Journal of Convex Analysis
Volume 13 (2006), No. 3+4, 647–685

On the Variational Behaviour of the Subhessians of the Lasry-Lions Envelope*

A. Eberhard
Mathematics Department, RMIT, GPO Box 2476V,
Melbourne, Victoria, Australia 3001
andy.eb@rmit.edu.au

R. Sivakumaran
Mathematics Department, RMIT, GPO Box 2476V,
Melbourne, Victoria, Australia 3001

R. Wenczel
Mathematics Department, RMIT, GPO Box 2476V,
Melbourne, Victoria, Australia 3001

Dedicated to the memory of Simon Fitzpatrick.

Received: February 21, 2005
Revised manuscript received: November 23, 2005

A full analysis is made of the effect that the Lasry-Lions double-envelope has on the subjet and associated subhessians of a prox-bounded, lower-semicontinuous function. This enables the (basic) subdifferential of nonsmooth analysis and the limiting subjet of viscosity-solution theory to be characterized via accumulation points of the first- and second-order derivative information of approximating functions provided by the Lasry-Lions double-envelope.

2000 Mathematics Subject Classification: 49J53, 49J52, 58C20, 49L25

1. Introduction

In this paper we provide an extension of some well-known results which relate an element of the viscosity subdifferential (or subjet) of the infimal convolution (regularisation) of a lower–semicontinuous, prox–bounded function, to an element of the original function at a perturbed base point (see [7, Appendix A], [19], [21]). This kind of result was pivotal to the development of general conditions for the comparison principle for fully-nonlinear elliptic partial differential equations. It has also found application in some areas of nonsmooth analysis where it is used to provide $C^{1,1}$ smooth approximations for the class of "prox–regular" functions ([24], [28], [30], [26] and [11]). We recall that a function is $C^{1,1}$ when its gradient exists and this gradient is also Lipschitz continuous. Motivated in part by the desire to study a $C^{1,1}$ approximation for a very general class of functions we consider the Lasry-Lions double–envelope and some related constructions, instead of the infimal convolution. It is well documented that the Lasry-Lions double–envelope produces a $C^{1,1}$–approximation of a prox–bounded function [22]. The effect this double–envelope has on the second-order (sub–)differential information of the original function is not well

*The research of the first and the last author was supported in part by a ARC Discovery Grant (number DP0451168).

ISSN 0944-6532 / $ 2.50

understood. There have been some attempts to quantify this [8] but in this paper we provide for the first time a full analysis of this problem.

There are some notable attempts to study limits of first-order (sub–)differentials provided by a family of convergent functions. Most of these studies either consider the closed, convex hull of the accumulation points provided by the first-order (sub–)derivatives of a convergent family or restrict consideration to a special class of functions. In [13, 14] the authors study locally Lipschitz functions on both Hilbert and Banach spaces, that are globally minorised by a translation of a negative multiple of the norm squared (i.e. the functions are prox-bounded). In a general Banach space the Clarke subdifferential is shown to be generated by the weak* convex hull of a weak* upper limit of the Clarke subdifferentials of the family of Lasry–Lions double envelopes. This work extended the earlier work of [4, 5] which was based on an interesting geometric approach. In [6] prox-regular, subdifferentially continuous functions on a Hilbert space are studied. A number of results first established in finite dimensions in [24] are shown to hold in infinite dimensions. In the main these concern the $C^{1,1}$ property of the infimal convolution of such functions, the hypomonotonicity of the f–attentive, ε-localization of the subdifferential and the characterisation and single-valuedness of the associated prox-mapping. The work of [20] is most similar to that embarked on here. The upper limit of the G–subdifferential (introduced in [16, 17]), for a uniformly convergent family of functions, is shown to contain the G–subdifferential of the limiting function. In the context of an Asplund space this result is used to show that the limiting Fréchet subdifferential of a lower semi–continuous, prox–bounded function may be characterised via limits of the Fréchet subdifferentials of its infimal convolutions. Although we restrict attention to finite dimensions we study the most general class of functions that is possible to consider: namely the class of lower semi–continuous, prox-bounded functions and consider the differential information provided by their Lasry–Lions approximations. We consider the most general result that is possible in this context, namely: the characterisation of the basic subdifferential (resp. the limiting subhessians) via sets of accumulation points of gradients (resp. Hessians) of the family of Lasry–Lions double envelopes. In the context of finite dimensions the work in this paper not only extends known results concerning the characterisation of the Clarke subdifferential using the Lasry–Lions approximation, to ones involving the characterisation of the (basic) subdifferential but also extend the analysis to the second-order level.

A convenient way to understand the presence of the $C^{1,1}$ property is to invoke a result which was probably first noted by Hiriart–Urruty and Plazanet [15]. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is locally para–convex around x if there exists $r > 0$ such that the function $f(\cdot) + \frac{r}{2}\|\cdot\|^2$ is convex on some neighbourhood of x. Similarly a function is locally para–concave when $-f$ is locally para–convex. In [15] it is noted that a function is locally $C^{1,1}$ precisely when it is simultaneously locally para–convex and para–concave. Consider a prox–bounded function f (i.e. a function globally minorized by some quadratic of the form $\alpha - \frac{r}{2}\|y - x\|^2$ for some $r > 0$ and $y \in \mathbb{R}^n$ and where the quantity $r(f)$ is the infimum over all such constants r). The infimal convolution

$$f_\lambda(x) := \inf_{y \in \mathbb{R}^n} \{f(y) + \frac{1}{2\lambda}\|x - y\|^2\}$$

is globally para–concave in that $x \mapsto f_\lambda(x) - \frac{1}{2\lambda}\|x\|^2$ is a continuous, concave function for

$0 < \lambda < r^{-1}$. Similarly the supremal deconvolution

$$f^{\mu}(x) := \sup_{y \in \mathbb{R}^n} \{f(y) - \frac{1}{2\mu}\|y - x\|^2\}$$

provides a para–convex function in that $x \mapsto f^{\mu}(x) + \frac{1}{2\mu}\|x\|^2$ is globally finite and convex when $-f$ is prox–bounded and $\mu > 0$ is sufficiently small. The combination of these two operations results in an approximation which inherits both properties and hence becomes $C^{1,1}$. One such object is the Lasry-Lions double-envelope: for $0 < \mu < \lambda < \bar{r}(f)^{-1}$ (where the quantity $\bar{r}(f) := \max\{0, r(f)\}$ and $\bar{r}(f)^{-1}$ is called the prox–threshold of f, see [30]) defined as

$$f_{\lambda,\mu}(x) = \sup_{w} \left\{ \inf_{u} \{f(u) + \frac{1}{2\lambda}\|u - w\|^2\} - \frac{1}{2\mu}\|w - x\|^2 \right\}.$$

In this paper we find it convenient to consider a related construction. The proximal hull $h_\lambda f$ of a function f is given by $h_\lambda f(x) = (f_\lambda)^{\lambda}(x)$ and corresponds to the supremum of all quadratics of the form $x \mapsto \alpha - \frac{1}{2\lambda}\|x - w\|^2$ majorized by f. It is a proper lower semi–continuous function when $\lambda < \bar{r}(f)^{-1}$. For $0 < \mu$ and $0 < \lambda < \bar{r}(f)^{-1}$ we define the μ–proximal hull of the infimal convolution f_λ by

$$f_{\lambda|\mu} := h_\mu(f_\lambda).$$

This is well defined irrespective of the relative magnitude of the parameters $\lambda, \mu > 0$. A number of equivalent formulas exist for this object and we refer the reader to [30] for a discussion of these. In particular we have $f_{\lambda|\mu} = (f_{\lambda+\mu})^{\mu} = f_{(\lambda+\mu),\mu}$.

In [21], [19] and [7] versions of the following result was observed. Let $\partial^{2,-} f(x)$ denote the subjet of viscosity-solution theory which consists of the collection of all $(\nabla\varphi(x), \nabla^2\varphi(x))$ where $\varphi \in C^2(\mathbb{R}^n)$ attains $\min_{y \in \mathbb{R}^n}(f(y) - \varphi(y))$ at x. Under only the assumption of $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ being prox–bounded and lower semi–continuous we have (for $0 < \lambda < \bar{r}(f)^{-1}$) that $(p, Q) \in \partial^{2,-} f_\lambda(\bar{x})$ implies

$$(p, Q) \in \partial^{2,-} f(\bar{x} - \lambda p) \quad \text{and} \quad f(\bar{x} - \lambda p) = f_\lambda(\bar{x}) - \frac{\lambda}{2}\|p\|^2. \tag{1}$$

When considering $\partial^{2,-} f_{\lambda|\mu}(x)$ this result is still useful when applied to the inner operation of $f_{\lambda|\mu} = (f_{\lambda+\mu})^{\mu}$. First we are faced with the problem of considering the subjet of the supremal deconvolution of a para–concave function $f_{\lambda+\mu}$. In this paper we give an explicit formula which relates the subjet of the deconvolution of a para–concave function to the subjet of an underlying para–concave function. We may then use these to extend (1) to where we use $f_{\lambda|\mu}$ instead of the infimal convolution of the function. Care must be taken when estimating the size of parameters used to obtained positive results. In particular we require the use of matrices $(\frac{1}{2}Q)_\mu$ that satisfy

$$\inf_{y \in \mathbb{R}^n} \{\langle \frac{1}{2}Qy, y\rangle + \frac{1}{2\mu}\|y - x\|^2\} = \langle (\frac{1}{2}Q)_\mu x, x\rangle.$$

These matrices only exist when $I + \mu Q \in \operatorname{int} \mathcal{P}(n)$, where $\mathcal{P}(n)$ denote the cone of positive–semidefinite $n \times n$ matrices (and of course $\operatorname{int}\mathcal{P}(n)$ corresponds to the positive–definite forms).

Amongst a number of other results we show the following: Let $0 < \mu$, λ where $\lambda+\mu < \frac{1}{2\bar{r}(f)}$ ($\bar{r}(f)$ the prox–threshold of f). Then, if $(p, Q) \in \partial^{2,-} f_{\lambda|\mu}(x)$ such that $I + \mu Q \in \operatorname{int} \mathcal{P}(n)$, then

$$\left(p, 2\left(\frac{1}{2}Q\right)_{\mu}\right) \in \partial^{2,-} f(x - \lambda p) \quad \text{with } f_{\lambda|\mu}(x) = f(x - \lambda p) + \frac{\lambda}{2}\|p\|^2.$$

This enables the proof of the following two very striking expressions, the first for the (basic) subdifferential of nonsmooth analysis [30];

$$\partial f(x) = \limsup \left\{\nabla f_{\lambda|\mu}(x') \mid x' \to x,\ f_{\lambda|\mu}(x') \to f(x) \text{ and } (\lambda, \mu) \downarrow (0, 0),\ \lambda,\mu > 0\right\}$$

and the second for the limiting subjet of viscosity-solution theory [7] and nonsmooth analysis [23], [18] and [9];

$$\underline{\partial}^2 f(x) = \limsup \Big\{(\nabla f_{\lambda|\mu}(x'), Q) \mid x' \in S_2(f_{\lambda|\mu}),\ x' \to x,\ f_{\lambda|\mu}(x') \to f(x); \\ Q \leq_{\mathcal{P}(n)} \nabla^2 f_{\lambda|\mu}(x') \text{ and } (\lambda, \mu) \downarrow (0, 0),\ \lambda,\mu > 0\Big\},$$

where $Q \leq_{\mathcal{P}(n)} A$ means $A - Q \in \mathcal{P}(n)$ and $S_2(f_{\lambda|\mu})$ denotes the points of second-order differentiability of $f_{\lambda|\mu}$ (which according to a classical result of Aleksandrov is a set of full Lebesgue measure).

2. Preliminaries

In this section we provide the relevant concepts and notation taken from variational and nonsmooth analysis [30] that we require in this paper. Readers familiar with the book [30] and the concept of viscosity subdifferentials [7], [18] and [23] may skip this section only to return to consult definitions for relevant notation. As usual we denote the gradient of a smooth function $\varphi : \mathbb{R}^n \to \mathbb{R}$ at a point $\bar{x}$ by $\nabla f(\bar{x})$ and its Hessian at $\bar{x}$ by $\nabla^2 f(\bar{x})$. When $\nabla f(x)$ exists for $x \in \Omega$ and $x \mapsto \nabla f(x)$ is Lipschitz continuous we will say f is $C^{1,1}$ on the set Ω (or $f \in C^{1,1}(\Omega)$). Endow $\mathbb{R}^n$ with a norm $\|\cdot\|$ and denote the open ball at $\bar{x}$ of radius $\delta > 0$ by $B_\delta(\bar{x}) := \{y \in \mathbb{R}^n \mid \|x - \bar{x}\| < \delta\}$. We deal exclusively with extended–real–valued functions $f : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ on a space of finite dimension n. Denote the vector space of all real symmetric matrices of dimension $n \times n$ by $\mathcal{S}(n)$ and endow it with the Frobenius inner product $\langle Q, M\rangle := \operatorname{trace} M^t Q$ for any $Q, M \in \mathcal{S}(n)$. We call matrices of the form $xx^t \in \mathcal{S}(n)$ rank–1 matrices and conveniently $\langle Q, xx^t\rangle = x^t Q x$ which we refer to as quadratic forms. A matrix is positive–semidefinite when $x^t Q x \geq 0$ for all x and we denote this conic subset of $\mathcal{S}(n)$ by $\mathcal{P}(n)$. Clearly $\operatorname{int} \mathcal{P}(n)$ corresponds to all positive–definite matrices. We will often have the need to consider certain special convex subsets of $\mathcal{S}(n)$. Given a convex subset $\mathcal{A} \subseteq \mathcal{S}(n)$ with $\mathcal{A} - \mathcal{P}(n) \subseteq \mathcal{A}$ we denote the (symmetric) rank–1 support by $q(\mathcal{A})(u) := \sup\{\langle Q, uu^t\rangle \mid Q \in \mathcal{A}\}$. The subset $\mathcal{A}$ is a called a rank–1 representer when $\mathcal{A} = \{Q \in \mathcal{P}(n) \mid \langle Q, uu^t\rangle \leq q(\mathcal{A})(u) \text{ for all } u\}$. There now exists an extensive literature regarding the properties of these sets [9], [10], [11] and [12]. The class of functions $q : \mathbb{R}^n \to \mathbb{R}^n$ that are generated as rank–1 supports as just described are characterized by the four properties (see [9]): properness (i.e. not identically $+\infty$ and never $-\infty$), lower semi–continuity, evenness (i.e. $q(u) = q(-u)$) and positive homogeneity degree 2 (i.e. $q(tu) = t^2 q(u)$ for $t > 0$). When these properties hold there exists a rank–1 representer $\mathcal{A} \subseteq \mathcal{S}(n)$ such that $q(u) = q(\mathcal{A})(u)$ for all u. Any regularization operation that preserves these four properties (such as the infimal convolution) can consequently be viewed as mapping rank–1 representers onto rank–1 representers. This fact is used frequently in this paper.

Definition 2.1. For a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, we say $x \in S_2(f)$ if $f(x)$ is finite and $\exists \delta > 0$ such that for all $y \in B_\delta(x)$

$$f(y) = f(x) + \langle \nabla f(x), y - x\rangle + \frac{1}{2}\langle \nabla^2 f(x)(y-x), y-x\rangle + o(\|y-x\|^2)$$

where $o(\cdot)$ is the usual small–order notation.

As is usual for $f : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ we have epi$f := \{(x,\alpha) \mid f(x) \le \alpha\}$. By $x_v \to^f x$ we mean $x_v \to x$ along with $f(x_v) \to f(x)$.

Definition 2.2. Let Ω be an open subset of $\mathbb{R}^n$.

1. A function $f : \Omega \to \overline{\mathbb{R}}$ is said to be twice sub-differentiable (or possess a subjet) at x if $f(x)$ is finite and

$$\partial^{2,-} f(x) = \{(\nabla\varphi(x), \nabla^2\varphi(x)) : f - \varphi \text{ has a local minimum at } x \text{ with } \varphi \in \mathcal{C}^2(\mathbb{R}^n)\} \ne \emptyset.$$

   We call the collection $\partial^{2,-} f(x,p) := \{Q \in \mathcal{S}(n) \mid (p,Q) \in \partial^{2,-} f(x)\}$ the subhessians of f at (x,p).
   Similarly f is said to be twice super-differentiable (or possess a superjet) at x if $f(x)$ is finite and

$$\partial^{2,+} f(x) = \{(D\varphi(x), D^2\varphi(x)) : f - \varphi \text{ has a local maximum at } x \text{ with } \varphi \in \mathcal{C}^2(\mathbb{R}^n)\} \ne \emptyset.$$

   We call the collection $\partial^{2,+} f(x,p) := \{Q \in \mathcal{S}(n) \mid (p,Q) \in \partial^{2,+} f(x)\}$ the superhessians of f at (x,p).
2. The limiting subjet (superjet) of f at x is defined to be respectively;

$$\underline{\partial}^2 f(x) = \limsup_{u \to^f x} \partial^{2,-} f(u) \quad \text{and} \quad \overline{\partial}^2 f(x) = \limsup_{u \to^f x} \partial^{2,+} f(u).$$

   Denote by $\underline{\partial}^2 f(x,p) = \{Q \in \mathcal{S}(n) \mid (p,Q) \in \underline{\partial}^2 f(x)\}$ the limiting subhessians of f and $\overline{\partial}^2 f(x,p) = \{Q \in \mathcal{S}(n) \mid (p,Q) \in \overline{\partial}^2 f(x)\}$ are called limiting superhessians of f.
3. $$\overline{D}^2 f(\bar{x},p) = \{Q \in \mathcal{S}(n) \mid Q = \lim_{n\to\infty} \nabla^2 f(x_n) \text{ where } \{x_n\} \subseteq S_2(f),\ x_n \to^f \bar{x} \text{ and } \nabla f(x_n) \to p\}.$$

The connection that the limiting subhessians have to limiting Hessians is important. The classical Aleksandrov theorem plays an important role in that it connects the existence of the limiting Hessian to convexity/concavity properties (and hence to generalized convexity). We state this theorem in the form provided by Rockafellar and Wets in [30, Theorem 13.51], (where we have used the equivalence between the so–called lower C^2 functions and locally para–convex functions). Recall that a function is locally para–convex around $\bar{x}$ if there exists a neighbourhood $B_\delta(\bar{x})$ and a constant $\lambda > 0$ such that $f(\cdot) + \frac{1}{2\lambda}\|\cdot\|^2$ is convex relative to $B_\delta(\bar{x})$.

Theorem 2.3. *Any locally para–convex function f on an open set O is twice differentiable almost everywhere in the sense that there exists a set N of measure zero and at all $\bar{x} \in O \cap N^c$ we have f differentiable at $\bar{x}$ and ∇f is differentiable at $\bar{x}$ relative to the domain $S_1(f)$ of existence of ∇f and there is a square symmetric matrix, denoted by $\nabla^2 f(\bar{x})$, such that*

$$\nabla f(x) = \nabla f(\bar{x}) + \nabla^2 f(\bar{x})(x - \bar{x}) + o(\|x - \bar{x}\|^2) \quad \text{for } x \in S_1(f).$$

Aleksandrov's theorem was originally stated in terms of finite convex functions and the existence of a quadratic (Taylor) expansion almost everywhere. This form is a consequence of [30, Corollary 13.42] which, for completeness, we stated next in a form sufficient for our purposes (recall that locally para–convex functions are examples of the so-called "prox–regular" functions).

Proposition 2.4. *If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is differentiable at $\bar{x}$ and locally para–convex around $\bar{x}$, then the following properties are equivalent:*

1. *f is twice differentiable at $\bar{x}$ in the sense that there is a square symmetric matrix, denoted by $\nabla^2 f(\bar{x})$, such that*

$$\nabla f(x) = \nabla f(\bar{x}) + \nabla^2 f(\bar{x})(x - \bar{x}) + o(\|x - \bar{x}\|^2) \quad \text{for all } x \in S_1(f);$$

2. *f has a quadratic (Taylor) expansion at $\bar{x}$, in the sense that there exists a square symmetric matrix, denoted by $\nabla^2 f(\bar{x})$, such that*

$$f(x) = f(\bar{x}) + \langle \nabla f(\bar{x}), x - \bar{x} \rangle + \frac{1}{2}\langle x - \bar{x}, \nabla^2 f(\bar{x})(x - \bar{x}) \rangle + o(\|x - \bar{x}\|^2).$$

We now define some fundamental notions of first–order sub–differentiability used in non-smooth analysis.

Definition 2.5.

1. A vector $y \in \mathbb{R}^n$ is called a proximal sub-gradient to f at $\overline{x}$ if $f(\overline{x})$ is finite and for some $c > 0$

$$f(x) \geq f(\overline{x}) + \langle y, x - \overline{x} \rangle - \frac{c}{2}\|x - \overline{x}\|^2$$

in a neighbourhood of $\overline{x}$. The set of all proximal sub-gradients to f at $\overline{x}$ is denoted $\partial_p f(\overline{x})$.

2. The limiting subdifferential is given by

$$\partial f(x) = \limsup_{x' \to^f x} \partial_p f(x') := \{\lim_{v \to \infty} z_v \mid z_v \in \partial_p f(x_v),\, x_v \to^f x\}.$$

Denote by $S_p(f) = \{x \in \mathbb{R}^n \mid \partial^{2,-} f(x) \neq \emptyset\}$.

Definition 2.6. Let $\{f, f^v : \mathbb{R}^n \to \overline{\mathbb{R}}, v \in W\}$ be a family of proper extended-real-valued functions, where W is a neighbourhood of w (in some topological space). Then the lower epi-limit $e\text{-li}_{v\to w} f^v$ is the function having as its epi-graph the outer limit of the sequence of sets epi f^v:

$$\text{epi}\,(e\text{-li}_{v\to w} f^v) := \limsup_{v\to w}(\text{epi} f^v).$$

The upper epi-limit $e\text{-ls}_{v\to w} f^v$ is the function having as its epigraph the inner limit of sets epif^v:

$$\text{epi}\,(e\text{-ls}_{v\to w} f^v) := \liminf_{v\to w}(\text{epi}\, f^v).$$

When these two functions are equal, the epi-limit function $e\text{-lim}_{v\to w} f^v$ is said to exist. In this case the sequence f^v is said to epi-converge to f.

Clearly as $e\text{-li}_{v\to w} f^v(x) \le e\text{-ls}_{v\to w} f^v(x)$ we have epi-convergence of f^v to f occuring when $e\text{-ls}_{v\to w} f^v(x) \le f(x)$ and $f(x) \le e\text{-li}_{v\to w} f^v(x)$ for all x. The upper and lower epi-limits of the sequence f^v may also be defined via composite limits (see [30]). In particular

$$\begin{aligned} e\text{-ls}_{v\to w} f^v(x) &= \sup_{\delta>0} \limsup_{v\to w} \inf_{x'\in B_\delta(x)} f^v(x') := \limsup_{v\to w} \inf_{x'\to x} f^v(x') \\ \text{and}\quad e\text{-li}_{v\to w} f^v(x) &= \sup_{\delta>0} \liminf_{v\to w} \inf_{x'\in B_\delta(x)} f^v(x') := \liminf_{v\to w} \inf_{x'\to x} f^v(x') \\ &\equiv \liminf_{\substack{v\to w \\ x'\to x}} f^v(x') \end{aligned}$$

Remark 2.7. In [9] it is shown that the subhessian is a rank–1 representer and rank–1 support to the subhessian could be characterized as a directional derivative i.e.

$$q(\partial^{2,-} f(\bar{x}, p))(u) = \liminf_{\substack{t\to 0 \\ u'\to u}} \Delta_2 f(\bar{x}, t, p, u') := f''_s(\bar{x}, p, u), \tag{2}$$

where $\Delta_2 f(\bar{x}, t, p, u) := \frac{2}{t^2}(f(\bar{x}+tu) - f(\bar{x}) - t\langle p, u\rangle)$. It also shown in [9] that $f''_s(\bar{x}, p, u) = \min\{f''_-(\bar{x}, p, u), f''_-(\bar{x}, p, -u)\}$, where

$$f''_-(\bar{x}, p, u) := \liminf_{\substack{t\downarrow 0 \\ u'\to u}} \Delta_2 f(\bar{x}, t, p, u').$$

We say $\{f^v\}$ is epi–lower semi–continuous at x with respect to f if $e\text{-li}_{v\to w} f^v(x) \ge f(x)$ and epi–upper semi–continuous at x with respect to f if $e\text{-ls}_{v\to w} f^v(x) \le f(x)$. For a family of indicator functions $\{\psi_{C^v}\}_{v\in W}$ we have epi–lower (epi–upper) semi–continuity at all x if and only if $\limsup_{v\to w} C^v \subseteq C$ ($\liminf_{v\to w} C^v \supseteq C$).

Remark 2.8. We may also define hypo $f :=\{(x, \alpha) : \alpha \le f(x)\}$ and

$$\begin{aligned} \text{hypo}\,(h\text{-li}_{v\to w} f^v) &:= \liminf_{v\to w}(\text{hypo}\, f^v) \\ \text{hypo}\,(h\text{-ls}_{v\to w} f^v) &:= \limsup_{v\to w}(\text{hypo}\, f^v) \qquad \text{etc.} \end{aligned}$$

Clearly

$$\begin{aligned} h\text{-ls}_{v\to w} f^v &= -\,(e\text{-li}_{v\to w}(-f^v)) \\ \text{and}\quad h\text{-li}_{v\to w} f^v &= -\,(e\text{-ls}_{v\to w}(-f^v)). \end{aligned}$$

The next result may be found in [18] and will be used latter.

Proposition 2.9 ([18]). *If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is proper and lower semi–continuous then we have*

$$\overline{D}^2 f(\bar{x}, p) - \mathcal{P}(n) \subseteq \underline{\partial}^2 f(\bar{x}, p). \tag{3}$$

If we assume in addition that f is continuous and para–concave around $\bar{x}$ then equality holds in (3).

3. Some Basic Properties of the Infimal Convolutions

In this section we survey some basic properties of the infimal convolution which are used repeatedly in the paper. A number of these properties are discussed in [30] to which we refer the reader for proofs when it is possible to do so. Proofs are provided when a suitable prior reference is lacking.

Definition 3.1. For any function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, the function $f^* : \mathbb{R}^n \to \overline{\mathbb{R}}$ defined by

$$f^*(u) := \sup_x \{\langle u, x\rangle - f(x)\}$$

is conjugate to f, while the function $f^{**} = (f^*)^*$ defined by

$$f^{**}(x) := \sup_u \{\langle u, x\rangle - f^*(u)\}.$$

Denote the infimal convolution (or regularization) of f by

$$f_\lambda(x) := \inf_{u\in\mathbb{R}^n} \left(f(u) + \frac{1}{2\lambda}\|x-u\|^2 \right) = \frac{1}{2\lambda}\|x\|^2 - \left(f(\cdot) + \frac{1}{2\lambda}\|\cdot\|^2 \right)^* (\frac{x}{\lambda}),$$

where $\|x\|$ denotes the Euclidean norm on $\mathbb{R}^n$. Also define

$$P_\lambda(x) = \operatorname{argmin}\left\{ f(\cdot) + \frac{1}{2\lambda}\|x - \cdot\|^2 \right\}.$$

We will drive $\lambda \to 0$ to approximate f, unlike other studies which use a kernel $\frac{\lambda}{2}\|x-u\|^2$ while driving $\lambda \to \infty$. In recent years the infimal convolution has been applied to the study of the differentiability of convex and nonconvex functions. This may be found in the work of Poliquin and Rockafellar [24], [25], Penot [23], Eberhard, Nyblom and Ralph, [9], [10] and many other authors.

In [30] the concept of "prox–bounded" is used which is equivalent to $f + \frac{r}{2}\|\cdot\|^2$ being bounded below. This is clearly the same as the assumption of f being quadratically minorized (by a quadratic of the form $\alpha - \frac{r}{2}\|\cdot\|^2$). Thus a sufficient condition for $f_\lambda > -\infty$ is $\lambda < (\max\{0, r\})^{-1}$ (and hence $P_\lambda(x) \neq \emptyset$). The infimal of all such r is denoted by $r(f)$. It is possible for $r(f) < 0$ and so we place $\bar{r}(f) := \max\{r(f), 0\}$ interpreting $1/0 = +\infty$ and $\lambda_f := (\bar{r}(f))^{-1}$ is called the proximal threshold for f. Thus when $r(f) < 0$ we have $f_\lambda > -\infty$ for all $\lambda > 0$.

We say a sequence $\{f^v\}_{v\in\mathbb{N}}$ of functions on $\mathbb{R}^n$ is eventually prox–bounded if there exists $\lambda > 0$ such that $\liminf_v f^v_\lambda(x) > -\infty$ for some x. We can define the threshold of eventual prox–boundedness as λ_{f^v} where λ_{f^v} is the supremum of all λ such that $\liminf_v f^v_\lambda(x) > -\infty$ for some x. Similarly we may place $\bar{r}(\{f^v\}) = (\lambda_{f^v})^{-1}$.

In [7] the following was observed. Under only the assumption of $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ being quadratically minorized and lower semi–continuous (see [9] for this version of the result) we have (for $0 < \lambda < \frac{1}{\bar{r}(f)} = \lambda_f$ the prox–threshold of f)

$$(p, Q) \in \partial^{2,-} f_\lambda(\bar{x}) \text{ implies both } (p, Q) \in \partial^{2,-} f(\bar{x} - \lambda p) \quad \text{and } f(\bar{x} - \lambda p) = f_\lambda(\bar{x}) - \frac{\lambda}{2}\|p\|^2. \tag{4}$$

We now consider the effect of adding a linear and a quadratic function before taking the infimal convolution. We provide a proof as we are lacking a prior reference.

Lemma 3.2. *For any function*

$$(f - \langle p, \cdot \rangle)_\lambda (x) = f_\lambda(x + \lambda p) - \langle p, x \rangle - \frac{\lambda}{2} \|p\|^2 . \tag{5}$$

Thus $f(\bar{x}) + (\lambda/2)\|\bar{p}\|^2 = f_\lambda(\bar{x} + \lambda\bar{p})$ *implies* $(f - \langle \bar{p}, \cdot \rangle)_\lambda (\bar{x}) = f(\bar{x}) - \langle \bar{p}, \bar{x} \rangle$, *in turn implying* $(\bar{p}, -\frac{1}{\lambda} I) \in \partial^{2,-} f(\bar{x})$. *In particular we have* $\bar{p} \in \partial_p f(\bar{x})$.

Proof. By direct calculation

$$\begin{aligned}
(f - \langle p, \cdot \rangle)_\lambda (x) &= \inf_w \left\{ f(w) - \langle p, w \rangle + \frac{1}{2\lambda}\|w - x\|^2 \right\} \\
&= \inf_w \left\{ f(w) + \frac{1}{2\lambda} \left(\|\lambda p\|^2 - 2\langle \lambda p, w - x \rangle + \|w - x\|^2 \right) \right\} \\
&\qquad -\langle p, x \rangle - \frac{\lambda \|p\|^2}{2} \\
&= \inf_w \left\{ f(w) + \frac{1}{2\lambda}\|w - (x + \lambda p)\|^2 \right\} - \langle p, x \rangle - \frac{\lambda \|p\|^2}{2} \\
&= f_\lambda(x + \lambda p) - \langle p, x \rangle - \frac{\lambda}{2} \|p\|^2 .
\end{aligned}$$

Now suppose $f(\bar{x}) + (\lambda/2)\|\bar{p}\|^2 = f_\lambda(\bar{x} + \lambda\bar{p})$. We deduce from (5) that

$$\begin{aligned}
f_\lambda(\bar{x} + \lambda\bar{p}) &= (f - \langle \bar{p}, \cdot \rangle)_\lambda (\bar{x}) + \langle \bar{p}, \bar{x} \rangle + \frac{\lambda}{2} \|\bar{p}\|^2 \\
&= f(\bar{x}) + (\lambda/2)\|\bar{p}\|^2
\end{aligned}$$

and so $(f - \langle \bar{p}, \cdot \rangle)_\lambda (\bar{x}) = f(\bar{x}) - \langle \bar{p}, \bar{x} \rangle$. By the definition of the infimal convolution we have then

$$f(\bar{x}) - \langle \bar{p}, \bar{x} \rangle \leq f(w) - \langle \bar{p}, w \rangle + \frac{1}{2\lambda} \|\bar{x} - w\|^2$$

$$\text{or} \quad f(w) \geq f(\bar{x}) + \langle \bar{p}, w - \bar{x} \rangle - \frac{1}{2\lambda} \|\bar{x} - w\|^2$$

and so $\bar{p} \in \partial_p f(\bar{x})$. □

We now consider the effect of adding a quadratic. We provide a proof as we are lacking a prior reference.

Proposition 3.3. *Suppose $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be lower semi–continuous and prox–bounded with prox–threshold $(r(f))^{-1}$.*

1. *If $r > r(f)$ then $h_r(x) := f(x) + \frac{r}{2}\|x - \bar{x}\|^2$ has $r(h_r) \le 0$, $h_r(\bar{x}) = f(\bar{x})$ and*
$$Q \in \partial^{2,-} f(\bar{x}, 0) \quad \text{implies} \quad Q + rI \in \partial^{2,-}(f + \frac{r}{2}\| \cdot - \bar{x}\|^2)(\bar{x}, 0).$$
2. *We have*
$$\begin{aligned} &\left(f(\cdot) + \frac{r}{2}\| \cdot - \bar{x}\|^2\right)_\lambda (x) \\ = \ & f_{\frac{\lambda}{1+\lambda r}}\left(\frac{x + \lambda r \bar{x}}{1 + \lambda r}\right) - \frac{1}{2\lambda}\left(\frac{1}{1 + \lambda r}\|x + \lambda r \bar{x}\|^2 - \|x\|^2\right) + \frac{r}{2}\|\bar{x}\|^2. \end{aligned} \tag{6}$$
3. *If $r(f) > 0$ we may take $r = r(f)$ in part 1.*

Proof. First take $r > r(f)$ and $0 < \varepsilon < r - r(f)$. As $r - \varepsilon > r(f)$ there exists β such that $f(x) \ge \beta - \frac{r-\varepsilon}{2}\|x\|^2$ for all x. Let $\alpha := \beta + \left(\frac{r}{2} - \frac{2r^2}{\varepsilon}\right)\|\bar{x}\|^2$ then consider

$$\begin{aligned} \alpha - \frac{r}{2}\|x - \bar{x}\|^2 &= \alpha - \frac{r}{2}\|\bar{x}\|^2 - \frac{r-\varepsilon}{2}\|x\|^2 + r\langle x, \bar{x}\rangle - \frac{\varepsilon}{2}\|x\|^2 \\ &\le \left(\alpha - \frac{r}{2}\|\bar{x}\|^2\right) - \frac{r-\varepsilon}{2}\|x\|^2 + \|x\|\left(r\|\bar{x}\| - \frac{\varepsilon}{2}\|x\|\right). \end{aligned} \tag{7}$$

If $\frac{2r}{\varepsilon}\|\bar{x}\| < \|x\|$ then $\left(r\|\bar{x}\| - \frac{\varepsilon}{2}\|x\|\right) < 0$ and we may drop the last term in (7). When $\|x\| \le \frac{2r}{\varepsilon}\|\bar{x}\|$ we have $\|x\|\left(r\|\bar{x}\| - \frac{\varepsilon}{2}\|x\|\right) \le \frac{2r^2}{\varepsilon}\|\bar{x}\|^2$. Thus

$$\alpha - \frac{r}{2}\|x - \bar{x}\|^2 \le \left(\alpha - \frac{r}{2}\|\bar{x}\|^2 + \frac{2r^2}{\varepsilon}\|\bar{x}\|^2\right) - \frac{r-\varepsilon}{2}\|x\|^2 = \beta - \frac{r-\varepsilon}{2}\|x\|^2 \le f(x)$$

for all x. Then it follows that

$$h_r(x) := f(x) + \frac{r}{2}\|x - \bar{x}\|^2 \ge \alpha.$$

Thus for all $r > r(f)$ we have $(h_r)_\lambda > -\infty$ for all $\lambda > 0$. Thus $r(h_r) \le 0$. The other statements in 1 follow easily from definitions.

Consider

$$\begin{aligned} &\left(f(\cdot) + \frac{r}{2}\| \cdot - \bar{x}\|^2 + \frac{1}{2\lambda}\| \cdot \|^2\right)^* \left(\frac{x}{\lambda}\right) \\ = \ & \sup_w \{\langle w, \frac{x}{\lambda}\rangle + \langle w, r\bar{x}\rangle - f(w) - \frac{1}{2}(r + \frac{1}{\lambda})\|w\|^2\} - \frac{r}{2}\|\bar{x}\|^2 \\ = \ & \sup_w \{\langle w, \frac{x}{\lambda} + r\bar{x}\rangle - \left(\frac{1}{2}(\frac{1 + r\lambda}{\lambda})\|w\|^2 + f(w))\right)\} - \frac{r}{2}\|\bar{x}\|^2 \\ = \ & \sup_w \{\langle w, \frac{x + \lambda r \bar{x}}{\lambda}\rangle - \left(\frac{1}{2}(\frac{1}{\frac{\lambda}{1+r\lambda}})\|w\|^2 + f(w))\right)\} - \frac{r}{2}\|\bar{x}\|^2 \\ = \ & \left(\frac{1}{2}(\frac{1}{\frac{\lambda}{1+r\lambda}})\| \cdot \|^2 + f(\cdot)\right)^* \left(\frac{\frac{x+\lambda r\bar{x}}{1+\lambda r}}{\frac{\lambda}{1+\lambda r}}\right) - \frac{r}{2}\|\bar{x}\|^2. \end{aligned} \tag{8}$$

Thus we have

$$\begin{aligned} f_{(\frac{\lambda}{1+\lambda r})}\left(\frac{x+\lambda r\bar{x}}{1+\lambda r}\right) &= \frac{1}{2(\frac{\lambda}{1+\lambda r})}\|\frac{x+\lambda r\bar{x}}{1+\lambda r}\|^2 - \left(f(\cdot)+\frac{1}{2(\frac{\lambda}{1+\lambda r})}\|\cdot\|^2\right)^*\left(\frac{x+\lambda r\bar{x}}{1+\lambda r}\right) \\ &= \frac{1}{2\lambda}\left(\frac{1}{1+\lambda r}\|x+\lambda r\bar{x}\|^2 - \|x\|^2\right) \\ &\quad +\frac{1}{2\lambda}\|x\|^2 - \left(f(\cdot)+\frac{r}{2}\|\cdot-\bar{x}\|^2+\frac{1}{2\lambda}\|\cdot\|^2\right)^*\left(\frac{x}{\lambda}\right) - \frac{r}{2}\|\bar{x}\|^2 \\ &= \frac{1}{2\lambda}\left(\frac{1}{1+\lambda r}\|x+\lambda r\bar{x}\|^2 - \|x\|^2\right) - \frac{r}{2}\|\bar{x}\|^2 \\ &\quad +(f(\cdot)+\frac{r}{2}\|\cdot-\bar{x}\|^2)_\lambda(x). \end{aligned} \tag{9}$$

From (9) we see that $(f(\cdot)+\frac{r(f)}{2}\|\cdot-\bar{x}\|^2)_\lambda(x) > -\infty$ whenever $f_{(\frac{\lambda}{1+\lambda r})}\left(\frac{x+\lambda r\bar{x}}{1+\lambda r}\right) > -\infty$ which only occurs when $\frac{\lambda}{1+\lambda r(f)} < \bar{r}(f)^{-1}$. But when $r(f) = \bar{r}(f) > 0$ this is equivalently expressed as $\frac{1}{\frac{1}{\lambda}+r(f)} < \frac{1}{r(f)}$ which holds for all $\lambda > 0$ thus $r(h_{r(f)}) \le 0$ and $h_{r(f)}$ is well defined. □

The supremal deconvolution is defined by

$$\begin{aligned} f^\lambda(x) &= \sup_w \left\{ f(w) - \frac{1}{2\lambda}\|w-x\|^2 \right\} \\ &= \left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)^*(\lambda^{-1}x) - \frac{1}{2\lambda}\|x\|^2. \end{aligned} \tag{10}$$

The proximal hull $h_\lambda f$ of a function f is given by $h_\lambda f(x) = (f_\lambda)^\lambda(x)$ and corresponds to the supremum of all quadratics of the form $x \mapsto \alpha - \frac{1}{2\lambda}\|x-w\|^2$ majorized by f. It is a proper lower semi–continuous function when $\lambda < \bar{\lambda} := \bar{r}(f)^{-1}$.

Definition 3.4.

1. The Lasry–Lions double–envelope for $0 < \mu < \lambda < \bar{\lambda} = \bar{r}(f)^{-1}$ (the prox–threshold of f) is defined by

$$\begin{aligned} f_{\lambda,\mu}(x) &:= \sup_w \left\{ f_\lambda(w) - \frac{1}{2\mu}\|x-w\|^2 \right\} \\ &= \sup_w \left\{ \inf_u \{f(u)+\frac{1}{2\lambda}\|u-w\|^2\} - \frac{1}{2\mu}\|w-x\|^2 \right\} \end{aligned}$$

2. For $0 < \mu$ and $0 < \lambda < \bar{\lambda} = \bar{r}(f)^{-1}$ the μ–proximal hull $h_\mu(f_\lambda)$ of the infimal convolution f_λ is denoted by $f_{\lambda|\mu}$.

We now turn our attention to the epi-convergence of the Lasry–Lions double envelope.

Proposition 3.5 ([30]). *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, lower–semicontinuous and prox–bounded with threshold $\bar{\lambda} := \bar{r}(f)^{-1}$. Then for every $\lambda \in (0, \bar{\lambda})$ the set $P_\lambda f(x)$ is nonempty* *and compact, while the value $f_\lambda(x)$ is finite and depends continuously on (λ, x), with*

$$f_\lambda(x) \uparrow f(x) \quad \text{for all } x \text{ as } \lambda \downarrow 0.$$

In fact, $f_{\lambda^\nu}(x^\nu) \to f(\bar{x})$ whenever $x^\nu \to \bar{x}$ and $\lambda^\nu \downarrow 0$ in $(0, \bar{\lambda})$ in such a way that the sequence $\{\|x^\nu - \bar{x}\|/\lambda^\nu\}_{\nu\in\mathbb{N}}$ is bounded.

Furthermore, if $w^\nu \in P_{\lambda^\nu} f(x^\nu)$, $x^\nu \to \bar{x}$ and $\lambda^\nu \to \lambda \in (0, \bar{\lambda})$, then the sequence $\{w^\nu\}_{\nu\in\mathbb{N}}$ is bounded and all its cluster points lie in $P_\lambda f(\bar{x})$.

The following results show the connection between the Lasry–Lions double envelope and the infimal convolution of proximal hull and the intermediacy property of the Lasry–Lions double envelope.

Proposition 3.6 ([30]). *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, lower semi–continuous and prox–bounded with $0 < \mu < \lambda < \bar{\lambda} = \bar{r}(f)^{-1}$. Then*

1. $f_{\lambda,\mu} = (h_\lambda f)_{\lambda-\mu} = h_\mu (f_{\lambda-\mu}) = f_{(\lambda-\mu)|\mu}$
2. $f_\lambda \le f_{\lambda,\mu} \le f_{\lambda-\mu} \le f$

The following epi-limit property appears as Proposition 7.4 in [30].

Proposition 3.7. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be lower semicontinuous and proper. The following property holds for any sequence $\{f^\nu\}_{\nu\in N}$ of functions on $\mathbb{R}^n$. If $f_1^\nu \le f^\nu \le f_2^\nu$ with $f_1^\nu \to^e f$ and $f_2^\nu \to^e f$, then $f^\nu \to^e f$.*

The Lasry-Lions double–envelopes then have $f_{\lambda,\mu} \to^e f$ as $\lambda \downarrow 0$ and $\mu \downarrow 0$ with $0 < \mu < \lambda$, because of the sandwiching $f_\lambda \le f_{\lambda,\mu} \le f_{\lambda-\mu}$ that was recalled in Propositions 3.6 and 3.7.

Lemma 3.8. *Suppose $f^v : \mathbb{R}^n \to \overline{\mathbb{R}}$ is eventually prox–bounded and epi–converges to f as $v \to w$ and $\mu < \lambda < \bar{r}(\{f^v\})^{-1}$ (the eventual uniform prox-threshold for the f^v). Then the family $\{f^v_{\lambda,\mu}\}$ converges continuously to $f_{\lambda,\mu}$ (i.e. $f^v_{\lambda,\mu}(x^v) \to f_{\lambda,\mu}(x)$ for all $x^v \to x$).*

Proof. First note that by Theorem 7.37 of [30] we have $f^v_\lambda \to f_\lambda$ uniformly on bounded sets for $\lambda < \bar{r}(\{f^v\})^{-1}$. Let $\mu < \bar{\mu} < \lambda < \bar{r}(\{f^v\})^{-1}$. Take β and b such that $f^v_{\lambda,\bar{\mu}}(b) \le \beta$ for all v and hence for all y

$$f^v_\lambda(y) \le \beta + \frac{1}{2\bar{\mu}}\|b - w\|^2.$$

Let $x^v \to x$ (and so $\{\|x^v\|\}_v$ is uniformly bounded) and choose ρ so that $\rho \ge \max_v \|b - x^v\|$. Now as $f^v_\lambda - \frac{1}{2\lambda}\|\cdot\|^2$ is concave and we may bound above

$$\begin{aligned} f^v_\lambda(y) - \frac{1}{2\mu}\|x^v - y\|^2 &\le \beta + \frac{1}{2\bar{\mu}}\|b - y\|^2 - \frac{1}{2\mu}\|x^v - y\|^2 \\ &\le \beta - \frac{1}{2}\left(\frac{\bar{\mu} - \mu}{\bar{\mu}\mu}\right)\|y - b\|^2 + \frac{1}{2\mu}\|x^v - b\|\left(\|y - b\| + \|x^v - y\|\right) \\ &\le \beta - \frac{1}{2}\left(\frac{\bar{\mu} - \mu}{\bar{\mu}\mu}\right)\|y - b\|^2 + \frac{\rho}{2\mu}\left(2\|y - b\| + \rho\right) := h(y). \end{aligned}$$

As h has bounded upper level–sets we have $\operatorname{argmax}(f_\lambda^v(\cdot) - \frac{1}{2\mu}\|x^v - \cdot\|^2) \neq \emptyset$ for all v and as both f_λ^v and $-\frac{1}{2\mu}\|x^v - \cdot\|^2$ uniformly converge on bounded sets we have that $f_\lambda^v(\cdot) - \frac{1}{2\mu}\|x^v - \cdot\|^2$ hypo-converges to $f_\lambda(\cdot) - \frac{1}{2\mu}\|x - \cdot\|^2$ as $v \to w$, since $x^v \to x$. Hence

$$\max_y\{f_\lambda^v(y) - \frac{1}{2\mu}\|x^v - y\|^2\} = f_{\lambda,\mu}^v(x^v) \to f_{\lambda,\mu}(x),$$

verifying continuous convergence. □

4. Subhessians of the Infimal Convolution

In recent years the infimal convolution has been applied to the study of the differentiability of convex and nonconvex functions. This may be found in the work of Poliquin and Rockafellar [24], [25], Penot [23], Eberhard, Nyblom and Ralph, [9], [10] and many other authors. One of the main motivations for their work stems from the following observation which was probably first observed by Hiriart–Urruty and Plazanet in [15] (an alternative proof of this result is provided in [11]).

Lemma 4.1. *A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is locally $C^{1,1}$ around $\bar{x}$ if and only if it is simultaneously locally para-convex and para-concave around $\bar{x}$, and finite valued.*

The infimal convolution produces a para–concave function and the supremal deconvolution results in a para–convex function. The Lasry–Lions double-envelope combines both in a way that is designed to produce a $C^{1,1}(\mathbb{R}^n)$ function.

Definition 4.2. Suppose that $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a lower semi-continuous function.

1. Let $q_\lambda(\mathcal{A})(u) = \inf_{w\in\mathbb{R}^n}\{q(\mathcal{A})(w) + \frac{1}{2\lambda}\|w - u\|^2\}$ and

$$\mathcal{A}_\lambda := \{Q \in \mathcal{S}(n) \mid \langle Q, uu^t\rangle \leq 2q_\lambda(\frac{1}{2}\mathcal{A})(u) \quad \text{for all } u \in \mathbb{R}^n\}, \tag{11}$$

   where $q(\mathcal{A})(w) := \sup_{Q\in\mathcal{A}}\langle w, Qw\rangle$.
2. Denote $\partial_\lambda^{2,-} f(x,p) := (\partial^{2,-} f(x,p))_\lambda$.

Such smoothing may be alternatively viewed in terms of infimal convolution smoothings of the associated quadratic forms rather than the smoothing of the rank-1 support. This follows from the observation that the form $q_Q(u) := \langle Q, uu^t\rangle$ has an infimal convolution characterised as follows (the star denotes the convex conjugate)

$$(q_Q)_\lambda(h) = \frac{1}{2\lambda}\|h\|^2 - \lambda q^*_{I+2\lambda Q}(h) \text{ where} \tag{12}$$

$$q^*_{I+2\lambda Q}(h) = \begin{cases} -\infty & \text{if } I + 2\lambda Q \notin \mathcal{P}(n) \\ q_{(I+2\lambda Q)^+}(h) & \text{if } I + 2\lambda Q \in \mathcal{P}(n)\setminus(\operatorname{int}\mathcal{P}(n)),\ h \in \operatorname{Im}(I + 2\lambda Q) \\ q_{(I+2\lambda Q)^{-1}}(h) & \text{if } I + 2\lambda Q \in \operatorname{int}\mathcal{P}(n) \\ +\infty & \text{if } h \notin \operatorname{Im}(I + 2\lambda Q) \end{cases}$$

and $\operatorname{Im}(I + 2\lambda Q)$ denotes the image or range and $(I + 2\lambda Q)^+$ the Moore-Penrose inverse. For $\lambda > 0$ sufficiently small the minimum in $\min_\eta\{\langle Q\eta, \eta\rangle + \frac{1}{2\lambda}\|h - \eta\|^2\}$ is achieved at a unique point η and is the solution to the equation $2Q\eta + \frac{1}{\lambda}(\eta - h) = 0$. That is a unique minimum occurs at $\eta = (I + 2\lambda Q)^{-1}h$ if and only if $I + 2\lambda Q$ is invertible which occurs if and only if $I + 2\lambda Q \in \operatorname{int}\mathcal{P}(n)$. Thus $I + 2\lambda Q \in \operatorname{int}\mathcal{P}(n)$ is also a sufficient condition for $(q_Q)_\lambda = q_{Q_\lambda}$. *When this occurs we will (loosely) say that Q_λ is a quadratic form.*

The next result first appeared in a looser form in [9, Proposition 8] (i.e. the result does not state how small λ must be to obtain the inclusions). As we require explicit bounds on how small λ must be taken in order to obtain certain results we provide a proof in an appendix.

Proposition 4.3. *Suppose that $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a lower semi-continuous function which has a global minimum at $\bar{x}$.*

1. *Suppose $I + \lambda Q \in \operatorname{int}\mathcal{P}(n)$. If $(0, Q) \in \partial^{2,-}f(\bar{x})$ then there exists $\varphi \in C^2(\mathbb{R}^n)$ with $\varphi(\cdot) \le f(\cdot)$ in a neighbourhood of $\bar{x}$ and with $\varphi(\bar{x}) = f(\bar{x})$. In addition $(0, Q) = (\nabla\varphi(\bar{x}), \nabla^2\varphi(\bar{x}))$ and $(0, 2(\frac{1}{2}Q)_\lambda) = (\nabla\varphi_\lambda(\bar{x}), \nabla^2\varphi_\lambda(\bar{x}))$.*
2. *Also we may write for all $h \in \mathbb{R}^n$ and $\lambda > 0$ with $I + \lambda Q \in \operatorname{int}\mathcal{P}(n)$*

$$\begin{aligned}\langle \nabla^2\varphi_\lambda(\bar{x})h, h\rangle &= \langle Q_\lambda, hh^t\rangle \\ &:= 2\inf_\eta\{\langle \frac{1}{2}Q\eta, \eta\rangle + \frac{1}{2\lambda}\|h-\eta\|^2\} = 2\langle \frac{1}{2}Q, hh^t\rangle_\lambda \le \langle Q, hh^t\rangle\end{aligned}$$

   which monotonically decreases as λ increases.
3. *Also, for $\lambda > 0$ with $I + \lambda Q \in \operatorname{int}\mathcal{P}(n)$ we have $2(\frac{1}{2}Q)_\lambda \in \partial^{2,-}f_\lambda(x, 0) \subseteq \partial^{2,-}f(x, 0)$.*

Addition of a quadratic may be made to simplify the proof since we may relate the subjet of the resultant function to that for the original function.

Proposition 4.4. *Suppose $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be lower semi–continuous and prox–bounded with prox–threshold $(\bar{r}(f))^{-1}$. For $r \ge \bar{r}(f)$ and $0 < \mu < \frac{1}{r}$ place $\lambda = \frac{\mu}{1-r\mu}$. Then $2\left(\frac{1}{2}(Q + rI)\right)_\lambda$ is a quadratic form if and only if $2(\frac{1}{2}Q)_\mu$ is a quadratic form. Then*

$$2(\frac{1}{2}(Q + rI))_\lambda \in \partial^{2,-}\left(f(\cdot) + \frac{r}{2}\|\cdot - \bar{x}\|^2\right)_\lambda(\bar{x}, 0) \quad \Leftrightarrow \quad 2(\frac{1}{2}Q)_\mu \in \partial^{2,-}f_\mu(\bar{x}, 0). \qquad (13)$$

Proof. First note that $\lambda = \frac{\mu}{1-r\mu}$ if and only if $\mu = \frac{\lambda}{1+r\lambda}$. We have $2\left(\frac{1}{2}(Q + rI)\right)_\lambda$ a quadratic form if and only if $I + \lambda(Q + rI) \in \operatorname{int}\mathcal{P}(n)$ which is equivalent to $I + \frac{\mu r}{1-r\mu}I + \frac{\mu}{1-r\mu}Q = \frac{1}{1-r\mu}(I + \mu Q) \in \operatorname{int}\mathcal{P}(n)$ or $(I + \mu Q) \in \operatorname{int}\mathcal{P}(n)$, a condition necessary and sufficient for $I + \mu Q$ to be as quadratic form. Note next that

$$\frac{x + \lambda r\bar{x}}{1 + \lambda r} = (1 - \mu r)x + \mu r\bar{x}.$$

Applying (6) we find

$$\begin{aligned}&\left(f(\cdot) + \frac{r}{2}\|\cdot - \bar{x}\|^2\right)_\lambda(x) \\ &= f_\mu\left((1-\mu r)x + r\mu\bar{x}\right) - \frac{1-r\mu}{2\mu}\left((\frac{1}{1-\mu r})\|(1-\mu r)x + r\mu\bar{x}\|^2 - \|x\|^2\right) \\ &= f_\mu\left((1-\mu r)x + r\mu\bar{x}\right) - \frac{1}{2\mu}\left(\|(1-\mu r)x + r\mu\bar{x}\|^2 - (1-\mu r)\|x\|^2\right).\end{aligned}$$

Now observe that

$$\nabla_x \left\{ \frac{1}{2\mu} \left(\|(1-\mu r)x + r\mu\bar{x}\|^2 - (1-\mu r)\|x\|^2 \right) \right\} = -r(1-r\mu)(x-\bar{x})$$

(which is zero at $\bar{x}$) and so

$$\nabla_x^2 \left\{ \frac{1}{2\mu} \left(\|(1-\mu r)x + r\mu\bar{x}\|^2 - (1-\mu r)\|x\|^2 \right) \right\} = -r(1-r\mu)I.$$

Place $h(x) := f_\mu\left((1-\mu r)x + r\mu\bar{x}\right)$ and note that $(1-r\mu)^2 M \in \partial^{2,-}h(\bar{x},0)$ when

$$h(x) - h(\bar{x}) \geq \frac{1}{2}\langle (1-r\mu)^2 M(x-\bar{x}), (x-\bar{x})\rangle + o(\|x-\bar{x}\|^2)$$

which may be rewritten with $y := (1-r\mu)x + r\mu\bar{x}$ (or $y - \bar{x} = (1-r\mu)(x-\bar{x})$) as

$$f_\mu(y) - f_\mu(\bar{x}) \geq \frac{1}{2}\langle M(y-\bar{x}), (y-\bar{x})\rangle + o(\|y-\bar{x}\|^2).$$

Thus it follows that $(1-r\mu)^2 M \in \partial^{2,-}h(\bar{x},0)$ if and only if $M \in \partial^{2,-}f_\mu(\bar{x},0)$ and hence

$$\partial^{2,-}\left(f(\cdot) + \frac{r}{2}\|\cdot - \bar{x}\|^2\right)_\lambda (\bar{x},0) = (1-\mu r)^2 \partial^{2,-} f_\mu(\bar{x},0) + r(1-r\mu)I. \tag{14}$$

Next note that $2(\frac{1}{2}(Q+rI))_\lambda$ corresponds to the quadratic form $2(\langle \frac{1}{2}Q, xx^t\rangle + \frac{r}{2}\|x\|^2)_\lambda$ and so we may apply (6) with $\bar{x} = 0$ to obtain

$$\begin{aligned}
&\langle 2(\frac{1}{2}(Q+rI))_\lambda, xx^t\rangle \\
= \;& 2\left(\langle \frac{1}{2}Q, ((1-r\mu)x)((1-r\mu)x)^t\rangle_\mu - \frac{1}{2\mu}(\|(1-r\mu)x\|^2 - (1-r\mu)\|x\|^2) \right) \\
= \;& 2\left((1-r\mu)^2 \langle \frac{1}{2}Q, xx^t\rangle_\mu - \frac{(1-r\mu)}{2\mu}((1-r\mu)\|x\|^2 - \|x\|^2) \right) \\
= \;& 2\langle \frac{1}{2}Q, xx^t\rangle_\mu (1-r\mu)^2 + r(1-r\mu)\|x\|^2 \\
= \;& \langle 2(\frac{1}{2}Q)_\mu (1-r\mu)^2 + r(1-r\mu)I, xx^t\rangle.
\end{aligned}$$

As this holds for all x we have

$$2(\frac{1}{2}(Q+rI))_\lambda = 2(\frac{1}{2}Q)_\mu (1-r\mu)^2 + r(1-r\mu)I. \tag{15}$$

On comparing (14) and (15) find that $2(\frac{1}{2}(Q+rI))_\lambda \in \partial^{2,-}\left(f + \frac{1}{2}\|\cdot - \bar{x}\|^2\right)(\bar{x},0)$ if and only if $2(\frac{1}{2}Q)_\mu \in \partial^{2,-}f_\mu(\bar{x},0)$. □

We may now remove the assumption that f take a global minimum at x when obtaining a result similar to Proposition 4.3.

Corollary 4.5. *Suppose that $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is lower semi-continuous and prox–bounded with a prox–threshold of $(\bar{r}(f))^{-1}$. Let $Q \in \partial^{2,-} f(\bar{x}, 0) \neq \emptyset$. Then*

1. *there exists an $r \geq \bar{r}(f)$ such that*

$$f(x) \geq f(\bar{x}) - \frac{r}{2}\|x - \bar{x}\|^2 \quad \text{for all } x \text{ and} \tag{16}$$

2. *for $\mu > 0$ with $I + \mu Q \in \operatorname{int} \mathcal{P}(n)$ and $0 < \mu < \frac{1}{r}$ we have $2(\frac{1}{2}Q)_\mu \in \partial^{2,-} f_\mu(\bar{x}, 0)$.*

Proof. By a translation we may place $\bar{x} = 0$. Take a fixed $X \in \partial^{2,-} f(0,0) \neq \emptyset$ then by [9, Proposition 6] for all $M, \gamma > 0$ there exists a function $\varepsilon : \mathbb{R}_+ \to \mathbb{R}$ such that $\lim_{t \downarrow 0} \varepsilon(t) = 0 = \varepsilon(0)$ and a function $\rho : y \mapsto \varepsilon(\|y\|)\|y\|^2 \in \mathcal{C}^2(\mathbb{R}^n)$ satisfying $(\nabla \rho(0), \nabla^2 \rho(0)) = (0,0)$ along with $\varepsilon(\|y\|) = c := \bar{r}(f) + \max\{\max\{0, -k\} \mid k \text{ is an eigenvalue of } X\}$ for all $y \notin B_M(0)$ and

$$f(y) - f(0) \geq \frac{1}{2}\langle X - \gamma I, yy^t \rangle - \varepsilon(\|y\|)\|y\|^2.$$

Thus we may take

$$r := \max\{\max\{0, -k\} \mid k \text{ is an eigenvalue of } X\} + 2\max\{c, \sup\{\varepsilon(y) \mid y \in B_M(0)\}\} + \gamma.$$

Take r to be the smallest such positive number that satisfies (16). Define the function

$$h_r(x) := f(x) + \frac{r}{2}\|x - \bar{x}\|^2.$$

Note that h_r has a global minimum at $\bar{x}$ and so $\bar{r}(h_r) = 0$ and $(h_r)_\lambda$ is defined for all $\lambda > 0$. Take $Q \in \partial^{2,-} f(\bar{x}, 0)$ and so by Proposition 3.3 we have $Q + rI \in \partial^{2,-}(h_r)(\bar{x}, 0)$. Now apply Proposition 4.3 assuming $I + \lambda(Q + rI) \in \operatorname{int} \mathcal{P}(n)$ and deduce that $2(\frac{1}{2}(Q + rI))_\lambda \in \partial^{2,-}(h_r)_\lambda(\bar{x}, 0)$. Now apply Proposition 4.4 to obtain that $2(\frac{1}{2}Q)_\mu \in \partial^{2,-} f_\mu(\bar{x}, 0)$, where $\mu = \frac{\lambda}{1 + r\lambda}$. Next note that

$$I + \lambda(Q + rI) \in \operatorname{int} \mathcal{P}(n) \quad \Leftrightarrow \quad I + \mu Q \in \operatorname{int} \mathcal{P}(n)$$

since $I + \mu Q = \frac{I + \lambda(Q + rI)}{1 + \lambda r}$. As $\lambda > 0$ the only restriction on μ is that implied by $\lambda = \frac{\mu}{1 - r\mu} > 0$ which gives $0 < \mu < \frac{1}{r}$. □

Remark 4.6. Having established that a value of r for which (16) holds we could take the smallest r and denote this number by $\bar{r}(f, \bar{x})$ ($\geq \bar{r}(f)$ by definition). This number is a property of the function f at $\bar{x}$ and not dependent on the choice of Q.

The following result is very useful in subsequent proofs (see [9, Proposition 12]). Note that the condition $I + \lambda Q \in \mathcal{P}(n)$ is necessary and sufficient to ensure $2(\frac{1}{2}Q)_\lambda$ is a quadratic form. Once again minor changes are required to that in [9, Proposition 12] so we provide a proof in the appendix.

Proposition 4.7. *Suppose that $\mathcal{A}$ is a rank–one representer with $-\mathcal{P}(n) \subseteq 0^+\mathcal{A}$. Then for $\lambda > 0$ such that $\mathcal{A}_\lambda \neq \emptyset$ we have*

$$\mathcal{A}_\lambda = \operatorname{cl}\{2(\frac{1}{2}Q)_\lambda \mid Q \in \mathcal{A} \text{ and } 2(\frac{1}{2}Q)_\lambda \text{ is a quadratic form}\} - \mathcal{P}(n).$$

The following result shows that the infimal smoothing of a function corresponds exactly to an infimal convolution smoothing of its subjets.

Proposition 4.8 ([9]). *Suppose that* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *is lower semi-continuous and prox–bounded.*

1. *Then for all* $\lambda > 0$ *such that* $\lambda < \bar{r}(f,x)^{-1}$*.*

$$\partial^{2,-} f_\lambda(x,0) = \partial^{2,-}_\lambda f(x,0).$$

2. *For the limiting subhessians*

$$\begin{aligned}\underline{\partial}^2 f(x,p) &= \limsup_{\lambda\downarrow 0} \underline{\partial}^2 (f - \langle p,\cdot\rangle)_\lambda(x,0) \\ &= \limsup_{\lambda\downarrow 0} \underline{\partial}^2_\lambda f(x,p) = \limsup_{\lambda\downarrow 0} \underline{\partial}^2 f_\lambda(x+\lambda p, p).\end{aligned} \tag{17}$$

Proof. We prove *1.* only as it is a slight variation of [9, Corollary 3] but contains a quantitative bound for the size of λ required. The inequality

$$2(\frac{1}{2} f''_s(x,0,\cdot))_\lambda(h) \geq (f_\lambda)''_s(x,0,h)$$

is proved in [9, Corollary 3] for all $\lambda > 0$. As this holds for all h we have by the properties of rank–1 support functions that

$$\partial^{2,-} f_\lambda(x,0) \subseteq \partial^{2,-}_\lambda f(x,0).$$

Now take $H \in \partial^{2,-}_\lambda f(x,0)$ and hence by Proposition 4.7 we have the existence of a sequence $\{Q_k\}_{k\in\mathbb{N}}$ such that $2(\frac{1}{2}Q_k)_\lambda \to P + H$ where $P \in \mathcal{P}(n)$, $2(\frac{1}{2}Q_k)_\lambda \in \partial^{2,-}_\lambda f(x,0)$ for all k and $Q_k \in \partial^{2,-} f(x,0)$. In particular since $2(\frac{1}{2}Q_k)_\lambda$ is a quadratic form we have $I + \lambda Q_k \in \operatorname{int}\mathcal{P}(n)$ for all k. Applying Corollary 4.5 we have $2(\frac{1}{2}Q_k)_\lambda \in \partial^{2,-} f_\lambda(x,0)$ for all k. As $\partial^{2,-} f_\lambda(x,0)$ is closed we have $H + P \in \partial^{2,-} f_\lambda(x,0)$ implying $H = H + P - P \in \partial^{2,-} f_\lambda(x,0)$. □

The assumption that $p = 0$ is finally removed. First note that $r \geq \bar{r}(f - \langle p,\cdot\rangle, x)$ if and only if $r \geq 0$ and for all y we have

$$f(y) \geq f(x) - \langle p, y - x\rangle + \frac{r}{2}\|y - x\|^2.$$

Denote the smallest of these positive constants by $\bar{r}(f,x,p) := \bar{r}(f - \langle p,\cdot\rangle, x)$.

Lemma 4.9. *Suppose* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *is lower semi–continuous, prox–bounded and* $(p,Q) \in \partial^{2,-} f(x)$*. Then for all* $\lambda > 0$ *such that* $\lambda < \bar{r}(f,x,p)^{-1}$ *and* $I + \lambda Q \in \operatorname{int}\mathcal{P}(n)$ *we have*

$$(p,Q) \in \partial^{2,-} f(x) \quad \Rightarrow \quad (p, 2(\frac{1}{2}Q)_\lambda) \in \partial^{2,-} f_\lambda(x + \lambda p)$$

$$\text{and} \quad (f - \langle p,\cdot\rangle)_\lambda(x) = f(x) - \langle p, x\rangle.$$

Proof. First note that $(0,Q) \in \partial^{2,-}(f - \langle p, \cdot\rangle)(x)$. Using Proposition 4.8 we have

$$\partial^{2,-}\left(f - \langle p,\cdot\rangle\right)_\lambda (x,0) = \partial^{2,-}_\lambda \left(f - \langle p,\cdot\rangle\right)(x,0)$$

and so for λ such that $I + \lambda Q \in \operatorname{int} \mathcal{P}(n)$ we have $2(\frac{1}{2}Q)_\lambda \in \partial^{2,-}_\lambda \left(f - \langle p,\cdot\rangle\right)(x,0)$ and so $2(\frac{1}{2}Q)_\lambda \in \partial^{2,-}\left(f - \langle p,\cdot\rangle\right)_\lambda (x,0)$. Now we use Lemma 3.2 which establishes that for any y we have $\left(f - \langle p,\cdot\rangle\right)_\lambda (y) = f_\lambda(y+\lambda p) - \langle p, y\rangle - \frac{\lambda}{2}\|p\|^2$. This yields

$$\begin{aligned}(0, 2(\frac{1}{2}Q)_\lambda) &\in \partial^{2,-}\left(f_\lambda(\cdot + \lambda p) - \langle p,\cdot\rangle - \frac{\lambda}{2}\|p\|^2\right)(x)\\ &= \partial^{2,-} f_\lambda(x+\lambda p) - (p,0)\\ \text{and so} \quad & (p, 2(\frac{1}{2}Q)_\lambda) \in \partial^{2,-} f_\lambda(x+\lambda p).\end{aligned}$$

To prove the final part we note that there exists a $\varphi \in C^2(\mathbb{R}^n)$ such that $\nabla\varphi(x) = 0$ and $\nabla^2\varphi(x) = 2(\frac{1}{2}Q)_\lambda$ with

$$\left(f - \langle p,\cdot\rangle\right)_\lambda (y) - \varphi(y) \geq \left(f - \langle p,\cdot\rangle\right)_\lambda (x) - \varphi(x)$$

for all y. This implies

$$\begin{aligned} f_\lambda(y+\lambda p) - \langle p,y\rangle - \frac{\lambda}{2}\|p\|^2 - \varphi(y) &\geq f_\lambda(x+\lambda p) - \langle p,x\rangle - \frac{\lambda}{2}\|p\|^2 - \varphi(x)\\ \text{or} \quad f_\lambda(y+\lambda p) - (\varphi(y) + \langle p,y\rangle) &\geq f_\lambda(x+\lambda p) - (\varphi(x) + \langle p,x\rangle)\end{aligned}$$

where $\nabla\left(\varphi(\cdot) + \langle p,\cdot\rangle\right)(x) = p$ and $\nabla^2\left(\varphi(\cdot) + \langle p,\cdot\rangle\right)(x) = 2(\frac{1}{2}Q)_\lambda$.
Finally apply (4) to $(p, 2(\frac{1}{2}Q)_\lambda) \in \partial^{2,-} f_\lambda(x+\lambda p)$ to get $(p, 2(\frac{1}{2}Q)_\lambda) \in \partial^{2,-} f(x)$ and

$$f(x) = f(x + \lambda p - \lambda p) = f_\lambda(x+\lambda p) - \frac{\lambda}{2}\|p\|^2$$

and so

$$\begin{aligned}\left(f - \langle p,\cdot\rangle\right)_\lambda (x) &= f_\lambda(x+\lambda p) - \langle p,x\rangle - \frac{\lambda}{2}\|p\|^2\\ &= f(x) - \langle p,x\rangle.\end{aligned}$$

□

5. The Subhessian of the Supremal Deconvolution

In this section we discuss the problem of relating the subhessians of the supremal deconvolution of a function to the elements of the subhessian of the function. A complete comparison is only possible when we are dealing with para–concave functions. Fortunately this is the situation that arises naturally in the analysis of the Lasry–Lions double envelope which is the subject of the next section.

From time to time we will use the Hilbert identity

$$\begin{aligned}&\frac{c_1}{2}\|w - y_1\|^2 + \frac{c_2}{2}\|w - y_2\|^2\\ &= \frac{c_1 c_2}{2(c_1+c_2)}\|y_1 - y_2\|^2 + \left(\frac{c_1+c_2}{2}\right)\left\|w - \frac{1}{c_1+c_2}(c_1 y_1 + c_2 y_2)\right\|^2. \end{aligned} \tag{18}$$

For future use we also note that (18) gives

$$\begin{aligned}(f_\lambda)_\eta(x) &= \inf_w \left\{ \inf_v \left\{ f(v) + \frac{1}{2\lambda}\|w-v\|^2 + \frac{1}{2\eta}\|w-x\|^2 \right\}\right\} \\ &= \inf_{w,v} \left\{ f(v) + \frac{1}{2(\lambda+\eta)}\|v-x\|^2 + \frac{\lambda+\eta}{2\lambda\eta}\left\| w - \frac{1}{\lambda+\eta}(\lambda v + \eta x)\right\|^2 \right\} \\ &= \inf_v \left\{ f(v) + \frac{1}{2(\lambda+\eta)}\|v-x\|^2 \right\} = f_{\lambda+\eta}(x).\end{aligned}$$

We shall require the following result which due to Penot [23]. It motivates the desire to study the sub/super–jet of the convex conjugate of certain functions.

Proposition 5.1. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ with $f(\bar{x}) < +\infty$.*

1. *Suppose A is a positive semi–definite superjet of f at $(\bar{x}, \bar{x}^*)$. If A is invertible and $f^*(\bar{x}^*) < +\infty$ then A^{-1} is a subjet of f^* at $(\bar{x}^*, \bar{x})$.*
2. *Suppose we assume f is proper, lower semi–continuous, convex and suppose that for some $\bar{x}^* \in \partial f(\bar{x})$ we have A a positive semi–definite subjet at $(\bar{x}, \bar{x}^*)$. If A is invertible then A^{-1} is a superjet of f^* at $(\bar{x}^*, \bar{x})$.*

We are going to connect this result to the subhessians of a general class of functions.

Lemma 5.2. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Assume $I - \lambda B \in \mathcal{P}(n)$. Then*

1. *$B \in \partial^{2,-} f(x, 0)$ if and only if $A := \frac{1}{\lambda} I - B \in \partial^{2,+}\left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)(x, \lambda^{-1}x)$ and*
2. *$A^{-1} = \left(\frac{1}{\lambda} I - B\right)^{-1} \in \partial^{2,-}\left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)^*(\lambda^{-1}x, x)$ if and only if $B + \lambda B^2 (I - \lambda B)^{-1} \in \partial^{2,-} f^\lambda(x, 0)$*

Proof. First note that when $A \in \partial^{2,+}\left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)(x, \lambda^{-1}x)$ then by definition we have for all y that

$$\frac{1}{2\lambda}\|y\|^2 - f(y) \le \frac{1}{2\lambda}\|x\|^2 - f(x) + \langle \lambda^{-1}x, y - x\rangle + \frac{1}{2}\langle A(y-x), y-x\rangle + o(\|y-x\|^2)$$

which is equivalent to

$$\begin{aligned} f(y) &\ge f(x) + \frac{1}{2\lambda}\|y\|^2 - \frac{1}{2\lambda}\|x\|^2 - \langle \lambda^{-1}x, y-x\rangle - \frac{1}{2}\langle A(y-x), y-x\rangle + o(\|y-x\|^2) \\ &= f(x) + \frac{1}{2\lambda}\|y-x\|^2 - \frac{1}{2}\langle A(y-x), y-x\rangle + o(\|y-x\|^2) \\ &= f(x) + \frac{1}{2}\langle \left(\frac{1}{\lambda} I - A\right)(y-x), y-x\rangle + o(\|y-x\|^2)\end{aligned}$$

which in turn is equivalent to $B := \frac{1}{\lambda} I - A \in \partial^{2,-} f(x, 0)$. This establishes the first part.

Next consider $A^{-1} \in \partial^{2,-}\left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)^*(\lambda^{-1}x, x)$ then by definition this means $\left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)^*(\lambda^{-1}x)$ finite, and for all y that

$$\begin{aligned}\left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)^*(\lambda^{-1}y) &\ge \left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)^*(\lambda^{-1}x) + \langle x, \lambda^{-1}(y-x)\rangle \\ &\quad + \frac{1}{2}\langle A^{-1}\lambda^{-1}(y-x), \lambda^{-1}(y-x)\rangle + o(\|\lambda^{-1}y - \lambda^{-1}x\|^2)\end{aligned}$$

which becomes on the use of (10)

$$\begin{aligned} f^\lambda(y) &\geq f^\lambda(x) + \langle \lambda^{-1}x, y-x\rangle + \frac{1}{2\lambda^2}\langle A^{-1}(y-x),(y-x)\rangle + o(\lambda^{-2}\|y-x\|^2) \\ &\quad - \left(\frac{1}{2\lambda}\|y\|^2 - \frac{1}{2\lambda}\|x\|^2\right) \\ &= f^\lambda(x) + \frac{1}{2\lambda^2}\langle A^{-1}(y-x),(y-x)\rangle - \frac{1}{2\lambda}\|y-x\|^2 + o(\|y-x\|^2) \\ &= f^\lambda(x) + \frac{1}{2}\langle\left(\frac{1}{\lambda^2}A^{-1} - \frac{1}{\lambda}I\right)(y-x),(y-x)\rangle + o(\|y-x\|^2) \end{aligned}$$

which is equivalent to $(0, \frac{1}{\lambda^2}A^{-1} - \frac{1}{\lambda}I) \in \partial^{2,-} f^\lambda(x)$.

Now using the fact $A = \frac{1}{\lambda}I - B$ we will show

$$\frac{1}{\lambda^2}A^{-1} - \frac{1}{\lambda}I = B + \lambda B^2\left(I - \lambda B\right)^{-1} \text{ when } (I-\lambda B)^{-1} \text{ exists.}$$

Indeed, $I = I$ can be written as

$$\begin{aligned} I &= I - \lambda B + \lambda B - \lambda^2 B^2 + \lambda^2 B^2 \\ &= (I+\lambda B)(I-\lambda B) + \lambda^2 B^2 \\ &= (I+\lambda B)(I-\lambda B) + \lambda^2 B^2 (I-\lambda B)^{-1}(I-\lambda B) \\ &= (I + \lambda B + \lambda^2 B^2(I-\lambda B)^{-1})(I-\lambda B) \end{aligned}$$

which implies that

$$(I-\lambda B)^{-1} = I + \lambda B + \lambda^2 B^2 (I-\lambda B)^{-1}$$

and so we get

$$\frac{1}{\lambda^2}A^{-1} - \frac{1}{\lambda}I = B + \lambda B^2 (I-\lambda B)^{-1}.$$

This completes the proof. □

The next result exploits the ability to rewrite the supremal deconvolution as a formula involving the convex conjugate of the function $\frac{1}{2\lambda}\|\cdot\|^2 - f$.

Theorem 5.3. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be prox–bounded and proper.*

1. *Then $B \in \partial^{2,-} f(x,0)$ implies $B + \lambda B^2\left(I-\lambda B\right)^{-1} \in \partial^{2,-} f^\lambda(x,0)$ if $0 < \lambda < \frac{1}{\bar{r}(f)}$ (the prox–threshold for f) is such that $\lambda^{-1}I - B$ is positive definite (or $I - \lambda B \in \operatorname{int}\mathcal{P}(n)$) and if $f^\lambda(x) < +\infty$.*
2. *Assume further that f is para–concave and $\lambda < \frac{1}{\bar{r}(f)}$ (the prox–threshold for f) is such that both $I - \lambda B$ is positive definite and $f - \frac{1}{2\lambda}\|\cdot\|^2$ is concave. Then $B + \lambda B^2\left(I-\lambda B\right)^{-1} \in \partial^{2,-} f^\lambda(x,0)$ implies $B \in \partial^{2,-} f(x,0)$. Also, we have $f^\lambda(x) = f(x)$.*

Proof. *Part 1.* Clearly if $B \in \partial^{2,-} f(x,0)$ then f is finite at x. Now by Lemma 5.2 we have $A := \frac{1}{\lambda}I - B \in \partial^{2,+}\left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)(x, \lambda^{-1}x)$ and when A is positive definite by Proposition 5.1 (part *1.*) we have (since $\left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)^*\left(\frac{x}{\lambda}\right) < +\infty$ as follows from $f^\lambda(x) < +\infty$ by using (10))

$$\left(\frac{1}{\lambda}I - B\right)^{-1} \in \partial^{2,-}\left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)^*(\lambda^{-1}x, x). \tag{19}$$

On applying Lemma 5.2 we arrive at $B + \lambda B^2 (I - \lambda B)^{-1} \in \partial^{2,-} f^\lambda(x, 0)$ as desired.

Part 2. Assume now that $\lambda > 0$ such that $I - \lambda B > 0$, $f - \frac{1}{2\lambda}\|\cdot\|^2$ concave, and that $B + \lambda B^2 (I - \lambda B)^{-1} \in \partial^{2,-} f^\lambda(x, 0)$ (so $f^\lambda(x) < +\infty$ by definition). Then

$$A^{-1} = (\frac{1}{\lambda}I - B)^{-1} \in \partial^{2,-}(\frac{1}{2\lambda}\|\cdot\|^2 - f)^*(\frac{x}{\lambda}, x) \text{ by Lemma 5.2 (part } \mathit{2.})$$

with $(\frac{1}{2\lambda}\|\cdot\|^2 - f)^*(\frac{x}{\lambda})$ finite.

Place $\psi := (\frac{1}{2\lambda}\|\cdot\|^2 - f)^*$. It is convex and lower semi–continuous. Also $\psi(\frac{x}{\lambda})$ finite, and $\psi(x^*) > -\infty$ for all x^* from the prox–bound on f (indeed, $\frac{1}{2\lambda}\|\cdot\|^2 - f \le (\frac{1}{2\lambda} + c)\|\cdot\|^2 - \alpha$ for some $c > 0$, $\alpha \in \mathbb{R}$, so $\psi(x^*) = (\frac{1}{2\lambda}\|\cdot\|^2 - f)^*(x^*) \ge ((\frac{1}{2\lambda} + c)\|\cdot\|^2)^*(x^*) + \alpha > -\infty$).

Hence ψ is also proper, and we may apply Proposition 5.1 (part *2.*) to ψ. Now $A^{-1} = (\frac{1}{\lambda}I - B)^{-1} \in \partial^{2,-}\psi(\frac{x}{\lambda}, x)$ and $A^{-1} > 0$ (since $\frac{I}{\lambda} - B > 0$ so $(\frac{I}{\lambda} - B)^{-1} > 0$ also). By the subjet property,

$$\psi(y) \ge \psi(\frac{x}{\lambda}) + \langle y - \frac{x}{\lambda}, x\rangle + \frac{1}{2}\langle (y - \frac{x}{\lambda}), A^{-1}(y - \frac{x}{\lambda})\rangle + o(\|y - \frac{x}{\lambda}\|^2)$$

locally near $\frac{x}{\lambda}$. By positive–definiteness of A^{-1}, there is $\zeta > 0$ such that $\langle (y - \frac{x}{\lambda}), A^{-1}(y - \frac{x}{\lambda})\rangle \ge \zeta\|y - \frac{x}{\lambda}\|^2$ for all y and so, in a small neighbourhood of $\frac{x}{\lambda}$, $\psi(y) \ge \psi(\frac{x}{\lambda}) + \langle x, y - \frac{x}{\lambda}\rangle$ for all y near $\frac{x}{\lambda}$ (and hence globally, since ψ convex). Thus $x \in \partial\psi(\frac{x}{\lambda})$ (convex subdifferential), and Proposition 5.1 (part *2.*) yields $A \in \partial^{2,+}\psi^*(x, \frac{x}{\lambda})$. Note that $\psi^* = (\frac{1}{2\lambda}\|\cdot\|^2 - f)^{**}$ is also proper convex, lower semi–continuous and so $\psi^* = \mathrm{cl}(\frac{1}{2\lambda}\|\cdot\|^2 - f)$. From the prox–bound on f, we know that $\frac{1}{2\lambda}\|\cdot\|^2 - f(\cdot) < +\infty$ always. From the finiteness of $\psi(\frac{x}{\lambda})$ and the Fenchel inequality, $(-\infty \ne)\ \langle \frac{x}{\lambda}, x^*\rangle - \psi(\frac{x}{\lambda}) \le \psi^*(x^*) \le \frac{1}{2\lambda}\|x^*\|^2 - f(x^*)$ for all x^* and so $\frac{1}{2\lambda}\|\cdot\|^2 - f$ is finite–valued, implying continuity (since convex). Thus $\psi^* = \mathrm{cl}(\frac{1}{2\lambda}\|\cdot\|^2 - f) = \frac{1}{2\lambda}\|\cdot\|^2 - f$. Hence $A \in \partial^{2,+}(\frac{1}{2\lambda}\|\cdot\|^2 - f)(x, \frac{x}{\lambda})$, and finally, from Lemma 5.2 (part *1.*) $B \in \partial^{2,-} f(x, 0)$.

For the claim that $f_\lambda(x) = f(x)$, note that we have (19) holding. Now when $f - \frac{1}{2\lambda}\|\cdot\|^2$ is concave by (19) we have $x \in \partial\left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)^*(\lambda^{-1}x) \ne \emptyset$ and via the Fenchel equality that

$$\left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)(x) + \left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)^*(\lambda^{-1}x) = \langle \lambda^{-1}x, x\rangle = \frac{1}{\lambda}\|x\|^2$$

$$\text{or} \quad \left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)^*(\lambda^{-1}x) - \frac{1}{2\lambda}\|x\|^2 = f(x).$$

Applying (10) we have $f^\lambda(x) = f(x)$. □

Next lemma allows the assumption that $p = 0$ to be removed.

Lemma 5.4. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Then for any x and $\bar{p}$ we have*

$$(f - \langle \bar{p}, \cdot \rangle)^{\lambda}(x) = f^{\lambda}(x - \lambda \bar{p}) - \langle \bar{p}, x \rangle + \frac{\lambda}{2}\|\bar{p}\|^2. \tag{20}$$

In particular when $(f - \langle \bar{p}, \cdot \rangle)^{\lambda}(x) = f(x) - \langle \bar{p}, x \rangle$ we have $f^{\lambda}(x - \lambda \bar{p}) + \frac{\lambda}{2}\|\bar{p}\|^2 = f(x)$.

Proof. By definition

$$\begin{aligned}
(f - \langle \bar{p}, \cdot \rangle)^{\lambda}(x) &= \left(\frac{1}{2\lambda}\|\cdot\|^2 - (f - \langle \bar{p}, \cdot \rangle)\right)^*(\lambda^{-1}x) - \frac{1}{2\lambda}\|x\|^2 \\
&= \sup_w \left(\langle \lambda^{-1}x, w \rangle - \langle \bar{p}, w \rangle - \left(\frac{1}{2\lambda}\|w\|^2 - f(w)\right)\right) - \frac{1}{2\lambda}\|x\|^2 \\
&= \sup_w \left(\langle \lambda^{-1}x - \bar{p}, w \rangle - \left(\frac{1}{2\lambda}\|w\|^2 - f(w)\right)\right) - \frac{1}{2\lambda}\|x\|^2 \\
&= \left(\frac{1}{2\lambda}\|\cdot\|^2 - f\right)^*(\lambda^{-1}(x - \lambda\bar{p})) - \frac{1}{2\lambda}\|x - \lambda\bar{p}\|^2 - \langle \bar{p}, x \rangle + \frac{\lambda}{2}\|\bar{p}\|^2 \\
&= f^{\lambda}(x - \lambda\bar{p}) - \langle \bar{p}, x \rangle + \frac{\lambda}{2}\|\bar{p}\|^2.
\end{aligned}$$

The last statement follows immediately. □

The previous results may be extended using the fact that

$$\partial^{2,-} f(\bar{x}, \bar{p}) = \partial^{2,-}(f - \langle \bar{p}, \cdot \rangle)(\bar{x}, 0).$$

Corollary 5.5. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be prox–bounded and proper.*

1. *Then $B \in \partial^{2,-} f(\bar{x}, \bar{p})$ implies $B + \lambda B^2 (I - \lambda B)^{-1} \in \partial^{2,-} f^{\lambda}(\bar{x} - \lambda\bar{p}, \bar{p})$ whenever λ is chosen such that $0 < \lambda < \frac{1}{r(f)}$, $(\lambda^{-1}I - B)$ is positive definite and $f^{\lambda}(\bar{x} - \lambda\bar{p}) < +\infty$. In particular we have that $\bar{p} \in \partial_p f(\bar{x})$ implies $\bar{p} \in \partial_p f^{\lambda}(\bar{x} - \lambda\bar{p})$.*
2. *Assume further that f is para–concave and λ is such that both $I - \lambda B$ is positive definite and $f - \frac{1}{2\lambda}\|\cdot\|^2$ is concave. Then $B + \lambda B^2 (I - \lambda B)^{-1} \in \partial^{2,-} f^{\lambda}(\bar{x} - \lambda\bar{p}, \bar{p})$ implies $B \in \partial^{2,-} f(\bar{x}, \bar{p})$ and $f^{\lambda}(x - \lambda\bar{p}) + \frac{\lambda}{2}\|\bar{p}\|^2 = f(x)$. In particular we have that $\bar{p} \in \partial_p f^{\lambda}(\bar{x} - \lambda\bar{p})$ implies $\bar{p} \in \partial_p f(\bar{x})$.*

Proof. The proof is immediate from Theorem 5.3 and Lemma 5.4. □

The expression for the matrices may be expressed in a more compact form by using the next result.

Proposition 5.6. *Given $Q \in \mathcal{S}(n)$ denote by $q_Q(h) = \langle Qh, h \rangle$ and $(q_Q)^{\lambda}(h) = \sup_{\eta}\{\langle Q\eta, \eta \rangle - \frac{1}{2\lambda}\|h - \eta\|^2\}$ for all $h \in \mathbb{R}^n$.*

1. *Let $\lambda > 0$ be such that $I - 2\lambda Q \in \operatorname{int} \mathcal{P}(n)$ then $(q_Q)^{\lambda}$ is a quadratic form. More precisely $(q_Q)^{\lambda} = q_{Q^{\lambda}}$ with*

$$Q^{\lambda} = Q(I - 2\lambda Q)^{-1} = (I - 2\lambda Q)^{-1}Q = \frac{1}{2\lambda}((I - 2\lambda Q)^{-1} - I).$$

2. *Given $P \in \mathcal{S}(n)$ then when $I + 2\lambda P \in \operatorname{int} \mathcal{P}(n)$ we have*

$$P_\lambda = P(I + 2\lambda P)^{-1} = (I + 2\lambda P)^{-1} P \in \mathcal{S}(n)$$

if and only if $(P_\lambda)^\lambda = P$.
Indeed $I - 2\lambda Q \in \operatorname{int} \mathcal{P}(n)$ is a necessary and sufficient condition for $(q_Q)^\lambda$ to be the quadratic form $q_{Q^\lambda}(h) = \langle Q^\lambda h, h\rangle$.

Proof. Note that when $I - 2\lambda Q \in \operatorname{int} \mathcal{P}(n)$ then the $\max_\eta\{\langle Q\eta, \eta\rangle - \frac{1}{2\lambda} \|h - \eta\|^2\}$ has a unique solution at η which satisfies the equation $2Q\eta - \lambda^{-1}(\eta - h) = 0$. That is $\eta = (I - 2\lambda Q)^{-1}h$. Thus

$$\begin{aligned}\langle Q, hh^t\rangle^\lambda &= \langle Q, (I - 2\lambda Q)^{-1} hh^t (I - 2\lambda Q)^{-1}\rangle - \frac{1}{2\lambda}\|(I - (I - 2\lambda Q)^{-1})h\|^2 \\ &= \langle (I - 2\lambda Q)^{-1} Q (I - 2\lambda Q)^{-1} - \frac{1}{2\lambda}(I - (I - 2\lambda Q)^{-1})^2, hh^t\rangle = \langle Q^\lambda, hh^t\rangle.\end{aligned}$$

Next we verify the other identities. Noting that $(I - (I - 2\lambda Q)^{-1}) = 2\lambda Q(I - 2\lambda Q)^{-1}$ we find that

$$\begin{aligned}Q^\lambda &= (I - 2\lambda Q)^{-1} Q (I - 2\lambda Q)^{-1} - 2\lambda (I - 2\lambda Q)^{-1} Q^2 (I - 2\lambda Q)^{-1} \\ &= (I - 2\lambda Q)^{-1} (I - 2\lambda Q) Q (I - 2\lambda Q)^{-1} = Q(I - 2\lambda Q)^{-1}.\end{aligned}$$

The last identity $Q^\lambda = (I - 2\lambda Q)^{-1} Q$ follows from the symmetry of Q. Clearly if $I - 2\lambda Q \notin \operatorname{int} \mathcal{P}(n)$ we cannot find $(I - 2\lambda Q)^{-1}$.

One can verify that $Q = P(I + 2\lambda P)^{-1}$ solves the equation $P = Q^\lambda$ for any λ i.e.

$$P = Q(I - 2\lambda Q)^{-1}$$

implies

$$\begin{aligned}& Q = (I - 2\lambda Q)P \\ \text{and so}\quad & P = 2\lambda QP + Q = Q\,(I + 2\lambda P) \\ \text{or}\quad & Q = P\,(I + 2\lambda P)^{-1}.\end{aligned}$$

The identity $P(I + 2\lambda P)^{-1} = (I + 2\lambda P)^{-1} P$ follows from symmetry again. □

The following related result was proved in [9].

Proposition 5.7 ([9]). *Given $Q \in \mathcal{S}(n)$ denote by $q_Q(\eta)$ the quadratic form $\langle Q\eta, \eta\rangle$ and by $(q_Q)_\lambda(h)$ the form $\inf_\eta\{\langle Q\eta, \eta\rangle + \frac{1}{2\lambda} \|h - \eta\|^2\}$.*

1. *Let $\lambda > 0$. The condition that $I + 2\lambda Q \in \operatorname{int} \mathcal{P}(n)$ is necessary and sufficient for $(q_Q)_\lambda$ to be a quadratic form. More precisely $(q_Q)_\lambda = q_{Q_\lambda}$ with*

$$Q_\lambda := \frac{1}{2\lambda}\left(I - (I + 2\lambda Q)^{-1}\right) = Q(I + 2\lambda Q)^{-1} = (I + 2\lambda Q)^{-1} Q.$$

2. *If for a given $P \in \mathcal{S}(n)$ we have $Q_\lambda = P$, then $I - 2\lambda P \in \operatorname{int} \mathcal{P}(n)$ and $Q = P^\lambda$ where $P^\lambda := \frac{1}{2\lambda}((I - 2\lambda P)^{-1} - I) = P\,(I - 2\lambda P)^{-1} \in \mathcal{S}(n)$.*

Remark 5.8. Assume $I - 2\lambda P$ is invertible. $I = I$ may be written as

$$\begin{aligned} I &= I - 2\lambda P + 2\lambda P - 4\lambda^2 P^2 + 4\lambda^2 P^2 \\ &= (I + 2\lambda P)(I - 2\lambda P) + 4\lambda^2 P^2 \\ &= (I + 2\lambda P)(I - 2\lambda P) + 4\lambda^2 P^2 (I - 2\lambda P)^{-1}(I - 2\lambda P) \\ &= [I + 2\lambda P + 4\lambda^2 P^2 (I - 2\lambda P)^{-1}](I - 2\lambda P) \end{aligned}$$

This implies that

$$(I - 2\lambda P)^{-1} = I + 2\lambda P + 4\lambda^2 P^2 (I - 2\lambda P)^{-1}.$$

and so we get

$$P + 2\lambda P^2 (I - 2\lambda P)^{-1} = \frac{1}{2\lambda}[(I - 2\lambda P)^{-1} - I] = P^\lambda \quad \text{(by Proposition 5.7)}$$

and so

$$\left(\frac{1}{2}B\right)^\lambda = \frac{1}{2}\left(B + \lambda B^2 (I - \lambda B)^{-1}\right)$$

$$\text{and so} \quad 2\left(\frac{1}{2}B\right)^\lambda = \left(B + \lambda B^2 (I - \lambda B)^{-1}\right) \in \partial^{2,-} f^\lambda(\bar{x} - \lambda \bar{p}, \bar{p})$$

for all λ sufficiently small so that $I - \lambda B \in \operatorname{int} \mathcal{P}(n)$ (under the conditions of Corollary 5.5).

Rewording Corollary 5.5 we obtain the main result of this section.

Theorem 5.9. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be prox–bounded and para–concave, and suppose $\lambda < \frac{1}{\bar{r}(f)}$ (the prox–threshold for f) is such that both $\lambda^{-1} I - B$ is positive definite and $f - \frac{1}{2\lambda}\|\cdot\|^2$ is concave. Then (whenever $f^\lambda(\bar{x} - \lambda\bar{p}) < +\infty$)*

$$2\left(\frac{1}{2}B\right)^\lambda \in \partial^{2,-} f^\lambda(\bar{x} - \lambda\bar{p}, \bar{p}) \quad \textit{if and only if } B \in \partial^{2,-} f(\bar{x}, \bar{p})$$

in which case we have $f^\lambda(\bar{x} - \lambda\bar{p}) + \frac{\lambda}{2}\|\bar{p}\|^2 = f(\bar{x})$. In particular we have $\bar{p} \in \partial_p f(\bar{x})$ if and only if $\bar{p} \in \partial_p f^\lambda(\bar{x} - \lambda\bar{p})$.

6. Subhessians of the Lasry–Lions Double Envelope

In this section we will provide a full analysis of the subjet and the associated subhessians of the Lasry–Lions double envelope. The following Lemma will be used in the proof of the next main theorem of the paper.

Lemma 6.1. *Let f be prox–bounded with prox–threshold $\bar{r}(f)^{-1}$.*

1. *For any positive λ such that $\lambda < \frac{1}{2\bar{r}(f)}$, we have f_λ prox–bounded, with $\lambda < \frac{1}{\bar{r}(f_\lambda)}$ (the prox–threshold of f_λ). That is $2\bar{r}(f) \geq \bar{r}(f_\lambda)$ for all $0 < \lambda < \frac{1}{2\bar{r}(f)}$.*
2. *If $0 < \lambda < \frac{1}{2\bar{r}(f,\bar{x},p)}$ then $0 < \lambda < \frac{1}{\bar{r}(f_\lambda,\bar{x}+\lambda p,p)}$ (i.e. $2\bar{r}(f,\bar{x},p) \geq \bar{r}(f_\lambda, \bar{x} + \lambda p, p)$ for all $0 < \lambda < \frac{1}{2\bar{r}(f,\bar{x},p)}$).*

Proof. Let $\lambda < \frac{1}{2\bar{r}(f,\bar{x},p)}$ so there is $r > 2\bar{r}(f,\bar{x},p)$ such that $\lambda < \frac{1}{r} < \frac{1}{2\bar{r}(f,\bar{x},p)}$. Assume

$$f(y) \geq f(\bar{x}) + \langle p, y-\bar{x}\rangle - \frac{r}{2}\|y-\bar{x}\|^2 \text{ as } r > 2\bar{r}(f,\bar{x},p).$$

This implies

$$\begin{aligned} f_\lambda(x) &= \inf_y\{f(y) + \frac{1}{2\lambda}\|x-y\|^2\} \\ &\geq f(\bar{x}) + \inf_y\{\frac{1}{2\lambda}\|x-y\|^2 + \langle p, y-\bar{x}\rangle - \frac{r}{2}\|y-\bar{x}\|^2\} \end{aligned} \tag{21}$$

In (21), the minimum occurs when $\frac{1}{\lambda}(y-x) - r(y-\bar{x}) = -p$ and simple calculations give $y = \frac{x-\lambda(r\bar{x}+p)}{1-\lambda r}$. We have

$$\begin{aligned} &\frac{1}{2\lambda}\left\|x - \frac{x-\lambda(r\bar{x}+p)}{1-\lambda r}\right\|^2 + \langle p, \frac{x-\lambda(r\bar{x}+p)}{1-\lambda r} - \bar{x}\rangle - \frac{r}{2}\left\|\frac{x-\lambda(r\bar{x}+p)}{1-\lambda r} - \bar{x}\right\|^2 \\ =\ & \frac{1}{2\lambda}\|(x-\bar{x}) - \frac{1}{1-\lambda r}(x - (\bar{x}+\lambda p))\|^2 \\ &+ \left(\frac{1}{1-\lambda r}\right)\langle p, x-(\bar{x}+\lambda p)\rangle - \frac{r}{2}\left(\frac{1}{1-\lambda r}\right)^2\|x-(\bar{x}+\lambda p)\|^2 \\ =\ & \frac{1}{2\lambda}\|\lambda p - (x-(\bar{x}+\lambda p))\left(\frac{\lambda r}{1-\lambda r}\right)\|^2 \\ &+ \left(\frac{1}{1-\lambda r}\right)\langle p, x-(\bar{x}+\lambda p)\rangle - \frac{r}{2}\left(\frac{1}{1-\lambda r}\right)^2\|x-(\bar{x}+\lambda p)\|^2 \\ =\ & \frac{\lambda}{2}\|p\|^2 - \left(\frac{\lambda r}{1-\lambda r}\right)\langle p, (x-(\bar{x}+\lambda p))\rangle \\ &+ \left(\frac{1}{1-\lambda r}\right)\langle p, x-(\bar{x}+\lambda p)\rangle + \left(\left(\frac{1}{2\lambda}\right)(\lambda r)^2 - \frac{r}{2}\right)\left(\frac{1}{1-\lambda r}\right)^2\|x-(\bar{x}+\lambda p)\|^2 \\ =\ & \frac{\lambda}{2}\|p\|^2 + \langle p, x-(\bar{x}+\lambda p)\rangle - \frac{r}{2}\left(\frac{1}{1-\lambda r}\right)\|x-(\bar{x}+\lambda p)\|^2. \end{aligned}$$

Noting that by Lemma 3.2 and the fact that $\lambda < \frac{1}{r} < \frac{1}{2\bar{r}(f,\bar{x},p)}$ implies $(p, -\frac{1}{\lambda}I) \in \partial^{2,-}f(\bar{x})$ we have

$$f_\lambda(\bar{x}+\lambda p) = f(\bar{x}) + \frac{\lambda}{2}\|p\|^2.$$

Thus

$$\begin{aligned} f_\lambda(x) &\geq \left(f(\bar{x}) + \frac{\lambda}{2}\|p\|^2\right) + \langle p, x-(\bar{x}+\lambda p)\rangle - \frac{r}{2}\left(\frac{1}{1-\lambda r}\right)\|x-(\bar{x}+\lambda p)\|^2 \\ &= f_\lambda(\bar{x}+\lambda p) + \langle p, x-(\bar{x}+\lambda p)\rangle - \frac{r}{2}\left(\frac{1}{1-\lambda r}\right)\|x-(\bar{x}+\lambda p)\|^2 \\ &\geq f_\lambda(\bar{x}+\lambda p) + \langle p, x-(\bar{x}+\lambda p)\rangle - r\|x-(\bar{x}+\lambda p)\|^2 \\ &\qquad \left(\text{since } \lambda < \frac{1}{2r} \text{ and hence } \left(\frac{1}{1-\lambda r}\right) < 2\right) \end{aligned}$$

for all x whence $\bar{r}(f_\lambda) \le \frac{r}{1-\lambda r}$. Since we may allow $r \to \bar{r}$, we obtain $\bar{r}(f_\lambda) \le \frac{\bar{r}}{1-\lambda \bar{r}}$. If $\bar{r} = 0$, then $\bar{r}(f_\lambda, \bar{x}, p) = 0$ also (for all $\lambda > 0$). So trivially, $\lambda < \frac{1}{\bar{r}(f_\lambda, \bar{x}, p)}$ for all $\lambda > 0$. If $\bar{r} > 0$, observe that a sufficient condition for $\lambda < \frac{1}{\bar{r}(f_\lambda, \bar{x}, p)}$ is that $\lambda < \frac{1-\lambda\bar{r}}{\bar{r}}$, which requires $2\lambda\bar{r} < 1$ or $\lambda < \frac{1}{2\bar{r}}$. Thus, for $\lambda < \frac{1}{2\bar{r}}$, get $\lambda < \frac{1-\lambda\bar{r}}{\bar{r}} \le \frac{1}{\bar{r}(f_\lambda, \bar{x}, p)}$. The first assertion follows by the same argument with β replacing $f(\bar{x})$ and $p = 0$. □

Remark 6.2. Since $f_{\lambda|\mu}(x) = (f_{\lambda+\mu})^\mu(x) \ge f_{\lambda+\mu}(x)$ we find that if $0 < \lambda + \mu < \frac{1}{2\bar{r}(f)}$ then $0 < \lambda + \mu < \frac{1}{\bar{r}(f_{\lambda|\mu})}$ etc.

Theorem 6.3. *Suppose $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a lower semi–continuous, prox–bounded and proper function and $0 < \lambda < \bar{\lambda} = (\bar{r}(f))^{-1}$ (the prox threshold of f). Then $h_\lambda f + \frac{1}{2\lambda}\|\cdot\|^2$ is proper, convex and for all $\eta > 0$ the following function is also convex and finite–valued:*

$$\left(h_\lambda f + \frac{1}{2\lambda}\|\cdot\|^2\right)_\eta (x) = (h_\lambda f)_{\left(\frac{1}{\lambda}+\frac{1}{\eta}\right)^{-1}} \left(\frac{\lambda x}{\lambda + \eta}\right) + \frac{1}{2(\lambda+\eta)}\|x\|^2.$$

In particular we have $(h_\lambda f)_{\left(\frac{1}{\lambda}+\frac{1}{\eta}\right)^{-1}}$ para–convex and hence $C^{1,1}$ for all $\eta > 0$. Hence $f_{\lambda,\mu}$ for all $0 < \mu < \lambda$ and $f_{\lambda|\mu}$ for all $0 < \mu, \lambda$, are $C^{1,1}(\mathbb{R}^n)$ along with

$$f_{\lambda+\mu} \le f_{\lambda|\mu} := h_\mu(f_\lambda) = (h_{\lambda+\mu} f)_\lambda = (f_{\lambda+\mu})^\mu \le f_\lambda \le f\ . \tag{22}$$

In particular $\{f_{\lambda|\mu}\}_{0<\mu<\lambda}$ epiconverges (monotonically upwards) to f_λ as $\mu \downarrow 0$. The following hold:

1. *Let $0 < \mu$, λ where $\lambda + \mu < \frac{1}{2\bar{r}(f,x,p)}$ and $(p, Q) \in \partial^{2,-} f(x)$. If also $I + (\lambda+\mu)Q \in \operatorname{int}\mathcal{P}(n)$, then*

$$\left(p, 2\left(\frac{1}{2}Q\right)_\lambda\right) \in \partial^{2,-} f_{\lambda|\mu}(x + \lambda p) \quad \text{and} \quad f_{\lambda|\mu}(x+\lambda p) = f(x) + \frac{\lambda}{2}\|p\|^2.$$

2. *Let $0 < \mu$, λ where $\lambda + \mu < \frac{1}{2\bar{r}(f)}$ ($\bar{r}(f)$ the prox–threshold of f). Then, if $(p, Q) \in \partial^{2,-} f_{\lambda|\mu}(x)$ such that $I + \mu Q \in \operatorname{int}\mathcal{P}(n)$, then*

$$\left(p, 2\left(\frac{1}{2}Q\right)_\mu\right) \in \partial^{2,-} f(x - \lambda p)$$

$$\text{with} \quad f_{\lambda|\mu}(x) = f(x - \lambda p) + \frac{\lambda}{2}\|p\|^2.$$

3. *Let $0 < \mu$, λ with $\lambda + \mu < \frac{1}{2\bar{r}(f)}$. Then, if $(p, Q) \in \underline{\partial}^2 f_{\lambda|\mu}(x)$ with $I + \mu Q \in \operatorname{int}\mathcal{P}(n)$, then*

$$\left(p, 2\left(\frac{1}{2}Q\right)_\mu\right) \in \underline{\partial}^2 f(x - \lambda p)$$

$$\text{with} \quad f_{\lambda|\mu}(x) = f(x - \lambda p) + \frac{\lambda}{2}\|p\|^2.$$

4.

$$\partial f(x) = \limsup \Big\{ \nabla f_{\lambda|\mu}(x') \mid x' \to x,\ f_{\lambda|\mu}(x') \to f(x) \ \textit{and} \ (\lambda,\mu) \downarrow (0,0),\ \lambda,\mu > 0 \Big\}$$

and

$$\underline{\partial}^2 f(x) = \limsup \Big\{ (\nabla f_{\lambda|\mu}(x'), Q) \mid x' \in S_2(f_{\lambda|\mu}),\ x' \to x,\ f_{\lambda|\mu}(x') \to f(x);\ Q \leq_{\mathcal{P}(n)} \nabla^2 f_{\lambda|\mu}(x') \ \textit{and} \ (\lambda,\mu) \downarrow (0,0),\ \lambda,\mu > 0 \Big\}.$$

Proof. Note that $(h_\lambda f + \frac{1}{2\lambda}\|\cdot\|^2)_\eta$ is uniformly bounded below by $\inf_\lambda (h_\lambda f + \frac{1}{2\lambda}\|\cdot\|^2) > -\infty$ for all η, and never takes the value $+\infty$ from properness of $h_\lambda f \leq f \not\equiv +\infty$ so that $(h_\lambda f + \frac{1}{2\lambda}\|\cdot\|^2)_\eta(x)$ is finite for all x and all $\eta > 0$. It is well known that (see page 495 of [30])

$$h_\lambda f(x) = \left(f + \frac{1}{2\lambda}\|\cdot\|^2 \right)^{**}(x) - \frac{1}{2\lambda}\|x\|^2.$$

Thus $h_\lambda f(x) + \frac{1}{2\lambda}\|x\|^2$ is convex being equal to the second convex conjugate of $f + \frac{1}{2\lambda}\|\cdot\|^2$. Therefore it is bounded below, since $f + \frac{1}{2\lambda}\|\cdot\|^2$ bounded below from the assumption of f. It is well known that the infimal convolution of a proper, convex function is also proper convex and finite as $0 < \lambda < \bar{\lambda} = \bar{r}(f)^{-1}$ ($\bar{r}(f)$ is the prox–threshold of f). Let us now explicitly calculate this infimal convolution with parameter η , utilizing (18) with $c_1 = \frac{1}{\lambda}$, $c_2 = \frac{1}{\eta}$ and $y_1 = 0$, $y_2 = x$,

$$\begin{aligned}
\left(h_\lambda f + \frac{1}{2\lambda}\|\cdot\|^2 \right)_\eta (x) &= \inf_w \left\{ h_\lambda f(w) + \frac{1}{2\lambda}\|w\|^2 + \frac{1}{2\eta}\|x - w\|^2 \right\} \\
&= \inf_w \left\{ h_\lambda f(w) + \frac{1}{2\left(\frac{1}{\lambda} + \frac{1}{\eta}\right)^{-1}} \left\| w - \left(\frac{\lambda}{\lambda+\eta}\right) x \right\|^2 \right\} \\
&\quad + \frac{1}{2(\lambda+\eta)}\|x\|^2 \\
&= (h_\lambda f)_{\left(\frac{1}{\lambda}+\frac{1}{\eta}\right)^{-1}} \left(\frac{\lambda x}{\lambda+\eta}\right) + \frac{1}{2(\lambda+\eta)}\|x\|^2.
\end{aligned}$$

As $\left(h_\lambda f + \frac{1}{2\lambda}\|\cdot\|^2\right)_\eta$ is convex it follows that $x \mapsto (h_\lambda f)_{\left(\frac{1}{\lambda}+\frac{1}{\eta}\right)^{-1}}(x)$ is para–convex and hence simultaneously para–convex and para–concave as well as finite–valued. By Lemma 4.1 it must be $C^{1,1}$. We may place $\left(\frac{1}{\lambda} + \frac{1}{\eta}\right)^{-1} = \lambda - \mu$ and solve to obtain $\eta = \frac{\lambda(\lambda-\mu)}{\mu} > 0$ for all $\mu < \lambda$. In this case we have $(h_\lambda f)_{\lambda-\mu} = f_{\lambda,\mu}$ being $C^{1,1}$. Indeed this is true if we replace λ by $\lambda + \mu$ and obtain

$$f_{\lambda+\mu} \leq (f_{\lambda+\mu})^\mu = (h_{\lambda+\mu} f)_\lambda = h_\mu(f_\lambda) \leq f_\lambda \leq f$$

and so (22) follows as $h_\mu(f_\lambda) := f_{\lambda|\mu}$.

Part 1. For any $\xi < \frac{1}{\bar{r}} = \frac{1}{\bar{r}(f,x,p)}$, recall that $(f - \langle p, \cdot \rangle)_\xi$ is para–concave (with $(f - \langle p, \cdot\rangle)_\xi - \frac{1}{2\xi}\|\cdot\|^2$ concave), prox–bounded and finite–valued. Now use Lemma 3.2 which establishes that for any y we have $(f - \langle p, \cdot\rangle)_\xi (y) = f_\xi(y+\xi p) - \langle p, y\rangle - \frac{\xi}{2}\|p\|^2$. Consequently we have $\xi < \frac{1}{\bar{r}(f_\xi, x+\xi p, p)}$ for all $\xi < \frac{1}{2\bar{r}}$ (as observed in Lemma 6.1) along with $f_\xi(\cdot + \xi p) - \frac{1}{2\xi}\|\cdot\|^2$ concave, prox–bounded and finite–valued.

Let $\lambda, \mu > 0$ such that $\lambda + \mu < \frac{1}{2\bar{r}}$. If $(p, Q) \in \partial^{2,-} f(x)$ and $I + (\lambda + \mu)Q > 0$ then $I + \lambda Q > 0$ and there follows from Lemma 4.9 that

$$\begin{aligned} B &:= 2(\frac{1}{2}Q)_{\lambda+\mu} \in \partial^{2,-} f_{\lambda+\mu}(x + (\lambda+\mu)p, p) \\ \text{with} &\quad (f - \langle p, \cdot\rangle)_{\lambda+\mu}(x) = f(x) - \langle p, x\rangle \end{aligned} \tag{23}$$

We seek now to apply Theorem 5.9 to the function $x \mapsto f_{\lambda+\mu}(x + (\lambda+\mu)p)$ using a "parameter" μ. Note that

$$f_{\lambda+\mu}(\cdot + (\lambda+\mu)p) - \frac{1}{2\mu}\|\cdot\|^2 = \left\{ f_{\lambda+\mu}(\cdot + (\lambda+\mu)p) - \frac{1}{2(\lambda+\mu)}\|\cdot\|^2 \right\} - \frac{\lambda}{2\mu(\lambda+\mu)}\|\cdot\|^2$$

and hence is concave. Also $(f_{\lambda+\mu})^\mu = f_{\lambda|\mu}$ has already been shown to be finite-valued. Further,

$$\begin{aligned} I - \mu B &= I - 2\mu \left(\frac{1}{2}Q\right)_{\lambda+\mu} = I - \mu Q \left(I + (\lambda+\mu)Q\right)^{-1} \\ &= \left(I + (\lambda+\mu)Q - \mu Q\right)\left(I + (\lambda+\mu)Q\right)^{-1} = \left(I + \lambda Q\right)\left(I + (\lambda+\mu)Q\right)^{-1}. \end{aligned}$$

So, $I - \mu B > 0$ since $I + \lambda Q > 0$ and $\left(I + (\lambda+\mu)Q\right)^{-1} > 0$. Also, from Lemma 6.1, $\mu < \mu + \lambda < \frac{1}{\bar{r}(f_{\lambda+\mu}, x+(\lambda+\mu)p, p)}$. Thus, the conditions of Theorem 5.9 are satisfied, giving

$$\begin{aligned} 2\left(\frac{1}{2}Q\right)_\lambda = 2(\frac{1}{2}B)^\mu &\in \partial^{2,-}\left(f_{\lambda+\mu}\right)^\mu \left(x + (\lambda+\mu)p - \mu p, p\right) \\ &= \partial^{2,-} f_{\lambda|\mu}(x + \lambda p, p) \end{aligned}$$

along with

$$\begin{aligned} f_{\lambda|\mu}(x + \lambda p) &= f_{\lambda+\mu}(x + (\lambda+\mu)p) - \frac{\mu}{2}\|p\|^2 \\ &= f_{\lambda+\mu}(x + (\lambda+\mu)p) - \frac{\lambda+\mu}{2}\|p\|^2 + \frac{\lambda}{2}\|p\|^2 \\ &= \left(f - \langle p, \cdot\rangle\right)_{\lambda+\mu}(x) + \langle p, x\rangle + \frac{\lambda}{2}\|p\|^2 \quad \text{from Lemma 3.2} \\ &= f(x) + \frac{\lambda}{2}\|p\|^2 \quad \text{from (23) above.} \end{aligned}$$

Part 2. Place $B := 2\left(\frac{1}{2}Q\right)_\mu$ so $2\left(\frac{1}{2}B\right)^\mu = Q$ as may be easily shown. Then $2\left(\frac{1}{2}B\right)^\mu \in \partial^{2,-}\left(f_{\lambda+\mu}\right)^\mu(x)$. Noting that $I - \mu B = (I + \mu Q)^{-1} > 0$, we may, as we did for part 1, apply Theorem 5.9 to $f_{\lambda+\mu}$ and parameter μ, to obtain

$$\begin{aligned} 2\left(\frac{1}{2}Q\right)_\mu &= B \in \partial^{2,-} f_{\lambda+\mu}(x + \mu p, p) \\ \text{with} \quad f_{\lambda|\mu}(x) + \frac{\mu}{2}\|p\|^2 &= f_{\lambda+\mu}(x + \mu p)\ . \end{aligned}$$

From (4) applied to f (and parameter $\lambda+\mu$) which is permitted since $\lambda+\mu<\frac{1}{\bar{r}(f)}$, we obtain

$$B \in \partial^{2,-} f(x+\mu p-(\lambda+\mu)p, p)=\partial^{2,-} f(x-\lambda p, p)$$

$$\text{with} \quad f(x-\lambda p)=f_{\lambda+\mu}(x+\mu p)-\frac{\lambda+\mu}{2}\|p\|^2,$$

yielding

$$\left(p, 2\left(\frac{1}{2}Q\right)_\mu\right) \in \partial^{2,-} f(x-\lambda p)$$

$$\begin{aligned} \text{and} \quad f_{\lambda|\mu}(x) &= f_{\lambda+\mu}(x+\mu p)-\frac{\mu}{2}\|p\|^2 \\ &= f(x-\lambda p)+\frac{\lambda}{2}\|p\|^2 . \end{aligned}$$

Part 3. $\exists (p_m, Q_m) \to (p, Q)$ with $(p_m, Q_m) \in \partial^{2,-} f_{\lambda|\mu}(x_m)$, $x_m \to x$ (and $f_{\lambda|\mu}(x_m) \to f_{\lambda|\mu}(x)$). As $I+\mu Q>0$, have $I+\mu Q_m>0$ for all large m, so by Part 2,

$$\left(p_m, 2\left(\frac{1}{2}Q_m\right)_\mu\right) \in \partial^{2,-} f(x_m-\lambda p_m)$$

$$\text{and} \quad f_{\lambda|\mu}(x_m)=f(x_m-\lambda p_m)+\frac{\lambda}{2}\|p_m\|^2 \tag{24}$$

By continuity of the inversion operation $A \mapsto A^{-1}$ in $\operatorname{int}\mathcal{P}(n)$, we have $2\left(\frac{1}{2}Q_m\right)_\mu \to 2\left(\frac{1}{2}Q\right)_\mu$ as $m \to \infty$. Since $x_m-\lambda p_m \to x-\lambda p$, we may let $m \to \infty$ in (24) to obtain

$$\begin{aligned} f_{\lambda|\mu}(x) &= \lim_{m\to\infty} f(x_m-\lambda p_m)+\frac{\lambda}{2}\|p_m\|^2 \qquad (25) \\ &\geq f(x-\lambda p)+\frac{\lambda}{2}\|p\|^2 \quad \text{as } f \text{ is l.s.c.} \end{aligned}$$

However,

$$\begin{aligned} f_{\lambda|\mu}(x) &= \left(f_{\lambda+\mu}\right)^\mu(x)=\left(\left(f_\lambda\right)_\mu\right)^\mu(x)=h_\mu(f_\lambda)(x) \leq f_\lambda(x) \\ &= \inf_w\left(f(w)+\frac{1}{2\lambda}\|x-w\|^2\right) \leq f(x-\lambda p)+\frac{1}{2\lambda}\|x-(x-\lambda p)\|^2 \\ &= f(x-\lambda p)+\frac{\lambda}{2}\|p\|^2. \end{aligned}$$

Combining with the inequality in (25) gives $f(x-\lambda p)+\frac{\lambda}{2}\|p\|^2=f_{\lambda|\mu}(x)$ and $\lim_{m\to\infty} f(x_m-\lambda p_m)=f(x-\lambda p)$. Thus

$$\left(p, 2\left(\frac{1}{2}Q\right)_\mu\right)=\lim_m\left(p_m, 2\left(\frac{1}{2}Q_m\right)_m\right) \in \overline{\lim}_{\substack{x'\to x-\lambda p \\ f(x')\to f(x-\lambda p)}} \partial^{2,-} f(x')=\underline{\partial}^2 f(x-\lambda p).$$

Part 4. For $(p, Q) \in \partial^{2,-} f(x)$, observe that as $\lambda \downarrow 0$ with $\lambda, \mu > 0$ we obtain from Part 1 that

$$f_{\lambda|\mu}(x + \lambda p) = f(x) + \frac{\lambda}{2} \|p\|^2$$

and $(p, 2(\frac{1}{2}Q)_\lambda) \to (p, Q)$ when $\lambda \downarrow 0$ with $\lambda > \mu > 0$. Thus

$$\limsup \left\{\partial^{2,-} f_{\lambda|\mu}(x') \mid x' \to x,\ f_{\lambda|\mu}(x') \to f(x) \text{ and } (\lambda, \mu) \downarrow (0,0),\ \lambda,\mu > 0\right\} \supseteq \partial^{2,-} f(x). \tag{26}$$

We may also use Proposition 2.9 (as $f_{\lambda|\mu}$ is continuous and para–concave) to deduce that

$$\overline{D}^2 f_{\lambda|\mu}(x', \nabla f_{\lambda|\mu}(x')) - \mathcal{P}(n) = \underline{\partial}^2 f_{\lambda|\mu}(x', \nabla f_{\lambda|\mu}(x')) \supseteq \partial^{2,-} f_{\lambda|\mu}(x', \nabla f_{\lambda|\mu}(x')). \tag{27}$$

Thus it follows from (26) and (27) via a sequence diagonalization argument that

$$\limsup \Big\{(\nabla f_{\lambda|\mu}(x'), Q) \mid x' \in S_2(f_{\lambda|\mu}),\ x' \to x,\ f_{\lambda|\mu}(x') \to f(x); \\ Q \leq_{\mathcal{P}(n)} \nabla^2 f_{\lambda|\mu}(x') \text{ and } (\lambda, \mu) \downarrow (0,0),\ \lambda,\mu > 0\Big\} \supseteq \partial^{2,-} f(x).$$

As the left hand side of this containment is invariant with respect to upper limits in $x' \to_f x$ we conclude that it also contains $\underline{\partial}^2 f(x)$.

We now show the reverse containment. Suppose there exists a sequence $x_n \in S_2(f_{\lambda_n|\mu_n})$ with $x_n \to x$, $f_{\lambda_n|\mu_n}(x_n) \to f(x)$ as $\lambda_n \downarrow 0$ and $\lambda_n > \mu_n > 0$ along with $(p_n, Q_n) := (\nabla f_{\lambda_n|\mu_n}(x_n), \nabla^2 f_{\lambda_n|\mu_n}(x_n)) \to (p, Q)$. Since

$$(f_{\lambda_n+\mu_n})^{\mu_n} = \left((f_{\lambda_n})_{\mu_n}\right)^{\mu_n} = h_{\mu_n}(f_{\lambda_n}) = f_{\lambda_n|\mu_n},$$

$(p_n, Q_n) \in \partial^{2,-} f_{\lambda_n|\mu_n}(x_n) = \partial^{2,-} (f_{\lambda_n+\mu_n})^{\mu_n}(x_n)$ and $f_{\lambda_n+\mu_n}$ is para–concave we may apply Theorem 5.9 to deduce that

$$(f_{\lambda_n+\mu_n})^{\mu_n}((x_n + \mu_n p_n) - \mu_n p_n) + \frac{\mu_n}{2}\|p_n\|^2 = f_{\lambda_n+\mu_n}(x_n + \mu_n p_n) \tag{28}$$

and

$$2\left(\frac{1}{2}Q_n\right)_{\mu_n} \in \partial^{2,-} f_{\lambda_n+\mu_n}(x_n + \mu_n p_n, p_n). \tag{29}$$

Theorem 5.9 is applicable provided that $2\left(\left(\frac{1}{2}Q_n\right)_{\mu_n}\right)^{\mu_n} = Q_n$, $I - 2\mu_n\left(\frac{1}{2}Q_n\right)_{\mu_n} \in \operatorname{int}\mathcal{P}(n)$ and $f_{\lambda_n+\mu_n} - \frac{1}{2\mu_n}\|\cdot\|^2$ is concave. The last of these conditions is always satisfied for $\lambda_n + \mu_n < (\bar{r}(f))^{-1}$ the prox–threshold of f since $f_{\lambda_n+\mu_n} - \frac{1}{2(\mu_n+\lambda_n)}\|\cdot\|^2$ is concave and thus so is

$$f_{\lambda_n+\mu_n} - \frac{1}{2\mu_n}\|\cdot\|^2 = \left\{f_{\lambda_n+\mu_n} - \frac{1}{2(\mu_n+\lambda_n)}\|\cdot\|^2\right\} - \frac{\lambda_n}{2(\lambda_n+\mu_n)\mu_n}\|\cdot\|^2.$$

Also, $f_{\lambda_n+\mu_n}$ is prox–bounded if $\lambda_n + \mu_n < \frac{1}{2r}$ (by Lemma 6.1) and is continuous. The first two conditions are satisfied as follows. Place $B_n := 2\left(\frac{1}{2}Q_n\right)_{\mu_n} = Q_n(1 + \mu_n Q_n)^{-1}$ the last equality following for all large n since $I + \mu_n Q_n \to I$ as $n \to \infty$ from boundedness of the Q_n. Note also that $(I + \mu_n Q_n)^{-1} \to I$ by continuity of the mapping $A \mapsto A^{-1}$ on $\operatorname{int} \mathcal{P}(n)$. We have $Q_n = 2\left(\frac{1}{2}B_n\right)^{\mu_n}$ and to apply Theorem 5.9, we need to check that $I - \mu_n B_n > 0$ eventually — indeed, $I - \mu_n B_n = I - \mu_n Q_n (I - \mu_n Q_n)^{-1} \to I$ as $\mu_n \to 0$.

Having established (29) we may now couple it with (4) (for n so large that $\lambda_n + \mu_n < \frac{1}{r}$) to obtain

$$\begin{aligned}\left(p_n, 2\left(\frac{1}{2}Q_n\right)_{\mu_n}\right) &\in \partial^{2,-} f(x_n + \mu_n p_n - (\lambda_n + \mu_n)p_n) \\ &= \partial^{2,-} f(x_n - \lambda_n p_n) \qquad (30)\end{aligned}$$

$$\text{and} \quad f(x_n - \lambda_n p_n) = f_{\lambda_n + \mu_n}(x_n + \mu_n p_n) - \frac{(\lambda_n + \mu_n)}{2}\|p_n\|^2.$$

Using (28) we have

$$\begin{aligned} f(x_n - \lambda_n p_n) &= f_{\lambda_n | \mu_n}(x_n) + \frac{\mu_n}{2}\|p_n\|^2 - \frac{(\lambda_n + \mu_n)}{2}\|p_n\|^2 \\ &= f_{\lambda_n | \mu_n}(x_n) - \frac{\lambda_n}{2}\|p_n\|^2. \qquad (31)\end{aligned}$$

It follows that when $f_{\lambda_n|\mu_n}(x_n) \to f(x)$ as $(\lambda_n, \mu_n) \downarrow (0,0)$, $\lambda_n, \mu_n > 0$ and $(p_n, Q_n) \to (p, Q)$ we have $f(x_n - \lambda_n p_n) \to f(x)$ and hence (since $2\left(\frac{1}{2}Q_n\right)_{\mu_n} \to Q$)

$$(p, Q) \in \limsup \left\{\partial^{2,-} f(x') \mid x' \to x,\ f(x') \to f(x)\right\} \subseteq \underline{\partial}^2 f(x).$$

This has shown that

$$\begin{aligned}\limsup \Big\{(\nabla f_{\lambda|\mu}(x'), \nabla^2 f_{\lambda|\mu}(x')) \mid x' \in S_2(f_{\lambda|\mu}),\ x' \to x, \\ f_{\lambda|\mu}(x') \to f(x);\ (\lambda, \mu) \downarrow (0,0),\ \lambda,\mu > 0\Big\} \subseteq \underline{\partial}^2 f(x).\end{aligned}$$

This implies

$$\begin{aligned}\limsup \Big\{(\nabla f_{\lambda|\mu}(x'), \nabla^2 f_{\lambda|\mu}(x')) \mid x' \in S_2(f_{\lambda|\mu}),\ x' \to x,\ f_{\lambda|\mu}(x') \to f(x);\ (\lambda, \mu) \downarrow (0,0); \\ \lambda,\mu > 0\Big\} - 0 \times \mathcal{P}(n) \subseteq \underline{\partial}^2 f(x) \text{ as } -(0 \times \mathcal{P}(n)) \subseteq \operatorname{rec} \underline{\partial}^2 f(x),\end{aligned}$$

for all x.

Now fix $\bar{x}$ and take limits $x \to \bar{x}$, $f(x) \to f(\bar{x})$ and note that

$$\limsup_{x \to^f \bar{x}} \underline{\partial}^2 f(x) = \underline{\partial}^2 f(\bar{x}).$$

Then

$$\begin{aligned}\limsup_{x \to^f \bar{x}} \Big\{ \limsup \Big\{(\nabla f_{\lambda|\mu}(x'), \nabla^2 f_{\lambda|\mu}(x')) \mid x' \in S_2(f_{\lambda|\mu}),\ x' \to x, \\ f_{\lambda|\mu}(x') \to f(x);\ (\lambda, \mu) \downarrow (0,0);\ \lambda, \mu > 0\Big\} - 0 \times \mathcal{P}(n)\Big\} \subseteq \underline{\partial}^2 f(\bar{x}). \qquad (32)\end{aligned}$$

Now the left hand side of (32) contains all sequences of the form

$$\left(\nabla f_{\lambda_n|\mu_n}(x_n), Q_n\right) \text{ with } Q_n \leq_{\mathcal{P}(n)} \nabla^2 f_{\lambda_n|\mu_n}(x_n)$$

for $x_n \in S_2(f_{\lambda_n|\mu_n})$, $x_n \to \bar{x}$, $f_{\lambda_n|\mu_n}(x_n) \to f(\bar{x})$ as $(\lambda_n, \mu_n) \downarrow (0,0)$ with $\lambda_n, \mu_n > 0$ and hence all accumulation points of such sequences. This implies equality in (32). □

As is usual we will require the characterization of superjets and so define

$$f^{\lambda|\mu}(x) = -\left(-f\right)_{\lambda|\mu}.$$

As $\partial^{2,+} f(x) = -\partial^{2,-}(-f)(x)$ the following is immediate from Theorem 6.3.

Theorem 6.4. *Suppose $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is an upper semi–continuous proper function (i.e. f is nowhere $+\infty$ and not identically $-\infty$) and $0 < \lambda < \bar{\lambda} = (\bar{r}(-f))^{-1}$ ($\bar{r}(-f)^{-1}$ is the prox–threshold of $-f$). Then $f^{\lambda|\mu}$ are $C^{1,1}(\mathbb{R}^n)$ for all $0 < \mu < \lambda$ along with*

$$f^{\lambda+\mu} \geq f^{\lambda|\mu} = \left(f^{\lambda+\mu}\right)_{\mu} \geq f^{\lambda} \ \geq f. \tag{33}$$

In particular $\left\{f^{\lambda|\mu}\right\}_{0<\mu<\lambda}$ hypoconverges (monotonically) to f^{λ} as $\mu \downarrow 0$. The following hold:

1. *Let $0 < \mu, \lambda$ where $\lambda + \mu < \frac{1}{2\bar{r}(-f,x)}$. Let $(p, Q) \in \partial^{2,+} f(x)$. If also $I - (\lambda+\mu)Q \in \operatorname{int}\mathcal{P}(n)$, Then*

$$(p, 2(\frac{1}{2}Q)^{\lambda}) \in \partial^{2,+} f^{\lambda|\mu}(x - \lambda p) \quad \textit{and} \quad f^{\lambda|\mu}(x - \lambda p) = f(x) - \frac{\lambda}{2}\|p\|^2.$$

2. *Let $0 < \mu$, λ where $\lambda + \mu < \frac{1}{2\bar{r}(-f)}$ (where $(\bar{r}(-f))^{-1}$ is the prox–threshold of $-f$). Then, if $(p, Q) \in \partial^{2,+} f^{\lambda|\mu}(x)$ such that $I - \mu Q \in \operatorname{int}\mathcal{P}(n)$, then*

$$\left(p, 2\left(\frac{1}{2}Q\right)^{\mu}\right) \in \partial^{2,+} f(x + \lambda p)$$

$$\textit{with} \quad f^{\lambda|\mu}(x) = f(x + \lambda p) - \frac{\lambda}{2}\|p\|^2.$$

3. *Let $0 < \mu$, $\lambda : \lambda + \mu < \frac{1}{2\bar{r}(-f)}$ (where $(\bar{r}(-f))^{-1}$ prox–threshold of $-f$). Then*

$$(p, Q) \in \overline{\partial}^2 f^{\lambda|\mu}(x) \quad \textit{and} \quad I - \mu Q > 0 \quad \Longrightarrow \quad \left(p, 2\left(\frac{1}{2}Q\right)^{\mu}\right) \in \overline{\partial}^2 f(x + \lambda p)$$

$$\textit{and} \quad f(x + \lambda p) = f^{\lambda|\mu}(x) + \frac{\lambda}{2}\|p\|^2.$$

4.

$$\overline{\partial}^2 f(x) = \limsup \Big\{(\nabla f^{\lambda|\mu}(x'), Q) \mid x' \in S_2(f^{\lambda|\mu}),\ x' \to x,$$
$$f^{\lambda|\mu}(x') \to f(x);\ Q \geq_{\mathcal{P}(n)} \nabla^2 f^{\lambda|\mu}(x') \ \textit{and}\ (\lambda, \mu) \downarrow (0,0);\ \lambda, \mu > 0\Big\}.$$

We conclude this section with a definition of a quantity which is suggested by these formula but we will leave its analysis to a later paper. The motivation for defining this quantity is to remove some of the negative semi–definite recession directions that occur in the limiting subhessian.

Definition 6.5. Suppose $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a lower semi–continuous, prox–bounded and proper function.

1. The generalized subhessians of f at x for $p \in \partial f(x)$ are denoted by:

$$\partial^2 f(x,p) = \limsup \Big\{ \nabla^2 f_{\lambda|\mu}(x') \mid x' \in S_2(f_{\lambda|\mu}),\ x' \to x,\ f_{\lambda|\mu}(x') \to f(x); \\ \nabla f_{\lambda|\mu}(x') \to p \text{ and } (\lambda,\mu) \downarrow (0,0),\ \lambda,\mu > 0 \Big\}.$$

2. The singular generalized subhessians of f at x for $p \in \partial f(x)$ are denoted by:

$$\partial^{2,\infty} f(x,p) = \limsup \Big\{ \gamma \nabla^2 f_{\lambda|\mu}(x') \mid x' \in S_2(f_{\lambda|\mu}),\ x' \to x,\ f_{\lambda|\mu}(x') \to f(x); \\ \nabla f_{\lambda|\mu}(x') \to p;\ 0 < \gamma \le \max\{\lambda,\mu\} \text{ and } (\lambda,\mu) \downarrow (0,0),\ \lambda,\mu > 0 \Big\}.$$

Of course a generalized super–hessian could be defined along similar lines.

7. Appendix

The proof of Proposition 4.3 is provided in this appendix. It is modeled on that given in [9, Proposition 8]. Throughout the proof it is assumed that $\bar{r}(f) = 0$ (as f has a global minimum at $\bar{x}$) and for a given $Q \in \partial^{2,-} f(\bar{x}, 0)$ we have chosen $\lambda > 0$ to satisfy $I + \lambda Q \in \operatorname{int} \mathcal{P}(n)$.

Proof of Proposition 4.3. Without loss of generality we may assume that $\bar{x} = 0$ and $f(0) = 0$ since $(f(\cdot - \bar{x}) - f(0))_\lambda(x) = f_\lambda(x - \bar{x}) - f(0)$. Now take any $(0, Q) \in \partial^{2,-} f(0)$. As $f \ge f(0)$ has a global minimum at 0, by Lemma 4.1 of [12] there exists a function $\hat{\varepsilon} : \mathbb{R}_+ \to \mathbb{R}$ such that $\lim_{t\downarrow 0} \hat{\varepsilon}(t) = 0 = \hat{\varepsilon}(0)$ and a function $r : y \mapsto \hat{\varepsilon}(\|y\|)\|y\|^2 \in \mathcal{C}^2(\mathbb{R}^n)$ satisfying $(\nabla r(0), \nabla^2 r(0)) = (0, 0)$ and

$$f(y) \ge \frac{1}{2}\langle Q, yy^t \rangle - \hat{\varepsilon}(\|y\|)\|y\|^2. \tag{34}$$

Next note that

$$0 < 1 + \lambda \min\{\mu \mid \mu \text{ is an eigenvalue of } Q\} \quad \Leftrightarrow \quad I + \lambda Q \in \operatorname{int} \mathcal{P}(n) \ .$$

Now fix $\lambda > 0$ such that $I + \lambda Q \in \operatorname{int} \mathcal{P}(n)$. Place

$$\max_{\eta \in \overline{B}_\delta(0)} \hat{\varepsilon}(\|\eta\|) = G(\delta) < \infty$$

which is finite as $r \mapsto \hat{\varepsilon}(r)$ is continuous and $G(\delta) \to 0$ as $\delta \to 0$. If necessary reduce δ so that

$$0 < 1 + \lambda(\min\{\mu \mid \mu \text{ is an eigenvalue of } Q\} - 2G(\delta)). \tag{35}$$

Let $b(\cdot)$ be a $\mathcal{C}^2$–smooth bump function with $b(x) = 1$ for $x \in B_{\frac{\delta}{2}}(0)$ and $b(x) = 0$ for $x \notin B_\delta(0)$. Place

$$\varepsilon(x) = b(x)\hat{\varepsilon}(x) + (1 - b(x))G(\delta)$$

$$\text{and} \quad \varphi(y) := \frac{1}{2}\langle Q, yy^t\rangle - \varepsilon(\|y\|)\|y\|^2.$$

Then clearly $\varphi \in \mathcal{C}^2(\mathbb{R}^n)$, $f(y) \geq \varphi(y)$ for $y \in B_\delta(0)$ and $f - \varphi$ takes a minimum over $B_\delta(0)$ at $\bar{x} = 0$ with $\varphi(0) = f(0) = 0$, $\nabla\varphi(0) = 0$ and $\nabla^2\varphi(0) = Q$.

Let us now recall that by the results of [2] the infimal convolution is locally Lipschitz in that

$$f_\lambda(y) - f_\lambda(x) \leq \lambda^{-1}\kappa_x\|y - x\|$$

where an explicit expression for κ_x is given by

$$\kappa_x = \left(\|y - x\| + \left[\frac{1 + 4\alpha_0\lambda}{1 - 4\alpha_0\lambda}\|x\|^2 + \frac{2\lambda}{1 - 4\alpha_0\lambda}\alpha_1\right]^{\frac{1}{2}}\right)$$

when $f(\cdot) + \alpha_0\|\cdot\|^2 + \alpha_1 \geq 0$ and $0 < \lambda < \frac{1}{4\alpha_0}$. In our case we may take $\alpha_0 = 0$, $\alpha_1 = f(0) = 0$, $x = 0$ and so write $\kappa_0 = \|y\|$ to obtain

$$f_\lambda(y) - f_\lambda(0) \leq \lambda^{-1}\|y\|^2, \tag{36}$$

for all $\lambda > 0$. Now consider $\eta \in P_\lambda(y)$. Then we have since $x = 0$ is a global minimum we have $f_\lambda(0) = f(0)$ and

$$f_\lambda(y) = f(\eta) + \frac{1}{2\lambda}\|y - \eta\|^2 \geq f(0) + \frac{1}{2\lambda}\|y - \eta\|^2$$

$$\text{implying} \quad \sqrt{2\lambda(f_\lambda(y) - f_\lambda(0))} \geq \|y - \eta\|.$$

Using (36) we obtain the bound for $\eta \in P_\lambda(y)$ to be $\sqrt{2}\|y\| \geq \|\eta - y\|$ and so for $y \in B_{\frac{\delta}{2\sqrt{2}}}(0)$ we have $\|\eta - y\| \leq \frac{\delta}{2}$ and hence $\eta \in B_\delta(0)$.

This implies (as $\varphi \leq f$ on $B_\delta(0)$)

$$f(\eta) + \frac{1}{2\lambda}\|y - \eta\|^2 \geq \varphi(\eta) + \frac{1}{2\lambda}\|y - \eta\|^2 \quad \text{for all } \eta \in P_\lambda(y) \subseteq B_\delta(0)$$

$$\text{giving} \quad f_\lambda(y) \geq \varphi_\lambda(y) \quad \text{for all } y \in B_{\frac{\delta}{2\sqrt{2}}}(0).$$

In order for φ_λ to be finite valued we require $\lambda < \bar{r}(\varphi)^{-1}$. It is easily seen that $\bar{r}(\varphi) = \max\{-\mu \mid \mu \text{ is an eigenvalue of } Q\} + 2G(\delta)$, since $\frac{1}{2}\langle Q - (2G(\delta) + \gamma)I, yy^t\rangle$ locally minorizes φ for all $\gamma > 0$. In particular $\lambda < \bar{r}(\varphi)^{-1}$ is equivalent to demanding (35). Thus $0 = f(0) \geq f_\lambda(0) = \varphi_\lambda(0) = \inf_\eta\{\varphi(\eta) + \frac{1}{2\lambda}\|\eta\|^2\} > -\infty$. We now show that we have $\varphi_\lambda(0) = 0$. Now

$$\varphi(x) = \frac{1}{2}\langle Qx, x\rangle - \varepsilon(\|x\|)\|x\|^2 \tag{37}$$

$$\text{implying} \quad \varphi_\lambda(x) = \inf_\eta\{\frac{1}{2}\langle Q\eta, \eta\rangle - \varepsilon(\|\eta\|)\|\eta\|^2 + \frac{1}{2\lambda}\|x - \eta\|^2\}.$$

We claim that $\eta \mapsto \frac{1}{2}\langle Q\eta,\eta\rangle - \varepsilon(\|\eta\|)\|\eta\|^2 + \frac{\lambda}{2}\|\eta\|^2$ has a minimum at $\eta = 0$ implying $\varphi_\lambda(0) \geq 0$.

By the definition of ε on $B_\delta(0)$ we have φ minorized by

$$\bar{p}(\eta) := \frac{1}{2}\langle Q\eta,\eta\rangle + (\frac{1}{2\lambda} - G(\delta))\|\eta\|^2$$

which by (35) is a positive, positive–definite form. Outside $B_\delta(0)$ we have $\bar{p} = \varphi$. Thus $\frac{1}{2}\langle Q\eta,\eta\rangle - \varepsilon(\|\eta\|)\|\eta\|^2 + \frac{1}{2\lambda}\|\eta\|^2$ is bounded below by a positive function $\bar{p}$ with a minimum of 0 at $\eta = 0$. Thus $\varphi_\lambda(0) = f_\lambda(0) = 0$ and so $f_\lambda - \varphi_\lambda$ has a local minimum (over $B_{\frac{\delta}{2\sqrt{2}}}$) at 0. Thus $(\nabla\varphi_\lambda(0), \nabla^2\varphi_\lambda(0)) \in \partial^{2,-} f_\lambda(0)$. Also (using the variable substitution $t\eta$ for η)

$$\begin{aligned}\frac{\varphi_\lambda(th)}{t} &= \left(\frac{1}{t}\right)\inf_\eta\{\varphi(\eta) + \frac{1}{2\lambda}\|th-\eta\|^2\}\\ &= \inf_\eta\{\frac{\varphi(t\eta)}{t} + \frac{t^2}{2t\lambda}\|h-\eta\|^2\} = \inf_\eta\{\frac{\varphi(t\eta)}{t} + \frac{t}{2\lambda}\|h-\eta\|^2\}\end{aligned}$$

and using

$$\begin{aligned}\varphi(t\eta) &= \varphi(0) + \langle\nabla\varphi(0), t\eta\rangle + \frac{1}{2}\langle\nabla^2\varphi(0)(t\eta), t\eta\rangle + \varepsilon(t\|\eta\|)t^2\|\eta\|^2\\ &= t^2(\frac{1}{2}\langle Q\eta,\eta\rangle + \varepsilon(t\|\eta\|)\|\eta\|^2)\end{aligned}$$

we get

$$\frac{\varphi_\lambda(th)}{t} = t\inf_\eta\{\frac{1}{2}\langle Q\eta,\eta\rangle + \varepsilon(t\|\eta\|)\|\eta\|^2 + \frac{1}{2\lambda}\|h-\eta\|^2\}. \tag{38}$$

Thus for the given λ we have

$$\begin{aligned}\lim_{t\downarrow 0} 2\frac{\varphi_\lambda(th)}{t^2} &= 2\lim_{t\downarrow 0}\left(\frac{1}{t^2}\right)\inf_\eta\{\varphi(\eta) + \frac{1}{2\lambda}\|th-\eta\|^2\}\\ &= 2\lim_{t\downarrow 0}\left(\frac{1}{t^2}\right)\inf_\eta\{\frac{1}{2}\langle Q(t\eta),(t\eta)\rangle + \varepsilon(t\|\eta\|)\|t\eta\|^2 + \frac{t^2}{2\lambda}\|h-\eta\|^2\}\\ &= 2\lim_{t\downarrow 0}\left(\inf_\eta\{\frac{1}{2}\langle Q\eta,\eta\rangle + \varepsilon(t\|\eta\|)\|\eta\|^2 + \frac{1}{2\lambda}\|h-\eta\|^2\}\right). \qquad (39)\end{aligned}$$

There is a compact set $C = \overline{B}_K(0)$ such that for all $t > 0$ sufficiently small we have $\inf_{\eta\in C} g_t(\eta) = \inf_\eta g_t(\eta)$ where $g_t(\eta) = \{\frac{1}{2}\langle Q\eta,\eta\rangle + \varepsilon(t\|\eta\|)\|\eta\|^2 + \frac{1}{2\lambda}\|h-\eta\|^2\}$. To see this note that

$$\max_{\eta\in\overline{B}_{\left(\frac{\delta}{t}\right)}(0)} \varepsilon(\|t\eta\|) = \max_{\eta\in\overline{B}_\delta(0)} \varepsilon(\|\eta\|) = G(\delta)$$

as before. Outside $\overline{B}_{\left(\frac{\delta}{t}\right)}(0)$ the function $\varepsilon(\|t\eta\|)$ is constant and so we have $\varepsilon(\|t\eta\|) = G(\delta)$. Thus

$$\begin{aligned}\frac{1}{2}\langle Q\eta,\eta\rangle + \varepsilon(t\|\eta\|)\|\eta\|^2 + \frac{1}{2\lambda}\|h-\eta\|^2 &\geq \frac{1}{2}\langle Q\eta,\eta\rangle + \varepsilon(t\|\eta\|)\|\eta\|^2 + \frac{1}{2\lambda}\left(\|\eta\|^2 - \langle h,\eta\rangle\right)\\ &\geq \bar{p}(\eta) - \frac{1}{2\lambda}\langle h,\eta\rangle := \bar{k}(\eta).\end{aligned}$$

As $\bar{k}(\eta) = \bar{p}(\eta) - \frac{1}{2\lambda}\langle h, \eta\rangle$ we have $\bar{k}$ coercive. Thus $\text{lev}_\alpha g_t := \{g_t \le \alpha\} \subseteq \text{lev}_\alpha \bar{k} \subseteq B_{\hat{K}}(0)$ for all $t > 0$ sufficiently small and some fixed $\hat{K} > 0$.

Now as $t \downarrow 0$ we have $g_t(\eta)$ converging to $g(\eta) = \frac{1}{2}\langle Q\eta, \eta\rangle + \frac{1}{2\lambda}\|h - \eta\|^2$ uniformly on bounded sets. It follows that $\{g_t(\cdot)\}_{t>0}$ epi-converges to $g(\cdot)$. Using Proposition 3.36 of [29] we have the marginal mapping $t \mapsto \inf_\eta g_t(\eta)$ continuous at $t = 0$.

From this we are first able to deduce by (38) that $\nabla\varphi_\lambda(0) = 0$. Using the continuity of the marginal mapping $t \mapsto \inf_\eta g_t(\eta)$ at $t = 0$ we have

$$\begin{aligned}\langle \nabla^2\varphi_\lambda(0)h, h\rangle &= 2\lim_{t\downarrow 0}\frac{\varphi_\lambda(th)}{t^2} = 2\lim_{t\downarrow 0}\inf_\eta g_t(\eta) = \inf_\eta g(\eta) \\ &= 2\inf_\eta\{\langle\frac{1}{2}Q\eta, \eta\rangle + \frac{1}{2\lambda}\|h - \eta\|^2\} = \langle 2(\frac{1}{2}Q)_\lambda, hh^t\rangle \\ \text{implying} \quad \langle \nabla^2\varphi_\lambda(0)h, h\rangle &= \inf_\eta\{\langle Q\eta, \eta\rangle + \frac{1}{\lambda}\|h - \eta\|^2\} \le \langle Q, hh^t\rangle.\end{aligned}$$

Since this is true for all h it follows that $\nabla^2\varphi_\lambda(0) = Q_\lambda$. Finally we see that if $\lambda_1 \le \lambda_2$ we have $\{\langle Q\eta, \eta\rangle + \frac{1}{\lambda_1}\|h - \eta\|^2\} \ge \{\langle Q\eta, \eta\rangle + \frac{1}{\lambda_2}\|h - \eta\|^2\}$ implying $\langle Q, hh^t\rangle_{\lambda_1} \ge \langle Q, hh^t\rangle_{\lambda_2}$. □

We will now provide a proof of Proposition 4.7 but require the following results in the proof. Once again for completeness we provide a proof that gives an explicit bound on $\lambda > 0$. Let $E(\mathcal{A}, u) := \{Q \in \mathcal{A} \mid \langle\frac{1}{2}Qu, u\rangle = q(\frac{1}{2}\mathcal{A})(u)\}$.

Proposition 7.1 ([12, Proposition 3.2]). *Suppose $\mathcal{A}$ is a non–empty rank-1 representer. If $Q \in E(\mathcal{A}, u)$ then when $I + \lambda Q \in \operatorname{int}\mathcal{P}(n)$ we have $Q_\lambda \in E(\mathcal{A}_\lambda, h_\lambda)$ where $h_\lambda = (I + \lambda Q)\, u \to u$ as $\lambda \to 0$ and $q(\mathcal{A}_\lambda)(h_\lambda) = q(\mathcal{A})(u) + \lambda\|Qu\|^2$.*

Proof. When $I + \lambda Q \in \operatorname{int}\mathcal{P}(n)$ we have Q_λ well defined. Since $\langle\frac{1}{2}Q, hh^t\rangle \le q(\frac{1}{2}\mathcal{A})(h)$ for all h, application of the infimal convolution to both sides of this inequality provides $\langle 2(\frac{1}{2}Q)_\lambda, uu^t\rangle \le 2q((\frac{1}{2}\mathcal{A})_\lambda)(u)$ for all u. Thus $Q_\lambda \in \mathcal{A}_\lambda$. The matrix $\langle\frac{1}{2}Q, \eta\eta^t\rangle + (2\lambda)^{-1}\|\eta\|^2$ is positive definite and so the problem $\inf_\eta\left\{\langle\frac{1}{2}Q, \eta\eta^t\rangle + \frac{1}{2\lambda}\|h - \eta\|^2\right\}$ has a unique solution at $\eta = (I + \lambda Q)^{-1} h$. In particular, for any fixed $h \in \mathbb{R}^n$, we have $h_\lambda := (I + \lambda Q)\, h$ has $h = (I + \lambda Q)^{-1} h_\lambda$. Thus

$$\begin{aligned}\langle 2(\frac{1}{2}Q)_\lambda, h_\lambda h_\lambda^t\rangle &= 2\left(\langle\frac{1}{2}Q, hh^t\rangle + \frac{1}{2\lambda}\|h_\lambda - h\|^2\right) \\ &= \langle Q, hh^t\rangle + \frac{1}{\lambda}\|\left(I - (I + \lambda Q)\right)h\|^2 = \langle Q, hh^t\rangle + \lambda\|Qh\|^2.\end{aligned}$$

When $Q \in E(\mathcal{A}, h)$ we have $q(\mathcal{A})(h) = \langle Q, hh^t\rangle$ and so

$$\begin{aligned}\langle Q_\lambda, h_\lambda h_\lambda^t\rangle &= \langle Q, hh^t\rangle + \lambda\|Qh\|^2 \\ &= 2\left(q(\frac{1}{2}\mathcal{A})(h) + \frac{1}{2\lambda}\|h_\lambda - h\|^2\right) \ge q(\mathcal{A}_\lambda)(h_\lambda),\end{aligned}$$

which implies $Q_\lambda \in E(\mathcal{A}_\lambda, h_\lambda)$. □

We may now provide a proof of Proposition 4.7. Recall that when $-\mathcal{P}(n) \subseteq 0^+\mathcal{A}$ we must have $\operatorname{int}\mathcal{A} \neq \emptyset$.

Proof of Proposition 4.7. When $\langle Q, hh^t\rangle_\lambda = \langle Q_\lambda, hh^t\rangle$ is a quadratic form using $\langle Q, uu^t\rangle \leq q(\mathcal{A})(u)$ for all u we have

$$2\langle (\frac{1}{2}Q)_\lambda, hh^t\rangle = 2\langle \frac{1}{2}Q, hh^t\rangle_\lambda \leq 2q_\lambda(\frac{1}{2}\mathcal{A})(h) = q(\mathcal{A}_\lambda)(h)$$

implying $2(\frac{1}{2}Q)_\lambda \in \mathcal{A}_\lambda$ and so $2(\frac{1}{2}Q)_\lambda - \mathcal{P}(n) \subseteq \mathcal{A}_\lambda$. We thus arrive at

$$\begin{aligned}\mathcal{A}_\lambda = (\mathcal{A}_\lambda)^1 &\supseteq \left(\{2(\frac{1}{2}Q)_\lambda \mid Q \in \mathcal{A} \text{ and } 2(\frac{1}{2}Q)_\lambda \text{ is a quadratic form}\} - \mathcal{P}(n)\right)^1 \\ &\supseteq \operatorname{cl}\{2(\frac{1}{2}Q)_\lambda \mid Q \in \mathcal{A} \text{ and } 2(\frac{1}{2}Q)_\lambda \text{ is a quadratic form}\} - \mathcal{P}(n)\end{aligned}$$

as $\mathcal{A}_\lambda$ is closed.

Now suppose there exists a $P \in \mathcal{A}_\lambda$ such that $2(\frac{1}{2}A)_\lambda = P$ for some A and $P \notin \{2(\frac{1}{2}Q)_\lambda \mid Q \in \mathcal{A}$ and $2(\frac{1}{2}Q)_\lambda$ is a quadratic form$\}$. Then by definition $A \notin \mathcal{A}$. That is there exists a u such that $\langle A, uu^t\rangle > q(\mathcal{A})(u)$. Place $h = (I + \lambda A)\,u$ then

$$\begin{aligned}\langle 2(\frac{1}{2}A)_\lambda, hh^t\rangle &= 2\left(\langle \frac{1}{2}A, uu^t\rangle + \frac{1}{2\lambda}\|h-u\|^2\right) > q(\mathcal{A})(u) + \frac{1}{\lambda}\|h-u\|^2 \\ &= 2\left(q(\frac{1}{2}\mathcal{A})(u) + \frac{1}{2\lambda}\|h-u\|^2\right) \geq 2q_\lambda(\frac{1}{2}\mathcal{A})(h) = q(\mathcal{A}_\lambda)(h)\end{aligned}$$

implying the contradiction $2(\frac{1}{2}A)_\lambda = P \notin \mathcal{A}_\lambda$. Hence when $P \in \mathcal{A}_\lambda$ is such that $2(\frac{1}{2}A)_\lambda = P$ for some A we have $A \in \mathcal{A}$.

Finally we note that if $P \in \operatorname{int}\mathcal{A}_\lambda$ then $\langle P, uu^t\rangle - (2\lambda)^{-1}\|u\|^2 < 2q_\lambda(\frac{1}{2}\mathcal{A})(u) - (2\lambda)^{-1}\|u\|^2 := h(u)$ where the dominating function h is concave. As $h(0) = 0$ we have $\langle P, uu^t\rangle - (2\lambda)^{-1}\|u\|^2 < h(u) \leq 0$ for all u and so $P <_{\mathcal{P}(n)} \lambda^{-1}I$ (in that $I - \lambda P \in \operatorname{int}\mathcal{P}(n)$). Thus we have $A = 2(\frac{1}{2}P)^\lambda$ such that $2(\frac{1}{2}A)_\lambda = P$ and so $A = 2(\frac{1}{2}P)^\lambda \in \mathcal{A}$ by the previous argument. Then

$$\operatorname{int}\mathcal{A}_\lambda \subseteq \{2(\frac{1}{2}Q)_\lambda \mid Q \in \mathcal{A} \text{ and } 2(\frac{1}{2}Q)_\lambda \text{ is a quadratic form}\} - \mathcal{P}(n) \tag{40}$$

implying $\mathcal{A}_\lambda = (\operatorname{int}\mathcal{A}_\lambda)^1 \subseteq \left(\{2(\frac{1}{2}Q)_\lambda \mid Q \in \mathcal{A} \text{ and } 2(\frac{1}{2}Q)_\lambda \text{ is a quadratic form}\} - \mathcal{P}(n)\right)^1$

and as $\operatorname{cl}\operatorname{int}\mathcal{A}_\lambda = \mathcal{A}_\lambda$ by convexity and the fact that

$$\operatorname{cl}\{2(\frac{1}{2}Q)_\lambda \mid Q \in \mathcal{A} \text{ and } 2(\frac{1}{2}Q)_\lambda \text{ is a quadratic form}\} - \mathcal{P}(n)$$

is closed (since $0^+\operatorname{cl}\{2(\frac{1}{2}Q)_\lambda \mid Q \in \mathcal{A}$ and $2(\frac{1}{2}Q)_\lambda$ is a quadratic form$\} \cap \mathcal{P}(n) = \{0\}$) the second equality follows. □

Acknowledgements. The authors would like to thank an anonymous referee for alerting us to the references [4, 5, 6, 13, 14, 20].

References

[1] A. Auslender: Stability in mathematical programming with nondifferentiable data, SIAM J. Control Optimization 22 (1984) 239–254.

[2] H. Attouch, R. J.-B. Wets: Quantitative stability of variational systems: I. The epigraphical distance, Trans. Amer. Math. Soc. 328 (1991) 695–729.

[3] G. Beer: Topologies on Closed and Convex Sets, Mathematics and its Applications 268, Kluwer Academic Publishers, Dordrecht (1993).

[4] J. Benoist: Sur la convergence de la dérivée de la régularisée de Lasry-Lions, C. R. Acad. Sci., Paris, Sér. I 315 (1992) 941–944.

[5] J. Benoist: Approximation and regularization of arbitrary sets in finite dimensions, Set-Valued Anal. 2 (1994) 95–115.

[6] F. Bernard, L. Thibault: Prox-regular functions in Hilbert spaces, J. Math. Anal. Appl. 303 (2005) 1–14.

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

[8] M. G. Crandall, M. Kocan, P. Soravia, A. Swiech: On the equivalence of various weak notions of solutions of elliptic PDE's with measurable ingredients, in: Progress in Elliptic and Parabolic Partial Differential Equations (Capri, 1994), A. Alvino et al. (ed.), Pitman Res. Notes Math. Ser. 350, Longman, Harlow (1996) 136–162.

[9] A. Eberhard, M. Nyblom, D. Ralph: Applying generalised convexity notions to jets, in: Generalised Convexity, Generalised Monotonicity: Recent Results, J.-P. Crouzeix et al. (ed.), Kluwer Academic Publishers, Dordrecht (1998) 111–157.

[10] A. Eberhard, M. Nyblom: Jets, generalized convexity, proximal normality and differences of functions, Nonlinear Anal. 34 (1998) 319–360.

[11] A. Eberhard: Prox-regularity and subjets, in: Optimization and Related Topics, A. Rubinov et al. (ed.), Applied Optimization 47, Kluwer Academic Publishers, Dordrecht (2001) 237–313.

[12] A. C. Eberhard, C. E. M. Pearce: A comparison of two approaches to second-order subdifferentiability concepts with applications to optimality conditions, in: Optimization and Control with Applications, L. Qi, K. Teo, X. Yang (eds.), Applied Optimization 96, Springer, New York (2005) 35–100.

[13] P. Georgiev, N. Zlateva: Lasry-Lions regularisations and reconstruction of subdifferentials, C. R. Acad. Bulg. Sci. 51(9-10) (1998) 9–12.

[14] P. Georgiev, N. Zlateva: Reconstruction of the Clarke subdifferential by Lasry-Lions regularization, J. Math. Anal. Appl. 248 (2000) 415–428.

[15] J.-B. Hiriart-Urruty, Ph. Plazanet: Moreau's theorem revisited, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 6, No. Suppl. (1989) 325–338.

[16] A. D. Ioffe: Approximate subdifferentials and applications 2: Functions on locally convex spaces, Mathematika 33 (1986) 111–128.

[17] A. D. Ioffe: Approximate subdifferentials and applications 3: Metric theory, Mathematika 36 (1989) 1–36.

[18] A. D. Ioffe, J.-P. Penot: Limiting subhessians, limiting subjets and their calculus, Trans. Amer. Math. Soc. 349 (1997) 789–807.

[19] R. Jensen, P. L. Lions, P. E Souganidis: A uniqueness result for the viscosity solutions of second-order fully nonlinear partial differential equations, Proc. Amer. Math. Soc. 102 (1988) 975–978.

[20] A. Jourani: Limit-superior of subdifferentials of uniformly convergent functions, Positivity 3 (1999) 33–47.

[21] P.-L. Lions, P. E. Souganidis: Viscosity solutions of stochastic differential equations, stochastic control and stochastic differential games, in: Stochastic Differential Systems, Stochastic Control Theory and Applications (Minneapolis, 1986), W. H. Fleming, P.-L. Lions (eds.), IMA Vol. Math. Appl. 10, Springer, New York (1988) 293–309.

[22] J.-M. Lasry, P.-L. Lions: A remark on regularization in Hilbert spaces, Isr. J. Math. 55 (1986) 257–266.

[23] J.-P. Penot: Sub-Hessians, super-Hessians and conjugation, Nonlinear Anal. 23 (1994) 689–702.

[24] R. A. Poliquin, R. T. Rockafellar: Prox-regular functions in variational analysis, Trans. Amer. Math. Soc. 348 (1996) 1805–1838.

[25] R. A. Poliquin, R. T. Rockafellar: Generalized Hessian properties of regularized nonsmooth functions, SIAM J. Optimization 6 (1996) 1121–1137.

[26] R. A. Poliquin, R. T. Rockafellar: Tilt stability of local minimum, SIAM J. Optimization 8 (1998) 287–299.

[27] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton (1970).

[28] R. T. Rockafellar: A derivative-coderivative inclusion in second-order nonsmooth analysis, Set-Valued Anal. 5 (1997) 89–105.

[29] R. T. Rockafellar, R. J.-B. Wets: Variational systems - An introduction, in: Multifunctions and Integrands, G. Salinetti (ed.), Lecture Notes in Math. 1091, Springer, Berlin (1984) 1–54.

[30] R. T. Rockafellar, R. J.-B. Wets: Variational Analysis, Comprehensive Studies in Mathematics 317, Springer, Berlin (1998).