Optimality, identifiability, and sensitivity

D. Drusvyatskiy · A. S. Lewis

Abstract Around a solution of an optimization problem, an "identifiable" subset of the feasible region is one containing all nearby solutions after small perturbations to the problem. A quest for only the most essential ingredients of sensitivity analysis leads us to consider identifiable sets that are "minimal". This new notion lays a broad and intuitive variational-analytic foundation for optimality conditions, sensitivity, and active set methods.

D. Drusvyatskiy
School of Operations Research and Information Engineering, Cornell University, Ithaca, New York, USA;
Tel.: (607) 255-4856
Fax: (607) 255-9129
E-mail: dd379@cornell.edu
`http://people.orie.cornell.edu/dd379/`.

A. S. Lewis
School of Operations Research and Information Engineering, Cornell University, Ithaca, New York, USA;
Tel.: (607) 255-9147
Fax: (607) 255-9129
E-mail: aslewis@orie.cornell.edu
`http://people.orie.cornell.edu/aslewis/`.

1 Introduction

Active set ideas permeate traditional nonlinear optimization. Classical problems involve a list of smooth nonlinear constraints: the active set for a particular feasible solution — the collection of binding constraints at that point — is crucial in first and second order optimality conditions, in sensitivity analysis, and for certain algorithms. Contemporary interest in more general constraints (such as semidefiniteness) suggests a reappraisal. A very thorough modern study of sensitivity analysis in its full generality appears in [2]. Approaches more variational-analytic in flavor appear in texts such as [29]. Our aim here is rather different: to present a simple fresh approach, combining wide generality with mathematical elegance.

Our approach has its roots in the notion of an "identifiable surface" [32], and its precursors [10, 6, 5, 4, 13, 1, 14]. In essence, the idea is extremely simple: given a critical point x for a function f, a set M is *identifiable* if any sequence of points approaching x that is approximately critical (meaning corresponding subgradients approach zero) must eventually lie in M. The terminology comes from the idea that an iterative algorithm that approximates x along with an approximate criticality certificate must "identify" M. To take the classical example where f is a pointwise maximum of smooth functions, around any critical point x, assuming a natural constraint qualification, we can define M as those points with the same corresponding "active set" of functions attaining the maximum.

Identifiable sets M are useful computationally because the problem of minimizing the function f near the critical point x is equivalent to minimizing the restriction of f to M, which may be an easier problem, and because the identifiability property allows convergent algorithms to find M — the motivation for active set methods. We show moreover how M is a natural tool for optimality conditions: under reasonable conditions, quadratic growth of f around x is equivalent to quadratic growth on M — a potentially easier condition to check.

Clearly the smaller the identifiable set M, the more informative it is. Ideal would be a "locally minimal identifiable set". We note that such sets may fail to exist, even for finite convex functions f. However, when a minimal identifiable set M does exist, we show that it is both unique (locally), and central to sensitivity analysis: it consists locally of all critical points of small linear perturbations to f. We show furthermore that, under reasonable conditions, variational analysis of f simplifies because, locally, the graph of its subdifferential mapping is influenced only by the restriction of f to M. One appealing consequence is a close relationship between minimal identifiable sets and critical cones appearing in the study of variational inequalities.

The case when an identifiable set M is in fact a manifold around the point x (as in the classical example above) is particularly interesting. Remarkably, this case is equivalent to a powerful but seemingly stringent list of properties known as "partial smoothness" [19], nondegeneracy and prox-regularity — related work on "$\mathcal{VU}$ algorithms" and "the fast track" appears in [20–22] and

[15,16]. By contrast, our approach here is to offer a concise mathematical development emphasizing how this important scenario is in fact very natural indeed.

The outline of the paper is as follows. In Section 2, we introduce the notion of identifiability for arbitrary set-valued mappings. Then in Section 3, we specialize this idea to subdifferential mappings, laying the foundation for the rest of the paper. Section 4 contains basic examples of identifiable sets, while Section 5 establishes a calculus of identifiability, which could be skipped at first reading. Arriving at our main results in Section 6, we study variational geometry of identifiable sets; this in particular allows us to establish a strong relationship between identifiable sets and critical cones in Section 7. Finally in Section 8, we consider optimality conditions in the context of identifiable sets, while in Section 9 we establish a relationship between identifiable manifolds and partial smoothness — one of our central original goals. In Section 10, results of Section 6 are generalized to a function setting.

2 Identifiability in set-valued analysis

A *set-valued mapping* G from $\mathbf{R}^n$ to $\mathbf{R}^m$, denoted by $G \colon \mathbf{R}^n \rightrightarrows \mathbf{R}^m$, is a mapping from $\mathbf{R}^n$ to the power set of $\mathbf{R}^m$. Thus for each point $x \in \mathbf{R}^n$, $G(x)$ is a subset of $\mathbf{R}^m$. The *domain*, *graph*, and *range* of G are defined to be

$$\operatorname{dom} G := \{x \in \mathbf{R}^n : G(x) \neq \emptyset\},$$
$$\operatorname{gph} G := \{(x,y) \in \mathbf{R}^n \times \mathbf{R}^m : y \in G(x)\},$$
$$\operatorname{rge} G := \bigcup_{x \in \mathbf{R}^n} G(x),$$

respectively. Observe that $\operatorname{dom} G$ and $\operatorname{rge} G$ are images of $\operatorname{gph} G$ under the projections $(x,y) \mapsto x$ and $(x,y) \mapsto y$, respectively.

For two sets M and Q in $\mathbf{R}^n$, we will say that the inclusion $M \subset Q$ *holds locally around a point* $\bar{x}$, if there exists a neighborhood U of $\bar{x}$ satisfying $M \cap U \subset Q \cap U$.

The key property that we explore in this work is that of finite identification.

Definition 2.1 (Identifiable sets) Consider a mapping $G \colon \mathbf{R}^n \rightrightarrows \mathbf{R}^m$. We say that a subset $M \subset \mathbf{R}^n$ is *identifiable* at $\bar{x}$ for $\bar{v}$, where $\bar{v} \in G(\bar{x})$, if the inclusion

$$\operatorname{gph} G \subset M \times \mathbf{R}^n \text{ holds locally around } (\bar{x}, \bar{v}).$$

Equivalently, a set M is identifiable at $\bar{x}$ for $\bar{v} \in G(\bar{x})$ if for any sequence $(x_i, v_i) \to (\bar{x}, \bar{v})$ in $\operatorname{gph} G$, the points x_i must lie in M for all sufficiently large indices i. Clearly M is identifiable at $\bar{x}$ for $\bar{v}$ if and only if the same can be said of $M \cap \operatorname{dom} G$. Hence we will make light of the distinction between two such sets.

Clearly $\operatorname{dom} G$ is identifiable at $\bar{x}$ for $\bar{v} \in G(\bar{x})$. More generally, if $\bar{v}$ lies in the interior of some set U, then $G^{-1}(U)$ is identifiable at $\bar{x}$ for $\bar{v}$. Sometimes

all identifiable subsets of $\operatorname{dom} G$ arise locally in this way. In particular, one can readily check that this is the case for any set-valued mapping G satisfying $G^{-1} \circ G = \mathrm{Id}$, an important example being the inverse of the projection map $G = P_Q^{-1}$ onto a nonempty, closed, convex set Q.

The "smaller" the set M is, the more interesting and the more useful it becomes. Hence an immediate question arises. When is the identifiable set M *locally minimal*, in the sense that for any other identifiable set M' at $\bar{x}$ for $\bar{v}$, the inclusion $M \subset M'$ holds locally around $\bar{x}$? The following notion will be instrumental in addressing this question.

Definition 2.2 (Necessary sets) Consider a set-valued mapping $G \colon \mathbf{R}^n \rightrightarrows \mathbf{R}^m$, a point $\bar{x} \in \mathbf{R}^n$, and a vector $\bar{v} \in G(\bar{x})$. We say that a subset $M \subset \mathbf{R}^n$, containing $\bar{x}$, is *necessary* at $\bar{x}$ for $\bar{v}$ if the function

$$x \mapsto d(\bar{v}, G(x)),$$

restricted to M, is continuous at $\bar{x}$.

Thus M is necessary at $\bar{x}$ for $\bar{v} \in G(\bar{x})$ if for any sequence $x_i \to \bar{x}$ in M, there exists a sequence $v_i \in G(x_i)$ with $v_i \to \bar{v}$. The name "necessary" arises from the following simple observation.

Lemma 2.3 *Consider a set-valued mapping $G \colon \mathbf{R}^n \rightrightarrows \mathbf{R}^m$, a point $\bar{x} \in \mathbf{R}^n$, and a vector $\bar{v} \in G(\bar{x})$. Let M and M' be two subsets of $\mathbf{R}^n$. Then the implication*

$$\left. \begin{array}{l} M \text{ is identifiable at } \bar{x} \text{ for } \bar{v} \\ M' \text{ is necessary at } \bar{x} \text{ for } \bar{v} \end{array} \right\} \Rightarrow M' \subset M \text{ locally around } \bar{x},$$

holds.

The following elementary characterization of locally minimal identifiable sets will be used extensively in the sequel, often without an explicit reference.

Proposition 2.4 (Characterizing locally minimal identifiability)
Consider a set-valued mapping $G \colon \mathbf{R}^n \rightrightarrows \mathbf{R}^m$ and a pair $(\bar{x}, \bar{v}) \in \operatorname{gph} G$. The following are equivalent.

1. *M is a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$.*
2. *There exists neighborhood V of $\bar{v}$ such that for any subneighborhood $W \subset V$ of $\bar{v}$, the representation*

$$M = G^{-1}(W) \text{ holds locally around } \bar{x}.$$

3. *M is a locally maximal necessary set at $\bar{x}$ for $\bar{v}$.*
4. *M is identifiable and necessary at $\bar{x}$ for $\bar{v}$.*

Proof $(1) \Rightarrow (2)$: Suppose M is a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$. Then there exists a neighborhood V of $\bar{v}$ satisfying

$$G^{-1}(V) \subset M \text{ locally around } \bar{x}.$$

In turn, the set $G^{-1}(V)$ is identifiable at $\bar{x}$ for $\bar{v}$. Consequently, by local minimality of M we deduce the equality $M = G^{-1}(V)$ locally around $\bar{x}$. Observe that the same argument is valid for any subneighborhood $W \subset V$ of $\bar{v}$. The result follows.

(2) $\Rightarrow$ (3) : We first prove that M is necessary at $\bar{x}$ for $\bar{v}$. Consider any sequence $x_i \to \bar{x}$ in M and, for the sake of contradiction, suppose that there do not exist vectors $v_i \in G(x_i)$ with $v_i \to \bar{v}$. Then there exists a real number $\epsilon > 0$ such that, restricting to a subsequence, we have $G(x_i) \cap B_\epsilon(\bar{v}) = \emptyset$. Consequently, we obtain $x_i \notin G^{-1}(V \cap B_\epsilon(\bar{v}))$ for all large indices i. We conclude that $G^{-1}(V \cap B_\epsilon(\bar{v}))$ and M do not coincide on any neighborhood of $\bar{x}$, which is a contradiction. Hence M is necessary at $\bar{x}$ for $\bar{v}$.

Now consider any other necessary set M' at $\bar{x}$ for $\bar{v}$ and a sequence $x_i \to \bar{x}$ in M'. By necessity of M', there exist vectors $v_i \in G(x_i)$ converging to $\bar{v}$. By the defining property of M, we deduce $x_i \in M$ for all large indices i. Hence the inclusion $M' \subset M$ holds locally around $\bar{x}$. We conclude that M is a locally maximal necessary set at $\bar{x}$ for $\bar{v}$.

(3) $\Rightarrow$ (4) : Consider any sequence $(x_i, v_i) \to (\bar{x}, \bar{v})$ in gph G. Observe that the set

$$M' := M \cup \bigcup_{i=1}^{\infty} \{x_i\},$$

is necessary at $\bar{x}$ for $\bar{v}$. Consequently, by local maximality of M, we have $x_i \in M$ for all large indices i. Thus M is identifiable at $\bar{x}$ for $\bar{v}$.

(4) $\Rightarrow$ (1) : This is immediate from Lemma 2.3.

Remark 2.5 It is clear from Proposition 2.4 that whenever locally minimal identifiable sets exist, they are locally unique. That is, if M_1 and M_2 are both locally minimal identifiable sets at $\bar{x}$ for $\bar{v} \in G(\bar{x})$, then we have $M_1 = M_2$ locally around $\bar{x}$.

The central goal in sensitivity analysis is to understand the behavior of solutions x, around $\bar{x}$, to the inclusion

$$v \in G(x),$$

as v varies near $\bar{v}$. Characterization 2 of Proposition 2.4 shows that a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$ is a locally minimal set that captures *all* the sensitivity information about the inclusion above.

This characterization yields a constructive approach to finding locally minimal identifiable sets. Consider any open neighborhoods $V_1 \supset V_2 \supset V_3 \supset \dots$, around $\bar{v}$ with the diameters of V_i tending to zero. If the chain

$$G^{-1}(V_1) \supset G^{-1}(V_2) \supset G^{-1}(V_3) \supset \dots,$$

stabilizes, in the sense that for all large indices i and j, we have $G^{-1}(V_i) = G^{-1}(V_j)$ locally around $\bar{x}$, then $G^{-1}(V_i)$ is a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$, whenever i is sufficiently large. Moreover, the locally minimal identifiable set at $\bar{x}$ for $\bar{v}$, if it exists, must arise in this way.

The following example shows that indeed a set-valued mapping can easily fail to admit a locally minimal identifiable set.

Example 2.6 (Failure of existence) Consider the mapping $G\colon \mathbf{R}^2 \rightrightarrows \mathbf{R}$, defined in polar coordinates, by

$$G(r,\theta) = \begin{cases} |\theta| & \text{if } r \neq 0, \ \theta \in [-\pi, \pi], \\ [-1, 1] & \text{if } r = 0. \end{cases}$$

Let $\bar{x}$ be the origin in $\mathbf{R}^2$ and $\bar{v} := 0 \in G(\bar{x})$. Observe that for $\epsilon \to 0$, the preimages

$$G^{-1}(-\epsilon, \epsilon) = \{(r, \theta) \in \mathbf{R}^2 : G(r, \theta) \cap (-\epsilon, \epsilon) \neq \emptyset\} = \{(r, \theta) : |\theta| < \epsilon\},$$

never coincide around $\bar{x}$. Consequently, there is no locally minimal identifiable set at $\bar{x}$ for $\bar{v}$.

Notwithstanding the previous example, locally minimal identifiable sets do often exist. Before proceeding, we recall the following two notions of continuity for set-valued mappings.

Definition 2.7 (Continuity) Consider a set-valued mapping $F\colon \mathbf{R}^n \rightrightarrows \mathbf{R}^m$.

1. F is *outer semicontinuous* at a point $\bar{x} \in \mathbf{R}^n$ if for any sequence of points $x_i \in \mathbf{R}^n$ converging to $\bar{x}$ and any sequence of vectors $v_i \in F(x_i)$ converging to $\bar{v}$, we must have $\bar{v} \in F(\bar{x})$.
2. F is *inner semicontinuous* at $\bar{x}$ if for any sequence of points x_i converging to $\bar{x}$ and any vector $\bar{v} \in F(\bar{x})$, there exist vectors $v_i \in F(x_i)$ converging to $\bar{v}$.

If both properties hold, then we say that F is *continuous* at $\bar{x}$. We will say that F is inner-semicontinuous at $\bar{x}$, *relative* to a certain set $Q \subset \mathbf{R}^n$, if the condition above for inner-semicontinuity is satisfied for sequences $x_i \to \bar{x}$ in Q.

Clearly inner-semicontinuous mappings always admit locally minimal identifiable sets.

Proposition 2.8 (Identifiability under continuity)
Consider a set-valued mapping $G\colon \mathbf{R}^n \rightrightarrows \mathbf{R}^m$ that is inner semicontinuous, relative to $\operatorname{dom} G$, at a point $\bar{x} \in \operatorname{dom} G$. Then $\operatorname{dom} G$ is a locally minimal identifiable set at $\bar{x}$ for any vector $\bar{v} \in G(\bar{x})$.

More interesting examples can be constructed by taking pointwise unions of maps admitting locally minimal identifiable sets.

Proposition 2.9 (Pointwise union) *Consider a finite collection of outer-semicontinuous mappings, $G_i\colon \mathbf{R}^n \rightrightarrows \mathbf{R}^m$, for $i = 1, \ldots, k$. Define the pointwise union mapping $G\colon \mathbf{R}^n \rightrightarrows \mathbf{R}^m$ to be*

$$G(x) = \bigcup_{i=1}^{m} G_i(x).$$

Fix a point $\bar{x} \in \mathbf{R}^n$ and a vector $\bar{v} \in G(\bar{x})$, and suppose that for each index i, satisfying $\bar{v} \in G_i(\bar{x})$, there exists a locally minimal identifiable set M_i (with respect to G_i) at $\bar{x}$ for $\bar{v}$. Then the set

$$M := \bigcup_{i:\bar{v} \in G_i(\bar{x})} M_i,$$

is a locally minimal identifiable set (with respect to G) at $\bar{x}$ for $\bar{v}$.

Proof This readily follows from Proposition 2.4.

In particular, locally minimal identifiable sets exist for *piecewise polyhedral* mappings. These are those mappings whose graphs can be decomposed into a union of finitely many convex polyhedra.

Example 2.10 (Piecewise polyhedral mappings)
Consider a piecewise polyhedral mapping $G \colon \mathbf{R}^n \rightrightarrows \mathbf{R}^m$, where $\operatorname{gph} G = \bigcup_{i=1}^k V_i$ and $V_i \subset \mathbf{R}^n$ are convex polyhedral sets. It is easy to check that set-valued mappings whose graphs are convex polyhedral are inner-semicontinuous on their domains. Fix a point $\bar{x} \in \mathbf{R}^n$ and a vector $\bar{v} \in G(\bar{x})$, and let $\pi \colon \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^n$ be the canonical projection onto $\mathbf{R}^n$. Consequently, by Propositions 2.8 and 2.9, the set

$$\bigcup_{i:(\bar{x},\bar{v}) \in V_i} \pi(V_i)$$

is a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$.

For the remainder of the current work, we will be investigating the notion of identifiability in the context of the workhorse of variational analysis, the *subdifferential* mapping (Definition 3.5).

3 Identifiability in variational analysis

3.1 Preliminaries from variational analysis

In this section, we briefly summarize some of the fundamental tools used in variational analysis and nonsmooth optimization. We refer the reader to the monographs Borwein-Zhu [3], Clarke-Ledyaev-Stern-Wolenski [7], Mordukhovich [23,24], and Rockafellar-Wets [30], for more details. Unless otherwise stated, we follow the terminology and notation of [30].

The functions that we will be considering will take their values in the extended real line $\overline{\mathbf{R}} := \mathbf{R} \cup \{-\infty, \infty\}$. We say that an extended-real-valued function is proper if it is never $\{-\infty\}$ and is not always $\{+\infty\}$.

For a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$, the *domain* of f is

$$\operatorname{dom} f := \{x \in \mathbf{R}^n : f(x) < +\infty\},$$

and the *epigraph* of f is

$$\text{epi } f := \{(x, r) \in \mathbf{R}^n \times \mathbf{R} : r \geq f(x)\}.$$

We will say that f is *lower-semicontinuous* (*lsc* for short) at a point $\bar{x}$ provided that the inequality $\liminf_{x \to \bar{x}} f(x) \geq f(\bar{x})$ holds. If f is lower-semicontinuous at every point, then we will simply say that f is lower-semicontinuous.

Throughout this work, we will only use Euclidean norms. Hence for a point $x \in \mathbf{R}^n$, the symbol $|x|$ will denote the standard Euclidean norm of x. Given a set $Q \subset \mathbf{R}^n$ and a point $\bar{x} \in Q$, we let $o(|x - \bar{x}|)$ for $x \in Q$ be shorthand for a function that satisfies $\frac{o(|x-\bar{x}|)}{|x-\bar{x}|} \to 0$ whenever $x \xrightarrow{Q} \bar{x}$ with $x \neq \bar{x}$.

For a set $Q \subset \mathbf{R}^n$ and a point $x \in \mathbf{R}^n$, the distance of x from Q is

$$d_Q(x) := \inf_{y \in Q} |x - y|,$$

and the projection of x onto Q is

$$P_Q(x) := \{y \in Q : |x - y| = d_Q(x)\}.$$

Normal cones are fundamental in variational geometry. The most intuitive type of a normal cone arises from the metric projection.

Definition 3.1 (Proximal normals) Consider a set $Q \subset \mathbf{R}^n$ and a point $\bar{x} \in Q$. The *proximal normal cone* to Q at $\bar{x}$, denoted $N_Q^P(\bar{x})$, consists of all vectors $v \in \mathbf{R}^n$ such that $\bar{x} \in P_Q(\bar{x} + \frac{1}{r}v)$ for some $r > 0$. In this case, $\{\bar{x}\} = P_Q(\bar{x} + \frac{1}{r'}v)$ for each real number $r' > r$.

Geometrically, a vector $v \neq 0$ is a proximal normal to Q at $\bar{x}$ precisely when there exists a ball touching Q at $\bar{x}$ such that v points from $\bar{x}$ towards the center of the ball. Furthermore, this condition amounts to

$$\langle v, x - \bar{x} \rangle \leq O(|x - \bar{x}|^2) \quad \text{as } x \to \bar{x} \text{ in } Q.$$

Relaxing the inequality above, one obtains the following notion.

Definition 3.2 (Frechét normals) Consider a set $Q \subset \mathbf{R}^n$ and a point $\bar{x} \in Q$. The *Frechét normal cone* to Q at $\bar{x}$, denoted $\hat{N}_Q(\bar{x})$, consists of all vectors $v \in \mathbf{R}^n$ such that

$$\langle v, x - \bar{x} \rangle \leq o(|x - \bar{x}|) \text{ as } x \to \bar{x} \text{ in } Q.$$

Note that both $N_Q^P(\bar{x})$ and $\hat{N}_Q(\bar{x})$ are convex cones, while $\hat{N}_Q(\bar{x})$ is also closed.

For a set $Q \in \mathbf{R}^n$, the set-valued mapping $x \mapsto \hat{N}_Q(x)$ is not outer-semicontinuous, and hence is not robust relative to perturbations in x. To correct for that, the following definition is introduced.

Definition 3.3 (Limiting normals) Consider a set $Q \subset \mathbf{R}^n$ and a point $\bar{x} \in Q$. The *limiting normal cone* to Q at $\bar{x}$, denoted $N_Q(\bar{x})$, consists of all vectors $v \in \mathbf{R}^n$ such that there are sequences $x_i \xrightarrow{Q} \bar{x}$ and $v_i \to v$ with $v_i \in \hat{N}_Q(x_i)$.

The limiting normal cone, as defined above, consists of limits of nearby Frechét normals. In fact, the same object arises if we only allow limits of nearby proximal normals. An important and favorable situation arises when the Frechét and limiting constructions coincide.

Definition 3.4 (Clarke regularity of sets) A set $Q \subset \mathbf{R}^n$ is said to be *Clarke regular* at a point $\bar{x} \in Q$ if it is locally closed at $\bar{x}$ and every limiting normal vector to Q at $\bar{x}$ is a Frechét normal vector, that is $N_Q(\bar{x}) = \hat{N}_Q(\bar{x})$.

We can study variational properties of functions by means of normal cones to their epigraphs.

Definition 3.5 (Subdifferentials) Consider a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a point $\bar{x} \in \mathbf{R}^n$ where f is finite. The *proximal*, *Frechét*, and *limiting subdifferentials* of f at $\bar{x}$, respectively, are defined by

$$\partial_P f(\bar{x}) = \{v \in \mathbf{R}^n : (v, -1) \in N^P_{\text{epi} f}(\bar{x}, f(\bar{x}))\},$$
$$\hat{\partial} f(\bar{x}) = \{v \in \mathbf{R}^n : (v, -1) \in \hat{N}_{\text{epi} f}(\bar{x}, f(\bar{x}))\},$$
$$\partial f(\bar{x}) = \{v \in \mathbf{R}^n : (v, -1) \in N_{\text{epi} f}(\bar{x}, f(\bar{x}))\},$$

while the *horizon subdifferential* is defined by

$$\partial^\infty f(\bar{x}) = \{v \in \mathbf{R}^n : (v, 0) \in N_{\text{epi} f}(\bar{x}, f(\bar{x}))\}.$$

For $\bar{x}$ such that $f(\bar{x})$ is not finite, we follow the convention that $\partial_P f(\bar{x}) = \hat{\partial} f(\bar{x}) = \partial f(\bar{x}) = \partial^\infty f(\bar{x}) = \emptyset$.

For convex functions f, the subdifferentials $\partial_p f$, $\hat{\partial} f$, and ∂f reduce to the classical convex subdifferential, while for smooth f they coincide with the gradient mapping ∇f. See for example [30, Exercise 8.8]. In this sense, these three subdifferentials generalize the classical gradient. The horizon subdifferential plays an entirely different role; it records horizontal normals to the epigraph of the function and is instrumental in establishing subdifferential calculus rules. See [30, Theorem 10.6].

Definition 3.6 (Clarke regularity of functions) A function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ is said to be *Clarke regular* at a point $\bar{x} \in \mathbf{R}^n$ if the epigraph epi f is Clarke regular at $(\bar{x}, f(\bar{x}))$.

Given any set $Q \subset \mathbf{R}^n$ and a mapping $F \colon Q \to \widetilde{Q}$, where $\widetilde{Q} \subset \mathbf{R}^m$, we say that F is $\mathbf{C}^p$-*smooth* $(p \geq 1)$ if for each point $\bar{x} \in Q$, there is a neighborhood U of $\bar{x}$ and a $\mathbf{C}^p$ mapping $\tilde{F} \colon \mathbf{R}^n \to \mathbf{R}^m$ that agrees with F on $Q \cap U$.

Definition 3.7 (Smooth Manifolds) We say that a set $M \subset \mathbf{R}^n$ is a $\mathbf{C}^p$-*submanifold* of dimension r if for each point $\bar{x} \in M$, there is an open neighborhood U around $\bar{x}$ and a function $F \colon U \to \mathbf{R}^{n-r}$ that is $\mathbf{C}^p$-smooth with $\nabla F(\bar{x})$ of full rank and satisfying $M \cap U = \{x \in U : F(x) = 0\}$. In this case, we call F a *local defining function* for M around $\bar{x}$.

A good source on smooth manifold theory is [18].

Theorem 3.8 ([30, Example 6.8]) *Any $\mathbf{C}^1$-manifold M is Clarke regular at every point $x \in M$ and the normal cone $N_M(x)$ is equal to the normal space to M at x, in the sense of differential geometry.*

We will have occasion to use the following simple result.

Proposition 3.9 *Consider a set $M \subset \mathbf{R}^n$ and a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ that is finite-valued and $\mathbf{C}^1$-smooth on M. Then, at any point $\bar{x} \in M$, we have*

$$\hat{\partial} f(\bar{x}) \subset \nabla g(\bar{x}) + \hat{N}_M(\bar{x}),$$

where $g \colon \mathbf{R}^n \to \mathbf{R}$ is any $\mathbf{C}^1$-smooth function agreeing with f on a neighborhood of $\bar{x}$ in M.

Proof Define a function $h \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ agreeing with f on M and equalling plus infinity elsewhere. We successively deduce

$$\hat{\partial} f(\bar{x}) \subset \hat{\partial} h(\bar{x}) = \hat{\partial}(g(\cdot) + \delta_M(\cdot))(\bar{x}) = \nabla g(\bar{x}) + \hat{N}_M(\bar{x}),$$

as we need to show.

For a set $Q \subset \mathbf{R}^n$, we let δ_Q denote a function that is 0 on Q and $+\infty$ elsewhere; we call δ_Q the *indicator function* of Q. Then for any point $\bar{x} \in Q$, we have $N_Q^P(\bar{x}) = \partial_P \delta_Q(\bar{x})$, $\hat{N}_Q(\bar{x}) = \hat{\partial} \delta_Q(\bar{x})$ and $N_Q(\bar{x}) = \partial \delta_Q(\bar{x})$.

3.2 Identifiability in variational analysis

We are now ready to define the appropriate notion of identifiability in the context of optimization.

Definition 3.10 (Identifiability for functions)
Consider a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$, a point $\bar{x} \in \mathbf{R}^n$, and a subgradient $\bar{v} \in \partial f(\bar{x})$. A set $M \subset \mathbf{R}^n$ is *identifiable at $\bar{x}$ for $\bar{v}$* if for any sequences $(x_i, f(x_i), v_i) \to (\bar{x}, f(\bar{x}), \bar{v})$, with $v_i \in \partial f(x_i)$, the points x_i must all lie in M for all sufficiently large indices i.

The definition above can be interpreted in the sense of Section 2. Indeed, consider a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a subgradient $\bar{v} \in \partial f(\bar{x})$, for some point $\bar{x} \in \mathbf{R}^n$. Define the set-valued mapping

$$G \colon \mathbf{R}^n \rightrightarrows \mathbf{R} \times \mathbf{R}^n,$$

$$x \mapsto \{f(x)\} \times \partial f(x).$$

Then M is identifiable (relative to f) at $\bar{x}$ for $\bar{v}$ if and only if it is identifiable (relative to G) at $\bar{x}$ for the vector $(f(\bar{x}), \bar{v})$. Here, we have to work with the mapping G, rather than the subdifferential mapping ∂f directly, so as to facilitate coherence between normal cone mappings and subdifferential mappings via epigraphical geometry. (See Proposition 3.16.) This slight annoyance can be avoided whenever f is subdifferentially continuous at $\bar{x}$ for $\bar{v}$.

Definition 3.11 (Subdifferential continuity) A function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ is *subdifferentially continuous at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$* if for any sequences $x_i \to \bar{x}$ and $v_i \to \bar{v}$, with $v_i \in \partial f(x_i)$, it must be the case that $f(x_i) \to f(\bar{x})$.

Subdifferential continuity of a function f at $\bar{x}$ for $\bar{v}$ was introduced in [26, Definition 1.14], and it amounts to requiring the function $(x, v) \mapsto f(x)$, restricted to $\mathrm{gph}\, \partial f$, to be continuous in the usual sense at the point $(\bar{x}, \bar{v})$. In particular, any lower-semicontinuous convex function is subdifferentially continuous [30, Example 13.30].

Similarly, we define necessary sets as follows.

Definition 3.12 (Necessity for functions) Consider a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$. A set $M \subset \mathbf{R}^n$ is *necessary at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$* if both the function f and the mapping

$$x \mapsto d(\bar{v}, \partial f(x)),$$

restricted to M, are continuous at $\bar{x}$.

Specializing the characterization in Proposition 2.4 to this setting, we obtain the following.

Proposition 3.13 (Characterizing locally minimal identifiability)
Consider a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$, a point $\bar{x} \in \mathbf{R}^n$, and a subgradient $\bar{v} \in \partial f(\bar{x})$. Then the following are equivalent.

1. *M is a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$,*
2. *There exists a neighborhood V of $\bar{v}$ and a real number $\epsilon > 0$ such that for any subneighborhood $W \subset V$ of $\bar{v}$ and a real number $0 < \epsilon' < \epsilon$, the presentation*

   $$M = (\partial f)^{-1}(W) \cap \{x \in \mathbf{R}^n : |f(x) - f(\bar{x})| < \epsilon'\} \text{ holds locally around } \bar{x}.$$

3. *M is a locally maximal necessary set at $\bar{x}$ for $\bar{v}$.*
4. *M is identifiable and necessary at $\bar{x}$ for $\bar{v}$*

Definition 3.14 (Identifiability for sets) Given a set $Q \subset \mathbf{R}^n$, we will say that a subset $M \subset Q$ is *identifiable (relative to Q)* at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$ if M is identifiable (relative to δ_Q) at $\bar{x}$ for $\bar{v} \in \partial \delta_Q(\bar{x})$. Analogous conventions will hold for necessary sets and locally minimal identifiable sets.

It is instructive to observe the relationship between identifiability and the metric projection in presence of convexity.

Proposition 3.15 (Identifiability for convex sets)
Consider a closed, convex set Q and a subset $M \subset Q$. Let $\bar{x} \in M$ and $\bar{v} \in N_Q(\bar{x})$. Then the following are equivalent.

1. *M is identifiable (relative to Q) at $\bar{x}$ for $\bar{v}$.*
2. *M is identifiable (relative to P_Q^{-1}) at $\bar{x}$ for $\bar{x} + \bar{v}$.*

Analogous equivalence holds for necessary sets.

Proof Suppose that M is identifiable (relative to Q) at $\bar{x}$ for $\bar{v}$. Consider a sequence $(x_i, y_i) \to (\bar{x}, \bar{x} + \bar{v})$ in gph P_Q^{-1}. Observe $x_i = P_Q(y_i)$ and the sequence $y_i - x_i \in N_Q(x_i)$ converges to $\bar{v}$. Consequently, the points x_i all eventually lie in M.

Conversely suppose that M is identifiable (relative to P_Q^{-1}) at $\bar{x}$ for $\bar{x} + \bar{v}$. Consider a sequence $(x_i, v_i) \to (\bar{x}, \bar{v})$ in gph N_Q. Then the sequence $(x_i, x_i + v_i) \in$ gph P_Q^{-1} converges to $(\bar{x}, \bar{x} + \bar{v})$. Consequently, we have $x_i \in M$ for all large i.

We leave the verification of the analogous equivalence for necessary sets to the reader.

Thus a subset M of a closed, convex set Q is identifiable at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$ if and only if the equality, $P_Q = P_M$, holds locally around $\bar{x} + \bar{v}$.

The following simple proposition establishes epigraphical coherence, alluded to above, between normal cone mappings and subdifferential mappings in the context of identifiability.

Proposition 3.16 (Epigraphical coherence)
Consider a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a subgradient $\bar{v} \in \partial f(\bar{x})$, for some point $\bar{x} \in \mathbf{R}^n$. Then $M \subset \mathrm{dom}\, f$ is an identifiable set (relative to f) at $\bar{x}$ for $\bar{v}$ if and only if $\mathrm{gph}\, f\big|_M$ is an identifiable set (relative to $\mathrm{epi}\, f$) at $(\bar{x}, f(\bar{x}))$ for $(\bar{v}, -1)$. Analogous statements hold for necessary, and consequently for locally minimal identifiable sets.

4 Basic Examples

In this section, we present some basic examples of identifiable sets.

Example 4.1 (Smooth functions) If U is an open set containing $\bar{x}$ and $f \colon U \to \mathbf{R}$ is $\mathbf{C}^1$-smooth. Then U is a locally minimal identifiable set at $\bar{x}$ for $\nabla f(\bar{x})$.

Example 4.2 (Smooth manifolds) If M is a $\mathbf{C}^1$ manifold, the M is a locally minimal identifiable set at any $x \in M$ for any $v \in N_M(x)$. This follows immediately by observing that the normal cone mapping $x \mapsto N_M(x)$ is inner-semicontinuous on M.

We define the *support* of any vector $v \in \mathbf{R}^n$, denoted supp v, to be the set consisting of all indices $i \in \{1, \dots, n\}$ such that $v_i \neq 0$. The *rank* of v, denoted rank v, is then the size of the support supp v.

Example 4.3 (Convex polyhedra) Let $Q \subset \mathbf{R}^n$ be a convex polyhedron. Example 2.10 shows that $M := N_Q^{-1}(\bar{v})$ (equivalently, $M := \mathrm{argmax}_{x \in Q} \langle \bar{v}, x \rangle$) is a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$.

More concretely, suppose that Q has the representation

$$Q = \{x \in \mathbf{R}^n : \langle a_i, x \rangle \leq b_i \text{ for all } i \in I\}, \tag{4.1}$$

for some index set $I = \{1, \ldots, m\}$ and vectors $a_1, \ldots, a_m \in \mathbf{R}^n$ and $b \in \mathbf{R}^m$. For any point $x \in \mathbf{R}^n$, define the active index set

$$I(x) := \{i \in I : \langle a_i, x \rangle = b_i\}.$$

Then we have $N_Q(x) = \operatorname{cone}\{a_i : i \in I(x)\}$ and consequently there exist multipliers $\bar{\lambda} \in \mathbf{R}_+^m$ with $\operatorname{supp}\bar{\lambda} \subset I(\bar{x})$ satisfying $\bar{v} = \sum_{i \in I(\bar{x})} \bar{\lambda}_i a_i$. Hence for any point $y \in Q$, the equivalence

$$y \in M \Longleftrightarrow \Big\langle \sum_{i \in I(\bar{x})} \bar{\lambda}_i a_i, y \Big\rangle = \Big\langle \sum_{i \in I(\bar{x})} \bar{\lambda}_i a_i, \bar{x} \Big\rangle$$

$$\Longleftrightarrow \sum_{i \in I(\bar{x})} \bar{\lambda}_i [\langle a_i, y \rangle - b_i] = 0,$$

holds. We deduce that M has the alternate description

$$M = \{x \in Q : \operatorname{supp}\bar{\lambda} \subset I(x)\}.$$

We should note that under a strict complementarity condition, $\bar{v} \in \operatorname{ri} N_Q(\bar{x})$, we may choose $\bar{\lambda}$ with $\operatorname{supp}\bar{\lambda} = I(\bar{x})$. Then M would consist of all points $x \in Q$, whose active index set $I(x)$ coincides with $I(\bar{x})$. It is then standard to check that M coincides with an affine subspace locally around $\bar{x}$.

Example 4.4 (Polyhedral functions) Analogously, we may analyse a convex polyhedral function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$, a function whose epigraph is a convex polyhedron. To be more precise, we may express f as

$$f(x) = \begin{cases} \max_{i \in I}\{\langle a_i, x \rangle + b_i\} & \text{whenever } \langle c_j, x \rangle \le d_j \text{ for all } j \in J, \\ \infty & \text{otherwise,} \end{cases}$$

for some index sets $I = \{1, \ldots, m\}$ and $J = \{1, \ldots, k\}$, vectors $a_i, c_j \in \mathbf{R}^n$, and real numbers b_i, d_j for $i \in I$ and $j \in J$. For any point $x \in \mathbf{R}^n$, define the active index sets

$$I(x) = \{i \in I : \langle a_i, x \rangle + b_i = f(x)\},$$
$$J(x) = \{j \in J : \langle c_j, x \rangle = d_j\}.$$

A straightforward computation shows

$$\partial f(x) = \operatorname{conv}\{a_i : i \in I(x)\} + \operatorname{cone}\{c_j : j \in J(x)\}.$$

Consider a pair $(\bar{x}, \bar{v}) \in \operatorname{gph}\partial f$. Then there exist multipliers $(\bar{\lambda}, \bar{\mu}) \in \mathbf{R}_+^m \times \mathbf{R}_+^k$ satisfying

$$\bar{v} = \sum_{i \in I} \bar{\lambda}_i a_i + \sum_{j \in J} \bar{\mu}_j c_j,$$

with $\sum_{i \in I} \bar{\lambda}_i = 1$, $\operatorname{supp}\bar{\lambda} \subset I(\bar{x})$, and $\operatorname{supp}\bar{\mu} \subset J(\bar{x})$. Applying the same argument as in Example 4.3 to $\operatorname{epi} f$, we deduce that the set

$$M = \{x \in \operatorname{dom} f : \operatorname{supp}\bar{\lambda} \subset I(x), \ \operatorname{supp}\bar{\mu} \subset J(x)\},$$

is a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$. Again we should note that a particularly nice situation occurs under a strict complementarity condition, $\bar{v} \in \text{ri} \, \partial f(\bar{x})$. In this case there exist multipliers $(\bar{\lambda}, \bar{\mu})$ so that $\text{supp} \, \bar{\lambda} = I(\bar{x})$ and $\text{supp} \, \bar{\mu} = J(\bar{x})$, and then M coincides with an affine subspace locally around $\bar{x}$.

Example 4.5 (Maximum function) In particular, consider the maximum function $\text{mx} \colon \mathbf{R}^n \to \mathbf{R}$, defined by

$$\text{mx}(x) := \max\{x_1, \ldots, x_n\}.$$

Then given a point $\bar{x}$ and a vector $\bar{v} \in \partial(\text{mx})(\bar{x})$, the set $M = \{x \in \mathbf{R}^n : \text{supp} \, \bar{v} \subset I(x)\}$, where

$$I(x) := \{i : x_i = \text{mx}(x)\},$$

is a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$. Alternatively, M admits the presentation

$$M = \{x \in \mathbf{R}^n : \text{mult} \, \text{mx}(x) \geq \text{rank} \, \bar{v}\} \quad \text{locally around } \bar{x},$$

where $\text{mult} \, \text{mx}(x)$ simply denotes the size of the set $I(x)$.

Generalizing beyond polyhedrality, we now consider the so-called *piecewise linear-quadratic functions*; these are those functions whose domain can be represented as the union of finitely many convex polyhedra, so that the function is linear or quadratic on each such set. Convex piecewise linear-quadratic functions are precisely the convex functions whose subdifferential mappings are piecewise polyhedral [28].

Proposition 4.6 (Piecewise linear-quadratic functions)
Consider a convex, piecewise linear-quadratic function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$. Then there exists a locally minimal identifiable set at any point $x \in \text{dom} \, f$ for any vector $y \in \partial f(x)$.

Proof Convex piecewise linear-quadratic functions have piecewise polyhedral subdifferential mappings [28]. Consequently, Example 2.10 shows that the mapping $x \mapsto \partial f(x)$ admits a locally minimal identifiable set at any point $x \in \mathbf{R}^n$ for any vector $v \in \partial f(x)$. Since piecewise linear-quadratic functions are lower-semicontinuous [30, Proposition 10.21], and lower-semicontinuous convex functions are subdifferentially continuous [30, Example 13.30], the result follows.

We now briefly consider the three standard convex cones of mathematical programming.

Example 4.7 (Non-negative Orthant) Consider a point $\bar{x} \in \mathbf{R}_+^n$ and a vector $\bar{v} \in N_{\mathbf{R}_+^n}(\bar{x})$. Then $M := \{x \in \mathbf{R}_+^n : x_i = 0 \text{ for each } i \in \text{supp} \, \bar{v}\}$ is a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$. Observe that M also admits the presentation

$$M = \{x \in \mathbf{R}_+^n : \text{rank} \, x + \text{rank} \, \bar{v} \leq n\} \quad \text{locally around } \bar{x}.$$

Example 4.8 (Lorentz cone) Consider the Lorentz cone

$$\mathcal{L}^n := \{(x, r) \in \mathbf{R}^n \times \mathbf{R} : r \geq |x|\}.$$

Observe that $\mathcal{L}^n$ coincides with the epigraph $\mathrm{epi}\, |\cdot|$. Let $\bar{x} = 0$ and consider any $v \in \partial |\cdot|(0)$ with $|v| = 1$. Then for any real $\epsilon > 0$, the set $M_\epsilon := \{x \in \mathbf{R}^n : \langle \frac{x}{|x|}, \bar{v} \rangle \leq \epsilon\}$ is identifiable at $\bar{x}$ for $\bar{v}$. In particular, for $n \geq 2$ and $\epsilon \neq \epsilon'$ the sets M_ϵ and $M_{\epsilon'}$ do not coincide on any neighborhood of $\bar{x}$, and consequently there is no locally minimal identifiable set at $\bar{x}$ for $\bar{v}$.

In what follows $\mathbf{S}^n$ will denote the space of $n \times n$ real symmetric matrices with the trace inner product while $\mathbf{S}^n_+$ will denote the convex cone of symmetric positive semi-definite matrices. With every matrix $X \in \mathbf{S}^n$ we will associate its largest eigenvalue, denoted by $\lambda_1(X)$. The multiplicity of $\lambda_1(X)$ as an eigenvalue of X will be written as $\mathrm{mult}\, \lambda_1(X)$. Finally $\mathbf{M}^{n \times m}$ will denote the space of $n \times m$ matrices with real entries. We defer the verification of the following two examples to a forthcoming paper [9]. We should also emphasize the intriguing parallel between these two examples and Examples 4.5 and 4.7.

Example 4.9 (Positive semi-definite cone) Consider a matrix $\bar{X} \in \mathbf{S}^n_+$ and a normal $\bar{V} \in N_{\mathbf{S}^n_+}(\bar{X})$. Then

$$M = \{X \in \mathbf{S}^n_+ : \mathrm{rank}\, X + \mathrm{rank}\, \bar{V} \leq n\},$$

is an identifiable set at $\bar{X}$ for $\bar{V}$. It is interesting to note that M may fail to be locally minimal in general. Indeed, it is possible that $\mathbf{S}^n_+$ admits no locally minimal identifiable set at $\bar{X}$ for $\bar{V}$. This can easily be seen from the previous example and the fact that $\mathbf{S}^2_+$ and $\mathcal{L}^2$ are isometrically isomorphic.

However, under the strict complementarity condition $\bar{V} \in \mathrm{ri}\, N_{\mathbf{S}^n_+}(\bar{X})$, we have $\mathrm{rank}\, \bar{X} + \mathrm{rank}\, \bar{V} = n$, and consequently M coincides with $\{X \in \mathbf{S}^n_+ : \mathrm{rank}\, X = \mathrm{rank}\, \bar{X}\}$ around $\bar{X}$. It is then standard that M is an analytic manifold around $\bar{X}$, and furthermore one can show that M is indeed a locally minimal identifiable set at $\bar{X}$ for $\bar{V}$. For more details see [9].

Example 4.10 (Maximum eigenvalue) Consider a matrix $\bar{X}$ and a subgradient $\bar{V} \in \partial \lambda_1(\bar{X})$, where $\lambda_1 : \mathbf{S}^n \to \mathbf{R}$ is the maximum eigenvalue function. Then
$$M := \{X \in \mathbf{S}^n : \mathrm{mult}\, \lambda_1(X) \geq \mathrm{rank}\, \bar{V}\},$$

is an identifiable set at $\bar{X}$ for $\bar{V}$. Again under a strict complementarity condition $\bar{V} \in \mathrm{ri}\, \partial \lambda_1(\bar{X})$, we have $\mathrm{rank}\, \bar{V} = \mathrm{mult}\, \lambda_1(\bar{X})$, and consequently M coincides with the manifold $\{X \in \mathbf{S}^n : \mathrm{mult}\, \lambda_1(X) = \mathrm{mult}\, \lambda_1(\bar{X})\}$ locally around $\bar{X}$. Furthermore under this strict complementarity condition, M is locally minimal. For more details see [9].

Example 4.11 (The rank function) Consider the rank function, denoted $\mathrm{rank} : \mathbf{M}^{n \times m} \to \mathbf{R}$. Then

$$M := \{X \in \mathbf{M}^{n \times m} : \mathrm{rank}\, X = \mathrm{rank}\, \bar{X}\}$$

is a locally minimal identifiable set at $\bar{X}$ for any $\bar{V} \in \partial(\text{rank})(\bar{X})$. To see this, observe that the equality

$$\text{epi rank} = \text{epi}(\text{rank } \bar{X} + \delta_M) \text{ holds locally around } (\bar{X}, \text{rank } \bar{X}).$$

Combining this with the standard fact that M is an analytic manifold verifies the claim.

In Examples 4.8 and 4.9, we already saw that there are simple functions $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ that do not admit a locally minimal identifiable set at some point $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$. However in those examples $\bar{v}$ was degenerate in the sense that $\bar{v}$ was contained in the relative boundary of $\partial f(\bar{x})$. We end this section by demonstrating that locally minimal identifiable sets may, in general, fail to exist even for subgradients $\bar{v}$ lying in the relative interior of the convex subdifferential $\partial f(\bar{x})$.

Example 4.12 (Failure of existence)
Consider the convex function $f \colon \mathbf{R}^2 \to \mathbf{R}$, given by

$$f(x, y) = \sqrt{x^4 + y^2}.$$

Observe that f is continuously differentiable on $\mathbf{R}^2 \setminus \{(0, 0)\}$, with

$$|\nabla f(x, y)|^2 = \frac{4x^6 + y^2}{x^4 + y^2},$$

and

$$\partial f(0, 0) = \{0\} \times [-1, 1].$$

We claim that f does not admit a locally minimal identifiable set at $(0, 0)$ for the vector $(0, 0) \in \partial f(0, 0)$. To see this, suppose otherwise and let M be such a set.

Consider the curves

$$L_n := \{(x, y) \in \mathbf{R}^2 : y = \frac{1}{n}x^2\},$$

parametrized by integers n. For a fixed integer n, consider a sequence of points $(x_i, y_i) \to (0, 0)$ in L_n. Then

$$\lim_{i \to \infty} |\nabla f(x_i, y_i)| = \frac{n^2}{n^4 + 1}.$$

Since M is necessary at $(0, 0)$ for $(0, 0)$, we deduce that for each integer n, there exists a real number $\epsilon_n > 0$ such that

$$\mathbf{B}_{\epsilon_n} \cap L_n \cap M = \{(0, 0)\}.$$

However observe $\lim_{n \to \infty} \frac{n^2}{n^4 + 1} = 0$. Therefore we can choose a sequence $(x_n, y_n) \in \mathbf{B}_{\epsilon_n} \cap L_n$, with $(x_n, y_n) \neq (0, 0)$, $(x_n, y_n) \to (0, 0)$, and the gradients $\nabla f(x_n, y_n)$ tending to $(0, 0)$. Since M is identifiable at $(0, 0)$ for $(0, 0)$, the points (x_n, y_n) lie in M for all large indices n, which is a contradiction.

5 Calculus of identifiability

To build more sophisticated examples, it is necessary to develop some calculus rules. Our starting point is the following intuitive chain rule.

Proposition 5.1 (Chain Rule) *Consider a function $f(x) := g(F(x))$ defined on an open neighborhood $V \subset \mathbf{R}^n$, where $F \colon V \to \mathbf{R}^m$ is a $\mathbf{C}^1$-smooth mapping and $g \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ is a lsc function. Suppose that at some point $\bar{x} \in \operatorname{dom} f$, the qualification condition*

$$\ker \nabla F(\bar{x})^* \cap \partial^\infty g(F(\bar{x})) = \{0\}, \tag{5.1}$$

is valid, and hence the inclusion

$$\partial f(\bar{x}) \subset \nabla F(\bar{x})^* \partial g(F(\bar{x})) \text{ holds.}$$

Consider a vector $\bar{v} \in \partial f(\bar{x})$ and the corresponding multipliers

$$\Lambda := \{y \in \partial g(F(\bar{x})) : \bar{v} = \nabla F(\bar{x})^* y\}.$$

Suppose that for each vector $y \in \Lambda$, there exists an identifiable set M_y (with respect to g) at $F(\bar{x})$ for y. Then the set

$$M := \bigcup_{y \in \Lambda} F^{-1}(M_y),$$

is identifiable (with respect to f) at $\bar{x}$ for $\bar{v}$.
If, in addition,

- *g is Clarke regular at all points in $\operatorname{dom} g$ around $F(\bar{x})$,*
- *the collection $\{M_y\}_{y \in S}$ is finite, and*
- *each set M_y is a locally minimal identifiable set (with respect to g) at $F(\bar{x})$ for y,*

then M is a locally minimal identifiable set (with respect to f) at $\bar{x}$ for $\bar{v}$.

Proof We first argue the identifiability of M. To this effect, consider any sequence $(x_i, f(x_i), v_i) \to (\bar{x}, f(\bar{x}), \bar{v})$, with $v_i \in \partial f(x_i)$. It is easy to see that the transversality condition

$$\ker \nabla F(x_i)^* \cap \partial^\infty g(F(x_i)) = \{0\}, \tag{5.2}$$

holds for all sufficiently large indices i. Then by [30, Theorem 10.6], we have

$$v_i \in \partial f(x_i) \subset \nabla F(x_i)^* \partial g(F(x_i)).$$

Choose a sequence $y_i \in \partial g(F(x_i))$ satisfying $v_i = \nabla F(x_i)^* y_i$. We claim that the sequence y_i is bounded. Indeed suppose otherwise. Then restricting to a subsequence, we can assume $|y_i| \to \infty$ and $\frac{y_i}{|y_i|} \to \tilde{y}$, for some nonzero vector $\tilde{y} \in \partial^\infty g(F(\bar{x}))$. Consequently

$$\nabla F(\bar{x})^* \tilde{y} = \lim_{i \to \infty} \nabla F(x_i)^* \frac{y_i}{|y_i|} = \lim_{i \to \infty} \frac{v_i}{|y_i|} = 0,$$

thus contradicting (5.2).

Now restricting to a subsequence, we may suppose that the vectors $y_i \in \partial g(F(x_i))$ converge to $\bar{y}$ for some vector $\bar{y} \in \partial g(F(\bar{x}))$. Furthermore, observe $\bar{y} \in \Lambda$. So for all sufficiently large indices i, the points $F(x_i)$ all lie in $M_{\bar{y}}$. Consequently the points x_i lie in M for all large indices i, and we conclude that M is identifiable (with respect to f) at $\bar{x}$ for $\bar{v}$.

Now suppose that g is Clarke regular at all points of dom g near $F(\bar{x})$, the collection $\{M_y\}_{y \in S}$ is finite, and each set M_y is a locally minimal identifiable set (with respect to g) at $F(\bar{x})$ for y. We now show that M is necessary (with respect to f) at $\bar{x}$ for $\bar{v}$. To this effect, consider a sequence $x_i \to \bar{x}$ in M. Then restricting to a subsequence, we may suppose that the points $F(x_i)$ all lie in $M_{\bar{y}}$ for some $\bar{y} \in \Lambda$. Consequently there exists a sequence $y_i \in \partial g(F(x_i))$ converging to $\bar{y}$. Hence we deduce

$$v_i := \nabla F(x_i)^* y_i \to \nabla F(\bar{x})^* \bar{y} = \bar{v}.$$

Since g is Clarke regular at all points of dom g near $F(\bar{x})$, by [30, Theorem 10.6], the inclusion $v_i \in \partial f(x_i)$ holds for all large i. Hence M is necessary (with respect to f) at $\bar{x}$ for $\bar{v}$.

Our goal now is to obtain a sum rule. The passage to this result though the chain rule is fairly standard. The first step is to deal with separable functions.

Proposition 5.2 (Separable functions)
Consider proper, lsc functions $f_i \colon \mathbf{R}^{n_i} \to \overline{\mathbf{R}}$, for $i = 1, \ldots, k$, and define

$$f(x_1, \ldots, x_k) = \sum_{i=1}^{k} f_i(x_i).$$

Suppose that $M_{\bar{v}_i} \subset \mathbf{R}^{n_i}$ is an identifiable set (with respect to f_i) at $\bar{x}_i$ for $\bar{v}_i \in \partial f_i(\bar{x}_i)$, for each $i = 1, \ldots, k$. Then the set

$$M := M_{\bar{v}_1} \times \ldots \times M_{\bar{v}_k},$$

is identifiable (with respect to f) at $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_k)$ for $\bar{v} = (\bar{v}_1, \ldots, \bar{v}_k)$. An analogous result holds for necessary sets.

Proof Clearly $M := M_{\bar{v}_1} \times \ldots \times M_{\bar{v}_k}$ is identifiable for the set-valued mapping

$$(x_1, \ldots, x_k) \mapsto \prod_{i=1}^{k} \{f_i(x_i)\} \times \partial f_i(x_i),$$

at $\bar{x}$ for $\prod_{i=1}^{k} (f_i(\bar{x}_i), \bar{v}_i)$. Furthermore lower-semicontinuity of the functions f_i readily implies that M is also identifiable for

$$(x_1, \ldots, x_k) \mapsto \{f(\bar{x})\} \times \prod_{i=1}^{k} \partial f_i(x_i),$$

at $\bar{x}$ for $(f(\bar{x}), \bar{v})$. Using the identity $\partial f(x_1, \ldots, x_k) = \prod_{i=1}^{k} \partial f_i(x_i)$, we deduce the result. The argument in the context of necessary sets is similar.

Corollary 5.3 (Sum Rule) *Consider proper, lsc functions $f_i\colon \mathbf{R}^n \to \overline{\mathbf{R}}$, for $i = 1, \ldots, k$, and define the sum $f(x) = \sum_{i=1}^k f_i(x)$. Assume at some point $\bar{x} \in \operatorname{dom} f$, the qualification condition*

$$\sum_{i=1}^k v_i = 0 \text{ and } v_i \in \partial f_i^\infty(\bar{x}) \text{ for each } i \quad \Longrightarrow \quad v_i = 0 \text{ for each } i.$$

Consider a vector $\bar{v} \in \partial f(\bar{x})$ and define the set

$$\Lambda = \{(v_1, \ldots, v_k) \in \prod_{i=1}^k \partial f_i(\bar{x}) : \bar{v} = \sum_{i=1}^k v_i\}.$$

For each $(v_1, \ldots, v_k) \in \Lambda$, let M_{v_i} be an identifiable set (with respect to f_i) at $\bar{x}$ for v_i. Then

$$M := \bigcup_{(v_1, \ldots, v_k) \in \Lambda} M_{v_1} \cap \ldots \cap M_{v_k},$$

is identifiable (with respect to f) at $\bar{x}$ for $\bar{v}$.
If, in addition,

- *each f_i is Clarke regular at all points in $\operatorname{dom} f_i$ around $\bar{x}$,*
- *the collection $\{M_{v_1} \times \ldots \times M_{v_k}\}_{(v_1, \ldots, v_k) \in \Lambda}$ is finite, and*
- *for each $(v_1, \ldots, v_k) \in \Lambda$, the set M_{v_i} is a locally minimal identifiable set (with respect to f_i) at $\bar{x}$ for v_i,*

then M is a locally minimal identifiable set (with respect to f) at $\bar{x}$ for $\bar{v}$.

Proof We may rewrite f in the composite form $g \circ F$, where $F(x) := (x, \ldots, x)$ and $g(x_1, \ldots, x_k) := \sum_{i=1}^k f_i(x_i)$ is separable. Then applying Proposition 5.1 and Proposition 5.2 we obtain the result.

In particular, we now obtain the following geometric version of the chain rule.

Proposition 5.4 (Sets with constraint structure)
Consider closed sets $Q \in \mathbf{R}^n$ and $K \in \mathbf{R}^m$, and a $\mathbf{C}^1$ smooth mapping $F\colon \mathbf{R}^n \to \mathbf{R}^m$. Define the set

$$L = \{x \in Q : F(x) \in K\}.$$

Consider a pair $(\bar{x}, \bar{v}) \in \operatorname{gph} N_L$ and suppose that the constraint qualification

$$\left.\begin{array}{r} y \in N_Q(\bar{x}), w \in N_K(F(\bar{x})) \\ y + \nabla F(\bar{x})^* w = 0 \end{array}\right\} \Longrightarrow (y, w) = (0, 0),$$

holds. Define the set

$$\Lambda = \{(y, w) \in N_Q(\bar{x}) \times N_K(F(\bar{x})) : y + \nabla F(\bar{x})^* w = \bar{v}\},$$

*and for each pair $(y, w) \in \Lambda$, let M_y be an identifiable set (relative to Q) at $\bar{x}$
for y and let K_w be an identifiable set (relative to K) at $F(\bar{x})$ for w. Then*

$$M := \bigcup_{(v,w) \in \Lambda} M_v \cap F^{-1}(K_w),$$

is identifiable (relative to L) at $\bar{x}$ for $\bar{v}$.
    If, in addition,

- *Q (respectively K) is Clarke regular at each of its point near $\bar{x}$ (respectively $F(\bar{x})$),*
- *the collection $\{M_y \times K_w\}_{(y,w) \in \Lambda}$ is finite,*
- *for each $(y, w) \in \Lambda$, the set M_y (respectively K_w) is a locally minimal identifiable set with respect to Q (respectively K) at $\bar{x}$ for y (respectively at $F(\bar{x})$ for w),*

then M is a locally minimal identifiable set (relative to L) at $\bar{x}$ for $\bar{v}$.

Proof Observe $\delta_L = \delta_Q + \delta_{F^{-1}(K)}$. Combining Proposition 5.1 and Corollary 5.3, we obtain the result.

Corollary 5.5 (Max-type functions)
*Consider $\mathbf{C}^1$-smooth functions $f_i \colon \mathbf{R}^n \to \overline{\mathbf{R}}$, for $i \in I := \{1, \ldots, m\}$, and let
$f(x) := \max\{f_1(x), f_2(x), \ldots, f_m(x)\}$. For any $x \in \mathbf{R}^n$, define the active set*

$$I(x) = \{i \in I : f(x) = f_i(x)\}.$$

Consider a pair $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$, and the corresponding set of multipliers

$$\Lambda = \{\lambda \in \mathbf{R}^m : \bar{v} = \sum_{i \in I(\bar{x})} \lambda_i \nabla f_i(\bar{x}), \ \operatorname{supp} \lambda \subset I(\bar{x})\}.$$

Then

$$M = \bigcup_{\lambda \in \Lambda} \{x \in \mathbf{R}^n : \operatorname{supp} \lambda \subset I(x)\},$$

is a locally minimal identifiable set (relative to f) at $\bar{x}$ for $\bar{v}$.

Proof This follows directly from Proposition 5.1 and Example 4.4 by writing f as the composition $\operatorname{mx} \circ F$, where $F(x) = (f_1(x), \ldots, f_m(x))$.

Corollary 5.6 (Smooth constraints)
Consider $\mathbf{C}^1$-smooth functions $g_i \colon \mathbf{R}^n \to \overline{\mathbf{R}}$, for $i \in I := \{1, \ldots, m\}$, and define the set
$$Q = \{x \in \mathbf{R}^n : g_i(x) \leq 0 \text{ for each } i \in I\}.$$

For any $x \in \mathbf{R}^n$, define the active set

$$I(x) = \{i \in I : g_i(x) = 0\}.$$

and suppose that for a certain pair $(\bar{x}, \bar{v}) \in \operatorname{gph} N_Q$, the constraint qualification

$$\sum_{i \in I(\bar{x})} \lambda_i \nabla g_i(\bar{x}) = 0 \ and \ \lambda_i \geq 0 \ for \ all \ i \in I(\bar{x}) \implies \lambda_i = 0 \ for \ all \ i \in I(\bar{x}),$$

holds. Then in terms of the Lagrange multipliers

$$\Lambda := \{\lambda \in \mathbf{R}^m : \bar{v} = \sum_{i \in I(\bar{x})} \lambda_i \nabla g_i(\bar{x}), \ \text{supp} \, \lambda \subset I(\bar{x})\},$$

the set

$$M = \bigcup_{\lambda \in \Lambda} \{x \in Q : g_j(x) = 0 \ for \ each \ j \in \text{supp} \, \lambda\},$$

is a locally minimal identifiable set (relative to Q) at $\bar{x}$ for $\bar{v}$.

Proof This follows immediately from Proposition 5.1 and Example 4.7.

We end the section by observing that, in particular, the chain rule, established in Proposition 5.1, allows us to consider the rich class of fully amenable functions, introduced in [25].

Definition 5.7 (Fully amenable functions)
A function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ is *fully amenable* at $\bar{x}$ if f is finite at $\bar{x}$, and there is an open neighborhood U of $\bar{x}$ on which f can be represented as $f = g \circ F$ for a $\mathbf{C}^2$-smooth mapping $F \colon V \to \mathbf{R}^m$ and a convex, piecewise linear-quadratic function $g \colon \mathbf{R}^m \to \overline{\mathbf{R}}$, and such that the qualification condition

$$\ker \nabla F(\bar{x})^* \cap \partial g^\infty(F(\bar{x})) = \{0\},$$

holds.

The qualification condition endows the class of fully amenable functions with exact calculus rules. Such functions are indispensable in nonsmooth second order theory. For more details, see [25].

Proposition 5.8 (Identifiable sets for fully amenable functions)
A function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ that is fully amenable at a point $\bar{x} \in \mathbf{R}^n$ admits a locally minimal identifiable set at $\bar{x}$ for any vector $\bar{v} \in \partial f(\bar{x})$.

Proof This follows immediately from Propositions 4.6 and 5.1.

6 Variational geometry of identifiable sets

In the previous sections, we have introduced the notions of identifiability, analyzed when locally minimal identifiable sets exist, developed calculus rules, and provided important examples. In this section, we consider the interplay between variational geometry of a set Q and its identifiable subsets M. Considering sets rather than functions has the advantage of making our arguments entirely geometric. We begin with the simple observation that locally minimal identifiable sets are locally closed.

Proposition 6.1 *Consider a closed set $Q \subset \mathbf{R}^n$ and a subset $M \subset Q$ that is a locally minimal identifiable set at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$. Then M is locally closed at $\bar{x}$*

Proof Suppose not. Then there exists a sequence $x_i \in (\operatorname{bd} M) \setminus M$ with $x_i \to \bar{x}$. Since M is identifiable at $\bar{x}$ for $\bar{v}$, there exists a neighborhood V of $\bar{v}$ satisfying $\operatorname{cl} V \cap N_Q(x_i) = \emptyset$ for all large indices i. Observe that for each index i, every point y sufficiently close to x_i satisfies $V \cap N_Q(y) = \emptyset$. Consequently, there exists a sequence $y_i \in Q$ converging to $\bar{x}$ with $V \cap N_Q(y) = \emptyset$, which contradicts the necessity of M at $\bar{x}$ for $\bar{v}$.

Recall that for a set $Q \subset \mathbf{R}^n$ and a subset $M \subset Q$, the inclusion $\hat{N}_Q(x) \subset N_M(x)$ holds for each point $x \in M$, while the analogous inclusion for the limiting normal cone may fail. This pathology does not occur for identifiable sets.

Proposition 6.2 *Consider a closed set Q and a set M that is identifiable at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$. Then the equation*

$$\operatorname{gph} N_Q \subset \operatorname{gph} N_M \quad \text{holds locally around } (\bar{x}, \bar{v}).$$

Proof Consider a sequence $(x_i, v_i) \in \operatorname{gph} N_Q$ converging to $(\bar{x}, \bar{v})$. Then for each i, there exists a sequence $(x_i^j, v_i^j) \in \operatorname{gph} \hat{N}_Q$ converging to (x_i, v_i). For sufficiently large indices i, the points x_i^j lie in M for all large j. For such indices we have $v_i^j \in \hat{N}_Q(x_i^j) \subset \hat{N}_M(x_i^j)$, and consequently $v_i \in N_M(x_i)$. This verifies the inclusion $\operatorname{gph} N_Q \subset \operatorname{gph} N_M$ locally around $(\bar{x}, \bar{v})$.

It turns out that in order to make further headway in studying properties of identifiable sets, one must impose the condition of *prox-regularity*. This concept has been discovered and rediscovered by various authors, notably by Federer [12] and Rockafellar-Poliquin [26]. We follow the development of [26] and [27].

Definition 6.3 (Prox-regularity) We say that a set $M \subset \mathbf{R}^n$ is *prox-regular* at $\bar{x} \in M$ if it is locally closed around $\bar{x}$ and there exists a neighborhood U of $\bar{x}$, such that the projection map P_M is single-valued on U.

Prox-regularity unifies the notions of convex sets and $\mathbf{C}^2$-manifolds. A proof may be found in [30, Example 13.30, Proposition 13.32].

Theorem 6.4 (Prox-regularity under convexity and smoothness)
Convex sets and $\mathbf{C}^2$-manifolds are prox-regular at each of their points.

It will be useful to consider a variant of prox-regularity where we consider localization with respect to direction [26, Definition 2.10].

Definition 6.5 (Directional prox-regularity for sets)
Consider a set $Q \subset \mathbf{R}^n$, a point $\bar{x} \in Q$, and a normal vector $\bar{v} \in N_Q(\bar{x})$. We

say that Q is *prox-regular at $\bar{x}$ for $\bar{v}$* if Q is locally closed at $\bar{x}$ and there exist real numbers $\epsilon > 0$ and $r > 0$ such that the implication

$$\left.\begin{array}{cc} x \in Q, & v \in N_Q(x) \\ |x - \bar{x}| < \epsilon, & |v - \bar{v}| < \epsilon \end{array}\right\} \Rightarrow P_{Q \cap B_\epsilon(\bar{x})}(x + r^{-1}v) = x,$$

holds.

Observe that if the implication above holds for some $r, \epsilon > 0$, then it also holds for any $r', \epsilon' > 0$ with $r' > r$ and $\epsilon' < \epsilon$.

In particular, a set Q is prox-regular at $\bar{x}$, in the sense of Definition 6.3, if and only of Q is prox-regular at $\bar{x}$ for every vector $v \in N_Q(\bar{x})$. Clearly, if Q is prox-regular at $\bar{x}$ for $\bar{v}$, then the equalities

$$\operatorname{gph} N_Q^P = \operatorname{gph} \hat{N}_Q = \operatorname{gph} N_Q \text{ hold locally around } (\bar{x}, \bar{v}).$$

The following characterization [26, Corollary 3.4] will be of some use for us.

Proposition 6.6 (Prox-regularity and monotonicity)
For a set $Q \subset \mathbf{R}^n$ and $\bar{x} \in Q$, with Q locally closed at $\bar{x}$, the following are equivalent.

1. *Q is prox-regular at $\bar{x}$ for $\bar{v}$.*
2. *The vector $\bar{v}$ is a proximal normal to Q at $\bar{x}$, and there exists a real number $r > 0$ satisfying*

$$\langle v_1 - v_0, x_1 - x_0 \rangle \geq -r|x_1 - x_0|^2,$$

for any pairs $(x_i, v_i) \in \operatorname{gph} N_Q$ (for $i = 0, 1$) near $(\bar{x}, \bar{v})$.

So Q is prox-regular at $\bar{x}$ for a proximal normal $\bar{v} \in N_Q^P(\bar{x})$ as long as $N_Q + rI$ has a monotone localization around $(\bar{x}, \bar{v} + r\bar{x})$, for some real number $r > 0$. We may talk about prox-regularity of functions by means of epigraphical geometry [26, Theorem 3.5].

Definition 6.7 (Directional prox-regularity for functions)
We say that a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ is *prox-regular at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$* if the epigraph $\operatorname{epi} f$ is prox-regular at $(\bar{x}, f(\bar{x}))$ for the vector $(\bar{v}, -1)$.

The following proposition shows that prox-regularity of an identifiable subset M of a set Q implies that Q itself is prox-regular.

Proposition 6.8 (Prox-regularity of identifiable sets)
Consider a closed set Q and a subset $M \subset Q$ that is identifiable at $\bar{x}$ for $\bar{v} \in \hat{N}_Q(\bar{x})$. In addition, suppose that M is prox-regular at $\bar{x}$ for $\bar{v}$. Then Q is prox-regular at $\bar{x}$ for $\bar{v}$.

Proof To show that Q is prox-regular at $\bar{x}$ for $\bar{v}$, we will utilize Proposition 6.6. To this end, we first claim that the inclusion $\bar{v} \in N_Q^P(\bar{x})$ holds. To see this, choose a sequence of real numbers $r_i \to \infty$ and a sequence of points

$$x_i \in P_Q(\bar{x} + r_i^{-1}\bar{v}).$$

We have
$$v_i := r_i(\bar{x} - x_i) + \bar{v} \in N_Q^P(x_i).$$

Clearly $x_i \to \bar{x}$. We now claim that the sequence v_i converges to $\bar{v}$. Observe by definition of x_i, we have
$$|(\bar{x} - x_i) + r_i^{-1}\bar{v}| \le |r_i^{-1}\bar{v}|.$$

Squaring and cancelling terms, we obtain
$$2\langle \bar{v}, \bar{x} - x_i \rangle \le -r_i|\bar{x} - x_i|^2.$$

Combining this with the inclusion $\bar{v} \in \hat{N}_Q(\bar{x})$, we deduce
$$\frac{o(\bar{x} - x_i)}{|\bar{x} - x_i|} \le 2\langle \bar{v}, \frac{\bar{x} - x_i}{|\bar{x} - x_i|} \rangle \le -r_i|\bar{x} - x_i|$$

We conclude $r_i(\bar{x} - x_i) \to 0$ and consequently $v_i \to \bar{v}$. Since M is identifiable at $\bar{x}$ for $\bar{v}$, we deduce $x_i \in M$ for all large indices i. In addition, since M is prox-regular at $\bar{x}$ for $\bar{v}$, we have
$$x_i = P_{M \cap B_\epsilon(\bar{x})}(\bar{x} + r_i^{-1}\bar{v}) = \bar{x},$$

for some $\epsilon > 0$ and for sufficiently large indices i. Hence the inclusion $\bar{v} \in N_Q^P(\bar{x})$ holds.

Now since M is prox-regular at $\bar{x}$ for $\bar{v}$ we deduce, using Proposition 6.6, that there exists a real number $r > 0$ satisfying
$$\langle v_1 - v_0, x_1 - x_0 \rangle \ge -r|x_1 - x_0|^2,$$

for any pairs $(x_i, v_i) \in \text{gph}\, N_M$ (for $i = 0, 1$) near $(\bar{x}, \bar{v})$.

By Proposition 6.2, we have
$$\text{gph}\, N_Q \subset \text{gph}\, N_M \quad \text{locally around } (\bar{x}, \bar{v}).$$

Recalling that $\bar{v}$ is a proximal normal to Q at $\bar{x}$ and again appealing to Proposition 6.6, we deduce that Q is prox-regular at $\bar{x}$ for $\bar{v}$.

Remark 6.9 In Proposition 6.8, we assumed that the vector $\bar{v}$ is a Frechét normal. Without this assumption, the analogous result fails. For instance, consider the sets
$$Q := \{(x, y) \in \mathbf{R}^2 : xy = 0\},$$
$$M := \{(x, y) \in \mathbf{R}^2 : y = 0\}.$$

Then clearly M is identifiable at $\bar{x} := (0, 0)$ for the normal vector $\bar{v} := (0, 1) \in N_Q(\bar{x})$. However, $\bar{v}$ is not a proximal normal.

The following result brings to the fore the insight one obtains by combining the notions of identifiability and prox-regularity. It asserts that given a prox-regular identifiable set M at $\bar{x}$ for $\bar{v} \in \hat{N}_Q(\bar{x})$, not only does the inclusion $\text{gph}\, N_Q \subset \text{gph}\, N_M$ hold locally around $(\bar{x}, \bar{v})$, but rather the two sets $\text{gph}\, N_Q$ and $\text{gph}\, N_M$ coincide around $(\bar{x}, \bar{v})$.

Proposition 6.10 (Reduction I) *Consider a closed set Q and let $M \subset Q$ be a set that is prox-regular at a point $\bar{x}$ for $\bar{v} \in \hat{N}_Q(\bar{x})$. Then M is identifiable at $\bar{x}$ for $\bar{v}$ if and only if*

$$\text{gph}\, N_Q = \text{gph}\, N_M \quad \text{locally around } (\bar{x}, \bar{v}).$$

Proof We must show that locally around $(\bar{x}, \bar{v})$, we have the equivalence

$$\text{gph}\, N_Q \subset M \times \mathbf{R}^n \Leftrightarrow \text{gph}\, N_Q = \text{gph}\, N_M.$$

The implication "$\Leftarrow$" is clear. Now assume $\text{gph}\, N_Q \subset M \times \mathbf{R}^n$ locally around $(\bar{x}, \bar{v})$. By prox-regularity, there exist real numbers $r, \epsilon > 0$ so that $P_Q(\bar{x} + r^{-1}\bar{v}) = \bar{x}$ (Proposition 6.8) and so that the implication

$$\left. \begin{array}{cc} x \in M, & v \in N_M(x) \\ |x - \bar{x}| < \epsilon, & |v - \bar{v}| < \epsilon \end{array} \right\} \Rightarrow P_{M \cap B_\epsilon(\bar{x})}(x + r^{-1}v) = x,$$

holds. By Proposition 6.2, it is sufficient to argue that the inclusion

$$\text{gph}\, N_M \subset \text{gph}\, N_Q \text{ holds locally around } (\bar{x}, \bar{v}).$$

Suppose this is not the case. Then there exists a sequence $(x_i, v_i) \to (\bar{x}, \bar{v})$, with $(x_i, v_i) \in \text{gph}\, N_M$ and $(x_i, v_i) \notin \text{gph}\, N_Q$. Let $z_i \in P_Q(x_i + r^{-1}v_i)$. We have

$$(x_i - z_i) + r^{-1}v_i \in N_Q^P(z_i),$$
$$x_i \neq z_i. \tag{6.1}$$

Observe $z_i \to \bar{x}$ by the continuity of the projection map. Consequently, by the finite identification property, for large indices i, we have $z_i \in M$ and

$$x_i + r^{-1}v_i \in z_i + N_Q^P(z_i) \subset z_i + N_M(z_i).$$

Hence $z_i = P_{M \cap B_\epsilon(\bar{x})}(x_i + r^{-1}v_i) = x_i$, for large i, thus contradicting (6.1).

Recall that Proposition 6.8 shows that prox-regularity of an identifiable subset $M \subset Q$ is inherited by Q. It is then natural to consider to what extent the converse holds. It clearly cannot hold in full generality, since identifiable sets may contain many extraneous pieces. However we will see shortly that the converse does hold for a large class of identifiable sets M, and in particular for ones that are locally minimal. The key tool is the following lemma, which may be of independent interest.

Lemma 6.11 (Accessibility) *Consider a closed set $Q \subset \mathbf{R}^n$ and a subset $M \subset Q$ containing a point $\bar{x}$. Suppose that for some vector $\bar{v} \in \hat{N}_Q(\bar{x})$, there exists a sequence*

$$y_i \in N_M^P(\bar{x}) \setminus N_Q^P(\bar{x}) \text{ with } y_i \to \bar{v}.$$

Then there exists a sequence $(x_i, v_i) \in \text{gph}\, N_Q^P$ converging to $(\bar{x}, \bar{v})$ with $x_i \notin M$ for each index i.

Proof For each index i, there exists a real $r_i > 0$ satisfying $P_M(\bar{x} + r_i^{-1} y_i) = \{\bar{x}\}$. Furthermore we can clearly assume $r_i \to \infty$. Define a sequence $(x_i, v_i) \in \text{gph } N_Q^P$ by

$$x_i \in P_Q(\bar{x} + r_i^{-1} y_i) \quad \text{and} \quad v_i := r_i(\bar{x} - x_i) + y_i.$$

Observe $x_i \notin M$ since otherwise we would have $x_i = \bar{x}$ and $y_i = r_i v_i \in N_Q^P(\bar{x})$, a contradiction. By continuity of the projection P_Q, clearly we have $x_i \to \bar{x}$. Now observe

$$|(\bar{x} - x_i) + r_i^{-1} y_i| \leq r_i^{-1} |y_i|.$$

Squaring and simplifying we obtain

$$r_i |\bar{x} - x_i| + 2 \Big\langle \frac{\bar{x} - x_i}{|\bar{x} - x_i|}, y_i \Big\rangle \leq 0.$$

Since $\bar{v}$ is a Frechét normal, we deduce

$$\liminf_{i \to \infty} \Big\langle \frac{\bar{x} - x_i}{|\bar{x} - x_i|}, y_i \Big\rangle = \liminf_{i \to \infty} \Big\langle \frac{\bar{x} - x_i}{|\bar{x} - x_i|}, \bar{v} \Big\rangle \geq 0.$$

Consequently we obtain $r_i |\bar{x} - x_i| \to 0$ and $v_i \to \bar{v}$, as claimed.

Proposition 6.12 (Reduction II) *Consider a closed set $Q \subset \mathbf{R}^n$, a point $\bar{x}$, and a normal $\bar{v} \in N_Q(\bar{x})$. Suppose $\text{gph } N_Q^P = \text{gph } N_Q$ locally around $(\bar{x}, \bar{v})$, and consider a set $M := N_Q^{-1}(V)$, where V is a convex, open neighborhood of $\bar{v}$. Then the equation*

$$\text{gph } N_Q = \text{gph } N_M \ \textit{holds locally around } (\bar{x}, \bar{v}).$$

Proof First observe that since M is identifiable at $\bar{x}$ for $\bar{v}$, applying Proposition 6.2, we deduce that the inclusion $\text{gph } N_Q \subset \text{gph } N_M$ holds locally around $(\bar{x}, \bar{v})$. To see the reverse inclusion, suppose that there exists a pair $(x, v) \in \text{gph } N_M^P$, arbitrarily close to $(\bar{x}, \bar{v})$, with $v \in V$ and $v \notin N_Q^P(x)$. By definition of M, we have $N_Q^P(x) \cap V \neq \emptyset$. Let z be a vector in this intersection, and consider the line segment γ joining z and v. Clearly the inclusion $\gamma \subset V \cap N_M^P(x)$ holds. Observe that the line segment $\gamma \cap N_Q^P(x)$ is strictly contained in γ, with z being one of its endpoints. Let w be the other endpoint of $\gamma \cap N_Q^P(x)$. Then we immediately deduce that there exists a sequence

$$y_i \in N_M^P(x) \setminus N_Q^P(x) \text{ with } y_i \to w.$$

Applying Lemma 6.11, we obtain a contradiction. Therefore the inclusion $\text{gph } N_Q \supset \text{gph } N_M^P$ holds locally around $(\bar{x}, \bar{v})$. Taking the closure the result follows.

In particular, we obtain the following essential converse of Proposition 6.8.

Proposition 6.13 (Prox-regularity under local minimality)
Consider a closed set Q and a subset $M \subset Q$ that is a locally minimal identifiable set at $\bar{x}$ for $\bar{v} \in \hat{N}_Q(\bar{x})$. Then Q is prox-regular at $\bar{x}$ for $\bar{v}$ if and only if M is prox-regular at $\bar{x}$ for $\bar{v}$.

Proof The implication $\Leftarrow$ was proven in Proposition 6.8. To see the reverse implication, first recall that M is locally closed by Proposition 6.1. Furthermore Propositions 3.13 shows that there exists an open convex neighborhood V of $\bar{v}$ so that M coincides locally with $N_Q^{-1}(V)$. In turn, applying Proposition 6.12 we deduce that the equation

$$\text{gph}\, N_Q = \text{gph}\, N_M \text{ holds locally around } (\bar{x}, \bar{v}).$$

Finally by Proposition 6.6, prox-regularity of Q at $\bar{x}$ for $\bar{v}$ immediately implies that M is prox-regular at $\bar{x}$ for $\bar{v}$.

We end this section by exploring the strong relationship between identifiable sets and the metric projection map. We begin with the following proposition.

Proposition 6.14 (Identifiability and the metric projection)
Consider a closed set Q and a subset $M \subset Q$. Let $\bar{x} \in M$ and $\bar{v} \in N_Q^P(\bar{x})$. Consider the following conditions.

1. *M is identifiable (relative to Q) at $\bar{x}$ for $\bar{v}$.*
2. *For all sufficiently small $\lambda > 0$, the set M is identifiable (relative to P_Q^{-1}) at $\bar{x}$ for $\bar{x} + \lambda\bar{v}$.*

Then the implication $1 \Rightarrow 2$ holds. If in addition Q is prox-regular at $\bar{x}$ for $\bar{v}$, then the equivalence $1 \Leftrightarrow 2$ holds.

Proof $1 \Rightarrow 2$: Recall that for all small $\lambda > 0$, we have $P_Q(\bar{x} + \lambda\bar{v}) = \bar{x}$. Fix such a real number λ and consider a sequence $(x_i, y_i) \to (\bar{x}, \bar{x} + \lambda\bar{v})$ in $\text{gph}\, P_Q^{-1}$. Observe $x_i \in P_Q(y_i)$ and the sequence $\lambda^{-1}(y_i - x_i) \in N_Q(x_i)$ converges to $\bar{v}$. Consequently, the points x_i all eventually lie in M.

Suppose now that Q is prox-regular at $\bar{x}$ for $\bar{v}$.

$2 \Rightarrow 1$: We may choose $\lambda, \epsilon > 0$ as in Definition 6.5, and satisfying $P_Q(\bar{x} + \lambda\bar{v}) = \bar{x}$. Consider a sequence $(x_i, v_i) \to (\bar{x}, \bar{v})$ in $\text{gph}\, N_Q$. Then the sequence $(x_i, x_i + \lambda v_i)$ converges to $(\bar{x}, \bar{x} + \lambda\bar{v})$ and lies in $\text{gph}\, P_Q^{-1}$ for all sufficiently large i. Consequently, we have $x_i \in M$ for all large i.

Assuming prox-regularity, a simple way to generate identifiable subsets $M \subset Q$ is by projecting open sets onto Q.

Proposition 6.15 (Projections of neighborhoods are identifiable)
Consider a set $Q \subset \mathbf{R}^n$ that is prox-regular at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$. Then for all sufficiently small $\lambda > 0$, if the inclusion $\bar{x} + \lambda\bar{v} \in \text{int}\, U$ holds for some set U, then $P_Q(U)$ is identifiable at $\bar{x}$ for $\bar{v}$.

Proof Suppose $(x_i, v_i) \to (\bar{x}, \bar{v})$ in $\text{gph}\, N_Q$. Then by prox-regularity for all sufficiently small $\lambda > 0$, we have $P_Q(x_i + \lambda v_i) = x_i$ and $x_i + \lambda v_i \in U$ for all large i. We deduce $x_i \in P_Q(U)$ for all large i, as we needed to show.

In fact, we will see shortly that under the prox-regularity assumption, *all* locally minimal identifiable sets arise in this way.

Lemma 6.16 *Consider a closed set $Q \subset \mathbf{R}^n$ and a subset M that is identifiable at $\bar{x}$ for $\bar{v} \in N_Q^P(\bar{x})$. Then for all sufficiently small $\lambda > 0$ and all $\epsilon > 0$, the inclusion*

$$\bar{x} + \lambda \bar{v} \in \operatorname{int} \{x + \lambda v : x \in M, v \in N_Q(x), |x - \bar{x}| < \epsilon, |v - \bar{v}| < \epsilon\},$$

holds.

Proof For all sufficiently small $\lambda > 0$, we have $P_Q(\bar{x} + \lambda \bar{v}) = \bar{x}$. Consider a sequence $z_i \to \bar{x} + \lambda \bar{v}$ and choose points

$$x_i \in P_Q(z_i).$$

Observe $x_i \to \bar{x}$ and the vectors $\lambda^{-1}(z_i - x_i) \in N_Q(x_i)$ converge to $\bar{v}$. Consequently, the points x_i lie in M for all large i, and hence $z_i = x_i + \lambda(\lambda^{-1}(z_i - x_i))$ lie in the desired set eventually.

Proposition 6.17 (Representing locally minimal identifiable sets)
Consider a set $Q \subset \mathbf{R}^n$ that is prox-regular at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$ and let M be a locally minimal identifiable set at $\bar{x}$ for $\bar{v}$. For $\lambda, \epsilon > 0$, define

$$U := \{x + \lambda v : x \in M, v \in N_Q(x), |x - \bar{x}| < \epsilon, |v - \bar{v}| < \epsilon\}.$$

Then for all sufficiently small $\lambda, \epsilon > 0$, we have $\bar{x} + \lambda \bar{v} \in \operatorname{int} U$ and M admits the presentation

$$M = P_Q(U) \text{ locally around } \bar{x}.$$

Proof Using prox-regularity and Lemma 6.11, we deduce that for all sufficiently small $\lambda, \epsilon > 0$ we have

$$\left. \begin{array}{cc} x \in Q, & v \in N_Q(x) \\ |x - \bar{x}| < \epsilon, & |v - \bar{v}| < \epsilon \end{array} \right\} \Rightarrow P_Q(x + \lambda v) = x,$$

and $\bar{x} + \lambda \bar{v} \in \operatorname{int} U$. Using the fact that M is locally minimal at $\bar{x}$ for $\bar{v}$, it is easy to verify that M and $P_Q(U)$ coincide locally around $\bar{x}$.

Finally, we record the following characterization of identifiable sets.

Proposition 6.18 (Characterization of identifiable sets)
Consider a closed set $Q \subset \mathbf{R}^n$, a subset $M \subset Q$, a point $\bar{x} \in M$, and a normal vector $\bar{v} \in \hat{N}_Q(\bar{x})$. Consider the properties:

1. *M is identifiable at $\bar{x}$ for $\bar{v}$.*
2. *Q is prox-regular at $\bar{x}$ for $\bar{v}$ and for all sufficiently small $\lambda > 0$ and all $\epsilon > 0$ the inclusion*

$$\bar{x} + \lambda \bar{v} \in \operatorname{int} \{x + \lambda v : x \in M, v \in N_Q(x), |x - \bar{x}| < \epsilon, |v - \bar{v}| < \epsilon\},$$

   holds.

Then the implication (2) $\Rightarrow$ (1) holds. If M is prox-regular at $\bar{x}$ for $\bar{v}$, then we have the equivalence (1) $\Leftrightarrow$ (2).

Proof $(2) \Rightarrow (1)$: By Proposition 6.15, for all sufficiently small $\lambda > 0$ and all $\epsilon > 0$, the set

$$P_Q\Big(\{x + \lambda v : x \in M, v \in N_Q(x), |x - \bar{x}| < \epsilon, |v - \bar{v}| < \epsilon\}\Big),$$

is identifiable at $\bar{x}$ for $\bar{v}$. Furthermore, for all sufficiently small $\lambda, \epsilon > 0$ this set is contained in M locally around $\bar{x}$. Consequently M is identifiable at $\bar{x}$ for $\bar{v}$.

Now suppose that M is prox-regular at $\bar{x}$ for $\bar{v}$.

$(1) \Rightarrow (2)$: This follows trivially from Proposition 6.8 and Lemma 6.16.

7 Identifiable sets and critical cones

In this section, we consider *critical cones*, a notion that has been instrumental in sensitivity analysis, particularly in connection with polyhedral variational inequalities. See [29, Section 2E] for example. We will see that there is a strong relationship between these objects and locally minimal identifiable sets. We begin with the notion of tangency.

Definition 7.1 (Tangent cones) Consider a set $Q \subset \mathbf{R}^n$ and a point $\bar{x} \in Q$. The *tangent cone* to Q at $\bar{x}$, written $T_Q(\bar{x})$, consists of all vectors w such that

$$w = \lim_{i \to \infty} \frac{x_i - \bar{x}}{\tau_i}, \quad \text{for some } x_i \xrightarrow{Q} \bar{x}, \ \tau_i \downarrow 0.$$

The tangent cone is always closed but may easily fail to be convex. For any cone $K \in \mathbf{R}^n$, we consider the polar cone

$$K^* := \{y : \langle y, v \rangle \leq 0 \text{ for all } v \in K\}.$$

It turns out that the sets $\operatorname{cl conv} T_Q(\bar{x})$ and $\hat{N}_Q(\bar{x})$ are dual to each other, that is the equation

$$\hat{N}_Q(\bar{x}) = T_Q(\bar{x})^*,$$

holds [30, Theorem 6.28]. Consequently if Q is locally closed at $\bar{x}$, then Q is Clarke regular at $\bar{x}$ if and only if the equation $N_Q(\bar{x}) = T_Q(\bar{x})^*$ holds.

A companion notion to tangency is smooth derivability.

Definition 7.2 (smooth derivability) Consider a set Q and a point $\bar{x} \in Q$. Then a tangent vector $w \in T_Q(\bar{x})$ is *smoothly derivable* if there exists a $\mathbf{C}^1$-smooth path $\gamma \colon [0, \epsilon) \to Q$ satisfying

$$w = \lim_{t \downarrow 0} \frac{\gamma(t) - \bar{x}}{t},$$

where $\epsilon > 0$ is a real number and $\gamma(0) = \bar{x}$. We will say that Q is *smoothly derivable* at $\bar{x}$ if every tangent vector $w \in T_Q(\bar{x})$ is smoothly derivable.

We should note that there is a related weaker notion of geometric derivability, where the path γ is not required to be $\mathbf{C}^1$-smooth. For more details see [30, Definition 6.1].

Most sets that occur in practice are smoothly derivable. In particular, any smooth manifold is smoothly derivable at each of its point, as is any semi-algebraic set $Q \subset \mathbf{R}^n$. We omit the proof of the latter claim, since it is a straightforward consequence of the curve selection lemma [31, Property 4.6] and the details needed for the argument would take us far off field. For a nice survey on semi-algebraic geometry, see [8].

We now arrive at the following central notion.

Definition 7.3 (Critical cones) For a set $Q \subset \mathbf{R}^n$ that is Clarke regular at a point $\bar{x} \in Q$, the *critical cone* to Q at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$ is the set

$$K_Q(\bar{x}, \bar{v}) := N_{N_Q(\bar{x})}(\bar{v}).$$

Because of the polarity relationship between normals and tangents, the critical cone $K_Q(\bar{x}, \bar{v})$ can be equivalently described as

$$K_Q(\bar{x}, \bar{v}) = T_Q(\bar{x}) \cap \bar{v}^\perp,$$

where $\bar{v}^\perp$ is the subspace perpendicular to $\bar{v}$. For more information about critical cones and their use in variational inequalities and complementarity problems, see [11].

Connecting the classical theory of critical cones to our current work, we will now see that critical cones provide tangential approximations to locally minimal identifiable sets. In what follows, we denote the closed convex hull of any set $Q \subset \mathbf{R}^n$ by $\overline{\mathrm{co}}\, Q$.

Proposition 7.4 (Critical cones as tangential approximations)
Consider a set Q that is Clarke regular at a point $\bar{x}$ and a locally minimal identifiable set M at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$. Suppose furthermore that M is prox-regular at $\bar{x}$ for $\bar{v}$ and is smoothly derivable at $\bar{x}$. Then the equation

$$\overline{\mathrm{co}}\, T_M(\bar{x}) = K_Q(\bar{x}, \bar{v}),$$

holds.

Proof Observe

$$K_Q(\bar{x}, \bar{v}) = N_{N_Q(\bar{x})}(\bar{v}) = N_{N_M(\bar{x})}(\bar{v}) = N_{\hat{N}_M(\bar{x})}(\bar{v}) = \overline{\mathrm{co}}\, T_M(\bar{x}) \cap \bar{v}^\perp,$$

where the second equality follows from Proposition 6.10 and the last equality follows from polarity of $\mathrm{cl\,conv}\, T_M(\bar{x})$ and $\hat{N}_M(\bar{x})$. Hence to establish the claim, it is sufficient to argue that every tangent vector $w \in T_M(\bar{x})$ is orthogonal to $\bar{v}$.

To this end, fix a vector $w \in T_M(\bar{x})$ and a $\mathbf{C}^1$-smooth path $\gamma \colon [0, \epsilon) \to Q$ satisfying

$$w = \lim_{t \downarrow 0} \frac{\gamma(t) - \bar{x}}{t},$$

where $\epsilon > 0$ is a real number and $\gamma(0) = \bar{x}$.

Let $t_i \in (0, \epsilon)$ be a sequence converging to 0 and define $x_i := \gamma(t_i)$. Observe that for each index i, the tangent cone $T_M(x_i)$ contains the line $\{\lambda\dot{\gamma}(t_i) : \lambda \in \mathbf{R}\}$. Since M is necessary at $\bar{x}$ for $\bar{v}$, there exist vectors $v_i \in N_Q(\gamma(t_i))$ with $v_i \to \bar{v}$. By Proposition 6.10, we have $v_i \in \hat{N}_M(\gamma(t_i))$ for all large i. For such indices, we have $\langle v_i, \dot{\gamma}(t_i) \rangle = 0$. Letting i tend to ∞, we deduce $\langle \bar{v}, \bar{w} \rangle = 0$, as we needed to show.

Classically, the main use of critical cones has been in studying polyhedral variational inequalities. Their usefulness in that regard is due to Proposition 7.5, stated below. We provide a simple proof of this proposition that makes it evident that this result is simply a special case of Proposition 6.10. This further reinforces the theory developed in our current work. For an earlier proof that utilizes representations of polyhedral sets, see for example [29, Lemma 2E.4].

Proposition 7.5 (Polyhedral reduction) *Consider a polyhedron $Q \subset \mathbf{R}^n$ and a normal vector $\bar{v} \in N_Q(\bar{x})$, for some point $\bar{x} \in Q$. Let $K := K_Q(\bar{x}, \bar{v})$. Then we have*

$$\mathrm{gph}\, N_Q - (\bar{x}, \bar{v}) = \mathrm{gph}\, N_K \ \textit{locally around } (0,0).$$

Proof By Example 4.4, the set $M := \mathrm{argmax}_{x \in Q}\langle x, \bar{v} \rangle$ is the locally minimal identifiable set at $\bar{x}$ for $\bar{v}$. Being polyhedral, M is smoothly derivable and it satisfies

$$\bar{x} + T_M(\bar{x}) = M \text{ locally around } \bar{x}.$$

In light of Proposition 7.4, we deduce $M - \bar{x} = K$ locally around 0.

Thus for all (u, w) sufficiently near $(0, 0)$ we have

$$\begin{aligned}
\bar{v} + u \in N_Q(\bar{x} + w) &\Longleftrightarrow \bar{v} + u \in N_M(\bar{x} + w) \\
&\Longleftrightarrow \bar{v} + u \in N_K(w) \\
&\Longleftrightarrow u \in N_K(w)
\end{aligned}$$

where the first equivalence follows from Proposition 6.10, and the last equivalence follows from the fact that $K \subset \bar{v}^\perp$ and so for all $w \in K$, the cone $N_K(w)$ contains the line spanned by $\bar{v}$.

Proposition 7.5 easily fails for nonpolyhedral sets. Indeed, in light of Proposition 7.4, this is to be expected since critical cones provide only tangential approximations to locally minimal identifiable sets. Such an approximation is exact only for polyhedral sets. Hence the theory of locally minimal identifiable sets (in particular, Proposition 6.10) extends Proposition 7.5 far beyond polyhedrality.

We end this section by showing how Proposition 7.4 can be extended even further to the situation when locally minimal identifiable sets do not even exist. Indeed, consider a set Q that is Clarke regular at a point $\bar{x}$, and let $\bar{v} \in N_Q(\bar{x})$. Consider a nested sequence of open neighborhoods V_i of $\bar{v}$ satisfying

$\bigcap_{i=1}^{\infty} V_i = \{\bar{v}\}$. One would then expect that, under reasonable conditions, the equality

$$K_Q(\bar{x}, \bar{v}) = \overline{\text{co}} \bigcap_{i=1}^{\infty} T_{N_Q^{-1}(V_i)}(\bar{x}),$$

holds. To put this in perspective, observe that if there exists a locally minimal identifiable set M at $\bar{x}$ for $\bar{v}$, then the sets $T_{N_Q^{-1}(V_i)}(\bar{x})$ are equal to $T_M(\bar{x})$ for all large i, and the equation above reduces to Proposition 7.4. More generally, the following is true.

Proposition 7.6 (Critical cones more generally) *Consider a set Q that is Clarke regular at a point $\bar{x}$, and let $\bar{v} \in N_Q(\bar{x})$. Consider a nested sequence of open neighborhoods V_i of $\bar{v}$ satisfying $\bigcap_{i=1}^{\infty} V_i = \{\bar{v}\}$ and the corresponding preimages $M_i := \hat{N}_Q^{-1}(V_i)$. Assume that each M_i is smoothly derivable at $\bar{x}$. Then the inclusion*

$$K_Q(\bar{x}, \bar{v}) \supset \overline{\text{co}} \bigcap_{i=1}^{\infty} T_{M_i}(\bar{x}), \tag{7.1}$$

holds. Assume in addition that each M_i is prox-regular at $\bar{x}$ for $\bar{v}$ and that the formula

$$\overline{\text{co}} \bigcap_{i=1}^{\infty} T_{M_i}(\bar{x}) = \bigcap_{i=1}^{\infty} \overline{\text{co}}\, T_{M_i}(\bar{x}), \tag{7.2}$$

holds. Then each M_i is an identifiable set at $\bar{x}$ for $\bar{v}$ and we have

$$K_Q(\bar{x}, \bar{v}) = \overline{\text{co}} \bigcap_{i=1}^{\infty} T_{M_i}(\bar{x}).$$

We omit the proof of the proposition above since it follows along the same lines as the proof of Proposition 7.4. In particular, let us note that (7.2) holds whenever the tangent spaces $T_{M_i}(\bar{x})$ all coincide for sufficiently large indices i or whenever all M_i are Clarke regular at $\bar{x}$.

8 Optimality conditions

In this section, we will see that the order of growth of a function f around a critical point (a point satisfying $0 \in \partial f(x)$) is dictated entirely by its order of growth around this point on a corresponding identifiable set. Here is a preliminary geometric result.

Proposition 8.1 (Restricted optimality) *Consider a closed set Q and a subset $M \subset Q$ that is identifiable at $\bar{x}$ for a proximal normal $\bar{v} \in N_Q^P(\bar{x})$. Then $\bar{x}$ is a (strict) local maximizer of the linear function $\langle \bar{v}, \cdot \rangle$ on M if and only if $\bar{x}$ is a (strict) local maximizer of $\langle \bar{v}, \cdot \rangle$ on Q.*

Proof One implication is clear. To establish the converse, suppose that $\bar{x}$ is a local maximizer of the linear function $\langle \bar{v}, \cdot \rangle$ on M. We will show that the inequality, $\langle \bar{v}, \bar{x} \rangle > \langle \bar{v}, x \rangle$, holds for all points $x \in Q \setminus M$ near $\bar{x}$. Indeed, suppose this is not the case. Then there exists a sequence $x_i \to \bar{x}$ in $Q \setminus M$ satisfying

$$\langle \bar{v}, \bar{x} \rangle \leq \langle \bar{v}, x_i \rangle. \tag{8.1}$$

Since $\bar{v}$ is a proximal normal, we deduce that there exists a real number $r > 0$ satisfying $P_Q(\bar{x} + r^{-1}\bar{v}) = \{\bar{x}\}$. Consider any points z_i with

$$z_i \in P_Q(x_i + r^{-1}\bar{v}).$$

Clearly we have $z_i \to \bar{x}$ and

$$(x_i - z_i) + r^{-1}\bar{v} \in N_Q(z_i).$$

Since M is identifiable at $\bar{x}$ for $\bar{v}$, we deduce $z_i \in M$ for all large indices i. Consequently, for such indices i, we have $x_i \neq z_i$.

Observe

$$|(x_i - z_i) + r^{-1}\bar{v}| \leq r^{-1}|\bar{v}|.$$

Squaring and canceling terms, we obtain

$$\langle \bar{v}, z_i - x_i \rangle \geq \frac{r}{2}|z_i - x_i|^2. \tag{8.2}$$

Consequently,

$$\frac{r}{2}|z_i - x_i|^2 \leq \langle \bar{v}, \bar{x} - x_i \rangle \leq 0,$$

which is a contradiction. Claim (1) now follows.

Recall that a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ is said to *grow quadratically* around $\bar{x}$ provided that the inequality

$$\liminf_{x \to \bar{x}} \frac{f(x) - f(\bar{x})}{|x - \bar{x}|^2} > 0,$$

holds. We now arrive at the main result of this section.

Proposition 8.2 (Order of growth) *Consider a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a set $M \subset \mathbf{R}^n$. Suppose that M is identifiable at $\bar{x}$ for $\bar{v} = 0 \in \partial_P f(\bar{x})$. Then the following are true.*

1. *$\bar{x}$ is a (strict) local minimizer of f restricted to M $\Leftrightarrow$ $\bar{x}$ is a (strict) local minimizer of the unrestricted function f.*
2. *More generally, consider a growth function $g \colon U \to \mathbf{R}$, defined on an open neighborhood U of 0, that is $\mathbf{C}^1$-smooth and satisfies*

$$f(\bar{x}) < f(x) - g(x - \bar{x}) \text{ for all } x \in M \text{ near } \bar{x},$$
$$g(0) = 0, \ \nabla g(0) = 0,$$

Then the above inequality, in fact, holds for all points $x \in \mathbf{R}^n$ near $\bar{x}$. In particular, the function f, restricted to M, grows quadratically near $\bar{x}$ if and only if the unrestricted function f grows quadratically near $\bar{x}$.

Proof We first prove claim (1). By Proposition 3.16, gph $f\big|_M$ is identifiable, with respect to epi f, at $(\bar{x}, f(\bar{x}))$ for $(0, -1)$. Now observe that $\bar{x}$ is a (strict) local minimizer of $f\big|_M$ if and only of $(\bar{x}, f(\bar{x}))$ is a (strict) local maximizer of the linear function, $(x, r) \mapsto -r$, on gph $f\big|_M$. Similarly $\bar{x}$ is a (strict) local minimizer of f if and only if $(\bar{x}, f(\bar{x}))$ is a (strict) local maximizer of the linear function, $(x, r) \mapsto -r$, on epi f. Combining these equivalences with Proposition 8.1 establishes the claim.

We now prove claim (2). Suppose that the growth condition is satisfied. Let $h := f - g(x - \bar{x})$. Since f is $\mathbf{C}^1$-smooth, $g(0) = 0$, and $\nabla g(0) = 0$, it easily follows that M is identifiable, now with respect to h, at $\bar{x}$ for $0 \in \partial_P h(\bar{x})$. Furthermore, the point $\bar{x}$ is a strict local minimizer of $h\big|_M$. Applying claim (1) of the current proposition, we deduce that $\bar{x}$ is a strict local minimizer of the unrestricted function h, that is

$$f(\bar{x}) = h(\bar{x}) < h(x) = f(x) - g(x - \bar{x}), \text{ for all } x \text{ near } \bar{x},$$

as we needed to show.

In particular, we obtain the following curious characterization of quadratic growth.

Corollary 8.3 (Refined optimality) *Consider a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a point $\bar{x}$ with $0 \in \partial_P f(\bar{x})$. Then f grows quadratically around $\bar{x}$ if and only if*

$$\liminf_{\substack{(x, f(x), v) \to (\bar{x}, f(\bar{x}), 0) \\ v \in \partial f(x)}} \frac{f(x) - f(\bar{x})}{|x - \bar{x}|^2} > 0. \tag{8.3}$$

Proof Clearly if f grows quadratically around $\bar{x}$, then (8.3) holds. Conversely, assume (8.3) holds and let V_i be a sequence of neighborhoods of 0 shrinking to 0 and let $\epsilon_i > 0$ be real number tending to 0. Then the sets

$$M_i := (\partial f)^{-1}(V_i) \cap \{x \in \mathbf{R}^n : |f(x) - f(\bar{x})| < \epsilon_i\},$$

are identifiable at $\bar{x}$ for 0. Furthermore, f restricted to M_i must grow quadratically around $\bar{x}$, for all sufficiently large indices i, since the alternative would contradict (8.3). Applying Proposition 8.2, we obtain the result.

9 Identifiable Manifolds

Consider a closed set Q and a normal vector $\bar{v} \in \hat{N}_Q(\bar{x})$, for a point $\bar{x} \in Q$. The inherent difficulty in analyzing properties of the optimization problem,

$$P(v) : \quad \max \quad \langle v, x \rangle,$$
$$\text{s.t.} \quad x \in Q,$$

such as dependence of the local maximizers of $P(v)$ on v or the order of growth of the function $x \mapsto \langle x, \bar{v} \rangle$ on Q near $\bar{x}$, stem entirely from the potential

nonsmoothness of Q. However, as we have seen in Proposition 6.10, the local geometry of gph N_Q is entirely the same as that of a prox-regular identifiable set M at $\bar{x}$ for $\bar{v}$. Thus, for instance, existence of an *identifiable manifold* M at $\bar{x}$ for $\bar{v}$ shows that the nonsmoothness of Q is not intrinsic to the problem at all. Our goal in this section is to investigate this setting. We begin with the following easy consequence of Proposition 6.10.

Proposition 9.1 *Consider a closed set $Q \subset \mathbf{R}^n$ and suppose that a subset $M \subset Q$ is a $\mathbf{C}^2$ identifiable manifold at $\bar{x}$ for $\bar{v} \in \hat{N}_Q(\bar{x})$. Then the following properties hold.*

1. *$\bar{v}$ lies in the interior of the cone $N_Q^P(\bar{x})$, relative to its linear span $N_M(\bar{x})$.*
2. *There exists an open neighborhood U of $\bar{x}$ and V of $\bar{v}$ such that the mapping $x \mapsto V \cap N_Q(x)$, restricted to M, is inner-semicontinuous at each $x \in U \cap M$.*

Proof To see the validity of the first claim, observe that if it did not hold, then we could choose a sequence of vectors v_i satisfying

$$v_i \to \bar{v}, \quad v_i \in N_M(\bar{x}), \quad v_i \notin N_Q(\bar{x}),$$

thus contradicting Proposition 6.10.

The second claim now easily follows from Proposition 6.10.

Consider a locally minimal identifiable subset $M \subset Q$ at $\bar{x}$ for $\bar{v} \in N_Q(\bar{x})$. Then M remains identifiable at x for $v \in N_Q(x)$, whenever the pair (x, v) is sufficiently close to $(\bar{x}, \bar{v})$. However under such perturbations, M might cease to be locally minimal, as one can see even from polyhedral examples. (Indeed when Q is a convex polyhedron, this instability occurs whenever the inclusion $\bar{v} \in \mathrm{rb}\, N_Q(\bar{x})$ holds.)

In the case of identifiable manifolds, the situation simplifies. Identifiable manifolds at $\bar{x}$ for $\bar{v} \in \hat{N}_Q(\bar{x})$ are automatically locally minimal, and furthermore they remain locally minimal under small perturbations to $(\bar{x}, \bar{v})$ in gph N_Q.

This important observation is summarized below.

Proposition 9.2 *Consider a closed set Q and a $\mathbf{C}^2$ identifiable manifold $M \subset Q$ at $\bar{x}$ for $\bar{v} \in \hat{N}_Q(\bar{x})$. Then M is automatically a locally minimal identifiable set at $x \in M$ for $v \in N_Q(x)$ whenever the pair (x, v) is near $(\bar{x}, \bar{v})$.*

Proof This follows directly from Proposition 2.4 and Proposition 9.1.

In particular, identifiable manifolds at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ are locally unique.

9.1 Relation to Partial Smoothness

In this section, we will relate identifiable manifolds to the notion of partial smoothness, introduced in [19]. The motivation behind partial smoothness

is two-fold. On one hand, it is an attempt to model an intuitive idea of a "stable active set". On the other hand, partial smoothness, along with certain nondegeneracy and growth conditions, provides checkable sufficient conditions for optimization problems to possess good sensitivity properties. Evidently, partial smoothness imposes conditions that are unnecessarily strong. We now describe a variant of partial smoothness that is localized in a directional sense. This subtle distinction, however, will be important for us.

Definition 9.3 (Directional Partial Smoothness) Consider a closed set $Q \subset \mathbf{R}^n$ and a $\mathbf{C}^2$-manifold $M \subset Q$. Then Q is *partly smooth with respect to M at $\bar{x} \in M$ for $\bar{v} \in N_Q(\bar{x})$* if

1. **(prox-regularity)** Q is prox-regular at $\bar{x}$ for $\bar{v}$.
2. **(sharpness)** span $\hat{N}_Q(\bar{x}) = N_M(\bar{x})$.
3. **(continuity)** There exists a neighborhood V of $\bar{v}$, such that the mapping, $x \mapsto V \cap N_Q(x)$, when restricted to M, is inner-semicontinuous at $\bar{x}$.

We arrive at the main result of this subsection.

Proposition 9.4 (Identifiable manifolds and partial smoothness) *Consider a closed set $Q \subset \mathbf{R}^n$ and a subset $M \subset Q$ that is a $\mathbf{C}^2$ manifold around a point $\bar{x} \in Q$. Let $\bar{v} \in \hat{N}_Q(\bar{x})$. Then the following are equivalent.*

1. *M is an identifiable manifold at $\bar{x}$ for $\bar{v}$.*
2. *We have*
$$\operatorname{gph} N_Q = \operatorname{gph} N_M \quad \text{locally around } (\bar{x}, \bar{v}).$$
3. *– Q is partly smooth with respect to M at $\bar{x}$ for $\bar{v}$.*
   – the strong inclusion $\bar{v} \in \operatorname{ri} \hat{N}_Q(\bar{x})$ holds.
4. *The set Q is prox-regular at $\bar{x}$ for $\bar{v}$, and for all sufficiently small real numbers $\lambda, \epsilon > 0$, the inclusion*
$$\bar{x} + \lambda \bar{v} \in \operatorname{int} \Big(\bigcup_{x \in M \cap B_\epsilon(\bar{x})} \big(x + N_Q(x)\big) \Big),$$

holds.

Proof The equivalence (1) $\Leftrightarrow$ (2) has been established in Proposition 6.10. The implication (1) $\Rightarrow$ (3) follows trivially from Propositions 6.8 and 9.1.

(3) $\Rightarrow$ (4): There exist real numbers $r, \epsilon > 0$ so that the implication

$$\left. \begin{array}{c} x \in Q, \quad v \in N_Q(x) \\ |x - \bar{x}| < \epsilon, \ |v - \bar{v}| < \epsilon \end{array} \right\} \Rightarrow P_{Q \cap B_\epsilon(\bar{x})}(x + r^{-1}v) = x,$$

holds.

For the sake of contradiction, suppose

$$\bar{x} + r^{-1}\bar{v} \in \operatorname{bd} \Big(\bigcup_{x \in M \cap B_\epsilon(\bar{x})} \big(x + N_Q(x)\big) \Big).$$

Then there exists a sequence of points $z_i \to \bar{x} + r^{-1}\bar{v}$ with

$$z_i \notin \bigcup_{x \in M \cap B_\epsilon(\bar{x})} \left(x + N_Q(x) \right),$$

for each index i. Let $x_i \in P_M(z_i)$. Observe

$$x_i \to \bar{x}, \quad z_i - x_i \in N_M(x_i).$$

Clearly,

$$z_i - x_i \notin N_Q^P(x_i) \subset N_M(x_i),$$

for large indices i. Hence, there exist separating vectors $a_i \in N_M(x_i)$ with $|a_i| = 1$ satisfying

$$\sup_{v \in N_Q^P(x_i)} \langle a_i, v \rangle \le \langle a_i, z_i - x_i \rangle = \langle a_i, z_i - \bar{x} \rangle + \langle a_i, \bar{x} - x_i \rangle.$$

We deduce,

$$\sup_{v \in N_Q^P(x_i)} \langle a_i, v \rangle \le \langle a_i, r(z_i - \bar{x}) \rangle + \langle a_i, r(\bar{x} - x_i) \rangle.$$

Passing to a subsequence, we may assume $a_i \to a$ for some nonzero vector $a \in N_M(\bar{x})$. Observe $\bar{v} + \delta a \in N_Q(x_i)$ for all small $\delta > 0$. Consequently for all sufficiently small $\delta > 0$, there exist vectors $v_i \in N_Q(x_i)$ with $v_i \to \bar{v} + \delta a$. Observe

$$\langle a_i, v_i \rangle \le \langle a_i, r(z_i - \bar{x}) \rangle + \langle a_i, r(\bar{x} - x_i) \rangle.$$

Letting i tend to ∞, we obtain

$$\langle a, \bar{v} + \delta a \rangle \le \langle a, \bar{v} \rangle,$$

which is a contradiction.

(4) $\Rightarrow$ (1): Choose $r, \epsilon > 0$ so as to ensure

$$\bar{x} + r^{-1}\bar{v} \in \mathrm{int} \left(\bigcup_{x \in M \cap B_\epsilon(\bar{x})} \left(x + N_Q(x) \right) \right).$$

Consider any sequence of points $x_i \in \mathbf{R}^n$ and vectors $v_i \in N_Q(x_i)$, with $x_i \to \bar{x}$ and $v_i \to \bar{v}$. Then for all large indices i, the inclusion

$$x_i + r^{-1}v_i \in \bigcup_{x \in M \cap B_\epsilon(\bar{x})} \left(x + N_Q(x) \right),$$

holds. Shrinking r and ϵ, from prox-regularity of Q, we deduce $x_i \in M$ for all large indices i. Hence M is identifiable at $\bar{x}$ for $\bar{v}$.

Some comments concerning characterization (4) of the previous proposition are in order. Consider a convex set Q containing a point $\bar{x}$, and let $\bar{v} \in N_Q(\bar{x})$ be a normal vector. Then the arguments $(3) \Rightarrow (4)$ and $(4) \Rightarrow (1)$ show that a manifold $M \subset Q$ is identifiable at $\bar{x}$ for $\bar{v}$ if and only if the inclusion

$$\bar{x} + \bar{v} \in \text{int} \Big(\bigcup_{x \in M} \big(x + N_Q(x) \big) \Big),$$

holds. The region $\bigcup_{x \in M} \big(x + N_Q(x) \big)$ is formed by attaching cones $N_Q(x)$ to each point $x \in M$. This set is precisely the set of points in $\mathbf{R}^n$ whose projections onto Q lie in M. Thus a manifold M is identifiable at $\bar{x}$ for $\bar{v}$ whenever the region $\bigcup_{x \in M} \big(x + N_Q(x) \big)$ is "valley-like" around $\bar{x} + \bar{v}$. We end the section with an observation relating identifiable manifolds to critical cones.

Corollary 9.5 *Consider a closed set $Q \subset \mathbf{R}^n$ that is Clarke regular at a point $\bar{x} \in Q$. Suppose that $M \subset Q$ is a $\mathbf{C}^2$ identifiable manifold at $\bar{x}$ for $\bar{v}$. Then the critical cone $K_Q(\bar{x}, \bar{v})$ coincides with the tangent space $T_M(\bar{x})$.*

Proof This is an immediate consequence of Proposition 7.4.

9.2 Projections onto identifiable manifolds

In this subsection, we show that $\mathbf{C}^2$ identifiable manifolds occur precisely as the images of $\mathbf{C}^1$-smooth projections with constant rank. Predecessor results of a similar flavour can be found in [17] and [32]. We will need the following classical result.

Proposition 9.6 *Consider a $\mathbf{C}^2$-smooth manifold $M \subset \mathbf{R}^n$ and a point $\bar{x} \in M$. Then there exists a neighborhood U of $\bar{x}$ so that the projection mapping P_M is a $\mathbf{C}^1$-smooth submersion on U. In particular it sends open subsets of U to open subsets of M.*

Proof Let V be a neighborhood of $\bar{x}$ so that

$$M \cap V = \{x \in V : h_i(x) = 0 \text{ for } i = 1, \ldots, m\},$$

where $h_i \colon V \to \mathbf{R}$ are $\mathbf{C}^2$-smooth defining functions and the gradients $\nabla h_i(\bar{x})$ are linearly independent. Since M is prox-regular at $\bar{x}$, the mapping $y \mapsto P_M(y)$ is single-valued near $\bar{x}$. Furthermore, the point $P_M(y)$ is precisely the point x solving the system

$$h_i(x) = 0, \text{ for } i = 1, \ldots, m,$$
$$x + \sum_{i=1}^{m} \lambda_i \nabla h_i(x) = y.$$

for some $\lambda \in \mathbf{R}^m$. Applying the inverse function theorem to the system above yields the result.

Proposition 9.7 (Smooth projections) *Consider a closed set $Q \subset \mathbf{R}^n$ and a $\mathbf{C}^2$ manifold $M \subset Q$ containing a point $\bar{x}$. Let $\bar{v} \in \hat{N}_Q(\bar{x})$ be a normal vector. Then the following are equivalent.*

1. *M is an identifiable manifold around $\bar{x}$ for $\bar{v}$.*
2. *The set Q is prox-regular at $\bar{x}$ for $\bar{v}$, and for all sufficiently small real numbers λ and all sufficiently small neighborhoods U of $\bar{x} + \lambda\bar{v}$,*
   - *The projection mapping P_Q is $\mathbf{C}^1$-smooth and has constant rank on the ball U.*
   - *The image $P_Q(U)$ coincides with M near $\bar{x}$.*

Proof (1) $\Rightarrow$ (2): This follows immediately from Propositions 6.8, 9.6, and 6.17.

(2) $\Rightarrow$ (1): This is immediate from Proposition 6.15.

10 The case of functions

Many of the results we have derived in the previous sections were stated in terms of sets. Their generalizations to the functional setting can be carried out in a standard way by appealing to Proposition 3.16. In this section, we record some of these results for ease of reference in future work. Most of the proofs are omitted.

Proposition 10.1 (Prox-regularity of identifiable sets) *Consider a lsc function $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a set $M \subset \mathbf{R}^n$ that is identifiable at $\bar{x}$ for $\bar{v} \in \hat{\partial}f(\bar{x})$. In addition, suppose that $f + \delta_M$ is prox-regular at $\bar{x}$ for $\bar{v}$. Then f is prox-regular at $\bar{x}$ for $\bar{v}$.*

In stating the next results, it will be convenient to define "subjets" associated to each subdifferential. Namely, for a function $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$, we define the *limiting subjet* to be the set

$$[\partial f] := \{(x, f(x), v) \in \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^n : v \in \partial f(x)\}.$$

Subjets corresponding to the other subdifferentials are defined analogously. These objects are useful for concisely describing variational properties of functions that are not necessarily subdifferentially continuous. On the other hand if f is subdifferentially continuous, as is often the case, then in the results that follow we may simply use subdifferential graphs, instead of subjets.

Proposition 10.2 (Reduction I) *Consider a lsc function $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a set $M \subset \mathbf{R}^n$ so that $f + \delta_M$ is prox-regular at a point $\bar{x}$ for $\bar{v} \in \hat{\partial}f(\bar{x})$. Then M is identifiable at $\bar{x}$ for $\bar{v}$ if and only if*

$$[\partial f] = [\partial(f + \delta_M)] \quad \text{locally around } (\bar{x}, f(\bar{x}), \bar{v}).$$

Proposition 10.3 (Reduction II) *Consider a lsc function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$, a point $\bar{x} \in \mathbf{R}^n$, and a subgradient $\bar{v} \in \partial f(\bar{x})$. Suppose*

$$[\partial^P f] = [\partial f] \ \text{locally around} \ (\bar{x}, f(\bar{x}), \bar{v}),$$

and consider a set $M := (\partial f)^{-1}(V) \cap \{x : |f(x) - f(\bar{x})| < \epsilon\}$, where V is a convex, open neighborhood of $\bar{v}$ and $\epsilon > 0$ is a real number. Then the equation

$$\mathrm{gph}\,[\partial f] = \mathrm{gph}\,[\partial(f + \delta_M)] \ \text{holds locally around} \ (\bar{x}, f(\bar{x}), \bar{v}).$$

Proposition 10.4 (Prox-regularity and local minimality)
Consider a lsc function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a set $M \subset \mathbf{R}^n$ that is a locally minimal identifiable set at $\bar{x}$ for $\bar{v} \in \hat{\partial} f(\bar{x})$. Then f is prox-regular at $\bar{x}$ for $\bar{v}$ if and only if $f + \delta_M$ is prox-regular at $\bar{x}$ for $\bar{v}$.

The following is a standard generalization of critical cones to the functional setting.

Definition 10.5 (Critical cones) For a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ that is Clarke regular at a point $\bar{x}$, the *critical cone* of f at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ is the set

$$K_f(\bar{x}, \bar{v}) := N_{\partial f(\bar{x})}(\bar{v}).$$

Definition 10.6 (Smooth derivability) A function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ is said to be *smoothly derivable* at a point $\bar{x} \in \mathbf{R}^n$ if $\mathrm{gph}\,f$ is smoothly derivable at $(\bar{x}, f(\bar{x})$.

Proposition 10.7 (Critical cones more generally) *Consider a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ that is Clarke regular at a point $\bar{x}$, and let $\bar{v} \in \partial f(\bar{x})$. Consider a nested sequence of open neighborhoods V_i of $\bar{v}$ satisfying $\bigcap_{i=1}^{\infty} V_i = \{\bar{v}\}$ and a sequence of reals $\epsilon_i \downarrow 0$, and the corresponding preimages $M_i := \partial f^{-1}(V_i) \cap \{x : |f(x) - f(\bar{x})| < \epsilon\}$. Assume that each $f + \delta_{M_i}$ is smoothly derivable at $\bar{x}$. Then the inclusion*

$$K_f(\bar{x}, \bar{v}) \supset \overline{\mathrm{co}} \bigcap_{i=1}^{\infty} T_{M_i}(\bar{x}), \tag{10.1}$$

holds. Assume in addition that each $f + \delta_{M_i}$ is prox-regular at $\bar{x}$ for $\bar{v}$ and that the formula

$$\overline{\mathrm{co}} \bigcap_{i=1}^{\infty} T_{M_i}(\bar{x}) = \bigcap_{i=1}^{\infty} \overline{\mathrm{co}}\, T_{M_i}(\bar{x}), \tag{10.2}$$

holds. Then each M_i is an identifiable set at $\bar{x}$ for $\bar{v}$ and we have

$$K_f(\bar{x}, \bar{v}) = \overline{\mathrm{co}} \bigcap_{i=1}^{\infty} T_{M_i}(\bar{x}).$$

In particular, the following is true.

Proposition 10.8 (Critical cones as tangential approximations)
Consider a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ that is Clarke regular at a point $\bar{x}$, and let M be a locally minimal identifiable set at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$. Suppose furthermore that $f + \delta_M$ is prox-regular at $\bar{x}$ for $\bar{v}$ and is smoothly derivable at $\bar{x}$. Then the equation

$$\overline{\mathrm{co}}\, T_M(\bar{x}) = K_f(\bar{x}, \bar{v}),$$

holds.

Definition 10.9 (Identifiable manifolds) Given a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$, a set $M \subset \mathbf{R}^n$ is a $\mathbf{C}^p$ *identifiable manifold* at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ provided that the following hold.

- M is a $\mathbf{C}^p$ manifold around $\bar{x}$ and the restriction $f\big|_M$ is $\mathbf{C}^p$-smooth near $\bar{x}$.
- M is identifiable at $\bar{x}$ for $\bar{v}$.

Proposition 10.10 *Consider a lsc function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a $\mathbf{C}^2$ identifiable manifold $M \subset \mathrm{dom}\, f$ at $\bar{x}$ for $\bar{v} \in \hat{\partial} f(\bar{x})$. Then M is automatically a locally minimal identifiable set at $x \in M$ for $v \in N_Q(x)$, whenever the triple $(x, f(x), v)$ is near $(\bar{x}, f(\bar{x}), \bar{v})$.*

A functional version of partial smoothness is as follows. For a convex sets $K \subset \mathbf{R}^n$, we let $\mathrm{par}\, K$ be parallel subspace to K, that is $\mathrm{par}\, K = \mathrm{span}\,\{K - x\}$, where x is any point of K.

Definition 10.11 (Directional partial smoothness) Consider a lsc function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a $\mathbf{C}^2$-manifold $M \subset \mathrm{dom}\, f$. Then f is *partly smooth with respect to M at $\bar{x} \in M$ for $\bar{v} \in \partial f(\bar{x})$* if

1. **(prox-regularity)** f is prox-regular at $\bar{x}$ for $\bar{v}$.
2. **(smoothness)** The restriction $f\big|_M$ is $\mathbf{C}^2$ smooth around $\bar{x}$.
3. **(sharpness)** $\mathrm{par}\, \hat{\partial} f = N_M(\bar{x})$.
4. **(continuity)** There exists a neighborhood V of $\bar{v}$, such that the mapping, $x \mapsto V \cap \partial f(x)$, when restricted to M, is inner-semicontinuous at $\bar{x}$.

Proposition 10.12 (Identifiable manifolds and partial smoothness)
Consider a lsc function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a set $M \subset \mathbf{R}^n$, so that M is a $\mathbf{C}^2$ manifold around $\bar{x}$ and f is $\mathbf{C}^2$-smooth on M around $\bar{x}$. Consider a subgradient $\bar{v} \in \hat{\partial} f(\bar{x})$. Then the following are equivalent.

1. *M is an identifiable manifold at $\bar{x}$ for $\bar{v}$.*
2. *We have*

$$[\partial f] = [\partial(f + \delta_M)] \quad \text{locally around } (\bar{x}, f(\bar{x}), \bar{v}).$$

3. *– f is partly smooth with respect to M at $\bar{x}$ for $\bar{v}$.*
   – the strong inclusion $\bar{v} \in \mathrm{ri}\, \hat{\partial} f(\bar{x})$ holds.

Given a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a parameter $\lambda > 0$, the *proximal mapping* $P_\lambda f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ is defined by

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

The results of Section 6, concerning projections, may be extended to the functional setting with the proximal mapping taking the place of the metric projection. However for the sake of brevity, we do not pursue this further.

References

1. F. Al-Khayyal and J. Kyparisis. Finite convergence of algorithms for nonlinear programs and variational inequalities. *J. Optim. Theory Appl.*, 70(2):319–332, 1991.
2. J.F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems.* Springer, New York, 2000.
3. J. M. Borwein and Q. J. Zhu. *Techniques of Variational Analysis.* Springer Verlag, New York, 2005.
4. J. V. Burke. On the identification of active constraints. II. The nonconvex case. *SIAM J. Numer. Anal.*, 27(4):1081–1103, 1990.
5. J. V. Burke and J. J. Moré. On the identification of active constraints. *SIAM J. Numer. Anal.*, 25(5):1197–1211, 1988.
6. P. H. Calamai and J. J. Moré. Projected gradient methods for linearly constrained problems. *Math. Programming*, 39(1):93–116, 1987.
7. F. H. Clarke, Yu. Ledyaev, R. I. Stern, and P. R. Wolenski. *Nonsmooth Analysis and Control Theory.* Texts in Math. 178, Springer, New York, 1998.
8. M. Coste. *An Introduction to Semialgebraic Geometry.* RAAG Notes, 78 pages, Institut de Recherche Mathématiques de Rennes, October 2002.
9. D. Drusvyatskiy and A. S. Lewis. Spectral lifts of identifiable sets. *in preparation*, 2012.
10. J. C. Dunn. On the convergence of projected gradient processes to singular critical points. *J. Optim. Theory Appl.*, 55(2):203–216, 1987.
11. F. Facchinei and J. Pang. *Finite dimensional variational inequalities and complementarity problems.* Springer Series in Operations Research, Springer-Verlag, New York, 2003.
12. H. Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93(3):pp. 418–491, 1959.
13. M. C. Ferris. Finite termination of the proximal point algorithm. *Math. Programming*, 50(3, (Ser. A)):359–366, 1991.
14. S. D. Flåm. On finite convergence and constraint identification of subgradient projection methods. *Mathematical Programming*, 57:427–437, 1992.
15. W. Hare and C. Sagastizábal. Benchmark of some nonsmooth optimization solvers for computing nonconvex proximal points. *Pac. J. Optim.*, 2(3):545–573, 2006.
16. W. Hare and C. Sagastizábal. Computing proximal points of nonconvex functions. *Math. Program.*, 116(1-2, Ser. B):221–258, 2009.
17. W. L. Hare and A. S. Lewis. Identifying active constraints via partial smoothness and prox-regularity. *J. Convex Anal.*, 11(2):251–266, 2004.
18. J. M. Lee. *Introduction to Smooth Manifolds.* Springer, New York, 2003.
19. A. S. Lewis. Active sets, nonsmoothness, and sensitivity. *SIAM J. on Optimization*, 13:702–725, August 2002.
20. R. Mifflin and C. Sagastizábal. Primal-dual gradient structured functions: second-order results; links to epi-derivatives and partly smooth functions. *SIAM J. Optim.*, 13(4):1174–1194 (electronic), 2003.

21. R. Mifflin and C. Sagastizábal. $\mathcal{VU}$-smoothness and proximal point results for some nonconvex functions. *Optim. Methods Softw.*, 19(5):463–478, 2004.
22. R. Mifflin and C. Sagastizábal. A $\mathcal{VU}$-algorithm for convex minimization. *Math. Program.*, 104(2-3, Ser. B):583–608, 2005.
23. B. S. Mordukhovich. *Variational Analysis and Generalized Differentiation I: Basic Theory.* Grundlehren der mathematischen Wissenschaften, Vol 330, Springer, Berlin, 2005.
24. B. S. Mordukhovich. *Variational Analysis and Generalized Differentiation II: Applications.* Grundlehren der mathematischen Wissenschaften, Vol 331, Springer, Berlin, 2005.
25. R. A. Poliquin and R. T. Rockafellar. Amenable functions in optimization. In *Nonsmooth optimization: methods and applications (Erice, 1991)*, pages 338–353. Gordon and Breach, Montreux, 1992.
26. R. A. Poliquin and R. T. Rockafellar. Prox-regular functions in variational analysis. *Trans. Amer. Math. Soc.*, 348:1805–1838, 1996.
27. R. A. Poliquin, R. T. Rockafellar, and L. Thibault. Local differentiability of distance functions. *Transactions of the American Mathematical Society*, 352(11):pp. 5231–5249, 2000.
28. S. M. Robinson. Some continuity properties of polyhedral multifunctions. In *Mathematical Programming at Oberwolfach*, volume 14 of *Mathematical Programming Studies*, pages 206–214. Springer Berlin Heidelberg, 1981. 10.1007/BFb0120929.
29. R. T. Rockafellar and A. L. Dontchev. *Implicit functions and solution mappings*. Monographs in Mathematics, Springer-Verlag, 2009.
30. R. T. Rockafellar and R. J-B. Wets. *Variational Analysis*. Grundlehren der mathematischen Wissenschaften, Vol 317, Springer, Berlin, 1998.
31. L. van den Dries and C. Miller. Geometric categories and o-minimal structures. *Duke Mathematical Journal*, 84:497–540, 1996.
32. S. J. Wright. Identifiable surfaces in constrained optimization. *SIAM J. Control Optim.*, 31:1063–1079, July 1993.