# A Merit Function Approach for Direct Search

S. Gratton[*] L. N. Vicente[†]

May 2, 2014

**Abstract**

In this paper it is proposed to equip direct-search methods with a general procedure to minimize an objective function, possibly non-smooth, without using derivatives and subject to constraints on the variables.

One aims at considering constraints, most likely nonlinear or non-smooth, for which the derivatives of the corresponding functions are also unavailable. The novelty of this contribution relies mostly on how relaxable constraints are handled. Such constraints, which can be relaxed during the course of the optimization, are taken care by a merit function and, if necessary, by a restoration procedure. Constraints that are unrelaxable, when present, are treated by an extreme barrier approach.

One is able to show that the resulting merit function direct-search algorithm exhibits global convergence properties for first-order stationary constraints. As in the progressive barrier method [6], we provide a mechanism to indicate the transfer of constraints from the relaxable set to the unrelaxable one.

**Keywords:** Derivative-free optimization, direct-search methods, constraints, merit function, penalty parameter, random directions, non-smooth optimization.

## 1 Introduction

Consider the problem

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in \Omega = \Omega_r \cap \Omega_{nr}. \end{array} \tag{1}$$

The feasible region of this problem is defined by relaxable and/or unrelaxable constraints. The non-relaxable constraints correspond to $\Omega_{nr} \subseteq \mathbb{R}^n$. Such constraints have to be satisfied at all iterations in an algorithmic framework for which the objective function is evaluated. Typically they are bounds or linear constraints but they can also include hidden constraints (constraints

[*]ENSEEIHT, INPT, rue Charles Camichel, B.P. 7122 31071, Toulouse Cedex 7, France (serge.gratton@enseeiht.fr).

[†]CMUC, Department of Mathematics, University of Coimbra, 3001-501 Coimbra, Portugal (lnv@mat.uc.pt). Support for this resarch was provided by FCT under grants PTDC/MAT/116736/2010 and PEst-C/MAT/UI0324/2011 and by the Réseau Thématique de Recherche Avancée, Fondation de Coopération Sciences et Technologies pour l'Aéronautique et l'Espace, under the grant ADTAO.

which are not part of the problem specification/formulation and their manifestation comes in the form of some indication that the objective function could not be evaluated). In contrast, relaxable constraints need only to be satisfied approximately or asymptotically. In our notation $\Omega_r$ is the set of relaxable constraints, which is assumed to take the form

$$\Omega_r \;=\; \{x \in \mathbb{R}^n : \, c_i(x) \leq 0, \; \forall i \in \mathcal{I}\}\,.$$

Finally, the objective function $f : \mathbb{R}^n \to \mathbb{R}$ and the relaxable constraint functions $c_i$ are only assumed to be locally Lipschitz continuous (in the sense that the convergence analysis holds if $f$ and the $c_i$'s are Lipschitz continuous near an accumulation point produced by the algorithm). Most of the globally convergent derivative-free approaches for handling nonlinear constrained problems have been of direct search or line search type[1].

Feasible methods may be the only option when all the constraints are unrelaxable ($\Omega_r = \mathbb{R}^n$). In addition they generate a sequence of feasible points, thus allowing the iterative process to be terminated prematurely with a guarantee of feasibility for the best point tested so far. One way of designing feasible methods is by means of the barrier function

$$f_{\Omega_{nr}}(x) \;=\; \begin{cases} f(x) & \text{if } x \in \Omega_{nr}, \\ +\infty & \text{otherwise.} \end{cases} \tag{2}$$

Following the notation in [5], we refer to such a barrier function as the extreme barrier function. It is not necessary to evaluate $f$ at infeasible points since the value of the extreme barrier function is set to $+\infty$ at such points. Direct-search methods take action solely based on function values comparisons and are thus appropriate to use in conjunction with an extreme barrier function. In the context of direct-search methods of directional type using such functions, there are two known ways of designing globally convergent algorithms. In any of the cases, one must use sets of directions whose union (after normalization if needed) is asymptotically dense in the unit sphere of $\mathbb{R}^n$, even if the objective function is smooth. The first approach requires only a simple decrease to accept new iterates but imposes integer requirements throughout the algorithm (and in particular in the generation of the directions). This approach is known as mesh adaptive direct-search (MADS) and has been developed by Audet and Dennis [5]. One can, however, relax such integer lattice requirements and freely generate the directions densely in the unit sphere at the price of imposing a sufficient decrease condition on the acceptance of new iterates (see Vicente and Custódio [25]) — in practice, sufficient decrease can be imposed as not to differ much from simple decrease. An alternative to extreme barrier when designing feasible methods is the use of projections onto the feasible set, although this might require the knowledge of the derivatives of the constraints and be expensive or unpractical in many instances (see Lucidi, Sciandrone, and Tseng [21] for such an approach).

In the case where there are no unrelaxable constraints, one can use a penalty term by adding to the objective function a measure of constraint violation multiplied by a penalty parameter, and thus allowing to start infeasible with respect to the relaxable constraints. In this vein, Lewis and Torczon [18] (see also [17]) suggested an approach based on an augmented Lagrangian method. They consider the solution of a sequence of subproblems where the augmented Lagrangian function takes into account only the nonlinear constraints and is minimized subject to the

[1]On the model-based trust-region side of optimization without derivatives, nonlinear constraints have been considered mostly in implementations (see [8, 9, 11, 12, 24]), and as far as we know no convergence theory has yet been developed.

remaining constraints (bounds on the variables or more general linear constraints). Each problem can then be approximately solved using an appropriate directional direct-search method. This application of augmented Lagrangian methods yields global convergence results to first-order stationary points of the same type of those obtained under the presence of derivatives. Diniz-Ehrhardt, Martínez, and Pedroso [15] studied a more general augmented Lagrangian setting where the problem constraints imposed as subproblem constraints are not necessarily of linear type. In turn, Liuzzi and Lucidi [19] and Liuzzi, Lucidi, and Sciandrone [20] developed and analyzed algorithms for inequality constrained problems, based on nonsmooth and smooth, respectively, penalty functions. They imposed sufficient decrease and handled bound and linear constraints separately, proving that a subset of the set of limit points of the sequence of iterates satisfy the first-order necessary conditions of the original problem. Martínez and Sobral [22] proposed an algorithm for problems with 'thin' constraints based on relaxing feasibility and performing a subproblem restoration procedure. Filter methods may also be appropriate to handle relaxable constraints, and the first step in this direction was done by Audet and Dennis [4]. The filter approach of Dennis, Price, and Coope [14] guarantees global convergence to a first-order stationary point by means of an envelope around the filter as means of measuring sufficient decrease.

The first general approach to consider both relaxable and unrelaxable constraints is called progressive barrier and has been suggested by Audet and Dennis [6], exhibiting some global convergence properties. It allows the handling of both types of constraints, by combining mesh adaptive direct search for unrelaxable constraints with non-dominance filter type concepts for the relaxable constraints (see the consequent developments in [7]). An interesting feature is that a constraint can be considered relaxable until it becomes feasible whereupon it is transferred to the set of unrelaxable constraints.

In this paper, we develop an alternative approach to progressive barrier [6], handling the relaxable constraints by means of a merit function instead of a filter. For such a purpose, we consider a constraint violation function of the type

$$g(x) \;=\; \sum_{i \in \mathcal{I}} \max(c_i(x), 0) \tag{3}$$

and the merit function

$$M(x;\mu) \;=\; f(x) + \mu g(x), \tag{4}$$

where $\mu \in \mathbb{R}$ is a positive penalty parameter. The merit function and the corresponding penalty parameter are only used in the evaluation of an already computed step, to decide whether it will be accepted or not. The merit function (4) using (3) is known in Nonlinear Programming (see [23, Section 17.2]) as the $\ell_1$ penalty function and has been extensively used in implementations (see [23, Section 17.5]).

Our treatment of the non-relaxable constraints will implicitly consider the use of extreme barrier functions of the type (2). In practice what we optimize is $f_{\Omega_{nr}}$ since the non-relaxable constraints restrict the evaluation of the objective function $f$. For generality, one considers here that $\Omega_{nr}$ also constrains the evaluation of the relaxable constraints, and thus implicitly consider $g_{\Omega_{nr}}$ instead of $g$ in our proposed algorithm. Due to the presence of (derivative-free) unrelaxable constraints and/or of the non-smoothness of the objective function, the directions used in the algorithm must be generated densely in the unit sphere of $\mathbb{R}^n$.

Our merit function approach has been designed in a simple and modular way. A successful iteration is defined by a sufficient decrease in the constraint violation measure (3) (sufficiently away from feasibility) or a sufficient decrease in the merit function (4) for an appropriate value of the penalty parameter. Whenever a sufficient decrease in the constraint violation measure (3) is observed at the expense of a significant increase in the objective function, a restoration of feasibility mode is entered with the single purpose of minimizing (3).

The paper is organized as follows. We start by describing the merit function algorithm in Section 2. The convergence theory of the proposed approach is then divided in four sections: Section 3 for the behavior of the step size parameter; Section 4 for the case where restoration is only entered a finite number of times; Section 5 for the case where restoration is entered but never left; Section 6 for the case where restoration is entered an infinite number of times. In Section 7 we discuss how the theory particularizes in the presence of smoothness. In Section 8 we show a few runs of the algorithm as a proof of concept. Finally, Section 9 contains some concluding remarks and Appendix A summarizes a few notions of Clarke non-smooth calculus needed in the paper.

## 2 A merit function algorithm

In our algorithm framework an iteration is considered successful in two situations. To describe them in some detail let us assume a given iterate $x_k$ and a step size $\alpha_k > 0$. Each iteration is divided in a search and a poll step, but the latter is the one responsible for the convergence properties of the algorithm (and thus we ignore the search step for most of this discussion). Let also $d$ be a direction considered in the poll step and $\rho(\alpha)$ a forcing function, i.e., a positive and non-decreasing function verifying $\lim_{\alpha\downarrow 0} \rho(\alpha)/\alpha = 0$. The directions used in the poll step belong to a set $D_k$ which does not necessarily have to span $\mathbb{R}^n$ with non-negative coefficients as it happens in traditional direct-search methods for smooth problems.

The first possibility of success is that a certain sufficient decrease in the constraint violation measure $g$ is attained ($g(x_k+\alpha_k d) < g(x_k)-\rho(\alpha_k)$) and one is sufficiently away from the feasible region $g(x_k) > C\rho(\alpha_k)$, for some constant $C > 1$.

The other situation where success is declared is when the merit function is sufficiently decreased ($M(x_k+\alpha_k d_k;\mu_k) < M(x_k;\mu_k)-\rho(\alpha_k)$) for a certain choice of the penalty parameter $\mu_k$. The update of the penalty parameter follows a classical one [23, Formula (18.33)] since what we use in (5) below is essentially the formula

$$\frac{[f(x_k + \alpha_k d_k) - f(x_k)]/\alpha_k}{\rho(\alpha_k)/\alpha_k},$$

where the nominator corresponds to $\nabla f(x_k)^\top d_k$ in the classical update (where $f$ is typically continuously differentiable) and the denominator replaces the value of $g(x_k)$ in the classical update (and we will observe later that when $\rho(\alpha_k)/\alpha_k$ goes to zero so does in principle $g(x_k)$, see Theorems 4.1, 5.1-ii, and 6.1 and their proofs). We summarize below the definition of a successful point (to be used in both search and poll steps).

**Begin (successful point).**
Given $x_k$ and $\alpha_k$, a point $y_k$ (either in the search or in the poll step) is successful if

$$g(y_k) \ < \ g(x_k) - \rho(\alpha_k) \quad \text{and} \quad g(x_k) \ > \ C\rho(\alpha_k)$$

or, if that is false, if

$$M(y_k;\mu_k) \ < \ M(x_k;\mu_k) - \rho(\alpha_k),$$

where

$$\mu_k \;=\; \max\left\{\bar{\mu}, \frac{f(y_k) - f(x_k)}{C\rho(\alpha_k)}\right\} \tag{5}$$

and $\bar{\mu} > 0$ and $C > 1$ are constants independent of $k$.
**End (successful point).**

However, before measuring success, our algorithm framework considers a phase to restore feasibility or decrease the amount of constraint violation. A Restoration is entered (in the poll step) when there exists a $d \in D_k$ such that $g(x_k + \alpha_k d) < g(x_k) - \rho(\alpha_k)$, $g(x_k) > C\rho(\alpha_k)$, and $M(x_k + \alpha_k d; \bar{\mu}) \geq M(x_k; \bar{\mu})$, for a sufficiently large value $\bar{\mu}$ of the penalty parameter. Notice that the first and third of these conditions imply

$$f(x_k + \alpha_k d) - f(x_k) \;\geq\; \bar{\mu}[g(x_k) - g(x_k + \alpha_k d)] \;>\; \bar{\mu}\rho(\alpha_k).$$

Thus, when Restoration is entered it is because a direction $d$ has been found for which $g$ is sufficiently reduced ($g(x_k + \alpha_k d) < g(x_k) - \rho(\alpha_k)$) at a point $x_k$ sufficiently away from being feasible ($g(x_k) > C\rho(\alpha_k)$) and for which $f$ has considerably increased ($f(x_k + \alpha_k d) - f(x_k) > \bar{\mu}\rho(\alpha_k)$). Restoration can also be entered in the search step and so we define below the notion of a Restoration identifier in general terms, to be used in both search and poll steps.

**Begin (Restoration identifier).**
Given $x_k$ and $\alpha_k$, a point $y_k$ is a Restoration identifier (either in the search or in the poll step) if

$$g(y_k) \;<\; g(x_k) - \rho(\alpha_k) \quad \text{and} \quad g(x_k) \;>\; C\rho(\alpha_k)$$

and

$$M(y_k; \bar{\mu}) \;\geq\; M(x_k; \bar{\mu}),$$

where $\bar{\mu} > 0$ and $C > 1$ are constants independent of $k$.
**End (Restoration identifier).**

Our merit function approach is described below in Algorithm 2.1. All directions in the sets $D_k$ for all $k$ are considered normalized.

**Algorithm 2.1 (A merit function algorithm (Main))**

**Initialization**

Choose $x_0 \in \Omega_{nr}$, $\alpha_0, \bar{\mu} > 0$, $C > 1$, $0 < \beta_1 \leq \beta_2 < 1$, and $\gamma \geq 1$.

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

1. **Search step (optional):** Evaluate the functions $f$ and $g$ at a finite number of points in $\Omega_{nr}$. Enter Restoration (with $k_r = k$) if any of those points is a **Restoration identifier**. Otherwise, if any of those points (say $x$) is a **successful point**, then set $x_{k+1} = x$, declare the iteration and the search step successful, and skip the poll step.
2. **Poll step:** Select a finite subset of directions $D_k$. If $x_k + \alpha_k d \notin \Omega_{nr}$ for all $d \in D_k$, the iteration is declared unsuccessful. Otherwise, remove from $D_k$ all directions $d$ such that $x_k + \alpha_k d \notin \Omega_{nr}$.

   If any of the points $x_k + \alpha_k d$, with $d \in D_k$, is a **Restoration identifier**, then enter Restoration (with $k_r = k$).

Otherwise, if there is a **successful point** of the form $x_k + \alpha_k d_k$ with $d_k \in D_k$, then $x_{k+1} = x_k + \alpha_k d_k$ and declare the iteration and the poll step successful.

Otherwise, declare the iteration unsuccessful and set $x_{k+1} = x_k$.

3. **Step size 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: $\alpha_{k+1} \in [\beta_1\alpha_k, \beta_2\alpha_k]$.

As we said before, if $g$ can be sufficiently reduced (sufficiently away from feasibility) while $f$ is considerably increased, we need to focus totally on a reduction of the constraint violation, and such procedure is described below in Algorithm 2.2. Restoration is then left when progress in the reduction of the constraint violation cannot be further achieved and such a considerable increase in $f$ is no longer observed (we will later see in Section 5 the appropriateness of such a leaving criterion).

**Algorithm 2.2 (A merit function algorithm (Restoration))**

**Initialization**

Start from $x_{k_r} \in \Omega_{nr}$ given from the Main algorithm and consider the same parameters as in there.

**For** $k = k_r, k_r + 1, k_r + 2, \ldots$

1. **Search step (optional):**
   Evaluate the function $g$ at a finite number of points in $\Omega_{nr}$. If any of those points (say $x$) is such that $g(x) < g(x_k) - \rho(\alpha_k)$ and $g(x_k) > C\rho(\alpha_k)$, then set $x_{k+1} = x$, declare the iteration and the search step successful, and skip the poll step.
2. **Poll step:** Select a finite subset of directions $D_k$. If $x_k + \alpha_k d \notin \Omega_{nr}$ for all $d \in D_k$, the iteration is declared unsuccessful. Otherwise, remove from $D_k$ all directions $d$ such that $x_k + \alpha_k d \notin \Omega_{nr}$.
   Declare the poll step and the iteration successful if there exists a $d_k \in D_k$ such that
   $$g(x_k + \alpha_k d_k) \; < \; g(x_k) - \rho(\alpha_k) \quad \text{and} \quad g(x_k) \; > \; C\rho(\alpha_k)$$
   In such a case, set $x_{k+1} = x_k + \alpha_k d_k$.
   Otherwise, declare the iteration unsuccessful and set $x_{k+1} = x_k$.
   Leave Restoration and return to the Main algorithm (starting at a new $(k+1)$-th iteration using $x_{k+1}$ and $\alpha_{k+1}$) if the iteration is unsuccessful and $M(x_k + \alpha_k d; \bar{\mu}) < M(x_k; \bar{\mu})$ for some $d \in D_k$.
3. **Step size parameter update:** As in Step 3 of the Main algorithm.

# 3 Step size behavior

As it is classic in direct-search methods or other techniques for derivative-free optimization, we start our analysis of global convergence by showing that the step size parameter approaches zero. We will do this under the condition that Restoration is not entered an infinite number of times (and postpone to Section 6 the analysis of this situation).

**Theorem 3.1** *Assume that $f$ is bounded below. Assume that Restoration is entered finitely many times.*

*Then,*

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

**Proof.** Suppose that there exists $\bar{k} \in \mathbb{N}$ and $\bar{\alpha} > 0$ such that $\alpha_k \geq \bar{\alpha}$ and the $k$-th iteration is a Main one for every $k \geq \bar{k}$.

Let us assume now that there exists an infinite subsequence $J_1$ of successful iterations after $\bar{k}$. We thus know that $x_k \in \Omega_{nr}$ $\forall k \in J_1$. In the derivation below we will omit the unsuccessful iterations, since at those iterations the iterates do not move.

If $[g(x_{k+1}) < g(x_k) - \rho(\alpha_k)$ and $g(x_k) > C\rho(\alpha_k)]$ is true for sufficiently large $k \in J_1$, then

$$g(x_{k+1}) \;<\; g(x_k) - \rho(\alpha_k) \;\leq\; g(x_k) - \rho(\bar{\alpha})$$

for those indices $k$, which renders a contradiction since $g$ is bounded below by 0.

Thus, there must exists an infinite subsequence $J_2 \subseteq J_1$ of iterates for which $M(x_{k+1}; \mu_k) < M(x_k; \mu_k) - \rho(\alpha_k)$. Here we consider two possibilities.

In the first case, all these iterates are such that $\mu_k = \bar{\mu}$ for sufficiently large $k$. In such an occurrence one has that

$$M(x_{k+1}; \bar{\mu}) \;<\; M(x_k; \bar{\mu}) - \rho(\alpha_k) \;\leq\; M(x_k; \bar{\mu}) - \rho(\bar{\alpha})$$

for all $k \in J_2$ sufficiently large. However, in the successful iterations where $[g(x_{k+1}) < g(x_k) - \rho(\alpha_k)$ and $g(x_k) > C\rho(\alpha_k)]$, since Restoration was not entered ($\bar{k}$ was considered sufficiently large for this purpose), one knows that $M(x_{k+1}; \bar{\mu}) < M(x_k; \bar{\mu})$. Thus, $M(x_k; \bar{\mu})$ tends to $-\infty$ which is a contradiction given the boundedness from below of both $f$ and $g$.

In the second possibility, there is an infinite number of iterations in $J_2$ such that

$$\mu_k \;=\; \frac{f(x_{k+1}) - f(x_k)}{C\rho(\alpha_k)}.$$

Let us choose just one of these iterations. For such an iteration $k$, either $g(x_{k+1}) \geq g(x_k) - \rho(\alpha_k)$ or $g(x_k) \leq C\rho(\alpha_k)$. Thus, either

$$f(x_{k+1}) - f(x_k) \;=\; \mu_k C\rho(\alpha_k) \;\geq\; \mu_k[g(x_k) - g(x_{k+1})]$$

(since $C > 1$) or

$$f(x_{k+1}) - f(x_k) \;=\; \mu_k C\rho(\alpha_k) \;\geq\; \mu_k g(x_k) \;\geq\; \mu_k[g(x_k) - g(x_{k+1})],$$

both leading to $M(x_{k+1}; \mu_k) \geq M(x_k; \mu_k)$ which contradicts $M(x_{k+1}; \mu_k) < M(x_k; \mu_k) - \rho(\alpha_k)$.

We have proved under the assumption of contradiction that one cannot have an infinity of successful iterations. But if all iterations are unsuccessful after a certain order that also contradicts the assumption of contradiction. $\square$

The following corollary organizes the relevant information regarding unsuccessful iterations and step size behavior for the analysis in the next sections.

**Corollary 3.1** *Assume that $f$ is bounded below. Assume that Restoration is entered finitely many times.*

*Then, there exists at least one refining subsequence of Main iterations (i.e., a subsequence $K$ composed of unsuccessful Main iterations for which $\alpha_k \to 0$ for $k \in K$).*

**Proof.** The proof can be found for instance in [13] but it is given here for completeness. From Theorem 3.1 we conclude that there must exist a subsequence $J$ of unsuccessful iterations (or unsuccessful poll steps). Thus, from the way we update the step size parameter, there must exist a subsequence of unsuccessful iterations $K \subset J$ such that $\alpha_{k+1} \to 0$ for $k \in K$. Since, $\alpha_k \leq (1/\beta_1)\alpha_{k+1}$ for $k \in K$, we obtain $\alpha_k \to 0$ for $k \in K$. $\square$

# 4 Convergence assuming restoration is never entered after a certain order

The analysis of global convergence of Algorithm 2.1 is made by inspecting the sign of appropriate Clarke directional derivatives. Let $h$ (e.g., $h = f, g$) be Lipschitz continuous near $x_*$ and restricted to $\Omega_{nr} \subseteq \mathbb{R}^n$. We will use the following definition of the Clarke generalized derivative of $h$ at $x_*$ along $d$

$$h^\circ(x_*; d) \;=\; \limsup_{\substack{x \to x_*,\, x \in \Omega_{nr} \\ t \downarrow 0,\, x + td \in \Omega_{nr}}} \frac{h(x+td) - h(x)}{t},$$

where $d$ must be in the hypertangent $T^H_{\Omega_{nr}}(x_*)$ cone to $\Omega_{nr}$ at $x_*$ (i.e., $d$ must be in the interior of the tangent cone $T^{Cl}_{\Omega_{nr}}(x_*)$ to $\Omega_{nr}$ at $x_*$). In the Appendix of this paper we provide the rigorous definitions of these derivatives as well as the definitions of tangent and hypertangent cones. We assume throughout this paper that the hypertangent $T^H_{\Omega_{nr}}(x_*)$ is nonempty.

The sign of the Clarke derivatives is then analyzed at limit points of refining subsequences along refining directions. As we said before, by a refining subsequence [3], we mean a subsequence of unsuccessful Main iterates for which the step-size parameter converges to zero. By a refining direction [5] (in $T^H_{\Omega_{nr}}(x_*)$) associated with a refining subsequence $K$ converging to $x_*$, one means a limit point of $\{d_k\}$ (in $T^H_{\Omega_{nr}}(x_*)$) where $k \in K$ is taken sufficiently large such that $x_k + \alpha_k d_k \in \Omega_{nr}$. Given that our working directions in the sets $D_k$'s are normalized so are the refining directions.

## 4.1 Results on feasibility

We start by considering the determination of feasibility. (Note that since $\Omega_{nr}$ is not necessarily by assumption a closed set, one must assume below that the limit point of a refining subsequence verifies the non-relaxable constraints.)

**Theorem 4.1** *Assume that $f$ is bounded below. Assume that Restoration is entered finitely many times.*

*Let $\{x_k\}_{k \in K}$ be a refined subsequence converging to $x_* \in \Omega_{nr}$ and assume that $d \in T^H_{\Omega_{nr}}(x_*)$ is a refining direction associated with $K$ and $x_*$. Assume that $g$ is Lipschitz continuous near $x_*$. Then either $g(x_*) = 0$ (implying $x_* \in \Omega_r$ and thus $x_* \in \Omega$) or $g^\circ(x_*; d) \geq 0$.*

**Proof.** By assumption there exists a subsequence $K_1 \subseteq K$ and a corresponding subsequence $\{d_k\}_{k\in K_1}$ of polling directions such that $\{d_k\}$ converges to $d \in T^H_{\Omega_{nr}}(x_*)$ in $K_1$ and $\alpha_k$ goes to zero in $K_1$. Thus, one must necessarily have that $x_k + \alpha_k d_k \in \Omega_{nr}$ for $k$ sufficiently large in $K_1$.

Since the iteration $k \in K_1$ is unsuccessful, $g(x_k + \alpha_k d_k) \geq g(x_k) - \rho(\alpha_k)$ or $g(x_k) \leq C\rho(\alpha_k)$, and then either there exists an infinite number of the first or of the second. In the latter case, it is then trivial to obtain $g(x_*) = 0$ from the fact that $\alpha_k \to 0$ in $K_1$ and the continuity of $g$. In the former case, there exists a subsequence $K_2 \subseteq K_1$ such that

$$\frac{g(x_k + \alpha_k d_k) - g(x_k)}{\alpha_k} \;\geq\; -\frac{\rho(\alpha_k)}{\alpha_k} \quad \forall k \in K_2.$$

On the other hand, from the definitions of lim sup and $K_2$,

$$\limsup_{\substack{x \to x_*, x \in \Omega_{nr} \\ t \downarrow 0, x+td \in \Omega_{nr}}} \frac{g(x+td) - g(x)}{t} \;\geq\; \limsup_{k \in K_2} \frac{g(x_k + \alpha_k d) - g(x_k)}{\alpha_k}.$$

Since $g$ is Lipschitz continuous near $x_*$ (with constant $L_g$),

$$\frac{g(x_k + \alpha_k d_k) - g(x_k)}{\alpha_k} - L_g\|d_k - d\| \;\leq\; \frac{g(x_k + \alpha_k d) - g(x_k)}{\alpha_k}.$$

One then obtains $g^\circ(x_*; d) \geq 0$ since both $\|d_k - d\|$ and $\rho(\alpha_k)/\alpha_k$ tend to zero in $K_2$. $\square$

By assuming that appropriate refining directions are dense in $T^{Cl}_{\Omega_{nr}}(x_*) \cap \{d \in \mathbb{R}^n : \|d\| = 1\}$, one can show that the limit point $x_*$ is Clarke stationary for the constraint violation problem

$$\begin{array}{ll} \min & g(x) \\ \text{s.t.} & x \in \Omega_{nr}. \end{array} \tag{6}$$

**Theorem 4.2** *Assume that $f$ is bounded below. Assume that Restoration is entered finitely many times.*

*Let $\{x_k\}_{k\in K}$ be a refined subsequence converging to $x_* \in \Omega_{nr}$. Assume that $g$ is Lipschitz continuous near $x_*$.*

*Assume that $T^{Cl}_{\Omega_{nr}}(x_*)$ has a non-empty interior.*

*Then either $g(x_*) = 0$ (implying $x_* \in \Omega_r$ and thus $x_* \in \Omega$) or if the set of refining directions associated with $K'$ (where $K'$ is formed by the indices in $K$ such that $g(x_k + \alpha_k d_k) \geq g(x_k) - \rho(\alpha_k)$) and $x_*$ is dense in $T^{Cl}_{\Omega_{nr}}(x_*) \cap \{d \in \mathbb{R}^n : \|d\| = 1\}$, then $g^\circ(x_*; v) \geq 0$ for all $v \in T^{Cl}_{\Omega_{nr}}(x_*)$, and $x_*$ is a stationary point of the constraint violation problem (6).*

**Proof.** Following the proof of Theorem 4.1, if there exists an infinite number of cases where $g(x_k) \leq C\rho(\alpha_k)$, then $g(x_*) = 0$.

Now, let $v$ be such that $v \in T^{Cl}_{\Omega_{nr}}(x_*)$ and $\|v\| = 1$. Then $v$ is the limit of a sequence $\mathcal{D}$ of refining directions $d$ associated with $K'$ and $x_*$ such that $d \in T^H_{\Omega_{nr}}(x_*)$. For each such $d$ one can apply the proof of Theorem 4.1 and obtain $g^\circ(x_*; d) \geq 0$. Thus, $g^\circ(x_*; v) = \lim_{d \in T^H_{\Omega_{nr}}(x_*), d \in \mathcal{D}} g^\circ(x_*; d) \geq 0$. The result then holds for non-normalized $v$'s given that $T^{Cl}_{\Omega_{nr}}(x_*)$ is a cone and the Clarke derivatives are homogeneous in their second arguments. $\square$

## 4.2 Results on optimality

We now move to an intermediate optimality result. One does not explicitly use $x_* \in \Omega_r$ in the proof, but one notes that $g^\circ(x_*; d) \le 0$ only describes the cone of first order linearized directions under the feasibility assumption $x_* \in \Omega_r$.

**Theorem 4.3** *Assume that $f$ is bounded below. Assume that Restoration is entered finitely many times.*

*Let $\{x_k\}_{k \in K}$ be a refined subsequence converging to $x_* \in \Omega$. Assume that $f$ and $g$ are Lipschitz continuous near $x_*$. Assume that $d \in T^H_{\Omega_{nr}}(x_*)$ is a refining direction associated with $K$ and $x_*$ such that $g^\circ(x_*; d) \le 0$. Then $f^\circ(x_*; d) \ge 0$.*

**Proof.** By assumption there exists a subsequence $K_1 \subseteq K$ and a corresponding subsequence $\{d_k\}_{k \in K_1}$ of polling directions such that $\{d_k\}$ converges to $d \in T^H_{\Omega_{nr}}(x_*)$ in $K_1$ and $\alpha_k$ goes to zero in $K_1$. Thus, one must necessarily have that $x_k + \alpha_k d_k \in \Omega_{nr}$ for $k$ sufficiently large in $K_1$.

Since the iteration $k \in K_1$ is unsuccessful, one is sure that $M(x_k + \alpha_k d_k; \mu_k) \ge M(x_k; \mu_k) - \rho(\alpha_k)$, where $\mu_k$ is given by (5).

If $\mu_k = [f(x_k + \alpha_k d_k) - f(x_k)]/[C\rho(\alpha_k)]$, then it is because $[f(x_k + \alpha_k d_k) - f(x_k)]/[C\rho(\alpha_k)] \ge \bar{\mu}$, and thus

$$\frac{f(x_k + \alpha_k d_k) - f(x_k)}{\alpha_k} \;\ge\; C\bar{\mu}\frac{\rho(\alpha_k)}{\alpha_k}. \tag{7}$$

If not, then $M(x_k + \alpha_k d_k; \bar{\mu}) \ge M(x_k; \bar{\mu}) - \rho(\alpha_k)$, and thus

$$\frac{f(x_k + \alpha_k d_k) - f(x_k)}{\alpha_k} \;\ge\; \bar{\mu}\frac{g(x_k) - g(x_k + \alpha_k d_k)}{\alpha_k} - \frac{\rho(\alpha_k)}{\alpha_k}. \tag{8}$$

On the other hand, from the definition of lim sup and the assumption $g^\circ(x_*; d) \le 0$,

$$\limsup_{k \in K_1} \frac{g(x_k + \alpha_k d) - g(x_k)}{\alpha_k} \;\le\; \limsup_{\substack{x \to x_*,\, x \in \Omega_{nr} \\ t \downarrow 0,\, x + td \in \Omega_{nr}}} \frac{g(x + td) - g(x)}{t} \;\le\; 0.$$

Since $g$ is Lipschitz continuous near $x_*$ and the fact that $d_k \to d$ (and using an argument already seen in the proof of Theorem 4.1),

$$\limsup_{k \in K_1} \frac{g(x_k + \alpha_k d_k) - g(x_k)}{\alpha_k} \;=\; \limsup_{k \in K_1} \frac{g(x_k + \alpha_k d) - g(x_k)}{\alpha_k} \;\le\; 0.$$

Thus, one can say that there exists $\{\epsilon_k\}$, with $\epsilon_k \to 0$, such that

$$\frac{g(x_k + \alpha_k d_k) - g(x_k)}{\alpha_k} \;\le\; \epsilon_k \quad \forall k \in K_1,$$

which then implies when (8) occurs

$$\frac{f(x_k + \alpha_k d_k) - f(x_k)}{\alpha_k} \;\ge\; -\bar{\mu}\epsilon_k - \frac{\rho(\alpha_k)}{\alpha_k}. \tag{9}$$

Now we know already that

$$\limsup_{\substack{x \to x_*, x \in \Omega_{nr} \\ t \downarrow 0, x+td \in \Omega_{nr}}} \frac{f(x+td)-f(x)}{t} \;\geq\; \limsup_{k \in K_1} \frac{f(x_k+\alpha_k d)-f(x_k)}{\alpha_k}$$

$$= \limsup_{k \in K_1} \frac{f(x_k+\alpha_k d_k)-f(x_k)}{\alpha_k}.$$

The proof is completed since the right-hand-sides of (7) and (9) tend to zero in $K_1$. □

Finally, we make use of the density of the refining directions in the set

$$T(x_*) \;=\; T^H_{\Omega_{nr}}(x_*) \cap \{d \in \mathbb{R}^n : \|d\| = 1,\; g^\circ(x_*; d) \leq 0\} \tag{10}$$

to derive the complete optimality result.

**Theorem 4.4** *Assume that $f$ is bounded below. Assume that Restoration is entered finitely many times.*

*Let $\{x_k\}_{k\in K}$ be a refined subsequence converging to $x_* \in \Omega$. Assume that $f$ and $g$ are Lipschitz continuous near $x_*$.*

*Assume that $T^H_{\Omega_{nr}}(x_*) \cap \{d \in \mathbb{R}^n : g^\circ(x_*; d) \leq 0\}$ has a non-empty interior.*

*If the set of refining directions associated with $K$ and $x_*$ is dense in $T(x_*)$, then $f^\circ(x_*; v) \geq 0$ for all $v \in T^{Cl}_{\Omega_{nr}}(x_*)$ such that $g^\circ(x_*; v) \leq 0$, and $x_*$ is a stationary point of (1).*

**Proof.** Let $v$ be such that $v \in T^{Cl}_{\Omega_{nr}}(x_*)$, $g^\circ(x_*; v) \leq 0$, and $\|v\| = 1$. Then $v$ is the limit of a sequence $\mathcal{D}$ of refining directions $d$ associated with $K$ and $x_*$ such that $d \in T^H_{\Omega_{nr}}(x_*)$ and $g^\circ(x_*; d) \leq 0$. For each such $d$ one can apply Theorem 4.3 and obtain $f^\circ(x_*; d) \geq 0$. Thus, $f^\circ(x_*; v) = \lim_{d \in T^H_{\Omega_{nr}}(x_*), d \in \mathcal{D}} f^\circ(x_*; d) \geq 0$. The result then holds for non-normalized $v$'s given that $T^{Cl}_{\Omega_{nr}}(x_*)$ is a cone and the Clarke derivatives are homogeneous in their second arguments. □

# 5 Never leaving restoration

The analysis of an infinite run of consecutive steps inside Restoration shows that such a behavior would lead to feasibility and optimality results similar as in the previous case. By a refining subsequence below, we now mean a subsequence of unsuccessful Restoration iterates for which the step-size parameter converges to zero. The definition of refining direction is the same as before. (Again, since $\Omega_{nr}$ is not necessarily by assumption a closed set, one must assume below that $x_*$ belongs to $\Omega_{nr}$.)

**Theorem 5.1** *Assume that $f$ is bounded below. Assume that Restoration is entered and never left.*

*(i) Then there exists a refining subsequence.*

*(ii) Let $\{x_k\}_{k\in K}$ be a refined subsequence converging to $x_* \in \Omega_{nr}$ and assume that $d \in T^H_{\Omega_{nr}}(x_*)$ is a refining direction associated with $K$ and $x_*$. Assume that $g$ is Lipschitz continuous near $x_*$. Then either $g(x_*) = 0$ (implying $x_* \in \Omega_r$ and thus $x_* \in \Omega$) or $g^\circ(x_*; d) \geq 0$.*

*(iii) Let $\{x_k\}_{k\in K}$ be a refined subsequence converging to $x_* \in \Omega$ and assume that $d \in T^H_{\Omega_{nr}}(x_*)$ is a refining direction associated with $K$ and $x_*$ such that $g^\circ(x_*; d) \leq 0$. Assume that $f$ is also Lipschitz continuous near $x_*$. Then $f^\circ(x_*; d) \geq 0$.*

**Proof.** (i) There must exist a refining subsequence $K$ within this call of the Restoration (this is essentially the argument of the third paragraph of the proof of Theorem 3.1). By assumption there exists a subsequence $K_1 \subseteq K$ and a corresponding subsequence $\{d_k\}_{k\in K_1}$ of polling directions such that $\{d_k\}$ converges to $d \in T^H_{\Omega_{nr}}(x_*)$ in $K_1$ and $\alpha_k$ goes to zero in $K_1$. Thus, one must necessarily have that $x_k + \alpha_k d_k \in \Omega_{nr}$ for $k$ sufficiently large in $K_1$.

(ii) Since the iteration $k \in K_1$ is unsuccessful in the Restoration, $g(x_k+\alpha_k d_k) \geq g(x_k)-\rho(\alpha_k)$ or $g(x_k) \leq C\rho(\alpha_k)$, and the proof follows an argument already seen (in the second paragraph of the proof of Theorem 4.1).

(iii) Since at the unsuccessful iteration $k \in K_1$, Restoration is not left, it must be because $M(x_k+\alpha_k d_k;\bar{\mu}) \geq M(x_k;\bar{\mu})$ for all $k \in K_1$, and the proof follows an argument also already seen (see the fourth paragraph of the proof of Theorem 4.3). $\square$

By assuming density of appropriate refining directions in certain cones, we could establish also stationary results for problems (1) and (6) as in Theorem 4.2 and 4.4, respectively

# 6 Entering and leaving restoration an infinite number of times

It remains to analyze the case when one enters (and thus leave) Restoration an infinite number of times. In this case the conditions under which the global convergence results are derived are not the ideal ones since we will have the need to assume that the search step is not performed (or not performed when it requires restoration) and that the step size is not increased (or not increased as frequently as it is decreased).

**Theorem 6.1** *Assume that $f$ is bounded below. Assume that Restoration is entered and left an infinite number of times.*

*Assume that $\alpha_k$ is never increased, that the search step is not applied in the Main algorithm, and that $\{x_k\}$ converges to $x_*$.*

*Let $d$ be a direction which is the limit point of $\{d_k\}$ for both the sequences where Restoration is entered and left.*

*Assume that $f$ and $g$ are Lipschitz continuous near $x_*$. Then $x_* \in \Omega_{nr}$ and either $g(x_*) = 0$ (implying $x_* \in \Omega_r$ and thus $x_* \in \Omega$) or $g^\circ(x_*;d) \geq 0$. Furthermore, $f^\circ(x_*;d) \geq 0$ if $g^\circ(x_*;d) \leq 0$.*

**Proof.** Let $J_1$ and $J_2$ be two subsequences of iterations where Restoration is entered and left respectively.

Since for $k \in J_2$ one knows that $\alpha_k$ is reduced and the step parameter is never increased, one obtains $\alpha_k \to 0$.

Also, by assumption there exists a subsequence $J_3 \subseteq J_2$ and a corresponding subsequence $\{d_k\}_{k\in J_3}$ of polling directions such that $\{d_k\}$ converges to $d \in T^H_{\Omega_{nr}}(x_*)$ in $J_3$ and $\alpha_k$ goes to zero in $J_3$. Thus, one must necessarily have that $x_k + \alpha_k d_k \in \Omega_{nr}$ for $k$ sufficiently large in $J_3$. Thus, from $g(x_k + \alpha_k d_k) \geq g(x_k) - \rho(\alpha_k)$ or $g(x_k) \leq C\rho(\alpha_k)$, for all $k \in J_3$, one concludes that $g^\circ(x_*;d) \geq 0$ or $g(x_*) = 0$.

Now, for $k \in J_1$, $M(x_k + \alpha_k d_k;\bar{\mu}) \geq M(x_k;\bar{\mu})$, and from this we conclude that $f^\circ(x_*;d) \geq 0$ if $g^\circ(x_*;d) \leq 0$. $\square$

To derive a result of the form of Theorem 4.4, one would need to impose that the directions used when entering Restoration are dense in the set (10).

To establish Theorem 6.1 we needed to make sure that $\alpha_k$ goes zero, and since we already had a subsequence of step size decreases, one way to ensure such a property was to rule out step size increases. Note also that we can allow search steps as long as they are skipped when require restoration.

An alternative to this result is to consider a certain maximum number $N$ of Restoration calls, after which one decides to unrelax the relaxable constraints. In this approach, at the $(N+1)$-th call to Restoration, one enters a slightly different Restoration algorithm with the single purpose of minimizing $g$ (i.e., Algorithm 2.2 without the condition of leaving Restoration). After such a call, if one arrives at a point where $g$ is zero, one redefines $\Omega_{nr}$ as the intersection of the originals $\Omega_{nr}$ and $\Omega_r$, and start from there an approach strictly based on the minimization of the extreme barrier function $f_{\Omega_{nr}}$. This procedure can be applied to the relaxable constraints $c_i(x) \leq 0$, $i \in \mathcal{I}$, individually.

# 7 Particularization to smoother settings

When $f$ is strictly differentiable at $x_*$ in the sense of Clarke [10], there exists $\nabla f(x_*)$ such that $f^\circ(x_*; d) = \langle \nabla f(x_*), d \rangle$ for all $d$. Furthermore, if the $c_i$'s are smoother (for instance continuously differentiable at $x_*$), then $g$ in (3) is regular [10], and its Clarke directional derivatives coincide with the traditional ones, i.e., $g^\circ(x^*; d) = g'(x^*; d)$. Thus, under these smoother assumptions, the results would read as follows: (i) $g'(x^*; d) \geq 0$ (in the relaxable constraints criticality result of Theorem 4.1); (ii) the projection of $\nabla f(x_*)$ is zero onto the set of directions $v$ such that $v \in T^{Cl}_{\Omega_{nr}}(x_*)$ and $g'(x^*; v) \leq 0$ (in the optimality result of Theorem 4.4).

When $f$ and $c_i$, $i \in \mathcal{I}$, are continuously differentiable and $\Omega_{nr} = \mathbb{R}^n$, there is no need to use sets of polling directions dense in the unit sphere. The algorithms (Main and Restoration) can then consider in this smooth setting, in their poll steps, directions belonging to positive spanning sets $D_k$. To better extend the result of Theorem 4.1 to such a setting one would have to consider a continuously differentiable version for $g$, such as

$$g(x) = \sum_{i \in \mathcal{I}} [\max(c_i(x), 0)]^2. \tag{11}$$

**Theorem 7.1** *Assume that $f$ is bounded below. Assume that Restoration is entered finitely many times.*

*Let $\{x_k\}_{k \in K}$ be a refined subsequence converging to $x_*$. Suppose that $D_k$ converges in $K$ to a positive spanning set $D_*$. Assume that $\Omega_{nr} = \mathbb{R}^n$, that $c_i$, $i \in \mathcal{I}$, are continuously differentiable at $x_*$, and that $g$ is given by (11). Then either $g(x_*) = 0$ (and thus $x_* \in \Omega$) or $\nabla g(x_*) = 0$.*

**Proof.** Since the iteration $k \in K$ is unsuccessful, $g(x_k + \alpha_k d_k) \geq g(x_k) - \rho(\alpha_k)$ for all $d \in D_k$ or $g(x_k) \leq C\rho(\alpha_k)$, and then either there exists an infinite number of the first or of the second. In the latter case, it is then trivial to obtain $g(x_*) = 0$ from the fact that $\alpha_k \to 0$ in $K$ and the continuity of $g$. In the former case, there exists a subsequence $K_1 \subseteq K$ such that

$$\frac{g(x_k + \alpha_k d) - g(x_k)}{\alpha_k} \;\geq\; -\frac{\rho(\alpha_k)}{\alpha_k} \quad \forall d \in D_k, \; \forall k \in K_1.$$

Applying the mean value theorem, for some $t_k^d \in (0, 1)$,

$$\langle \nabla g(x_k + t_k^d \alpha_k d), d \rangle \;\geq\; -\frac{\rho(\alpha_k)}{\alpha_k} \quad \forall d \in D_k, \; \forall k \in K_1,$$

which then implies $\langle \nabla g(x_*), d\rangle \geq 0$ for all $d \in D_*$, and thus $\nabla g(x_*) = 0$. $\square$

Theorem 4.3 can also be adapted to the continuously differentiable case.

**Theorem 7.2** *Assume that $f$ is bounded below. Assume that Restoration is entered finitely many times.*

*Let $\{x_k\}_{k\in K}$ be a refined subsequence converging to $x_* \in \Omega$. Assume that $\Omega_{nr} = \mathbb{R}^n$ and that $f$, $c_i$, $i \in \mathcal{I}$, are continuously differentiable at $x_*$. Let $g$ be given by (3) or (11). Suppose that $D_k$ converges to a set $D_*$ containing positive generators for*

$$G(x_*) \;=\; \{v \in \mathbb{R}^n : g'(x^*; v) \leq 0\} \;=\; \{v \in \mathbb{R}^n : \langle \nabla c_i(x_*), v\rangle \leq 0 \textit{ when } c_i(x_*) = 0\}. \qquad (12)$$

*Then the projection of $\nabla f(x_*)$ onto $G(x_*)$ is zero.*

**Proof.** The proof of Theorem 4.3 shows that for all limit points $d$ of polling directions, if $d \in G(x_*)$, then $\langle \nabla f(x_*), d\rangle \geq 0$. Thus, for all positive generators of $G(x_*)$ in $D_*$, $\langle \nabla f(x_*), d\rangle \geq 0$, and this implies the result. $\square$

# 8 Numerical illustration

We illustrate the performance of the merit function algorithm on three test problems, which were also tested in [6] to assess the progressive barrier method. The first two problems are defined by a simple algebraic formulation whereas the third one comes from an application.

A simple implementation of Algorithm 2.1 was made in MATLAB without any parameter tuning. The step size updating parameters were set to $\alpha_0 = 1$, $\beta_1 = \beta_2 = 0.5$, and $\gamma = 2$. The forcing function was set chosen as $\rho(\alpha) = \min\{10^{-5}, 10^{-5}\alpha_k^2\}$. For the update of the penalty parameter we picked $\bar{\mu} = \max\{10, g(x_0)\}$ and $C = 100$. No search step was attempted. The measure of constraint violation was the non-smooth one (3). As for the polling directions, those were randomly generated each step with norm one. We show results for $|D_k| = n/2, n+1, 2n$. There is no guarantee, even in the cases $|D_k| = n+1, 2n$, of having computed a positive spanning set, but one knows that that is not required in the convergence theory. A study of random positive spanning sets is out of the scope of this paper. The results presented are the average of 40 runs (corresponding to 40 values of the seed of the MATLAB random generator `randn`).

In the first problem [5], one minimizes a linear function in a convex domain:

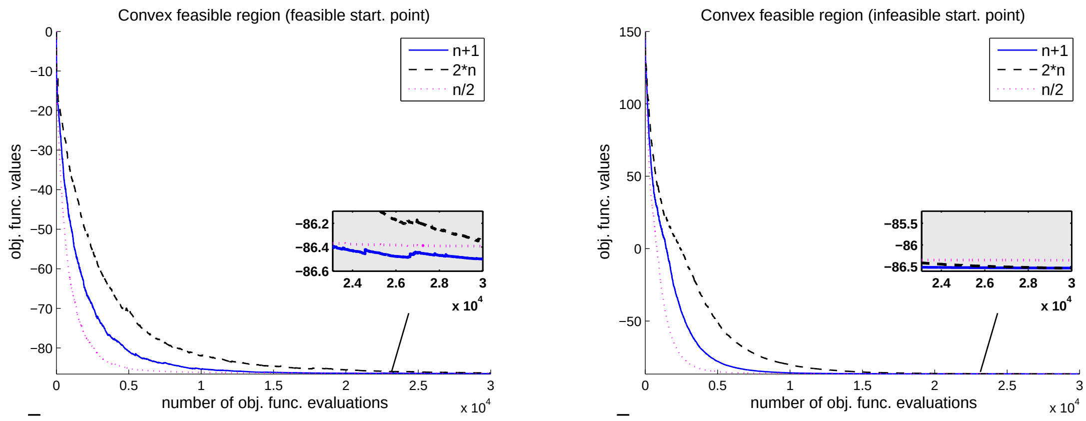
$$\begin{array}{ll} \min & \displaystyle\sum_{i=1}^{n} x_i \\ \text{s.t.} & \displaystyle\sum_{i=1}^{n} x_i^2 \leq 3n. \end{array} \qquad (13)$$

Two starting points are considered, one feasible $(0,\ldots,0)^\top$ and the other infeasible $(3,\ldots,3)^\top$. There is a single (global) solution $(-\sqrt{3},\ldots,-\sqrt{3})^\top$, with optimal value $-\sqrt{3}n$. In the second problem [6], the objective is still linear but the feasible region is non-convex:

$$\begin{array}{ll} \min & x_n \\ \text{s.t.} & \displaystyle\sum_{i=1}^{n} (x_i - 1)^2 \leq n^2 \leq \sum_{i=1}^{n} (x_i + 1)^2. \end{array} \qquad (14)$$

Two starting points are also considered, one feasible $(n, 0, \ldots, 0)^\top$ and the other infeasible $(n, 0, \ldots, 0, -n)^\top$. There is a single (global) solution $(1, \ldots, 1, 1-n)^\top$, with optimal value $1-n$.

The results for problems (13)–(14) are depicted in Figures 1–2 for the case $n = 50$. One can see that convergence is attained in all the cases and that the results must be considered good when compared to those reported in [6]. One observes the non-monotonicity in the value of the objective function (especially when starting infeasible), while reaching feasibility or within the compromise promoted by the merit function. This effect is even visible while approaching the minimizer (which lies at the boundary) for problem (13). One also observes that most of the progress is made within the first 10000 function evaluations (5000 for $|D_k| = n/2$), which is reasonable given the size of the problem and the lack of modeling. The version $|D_k| = n/2$ seems to be the less robust for these test problems. In addition, the number of iterations is much lower (most of the cases below 1000 and never exceeding 2000 for the chosen budget size) meaning that the parallelization of the algorithm would bring significant gains in the overall computational time.



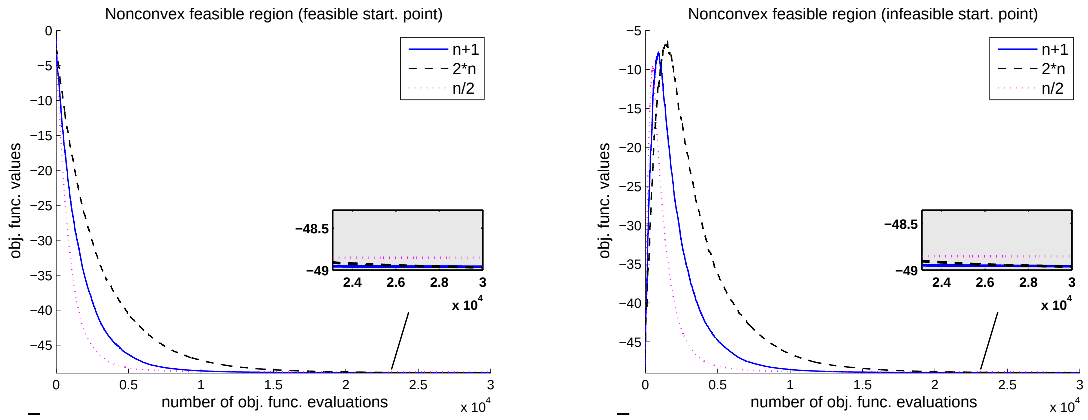

(a) Starting feasible. (b) Starting infeasible.

Figure 1: Two runs of Algorithm 2.1 on problem (13) when $n = 50$ (and a budget of $600n$ is given). The optimal value is approximately 86.6025. On the left (resp. on the right) the starting point is feasible (resp. infeasible).

We also ran the code on the truth model of a problem defined by the optimization of a styrene process production process (see [2]). The problem has 8 variables, 4 unrelaxable constraints (of the type yes-no), and 7 relaxable constraints. The variables have lower and upper bounds ($x_i \in [0, 100]$, $i = 1, \ldots, n$) which were treated by us as unrelaxable constraints. We interfaced the C++ code available in NOMAD [1] for this problem to our MATLAB optimizer. We considered the two initial points also used in [6], namely

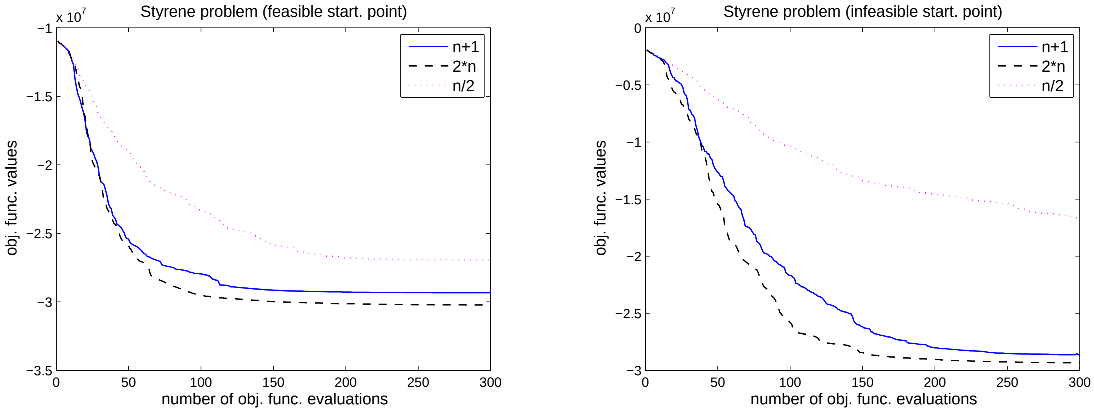
$$\begin{aligned} x_0 &= 100[0.54, 0.66, 0.86, 0.08, 0.29, 0.51, 0.32, 0.15]^\top \quad \text{(feasible for the relaxable const.)} \\ x_0 &= 100[0.44, 0.99, 0.76, 0.39, 0.39, 0.48, 0.43, 0.05]^\top \quad \text{(infeasible for the relaxable const.).} \end{aligned}$$

The plots in Figure 3 depict the performance of the algorithm for these two starting points when using $n/2$, $n+1$, and $2n$ polling directions. Again the version $|D_k| = n/2$ appeared as the less robust one. One can see that the results for this third problem must also be considered good when compared to those reported in [6].




(a) Starting feasible. (b) Starting infeasible.

Figure 2: Two runs of Algorithm 2.1 on problem (14) when $n = 50$ (and a budget of $600n$ is given). The optimal value is $-49$. On the left (resp. on the right) the starting point is feasible (resp. infeasible).




(a) Starting feasible. (b) Starting infeasible.

Figure 3: Two runs of Algorithm 2.1 on the styrene problem for a budget of 300 function evaluations. On the left (resp. on the right) the starting point is feasible (resp. infeasible) with respect to the unrelaxable constraints.

Finally, we point out that, for all the instances run, the returned points were always feasible with respect to the relaxable constraints and that the update of the penalty parameter has never posed any problem of scaling or magnitude. Restoration was only entered a negligible number of times.

# 9 Concluding remarks

We have introduced a globalization procedure to include relaxable constraints in direct-search methods, allowing starting infeasible with respect to these constraints. The procedure introduced requires the evaluation of a merit function for the purposes of measuring success of an iteration.

The penalty parameter present in the merit function does not, thus, play any explicit role in the computation of the step. It is also important to stress that no type of boundedness of the penalty parameter was assumed to derive the global convergence results. We included a scheme to restore feasibility associated with these constraints (or just to significantly reduce such a constraint violation) as it seemed to us as a potentially useful tool.

The convergence analysis is organized depending on the number of times Restoration is entered. When Restoration is entered finitely often, we showed in Theorem 4.2 that the limit points of certain subsequences of iterates (called refining and composed of unsuccessful iterations for which $\alpha_k$ goes to zero) are either feasible or Clarke stationary for the constraint violation problem (6). Then, we showed in Theorem 4.4 that such limit points, when feasible to the original problem, are Clarke stationary for (1). Our theory provides similar results when Restoration is entered but never left (see Section 5). The remaining case is when Restoration is entered and left an infinite number of times (Section 6). Here, to guarantee the same type of results we required the algorithm to meet two additional criteria related to the application of the search step and the update of the step size $\alpha_k$ in successful iterations.

As a referee pointed out to us, our algorithmic framework could be simplified by setting $\mu_k = \bar{\mu}$ without affecting the theoretical properties. Having $\mu_k = \bar{\mu}$ would implicitly maintain the presence of a penalty parameter. For sake of generality and algorithmic flexibility, we maintained the more general penalty parameter update.

A number of issues remain to be better investigated, in particular how our approach would rank in a comprehensive numerical comparison of the existing direct-search type methods for nonlinear constrained derivative-free optimization. The few numerical tests made until now are relatively promising and indicated the need to a better understanding of the use of random directions and random positive spanning sets in direct search, a study which we are currently undertaking. Other algorithmic options are likely to have also a significant impact like the application of a search step, the choice of parameters such as the initial threshold $\bar{\mu}$ for the penalty parameter, and the resetting of the step size before and after a change in optimization state (such as the restoration).

# A Cones and derivatives in the constrained case

A vector is said tangent to $\Omega_{nr}$ at $x$ if it satisfies the following definition.

**Definition A.1** *A vector $d \in \mathbb{R}^n$ is said to be a Clarke tangent vector to the set $\Omega_{nr} \subseteq \mathbb{R}^n$ at the point $x$ in the closure of $\Omega_{nr}$ if for every sequence $\{y_k\}$ of elements of $\Omega_{nr}$ that converges to $x$ and for every sequence of positive real numbers $\{t_k\}$ converging to zero, there exists a sequence of vectors $\{w_k\}$ converging to $d$ such that $y_k + t_k w_k \in \Omega_{nr}$.*

The Clarke tangent cone to $\Omega_{nr}$ at $x$, denoted by $T^{Cl}_{\Omega_{nr}}(x)$, is then defined as the set of all Clarke tangent vectors to $\Omega_{nr}$ at $x$. The Clarke tangent cone generalizes the tangent cone in Nonlinear Programming [23], but one can think about the latter one for gaining the necessary geometric motivation.

Given $x_* \in \Omega_{nr}$ and $d \in T^{Cl}_{\Omega_{nr}}(x)$, one is not sure that $x + td \in \Omega_{nr}$ for $x \in \Omega_{nr}$ arbitrarily close to $x_*$. Thus, for this purpose, one needs to consider directions in the interior of the Clarke tangent cone. The hypertangent cone appears then as the interior of the Clarke tangent cone (when such interior is nonempty, as we assume in this paper). In the sequel, $B(z;r)$ denotes $\{w \in \mathbb{R}^n : \|w - z\| < r\}$.

**Definition A.2** *A vector $d \in \mathbb{R}^n$ is said to be a hypertangent vector to the set $\Omega_{nr} \subseteq \mathbb{R}^n$ at the point $x$ in $\Omega_{nr}$ if there exists a scalar $\epsilon > 0$ such that*

$$y + tw \in \Omega_{nr}, \quad \forall y \in \Omega_{nr} \cap B(x;\epsilon), \quad w \in B(d;\epsilon), \quad \textit{and} \quad 0 < t < \epsilon.$$

The hypertangent cone to $\Omega_{nr}$ at $x$, denoted by $T^H_{\Omega_{nr}}(x)$, is then the set of all hypertangent vectors to $\Omega_{nr}$ at $x$. The closure of the hypertangent cone is the Clarke tangent one (when the former is nonempty).

If we assume that $h$ is Lipschitz continuous near $x_*$, we can define the Clarke-Jahn generalized derivative along directions $d$ in the hypertangent cone to $\Omega_{nr}$ at $x_*$,

$$\begin{aligned} h^\circ(x_*;d) &= \limsup_{\substack{x \to x_*, x \in \Omega_{nr} \\ t \downarrow 0, x+td \in \Omega_{nr}}} \frac{h(x+td)-h(x)}{t} \\ &= \lim_{\epsilon \downarrow 0} \sup_{\substack{x \in B(x_*;\epsilon) \cap \Omega_{nr} \\ t \in (0,\epsilon), x+td \in \Omega_{nr}}} \left\{ \frac{h(x+td)-h(x)}{t} \right\}. \end{aligned}$$

These derivatives are essentially the Clarke generalized directional derivatives [10], generalized by Jahn [16] to the constrained setting. Given a direction $v$ in the tangent cone, one can consider the Clarke-Jahn generalized derivative to $\Omega_{nr}$ at $x_*$ as the limit $h^\circ(x_*;v) = \lim_{d \in T^H_{\Omega_{nr}}(x_*), d \to v} h^\circ(x_*;d)$ (see [5]).

The point $x_*$ is considered stationary for problem (1) when $\Omega = \Omega_{nr}$ if $f^\circ(x_*;v) \geq 0$, $\forall v \in T^{Cl}_{\Omega_{nr}}(x_*)$.

When $\Omega_r \neq \mathbb{R}^n$, then the point $x_*$ is considered stationary for problem (1) if $f^\circ(x_*;v) \geq 0$, $\forall v \in T^{Cl}_{\Omega_{nr}}(x_*) \cap \{d \in \mathbb{R}^n : g^\circ(x_*;d) \leq 0\}$.


## Acknowledgements

The authors thank the two referees, in particular one of them, for their constructive comments and suggestions which led to an improved version of the paper. We are also grateful to A. Ismael F. Vaz for his help in running the application problem.


# References


[1] M.A. Abramson, C. Audet, G. Couture, J.E. Dennis, Jr., S. Le Digabel, and C. Tribes. The NOMAD project. Software available at `http://www.gerad.ca/nomad`.

[2] C. Audet, V. Béchard, and S. Le Digabel. Nonsmooth optimization through mesh adaptive direct search and variable neighborhood search. *J. Global Optim.*, 41:299–318, 2008.

[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. A pattern search filter method for nonlinear programming without derivatives. *SIAM J. Optim.*, 14:980–1010, 2004.

[5] C. Audet and J. E. Dennis Jr. Mesh adaptive direct search algorithms for constrained optimization. *SIAM J. Optim.*, 17:188–217, 2006.

[6] C. Audet and J. E. Dennis Jr. A progressive barrier for derivative-free nonlinear programming. *SIAM J. Optim.*, 20:445–472, 2009.

[7] C. Audet, J. E. Dennis Jr., and S. Le Digabel. Globalization strategies for Mesh Adaptive Direct Search. *Comput. Optim. Appl.*, 46:193–215, 2010.

[8] F. V. Berghen. *CONDOR: A Constrained, Non-Linear, Derivative-Free Parallel Optimizer for Continuous, High Computing Load, Noisy Objective Functions.* PhD thesis, Université Libre de Bruxelles, Belgium, 2004.

[9] R. Brekelmans, L. Driessen, H. Hamers, and D. den Hertog. Constrained optimization involving expensive function evaluations: A sequential approach. *European J. Oper. Res.*, 160:121–138, 2005.

[10] F. H. Clarke. *Optimization and Nonsmooth Analysis.* John Wiley & Sons, New York, 1983. Reissued by SIAM, Philadelphia, 1990.

[11] B. Colson. *Trust-Region Algorithms for Derivative-Free Optimization and Nonlinear Bilevel Programming.* PhD thesis, Département de Mathématique, FUNDP, Namur, Belgium, 2003.

[12] A. R. Conn, K. Scheinberg, and Ph. L. Toint. A derivative free optimization algorithm in practice. In *Proceedings of the 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, St. Louis, Missouri, September 2-4*, 1998.

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

[14] J. E. Dennis Jr., C. J. Price, and I. D. Coope. Direct search methods for nonlinearly constrained optimization using filters and frames. *Optim. Eng.*, 5:123–144, 2004.

[15] M. A. Diniz-Ehrhardt, J. M. Martínez, and L. G. Pedroso. Derivative-free methods for nonlinear programming with general lower-level constraints. *Comput. Appl. Math.*, 30:19–52, 2011.

[16] J. Jahn. *Introduction to the Theory of Nonlinear Optimization.* Springer-Verlag, Berlin, 1996.

[17] T. G. Kolda, R. M. Lewis, and V. Torczon. A generating set direct search augmented Lagrangian algorithm for optimization with a combination of general and linear constraints. Technical Report SAND2006-5315, Sandia National Laboratories, USA, 2006.

[18] R. M. Lewis and V. Torczon. A globally convergent augmented Lagrangian pattern search algorithm for optimization with general constraints and simple bounds. *SIAM J. Optim.*, 12:1075–1089, 2002.

[19] G. Liuzzi and S. Lucidi. A derivative-free algorithm for inequality constrained nonlinear programming via smoothing of an $\ell_\infty$ penalty function. *SIAM J. Optim.*, 20:1–19, 2009.

[20] G. Liuzzi, S. Lucidi, and M. Sciandrone. Sequential penalty derivative-free methods for nonlinear constrained optimization. *SIAM J. Optim.*, 20:2614–2635, 2010.

[21] S. Lucidi, M. Sciandrone, and P. Tseng. Objective-derivative-free methods for constrained optimization. *Math. Program.*, 92:37–59, 2002.

[22] J. M. Martínez and F. N. C. Sobral. Derivative-free constrained optimization on thin domains. *J. Global Optim.*, 56:1217–1232, 2013.

[23] J. Nocedal and S. J. Wright. *Numerical Optimization.* Springer-Verlag, Berlin, second edition, 2006.

[24] M. J. D. Powell. A direct search optimization method that models the objective and constraint functions by linear interpolation. In S. Gomez and J.-P. Hennart, editors, *Advances in Optimization and Numerical Analysis, Proceedings of the Sixth Workshop on Optimization and Numerical Analysis, Oaxaca, Mexico*, volume 275 of *Math. Appl.*, pages 51–67. Kluwer Academic Publishers, Dordrecht, 1994.

[25] L. N. Vicente and A. L. Custódio. Analysis of direct searches for discontinuous functions. *Math. Program.*, 133:299–325, 2012.