SIAM J. OPTIM.
Vol. 20, No. 5, pp. 2614–2635

SEQUENTIAL PENALTY DERIVATIVE-FREE METHODS FOR NONLINEAR CONSTRAINED OPTIMIZATION*

GIAMPAOLO LIUZZI†, STEFANO LUCIDI‡, AND MARCO SCIANDRONE§

Abstract. We consider the problem of minimizing a continuously differentiable function of several variables subject to smooth nonlinear constraints. We assume that the first order derivatives of the objective function and of the constraints can be neither calculated nor explicitly approximated. Hence, every minimization procedure must use only a suitable sampling of the problem functions. These problems arise in many industrial and scientific applications, and this motivates the increasing interest in studying derivative-free methods for their solution. The aim of the paper is to extend to a derivative-free context a sequential penalty approach for nonlinear programming. This approach consists in solving the original problem by a sequence of approximate minimizations of a merit function where penalization of constraint violation is progressively increased. In particular, under some standard assumptions, we introduce a general theoretical result regarding the connections between the sampling technique and the updating of the penalization which are able to guarantee convergence to stationary points of the constrained problem. On the basis of the general theoretical result, we propose a new method and prove its convergence to stationary points of the constrained problem. The computational behavior of the method has been evaluated both on a set of test problems and on a real application. The obtained results and the comparison with other well-known derivative-free software show the viability of the proposed sequential penalty approach.

Key words. derivative-free optimization, nonlinear programming, sequential penalty functions

AMS subject classifications. 65K05, 90C30, 90C56

DOI. 10.1137/090750639

1. Introduction.

In this paper we consider the nonlinear constrained minimization problem

$$\begin{aligned} &\min f(x),\\ &g(x) \le 0,\\ &l \le x \le u, \end{aligned} \tag{1}$$

where $f : R^n \to R$, $g : R^n \to R^m$, and $l, u \in \Re^n$, with $l < u$, are vectors of lower and upper bounds on the variables $x \in \Re^n$. We denote by X the set defined by simple bounds on the variables, that is,

$$X = \{x \in \Re^n : l \le x \le u\},$$

and by $\mathcal{F}$ the feasible set of problem (1), namely,

$$\mathcal{F} = \{x \in \Re^n : g(x) \le 0\} \cap X.$$

We note that, by definition, X is a compact set. Furthermore, we assume that f and g are continuously differentiable functions even though their derivatives can be neither

*Received by the editors February 24, 2009; accepted for publication (in revised form) April 27, 2010; published electronically July 13, 2010. This work has been supported by MIUR-PRIN 20079PLLN7 Research Program on Nonlinear Optimization, Variational Inequalities, and Equilibrium Problems.

http://www.siam.org/journals/siopt/20-5/75063.html

†Consiglio Nazionale delle Ricerche, Istituto di Analisi dei Sistemi ed Informatica "A. Ruberti," Viale Manzoni 30, 00185 Rome, Italy (liuzzi@iasi.cnr.it).

‡"Sapienza" Università di Roma, Dipartimento di Informatica e Sistemistica "A. Ruberti," Via Ariosto 25, 00185 Rome, Italy (lucidi@dis.uniroma1.it).

§Università di Firenze, Dipartimento di Sistemi e Informatica, Via di Santa Marta 3, 50139 Firenze, Italy (sciandro@dsi.unifi.it).

calculated nor explicitly approximated. In many engineering problems, the values of the functions defining the objective and constraints of the problem are computed by means of complex simulation programs. For this reason, their analytic expressions are not available. We refer the reader to the survey paper [14] for a detailed discussion on this issue.

In the literature, many approaches for derivative-based nonlinear programming have been extended to a derivative-free context. In [16, 15] a pattern search algorithm is used within a sequential augmented Lagrangian approach. In [4] the filter method proposed in [11] is adapted to include a pattern search minimization strategy. In [5] a so-called *extreme barrier* approach is employed, whereas in [6] the use of a progressive barrier approach to the problem is proposed. Finally, in [18] a derivative-free line search technique is used to minimize a smoothed ℓ_∞ exact penalty function. It is worth noting that the methods proposed in [4, 5, 6] do not assume that the functions f and g are differentiable, whereas in [16, 15, 18] the functions f and g are assumed to have continuous second derivatives.

In order to increase the tools available for the solution of constrained problems when derivatives are not available, we extend the sequential penalty approach to a derivative-free context.

When derivatives are available, the sequential penalty approach consists in solving the original problem by a sequence of approximate minimizations of a merit function where the objective function is augmented by a term that penalizes constraint violation. Every minimization is carried out with a given degree of approximation which is increased more and more during the optimization process. After any such approximate minimization the penalization is increased by a simple updating rule. By exploiting information on the derivatives, it is possible to tie the precision level to the penalty parameter updating in such a way that convergence to stationary points of problem (1) can be guaranteed [10, 8]. In a derivative-free context, a possible way to overcome the lack of derivative information can be to use some suitable sampling of the problem functions. To the best of our knowledge, derivative-free methods embedded in a sequential penalty framework have never been proposed. In fact, to enforce convergence to stationary points, this requires a suitable combination of the penalty parameter updating and the sampling technique.

In this paper, under some standard assumptions in a constrained context, we introduce a general theoretical result regarding the connections between the mentioned sampling technique and the updating of the penalization which are able to guarantee convergence to stationary points of the constrained problem of a sequential penalty-based model algorithm. On the basis of this general result, it is possible to define new derivative-free methods using different sampling strategies and prove their global convergence. The interested reader can find in [19] the definition of different methods and their convergence analysis based on the general theoretical result of this paper, Proposition 4. In this paper we focus on a line search–based algorithm which appears to be the most promising method.

The paper is organized as follows. In section 2 we introduce the basic assumptions required to prove the general convergence result, and we report some preliminary results. In section 3 we describe in more detail the sequential penalty approach proposed, and we prove the main convergence result of the paper. In section 4 we introduce the new derivative-free method which belongs to the class of derivative-free line search algorithms [20] and prove its convergence to KKT points of problem (1). In section 5 we show the results of numerical experiments performed both on a set of test problems and on a real application problem. We also present the comparison with two well-known derivative-free algorithms. Finally, in section 6 we report some conclusions. The paper also includes two appendices. Appendix A is concerned with the proof of two technical propositions whose results are needed to prove the main convergence theorem of this paper. In Appendix B we report the complete numerical results.

2. Notation and preliminary results. In this section we introduce some useful notation and assumptions that will be used throughout the paper.

Given a vector $v \in \Re^n$, a subscript will be used to denote either one of its components (v_i) or the fact that it is an element of an infinite sequence of vectors (v_k). To avoid possible misunderstanding or ambiguities, the ith component of a vector will be denoted by $(v)_i$. We denote by v^j the generic jth element of a finite set of vectors. Given two vectors $a, b \in \Re^n$, we denote by $y = \max\{a, b\}$ the vector such that $y_i = \max\{a_i, b_i\}$, $i = 1, \dots, n$. Furthermore, given a vector v, we denote $v^+ = \max\{0, v\}$.

Definition 1 (cone of feasible directions). *Given a point $x \in X$, let*

$$D(x) = \{d \in \Re^n : d_i \geq 0 \text{ if } x_i = l_i,\ d_i \leq 0 \text{ if } x_i = u_i,\ i = 1, \dots, n\}$$

be the cone of feasible directions at x with respect to the simple bound constraints.

Let $L(x, \lambda)$ be the Lagrangian function associated with the nonlinear constraints of problem (1),

$$L(x, \lambda) = f(x) + \lambda^T g(x).$$

We recall the Mangasarian–Fromovitz constraint qualification (MFCQ).

Definition 2. *A point $x \in X$ is said to satisfy the MFCQ if there exists a vector $\hat{d} \in D(x)$ such that*

$$\nabla g_l(x)^T \hat{d} < 0 \quad \forall l \in I^+(x),$$

where $I^+(x) = \{i : g_i(x) \geq 0\}$.

The following proposition is a well-known result (see, for instance, [8]) which states necessary optimality conditions for problem (1).

Proposition 1. *Let $x^\star \in \mathcal{F}$ be a local minimum of problem (1). Then, there exists a vector $\lambda^\star \in R^m$ such that*

$$\nabla_x L(x^\star, \lambda^\star)^T (x - x^\star) \geq 0 \qquad \forall x \in X, \tag{2}$$

$$(\lambda^\star)^T g(x^\star) = 0, \qquad \lambda^\star \geq 0. \qquad \square \tag{3}$$

Definition 3 (stationary point). *A point $x^\star \in \mathcal{F}$ is said to be a stationary point for problem (1) if a vector $\lambda^\star \in R^m$ exists such that (2) and (3) are satisfied.*

We recall a result from [17] concerning the set $D(x)$.

Proposition 2. *Let $\{x_k\}$ be a sequence of points such that $x_k \in X$ for all k. Assume further that $x_k \to \bar{x}$ for $k \to \infty$. Then, given any direction $\bar{d} \in D(\bar{x})$, there exists a scalar $\bar{\beta} > 0$ such that, for sufficiently large k, we have*

$$x_k + \beta \bar{d} \in X \quad \forall \beta \in [0, \bar{\beta}].$$

As an immediate consequence, we have the following corollary.

COROLLARY 1. *Let $\{x_k\}$ be a sequence of points such that $x_k \in X$ for all k, and $x_k \to \bar{x}$ for $k \to \infty$. Then*

$$D(\bar{x}) \subseteq D(x_k)$$

for k sufficiently large.

Now we define the set of unit vectors

$$D = \{\pm e^1, \dots, \pm e^n\},$$

where e^i, $i = 1, \dots, n$, is the ith unit coordinate vector. In particular, in the following proposition, we show that set D contains the generators of the cone of feasible directions $D(x)$ at any point $x \in X$.

PROPOSITION 3. *Let $x \in X$. We have*

$$cone\{D \cap D(x)\} = D(x). \tag{4}$$

Proof. Given $x \in X$, let us consider $d \in D(x)$. We can write

$$d = \sum_{i \in I} d_i e^i - \sum_{j \in J} |d_j| e^j, \tag{5}$$

where $I = \{i \in \{i, \dots, n\} : d_i > 0\}$, $J = \{i \in \{1, \dots, n\} : d_i < 0\}$. Since $d \in D(x)$, we have

$$x_i < u_i \quad \forall i \in I,$$

$$x_j > l_j \quad \forall j \in J,$$

so that

$$e^i \in D(x) \quad \forall i \in I \qquad \text{and} \qquad -e^j \in D(x) \quad \forall j \in J. \tag{6}$$

Then, the thesis follows from (5) and (6). ☐

3. Penalty function and convergence conditions. In order to solve problem (1), it is possible to augment the objective function by adding terms that are able to penalize constraints violation; namely,

$$Q(x;\epsilon) = f(x) + \frac{1}{\epsilon}\left(\sum_{j=1}^{m}\left[g_j(x)^+\right]^2 + \sum_{i=1}^{n}\left[(x_i - u_i)^+\right]^2 + \sum_{i=1}^{n}\left[(l_i - x_i)^+\right]^2\right).$$

In [10], function $Q(x;\epsilon)$ has been used to define an algorithm for the solution of problem (1). In particular, in [10] it has been proved that if one is able to find a global minimizer x_k^* of the penalty function in correspondence with each penalty parameter of a sequence $\{\epsilon_k\}$ such that $\epsilon_k \to 0$, then the sequence $\{x_k^*\}$ converges to a global minimizer x^* of the original constrained problem.

In more practical terms, in [8, 21] convergence to a stationary point of problem (1) has been proved under suitable regularity assumptions provided that one is able to find an approximate stationary point of $Q(x;\epsilon_k)$, for every k, with higher and higher precision. More precisely, if $\{x_k\}$ is a sequence of points satisfying

$$\|\nabla Q(x_k;\epsilon_k)\| \le \tau_k,$$

where $\{\tau_k\}$ is a sequence of scalars such that $0 < \tau_{k+1} < \tau_k$ for all k, and $\tau_k \to 0$, then, provided that $\{x_k\}$ (or, at least, a subsequence) converges to a point $\tilde{x}$ where the gradients of the active constraints are linearly independent, $\tilde{x}$ is stationary for problem (1).

In this paper we extend the preceding approach to the case where we cannot use any derivative information on the objective and nonlinear constraint functions defining problem (1). Bound constraints on the variables are handled explicitly since their gradients and structure are perfectly known. Hence, we introduce the sequential penalty function [10]

$$P(x;\epsilon) = f(x) + \frac{1}{\epsilon}\sum_{j=1}^{m}\left[g_j(x)^+\right]^q,$$

where $q > 1$ and only the nonlinear constraints have been taken into account, and consider the problem [7, 8]

$$\min_{l \le x \le u} P(x;\epsilon). \tag{7}$$

For every fixed value of the penalty parameter ϵ, function $P(x;\epsilon)$ is continuously differentiable under the stated assumptions.

Derivative-free methods are based on a suitable sampling technique along a set of directions that are able to convey, in the limit, sufficient knowledge of the problem functions to recover first order information. However, in a constrained context, in which the penalty parameter has to be updated and progressively driven to zero, the updating rule must be connected with the sampling technique. Roughly speaking, the penalty parameter must converge to zero more slowly than the maximum stepsize used by the sampling scheme.

The proposition that follows states a general result that can be used to prove convergence toward stationary points of the sequence of iterates produced by a derivative-free algorithm used to approximately minimize the penalty function $P(x;\epsilon)$ on the set X. Namely, the proposition gives sufficient conditions on the sampling technique performed by the derivative-free algorithm and on the updating of the penalty parameter that are able to guarantee convergence toward a stationary point of problem (1).

PROPOSITION 4. *Let $\{\epsilon_k\}$ be a bounded sequence of positive penalty parameters. Let $\{x_k\}$ be a sequence of points such that $x_k \in X$ for all k, and let $\bar{x}$ be a limit point of a subsequence $\{x_k\}_K$ for some infinite set $K \subseteq \{0,1,\ldots\}$. Suppose that $\bar{x}$ satisfies the MFCQ and that for each $k \in K$ sufficiently large*

(i) *for all $d^i \in D \cap D(\bar{x})$ there exist vectors y_k^i and scalars $\xi_k^i > 0$ such that*

$$y_k^i + \xi_k^i d^i \in X, \tag{8}$$

$$P(y_k^i + \xi_k^i d^i;\epsilon_k) \ge P(y_k^i;\epsilon_k) - o(\xi_k^i), \tag{9}$$

$$\lim_{k\to\infty, k\in K} \frac{\max\{\xi_k^i, \|x_k - y_k^i\|\}}{\epsilon_k} = 0; \tag{10}$$

(ii) *and*

$$\lim_{k\to\infty, k\in K} \epsilon_k \|g^+(x_k)\| = 0. \tag{11}$$

Then $\bar{x}$ is a stationary point for problem (1).

Proof. The proof can be divided into the following three main parts.

(a) In the first part, the mean-value theorem is applied to condition (9), and some relations are derived which will be used further in the proof.

(b) The second part proves that the limit point $\bar{x}$ is feasible for problem (1).

(c) In the last part of the proof, we introduce multiplier functions

$$\lambda_l(x;\epsilon) = \frac{q}{\epsilon}\max\{g_l(x),0\}^{q-1}, \qquad l = 1,\dots,m,$$

and we show that $\nabla L(\bar{x},\bar{\lambda})^T d^i \geq 0$ for all $d^i \in D \cap D(\bar{x})$, so that the thesis follows from Proposition 3. In this part of the proof we exploit a technical result (Proposition 6 in Appendix A) concerning the boundedness of the sequences $\{\lambda_l(x_k;\epsilon_k)\}$.

Part (a). Let us denote $\bar{D} = D \cap D(\bar{x})$. By applying the mean-value theorem to (9), we can write

$$-o(\xi_k^i) \leq P(y_k^i + \xi_k^i d^i;\epsilon_k) - P(y_k^i;\epsilon_k) = \xi_k^i \nabla P(u_k^i;\epsilon_k)^T d^i \qquad \forall d^i \in \bar{D},$$

where $u_k^i = y_k^i + t_k^i \xi_k^i d^i$, with $t_k^i \in (0,1)$. Thus, we have

$$-\frac{o(\xi_k^i)}{\xi_k^i} \leq \nabla P(u_k^i;\epsilon_k)^T d^i \qquad \forall d^i \in \bar{D}.$$

By considering the expression of $P(x;\epsilon)$, we can write

$$\begin{aligned}(12)\quad & \nabla P(u_k^i;\epsilon_k)^T d^i \\ & = \left(\nabla f(u_k^i) + \frac{q}{\epsilon_k}\sum_{l=1}^m \max\{g_l(u_k^i),0\}^{q-1}\nabla g_l(u_k^i)\right)^T d^i \geq -\frac{o(\xi_k^i)}{\xi_k^i} \qquad \forall d^i \in \bar{D}.\end{aligned}$$

Recalling that $u_k^i = y_k^i + t_k^i \xi_k^i d^i$, with $t_k^i \in (0,1)$, we have that, for all i such that $d^i \in \bar{D}$,

$$\lim_{k\to\infty, k\in K} u_k^i = \bar{x}. \tag{13}$$

Part (b). We prove that $g(\bar{x}) \leq 0$. We consider the sequence of positive penalty parameters $\{\epsilon_k\}$. If this sequence is bounded away from zero, then limit (11) implies that

$$\lim_{k\to\infty, k\in K} \|g^+(x_k)\| = \|g^+(\bar{x})\| = 0.$$

If, on the contrary, we have that

$$\lim_{k\to\infty, k\in K} \epsilon_k = 0,$$

recalling assumption (i), multiplying relation (12) by ϵ_k, and taking the limit, we obtain

$$\left(q\sum_{l=1}^m \max\{g_l(\bar{x}),0\}^{q-1}\nabla g_l(\bar{x})\right)^T d^i \geq 0 \qquad \forall d^i \in \bar{D}. \tag{14}$$

Recall that, by assumption, $\bar{x}$ satisfies the MFCQ (see Definition 2), and let $\hat{d} \in D(\bar{x})$ be the direction considered in Definition 2; then we can write

$$\sum_{l=1}^{m} \max\{g_l(\bar{x}), 0\}^{q-1} \nabla g_l(\bar{x})^T \hat{d} \le 0. \tag{15}$$

From Proposition 3 we have

$$\hat{d} = \sum_{i:d^i \in \bar{D}} \hat{\beta}_i d^i, \tag{16}$$

where $\hat{\beta}_i \ge 0$, so that using (14) and (16), we obtain

$$\begin{aligned}\left(q \sum_{l=1}^{m} \max\{g_l(\bar{x}), 0\}^{q-1} \nabla g_l(\bar{x})^T\right) \hat{d} \\ = \sum_{i:d^i \in \bar{D}} \hat{\beta}_i \left(q \sum_{l=1}^{m} \max\{g_l(\bar{x}), 0\}^{q-1} \nabla g_l(\bar{x})\right)^T d^i \ge 0.\end{aligned} \tag{17}$$

From (15) and (17) it follows that

$$0 \le \sum_{l=1}^{m} \max\{g_l(\bar{x}), 0\}^{q-1} \left(\nabla g_l(\bar{x})^T \hat{d}\right) = \sum_{l \in I^+(\bar{x})} \max\{g_l(\bar{x}), 0\}^{q-1} \left(\nabla g_l(\bar{x})^T \hat{d}\right) \le 0.$$

Therefore, since $\bar{x}$ satisfies the MFCQ, we obtain $g(\bar{x}) \le 0$.

Part (c). For $l = 1, \dots, m$ set

$$\lambda_l(x; \epsilon) = \frac{q}{\epsilon} \max\{g_l(x), 0\}^{q-1}.$$

By Proposition 6, there exists a subset of K, which we relabel again K, such that

$$\lim_{k \to \infty, k \in K} \lambda_l(x_k; \epsilon_k) = \bar{\lambda}_l \ge 0, \qquad l = 1, \dots, m,$$

where $\bar{\lambda}_l = 0$ for $l \notin I^+(\bar{x})$. By simple manipulations, (12) can be rewritten as

$$\begin{aligned}\Bigg(\nabla f(u_k^i) + \sum_{l=1}^{m} \nabla g_l(u_k^i) \lambda_l(x_k; \epsilon_k) \\ + \sum_{l=1}^{m} \nabla g_l(u_k^i) \left(\lambda_l(u_k^i; \epsilon_k) - \lambda_l(x_k; \epsilon_k)\right)\Bigg)^T d^i \ge -\frac{o(\xi_k^i)}{\xi_k^i} \qquad \forall i : d^i \in \bar{D}.\end{aligned} \tag{18}$$

Taking the limits for $k \to \infty$ and $k \in K$ in relation (18) and recalling (57) from the proof of Proposition 6 previously invoked, we obtain

$$\left(\nabla f(\bar{x}) + \sum_{l=1}^{m} \nabla g_l(\bar{x}) \bar{\lambda}_l\right)^T d^i \ge 0 \qquad \forall i : d^i \in \bar{D}.$$

Recalling that $\bar{D} = D \cap D(\bar{x})$, from Proposition 3 we get

$$\nabla L(\bar{x}, \bar{\lambda})^T d \ge 0 \qquad \forall d \in D(\bar{x}),$$

which concludes the proof. $\square$

Following [14], it can be shown that

$$\max_{i:d^i\in D\cap D(\bar{x})}\{\xi_k^i, \|x_k - y_k^i\|\}$$

bounds a measure of stationarity of the current iterate x_k for problem (7). Hence, limit (10) amounts to requiring that the current measure of stationarity goes to zero faster than the penalty parameter ϵ_k.

4. A derivative-free method for problems with bound constraints. This section is devoted to the introduction and analysis of a derivative-free method for the solution of problem (1). More precisely, we propose a derivative-free line search–type [20] algorithm.

As we will show, the theoretical convergence analysis can be derived from the general result of Proposition 4.

The proposed algorithm uses a line search technique which, roughly speaking, performs an approximate minimization of the penalty function along the promising search directions. In this way it is possible to probe and exploit the sensitivity of the objective function along the considered direction. To this aim we compute different stepsizes on each search direction. In particular, at every iteration we compute the following quantities:

- $\bar{\alpha}^i$, $i = 1, \dots, n$, which represent the maximum steplengths that can be taken along the directions without leaving set X;
- α_k^i, $i = 1, \dots, n$, which are the results of the approximate minimizations of the penalty function along the directions d_k^i;
- $\tilde{\alpha}_k^i$, $i = 1, \dots, n$, which record the result of the line searches at the preceding iteration and are used as initial stepsizes for the line searches at the current iteration.

The following proposition shows the well-definedness of Algorithm DFL and gives some preliminary properties of the produced sequences.

PROPOSITION 5. *Let* $\{x_k\}$, $\{\epsilon_k\}$, $\{\tilde{\alpha}_k^i\}$, *and* $\{\alpha_k^i\}$, $i = 1, \dots, n$, *be the sequences produced by Algorithm* DFL. *Then the following hold:*

(i) *Algorithm* DFL *is well defined.*

(ii) *If the monotonically nonincreasing sequence of positive numbers* $\{\epsilon_k\}$ *is such that*

$$\lim_{k\to\infty} \epsilon_k = \bar{\epsilon} > 0, \tag{19}$$

then

$$\lim_{k\to\infty} \alpha_k^i = 0 \qquad \textit{for } i = 1, \dots, n, \tag{20}$$

$$\lim_{k\to\infty} \tilde{\alpha}_k^i = 0 \qquad \textit{for } i = 1, \dots, n. \tag{21}$$

(iii) *If the monotonically nonincreasing sequence of positive numbers* $\{\epsilon_k\}$ *is such that*

$$\lim_{k\to\infty} \epsilon_k = 0, \tag{22}$$

then

$$\lim_{k\to\infty, k\in K} \alpha_k^i = 0 \qquad \textit{for } i = 1, \dots, n, \tag{23}$$

Algorithm DFL.

Data. $x_0 \in X$, $\epsilon_0 > 0$, $\gamma > 0$, $\theta \in (0,1)$, $p > 1$, $\tilde{\alpha}_0^i > 0$, a sequence of positive numbers $\eta_k \to 0$, and set $d_0^i = e^i$ for $i = 1, \dots, n$.

Step 1. *(Minimization on the cone$\{D\}$)*

Step 1.1. Set $i = 1$, $y_k^i = x_k$.

Step 1.2. Compute $\bar{\alpha}^i$ such that $y_k^i + \bar{\alpha}^i d_k^i \in X$, and set $\hat{\alpha}_k^i = \min\{\bar{\alpha}^i, \tilde{\alpha}_k^i\}$:

If $\hat{\alpha}_k^i > 0$ and $P(y_k^i + \hat{\alpha}_k^i d_k^i; \epsilon_k) \le P(y_k^i; \epsilon_k) - \gamma (\hat{\alpha}_k^i)^2$, compute α_k^i by the *Expansion Step*($\bar{\alpha}^i, \hat{\alpha}_k^i, y_k^i, d_k^i, \gamma; \alpha_k^i$); set $\tilde{\alpha}_{k+1}^i = \alpha_k^i$, $d_{k+1}^i = d_k^i$ and go to Step 1.5.

Step 1.3. Compute $\bar{\alpha}$ such that $y_k^i - \bar{\alpha}^i d_k^i \in X$, and set $\hat{\alpha}_k^i = \min\{\bar{\alpha}^i, \tilde{\alpha}_k^i\}$:

If $\hat{\alpha}_k^i > 0$ and $P(y_k^i - \hat{\alpha}_k^i d_k^i; \epsilon_k) \le P(y_k^i; \epsilon_k) - \gamma (\hat{\alpha}_k^i)^2$, compute α_k^i by the *Expansion Step*($\bar{\alpha}^i, \hat{\alpha}_k^i, y_k^i, -d_k^i, \gamma; \alpha_k^i$); set $\tilde{\alpha}_{k+1}^i = \alpha_k^i$, $d_{k+1}^i = -d_k^i$, and go to Step 1.5.

Step 1.4. Set $\alpha_k^i = 0$ and $\tilde{\alpha}_{k+1}^i = \theta \tilde{\alpha}_k^i$.

Step 1.5. Set $y_k^{i+1} = y_k^i + \alpha_k^i d_k^i$. **If** $i < n$, set $i = i + 1$ and go to Step 1.2.

Step 2. **If** $\max_{i=1,\dots,n}\{\tilde{\alpha}_k^i, \alpha_k^i\} \le \epsilon_k^p$ and $(\|g^+(x_k)\| > \eta_k)$, choose $\epsilon_{k+1} = \theta \epsilon_k$

Else set $\epsilon_{k+1} = \epsilon_k$.

Step 3. Find $x_{k+1} \in X$ such that $P(x_{k+1}; \epsilon_k) \le P(y_k^{i+1}; \epsilon_k)$. Set $k = k + 1$ and go to Step 1.

Expansion Step ($\bar{\alpha}, \hat{\alpha}, y, p, \gamma; \alpha$).

Data. $\delta \in (0,1)$.

Step 1. Set $\alpha = \hat{\alpha}$.

Step 2. Let $\tilde{\alpha} = \min\{\bar{\alpha}, (\alpha/\delta)\}$.

Step 3. **If** $\alpha = \bar{\alpha}$ or $P(y + \tilde{\alpha} p; \epsilon_k) > P(y; \epsilon_k) - \gamma \tilde{\alpha}^2$, return.

Step 4. Set $\alpha = \tilde{\alpha}$ and go to Step 2.

$$\lim_{k \to \infty, k \in K} \tilde{\alpha}_k^i = 0 \qquad \textit{for } i = 1, \dots, n, \tag{24}$$

where $K = \{k : \epsilon_{k+1} < \epsilon_k\}$.

Proof. In order to prove that Algorithm DFL is well defined, we have to ensure that the Expansion Step, when performed along a direction d_k^i, with $i \in \{1, \dots, n\}$, terminates in a finite number j of steps. This is clearly true since, by the instructions of the Expansion Step,

$$x_k + \delta^{-j} \alpha d_k^i \in X \qquad \forall j,$$

and X is a compact set.

Now we prove assertion (ii). By (19), a $\bar{k} \ge 0$ exists such that

$$\epsilon_{k+1} = \epsilon_k = \epsilon_{\bar{k}} = \bar{\epsilon} \qquad \forall k \ge \bar{k}.$$

For every $i = 1, \dots, n$ we prove (20) by splitting the iteration sequence $\{k\}$ into two parts, K' and K''. We identify with K' those iterations where

$$\alpha_k^i = 0 \tag{25}$$

and with K'' those iterations where $\alpha_k^i \neq 0$ is produced by the Expansion Step. Then the instructions of the algorithm imply

$$P(x_{k+1};\bar\epsilon) \le P(y_k^i + \alpha_k^i d_k^i;\bar\epsilon) \le P(y_k^i;\bar\epsilon) - \gamma(\alpha_k^i)^2\|d_k^i\|^2 \le P(x_k;\bar\epsilon) - \gamma(\alpha_k^i)^2\|d_k^i\|^2. \tag{26}$$

Taking into account the compactness assumption on X, it follows from (26) that $\{P(x_k;\bar\epsilon)\}$ tends to a limit $\bar P$. If K'' is an infinite subset, recalling that $\|d_k^i\| = 1$, we obtain

$$\lim_{k\to\infty, k\in K''} \alpha_k^i = 0. \tag{27}$$

Therefore, (25) and (27) imply (20).

In order to prove (21), for each $i \in \{1,\dots,n\}$ we split the iteration sequence $\{k\}$ into two parts, K_1 and K_2. We identify with K_1 those iterations where the Expansion Step has been performed using the direction d_k^i, for which we have

$$\tilde\alpha_{k+1}^i = \alpha_k^i. \tag{28}$$

We denote by K_2 those iterations where we have failed in decreasing the objective function along the directions d_k^i and $-d_k^i$. By the instructions of the algorithm it follows that for all $k \in K_2$

$$\tilde\alpha_{k+1}^i \le \theta\tilde\alpha_k^i, \tag{29}$$

where $\theta \in (0,1)$.

If K_1 is an infinite subset, from (28) and (20) we get that

$$\lim_{k\to\infty, k\in K_1} \tilde\alpha_{k+1}^i = 0. \tag{30}$$

Now, let us assume that K_2 is an infinite subset. For each $k \in K_2$, let m_k (we omit the dependence on i) be the biggest index such that $m_k < k$ and $m_k \in K_1$. Then we have

$$\tilde\alpha_{k+1}^i \le \theta^{(k+1-m_k)}\tilde\alpha_{m_k}^i \le \tilde\alpha_{m_k}^i \tag{31}$$

(we can assume $m_k = 0$ if the index m_k does not exist, that is, K_1 is empty).

As $k \to \infty$ and $k \in K_2$, either $m_k \to \infty$ (namely, K_1 is an infinite subset) or $(k+1-m_k) \to \infty$ (namely, K_1 is finite). Hence, if K_2 is an infinite subset, (31) together with (30), or the fact that $\theta \in (0,1)$, yields

$$\lim_{k\to\infty, k\in K_2} \tilde\alpha_{k+1}^i = 0, \tag{32}$$

so that (21) is proved, and this concludes the proof of point (ii).

Point (iii). If (22) holds, there must exist an infinite subset $K \subseteq \{0,1,\dots\}$ such that $\epsilon_{k+1} = \theta\epsilon_k < \epsilon_k$ for all $k \in K$.

Since, from the instructions at Step 2 of the algorithm, the penalty parameter is only updated when the test

$$\max_{i=1,\dots,n} \{\tilde{\alpha}_k^i, \alpha_k^i\} \le \epsilon_k^p \tag{33}$$

is satisfied, the proof of point (iii) follows from (33) and (22). ∎

Now we prove the main convergence result concerning Algorithm DFL.

THEOREM 1. *Let* $\{x_k\}$ *be the sequence generated by Algorithm* DFL. *Assume that every limit point of the sequence* $\{x_k\}$ *satisfies the MFCQ; then there exists a limit point* $\bar{x}$ *of the sequence* $\{x_k\}$ *which is a stationary point of problem* (1).

Proof. We recall that, by the instructions of Algorithm DFL, at every iteration k, the following set of directions is considered:

$$D_k = \{d_k^1, -d_k^1, \dots, d_k^n, -d_k^n\} = \{\pm e^1, \dots, \pm e^n\} = D.$$

At every iteration k, Algorithm DFL extracts information on the behavior of the penalty function along both d_k^i and $-d_k^i$.

In particular, along all d_k^i, $i = 1, \dots, n$, the algorithm identifies the following circumstances:

If the initial stepsize $\tilde{\alpha}_k^i$ fails to produce a decrease of the penalty function, we have either

$$y_k^i + \tilde{\alpha}_k^i d_k^i \notin X \tag{34}$$

or

$$P(y_k^i + \tilde{\alpha}_k^i d_k^i; \epsilon_k) > P(y_k^i; \epsilon_k) - \gamma(\tilde{\alpha}_k^i)^2. \tag{35}$$

If, instead, the initial stepsize $\tilde{\alpha}_k^i$ produces a decrease of the penalty function, we have both

$$y_k^i + \tilde{\alpha}_k^i d_k^i \in X \tag{36}$$

and

$$P(y_k^i + \tilde{\alpha}_k^i d_k^i; \epsilon_k) \le P(y_k^i; \epsilon_k) - \gamma(\tilde{\alpha}_k^i)^2, \tag{37}$$

and a stepsize α_k^i is produced by the line search such that either

$$y_k^i + \frac{\alpha_k^i}{\delta} d_k^i \notin X \tag{38}$$

or

$$P\left(y_k^i + \frac{\alpha_k^i}{\delta} d_k^i; \epsilon_k\right) > P(y_k^i; \epsilon_k) - \gamma\left(\frac{\alpha_k^i}{\delta}\right)^2. \tag{39}$$

Regarding the behavior of the penalty function along the opposite direction $-d_k^i$, if (34) or (35) holds, the algorithm investigates along the direction $-d_k^i$. Similarly to the analysis along d_k^i, it determines that the initial stepsize $\tilde{\alpha}_k^i$ satisfies either

$$y_k^i + \tilde{\alpha}_k^i(-d_k^i) \notin X \tag{40}$$

or

$$P(y_k^i + \tilde{\alpha}_k^i(-d_k^i); \epsilon_k) > P(y_k^i; \epsilon_k) - \gamma(\tilde{\alpha}_k^i)^2, \tag{41}$$

or it computes a stepsize α_k^i such that either

$$y_k^i + \frac{\alpha_k^i}{\delta}(-d_k^i) \notin X \tag{42}$$

or

$$P\left(y_k^i + \frac{\alpha_k^i}{\delta}(-d_k^i); \epsilon_k\right) > P(y_k^i; \epsilon_k) - \gamma\left(\frac{\alpha_k^i}{\delta}\right)^2. \tag{43}$$

If (36) and (37) hold, the algorithm does not consider the opposite direction $-d_k^i$ directly, but it can extract information on the behavior of P along $-d_k^i$ by using relation (37). In fact, by setting $\tilde{y}_k^i = y_k^i + \tilde{\alpha}_k^i d_k^i$, relation (37) can be rewritten as

$$P(\tilde{y}_k^i + \tilde{\alpha}_k^i(-d_k^i); \epsilon_k) \geq P(\tilde{y}_k^i; \epsilon_k) - \gamma(-(\tilde{\alpha}_k^i)^2). \tag{44}$$

Now let us consider the (sub)sequence $\{x_k\}_K$, where

$$K = \{0, 1, 2, \ldots\} \quad \text{if } \lim_{k\to\infty} \epsilon_k = \bar{\epsilon} > 0,$$
$$K = \{k : \epsilon_{k+1} < \epsilon_k\} \quad \text{if } \lim_{k\to\infty} \epsilon_k = 0.$$

The instructions of Algorithm DFL imply that $x_k \in X$, for all k, so that the sequence $\{x_k\}_K$ admits limit points. Then, let $\bar{x} \in X$ be a limit point of $\{x_k\}_K$. By using Proposition 5 we have

$$\lim_{k\to\infty, k\in K} \alpha_k^i = 0 \qquad \text{for } i = 1, \ldots, n, \tag{45}$$

$$\lim_{k\to\infty, k\in K} \tilde{\alpha}_k^i = 0 \qquad \text{for } i = 1, \ldots, n. \tag{46}$$

By recalling the definitions of the search direction d_k^i, $i = 1, \ldots, n$, we obtain

$$D \cap D(\bar{x}) \subseteq \{d_k^1, -d_k^1, \ldots, d_k^n, -d_k^n\}. \tag{47}$$

Now by using (45)–(47) and Proposition 2, we have that, for sufficiently large k and for all $d_k^i \in D \cap D(\bar{x})$, neither (34) nor (38) can happen and that, for sufficiently large k and for all $-d_k^i \in D \cap D(\bar{x})$, neither (40) nor (42) can happen.

Now we prove the theorem by showing that all the requirements of Proposition 4 hold. Let us consider all the directions $d^i \in D \cap D(\bar{x})$.

If $d^i = d_k^i$, assumptions (8), (9) of Proposition 4 follow, for sufficiently large k, by setting $\xi_k^i = \tilde{\alpha}_k^i$, $y_k^i = y_k^i$, and $o(\xi_k^i) = \gamma(\tilde{\alpha}_k^i)^2$ if (35) holds or by setting $\xi_k^i = \frac{\alpha_k^i}{\delta}$, $y_k^i = y_k^i$, and $o(\xi_k^i) = \gamma(\frac{\alpha_k^i}{\delta})^2$ if (39) holds.

If $d^i = -d_k^i$, assumptions (8), (9) of Proposition 4 follow, for sufficiently large k, by setting $\xi_k^i = \tilde{\alpha}_k^i$, $y_k^i = y_k^i$, and $o(\xi_k^i) = \gamma(\tilde{\alpha}_k^i)^2$ if (41) holds; by setting $\xi_k^i = \frac{\alpha_k^i}{\delta}$, $y_k^i = y_k^i$, and $o(\xi_k^i) = \gamma(\frac{\alpha_k^i}{\delta})^2$ if (43) holds; or by setting $\xi_k^i = \tilde{\alpha}_k^i$, $y_k^i = \tilde{y}_k^i$, and $o(\xi_k^i) = -\gamma(\tilde{\alpha}_k^i)^2$ if (44) holds.

Now, from the definitions of ξ_k^i, y_k^i and the fact that $y_k^i = x_k + \sum_{j=1}^{i-1} \alpha_k^j d_k^j$, we obtain

$$\frac{\max\{\xi_k^i, \|x_k - y_k^i\|\}}{\epsilon_k} \leq \frac{1}{\delta\epsilon_k} \max\left\{\max_{i=1,\ldots,n}\{\tilde{\alpha}_k^i, \alpha_k^i\}, \sum_{j=1}^{n} \alpha_k^j\right\} \quad \forall i : d^i \in D \cap D(\bar{x}). \tag{48}$$

If $K = \{0, 1, 2, \ldots\}$ and $\lim_{k\to\infty} \epsilon_k = \bar{\epsilon} > 0$, point (10) of Proposition 4 follows from (45), (46), and (48).

If $K = \{k : \epsilon_{k+1} < \epsilon_k\}$, the instructions of Step 2 imply that, for all $k \in K$,

$$\frac{\max_{i=1,\ldots,n}\{\tilde{\alpha}_k^i, \alpha_k^i\}}{\epsilon_k} \leq \max_{i=1,\ldots,n}\{\tilde{\alpha}_k^i, \alpha_k^i\}^{\frac{p-1}{p}}. \tag{49}$$

Then by (48) and (49) we obtain

$$\frac{\max\{\xi_k^i, \|x_k - y_k^i\|\}}{\epsilon_k} \leq \frac{1}{\delta} \max\left\{\max_{i=1,\ldots,n}\{\tilde{\alpha}_k^i, \alpha_k^i\}^{\frac{p-1}{p}}, \sum_{j=1}^{n} \alpha_k^j\right\} \quad \forall i : d^i \in D \cap D(\bar{x}), \tag{50}$$

which, along with (45) and (46), shows that (10) of Proposition 4 holds.

Finally, (11) of Proposition 4 follows from the updating rule of the penalty parameter ϵ_k at Step 2 of Algorithm DFL. ∎

By the proof of Theorem 1, the following corollary better characterizes the accumulation points of $\{x_k\}$ which correspond to KKT points of problem (1).

COROLLARY 2. *Let $\{x_k\}$, $\{\epsilon_k\}$ be the sequences produced by Algorithm* DFL*, and assume that every limit point of $\{x_k\}$ satisfies the MFCQ.*

(i) *If $\lim_{k\to\infty} \epsilon_k = \bar{\epsilon} > 0$, then every accumulation point of $\{x_k\}$ is stationary for problem* (1).

(ii) *If $\lim_{k\to\infty} \epsilon_k = 0$, then every accumulation point of $\{x_k\}_K$ is stationary for problem* (1)*, where $K = \{k : \epsilon_{k+1} < \epsilon_k\}$.*

5. Numerical experiments. In this section we report the numerical performance of the proposed sequential penalty derivative-free Algorithm DFL both on a set of academic test problems and on a real application arising in the optimal design of an interplanetary trajectory for a space mission. The experimentation on smooth academic problems has been conducted mainly to evaluate the influence of the exponent q of the penalty terms

$$\left[g_j(x)^+\right]^q, \qquad j = 1, \ldots, m.$$

Since quadratic penalty function methods are subject to ill-conditioning, we decided to experiment with different values of q with $1 < q \leq 2$.

The proposed method has been implemented in double precision Fortran 90, and all the experiments have been conducted by choosing the following values for the parameters defining Algorithm DFL: $\gamma = 10^{-6}$, $\theta = 0.5$, $p = 2$,

$$\tilde{\alpha}_0^i = \max\left\{10^{-3}, \min\{1, |(x_0)^i|\}\right\}, \qquad i = 1, \ldots, n.$$

Concerning the penalty parameter, in the implementation of Algorithm DFL we use a vector of penalty parameters $\epsilon \in \Re^m$ and choose

$$(\epsilon_0)^j = \begin{cases} 10^{-3} & \text{if } g_j(x_0)^+ < 1, \\ 10^{-1} & \text{otherwise,} \end{cases} \qquad j = 1, \dots, m. \tag{51}$$

In order to preserve all the theoretical results, the test at Step 2 of Algorithm DFL, $\max_{i=1,\dots,n}\{\tilde{\alpha}_k^i, \alpha_k^i\} \le \epsilon_k^p$, has been replaced by

$$\max_{i=1,\dots,n}\{\tilde{\alpha}_k^i, \alpha_k^i\} \le \max_{i=1,\dots,m}\{(\epsilon_k)^i\}^p.$$

As termination criterion, we stop the algorithm whenever $\max_{i=1,\dots,n}\{\tilde{\alpha}_k^i, \alpha_k^i\} \le 10^{-5}$. As a consequence of this stopping condition and of the initialization (51), we have that the final values of the penalty parameters are greater than 10^{-5}. Finally, we allow a maximum of 5000 function evaluations.

5.1. Results of test problems. We selected a set of 50 test problems from the well-known collections [12, 23]. In Table 3 we report the details of the selected test problems. Namely, for each problem we indicate by n the number of variables, by m the number of nonlinear plus general linear constraints, and by $\bar{n}$ the number of bound constraints on the variables; f_0 denotes the value of the objective function on the initial point, that is, $f_0 = f(x_0)$; and finally, viol_0 is a measure of the infeasibility on the initial point, that is, $\text{viol}_0 = \sum_{j=1}^m g_j(x_0)^+$. In the table we indicate (by an "$*$" symbol after the name) the problems whose initial points are infeasible with respect to the bound constraints. In those cases we obtained an initial point by projecting the provided point onto the set defined by the bound constraints.

First, in order to assess the influence of the exponent q of the penalty terms, we compare two versions of the code with $q = 2$ and $q = 1.1$, respectively. We report in Table 4 the final objective function value (f^*) and the final constraint violation (viol^*) for the two versions of the code.

By considering the results reported in Table 4 we note that the algorithm with $q = 1.1$ solves 44 problems out of 50 with a final constraint violation $\text{viol}^* < 10^{-4}$, whereas, when $q = 2$, 31 problems are solved with $\text{viol}^* < 10^{-4}$. However, we remark that the case $q = 2$ is slightly more efficient in terms of final function value f^*, possibly at the expense of increased constraint violations.

We also performed some experiments with $q = 1$ (even though in this last case the penalty function is not differentiable and the global convergence theory developed does not hold). The obtained results (which are not reported here) show that this version of the algorithm is slightly worse than that with $q = 1.1$ in terms of final constraint violation.

On the basis of these results it seems that the choice with $q = 1.1$ leads to a more efficient version of the algorithm, at least in terms of the number of function evaluations and obtained feasibility.

Furthermore, we ran on the same set of test problems other available derivative-free optimization software: NOMAD [2, 5, 6, 1], which works directly with function values, and COBYLA [22], which constructs models of the objective and constraint functions. Both codes were run using their default parameter settings, except for the maximum number of function evaluations, which was set to 5000. In Appendix B we report Table 5 with the complete results of the two codes. In order to help the reader evaluate the performance of the algorithms, we present in Table 1 some cumulative results: the number of test problems where they converge to a feasible point, that is,

Table 1
Cumulative results of 50 *test problems.*

	DFL	NOMAD	COBYLA
Number of feasible problems	44	30	36
Number of function evaluations	360.4	446.1	447.9

Table 2
Cumulative results of 1000 *runs on the real application problem.*

	DFL	NOMAD	COBYLA
Number of feasible problems	808	841	489
Number of function evaluations	403.9	3231.3	11305
Best feasible function value	6.31	5.70	5.48

a point with a feasibility violation less than 10^{-4}, and the average number of function evaluations over the set of 23 problems where all the methods converge to a feasible point.

On the basis of the above experimentation, we can say that the proposed method has a satisfying computational behavior when compared with other well-known derivative-free optimization software.

5.2. Results of a real application problem. We have considered a real application belonging to a family of hard global optimization problems arising in the definition of a trajectory for a space vehicle. Important references on the subject can be found in [3, 13]. These problems have relatively few variables (a few tens at most), and do not have derivative information since function evaluations require the numerical solution of a system of differential equations. The Advanced Concepts Team at the European Space Agency (ESA) maintains a web site where many instances of trajectory optimization problems are available in the form of MATLAB or C source code [9].

We focus on the Cassini spacecraft trajectory design problem, characterized by six variables, four black box nonlinear constraints, and lower and upper bounds on the variables.

Global optimization issues are out of the scope of this paper, so we limited our experiments to performing 1000 runs for each of the three codes (DFL, NOMAD, COBYLA) starting from random unfeasible points. The obtained results are shown in Table 2, where we report for each algorithm the number of instances (success runs) where they converge to a feasible point, the average number of function evaluations over the success runs, and the best feasible function value attained.

Regarding the number of feasible points found by the codes, both NOMAD and DFL outperform COBYLA with NOMAD being slightly more efficient than DFL. In terms of computational burden, the best performing code is DFL. Finally, the best code with respect to the best feasible objective function value is COBYLA.

The results point out that each method has some good feature which may be useful for the considered application. Hence, the proposed sequential penalty method DFL exhibits good performance as a local optimizer applied to a real-world problem. Furthermore, it could be a valuable option for embedding in a global optimization framework.

6. Concluding remarks. In this paper we have extended a sequential penalty approach for nonlinear programming to a derivative-free context. The extension is not immediate in that the global convergence relies on a suitable connection between the sampling technique and the updating rule for the penalty parameter. The main result of this paper is a general convergence theorem which, under mild assumptions, states sufficient conditions on the sampling technique performed by the derivative-free algorithm and on the updating rule of the penalty parameter that are able to guarantee convergence toward a stationary point of problem (1). Moreover, we have presented an algorithm based on a derivative-free line search strategy whose convergence proof has been derived from the general convergence result. We remark that the generality of Proposition 4 allows us to define other convergent derivative-free methods based on different sampling strategies [19]. Finally, the numerical experiments carried out both on standard test problems and on a real application problem show the effectiveness of the proposed method compared to that of other well-known derivative-free methods.

Appendix A. Technical results. First we state a technical result concerning a property of sequences of nonzero scalars which will be used in the proof of the next proposition.

LEMMA 1. *Let $\{a_k^i\}$, $i = 1, \ldots, p$, be sequences of nonzero scalars. There exist an index $i^\star \in \{1, \ldots, p\}$ and an infinite subset $K \subseteq \{0, 1, \ldots\}$ such that*

$$\lim_{k\to\infty, k\in K} \frac{a_k^i}{|a_k^{i^\star}|} = z_i, \qquad |z_i| < +\infty, \qquad i = 1, \ldots, p. \tag{52}$$

Proof. The assertion is true if $p = 1$. We prove the thesis by induction on p. Suppose that there exist an integer $\hat{\imath} \in \{1, \ldots, p-1\}$ and a subset $\hat{K} \subseteq \{0, 1, \ldots\}$ such that

$$\lim_{k\to\infty, k\in \hat{K}} \frac{a_k^i}{|a_k^{\hat{\imath}}|} = z_i, \qquad |z_i| < +\infty, \qquad i = 1, \ldots, p-1. \tag{53}$$

Two cases can occur:

(i) the sequence $\{\frac{a_k^p}{|a_k^{\hat{\imath}}|}\}_{\hat{K}}$ is bounded;

(ii) there exists at least one unbounded subsequence $\{\frac{a_k^p}{|a_k^{\hat{\imath}}|}\}_{K_1}$, with $K_1 \subseteq \hat{K}$.

In case (i) we can extract a convergent subsequence $\{\frac{a_k^p}{|a_k^{\hat{\imath}}|}\}_{K_2}$, $K_2 \subseteq \hat{K}$, and the thesis is proved taking $i^\star = \hat{\imath}$ and $K = K_2$.

In case (ii) we have

$$\lim_{k\to\infty, k\in K_1} \frac{a_k^{\hat{\imath}}}{a_k^p} = 0.$$

Then we can write for $i = 1, \ldots, p-1$

$$\lim_{k\to\infty, k\in K_1} \frac{a_k^i}{a_k^p} = \frac{a_k^i}{a_k^{\hat{\imath}}} \frac{a_k^{\hat{\imath}}}{a_k^p} = 0,$$

from which the thesis is proved taking $i^\star = p$ and $K = K_1$. ∎

PROPOSITION 6. *Let the assumptions of Proposition* 4 *be satisfied, and define*

$$\lambda_l(x; \epsilon) = \frac{q}{\epsilon} \max\{g_l(x), 0\}^{q-1}, \qquad l = 1, \ldots, m.$$

Then the sequences $\{\lambda_l(x_k; \epsilon_k)\}$, $l = 1, \ldots, m$, are bounded.

Proof. Let us denote $\bar{D} = D \cap D(\bar{x})$. By applying the mean-value theorem to (9), we can write

$$-o(\xi_k^i) \le P(y_k^i + \xi_k^i d^i; \epsilon_k) - P(y_k^i; \epsilon_k) = \xi_k^i \nabla P(u_k^i; \epsilon_k)^T d^i \qquad \forall d^i \in \bar{D},$$

where $u_k^i = y_k^i + t_k^i \xi_k^i d^i$, with $t_k^i \in (0,1)$. Thus, we have

$$-\frac{o(\xi_k^i)}{\xi_k^i} \le \nabla P(u_k^i; \epsilon_k)^T d^i \qquad \forall d^i \in \bar{D}.$$

By considering the expression of $P(x;\epsilon)$, we can write

$$\begin{aligned} (54)\quad & \nabla P(u_k^i;\epsilon_k)^T d^i \\ &= \left(\nabla f(u_k^i) + \frac{q}{\epsilon_k} \sum_{l=1}^m \max\{g_l(u_k^i), 0\}^{q-1} \nabla g_l(u_k^i) \right)^T d^i \ge -\frac{o(\xi_k^i)}{\xi_k^i} \qquad \forall d^i \in \bar{D}. \end{aligned}$$

Recalling that $u_k^i = y_k^i + t_k^i \xi_k^i d^i$, with $t_k^i \in (0,1)$, we have that

$$\lim_{k\to\infty, k\in K} u_k^i = \bar{x}. \tag{55}$$

By recalling the expression of $\lambda_l(x;\epsilon)$, $l = 1,\dots,m$, we can rewrite relation (54) as

$$\left(\nabla f(u_k^i) + \sum_{l=1}^m \lambda_j(u_k^i;\epsilon_k) \nabla g_j(u_k^i) \right)^T d^i \ge -\frac{o(\xi_k^i)}{\xi_k^i} \qquad \forall i : d^i \in \bar{D}. \tag{56}$$

First we prove that

$$\lim_{k\to\infty, k\in K} \left| \lambda_l(u_k^i;\epsilon_k) - \lambda_l(x_k;\epsilon_k) \right| = 0, \qquad l = 1,\dots,m \quad \forall\ i : d^i \in \bar{D}. \tag{57}$$

In fact,

$$\begin{aligned} (58)\quad & \left| \max\left\{ \frac{g_j(u_k^i)}{\epsilon_k}, 0 \right\}^{q-1} - \max\left\{ \frac{g_j(x_k)}{\epsilon_k}, 0 \right\}^{q-1} \right| \\ &= \left| \max\left\{ \frac{g_j(x_k)}{\epsilon_k} + \frac{1}{\epsilon_k} \nabla g_j(\tilde{u}_k^{i,j})(u_k^i - x_k), 0 \right\}^{q-1} - \max\left\{ \frac{g_j(x_k)}{\epsilon_k}, 0 \right\}^{q-1} \right| \\ &\le \left| \max\left\{ \frac{g_j(x_k)}{\epsilon_k}, 0 \right\}^{q-1} + \max\left\{ \frac{1}{\epsilon_k} \nabla g_j(\tilde{u}_k^{i,j})(u_k^i - x_k), 0 \right\}^{q-1} - \max\left\{ \frac{g_j(x_k)}{\epsilon_k}, 0 \right\}^{q-1} \right| \\ &= \max\left\{ \frac{1}{\epsilon_k} \nabla g_j(\tilde{u}_k^{i,j})(u_k^i - x_k), 0 \right\}^{q-1} \le \frac{\|\nabla g_j(\tilde{u}_k^{i,j})\|^{q-1} \|(u_k^i - x_k)\|^{q-1}}{\epsilon_k^{q-1}} \\ &\le c_1 \frac{\max_{i:d^i\in\bar{D}} \{(\xi_k^i)^{q-1}, \|y_k^i - x_k\|^{q-1}\}}{\epsilon_k^{q-1}}, \end{aligned}$$

where $\tilde{u}_k^{i,j} = u_k^i + \tilde{t}_k^{i,j} x_k$ with $\tilde{t}_k^{i,j} \in (0,1)$. Hence, by recalling assumption (10), (57) is proved.

By simple manipulations (56) can be rewritten as

$$\begin{aligned} & \Bigg(\nabla f(u_k^i) + \sum_{l=1}^m \nabla g_l(u_k^i) \lambda_l(x_k;\epsilon_k) \\ (59)\quad & + \sum_{l=1}^m \nabla g_l(u_k^i)(\lambda_l(u_k^i;\epsilon_k) - \lambda_l(x_k;\epsilon_k)) \Bigg)^T d^i \ge -\frac{o(\xi_k^i)}{\xi_k^i} \qquad \forall i : d^i \in \bar{D}. \end{aligned}$$

Now we show that the sequences $\{\lambda_l(x_k;\epsilon_k)\}$, $l=1,\dots,m$, are bounded. Let

$$\{a_k^1,\dots,a_k^m\}=\{\lambda_1(x_k;\epsilon_k),\dots,\lambda_m(x_k;\epsilon_k)\}.$$

By contradiction let us assume that there exists at least an index $h\in\{1,\dots,m\}$ such that

$$\lim_{k\to\infty,k\in K}|a_k^h|=+\infty.$$

From Lemma 1 we get that there exist an infinite subset (again, relabelled K) and an index $s\in\{1,\dots,m\}$ such that for $k\to\infty,k\in K$, and

$$\lim_{k\to\infty,k\in K}\frac{a_k^i}{|a_k^s|}=z_i,\quad |z_i|<+\infty,\quad i+1,\dots,m. \tag{60}$$

Note that

$$z_s=1 \qquad \text{and} \qquad |a_k^s|\to+\infty. \tag{61}$$

Dividing relation (59) by $|a_k^s|$, we have

$$\left(\frac{\nabla f(u_k^i)}{|a_k^s|}+\sum_{l=1}^m\frac{\nabla g_l(u_k^i)a_k^l}{|a_k^s|}\right. \\ \left.+\sum_{l=1}^m\nabla g_l(u_k^i)\frac{\lambda_l(u_k^i;\epsilon_k)-\lambda_l(x_k;\epsilon_k)}{|a_k^s|}\right)^T d^i\ge-\frac{o(\xi_k^i)}{\xi_k^i|a_k^s|}\qquad\forall i:d^i\in\bar D, \tag{62}$$

Taking the limits for $k\to\infty$ and $k\in K$, recalling that $|a_k^s|\to\infty$, and using (57), (60), and (55), we obtain

$$\sum_{l=1}^m z_l\nabla g_l(\bar x)^Td^i\ge 0\qquad\forall i:d^i\in\bar D. \tag{63}$$

Recall that, by assumption, $\bar x$ satisfies the MFCQ, and let $\hat d\in D(\bar x)$ be the direction considered in Definition 2. From Proposition 3 we have that

$$\hat d=\sum_{i:d^i\in\overline{D}}\hat\beta_i d^i. \tag{64}$$

From (64) and (63) we obtain

$$\sum_{l=1}^m z_l\nabla g_l(\bar x)^T\hat d=\sum_{i:d^i\in\bar D}\hat\beta_i\left(\sum_{l=1}^m z_l\nabla g_l(\bar x)\right)^T d^i\ge 0. \tag{65}$$

Now suppose $z_h>0$ for some $h\in\{1,\dots,m\}$. Note that we have $z_i=0$ for all $i\notin I^+(\bar x)$. We can write

$$\sum_{l=1}^m z_l\nabla g_l(\bar x)^T\hat d\le z_h\nabla g_h(\bar x)^T\hat d<0,$$

which contradicts (65) and concludes the proof. ∎

Appendix B. Detailed numerical results. In this section we report three tables concerning our numerical experimentation. Table 3 reports the relevant details of the selected test problems. Table 4 shows the comparison of two versions of DFL with $q=2$ and $q=1.1$, respectively. Finally, Table 5 reports the complete results of the comparison of DFL with NOMAD and COBYLA.

TABLE 3
Test problem characteristics.

Problem	n	m	$\bar{n}$	f_0	viol_0
HS 14	2	3	0	1.00E+00	4.00E+00
HS 15	2	2	1	9.09E+02	3.00E+00
HS 16	2	2	3	9.09E+02	1.00E+00
HS 18	2	2	4	4.04E+00	2.10E+01
HS 19	2	2	4	-1.81E+03	1.20E+02
HS 20	2	3	2	5.85E+01	0.00E+00
HS 21	2	1	4	-9.90E+01	0.00E+00
HS 22	2	2	0	1.00E+00	4.00E+00
HS 23	2	5	4	1.00E+01	2.00E+00
HS 30	3	1	6	3.00E+00	0.00E+00
HS 31	3	1	6	1.90E+01	0.00E+00
HS 39	4	4	0	-2.00E+00	1.00E+01
HS 40	4	6	0	-4.10E-01	2.90E-01
HS 42	4	4	0	1.40E+01	1.00E+00
HS 43	4	3	8	0.00E+00	0.00E+00
HS 60	3	2	6	1.00E+00	1.80E+01
HS 64	3	1	3	2.66E+05	1.60E+02
HS 65	3	1	6	1.14E+02	0.00E+00
HS 72	4	2	8	5.00E+00	7.50E+00
HS 74	4	8	8	0.00E+00	8.00E+02
HS 75	4	8	8	0.00E+00	8.00E+02
HS 78	5	6	0	-6.00E+00	3.60E+00
HS 79	5	6	0	1.00E+00	7.80E+00
HS 80	5	6	10	3.36E-04	4.00E+00
HS 83	5	6	10	-3.22E+04	3.20E+00
HS 95	6	4	12	0.00E+00	7.80E+01
HS 96	6	4	12	0.00E+00	1.20E+02
HS 97	6	4	12	0.00E+00	7.80E+01
HS 98	6	4	12	0.00E+00	1.70E+02
HS 100	7	4	14	7.14E+02	0.00E+00
HS 101	7	6	14	2.21E+03	3.70E+02
HS 104	8	6	16	3.66E+00	4.20E-01
HS 106	8	6	16	1.50E+04	6.20E+04
HS 107	9	6	8	4.85E+03	8.00E-01
HS 113	10	8	0	7.53E+02	0.00E+00
HS 114	10	14	20	-8.72E+02	4.40E-01
HS 116	13	15	26	2.50E+02	8.10E+01
HS 223	2	2	4	-1.00E-01	0.00E+00
HS 225	2	5	0	1.00E+01	2.00E+00
HS 228	2	2	0	0.00E+00	0.00E+00
HS 230	2	2	0	0.00E+00	1.00E+00
HS 263	4	6	0	-1.00E+01	1.10E+03
HS 315	2	3	0	9.00E-01	0.00E+00
HS 323	2	2	2	5.00E+00	0.00E+00
HS 343	3	2	6	-3.88E+00	3.60E-01
HS 365*	7	5	4	6.00E+00	1.80E+00
HS 369*	8	6	16	2.10E+03	3.80E+01
HS 372*	9	12	6	7.14E+05	9.80E+01
HS 373	9	12	0	7.53E+05	2.60E+02
HS 374	10	35	0	1.00E-01	8.10E-01

Table 4
Comparison between the two versions of DFL *with $q = 2$ and $q = 1.1$, respectively.*

Problem	$q = 2$			$q = 1.1$		
	nF	f^*	viol*	nF	f^*	viol*
HS 14	5002	9.51E+00	1.20E-05	140	1.93E+00	0.00E+00
HS 15	83	3.04E+02	3.50E-03	86	3.07E+02	0.00E+00
HS 16	151	2.50E-01	0.00E+00	151	2.50E-01	0.00E+00
HS 18	136	7.25E+00	1.90E-06	136	7.25E+00	0.00E+00
HS 19	5001	-6.98E+03	6.50E-03	871	-6.96E+03	0.00E+00
HS 20	80	4.02E+01	3.60E-04	78	4.02E+01	0.00E+00
HS 21	80	-1.00E+02	0.00E+00	80	-1.00E+02	0.00E+00
HS 22	109	1.00E+00	3.90E-06	120	1.00E+00	1.10E-16
HS 23	122	2.00E+00	1.10E-05	113	2.00E+00	0.00E+00
HS 30	134	1.00E+00	0.00E+00	134	1.00E+00	0.00E+00
HS 31	5001	6.01E+00	3.40E-05	134	1.00E+01	0.00E+00
HS 39	5001	-1.43E-01	3.90E-06	164	-6.35E-05	1.20E-10
HS 40	294	-2.26E-01	4.00E-06	294	-2.33E-01	1.30E-06
HS 42	5003	1.40E+01	1.70E-05	219	1.40E+01	0.00E+00
HS 43	5002	-2.42E+01	1.00E-05	247	-2.26E+01	0.00E+00
HS 60	5002	1.90E+01	8.10E-06	200	2.39E+01	6.30E-08
HS 64	627	6.28E+03	1.10E-02	1275	6.72E+03	0.00E+00
HS 65	164	5.22E+00	7.50E-06	131	5.23E+00	0.00E+00
HS 72	709	3.92E+02	2.50E-02	564	7.22E+02	1.50E-04
HS 74	1109	5.26E+03	2.70E-02	1225	5.28E+03	7.20E-05
HS 75	1148	5.26E+03	9.70E-02	1225	5.28E+03	7.20E-05
HS 78	359	-1.16E-01	1.40E-06	322	-8.52E-01	2.60E-06
HS 79	5001	9.98E+00	1.20E-04	321	1.06E+02	1.80E-06
HS 80	313	1.00E+00	2.40E-06	321	1.00E+00	2.80E-06
HS 83	1461	-3.07E+04	4.00E-03	265	-3.00E+04	0.00E+00
HS 95	181	3.72E+00	2.40E+01	181	3.72E+00	2.40E+01
HS 96	166	3.92E+00	6.40E+01	166	3.92E+00	6.40E+01
HS 97	181	3.72E+00	2.40E+01	181	3.72E+00	2.40E+01
HS 98	166	3.92E+00	1.20E+02	166	3.92E+00	1.20E+02
HS 100	5004	6.83E+02	1.40E-05	397	6.85E+02	0.00E+00
HS 101	5003	1.97E+03	2.80E-02	688	2.39E+03	0.00E+00
HS 104	5002	4.02E+00	1.50E-05	440	4.14E+00	0.00E+00
HS 106	629	1.29E+04	0.00E+00	629	1.29E+04	0.00E+00
HS 107	3425	4.76E+03	2.50E-02	1102	6.99E+03	0.00E+00
HS 113	5002	2.83E+01	3.30E-06	581	2.83E+01	0.00E+00
HS 114	5003	-1.56E+03	1.70E-01	651	-9.31E+02	2.50E-07
HS 116	5001	5.00E+01	6.20E-02	656	1.21E+02	1.00E-01
HS 223	142	-8.34E-01	3.20E-06	105	-1.77E-01	0.00E+00
HS 225	122	2.00E+00	1.10E-05	113	2.00E+00	0.00E+00
HS 228	93	-3.00E+00	0.00E+00	93	-3.00E+00	0.00E+00
HS 230	414	3.75E-01	2.60E-06	153	9.92E-01	0.00E+00
HS 263	334	6.00E+00	1.50E-06	295	1.45E+00	1.20E-06
HS 315	104	-4.50E-01	3.40E-06	103	-4.50E-01	0.00E+00
HS 323	5003	3.92E+00	0.00E+00	195	4.96E+00	0.00E+00
HS 343	133	-2.86E+00	0.00E+00	133	-2.86E+00	0.00E+00
HS 365	5003	0.00E+00	1.00E+00	5003	0.00E+00	1.00E+00
HS 369	416	2.10E+03	0.00E+00	509	2.10E+03	0.00E+00
HS 372	5002	1.13E+05	1.20E-01	1226	5.47E+04	0.00E+00
HS 373	5002	1.56E+05	2.30E-01	836	1.57E+05	8.50E-07
HS 374	5002	3.06E-01	6.70E-06	569	1.32E+00	0.00E+00

Table 5
Complete results of test problems.

Problem	DFL			NOMAD			COBYLA		
hs14	140	1.93E+00	0.00E+00	132	1.58E+00	0.00E+00	21	5.63E+00	2.50E-01
hs15	86	3.07E+02	0.00E+00	95	3.07E+02	0.00E+00	79	1.03E+02	6.18E-01
hs16	151	2.50E-01	0.00E+00	290	4.24E-01	0.00E+00	31	1.00E+00	6.18E-10
hs18	136	7.25E+00	0.00E+00	211	5.35E+00	0.00E+00	23	2.07E+01	4.52E+00
hs19	871	-6.96E+03	0.00E+00	97	-6.69E+03	0.00E+00	59	-5.78E+02	1.16E+01
hs20	78	4.02E+01	0.00E+00	89	4.03E+01	0.00E+00	19	6.50E+00	5.00E-01
hs21	80	-1.00E+02	0.00E+00	107	-1.00E+02	0.00E+00	19	-9.93E+01	8.33E-01
hs22	120	1.00E+00	1.10E-16	122	1.00E+00	0.00E+00	49	8.25E-01	3.93E-11
hs23	113	2.00E+00	0.00E+00	162	2.04E+00	0.00E+00	21	2.00E+00	0.00E+00
hs30	134	1.00E+00	0.00E+00	82	1.00E+00	0.00E+00	68	1.00E+00	0.00E+00
hs31	134	1.00E+01	0.00E+00	419	6.09E+00	0.00E+00	77	6.00E+00	2.84E-11
hs39	164	-6.35E-05	1.20E-10	191	1.35E-01	4.29E-01	125	-1.00E+00	2.52E-10
hs40	294	-2.33E-01	1.30E-06	161	-3.10E-01	2.75E-01	90	-2.50E-01	2.56E-10
hs42	219	1.40E+01	0.00E+00	162	1.43E+01	5.91E-02	100	1.39E+01	2.79E-10
hs43	247	-2.26E+01	0.00E+00	577	-4.40E+01	0.00E+00	109	-4.40E+01	2.13E-10
hs60	200	2.39E+01	6.30E-08	112	1.10E+00	3.31E-03	70	3.26E-02	2.62E-10
hs64	1275	6.72E+03	0.00E+00	26	3.50E+06	0.00E+00	387	6.30E+03	4.80E-15
hs65	131	5.23E+00	0.00E+00	427	9.67E-01	0.00E+00	80	9.54E-01	2.45E-10
hs72	564	7.22E+02	1.50E-04	83	8.50E+04	0.00E+00	968	7.28E+02	8.27E-17
hs74	1225	5.28E+03	7.20E-05	397	5.20E+03	5.18E-01	5000	2.35E+03	2.54E+02
hs75	1225	5.28E+03	7.20E-05	355	5.12E+03	1.25E+00	5000	2.35E+03	2.54E+02
hs78	322	-8.52E-01	2.60E-06	211	-8.20E-01	3.33E+00	92	-2.92E+00	6.25E-10
hs79	321	1.06E+02	1.80E-06	223	1.13E+01	1.32E+00	96	8.23E-02	1.07E-09
hs80	321	1.00E+00	2.80E-06	295	2.26E-01	4.99E-03	77	5.40E-02	1.39E-10
hs83	265	-3.00E+04	0.00E+00	982	-3.06E+04	0.00E+00	73	-3.07E+04	1.85E-13
hs95	181	3.72E+00	2.40E+01	124	2.49E-01	5.14E+01	50	-3.62E+00	2.18E-02
hs96	166	3.92E+00	6.40E+01	124	2.49E-01	1.08E+02	139	5.27E+00	3.17E-02
hs97	181	3.72E+00	2.40E+01	124	2.49E-01	5.14E+01	50	-3.62E+00	2.18E-02
hs98	166	3.92E+00	1.20E+02	124	2.49E-01	1.85E+02	56	-9.84E+00	5.68E-02
hs100	397	6.85E+02	0.00E+00	1102	6.84E+02	0.00E+00	288	6.81E+02	2.05E-09
hs101	688	2.39E+03	0.00E+00	1772	1.23E+03	0.00E+00	1495	3.00E+03	2.24E-01
hs100	440	4.14E+00	0.00E+00	1082	4.10E+00	0.00E+00	2329	3.95E+00	6.10E-11
hs106	629	1.29E+04	0.00E+00	1652	6.64E+03	1.13E-01	5000	7.89E+03	8.68E-02
hs107	1102	6.99E+03	0.00E+00	602	2.54E+12	0.00E+00	177	5.06E+03	7.61E-11
hs113	581	2.83E+01	0.00E+00	998	8.44E+01	0.00E+00	278	2.43E+01	7.99E-11
hs114	651	-9.31E+02	2.50E-07	717	-9.93E+02	2.79E-02	3395	-1.55E+03	9.04E-14
hs116	656	1.21E+02	1.00E-01	1400	8.07E+01	1.07E-02	5000	2.46E+02	8.79E-07
hs223	105	-1.77E-01	0.00E+00	176	-2.82E-01	0.00E+00	26	-8.34E-01	0.00E+00
hs225	113	2.00E+00	0.00E+00	231	2.27E+01	0.00E+00	21	2.00E+00	0.00E+00
hs228	93	-3.00E+00	0.00E+00	281	-2.23E+00	0.00E+00	50	-3.00E+00	2.53E-10
hs230	153	9.92E-01	0.00E+00	198	1.33E+00	0.00E+00	16	3.75E-01	0.00E+00
hs263	295	1.45E+00	1.20E-06	202	2.61E+00	1.08E+00	136	-1.00E+00	1.01E-10
hs315	103	-4.50E-01	0.00E+00	333	2.17E-01	0.00E+00	19	-8.00E-01	0.00E+00
hs323	195	4.96E+00	0.00E+00	86	4.47E+00	0.00E+00	52	3.80E+00	8.99E-11
hs343	133	-2.86E+00	0.00E+00	343	-5.68E+00	0.00E+00	69	-5.68E+00	6.90E-12
hs365	5003	0.00E+00	1.00E+00	52	6.00E+00	1.41E+00	574	1.24E+02	1.21E-11
hs369	509	2.10E+03	0.00E+00	399	2.10E+03	0.00E+00	824	2.10E+03	0.00E+00
hs372	1226	5.47E+04	0.00E+00	935	7.23E+05	0.00E+00	5000	2.48E+05	4.74E-06
hs373	836	1.57E+05	8.50E-07	545	7.27E+05	2.83E+02	5000	6.17E+05	4.89E+01
hs374	569	1.32E+00	0.00E+00	408	2.35E+00	0.00E+00	257	2.33E-01	5.48E-10

Acknowledgment. We thank two anonymous referees whose many comments and suggestions greatly helped us to improve this paper.

REFERENCES

[1] M. A. Abramson, C. Audet, G. Couture, J. E. Dennis, Jr., S. Le Digabel, and Q. Reynaud, *The NOMAD Project*, http://www.gerad.ca/nomad.

[2] M. A. Abramson, C. Audet, J. E. Dennis, Jr., and S. Le Digabel, *OrthoMADS: A deterministic MADS instance with orthogonal directions*, SIAM J. Optim., 20 (2009), pp. 948–966.

[3] B. Addis, A. Cassioli, M. Locatelli, and F. Schoen, *Global optimization for the design of space trajectories*, Comput. Optim. and Appl., to appear.

[4] C. Audet and J. E. Dennis, Jr., *A pattern search filter method for nonlinear programming without derivatives*, SIAM J. Optim., 14 (2004), pp. 980–1010.

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

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

[7] D. P. Bertsekas, *Constrained Optimization and Lagrangian Multiplier Methods*, Academic Press, New York, 1982.

[8] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, MA, 1999.

[9] ESA Advanced Concepts Team, *GOTP Database: Global Optimisation Trajectory Problems and Solutions*, http://www.esa.int/gsp/ACT/inf/op/globopt/mgadsm.htm.

[10] A. V. Fiacco and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley and Sons, New York, 1968.

[11] R. Fletcher and S. Leyffer, *Nonlinear programming without a penalty function*, Math. Program., 91 (2002), pp. 239–269.

[12] W. Hock and K. Schittkowski, *Test Examples for Nonlinear Programming Codes*, Lecture Notes in Econom. and Math. Systems 187, Springer-Verlag, Berlin, New York, 1981.

[13] D. Izzo, V. Becerra, D. Myatt, S. Nasuto, and J. Bishop, *Search space pruning and global optimisation of multiple gravity assist spacecraft trajectories*, J. Global Optim., 38 (2007), pp. 283–296.

[14] T. G. Kolda, R. M. Lewis, and V. Torczon, *Optimization by direct search: New perspectives on some classical and modern methods*, SIAM Rev., 45 (2003), pp. 385–482.

[15] 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, Albuquerque, NM, Livermore, CA, 2006.

[16] 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 (2002), pp. 1075–1089.

[17] C.-J. Lin, S. Lucidi, L. Palagi, A. Risi, and M. Sciandrone, *Decomposition algorithm model for singly linearly-constrained problems subject to lower and upper bounds*, J. Optim. Theory Appl., 141 (2009), pp. 107–126.

[18] G. Liuzzi and S. Lucidi, *A derivative-free algorithm for inequality constrained nonlinear programming via smoothing of an ℓ_∞ penalty function*, SIAM J. Optim., 20 (2009), pp. 1–29.

[19] G. Liuzzi, S. Lucidi, and M. Sciandrone, *Sequential Penalty Derivative-free Methods for Nonlinear Constrained Optimization*, Technical report RT 1/2009, Dipartimento di Sistemi e Informatica, Università di Firenze, Firenze, Italy, 2009.

[20] S. Lucidi and M. Sciandrone, *On the global convergence of derivative-free methods for unconstrained optimization*, SIAM J. Optim., 13 (2002), pp. 97–116.

[21] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Ser. Oper. Res., Springer-Verlag, New York, 1999.

[22] M. J. D. Powell, *A direct search optimization method that models the objective and constraint functions by linear interpolation*, in Advances in Optimization and Numerical Analysis, S. Gomez and J-P. Hennart, eds., Kluwer Academic Publishers, Dordrecht, The Netherlands, 1994, pp. 51–67.

[23] K. Schittkowski, *More Test Examples for Nonlinear Programming Codes*, Lecture Notes in Econom. and Math. Systems 282, Springer-Verlag, Berlin, 1987.