Nonlinear Analysis 75 (2012) 1231–1245

Contents lists available at SciVerse ScienceDirect

Nonlinear Analysis

journal homepage: www.elsevier.com/locate/na

The dimension of semialgebraic subdifferential graphs

D. Drusvyatskiy [a,*], A.D. Ioffe [b], A.S. Lewis [a]

[a] School of Operations Research and Information Engineering, Cornell University, Ithaca, NY, USA
[b] Department of Mathematics, Technion-Israel Institute of Technology, Haifa, 32000, Israel

ARTICLE INFO

Communicated by Ravi Agarwal

Keywords:
Set-valued map
Subdifferential
Semi-algebraic
Stratification
Dimension

ABSTRACT

A corollary of a celebrated theorem of Minty is that the subdifferential graph of a closed convex function on $\mathbf{R}^n$ has uniform local dimension n. In contrast, there exist nonconvex closed functions whose subdifferentials have large graphs. We consider how far Minty's corollary extends to functions that are nonconvex but semi-algebraic.

1. Introduction

This paper deals with three standard subdifferentials. We begin by fixing some notation. The functions that we will be considering will be allowed to take values in the extended real line $\overline{\mathbf{R}} := \mathbf{R} \cup \{-\infty\} \cup \{+\infty\}$. 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 point $\bar{x} \in \mathbf{R}^n$, we let $o(|x-\bar{x}|)$ be shorthand for a function that satisfies $\frac{o(|x-\bar{x}|)}{|x-\bar{x}|} \to 0$ whenever $x \to \bar{x}$ with $x \neq \bar{x}$.

Consider a function $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ and a point $\bar{x}$ with $f(\bar{x})$ finite. The set $\hat{\partial} f(\bar{x})$ consists of all vectors $v \in \mathbf{R}^n$ satisfying

$$f(x) \geq f(\bar{x}) + \langle v, x-\bar{x}\rangle + o(|x-\bar{x}|),$$

and $\partial f(\bar{x})$ is the set of all vectors $v \in \mathbf{R}^n$ such that there are sequences $x_i \in \mathbf{R}^n$ and $v_i \in \hat{\partial} f(x_i)$, with $(x_i, f(x_i), v_i) \to (\bar{x}, f(\bar{x}), \bar{v})$.

Now assuming (for technical reasons) that f is locally Lipschitz continuous around $\bar{x}$, we consider

$$\overline{\partial} f(\bar{x}) = \operatorname{conv} \partial f(\bar{x})$$

(a set that can be easier to approximate numerically—see for example [1]).

For x such that $f(x)$ is not finite, we follow the convention that $\hat{\partial} f(x) = \partial f(x) = \overline{\partial} f(x) = \emptyset$. The sets $\hat{\partial} f(x)$, $\partial f(x)$, $\overline{\partial} f(x)$ are known as *subdifferentials*.

The three subdifferentials above appear under a variety of names in standard recent books on the subject. The subdifferential $\hat{\partial} f(x)$ appears under the name of *Fréchet subdifferential* in Borwein–Zhu [2, page 39] and Mordukhovich [3, page 90], while in Rockafellar–Wets [4, page 301], its elements are called *regular subgradients*. Other names include the *viscosity subdifferential* [2, page 39] and the *presubdifferential* [3, page 90]. By contrast, elements of $\partial f(x)$ are called *(general) subgradients* in Rockafellar–Wets [4, page 301]; the set is called the *limiting subdifferential* in Borwein–Zhu [2, page 196], and in

* Correspondence to: 206 Rhodes Hall, Cornell University, Ithaca, NY 14853, USA. Tel.: +1 718 865 6367; fax: +1 607 255 9129.
E-mail addresses: dd379@cornell.edu (D. Drusvyatskiy), ioffe@math.technion.ac.il (A.D. Ioffe).
URL: http://people.orie.cornell.edu/~aslewis/ (A.S. Lewis).

0362-546X/$ – see front matter
doi:10.1016/j.na.2011.07.040

Clarke–Ledyaev–Stern-Wolenski [5, page 61], while Mordukhovich [3, page 83] refers to it as the *(basic, limiting) subdifferential*. Finally, for locally Lipschitz f, the set $\overline{\partial} f(x)$ is called the *Clarke subdifferential* in Borwein–Zhu [2, page 189] and Mordukhovich [3, page 236], while in Clarke–Ledyaev–Stern–Wolenski [5, page 72] and Mordukhovich [3, page 236] it is referred to as the *generalized gradient*; in Rockafellar–Wets [4] elements of this set are called *Clarke subgradients* [4, page 336], or *Clarke convexified subgradients* [4, page 733] (corresponding to the terminology "convexified normal cone" [4, page 225]—see Definition 2.3). Historical commentaries on these subdifferentials may be found in the books by Rockafellar–Wets [4, pages 234–235 and 344–347], Mordukhovich [3, pages 132–155], and Borwein–Zhu [2, pages 43 and 207].

A principle goal of variational analysis and nonsmooth optimization (and of critical point theory) is to study generalized critical points of extended-real-valued functions f on $\mathbf{R}^n$. These are the points x where a generalized subdifferential, such as $\hat{\partial} f(x)$, $\partial f(x)$, or $\overline{\partial} f(x)$, contains the zero vector. Generalized critical points of smooth functions are, in particular, critical points in the classical sense, while critical points of convex functions are simply their minimizers. More generally, one could consider the perturbed function $x \mapsto f(x) - \langle v, x\rangle$, for some fixed vector $v \in \mathbf{R}^n$. Then a point x is critical precisely when the pair (x, v) lies in the graph of the subdifferential mapping. Hence, it is natural to try to understand geometric properties of subdifferential graphs.

In particular, an interesting question in this area is to understand the "size" of the subdifferential graph. For instance, for a smooth function defined on $\mathbf{R}^n$, the graph of the subdifferential mapping is an n-dimensional surface. Minty [6] famously showed that the subdifferential graph of a lower semicontinuous, convex function defined on $\mathbf{R}^n$ is Lipschitz homeomorphic to $\mathbf{R}^n$. In fact, he provided explicit Lipschitz homeomorphisms that are very simple in nature. More generally in [7], Poliquin and Rockafellar used Minty's theorem to show that an analogous result holds for "prox-regular functions", unifying the smooth and the convex cases. Hence, we would expect that for a nonpathological function, the subdifferential graph should have the same dimension, in some sense, as the space that the function is defined on. A limiting feature of Poliquin's and Rockafellar's approach is that their arguments rely on convexity, or rather the related notion of maximal monotonicity. Hence their techniques do not seem to extend to a larger class of functions.

From a practical point of view, the size of the subdifferential graph may have important algorithmic applications. For instance, Robinson [8] shows computational promise for functions defined on $\mathbf{R}^n$ whose subdifferential graphs are locally homeomorphic to an open subset of $\mathbf{R}^n$. In particular, due to Minty's result, Robinson's techniques are applicable for lower semicontinuous, convex functions. When can we then be sure that the dimension of the subdifferential graph is the same as the dimension of the domain space?

It is well-known that for general functions, even ones that are Lipschitz continuous, the subdifferential graph can be very large. For instance, there is a 1-Lipschitz function $f : \mathbf{R} \to \mathbf{R}$, such that the subdifferential $\overline{\partial} f(x)$ is the unit interval $[-1, 1]$ at every point x. Furthermore, this behavior is typical [9] and such pathologies are not particular to $\overline{\partial} f$ [10,11].

These pathological functions, however, do not normally appear in practice. As a result, the authors of [12] were led to consider *semi-algebraic* functions, those functions whose graphs are defined by finitely many polynomial equalities and inequalities. They showed that for a proper, semi-algebraic function on $\mathbf{R}^n$, any reasonable subdifferential has a graph that is, in a precise mathematical sense, exactly n-dimensional. The authors derived a variety of applications for generic semi-algebraic optimization problems.

The dimension of a semi-algebraic set, as discussed in [12], is a global property governed by the maximal size of any part of this set. In particular, the result above does not rule out that some parts of the subdifferential graph may be small. In fact, in the case of the subdifferential mapping $\overline{\partial} f$ this can happen! It is the aim of our current work to elaborate on this phenomenon and to show that it does not occur in the case of ∂f. Specifically, we will show that for a lower semicontinuous, semi-algebraic function f on $\mathbf{R}^n$, the graph of the subdifferential ∂f has local dimension n, uniformly over the whole set. Surprisingly, as we noted, this type of result does not hold for $\overline{\partial} f$. That is, even for the simplest of examples, the graph of the subdifferential $\overline{\partial} f$ may be small in some places, despite being a larger set than the graph of ∂f.

To be concrete, we state our results for semi-algebraic functions. Analogous results, with essentially identical proofs, hold for functions definable in an "o-minimal structure" and, more generally, for "tame" functions. In particular, our results hold for globally subanalytic functions, discussed in [13]. For a quick introduction to these concepts in an optimization context, see [14].

2. Preliminaries

2.1. Variational analysis

In this section, we summarize some of the fundamental tools used in variational analysis and nonsmooth optimization. We refer the reader to the monographs Borwein–Zhu [2], Mordukhovich [3], Clarke–Ledyaev–Stern–Wolenski [5], and Rockafellar–Wets [4], for more details.

We say that an extended-real-valued function is proper if it is never $-\infty$ and is not always $+\infty$. For a function $f : \mathbf{R}^n \to \overline{\mathbf{R}}$, we define the *domain* of f to be

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

and we define the *epigraph* of f to be the set

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

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

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

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

We now define "subjets" corresponding to each subdifferential.

Definition 2.1. Consider a function $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$. We define

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

We make analogous definitions for $[\partial f]$ and for $[\overline{\partial} f]$ when f is locally Lipschitz continuous.

The following is a standard result in subdifferential calculus.

Proposition 2.2 ([4, Exercise 10.10]). *Consider a function $f_1\colon \mathbf{R}^n \to \overline{\mathbf{R}}$ that is locally Lipschitz around a point $\bar{x} \in \mathbf{R}^n$ and a function $f_2\colon \mathbf{R}^n \to \overline{\mathbf{R}}$ that is lower semi-continuous and proper with $f_2(\bar{x})$ finite. Then the inclusion*

$$\partial(f_1 + f_2)(\bar{x}) \subset \partial f_1(\bar{x}) + \partial f_2(\bar{x}),$$

holds.

We can now talk about restrictions of subjets. Given a function $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a set $M \subset \mathbf{R}^n$, we define the restriction of $[\partial f]$ to M to be the set $[\partial f]|_M := [\partial f] \cap (M \times \mathbf{R} \times \mathbf{R}^n)$. Analogous notation will be used for restrictions of the subjet $[\hat{\partial} f]$. Observe that in general, the set $[\partial f]|_M$ is not a subjet of any function. More generally, for a set $F \subset \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^n$ and a set $M \subset \mathbf{R}^n$, we let $F|_M := F \cap (M \times \mathbf{R} \times \mathbf{R}^n)$.

An open ball of radius r around a point $x \in \mathbf{R}^n$ will be denoted by $B_r(x)$, while the closed ball of radius r around a point $x \in \mathbf{R}^n$ will be denoted by $\bar{B}_r(x)$. The open and the closed unit balls will be denoted by $\mathbf{B}$ and $\overline{\mathbf{B}}$, respectively. Consider a set $M \subset \mathbf{R}^n$. We denote the topological closure, interior, and boundary of M by cl M, int M, and bd M, respectively. We define the *indicator function* of M, $\delta_M\colon \mathbf{R}^n \to \overline{\mathbf{R}}$, to be 0 on M and $+\infty$ elsewhere. Indicator functions allow us to translate analytic information about functions to geometric information about sets. In this spirit, we now introduce *normal cones*, which are the geometric analogues of subdifferentials.

Definition 2.3. Consider a set $M \subset \mathbf{R}^n$ and a point $x \in \mathbf{R}^n$. We define $\hat{N}_M(x) := \hat{\partial}\delta_M(x)$, $N_M(x) := \partial\delta_M(x)$ and $\overline{N}_M(x) := \operatorname{cl}\operatorname{conv}\big(N_M(x)\big)$.

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 $\mathbf{C}^1$-*smooth* if for each point $\bar{x} \in Q$, there is a neighborhood U of $\bar{x}$ and a $\mathbf{C}^1$ mapping $\hat{F}\colon \mathbf{R}^n \to \mathbf{R}^m$ that agrees with F on $Q \cap U$. Henceforth, the word *smooth* will always mean $\mathbf{C}^1$-smooth. Since we will not need higher orders of smoothness in our work, no ambiguity should arise. If a smooth function F is bijective and its inverse is also smooth, then we say that F is a *diffeomorphism*. More generally, we have the following definition.

Definition 2.4. Consider sets $Q \subset \mathbf{R}^n$, $\widetilde{Q} \subset \mathbf{R}^m$, and a mapping $F\colon Q \to \widetilde{Q}$. We say that F is a *local diffeomorphism* around a point $\bar{x} \in Q$ if there exists a neighborhood U of $\bar{x}$ such that the restriction

$$F|_{Q \cap U}\colon Q \cap U \to F(Q \cap U), \tag{1}$$

is a diffeomorphism. Now consider another set $K \subset \mathbf{R}^m$. We say that F is a local diffeomorphism around $\bar{x}$ onto K if there exists a neighborhood U of $\bar{x}$ such that the mapping in (1) is a diffeomorphism and $K = F(Q \cap U)$.

We now recall the notion of a *manifold*.

Definition 2.5. Consider a set $M \subset \mathbf{R}^n$. We say that M is a *manifold* of dimension r if for each point $\bar{x} \in M$, there is an open neighborhood U around $\bar{x}$ such that $M \cap U = F^{-1}(0)$, where $F\colon U \to \mathbf{R}^{n-r}$ is a $\mathbf{C}^1$ smooth map with $\nabla F(\bar{x})$ of full rank. In this case, we call F a *local defining function* for M around $\bar{x}$.

Strictly speaking, what we call a manifold is usually referred to as a $\mathbf{C}^1$-*submanifold* of $\mathbf{R}^n$. For a manifold $M \subset \mathbf{R}^n$ and a point $x \in M$, the three normal cones, defined above, coincide and are equal to the normal space, in the sense of differential geometry. For more details, see for example [4, Example 6.8]. For a smooth map $F\colon M \to N$, where M and N are manifolds, we say that F has *constant rank* if its derivative has constant rank throughout M.

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

$$d_M(x) = \inf_{y \in M} |x - y|,$$

and the projection of x onto M is

$$P_M(x) = \{y \in M : |x - y| = d_M(x)\}.$$

Finally, we will need the following result.

Theorem 2.6 *([4, Example 10.32]). For a closed set $M \subset \mathbf{R}^n$, the inclusion*

$$\partial[d_M^2](x) \subset 2[x - P_M(x)],$$

holds for all $x \in \mathbf{R}^n$.

2.2. Semi-algebraic geometry

A *semi-algebraic* set $S \subset \mathbf{R}^n$ is a finite union of sets of the form

$$\{x \in \mathbf{R}^n : P_1(x) = 0, \ldots, P_k(x) = 0, Q_1(x) < 0, \ldots, Q_l(x) < 0\},$$

where $P_1, \ldots, P_k$ and $Q_1, \ldots, Q_l$ are polynomials in n variables. In other words, S is a union of finitely many sets, each defined by finitely many polynomial equalities and inequalities. A map $F : \mathbf{R}^n \rightrightarrows \mathbf{R}^m$ is said to be *semi-algebraic* if $\operatorname{gph} F \subset \mathbf{R}^{n+m}$ is a semi-algebraic set. Semi-algebraic sets enjoy many nice structural properties. We discuss some of these properties in this section. For more details, see the monographs of Basu–Pollack–Roy [15], Lou van den Dries [16], and Shiota [13]. For a quick survey, see the article of van den Dries–Miller [17] and the surveys of Coste [18,19]. Unless otherwise stated, we follow the notation of [17,18].

A fundamental fact about semi-algebraic sets is provided by the Tarski–Seidenberg Theorem [18, Theorem 2.3]. Roughly speaking, it states that a linear projection of a semi-algebraic set remains semi-algebraic. From this result, it follows that a great many constructions preserve semi-algebraicity. In particular, for a semi-algebraic function $f : \mathbf{R}^n \to \overline{\mathbf{R}}$, it is easy to see that the set-valued mappings $\hat{\partial} f$, ∂f, $\overline{\partial} f$ along with the subjets $[\hat{\partial} f]$, $[\partial f]$, $[\overline{\partial} f]$ are semi-algebraic. See for example [14, Proposition 3.1].

Definition 2.7. Given finite collections $\{B_i\}$ and $\{C_j\}$ of subsets of $\mathbf{R}^n$, we say that $\{B_i\}$ is *compatible* with $\{C_j\}$ if for all B_i and C_j, either $B_i \cap C_j = \emptyset$ or $B_i \subset C_j$.

Definition 2.8. Consider a semi-algebraic set Q in $\mathbf{R}^n$. A *stratification* of Q is a finite partition of Q into disjoint, connected, semi-algebraic manifolds M_i (called strata) with the property that for each index i, the intersection of the closure of M_i with Q is the union of some M_j's.

The most striking and useful fact about semi-algebraic sets is that stratifications of semi-algebraic sets always exist. In fact, a more general result holds, which is the content of the following theorem.

Theorem 2.9 *([17, Theorem 4.8]). Consider a semi-algebraic set S in $\mathbf{R}^n$ and a semi-algebraic map $f : S \to \mathbf{R}^m$. Then there exists a stratification $\mathcal{A}$ of S and a stratification $\mathcal{B}$ of $\mathbf{R}^m$ such that for every stratum $M \in \mathcal{A}$, we have that the restriction $f|_M$ is smooth, $f(M) \in \mathcal{B}$, and f has constant rank on M. Furthermore, if $\mathcal{A}'$ is some other stratification of S, then we can ensure that $\mathcal{A}$ is compatible with $\mathcal{A}'$.*

Definition 2.10. Let $A \subset \mathbf{R}^n$ be a nonempty semi-algebraic set. Then we define the *dimension* of A, $\dim A$, to be the maximal dimension of a stratum in any stratification of A. We adopt the convention that $\dim \emptyset = -\infty$.

It can be easily shown that the dimension does not depend on the particular stratification. The dimension is a very well behaved quantity, which is the content of the following proposition. See [16, Chapter 4] for more details.

Theorem 2.11. *Let A and B be nonempty semi-algebraic sets in $\mathbf{R}^n$. Then the following hold.*

1. *If $A \subset B$, then* $\dim A \le \dim B$.
2. $\dim A = \dim \operatorname{cl} A$.
3. $\dim(\operatorname{cl} A \setminus A) < \dim A$.
4. *If $f : A \to \mathbf{R}^n$ is a semi-algebraic mapping, then* $\dim f(A) \le \dim A$. *If f is one-to-one, then* $\dim f(A) = \dim A$. *In particular, semi-algebraic homeomorphisms preserve dimension.*
5. $\dim A \cup B = \max\{\dim A, \dim B\}$.
6. $\dim A \times B = \dim A + \dim B$.

Observe that the dimension of a semi-algebraic set only depends on the maximal dimensional manifold in a stratification. Hence, the dimension is a somewhat crude measure of the size of the semi-algebraic set. In particular, it does not provide much insight into what the set looks like locally around each of its point. Hence, this motivates a localized notion of dimension.

Definition 2.12. Consider a semi-algebraic set $Q \subset \mathbf{R}^n$ and a point $\bar{x} \in Q$. We let the *local dimension* of Q at $\bar{x}$ be

$$\dim_Q(\bar{x}) := \inf_{r>0} \dim(Q \cap B_r(\bar{x})).$$

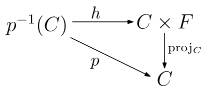
In fact, it is not hard to see that there exists a real number $\bar{r} > 0$ such that for every real number $0 < r < \bar{r}$, we have $\dim_Q(\bar{x}) = \dim(Q \cap B_r(\bar{x}))$.

The following is now an easy observation.

Proposition 2.13 (*[18, Exercise 3.19]*). *For any semi-algebraic set $Q \subset \mathbf{R}^n$, we have the identity*

$$\dim Q = \max_{x \in Q} \dim_Q(x).$$

Definition 2.14. Let $A \subset \mathbf{R}^n$ be a semi-algebraic set. A continuous semi-algebraic mapping $p: A \to \mathbf{R}^m$ is *semi-algebraically trivial* over a semi-algebraic set $C \subset \mathbf{R}^m$ if there is a semi-algebraic set F and a semi-algebraic homeomorphism $h: p^{-1}(C) \to C \times F$ such that $p|_{p^{-1}(C)} = \text{proj}_C \circ h$, or in other words the following diagram commutes:



We call h a *semi-algebraic trivialization* of p over C.

Henceforth, we use the symbol $\cong$ to indicate that two semi-algebraic sets are semi-algebraically homeomorphic.

Remark 2.15. If p is trivial over some semi-algebraic set C, then we can decompose $p|_{p^{-1}(C)}$ into a homeomorphism followed by a simple projection. Also, since the homeomorphism h in the definition is surjective and $p|_{p^{-1}(C)} = \text{proj}_C \circ h$, it easily follows that for any point $c \in C$, we have $p^{-1}(c) \cong F$ and $p^{-1}(C) \cong C \times p^{-1}(c)$.

Definition 2.16. In the notation of Definition 2.14, a trivialization h is *compatible* with a semi-algebraic set $B \subset A$ if there is a semi-algebraic set $H \subset F$ such that $h(B \cap p^{-1}(C)) = C \times H$.

If h is a trivialization over C then, certainly, for any set $B \subset A$ we know h restricts to a homeomorphism from $B \cap p^{-1}(C)$ to $h(B \cap p^{-1}(C))$. The content of the definition above is that if p is compatible with B, then h restricts to a homeomorphism between $B \cap p^{-1}(C)$ and the product $C \times H$ for some semi-algebraic set $H \subset F$.

The following is a remarkably useful theorem [18, Theorem 4.1].

Theorem 2.17 (*Hardt Triviality*). *Let $A \subset \mathbf{R}^n$ be a semi-algebraic set and $p: A \to \mathbf{R}^m$, a continuous semi-algebraic mapping. Then, there is a finite partition of the image $p(A)$ into semi-algebraic sets $C_1, \ldots, C_k$ such that p is semi-algebraically trivial over each C_i. Moreover, if Q is a semi-algebraic subset of A, we can require each trivialization $h_i: p^{-1}(C_i) \to C_i \times F_i$ to be compatible with Q.*

For an application of Hardt triviality to semi-algebraic set-valued analysis, see [12, Section 2.2]. The following proposition is a simple consequence of Hardt triviality.

Proposition 2.18. *Consider semi-algebraic sets M and Q satisfying $M \subset Q \subset \mathbf{R}^n$. Assume that there exists a continuous mapping $p: Q \to \mathbf{R}^m$, for some positive integer m, such that for each point x in the image $p(Q)$ we have $\dim p^{-1}(x) = \dim(p^{-1}(x) \cap M)$. Then M and Q have the same dimension.*

Proof. Applying Theorem 2.17 to the map p, we partition the image $p(Q)$ into finitely many disjoint sets $C_1, \ldots, C_k$ such that for each index i, we have the relations

$$p^{-1}(C_i) \cong C_i \times p^{-1}(c),$$
$$p^{-1}(C_i) \cap M \cong C_i \times (p^{-1}(c) \cap M),$$

where c is any point in C_i. Since by assumption, the equation $\dim p^{-1}(x) = \dim(p^{-1}(x) \cap M)$ holds for all points x in the image $p(Q)$, we deduce

$$\dim p^{-1}(C_i) = \dim(p^{-1}(C_i) \cap M),$$

for each index i. Thus

$$\begin{aligned}\dim Q &= \dim \bigcup_i p^{-1}(C_i) = \max_i \dim p^{-1}(C_i) = \max_i \dim(p^{-1}(C_i) \cap M) \\ &= \dim \bigcup_i (p^{-1}(C_i) \cap M) = \dim M,\end{aligned}$$

as we needed to show. □

We will have occasion to use the following simple proposition.

Proposition 2.19 (*[19, Theorem 3.18]*). *Consider a semi-algebraic, set-valued mapping $F: \mathbf{R}^n \rightrightarrows \mathbf{R}^m$. Suppose there exists an integer k such that the set $F(x)$ is k-dimensional for each point $x \in \operatorname{dom} F$. Then the equality,*

$$\dim \operatorname{gph} F = \dim \operatorname{dom} F + k,$$

holds.

3. Main results

In our current work, we build on the following theorem. This result and its consequences for generic semi-algebraic optimization problems are discussed extensively in [12].

Theorem 3.1 (*[12, Theorem 3.6]*). *Let $f: \mathbf{R}^n \to \overline{\mathbf{R}}$ be a proper semi-algebraic function. Then the graphs of the subdifferential mappings $\hat{\partial} f$ and ∂f have dimension exactly n.*

In fact, Theorem 3.1 also holds for $\overline{\partial} f$. For more details see [12].

To motivate our current work, consider a manifold $M \subset \mathbf{R}^n$. The set,

$$\operatorname{gph} N_M = \{(x, y) \in \mathbf{R}^n \times \mathbf{R}^n : y \in N_M(x)\},$$

is the normal bundle of M, and as such, $\operatorname{gph} N_M$ is itself a manifold of dimension n [20, Proposition 10.18]. In particular, $\operatorname{gph} N_M$ is n-dimensional, *locally* around each of its points. This suggests that perhaps Theorem 3.1 may be strengthened to pertain to the local dimension of the graph of the subdifferential. Indeed, this is the case. In fact, we will prove something stronger.

Let $f: \mathbf{R}^n \to \overline{\mathbf{R}}$ be a lower semicontinuous, proper, semi-algebraic function. Observe that the sets $\operatorname{gph} \partial f$ and $[\partial f]$ are in semi-algebraic bijective correspondence, via the map $(x, v) \longmapsto (x, f(x), v)$, and hence these two sets have the same dimension. Thus by Theorem 3.1, the dimension of the subjet $[\partial f]$ is exactly n. Combining this observation with Proposition 2.13, we deduce that the local dimension of $[\partial f]$ at each of its points is at most n. In this work, we prove that, remarkably, the local dimension of $[\partial f]$ at each of its points is exactly n (Theorem 3.8). From this result, it easily follows that the local dimension of $\operatorname{gph} \partial f$ at each of its points is exactly n as well. Analogous result holds for the subjet $[\hat{\partial} f]$.

The proof of Theorem 3.8 relies on a very general accessibility result, which we establish in Lemma 3.2. This result, in fact, holds in the absence of semi-algebraicity. In Remark 3.10, we provide a simple example illustrating that the assumption of lower-semicontinuity is necessary for our conclusions to hold. Then in Section 3.2, we show that the graph of the mapping $\overline{\partial} f$ may have small local dimension at some of its points. Thus, the analogue of Theorem 3.8 fails for the subdifferential $\overline{\partial} f$. This further illustrates the subtlety involved when analyzing local dimension.

3.1. Geometry of subdifferential mappings

Consider a proper, lower-semicontinuous, semi-algebraic function $f: \mathbf{R}^n \to \overline{\mathbf{R}}$. In this subsection, we aim to prove that the subjet $[\hat{\partial} f]$, and consequently $[\partial f]$, has local dimension n around each of its points (Theorem 3.8). Before proceeding with the technical details of the argument, it is instructive to first describe the following situation, which will be of great importance.

Consider an m-dimensional semi-algebraic manifold $M \subset \operatorname{dom} \hat{\partial} f$, with f smooth on M. It is not difficult to see that for any point $x \in M$, a translate of $\hat{\partial} f(x)$ is contained in the normal space $N_M(x)$. Therefore, the inequality

$$\dim \hat{\partial} f(x) \leq n - m, \tag{2}$$

holds. This inequality can easily be strict for every point $x \in M$. (For example, consider the function f on $\mathbf{R}^2$, where $f(x, y) = (|x| + |y|)^2$, and a singleton set $M := \{(0, 0)\}$.) In light of this, suppose that the strict inequality, $\dim [\hat{\partial} f]|_M < n$, holds. Then validity of Theorem 3.8 would immediately imply the inclusion $[\hat{\partial} f]|_M \subset \operatorname{cl} [\hat{\partial} f]|_{M^c}$.

In fact, in Corollary 3.3 we prove this consequence directly, that is without using Theorem 3.8. The key ingredient in the proof of this corollary is a powerful accessibility result, which we establish in Lemma 3.2. Armed with Corollary 3.3 and stratification techniques, we then prove Theorem 3.8 in its full generality. We now begin the formal development.

Lemma 3.2 (*Accessibility*). *Let $f: \mathbf{R}^n \to \overline{\mathbf{R}}$ be a lower semicontinuous function and $M \subset \mathbf{R}^n$ a set on which f is finite. Suppose that M is locally closed around a point $\bar{x} \in M$ and consider a triple $(\bar{x}, f(\bar{x}), \bar{v}) \in [\hat{\partial} f]|_M$. Suppose that there exists a sequence of real numbers $m_i \to \infty$ such that*

$$\bar{v} \in \operatorname{bd} \bigcup_{x \in M} \partial \left(f(\cdot) + \frac{1}{2} m_i |\cdot - \bar{x}|^2 \right)(x),$$

for each i. Then the inclusion $(\bar{x}, f(\bar{x}), \bar{v}) \in \operatorname{cl}[\hat{\partial} f]|_{M^c}$ holds. That is there exist sequences x_i and v_i, with $v_i \in \hat{\partial} f(x_i)$ and $x_i \notin M$, such that $(x_i, f(x_i), v_i)$ converges to $(\bar{x}, f(\bar{x}), \bar{v})$.

Proof. We first prove the lemma for the special case when $(\bar{x}, f(\bar{x}), \bar{v}) = (0, 0, 0)$. The general result will then easily follow.

Thus, assume that there exists a sequence of real number m_i with $m_i \to \infty$, such that the inclusion

$$0 \in \operatorname{bd} \bigcup_{x \in M} m_i x + \partial f(x), \tag{3}$$

holds. We must show that there exists a sequence $(x_i, f(x_i), v_i) \in [\hat{\partial} f]|_{M^c}$ converging to $(0, 0, 0)$.

We make some simplifying assumptions. Since f is lower semicontinuous and M is locally closed around the point 0, there exists a real number $r > 0$ such that $f|_{r\overline{\mathbf{B}}} \geq -1$ and the set $M \cap \frac{1}{2} r\overline{\mathbf{B}}$ is closed.

Claim. *Without loss of generality, we can replace the function f by $f_o := f + \delta_{r\overline{\mathbf{B}}}$ and the set M by $M_o := M \cap \frac{1}{2} r\overline{\mathbf{B}}$.*

Proof. Observe $(0, 0, 0) \in [\hat{\partial} f_o]|_{M_o}$. Furthermore, we have

$$[\partial f_o]|_{M_o} = [\partial (f + \delta_{r\overline{\mathbf{B}}})]|_{M \cap \frac{1}{2} r\overline{\mathbf{B}}} = [\partial f]|_{M \cap \frac{1}{2} r\overline{\mathbf{B}}} \subset [\partial f]|_M.$$

Combining this with (3), we obtain

$$0 \in \operatorname{bd} \bigcup_{x \in M_o} m_i x + \partial f_o(x).$$

Consequently, if we replace the function f by f_0 and the set M by M_0, then the requirements of the lemma will still be satisfied. Now suppose that with this replacement, the result of the lemma holds. Then there exists a sequence $(x_i, f(x_i), v_i) \in [\hat{\partial} f_o]|_{M_o^c}$ converging to $(0, 0, 0)$. For indices i satisfying $|x_i| < \frac{1}{2} r$, we have $x_i \notin M$ and $(x_i, f(x_i), v_i) \in [\hat{\partial} f]$. Thus restricting to large enough i, we obtain a sequence $(x_i, f(x_i), v_i) \in [\hat{\partial} f]|_{M^c}$ converging to $(0, 0, 0)$, as claimed. Therefore, without loss of generality, we can replace the function f by f_o and the set M by M_o. □

Thus to summarize, we have

$$(\bar{x}, f(\bar{x}), \bar{v}) = (0, 0, 0), \qquad f|_{r\overline{\mathbf{B}}} \geq -1, \qquad M \subset \frac{1}{2} r\overline{\mathbf{B}},$$
$$M \text{ is closed}, \ f(x) = +\infty \quad \text{for } x \notin r\overline{\mathbf{B}}.$$

We now define a certain auxiliary sequence of vectors y_i, which will allow us to construct the sequence $(x_i, f(x_i), v_i)$ that we seek. To this end, let y_i be a sequence satisfying $y_i \to 0$ and

$$y_i \notin \bigcup_{x \in M} m_i x + \partial f(x), \tag{4}$$

for each index i. By (3), such a sequence can easily be constructed. The motivation behind our choice of the sequence y_i will soon become apparent.

The key idea now is to consider the following sequence of minimization problems.

$$P(i) : \min_{x \in \mathbf{R}^n} \langle -y_i, x \rangle + m_i (d_M^2(x) + |x|^2) + f(x).$$

Since f is lower semi-continuous and the domain of f is bounded, we conclude that there exists a minimizer x_i for the problem $P(i)$. For each index i, we have

$$\begin{aligned} y_i \in \partial [m_i (d_M^2(\cdot) + |\cdot|^2) + f(\cdot)](x_i) &\subset \partial [m_i (d_M^2(\cdot) + |\cdot|^2)](x_i) + \partial f(x_i) \\ &\subset m_i (x_i - P_M(x_i)) + m_i x_i + \partial f(x_i), \end{aligned} \tag{5}$$

where the inclusions follow from Proposition 2.2 and Theorem 2.6. We claim

$$x_i \notin M, \tag{6}$$

for each index i. Indeed, if it were otherwise, from (5) we would have

$$y_i \in m_i x_i + \partial f(x_i) \subset \bigcup_{x\in M} m_i x + \partial f(x),$$

thus contradicting our choice of the vector y_i.

Now from (5), let $z_i \in P_M(x_i)$ be a vector satisfying

$$v_i := y_i - m_i(x_i - z_i) - m_i x_i \in \partial f(x_i). \tag{7}$$

Our immediate goal is to show that the sequence $(x_i, f(x_i), v_i) \in [\partial f]|_{M^c}$ converges to $(0, 0, 0)$. To that end, evaluating the value function of $P(i)$ at 0, we obtain

$$0 \geq \langle -y_i, x_i\rangle + m_i(d_M^2(x_i) + |x_i|^2) + f(x_i).$$

From (6), we deduce $x_i \neq 0$, and combining this with the inequality above, we obtain

$$|y_i| \geq \left\langle y_i, \frac{x_i}{|x_i|}\right\rangle \geq m_i \frac{d_M^2(x_i)}{|x_i|} + m_i|x_i| + \frac{f(x_i)}{|x_i|}.$$

Since $y_i \to 0$, $m_i \to \infty$, and the function f is bounded below, it is easy to see that x_i converges to 0. Furthermore, since we have $0 \in \hat{\partial} f(0)$, we deduce

$$\frac{f(x_i)}{|x_i|} \geq \frac{o(|x_i|)}{|x_i|}.$$

In particular, we conclude $m_i|x_i| \to 0$ and $f(x_i) \to 0$. Since $d_M(x_i) \leq |x_i|$, we deduce $m_i d_M(x_i) \to 0$. Hence from (7), we obtain

$$|v_i| \leq |y_i| + m_i d_M(x_i) + m_i|x_i| \to 0.$$

Thus we have produced a sequence $(x_i, f(x_i), v_i) \in [\partial f]|_{M^c}$ converging to $(0, 0, 0)$ with $x_i \notin M$ for each index i. We are almost done. The trouble is that the inclusion $v_i \in \partial f(x_i)$ holds, rather than $v_i \in \hat{\partial} f(x_i)$. However, this can be dealt with easily. Since M is closed, it is easy to see that we can perturb the triples $(x_i, f(x_i), v_i)$, to obtain a sequence $(x_i', f(x_i'), v_i') \in [\hat{\partial} f]$ converging to $(0, 0, 0)$, still satisfying $x_i' \notin M$ for each index i. This completes the proof for the case when $(\bar{x}, f(\bar{x}), \bar{v}) = (0, 0, 0)$.

Finally, we prove that the lemma holds when $(\bar{x}, f(\bar{x}), \bar{v}) \neq (0, 0, 0)$. Suppose that the point $(\bar{x}, f(\bar{x}), \bar{v})$, the set M, and the function f satisfy the requirements of the lemma. Now, consider the function $g(x) := f(x+\bar{x}) - \langle \bar{v}, x\rangle - f(\bar{x})$ and the set $N := M - \bar{x}$. We will show that the function g, the set N, and the triple $(0, 0, 0)$ also satisfy the requirements of the lemma. To this end, observe $0 \in N$ and $g(0) = 0$. It is easy to verify the equivalence,

$$v \in \hat{\partial} g(x) \Leftrightarrow v + \bar{v} \in \hat{\partial} f(x + \bar{x}).$$

Hence, clearly, $(0, 0, 0) \in [\hat{\partial} g]|_N$. Furthermore, the equation

$$\bigcup_{x\in N} \partial \left(g(\cdot) + \frac{1}{2} m|\cdot|^2\right)(x) = -\bar{v} + \bigcup_{x\in M} \partial \left(f(\cdot) + \frac{1}{2} m|\cdot - \bar{x}|^2\right)(x),$$

holds. Consequently, we deduce $0 \in \operatorname{bd} \bigcup_{x\in N} \partial \left(g(\cdot) + \frac{1}{2} m|\cdot|^2\right)(x)$. We can now apply the lemma to the triple $(0, 0, 0)$, the function g, and the set N. Thus there exists a sequence $(x_i, f(x_i), v_i) \in [\hat{\partial} g]|_{N^c}$ with $(x_i, g(x_i), v_i) \to (0, 0, 0)$. Now observe that the sequence $(x_i + \bar{x}, f(x_i + \bar{x}), v_i + \bar{v})$ lies in $[\hat{\partial} f]|_{M^c}$ and converges to $(\bar{x}, f(\bar{x}), \bar{v})$, and hence the lemma follows. □

In the semi-algebraic setting, Lemma 3.2 yields the following important corollary. This corollary, as alluded to in the beginning of this subsection, will be crucial for proving our main result (Theorem 3.8).

Corollary 3.3. *Consider a lower semicontinuous, semi-algebraic function $f: \mathbf{R}^n \to \overline{\mathbf{R}}$ and a semi-algebraic manifold $M \subset \mathbf{R}^n$ such that $f|_M$ is finite. Assume $\dim[\partial f]|_M < n$. Then the inclusion*

$$[\partial f]|_M \subset \operatorname{cl}[\hat{\partial} f]|_{M^c},$$

holds. That is, for any triple $(\bar{x}, f(\bar{x}), \bar{v})$ in the restricted subjet $[\partial f]|_M$, there exist sequences x_i and v_i, with $v_i \in \hat{\partial} f(x_i)$ and $x_i \notin M$, such that $(x_i, f(x_i), v_i) \to (\bar{x}, f(\bar{x}), \bar{v})$.

Proof. Consider an arbitrary triple $(\bar{x}, f(\bar{x}), \bar{v}) \in [\hat{\partial} f]|_M$ and let m be a real number. Observe that the map

$$\phi: [\partial f]|_M \to \operatorname{gph}(m(\cdot - \bar{x}) + \partial f(\cdot))|_M$$
$$(x, y, v) \mapsto (x, m(x - \bar{x}) + v)$$

is bijective. Thus we deduce

$$\dim \operatorname{gph} (m(\cdot - \bar{x}) + \partial f(\cdot))|_M = \dim[\partial f]|_M < n.$$

Hence, the set

$$\bigcup_{x\in M} m(x-\bar{x}) + \partial f(x) = \bigcup_{x\in M} \partial \left(f(\cdot) + \frac{1}{2} m|\cdot - \bar{x}|^2 \right)(x),$$

has dimension strictly less than n, and in particular has empty interior. Therefore, we have

$$\bar{v} \in \operatorname{bd} \bigcup_{x\in M} \partial \left(f(\cdot) + \frac{1}{2} m|\cdot - \bar{x}|^2 \right)(x),$$

for any real number m. Noting that any manifold is locally closed around each of its point and using Lemma 3.2, we deduce that the inclusion $(\bar{x}, f(\bar{x}), \bar{v}) \in \operatorname{cl} [\hat{\partial} f]|_{M^c}$ holds. Consequently we obtain

$$[\hat{\partial} f]|_M \subset \operatorname{cl} [\hat{\partial} f]|_{M^c}. \tag{8}$$

Now consider a triple $(\bar{x}, f(\bar{x}), \bar{v}) \in [\partial f]|_M$. Then there exists a sequence $(x_i, f(x_i), v_i) \in [\hat{\partial} f]$ converging to $(\bar{x}, f(\bar{x}), \bar{v})$. If there is a subsequence contained in M^c, then we are done. If not, then the whole sequence eventually lies in M, and then from (8) the result follows. □

We now establish a few simple propositions concerning local dimension.

Proposition 3.4. *Consider a semi-algebraic set $Q \subset \mathbf{R}^n$ and a point $\bar{x} \in Q$. Let $\{M_i\}$ be any stratification of Q. Then we have the identity*

$$\dim_Q(\bar{x}) = \max_i \{\dim M_i : \bar{x} \in \operatorname{cl} M_i\}.$$

Proof. Since there are finitely many strata, there exists some real number $\epsilon > 0$ such that for any $0 < r < \epsilon$, we have

$$Q \cap B_r(\bar{x}) = \bigcup_{i:\ \bar{x}\in \operatorname{cl} M_i} M_i \cap B_r(\bar{x}).$$

Hence, we deduce

$$\begin{aligned} \dim(Q \cap B_r(\bar{x})) &= \max_i \{\dim(M_i \cap B_r(\bar{x})) : \bar{x} \in \operatorname{cl} M_i\} \\ &= \max_i \{\dim M_i : \bar{x} \in \operatorname{cl} M_i\}, \end{aligned}$$

where the last equality follows since the inclusion $\bar{x} \in \operatorname{cl} M_i$ implies that $M_i \cap B_r(\bar{x})$ is a nonempty open submanifold of M_i, and hence has the same dimension as M_i. Letting $r \to 0$ yields the result. □

Definition 3.5. Given a stratification $\{M_i\}$ of a semi-algebraic set $Q \subset \mathbf{R}^n$, we will say that a stratum M is *maximal* if it is not contained in the closure of any other stratum.

Remark 3.6. Using the defining property of a stratification, we can equivalently say that given a stratification $\{M_i\}$ of a semi-algebraic set $Q \subset \mathbf{R}^n$, a stratum M is maximal if and only if it is disjoint from the closure of any other stratum.

Proposition 3.7. *Consider a stratification $\{M_i\}$ of a semi-algebraic set $Q \subset \mathbf{R}^n$. Then given any point $\bar{x} \in Q$, there exists a maximal stratum M satisfying $\bar{x} \in \operatorname{cl} M$ and $\dim M = \dim_Q(\bar{x})$.*

Proof. By Proposition 3.4, we have the identity

$$\dim_Q(\bar{x}) = \max_i \{\dim M_i : \bar{x} \in \operatorname{cl} M_i\}.$$

Let M be a stratum achieving this maximum. If there existed a stratum M_i satisfying $M \subset \operatorname{cl} M_i$, then we would have $\dim M < \dim M_i$ and $\bar{x} \in \operatorname{cl} M \subset \operatorname{cl} M_i$, thus contradicting our choice of M. Therefore, we conclude that M is maximal. □

We are now ready to prove the main result of this section.

Theorem 3.8. *Let $f: \mathbf{R}^n \to \overline{\mathbf{R}}$ be a proper lower semicontinuous, semi-algebraic function. Then the subjet $[\hat{\partial} f]$ has local dimension n around each of its points. The same holds for the subjet $[\partial f]$.*

Proof. We first prove the claim for the subjet $[\hat{\partial} f]$ and then the case of $[\partial f]$ will be an easy consequence. Observe that the sets $\operatorname{gph} \hat{\partial} f$ and $[\hat{\partial} f]$ are in semi-algebraic bijective correspondence, via the map $(x, v) \longmapsto (x, f(x), v)$, and hence these two sets have the same dimension. Combining this observation with Theorem 3.1, we deduce that the dimension of $[\hat{\partial} f]$ is n. Thus the local dimension of $[\hat{\partial} f]$ at any point is at most n. We must now establish the reverse inequality.

Consider the subjet $[\hat{\partial} f]$ and the projection map $\pi : [\hat{\partial} f] \to \mathbf{R}^n$, which projects onto the first n coordinates. Applying Theorem 2.9 to π, we obtain a finite partition of $[\hat{\partial} f]$ into disjoint semi-algebraic manifolds $\{M_i\}$ and a finite partition of the image $\pi([\hat{\partial} f])$ into disjoint semi-algebraic manifolds $\{L_j\}$, such that for each index i, we have $\pi(M_i) = L_j$ for some index j.

Assume that the statement of the theorem does not hold. Thus there exists some point in the subjet $[\hat{\partial} f]$ at which $[\hat{\partial} f]$ has local dimension strictly less than n. Therefore, by Proposition 3.7, there is a maximal stratum M with $\dim M < n$. We now focus on this stratum.

Lemma 3.9.

$$\dim [\partial f]|_{\pi(M)} < n.$$

Proof. For each $x \in \pi(M)$, the set $M \cap \pi^{-1}(x)$ is open relative to $\pi^{-1}(x)$, since the alternative would contradict maximality of M. Thus

$$\dim(M \cap \pi^{-1}(x)) = \dim \pi^{-1}(x),$$

for each $x \in \pi(M)$. Therefore the sets M and $[\hat{\partial} f]|_{\pi(M)}$, along with the projection map π, satisfy the assumptions of Proposition 2.18. Hence we deduce $\dim [\hat{\partial} f]|_{\pi(M)} = \dim M < n$. Observe $[\partial f] \setminus [\hat{\partial} f] \subset (\operatorname{cl} [\hat{\partial} f]) \setminus [\hat{\partial} f]$. Hence as a direct consequence of Theorem 3.1, we see $\dim([\partial f] \setminus [\hat{\partial} f])|_{\pi(M)} \le \dim((\operatorname{cl} [\hat{\partial} f]) \setminus [\hat{\partial} f]) < n$. Thus we conclude $\dim [\partial f]|_{\pi(M)} < n$, as was claimed. □

Consider an arbitrary point $\bar{x} \in \pi(M)$ with $(\bar{x}, f(\bar{x}), \bar{v}) \in M$. Combining Corollary 3.3 and Lemma 3.9, we deduce that there exists a sequence $(x_i, f(x_i), v_i) \in [\hat{\partial} f]$ converging to $(\bar{x}, f(\bar{x}), \bar{v})$ where $x_i \notin \pi(M)$. Since there are finitely many strata, we conclude that the point $(\bar{x}, f(\bar{x}), \bar{v}) \in M$ is in the closure of some stratum other than M, thus contradicting maximality of M. Thus the subjet $[\hat{\partial} f]$ has local dimension n around each of its points.

Now for the subjet $[\partial f]$, observe that for any real number $r > 0$, we have $B_r(x, f(x), v) \cap [\hat{\partial} f] \neq \emptyset$. Hence it easily follows that $[\partial f]$ has local dimension n around each of its points as well. □

Remark 3.10. If a semi-algebraic function $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ is not lower semicontinuous, then the result of Theorem 3.8 can easily fail. For instance, consider the set $S := \{x \in \mathbf{R}^2 : |x| < 1\} \cup \{(1, 0)\}$. The local dimension of $[\partial \delta_S]$ at $((1, 0), 0, (1, 0))$ is one, rather than two.

3.2. Geometry under convexification

Consider a locally Lipschitz function $f : \mathbf{R}^n \to \mathbf{R}$, and let Ω be the set of points at which f is differentiable. By Rademacher's theorem, Ω has full measure, and furthermore it is well known that the representation

$$\overline{\partial} f(x) = \operatorname{conv}\{\lim_{i\to\infty} \nabla f(x_i) : x_i \to x, x_i \in \Omega\},$$

is valid for each point $x \in \mathbf{R}^n$.

It has been shown in [12, Theorem 3.6] that if f is semi-algebraic, then the global dimension of the set $\operatorname{gph} \overline{\partial} f$ is n. Since at each point x, the subdifferential $\overline{\partial} f(x)$ contains both $\hat{\partial} f(x)$ and $\partial f(x)$, it is tempting to think, in light of Theorem 3.8, that the set $\operatorname{gph} \overline{\partial} f$ should have local dimension n around each of its points, as well.

It can be shown that this indeed is the case when $n \le 2$. In fact, this even holds for semi-linear functions for arbitrary n. (Semi-linear function are those functions whose domains can be decomposed into finitely many convex polyhedra so that the restriction of the function to each polyhedron is affine.) However for $n \ge 3$, as soon as we allow the function f to have any curvature at all, the conjecture is decisively false. Consider the following illustrative example.

Example 3.11. Consider the function $f : \mathbf{R}^3 \to \mathbf{R}$, defined by

$$f(x, y, z) = \begin{cases} \min\{x, y, z^2\}, & \text{if } (x, y, z) \in \mathbf{R}^3_+ \\ \min\{-x, -y, z^2\}, & \text{if } (x, y, z) \in \mathbf{R}^3_- \\ 0, & \text{otherwise.} \end{cases}$$

It is standard to verify that f is locally Lipschitz continuous and semi-algebraic. Let $\Gamma := \operatorname{conv}\{(1, 0, 0), (0, 1, 0), (0, 0, 0)\}$. Consider the set of points $\Omega \subset \mathbf{R}^3$ where f is differentiable. Then we have

$$\operatorname{conv}\{\lim_{i\to\infty} \nabla f(\gamma_i) : \gamma_i \to (0, 0, 0), \gamma_i \in \Omega \cap \mathbf{R}^3_+\} = \operatorname{conv}\{(1, 0, 0), (0, 1, 0), (0, 0, 0)\} = \Gamma,$$

and

$$\operatorname{conv}\{\lim_{i\to\infty} \nabla f(\gamma_i) : \gamma_i \to (0,0,0),\ \gamma_i \in \Omega \cap \mathbf{R}^3_-\} = \operatorname{conv}\{(-1,0,0),(0,-1,0),(0,0,0)\} = -\Gamma.$$

In particular, we deduce $\overline{\partial} f(0,0,0) = \operatorname{conv}\{\Gamma \cup -\Gamma\}$. Hence the subdifferential $\overline{\partial} f(0,0,0)$ has dimension two.

Let $((x_i,y_i,z_i),v_i) \in \operatorname{gph} \overline{\partial} f|_{\mathbf{R}^3_+}$ be a sequence converging to $((0,0,0),\bar{v})$, for some vector $\bar{v} \in \mathbf{R}^3$. Observe $v_i \in \operatorname{conv}\{(1,0,2z_i),(0,1,2z_i),(0,0,0)\}$. Hence, we must have $\bar{v} \in \Gamma$. Now consider a sequence $((x_i,y_i,z_i),v_i) \in \operatorname{gph} \overline{\partial} f|_{\mathbf{R}^3_-}$ converging to $((0,0,0),\bar{v})$, for some vector $\bar{v} \in \mathbf{R}^3$. A similar argument as above yields the inclusion $\bar{v} \in -\Gamma$. This implies that for any vector $\bar{v}$ in $\overline{\partial} f(0,0,0) \setminus (\Gamma \cup -\Gamma)$, there does not exist a sequence $((x_i,y_i,z_i),v_i) \in \operatorname{gph} \overline{\partial} f$ converging to $((0,0,0),\bar{v})$. Therefore for such a vector $\bar{v}$, there exists an open ball $B_\epsilon((0,0,0),\bar{v})$ such that $B_\epsilon((0,0,0),\bar{v}) \cap \operatorname{gph} \overline{\partial} f \subset \{(0,0,0)\} \times \overline{\partial} f(0,0,0)$. Thus the local dimension of $\operatorname{gph} \overline{\partial} f$ around the pair $((0,0,0),\bar{v})$ is two, instead of three.

3.3. Composite optimization

Consider a composite optimization problem

$$\min_x\ g(F(x)),$$

where $g\colon \mathbf{R}^m \to \overline{\mathbf{R}}$ is a lower semicontinuous, semi-algebraic function and $F\colon \mathbf{R}^n \to \mathbf{R}^m$ is a smooth, semi-algebraic mapping. It is often computationally more convenient to replace the criticality condition $0 \in \partial(g \circ F)(x)$ with the potentially different condition $0 \in \nabla F(x)^* \partial g(F(x))$, related to the former condition by an appropriate chain rule. See for example the discussion of Lagrange multipliers [21]. Thus it is interesting to study the graph of the set-valued mapping $x \mapsto \nabla F(x)^*\partial g(F(x))$. In fact, it is shown in [12, Theorem 5.3] that the dimension of the graph of this mapping is at most n. Furthermore, under some assumptions, such as the set $F^{-1}(\operatorname{dom} \partial g)$ having a nonempty interior for example, this graph has dimension exactly n.

In the spirit of our current work, we ask whether under reasonable conditions, the graph of the mapping $x \mapsto \nabla F(x)^*\partial g(F(x))$ has local dimension n around each of its points. In fact, the answer is no. That is, subdifferential calculus does not preserve local dimension. As an illustration, consider the following example.

Example 3.12. Observe that for a lower semicontinuous function f, if we let $F(x) = (x,x)$ and $g(x,y) = f(x) + f(y)$, then we obtain $\nabla F(x)^*\partial g(F(x)) = \partial f(x) + \partial f(x)$. Now let the function $f\colon \mathbf{R} \to \mathbf{R}$ be $f(x) = -|x|$. Then we have

$$\partial f(x) = \begin{cases} 1, & x < 0 \\ \{-1, 1\}, & x = 0 \\ -1, & x > 0. \end{cases}$$

The set $\operatorname{gph} \partial f$ has local dimension 1 around each of its point, as is predicted by Theorem 3.8. However, the graph of the mapping $x \mapsto \partial f(x) + \partial f(x)$ has an isolated point at $(0,0)$, and hence this graph has local dimension zero around this point, instead of one. Furthermore, using Theorem 3.8, we can now conclude that the mapping $x \mapsto \partial f(x) + \partial f(x)$ is not the subdifferential mapping of any semi-algebraic, lower semicontinuous function.

4. Consequences

In this section, we present some consequences of Theorem 3.8. Specifically, in Section 4.1 we develop a nonconvex, semi-algebraic analog of Minty's Theorem, and in Section 4.2 we derive certain sensitivity information about variational problems, using purely dimensional considerations. Both of these results illustrate that local dimension shows the promise of being a powerful, yet simple to use, tool in semi-algebraic optimization.

4.1. Analogue of Minty's Theorem

The celebrated theorem of Minty states that for a proper, lower semicontinuous, convex function $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$, the set $\operatorname{gph} \partial f$ is Lipschitz homeomorphic to $\mathbf{R}^n$ [6]. In fact, for each real number $\lambda > 0$, the so called Minty map $(x,y) \mapsto \lambda x + y$ is such a homeomorphism. For nonconvex functions, Minty's theorem easily fails. However, one may ask if for a nonconvex, lower semicontinuous function $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$, a Minty type result holds locally around many of the points in the set $\operatorname{gph} \partial f$. In general, nothing like this can hold either. However, in the semi-algebraic setting, Theorem 3.8 does provide an affirmative answer.

Proposition 4.1. *If $Q \subset \mathbf{R}^p$ has local dimension q around every point, then it is locally diffeomorphic to $\mathbf{R}^q$ around every point in a dense semi-algebraic subset.*

Proof. Applying Theorem 2.9, we obtain a stratification $\{M_i\}$ of Q. Let D be the union of the maximal strata in the stratification. By Proposition 3.7, we see that D is dense in Q. Now consider an arbitrary point $x \in D$ and let M be the maximal stratum containing this point. Since Q has local dimension q around x, we deduce that the manifold M has dimension q. By maximality of M, there exists a real number $r > 0$ such that $B_r(x) \cap Q = B_r(x) \cap M$, and hence Q is locally diffeomorphic to $\mathbf{R}^q$ around x, as we claimed. □

Consider a lower semicontinuous, semi-algebraic function $f: \mathbf{R}^n \to \overline{\mathbf{R}}$. Combining Proposition 4.1 and Theorem 3.8, we see that gph ∂f is locally diffeomorphic to $\mathbf{R}^n$ around every point in a dense semi-algebraic subset. In fact, we can significantly strengthen Proposition 4.1. Shortly, we will show that we can choose the local diffeomorphisms of Proposition 4.1 to have a very simple form that is analogous to the Minty map.

We will say that a certain property holds for a *generic* vector $v \in \mathbf{R}^n$ if the set of vectors for which this property does not hold is a semi-algebraic set of dimension strictly less than n. In the semi-algebraic setting, this notion coincides with the measure-theoretic concept of "almost everywhere". For a more in-depth discussion of generic properties in the semi-algebraic setting, see for example [22,12].

Definition 4.2. For a set $Q \subset \mathbf{R}^n$ and a map $\phi: Q \to \mathbf{R}^m$, we say that ϕ is *finite-to-one* if for every point $x \in \mathbf{R}^m$, the set $\phi^{-1}(x)$ consists of finitely many points.

We need the following proposition, which is essentially equivalent to [17, Theorem 4.9,]. We sketch a proof below, for completeness.

Proposition 4.3. *Let $Q \subset \mathbf{R}^n \times \mathbf{R}^n$ be a semi-algebraic set having dimension no greater than n. Then for a generic matrix $A \in \mathbf{R}^{n\times n}$, the map*

$$\phi_A: Q \to \mathbf{R}^n,$$
$$(x, y) \mapsto Ax + y,$$

is finite-to-one.

Proof. Let $I \in \mathbf{R}^{n\times n}$ be the identity matrix and consider the matrix $[A, I]$. Let L denote the nullspace of $[A, I]$. It is standard to check the equivalence

$$Ax + y = b \Leftrightarrow \pi_{L^\perp}(x, y) = \pi_{L^\perp}(0, b), \tag{9}$$

where $\pi_{L^\perp}$ denotes the orthogonal projection onto $L^\perp$. Recall that each element of a dense collection of n dimensional subspaces of $\mathbf{R}^n \times \mathbf{R}^n$ can be written uniquely as rge $[A, I]^T$ for some matrix A. From [17, Theorem 4.9], we have that for a generic n-dimensional subspace U of $\mathbf{R}^n \times \mathbf{R}^n$, the orthogonal projection map $\pi_U: Q \to U$ is finite-to-one. Hence, we deduce that for a generic matrix $A \in \mathbf{R}^{n\times n}$, the corresponding projection map $\pi_{L^\perp}$ is finite-to-one. Combining this with (9), the result follows. □

Proposition 4.4. *Consider a semi-algebraic set $Q \subset \mathbf{R}^n$ and a continuous, semi-algebraic mapping $p: Q \to \mathbf{R}^m$ that is finite-to-one. Then there exists a stratification of Q such that for each stratum M, the map $p|_M$ is a diffeomorphism onto its image.*

Proof. Applying Theorem 2.17, we obtain a partition of the image $p(Q)$ into semi-algebraic sets C_i such that the map p is semi-algebraically trivial over each C_i. Thus for each index i, and any point $c \in C_i$, there is a semi-algebraic homeomorphism $h: p^{-1}(C_i) \to C_i \times p^{-1}(c)$, such that the diagram,

$$\begin{array}{ccc} p^{-1}(C_i) & \xrightarrow{h} & C_i \times p^{-1}(c) \\ & \searrow^{p} & \downarrow \mathrm{proj}_{C_i} \\ & & C_i \end{array}$$

commutes.

Fix some index i. We will now show that the map p is injective on any connected subset of $p^{-1}(C_i)$. To this effect, consider a connected subset $M \subset p^{-1}(C_i)$. Observe that the set $h(M)$ is connected. Since $p^{-1}(c)$ is a finite set, we deduce that there exists a point $v \in p^{-1}(c)$ such that the inclusion,

$$h(M) \subset C_i \times \{v\} \tag{10}$$

holds. Now given any two distinct points $x, y \in M$, since h is a homeomorphism, we have $h(x) \neq h(y)$. Combining this with (10), we deduce $p(x) = \mathrm{proj}_{C_i} \circ h(x) \neq \mathrm{proj}_{C_i} \circ h(y) = p(y)$, as we needed to show.

Applying Theorem 2.9 to the map p, we obtain a finite partition of Q into connected, semi-algebraic manifolds $\{M_i\}$ compatible with $\{p^{-1}(C_i)\}$, such that for each stratum M_i, the map $p|_{M_i}$ is smooth and p has constant rank on M_i. Fix a stratum M. Since M is connected, it follows from the argument above that p is injective on M. Combining this observation with the fact that p has constant rank on M, we deduce that $p|_M$ is a diffeomorphism onto its image. □

We are now ready for the main result of this subsection.

Theorem 4.5. *Consider a semi-algebraic set $Q \subset \mathbf{R}^n \times \mathbf{R}^n$ that has local dimension n around every point. Then for a generic matrix $A \in \mathbf{R}^{n\times n}$, the map*

$$\phi_A: Q \to \mathbf{R}^n,$$
$$(x, y) \mapsto Ax + y,$$

is a local diffeomorphism of Q onto an open subset of $\mathbf{R}^n$, around every point in a dense semi-algebraic subset of Q.

Proof. By Proposition 4.3, we have that for a generic matrix $A \in \mathbf{R}^{n\times n}$, the map ϕ_A is finite-to-one. Fix such a matrix A. Consider the stratification guaranteed to exist by applying Proposition 4.4 to the map ϕ_A, and let D_A be the union of the maximal strata in this stratification. By Proposition 3.7, we see that D_A is dense in Q. Consider a point $(\bar{x}, \bar{y}) \in D_A$, which is contained in some maximal stratum M. Since the set Q has local dimension n around each of its points, we deduce that the stratum M is n-dimensional. Recall that the mapping $\phi_A|_M$ is a diffeomorphism onto its image. By maximality of M, there is a real number $\epsilon > 0$ such that $B_\epsilon(\bar{x}, \bar{y}) \cap M = B_\epsilon(\bar{x}, \bar{y}) \cap Q$ and hence the restricted mapping $\phi_A|_{B_\epsilon(\bar{x},\bar{y})\cap Q}$ is a diffeomorphism onto its image. Consequently the image $\phi_A(B_\epsilon(\bar{x}, \bar{y}) \cap Q)$ is an n-dimensional submanifold of $\mathbf{R}^n$, and hence is an open subset of $\mathbf{R}^n$. □

As a direct consequence of Theorems 4.5 and 3.8, we obtain

Corollary 4.6. *Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ be a lower semicontinuous, semi-algebraic function. Then for a generic matrix $A \in \mathbf{R}^{n\times n}$, the map*

$$\begin{aligned} \phi_A : \operatorname{gph} \partial f &\to \mathbf{R}^n, \\ (x, y) &\mapsto Ax + y, \end{aligned}$$

is a local diffeomorphism of $\operatorname{gph} \partial f$ onto an open subset of $\mathbf{R}^n$ around every point in a dense semi-algebraic subset of $\operatorname{gph} \partial f$. The analogous statement holds in the case of $\hat{\partial} f$.

4.2. Sensitivity

Proposition 4.7. *Consider a semi-algebraic set Q and a finite-to-one, continuous, semi-algebraic map $\phi : Q \to \mathbf{R}^m$. Then the map ϕ does not decrease local dimension, that is*

$$\dim_Q(x) \le \dim_{\operatorname{rge}\phi} \phi(x),$$

for any point $x \in Q$. In particular, semi-algebraic homeomorphisms preserve local dimension.

Proof. By Proposition 4.4, there exists a stratification of Q into semi-algebraic manifolds $\{M_i\}$, such that for each maximal stratum M, the restriction $\phi|_M$ is a diffeomorphism onto its image. Fix some point $x \in Q$. By Proposition 3.7, there is a maximal stratum M satisfying $x \in \operatorname{cl} M$ and $\dim M = \dim_Q(x)$. Now since $\phi|_M$ is a diffeomorphism onto its image, we deduce that the manifold $\phi(M)$ has dimension $\dim_Q(x)$. By continuity of ϕ, we have $\phi(x) \in \operatorname{cl} \phi(M)$. Hence,

$$\dim_{\operatorname{rge}\phi} \phi(x) \ge \dim \phi(M) = \dim_Q(x),$$

as we needed to show. □

Proposition 4.8. *Let $Q \subset \mathbf{R}^n \times \mathbf{R}^n$ be a semi-algebraic set and suppose that Q has local dimension n at a point $(\bar{x}, \bar{y})$. Consider the following parametric system, parametrized by matrices $A \in \mathbf{R}^{n\times n}$ and vectors $b \in \mathbf{R}^n$.*

$$\begin{aligned} P(A, b) : \quad & (x, y) \in Q, \\ & Ax + y = b. \end{aligned}$$

Define the solution set, $S(A, b)$, to be the set of all pairs (x, y) solving $P(A, b)$. Suppose that we have $(\bar{x}, \bar{y}) \in S(\bar{A}, \bar{b})$, for some matrix $\bar{A}$ and vector $\bar{b}$. Fix some precision parameter $\epsilon > 0$, and let $\Omega \subset \mathbf{R}^{n\times n} \times \mathbf{R}^n$ be the set of parameters (A, b), for which the solution set $S(A, b)$ is finite and the intersection $S(A, b) \cap B_\epsilon(\bar{x}, \bar{y})$ is nonempty. Then for any real number $\delta > 0$, the set $\Omega \cap B_\delta(\bar{A}, \bar{b})$ has dimension $n^2 + n$, and in particular has strictly positive measure.

Proof. By Proposition 4.3, for a generic matrix $A \in \mathbf{R}^{n\times n}$ the map

$$\begin{aligned} &\phi_A : Q \to \mathbf{R}^n, \\ &(x, y) \mapsto Ax + y, \end{aligned}$$

is finite-to-one. Denote this generic collection of matrices by Σ. Let $Q' := Q \cap B_\epsilon(\bar{x}, \bar{y})$. Observe that for each matrix $A \in \Sigma$, the restriction $\phi_A|_{Q'}$ is still finite-to-one. For notational convenience, we will abuse notation slightly and we will always use the symbol ϕ_A to mean the restriction of ϕ_A to Q', that is we now have $\phi_A : Q' \to \mathbf{R}^n$.

Fix some arbitrary real numbers $\delta, \gamma > 0$, and let $N_{\delta,\gamma}(\bar{A}, \bar{b}) := B_\delta(\bar{A}) \times B_\gamma(\bar{b})$. We will show that the set $\Omega \cap N_{\delta,\gamma}(\bar{A}, \bar{b})$ has dimension $n^2 + n$. To this effect, observe that the inclusion,

$$\Omega \cap N_{\delta,\gamma}(\bar{A}, \bar{b}) \supset \{(A, b) \in \mathbf{R}^{n\times n} \times \mathbf{R}^n : A \in \Sigma \cap B_\delta(\bar{A}),\ b \in \operatorname{rge} \phi_A \cap B_\gamma(\bar{b})\}, \tag{11}$$

holds. The set on the right hand side of (11) is exactly the graph of the set-valued mapping,

$$\begin{aligned} &F : \Sigma \cap B_\delta(\bar{A}) \rightrightarrows \mathbf{R}^n, \\ &A \mapsto \operatorname{rge} \phi_A \cap B_\gamma(\bar{b}). \end{aligned}$$

Thus, in order to complete the proof, it is sufficient to show that $\operatorname{gph} F$ has dimension $n^2 + n$. We will do this by showing that both the domain and the values of F have large dimension.

First, we analyze the domain of F. Consider any matrix $A \in \Sigma \cap B_\delta(\bar{A})$. We have

$$|\phi_A(\bar{x}, \bar{y}) - \bar{b}| = |(A\bar{x} + \bar{y}) - (\bar{A}\bar{x} + \bar{y})| \le |A - \bar{A}|\,|\bar{x}|.$$

So by shrinking δ, if necessary, we can assume $|\phi_A(\bar{x}, \bar{y}) - \bar{b}| < \gamma$. Hence, we deduce

$$\phi_A(\bar{x}, \bar{y}) \in \operatorname{rge} \phi_A \cap B_\gamma(\bar{b}). \tag{12}$$

In particular, we deduce that F is nonempty valued on $\Sigma \cap B_\delta(\bar{A})$. Combining this with the fact that the set Σ is generic, we obtain

$$\dim \operatorname{dom} F = \dim \Sigma \cap B_\delta(\bar{A}) = n^2. \tag{13}$$

We now analyze the set $F(A)$. Since the continuous map ϕ_A is finite-to-one and Q' has local dimension n at the point $(\bar{x}, \bar{y})$, appealing to Proposition 4.7, we obtain

$$\dim_{\operatorname{rge} \phi_A} \phi_A(\bar{x}, \bar{y}) = n. \tag{14}$$

From (12) and (14), we obtain

$$\dim F(A) = \dim \operatorname{rge} \phi_A \cap B_\gamma(\bar{b}) = n, \tag{15}$$

for all matrices $A \in \Sigma \cap B_\delta(\bar{A})$. Finally combining (13), (15), and Proposition 2.19, we deduce

$$\dim \operatorname{gph} F = \dim \operatorname{dom} F + \dim F(A) = n^2 + n,$$

thus completing the proof. □

Thus we have the following corollary.

Corollary 4.9. *Let $f: \mathbf{R}^n \to \overline{\mathbf{R}}$ be a lower semicontinuous, semi-algebraic function. Consider the following parametric system, parametrized by matrices $A \in \mathbf{R}^{n\times n}$ and vectors $b \in \mathbf{R}^n$.*

$$P(A, b): \quad \begin{aligned} & y \in \partial f(x), \\ & Ax + y = b. \end{aligned}$$

Define the solution set, $S(A, b)$, to be the set of all pairs (x, y) solving $P(A, b)$. Suppose that we have $(\bar{x}, \bar{y}) \in S(\bar{A}, \bar{b})$, for some matrix $\bar{A}$ and vector $\bar{b}$. Fix some precision parameter $\epsilon > 0$, and let $\Omega \subset \mathbf{R}^{n\times n} \times \mathbf{R}^n$ be the set of parameters (A, b), for which the solution set $S(A, b)$ is finite and the intersection $S(A, b) \cap B_\epsilon(\bar{x}, \bar{y})$ is nonempty. Then for any real number $\delta > 0$, the set $\Omega \cap B_\delta(\bar{A}, \bar{b})$ has dimension $n^2 + n$, and in particular has strictly positive measure.

To clarify Corollary 4.9, consider a solution $(\bar{x}, \bar{y})$ to the system $P(\bar{A}, \bar{b})$. Then the content of Corollary 4.9 is that under small random (continuously distributed) perturbations to the pair $(\bar{A}, \bar{b})$, with positive probability the perturbed system $P(A, b)$ has a strictly positive and finite number of solutions arbitrarily close to $(\bar{x}, \bar{y})$.

Acknowledgments

Much of the current work has been done while the first and second authors were visiting CRM (Centre de Recerca Matemàtica) at Universitat Autònoma de Barcelona. The concerned authors would like to acknowledge the hosts for their hospitality. We thank Aris Daniilidis and Jérôme Bolte for fruitful discussions, and we also thank C.H. Jeffrey Pang for providing the illustrative Example 3.11.

The work of Dmitriy Drusvyatskiy on this paper has been partially supported by the NDSEG grant from the Department of Defense. The second author's research was supported in part by the US–Israel Binational Scientific Foundation Grant 2008261. The third author's research was supported in part by National Science Foundation Grant DMS-0806057 and by the US–Israel Binational Scientific Foundation Grant 2008261.

References

[1] J.V. Burke, A.S. Lewis, M.L. Overton, Approximating subdifferentials by random sampling of gradients, Mathematics of Operations Research 27 (3) (2002) 567–584.
[2] J.M. Borwein, Q.J. Zhu, Techniques of Variational Analysis, Springer-Verlag, New York, 2005.
[3] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation I: Basic Theory, in: Grundlehren der Mathematischen Wissenschaften, vol. 330, Springer, Berlin, 2006.
[4] R.T. Rockafellar, R.J-B. Wets, Variational Analysis, in: Grundlehren der Mathematischen Wissenschaften, vol. 317, Springer, Berlin, 1998.
[5] F.H. Clarke, Yu. Ledyaev, R.I. Stern, P.R. Wolenski, Nonsmooth Analysis and Control Theory, in: Texts in Math., vol. 178, Springer, New York, 1998.
[6] G.J. Minty, Monotone (nonlinear) operators in Hilbert space, Duke Mathematical Journal 29 (1962) 341–346.

[7] R.A. Poliquin, R.T. Rockafellar, Prox-regular functions in variational analysis, Transactions of the American Mathematical Society 348 (1996) 1805–1838.
[8] S.M. Robinson, Equations on monotone graphs, Preprint, 2011.
[9] J.M. Borwein, X. Wang, Lipschitz functions with maximal clarke subdifferentials are generic, Proceedings of the American Mathematical Society 128 (11) (2000) 3221–3229.
[10] J. Benoist, Intégration du sous-différentiel proximal: un contre-exemple, Comptes Rendus de l'Académie des Sciences Série 1, Mathématique 325 (1997) 867–870.
[11] J.M. Borwein, R. Girgensohn, X. Wang, On the construction of Hölder and proximal subderivatives, Canadian Mathematical Bulletin 41 (4) (1998) 497–507.
[12] D. Drusvyatskiy, A.S. Lewis, Semi-algebraic functions have small subdifferentials, Mathematical Programming, Series B (2010) (in press).
[13] M. Shiota, Geometry of Subanalytic and Semialgebraic Sets, Birkhäuser Boston Inc., Cambridge, MA, USA, 1997.
[14] A.D. Ioffe, An invitation to tame optimization, SIAM Journal on Optimization 19 (4) (2009) 1894–1917.
[15] S. Basu, R. Pollack, M. Roy, Algorithms in Real Algebraic Geometry, in: Algorithms and Computation in Mathematics, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
[16] L. van den Dries, Tame Topology and *O*-Minimal Structures, in: LMS Lecture Note Series, vol. 248, Cambridge University Press, Cambridge, 1998.
[17] L. van den Dries, C. Miller, Geometric categories and *O*-minimal structures, Duke Mathematical Journal 84 (1996) 497–540.
[18] M. Coste, An introduction to semialgebraic geometry, Institut de Recherche Mathématiques de Rennes (2002) RAAG Notes, 78 pages.
[19] M. Coste, An introduction to *O*-minimal geometry, Institut de Recherche Mathématiques de Rennes (1999) RAAG Notes, 81 pages.
[20] J.M. Lee, Introduction to Smooth Manifolds, Springer, New York, 2003.
[21] T.R. Rockafellar, Lagrange multipliers and optimality, SIAM Review 35 (2) (1993) 183–238.
[22] J. Bolte, A. Daniilidis, A.S. Lewis, Generic optimality conditions for semi-algebraic convex programs, Mathematics of Operations Research (2011) (in press).