INFORMS Journal on Computing
Vol. 21, No. 2, Spring 2009, pp. 257–267
ISSN 1091-9856 | EISSN 1526-5528 | 09 | 2102 | 0257

DOI 10.1287/ijoc.1080.0290

A Complementarity Constraint Formulation of Convex Multiobjective Optimization Problems

Sven Leyffer
Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois 60439,
leyffer@mcs.anl.gov

We propose a new approach to convex nonlinear multiobjective optimization that captures the geometry of the Pareto set by generating a discrete set of Pareto points optimally. We show that the problem of finding a maximally uniform representation of the Pareto surface can be formulated as a mathematical program with complementarity constraints. The complementarity constraints arise from modeling the set of Pareto points, and the objective maximizes some quality measure of this discrete set. We present encouraging numerical experience on a range of test problems collected from the literature.

Key words: multiobjective optimization; nonlinear programming; complementarity constraints; mathematical program with complementarity constraints
History: Accepted by John W. Chinneck, former Area Editor for Modeling: Methods and Analysis; received January 2007; revised August 2007, February 2008; accepted March 2008. Published online in *Articles in Advance* September 17, 2008.

1. Introduction

We consider the solution of nonlinear multiobjective optimization problems (MOOPs). MOOPs arise in engineering and economic applications with multiple competing objectives. Applications include the construction of structures to minimize total mass and maximize stiffness, design problems with multiple loading cases, and airplane design to maximize fuel efficiency and minimize cabin noise; see the recent monographs (Ehrgott 2005, Hillermeier 2001, Miettinen 1999, Rustem 1998, Stadler 1988, Steuer 1986).

The multiobjective optimization problem is formally defined as

$$\text{(MOOP)}\quad \begin{cases} \underset{x\ge 0}{\text{minimize}} & f(x) \\ \text{subject to} & c(x)\ge 0, \end{cases}$$

where $x\in\mathbb{R}^n$. We assume that the objective functions $f(x)=(f_1(x),\ldots,f_p(x))$: $\mathbb{R}^n\to\mathbb{R}^p$ and that the constraints $c(x)=(c_1(x),\ldots,c_m(x))$: $\mathbb{R}^n\to\mathbb{R}^m$ are twice continuously differentiable. We denote the feasible set by

$$\mathcal{F}:=\{x\ge 0\colon\, c(x)\ge 0\}$$

and assume that it is nonempty.

We present a new approach to nonlinear multiobjective optimization that captures the geometry of the Pareto set by generating a discrete set of Pareto points that maximizes the uniformity of the representation of the Pareto set. We show that the problem of finding an optimal discrete representation of the Pareto set can be formulated as a bilevel optimization problem. If MOOP is convex, then we show how to solve the bilevel problem as a mathematical program with complementarity constraints (MPCCs) by taking advantage of recent progress on the solution of MPCCs.

This paper is organized as follows. In the remainder of this section we briefly review optimality conditions for MOOPs, discuss existing solution methods, and motivate our approach with a small example. In §2 we formally introduce our new approach and derive some theoretical properties of our formulation. In §3 we describe a random MOOP generator and a collection of test problems from the literature, and we present our numerical results. In §4 we briefly examine open questions and suggest some future lines of research.

1.1. Introduction to Multiobjective Optimization

We start by reviewing some basic concepts of MOOPs that will be used throughout the paper. Let x_k^* denote a solution to the single-objective nonlinear program (NLP) given by

$$\begin{cases} \underset{x\ge 0}{\text{minimize}} & f_k(x) \\ \text{subject to} & c(x)\ge 0, \end{cases} \tag{1}$$

and define the *payoff matrix* $Z\in\mathbb{R}^{p\times p}$ as $Z_{ij}:=f_i(x_j^*)$, which provides useful information on the trade-offs between the multiple objectives. Note that the minima of each single-objective NLP (1) are the diagonal of entries of Z, also referred to as *ideal values*. We define

$$\underline{z}^* := (f_1(x_1^*), \dots, f_p(x_p^*)) \quad \text{and}$$
$$\bar{z}^* := \Big(\max_{i\neq 1} f_1(x_i^*), \dots, \max_{i\neq p} f_p(x_i^*)\Big), \tag{2}$$

and note that the ideal values $\underline{z}^*$ and the maximum values $\bar{z}^*$ approximate the range of the objective values.

Optimality conditions for MOOPs are given by Miettinen and Mäkelä (1998a), based on normal cones and Clarke's generalized gradients (Clarke 1983).

Definition 1.1 (Miettinen and Mäkelä 1998b). Let $x^* \in \mathcal{F}$ be a feasible point with the corresponding criterion vector $z^* = f(x^*)$.

1. (x^*, z^*) is *globally Pareto-optimal* if there exists no $x \in \mathcal{F}$, $x \neq x^*$, with $f_k(x) \leq f_k(x^*)$ for all $k = 1, \dots, p$, and $f_r(x) < f_r(x^*)$ for at least one index $1 \leq r \leq p$.
2. (x^*, z^*) is *locally Pareto-optimal* if there exists a $\delta > 0$ such that $x^* \in \mathcal{F}$ is globally Pareto-optimal in $\mathcal{F} \cap B(x^*, \delta)$, where $B(x^*, \delta)$ is a ball of radius δ around x^*.
3. We designate the set of all Pareto points as $\mathcal{P} := \{z^*: (x^*, z^*) \text{ is a Pareto point}\}$.
4. MOOP is said to be convex if the functions $f(x)$ are convex and the constraint functions $c(x)$ are concave (i.e., the feasible set is convex).

The following result gives a necessary condition for local Pareto optimality.

Theorem 1.2 (Miettinen and Mäkelä 1998b). *Let $x^* \in \mathcal{F}$ be a feasible point at which Cottle's constraint qualification holds. A necessary condition for $z^* = f(x^*)$ to be locally Pareto-optimal is that there exist multipliers $w \geq 0$, $w \neq 0$, and $y \geq 0$ such that*

$$0 = \sum_{k=1}^{p} w_k \nabla f_k(x^*) - \sum_{j=1}^{m} y_j \nabla c_j(x^*), \tag{3}$$

and $y_j c_j(x^*) = 0$ for all $j = 1, \dots, m$. If MOOP is convex, then this condition is also sufficient.

1.2. Solution Methods for MOOPs

Here, we briefly review two techniques for finding a single Pareto point. Other techniques can be found in recent monographs by Ehrgott (2005) and Miettinen (1999). Both techniques form the basis of our approach to finding multiple Pareto points. The first technique forms a convex combination of the objective functions and solves the following NLP:

$$(\mathrm{SUM}(w)) \quad \begin{cases} \underset{x\geq 0}{\text{minimize}} & \displaystyle\sum_{k=1}^{p} w_k f_k(x) \\ \text{subject to} & c(x) \geq 0, \end{cases}$$

where the weights $w_k \geq 0$, $k = 1, \dots, p$, with $\sum w_k = 1$. By varying the weights we can identify different Pareto points. We are grateful to an anonymous referee for pointing out that SUM(w) may generate weakly dominated solutions. Only if the MOOP is convex can SUM(w) generate all Pareto points by varying w.

The second technique is related to goal programming and classification techniques. It minimizes one objective subject to achieving a given goal on all other objectives. Without loss of generality, we let $f_1(x)$ be the objective that is minimized, and we denote the goals by $z \in \mathbb{R}^{p-1}$ and solve the following NLP:

$$(\mathrm{GOAL}(z)) \quad \begin{cases} \underset{x\geq 0}{\text{minimize}} & f_1(x) \\ \text{subject to} & f_k(x) \leq z_k,\ k = 2, \dots, p, \\ & c(x) \geq 0. \end{cases}$$

Clearly, the goals should be chosen to lie between $\underline{z}^*$ and $\bar{z}^*$, although not all choices of z give rise to a feasible problem GOAL(z). We show in the next section that GOAL(z) gives rise to Pareto points. In contrast to SUM(w), however, all feasible choices of target z generate a Pareto point, and all Pareto points can be found by varying z.

1.3. Motivation of New Approach

One way in which we can obtain a discrete description of the Pareto set, $\mathcal{P}$, is to solve SUM(w) or GOAL(z) repeatedly for different weights or goals. However, choosing the weights and goals is not straightforward. For example, Das and Dennis (1998) have observed that a uniform distribution of weights does not provide a uniform description of the Pareto set. Figure 1 shows two discrete descriptions of the Pareto set of three objective functions. The first description (circles) was generated from a uniform distribution of the goals, while the second description (boxes) was generated by maximizing the uniformity of the representation. The figure shows two viewpoints of the same three-dimensional (3-D) Pareto set and shows that the optimized description provides a better description of the Pareto set.

We close this section by summarizing our main assumptions.

Assumptions 1.3. *Throughout, we make the following assumptions*:

A1. *The problem functions $f(x)$ and $c(x)$ are twice continuously differentiable.*

A2. *The feasible set $\mathcal{F} := \{x \mid x \geq 0 \text{ and } c(x) \geq 0\}$ is not empty and bounded.*

A3. *Any local solution to SUM(w) and GOAL(z) satisfies the Mangasarian-Fromowitz constraint qualification and a second-order sufficient condition.*

A4. *The functions $f(x)$ and $c(x)$ are convex.*

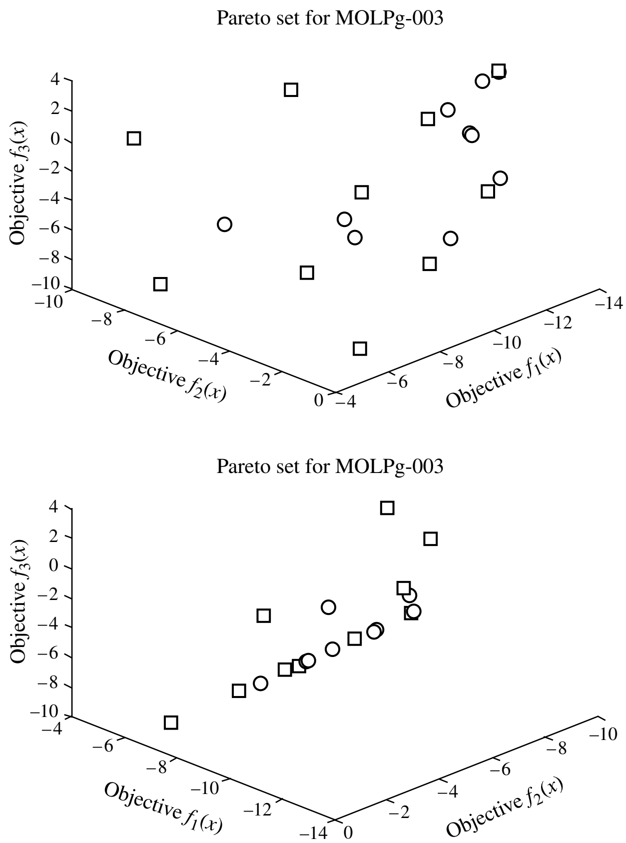

Figure 1 Two Perspectives of the Same Representations of a Pareto Set

Note. The circles show a uniform Pareto set, and the boxes show the maximally uniform Pareto set.

Assumptions A1 to A3 are relatively weak and simply ensure that any single-objective NLP is tractable and can be solved by using standard NLP techniques. We could replace the boundedness assumption A2 by an assumption on the boundedness of level sets of an exact penalty function. The most restrictive assumption is Assumption A4. The main reason for this assumption is that we replace the NLPs SUM(w) and GOAL(z) by their respective first-order conditions, which are necessary and sufficient, if the NLPs are convex. We note that convexity does not imply the second-order sufficient condition.

2. Optimal Representation of the Pareto Surface

In this section we present a new approach to finding a discrete representation of the Pareto set, $\mathcal{P}$, that is optimal in a certain sense. We start by reviewing three quality measures of a discrete representation of the Pareto set proposed by Sayin (2000) and show that they lead to a bilevel problem whose solution corresponds to an optimal representation of the Pareto set. We also derive a complementarity constraint formulation by replacing the lower-level problems by their first-order conditions.

2.1. Bilevel Formulation of MOOPs

Sayin (2000) introduces three quality measures of a discrete representation of the Pareto set: cardinality, coverage error, and uniformity of the representation. We assume here that the cardinality is user defined and is fixed. The coverage error for a discrete representation $\mathcal{D} \subset \mathcal{P}$ of the Pareto set, $\mathcal{P}$, is defined as

$$\epsilon = \max_{v \in \mathcal{P}} \min_{u \in \mathcal{D}} \|u - v\|,$$

where $\|\cdot\|$ is any norm in $\mathbb{R}^p$. Unfortunately, to compute this measure, we require explicit knowledge of the Pareto set, $\mathcal{P}$. We therefore believe that coverage error is not a practical measure of quality. However, the final quality measure introduced by Sayin (2000)—namely, the uniformity of the representation—can be used to derive optimally uniform approximations of the Pareto set. Uniformity of representation is defined as the largest η such that

$$\eta \leq \min_{u,\, v \in \mathcal{D},\, u \neq v} \|u - v\|. \tag{4}$$

Next, we show that the problem of finding a maximal uniform representation of the Pareto set, $\mathcal{P}$, can be formulated as a bilevel programming problem. We consider *any* single-objective approach such as SUM(w) or GOAL(z) and consider the parameters w or z as variables that are optimally determined within a bilevel optimization problem. The key idea is to simultaneously determine $q \geq 2$ Pareto points x_l, for $l = 1, \ldots, q$, and their corresponding parameters w_l (or z_l) such that the Pareto points maximize the uniformity of the presentation of the Pareto set. The upper level controls the parameters w_l (or z_l), while the lower level corresponds to q single-objective NLPs given by SUM(w_l) or GOAL(z_l).

Figure 2 provides a graphical illustration of our approach. There are two objective functions, and the solid line shows the Pareto set. We are seeking a given number of discrete points such that the pairwise distances between the Pareto points is maximized, illustrated by the circles around each Pareto point. Here, we maximize η subject to the constraints $\eta \leq \eta_{lk}$, where $\eta_{lk} = \|f(x_l) - f(x_k)\|$ and x_l are Pareto points characterized by solving SUM(w) or GOAL(z).

Formally, we consider the problem of finding a given number $q \geq 2$ of Pareto points that maximize the uniformity of the discrete representation of the Pareto set. We start by deriving a problem to find an optimal representation of the Pareto set based on the convex combination problem SUM(w). Let $w := (w_1, \ldots, w_q)^T$ denote the weights to be determined,

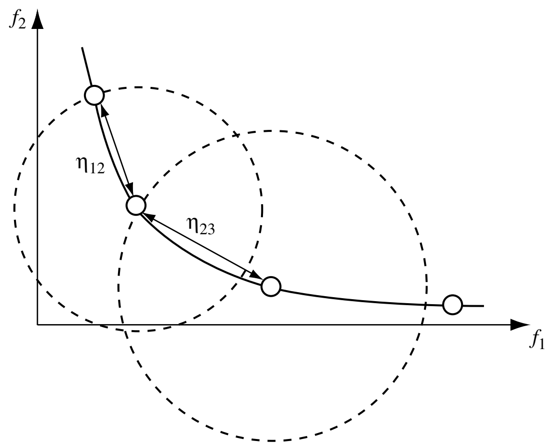

Figure 2 **Maximizing Distances Between Pareto Points**

and let $x := (x_1, \ldots, x_q)^T$ denote the corresponding Pareto points (one copy for each Pareto point). The problem of maximizing the uniformity of the discrete representation of the Pareto set can then be formulated as the following bilevel optimization problem (in the remainder we choose the ℓ_2 norm for simplicity):

$$\begin{cases} \underset{x,w,\eta}{\text{maximize}} & \eta \\ \text{subject to} & \eta \le \|f(x_l) - f(x_k)\|_2^2 \\ & \qquad \forall 1 \le k,\ l \le q,\ k \ne l, \\ & w_k \ge 0, \text{ and } e^T w_k = 1, \quad \forall k = 1, \ldots, q, \\ & x_k \text{ solves SUM}(w_k). \end{cases} \tag{5}$$

The aim in (5) is to find $q \ge 2$ Pareto points such that the smallest distance between any two function values f_k is pushed as far apart as possible while remaining within the Pareto set. As is customary in bilevel optimization, we refer to w and η as the control, or upper-level, variables and to x as the state, or lower-level, variables. We note that even though MOOP is convex, the bilevel problem is in general nonconvex, and the task of finding a global solution is daunting. However, we present numerical evidence in §3 that even local solutions of (5) provide improved representations of the Pareto set.

One disadvantage of (5) is the lack of general-purpose solvers for bilevel optimization problems. To develop a practical technique for solving (5), we therefore replace the constraint "x_k solves SUM(w_k)" by its first-order conditions and exploit recent advances in the development of robust solvers for mathematical programs with complementarity constraints.

Under Assumptions A1–A4, it follows that the first-order conditions for SUM(w_k) are necessary and sufficient. We can therefore equivalently replace (5) by the following mathematical program with complementarity constraints (MPCC):

$$\begin{cases} \underset{x,\,y,\,w\ge 0,\,\eta}{\text{maximize}} & \eta \\ \text{subject to} & \eta \le \|f(x_k) - f(x_l)\|_2^2 \\ & \qquad \forall 0 \le k,\ l \le q,\ k \ne l, \\ & e^T w_l = 1 \quad \forall l = 1, \ldots, q, \\ & 0 \le x_l \perp \nabla(w_l^T f(x_l)) - \nabla c(x_l) y_l \ge 0 \\ & \qquad \forall l = 1, \ldots, q, \\ & 0 \le y_l \perp c(x_l) \ge 0 \quad \forall l = 1, \ldots, q, \end{cases} \tag{6}$$

where the last two sets of constraints are complementarity constraints. The notation $0 \le u \perp v \ge 0$ means that the two vectors $u, v \ge 0$ and that, in addition, $u^T v \le 0$; that is, a component i of $u_i = 0$ or the corresponding component of $v_i = 0$. We note that the dimension of (6) is roughly q times the dimension of SUM(w_l) (plus $q \times p$ weights), as every Pareto point requires a new copy of the primal and dual variables x and y. We can remove one component of each w_l and the constraints $e^T w_l = 1$ if we replace the first-order condition by

$$0 \le x_l \perp \nabla((1, \hat{w}_l)^T f(x_l)) - \nabla c(x_l) y_l \ge 0 \qquad \forall l = 1, \ldots, q, \tag{7}$$

where $\hat{w}_l \in \mathbb{R}^{p-1}$ are the weights on the remaining objectives. This formulation has the advantage that it removes one bilinearity from the first-order condition. We note, however, that (6) and (7) are not equivalent, because the latter overemphasizes the first objective.

An alternative MPCC is obtained by using the first-order conditions of GOAL(z). In this case, we are looking for goals $z = (z_1, \ldots, z_q)$ and corresponding multipliers $u = (u_1, \ldots, u_q)$ that solve

$$\begin{cases} \underset{x,\,y,\,z,\,u,\,\eta}{\text{maximize}} & \eta \\ \text{subject to} & \eta \le \|f(x_k) - f(x_l)\|_2^2 \\ & \qquad \forall 0 \le k,\ l \le q,\ k \ne l, \\ & 0 \le x_l \perp \nabla((1, u_l)^T f(x_l)) - \nabla c(x_l) y_l \ge 0 \\ & \qquad \forall l = 1, \ldots, q, \\ & 0 \le y_l \perp c(x_l) \ge 0 \quad \forall l = 1, \ldots, q, \\ & 0 \le u_l \perp z_l - \hat{f}(x_l) \ge 0, \end{cases} \tag{8}$$

where $\hat{f}(x_l) = (f_2(x_l), \ldots, f_p(x_l))$. We note that even if the MOOP is linear, the MPCCs (6) and (8) are nonconvex optimization problems because of the presence of the complementarity constraints and the upper bound on $\eta \le \|f(x_k) - f(x_l)\|_2^2$. Thus, in practice we can at best hope to find a local solution.

Numerical experience has shown that it can be advantageous to work with a componentwise definition of η. Thus, the goal programming version becomes

$$\begin{cases} \underset{x,y,z,u,\eta}{\text{maximize}} & \sum_{i=1}^{p} \eta_i \\ \text{subject to} & \eta_i \le (f_i(x_k) - f_i(x_l))^2 \\ & \quad \forall 0 \le k, l \le q, k \ne l \text{ and } \forall i = 1, \dots, p, \\ & 0 \le x_l \perp \nabla((1, u_l)^T f(x_l)) - \nabla c(x_l) y_l \ge 0 \\ & \qquad\qquad \forall l = 1, \dots, q, \\ & 0 \le y_l \perp c(x_l) \ge 0 \quad \forall l = 1, \dots, q, \\ & 0 \le u_l \perp \hat{f}(x_l) \le z_l. \end{cases} \tag{9}$$

Similarly we can define componentwise versions with the first-order conditions of SUM(w_k). This new MPCC approach can be generalized easily by using other single-objective characterizations of Pareto points. Many algorithmic choices and variants are possible and can be used to tackle multiobjective optimization problems within the framework of equilibrium constraints.

2.2. Theoretical Foundation of New Approach

We start by recalling that under Assumptions A1–A4, the first-order conditions of SUM(w) and GOAL(z) characterize a Pareto point. This result is a direct corollary of Theorem 1.2.

Corollary 2.1. *Let Assumptions* A1–A4 *hold. Then it follows that* (x^*, y^*) *is a Pareto point if*

1. (x^*, y^*) *solves the first-order conditions of SUM*(w) *for some weights* $w \ge 0$ *with* $e^T w = 1$, *or*
2. (x^*, y^*, u^*) *solves the first-order conditions of GOAL*(z) *for some goals* z.

Clearly, the solution of the bilevel program (5) gives rise to a set of Pareto points.

Proposition 2.2. *Let Assumptions* A1–A4 *hold. Then it follows that if*

1. $(x_k^*, y_k^*, w_k^*, \eta^*)$ *solves problem* (6), *then* (x_k^*, f_k^*) *are Pareto points of MOOP;*
2. $(x_k^*, y_k^*, u_k^*, z_k^*, \eta^*)$ *solves problem* (8), *then* (x_k^*, f_k^*) *are Pareto points of MOOP.*

Moreover, in each case, if η^* *is the global maximizer, then* η^* *maximizes the uniformity of the discrete representation of the Pareto set.*

What makes this new approach practical is the fact that the MPCCs can be solved reliably and efficiently as nonlinear programs (NLPs) (Anitescu 2005, Fletcher et al. 2006). For example, a suitable NLP formulation of the MPCC (6) is given by

$$\begin{cases} \underset{x,y,w,s,t\ge 0,\eta}{\text{maximize}} & \eta \\ \text{subject to} & \eta \le \|f(x_k) - f(x_l)\|_2^2 \quad \forall 0 \le k, l \le q, k \ne l, \\ & e^T w_l = 1 \quad \forall l = 1, \dots, q, \\ & s_l = \nabla(w_l^T f(x_l)) - \nabla c(x_l) y_l \quad \forall l = 1, \dots, q, \\ & x_l \ge 0, \quad s_l \ge 0, \quad x_l^T s_l \le 0 \quad \forall l = 1, \dots, q, \\ & t_l = c(x_l) \quad \forall l = 1, \dots, q, \\ & y_l \ge 0, \, t_l \ge 0, \, y_l^T t_l \le 0 \quad \forall l = 1, \dots, q, \end{cases} \tag{10}$$

where we have introduced slacks to obtain a numerically favorable formulation. It is well known that (10) violates the Mangasarian-Fromowitz constraint qualification at any feasible point (Chen and Florian 1995) because of the presence of the bilinear terms $x_l^T s_l \le 0$ and $y_l^T t_l \le 0$. Recently, however, Fletcher et al. (2006) have shown that any stationary point of the NLP (10) is a strongly stationary point (Scheel and Scholtes 2000) of the MPCC (6) and vice versa. This fact has been used to show that standard NLP solvers can tackle MPCCs reliably and efficiently provided an MPCC-LICQ holds (Anitescu 2005, Benson et al. 2006, Fletcher et al. 2006, Fletcher and Leyffer 2004, Leyffer 2003, Leyffer et al. 2006, Liu et al. 2006, Raghunathan and Biegler 2005). We note that similar results hold for other nonlinear formulations of the complementarity conditions (Leyffer 2006).

Next, we analyze the complementarity constraints further. We prove that if the single-objective NLP satisfies the linear-independence constraint qualification and a second-order sufficient condition, then the constraint normals of the first-order conditions are linearly independent. We state this result in a slightly more general form.

Proposition 2.3. *Consider the general single-objective nonlinear program*

$$\underset{x}{\text{minimize}}\, F(x) \quad \text{subject to} \quad G(x) \ge 0, \tag{11}$$

where F: $\mathbb{R}^n \to \mathbb{R}$ *and* G: $\mathbb{R}^n \to \mathbb{R}^m$ *are twice continuously differentiable. Let* x^* *be a solution that satisfies the linear-independence constraint qualification and a second-order sufficient condition. Then it follows that the active constraint normals of the mixed complementarity problem corresponding to the first-order conditions of* (11),

$$\begin{gathered} \nabla F(x) + \nabla G(x)^T y = 0, \\ 0 \le y \perp G(x) \ge 0, \end{gathered} \tag{12}$$

are linearly independent.

Proof. Let (x^*, y^*) be the optimal primal-dual solution of (11). The active constraints are defined as

$$\mathcal{A}^* := \{i\colon G_i(x^*) = 0\}.$$

We introduce the following notation,

$$A_* := [\nabla G_i^*]_{i\in\mathcal{A}^*}, \quad W_* := \nabla^2 F^* + \sum_{i=1}^{m} y_i^* \nabla^2 G_i^*, \quad \text{and}$$
$$I_* := [e_i]_{i\notin\mathcal{A}^*},$$

to denote the various parts of the active constraint normals. To prove the result, we need to show that the basis matrix

$$B_* := \begin{bmatrix} W_* & A_* \\ A_*^T & 0 \\ 0 & I_* \end{bmatrix}$$

has linearly independent columns.

The proof is by contradiction, and we assume that there exists $s = (s_x, s_y) \neq 0$ such that $B_* s = 0$. The last two equations imply that $s_y = 0$ and that $A_*^T s_x = 0$, and we are left with $W_* s_x = 0$. Premultiplying by s_x gives $s_x^T W_* s_x = 0$, which contradicts the second-order sufficient condition, namely, that $s_x^T W_* s_x > 0$ for all $s_x \neq 0$ such that $A_*^T s_x = 0$. Thus, the columns of B_* are linearly independent. □

Proposition 2.3 ensures that the complementarity constraints of (6) and (8) satisfy an MPCC-LICQ condition (Scheel and Scholtes 2000) whenever the underlying single-objective NLP satisfies an LICQ and a second-order sufficient condition. However, this result does not prove that the MPCC (6) or (8) satisfies an MPCC-LICQ, because degeneracy may exists in the constraints that define the maximum uniformity, η.

One limitation of our approach is the fact that even linear MOOPs such as

$$\begin{cases} \underset{x}{\text{maximize}} & C^T x \\ \text{subject to} & A^T x \geq b, \\ & x \geq 0, \end{cases}$$

lead to nonconvex NLP formulations. The reason for the nonconvexity of (6) is the presence of the constraints $\eta \leq \|C^T x_k - C^T x_l\|_2^2$ and the presence of the complementarity constraints. Thus, in general, we cannot expect to find the global minimum of (6). However, numerical experience presented in the next section shows that our approach is promising.

Another limitation of our approach is the requirement that the MOOP be convex (Assumption A4). Consider the MOOP

$$\underset{x}{\text{minimize}}[(x^2-1)^2, (x^2-4)^2]. \tag{13}$$

It follows that $f_1(x) = (x^2 - 1)^2$ has two minimizers at $x = \pm 1$ and a maximum at $x = 0$. Likewise, $f_2(x) = (x^2 - 4)^2$ has two minimizers at $x = \pm 2$ and a maximum at $x = 0$. However, the MPCC (6) cannot distinguish between minima and maxima. For this example, the MPCC approach generates the two "Pareto" points $x_1 = 1$ and $x_2 = 0$, which maximize the uniformity of representation. However, the second Pareto point, $x_2 = 0$, clearly corresponds to a maximizer of the lower-level problem. Thus, for nonconvex MOOPs, the MPCC approach has a bias toward generating both minima and maxima of the MOOP, because such a choice of controls maximizes the separation between the objective values. Note that we could still use the bilevel formulation (5), but that would rule out the use of standard NLP solvers.

3. Numerical Experience

This section presents our numerical results. To test our approach, we have collected test problems from the literature and generated random quadratic MOOPs. All test problems and the random generator are available at http://www.mcs.anl.gov/~leyffer/MOOP/.

3.1. Obtaining Good Starting Points

Early numerical experience showed that the NLP solvers may fail to find a feasible point to the MPCC formulations (6) or (8). The failures were caused by the nonconvexity of the MPCC formulation, which can cause the NLP solvers to converge to a local minimum of the constraint violation.

Hence we have adopted the following strategy for finding initial feasible points. We first fix the weights, or goals, and solve the resulting NCP using `PATH` (Dirkse and Ferris 1995, Ferris and Munson 2000). This is a standard strategy for solving complex MPCCs and is readily implemented in `AMPL` (Fourer et al. 2003) by using the named model facility.

Another difficulty that arose for some problems is that different weights can give rise to the same Pareto point. Unfortunately, this corresponds to a stationary point of the MPCC (6) and (8) with $\eta^* = 0$. Thus we ran the NCP solver for different choices of weights until we found a set of Pareto points with $\eta \neq 0$. This initial NCP solution also provides an initial guess at the maximum uniformity. The results for the start-up with PATH and their computational cost are included in Table 3.

3.2. Description of Test Problems and Solvers

Table 1 shows the name of the test problem, the number of variables n, the number of constraints m, the number of objectives p, the type of objectives and constraints, and whether or not the problem is convex (C) in the final column. We note that our collection contains the nonconvex problems ABC-comp, ex002, and ex004.

Table 1 Multiobjective Optimization Problem Characteristics

Name	n	m	p	Source	Objective	Constraints	C
ABC-comp	2	3	2	Hwang and Masud (1979)	Quadratic	Bilinear	N
ex001	5	3	2	Das and Dennis (1997)	Quadratic	Quadratic	Y
ex002	5	2	2	Wang and Renaud (1999)	Quadratic	Nonlinear	N
ex003	2	2	2	Tappeta and Renaud (1999)	Quadratic	Nonlinear	Y
ex004	2	3	2	Oliveira and Ferreira (2000)	Nonlinear	Linear	N
ex005	2	0	2	Hwang and Masud (1979)	Nonlinear	Bounds	Y
hs05x	5	3	3	Hock and Schittkowski (1981)	Quadratic	Linear	Y
liswetm	7	5	2	Li and Swetits (1993)	Quadratic	Linear	Y
MOLPg-1	8	8	3	Steuer (1986)	Linear	Linear	Y
MOLPg-2	12	16	3	Steuer (1986)	Linear	Linear	Y
MOLPg-3	10	14	3	Steuer (1986)	Linear	Linear	Y
MOQP[01-03]	20	10	3		Quadratic	Linear	Y

Problems hs05x and liswetm are constructed from several academic NLP test problems that have the same constraints and different objective functions. We have also written a random MOOP generator that generates multiobjective quadratic programs with linear constraints. The generator is written in `matlab` and generates large sparse problems that are output in `AMPL` format. The Hessian matrix is forced to be positive definite by adding a suitably large multiple of the identity to the diagonal. This ensures that the resulting MOOPs are convex.

Table 2 shows the size of the NCP and the various MPCC formulations for $q = 10$ Pareto points. Here, n, m, and r refer to the number of variables, the number of constraints, and the number of complementarity conditions, respectively. As expected, the growth in terms of the number of variables compared with the NCP formulation is modest, while the increase in the number of constraints corresponds to the addition of the constraints $\eta \leq \cdots$, which is of order q^2. We also note that formulation (8) gives rise to the largest MPCCs because we have added multipliers of the goal constraints.

We note that the problem sizes differ for the NCP and MPCC formulations. The reason is that the NCP fixes the upper-level variables, w_l, to obtain a feasible solution. The differences between the MPCC formulations are due to the fact that in (7) we have fixed one weight from (6) to one. Moreover, (8) contains variables corresponding to the objective multipliers u_l, in addition to the goals z_l.

The problems are formulated in `AMPL`, and the initial NCPs are solved by using PATH. PATH implements a generalized Newton method that solves a linear complementarity problem to compute the search direction. The MPCCs are solved by using filterSQP (Fletcher and Leyffer 2002, 2004), which automatically reformulates the complementarity constraints as nonlinear equations. This solver implements a sequential quadratic programming algorithm with a filter to promote global convergence (Fletcher et al. 2002).

3.3. Detailed Numerical Results

Table 3 summarizes our numerical experience. The table shows the number of major (Newton) iterations and the final value of η, which can be taken as an indication of the quality of the computed representation of the Pareto set. We provide results only for the NCP version of SUM(w); results for the other formulations are similar, and we merely mention the NCP run to illustrate the start-up cost (the PATH

Table 2 Characteristics of NCP and MPCC Formulations

	NCP			(6)			(7)			(8)		
Name	n	m	r	n	m	r	n	m	r	n	m	r
ABC-comp	51	51	50	71	105	50	61	95	50	71	150	60
ex001	80	80	10	101	135	10	91	125	10	101	180	20
ex002	70	70	50	91	134	50	81	115	50	91	170	60
ex003	40	40	40	61	104	40	51	94	40	61	140	50
ex004	40	40	40	61	104	40	51	94	40	51	130	40
ex005	20	20	20	41	84	20	31	74	20	41	120	30
hs05x	80	80	50	111	135	50	101	170	50	121	190	70
liswetm	121	121	50	141	184	50	131	174	50	141	220	60
MOLPg-1	160	160	160	191	260	160	181	250	160	201	290	160
MOLPg-2	291	291	280	321	390	280	311	380	280	331	420	280
MOLPg-3	261	261	240	291	360	240	281	350	240	301	390	240
MOQP[01-03]	311	311	300	341	410	300	331	400	300	351	420	320

Table 3 Numerical Results for NCP and MPCC Formulations

	NCP/PATH		(6)		(7)		(8)	
Name	Iteration	η^*	Iteration	η^*	Iteration	η^*	Iteration	η^*
ABC-comp	5	1.154	44	28.43	11	28.43	36	[I]
ex001	4	1.210E−2	28	1.648	14	1.648	4	1.648
ex002	22	4.723E−7	21	3.245E−6	78	2.107 [L]	45	3.206 [L]
ex003	23	1.944E−6	22	2.912E−4	11	4.478E−1	1	8.449 E−2
ex004	6	3.577E−2	14	6.441E−1	49	6.441E−1	13	8.150E−1 [L]
ex005	2	8.702E−5	405	1.656E−2	1,000	1.540E−2	9	1.496E−1
hs05x	1	1.615E−1	237	323.2	374	323.3	193	316.5
listwetm	0	1.847E−2	48	2.830E−1	217	2.830E−1	62	2.210E−4
MOLPg-1	7	0	1	0	1	0	10	52.11
MOLPg-2	5	0	4	0	6	0	18	3.623
MOLPg-3	7	0	5	0	7	0	24	15.74
MOQP-01	6	243.3	296	5,466	897	5,622	452	3,117
MOQP-02	[S]		262	4,046	1,000	[I]	707	5,235
MOQP-03	9	69.66	1,000	[I]	1,000	[I]	488	1,160

solve) and the improvement in uniformity that can be achieved. The iteration limit for all solvers is 1,000 major iterations.

Failures of the solvers are indicated by the following: [S] indicates termination with segmentation fault; [I] means that the solver failed to find a feasible point. Unfortunately, this latter outcome is difficult to avoid because the MPCC are nonconvex. Some of these failures may be due to the relative immaturity of the computational tools for solving MPCCs. In our experience, warm-starting the MPCCs from a solution of the initial NCP greatly improves the likelihood of finding a feasible MPCC solution. Runs for which the solver failed to converge within the limit of 1,000 iterations are identified by 1,000 in the Iteration column. We note that for only two problems does the solver fail in this way, namely, ex005 and MOQP-3.

The results for the nonconvex MOOPs are interesting. As indicated earlier (see (13)), nonconvex MOOPs can have the undesirable effect of increasing η by placing points x_k at local maxima. This corresponds to a failure of the formulation, because the first-order condition cannot capture the difference between minima and maxima. We indicate this failure by [L]. It occurs on the two nonconvex examples (ex002 and ex004) and corresponds to a failure of the Karush-Kuhn-Tucker (KKT) conditions to characterize local minima. Despite this shortcoming, we are able to find valid approximations of the Pareto set for the nonconvex example ABC-comp.

We also observe that the approach can generate weakly dominated Pareto points. Figure 3 shows the computed Pareto set for ex004. On the left is the Pareto set for the uniform weights; on the right is the Pareto set with optimal weights, which clearly provides greater uniformity. We note, however, that the MPCC approach identifies one weakly dominated Pareto point, namely, the point $(2, 6)$, in the right plot, which is clearly dominated by the lower point.

The results of MOLPg-* for NCP, (6), and (7) are also of interest. In these cases, the optimal uniformity is $\eta^* = 0$, which corresponds to two or more coalescing Pareto points. Typically, however, the MPCC approaches are able to improve the uniformity by orders of magnitude compared with the uniform representation corresponding to NCP. In our experiments, (7) obtained better uniformity than (6) on two examples: ex003 and MOQP-01. We believe that there may be numerical reasons that make (7) preferable on some examples.

The results in Table 3 show that the MPCC formulation based on goal programming, (8), is clearly superior to the other two formulations: the formulation based on goal programming is the only formulation that achieves positive separation between all Pareto points for the MOLPg problems. The better performance of the goal-programming-based approach (8) is not surprising, given its superior theoretical properties. We also note that in our experiments we arbitrarily fixed the first objective as the main objective in GOAL(z). The approach may be even more robust if we allow ourselves to cycle through the objectives in turn.

We are interested in discovering how close the MPCC approach gets to the global maximum. Unfortunately, it would be prohibitive to run global optimization software on these test problems. Thus, we have conducted a small experiment by running the global optimization solver `Baron` (Tawarmalani and Sahinidis 2002, Sahinidis 2000) on the smallest MOOP, namely, the GOAL(z) formulation of ex005 for a reduced number of Pareto points ($q = 8$). `Baron` is a branch-and-reduce solver that generates valid bounds by constructing (local) outer approximations of the nonconvex functions that are refined in a branch-and-bound tree search. Even though `Baron` did not find the global maximum in 36,000 seconds CPU time, the

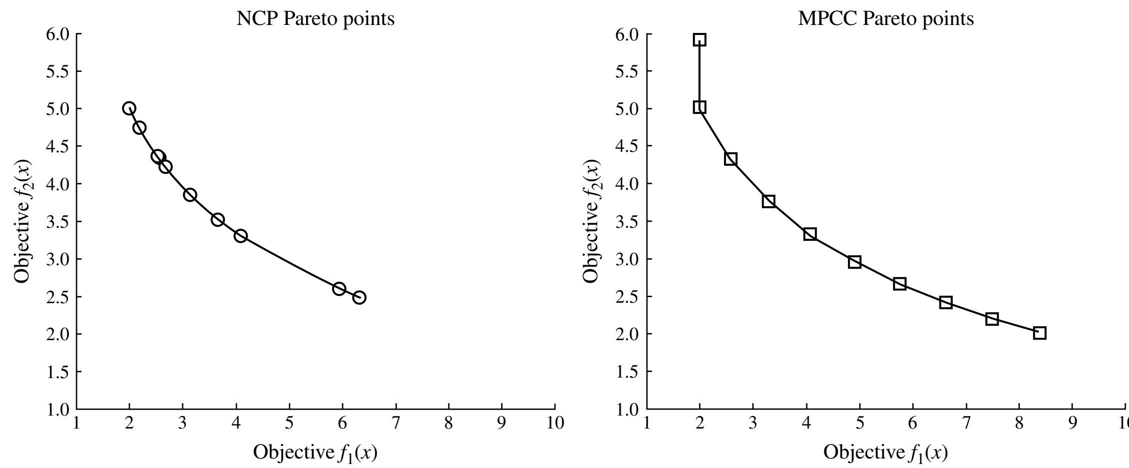

Figure 3 **NCP (Left) and MPCC (Right) Pareto Sets for ex004**
Note. The latter has one spurious point.

results are still of interest. In particular, Baron gives the following valid bounds on the maximum: $0.3032 \le \eta^* \le 0.7055$. Table 4 gives the details of the runs, where we have added the results of the uniform distribution for comparison.

The approximate Pareto set for $q = 8$ is shown in Figure 4. The left plot shows the Pareto set of the NCP approach, which does not attempt to maximize uniformity. The plot on the right compares the Pareto sets obtained with the MPCC approach (crosses) and the global optimization solver Baron (circles). Surprisingly, the local MPCC solver finds a better solution than the global optimization approach. Of course, it is not clear whether even this solution is a global maximum, but the bounds obtained from Baron are encouraging. We note also that Baron finds its lower bound (candidate solution) after 2.54 seconds and spends the remainder of the time searching the tree for a better solution.

4. Conclusions and Outlook

We have presented a new approach to solving multiobjective optimization problems that approximates a maximally uniform representation of the Pareto set. We show how this problem can be formulated as a mathematical program with complementarity constraints, and we present three formulations based on convex sum and goal-programming single-objective formulations of MOOP. Preliminary numerical results are encouraging, especially for the approach based on goal programming.

Table 4 **Comparison of Solution for ex005 from NCP, MPCC, and Baron**

	NCP	MPCC	Baron
CPU time (s)	0.0	0.04	36,000
Iterations	14	33	80,488
Uniformity η^*	4.91E−4	0.3580	0.3032

Our new MPCC approach can be generalized easily by using other single-objective characterizations of Pareto points. Many algorithmic choices and variants are possible and can be used to tackle multiobjective optimization problems within the framework of equilibrium constraints. More numerical experience is needed to decide which of these schemes works best under which circumstances.

Important open questions do remain, however. For example, the reformulation requires the user to form the first-order conditions of a single-objective formulation of MOOP, a process that (from our experience) is prone to error. In addition, the first-order conditions are necessary and sufficient only if the MOOP is convex. We have observed examples where a lack of convexity results in spurious Pareto points being found by our approach.

Some of these limitations can be overcome by better MPCC solvers that preserve local minima. However, such an approach would make it harder to exploit the available NLP solver technology. The requirement that the user form first-order conditions can be overcome by developing extensions to AMPL that allow bilevel optimization models. This is a nontrivial task, however, because AMPL would then have to provide derivatives up to third order for the Hessian matrices used in the NLP solvers.

Another limitation of our approach is the $\mathcal{O}(q^2)$ number of constraints that define the uniformity η in (6) and (8). This limits the applicability of our approach to a mere 10 Pareto points. If more Pareto points were needed, then we could apply ideas similar to domain decomposition to partition the objective space, and then apply our approach to each partition.

Ultimately, we believe that our technique can be incorporated into interactive MOOP solution approaches such as www-nimbus (Miettinen and Mäkelä

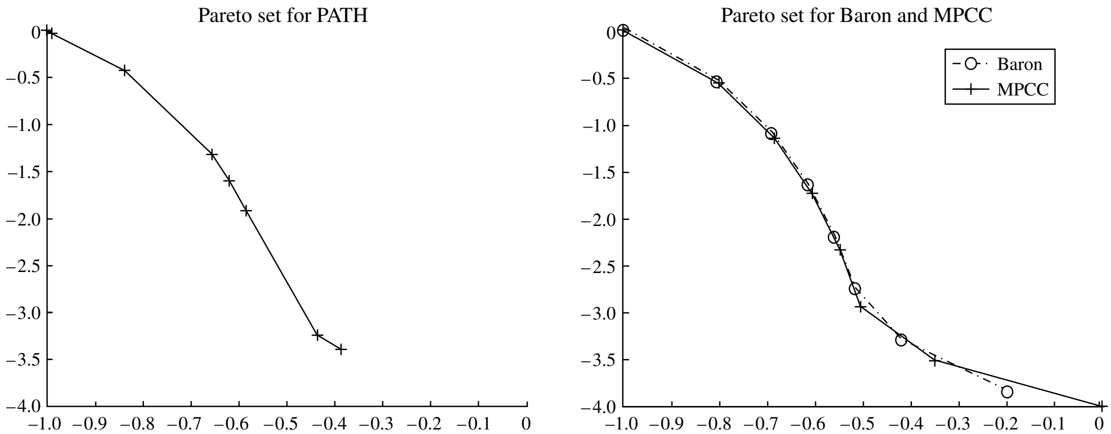

Figure 4 **NCP (Left) and Baron and MPCC (Right) Pareto Sets for ex005**

2000). The advantage of our approach is that it provides a broader picture of the Pareto set. By allowing the user to interact with this representation, we believe that our approach can be made more robust and less susceptible to problems caused by nonconvexities.

Acknowledgments

The author is grateful to the area editor and to two anonymous referees for their insightful comments that improved the presentation of the manuscript. This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy, under Contracts DE-AC02-06CH11357 and DE-FG02-05ER25694.

References

Anitescu, M. 2005. Global convergence of an elastic mode approach for a class of mathematical programs with complementarity constraints. *SIAM J. Optim.* **16**(1) 120–145.

Benson, H., A. Sen, D. Shanno, R. Vanderbei. 2006. Interior-point algorithms, penalty methods and equilibrium problems. *Computational Optim. Appl.* **34**(2) 155–182.

Chen, Y., M. Florian. 1995. The nonlinear bilevel programming problem: Formulations, regularity and optimality conditions. *Optimization* **32** 193–209.

Clarke, F. H. 1983. *Optimization and Nonsmooth Analysis*. Wiley, New York. Reprinted by SIAM, 1990.

Das, I., J. E. Dennis. 1997. A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. *Structural Multidisciplinary Optim.* **14**(1) 63–69.

Das, I., J. E. Dennis. 1998. Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems. *SIAM J. Optim.* **8**(3) 631–657.

Dirkse, S. P., M. C. Ferris. 1995. The path solver: A nonmonotone stabilization scheme for mixed complementarity problems. *Optim. Methods Software* **5** 123–156.

Ehrgott, M. 2005. *Multicriteria Optimization, Lecture Notes in Economics and Mathematical Systems*, Vol. 491, 2nd ed. Springer Verlag, New York.

Ferris, M. C., T. S. Munson. 2000. Complementarity problems in GAMS and the PATH solver. *J. Econom. Dynam. Control* **24**(2) 165–188.

Fletcher, R., S. Leyffer. 2002. Nonlinear programming without a penalty function. *Math. Programming* **91** 239–270.

Fletcher, R., S. Leyffer. 2004. Solving mathematical program with complementarity constraints as nonlinear programs. *Optim. Methods Software* **19**(1) 15–40.

Fletcher, R., S. Leyffer, P. L. Toint. 2002. On the global convergence of a filter-SQP algorithm. *SIAM J. Optim.* **13**(1) 44–59.

Fletcher, R., S. Leyffer, D. Ralph, S. Scholtes. 2006. Local convergence of SQP methods for mathematical programs with equilibrium constraints. *SIAM J. Optim.* **17**(1) 259–286.

Fourer, R., D. M. Gay, B. W. Kernighan. 2003. *AMPL: A Modelling Language for Mathematical Programming*, 2nd ed. Books/Cole—Thomson Learning, Pacific Grove, CA.

Hillermeier, C. 2001. *Nonlinear Multiobjective Optimization*. Birkhäuser Verlag, Berlin.

Hock, W., K. Schittkowski. 1981. *Test Examples for Nonlinear Programming Codes*. Springer-Verlag, Berlin.

Hwang, C.-L., A. S. M. Masud. 1979. *Multiple Objective Decision Making—Methods and Applications, Lecture Notes in Economics and Mathematical Systems*, Vol. 164. Springer-Verlag, Berlin.

Leyffer, S. 2003. Mathematical programs with complementarity constraints. *SIAG/OPT Views-and-News* **14**(1) 15–18.

Leyffer, S. 2006. Complementarity constraints as nonlinear equations: Theory and numerical experience. S. Dempe, V. Kalashnikov, eds. *Optimization and Multivalued Mappings*. Springer-Verlag, Berlin, 169–208.

Leyffer, S., G. Lopez-Calva, J. Nocedal. 2006. Interior methods for mathematical programs with complementarity constraints. *SIAM J. Optim.* **17**(1) 52–77.

Li, W., J. Swetits. 1993. A Newton method for convex regression, data smoothing and quadratic programming with bound constraints. *SIAM J. Optim.* **3**(3) 466–488.

Liu, X., G. Perakis, J. Sun. 2006. A robust SQP method for mathematical programs with linear complementarity constraints. *Computational Optim. Appl.* **34**(1) 5–33.

Miettinen, K. 1999. *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, Boston.

Miettinen, K., M. M. Mäkelä. 1998a. Proper Pareto optimality in nonconvex problems—Characterization with tangent and normal cones. Report 11/1998, Department of Mathematics, University of Jyväskylä, Jyväskylä, Finland.

Miettinen, K., M. M. Mäkelä. 1998b. Theoretical and computational comparison of multiobjective optimization methods NIMBUS and RD. Report 5/1998, Department of Mathematics, University of Jyväskylä, Jyväskylä, Finland.

Miettinen, K., M. M. Mäkelä. 2000. Interactive multiobjective optimization system WWW-NIMBUS on the Internet. *Comput. Oper. Res.* **27** 709–723.

Oliveira, S. L. C., P. A. V. Ferreira. 2000. Bi-objective optimization with multiple decision-makers: A convex approach to attain majority solutions. *JORS* **51** 333–340.

Raghunathan, A., L. T. Biegler. 2005. An interior point method for mathematical programs with complementarity constraints (MPCCs). *SIAM J. Optim.* **15**(3) 720–750.

Rustem, B. 1998. *Algorithms for Nonlinear Programming and Multiple Objective Decision*. Wiley, Chichester, UK.

Sahinidis, N. V. 2000. *Branch and Reduce Optimization Navigator*. Version 4.0 ed. University of Illinois at Urbana-Champaign, Urbana.

Sayin, S. 2000. Measuring the quality of discrete representations of efficient sets in multiple objective mathematical programming. *Math. Programming* **87**(3) 543–560.

Scheel, H., S. Scholtes. 2000. Mathematical program with complementarity constraints: Stationarity, optimality and sensitivity. *Math. Oper. Res.* **25** 1–22.

Stadler, W., ed. 1988. *Multicriteria Optimization in Engineering and in the Sciences*. Plenum Press, New York.

Steuer, R. E. 1986. *Multiple Criteria Optimization: Theory, Computation and Applications*. John Wiley & Sons, New York.

Tappeta, R. V., J. E. Renaud. 1999. Interactive multiobjective optimization procedure with local preferences. *Proc. 3rd WCSMO, Buffalo, NY*, http://www.eng.buffalo.edu/Research/MODEL/wcsmo3/proceedings/.

Tawarmalani, M., N. V. Sahinidis. 2002. *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software, and Applications, Nonconvex Optimization and Its Applications*, Vol. 65. Kluwer Academic Publishers, Boston.

Wang, J.-F., J. E. Renaud. 1999. Automatic differentiation in multi-objective collaborative optimization. *Proc. 3rd WCSMO, Buffalo, NY*, http://www.eng.buffalo.edu/Research/MODEL/wcsmo3/proceedings/.