Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 09–38

ANALYSIS OF DIRECT SEARCHES FOR NON-LIPSCHITZIAN FUNCTIONS

L. N. VICENTE AND A. L. CUSTÓDIO

Abstract: It is known that the Clarke generalized directional derivative is nonnegative along the limit directions generated by directional direct-search methods at a limit point of certain subsequences of unsuccessful iterates, if the function being minimized is Lipschitz continuous near the limit point.

In this paper we generalize this result for non-Lipschitzian functions using Rockafellar generalized directional derivatives (upper subderivatives). We show that Rockafellar derivatives are also nonnegative along the limit directions of those subsequences of unsuccessful iterates when the function values converge to the function value at the limit point. This result is obtained assuming that the function is directionally Lipschitzian with respect to the limit direction.

It is also possible under appropriate conditions to establish more insightful results by showing that the sequence of points generated by these methods eventually approaches the limit point along the locally best branch or step function (when the number of steps is equal to two).

The results of this paper are presented for constrained optimization and illustrated numerically.

Keywords: Direct-search methods, directionally Lipschitzian functions, discontinuities, lower semicontinuity, Rockafellar directional derivatives, nonsmooth calculus, Lipschitz extensions.
AMS Subject Classification (2000): 90C30, 90C56.

1. Introduction

In this paper we consider a constrained minimization problem posed as

$$\begin{array}{ll} \min & f(x), \\ \text{s.t.} & x \in \Omega, \end{array} \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ represents a nonsmooth, extended-real-valued objective function and $\Omega \subseteq \mathbb{R}^n$ denotes a nonempty constrained or feasible region. Our interest relies on the solution of problem (1) by derivative-free methods, i.e., by methods which make no use of derivatives of the objective function (or of the functions defining the feasible region), and in particular by direct-search methods.

Date: October 12, 2009.
Support for this work was provided by FCT under the grant PTDC/MAT/64838/2006.

Direct-search methods (DSM) can be classified as either directional or simplicial [6, Chapters 7 and 8]. In this paper we are interested in directional DSM and will consider their iterations organized around a search step (optional) and a poll step. We will essentially concentrate on the poll step since it is responsible for the global convergence properties of the resulting algorithm. A poll step consists of evaluating the objective function at a set of points defined by a positive spanning set or, in some methods different from the ones studied in our paper, defined by a set of positive generators of a cone related to the constraints. A successful poll step occurs when at least one poll point exhibits a function value lower than the current one.

A number of directional DSM consider a finite number of such sets of directions and are referred to as pattern search or generalized pattern search. Although some of our basic results apply to these methods, we will focus on those directional DSM, like mesh adaptive direct search (MADS) [4] and generating set search [9] (under sufficient decrease), which are entitled to use an infinite number of poll directions.

It is possible to prove for these classes of methods the existence of a subsequence of unsuccessful iterates (i.e., unsuccessful poll steps) converging to a limit point and driving the step size or mesh size parameter to zero. At these refining subsequences one can consider limits of normalized poll directions which are then called refining directions. Audet and Dennis [4] have proved that if the objective function is Lipschitz continuous near the limit point x_*, then the Clarke-Jahn directional derivative is nonnegative along an appropriate refining direction v:

$$f_C^\circ(x_*;v) \;=\; \limsup_{\substack{x' \to x_*,\, x' \in \Omega \\ t \downarrow 0,\, x' + tv \in \Omega}} \frac{f(x'+tv) - f(x')}{t} \;\geq\; 0. \tag{2}$$

This derivative is essentially the Clarke generalized directional derivative [5] extended by Jahn [8] to the constrained setting. The refining direction must belong to the interior of the hypertangent cone to Ω at x_*. If the corresponding set of refining directions for x_* is dense in the unit sphere, then these derivatives are proved to be nonnegative for all directions in the tangent cone to Ω at x_* (i.e., for all directions on the closure of the hypertangent cone). A similar result had already been proved for unconstrained optimization and generalized pattern search [3].

In this paper we will show that this result can be extended to the Rockafellar upper subderivative [11] (generalized by us to the constrained case),

$$f^{\uparrow}(x_*; v) \;=\; \limsup_{\substack{x' \to_f x_*,\, x' \in \Omega \\ t \downarrow 0}} \;\; \inf_{\substack{v' \to v \\ x' + tv' \in \Omega}} \; \frac{f(x' + tv') - f(x')}{t} \;\geq\; 0, \quad (3)$$

whenever, at the point x_*, the function f is lower semicontinuous and directionally Lipschitzian with respect to a direction v belonging to the interior of the hypertangent cone $H_\Omega(x_*)$. The notation $x' \to_f x$ represents $x' \to x$ and $f(x') \to f(x)$. The function f is said to be directionally Lipschitzian at x with respect to v if

$$f^{\circ\circ}(x; v) \;=\; \limsup_{\substack{x' \to_f x,\, x' \in \Omega \\ t \downarrow 0}} \;\; \sup_{\substack{v' \to v \\ x' + tv' \in \Omega}} \; \frac{f(x' + tv') - f(x')}{t} \;<\; +\infty.$$

Examples of directionally Lipschitzian functions are given in [11, Section 6].

In this paper we will also show that when f is lower semicontinuous at a point x and directionally Lipschitzian at the point with respect to $v \in \mathrm{int}(H_\Omega(x))$, one has $f^{\uparrow}(x; v) = f^{\circ\circ}(x; v) = f^{\circ}_R(x; v)$, where

$$f^{\circ}_R(x; v) \;=\; \limsup_{\substack{x' \to_f x,\, x' \in \Omega \\ t \downarrow 0,\, x' + tv \in \Omega}} \; \frac{f(x' + tv) - f(x')}{t}. \quad (4)$$

This result was originally established by Rockafellar [11] for the case $\Omega = \mathbb{R}^n$. Also by extending the results of Rockafellar [11] for the constrained setting $\Omega \neq \mathbb{R}^n$, we will show, under appropriate conditions, that the upper directional derivative $f^{\uparrow}(x; v)$, when v is in the tangent cone $T_\Omega(x)$, is the limit inferior of derivatives $f^{\circ}_R(x; w)$ where $w \in \mathrm{int}(H_\Omega(x))$. This analysis will allow us then to state a result for directions in the tangent cone $T_\Omega(x_*)$ but not necessarily in the interior $\mathrm{int}(H_\Omega(x_*))$ of the hypertangent one.

These results apply to discontinuous functions but they do not provide information about the ability of the algorithms to locally identify the best branch or step function. It is possible, however, to prove that the algorithms have the capability to generate an infinite number of iterates in such a step, provided the number of steps is two around the limit point and the function has some continuity properties in each step (essentially the step domains must have nonempty interiors and one must be able to extend the function, in a Lipschitz continuous way, from a step domain to a neighborhood of the limit point).

This paper follows the line of others where nonsmooth calculus (in particular Clarke calculus) has been used to analyze the asymptotic properties of the sequence of iterates generated by DSM of directional type (besides the above cited papers [3, 4], see also [2, 7]).

We organize the material of this paper in the following way. In Section 2 we describe the algorithmic setting for direct-search methods. Then, in Section 3, we gather the necessary material about the globalization strategies that we consider and about the notions of refining subsequences and directions. The main asymptotic results of this paper are contained in Section 4 for functions directionally Lipschitzian with respect to certain directions. We leave to an appendix all the auxiliary nonsmooth calculus background needed for these results. In Section 5 we study the behavior of the algorithm for step discontinuities. We illustrate a number of our results and assumptions in Section 6, numerically and for problems in two dimensions. The paper is concluded in Section 7 with some final remarks.

2. Algorithmic framework

Our algorithmic description follows the one in [6, Chapter 7] for the unconstrained case. This framework will encompass both the MADS methodology (based on integer lattices and where a simple decrease on the function value suffices to identify a new iterate) and general directional DSM based on randomly generated normalized directions and sufficient decrease for acceptance of new iterates. Each iteration of the algorithm is organized around a search step (optional) and a poll step. The evaluation process of the poll step is opportunistic moving to a poll point once simple or sufficient decrease is found, depending on the variant being used.

As we will see later in the convergence theory, the set of directions used for polling is not required to positively span $\mathbb{R}^n$ (although for coherence with the smooth case we will write it so in the algorithm below) and not necessarily drawn from a finite set of directions. The algorithm requires an initial feasible point with finite objective function value.

To make the algorithmic description shorter we will make use of the extreme barrier function

$$f_\Omega(x) = \begin{cases} f(x) & \text{if } x \in \Omega, \\ +\infty & \text{otherwise.} \end{cases}$$

Following the terminology in [9], $\rho : (0, +\infty) \to (0, +\infty)$ will represent a forcing function, i.e., a continuous and non increasing function satisfying $\rho(t)/t \to 0$ when $t \downarrow 0$. Typical examples of forcing functions are $\rho(t) = t^{1+a}$, for $a > 0$. To write the algorithm in general terms we will use $\bar{\rho}(\cdot)$ to either represent a forcing function $\rho(\cdot)$ or the constant, zero function.

Algorithm 2.1 (Directional direct-search method).

Initialization

Choose $x_0 \in \Omega$ with $f(x_0) < +\infty$, $\alpha_0 > 0$, $0 < \beta_1 \leq \beta_2 < 1$, and $\gamma \geq 1$. Let $\mathcal{D}$ be a (possibly infinite) set of positive spanning sets.

For $k = 0, 1, 2, \ldots$

(1) **Search step:** Try to compute a point with $f_\Omega(x) < f(x_k) - \bar{\rho}(\alpha_k)$ by evaluating the function f at a finite number of points. If such a point is found then set $x_{k+1} = x$, declare the iteration and the search step successful, and skip the poll step.

(2) **Poll step:** Choose a positive spanning set D_k from the set $\mathcal{D}$. Order the set of poll points $P_k = \{x_k + \alpha_k d : d \in D_k\}$. Start evaluating f_Ω at the poll points following the chosen order. If a poll point $x_k + \alpha_k d_k$ is found such that $f_\Omega(x_k + \alpha_k d_k) < f(x_k) - \bar{\rho}(\alpha_k)$ then stop polling, set $x_{k+1} = x_k + \alpha_k d_k$, and declare the iteration and the poll step successful. Otherwise declare the iteration (and the poll step) unsuccessful and set $x_{k+1} = x_k$.

(3) **Mesh parameter update:** If the iteration was successful then maintain or increase the step size parameter: $\alpha_{k+1} \in [\alpha_k, \gamma\alpha_k]$. Otherwise decrease the step size parameter: $\alpha_{k+1} \in [\beta_1\alpha_k, \beta_2\alpha_k]$.

3. Behavior of the step size

The global convergence of directional DSM is heavily based on the analysis of the behavior of the step size parameter α_k which must approach zero as an indication of some form of stationarity. There are essentially two known ways of enforcing the existence of a subsequence of step size parameters converging to zero in DSM of directional type. One way is by ensuring that all new iterates lie on an integer lattice (rigorously speaking only when the step size is bounded away from zero). The other form consists of imposing a sufficient decrease on the acceptance of new iterates. In the former case we need the iterates to lie in a bounded set and in the latter situation the objective function must be bounded below.

Assumption 3.1. *The level set* $L(x_0) = \{x \in \Omega : \ f(x) \leq f(x_0)\}$ *is bounded. The function* f *is bounded below in* $L(x_0)$*.*

3.1. Integer lattices (MADS). Generalized pattern search makes use of a finite set of directions $\mathcal{D} = D$ which satisfy appropriate integrality requirements for globalization by integer lattices.

Assumption 3.2. *The set* D *of positive spanning sets is finite and the elements of* D *are of the form* $G\bar{z}_j$*,* $j = 1, \ldots, |D|$*, where* $G \in \mathbb{R}^{n \times n}$ *is a nonsingular matrix and each* $\bar{z}_j$ *is a vector in* $\mathbb{Z}^n$*.*

Given the type of non-smoothness and discontinuity of the objective function which we would like to consider in this paper, we need to make use of an infinite set of directions $\mathcal{D}$ dense (after normalization) in the unit sphere. MADS makes use of such a set of directions but, since it is also based on globalization by integer lattices, the set $\mathcal{D}$ must then be generated from a finite set D satisfying Assumption 3.2 (which will be guaranteed by the first requirement of the next assumption).

Assumption 3.3. *Let* D *represent a finite set of positive spanning sets satisfying Assumption 3.2.*

The set $\mathcal{D}$ *is so that the elements* $d_k \in D_k$ *satisfy the following conditions:*

1. d_k *is a nonnegative integer combination of the columns of* D*.*
2. *The distance between* x_k *and the point* $x_k + \alpha_k d_k$ *tends to zero if and only if* α_k *does:*

$$\lim_{k \in K} \alpha_k \|d_k\| \;=\; 0 \iff \lim_{k \in K} \alpha_k \;=\; 0,$$

for any infinite subsequence K*.*

3. *The limits of all convergent subsequences of* $\bar{D}_k = \{d_k / \|d_k\| : \ d_k \in D_k\}$ *are positive spanning sets for* $\mathbb{R}^n$*.*

In addition, the update of the step size parameter must conform to some form of integrality.

Assumption 3.4. *The step size parameter is updated as follows: Choose a rational number $\tau > 1$, a nonnegative integer $m^+ \geq 0$, and a negative integer $m^- \leq -1$. If the iteration is successful, the step size parameter is maintained or increased by taking $\alpha_{k+1} = \tau^{m_k^+}\alpha_k$, with $m_k^+ \in \{0, \ldots, m^+\}$. Otherwise, the step size parameter is decreased by setting $\alpha_{k+1} = \tau^{m_k^-}\alpha_k$, with $m_k^- \in \{m^-, \ldots, -1\}$.*

Note that these rules respect those of Algorithm 2.1 by setting $\beta_1 = \tau^{m^-}$, $\beta_2 = \tau^{-1}$, and $\gamma = \tau^{m^+}$.

Finally, the search step is restricted to points in a previously (implicitly defined) mesh or grid.

Assumption 3.5. *The search step in Algorithm 2.1 only evaluates points in*

$$M_k \;=\; \bigcup_{x \in S_k} \{x + \alpha_k D z : z \in \mathbb{N}_0^{|D|}\},$$

where S_k is the set of all the points evaluated by the algorithm previously to iteration k.

The following result was originally proved by Torczon [12] for pattern search and extended later to generalized pattern search [3] and MADS [4].

Theorem 3.1. *Let Assumption 3.1 hold. Algorithm 2.1 under Assumptions 3.3–3.5 and $\bar{\rho}(\cdot) = 0$ (MADS) generates a sequence of iterates satisfying*

$$\liminf_{k \to +\infty} \alpha_k \;=\; 0.$$

3.2. Sufficient decrease. An alternative to the use of integer lattices is to impose sufficient rather than simple decrease as a criterion for accepting new iterates. This can be simply achieved by selecting $\bar{\rho}(\cdot)$ as a forcing function in Algorithm 2.1. The following result is relatively classic in nonlinear optimization when using some form of sufficient decrease. It is proved in [9] and in [6, Section 7.7] in the context of directional DSM for unconstrained optimization.

Theorem 3.2. *Let Assumption 3.1 hold. Algorithm 2.1, when $\bar{\rho}(\cdot)$ is a forcing function, generates a sequence of iterates satisfying*

$$\liminf_{k \to +\infty} \alpha_k \;=\; 0.$$

Note that such a result is derived under no assumptions whatsoever on the set of directions $\mathcal{D}$. Thus, we are free to use a normalized set of directions $\mathcal{D}$ dense in the unit sphere.

3.3. Refining subsequences and directions. The type of stationarity results which can be derived for DSM of directional type are established at limit points of the so-called refining subsequences (a concept formalized in [3]).

Definition 3.1. *A subsequence $\{x_k\}_{k \in K}$ of iterates corresponding to unsuccessful poll steps is said to be a refining subsequence if $\{\alpha_k\}_{k \in K}$ converges to zero.*

One can ensure for the two algorithmic settings of this paper (Sections 3.1 and 3.2) the existence of a convergent refining subsequence. Such a result is a simple and known consequence of Assumption 3.1, Theorems 3.1 or 3.2, and the scheme that updates the step size parameter (see, e.g., [6, Section 7.3]).

Theorem 3.3. *Let Assumption 3.1 hold. Consider a sequence of iterates generated by Algorithm 2.1 under the scenarios of either Section 3.1 (MADS) or Section 3.2 (sufficient decrease). Then there is at least a convergent refining subsequence $\{x_k\}_{k \in K}$.*

The type of directions along which appropriate directional derivatives will be proved nonnegative are the so-called refining directions (a notion formalized in [4]).

Definition 3.2. *Let x_* be the limit point of a convergent refining subsequence. If the limit $\lim_{k \in L} d_k/\|d_k\|$ exists, where $L \subseteq K$ and $d_k \in D_k$, and if $x_k + \alpha_k d_k \in \Omega$, for sufficiently large $k \in L$, then this limit is said to be a refining direction for x_*.*

Some of the results of this paper will require for the refining subsequence under consideration that the associated set of refining directions for x_* is dense in the unit sphere (an assumption stronger than just saying that the normalized set of directions $\mathcal{D}$ is dense in the unit sphere).

4. Results for directionally Lipschitzian functions

Our first convergence result addresses the case where a refining direction is in the interior of the hypertangent cone to Ω at the limit point.

Definition 4.1. *A vector v is said to be hypertangent to Ω at $x \in \Omega$ if there exists an $\epsilon > 0$ such that*

$$x' + tv \in \Omega \quad \textit{for all} \quad x' \in \Omega \cap B(x;\epsilon) \quad \textit{and} \quad t \in (0,\epsilon).$$

The set of all hypertangent vectors to Ω at x is called the hypertangent cone to Ω at x and is represented by $H_\Omega(x)$.

The hypertangent cone is convex and contains the origin (see [11, Corollary 2]).

It seems that there is no universal definition of hypertangency in the literature. In the paper [4], where MADS was proposed, the authors have used a different definition. In their case, a vector v is said to be hypertangent to Ω at x if there exists an $\epsilon > 0$ such that

$$x' + tv' \in \Omega \quad \text{for all} \quad x' \in \Omega \cap B(x;\epsilon), \quad v' \in B(v;\epsilon), \quad \text{and} \quad t \in (0,\epsilon).$$

Rockafellar [11] calls these vectors those with respect to Ω is epi-Lipschitzian at x and proves (a fact that results from [11, Corollary 2], see also Proposition A.2 in the Appendix of our paper) that they coincide with those in the interior of the hypertangent cone of Definition 4.1. In summary, the interior of the hypertangent cone of Definition 4.1 is the hypertangent cone used in [4]. We chose to follow Definition 4.1 since this was the one adopted by Rockafellar [11] and our analysis is heavily based on his.

As we have seen before, the existence of a convergent refining subsequence $\{x_k\}_{k\in K}$ is guaranteed by Theorem 3.3. It is then possible to state this condition as an assumption for deriving asymptotic results at limit points.

Theorem 4.1. *Consider a refining subsequence $\{x_k\}_{k\in K}$ converging to $x_* \in \Omega$ and a refining direction v for x_* in $\mathrm{int}(H_\Omega(x_*))$. Assume that f is lower semicontinuous at x_* and directionally Lipschitzian at x_* with respect to v. Assume further that $\lim_{k\in K} f(x_k) = f(x_*)$. Then $f^\uparrow(x_*;v) = f^{\circ\circ}(x_*;v) = f^\circ_R(x_*;v) \geq 0$.*

Proof: Since f is directionally Lipschitzian at x_* with respect to v, we have that $f^{\circ\circ}(x_*; v) < +\infty$. Now let $v = \lim_{k \in L} d_k/\|d_k\|$, with $L \subseteq K$. Thus,

$$
\begin{aligned}
f^{\circ\circ}(x_*; v) &= \limsup_{\substack{x' \to_f x_*,\, x' \in \Omega \\ t \downarrow 0}} \ \sup_{\substack{v' \to v \\ x' + tv' \in \Omega}} \frac{f(x' + tv') - f(x')}{t} \\
&\geq \limsup_{k \in L} \frac{f(x_k + \alpha_k \|d_k\| (d_k/\|d_k\|)) - f(x_k)}{\alpha_k \|d_k\|} \\
&= \limsup_{k \in L} \frac{f(x_k + \alpha_k d_k) - f(x_k) + \bar{\rho}(\alpha_k)}{\alpha_k \|d_k\|} - \frac{\bar{\rho}(\alpha_k)}{\alpha_k \|d_k\|} \ \geq \ 0.
\end{aligned}
$$

The first inequality follows from $\{x_k\}_{k \in L}$ being a feasible refining subsequence with $\lim_{k \in L} f(x_k) = f(x_*)$ and the fact that $x_k + \alpha_k d_k$ is feasible for $k \in L$ sufficiently large. The limit $\lim_{k \in L} \bar{\rho}(\alpha_k)/(\alpha_k \|d_k\|)$ is 0 for both globalization strategies (Sections 3.1 and 3.2). In the case of MADS (Section 3.1), one uses $\bar{\rho}(\cdot) = 0$. When imposing sufficient decrease (Section 3.2), since $\|d_k\| = 1$ for all k, this limit follows from the properties of the forcing function.

The fact that $f^{\uparrow}(x_*; v) = f^{\circ\circ}(x_*; v) = f_R^{\circ}(x_*; v)$ is showed in the Appendix (Theorem A.1). ■

Now we need to address the case where the directions are in the tangent cone to Ω at the limit point but not necessarily in its interior.

Definition 4.2. *A vector v is said to be tangent to Ω at x if for all sequences $\{y_k\} \subset \Omega$ converging to x and for all sequences $\{t_k\}$ with $t_k \downarrow 0$, there exists a sequence of vectors $\{w_k\}$ converging to v such that $y_k + t_k w_k \in \Omega$ for all k.*

The set of all tangent vectors to Ω at x is called the tangent cone to Ω at x and is represented by $T_\Omega(x)$.

The tangent cone $T_\Omega(x)$ is the closure of both $H_\Omega(x)$ and $\mathrm{int}(H_\Omega(x))$. It can be also defined by the limit inferior of a multifunction (see the Appendix for details).

Theorem 4.2. *Consider a refining subsequence $\{x_k\}_{k \in K}$ converging to $x_* \in \Omega$. Let v be in $T_\Omega(x_*)$ (but not necessarily in* $\mathrm{int}(H_\Omega(x_*))$*). Assume that f is lower semicontinuous at x_* and directionally Lipschitzian at x_* with respect to v. Assume further that $\lim_{k \in K} f(x_k) = f(x_*)$.*

If f is directionally Lipschitzian with respect to all directions in the intersection of a ball centered at v with $\mathrm{int}(H_\Omega(x_*))$ *and the set of refining directions for x_* is dense in this intersection, then $f^\uparrow(x_*;v) \geq 0$.*

Proof: First we apply Theorem 4.1 to obtain that $f_R^\circ(x_*;w) \geq 0$ for all the refining directions w in the intersection of the ball centered at v with $\mathrm{int}(H_\Omega(x_*))$. Then, from the result proved in the Appendix (Theorem A.1)

$$f^\uparrow(x_*;v) \;=\; \liminf_{\substack{w \to v \\ w \in \mathrm{int}(H_\Omega(x_*))}} f_R^\circ(x_*;w) \;\geq\; 0.$$

■

5. Results for non-Lipschitzian functions

We are now interested in studying the behavior of directional DSM when the objective function is defined by several branches or steps, in particular to know if the algorithm can identify the locally best step. We will give an affirmative answer provided the number of steps is two, the step domains have nonempty interiors, and their borders exhibit a minimum of regularity (the exterior cone property).

The condition stated below covers a wide range of discontinuities.

Assumption 5.1. *The function f is such that there exists a neighborhood B of x_* (a limit point of a refining subsequence) which admits a finite partition*

$$B \;=\; \bigcup_{i=1}^{n_B} B_i,$$

such that, for all $i \in \{1,\ldots,n_B\}$,

(1) $\mathrm{int}(B_i) \neq \emptyset$,
(2) $\mathrm{cl}(B_i)$ *has the exterior cone property (see Definition 5.1),*
(3) f *is Lipschitz continuous in* $\mathrm{int}(B_i)$ *and can be continuously extended to ∂B_i.*

It can be easily seen that if we extend a Lipschitz continuous function in the interior of a set to the boundary, in a continuous way, the extension is also Lipschitz continuous on the closure of the set.

Proposition 5.1. *Let f be a locally Lipschitz function in* $\mathrm{int}(S)$, *with Lipschitz constant L. The continuous extension of f to* $\mathrm{cl}(S)$ *is locally Lipschitz continuous with Lipschitz constant $3L+2$.*

Proof: Let $x \in \partial S$ and consider a neighborhood N of x. The value of the extended function in any point $y \in \partial S \cap N$ can be given as the limit of $\{f(y^j)\}$ for any sequence of points $\{y^j\} \subseteq \mathrm{int}(S) \cap N$ converging to y.

Let us consider two points z and w in $\partial S \cap N$. Let $\epsilon \leq \|z - w\|$. Then, there exist $z_\epsilon, w_\epsilon \in \mathrm{int}(S) \cap N$, with $\|z - z_\epsilon\| \leq \epsilon$ and $\|w - w_\epsilon\| \leq \epsilon$, such that:

$$\begin{aligned} |f(z) - f(w)| &\leq |f(z) - f(z_\epsilon)| + |f(z_\epsilon) - f(w_\epsilon)| + |f(w_\epsilon) - f(w)| \\ &\leq \epsilon + L\|z_\epsilon - w_\epsilon\| + \epsilon \ \leq \epsilon + 3L\|z - w\| + \epsilon \\ &\leq (2 + 3L)\|z - w\|. \end{aligned}$$

The case where one point is in $\mathrm{int}(S) \cap N$ and the other in $\partial S \cap N$ can be proved analogously (the result is trivial when both points are in $\mathrm{int}(S) \cap N$). ■

The precise form of the exterior cone property which we will use is stated below.

Definition 5.1. *A set S has the exterior cone property if at any point $z \in \partial S$ there exist a cone C_z with nonempty interior, an angle $\theta_z > 0$, and a neighborhood N_z of z such that $E_z = N_z \cap \{z' = z + c, c \in C_z, c \neq 0\} \subset \mathbb{R}^n \backslash S$ and the angle between all the vectors in $E_z - \{z\}$ and all the vectors in $S_z - \{z\}$, with $S_z = S \cap N$, is larger than θ_z.*

We will also need the following auxiliary result.

Proposition 5.2. *Let S be a set with the exterior cone property and g a function Lipschitz continuous in S. Let also $z \in \partial S \cap S$.*

Then there exists an extension $\tilde{g}$ of g from S to $\mathbb{R}^n$ which is Lipschitz continuous in $\mathbb{R}^n$ and locally strictly decreasing along all directions emanating from z and belonging to a cone with nonempty interior.

Proof: Consider the sets C_z, N_z, E_z, and S_z as in Definition 5.1. Define an auxiliary function $\tilde{g}$ which coincides with g in S_z and is linear and strictly decreasing from z to the interior of E_z. We will show first that this function is Lipschitz continuous in $S_z \cup E_z$. Let L_1 be the Lipschitz constant of $g = \tilde{g}$ in S_z and L_2 the Lipschitz constant of $\tilde{g}$ in E_z. Now consider a point $y \in S_z$

and a point $w \in E_z$. One can derive that

$$\begin{array}{rcl} |\tilde{g}(y) - \tilde{g}(w)| & \leq & |g(y) - g(z)| + |\tilde{g}(w) - \tilde{g}(z)| \\ & \leq & \max\{L_1, L_2\} \left(\|y - z\| + \|w - z\|\right) \\ & \leq & \max\{L_1, L_2\} M \|y - w\|, \end{array}$$

where the last equality follows from $\|y - z\| + \|w - z\| \leq M\|y - w\|$, with $M = \sqrt{2}\pi/\theta_z$. The cases where both points lie in S_z or E_z are straightforward to analyze. We then obtain that $|\tilde{g}(y) - \tilde{g}(w)| \leq \max\{L_1, L_2\} \max\{1, M\}\|y - w\|$, for all y and w in $S_z \cup E_z$.

It is known that any Lipschitz function in a set can be extended to the whole space with the same Lipschitz constant (see [10, Theorem 1]). Thus, one can now extend $\tilde{g}$ from $S_z \cup E_z$ to $\mathbb{R}^n$, and in particular to N_z, with the same Lipschitz constant. ■

We are now ready for the main result of this section. Recall that the existence of a convergent refining subsequence $\{x_k\}_{k \in K}$ is guaranteed by Theorem 3.3.

Theorem 5.1. *Consider a refining subsequence* $\{x_k\}_{k \in K}$ *converging to* $x_* \in \Omega$. *Assume that* f *is lower semicontinuous at* x_* *and satisfies Assumption 5.1. Let the corresponding set of refining directions for* x_* *be dense in the unit sphere.*

If x_* *belongs to the interior of a partition set in* $\{B_1, \ldots, B_{n_B}\}$, *then* $f_C^\circ(x_*; v) \geq 0$ *for all refining directions* $v \in T_\Omega(x_*)$.

Otherwise, there exists a subsequence $K' \subset K$ *and a partition set* $B' \in \{B_1, \ldots, B_{n_B}\}$ *such that* $\{x_k\}_{k \in K'} \subset \mathrm{cl}(B')$ *and there is an infinite number of poll points corresponding to iterates in* K' *belonging to both* $\mathrm{int}(B')$ *and* $\mathbb{R}^n \setminus \mathrm{cl}(B')$.

Proof: The proof is done for the case of MADS (Section 3.1) but the case of sufficient decrease (Section 3.2) is obtained from this one with minor modifications.

Consider first the neighborhood B guaranteed by Assumption 5.1. If x_* belongs to $\mathrm{int}(B_l)$, for some $l \in \{1, \ldots, n_B\}$, the Lipschitz continuity of f near x_* would allows to apply the known results from [4].

So, let us assume that x_* belongs to the boundary of B_l, for some $l \in \{1, \ldots, n_B\}$. Since the partition is finite, by passing at a subsequence $K_1 \subset K$ if necessary, one can state the existence of an $i \in \{1, \ldots, n_B\}$ such that $x_* \in \partial B_i$ and $\{x_k\}_{k \in K_1} \subset \mathrm{cl}(B_i)$ with $K_1 \subset K$.

By using Assumption 5.1 and Proposition 5.1, we can extend f from B_i to $\mathrm{cl}(B_i)$ in a continuous way and ensuring that the extended function $\bar{f}$ is Lipschitz continuous in $\mathrm{cl}(B_i)$.

Let us assume that all poll points associated with the refining subsequence belong to $\mathrm{cl}(B_i)$. We will see that this leads us to a contradiction. So, let us assume that there exists a $\bar{k} \in K_1$ such that $x_k + \alpha_k d \in \mathrm{cl}(B_i)$ for all $k \in K_1$ with $k \geq \bar{k}$ and for all $d \in D_k$. We now apply Proposition 5.2 using $g = \bar{f}$ and $S = \mathrm{cl}(B_i)$. Let $\tilde{f}$ be the extended function (and L its Lipschitz constant). We then obtain that

$$
\begin{aligned}
& \tilde{f}_C^\circ(x_*; v) \\
& \quad \geq \limsup_{\substack{k \to +\infty \\ k \in K_1}} \frac{\tilde{f}(x_k + \alpha_k d_k) - \tilde{f}(x_k)}{\alpha_k \|d_k\|} + \frac{\tilde{f}(x_k + \alpha_k \|d_k\| v) - \tilde{f}(x_k + \alpha_k d_k)}{\alpha_k \|d_k\|} \\
& \quad \geq \limsup_{\substack{k \to +\infty \\ k \in K_1}} \frac{f(x_k + \alpha_k\, d_k) - f(x_k)}{\alpha_k \|d_k\|} - \frac{L \alpha_k \|d_k\| \|v - (d_k / \|d_k\|)\|}{\alpha_k \|d_k\|} \\
& \quad \geq 0,
\end{aligned}
$$

for all refining directions v, which is a contradiction since these directions are dense in the unit sphere and $\tilde{f}$ is locally strictly decreasing from x_* along all directions in a cone of nonempty interior.

So, one can build a sequence of points $K_2 \subset K_1$ for which there exists $d_k \in D_k$ such that

$$
f_\Omega(x_k + \alpha_k d_k) \;\geq\; f(x_k), \quad x_k + \alpha_k d_k \notin \mathrm{cl}(B_i), \tag{5}
$$

for all $k \in K_2$.

Given that B_i has a nonempty interior and that the set of refining directions for x_* is dense in the unit sphere, there must exist an infinite number of poll points associated with a subsequence $K_3 \subset K_2$ belonging to $\mathrm{int}(B_i)$.

The proof is completed by setting $B' = B_i$ and $K' = K_3$. ∎

When the number of steps is equal to two ($n_B = 2$) it is possible to prove a stronger result.

Corollary 5.1. *Under the assumptions of Theorem 5.1 and when $n_B = 2$, there exists a subsequence $K_* \subset K$ and a partition set $B_* \in \{B_1, B_2\}$ such that, when x_* is in the border of the two partition sets,*

(1) *B_* satisfies the properties stated for B' in Theorem 5.1,*
(2) *B_* is the partition set where the lowest values of f are attained around x_*,*
(3) $\lim_{k \in K_*} f(x_k) = f(x_*)$,
(4) *the function f can be extended from B_* to a neighborhood of x_* in a Lipschitz continuously way so that $F_C^\circ(x_*; v) \geq 0$ for all refining directions $v \in T_\Omega(x_*)$, where F denotes the extended function (in particular this property holds for all refining directions $v \in T_{\Omega \cap B_*}(x_*)$).*

Proof: The proof of the corollary is a continuation of the proof of the Theorem 5.1.

First we note that $\{f(x_k)\}$ is decreasing and bounded below and thus it converges, say to f_*. Since f is lower semicontinuous, $f(x_*) \leq f_*$.

We can now show that it is along B' (see the proof of Theorem 5.1) that the value of f is attained, i.e., that $f_* = \lim_{k \in K_3} f(x_k) = f(x_*)$. If this was not true, then there would exist an $\epsilon > 0$ and a bordering B'' (since $n_B = 2$, the remaining one) and a neighborhood N of x_* for which $f(y) > f(z) + \epsilon$, for all $y \in B' \cap N \cap \Omega$ and $z \in B'' \cap N \cap \Omega$. But this contradicts (5).

From (5), we also conclude that $B_* = B'$ is the partition set where f attains the lowest feasible values around x_*.

Finally, the last point of the corollary can be easily proved by extending f from $\mathrm{cl}(B_*)$ to B in a Lipschitz continuous form. ■

Other forms of results could have been stated at the fourth point of this corollary. For instance, we could have mixed here the analysis of this and the previous section, and consider vectors with respect to which the objective function is directionally Lipschitzian at x_* and then restrict these vectors to the hypertangent and tangent cones to $\Omega \cap B_*$ at this point (as in Theorems 4.1 and 4.2).

6. Numerical illustrations

To illustrate the ability of Algorithm 2.1 in finding local minimizers for lower semicontinuous functions we ran some examples in MATLAB. We included examples which violate some of the assumptions required to ensure convergence. Five problems of the form (1) were considered, where $\Omega = [-1, 1] \times [-1, 1]$ was partitioned into a finite number of disjoint subsets $\Omega = \bigcup_{i=1}^{n_B} \Omega_i$, with $n_B = 2$ in four of the cases and $n_B = 4$ in the last problem. The minimizer, when exists, is unique and corresponds to $x_* = (0, 0)$. Figures 1-5 depict plots of each one of the functions considered.

Problem 1: A lower semicontinuous function of the form

$$f_1(x) \;=\; \begin{cases} x^2+y^2 & \text{if } \frac{x}{2} \leq y \leq 2x, \\ 10+x^2+y^2 & \text{otherwise,} \end{cases}$$

and steps Ω_1 and Ω_2 with nonempty interior. Near x_*, the distance between function values in the two steps remains constant.

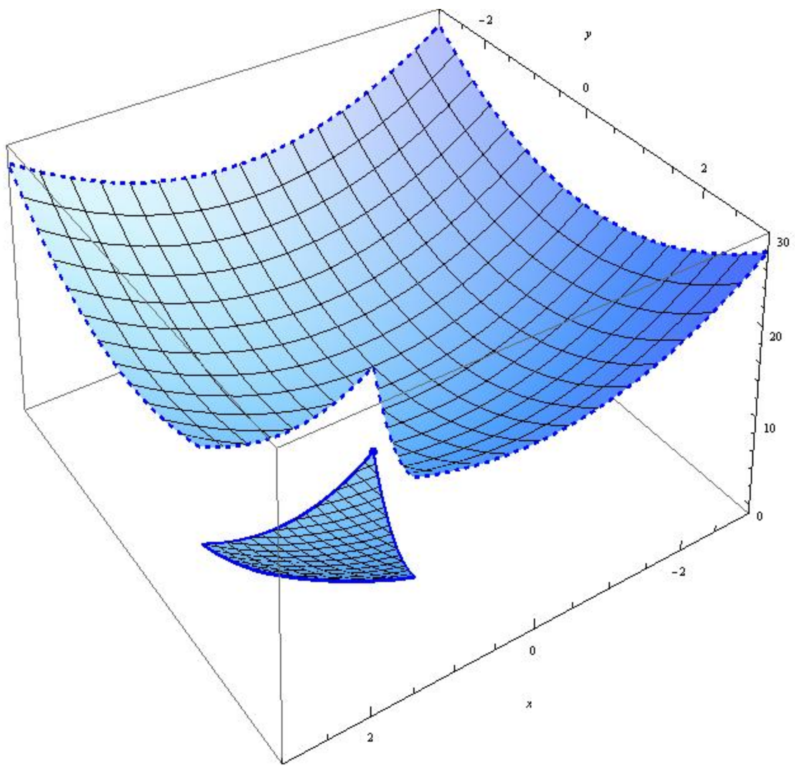

FIGURE 1. Plot of function f_1.

Problem 2: A lower semicontinuous function of the form

$$f_2(x) = \begin{cases} 10x^2+10y^2 & \text{if } x<0, \\ 10x^2+y^2 & \text{otherwise,} \end{cases}$$

and steps Ω_1 and Ω_2 with nonempty interior. Near x_*, the distance between function values in the steps converges to zero.

Problem 3: An upper semicontinuous function of the form

$$f_3(x) \;=\; \begin{cases} x^2+y^2 & \text{if } \frac{x}{2} < y < 2x, \\ 10+x^2+y^2 & \text{otherwise,} \end{cases}$$

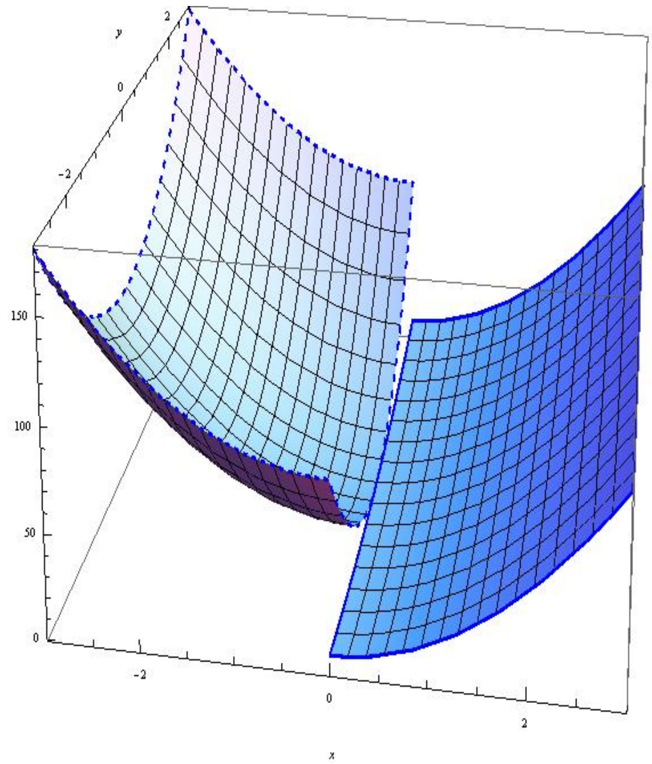

FIGURE 2. Plot of function f_2.

and steps Ω_1 and Ω_2 with nonempty interior. The upper semicontinuity, prevents $x_* = (0, 0)$ from being considered as a local minimizer.

Problem 4: A lower semicontinuous function of the form

$$f_4(x) = \begin{cases} x^2 + y^2 & \text{if } y = 2x, \\ 10 + x^2 + y^2 & \text{otherwise,} \end{cases}$$

where one of the steps has empty interior.

Problem 5: A lower semicontinuous function of the form

$$f_5(x) = \begin{cases} x^2 + y^2 & \text{if } \frac{x}{2} \leq y \leq 2x, \\ 5 + x^2 + y^2 & \text{if } x \leq 0 \wedge y \leq 0 \wedge (x, y) \neq (0, 0), \\ 10 + x^2 + y^2 & \text{if } y < \frac{x}{2} \wedge x > 0, \\ 15 + x^2 + y^2 & \text{otherwise,} \end{cases}$$

and steps Ω_i, $i \in \{1, 2, 3, 4\}$, with nonempty interior. Near x_*, the distance between function values in any of the steps remains constant. The number of steps considered exceeds two, which violates one of the conditions required in Section 5 to establish the asymptotic results.

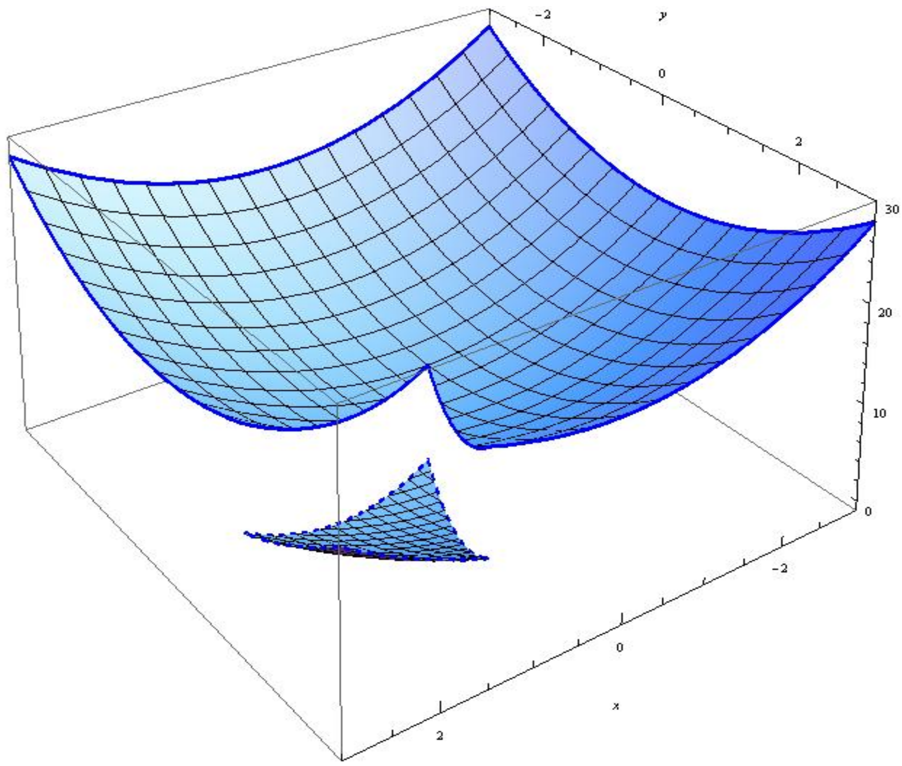

FIGURE 3. Plot of function f_3.

We tested `NOMADm` [1], version 4.6, a Matlab implementation of MADS (which fits into Algorithm 2.1, see Section 3.1) and a very simple implementation of Algorithm 2.1 with a globalization strategy based on sufficient decrease (as in Section 3.2). In both algorithms, the search step was empty, the initial step size parameter was set to one, and the run stopped once the step size reached the threshold 10^{-7}. In the implementation of the variant which requires sufficient decrease, the poll set D_k was set equal to $[Q_k\ -Q_k]$, where Q_k is an orthogonal matrix computed by randomly generating the first column. In MADS, the positive spanning set considered corresponds to the implementation LTMADS, with a total of $2n$ directions. Since our main concern is (proper) convergence rather than efficiency, the poll points were evaluated following the consecutive order of storage. As a forcing function, in the case of the sufficient decrease variant, we considered $\rho(t) = t^2$ (other variants were tested, but with worse results in what concerns the total number of function evaluations required).

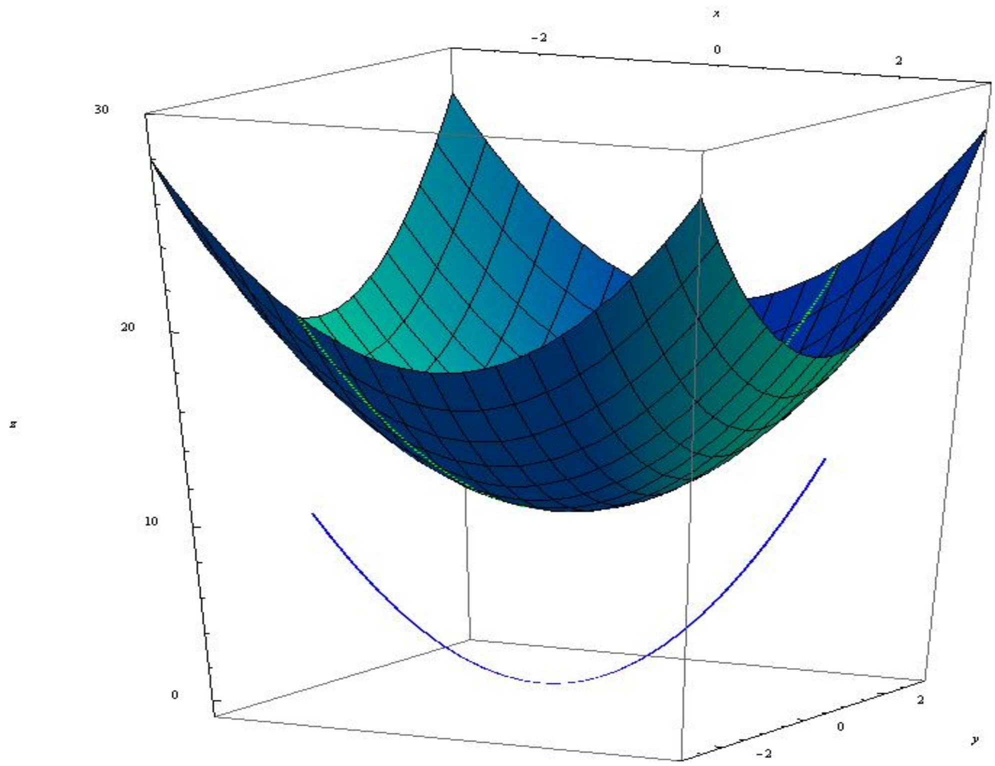

FIGURE 4. Plot of function f_4.

Given the random behavior of both algorithms, a sequence of 10 runs was considered for each problem. The initial point was set to $x_0 = (-0.4, -0.5)$. A summary of the computational experiments is reported in Table 1.

	Algorithm 2.1			
	MADS		**Suff. Decrease**	
function	**#failures**	**#fevals**	**#failures**	**#fevals**
f_1	0	233	0	175.6
f_2	0	193.9	0	494.6
f_3	0	228.8	0	175.6
f_4	10	173.7	10	144.4
f_5	2	220.6	1	177.7

TABLE 1. Number of failures in identifying a local minimizer and corresponding average number of function evaluations required. See the text for a description about the type of successes obtained for f_3.

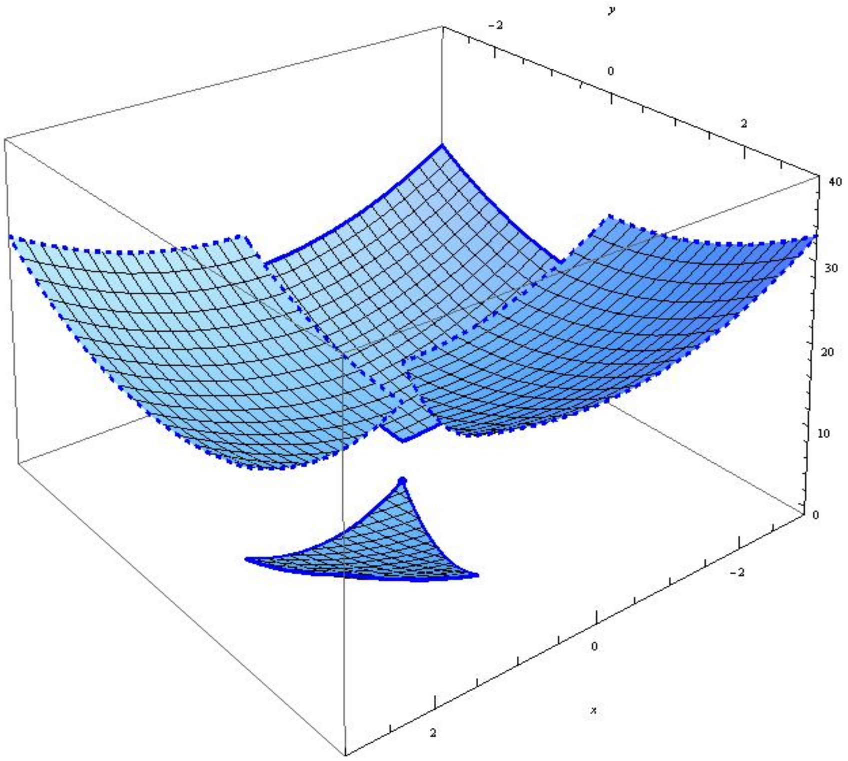

FIGURE 5. Plot of function f_5.

When any of the algorithms failed to converge to the function minimizer, the final iterate corresponded generally to a point near the minimizer of the function when restricted to a higher step. The exception occurred with f_4 where there were cases of convergence to points lying on the line of discontinuity.

In any of the 10 runs performed for each algorithm, the final iterate computed for the function f_3 corresponded to a function value near zero. Even if this point can be considered as a good numerical value for minimization purposes, we note that it is not the function minimizer, given the upper semicontinuity. By changing the initial point provided to the optimizer to $x_0 = (-0.5, -0.5)$, we observed cases of convergence to the minimum value of f_3 when restricted to the highest step (a behavior similar to what was noticed for the other functions in case of failure of convergence).

In order to access the dependency of the results from the initial point provided to the methods, we considered a grid of 100 points, equally spaced in $\Omega = [-1, 1] \times [-1, 1]$, and run Algorithm 2.1 with a globalization strategy based on sufficient decrease. For each of the five problems and for each of the initial points, we ran the algorithm 10 times, yielding a total of 1000 runs for each problem. The number of failures in detecting the minimizer is reported in Table 2. It is possible to state with some level of certainty that there were basically no failures for the first three problems and that the failures observed in Table 1 for last two were not by chance.

function	**#failures**
f_1	2
f_2	0
f_3	2
f_4	1000
f_5	61

Table 2. Number of failures of Algorithm 2.1 (sufficient decrease variant, see Section 3.2), for a sequence of 1000 runs, starting from different initial points. See the text for a description about the type of successes obtained for f_3.

The numerical results support the theoretical analysis developed in the previous sections. Failures in locating the minimizer occur only when at least one of the assumptions required for establishing convergence is violated.

7. Final remarks

In this paper we tried to shed some light on the convergence properties of direct-search methods (DSM) of directional type for lower semicontinuous functions not necessarily Lipschitz continuous. We divided our analysis into two main parts. In the first part, we derive results for refining directions with respect to which the function is directionally Lipschitzian at the limit point of the underlying refining subsequence. These results were derived for the constrained case which forced us to redo the analysis in [11] for upper subderivatives in the presence of constraints.

In the second part of the analysis, we considered a class of discontinuous functions and showed that when the number of branches or steps is two and the function has some continuity properties in each step, these DSM identify the best local step around the limit point. The problem in extending this result to more than two local steps or branches lies on the fact that the speed at which the poll points approach the border of a step domain can be slower than the speed at which these points approach the iterates. We were able to prove, by extending continuously the function and taking limits, that an infinite number of poll points jump out of the step domain. However, they could only visit a neighbor step and thus one can only infer results when the number of steps is equal to two.

Appendix A

In this section we provide the rigorous definitions of the various generalized directional derivatives used throughout this paper.

Definition of upper subderivative. The upper subderivative (3) was defined by Rockafellar [11] for the case $\Omega = \mathbb{R}^n$. To extend it to the constrained case $\Omega \neq \mathbb{R}^n$, let $g(s, y)$ be an extended-real-valued function defined on $(\mathbb{R}^n \times \mathbb{R} \times [0, +\infty)) \times \mathbb{R}^n$. Let also $s \in S \subset \mathbb{R}^n \times \mathbb{R} \times [0, +\infty)$. Define

$$h(s, y) \;=\; \limsup_{\substack{s' \to s \\ s' \in S}} \; \inf_{\substack{y' \to y \\ y' \in \Gamma_\Omega(p(s'))}} \; g(s', y')$$

as

$$\sup_{Y \in N(y)} \inf_{U \in N(s)} \sup_{s' \in S \cap U} \inf_{y' \in \Gamma_\Omega(p(s')) \cap Y} g(s', y'),$$

where $N(y)$ and $N(s)$ denote, respectively, a family of sufficiently small neighborhoods around y and s, $p(\cdot)$ denotes the projection from $\mathbb{R}^n \times \mathbb{R} \times [0, +\infty)$ onto $\mathbb{R}^n \times [0, +\infty)$, and

$$\Gamma_\Omega(x, t) \;=\; \begin{cases} t^{-1}(\Omega - x) & \text{if } t > 0, \\ \mathbb{R}^n & \text{if } t = 0. \end{cases}$$

To define the upper subderivative $f^{\uparrow}(x; v)$ one proceeds similarly as in [11] and chooses

$$g(x', \alpha', t, v') \;=\; \begin{cases} \frac{f(x' + tv') - \alpha'}{t} & \text{if } t > 0, \\ -\infty & \text{if } t = 0, \end{cases} \tag{6}$$

$s = (x, f(x), 0)$, $s' = (x', \alpha', t)$, $y = v$, and $y' = v'$. In the constrained case, however, one has now $S = \mathrm{epi}(f)(\Omega) \times [0, +\infty)$. These choices result then in the definition $f^{\uparrow}(x; v) = h((x, f(x), 0), v)$. We use the following expression to more easily grasp the essential of the definition of the upper subderivative $f^{\uparrow}(x;v)$:

$$f^{\uparrow}(x;v) \;=\; \limsup_{\substack{(x',\alpha')\downarrow_f x,\, x'\in\Omega \\ t\downarrow 0}} \;\; \inf_{\substack{v'\to v \\ x'+tv'\in\Omega}} \;\; \frac{f(x'+tv')-\alpha'}{t}.$$

The notation $(x',\alpha')\downarrow_f x$ represents $(x',\alpha')\to(x,f(x))$ with $\alpha'\geq f(x')$.

When f is lower semicontinuous at x, the derivative $f^{\uparrow}(x;v)$ can be equivalently defined by

$$f^{\uparrow}(x;v) \;=\; \limsup_{\substack{x'\to_f x,\, x'\in\Omega \\ t\downarrow 0}} \;\; \inf_{\substack{v'\to v \\ x'+tv'\in\Omega}} \;\; \frac{f(x'+tv')-f(x')}{t},$$

where, recall, $x'\to_f x$ represents $x'\to x$ and $f(x')\to f(x)$.

A characterization of the epigraph of the upper subderivative. The following proposition extends [11, Proposition 1] to the constrained case $\Omega\neq\mathbb{R}^n$.

Proposition A.1. *For each $s'\in S\subset\mathbb{R}^n\times\mathbb{R}\times[0,+\infty)$, let $\Gamma(s')$ denote the set in $\mathbb{R}^n\times\mathbb{R}$ which is the epigraph of $y\to g(s',y)$ restricted to $\Gamma_\Omega(p(s'))$:*

$$\Gamma(s') \;=\; \mathrm{epi}(g(s',\cdot))(\Gamma_\Omega(p(s'))).$$

Let also

$$\Delta(s) \;=\; \liminf_{\substack{s'\to s \\ s'\in S}} \;\Gamma(s').$$

Then $\Delta(s)$ is the epigraph of $y\to h(s,y)$ restricted to $T_\Omega(s)$:

$$\mathrm{epi}(h(s,\cdot))(T_\Omega(s)) \;=\; \Delta(s).$$

Proof: From its definition, the point (y,β) is in $\Delta(s)$ if and only if

$$\begin{gathered}\forall Y\in N(y),\forall\epsilon>0,\exists U\in N(s):\forall_{s'\in S\cap U}\exists(y',\beta'):\\ y'\in\Gamma_\Omega(p(s'))\cap Y,\beta'\in(\beta-\epsilon,\beta+\epsilon),g(s',y')\leq\beta'.\end{gathered}$$

which is equivalent to

$$\begin{gathered}\forall Y\in N(y),\forall\epsilon>0,\exists U\in N(s):\forall_{s'\in S\cap U}\exists y':\\ y'\in\Gamma_\Omega(p(s'))\cap Y,g(s',y')\leq\beta+\epsilon.\end{gathered}$$

Thus, (y, β) is in $\Delta(s)$ if and only if $y \in T_\Omega(s)$ and

$$\forall Y \in N(y), \forall \epsilon > 0, \exists U \in N(s) : \sup_{s' \in S \cap U} \inf_{y' \in \Gamma_\Omega(p(s')) \cap Y} g(s', y') \leq \beta + \epsilon.$$

This last condition is the same as saying that $h(s, y) \leq \beta$. ■

Note that the epigraph of $y \to g(s', y)$ restricted to $\Gamma_\Omega(p(s'))$ in the case (6) considered for the upper subderivatives is:

$$\Gamma(x', \alpha', t) \;=\; \begin{cases} t^{-1}\left(\mathrm{epi}(f)(\Omega) - (x', \alpha')\right) & \text{if } t > 0, \\ \mathbb{R}^n \times \mathbb{R} & \text{if } t = 0. \end{cases}$$

Thus, from Proposition A.1,

$$\liminf_{\substack{(x', \alpha') \downarrow_f x, x' \in \Omega \\ t \downarrow 0}} \Gamma(x', \alpha', t) \;=\; \mathrm{epi}(f^\uparrow(x; \cdot))(T_\Omega(x)).$$

On the other hand, from the definition of tangent cone

$$\liminf_{\substack{(x', \alpha') \downarrow_f x, x' \in \Omega \\ t \downarrow 0}} \Gamma(x', \alpha', t) \;=\; T_{\mathrm{epi}(f)(\Omega)}(x, f(x)).$$

Thus,

$$\mathrm{epi}(f^\uparrow(x; \cdot))(T_\Omega(x)) \;=\; T_{\mathrm{epi}(f)(\Omega)}(x, f(x)). \tag{7}$$

The relation (7) extends, to the constrained case, the part of [11, Theorem 2] which we need for what comes in Theorem A.1 below.

Definitions of other generalized directional derivatives. To define the generalized directional derivative $f_R^\circ(x; v)$ introduced in (4), in the constrained case, one first considers

$$h(s, y) \;=\; \limsup_{\substack{s' \to s \\ s' \in S \,:\, y \in \Gamma_\Omega(p(s'))}} g(s', y)$$

as

$$\inf_{U \in N(s)} \sup_{s' \in S \cap U : y \in \Gamma_\Omega(p(s'))} g(s', y).$$

The derivative is then defined as $f_R^\circ(x; v) = h((x, f(x), 0), v)$ by setting g as in (6), $s = (x, f(x), 0)$, $s' = (x', \alpha', t)$, and $y = v$ and, given the constrained case, $S = \mathrm{epi}(f)(\Omega) \times [0, +\infty)$. We will also use a more friendly description for this definition:

$$f_R^\circ(x; v) \;=\; \limsup_{\substack{(x', \alpha') \downarrow_f x, x' \in \Omega \\ t \downarrow 0, x' + tv \in \Omega}} \frac{f(x' + tv) - \alpha'}{t}.$$

When f is lower semicontinuous at x, the derivative $f_R^\circ(x; v)$ can be equivalently defined by

$$f_R^\circ(x; v) \;=\; \limsup_{\substack{x' \to_f x, x' \in \Omega \\ t \downarrow 0, x' + tv \in \Omega}} \frac{f(x' + tv) - f(x')}{t}.$$

Finally, if f is Lipschitz continuous near x, this derivative coincides with the Clarke-Jahn generalized directional derivative (2):

$$f_R^\circ(x; v) \;=\; f_C^\circ(x; v) \;=\; \limsup_{\substack{x' \to x, x' \in \Omega \\ t \downarrow 0, x' + tv \in \Omega}} \frac{f(x' + tv) - f(x')}{t}.$$

A characterization for the upper subderivatives. We reproduce below, in the space $\mathbb{R}^m$, the part of [11, Corollary 2] which will be needed later. Recall, from Definition 4.1, the notion of a vector hypertangent to a set at a point of the set and, from the discussion after this definition, the concept of a set epi-Lipschitzian with respect to a vector at a point of the set.

Proposition A.2. *Let $C \subset \mathbb{R}^m$ and $y \in C$. If C is epi-Lipschitzian at y with respect to some w, then the vectors w with this property form $\mathrm{int}(H_C(y))$ and one has $T_C(y) = \mathrm{cl}(H_C(y))$.*

Finally, we prove the results needed for Theorems 4.1 and 4.2.

Theorem A.1. *Let f be an extended-real-valued function and x a point in Ω with $f(x) < +\infty$.*

The function f is directionally Lipschitzian at x with respect to all vectors in

$$\{v' \in \mathrm{int}(H_\Omega(x)) : \; f_R^\circ(x; v') < \infty\}. \tag{8}$$

In this set, $f^\uparrow(x; \cdot)$ is continuous and

$$f^\uparrow(x; \cdot) \;=\; f^{\circ\circ}(x; \cdot) \;=\; f_R^\circ(x; \cdot).$$

Further, for all vectors in $v \in T_\Omega(x)$ which are approachable from (8), one has

$$f^{\uparrow}(x;v) \;=\; \liminf_{\substack{v' \to v \\ v' \in \operatorname{int}(H_\Omega(x)) \\ f^{\circ\circ}(x;v') < +\infty}} f^\circ_R(x;v'). \tag{9}$$

Proof: Let us apply Proposition A.2 to $C = \operatorname{epi}(f)(\Omega)$ at $(x, f(x)) \in (\Omega, \mathbb{R})$. The hypertangent cone $H_{\operatorname{epi}(f)(\Omega)}$ consists of all vectors (v, β) such that

$$\begin{aligned} \exists \epsilon > 0 : \;\; & (x', \alpha') + t(v, \beta) \in \operatorname{epi}(f)(\Omega), \\ & \forall t \in (0, \epsilon),\; (x', \alpha') \in \operatorname{epi}(f)(\Omega) \cap B((x, f(x)); \epsilon). \end{aligned}$$

Note that $(x', \alpha') + t(v, \beta) \in \operatorname{epi}(f)(\Omega)$ is the same as

$$x' + tv \in \Omega \quad \text{and} \quad \frac{f(x' + tv) - \alpha'}{t} \le \beta.$$

Thus, one can see directly from this definition that

$$\begin{aligned} & \operatorname{int}(H_{\operatorname{epi}(f)(\Omega)}(x, f(x))) \\ = \;\; & \left\{(v, \beta) :\; v \in \operatorname{int}(H_\Omega(x)),\; f^\circ_R(x; \cdot) < \beta \text{ in a neigh. of } v\right\}. \end{aligned}$$

The set of vectors for which $\operatorname{epi}(f)(\Omega)$ is epi-Lipschitzian w.r.t., at the point $(x, f(x))$, is

$$\left\{(v, \beta) :\; v \in \operatorname{int}(H_\Omega(x)),\; f^{\circ\circ}(x; v) < \beta\right\}.$$

Note that this set is the same as

$$\left\{(v, \beta) :\; v \in \operatorname{int}(H_\Omega(x)),\; f^{\circ\circ}(x; \cdot) < \beta \text{ in a neigh. of } v\right\}$$

and, from Proposition A.2, we also obtain

$$\begin{aligned} & \operatorname{int}(H_{\operatorname{epi}(f)(\Omega)}(x, f(x))) \\ = \;\; & \left\{(v, \beta) :\; v \in \operatorname{int}(H_\Omega(x)),\; f^{\circ\circ}(x; \cdot) < \beta \text{ in a neigh. of } v\right\}. \end{aligned}$$

Thus,

$$\begin{aligned} & \left\{(v, \beta) :\; v \in \operatorname{int}(H_\Omega(x)),\; f^\circ_R(x; \cdot) < \beta \text{ in a neigh. of } v\right\} \\ = \;\; & \left\{(v, \beta) :\; v \in \operatorname{int}(H_\Omega(x)),\; f^{\circ\circ}(x; \cdot) < \beta \text{ in a neigh. of } v\right\}. \end{aligned}$$

or, in other words,

$$\operatorname{int}\left[\operatorname{epi}(f^\circ_R(x; \cdot))(\operatorname{int}(H_\Omega(x)))\right] \;=\; \operatorname{int}\left[\operatorname{epi}(f^{\circ\circ}(x; \cdot))(\operatorname{int}(H_\Omega(x)))\right].$$

We conclude that $f^\circ_R(x; \cdot)$ and $f^{\circ\circ}(x; \cdot)$ coincide in $\operatorname{int}(H_\Omega(x))$.

One can see from its epigraph that $f_R^\circ(x;\cdot)$ is convex. Also, since $f_R^\circ(x;\cdot)$ is bounded above in a neighborhood of at least a point in $\mathrm{int}(H_\Omega(x))$ and since it is convex, it is necessarily continuous in

$$\{v' \in \mathrm{int}(H_\Omega(x)) : \; f_R^\circ(x;v') < \infty\}.$$

We have seen in this proof that the epigraph of

$$v \to \liminf_{\substack{v' \to v \\ v' \in \mathrm{int}(H_\Omega(x)) \\ f^{\circ\circ}(x;v') < +\infty}} f_R^\circ(x;v')$$

in $T_\Omega(x)$ is the closure of $H_{\mathrm{epi}(f)(\Omega)}(x, f(x))$. On the other hand, from Proposition A.2 and (7) and we know that

$$\mathrm{cl}[H_{\mathrm{epi}(f)(\Omega)}(x, f(x))] \;=\; T_{\mathrm{epi}(f)(\Omega)}(x, f(x)) \;=\; \mathrm{epi}(f^\uparrow(x;\cdot))(T_\Omega(x)).$$

Thus, we have established (9) for all vectors in $v \in T_\Omega(x)$ which are approachable from (8). This also shows that $f^\uparrow(x;v) = f_R^\circ(x;v)$ in $\mathrm{int}(H_\Omega(x))$ whenever $f^\uparrow(x;v) < +\infty$. ■

Acknowledgements. We are grateful to José Francisco Rodrigues (University of Lisbon) for having drawn our attention to the extension result in [10]. We also thank Boris Mordukhovich (Wayne State University) and Rui Loja Fernandes (Technical University of Lisbon) for interesting discussions on related topics.

References

[1] NOMADm. http://www.gerad.ca/NOMAD/Abramson/nomadm.html.
[2] M. A. Abramson and C. Audet. Convergence of mesh adaptive direct search to second-order stationary points. *SIAM J. Optim.*, 17:606–619, 2006.
[3] C. Audet and J. E. Dennis Jr. Analysis of generalized pattern searches. *SIAM J. Optim.*, 13:889–903, 2002.
[4] C. Audet and J. E. Dennis Jr. Mesh adaptive direct search algorithms for constrained optimization. *SIAM J. Optim.*, 17:188–217, 2006.
[5] F. H. Clarke. *Optimization and Nonsmooth Analysis*. John Wiley & Sons, New York, 1983. Reissued by SIAM, Philadelphia, 1990.
[6] A. R. Conn, K. Scheinberg, and L. N. Vicente. *Introduction to Derivative-Free Optimization*. MPS-SIAM Series on Optimization. SIAM, Philadelphia, 2009.
[7] A. L. Custódio, J. E. Dennis Jr., and L. N. Vicente. Using simplex gradients of nonsmooth functions in direct search methods. *IMA J. Numer. Anal.*, 28:770–784, 2008.
[8] J. Jahn. *Introduction to the Theory of Nonlinear Optimization*. Springer-Verlag, Berlin, 1996.
[9] T. G. Kolda, R. M. Lewis, and V. Torczon. Optimization by direct search: New perspectives on some classical and modern methods. *SIAM Rev.*, 45:385–482, 2003.

[10] E. J. McShane. Extension of range of functions. *Bull. Amer. Math. Soc.*, 40:837–842, 1934.
[11] R. T. Rockafellar. Generalized directional derivatives and subgradients of nonconvex functions. *Can. J. Math.*, 32:257–280, 1980.
[12] V. Torczon. On the convergence of pattern search algorithms. *SIAM J. Optim.*, 7:1–25, 1997.

L. N. Vicente
CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal (lnv@mat.uc.pt).

A. L. Custódio
CMA-UNL, Department of Mathematics, FCT-UNL, Quinta da Torre, 2829-516 Caparica, Portugal (alcustodio@fct.unl.pt).