cannot be used to model digital image related functions as functions representing images may have jumps [37], which are associated with the image features. However, from the human vision perspective, it is hard to distinguish between a jump discontinuity, where
values change abruptly, and a continuous function with sharp changes within a very small transition layer. By the application of our compensated convex based average transforms we are adopting the latter point of view. A comprehensive study of this theory applied to image inpainting can be found in [119] [121], where we also establish error estimates for our inpainting method and compare with the error analysis for image inpainting discussed in [36]. We note that for the relaxed Dirichlet problem of the minimal graph [36] or of the TV model used in [36], as the boundary value of the solution does not have to agree with the original boundary value, extra jumps can be introduced along the boundary. By comparison, since our average approximation is continuous, it will not introduce such a jump discontinuity at the boundary.
"Those who fall in love with practice without science are like a sailor who enters a ship without a helm or a compass, and who never can be certain whither he is going."

Leonardo da Vinci

"Those who fall in love with practice without science are like a sailor who enters a ship without a helm or a compass, and who never can be certain whither he is going."

Leonardo da Vinci

Figure 28 Inpainting of a text overprinted on an image: (a) Input image. (b) Restored image \( A^M_\lambda(f_K) \) using Algorithm 2. PSNR = 39.122 dB. Parameters: \( \lambda = 18 \) and \( M = 1 \cdot 10^5 \). Total number of iterations: 19. (c) Restored image by the AMLE method described in [95, 83]. PSNR = 36.406 dB. Total number of iterations: 5247. (d) Restored image by the Split Bregman inpainting method described in [52]. PSNR = 39.0712 dB. Total number of iterations: 19.

To assess the performance of our reconstruction compared to state-of-art inpainting methods, we consider synthetic example where we are given an image \( f \) and we overprint some text on it. The problem is then removing the text overprinted on the image displayed in Figure 28(a) and how close we can get to the original image \( f \). If we denote by \( P \) the set of pixels containing the overprinted text, and by \( \Omega \) the domain of the whole image, then \( K = \Omega \setminus P \) is the set of the true pixels and the inpainting problem is in fact the problem of reconstructing the image over \( P \) from knowing \( f_K \), if we denote by \( f \) the original image values. we compare our method with the total variation based image inpainting method solved by the split Bregman method described in
and with the AMLE inpainting reported in [95]. The restored image $A_M^\lambda(f_K)$ obtained by our compensated convexity method is displayed in Figure 28(b), the restored image by the AMLE method is shown in Figure 28(d) whereas 28(c) presents the restored image by the the split Bregman inpainting method. All the restored images look visually quite good. However, if we use the PSNR as a measure of the quality of the restoration, we find that $A_M^\lambda(f_K)$ has a value of PSNR equal to 39.122 dB, the split Bregman inpainting restored image gives a value for PSNR = 39.071 dB, whereas the AMLE restored image has PSNR equal to 36.406 dB. To assess how well $A_M^\lambda(f_K)$ is able to preserve image details and not to introduce unintended effects such as image blurring and staircase effects, Figure 29 displays details of the original image and of the restored images by the three methods. Once again, the good performance of $A_M^\lambda(f_K)$ can be appreciated visually.

![Figure 29](image)

**Figure 29** Comparison of a detail of the original image with the corresponding detail of the restored images according to the compensated convexity method and the TV-based method. Lips detail of the original image (a) without and (b) with overprinted text. Lips detail of the (c) restored image $A_M^\lambda(f_K)$ using Algorithm 2; (d) AMLE-based restored image; (d) TV-based restored image.

We conclude this section with two real-world applications, where we actually do not know the true background picture $f$, thus the assessment of the inpainting must simply rely on the visual quality of the approximation. Figure 30 compares the results of the Average compensated approximation and of the TV-based approximation in the case of the restoration of an image containing a scratch, whereas Figure ?? refers to the removal of an unwanted thin object from a picture. For both the examples, the two approximations yield qualitatively good results.
Figure 30 Restoration of an old image. (a) Input image with the scratch. (b) Input image with manual definition of the mask, given by the domain to repair. (c) Restored image $A_M^M(f_R)$ with $\lambda = 15, M = 10^6$. (d) TV-based restored image.
Figure 31  Removal of a thin object from a picture. (a) Input image. (b) Input image with manual definition of the mask, given by the domain to be inpainted. (c) Restored image $A^M_\lambda(f_K)$ with $\lambda = 15, M = 10^6$. (d) TV-based restored image.
Acknowledgements

AO acknowledges the partial financial support of the Argentinian Research Council (CONICET) through the project PIP 11220170100100CO, the National University of Tucumán through the project PIUNT CX-E625 and the FonCyT through the project PICT 2016 201-0105 Prestamo Bid. EC is grateful for the financial support of the College of Science, Swansea University, and KZ wishes to thank University of Nottingham for its support.

References

[1] Aichholzer O., Aigner W., Aurenhammer F., Hackl T., Jüttler B., Rabl M., Medial axis computation for planar free-form shapes, Comput. Aided Design 41 (2009) 339–349.

[2] Albano P., The regularity of the distance function propagates along minimizing geodesics, Nonlinear Anal. 95 (2014) 308–312.

[3] Albano P., Cannarsa P., Structural properties of singularities of semiconcave functions, Ann. Scuola Norm. Sup. Pisa Cl. Sci. 28 (1999) 719–740.

[4] Albano P., Cannarsa P., Nguyen K. T., Sinestrari C., Singular gradient flow of the distance function and homotopy equivalence, Math. Ann. 356 (2013) 23–43.

[5] Alberti G., On the structure of singular sets of convex functions, Calculus of Variations and Partial Differential Equations 2 (1994) 17–27.

[6] Alberti G., Ambrosio L., Cannarsa P., On the singularities of convex functions, Manuscripta Math. 76 (1992) 421–435.

[7] Almansa A., Cao F., Gousseau , Rougé B., Interpolation of digital elevation models using AMLE and related methods, IEEE Trans. Geoscience and Remote Sensing 40 (2002) 314–325.

[8] Alvarez L., Guichard F., Lions P.L., Morel J.M., Axioms and fundamental equations in image processing. Arch. Ration. Mech. Anal. 123 (1993) 199–257.

[9] Ambrosio L., Fusco N., Pallara D., Functions of Bounded Variation and Free Discontinuity Problems, Clarendon Press, UK, 2000.

[10] Ambrosio L., Tilli P., Topics on Analysis in Metric Spaces, Oxford Univ. Press, 2004.

[11] Amenta N., Bern M., Surface reconstruction by Voronoi filtering, Discrete Comput. Geom. 22 (1999) 481–504.

[12] Amenta N., Choi S., Kolluri R., The power crust, unions of balls, and the medial axis transform, Comp. Geom-Theor. Appl. 19 (2001) 127–153.

[13] Astola J., Kuosmanen P., Fundamentals of Nonlinear Digital Filtering. CRC Press, Boca Raton, 1997.

[14] Attali D., Lieutier A., Optimal reconstruction might be hard, Discrete Comput. Geom. 49 (2013) 133–156.

[15] Attali D., Montanvert A., Computing and simplifying 2D and 3D semicontinuous skeletons of 2D and 3D shapes, Comput. Vis. Image Und. 67 (1997) 261–273.
[16] Attouch H., Aze D., Approximations and regularizations of arbitrary functions in Hilbert spaces by the Lasry-Lions methods, Anal. Non-Lin. H. Poincaré Inst. 10 (1993) 289-312.

[17] Aubert G., Kornprobst P., Mathematical Problems in Image Processing. Partial Differential Equations and the Calculus of Variations. Springer Berlin 2006

[18] Aussel D., Daniilidis A., Thibault L., Subsmooth sets: functional characterizations and related concepts, Trans. Am. Math. Soc. 357 (2005) 1275–1301.

[19] Ball J., The calculus of variations and materials science Quarterly of Applied Mathematics LVI (1998) 719–740.

[20] Ball J. M., James R. D., Fine phase mixtures as minimizers of energy, Arch. Rational Mech. Anal. 100 (1987) 13–52.

[21] Ball J. M., James R. D., Proposed experimental tests of a theory of fine microstructures and the two-well problem, Phil. Royal Soc. Lond. A 338 (1992) 389–450.

[22] Barber, C. B., Dobkin, D. P., Huhdanpaa, H., The quickhull algorithm for convex hulls. ACM Transactions on Mathematical Software 22 (1996) 469–483.

[23] Bauschke, H. H., Combettes, P. L., Convex Analysis and Monotone Operator Theory in Hilbert Spaces Springer Verlag, Berlin, 2nd ed., 2017.

[24] Blum H., A transformation for extracting new descriptors of shape, Prop. Symp. Models for the Perception of Speech and Visual Form (W. W. Dunn ed.), MIT Press (1967) 362–380.

[25] Boissonnat J.D., Cazals F., Smooth surface reconstruction via natural neighbor interpolation of distance functions, ACM Symposium on Computational Geometry (2000) 223–232.

[26] Brenier Y., Un algorithme rapide pour le calcul de transformées de Legendre-Fenchel discrètes. C.R. Acad. Sci. Paris Sér. I Math. 308 (1989) 587–589.

[27] Brighi B., Chipot M., Approximated convex envelope of a function. SIAM Journal on Numerical Analysis 31 (1994) 128–148.

[28] Cai J.-F., Chan R., Morini B., Minimization of an edge-preserving regularization functional by conjugate gradient type methods. In: Image Processing Based on Partial Differential Equations, Tai X.-C., Lie K.-A., Chan T. F., Osher S., Eds., Springer (2005) 109–122.

[29] Calabi L., Hartnett W. E., Shape recognition, prairie fires, convex deficiencies and skeletons, The American Mathematical Monthly 75 (1968) 335–342.

[30] Cannarsa P., Peirone R., Unbounded components of the singular set of the distance function in $\mathbb{R}^n$, Transactions of the American Mathematical Society 353 (2001) 4567–4581.

[31] Cannarsa P., Sinestrari C., Semiconcave Functions, Hamilton-Jacobi Equations and Optimal Control, Birkhäuser, Boston, 2004.

[32] Carlsson M., On convex envelopes and regularization of non-convex functionals without moving global minima. Journal of Optimization Theory and Applications 183 (2019) 66–84.

[33] Caselles V., Morel J.-M., Sbert C., An axiomatic approach to image interpolation, IEEE Trans. Image Processing 7 (1998) 376–386.
[34] Chan R.H., Ho C.-W., Nikolova M., Salt-and-pepper noise removal by median-type noise
detectors and detail-preserving regularization. IEEE Transactions on Image Processing 14 (2005)
1479–1485.

[35] Chan T. M., Optimal output-sensitive convex hull algorithms in two and three dimensions.
Discrete and Computational Geometry 16 (1996) 361–368.

[36] Chan T. F., Kang S. H., Error analysis for image inpainting, J. Math. Imag. Vis. 26 (2006)
85–103.

[37] Chan T., Shen J., Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic
Methods, SIAM, 2005.

[38] Chazal F., Soufflet R., Stability and finiteness properties of medial axis and skeleton, J. Control
Dyn. Sys. 10 (2004) 149–170.

[39] Choi H. I., Choi S. W., Moon H. P., Mathematical theory of medial axis transform, Pacific J.
Math. 181 (1997) 57–88.

[40] Choi S. W., Seidel H.-P., Linear one-sided stability of MAT for weakly injective 3D domain,
Comput. Aided Design 36 (2004) 95–109.

[41] Clarke F.H., Ledyaev Yu.S., Stern R.J., Wolenski P.R., Nonsmooth Analysis and Control
Theory, Springer-Verlag, New York, 1998.

[42] Contento L., Ern A., Vermiglio R., A linear-time approximate convex envelope algorithm using
the double Legendre-Fenchel transform with application to phase separation. Computational
Optimization and Applications volume 60 (2015) 231–261.

[43] Corrias L., Fast Legendre-Fenchel transform and applications to Hamilton-Jacobi equations
and conservation laws. SIAM Journal on Numerical Analysis 33 (1996) 1534–1558.

[44] Crandall M. G., Ishii H., Lions P.-L., User’s guide to viscosity solutions of second order partial
differential equations, Bull. Am. Math. Soc. 27 (1992) 1–67.

[45] Dacorogna B., Direct Methods in the Calculus of Variations, Springer–Verlag, New York, 2008,
2nd Ed.

[46] DeVore R.A., Lorentz G.G., Constructive Approximation, Springer-Verlag, Berlin, 1993.

[47] Dey T. K., Curve and Surface Reconstruction, Cambridge University Press, 2006.

[48] Dey T. K., Zhao W., Approximating the medial axis from the Voronoi diagram with a conver-
gence guarantee, Algorithmica 38 (2004) 356–366.

[49] Felzenszwalb P. F., Huttenlocher D. P., Distance transforms of sampled functions. Theory of
Computing 8 (2012) 415–428.

[50] Firoozye, N. B., Optimal use of the translation method and relaxations of variational problems.
Communications on Pure and Applied Mathematics 44 (1991) 643–678.

[51] Gesch D., Evans G., Mauck J., Hutchinson J., Carwell Jr W. J., The national map elevation,
U.S. Geological Survey Fact Sheet 3053, 2009.
[52] Getreuer P., Total variation inpainting using split Bregman. Image Processing On Line 2 (2012) 147–157.

[53] Giusti E., Minimal Surfaces and Functions of Bounded Variation, Birkhäuser Verlag, Switzerland, 1984.

[54] Grabovsky, Y., Bounds and extremal microstructures for two-component composites: A unified treatment based on the translation method. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 452 (1996) 919–944.

[55] Hare W.L., A proximal average for nonconvex functions: A proximal stability perspective. SIAM Journal on Optimization 20 (2009) 650–666.

[56] Hartman P., On functions representable as a difference of convex functions, Pacific J. Math. 9 (1959), 707–713.

[57] Helluy P., Mathis H., Pressure laws and fast Legendre transform. Mathematical Models and Methods in Applied Sciences 21 (2011) 745–775.

[58] Hiriart-Urruty J.-B., Lemaréchal C., Fundamentals of Convex Analysis. Springer, 2001.

[59] Hwang H, Haddad R.A., Adaptive median filters: New algorithms and results. IEEE Transactions on Image Processing 4 (1995) 499–502.

[60] Jackway P. T., Morphological scale-space, IAPR International Conference on Pattern Recognition, IEEE Computer Society Press, Los Alamitos, CA (1992) 252–255.

[61] Jackway P.T., Properties of multiscale morphological smoothing by poweroids. Pattern Recognit. Lett. 15 (1994) 135–140.

[62] Jackway P.T., On dimensionality in multiscale morphological scale-space with elliptic poweroid structuring functions. J. Vis. Commun. Image Represent. 6 (1995) 189–195.

[63] Jackway P.T., Deriche M., Scale–space properties of the multiscale morphological dilation–erosion. IEEE Trans. Pattern Anal. Mach. Intell. 18 (1996) 38–51

[64] Kimmel R., Shaked D., Kiryati N., Bruckstein A., Skeletonization via distance maps and level sets, Comput. Vis. Image Und. 62 (1995) 382–391.

[65] Koenderink J.J., The structure of images. Biol. Cybernet. 50 (1984) 363–370.

[66] Lasry J.M., Lions P.L., A remark on regularization in Hilbert Spaces, Israel Math. J. 55 (1986) 257-266.

[67] Lieutier A., Any open bounded subset of $\mathbb{R}^n$ has the same homotopic type as its medial axis, Comput. Aided Design 36 (2004) 1029–1046.

[68] Lindeberg T., Scale-Space Theory in Computer Vision. Kluwer, Boston, 1994.

[69] Lindeberg T., Generalized Gaussian scale-space axiomatics comprising linear scale-space, affine scale-space and spatio-temporal scale-space. J. Math. Imaging Vis. 40 (2011) 36–81.

[70] Lucet Y., Faster than the Fast Legendre-Transform, the linear-time Legendre Transform. Numerical Algorithms 16 (1997) 171–185.
[71] Lucet Y., Fast Moreau envelope computation I: numerical algorithms. Numerical Algorithms 43 (2006) 235–249.

[72] Lucet Y., What shape is your conjugate? A survey of computational convex analysis and its applications. SIAM Review 52 (2010) 505–542.

[73] Mantegazza C., Mennucci A. C., Hamilton-Jacobi equations and distance functions on Riemannian manifolds. Appl. Math. Optim. 47 (2003) 1–25.

[74] Maragos P., Schafer R., Morphological filters-Part I: their settheoretic analysis and relations to linear shift-invariant filters. IEEE Trans. Acoust. Speech Signal Process. 35 (1987) 1153–1169.

[75] Matheron G., Examples of topological properties of skeletons, J. Serra (Ed), Image Analysis and Mathematical Morphology, Part II, Academic Press, 1988.

[76] Milton, G. W., On characterizing the set of possible effective tensors of composites: The variational method and the translation method. Communications on Pure and Applied Mathematics 43 (1990) 63–125.

[77] Moreau J.-J., Proximaté dualité dans un espace Hilbertien, Bull. Soc. Math. Fr. 93 (1965) 273–299.

[78] Moreau J.-J., Fonctionnelles Convexes, Lecture Notes, Collège de France, 1966.

[79] Okabe A., Boots B., Sugihara K., Chiu S. N., Spatial Tessellations–Concepts and Applications of Voronoi Diagrams. John Wiley & Sons, Second Edition, 2000.

[80] Oberman A. M., Computing the convex envelope using a nonlinear partial differential equation. Math. Models Methods Appl. Sci. 18 (2008) 759–780.

[81] Ogniewicz R. L., Kübler O., Hierarchic Voronoi skeletons, Pattern Recogn. 28 (1995) 343–359.

[82] Parikh N., Boyd S., Proximal algorithms, Foundations and Trends in Optimization 1 (2013) 123–231.

[83] Parisotto S., Schönlieb B.-C., MATLAB Codes for the Image Inpainting Problem, GitHub repository, MATLAB Central File Exchange, September, 2016.

[84] Parisotto S., Lellmann J., Masnou S., Schönlieb C.-B., Higher-order total directional variation. Part I: Imaging applications. Available at https://arxiv.org/abs/1812.05023.

[85] Patrikalakis N. M., Maekawa T., Shape Interrogation for Computer Aided Design and Manufacturing, Springer, Heidelberg, 2002.

[86] Preparata F. P., Shamos M., Computational Geometry. An Introduction, Springer Verlag, Berlin, 1985.

[87] Reshetnyak Y. G., On a generalization of convex surfaces, Mat. Sbornik 40 (1956) 381–398.

[88] Rockafellar R. T., Convex Analysis, Princeton Univ. Press, New Jersey, 1970.

[89] Rockafellar R. T., Clarke’s tangent cones and boundaries of closed sets in \( \mathbb{R}^n \), Nonlinear Analysis: Theory, Methods & Applications 3 (1979) 145–154.

[90] Rockafellar R. T., Wets R. J-B., Variational Analysis, Springer, Berlin, 1998.
Rolewicz S., On $\gamma$-paraconvex multifunctions, Math. Jpn. 24 (1979) 293–300.

Rolewicz S., On $\alpha(\cdot)$-paraconvex and strongly $\alpha$-paraconvex functions, Control Cyber. 29 (2000) 367–377.

Schavemaker J.G.M., Reinders M.J.T., Gerbrands J.J., Backer E., Image sharpening by morphological filtering, Pattern Recognition 33 (2000) 997–1012.

Schmidt M., Weickert J., Morphological counterparts of linear shift-invariant scale-spaces, Journal of Mathematical Imaging and Vision, 56 (2016) 352–366.

Schönlieb C.-B., Partial Differential Equation Methods for Image Inpainting, Cambridge University Press, 2015.

Serra J., Image Analysis and Mathematical Morphology, Academic Press, London, 1982.

Sherbrooke E. C., Patrikalakis N. M., Wolter F. E., Differential and topological properties of medial axis transforms, Graph. Model Im. Proc. 58 (1996) 574–592.

Shih F.Y., Image Processing and Mathematical Morphology, CRC Press, Boca Raton, USA, 2009.

Shih F.Y., O. Mitchell O., A mathematical morphology approach distance transformation, IEEE Transactions on Image Processing 1 (1992) 197–204.

Siddiqi K., Pizer S. M. (Eds), Medial Representations, Springer, New York, 2008.

Smith S., Brady J., SUSAN—a new approach to low-level image processing. International Journal of Computer Vision 23 (1997) 45–78.

Soille P., Morphological Image Analysis, Springer, Berlin, 2nd Ed, 2004.

SRTM and Landcover Download site, http://ve2dbe.com/geodata/, Accessed: 2020-09-30.

Strömberg T., On regularization in Banach spaces, Arkiv för Matematik 34 (1996) 383–406.

Tartar, L., Estimations fines de coefficients homogénéisés. In: Ennio De Giorgi Colloquium (P. Krée, ed.), 168–187. Res. Notes Math. 125 London, Pitman, 1985.

Van Den Boomgaard R., Mathematical Morphology: Extensions towards Computer Vision, PhD Thesis, University of Amsterdam, The Netherlands, 1992.

Van den Boomgaard R., The morphological equivalent of the Gauss convolution. Nieuw Archief VoorWiskunde 10 (1992) 219–236.

Van den Boomgaard R., Smelders A. W. M., The morphological structure of images, IAPR International Conference on Pattern Recognition, IEEE Computer Society Press, Los Alamitos, CA (1992) 268–271.

Vese L., A method to convexify functions via curve evolution. Communications on Partial Differential Equations 24 (1999) 1573–1591.

Vese L., Le Guyader C., Variational Methods in Image Processing Chapman & Hall 2015.

Vial J.-P., Strong and weak convexity of sets and functions, Math. Oper. Research. 8 (1983) 231–259.
[112] Weickert J., Anisotropic Diffusion in Image Processing. Teubner, Stuttgart, 1998.

[113] Weickert J., Ishikawa S., Imiya A., Linear scale-space has first been proposed in Japan. J. Math. Imaging Vis. 10 (1999) 237–252.

[114] Wolter F. E., Cut locus and medial axis in global shape interrogation and representation, MIT, Dept. Ocean Engineering, Design Laboratory Memorandum Issue 92-2 (1993).

[115] Zhang K., On the structure of quasiconvex hulls , Anal. Non-Lin. H. Poincaré Inst. 15 (1998) 663–686.

[116] Zhang K., Compensated convexity and its applications, Anal. Non-Lin. H. Poincaré Inst. 25 (2008) 743–771.

[117] Zhang K., Convex analysis based smooth approximations of maximum functions and squared-distance functions, Journal of Nonlinear Convex Analysis 9 (2008) 379–406.

[118] Zhang K., Crooks E., Orlando A., Compensated convexity, multiscale medial axis maps and sharp regularity of the squared distance function, SIAM Journal on Mathematical Analysis 47 (2015) 4289–4331.

[119] Zhang Z., Crooks E., Orlando A., Compensated convexity methods for approximations and interpolations of sampled functions in Euclidean spaces: Theoretical Foundations. SIAM Journal on Mathematical Analysis 48 (2016) 4126–4154.

[120] Zhang Z., Crooks E., Orlando A., Compensated convex transforms and geometric singularity extraction from semiconvex functions (in Chinese). Scientia Sinica Mathematica 46 (2016) 1–22. DOI: 10.1360/N012015-00339, (revised English version available at https://arxiv.org/abs/1610.01451)

[121] Zhang Z., Crooks E., Orlando A., Compensated convexity methods for approximations and interpolations of sampled functions in Euclidean Spaces: Applications to contour lines, sparse data and inpainting. SIAM J. Imaging Sciences 11 (2018) 2368–2428.

[122] Zhang K., Orlando A., Crooks E., Compensated convexity and Hausdorff stable geometric singularity extractions, Mathematical Models and Methods in Applied Sciences 25 (2015) 747–801, DOI: 10.1142/S0218202515500189.

[123] Zhang K., Orlando A., Crooks E., Compensated convexity and Hausdorff stable extraction of intersections for smooth manifolds, Mathematical Models and Methods in Applied Sciences 25 (2015) 839–873, DOI: 10.1142/S0218202515500207.

[124] Zhang K., Orlando A., Crooks E., Compensated Convexity on Bounded Domains, Mixed Moreau Envelopes and Computational Methods. Available at https://arxiv.org/abs/1907.02286