AN ADAPTIVE PRIMAL-DUAL WARM-START TECHNIQUE FOR QUADRATIC MULTIOBJECTIVE OPTIMIZATION

DANIEL MOLZ, CHRISTOPH HEERMANN, & JÖRG FLIEGE*
SCHOOL OF MATHEMATICS
UNIVERSITY OF BIRMINGHAM
BIRMINGHAM
U.K.

EMAIL FLIEGEJ@BHAM.AC.UK, TEL +44-121-414-6200, FAX +44-121-414-3389

Abstract. We present a new primal-dual algorithm for convex quadratic multicriteria optimization. The algorithm is able to adaptively refine the approximation to the set of efficient points by way of a warm-start interior-point scalarization approach. Results of this algorithm when applied on a three-criteria real-world power plant optimization problem are reported, thereby illustrating the feasibility of this approach when used in practice.

Keywords. warm-start, interior-point method, multicriteria optimization, multiobjective optimization.

AMS Classification. 90C51, 90C29.

1. Introduction.

Multicriteria optimization problems are a class of difficult optimization problems in which several different objective functions have to be taken care of at the same time. It will usually be the case that no single point will minimize all of the several objective functions given at once. Therefore, we are in search for so-called *efficient* points, i. e. feasible points for which there does not exist a different feasible point with the same or smaller objective function values such that there is a strict decrease in at least one objective function value. Since two different efficient points will usually be not only quite different from each other in terms of objective function values, but also incomparable with each other, we have to gain as much information as possible about the solution set of a given problem, preferably by constructing a well-defined approximation to it. This is the subject of this paper.

Applications of multicriteria optimization can be found in various areas, e. g. in engineering design [9, 8, 27], space exploration [32], antenna design [26, 25], management science [10, 3, 20, 30, 34, 1], environmental analysis [29, 11, 12], cancer treatment planning [23], bilevel programming [16], location science [4], statistics [5], etc. A further application, power plant optimization, is described in Subsection 5.2.

The rest of this paper is as follows. In Section 2, we review in short the problem of solving one single-criteria convex-quadratic optimization problem by an interior-point method, namely by an infeasible point method. Sections 3 and 4, containing the main theoretical results of this paper, consider perturbed optimization problems and a strategy to compute a *warm-start point*, i. e. a starting point for the new, perturbed problem, computed out of an iteration point for the unperturbed problem. Armed with this technique, we are ready to tackle our main problem. After explaining in short what multicriteria optimization is and where the main difficulties lie (Section 5), we come to the main part of the paper, Section 6. There, we describe a new efficient adaptive interior-point technique for solving convex quadratic multicriteria problems and show how it can be applied to a real-world multicriteria optimization problem from power plant control.

2. The Interior-Point Algorithm.

*Corresponding author.

2.1. The Problem. Let there be given a primal quadratic optimization problem (PQP) of the form

$$\begin{array}{ll} \min & \frac{1}{2}x^T Qx + c^T x \\ \text{s. t.} & Ax = b, \\ & x \geq 0, \end{array} \tag{2.1}$$

with $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times n}$ and a symmetric matrix $Q \in \mathbb{R}^{n\times n}$. The vector $x \in \mathbb{R}^n$ represents the primal variables.

A dual problem (DQP) to (PQP) is

$$\begin{array}{ll} \max & -\frac{1}{2}x^T Qx + b^T \lambda \\ \text{s. t.} & -Qx + A^T\lambda + s = c, \\ & s \geq 0, \end{array} \tag{2.2}$$

with the dual variables $s \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^m$. The set of primal-dual feasible 3-tuple $w = (x, \lambda, s)$ is given by

$$\Omega := \{w = (x, \lambda, s) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \mid Ax = b,\ -Qx + A^T\lambda + s = c,\ (x, s) \geq 0\}.$$

For a primal-dual feasible point $w \in \Omega$, the difference between the primal and the dual objective function values is given by

$$s^T x = c^T x + x^T Qx - b^T \lambda \geq 0$$

and is called *duality gap.* Denote by

$$\Omega^0 := \{w = (x, \lambda, s) \in \Omega \mid (x, s) > 0\}$$

the set of *strictly feasible points.*

In the rest of the paper, we make the following three assumptions.

1. The set of primal-dual feasible points is nonempty: $\Omega \neq \emptyset$.
2. The constraint matrix A has full row rank.
3. The matrix Q is positive semidefinite.

Clearly, under these assumptions, a primal-dual point $w = (x, \lambda, s)$ is optimal for (PQP) as well as (DQP) (i. e. x is optimal for (PQP) and (λ, s) is optimal for (DQP)) if the KKT-conditions hold:

$$F(x, \lambda, s) := \begin{bmatrix} -Qx + A^T\lambda + s - c \\ Ax - b \\ SXe \end{bmatrix} = 0, \qquad (x, s) \geq 0. \tag{2.3}$$

Here, as usual, $e := (1, 1, ..., 1)^T \in \mathbb{R}^n$, $S := \text{diag}(s) \in \mathbb{R}^{n\times n}$ and $X := \text{diag}(x) \in \mathbb{R}^{n\times n}$.

2.2. The Algorithm. We now perturb the right hand side of the system of equations $F(x, \lambda, s) = 0$ in the usual way by considering a parameter $\tau > 0$ and the perturbed system

$$F_\tau(x, \lambda, s) := \begin{bmatrix} -Qx + A^T\lambda + s - c \\ Ax - b \\ SXe - \tau e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \qquad (x, s) > 0. \tag{2.4}$$

A Newton step for the nonlinear system of equations $F_\tau(x, \lambda, s) = 0$ amounts in solving the linear system

$$\begin{bmatrix} -Q & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -(SXe - \sigma\mu e) \end{bmatrix}. \tag{2.5}$$

Here, we have used the abbreviations $r_b := Ax - b$ and $r_c := -Qx + A^T\lambda + s - c$. (Note that for these abbreviations, r_b as well as r_c depend on (x, λ, s).) Moreover, we used $\tau = \sigma\mu$ with the duality measure

$$\mu = \frac{x^T s}{n} \tag{2.6}$$

and the centering parameter $\sigma \in]\,0, 1[$. This parameter weights the competing aims of convergence of μ to zero and closeness to the central path. Denote by $\Delta w = (\Delta x, \Delta\lambda, \Delta s)$ the solution of the linear system (2.5).

Basically, we want to execute Newton steps for F_τ with the parameter $\tau > 0$ converging to 0. For prespecified parameters $\gamma \in]\,0, 1\,[$, $\beta \geq 1$, and a starting point (x^0, λ^0, s^0) defining residuals r_b^0 and r_c^0 and a duality measure μ_0, all iterates of the algorithm presented below will lie in the set

$$\mathcal{N}_{-\infty}(\gamma, \beta) := \left\{ (x, \lambda, s) \;\middle|\; \begin{array}{l} \|(r_b, r_c)\|_2 \leq \dfrac{\|({r_b}^0, {r_c}^0)\|_2}{\mu_0}\beta\mu, \; (x, s) > 0, \\ \qquad\qquad x_i s_i \geq \gamma\mu \quad \forall\, i = 1, 2, \ldots, n \end{array} \right\}. \tag{2.7}$$

Here, the inequality $x_i s_i \geq \gamma\mu$ serves the purpose of hindering some products $x_i s_i$ to converge faster to zero than other ones. Moreover, due to $\beta \geq 1$, we have $(x^0, \lambda^0, s^0) \in \mathcal{N}_{-\infty}(\gamma, \beta)$.

In order to stay within the set $\mathcal{N}_{-\infty}(\gamma, \beta)$ with all iterates we will introduce a step length $\alpha > 0$. (Indeed, full Newton steps for F_τ might lead outside of the set.) We will choose α in such a way that with $w^k \in \mathcal{N}_{-\infty}(\gamma, \beta)$ and $w^{k+1} = w^k + \alpha\Delta w$ we still have $w^{k+1} \in \mathcal{N}_{-\infty}(\gamma, \beta)$. For this, define

$$(x(\alpha), \lambda(\alpha), s(\alpha)) := (x, \lambda, s) + \alpha(\Delta x, \Delta\lambda, \Delta s)$$

as well as

$$\mu(\alpha) := x(\alpha)^T s(\alpha)/n.$$

Now we are ready to describe the primal-dual infeasible-point long-step algorithm for quadratic problems, see Algorithm 1, p. 4.

Remark 1. In case of $r_b = 0$ as well as $r_c = 0$, i. e. if the starting point fulfills already the primal-dual equality constraints, we have a *feasible point algorithm*. In this case, all iterates generated are strictly feasible. Such an algorithm for the problem considered here is discussed in [15]. However, finding such a starting point turns out to be rather difficult. Indeed, there exist problems with $\Omega \neq \emptyset$ but $\Omega^0 = \emptyset$. In the algorithm above, we do not need such a feasible starting point, just a point with $(x, s) > 0$.

With respect to convergence, we just note that we can use the convergence proof of Wright [35, p. 110ff] for algorithm IPF (stated there for linear optimization problems)

Input: Problem data for problem (PQP): A, b, c, Q.

Choose $\varepsilon > 0$, $\beta \geq 1$, $\gamma \in]\,0,1\,[$, $\sigma_{\min}$ and $\sigma_{\max}$ with $0 < \sigma_{\min} < \sigma_{\max} < 1/2$ and $(x^0, \lambda^0, s^0) \in \mathcal{N}_{-\infty}(\gamma, \beta)$

$k := 0$

while

$$\mu_k := \frac{(x^k)^T s^k}{n} > \varepsilon \tag{2.8}$$

do

Choose $\sigma_k \in]\sigma_{\min}, \sigma_{\max}]$ and compute r_c^k, r_b^k as well as r_{xs}^k by

$$\begin{array}{rcl} r_c^k & := & -Qx^k + A^T\lambda^k + s^k - c \\ r_b^k & := & Ax^k - b \\ r_{xs}^k & := & S^k X^k e - \sigma_k \mu_k e. \end{array} \tag{2.9}$$

Compute $(\Delta x^k, \Delta\lambda^k, \Delta s^k)$ by solving

$$\begin{bmatrix} -Q & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta\lambda^k \\ \Delta s^k \end{bmatrix} = \begin{bmatrix} -r_c^k \\ -r_b^k \\ -r_{xs}^k \end{bmatrix}. \tag{2.10}$$

Choose α_k as the largest $\alpha \in [\,0,1\,]$ such that

$$(x^k(\alpha), \lambda^k(\alpha), s^k(\alpha)) \in \mathcal{N}_{-\infty}(\gamma, \beta) \tag{2.11}$$

as well as the Armijo condition

$$\mu_k(\alpha) \leq (1 - 0.01\alpha)\mu_k \tag{2.12}$$

holds.

Define

$$(x^{k+1}, \lambda^{k+1}, s^{k+1}) := (x^k(\alpha_k), \lambda^k(\alpha_k), s^k(\alpha_k)).$$

$k := k + 1$

endw

Output: (x^k, λ^k, s^k)

Algorithm 1: Algorithm QIP for solving a convex-quadratic optimization problem subject to linear equality and inequality constraints in standard form.

almost verbatim. The only difference of any interest is that for the pair $(\overline{x}, \overline{s})$ defined in Lemma 6.3 of [35] the equality $\overline{x}^T\overline{s} = 0$ does not hold. Instead, we have $\overline{x}^T\overline{s} \geq 0$. This, however, is sufficient for the rest of the reasoning to work. We summarize the main results in the following two theorems, noting again that all iterates can be infeasible.

THEOREM 1 (Convergence). *Let $\{(x^k, \lambda^k, s^k)\}$ be a sequence constructed by Algorithm QIP. Then,*

1. *the sequence $\{\mu_k\}$ of duality measures converges Q-linearly to zero, and*
2. *the sequence $\{\|(r_b^k, r_c^k)\|_2\}$ of residuals converges Q-linearly to zero.*

THEOREM 2 (Complexity). *Let $\varepsilon > 0$ be given. Suppose that for the starting point (x^0, λ^0, s^0) we have that*

$$(x^0, \lambda^0, s^0) = \zeta(e, 0, e) \quad \textit{and} \quad \zeta \geq \|(x^*, s^*)\|_\infty$$

holds for a solution (x^*, λ^*, s^*) of (PQP) and (DQP). Moreover, let constants $C, \kappa > 0$ be given such that

$$\zeta^2 \leq \frac{C}{\varepsilon^\kappa}.$$

Then, there exists an index K with

$$K = \mathcal{O}(n^2 |\log \varepsilon\,|),$$

such that for the iterates (x^k, λ^k, s^k) constructed by algorithm QIP *we have that*

$$\mu_k \leq \varepsilon \qquad \textit{for all} \quad k \geq K$$

holds.

3. Warm-start Points. Let us consider the primal and the dual problem (PQP) and (DQP) from Subsection 2.1. Both problems, (PQP) as well as (DQP), can be described in a unique way by the 4-tuple

$$d := (A, b, Q, c). \tag{3.1}$$

We define the norm of such a data instance d by the maximum of the 2-norm of the components,

$$\|d\|_2 := \max\{\|A\|_2, \|b\|_2, \|Q\|_2, \|c\|_2\}. \tag{3.2}$$

(Of course, the matrix norms are matrix norms induced by the Euclidean norms in the corresponding vector spaces.)

Let us now consider the data instance $d = (A, b, Q, c)$ as well as the perturbed instance $\tilde{d} := d + \Delta d$ with the perturbation $\Delta d = (0, 0, \Delta Q, \Delta c)$. (We will see later, in Section 5, that we are in exactly such a situation if we want to solve a multicriteria optimization problem.) If we solve the problem described by d with algorithm QIP, all iterates will lie in $\mathcal{N}_{-\infty}(\gamma, \beta)$. If we want to solve the problem represented by $d + \Delta d$ with algorithm QIP, too, we have to use a similar set for the corresponding sequence of iterates.

REMARK 2. To differentiate between variables and parameters for the original problem with data instance d from variables and parameters for the perturbed problem with data instance $\tilde{d} := d + \Delta d$, we will use a tilde ($\tilde{\ }$) on all variables and parameters for the latter to signify the perturbation of the data.

Now let $w = (x, \lambda, s)$ be a primal-dual strictly feasible point of the original problem. We want to construct a *warm-start point*

$$\tilde{w} := w + \Delta w = (x + \Delta x, \lambda + \Delta\lambda, s + \Delta s)$$

of the perturbed problem. This warm start point should have the same residuals $(\tilde{r}_b, \tilde{r}_c)$ as those given by w, i. e. (r_b, r_c), the residuals of w for the original problem. For the problem instance $d + \Delta d$ and the residuals $(\tilde{r}_b, \tilde{r}_c)$, we consider the sets

$$\begin{aligned}\tilde{\Omega} &:= \{(x, \lambda, s) \mid Ax - b = \tilde{r}_b,\ -(Q + \Delta Q)x + A^T\lambda + s - (c + \Delta c) = \tilde{r}_c, \\ &\qquad (x, s) \geq 0\}, \\ \tilde{\Omega}^0 &:= \{(x, \lambda, s) \in \tilde{\Omega} \mid (x, s) > 0\}\end{aligned}$$

as well as

$$\tilde{\mathcal{N}}_{-\infty}(\gamma,\beta) := \Big\{ (x,\lambda,s) \in \tilde{\Omega}^0 \;\Big|\; \|(\tilde{r}_b,\tilde{r}_c)\|_2 \le \frac{\|(r_b^0,\ r_c^0)\|_2}{\mu_0}\beta\tilde{\mu}, \quad x_i s_i \ge \gamma\tilde{\mu}, \quad i=1,2,\dots,n \Big\}. \tag{3.3}$$

Note that the residuals r_b^0 and r_c^0 and the duality gap μ_0 stem from the starting point of the original problem represented by d. Now we are in search of a corrector step $\Delta w = (\Delta x, \Delta\lambda, \Delta s)$ such that the warm start point $\tilde{w}$ defined by $\tilde{w} = w + \Delta w$ is in the set $\tilde{\mathcal{N}}_{-\infty}(\gamma,\beta)$.

With the assumptions from above, we can make the following observations. Due to $\tilde{r}_b = r_b$, we have

$$\begin{aligned} \tilde{r}_b = A\tilde{x} - b = A(x+\Delta x) - b = \underbrace{Ax - b}_{= r_b} + A\Delta x \\ \iff A\Delta x = 0. \end{aligned} \tag{3.4}$$

Moreover, with $\tilde{r}_c = r_c$, it follows that

$$\begin{aligned} \tilde{r}_c &= -(Q+\Delta Q)\tilde{x} + A^T\tilde{\lambda} + \tilde{s} - (c+\Delta c) \\ &= -(Q+\Delta Q)(x+\Delta x) + A^T(\lambda+\Delta\lambda) + (s+\Delta s) - (c+\Delta c) \\ &= \underbrace{-Qx + A^T\lambda + s - c}_{= r_c} - (Q+\Delta Q)\Delta x + A^T\Delta\lambda + \Delta s - \Delta c - \Delta Q x \\ \iff &\quad -(Q+\Delta Q)\Delta x + A^T\Delta\lambda + \Delta s = \Delta c + \Delta Q x. \end{aligned} \tag{3.5}$$

In addition, we want that the duality gap of $\tilde{w}$ is at most as large as w. We will show in the proof of Theorem 3 (p. 7) that this can be achieved by

$$S\Delta x + X\Delta s = 0. \tag{3.6}$$

Taking (3.4), (3.5), and (3.6) together, we see that we need to consider the following system of linear equations:

$$\begin{bmatrix} -\tilde{Q} & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} \Delta c + \Delta Q x \\ 0 \\ 0 \end{bmatrix}. \tag{3.7}$$

Here,

$$\tilde{Q} := Q + \Delta Q.$$

Solving (3.7) for $\Delta w = (\Delta x, \Delta\lambda, \Delta s)$, we get

$$\Delta\lambda = (AM^{-1}A^T)^{-1}AM^{-1}(\Delta c + \Delta Q x), \tag{3.8}$$

$$\Delta s = D^{-2}M^{-1}(\Delta c + \Delta Q x - A^T\Delta\lambda), \tag{3.9}$$

$$\Delta x = -D^2\Delta s. \tag{3.10}$$

Here, we have used the abbreviations

$$D^2 := S^{-1}X$$

and

$$M := \tilde{Q} + D^{-2}.$$

Note, however, that in order for the inclusion

$$\tilde{w} \in \tilde{\mathcal{N}}_{-\infty}(\gamma, \beta) \tag{3.11}$$

to hold, we need strict positivity of all components of $\tilde{x}$ as well as $\tilde{s}$. Unfortunately, this is not necessarily true for $\tilde{w} = w + \Delta w$. Using a step length $\alpha \neq 1$ like in algorithm QIP does not help, either, since only full steps can assure that the residuals stay the same. We are therefore in search for sufficient conditions for (3.11). This is the subject of the next section.

4. Warmstart Criteria. We follow the strategy outlined before in Nunez & Freund [33], Yıldırım & Wright [36], Fliege & Heseler [15], and Heermann, Weyers, and Fliege [7].

4.1. Necessary and Sufficient Conditions. A necessary and sufficient condition for strict positivity of $\tilde{x}$ and $\tilde{s}$ can be found by taking a closer look at (3.10). Indeed, decomposing this equation in its actual components, we get

$$s_i \Delta x_i + x_i \Delta s_i = 0, \qquad i = 1, 2, ..., n,$$

and using $(x, s) > 0$, we see that this is equivalent to

$$\frac{\Delta x_i}{x_i} + \frac{\Delta s_i}{s_i} = 0, \qquad i = 1, 2, ..., n. \tag{4.1}$$

Therefore, $\tilde{x}$ and $\tilde{s}$ are strictly positive componentwise if and only if

$$\frac{|\Delta x_i|}{x_i} < 1 \text{ and } \frac{|\Delta s_i|}{s_i} < 1, \qquad i = 1, 2, ..., n. \tag{4.2}$$

Combining (4.1) and (4.2) we arrive at

$$\left|\frac{\Delta x_i}{x_i}\right| = \left|\frac{\Delta s_i}{s_i}\right| < 1, \qquad i = 1, 2, ..., n,$$

i. e.

$$\left\|X^{-1}\Delta x\right\|_\infty = \left\|S^{-1}\Delta s\right\|_\infty < 1. \tag{4.3}$$

This, now, is a necessary and sufficient condition that a feasible warm-start point has to fulfill.

4.2. Sufficient Conditions. Unfortunately, before we can check (4.3), we have to compute $\Delta\lambda$ as well as Δs, and it might be argued that this is slightly inefficient: in such a scheme, we first compute a warm-start point, then we check for feasibility. The next theorem is a step further on in our search for simple sufficient conditions for feasibility of a warm-start point.

THEOREM 3. *Let d be a problem instance and w be a point for this instance with $(x, s) > 0$ and residuals $(r_b, r_c) \neq 0$. Let a perturbation $\Delta d = (0, 0, \Delta Q, \Delta c)$ be given, let $\Delta w = (\Delta x, \Delta\lambda, \Delta s)$ be a solution to (3.7) and set $\tilde{w} = (\tilde{x}, \tilde{\lambda}, \tilde{s}) = w + \Delta w$. With*

$$T := I - A^T(AM^{-1}A^T)^{-1}AM^{-1},$$

suppose that

$$\left\| \begin{bmatrix} \Delta c \\ \Delta Q x \end{bmatrix} \right\|_\infty < \frac{1}{\|X^{-1}M^{-1}(T,T)\|_\infty}, \tag{4.4}$$

holds. Then,

$$\tilde{w} \in \tilde{\Omega}^0$$

as well as

$$\tilde{x}^T \tilde{s} \leq x^T s \tag{4.5}$$

follows.

Proof. (Cmp. Yıldırım & Wright [37], Proposition 5.1, p. 797f.) Using (3.8) and (3.9) we arrive at the following chain of equalities and inequalities.

$$\begin{aligned}
&\left\|S^{-1}\Delta s\right\|_\infty \\
=& \left\|S^{-1}D^{-2}M^{-1}[\Delta c + \Delta Qx - A^T\Delta\lambda]\right\|_\infty \\
=& \left\|X^{-1}M^{-1}\left[\Delta c + \Delta Qx - A^T\left((AM^{-1}A^T)^{-1}AM^{-1}(\Delta c + \Delta Qx)\right)\right]\right\|_\infty \\
=& \left\|X^{-1}M^{-1}\left[\left(\Delta c - A^T(AM^{-1}A^T)^{-1}AM^{-1}\Delta c\right)\right.\right. \\
&\left.\left. + \left(\Delta Qx - A^T(AM^{-1}A^T)^{-1}AM^{-1}\Delta Qx\right)\right]\right\|_\infty \\
\leq& \left\|X^{-1}M^{-1}\left(I - A^T(AM^{-1}A^T)^{-1}AM^{-1}\right)[I,I]\right\|_\infty \left\| \begin{bmatrix} \Delta c \\ \Delta Qx \end{bmatrix} \right\|_\infty \\
=& \left\|X^{-1}M^{-1}(T,T)\right\|_\infty \left\| \begin{bmatrix} \Delta c \\ \Delta Qx \end{bmatrix} \right\|_\infty .
\end{aligned}$$

Therefore, as long as

$$\left\| \begin{bmatrix} \Delta c \\ \Delta Q x \end{bmatrix} \right\|_\infty < \left\|X^{-1}M^{-1}(T,T)\right\|_\infty^{-1} \tag{4.6}$$

holds, we have $\left\|S^{-1}\Delta s\right\|_\infty < 1$ and therefore $\tilde{w} \in \tilde{\Omega}^0$.

It remains to show (4.5). We have

$$x^T\Delta s + s^T\Delta x = 0.$$

Due to (4.1), we know that Δx_i and Δs_i have different signs for all $i = 1, 2, \ldots, n$. This results in $(\Delta x)^T\Delta s \leq 0$, which in turn leads to

$$\tilde{x}^T\tilde{s} = (x + \Delta x)^T(s + \Delta s) = x^T s + x^T\Delta s + s^T\Delta x + (\Delta x)^T\Delta s \leq x^T s.$$

□

Of course, we need that the warm-start point generated by (3.7) is not only strictly feasible, but also in $\tilde{\mathcal{N}}_{-\infty}(\gamma,\beta)$ for some $\gamma \in]\,0,1\,[$ and a $\beta \geq 1$. More precisely, we need that

$$\tilde{x}_i\tilde{s}_i \geq \gamma\tilde{\mu}, \qquad i = 1, 2, \ldots, n,$$

and

$$\|(\tilde{r}_b, \tilde{r}_c)\|_2 \leq \|(r_b^0, r_c^0)\|_2 \frac{\beta\tilde{\mu}}{\mu_0}$$

holds. Before taking a closer look at these inequalities, we consider the following lemma.

LEMMA 4. *Under the assumptions of Theorem 3, define*

$$\theta := 1 - \left\| S^{-1} \Delta s \right\|_\infty . \tag{4.7}$$

Then,

$$\tilde{\mu} \geq \theta \mu. \tag{4.8}$$

Proof. Due to (4.3) and (4.7), we have $\theta > 0$. Moreover, (4.7) can be written as

$$\tilde{x}_i := x_i + \Delta x_i \geq \theta x_i \quad \text{and} \quad \tilde{s}_i := s_i + \Delta s_i \geq \theta s_i, \quad i = 1, 2, \dots, n.$$

According to the proof of Theorem 3, we have that

$$\Delta x_i \Delta s_i \leq 0 \quad \text{for all} \quad i = 1, 2, \dots, n.$$

On the one hand, assuming $\Delta x_i \geq 0$ leads immediately to $\tilde{x}_i \geq x_i$. With (4.7), we get

$$\tilde{x}_i \tilde{s}_i \geq x_i \tilde{s}_i \geq \theta x_i s_i.$$

On the other hand, assuming $\Delta s_i \geq 0$ we get in an analogous way

$$\tilde{x}_i \tilde{s}_i \geq \tilde{x}_i s_i \geq \theta x_i s_i.$$

Taking both cases together, we arrive at

$$\tilde{\mu} \geq \theta \mu.$$

□

COROLLARY 5. *Let all assumptions of Theorem 3 hold and let a* $\theta \in]\,0, 1\,[$ *be given with*

$$\theta \leq 1 - \left\| S^{-1} \Delta s \right\|_\infty . \tag{4.9}$$

Suppose now that

$$\left\| \begin{bmatrix} \Delta c \\ \Delta Q x \end{bmatrix} \right\|_\infty \leq \frac{1 - \theta}{\| X^{-1} M^{-1}(T, T) \|_\infty}$$

holds. Then,

$$\tilde{w} \in \tilde{\Omega}^0$$

as well as $\tilde{\mu} \leq \mu$. *Moreover, if for some parameters* $\gamma_0 \in]\,0, 1\,[$ *and* $\beta_0 \geq 1$ *we have* $w \in \mathcal{N}_{-\infty}(\gamma_0, \beta_0)$, *then*

$$\tilde{w} \in \tilde{\mathcal{N}}_{-\infty}(\theta \gamma_0, \beta_0 / \theta).$$

Proof. The decrease of the duality gap as well as $\tilde{w} \in \tilde{\Omega}^0$ follows with Theorem 3. Now let $w \in \mathcal{N}_{-\infty}(\gamma_0, \beta_0)$. Using (4.9) and (4.5), we get

$$\tilde{x}_i \tilde{s}_i \geq \theta x_i s_i \geq \theta \gamma_0 \mu \geq \theta \gamma_0 \tilde{\mu}.$$

Moreover, due to (4.8), we have

$$\|(\tilde{r}_b, \tilde{r}_c)\|_2 = \|(r_b, r_c)\|_2 \leq \|(r_b^0, r_c^0)\|_2 \frac{\beta_0 \mu}{\mu_0} \leq \|(r_b^0, r_c^0)\|_2 \frac{\frac{\beta_0}{\theta}\tilde{\mu}}{\mu_0}.$$

As a consequence, $\tilde{w} \in \tilde{\mathcal{N}}_{-\infty}(\theta\gamma_0, \beta_0/\theta)$. □

Up to now, we have just found criteria which make it possible to check if, starting with a prespecified perturbation, a given point (i. e. an iterate) can be used to construct a warm-start point. For the complexity analysis still to follow (see Subsection 4.3), we want to couple the size of a possible perturbation with the duality gap of a given point. We prepare the road with some concepts and some preliminary results.

The set of data instances for which there exists a strictly feasible point is denoted by

$$\mathcal{L} = \{(A, b, Q, c) \mid \exists\, x, \lambda, s : (x, s) > 0, Ax = b, -Qx + A^T\lambda + s = c\}.$$

Let the complement of $\mathcal{L}$ be denoted by $\mathcal{L}^{\mathcal{C}}$. The set $\mathcal{L}^{\mathcal{C}}$ contains all those data instances for which either (PQP) or (DQP) or both do not have a strictly feasible point. Denote the boundary between $\mathcal{L}$ and $\mathcal{L}^{\mathcal{C}}$ by

$$\mathcal{B} := \mathrm{cl}(\mathcal{L}) \cap \mathrm{cl}(\mathcal{L}^{\mathcal{C}}).$$

Due to $(0, 0, 0, 0) \in \mathcal{B}$ we have $\mathcal{B} \neq \emptyset$. A data instance $d \in \mathcal{B}$ is called *ill-posed*: an arbitrary small perturbation Δd can result in $d + \Delta d \in \mathcal{L}$ or $d + \Delta d \in \mathcal{L}^{\mathcal{C}}$. The *distance to ill-posedness* is defined by

$$\rho(d) := \inf\{\|\Delta d\|_2 \mid d + \Delta d \in \mathcal{B}\}.$$

At last, the *condition number* of a data instance d is defined by

$$\mathcal{C}(d) := \frac{\|d\|_2}{\rho(d)}$$

(resp. $\mathcal{C}(d) = \infty$ in case of $\rho(d) = 0$).

REMARK 3. *Since for $\Delta d = -d$ we have $d + \Delta d \in \mathcal{B}$, we always have $\rho(d) \leq \|d\|_2$ and $\mathcal{C}(d) \geq 1$.*

A sufficient condition for the infeasibility of a convex-quadratic problem is given in the following lemma. As usual, all inequalities there between vectors are to be understood componentwise.

LEMMA 6. *Let there be given matrices $Q \in \mathbb{R}^{n\times n}$ and $A \in \mathbb{R}^{m\times n}$ as well as a vector $c \in \mathbb{R}^n$. Then, the systems*

$$A^T\lambda < c + Qx \tag{4.10}$$

and

$$\begin{array}{rcl} Ax & = & 0, \\ x & \geq & 0, \\ c^Tx + x^TQx & \leq & 0, \\ x & \neq & 0 \end{array} \tag{4.11}$$

cannot be solved simultaneously by vectors $x \in \mathbb{R}^n$ *and* $\lambda \in \mathbb{R}^m$.

Proof. Suppose that both systems can be solved. Then (4.10) is equivalent to

$$c > A^T \lambda - Qx.$$

Substituting this in the third block of (4.11), we get

$$0 \geq c^T x + x^T Q x > \lambda^T \underbrace{Ax}_{=0} - x^T Q^T x + x^T Q x = 0,$$

which is a contradiction. Therefore, the conclusion holds. □

LEMMA 7. *Let there be given a data instance* $d = (A, b, Q, c) \in \mathcal{L}$ *and a point* $w = (x, \lambda, s)$ *with residuals* (r_b, r_c). *Define the data instance* $\hat{d}$ *by* $\hat{d} := (A, b + r_b, Q, c + r_c)$. *Then, it follows that*

$$\|x\|_2 \leq \frac{\max\left\{\|\hat{b}\|_2, |\hat{c}^T x + x^T Q x|, \|P\|_2\right\}}{\rho(\hat{d})}. \tag{4.12}$$

holds for all positive semidefinite matrices $P \in \mathbb{R}^{n \times n}$.

Proof. We will modify the idea of Nunez & Freund [33, p. 11f]. Let there be given the data instance d and the point $w = (x, \lambda, s)$ with $x \neq 0$. The residuals are given by $r_b = Ax - b$ and $r_c = -Qx + A^T \lambda + s - c$. Define

$$\hat{b} := b + r_b \quad \text{and} \quad \hat{c} := c + r_c.$$

Then, w is feasible for the data instance $\hat{d} = (A, \hat{b}, Q, \hat{c})$. Now consider the perturbation $\Delta d = (\Delta A, 0, \Delta Q, \Delta c)$ defined by

$$\begin{aligned}
\Delta A &:= -\hat{b} x^T \frac{1}{\|x\|_2^2}, \\
\Delta c &:= \frac{-|\hat{c}^T x + x^T Q x|}{\|x\|_2^2} x, \\
\Delta Q &:= -\frac{1}{\|x\|_2} P
\end{aligned}$$

for some positive semidefinite matrix P. Then,

$$(A + \Delta A)x = Ax - \hat{b} \frac{x^T x}{\|x\|_2^2} = Ax - \hat{b} = 0$$

and

$$\begin{aligned}
&(\hat{c} + \Delta c)^T x + x^T (Q + \Delta Q) x \\
&= \hat{c}^T x + \Delta c^T x + x^T Q x + x^T \Delta Q x \\
&= \hat{c}^T x + x^T Q x - |\hat{c}^T x + x^T Q x| - \frac{1}{\|x\|_2} x^T P x \\
&\leq 0.
\end{aligned}$$

Using Lemma 6, we see that there does not exist a λ with

$$(A + \Delta A)^T \lambda < \hat{c} + \Delta c + (Q + \Delta Q) x$$

and the dual problem to the data instance $\hat{d} + \Delta d$ is infeasible. Therefore,

$$\begin{aligned}\rho(\hat{d}) \leq \|\Delta d\|_2 &= \max\{\|\Delta A\|_2, \|\Delta c\|_2, \|\Delta Q\|_2\} \\ &= \frac{\max\{\|\hat{b}\|_2, |\hat{c}^T x + x^T Q x|, \|P\|_2\}}{\|x\|_2},\end{aligned}$$

and the conclusion follows with

$$\|x\|_2 \leq \frac{\max\{\|\hat{b}\|_2, |\hat{c}^T x + x^T Q x|, \|P\|_2\}}{\rho(\hat{d})}. \tag{4.13}$$

□

COROLLARY 8. *Let there be given a data instance $d = (A, b, Q, c) \in \mathcal{L}$ and a point $w = (x, \lambda, s)$ with residuals (r_b, r_c). Define the data instance $\hat{d}$ by $\hat{d} := (A, b + r_b, Q, c + r_c)$. Then, it follows that*

$$\|x\|_2 \leq \mathcal{C}(\hat{d}) + \frac{|\hat{c}^T x + x^T Q x|}{\rho(\hat{d})}.$$

Proof. Use $P = 0$ in the last lemma. Then $\Delta Q = 0$. With (4.13) it follows that

$$\begin{aligned}\|x\|_2 &\leq \frac{\|\hat{b}\|_2 + |\hat{c}^T x + x^T Q x|}{\rho(\hat{d})} \\ &\leq \frac{\|\hat{d}\|_2}{\rho(\hat{d})} + \frac{|\hat{c}^T x + x^T Q x|}{\rho(\hat{d})} \\ &= \mathcal{C}(\hat{d}) + \frac{|\hat{c}^T x + x^T Q x|}{\rho(\hat{d})}.\end{aligned}$$

□

COROLLARY 9. *Let there be given a data instance $d = (A, b, Q, c) \in \mathcal{L}$ and a point $w = (x, \lambda, s)$ strictly feasible for d. Then,*

$$\|x\|_2 \leq \frac{\max\left\{\|b\|_2, |c^T x + x^T Q x|\right\}}{\rho(d)} \leq \mathcal{C}(d) + \frac{|c^T x + x^T Q x|}{\rho(d)}.$$

Now we are almost ready to couple the size of a given perturbation with a given duality gap. We just need one more technical result.

LEMMA 10. *Let there be given symmetric and positive semidefinite matrices $M_j \in \mathbb{R}^{n \times n}$ $(j = 1, \ldots, p)$. For $\alpha \in \mathbb{R}^p$, $\alpha \geq 0$, define $M(\alpha) := \sum_{j=1}^{p} \alpha_j M_j$. If A has full row rank, then*

$$\chi(A) := \sup_{\substack{\alpha \geq 0: \\ M(\alpha)\ p.\ d.}} \|(A(M(\alpha))^{-1} A^T)^{-1} A (M(\alpha))^{-1}\|_\infty < \infty$$

holds. Here, "p. d." stands for positive definite.

Proof. This follows directly with inequality (5.6) from the proof of Corollary 5.2 in Forsgren and Sporre [17]. Note that this Corollary 5.2 is based directly on Theorem 5.1 in the same paper. □

To use the result above, we assume in what follows that we have given a data instance $d = (A, b, Q, c)$ and a perturbation of the primal objective function, i. e.

a data perturbation of the form $\Delta d = (0, 0, \Delta Q, \Delta c)$. Furthermore, let us assume that we have given symmetric positive definite matrices $Q_1, Q_2, \ldots, Q_p$ and a vector $\alpha \in \mathbb{R}^p$, $\alpha \geq 0$ such that

$$Q + \Delta Q = \sum_{i=1}^{p} \alpha_i Q_i.$$

We will see in Section 5 that these additional assumptions fit perfectly into the framework of multicriteria optimization.

THEOREM 11. *Let there be given a data instance $d = (A, b, Q, c) \in \mathcal{L}$ and a perturbation $\Delta d = (0, 0, \Delta Q, \Delta c)$. Let $w = (x, \lambda, s)$ be a point with $(x, s) > 0$ and residuals (r_b, r_c) (with respect to d). Define the data instance $\hat{d}$ by $\hat{d} := (A, b+r_b, Q, c+r_c)$. Let A have full row rank and define*

$$\psi(A) := 1 + \|A^T\|_\infty \chi(A).$$

Suppose that $w \in \mathcal{N}_{-\infty}(\gamma, \beta)$ for some $\beta, \gamma > 0$. Then,

$$\frac{1}{\|X^{-1}M^{-1}(T,T)\|_\infty} \geq \frac{\gamma\mu}{2n^{1/2}\left(\mathcal{C}(\hat{d}) + \frac{|\hat{c}^T x + x^T Qx|}{\rho(\hat{d})}\right)\psi(A)}. \tag{4.14}$$

Proof. The proof follows exactly the lines of a similar result in Fliege & Heseler [15, p. 14f] and is therefore omitted here. □

We are now ready to state the main result, connecting the size of the duality gap with the size of a perturbation and vice versa.

Define

$$\|\Delta d\|_\infty := \max\{\|\Delta c\|_\infty, \|\Delta Q\|_\infty\} \tag{4.15}$$

as well as

$$\xi := 1 + \frac{|\hat{c}^T x + x^T Qx|}{\|\hat{d}\|_2} \geq 1.$$

Then,

$$\|x\|_2 \leq \xi\mathcal{C}(\hat{d}) = \mathcal{C}(\hat{d}) + \mathcal{C}(\hat{d})\frac{|\hat{c}^T x + x^T Qx|}{\|\hat{d}\|_2} = \mathcal{C}(\hat{d}) + \frac{|\hat{c}^T x + x^T Qx|}{\rho(\hat{d})}. \tag{4.16}$$

THEOREM 12. *Let there be given parameters γ and γ_0 with $0 < \gamma < \gamma_0 < 1$ and define $\theta := \gamma/\gamma_0$. Define ξ as above. Let w and $\tilde{w}$ as well as β be as in Corollary 5. If*

$$\mu \geq \frac{2\|\Delta d\|_\infty}{\gamma_0(1-\theta)}\xi n^{1/2}\mathcal{C}(\hat{d})\left(\mathcal{C}(\hat{d}) + \frac{|\hat{c}^T x + x^T Qx|}{\rho(\hat{d})}\right)\psi(A)$$

holds, then $\tilde{w} \in \tilde{\mathcal{N}}_{-\infty}(\gamma, \beta)$.

Proof. See Fliege & Heseler [15, p. 15f]. □

4.3. Complexity. Let there be given a problem in form of a data instance d as well as a primal-dual starting point (x^0, λ^0, s^0), possibly infeasible for d. Denote, as usual, the residuals by (r_b^0, r_c^0) and let μ_0 be the duality measure at the starting point. Moreover, let $\varepsilon > 0$ be given. Suppose furthermore that we have already solved that problem with algorithm QIP by computing iterates $w^k = (x^k, \lambda^k, s^k)$, $k = 1, 2, 3, \ldots$. Our complexity analysis has shown that

$$\mu_k \leq \|d\|_2 \varepsilon$$

holds for

$$k \geq K = \mathcal{O}\left(n^2 \log \frac{\mu_0}{\|d\|_2 \varepsilon}\right).$$

Now suppose that we perturb our problem d to $d + \Delta d$ and construct a warm-start point by our warm-start strategy out of the iterate w^j. This warm-start point is then used by algorithm QIP to solve the perturbed problem. Clearly, after

$$k \geq K_{\text{warm}} = \mathcal{O}\left(n^2 \log \frac{\mu_j}{\|d + \Delta d\|_2 \varepsilon}\right)$$

iterations we have that

$$\tilde{\mu}_k \leq \|d + \Delta d\|_2 \varepsilon$$

holds. Here, $\tilde{\mu}_k$ are the duality measures for the perturbed problem, as computed by algorithm QIP. Furthermore, if $\|\Delta d\|_2 \leq \|d\|_2/2$ holds, we can use the estimate

$$\frac{1}{\|d + \Delta d\|_2} \leq \frac{1}{\|d\|_2 - \|\Delta d\|_2} \leq \frac{2}{\|d\|_2}$$

to conclude

$$K_{\text{warm}} = \mathcal{O}\left(n^2 \log \frac{\mu_j}{\|d\|_2 \varepsilon}\right).$$

5. Multicriteria Optimization. In this section, we give a short introduction to multicriteria optimization. We follow roughly the chain of arguments presented in [15] and repeat the main points for the sake of completeness here.

5.1. The Problem. Let there be given $p > 1$ convex quadratic objective functions of the form

$$f_i(x) = \frac{1}{2} x^T Q_i x + c_i^T x, \qquad i = 1, \ldots, p, \tag{5.1}$$

with positive semidefinite matrices $Q_i \in \mathbb{R}^{n \times n}$ and vectors $c_i \in \mathbb{R}^n$ for all i. Moreover, let

$$G := \{x \in \mathbb{R}^n \mid Ax = b\,,\, x \geq 0\} \tag{5.2}$$

be the set of feasible points. We are interested in minimizing simultaneously the functions

$$f_1, \ldots, f_p : G \longrightarrow \mathbb{R} \tag{5.3}$$

on the set G in a sense specified as follows. Define the function $f : G \longrightarrow \mathbb{R}^p$ by $f = (f_1, \ldots, f_p)^T$. The element $y^* \in f(G)$ is called *efficient*, if and only if there is no other $y \in f(G)$ with

$$y_i \leq y_i^* \quad \forall\, i \in \{1, 2, \ldots, p\}$$

and

$$y_k < y_k^* \quad \text{for at least one } k \in \{1, 2, \ldots, p\}.$$

The set of all efficient points of the set $f(G)$ is called the *efficient set*, $E(f(G))$. With the definition of efficiency as above, it becomes clear that in multicriteria optimization we are in search for the whole set $E(f(G))$ and, obviously, for the corresponding set of optimal decision variables $f^{-1}(E(f(G)))$. For typical examples for this type of problem we refer to, e. g., [15]. Note that two efficient points $f(x^{(1)}), f(x^{(2)}) \in E(f(G))$ $(x^{(1)}, x^{(2)} \in G)$ with $f(x^{(1)}) \neq f(x^{(2)})$ are incomparable to each other with respect to the order defined above. By their very nature of being efficient, there exist two indices $i, j \in \{1, \ldots, p\}$ such that $f_i(x^{(1)}) < f_i(x^{(2)})$ and $f_j(x^{(2)}) < f_j(x^{(1)})$. Therefore, just one efficient point cannot capture the possible optimal alternatives we face when solving a multicriteria optimization problem. This clearly shows that human decision makers need information about the whole set $E(f(G))$.

5.2. An Application: Power Plant Optimization. Suppose an energy producer is operating a number of power plants. Subject to various technical constraints, the producer would certainly like to

1. minimize the total cost for producing electricity,
2. minimize the wear and tear of the power plants,
3. minimize the error in meeting the forecast demand.

This is clearly a three-criteria optimization problem, currently only studied as a one- or two-criteria optimization problem [2, 7, 21]. We will model this problem as follows, following closely [2]. Let K be the number of power plants, numbered by $i = 1, \ldots, K$. Index by $t = 1, \ldots, T$ the future time steps for which a demand forecast is given. We assume that between time steps demand and production varies linearly, an assumption usually well met in practice for appropriately chosen time steps [2, 28]. With this, we can use as decision variables $x_{i,t}$ the production (in megawatt) of power plant no. i at time t $(i = 1, \ldots, K,\ t = 1, \ldots, T)$. According to [2, 28], the first objective can be well approximated by

$$f_1((x_{i,t})_{i,t}) := \sum_{i=1}^{K} \sum_{t=1}^{T} p_{i,t} x_{i,t} + q_{i,t} x_{i,t}^2,$$

where $p_{i,t}, q_{i,t} \geq 0$ are known constants $(i = 1, \ldots, k,\ t = 1, \ldots, T)$. Using the vectors $x_t := (x_{1,t}, \ldots x_{K,t})^\top \in \mathbb{R}^K$ $(t = 1, \ldots, T)$ as well as the vector $x_0 \in \mathbb{R}^K$ of current power production, total wear and tear is given by

$$f_2((x_{i,t})_{i,t}) := \sum_{t=0}^{T-1} \|x_{t+1} - x_t\|_2^2.$$

Finally, if $y_t \in \mathbb{R}_+$ $(t = 1, \ldots, T)$ denotes the demand forecast for the time t, the third criterion to be minimized can be written as

$$f_3((x_{i,t})_{i,t}) := \sum_{t=1}^{T} \left(\sum_{i=1}^{K} x_{i,t} - y_t \right)^2.$$

Clearly, all three objectives are convex-quadratic. Constraints are of the form

$$x_{i,t}^{\min} \leq x_{i,t} \leq x_{i,t}^{\max}$$

for prespecified plant-specific constants $x_{i,t}^{\min}, x_{i,t}^{\max} \geq 0$ ($i = 1, \ldots, K$, $t = 1, \ldots, T$) as well as

$$\dot{x}_{i,t}^{\min} \leq x_{i,t} - x_{i,t-1} \leq \dot{x}_{i,t}^{\max}$$

for prespecified plant-specific constants $\dot{x}_{i,t}^{\min}, \dot{x}_{i,t}^{\max} \in \mathbb{R}$ ($i = 1, \ldots, K$, $t = 1, \ldots, T$). Usually, neither K nor T are large, i. e. the problems considered are not large-scale. Nevertheless, these problems are challenging for two reasons. First, as it will be explained in Subsection 5.3 below, solving a multicriteria problem amounts to solving many classical single-criteria optimization problems, typically hundreds or thousands of them. Second, for the application considered here, demand forecasts are constantly updated, so we face a real-time optimization problem. Problems of this type must be solved in no more than a few minutes [2], and decision makers need good knowledge of the set of efficient points to make a meaningful and well-informed decision.

5.3. Scalarization. It is well-known that we can find a point close to $E(f(G))$ of the problem specified by (5.3) by solving the single-objective optimization problem

$$\begin{aligned} &\min \omega^T f(x) \\ &\text{s. t. } x \in G, \end{aligned} \tag{5.4}$$

with ω an arbitrary weight vector from the set

$$Z := \left\{ \omega \in \mathbb{R}^p \;\middle|\; \sum_{i=1}^{p} \omega_i = 1,\ \omega_i > 0\ \forall\, i \in \{1, 2, \ldots, p\} \right\}. \tag{5.5}$$

This approach is often called *scalarization.* (For a discussion of this and other scalarization techniques see e. g. [19, 24, 22, 13].) Indeed, defining the set of *properly efficient points* P by

$$P(f(G)) := \left\{ f(x^*) \;\middle|\; \omega \in Z,\ x^* \in G,\ \omega^T f(x^*) = \min_{x \in G} \omega^T f(x) \right\},$$

it can be shown [18, 31] that

$$P(f(G)) \subseteq E(f(G)) \subseteq \operatorname{cl}(P(f(G))) \tag{5.6}$$

holds. Here, $\operatorname{cl}(\cdot)$ is the closure operator. In fact, this result holds for arbitrary functions $f : G \longrightarrow \mathbb{R}^p$ as long as the set $f(G) + \mathbb{R}^p_+$ is closed and convex. Since we can not distinguish numerically between a set and its closure, we can, therefore, replace E by P in all applications involving convex functions. Turning our attention to (5.3), (5.2), and (5.1), we see that we have to consider several scalar problems of the form

$$\begin{aligned} &\min \frac{1}{2} x^T Q x + c^T x \\ &\text{s. t. } Ax = b, \\ &\qquad x \geq 0, \end{aligned} \tag{5.7}$$

where $Q = \sum_{i=1}^{p} \omega_i Q_i$, $c = \sum_{i=1}^{p} \omega_i c_i$, and $\omega = (\omega_1, \dots, \omega_p)^T \in Z$ is a given parameter or weight vector. As a consequence of the discussion above, we are able to approximate the set $E(f(G))$ as well as $f^{-1}(E(f(G)))$ by solving optimization problems of the form (5.7). These ersatz problems are defined by choosing different weights ω, see (5.4).

The basic idea is now as follows. Our aim is to compute a discrete approximation of the set of efficient points. We have to solve a standard scalar optimization problem for each efficient point we want to compute. The different optimization problems we have to consider can be viewed as perturbations of each other, with vectors of weights $\omega \in Z$ serving as parameters defining the perturbations. We propose to use an adaptive discretization technique for the set of weights Z. Basically, we want to use more parameters in those regions of the parameter space where weight vectors which are close together result in efficient points whose images in the image space $\mathbb{R}^p$ are far apart from each other. But in contrast to [15], where efficient points were calculated first and then a new weight was chosen adaptively by using information about the last optimal function values computed, we will now introduce more parameters $\omega \in Z$ (i. e. more scalar problems) during the solution process of other scalar problems, before these other problems are completely solved.

Furthermore, to save work when computing the new efficient points (i. e. when solving the new optimization problems), we propose to use a warm-start strategy. With such a strategy, points from the iteration history of scalar problems already solved are used as starting points for the optimization problems currently under consideration.

6. The EffSet-Algorithm.

6.1. The Basic Idea. When solving a multicriteria optimization problem, we have to solve many standard scalar problems, each of them a perturbation of each other one. Of course, we can solve one scalarized problem first, use one of the iterates computed to generate a warm-start point for the next scalarized problem considered, etc. This is basically the idea outlined in Fliege & Heseler [15]. However, in the algorithm presented below we want to do warm-starts as early as possible, thereby considering many scalarized problems simultaneously. In this way, we will be able to generate an approximation of the set of efficient points even if none of the scalar problems considered is solved up to a prespecified accuracy, yet. (I. e. even if all duality gaps of all scalar problems considered are still rather large.) A similar idea, albeit in a purely primal framework, has been pursued in [14].

Figure 6.1 (p. 18), taken from [7] illustrates the idea for bicriteria problems. The basic idea for bicriteria problems is explained below.

Suppose we start with a prespecified scalarization parameter ω, defining a problem with data instance denoted by $d(\omega)$. Choose a starting point $(x^0(\omega), \lambda^0(\omega), s^0(\omega))$ for algorithm QIP. In Figure 6.1, the image of this point under two real-valued objective functions is shown in the upper right hand corner. Executing one step with algorithm QIP results in the point $(x^1, \lambda^1, s^1)(\omega)$. Next, we choose two new scalarization parameters ω_l and ω_r and define the corresponding data instances $d(\omega_l)$ and $d(\omega_r)$. These two data instances can be seen as perturbations of $d(\omega)$ simply by noting $d(\omega_l) = d(\omega) + (d(\omega_l) - d(\omega))$ resp. $d(\omega_r) = d(\omega) + (d(\omega_r) - d_(\omega))$. Accordingly, we can try to construct warm-start points based on the information given by $(x^1(\omega), \lambda^1(\omega), s^1(\omega))$. One additional step from each of these warm-start points with algorithm QIP leads to $(x^2(\omega_l), \lambda^2(\omega_l), s^2(\omega_l))$ and $(x^2(\omega_l), \lambda^2(\omega_l), s^2(\omega_l))$, whose images under the two objectives are depicted in see Figure 6.1. Moreover, one additional step with algorithm

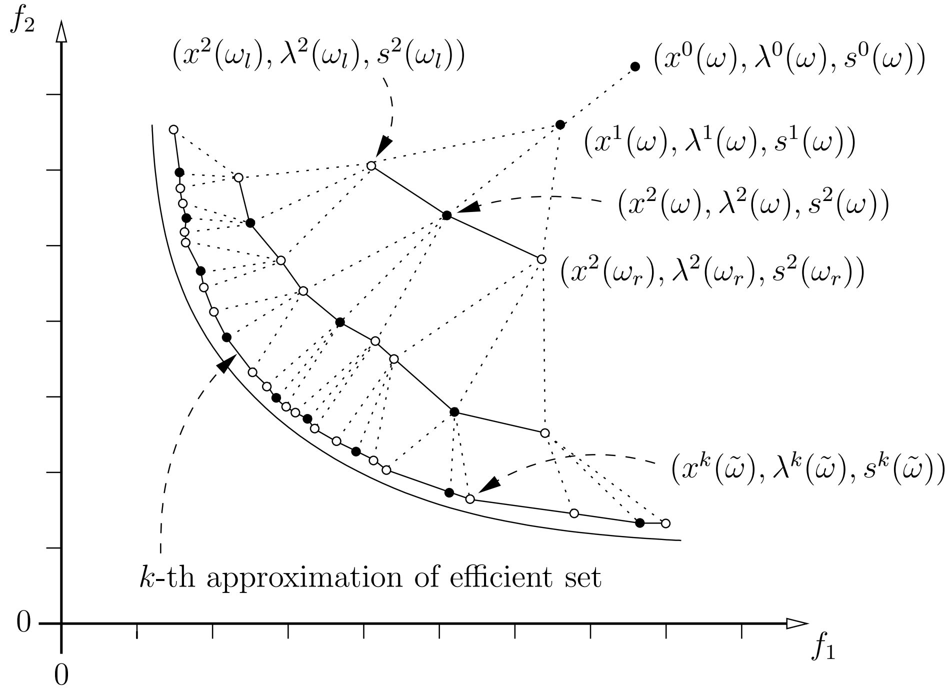

FIG. 6.1. *Approximation of the image of the set of efficient points by early warm-starts. An explanation is given in the text.*

QIP for the problem with parameter ω from $(x^1(\omega), \lambda^1(\omega), s^1(\omega))$ leads to the point $(x^2(\omega), \lambda^2(\omega), s^2(\omega))$. This scheme can now be applied recursively for each of the three points $(x^2(\omega_l), \lambda^2(\omega_l), s^2(\omega_l))$, $(x^2(\omega), \lambda^2(\omega), s^2(\omega))$, $(x^2(\omega_r), \lambda^2(\omega_r), s^2(\omega_r))$, until the images of neighboring iterates are closer to each other than a prespecified distance. A precise statement of the algorithm can be found in the next subsection.

In what follows, we will consider multicriteria optimization problems with three criteria of the form

$$\begin{aligned} \min \;\; f(x) &= \begin{bmatrix} (1/2)x^T Q_1 x + c_1^T x \\ (1/2)x^T Q_2 x + c_2^T x \\ (1/2)x^T Q_3 x + c_3^T x \end{bmatrix} \\ \text{subject to} \quad & x \in G \end{aligned}$$

with $G = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\}$. An efficient point to this problem will be approximated by solving a scalarized problem of the form

$$\min_{x \in G} \; \omega_1 f_1(x) + \omega_2 f_2(x) + \omega_3 f_3(x), \tag{6.1}$$

parameterized with $\omega_1, \omega_2, \omega_3 > 0$. In contrast to the rather simple twocriteria case that has been outlined above, we have to consider a triangulation of the parameter space given by the set $Z \subset \mathbb{R}^3$, see (5.5). This is the subject of the next subsection.

6.2. A Triangularization Strategy. For a given parameter $\omega \in Z$, let $x^k(\omega)$ be the primal iterate no. k of the algorithm described below. We will then use the notation

$$\begin{bmatrix} f_1(x^k(\omega)) \\ f_2(x^k(\omega)) \\ f_3(x^k(\omega)) \end{bmatrix} = f(x^k(\omega)) =: f^k(\omega)$$

as well as $w^k(\omega) := (x^k(\omega), \lambda^k(\omega), s^k(\omega))$ for the corresponding primal-dual iterate. With this, a triangle described by its vertices $(\omega^1, \omega^2, \omega^3) \in Z$ induces corresponding triangles in the image space of the multicriteria problem given by its three vertices $(f^k(\omega^1), f^k(\omega^2), f^k(\omega^3))$.

We now have two quality measures for the approximation to the set of efficient points to be constructed: the duality gaps of the primal-dual pairs found up to now (or, more precisely, the maximum of all duality gaps for the current iterate) and the sizes of the triangles in the image space (or, more precisely, the maximum size of these triangle sizes for the current iterate).

We will make use of the following notation. At iteration k, we denote the set of scalarized problems considered by $P(k)$, and we identify these problems with their weight vectors, i. e. $P(k) \subset Z$. Moreover, we will identify the triangulation of the set of weights Z considered at iteration no. k with a finite set $T(k) \subset Z^3$ of vertices of the corresponding triangles. This set induces a corresponding set of vertices of triangles $F(k) \subset (\mathbb{R}^3)^3$ in the image space of f, i. e. for $(\omega_1, \omega_2, \omega_3) \in T(k)$ we have, with the notation as above, $(f^k(\omega^1), f^k(\omega^2), f^k(\omega^3)) \in F(k)$.

Algorithm EffSet is depicted on p. 21. Some remarks on its workings are in order.

At each iteration, the algorithm proceeds as follows. If the area of a triangle in $F(k)$ is found to be larger than a prespecified accuracy δ, this triangle (and the corresponding one in the parameter space) has to be subdivided. For this, we chose the following strategy. Let ℓ be the length of the longest edge of such a triangle. Then we will split all edges of the triangle whose lengths are longer than $\ell/4$. The basic strategy for doing this is as follows.

Let such a triangle edge that has to be split be given. Denote the corresponding endpoints by $f^k(\omega_1)$ and $f^k(\omega_2)$. Define $\tilde{\omega} := (\omega_1 + \omega_2)/2$. Choose one of the problems with parameter ω_1 or ω_2 and do a warm-start for the problem with parameter $\tilde{\omega}$, making use of either $w^k(\omega_1)$ or $w^k(\omega_2)$. If such a warm-start is successful, i. e. if, say, $w^k(\omega_1) + \Delta w \in \tilde{\mathcal{N}}_{-\infty}(\gamma, \beta)$ for the Δw computed, denote by $w^k(\tilde{\omega})$ the warm-start point computed. In case a warm-start is not successful, i. e. if $w^k(\omega_1) + \Delta w \notin \tilde{\mathcal{N}}_{-\infty}(\gamma, \beta)$, a simple backtracking strategy is used to compute parameters of the form $q^j \tilde{\omega} + (1 - q^j)\omega$ ($q \in]0, 1[$ prespecified, $j \in \mathbb{N}$) from which further warm-start attempts were made. Only if this is unsuccessful, too, a point $w^k(\tilde{\omega})$ is computed by k iterations with algorithm QIP from the standard starting point, i. e. by a cold-start for the corresponding problem. This warm-start strategy is depicted in the subroutine 2, see p. 20.

Finally, the original triangle in the parameter space is subdivided into two, three, or four smaller triangles, depending on the number of edges that have been divided. All bordering triangles are also subdivided to keep the whole triangulation regular. The same happens in the image space.

In an actual implementation of this algorithm, a lower bound on the size of triangles in the parameter space should also be considered, i. e. if the area of a triangle represented by $(\omega^1, \omega^2, \omega^3) \in Z^3$ is below a certain strictly positive bound, such a triangle should not be subdivided any more, even if $\text{area}(\text{conv}\{f^k(\omega^1), f^k(\omega^2), f^k(\omega^3)\})$ is still considered as too large. In that way, the finiteness of the algorithm follows

Input: $\omega, \bar{\omega} \in Z$, $d_\omega = (A, b, Q, c)$, $w = (x, \lambda, s)$.

Choose $q \in]0, 1[$ and $\varrho > 0$.

$j := 0$

repeat

Define

$$\tilde{\omega} := q^j \bar{\omega} + (1 - q^j)\omega,$$
$$\tilde{Q} := \sum_{i=1}^{3} \tilde{\omega}_i Q_i,$$
$$\Delta Q := \tilde{Q} - Q,$$
$$\Delta c := \sum_{i=1}^{3} (\tilde{\omega}_i - \omega_i) c_i.$$

Solve

$$\begin{bmatrix} -\tilde{Q} & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} \Delta c + \Delta Q x \\ 0 \\ 0 \end{bmatrix}.$$

Set

$$w(\tilde{\omega}) := w(\omega) + (\Delta x, \Delta \lambda, \Delta s).$$

$j := j + 1$

until $w(\tilde{\omega}) \in \tilde{\mathcal{N}}_{-\infty}(\gamma, \beta)$ *or* $q^j \|\omega\| < \varrho$

if $w(\tilde{\omega}) \notin \tilde{\mathcal{N}}_{-\infty}(\gamma, \beta)$ **then**

Define the problem data $\tilde{d}$ for the weight $\tilde{\omega} := \bar{\omega}$ and execute k steps of algorithm QIP to generate $w(\tilde{\omega})$.

endif

Output: $w(\tilde{\omega})$, $\tilde{\omega}$, $\tilde{d} = (A, b, \tilde{Q}, c + \Delta c)$.

Algorithm 2: Warmstart

directly from the standard convergence properties of algorithm QIP and [13, Theorem 4.8], as long as at least one of the criteria functions is strictly convex.

7. Numerical Results. The implementation of algorithm EffSet was carried out in MATLAB, Version 7.0 (R14). We tested the algorithm on the problems described in Subsection 5.2. We have been provided 1201 problems from a power plant optimization scenario with $K = 14$ and $t_{\max} = 4$ [28]. Table 7.1 (p. 21) contains the most important parameter values chosen. Our update rule for the parameter σ follows the strategy of Wright [35].

Figures 7.1 and 7.2 show a typical computed approximation of the image of the set of efficient points and the corresponding discretization of the parameter space, respectively. Note that in Figure 7.1, each axis is scaled in such a way that 0 represents the smallest function value occurring in an efficient point and 1 represents the largest function value occurring in an efficient point. The discretization in the parameter space also provides an excellent visual clue on the sensitivity of the scalarized problems

Choose $\varepsilon, \delta > 0$, $\gamma \in]0,1[$ and $\beta \geq 1$. Choose $k, k_{\max} \in \mathbb{N}$.
With $p^1 := (1,0,0)^T$, $p^2 := (0,1,0)^T$, $p^3 := (0,0,1)^T$, $p^4 := (1/3,1/3,1/3)^T$, $P(k) := \{p^1, p^2, p^3, p^4\}$, set $T(k) := \{(p^1,p^2,p^4),(p^1,p^3,p^4),(p^2,p^3,p^4)\}$.
foreach *problem represented by* $\omega \in P(k)$ **do**
  Execute k steps of algorithm QIP from the standard starting point (cold start).
endfch
while $(\max\{\mu_k(\omega) \mid \omega \in P(k)\} \geq \varepsilon$ *and* $\max\{\text{area}(t) \mid t \in F(k)\} \geq \delta)$ *or* $k \leq k_{\max}$ **do**
  foreach $t \in F(k)$ **do**
    if $\text{area}(t) > \delta$ **then**
      forall the *edges* e *of* t **do**
        if $\text{length}(e) > \max_e \text{length}(e)/4$ **then**
          Denote the parameters corresponding to one of the endpoints of e by $\omega \in Z$ and let $\bar{\omega} \in Z$ be the midpoint of e.
          $(w^k(\tilde{\omega}),\, \tilde{\omega},\, \tilde{d}) :=$ `WarmStart` $(\omega, \bar{\omega}, d_\omega, w^k(\omega))$
          Define
$$P(k) := P(k) \cup \{\tilde{\omega}\}.$$
        endif
      endfall
      Subdivide t and update $T(k)$ accordingly.
    endif
  endfch
  foreach *problem represented by* $\omega \in P(k)$ **do**
    Execute one step of algorithm QIP from $w^k(\omega)$ to generate $w^{k+1}(\omega)$ and new dual gaps $\mu_{k+1}(\omega)$.
  endfch
  foreach $t \in F(k)$ **do**
    update information on t according to the subdivision of the parameter space.
  endfch
  $k := k + 1$
endw

Algorithm 3: EffSet

parameter	value	parameter	value
β	1.2	ζ	400
$\bar{\beta}$	12	ϱ	$2\sqrt{\texttt{eps}}$
γ	10^{-4}	$\sigma_{\max}$	0.1
$\bar{\gamma}$	10^{-5}	$\sigma_{\min}$	0.06
δ_P	$5 \cdot 10^{-16}$	k	5
ε	$\sqrt{\texttt{eps}}$	q	0.8

TABLE 7.1
Parameter values for algorithm EffSet and QIP

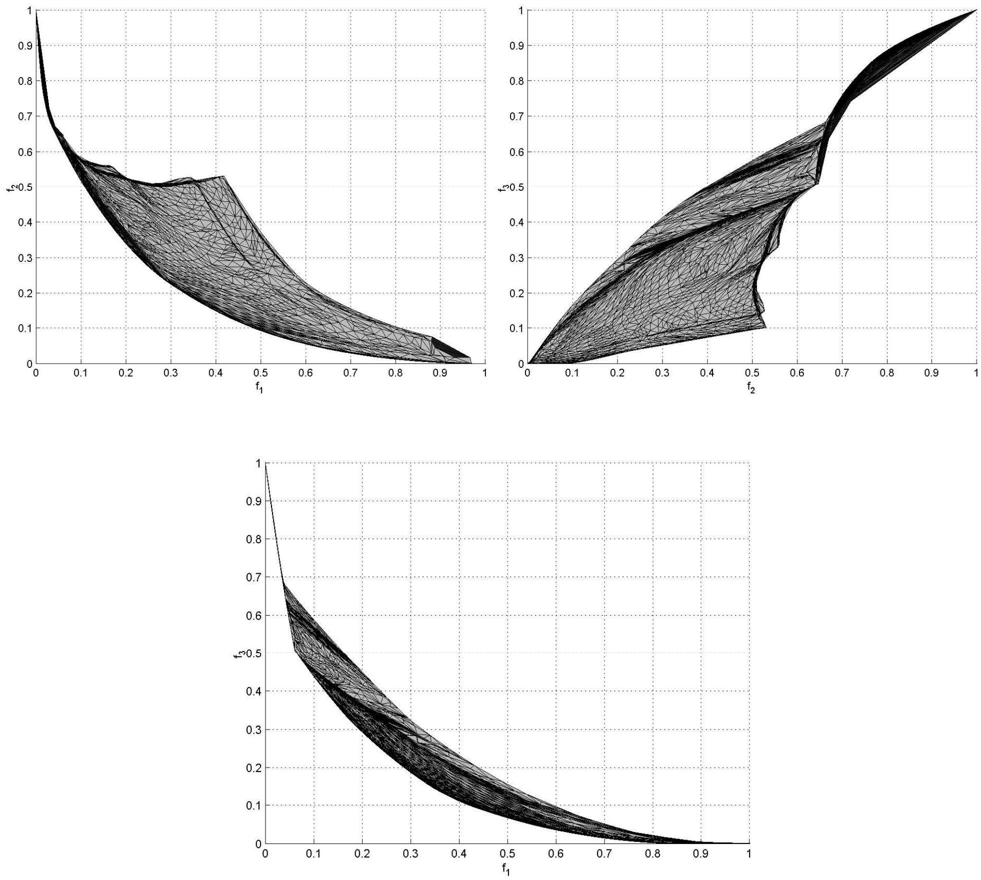

FIG. 7.1. *Computed approximation to the image of the set of efficient points, seen from three different view points, for a typical power plant optimization problem. Axes have been scaled.*

with respect to the scalarization parameter $\omega \in Z$.

For comparison purposes, we replaced line 12 of Algorithm EffSet with a corresponding *cold start strategy*, namely line 10 of Algorithm WarmStart, i. e. we assume for comparison purposes that the warm start strategy always fails.

The corresponding results can be seen in Table 7.2, p. 23. The most striking difference between the two strategies is the number of linear systems that had to be solved per efficient point (line 4 of Table 7.2). As it can be seen, this number increases by ca. 250% if only cold-starts are used instead of the warm-start strategy. While the number of iterations for both approaches is about the same, the total number of linear systems that have to be solved increases by ca. 24% when only cold-starts are used. Moreover, the warm-start strategy produces, on average, more than twice as many efficient points, resulting in an approximation of the efficient set of much higher quality than the cold start approach. It can also be seen that the warm-start strategy failed for only about 6% of all efficient points computed, and only then the corresponding cold-start from line 10 of Algorithm 2 had to be executed. These results clearly illustrate the efficiency of the method proposed.

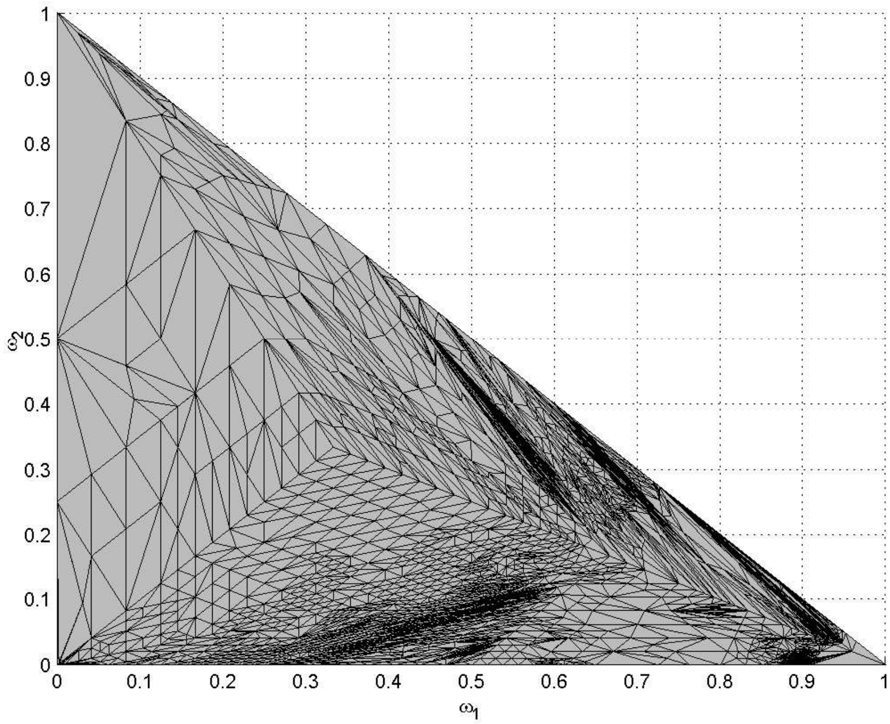

FIG. 7.2. *Discretization of the parameter space for the problem whose solution is depicted in Figure 7.1. Since $\omega_3 = 1 - \omega_1 - \omega_2$, only the (ω_1, ω_2)-plane is depicted.*

	average	max	min	average	max	min
# iterations	39	133	24	40	122	28
# efficient points	967	6259	157	435	2987	88
# linear systems considered	8574	66410	1711	10586	72617	2236
# linear systems / # points	9.71	22.05	5.79	24.81	27.11	22.61
# cold starts / # eff. pts	0.06	0.75	0.0	1.0	1.0	1.0
running time [s]	171.9	1255.1	32.6	189.3	1361.9	39

TABLE 7.2

Results for algorithm **EffSet** *as well as for a cold start approach for the power plant optimization problems considered. Reported are average, maximal, and minimal values over all problems.*

REFERENCES

[1] Warren Adams, Marie Coffin, Peter Kiessler, K. B. Kulasekera, Matthew J. Saltzmann, and Margaret Wiecek. An optimization-based framework for affordable science and technology decisions at the office of naval research. Clemson report, Clemson University, Clemson, South Carolina, October 16 1999.

[2] Jens Albrecht. *Optimal Forward Planning Control Strategies to Improve Power Plant Control (in german)*. PhD thesis, Faculty of Electrical Engineering, University of Dortmund, Germany, 1997. Published in Fortschrittsberichte VDI, Series 8, No. 616.

[3] U. Bagchi. Simultaneous minimization of mean and variation of flow time and waiting time in single machine systems. *Operations Research*, 37:118–125, 1989.

[4] E[milio] Carrizosa, E. Conde, M[anuel] Munoz, and J[usto] Puerto. Planar point-objective location problems with nonconvex constraints: a geometrical construction. *Journal of Global Optimization*, 6:77–86, January 1995.

[5] E[milio] Carrizosa and J. B. G. Frenk. Dominating sets for convex functions with some applications. Preprint TI 96-156/9, Tinbergen Institute, The Netherlands, 1996. Published as [6].

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

[7] Bernd Weyers Christoph Heermann and Jörg Fliege. A new adaptive algorithm for convex quadratic multicriteria optimization. In Jürgen Branke and Kaisa Miettinen, editors, *Practical Approaches to Multi-Objective Optimization*, Schloß Dagstuhl, Germany, 2005. Dagstuhl Seminar Proceedings. Electronically available from `http://drops.dagstuhl.de/opus/volltexte/2005/238/`.

[8] Indraneel Das and J. E. Dennis. Normal-Boundary Intersection: A New Method for Generating Pareto Optimal Points in Nonlinear Multicriteria Optimization Problems. *SIAM Journal on Optimization*, 8(3):631–657, 1998.

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

[10] G. Evans. Overview of techniques for solving multiobjective mathematical programs. *Management Science*, 30:1268–1282, 1984.

[11] Jörg Fliege. Project report *Research in Brussels: Atmospheric Dispersion Simulation*. Report BEIF/116, Vakgroep Beleidsinformatica, Department of Management Informatics, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussel, Belgium, October 1999. Also published by the Region Bruxelles-Capital, Belgium. Available electronically via `http://www.mathematik.uni-dortmund.de/user/lsx/fliege/olaf/`.

[12] Jörg Fliege. OLAF — a general modeling system to evaluate and optimize the location of an air polluting facility. *ORSpektrum*, 23(1):117–136, 2001.

[13] Jörg Fliege. Gap-Free Computation of Pareto-Points by Quadratic Scalarizations. *Mathematical Methods of Operations Research*, 59, 2004.

[14] Jörg Fliege. An Efficient Interior-Point Method for Convex Multicriteria Optimization Problems. Preprint 29/2005, The University of Birmingham, School of Mathematics, The University of Birmingham, Edgbaston, B15 2TT, Birmingham, U.K., 2005. To appear in *Mathematics of Operations Research*. Available at `http://www.optimization-online.org`.

[15] Jörg Fliege and Andree Heseler. Constructing approximations to the efficient set of convex quadratic multiobjective problems. Ergebnisberichte Angewandte Mathematik 211, Universität Dortmund, Fachbereich Mathematik, Universität Dortmund, 44221 Dortmund, Germany, January 2002. Submitted. Available at `http://www.optimization-online.org`.

[16] Jörg Fliege and Luis N. Vicente. A Multicriteria Approach to Bilevel Optimization. Ergebnisberichte Angewandte Mathematik 227, Universität Dortmund, Fachbereich Mathematik, Universität Dortmund, 44221 Dortmund, Germany, February 2003. To appear in *Journal of Optimization Theory and Applications*. Available at `http://www.optimization-online.org`.

[17] Anders Forsgren and Göran Sporre. On weighted linear least-squares problems related to interior methods for convex quadratic programming. *SIAM Journal On Matrix Analysis and Applications*, 23(1):42–56, 2001.

[18] Arthur M. Geoffrion. Proper Efficiency and the Theory of Vector Maximization. *Journal of Optimization Theory and Applications*, 22:618–630, 1968.

[19] A[lfred] Göpfert and R[einhard] Nehse. *Vektoroptimierung — Theorie, Verfahren und Anwendungen*, volume 74 of *Mathematisch-Naturwissenschaftliche Bibliothek*. BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1990.

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

[21] Andree Heseler. *Warmstart-Strategies in Parametric Optimization with Application in Multicriteria Optimization (in german)*. PhD thesis, Department of Mathematics, University of Dortmund, Germany, 2005.

[22] Claus Hillermeier. *Nonlinear Multiobjective Optimization: a Generalized Homotopy Approach*, volume 25 of *International Series of Numerical Mathematics*. Birkäuser Verlag, Basel — Boston — Berlin, 2001.

[23] Angelika Hutterer and Johannes Jahn. Optimization of the location of antennas for treatment planning in hyperthermia. Preprint 265, Institut für Angewandte Mathematik, Universität Erlangen-Nürnberg, Martensstraße 3, D-91058 Erlangen, June 15 2000.

[24] Johannes Jahn. *Mathematical Vector Optimization in Partially Ordered Linear Spaces.* Verlag Peter Lang, Frankfurt, 1986.

[25] Johannes Jahn, Andreas Kirsch, and Carmen Wagner. Optimization of rod antennas of mobile phones. *Mathematical Methods of Operations Research*, 59:37–51, 2004.

[26] A. Jüschke, J. Jahn, and A. Kirsch. A bicriterial optimization problem of antenna design. *Computational Optimization and Applications*, 7:261–276, 1997.

[27] Renata Kasperska. Bi-criteria Optimization Of Open Cross-Section Of The Thin-Walled Beams With Flat Flanges. Presentation at 75th Annual Scientific Conference of the Gesellschaft für Angewandte Mathematik und Mechanik e. V., Dresden, Germany, March 2004.

[28] H. Kiendl and D. Schauten. Private communication, April 2001.

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

[30] D. Prabuddha, J. B. Gosh, and C. E. Wells. On the minimization of completion time variance with a bicriteria extension. *Operations Research*, 40:1148–1155, 1992.

[31] Yoshikazu Sawaragi, Hirotaka Nakayama, and Tetsuzo Tanino. *Theory of Multiobjective Optimization*, volume 176 of *Mathematics in Science and Engineering.* Academic Press, Inc., Orlando, San Diego, New York, London, 1985.

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

[33] Manuel A. Nunez und Robert M. Freund. Condition measures and properties of the central trajectory of a linear program. *Mathematical Programming*, 83:1–28, 1998.

[34] D[ouglas] J[ohn] White. Epsilon-dominating solutions in mean-variance portfolio analysis. *European Journal of Operational Research*, 105:457–466, 1998.

[35] Stephen J. Wright. *Primal-dual interior-point methods.* SIAM, Philadelphia, 1997.

[36] E. Alper Yıldırım and S. J. Wright. Warm-start strategies in interior-point methods for linear programming. Technical Report 1258, School of Operations Research and Industrial Engineering, Cornell University, 2000.

[37] E. Alper Yıldırım and S. J. Wright. Warm-start strategies in interior-point methods for linear programming. Technical Report 1258, School of Operations Research and Industrial Engineering, Cornell University, 2000.