3-1-2010

First-Order and Second-Order Optimality Conditions for Nonsmooth Constrained Problems via Convolution Smoothing

Andrew C. Eberhard
Royal Melbourne Institute of Technology, Australia, andy.eb@rmit.edu.au

Boris S. Mordukhovich
Wayne State University, boris@math.wayne.edu

FIRST-ORDER AND SECOND-ORDER OPTIMALITY CONDITIONS FOR NONSMOOTH CONSTRAINED PROBLEMS VIA CONVOLUTION SMOOTHING

ANDREW C. EBERHARD and BORIS S. MORDUKHOVICH

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics
Research Report

2010 Series
#3

This research was partly supported by the USA National Science Foundation

FIRST-ORDER AND SECOND-ORDER OPTIMALITY CONDITIONS FOR NONSMOOTH CONSTRAINED PROBLEMS VIA CONVOLUTION SMOOTHING[1]

ANDREW C. EBERHARD[2] and BORIS S. MORDUKHOVICH[3]

Dedicated to Franco Giannessi in honor of his 75th birthday

Abstract. This paper mainly concerns deriving first-order and second-order necessary (and partly sufficient) optimality conditions for a general class of constrained optimization problems via smoothing regularization procedures based on infimal-like convolutions/envelopes. In this way we obtain first-order optimality conditions of both lower subdifferential and upper subdifferential types and then second-order conditions of three kinds involving, respectively, generalized second-order directional derivatives, graphical derivatives of first-order subdifferentials, and second-order subdifferentials defined via coderivatives of first-order constructions.

Keywords. Variational analysis, constrained optimization, generalized differentiation, first-order and second-order optimality conditions

AMS Subject Classification. 49J52, 49J53, 90C30, 90C26, 26A27

1 Introduction

In this paper we pay the main attention to the study of the following general problem of constrained optimization in finite-dimensional spaces:

$$\text{minimize} \quad f(x) \quad \text{subject to } x \in F^{-1}(\Omega), \tag{1.1}$$

where $f\colon I\!R^n \to \overline{I\!R} := (-\infty, \infty]$ is a proper ($\not\equiv \infty$) extended-real-valued function, $F\colon I\!R^n \to I\!R^m$ is a vector-valued mapping, $\Omega \subset I\!R^m$ is a nonempty subset, and $F^{-1}(\Omega)$ is the inverse image/preimage of Ω under F defined by

$$F^{-1}(\Omega) := \left\{x \in I\!R^n \middle|\, F(x) \in \Omega\right\}. \tag{1.2}$$

Note that we confine ourselves to the finite-dimensional setting just for simplicity; most of the results obtained in the paper can be extended to infinite dimensions using the techniques developed below and tools of infinite-dimensional variational

[1]This research was supported by the Australian Research Council under Discovery Grant DP0664423.

[2]School of Mathematical and Geospatial Sciences, Royal Melbourne Institute of Technology, Melbourne, Victoria 3001, Australia. Email: andy.eb@rmit.edu.au.

[3]Department of Mathematics, Wayne State University, Detroit, Michigan 48202, USA. Email: boris@math.wayne.edu. Research of this author was partly supported by the US National Science Foundation under Grant DMS-0603846.

analysis and generalized differentiation presented in [13, 14]. Furthermore, the constraint mapping F in (1.1) can be set-valued, in which case the preimage (1.2) is replaced by $F^{-1}(\Omega) := \{x|\ F(x) \cap \Omega \neq \emptyset\}$. But already in finite dimensions the class of constrained optimization problems (1.1) is fairly general including, in particular, problems with conventional equality and inequality constraints and much more; see, e.g., [14, 17].

Unless otherwise stated, we assume that the cost function f is lower semicontinuous, the constraint mapping F is continuous, and the set Ω in (1.1) is closed.

Our major goal in what follows is to study *optimality conditions* via certain *regularization procedures* that approximate the original nonsmooth constrained optimization problem (1.1) by a parametric family of unconstrained minimization problems. Define, for all $\lambda > 0$ sufficiently small, the *infimal convolution/envelope* of the constrained problem (1.1) by

$$f_\lambda(x) := \inf_{u \in F^{-1}(\Omega)} \left(f(u) + \frac{1}{2\lambda}\|x - u\|^2 \right). \tag{1.3}$$

It has been well recognized in variational analysis that envelopes of type (1.3) have a number of remarkable properties important for various approximation and numerical techniques while dealing with nonsmooth optimization problems; see, e.g., [1, 2, 3, 6, 7, 8, 9, 10, 16, 17] and the references therein. We intend to provide a better understanding of the underlying necessary (and partly sufficient) optimality conditions appearing in such regularizations and being useful from both qualitative and algorithmic viewpoints. While restricting our attention to finite dimensions, we keep our functions as general as possible within this framework.

2 Preliminaries in Generalized Differentiation

Here we introduce and discuss the main generalized differential constructions used in this paper. Our notation is standard corresponding to the basic monographs [13, 17] on variational analysis and generalized differentiation. The reader can consult with these texts for more details.

Recall that the *Painlevé-Kuratowski upper/outer limit* of a set-valued mapping $G\colon I\!R^n \rightrightarrows I\!R^m$ as $x \to \bar{x} \in \operatorname{dom} G := \{x \in I\!R^n|\ G(x) \neq \emptyset\}$ is defined by

$$\operatorname*{Lim\,sup}_{x \to \bar{x}} G(x) := \Big\{ y \in I\!R^m \Big|\ \ \exists\, x_k \to \bar{x},\ y_k \to y \ \text{ as } \ k \to \infty \ \text{ such that } \ y_k \in G(x_k) \ \text{ for all } \ k \in I\!N := \{1, 2, \ldots\} \Big\}. \tag{2.1}$$

Given further an extended-real-valued function $g\colon I\!R^n \to [-\infty, \infty]$ finite at $\bar{x}$, define the *regular/Fréchet subdifferential* of g at $\bar{x}$ by

$$\widehat{\partial} g(\bar{x}) := \left\{ x^* \in I\!R^n \,\Big|\, \liminf_{\substack{t \downarrow 0 \\ u \to v}} \frac{1}{t}\left(g(\bar{x} + tu) - g(\bar{x})\right) \geq \langle x^*, v\rangle \text{ for all } v \in I\!R^n \right\} \tag{2.2}$$

along with the corresponding *upper subdifferential* (or *superdifferential*) given by

$$\widehat{\partial}^{+}g(\bar{x}) := -\widehat{\partial}\,(-g)\,(\bar{x}). \tag{2.3}$$

By [17, Proposition 8.5] regular subgradients admits the following smooth variational description: $x^* \in \widehat{\partial}g(\bar{x})$ if and only there is a C^1 function $s\colon I\!R^n \to I\!R$ such that $s\,(\bar{x}) = g\,(\bar{x})$, $s\,(x) \le g\,(x)$ for all $x \in I\!R^n$, and $\nabla s\,(\bar{x}) = x^*$.

The *proximal subdifferential* $\partial_p g(\bar{x})$ of g at $\bar{x}$ is the collection of proximal subgradients $x^* \in \partial_p g(\bar{x})$ defined as follows: there is $r \ge 0$ such that

$$g(x) \ge g(\bar{x}) + \langle x^*, x - \bar{x}\rangle - \frac{r}{2}\|x - \bar{x}\|^2 \tag{2.4}$$

for all x around $\bar{x}$. By [17, Proposition 8.46] we have that $x^* \in \partial_p g\,(\bar{x})$ if and only if there is a C^2 function $s\colon I\!R^n \to I\!R$ with $s\,(\bar{x}) = g\,(\bar{x})$, $s\,(x) \le g\,(x)$ for all $x \in I\!R^n$, and $\nabla s\,(\bar{x}) = x^*$. Consequently the proximal subdifferential is a convex (may not closed) subset of the regular subdifferential, i.e., $\partial_p g\,(\bar{x}) \subset \widehat{\partial}g(\bar{x})$.

Fundamental concepts of generalized differentiation in variational analysis are obtained by limiting procedures. In this way the *basic/limiting/Mordukhovich subdifferential* of g at $\bar{x}$ is defined by the formulas

$$\partial g(\bar{x}) := \operatorname*{Lim\,sup}_{x \xrightarrow{g} \bar{x}} \widehat{\partial}g(x) = \operatorname*{Lim\,sup}_{x \xrightarrow{g} \bar{x}} \partial_p g(x) \tag{2.5}$$

equivalent in finite dimensions, where the notation $x \xrightarrow{g} \bar{x}$ in the Painlevé-Kuratowski outer limit (2.1) means that $x \to \bar{x}$ with $g(x) \to g(\bar{x})$. The corresponding *singular/horizontal subdifferential* of g at $\bar{x}$ is given by

$$\partial^\infty g(\bar{x}) := \operatorname*{Lim\,sup}_{x \xrightarrow{g} \bar{x},\, \lambda\downarrow 0} \lambda\widehat{\partial}g(x) = \operatorname*{Lim\,sup}_{x \xrightarrow{g} \bar{x},\, \lambda\downarrow 0} \lambda\partial_p g(x). \tag{2.6}$$

In this paper we use two constructions of normal cones to sets generated by the the regular and basic subdifferentials. Given $\Omega \subset I\!R^n$ and $\bar{x} \in \Omega$, define respectively

$$\widehat{N}_\Omega(\bar{x}) := \widehat{\partial}\delta_\Omega(\bar{x}) \quad\text{and}\quad N_\Omega(\bar{x}) := \partial\delta_\Omega(\bar{x}), \tag{2.7}$$

where δ_Ω stands for the indicator function of Ω equal 0 on the set and ∞ outside. We clearly have $\widehat{N}_\Omega(\bar{x}) \subset N_\Omega(\bar{x})$ and say that Ω is *normally regular* at $\bar{x}$ if the latter inclusion holds as equality. A number of sufficient conditions for this property as well as its applications can be found in [13, 17].

Let us next discuss second-order generalized differential constructions employed in what follows. The *jet* of $g\colon I\!R^n \to \overline{I\!R}$ at $\bar{x} \in \operatorname{dom} g := \{x \in I\!R^n |\ g(x) < \infty\}$ is

$$\partial^{2,-}g(\bar{x}) := \Big\{ (\nabla\varphi(\bar{x}), \nabla^2\varphi(\bar{x}))\Big|\ \begin{array}{l} g - \varphi \text{ has a local minimum at } \bar{x} \\ \text{with } \varphi \in C^2(I\!R^n)\Big\}. \end{array} \tag{2.8}$$

The jet construction (2.8) first appeared and has been proved to be very useful in the theory of viscosity solutions of second-order partial differential equations [4].

Second-order subdifferentials of another type defined via *graphical derivatives* and *coderivatives* of first-order subdifferentials appeared in optimization; cf. [7, 11, 13, 15, 17]. In this paper we use the following constructions of this type given by

$$D\left(\partial_p g\right)(\bar{x},\bar{x}^*)(w) := \left\{z \in I\!R^n \,\middle|\, (w,z) \in T_{\partial_p g}(\bar{x},\bar{x}^*)\right\}, \tag{2.9}$$

$$\widehat{D}^*\left(\partial_p g\right)(\bar{x},\bar{x}^*)(w) := \left\{z \in I\!R^n \,\middle|\, (z,-w) \in \left(T_{\partial_p g}(\bar{x},\bar{x}^*)\right)^{\circ}\right\}, \tag{2.10}$$

where $(\bar{x},\bar{x}^*) \in \operatorname{gph}\partial_p g$, where $^\circ$ stands for the polar of sets, and where

$$T_{\partial_p g}(\bar{x},\bar{x}^*) := \operatorname*{Lim\,sup}_{t\downarrow 0} \frac{\operatorname{gph}\partial_p g - (\bar{x},\bar{x}^*)}{t}$$

is the *Bouligand-Severi contingent cone* to the graph of $\partial_p g\colon I\!R^n \rightrightarrows I\!R^n$.

Finally in this section, recall the notions of the first-order and second-order *Dini-Hadamard directional derivatives* of g at $\bar{x}$ in the direction $u \in I\!R^n$ defined by

$$g'_-(\bar{x};v) := \liminf_{u\to v,\, t\downarrow 0} \frac{1}{t}\Big(g(\bar{x}+tu) - g(\bar{x})\Big), \tag{2.11}$$

$$g''_-(\bar{x},\bar{x}^*;v) := \liminf_{u\to v,\, t\downarrow 0} \frac{2}{t^2}\Big(g(\bar{x}+tu) - g(\bar{x}) - t\langle\bar{x}^*,u\rangle\Big), \tag{2.12}$$

respectively, where $\bar{x}^* \in \partial_p g(\bar{x})$ in (2.12).

Throughout the paper we use standard notation of variational analysis [13, 17]. Recall that $I\!B$ is the closed unit ball of the spaces in question, $B_\gamma(\bar{x})$ stands for the ball centered at $\bar{x}$ with radius $\gamma > 0$, A^T signifies the matrix transposition, and $d(\cdot;\Omega)$ denotes the Euclidean distance function.

3 Minimization of Infimal Convolutions

Let us first discuss here some characteristic properties of the infimal convolution and then collect the results relating to their minimization needed in what follows.

Recall that a function $g\colon I\!R^n \mapsto \overline{I\!R}$ is *quadratically minorized* (or *prox-bounded* [16, 17]) if there are $\alpha \in I\!R$ and $r \geq 0$ such that

$$g(x) \geq \alpha - \frac{r}{2}\|x\|^2 \quad \text{for all} \;\; x \in I\!R^n,$$

i.e., the function $g + \frac{r}{2}\|\cdot\|^2$ is bounded from below. The *lower λ-quadratic/Moreau envelope* (or *infimal convolution*) of $g\colon I\!R^n \to \overline{I\!R}$ is defined by

$$e_\lambda(g)(x) := \inf_{u\in I\!R^n}\left\{g(u) + \frac{1}{2\lambda}\|u-x\|^2\right\}, \tag{3.1}$$

and the corresponding *proximal mapping* $P_\lambda(g)\colon I\!R^n \rightrightarrows I\!R^n$ is given by

$$P_\lambda\left(g\right)\left(x\right) := \text{argmin}\left\{g\left(\cdot\right) + \frac{1}{2\lambda}\left\|\cdot - x\right\|^2\right\}. \tag{3.2}$$

Both quantities in (3.1) and (3.2) are well defined for $0 < \lambda < \frac{1}{r}$; see [17, Exercise 1.24]. Furthermore, $g_\lambda > -\infty$ and thus $P_\lambda(g)(x) \neq \emptyset$ if $\lambda < (\max\{0, r\})^{-1}$. The infimal of all such r is denoted by $r(g)$. When $r(g) < 0$, we have $g_\lambda > -\infty$ for all $\lambda > 0$. Taking the latter into account, we denote $\bar{r}(g) := \max\{r(g), 0\}$ with the convention $1/0 = \infty$. The quantity $\lambda_g := (\bar{r}(g))^{-1}$ is called the *proximal threshold* for the function g.

It is easy to check that there is $M > 0$ and $\gamma > 0$ such that $\|y\| \leq M$ for all $y \in P_\lambda(g)(x)$ and all $x \in B_\gamma(\bar{x})$ satisfying

$$\left|e_\lambda\left(g\right)\left(x\right) - e_\lambda\left(g\right)\left(u\right)\right| \leq \frac{M}{\lambda}\|x - u\| \text{ whenever } x, u \in B_\gamma(\bar{x}). \tag{3.3}$$

In particular, this implies that the directional derivative $(-e_\lambda\left(g\right))'(\bar{x};\cdot)$ of the infimal convolution (3.1) always exists and is finite-valued.

Observe further that the quadratically shifted function

$$\begin{aligned} e_\lambda\left(g\right)\left(x\right) - \frac{1}{2\lambda}\left\|x\right\|^2 &= \inf_{u \in I\!R^n}\left\{g\left(u\right) + \frac{1}{2\lambda}\left\|u - x\right\|^2 - \frac{1}{2\lambda}\left\|x\right\|^2\right\} \\ &= \inf_{u \in I\!R^n}\left\{g\left(u\right) + \frac{1}{2\lambda}\left\|u\right\|^2 - \frac{1}{\lambda}\langle u, x\rangle\right\} \end{aligned}$$

is concave as the infimum of affine functions. Thus, by definition, the infimal convolution $e_\lambda\left(g\right)$ is a *paraconcave* function on $I\!R^n$. It easily follows from paraconcavity and basic convex analysis that

$$\widehat{\partial}^+ e_\lambda\left(g\right)\left(\bar{x}\right) = -\partial_p\left\{-e_\lambda\left(g\right)\right\}\left(\bar{x}\right) \neq \emptyset \text{ whenever } \bar{x} \in I\!R^n$$

for the upper subdifferential (2.3) of the infimal convolution. This fundamental fact is crucial for the further consideration of this paper.

In what follows we apply this fact to the infimal convolution

$$f_\lambda\left(x\right) = e_\lambda\left(f + \delta_{F^{-1}(\Omega)}\right)\left(x\right) \tag{3.4}$$

constructed via the initial data of the original optimization problem (1.1). Clearly (3.4) agrees with the constrained convolution form (1.3) of problem (1.1). It follows from (3.4) and the discussion above that the envelope f_λ is paraconcave, that at every point $\bar{x}$ its directional derivative exists and is finite-valued on $I\!R^n$, and that

$$\widehat{\partial}^+ f_\lambda(\bar{x}) = -\partial_p(-f_\lambda)(\bar{x}) \neq \emptyset \text{ for all } \bar{x} \in I\!R^n. \tag{3.5}$$

Let us next derive some useful relationships between optimal solutions to the original problem (1.1) and to the corresponding infimal convolutions. Note that if $\bar{x} \in \text{argmin}\left\{f + \delta_{F^{-1}(\Omega)}\right\}$, then the envelope $f_\lambda(\bar{x})$ is well defined and prox-bounded being minorized by the quadratic function $f(\bar{x}) - \frac{r}{2}\|x - \bar{x}\|^2$ for any $r > 0$.

Proposition 3.1 (optimal solutions for the original and convolution problems). *If* $\bar{x} \in \operatorname{argmin}\{f + \delta_{F^{-1}(\Omega)}\}$ *in* (1.1), *then* $f_\lambda(\bar{x}) = f(\bar{x})$ *for all* $\lambda > 0$. *Furthermore, we have the inclusions*

$$\operatorname{argmin}\left\{f + \delta_{F^{-1}(\Omega)}\right\} \subset \operatorname{argmin}\left\{f_\lambda + \delta_{F^{-1}(\Omega)}\right\} \subset \operatorname{argmin} f_\lambda. \tag{3.6}$$

Proof. Let us show first that $f_\lambda(\bar{x}) = f(\bar{x})$ whenever $\bar{x} \in \operatorname{argmin}\left\{f + \delta_{F^{-1}(\Omega)}\right\}$ and $\lambda > 0$. Since $\bar{x} \in F^{-1}(\Omega)$, we have

$$\begin{aligned} f_\lambda(\bar{x}) &= \min_{u \in F^{-1}(\Omega)} \left(f(u) + \frac{1}{2\lambda}\|\bar{x} - u\|^2\right) \\ &\le \left(f(u) + \frac{1}{2\lambda}\|\bar{x} - u\|^2\right)|_{u=\bar{x}} = f(\bar{x}) \end{aligned}$$

for all $\lambda > 0$. Moreover, it follows from $\bar{x} \in \operatorname{argmin}\left\{f + \delta_{F^{-1}(\Omega)}\right\}$ that

$$\begin{aligned} f_\lambda(\bar{x}) &= \min_{u \in F^{-1}(\Omega)} \left(f(u) + \frac{1}{2\lambda}\|\bar{x} - u\|^2\right) \\ &\ge \min_{u \in F^{-1}(\Omega)} \left(f(\bar{x}) + \frac{1}{2\lambda}\|\bar{x} - u\|^2\right) \\ &= f\left(\bar{x}\right) + \min_{u \in F^{-1}(\Omega)} \left(\frac{1}{2\lambda}\|\bar{x} - u\|^2\right) = f\left(\bar{x}\right). \end{aligned}$$

Combining the relationships above gives us $f_\lambda(\bar{x}) = f(\bar{x}) > -\infty$.

Observe further that when $\bar{x} \in \operatorname{argmin}\left\{f + \delta_{F^{-1}(\Omega)}\right\}$, we have

$$f_\lambda(x) = \min_{u \in F^{-1}(\Omega)} \left(f(u) + \frac{1}{2\lambda}\|x - u\|^2\right) \ge \min_{u \in F^{-1}(\Omega)} f(u) = f(\bar{x}) = f_\lambda(\bar{x})$$

for any $x \in I\!R^n$. Since $\bar{x} \in F^{-1}(\Omega)$, the latter implies that $\bar{x} \in \operatorname{argmin}\left\{f_\lambda + \delta_{F^{-1}(\Omega)}\right\}$. This justifies the first inclusion in (3.6); the second one is obvious. $\triangle$

The next result provides an important in what follows necessary condition for optimal solutions to the original problem (1.1) via the stationary condition for the infimal convolution (3.4) that occurs to be differentiable at minimal points. It is based on paraconcavity of f_λ and the upper subdifferential property (3.5).

Theorem 3.2 (stationary condition via smoothing infimal convolutions). *Let* $\bar{x}$ *be an optimal solution to the original problem* (1.1), *i.e.*, $\bar{x} \in \operatorname{argmin}\{f + \delta_{F^{-1}(\Omega)}\}$. *Then the infimal convolution* f_λ *is differentiable at* $\bar{x}$ *and we have the stationary condition* $\nabla f_\lambda(\bar{x}) = 0$ *for all* $\lambda > 0$ *sufficiently small.*

Proof. As mentioned above, we have from $\bar{x} \in \operatorname{argmin}\left\{f + \delta_{F^{-1}(\Omega)}\right\}$ that the infimal convolution f_λ is well-defined and paraconcave for sufficiently small $\lambda > 0$. Furthermore, it follows from Proposition 3.1 that $\bar{x}$ is a minimizer for f_λ. Thus

$0 \in \widehat{\partial} f_\lambda(\bar{x})$ by an Fréchet subdifferential counterpart of the Fermat stationary rule, which can be easily seen by definition (2.2). On the other hand, we have from the paraconcavity of f_λ and condition (3.5) that $\widehat{\partial}^+ f_\lambda(\bar{x}) \neq \emptyset$. Employing now [13, Proposition 1.87], conclude that f_λ is in fact (Fréchet) differentiable at $\bar{x}$ with

$$\widehat{\partial} f_\lambda(\bar{x}) = \widehat{\partial}^+ f_\lambda(\bar{x}) = \{\nabla f_\lambda(\bar{x})\}.$$

Thus $\nabla f_\lambda(\bar{x}) = 0$ for all small $\lambda > 0$, which completes the proof. $\triangle$

4 First-Order Optimality Conditions

It is shown in [12, Theorem 3.1(i)] (see also [14, Proposition 5.2]) that optimal solutions to the constrained *minimization* problem (1.1) satisfy the following *upper* subdifferential necessary optimality condition:

$$-\widehat{\partial}^+ f(\bar{x}) \subset \widehat{N}_{F^{-1}(\Omega)}(\bar{x}), \tag{4.1}$$

in terms of the Fréchet/regular normal cone to $F^{-1}(\Omega)$ at $\bar{x}$ defined in (2.7). As shown in [12, 14], condition (4.1) is generally independent of rather conventional "lower" subdifferential necessary optimality conditions while providing more selective information to single out nonoptimal solutions in certain classes of minimization problems. However, a major drawback of (4.1) is that the upper subdifferential $\widehat{\partial} f(\bar{x})$ may be empty for important classes of cost functions in (4.1), which happens, e.g., when f is convex. In such situations the optimality condition (4.1) is trivial.

In this section we show that the upper subdifferential optimality condition (4.1) can be replaced but its *convolution upper subdifferential* counterpart

$$-\widehat{\partial}^+ f_\lambda(\bar{x}) \subset \widehat{N}_{F^{-1}(\Omega)}(\bar{x}) \ \text{ for all small } \ \lambda > 0, \tag{4.2}$$

which has at least two advantages in comparison with (4.1):

(a) the necessary optimality condition (4.2) always implies that of (4.1);
(b) the upper subdifferential $\widehat{\partial}^+ f_\lambda(\bar{x})$ is always *nonempty*.

We show furthermore that the upper subdifferential convolution condition implies more conventional first-order necessary conditions in the lower subdifferential form. Note that, although the results below are formulated for global minimizers, they can be easily extended for *local* ones by restricting the cost function to a small ball around the local minimizer under consideration.

Theorem 4.1 (upper subdifferential optimality conditions via convolutions). *Let $\bar{x} \in \operatorname{argmin}_{x \in F^{-1}(\Omega)} f(x)$. Then for all small $\lambda > 0$ we have*

$$\emptyset \neq -\widehat{\partial}^+ f_\lambda(\bar{x}) \subset \widehat{N}_{F^{-1}(\Omega)}(\bar{x}) \subset N_{F^{-1}(\Omega)}(\bar{x}), \tag{4.3}$$

where $N_{F^{-1}(\Omega)}(\bar{x})$ stands for the basic normal cone to $F^{-1}(\Omega)$ at $\bar{x}$ defined in (2.7).

Proof. For completeness and the reader's convenience, let us first present, as a part of the proof of this theorem, an alternative *convolution* proof of the upper subdifferential condition (4.1) in an equivalent form that is of its own interest:

$$-\widehat{\partial}^+ f(\bar{x}) \subset \Big(\operatorname*{Lim\,sup}_{t\downarrow 0} \frac{F^{-1}(\Omega)-\bar{x}}{t}\Big)^\circ. \tag{4.4}$$

By Proposition 3.1 we have $f_\lambda(\bar{x}) = f(\bar{x})$ for all $\lambda > 0$. Picking any $x^* \in \widehat{\partial}^+ f(\bar{x})$ and employing the aforementioned smooth variational description of the regular subgradient $-x^* \in \widehat{\partial}(-f)(\bar{x})$, find a smooth function $s\colon I\!R^n \to I\!R$ such that $\nabla s(\bar{x}) = x^*$, $s(\bar{x}) = f(\bar{x})$, and $s(x) \ge f(x)$ for all $x \in I\!R^n$. Thus $s(\bar{x}) = f_\lambda(\bar{x})$ and

$$s(x) + \delta_{F^{-1}(\Omega)} \ge f(x) + \delta_{F^{-1}(\Omega)} \ge f_\lambda(x), \quad x \in I\!R^n.$$

By Theorem 3.2 the envelope f_λ is differentiable at $\bar{x}$ for all $\lambda > 0$ sufficiently small. Thus its classical directional derivative exists and reduces to $(f_\lambda)'_-(\bar{x};v)$ in (2.11) for any $v \in I\!R^n$. This gives the relationships

$$\begin{aligned}(-f_\lambda)'_-(\bar{x};v) &\ge -\liminf_{t\downarrow 0}\left[\frac{1}{t}\big(s(\bar{x}+tv) - s(\bar{x})\big) + \delta_{F^{-1}(\Omega)}(\bar{x}+tv)\right]\\ &= -\Big\{\langle \nabla s(\bar{x}), v\rangle + \delta_{T_{F^{-1}(\Omega)}(\bar{x})}(v)\Big\}, \quad v \in I\!R^n\end{aligned}$$

where $T_{F^{-1}(\Omega)}(\bar{x})$ stands for the contingent cone defined in Section 2. Hence

$$(-f_\lambda)'_-(\bar{x};v) + \delta_{T_{F^{-1}(\Omega)}(\bar{x})}(v) \ge \langle -\nabla s(\bar{x}), v\rangle \text{ for all } v \in I\!R^n. \tag{4.5}$$

Since $\nabla f_\lambda(\bar{x}) = 0$ by Theorem 3.2, we have $(-f_\lambda)'_-(\bar{x};v) = 0$ and then

$$-x^* = -\nabla s(\bar{x}) \in \big(T_{F^{-1}(\Omega)}(\bar{x})\big)^\circ$$

by (4.5), which thus justifies (4.4). Note that (4.4) is equivalent to (4.2) due to the well-known duality correspondence between the contingent cone and Fréchet normal cone to sets in finite dimensions; see, e.g., [13, Theorem 1.10].

To justify further the optimality conditions in (4.3), observe that $\widehat{\partial}^+ f_\lambda(\bar{x}) \ne \emptyset$ follows from (3.5) due to the *paraconcavity* of the infimal convolution f_λ, while the second inclusion in (4.3) follows form the observation that $\widehat{\partial} g(\bar{x}) \subset \partial g(\bar{x})$ for any function $g\colon I\!R^n \to \overline{I\!R}$. Thus it remains to check the first inclusion in (4.3), which reduces to (4.2). The latter readily follows from Proposition 3.1 by applying condition (4.1), or its equivalent (4.4), to the infimal convolution f_λ. $\triangle$

Remark 4.2 (relationship between upper subdifferential and convolution upper subdifferential conditions). Let us show that the convolution upper subdifferential condition (4.2) always implies the upper subdifferential one (4.1). Suppose without loss of generality that $\widehat{\partial}^+ f(\bar{x}) \ne \emptyset$. Since by definition (1.3) $f_\lambda(x) \le f(x)$ for all x and $f_\lambda(\bar{x}) = f(\bar{x})$ by Proposition 3.1, we have

$$\liminf_{\substack{t\downarrow 0\\ u\to v}} \frac{1}{t}\Big((-f)(\bar{x}+tu) - (-f)(\bar{x})\Big) \le \liminf_{\substack{t\downarrow 0\\ u\to v}} \frac{1}{t}\Big((-f_\lambda)(\bar{x}+tu) - (-f_\lambda)(\bar{x})\Big)$$

for $\lambda > 0$ and consequently $\widehat{\partial}(-f)(\bar{x}) \subset \widehat{\partial}(-f_\lambda)(\bar{x})$ by (2.2). It gives therefore that $\widehat{\partial}^+ f(\bar{x}) \subset \widehat{\partial}^+ f_\lambda(\bar{x})$, which justifies the implication (4.2) $\Longrightarrow$ (4.3).

It is worth mentioning that for *any convex function* $f\colon I\!\!R^n \to I\!\!R$ the upper subdifferential condition (4.3) is always trivial while the convolution one (4.2) provides nontrivial information, since $\widehat{\partial}^+ f_\lambda(\bar{x}) \neq \emptyset$ by the results above.

Next we use the convolution upper subdifferential condition (4.2) for deriving some *lower subdifferential* optimality conditions for the minimization problem (1.1). The following proposition establishes a relationship between the (lower) basic subdifferential (2.5) and the upper subdifferential (2.3) of infimal convolutions.

Proposition 4.3 (basic and upper subdifferentials of infimal convolutions). *Let $f\colon I\!\!R^n \to \overline{I\!\!R}$ be finite at $\bar{x}$. Then for all $\lambda > 0$ sufficiently small we have*

$$\partial f_\lambda(\bar{x}) \subset \widehat{\partial}^+ f_\lambda(\bar{x}). \tag{4.6}$$

This implies the convolution lower subdifferential optimality conditions

$$-\partial f_\lambda(\bar{x}) \subset \widehat{N}_{F^{-1}(\Omega)}(\bar{x}) \subset N_{F^{-1}(\Omega)}(\bar{x}) \tag{4.7}$$

when $\bar{x} \in \operatorname{argmin}_{x \in F^{-1}(\Omega)} f(x)$.

Proof. For all small $\lambda > 0$ consider

$$u_\lambda \in \operatorname{argmin}\left\{ f(\cdot) + \delta_{F^{-1}(\Omega)}(\cdot) + \frac{1}{2\lambda}\|\cdot - \bar{x}\|^2 \right\} := P_\lambda. \tag{4.8}$$

Whenever $t > 0$ we have the relationships

$$\begin{aligned}\frac{1}{t}\left(f_\lambda(\bar{x}+tu) - f_\lambda(\bar{x})\right) \quad &\le \frac{1}{t}\left(\frac{1}{2\lambda}\left(\|\bar{x}+tu-u_\lambda\|^2 - \|\bar{x}-u_\lambda\|^2\right)\right)\\ &= \left\langle \frac{1}{\lambda}(\bar{x}-u_\lambda), u\right\rangle + \frac{t}{2\lambda}\|u\|^2 \quad \text{for all } \ u \in I\!\!R^n.\end{aligned}$$

Passing above to the 'lim sup' as $t \downarrow 0$ and $u \to v$, we get by (2.2) and (2.3) that $\frac{1}{\lambda}(\bar{x}-u_\lambda) \in \widehat{\partial}^+ f_\lambda(\bar{x})$. The convexity of the set $\widehat{\partial}^+ f_\lambda(\bar{x})$ implies that

$$\operatorname{co}\left\{ \frac{1}{\lambda}(\bar{x}-u_\lambda) \;\Big|\; u_\lambda \in P_\lambda \right\} \subset \widehat{\partial}^+ f_\lambda(\bar{x}). \tag{4.9}$$

Without loss of generality, we may assume in this setting that f is prox-bounded with $f(\bar{x}) = 0$. Then the result of [16, Proposition 4.3(a)] ensures that

$$\partial f_\lambda(\bar{x}) \subset \operatorname{co}\left\{ \frac{1}{\lambda}(\bar{x}-u_\lambda) \;\Big|\; u_\lambda \in P_\lambda \right\}$$

for all $\lambda > 0$ sufficiently small. By (4.9) the latter gives (4.6). The lower subdifferential optimality conditions (4.7) follow now from (4.6) and the upper subdifferential convolution conditions of Theorem 4.1. $\triangle$

By passing to the limit as $\lambda \downarrow 0$, we get the following consequences of Theorem 4.1 and Proposition 4.3.

Corollary 4.4 (upper and lower subdifferential limiting convolution conditions). *Let $\bar{x} \in \operatorname{argmin}_{x \in F^{-1}(\Omega)} f(x)$. Then we have the inclusions*

$$-\Big[\Big(\operatorname*{Lim\,sup}_{\lambda \downarrow 0} \widehat{\partial}^+ f_\lambda(\bar{x})\Big) \bigcup \Big(\operatorname*{Lim\,sup}_{\lambda \downarrow 0} \lambda \widehat{\partial}^+ f_\lambda(\bar{x})\Big)\Big] \subset \widehat{N}_{F^{-1}(\Omega)}(\bar{x}),$$

$$-\Big[\Big(\operatorname*{Lim\,sup}_{\lambda \downarrow 0} \widehat{\partial} f_\lambda(\bar{x})\Big) \bigcup \Big(\operatorname*{Lim\,sup}_{\lambda \downarrow 0} \lambda \widehat{\partial} f_\lambda(\bar{x})\Big)\Big] \subset \widehat{N}_{F^{-1}(\Omega)}(\bar{x}).$$

Proof. Both inclusions follow in the same way from the upper subdifferential (4.3) and lower subdifferential (4.7) convolution conditions, respectively, by passing to the limit as $\lambda \downarrow 0$ and taking into account that $\widehat{N}_{F^{-1}(\Omega)}(\bar{x})$ is a closed cone. $\triangle$

5 Infimal Convolutions and Optimality Conditions under Qualification Conditions

In this section we first study limiting behavior of infimal convolutions and the corresponding minimizers under appropriate *qualification conditions* imposed on the initial data of (1.1) via the singular subdifferential of f and the basic normal cone to Ω. Then we use certain constraint qualifications to derive necessary optimality conditions for (1.1) in terms of basic subgradients of the cost function.

Let us start with clarification of the convolution limiting behavior.

Theorem 5.1 (limiting behavior of infimal convolutions and their minimizers). *Let the cost function f in (1.1) be prox-bounded and continuous around some point $\bar{x}$, and let the constraint mapping F be continuously differentiable around $\bar{x}$. Assume in addition that the following qualification conditions are satisfied:*

$$N_\Omega(F(\bar{x})) \cap \big\{ y \in I\!R^m \big|\, \nabla F(\bar{x})^T y = 0 \big\} = \{0\}, \tag{5.1}$$

$$\partial^\infty f(\bar{x}) \cap \big\{ -\nabla F(\bar{x})^T y \big|\, y \in N_\Omega(F(\bar{x})) \big\} = \{0\}. \tag{5.2}$$

Then there is $\gamma > 0$ such that for all $\lambda > 0$ sufficiently small we have

$$f_\lambda(x) = \min_{u \in I\!R^n} \Big(f(u) + \frac{1}{\sqrt{\lambda}} d(F(u);\Omega) + \frac{1}{2\lambda}\|x-u\|^2 \Big), \quad x \in B_\gamma(\bar{x}), \tag{5.3}$$

and then the limiting relationship

$$\lim_{\lambda \downarrow 0} f_\lambda(x) = f(x) + \delta_{F^{-1}(\Omega)}(x) \tag{5.4}$$

whenever $x \in B_\gamma(\bar{x})$. Selecting arbitrary minimizers

$$u_\lambda(x) \in \operatorname{argmin} \Big(f(\cdot) + \frac{1}{\sqrt{\lambda}} d(F(u),;\Omega) + \frac{1}{2\lambda}\|x - \cdot\|^2 \Big), \quad x \in B_\gamma(\bar{x}), \tag{5.5}$$

we conclude that $u_\lambda(x) \in B_\gamma(\bar{x}) \cap F^{-1}(\Omega)$ for all $x \in B_{\gamma/2}(\bar{x})$ and all $\lambda > 0$ sufficiently small, along with the two convergence relationships:

$$
\begin{aligned}
&u_\lambda(x) \to x \quad \textit{as} \quad \lambda \downarrow 0, \ \ x \in B_\gamma(\bar{x}), \\
&u_\lambda(x) \to \bar{x} \quad \textit{as} \quad \lambda \downarrow 0 \ \ \textit{and} \ \ x \to \bar{x}.
\end{aligned} \tag{5.6}
$$

Proof. Our arguments below show the fulfillment of representation (5.3) of the infimal convolution (1.3), for all small numbers $\gamma > 0$, and simultaneously justify the subsequent limiting relationship (5.4) under the imposed qualification conditions. Furthermore, in the process of the proof we also justify all the conclusions made for the argminimum functions $u_\lambda(x)$ defined in (5.5).

Observe that the claimed formula (5.3) can be rewritten as

$$
f_\lambda(x) = e_\lambda\left(f(x) + \frac{1}{\sqrt{\lambda}} d\big(F(x);\Omega\big)\right), \quad x \in B_\gamma(\bar{x}),
$$

by the envelop definition (3.1), with the additional conclusion that the minimum in (5.3) is realized as reflected by the notation therein. Denoting

$$
h_\lambda(x) := \inf_u \left(f(u) + \frac{1}{\sqrt{\lambda}} d\big(F(u);\Omega\big) + \frac{1}{2\lambda}\|x-u\|^2\right), \tag{5.7}
$$

we easily get the relationships

$$
h_\lambda(x) = e_\lambda\left(f(x) + \frac{1}{\sqrt{\lambda}} d\big(F(x);\Omega\big)\right) \le \inf_{u \in F^{-1}(\Omega)} \left(f(u) + \frac{1}{2\lambda}\|x-u\|^2\right) = f_\lambda(x)
$$

for all $x \in I\!R^n$ and $\lambda > 0$. Since $h_\lambda(x)$ is increasing with respect to λ, for any sequence of $\lambda_k \downarrow 0$, there is an extended-real-valued function $h\colon I\!R^n \to \overline{I\!R}$ such that $h_{\lambda_k}(x) \uparrow h(x)$ as $k \to \infty$ for each $x \in I\!R^n$, and we have

$$
h_\lambda(x) \le f_\lambda(x) \le f(x) < \infty \ \text{ whenever } \ x \in F^{-1}(\Omega). \tag{5.8}
$$

If $x \notin F^{-1}(\Omega)$, then clearly $h(x) = \infty$.

To proceed further, observe that the qualification condition (5.1) ensures that the constraint system $F(x) - \Omega$ is *metrically regular* around $(\bar{x}, 0)$ due to the coderivative criterion of metric regularity; see, e.g., [17, Theorem 9.43 and Example 9.44]. The latter provides the existence of $\kappa > 0$ and $\gamma > 0$ such that

$$
d\big(x; F^{-1}(\Omega)\big) \le \kappa\, d\big(F(x);\Omega\big) \ \text{ for all } \ x \in B_{2\gamma}(\bar{x}). \tag{5.9}
$$

Apply now [1, Proposition 1] and get for all $\lambda > 0$ sufficiently small that

$$
\begin{aligned}
f_\lambda(x) &= e_\lambda\left(f + \frac{1}{\sqrt{\lambda}} d\big(F(\cdot);\Omega\big)\right)(x) \\
&= \inf_{\|u-x\|<\gamma} \left\{f(u) + \frac{1}{\sqrt{\lambda}} d\big(F(u);\Omega\big) + \frac{1}{2\lambda}\|x-u\|^2\right\}.
\end{aligned} \tag{5.10}
$$

Substituting (5.10) and (5.9) into (5.8) gives us

$$f(x) \geq h_\lambda(x) \geq \inf_u \left(f(u) + \frac{1}{\kappa\sqrt{\lambda}} d(u, F^{-1}(\Omega)) + \frac{1}{2\lambda}\|x - u\|^2 \right)$$

whenever $x \in B_\gamma(\bar{x}) \cap F^{-1}(\Omega)$. Hence for such x we have

$$f(x) \geq h_\lambda(x) \geq \inf_u \left(f(u) + \frac{1}{2\lambda}\|x - u\|^2 \right) = e_\lambda(f)(x).$$

Since $e_\lambda(f)(x) \uparrow f(x)$ as $\lambda \downarrow 0$, the latter implies that $f(x) = h(x)$ for all vectors $x \in B_\gamma(\bar{x}) \cap F^{-1}(\Omega)$, and thus

$$h(x) = f(x) + \delta_{F^{-1}(\Omega)}(x), \quad x \in B_\gamma(\bar{x}).$$

Observe next that the minimum is achieved in the infimal convolution (5.7) for all $\lambda > 0$ sufficiently small due to the assumed lower semicontinuity and prox-boundedness of f and the facts presented at the beginning of Section 3. Let us justify the existence of $\gamma > 0$ such that all the minimizers $u_\lambda(x)$ of $h_\lambda(x)$ for $x \in B_{\gamma/2}(\bar{x}) \cap F^{-1}(\Omega)$ lie inside of the ball $B_\gamma(\bar{x})$. We first show that for every $\varepsilon > 0$ there are $\gamma > 0$ and $\widehat{\lambda} > 0$ such that

$$\lambda^{-1}\|u_\lambda(x) - x\|^2 \leq \varepsilon \text{ for all } x \in B_\gamma(\bar{x}) \cap F^{-1}(\Omega) \text{ and } 0 < \lambda < \widehat{\lambda}. \tag{5.11}$$

Assuming by the contrary that (5.11) fails, find $\varepsilon > 0$ and sequences $\lambda_k \downarrow 0$ and $x_k \to \bar{x}$ as $k \to \infty$ with $x_k \in F^{-1}(\Omega)$ satisfying

$$\lambda_k^{-1}\|u_{\lambda_k}(x_k) - x_k\|^2 > \varepsilon \text{ for all } k \in I\!N. \tag{5.12}$$

Denoting $u_{\lambda_k}(x_k) =: u_k$, we get from the above that

$$\begin{aligned} h_{4\lambda_k}(x_k) &\leq f(u_k) + \frac{1}{2\sqrt{\lambda_k}} d(F(u_k); \Omega) + \frac{1}{8\lambda_k}\|x_k - u_k\|^2 \\ &= h_{\lambda_k}(x_k) - \frac{1}{2}\left(\frac{1}{\sqrt{\lambda_k}} d(F(u_k); \Omega) + \frac{3}{4\lambda_k}\|x_k - u_k\|^2 \right), \end{aligned}$$

which imply the estimates

$$2\left(h_{\lambda_k}(x_k) - h_{\lambda_k/4}(x_k)\right) \geq \frac{1}{\sqrt{\lambda_k}} d(F(u_k); \Omega) + \frac{3}{4\lambda_k}\|x_k - u_k\|^2 \geq 0 \tag{5.13}$$

for all $k \in I\!N$. Taking into account that $x_k \to \bar{x}$, that the function $h(x)$ is continuous on $F^{-1}(\Omega) \cap B_\gamma(\bar{x})$, and that the pointwise convergence of an increasing sequence of functions implies its epi-convergence (see, e.g., [17, Theorem 7.10]), we arrive at

$$\liminf_{k \to \infty} h_{\lambda_k/4}(x_k) \geq h(\bar{x}), \quad \text{and thus}$$

$$0 \leq \liminf_{k \to \infty} \left(h_{\lambda_k}(x_k) - h_{\lambda_k/4}(x_k)\right) \leq \liminf_{k \to \infty} \left(h(x_k) - h(\bar{x})\right) \leq 0.$$

The latter yields by (5.13) that $\lambda_k^{-1}\|x_k - u_k\|^2 \to 0$, which contradicts (5.12). Thus we have (5.11) that implies the first convergence in (5.6). Letting further $0 < \lambda < \min\{\gamma^2/(4\varepsilon), \widehat{\lambda}\}$ in (5.11) gives us

$$\|u_\lambda(x) - \bar{x}\| \le \|u_\lambda(x) - x\| + \|x - \bar{x}\| \le \sqrt{\lambda\varepsilon} + \gamma/2 < \gamma,$$

which justifies the second convergence in (5.6) as well.

To proceed with the proof of representation (5.3), fix λ and x as above and employ the stationary condition in terms of basic subgradients (2.5) to the optimal solution $u_\lambda(x)$ of the unconstrained optimization problem (5.7). This gives

$$0 \in \partial_u\left(f(\cdot) + \frac{1}{\sqrt{\lambda}}d(F(\cdot;\Omega)) + \frac{1}{2\lambda}\|x - \cdot\|^2\right)(u_\lambda(x)). \tag{5.14}$$

Using now in (5.14) the basic subdifferential sum and chain rules (see, e.g., [13, Subsection 3.2.1] and [17, Section 10B]) and taking into account that the composite distance function in (5.7) is locally Lipschitzian, we get the inclusion

$$0 \subset \partial f(u_\lambda(x)) + \frac{1}{\sqrt{\lambda}}\nabla F(u_\lambda(x))^T \partial d(F(u_\lambda(x);\Omega) + \frac{1}{\lambda}(u_\lambda(x) - x). \tag{5.15}$$

It follows from [13, Theorem 1.97] and [17, Example 8.53] that

$$\partial d(F(u);\Omega) = \begin{cases} N_\Omega((F(u)) \cap I\!B & \text{if } F(u) \in \Omega, \\ \dfrac{F(u) - \Pi_\Omega(F(u))}{d(F(u);\Omega)} & \text{if } F(u) \notin \Omega, \end{cases} \tag{5.16}$$

where $\Pi_\Omega(v) := \{w \in \Omega |\ \|w - v\| = d(v;\Omega)\}$ stands for the Euclidean projector of v into Ω. Multiplying both parts of inclusion (5.14) by $\sqrt{\lambda}$, we find $y_\lambda \in \partial d(F(u_\lambda(x);\Omega)$ and $z_\lambda \in \partial f(u_\lambda(x))$ such that

$$\nabla F(u_\lambda(x))^T y_\lambda = -\sqrt{\lambda}z_\lambda - \frac{1}{\sqrt{\lambda}}(u_\lambda(x) - x). \tag{5.17}$$

Let us show that (5.17) yields $\|\sqrt{\lambda}z_\lambda\| \to 0$ as $\lambda \downarrow 0$. Assume the contrary and then find a sequence of $\lambda_k > 0$ such that

$$\|\sqrt{\lambda_k}z_{\lambda_k}\| \to \nu > 0 \quad \text{as} \quad k \to \infty. \tag{5.18}$$

The second relationship in (5.6) implies (by using the diagonal process without relabeling) that $u_{\lambda_k}(x_k) \to \bar{x}$ as $k \to \infty$ along some sequence $x_k \to \bar{x}$. Recalling now that $\|y_\lambda\| \le 1$ for all $\lambda > 0$ by the Lipschitz continuity of the distance function with constant one, we get from (5.18) that the sequence $\{y_{\lambda_k}/\|\sqrt{\lambda_k}z_{\lambda_k}\|\}$ is bounded, and hence it contains a convergence subsequence. This allows us to conclude, by the subdifferential formulas (5.16) and the closed-graph property of the basic normal cone, that there is $y \in I\!R^n$ satisfying the relationships

$$\frac{y_{\lambda_k}}{\|\sqrt{\lambda_k}z_{\lambda_k}\|} \to y \in N_\Omega(F(\bar{x})) \quad \text{as} \quad k \to \infty. \tag{5.19}$$

Using further the above construction of z_λ, the closed graph property of the basic subdifferential (2.5), and the definition of its singular counterpart (2.6), we find $z \in I\!R^n$ with $\|z\| = 1$ satisfying the relationships

$$\frac{z_{\lambda_k}}{\|z_{\lambda_k}\|} \to z \in \partial^\infty f(\bar{x}) \quad \text{as} \quad k \to \infty. \tag{5.20}$$

Now consider equality (5.17) along the selected sequences $\{\lambda_k, x_k, y_{\lambda_k}, z_{\lambda_k}\}$ and divide its both sides by $\|\sqrt{\lambda_k} z_{\lambda_k}\|$. Passing to the limit as $k \to \infty$ and using conditions (5.11), (5.19), and (5.20) along these sequences, we get

$$0 \neq z = -y \in \partial^\infty f(\bar{x}) \cap \big\{ -\nabla F(\bar{x})^T y \big|\, y \in N_\Omega\big(F(\bar{x})\big)\big\},$$

which contradicts the qualification assumption (5.2). Taking (5.11) into account, we thus arrive at the limiting relationship

$$\lim_{\lambda \downarrow 0} \left[\sqrt{\lambda} z_\lambda - \frac{1}{\sqrt{\lambda}}\big(u_\lambda(x) - x\big)\right] = 0 \quad \text{for all} \quad x \in B_\gamma(\bar{x}) \cap F^{-1}(\Omega). \tag{5.21}$$

To complete the proof of the theorem, it remains to show that under the qualification condition (5.1) we have the inclusion

$$u_\lambda(x) \in F^{-1}(\Omega) \quad \text{for all} \quad x \in B_\gamma(\bar{x}) \quad \text{and small} \quad \lambda > 0 \tag{5.22}$$

whenever $u_\lambda(x)$ satisfies (5.17), which is necessary for the optimal solutions. Indeed, the closed-graph property of the basic normal cone yields that the qualification condition (5.1) holds true in a neighborhood of the reference point $\bar{x}$. Furthermore, the subdifferential formula (5.16) implies that $\|y\| = 1$ for all $y \in \partial d(F(u); \Omega)$ with $u \notin F^{-1}(\Omega)$. This means that assuming the opposite to (5.22) leads us to a contradiction between the qualification condition (5.1) from one side and the relationships in (5.17) and (5.21) from the other. Thus we arrive at the inequality

$$\begin{aligned} &\min_{u \in I\!R^n} \Big(f(u) + \frac{1}{\sqrt{\lambda}} d\big(F(u); \Omega\big) + \frac{1}{2\lambda}\|x - u\|^2\Big) \\ &\geq \min_{u \in F^{-1}(\Omega)} \left(f(u) + \frac{1}{2\lambda}\|x - u\|^2\right) = f_\lambda(x) \end{aligned} \tag{5.23}$$

for all $x \in B_\gamma(\bar{x}) \cap F^{-1}(\Omega)$ and $\lambda > 0$ sufficiently small. Combining (5.8) and (5.23) gives representation (5.3) and completes the proof of the theorem. $\triangle$

Remark 5.2 (special cases and extensions of Theorem 5.1).

(i) Observe that the first qualification condition (5.1) in Theorem 5.1 surely holds if the Jacobian matrix $\nabla F(\bar{x})$ is of full rank, which is equivalent to the metric regularity/linear openness of the constraint mapping F around $\bar{x}$; see, e.g., [13, Theorem 1.57]. The second qualification condition (5.2) of this theorem is automatic if the cost function f is locally Lipschitzian around $\bar{x}$. This follows from the singular subdifferential characterization $\partial^{\infty} f(\bar{x}) = \{0\}$ of the local Lipschitzian property; see [13, Theorem 3.52] and [17, Theorem 9.13].

(ii) As follows from the proof of Theorem 5.1 and the results of [13, 17] applied therein, this theorem can be extended to large classes of nonsmooth and set-valued constraint mappings F. Furthermore, applications of the corresponding results from [13] allows us to obtain extensions of Theorem 5.1 to constrained optimization problems in infinite-dimensional spaces under additional "sequential normal compactness" conditions that are automatic in finite dimensions.

Next let us employ the constraint qualification (5.1) of Theorem 5.1 to derive new lower subdifferential optimality conditions for (1.1)—from the underlying upper subdifferential one—in terms of infimal convolutions. We show furthermore that the convolution conditions obtained imply, under the qualification condition (5.2), a more conventional optimality condition in terms of basic subgradients of the cost function f. As in the other necessary optimality conditions studied in this paper, we consider for simplicity only the case of global minimizers for (1.1).

Theorem 5.3 (lower subdifferential optimality conditions under constraint qualifications). *Let $\bar{x} \in \operatorname{argmin}_{x \in F^{-1}(\Omega)} f(x)$. Assume that the constraint qualification* (5.1) *is satisfied. Then we have the set inclusion*

$$-\partial f_{\lambda}(\bar{x}) \subset \nabla F(\bar{x})^T N_{\Omega}(F(\bar{x})) \tag{5.24}$$

for all $\lambda > 0$ sufficiently small. If in addition the qualification condition (5.2) *also holds and the set Ω is normally regular at $F(\bar{x})$, then*

$$0 \in \partial f(\bar{x}) + \nabla F(\bar{x})^T N_{\Omega}(F(\bar{x})). \tag{5.25}$$

Proof. It follows from assertion (4.7) of Proposition 4.3, obtained in turn from the upper subdifferential condition of Theorem 4.1, that

$$-\partial f_{\lambda}(\bar{x}) \subset N_{F^{-1}(\Omega)}(\bar{x}) \tag{5.26}$$

while $\lambda > 0$ is sufficiently small. Since

$$N_{F^{-1}(\Omega)}(\bar{x}) \subset \nabla F(\bar{x})^T N_{\Omega}(F(\bar{x})) \tag{5.27}$$

under the constraint qualification (5.1) (see, e.g., [13, Corollary 3.42]), inclusion (5.26) implies that the convolution condition (5.24) holds for all small $\lambda > 0$.

Let us show next that the convolution condition (5.24) implies that of (5.25) under the additional assumptions made. It follows from (5.24) that for any small $\lambda > 0$ and any $x^*_{\lambda} \in \partial f_{\lambda}(\bar{x})$ we have

$$-x^*_{\lambda} \in F(\bar{x})^T N_{\Omega}(F(\bar{x})). \tag{5.28}$$

By the proximal representation (2.5) of the basic subdifferential and taking into account the subsequent passage to the limit, it is possible to assume without loss of generality that $x_\lambda^* \in \partial_p f_\lambda(\bar{x})$ in (5.28) and as mentioned above, we may always suppose that the function $f + \delta_{F^{-1}(\Omega)}$ is prox-bounded. Applying [4, Lemma A5] in this case allows us to deduce from $x_\lambda^* \in \partial_p f_\lambda(\bar{x})$ that

$$x_\lambda^* \in \partial_p\big(f + \delta_{F^{-1}(\Omega)}\big)(\bar{x} - \lambda x_\lambda^*) \ \text{ and } \ \big(f + \delta_{F^{-1}(\Omega)}\big)(\bar{x} - \lambda x_\lambda^*) = f_\lambda(\bar{x}) - \frac{\lambda}{2}\|x_\lambda^*\|^2$$

for all λ sufficiently small. Taking now a sequence $\lambda_k \downarrow 0$ as $k \to \infty$ and using the relationships above, we find a sequence $\{x_k^*\}$ with $x_k^* := x_{\lambda_k}^*$ such that

$$\begin{cases} 0 \in x_k^* + F(\bar{x})^T N_\Omega\big(F(\bar{x})\big), \ \ x_k^* \in \partial_p\big(f + \delta_{F^{-1}(\Omega)}\big)(\bar{x} - \lambda_k x_k^*), \\ \\ \text{and } \ \big(f + \delta_{F^{-1}(\Omega)}\big)(\bar{x} - \lambda_k x_k^*) = f_{\lambda_k}(\bar{x}) - \dfrac{\lambda_k}{2}\|x_k^*\|^2 \ \text{ for all } \ k \in I\!N. \end{cases} \tag{5.29}$$

Assume first that the sequence $\{x_k^*\}$ is bounded and select its subsequence (with no relabeling), which converges to some $\bar{x}^*$. Then we have from (5.29) that

$$-\bar{x}^* \in F(\bar{x})^T N_\Omega\big(F(\bar{x})\big) \ \text{ and } \ \bar{x}^* \in \partial\big(f + \delta_{F^{-1}(\Omega)}\big)(\bar{x}). \tag{5.30}$$

Employing the subdifferential sum rule for basic subgradients from [13, Theorem 3.36] to the second inclusion in (5.30) and then the one in (5.27) to $N_{F^{-1}(\Omega)}(\bar{x})$, we get under the qualification conditions (5.1) and (5.2) that

$$\bar{x}^* \in \partial f(\bar{x}) + F(\bar{x})^T N_\Omega\big(F(\bar{x})\big). \tag{5.31}$$

Summing up the latter with the first inclusion in (5.30) and taking into account the convexity of the cone $N_\Omega(\bar{x})$ (equivalent to the assumed normal regularity) gives us

$$0 \in \partial f(\bar{x}) + F(\bar{x})^T N_\Omega\big(F(\bar{x})\big) + F(\bar{x})^T N_\Omega\big(F(\bar{x})\big) \subset \partial f(\bar{x}) + F(\bar{x})^T N_\Omega\big(F(\bar{x})\big),$$

which justifies (5.25) provided that the above sequence $\{x_k^*\}$ is bounded.

Assume next that the sequence $\{x_k^*\}$ is unbounded and then arrive at a contradiction. Indeed, it follows from (5.29) that $\bar{x} - \lambda_k x_k^* \in P_{\lambda_k}$ for the projection set P_λ defined in (4.8). Hence using the prox-boundedness and [17, Theorem 1.25] we have $\bar{x} - \lambda_k x_k^* \to \bar{x}$ and $\lambda_k x_k^* \to 0$ as $k \to \infty$, and thus

$$f(\bar{x}) \le \liminf_k \big(f + \delta_{F^{-1}(\Omega)}\big)(\bar{x} - \lambda_k x_k^*) \le \liminf_k e_{\lambda_k}\big(f + \delta_{F^{-1}(\Omega)}\big)(\bar{x}) = f(\bar{x}),$$

which allow us to select a further subsequence to ensure that

$$\big(f + \delta_{F^{-1}(\Omega)}\big)(\bar{x} - \lambda_k x_k^*) \to f(\bar{x}) \ \text{ as } \ k \to \infty.$$

Using (5.29) again, we get that $\sqrt{\lambda_k}\|x_k^*\| \to 0$ as $k \to \infty$, and hence there is $x^* \in I\!R^n$ with $\|x^*\| = 1$ such that

$$\frac{x_k^*}{\|x_k^*\|} \to x^* \in \partial^\infty\big(f + \delta_{F^{-1}(\Omega)}\big)(\bar{x}). \tag{5.32}$$

Note also that $-x^* \in F(\bar{x})^T N_\Omega(F(\bar{x}))$ by (5.29). Applying further to (5.32) the subdifferential sum rule for singular subgradients from [13, Theorem 3.36], which holds under the same qualification condition (5.2), we get

$$x^* \in \partial^\infty f(\bar{x}) + N_{F^{-1}(\Omega)}(\bar{x}) \subset \partial^\infty f(\bar{x}) + F(\bar{x})^T N_\Omega\big(F(\bar{x})\big).$$

Similarly to the case of (5.31) this implies the relationships

$$0 \neq x^* \in \partial^\infty f(\bar{x}) \cap \big\{ -\nabla F(\bar{x})^T y \big|\ y \in N_\Omega\big(F(\bar{x})\big)\big\}$$

in contradiction to (5.2) and thus completes the proof of the theorem. $\triangle$

Remark 5.4 (discussions on lower subdifferential optimality conditions).
(i) There are two major issues worth emphasizing in the new convolution optimality condition (5.24) in comparison with known lower subdifferential ones for problems of constrained minimization:

—it holds as a set inclusion for *all* subgradients of $\partial f_\lambda(\bar{x})$ and all small $\lambda > 0$, not just for an element of $\partial f(\bar{x})$ as in (5.25);

—it does not required any qualification condition imposed on the cost function f as for the more conventional one in (5.25).

(ii) The lower subdifferential condition (5.25) holds under the qualification assumptions (5.1) and (5.2) with no normal regularity requirement on Ω; see, e.g., [14, Theorem 5.24]. The purpose of this part in the proof of Theorem 5.3 is to illustrate relationships between such optimality conditions of the conventional lower subdifferential type and the convolution optimality condition (5.25) that closely connected to the upper subdifferential condition of Theorem 4.1.

(iii) The results obtained in Theorem 5.3 can be extended by applying similar arguments to more general optimization problems in both finite-dimensional and infinite-dimensional spaces; cf. the relevant discussions in Remark 5.2.

6 Second-Order Optimality Conditions

This section is devoted to applying infimal convolutions to the study of *second-order* necessary conditions and sufficient conditions for optimality in minimization problems. For these purposes we use in what follows the second-order generalized differential constructions defined in Section 2.

Let us first define three kinds of second-order optimality conditions for (unconstrained) extended-real-valued functions studied in [7] from the viewpoint of sufficient conditions for optimality.

Definition 6.1 (classification of second-order optimality conditions). *Take $f\colon I\!R^n \to \overline{I\!R}$ and assume that the first-order stationary condition $0 \in \partial_p f(\bar{x})$ is satisfied at some point $\bar{x} \in \operatorname{dom} f$. Then we say that:*

(i) *The necessary (resp. sufficient) condition of the* FIRST KIND *holds at $\bar{x}$ when (resp. there exists $\beta > 0$ such that)*

$$f''_-(\bar{x}, 0; v) \geq 0 \ (resp.\ \geq \beta)\ \ for\ all\ \ v \in I\!R^n\ \ with\ \ \|v\| = 1.$$

(ii) *The necessary (resp. sufficient) condition of the* SECOND KIND *holds at* $\bar{x}$ *when (resp. there exists* $\beta > 0$ *such that)*

$$\forall v \in \operatorname{Dom} D\left(\partial_p f\right)(\bar{x}, 0)(\cdot) \text{ with } \|v\| = 1$$
$$\exists z \in D\left(\partial_p f\right)(\bar{x}, 0)(v) \text{ such that } \langle z, v\rangle \geq 0 \ \text{(resp. } \geq \beta\text{)}.$$

(iii) *The necessary (resp. sufficient) condition of the* THIRD KIND *holds at* $\bar{x}$ *when (resp. there exists* $\beta > 0$ *such that)*

$$\forall v \in I\!R^n \text{ with } \|v\| = 1,\ \forall x^* \in \widehat{D}^*\left(\partial_p f\right)(\bar{x}, 0)(v) \text{ we have } \langle x^*, v\rangle \geq 0 \ (\text{resp. } \geq \beta).$$

The study of of the first kind of optimality conditions were initiated in [2] and later were continued in [18, 19, 20]. It is worth noting that the sufficient optimality condition of the first kind is equivalent to the concept of a strict local minimum of order two; cf. [2, 18, 20].

Definition 6.2 (strict minimizers of order two). *Let* $f\colon I\!R^n \to \overline{I\!R}$ *be finite in* $\bar{x} \in U \subset I\!R^n$. *We say that:*

(i) *The point* $\bar{x}$ *is a* STRICT MINIMIZER *of* ORDER TWO *for* f *relative to* U *if there is* $\beta > 0$ *such that*

$$f(x) \geq f(\bar{x}) + \frac{\beta}{2}\|x - \bar{x}\|^2 \quad \text{for all} \quad x \in U.$$

(ii) *The point* $\bar{x}$ *is a* STRICT LOCAL MINIMIZER *of* ORDER TWO *for* f *if* $\bar{x}$ *is a strict minimizer of order two for* f *relative to* $B_\gamma(\bar{x})$ *with some* $\gamma > 0$.

(iii) *A number* $\bar{\beta} > 0$ *is a* THRESHOLD VALUE *for a strict local minimizer* $\bar{x}$ *of order two for the function* f *if whenever* $0 < \beta < \bar{\beta}$ *there is* $\gamma(\beta) > 0$ *such that*

$$f(x) \geq f(\bar{x}) + \frac{\beta}{2}\|x - \bar{x}\|^2 \text{ for all } x \in B_{\gamma(\beta)}(\bar{x}).$$

Together with the quadratic estimates in Definition 6.2 we also consider those in the prox-boundedness form of Section 3:

$$f(x) \geq f(\bar{x}) - \frac{r}{2}\|x - \bar{x}\|^2, \quad x \in I\!R^n, \tag{6.1}$$

and denote by $r(f, \bar{x})$ the infimum of $r \in I\!R$ for which (6.1) holds. As discussed in Section 3, the quadratically shifted convolutions $e_\lambda(f) - \frac{1}{\lambda}\|\cdot\|^2$ are always concave being well-defined (finite) for all $0 < \lambda < (\max\{0, r(f, \bar{x})\})^{-1}$. Furthermore, by [7, Theorem 21] we have the equality

$$e_\lambda\Big(\frac{1}{2}f''_-(\bar{x}, 0; \cdot)\Big)(v) = \frac{1}{2}\big(e_\lambda(f)\big)''_-(\bar{x}, 0; v), \quad v \in I\!R^n, \tag{6.2}$$

for such λ, and hence the function $(e_\lambda(f))''_-(\bar{x}, 0; \cdot) - \frac{1}{\lambda}\|\cdot\|^2$ is also concave.

Based on the paraconcavity of infimal convolutions, the equivalence between the second-order sufficient conditions of all the three kinds in Definition 6.1 is established in [7, Theorem 66] under natural assumptions. In fact, the proof given in [7] allows us to reveal the equivalence between the necessary optimality conditions of Definition 6.1 as well; thus we come up to the following results.

Theorem 6.3 (equivalence between second-order optimality conditions). *Let $f\colon I\!R^n \to \overline{I\!R}$ be prox-bounded, and let $0 \in \partial_p f(\bar{x})$. Assume also that there exists $\mu > 0$ for which both functions $f - \frac{\mu}{2}\|\cdot\|^2$ and $f''_-(\bar{x}, 0, \cdot) - \mu\|\cdot\|^2$ are concave. Then the following assertions hold:*

(i) *All the necessary optimality conditions of Definition* 6.1 *are equivalent.*

(ii) *All the sufficient conditions of Definition* 6.1 *are equivalent with the same number $\beta > 0$ therein.*

(iii) *The coderivative $\widehat{D}^*(\partial_p f)(\bar{x}, 0)(\cdot)$ is a bounded set if it is nonempty and* $\mathrm{Dom}(D(\partial_p f)(\bar{x}, 0)(\cdot) = I\!R^n$.

Denote further 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 $\langle Q, M\rangle := \text{trace } M^T Q$ for any $Q, M \in \mathcal{S}(n)$. The conic subset of $\mathcal{S}(n)$ is denoted by $\mathcal{P}(n)$, and it corresponds to all the positively definite matrices. Recall [5, Proposition 8] that the condition $I + 2\lambda Q \in \text{int } P(n)$ with the unit matrix I is necessary and sufficient condition for the relationship $(q_Q)_\lambda = q_{Q_\lambda}$ with some $Q_\lambda \in S(n)$, where

$$(q_Q)_\lambda(x) := \inf_u \left\{ \langle Qu, u\rangle + \frac{1}{2\lambda}\|x - u\|^2 \right\} \quad \text{and} \quad q_{Q_\lambda}(x) := \langle Q_\lambda x, x\rangle.$$

In this case we associate Q_λ with a quadratic form.

The following result from [6] concerning jets (2.8) of infimal convolutions is useful in the subsequent analysis.

Proposition 6.4 (infimal convolutions of jets). *Let $f\colon I\!R^n \to I\!R$ be prox-bounded with the proximal threshold $(\bar{r}(f))^{-1}$. Then for every $Q \in \partial^{2,-} f(\bar{x}, 0)$ there is $r \geq (\bar{r}(f))^{-1}$ such that whenever $0 < \lambda < \frac{1}{r}$ with $I + \lambda Q \in \text{int } P(n)$ we have*

$$2\left(\frac{1}{2}Q\right)_\lambda \in \partial^{2,-}(e_\lambda f)(\bar{x}, 0) \subset \partial^{2,-} f(\bar{x}, 0).$$

We now return to the original minimization problem (1.1) and apply the above constructions and results to the extended-real-valued function $f + \delta_{F^{-1}(\Omega)}$. The next theorem provides second-order necessary optimality conditions for problem (1.1) via its infimal convolutions (1.3) and also establishes convolution characterizations of strict local minimizers of order two.

Theorem 6.5 (second-order optimality conditions via infimal convolutions). *Given $f\colon I\!\!R^n \to \overline{I\!\!R}$ finite at $\bar{x}$ and given $\lambda > 0$ sufficiently small, the following assertions hold:*

(i) *If $\bar{x} \in \operatorname{argmin}_{F^{-1}(\Omega)} f$, then $0 \in \partial_p f_\lambda(\bar{x})$ and all three kinds of second-order necessary optimality conditions of Definition 6.1 are satisfied for f_λ at $\bar{x}$.*

(ii) *The point $\bar{x}$ with $0 \in \partial_p f_\lambda(\bar{x})$ is a strict local minimizer of order two for problem (1.1) if and only if at least one of the three kinds of the second-order sufficient condition in Definition 6.1 is satisfied for f_λ at $\bar{x}$. Moreover, we can take the same value of $\beta > 0$ in all these sufficient conditions.*

Proof. To justify the second-order necessary optimality conditions in (i), observe that by Proposition 3.1 we have $\bar{x} \in \operatorname{argmin}_{x \in I\!\!R^n} f_\lambda$ and hence $0 \in \partial_p f_\lambda(\bar{x})$ for all small λ. It follows further from (6.2) that both functions $f_\lambda - \frac{1}{2\lambda}\|\cdot\|^2$ and $(f_\lambda)''_-(\bar{x},0,\cdot) - \frac{1}{\lambda}\|\cdot\|^2$ are concave. Thus all the three necessary optimality conditions are equivalent for f_λ by Theorem 6.3. Considering now the second-order necessary condition of the first kind for f_λ at $\bar{x}$ and taking into account that $f_\lambda(x) \geq f_\lambda(\bar{x})$, we arrive at the inequality

$$(f_\lambda)''(\bar{x}, 0; v) = \liminf_{u \to v, t \downarrow 0} \frac{2}{t^2}\Big(f(\bar{x} + tu) - f(\bar{x}) - t\langle 0, u\rangle\Big) \geq 0,$$

which implies the fulfillment of all the three second-order necessary optimality conditions for the infimal convolutions f_λ at $\bar{x}$ whenever $\lambda > 0$ is sufficiently small and thus justifies assertion (i).

To prove (ii), let $\bar{x} \in F^{-1}(\Omega)$ be a strict local minimizer of order two for (1.1) and find $\gamma > 0$ such that

$$\left[f(x) + \delta_{F^{-1}(\Omega)}(x)\right] - \left[f(\bar{x}) + \frac{\beta}{2}\|x - \bar{x}\|^2\right] \geq 0 \ \text{ for all } \ x \in B_\gamma(\bar{x}).$$

By the jet definition (2.8) we have $(0, \beta I) \in \partial^{2,-}\left(f + \delta_{F^{-1}(\Omega)}\right)(\bar{x})$ and then

$$2\Big(\frac{\beta}{2}I\Big)_\lambda \in \partial^{2,-} e_\lambda\left(f + \delta_{F^{-1}(\Omega)}\right)(\bar{x}, 0) = \partial^{2,-} f_\lambda(\bar{x}, 0) \tag{6.3}$$

by Proposition 6.4. It is easy to see that $2(\frac{\beta}{2}I)_\lambda = \frac{\beta}{(1+\lambda\beta)}I$. Thus (6.3) gives

$$\frac{\beta}{(1+\lambda\beta)}I \in \partial^{2,-} f_\lambda(\bar{x}, 0),$$

which ensures in turn by (2.8) the existence of $\varphi \in C^2(I\!\!R^n)$ and $\gamma > 0$ such that

$$\left(0, \frac{\beta}{(1+\lambda\beta)}I\right) = \left(\nabla\varphi(\bar{x}), \nabla^2\varphi(\bar{x})\right) \ \text{ and } \ f_\lambda(x) - \varphi(x) \geq f_\lambda(\bar{x}) - \varphi(\bar{x})$$

when $x \in B_\gamma(\bar{x})$. By the Taylor expansion for φ we deduce from the above that

$$\begin{aligned} f_\lambda(x) &\geq f_\lambda(\bar{x}) + \frac{1}{2}\frac{\beta}{(1+\lambda\beta)}\langle I(x-\bar{x}), (x-\bar{x})\rangle + o\left(\|(x-\bar{x})\|^2\right) \\ &= f_\lambda(\bar{x}) + \frac{1}{2}\frac{\beta}{(1+\lambda\beta)}\|(x-\bar{x})\|^2 + o\left(\|(x-\bar{x})\|^2\right). \end{aligned}$$

This implies that whenever $0 < \beta' < \frac{\beta}{(1+\lambda\beta)}$ there is $0 < \gamma' < \gamma$ such that

$$\begin{aligned} f_\lambda(x) &:= \inf_{u \in F^{-1}(\Omega)} \left(f(u) + \frac{1}{2\lambda}\|x-u\|^2\right) \\ &\geq f_\lambda(\bar{x}) + \frac{\beta'}{2}\|x-\bar{x}\|^2 \quad \text{for all } x \in B_{\gamma'}(\bar{x}) \cap F^{-1}(\Omega), \end{aligned} \tag{6.4}$$

and thus $\bar{x}$ is a strict local minimizer of order two for the convolutions f_λ as $\lambda > 0$.

We have from Theorem 6.3 that all the three kinds of the second-order sufficient conditions of Definition 6.1 hold for f_λ at $\bar{x}$ if one of them holds. Moreover, the same threshold value $\frac{\beta}{2(1+\lambda\beta)}$ can be used in each sufficient condition. It follows from (6.4) by definition (2.12) that

$$(f_\lambda)''_-(\bar{x}, 0; v) \geq \beta' > 0 \quad \text{for all } v \in \mathbb{R}^n \text{ with } \|v\| = 1,$$

i.e., the sufficient condition of the first kind is satisfied for f_λ at $\bar{x}$. This justifies the "only if" part in assertion (ii) of the theorem.

It remains to prove the "if" part therein. Assuming that one of the sufficient conditions holds for f_λ at $\bar{x}$, we get from Theorem 6.3 that all three of them holding. It follows from the results in [2, 18] that the validity of the sufficient condition of the first kind ensures that the $\bar{x}$ is a strong minimizer of order two for the convolutions f_λ as $\lambda > 0$. Then there are $\beta > 0$ and $\gamma > 0$ such that

$$f(x) \geq f_\lambda(x) \geq f_\lambda(\bar{x}) + \frac{\beta}{2}\|x-\bar{x}\|^2 \quad \text{whenever } x \in B_\gamma(\bar{x}) \tag{6.5}$$

when $(f_\lambda)''_-(\bar{x}, 0; v) > \beta > 0$ for all $v \in \mathbb{R}^n$ with $\|v\| = 1$. Let us finally show that $f(\bar{x}) = f_\lambda(\bar{x})$ if $\lambda > 0$ is sufficiently small, and thus (6.5) is a strict local minimizer of order two for the original problem (1.1).

To proceed, observe from Proposition 6.4 the implication

$$\left[(f_\lambda)''_-(\bar{x}, 0; v) > \beta\right] \Longrightarrow \left[\beta I \in \partial^{2,-} f_\lambda(\bar{x}, 0) \subset \partial^{2,-}\left(f + \delta_{F^{-1}(\Omega)}\right)(\bar{x}, 0)\right].$$

Applying now the same arguments as in the above proof of (6.4), we conclude that for any $0 < \beta' < \beta$ there is $\gamma' > 0$ such that

$$f(x) \geq f(\bar{x}) + \frac{\beta'}{2}\|x-\bar{x}\|^2 \quad \text{whenever } x \in B_{\gamma'}(\bar{x}) \cap F^{-1}(\Omega).$$

The latter yields by [1, Proposition 1] that

$$
\begin{aligned}
f_\lambda(\bar{x}) = e_\lambda\left(f+\delta_{F^{-1}(\Omega)}\right)(\bar{x}) &= \inf_{u\in F^{-1}(\Omega),\,\|u-\bar{x}\|<\gamma'}\left\{f(u)+\frac{1}{2\lambda}\|u-\bar{x}\|^2\right\}\\
&\geq \inf_{u\in F^{-1}(\Omega),\,\|u-\bar{x}\|<\gamma'}\left\{f(\bar{x})+\frac{1}{2\lambda}\|u-\bar{x}\|^2\right\} = f(\bar{x})
\end{aligned}
$$

for all $\lambda>0$ sufficiently small. This completes the proof of the theorem. △

Note that the equivalence between all the three kinds of second-order sufficient conditions for infimal convolutions has been largely exploited in [7] to study second-order sufficient conditions for strict minimizers of order two in general problems of minimizing nonsmooth functions. In the rest of the paper we focus on developing this convolution approach to the study of *necessary* second-order optimality conditions of all the three types in Definition 6.1 for the original problem (1.1). We need the following result from [7, Lemma 42] concerning the linear transformation $L_\lambda\colon (x,y)\mapsto(x+\lambda y,y)$ and its inverse $L_\lambda\colon (x,y)\mapsto(x-\lambda y,y)$ as $\lambda>0$.

Lemma 6.6 (linear transformations of normal cones to subdifferential graphs). *Let $f\colon \mathbb{R}^n\to\overline{\mathbb{R}}$ be prox-bounded, and let $(\bar{x},\bar{y})\in \operatorname{gph}\partial_p e_\lambda(f)$. Then for any $0<\lambda<\bar{r}(f)^{-1}$ we have the inclusion*

$$
\widehat{N}_{\operatorname{gph}\partial_p f}\left(L_\lambda^{-1}(\bar{x},\bar{y})\right)\subset L_\lambda^T\left(\widehat{N}_{\operatorname{gph}\partial_p e_\lambda(f)}(\bar{x},\bar{y})\right).
$$

Now we are ready to justify the second-order necessary optimality conditions of the first and third kinds in Definition 6.1 for the original problem (1.1) considering the case of global minimizers with no loss of generality.

Theorem 6.7 (second-order necessary optimality conditions of the first and third kinds for the original problem). *Let $\bar{x}\in \operatorname{argmin}_{x\in F^{-1}(\Omega)} f(x)$. Then the second-order necessary optimality conditions of the first and third kinds in problem (1.1) hold for $\bar{x}$.*

Proof. As before, we may always assume that the function $f+\delta_{F^{-1}(\Omega)}$ is prox-bounded. Then applying Theorem 6.5(i) and the second-order convolution relationship (6.2) gives us

$$
\begin{aligned}
\frac{1}{2}(f_\lambda)''_-(\bar{x},0;v) &:= \frac{1}{2}\Big(\left(f+\delta_{F^{-1}(\Omega)}\right)_\lambda\Big)''_-(\bar{x},0;v)\\
&= \frac{1}{2}\Big(e_\lambda\left(f+\delta_{F^{-1}(\Omega)}\right)\Big)''_-(\bar{x},0;v)\geq 0
\end{aligned}
$$

for all $v\in\mathbb{R}^n$ with $\|v\|=1$ and all $0<\lambda<(\max\{0,r(f+\delta_{F^{-1}(\Omega)},\bar{x})\})^{-1}$. Since by Theorem 6.5(i) the necessary optimality condition of the first kind holds for the convolutions f_λ when $\lambda > 0$ is sufficiently small and by applying (6.2) to the function $f + \delta_{F^{-1}(\Omega)}$, we get for all $v \in \mathbb{R}^n$ with $\|v\| = 1$ that

$$\begin{aligned}\frac{1}{2}\left(f + \delta_{F^{-1}(\Omega)}\right)''_-(\bar{x}, 0; v) &\geq e_\lambda\left(\frac{1}{2}\left(f + \delta_{F^{-1}(\Omega)}\right)''_-(\bar{x}, 0, \cdot)\right)(v)\\ &= \frac{1}{2}\Big(e_\lambda\left(f + \delta_{F^{-1}(\Omega)}\right)\Big)''_-(\bar{x}, 0; v)\\ &= \frac{1}{2}(f_\lambda)''_-(\bar{x}, 0; v) \geq 0,\end{aligned}$$

which justifies the validity of the necessary optimality condition of the first kind for the minimizer $\bar{x}$ in the original problem (1.1).

Theorem 6.5 ensures also that the necessary optimality condition of the third kind holds for infimal convolutions. Then (3.4) yields that

$$0 \in \partial_p e_\lambda\left(f + \delta_{F^{-1}(\Omega)}\right)(\bar{x}) \qquad \text{and}$$

$$\langle z, w\rangle \geq 0 \ \text{ for all } \ (z, -w) \in \widehat{N}_{\operatorname{gph} \partial_p e_\lambda(f + \delta_{F^{-1}(\Omega)})}(\bar{x}, 0). \tag{6.6}$$

Now pick an arbitrary element $(z, -u) \in \widehat{N}_{\operatorname{gph} \partial_p(f + \delta_{F^{-1}(\Omega)})}(\bar{x}, 0)$ independent of λ and observe that by Lemma 6.6 we have

$$(z, -u) \in \widehat{N}_{\operatorname{gph} \partial_p(f + \delta_{F^{-1}(\Omega)})}(\bar{x}, 0) \subset L_\lambda^T\left(\widehat{N}_{\operatorname{gph} \partial_p e_\lambda(f + \delta_{F^{-1}(\Omega)})}(\bar{x}, 0)\right)$$

for all $\lambda > 0$. It holds furthermore that

$$\left(L_\lambda^T\right)^{-1}(z, -u) = \big(z, -(u + \lambda z)\big) \in \widehat{N}_{\operatorname{gph} \partial_p e_\lambda(f + \delta_{F^{-1}(\Omega)})}(\bar{x}, 0).$$

Thus it follows from (6.6) that

$$\langle z, u + \lambda z\rangle = \langle z, u\rangle + \lambda\|z\|^2 \geq 0 \ \text{ if } \ (z, -u) \in \widehat{N}_{\operatorname{gph} \partial_p(f + \delta_{F^{-1}(\Omega)})}(\bar{x}, 0) \tag{6.7}$$

whenever $\lambda > 0$. Since (z, u) is independent of λ, we can pass to the limit in (6.7) as $\lambda \downarrow 0$ and arrive in this way at the necessary optimality condition of the third kind for the minimizer $\bar{x}$ in (1.1) by taking into account the coderivative definition (2.10). This completes the proof of the theorem. △

Let us finally justify the validity of the necessary optimality condition of the second kind for problem (1.1). In contrast to the necessary conditions of the first and third kinds established in Theorem 6.7 in the general setting, the second kind of necessary conditions require some additional assumptions.

Proposition 6.8 (necessary optimality conditions of the second kind for the original problem). *Taking $\bar{x} \in \operatorname{argmin}_{x \in F^{-1}(\Omega)} f(x)$, assume that*

$$\operatorname{Dom} D(\partial_p f)(\bar{x}, 0)(\cdot) \subset \operatorname{Dom} \partial f''_-(\bar{x}, 0; \cdot) \tag{6.8}$$

and that for any $v \in I\!R^n$ with $D(\partial_p f)(\bar{x},0)(v) \neq \emptyset$ and $\|v\| = 1$ the supremum in

$$\sup \{ \langle z, v \rangle \mid z \in D(\partial_p f)(\bar{x},0)(v) \}$$

is realized. Then the necessary condition of the second kind for (1.1) *holds at* $\bar{x}$.

Proof. Supposing without loss of generality that the function $f + \delta_{F^{-1}(\Omega)}$ is prox-bounded, we get from [7, Proposition 45] that

$$f''_{-}(\bar{x},u;v) \leq \sup \{ \langle z, v \rangle \mid z \in D(\partial_p f)(\bar{x},u)(v) \}. \tag{6.9}$$

It is easy to see that assumption (6.9) can be equivalently written as the inclusion

$$\{ v \in \operatorname{Dom} D(\partial_p f)(\bar{x},0)(\cdot),\ \|v\| = 1 \} \subset \{ v \mid \partial f''_{-}(\bar{x},0;\cdot)(v) \neq \emptyset,\ \|v\| = 1 \}.$$

Since $\bar{x}$ is a solution to (1.1), it satisfies by Theorem 6.7 the necessary optimality condition of the first kind in Definition 6.1(i) for the original problem (1.1). Hence

$$f''_{-}(\bar{x},0;v) \geq 0 \text{ for all } v \in I\!R^n \text{ with } \|v\| = 1,$$

which implies the inequality

$$\sup \{ \langle z, v \rangle \mid z \in D(\partial_p f)(\bar{x},0)(v) \} \geq 0$$

Thus we arrive at necessary optimality condition of the second kind in Definition 6.1(ii) by the assumptions made. △

References

[1] H. Attouch and D. Azé, Approximation and regularization of arbitrary functions in Hilbert spaces by the Lasry-Lions method, Ann. Inst. H. Poincaré. Anal. Non Linéare10 (1993), 289–312.

[2] A. Auslender, Stability in mathematical programming with nondifferentiable data, SIAM J. Control Optim. 22 (1984), pp 239–254.

[3] R. Correa and C. Lemarechal (1993), Convergence of some algorithms for convex minimization, Math. Program. 6 (1993), 261–275.

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

[5] A. C. Eberhard, M. Nyblom and D. Ralph, Applying generalized convexity notions to jets, in: Generalized Convexity, Generalized Monotonicity: Recent Results (J.-P. Crouzeix et al., eds.), pp. 111–157, Kluwer, Dordrecht, 1998.

[6] A. C. Eberhard, R. Sivakumaran and R. Wenczel, On the variational behavior of the subhessians of the Lasry–Lions envelope, J. Convex Anal. 13 (2006), 647–685.

[7] A. C. Eberhard and R. Wenczel, Some sufficient optimality conditions in nonsmooth analysis, SIAM J. Optim. 20 (2009), 251–296.

[8] M. Fukushima and L. Qi, A globally and superlinearly convergent algorithm for nonsmooth convex minimization, SIAM J. Optim. 6 (1996), 1106–1120.

[9] E. Karas, A. Ribeiro, C. Sagastizábal and M. Solodov, (2009), A bundle-filter method for nonsmooth convex constrained optimization, Math. Program. 116 (2009), 297–320.

[10] R. Mifflin, A quasi-second-order proximal bundle algorithm, Math. Program. 73 (1996), 51–72.

[11] B. S. Mordukhovich, Sensitivity analysis in nonsmooth optimization, in Theoretical Aspects of Industrial Design (D. Field and V. Komkov, eds.), SIAM Proc. Appl. Math. 58, pp. 32–46, SIAM Publ., Philadelphia, 1992.

[12] B. S. Mordukhovich, Necessary conditions in nonsmooth minimization via lower and upper ssubgradients, Set-Valued Anal. 12 (2004), 163–193.

[13] B. S. Mordukhovich, Variational Analysis and Generalized Differentiation, I: Basic Theory, Grundlehren Series (Fundamental Principles of Mathematical Sciences) 330, Springer, Berlin, 2006.

[14] B. S. Mordukhovich, Variational Analysis and Generalized Differentiation, II: Applications, Grundlehren Series (Fundamental Principles of Mathematical Sciences) 331, Springer, Berlin, 2006.

[15] B. S. Mordukhovich and J. V. Outrata, On second-order subdifferentials and their applications, SIAM J. Optim. 12 (2001), 139–169.

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

[17] R. T. Rockafellar and R. J-B. Wets, Variational Analysis, Grundlehren Series (Fundamental Principles of Mathematical Sciences) 317, Springer, Berlin, 2006.

[18] M. Studniarski, Necessary and sufficient conditions for isolated local minima of nonsmooth functions, SIAM J. Control Optim. 24 (1986), 1044–1049.

[19] D. Ward, Characterizations of strict local minima and necessary conditions for weak sharp minima, J. Optim. Theory Appl. 80 (1994), 551–571.

[20] D. Ward, A comparison of second-order epiderivatives: calculus and optimality conditions, J. Math. Anal. Appl. 193 (1995), 465–482.