Available online at www.sciencedirect.com

ScienceDirect

Ann. I. H. Poincaré – AN 25 (2008) 743–771

ANNALES DE L'INSTITUT HENRI POINCARÉ
ANALYSE NON LINÉAIRE

www.elsevier.com/locate/anihpc

Compensated convexity and its applications

Kewei Zhang

Department of Mathematics, University of Wales Swansea, Singleton Park, Swansea, SA2 8PP, UK

Received 12 March 2006; accepted 2 August 2007

Available online 2 October 2007

Dedicated to Professor Kung-Ching Chang on the occasion of his 70th birthday

Abstract

We introduce the notions of lower and upper quadratic compensated convex transforms $C^l_{2,\lambda}(f)$ and $C^u_{2,\lambda}(f)$ respectively and the mixed transforms by composition of these transforms for a given function $f:\mathbb{R}^n \mapsto \mathbb{R}$ and for possibly large $\lambda > 0$. We study general properties of such transforms, including the so-called 'tight' approximation of $C^l_{2,\lambda}(f)$ to f as $\lambda \to +\infty$ and compare our transforms with the well-known Moreau–Yosida regularization (Moreau envelope) and the Lasry–Lions regularization. We also study analytic and geometric properties for both the quadratic lower transform $C^l_{2,\lambda}(\mathrm{dist}^2(x,K))$ of the squared-distance function to a compact set K and the quadratic upper transform $C^u_{2,\lambda}(f)$ for any convex function f of at most quadratic growth. We show that both $C^l_{2,\lambda}(\mathrm{dist}^2(x,K))$ and $C^u_{2,\lambda}(f)$ are $C^{1,1}$ approximations of the original functions for large $\lambda > 0$ and $C^u_{2,\lambda}(f)$ remains convex. Explicitly calculated examples of quadratic transforms are given, including the lower transform of squared distance function to a finite set and upper transform for some non-smooth convex functions in mathematical programming.

Keywords: Convex functions; Convex envelope; Weak lower semi-continuous functions; Moreau envelope; Squared-distance functions; Tight approximation; $C^{1,1}$-smoothing; Quasi-convexity; Maximum function

1. Introduction

In this paper we introduce the notions of *compensated convex transforms*. In particular, we present a systematic study of the properties for the quadratic compensated convex transforms and apply them to various problems in applied analysis. The lower quadratic compensated convex transforms have been used in several places in the context of quasi-convex functions, the quasi-convex hull and gradient Young measures in the calculus of variations [52–54,56].

Definition 1.1. Suppose that $f:\mathbb{R}^n \mapsto \mathbb{R}\cup\{+\infty\}$ satisfies

$$f(x) \geqslant -C_f|x|^2 - C_1, \quad x \in \mathbb{R}^n, \tag{1.1}$$

for some constants $C_f > 0$, $C_1 > 0$, then the *quadratic lower compensated convex transform* (for short: *lower transform*) for f is defined by

$$C^l_{2,\lambda}\big(f(x)\big) = C\big[f(x) + \lambda|x|^2\big] - \lambda|x|^2, \quad x \in \mathbb{R}^n,\ \text{for } \lambda > C_f, \tag{1.2}$$

E-mail address: k.zhang@swansea.ac.uk.

0294-1449/$ – see front matter
doi:10.1016/j.anihpc.2007.08.001

where $C[f(x)+\lambda|x|^2]$ is the value of the convex envelope of the function $y\mapsto f(y)+\lambda|y|^2$ at $x\in\mathbb{R}^n$.

If $f:\mathbb{R}^n\mapsto\mathbb{R}$ satisfies the condition

$$f(x)\leqslant C_f|x|^2+C_1,\quad x\in\mathbb{R}^n, \tag{1.3}$$

we define its *quadratic upper compensated convex transform* (for short: *upper transform*) by

$$C^u_{2,\lambda}\big(f(x)\big):=-C^l_{2,\lambda}\big(-f(x)\big),\quad x\in\mathbb{R}^n,\ \text{for } \lambda>C_f. \tag{1.4}$$

If $f:\mathbb{R}^n\mapsto\mathbb{R}$ satisfies the growth condition

$$\big|f(x)\big|\leqslant C_f|x|^2+C_1,\quad x\in\mathbb{R}^n, \tag{1.5}$$

we define the two *quadratic mixed compensated convex transforms* (*mixed transforms* for short) by

$$C^{u,l}_{2,\lambda,\tau}\big(f(x)\big):=C^u_{2,\lambda}\big[C^l_{2,\tau}\big(f(x)\big)\big],\quad C^{l,u}_{2,\lambda,\tau}\big(f(x)\big):=C^l_{2,\lambda}\big[C^u_{2,\tau}\big(f(x)\big)\big],\quad x\in\mathbb{R}^n,\ \lambda,\tau>C_f. \tag{1.6}$$

Obviously, for general $p>1$, we may define the p-compensated convex transforms

$$C^l_{p,\lambda}\big(f(x)\big)=C\big[f(x)+\lambda|x|^p\big]-\lambda|x|^p,\quad\text{and}\quad C^u_{p,\lambda}(f)=\lambda|x|^p-C\big[-f(x)+\lambda|x|^p\big]$$

respectively by using the convex function $\lambda|x|^p$. Later in this paper, we also use more general notions of quadratic compensated convex transforms to serve our purpose for deriving explicit approximate functions. For example (see Section 5), we take the simple 'anisotropic' convex quadratic function $g_\lambda(x,y)=\lambda|x|^2+(1+\lambda)|y|^2$, $(x,y)\in\mathbb{R}^n\times\mathbb{R}^m$, $\lambda>0$, and define the g_λ-lower transforms for $f(x,y)$ by $C^l_{g_\lambda}(f(x,y))=C[f(x,y)+g_\lambda(x,y)]-g_\lambda(x,y)$.

Throughout this paper we only consider functions which are everywhere finite with $p=2$. However, many results in Section 2 can be extended to $f:\mathbb{R}^n\mapsto\mathbb{R}\cup\{+\infty\}$ or $f:\mathbb{R}^n\mapsto\mathbb{R}\cup\{-\infty\}$, and to the case of general $p>1$. Clearly, the case $p\neq 2$ is more technical and requires estimates of the function $x\mapsto|x|^p$ [48]. We will present such generalizations elsewhere.

From the definition of the lower quadratic transform, it is easy to see that for a convex function f, such a transform has no effect on f, that is $C^l_{2,\lambda}(f)=f$. Also note that even if f is of class C^∞, the convex envelope $C(f)$ cannot in general be any better than $C^{1,1}$ [23,11,20]. However we can show (Theorem 4.1) that for any convex function of at most quadratic growth, the quadratic upper transform $C^u_{2,\lambda}(f)$ is a convex $C^{1,1}$ approximation of f when $\lambda\to+\infty$. Also we can calculate $C^u_{2,\lambda}(f)$ explicitly for some non-smooth convex functions f widely used in convex programming. For example, the maximum function in mathematical programming [4,34,6,13,16,37,41] is defined by

$$f(x)=\max_{1\leqslant i\leqslant n}x_i,\quad x=(x_1,\dots,x_n)\in\mathbb{R}^n. \tag{1.7}$$

By applying the quadratic upper transform to f, we show that (see Theorem 5.1 below)

$$C^u_{2,\lambda}\big(f(x)\big)=\lambda|x|^2-\operatorname{dist}^2\left(x,C\left(\frac{K_n}{2\lambda}\right)\right)+\frac{1}{4\lambda}, \tag{1.8}$$

where $K_n=\{e_i,\ i=1,2,\dots,n\}$ consists of the standard Euclidean basis of $\mathbb{R}^n$ while $C(K^n/2\lambda)$ is the convex hull of $K^n/2\lambda:=\{e_i/(2\lambda),\ i=1,2,\dots,n\}$. In Section 5, among other examples, we will show for the maximum function f that the geometrically simple function $C^u_{2,\lambda}(f)$ given by (1.8) is a convex $C^{1,1}$ approximation of f as $\lambda\to+\infty$ satisfying the uniform error estimate $0\leqslant f(x)-C^u_{2,\lambda}(f(x))\leqslant 1/(2\lambda)$.

We are interested in the study of $C^l_{2,\lambda}(f)$ for both small and large $\lambda>0$ when it can be defined. If f is lower semi-continuous and maps bounded sets to bounded sets, we establish the following properties for the lower compensated convex transform $C^l_{2,\lambda}(f)$ in Section 2:

(i) the *continuity property* that $\lambda\mapsto C^l_{2,\lambda}(f)$ is continuous;
(ii) the *monotonicity preserving property* that if f is 'monotone' then $C^l_{2,\lambda}(f)$ has the same monotonicity property;
(iii) the *tight approximation theorem* in the sense that $\lim_{\lambda\to+\infty}C^l_{2,\lambda}(f)=f$, and at every point x where the original function f is of $C^{1,1}$ in a neighborhood of x, the value $f(x)$ can be reached by $C^l_{2,\lambda}(f(x))$ in finite 'time', that is $C^l_{2,\lambda}(f(x))=f(x)$ when λ is greater than certain constant depending on f and the size of the neighborhood;

(vi) the *locality property* that for any given ball $B(x,\delta)$, the value of $C^l_{2,\lambda}(f)(x)$ depends only on the value of $f(y)+\lambda|y|^2$ in $B(x,\delta)$ when λ is greater than a constant depending on f and δ.

The well-known Moreau–Yosida regularization (or Moreau envelope) [31,32,47,1] is defined for $f:\mathbb{R}^n\mapsto\mathbb{R}\cup\{+\infty\}$ satisfying (1.1) and for small $\lambda>0$ by

$$f_\lambda(x)=\inf_{y\in\mathbb{R}^n}\left\{f(y)+\frac{1}{2\lambda}|x-y|^2\right\},\quad 0<\lambda<\frac{1}{2C_f}.$$

For convex functions, a classical result due to Moreau [31] (also see [47,3]) says that f_λ is a $C^{1,1}$ function and as $\lambda\to 0_+$, $f_\lambda\to f$.

The Lasry–Lions regularizations [27,1] are based on the Moreau envelope and are defined originally [27] in the space $BUC(H)$ of real-valued, bounded uniformly continuous functions in a Hilbert space H by

$$(f_\lambda)^\tau(x)=\sup_{y\in\mathbb{R}^n}\inf_{u\in\mathbb{R}^n}\left(f(u)+\frac{1}{2\lambda}|u-y|^2-\frac{1}{2\tau}|y-x|^2\right),$$
$$(g^\tau)_\lambda(x)=\inf_{y\in\mathbb{R}^n}\sup_{u\in\mathbb{R}^n}\left(g(u)-\frac{1}{2\tau}|u-y|^2+\frac{1}{2\lambda}|y-x|^2\right)$$

for small $\lambda>0$ and $0<\tau<\lambda$. These notions have since been generalized to functions in normed linear spaces which are quadratic minorized and majorized in the sense of (1.1) and (1.3) respectively. It is known [1] that the Lasry–Lions regularizations are $C^{1,1}$ approximations of the original function in Hilbert spaces. For the sake of convenience, when we compare the two different methods, we make the following change of notation from now on for the Moreau envelope and Lasry–Lions regularization as

$$M_{2\lambda}\big(f(x)\big)=f_{2/\lambda}(x)=\inf_{y\in\mathbb{R}^n}\big(f(y)+\lambda|y-x|^2\big),$$
$$M^{2\mu}\big(f(x)\big)=f^{2/\mu}(x)=\inf_{y\in\mathbb{R}^n}\big(f(y)-\mu|y-x|^2\big)$$

for large $\lambda>0$, $\mu>0$ so that $(f_{2/\lambda})^{2/\mu}(x)=M^{2\mu}(M_{2\lambda}(f(x)))$.

We compare our quadratic transforms with Moreau envelopes by showing that in general, our quadratic transforms are tighter approximations to the original function than the Moreau envelopes. More precisely, we have (Theorem 2.5)

$$M_{2\lambda}(f)\leqslant C^l_{2,\lambda}(f)\leqslant f\leqslant C^u_{2,\mu}(f)\leqslant M^{2\mu}(f).\tag{1.9}$$

We also use several explicitly calculated simple one-dimensional examples to compare our quadratic transforms and the mixed transforms with Moreau-envelopes and Lasry–Lions regularizations. One of such example (Example 5.4) is the distance function $x\mapsto \operatorname{dist}(x,\{-1,1\})$ (see [1, Example 2.4]). This example indicates that the behavior of our mixed transforms are more predictable than that of the Lasry–Lions regularizations.

Before we briefly describe our applications of compensated convex transforms to the squared distance functions and convex functions in Sections 3 and 4 respectively, let us provide some motivations for the notions of compensated convex transforms.

(i) Given a function f bounded below, the convex envelope $C(f)$ can be considered, geometrically, as a poor approximation of f from below, as the gap between f and $C(f)$ can be very large. To fill this gap, we may, at least intuitively, try to compensate f by a 'well behaved' convex function such as $\lambda|x|^2$ to strengthen f by the re-enforced function $f(x)+\lambda|x|^2$. The convex envelope $C[f(x)+\lambda|x|^2]$ is obviously below $f(x)+\lambda|x|^2$. Then we partially remove the effect of $\lambda|x|^2$ from the convex envelope to obtain $C^l_{2,\lambda}(f)$ which lies between $C(f)$ and f: $C(f)\leqslant C^l_{2,\lambda}(f)\leqslant f$. Thus we are led naturally to the definition of $C^l_{2,\lambda}(f)$ which is obviously a better 'convex' approximation to f than $C(f)$.

(ii) Historically, a more direct motivation of (lower) compensated convex transforms (for small $\lambda>0$) is from the *translation method* [44,26,28,22,18,24,33,19] in the application of compensated compactness [43] to derive optimal bounds of effective moduli for composite materials and in the study of material micro-structure [9,10] and related quasi-convex functions and quasi-convex envelopes in the calculus of variations. Given a function

$F : M^{N\times n} \mapsto \mathbb{R}$, where $M^{N\times n}$ is the linear space of real $N \times n$ matrices, one can bound F from below by a quasi-convex 'translation bound' $C[F(x) - G(X)] + G(X)$, where G is a quasi-convex function in the sense of Morrey [30,5]. Although the definition of translation bound does not appear to be the same as lower compensated convex transforms, when the method is applied to bound the quasi-convex envelope for certain squared-distance functions [52–54,56], the method leads to a version of the quadratic lower transform of the squared-distance function for relatively small $\lambda > 0$.

(iii) Analytically, it has been observed recently that convex and quasi-convex envelopes have certain smoothing effects upon the original functions [8,20,23,11]. An example is that the quasi-convex envelope $Q(\mathrm{dist}^p(X,K))$ of the p-distance function to a compact set $K \subset M^{N\times n}$ is of class $C^{1,1}_{\mathrm{loc}}$ if $p > 2$ and of class $C^{1,p-1}$ if $1 < p \leqslant 2$. This result remains true if one replaces the quasi-convex envelope by the convex envelope $C(\mathrm{dist}^p(X,K))$. Thus by borrowing the analytic methods from [11] one is naturally led to predict that compensated convex transforms could be used as a smoothing tool which might be calculated explicitly in many applicable models.

In Section 3 we apply the quadratic lower transforms $C^l_{2,\lambda}(\mathrm{dist}^2(x,K))$ to the Euclidean squared-distance function $\mathrm{dist}^2(\cdot,K)$ for a compact subset K of $\mathbb{R}^n$. This is motivated from many applications in various areas of mathematics [29,9,10,7,51–54,56] and computing sciences [14,15,46,45,12,36,40]. Our main results there are concerned with the smoothness property of $C^l_{2,\lambda}(\mathrm{dist}^2(\cdot,K))$, the quasi-convex hull of compact sets in subspaces of $M^{N\times n}$ without rank-one matrices and the effect of lower transforms on the 'medial-axis' of a Euclidean domain.

Section 4 is devoted to the study of the smoothing effect of the upper transform $C^u_{2,\lambda}(f)$ on convex functions satisfying the quadratic growth condition (1.5), that is, $|f(x)| \leqslant C_f|x|^2 + C_1$. We show (Theorem 4.1), for convex functions satisfying (1.5) that $C^u_{2,\lambda}(f)$ is both convex and is of $C^{1,1}$ in $\mathbb{R}^n$ whenever $\lambda > C_f$. Based on this, we consider the mixed transforms $C^u_{2,\tau}[C^l_{2,\lambda}(f)]$ and $C^l_{2,\tau}[C^u_{2,\lambda}(f)]$. We show that one can approximate every given continuous function f satisfying (1.5) by a sequence of $C^{1,1}$ functions $C^u_{2,\tau_j}[C^l_{2,\lambda_j}(f)]$ uniformly on compact sets as $\tau_j \geqslant \lambda_j \to +\infty$.

In Section 5, we present two sets of calculated examples. The first consists of lower and upper transforms for one-dimensional functions and we compare them with the Moreau envelopes and Lasry–Lions regularizations for these functions. The second is on more specific functions defined in $\mathbb{R}^n$. We calculate the upper transform $C^u_{2,\lambda}(f)$ for non-smooth convex functions arising from convex programming, such as the maximum function (1.8), and we also use an anisotropic lower transform to obtain a $C^{1,1}$ approximation of the squared-distance functions to finite sets.

2. Basic properties of quadratic compensated convex transforms

In this section we establish some basic properties of quadratic compensated convex transforms mentioned in Section 1. Let us first provide some preliminaries concerning the convex envelope $C(f)$ (see [38,39,21].

For a function $f:\mathbb{R}^n \mapsto \mathbb{R}\cup\{+\infty\}$ bounded below, we define its convex envelope $C(f)$ by

$$C\big(f(x)\big) = \sup\big\{g(x),\ g:\mathbb{R}^n \mapsto \mathbb{R} \text{ convex},\ g(y) \leqslant f(y),\ y\in\mathbb{R}^n\big\}. \tag{2.1}$$

An equivalent definition of $C(f(x))$ [39] is that we can replace convex functions in the definition by affine functions, that is,

$$C\big(f(x)\big) = \sup\big\{l(x),\ l:\mathbb{R}^n \mapsto \mathbb{R} \text{ affine},\ l(y) \leqslant f(y),\ y\in\mathbb{R}^n\big\}. \tag{2.2}$$

Therefore, for a function $f \in C^1(\mathbb{R}^n)$, f is convex if and only if

$$f(y) \geqslant f(x) + Df(x)\cdot(y-x), \quad x, y \in \mathbb{R}^n.$$

Furthermore [39, Cor. 17.1.5],

$$C\big(f(x)\big) = \inf\left\{\sum_{i=1}^{n+1}\lambda_i f(x_i),\ \lambda_i \geqslant 0,\ \sum_{i=1}^{n+1}\lambda_i = 1,\ \sum_{i=1}^{n+1}\lambda_i x_i = x\right\}. \tag{2.3}$$

Let $f:\mathbb{R}^n \mapsto \mathbb{R}\cup\{+\infty\}$ be a function, we define its epi-graph [39] by

$$\mathrm{epi}(f) = \big\{(x,t)\in\mathbb{R}^{n+1},\ t \geqslant f(x)\big\}.$$

Remark 2.1. If $f:\mathbb{R}^n \mapsto \mathbb{R}$ is of super-linear growth in the sense that

$$\lim_{|x|\to\infty} \frac{f(x)}{|x|} \to +\infty, \tag{2.4}$$

then we see that there are $x_1,\dots,x_{n+1} \in \mathbb{R}^n$ such that $C(f(x)) = \sum_{i=1}^{n+1} \lambda_i f(x_i)$ for some $\lambda_i \geqslant 0$, $\sum_{i=1}^{n+1} \lambda_i = 1$ with $\sum_{i=1}^{n+1} \lambda_i x_i = x$. This statement should be well known. However, the author cannot find a reference. It can be proved as follows. Since it is known [39, p. 36] that $\mathrm{epi}(C(f)) = C(\mathrm{epi}(f))$ and $(C(f(x)), x) \in \partial C(\mathrm{epi}(f))$ – the boundary of $C(\mathrm{epi}(f))$ and $C(f)$ is of super-linear growth, we see that $\mathrm{epi}(C(f))$ does not contain any extreme rays hence by [39, Th. 18.5], $C(\mathrm{epi}(f))$ is the convex hull of the extreme points of $\mathrm{epi}(f)$. If we let $E \subset \mathbb{R}^{n+1}$ be the supporting plane passing through $(C(f(x)), x)$, then the affine dimension of $\mathrm{epi}(C(f)) \cap E$ is at most n. The point $(C(f(x)), x)$ is then a convex combination of points by the extreme points in $\mathrm{epi}(f) \cap E$. Therefore by Carathéodory's theorem [39, Th. 17.1] (also see the proofs of [39, Th. 17.3, Th. 17.5]), there are at most $n+1$ points $x_1,\dots,x_{n+1} \in \mathbb{R}^n$ such that $C(f(x)) = \sum_{i=1}^{n+1} \lambda_i f(x_i)$. Also under assumption (2.4), we may claim that there is an affine function l_x such that $l_x(y) \leqslant f(y)$ and $l_x(x) = C(f(x))$.

The following are some properties of compensated convex transforms. Theorem 2.1 presents some simple properties of lower quadratic transforms. Theorem 2.2 says that $C^l_{2,\lambda}(f)$ preserves 'monotonicity' of f. We call Theorem 2.3 the Recovery Theorem or Approximation Theorem. We show in this theorem that $C^l_{2,\lambda}(f)$ converges to the original function as $\lambda \to +\infty$. Furthermore, at points $x \in \mathbb{R}^n$ where f is of $C^{1,1}$ near x, $C^l_{2,\lambda}(f(x)) = f(x)$ for large $\lambda > 0$ depending on the local behavior of f near x. Theorem 2.4 is the Locality Theorem which says that for large $\lambda > 0$, the value of $C^l_{2,\lambda}(f(x))$ depends only on that of $f(y) + \lambda|y|^2$ in a small neighborhood of x.

Theorem 2.1. *Let $f:\mathbb{R}^n \mapsto \mathbb{R}$ be a real-valued function.*

(i) *Suppose f satisfies* (1.1)*, then for $\lambda > C_f$, the mapping $\lambda \mapsto C^l_{2,\lambda}(f(x))$ is bounded above by $f(x)$, that is, $C_{2,\lambda}(f(x)) \leqslant f(x)$ for $x \in \mathbb{R}^n$, and is non-decreasing in the sense that*

$$C^l_{2,\lambda}\big(f(x)\big) \leqslant C_{2,\tau}\big(f(x)\big) \quad \text{for } x \in \mathbb{R}^n;\ \lambda < \tau.$$

Furthermore, if $f:\mathbb{R}^n \mapsto \mathbb{R}$ is bounded below, then $C^l_{2,0}(f) = C(f)$.

(ii) *If $f \leqslant g$ in $\mathbb{R}^n$ and satisfy* (1.1)*, then $C^l_{2,\lambda}(f(x)) \leqslant C^l_{2,\lambda}(g(x))$ for $x \in \mathbb{R}^n$ and $\lambda > \max\{C_f, C_g\}$.*

(iii) *Suppose f satisfies* (1.1)*, then for $\lambda > \tau > C_f$,*

$$C^l_{2,\lambda}\big[C^l_{2,\tau}(f)\big] = \begin{cases} C^l_{2,\lambda}(f), & \tau \geqslant \lambda,\\ C^l_{2,\tau}(f), & \tau \leqslant \lambda. \end{cases}$$

Suppose f satisfies (1.3)*, then for $\tau, \lambda > C_f$,*

$$C^u_{2,\lambda}\big[C^u_{2,\tau}(f)\big] = \begin{cases} C^u_{2,\lambda}(f), & \tau \geqslant \lambda,\\ C^u_{2,\tau}(f), & \tau \leqslant \lambda. \end{cases}$$

(iv) *Suppose f satisfies* (1.5)*, then for $\tau > \lambda > C_f$,*

$$C^u_{2,\tau}\big[C^l_{2,\lambda}\big(f(x)\big)\big] = C^u_{2,\tau+\lambda}\big[C\big(f(x) + \lambda|x|^2\big)\big] - \lambda|x|^2,$$
$$C^l_{2,\tau}\big[C^u_{2,\lambda}\big(f(x)\big)\big] = C^l_{2,\tau+\lambda}\big[-C\big(-f(x) + \lambda|x|^2\big)\big] + \lambda|x|^2.$$

(v) (*Rotation Property*) *Suppose f satisfies* (1.1)*. Let Q be an $n\times n$ orthogonal matrix and let $g_Q(x) = f(Qx)$ for $x \in \mathbb{R}^n$, then for $\lambda > C_f$,*

$$C^l_{2,\lambda}\big(f(Qx)\big) = C^l_{2,\lambda}\big(g_Q(x)\big).$$

(vi) (*Translation Property*) *Suppose f satisfies* (1.1)*. Let $h \in \mathbb{R}^n$ be fixed and let $g_h(x) = f(x+h)$ for $x \in \mathbb{R}^n$, then*

$$C^l_{2,\lambda}\big(f(x+h)\big) = C^l_{2,\lambda}\big(g_h(x)\big).$$

(vii) *If f is bounded on every bounded set, then $(x,\lambda)\mapsto C_{2,\lambda}(f(x))$ is locally Lipschitz for (λ,x), with $\lambda > C_f$ and $x\in\mathbb{R}$.*

Properties of quadratic transforms listed in Theorem 2.1 will be used in later applications.

Given $h=(h_1,\dots,h_n)\in\mathbb{R}^n$, we say that h is *non-negative* and write $h\geqslant 0$ if $h_i\geqslant 0$ for $i=1,\dots,n$. A function $f:\mathbb{R}^n\mapsto\mathbb{R}$ is *monotone increasing* (respectively, *decreasing*) if for any $x,h\in\mathbb{R}^n$ with $h\geqslant 0$, $f(x+h)\geqslant f(x)$ (respectively, $f(x+h)\leqslant f(x)$). A simple example of convex and monotone increasing function is the maximum function (1.7). We have

Theorem 2.2. *Suppose $f:\mathbb{R}^n\mapsto\mathbb{R}$ is a monotone increasing function (respectively, decreasing function) satisfying* (1.1). *Then $x\mapsto C^l_{\lambda,2}(f(x))$ is monotone increasing (respectively, monotone decreasing) in $\mathbb{R}^n$ when $\lambda > C_f$.*

The following result is concerned with the effect of continuity, smoothness and local geometry of the original function f on the speed of convergence of $C^l_{2,\lambda}(f)$ to f as $\lambda\to+\infty$.

Theorem 2.3. *(Recovery/Approximation Theorem). Suppose $f:\mathbb{R}^n\mapsto\mathbb{R}$ satisfies* (1.1).

(i) *If $f:\mathbb{R}^n\mapsto\mathbb{R}$ is lower semi-continuous, then for every $x\in\mathbb{R}^n$,*

$$\lim_{\lambda\to+\infty} C_{2,\lambda}(f)(x)=f(x). \tag{2.6}$$

(ii) *Assume that $x_0\in\mathbb{R}^n$ is a local minimum point of f, then there is some $\lambda_{x_0}>0$ such that $C_{2,\lambda}(f)(x_0)=f(x_0)$ whenever $\lambda\geqslant\lambda_{x_0}$.*

(iii) *If $f:\mathbb{R}^n\mapsto\mathbb{R}$ is continuous in $\mathbb{R}^n$, then* (2.6) *holds uniformly on any compact subset of $\mathbb{R}^n$.*

(iv) *If for some $x_0\in\mathbb{R}^n$, $f\in C^{1,1}(\bar B(x_0,\delta))$ for some $\delta>0$, then*

$$C_{2,\lambda}(f)(x_0)=f(x_0),\quad \textit{whenever}\quad \lambda>\max\left\{C(x_0),\ 4C'_f+\frac{1}{\delta^2}\big(2C'_f+1+\big|Df(x_0)\big|\big)\right\}. \tag{2.7}$$

(v) *If $f\in C^{1,1}(\mathbb{R})$ and $L>0$ be such that $|Df(y)-Df(x)|\leqslant L|y-x|$ for all $x,y\in\mathbb{R}^n$, then $C_{2,\lambda}(f)(x)=f(x)$ for all $x\in\mathbb{R}^n$ whenever $\lambda\geqslant L$.*

Remark 2.2. Items (iv)–(vi) in Theorem 2.3 show that the analytic property that $f\in C^{1,1}_{\mathrm{loc}}$ guarantees the finite time attainment of $f(x)$ by quadratic lower transforms. We may therefore view our quadratic lower compensated convex transforms as 'tight approximations' for the original function. This can happen, if for example, the original function is piecewise smooth. The 'tightness' of $C^l_{2,\lambda}(f)$ related to f is a feature of compensated convex transforms which is not shared by other well-known smooth approximations such as the mollified smoothing or Moreau–Yosida regularization. Formula (2.7) provides a precise estimate of λ for which $C^l_{2,\lambda}(f(x_0))$ reaches $f(x_0)$.

Item (ii) shows that the geometric property that x_0 is a local minimum points also implies that $C_{2,\lambda}(f)(x_0)$ reaches $f(x_0)$ in finite time without any continuity assumption. This fact indicates that the quadratic lower compensated convex transform does not improve the regularity of f at such points for large $\lambda>0$.

Theorem 2.4. *(Locality Property). Suppose $f:\mathbb{R}^n\mapsto\mathbb{R}$ satisfies* (1.1) *and is lower semi-continuous. Furthermore, f is bounded on any bounded set. Then for any given $R>0$ and any $\delta>0$, there is some $\Lambda>0$ such that if*

$$\sum_{i=1}^{n+1}\tau_i\big[f(x_i)+\lambda|x_i|^2\big]=C\big[f(x)+\lambda|x|^2\big],\quad \sum_{i=1}^{n+1}\tau_i=1,\quad \tau_i\geqslant 0,\quad i=1,2,\dots,n+1,$$

then $x_i\in B(x,\delta)$ for all $x\in\bar B(0,R)$ whenever $\lambda>\Lambda$. More precisely, the above statement holds if $\lambda>C'_f$ and

$$0<\delta\leqslant\left[\frac{2(M_R+2C'_f+C'_f(R+1)^2)}{\lambda-C'_f}\right]^{1/2} \tag{2.8}$$

where $M_R=\sup\{f(x),x\in\bar B(0,R+1)\}$ and $C'_f>0$ is a constant such that $f(x)\geqslant -C'_f(|x|^2+1)$ for all $x\in\mathbb{R}^n$ which is a consequence of (1.1).

Formula (2.8) provides a precise estimate of the radius δ of the ball where the value of $C^l_{2,\lambda}(f(x_0))$ depends on $f(x)+\lambda|x|^2$ in terms of $\lambda > C_f$.

Our last result in this section provides a connection between quadratic transforms and Moreau envelopes.

Theorem 2.5. *Suppose $f:\mathbb{R}^n \mapsto \mathbb{R}\cup\{+\infty\}$ is not identically $+\infty$ and satisfies* (1.1). *Then $M_{2\lambda}(f) \leqslant C^l_{2,\lambda}(f) \leqslant f$ for $\lambda > C_f$, where $C_f > 0$ is the constant given by* (1.1).

Remark 2.3. Once we have established Theorem 2.5, we see that if $f:\mathbb{R}^n \mapsto \mathbb{R}\cup\{-\infty\}$ is not identically $-\infty$ and satisfies (1.1), then for $\mu > C_f$ given by (1.1), $M_{2\mu}(f) \geqslant C^u_{2,\lambda}(f) \geqslant f$. Thus if $f:\mathbb{R}^n \mapsto \mathbb{R}$ satisfies (1.5), then (1.9) holds, that is,

$$M_{2\lambda}(f) \leqslant C^l_{2,\lambda}(f) \leqslant f \leqslant C^u_{2,\lambda}(f) \leqslant M_{2\lambda}(f).$$

Proof of Theorem 2.1. Most of the statements in item (i) are either obvious or direct consequences of [39, Th. 10.4]. Let us prove that $C^l_{2,\lambda}(f)(x)$ is non-decreasing in λ. By definition, we need to show that

$$C\big[f(x)+\lambda|x|^2\big]-\lambda|x|^2 \leqslant C\big[f(x)+\mu|x|^2\big]-\mu|x|^2, \quad \lambda<\mu.$$

This is equivalent to

$$C\big[f(x)+\lambda|x|^2\big]+(\mu-\lambda)|x|^2 \leqslant C\big[f(x)+\mu|x|^2\big].$$

As $C[f(x)+\lambda|x|^2]+(\mu-\lambda)|x|^2$ is convex and

$$C\big[f(x)+\lambda|x|^2\big]+(\mu-\lambda)|x|^2 \leqslant f(x)+\lambda|x|^2+(\mu-\lambda)|x|^2 = f(x)+\mu|x|^2,$$

we see, due to the fact that convex envelope $C(f)$ is the largest convex function less than or equal to f, that the result follows.

Since (ii) is easy to derive from definition, we now turn to the proofs (iii)–(iv).

Proof of Theorem 2.1(iii). If $\tau \geqslant \lambda$, we have, by (i) that $C^l_{2,\tau}(f) \leqslant f$, hence $C^l_{2,\lambda}[C^l_{2,\tau}(f)] \leqslant C^l_{2,\lambda}(f)$. We only need to show that the opposite inequality also holds. By definition, we need to prove that

$$C^l_{2,\lambda}\big[C^l_{2,\tau}\big(f(x)\big)\big] = C\big\{\big[C\big[f(x)+\tau|x|^2\big]-\tau|x|^2\big]+\lambda|x|^2\big\}-\lambda|x|^2 \geqslant C\big[f(x)+\lambda|x|^2\big]-\lambda|x|^2,$$

which is obviously equivalent to

$$C\big\{\big[C\big[f(x)+\tau|x|^2\big]-(\tau-\lambda)|x|^2\big]\big\} \geqslant C\big[f(x)+\lambda|x|^2\big].$$

Thus we only need to prove that

$$\big[C\big[f(x)+\tau|x|^2\big]-(\tau-\lambda)|x|^2\big] \geqslant C\big[f(x)+\lambda|x|^2\big].$$

The last inequality is equivalent to $C^l_{2,\tau}(f(x)) \geqslant C^l_{2,\lambda}(f(x))$ which is known by (i).

If $\tau \leqslant \lambda$, we have

$$C^l_{2,\tau}\big[f(x)\big]+\lambda|x|^2 = C\big[f(x)+\tau|x|^2\big]+(\lambda-\tau)|x|^2$$

which is a convex function already. Thus

$$C\big[C^l_{2,\tau}\big[f(x)\big]+\lambda|x|^2\big] = C\big[f(x)+\tau|x|^2\big]+(\lambda-\tau)|x|^2,$$

so that

$$C^l_{2,\lambda}\big[C^l_{2,\tau}\big(f(x)\big)\big] = C\big[f(x)+\tau|x|^2\big]-\tau|x|^2 = C^l_{2,\tau}\big(f(x)\big).$$

The proof of the other statement is similar. $\square$

Proof of Theorem 2.1(iv). The conclusions follow from direct calculations as

$$\begin{aligned}C^u_{2,\tau}\big[C^l_{2,\lambda}\big(f(x)\big)\big] &= \tau|x|^2 - C\big[\tau|x|^p - C^l_{2,\lambda}\big(f(x)\big)\big] = \tau|x|^2 - C\big[\tau|x|^2 - \big[C\big(f(x)+\lambda|x|^2\big) - \lambda|x|^2\big]\big]\\ &= \big\{(\tau+\lambda)|x|^2 - C\big[(\tau+\lambda)|x|^2 - C\big[f(x)+\lambda|x|^2\big]\big]\big\} - \lambda|x|^2\\ &= C^u_{2,\tau+\lambda}\big[C\big(f(x)+\lambda|x|^2\big)\big] - \lambda|x|^2.\end{aligned}$$

The proof of the other equality is similar. □

Proof of Theorem 2.1(v)–(vii). We first notice that

$$C^l_{2,\lambda}\big(f(Qx)\big) = C^l_{2,\lambda}\big(f(y)\big)\big|_{y=Qx}, \quad \text{and} \quad C^l_{2,\lambda}\big(f(x+h)\big) = C^l_{2,\lambda}\big(f(y)\big)\big|_{y=x+h}.$$

Thus by definition,

$$C\big(f(Qx)\big) = \inf\left\{\sum_{i=1}^{n+1}\tau_i\big(f(x_i)+\lambda|x_i|^2\big),\ \tau_i \geqslant 0,\ \sum_{i=1}^{n+1}\tau_i = 1,\ \sum_{i=1}^{n+1}\tau_i x_i = Qx\right\}.$$

Let $y_i = Q^{-1}x_i$, $i=1,\dots,n$, we have

$$\begin{aligned}C\big(f(Qx)\big) &= \inf\left\{\sum_{i=1}^{n+1}\tau_i\big(f\big(Q(Q^{-1}x_i)\big)+\lambda|Q^{-1}x_i|^2\big),\ \tau_i \geqslant 0,\ \sum_{i=1}^{n+1}\tau_i = 1,\ \sum_{i=1}^{n+1}\tau_i(Q^{-1}x_i) = x\right\}\\ &= \inf\left\{\sum_{i=1}^{n+1}\tau_i\big(f\big(Q(y_i)\big)+\lambda|y_i|^2\big),\ \tau_i \geqslant 0,\ \sum_{i=1}^{n+1}\tau_i = 1,\ \sum_{i=1}^{n+1}\tau_i(y_i) = x\right\} = C\big(g_Q(x)\big),\end{aligned}$$

as $Q^{-1}:\mathbb{R}^n \mapsto \mathbb{R}^n$ is both onto and isometric. The proof for (v) is finished.

Next we prove (vi). We have

$$C\big(f(x+h)\big) = \inf\left\{\sum_{i=1}^{n+1}\tau_i\big(f(x_i)+\lambda|x_i|^2\big),\ \tau_i \geqslant 0,\ \sum_{i=1}^{n+1}\tau_i = 1,\ \sum_{i=1}^{n+1}\tau_i x_i = x+h\right\}.$$

Similar to the proof of (v), we define $y_i = x_i - h$, $i=1,\dots,n$, and notice that the mapping $x \mapsto x-h$ is onto in $\mathbb{R}^n$, thus

$$\begin{aligned}C\big(f(x+h)\big) &= \inf\left\{\sum_{i=1}^{n+1}\tau_i\big(f(h+y_i)+\lambda|h+y_i|^2\big),\ \tau_i \geqslant 0,\ \sum_{i=1}^{n+1}\tau_i = 1,\ \sum_{i=1}^{n+1}\tau_i y_i = x\right\}\\ &= \inf\left\{\sum_{i=1}^{n+1}\tau_i\big(g_h(y_i)+\lambda|y_i|^2\big),\ \tau_i \geqslant 0,\ \sum_{i=1}^{n+1}\tau_i = 1,\ \sum_{i=1}^{n+1}\tau_i y_i = x\right\} + \lambda\big(|h|^2 + 2x\cdot h\big)\\ &= C\big(g_h(x)\big) + \lambda\big(|h|^2 + 2x\cdot h\big).\end{aligned}$$

Therefore

$$C^l_{2,\lambda}\big(f(x+h)\big) = C\big(g_h(x)\big) + \lambda\big(|h|^2+2x\cdot h\big) - \lambda|x+h|^2 = C^l_{2,\lambda}\big(g_h(x)\big).$$

Finally, we prove (vii). Fix $x \in \mathbb{R}^n$ and $\lambda_0 > C_f$ as required in (1.3). We now show that $(y,\lambda) \mapsto C[f(y)+\lambda|y|^2]$ is Lipschitz continuous in the domain $\bar{B}(x,1)\times[\lambda_0,\lambda_0+\tau]$ for any fixed $\tau > 0$. Since $y \mapsto C[f(y)+\lambda|y|^2]$ is uniformly bounded and Lipschitz in $\bar{B}(x,1)$, for $\lambda \in [\lambda_0,\lambda_0+\tau]$, we have, by [39, Th. 10.6] that $y \mapsto C[f(y)+\lambda|y|^2]$ is uniformly Lipschitz in $\bar{B}(x,1)$ with respect to $\lambda \in [\lambda_0,\lambda_0+\tau]$. Let $M > 0$ be the corresponding Lipschitz constant. Since for each $(y,\lambda) \in \bar{B}(x,1)\times[\lambda_0,\lambda_0+\tau]$, there is a supporting plane $E_{(y,\lambda)}$ of $\mathrm{epi}(C(f(\cdot)+\lambda|\cdot|^2))$ passing through $C[f(y)+\lambda|y|^2], y)$ and by Remark 2.1, $K_{(y,\lambda)} := E_{(y,\lambda)} \cap \mathrm{epi}(f(\cdot)+\lambda|\cdot|^2)$ is compact for each $(y,\lambda) \in \bar{B}(x,1)\times[\lambda_0,\lambda_0+\tau]$. We denote by $P:\mathbb{R}^{n+1}\mapsto\mathbb{R}^n$ the projection $P(y,\lambda) = y$. Let $l_{(y,\lambda)}(\cdot)$ be the affine function representing the plane $E_{(y,\lambda)}$ and we may write it by $l_{(y,\lambda)}(z) = a(y,\lambda)\cdot(z-y) + C[f(y)+\lambda|y|^2]$ with $a(y,\lambda) \in \mathbb{R}^n$. Then

$$l_{(y,\lambda)}(y) = C\big[f(y) + \lambda|y|^2\big], \quad l_{(y,\lambda)}(z) \leqslant f(z) + \lambda|z|^2 \quad (z \in \mathbb{R}^n),$$
$$l_{(y,\lambda)}(z) = f(z) + \lambda|z|^2, \quad z \in P\big[E_{(y,\lambda)} \cap \operatorname{epi}\big(f(\cdot) + \lambda|\cdot|^2\big)\big].$$

Now we claim that $E_{(y,\lambda)} \cap \operatorname{epi}(f(\cdot) + \lambda|\cdot|^2)]$ is uniformly bounded for all $(y,\lambda) \in \bar{B}(x,1) \times [\lambda_0, \lambda_0 + \tau]$. If the claim is not true, there is a sequence $(y_j, \lambda_j) \in \bar{B}(x,1) \times [\lambda_0, \lambda_0 + \tau]$, $j = 1, 2, \ldots$ such that there are some $x_j \in \mathbb{R}^n$ satisfying $l_{(y_j,\lambda_j)}(x_j) = f(x_j) + \lambda_j|x_j|^2$ and $|x_j| \to +\infty$. Thus

$$\begin{aligned} &a(y_j,\lambda_j)\cdot(x_j - y_j) + C\big[f(y_j) + \lambda_j|y_j|^2\big] = f(x_j) + \lambda_j|x_j|^2, \\ &a(y_j,\lambda_j)\cdot(z - y_j) + C\big[f(y_j) + \lambda_j|y_j|^2\big] \leqslant f(z) + \lambda_j|z|^2, \quad z \in \mathbb{R}^n. \end{aligned} \tag{2.9}$$

We write $a_j = a(y_j,\lambda_j)$ for short. Now as $|x_j| \to +\infty$, clearly by (1.1), (2.9) and the fact that $C[f(y_j) + \lambda_j|y_j|^2]$ is uniformly bounded, we have that $|a(y_j,\lambda_j)| \to +\infty$ as $j \to \infty$. On the other hand, if we let $z = a_j/|a_j| - y_j$, then (2.7) implies that

$$|a_j| + C\big[f(y_j) + \lambda_j|y_j|^2\big] \leqslant f\left(\frac{a_j}{|a_j|} - y_j\right) + \lambda_j\left|\frac{a_j}{|a_j|} - y_j\right|^2,$$

which implies that $|a_j|$ is bounded. The claim above is then proved by this contradiction. Let

$$M' = \sup\big\{|z|^2,\ z \in P\big[E_{(y,\lambda)} \cap \big(\operatorname{epi}\big(f(\cdot) + \lambda|\cdot|^2\big)\big)\big],\ (x,\lambda) \in \bar{B}(x,1) \times [\lambda_0, \lambda_0 + \tau]\big\},$$

and consider for $(y_1,\lambda_1), (y_2,\lambda_2) \in \bar{B}(x,1) \times [\lambda_0, \lambda_0 + \tau]$ with $\lambda_1 \leqslant \lambda_2$, we have

$$C\big[f(y_1) + \lambda_1|y_1|^2\big] = \sum_{i=1}^{n+1} \eta_i\big(f(z_i) + \lambda_1|z_i|^2\big)$$

for some $\eta_i \geqslant 0$, $\sum_{i=1}^{n+1} \eta_i = 1$, $\sum_{i=1}^{n+1} \eta_i z_i = y_1$ and $z_i \in P[E_{(y_1,\lambda_1)} \cap (\operatorname{epi}(f(\cdot) + \lambda_1|\cdot|^2))]$. Therefore we have, on one hand, obviously

$$C\big[f(y_1) + \lambda_1|y_1|^2\big] \leqslant C\big[f(y_1) + \lambda_2|y_1|^2\big].$$

On the other hand

$$\begin{aligned} C\big[f(y_1) + \lambda_1|y_1|^2\big] &= \sum_{i=1}^{n+1} \eta_i\big(f(z_i) + \lambda_1|z_i|^2\big) = \sum_{i=1}^{n+1} \eta_i\big(f(z_i) + \lambda_2|z_i|^2\big) - (\lambda_2 - \lambda_1)\left[\sum_{i=1}^{n+1} \eta_i|z_i|^2\right] \\ &\geqslant C\big[f(y_1) + \lambda_2|y_1|^2\big] - (\lambda_2 - \lambda_1)M'. \end{aligned}$$

Thus

$$\big|C\big[f(y_1) + \lambda_2|y_1|^2\big] - C\big[f(y_1) + \lambda_1|y_1|^2\big]\big| \leqslant M'|\lambda_2 - \lambda_1|.$$

Consequently

$$\begin{aligned} &\big|C\big[f(y_2) + \lambda_2|y_1|^2\big] - C\big[f(y_1) + \lambda_1|y_1|^2\big]\big| \\ &\quad \leqslant \big|C\big[f(y_2) + \lambda_2|y_1|^2\big] - C\big[f(y_1) + \lambda_2|y_1|^2\big]\big| + \big|C\big[f(y_1) + \lambda_2|y_1|^2\big] - C\big[f(y_1) + \lambda_1|y_1|^2\big]\big| \\ &\quad \leqslant M|y_2 - y_1| + M'|\lambda_2 - \lambda_1| \leqslant \sqrt{M^2 + \big(M'\big)^2}\sqrt{|y_2 - y_1|^2 + |\lambda_2 - \lambda_1|^2}. \end{aligned}$$

The proof is finished. □

Proof of Theorem 2.2. Suppose f is monotone increasing and fix $h \in \mathbb{R}^n$, $h \geqslant 0$. Let $g_h(x) = f(x + h)$, we have $f(x) \leqslant g_h(x)$ for all $x \in \mathbb{R}^n$, hence for any fixed λ satisfying (1.1),

$$C\big[f(x) + \lambda|x|^2\big] \leqslant g_h(x) + \lambda|x|^2 = \big[f(x+h) + \lambda|x+h|^2\big] - 2\lambda x\cdot h - 2\lambda|h|^2.$$

Similar to the proof of Theorem 2.1(iv), we may define the affine function in x by

$$l_h(x) := -2\lambda x\cdot h - \lambda|h|^2.$$

It is also easy to see from (2.2) that

$$C\big[f(x+h)+\lambda|x+h|^2\big]=C_x\big[f(x+h)+\lambda|x+h|^2\big]$$

where the convex envelope on the right hand side of the above is taken with respect to the x-variable. Thus

$$C\big[f(x)+\lambda|x|^2\big]\leqslant C\big[f(x+h)+\lambda|x+h|^2\big]-2x\cdot h-2|h|^2,$$

which implies that

$$\begin{aligned}C^l_{2,\lambda}\big(f(x)\big) &\leqslant C\big[f(x+h)+\lambda|x+h|^2\big]-\lambda|x|^2-2\lambda x\cdot h-2\lambda|h|^2\\ &=C\big[f(x+h)+\lambda|x+h|^2\big]-\lambda|x+h|^2=C^l_{2,\lambda}\big[f(x+h)\big].\end{aligned}$$

The proof for the case where f is monotone decreasing is similar. □

Proof of Theorem 2.3(i). From the definition of $C^l_{2,\lambda}(f)$, we see that $-\infty<C^l_{2,\lambda}(f)(x)\leqslant f(x)$ for all $\lambda>C_f$ and $x\in\mathbb{R}^n$. We fix $x\in\mathbb{R}^n$ and take any $\epsilon>0$, we try to prove that for sufficiently large $\lambda>0$, we have

$$f(x)\leqslant C^l_{2,\lambda}(f)(x)+\epsilon,\quad \text{equivalently,}\quad f(x)+\lambda|x|^2-\epsilon<C\big[f(x)+\lambda|x|^2\big]. \tag{2.10}$$

Now we take (2.2) as the definition of the convex envelope. We define

$$l_{\lambda,x}(y)=f(x)-\frac{\epsilon}{2}+\lambda|x|^2+2\lambda x\cdot(y-x). \tag{2.11}$$

Clearly, $l_{\lambda,x}(x)=f(x)+\lambda|x|^2-\epsilon/2$. If we can show that

$$l_{\lambda,x}(y)\leqslant f(y)+\lambda|y|^2,\quad y\in\mathbb{R}^n, \tag{2.12}$$

when $\lambda>0$ is sufficiently large, then (2.10) follows and we may conclude our proof. If (2.12) is not true for some $\epsilon>0$, there are $y_j\in\mathbb{R}^n$, $j=1,2,\dots$, such that $l_j(y_j)>f(y_j)+j|y_j|^2$, that is,

$$f(x)-\frac{\epsilon}{2}+j|x|^2+2jx\cdot(y_j-x)>f(y_j)+j|y_j|^2. \tag{2.13}$$

Now for any $0<\delta<1$, we claim that for sufficiently large $j>0$, $|y_j-x|\leqslant\delta$. If this is not the case, there is a subsequence (y_{j_k}) of (y_j), such that $|y_{j_k}-x|>\delta$, $k=1,2,\dots$. Since (2.13) implies that

$$\begin{aligned}f(x)-\frac{\epsilon}{2}&>f(y_{j_k})+j_k\big(|y_{j_k}|^2-|x|^2-2x\cdot(y_{j_k}-x)\big)\\ &\geqslant -C'_f\big(|y_{j_k}|^2+1\big)+j_k\big(|y_{j_k}|^2-|x|^2-2x\cdot(y_{j_k}-x)\big)=-C'_f\big(|y_{j_k}|^2+1\big)+j_k|y_{j_k}-x|^2\\ &\geqslant -2C'_f\big(|y_{j_k}-x|^2+2|x|^2\big)-C_f+j_k|y_{j_k}-x|^2=(j_k-2C'_f)|y_{j_k}-x|^2-2C'_f|x|^2-C_f\\ &\geqslant (j_k-2C'_f)\delta^2-2C'_f|x|^2-C_f.\end{aligned} \tag{2.14}$$

We see that (2.14) cannot hold for large $k>0$.

Now by the lower semi-continuity of f at x, we have, for $\epsilon/4>0$, there is a $\delta>0$, such that $f(y)-f(x)\geqslant\epsilon/4$, whenever $|y-x|\leqslant\delta$. For this $\delta>0$, we see that for sufficiently large $j>0$, $|y_j-x|\leqslant\delta$. Thus for sufficiently large $j>0$, (2.3) implies that

$$-\frac{\epsilon}{2}>f(y_j)-f(x)+j\big(|y_j|^2-j|x|^2-2x\cdot(y_j-x)\big)\geqslant-\frac{\epsilon}{4},$$

and we reach a contradiction. So item (i) is proved. □

Proof of Theorem 2.3(ii). Now we define

$$l_\lambda(y)=f(x_0)+\lambda|x_0|^2+2\lambda x\cdot(y-x_0). \tag{2.15}$$

If we can show that

$$l_\lambda(y)\leqslant f(y)+\lambda|y|^2,\quad y\in\mathbb{R}^n, \tag{2.16}$$

for sufficiently large λ, then, as $l_\lambda(x_0) = f(x_0) + \lambda|x_0|^2$, we may conclude that

$$C\big[f(x_0) + \lambda|x_0|^2\big] = f(x_0) + \lambda|x_0|^2,$$

the proof is then complete. Similar to the proof of (i), we can show that for any $\delta > 0$, (2.16) holds for sufficiently large $\lambda > 0$ when $|y - x_0| \geqslant \delta$. Note that in the proof of (i) for the case $|y - x_0| \geqslant \delta$, we have only used (1.1). We did not use the lower semi-continuity property which is not assumed in the present case.

Now we only need to show that (2.16) remains true when $|y - x_0| \leqslant \delta$. This is easy to proof. Since x_0 is a local minimum point of f, there is a $\delta > 0$, such that $f(y) \geqslant f(x_0)$ whenever $|y - x_0| \leqslant \delta$. Inequality (2.16) is equivalent to

$$f(y) - f(x_0) + \lambda\big[|y|^2 - |x_0|^2 + 2x\cdot(y - x_0)\big] \geqslant 0,$$

which is obviously true for all $\lambda \geqslant 0$ due to $|y - x_0| \leqslant \delta$. □

The proof of Theorem 2.3(iii) is very similar to that of (i) with some minor changes. We leave the proof to interested readers.

Proof of Theorem 2.3(iv). Consider the affine function $l(y)$ which defines the tangent plane of the graph of $f(x) + \lambda|x|^2$ at x_0. We have

$$l_\lambda(y) = f(x_0) + \lambda|x_0|^2 + \big(Df(x_0) + \lambda 2x_0\big)\cdot(y - x_0). \tag{2.17}$$

We only need to show that for sufficiently large $\lambda > 0$, $l(y) \leqslant f(y) + \lambda|y|^2$ for all $y \in \mathbb{R}^n$.

Since $f \in C^{1,1}(\bar{B}(x_0,\delta)$, there is some $C(x_0) > 0$ such that $|Df(x) - Df(x_0)| \leqslant C(x_0)|x - x_0|$ whenever $|x - x_0| \leqslant \delta$. We first show by a direct argument, rather than quoting the proof of (i) that $l_\lambda(y) \leqslant f(y) + \lambda|y|^2$ when $|y - x_0| \geqslant \delta$, that is

$$f(x_0) + \lambda|x_0|^2 + \big(Df(x_0) + 2\lambda x_0\big)\cdot(y - x_0) \leqslant f(y) + \lambda|y|^2,$$

or equivalently,

$$\big[f(y) - f(x_0) - Df(x_0)\cdot(y - x_0)\big] + \lambda\big[|y|^2 - |x_0|^2 - 2x_0\cdot(y - x_0)\big] \geqslant 0. \tag{2.18}$$

We have

$$\begin{aligned}
&\big[f(y) - f(x_0) - Df(x_0)\cdot(y - x_0)\big] + \lambda\big[|y|^2 - |x_0|^2 - 2x_0\cdot(y - x_0)\big]\\
&\quad \geqslant -C'_f\big(|y|^2 + 1\big) - f(x_0) - Df(x_0)\cdot(y - x_0) + \lambda|y - x_0|^2\\
&\quad \geqslant -2C'_f\big(|y - x_0|^2 + |x_0|^2 + 1\big) - f(x_0) - \frac{\lambda}{2}|y - x_0|^2 - \frac{1}{2\lambda}\big|Df(x_0)\big|^2 + \lambda|y - x_0|^2\\
&\quad \geqslant \left(\lambda - 2C'_f - \frac{1}{2}\right)\delta^2 - 2C'_f\big(|x_0|^2 + 1\big) - \frac{1}{2}\big|Df(x_0)\big|^2 \geqslant 0,
\end{aligned}$$

if we take λ such that

$$\lambda - 2C'_f - \frac{1}{2} \geqslant \frac{1}{\delta^2}\left(2C'_f\big(|x_0|^2 + 1\big) + \frac{1}{2\lambda}\big|Df(x_0)\big|^2\right),$$

which is satisfied if

$$\lambda \geqslant 4C'_f + \frac{1}{\delta^2}\big(2C'_f + 1 + \big|Df(x_0)\big|\big).$$

If $|y - x_0| < \delta$, we have, in (2.18) that

$$\begin{aligned}
&\big[f(y) - f(x_0) - Df(x_0)\cdot(y - x_0)\big] + \lambda\big[|y|^2 - |x_0|^2 - 2x_0\cdot(y - x_0)\big]\\
&\quad = \int_0^1 \big[Df\big(x_0 + t(y - x_0)\big) - Df(x_0)\big]\cdot(y - x_0)\,dt + \lambda|y - x_0|^2 \geqslant -C(x_0)|y - x_0|^2 + \lambda|y - x_0|^2 \geqslant 0,
\end{aligned}$$

when $\lambda > C(x_0)$. Thus $C^l_{2,\lambda}(f(x_0)) = f(x_0)$ whenever

$$\lambda > \max\left\{C(x_0),\ 4C'_f + \frac{1}{\delta^2}\big(2C'_f + 1 + \big|Df(x_0)\big|\big)\right\}$$

as claimed in Theorem 2.3(iv). $\square$

Proof of Theorem 2.3(v). By the assumption that $|Df(x) - Df(y)| \leqslant L|x-y|$ for all $x, y \in \mathbb{R}^n$ and proofs of items (iii) and (iv), we see that the conclusion follows if we can prove that

$$\big[f(y) - f(x) - Df(x)\cdot(y-x)\big] + \lambda\big[|y|^2 - |x|^2 - 2x\cdot(y-x)\big] \geqslant 0,$$

for all $x, y \in \mathbb{R}^n$ when $\lambda \geqslant L$. We have

$$\big[f(y) - f(x) - Df(x)\cdot(y-x)\big] + \lambda\big[|y|^2 - |x|^2 - 2x\cdot(y-x)\big]$$
$$= \int_0^1 \big[Df\big(x + t(y-x)\big) - Df(x)\big]\cdot(y-x)\,dt + \lambda|y-x|^2 \geqslant -L|y-x|^2 + \lambda|y-x|^2 \geqslant 0,$$

when $\lambda \geqslant L$. The proof is complete. $\square$

Proof of Theorem 2.4. For a given $R > 0$, let $\bar{B}(0,R) \subset \mathbb{R}^n$ be the closed ball centered at 0 with radius R and let $M_R > 0$ be such that $f(x) \leqslant M_R$ for all $x \in \bar{B}(0, R+1)$. By (1.1) we see that there is some $C'_f > 0$ such that

$$f(x) \geqslant -C'_f\big(|x|^2 + 1\big), \quad x \in \mathbb{R}^n. \tag{2.19}$$

Now given any $0 < \delta < 1/2$, we show that for sufficiently large $\lambda > C$, $x_i \in B(x,\delta)$, $i = 1, \ldots, n+1$, be such that

$$\sum_{i=1}^{n+1} \tau_i\big[f(x_i) + \lambda|x_i|^2\big] = C\big[f(x) + \lambda|x|^2\big], \quad \sum_{i=1}^{n+1}\tau_i = 1, \quad \tau_i \geqslant 0, \quad i = 1, \ldots, n+1.$$

If there is some x_i in the above representation such that $x_i \notin B(x,\delta)$, we consider the convex function $C[f(y) + \lambda|y|^2]$. We see that on the line segment $J = \{y = tx_i + (1-t)x, 0 \leqslant t \leqslant 1\}$, $C[f(y) + \lambda|y|^2]$ is an affine function. Let x_0 be the unique point in $J \cap \partial B(x,\delta)$ and let $z = (x + x_0)/2$, then we have $x_0, z \in \bar{B}(0, R+1)$, $|x - z| = |x_0 - z| = \delta/2$ and

$$C\big[f(z) + \lambda|z|^2\big] = \frac{1}{2}\big(C\big[f(x) + \lambda|x|^2\big] + C\big[f(x_0) + \lambda|x_0|^2\big]\big). \tag{2.20}$$

Thus on one hand,

$$C\big[f(z) + \lambda|z|^2\big] \leqslant M_R + \lambda|z|^2. \tag{2.21}$$

On the other hand, we have

$$\frac{1}{2}\big(C\big[f(x) + \lambda|x|^2\big] + C\big[f(x_0) + \lambda|x_0|^2\big]\big) \geqslant \frac{1}{2}(\lambda - C'_f)\big(|x|^2 + |x_0|^2\big) - 2C'_f$$
$$\geqslant \frac{1}{2}(\lambda - C'_f)\big[\big(|z|^2 + 2z\cdot(x-z) + |x-z|^2\big) + \big(|z|^2 + 2z\cdot(x_0 - z) + |x_0 - z|^2\big)\big] - 2C'_f$$
$$\geqslant (\lambda - C'_f)|z|^2 + (\lambda - C'_f)\frac{\delta^2}{2}\chi - 2C'_f. \tag{2.22}$$

Consequently, by (2.21),

$$(\lambda - C'_f)|z|^2 + (\lambda - C'_f)\frac{\delta^2}{2} - 2C'_f \leqslant M_R + \lambda|z|^2,$$

which leads to

$$(\lambda - C'_f)\frac{\delta^2}{2} \leqslant M_R + 2C'_f + C'_f|z|^2 \leqslant M_R + 2C'_f + C'_f(R+1)^2.$$

This last inequality cannot hold for large $\lambda > C'_f$. Also from the last inequality we obtain the estimate for δ given by (2.8). The proof is finished. □

Proof of Theorem 2.5. The inequality $C^l_{2,\lambda}(f) \leqslant f$ follows from Theorem 2.1(i). Fix $x \in \mathbb{R}^n$ and let $x_i \in \mathbb{R}^n$, $\tau_i \geqslant 0$, $i = 1, 2, \ldots, n+1$ such that $\sum_{i=1}^{n+1} \tau_i = 1$, $\sum_{i=1}^{n+1} \tau_i x_i = x$. The inequality $C^l_{2,\lambda}(f(x)) \geqslant M_{2\lambda}(f(x))$ can then be established by the following argument:

$$\begin{aligned} C^l_{2,\lambda}\big(f(x)\big) &= \inf\left\{\sum_{i=1}^{n+1} \tau_i\big(f(x_i) + \lambda|x_i|^2\big)\right\} - \lambda|x|^2 = \inf\left\{\sum_{i=1}^{n+1} \tau_i\big(f(x_i) + \lambda|x_i - x|^2\big)\right\} \\ &\geqslant \inf\big\{\min_i f(x_i) + \lambda|x_i - x|^2\big\} = M_{2\lambda}\big(f(x)\big). \qquad \square \end{aligned}$$

Remark 2.4. The above inequality shows that our quadratic transforms are more accurate approximations of the original function than the Moreau envelopes. However, this better accuracy also leads to some disadvantages. Recall that for $f : \mathbb{R}^n \mapsto \mathbb{R} \cup \{+\infty\}$ satisfying (1.1), the Lasry–Lions regularity $M^{2\mu}(M_{2\lambda}(f))$ can be defined as an everywhere finite function. This is due to the fact that if f is not identically $+\infty$, the Moreau envelope $M_{2\lambda}(f)$ is everywhere finite and is of quadratic growth. On the other hand, the lower quadratic transform $C^l_{2,\lambda}(f)$ can still take $+\infty$ as its value, which makes the mixed transform $C^u_{2,\mu}[C^l_{2,\lambda}(f)]$ not well-defined. For example, if f is convex and takes $+\infty$ as its value somewhere, then as $C^l_{2,\lambda}(f) = f$, we cannot define the above mixed transform.

3. Quadratic lower transforms and the squared-distance function

Let $K \subset \mathbb{R}^n$ be a non-empty compact set. In this section we study analytic and geometric properties of the quadratic lower transform $C^l_{2,\lambda}(\mathrm{dist}^2(x, K))$ for $0 < \lambda < +\infty$. By a direct application of [11, Th. 5.5], we have the following regularity theorem.

Theorem 3.1. *Suppose $K \subset \mathbb{R}^n$ is non-empty and compact. Then $C^l_{2,\lambda}(\mathrm{dist}^2(x, K)) \in C^{1,1}(\mathbb{R}^n)$ and the Lipschitz constant for the gradient $DC^l_{2,\lambda}(\mathrm{dist}^2(x, K))$ is bounded above by $8 + 10\lambda$. Furthermore, $C^l_{2,\lambda}(\mathrm{dist}^2(x, K)) = 0$ if and only if $x \in K$ and $\lambda > 0$.*

Remark 3.1. The regularity of $C^l_{2,\lambda}(\mathrm{dist}^2(x, K))$ is a direct consequence of [11, Th. C, Th. 5.5] as the function $f_\lambda(x) = \mathrm{dist}^2(x, K) + \lambda|x|^2$ satisfies a stronger version of the upper subdifferentiability of $f_\lambda(\cdot)$:

$$f(x+y) - f(x) - u \cdot y \leqslant C \max\big\{1, f(x)\big\}|y|^2$$

for some $u \in \mathbb{R}^n$ and $|y| \leqslant 1$ and other requirements in [11, Prop. 3.7, Th. 3.5]. We will come to back to this for a more detailed study in Section 4.

Proof of Theorem 3.1. We only prove the second statement, that is, $C^l_{2,\lambda}(\mathrm{dist}^2(x, K)) = 0$ if and only if $x \in K$. For every $x \in \mathbb{R}^n$, clearly $0 \leqslant C^l_{2,\lambda}(\mathrm{dist}^2(x, K)) \leqslant \mathrm{dist}^2(x, K)$. Therefore if $x \in K$, we have $C^l_{2,\lambda}(\mathrm{dist}^2(x, K)) = 0$. Now assume that $C^l_{2,\lambda}(\mathrm{dist}^2(x_0, K)) = 0$ for some $x_0 \in \mathbb{R}^n$, then by Carathéodory's theorem, there are at most $n+1$ points $x_1, \ldots, x_{n+1} \in \mathbb{R}^n$ such that

$$\lambda|x_0|^2 = C\big[\mathrm{dist}^2(x_0, K) + \lambda|x_0|^2\big) = \sum_{i=1}^{n+1} \tau_i\big[\mathrm{dist}^2(x_i, K) + \lambda|x_i|^2\big] \geqslant \lambda \sum_{i=1}^{n+1} \tau_i \lambda|x_i|^2$$

with $\tau_i \geqslant 0$, $i = 1, \ldots, n+1$, $\sum_{i=1}^{n+1} \tau_i = 1$ and $\sum_{i=1}^{n+1} \tau_i x_i = x_0$. As the function $x \mapsto \lambda|x|^2$ is strictly convex, we see that $x_i = x_0$ for all $i = 1, \ldots, n+1$, hence $\mathrm{dist}^2(x_0, K) = 0$. □

The following results concerning the squared distance function to a compact set will be used repeatedly later.

Lemma 3.1. *Suppose $K \subset \mathbb{R}^n$ be a compact set, then*

(i) $C[\operatorname{dist}^2(x,K)] = \operatorname{dist}^2(x, C(K))$ *for all* $x \in \mathbb{R}^n$;
(ii) $0 \leqslant \operatorname{dist}^2(x,K) - \operatorname{dist}^2(x, C(K)) \leqslant \operatorname{diam}^2(K)$ *for all* $x \in \mathbb{R}^n$, *where* $\operatorname{diam}(K)$ *is the diameter of the set* K *defined by* $\operatorname{diam}(K) = \sup\{|x-y|, x, y \in K\}$.
(iii) *If* $K \subset \mathbb{R}^n$ *is convex and compact, then*

$$\operatorname{dist}^2(x,K) = \left|x - P_K(x)\right|^2, \quad x \in \mathbb{R}^n, \tag{3.1}$$

where $P_K(x) \in K$ *is the unique point in* K *such that* (3.1) *holds* [25,21]. *Furthermore,* $P_K : \mathbb{R}^n \mapsto K$ *is continuous and* $D\operatorname{dist}^2(x,K) = 2(x - P_K(x))$, *for all* $x \in \mathbb{R}^n$.

Proof of Lemma 3.1. Item (i) is well known. Even for the more general case of quasi-convex envelope of the p-distance function to a closed set $K \subset M^{N\times n}$ in the calculus of variations, we have that $Q[\operatorname{dist}^p(X,K)] = \operatorname{dist}^p(X, Q_p(K)]$ where $Q_p(K)$ is the quasi-convex hull if K is compact and while $Q_p(K)$ is the p-quasi-convex hull of K if K is unbounded [49]. For the convenience of the reader, we give a proof here.

Clearly, $\operatorname{dist}^2(x,K) \geqslant C\operatorname{dist}^2(x,K) \geqslant \operatorname{dist}^2(x, C(K))$. We prove the opposite inequality. Let $\operatorname{dist}^2(x, C(K)) = |x - x_0|^2$ for some $x_0 \in C(K)$. By Carathéodory's theorem [39], there are at most $n+1$ points $x_1, \dots, x_{n+1} \in K$, such that $x_0 = \sum_{i=1}^{n+1} \lambda_i x_i$ with $\lambda_i \geqslant 0$, $\sum_{i=1}^{n+1} \lambda_i = 1$. Thus by the convexity of $C\operatorname{dist}^2(\cdot, K)$, we have

$$\begin{aligned} C\operatorname{dist}^2(x,K) &= C\operatorname{dist}^2\left(x + \sum_{i=1}^{n+1} \lambda_i(x_i - x_0), K\right) \leqslant \sum_{i=1}^{n+1} \lambda_i C\operatorname{dist}^2(x + x_i - x_0, K) \\ &\leqslant \sum_{i=1}^{n+1} \lambda_i \operatorname{dist}^2(x + x_i - x_0, K) \leqslant \sum_{i=1}^{n+1} \lambda_i |x + x_i - x_0 - x_i|^2 = |x - x_0|^2 = \operatorname{dist}^2\bigl(x, C(K)\bigr). \end{aligned}$$

The proof of (i) is finished.

Item (ii) is a consequence of the Pythagorean theorem. Without loss of generality, we may assume that the affine dimension of $C(K)$ is n [39]. Clearly $\operatorname{dist}^2(x, C(K)) \leqslant \operatorname{dist}^2(x,K)$. Now given $x \in \mathbb{R}^n$ and $\operatorname{dist}^2(x, C(K)) < \operatorname{dist}^2(x,K)$ and $x \notin C(K)$, let $x_0 \in C(K) \setminus K$ be the unique point such that $\operatorname{dist}^2(x, C(K)) = |x - x_0|^2$. Note that for this particular x, $\operatorname{dist}^2(x, C(K)) = \operatorname{dist}^2(x, \partial C(K))$. Let $E \subset \mathbb{R}^n$ be a supporting plane of $C(K)$ passing through x_0 with the smallest dimension, then we see that $x - x_0$ is perpendicular to E. Now we take $x_1 \in E \cap K$ then we have

$$\operatorname{dist}^2(x,K) - \operatorname{dist}^2\bigl(x, C(K)\bigr) \leqslant |x - x_1|^2 - |x - x_0|^2 = |x_1 - x_0|^2 \leqslant \operatorname{diam}^2\bigl(C(K)\bigr) = \operatorname{diam}^2(K).$$

The proof for (ii) is finished.

Item (iii) is well known [21,29,46]. The proof is also easy due to the nearest point property for compact convex sets [25]. So it is easy to see that $P_K(\cdot)$ is continuous [21]. The formula for the gradient follows from the continuity of P_K and the definition of the squared distance function. □

We need the following two lemmas concerning the geometric properties of $C^l_{2,\lambda}(\operatorname{dist}^2(x,K))$.

Lemma 3.2. *Let* $K = (B(0,\rho))^c = \{x \in \mathbb{R}^n,\ |x| \geqslant \rho\}$ *be the complement of the open ball* $B(0,\rho)$ *with* $\rho > 0$, *then*

$$C^l_{2,\lambda}\bigl(\operatorname{dist}^2(x,K)\bigr) = \begin{cases} \frac{\lambda}{1+\lambda}\rho^2 - \lambda|x|^2, & |x| \leqslant \frac{\rho}{1+\lambda}, \\ \operatorname{dist}^2(x,K), & |x| \geqslant \frac{\rho}{1+\lambda}; \end{cases} \tag{3.2}$$

The proof of Lemma 3.2 involves only simple calculations and is left to the reader.

Lemma 3.3. *Let* $n \geqslant 2$ *and let* $\{e_1, \dots, e_n\}$ *be the standard Euclidean basis of* $\mathbb{R}^n$, *where* e_i *is the vector with its* i*th component* 1 *and others zero. Let* $K = \{-\alpha e_1, \alpha e_1\}$, *where* $\alpha > 0$. *We write* $y = e_2 y_2 + \dots + e_n y_n \in \mathbb{R}^{n-1}$ *and* $xe_1 + y := (x, y) = (x, y_2, \dots, y_n) \in \mathbb{R}^n$ *with* $x, y_i \in \mathbb{R}$, $2 \leqslant i \leqslant n$. *Then for every* $\lambda > 0$, *we have*

$$C^l_{2,\lambda}\bigl(\operatorname{dist}^2\bigl((x,y),K\bigr)\bigr) = \begin{cases} \frac{\lambda}{1+\lambda}\alpha^2 - \lambda x^2 + |y|^2, & |x| \leqslant \frac{\alpha}{1+\lambda}, \\ \operatorname{dist}^2((x,y),K), & |x| \geqslant \frac{\alpha}{1+\lambda}. \end{cases} \tag{3.3}$$

In particular,

$$C^l_{2,\lambda}\big(\mathrm{dist}^2\big((0,y),K\big)\big)=\frac{\lambda}{1+\lambda}\alpha^2+|y|^2<\alpha^2+|y|^2=\mathrm{dist}^2\big((0,y),K\big).$$

Proof of Lemma 3.3. We prove the first statement. The second follows from a direct calculation.

The explicit formula of $C^l_{2,\lambda}(\mathrm{dist}^2((x,y),K))$, where $K=\{\alpha e_1,-\alpha e_1\}$ can be obtained as follows. Fix $y\in\mathbb{R}^{n-1}$, we see that $C(\mathrm{dist}^2((x,y),K)+\lambda|(x,y)|^2)$ must be convex in x. As

$$\mathrm{dist}^2\big((x,y),K\big)+\lambda\big|(x,y)\big|^2=\min\big\{(x-\alpha)^2,(x+\alpha)^2\big\}+\lambda x^2+(1+\lambda)|y|^2.$$

For the fixed y, the convex envelope of the above function in x can be easily calculated as

$$C_{e_1}\big[\mathrm{dist}^2\big((x,y),K\big)+\lambda\big|(x,y)\big|^2\big]=(1+\lambda)|y|^2+\begin{cases}\frac{\lambda}{1+\lambda}\alpha^2, & |x|\leqslant\frac{\alpha}{1+\lambda},\\ \mathrm{dist}^2(x,\{\alpha,-\alpha\})+\lambda x^2, & \end{cases}$$

where $C_{e_1}[g(x,y)]$ is the convex envelope of $g(x,y)$ along the one-dimensional space span$[e_1]$. We see that

$$C\big[\mathrm{dist}^2\big((x,y),K\big)+\lambda\big|(x,y)\big|^2\big]\leqslant C_{e_1}\,\mathrm{dist}^2\big((x,y),K\big)+\lambda\big|(x,y)\big|^2\big]\leqslant\mathrm{dist}^2\big((x,y),K\big)+\lambda\big|(x,y)\big|^2,$$

while $C_{e_1}[\mathrm{dist}^2((x,y),K)+\lambda|(x,y)|^2]$ given above is already convex. Thus

$$C\big[\mathrm{dist}^2\big((x,y),K\big)+\lambda\big|(x,y)\big|^2\big]=C_{e_1}\,\mathrm{dist}^2\big((x,y),K\big)+\lambda\big|(x,y)\big|^2$$

so that $C^l_{2,\lambda}(\mathrm{dist}^2((x,y),K))$ is given by (3.3). □

Remark 3.2. The quasi-convex envelope $Q(f)(X)$ for $f(X)=\mathrm{dist}^2(X,\{A,B\})$ for $X,A,B\in M^{N\times n}$ was first calculated by Kohn [22] using a Fourier series method and by Pipkin [35] applying a translation type method. The function $Qf(X)$ was used in several places in the study of quasi-convex functions and quasi-convex hulls [52–56] and explicit calculations using a version of the quadratic lower compensated convex transform for small $\lambda>0$ were used in [53,54,56] to analyze geometric properties for quasi-convex functions $Q[\mathrm{dist}^2(X,K)]$ when $K\subset E$ is a compact subset of a linear subspace E of $M^{N\times n}$ without rank-one matrices. In particular, the case when K is a finite set was studied in [56].

Next we study the local behavior of $C^l_{2,\lambda}(\mathrm{dist}^2(x,K))$ near a compact set $K\subset\mathbb{R}^n$ under certain regularity of K.

It is well known that a closed set $K\subset\mathbb{R}^n$ is convex if and only if K has the *nearest point property* [25,21] that for every $y\in\mathbb{R}^n$, there is a unique $x\in K$ such that $\mathrm{dist}(y,K)=|y-x|$. We have

Definition 3.1. A non-empty closed set $K\subset\mathbb{R}^n$ is said to have the *uniform local nearest point property* if there is an $\epsilon>0$ such that every $y\in K_\epsilon=\{y\in\mathbb{R}^n,\mathrm{dist}(y,K)\leqslant\epsilon\}$ has the nearest point property. The set $K\subset\mathbb{R}^n$ is said to have the *strong local nearest point property* if K has uniform local nearest point property for some $\epsilon>0$ and for every x in the level set $K^\epsilon:=\{z\in\mathbb{R}^n,\mathrm{dist}(z,K)=\epsilon\}$, if we denote by $y(x)\in K$ the unique point in K such that $|x-y(x)|=\epsilon$, then

$$\left\{z=y(x)+t\frac{x-y(x)}{|x-y(x)|},\ 0<t\leqslant\delta,\ x\in K^\epsilon\right\}=K_\delta\setminus K,$$

for all $0<\delta<\epsilon$.

There are many examples of closed sets possessing the strong local nearest point property. Smooth (C^2) compact manifolds without boundary in $\mathbb{R}^n$ [29,46], finite sets, objects obtained by rotating the graph of a non-negative C^2 function $x_2=f(x_1)$ in $\mathbb{R}^2\subset\mathbb{R}^2\times\mathbb{R}^n$ ($n>2$) with respect to the x_1-axis are some of the examples. However, I do not know whether the uniform local nearest point property implies the strong local nearest point property in general.

Theorem 3.4. *Suppose $K\subset\mathbb{R}^n$ is non-empty, closed and satisfies the strong local nearest point property for some $\epsilon>0$. Then for every $\lambda>0$,*

$$C^l_{2,\lambda}\big(\mathrm{dist}^2(x,K)\big)=\mathrm{dist}^2(x,K)\quad \textit{for } x\in K_\delta \textit{ with } \delta\leqslant\frac{\lambda}{1+\lambda}\epsilon.$$

Proof. Let $z\in K_\delta\setminus K$. By our assumption on K, there is some $x\in K_\epsilon$ and $y(x)\in K$ such that $z=y(x)+t(x-y(x))/|x-y(x)|$ for some $0<t\leqslant\delta$. Clearly $y(x)$ is also the unique point in K such that $|z-y(x)|=\mathrm{dist}(z,K)=t$ holds. Now we consider the function $y\mapsto \mathrm{dist}^2(y,B^c(x,\epsilon))$. By a simple translation and by replacing ρ by ϵ, we obtain from Lemma 3.2 that

$$C^l_{2,\lambda}\big(\mathrm{dist}^2\big(y,B^c(x,\epsilon)\big)\big)=\mathrm{dist}^2\big(y,B^c(x,\epsilon)\big),\quad \text{whenever } |y-x|\geqslant\frac{\epsilon}{1+\lambda}.$$

Now since $|z-x|=\epsilon-t\geqslant\epsilon/(1+\lambda)$, we have

$$C^l_{2,\lambda}\big(\mathrm{dist}^2\big(z,B^c(x,\epsilon)\big)\big)=\mathrm{dist}^2\big(z,B^c(x,\epsilon)\big)=t^2=\mathrm{dist}^2(z,K).$$

On the other hand, since $K\subset B^c(x,\epsilon)$ we have $\mathrm{dist}^2(y,B^c(x,\epsilon))\leqslant\mathrm{dist}^2(y,K)$ for all $y\in\mathbb{R}^n$ so that by Theorem 1.1, we have that

$$C^l_{2,\lambda}\big(\mathrm{dist}^2\big(y,B^c(x,\epsilon)\big)\big)\leqslant C^l_{2,\lambda}\big(\mathrm{dist}^2(y,K)\big)\leqslant\mathrm{dist}^2(y,K)$$

for all $y\in\mathbb{R}^n$. Thus at $y=z$ we have $C^l_{2,\lambda}(\mathrm{dist}^2(z,K))\leqslant\mathrm{dist}^2(z,K)$. □

Next we apply Theorem 3.4 to the study of the quasi-convex sets in the calculus of variations. Let $M^{N\times n}$ be the space of all real $N\times n$ matrices equipped with the Euclidean norm. A function $f:M^{N\times n}\mapsto\mathbb{R}$ is called quasi-convex [30,5] if for every $X\in M^{N\times n}$, every open set $\Omega\subset\mathbb{R}^n$ and every $\phi\in C^1_0(\Omega,\mathbb{R}^N)$, we have $\int_\Omega f(X+D\phi(x))-f(X)\,dx\geqslant 0$. Quasi-convexity is a fundamental concept in the vectorial calculus of variations concerning the weak lower semi-continuity if variational integrals [2]. Similar to convex envelope for a non-convex function, we may define the quasi-convex envelope [17] of a given function $f:M^{N\times n}\mapsto\mathbb{R}$ bounded below as $Q(f(X))=\sup\{g(X),\ g\leqslant f,\ g \text{ quasi-convex}\}$. The sub-level set of a quasi-convex function is called a quasi-convex set [42,50]. Since quasi-convex sets need not to be convex, a topological question arises as how complicated a quasi-convex set can be. Suppose $E\subset M^{N\times n}$ is a linear subspace without rank-one matrices, it is well known [7,56] that any compact set $K\subset E$ is a quasi-convex set. This indicates that any topological structure that can be embedded in E as a closed set can be realized by a quasi-convex set. It is known [53] that every linear subspace E without rank-one matrices can be extended to a maximal subspace $E_{\max}$ satisfying the same restriction and $\dim(E_{\max})\geqslant(N-1)(n-1)$. Therefore any compact set in $\mathbb{R}^{(N-1)(m-1)}$ can be embedded in $M^{N\times n}$ as a quasi-convex set. On the other hand, it is easy to see that if any topological embedding of a codimension one compact manifold without boundary that divide $M^{N\times n}$ into disconnected domains cannot be a quasi-convex set. Our question now is whether a tubular neighborhood in $M^{N\times n}$ of a 'smooth' quasi-convex set $K\subset E$ remains a 'quasi-convex body'.

Now given a linear subspace without rank-one matrices. Let P_E and $P_{E^\perp}$ be the orthogonal projections from $M^{N\times n}$ to E and its orthogonal complement $E^\perp$. Given $a\in\mathbb{R}^n$, $b\in\mathbb{R}^N$ (viewing as column vectors) we denote by $a\otimes b=ab^T\in M^{N\times n}$ the tensor product of a and b. Clearly every rank-one matrix can be written as the tensor product of some a and b. We define $\lambda_E>0$ by

$$\frac{1}{\lambda_E}=\sup_{|a|=|b|=1,a\in\mathbb{R}^n,b\in\mathbb{R}^N}\frac{|P_E(a\otimes b)|^2}{|P_{E^\perp}(a\otimes b)|^2}.$$

It is easy to see that $0<\lambda_E<\infty$ and λ_E is the largest positive number such that

$$\big|P_{E^\perp}(a\otimes b)\big|^2\geqslant\lambda_E\big|P_E(a\otimes b)\big|^2,\quad a\in\mathbb{R}^n,\ b\in\mathbb{R}^N.$$

It is also easy to see that [54,56]

$$q(X)=\big|P_{E^\perp}(X)\big|^2-\lambda_E\big|P_E(X)\big|^2,\quad X\in\mathbb{M}^{N\times n}$$

is a rank-one convex quadratic form. We have

Theorem 3.5. *Suppose $K\subset E\subset M^{N\times n}$ with E a subspace without rank-one matrices and K has the strong local nearest point property for some $\epsilon>0$. Then the closed δ-neighborhood K_δ of K in $M^{N\times n}$ defined by $K_\delta=\{X\in M^{N\times n},\ \mathrm{dist}(X,K)\leqslant\delta\}$ is a quasi-convex set when $\delta\leqslant\epsilon\lambda_E/(1+\lambda_E)$.*

Proof of Theorem 3.5. Since $K \subset E$, we have $\operatorname{dist}^2(X,K) = \operatorname{dist}^2(P_E(X),K) + |P_{E^\perp}(X)|^2$. We define

$$\begin{aligned} F(X) &= C\big[\operatorname{dist}^2\big(P_E(X),K\big) + \lambda_E\big|P_E(X)\big|^2\big] + \big[\big|P_{E^\perp}(X)\big|^2 - \lambda_E\big|P_E(X)\big|^2\big] \\ &= C^l_{2,\lambda_E}\big[\operatorname{dist}^2\big(P_E(X),K\big)\big] + \big|P_{E^\perp}(X)\big|^2, \end{aligned}$$

where $C^l_{2,\lambda_E}[\operatorname{dist}^2(P_E(X),K)]$ is taken in E identified as an Euclidean space and evaluated at $P_E(X) \in E$. Clearly, we see that $\operatorname{dist}^2(X,K) \geqslant F(X)$ for $X \in M^{N\times n}$ and we also see that $F(X)$ is quasi-convex as $C[\operatorname{dist}^2(P_E(X),K) + \lambda_E|P_E(X)|^2]$ is convex and $|P_{E^\perp}(X)|^2 - \lambda_E|P_E(X)|^2$ is a rank-one convex quadratic form.

The conclusion then follows from Theorem 3.3 with $\delta_2 \leqslant \lambda_E\epsilon/(1+\lambda_E)$ and if $X \in K_\delta$, $P_E(X) \in P_E(K_\delta) \subset E$. □

Remark 3.3. When $K \subset E$ is a finite set Theorem 3.5 was essentially proved in [56]. From the recovery/approximation theorem (Theorem 2.3) we see that $C^l_{2,\lambda}(\operatorname{dist}^2(x,K))$ converges to $\operatorname{dist}^2(x,K)$ uniformly on compact sets. By Theorem 3.1 we also see that $C^l_{2,\lambda}(\operatorname{dist}^2(x,K))$ is locally a $C^{1,1}$ function for any $\lambda > 0$. It was established in [51] that the squared-distance function $X \mapsto \operatorname{dist}^2(X,K)$ in $M^{N\times n}$ is quasi-convex (rank-one convex respectively) if and only if the compact set $K \subset M^{N\times n}$ is convex. The key idea in [51] is to use the nearest point property and a version of Lemma 3.3.

By using the regularity results on $Q\operatorname{dist}^2(X,K)$, Ball, Kirchheim and Kristensen gave a pure 'analytic' proof of this result in [11, Th. 5.5]. If a compact set $K \subset M^{N\times n}$ is not convex, $\operatorname{dist}^2(X,K)$ is not locally $C^{1,1}$ while $Q\operatorname{dist}^2(X,K)$ belongs to $C^{1,1}(M^{N\times n})$, the conclusion follows. However, if one replace $\operatorname{dist}^2(X,K)$ by $C^l_{2,\lambda}(\operatorname{dist}^2(X,K))$ for some large $\lambda > 0$, then the 'geometric' argument in [51] still works while the 'analytic' argument [11, Th. 5.5] based on smoothness of the function no longer applies. More precisely, we have

Theorem 3.6. *Let* $\min\{N,n\} \geqslant 2$. *For every non-convex compact set* $K \subset M^{N\times n}$*, there is some* $\Lambda(K) > 0$ *such that the* $C^{1,1}$ *function* $X \mapsto C^l_{2,\lambda}(\operatorname{dist}^2(X,K))$ *is not quasi-convex whenever* $\lambda > \Lambda(K)$.

Proof of Theorem 3.6. We combine the method of proof in [51], Lemmas 3.2 and 3.3. If K is not convex, by the nearest point property for convex sets [25], there is some point $X_0 \in M^{N\times n}$ and $X_1, X_2 \in K$, $X_1 \neq X_2$ such that $\operatorname{dist}(X_0,K) = |X_0 - X_1| = |X_0 - X_2| =: r > 0$. By a simple translation of the origin to $(X_1+X_2)/2$ we may assume that $X_1 = \alpha A_0$, $X_2 = -\alpha A_0$ with $\operatorname{rank}(A_0) > 1$, $|A_0| = 1$ and $\alpha > 0$. Let $E = \operatorname{span}[A_0]$ be the one-dimensional subspace of $M^{N\times n}$ spanned by A_0, P_E and $P_{E^\perp}$ the orthogonal projections to E and its complement $E^\perp$. Now we set $K_0 = \{\alpha A_0, -\alpha A_0\}$ and take $\Lambda(K) = n - 1 + n|P_{E^\perp}(X_0)|^2/\alpha^2$, then we have $\operatorname{dist}^2(X,K) \leqslant \operatorname{dist}(X,K_0)$ for $X \in M^{N\times n}$ and $\operatorname{dist}^2(X_0,K) \leqslant \operatorname{dist}^2(X_0,K_0)$. Thus by Lemma 3.2, we have, for $\lambda > \Lambda(K)$ that

$$\begin{aligned} \frac{\lambda}{1+\lambda}r^2 - \lambda|X - X_0|^2 &= C^l_{2,\lambda}\operatorname{dist}^2\big(X, \big(\bar B(X_0,r)\big)^c\big) \\ &\leqslant C^l_{2,\lambda}\big[\operatorname{dist}^2(X,K)\big] \leqslant C^l_{2,\lambda}\big[\operatorname{dist}^2(X,K_0)\big] \leqslant \operatorname{dist}^2(X,K_0). \end{aligned}$$

If $C^l_{2,\lambda}[\operatorname{dist}^2(X,K)]$ is quasi-convex, then by the definition of quasi-convex envelope, we have

$$C^l_{2,\lambda}\big[\operatorname{dist}(X,K)\big] \leqslant Q\big[\operatorname{dist}^2(X,K_0)\big].$$

Now it is known [22,51,54,56] that

$$\begin{aligned} Q\big[\operatorname{dist}^2(X,K_0)\big] &= C^l_{2,\lambda^*}\big[\operatorname{dist}^2(X,K_0)\big] \\ &= (1+\lambda^*)\operatorname{dist}^2\left(P_E(X), \frac{1}{1+\lambda^*}C(K_0)\right) + \alpha^2\frac{\lambda^*}{1+\lambda^*} - \lambda\big|P_E(X)\big|^2 + \big|P_{E^\perp}(X)\big|^2, \end{aligned}$$

where $\lambda^* = (1-\lambda_{\max})/\lambda_{\max}$ with $\lambda_{\max}$ the largest eigenvalue of $A_0^T A_0$. Note that since $|A_0| = 1$, we have $1/n \leqslant \lambda_{\max} < 1$ as A_0 is not a rank-one matrix. Thus $\lambda^* \leqslant n-1$. Now we have

$$C^l_{2,\lambda}\big[\operatorname{dist}^2(X,K)\big] \leqslant Q\operatorname{dist}^2(X,K_0).$$

In particular, at X_0, by noticing that $\alpha^2 + |P_{E^\perp}(X_0)|^2 = r^2$, we have $P_E(X_0) = 0$ and

$$\frac{\lambda}{1+\lambda}r^2 \leqslant C^l_{2,\lambda}\big[\operatorname{dist}^2(X_0,K_0)\big] =\leqslant Q\operatorname{dist}^2(X_0,K_0) = \alpha^2\frac{\lambda^*}{1+\lambda^*} + \big|P_{E^\perp}(X_0)\big|^2.$$

This implies that

$$\lambda < \lambda^* + (1+\lambda^*)\frac{|P_{E^\perp}(X_0)|^2}{\alpha^2} \leqslant n-1+n\frac{|P_{E^\perp}(X_0)|^2}{\alpha^2} = \Lambda(K).$$

This contradict to our choice of $\lambda > \Lambda(K)$. □

Next we consider the notion of medial axis related to global shape interrogation and representation in the theory of computer aided design. For bounded domain $\Omega \subset \mathbb{R}^2$, the earlier definition of the *medial axis* was given by Blum [12] who generalized the so-called Voronoi diagram [45] for finite sets. We take the definition by Wolter [46]. Given a bounded open domain $\Omega \subset \mathbb{R}^n$ whose boundary $\partial\Omega$ is an $(n-1)$-dimensional topological manifold, the *medial axis* M_Ω is the closure of the set $M^0_\Omega \subset \Omega$, where M^0_Ω consists all of the points $x \in \Omega$ and a corresponding radius $r(x) > 0$ such that $B(x, r(x)) \subset \Omega$ is the maximal open ball in the sense that any open ball B containing $B(x, r(x))$ cannot be completely contained in Ω. An simple example of points $x \in M^0_\Omega$ is that $\operatorname{dist}(x, \partial\Omega)$ is reached at least two different points on $\partial\Omega$. The set $\{(x, r(x)), x \in M_\Omega\}$ is called the *medial axis transform* of Ω.

Next we define the *cut-locus* for Ω under the same assumption as above. We call $x \in \Omega$ is a non-extender [46] if for any $y \in \partial\Omega$ such that $\operatorname{dist}(x, \partial\Omega) = |y - x|$, $\operatorname{dist}(y + t(x - y), \partial\Omega) < t|x - y|$ for any $t > 1$. The cut-locus C_Ω for Ω is defined as the closure of the set of all non-extenders for Ω [46]. It was proved in [46], among other results, that $M_\Omega = C_\Omega$ and $\bar\Omega = \bigcup_{x\in M_\Omega} \bar B(x, r(x))$ under the above assumption on Ω. It was also established in [46] that if $\partial\Omega$ is smooth (C^2), then M_Ω is a deformation retraction of $\bar\Omega$.

For a finite set $K \subset \mathbb{R}^n$ we may take $\Omega = \mathbb{R}^n \setminus K$ and define the medial axis of K by $M_K := M_\Omega$. It is well known that M_K is the so-called Voronoi diagram [46,45]. In this special case $M_\Omega = M^0_\Omega$, that is, for every point in $x \in M_\Omega$ there are at least two distinct points $y_1, y_2 \in K$ such that $\operatorname{dist}(x, K) = |x - y_1| = |x - y_2|$.

Our aim in the last part of this section is to give a geometric description of the approximation $C^l_{2,\lambda}(\operatorname{dist}^2(x, \partial\Omega))$ for a bounded open domain $\Omega \subset \mathbb{R}^n$ and for a finite set K. The main tools we use are Lemmas 3.2 and 3.3. We have

Theorem 3.7. *For any $\lambda > 0$, let*

$$r_2(x,\lambda) = \frac{r(x)}{1+\lambda}, \qquad M^{(2)}_\lambda = \bigcup_{x\in M_\Omega\cap\Omega} B\left(x, \frac{r(x)}{1+\lambda}\right).$$

Let $x \in M_\Omega \cap \Omega$ and $K(x) = \{y + x \in \partial\Omega, |y| = r(x)\}$. For such an x, we define

$$K_{2,\lambda}(x) = \left\{x + \frac{y}{1+\lambda},\ x+y \in \partial\Omega,\ |y| = r(x)\right\},$$
$$M^*_\Omega = \big\{x \in M_\Omega \cap \Omega,\ K(x) \text{ contains at least two elements}\big\}.$$

Then

(i)
$$C^l_{2,\lambda}\big(\operatorname{dist}^2(x,\partial\Omega)\big) = \operatorname{dist}^2(x,\partial\Omega), \quad x \in \Omega \setminus M^{(2)}_\lambda,$$
$$C^l_{2,\lambda}\big(\operatorname{dist}^2(y,\partial\Omega)\big) < \operatorname{dist}^2(y,\partial\Omega), \quad y \in M^*_\Omega, \quad \lambda > 0;$$

(ii) *For each $x \in M_\Omega \cap \Omega$,*
$$C^l_{2,\lambda}\big(\operatorname{dist}^2(y,\partial\Omega)\big) = \frac{1}{1+\lambda}r^2(x) - \lambda|y-x|^2, \quad y \in C\big[K_{2,\lambda}(x)\big].$$

(iii) *If we further assume that $x \in M_\Omega \cap \Omega \cap C(K(x))$, and for every line passing through x, the intersection of the line and $C(K(x))$ contains more than one element, then x is a stationary point of $C^l_{2,\lambda}(\operatorname{dist}^2(x,\partial\Omega))$, that is, $DC^l_{2,\lambda}(\operatorname{dist}^2(x,\partial\Omega)) = 0$ for all $\lambda > 0$.*
The above condition is satisfied if $x \in M_\Omega \cap \Omega$ is an interior point of $C(K(x))$. In this case x is also a local maximum point of $C^l_{2,\lambda}(\operatorname{dist}^2(x,\partial\Omega))$.

Proof of Theorem 3.7. We prove (i) first. Fix $y \in \Omega \setminus M_\lambda^{(2)}$, then there is a unique $x_0 \in \partial\Omega$ such that $\operatorname{dist}(y, \partial\Omega) = |x - x_0|$. Now we consider the ray $l = \{x_0 + t(x - x_0), t \geqslant 0\}$ and let

$$w = x_0 + t_0(x - x_0), \quad \text{with} \quad 1 < t_0 = \sup\{t, x_0 + t(x - x_0) \notin M_\Omega)\}.$$

Such $t_0 > 1$ exists because Ω is bounded and $x \in \Omega \setminus M_\Omega$, there is some $t > 1$ such that $x_0 + t(x - x_0)$ is a non-extender.

Now we have for all $y \in \mathbb{R}^n$,

$$\operatorname{dist}^2\big(y, B\big(w, r(w)\big)^c\big) \leqslant \operatorname{dist}^2(y, \partial\Omega), \qquad C_{2,\lambda}^l\big(\operatorname{dist}^2\big(y, B\big(w, r(w)^c\big)\big)\big) \leqslant C_{2,\lambda}^l\big(\operatorname{dist}^2(y, \partial\Omega)\big).$$

By Lemma 3.2(b), we have, by a simple translation of the origin,

$$C_{2,\lambda}^l\big(\operatorname{dist}^2\big(y, B\big(w, r(w)\big)^c\big)\big) = \begin{cases} \frac{\lambda}{r^2(w)} 1 + \lambda - \lambda|y - w|^2, & |y - w| \leqslant \frac{r(w)}{1+\lambda}, \\ \operatorname{dist}^2(y, B(w, r(w))^c), & |y - w| \geqslant \frac{r(w)}{1+\lambda}. \end{cases}$$

As $x \in B(w, r(w)) \setminus B(w, r(w)/(1 + \lambda)$, we have

$$C_{2,\lambda}^l\big(\operatorname{dist}^2\big(x, B\big(w, r(w)^c\big)\big)\big) = \operatorname{dist}^2\big(w, B\big(w, r(w)^c\big)\big) = |x - x_0|^2 = \operatorname{dist}^2(x, \partial\Omega).$$

Next we show that

$$C_{2,\lambda}^l\big(\operatorname{dist}^2(x, \partial\Omega)\big) < \operatorname{dist}^2(x, \partial\Omega)$$

for $x \in M_\Omega^*$. Taking any two points $x_1, x_2 \in K(x)$ with $x_1 \neq x_2$. By a simple translation of the origin to $(x_1 + x_2)/2$ followed by a rotation, we may assume that $x_1 = (-\alpha, 0)$, $x_2 = (\alpha, 0)$ with $\alpha > 0$ and $x = (0, y)$ with $y \in \mathbb{R}^{n-1}$. By Lemma 3.3, we have $x = (0, y)$,

$$C_{2,\lambda}^l\big(\operatorname{dist}^2(x, \partial\Omega)\big) \leqslant C_{2,\lambda}^l\big(\operatorname{dist}^2\big(x, \big\{(-\alpha, 0), (\alpha, 0)\big\}\big)\big) < \operatorname{dist}^2\big(x, \big\{(-\alpha, 0), (\alpha, 0)\big\}\big) = \operatorname{dist}^2(x, \partial\Omega).$$

Now we prove (ii). Firstly, we give an upper bound of $C_{2,\lambda}^l(\operatorname{dist}^2(y, \partial\Omega))$. Since

$$\operatorname{dist}^2(y, \partial\Omega) \leqslant \operatorname{dist}^2\big(y, K(x)\big), \quad y \in \mathbb{R}^n,$$

we have

$$C_{2,\lambda}^l\big(\operatorname{dist}^2(y, \partial\Omega)\big) \leqslant C_{2,\lambda}^l\big(\operatorname{dist}^2\big(y, K(x)\big)\big), \quad y \in \mathbb{R}^n.$$

By noticing that $|z - x| = r(x)$ for all $z \in K(x)$

$$\begin{aligned}
\operatorname{dist}^2\big(y, K(x)\big) + \lambda|y|^2 &= \inf_{z \in K(x)} |y - z|^2 + \lambda|y|^2 \\
&= \Big[\inf_{z \in K(x)} \big(|y - x|^2 - 2(y - x) \cdot (z - x) + |z - x|^2\big)\Big] + \lambda|y - x|^2 \\
&\quad + 2\lambda x \cdot (y - x) + \lambda|x|^2 \\
&= \Big[\inf_{z \in K(x)} \big((1 + \lambda)|y - x|^2 - 2(y - x) \cdot (z - x) + |z - x|^2\big)\Big] + 2\lambda x \cdot (y - x) + \lambda|x|^2 \\
&= (1 + \lambda)\left[\inf_{z \in K(x)} \left|(y - x) - \frac{z - x}{1 + \lambda}\right|^2\right] + \frac{\lambda r^2(x)}{1 + \lambda} + 2\lambda x \cdot (y - x) + \lambda|x|^2 \\
&= (1 + \lambda) \operatorname{dist}^2\big(y, K_{2,\lambda}(x)\big) + l(y),
\end{aligned}$$

where $l(y) = \lambda r^2(x)/(1 + \lambda) + 2\lambda x \cdot (y - x) + \lambda|x|^2$ is an affine function of $y \in \mathbb{R}^n$. Thus

$$C\big[(1 + \lambda) \operatorname{dist}^2\big(y, K(x)\big) + \lambda|y|^2\big] = (1 + \lambda) \operatorname{dist}^2\big(y, C\big[K_{2,\lambda}(x)\big]\big) + l(y),$$

so that

$$
\begin{aligned}
C^l_{2,\lambda}\big[\operatorname{dist}^2\big(y,K(x)\big)\big] &= (1+\lambda)\operatorname{dist}^2\big(y,C\big[K_{2,\lambda}(x)\big]\big)+l(y)-\lambda|y|^2\\
&= (1+\lambda)\operatorname{dist}^2\left(y,\frac{\lambda r^2(x)}{1+\lambda}\right)+\frac{\lambda r^2(x)}{1+\lambda}+2\lambda x\cdot(y-x)+\lambda|x|^2-\lambda|y-x|^2\\
&\quad -2\lambda x\cdot(y-x)-\lambda|x|^2\\
&= (1+\lambda)\operatorname{dist}^2\big(y,C\big[K_{2,\lambda}(x)\big]\big)+\frac{\lambda r^2(x)}{1+\lambda}-\lambda|y-x|^2.
\end{aligned}
$$

Note that $C[K_{2,\lambda}(x)]\subset \bar B(x,r(x)/(1+\lambda))$, hence we have, for $y\in C[K_{2,\lambda}(x)]$ that

$$
C^l_{2,\lambda}\big[\operatorname{dist}^2\big(y,K(x)\big)\big]=\frac{\lambda r^2(x)}{1+\lambda}-\lambda|y-x|^2=C^l_{2,\lambda}\big(\operatorname{dist}^2\big(y,B\big(x,r(x)^c\big)\big)\big).
$$

Therefore

$$
C^l_{2,\lambda}\big(\operatorname{dist}^2(y,\partial\Omega)\big)=C^l_{2,\lambda}\big[\operatorname{dist}^2\big(y,K(x)\big)\big]=C^l_{2,\lambda}\big(\operatorname{dist}^2\big(y,B\big(x,r(x)^c\big)\big)\big)=\frac{\lambda r^2(x)}{1+\lambda}-\lambda|y-x|^2. \tag{3.5}
$$

Finally, we prove (iii). By Theorem 3.1, we have $C^l_{2,\lambda}(\operatorname{dist}^2(\cdot,\partial\Omega))\in C^{1,1}(\bar\Omega)$. Now we assume that $x\in M_\Omega\cap\Omega\cap C(K(x))$, and for every line passing through x, the intersection of the line and $C(K(x))$ contains more than one element. Then for any e_i among the standard Euclidean basis, we have either te_i+x or $-te_i\in C(K_{2,\lambda}(x))$ for sufficiently small $t>0$. Thus

$$
\frac{d}{ds}C^l_{2,\lambda}\big(\operatorname{dist}^2(x+se_i,\partial\Omega)\big)\big|_{s=0}=\frac{d}{ds}\left(\frac{\lambda r^2(x)}{1+\lambda}-\lambda|x+se_i-x|^2\right)\bigg|_{s=0}=0.
$$

Therefore

$$
\frac{\partial}{\partial x_i}C^l_{2,\lambda}\big(\operatorname{dist}^2(x,\partial\Omega)\big)=0,\quad i=1,\dots,n,\quad\text{hence}\quad DC^l_{2,\lambda}\big(\operatorname{dist}^2(x,\partial\Omega)\big)=0.
$$

When $x\in M_\Omega\cap\Omega\cap C(K(x))$ is an interior point of $C(K(x))$, then in a the neighborhood $C[K_{2,\lambda}(x)]$ of x, by (3.5) we have

$$
C^l_{2,\lambda}\big(\operatorname{dist}^2(y,\partial\Omega)\big)=\frac{\lambda r^2(x)}{1+\lambda}-\lambda|y-x|^2,\quad y\in C\big[K_{2,\lambda}(x)\big].
$$

Thus x is a local maximum point of $C^l_{2,\lambda}(\operatorname{dist}^2(y,\partial\Omega))$. □

4. Quadratic upper transforms for convex functions

In this section we establish the $C^{1,1}$ smoothness and convexity for the quadratic upper compensated convex transform $C^u_{2,\lambda}(f(x))$ for $f:\mathbb{R}^n\mapsto\mathbb{R}$ satisfying the growth condition (1.5), that is,

$$
\big|f(x)\big|\leqslant C|x|^2+C_1,\quad x\in\mathbb{R}^n, \tag{4.1}
$$

where $C>0$ and $C_1>0$ are constants. As a consequence, we also prove a general $C^{1,1}$ approximation theorem on compact subsets of $\mathbb{R}^n$ for continuous functions satisfying (4.1).

Theorem 4.1. *Suppose $f:\mathbb{R}^n\mapsto\mathbb{R}$ is convex and satisfies* (4.1). *Then for $\lambda>C$,*

(i) *$x\mapsto C^u_{2,\lambda}(f(x))$ is convex in $\mathbb{R}^n$;*
(ii) *$C^u_{2,\lambda}(f(x))\in C^{1,1}(\mathbb{R}^n)$, and more precisely, the gradient of $C^u_{2,\lambda}(f(\cdot))$ satisfies*

$$
\big|D\big[C^u_{2,\lambda}\big(f(x)\big)\big]-D\big[C^u_{2,\lambda}\big(f(y)\big)\big]\big|\leqslant 8\lambda|x-y|,\quad x,y\in\mathbb{R}^n. \tag{4.2}
$$

Corollary 4.1. *Suppose $f:\mathbb{R}^n\mapsto\mathbb{R}$ is continuous and satisfies* (4.1). *Then, there are $\tau_j>\lambda_j\to+\infty$ as $j\to\infty$ such that $C^u_{2,\tau_j}[C^l_{2,\lambda_j}(f(\cdot))]\in C^{1,1}(\mathbb{R}^n)$ and on every compact set, $C^u_{2,\tau_j}[C^l_{2,\lambda_j}(f)]\to f$ uniformly as $j\to\infty$.*

Now we turn to the proof of Theorem 4.1. It requires some careful calculations following the line of proof for [11, Prop. 3.7] and also some ingredient in [23]. However, certain heavy machinery used in [11] for the more difficult quasi-convex functions can be simplified for convex functions and the results are better as estimates for quasi-convex/rank-one convex functions in [11] cannot be very sharp because of the restrictions along rank-one directions. On the other hand, as the results in [11] apply to the convex case, we just follow the line of their proof. We need the following preparations.

A special case of the combination of [11, Th. 3.1, Lemma 3.2] can be stated as follows.

Lemma 4.1. *Suppose $f : M^{N\times n} \mapsto \mathbb{R}$ is bounded below, continuous and upper semi-differentiable. Assume that for some $p \in [0,\infty)$,*

$$\liminf_{|X|\to\infty} \frac{f(X)}{|X|^p} > 0 \quad \text{and} \quad \limsup_{|X|\to\infty} \frac{f(X)}{|X|^{p+1}} = +\infty. \tag{4.3}$$

Then $Q(f)$ is a C^1 function.

In our present situation, as in $\mathbb{R}^n$, $C(f_\lambda) = Q(f_\lambda)$ [5,17], where $f_\lambda(x) := \lambda|x|^2 - f(x)$ with f the convex function in Theorem 4.1 satisfying (4.1). Since f is convex, for any $u \in \partial f(x)$, where $\partial f(x)$ is the subdifferential [39,21] of f, we have, as $f(x+y) - f(x) - u\cdot y \geqslant 0$, hence

$$\big[\lambda|x+y|^2 - f(x+y)\big] - \big[\lambda|x|^2 - f(x)\big] - [2\lambda x\cdot y - u\cdot y] = \lambda|y|^2 - \big[f(x+y) - f(x) - u\cdot y\big] \leqslant \lambda|y|^2. \tag{4.4}$$

This implies that $f_\lambda(x)$ above is upper semi-differentiable. Furthermore, for $p = 2$ in Lemma 4.1, as $\lambda > C$, where $C > 0$ is the bound given by (4.1), we have

$$\frac{(\lambda - C)|x|^2 - C}{|x|^2} \leqslant \frac{f_\lambda(x)}{|x|^2} \leqslant \frac{(\lambda + C)|x|^2 + C}{|x|^2}, \qquad \frac{(\lambda - C)|x|^2 - C}{|x|^3} \leqslant \frac{f_\lambda(x)}{|x|^2} \leqslant \frac{(\lambda + C)|x|^2 + C}{|x|^3}.$$

Therefore (4.3) holds. By Lemma 4.1, we have $C(f_\lambda) \in C^1(\mathbb{R}^n)$. However, our first aim is to show that

$$C\big(f_\lambda(x+y)\big) - C\big(f_\lambda(x)\big) - DC\big(f_\lambda(x)\big)\cdot y \leqslant \lambda|y|^2, \quad x, y \in \mathbb{R}^n. \tag{4.5}$$

Next we have (see [11,23, Cor. 2.5])

Lemma 4.2. *Suppose $h : B(x_0, r) \mapsto \mathbb{R}$ is convex and $g : B(x_0, r) \mapsto \mathbb{R}$ is upper semi-differentiable at x_0 such that $h \leqslant g$ on $B(x_0, r)$ and $h(x_0) = g(x_0)$. Then h and g are both differentiable at x_0 and $Dh(x_0) = Dg(x_0)$.*

Lemma 4.3. *Under the assumptions of Theorem* 4.1*, inequality* (4.5) *holds.*

Proof. We follow the proof of [11, Prop. 3.7] in the case of $M^{1\times n} = \mathbb{R}^n$, where quasi-convexity and convexity are equivalent.

We may add the positive constant $C_1 + 1 > 0$ to $f_\lambda(x) = \lambda|x|^2 - f(x)$ by letting $g_\lambda = f_\lambda + C_1 + 1$ such that $g_\lambda \geqslant 1$. We now fix x_0 and follow [11] as in their proof of [11, Th. 3.1]. There is a sub-probability measure ν [11, Prop. 3.6] (in our case, ν is a probability measure supported at most $n+1$ points) satisfying

$$g_\lambda = C(g_\lambda), \quad \nu\text{-a.e.}, \qquad Dg_\lambda = DC(g_\lambda) \quad \nu\text{-a.e.},$$
$$DC\big(C(g_\lambda)\big)(x_0) = \int Dg_\lambda\, d\nu, \qquad \int g_\lambda\, d\nu \leqslant C\big(g_\lambda(x_0)\big)$$

and for all y,

$$C\big(g_\lambda(x_0+y)\big) - C\big(g_\lambda(x_0)\big) - DC\big(g_\lambda(x_0)\big)\cdot y \leqslant \int \big[C\big(g_\lambda(x+y)\big) - C\big(g_\lambda(x)\big) - DC\big(g_\lambda(x)\big)\cdot y\big]\, d\nu(x).$$

Now for ν-a.e. x,

$$C\big(g_\lambda(x+y)\big)-C\big(g_\lambda(x)\big)-DC\big(g_\lambda(x)\big)\cdot y\leqslant g_\lambda(x+y)-g_\lambda(x)-Dg_\lambda(x)\cdot y,$$

hence by (4.4) we have, for all y,

$$0\leqslant C\big(g_\lambda(x_0+y)\big)-C\big(g_\lambda(x_0)\big)-DC\big(g_\lambda(x_0)\big)\cdot y\leqslant \lambda|y|^2$$

as $C(g_\lambda(x))$ is convex. Since $C(g_\lambda(\cdot))=C(f_\lambda(\cdot))+C_1+1$, we see that (4.5) holds. □

In order to prove Theorem 4.1 based on Lemma 4.3, we need another preparation. We denote by $\mathrm{lip}(g; B(x_0,r))$ the Lipschitz constant for a Lipschitz function and we define the oscillation of a function g in a set S by $\mathrm{osc}(g; S)=\sup\{|g(x)-g(y)|, x, y\in S\}$. By slightly modifying [11, Lemma 2.2], we have (see also [23])

Lemma 4.4. *If $g : B(x_0, 2r)\mapsto\mathbb{R}$ is convex, then*

$$\mathrm{lip}\big(g, B(x_0,r)\big)\leqslant\frac{\mathrm{osc}(g; B(x_0,2r))}{r}.$$

Note that if g is differentiable in $B(x_0,r)$, we have $|Dg(x)|\leqslant \mathrm{lip}(g, B(x_0,r))\leqslant \mathrm{osc}(g; B(x_0,2r))/r$.

Proof of Theorem 4.1. By definition, $C^u_{2,\lambda}(f(x))=\lambda|x|^2-C(f_\lambda(x))$, where $f_\lambda(x)=\lambda|x|^2-f(x)$. By Lemma 4.3, we see that, for every $x, y\in\mathbb{R}^n$,

$$\begin{aligned}&C^u_{2,\lambda}\big(f(x+y)\big)-C^u_{2,\lambda}\big(f(x)\big)-D\big[C^u_{2,\lambda}\big(f(x)\big)\big]\cdot y\\&\quad=\lambda|y|^2-\big[C\big(f_\lambda(x+y)\big)-C\big(g_\lambda(x)\big)-DC\big(g_\lambda(x)\big)\cdot y\big]\geqslant 0.\end{aligned}$$

Thus $x\mapsto C^u_{2,\lambda}(f(x))$ is convex. Also by Lemma 4.3, we see that

$$\big|C^u_{2,\lambda}\big(f(x+y)\big)-C^u_{2,\lambda}\big(f(x)\big)-D\big[C^u_{2,\lambda}\big(f(x)\big)\big]\cdot y\big|\leqslant\lambda|y|^2. \tag{4.6}$$

Now we fix x and assume that $0<|y|\leqslant r$. Define

$$h(y)=C^u_{2,\lambda}\big(f(x+y)\big)-C^u_{2,\lambda}\big(f(x)\big)-D\big[C^u_{2,\lambda}\big(f(x)\big)\big]\cdot y.$$

Since $y\mapsto h(y)$ is convex, we have

$$\mathrm{lip}\big(h; B(0,r)\big)\leqslant\frac{\mathrm{osc}(h; B(0,2r))}{r},\quad \text{and by (4.6)}\quad \mathrm{osc}\big(h; B(0,2r)\big)\leqslant 8\lambda r^2.$$

Because $Dh(y)$ exists, we have

$$\big|Dh(y)\big|\leqslant\mathrm{lip}\big(h; B\big(0,|y|\big)\big)\leqslant 8\lambda|y|.$$

Since $Dh(y)=DC^u_{2,\lambda}(f(x+y))-DC^u_{2,\lambda}(f(x))$, we see that

$$\big|DC^u_{2,\lambda}\big(f(x+y)\big)-DC^u_{2,\lambda}\big(f(x)\big)\big|\leqslant 8\lambda|y|$$

for any $y\neq 0$. The proof is finished. □

Remark 4.1. If we apply the arguments in the proof of Theorem 4.1 to $C(f_\lambda)$, we can also see that

$$\big|D\big(f_\lambda(x+y)\big)-D\big(f_\lambda(x)\big)\big|\leqslant 8\lambda|y|,\quad x, y\in\mathbb{R}^n.$$

Proof of Corollary 4.1. For each fixed integer $j>0$, by Theorem 2.3 there is some $\lambda_j>0$ such that

$$\big|C^l_{2,\lambda_j}\big(f(x)\big)-f(x)\big|\leqslant\frac{1}{j},\quad x\in\bar{B}(0,j). \tag{4.7}$$

Now we take $\tau_j>\lambda_j$ and consider $C^u_{2,\tau_j}[C^l_{2,\lambda_j}(f(x))]$. By Theorem 2.1(iv),

$$C^u_{2,\tau_j}\big[C^l_{2,\lambda_j}\big(f(x)\big)\big]=C^u_{2,\tau_j+\lambda_j}\big[C\big(f(x)+\lambda_j|x|^2\big)\big]-\lambda_j|x|^2. \tag{4.8}$$

For the fixed λ_j, when τ_j is sufficiently large, we see that

$$\left|C^u_{2,\tau_j}\left[C^l_{2,\lambda_j}\big(f(x)\big)\right] - C^l_{2,\lambda_j}\big(f(x)\big)\right| \leqslant \frac{1}{j}, \quad x \in \bar{B}(0, j). \tag{4.9}$$

By (4.8) and Theorem 4.1, we see that $C^u_{2,\tau_j}[C^l_{2,\lambda_j}(f(\cdot)) \in C^{1,1}(\mathbb{R}^n)$ and is convex with the Lipschitz constant for the gradient at most $8\tau_j + 10\lambda_j$. Combining (4.7) and (4.9), we see that

$$\left|C^u_{2,\tau_j}\left[C^l_{2,\lambda_j}\big(f(x)\big)\right] - f(x)\right| \leqslant \frac{2}{j}, \quad x \in \bar{B}(0, j).$$

The conclusion then follows. $\square$

5. Examples

In this section we give some calculated examples for lower, upper and mixed quadratic transforms. The first set of examples consists of some functions of one variable. We compare the quadratic transforms of the functions with their Moreau envelopes and sometimes with their Lasry–Lions regularization. Then we consider examples of upper transforms for maximum-like functions and lower transforms for the squared distance function to a finite set.

Example 5.1. Consider the signature function $f(x) = \operatorname{sign}(x)$ defined by $\operatorname{sign}(x) = x/|x|$ if $x \neq 0$ and $\operatorname{sign}(x) = 0$ if $x = 0$. We have

$$C^l_{2,\lambda}\big(f(x)\big) = \begin{cases} -1, & x \leqslant 0, \\ 2\sqrt{2\lambda}x - \lambda x^2 - 1, & 0 \leqslant x \leqslant \sqrt{2/\lambda}, \\ 1, & x \geqslant \sqrt{2/\lambda}. \end{cases}$$

$$M_{2\lambda}\big(f(x)\big) = \begin{cases} -1, & x \leqslant 0, \\ \lambda x^2 - 1, & 0 \leqslant x \leqslant \sqrt{2/\lambda}, \\ 1, & x \geqslant \sqrt{2/\lambda}. \end{cases}$$

The modification of the original function for both the lower transform and the Moreau envelope occurs in the same interval $[0, \sqrt{2/\lambda}\,]$. Observe that $C^l_{2,\lambda}(f)$ is locally $C^{1,1}$ near $x = \sqrt{2/\lambda}$ while $M_{2\lambda}(f)$ is and is locally $C^{1,1}$ near $x = 0$. These indicate that $M_{2\lambda}(f)$ smooths the lower part of f at $x = 0$ and connect to the upper half in a non-differentiable manner while $C^l_{2,\lambda}(f)$ starts from the lower half at $x = 0$ with and tries to connect the upper half smoothly.

Example 5.2. Let $f_\alpha(x) = 1/|x|^\alpha$, $x \neq 0$ and $f(0) = +\infty$ with $\alpha > 0$. We can easily derive that

$$C^l_{2,\lambda}\big(f_\alpha(x)\big) = \begin{cases} \lambda^{\frac{\alpha}{2+\alpha}}\left(\frac{\alpha}{2}\right)^{\frac{2}{2+\alpha}} + \left(\frac{2\lambda}{\alpha}\right)^{\frac{\alpha}{2+\alpha}} - \lambda x^2, & |x| \leqslant \left(\frac{\alpha}{2\lambda}\right)^{\frac{1}{2+\alpha}}, \\ \frac{1}{|x|^\alpha}, & |x| \geqslant \left(\frac{\alpha}{2\lambda}\right)^{\frac{1}{2+\alpha}}. \end{cases}$$

Clearly, $C^l_{2,\lambda}(f_\alpha(x))$ modifies f near $x = 0$ by a simple quadratic function and is a tight approximation of $f_\alpha(x)$. On the other hand, the calculation of the Moreau envelope of $f_\alpha(\cdot)$ is non-trivial for any of $\alpha > 0$. Even for $\alpha = 1$, the calculation of the Moreau envelope $F_{2\lambda}(f_1(x) = \inf_{y\in\mathbb{R}}(\lambda(y - x)^2 + f_1(y))$ involves the solutions for the cubic equation $y^3 - xy^2 - 1/(2\lambda) = 0$ and the resulting expression for $F_{2\lambda}(f_1(x))$ is very complicated. In general, $F_{2\lambda}(f_\alpha(x))$ can only be obtained numerically.

To compare with the Moreau envelope, we have to use the very special example $g(x) = -\log|x|$, $x \neq 0$ and $g(0) = +\infty$ for which the calculation of Moreau envelope is not very involved. We have

$$C^l_{2,\lambda}\big(g(x)\big) = \begin{cases} \frac{1}{2}(1 + \log(2\lambda)) - \lambda x^2, & |x| \leqslant 1/(2\lambda)^{1/2}, \\ -\log|x|, & |x| \geqslant 1/(2\lambda)^{1/2}. \end{cases}$$

The effect of $C^l_{2,\lambda}(g(x))$ is again on the points with large second order derivative and it is easy to see that the formula gives a simple approximation of $g(\cdot)$.

The Moreau envelope is given by

$$F_{2\lambda}(x) := \inf_{y\in\mathbb{R}}\big(\lambda(y-x)^2+g(x)\big)$$
$$=\begin{cases}\frac{\lambda}{4}(\sqrt{x^2+2/\lambda}-x)^2-\log\big(\frac{\sqrt{x^2+2/\lambda}-x}{2}\big), & x\geqslant 0,\\ \frac{\lambda}{4}(\sqrt{x^2+2/\lambda}+x)^2-\log\big(\frac{\sqrt{x^2+2/\lambda}+x}{2}\big), & x\leqslant 0.\end{cases}$$

The effect of Moreau envelope is everywhere and is less as obvious to see that it is an approximation of the simple function $-\log|x|$.

Example 5.3. Consider the simple function $f(x)=x^2$, we have $M_{2\lambda}(f(x))=\lambda x^2/(1+\lambda)$. Clearly, $M_{2\lambda}(f(x))$ does not converge uniformly to f in $\mathbb{R}$ and is not a tight approximation. On the other hand, we have, by Theorem 2.3(v) that $C^u_{2,\lambda}(f(x))=f(x)$ for all $x\in\mathbb{R}$ when $\lambda>2$. We see in this simple example that $M_{2\lambda}(f)$ is a much poorer $C^{1,1}$ approximation of the convex function f than $C^u_{2,\lambda}(f)$.

The following example will show that the Moreau envelope and Lasry–Lions regularization are much less predictable than the quadratic transforms and the mixed transforms. Since we are only interested in the approximation and smoothing effects of the transforms upon the original function, we only consider the transforms with large parameters $\lambda>0$ and $\mu>0$.

Example 5.4. Consider $f(x)=\operatorname{dist}(x,\{-1,1\})=\min\{|x-1|,|x+1|\}$, $x\in\mathbb{R}$. We have

$$C^l_{2,\lambda}\big(f(x)\big)=\begin{cases}1-\frac{1}{4\lambda}-\lambda x^2, & |x|\leqslant\frac{1}{2\lambda},\\ f(x), & |x|\geqslant\frac{1}{2\lambda};\ (\lambda\geqslant 1/2);\end{cases}$$
$$C^u_{2,\mu}\big(f(x)\big)=\begin{cases}\mu(|x|-1)^2+\frac{1}{4\mu}, & ||x|-1|\leqslant\frac{1}{2\mu},\\ f(x), & \text{otherwise};\ (\mu\geqslant 1/2).\end{cases}$$

One of the mixed transform for $\lambda>1$ and $\mu>1$ is given by

$$C^u_{2,\mu}\big[C^l_{2,\lambda}\big(\operatorname{dist}\big(x,\{-1,1\}\big)\big)\big]=\begin{cases}1-\frac{1}{4\lambda}-\lambda x^2, & |x|\leqslant\frac{1}{2\lambda},\\ \mu(|x|-1)^2+\frac{1}{4\mu}, & ||x|-1|\leqslant\frac{1}{2\mu},\\ \operatorname{dist}(x,\{-1,1\}), & \text{otherwise}.\end{cases}$$

Due to the locality property of quadratic transforms, we see that in this example $C^u_{2,\mu}[C^l_{2,\lambda}(f(x)]=C^l_{2,\mu}[C^u_{2,\lambda}(f(x)]$ for large $\lambda>0$ and $\mu>0$ due to the fact that the lower and upper transforms act on different part of the graph respectively, upper near non-smooth local convex points -1 and 1 while lower near the concave point 0.

On the other hand, the Moreau–Yosida regularization $M_{2\lambda}(f(x))$ and the Lasry–Lions regularization $M^{2\tau}(M_{2\lambda})$ are given by [1] as (under our notation)

$$M_{2\lambda}\big(f(x)\big)=\begin{cases}1-|x|-\frac{2}{\lambda}, & |x|\leqslant 1-\frac{2}{\lambda},\\ \frac{(|x|^2-1)^2}{4/\lambda}, & 1-\frac{2}{\lambda}\leqslant|x|\leqslant 1+\frac{2}{\lambda},\\ |x|-1-\frac{1}{\lambda}, & |x|\geqslant 1+\frac{2}{\lambda};\\ (2<\lambda).\end{cases}$$

$$M^{2\mu}\big(M_{2\lambda}\big(f(x)\big)\big)=\begin{cases}1-\frac{1}{\lambda}-\frac{|x|^2}{4/\mu}, & |x|\leqslant 2/\mu,\\ 1-|x|-\big(\frac{1}{\lambda}-\frac{1}{\mu}\big), & \frac{2}{\mu}\leqslant|x|\leqslant 1-\big(\frac{2}{\lambda}-\frac{2}{\mu}\big),\\ \frac{(|x|-1)^2}{4(\frac{1}{\lambda}-\frac{1}{\mu})}, & 1-\big(\frac{2}{\lambda}-\frac{2}{\mu}\big)\leqslant|x|\leqslant 1+\big(\frac{2}{\lambda}-\frac{2}{\mu}\big),\\ |x|-1-\big(\frac{1}{\lambda}-\frac{1}{\mu}\big), & |x|\geqslant\big(\frac{2}{\lambda}-\frac{2}{\mu}\big);\\ (0<2/\mu<\min\big\{2/\lambda,\frac{1-2/\lambda}{2}\big\}.\end{cases}$$

From the above explicit calculations, we see that at least for one-dimensional piecewise affine functions, the mixed transform can be easily predicted as to attach quadratic arcs in neighborhoods of points where the function is not differentiable. The graph of the mixed transforms can easily be sketched by hand. We may imagine that for large

$\mu > \lambda > 0$, the behavior of the mixed transform $C^u_{2,\mu}(C^l_{2,\lambda}(f(\cdot))$ on higher-dimensional piecewise affine functions would be similar and the mixed transform equals the original function outside a neighborhood of points where f is not differentiable due to the tightness of the quadratic transforms. On the other hand, the Moreau–Yosida and Lasry–Lions regularization are much less predictable.

Given a subset K of $\mathbb{R}^n$, we write $tK = \{tx, x \in K\}$ for $t \in \mathbb{R}$ and $K + y = \{x + y, x \in K\}$ for $y \in \mathbb{R}^n$. As before, we denote by $C(K)$ the convex hull of K. Recall the definition of the maximum function

$$f(x) = \max_{1\leqslant i\leqslant n} x_i, \quad x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n. \tag{5.1}$$

Note that f is convex, 'monotone increasing' and is of linear growth.

Let $e_1, e_2, \dots, e_n$ be the standard Euclidean basis of $\mathbb{R}^n$. Let $K_n = \{e_1, e_2, \dots, e_n\}$, then

Theorem 5.1. *For every $\lambda > 0$,*

$$C^u_{2,\lambda}\big(f(x)\big) = \lambda|x|^2 - \lambda\,\mathrm{dist}^2\left(x, C\left(\frac{K_n}{2\lambda}\right)\right) + \frac{1}{4\lambda}, \quad x \in \mathbb{R}^n. \tag{5.2}$$

Furthermore, we have the following uniform bounds

$$0 \leqslant C^u_{2,\lambda}\big(f(x)\big) - f(x) \leqslant \frac{1}{2\lambda}, \quad \big|DC^u_{2,\lambda}\big(f(x)\big)\big| \leqslant 1, \quad x \in \mathbb{R}^n. \tag{5.3}$$

Without direct calculations, we can draw the following conclusions for $C^u_{2,\lambda}(f(x))$ from our general theory for quadratic compensated convex transforms.

(i) The quadratic upper transform $C^u_{2,\lambda}(f(x))$ is convex and belongs to $C^{1,1}(\mathbb{R}^n)$ (Theorem 4.1).
(ii) The function $x \mapsto C^u_{2,\lambda}(f(x))$ is monotone increasing (Theorem 2.2) and $f(x) \leqslant C^u_{2,\lambda}(f(x))$ (Theorem 2.1).
(iii) Theorem 2.3 implies that $\lim_{\lambda\to+\infty} C^u_{2,\lambda}(f(\cdot)) = f(\cdot)$ uniformly on any compact set of $\mathbb{R}^n$.

A new feature of Theorem 5.1 is that we have the uniform estimate of approximation of f by its quadratic upper transform (5.2) which is independent of the dimension n. This is in contrast with the classical entropy regularization of the maximum function $G_\epsilon(x) = \epsilon\log(\sum_{i=1}^n \exp(x_i/\epsilon)$ which has the uniform error estimate depending on the dimension n [4,6,13,41,34]: $f(x) \leqslant G_\epsilon(x) \leqslant f(x) + \epsilon\log n$.

Proof of Theorem 5.1. For any $\lambda > 0$, we have

$$\begin{aligned}\lambda|x|^2 - f(x) &= \lambda|x|^2 - \max_{1\leqslant i\leqslant n} x_i = \min_{1\leqslant i\leqslant n}\big(\lambda|x|^2 - x_i\big) = \min_{1\leqslant i\leqslant n}\left[\left(\lambda\sum_{j\neq i}^n x_j^2\right) + \lambda\left(x_i - \frac{1}{2\lambda}\right)^2 - \frac{1}{4\lambda}\right]\\ &= \min_{1\leqslant i\leqslant n}\left(\left|x - \frac{e_i}{2\lambda}\right|^2\right) - \frac{1}{4\lambda} = \lambda\,\mathrm{dist}^2\left(x, \frac{K_n}{2\lambda}\right) - \frac{1}{4\lambda}.\end{aligned} \tag{5.4}$$

Thus firstly we have

$$f(x) = \lambda|x|^2 - \big(\lambda|x|^2 - f(x)\big) = \lambda|x|^2 - \lambda\,\mathrm{dist}^2\left(x, \frac{K_n}{2\lambda}\right) + \frac{1}{4\lambda}. \tag{5.5}$$

Secondly, we see, by (5.4) and Lemma 3.1 that

$$C\big[\lambda|x|^2 - f(x)\big] = C\left[\mathrm{dist}^2\left(x, \frac{K_n}{2\lambda}\right) - \frac{1}{4\lambda}\right] = \lambda\,\mathrm{dist}^2\left(x, C\left(\frac{K_n}{2\lambda}\right)\right) - \frac{1}{4\lambda},$$

so that

$$C^u_{2,\lambda}\big(f(x)\big) = \lambda|x|^2 - C\big[\lambda|x|^2 - f(x)\big] = \lambda|x|^2 - \lambda\,\mathrm{dist}^2\left(x, C\left(\frac{K_n}{2\lambda}\right)\right) + \frac{1}{4\lambda}. \tag{5.6}$$

Thus (5.2) is established. Next we establish estimates (5.3). By (5.5), (5.6) and Lemma 3.1, we first observe that $\operatorname{diam}(K_n) = \sqrt{2}$ and

$$\begin{aligned}
0 &\leqslant C^u_{2,\lambda}\big(f(x)\big) - f(x) \\
&= \left[\lambda|x|^2 - \lambda \operatorname{dist}^2\left(x, C\left(\frac{K_n}{2\lambda}\right)\right) + \frac{1}{4\lambda}\right] - \left[\lambda|x|^2 - \lambda \operatorname{dist}^2\left(x, \frac{K_n}{2\lambda}\right) + \frac{1}{4\lambda}\right]
\end{aligned}$$

$$\lambda\left[\operatorname{dist}^2\left(x, \frac{K_n}{2\lambda}\right) - \operatorname{dist}^2\left(x, C\left(\frac{K_n}{2\lambda}\right)\right)\right] \leqslant \lambda \operatorname{diam}^2\left(\frac{K_n}{2\lambda}\right) = \lambda\left(\frac{\sqrt{2}}{2\lambda}\right)^2 = \frac{2}{\lambda}.$$

Thus the first estimate concerning the error of the approximation is proved. The second estimate on gradient follows from Lemma 3.1 as

$$DC^u_{2,\lambda}\big(f(x)\big) = 2\lambda x - 2\lambda\big(x - P_{C(K_n/(2\lambda))}(x)\big) = 2\lambda P_{C(K_n/(2\lambda))}(x),$$

where $P_{C(K_n/(2\lambda))}(x)$ is the unique nearest point in $C(K_n/(2\lambda))$ such that $\operatorname{dist}(x, C(K_n/(2\lambda))) = |x - P_{C(K_n/(2\lambda))}(x)|$. Thus

$$\big|DC^u_{2,\lambda}\big(f(x)\big)\big| = 2\lambda\big|P_{C(K_n/(2\lambda))}(x)\big| \leqslant 2\lambda\left|C\left(\frac{K_n}{2\lambda}\right)\right| = 1. \qquad \square$$

Next we consider the support function [39] for a given compact set $K \subset \mathbb{R}^n$ defined by $\phi(x) = \max\{y \cdot x),\ y \in K\}$. Note that if $K = K_n$ where K_n is the set of standard Euclidean basis defined in Theorem 5.1, the resulting support function is the maximum function. Since the set K can be non-convex and not necessarily finite now, the smoothing method using (5.4) may lead to very complicated calculations. To derive a simple approximation for ϕ, we perturb ϕ slightly and calculate the quadratic upper transform for the perturbed function. Let $|K| = \sup\{|y|, y \in K\}$, we consider, for $\lambda > 0$

$$\phi_\lambda(x) = \max_{y\in K}\left(y \cdot x - \frac{|y|^2}{4\lambda}\right).$$

Clearly, we have the error estimate

$$0 \leqslant \phi_\lambda(x) - \phi(x) \leqslant \frac{|K|^2}{4\lambda}. \tag{5.7}$$

Now we calculate $C^u_{2,\lambda}(\phi_\lambda(x))$. Firstly we have

$$\lambda|x|^2 - \phi_\lambda(x) = \lambda \min_{y\in K}\left(|x|^2 - \frac{y\cdot x}{\lambda} + \frac{|y|^2}{(2\lambda)^2}\right) = \lambda \min_{y\in K}\left|x - \frac{y}{2\lambda}\right|^2 = \lambda \operatorname{dist}^2\left(x, \frac{K}{2\lambda}\right).$$

Thus $C[\lambda|x|^2 - \phi_\lambda(x)] = \lambda \operatorname{dist}^2(x, C(K/\lambda))$ and

$$C^u_{2,\lambda}\big(\phi_\lambda(x)\big) = \lambda|x|^2 - \lambda \operatorname{dist}^2\left(x, C\left(\frac{K}{\lambda}\right)\right). \tag{5.8}$$

We have

Theorem 5.2. *For $\lambda > 0$, $C^u_{2,\lambda}(\phi_\lambda(x))$ given by* (5.8) *is convex and belongs to $C^{1,1}(\mathbb{R}^n)$ with the Lipschitz constant bounded above by* 8λ. *Furthermore*

$$\big|C^u_{2,\lambda}\big(\phi_\lambda(x)\big) - \phi(x)\big| \leqslant \frac{\operatorname{diam}(K))}{2\lambda} + \frac{|K|^2}{4\lambda}, \quad \big|DC^u_{2,\lambda}\big(\phi_\lambda(x)\big)\big| \leqslant |K|. \tag{5.9}$$

Proof of Theorem 5.2. The error estimate $0 \leqslant C^u_{2,\lambda}(\phi_\lambda(x)) - \phi_\lambda(x) \leqslant \operatorname{diam}^2(K)/(2\lambda)$ follows from a similar calculation as in Theorem 5.1 which, combined with (5.7) implies the required estimate. The convexity and $C^{1,1}$-smoothness of $C^u_{2,\lambda}(\phi_\lambda(x))$ are both direct consequence of Theorem 4.1. $\square$

Remark 5.1. It is interesting to compare the upper transforms of the maximum-like functions obtained in Theorems 5.1 and 5.2 with the Moreau envelope of the same functions. Let $\phi(x)$ be the function in Theorem 5.2, as it is well known and easy to check that $\max_{y\in K} x\cdot y = \max y\in C(K) x\cdot y$, we have

$$M_{2\lambda}\big(\phi(x)\big)=\min_{y\in\mathbb{R}^n}\Big(\max_{y\in C(K)} x\cdot y+\lambda|y-x|^2\Big)=\max_{z\in C(K)}\Big[\min_{y\in\mathbb{R}^n}\big(y\cdot z+\lambda|y-x|^2\big)\Big],$$

$$\max_{z\in C(K)}\left(x\cdot z-\frac{|z|^2}{4\lambda}\right)=\lambda|x|^2-\lambda\min_{z\in C(K)}\left|x-\frac{z}{2\lambda}\right|^2=\lambda|x|^2-\lambda\,\mathrm{dist}^2\left(xC\left(\frac{K}{2\lambda}\right)\right)=C^u_{2,\lambda}\big(\phi_\lambda(x)\big).$$

The exchange of min and max follows from [39, § 37]. If we take ϕ to be the maximum function $f(x)$ in Theorem 5.1, we have $\phi_\lambda(x)=f(x)-1/(4\lambda)$, we have

$$M_{2\lambda}\big(f(x)\big)=C^u_{2,\lambda}\big(f(x)-1/(4\lambda)\big)=C^u_{2,\lambda}\big(f(x)\big)-\frac{1}{4\lambda},\quad\text{hence}\quad C^u_{2,\lambda}\big(f(x)\big)=M_{2\lambda}\big(f(x)\big)+\frac{1}{4\lambda}.$$

Thus the lift of the $C^{1,1}$-approximation $M_{2\lambda}(f(x))$ by $1/(4\lambda)$ gives the tight $C^{1,1}$-approximation $C^u_{2,\lambda}(f(x))$.

From the calculations in the proof of Theorem 3.7 we realize that in general it is difficult to find explicit quadratic lower transforms for the squared-distance function to a general compact set K even if K is finite. In order to overcome this difficulty, at least for a finite set, we modify our quadratic lower transform and derive an explicit $C^{1,1}$ lower approximations of $\mathrm{dist}^2(\cdot,K)$ for finite sets.

Our example is concerned with a smooth approximation of a squared-distance-like function in the form

$$G(x)=\min_{1\leqslant i\leqslant m}\big(|x-x_i|^2+b_i\big),\quad x\in\mathbb{R}^n$$

with $K_m=\{y_i,1\leqslant i\leqslant m\}$ a finite set. Without loss of generality we assume that $b_i\geqslant 0$. When all b_i's are zero, $G(x)$ is exactly the squared distance function to K_m. Note that even in the special case that $G(x)=\mathrm{dist}^2(x,K_m)$, in general the explicit formula for $C^l_{2,\lambda}(\mathrm{dist}^2(x,K_m))$ could be difficult to calculate. However, by modifying the quadratic lower transform and by a dimension reduction method, we can still find an explicit $C^{1,1}$ approximation for $G(\cdot)$.

Now we write $G(x)$ as

$$G(x)=\min_{1\leqslant i\leqslant m}\big(|x|^2-2x\cdot y_i+|x_i|^2+b_i\big)=\min_{1\leqslant i\leqslant m}\left(\left(|x|^2-2x\cdot y_i+\frac{|y_i|^2}{1+\lambda}\right)+\left(\frac{\lambda|y_i|^2}{1+\lambda}b_i\right)\right)$$

for $\lambda>0$ and consider for $(x,t)\in\mathbb{R}^n\times\mathbb{R}^m$ the function

$$F_\lambda(x,t)=\min_{1\leqslant i\leqslant m}\left(\left(|x|^2-2x\cdot y_i+\frac{|y_i|^2}{1+\lambda}\right)-2t_i\right).$$

If we define $\overline{|y|}^2=(|y_1|^2,\dots,|y_m|^2)$ and $b=(b_1,\dots,b_m)\in\mathbb{R}^m$, then

$$G(x)=F\big(x,-\lambda\overline{|y|}^2/\big(2(1+\lambda)\big)-b/2\big).$$

Now we take the anisotropic quadratic lower transform $C^l_{g_\lambda}(F(x,t))$ against

$$g_\lambda(x,t)=\lambda|x|^2+(1+\lambda)|t|^2,\quad (x,t)\in\mathbb{R}^{n+m},$$

we first have

$$\begin{aligned}F_\lambda(x,t)&+g_\lambda(x,t)\\&=\min_{1\leqslant i\leqslant m}\left(\left((1+\lambda)|x|^2-2x\cdot y_i+\frac{|y_i|^2}{1+\lambda}\right)+\left((1+\lambda)|t|^2-2t_i+\frac{1}{1+\lambda}\right)-\frac{1}{1+\lambda}\right)\\&=(1+\lambda)\,\mathrm{dist}^2\left((x,t),\frac{\hat K_m}{1+\lambda}\right)-\frac{1}{1+\lambda},\end{aligned}$$

where $\hat K_m=\{(y_i,e_i),i=1,\dots,m\}$ with $e_i\in\mathbb{R}^m$ the ith unit vector of the standard Euclidean basis of $\mathbb{R}^m$. Now clearly

$$C\big[F_\lambda(x,t)+g_\lambda(x,t)\big]=(1+\lambda)\,\mathrm{dist}^2\left[(x,t),\frac{C(\hat K_m)}{1+\lambda}\right]-\frac{1}{1+\lambda},$$

hence

$$C^l_{g_\lambda}\big(F(x,t)\big)=(1+\lambda)\,\mathrm{dist}^2\left[(x,t),\frac{C(\hat K_m)}{1+\lambda}\right]-g_\lambda(x,t)-\frac{1}{1+\lambda}.$$

Theorem 5.3. *Let*

$$G_\lambda(x)=C^l_{g_\lambda}\left(F\left(x,-\frac{\lambda|\overline{y}|^2}{2(1+\lambda)}-\frac{b}{2}\right)\right),$$

then $G_\lambda\in C^{1,1}(\mathbb{R}^n)$ with the Lipschitz constant of the gradient bounded above by $16(\lambda+1)$ and

$$0\leqslant G(x)-G_\lambda(x)\leqslant\frac{\mathrm{diam}^2(K_m)}{1+\lambda},\quad x\in\mathbb{R}^n,\tag{5.10}$$

$$\big|DG_\lambda(x)\big|\leqslant 2|K_m|,\quad x\in\mathbb{R}^n.\tag{5.11}$$

The Lipschitz constant $16(\lambda+1)$ is a consequence of [11, Th. 5.5]. It also follows from the proof of Theorem 4.1. The error estimate (5.10) is a direct consequence of Lemma 3.1 as

$$\begin{aligned}0&\leqslant F(t,x)-C^l_{g_\lambda}\big(F(x,t)\big)\\&=(1+\lambda)\left(\mathrm{dist}^2\left[(x,t),\frac{K_m}{1+\lambda}\right]-\mathrm{dist}^2\left[(x,t),\frac{C(K_m)}{1+\lambda}\right]\right)\\&\leqslant(1+\lambda)\,\mathrm{diam}^2\left(\frac{K_m}{1+\lambda}\right)=\frac{\mathrm{diam}(K_m)}{1+\lambda}.\end{aligned}\tag{5.12}$$

The gradient estimate (5.11) follows from Lemma 3.1. Let $P_1(x,t)=x$ be the projection from $\mathbb{R}^n\times\mathbb{R}^m$ to $\mathbb{R}^n$ then

$$D_xC^l_{g_\lambda}\big(F(x,t)\big)=2(1+\lambda)\big(x-P_1\big(P_{C(\hat K_m/(1+\lambda))}(x,t)\big)\big)-2(1+\lambda)x=-2P_1\big(P_{C(\hat K_m)}(x,t)\big),$$

where $D_xC^l_{g_\lambda}(F(x,t))$ is the gradient against the x-variable and $P_{C(\hat K_m/(1+\lambda))}(x,t)$ is the nearest point from (x,t) to $C(\hat K_m/(1+\lambda))$. The conclusion then follows. □

Acknowledgement

I wish to thank the referee for several very helpful suggestions.

References

[1] H. Attouch, D. Aze, Approximations and regularization of arbitrary functions in Hilbert spaces by the Lasry–Lions method, Anal. Non-Lin. H. Poincaré Inst. 10 (1993) 289–312.
[2] E. Acerbi, N. Fusco, Semicontinuity problems in the calculus of variations, Arch. Rational Mech. Anal. 86 (1984) 125–145.
[3] H. Attouch, Variational Convergence for Functions and Operators, Applicable Mathematics Series, Pitman, London, 1984.
[4] A. Auslander, Penalty and barrier methods: a unified framework, SIAM J. Optim. 10 (1999) 211–230.
[5] J.M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, Arch. Rational Mech. Anal. 63 (1977) 337–403.
[6] D.P. Bertsekas, Constrained Optimization and Lagrangian Multiplier Methods, Academic Press, 1982.
[7] K. Bhattacharya, N.B. Firoozy, R.D. James, R.V. Kohn, Restrictions on microstructures, Proc. Royal Soc. Edinb. A 124 (1994) 843–878.
[8] 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.
[9] J.M. Ball, R.D. James, Fine phase mixtures as minimizers of energy, Arch. Rational Mech. Anal. 100 (1987) 13–52.
[10] J.M. Ball, R.D. James, Proposed experimental tests of a theory of fine microstructures and the two-well problem, Phil. Royal Soc. Lond. A 338 (1992) 389–450.
[11] J.M. Ball, B. Kirchheim, J. Kristensen, Regularity of quasi-convex envelopes, Calc. Var. Partial Differ. Equ. 11 (2000) 333–359.
[12] H. Blum, A transformation for extracting new descriptors of shape, in: W.W. Dunn (Ed.), Prop. Symp. Models for the Perception of Speech and Visual Form, MIT Press, 1967, pp. 362–380.
[13] A. Ben-Tal, M. Teboulle, A smoothing technique for nondifferentiable optimization problems, in: Optimization, Lect. Notes in Math., vol. 1405, Springer, 1989, pp. 1–11.

[14] H.I. Choi, S.W. Choi, H.P. Moon, Mathematical theory of medial axis transform, Pacific J. Math. 181 (1997) 57–88.
[15] G.M. Crippen, T.F. Havel, Distance Geometry and Molecular Conformation, John Wiley & Sons, 1988.
[16] X. Chen, H. Qi, L. Qi, K.-K. Teo, Smooth convex approximation to the maximum eigenvalue function, J. Global Optim. 30 (2004) 253–270.
[17] B. Dacorogna, Direct Methods in the Calculus of Variations, Springer, 1989.
[18] N.B. Firoozye, Optimal use of the translation method and relaxations of variational problems, Comm. Pure Appl. Math. 44 (1991) 643–678.
[19] Yu. Grabovsky, Bounds and extreme microstructures for two-component composites: a unified treatment based on the translation method, Proc. Royal Soc. Lond. A 452 (1996) 919–944.
[20] Griewank, P.J. Rabier, On the smoothness of convex envelopes, Trans. AMS 322 (1990) 691–709.
[21] J.-B. Hiriart-Urruty, C. Lemaréchal, Fundamentals of Convex Analysis, Springer, 2001.
[22] R.V. Kohn, The relaxation of a double-well energy, Cont. Mech. Therm. 3 (1991) 981–1000.
[23] B. Kirchheim, J. Kristensen, Differentiability of convex envelopes, C. R. Acad. Sci., Paris, Sr. I, Math. 333 (2001) 725–728.
[24] R.V. Kohn, G.W. Milton, On bounding the effective conductivity of anisotropic composites, in: J.L. Ericksen, D. Kinderlehrer, R.V. Kohn, P.L. Lions (Eds.), Homogenization and Effective Moduli of Materials and Media, Springer, New York, 1986, pp. 97–125.
[25] S.R. Lay, Convex Sets and Their Applications, John Wiley & Sons, 1982.
[26] K.A. Lurie, A.V. Cherkaev, Exact estimates of the conductivity of composites formed two isotropically conducting media taken in prescribed proportion, Proc. Royal Soc. Edinb. A 99 (1984) 71–87.
[27] J.-M. Lasry, P.-L. Lions, A remark on regularization in Hilbert spaces, Israel J. Math. 55 (1986) 257–266.
[28] G. Milton, On characterizing the set of possible effective tensors of composites: the variational method and the translation method, Comm. Pure Appl. Math. XLIII (1990) 63–125.
[29] C. Mantegazza, A.C. Mennacci, Hamilton–Jacobi equations and distance functions on Riemannian manifolds, Appl. Math. Optim. 47 (2002) 1–25.
[30] C.B. Morrey Jr., Multiple Integrals in the Calculus of Variations, Springer, 1966.
[31] J.J. Moreau, Proximité dualité dans un espace Hilbertien, Bull. Soc. Math. Fr. 93 (1965) 273–299.
[32] J.J. Moreau, Fonctionnelles convexes, Lecture Notes, Collège de France, 1967.
[33] V. Nesi, Bounds for the effective conductivity of two-dimensional composites made of $n \geqslant 3$ isotropic phases in prescribed volume fraction: the weighted translation method, Proc. Royal Soc. Edinb. A 125 (1995) 1219–1239.
[34] Yu. Nesterov, Smooth minimization of non-smooth functions, Math. Program. Ser. A 103 (2005) 127–152.
[35] A.C. Pipkin, Elastic materials with two preferred states, Q. J. Mech. Appl. Math. 44 (1991) 1–15.
[36] N.M. Patrikalakis, T. Maekawa, Shape interrogation for computer aided design and manufacturing, Springer, 2002.
[37] H.-D. Qi, P. Tseng, Analysis of piecewise smooth functions and almost smooth functions, 21 (2002) 45–66.
[38] R.T. Rockafellar, Minimax theorems and conjugate saddle functions, Math. Scand. 14 (1964) 151–173.
[39] R.T. Rockafellar, Convex Analysis, Princeton Univ. Press, 1966.
[40] J. Serra, Image Analysis and Mathematical Morphology, Academic Press, 1982.
[41] R.L. Sheu, J.Y. Lin, Solving continuous min-max problems by an iterative entropic regularization method, J. Optim. Theory Appl. 121 (2004) 597–612.
[42] V. Šverák, On the problem of two wells, in: Microstructure and Phase Transition, IMA Vol. Math. Appl., vol. 54, Springer, New York, 1993, pp. 183–189.
[43] L. Tartar, Compensated compactness and applications to partial differential equations, in: R.J. Knops (Ed.), Nonlinear Analysis and Mechanics, Heriot-Watt Symp. IV, Pitman, 1979.
[44] L. Tartar, Estimations fine des coefficients homogénéisés, in: P. Krée (Ed.), Ennio de Giorgi's Colloquium, Pitman, 1985, pp. 168–187.
[45] G. Voronoi, Nouvelles applications des parametres continus a la theorie des formes quadratiques, J. Reine Angew. Math. 134 (1908) 198–287.
[46] F.E. Wolter, Cut locus and medial axis in global shape interrogation and representation, ftp://ftp.gdv.uni-hannover.de/www/paper/ReportMIT93.pdf, MIT, Dept. Ocean Engineering, Design Laboratory Memorandum, 1993, no. 92-2.
[47] K. Yosida, Functional Analysis, Springer, 1971.
[48] K. Zhang, Energy minimizers in nonlinear elasticity and the implicit function theorem, Arch. Rational Mech. Anal. 114 (1991) 95–117.
[49] K. Zhang, On some quasi-convex functions with linear growth, J. Convex Anal. 5 (1998) 133–146.
[50] K. Zhang, On various semi-convex hulls in the calculus of variations, Cal. Var. PDEs 6 (1998) 143–160.
[51] K. Zhang, On various semi-convex relaxations of the squared-distance function, Proc. Royal Soc. Edinb. A 129 (1999) 1309–1323.
[52] K. Zhang, A two-well structure and intrinsic mountain pass points, Calc. Var. PDEs 13 (2001) 231–264.
[53] K. Zhang, Maximal extension for linear spaces of real matrices with large rank, Proc. Royal Soc. Edinb. A 131 (2001) 1481–1491.
[54] K. Zhang, An elementary derivation of the generalized Kohn–Strang relaxation formulae, J. Convex Anal. 9 (2002) 269–285.
[55] K. Zhang, On conditions for equality of relaxations in the calculus of variations, J. Nonl. Convex Anal. 3 (2002) 145–154.
[56] K. Zhang, On separation of gradient Young measures, Calc. Var. PDEs 17 (2003) 85–103.