Open access • Journal Article • DOI:10.1137/110852103

Partial Smoothness, Tilt Stability, and Generalized Hessians — Source link

Adrian S. Lewis, S. Zhang

Institutions: Cornell University

Published on: 29 Jan 2013 - Siam Journal on Optimization (Society for Industrial and Applied Mathematics)

Topics: Variational analysis and Hessian matrix

PARTIAL SMOOTHNESS, TILT STABILITY, AND GENERALIZED HESSIANS*

A. S. LEWIS† AND S. ZHANG‡

Abstract

We compare two recent variational-analytic approaches to second-order conditions and sensitivity analysis for nonsmooth optimization. We describe a broad setting where computing the generalized Hessian of Mordukhovich is easy. In this setting, the idea of tilt stability introduced by Poliquin and Rockafellar is equivalent to a classical smooth second-order condition.

Key words: variational analysis, nonsmooth optimization, second-order, sensitivity analysis, prox-regular, subdifferential continuity, partial smoothness, generalized Hessian, tilt stability

1 Introduction

The distinction between active and inactive constraints is fundamental throughout optimization, underlying optimality conditions, sensitivity analysis, and algorithm design. The notion of "partial smoothness" [8] (along with analogues such as "identifiable surfaces" [21] and "$\mathcal{UV}$ decompositions" [9]) captures the essential geometry associated with activity, and in a fashion suitable for generalization beyond classical nonlinear programming into such domains as semidefinite programming. Partial smoothness illustrates well the power of modern variational analysis as a unifying language for concrete optimization. It is, furthermore, a generic property in concrete settings such as semi-algebraic convex optimization [2].

The partly smooth setting allows intuitive and appealing statements of second-order optimality conditions and associated sensitivity analysis around a "nondegenerate" critical point (where the subdifferential contains zero in its relative interior) [8, 5]. In this case the second-order conditions boil down to the classical smooth case, resulting in the idea of a "strong critical point". Much more general second-order variational analysis is available: see for the example the monographs [20, 3, 11]. A particularly attractive approach is via Mordukhovich's generalized Hessian [11]. That particular theoretical development is natural and compelling, relying simply on two sequential applications of the normal cone construction basic to variational analysis, but computing the generalized Hessian in general can be hard.

Despite computational challenges, the generalized Hessian is clearly a fundamental tool. In particular, [18] considers one of the most basic questions of sensitivity analysis: under what conditions does a local minimizer of a function depend in a Lipschitz fashion on linear perturbations to the function? Assuming the function is both "prox-regular" and "subdifferentially continuous" (as holds, for example, for a composition of a continuous convex function with a C^2-smooth map), this "tilt stability" property turns out to be equivalent to positive-definiteness of the generalized Hessian [18].

We prove two main results. We first show that, for partly smooth, prox-regular, subdifferentially continuous functions, the generalized Hessian is easy to compute at a nondegenerate critical point. Then, as a simple consequence using the characterization of [18], we show that, in this setting, strong criticality is actually equivalent to tilt stability.

*This research was supported in part by National Science Foundation Grant DMS-0806057.
†ORIE, Cornell University, Ithaca, NY 14853, U.S.A. (aslewis@orie.cornell.edu, people.orie.cornell.edu/~aslewis).
‡ORIE, Cornell University, Ithaca, NY 14853, U.S.A. (sz254@cornell.edu).

2 Generalized Hessian mappings of simple nonsmooth Functions

Unless otherwise stated, we follow the notation and terminology of [20]. In particular, $\overline{\mathbb{R}}$ denotes the extended reals, $\partial f(x)$ denotes the set of subgradients of a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ at a point $x \in \mathbb{R}^n$, and $N_S(x)$ denotes the normal cone to set $S \subset \mathbb{R}^n$ at a point $x \in \mathbb{R}^n$. We denote the graph of a set-valued mapping F by gph F.

The concept of tilt stability, introduced in [18], characterizes the case where the minimizing point of a function by adding a small linear term shifts in a Lipschitzian manner and is locally unique.

Definition 2.1. A point $\bar{x}$ will be said to give a *tilt stable* local minimum of the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ if $f(\bar{x})$ is finite and there exists a $\delta > 0$ such that the mapping

$$M : v \mapsto \operatorname{argmin}_{|x-\bar{x}|\le\delta}\{ f(x) - f(\bar{x}) - \langle v, x - \bar{x}\rangle \},$$

is single-valued and Lipschitzian on some neighborhood of $v = 0$, with $M(0) = \bar{x}$.

For a C^2-smooth function f with $\nabla f(\bar{x}) = 0$, the point $\bar{x}$ gives a tilt stable local minimum of f if and only if $\nabla^2 f(\bar{x})$ is positive definite, according to [18, Prop. 1.2]. This fact has been extended to nonsmooth functions in terms of the positivity of a certain generalized Hessian mapping [18].

Definition 2.2. For any point $\bar{x}$ and any subgradient $\bar{v} \in \partial f(\bar{x})$, define the *generalized Hessian mapping* $\partial^2 f(\bar{x}|\bar{v}) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ by

$$\partial^2 f(\bar{x}|\bar{v}) : w \mapsto \{ z \mid (z, -w) \in N_{\mathrm{gph}\,\partial f}(\bar{x}, \bar{v}) \}.$$

For a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ having $0 \in \partial f(\bar{x})$, [18, Thm. 1.3] shows that under certain assumptions, the point $\bar{x}$ gives a tilt stable local minimum of f if and only if the mapping $\partial^2 f(\bar{x}|0)$ is positive definite in the sense that

$$\langle z, w\rangle > 0 \text{ whenever } z \in \partial^2 f(\bar{x}|0)(w),\ w \neq 0.$$

To compute the generalized Hessian mapping, it is sufficient to know $N_{\mathrm{gph}\,\partial f}$. Let's introduce the definition of a manifold first.

Definition 2.3. We say that a set $\mathcal{M} \subset \mathbb{R}^n$ is a C^2*-smooth manifold* of codimension m around a point $\bar{x} \in \mathbb{R}^n$ if $\bar{x} \in \mathcal{M}$ and there is an open set $V \subset \mathbb{R}^n$ such that

$$\mathcal{M} \cap V = \{ x \in V \mid \Phi_i(x) = 0,\ i = 1, \cdots, m \}$$

where Φ_i are C^2-smooth functions with $\nabla\Phi_i(\bar{x})$ linearly independent.

In this case, it is well known that the tangent space to $\mathcal{M}$ at $\bar{x}$ is given by

$$T_{\mathcal{M}}(\bar{x}) = \{\nabla\Phi_i(\bar{x})\}^{\perp}$$

and the normal space to $\mathcal{M}$ at $\bar{x}$ is

$$N_{\mathcal{M}}(\bar{x}) = \Big\{ \sum_i \lambda_i \nabla\Phi_i(\bar{x}) \mid \lambda \in \mathbb{R}^m \Big\}.$$

We call $\Phi_i(x) = 0$ local equations for $\mathcal{M}$.

Our immediate aim is to compute the normal cone to the graph of the normal cone mapping $N_{\mathcal{M}}$. An explicit formula follows from [6, Thm. 3.1], [16, Thm. 7], and [14, Thm. 3.1]- see also [13, Thm. 3.4] and [12, Thm. 1.127]. Here, for completeness and to fix our later notation, we give a self-contained classical approach.

Definition 2.4. When $F : U \to \mathbb{R}^m$ is a C^1-smooth mapping of an open set $U \subset \mathbb{R}^n$, the *rank* of F at a point $x \in U$ is defined as the dimension of range of the gradient $\nabla F(x)$.

The next result shows that functions of constant rank have simple structure.

Theorem 2.5 (constant rank)**.** Suppose $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$ are open sets and $F : U \to V$ is a smooth map with constant rank k. For any point $p \in U$, there exist open sets $U_0 \subset U$ containing p, $V_0 \subset V$ containing $F(p)$ and diffeomorphisms $\varphi : U_0 \to \varphi(U_0)$, $\psi : V_0 \to \psi(V_0)$, with $F(U_0) \subset V_0$, such that

$$\psi \circ F \circ \varphi^{-1}(x_1, \cdots, x_k, x_{k+1}, \cdots, x_m) = (x_1, \cdots, x_k, 0, \cdots, 0).$$

Proof. See [7, Thm. 7.8]. □

Note that the above theorem is also true for C^k-smooth functions ($k \geq 1$), in which case φ and ψ are C^k diffeomorphisms. The following is standard, but we include a proof for convenience.

Proposition 2.6 (Immersion)**.** If $\mathcal{M}$ is a C^2-smooth manifold of codimension m around a point $\bar{x}$, then there exist an open set $U \subset \mathbb{R}^{n-m}$ and an injective C^2-smooth mapping $G : U \to \mathbb{R}^n$ with $G(U) = \mathcal{M}$ locally around $\bar{x}$.

Proof. Since $\mathcal{M}$ is a C^2-manifold of codimension m around $\bar{x}$, then there exists an open set $V \subset \mathbb{R}^n$ such that

$$\mathcal{M} \cap V = \{\, x \in V \mid \Phi_i(x) = 0,\ i = 1, \cdots, m \,\},$$

where Φ_i are C^2-smooth with $\nabla\Phi_i(\bar{x})$ linearly independent. Shrinking V if necessary, we can assume that $\nabla\Phi_i(x)$ are linearly independent for all $x \in V$. The Implicit Function Theorem is stated as follows: Let $F : \mathbb{R}^{n+m} \to \mathbb{R}^m$ be a continuously differentiable function, and let $\mathbb{R}^{n+m}$ have coordinates (x, y). Fix a point $(a_1, \cdots, a_n, b_1, \cdots, b_m) = (a, b)$ with $F(a, b) = c$, where $c \in \mathbb{R}^m$. If the matrix $[(\partial F_i/\partial y_j)(a, b)]$ is invertible, then there exists an open set U containing a, an open set V containing b, and a unique continuously differentiable function $g : U \to V$ such that

$$\{\, (x, g(x)) \mid x \in U \,\} = \{\, (x, y) \in U \times V \mid F(x, y) = c \,\}.$$

According to the Implicit Function Theorem, without loss of generality there exist open sets $U \subset \mathbb{R}^{n-m}$, $W \subset R^m$ and a C^2-smooth function $g : U \to W$ with $\bar{x} \in U \times W \subset V$ such that

$$\{\, (u, g(u)) \in U \times W \,\} = \{\, (u, w) \in U \times W \mid \Phi_i(u, w) = 0\ , i = 1, \cdots, m \,\}.$$

Then define an injective function $G : U \to \mathbb{R}^n$ by

$$G(u) = (u, g(u)).$$

It is easy to check that $G(U) = \mathcal{M}$ locally around $\bar{x}$. □

Proposition 2.7 (Tangents to immersions)**.** Let $U \subset \mathbb{R}^m$ be an open set with a point $\bar{u} \in U$ and $G : U \to \mathbb{R}^n$ be C^k-smooth with $\nabla G(\bar{u})$ full rank. Then there exists an open set $U_0 \subset U$ containing $\bar{u}$ such that $G(U_0)$ is a C^k-manifold around $G(\bar{u})$ and $T_{G(U_0)}(G(u)) = R(\nabla G(u))$ for all $u \in U_0$.

Proof. Since $G : U \to \mathbb{R}^n$ is C^k-smooth with $\nabla G(\bar{u})$ full rank, then G is of constant rank m around $\bar{u}$. According to Theorem 2.5, there exist open sets $U_0 \subset \mathbb{R}^m$ containing $\bar{u}$, $V_0 \subset \mathbb{R}^n$ containing $G(\bar{u})$ and diffeomorphisms $\varphi : U_0 \to \varphi(U_0)$, $\psi : V_0 \to \psi(V_0)$, with $U_0 \subset U$ and $G(U_0) \subset V_0$, such that

$$\psi \circ G \circ \varphi^{-1}(x_1, \cdots, x_m) = (x_1, \cdots, x_m, 0, \cdots, 0).$$

Hence,

$$G(U_0) \cap V_0 = \{x \in V_0 : \psi_i(x) = 0,\ i = m + 1, \cdots, n\}$$

where $\nabla\psi_i(x)$ are linearly independent on V_0. Therefore $G(U_0)$ is a manifold around $G(\bar{u})$. Hence

$$T_{G(U_0)}(G(u)) = \mathrm{Ker}(\nabla\Phi(G(u)))$$

where $\Phi(x) = (\psi_{m+1}, \cdots, \psi_n)$. Since $\Phi \circ G(u) = 0$ for any $u \in U_0$, then by chain rule we get

$$\nabla\Phi(G(u))\nabla G(u) = 0.$$

Therefore $R(\nabla G(u)) \subset \mathrm{Ker}(\nabla\Phi(G(u)))$. Since $\dim(\nabla G(u))$=$\dim(\mathrm{Ker}(\nabla\Phi(G(u))))$=$m$, then

$$T_{G(U_0)}(G(u)) = \mathrm{Ker}(\nabla\Phi(G(u))) = R(\nabla G(u)).$$

□

Theorem 2.8 (Normals to the normal bundle)**.** Suppose a point $\bar{x} \in V \subset \mathbb{R}^n$ where V is an open set and $\Phi_i : V \to \mathbb{R}$ $(i = 1, \cdots, m)$ are C^2-smooth functions with $\nabla\Phi_i(\bar{x})$ linearly independent. Then there exists an open set $V' \subset V$ containing $\bar{x}$ such that

$$\mathcal{M} = \{ x \in V' \mid \Phi_i(x) = 0,\ i = 1, \cdots, m \}$$

is a C^2-smooth manifold around x with

$$T_{\mathcal{M}}(x) = \{\nabla\Phi_i(x)\}^{\perp} \text{ and } N_{\mathcal{M}}(x) = \Big\{ \sum_i \lambda_i \nabla\Phi_i(x) \mid \lambda \in \mathbb{R}^m \Big\} \tag{1}$$

for any $x \in \mathcal{M}$. Furthermore, the *normal bundle* gph $N_{\mathcal{M}}$ is a C^1-smooth manifold around $(x, \sum \lambda_i \nabla\Phi_i(x))$ and

$$N_{\text{gph } N_{\mathcal{M}}}\Big(x, \sum_i \lambda_i \nabla\Phi_i(x)\Big) = \Big\{ (z, w) \mid w \in T_{\mathcal{M}}(x),\ z + \sum_i \lambda_i \nabla^2\Phi_i(x) w \in N_{\mathcal{M}}(x) \Big\}$$

for any $x \in \mathcal{M}$ and $\lambda \in \mathbb{R}^m$.

Proof. Since $\mathcal{M}$ is a C^k-smooth manifold of codimension m, we can choose $G : U \to \mathbb{R}^n$ with $G(\bar{u}) = \bar{x}$ as in Proposition 2.6. According to the proof of Proposition 2.6, it is easy to deduce that $\nabla G(u)$ is full rank for any $u \in U$. Moreover, (1) holds. Define the following C^1-smooth function $F : U \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^n$ by

$$F(u, \lambda) = \Big(G(u), \sum \lambda_i \nabla\Phi_i(G(u))\Big) \text{ where } u \in U,\ \lambda \in \mathbb{R}^m.$$

Let's compute $T_{\text{gph } N_{\mathcal{M}}}(x, \sum_i \lambda_i \nabla\Phi_i(x))$ first. Since

$$\nabla F(u, \lambda) = \begin{pmatrix} \nabla G(u) & 0 & \cdots & 0 \\ \sum \lambda_i \nabla^2\Phi_i(G(u))\nabla G(u) & \nabla\Phi_1(G(u)) & \cdots & \nabla\Phi_m(G(u)) \end{pmatrix}$$

has full rank for any $(u, \lambda) \in U \times \mathbb{R}^m$, then there exists an open set $U_0 \times W_0 \subset U \times \mathbb{R}^m$ such that locally around $(x, \sum_i \lambda_i \nabla\Phi_i(x))$, the set $F(U_0 \times W_0) = \text{gph } N_{\mathcal{M}}$ is a C^1-smooth manifold by Proposition 2.7. Moreover, we have that

$$\begin{aligned} T_{\text{gph } N_{\mathcal{M}}}\Big(x, \sum_i \lambda_i \nabla\Phi_i(x)\Big) &= R(\nabla F(u, \lambda)) \\ &= \Big\{\Big(\nabla G(u) w, \sum \lambda_i \nabla^2\Phi_i(G(u))\nabla G(u) w + \sum_i z_i \nabla\Phi_i(G(u))\Big) \mid w \in \mathbb{R}^{n-m}, z \in \mathbb{R}^m \Big\} \\ &= \Big\{ (z, w) \mid z \in T_{\mathcal{M}}(x), w - \sum_i \lambda_i \nabla^2\Phi_i(x) z \in N_{\mathcal{M}}(x) \Big\}. \end{aligned}$$

Since gph $N_{\mathcal{M}}$ is a C^1-smooth manifold around $(x, \sum_i \lambda_i \nabla\Phi_i(x))$, then $N_{\text{gph } N_{\mathcal{M}}}(x, \sum_i \lambda_i \nabla\Phi_i(x)) = T_{\text{gph } N_{\mathcal{M}}}(x, \sum_i \lambda_i \nabla\Phi_i(x))^{\perp}$. We can calculate this set from the fact that for any linear map A and a linear subspace S

$$\{ x \mid Ax \in S \}^{\perp} = A^* S^{\perp}.$$

In this case,

$$A = \begin{pmatrix} I & 0 \\ -\sum_i \lambda_i \nabla^2\Phi_i(x) & I \end{pmatrix} \text{ and } S = \{ (u, v) \mid u \in T_{\mathcal{M}}(x), v \in N_{\mathcal{M}}(x) \}.$$

Therefore

$$N_{\text{gph } N_{\mathcal{M}}}\Big(x, \sum_i \lambda_i \nabla\Phi_i(x)\Big) = \Big\{ (z, w) \mid w \in T_{\mathcal{M}}(x),\ z + \sum_i \lambda_i \nabla^2\Phi_i(x) w \in N_{\mathcal{M}}(x) \Big\}.$$

□

Note the classic definition of a manifold is via "coordinate charts." Then the manifold $\mathcal{M} \subset \mathbb{R}^n$ defined by Definition 2.3 can be identified as an embedded submanifold of $\mathbb{R}^n$ according to [7, Prop. 8.12]. In this setting, Proposition 2.6 and 2.7 are standard results in smooth manifold theory.

Corollary 2.9 (Generalized Hessians: smooth case)**.** Suppose a point $\bar{x} \in V \subset \mathbb{R}^n$ where V is an open set and $\Phi_i : V \to \mathbb{R}$ $(i = 1, \cdots, m)$ are C^2-smooth with $\nabla\Phi_i(\bar{x})$ linearly independent. Then there exists an open set $V' \subset V$ containing $\bar{x}$ such that

$$\mathcal{M} = \{\ x \in V' \mid \Phi_i(x) = 0,\ i = 1, \cdots, m\ \}$$

is a C^2-smooth manifold around $\bar{x}$ with the following property. Suppose $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a C^2-smooth function around $\bar{x}$ with $0 \in \partial(h + \delta_{\mathcal{M}})(\bar{x})$. Then there exists a unique $\bar{\lambda} \in \mathbb{R}^m$ such that the *Lagrangian* $L = h + \sum_i \bar{\lambda}_i \Phi_i$ satisfies $\nabla L(\bar{x}) = 0$ and

$$\partial^2(h + \delta_{\mathcal{M}})(\bar{x}|0)(w) = \begin{cases} \nabla^2 L(\bar{x})w + N_{\mathcal{M}}(\bar{x}) & w \in T_{\mathcal{M}}(\bar{x}) \\ \emptyset & w \notin T_{\mathcal{M}}(\bar{x}). \end{cases}$$

Proof. Since $0 \in (h + \delta_{\mathcal{M}})(\bar{x})$ and $\nabla\Phi_i(\bar{x})$ are linearly independent, then there exists a unique $\bar{\lambda} \in \mathbb{R}^m$ such that $-\nabla h(\bar{x}) = \sum_i \bar{\lambda}_i \nabla\Phi_i(\bar{x})$. According to [18, Prop. 4.1], we have that for any $\bar{x} \in \mathcal{M}$ and $w \in \mathbb{R}^n$

$$\partial^2(h + \delta_{\mathcal{M}})(\bar{x}|0)(w) = \nabla^2 h(\bar{x})w + \partial^2\delta_{\mathcal{M}}(\bar{x}| - \nabla h(\bar{x}))(w).$$

Since

$$\partial^2\delta_{\mathcal{M}}(\bar{x}| - \nabla h(\bar{x})) : w \mapsto \{\ z \mid (z, -w) \in N_{\text{gph } N_{\mathcal{M}}}(\bar{x}, -\nabla h(\bar{x}))\ \},$$

then this problem boils down to computing the normal cone of gph $N_{\mathcal{M}}$ at $(\bar{x}, -\nabla h(\bar{x}))$. According to Proposition 2.8, we have that for any $w \in T_{\mathcal{M}}(\bar{x})$

$$\partial^2\delta_M(\bar{x}| - \nabla h(\bar{x}))(w) = \partial^2\delta_M\Big(\bar{x}| \sum_i \bar{\lambda}_i \nabla\Phi_i(\bar{x})\Big)(w) = \sum_i \bar{\lambda}_i \nabla^2\Phi_i(\bar{x})w + N_{\mathcal{M}}(\bar{x}).$$

Hence

$$\partial^2(h + \delta_{\mathcal{M}})(\bar{x}|0)(w) = \nabla^2 h(\bar{x})w + \partial^2\delta_{\mathcal{M}}(\bar{x}| - \nabla h(\bar{x}))(w) = \begin{cases} \nabla^2 L(\bar{x})w + N_{\mathcal{M}}(\bar{x}) & w \in T_{\mathcal{M}}(\bar{x}) \\ \emptyset & w \notin T_{\mathcal{M}}(\bar{x}). \end{cases}$$

□

Since $N_{\text{gph } N_{\mathcal{M}}}$ is only determined by the geometry of $\mathcal{M}$, we can use intrinsic geometric objects to formulate it. Next, we will introduce the concept of covariant derivative and Hessian.

Definition 2.10. Let a C^2-smooth manifold $\mathcal{M} \subset \mathbb{R}^n$ contain a point $\bar{x}$. We say a function $f : \mathcal{M} \to \mathbb{R}$ is C^2-*smooth around* $\bar{x}$ if there exists a *representative function* $h : \mathbb{R}^n \to \mathbb{R}$ which is C^2-smooth around $\bar{x}$ with $h|_{\mathcal{M}} = f|_{\mathcal{M}}$ locally around $\bar{x}$.

Let $\mathcal{M}$ be a C^2-smooth manifold around $\bar{x}$. Then the projection mapping $u \mapsto P_{\mathcal{M}}(\bar{x} + u)$ is well-defined and C^2-smooth around 0 on $T_{\mathcal{M}}(\bar{x})$ as proved in [10].

Definition 2.11. Suppose $\mathcal{M} \subset \mathbb{R}^n$ is a C^2-smooth manifold around a point $\bar{x}$ and a function $f : \mathcal{M} \to \mathbb{R}$ is C^2-smooth around $\bar{x}$. Then the *covariant derivative* $\nabla f_{\mathcal{M}}(\bar{x}) \in T_{\mathcal{M}}(\bar{x})$ is defined by

$$\langle \nabla_{\mathcal{M}} f(\bar{x}), u \rangle = \frac{d}{dt} f(P_{\mathcal{M}}(\bar{x} + tu))|_{t=0} \text{ for all } u \in T_{\mathcal{M}}(\bar{x}),$$

and the *covariant Hessian* $\nabla^2_{\mathcal{M}} f(\bar{x}) : T_{\mathcal{M}}(\bar{x}) \times T_{\mathcal{M}}(\bar{x}) \to \mathbb{R}$ is the unique self-adjoint and bilinear map satisfying

$$\langle \nabla^2_{\mathcal{M}} f(\bar{x})u, u \rangle = \frac{d^2}{dt^2} f(P_{\mathcal{M}}(\bar{x} + tu))|_{t=0} \text{ for all } u \in T_{\mathcal{M}}(\bar{x}).$$

This definition agrees with the classic definition of covariant derivative and Hessian using geodesics as proved in [10]. Suppose the function $f : \mathcal{M} \to \mathbb{R}$ is C^2-smooth around $\bar{x}$. Let h be any C^2-smooth representative of f around $\bar{x}$ and C^2-smooth functions Φ_i define local equations for $\mathcal{M}$. If $\nabla h(\bar{x}) \in N_{\mathcal{M}}(\bar{x})$, then using the notation of Corollary 2.9, there exists a unique $\bar{\lambda}$ such that $\nabla h(\bar{x}) + \sum \bar{\lambda}_i \nabla\Phi_i(\bar{x}) = 0$. Furthermore, the following results have been showed in [10]:

$$\nabla_{\mathcal{M}} f(\bar{x}) = P_{T_{\mathcal{M}}(\bar{x})} \nabla h(\bar{x}) \text{ and } \nabla^2_{\mathcal{M}} f(\bar{x}) = P_{T_{\mathcal{M}}(\bar{x})} \nabla^2 L(\bar{x}) P_{T_{\mathcal{M}}(\bar{x})}.$$

Theorem 2.12 (Generalized and covariant Hessians)**.** Suppose $\mathcal{M} \subset \mathbb{R}^n$ is a C^2-smooth manifold around a point $\bar{x}$ and the function $f : \mathcal{M} \to \mathbb{R}$ is C^2-smooth around $\bar{x}$. Define the function $\tilde{f} : \mathbb{R}^n \to \overline{\mathbb{R}}$ by

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \mathcal{M} \\ +\infty & x \notin \mathcal{M}. \end{cases}$$

Then

$$0 \in \partial \tilde{f}(\bar{x}) \Leftrightarrow \nabla f_{\mathcal{M}}(\bar{x}) = 0$$

and in that case

$$\partial^2 \tilde{f}(\bar{x}|0)(w) = \begin{cases} \nabla^2_{\mathcal{M}} f(\bar{x})w + N_{\mathcal{M}}(\bar{x}) & w \in T_{\mathcal{M}}(\bar{x}) \\ \emptyset & w \notin T_{\mathcal{M}}(\bar{x}). \end{cases}$$

Proof. Let h be a C^2-smooth representative of f around $\bar{x}$. Then we have

$$\nabla_{\mathcal{M}} f(\bar{x}) = 0 \Leftrightarrow \nabla h(\bar{x}) \in N_{\mathcal{M}}(\bar{x}) \Leftrightarrow 0 \in \partial \tilde{f}(\bar{x}).$$

Let $\bar{\lambda}$ be the unique multiplier satisfying $\nabla h(\bar{x}) + \sum_i \bar{\lambda}_i \nabla \Phi_i(\bar{x}) = 0$. Since $\tilde{f}(x) = h(x) + \delta_{\mathcal{M}}(x)$, then for any $w \in T_{\mathcal{M}}(\bar{x})$ we have, by Corollary 2.9

$$\begin{aligned} \partial^2 \tilde{f}(\bar{x}|0)(w) &= \partial^2 (h + \delta_{\mathcal{M}})(\bar{x}|0)(w) \\ &= \nabla^2 h(\bar{x})w + \sum_i \bar{\lambda}_i \nabla^2 \Phi_i(\bar{x})w + N_{\mathcal{M}}(\bar{x}) \\ &= P_{T_{\mathcal{M}}(\bar{x})}\Big(\nabla^2 h(\bar{x}) + \sum_i \bar{\lambda}_i \nabla^2 \Phi_i(\bar{x})\Big) P_{T_{\mathcal{M}}(\bar{x})} w + N_{\mathcal{M}}(\bar{x}) \\ &= \nabla^2_{\mathcal{M}} f(\bar{x})w + N_{\mathcal{M}}(\bar{x}). \end{aligned}$$

The result follows. □

We refer to functions of the form $\tilde{f}$ as *extended-C^2-smooth* at $\bar{x}$. The above theorem gives us some indication of how to calculate a generalized Hessian mapping. The smooth manifold $\mathcal{M}$ simplifies the calculation. "Partial smoothness", which was introduced in [8], gives some underlying smooth structure for a non-smooth function. In this paper, we are going to show that for a partly smooth function relative to manifold $\mathcal{M}$, the local geometry of gph $\partial f(x)$ is determined by the restriction of f to $\mathcal{M}$, under certain assumptions. In this way, we can extend Theorem 2.9 to partly smooth functions.

3 Definitions and results

Definition 3.1. Suppose $C \subset \mathbb{R}^n$ is a nonempty convex set. The subspace parallel to the set C, denoted by par C, is defined by

$$\text{par}\, C = \text{aff}\, C - x \quad \text{for any } x \in C,$$

where aff C is the affine span of C.

Definition 3.2. Suppose that the set $\mathcal{M} \subset \mathbb{R}^n$ contains the point $\bar{x}$. The function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is C^2*-partly smooth* at $\bar{x}$ relative to $\mathcal{M}$ if $\mathcal{M}$ is a C^2-smooth manifold around $\bar{x}$ and the following four properties hold:

1. **(restricted smoothness)** $f|_{\mathcal{M}}$ is C^2-smooth around $\bar{x}$;
2. **(regularity)** at every point close to $\bar{x}$ in $\mathcal{M}$, the function f is subdifferentially regular and has a subgradient;
3. **(normal sharpness)** $N_{\mathcal{M}}(\bar{x}) = \text{par}\, \partial f(\bar{x})$;
4. **(subgradient continuity)** the subdifferential map ∂f is continuous at $\bar{x}$ relative to $\mathcal{M}$.

Definition 3.3. Let f be a C^2-partly smooth function at a point $\bar{x}$ relative to a C^2-smooth manifold $\mathcal{M}$. Then we call $\bar{x}$ is a *strong critical point* of f relative to $\mathcal{M}$ if

$$0 \in \text{ri}\, \partial f(\bar{x})$$

and there exists $\epsilon > 0$ such that

$$f(x) \geq f(\bar{x}) + \epsilon |x - \bar{x}|^2$$

for all points $x \in \mathcal{M}$ near $\bar{x}$.

Given certain assumptions, critical points of parametric partly smooth functions are stable.

Theorem 3.4 (Strong critical points with parameters)**.** Suppose the set $Q \subset \mathbb{R}^m \times \mathbb{R}^n$ is a C^2-smooth manifold containing the point $(\bar{y}, \bar{x})$ and satisfies the condition

$$(w, 0) \in N_Q(\bar{y}, \bar{x}) \Rightarrow w = 0.$$

For each $y \in \mathbb{R}^m$ we define the set

$$Q_y = \{x \in \mathbb{R}^n : (y, x) \in Q\}.$$

Given any function $p : \mathbb{R}^m \times \mathbb{R}^n \to \overline{\mathbb{R}}$, define a function $p_y : \mathbb{R}^n \to \overline{\mathbb{R}}$ by

$$p_y(x) = p(y, x) \text{ for } y \in \mathbb{R}^m \text{ and } x \in \mathbb{R}^n.$$

Suppose the function p is C^2-partly smooth relative to Q. If $\bar{x}$ is a strong critical point of the function $p_{\bar{y}}$ relative to the set $Q_{\bar{y}}$, then there are open neighborhoods $U \subset \mathbb{R}^n$ of $\bar{x}$ and $V \subset \mathbb{R}^m$ of $\bar{y}$ and a C^1-smooth function $\Psi : V \to U$ satisfying $\Psi(\bar{y}) = \bar{x}$, and with the following properties, for all vectors $y \in V$:

1. for all vectors $y \in V$ the set $Q_y \cap U$ is a C^2-smooth manifold;
2. for all vectors $y \in V$ the function p_y is C^2-partly smooth relative to $Q_y \cap U$;
3. the function $p_y|_{Q_y \cap U}$ has a unique critical point $\Psi(y)$;
4. $\Psi(y)$ is a strong critical point of the function p_y relative to $Q_y \cap U$.

Proof. See [8, Thm. 5.2, 5.3, 5.7] . □

The concept of prox-regularity extends properties of convexity to a broader class of functions. It is essential for partly smooth functions to locally identify their manifolds uniquely.

Definition 3.5. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is *prox-regular* at a point $\bar{x}$ for a subgradient $\bar{v} \in \partial f(\bar{x})$ if f is finite at $\bar{x}$, locally lower semi-continuous at $\bar{x}$, and there exist $r > 0$ and $\epsilon > 0$ such that

$$f(x') > f(x) + \langle v, x' - x \rangle - \frac{r}{2}|x' - x|^2 \text{ for } x' \neq x \text{ when}$$

$$|x' - \bar{x}| < \epsilon,\ |x - \bar{x}| < \epsilon,\ |f(x) - f(\bar{x})| < \epsilon,\ |v - \bar{v}| < \epsilon,\ v \in \partial f(x).$$

More precisely, we say f is *prox-regular* at $\bar{x}$ for $\bar{v}$ with respect to ϵ and r. Further, f is *prox-regular* at $\bar{x}$ if it is prox-regular at $\bar{x}$ for every $\bar{v} \in \partial f(\bar{x})$. A set S is *prox-regular* at $\bar{x}$ for $\bar{v} \in N_S(\bar{x})$ if its indicator function δ_S is prox-regular at $\bar{x}$ for $\bar{v} \in \partial \delta_S(\bar{x})$.

Proposition 3.6. Suppose the set $S \subset \mathbb{R}^n$ is closed. Then S is prox-regular at the point $\bar{x} \in S$ if and only if the projection mapping P_S is single-valued near $\bar{x}$.

Proof. See [19, Thm. 1.3]. □

Definition 3.7. For a proper, lower semi-continuous function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and parameter value $\lambda > 0$, the *proximal mapping* $P_\lambda f$ is defined by

$$P_\lambda f(x) := \text{argmin}_w \Big\{ f(w) + \frac{1}{2\lambda}|w - x|^2 \Big\}.$$

Definition 3.8. For $\epsilon > 0$, the *f-attentive ϵ-localization* of ∂f around $(\bar{x}, \bar{v})$ is a (generally set-valued) mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ defined by

$$T(x) = \begin{cases} \{ v \in \partial f(x) \mid |v - \bar{v}| < \epsilon \} & \text{if } |x - \bar{x}| < \epsilon \text{ and } |f(x) - f(\bar{x})| < \epsilon, \\ \emptyset & \text{otherwise.} \end{cases}$$

Definition 3.9. For a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, a set $V \subset \mathbb{R}^n$ is called an *f-attentive neighborhood* of $\bar{x}$ if there exists a $\delta > 0$ such that

$$\{ x \in \mathbb{R}^n \mid |x - \bar{x}| < \delta, \ |f(x) - f(\bar{x})| < \delta \} \subset V.$$

Definition 3.10. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is *subdifferentially continuous* at a point $\bar{x}$ for $\bar{v}$, where $\bar{v} \in \partial f(\bar{x})$, if for every $\delta > 0$ there exists $\epsilon > 0$ such that $|f(x) - f(\bar{x})| < \delta$ whenever $|x - \bar{x}| < \epsilon$ and $|v - \bar{v}| < \epsilon$ with $v \in \partial f(x)$.

Proposition 3.11 (Prox-regularity and proximal mapping)**.** Suppose that the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is prox-regular at $\bar{x} = 0$ for $\bar{v} = 0$ with respect to ϵ and r. In particular suppose f satisfies the following assumption:

$$\begin{cases} f \text{ is locally lower-semicontinuous at } 0 \text{ with } f(0) = 0, \text{ and} \\ r > 0 \text{ is such that } f(x) > -\frac{r}{2}|x|^2 \text{ for all } x \neq 0. \end{cases}$$

(which implies that $P_\lambda f(0) = \{0\}$ when $\lambda \in (0, \frac{1}{r})$). Let T be the f-attentive ϵ-localization T of ∂f around $(0, 0)$. Then for each $\lambda \in (0, \frac{1}{r})$ there is a neighborhood X of $\bar{x} = 0$ such that, on X, the mapping $P_\lambda f$ is single-valued and continuous and

$$P_\lambda f(x) = (I + \lambda T)^{-1}(x).$$

Proof. See [17, Thm. 4.4]. □

Lemma 3.12. Suppose the function f is extended-C^2-smooth at $\bar{x}$. Then f is subdifferentially regular, prox-regular and subdifferentially continuous at $\bar{x}$.

Proof. See [17, Ex. 2.8]. □

Note that function f being prox-regular at $\bar{x}$ doesn't imply that f is subdifferentially regular at $\bar{x}$. Here is an example. Let $f(x, y) = (x - |y|)^{\frac{1}{3}}$. Since there is no subgradient at $(0, 0)$, then f is prox-regular there. However, epi f is not Clarke regular at $(0, 0, 0)$ which implies f is not subdifferentially regular at $(0, 0)$.

4 Identification for functions

A partly smooth function has a smooth structure on its corresponding manifold. [5, Thm. 5.3] gives a nice "identification" property for partly smooth, prox-regular functions. Though this theorem is true, its proof is flawed because it depends on the assumption that the prox-regularity of a function implies the prox-regularity of its epigraph. We will prove this theorem by using proximal mappings in this section. First, let's see an example which shows that the prox-regularity of a function isn't equivalent to the prox-regularity of its epigraph.

Example 4.1 (Prox-regularity of functions versus epigraphs)**.** Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(2^n) = \sqrt{2^n}$ for any $n \in \mathbb{Z}$, f affine on $[2^n, 2^{n+1}]$, $f(0) = 0$ and $f(x) = f(-x)$ for any x. First note that

$$\begin{aligned} \partial f(\pm 2^n) &= \Big\{ \pm \frac{1}{\sqrt{2^{n-1}} + \sqrt{2^n}}, \pm \frac{1}{\sqrt{2^n} + \sqrt{2^{n+1}}} \Big\}, \\ \partial f(\pm x) &= \pm \frac{1}{\sqrt{2^n} + \sqrt{2^{n+1}}}, \ x \in (2^n, 2^{n+1}), \\ \partial f(0) &= (-\infty, +\infty). \end{aligned}$$

Next, we are going to prove that f is prox-regular at 0 for any $v \in \partial f(0)$. It is equivalent to show that there exist $\epsilon > 0$ and $r > 0$ such that

$$f(x') > f(x) + \langle u, x' - x \rangle - \frac{r}{2}|x' - x|^2 \text{ for } x \neq x' \text{ when}$$
$$|x'| < \epsilon, \ |x| < \epsilon, \ |f(x)| < \epsilon, \ |v - u| < \epsilon, \ u \in \partial f(x).$$

For any $x \to 0$ and $u \in \partial f(x)$ we have that $|u| \to +\infty$. Since $|v - u| < \epsilon$ and $|x| < \epsilon$, then x has to be 0 when ϵ is small. Hence we just have to prove

$$f(x') > \langle u, x' \rangle - \frac{r}{2}|x'|^2.$$

By the definition of f, we know that $f(x') > |\langle \frac{1}{\sqrt{2^{n-1}}+\sqrt{2^n}}, x' \rangle|$. Thus f is prox-regular at 0. However, epi f is not prox-regular at $(0, 0)$. If so, there should be a neighborhood V of $(0, 0)$ such that the projection mapping $P_{\text{epi } f}$ is single-valued on V by Proposition 3.6. However, $P_{\text{epi } f}$ is not single-valued around $(\pm 2^n, \sqrt{2^n})$ for any $n \in \mathbb{Z}$. Thus epi f is not prox-regular at $(0, 0)$.

Lemma 4.2. Suppose the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is C^2-partly smooth at a point $\bar{x}$ relative to a C^2-smooth manifold $\mathcal{M}$. Then $f + \delta_{\mathcal{M}}$ is prox-regular at $\bar{x}$ and $\partial f(\bar{x}) \subset \partial(f + \delta_{\mathcal{M}})(\bar{x})$.

Proof. Let h be any C^2-smooth representative of f around $\bar{x}$. Since $f \le f + \delta_{\mathcal{M}} = h + \delta_{\mathcal{M}}$ and $f + \delta_{\mathcal{M}}$ is extended-C^2-smooth at $\bar{x}$, so $f + \delta_{\mathcal{M}}$ is prox-regular at $\bar{x}$, and $\hat{\partial} f(\bar{x}) \subset \hat{\partial}(f + \delta_{\mathcal{M}})(\bar{x})$. The result follows since f and $h + \delta_{\mathcal{M}}$ are both regular at $\bar{x}$. □

Proposition 4.3 (Subdifferential smoothness)**.** Suppose that the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is C^2-partly smooth at a point $\bar{x}$ relative to a C^2-smooth manifold $\mathcal{M}$ with $\bar{y} \in \text{ri } \partial f(\bar{x})$. Let h be any C^2-smooth representative of f around $\bar{x}$. Then

$$\text{gph } \partial f \cap (\mathcal{M} \times \mathbb{R}^n) = \text{gph } (\nabla h + N_{\mathcal{M}}) \cap (\mathcal{M} \times \mathbb{R}^n) \text{ locally around } (\bar{x}, \bar{y}).$$

Proof. According to [8, Prop. 2.4], we know

$$\partial f(x) \subset \text{aff } \partial f(x) = \nabla h(x) + N_{\mathcal{M}}(x)$$

for any x close to $\bar{x}$ in $\mathcal{M}$. Thus

$$\text{gph } \partial f \cap (\mathcal{M} \times \mathbb{R}^n) \subset \text{gph } (\nabla h + N_{\mathcal{M}}) \cap (\mathcal{M} \times \mathbb{R}^n) \text{ locally around } (\bar{x}, \bar{y}).$$

Next, we claim the reverse inclusion: given (x, y) is close to $(\bar{x}, \bar{y})$, then $y \in \nabla h(x) + N_{\mathcal{M}}(x)$ implies $y \in \partial f(x)$. If this is not true, then there exist sequences $x_n \to \bar{x}$ in $\mathcal{M}$ and $y_n \in \text{aff } \partial f(x_n) \to \bar{y}$ with $y_n \notin \partial f(x_n)$. Since f is regular at x when x is close to $\bar{x}$ in $\mathcal{M}$, then $\partial f(x)$ is closed and convex. According to the Separation Theorem, for all large n there exists a unit vector $z_n \in \text{par } \partial f(x_n) = N_{\mathcal{M}}(x_n)$ such that

$$\langle z_n, u \rangle \ge \langle z_n, y_n \rangle$$

for all $u \in \partial f(x_n)$. Passing to a subsequence if necessary, we can assume z_n approaches a unit vector z. Since ∂f is continuous at $\bar{x}$ relative to $\mathcal{M}$, then $\partial f(x_n)$ converges to $\partial f(\bar{x})$. Also, $N_{\mathcal{M}}(x_n)$ converges to $N_{\mathcal{M}}(\bar{x})$. As a result, we have

$$z \in N_{\mathcal{M}}(\bar{x}) \text{ and } \langle z, u \rangle \ge \langle z, \bar{y} \rangle$$

for any $u \in \partial f(\bar{x})$. To see this, choose $u_n \to u$ satisfying $u_n \in \partial f(x_n)$, note $\langle z_n, u_n \rangle \ge \langle z_n, y_n \rangle$, and take limits. This shows that $\bar{y}$ is separated from the convex set $\partial f(\bar{x})$ in its affine span. But this contradicts the fact that $\bar{y} \in \text{ri } \partial f(\bar{x})$. The result follows. □

Corollary 4.4 (Set version of subdifferential smoothness)**.** Suppose a set $S \subset \mathbb{R}^n$ is partly smooth at a point $\bar{x}$ relative to a C^2-smooth manifold $\mathcal{M}$ with $\bar{y} \in \text{ri } N_S(\bar{x})$. Then

$$\text{gph } N_S \cap (\mathcal{M} \times \mathbb{R}^n) = \text{gph } N_{\mathcal{M}} \cap (\mathcal{M} \times \mathbb{R}^n) \text{ locally around } (\bar{x}, \bar{y}).$$

Proposition 4.5 (Extended-smooth reduction)**.** Suppose that the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is C^2-partly smooth at a point $\bar{x}$ relative to a C^2-smooth manifold $\mathcal{M}$ with $0 \in \text{ri } \partial f(\bar{x})$, and f is prox-regular at $\bar{x}$ for 0. Then if $\lambda > 0$ is sufficiently small, there exists a neighborhood V of $\bar{x}$ on which the proximal mappings $P_\lambda f$ and $P_\lambda(f + \delta_{\mathcal{M}})$ agree.

Proof. Without loss of generality, let $\bar{x} = 0$. According to Lemma 4.2, we know $f + \delta_M$ is also prox-regular at 0. We can choose ϵ and r such that f and $f + \delta_{\mathcal{M}}$ are both prox-regular at 0 for 0 with respect to r and ϵ, in particular with the assumption in Proposition 3.11 holding. Let T be the f-attentive ϵ-localization of ∂f around $(\bar{x}, 0)$. For any $\lambda \in (0, 1/r)$ there exists a neighborhood X of $\bar{x} = 0$ such that both $P_\lambda f$ and $P_\lambda(f + \delta_{\mathcal{M}})$ are single-valued and continuous, by Proposition 3.11. In order to prove this proposition, it is sufficient to prove for any $x_n \to \bar{x}$ we have $P_\lambda f(x_n) = P_\lambda(f + \delta_{\mathcal{M}})(x_n)$ for large n. Let h be any C^2-smooth representative of f on $\mathcal{M}$ and define

$$w_n = P_\lambda(f + \delta_{\mathcal{M}})(x_n) = \operatorname{argmin}_x \Big\{ h(x) + \delta_{\mathcal{M}}(x) + \frac{1}{2\lambda}|x - x_n|^2 \Big\} \in \mathcal{M}.$$

Since the assumption in Proposition 3.11 holds for $f + \delta_{\mathcal{M}}$, we have $P_\lambda(f + \delta_{\mathcal{M}})(\bar{x}) = \bar{x}$. Moreover, the continuity of $P_\lambda(f + \delta_{\mathcal{M}})$ implies $w_n \to \bar{x}$. Consequently, $x_n - w_n \to 0$. Since w_n minimizes $h(x) + \delta_{\mathcal{M}}(x) + \frac{1}{2\lambda}|x - x_n|^2$, then

$$0 \in \partial(h(w_n) + \delta_{\mathcal{M}}(w_n) + \frac{1}{2\lambda}|w_n - x_n|^2) = \nabla h(w_n) + N_{\mathcal{M}}(w_n) + \frac{1}{\lambda}(w_n - x_n)$$

or equivalently

$$\frac{1}{\lambda}(x_n - w_n) \in \nabla h(w_n) + N_{\mathcal{M}}(w_n).$$

Since $0 \in \operatorname{ri} \partial f(\bar{x})$ and $\frac{1}{\lambda}(x_n - w_n) \to 0$, then by Proposition 4.3 we know

$$\frac{1}{\lambda}(x_n - w_n) \in \partial f(w_n) \text{ for large } n,$$

which also implies

$$\frac{1}{\lambda}(x_n - w_n) \in T(w_n) \text{ for large } n,$$

since $w_n \to \bar{x}$ in $\mathcal{M}$, so $f(w_n) \to f(\bar{x})$. Thus

$$x_n \in (I + \lambda T)(w_n) \text{ for all large } n,$$

from which we get

$$w_n \in (I + \lambda T)^{-1}(x_n) = P_\lambda f(x_n) \text{ for all large } n$$

by Proposition 3.11. Hence $P_\lambda f(x_n) = P_\lambda(f + \delta_{\mathcal{M}})(x_n)$ for all large n. □

If $0 \in \operatorname{ri} \partial f(\bar{x})$ doesn't hold, the above result can fail. Here is an example.

Example 4.6. Define the function f as follows:

$$f(x) = \begin{cases} +\infty & x \in (-\infty, 0) \\ 0 & x \in [0, \infty). \end{cases}$$

It is easy to see that f is prox-regular at x for all $x \in [0, \infty)$, and partly smooth at 0 relative to $\mathcal{M} = \{0\}$. Since $\partial f(0) = (-\infty, 0]$, then 0 doesn't lie in the interior of $\partial f(0)$. For any small $\lambda > 0$,

$$P_\lambda f(x) = \operatorname{argmin}_w \{ f(w) + \frac{1}{2\lambda}|x - w|^2 \} = x \text{ for alll } x > 0.$$

Corollary 4.7 (Set version of extended-smooth reduction)**.** Let $\mathcal{M}$ be a C^2-smooth manifold around a point $\bar{x}$. Suppose a set S is partly smooth at $\bar{x}$ relative to $\mathcal{M}$, and S is prox-regular at $\bar{x}$ for $\bar{v} \in \operatorname{ri} N_S(\bar{x})$. Suppose $\lambda > 0$ is sufficiently small. Then for x sufficiently close to $\bar{x}$, the projections $P_S(x + \lambda\bar{v})$ lies in $\mathcal{M}$.

Proof. Applying Proposition 4.5 to $f = \delta_S - \langle \bar{v}, \cdot \rangle$. □

Corollary 4.8 (Active manifold as proximal range)**.** Under the same assumption as Proposition 4.5, the set $P_\lambda f(V)$ is a neighborhood of $\bar{x}$ in $\mathcal{M}$ for any sufficiently small neighborhood V of $\bar{x}$.

Proof. By Proposition 4.5, it is sufficient to prove that for any $x_n \to \bar{x}$ in $\mathcal{M}$, there exists $w_n \to \bar{x}$ with $P_\lambda f(w_n) = x_n$ for large n. Since f is partly smooth at $\bar{x}$ relative to $\mathcal{M}$, then there exists $y_n \in \partial f(x_n) \to 0$. For large n, we have $y_n \in T(x_n)$. So $x_n + \lambda y_n \in (I + \lambda T)(x_n)$, which implies $x_n = P_\lambda f(x_n + \lambda y_n)$. Let $w_n = x_n + \lambda y_n$. The result follows. □

Corollary 4.9 (Set version of active manifold as proximal range)**.** Under the same assumption as Corollary 4.7, for any sufficiently small neighborhood V of $\bar{x}$, the projection $P_S(V + \lambda\bar{v})$ is a neighborhood of $\bar{x}$ in $\mathcal{M}$.

Proof. Apply Corollary 4.8. □

Theorem 4.10 (Identification)**.** Let the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be C^2-partly smooth at a point $\bar{x}$ relative to a C^2-smooth manifold $\mathcal{M}$ and prox-regular at $\bar{x}$ for $\bar{y} \in \text{ri } \partial f(\bar{x})$. Suppose $x_k \to \bar{x}$ and $f(x_k) \to f(\bar{x})$. Then

$$x_k \in \mathcal{M} \text{ for all large } k$$

if and only if

$$\text{dist}(\bar{y}, \partial f(x_k)) \to 0.$$

Proof. By subtracting an affine function from f, we can assume $\bar{x} = 0$, $\bar{y} = 0$ and $f(\bar{x}) = 0$ without loss of generality. Since f is prox-regular at 0 for 0, then there exist $\epsilon > 0$ and $r > 0$ such that

$$f(x') > f(x) + \langle v, x' - x\rangle - \frac{r}{2}|x - x'|^2 \text{ for } x' \neq x$$

whenever $|x| < \epsilon, |f(x)| < \epsilon, |x'| < \epsilon, |v| < \epsilon$ and $v \in \partial f(x)$. Letting $x = 0, v = 0$, we have

$$f(x') > -\frac{r}{2}|x'|^2$$

for any $|x'| < \epsilon$ and $x' \neq x$. Since we are only interested in the local geometry of epi f around $(0, 0)$, then we can add to f the indicator of a compact neighborhood $\overline{B}_{\frac{\epsilon}{2}}(0)$, which is a closed ball centered at 0 with radius $\frac{\epsilon}{2}$, to make the assumption in Proposition 3.11 hold for f: if this proposition is true for $f + \delta_{\overline{B}_{\frac{\epsilon}{2}}(0)}$, it is also true for f. To sum up, we can assume f satisfies the assumption in Proposition 3.11 without loss of generality. Fix λ such that Proposition 4.5 holds for f. Let T be the f-attentive ϵ-localization of ∂f. If $\text{dist}(0, \partial f(x_k)) \to 0$, then there exists a sequence of $y_k \to 0$ with $y_k \in \partial f(x_k)$. Then we have

$$\frac{1}{\lambda}((x_k + \lambda y_k) - x_k) \in \partial f(x_k),$$

which implies

$$\frac{1}{\lambda}((x_k + \lambda y_k) - x_k) \in T(x_k) \text{ for large } k.$$

Thus

$$x_k + \lambda y_k \in (I + \lambda T)(x_k) \text{ for large } k.$$

Consequently

$$x_k = (I + \lambda T)^{-1}(x_k + \lambda y_k) = P_\lambda f(x_k + \lambda y_k) \in \mathcal{M} \text{ for large } k$$

by Proposition 4.5. Thus the result follows since the converse is immediate by partial smoothness. □

Corollary 4.11 (Identification for sets)**.** Let the set S be C^2-partly smooth at the point $\bar{x}$ relative to the C^2-smooth manifold $\mathcal{M}$ and prox-regular there for $\bar{n} \in \text{ri } N_S(\bar{x})$. If the sequence $\{x_k\} \in S$ satisfies $x_k \to \bar{x}$, then

$$\text{dist}(\bar{n}, N_S(x_k)) \to 0 \Leftrightarrow x_k \in \mathcal{M} \text{ for all large } k.$$

Proof. The result follows by applying Theorem 4.10 to the indicator function δ_S. □

Corollary 4.12 (Uniqueness of active manifold)**.** Consider a set S that is prox-regular at a point $\bar{x}$ for $\bar{n} \in \text{ri } N_S(\bar{x})$ and C^2-partly smooth there relative to each of the two C^2-smooth manifolds $\mathcal{M}_1$ and $\mathcal{M}_2$. Then near $\bar{x}$ we have $\mathcal{M}_1 \equiv \mathcal{M}_2$.

Proof. If this is not true, then there exists a sequence of points x_k converging to $\bar{x}$ such that $x_k \in \mathcal{M}_1 \setminus \mathcal{M}_2$. Since S is partly smooth relative to $\mathcal{M}_1$, then the normal cone $N_S(x_k) \to N_S(\bar{x})$. Hence $\text{dist}(\bar{n}, N_S(x_k)) \to 0$. Applying Corollary 4.11 to δ_S with $\mathcal{M} \equiv \mathcal{M}_2$ implies $x_k \in \mathcal{M}_2$ for all large k, which is contradictory to $x_k \notin \mathcal{M}_2$. Thus the result follows. □

The definition of strong critical points demands quadratic growth along the manifold. Under the assumption of prox-regularity, strong critical points of such functions are actually locally quadratic minimizers. [5, Thm. 6.2] gives a proof, requiring such functions to be prox-regular at the local minimizer. In this paper, we use another approach to prove this with a more natural, slightly weaker assumption, only requiring such functions to be prox-regular at the minimizer for the subgradient 0.

Proposition 4.13 (Sufficient optimality conditions)**.** Suppose the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is C^2-partly smooth at the point $\bar{x}$ relative to the C^2-smooth manifold $\mathcal{M}$ and prox-regular there for $0 \in \operatorname{ri} \partial f(\bar{x})$.

1. $\bar{x}$ is a strict local minimizer of the restricted function $f|_{\mathcal{M}} \Leftrightarrow \bar{x}$ is in fact an unconstrained strict local minimizer of f.
2. $\bar{x}$ is a strong critical point of f relative to $\mathcal{M} \Leftrightarrow f$ grows at least quadratically near $\bar{x}$.

Proof. One direction of both cases is obvious. Let's prove the other direction. First we are going to prove that $\bar{x}$ being a strict local minimizer of the restricted function $f|_{\mathcal{M}}$ is equivalent to $\bar{x}$ being an unconstrained strict local minimizer of f. Without loss of generality, let $\bar{x} = 0$, $f(\bar{x}) = 0$ and f satisfy the assumption in Proposition 3.11. We are going to prove this proposition by contradiction. Suppose there exists a sequence $x_k \notin \mathcal{M} \to \bar{x}$ with

$$f(x_k) \leq f(\bar{x}) \text{ for all } k.$$

For large k, we know that x_k lies in the f-attentive neighborhood of $\bar{x}$ in Proposition 4.5. Hence $x_k \neq y_k = P_\lambda f(x_k) \in \mathcal{M}$ and $y_k \to P_\lambda f(\bar{x}) = \bar{x}$. Then we have

$$\begin{aligned} f(\bar{x}) &\geq f(x_k) \\ &\geq \min_w \Big\{ f(w) + \frac{1}{2\lambda}|x_k - w|^2 \Big\} \\ &= f(y_k) + \frac{1}{2\lambda}|y_k - x_k|^2 \\ &> f(\bar{x}) + \frac{1}{2\lambda}|y_k - x_k|^2. \end{aligned}$$

Consequently, we get a contradiction

$$0 > \frac{1}{2\lambda}|y_k - x_k|^2.$$

Next we are going to prove case (2). Since f grows quadratically at $\bar{x}$ relative to $\mathcal{M}$, then there exists a $\delta > 0$ such that $f(x) > \delta|x - \bar{x}|^2$ around $\bar{x}$ relative to $\mathcal{M}$. Define h by $h(x) = f(x) - \delta|x - \bar{x}|^2$. Since $\delta|x - \bar{x}|^2$ is C^2-smooth, then h is also prox-regular at $\bar{x}$ for $0 \in \operatorname{ri} \partial h(\bar{x})$ and partly smooth at $\bar{x}$ relative to $\mathcal{M}$. Moreover, we know that $h(x) > h(\bar{x})$ locally around $\bar{x}$ restricted to $\mathcal{M}$. According to case (1), we know that $h(x) > h(\bar{x})$ locally around $\bar{x}$. Then the second case follows. □

5 Calculation of generalized Hessian mappings

In general it may be hard to compute the generalized Hessian mapping. Our goal is to analyze the generalized Hessian mapping in the easier special case of partly smooth and prox-regular functions. Given these assumptions plus subdiffential continuity property, Theorem 4.10 guarantees that the local geometry of gph ∂f is determined by $f|_{\mathcal{M}}$. This smooth structure simplifies the computation of the generalized Hessian mapping and also gives a geometry explanation of the second condition in Theorem 6.1.

Proposition 5.1 (Subdifferential localization and active manifolds)**.** Suppose the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is C^2-partly smooth at the point $\bar{x}$ relative to the C^2-smooth manifold $\mathcal{M}$, and both prox-regular and subdifferentially continuous at $\bar{x}$ for $\bar{y} \in \operatorname{ri} \partial f(\bar{x})$. Then

$$\operatorname{gph} \partial f \subset \mathcal{M} \times \mathbb{R}^n$$

locally around $(\bar{x}, \bar{y})$.

Proof. Since$f$ is subdifferentially continuous at $\bar{x}$ for $\bar{y} \in \partial f(\bar{x})$, then $(x_n, y_n) \to (\bar{x}, \bar{y})$ with $y_n \in \partial f(x_n)$ implies $f(x_n) \to f(\bar{x})$. According to Theorem 4.10, we know $x_n \in \mathcal{M}$ for all large n, so the result follows. □

Corollary 5.2 (Smooth reduction for subdifferential localization)**.** Suppose the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is C^2-partly smooth at the point $\bar{x}$ relative to the C^2-smooth manifold $\mathcal{M}$, and both prox-regular and subdifferentially continuous at $\bar{x}$ for $\bar{y} \in \text{ri}\, \partial f(\bar{x})$. Let h be any C^2-smooth representative of f around $\bar{x}$. Then

$$\text{gph}\, \partial f = \text{gph}\, (\nabla h + N_{\mathcal{M}}) \cap (\mathcal{M} \times \mathbb{R}^n) = \text{gph}\, \partial (f + \delta_{\mathcal{M}})$$

locally around $(\bar{x}, \bar{y})$.

Proof. This result is easily derived from Proposition 4.3 and 5.1. □

The following result gives a formula for the generalized Hessian mapping for partly smooth and prox-regular functions.

Theorem 5.3 (Generalized and covariant Hessians)**.** Suppose that the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is C^2-partly smooth at the point $\bar{x}$ relative to the C^2-smooth manifold $\mathcal{M}$ and both prox-regular and subdifferentially continuous at $\bar{x}$ for $0 \in \text{ri}\, \partial f(\bar{x})$. Then

$$\partial^2 f(\bar{x}|0)(w) = \begin{cases} \nabla^2_{\mathcal{M}} f(\bar{x})w + N_{\mathcal{M}}(\bar{x}) \text{ for } w \in T_{\mathcal{M}}(\bar{x}) \\ \emptyset \text{ for } w \notin T_{\mathcal{M}}(\bar{x}). \end{cases}$$

Proof. According to Corollary 5.2, we have that $\partial^2 f(\bar{x}|0) = \partial^2 (f + \delta_{\mathcal{M}})(\bar{x}|0)$. Then by Theorem 2.12, the result follows. □

Corollary 5.4. Suppose that the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is C^2-partly smooth at the point $\bar{x}$ relative to the C^2-smooth manifold $\mathcal{M}$ and both prox-regular and subdifferentially continuous at $\bar{x}$ for $\bar{v} \in \text{ri}\, \partial f(\bar{x})$. Let $\tilde{f}(x) = f(x) - \langle \bar{v}, x \rangle$. Then

$$\partial^2 f(\bar{x}|\bar{v})(w) = \begin{cases} \nabla^2_{\mathcal{M}} \tilde{f}(\bar{x})w + N_{\mathcal{M}}(\bar{x}) \text{ for } w \in T_{\mathcal{M}}(\bar{x}) \\ \emptyset \text{ for } w \notin T_{\mathcal{M}}(\bar{x}). \end{cases}$$

Proof. First note that

$$\partial^2 \tilde{f}(\bar{x}|0)(w) = \partial^2 (f - \langle \bar{v}, \cdot \rangle)(\bar{x}|0)(w) = \partial^2 f(\bar{x}|\bar{v})(w) \text{ for all } w.$$

Furthermore, we know $\tilde{f}$ is partly smooth at $\bar{x}$ relative to $\mathcal{M}$ and both prox-regular and subdifferentially continuous at $\bar{x}$ for $0 \in \text{ri}\, \partial \tilde{f}(\bar{x})$. According to Theorem 5.3, we have

$$\partial^2 f(\bar{x}|\bar{v})(w) = \partial^2 \tilde{f}(\bar{x}|0)(w) = \begin{cases} \nabla^2_{\mathcal{M}} \tilde{f}(\bar{x})w + N_{\mathcal{M}}(\bar{x}) \text{ for } w \in T_{\mathcal{M}}(\bar{x}) \\ \emptyset \text{ for } w \notin T_{\mathcal{M}}(\bar{x}). \end{cases}$$

□

Without subdifferential continuity, the above result will fail in general.

Example 5.5. Define the function $f : \mathbb{R} \to \mathbb{R}$ as follows:

$$f(x) = \begin{cases} 1 & x \in (-\infty, 0) \\ x & x \in [0, \infty) \end{cases}$$

It is easy to check that f is prox-regular at 0 with $0 \in \text{ri}\, \partial f(0)$, and partly smooth relative to the manifold $\mathcal{M} = \{0\}$. However, the function f is not subdifferentially continuous at 0 for $0 \in \partial f(0)$. Then $\text{gph}\, \partial f \neq \text{gph}\, \partial (f + \delta_{\mathcal{M}})$ locally around $(0, 0)$.

Note Corollary 2.9 gives a more concrete description of the generalized Hessian in terms of a smooth representative of f and smooth equations for $\mathcal{M}$. Next, we will use Theorem 5.3 to calculate the generalized Hessian mapping for the maximum eigenvalue function. We will use $\mathcal{U}$-Lagrangian in this example. Let's introduce the definition first. (cf. [10]).

Definition 5.6. Suppose a convex function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is C^2-partly smooth at a point $\bar{x}$ relative to a C^2-smooth manifold $\mathcal{M}$. Let $\mathcal{U}(\bar{x}) = T_{\mathcal{M}}(\bar{x})$ and $\mathcal{V}(\bar{x}) = N_{\mathcal{M}}(\bar{x})$. Given $\bar{g} \in \partial f(\bar{x})$, then the *$\mathcal{U}$-Lagrangian* of f is the function $L^f_{\mathcal{U}}(\bar{x}; \bar{g}; \cdot) : \mathcal{U}(\bar{x}) \to \mathbb{R}$ defined by

$$L^f_{\mathcal{U}}(\bar{x}; \bar{g}; u) = \inf_{v \in \mathcal{V}(\bar{x})} \{\, f(\bar{x} + u + v) - \bar{g}^T v \,\}.$$

Let $g_{\mathcal{U}}(\bar{x}) = \nabla_{\mathcal{M}} f(\bar{x})$. According to [10], we have $g_{\mathcal{U}}(\bar{x}) = \nabla_u L^f_{\mathcal{U}}(\bar{x}; g_{\mathcal{U}}(\bar{x}); 0)$ and $\nabla^2_{\mathcal{M}} f(\bar{x}) = \nabla^2_{uu} L^f_{\mathcal{U}}(\bar{x}; g_{\mathcal{U}}(\bar{x}); 0)$.

Example 5.7. Let $\mathbb{S}^n$ be the space consisting of the n-by-n real symmetric matrices. Suppose the function $\lambda_1(X) : \mathbb{S}^n \to \mathbb{R}$ maps every real symmetric matrix to its maximum eigenvalue. According to [8, Exp. 3.6], we have the following results:

1. λ_1 is partly smooth relative to the manifold

$$\mathcal{M}_m = \{\, X \in \mathbb{S}^n : \lambda_1(X) \text{ has multiplicity } m \,\} \quad (1 \le m \le n).$$

2. λ_1 is a finite-valued convex function. Hence λ_1 is prox-regular and subdifferentially continuous everywhere.

3. There is an $n \times m$ matrix $Q(X)$, depending continuously on $X \in \mathcal{M}_m$, whose columns are a basis for the eigenspace of X corresponding to $\lambda_1(X)$, and then we have

$$\begin{aligned} N_{\mathcal{M}_m}(X) &= Q(X)\{W \in \mathbb{S}^m : \text{trace } W = 0\}Q(X)^T, \\ \partial\lambda_1(X) &= Q(X)\{W \in \mathbb{S}^m_+ : \text{trace } W = 1\}Q(X)^T, \end{aligned}$$

where $\mathbb{S}^m_+$ denotes the positive semidefinite matrices.

Now suppose $\bar{X} \in \mathcal{M}_m$ and $\bar{G} \in \text{ri } \partial\lambda_1(\bar{X})$. Let $\mu(X) = \lambda_1(X) - \langle \bar{G}, X\rangle$. According to Theorem 5.3, we have

$$\partial^2\lambda_1(\bar{X}|\bar{G})(W) = \partial^2\mu(\bar{X}|0)(W) = \begin{cases} \nabla^2_{\mathcal{M}_m}\mu(\bar{X})W + N_{\mathcal{M}_m}(\bar{X}) \text{ for } W \in T_{\mathcal{M}_m}(\bar{X}) \\ \emptyset \text{ for } W \notin T_{\mathcal{M}_m}(\bar{X}). \end{cases}$$

By definition

$$L^{\lambda_1}_{\mathcal{U}}(\bar{X}; \bar{G}; U) = \inf_{V \in \mathcal{V}(\bar{X})}\{\, \lambda_1(\bar{X} + U + V) - \langle \bar{G}, V\rangle \,\} \ \text{ for } U \in T_{\mathcal{M}_m}(\bar{X}), \mathcal{V}(\bar{X}) = N_{\mathcal{M}_m}(\bar{X}).$$

Since $0 \in \partial\mu(\bar{X})$, we have $\nabla_{\mathcal{M}_m}\mu(\bar{X}) = 0$ and

$$L^{\mu}_{\mathcal{U}}(\bar{X}; 0; U) = \inf_{V \in \mathcal{V}(\bar{X})}\{\, \lambda_1(\bar{X} + U + V) - \langle \bar{G}, \bar{X} + U + V\rangle\} \ \text{ for } U \in T_{\mathcal{M}_m}(\bar{X}), \mathcal{V}(\bar{X}) = N_{\mathcal{M}_m}(\bar{X}).$$

Note that $L^{\mu}_{\mathcal{U}}(\bar{X}; 0; U) = L^{\lambda_1}_{\mathcal{U}}(\bar{X}; \bar{G}; U) - \langle \bar{G}, \bar{X} + U\rangle$. Then we have $\nabla^2_{\mathcal{M}_m}\mu(\bar{X}) = \nabla^2_{UU} L^{\mu}_{\mathcal{U}}(\bar{X}; 0; 0) = \nabla^2_{UU} L^{\lambda_1}_{\mathcal{U}}(\bar{X}; \bar{G}; 0)$. According to [15, Thm. 4.12], we have

$$\nabla^2_{UU} L^{\lambda_1}_{\mathcal{U}}(\bar{X}; \bar{G}; 0) = \text{proj}_{T_{\mathcal{M}_m}(\bar{X})} \circ H(\bar{X}, \bar{G}) \circ \text{proj}^*_{T_{\mathcal{M}_m}(\bar{X})},$$

where $H(\bar{X}, \bar{G})$ is the symmetric linear operator on $\mathbb{S}^n$ defined by

$$H(\bar{X}, \bar{G}) \cdot Y = \bar{G}Y[\lambda_1(\bar{X})I_n - \bar{X}]^\dagger + [\lambda_1(\bar{X})I_n - \bar{X}]^\dagger Y\bar{G} \text{ for all } Y \in \mathbb{S}^n.$$

($[\lambda_1(\bar{X})I_n - \bar{X}]^\dagger$ is the corresponding generalized inverse.)

For all $W \in T_{\mathcal{M}_m}(\bar{X})$, we therefore have

$$H(\bar{X}, \bar{G}) \cdot W = \bar{G}W[\lambda_1(\bar{X})I_n - \bar{X}]^\dagger + [\lambda_1(\bar{X})I_n - \bar{X}]^\dagger W\bar{G}.$$

Hence we have

$$\partial^2\lambda_1(\bar{X}|\bar{G})(W) = \begin{cases} \bar{G}W[\lambda_1(\bar{X})I_n - \bar{X}]^\dagger + [\lambda_1(\bar{X})I_n - \bar{X}]^\dagger W\bar{G} + N_{\mathcal{M}_m}(\bar{X}) \text{ for } W \in T_{\mathcal{M}_m}(\bar{X}) \\ \emptyset \text{ for } W \notin T_{\mathcal{M}_m}(\bar{X}). \end{cases}$$

6 Stability and partial smoothness

The following theorem in [18] gives a generalized Hessian characterization for tilt stability.

Theorem 6.1. For a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ having $0 \in \partial f(\bar{x})$ and such that f is both prox-regular and subdifferentially continuous at $\bar{x}$ for 0, the point $\bar{x}$ gives a tilt stable local minimum of f if and only if the mapping $\partial^2 f(\bar{x}|0)$ is positive definite in the sense that

$$\langle z, w\rangle > 0 \text{ whenever } z \in \partial^2 f(\bar{x}|0)(w), w \neq 0.$$

In this case, the mapping M from Definition 2.1 and $(\partial f)^{-1}$ have locally identical graphs around the point $(0, \bar{x})$.

Proof. See [18, Thm. 1.3]. □

With the assumption of Theorem 6.1, suppose in addition that f is C^2-partly smooth at $\bar{x}$ relative to the C^2-smooth manifold $\mathcal{M}$. Then, by combining the result above with our Hessian calculations in the previous section, we easily deduce the equivalence of the following properties.

(a) The point $\bar{x}$ is a tilt stable local minimum of the function f.

(b) The point $\bar{x}$ is a tilt stable local minimum of the function $f + \delta_{\mathcal{M}}$.

(c) The point $\bar{x}$ is a strong critical point of f relative to $\mathcal{M}$.

To see this note that $\partial^2 f(\bar{x}|0) = \partial^2(f + \delta_{\mathcal{M}})(\bar{x}|0)$ by Corollary 5.2, so (a) and (b) are equivalent by Theorem 6.1. We also know that (b) is equivalent to $\partial^2(f + \delta_{\mathcal{M}})(\bar{x}|0)$ being positive definite, which is also equivalent to

$$\langle \nabla^2 L(\bar{x})w, w\rangle > 0 \text{ for any } 0 \neq w \in T_{\mathcal{M}}(\bar{x})$$

with L the Lagrangian of Corollary 2.9. This in turn is equivalent to $\bar{x}$ being a strong critical point of f relative to $\mathcal{M}$, according to [8, p. 25]. Therefore the result follows.

With a little extra care, we can dispense with the assumption of subdifferential continuity. We use the following easy tool.

Proposition 6.2 (Local minimizers and perturbation)**.** Suppose the point $\bar{x}$ gives a tilt stable local minimum of the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. If a sequence of points $v_k \in \mathbb{R}^n \to 0$, the mapping M in Definition 2.1 satisfies

$$M(v_k) \to \bar{x} \text{ and } f(M(v_k)) \to f(\bar{x}).$$

Proof. Since M is Lipschitz at 0, then $x_k := M(v_k) \to \bar{x}$. Note $f(x) > f(\bar{x})$ for any $\bar{x} \neq x \in B_\delta(\bar{x})$. Suppose $f(x_k) \to f(\bar{x})$ is not true. Without loss of generality, we can assume that there exists an $\epsilon > 0$ such that $|f(x_k) - f(\bar{x})| > \epsilon$ for all large k. Since $\bar{x}$ is a strict local minimizer, then

$$f(x_k) > f(\bar{x}) + \epsilon.$$

Take limits on both sides. We get

$$\text{liminf}_{k\to\infty} f(x_k) \geq f(\bar{x}) + \epsilon,$$

which is contradictory to the fact that $f(x)$ is locally l.s.c at $\bar{x}$. Therefore $f(x_k) \to f(\bar{x})$. □

We now have our main result.

Theorem 6.3 (Strong criticality point and tilt stability)**.** Suppose the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is C^2-partly smooth at the point $\bar{x}$ relative to the C^2-smooth manifold $\mathcal{M}$, and prox-regular at $\bar{x}$ for $0 \in \text{ri}\, \partial f(\bar{x})$. Then the following are equivalent

(a) The point $\bar{x}$ is a tilt stable local minimum of the function f;

(b) the point $\bar{x}$ is a tilt stable local minimum of the function $f + \delta_{\mathcal{M}}$;

(c) the point $\bar{x}$ is a strong critical point of f relative to $\mathcal{M}$;

(d) the function f grows quadratically near $\bar{x}$.

Proof. By Proposition 4.13, we know (c)⇔(d). Since $f + \delta_{\mathcal{M}}$ is both prox-regular and subdifferentially continuous at $\bar{x}$ for 0, then we know that

$$\langle \nabla^2 L(\bar{x})w, w\rangle > 0 \text{ for any } w \in T_{\mathcal{M}}(\bar{x})$$

by Theorem 6.1. This is also equivalent to $\bar{x}$ being a strong critical point of f relative to $\mathcal{M}$ by previous argument. Therefore (b)⇔(c). Since f is partly smooth at $\bar{x}$ relative to $\mathcal{M}$ and prox-regular at $\bar{x}$ for $0 \in \text{ri}\, \partial f(\bar{x})$, then for any $(x_k, v_k) \to (\bar{x}, 0)$ with $v_k \in \partial f(x_k)$, we have $x_k = M(v_k)$ and $f(x_k) \to f(\bar{x})$ for large k by Proposition 6.2. Hence $x_k = M(v_k) \in \mathcal{M}$ for all large k, according to Theorem 4.10. Therefore for all large k, we have

$$\begin{aligned} M(v_k) &= \text{argmin}_{|x-x_k|\le\delta} \{ f(x) - f(\bar{x}) - \langle v_k, x-\bar{x}\rangle \} \\ &= \text{argmin}_{|x-x_k|\le\delta} \{ f(x) + \delta_{\mathcal{M}}(x) - f(\bar{x}) - \delta_{\mathcal{M}}(\bar{x}) - \langle v_k, x-\bar{x}\rangle \}. \end{aligned}$$

Hence the point $\bar{x}$ gives a tilt stable local minimum of f if and only if $\bar{x}$ gives a tilt stable local minimum of $f + \delta_{\mathcal{M}}$. In other words, we have (a)⇔(b). Then the theorem follows. □

Note that it is possible to give a direct proof of the above theorem without using generalized Hessian mappings.

7 Strong metric regularity and tilt stability

In this section, we first note that tilt stability is equivalent to "strong metric regularity" of the subdifferential.[1]

Definition 7.1. A set-valued mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *strongly metrically regular* at $\bar{x}$ for $\bar{v}$ if S^{-1} has a Lipschitz continuous single-valued localization around $\bar{v}$ for $\bar{x}$. (cf. [4])

Proposition 7.2 (Strong metric regularity)**.** Suppose the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is locally lower semicontinuous at $\bar{x}$. Then the following are equivalent:

1. The point $\bar{x}$ gives a tilt stable local minimum for the function f.
2. The point $\bar{x}$ is a strict local minimizer of the function f and the subgradient mapping ∂f is strongly metrically regular at $\bar{x}$ for 0.

Proof. (1) ⇒ (2) Suppose the point $\bar{x}$ gives a tilt stable local minimum to the function f. Then we know

$$M(v) : v \mapsto \text{argmin}_{|x-\bar{x}|\le\delta}\{ f(x) - f(\bar{x}) - \langle v, x-\bar{x}\rangle \}$$

is single-valued and Lipschitz continuous on around 0 with $M(0) = \bar{x}$. Note that $M(v) = (\partial f)^{-1}(v) \cap \overline{B}_\delta(\bar{x})$, where $\overline{B}_\delta(\bar{x}) = \{ x \mid |x-\bar{x}| \le \delta \}$, for any v close to 0. Hence ∂f is strongly metrically regular at $\bar{x}$ for 0.

(2) ⇒ (1) Since $\bar{x}$ is a strict local minimum , then there exists a $\delta > 0$ such that $f(x) > f(\bar{x})$ for any $\bar{x} \ne x \in \overline{B}_\delta(\bar{x})$. We claim that if $v_k \to 0$ and x_k minimizes $f(x) - \langle v_k, x\rangle$ over $\overline{B}_\delta(\bar{x})$, then $x_k \to \bar{x}$. Suppose the claim is not true. Then, there exists an $\epsilon > 0$ such that there are sequences $v_k \to 0$ and x_k minimizing $f(x) - \langle v_k, x\rangle$ over $\overline{B}_\delta(\bar{x})$ with $|x_k - \bar{x}| > \epsilon$. So

$$f(x_k) - \langle v_k, x_k\rangle \le f(\bar{x}) - \langle v_k, \bar{x}\rangle.$$

Without loss of generality, choose a subsequence of x_r which converges to $\hat{x}$. Since f is locally lower semicontinuous at $\bar{x}$, we have

$$f(\hat{x}) \le f(\bar{x})$$

by taking limits on both sides. We get a contradiction. Next we define the following mapping

$$M(v) : v \mapsto \text{argmin}_{|x-\bar{x}|\le\delta}\{ f(x) - f(\bar{x}) - \langle v, x-\bar{x}\rangle \} \text{ with } M(0) = \bar{x}.$$

According to the claim, we know that $M(v)$ should lie in the interior of $B_\delta(\bar{x})$ for small v. Therefore $M(v)$ are also critical points of $f(x) - \langle v, x\rangle$ for all small v . Since ∂f is strongly metrically regular at $\bar{x}$ for 0, then $M(v)$ is single-valued and Lipschitz continuous around 0. Therefore $\bar{x}$ gives a tilt stable local minimum of the function f. □

[1]After completing an initial version of this work, the authors became aware of recent work analogous to Proposition 7.2 below-see [14, Cor. 5.3].

[1] showed that for a proper lower semicontinuous convex function in a Hilbert space, the strong metric regularity property of its subdifferential is equivalent to a quadratic growth condition involving the function.

Theorem 7.3. Suppose $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a proper lower semicontinuous convex function. Then ∂f is strongly metrically regular at $\bar{x}$ for $\bar{v}$ if and only if there exist neighborhoods X of $\bar{x}$ and V of $\bar{v}$ and a positive constant c such that, for any $v \in V$ there is $\tilde{x} \in \mathbb{R}^n$ such that $\partial^{-1} f(v) = \{\tilde{x}\}$ and

$$f(x) \geq f(\tilde{x}) - \langle v, \tilde{x} - x \rangle + c|x - \tilde{x}|^2 \text{ whenever } x \in X.$$

Proof. See [1, Cor. 3.9]. □

Theorem 6.3 shows that tilt stability is equivalent to a quadratic growth condition for prox-regular and partly smooth functions, which is also equivalent to the strong metric regularity of the subdifferential by Proposition 7.2. On the other hand, Theorem 7.3 implies that strong metric regularity of the subdifferential is equivalent to a quadratic growth condition, for convex functions. In this sense, Proposition 7.2 is an analogue of Theorem 7.3 for a broader class of functions.

Acknowledgement

We would like to thank Dmitriy Drusvyatskiy for pointing out the simple proof of Corollary 4.13, part 2.

References

[1] F. J. A. Artacho and M. H. Geoffroy, *Characterization of metric regularity of subdifferentials*, Convex Anal., 15 (2008), pp. 365–380.

[2] J. Bolte, A. Daniilidis, and A. S. Lewis, *Tame mappings are semismooth*, Math. Program., Vol. 117 (2009).

[3] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[4] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solutions Mappings*, Springer, New York, 2009.

[5] W. L. Hare and A. S. Lewis, *Identifying active constraints via partial smoothness and prox-regularity*, Convex Anal., 11 (2004), pp. 251–266.

[6] R. Henrion, J. Outrata, and T. Surowiec, *On the co-derivative of normal cone mappings to inequality systems*, Nonlinear Anal., 71 (2009), pp. 1213–1226.

[7] J. M. Lee, *Introduction to Smooth Manifolds*, Springer, New York, 2006.

[8] A. S. Lewis, *Active sets, nonsmoothness and sensitivity*, SIAM J. Optim., 13 (2003), pp. 702–725.

[9] R. Mifflin and C. Sagastizábal, *A $\mathcal{VU}$-algorithm for convex minimization*, Math. Program., 104 (2005), pp. 583–608.

[10] S. A. Miller and J. Malick, *Newton methods for nonsmooth convex minimization: connections among $\mathcal{U}$-lagrangian, Riemannian Newton and SQP methods*, Math. Progam., 104 (2005), pp. 609–633.

[11] B. S. Mordukhovich, *Sensitivity analysis in nonsmooth optimization*, SIAM J. Appl. Math., 58 (1992), pp. 32–46.

[12] ——, *Variational Analysis and Generalized Differentiation*, Springer, Berlin, 2006.

[13] B. S. Mordukhovich and J. Outrata, *On second-order subdifferentials and their applications*, SIAM J. Optim., 1 (2001), pp. 139–169.

[14] B. S. Mordukhovich and R. T. Rockafellar, *Second-order subdifferential calculus with applications to optimization*, (2011). Manuscript.

[15] F. Oustry, *The $\mathcal{U}$-lagrangian of the maximum eigenvalue function*, SIAM J. Optim., 9 (1999), pp. 526–549.

[16] J. V. Outrata and H. Ramírez C., *On the Aubin property of critical points to perturbed second-order cone programs*, SIAM J. Optim., 21 (2011), pp. 798–823.

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

[18] ——, *Tilt stability of a local minimum*, SIAM J. Optim., 8 (1998), pp. 287–299.

[19] R. A. Poliquin, R. T. Rockafellar, and L. Thibault, *Local differentiability of distance functions*, Trans. Amer. Math. Soc., 352 (2000), pp. 5231–5249.

[20] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1998.

[21] S. J. Wright, *Identifiable surfaces in constrained optimization*, SIAM J. Control Optim., 31 (1993), pp. 1063–1079.