A barrier-type method for multiobjective optimization[*]

Ellen H. Fukuda[†]   L. M. Graña Drummond[‡]   Fernanda M. P. Raupp[§]

March 29, 2018

Abstract

For solving constrained multicriteria problems, we introduce the multiobjective barrier method (MBM), which extends the scalar-valued internal penalty method. This multiobjective version of the classical method also requires a penalty barrier for the feasible set and a sequence of nonnegative penalty parameters. Differently from the single-valued procedure, MBM is implemented by means of an auxiliary "monotonic" real-valued mapping, which may be chosen in a quite large set of functions. Here, we consider problems with continuous objective functions, where the feasible sets are defined by finitely many continuous inequalities. Under mild assumptions, and depending on the monotonicity type of the auxiliary function, we establish convergence to Pareto or weak Pareto optima. Finally, we also propose an implementable version of MBM for seeking local optima and analyze its convergence to Pareto or weak Pareto solutions.

Keywords: Barrier methods, multiobjective optimization, Pareto optimality, penalty methods.

1 Introduction

Practical issues on different areas, such as statistics [4], engineering [5, 19], environmental analysis [18], space exploration [20], management science [15, 21] and design [8] can be modeled as constrained multicriteria minimization problems. There are many procedures for solving these problems, including scalarization techniques and heuristics. Among the scalarization approach, the weighting method is possibly the most well-known. It basically replaces the original multiobjective problem into a minimization of some convex combination of the objectives. Its main drawback is the fact that we do not know a priori which are

1

the weights that do not lead to unbounded scalar problems. Heuristics may neither guarantee convergence. To overcome such drawbacks, extensions to the vector-valued setting of classical real-valued methods have been proposed in recent years [10].

In this work, we propose an extension of the *internal penalty* (or *barrier-type*) method [2, 1], which is a well-known technique for constrained scalar optimization problems with continuous objective and constraints functions. As we know, the scalar-valued method consists in sequentially minimizing, without constraints, the objective with the addition of positive multiples of a feasible set barrier. That is to say, one substitutes a constrained problem by a sequence of unconstrained ones, for which we do have efficient solving techniques. Under reasonable assumptions, the sequence of those minimizers converges to an optimum for the original constrained problem.

Here, we present a vector-valued version of this procedure, that we call *multiobjective barrier method* (MBM). Each iteration requires solving an unconstrained scalar-valued problem. As in the classical barrier method, the iterates generated are feasible, and the penalized unconstrained objective values decreases iteratively. Moreover, all accumulation point, if any, are optimal solutions of the original problem. Furthermore, the convergence conditions are generalizations of those required in the real-valued case, and no additional hypotheses are needed. Besides the nonincreasing convergent to zero parameter sequence of positive real numbers and the barrier function, the vector-valued method uses an auxiliary monotonic function (whose presence in the scalar case is immaterial) and, depending on its type of monotonicity, the method converges to Pareto or weak Pareto optimal solutions.

MBM shares some properties with other classical scalar-valued methods extensions, as the steepest descent [7, 14], the projected gradient [12, 9], the Newton [6, 13], the proximal point [3], and the external penalty methods [11]. More precisely, all iterates of these procedures are computed by solving scalar optimization subproblems, and each one of them could be obtained by the application of the corresponding scalar-valued method to a certain (a priori unknown) linear combination of the objectives. These methods (including MBM) seek just a single optimal point and the convergence results are natural extensions of their scalar correlatives. Nevertheless, as shown through numerical experiments (see, for instance, [6]), by initializing them with randomly chosen points, in some cases, one can expect to obtain quite good approximations of the optimal set.

The outline of this article is the following. In Section 2, we introduce the problem and the notion of multiobjective barrier function, and give some examples. In Section 3, we define the auxiliary functions, the multicriteria barrier method, and make some comments about the approach. In Section 4, we analyze the convergence of MBM, and verify that the results are extensions of the classical single objective case. In Section 5, we exhibit a very simple instance of the multiobjective constrained problem for which the weighting method fails to obtain optimal solutions for an arbitrary large family of parameters, while MBM furnishes a Pareto optimal point for the problem. We also show an (*ad hoc*) example in which, by varying a parameter in the barrier, one can retrieve the whole Pareto optimal set. Finally, in Section 6 we draw some conclusions and mention some possible continuations for this work.

2 The multicriteria problem and the internal penalty function

We start this section with some notations which will be used in the whole paper. The Euclidean inner product is written as $\langle \cdot, \cdot \rangle$. For any matrix A, its transpose is denoted by $A^\top$. A vector $x \in \mathbb{R}^n$ with entries $x_i \in \mathbb{R}$, $i = 1, \ldots, n$, is written as $(x_1, \ldots, x_n)^\top$. A sequence of vectors $y^1, y^2, \ldots$ is denoted by $\{y^k\}$. Also, we define the vector with all ones with appropriate dimension as $e := (1, \ldots, 1)^\top$. Moreover, the gradient of a function $\zeta \colon \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$ is denoted by $\nabla \zeta(x)$.

Now, consider $\mathbb{R}^m$ endowed with the partial order induced by the Paretian cone $\mathbb{R}^m_+ = \mathbb{R}_+ \times \cdots \times \mathbb{R}_+$, where $\mathbb{R}_+ = [0, +\infty)$, given by

$$u \leq v \quad \text{if} \quad u_j \leq v_j \quad \text{for all} \quad j = 1, \ldots, m,$$

and with the following stronger relation:

$$u < v \quad \text{if} \quad u_j < v_j \quad \text{for all} \quad j = 1, \ldots, m,$$

with $u, v \in \mathbb{R}^m$.

Given continuous functions $f \colon \mathbb{R}^n \to \mathbb{R}^m$ and $g \colon \mathbb{R}^n \to \mathbb{R}^p$, define

$$D := \{x \in \mathbb{R}^n : g(x) \leq 0\}$$

and consider the following constrained vector-valued optimization problem:

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & x \in D, \end{aligned} \tag{1}$$

understood in the Pareto or weak Pareto sense. Recall that a point $x^* \in D$ is a *weak Pareto optimal* solution of the above problem if there does not exist $x \in D$ such that $f(x) < f(x^*)$. Also, $x^* \in D$ is a *Pareto optimal* solution of (1) if there does not exist $x \in D$ such that $f(x) \leq f(x^*)$ with $f_{j_0}(x) < f_{j_0}(x^*)$ for at least one index $j_0 \in \{1, \ldots, m\}$. Clearly, a Pareto optimum is a weak Pareto optimal solution.

From now on, we assume the existence of a Slater point, i.e., we suppose that

$$D^o := \{x \in \mathbb{R}^n : g(x) < 0\} \neq \emptyset. \tag{2}$$

We intend to find Pareto or weak Pareto solutions of problem (1) by means of a barrier-type method for multiobjective optimization, which produces its iterates in D^o and such that, under certain mild conditions, will approximate optimal points laying outside D^o.

In our setting, we define an internal penalty function for the feasible set D as follows.

Definition 2.1. A *multiobjective barrier function* for the set D is a continuous function $B \colon D^o \to \mathbb{R}^m_+$ such that for any sequence $\{z^k\} \subset D^o$ with $g_i(z^k) \to 0$ for an index $i \in \{1, ..., p\}$, then $\lim_{k \to \infty} B_{j_i}(z^k) = +\infty$ for at least one index $j_i \in \{1, \ldots, m\}$.

Note that, when $m = 1$, we retrieve the classical notion of barrier function used in scalar-valued optimization. See, for instance [2, Section 5.1] or [1, Section 9.4].

2.1 Examples of multiobjective barrier functions

We now give some examples of multiobjective barrier functions for feasible sets D given by finitely many scalar inequalities, which are extensions of the inverse and the logarithmic barriers for real-valued optimization.

Example 2.2. If $p \leq m$, for $x \in D^o$, we define

$$B(x) := \Big(\frac{1}{-g_1(x)}, \ldots, \frac{1}{-g_p(x)}, 0, \ldots, 0\Big)^\top.$$

If $p > m$, for $x \in D^o$, we can take

$$B(x) := \Big[\sum_{i=1}^p \frac{1}{-g_i(x)}\Big] e,$$

recalling that $e = (1, \ldots, 1)^\top \in \mathbb{R}^m$. In this case, we can also take, for example,

$$B(x) := \Big(\frac{1}{-g_1(x)}, \ldots, \frac{1}{-g_{m-1}(x)}, \sum_{i=m}^p \frac{1}{-g_i(x)}\Big)^\top.$$

Example 2.3. Let us consider the case in which $-\log(-g_i(x))$ replaces $\frac{1}{-g_i(x)}$ in the above example. Of course, functions as

$$B(x) := \Big[-\sum_{i=1}^p \log(-g_i(x))\Big] e, \tag{3}$$

may fail to be nonnegative. For our purposes, under certain circumstances, the use of such B is equivalent to the use of a positive function of the form $B - \rho e$, where $\rho \in \mathbb{R}$ is constant. Indeed, as we will see in Section 3, the strategy we are about to propose, which is an extension of the classical barrier method, basically consists of somehow minimizing the objective function f with the addition of a positive multiple of the barrier B. Therefore, the presence of the constant ρ does not have any kind of influence when a minimizer of this perturbed objective function is computed.

For instance, if D is bounded, then B given in (3) is bounded from below, say $B(x) > \rho e$ for all $x \in D^o$, for some constant ρ. So, $\bar{B}(x) := B(x) - \rho e > 0$ for all $x \in D^o$ and, for $\tau > 0$, the minimizers of $f(x) + \tau \bar{B}(x) = f(x) + \tau(B(x) - \rho e)$ are the same as those of $f(x) + \tau B(x)$. (To be more precises, the minimization of the penalized objective will be carried on in $\mathbb{R}$, since, as we will see in Subsection 3.2, this perturbation of f will be composed with a real-valued monotonic function.)

3 A barrier-type method for multiobjective optimization

In this section, we propose a barrier-type method for the constrained multiobjective problem (1). We begin by presenting some functions that will be used in the method.

3.1 Auxiliary functions

Here we introduce the auxiliary functions necessary for the multiobjective-type barrier method. Essentially, they are continuous and have some kind of monotonic behavior. First, let us recall a couple of well-known notions (see [17]).

For a closed set $C \subset \mathbb{R}^m$, a function $\Phi \colon C \to \mathbb{R}$ is called *strictly increasing* or *weakly-increasing* (*w-increasing*) in C if, for all $u, v \in C$,

$$u < v \quad \Rightarrow \quad \Phi(u) < \Phi(v),$$

and Φ is called *strongly increasing* (*s-increasing*) in C if for all $u, v \in C$,

$$u \leq v \text{ and } u_{j_0} < v_{j_0} \text{ for at least one } j_0 \quad \Rightarrow \quad \Phi(u) < \Phi(v).$$

Clearly, an *s*-increasing function is also *w*-increasing. Note that, for $n = m = 1$, s and w-increasing functions are both increasing.

A simple example of a *w*-increasing function, which is not *s*-increasing, in $C = \mathbb{R}^m$ is $\Phi \colon \mathbb{R}^m \to \mathbb{R}$, given by $\Phi(u) := \max_{i=1,\ldots,m}\{u_i\}$. A general example of *w*-increasing function which is not *s*-increasing is the following: $\Phi(u) := \sum_{i=1}^{m} \phi_i(u_i)$, where $\phi_i \colon \mathbb{R} \to \mathbb{R}$ is nondecreasing for $i = 1, \ldots, m$ and there exists at least one index i_0 such that ϕ_{i_0} is increasing and not all ϕ_i's are increasing functions.

An example of an *s*-increasing bounded function is $\Phi \colon \mathbb{R}^m \to \mathbb{R}$, given by $\Phi(u) := \sum_{i=1}^{m} \arctan(u_i)$. More generally, $\Phi(u) := \sum_{i=1}^{m} \phi_i(u_i)$, where $\phi_i \colon \mathbb{R} \to \mathbb{R}$ is increasing for $i = 1, \ldots, m$, is an *s*-increasing function (not necessarily bounded). For other examples of such auxiliary functions, we refer to [11, Section 3].

It is also easy to see that if the function Φ is continuous and *w*-increasing in C, then we have

$$u \leq v \quad \Rightarrow \quad \Phi(u) \leq \Phi(v) \quad \text{for any } u, v \in C. \tag{4}$$

For the sake of completeness, we state a simple result for these type of monotonic functions which relates optimality for the scalar-valued problem $\min_{x \in D} \Phi(f(x))$ to weak Pareto and Pareto optimality for the vector-valued problem (1). Note that the continuity (of f or of Φ) is not needed.

Lemma 3.1. (a) If Φ is a *w*-increasing function in $f(D)$ and $x^* \in \operatorname{argmin}_{x \in D} \Phi(f(x))$, then x^* is a weak-Pareto optimal solution for problem (1).

(b) If Φ is an *s*-increasing function in $f(D)$ and $x^* \in \operatorname{argmin}_{x \in D} \Phi(f(x))$, then x^* is a Pareto optimal solution for problem (1).

Proof. The results follow from [17, Lemmas 5.14 and 5.24]. $\qquad \square$

3.2 The algorithm

Now, we define the *multicriteria barrier method* for solving problem (1), an extension of the classical procedure for scalar-valued constrained optimization.

Algorithm 3.2. The multicriteria barrier method (MBM):

Take $B\colon D^o \to \mathbb{R}^m_+$ a multiobjective barrier for $D \subseteq \mathbb{R}^n$, $\Phi\colon \mathbb{R}^m \to \mathbb{R}$ a w-increasing continuous auxiliary function, and $\{\tau_k\} \subset \mathbb{R}_{++} = (0, +\infty)$ a sequence such that $\tau_{k+1} < \tau_k$ for all k and $\tau_k \to 0$.

The method is iterative and generates a sequence $\{x^k\} \subset \mathbb{R}^n$ by solving the following scalar-valued optimization problems:

$$x^k \in \operatorname*{argmin}_{x \in D^o} \Phi\Big(f(x) + \tau_k B(x) \Big), \quad k = 1, 2, \dots \tag{5}$$

The strong version of the method is formally identical to the weak one, but with a continuous s-increasing auxiliary function $\Phi\colon \mathbb{R}^m \to \mathbb{R}$.

From now on, for problem (1), understood in both the weak Pareto or Pareto senses, we assume that

$$\inf_{x \in D^o} \Phi(f(x)) = \inf_{x \in D} \Phi(f(x)) \in \mathbb{R}, \tag{6}$$

where, according to the case we seek Pareto or weak Pareto optima, Φ is, respectively, a continuous strongly or weakly increasing function. Observe that, in any case, by virtue of the monotonic behavior of Φ, this condition is an extension of the classical hypotheses needed in the scalar case (i.e., $m = 1$), which is given by $\inf_{x \in D^o} f(x) = \inf_{x \in D} f(x)) > -\infty$.

Let us now list some comments and observations concerning both versions of multicriteria barrier method.

1. In order to guarantee the existence of x^k for all k, we need some additional assumptions. One possibility is to apply the method for any functions f, Φ and B such that $x \mapsto \Phi\big(f(x) + \tau_k B(x) \big)$ is "coercive" in D^o for all k, so it diverges to $+\infty$ whenever the argument approaches more and more to the boundary of D. In such case, Weierstrass theorem shows that $\operatorname{argmin}_{x \in D^o} \Phi\big(f(x) + \tau_k B(x) \big) \neq \emptyset$ for all k.

2. Observe that whenever the data f and g are smooth, it is convenient to choose B and Φ as smooth functions too, so that the subproblems (5) can be solved by means of first-order scalar methods.

3. When $m = 1$, any auxiliary function $\Phi\colon \mathbb{R} \to \mathbb{R}$ (weak or strong) is increasing, so, in this case, the update (5) generates the same sequence as the classical scalar-valued barrier method.

4. In the scalar-valued setting, the behavior of the barrier B close to the border of the feasible region seems a natural assumption for the welldefinedness of the method. Moreover, it is natural to expect that the presence of the barrier forces any unconstrained procedure used to compute the iterates to not choose x^k too close to the boundary of D. In the multiobjective case, due to the presence of Φ, which may be bounded, this is not so obvious; however, taking an auxiliary function Φ, which satisfies the condition $u_i \le \Phi(u)$ for all $u \in \mathbb{R}^m$ and all $i = 1, 2, \dots, m$ (see examples of w or s-type functions in [11]), it

becomes clear that one can also expect this kind of behavior for any unconstrained procedure used to compute x^k. Anyway, for any m, these subroutines must not be completely unconstrained, because they have to guarantee the feasibility of all x^k.

5. Let ∂D be the boundary of set D. Since, in general, $\operatorname{argmin}_{x \in D} \Phi(f(x)) \subset \partial D$, at least in the scalar-valued case, it is natural to ask the parameter sequence $\tau_k \to 0$, in order to compensate the fact that $B(z^k) \to +\infty$ for sequences $\{z^k\}$ which approach ∂D. In the multiobjective case, since Φ may be bounded, this condition does not seem so natural; nevertheless, if, again, one chooses an auxiliary function Φ such that $u_i \leq \Phi(u)$ for all $u \in \mathbb{R}^m$ and all $i = 1, 2, \ldots, m$, we see that this prescription is also very reasonable.

6. This method inherits some features of its real-valued counterpart. Firstly, we mention a drawback of MBM, namely that it does not have any kind of "memory", i.e., the former iterate needs not to be used to compute the current one. However, it seems natural to obtain x^k by initializing the subroutine used to solve subproblem (5) with the previous iterate x^{k-1}. Secondly, a benefit of applying MBM is to replace a constrained vector-valued problem with a sequence of unconstrained scalar-valued ones.

7. Note that when $\Phi(u) = \max\{u_i\}$ and $B = (\hat{B}, \ldots, \hat{B})^\top$, where $\hat{B}$ is a real-valued barrier function for D, the use of MBM to (1) reduces to the application of the classical scalar barrier method to problem $\min_{x \in D} \max_{1 \leq i \leq m} \{f_i(x)\}$. One may then ask why we should use MBM. An answer to this question is that, MBM has more degrees of freedom: we do not always need to choose that specific Φ and neither a barrier of the type $B = (\hat{B}, \ldots, \hat{B})^\top$. Moreover, the choice of other barriers may be very useful whenever some of the given data components are in quite different scales. Indeed, in order to compensate this drawback, MBM can be implemented with appropriate barrier components B_i. Also, as we have already mentioned in item 2, when (1) is smooth, a max type auxiliary function is definitely not the best choice for Φ, since it will produce nonsmooth subproblems.

8. As in other extensions of classical scalar methods to the vectorial setting, under certain regularity conditions, all iterates of MBM are implicitly obtained by the application of the corresponding real-valued algorithm to a certain weighted scalarization. Let us see this assertion. Assume that f and $B = (\hat{B}, \ldots, \hat{B})^\top$ are differentiable $\mathbb{R}_+^m$-convex functions (i.e., f_j and $B_j = \hat{B}$ are convex for all j), with $\hat{B}$ a scalar-valued penalty for D. Let Φ be defined by $\Phi(u) = \max_{i=1,\ldots,m}\{u_i\}$ for all $u \in \mathbb{R}^m$. It is a well-known fact (see, for example, [16]) that x^k is an unconstrained minimizer of $\max_{i=1,\ldots,m}\{f_i(x) + \tau_k B_i(x)\}$ if and only if there exist positive scalars $\alpha_j = \alpha_j^{(k)}$, $j \in I(x^k)$, where

$$I(x^k) = \left\{ j \colon f_j(x^k) + \tau_k B_j(x^k) = \max_{i=1,\ldots,m} \{f_i(x^k) + \tau_k B_i(x^k)\} \right\},$$

such that

$$\sum_{j \in I(x^k)} \alpha_j = 1 \quad \text{and} \quad \sum_{i \in I(x^k)} \alpha_i \Big(\nabla f_i(x^k) + \tau_k \nabla B_i(x^k) \Big) = 0.$$

Taking $\alpha_j = 0$ for all $j \in \{1, \ldots, m\} \setminus I(x^k)$, we have

$$\sum_{i=1}^{m} \alpha_i \left(\nabla f_i(x^k) + \tau_k \nabla B_i(x^k) \right) = 0. \tag{7}$$

But, under our assumptions, this equality is also a necessary and sufficient optimality condition for the penalized scalar-valued convex function

$$x \mapsto \sum_{i=1}^{m} \alpha_i \left(f_i(x) + \tau_k B_i(x) \right) = \langle \alpha, f(x) \rangle + \tau_k \hat{B}(x),$$

where we used that $B_i = \hat{B}$ for $i = 1, \ldots, m$, $\alpha := (\alpha_1, \ldots, \alpha_m)^\top$ and $\sum_{i=1}^{m} \alpha_i = 1$. That is to say,

$$x^k \in \operatorname*{argmin}_{x \in \mathbb{R}^n} \langle \alpha, f(x) \rangle + \tau_k \hat{B}(x),$$

which in turn means that $\{x^k\}$ can be obtained via the application of the classical (scalar) barrier method to the real-valued function $x \mapsto \langle \alpha, f(x) \rangle$, a weighted scalarization of the vector-valued objective f, with weighting vector given by $\alpha = \alpha^k \in \mathbb{R}_+^m$, using the scalar-valued barrier $\hat{B}(x)$ for D and $\{\tau_k\}$ as the parameter sequence. Of course, a priori, we do not know the nonnegative weights $\alpha_1, \ldots, \alpha_m$ which add up one and satisfy the Lagrangian equation (7).

4 Convergence analysis of MBM

We begin this section by showing some of elementary properties of sequences produced by both versions of MBM which will be needed in the sequel.

Lemma 4.1. *Let $\{x^k\} \subset \mathbb{R}^n$ be a sequence generated by the weak or strong version of MBM, implemented with a barrier function $B \colon D^o \to \mathbb{R}_+^m$, a parameters sequence $\{\tau_k\} \subset \mathbb{R}_{++}$ and an auxiliary function $\Phi \colon \mathbb{R}^m \to \mathbb{R}$. Define $\Phi^* := \inf_{x \in D} \Phi(f(x))$ and $\Phi_k := \Phi(f(x^k) + \tau_k B(x^k))$ for all k. Then, the following statements hold:*

(a) For all $k = 1, 2, \ldots,$ we have

$$\Phi^* \le \Phi_{k+1} \le \Phi_k.$$

(b) There exists $\eta \in \mathbb{R}$ such that

$$\lim_{k \to \infty} \Phi_k = \eta.$$

In particular, $\Phi^* \le \eta$.

Proof. (a) Using the definition of Φ_k and the fact that $x^k \in D^o$ for all k, we obtain

$$
\begin{aligned}
\Phi^* &= \inf_{x \in D} \Phi(f(x)) \\
&= \inf_{x \in D^o} \Phi(f(x)) \\
&\leq \Phi(f(x^{k+1})) \\
&\leq \Phi(f(x^{k+1}) + \tau_{k+1} B(x^{k+1})) \\
&= \Phi_{k+1} \\
&= \min_{x \in D^o} \Phi\big(f(x) + \tau_{k+1} B(x)\big) \\
&\leq \Phi\big(f(x^k) + \tau_{k+1} B(x^k)\big) \\
&\leq \Phi\big(f(x^k) + \tau_k B(x^k)\big) \\
&= \Phi_k,
\end{aligned}
$$

where the second equality follows from (6), the second and the last inequalities follow from the facts that Φ is w-increasing, $0 \leq B(x)$ for any x, $0 < \tau_{k+1} < \tau_k$ for all k and (4). Note that this proof holds for Φ weakly or strongly increasing.

(b) By item (a) and (6), we have that $\{\Phi_k\} \subset \mathbb{R}$ is a nonincreasing bounded from below sequence, so, as $k \to \infty$, it converges to some $\eta \in \mathbb{R}$.

$\square$

Now, we state and prove an extension of a classical convergence result for real-valued optimization, which establishes that accumulation points, if any, of a sequence generated by the barrier method are optima for the original constrained minimization problem.

Theorem 4.2. *Let $\{x^k\} \subset \mathbb{R}^n$ be a sequence generated by the weak or strong version of MBM, implemented with an auxiliary function $\Phi \colon \mathbb{R}^m \to \mathbb{R}$, a multiobjective barrier function $B \colon D^o \to \mathbb{R}^m_+$ and a sequence of parameters $\{\tau_k\} \subset \mathbb{R}_{++}$. If $\bar{x} \in D$ is an accumulation point of $\{x^k\}$, then $\bar{x} \in \operatorname{argmin}_{x \in D} \Phi(f(x))$. Moreover, if Φ is w-increasing, then $\bar{x}$ is a weak Pareto optimum for problem (1); if Φ is s-increasing, then $\bar{x}$ is a Pareto optimal solution for (1).*

Proof. Let us first consider the case in which MBM is implemented with a weakly-increasing auxiliary function Φ. From Lemma 4.1, there exists a real number $\eta \in \mathbb{R}$ such that $\Phi_k \to \eta$, where $\Phi_k = \Phi(f(x^k) + \tau_k B(x^k))$. As in the same lemma, call $\Phi^* = \inf_{x \in D} \Phi(f(x))$. We claim that $\Phi^* = \eta$. If not, then, again by Lemma 4.1, we have

$$
\Phi^* < \eta. \tag{8}
$$

Since, by (6), $\Phi^* = \inf_{x \in D} \Phi(f(x)) = \inf_{x \in D^o} \Phi(f(x))$ and $\Phi^* < \eta$, there exists $\tilde{x} \in D^o$ such that $\Phi^* < \Phi(f(\tilde{x})) < \eta$. Hence, from the continuity of Φ and the fact that $\tau_k \to 0$, for k large enough we have that

$$
\Phi(f(\tilde{x}) + \tau_k B(\tilde{x})) < \eta.
$$

9

From the fact that $\tilde{x} \in D^o$, it follows that $\Phi_k = \min_{x \in D^o} \Phi(f(x) + \tau_k B(x)) \leq \Phi(f(\tilde{x}) + \tau_k B(\tilde{x}))$, which combined with the above inequality leads to

$$\Phi_k < \eta$$

for k large enough, in contradiction with the results of Lemma 4.1.

Whence (8) does not hold and so

$$\Phi^* = \eta. \tag{9}$$

Now let $\{x^{k_j}\}$ be a subsequence of $\{x^k\}$ such that $\lim_{j \to \infty} x^{k_j} = \bar{x}$. Since $x^{k_j} \in D^o$ for all $j = 1, 2, \ldots$, we have $g(x^{k_j}) < 0$ for all j and, by the continuity of g, $g(\bar{x}) \leq 0$, i.e., $\bar{x} \in D$. Assume that $\bar{x} \notin \operatorname{argmin}_{x \in D} \Phi(f(x))$; from the feasibility of $\bar{x}$ and the definition of Φ^*, this means that

$$\Phi(f(\bar{x})) > \Phi^*. \tag{10}$$

From (9) we get that

$$\lim_{j \to \infty} \Phi(f(x^{k_j}) + \tau_{k_j} B(x^{k_j})) - \Phi^* = \lim_{j \to \infty} \Phi_{k_j} - \Phi^* = \eta - \Phi^* = 0. \tag{11}$$

But, combining the fact that $\tau_{k_j} B(x^{k_j}) \geq 0$ for all $j = 1, 2, \ldots$ with (4), and using the continuity of $\Phi(f(\cdot))$, the definition of Φ^*, as well as (10), we also get

$$\lim_{j \to \infty} \Phi(f(x^{k_j}) + \tau_{k_j} B(x^{k_j})) - \Phi^* \geq \lim_{j \to \infty} \Phi(f(x^{k_j})) - \Phi^* = \Phi(f(\bar{x})) - \Phi^* > 0, \tag{12}$$

which contradicts (11).

Hence (10) is not true and so $\Phi(f(\bar{x})) = \Phi^* = \inf_{x \in D} \Phi(f(x))$, that is to say $\bar{x} \in \operatorname{argmin}_{x \in D} \Phi(f(x))$ and the result follows from item (a) of Lemma 3.1. The proof for the case in which MBM is implemented with an s-increasing Φ is almost identical, and it concludes using item (b) of Lemma 3.1. $\square$

The following result establishes that both versions of MBM are convergent whenever $\Phi \circ f := \Phi(f(\cdot))$ has a strict minimizer in D, which happens, for instance, if this composition is strictly convex and coercive.

Corollary 4.3. *Let $\{x^k\} \subset \mathbb{R}^n$ be generated by the weak or strong version of MBM, with auxiliary function $\Phi \colon \mathbb{R}^m \to \mathbb{R}$ such that $x \mapsto \Phi(f(x))$ has a strict minimizer $\bar{x}$ in D, i.e., $\operatorname{argmin}_{x \in D} \Phi(f(x)) = \{\bar{x}\}$. If we assume that $\{x^k\}$ has an accumulation point, then $x^k \to \bar{x}$. Furthermore, if Φ is w-increasing, then $\bar{x}$ is a weak Pareto optimal solution for (1); if Φ is s-increasing, then $\bar{x}$ is a Pareto optimum for (1).*

Proof. Let us just prove the case in which $\{x^k\}$ is generated by the weak version of MBM. From Theorem 4.2 and the strict optimality of $\bar{x} \in D$, any subsequential limit of $\{x^k\}$ is equal to $\bar{x}$ and so $x^k \to \bar{x}$, a weak Pareto optimal solution for (1). The convergence result for the strong version of MBM also follows from Theorem 4.2. $\square$

4.1 Local implementation of MBM and its convergence

We now sketch a practical implementation of both, strong and weak, versions of MBM. Take $B\colon \mathbb{R}^n \to \mathbb{R}^m_+$, a barrier for D, a w-increasing continuous auxiliary function $\Phi\colon \mathbb{R}^m \to \mathbb{R}$, a decreasingly convergent to zero sequence of positive penalty parameters $\{\tau_k\}$ and $V \subset \mathbb{R}^n$, a compact set with nonempty interior (e.g., a closed ball with positive radius). Let

$$x^k \in \operatorname*{argmin}_{x \in D^o \cap \mathrm{int}(V)} \Phi\big(f(x) + \tau_k B(x)\big), \quad k = 1, 2, \dots \tag{13}$$

where $\mathrm{int}(V)$ stands for the topological interior of V.

The strong version of this local implementation of the method is formally identical to the weak one, but with a s-increasing continuous auxiliary function Φ instead of a w-increasing one.

From now on, we assume that

$$\inf_{x \in D^o \cap \mathrm{int}(V)} \Phi(f(x)) = \inf_{x \in D \cap V} \Phi(f(x)) \in \mathbb{R}, \tag{14}$$

where Φ is a strongly or weakly-increasing continuous function if we are, respectively, seeking Pareto or weak Pareto optima.

Assuming the existence of a strict local minimizer of $\Phi(f(\cdot))$ within D (i.e., a feasible $\bar{x}$ such that $\Phi(f(\bar{x})) < \Phi(f(x))$ for any feasible x in a vicinity of $\bar{x}$), we will prove that both the weak and the strong versions of the local MBM are fully convergent to a weak Pareto and a Pareto optimal solutions, respectively.

Theorem 4.4. *Suppose that $\Phi\colon \mathbb{R}^m \to \mathbb{R}$ is a weakly or strongly increasing continuous function and that $\bar{x} \in D$ is a strict local minimizer of $\Phi(f(\cdot))$ in D, say $\bar{x} = \operatorname{argmin}_{x \in D \cap U} \Phi\big(f(x)\big)$, for some vicinity $U \subset \mathbb{R}^n$ of $\bar{x}$. Let $\{x^k\} \subset \mathbb{R}^n$ be a sequence generated by the local MBM version, implemented with a barrier function $B\colon \mathbb{R}^n \to \mathbb{R}^m_+$, a parameters sequence $\{\tau_k\} \subset \mathbb{R}_{++}$, the auxiliary function Φ and V, a compact subset of U with nonempty interior. Then $x^k \to \bar{x}$ and, moreover, $\bar{x}$ is a weak Pareto optimal solution for problem (1). If Φ is a s-increasing auxiliary function, then $x^k \to \bar{x}$ and $\bar{x}$ is a Pareto optimum for problem (1).*

Proof. Since V is compact, the sequence $\{x^k\} \subset V$ has an accumulation point $\tilde{x} \in V$. Note that condition (14) allows us to prove a local version of Lemma 4.1, so as in the proof of Theorem 4.2, we see that $\tilde{x} \in \operatorname{argmin}_{x \in D \cap V} \Phi(f(x))$. Since $\bar{x}$ is the unique minimizer of $\Phi(f(\cdot))$ within $D \cap U$ and $V \subset U$, we conclude that $\tilde{x} = \bar{x}$; therefore, the unique accumulation point of $\{x^k\}$ is $\bar{x}$ and the proof is complete. Now, from Lemma 3.1, $\bar{x}$ is a weak Pareto or a Pareto optimum, whether Φ is, respectively, a w-increasing or a s-increasing auxiliary function. $\qquad\square$

Local convergence may also be true under a weaker condition than the existence of a strict local minimizer, namely whenever $S_U := \operatorname{argmin}_{x \in D \cap U} \Phi(f(x))$ is an isolated set of $S := \bigcup_{W \subset \mathbb{R}^n} S_W$, where $S_W := \{\bar{x} \in \mathbb{R}^n\colon \Phi(f(\bar{x})) = \inf_{x \in D \cap W} \Phi(f(x))\}$, with W compact subset of $\mathbb{R}^n$ such that $\mathrm{int}(W) \neq \emptyset$, which means that there exists a closed set $G \subset \mathbb{R}^n$, such that $\emptyset \neq S_U \subset \mathrm{int}(G)$ and $G \setminus S_U \subset \mathbb{R}^n \setminus S$.

5 MBM applications

Let us now show a very simple *ad hoc* instance of problem (1) where the weighting method, a quite popular strategy for solving multicriteria problems, fails to furnish optima in an arbitrary large set of weights, while MBM provides an optimal solution.

Recall that the weighting method consists of minimizing a convex combination of the objective components on the feasible set. So for $m = 2$, it just requires a single weight $\alpha \in [0, 1]$ (the other one is $1 - \alpha$). For a given tolerance $\varepsilon \in (0, 1)$, we will show that, when applied to (1) for a certain $f = (f_1, f_2)^\top$ and a particular constraint set D, this method fails to have optimal solutions in a set of parameters whose length is greater than $1 - \varepsilon$.

Example 5.1. Let $\varepsilon \in (0, 1)$. Consider $f \colon \mathbb{R} \to \mathbb{R}^2$ be given by $f(t) := (t, -at)^\top$, where $a > \frac{1-\varepsilon}{\varepsilon}$, with $D := \{t \in \mathbb{R} \colon t \geq 0\}$. So, the multiobjective problem can be written as

$$\min_{-t \leq 0} \ f(t). \tag{15}$$

The set of Pareto (weak or not) optima is $S = D = [0, +\infty)$. The weighting method prescribes to solve the following scalar optimization problem:

$$\min_{t \geq 0} \ \big\langle (\alpha, 1 - \alpha), (t, -at) \big\rangle = \min_{t \geq 0} \ [\alpha(1 + a) - a]t.$$

Clearly, the above problem has an optimal solution if, and only if, $\alpha(1 + a) - a \geq 0$, which, in turn, is equivalent to $\alpha \geq \frac{a}{1+a}$. This means, that the weighting method works properly only in an interval with length $1 - \frac{a}{1+a} = \frac{1}{1+a}$.

Therefore, since $a > \frac{1-\varepsilon}{\varepsilon}$, the weighting method gives us an optimum in an interval of weights of length smaller than $\varepsilon > 0$ or, in other words, it fails taking the parameter α in a subinterval of $[0, 1]$ with length greater than $1 - \varepsilon$.

On the other hand, if we implement MBM with $B \colon D^o \to \mathbb{R}_+^2$, given by $B(t) := (\frac{1}{t}, \frac{1}{t})^\top$, where $D^o = \{t \in \mathbb{R} \colon t > 0\}$, $\tau_k = \frac{1}{k}$ for all $k \in \mathbb{N}$ and $\Phi \colon \mathbb{R}^2 \to \mathbb{R}$, given by $\Phi(u) := \max\{u_1, u_2\}$, it is easy to see that

$$t_k := k^{-1/2} \in \operatorname*{argmin}_{t \in D^o} \ \Phi\big(f(t) + \tau_k B(t)\big) = \operatorname*{argmin}_{t \in D^o} \max\left\{t + \frac{1}{kt}, -at + \frac{1}{kt}\right\} = t + \frac{1}{kt}.$$

Note that $\Phi(f(t)) = \max\{t, -at\} = t$ for all $t \geq 0$, so we have that $\operatorname{argmin}_{t \in D} \Phi(f(t)) = \{0\}$. Since $-\infty < 0 = \inf_{t \in D} \Phi(f(t)) = \inf_{t \in D^o} \Phi(f(t))$ and also $t_k \to 0$, by Corollary 4.3, we have that MBM furnishes $t^* = 0$, an optimal point for problem (15).

We now exhibit a very simple instance of problem (1), for which there exists a family of auxiliary functions $\{\Phi_\omega\}_{\omega \in \Omega}$, $\Omega \subset \mathbb{R}^m$, such that, using local MBM implemented with these functions, any multiobjective barrier function for the feasible set and any positive and decreasing to zero parameters sequence, by varying $\omega \in \Omega$, we can retrieve the whole optimal set.

First, observe that, for any $\omega \in \mathbb{R}^m$, the function $\Phi_\omega \colon \mathbb{R}^m \to \mathbb{R}$, given by $\Phi_\omega(u) := \max_{i=1,\ldots,m}\{u_i + \omega_i\}$ is w-increasing and continuous, so it can be used as an auxiliary function.

Example 5.2. Consider $n = 1$, $m = 2$, $D = [-2, +\infty)$ and $f \colon \mathbb{R} \to \mathbb{R}^2$ defined by $f_1(t) := t^2 + 1$, $f_2(t) := t^2 - 2t + 1$. Clearly, in the interval $[0, 1]$, whenever f_2 decreases, f_1 increases and, in D, this happens only in this interval; that is to say, $[0, 1]$ is the optimal set for problem (1) with these data.

We will apply MBM with $\Phi_\omega(u) = \max_{i=1,2}\{u_i + \omega_i\}$, $\omega \in \mathbb{R}^2$, a continuous w-increasing function. First of all, note that $f_1(t) \leq f_2(t)$ if and only if $t \leq 0$, so $t \mapsto \Phi_\omega(f(t))$, with $\omega = (0, 0)^\top$, has a strict minimizer at $t = 0$. Now, let us investigate $\operatorname{argmin}_{t \in D} \Phi_\omega(f(t))$ not just for $\omega = (0, 0)^\top$, but for $\omega_\alpha := (\alpha, 0)^\top$, with $\alpha \in [-2, 0]$.

It is easy to see that $x \mapsto \Phi_{\omega_\alpha}(f(x))$ has a unique minimizer in $[0, 1]$ at the sole point $t_\alpha \in [0, 1]$ where $f_1(t_\alpha) + \alpha = f_2(t_\alpha)$, that is to say $t_\alpha = -\frac{\alpha}{2}$. Therefore,

$$\bigcup_{-2 \leq \alpha \leq 0} \operatorname{argmin}_{t \in D} \Phi_{\omega_\alpha}(f(t)) = [0, 1].$$

Applying Theorem 4.4, with the auxiliary function Φ_{ω_α}, where $\omega_\alpha \in \Omega := [-2, 0] \times \{0\}$, any multiobjective barrier function B for D, as well as any $V_\alpha \subset \mathbb{R}$ compact vicinity of t_α and any parameter sequence $\{\tau_k\} \subset \mathbb{R}_{++}$ such that $\tau_{k+1} < \tau_k$ and $\tau_k \to 0$, the generated sequence $\{t^k\}$ converges to t_α. This means that, using MBM with all auxiliary functions of the family $\{\Phi_{\omega_\alpha}\}_{\alpha \in [-2,0]}$, we retrieve the whole optimal set of the multiobjective problem $\min_{-t \leq 2} f(t) = (t^2 + 1, t^2 - 2t + 1)^\top$.

Of course, this is also an *ad hoc* example, but it may be useful in order to investigate when do we have auxiliary functions families such that by varying the parameters we can obtain the whole optimal set by means of the sequences produced by MBM.

6 Final remarks

In this work we developed an extension for the vector-valued setting of the classical internal penalty method for single optimization. As expected, the convergence results are generalizations of those which hold when the method is applied to scalar problems. For future research, we leave the study of conditions which guarantee the existence of auxiliary functions families such that by varying the parameters, the whole Pareto (weak or not) frontier can be generated applying MBM implemented with all those functions. Example 5.1 together with Theorem 4.4 may shed some light on that matter. Indeed, the fact that MBM only converges to minimizers of the scalar representations induced by the auxiliary functions suggest that this apparent drawback of the method can be useful in order to study necessary or sufficient conditions for the existence of such families. Another interesting thing to analyze in the multicriteria setting is the developing of mixed interior-exterior penalty methods. Finally, the existence of a generalization of MBM to the vector optimization case, or even to the variable order setting, may deserve some attention.

References

[1] M. S. Bazaraa, Sherali H. D., and C. M. Shetty. *Nonlinear Programming: Theory and Algorithms.* John Wiley & Sons, Inc, New York, 3 edition, 2006.

[2] D. P. Bertsekas. *Nonlinear Programming.* Athena Scientific, 3 edition, 2016.

[3] H. Bonnel, A. N. Iusem, and B. F. Svaiter. Proximal methods in vector optimization. *SIAM Journal on Optimization*, 15(4):953–970, 2005.

[4] E. Carrizosa and J. B. G. Frenk. Dominating sets for convex functions with some applications. *Journal of Optimization Theory and Applications*, 96(2):281–295, 1998.

[5] H. Eschenauer, J. Koski, and A. Osyczka. *Multicriteria Design Optimization.* Springer, Berlin, 1990.

[6] J. Fliege, L. M. Graña Drummond, and B. F. Svaiter. Newton's method for multiobjective optimization. *SIAM Journal on Optimization*, 20(2):602–626, 2009.

[7] J. Fliege and B. F. Svaiter. Steepest descent methods for multicriteria optimization. *Mathematical Methods of Operations Research*, 51(3):479–494, 2000.

[8] Y. Fu and U.M. Diwekar. An efficient sampling approach to multiobjective optimization. *Annals of Operations Research*, 132(1-4):109–134, 2004.

[9] E. H. Fukuda and L. M. Graña Drummond. On the convergence of the projected gradient method for vector optimization. *Optimization*, 60(8-9):1009–1021, 2011.

[10] E. H. Fukuda and L. M. Graña Drummond. A survey on multiobjective descent methods. *Pesquisa Operacional*, 34:585–620, 2014.

[11] E. H. Fukuda, L. M. Graña Drummond, and F. M. P Raupp. An external penalty-type method for multicriteria. *TOP*, 24:493–513, 2016.

[12] L. M. Graña Drummond and A. N. Iusem. A projected gradient method for vector optimization problems. *Computational Optimization and Applications*, 28(1):5–29, 2004.

[13] L. M. Graña Drummond, F. M. P. Raupp, and B. F. Svaiter. A quadratically convergent Newton method for vector optimization. *Optimization*, 63(5):661–677, 2014.

[14] L. M. Graña Drummond and B. F. Svaiter. A steepest descent method for vector optimization. *Journal of Computational and Applied Mathematics*, 175(2):395–414, 2005.

[15] M. Gravel, J. R Martel, W. Price, and R. Tremblay. A multicriterion view of optimal ressource allocation in job-shop production. *European Journal of Operations Research*, 61:230–244, 1992.

[16] J. B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of Convex Analysis*. Springer Verlag, Heidelberg, 2001.

[17] J. Jahn. *Vector Optimization – Theory, Applications and Extensions*. Springer, Erlangen, 2003.

[18] T. M. Leschine, H. Wallenius, and W.A. Verdini. Interactive multiobjective analysis and assimilative capacity-based ocean disposal decisions. *European Journal of Operational Research*, 56:278–289, 1992.

[19] Y. Sun, D. W. K. Ng, and R. Schober. Multi-objective optimization for power efficient full-duplex wireless communication systems. In *2015 IEEE Global Communications Conference*, pages 1–6, 2015.

[20] M. Tavana. A subjective assessment of alternative mission architectures for the human exploration of mars at NASA using multicriteria decision making. *Computational and Operations Research*, 31:1147–1164, 2004.

[21] D. J. White. Epsilon-dominating solutions in mean-variance portfolio analysis. *European Journal on Operations Research*, 105:457–466, 1998.