Cronfa - Swansea University Open Access Repository

_____

This is an author produced version of a paper published in :
SIAM Journal on Mathematical Analysis

_____

Cronfa URL for this paper:

http://cronfa.swan.ac.uk/Record/cronfa30572

_____

Paper:

Zhang, K., Crooks, E. & Orlando, A. (2016). Compensated Convexity Methods for Approximations and Interpolations of Sampled Functions in Euclidean Spaces: Theoretical Foundations. *SIAM Journal on Mathematical Analysis, 48*(6), 4126-4154.

http://dx.doi.org/10.1137/15M1045673

_____

Compensated Convexity Methods
for Approximations and Interpolations
of Sampled Functions in Euclidean Spaces:
Theoretical Foundations

Kewei Zhang,[*] Elaine Crooks[†] and Antonio Orlando[‡]

Abstract

We introduce Lipschitz continuous and $C^{1,1}$ geometric approximation and interpolation methods for sampled bounded uniformly continuous functions over compact sets and over complements of bounded open sets in $\mathbb{R}^n$ by using compensated convex transforms. Error estimates are provided for the approximations of bounded uniformly continuous functions, of Lipschitz functions, and of $C^{1,1}$ functions. We also prove that our approximation methods, which are differentiation and integration free and not sensitive to sample type, are stable with respect to the Hausdorff distance between samples.

1 Introduction

In this paper we apply compensated convex transforms [47, 48, 49, 50] to define Lipschitz continuous and smooth ($C^{1,1}$) geometric approximations and interpolations for bounded real-valued functions sampled from either a compact set K in $\mathbb{R}^n$ or the complement $K = \mathbb{R}^n \setminus \Omega$ of a bounded open set Ω. The former is motivated by approximating or interpolating sparse data or contour lines and the latter by the so-called inpainting problem in image processing [17], where some parts of the image content are missing and the aim is to use other parts of the image to repair or reconstruct the missing parts. We first define two one-sided approximations, called upper and lower approximations, from above and below the graph of the sampled function respectively, and then an average approximation. By using mixed compensated convex transforms [47], we will also define a smooth ($C^{1,1}$) average approximation. Our central aim here is to develop a mathematical theory for these average approximations. Applications of this theory to level-set reconstruction, scattered data interpolation and inpainting will be presented, together with some prototype examples, in a follow-on paper [51].

---

[*]School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

[†]Department of Mathematics, Swansea University, Singleton Park, Swansea, SA2 8PP, UK

[‡]CONICET, FACET, Universidad Nacional de Tucumán, Argentina

Before relating our results to previous work on approximations and interpolations of sampled functions, we first recall the notions of quadratic compensated convex transforms of bounded functions and present our definitions of upper, lower and average approximations (note that compensated convex transforms can be defined under more general growth conditions than those given here [47]).

Suppose $f : \mathbb{R}^n \mapsto \mathbb{R}$ is bounded. The quadratic lower and upper compensated convex transform [47] (lower and upper transforms for short) are defined for each $\lambda > 0$ by

$$C_\lambda^l(f)(x) = \mathsf{co}[\lambda|\cdot|^2 + f](x) - \lambda|x|^2, \quad \text{resp.} \quad C_\lambda^u(f)(x) = \lambda|x|^2 - \mathsf{co}[\lambda|\cdot|^2 - f](x), \qquad x \in \mathbb{R}^n, \quad (1.1)$$

where $|x|$ is the standard Euclidean norm of $x \in \mathbb{R}^n$ and $\mathsf{co}[g]$ denotes the convex envelope [27, 37] of a function $g : \mathbb{R}^n \mapsto \mathbb{R}$ that is bounded below. For given $\lambda > 0$ and $\tau > 0$, two quadratic mixed compensated convex transforms [47] (mixed transforms for short) are defined, respectively, by $C_\tau^u(C_\lambda^l(f))$ and $C_\tau^l(C_\lambda^u(f))$.

One key property of the compensated convex transforms, established in [47], is that $C_\lambda^l(f)$ (respectively, $C_\lambda^u(f)$) realises a 'tight' approximation of f from below (respectively, from above), in the sense that if f is $C^{1,1}$ in a neighbourhood of some x_0, then there is a finite $\Lambda > 0$, such that $f(x_0) = C_\lambda^l(f)(x_0)$ (respectively, $f(x_0) = C_\lambda^u(f)(x_0)$) whenever $\lambda \geq \Lambda$. A second important property is that of locality. Since the definitions(1.1) involve the evaluation of the convex envelope of functions [27, 37], one might think that these notions are global in nature, that is, the values of these transforms at a given point might involve values of the original function far away from the point. However, the locality property for compensated convex transforms [48, Theorem 3.10] states that if f is bounded, i.e., $|f(x)| \leq M$ in $\mathbb{R}^n$ for some $M > 0$, then the values of $C_\lambda^l(f)(x_0)$ and $C_\lambda^u(f)(x_0)$ depend only on the values of f in the closed ball $\bar{B}(x_0; R)$ with $R = 2\sqrt{2}\sqrt{M/\lambda}$. As a result, these apparently global transforms are, in fact, local.

In this paper, we mainly consider two types of data sets in $\mathbb{R}^n$, given that the typical applications we have in mind are approximation of sparse data and of contour lines, and inpainting of damaged images. We therefore assume in the following that, unless otherwise specified, $K \subset \mathbb{R}^n$ is either a compact set or the complement of a bounded open set $\Omega \subset \mathbb{R}^n$, i.e. $K = \mathbb{R}^n \setminus \Omega$. We denote by $f : \mathbb{R}^n \mapsto \mathbb{R}$ the underlying function to be approximated. The function $f_K : K \subset \mathbb{R}^n \mapsto \mathbb{R}$ is our sampled function defined by $f_K(x) = f(x)$ for $x \in K$, and $\Gamma_{f_K} := \{(x, f_K(x)), x \in K\}$ is its graph.

Let $K \subset \mathbb{R}^n$ be a non-empty closed set and suppose that for some constant $A_0 > 0$, $|f_K(x)| \leq A_0$ for all $x \in K$. Given $M > 0$, we define two functions extending f_K to $\mathbb{R}^n \setminus K$, namely

$$
\begin{aligned}
f_K^{-M}(x) &= f(x)\chi_K(x) - M\chi_{\mathbb{R}^n \setminus K} &&= \begin{cases} f_K(x), & x \in K, \\ -M, & x \in \mathbb{R}^n \setminus K; \end{cases} \\
f_K^M(x) &= f(x)\chi_K(x) + M\chi_{\mathbb{R}^n \setminus K} &&= \begin{cases} f_K(x), & x \in K, \\ M, & x \in \mathbb{R}^n \setminus K, \end{cases}
\end{aligned}
\qquad (1.2)
$$

where χ_G denotes the characteristic function of a set G.

Definition 1.1. *For $M > 0$, the **upper compensated convex approximation** with scale $\lambda > 0$ for the sampled function $f_K : K \to \mathbb{R}$ is defined by*

$$U_\lambda^M(f_K)(x) = C_\lambda^u(f_K^{-M})(x), \quad x \in \mathbb{R}^n. \qquad (1.3)$$

*The **lower compensated convex approximation** with scale $\lambda > 0$ for the sampled function*

2

$f_K : K \to \mathbb{R}$ *is defined by*

$$L_\lambda^M(f_K)(x) = C_\lambda^l(f_K^M)(x), \quad x \in \mathbb{R}^n. \tag{1.4}$$

The **average compensated convex approximation** *with scale $\lambda > 0$ for the sampled function $f_K : K \to \mathbb{R}$ is defined by*

$$A_\lambda^M(f_K)(x) = \frac{1}{2}\left(C_\lambda^l(f_K^M)(x) + C_\lambda^u(f_K^{-M})(x)\right), \quad x \in \mathbb{R}^n. \tag{1.5}$$

The **mixed average compensated convex approximation** *with scales $\lambda > 0$ and $\tau > 0$ for the sampled function $f_K : K \to \mathbb{R}$ is defined by*

$$(SA)_{\tau,\lambda}^M(f_K)(x) = \frac{1}{2}(C_\tau^u(C_\lambda^l(f_K^M))(x) + C_\tau^l(C_\lambda^u(f_K^{-M}))(x), \quad x \in \mathbb{R}^n. \tag{1.6}$$

In the following, we refer to the approximations in Definition 1.1, for short, as the upper, lower, average and mixed approximations.

Note that since the mixed compensated convex transforms are $C^{1,1}$ functions [47, Theorem 2.1(iv) and Theorem 4.1(ii)], the mixed average approximation $(SA)_{\tau,\lambda}^M$ is a smooth version of our average approximation. Also, for a bounded function $f : \mathbb{R}^n \mapsto \mathbb{R}$, satisfying $|f(x)| \leq M$, $x \in \mathbb{R}^n$ for some constant $M > 0$, we have the following estimates [48, Theorem 3.13]

$$0 \leq C_\tau^u(C_\lambda^l(f))(x) - C_\lambda^l(f)(x) \leq \frac{16M\lambda}{\tau}, \quad 0 \leq C_\lambda^u(f)(x) - C_\tau^l(C_\lambda^u(f))(x) \leq \frac{16M\lambda}{\tau}$$

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

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

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

Remark 1.2. *We can additionally consider the families of average approximations*

$$A_{\lambda,s}^M(f_K)(x) = sC_\lambda^l(f_K^M)(x) + (1-s)C_\lambda^u(f_K^{-M})(x), \quad s \in [0, 1]$$

and

$$(SA)_{\tau,\lambda,s}^M(f_K)(x) = sC_\tau^u(C_\lambda^l(f_K^M))(x) + (1-s)C_\tau^l(C_\lambda^u(f_K^{-M}))(x), \quad s \in [0, 1].$$

These more general average approximations give some flexibility when dealing with sets which are not graphs of single-valued functions. For instance, suppose $X \subset \mathbb{R}^n \times \mathbb{R}$ is a finite set. Let $K = \mathbb{P}_{\mathbb{R}^n}(X) := \{x_1, \ldots, x_n\}$ be the orthogonal projection of X to $\mathbb{R}^n$, and for $x \in K$, define

$$\check{f}_K(x) = \inf\{v, (x, v) \in X\}, \qquad \hat{f}_K(x) = \sup\{v, (x, v) \in X\}.$$

Then $\check{f}_K(x) \leq \hat{f}_K(x)$ and $\check{f}_K$, $\hat{f}_K$ are both single-valued functions. We can then define

$$A_{\lambda,s}^M(X)(x) := sC_\lambda^l(\check{f}_K^M)(x) + (1-s)C_\lambda^u(\hat{f}_K^{-M})(x) \tag{1.7}$$

for suitable M, and optimise $A_{\lambda,s}^M(X)$ with respect to $s \in [0, 1]$ to find a good approximation of the set X by a single-valued function. For example, we may consider the following nonlinear least

square approximation of the data set by the family of functions $A_{\lambda,s}^M(X)$,

$$\inf_{s\in[0,1]} \sum_{i=1}^{n} \max\left\{|A_{\lambda,s}^M(X)(x_i) - v|^2, \ (x_i, v) \in X\right\}.$$

However, we do not explore this further here, instead focussing on our basic average approximation $A_\lambda^M(f_K)$ and the mixed approximation $(SA)_{\tau,\lambda}^M$.

If we consider the special case where K is a finite set, the average approximation $A_\lambda^M(f_K)$ defines an approximation for the scattered data $\Gamma_{f_K} = \{(x, f_K(x)), \ x \in K\}$. Moreover, although our extended functions are defined in the whole space $\mathbb{R}^n$, when K is compact we are interested only in the values of our average approximation $A_\lambda^M(f_K)(x)$ for x in the convex hull $\mathsf{co}[K]$ of the sampled set K. If K is the complement of a bounded open set $\Omega \subset \mathbb{R}^n$, we will consider the values of $A_\lambda^M(f_K)(x)$ for x in the whole space $\mathbb{R}^n$ or in a large domain containing $\bar{\Omega}$.

Theoretically, we may also set $M = +\infty$ and consider the following functions, which are commonly used in convex analysis, in place of (1.2):

$$f_K^{-\infty}(x) = \begin{cases} f(x), & x \in K, \\ -\infty, & x \in \mathbb{R}^n \setminus K; \end{cases} \qquad f_K^{+\infty}(x) = \begin{cases} f(x), & x \in K, \\ +\infty, & x \in \mathbb{R}^n \setminus K. \end{cases} \tag{1.8}$$

This method of extension can help to establish better approximation results than those obtained using f_K^{-M} and f_K^M (compare Theorem 3.6 with Theorem 3.7). Note, however, that the corresponding average approximation,

$$A_\lambda^\infty(f_K)(x) := \frac{1}{2}\left(C_\lambda^l(f_K^{+\infty})(x) + C_\lambda^u(f_K^{-\infty})(x)\right), \quad x \in \mathbb{R}^n,$$

is not Hausdorff stable with respect to sample sets in the sense introduced in Section 4, in contrast to the basic average approximation $A_\lambda^M(f_K)$.

We turn now to some background and motivation. Selected recent developments on approximation and interpolation methods are discussed in [30]. The literature on approximation and interpolation theory for sampled functions over the real line $\mathbb{R}$ by polynomials and other functions is quite rich [44, 18]. When $n \geq 2$, however, many methods for $\mathbb{R}$ no longer apply directly to $\mathbb{R}^n$. In particular, there is no direct construction of interpolations for randomly placed sample functions based on one-dimensional interpolation methods. For scattered data, Delaunay triangulation-based direct spline designs have been widely used in computational geometry [35, 21]. Thin plate spline methods, variational methods, which are related to radial basis function methods, and more general kernel methods, have been used extensively in applications [45, 14, 30], and morphological reconstruction methods based on level sets using geodesic distance [43, see Ch.6.4 and Ch.7.1.2] have also been developed. Nonlinear partial differential equations and variational methods using various total variation (TV) based models [40] have been used for image reconstruction problems, salt & pepper noise reduction [16] and image inpainting [10, 17]. Although there is a well-developed mathematical theory on the existence and uniqueness of their weak solutions [22, 2, 3, 8], the quantitative effectiveness of such methods is mostly assessed on the basis of numerical experiments.

Note that many methods are sensitive to the type of data to be interpolated or approximated, that is, to the sample type. The spline function interpolation and finite element based methods design interpolations require precise knowledge of the sample locations. In this case, Delaunay triangulation or other types of decomposition typically must be constructed first [35]. The radial

basis function method relies on solving systems of linear equations [45]. In order to apply any of these methods to interpolate or approximate data sets, one has to assume that the data set is scattered, that is, the set is finite and the points are isolated. If the data set is given by contour lines (or by level sets), further discretisation is required before such methods can be used.

Our approach, on the other hand, is not sensitive to data types. We only assume the underlying function to be bounded and uniformly continuous, and the sample sets to be compact or to be the complement of a bounded open set. In the digital setting, the data are always finite sets, and in a 'point cloud', a line can be formed by discrete points next to each other, which, by definition, should not be thought of as scattered data. Therefore further down sampling might be needed in order to apply spline or radial basis function methods. But our average approximation $A_\lambda^M(f_K)$, on the contrary, applies directly to these data sets. In addition, collected data are bounded in a given window, and thus the assumption of boundedness of the underlying functions covers most situations in applications. It should be noted that the idea of using averages for approximations is natural and has also been used before by several authors, for example [7, 6, 5] introduce the notion of proximal average, a parametrized convex function that provides a continuous transformation of a convex function into another. In [25] this transformation has also been applied to non-convex functions and with non-quadratic weights by exploiting its relationship with the Moreau envelopes, and has been used as a fundamental tool to justify the application of parallel proximal algorithms in nonsmooth optimization [46, 36].

The exact form of an interpolation is of interest but is often not known. An advantage of Delaunay triangulation-based spline interpolation methods is that for simple geometric examples, one can describe precisely what the interpolation is, in contrast to, for instance, radial basis function and partial differential equation based methods. Although we do not deliberately design the form of our interpolations, it can be shown that our average approximation $A_\lambda^M(f_K)$ produces particular forms for us automatically. For example, if K is finite and $\lambda > 0$, $M > 0$ are large, we will prove in a follow-on paper [51] that $A_\lambda^M(f_K)(x)$ is a piecewise affine interpolation from K to $\mathsf{co}[K]$. We can also give explicit calculations of our approximations in some other simple geometric cases.

A further 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 [1], 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 our method is that for a bounded uniformly continuous function f, and for fixed $M > 0$ and $\lambda > 0$, the mapping $K \mapsto A_\lambda^M(f_K)(x)$ 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 $A_\lambda^M(f_E)(x)$ is close to $A_\lambda^M(f_K)(x)$ uniformly with respect to $x \in \mathbb{R}^n$. As far as we know, not many known interpolation/approximation methods share such a property.

To shed further light on the theory we develop, it is worth observing the connection between the compensated convex transforms and our proposed average approximation on the one hand, and the critical mixed Moreau envelopes and mathematical morphology on the other hand. The lower and upper transforms can be viewed as 'one-step' morphological opening and closing, respectively [48]. They in fact coincide with the critical mixed Moreau envelopes, that is,

$$C_\lambda^l(f)(x) = M^\lambda(M_\lambda(f))(x) \quad \text{and} \quad C_\lambda^u(f)(x) = M_\lambda(M^\lambda(f))(x), \tag{1.9}$$

where

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

are the lower and upper Moreau envelopes [33, 34, 31, 4, 15], respectively. If we denote by $b_\lambda(x) = -\lambda|x|^2$ the quadratic structuring function, introduced for the first time in [28, 11, 12, 29], then with the notation of [41, 43], we have[1]

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

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

that is, the Moreau lower and upper envelopes can be viewed as 'greyscale' erosion and dilation with quadratic structuring function, respectively [11, 32]. Compared with (1.9), we thus have

$$C_\lambda^l(f) = (f \ominus b_\lambda) \oplus b_\lambda \quad \text{and} \quad C_\lambda^u(f) = (f \oplus b_\lambda) \ominus b_\lambda,$$

and hence, using the definition of $A_\lambda^M(f_K)$, it follows that

$$A_\lambda^M(f_K) = \frac{1}{2}\left((f_K^{-M} \oplus b_\lambda) \ominus b_\lambda + (f_K^M \ominus b_\lambda) \oplus b_\lambda\right).$$

Given such an interpretation for $A_\lambda^M(f_K)$, the properties of $A_\lambda^M(f_K)$ could therefore also be analysed by tools from the theory of morphological filtering [42, 24].

The plan of the rest of this paper is as follows. In Section 2, we introduce notation and recall some useful results from convex analysis and compensated convexity theory. Our main sources of references for convex analysis are [27, 37, 39] whereas for the properties of compensated convex transforms we refer to [47, 48, 49, 50]. In Section 3 we state our general approximation/interpolation theorems for a compact sample set $K \subset \mathbb{R}^n$ (Theorem 3.6) and for $K = \mathbb{R}^n \setminus \Omega$ with Ω a bounded open set (Theorem 3.7). We consider uniformly continuous, Lipschitz and $C^{1,1}$ functions $f : \mathbb{R}^n \mapsto \mathbb{R}$ as the underlying functions to be approximated. We show that when $M > 0$ is sufficiently large, $A_\lambda^M(f_K)$ approaches f_K in K as $\lambda \to \infty$. If f is a $C^{1,1}$ function, we also show that $A_\lambda^M(f_K)$ is an interpolation of f in the convex hull $\mathsf{co}[K]$ of K when $\lambda > 0$ is large enough. For points x in $\mathsf{co}[K] \setminus K$, we introduce the notion of convex density radius $r_c(x)$ which is the smallest radius of a closed ball $\bar{B}(x; r_c(x))$ such that x is in the convex hull of $K \cap \bar{B}(x; r_c(x))$. We use $r_c(x)$ to bound the errors of our approximations $|A_\lambda^\infty(f_K)(x) - f(x)|$. For a finite $M > 0$ and for a compact sample set K, we extend f_K to be a constant c_0 outside a large ball $B(0; R)$ containing K and define $K_R = K \cup B^c(0; R)$. We then prove similar error estimates to those obtained for $A_\lambda^\infty(f_K) - f$ (Corollary 3.9). For example, for a bounded uniformly continuous function f,

$$|A_\lambda^M(f_{K_R})(x) - f(x)| \leq \omega(r_c(x) + a/\lambda + \sqrt{2b/\lambda}), \quad x \in \mathbb{R}^n,$$

where $\omega : [0, +\infty) \mapsto [0, +\infty)$ is the least concave majorant of the modulus of continuity of the function f [20] which satisfies $\omega(t) \leq at + b$ for $t \geq 0$ and some constants $a > 0$ and $b \geq 0$. Better estimates are also established for Lipschitz functions and for $C^{1,1}$ functions.

---

[1]In convex analysis, the infimal convolution of f with g is denoted in [37] as $f \Box g$ and is defined as $(f \Box g)(x) = \inf_y \{f(y) + g(x - y)\}$, whereas in [26] the deconvolution of f with g is denoted as $f \ominus g$ and, under the condition that for some $x_0 \in \mathbb{R}^n$ and $r \in \mathbb{R}$, we have $f(x) \leq g(x - x_0) + r$ for all $x \in \mathbb{R}^n$, is defined as $(f \ominus g)(x) = \sup_y \{f(x - y) - g(y)\}$.

Thus $M_\lambda(f)$ is the inf-convolution of f with $-b_\lambda$ whereas $M^\lambda(f)$ is the deconvolution of f with b_λ.

In Section 4 we state the Hausdorff stability property for the average approximation $A_\lambda^M(f_K)$ of bounded uniformly continuous functions, showing that given two non-empty closed sets K and E, for fixed $M > 0$ and $\lambda > 0$, $|A_\lambda^M(f_K)(x) - A_\lambda^M(f_E)(x)|$ is uniformly small in $\mathbb{R}^n$ with explicit estimates whenever K and E are closed. For a bounded Lipschitz function $f : \mathbb{R}^n \mapsto \mathbb{R}$ with $|f(x)| \le A_0 < M$ for some constant $A_0 > 0$ and all $x \in \mathbb{R}^n$, the mapping $K \mapsto A_\lambda^M(f_K)(x)$ is Lipschitz continuous with respect to the Hausdorff metric, uniformly with respect to $x \in \mathbb{R}^n$. This result generalises an earlier Hausdorff-Lipschitz continuity result for the upper transform of characteristic functions $K \mapsto C_\lambda^u(\chi_K)$ established in [48, Theorem 5.5]. We conclude Section 4 by proving regularity properties of our approximations. For example, we show that $A_\lambda^M(f_K)$ is a globally Lipschitz function in $\mathbb{R}^n$ and give an explicit estimate of its Lipschitz constant.

The proofs of our main results are presented in Section 5.

In the follow-on paper [51] we will present some applications of the theory developed here, such as interpolation and approximation of scattered data and for contour lines. We will also give some prototype examples with analytical expressions of our approximations, and numerical experiments on salt-and-pepper denoising, inpainting and contour-line based reconstructions.

2   Notation and Preliminaries

In this section we collect basic results and definitions from convex analysis, referring to [27, 37, 39] for further references and proofs, and recall the notion of the Hausdorff distance between two non-empty sets. We then list some selected basic properties of compensated convex transforms [47, 48, 49, 50] that will be needed in the following.

Proposition 2.1. *Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be coercive in the sense that $f(x)/|x| \to \infty$ as $|x| \to \infty$, and $x_0 \in \mathbb{R}^n$. Then*

(i) The value $\mathsf{co}\,[f]\,(x_0)$ of the convex envelope of f at $x_0 \in \mathbb{R}^n$ is given by

$$\mathsf{co}\,[f]\,(x_0) = \inf_{i=1,\ldots,n+1} \left\{ \sum_{i=1}^{n+1} \lambda_i f(x_i) : \sum_{i=1}^{n+1} \lambda_i = 1, \ \sum_{i=1}^{n+1} \lambda_i x_i = x_0, \ \lambda_i \ge 0, \ x_i \in \mathbb{R}^n \right\}. \quad (2.1)$$

If, in addition, f is lower semicontinuous, the infimum is attained by some (λ_i^*, x_i^*) for $i = 1, 2, \ldots, n+1$ with $(x_i^*, f(x_i^*))$'s lying in the intersection of a supporting plane of the epigraph of f, $\mathrm{epi}(f)$, and $\mathrm{epi}(f)$ (see [9, Lemma 3.3(ii)], [23, Theorem 2.1], and [39, Corollary 3.47]). In this case,

$$\mathsf{co}\,[f]\,(x_0) = \sum_{i=1}^{n+1} \lambda_i^* f(x_i^*). \quad (2.2)$$

(ii) The value $\mathsf{co}\,[f]\,(x_0)$, for f taking only finite values, can also be obtained as follows:

$$\mathsf{co}\,[f]\,(x_0) = \sup \{\ell(x_0) : \ \ell \ \text{affine} \quad \text{and} \quad \ell(y) \le f(y) \ \text{for all} \ y \in \mathbb{R}^n\} \quad (2.3)$$

with the $\sup$ attained by an affine function $\ell^* \in \mathrm{Aff}(\mathbb{R}^n)$.

(iii) If f is differentiable at x_0 and

$$f(x) \ge f(x_0) + Df(x_0) \cdot (x - x_0) \quad \text{for all } x \in \mathbb{R}^n, \quad (2.4)$$

then $\mathsf{co}[f](x_0) = f(x_0)$.

Next we recall the definition of Hausdorff distance between two non-empty sets [1], which measures how far the sets are from each other. To do so, we first need the notion of δ-neighbourhood of a set, and also define the diameter of a set.

Definition 2.2. *Given a non-empty subset $E \subset \mathbb{R}^n$ and $\delta > 0$, we define the δ-neighbourhood E^δ of E by*

$$E^\delta = \{x \in \mathbb{R}^n, \ \mathrm{dist}(x; E) < \delta\}.$$

where $\mathrm{dist}(x; E) = \inf\{|x - y|, \ y \in E\}$, and the diameter of E by

$$d_E := \sup\{|x - y|, \ x, \ y \in E\}.$$

Definition 2.3. *Let E, F be non-empty subsets of $\mathbb{R}^n$. The Hausdorff distance between E and F is defined by*

$$\mathrm{dist}_{\mathcal{H}}(E, F) = \inf\left\{\delta > 0 : F \subset E^\delta \ and \ E \subset F^\delta\right\}. \tag{2.5}$$

For general closed sets K, $G \subset \mathbb{R}^n$, if there is some $\delta > 0$ such that $K \subset G^\delta$, $G \subset K^\delta$, then the Hausdorff distance between F and G is finite and is given by (2.5). Otherwise we say that $\mathrm{dist}_{\mathcal{H}}(K, G) = +\infty$.

We now list some properties of the quadratic compensated convex transforms. Recall first the following ordering properties [47]:

$$C_\lambda^l(f)(x) \leq f(x) \leq C_\lambda^u(f)(x), \quad x \in \mathbb{R}^n, \tag{2.6}$$

whereas for $f \leq g$ in $\mathbb{R}^n$, we have that

$$C_\lambda^l(f)(x) \leq C_\lambda^l(g)(x) \quad \text{and} \quad C_\lambda^u(f)(x) \leq C_\lambda^u(g)(x), \quad x \in \mathbb{R}^n. \tag{2.7}$$

Furthermore, the compensated convex transforms are affine invariant [48], that is,

$$C_\lambda^l(f + \ell) = C_\lambda^l(f) + \ell \quad \text{and} \quad C_\lambda^u(f + \ell) = C_\lambda^u(f) + \ell \tag{2.8}$$

where ℓ is any affine function, and we also have [47, Theorem 2.1(iii)]

$$C_\tau^u(C_\lambda^u(f)) = \begin{cases} C_\lambda^u(f) & \text{if } \tau \geq \lambda, \\ C_\tau^u(f) & \text{if } \tau \leq \lambda; \end{cases} \quad \text{and} \quad C_\tau^l(C_\lambda^l(f)) = \begin{cases} C_\lambda^l(f) & \text{if } \tau \geq \lambda, \\ C_\tau^l(f) & \text{if } \tau \leq \lambda. \end{cases} \tag{2.9}$$

The following translation-invariance property will often be used in our proofs, since it allows us to refer our arguments to the point $x_0 = 0$ without loss of generality.

Proposition 2.4. *(Translation-invariance property) For any $f : \mathbb{R}^n \mapsto \mathbb{R}$ bounded below and for any affine function $\ell : \mathbb{R}^n \mapsto \mathbb{R}$, $\mathsf{co}[f + \ell] = \mathsf{co}[f] + \ell$. Consequently, both $C_\lambda^u(f)$ and $C_\lambda^l(f)$ are translation invariant against the weight function, that is:*

$$C_\lambda^l(f)(x) = \mathsf{co}\left[\lambda|(\cdot) - x_0|^2 + f\right](x) - \lambda|x - x_0|^2,$$

$$C_\lambda^u(f)(x) = \lambda|x - x_0|^2 - \mathsf{co}\left[\lambda|(\cdot) - x_0|^2 - f\right](x)$$

8

for all $x \in \mathbb{R}^n$ and for every fixed x_0. Hence, at x_0,

$$C_\lambda^l(f)(x_0) = \mathsf{co}[\lambda|(\cdot) - x_0|^2 + f](x_0), \quad C_\lambda^u(f)(x_0) = -\mathsf{co}[\lambda|(\cdot) - x_0|^2 - f](x_0).$$

For some theoretical developments and proofs, it can be convenient to view the lower and upper compensated convex transforms as parametrized semiconvex and semiconcave envelopes, respectively. We recall the following definition from [15, 19].

Definition 2.5. *A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is called 2λ-semiconvex (respectively, 2λ-semiconcave) if $x \mapsto f(x) + \lambda|x|^2$ (respectively, if $x \mapsto -f(x) + \lambda|x|^2$) is convex.*

Remark 2.6. *In convex analysis, the semiconvexity property as given by Definition 2.5 is sometimes also referred to as the uniform lower-C^2 property; compare Definition 2.5 with that of lower-C^2 in [13, page 228]. Such functions enjoy local regularity properties; note, for instance, the characterization of the locally Lipschitz functions as locally lower-C^2 - see [38, Theorem 6] and [15, Theorem 2.1.7].*

In our approximation theorems for bounded and uniformly continuous functions f, we make use of the modulus of continuity of f, which is defined as follows [20].

Definition 2.7. *Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a bounded and uniformly continuous function in $\mathbb{R}^n$. Then,*

$$\omega_f : t \in [0, \infty) \mapsto \omega_f(t) = \sup\left\{|f(x) - f(y)| : x, y \in \mathbb{R}^n \text{ and } |x - y| \le t\right\} \tag{2.10}$$

is called the modulus of continuity of f.

The modulus of continuity of f has the following properties.

Proposition 2.8. *Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a bounded and uniformly continuous function in $\mathbb{R}^n$. Then the modulus of continuity ω_f of f satisfies the following properties:*

(i) $\quad \omega_f(t) \to \omega_f(0) = 0, \text{ as } t \to 0;$

(ii) $\quad \omega_f \text{ is non-negative and non-decreasing continuous function on } [0, \infty);$ $\tag{2.11}$

(iii) $\quad \omega_f \text{ is subadditive: } \omega_f(t_1 + t_2) \le \omega_f(t_1) + \omega_f(t_2) \text{ for all } t_1, t_2 \ge 0.$

Any function ω defined on $[0, \infty)$ and satisfying $(2.11)(i), (ii), (iii)$ is called *a modulus of continuity*. A modulus of continuity ω can be bounded from above by an affine function (see [20, Lemma 6.1]), that is, there exist constants $a > 0$ and $b \ge 0$ such that

$$\omega(t) \le at + b \quad (\text{for all } t \ge 0). \tag{2.12}$$

As a result, given ω_f, one can define the least concave majorant of ω_f, which we denote by ω, which is also a modulus of continuity with the property (see [20])

$$\frac{1}{2}\omega(t) \le \omega_f(t) \le \omega(t) \quad (\text{for all } t \in [0, \infty)). \tag{2.13}$$

3 Approximations and Interpolations

In this section we consider the general theory of our interpolation and approximation problem when either $K \subset \mathbb{R}^n$ is compact or $K = \mathbb{R}^n \setminus \Omega$ where $\Omega \subset \mathbb{R}^n$ is a bounded open set.

Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a bounded function and denote by $f_K : K \subset \mathbb{R}^n \mapsto \mathbb{R}$ the restriction of f to K. A function $g : \mathsf{co}[K] \subset \mathbb{R}^n \mapsto \mathbb{R}$ is said to be an *interpolation* of f_K if $g = f$ in K, while for $\lambda > 0$, a family of functions $g_\lambda : \mathsf{co}[K] \subset \mathbb{R}^n \mapsto \mathbb{R}$ is said to *approximate* f if $\lim_{\lambda \to +\infty} g_\lambda = f$ uniformly in K.

We will see that the precise approximation and interpolation properties of f_K depend on the smoothness of the function f under consideration.

The following is a first simple observation.

Proposition 3.1. *Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a bounded 2λ-semiconvex (respectively, 2λ-semiconcave) function and $K \subset \mathbb{R}^n$ a non-empty closed set. If $|f(x)| < M$ for all $x \in \mathbb{R}^n$, then for any $\tau \geq \lambda$, $C^l_\tau(f^M_K)$ (respectively, $C^u_\tau(f^{-M}_K)$) is an interpolation of f_K, that is,*

$$C^l_\tau(f^M_K)(x) = f(x) \quad (\text{respectively, } C^u_\tau(f^{-M}_K)(x) = f(x)), \quad x \in K.$$

In general, if we are given sample values only in a closed set without any knowledge of the underlying function, we do not know whether or not our transforms are approximations of the original function. However, for any bounded function $f : \mathbb{R}^n \mapsto \mathbb{R}$, we have [48, Prop. 3.1]

$$C^l_\lambda(f)(x) = C^l_\lambda(\underline{f})(x), \quad \text{and} \quad C^u_\lambda(f)(x) = C^u_\lambda(\overline{f})(x) \tag{3.1}$$

for all $x \in \mathbb{R}^n$, where $\overline{f}$ and $\underline{f}$ are the upper and lower semicontinuous closures of f, respectively, and

$$\lim_{\lambda \to \infty} C^u_\lambda(f)(x) = \overline{f}(x), \qquad \lim_{\lambda \to \infty} C^l_\lambda(f)(x) = \underline{f}(x) \tag{3.2}$$

for all $x \in \mathbb{R}^n$. As a result, we have the following general approximation theorem.

Theorem 3.2. *Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be bounded, such that $|f(x)| < M$ for all $x \in \mathbb{R}^n$, and let $K \subset \mathbb{R}^n$ be non-empty compact set. Then for all $x \in \mathbb{R}^n$,*

$$\lim_{\lambda \to +\infty} C^u_\lambda(f^{-M}_K)(x) = \overline{f^{-M}_K}(x), \qquad \lim_{\lambda \to +\infty} C^l_\lambda(f^M_K)(x) = \underline{f^M_K}(x),$$
$$\lim_{\lambda \to +\infty} A^M_\lambda(f_K)(x) = \frac{1}{2}(\underline{f^M_K}(x) + \overline{f^{-M}_K}(x)), \tag{3.3}$$

and if $f : \mathbb{R}^n \mapsto \mathbb{R}$ is continuous, then for all $x \in K$,

$$\lim_{\lambda \to +\infty} C^u_\lambda(f^{-M}_K)(x) = f(x), \qquad \lim_{\lambda \to +\infty} C^l_\lambda(f^M_K)(x) = f(x),$$
$$\lim_{\lambda \to +\infty} A^M_\lambda(f_K)(x) = f(x), \tag{3.4}$$

and the convergence in (3.4) is uniform on K.

Note that the equalities $\overline{f^{-M}_K} = (\overline{f})^{-M}_K$ and $\underline{f^M_K} = (\underline{f})^M_K$ do not hold in general. For example, in $\mathbb{R}$, if we define $f(x) = 1$ if x is rational, $f(x) = -1$ if x is irrational and take $M = 2$, then we have $\overline{f} \equiv 1$ and $\underline{f} \equiv -1$. But if we choose K to be a finite set of rational numbers, then $\underline{f^M_K} = \chi_K + 2\chi_{\mathbb{R} \setminus K}$, whereas $(\underline{f})^M_K = -\chi_K + 2\chi_{\mathbb{R} \setminus K}$.

Note also that Theorem 3.2 suggests that we can apply our approximation methods to data sets which may not define a function, as discussed in Remark 1.2.

The following proposition provides conditions sufficient to ensure that our average approximation does not attain the value M or $-M$.

Proposition 3.3. *Let $K \subset \mathbb{R}^n$ be a non-empty compact set and denote by d_K the diameter of K. Suppose $f_K : K \subset \mathbb{R}^n \mapsto \mathbb{R}$ is bounded, with $|f_K(x)| \leq A_0$ for all $x \in K$. Then for $\lambda > 0$ and $M > A_0 + \lambda d_K^2$,*

$$-A_0 \leq C_\lambda^l(f_K^M)(x) < M, \quad -M < C_\lambda^u(f_K^{-M})(x) \leq A_0, \quad -M < A_\lambda^M(f_K)(x) < M,$$

for all $x \in \mathsf{co}[K]$.

Next we state our weak maximum principle. To make our statement simpler, we assume that the median of values of f_K over K is zero, which in practice can be easily satisfied by a simple translation of values.

Theorem 3.4. *Let $K \subset \mathbb{R}^n$ be a non-empty compact set. Suppose $f_K : K \mapsto \mathbb{R}$ is bounded and assume that*

$$m(f_K) := \frac{1}{2}\Big(\sup\{f_K(x),\ x \in K\} + \inf\{f_K(x),\ x \in K\}\Big) = 0. \tag{3.5}$$

Then

$$\inf\{f_K(y),\ y \in K\} \leq A_\lambda^\infty(f_K)(x) \leq \sup\{f_K(y),\ y \in K\}$$

for all $x \in \mathsf{co}[K]$ and $\lambda > 0$.

Before stating the error estimates for our approximations, we introduce the notions of density radius and convex density radius.

Definition 3.5. *Suppose $K \subset \mathbb{R}^n$ is a non-empty closed set, and denote by $\mathrm{dist}(x;\ K)$ the Euclidean distance of x to K.*

(i) *For $x \in \mathsf{co}[K]$, the density radius $r_d(x)$ of x with respect to K is just the Euclidean distance of x to K, i.e. we set $r_d(x) = \mathrm{dist}(x;\ K)$, whereas the density radius of K in $\mathsf{co}[K]$ is defined as*

$$r_d(K) = \sup\{r_d(x),\ x \in \mathsf{co}[K]\}.$$

(ii) *For $x \in \mathsf{co}[K]$, consider the balls $B(x;\ r)$ such that $x \in \mathsf{co}[\bar{B}(x;\ r) \cap K]$. The convex density radius of x with respect to K is defined as follows*

$$r_c(x) = \inf\{r \geq 0,\ x \in \mathsf{co}[\bar{B}(x;\ r) \cap K]\},$$

whereas the convex density radius of K in $\mathsf{co}[K]$ is defined by

$$r_c(K) = \sup\{r_c(x),\ x \in \mathsf{co}[K]\}.$$

From the definition above, we see that if K is compact or K^c is a bounded open set, $r_c(K)$ is finite. The convex density radius is zero if K is convex. If A and B are two compact sets such that $K \subset A \subset B \subset \mathsf{co}[K]$, then $r_c(A) \geq r_c(B)$. Also, the smaller $r_c(K)$ is, the denser the set K is in $\mathsf{co}[K]$. In general, if K is compact, $r_c(K)$ can be as large as the diameter of K. In this case, in order to make the convex density radius of K small, we require not only the density radius $r_d(K)$ of K in $\mathsf{co}[K]$ to be small but also that K is 'dense' in $\partial \mathsf{co}[K]$, the relative boundary of $\mathsf{co}[K]$. If K^c is bounded, then $r_c(K)$ can be as large as the diameter of K^c.

The following is a simple illustrative example for the case K compact. Consider the box $D = \{(x, y) \in \mathbb{R}^2,\ |x| < 1,\ |y| < 1\}$. For any $\delta > 0$, let $K_0 \subset D$ be a finite set with $r_d(K_0) < \delta$, so that

$\bar{D} \subset K_0^\delta$, and let $K = K_0 \cup \{\pm 1, \pm 1\}$. Then $r_d(K_0) < \delta$, whereas $r_c(K) = 1$ if we consider, say, the point $(1, 0) \in \mathsf{co}[K] = \bar{D}$.

We now formulate error estimates for our average approximations. Consider first the case when K is compact and $M = +\infty$. The estimates are expressed in terms of the modulus of continuity of the underlying uniformly continuous function f and the convex density radius. As special cases, we also consider bounded Lipschitz functions and $C^{1,1}$ functions.

Theorem 3.6. *Suppose $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a bounded uniformly continuous function satisfying $|f(x)| \leq A_0$ for some constant $A_0 > 0$ and all $x \in \mathbb{R}^n$, and let $K \subset \mathbb{R}^n$ be a non-empty compact set.*

(i) *Denote by ω the least concave majorant of the modulus of continuity ω_f of f. Let $a \geq 0$, $b \geq 0$ be such that $\omega(t) \leq at + b$ for $t \geq 0$. Then for all $\lambda > 0$ and $x \in \mathsf{co}[K]$,*

$$|A_\lambda^\infty(f_K)(x) - f(x)| \leq \omega \left(r_c(x) + \frac{a}{\lambda} + \sqrt{\frac{2b}{\lambda}} \right), \tag{3.6}$$

where $r_c(x) \geq 0$ is the convex density radius of x with respect to K.

(ii) *If we further assume that f is a globally Lipschitz function with Lipschitz constant $L > 0$, then for all $\lambda > 0$ and $x \in \mathsf{co}[K]$,*

$$|A_\lambda^\infty(f_K)(x) - f(x)| \leq Lr_c(x) + \frac{L^2}{\lambda}. \tag{3.7}$$

(iii) *If we further assume that f is a $C^{1,1}$ function satisfying $|Df(x) - Df(y)| \leq L|x - y|$ for $x, y \in \mathbb{R}^n$ and for some fixed $L > 0$, then for all $\lambda > L$ and $x \in \mathsf{co}[K]$,*

$$|A_\lambda^\infty(f_K)(x) - f(x)| \leq \frac{L}{4} \left(\frac{\lambda + L/2}{\lambda - L/2} + 1 \right) r_c^2(x). \tag{3.8}$$

Furthermore, in case (iii), $A_\lambda^\infty(f_K)$ is an interpolation of f_K^∞ in $\mathsf{co}[K]$.

Next we consider the case when $\Omega \subset \mathbb{R}^n$ is a non-empty bounded open set and define $K = \Omega^c := \mathbb{R}^n \setminus \Omega$. Clearly, $\mathsf{co}[K] = \mathbb{R}^n$ for such K. We then have the following estimate of the average approximation $A_\lambda^M(f_K)$.

Theorem 3.7. *Suppose $f : \mathbb{R}^n \mapsto \mathbb{R}$ is bounded and uniformly continuous, satisfying $|f(x)| \leq A_0$ for some constant $A_0 > 0$ and all $x \in \mathbb{R}^n$. Let $\Omega \subset \mathbb{R}^n$ be a bounded open set and $K = \Omega^c$. Denote by d_Ω the diameter of Ω.*

(i) *Let ω be the least concave majorant of the modulus of continuity ω_f of f. Assume $a \geq 0$, $b \geq 0$ are such that $\omega(t) \leq at + b$ for $t \geq 0$. Then for $\lambda > 0$, $M > A_0 + \lambda d_\Omega^2$ and all $x \in \mathbb{R}^n$, we have*

$$|A_\lambda^M(f_K)(x) - f(x)| \leq \omega \left(r_c(x) + \frac{a}{\lambda} + \sqrt{\frac{2b}{\lambda}} \right), \tag{3.9}$$

where $r_c(x) \geq 0$ is the convex density radius of x with respect to K.

(ii) *If we further assume that f is a globally Lipschitz function with Lipschitz constant $L > 0$, then for $\lambda > 0$, $M > A_0 + \lambda d_\Omega^2$ and all $x \in \mathbb{R}^n$, we have*

$$|A_\lambda^M(f_K)(x) - f(x)| \leq Lr_c(x) + \frac{L^2}{\lambda}. \tag{3.10}$$

(iii) If we further assume that f is a $C^{1,1}$ function such that $|Df(x) - Df(y)| \leq L|x - y|$ for all x, $y \in \mathbb{R}^n$, where $L > 0$ is a constant, then for $\lambda > L$, $M > A_0 + \lambda d_\Omega^2$ and all $x \in \mathbb{R}^n$, we have

$$|A_\lambda^M(f_K)(x) - f(x)| \leq \frac{L}{4}\left(\frac{\lambda + L/2}{\lambda - L/2} + 1\right) r_c^2(x). \tag{3.11}$$

Furthermore, in case (iii), $A_\lambda^M(f_K)$ is an interpolation of f_K in $\mathbb{R}^n$.

Remark 3.8. *Theorem 3.6 can be used for the solution of practical problems such as salt & pepper noise removal, in which case K is the compact set given by the part of the image which is noise free. On the other hand, Theorem 3.7 can be applied, for instance, to inpainting of damaged images, where Ω will be the domain to be inpainted using information about f_K, with $K = \Omega^c$. We will discuss these applications of the theory developed here in our follow-on paper [51].*

The following corollary of Theorem 3.7 can be thought of as an extension of Theorem 3.6, which concern $A_\lambda^\infty(f_K)$, to the case of finite $M > 0$, under an extra restriction.

Corollary 3.9. *Suppose $f : \mathbb{R}^n \mapsto \mathbb{R}$ is bounded and uniformly continuous, with $|f(x)| \leq A_0$ for some constant $A_0 > 0$ and all $x \in \mathbb{R}^n$. Assume that $f(x) = c_0$ for $|x| \geq r > 0$, where $c_0 \in \mathbb{R}$ and $r > 0$ are constants. Let $K \subset \mathbb{R}^n$ be a non-empty compact set satisfying $K \subset \bar{B}(0; r)$. For $R > r$, define $K_R := K \cup B^c(0; R)$.*

(i) Let ω be the least concave majorant of the modulus of continuity ω_f of f. Assume $a \geq 0$, $b \geq 0$ are such that $\omega(t) \leq at + b$ for $t \geq 0$. Then for $\lambda > 0$, $M > A_0 + \lambda(R + r)^2$ and all $x \in \mathsf{co}[K]$, we have

$$|A_\lambda^M(f_{K_R})(x) - f(x)| \leq \omega\left(r_c(x) + \frac{a}{\lambda} + \sqrt{\frac{2b}{\lambda}}\right). \tag{3.12}$$

(ii) If we further assume that f is a globally Lipschitz function with Lipschitz constant $L > 0$, then for $\lambda > 0$, $M > A_0 + \lambda(R + r)^2$ and all $x \in \mathsf{co}[K]$, we have

$$|A_\lambda^M(f_{K_R})(x) - f(x)| \leq Lr_c(x) + \frac{L^2}{\lambda}. \tag{3.13}$$

(iii) If we further assume that f is a $C^{1,1}$ function such that $|Df(x) - Df(y)| \leq L|x - y|$ for all x, $y \in \mathbb{R}^n$ and $L > 0$ is a constant, then for $\lambda > L$, $M > A_0 + \lambda(R + r)^2$ and all $x \in \mathsf{co}[K]$, we have

$$|A_\lambda^M(f_{K_R})(x) - f(x)| \leq \frac{L}{4}\left(\frac{\lambda + L/2}{\lambda - L/2} + 1\right) r_c^2(x). \tag{3.14}$$

Furthermore, in case (iii), $A_\lambda^M(f_{K_R})$ is an interpolation of f_K in $\mathbb{R}^n$.

Remark 3.10. *Corollary 3.9 can be viewed as an extrapolation result for bounded uniformly continuous functions and for globally Lipschitz functions defined on a compact set. For example, we can define f to be zero outside a large ball containing K and then apply Theorem 3.7. Another reason for such extensions is that if we simply replace $+\infty$ by a finite $M > 0$ in Theorem 3.6, we are not able to obtain an error estimate independent of M, particularly near the boundary of $\mathsf{co}[K]$.*

4 Hausdorff Stability and Regularity

In this section we establish stability and regularity results for our approximations. The stability properties will be expressed in terms of a notion of Hausdorff continuity, and we first introduce a definition of Hausdorff continuity with respect to closed samples for transforms of bounded functions on $\mathbb{R}^n$.

Definition 4.1. *Let $\mathcal{B}(\mathbb{R}^n)$ be the class of bounded real-valued functions from $\mathbb{R}^n$ to $\mathbb{R}$ and choose a fixed $f \in \mathcal{B}(\mathbb{R}^n)$ that is uniformly continuous. A transform $\mathcal{T} : \mathcal{B}(\mathbb{R}^n) \to \mathcal{B}(\mathbb{R}^n)$ is said to be Hausdorff continuous with respect to closed sample sets at f if the mapping $K \mapsto \mathcal{T}(f\chi_K)$ is Hausdorff continuous at each non-empty closed set $K_0 \subset \mathbb{R}^n$, in the sense that for every $\epsilon > 0$, there exists $\delta > 0$ such that*
$$|\mathcal{T}(f\chi_K)(x) - \mathcal{T}(f\chi_{K_0})(x)| < \epsilon$$
for all $x \in \mathbb{R}^n$ whenever K is a non-empty closed set with $\mathrm{dist}_{\mathcal{H}}(K, K_0) < \delta$, and to be uniformly Hausdorff continuous with respect to closed sample sets at f if δ can be chosen independently of K_0.

A transform $\mathcal{T} : \mathcal{B}(\mathbb{R}^n) \to \mathcal{B}(\mathbb{R}^n)$ is said to be Hausdorff-Lipschitz continuous with respect to closed sample sets at f if the mapping $K \mapsto \mathcal{T}(f\chi_K)$ is Hausdorff-Lipschitz continuous, in the sense that there exists $L > 0$ such that

$$|\mathcal{T}(f\chi_K)(x) - \mathcal{T}(f\chi_G)(x)| \leq L \, \mathrm{dist}_{\mathcal{H}}(K, G)$$

for all $x \in \mathbb{R}^n$ whenever $K, G \subset \mathbb{R}^n$ are closed sets with $\mathrm{dist}_{\mathcal{H}}(K, G) < \infty$.

Remark 4.2. *It is well known that the Euclidean distance function to a non-empty closed set K, i.e. the mapping $K \mapsto \mathrm{dist}(\cdot; K)$, is Hausdorff-Lipschitz continuous in the sense that $|\mathrm{dist}(x; K) - \mathrm{dist}(x; G)| \leq \mathrm{dist}_{\mathcal{H}}(K, G)$ for all $x \in \mathbb{R}^n$ and non-empty closed sets K, G, and this is, to our knowledge, the only well-known example of a function satisfying a Hausdorff-Lipschitz property. A further example, which we will extend here, is given in [48, Theorem 5.5], where it is shown that the mapping $K \mapsto C_\lambda^u(\chi_K)$ is Hausdorff-Lipschitz continuous when K is compact.*

Our first objective is to show that the mappings $K \mapsto L_\lambda^M(f_K)$, $K \mapsto U_\lambda^M(f_K)$ and $K \mapsto A_\lambda^M(f_K)$ are uniformly Hausdorff continuous for every bounded uniformly continuous function f with $\sup_{x \in \mathbb{R}^n} |f(x)| < M$.

Lemma 4.3. *Suppose $f : \mathbb{R}^n \mapsto \mathbb{R}$ is bounded and uniformly continuous, with $\sup_{\mathbb{R}^n} |f(x)| \leq A_0$ for some constant $A_0 > 0$, and let $M > A_0$. Then for any fixed $\lambda > 0$ and any non-empty closed set $K \subset \mathbb{R}^n$,*

$$C_\lambda^l(f_K^M) = M - C_\lambda^u((M - f)\chi_K) \quad and \quad C_\lambda^u(f_K^{-M}) = -M + C_\lambda^u((M + f)\chi_K).$$

Now by the assumption that $\sup_{\mathbb{R}^n} |f(x)| \leq A_0 < M$, both $M + f$ and $M - f$ are strictly positive in $\mathbb{R}^n$. Hence, by Definition 1.1 and Lemma 4.3, the Hausdorff continuity of the mappings $K \mapsto L_\lambda^M(f_K)$, $K \mapsto U_\lambda^M(f_K)$ and $K \mapsto A_\lambda^M(f_K)$ reduces to the Hausdorff continuity of $K \mapsto C_\lambda^u(g_K)$ for uniformly continuous functions $g : \mathbb{R}^n \mapsto \mathbb{R}$ satisfying

$$0 < M - A_0 \leq g(x) < M \quad \text{for all } x \in \mathbb{R}^n. \tag{4.1}$$

We will thus extend [48, Theorem 5.5], that proved Hausdorff-Lipschitz continuity of $K \mapsto C_\lambda^u(\chi_K)$ corresponding to the special case $f \equiv 1$ in Definition 4.1, to the general case of any bounded

uniformly continuous function f. In the terminology of Definition 4.1, we will show that the upper transform C_λ^u is uniformly Hausdorff continuous with respect to closed sample sets at each bounded uniformly continuous function f, and is Hausdorff-Lipschitz continuous with respect to closed sample sets at each such f that is also globally Lipschitz continuous.

Motivated by the analysis in [48], we introduce a squared distance-like function $D_{\lambda,f}^2$, the upper transform of which is equal to the upper transform $C_\lambda^u(f_K)$ of f_K and which proves to be a useful tool in the following.

Definition 4.4. *For $f : \mathbb{R}^n \mapsto \mathbb{R}$ with $0 < f \le M$, we define the following distance-like functions for a closed set $K \subset \mathbb{R}^n$:*

$$d_{\lambda,f}(x,\,K) = \inf\left\{|y - x| - \sqrt{\frac{f(y)}{\lambda}},\ y \in K\right\}, \quad x \in \mathbb{R}^n; \tag{4.2}$$

and

$$D_{\lambda,f}(x,\,K) = -\sqrt{\lambda}\min\{0,\,d_{\lambda,f}(x,\,K)\}, \quad x \in \mathbb{R}^n. \tag{4.3}$$

Remark 4.5. *In the definition of $d_{\lambda,f}(x,\,K)$, if f is continuous on $\mathbb{R}^n$ and K is closed, the minimum in (4.2) is attained, that is, for every $x \in \mathbb{R}^n$, there exists $x^K \in K$ such that $d_{\lambda,f}(x,\,K) = |x^K - x| - \sqrt{f(x^K)/\lambda}$. Thus if f is continuous, the 'inf' in (4.2) can be replaced by 'min'.*

In Theorem 4.10 below, we will follow an indirect approach to proving the Hausdorff continuity of $C_\lambda^u(f_K)$ that exploits the squared function $D_{\lambda,f}^2(x;\,K)$. Note that it is also possible to give a direct proof of Hausdorff continuity, avoiding use of $D_{\lambda,f}^2(x;\,K)$, which yields a weaker result, namely that for every $\epsilon > 0$, there exists $\delta > 0$ such that $|C_\lambda^u(f_K)(x)) - C_\lambda^u(f_E)(x))| < \epsilon$ whenever $\text{dist}_{\mathcal{H}}(K, E) < \delta$. Additionally, we can derive a Hausdorff continuity result using the Hausdorff continuity of the Moreau envelopes, since $C_\lambda^u(f) = M_\lambda(M^\lambda(f)))$ [48], and it can be shown that

$$|M_\lambda(f_K)(x) - M_\lambda(f_E)(x)| \le 2\lambda\left(\text{dist}_{\mathcal{H}}(K,E) + \sqrt{\frac{2M}{\lambda}}\right)\text{dist}_{\mathcal{H}}(K,E) + \omega(\text{dist}_{\mathcal{H}}(K,E)),$$

$$|M^\lambda(f_K)(x) - M^\lambda(f_E)(x)| \le 2\lambda\left(\text{dist}_{\mathcal{H}}(K,E) + \sqrt{\frac{2M}{\lambda}}\right)\text{dist}_{\mathcal{H}}(K,E) + \omega(\text{dist}_{\mathcal{H}}(K,E)),$$

from which a version of Hausdorff continuity of $C_\lambda^u(f_K)$ follows.

Note that the function $D_{\lambda,f}(x;\,K)$ defined in (4.3) is a generalisation of $D_\lambda(x;\,K)$ for the characteristic function χ_K introduced in [48, Definition 5.1], since if we take $f \equiv 1$ in (4.3), we have

$$D_\lambda(x;\,K) = \max\left\{0,\,1 - \sqrt{\lambda}\text{dist}(x;\,K)\right\} = \max\left\{0,\,\sqrt{\lambda}\left(\frac{1}{\sqrt{\lambda}} - \text{dist}(x;\,K)\right)\right\}$$

$$= \sqrt{\lambda}\max\left\{0,\,-\left(\min_{y\in K}|y - x| - \frac{1}{\sqrt{\lambda}}\right)\right\}$$

$$= -\sqrt{\lambda}\min\left\{0,\,\min_{y\in K}\left\{|y - x| - 1/\sqrt{\lambda}\right\}\right\} = D_{\lambda,f}(x;\,K).$$

As well as being a tool to investigate the stability of the upper compensated convex transform of characteristic functions, the geometry-based function $D_\lambda(x; K)$ has also, for instance, been used to find geometric features such as interior corners [52]. Hence our generalised function $D_{\lambda,f}(x, K)$ might also have other applications which we will explore elsewhere.

We start by stating a few preliminary lemmas, the proofs of which are given in Section 5.

Lemma 4.6. *Suppose* $f : \mathbb{R}^n \mapsto \mathbb{R}$ *is bounded and uniformly continuous such that for some constant* $M > 0$, $0 < f(x) \leq M$ *for all* $x \in \mathbb{R}^n$. *Let* ω *be the least concave majorant of the modulus of continuity of* $\sqrt{f}$, *which is itself a modulus of continuity. Let* $K, E \subset \mathbb{R}^n$ *be non-empty closed sets with* $\mathrm{dist}_{\mathcal{H}}(K, E) < +\infty$. *Then for all* $x \in \mathbb{R}^n$,

$$|d_{\lambda,f}(x, K) - d_{\lambda,f}(x, E)| \leq \mathrm{dist}_{\mathcal{H}}(K, E) + \frac{\omega(\mathrm{dist}_{\mathcal{H}}(K, E))}{\sqrt{\lambda}}.$$

Lemma 4.7. *Under the assumptions of Lemma 4.6, we have*

$$|D_{\lambda,f}^2(x, K) - D_{\lambda,f}^2(x, E)| \leq 2\sqrt{\lambda M}\mathrm{dist}_{\mathcal{H}}(K, E) + 2\sqrt{M}\omega(\mathrm{dist}_{\mathcal{H}}(K, E)), \quad x \in \mathbb{R}^n. \qquad (4.4)$$

Lemma 4.8. *Suppose* $\alpha > 0$ *is a constant and* $x_0 \in \mathbb{R}^n$, *then for* $\lambda > 0$,

$$C_\lambda^u(\alpha\chi_{\{x_0\}})(x) = \begin{cases} \lambda(|x - x_0| - \sqrt{\alpha/\lambda})^2, & |x - x_0| \leq \sqrt{\alpha/\lambda}, \\ 0, & |x - x_0| \geq \sqrt{\alpha/\lambda}. \end{cases}$$

Lemma 4.9. *Suppose* f *satisfies the assumptions of Lemma 4.6 and* $K \subset \mathbb{R}^n$ *is closed. Then for* $\lambda > 0$ *and for all* $x \in \mathbb{R}^n$,

$$C_\lambda^u(f\chi_K)(x) = C_\lambda^u(D_{\lambda,f}^2(\cdot\,;\,K))(x).$$

We are now in a position to state our key result on the Hausdorff stability of the upper compensated convex transform with respect to closed sample sets at a bounded uniformly continuous positive function f.

Theorem 4.10. *Suppose* $f : \mathbb{R}^n \mapsto \mathbb{R}$ *is bounded and uniformly continuous, with* $0 < f(x) \leq M$ *for some constant* $M > 0$ *and all* $x \in \mathbb{R}^n$. *Let* ω *be the least concave majorant of the modulus of continuity of* $\sqrt{f}$, *which is itself a modulus of continuity. Let* $K, E \subset \mathbb{R}^n$ *be non-empty closed sets with* $\mathrm{dist}_{\mathcal{H}}(K, E) < +\infty$. *Then for all* $x \in \mathbb{R}^n$, $C_\lambda^u(f\chi_K)$ *is Hausdorff continuous in the sense that*

$$|C_\lambda^u(f\chi_K)(x) - C_\lambda^u(f\chi_E)(x)| \leq 2\sqrt{\lambda M}\mathrm{dist}_{\mathcal{H}}(K, E) + 2\sqrt{M}\,\omega(\mathrm{dist}_{\mathcal{H}}(K, E)). \qquad (4.5)$$

Corollary 4.11. *Under the assumptions of Theorem 4.10, if we further assume that* $f : \mathbb{R}^n \mapsto \mathbb{R}$ *is a globally Lipschitz continuous function satisfying* $|f(x) - f(y)| \leq L|x - y|$ *and* $0 < \alpha \leq f \leq M$, *then for all* $x \in \mathbb{R}^n$, $C_\lambda^u(f\chi_K)$ *is Hausdorff-Lipschitz continuous in the sense that*

$$|C_\lambda^u(f\chi_K)(x) - C_\lambda^u(f\chi_E)(x)| \leq \left(2\sqrt{\lambda M} + L\sqrt{\frac{M}{\alpha}}\right) \mathrm{dist}_{\mathcal{H}}(K, E). \qquad (4.6)$$

16

We can now easily state the Hausdorff stability theorem for our approximations.

Theorem 4.12. *Suppose $f : \mathbb{R}^n \mapsto \mathbb{R}$ is bounded and uniformly continuous (respectively, globally Lipschitz continuous) and $|f(x)| \leq A_0$ for $x \in \mathbb{R}^n$. Then for $M > A_0$, the mappings $K \mapsto L_\lambda^M(f_K)$, $K \mapsto U_\lambda^M(f_K)$, $K \mapsto A_\lambda^M(f_K)$ and $K \mapsto (SA)_\lambda^M(f_K)$ are uniformly Hausdorff continuous (respectively, Hausdorff-Lipschitz continuous).*

We conclude this section by stating the regularity, or smoothness, of our approximations. Since our upper, lower and average approximations are globally Lipschitz functions on $\mathbb{R}^n$ and our mixed approximation is a $C^{1,1}$ function, we have the following.

Theorem 4.13. *Let $K \subset \mathbb{R}^n$ be a non-empty closed set and $f_K : K \subset \mathbb{R}^n \mapsto \mathbb{R}$ be a bounded function with $|f(x)| < M$ for all $x \in K$. Suppose $\lambda > 0$ and $\tau > 0$ are fixed. Then*

(i) *$U_\lambda^M(f_K)$, $L_\lambda^M(f_K)$ and $A_\lambda^M(f_K)$ are globally Lipschitz continuous on $\mathbb{R}^n$, with the Lipschitz constant bounded above by $8\sqrt{M\lambda}$;*

(ii) *$(SA)_{\lambda,\tau}^M(f_K)$ is a $C^{1,1}$ function on $\mathbb{R}^n$, and satisfies*

$$|(SA)_{\lambda,\tau}^M(f_K)(x) - A_\lambda^M(f_K)(x)| \leq \frac{16M\lambda}{\tau} \quad \text{for all } x \in \mathbb{R}^n. \tag{4.7}$$

5 Proofs of the Main Results

Proof of Proposition 3.1: Since f is 2λ-semiconvex, $C_\tau^l(f)(x) = f(x)$ for $x \in \mathbb{R}^n$ and $\tau \geq \lambda$. As $f(x) \leq f_K^M(x)$ for $x \in \mathbb{R}^n$, we have, for any $x \in K$,

$$f(x) = C_\tau^l(f)(x) \leq C_\tau^l(f_K^M)(x) \leq f_K^M(x) = f(x),$$

where we have applied the ordering property (2.6) to show that $C_\tau^l(f)(x) \leq C_\tau^l(f_K^M)(x)$ and (2.7) to state that $C_\tau^l(f_K^M)(x) \leq f_K^M(x)$. Thus $C_\tau^l(f_K^M)$ is an interpolation of f_K. Similarly, if f is 2λ-semiconcave, $C_\tau^u(f_K^{-M})$ is an interpolation of f_K. $\qquad\square$

Proof of Theorem 3.2: The first part is immediate from (3.1) and (3.2). If f is continuous, it follows from [47, Theorem 2.3(iii)] that, uniformly on any compact set,

$$\lim_{\lambda \to +\infty} C_\lambda^l(f)(x) = f(x), \qquad \lim_{\lambda \to +\infty} C_\lambda^u(f)(x) = f(x),$$

whereas the ordering properties (2.6) and (2.7) imply that

$$C_\lambda^l(f) \leq C_\lambda^l(f_K^M) \leq f_K^M, \quad C_\lambda^u(f) \geq C_\lambda^u(f_K^{-M}) \geq f_K^{-M}.$$

Since $f_K^M = f = f_K^{-M}$ on K, it follows that

$$\lim_{\lambda \to +\infty} C_\lambda^l(f_K^M)(x) = f(x), \quad \lim_{\lambda \to +\infty} C_\lambda^u(f_K^{-M})(x) = f(x), \quad \lim_{\lambda \to +\infty} A_\lambda^M(f_K)(x) = f(x)$$

uniformly on K, as required. $\qquad\square$

Proof of Proposition 3.3: By Proposition 2.4, without loss of generality, we may assume that $x = 0$. Taking the constant function $\ell(y) = -A_0$ for $y \in \mathbb{R}^n$, we see that $-A_0 = \ell(y) \leq f_K^M(y) + \lambda|y|^2$ so that $-A_0 \leq C_\lambda^l(f_K^M)(0)$.

Since $0 \in \mathsf{co}[K]$, by Proposition 2.1, there exist $x_i, \ldots, x_{n+1} \in \mathbb{R}^n$ with $\lambda_i \geq 0$ for $i = 1, \ldots, n+1$ such that $\sum_{i=1}^{n+1} \lambda_i = 1$ and $\sum_{i=1}^{n+1} \lambda_i x_i = 0$. We then have

$$C_\lambda^l(f_K^M)(0) = \mathsf{co}[f_K^M + \lambda | \cdot |^2](0) \leq \sum_{i=1}^{n+1} \lambda_i(f_K^M(x_i) + \lambda |x_i|^2) \leq \sum_{i=1}^{n+1} \lambda_i(A_0 + \lambda d_K^2) < M.$$

The proof for the upper transform follows similar arguments. □

Proof of Theorem 3.4: Let $\sup_K f = A_0$, so that by our assumption (3.5), $\inf_K f = -A_0$. Fix $x \in \mathsf{co}[K]$. By Proposition 2.4, without loss of generality, we assume that $x = 0$. Notice that K is compact, $C_\lambda^l(f_K^\infty)(0) = C_\lambda^l(\underline{f_K^\infty})(0)$, and $\underline{f_K^\infty}$ is lower semicontinuous. Also, $C_\lambda^u(f_K^{-\infty})(0) = C_\lambda^u(\overline{f_K^{-\infty}})(0)$, and $\overline{f_K^{-\infty}}$ is upper semicontinuous. Thus, by Proposition 2.1, there are two finite generating sets $K_l = \{x_i^-\}_{i=1}^{k_-} \subset K$ and $K_u = \{x_i^+\}_{i=1}^{k_+} \subset K$, two sets of positive numbers $\Lambda_l = \{\lambda_i^-\}_{i=1}^{k_-}$ and $\Lambda_u = \{\lambda_i^+\}_{i=1}^{k_+}$ satisfying $1 \leq k_-, k_+ \leq n+1$, $\sum_{i=1}^{k_-} \lambda_i^- = 1$, $\sum_{i=1}^{k_-} \lambda_i^- x_i^- = 0$, $\sum_{i=1}^{k_+} \lambda_i^+ = 1$, $\sum_{i=1}^{k_+} \lambda_i^+ x_i^+ = 0$, such that

$$\mathsf{co}[\underline{f_K^\infty} + \lambda | \cdot |^2](0) = \sum_{i=1}^{k_-} \lambda_i^-[\underline{f_K^\infty}(x_i^-) + \lambda |x_i^-|^2]$$
$$= \inf \left\{ \sum_{i=1}^{n+1} \lambda_i[\underline{f_K^\infty}(x_i) + \lambda |x_i|^2], \ x_i \in K_l \cup K_u, \ \lambda_i \geq 0, \ \sum_{i=1}^{n+1} \lambda_i = 1, \ \sum_{i=1}^{n+1} \lambda_i x_i = 0 \right\}$$
$$\geq B_0 - A_0,$$

where $B_0 = \lambda \inf \left\{ \sum_{i=1}^{n+1} \lambda_i |x_i|^2, \ x_i \in K_l \cup K_u, \ \lambda_i \geq 0, \ \sum_{i=1}^{n+1} \lambda_i = 1, \ \sum_{i=1}^{n+1} \lambda_i x_i = 0 \right\}$. Likewise

$$\mathsf{co}[\underline{f_K^\infty} + \lambda | \cdot |^2](0) \leq B_0 + A_0,$$

and thus

$$B_0 - A_0 \leq C_\lambda^l(f_K^\infty)(0) \leq B_0 + A_0.$$

On the other hand, we also have, by Proposition 2.1, that

$$\mathsf{co}[\lambda | \cdot |^2 - \overline{f_K^{-\infty}}](0) = \sum_{i=1}^{k_+} \lambda_i^+[\lambda |x_i^+|^2 - \overline{f_K}(x_i^+)]$$
$$= \inf \left\{ \sum_{i=1}^{n+1} \lambda_i[\lambda |x_i|^2 - \overline{f_K^{-\infty}}(x_i)], \ x_i \in K_l \cup K_u, \ \lambda_i \geq 0, \ \sum_{i=1}^{n+1} \lambda_i = 1, \ \sum_{i=1}^{n+1} \lambda_i x_i = 0 \right\}$$
$$\geq B_0 - A_0,$$

and similarly

$$\mathsf{co}[\lambda | \cdot |^2 - \overline{f_K^{-\infty}}](0) \leq B_0 + A_0,$$

so since $C_\lambda^u(f_K^{-\infty})(0) = -\mathsf{co}[\lambda | \cdot |^2 - \overline{f_K^{-\infty}}](0)$, we obtain

$$-B_0 - A_0 \leq C_\lambda^u(f_K^{-\infty})(0) \leq -B_0 + A_0.$$

Thus

$$-A_0 \leq A_\lambda^\infty(f_K)(0) \leq A_0,$$

which concludes the proof. $\qquad\square$

Proof of Theorem 3.6. *Part (i)*: By Proposition 2.4, without loss of generality we again assume that $x = 0$. Since both $y \mapsto \lambda|y|^2 + f_K^\infty(y)$ and $y \mapsto \lambda|y|^2 - f_K^{-\infty}(y)$ are coercive and lower semicontinuous, we have, by Proposition 2.1, that

$$C_\lambda^l(f_K^\infty)(0) = \sum_{j=1}^{k_l} \lambda_j^l(\lambda|x_j^l|^2 + f(x_j^l)), \quad -C_\lambda^u(f_K^{-\infty})(0) = \sum_{j=1}^{k_u} \lambda_j^u(\lambda|x_j^u|^2 - f(x_j^u)),$$

where $2 \le k_l,\ k_u \le n + 1$, $\lambda_j^l > 0$, $x_j^l \in K$, $j = 1,\ldots,k_l$, $\sum_{j=1}^{k_l} \lambda_j^l = 1$, $\sum_{j=1}^{k_l} \lambda_j^l x_j^l = 0$; $\lambda_j^u > 0$, $x_j^u \in K$, $j = 1,\ldots,k_u$, $\sum_{j=1}^{k_l} \lambda_j^u = 1$, $\sum_{j=1}^{k_l} \lambda_j^u x_j^u = 0$.
We also define

$$B_0 = \min\left\{\sum_{k=1}^{n+1} \lambda_k|x_k|^2,\ \lambda_k \ge 0,\ x_k \in K,\ k = 1, 2, \ldots, n+1, \sum_{k=1}^{n+1} \lambda_k = 1,\ \sum_{k=1}^{n+1} \lambda_k x_k = 0\right\}$$

$$= \sum_{k=1}^{m^*} \lambda_k^*|x_k^*|^2,$$

for some $2 \le m^* \le n + 1$, $\lambda_j^* > 0$, $x_j^* \in K^*$ for $j = 1, 2, \ldots, m^*$, $\sum_{j=1}^{m^*} \lambda_j^* = 1$ and $\sum_{j=1}^{m^*} \lambda_j^* x_j^* = 0$, and let

$$C_0 = \min\left\{\sum_{k=1}^{n+1} \lambda_k|x_k|^2,\ \lambda_k \ge 0,\ x_k \in \bar{B}_{r_c(0)}(0) \cap K,\ k = 1, \ldots, n+1, \sum_{k=1}^{n+1} \lambda_k = 1,\ \sum_{k=1}^{n+1} \lambda_k x_k = 0\right\}$$

$$= \sum_{k=1}^{n+1} \lambda_k^r|x_k^r|^2.$$

Clearly $C_0 \le r_c^2(0)$, and by definition,

$$B_0 = \sum_{k=1}^{m^*} \lambda_k^*|x_k^*|^2 \le C_0 \le r_c^2(0).$$

By the Cauchy-Schwarz inequality, we also have

$$\sum_{k=1}^{m^*} \lambda_k^*|x_k^*| \le r_c(0).$$

Now

$$\begin{aligned}
C_\lambda^l(f_K^\infty)(0) &\le \sum_{k=1}^{m^*} \lambda_k^*(\lambda|x_k^*|^2 + f(x_k^*)) = \lambda B_0 + f(0) + \sum_{k=1}^{m^*} \lambda_k^*(f(x_k^*) - f(0)) \\
&\le \lambda B_0 + f(0) + \sum_{k=1}^{m^*} \lambda_k^*\omega(|x_k^*|) \le \lambda B_0 + f(0) + \omega\left(\sum_{k=1}^{m^*} \lambda_k^*|x_k^*|\right) \\
&\le \lambda B_0 + f(0) + \omega(r_c(0)),
\end{aligned} \tag{5.1}$$

since ω is non-decreasing and concave. Furthermore, we also have

$$\lambda B_0 + f(0) + \omega(r_c(0)) \leq \lambda r_c^2(0) + f(0) + \omega(r_c(0)), \qquad (5.2)$$

and

$$
\begin{aligned}
C_\lambda^l(f_K^\infty)(0) = \sum_{j=1}^{k_l} \lambda_j^l(\lambda|x_j^l|^2 + f(x_j^l)) &\geq f(0) + \sum_{j=1}^{k_l} \lambda_j^l(\lambda|x_j^l|^2 - |f(x_j^l) - f(0)|) \\
&\geq f(0) + \sum_{j=1}^{k_l} \lambda_j^l(\lambda|x_j^l|^2 - \omega(|x_j^l|)) \geq f(0) + \sum_{j=1}^{k_l} \lambda_j^l(\lambda|x_j^l|^2 - a|x_j^l| - b).
\end{aligned}
\qquad (5.3)
$$

By comparing (5.1), (5.2) with (5.3), it follows that

$$f(0) + \sum_{j=1}^{k_l} \lambda_j^l(\lambda|x_j^l|^2 - a|x_j^l| - b) \leq \lambda r_c^2(0) + f(0) + \omega(r_c(0)),$$

and hence

$$\sum_{j=1}^{k_l} \lambda_j^l \left(|x_j^l| - \frac{a}{2\lambda}\right)^2 \leq r_c^2(0) + \frac{\omega(r_c(0))}{\lambda} + \frac{a^2}{4\lambda^2} + \frac{b}{\lambda} \leq r_c^2(0) + \frac{a}{\lambda} + \frac{a^2}{4\lambda^2} + \frac{2b}{\lambda} = \left(r_c(0) + \frac{a}{2\lambda}\right)^2 + \frac{2b}{\lambda}.$$

Here we have used the fact that $\omega(t) \leq at + b$ for $t \geq 0$. Thus by the Cauchy-Schwarz inequality,

$$\sum_{j=1}^{k_l} \lambda_j^l \left||x_j^l| - \frac{a}{2\lambda}\right| \leq \left(\left(r_c(0) + \frac{a}{2\lambda}\right)^2 + \frac{2b}{\lambda}\right)^{1/2} \leq r_c(0) + \frac{a}{2\lambda} + \sqrt{2b/\lambda},$$

so that

$$\sum_{j=1}^{k_l} \lambda_j^l |x_j^l| \leq r_c(0) + \frac{a}{\lambda} + \sqrt{2b/\lambda}.$$

Now

$$
\begin{aligned}
C_\lambda^l(f_K^\infty(0)) = \sum_{j=1}^{k_l} \lambda_j^l(\lambda|x_j^l|^2 + f(x_j^l)) &\geq f(0) + \sum_{j=1}^{k_l} \lambda_j^l(\lambda|x_j^l|^2 - |f(x_j^l) - f(0)|) \\
&\geq f(0) + \lambda B_0 - \sum_{j=1}^{k_l} \lambda_j^l \omega(|x_j^l|) \geq f(0) + \lambda B_0 - \omega\left(\sum_{j=1}^{k_l} \lambda_j^l |x_j^l|\right) \\
&\geq f(0) + \lambda B_0 - \omega\left(r_c(0) + a/\lambda + \sqrt{2b/\lambda}\right),
\end{aligned}
\qquad (5.4)
$$

and by combining (5.1) and (5.4), we obtain

$$f(0) + \lambda B_0 - \omega\left(r_c(0) + a/\lambda + \sqrt{2b/\lambda}\right) \leq C_\lambda^l(f_K^\infty)(0) \leq \lambda B_0 + f(0) + \omega(r_c(0)).$$

Similarly, we can prove

$$f(0) - \lambda B_0 - \omega(r_c(0)) \leq C_\lambda^u(f_K^{-\infty})(0) \leq f(0) + \omega\left(r_c(0) + a/\lambda + \sqrt{2b/\lambda}\right) - \lambda B_0,$$

20

and thus

$$|A_\lambda^\infty(f_K)(0) - f(0)| \leq \frac{1}{2}\left(\omega(r_c(0)) + \omega\left(r_c(0) + a/\lambda + \sqrt{2b/\lambda}\right)\right).$$

The proof of Part (i) is thus complete.

Part (ii): We only need to note that in this case, $\omega(t) = Lt$ for $t \geq 0$, taking $a = L$ and $b = 0$. The result then follows.

Part (iii): By Proposition 2.4 we again assume that $x = 0$. The proof is similar to that of *Part (i)*, and in the following we use the same notation as in the proof of *Part (i)* for λ_i^l, x_i^l, λ_j^r, x_j^r and λ_k^*, x_k^*. Thus since

$$B_0 := \sum_{k=1}^{m^*} \lambda_k^* |x_k^*|^2 \leq \sum_{k=1}^{n+1} \lambda_k^r |x_k^r|^2 \leq r_c^2(0),$$

we have

$$C_\lambda^l(f_K^\infty)(0) \leq \sum_{k=1}^{m^*} \lambda_k^*(f(x_k^*) + \lambda|x_k^*|^2) = \lambda B_0 + f(0) + \sum_{k=1}^{m^*} \lambda_k^*(f(x_k^*) - f(0) - Df(0) \cdot x_k^*)$$

$$\leq \lambda B_0 + f(0) + \frac{L}{2}\sum_{k=1}^{m^*} \lambda_k^*|x_k^*|^2 \leq \lambda B_0 + f(0) + \frac{L}{2}r_c^2(0) \leq f(0) + \left(\frac{L}{2} + \lambda\right)r_c^2(0),$$

(5.5)

and also

$$C_\lambda^l(f_K^\infty)(0) = \sum_{i=1}^{k_l} \lambda_i^l(f(x_i^l) + \lambda|x_i^l|^2) = f(0) + \lambda\sum_{i=1}^{k_l} \lambda_i^l|x_i^l|^2 + \sum_{i=1}^{k_l} \lambda_i^l(f(x_i^l) - f(0) - Df(0) \cdot x_i^l)$$

$$\geq f(0) + \lambda\sum_{i=1}^{k_l} \lambda_i^l|x_i^l|^2 - \frac{L}{2}\sum_{i=1}^{k_l} \lambda_i^l|x_i^l|^2 = f(0) + \left(\lambda - \frac{L}{2}\right)\sum_{i=1}^{k_l} \lambda_i^l|x_i^l|^2.$$

(5.6)

By comparing (5.5) and (5.6), we then obtain

$$f(0) + \left(\lambda - \frac{L}{2}\right)\sum_{i=1}^{k_l} \lambda_i^l|x_i^l|^2 \leq f(0) + \left(\frac{L}{2} + \lambda\right)r_c^2(0),$$

so that

$$\sum_{i=1}^{k_l} \lambda_i^l|x_i^l|^2 \leq \frac{\lambda + \frac{L}{2}}{\lambda - \frac{L}{2}}r_c^2(0).$$

Thus from (5.6), we have

$$C_\lambda^l(f_K^\infty)(0) = \sum_{i=1}^{k_l} \lambda_i^l(f(x_i^l) + \lambda|x_i^l|^2)$$

$$= f(0) + \lambda \sum_{i=1}^{k_l} \lambda_i^l|x_i^l|^2 + \sum_{i=1}^{k_l} \lambda_i^l(f(x_i^l) - f(0) - Df(0) \cdot x_i^l) \qquad (5.7)$$

$$\geq f(0) + \lambda B_0 - \frac{L}{2}\sum_{i=1}^{k_l} \lambda_i^l|x_i^l|^2 \ \geq \ f(0) + \lambda B_0 - \frac{L}{2}\left(\frac{\lambda + \frac{L}{2}}{\lambda - \frac{L}{2}}\right)r_c^2(0).$$

By combining (5.5) and (5.7), we finally get

$$f(0) + \lambda B_0 - \frac{L}{2}\left(\frac{\lambda + \frac{L}{2}}{\lambda - \frac{L}{2}}\right)r_c^2(0) \leq C_\lambda^l(f_K^\infty)(0) \leq f(0) + \lambda B_0 + \frac{L}{2}r_c^2(0).$$

Similarly we can show that

$$f(0) - \lambda B_0 - \frac{L}{2}r_c^2(0) \leq C_\lambda^u(f_K^{-\infty})(0) \leq f(0) - \lambda B_0 + \frac{L}{2}\left(\frac{\lambda + \frac{L}{2}}{\lambda - \frac{L}{2}}\right)r_c^2(0).$$

The conclusion then follows. $\qquad\square$

Remark 5.1. *From the proof of Part (i) of Theorem 3.6 we observe that for a finite $M > 0$, if $C_\lambda^l(f_K^M)(0)$ can be calculated by using values of f in K ($0 \in K$), that is,*

$$C_\lambda^l(f_K^M)(0) = \sum_{i=1}^{n+1} \lambda_i\big(f(x_i) + \lambda|x_i|^2\big),$$

with $\lambda_i \geq 0$, $x_i \in K$, $\sum_{i=1}^{n+1} \lambda_i = 1$ and $\sum_{i=1}^{n+1} \lambda_i x_i = 0$, and if a similar result holds for the upper transform, then the arguments of the proof of Part (i) can go through without any changes. However, it is possible that one of the x_i's does not belong to K. In this case the situation is more complicated. In fact, we do not know whether Part (i) still holds for a finite $M > 0$. However, if we extend f_K outside a large ball as zero, we can still derive error bounds (see Corollary 3.9).

Proof of Theorem 3.7. *Part (i)*: We first give estimates for $C_\lambda^l(f_K^M)(x)$. Without loss of generality we assume that $x = 0$. Since $y \mapsto f_K^M(y) + \lambda|y|^2$ is lower semicontinuous and coercive, there are $x_i \in \mathbb{R}^n$, $\lambda_i > 0$ for $i = 1, 2, \ldots, m \leq n + 1$ such that, $\sum_{i=1}^m \lambda_i = 1$, $\sum_{i=1}^m \lambda_i x_i = 0$ and

$$\mathsf{co}[f_K^M + \lambda|\cdot|^2](0) = \sum_{i=1}^m \lambda_i\big(f_K^M(x_i) + \lambda|x_i|^2\big).$$

This implies that there is an affine function $\ell(y)$ such that $\ell(y) \leq f_K^M(y) + \lambda|y|^2$ for $y \in \mathbb{R}^n$ and $\ell(x_i) = f_K^M(x_i) + \lambda|x_i|^2$.

We first show that $x_i \in K$ for $i = 1, 2, \ldots, k$. If this is not the case, there is some $1 \leq i_0 \leq k$ such that $x_{i_0} \in \Omega$. Since $\ell(x_{i_0}) = f_K^M(x_{i_0}) + \lambda|x_{i_0}|^2$, $\ell(y)$ is an affine support function of $M + \lambda|y|^2$ at x_{i_0}, and hence is the unique tangent plane of the function $M + \lambda|y|^2$. Thus $\ell(y) = M + \lambda|x_{i_0}|^2 + 2\lambda x_{i_0} \cdot y$.

22

If $x_{i_0} = 0$, $\ell(0) = M$, which contradicts the assumption that $M > 2A_0 + \lambda d_\Omega^2$. If $x_{i_0} \neq 0$ and $x_{i_0} \in \Omega$, then since Ω is a bounded domain, there are two points x'_{i_0}, $x''_{i_0} \in \partial\Omega$ and some $0 < \alpha < 1$, such that $x_{i_0} = \alpha x'_{i_0} + (1 - \alpha)x''_{i_0}$. We also have

$$\alpha\big(f_K^M(x'_{i_0}) + \lambda|x'_{i_0}|^2\big) + (1 - \alpha)\big(f_K^M(x''_{i_0}) + \lambda|x''_{i_0}|^2\big) = \alpha\big(f(x'_{i_0}) + \lambda|x'_{i_0}|^2\big) + (1 - \alpha)\big(f(x''_{i_0}) + \lambda|x''_{i_0}|^2\big)$$
$$\leq A_0 + \lambda d_\Omega^2$$
$$< M \leq M + \lambda|x_{i_0}|^2 = f_K^M(x_{i_0}) + \lambda|x_{i_0}|^2.$$

Here we have used the fact that x'_{i_0}, $x''_{i_0} \in \partial\Omega$ and $0 \in \Omega$, so that $|x'_{i_0} - 0| \leq d_\Omega$ and $|x''_{i_0} - 0| \leq d_\Omega$. Thus

$$1 = \sum_{i=1}^{m} \lambda_i = \left(\sum_{i=1, i\neq i_0}^{m} \lambda_i\right) + \alpha\lambda_{i_0} + (1 - \alpha)\lambda_{i_0},$$

$$0 = \sum_{i=1}^{m} \lambda_i x_i = \left(\sum_{i=1, i\neq i_0}^{m} \lambda_i x_i\right) + \lambda_{i_0}\alpha x'_{i_0} + \lambda_{i_0}(1 - \alpha)x''_{i_0},$$

and

$$\mathsf{co}[f_K^M + \lambda|\cdot|^2](0) = \sum_{i=1}^{m} \lambda_i[f_K^M(x_i) + \lambda|x_i|^2]$$
$$> \left(\sum_{i=1, i\neq i_0}^{m} \lambda_i[f_K^M(x_i) + \lambda|x_i|^2]\right) + \alpha\lambda_{i_0}\big(f_K^M(x'_{i_0}) + \lambda|x'_{i_0}|^2\big) + (1 - \alpha)\lambda_{i_0}\big(f_K^M(x''_{i_0}) + \lambda|x''_{i_0}|^2\big).$$

But this contradicts the definition of the convex envelope. So $x_i \in K$ for all $i = 1, 2, \ldots, k$. The rest of the proof of *Part (i)* then follows from a similar argument to that for *Part (i)* of Theorem 3.6.

For *Part (ii)* and *Part (iii)*, we can use similar arguments to the proof of *Part (i)* to show that all x_i's are in K, so that the conclusions then follow from *Part (ii)* and *Part (iii)* of Theorem 3.6, respectively. $\qquad\square$

Proof of Corollary 3.9: The proof is very similar to that of Theorem 3.7 and is left to interested readers. $\qquad\square$

Proof of Lemma 4.3: This lemma is a direct consequence of the definitions (1.2) of f_K^M, f_K^{-M} and the definition of the upper and lower compensated convex transforms (1.1). $\qquad\square$

Proof of Lemma 4.6: Fix $x \in \mathbb{R}^n$. For every $\delta > \mathrm{dist}_\mathcal{H}(K, E)$, by Remark 4.5, there is some $x^E \in E$, such that $d_{\lambda, f}(x, E) = |x^E - x| - \sqrt{f(x^E)/\lambda}$. For $x^E \in E$, there is some $x^K \in K$ such that $|x^K - x^E| < \delta$. Thus

$$d_{\lambda, f}(x, K) - d_{\lambda, f}(x, E) \leq |x^K - x| - \sqrt{f(x^K)/\lambda} - |x^E - x| + \sqrt{f(x^E)/\lambda}$$

$$\leq |x^K - x^E| + \frac{1}{\sqrt{\lambda}}\omega(|x^K - x^E|)$$

$$\leq \delta + \frac{\omega(\delta)}{\sqrt{\lambda}}$$

for all $\delta > \text{dist}_{\mathcal{H}}(K, E)$. Hence,

$$d_{\lambda,f}(x,\,K) - d_{\lambda,f}(x,\,E) \le \text{dist}_{\mathcal{H}}(K, E) + \frac{\omega(\text{dist}_{\mathcal{H}}(K, E))}{\sqrt{\lambda}}\,.$$

Similarly, we can show that

$$d_{\lambda,f}(x,\,E) - d_{\lambda,f}(x,\,K) \le \text{dist}_{\mathcal{H}}(K, E) + \frac{\omega(\text{dist}_{\mathcal{H}}(K, E))}{\sqrt{\lambda}}\,,$$

and conclusion then follows. $\qquad\square$

Proof of Lemma 4.7: We have

$$|D^2_{\lambda,f}(x,\,K) - D^2_{\lambda,f}(x,\,E)| \le (|D_{\lambda,f}(x,\,K)| + |D_{\lambda,f}(x,\,E)|)|D_{\lambda,f}(x,\,K) - D_{\lambda,f}(x,\,E)|\,.$$

By definition of $D_{\lambda,f}(x,\,K)$, we then have if $\min_{y\in K}(|y-x| - \sqrt{f(y)/\lambda}) > 0$,

$$|D_{\lambda,f}(x,\,K)| = 0\,,$$

and if $\min_{y\in K}(|y-x| - \sqrt{f(y)/\lambda}) = |x^K - x| - \sqrt{f(x^K)/\lambda} < 0$ for some $x^K \in K$, then

$$|D_{\lambda,f}(x,\,K)| = \sqrt{\lambda}\big||x^K - x| - \sqrt{f(x^K)/\lambda}\big| = \sqrt{\lambda}(\sqrt{f(x^K)/\lambda} - |x^K - x|) \le \sqrt{f(x^K)} \le \sqrt{M}\,.$$

Similarly, we have

$$|D_{\lambda,f}(x,\,E)| \le \sqrt{M}\,.$$

Next, by the formula $\min\{0,\,a\} = (a - |a|)/2$ for $a \in \mathbb{R}$, we have

$$|D_{\lambda,f}(x,\,K) - D_{\lambda,f}(x,\,E)| = \frac{\sqrt{\lambda}}{2}\Big|d_{\lambda,f}(x,K) - |d_{\lambda,f}(x,K)| - (d_{\lambda,f}(x,E) - |d_{\lambda,f}(x,E)|)\Big|$$

$$\le \sqrt{\lambda}\Big|d_{\lambda,f}(x,K) - d_{\lambda,f}(x,E)\Big|$$

$$\le \sqrt{\lambda}\Big(\text{dist}_{\mathcal{H}}(K, E) + \frac{\omega(\text{dist}_{\mathcal{H}}(K, E))}{\sqrt{\lambda}}\Big)\,.$$

Thus we obtain

$$|D^2_{\lambda,f}(x,\,K) - D^2_{\lambda,f}(x,\,E)| \le 2\sqrt{\lambda M}\,\text{dist}_{\mathcal{H}}(K, E) + 2\sqrt{M}\omega(\text{dist}_{\mathcal{H}}(K, E))\,,$$

which completes the proof. $\qquad\square$

Proof of Lemma 4.8: The proof of this lemma is an easy exercise and is omitted here. $\qquad\square$

Proof of Lemma 4.9: We first show that

$$f(x)\chi_K(x) \le D^2_{\lambda,f}(x,\,K) \tag{5.8}$$

for all $x \in \mathbb{R}^n$, so that by (2.7),

$$C^u_\lambda(f\chi_K)(x) \le C^u_\lambda(D^2_{\lambda,f}(\cdot,\,K))(x) \tag{5.9}$$

for all $x \in \mathbb{R}^n$. If $x \notin K$, clearly, $\chi_K(x)f(x) = 0 \le D_{\lambda,f}^2(x, K)$. If $x \in K$, since

$$d_{\lambda,f}(x, K) = \min_{y \in K}(|y - x| - \sqrt{f(y)/\lambda}) \le -\sqrt{f(x)/\lambda} < 0,$$

we have

$$D_{\lambda,f}(x, K) = -\sqrt{\lambda}\min\{0, d_{\lambda,f}(x, K)\} = -\sqrt{\lambda}\, d_{\lambda,f}(x, K) \ge \sqrt{\lambda}\sqrt{f(x)/\lambda} = \sqrt{f(x)},$$

and thus $D_{\lambda,f}^2(x, K) \ge f(x)$. Therefore (5.8) holds for all $x \in \mathbb{R}^n$, from which (5.9) follows.

Next we show that the opposite inequality, $C_\lambda^u(D_{\lambda,f}^2(\cdot, K))(x) \le C_\lambda^u(f\chi_K)(x)$, also holds. If

$$d_{\lambda,f}(x, K) = \min_{y \in K}(|y - x| - \sqrt{f(y)/\lambda}) > 0, \tag{5.10}$$

then by definition, $D_{\lambda,f}(x, K) = 0$, and hence $D_{\lambda,f}^2(x, K) = 0$. We show in this case that

$$C_\lambda^u(D_{\lambda,f}^2(\cdot, K))(x) = 0 \le C_\lambda^u(f\chi_K)(x).$$

We will consider the function $z \mapsto \lambda|z - x|^2 - D_{\lambda,f}^2(z, K)$ for $z \in \mathbb{R}^n$ and show that the value of the convex envelope of this function at x is zero. Consider the affine function $\ell(z) = 0$ for all $z \in \mathbb{R}^n$ and show that

$$0 = \ell(x) = \left(\lambda|z - x|^2 - D_{\lambda,f}^2(z, K)\right)|_{z=x}, \tag{5.11}$$

and

$$0 = \ell(z) \le \lambda|z - x|^2 - D_{\lambda,f}^2(z, K), \quad z \in \mathbb{R}^n. \tag{5.12}$$

Equality (5.11) is obvious as $[\lambda|z - x|^2 - D_{\lambda,f}^2(z, K)]|_{z=x} = -D_{\lambda,f}^2(x, K) = 0$. Now we prove (5.12), that is, $0 \le \lambda|z - x|^2 - D_{\lambda,f}^2(z, K)$, which is equivalent to

$$D_{\lambda,f}^2(z, K) \le \lambda|z - x|^2, \quad z \in \mathbb{R}^n. \tag{5.13}$$

If $d_{\lambda,f}(z, K) \ge 0$, then $D_{\lambda,f}^2(z, K) = 0$, hence (5.13) holds. If $d_{\lambda,f}(z, K) < 0$, then $D_{\lambda,f}^2(z, K) = \lambda d_{\lambda,f}^2(z, K)$. We need to show that $\lambda(\min_{y \in K}(|y - z| - \sqrt{f(y)/\lambda})^2 \le \lambda|z - x|^2$, which is equivalent to $-\min_{y \in K}(|y - z| - \sqrt{f(y)/\lambda}) \le |z - x|$, which is in turn equivalent to

$$|z - x| + \min_{y \in K}(|y - z| - \sqrt{f(y)/\lambda}) \ge 0. \tag{5.14}$$

By the triangle inequality and (5.10), we have

$$|z - x| + \min_{y \in K}(|y - z| - \sqrt{f(y)/\lambda}) = \min_{y \in K}(|z - x| + |y - z| - \sqrt{f(y)/\lambda})$$

$$\ge \min_{y \in K}(|y - x| - \sqrt{f(y)/\lambda}) = d_{\lambda,f}(x, K) > 0.$$

Thus (5.12) holds. Therefore

$$0 = \mathsf{co}[\lambda|\cdot - x|^2 - D_{\lambda,f}^2(\cdot, K)](x) = -C_\lambda^u(D_{\lambda,f}^2(\cdot, K))(x),$$

25

which implies

$$D^2_{\lambda,f}(x,K) \le C^u_\lambda(D^2_{\lambda,f}(\cdot,K))(x) = 0 \le C^u_\lambda(f\chi_K)(x). \tag{5.15}$$

Finally, we consider the case

$$d_{\lambda,f}(x,K) = \min_{y\in K}(|y-x| - \sqrt{f(y)/\lambda}) = |x^K - x| - \sqrt{f(x^K)/\lambda} < 0, \tag{5.16}$$

where $x^K \in K$ is the minimum point. Now we consider the function $f(y)\chi_{\{x^K\}}(y)$ for $y \in \mathbb{R}^n$. By Lemma 4.8, we have

$$C^u_\lambda(f\chi_{\{x^K\}})(y) = \begin{cases} \lambda\left(|y-x^K| - \sqrt{f(x^K)/\lambda}\right)^2, & |y-x^K| \le \sqrt{f(x^K)/\lambda}, \\ 0, & |y-x^K| \ge \sqrt{f(x^K)/\lambda}. \end{cases}$$

In particular, since $x^K \in K$, we have $f(y)\chi_{\{x^K\}}(y) \le f(y)\chi_K(y)$ for all $y \in \mathbb{R}^n$, so that by (2.7),

$$C^u_\lambda(f\chi_{\{x^K\}})(y) \le C^u_\lambda(f\chi_K)(y) \quad \text{for all } y \in \mathbb{R}^n.$$

By our assumption (5.16), we also have $|x^K - x| < \sqrt{f(x^K)/\lambda}$, thus

$$C^u_\lambda(f\chi_{\{x^K\}})(x) = \lambda\left(|x-x^K| - \sqrt{f(x^K)/\lambda}\right)^2 = D^2_{\lambda,f}(x,K)$$

as $d_{\lambda,f}(x,K) < 0$. Thus, in this case, $D^2_{\lambda,f}(x,K) = C^u_\lambda(f\chi_{\{x^K\}})(x) \le C^u_\lambda(f\chi_K)(x)$. By combining this case and (5.15), we have, for all $x \in \mathbb{R}^n$, that $D^2_{\lambda,f}(x,K) \le C^u_\lambda(f\chi_K)(x)$, so that

$$C^u_\lambda(D^2_{\lambda,f}(\cdot,K))(x) \le C^u_\lambda(C^u_\lambda(f\chi_K))(x) = C^u_\lambda(f\chi_K)(x).$$

Since the opposite inequality (5.9) also holds, we have

$$C^u_\lambda(D^2_{\lambda,f}(\cdot,K))(x) = C^u_\lambda(f\chi_K)(x)$$

for all $x \in \mathbb{R}^n$, which completes the proof. $\qquad\square$

Proof of Theorem 4.10: By Lemma 4.9, we only need to prove

$$|C^u_\lambda(D^2_{\lambda,f}(\cdot,K))(x) - C^u_\lambda(D^2_{\lambda,f}(\cdot,E))(x)| \le 2\sqrt{\lambda M}\text{dist}_\mathcal{H}(K,E) + 2\sqrt{M}\omega(\text{dist}_\mathcal{H}(K,E)). \tag{5.17}$$

By Lemma 4.7 we have, for all $x \in \mathbb{R}^n$ that

$$|D^2_{\lambda,f}(x,K) - D^2_{\lambda,f}(x,E)| \le 2\sqrt{\lambda M}\text{dist}_\mathcal{H}(K,E) + 2\sqrt{M}\omega(\text{dist}_\mathcal{H}(K,E)).$$

Thus

$$D^2_{\lambda,f}(x,E) - 2\sqrt{\lambda M}\text{dist}_\mathcal{H}(K,E) - 2\sqrt{M}\omega(\text{dist}_\mathcal{H}(K,E)) \le D^2_{\lambda,f}(x,K)$$

$$\le D^2_{\lambda,f}(x,E) + 2\sqrt{\lambda M}\text{dist}_\mathcal{H}(K,E) + 2\sqrt{M}\omega(\text{dist}_\mathcal{H}(K,E))$$

for all $x \in \mathbb{R}^n$. By the ordering and the affine covariance properties of compensated convex trans-

forms, we have

$$C_\lambda^u(D_{\lambda,f}^2(\cdot, E))(x) - 2\sqrt{\lambda M}\text{dist}_\mathcal{H}(K, E) - 2\sqrt{M}\omega(\text{dist}_\mathcal{H}(K, E)) \le C_\lambda^u(D_{\lambda,f}^2(\cdot, K))(x)$$

$$\le C_\lambda^u(D_{\lambda,f}^2(\cdot, E))(x) + 2\sqrt{\lambda M}\text{dist}_\mathcal{H}(K, E) + 2\sqrt{M}\omega(\text{dist}_\mathcal{H}(K, E)).$$

Hence (5.17) follows. $\qquad\square$

Proof of Corollary 4.11: This is a direct consequence of Theorem 4.10 where $\omega(t) = Lt/(2\sqrt{\alpha})$ for $t > 0$, with $L \ge 0$ the Lipschitz constant of f, since $|\sqrt{f(x)} - \sqrt{f(y)}| \le L|x - y|/(2\sqrt{\alpha})$. $\qquad\square$

Proof of Theorem 4.12: The cases of $L_\lambda^M(f_K)$, $U_\lambda^M(f_K)$ and $A_\lambda^M(f_K)$ are direct consequences Definition 1.1, Lemma 4.3, Theorem 4.10 and Corollary 4.11. Note that for the Hausdorff-Lipschitz continuity properties, the assumption that $M > A_0$ ensures that the uniform positivity assumption in Corollary 4.11 is satisfied by both $M - f$ and $M + f$. For the mixed average approximation $(SA)_{\lambda,\tau}^M(f_K) = \frac{1}{2}(C_\tau^u(C_\lambda^l(f_K^M)) + C_\tau^l(C_\lambda^u(f_K^{-M})))$, we use (2.7). Since $|C_\lambda^l(f_G^M)(x) - C_\lambda^l(f_K^M)(x)| < \epsilon$ for all $x \in \mathbb{R}^n$ with $\epsilon = 2\sqrt{\lambda M}\text{dist}_\mathcal{H}(G, K) + 2\sqrt{M}\omega(\text{dist}_\mathcal{H}(G, K))$, we have

$$C_\lambda^l(f_K^M)(x) - \epsilon < C_\lambda^l(f_G^M)(x) < C_\lambda^l(f_K^M)(x) + \epsilon,$$

and hence

$$|C_\tau^u(C_\lambda^l(f_G^M))(x) - C_\tau^u(C_\lambda^l(f_K^M))(x)| < \epsilon,$$

since $C_\tau^u(C_\lambda^l(f_K^M) \pm \epsilon) = C_\tau^u(C_\lambda^l(f_K^M)) \pm \epsilon$. Similarly,

$$|C_\tau^l(C_\lambda^u(f_G^{-M}))(x) - C_\tau^l(C_\lambda^u(f_K^{-M}))(x)| < \epsilon$$

since $|C_\lambda^u(f_G^{-M})(x) - C_\lambda^u(f_K^{-M})(x)| < \epsilon$. The proof for $(SA)_{\lambda,\tau}^M(f_K)$ then follows. The proof for the Lipschitz case is similar, using arguments from Lemma 4.3 and Corollary 4.11. $\qquad\square$

Proof of Theorem 4.13: Part (i) and the error estimate (4.7) follow from [48, Theorem 3.13]. The fact that mixed transforms are $C^{1,1}$ is a consequence of [47, Theorem 2.1(iv), Theorem 4.1(ii)]. Note that this latter regularity result also follows from the fact that if g is both 2λ-semiconvex and 2λ-semiconcave, then g is a $C^{1,1}$ function [15, Corollary 3.3.8]. $\qquad\square$

References

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

[2] F. Andreu, C. Ballester, V. Caselles, J. M. Mazón, The Dirichlet problem for the total variation flow, J. Functional Anal. 180 (2001) 347-403.

[3] F. Andreu, C. Ballester, V. Caselles, J. M. Mazón, Minimizing total variation flow, Diff. Integr. Equations. 14 (2001) 321-360.

[4] H. Attouch, D. Aze, Approximations and regularizations of arbitrary functions in Hilbert spaces by the Lasry-Lions methods, Anal. Non-Lin. H. Poincaré Inst. 10 (1993) 289-312.

[5] H. H. Bauschke, Y. Lucet, S. Wang, Primal-Dual symmetric intrinsic methods for finding antiderivatives of cyclically monotone operators, SIAM Journal of Control and Optimization 46 (2007) 2031-2051.

[6] H. H. Bauschke, Y. Lucet, M. Trienis, How to transform one convex function continuously into another, SIAM Review 50 (2008) 115-132.

[7] H. H. Bauschke, R. Goebel, Y. Lucet, S. Wang, The proximal average: Basic theory, SIAM Jour. of Optim. 19 (2008) 766-785.

[8] G. Bellettini, V. Caselles, M. Novaga, The total variation flow in $\mathbb{R}^N$, J. Diff. Equations 184 (2002) 475-525.

[9] J. Benoist, J.-B. Hiriart-Urruty, What is the subdifferential of the closed convex hull of a function? SIAM J. Math. Anal. 27 (1996) 1661-1679

[10] M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester, Image inpainting, In: SIG-GRPAH 2000.

[11] R. Van Den Boomgaard, H. J. A. M. Heijmans, Morphological scale-space operators: An algebraic framework, in J. Goutsias, L. Vincent, D. S. Bloomberg (Eds), Mathematical Morphology and its Applications to Image and Signal Processing, Volume 18, 2000, 282-290

[12] R. Van den Boomgaard, A. W. M. Smelders, The morphological structure of images, in Proceedings 11th IAPR International Conference on Pattern Recognition. The Hague, The Netherlands: IEEE Computer Society Press, Los Alamitos, CA, (1992), 268-271.

[13] M. Bounkhel, Regularity Concepts in Nonsmooth Analysis: Theory and Applications, Springer Optimization and Its Applications 59, Springer-Verlag Berlin, 2012.

[14] M. D. Buhmann, Radial Basis Functions, Cambridge Univ. Press, 2004.

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

[16] R. H. Chan, C.-W. Ho, M. Nikolova, Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization, IEEE Trans. Image Processing 14 (2005) 1479-1485.

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

[18] E. E. W. Cheney, W. W. A. Light, A Course in Approximation Theory, American Mathematical Society, 2000.

[19] F. H. Clarke, R. J. Stern, P. R. Wolenski, Proximal smoothness and the lower-C^2 property, J. Convex Anal. 2 (1995) 117-144.

[20] R. A. DeVore and G. G. Lorentz, Constructive Approximation, Springer, 1993.

[21] H. Edelsbrunner, Algorithms in Combinatorial Geometry, Springer, 1987.

[22] M. Giaquinta, G. Modica, J. Soucek, Functionals with linear growth in the calculus of variations I, Comment. Math. Univ. Carolinae 20 (1979) 143-156.

[23] A. Griewank, P.J. Rabier, On the smoothness of convex envelopes, Trans. Amer. Math. Soc. 322 (1990) 691-709.

[24] F. Guichard, P. Maragos, J.-M. Morel, Partial differential equations for morphological operators, In: M. Bilodeau, F. Meyer, M. Schmitt (Eds), Space, Structure and Randomness, Lecture Notes in Statistics 183 (2005) 369-390, Springer Verlag

[25] W.L. Hare, A proximal average for nonconvex functions: A proximal stability perspective, SIAM J. Optim. 20 (2009) 650-666

[26] J.-B. Hiriart-Urruty, The deconvolution operation in convex analysis: An introduction, Cybernetics and Systems Analysis 30 (1994) 555-560.

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

[28] P. T. Jackway, Morphological scale-space, in Proceedings 11th IAPR International Conference on Pattern Recognition. The Hague, The Netherlands: IEEE Computer Society Press, Los Alamitos, CA, (1992) 252-255.

[29] P. T. Jackway, M. Deriche, Scale-space properties of the multiscale morphological dilation-erosion, IEEE Transactions on Pattern Analysis and Machine intelligence 18 (1996) 38-51

[30] K. Jetter, M. D. Buhmann, W. Haussman, R. Schaback (eds) Topics in Multivariate Approximation and Interpolation, Elsevier, 2006.

[31] J. M. Lasry and P. L. Lions, A remark on regularization in Hilbert spaces, Israel Math. Journal 55 (1996) 257-266.

[32] Y. Lucet, What shape is your conjugate? A survey of computational convex analysis and its applications, SIAM J. Optim. 20 (2009) 216-250

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

[34] J.-J. Moreau, Fonctionnelles Convexes. Séminaire "Sur les équations aux dérivées partielles". Lecture Notes, Collége de France, 1966.

[35] A. Okabe, B. Boots, K. Sugihara, S. N. Chiu, Spatial Tessellations, 2nd edition, John Wdey & Sons, 2000.

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

[37] R. T. Rockafellar, Convex Analysis, Princeton University Press, 1970.

[38] R. T. Rockafellar, Favorable classes of Lipschitz-continuous functions in subgradient optimization. Progress in nondifferentiable optimization, 125-143, IIASA Collaborative Proc. Ser. CP-82, 8, Internat. Inst. Appl. Systems Anal., Laxenburg, 1982

[39] R. T. Rockafellar, R. Wets, Variational Analysis, Springer Verlag Berlin, 1998.

[40] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1992) 259-268.

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

[42] J. Serra, L. Vincent, An overview of morphological filtering, Circuits Systems Signals Process 11 (1992) 47-108

[43] P. Soille, Morphological Image Analysis: Principles and Applications, Springer, Berlin, 2nd Ed., 2004

[44] A. F. Timan, Theory of Approximation of Functions of a Real Variable, Dover Publishing, 1994.

[45] H. Wendland, Scattered Data Approximation, Cambridge University Press, 2005.

[46] Y. Yu, Better approximation and faster algorithms using the proximal average, in Advances in Neural Information Processing Systems 26 (NIPS13), Proceedings of the Conference Neural Information Processing Systems 2013, Eds. C.J.C. Burgers et al, 2013

[47] K. Zhang, Compensated convexity and its applications, Anal. Nonlin. H. Poincare Inst. 25 (2008) 743- 771.

[48] K. Zhang, A. Orlando, E. Crooks, Compensated convexity and Hausdorff stable geometric singularity extractions, Math. Models Methods Appl. Sci. 25 (2015) 747-801

[49] K. Zhang, A. Orlando, E. Crooks, Compensated convexity and Hausdorff stable extraction of intersections for smooth manifolds, Math. Models Methods Appl. Sci. 25 (2015) 839-873

[50] K. Zhang, E. Crooks, A. Orlando, Compensated convexity, multiscale medial axis maps and sharp regularity of the squared distance function, SIAM J. Math. Anal. 47 (2015) 4289-4331

[51] K. Zhang, E. Crooks, A. Orlando, Compensated convexity methods for approximations and interpolations of sample functions in Euclidean spaces: Applications to sparse data, contour lines and inpainting. In preparation.

[52] K. Zhang, E. Crooks, A. Orlando, Compensated convexity and Hausdorff stable extraction of geometric intersections and interior corners. In preparation.