Les Cahiers du GERAD ISSN: 0711–2440

Convergence towards a local minimum by direct search methods with a covering step

P.-Y. Bouchet, C. Audet, L. Bourdin

G–2024–05

January 2024

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2024
– Bibliothèque et Archives Canada, 2024

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2024
– Library and Archives Canada, 2024

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Convergence towards a local minimum by direct search methods with a covering step

Pierre-Yves Bouchet [a, b]
Charles Audet [a, b]
Loïc Bourdin [c]

[a] *Département de mathématiques et de génie industriel, Polytechnique Montréal, Montréal (Qc), Canada, H3T 1J4*

[b] *GERAD, Montréal (Qc), Canada, H3T 1J4*

[c] *XLIM Research Institute, University of Limoges, 87060 Limoges, France*

pierre-yves.bouchet@polymtl.ca
charles.audet@gerad.ca
loic.bourdin@unilim.fr

January 2024
Les Cahiers du GERAD
G–2024–05

Abstract : This paper introduces a new step to the *Direct Search Method* (DSM) to strengthen its convergence analysis. By design, this so-called *covering step* may ensure that for all refined points of the sequence of incumbent solutions generated by the resulting cDSM (*covering DSM*), the set of all evaluated trial points is dense in a neighborhood of that refined point. We prove that this additional property guarantees that all refined points are local solutions to the optimization problem. This new result holds true even for discontinuous objective function, under a mild assumption that we discuss in details. We also provide a practical construction scheme for the covering step that works at low additional cost per iteration.

Keywords: Discontinuous optimization, nonsmooth optimization, derivative-free optimization, Direct Search Method, convergence, local solution

Résumé : Cet article propose une nouvelle étape à ajouter à chaque itération de la *Méthode de Recherche Directe* (*Direct Search Method* (DSM) en anglais) pour renforcer son analyse de convergence. Cette nouvelle étape, nommée *covering step*, peut garantir par construction que pour tout point raffiné de la suite de points générée par cDSM (*covering DSM*), un ensemble dense de points est évalué dans un voisinage de ce point raffiné. Nous prouvons que cette propriété permet de certifier l'optimalité locale de tous les points raffinés. Ce nouveau résultat est valide pour des fonctions objectif potentiellement discontinues, sous une hypothèse légère que nous discutons en détails. Nous proposons également un schéma de construction pratique pour la covering step qui mène à un faible surcoût par itération.

Mots clés : Optimisation discontinue, optimisation non lisse, optimisation sans dérivées, Direct Search Methods, convergence, solution locale

Acknowledgements: We thank Dr. Solène Kojtych for her insightful comments when we were investigating the properties of the revealing step. She deserves credit for Theorem 1 since she also foresaw that the revealing step could strengthen the convergence analysis of DSM. This work was funded by Audet's NSERC Canada Discovery Grant 2020–04448.

1 Introduction

Consider the generic optimization problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x), \tag{$\mathbf{P}$}$$

where $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is possibly discontinuous and has a nonempty effective domain $X \subseteq \mathbb{R}^n$. The *Direct Search Method* (DSM) addresses Problem ($\mathbf{P}$) by generating sequences $(x^k)_{k\in\mathbb{N}}$ of incumbent solutions and $(\delta^k)_{k\in\mathbb{N}}$ of poll radii, where $(x^k, \delta^k) \in X \times \mathbb{R}^*_+$ for all $k \in \mathbb{N}$. The literature about DSM extracts *refining subsequences* from $(x^k, \delta^k)_{k\in\mathbb{N}}$ and studies the properties of their associated *refined points*. A subsequence $(x^k, \delta^k)_{k\in K^*}$, where $K^* \subseteq \mathbb{N}$ is infinite, is said to be *refining* [3, 4] if all iterations indexed by $k \in K^*$ *fail* (that is, $x^{k+1} = x^k$), $(\delta^k)_{k\in K^*}$ converges to zero and $(x^k)_{k\in K^*}$ converges to a limit x^* named the *refined point*. It is proved in [18] that, for all sets $K^* \subseteq \mathbb{N}$ indexing a refining subsequence, the corresponding refined point x^* satisfies a necessary optimality condition expressed in term of the Rockafellar derivative [17, 18] of f at x^*, provided that $(f(x^k))_{k\in K^*}$ converges to $f(x^*)$. Then, our previous work [2] extends [18] to ensure this last requirement under the assumption that all refining subsequences admit the same refined point. These results are valid for two standard classes of DSM: the *mesh*-based method [6, Part 3] and the *sufficient decrease*-based method [9, Section 7.7].

The main goal of this paper is to strengthen the above convergence analysis, via the addition of a new step to the DSM to guarantee that all refined points are local minima. First, we introduce the covering DSM (cDSM), relying on the so-called covering step which aims to ensure that for all refined points, the set of all evaluated trial points is dense in a neighborhood of that refined point. We refer to Property 1 for details. Second, we prove that this property implies that, for all $K^* \subseteq \mathbb{N}$ indexing a refining subsequence with refined point denoted by x^*, either $x^* \in X$ is a local solution to Problem ($\mathbf{P}$), or $x^* \notin X$ is such that $(f(x^k))_{k\in K^*}$ converges to the infimum of f over a neighborhood of x^*. This result is formalized in Theorem 1. Third, we propose a practical construction scheme for the covering step to indeed ensure Property 1 by design. This scheme fits in both the mesh-based cDSM and the sufficient decrease-based cDSM, and the additional cost per iteration it induces is low. Finally, this paper has an auxiliary goal in proving that our framework is tight and broader than others in the literature. Our assumptions cannot be relaxed in general, and they are weaker than in former work.

The covering step generalizes the revealing step from [1, 2]. Actually, our initial motivation was to better study this revealing step. Its goal in [1] is to reveal local discontinuities, but we observe in [2] that when the DSM generates a unique refined point, the revealing step provides the density of the trial points in a neighborhood of the refined point. However, [1, 2] fail to deduce the local optimality of the refined point. Thus, the current work originally aimed to state this property. Yet, we eventually found that the revealing step admits a generalization providing the density of the set of trial points around an arbitrary number of refined points. The formalization of this generalization and the study of its properties constitute the core of the present work. Note that we decided to change the terminology because, in comparison with the name revealing, we believe that the name covering better captures what this step actually does.

This paper is organized as follows. Section 2 formalizes the cDSM and states its convergence analysis in Theorem 1. Section 3 proves Theorem 1. Section 4 provides a construction scheme for the covering step. Section 5 discusses our assumptions about f. Finally, Section 6 identifies some extensions of our work. Appendix A contains proofs of auxiliary results used in the paper.

Notation: For all $r > 0$ and all $x \in \mathbb{R}^n$, we denote by $\mathcal{B}_r(x)$ the open ball of radius r in infinite norm centered at x, and by $\mathcal{B}_r \triangleq \mathcal{B}_r(0)$. For all $x \in \mathbb{R}^n$ and all $\mathcal{S} \subseteq \mathbb{R}^n$, we denote by $\{x\}+\mathcal{S} \triangleq \{x+s : s \in \mathcal{S}\}$ and $f(\mathcal{S}) \triangleq \{f(s) : s \in \mathcal{S}\}$ and $\operatorname{dist}(x, \mathcal{S}) \triangleq \inf\{\|x - s\| : s \in \mathcal{S}\}$. For all $\delta \in \mathbb{R}^*_+$ and all $\mathcal{S} \subseteq \mathbb{R}$, we define $\delta\mathcal{S} \triangleq \{\delta s : s \in \mathcal{S}\}$. For all $\mathcal{S} \subseteq \mathbb{R}^n$, $\mathcal{S}$ is said to be *ample* if $\mathcal{S} \subseteq \operatorname{cl}(\operatorname{int}(\mathcal{S}))$, or *locally thin* if there exists $\mathcal{N} \subseteq \mathbb{R}^n$ open so that $\mathcal{S} \cap \mathcal{N} \neq \emptyset = \operatorname{int}(\mathcal{S}) \cap \mathcal{N}$. Note that a set is either ample or locally thin (see Proposition 7.a) in Appendix A), and that the definition of an ample set is equivalent to that of a *semi-open* set introduced in [15]. We denote by $\mathbb{S}^n$ the unit sphere of $\mathbb{R}^n$.

2 Formal `covering` step and convergence result of `cDSM`

This section formalizes the `cDSM` as Algorithm 1 and its convergence analysis in Theorem 1. Theorem 1 is based on Property 1 regarding a *dense covering property* (DCP) provided by the `covering` step and on Assumption 1 regarding the objective function f. The proof of Theorem 1 follows in Section 3.

The `cDSM` matches the usual `DSM` in most of its aspects. The only novelty lies in the `covering` step, with its parameter $r \in \mathbb{R}^*_+$, detailed in Section 4. The other steps and parameters may be designed as in all mesh-based `DSM` [6, Part 3] and all sufficient decrease-based `DSM` [9, Section 7.7].

Algorithm 1 `cDSM` (`covering DSM`) **solving Problem (P).**

Initialization:

set a covering radius $r \in \mathbb{R}^*_+$ and the trial points history as $\mathcal{V}^0 \triangleq \emptyset$;

set the incumbent solution and poll radius as $(x^0, \delta^0) \in X \times \mathbb{R}^*_+$, and set $\underline{\delta}^0 \triangleq \delta^0$;

set $\mathcal{M} : \mathbb{R}^*_+ \to 2^{\mathbb{R}^n}$ and $\rho : \mathbb{R}^*_+ \to \mathbb{R}_+$ and $\Lambda \subset]0, 1[$ and $\Upsilon \subset [1, +\infty[$ according to one of

- the *mesh-based* `DSM`: $\mathcal{M}(\nu) \triangleq \min\{\nu, \frac{\nu^2}{\delta^0}\} M\mathbb{Z}^p$ for all $\nu \in \mathbb{R}^*_+$, where $p > n$ and $M \in \mathbb{R}^{n \times p}$ positively spans $\mathbb{R}^n$, and $\rho(\cdot) \triangleq 0$, and $\Lambda \subseteq \{\tau^\ell : \ell \in [\![1, m]\!]\}$ and $\Upsilon \subseteq \{\tau^\ell : \ell \in [\![-m, 0]\!]\}$ where $\tau \in \mathbb{Q} \cap]0, 1[$ and $m \in \mathbb{N}^*$;
- the *sufficient decrease-based* `DSM`: $\mathcal{M}(\cdot) \triangleq \mathbb{R}^n$, and ρ increasing with $0 < \rho(\nu) \in o(\nu)$ as $\nu \searrow 0$, and $\Lambda \subseteq [\underline{\lambda}, \overline{\lambda}]$ and $\Upsilon \subseteq [\underline{\upsilon}, \overline{\upsilon}]$ where $0 < \underline{\lambda} \leq \overline{\lambda} < 1 \leq \underline{\upsilon} \leq \overline{\upsilon} < +\infty$.

For $k \in \mathbb{N}$ **do**:

`search` **step**:

set $\mathcal{D}^k_{\mathtt{S}} \subseteq \mathcal{M}(\underline{\delta}^k)$ finite; if $\mathcal{T}^k_{\mathtt{S}} \triangleq \{x^k\} + \mathcal{D}^k_{\mathtt{S}}$ is nonempty, then set $t^k_{\mathtt{S}} \in \operatorname{argmin} f(\mathcal{T}^k_{\mathtt{S}})$;

if also $f(t^k_{\mathtt{S}}) < f(x^k) - \rho(\underline{\delta}^k)$, then set $t^k \triangleq t^k_{\mathtt{S}}$ and $\mathcal{T}^k_{\mathtt{C}} = \mathcal{T}^k_{\mathtt{P}} \triangleq \emptyset$ and go to the `update` step;

`covering` **step**:

set $\mathcal{D}^k_{\mathtt{C}} \subseteq \mathcal{M}(\underline{\delta}^k) \cap \mathrm{cl}(\mathcal{B}_r)$ finite; if $\mathcal{T}^k_{\mathtt{C}} \triangleq \{x^k\} + \mathcal{D}^k_{\mathtt{C}}$ is nonempty, then set $t^k_{\mathtt{C}} \in \operatorname{argmin} f(\mathcal{T}^k_{\mathtt{C}})$;

if also $f(t^k_{\mathtt{C}}) < f(x^k) - \rho(\underline{\delta}^k)$, then set $t^k \triangleq t^k_{\mathtt{C}}$ and $\mathcal{T}^k_{\mathtt{P}} \triangleq \emptyset$ and go to the `update` step;

`poll` **step**:

set $\mathcal{D}^k_{\mathtt{P}} \subseteq \mathcal{M}(\underline{\delta}^k) \cap \mathrm{cl}(\mathcal{B}_{\delta^k})$ a positive basis of $\mathbb{R}^n$; set $\mathcal{T}^k_{\mathtt{P}} \triangleq \{x^k\} + \mathcal{D}^k_{\mathtt{P}}$; set $t^k_{\mathtt{P}} \in \operatorname{argmin} f(\mathcal{T}^k_{\mathtt{P}})$;

if $f(t^k_{\mathtt{P}}) < f(x^k) - \rho(\underline{\delta}^k)$, then set $t^k \triangleq t^k_{\mathtt{P}}$, otherwise set $t^k \triangleq x^k$;

`update` **step**:

set $\mathcal{V}^{k+1} \triangleq \mathcal{V}^k \cup \mathcal{T}^k$, where $\mathcal{T}^k \triangleq \mathcal{T}^k_{\mathtt{S}} \cup \mathcal{T}^k_{\mathtt{C}} \cup \mathcal{T}^k_{\mathtt{P}}$; set $x^{k+1} \triangleq t^k$;

set δ^{k+1} as $\delta^{k+1} \in \delta^k \Upsilon$ if $t^k \neq x^k$ and $\delta^{k+1} \in \delta^k \Lambda$ otherwise; set $\underline{\delta}^{k+1} \triangleq \min_{\ell \leq k+1} \delta^k$.

A carefully constructed `covering` step ensures that all executions of the `cDSM` satisfy the next Property 1. A practical construction scheme to indeed meet this property is stated in Section 4.3.

Property 1 (*dense covering property* provided by the `covering` step)**.** The `covering` step ensures that the trial points history $\mathcal{V} \triangleq \cup_{k \in \mathbb{N}} \mathcal{V}^k$ generated by Algorithm 1 satisfies, for all refined points x^*,

$$\mathcal{B}_r(x^*) \ \subseteq\ \mathrm{cl}\,(\mathcal{V})\,. \tag{DCP}$$

Theorem 1 also requires the following assumptions about f. Assumptions 1.a) and 1.b) match usual assumptions used in the literature, while the unusual Assumption 1.c) is discussed in Section 5.

Assumption 1 (on the objective function f)**.** The objective function f in Problem (**P**) is such that

a) f is bounded below and has bounded sublevel sets;

b) the restriction $f_{|X} : X \to \mathbb{R}$ is lower semicontinuous;

c) X admits a partition $X = \sqcup_{i=1}^N X_i$ (where $N \in \mathbb{N}^* \cup \{+\infty\}$) such that, for all $i \in [\![1, N]\!]$, X_i is an *ample continuity set of* f (that is, X_i is ample and the restriction $f_{|X_i} : X_i \to \mathbb{R}$ is continuous).

Theorem 1. Under Assumption 1, Algorithm 1 generates at least one refining subsequence and, if Property 1 holds, then for all $K^* \subseteq \mathbb{N}$ indexing a refining subsequence, the refined point x^* satisfies

$$\lim_{k \in K^*} f(x^k) \ = \ \begin{cases} \min f(\mathcal{B}_r(x^*)) = f(x^*) & \text{if } x^* \in X, \\ \inf f(\mathcal{B}_r(x^*)) & \text{if } x^* \notin X. \end{cases}$$

Before proving Theorem 1 in Section 3, several comments are in order. First, the `covering` step fits in the framework of the `search` step. Yet, to the best of our knowledge, no `DSM` from the literature ensures Property 1, except instances of the `cDSM` relying on the scheme for the `covering` step we propose in Section 4. Second, when Assumption 1 holds, Property 1 ensures that all refined points are local solutions to Problem (**P**), in a generalized sense enclosing the case where some refined points lie outside the effective domain X. Last, in practice the `cDSM` likely generates exactly one refined point which moreover lies in X. This point is therefore a local solution to Problem (**P**) in the usual sense.

3 Proof of Theorem 1

Preliminary results. The `covering` step is a specific `search` step, thus Algorithm 1 inherits all the properties of the usual `DSM`. Hence, the next Proposition 1 holds. It is stated as [6, Theorem 8.1] for the mesh-based `DSM` and as [9, Corollary 7.2] for the sufficient decrease-based `DSM`.

Proposition 1. Under Assumption 1.a), Algorithm 1 generates at least one refining subsequence.

The rest of this section relies on the following unusual topological concepts. A set $\mathcal{S} \subseteq \mathbb{R}^n$ is said to be *ample* if $\mathcal{S} \subseteq \mathrm{cl}(\mathrm{int}(\mathcal{S}))$, and a set $\mathcal{S}_1 \subseteq \mathbb{R}^n$ is said to *have a dense intersection with another set* $\mathcal{S}_2 \subseteq \mathbb{R}^n$ if $\mathcal{S}_2 \subseteq \mathrm{cl}(\mathcal{S}_1 \cap \mathcal{S}_2)$. We leave some related properties to Proposition 7 in Appendix A.

The next Lemma 1 highlights what Property 1 provides. It leads to Proposition 2, which settles the ground for the proof of Theorem 1.

Lemma 1. Under Assumption 1.c), if Algorithm 1 satisfies Property 1, then for all refined points x^*, all $i \in [\![1, N]\!]$ and all $x \in \mathcal{B}_r(x^*) \cap X_i$, the set $\mathcal{V} \setminus \{x\}$ has a dense intersection with $\mathcal{B}_r(x^*) \cap X_i$.

Proof. Let x^* be a refined point generated by Algorithm 1 satisfying Property 1 and let $i \in [\![1, N]\!]$. Then $\mathcal{V}$ has a dense intersection with $\mathcal{B}_r(x^*)$, via Property 1 and Proposition 7.c) applied to $\mathcal{S}_1 \triangleq \mathcal{V}$ and $\mathcal{S}_2 \triangleq \mathrm{cl}(\mathcal{V})$ and $\mathcal{S}_3 \triangleq \mathcal{B}_r(x^*)$. Also, X_i is ample by Assumption 1.c) so $\mathcal{B}_r(x^*) \cap X_i$ is an ample subset of $\mathcal{B}_r(x^*)$ by Proposition 7.b). It follows that $\mathcal{V}$ has a dense intersection with $\mathcal{B}_r(x^*) \cap X_i$, from Proposition 7.c) applied to $\mathcal{S}_1 \triangleq \mathcal{V}$, $\mathcal{S}_2 \triangleq \mathcal{B}_r(x^*)$ and $\mathcal{S}_3 \triangleq \mathcal{B}_r(x^*) \cap X_i$. The claim follows from Proposition 7.d) applied to $\mathcal{S}_1 \triangleq \mathcal{V}$ and $\mathcal{S}_2 \triangleq \mathcal{B}_r(x^*) \cap X_i$. □

Proposition 2. Under Assumption 1, if Algorithm 1 satisfies Property 1, then for all refined points x^*, we have $\lim_{k\in\mathbb{N}} f(x^k) \leq f(x)$ for all $x \in \mathcal{B}_r(x^*)$.

Proof. First, $f^* \triangleq \lim_{k\in\mathbb{N}} f(x^k) \in \mathbb{R}$ exists since $(f(x^k))_{k\in\mathbb{N}}$ is bounded below by Assumption 1.a) and decreasing by construction. Second, let x^* be a refined point generated by Algorithm 1 satisfying Property 1 and take $x \in \mathcal{B}_r(x^*)$. The result holds if $x \notin X$, so assume that $x \in X_i$ for some $i \in [\![1, N]\!]$. Let $(y_\ell)_{\ell\in\mathbb{N}}$ converging to x with $y_\ell \in \mathcal{B}_r(x^*) \cap X_i \cap \mathcal{V} \setminus \{x\}$ for all $\ell \in \mathbb{N}$ (it exists by Lemma 1). Let $\kappa(\ell) \triangleq \min\{k \in \mathbb{N} : y_\ell \in \mathcal{T}^k\}$ for all $\ell \in \mathbb{N}$. Then $(\kappa(\ell))_{\ell\in\mathbb{N}}$ diverges to $+\infty$, since $\mathcal{T}^k$ is finite for all $k \in \mathbb{N}$ and every subsequence of $(y_\ell)_{\ell\in\mathbb{N}}$ takes infinitely many values (since $(y_\ell)_{\ell\in\mathbb{N}}$ converges to x with $y_\ell \neq x$ for all $\ell \in \mathbb{N}$). Also, for all $\ell \in \mathbb{N}$ we have $f(t^{\kappa(\ell)}) \geq f(x^{\kappa(\ell)}) - \rho(\underline{\delta}^{\kappa(\ell)})$ if iteration $\kappa(\ell)$ fails and $f(t^{\kappa(\ell)}) = f(x^{\kappa(\ell)+1})$ otherwise, and $f(y_\ell) \geq f(t^{\kappa(\ell)})$ by construction. Hence, for all $\ell \in \mathbb{N}$ we have $f(y_\ell) + \rho(\underline{\delta}^{\kappa(\ell)}) \geq f(x^{\kappa(\ell)+1})$. The result follows by taking $\ell \to +\infty$, since $(f(y_\ell))_{\ell\in\mathbb{N}}$ converges to $f(x)$ by continuity of $f_{|X_i}$ and $(f(x^{\kappa(\ell)+1}))_{\ell\in\mathbb{N}}$ converges to f^* and $(\rho(\underline{\delta}^{\kappa(\ell)}))_{\ell\in\mathbb{N}}$ converges to 0. □

Proof of Theorem 1. Consider that Assumption 1 is satisfied. Proposition 1 states that at least one refining subsequence is generated. Assume that Property 1 holds. Let $K^* \subseteq \mathbb{N}$ indexing a refining subsequence, x^* denoting its refined point, and let $f^* \triangleq \lim_{k\in K^*} f(x^k) = \lim_{k\in\mathbb{N}} f(x^k)$. Let us show that $f^* = \inf f(\mathcal{B}_r(x^*))$. We have $f^* \geq \inf f(\mathcal{B}_r(x^*))$ since $x^k \in \mathcal{B}_r(x^*)$ for all $k \in K^*$ large enough, and $f^* \leq \inf f(\mathcal{B}_r(x^*))$ is proved by contradiction: if $f^* > \inf f(\mathcal{B}_r(x^*))$, then there exists $x^\sharp \in \mathcal{B}_r(x^*)$ such that $f^* > f(x^\sharp)$, but then $f^* > f(x^\sharp) \geq f^*$ by Proposition 2. Note that this already concludes the proof of the case where $x^* \notin X$. Now assume that $x^* \in X$. Then $f^* \geq f(x^*)$ by Assumption 1.b) and $f^* \leq f(x^*)$ by Proposition 2, so $f(x^*) = f^* = \inf f(\mathcal{B}_r(x^*)) = \min f(\mathcal{B}_r(x^*))$. □

4 Discussion on the `covering` step and Property 1

This section discusses the novel algorithmic aspect of the `cDSM`, that is, the `covering` step with Property 1 it ensures *a posteriori* (that is, after the execution of `cDSM`). Section 4.1 highlights the differences between the `covering` step and the `revealing` step it is inspired from. Section 4.2 provides a sufficient condition about a construction scheme for the `covering` step to guarantee *a priori* (that is, prior to the execution of `cDSM`) that an execution of the `cDSM` will satisfy Property 1. It is presumably easier to check this condition a priori than to verify a posteriori that Property 1 holds. Section 4.3 provides a tractable construction scheme for the `covering` step which checks this sufficient condition.

4.1 The `revealing` step from prior work may not provide Property 1 a posteriori

As stated in Section 1, the `covering` step is inspired from the `revealing` step in a `DSM` that we hereafter call the `revealing DSM` (`rDSM`) [2, Algorithm 1]. Nevertheless, we show in this section that the `rDSM` may fail to ensure Property 1 when more than one refined point is generated.

Expressed in our notation, the `revealing` step in [2] relies on a sequence $(\mathcal{D}_\ell)_{\ell\in\mathbb{N}}$ with $\mathcal{D}_\ell \subset \mathrm{cl}(\mathcal{B}_r)$ finite for all $\ell \in \mathbb{N}$ and such that $\mathcal{B}_r \subseteq \mathrm{cl}(\cup_{\ell\in\mathbb{N}}\mathcal{D}_\ell)$. For all $k \in \mathbb{N}$, $\mathcal{D}^k_{\mathtt{C}}$ is the rounding of $\mathcal{D}_{\ell(k)}$ onto $\mathcal{M}(\underline{\delta}^k)$, where $\ell(k)$ denotes the number of iterations indices $u < k$ such that $\underline{\delta}^{u+1} < \underline{\delta}^u$. Thus, the `revealing` step ensures that

$$\mathcal{B}_r \subseteq \mathrm{cl}\left(\underset{k\in\mathbb{N}}{\cup}\mathcal{D}^k_{\mathtt{C}}\right)$$

when the `rDSM` generates at least one refining subsequence, which is certified a priori by Proposition 1. Nevertheless, this property states only the dense intersection of the trial directions history with $\mathcal{B}_r$. Property 1 follows when the `rDSM` generates exactly one refined point [2, Lemma 2]. However, this fails when more than one refined point exist. The following example illustrates this observation.

Example 1. Consider the objective function $f : \mathbb{R}^2 \to \mathbb{R}$ defined as $f(x) \triangleq \|x\|_\infty$ for all $x \in \mathbb{R}^2$ and the algorithmic parameters $r \triangleq 1$, $x^0 \triangleq -3\mathbb{1}$ where $\mathbb{1} \triangleq (1,1)$, $\delta^0 \triangleq 1$, $\mathcal{M}(\nu) \triangleq \min\{\nu,\nu^2\}\mathbb{Z}^2$ and $\rho(\nu) \triangleq 0$ for all $\nu \in \mathbb{R}^*_+$, $\Lambda \triangleq \{\frac{1}{2}\}$ and $\Upsilon \triangleq \{1\}$. Define the `poll` step as $\mathcal{D}^k_{\mathtt{P}} \triangleq \{(\pm\delta^k,0),(0,\pm\delta^k)\}$ for all $k \in \mathbb{N}$. Define the `revealing` step such that the sequence $(\mathcal{D}_\ell)_{\ell\in\mathbb{N}}$ satisfies

$$\forall q \in \mathbb{N}, \qquad \mathcal{D}_{4q} \subset \mathbb{R}_+ \times \mathbb{R}_+, \quad \mathcal{D}_{4q+1} \subset \mathbb{R}_- \times \mathbb{R}_+, \quad \mathcal{D}_{4q+2} \subset \mathbb{R}_- \times \mathbb{R}_-, \quad \mathcal{D}_{4q+3} \subset \mathbb{R}_+ \times \mathbb{R}_-.$$

Define the `search` step at each iteration $k \in \mathbb{N}$ so that $\mathcal{T}^k_{\mathtt{S}} \triangleq \emptyset$ if $k \notin 3\mathbb{N}$, and so that $\mathcal{T}^{3q}_{\mathtt{S}} \triangleq \{t^{3q}_{\mathtt{S}}\}$ for all $k = 3q \in 3\mathbb{N}$, where

$$t^{3q}_{\mathtt{S}} \triangleq (-1)^q\left(1+2^{-q}\right)\mathbb{1}.$$

This instance of `rDSM` has a predictable behavior. Using induction, one can show that

$$\forall q \in \mathbb{N}, \quad \left\{\begin{array}{lllll} x^{3q} = (-1)^{q-1}\left(1+2^{-(q-1)}\right)\mathbb{1}, & \delta^{3q} = 4^{-q}, & \mathcal{D}^{3q}_{\mathtt{C}} = \mathcal{D}_{2q}, & \texttt{search}\text{ success,} \\ x^{3q+1} = (-1)^q \left(1+2^{-q}\right)\mathbb{1}, & \delta^{3q+1} = 4^{-q}, & \mathcal{D}^{3q+1}_{\mathtt{C}} = \mathcal{D}_{2q}, & \text{iteration fails,} \\ x^{3q+2} = (-1)^q \left(1+2^{-q}\right)\mathbb{1}, & \delta^{3q+2} = \frac{1}{2}4^{-q}, & \mathcal{D}^{3q+2}_{\mathtt{C}} = \mathcal{D}_{2q+1}, & \text{iteration fails.} \end{array}\right.$$

Thus, there are two refined point, $x^*_+ = \mathbb{1}$ and $x^*_- = -\mathbb{1}$, and none is a local minimizer of f. However, this does not contradict Theorem 1 because Property 1 is not satisfied. Indeed, we have

$$\mathrm{cl}(\mathcal{V}) \cap \,]-1,1[^{\,2} = \emptyset \quad \text{while} \quad \mathcal{B}_r(x^*_+) = \,]1-r,1+r[^{\,2} \quad \text{and} \quad \mathcal{B}_r(x^*_-) = \,]-1-r,-1+r[^{\,2}, \quad \forall r \in \mathbb{R}^*_+.$$

Example 1 shows that the `revealing` step from [2] focuses only on the asymptotic density of the trial directions, and that this may not translate to the density of the trial points in some neighborhoods of the refined points. The original instance of the `revealing` step provided in [1] differs from whose in [2] (in [1], the `revealing` step draws at each iteration a direction in $\mathrm{cl}(\mathcal{B}_r)$ according to the independent uniform distribution), but it also fails to provide Property 1 when more than one refined point exists. Accordingly, in Section 4.2 we study schemes for the `covering` step that focus instead on satisfying directly the dense intersection of the trial points with a neighborhood of all refined points.

4.2 Sufficient condition to ensure a priori that Property 1 will hold

This section focuses on `covering` step instances relying only, at each iteration $k \in \mathbb{N}$, on the current couple $(x^k, \underline{\delta}^k)$, and the current history $\mathcal{V}^k = \cup_{\ell<k}\mathcal{T}^\ell$. We prove in Proposition 3 that if the `covering` step relies on a `covering` oracle from Definition 1, then all executions of the `cDSM` satisfy Property 1.

Definition 1 (`covering` oracle)**.** Given a `covering` radius $r \in \mathbb{R}^*_+$ and a mesh $\mathcal{M} : \mathbb{R}^*_+ \to 2^{\mathbb{R}^n}$ defined either as in the mesh-based `DSM` or the sufficient decrease-based `DSM`, a function $\mathbb{O} : \mathbb{R}^n \times \mathbb{R}^*_+ \times 2^{\mathbb{R}^n} \to 2^{\mathbb{R}^n}$ is said to be a *`covering` oracle* if $\mathbb{O}(y, \nu, \mathcal{S}) \subseteq \mathcal{M}(\nu) \cap \mathrm{cl}(\mathcal{B}_r)$ for all points $y \in \mathbb{R}^n$, all radii $\nu \in \mathbb{R}^*_+$ and all sets $\mathcal{S} \subseteq \mathbb{R}^n$, and if

$$\lim_{k\in\mathbb{N}} \max_{d\in\mathcal{M}(\nu^k)\cap\mathrm{cl}(\mathcal{B}_r)} \mathrm{dist}\left(y^k + d,\ \mathcal{S}^k\right) \quad = \quad 0$$

for all sequences $(y^k, \nu^k, \mathcal{S}^k)_{k\in\mathbb{N}}$ of elements of $\mathbb{R}^n \times \mathbb{R}^*_+ \times 2^{\mathbb{R}^n}$ such that $(y^k)_{k\in\mathbb{N}}$ converges and $(\nu^k)_{k\in\mathbb{N}}$ converges to 0 and $\mathcal{S}^{k+1} \supseteq \mathcal{S}^k \cup (\{y^k\} + \mathbb{O}(y^k, \nu^k, \mathcal{S}^k))$ for all $k \in \mathbb{N}$.

Proposition 3. If Algorithm 1 has its `covering` step constructed as

$$\forall k \in \mathbb{N}, \qquad \mathcal{D}^k_{\mathtt{C}} \triangleq \mathbb{O}\left(x^k, \underline{\delta}^k, \mathcal{V}^k\right),$$

where $\mathbb{O}$ satisfies Definition 1, then Property 1 holds a posteriori for all executions of Algorithm 1.

Proof. Consider the framework of Proposition 3. Denote by $\mathcal{M}_r(y,\nu) \triangleq \{y\} + (\mathcal{M}(\nu) \cap \mathrm{cl}(\mathcal{B}_r))$ for all $(y,\nu) \in \mathbb{R}^n \times \mathbb{R}^*_+$. Let $K^* \subseteq \mathbb{N}$ indexing a refining subsequence, with refined point x^*. We state Property 1 by checking that $\mathcal{V} \cap \mathcal{B}_\varepsilon(x) \neq \emptyset$, for all $\varepsilon > 0$ and all $x \in \mathcal{B}_r(x^*)$. Let $\varepsilon > 0$ and $x \in \mathcal{B}_r(x^*)$. By construction of the `covering` step, we have $\mathcal{V}^{k+1} \supseteq \mathcal{V}^k \cup \mathcal{T}^k_{\mathtt{C}} = \mathcal{V}^k \cup (\{x^k\} + \mathbb{O}(x^k, \underline{\delta}^k, \mathcal{V}^k))$ for all $k \in K^*$. Then, we may apply the definition of the `covering` oracle with $(x^k, \underline{\delta}^k, \mathcal{V}^k)_{k\in K^*}$ as the sequence $(y^k, \nu^k, \mathcal{S}^k)_{k\in\mathbb{N}}$. This leads to $\mathrm{dist}(\mathcal{M}_r(x^k, \underline{\delta}^k), \mathcal{V}^k) < \frac{\varepsilon}{2}$ for all $k \in K^*$ large enough. By construction of $\mathcal{M}(\underline{\delta}^k)$, we also have $\mathrm{dist}(x, \mathcal{M}_r(x^k, \underline{\delta}^k)) < \frac{\varepsilon}{2}$ for all $k \in K^*$ large enough. Then, let $k \in K^*$ satisfying these two conditions, define y as the rounding of x onto $\mathcal{M}_r(x^k, \underline{\delta}^k)$ and $c \in \mathcal{V}$ as the rounding of y onto $\mathcal{V}^k$. Thus, $\|x - c\| \leq \|x - y\| + \|y - c\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, so $c \in \mathcal{B}_\varepsilon(x) \cap \mathcal{V}$. □

Proposition 3 depends only on the oracle $\mathbb{O}$ driving the `covering` step, so it may be checked a priori. In contrast, Property 1 must be checked a posteriori, as this requires the whole sequence $(x^k, \delta^k, \mathcal{V}^k)_{k\in\mathbb{N}}$ to identify the refined points and check (DCP). Thus, checking Proposition 3 prior to executing the `cDSM` is presumably a good strategy to ensure that Property 1 will be satisfied. The construction scheme for the `covering` step we introduce in Section 4.3 relies on Definition 1 and Proposition 3.

4.3 Construction of a `covering` step instance usable in practice

In practice, an instance of the `covering` step must satisfy two criteria. The first is that the resulting trial points history $\mathcal{V}$ must satisfy Property 1. The second is that the number of `covering` trial points must be small at each iteration. This section proposes a practical scheme that meets these two criteria. We follow the guideline from Section 4.2, that is, $\mathcal{D}^k_{\mathtt{C}} \triangleq \mathbb{O}(x^k, \underline{\delta}^k, \mathcal{V}^k)$ at each $k \in \mathbb{N}$, where $\mathbb{O}$ is a `covering` oracle from Definition 1. Our scheme designs a tractable expression for $\mathbb{O}$.

The baseline scheme we suggest for the `covering` step relies on the following oracle:

$$\forall y \in \mathbb{R}^n, \quad \forall \nu \in \mathbb{R}^*_+, \quad \forall \mathcal{S} \subseteq \mathbb{R}^n, \qquad \mathbb{O}(y,\nu,\mathcal{S}) \triangleq \operatorname*{argmax}_{d\in\mathcal{M}(\nu)\cap\mathrm{cl}(\mathcal{B}_r)} \mathrm{dist}\left(y + d, \mathcal{S}\right). \tag{1}$$

Proposition 8 states that Oracle 1 is indeed a `covering` oracle. Then, a `covering` step instance relying on Oracle 1 ensures Property 1, as claimed by Proposition 3. This instance selects $\mathcal{D}^k_{\mathtt{C}} \triangleq \mathbb{O}(x^k, \underline{\delta}^k, \mathcal{V}^k)$ at each $k \in \mathbb{N}$ as the set of all directions $d^k_{\mathtt{C}} \in \mathcal{M}(\underline{\delta}^k) \cap \mathrm{cl}(\mathcal{B}_r)$ such that $t^k_{\mathtt{C}} \triangleq x^k + d^k_{\mathtt{C}}$ is the farthest possible from the set $\mathcal{V}^k$ of all past trial points. Moreover, Property 1 remains valid if we actually compute only one such direction. This computation is costly when k is large, but it may be alleviated.

First, in the mesh-based `cDSM`, computing Oracle 1 is a combinatorial problem, since $\mathcal{M}(\nu)$ is a discrete set for all $\nu \in \mathbb{R}^*_+$. It may be solved using a *distance transform algorithm*, such as whose in [16], which works in a number of operations linear with the cardinality of $\mathcal{M}(\nu) \cap \mathrm{cl}(\mathcal{B}_r)$ and most of them are feasible in parallel. In the sufficient decrease-based `cDSM` (where $\mathcal{M}(\nu) \triangleq \mathbb{R}^n$ for all $\nu \in \mathbb{R}^*_+$), computing Oracle 1 is a continuous and piecewise smooth problem. This problem admits surrogates, such as $\overline{\mathbb{O}}(y, \gamma, \mathcal{S}) \triangleq \mathrm{argmax}_{d \in \mathrm{cl}(\mathcal{B}_r)} \Sigma_{s \in \mathcal{S}} \frac{-1}{\|y+d-s\|}$. Also, at each iteration $k \in \mathbb{N}^*$ the computation of $\mathcal{D}^k_{\mathtt{C}}$ may start from the point $t^{k-1}_{\mathtt{C}}$ calculated at the preceding iteration.

Second, in practice we may alter Oracle 1 as

$$\forall y \in \mathbb{R}^n, \quad \forall \nu \in \mathbb{R}^*_+, \quad \forall \mathcal{S} \subseteq \mathbb{R}^n, \qquad \mathbb{O}(y, \nu, \mathcal{S}) \triangleq \underset{d \in \mathcal{M}(\nu) \cap \mathrm{cl}(\mathcal{B}_r)}{\mathrm{argmax}} \mathrm{dist}\left(y + d, \mathcal{S} \cap \mathrm{cl}(\mathcal{B}_r)\right). \tag{2}$$

This reduces the number of elements to consider in the computation of the point-set distance, and this alteration remains a `covering` oracle (the proof of Proposition 8 admits an adaptation to this oracle).

Third, we may also consider the following alteration of Oracle 2. Let $\alpha \in]0, 1]$ and consider

$$\mathbb{O}_\alpha(y, \nu, \mathcal{S}) \triangleq \left\{ d_\alpha \in \mathcal{M}(\nu) \cap \mathrm{cl}(\mathcal{B}_r) : \frac{\mathrm{dist}\left(y + d_\alpha, \mathcal{S} \cap \mathrm{cl}(\mathcal{B}_r)\right)}{\underset{d \in \mathcal{M}(\nu) \cap \mathrm{cl}(\mathcal{B}_r)}{\max} \mathrm{dist}\left(y + d, \mathcal{S} \cap \mathrm{cl}(\mathcal{B}_r)\right)} \geq \alpha \right\}, \tag{3}$$

for all $(y, \nu, \mathcal{S}) \in \mathbb{R}^n \times \mathbb{R}^*_+ \times 2^{\mathbb{R}^n}$. Oracle 3 remains a `covering` oracle (the proof of Proposition 8 may be adapted accordingly). Oracle 3 usually contains numerous elements, but recall that in practice we do not need to compute more than one. It is presumably easier to compute an element of the set defined by Oracle 3 than one of the set defined by Oracle 1, especially when α is chosen close to 0. A simple heuristic approach to localize such an element d_α may use a grid search on a grid thin enough, or some *space-filling sequences* such as the Halton sequence [12].

Last, let us stress that in practice, we may perform a `revealing` step such as in [1, 2] instead of a `covering` step relying on a `covering` oracle. Indeed, in practice the `revealing` step usually ensures Property 1. Moreover, the computational cost required by the `revealing` step is almost null. Nevertheless, despite its more expensive cost, Oracle 1 ensures that the trial points are well spread in a neighborhood of the current incumbent solution at each iteration. This contrasts with the `revealing` step, which offers no such guarantee. In addition, in a blackbox context the cost to compute Oracle 1 may be negligible anyway, since the bottleneck in this context is the cost to evaluate $f(x)$ for all $x \in X$, and this computation involves no call to the objective function. For comparison, our scheme constructs a `covering` step instance evaluating at most 1 point per iteration while the `poll` step considers at least $n + 1$ points per iteration since it relies on a positive basis of $\mathbb{R}^n$.

Let us discuss also the `covering` radius r. All $r \in \mathbb{R}^*_+$ are accepted, but fine-tuning this value in practice is a problem-dependent concern. Indeed, the smaller r is, the faster the `covering` step covers $\mathcal{B}_r(x^*)$ well, while in contrast, the larger r is, the more likely the `cDSM` escapes poor local solutions. Following this observation, two simple choices when no information about Problem (**P**) is available are $r \triangleq \frac{\delta^0}{10}$ or $r \triangleq \delta^0$. We may also consider a sequence $(r^k)_{k \in \mathbb{N}}$ instead of a fixed r, provided that $\underline{r} \triangleq \inf_{k \in \mathbb{N}} r^k > 0$. In that case, replace $\mathcal{B}_r(x^*)$ by $\mathcal{B}_{\underline{r}}(x^*)$ in Property 1 and Theorem 1. This alteration of Theorem 1 is proved by a rewording of Section 3 where $\mathcal{B}_r(x^*)$ is replaced accordingly.

5 Discussion on Assumption 1.c)

This section discusses our novel assumption describing the continuity sets of f, that is, Assumption 1.c). In Section 5.1, we prove that Assumption 1.c) is strictly weaker than similar assumptions considered in former work [1, 2, 18]. In Section 5.2, we show that Assumption 1.c) is tight.

5.1 Comparison of Assumption 1.c) with similar assumptions in prior work

In this section, we compare Assumption 1.c) to similar assumptions considered by former work [1, 2, 18]. Precisely, we show that Assumption 1.c) is strictly weaker than either [1, Assumption 4.4] and [2, Assumption 1]. The work [18] is not considered since [2] is an extension of it.

First, let us compare Assumption 1.c) to [1, Assumption 4.4], recalled below as Assumption 2. We prove in Proposition 4 that Assumption 1.c) is strictly weaker than Assumption 2.

Assumption 2 (Assumption 4.4 in [1])**.** There exists $N \in \mathbb{N}^* \cup \{+\infty\}$ nonintersecting open sets X_i such that $\mathrm{cl}(X) = \cup_{i=1}^N \mathrm{cl}(X_i)$ and $f_{|X_i}$ is continuous for all $i \in [\![1,N]\!]$ and, for all $x \in X$, there exists $j \in [\![1,N]\!]$ such that $x \in \mathrm{cl}(X_j)$ and $f_{|X_j \cup \{x\}}$ is continuous.

Proposition 4. Assumption 1.c) is strictly weaker than Assumption 2.

Proof. Suppose that Assumption 2 holds. Denote by $(X_i)_{i=1}^N$ the family it provides. For all $i \in [\![1,N]\!]$, let $\mathrm{cl}_f(X_i) \triangleq \{x \in \mathrm{cl}(X_i) : f_{|X_i \cup \{x\}} \text{ is continuous}\}$. Let $I(x) \triangleq \min\{i \in [\![1,N]\!] : x \in \mathrm{cl}_f(X_i)\}$ for all $x \in X$. Then let $Y_i \triangleq \{x \in X : I(x) = i\}$ for every $i \in [\![1,N]\!]$. Thus $(Y_i)_{i=1}^N$ passes all the requirements in Assumption 1.c) (see Proposition 9), so Assumption 1.c) is weaker than Assumption 2. Now, to prove that Assumption 2 is not weaker than Assumption 1.c), consider the case

$$f : \left\{ \begin{array}{ccl} X \triangleq [-1,1] \setminus \{0\} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{1}{i} \text{ if } |x| \in \left]\dfrac{1}{i+1}, \dfrac{1}{i}\right] \text{ for some } i \in \mathbb{N}^*. \end{array} \right.$$

The continuity sets of f are $X_i \triangleq [\frac{-1}{i}, \frac{-1}{i+1}[\cup]\frac{1}{i+1}, \frac{1}{i}]$ for all $i \in \mathbb{N}^*$. Assumption 1.c) holds since X_i is ample for all $i \in \mathbb{N}^*$. Nevertheless Assumption 2 does not hold since the continuity sets must be adapted as $Y_i \triangleq \mathrm{int}(X_i)$ for all $i \in \mathbb{N}^*$ to be open, but then $\mathrm{cl}(X) = [-1,1] \neq \cup_{i=1}^\infty \mathrm{cl}(Y_i) = [-1,1] \setminus \{0\}$. ☐

Second, let us compare Assumption 1.c) to [2, Assumption 1], reformulated below in Assumption 3. We prove in Proposition 5 that Assumption 1.c) is strictly weaker than Assumption 3.

Assumption 3 (Global reformulation of Assumption 1 in [2])**.** The set X admits a partition $X = \sqcup_{i=1}^N X_i$ (where $N \in \mathbb{N}^* \cup \{+\infty\}$) such that, for all $i \in [\![1,N]\!]$, X_i is a *continuity set of f with the interior cone property* (that is, X_i satisfies Definition 2 below and $f_{|X_i} : X_i \to \mathbb{R}$ is continuous).

Definition 2 (Interior cone property and exterior cone property)**.** A set $\mathcal{S} \subseteq \mathbb{R}^n$ is said to have the *interior cone property* (ICP) if

$$\forall x \in \partial\mathcal{S}, \quad \exists \left\{ \begin{array}{l} \mathcal{U} \subseteq \mathbb{S}^n \text{ nonempty, open in the topology induced by } \mathbb{S}^n \\ \mathcal{K} \triangleq \mathbb{R}_+^* \mathcal{U} \text{ the cone generated by } \mathcal{U}, \quad \mathcal{K}_x \triangleq \{x\} + \mathcal{K} \\ \mathcal{O} \text{ an open neighborhood of 0 in } \mathbb{R}^n, \quad \mathcal{O}_x \triangleq \{x\} + \mathcal{O} \end{array} \right. \quad : \quad (\mathcal{K}_x \cap \mathcal{O}_x) \subseteq \mathcal{S}.$$

Similarly, $\mathcal{S}$ is said to have the *exterior cone property* (ECP) if $(\mathbb{R}^n \setminus \mathcal{S})$ has the ICP.

Assumption 3 differs from [2, Assumption 1] in three aspects. Let us stress those and argue that they do not spoil the assumption. First, Assumption 3 is a global statement, while [2, Assumption 1] states a local property but for each $x \in X$. Second, [2, Assumption 1] requires Lipschitz-continuity of f on each continuity set, while Assumption 3 calls for continuity only. Indeed [2] requires Lipschitz-continuity only to evaluate some generalized derivatives. We do not consider those in the current work, so we weaken the assumption accordingly. Third, [2, Assumption 1] requires an *exterior* cone property for $X \setminus X_i$ for all $i \in [\![1, N]\!]$, while Assumption 3 demands an *interior* cone property for X_i. Moreover the exterior cone property required in [2, Assumption 1] is [18, Definition 4.1], which differs from Definition 2. However, these two approaches are equivalent (see Proposition 10). Hence, Assumption 3 and [2, Assumption 1] differ only in the nature of the continuity of f on each continuity set.

Proposition 5. Assumption 1.c) is strictly weaker than Assumption 3.

Proof. For all $\mathcal{S} \subseteq \mathbb{R}^n$, if $\mathcal{S}$ has the ICP, then $\mathcal{S}$ is ample, but the reciprocal implication may fail (see Proposition 11). The result follows directly. $\square$

5.2 Tightness of Assumption 1.c)

In this section, we show that Assumption 1.c) is tight, in the sense that the conclusion of Theorem 1 may not hold if Assumption 1.c) is not satisfied. Relaxing it as a local property is possible, but then the convergence analysis relies on some information accessible only a posteriori (after the optimization process ends). Hence our framework cannot be broadened using only information available a priori.

Proposition 6. If Assumption 1.c) does not hold, then the conclusion of Theorem 1 may not hold.

Proof. Let us develop a counterexample. Let $n = 2$ and $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$\forall x \in X \triangleq \mathbb{R}^2, \qquad f(x) \triangleq \begin{cases} |x_1 - 1| + |x_2| - 1, & \text{if } x \in X_1 \triangleq (\mathbb{R}_- \times \mathbb{R}) \cup (\mathbb{R}_+^* \times \{0\}), \\ |x_1| + |x_2|, & \text{if } x \in X_2 \triangleq X \setminus X_1. \end{cases}$$

Assumptions 1.a) and 1.b) hold, but Assumption 1.c) does not since X_1 is locally thin and thus not ample. Consider an instance of Algorithm 1 such that $(\mathcal{D}^k_{\mathtt{S}} \cup \mathcal{D}^k_{\mathtt{C}} \cup \mathcal{D}^k_{\mathtt{P}}) \cap (\mathbb{R}_+^* \times \{0\}) = \emptyset$ for all $k \in \mathbb{N}$ and satisfying Property 1 and starting from the origin. This instance remains at the origin, since it evaluates only points in $\mathrm{cl}(\mathrm{int}(X_1)) \cup X_2 = \mathbb{R}^2 \setminus (\mathbb{R}_+^* \times \{0\})$ and the origin is the global minimizer of the restriction of f to $\mathrm{cl}(\mathrm{int}(X_1)) \cup X_2$. In that situation, the origin is a refined point lying in X but is not a local minimizer of f, which contradicts the conclusion of Theorem 1. $\square$

Assumption 1.c) allows a broad class of discontinuous functions, as it only rejects discontinuous functions for which at least one of the continuity sets is locally thin. It is possible to relax it as a local assumption holding only near some refined points and to adapt Theorem 1 accordingly. We leave this theorem as Assertion 1, although its proof is similar to whose provided in Section 3. Nevertheless, Assertion 1 is only usable a posteriori, since it is impossible to determine the refined points a priori.

Assertion 1. Under Assumptions 1.a) and 1.b), Algorithm 1 generates at least one refining subsequence and, if Property 1 holds, then for all $K^* \subseteq \mathbb{N}$ indexing a refining subsequence, with refined point x^*,
a) if $\mathcal{X}_i$ is ample for all $i \in [\![1, N]\!]$, then $\lim_{k \in K^*} f(x^k) = \inf f(\mathcal{B}_r(x^*))$;
b) if moreover $x^* \in \mathcal{X}_i$ for some $i \in [\![1, N]\!]$, then $\lim_{k \in K^*} f(x^k) = \min f(\mathcal{B}_r(x^*)) = f(x^*)$;
where we denote by $(\mathcal{X}_i)_{i=0}^N$ a partition of $\mathcal{B}_r(x^*)$, for some $N \in \mathbb{N}^* \cup \{+\infty\}$, such that $\mathcal{X}_0 \triangleq \mathcal{B}_r(x^*) \setminus X$ and $\mathcal{X}_i$ is a continuity set of f for all $i \in [\![1, N]\!]$.

6 General comments and main extensions for future work

Theorem 1 is stronger than most results from the literature about DSM in two aspects. First, Theorem 1 ensures the local optimality of all refined points, while usual results claim only some necessary local optimality conditions. Second, Theorem 1 holds for all refining subsequence. In contrast, the literature usually consider only refining subsequences such that the set of associated refined directions is dense in the unit sphere, and no ways to identify such a refining subsequence are provided.

Assumption 1.b), the assumption that f is lower semicontinuous, may be relaxed. The proof of Theorem 1 highlights that if only Assumption 1.b) fails, the equality $\lim_{k \in K^*} f(x^k) = \inf f(\mathcal{B}_r(x^*))$

holds for all $K^* \subseteq \mathbb{N}$ indexing a refining subsequence, with refined point denoted by x^*. Nevertheless, assuming the lower semicontinuity of f at x^* recovers $\inf f(\mathcal{B}_r(x^*)) = \min f(\mathcal{B}_r(x^*)) = f(x^*)$.

The `cDSM` currently requires some structure on X to select $x^0 \in X$. Following the terminology in [14], this is manageable when Problem (**P**) is the *extreme barrier* reformulation of a constrained problem with quantifiable constraints. Some two-phase algorithms exist [6, Algorithm 12.1], where the first phase minimizes the *constraints violation function* [10] to identify a feasible point.

Only the `covering` step matters to establish Theorem 1, in the sense that a variant of Algorithm 1 performing only the `covering` step and the `update` step, and with either $\rho(\nu) \triangleq 0$ and $\mathcal{M}(\nu) \triangleq \mathbb{R}^n$ for all $\nu \in \mathbb{R}^*_+$, still generates a refining subsequence and thus satisfies Theorem 1. This shares similarity with derivative-free line search algorithms [8, 11]. Similarly, all algorithms generating a refining subsequence may satisfy Theorem 1, when enhanced with the `covering` step. Then, future work may add a `covering` step into such algorithms. For example, Bayesian-based methods, model-based methods, and most methods listed in the survey [13]. This also includes algorithms handling constraints via advanced techniques, such as the *progressive barrier* [5] which iteratively reduces the infeasibility of the incumbent solution.

Nevertheless, the `search` and `poll` steps are important in practice. The `search` step allows for global exploration, and the `poll` step usually contributes to many successful iterations. In contrast, the current purpose of the `covering` step is to ensure the asymptotic Property 1, so a poor instance may be inefficient in finite time. Our scheme for the `covering` step discussed in Section 4.3 ensures that the trial points are well spread in a neighborhood of the current incumbent solution at each iteration. Yet, it is only a baseline, so additional investigations and careful implementations may identify more efficient schemes. A future work may verify these suppositions and quantify the relevance of each step during the entire optimization process. We may also study how well the `cDSM` performs when Assumption 1.c) holds but its stronger variant Assumption 3 does not. Presumably, the interior cone property provided by Assumption 3 is important for practical efficiency.

With an involved construction scheme, the `covering` step may help to capture and exploit the local shape of f and to evaluate some neighbors of the current incumbent in potential areas missed by the `poll` step. In this aspect, a promising `covering` oracle may derive from *model-based techniques* [6, Part 4], or from an expected improvement [19]. Indeed, they aim to identify some points that are relevant candidates (see also [7]) or that may help to gather the local structure of f.

The `covering` step is also compatible with the `DiscoMads` algorithm [1], which designs the original `revealing` step for the purpose to detect discontinuities and repel its incumbent solution from those. This discontinuities detection is more accurate with the `covering` step than with the `revealing` step, since the latter may fail to detect discontinuities when more than one refined point is eventually generated. Also, when the `covering` step relies on Oracle 1, it presumably has better practical guarantees to efficiently detect discontinuities than both the the original `revealing` step from [1] and the adapted `revealing` step from [2]. Then, for reliability reasons, it may be safer to use the `DiscoMads` algorithm with a `covering` step instead of a `revealing` step.

We conclude this paper with the next Table 1. It summarizes the conceptual differences between the usual `DSM` and our `cDSM`, and it highlights the aforementioned ideas to extend the `cDSM`.

Table 1: Differences between DSM and cDSM, and possible extension of cDSM with new covering step goals.

step	method		
	DSM	cDSM	cDSM with involved covering step design
search	Optional. Allows the use of heuristics and globalization strategies.		
poll	Required. Converges towards a refined point satisfying necessary optimality conditions.	Optional. But in practice, it performs well in converging towards a good refined point.	
covering	Undefined.	Required. Safeguard to asymptotically ensure that all refined points are local solutions, low cost per iteration, but may be inefficient in practice.	Required. Asymptotically ensures that all refined points are local solutions; also rapidly collects local information about the objective function.

Appendix

A Proofs of auxiliary results

Proposition 7. Let $\mathcal{S}_1 \subseteq \mathbb{R}^n$, $\mathcal{S}_2 \subseteq \mathbb{R}^n$ and $\mathcal{S}_3 \subseteq \mathbb{R}^n$. Then,

a) $\mathcal{S}_1$ is locally thin if and only if it is not ample;

b) if $\mathcal{S}_1$ is ample and $\mathcal{S}_2$ is open, then $\mathcal{S}_1 \cap \mathcal{S}_2$ is ample;

c) if $\mathcal{S}_1$ has a dense intersection with $\mathcal{S}_2$ and $\mathcal{S}_3$ is an ample subset of $\mathcal{S}_2$, then $\mathcal{S}_1$ has a dense intersection with $\mathcal{S}_3$;

d) if $\mathcal{S}_1$ has a dense intersection with $\mathcal{S}_2$ and $\mathcal{S}_2$ is ample, then, for all $x \in \mathcal{S}_2$, $\mathcal{S}_1 \setminus \{x\}$ has a dense intersection with $\mathcal{S}_2$.

Proof. Let us prove the first statement. Let $\mathcal{S}_1 \subseteq \mathbb{R}^n$ be locally thin and let us prove by contradiction that it is not ample. Assume that $\mathcal{S}_1$ is ample. Define $\mathcal{N} \subseteq \mathbb{R}^n$ open such that $\mathcal{S}_1 \cap \mathcal{N} \neq \emptyset = \mathrm{int}(\mathcal{S}_1) \cap \mathcal{N}$, and take $x \in \mathcal{S}_1 \cap \mathcal{N}$. Then $x \in \mathrm{cl}(\mathrm{int}(\mathcal{S}_1)) \cap \mathcal{N}$ and thus there exists $\varepsilon > 0$ such that $\mathcal{B}_\varepsilon(x) \subseteq \mathcal{N}$ and $\mathcal{B}_\varepsilon(x) \cap \mathrm{int}(\mathcal{S}_1) \neq \emptyset$. Hence $\mathrm{int}(\mathcal{S}_1) \cap \mathcal{N} \neq \emptyset$ which raises a contradiction. Reciprocally, let $\mathcal{S}_1 \subseteq \mathbb{R}^n$ be not ample. Then $\mathcal{N} \triangleq \mathbb{R}^n \setminus \mathrm{cl}(\mathrm{int}(\mathcal{S}_1))$ is open, and $\mathcal{S}_1 \cap \mathcal{N} \neq \emptyset = \mathrm{int}(\mathcal{S}_1) \cap \mathcal{N}$ by construction. Then $\mathcal{S}_1$ is locally thin.

Now let us prove the second statement. Let $\mathcal{S}_1 \subseteq \mathbb{R}^n$ be ample and $\mathcal{S}_2 \subseteq \mathbb{R}^n$ be open, and let $x \in \mathcal{S}_1 \cap \mathcal{S}_2$. We have $x \in \mathcal{S}_1 \subseteq \mathrm{cl}(\mathrm{int}(\mathcal{S}_1))$, so there exists $(x^k)_{k\in\mathbb{N}}$ converging to x with $x^k \in \mathrm{int}(\mathcal{S}_1)$ for all $k \in \mathbb{N}$. Since $x^k \to x \in \mathcal{S}_2$ with $\mathcal{S}_2$ open, it holds that $x^k \in \mathcal{S}_2$ for all k large enough. Finally, we have $x^k \in \mathrm{int}(\mathcal{S}_1) \cap \mathcal{S}_2 = \mathrm{int}(\mathcal{S}_1) \cap \mathrm{int}(\mathcal{S}_2) = \mathrm{int}(\mathcal{S}_1 \cap \mathcal{S}_2)$ for all k large enough. We deduce that $x \in \mathrm{cl}(\mathrm{int}(\mathcal{S}_1 \cap \mathcal{S}_2))$.

Next, let us prove the third statement. Let $\mathcal{S}_1 \subseteq \mathbb{R}^n$ having a dense intersection with $\mathcal{S}_2 \subseteq \mathbb{R}^n$ and let $\mathcal{S}_3 \subseteq \mathcal{S}_2$ be ample, and let $x \in \mathcal{S}_3$. We have $x \in \mathrm{cl}(\mathrm{int}(\mathcal{S}_3))$ so there exists $(x^k)_{k\in\mathbb{N}}$ converging to x with $x^k \in \mathrm{int}(\mathcal{S}_3) \subseteq \mathcal{S}_2 \subseteq \mathrm{cl}(\mathcal{S}_1 \cap \mathcal{S}_2)$ for all $k \in \mathbb{N}$. Then, for all $k \in \mathbb{N}$, there exists $(x^k_\ell)_{\ell\in\mathbb{N}}$ converging to x^k with $x^k_\ell \in \mathcal{S}_1 \cap \mathcal{S}_2 \cap \mathrm{int}(\mathcal{S}_3)$ for all $\ell \in \mathbb{N}$. For all $k \in \mathbb{N}$, let $\ell(k) \in \mathbb{N}$ be such that $\|x^k_{\ell(k)} - x^k\| \leq 2^{-k}$. It follows that $(x^k_{\ell(k)})_{k\in\mathbb{N}}$ converges to x and $x^k_{\ell(k)} \in \mathcal{S}_1 \cap \mathcal{S}_2 \cap \mathrm{int}(\mathcal{S}_3) \subseteq \mathcal{S}_1 \cap \mathcal{S}_3$ for all $k \in \mathbb{N}$. Hence, we get that $x \in \mathrm{cl}(\mathcal{S}_1 \cap \mathcal{S}_3)$.

Last, let us prove the fourth statement. Let $\mathcal{S}_1 \subseteq \mathbb{R}^n$ and let $\mathcal{S}_2 \subseteq \mathbb{R}^n$ be ample. Assume that $\mathcal{S}_1$ has a dense intersection with $\mathcal{S}_2$ and let $x \in \mathcal{S}_2$. Let $y \in \mathcal{S}_2$. Then $y \in \mathrm{cl}(\mathrm{int}(\mathcal{S}_2))$, so there exists $(y^k)_{k\in\mathbb{N}}$ converging to y with $y \neq y^k \in \mathrm{int}(\mathcal{S}_2) \subseteq \mathcal{S}_2$ for all $k \in \mathbb{N}$. Let $\varepsilon^k \triangleq \|y^k - y\| > 0$ for all $k \in \mathbb{N}$. Then $(\varepsilon^k)_{k\in\mathbb{N}}$ converges to 0. Now, remark that by dense intersection of $\mathcal{S}_1$ with $\mathcal{S}_2$, for all $k \in \mathbb{N}$ there exists $z^k \in \mathcal{S}_1 \cap \mathcal{S}_2 \cap \mathcal{B}_{\varepsilon^k}(y^k)$. Finally, let $\kappa \triangleq \min\{p \in \mathbb{N} : \max_{k\geq p}\|y^k - y\| \leq \frac{1}{2}\|y - x\|\}$ if $y \neq x$

and $\kappa \triangleq 0$ if $y = x$. Hence, for all $k \geq \kappa$ we have $x \notin \mathcal{B}_{\varepsilon^k}(y^k)$, and thus $z^k \in \mathcal{S}_1 \cap \mathcal{S}_2 \cap \mathcal{B}_{\varepsilon^k}(y^k) \setminus \{x\}$. Since $(z^k)_{k\in\mathbb{N}}$ converges to y, it follows that $y \in \mathrm{cl}((\mathcal{S}_1 \setminus \{x\}) \cap \mathcal{S}_2)$ as desired. □

Proposition 8. The function $\mathbb{O}$ defined as Oracle 1 satisfies Definition 1.

Proof. Consider $\mathbb{O}$ from Equation (1). The first requirement in Definition 1 is satisfied, since by construction $\mathbb{O}(y, \nu, \mathcal{S}) \subseteq \mathcal{M}(\nu) \cap \mathrm{cl}(\mathcal{B}_r)$ for all $y \in \mathbb{R}^n$, all $\nu \in \mathbb{R}^*_+$ and all $\mathcal{S} \subseteq \mathbb{R}^n$. To verify the second requirement, let $y \in \mathbb{R}^n$ and let $(y^k, \nu^k, \mathcal{S}^k) \in \mathbb{R}^n \times \mathbb{R}^*_+ \times 2^{\mathbb{R}^n}$ for all $k \in \mathbb{N}$ such that $(y^k)_{k\in\mathbb{N}}$ converges to y and $(\nu^k)_{k\in\mathbb{N}}$ converges to 0 and $\mathcal{S}^{k+1} \supseteq \mathcal{S}^k \cup (\{y^k\} + \mathbb{O}(y^k, \nu^k, \mathcal{S}^k))$ for all $k \in \mathbb{N}$. Then define $\kappa \in \mathbb{N}$ such that $y^k \in \mathcal{B}_r(y)$ for all $k \geq \kappa$. Define also $t^k \triangleq y^k + d^k \in \mathcal{S}^{k+1}$ for all $k \in \mathbb{N}$, where $d^k \in \mathbb{O}(y^k, \nu^k, \mathcal{S}^k)$. Thus, $\max_{d\in\mathcal{M}(\nu^k)\cap\mathrm{cl}(\mathcal{B}_r)} \mathrm{dist}(y^k + d, \mathcal{S}^k) = \mathrm{dist}(t^k, \mathcal{S}^k) \leq \min_{\ell<k} \left\|t^k - t^\ell\right\|$ for all $k \in \mathbb{N}^*$ by construction of $\mathbb{O}$. Also, by construction $(t^k)_{k\geq\kappa}$ lies in the compact set $\mathrm{cl}(\mathcal{B}_{2r}(y))$, and then $\min_{\ell<k} \left\|t^k - t^\ell\right\| \to 0$ as $k \to +\infty$. Indeed, assuming that this is not true, there exists $\varepsilon > 0$ and $K \subseteq \mathbb{N}$ infinite such that $\min_{\ell<k} \left\|t^k - t^\ell\right\| \geq \varepsilon$ for all $k \in K$, but this contradicts the Bolzano-Weierstrass theorem applied to $(t^k)_{k\in K}$. Hence the second requirement in Definition 1 also holds. □

Proposition 9. Under Assumption 2, consider the family $(X_i)_{i=1}^N$ it provides. For all $i \in [\![1, N]\!]$, denote by $\mathrm{cl}_f(X_i) \triangleq \{x \in \mathrm{cl}(X_i) : f_{|X_i\cup\{x\}} \text{ is continuous}\}$. Define $I(x) \triangleq \min\{i \in [\![1, N]\!] : x \in \mathrm{cl}_f(X_i)\}$ for all $x \in X$. Define $Y_i \triangleq \{x \in X : I(x) = i\}$ for all $i \in [\![1, N]\!]$. Then, $X = \sqcup_{i=1}^N Y_i$ and Y_i is an ample continuity set of f for all $i \in [\![1, N]\!]$.

Proof. Consider the notation from Proposition 9. By design, the sets $(Y_i)_{i=1}^N$ are pairwise disjoint and their union covers X, so $X = \sqcup_{i=1}^N Y_i$. Then we prove that for all $i \in [\![1, N]\!]$, Y_i is ample and $f_{|Y_i}$ is continuous. Let $i \in [\![1, N]\!]$. First, the properties of the sets $(X_i)_{i=1}^N$ and the construction of I lead to

$$\mathrm{int}(X_i) \underbrace{=}_{X_i \text{ open}} X_i \underbrace{\subseteq}_{I(X_i)=\{i\}} Y_i \underbrace{\subseteq}_{\substack{x\notin \mathrm{cl}_f(X_i) \implies I(x)\neq i \\ x\in \mathrm{cl}_f(X_i) \implies I(x)\leq i}} \mathrm{cl}_f(X_i) \underbrace{\subseteq}_{\text{by construction}} \mathrm{cl}(X_i).$$

Then, $\mathrm{int}(Y_i) \supseteq \mathrm{int}(X_i)$ and $Y_i \subseteq \mathrm{cl}(\mathrm{int}(X_i))$ so Y_i is ample. Moreover, let $x \in Y_i$ and let $(x^k)_{k\in\mathbb{N}}$ converging to x with $x^k \in Y_i$ for all $k \in \mathbb{N}$. For all $k \in \mathbb{N}$, $f_{|X_i\cup\{x^k\}}$ is continuous so there exists $(x^k_\ell)_{\ell\in\mathbb{N}}$ converging to x^k such that $x^k_\ell \in X_i$ for all $\ell \in \mathbb{N}$ and $(f(x^k_\ell))_{\ell\in\mathbb{N}}$ converges to $f(x^k)$. Let $\ell(k) \in \mathbb{N}$ such that $|f(x^k_{\ell(k)}) - f(x^k)| \leq 2^{-k}$ and $\|x^k_{\ell(k)} - x^k\| \leq 2^{-k}$. Then $(x^k_{\ell(k)})_{k\in\mathbb{N}}$ converges to x and $x^k_{\ell(k)} \in X_i$ for all $k \in \mathbb{N}$, so $(f(x^k_{\ell(k)}))_{k\in\mathbb{N}}$ converges to $f(x)$ by continuity of $f_{|X_i\cup\{x\}}$. Hence $(f(x^k))_{k\in\mathbb{N}}$ converges to $f(x)$. Thus $f_{|Y_i}$ is continuous at x, as desired. □

Proposition 10. A set has the interior cone property from Definition 2 if and only if its complement has the exterior cone property from [18, Definition 4.1] (quoted below).

> [**18, Definition 4.1**] A set $\mathcal{S} \subseteq \mathbb{R}^n$ is said to *have the exterior cone property* if at all points $x \in \partial\mathcal{S}$ there exists a cone $\mathcal{K}_x \triangleq \{x\} + \mathbb{R}^*_+\mathcal{U}$ (with $\emptyset \neq \mathcal{U} \subseteq \mathbb{S}^n$ open in induced topology) emanating from x, a neighborhood $\mathcal{O}_x$ of x and an angle $\theta > 0$ such that $\mathcal{E}_x \subseteq \mathcal{S}^c$ and $\Theta(e-x, a-x) \geq \theta$ for all $(e, a) \in \mathcal{E}_x \times \mathcal{S}_x$, where $\Theta(\cdot,\cdot)$ computes the unsigned internal angle between two vectors and $\mathcal{S}^c \triangleq (\mathbb{R}^n \setminus \mathcal{S})$ and $\mathcal{E}_x \triangleq (\mathcal{K}_x \cap \mathcal{O}_x)$ and $\mathcal{S}_x \triangleq (\mathcal{S} \cap \mathcal{O}_x \setminus \{x\})$.

Proof. Denote by (ICP) and (ECP) the interior and exterior cone properties stated in Definition 2, and by $[ECP]$ the exterior cone property stated in [2, Definition 4.1]. A set has the (ICP) if and only if its complement has the (ECP), so we only need to prove that the (ECP) is equivalent to the $[ECP]$. Denote by $[x, y] \triangleq \{x + t(y - x) : t \in [0, 1]\}$, for all $(x, y) \in (\mathbb{R}^n)^2$. Let $\mathcal{S} \subseteq \mathbb{R}^n$ and $\mathcal{S}^c \triangleq (\mathbb{R}^n \setminus \mathcal{S})$.

Assume that $\mathcal{S}$ has the $[ECP]$. For all $x \in \partial\mathcal{S}^c = \partial\mathcal{S}$, the $[ECP]$ of $\mathcal{S}$ applied to x provides the cone $\mathcal{K}_x$ and the neighborhood $\mathcal{O}_x$ such that $(\mathcal{K}_x \cap \mathcal{O}_x) \subseteq \mathcal{S}^c$. Then $\mathcal{S}^c$ has the (ICP), so $\mathcal{S}$ has the (ECP). Thus the $[ECP]$ implies the (ECP).

Assume that $\mathcal{S}$ has the (ECP). Let $x \in \partial\mathcal{S}$. The (ECP) of $\mathcal{S}$ applied to x provides $\emptyset \neq \mathcal{U} \subsetneq \mathbb{S}^n$ open in $\mathbb{S}^n$ and $\mathcal{K} \triangleq \mathbb{R}^*_+\mathcal{U}$ and the neighborhood $\mathcal{O}$ of 0 such that $(\mathcal{K}_x \cap \mathcal{O}_x) \subseteq \mathcal{S}^c$, where $\mathcal{K}_x \triangleq \{x\} + \mathcal{K}$ and $\mathcal{O}_x \triangleq \{x\} + \mathcal{O}$. Let $\theta > 0$ small enough so that the open set $\mathcal{U}' \triangleq \{u \in \mathcal{U} : \Theta(u, v) > \theta, \ \forall v \in \mathbb{S}^n \setminus \mathcal{U}\}$ is nonempty. Define $\mathcal{K}' \triangleq \mathbb{R}^*_+\mathcal{U}'$ and $\mathcal{K}'_x \triangleq \{x\} + \mathcal{K}' \subseteq \mathcal{K}_x$ and $\mathcal{E}_x \triangleq (\mathcal{K}'_x \cap \mathcal{O}_x)$ and $\mathcal{S}_x \triangleq (\mathcal{S} \cap \mathcal{O}_x \setminus \{x\})$. Then $\mathcal{E}_x \subseteq \mathcal{S}^c$ and θ satisfies the requirement in the $[ECP]$. Indeed, $\mathcal{E}_x \subseteq \mathcal{K}_x$ while $\mathcal{S}_x \cap \mathcal{K}_x = \emptyset$, thus for all $(e, a) \in \mathcal{E}_x \times \mathcal{S}_x$ there exists $y \in \partial\mathcal{K}_x \cap [e, a] \neq \emptyset$. Since e and y and a belong to the same line, we get $\Theta(e - x, a - x) = \Theta(e - x, y - x) + \Theta(y - x, a - x) = \Theta(\frac{e-x}{\|e-x\|}, \frac{y-x}{\|y-x\|}) + \Theta(y - x, a - x) \geq \theta + 0$. The first term is greater than θ since $\frac{e-x}{\|e-x\|} \in \mathcal{U}'$ while $\frac{y-x}{\|y-x\|} \in \mathbb{S}^n \setminus \mathcal{U}$, and the second term is positive by definition of Θ. Thus, $\mathcal{S}$ has the $[ECP]$ at x. Hence the (ECP) implies the $[ECP]$. ☐

Proposition 11. If $\mathcal{S} \subseteq \mathbb{R}^n$ has the ICP from Definition 2, then $\mathcal{S}$ is ample. The reciprocal is not true.

Proof. Let $\mathcal{S} \subseteq \mathbb{R}^n$ having the ICP. Let $x \in \mathcal{S}$. The ICP provides $\mathcal{K}_x$ and $\mathcal{O}_x$ satisfying $(\mathcal{K}_x \cap \mathcal{O}_x) \subseteq \mathcal{S}$ and $x \in \mathrm{cl}(\mathcal{K}_x \cap \mathcal{O}_x)$. Moreover, $\mathcal{K}$ is open as the image of $\mathbb{R}^*_+ \times \mathcal{U}$ (an open subset of $\mathbb{R}^*_+ \times \mathbb{S}^n$) by the homeomorphism $(\lambda, u) \in \mathbb{R}^*_+ \times \mathbb{S}^n \mapsto \lambda u \in \mathbb{R}^n \setminus \{0\}$. Thus, $\mathcal{K}_x \cap \mathcal{O}_x$ is also open, so $(\mathcal{K}_x \cap \mathcal{O}_x) \subseteq \mathrm{int}(\mathcal{S})$. Hence, $x \in \mathrm{cl}(\mathrm{int}(\mathcal{S}))$, which proves the direct implication. To observe that the reciprocal implication fails, consider $n \triangleq 2$ and $\mathcal{S} \triangleq \mathrm{epi}(\sqrt{|\cdot|})$. Then $\mathcal{S}$ is ample but the ICP fails at $x = (0, 0) \in \mathrm{cl}(\mathcal{S})$. ☐

References

[1] C. Audet, A. Batailly, and S. Kojtych. Escaping unknown discontinuous regions in blackbox optimization. SIAM Journal on Optimization, 32(3):1843–1870, 2022.

[2] C. Audet, P.-Y. Bouchet, and L. Bourdin. Counterexample and an additional revealing poll step for a result of "analysis of direct searches for discontinuous functions". Mathematical Programming, 2024.

[3] C. Audet and J.E. Dennis, Jr. Analysis of generalized pattern searches. SIAM Journal on Optimization, 13(3):889–903, 2003.

[4] C. Audet and J.E. Dennis, Jr. Mesh Adaptive Direct Search Algorithms for Constrained Optimization. SIAM Journal on Optimization, 17(1):188–217, 2006.

[5] C. Audet and J.E. Dennis, Jr. A Progressive Barrier for Derivative-Free Nonlinear Programming. SIAM Journal on Optimization, 20(1):445–472, 2009.

[6] C. Audet and W. Hare. Derivative-Free and Blackbox Optimization. Springer Series in Operations Research and Financial Engineering. Springer, Cham, Switzerland, 2017.

[7] C. Audet and W. Hare. Model-based methods in derivative-free nonsmooth optimization. In A.M. Bagirov, M. Gaudioso, N. Karmitsa, M.M. Mäkelä, and S. Taheri, editors, Numerical nonsmooth optimization, chapter 15. Springer, 2020.

[8] A.S. Berahas, L. Cao, and K. Scheinberg. Global convergence rate analysis of a generic line search algorithm with noise. SIAM Journal on Optimization, 31(2):1489–1518, 2021.

[9] A.R. Conn, K. Scheinberg, and L.N. Vicente. Introduction to Derivative-Free Optimization. MOS-SIAM Series on Optimization. SIAM, Philadelphia, 2009.

[10] R. Fletcher and S. Leyffer. Nonlinear programming without a penalty function. Mathematical Programming, Series A, 91:239–269, 2002.

[11] S. Gratton, C.W. Royer, L.N. Vicente, and Z. Zhang. Direct search based on probabilistic descent. SIAM Journal on Optimization, 25(3):1515–1541, 2015.

[12] J.H. Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numerische Mathematik, 2(1):84–90, 1960.

[13] J. Larson, M. Menickelly, and S.M. Wild. Derivative-free optimization methods. Acta Numerica, 28:287–404, 2019.

[14] S. Le Digabel and S.M. Wild. A Taxonomy of Constraints in Simulation-Based Optimization. Technical Report G-2015-57, Les cahiers du GERAD, 2015. To appear in Optimization and Engineering.

[15] N. Levine. Semi-open sets and semi-continuity in topological spaces. American Mathematical Monthly, 70:36–41, 1963.

[16] A. Meijster, J.B.T.N. Roerdink, and W.H. Hesselink. A General Algorithm for Computing Distance Transforms in Linear Time, pages 331–340. Springer US, Boston, MA, 2000.

[17] R.T. Rockafellar. Generalized directional derivatives and subgradients of nonconvex functions. Canadian Journal of Mathematics. Journal Canadien de Mathématiques, 32(2):257–280, 1980.

[18] L.N. Vicente and A.L. Custódio. Analysis of direct searches for discontinuous functions. Mathematical Programming, 133(1-2):299–325, 2012.

[19] D. Zhan and H. Xing. Expected improvement for expensive optimization: a review. Journal of Global Optimization, 78(3):507–544, 2020.