ORIGINAL PAPER

CrossMark

Bilevel optimization with a multiobjective problem in the lower level

Roberto Andreani[1] · Viviana A. Ramirez[2] · Sandra A. Santos[1] (ID) **·
Leonardo D. Secchin[3]**

Abstract

Bilevel problems model instances with a hierarchical structure. Aiming at an efficient
solution of a constrained multiobjective problem according with some pre-defined
criterion, we reformulate this semivectorial bilevel optimization problem as a clas-
sic bilevel one. This reformulation intents to encompass all the objectives, so that the
properly efficient solution set is recovered by means of a convenient weighted-sum
scalarization approach. Inexact restoration strategies potentially take advantage of
the structure of the problem under consideration, being employed as an alternative to

✉ Sandra A. Santos
 sandra@ime.unicamp.br

 Roberto Andreani
 andreani@ime.unicamp.br

 Viviana A. Ramirez
 viviana.ramirez@crub.uncoma.edu.ar

 Leonardo D. Secchin
 leonardo.secchin@ufes.br

[1] Institute of Mathematics, University of Campinas, Rua Sergio Buarque de Holanda, 651,
 Campinas, 13083-859, São Paulo, Brazil

[2] Academic Regional Center of Bariloche, National University of Comahue, Quintral 1250,
 San Carlos de Bariloche, Rio Negro, Argentina

[3] Department of Applied Mathematics, Federal University of Espírito Santo, Rodovia BR 101,
 Km 60, 29932-540, São Mateus, Espírito Santo, Brazil

the Karush-Kuhn-Tucker reformulation of the bilevel problem. Genuine multiobjective problems possess inequality constraints in their modeling, and these constraints generate theoretical and practical difficulties to our lower level problem. We handle these difficulties by means of a perturbation strategy, providing the convergence analysis, together with enlightening examples and illustrative numerical tests.

Keywords Bilevel optimization · Inexact restoration · Multiobjective optimization · KKT reformulation · Numerical experiments

Mathematics Subject Classification (2010) 90C29 · 65K05 · 49M37

1 Introduction

This work addresses a bilevel mathematical programming problem in which the lower level problem is a multiobjective one, the so-called *semivectorial bilevel optimization problem* (cf. [4, 5]). An example that models this situation is as follows: given a certain decision, that one considers as acceptable, what is the closest and feasible decision that takes into consideration other objectives, also of interest? The general problem is formulated below

$$\min F(x) \text{ such that } x \text{ belongs to the set of efficient solutions of problem MOP} \quad (1)$$

where MOP is the following multiobjective problem

$$
\begin{aligned}
\min \ & \{f_1(x), \ldots, f_p(x)\} \\
\text{s.t.} \quad & h(x) = 0 \\
& g(x) \leq 0 \\
& x \in X.
\end{aligned}
\quad (2)
$$

We assume that $X \subset \mathbb{R}^n$ is compact, $F : \mathbb{R}^n \to \mathbb{R}$, $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, p$, $h : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^n \to \mathbb{R}^q$ are twice continuously differentiable functions.

As (1)–(2) do not have the classic mathematical programming problem format, we start by considering a weighted-sum scalarization problem associated with (2) as follows. Given $w_i \geq 0$, $i = 1, \ldots, p$, such that $\sum_{i=1}^p w_i = 1$, let

$$
\begin{aligned}
\min_x \ & \sum_{i=1}^p w_i f_i(x) \\
\text{s.t.} \quad & h(x) = 0 \\
& g(x) \leq 0 \\
& x \in X.
\end{aligned}
\quad (3)
$$

Notice that the set of weak Pareto optimal solutions of problem (2) equals the union of the set of optimal solutions problem (3) for $w \in \mathbb{R}^p$ such that $\sum_{i=1}^p w_i = 1$

and $w_i \geq 0$ for all i (see, e.g. [10]). Hence, the following problem can be used to replace (1)–(2)

$$
\begin{aligned}
&\min_{x,w} \ F(x) \\
&\text{s.t.} \ \ \sum_{i=1}^{p} w_i = 1 \\
&\qquad w_i \geq 0, i = 1, \dots, p \\
&\qquad x \in \operatorname{argmin} \left\{
\begin{array}{l}
\min \ \sum_{i=1}^{p} w_i f_i(x) \\
\text{s.t.} \quad h(x) = 0 \\
\qquad\quad g(x) \leq 0 \\
\qquad\quad x \in X.
\end{array}
\right.
\end{aligned}
\tag{4}
$$

For ensuring that all the objectives play some role in the optimization process, we consider positive lower bounds $\varepsilon_i, i = 1, \dots, p$ for the weights. In the literature, such solutions are related to the concept of *properly efficient* solutions, a nomenclature that dates back to Kuhn and Tucker [20] and Geoffrion [16] (see also [2, 6]).

This notion eliminates efficient points of anomalous type, in which a small gain for one function provides a large loss for another. Moreover, for any fixed $w > 0$, local solutions of (3) are locally properly efficient points of (2) [16, Corollaries 5 and 7]. Reciprocally, in case all the functions f_i are convex and the feasible set of problem (2) is convex, then any properly efficient point of (2), in Geoffrion's sense, is a global solution of (3) for some $w > 0$ [16, Theorem 2].

Such a discussion motivates the reformulation we are actually considering for addressing problem (1)–(2), as follows

$$
\begin{aligned}
&\min_{w,x} \ F(x) \\
&\text{s.t.} \ \ \sum_{i=1}^{p} w_i = 1 \\
&\qquad w_i \geq \varepsilon_i, i = 1, \dots, p \\
&\qquad x \in \operatorname{argmin} \left\{
\begin{array}{l}
\min \ \sum_{i=1}^{p} w_i f_i(x) \\
\text{s.t.} \quad h(x) = 0 \\
\qquad\quad g(x) \leq 0 \\
\qquad\quad x \in X.
\end{array}
\right.
\end{aligned}
\tag{5}
$$

A well-known strategy for addressing problem (5) is to replace its lower level problem by the associate Karush-Kuhn-Tucker (KKT) conditions. This strategy, however, does not take into account the minimization structure inherent to the lower level problem. Alternatively, we have adopted an Inexact Restoration (IR)-based strategy in which the optimization problem is addressed in two phases, first pursuing feasibility and then optimality, keeping a certain control on the feasibility already achieved. Consequently, our approach takes advantage of the intrinsic minimization structure of the problem, particularly at the feasibility phase, so that better solutions might be expected. Moreover, at the feasibility phase of the IR strategy, the user is free to adopt the method of his choice, as long as the restored iterates fulfill some mild assumptions [21, 22].

Concerning the resolution of the original bilevel problem (1)–(2) by its weighted reformulation (5), when, for a given vector of weights, problem (3) is convex, it is

well-known that the KKT conditions are sufficient for its optimality. Thus, we can expect, in such a case, that an optimization algorithm applied to (3) always recover properly efficient solutions of problem (2) (at least if a constraint qualification holds). In the general case, only global solutions of the weighted problem (3) are guaranteed to be properly efficient solutions of (2) [16]. Thus, optimization algorithms may fail to get efficient solutions of (2) when they converge to a nonglobal stationary point of the weighted problem. In this sense, we reinforce that IR strategies can benefit from treating the lower level weighted problem directly, without reformulating it, because in this case, algorithms have more chances to reach global solutions (for instance, when they use second order information as sequential quadratic programming techniques do). Nevertheless, except for convex problems, global optimal solutions are not ensured. In the worst case, the IR strategy is likely to obtain local solutions [24], and hence, properly local efficient solutions [17], whereas the KKT strategy does not distinguish minimizers from maximizers.

One should also notice that, at the optimality phase of the IR strategy, the KKT reformulation takes place as an auxiliary tool. The KKT reformulation of the lower level problem within the bilevel context, on the other hand, is a basic tool for addressing the problem, as it is the IR strategy. Along this perspective, it is worth mentioning a software environment for solving bilevel problems using inexact restoration, presented in [25].

The optimization over efficient sets has been previously analyzed, but mainly considering linear optimization, linear bilevel problems, and multicriteria optimization with polyhedral feasible sets. Benson [3] has studied the problem of optimizing a linear function over the set of efficient solutions for a vector optimization problem. The linear objective function measures the importance of, or discriminates, the efficient alternatives that are available. He established necessary and sufficient conditions for efficient and for arbitrary solutions of the underlying vector optimization problem to be optimal solutions for the problem upon analysis. By taking advantage of the equivalence between efficient points and optimal solutions of a linear programming problem, Fülöp [12] pointed out that a pure linear bilevel problem might be addressed instead. In the general nonlinear case, a similar equivalence is not expected. A related and recent work is [8], in which the authors have addressed linearly equality constrained and bound constrained multiobjective instances in the lower level by means of a flexible IR strategy.

In preliminary numerical results involving multiobjective problems with inequality constraints, we have observed that algorithms applied to the lower level problem might get trapped at undesired boundary points. Despite being legitimate KKT points, and strong regular points of the MPEC reformulation, these points belong to an active inequality constraint, with associate positive multiplier. Example 2 exhibits a typical situation. As inequality constrained problems are usually present in general multiobjective problems, and motivated by this potential inconvenient, we have devised an IR strategy in which the inequality constraints of the lower level problem are perturbed. The associated convergence analysis is provided, together with a set of illustrative numerical results.

The text is organized as follows. In Section 2, we provide an overview of the IR strategy employed in this work, with examples that highlight our contributions.

The convergence analysis is provided in Section 3. Section 4 contains the numerical results, and in Section 5, we present final remarks and future work. The test problems are described in the Appendix.

Notation. The symbols $\|\cdot\|$, $\|\cdot\|_\infty$ and $|\cdot|$ denote, respectively, the Euclidean, the supremum and an arbitrary norm. We also use the notation $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0\}$.

2 Inexact restoration for bilevel programming

For the sake of simplicity, in this section, we deal with the problem as follows:

$$\text{P: } \min_x \quad F(x)$$
$$\text{s.t. } x \in \mathscr{X}^*$$

where $F : \mathbb{R}^n \to \mathbb{R}$ and $\mathscr{X}^*$ is the set of efficient solutions of the lower level multiobjective problem:

$$\min \{f_1(x), \ldots, f_p(x)\}$$
$$\text{s.t.} \quad h(x) = 0$$
$$0 \leq x \in X$$

with $f_1, \ldots, f_p : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^m$. We assume that all functions are twice continuously differentiable.

Given tolerances $\varepsilon_i \in (0, 1)$, $i = 1, \ldots, p$, we can state P as follows:

$$\min_{w,x} F(x)$$
$$\text{s.t.} \quad w \in W \tag{6}$$
$$x \in \text{argmin} \left\{ \begin{array}{l} \min_x \ \sum_{i=1}^p w_i f_i(x) \\ \text{s.t.} \quad h(x) = 0 \\ \phantom{\text{s.t.}} \quad 0 \leq x \in X \end{array} \right\}$$

where

$$W = \left\{ w \in \mathbb{R}^p \mid \sum_{i=1}^p w_i = 1, \quad \varepsilon_i \leq w_i \leq 1 - \varepsilon_i, \quad i = 1, \ldots, p \right\}.$$

Note that W is a convex and compact set. Here, we have $X \equiv \mathbb{R}^n$, so that $x \in X$ trivially holds. Nevertheless, such a pertinence is kept because, in the general case, a bounding compact set X is required. We emphasize that the feasible set of the lower level problem of (6) was described with nonlinear equalities and simple nonegative variables for simplifying the presentation. Compared with the feasible set of problem (2), one might assume that slack variables were added to the nonlinear inequality constraints, and all variables are nonnegative, so that the technicalities of dealing with some free variables and the associated cumbersome notation are avoided.

We define the following:

$$C(w, x, \mu, \gamma) = \begin{bmatrix} \sum_{i=1}^{p} w_i \nabla_x f_i(x) + \nabla_x h(x)\mu - \gamma \\ h(x) \\ \gamma_1 x_1 \\ \vdots \\ \gamma_n x_n \end{bmatrix} \in \mathbb{R}^{2n+m} \qquad (7)$$

as in [1], where $(\mu, \gamma) \in \Delta \subset \mathbb{R}^m \times \mathbb{R}^n_+$. The KKT conditions for the lower level problem, parametrized in w, are as follows:

$$C(w, x, \mu, \gamma) = 0, \quad 0 \le x \in X, \quad \gamma \ge 0 \qquad (8)$$

(thus, Δ plays the role of the set of multipliers for this parametrized problem, which in principle, can be taken as $\mathbb{R}^m \times \mathbb{R}^n_+$). We also define the following:

$$L(w, x, \mu, \gamma, \lambda) = F(x) + C(w, x, \mu, \gamma)^T \lambda$$

for $\lambda \in \mathbb{R}^{2n+m}$. We write shortly $s = (w, x, \mu, \gamma) \in W \times X \times \Delta$, and the Jacobian of C is given by the following:

$$C'(s) = \begin{bmatrix} \nabla_x f_1 & \cdots & \nabla_x f_p & \sum_{i=1}^{p} w_i \nabla^2_{xx} f_i + \sum_{i=1}^{m} \mu_i \nabla^2_{xx} h_i & \nabla h & -I_n \\ 0 & \cdots & 0 & \nabla h^T & 0 & 0 \\ 0 & \cdots & 0 & \mathrm{diag}(\gamma) & 0 & \mathrm{diag}(x) \end{bmatrix}.$$

$$(9)$$

In the following, we adapt the inexact restoration method presented in [1] to our purposes. The parameters $\eta > 0$, $M > 0$, $\theta_{-1} \in (0, 1)$, $\delta_{\min} > 0$, $\tau_1 > 0$, $\tau_2 > 0$ are given, as well as initial approximations $s^0 \in W \times X \times \Delta$, $\lambda^0 \in \mathbb{R}^{2n+m}$, and a sequence $\{\omega^k\}$ such that $\sum_{k=0}^{\infty} \omega^k < \infty$.

The steps for obtaining $s^{k+1} = (w^{k+1}, x^{k+1}, \mu^{k+1}, \gamma^{k+1})$ and λ^{k+1} are as follows.

An interesting issue in multiobjective optimization is the effect that constraints may have on the efficient set of a problem. Contrary to what we can expect, the efficient set of a constrained multiobjective problem

$$\min\{f_1(x), \ldots, f_p(x)\} \quad \text{s.t.} \quad x \in X$$

is not necessarily composed only by the efficient points of the unconstrained problem

$$\min\{f_1(x), \ldots, f_p(x)\}$$

that lies in the feasible region X. There may be additional efficient points at the border of X. In fact, a border point such that all non-ascending paths for all objective f_i are infeasible is efficient (see Fig. 1). The next example illustrates an instance for which the optimal solution of the bilevel problem (1) is located at the border of the lower level feasible set. It is based on a problem of [27], where the authors have not addressed this issue, missing efficient points from their analysis.

Algorithm 1 Inexact restoration

1 Define $\theta_k^{\min} = \min\{1, \theta_{k-1}, \ldots, \theta_{-1}\}, \theta_k^{\text{large}} = \min\{1, \theta_k^{\min} + \omega^k\}$ and $\theta_{k,-1} = \theta_k^{\text{large}}$.

2 **(Restoration phase)** Find an approximate minimizer $\overline{x}$ and multipliers $(\overline{\mu}, \overline{\gamma}) \in \Delta$ for the problem

$$
\min_x \ \sum_{i=1}^{p} w_i^k f_i(x)
$$
$$
\text{s.t.} \quad h(x) = 0
$$
$$
0 \leq x \in X,
$$

and define $z^k = (w^k, \overline{x}, \overline{\mu}, \overline{\gamma})$.

3 **(Tangent direction)** Compute

$$
d_{\tan}^k = P_k[z^k - \eta \nabla_s L(z^k, \lambda^k)] - z^k
$$

where $P_k[\cdot]$ is the orthogonal projection on

$$
\pi_k = \{s \in W \times X \times \Delta \mid C'(z^k)(s - z^k) = 0\}.
$$

$P_k[z^k - \eta \nabla_s L(z^k, \lambda^k)]$ is a solution of the problem

$$
\min_{y \in W \times X \times \Delta} \ \frac{1}{2} \left\| y - [z^k - \eta \nabla_s L(z^k, \lambda^k)] \right\|^2
$$
$$
\text{s.t.} \quad C'(z^k)(y - z^k) = 0
$$

(this solution is unique if X and Δ are convex). If $z^k = s^k$ and $d_{\tan}^k = 0$ stop and return x^k as a solution of P. Otherwise, set $i \leftarrow 0$, and choose $\delta_{k,0} \geq \delta_{\min}$.

4 **(Minimization phase)** If $d_{\tan}^k = 0$ take $v^{k,i} = z^k$. Otherwise, take $t_{\text{break}}^{k,i} = \min\{1, \delta_{k,i}/\|d_{\tan}^k\|\}$ and find $v^{k,i} \in \pi_k$ such that, for some $0 < t \leq t_{\text{break}}^{k,i}$, we have

$$
L(v^{k,i}, \lambda^k) \leq \max\{L(z^k + td_{\tan}^k, \lambda^k), \ L(z^k, \lambda^k) - \tau_1 \delta_{k,i}, \ L(z^k, \lambda^k) - \tau_2\}
$$

with $\gamma \geq 0$ (γ of $v^{k,i}$) and $\|v^{k,i} - z^k\|_\infty \leq \delta_{k,i}$.

5 If $d_{\tan}^k = 0$ define $\lambda^{k,i} = \lambda^k$. Otherwise, take $\lambda^{k,i} \in \mathbb{R}^{2n+m}$ such that $|\lambda^{k,i}| \leq M$.

6 For all $\theta \in [0, 1]$ define

$$
Pred_{k,i}(\theta) = \theta \left[L(s^k, \lambda^k) - L(v^{k,i}, \lambda^k) - C(z^k)^T(\lambda^{k,i} - \lambda^k) \right] +
$$
$$
(1 - \theta) \left[\|C(s^k)\| - \|C(z^k)\| \right].
$$

Take $\theta_{k,i}$ as the maximum $\theta \in [0, \theta_{k,i-1}]$ that satisfies

$$
\text{Pred}_{k,i}(\theta) \geq \frac{1}{2} \left[\|C(s^k)\| - \|C(z^k)\| \right],
$$

and define $\text{Pred}_{k,i} = \text{Pred}_{k,i}(\theta_{k,i})$.

7 Calculate

$$
\text{Ared}_{k,i} = \theta_{k,i} \left[L(s^k, \lambda^k) - L(v^{k,i}, \lambda^{k,i}) \right] + (1 - \theta_{k,i}) \left[\|C(s^k)\| - \|C(v^{k,i})\| \right].
$$

If

$$
\text{Ared}_{k,i} \geq 0.1 \text{Pred}_{k,i}
$$

set

$$
s^{k+1} = v^{k,i}, \quad \lambda^{k+1} = \lambda^{k,i}, \quad \theta_k = \theta_{k,i}, \quad \delta_k = \delta_{k,i},
$$
$$
\text{Ared}_k = \text{Ared}_{k,i}, \quad \text{Pred}_k = \text{Pred}_{k,i}
$$

and finish the current kth iteration.

Otherwise, choose $\delta_{k,i+1} \in [0.1\delta_{k,i}, 0.9\delta_{k,i}]$, set $i \leftarrow i + 1$ and go to step 4.

Example 1 Let us consider the nonconvex constrained multiobjective problem F3 of [27] as follows:

$$\min \ \{(\tilde{x}_1 - 2)^2 + (\tilde{x}_2 - 1)^2 + 2, \ 9\tilde{x}_1 - (\tilde{x}_2 - 1)^2\}$$
$$\text{s.t.} \ \ \tilde{x}_1^2 + \tilde{x}_2^2 \leq 225$$
$$\tilde{x}_1 - 3\tilde{x}_2 \leq -10.$$

The search space adopted by the authors is the box $[-20, 20]^2$. In order to turn all variables into nonnegative, we make the transformation $x_i = \tilde{x}_i + 20$, $i = 1, 2$, and define the problem

$$\min \ F(x_1, x_2) = (x_1 - 20)^2 + (x_2 - 20)^2$$
$$\text{s.t.} \ \ x \in \mathscr{X}^*, \tag{10}$$

where $\mathscr{X}^*$ is the set of efficient solutions of

$$\min \ \{(x_1 - 22)^2 + (x_2 - 21)^2 + 2, \ 9x_1 - (x_2 - 21)^2 - 180\}$$
$$\text{s.t.} \ \ (x_1 - 20)^2 + (x_2 - 20)^2 \leq 225$$
$$x_1 - 3x_2 \leq -50$$
$$0 \leq x_1, x_2 \leq 40. \tag{11}$$

In the interior of the feasible region, we analyze the slopes of the two objective functions of (11) to conclude that the interval $x_1 = 17.5$, $22.5 \leq x_2 \leq 34.79$ is efficient. This interval corresponds to the intersection of the efficient set of the unconstrained problem with the feasible region of (11), and the authors of [27] claim that it is the entire efficient set of (11). However, there are other efficient points at the border of feasible set. Geometrically, these points are those in which there is no feasible non-ascending direction that decreases at least one objective function of (11) (see Fig. 1). With a simple geometric argument, it is easy to see that $\overline{x} = (19, 23)$ is the optimal solution of (10), for which $F(\overline{x}) = 10$ and the associated vector of weights $\overline{w} = (23/37, 14/37)$.

It is worth mentioning that, in our computational tests, the IR strategy was able to reach the optimal solution $\overline{x}$ located at the border of the feasible set, whereas the KKT reformulation does not converge (see the results of instance P1.1 in Section 4).

2.1 A perturbed inexact restoration strategy

We start this section with an example showing that Algorithm 1 can converge to an undesirable feasible point of the multiobjective problem. The key to the failure lies in the fact that, for certain weights, the restoration phase can return a (first-order) solution that is undesirable for the upper level problem. Despite the fact that the weighted-sum scalarized version of the lower level problem is a strictly convex quadratic problem, the convexity does not avoid this potential undesired situation.

Example 2 Consider the bidimensional problem

$$\min \ F(x_1, x_2) = x_1^2 + x_2^2$$
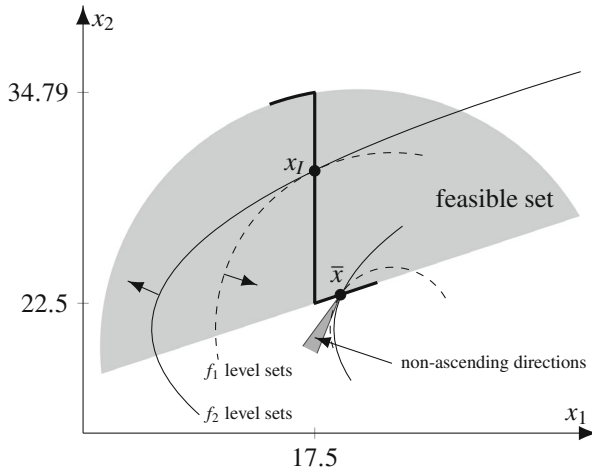$$\text{s.t.} \ \ (x_1, x_2) \in \mathscr{X}^*, \tag{12}$$

Fig. 1 Geometry of example 1. The arrows at the level sets of the objective functions indicate the corresponding decreasing directions. The vertical continuous line is the set of efficient points in which the level sets of f_1 and f_2 have the same slope. The continuous lines at the border of the feasible set represent the efficient points such that all non-ascending directions for both objectives are infeasible. In particular, x_I and $\overline{x}$ are efficient points of each one of these two types, respectively

where $\mathscr{X}^*$ is the set of efficient solutions of the lower level multiobjective problem

$$
\begin{aligned}
\min \ & \{f_1(x), \ldots, f_p(x)\} \\
\text{s.t.} \quad & (2/5)x_1 + x_2 \le 2 \\
& 5x_1 - x_2 \le 10 \\
& x_1, x_2 \ge 0,
\end{aligned}
\tag{13}
$$

and, for $1 \le i \le p$,

$$
\begin{aligned}
f_i(x_1, x_2) &= (x_1 - a_i)^2 + (x_2 - b_i)^2, \\
a_i &= 1 + 2\cos\theta_i, \\
b_i &= 1 + 2\sin\theta_i, \\
\theta_i &= \frac{2\pi}{p}(i - 1).
\end{aligned}
$$

We can take X as a bounding box that strictly contains the compact feasible region of (13), and thus we will omit it. We consider the case $p = 3$, for which $a = (3, 0, 0)^T$ and $b = (1, 1 + \sqrt{3}, 1 - \sqrt{3})^T$. It is possible to show that the efficient solutions of the unconstrained multiobjective problem such that all objective functions have the form $\|x - c\|^2$ is the convex hull of the centers cs. Then, the set of efficient solutions $\mathscr{X}^*$ of (13) (i.e., the feasible region of (12)) is the intersection of the convex hull of (a_i, b_i), $i = 1, 2, 3$, with the region delimited by the constraints of the lower level problem (13) (see Fig. 2). Hence, the origin is the unique global optimal solution of (12).
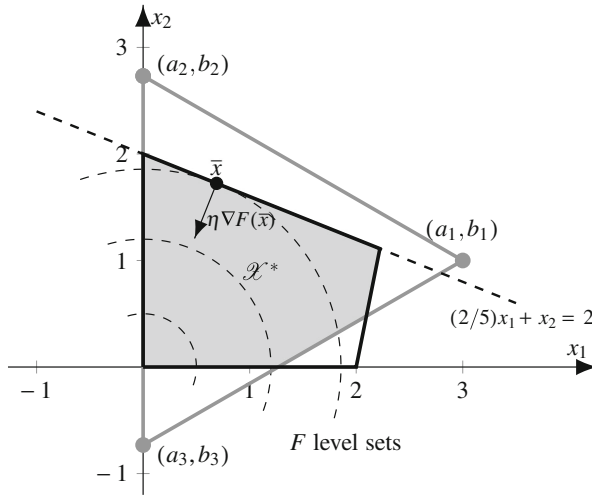
Fig. 2 Geometry of example 2. The feasible set $\mathscr{X}^*$ of (12) is the intersection of the convex hull of (a_i, b_i), $i = 1, 2, 3$, with the region delimited by the constraints of the lower level problem. The lower level KKT point $\bar{x}$ is an undesired solution for problem (12)

Adding slack variables, we write the weighted-sum scalarized version of problem (12) as follows:

$$
\begin{aligned}
\min \ &F(x_1, x_2) = x_1^2 + x_2^2 \\
\text{s.t.} \ &\textstyle\sum_{i=1}^3 w_i = 1 \\
&w_i \geq \varepsilon_i = \varepsilon, \quad i = 1, 2, 3 \\
&x \in \operatorname{argmin} \left\{
\begin{array}{l}
\min \ \sum_{i=1}^3 w_i f_i(x_1, x_2) \\
\text{s.t.} \ (2/5)x_1 + x_2 + x_3 = 2 \\
\phantom{\text{s.t.} \ } 5x_1 - x_2 + x_4 = 10 \\
\phantom{\text{s.t.} \ } x_i \geq 0, i = 1, \dots, 4.
\end{array}
\right\}.
\end{aligned}
\tag{14}
$$

We have that $\bar{x} = (20/29, 50/29, 0, 240/29)^T$ is a KKT point of the lower level problem for certain feasible weights w for problem (14). In fact, the KKT conditions of the lower level problem are as follows:

$$
\sum_{i=1}^3 2w_i \begin{bmatrix} x_1 - a_i \\ x_2 - b_i \\ 0 \\ 0 \end{bmatrix} + \mu_1 \begin{bmatrix} 2/5 \\ 1 \\ 1 \\ 0 \end{bmatrix} + \mu_2 \begin{bmatrix} 5 \\ -1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},
\tag{15}
$$

$$
(2/5)x_1 + x_2 + x_3 = 2, \tag{16}
$$

$$
5x_1 - x_2 + x_4 = 10, \tag{17}
$$

$$
\gamma_i x_i = 0, \ i = 1, \dots, 4, \tag{18}
$$

$$
x \geq 0, \ \mu \in \mathbb{R}^2, \ \gamma \in \mathbb{R}_+^4. \tag{19}
$$

At $\overline{x}$, (16)–(17) hold, $\gamma_1 = \gamma_2 = \gamma_4 = 0$ by (18), and $\mu_1 = \gamma_3 \geq 0$, $\mu_2 = \gamma_4 = 0$ by (15) and (19). Furthermore, if $w_1 \neq 20/87$, the expression (15) implies $\gamma_3 = \mu_1 > 0$, since $\mu_2 = \gamma_1 = 0$ and

$$\sum_{i=1}^{3} w_i(\overline{x}_1 - a_i) = \underbrace{(w_1 + w_2 + w_3)}_{1}\overline{x}_1 - w_1 a_1 - \underbrace{w_2 a_2}_{0} - \underbrace{w_3 a_3}_{0} = \overline{x}_1 - w_1 a_1 \neq 0.$$

Thus, by expression (15), $\overline{x}$ is a KKT point of the lower level problem for all feasible w such that

$$\begin{cases} 15w_1 - 2\sqrt{3}w_2 + 2\sqrt{3}w_3 = 2 \\ w_1 \neq 20/87. \end{cases} \qquad (20)$$

Evidently, the system (20) has feasible solutions for all $\varepsilon < 20/87 \approx 0.23$.

Now, take $\overline{w} \in \mathbb{R}^3$, feasible for (14) and satisfying (20). We will show that the inexact restoration method can converge to $\overline{x}$, an undesirable solution for problem (12) (see Fig. 2). In fact, since $\overline{x}$ is a KKT point of the lower level problem with $w = \overline{w}$, there are multipliers $(\overline{\mu}, \overline{\gamma})$ satisfying (15)–(19). As we have already mentioned, $\overline{\gamma}_3 > 0$. We denote $\overline{z} = (\overline{w}, \overline{x}, \overline{\mu}, \overline{\gamma})$. Starting Algorithm 1 with $s^0 = \overline{z}$ and $\lambda^0 = 0$, we can assume that step 2 returns $z^0 = \overline{z}$. Step 3 gives

$$\overline{\gamma}_3(x_3 - \overline{x}_3) + \overline{x}_3(\gamma_3 - \overline{\gamma}_3) = 0 \implies \overline{\gamma}_3 x_3 = 0,$$

that implies $x_3 = 0$, since $\overline{\gamma}_3 > 0$. Thus, the x-space of π_0 is contained in the line $(2/5)x_1 + x_2 = 2$ (indeed, it is exactly this line, see Fig. 2). It is easy to verify that $\nabla_x F(\overline{x})$ is orthogonal to the line $(2/5)x_1 + x_2 = 2$, and hence

$$d_{\text{tan}}^0 = P_0[\overline{z} - \eta \nabla_s L(\overline{z}, \lambda^0)] - \overline{z} = P_0[\overline{z} - \eta \nabla_s F(\overline{x})] - \overline{z} = 0.$$

Therefore, Algorithm 1 stops declaring $\overline{x}$ as a solution of problem (12).

In example 2, the x-space of π_0 at $\overline{x}$ is a border line of the lower level feasible region, which is orthogonal to $\nabla_x F(\overline{x})$, resulting in a premature convergence. If we "balance" complementarity by requiring that $\gamma_i x_i = \xi > 0$, the current point x will be in the interior of the lower level feasible region. In particular, we avoid that Algorithm 1 converges to $\overline{x}$ in such an example. Evidently, during the algorithm, ξ must vanish in order to recover the original KKT system. We call this strategy *perturbed inexact restoration*, clearly inspired by interior-point methods.

More specifically, at iteration $k \geq 1$, we make, for all i, $\gamma_i^k x_i^k = \xi^k$ where

$$\xi^k = \min\left\{\tilde{\xi}^k, \ \xi_{\max}, \ \xi_{\text{dec}} \cdot \xi^{k-1}\right\}, \qquad (21)$$

$$\tilde{\xi}^k = \frac{1}{k^2}\left(\frac{\gamma^{k-1} \cdot x^{k-1}}{n}\right), \qquad (22)$$

being $\xi_{\max} > 0$ the maximum allowed perturbation and $\xi_{\text{dec}} \in (0, 1)$ a parameter to ensure that the perturbation decreases between consecutive iterations. We also define $\xi^0 = \xi_{\max}$. The expression (22), motivated by convergence results in [11], has shown better numerical performance for our tests than the update $\tilde{\xi}^k = \gamma^{k-1} \cdot x^{k-1}/n$,

usually employed in interior-point methods. The restoration phase of Algorithm 1 turns into the resolution of

$$C_\xi(w, x, \mu, \gamma) = \begin{bmatrix} \sum_{i=1}^{p} w_i \nabla_x f_i(x) + \nabla_x h(x)\mu - \gamma \\ h(x) \\ \gamma_1 x_1 - \xi \\ \vdots \\ \gamma_n x_n - \xi \end{bmatrix} = 0, \quad 0 \le x \in X, \quad \gamma \ge 0$$

(23)

with $\xi = \xi^k$ and $w = w^k$ (actually, $(\mu, \lambda) \in \Delta$). We note that $C'_{\xi^k} = C'$, and hence the optimization problem in step 3 of Algorithm 1 remains the same, whereas steps 4 to 7 are adapted in an obvious way.

A solution of (23) can be recovered if the optimal functional value of the problem

$$\begin{aligned} \min_{x, \mu, \gamma} \quad & \left\| \sum_{i=1}^{p} w_i \nabla_x f_i(x) + \nabla_x h(x)\mu - \gamma \right\|^2 \\ \text{s.t.} \quad & h(x) = 0 \\ & x_i \gamma_i = \xi, \quad i = 1, \ldots, n \\ & 0 \le x \in X, \quad (\mu, \gamma) \in \Delta \end{aligned}$$

(24)

is zero. If we cannot obtain such a solution, we proceed with a stationary point of this problem.

The perturbed strategy avoids convergence to an undesirable point in the example. However, when the perturbation ξ is positive we do not have a feasible point (or at least, a lower level KKT point). The next example shows that if we simply adopt the same stopping criteria "$z^k = s^k$ and $d_{\tan}^k = 0$" of Algorithm 1 to our perturbed inexact restoration strategy, it may stop with a positive perturbation.

Example 3 Given a fixed $\alpha > 0$, let us consider the problem

$$\begin{aligned} \min \quad & F(x) = \tfrac{1}{2}x^2 - \sqrt{\alpha}x \\ \text{s.t.} \quad & w_1 + w_2 = 1 \\ & w_1, w_2 \ge \varepsilon \\ & x \in \operatorname{argmin} \left\{ \begin{aligned} \min \quad & w_1\left(\tfrac{1}{2}x^2\right) + w_2\left(\tfrac{1}{2}x^2\right) \\ \text{s.t.} \quad & x \ge 0. \end{aligned} \right\}. \end{aligned}$$

System (23) takes the form

$$C_\xi(w, x, \gamma) = \begin{bmatrix} (w_1 + w_2)x - \gamma \\ \gamma x - \xi \end{bmatrix} = 0, \quad 0 \le x, \quad \gamma \ge 0$$

(we may suppose that the compact set X is large enough to omit it), which yields $x = \gamma = \sqrt{\xi}$. As $C_\xi(w, \sqrt{\xi}, \sqrt{\xi}) = 0$, and we can assume that this point is the one obtained by the restoration phase (i.e., $z^0 = (1/2, 1/2, \sqrt{\alpha}, \sqrt{\alpha})$). Thus, if we initialize the method with $\xi^0 = \alpha$, $x^0 = \gamma^0 = \sqrt{\alpha}$, $w_1^0 = w_2^0 = 1/2$ (i.e., $s^0 = z^0$) and $\lambda^0 = 0$, then $\nabla_s L(z^0, \lambda^0) = (0, 0, x^0 - \sqrt{\alpha}, 0) = 0$,

$$C'_\xi(w, x, \gamma) = \begin{bmatrix} w_1 x & w_2 x & w_1 + w_2 & -1 \\ 0 & 0 & \gamma & x \end{bmatrix}$$

(see expression (9)), and the problem of step 3 in the first outer iteration takes the form

$$\min_{w_1, w_2, x, \gamma} \frac{1}{2} \left\| (w_1, w_2, x, \gamma) - \left(\frac{1}{2}, \frac{1}{2}, \sqrt{\alpha}, \sqrt{\alpha} \right) \right\|_2^2$$

$$\text{s.t.} \quad \begin{bmatrix} \frac{1}{2}\sqrt{\alpha} & \frac{1}{2}\sqrt{\alpha} & 1 & -1 \\ 0 & 0 & \sqrt{\alpha} & \sqrt{\alpha} \end{bmatrix} \begin{bmatrix} w_1 - 1/2 \\ w_2 - 1/2 \\ x - \sqrt{\alpha} \\ \gamma - \sqrt{\alpha} \end{bmatrix} = 0$$

$$w_1 + w_2 = 1$$
$$\epsilon_i \le w_i \le 1 - \epsilon_i, \quad i = 1, 2.$$

The unique solution is the point $z^0 = \left(1/2, 1/2, \sqrt{\alpha}, \sqrt{\alpha} \right)$, exactly the initial point s^0. Therefore, $z^0 = s^0$ and $d_{\tan}^0 = 0$, which implies a premature convergence to the nonfeasible point s^0 if no new conditions are imposed.

In order to guarantee a correct convergence of our perturbed inexact restoration algorithm, we reduce the perturbation ξ whenever there is lack of progress ($z^k = s^k$ and $d_{\tan}^k = 0$) with a positive ξ. In this case, we go back to the restoration phase with a smaller ξ. This ensures that the original KKT system will be recovered as desired. We summarize the new step 3 below.

3' *(Tangent direction)* Compute

$$d_{\tan}^k = P_k[z^k - \eta \nabla_s L(z^k, \lambda^k)] - z^k$$

where $P_k[\cdot]$ is the orthogonal projection on

$$\pi_k = \{ s \in W \times X \times \Delta \mid C'_{\xi^k}(z^k)(s - z^k) = 0 \}.$$

If $z^k \ne s^k$ or $d_{\tan}^k \ne 0$, do $i \leftarrow 0$, choose $\delta_{k,0} \ge \delta_{\min}$.
If $z^k = s^k$, $d_{\tan}^k = 0$ and $\xi^k > 0$, do

$$\xi^{k+1} = \xi_{\dec} \cdot \xi^k, \quad s^{k+1} = s^k, \quad \lambda^{k+1} = \lambda^k, \quad \theta_k = \theta_{k-1},$$

finish the current iteration k and go to step 2.
If $z^k = s^k$, $d_{\tan}^k = 0$ and $\xi^k = 0$, stop and return x^k as a solution of problem P.

There might be other ways to overcome an undesirable premature convergence, but certainly this choice is valid. Also, we did not observe this occurrence in our numerical tests, so possibly these cases are rare. Furthermore, a few modifications in the original inexact restoration method are made, and the already established convergence results are easily adapted to our perturbed version, as we discuss next.

3 Convergence results for the IR strategies

In this section, we stated convergence results for the plain and the perturbed inexact restoration strategies presented in the previous section. In [1], convergence results for an inexact restoration method applied to bilevel programming were established.

The authors considered the method with Lagrangian projection, i.e., with tangent direction as follows:

$$d_{\text{tan}}^k = P_k[z^k - \eta \nabla_s L(z^k, \lambda^k)] - z^k.$$

However, in standard nonlinear optimization, the convergence to points that satisfy the so-called *Approximate Gradient Projection* (AGP) [23], a sequential optimality condition, has been proved only with the tangent direction that uses the objective function [22], i.e. directions defined by the following:

$$d_{\text{tan,F}}^k = P_k[z^k - \eta \nabla_s F(z^k)] - z^k.$$

In fact, a point x^* that conforms to the AGP necessary optimality condition drives to zero the tangent direction defined by means of the objective function. In this sense, the AGP condition is closely linked to inexact restoration that uses $d_{\text{tan,F}}^k$. The use of the Lagrangian in the tangent directions was proposed in [21] and aims to avoid Marato's effect.

We have as follows:

$$d_{\text{tan,F}}^k = P_k[z^k - \eta \nabla_s L(z^k, 0)] - z^k,$$

and the inexact restoration method with $d_{\text{tan,F}}^k$ is exactly the Algorithm 1 setting $\lambda^k = 0$. Of course, in this case, step 5 can be eliminated. In the next sections, we will analyze the following variants of the inexact restoration strategy:

- IR$_L$: Inexact restoration with d_{tan}^k (Algorithm 1).
- IR$_L^\xi$: Perturbed inexact restoration with d_{tan}^k.
- IR$_F$: Inexact restoration with $d_{\text{tan,F}}^k$. This variant corresponds to Algorithm 1 setting $\lambda^k = 0$ for all k.
- IR$_F^\xi$: Perturbed inexact restoration with $d_{\text{tan,F}}^k$.

3.1 General assumptions

We make assumptions on the weighted bilevel problem (6) in order to guarantee convergence for all inexact restoration variants, following [1] and previous results for standard nonlinear programming [21, 22]. For non-perturbed variants, assumptions are easily adapted making $\xi = 0$. The hypothesis are as follows:

A1. The sets X and Δ are compact and convex.
A2. There exists $L_1 > 0$ such that, for all $(w, x), (\overline{w}, \overline{x}) \in W \times X$, $(\mu, \gamma), (\overline{\mu}, \overline{\gamma}) \in \Delta$ and $\xi \in [0, \xi_{\max}]$,

$$|C'_\xi(w, x, \mu, \gamma) - C'_\xi(\overline{w}, \overline{x}, \overline{\mu}, \overline{\gamma})| \leq L_1 |(w, x, \mu, \gamma) - (\overline{w}, \overline{x}, \overline{\mu}, \overline{\gamma})|.$$

A3. There exists $L_2 > 0$ such that, for all $x, \overline{x} \in X$,

$$|\nabla F(x) - \nabla F(\overline{x})| \leq L_2 |x - \overline{x}|.$$

A4. There exists $r \in [0, 1)$, independently of k, such that the point $z^k = (w^k, \overline{x}, \overline{\mu}, \overline{\gamma})$ obtained at the restoration phase satisfies

$$|C_{\xi^k}(z^k)| \leq r |C_{\xi^k}(s^k)|$$

where $s^k = (w^k, x^k, \mu^k, \gamma^k)$ is the current iterate. Moreover, if $C_{\xi^k}(s^k) = 0$ then $z^k = s^k$.

The first three assumptions (A1 to A3) correspond exactly to those made in [21]. As $C'_\xi = C'$ for all $\xi \geq 0$ (see expressions (7) and (9)), assumption A2 is independent of the perturbation. We observe that a way to guarantee A2 is imposing, together with A1, the Lipschitz continuity of the functions $\nabla^2 f_i$ and $\nabla^2 h_i$. This is the assumption adopted in [1].

3.2 Well-definiteness

Theorem 1 *Under assumptions A1, A2, A3, and A4, all inexact restoration algorithms are well defined.*

Proof In the $\mathrm{IR}_\mathrm{L}^\xi$ variant, if $C_{\xi^k}(s^k) \neq 0$ or ($C_{\xi^k}(s^k) = 0$ and $d_\mathrm{tan}^k \neq 0$) the proof follows Theorem 4.1 of [21]. If $C_{\xi^k}(s^k) = 0$ and $d_\mathrm{tan}^k = 0$, the current iteration finishes either with $\xi^k > 0$, going back to the restoration phase, or with $\xi^k = 0$, and the algorithm stops.

The other algorithms are also well defined, since they correspond to $\mathrm{IR}_\mathrm{L}^\xi$ setting $\lambda^k = 0$ and/or $\xi^k = 0$. □

3.3 Convergence to feasible points

If the $\mathrm{IR}_\mathrm{L}^\xi$ method stops at iteration k, then $C_{\xi^k}(s^k) = 0$ with $\xi^k = 0$, i.e., s^k is a feasible point of the KKT reformulation of the bilevel problem (6). The case of possibly infinite number of iterations are treated in the next theorem, which is an obvious adaptation of Theorems 3.4 and 3.5 of [22].

Theorem 2 *Suppose that assumptions A1, A2, A3, and A4 are valid. If the $\mathrm{IR}_\mathrm{L}^\xi$ algorithm generates an infinite sequence, then*

$$\lim_{k \to \infty} Ared_k = 0 \quad and \quad \lim_{k \to \infty} |C_{\xi^k}(s^k)| = 0.$$

The same occurs when $\xi^k = 0$ and/or $\lambda^k = 0$ for all k, that is, analogous results are valid for each variant of the inexact restoration algorithm.

Observe that a feasible limit point is obtained in the perturbed variants by Theorem 2, $\xi^k \to 0$ and assumption A1.

3.4 Convergence to optimality

As we have already mentioned, the classical inexact restoration method for standard nonlinear programming converges to AGP points [22] when the tangent directions are defined by the upper level objective function. In our case, the AGP condition is related to the vector $d_\mathrm{tan,F}(s)$ defined by

$$d_\mathrm{tan,F}(s) = P_\pi[s - \eta \nabla_s F(s)] - s$$

for any $s = (w, x, \mu, \gamma) \in W \times X \times \Delta$ (the domain of F is extended to the s-space in a natural way). Clearly, concepts based on AGP will be closely related to the IR_F and IR_F^ξ variants. First, let us consider the KKT reformulation of the weighted-sum reformulation (6) of the bilevel problem P,

$$
\begin{aligned}
\min_{w,x,\mu,\gamma} \quad & F(x) \\
\text{s.t.} \quad & C(w, x, \mu, \gamma) = 0, \quad x \geq 0 \\
& (w, x, \mu, \gamma) \in W \times X \times \Delta.
\end{aligned}
\tag{25}
$$

Theorem 2 ensures that all inexact restoration strategies converge to a feasible point of (25). We then define two optimality conditions for the bilevel programming based on the AGP condition as follows [1]:

– *Weak AGP*. A feasible point s^* of problem (25) satisfies the *weak AGP* condition of problem (6) if there exists a sequence $\{s^k\}$ converging to s^* such that $d_{\tan,F}(s^k) \to 0$.
– *Strong AGP*. A point $s^* = (w^*, x^*, \mu^*, \gamma^*)$ satisfies the *strong AGP* condition of problem (6) if it conforms to the weak AGP condition, $C(s^*) = 0$ and (w^*, x^*) is feasible for (6).

Weak AGP consists exactly in the classical AGP condition for problem (25). Strong AGP requires additionally that the primal vector (w^*, x^*) is a global optimal solution of the lower level problem of the original weighted bilevel problem (6). As the AGP condition is a legitimate necessary optimality condition [23] for standard nonlinear programming, weak and strong AGP are necessary optimality conditions for the bilevel problem (6).

A usual sufficient assumption to obtain convergence results for inexact restoration methods, incorporating the pertubation in a natural way, is the following [1, 21, 22]:

A5. There exists $\beta > 0$, independently of k, such that

$$
\|s^k - z^k\| \leq \beta |C_{\xi^k}(s^k)|.
$$

The next theorem is valid for all inexact restoration strategies, and it is analogous to Theorem 4.3 of [1]. We only state the result for the IR_L^ξ variant, since it encompasses the other ones.

Theorem 3 *Suppose that assumptions A1, A2, A3, A4, and A5 hold. If $\{s^k\}$ is an infinite sequence generated by IR_L^ξ and $\{z^k\}$ is the sequence defined at the restoration phase, then*

1. $\left| C_{\xi^k}(s^k) \right| \to 0$.
2. *There exists a limit point s^* of $\{s^k\}$.*
3. *Every limit point of $\{s^k\}$ is a feasible point of the KKT reformulation (25).*
4. *If, for all w, a global solution of the lower level problem is found then any limit point (w^*, x^*) is feasible for the weighted bilevel problem (6).*
5. *If s^* is a limit point of $\{s^k\}$, there exists an infinite set $K \subset \mathbb{N}$ such that*

$$
\lim_{k \in K} s^k = \lim_{k \in K} z^k = s^*, \quad C(s^*) = 0 \quad and \quad \lim_{k \in K} d_{tan}^k = 0.
$$

Proof The first two items follow from Theorem 2 and the assumption A1. Consequently, the third and fourth items are valid. The fifth item follows from assumption A5, from the first item and from arguments similar to those of Theorem 5.2 of [21]. □

Item 5 of Theorem 3 implies the fulfillment of AGP-type necessary optimality conditions in the IR_F and IR_F^ξ variants. We summarize this in the next result.

Corollary 1 *Suppose that assumptions A1, A2, A3, A4, and A5 hold. If $\{s^k\}$ is an infinite sequence generated by the IR_F or the IR_F^ξ method, then*

1. *Every limit point s^* is a weak AGP point.*
2. *If, for all w, a global solution of the lower level problem is found then any limit point s^* is a strong AGP point.*

3.5 Sufficient conditions for the assumptions

Compactness of X is a natural way to ensure the well-definiteness of optimization algorithms in the literature. In this sense, assumption A1 is a slightly stringent condition. Assumptions A2 and A3 are natural in the context of inexact restoration methods [21, 22] (and hence, we assume assumptions A1 to A3 from now on). However, assumptions A4 and A5 are related to sequences generated by the IR algorithms. Thus, it is interesting to establish sufficient conditions that ensure their validity and that do not depend on the progress of the methods. To verify A4, it is convenient that the lower level KKT perturbed system (23) has a solution. Particularly in the case of $\xi = 0$, it is equivalent to require that the lower level has a first-order stationary point. Following [1], we consider the next hypotheses throughout this subsection:

A6. For each $\xi \in [0, \xi_{\max}]$, every solution $s = (w, x, \mu, \gamma)$ of (23) is such that the gradients of the active lower level constraints

$$\nabla h_i(x), i = 1, \ldots, m, \quad \text{and} \quad -e_j, \text{ for } j \text{ such that } x_j = 0,$$

where e_j is the jth canonical vector of $\mathbb{R}^n$, are linearly independent.

A7. For each $\xi \in [0, \xi_{\max}]$, every solution $s = (w, x, \mu, \gamma)$ of (23) is such that the matrix

$$H(w, x, \mu) = \sum_{i=1}^{p} w_i \nabla_{xx}^2 f_i(x) + \sum_{i=1}^{m} \mu_i \nabla_{xx}^2 h_i(x)$$

is positive definite in the set

$$Z(x, \gamma) = \{d \in \mathbb{R}^n \mid \nabla h(x)^T d = 0, d_j = 0 \text{ for all } j \text{ such that } \gamma_j = 0\}.$$

A8. For $\xi = 0$, every solution $s = (w, x, \mu, \gamma)$ of (23) satisfies $x_i + \gamma_i > 0$ for all i (note that this is trivially satisfied if $\xi > 0$).

When $\xi = 0$, assumption A6 implies the well-known Linear Independence Constraint Qualification (LICQ) at every solution of the lower level problem. Thus, Δ may be naturally assumed a compact set, as in assumption A1. Assumption A8 resembles the usual strict complementarity of nonlinear programming.

We proceed analogously to the discussion made in [1]. Let us consider the matrix

$$D'(\xi, w, x, \mu, \gamma) = \begin{bmatrix} H(w, x, \mu) & \nabla h(x) & -I_n \\ \nabla h(x)^T & 0 & 0 \\ \text{diag}(\gamma) & 0 & \text{diag}(x) \end{bmatrix}.$$

We observe that D' is a submatrix of C' (or C'_ξ, see expression (9)), and does not depend on ξ.

Lemma 1 *Let $\xi \geq 0$. The matrix $D'(\xi, w, x, \mu, \gamma)$ is non-singular for any solution (w, x, μ, γ) of (23).*

Proof For $\xi = 0$, this is exactly Lemma 4.4 of [1]. For $\xi > 0$, the proof is the same since D' is constant with respect to ξ. $\qquad\square$

Let $D(\xi, w, x, \mu, \gamma) = C_\xi(w, x, \mu, \gamma)$ be defined on $[0, \xi_{\max}] \times W \times X \times \Delta$. As $D(\xi, w, x, \mu, \gamma)$ is continuous with respect to (ξ, x, μ, γ), it is possible to choose, for each $w \in W$, a solution $u_\xi(w) = (x_\xi(w), \mu_\xi(w), \gamma_\xi(w))$ of (23) such that the function $v(\xi, w) = u_\xi(w)$ is continuous on $[0, \xi_{\max}] \times W$. From now on, let the function $v(\xi, w)$ be fixed. By Lemma 1, we can define a function Υ over the set $[0, \xi_{\max}] \times W$ by

$$\Upsilon(\xi, w) = \left[D'(\xi, w, v(\xi, w)) \right]^{-1}. \tag{26}$$

Let $V(v(\xi, w), \alpha) = \{v \in [0, \xi_{\max}] \times X \times \Delta \mid \|v - v(\xi, w)\| \leq \alpha\}$.

Lemma 2 *There exist $\alpha > 0$ and $\beta > 0$ such that, for all $(\xi, w) \in [0, \xi_{\max}] \times W$, it holds $|\Upsilon(\xi, w)| \leq \beta$, and Υ coincides with the local inverse operator of $D'(\xi, w, \cdot)$ for all $v \in V(v(\xi, w), \alpha)$.*

Proof As $D'(\xi, w, v)$ is continuous on (w, v) and $v(\xi, w)$ on (ξ, w), $\Upsilon(\xi, w)$ is continuous with respect to $(\xi, w) \in [0, \xi_{\max}] \times W$. As this set is compact, there exists $\beta > 0$ such that $|\Upsilon(\xi, w)| \leq \beta$ for all $(\xi, w) \in [0, \xi_{\max}] \times W$. Thus, the first claim was proved. The second statement can be obtained in the same way as in Lemma 4.5 of [1], by choosing $x = (\xi, w)$ in its proof. $\qquad\square$

It is worth noticing that the radius α, cf. the second claim of Lemma 2, does not depend on ξ. This independence allows us to obtain local and global error bounds on the function D, analogous to Lemmas 4.6 and 4.7 of [1], respectively.

Finally, we state that A4 and A5 hold under assumptions A6 to A8 of this subsection. The next theorem summarizes this fact, and it can be proved by straightforward adaptations on Theorem 4.8 of [1].

Theorem 4 *Let $\xi \in [0, \xi_{\max}]$ and $r \in [0, 1)$. Let $(w, u_\xi) \in W \times X \times \Delta$ be such that $C_\xi(w, u_\xi) \neq 0$. If the assumptions A6–A8 hold, there exist $\beta > 0$, independent of w and ξ, and $\overline{u}_\xi = (\overline{x}, \overline{\mu}, \overline{\gamma}) \in X \times \Delta$ such that*

$$|C_\xi(w, \overline{u}_\xi)| \leq r |C_\xi(w, u_\xi)|$$

and

$$\|(w, u_\xi) - (w, \overline{u}_\xi)\| \leq \beta |C_\xi(w, u_\xi)|.$$

Proof The statement follows from the proof of Theorem 4.8 of [1], changing x to (ξ, w) and $C(x, \cdot)$ to $D(\xi, w, \cdot) = C_\xi(w, \cdot)$. $\qquad\qquad\square$

4 Numerical experiments

First of all, we describe details of our implementation of Algorithm 1, namely the strategy IR_L. The same observations are valid for the other variants.

– We use the stabilized sequential quadratic programming (SSQP) method implemented in the package WORHP [7] to solve the problem of step 2 (including (24)) and to compute the projection in step 3. WORHP is a robust and efficient software library, especially designed for large problems.

– Theoretical convergence results require assumption A5 whenever s^k is not a lower level stationary point. In [1], the authors proposed to monitor the growth of the quotient $\|z^k - s^k\|/\|C(s^k)\|$, comparing it to the quotient of the previous iteration (note that the theory does not require any relationship between iterations). Instead, we simply verify if this quotient exceeds 10^8. If this is the case, we stop the algorithm declaring failure. In our experiments this situation never happened, and we will not make further comments along this line.

– We consider that the condition $d_{\tan}^k = 0$ in step 4 was satisfied if $\|d_{\tan}^k\|_\infty \leq \varepsilon_{\text{opt}}$ for a given tolerance to optimality $\varepsilon_{\text{opt}} > 0$. If this condition does not hold, we should compute $v^{k,i} \in \pi_k$ such that, for some $0 < t \leq t_{\text{break}}^{k,i}$,

$$L(v^{k,i}, \lambda^k) \leq \max\{L(z^k + td_{\tan}^k, \lambda^k), L(z^k, \lambda^k) - \tau_1\delta_{k,i}, L(z^k, \lambda^k) - \tau_2\}$$

with $\gamma \geq 0$ and $\|v^{k,i} - z^k\|_\infty \leq \delta_{k,i}$. Initially, we calculate

$$L^* = \max\{L(z^k + t_{\text{break}}^{k,i} d_{\tan}^k, \lambda^k), L(z^k, \lambda^k) - \tau_1\delta_{k,i}, L(z^k, \lambda^k) - \tau_2\}$$

and we consider the optimization problem

$$\begin{aligned}
\min_{v=(w,x,\mu,\gamma)} \quad & L(v, \lambda^k) \\
\text{s.t.} \quad & C'(z^k)(v - z^k) = 0, \quad x \geq 0 \\
& v \in W \times X \times \Delta \\
& \|v - z^k\|_\infty \leq \delta_{k,i}.
\end{aligned}$$

A stationary point v^* of this problem is potentially a good point for step 4. We then solve it by the SSQP method of the WORHP package, initialized at z^k as proposed in [1]. If the SSQP method fails, or if the "optimal" value $L(v^*, \lambda^k)$ found is not sufficiently small (i.e., $L(v^*, \lambda^k) > L^*$), then we reduce the neighborhood size $\delta_{k,i}$ by the factor 0.8. This reduction anticipates the one that will be done in step 7. So, we recalculate $t_{\text{break}}^{k,i}$ and test if $L(v^{k,i} + t_{\text{break}}^{k,i} d_{\tan}^k, \lambda^k) \leq L^*$ (this represents a step along the descent direction $d_{\tan}^k$). We make these reductions until $\delta_{k,i} \geq \delta_{\min}$ for a given parameter $\delta_{\min} > 0$. In case $\delta_{k,i} < \delta_{\min}$ we stop,

declaring failure. It should be mentioned that if the SSQP strategy fails, we never reapply it in the same iteration of the inexact restoration method because there is no reason to expect that this strategy can reach a good point with a smaller neighborhood radius.

- In step 5, it is important that $\lambda^{k,i}$ is a good estimative for a solution of $\nabla_v L(v^{k,i}, \lambda) = 0$ (or equivalently, a good approximation for a solution of the linear system $C'(v^{k,i})\lambda = -\nabla_v F(x)$). If the SSQP method has success in step 4, we take $\lambda^{k,i}$ as the computed multipliers. Otherwise, i.e., if $v^{k,i}$ was obtained by backtracking along direction $d_{\tan}^k$, we try to find a least squares solution of the overdetermined linear system

$$C'(v^{k,i})\lambda = -\nabla_v F(x)$$

by the orthogonal factorization method implemented in the MA49 of HSL [18]. In case of failure, we set $\lambda^{k,i}$ as the last $\lambda^{k,i}$ computed throughout the process. In all cases, we project $\lambda^{k,i}$ onto the box $[-M, M]^{2n+m}$, and thus $\|\lambda^{k,i}\|_\infty \leq M$.

- Finally, we make double-precision floating point corrections for the sums $\sum_{i=1}^p w_i$ of s^k, z^k, $v^{k,i}$ (to 1) and $d_{\tan}^k$ (to 0). These corrections aim to avoid numerical errors.

We compared the IR strategy (with and without perturbation) with the classical KKT reformulation (25). The resulting problem is a special case of Mathematical Problem with Complementarity Constraints (MPCC), in which the SSQP strategy is suitable [13–15]. Thus, we solve it by the WORHP package. Additionally, we have considered the homotopic strategy for solving MPCCs, developed by Scholtes in [26]. Roughly speaking, this strategy solves a sequence of MPCCs with perturbed complementarity. In our case, it consists in minimizing the upper level objective function F over a system similar to (23), in which $\gamma_i x_i \leq \xi$ rather than $\gamma_i x_i = \xi$, for $\xi \downarrow 0$. Scholtes proposed a simple update rule for the perturbation, namely, take $\xi^0 = 1$ and successively $\xi^k = \xi^{k-1}/10$ as long as $\xi^k \geq 10^{-16}$. It is worth noticing that, although we might have employed our updating rule (21), (22), we observed numerically that Scholtes' rule performs sightly better for his method. We refer to Scholtes' algorithm as KKT$^\xi$.

The WORHP package permits several configurations; among them we highlight:

- The globalization strategy is a line search with an augmented Lagrangian merit function. This is standard in WORHP.
- Quadratic subproblems are solved by an interior-point method.
- The BFGS updates for Hessian approximations are employed. This is standard in WORHP.
- Symmetric linear systems are solved by the multifrontal strategy MA97 of HSL [18].

We have addressed 13 test problems. Some of them were built using existing techniques or multiobjective instances in the literature, and some are completely new. Details are in the Appendix. Varying the starting point x^0, the initial weights w^0 and the dimensions p and n, we define 69 different instances, that are summarized in Tables 1 and 2.

Table 1 Instances of problems 1 to 8

Instance	p	n	x^0	w^0	F^*
P1.1	2	2	(10, 10)	(0.5, 0.5)	1.00e+01
P2.1	2	2	(10, 10)	(0.5, 0.5)	9.50e−01
P3.1	2	2	(−2, 20)	(0.8, 0.2)	0.00e+00
P3.2	2	2	(−10, 10)	(0.4, 0.6)	0.00e+00
P4.1	2	2	(−10, 10)	(0.5, 0.5)	0.00e+00
P4.2	2	2	(−10, 10)	(0.4, 0.6)	0.00e+00
P5.1	2	2	(−10, 10)	(0.5, 0.5)	5.01e−01
P5.2	2	2	(40, 100)	(0.1, 0.9)	5.01e−01
P6.1	2	2	(−10, 10)	(0.5, 0.5)	0.00e+00
P6.2	2	2	(10, 20)	(0.73, 0.27)	0.00e+00
P6.3	2	2	(10, 20)	(0.23, 0.77)	0.00e+00
P7.1	3	2	(10, 10)	(0.2, 0.3, 0.5)	0.00e+00
P7.2	3	2	(0.68, 1.72)	(0.2497, 0.6272, 0.1231)	0.00e+00
P7.3	3	2	(0.68, 1.72)	(0.3, 0.6, 0.1)	0.00e+00
P7.4	6	2	(−10, −10)	$w_i^0 = 1/6$	0.00e+00
P7.5	100	2	(9/10, 9/10)	$w_i^0 = 1/100$	0.00e+00
P7.6	200	2	(10, 1)	$w_i^0 = 1/200$	0.00e+00
P8.1	2	2	(20, 10)	(0.9, 0.1)	0.00e+00
P8.2	2	2	(10, 40)	(0.45, 0.55)	0.00e+00
P8.3	2	2	$(10^{-4}, 10)$	(0.8, 0.2)	0.00e+00
P8.4	2	6	$x_{2i}^0 = 20, x_{2i+1}^0 = 10$	(0.9, 0.1)	0.00e+00
P8.5	2	6	$x_{2i}^0 = 10, x_{2i+1}^0 = 40$	(0.45, 0.55)	0.00e+00
P8.6	2	6	$x_{2i}^0 = 10^{-4}, x_{2i+1}^0 = 10$	(0.82, 0.18)	0.00e+00
P8.7	2	20	$x_{2i}^0 = 20, x_{2i+1}^0 = 10$	(0.9, 0.1)	0.00e+00
P8.8	2	20	$x_{2i}^0 = 10, x_{2i+1}^0 = 40$	(0.45, 0.55)	0.00e+00
P8.9	2	20	$x_{2i}^0 = 10^{-4}, x_{2i+1}^0 = 10$	(0.82, 0.18)	0.00e+00
P8.10	2	30	$x_{2i}^0 = 20, x_{2i+1}^0 = 10$	(0.9, 0.1)	0.00e+00
P8.11	2	30	$x_{2i}^0 = 10, x_{2i+1}^0 = 40$	(0.45, 0.55)	0.00e+00

The experiments were run in a computer equipped with Intel(R) Core(TM) i7-6600U 2.60Ghz processor, 8Gb RAM, and GNU/Linux Ubuntu 64 bits operational system. The main part of the code was written in Fortran 90 language and compiled with gfortran 4.8 version. Another small part was written in C language, especially the AMPL interface, which is able to model multiobjective problems. For the inexact restoration strategies, we set the tolerance for optimality $\varepsilon_{\mathrm{opt}} = 10^{-4}$ (step 3 of Algorithm 1) and the parameters $\eta = 0.1$, $\theta_{-1} = 0.5$, $r = 0.5$, $\varepsilon_i = 10^{-4} \ \forall i$, $M = 10^{16}$, $\delta_{\min} = \varepsilon_{\mathrm{opt}}/2$, $\tau_1 = 10^{-2}$, $\tau_2 = 10^{-3}$, $\omega^k = 1/k^2$, $\lambda^0 = 0$. For all subproblems (steps 2, 3, and 4), we set the optimality tolerance of the SSQP

Table 2 Instances of problems 9 to 13

Instance	p	n	x^0	w^0	F^*
P9.1	2	2	$x^0_{2i} = 20, x^0_{2i+1} = 10$	(0.9, 0.1)	0.00e+00
P9.2	2	2	$x^0_{2i} = 10, x^0_{2i+1} = 40$	(0.45, 0.55)	0.00e+00
P9.3	2	2	$(10^{-4}, 10)$	(0.82, 0.18)	0.00e+00
P9.4	2	6	$x^0_{2i} = 20, x^0_{2i+1} = 10$	(0.9, 0.1)	0.00e+00
P9.5	2	6	$x^0_{2i} = 10, x^0_{2i+1} = 40$	(0.45, 0.55)	0.00e+00
P9.6	2	6	$x^0_{2i} = 10^{-4}, x^0_{2i+1} = 10$	(0.82, 0.18)	0.00e+00
P9.7	2	20	$x^0_{2i} = 20, x^0_{2i+1} = 10$	(0.9, 0.1)	0.00e+00
P9.8	2	20	$x^0_{2i} = 10, x^0_{2i+1} = 40$	(0.45, 0.55)	0.00e+00
P9.9	2	20	$x^0_{2i} = 10^{-4}, x^0_{2i+1} = 10$	(0.82, 0.18)	0.00e+00
P9.10	2	30	$x^0_{2i} = 20, x^0_{2i+1} = 10$	(0.9, 0.1)	0.00e+00
P9.11	2	30	$x^0_{2i} = 10, x^0_{2i+1} = 40$	(0.45, 0.55)	0.00e+00
P10.1	2	2	(20, 10)	(0.9, 0.1)	2.12e−03
P10.2	2	2	(10, 40)	(0.45, 0.55)	2.12e−03
P10.3	2	2	$(10^{-4}, 10)$	(0.82, 0.18)	2.12e−03
P10.4	2	6	$x^0_{2i} = 20, x^0_{2i+1} = 10$	(0.9, 0.1)	2.12e−03
P10.5	2	6	$x^0_{2i} = 10, x^0_{2i+1} = 40$	(0.45, 0.55)	2.12e−03
P10.6	2	6	$x^0_{2i} = 10^{-4}, x^0_{2i+1} = 10$	(0.82, 0.18)	2.12e−03
P10.7	2	20	$x^0_{2i} = 20, x^0_{2i+1} = 10$	(0.9, 0.1)	2.12e−03
P10.8	2	20	$x^0_{2i} = 10, x^0_{2i+1} = 40$	(0.45, 0.55)	2.12e−03
P10.9	2	20	$x^0_{2i} = 10^{-4}, x^0_{2i+1} = 10$	(0.82, 0.18)	2.12e−03
P10.10	2	30	$x^0_{2i} = 20, x^0_{2i+1} = 10$	(0.9, 0.1)	2.12e−03
P10.11	2	30	$x^0_{2i} = 10, x^0_{2i+1} = 40$	(0.45, 0.55)	2.12e−03
P11.1	2	2	(20, 10)	(0.9, 0.1)	0.00e+00
P11.2	2	2	(10, 40)	(0.45, 0.55)	0.00e+00
P11.3	2	2	$(10^{-4}, 10)$	(0.82, 0.18)	0.00e+00
P11.4	2	6	$x^0_{2i} = 20, x^0_{2i+1} = 10$	(0.9, 0.1)	0.00e+00
P11.5	2	6	$x^0_{2i} = 10, x^0_{2i+1} = 40$	(0.45, 0.55)	0.00e+00
P11.6	2	6	$x^0_{2i} = 10^{-4}, x^0_{2i+1} = 10$	(0.82, 0.18)	0.00e+00
P11.7	2	20	$x^0_{2i} = 20, x^0_{2i+1} = 10$	(0.9, 0.1)	0.00e+00
P11.8	2	20	$x^0_{2i} = 10, x^0_{2i+1} = 40$	(0.45, 0.55)	0.00e+00
P11.9	2	20	$x^0_{2i} = 10^{-4}, x^0_{2i+1} = 10$	(0.82, 0.18)	0.00e+00
P11.10	2	30	$x^0_{2i} = 20, x^0_{2i+1} = 10$	(0.9, 0.1)	0.00e+00
P11.11	2	30	$x^0_{2i} = 10, x^0_{2i+1} = 40$	(0.45, 0.55)	0.00e+00
P12.1	2	2	(10, 10)	(0.75, 0.25)	0.00e+00
P12.2	2	2	$(-10, -10)$	(0.35, 0.65)	0.00e+00
P12.3	2	2	(0.5, 0.5)	(0.5, 0.5)	0.00e+00
P12.4	6	6	$x^0_i = 1/2$	$w^0_i = 1/6$	0.00e+00
P13.1	2	2	(0.7, 0.7)	(0.5, 0.5)	0.00e+00
P13.2	2	2	(10, 10)	(0.75, 0.25)	0.00e+00
P13.3	2	2	(10, 10)	(0.25, 0.75)	0.00e+00
P13.4	6	6	$x^0_i = 7/10$	$w^0_i = 1/6$	0.00e+00

method equal to $\varepsilon_{\mathrm{SQP}} = 10^{-6}$. We do not perturb the complementarity for more than $\xi_{\max} = 10^{-2}$ in IR_F^ξ and IR_L^ξ algorithms, while the forced decrease factor ξ_{dec} is set to 0.8 (see expression (21)). We initialize step 3 with $\delta_{k,0} = 10$ (according with [1]) and in step 7 we make $\delta_{k,i+1} = \delta_{k,i}/2$.

We have compared the inexact restoration strategies with the KKT reformulation. Tables 3 and 4 show the results. Columns with "F" contain the final functional value obtained by the method. Bold values represent the best value of F, while italic values indicate that the method reached the optimal value. We decide that two methods

Table 3 Numerical results of problems 1 to 8

Inst.	IR_F F ; It	IR_F^ξ F ; It	IR_L F ; It	IR_L^ξ F ; It	KKT F/status	KKT$^\xi$ F/status
P1.1	*1.00e+01* ; 7	*1.00e+01* ; 8	1.03e+01 ; 5	1.44e+01 ; 8	infeas	2.25e+02
P2.1	1.10e+00 ; 2	1.08e+00 ; 5	1.10e+00 ; 2	1.08e+00 ; 5	infeas	*9.03e−01*
P3.1	*7.95e−10* ; 5	*2.19e−15* ; 5	*1.06e−14* ; 4	*2.12e−15* ; 5	*4.82e−30*	*7.44e−16*
P3.2	*1.92e−12* ; 2	*2.99e−08* ; 5	*1.92e−12* ; 2	*2.99e−08* ; 5	*3.15e−14*	*1.68e−21*
P4.1	2.00e+00 ; 3	*1.12e−18* ; 6	2.00e+00 ; 3	*1.44e−20* ; 7	2.00e+00	2.00e+00
P4.2	2.00e+00 ; 2	*2.50e−24* ; 6	2.00e+00 ; 2	2.40e−01 ; 6	2.00e+00	2.00e+00
P5.1	2.00e+00 ; 2	*5.01e−01* ; 5	2.00e+00 ; 2	*5.01e−01* ; 5	1.44e+00	1.44e+00
P5.2	*5.01e−01* ; 3	*5.01e−01* ; 5	*5.01e−01* ; 3	*5.01e−01* ; 5	1.44e+00	1.44e+00
P6.1	*3.32e−06* ; 3	*3.27e−06* ; 5	*3.32e−06* ; 3	4.92e−01 ; 5	7.07e+02	infeas
P6.2	*3.28e−06* ; 4	*3.28e−06* ; 6	*3.28e−06* ; 4	2.12e+00 ; 5	max it	infeas
P6.3	*3.25e−06* ; 3	*3.26e−06* ; 6	*3.25e−06* ; 3	*3.26e−06* ; 6	infeas	infeas
P7.1	*9.00e−08* ; 2	*9.03e−08* ; 5	*9.00e−08* ; 2	*9.03e−08* ; 5	4.00e+00	4.00e+00
P7.2	3.45e+00 ; 2	*8.97e−08* ; 5	3.45e+00 ; 2	*1.10e−07* ; 5	4.00e+00	4.00e+00
P7.3	3.45e+00 ; 2	*8.97e−08* ; 5	3.45e+00 ; 2	*1.52e−06* ; 5	4.00e+00	error
P7.4	*3.55e−12* ; 2	*1.19e−19* ; 5	*3.55e−12* ; 2	*1.66e−24* ; 5	4.00e+00	4.00e+00
P7.5	*1.21e−09* ; 3	*1.16e−19* ; 5	*1.21e−09* ; 3	*1.66e−24* ; 5	infeas	infeas
P7.6	*1.08e−09* ; 3	*1.34e−19* ; 5	*1.08e−09* ; 3	*1.70e−24* ; 5	infeas	infeas
P8.1	1.00e+00 ; 2	*1.32e−20* ; 6	1.00e+00 ; 2	*1.78e−20* ; 6	*8.92e−15*	*9.26e−09*
P8.2	*2.76e−21* ; 2	*3.72e−09* ; 5	*2.76e−21* ; 2	*3.72e−09* ; 5	infeas	*5.99e−22*
P8.3	1.00e+00 ; 2	*9.41e−19* ; 5	1.00e+00 ; 2	*9.40e−19* ; 5	*3.07e−11*	*4.97e−10*
P8.4	1.00e+00 ; 2	*7.44e−09* ; 4	1.00e+00 ; 2	*8.28e−09* ; 4	*5.17e−08*	*1.06e−08*
P8.5	*1.68e−16* ; 2	*6.57e−09* ; 5	*1.68e−16* ; 2	*6.57e−09* ; 5	infeas	*1.19e−10*
P8.6	1.00e+00 ; 2	*1.49e−19* ; 5	1.00e+00 ; 2	*1.46e−19* ; 5	*5.66e−09*	*2.00e−07*
P8.7	1.00e+00 ; 2	*1.41e−20* ; 5	1.00e+00 ; 2	*2.92e−18* ; 5	infeas	*7.79e−08*
P8.8	*8.96e−19* ; 2	*6.45e−08* ; 5	*8.96e−19* ; 2	*6.45e−08* ; 5	infeas	*5.91e−07*
P8.9	1.00e+00 ; 2	*2.21e−20* ; 5	1.00e+00 ; 2	*3.45e−16* ; 5	*2.96e−13*	4.49e−01
P8.10	1.00e+00 ; 2	*2.41e−20* ; 5	1.00e+00 ; 2	*7.10e−18* ; 5	infeas	*9.41e−09*
P8.11	*2.02e−20* ; 2	*4.82e−08* ; 5	*2.02e−20* ; 2	*4.82e−08* ; 5	*1.26e−09*	*1.64e−07*

Table 4 Numerical results of problems 9 to 13

Inst.	IR_F F ; It	IR_F^ξ F ; It	IR_L F ; It	IR_L^ξ F ; It	KKT F / Status	KKT^ξ F / Status
P9.1	1.00e+00 ; 2	*1.11e−21* ; 6	1.00e+00 ; 2	*1.10e−21* ; 6	*2.16e−15*	*3.62e−11*
P9.2	*1.24e−21* ; 2	1.00e+00 ; 5	*1.24e−21* ; 2	*3.46e−23* ; 5	*2.02e−10*	*5.00e−11*
P9.3	1.00e+00 ; 2	1.00e+00 ; 5	1.00e+00 ; 2	1.00e+00 ; 5	*7.61e−24*	*2.73e−22*
P9.4	1.00e+00 ; 2	1.00e+00 ; 5	1.00e+00 ; 2	1.00e+00 ; 5	*2.17e−12*	*2.70e−09*
P9.5	*6.35e−22* ; 2	1.00e+00 ; 6	*6.35e−22* ; 2	1.00e+00 ; 6	*6.53e−08*	*3.87e−11*
P9.6	1.00e+00 ; 2	*5.10e−09* ; 4	1.00e+00 ; 2	*5.10e−09* ; 4	infeas	*5.89e−10*
P9.7	1.00e+00 ; 2	1.00e+00 ; 5	1.00e+00 ; 2	1.00e+00 ; 5	*1.76e−09*	*3.28e−09*
P9.8	1.00e+00 ; 2	1.00e+00 ; 5	1.00e+00 ; 2	1.00e+00 ; 5	*1.79e−07*	*3.64e−08*
P9.9	1.00e+00 ; 2	*1.10e−08* ; 4	1.00e+00 ; 2	*1.10e−08* ; 4	infeas	*5.00e−08*
P9.10	1.00e+00 ; 2	1.00e+00 ; 5	1.00e+00 ; 2	1.00e+00 ; 5	*9.80e−08*	*4.26e−07*
P9.11	1.00e+00 ; 2	1.00e+00 ; 6	1.00e+00 ; 2	1.00e+00 ; 6	*2.03e−08*	*2.30e−09*
P10.1	1.00e+00 ; 2	1.00e+00 ; 5	1.00e+00 ; 2	1.00e+00 ; 5	*2.39e−03*	5.62e−02
P10.2	6.17e−02 ; 2	1.00e+00 ; 6	6.17e−02 ; 2	1.00e+00 ; 6	*2.29e−03*	2.00e−02
P10.3	1.00e+00 ; 2	1.00e+00 ; 4	1.00e+00 ; 2	1.00e+00 ; 4	*2.39e−03*	1.11e−01
P10.4	1.00e+00 ; 2	2.56e−03 ; 5	1.00e+00 ; 2	2.56e−03 ; 5	*2.02e−03*	2.17e−03
P10.5	6.17e−02 ; 2	9.65e−01 ; 5	6.17e−02 ; 2	9.65e−01 ; 5	2.29e−03	*2.17e−03*
P10.6	1.00e+00 ; 2	2.40e−03 ; 4	1.00e+00 ; 2	2.40e−03 ; 4	2.39e−03	*2.17e−03*
P10.7	1.00e+00 ; 2	1.00e+00 ; 5	1.00e+00 ; 2	9.65e−01 ; 5	infeas	*2.17e−03*
P10.8	9.65e−01 ; 3	2.29e−03 ; 5	9.65e−01 ; 3	2.29e−03 ; 5	2.09e−02	*2.17e−03*
P10.9	1.00e+00 ; 2	1.00e+00 ; 5	1.00e+00 ; 2	9.65e−01 ; 6	6.81e−01	*2.17e−03*
P10.10	1.00e+00 ; 2	1.00e+00 ; 4	1.00e+00 ; 2	1.00e+00 ; 5	*2.02e−03*	*2.17e−03*
P10.11	7.02e−01 ; 3	9.65e−01 ; 5	7.02e−01 ; 3	9.65e−01 ; 5	2.29e−03	*2.17e−03*
P11.1	1.99e+00 ; 2	*5.82e−24* ; 5	1.99e+00 ; 2	1.00e+00 ; 5	infeas	2.50e+01
P11.2	*4.95e+00* ; 2	2.47e+01 ; 5	*4.95e+00* ; 2	2.47e+01 ; 5	infeas	2.50e+01
P11.3	*9.89e+00* ; 2	1.58e+01 ; 5	*9.89e+00* ; 2	1.58e+01 ; 5	infeas	2.50e+01
P11.4	5.11e+01 ; 2	1.06e+01 ; 8	5.11e+01 ; 2	*8.90e+00* ; 6	1.14e+02	2.50e+01
P11.5	5.11e+01 ; 2	1.24e+02 ; 6	5.11e+01 ; 2	1.24e+02 ; 6	1.05e+02	*2.50e+01*
P11.6	5.11e+01 ; 2	3.93e+01 ; 6	5.11e+01 ; 2	*2.05e+00* ; 5	infeas	2.50e+01
P11.7	4.28e+02 ; 2	3.11e+02 ; 6	4.28e+02 ; 2	*1.95e+00* ; 6	3.66e+02	2.50e+01
P11.8	4.70e+02 ; 2	4.60e+02 ; 5	4.70e+02 ; 2	1.94e+02 ; 6	3.98e+02	*2.50e+01*
P11.9	2.50e+02 ; 2	1.04e+02 ; 6	2.50e+02 ; 2	*2.14e+00* ; 6	infeas	2.50e+01
P11.10	7.17e+02 ; 2	7.07e+02 ; 7	7.17e+02 ; 2	4.31e+02 ; 6	7.10e+02	*2.50e+01*
P11.11	7.17e+02 ; 2	error	7.17e+02 ; 2	4.38e+02 ; 124	6.89e+02	*2.50e+01*
P12.1	*1.80e−19* ; 3	*4.80e−09* ; 5	*1.80e−19* ; 3	*4.80e−09* ; 5	*2.07e−11*	*0.00e+00*
P12.2	1.00e+00 ; 2	*1.13e−08* ; 5	1.00e+00 ; 2	*1.13e−08* ; 5	1.05e−01	*9.02e−19*
P12.3	1.00e+00 ; 2	*7.77e−09* ; 5	1.00e+00 ; 2	*7.77e−09* ; 5	*5.65e−11*	*1.93e−08*
P12.4	*3.24e−07* ; 6	1.00e+00 ; 6	3.00e+00 ; 4	1.00e+00 ; 7	*3.07e−08*	*1.63e−08*
P13.1	*2.72e−08* ; 4	6.46e−03 ; 5	*2.72e−08* ; 4	6.46e−03 ; 5	5.71e−03	5.91e−03
P13.2	*2.71e−08* ; 4	*4.43e−08* ; 6	*2.71e−08* ; 4	*4.43e−08* ; 6	*3.66e−08*	5.53e−03

Table 4 (continued)

Inst.	IR_F F ; It	IR_F^{ξ} F ; It	IR_L F ; It	IR_L^{ξ} F ; It	KKT F / Status	KKT^{ξ} F / Status
P13.3	6.76e−03 ; 2	*8.81e−08* ; 7	6.76e−03 ; 2	*2.95e−08* ; 8	infeas	5.53e−03
P13.4	*4.86e−07* ; 5	*1.12e−05* ; 16	4.00e+00 ; 3	1.44e+00 ; 6	3.68e−01	1.00e+00

reached the same functional value or a method attains the optimal value whenever the difference between the functional values is at most 10^{-4}. The term "It" indicates the number of outer iterations. We do not report the number of iterations for the KKT reformulation because its not comparable with the outer iterations of inexact restoration techniques. On the other hand, the number of iterations of all IR strategies are comparable. The column "F/Status" also reports the answer of the method for the KKT reformulation and Scholtes' strategy: the value of F to indicate convergence, "max it" if the maximum number of iterations (1000) is reached and "infeas" if the method converges to a point that does not conform to the first-order stationary conditions of the lower level problem. We note that all IR strategies declare convergence within 150 iterations, except for IR_F^{ξ} on P11.11 instance, where a fatal error occurred.

Analyzing the highlighted best and optimal values of F of both tables, for problems 1 to 8 (Table 3), the inexact restoration strategy, particulary its perturbed version IR_F^{ξ}, overcame the KKT reformulation. Actually, for problems 1 to 8, the strategy IR_F^{ξ} had the best performance, and IR_L^{ξ} was the second-best strategy. These problems have nonlinear constraints, whereas the remaining have just bound constraints. For problems 9 to 13 (Table 4), among the IR strategies, IR_L^{ξ} stood out, but, overall, the KKT reformulation had a better performance, even improved by the perturbed KKT reformulation approach. Hence, we have observed that the inexact restoration strategy suits better for problems that have more complicated constraints.

All in all, the perturbed inexact restoration proved effective for problems 1 to 5, 7, 8, 12, and 13. It is worth mentioning that the set of efficient solutions of problem 11 is composed by several discontinuous parts (see Table 5 and the discussion in the Appendix), what possibly justifies the poor performance of the KKT reformulation, never reaching an optimal value for any instance of this problem. Scholtes' method, on the other hand, reached the best value for F at four instances, and a similar performance was attained by the IR_L^{ξ} version.

Despite the good behavior of the inexact restoration method, especially its perturbed version, we have not obtained practical evidences to affirm that the IR approach *always* overcame the KKT-based strategies. Indeed, for problem 10, the KKT reformulation and Scholtes' method obtained the best results. Nevertheless, for 19 out of the 69 total tested instances, the KKT reformulation could not achieve a first-order stationary point, being these "infeas" occurrences connected with the lack of robustness of such a strategy. Among these failures, we highlight instance P1.1, thoroughly discussed as example 1. For this particular instance, as can be seen in Table 3, the lack of preference of the KKT-based reformulations towards optimizers

clearly leaves such strategies in disadvantage in comparison with the four IR-based strategies.

5 Final remarks

We have applied inexact restorationstrategies to address bilevel programming problems in which the lower level is a multiobjective constrained problem. Compared with the traditional strategy of reformulating the bilevel problem and replacing the lower level optimization by its first-order optimality conditions, our approaches have shown better computational performance, especially in terms of robustness. For problems P1 to P7, similar conclusions can be reached in comparison with Scholtes' regularization method for MPCCs [26]. From the results of instances P7.1–P7.6 (Table 3), one can notice that employing pertubations within the IR strategy has shown the expected effect. Indeed, following our primary motivation, the generated sequences managed to escape from undesired KKT points. For these instances, however, Scholtes's method was not as successful as the perturbed IR strategies. On one hand, the theoretical convergence analysis of the employed IR must encompass the optimality conditions of the lower level problem and its relationship with the upper level problem. On the other hand, the proposed approach allows a practical deep search into the lower level problem, exploiting its original and intrinsic structure. Moreover, nonlinear problems were solved, with general inequalities in the description of the feasible set of the multiobjective lower level problem, tackled by a perturbation scheme, with promising results. The weighted-sum scalarization reformulation allowed us to put our original and non-standard problem as a classical bilevel one. Distinct possibilities for handling the lower level optimization problem, as well as a wider set of computational experiments are objects of future research.

Appendix: Test Problems

Problem 1. This is the problem (10)–(11) of example 1 on page 8.

Problem 2. We define problem 2 as the minimization of $F(x_1, x_2) = x_1^2 + x_2^2$ over the set $\mathscr{X}^*$ of efficient solutions of the multiobjective constrained problem [28]

$$
\begin{aligned}
&\min \ \{x_1, : x_2\} \\
&\text{s.t.} \ \ 1 + 0.1 \cos\left(16 \arctan \tfrac{x_2}{x_1}\right) - x_1^2 - x_2^2 \leq 0 \\
&\qquad \left(x_1 - \tfrac{1}{2}\right)^2 + \left(x_2 - \tfrac{1}{2}\right)^2 \leq \tfrac{1}{2} \\
&\qquad 0 \leq x_1, x_2 \leq \pi.
\end{aligned}
$$

$\mathscr{X}^*$ is discontinuous and nonconvex. There are four optimal solutions with functional value $F^* \approx 0.95$ (see Fig. 3).
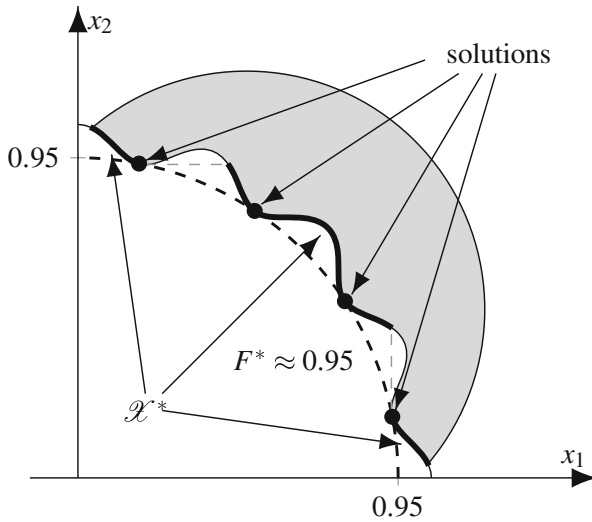
Fig. 3 Geometry of problem 2. The feasible set $\mathscr{X}^*$ is discontinuous and nonconvex. There are four solutions, with optimal value $F^* \approx 0.95$

Problem 3. Problem 3 consists in minimizing $F(x_1, x_2) = x_1^2 + x_2^2$ over the set of efficient solutions of

$$\min \{-x_1, \ x_1 + x_2^2\}$$
$$\text{s.t.} \quad x_1^2 - x_2 \leq 0$$
$$x_1 + 2x_2 \leq 3$$
$$x_1, x_2 \geq 0.$$

The feasible set is $\mathscr{X}^* = \{(x_1, x_2) \in \mathbb{R}^2; \ x_2 = x_1^2, : 0 \leq x_1 \leq 2\}$ and $(0, 0)^T$ is the optimal solution.

Problem 4. The lower level multiobjective problem is exactly the same of the problem 3, and thus the feasible set $\mathscr{X}^*$ is the same. The upper level objective function of problem 4 is $F(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 1)^2$, and $(1, 1)^T$ is the optimal solution.

Problem 5. This is a slight modification of problem 4. It consists in minimizing $F(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 1)^2$ over the set of efficient solutions of

$$\min \{x_1, \ -x_1 + x_2^2\}$$
$$\text{s.t.} \quad x_1^2 - x_2 \leq 0$$
$$-x_1 + 2x_2 \leq 3$$
$$x_1, x_2 \geq 0.$$

The feasible set is $\mathscr{X}^* = \{(x_1, x_2) \in \mathbb{R}^2; x_2 = x_1^2, 0 \leq x_1 \leq 2^{-2/3}\}$ and the optimal solution is $(2^{-2/3}, 2^{-4/3})^T \approx (0.63, 0.40)^T$.

Problem 6. The lower level problem is

$$\min \left\{ x_1^2 + \tfrac{1}{4}\left(x_2 - \tfrac{9}{2}\right)^2, x_2^2 + \tfrac{1}{4}\left(x_1 - \tfrac{9}{2}\right)^2 \right\}$$
$$\text{s.t.} \qquad -x_1 - x_2 + 2 \leq 0$$
$$x_1, x_2 \geq 0.$$

Geometrically, the set $\mathscr{X}^*$ of efficient solutions of this multiobjective problem is composed by the simultaneously tangential points of the level sets of both objective functions, when it is contained in the semispace defined by $x_1 + x_2 \geq 2$, and by the line $x_1 + x_2 = 2$ otherwise. More specifically, $\mathscr{X}^*$ is the union of the curves

$$C_1 : \left\{ \left(\tfrac{9t}{2t-8}, \tfrac{9}{2-8t} \right) ; -\tfrac{1}{2} \leq t \leq 0 \right\},$$
$$C_2 : \left\{ (t, 2-t) ; \tfrac{1}{2} \leq t \leq \tfrac{3}{2} \right\},$$
$$C_3 : \left\{ \left(\tfrac{9}{2-8t}, \tfrac{9t}{2t-8} \right) ; -\tfrac{1}{2} \leq t \leq 0 \right\}.$$

The upper level objective function is $F(x_1, x_2) = (x_1 - 9/2)^2 + x_2^2$, for which $(9/2, 0)^T$ is the optimal solution. Figure 4 illustrates its geometry.

Problem 7. This is the problem (12)–(13) of example 2 on page 9.

Problems 8 to 11 (ZDT family instances). Multiobjective problems of [29] (here named ZDT problems) are constructed with three functions f_1, g, h in the following way:

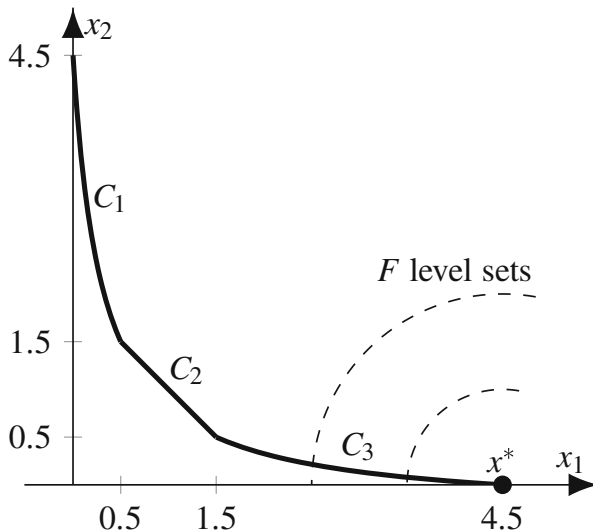$$\min \{f_1(x_1), f_2(x_1, \ldots, x_n)\}$$

Fig. 4 Geometry of problem 6. The feasible set is the union of curves C_1, C_2, and C_3. The optimal solution is $x^* = (9/2, 0)^T$

where

$$f_2(x_1, \ldots, x_n) = g(x_2, \ldots, x_n) h(f_1(x_1), g(x_2, \ldots, x_n))$$

(see also [9]). We were inspired in four combinations suggested in Definition 4 of [29]. For each combination, we minimize the following:

$$F(x_1, x_2, \ldots, x_n) = (x_1 - 1)^2 + \sum_{i=2}^{n} x_i^2$$

over the set of efficient solutions defined by the minimization of functions f_1 and f_2 on certain box constraints. We summarized the properties of these test problems in Table 5. For these problems, the set of efficient solutions is split into discontinuous components, one for each (local) minimum of g, formed with all possible values of f_1 (see Section 5.1 of [9]). Thus, problems 8 to 10 have a unique set of efficient solutions component, for which $x_2 = \cdots = x_n = 0$ (i.e. $g(x_2, \ldots, x_n) = 1$), while Problem 11 has several components. For problems 8, 9, and 11, the possible values of $x_1 = f_1(x_1)$ are the whole interval $[0, 1]$. For problem 10, the possible values of x_1 are the intervals for which h decreases (see Fig. 5). Hence, $(1, 0, \ldots, 0)$ is the optimal solution of problems 8, 9, and 11, whereas $(0.954, 0 \ldots, 0)^T$ is the (approximated) optimal solution of problem 10. Problem 11 has 21^{n-1} sets of efficient solutions components, one for each combination of the local minimizers of $x_i^2 - 2\cos(4\pi x_i)$, $i = 2, \ldots, n$ (see Fig. 6).

Table 5 ZDT family instances.

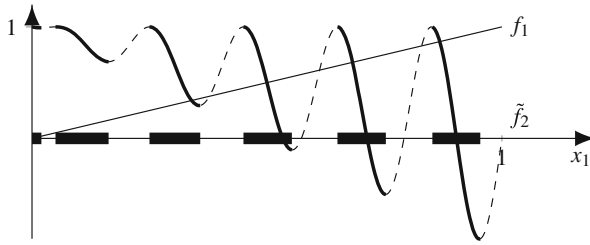| Problem ID	f_1, g, h, box constraints	Set of efficient solutions
Problem 8	$f_1 = x_1$	$x_2 = \cdots = x_n = 0$
	$g = 1 + 9 \sum_{i=2}^{n} x_i / (n-1)$	$x_1 \in [0, 1]$
	$h = 1 - f_1/g$	
	$x \in [0, 1]^n$	
Problem 9	$f_1 = x_1$	$x_2 = \cdots = x_n = 0$
	$g = 1 + 9 \sum_{i=2}^{n} x_i / (n-1)$	$x_1 \in [0, 1]$
	$h = 1 - (f_1/g)^2$	
	$x \in [0, 1]^n$	
Problem 10	$f_1 = x_1$	$x_2 = \cdots = x_n = 0$
	$g = 1 + 9 \sum_{i=2}^{n} x_i / (n-1)$	x_1 in the intervals
	$h = 1 - f_1/g(1 - \sin(10\pi f_1))$	in which h decreases
	$x \in [0, 1]^n$	
Problem 11	$f_1 = x_1$	several efficient
	$g = 2n - 1 + \sum_{i=2}^{n} (x_i^2 - 2\cos(4\pi x_i))$	components
	$h = 1 - (f_1/g)^2$	
	$x_1 \in [0, 1], \ x_2, \ldots, x_n \in [-5, 5]$	

Fig. 5 Geometry of problem 10. In the set of efficient solutions, $g(x_2^*, \ldots, x_n^*) = 1$ and x_1 lies on the intervals for which $\tilde{f}_2(x_1) = f_2(x_1, x_2^*, \ldots, x_n^*) = 1 - x_1(1 - \sin(10\pi x_1))$ decreases (huge lines). These are the non-dominated values corresponding to the two objectives $f_1(x_1) = x_1$ and $\tilde{f}_2(x_1)$

Problems 12 and 13 (WFG family instances). Based on the Walking Fish Group (WFG) Toolkit [19], we consider two instances, for which particular choices of the parameters were made to come up with smooth functions:

– **Problem 12**, that consists in minimizing $F(x) = \sum_{i=1}^{n} x_i^2$, over the set of efficient solutions defined by the minimization of

$$f_1(x) = \frac{x_n}{2n} + 2 \prod_{i=1}^{n-1} \left[1 - \cos\left(\frac{\pi x_i}{4i}\right) \right],$$

$$f_j(x) = \frac{x_n}{2n} + 2j \left[1 - \sin\left(\frac{\pi x_{n-j+1}}{4(n-j+1)}\right) \right] \prod_{i=1}^{n-j} \left[1 - \cos\left(\frac{\pi x_i}{4i}\right) \right],$$

$$j = 2, \ldots, n-1,$$

$$f_n(x) = \frac{x_n}{2n} + 2n \left[1 - \frac{x_1}{2} - \frac{\cos(5\pi x_1 + \pi/2)}{10\pi} \right]$$
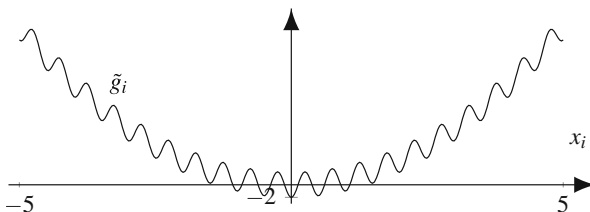
on $[0, 1]^n$. The optimal solution is the origin.

Fig. 6 Geometry of problem 11. There is one set of efficient solutions component for each combination of local minimizers of $\tilde{g}_i(x_i) = x_i^2 - 2\cos(4\pi x_i)$, $i = 2, \ldots, n$

– **Problem 13**, that consists in minimizing $F(x) = \sum_{i=1}^{n} x_i^2$, over the set of efficient solutions defined by the minimization of

$$f_1(x) = \frac{x_n}{2n} + 2\prod_{i=1}^{n-1}\left[1 - \cos\left(\frac{\pi x_i}{4i}\right)\right],$$

$$f_j(x) = \frac{x_n}{2n} + 2j\left[1 - \sin\left(\frac{\pi x_{n-j+1}}{4(n-j+1)}\right)\right]\prod_{i=1}^{n-j}\left[1 - \cos\left(\frac{\pi x_i}{4i}\right)\right],$$

$$j = 2, \ldots, n-1,$$

$$f_n(x) = \frac{x_n}{2n} + 2n\left[1 - \frac{x_1}{2}\cos^2\left(\frac{5\pi x_1}{2}\right)\right]$$

on $[0, 1]^n$. Again, the optimal solution is the origin.

References

1. Andreani, R., Castro, S.L.C., Chela, J.L., Friedlander, A., Santos, S.A.: An inexact-restoration method for nonlinear bilevel programming problems. Comput. Optim. Appl. **43**(3), 307–328 (2009)
2. Benson, H.P.: An improved definition of proper efficiency for vector maximization with respect to cones. J. Math. Anal. Appl. **71**(1), 232–241 (1979)
3. Benson, H.P.: Optimization over the efficient set. J. Math. Anal. Appl. **98**(2), 562–580 (1984)
4. Bonnel, H., Morgan, J.: Semivectorial bilevel optimization problem: penalty approach. J. Optim. Theory Appl. **131**(3), 365–382 (2006)
5. Bonnel, H., Morgan, J.: Optimality conditions for semivectorial bilevel convex optimal control problems, Proceedings in Mathematics & Statistics, vol. 50, pp. 5–78. Springer, Heidelberg (2013)
6. Borwein, J.: Proper efficient points for maximizations with respect to cones. SIAM J. Control. Optim. **15**(1), 57–63 (1977)
7. Bskens, C., Wassel, D.: The ESA NLP Solver WORHP, chap. 8. Springer. Springer Optimization and Its Applications **73**, 85–110 (2012)
8. Bueno, L.F., Haeser, G., Martínez, J.M.: An inexact restoration approach to optimization problems with multiobjective constraints under weighted-sum scalarization. Optim. Lett. **10**(6), 1315–1325 (2016)
9. Deb, K.: Multi-objective genetic algorithms: problem difficulties and construction of test problems. Evol. Comput. **7**(3), 205–230 (1999)
10. Ehrgott, M.: Multicriteria Optimization. Springer, Berlin (2005)
11. El-Bakry, A.S., Tapia, R.A., Tsuchiya, T., Zhang, Y.: On the formulation and theory of the newton interior-point method for nonlinear programming. J. Optim. Theory Appl. **89**(3), 507–541 (1996)
12. Fülöp, J.: On the equivalence between a linear bilevel programming problem and linear optimization over the efficient set. Technical report WP93-1, Laboratory of Operations Research and Decision Systems, Computer and Automation Institute, Hungarian Academy of Sciences (1993)
13. Fletcher, R., Leyffer, S.: Numerical experience with solving MPECs as NLPs. Tech. rep., University of Dundee Report NA pp. 210 (2002)
14. Fletcher, R., Leyffer, S.: Solving mathematical programs with complementarity constraints as nonlinear programs. Optimization Methods and Software **19**(1), 15–40 (2004)
15. Fletcher, R., Leyffer, S., Ralph, D., Scholtes, S.: Local convergence of SQP methods for mathematical programs with equilibrium constraints. SIAM J. Optim. **17**(1), 259–286 (2006)
16. Geoffrion, A.M.: Proper efficiency and the theory of vector maximization. J. Math. Anal. Appl. **22**(3), 618–630 (1968)

17. Guo, X.L., Li, S.J.: Optimality conditions for vector optimization problems with difference of convex maps. J. Optim. Theory Appl. **162**(3), 821–844 (2014)
18. HSL. A collection of Fortran codes for large scale scientific computation. Available at http://www.hsl.rl.ac.uk/
19. Huband, S., Barone, L., While, L., Hingston, P.: A scalable multi-objective test problem toolkit, pp. 280–295. Springer, Berlin (2005)
20. Kuhn, H.W., Tucker, A.W.: Nonlinear programming. In: Neyman, J. (ed.) Proceedings of the second Berkeley symposium on mathematical statistics and probability, pp. 481–492. University of California Press, Berkeley (1951)
21. Martínez, J.M.: Inexact-restoration method with Lagrangian tangent decrease and new merit function for nonlinear programming. J. Optim. Theory Appl. **3**(1), 39–58 (2001)
22. Martínez, J.M., Pilotta, E.A.: Inexact-restoration algorithm for constrained optimization. J. Optim. Theory Appl. **104**(1), 135–163 (2000)
23. Martínez, J.M., Svaiter, B.F.: A practical optimality condition without constraint qualifications for nonlinear programming. J. Optim. Theory Appl. **118**(1), 117–133 (2003)
24. Miettinen, K.M.: Nonlinear Multiobjective Optimization. Kluwer Academic Publishers, Boston (1999)
25. Pilotta, E.A., Torres, G.A.: An inexact restoration package for bilevel programming problems. Appl. Math. **15**(10A), 1252–1259 (2012)
26. Scholtes, S.: Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. SIAM J. Optim. **11**(4), 918–936 (2001)
27. Srinivas, N., Deb, K.: Muiltiobjective optimization using nondominated sorting in genetic algorithms. Evol. Comput. **2**(3), 221–248 (1994)
28. Tanaka, M., Watanabe, H., Furukawa, Y., Tanino, T.: GA-based decision support system for multicriteria optimization. In: IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21St Century, vol. 2, pp. 1556–1561 (1995)
29. Zitzler, E., Deb, K., Thiele, L.: Comparison of multiobjective evolutionary algorithms: empirical results. Evol. Comput. **8**(2), 173–195 (2000)