BERGISCHE
UNIVERSITÄT
WUPPERTAL

Solution Approaches for Multiobjective Convex Quadratic and Nonlinear Optimization Problems

Dissertation zur Erlangung des Doktorgrades (Dr. rer .nat.)
Fakultät für Mathematik und Naturwissenschaften
Bergische Universität Wuppertal

vorgelegt von
Marco Milano
aus Ratingen

Erstgutachterin:
Prof. Dr. Kathrin Klamroth

Zweitgutachterin:
Prof. Dr. Margaret Wiecek

Wuppertal, Juli 2019

The PhD thesis can be quoted as follows:

urn:nbn:de:hbz:468-20191029-103617-9
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20191029-103617-9]

DOI: 10.25926/k6wa-q019
[https://doi.org/10.25926/k6wa-q019]

Contents

Chapter 1

Introduction

Many decisions in life involve dealing with multiple conflicting objectives. For example, when selecting a portfolio out of a set of given market options, we would like to maximize our expected return while minimizing the risks involved [59]. In general, there does not exist a solution that satisfies all objectives simultaneously. In order to make an informed decision we need to have information about the alternatives that are available to us.
The field of multiobjective optimization aims to, among other things, provide theoretical frameworks and methods to provide information about the possible alternatives. Such alternatives are often characterized by the concept of Pareto-optimality which was introduced by Francis Edgeworth (1945-1926) and Vilfredo Pareto (1848-1923) over one hundred years ago: At a Pareto-optimal solution no objective can be improved without sacrificing another. For a decision maker it can be helpful to know the whole set of Pareto-optimal solutions or an approximation of such a set.
The portfolio selection problem introduced by Markowitz [59] is a biobjective optimization problem with a quadratic and a linear objective function. If the decision maker can articulate a preference wrt. the ratio of expected return and risk then a solution can be determined that optimizes the weighted sum of the objectives. The resulting problem is a singleobjective problem and can then be solved with methods from the field of quadratic optimization.

1.1 Outline

In this thesis we consider several classes of multiobjective optimization problems, such as multiobjective continuous optimization problems, multiobjective convex quadratic and convex piecewise-linear optimization problems.

Chapter 2 introduces notation, concepts and results that are important for this thesis. In Section 2.1 optimality criteria for singleobjective optimization problems are summarized. Linear complementarity problems are also introduced as a solution method for convex quadratic optimization problems.
Multiobjective optimization problems are discussed in Section 2.2 and a summary of scalarization techniques is given.

Chapter 3 reviews a class of multiobjective descent methods based on the steepest descent method by Fliege et. al. [30]. Methods of this type can be interpreted as extensions of well-kown singleobjective descent algorithms, such as the steepest descent method and Newtons method.
In the first part of Chapter 3 we consider the generalization of singleobjective descent methods to the multiobjective case by reviewing different choices for descent directions and step sizes.
Afterwards we consider different choices for descent directions and review the convergence properties of the resulting multiobjective descent algorithm. Using illustrative examples we conclude with a comparison of multiobjective descent algorithms and the application of the weighted sum scalarization.

Chapter 4 considers multiobjective convex quadratic optimization problems. First, two similar solution approaches from the literature are discussed that provide an analytical representation of the efficient set: A weight space decomposition by active sets as suggested by Goh and Yang [38] and a parametric approach by Adelgren [1]. We show that these approaches are, in fact, very similar and can be used in conjunction.
Afterwards, we consider special cases of multiobjective convex quadratic optimization problems that are easier to solve since only polyhedral set have to be considered. Furthermore, we discuss the relationship of the weight space decomposition by active sets and the parameter space of the e-constraint problem. Finally, an application to multiobjective location theory is discussed.

Chapter 5 reviews an outer approximation method proposed by Oberdieck and Pistikopoulos [67] who construct a multiobjective convex piecewise-linear optimization problem to approximate the weight space decomposition of a multiobjective convex optimization problem. In order to discuss this approach we consider the weight space decomposition by active sets for multiobjective convex quadratic optimization problems. In particular, we discuss how the weight space decomposition for multiobjective convex piecewise-linear optimization problems can be computed using multiobjective linear programming techniques.
Furthermore, we consider results from the field of approximation of convex bodies by convex polyhedra in order to provide a convergence result for the approach by Oberdieck and Pistikopoulos [67]. We conclude with an illustrative example.

Chapter 6 concludes the results from Chapters 3, 4 and Chapter 5.

1.2 Acknowledgments

From the beginning of my studies many people provided me with support in countless ways. This section is dedicated to those people.

Foremost, I would like to thank my supervisor Kathrin Klamroth, who sparked my interest in the fields of optimization, provided me with many insights into the topic, arranged my semester abroad in Nagoya, Japan, and offered me a position at the University of Wuppertal.

I also want to thank Margaret Wiecek for being the second reviewer of this thesis and for her invaluable insights into the topic of multiobjective optimization and for our interesting discussions during her stay in Wuppertal.

The time I have spent in the working group was wonderful, especially because of the friendly atmosphere created by the former and current members:
Peter Beisel, Margareta Heilmann, Michael Stiglmayr, Teresa Schnepper, Britta Schulze, Renaud Lacour, Kerstin Dächert, Markus Kaiser, Onur Doganay, Konstantin Kraus, Johanna Schultes, Julius Bauß, Tobias Suszka and Jonas Kremer.

I am also grateful for the support I received from friends and family, particularly from my parents Karin Milano and Giuseppe Milano, from my sister Romina Milano and from my close friends, Tim Diaubalick, Marie-Laure Schnieringer, Mirjam Pützer and Stephanie Warnicke.

Notation

All Chapters	
Ordering of Vectors in $\mathbb{R}^m$, $m \geq 2$	
$a \leqq b$	$a_i \leq b_i \forall i = 1, \dots, m$
$a \leq b$	$a \leqq b$ and $a \neq b$
$a < b$	$a_i < b_i \forall i = 1, \dots, m$
cl	Closure of a set
bd	Boundary of a set
X_E	Set of efficient solutions
Y_{nd}	Set of nondominated vectors
y^N, y^I	Nadir and ideal point
Λ	Weight space, $\Lambda = \left\{\lambda \in \mathbb{R}^m \ : \ \lambda \geq 0, \ \sum_{i=1}^m \lambda_i = 1\right\}$
W	W-Parameterization: $W = \{w \in \mathbb{R}^m, \ w \geq 0, \ w_1 = 1\}$
S	Feasible set
$X_{\text{opt}}(\lambda)$	Set of optimal solutions of the weighted sum problem for $\lambda \in \Lambda$
$\bar{x}(\lambda)$	Unique optimal solution of the weighted sum problem
Chapter 3	
$D(x)$	Descent cone in $x \in \mathbb{R}^n$
Chapter 4	
(MCP)	Multiobjective convex optimization problems in canonical form
(MLP)	Multiobjective linear optimization problem
(MQP)	Multiobjective convex quadratic optimization problems in canonical form
(gMQP)	Multiobjective convex quadratic optimization problems in general form
(MMLQP)	Multiobjective convex mixed linear-quadratic optimization problems in general form
(pLCP)	Parametric linear complementarity problem corresponding to (MQP)
(pmLCP)	Parametric mixed-linear complementarity problem corresponding to (gMQP)
$\mathscr{A} = (\mathcal{I}, \mathcal{J})$	Active sets $\mathscr{A}$ with index sets: $\mathcal{I}$ for lower bounds, $\mathcal{J}$ for linear inequality constraints.
$\Lambda^A(\mathscr{A})$	Weight cell for efficient active set of (MQP) or (gMQP)
$\Lambda^B(B)$	Weight cell for complementary basis of (pLCP) or (pmLCP)
$\mathfrak{h}_B^k$	Hypersurfaces that define weight cells $\Lambda^B(B)$
Chapter 5	
$\Lambda^C(\mathscr{A})$	Weight Cell for active set $\mathscr{A}$
η	Approximation error
$d_H(C_1, C_2)$	Hausdorff-distance between sets C_1 and C_2
$d(x, y)$	Euclidean distance between two vectors $x, y \in \mathbb{R}^n$
Epi$f$	Epigraph of function f
$\mathcal{C}$	Bounded epigraph
$\mathcal{T}(x, Z)$	Approximation function of $f(x)$
$P(Z)$	Bounded epigraph of $\mathcal{T}(x, Z)$

Chapter 2

Preliminaries

In this thesis the following notation for inequalities of vectors is used for $a, b \in \mathbb{R}^m$ with $m \geq 2$:

- $a \leqq b \Leftrightarrow a_i \leq b_i \forall i = 1, \dots, m$
- $a \leq b \Leftrightarrow a \leqq b$ and $a \neq b$
- $a < b \Leftrightarrow a_i < b_i \forall i = 1, \dots, m$

Note that when comparing two scalars $a, b \in \mathbb{R}$ that $a \leq b$ is used in canonical way.

2.1 Singleobjective Optimization

We refer to the books by Geiger and Kanzow [36, 37] and Nocedal and Wright [66] for a detailed introduction into the field of nonlinear optimization.

Consider an optimization problem of the following form:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & g(x) \leqq 0 \\ & h(x) = 0 \end{aligned} \tag{2.1}$$

with continuously differentiable functions $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^p$ and $h : \mathbb{R}^n \to \mathbb{R}^q$. The feasible set is given by

$$S := \{x \in \mathbb{R}^n \; : \; g(x) \leqq 0, \; h(x) = 0\}.$$

Fritz-John and KKT points are used as optimality conditions for many differentiable continuous optimization problems.

Definition 2.1. [37] A point $x \in \mathbb{R}^n$ satisfies the **linear constraint independence constraint qualification** (LICQ) if the set

$$\left\{\nabla g_j(x) \; : \; j \in \{i \in \{1, \dots, p\} \; : \; g_i(x) = 0\}\right\} \cup \{\nabla h_l(x) \; : \; l \in \{1, \dots, q\}\} \tag{2.2}$$

is linearly independent.
The point x is then called **regular point** of (2.1).

Definition 2.2. [37] A point $x \in \mathbb{R}^n$ is called **Fritz-John point** of (2.1) if there exist $u_0 \in \mathbb{R}$, $\pi \in \mathbb{R}^p$ and $\mu \in \mathbb{R}^q$ such that

$$\begin{aligned} & u_0 \nabla f(\bar{x}) + \sum_{j=1}^{p} \pi_j \nabla g_j(\bar{x}) + \sum_{l=1}^{q} \mu_l \nabla h_l(\bar{x}) = 0 \\ & g(\bar{x}) \leqq 0, h(\bar{x}) = 0 \\ & u_0 \geq, \ \pi \geqq 0 \\ & \pi_j(g_j(\bar{x})) = 0 \ j = 1, \ldots, p. \end{aligned}$$

Definition 2.3. [37] A point $\bar{x} \in \mathbb{R}^n$ is called **KKT point** of (2.1) if there exist $\pi \in \mathbb{R}^p$ and $\mu \in \mathbb{R}^q$ such that the KKT conditions are satisfied:

$$\begin{aligned} & \nabla f(\bar{x}) + \sum_{j=1}^{p} \pi_j \nabla g_j(\bar{x}) + \sum_{l=1}^{q} \mu_l \nabla h_l(\bar{x}) = 0 \\ & g(\bar{x}) \leqq 0, h(\bar{x}) = 0 \\ & \pi \geqq 0 \\ & \pi_j(g_j(\bar{x})) = 0 \ j = 1, \ldots, p. \end{aligned}$$

The following results provide necessary optimality conditions for (2.1):

Theorem 2.4. [37] Let $\bar{x}$ be a local minimum of (2.1). Then the following statements hold:

1. $\bar{x}$ is a Fritz-John point of (2.1).
2. If $\bar{x}$ is a regular point of (2.1), then $\bar{x}$ is a KKT point of (2.1).

For problems with a convex objective function $f(x)$ and a convex feasible sets the KKT conditions are also sufficient optimality conditions:

Theorem 2.5. [37] Let $f(x)$ and $g(x)$ be convex functions and let $h(x) = Hx - h$ for some $H \in \mathbb{R}^{q \times n}$ and $h \in \mathbb{R}^q$. Let $\bar{x}$ be a regular point of (2.1). If $\bar{x}$ is a KKT point, then $\bar{x}$ is an optimal solution of (2.1).

2.1.1 Quadratic Optimization Problems

Quadratic optimization problems arise from applications [59] as well as subproblems of nonlinear optimization methods [37].

Consider a convex quadratic programming problem in canonical form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) = \tfrac{1}{2} x^T Q x + c^T x \\ \text{s.t.} \quad & Ax \geqq b, \ x \geqq 0 \end{aligned} \tag{2.3}$$

with $Q \in \mathbb{R}^{n \times n}$ positive semidefinite, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^p$.

Corollary 2.6. [37] A regular point $\bar{x} \in \mathbb{R}^n$ of (2.3) is an optimal solution of (2.3) if and only if there exist $\pi \in \mathbb{R}^p$ and $y \in \mathbb{R}^n$ such that

$$\begin{array}{ll} Q\bar{x} + c - A^T\pi - y = 0 & \\ Ax \geqq b,\ x \geqq 0 & \\ \pi \geqq 0,\ y \geqq 0 & \\ \pi_j(A_{j\bullet}x - b_j) = 0 & \forall j = 1\dots,p \\ y_i x_i = 0 & \forall i = 1,\dots,n \end{array} \tag{2.4}$$

Consider a convex quadratic programming problem in general form

$$\begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & f(x) = \frac{1}{2}x^TQx + c^Tx \\ \text{s.t.} & Ax \geqq b,\ x_{\mathcal{I}_+} \geqq 0, Hx = h \end{array} \tag{2.5}$$

with $Q \in \mathbb{R}^{n\times n}$ positive semidefinite, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{p\times n}$, $b \in \mathbb{R}^p$, $\mathcal{I}_+ \subseteq \{1,\dots,n\}$, $H \in \mathbb{R}^{q\times n}$, $h \in \mathbb{R}^q$.

Corollary 2.7. [37] A regular point $\bar{x} \in \mathbb{R}^n$ of (2.5) is optimal solution of (2.5) if and only if there exist $\pi \in \mathbb{R}^p$, $y \in \mathbb{R}^{|\mathcal{I}_+|}$ and $\mu \in \mathbb{R}^q$ such that

$$\begin{array}{ll} Q\bar{x} + c - A^T\pi - (I_{\mathcal{I}_+\bullet})^T y_{\mathcal{I}_+} + H^T\mu = 0 & \\ Ax \geqq b,\ x_{\mathcal{I}_+} \geqq 0, Hx = h & \\ \pi \geqq 0,\ y_{\mathcal{I}_+} \geqq 0 & \\ \pi_j(A_{j\bullet}x - b_j) = 0 & \forall j = 1,\dots,p \\ y_i x_i = 0 & \forall i \in \mathcal{I}_+ \end{array}$$

Quadratic programming problems can be solved with a variety of method such as active set methods [37, 82] and interior point methods [66]. An alternative approach is given by linear complementarity problems.

2.1.2 Linear Complementarity Problems

For a more thorough investigation of linear complementarity problems we refer to the book by Cottle [14]. The KKT conditions (2.4) for quadratic optimization problems in canonical form (2.3) can be written as a so called linear complementarity problem [14] by introducing an additional variable $s \in \mathbb{R}^p$ and an additional equation $s = Ax - b$:

$$\begin{bmatrix} I_n & 0 & -Q & A^T \\ 0 & I_p & -A & 0 \end{bmatrix} \begin{pmatrix} y \\ s \\ x \\ \pi \end{pmatrix} = \begin{pmatrix} c \\ -b \end{pmatrix}$$

$$s \geqq 0,\ y \geqq 0,\ \pi \geqq 0,\ x \geqq 0$$
$$s_j\pi_j = 0\ \forall j = 1,\dots,p$$
$$y_i x_i = 0\ \forall i = 1,\dots,n$$

Consider a linear complementarity problem of the form for some matrix $K \in \mathbb{R}^{r\times r}$.

$$\begin{array}{l} \begin{bmatrix} I_r & -K \end{bmatrix}(u,v) = q \\ u \geqq 0, v \geqq 0 \\ u_i v_i = 0\ \forall i = 1,\dots,r \end{array} \tag{LCP}$$

with $\mathrm{M} = \begin{bmatrix} I_r & -K \end{bmatrix}$ and $z = (u,v)$. We denote the complementary variable of z_i by $\hat{z}_i$ for $i = 1,\dots,r$. For example if $z_j = u_i$ then $\hat{z}_j = v_i$.

Definition 2.8. [14] Let $B \subset \{u_1, \ldots, u_r\} \cup \{v_1, \ldots, v_r\}$ be a set of variables with $|B| = r$. The matrix M_B consists of the columns of M corresponding to the variables in B.

1. B is called a **basis** of (LCP), if the matrix M_B is regular.
2. B is called **complementary**, if for every $i = 1, \ldots, r$ either $u_i \in B$ or $v_i \in B$.
3. The complement of B is denoted by N.

The variables in B (N) are called **basic** (**nonbasic**) variables, respectively. Let B be a complementary basis. Setting $u_i = 0$ for every $u_i \notin B$ and $v_i = 0$ for every $v_i \notin B$ reduces the linear complementarity problem (LCP) to

$$\begin{aligned} M_B(u_B, v_B) &= q \\ u_B \geqq 0,\ v_B &\geqq 0 \end{aligned} \tag{2.6}$$

where $u_B := \{u_i \in B \ : \ i = 1, \ldots, r\}$ and $v_B := \{v_i \in B \ : \ i = 1, \ldots, r\}$. The nonnegativity of the nonbasic variables and the complementarity condition is satisfied now. Since M_B is regular we can write the solution of (2.6) as

$$q_B := M_B^{-1} q \tag{2.7}$$

Definition 2.9. [14] Let B be a complementary basis of (LCP). B is called **feasible complementary basis** of (LCP) if $q_B := M_B^{-1} q \geqq 0$. The vector q_B is referred to as the value of the basic variables or basic value.

In order to solve (LCP) we want to find a feasible complementary basis of (LCP). This is the fundamental idea for pivoting algorithms for solving linear complementarity problems. In order to guarantee that (LCP) has a feasible complementary basis we study a special class of matrices:

Definition 2.10. [15] A matrix $K \in \mathbb{R}^{r \times r}$ is called

1. **column sufficient**, if the following implication is satisfied:

$$(z_i(Kz)_i \leq 0 \ \forall i = 1, \ldots, r) \Rightarrow z_i(Kz)_i = 0 \ \forall i = 1, \ldots, r.$$

2. **row sufficient**, if K^T is column sufficient.
3. **sufficient**, if K is column and row sufficient.

Proposition 2.11. [13] If K is positive semidefinite, then K is sufficient.

Proposition 2.12. [66] Let $Q \in \mathbb{R}^{n \times n}$ be positive definite and let $A \in \mathbb{R}^{p \times n}$ be of full rank. Then the following matrices are regular matrices

$$\tilde{K} = \begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix} \text{ and } K = \begin{bmatrix} Q & -A^T \\ A & 0 \end{bmatrix},$$

and K is positive definite.

Proposition 2.13. [15] The following statements are equivalent:

(a) K is sufficient.

(b) The problem (LCP) has a (possibly empty) convex solution set for every right-hand-side q.

To find the solution, we start with an arbitrary basis and pivot variables in and out of the basis until a basis with basic value $q_B \geqq 0$ is found. There are two kinds of pivots [14]:

- diagonal pivot: pivot u_i and v_i (one complementary pair)
- exchange pivot: two successive diagonal pivot with two different complementary pairs. This is the pivot used in the simplex method for solving linear programming problems [37].

Definition 2.14. [14] Let $B = (z_1, \ldots, z_k, \ldots, z_r)$ be a complementary variable set.

- The diagonal pivot is defined as

$$\text{diag}(B, k) := (z_1, \ldots, \hat{z}_k, \ldots, z_r)$$

- The exchange pivot is defined as

$$\text{exch}(B, k, l) := (z_1, \ldots, \hat{z}_k, \ldots, \hat{z}_l, \ldots, z_r)$$

where $\hat{z}_k$ is the complementary variable wrt. variable z_k.

Theorem 2.15. [13] Let (LCP) be a feasible linear complementarity problem with a sufficient matrix K and let B be a complementary basis of (LCP). Then there exists a sequence of complementary bases connected by diagonal and exchange pivots that ends in a feasible complementary basis.

Definition 2.16. [13]

1. Let B be a set of complementary variables. The **complementary cone** is defined as

$$C(B) = \text{cone}\,(M_B) = \left\{ \sum_{k \in B} \alpha_k M_{\bullet k} \ : \ \alpha \geqq 0 \right\}.$$

2. Two complementary cones $C(B_1)$ and $C(B_2)$ are called **adjacent** if $\dim(C(B_1) \cap C(B_2)) = r - 1$.

Proposition 2.17. [13, 1]

1. If B is a complementary basis. Then

$$C(B) = \left\{ a \in \mathbb{R}^r \ : \ M_B^{-1} a \geqq 0 \right\}.$$

2. $C(B)$ is full dimensional if and only if M_B has full rank.

Proposition 2.18. [13] If K is sufficient, then the relative interior of any two distinct complementary cones are disjoint.

Theorem 2.19. [13] Let B and B' be two complementary bases of (LCP). If $C(B)$ and $C(B')$ are adjacent, then $|B \cap B'| \geq r - 2$.

The dictionary of a complementary basis B of (LCP) is defined as the $r \times r$ matrix

$$T(B) := -M_B^{-1} M_N.$$

The working tableau wrt. a complementary basis B of (LCP) is defined as the $r \times (2r+1)$ matrix

$$\begin{bmatrix} T(B) & I_r & q_B \end{bmatrix} \tag{2.8}$$

Since we are only interested in complementary bases we use the following indexing of the variables and the dictionary: The i-th entry of B corresponds to the variable of the i-th complementary pair, i.e. $B(i) = u_i$ or v_i. Similar to the simplex tableau [37] the tableau is often shown with the basic value as an additional column. The resulting structure of the dictionary is visualized in Figure 2.1.

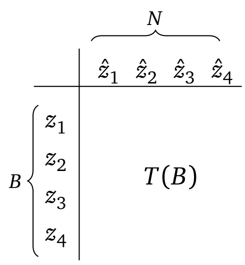

Figure 2.1: Indexing of the dictionary.

We will shortly discuss how a diagonal pivot changes the working tableau.
The new dictionary and working tableau after a diagonal pivot can be computed using elementary row transformations of the working tableau [14]. Consider performing a diagonal pivot $B' = \text{diag}(B, k)$ with $T_{kk} \neq 0$. A column for the basic variable z_k is included to visualize the basis exchange.

$$\begin{array}{c|ccccc:c:c}
 & \hat{z}_1 & \dots & \hat{z}_k & \dots & \hat{z}_r & z_k & \\
\hline
z_1 & T_{11} & \dots & T_{1k} & \dots & T_{1r} & 0 & (q_B)_1 \\
\vdots & \vdots & & \vdots & & \vdots & \vdots & \vdots \\
z_k & T_{k1} & \dots & (T_{kk}) & \dots & T_{kr} & 1 & (q_B)_k \\
\vdots & \vdots & & \vdots & & \vdots & \vdots & \vdots \\
z_r & T_{r1} & \dots & T_{rk} & \dots & T_{rr} & 0 & (q_B)_r
\end{array}$$

After one elementary row transformation we have the following working tableau [14]:

$$\begin{array}{c|ccccc:c:c}
 & \hat{z}_1 & \dots & \hat{z}_k & \dots & \hat{z}_r & z_k & \\
\hline
z_1 & T_{11} - T_{1k}\frac{T_{k1}}{T_{kk}} & \dots & 0 & \dots & T_{1r} - T_{1k}\frac{T_{kr}}{T_{kk}} & -\frac{T_{1k}}{T_{kk}} & (q_B)_1 - \frac{T_{1k}}{T_{kk}}(q_B)_k \\
\vdots & \vdots & & \vdots & & \vdots & \vdots & \vdots \\
\hat{z}_k & \frac{T_{k1}}{T_{kk}} & \dots & 1 & \dots & \frac{T_{kr}}{T_{kk}} & \frac{1}{T_{kk}} & \frac{(q_B)_k}{T_{kk}} \\
\vdots & \vdots & & \vdots & & \vdots & \vdots & \vdots \\
z_r & T_{r1} - T_{rk}\frac{T_{k1}}{T_{kk}} & \dots & 0 & \dots & T_{rr} - T_{rk}\frac{T_{kr}}{T_{kk}} & -\frac{T_{rk}}{T_{kk}} & (q_B)_r - \frac{T_{rk}}{T_{kk}}(q_B)_k
\end{array}$$

Consider performing an exchange pivot with z_k and z_l, where $T_{kk} = 0$. First, z_k is exchanged with $\hat{z}_l$:

$$
\begin{array}{c|ccccccc:cc:c}
 & \hat{z}_1 & \dots & \hat{z}_k & \dots & \hat{z}_l & \dots & \hat{z}_r & z_k & z_l & \\
\hline
z_1 & T_{11} & \dots & T_{1k} & \dots & T_{1l} & \dots & T_{1r} & 0 & 0 & (q_B)_1 \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots & \vdots & \vdots \\
z_k & T_{k1} & \dots & 0 & \dots & (T_{kl}) & \dots & T_{kr} & 1 & 0 & (q_B)_k \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots & \vdots & \vdots \\
z_l & T_{l1} & \dots & T_{lk} & \dots & T_{ll} & \dots & T_{lr} & 0 & 1 & (q_B)_l \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots & \vdots & \vdots \\
z_r & T_{r1} & \dots & T_{rk} & \dots & T_{rl} & \dots & T_{rr} & 0 & 0 & (q_B)_r
\end{array}
$$

$$
\begin{array}{c|ccccccc:cc:c}
 & \hat{z}_1 & \dots & \hat{z}_k & \dots & \hat{z}_l & \dots & \hat{z}_r & z_k & z_l & \\
\hline
z_1 & T_{11} - T_{1l}\frac{T_{k1}}{T_{kl}} & \dots & T_{1k} & \dots & 0 & \dots & T_{1r} - T_{1l}\frac{T_{kr}}{T_{kl}} & -\frac{T_{1l}}{T_{kl}} & 0 & (q_B)_1 - \frac{T_{1l}}{T_{kl}}(q_B)_k \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots & \vdots & \vdots \\
\hat{z}_l & \frac{T_{k1}}{T_{kl}} & \dots & 0 & \dots & 1 & \dots & \frac{T_{kr}}{T_{kl}} & \frac{1}{T_{kl}} & 0 & \frac{(q_B)_k}{T_{kl}} \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots & \vdots & \vdots \\
z_l & T_{l1} - T_{ll}\frac{T_{k1}}{T_{kl}} & \dots & (T_{lk}) & \dots & 0 & \dots & T_{lr} - T_{ll}\frac{T_{kr}}{T_{kl}} & -\frac{T_{ll}}{T_{kl}} & 1 & (q_B)_l - \frac{T_{ll}}{T_{kl}}(q_B)_k \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots & \vdots & \vdots & \vdots \\
z_r & T_{r1} - T_{rl}\frac{T_{k1}}{T_{kl}} & \dots & T_{rk} & \dots & 0 & \dots & T_{rr} - T_{rl}\frac{T_{kr}}{T_{kl}} & -\frac{T_{rl}}{T_{kl}} & 0 & (q_B)_r - \frac{T_{rl}}{T_{kl}}(q_B)_k
\end{array}
$$

After exchanging z_l with $\hat{z}_k$ the basic value of B after the exchange pivot is given by [14]:

$$
\begin{pmatrix} z_1 \\ \vdots \\ \hat{z}_k \\ \vdots \\ \hat{z}_l \\ \vdots \\ z_r \end{pmatrix}
=
\begin{pmatrix}
(q_B)_1 - \frac{T_{1l}}{T_{kl}}(q_B)_k - \frac{T_{1k}}{T_{lk}}\left((q_B)_l - \frac{T_{ll}}{T_{kl}}(q_B)_k\right) \\
\vdots \\
\frac{(q_B)_k}{T_{kl}} \\
\vdots \\
\frac{1}{T_{lk}}\left((q_B)_l - \frac{T_{ll}}{T_{kl}}(q_B)_k\right) \\
\vdots \\
(q_B)_r - \frac{T_{rl}}{T_{kl}}(q_B)_k - \frac{T_{rk}}{T_{lk}}\left((q_B)_l - \frac{T_{ll}}{T_{kl}}(q_B)_k\right)
\end{pmatrix}
$$

Proposition 2.20. [13] Let B be a complementary basis of (LCP) and let $k \in \{1, \dots, r\}$ be given. $B' = \operatorname{diag}(B, k)$ is a complementary basis of (LCP) if and only if $T_{kk}(B) \neq 0$.

Theorem 2.21. [13] Let K be a sufficient matrix and let B be a complementary basis of (LCP). Let $k, l \in \{1, \dots, r\}$ be with $k \neq l$ and let $B' = \operatorname{exch}(B, k, l)$. Then

$$\dim(C(B \setminus \{z_k\}) \cap C(B')) = r - 1 \Leftrightarrow T_{k,k}(B) = 0 \text{ and } T_{l,k}(B) < 0$$

Proposition 2.20 and Theorem 2.21 show that the dictionary $T(B)$ can be used to check whether a pivot will lead to a basis of the linear complementarity problem. The following section discusses one pivoting method to solve linear complementarity problems with sufficient matrices.

2.1.3 The Criss-Cross Method

Two popular pivoting algorithms for solving linear complementarity problems with sufficient matrices are Lehmke's method [57] and the criss-cross method [3].

Algorithm 2.1: Criss-Cross Method [3]

```
Choose a complementary Basis B and compute T(B) and q_B
while q_B ≱ 0 do
    Choose any k ∈ {1,...,r} with (q_B)_k < 0
    if T_kk(B) < 0 then
        Perform a diagonal privot with z_k and update T(B) and q_B
    else
        if {l ∈ {1,...,r} \ {k} : T_kl(B) < 0} = ∅ then
            STOP, the linear complementarity problem is infeasible.
        else
            Choose any l ∈ {1,...,r} \ {k} with T_kl(B) < 0.
            Perform a exchange pivot with z_k and z_l. Update T(B) and q_B
```

Using a pivoting strategy to avoid cycling (for example a variant from Akkeles et. al. [3]) Algorithm 2.1 will either end with a feasible basis of the linear complementarity problem or detect infeasibility after a finite number of iterations.

Theorem 2.22. [3] Let K be positive semidefinite. Then Algorithm 2.1 (with an appropriate pivoting strategy) computes a feasible complementary basis or determines that the linear complementarity problem is infeasible in a finite number of iterations.

2.2 Multiobjective Optimization Theory

For a general introduction to the topic of multiobjective optimization we refer to the books by Ehrgott [23] and Miettinen [62].
A multiobjective optimization problem is defined by m objective functions $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1,\dots,m$, and a feasible set $S \subseteq \mathbb{R}^n$:

$$\operatorname*{vmin}_{x\in S} f_i(x) \quad i = 1,\dots,m \tag{2.9}$$

Alternatively we define the vector-valued objective function $f(x)$, $f : \mathbb{R}^n \to \mathbb{R}^m$, component wise by

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}.$$

Definition 2.23. [23] A feasible solution $x \in S$ is called **efficient** or Pareto-optimal, if there does not exist any $x' \in S$ such that

$$f(x') \leq f(x)$$

In this case $f(x)$ is called **nondominated**. The set of efficient solutions is called the **efficient set** and is denoted by X_E. The set of nondominated points is denoted by Y_{nd}.
A feasible solution $x \in S$ is called **weakly efficient**, if there does not exist any $x' \in S$ such that

$$f(x') < f(x)$$

$f(x)$ is then called **weakly nondominated**.

Definition 2.24. Let Y_{nd} be the nondominated set of (2.9). The **ideal point** $y^I \in \mathbb{R}^m$ is defined by

$$y_i^I = \min\{y_i,\ y \in Y_{\text{nd}}\},\ i = 1,\ldots,m.$$

The **nadir point** $y^N \in \mathbb{R}^m$ is defined by

$$y_i^N = \max\{y_i,\ y \in Y_{\text{nd}}\},\ i = 1,\ldots,m.$$

The ideal point can be determined by computing the lexicographic minima of (2.9). The nadir point in general assumes knowledge of the nondominated set, and can thus only be computed efficiently in the biobjective case [23].
A common approach for computing efficient solutions of multiobjective optimization problems is to solve a scalarized problem. Two of such approaches will be used in this thesis: The weighted sum and the e-constraint scalarization [23].

2.2.1 The Weighted Sum Problem

The weighted sum problem of (2.9) for $\lambda \in \mathbb{R}^m$, $\lambda \geq 0$ is given by

$$\min_{x \in S} \sum_{i=1}^{m} \lambda_i f_i(x) \tag{2.10}$$

Theorem 2.25. [23] Let $\bar{x}$ be an optimal solution of (2.10) for $\lambda \in \mathbb{R}^m,\ \lambda \geq 0$.

1. $\bar{x}$ is weakly efficient.
2. If $\lambda > 0$, then $\bar{x}$ is efficient.
3. If $\bar{x}$ is the unique optimal solution of (2.10) then $\bar{x}$ is efficient.

Theorem 2.26. [23] Let all objective functions $f_i(x)$, $i = 1,\ldots,m$, be convex and let the feasible set S be convex.
For every efficient point $x \in X_E$ there exists $\lambda \in \mathbb{R}^m,\ \lambda \geq 0$ such that x is an optimal solution of the weighted sum problem (2.10) for λ.

The weighted sumscalarization can be used to compute all (weakly) efficient solutions of a multiobjective convex optimization problem (Theorem 2.25 and Theorem 2.26). In order to compute all efficient solution it is enough to consider the set following $(m-1)$-dimensional set [23]:

$$\Lambda = \left\{ \lambda \in \mathbb{R}^m \ :\ \lambda \geqq 0,\ \sum_{i=1}^{m} \lambda_i = 1 \right\}$$

Λ is called the **weight space** and can be parameterized by $m-1$ variables in multiple ways [23]:

2.2.1.1 Parameterization of the Weight Space

The most common way is to use the following parameterization [23]:

$$\bar{\Lambda} = \left\{ \lambda \in \mathbb{R}^m \ :\ \lambda \geqq 0,\ \lambda_m = 1 - \sum_{i=1}^{m-1} \lambda_i \right\}$$

A different parameterization has advantages in some situations:

$$W = \left\{ w \in \mathbb{R}^m \ : \ w = (1, w_2, \ldots, w_m)^T, \ w \geqq 0 \right\}.$$

Note that Λ is bounded, while W is unbounded. An example for a subset of weights W_1 in W and the corresponding counterpart Λ_1 in Λ can be seen in Figure 2.2. There exists a one-to-one correspondence between the points in $\bar{\Lambda}$ and points and rays in W.

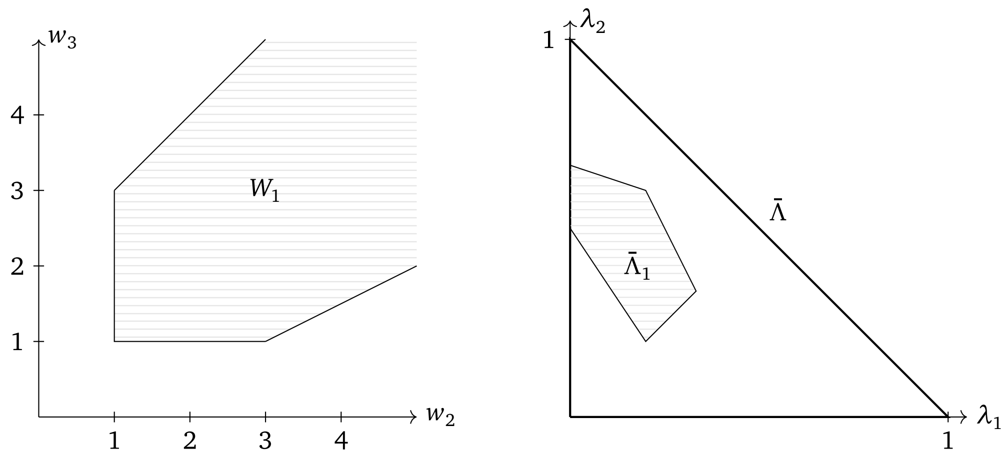

Figure 2.2: Parameterization of a set in W and Λ for $m = 3$.

2.2.2 The e-constraint Scalarization

The e-constraint scalarization problem of (2.9) is given by

$$\begin{array}{lll} \min\limits_{x \in S} & f_m(x) & \\ \text{s.t.} & f_i(x) \leq \varepsilon_i & i = 1, \ldots, m-1 \end{array} \tag{2.11}$$

The e-constraint scalarization problem can also be defined with any $f_i(x)$, $i = 1, \ldots, m$, as the objective function. Note that (2.11) can be infeasible, depending on the choice of ε.

Theorem 2.27. [23] Let $\bar{x}$ be an optimal solution of (2.11) for $\varepsilon \in \mathbb{R}^{m-1}$. Then the following statements hold:

1. $\bar{x}$ is weakly efficient.
2. If $\bar{x}$ is the unique optimal solution of (2.11) then $\bar{x}$ is efficient.

Theorem 2.28. [23] If $x \in X_E$ is an efficient solution, then there exists $\varepsilon \in \mathbb{R}^{m-1}$ such that x is the an optimal solution of the e-constraint scalarization problem.

Definition 2.29. [23] Let $f : \mathbb{R}^n \to \mathbb{R}^m$, $S \subseteq \mathbb{R}^n$ and $y \in \mathbb{R}^m$ be given. The level set of f on S at level y is defined as

$$\mathcal{L}(f, y, S) := \{x \in S \ : \ f(x) \leqq y\}$$

and the corresponding level curve is defined as

$$\mathcal{L}_=(f, y, S) := \{x \in S \ : \ f(x) = y\}$$

Furthermore, for $S = \mathbb{R}^n$ we denote $\mathcal{L}(f, y, \mathbb{R}^n)$ by $\mathcal{L}(f, y)$ and $\mathcal{L}_=(f, y, \mathbb{R}^n)$ by $\mathcal{L}_=(f, y)$.

Theorem 2.30. [24] Let $\bar{x} \in S$ be given with $y = f(\bar{x})$. Then $\bar{x}$ is efficient for (2.9) if and only if

$$\mathcal{L}(f, y, S) = \mathcal{L}_=(f, y, S).$$

Proposition 2.31. [23] If all objective functions $f_i(x)$ are convex and S is convex, then the efficient set is connected.

2.2.3 Multiobjective Linear Programming Problems

Consider a multiobjective linear programming problem in standard form:

$$\begin{aligned} \operatorname*{vmin}_{x \in \mathbb{R}^n} \quad & Cx \\ \text{s.t.} \quad & Ax = b, \ x \geqq 0 \end{aligned} \tag{MLP}$$

with cost matrix $C \in \mathbb{R}^{m \times n}$, matrix $A \in \mathbb{R}^{p \times n}$ of full rank, vector $b \in \mathbb{R}^p$ with $p \leq n$. The feasible set is denoted by $S := \{x \in \mathbb{R}^n \ : \ Ax = b, \ x \geqq 0\}$.
The weighted sum problem of (MLP) is given by

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \lambda^T Cx \\ \text{s.t.} \quad & Ax = b, \ x \geqq 0 \end{aligned} \tag{WLP}$$

Definition 2.32. [36] Let $B \subseteq \{1, \ldots, n\}$ be an index set and let A_B be the matrix containing the columns of A corresponding to B. B is called **feasible basis** of (MLP) if the matrix A_B is regular and

$$x_B := (A_B)^{-1} b \geqq 0. \tag{2.12}$$

Let N be the complement of B, i.e. $N = \{1, \ldots, n\} \setminus B$. The vector (x_B, x_N) with $x_N = 0$ and x_B defined as in (2.12) is called **basic feasible solution** of (MLP).

Definition 2.33. [27] A feasible basis $B \subseteq \{1, \ldots, n\}$ is called **efficient basis** of (MLP) if there exists $\lambda \in \Lambda$ such that (x_B, x_N) with $x_B := (A_B)^{-1} b$ and $x_N = 0$ is an optimal solution of the weighted sum scalarization problem (WLP) for λ.
The corresponding basic feasible solution $x = (x_B, x_N)$ is then called **efficient basic solution**.

Let $\mathcal{E}_E$ be the set of efficient basic solution of the multiobjective linear programming problem (MLP).

Definition 2.34. [48, 23] Let $x \in \mathcal{E}_E$ be an efficient basic solution of (MLP). Then the corresponding weight cell is defined by

$$\Lambda(x) = \left\{\lambda \in \Lambda \ : \lambda^T C x \leq \lambda^T C x' \ \forall x' \in S\right\}.$$

Theorem 2.35. [48, 23] Let $\mathcal{E}_E$ be the set of efficient basic feasible solution of the multiobjective linear programming problem (MLP). Then

$$\bigcup_{x \in \mathcal{E}_E} \Lambda(x) = \Lambda.$$

Chapter 3

Multiobjective Nonlinear Optimization and Descent Methods

Multiobjective nonlinear optimization problems have been discussed with the emergence of the field of multiobjective optimization and vector optimization. Two common solution approaches for these kind of problems are scalarization techniques and heuristic methods, such as evolutionary algorithms.

Scalarization approaches formulate a related singleobjective problem that can be solved with common scalar optimization methods and yield an optimal solution that is feasible and at least weakly efficient for the multiobjective problem. Some scalarizations have parameters that can be chosen to compute different efficient points. For these scalarizations also parametric optimization can be applied to compute an analytical description of the efficient set. Moreover, a variety of methods use preference information from a decision maker to provide solutions which are as close as possible to the preferences given, for example interactive methods and goal programming.
While evolutionary algorithms can be applied to scalarized problems, the direct application to the multiobjective case has the advantage that the final population computed by evolutionary methods already consists of multiple possibly efficient solutions [83].

This chapter reviews an alternative approach introduced by Fliege and Svaiter [30] that is different from the approaches discussed above in multiple ways. First, we discuss an optimality criterion for locally efficient points using descent directions. Many of the results in this section are generalizations of scalar nonlinear optimization theory for which a variety of books are available, for example the books from Carl Geiger and Christian Kanzow [36, 37], and the books about multiobjective optimization by Matthias Ehrgott [23] and Kaisa Miettinen [62].
After building a theoretical foundation a class of multiobjective descent methods is introduced. Several multiobjective descent methods that were discussed in the literature can be understood as special cases of this type of descent algorithm by changing a matrix that acts as a parameter in the algorithms. After a summary of the theoretical properties this method is compared to the weighted sum approach.

For this chapter we consider unconstrained multiobjective optimization problems in the following form

$$\operatorname*{vmin}_{x\in\mathbb{R}^n} \quad f(x) \tag{NLP}$$

with continuously differentiable objective functions $f : \mathbb{R}^n \to \mathbb{R}^m$, where

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}.$$

We have to consider a weaker definition of efficiency and nondominance for the nonlinear case. For this consider the neighborhood of $\bar{x}$ in $\mathbb{R}^n$ with radius ε defined as $U_\varepsilon(\bar{x}) := \{x \in \mathbb{R}^n \ : \ \|x - \bar{x}\| \leq \varepsilon\}$.

Definition 3.1. [62] Let $\bar{x} \in \mathbb{R}^n$ be given. $\bar{x}$ is called

1. **locally efficient**, if there exists $\varepsilon > 0$ such that there does not exists any $x \in U_\varepsilon(\bar{x})$ such that $f(x) \leq f(\bar{x})$.
2. **locally weakly efficient**, if there exists $\varepsilon > 0$ such that there does not exists any $x \in U_\varepsilon(\bar{x})$ such that $f(x) < f(\bar{x})$.

In Section 3.1 we consider an optimality criterion for efficient solutions of multiobjective nonlinear optimization problems. An extension of singleobjective descent methods to the multiobjective case proposed by Fliege et. al. [30] is reviewed in Sections 3.2 and 3.3 and a convergence result from the literature [34] is stated. We consider different choices for the search direction in Sections 3.3.1, 3.3.2 and 3.3.3 that can be understood as extensions of the steepest descent method, Newton method and Quasi-Newton method from the singleobjective to the multiobjective case, respectively.
In Section 3.3.4 descent directions are considered that use a weighted sum of the gradients of the objective functions. A compromise descent method is proposed that chooses descent directions by normalizing the gradients of the objective functions. The differences between the methods discussed in this chapter are visualized with a concrete example in Section 3.4. Finally, we compare the application of a singleobjective descent method to the weighted sum problem and underline the differences between these approaches in Section 3.4.1.

3.1 Descent Directions and Optimality Conditions

In this section optimality conditions for the multiobjective nonlinear optimization problem (NLP) are reviewed.

Definition 3.2. [30] Let $x \in \mathbb{R}^n$. A vector $d \in \mathbb{R}^n$ is called (multiobjective) **descent direction** of f in x if

$$\exists t_0 > 0 \ : \ f(x+td) < f(x) \ \forall t \in (0, t_0].$$

Proposition 3.3. [30] Let $x \in \mathbb{R}^n$ be a point. If $d \in \mathbb{R}^n$ satisfies

$$\nabla f_i(x)^T d < 0 \ \forall i = 1, \dots, m,$$

then d is a descent direction of f in x.

Proof. This result is a direct generalization of scalar nonlinear optimization theory (see for example Geiger and Kanzow [37]). □

We denote the descent cone in a point $x \in \mathbb{R}^n$ by

$$D(x) = \left\{ d \in \mathbb{R}^n \ : \ \nabla f_i(x)^T d < 0 \ \forall i = 1, \dots, m \right\}.$$

Figure 3.1 illustrates the gradients of two objective functions and descent cone $D(x)$ for $n = m = 2$.

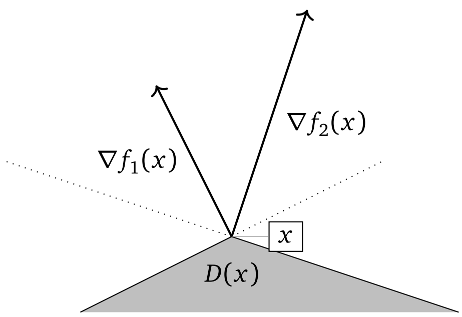

Figure 3.1: Descent cone for $m = 2$ at a point $x \in S$.

Proposition 3.4 (Necessary Optimality Conditions with Cones). [30] If x is a (locally) weakly efficient point of (NLP) then

$$D(x) = \emptyset.$$

Now we formulate an algebraic optimality condition that can either be derived from scalar optimization theory which is applied to the weighted sum problem of (NLP) or directly from the descent cone $D(x)$.

Lemma 3.5 (Necessary Optimality Condition). [62] Let $\bar{x} \in \mathbb{R}^n$ with $D(\bar{x}) = \emptyset$. Then there exists $\lambda \in \mathbb{R}^m$ with $\lambda \geq 0$ such that

$$\sum_{i=1}^{m} \lambda_i \nabla f_i(x) = 0.$$

Definition 3.6. [30] $x \in \mathbb{R}^n$ is called **critical** (or Pareto-critical) if there exists $\lambda \in \Lambda$ such that

$$\sum_{i=1}^{m} \lambda_i \nabla f_i(x) = 0.$$

Lemma 3.7 (Sufficient Optimality Condition for Convex Problems). [62] Let f be convex on $\mathbb{R}^n$. Then any critical point of (NLP) is weakly efficient for (NLP).

A general descent algorithm consists of two main steps. In each iteration a descent direction $d \in \mathbb{R}^n$ at the current iterate $x \in \mathbb{R}^n$ is chosen, if possible, and then a step size $t > 0$ is determined such that $f(x + td) < f(x)$. The algorithm terminates if no descent direction can be computed.

In Section 3.2 a generalization of the well-known Armijo step size to the multiobjective case [30] is reviewed followed by a short discussion about other choices of the step size. Thereafter several choices for the descent direction are introduced in Section 3.3.

3.2 Step Sizes

Assume a descent direction $d \in D(x)$ is given. A sufficient descent in the scalar case can be achieved by choosing a step size t by the **Armijo rule** which requires that t is chosen satisfying the following equation for some parameter $\sigma \in (0,1)$ [36]:

$$f(x+td) < f(x) + \sigma t \nabla f(x)^T d \tag{3.1}$$

The Armijo rule can be generalized to the multiobjective case by applying it for each objective function separately [30].

Definition 3.8. Let $d \in \mathbb{R}^n$ be a descent direction of $f(x)$ in x. $t > 0$ is called **Armijo step-size** for $\sigma \in (0,1)$ if

$$f_i(x+td) < f_i(x) + \sigma t \nabla f_i(x)^T d \;\; \forall i = 1, \dots, m.$$

The existence of a step size satisfying the Armijo rule can be shown in a similar way as in the scalar case:

Proposition 3.9. [36, 30, 31] Let $x \in \mathbb{R}^n$ be a point and let $d \in \mathbb{R}^n$ be a descent direction of f in x and let $\sigma \in (0,1)$. Then there exists some $t_0 > 0$ such that

$$f_i(x+td) < f_i(x) + \sigma t \nabla f_i(x)^T d \;\; \forall i = 1, \dots, m$$

holds for all $t \in (0, t_0]$.

Proof. A proof can be found in [30]. The statement follows directly from the scalar results (see for example [36]). □

To compute an Armijo step size the following algorithm can be used (with a step parameter $\beta \in (0,1)$) [30]:

Algorithm 3.1: Computing an Armijo step size

Input: Input f, point $x \in \mathbb{R}^n$, descent direction $d \in \mathbb{R}^n$, parameters $\sigma \in (0,1)$ and $\beta \in (0,1)$.
Set $t = 1$.
while $\exists i \in \{1, \dots, m\}$ *with* $f_i(x+td) \geq f_i(x) + \sigma t \nabla f_i(x)^T d$ **do**
  Set $t = \beta t$

The Armijo rule ensures a sufficient descent but may result in relatively small step sizes. In singleobjective optimization the **Wolfe-Powell step size** is used to avoid this behavior, by demanding that the slope of $f(x+td)$ at $t>0$ is not as steep as at $t=0$ [36]. The Wolfe-Powell conditions with $\rho \in (0,1)$ are given by

$$\begin{aligned} & f(x+td) \leq f(x) + \sigma t \nabla f(x)^T d \text{ (Armijo rule)} \\ & \nabla f(x+td)^T d \geq \rho \nabla f(x)^T d \end{aligned} \tag{3.2}$$

The set of step sizes satisfying the Wolfe-Powell conditions is denoted by T_{WP}. Figure 3.2 illustrates the definition of step sizes satisfying the Wolfe-Powell conditions. The dashed line is the graph of $f(x)+t\sigma\nabla f(x)^T d$ and is often referred to as the **Armijo-Goldstein line** [36]. At $t=\frac{1}{2}$ the tangent has the slope $\rho\nabla f(x)^T d$.

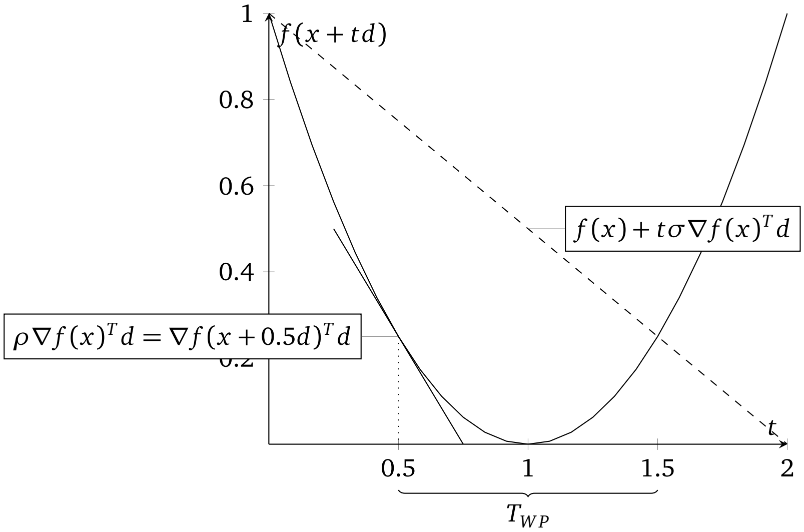

Figure 3.2: Wolfe-Powell step sizes.

The existence of a Wolfe-Powell step size is guaranteed in the singleobjective case if f is bounded from below [36]. Unfortunately this result can not be generalized to the multiobjective case by demanding that t is a Wolfe-Powell step size for every objective function as the following example shows:

Example 3.10. Consider the following biobjective problem:

$$\text{vmin}_{x\in\mathbb{R}} \quad \begin{aligned} f_1(x) &= 3x^2 - x + 1 \\ f_2(x) &= x^2 - 2x + 1 \end{aligned}$$

at the point $x=0$ with the descent direction $d=1$ and parameters $\sigma=\frac{1}{4}$ and $\rho=\frac{1}{2}$. On the one hand, to satisfy the Armijo rule for f_1 the step size t has to satisfy

$$f_1(x+td) \leq f_1(x) + \sigma t \nabla f_1(x)^T d \Leftrightarrow t \leq \frac{1}{4}.$$

On the other hand, t has to satisfy the Wolfe-Powell conditions, specifically

$$\nabla f_2(x+td) \geq \rho \nabla f_2(x)^T d \Leftrightarrow t \geq \frac{1}{2},$$

which contradicts the Armijo rule for f_1. Thus, there is no Wolfe-Powell step size of f in x in direction d.

Other choices for the step size cannot be extended to the multiobjective case for similar reasons. Consider for example the exact and Curry step size [36]:

$$t_e = \arg\min_{t>0} f(x+td) \text{ and } t_c = \min\left\{t > 0 \ : \ \nabla f(x+td)^T d = 0\right\}.$$

It is not clear how to extend these definitions to the multiobjective case. One could choose t such that $f(x+td)$ is a critical point of

$$\operatorname{vmin}_{t>0} f(x+td)$$

but it is not guaranteed that t yields a descent in every objective.

For these reasons we will only consider the Armijo rule for the selection of step sizes.

3.3 Search Directions

For a given point $x \in \mathbb{R}^n$ we want to determine if x is critical or, if not, compute a descent direction. Demanding that d is a descent direction is equivalent to

$$\max_{i=1,\dots,m} \nabla f_i(x)^T d < 0. \tag{3.3}$$

In accordance with the concept of steepest descent in scalar nonlinear optimization one might consider minimizing (3.3) in the following way:

$$\min_{d\in\mathbb{R}^n} \quad \max_{i=1,\dots,m} \nabla f_i(x)^T d \tag{3.4}$$

Note that if x is critical then the optimal solution of (3.4) is $d = 0$. But (3.4) is unbounded if x is not critical [30]. Some feasible solutions of (3.4) are descent directions but not all. In order to compute a particular descent direction we introduce a normalization term. There are many options available, including a linear normalization term [30]. For the discussion of known methods we will only consider a quadratic normalization term of the form $d^T H^i d$ for every objective $i = 1, \dots, m$, where each H^i is a symmetric positive definite matrix that is chosen either constant for every iteration or chosen differently in every iteration. We will see that the choice of H^i leads to descent algorithms motivated by different methods in scalar nonlinear optimization.

$$\min_{d\in\mathbb{R}^n} \quad \max_{i=1,\dots,m} \nabla f_i(x)^T d + \frac{1}{2} d^T H^i d \tag{3.5}$$

Since the optimization problem (3.5) has an objective function that is not differentiable everywhere we consider the differentiable formulation as a convex nonlinear optimization problem with quadratic constraints by introducing an additional variable τ:

$$\begin{array}{lll} \min\limits_{d\in\mathbb{R}^n,\tau\in\mathbb{R}} & \tau & \\ \text{s.t.} & \nabla f_i(x)^T d + \frac{1}{2} d^T H^i d \quad \leq \quad \tau & \forall i = 1,\dots,m \end{array} \tag{3.6}$$

Notice that the variable τ is used to ensure that any optimal solution (d, τ) of (3.6) with $\tau < 0$ is a descent direction.
The following results for the search direction problem (3.6) have been shown for special choices of the normalization term [29, 70] in the literature:

Proposition 3.11. [29, 70] Let $x \in \mathbb{R}^n$. Then the following statements hold:

1. (3.6) has a unique optimal solution (d^*, τ^*) with $\tau^* \leq 0$.
2. $\tau^* = 0 \Leftrightarrow d^* = 0$
3. If $\tau^* < 0$, then d^* is a descent direction of (NLP).

Proof. We follow the proof from [70] with the appropriate choice of the matrices H^i for $i = 1, \ldots, m$.

1. $(d, \tau) = (0, 0)$ is feasible for (3.6) and provides an upper bound for the optimal objective value of (3.6). The nondifferentiable formulation (3.5) has a strictly convex objective function (since all H^i, $i = 1, \ldots, m$, are positive definite) and linear inequality constraints. Thus, (3.6) has a unique optimal solution with $\tau^* \leq 0$.

2. We consider the two directions separately:

   '⇒': Let (d, τ^*) be an optimal solution of (3.6) with $\tau^* = 0$. Since d is feasible for (3.6) it must satisfy
   $$\nabla f_i(x)^T d + \frac{1}{2} d^T H^i d \leq \tau^* = 0 \ \forall i = 1, \ldots, m,$$
   which is equivalent to
   $$\frac{1}{2} d^T H^i d \leq -\nabla f_i(x)^T d \ \forall i = 1, \ldots, m. \tag{3.7}$$
   If there exists $i \in \{1, \ldots, m\}$ with $\nabla f_i(x) = 0$, then we can observe form (3.7) that
   $$\frac{1}{2} d^T H^i d \leq 0,$$
   which can only hold if $d = 0$ since H^i is positive definite for all $i = 1, \ldots, m$.
   Now, consider the case where $\nabla f_i(x) \neq 0$ for all $i = 1, \ldots, m$. Let $\alpha \in (0, 1)$ be given and observe that $(\alpha d, \tau^*)$ is a feasible solution of (3.6). Assume that $d \neq 0$. Using (3.7) we can see the following:
   $$\begin{aligned} \nabla f_i(x)^T (\alpha d) + \frac{1}{2} (\alpha d)^T H^i (\alpha d) &= \alpha \nabla f_i(x)^T d + \alpha^2 \underbrace{\frac{1}{2} d^T H^i d}_{\leq -\nabla f_i(x)^T d} \\ &\leq \underbrace{(\alpha - \alpha^2)}_{\in(0,1)} \underbrace{\nabla f_i(x)^T d}_{<0} \\ &< 0 \ \forall i = 1, \ldots, m \end{aligned}$$
   Which is a contradiction to the assumption that $\tau^* = 0$. Hence, $d = 0$.

   '⇐': Let (d, τ^*) be a optimal solution of (3.6) with $d = 0$. All constraints of (3.6) demand that $\tau^* \geq 0$. Thus, $\tau^* = \min(\{t \ : \ t \geq 0\}) = 0$.

3. Let (d, τ^*) be a optimal of (3.6) with $\tau^* < 0$. By construction d is a feasible descent direction of (NLP).

□

The following property has been shown by Fliege et. al. [29] for the choice $H^i = \nabla^2 f_i(x)$ and strongly convex objective functions $f_i(x)$ for $i = 1, \ldots, m$. Povalej [70] extended this result to the general case where all H^i are chosen positive definite.

Proposition 3.12. [29, 70] Let $x \in \mathbb{R}^n$ be given and let H^i be positive definite for all $i = 1, \ldots, m$. Let (d^*, τ^*) be the optimal solution of (3.6). Then there exists $\lambda \in \Lambda$ such that

$$d^* = -\left[\sum_{i=1}^{m} \lambda_i H^i\right]^{-1} \sum_{i=1}^{m} \lambda_i \nabla f_i(x). \tag{3.8}$$

Proof. We follow the proof in [70] with the appropriate choice of the matrices H^i for $i = 1, \ldots, m$.: According to Proposition 3.11 the search direction problem (3.6) has an optimal solution for every $x \in \mathbb{R}^n$. Then (d^*, τ^*) is a Fritz-John point of (3.6) (see Theorem 2.4). Thus, there exists $\lambda \geqq 0$ such that

$$\sum_{i=1}^{m} \lambda_i(\nabla f_i(x) + H^i d^*) = 0, \ \sum_{i=1}^{m} \lambda_i = 1 \text{ and} \tag{3.9}$$

$$\left(\nabla f_i(x)^T d^* + \frac{1}{2}(d^*)^T H^i d^* - \tau^*\right) \lambda_i = 0 \ \forall i = 1, \ldots, m. \tag{3.10}$$

Notice that the matrix

$$\sum_{i=1}^{m} \lambda_i H^i$$

is positive definite for all $\lambda \geq 0$. Thus, using (3.9) we get

$$\begin{aligned}\sum_{i=1}^{m} \lambda_i(\nabla f_i(x) + H^i d^*) = 0 &\Leftrightarrow \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + \sum_{i=1}^{m} \lambda_i H^i d^* = 0 \\ &\Leftrightarrow \sum_{i=1}^{m} \lambda_i H^i d^* = -\sum_{i=1}^{m} \lambda_i \nabla f_i(x) \\ &\Leftrightarrow d^* = -\left[\sum_{i=1}^{m} \lambda_i H^i\right]^{-1} \sum_{i=1}^{m} \lambda_i \nabla f_i(x).\end{aligned}$$

□

Theorem 3.13. [29, 70] Let $\bar{x} \in \mathbb{R}^n$ be given and let (d^*, τ^*) be the unique optimal solution of (3.6). Then the following statements are equivalent:

1. $\bar{x}$ is a critical point of (NLP).
2. $\tau^* = 0$

Proof. We follow the proof from [70] with the appropriate choice of the matrices H^i for $i = 1, \ldots, m$.

'⇒' This statement follows directly from Proposition 3.11.

'⇐' Proposition 3.12 shows that if (d, τ^*) is the optimal solution of (3.6) then there exists $\lambda \in \Lambda$ such that

$$\sum_{i=1}^{m} \lambda_i(\nabla f_i(\bar{x}) + H^i d) = 0 \tag{3.11}$$

Additionally, we know from the second part of Proposition 3.11 that if $\tau^* = 0$ holds then $d = 0$ and the terms $H^i d$ vanish for all $i = 1, \ldots, m$ in (3.11). Hence, we know there exists $\lambda \geq 0$ and such that

$$\sum_{i=1}^{m} \lambda_i \nabla f_i(\bar{x}) = 0,$$

which is equivalent to the condition for a critical point $\bar{x}$ of (NLP).

□

Proposition 3.14. [8] Let $A \in \mathbb{R}^{n\times n}$ be a positive definite matrix and let $\boldsymbol{a}_{\min}(A)$ and $\boldsymbol{a}_{\max}(A)$ be the smallest and largest eigenvalues of A, respectively. Then

$$\boldsymbol{a}_{\min}(A)\, x^T x \leq x^T A x \leq \boldsymbol{a}_{\max}(A)\, x^T x \;\forall x \in \mathbb{R}^n.$$

The following proposition is an extension of Lemma 3.2 in Fliege et. al. [29] and Lemma 2 from Povalej [70] to allow for a broader choice of matrices H^i in the normalization term $\frac{1}{2} d^T H^i d$ for $i = 1, \dots, m$.

Proposition 3.15. [29, 34] Let $X \subset \mathbb{R}^n$ be a compact set and let $H^i(x)$ be symmetric positive matrices such that the following conditions are satisfied:

1. For every $\theta > 0$ there exists $\delta > 0$ such that for all $x, x' \in X$ the following statement holds:

$$\|x - x'\| \leq \delta \Rightarrow \|H^i(x) - H^i(x')\| \leq \theta \;\forall i = 1, \dots, m. \tag{3.12}$$

2. There exists a constant $\hat{\boldsymbol{a}} > 0$ such that

$$\hat{\boldsymbol{a}} = \min_{x\in X,\ \|v\|=1,\ i=1,\dots,m} v^T H^i(x) v. \tag{3.13}$$

Let $(d(x), \tau(x))$ be the optimal solution of (3.6) for $x \in X$. Then the following statements hold:

1. The mapping $x \rightarrow \tau(x)$ is continuous and $d(x)$ is bounded on X.
2. The mapping $x \rightarrow d(x)$ is continuous on X.

Proof. This proof is very similar to the proof of Proposition 5.5 of the survey paper by Fukuda et. al. [34]. Since each objective function f_i is continuously differentiable and the set X is compact there exists a finite constant $\gamma > 0$ such that $\|\nabla f_i(x)\| \leq \gamma$ for all $x \in X$ and all $i = 1, \dots, m$.
Let $x \in X$ be given. Observe from (3.7) that

$$\frac{1}{2} d(x)^T H^i d(x) \leq -\nabla f_i(x)^T d(x) \;\forall i = 1, \dots, m.$$

Using the bounds $\hat{\boldsymbol{a}}$ and γ and the Cauchy-Schwarz inequality we see that

$$\frac{1}{2} \hat{\boldsymbol{a}} \|d(x)\|^2 \leq \|\nabla f_i(x)\| \cdot \|d(x)\| \leq \gamma \|d(x)\|$$

holds for every $i = 1, \dots, m$ and $x \in X$. Which shows that $\|d(x)\|$ is bounded by $\kappa = 2\frac{\gamma}{\hat{\boldsymbol{a}}}$, i.e.

$$\|d(x)\| \leq \kappa = 2\frac{\gamma}{\hat{\boldsymbol{a}}} \tag{3.14}$$

Using Proposition 3.12 we know there exists $\lambda \in \Lambda$ such that

$$\sum_{i=1}^{m} \lambda_i \nabla f_i(x) = -\sum_{i=1}^{m} \lambda_i H^i(x) d(x) \tag{3.15}$$

and such that the complementarity condition (3.10) is satisfied. Notice that $\lambda_j > 0$ for some $j \in \{1, \dots, m\}$ implies that

$$\tau(x) = \nabla f_j(x)^T d(x) + \frac{1}{2} d(x)^T H^j(x) d(x).$$

Using this fact and (3.15) and (3.13) we see that

$$\begin{aligned}\tau(x) &= \sum_{i=1}^{m}\Big(\lambda_i \nabla f_i(x)^T d(x) + \frac{1}{2} d(x)^T H^i(x) d(x)\Big)\\ &= \underbrace{\Big(\sum_{i=1}^{m}\lambda_i \nabla f_i(x)\Big)^T}_{=-\sum_{i=1}^{m}\lambda_i H^i(x)d(x)} d(x) + \frac{1}{2}\sum_{i=1}^{m}\lambda_i d(x)^T H^i(x) d(x)\\ &= -\frac{1}{2} d(x)^T \sum_{i=1}^{m}\lambda_i H^i(x) d(x). \end{aligned} \tag{3.16}$$

Using Proposition 3.14 and (3.13) notice that

$$\sum_{i=1}^{m}\lambda_i d(x)^T H^i(x) d(x) \geq \sum_{i=1}^{m}\lambda_i \underbrace{\boldsymbol{a}_{\min}(H^i(x))}_{\geq \hat{\boldsymbol{a}}} \|d(x)\|^2 \geq \sum_{i=1}^{m}\lambda_i \hat{\boldsymbol{a}} \|d(x)\|^2 = \hat{\boldsymbol{a}} \|d(x)\|^2. \tag{3.17}$$

Applying (3.17) to (3.16) leads to

$$\tau(x) = -\frac{1}{2} d(x)^T \sum_{i=1}^{m}\lambda_i H^i(x) d(x) \leq -\frac{\hat{\boldsymbol{a}}}{2} \|d(x)\|^2, \tag{3.18}$$

which is equivalent to

$$\|d(x)\|^2 \leq \frac{2}{\hat{\boldsymbol{a}}} \|\tau(x)\|. \tag{3.19}$$

1. Let $\delta > 0$ and $x, y \in X$ be given such that $\|x - y\| \leq \delta$. Let $i_0 \in \{1, \dots, m\}$ such that

$$\max_{i=1,\dots,m} \nabla f_i(y)^T d(y) + \frac{1}{2} d(y)^T H^i(y) d(y) = \nabla f_{i_0}(y)^T d(y) + \frac{1}{2} d(y)^T H^{i_0}(y) d(y).$$

Notice that

$$\begin{aligned}\tau(y) &= \nabla f_{i_0}(y)^T d(y) + \frac{1}{2} d(y)^T H^{i_0}(y) d(y)\\ &\leq \nabla f_{i_0}(y)^T d(x) + \frac{1}{2} d(x)^T H^{i_0}(y) d(x)\\ &= \nabla f_{i_0}(y)^T d(x) - \nabla f_{i_0}(x)^T d(x) + \nabla f_{i_0}(x)^T d(x) + \frac{1}{2} d(x)^T H^{i_0}(y) d(x)\\ &\qquad + \frac{1}{2} d(x)^T H^{i_0}(x) d(x) - \frac{1}{2} d(x)^T H^{i_0}(x) d(x)\\ &= \left(\nabla f_{i_0}(y) - \nabla f_{i_0}(x)\right)^T d(x) + \frac{1}{2} d(x)^T \left(H^{i_0}(y) - H^{i_0}(x)\right) d(x) + \underbrace{\nabla f_{i_0}(x)^T d(x) + \frac{1}{2} d(x)^T H^{i_0}(x) d(x)}_{\leq \tau(x)}\\ &\leq \left\|\nabla f_{i_0}(y) - \nabla f_{i_0}(x)\right\| \|d(x)\| + \frac{1}{2}\|d(x)\|^2 \left\|H^{i_0}(y) - H^{i_0}(x)\right\| + \tau(x). \end{aligned}$$

By interchanging the roles of x and y and using (3.14) and (3.12) we see that

$$\|\tau(y) - \tau(x)\| \leq \underbrace{\left\|\nabla f_{i_0}(y) - \nabla f_{i_0}(x)\right\|}_{\leq \gamma} \underbrace{\|d(x)\|}_{\leq \kappa} + \frac{1}{2} \underbrace{\|d(x)\|^2}_{\leq \kappa^2} \underbrace{\left\|H^{i_0}(y) - H^{i_0}(x)\right\|}_{\leq \theta},$$

which shows that the mapping $x \to \tau(x)$ is continuous on X.

2. We will now show that the mapping $x \to d(x)$ is continuous as well. Let $x, y \in X$ be defined as above.

   Using the Cauchy-Schwarz inequality we see that

$$(\nabla f_i(y) - \nabla f_i(x))^T d(x) \le \|\nabla f_i(y) - \nabla f_i(x)\| \, \|d(x)\|$$
$$\Leftrightarrow (\nabla f_i(x) - \nabla f_i(y))^T d(x) \ge - \underbrace{\|\nabla f_i(y) - \nabla f_i(x)\|}_{\le \gamma} \, \|d(x)\|$$

   and therefore

$$(\nabla f_i(x) - \nabla f_i(y))^T d(x) \ge -\gamma \, \|d(x)\| . \tag{3.20}$$

   Using (3.20) we observe that

$$\begin{aligned}
\tau(x) &= \max_{i=1,\ldots,m} \left(\nabla f_i(x)^T d(x) + \frac{1}{2} d(x)^T H^i(x) d(x) \right) \\
&= \max_{i=1,\ldots,m} \Big(\nabla f_i(x)^T d(x) - \nabla f_i(y)^T d(x) + \nabla f_i(y)^T d(x) \\
&\qquad + \frac{1}{2} d(x)^T H^i(x) d(x) - \frac{1}{2} d(x)^T H^i(y) d(x) + \frac{1}{2} d(x)^T H^i(y) d(x) \Big) \\
&= \max_{i=1,\ldots,m} \Big((\nabla f_i(x) - \nabla f_i(y))^T d(x) + \frac{1}{2} d(x)^T (H^i(x) - H^i(y)) d(x) \\
&\qquad + \nabla f_i(y)^T d(x) + \frac{1}{2} d(x)^T H^i(y) d(x) \Big) \\
&\ge -\gamma \, \|d(x)\| - \frac{1}{2} \theta \, \|d(x)\|^2 + \max_{i=1,\ldots,m} \left(\nabla f_i(y)^T d(x) + \frac{1}{2} d(x)^T H^i(y) d(x) \right).
\end{aligned} \tag{3.21}$$

   Consider the following functions for $i = 1, \ldots, m$ and $z \in \mathbb{R}^n$:

$$g_i(z) := \nabla f_i(y)^T z + \frac{1}{2} z^T H^i(y) z.$$

   Notice that $\boldsymbol{a}_{\min}(\nabla^2 g_i(z)) \ge \hat{\boldsymbol{a}}$ for all $i = 1, \ldots, m$ and $z \in X$. Thus, $g_i(z)$ is strongly convex with modulus $\hat{\boldsymbol{a}}$, which is also the case for $\max_{i=1,\ldots,m} g_i(z)$ for all $z \in X$. Now, consider the following inequality using the strong convexity and the fact that $\max_{i=1,\ldots,m} g_i(z)$ attains its minimum at $d(y)$:

$$\begin{aligned}
&\max_{i=1,\ldots,m} \left(\nabla f_i(y)^T d(x) + \frac{1}{2} d(x)^T H^i(y) d(x) \right) \\
\ge &\max_{i=1,\ldots,m} \left(\nabla f_i(y)^T d(y) + \frac{1}{2} d(y)^T H^i(y) d(y) \right) + \frac{\hat{\boldsymbol{a}}}{2} \|d(x) - d(y)\|^2 \\
= &\tau(y) + \frac{\hat{\boldsymbol{a}}}{2} \|d(x) - d(y)\|^2
\end{aligned} \tag{3.22}$$

   Applying (3.22) to (3.21) leads to the following inequality:

$$\tau(x) \ge -\gamma \, \|d(x)\| - \frac{1}{2} \theta \, \|d(x)\|^2 + \tau(y) + \frac{\hat{\boldsymbol{a}}}{2} \|d(x) - d(y)\|^2$$

   Which is equivalent to

$$\tau(x) - \tau(y) \ge -\gamma \, \|d(x)\| - \frac{1}{2} \theta \, \|d(x)\|^2 + \frac{\hat{\boldsymbol{a}}}{2} \|d(x) - d(y)\|^2 . \tag{3.23}$$

   Now, let $(x^k)_k$ be a sequence with

$$\lim_{k \to \infty} x^k = y.$$

The mapping $x \to \tau(x)$ is continuous and thus $(\tau(x^k))_k$ converges against $\tau(y)$. Using (3.19) and (3.23) it follows that the sequence $(d(x^k))_k$ also converges, with

$$\lim_{k\to\infty} d(x^k) = d(y),$$

showing that the mapping $x \to d(x)$ is continuous on X.

□

The assumptions of Proposition 3.15 are for example satisfied for any constant choice of H^i.
Using Proposition 3.11 and Theorem 3.13 we can now check whether a given point is critical or not, and if not, compute a descent direction. As a stopping criterion the optimal solution value τ^* of the search direction problem (3.6) or the norm of d^* of the corresponding optimal solution of (3.6) can be checked.
Using a choice of positive definite matrices H^i for $i = 1, \dots, m$ that satisfy the assumptions of Proposition 3.15 we can now formulate a general multiobjective descent method with Armijo step size.

Algorithm 3.2: General Multiobjective Descent Method

Choose a point $x^0 \in \mathbb{R}^n$ and $\sigma \in (0,1)$. Set $k = 0$.
while *Stopping criterion is violated at* x^k **do**
  Choose a descent direction $d^k \in \mathbb{R}^n$ as the solution of (3.6)
  Compute a step size t_k satisfying the Armijo rule (3.1) using Algorithm 3.1
  Iterate: $x^{k+1} = x^k + t_k d^k$
  Set $k = k + 1$

Theorem 3.16. [30, 70] Let $(x^k)_k$ be a sequence produced by Algorithm 3.2 with matrices $H^i(x^k)$ satisfying the assumptions of Proposition 3.15 on a compact set $X \subseteq \mathbb{R}^n$ with $x^k \in X$ for all $k \in \mathbb{N}$. If the level sets $\mathcal{L}(f, f(x^1))$ are bounded then every accumulation point of $(x^k)_k$ is a critical solution of (NLP).

Proof. The proof for the choice $H^i = I_n$ can be found in Fliege and Svaiter [36]. The proof can be extended to cover the general case for positive definite matrices $H^i(x)$ [34]:
Let $\tau(x^k)$ be the optimal objective value of the search direction problem (3.6) in iteration k. First, notice that the sequence $(f_i(x^k))_k$ is strictly decreasing for all $i = 1, \dots, m$. Let $\bar{x}$ be an accumulation point of $(x^k)_k$. Since all functions f_i are continuous functions we can see that

$$\lim_{k\to\infty} f(x^k) = f(\bar{x}) \text{ and } \lim_{k\to\infty} \|f(x^k) - f(x^{k+1})\| = 0.$$

Since the Armijo rule is satisfied in every iteration we also have in each iteration k

$$f_i(x^k) - f_i(x^{k+1}) \geq -t_k \sigma \nabla f_i(x^k)^T d^k \geq 0 \ \forall i = 1, \dots, m$$

and thus

$$\lim_{k\to\infty} t_k \nabla f_i(x^k)^T d^k = 0 \ \forall i = 1, \dots, m \tag{3.24}$$

Consider two cases:

- $\limsup_{k\to\infty} t_k > 0$: In this case there exists a converging subsequence $(x^l)_l$ of $(x^k)_k$ such that $\lim_{l\to\infty} x^l = \bar{x}$ and $\lim_{l\to\infty} t^l = \bar{t} \geq 0$. For every $i = 1, \dots, m$ we have that

$$0 = \lim_{l\to\infty} \nabla f_i(x^l)^T d^l \leq \lim_{l\to\infty} \left[\max_{i=1,\dots,m} \nabla f_i(x^l)^T d^l + \frac{1}{2} (d^l)^T H^i(x) d^l \right] \leq \lim_{l\to\infty} \tau(x^l) = \tau(\bar{x}).$$

  Using that $\tau(x) \leq 0$ for all $x \in \mathbb{R}^n$ (see Proposition 3.11) it follows that $\tau(\bar{x}) = 0$ and using Theorem 3.13 we can conclude that $\bar{x}$ is critical.

- $\limsup_{k\to\infty} t_k = 0$: We know from Proposition 3.15 that the mapping $x \to d(x)$ is continuous and, thus, that the sequence $d(x^k)$ is bounded and has a converging subsequence. Let $(x^l)_l$ be such a subsequence with $(d(x^l))_l$ converging to $\bar{d}$ and $\lim_{l\to\infty} t_l = 0$. Notice that for all $l = 1, \dots$

$$\max_{i=1,\dots,m} \nabla f_i(x^l)^T d^l \le \max_{i=1,\dots,m} \nabla f_i(x^l)^T d^l + \frac{1}{2}(d^l)^T H^i(x) d^l \le \tau(x^l) < 0$$

and in the limit $l \to \infty$

$$\max_{i=1,\dots,m} \nabla f_i(x^l)^T \bar{d} \le \tau(\bar{x}) \le 0 \tag{3.25}$$

Let r be an arbitrary but fixed positive integer. Since the sequence $(t_l)_l$ converges to 0 we have for large enough l that

$$t_l < \beta^r$$

which shows that the Armijo rule is violated for $t = \beta^r$ at some x^l, i.e. for all sufficiently large l there exists $i \in \{1, \dots, m\}$ such that

$$f_i(x^l + \beta^r d^l) \ge f_i(x^l) + \sigma\beta^r \nabla f_i(x^l)^T d^l.$$

Considering an appropriate subsequence and after taking the limit we see that that for some $i \in \{1, \dots, m\}$:

$$f_i(\bar{x} + \beta^r \bar{d}) \ge f_i(\bar{x}) + \sigma\beta^r \nabla f_i(\bar{x})^T \bar{d}$$

which is true for any integer $r > 0$. Using Proposition 3.9 it follows that

$$\max_{i=1,\dots,m} \nabla f_i(x)^T \bar{d} \ge 0.$$

Using (3.25) we conclude that $\tau(\bar{x}) = 0$ and thus show that $\bar{x}$ is critical.

□

3.3.1 Steepest Descent Method

One of the first multiobjective descent methods discussed in the literature was a generalization of the steepest descent method from scalar nonlinear optimization by Fliege and Svaiter [30]. This method will be introduced now and we will discuss some general properties of Algorithm 3.2 using the steepest descent variant. Setting $H^i = I_n$ for $i = 1, \dots, m$ results in the following search direction problem:

$$\begin{array}{lll} \min\limits_{d\in\mathbb{R}^n,\tau\in\mathbb{R}} & \tau & \\ \text{s.t.} & \nabla f_i(x)^T d + \frac{1}{2}\underbrace{d^T I_n d}_{=\|d\|_2^2} \le \tau & \forall i = 1, \dots, m \end{array} \tag{3.26}$$

Notice that the regularization term is now independent of the index $i \in \{1, \dots, m\}$ and the search direction problem (3.26) can be reformulated as a convex quadratic optimization problem:

$$\begin{array}{lll} \min\limits_{d\in\mathbb{R}^n,\tau\in\mathbb{R}} & \tau + \frac{1}{2}\|d\|_2^2 & \\ \text{s.t.} & \nabla f_i(x)^T d \le \tau & \forall i = 1, \dots, m \end{array} \tag{3.27}$$

Since (3.27) is a strictly convex quadratic programming problem we can also consider it's quadratic dual problem [30]:

$$\begin{array}{ll} \max\limits_{\lambda\in\mathbb{R}^m} & -\frac{1}{2}\|\sum_{i=1}^m \lambda_i \nabla f_i(x)\|_2^2 \\ \text{s.t.} & \sum_{i=1}^m \lambda_i = 1,\ \lambda \geqq 0 \end{array} \tag{3.28}$$

The dual formulation (3.28) has been used by Desideri [19, 20, 17, 18] in a multiobjective descent algorithm. It is interesting to see that in fact Algorithm 3.2 chooses locally a steepest descent of a weighted sum of the gradients for changing weights which are not set in advance. For $m = 1$ the method is the same as the well known steepest descent method [30].

As expected from a descent algorithm the choice of initial points is very important. Since Algorithm 3.2 compute a single critical point the algorithm has to be called from different initial points to compute multiple possibly efficient points. However, it can not be guaranteed that different initial points lead to different final solutions.
Figure 3.3 shows decision space on the left and objective space on the right for a quadratic biobjective problem.

Example 3.17. Consider the following biobjective nonlinear optimization problem:

$$\begin{aligned} \text{vmin}_{x\in\mathbb{R}^2} \quad & f_1(x) = \tfrac{1}{2}(x_1^2 + x_2^2) \\ & f_2(x) = \tfrac{1}{4}(x_1 - 2)^4 + \tfrac{1}{2}(x_2 - 2)^2; \end{aligned} \tag{3.29}$$

In Figure 3.3 sequences generated by Algorithm 3.2 using steepest descent directions for different initial points on a circle around the point $(1,1)^T$ are depicted.

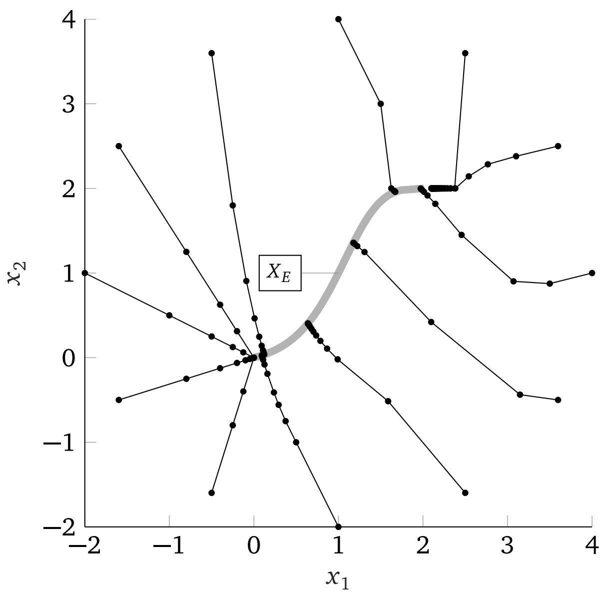

Figure 3.3: Iteration paths for Example 3.17 with multiobjective steepest descent method.

We can observe that the step sizes are small close to the efficient set which can be explained with the Armijo step size since both objectives have to be decreased simultaneously. Furthermore, we can see that the final iterates are unevenly distributed over the efficient set. In particular many final iterates lie in the minima of $f_1(x)$ or $f_2(x)$.

To improve the convergence close to the efficient set Fliege et. al. [29] proposed the multiobjective Newton method:

3.3.2 Newton Method

Let f_i be twice continuously differentiable on $\mathbb{R}^n$ for $i = 1, \dots, m$. Let $x^0 \in S$ and let each f_i be strictly convex on the level set $\left\{x \in \mathbb{R}^n \ : \ f_i(x) \leq f_i(x^0)\right\}$ for all $i = 1, \dots .m$. In this case it is possible to choose $H^i = \nabla^2 f_i(x)$, since all Hessian matrices are positive definite and the assumptions of Proposition 3.15 are satisfied.
The search direction problem for the Newton method [29] is then:

$$\begin{aligned} \min_{d\in\mathbb{R}^n,\tau\in\mathbb{R}} \quad & \tau \\ \text{s.t.} \quad & \nabla f_i(x)^T d + \tfrac{1}{2} d^T \nabla^2 f_i(x) d \quad \leq \quad \tau \qquad \forall i = 1, \dots, m \end{aligned} \tag{3.30}$$

In comparison to the steepest descent method we have to solve a nonlinear problem with quadratic constraints, which is more time expensive than solving the quadratic problem with linear constraints in the case of the steepest descent method. Notice that the matrices $H^i = \nabla^2 f_i(x)$ depend now on the current iterate $x \in \mathbb{R}^n$ for every $i = 1, \dots, m$.

The multiobjective Newton method can also be applied to nonconvex problems by first checking whether all Hessian matrices $\nabla^2 f_i(x)$ are positive definite for all $i = 1, \dots, m$ at the current iterate $x \in \mathbb{R}^n$. If not the direction of steepest descent as defined in Section 3.3.1 can be chosen, similar to the globalized Newton method for singleobjective optimization [36].

Algorithm 3.3: Globalized Multiobjective Newton Descent Method

Choose a point $x^0 \in \mathbb{R}^n$ and $\sigma \in (0, 1)$. Set $k = 0$
while *Stopping criterion is violated at* x^k **do**
  if $\nabla^2 f_i(x^k)$ *is positive definite for all* $i = 1, \dots, m$ **then**
    Compute a descent direction $d^k \in \mathbb{R}^n$ as the solution of (3.30)
  else
    Compute a steepest descent direction $d^k \in \mathbb{R}^n$ as the solution of (3.6)
  Compute a step size t_k satisfying the Armijo rule using Algorithm 3.1
  Iterate: $x^{k+1} = x^k + t_k d^k$
  $k = k + 1$

Under appropriate assumptions (for example Lipschitz-continuity of the Hessian matrices $\nabla^2 f_i(x)$) and similar to the scalar case superlinear and quadratic convergence of the sequence computed by Algorithm 3.2 was shown by Fliege et. al. [29].

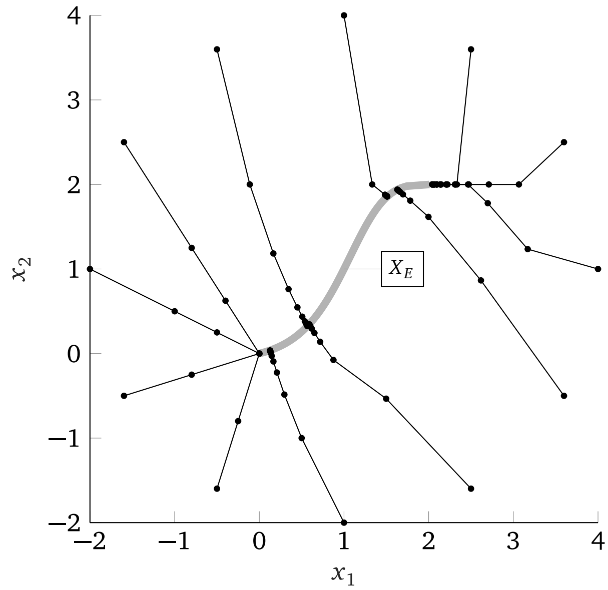

Figure 3.4: Iteration paths for Example 3.17 with multiobjective Newton method.

One numerical example that compares the convergence rate of the multiobjective steepest descent method (Algorithm 3.2) and multiobjective Newton method (Algorithm 3.3) can be found in Example 3.23.

Figure 3.4 shows iteration sequences of Algorithm 3.3 for Example 3.17. We can indeed observe that the iteration paths have slightly less iterations as was the case for the multiobjective steepest descent method (cp. Figure 3.3).
Computing the Hessian matrices may be too time-expensive in practice, for this reason we will also consider approximations of the Hessian matrices as a choice for H^i for $i = 1, \dots, m$.

3.3.3 Quasi-Newton Methods

Quasi-Newton methods for singleobjective optimization problems generate a sequence of positive definite matrices H_k, updating the matrices after every iteration of the descent method. In the multiobjective case we update m such matrices H_k^i for $i = 1, \dots, m$. An application of Quasi-Newton methods to the multiobjective case has been proposed by Povalej [70] building on the multiobjective Newton method by Fliege et. al. [29] where the Hessian matrices are updated separately for each objective function $f_i(x)$, $i = 1, \dots, m$. Let $(H_k^i)_k$ be a sequence of positive definite matrices for $i = 1, \dots, m$. The **Quasi-Newton equation** or secant equation [36] for one fixed $i \in \{1, \dots, m\}$ is given by

$$H_{k+1}^i(x^{k+1} - x^k) = \nabla f_i(x^{k+1}) - \nabla f_i(x^k).$$

For this section we will use the following abbreviations also found in Geiger and Kanzow [36]:

$$H^i = H_k^i,\ H_+^i = H_{k+1}^i,\ s = x^{k+1} - x^k,\ y^i = \nabla f_i(x^{k+1}) - \nabla f_i(x^k)\ \forall i = 1, \dots, m$$

With this notation the Quasi-Newton equation is $H_+^i s = y^i$.

Proposition 3.18. [36] Let $s, y^i \in \mathbb{R}^n$, $i = 1, \dots, m$, with $s \neq 0$. There exist positive definite matrices Q^i satisfying $Q^i s = y^i$ if and only if $s^T y^i > 0$ for all $i = 1, \dots, m$.

Proof. This is a direct extension to the multiobjective case of Lemma 11.5 in [36]. □

Hence, if $s^T y^i > 0$ is violated for some $i \in \{1, \dots, m\}$ the Quasi-Newton equation cannot be solved by positive definite matrices. The condition $s^T y^i > 0$ is referred to as the **curvature condition** [70].
There are several Quasi-Newton strategies for choosing H_+. One of the most successful ones is the BFGS (or Broydon-Fletcher-Goldfarb-Shanno) formula [36]:

$$H_+^i = H^i + \frac{y^i (y^i)^T}{s^T y^i} - \frac{H^i s s^T H^i}{s^T s}\ \forall i = 1, \dots, m \tag{3.31}$$

The curvature condition is always satisfied for strictly convex functions. For more general cases the method can be globalized by checking the curvature condition in each iteration. If the condition is violated, let's say in iteration k, it is not clear how to choose the next set of matrices. In the scalar case the steepest descent method is usually used when the curvature condition is not met. In accordance to this we will set all matrices H_{k+1}^i to I_n for $i = 1, \dots, m$. It should be noted that other choices are possible, see for example [36].
Povalej [70] showed that the assumptions of Proposition 3.15 are satisfied if the objective functions are strictly convex and twice continuously differentiable.
The resulting global Quasi-Newton method has some advantages over the multiobjective Newton method since the possibly time-consuming computation of the Hessian matrices can be avoided.
One numerical example that compares the convergence rate of the multiobjective Newton method (Algorithm 3.3) and the multiobjective Quasi-Newton method (Algorithm 3.4) can be found in Example 3.23.
Iteration paths for Example 3.17 are depicted in Figure 3.5. The iteration paths are similar to those of the multiobjective Newton methods (see Figure 3.4).

Algorithm 3.4: Globalized Multiobjective Quasi-Newton Descent Method

Choose a point $x^0 \in \mathbb{R}^n$ and $\sigma \in (0,1)$. Set $k=0$ and $H_0^i = I_n,\ \forall i = 1,\dots,m$.
while *Stopping criterion is violated at* x^k **do**
  Compute a descent direction $d^k \in \mathbb{R}^n$ as the solution of (3.6)
  Compute a step size t_k satisfying the Armijo rule using Algorithm 3.1
  Iterate: $x^{k+1} = x^k + t_k d^k$
  if *the curvature condition is satisfied* **then**
    Compute $H_{k+1}^i\ \forall i = 1,\dots,m$ according to the BFGS formula (3.31)
  else
    Set $H_{k+1}^i = I_n\ \forall i = 1,\dots,m$
  $k = k+1$

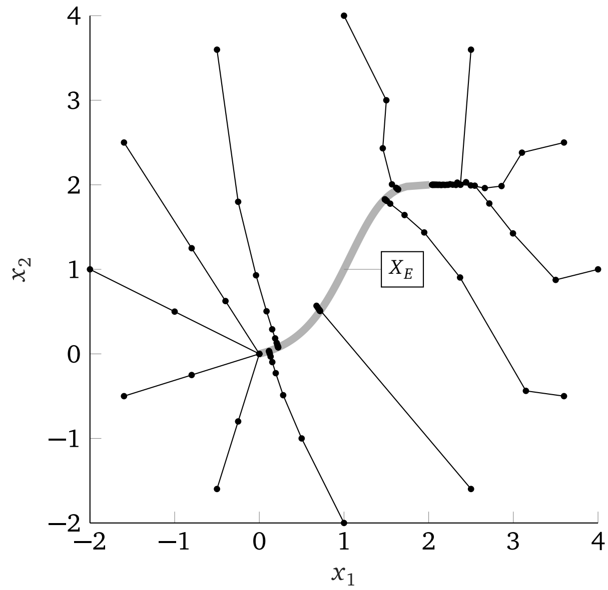

Figure 3.5: Iteration paths for Example 3.17 with multiobjective BFGS method.

3.3.4 Weight-based Descent Methods

In Proposition 3.12 it was shown that there exists $\lambda \in \Lambda$ such that the optimal solution of (3.6) is given by [30, 29, 70]

$$d = -\left[\sum_{i=1}^{m} \lambda_i H^i\right]^{-1} \left(\sum_{i=1}^{m} \lambda_i \nabla f_i(x)\right). \tag{3.32}$$

In the case of steepest descent (i.e. $H^i = I_n,\ i = 1,\dots,n$) we get

$$d_\Lambda(\lambda) = -\sum_{i=1}^{m} \lambda_i \nabla f_i(x). \tag{3.33}$$

Of course, $d_\Lambda(\lambda)$ is, in general, not a descent direction for every $\lambda \in \Lambda$. To avoid solving the quadratic optimization problem (3.6) to compute a search direction one might also consider choosing a weight $\lambda \in \Lambda$ and use $d_\lambda(\lambda)$ to compute a direction.

Definition 3.19. Let H^i be positive definite matrices. $\lambda \in \Lambda$ is called a **feasible weight**, if $d_\lambda(\lambda)$ is a descent direction of (NLP). The set

$$\Lambda_d(x) = \left\{\lambda \in \Lambda \ : \ \nabla f_i(x)^T d_\Lambda(\lambda) < 0 \ \forall \ i = 1, \ldots, m\right\}$$

is called the **set of feasible weights**.

The set of feasible weights are illustrated in Figure 3.6 for Example 3.17. For a set of points $X \subset \mathbb{R}^n$ on a grid the interval

$$\{\lambda_1 \ : \ (\lambda_1, 1-\lambda_1) \in \Lambda_d(x)\}$$

is depicted in black. We can observe that the set of feasible weights gets smaller the closer the iterate is to a critical point.

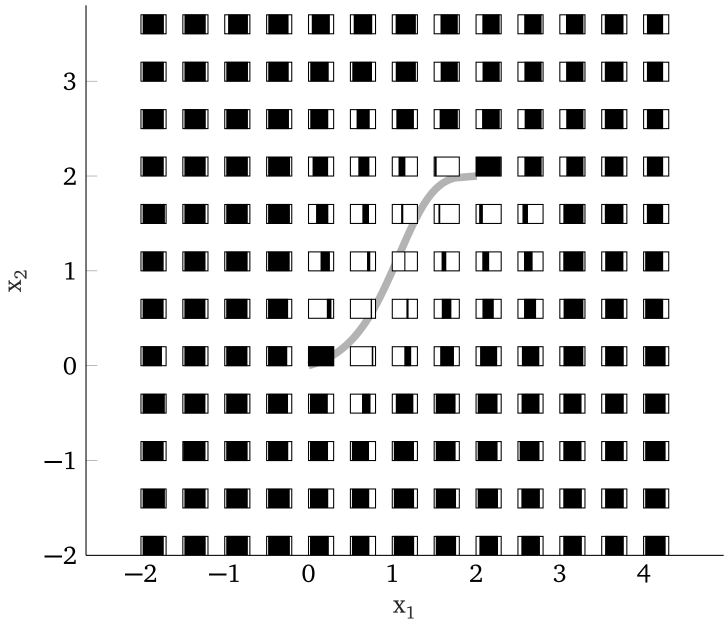

Figure 3.6: Feasible weights for points of Example 3.17

Proposition 3.20. The set of feasible weights $\Lambda_d(x)$ is a convex set

Proof. Note that

$$\begin{aligned}\Lambda_d(x) &= \left\{\lambda \in \Lambda \ : \ \nabla f_i(x)^T d_\Lambda(\lambda) < 0 \ \forall \ i = 1, \ldots, m\right\} \\ &= \left\{\lambda \in \Lambda \ : \ -\nabla f_i(x)^T \sum_{j=1}^m \lambda_j \nabla f_j(x) < 0 \ \forall \ i = 1, \ldots, m\right\} \\ &= \left\{\lambda \in \Lambda \ : \ \sum_{j=1}^m \lambda_j \nabla f_i(x)^T \nabla f_j(x) > 0 \ \forall \ i = 1, \ldots, m\right\}\end{aligned}$$

Let $\lambda^1,\ \lambda^2 \in \Lambda_d(x)$ be two feasible weights and $\alpha \in [0,1]$. Then

$$\begin{aligned}
\sum_{j=1}^{m} \lambda_j \nabla f_i(x)^T \nabla f_j(x) &= \sum_{j=1}^{m} \left(\alpha\lambda_j^1 + (1-\alpha)\lambda_j^2\right) \nabla f_i(x)^T \nabla f_j(x) \\
&= \alpha \underbrace{\sum_{j=1}^{m} \lambda_j^1 \nabla f_i(x)^T \nabla f_j(x)}_{>0} + (1-\alpha) \underbrace{\sum_{j=1}^{m} \lambda_j^2 \nabla f_i(x)^T \nabla f_j(x)}_{>0} \\
&> 0
\end{aligned}$$

Which shows that $(\alpha\lambda^1 + (1-\alpha)\lambda^2) \in \Lambda_d(x)$. □

3.3.4.1 Compromise Descent

Consider normalizing the gradients of the objective functions:

Definition 3.21. Let x be a point where $\|\nabla f_i(x)\| > \varepsilon$ for some $\varepsilon > 0$ for all $i = 1, \dots, m$. Then the **compromise direction** d^c is defined as:

$$d^c(x) = -\frac{1}{m} \sum_{i=1}^{m} \frac{\nabla f_i(x)}{\|\nabla f_i(x)\|} \tag{3.34}$$

Remark 3.22. 1. d^c is not necessarily a descent direction in x. It is easy to find examples where $n < m$. For example

$$Jf(x) = \begin{pmatrix} -2 & 4 \\ -1 & -3 \\ 1 & -6 \end{pmatrix},\ d^c(x) = \begin{pmatrix} 0.5990 \\ 1.0407 \end{pmatrix},\ Jf(x)d^c(x) = \begin{pmatrix} 2.9645 \\ -3.7210 \\ -5.6449 \end{pmatrix}$$

But note that there does exists a descent direction:

$$Jf(x)\begin{pmatrix} 0.7339 \\ 0.2202 \end{pmatrix} = \begin{pmatrix} -0.5872 \\ -1.3945 \\ -0.5872 \end{pmatrix}.$$

2. For $m = 2$ we can prove, that at least $\nabla f_i(x)^T d^c \le 0$ holds for $i = 1, 2$:

Using Cauchy-Schwarz we get:

$$\begin{aligned}
|\langle \nabla f_1(x), \nabla f_2(x) \rangle|^2 &\le \langle \nabla f_1(x), \nabla f_1(x) \rangle \langle \nabla f_2(x), \nabla f_2(x) \rangle \\
&\le \|\nabla f_1(x)\|^2 \cdot \|\nabla f_2(x)\|^2 \\
&\le 1
\end{aligned}$$

and

$$|\langle \nabla f_1(x), \nabla f_2(x) \rangle|^2 \le 1 \Leftrightarrow -1 \le \langle \nabla f_1(x), \nabla f_2(x) \rangle \le 1$$

$$\begin{aligned}
\langle \nabla f_1(x), -\nabla f_1(x) - \nabla f_2(x) \rangle &= -\underbrace{\langle \nabla f_1(x), \nabla f_1(x) \rangle}_{=\|\nabla f_1(x)\|^2 = 1} - \underbrace{\langle \nabla f_1(x), \nabla f_2(x) \rangle}_{\ge -1} \\
&\le 0
\end{aligned}$$

Similar equations hold true for $\nabla f_2(x)$.

3. The compromise descent direction can be obtained by a weighted sum of the gradients with weights

$$\lambda_j = \left(\sum_{i=1}^{m} \frac{1}{\|\nabla f_i(x)\|} \right)^{-1} \frac{1}{\|\nabla f_j(x)\|}$$

To use the compromise direction, we first check if any $\nabla f_i(x)$ vanishes. If yes, then x is critical. Otherwise, we compute the compromise direction and check weather it is a descent direction. If $d^c(x)$ is a descent direction then we multiply $d^c(x)$ by $\max_{i=1,\dots,m} \|\nabla f_i(x)\|$ since $d^c(x)$ is normalized and using $d^c(x)$ would result in small steps. If $d^c(x)$ is not a descent direction, we use another search direction problem, for example steepest descent direction given by (3.27).
The advantage of this procedure is that no optimization problem has to be solved, if the compromise direction is a descent direction.

Algorithm 3.5: Multiobjective Compromise Descent

Choose a point $x^0 \in \mathbb{R}^n$ and $\sigma \in (0, 1)$. Set $k = 0$
while *Stopping criterion is violated at* x^k **do**
  Compute the compromise descent direction $d^c(x^k)$ via (3.34)
  if $d^c(x^k)$ *is descent direction in* x^k **then**
    Set $d^k = d^c(x^k) \cdot \max_{i=1,\dots,m} \left\|\nabla f_i(x^k)\right\|$
  else
    Compute a steepest descent direction $d^k \in \mathbb{R}^n$ as the solution of (3.6)
  Compute a step size t_k satisfying the Armijo rule using Algorithm 3.1
  Iterate: $x^{k+1} = x^k + t_k d^k$
  $k = k + 1$

The iteration paths of the compromise descent method for Example 3.17 are shown in Figure 3.7. We can observe that the normalization of the gradients leads to a more even distribution of the final solutions.

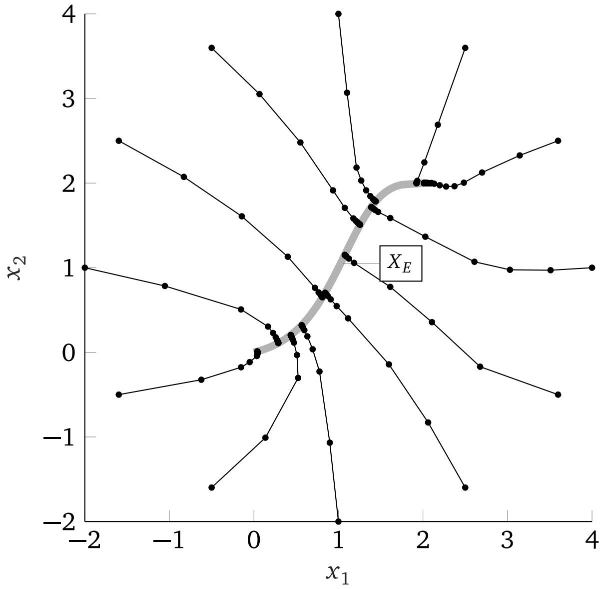

Figure 3.7: Iteration paths for Example 3.17 with compromise descent method.

3.4 Numerical Experiment

To visualize differences of the methods discussed in this chapter we consider the following triobjective convex optimization problem:

Example 3.23. Fliege et. al. [29] suggested the following triobjective convex problem to illustrate the convergence properties of the multiobjective descent methods reviewed in this chapter:

$$f_1(x) = \frac{1}{n^2}\sum_{j=1}^{n} j(x_j - j)^4$$
$$f_2(x) = \exp\left(\sum_{j=1}^{n}\frac{x_j}{n}\right) + \sum_{j=1}^{n}(x_j)^2$$
$$f_3(x) = \frac{1}{n(n+1)}\sum_{j=1}^{n} j(n-j+1)\exp(-x_j)$$

This nonlinear triobjective optimization problem can be scaled with the number of variables. For this numerical experiment we consider $n = 100$ variables.
The algorithm stopped when the norm of the descent direction d satisfies $\|d\| \leqq 0.001$. The parameters σ and β for the Armijo step size were each set to 0.5. A set of 200 random points in the box spanned by $(-2,\ldots,-2)$ and $(2,\ldots,2)$ was randomly chosen. All methods were tested starting from this set of initial points. The Algorithms were implemented in Matlab 2018a using the included solvers fmincon and quadprog to solve the search direction problems. The computations were executed on an Intel Core i5 processor with 3.2 GHz and 8 GB of ram. Table 3.1 shows the maximum and average number of iterations (which also coincides with the number of computations of the Hessian matrix for the newton method), the average time for reaching the final iterate and the average number of function evaluations (where one evaluation is an evaluation of the m-dimensional function $f(x)$). We observe that the bulk of the function evaluations occurs during the computation of the Armijo step sizes.
The average runtimes depend highly on the performance of the solvers used and should only be understood in this context. We can observe a distinct reduction in the number of iterations when using Newton and BFGS instead of the steepest descent method. This is expected and in accordance with results from scalar nonlinear optimization [36].
The search direction problems for the Newton and BFGS variants have to be solved with fmincon, which seems to be slower than quadprog in the case of the steepest descent method.
The BFGS method performs similar to the Newton method. Compromise descent showed the lowest average times as in most iterations no search direction problem has to be solved.

Method	Max It.	Avg. It.	Avg. time	Avg. feval
Steepest Descent	121	89.91	0.15133	351.54
Newton	15	10.09	0.18721	48.62
BFGS	16	15.135	0.1380	59.53
Compromise Descent	28	22.31	0.0047889	140.91

Table 3.1: Numerical results for Example 3.23

It is difficult to assess the performance of these algorithms in general. This requires extensive numerical testing on different types of problems. For instance, Huber et. al. [47] provide a review of scalable test problems.

3.4.1 Comparison of Multiobjective Descent Method and Weighted Sum

An alternative approach for solving nonlinear multiobjective problems is to apply a scalarization method and solve the scalarized problem for predetermined parameters. As an example, we will discuss the weighted sum method in combination with Algorithm 3.2 for the special case $m = 1$.

Choose a fixed $\lambda \geq 0$ and use a scalar descent method to minimize $\lambda^T f(x)$ [23].

Algorithm 3.6: Weighted Sum Descent Method

Choose a point $x^0 \in \mathbb{R}^n$ and $\sigma \in (0,1)$. Set $k = 0$.
while *Stopping criterion is violated at* x^k **do**
  Choose a descent direction $d^k \in \mathbb{R}^n$ of $\lambda^T f(x)$
  Compute a step size t_k satisfying the Armijo rule for $\lambda^T f(x)$ using Algorithm 3.1
  Iterate: $x^{k+1} = x^k + t_k d^k$
  $k = k + 1$

Example 3.24. For the illustration of some properties of the weighted sum descent method and the multiobjective steepest descent method consider the following nonconvex nonlinear biobjective problem:

$$\begin{array}{llll} \operatorname{vmin}\limits_{x \in \mathbb{R}^2} & f_1(x) & = & \cos(\pi x_1) + \cos(\pi x_2) \\ & f_2(x) & = & \frac{1}{2}(x_1 - 2)^2 + \frac{1}{2}(x_2 - 2)^2 \end{array} \tag{3.35}$$

The second objective, f_2, is strictly convex, but the first objective is periodic and has global minima at $(2a_1+1, 2a_2+1)^T$, global maxima at $(2a_1, 2a_2)^T$ and saddle points at $(2a_1, 2a_2+1)^T$ and $(2a_1+1, 2a_2)^T$ for every pair of integers a_1 and a_2. f_2 has a global minimum at $(2,2)^T$.
Every stationary point of f_1, i.e. every point with integer coordinates, is also stationary for (3.35) (the multiobjective problem) since $\nabla f_1(x) = 0$ is a sufficient condition for critical points of f. The efficient set of (3.35) is given by

$$\operatorname{conv}(\{(1,1)^{\mathrm{T}}, (3,3)^{\mathrm{T}}\}) \cup \operatorname{conv}(\{(1,3)^{\mathrm{T}}, (3,1)^{\mathrm{T}}\}).$$

To see this, consider that the minimum of $f_2(x)$ is the point $(2,2)^T$ and the closest global minima of $f_1(x)$ are the points $(1,1)^T$, $(-1,1)^T$, $(-1,-1)^T$ and $(1,-1)^T$.
Due to the periodic nature of f_1, (3.35) has many locally efficient sets. But not all stationary points of (3.35) are locally weakly efficient. For example, all saddle points of f_1 are not locally weakly efficient for (3.35).
To compare multiobjective descent methods and weighted sum descent methods we consider here the weighted sum for (the arbitrary choice) $\lambda = (0.5, 0.5)^T$ and the steepest descent variant in both cases. Figure 3.8 shows iteration paths, starting from initial points on the boundary of the box with corner points $(-4.5, -4.5)$ and $(6.5, 6.5)$ of the multiobjective descent method (left) and the weighted sum descent (right). The level lines are those of the weighted sum function $0.5 f_1(x) + 0.5 f_2(x)$.
We can observe that the multiobjective descent method gets stuck in stationary points (which are often saddle points of f_1), whereas the weighted sum descent method can sometimes, of course not in general, still converge to an efficient point. This behavior of multiobjective descent methods is a major drawback.

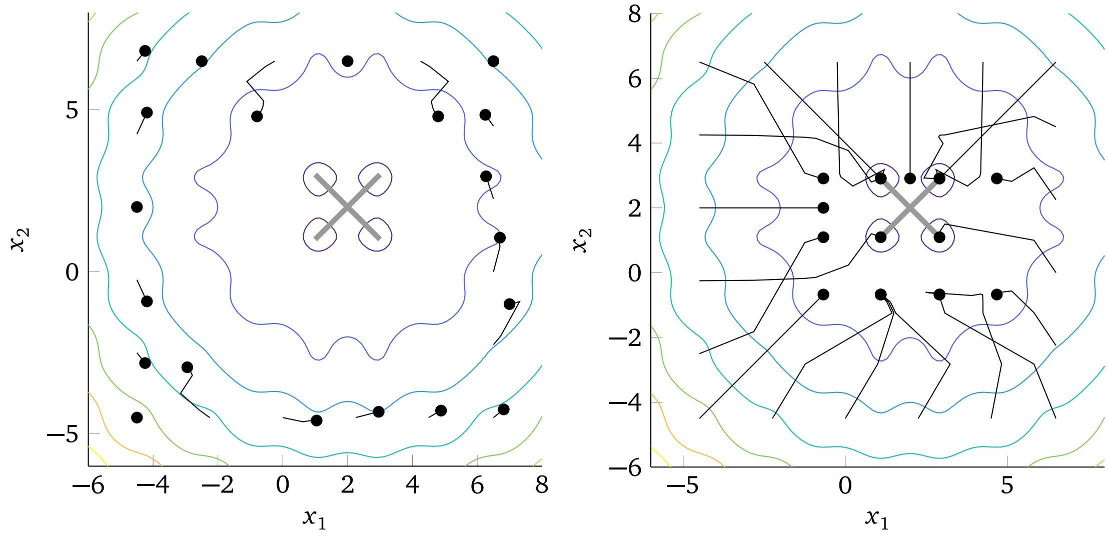

Figure 3.8: Iteration paths and final points multiobjective steepest descent (left) and weighted sum descent (right) of Example 3.24.

We demonstrated using Example 3.24 that the multiobjective steepest descent algorithm has some additional challenges regarding critical points which are not locally efficient. In the following we will discuss the multiobjective Newton method in this context.

Some of the properties of descent methods in scalar nonlinear optimization can be generalized to the multiobjective case in a way that is useful for multiobjective descent methods. One crucial property for some methods, for example the globalized Newton method, however, do not remain true in the multiobjective case. In the scalar case, we have the following:

Lemma 3.25. [36] Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable and $\bar{x} \in \mathbb{R}^n$ such that $\nabla^2 f(\bar{x})$ is positive definite. Then there exist constants $\delta > 0$ and $\alpha > 0$ such that

$$\alpha \|d\|^2 \leq d^T \nabla^2 f(x) d$$

for every $x \in \mathbb{R}^n$ with $\|x - \bar{x}\| \leq \delta$ and all $d \in \mathbb{R}^n$.

Lemma 3.25 shows that in a sufficiently small neighborhood of $\bar{x}$, the Hessian matrix of $f(x)$ is positive definite. This result is crucial for the convergence results for the globalized Newton methods as it guarantees that the Newton direction is well-defined close to a local minimum and thus the globalized Newton method will not use steepest descent directions close to the minimum and thus not have zigzagging behavior [36].

In the multiobjective case Lemma 3.25 does not apply in general. It does however apply to the weighted sum problem.

Example 3.26. Consider the following biobjective optimization problem:

$$\operatorname*{vmin}_{x\in\mathbb{R}^n} \; f_1(x) = x^3 - 3x^2 \qquad f_2(x) = x^2 \tag{3.36}$$

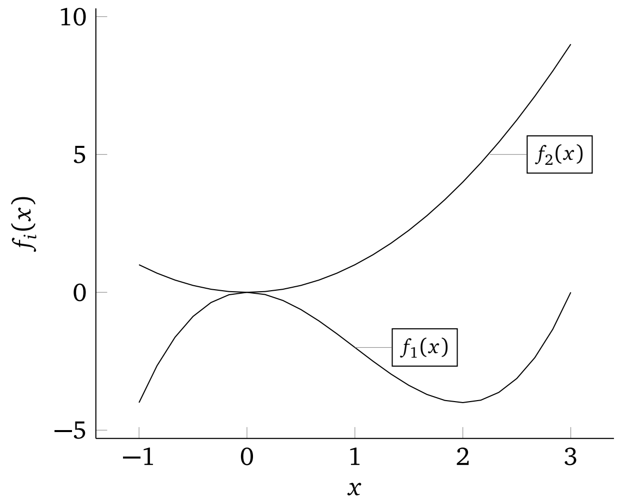

Figure 3.9: Function Graphs for Example 3.26.

The function graphs of $f_1(x)$ and $f_2(x)$ are shown in Figure 3.9. Notice that every $x \in (0,2)$ is locally efficient for (3.36). The second derivative of $f_1(x)$ is given by $6x - 6$ which is strictly positive only on the interval $(1,\infty)$. Thus for every $x \in (0,1)$ the second derivative of f_1 is negative and the multiobjective newton method can not use the newton direction.

Conclusion

We have reviewed a class of descent methods for multiobjective nonlinear optimization problems that generalizes methods such as the steepest descent method and the Newton method from the singleobjective to the multiobjective case. We have seen that these variants share some of the properties with their singleobjective counterparts, such as general convergence results in the convex case.
We introduced the compromise descent method and compared the methods using a larger dimensional triobjective convex optimization problem which visualizes the theoretical results from the literature [29] . We compared the multiobjective steepest descent method to the application of the singleobjective steepest descent method to the weighted sum problem for one nonconvex problem and observed that multiobjective descent methods have some additional challenges with critical solutions that are not locally efficient.

The advantage of multiobjective descent methods is that no preference information, like a weight for the weighted sum problem, has to be provided. However, we have also observed that the multiobjective descent methods converge favorably to particular critical points depending on the initial solution. The choice of the initial solution for multiobjective descent methods is an interesting topic for future research. Additionally, we have only discussed one weight-based descent methods. Since the set of feasible weights is convex one might formulate an algorithm to find feasible weights in each iteration in an efficient way, leading to a new variant of the multiobjective descent methods.

Chapter 4

Multiobjective Convex Quadratic Programming

Multiobjective convex quadratic optimization problems with linear constraints have been discussed in the literature for quite some time. Actually, some solution techniques for quadratic programming with linear constraints [60] were introduced alongside biobjective quadratic optimization in the context of portfolio optimization [59] in the 1950s. Other applications for multiobjective quadratic programming have arisen, for example in the fields of location analysis and radio therapy treatment planning [9].

The goal of many methods in portfolio optimization is the determination of efficient solutions [59] and the corresponding weights representing these solutions. This can, for example, be achieved by a parametric approach using the weighted sum scalarization, where the weights are incrementally changed to explore the whole efficient set. This approach for biobjective problems was extended by Goh and Yang [38] to the more general multiobjective case.
Using the KKT conditions of the weighted sum scalarization, multiobjective quadratic programming problems can be interpreted as parametric linear complementarity problems for which the weights are a parameter. Many solution techniques have been discussed for such problems where the parameter is only found on the right-hand-side of a system of linear equations [2, 13, 44, 46].
Adelgren [1] considered multiparametric linear complementarity problems with parameters in more positions and his results can also be applied to multiobjective quadratic programming problems.

In this chapter we focus on multiobjective convex quadratic programming problems with linear constraints and strictly convex objective functions.

$$\begin{array}{ll} \text{vmin} & f_i(x) = \frac{1}{2}x^T Q^i x + (c^i)^T x \quad , \ i = 1, \dots, m \\ \text{s.t.} & Ax \geqq b, \ Hx = h \end{array} \tag{4.1}$$

with $Q^i \in \mathbb{R}^{n\times n}$ symmetric positive definite **objective matrices** and vectors $c^i \in \mathbb{R}^n$ for all $i = 1, \dots, m$. The linear constraints are given by matrices $A \in \mathbb{R}^{p\times n}$ and $H \in \mathbb{R}^{q\times n}$ and vectors $b \in \mathbb{R}^p$ and and $h \in \mathbb{R}^q$. We assume that the matrix H has full rank.
The feasible set is denoted by $S = \{x \in \mathbb{R}^n \ : \ Ax \geqq b, \ Hx = h\}$ and throughout this chapter we assume that S is not empty.
As the weighted sum will be used throughout this chapter we will use the following notation for $\lambda \in \Lambda$:

$$Q(\lambda) := \sum_{i=1}^{m} \lambda_i Q^i \text{ and } c(\lambda) = \sum_{i=1}^{m} \lambda_i c^i$$

Notice that $Q(\lambda)$ is a symmetric and positive definite matrix for every $\lambda \geq 0$.

This chapter consists of the following sections:

- In Section 4.1 we consider unconstrained and equality-constrained multiobjective convex quadratic optimization problems. An optimality condition derived for this case will play an important role in the following sections.

- In Section 4.2 two solution approaches by Goh and Yang [38] and Adelgren [1] for multiobjective convex quadratic optimization problems in canonical form are reviewed. We introduce the weight space decomposition by efficient complementary bases and efficient active sets.

- In Section 4.3 an algorithm is proposed for the computation of all efficient complementary bases and a set of test instances is solved. The two-phase approach proposed by Adelgren [1] is reviewed and compared to the algorithm discussed in Section 4.3.

- In Section 4.4 the results from Section 4.2 are generalized for multiobjective convex optimization problems in general form with unbounded variables and linear equality constraints.

- In Section 4.5 we shortly consider the regularization of positive semidefinite matrices and provide a justification for considering only multiobjective strictly convex quadratic optimization problems.

- In Sections 4.6, 4.7 and 4.8 particularly structured multiobjective convex quadratic optimization problem are considered that have a polyhedral weight space decomposition. The first type of problem (discussed in Section 4.6) consists of one convex quadratic and $m-1$ linear objectives. This case has been considered by Hirschberger et. al. [45, 46].
  In Section 4.7 a very similar case to the previous one is discussed where the objective matrices are positive multiples of each other.
  In Section 4.8 we consider multiobjective convex optimization problems with diagonal objective matrices and lower and upper bounds. For this case we show that the weight space is an arrangement of hyperplanes and provide an upper bound for the number of efficient active sets.

- In Section 4.9 a parameter space decomposition for the e-constraint problem is introduced.

- In Section 4.10 we discuss an application of the results from Section 4.7 and Section 4.9 for a problem in the field of location analysis.

4.1 The Efficient Set of Unconstrained and Equality-constrained Multiobjective Convex Quadratic Programming Problems

First, we consider multiobjective unconstrained nonlinear optimization problems with strictly convex quadratic objective functions:

$$\operatorname*{vmin}_{x\in\mathbb{R}^n} \quad f_i(x) = \tfrac{1}{2}x^T Q^i x + (c^i)^T x \quad , \ i = 1,\ldots,m \tag{4.2}$$

with symmetric and positive definite objective matrices Q^i and linear cost vectors $c^i \in \mathbb{R}^n$ for $i = 1,\ldots,m$.
The weighted sum scalarization problem of (4.2) can be written as

$$\min_{x\in\mathbb{R}^n} \ Q(\lambda)x + c(\lambda). \tag{4.3}$$

Corollary 4.1. [23] $x \in \mathbb{R}^n$ is efficient for (4.2) if and only if there exists $\lambda \in \Lambda$ such that

$$x = -Q(\lambda)^{-1}c(\lambda).$$

Proof. All objective functions are stricly convex and thus the KTT conditions are necessary and sufficient for weak efficiency (see Theorem 2.4 and Theorem 2.5). $Q(\lambda)x + c(\lambda)$ is a strictly convex function for every $\lambda \geq 0$ and thus has a unique global minimum for every $\lambda \geq 0$. The KKT conditions for (4.2) with weights $\lambda \geq 0$ are given by

$$\sum_{i=1}^{m} \lambda_i \nabla f_i(x) = 0$$

which is equivalent to

$$\sum_{i=1}^{m} \lambda_i (Q^i x + c^i) = 0 \Leftrightarrow Q(\lambda)x + c(\lambda) = 0 \Leftrightarrow Q(\lambda)x = -c(\lambda) \Leftrightarrow x = -Q(\lambda)^{-1}c(\lambda).$$

□

A parametric representation of the efficient set X_E can be defined using the optimal solution $\bar{x}(\lambda)$ of the weighted sum problem (4.3):

$$\bar{x}(\lambda) := -Q(\lambda)^{-1}c(\lambda)$$

in the following way:

$$X_E = \{\bar{x}(\lambda) \ : \ \lambda \in \Lambda\}$$

4.1.1 Linear Equality Constraints

In this section we generalize the properties of unconstrained multiobjective quadratic programming problems to multiobjective quadratic programming problems with q linear equality constraints in the form:

$$\begin{aligned} \operatorname*{vmin}_{x\in\mathbb{R}^n} \quad & \tfrac{1}{2}x^T Q^i x + (c^i)^T x \quad i = 1,\ldots,m \\ \text{s.t.} \quad & Hx = h \end{aligned} \tag{4.4}$$

with a full-rank matrix $H \in \mathbb{R}^{q\times n}$, $h \in \mathbb{R}^q$ and the corresponding weighted sum problem with weights $\lambda \in \Lambda$

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & \tfrac{1}{2}x^T Q(\lambda)x + c(\lambda)^T x \\ \text{s.t.} \quad & Hx = h \end{aligned} \tag{4.5}$$

From Theorem 2.4 we know that x is a KKT point of (4.5) with multipliers $\mu \in \mathbb{R}^q$ if and only if (x,μ) is a solution of

$$\begin{bmatrix} Q(\lambda) & -H^T \\ H & 0 \end{bmatrix} \begin{pmatrix} x \\ \mu \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ h \end{pmatrix}. \tag{4.6}$$

Proposition 2.12 shows that if H has full rank then there exists a unique solution of (4.6) for every $\lambda \in \Lambda$.
Notice that for a fixed λ the equivalent system of equations (4.6) can be interpreted as the necessary optimality condition for a $n+q$-dimensional singleobjective unconstrained quadratic programming problem

$$\min_{z \in \mathbb{R}^{n+q}} \frac{1}{2} z^T K(\lambda) z + \begin{pmatrix} -c(\lambda) \\ h \end{pmatrix}^T z \tag{4.7}$$

with $K(\lambda) = \begin{bmatrix} Q(\lambda) & -H^T \\ H & 0 \end{bmatrix}$ and $z = (x,\mu)^T$.
Using normalized weights $\lambda \in \Lambda$ we see that

$$\sum_{i=1}^m \lambda_i \begin{bmatrix} Q^i & -H^T \\ H & 0 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m \lambda_i Q^i & -\sum_{i=1}^m \lambda_i H^T \\ \sum_{i=1}^m \lambda_i H & 0 \end{bmatrix} = \begin{bmatrix} Q(\lambda) & -H^T \\ H & 0 \end{bmatrix}$$

and

$$\sum_{i=1}^m \lambda_i \begin{pmatrix} -c^i \\ h \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ h \end{pmatrix}.$$

Hence, the optimality condition for an efficient point of (4.4) can be interpreted as the optimality conditions of the larger dimensional unconstrained problem (4.7).
The efficient set of (4.4) can be parameterized in the following way:

Corollary 4.2. [23] Let H have full rank. $x \in \mathbb{R}^n$ is efficient for (4.4) if and only if there exists $\lambda \in \Lambda$ and $\mu \in \mathbb{R}^q$ such that

$$\begin{pmatrix} x \\ \mu \end{pmatrix} = \begin{bmatrix} Q(\lambda) & -H^T \\ H & 0 \end{bmatrix}^{-1} \begin{pmatrix} -c(\lambda) \\ h \end{pmatrix}.$$

Proof. Since H has full rank we can apply Proposition 2.12 and show that

$$K(\lambda) = \begin{bmatrix} Q(\lambda) & -H^T \\ H & 0 \end{bmatrix}$$

is regular and positive definite. Thus the necessary optimality condition is satisfied at exactly one point and the inverse of $K(\lambda)$ is defined. If H has full rank then the LICQ are satisfied and we can apply Theorem 2.4. □

4.2 Multiobjective Convex Quadratic Problems in Canonical Form

In this section we will discuss the properties of multiobjective convex quadratic programming problems with linear inequality constraints and nonnegative variables. First, we will introduce a parametric approach with linear complementarity problems using the weighted sum scalarization and investigate properties of the weight space decomposition. For this, we will use results from the field of parametric linear complementarity problems [1, 13].
Afterwards, we will connect the parametric approach using linear complementarity problems with the concept of efficient active sets introduced by Goh and Yang [38].

We will now consider multiobjective problems in canonical form. Multiobjective convex quadratic optimization problems in general form are discussed in Section 4.4.

$$\begin{array}{ll} \text{vmin} & f_i(x) = \frac{1}{2}x^T Q^i x + (c^i)^T x \quad i = 1,\dots,m \\ \text{s.t.} & Ax \geqq b,\ x \geqq 0 \end{array} \tag{MQP}$$

with $Q^i \in \mathbb{R}^{n\times n}$ symmetric positive definite and $c^i \in \mathbb{R}^n$ for all $i = 1,\dots,m$, $A \in \mathbb{R}^{p\times n}$ and $b \in \mathbb{R}^p$.
Just like before, the feasible set is denoted by $S = \{x \in \mathbb{R}^n \ :\ Ax \geqq b,\ x \geqq 0\}$ and we assume that the feasible set S is not empty.
The weighted sum scalarization problem of (MQP) for $\lambda \in \Lambda$ is given by

$$\begin{array}{ll} \min & \frac{1}{2}x^T Q(\lambda) x + c(\lambda)^T x \\ \text{s.t.} & Ax \geqq b,\ x \geqq 0 \end{array} \tag{WQP}$$

The optimal solution of the weighted sum scalarization problem (WQP) for $\lambda \in \Lambda$ is denoted by $\bar{x}(\lambda)$.

Theorem 4.3 (Optimality Conditions for (MQP))**.** [23] Let $x \in S$ be a regular solution of (MQP). Then x is efficient for (MQP) if and only if there exist $\lambda \in \Lambda$, $\pi \in \mathbb{R}^p$ and $y \in \mathbb{R}^n$ such that

$$Q(\lambda)x + c(\lambda) - A^T\pi - y = 0 \tag{4.8}$$
$$Ax - b \geqq 0,\ x \geqq 0 \tag{4.9}$$
$$\pi \geqq 0,\ y \geqq 0 \tag{4.10}$$
$$x_i y_i = 0 \forall i = 1,\dots,n \tag{4.11}$$
$$(A_{j\bullet}x - b_j)\pi_j = 0 \forall j = 1,\dots,p \tag{4.12}$$

Proof. '⇒': If there exist $\bar{\lambda} \in \Lambda$, $\pi \in \mathbb{R}^p$ and $y \in \mathbb{R}^n$ such that the KKT conditions for (WQP) are satisfied then x is an optimal solution of (WQP) for $\bar{\lambda}$ (see Corollary 2.6). x is also the unique optimal solution of (WQP) for $\bar{\lambda}$ because the objective function of (WQP) is strictly convex for every $\lambda \in \Lambda$. Using Theorem 2.25 we see that x is an efficient solution of (MQP).

'⇐': If x is an efficient solution of (MQP) then according to Theorem 2.26 there exists $\bar{\lambda} \in \Lambda$ such that x is the optimal solution of (WQP) for $\bar{\lambda}$. Since x is a regular point of (WQP) x is also a KKT point of (WQP) for $\bar{\lambda}$ (see Corollary 2.6).

□

Equation (4.9) and (4.10) ensure primal and dual feasibility. The nonlinear conditions (4.11) and (4.12) are referred to as complementarity conditions. With the addition of slack variables $s \in \mathbb{R}^p$ with $s = Ax - b$ we can rewrite the KKT conditions as a parametric linear complementarity system (cp. Section 2.1.2) [14]:

$$\underbrace{\begin{bmatrix} I_n & 0 & -Q(\lambda) & A^T \\ 0 & I_p & -A & 0 \end{bmatrix}}_{=:M(\lambda)} \begin{pmatrix} y \\ s \\ x \\ \pi \end{pmatrix} = \underbrace{\begin{pmatrix} c(\lambda) \\ -b \end{pmatrix}}_{=:q(\lambda)}$$
$$s \geqq 0,\ y \geqq 0,\ \pi \geqq 0,\ x \geqq 0$$
$$s_j \pi_j = 0\ \forall j = 1, \dots, p$$
$$y_i x_i = 0\ \forall i = 1, \dots, n$$

With $u = (y, s)$, $v = (x, \pi)$ the parametric linear complementarity problem can be written as

$$\begin{aligned} M(\lambda)(u, v) &= q(\lambda) \\ u \geqq 0, v &\geqq 0 \\ u_i v_i &= 0 \forall i = 1, \dots, r \end{aligned} \tag{pLCP}$$

with $r = n + p$ and $M(\lambda) \in \mathbb{R}^{r \times 2r}$.
For a fixed weight $\lambda \in \Lambda$ we can solve this linear complementarity problem using Algorithm 2.1 and compute a feasible complementary basis B. If we change λ the basis B may stay feasible or be infeasible for the new weight. In the following we will discuss how to determine weights for which B is a feasible complementary basis.
Recall from Section 2.1.2 that, for a fixed $\lambda \in \Lambda$, the basic value $q_B(\lambda)$ for a complementary basis B of (pLCP) is given by

$$q_B(\lambda) = (M_B(\lambda))^{-1} q(\lambda).$$

Definition 4.4. A complementary basis B is called **efficient**, if there exists $\lambda \in \Lambda$ such that B is a feasible complementary basis with parameter λ, i.e. $q_B(\lambda) \geqq 0$.
The **set of efficient bases** is denoted by B_{eff}.

For every basis B there may be more than one weight where B is a feasible basis of the parametric linear complementarity problem. The set of parameters for which a given complementary basis is feasible is also referred to as invariancy region [2] or critical domains [13] in the literature.

Definition 4.5. Let B be a complementary basis of (pLCP). The set

$$\Lambda^B(B) := \{\lambda \in \Lambda \ :\ q_B(\lambda) \geqq 0\} \tag{4.13}$$

is called the **weight cell** of B.

Before we investigate the properties of the weight cells $\Lambda^B(B)$ for an efficient complementary basis B we consider an alternative approach introduced by Goh and Yang [38] using active sets in the next section.

4.2.1 Efficient Active Sets

An alternative way to describe the weight cells is an extension of the active set method for singleobjective optimization problems. We refer to the book by Geiger and Kanzow [37] for a general introduction for active set methods in singleobjective optimization.

We review the approach proposed by Goh and Yang [38] to provide an analytic description of the efficient set of (MQP). We will discuss this active set approach and show connections to the parametric linear complementarity system (pLCP) in this section.

Definition 4.6. Let two index sets $\mathcal{I} \subset \{1, \ldots, n\}$ and $\mathcal{J} \subset \{1, \ldots, p\}$ be given. Then $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ is called **active set**.

We denote the complements of $\mathcal{I}$ and $\mathcal{J}$ by $\bar{\mathcal{I}}$ and $\bar{\mathcal{J}}$, respectively. The matrix $I_{\mathcal{I}\bullet}$ denotes the matrix containing the rows of the identity matrix I_n corresponding to the entries in $\mathcal{I}$. Similarly, $(A_{\mathcal{J}\bullet})$ denotes the matrix consisting of the rows of A with index $j \in \mathcal{J}$.

Definition 4.7. An active set $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ is called **efficient active set** of (MQP), if there exists $\lambda \in \Lambda$ such that there exists a solution $(\bar{x}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}})$ of

$$\underbrace{\begin{bmatrix} Q(\lambda) & -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T \\ -I_{\mathcal{I}\bullet} & 0 & 0 \\ -A_{\mathcal{J}\bullet} & 0 & 0 \end{bmatrix}}_{=:K(\lambda)} \begin{pmatrix} x \\ y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \\ -b_{\mathcal{J}} \end{pmatrix} \tag{4.14}$$

that satisfies $\bar{x} \in S$ and $(\bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}) \geqq 0$.
The systems of equations (4.14) is called the **active set system**.
The set of efficient active sets of (MQP) is denoted by $\mathscr{A}_{\text{eff}}$.

Definition 4.8. An active set of (MQP) is called **regular** if the system of equations (4.14) has a unique solution for all $\lambda \in \Lambda$.

Notice that the active set system (4.14) is similar to the optimality condition discussed in Corollary 4.2 for a multiobjective quadratic optimization problem with the linear equality constraints $x_{\mathcal{I}} = 0$ and $A_{\mathcal{J}\bullet}x = b_{\mathcal{J}}$.

Proposition 4.9. [38] Let $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ be an efficient regular active set. Let $(\bar{x}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}})$ be the unique solution of (4.14) with $\bar{x} \in S$, $\bar{y}_{\mathcal{I}} \geqq 0$ and $\bar{\pi}_{\mathcal{J}} \geqq 0$. Then $\bar{x}$ is an efficient solution of (MQP).

Proof. Since $\mathscr{A}$ is regular we know that the LICQ are satisfied at $\bar{x}$. Setting $\bar{y}_{\bar{\mathcal{I}}} = 0$ and $\bar{\pi}_{\bar{\mathcal{J}}} = 0$ we see that $(\bar{x}, \bar{y}, \bar{\pi})$ satisfies the conditions of Theorem 4.3. Thus, $\bar{x}$ is efficient. □

Now we will discuss under which conditions the active set system (4.14) has a unique solution.

Proposition 4.10. Let $(\bar{x}, y_{\mathcal{I}}, \pi_{\mathcal{J}})$ be a solution of (4.14) for a given active set $\mathscr{A} = (\mathcal{I}, \mathcal{J})$. If the LICQ holds at $\bar{x}$ then $\mathscr{A}$ is regular.

Proof. The matrix $Q(\lambda)$ is positive definite for all $\lambda \in \Lambda$ and the matrix

$$[I_{\mathcal{I}\bullet}, A_{\mathcal{J}\bullet}]$$

has full rank, since the LICQ are satisfied for $\bar{x}$. Using Proposition 2.12 we see that the matrix

$$K(\lambda) = \begin{bmatrix} Q(\lambda) & -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T \\ -I_{\mathcal{I}\bullet} & 0 & 0 \\ -A_{\mathcal{J}\bullet} & 0 & 0 \end{bmatrix}$$

is regular for every $\lambda \in \Lambda$. □

Proposition 4.11. Let $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ be an active set. If (4.14) has a unique solution for one $\bar{\lambda} \in \Lambda$ then $\mathscr{A}$ is regular.

Proof. Let $\bar{\lambda} \in \Lambda$ such that $K(\bar{\lambda})$ is regular. Consider the rank of the matrix $K(\bar{\lambda})$ as defined in (4.14):

$$K(\bar{\lambda}) = \begin{bmatrix} Q(\bar{\lambda}) & -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T \\ -I_{\mathcal{I}\bullet} & 0 & 0 \\ -A_{\mathcal{J}\bullet} & 0 & 0 \end{bmatrix}$$

In order for $K(\bar{\lambda})$ to be of full rank the the columns of $K(\lambda)$ have to be linearly independent. Thus, the matrix

$$[(I_{\mathcal{I}\bullet})^T, (A_{\mathcal{J}\bullet})^T]$$

has to have full rank. Using Proposition 2.12 we see that $K(\lambda)$ is regular for every $\lambda \in \Lambda$.

□

As in the case of efficient complementary bases we consider a decomposition of the weight space by efficient active sets. The following definition was first considered by Goh and Yang [38].

Definition 4.12. Let $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ be an efficient active set. Then the corresponding weight cell $\Lambda^A(\mathscr{A})$ is defined as:

$$\Lambda^A(\mathscr{A}) = \left\{\lambda \in \Lambda \ : \ \exists \text{ solution } (\bar{x}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}) \text{ of (4.14) for } \lambda \text{ satisfying } \bar{x} \in S \text{ and } (\bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}) \geqq 0\right\}.$$

We will now investigate the relationship between efficient active sets of (MQP) and efficient complementary bases of (pLCP).

Definition 4.13. Let $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ be an active set of (MQP). Then the corresponding complementary index set $B_{\mathscr{A}}$ of $\mathscr{A}$ is given by

$$B_{\mathscr{A}}(j) = \begin{cases} \pi_j & \text{if } 1 \leq j \leq p \text{ and } j \in \mathcal{J} \\ s_j & \text{if } 1 \leq j \leq p \text{ and } j \notin \mathcal{J} \\ y_{j-p} & \text{if } j > p \text{ and } (j-p) \in \mathcal{I} \\ x_{j-p} & \text{otherwise} \end{cases} \tag{4.15}$$

Proposition 4.14. Let $\mathscr{A}$ be an efficient regular active set of (MQP). Then $B_{\mathscr{A}}$ is efficient complementary basis of (pLCP) and

$$\Lambda^A(\mathscr{A}) = \Lambda^B(B_{\mathscr{A}}).$$

Proof. Let $(\bar{x}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}})$ be the unique solution of (4.14) for $\mathscr{A}$ and one $\lambda \in \Lambda^A(\mathscr{A})$, i.e.

$$\begin{bmatrix} Q(\lambda) & -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T \\ -I_{\mathcal{I}\bullet} & 0 & 0 \\ -A_{\mathcal{J}\bullet} & 0 & 0 \end{bmatrix} \begin{pmatrix} \bar{x} \\ \bar{y}_{\mathcal{I}} \\ \bar{\pi}_{\mathcal{J}} \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \\ -b_{\mathcal{J}} \end{pmatrix}$$

and $\bar{x} \in S$, $\bar{y}_{\mathcal{I}} \geqq 0$ and $\bar{\pi}_{\mathcal{J}} \geqq 0$. Let $\bar{s} \in \mathbb{R}^p$ be defined as $\bar{s} := A\bar{x} - b$. Note that $\bar{s}_{\mathcal{J}} = 0$ and $\bar{s}_{\bar{\mathcal{J}}} \geq 0$. Notice that $(\bar{x}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}, \bar{s}_{\bar{\mathcal{J}}})$ is a solution of the following system of equations:

$$\begin{bmatrix} Q(\lambda) & -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T & 0 \\ -I_{\mathcal{I}\bullet} & 0 & 0 & 0 \\ -A & 0 & 0 & I_{\bullet\bar{\mathcal{J}}} \end{bmatrix} \begin{pmatrix} x \\ y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \\ s_{\bar{\mathcal{J}}} \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \\ -b \end{pmatrix} \tag{4.16}$$

The matrix $I_{\bullet\bar{\mathcal{J}}}$ denotes the matrix containing the columns with indixes $j \in \bar{\mathcal{J}}$ of the identity matrix I_p. Notice that (4.16) has a unique solution. Since $\bar{x}_{\mathcal{I}} = 0$ we can remove the second equation and the variables $x_{\mathcal{I}}$ from (4.16) and reduce the system to the following:

$$\begin{bmatrix} (Q(\lambda))_{\bullet\bar{\mathcal{I}}} & -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T & 0 \\ -A_{\bullet\bar{\mathcal{I}}} & 0 & 0 & I_{\bullet\bar{\mathcal{J}}} \end{bmatrix} \begin{pmatrix} x_{\bar{\mathcal{I}}} \\ y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \\ s_{\bar{\mathcal{J}}} \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ -b \end{pmatrix}. \tag{4.17}$$

Since $(\bar{x}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}, \bar{s}_{\bar{\mathcal{J}}})$ is the unique solution of (4.16) with $\bar{x}_{\mathcal{I}} = 0$ we know that $(\bar{x}_{\bar{\mathcal{I}}}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}, \bar{s}_{\bar{\mathcal{J}}})$ is the unique solution of (4.17). By rearranging the variables and multiplying the first equation by -1 we observe that the resulting system of equation (4.18)

$$\begin{bmatrix} (I_{\mathcal{I}\bullet})^T & 0 & -(Q(\lambda))_{\bullet\bar{\mathcal{I}}} & (A_{\mathcal{J}\bullet})^T \\ 0 & I_{\bullet\bar{\mathcal{J}}} & -A_{\bullet\bar{\mathcal{I}}} & 0 \end{bmatrix} \begin{pmatrix} y_{\mathcal{I}} \\ s_{\bar{\mathcal{J}}} \\ x_{\bar{\mathcal{I}}} \\ \pi_{\mathcal{J}} \end{pmatrix} = \begin{pmatrix} c(\lambda) \\ -b \end{pmatrix} \tag{4.18}$$

is identical to

$$M_B(\lambda)(y_{\mathcal{I}}, s_{\bar{\mathcal{J}}}, x_{\bar{\mathcal{I}}}, \pi_{\mathcal{J}}) = q(\lambda)$$

for $B = B_{\mathscr{A}}$. Since $(\bar{x}_{\bar{\mathcal{I}}}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}, \bar{s}_{\bar{\mathcal{J}}}) \geqq 0$ we know that $\lambda \in \Lambda^B(B_{\mathscr{A}})$. □

Definition 4.15. Let B be a complementary basis of (pLCP). Then the corresponding active set $\mathscr{A}_B = (\mathcal{I}_B, \mathcal{J}_B)$ of B is given by

$$\mathcal{I}_B = \left\{ j \ : \ y_j \in B \right\} \text{ and } \mathcal{J}_B = \left\{ j \ : \ \pi_j \in B \right\}. \tag{4.19}$$

Proposition 4.16. Let B be an efficient complementary basis of (pLCP). Then $\mathscr{A}_B$ is an efficient regular active set of (MQP) and

$$\Lambda^B(B) = \Lambda^A(\mathscr{A}_B).$$

Proof. We follow the proof of Proposition 4.14 in reverse:
Consider the index sets $\mathcal{I} = \mathcal{I}_B$ and $\mathcal{J} = \mathcal{J}_B$ as defined in (4.19) and the corresponding complements $\bar{\mathcal{I}} = \{1, \ldots, n\} \setminus \mathcal{I}$ and $\bar{\mathcal{J}} = \{1, \ldots, p\} \setminus \mathcal{J}$. Let $(\bar{y}, \bar{s}, \bar{x}, \bar{\pi}) \geqq 0$ be a solution of the linear complementarity problem (pLCP) for $\lambda \in \Lambda$ and complementary basis B, i.e. the vector $z_B := (\bar{y}_{\mathcal{I}}, \bar{s}_{\bar{\mathcal{J}}}, \bar{x}_{\bar{\mathcal{I}}}, \pi_{\mathcal{J}})$ is the unique solution to the system of linear equations

$$M_B(\lambda) z_B = q(\lambda).$$

Consider the equality constraint in (pLCP):

$$\begin{bmatrix} (I_{\mathcal{I}\bullet})^T & 0 & -(Q(\lambda))_{\bullet\bar{\mathcal{I}}} & (A_{\mathcal{J}\bullet})^T \\ 0 & I_{\bullet\bar{\mathcal{J}}} & -A_{\bullet\bar{\mathcal{I}}} & 0 \end{bmatrix} \begin{pmatrix} y_{\mathcal{I}} \\ s_{\bar{\mathcal{J}}} \\ x_{\bar{\mathcal{I}}} \\ \pi_{\mathcal{J}} \end{pmatrix} = \begin{pmatrix} c(\lambda) \\ -b \end{pmatrix}. \tag{4.20}$$

Since $\bar{x}_{\mathcal{I}} = 0$ we can add $x_{\mathcal{I}}$ and the constraint $x_{\mathcal{I}} = 0$ to the system (4.20) without changing the solution set of (4.20):

$$\begin{bmatrix} (I_{\mathcal{I}\bullet})^T & 0 & -(Q(\lambda)) & (A_{\mathcal{J}\bullet})^T \\ 0 & I_{\bullet\bar{\mathcal{J}}} & -A & 0 \\ 0 & 0 & I_{\mathcal{I}\bullet} & 0 \end{bmatrix} \begin{pmatrix} y_{\mathcal{I}} \\ s_{\bar{\mathcal{J}}} \\ x \\ \pi_{\mathcal{J}} \end{pmatrix} = \begin{pmatrix} c(\lambda) \\ -b \\ 0 \end{pmatrix} \tag{4.21}$$

The system of equations (4.21) has a unique solution $(\bar{y}_{\mathcal{I}}, \bar{s}_{\bar{\mathcal{J}}}, \bar{x}, \bar{\pi}_{\mathcal{J}})$ with $\bar{s}_{\bar{\mathcal{J}}} \geqq 0$. Then, $\bar{x}, \bar{y}_{\mathcal{I}}, \pi_{\mathcal{J}}$ are a solution of

$$\begin{bmatrix} Q(\lambda) & -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T \\ -I_{\mathcal{I}\bullet} & 0 & 0 \\ -A_{\mathcal{J}\bullet} & 0 & 0 \end{bmatrix} \begin{pmatrix} \bar{x} \\ \bar{y}_{\mathcal{I}} \\ \bar{\pi}_{\mathcal{J}} \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \\ -b_{\mathcal{J}} \end{pmatrix}$$

with $\bar{x} \in S$, $\bar{y}_{\mathcal{I}} \geqq 0$ and $\bar{\pi}_{\mathcal{J}} \geqq 0$.

□

Corollary 4.17. The set of efficient complementary bases of (pLCP) is connected by diagonal pivots and every efficient active set of (MQP) is regular if one of the following conditions holds:

1. Every efficient solution of (MQP) satisfies the LICQ.
2. The multiobjective problem (MQP) is only constrained by lower and upper bounds, i.e. the feasible set if given by

$$S = \left\{ x \in \mathbb{R}^n \ : 0 \leqq x,\ x_j \leq u_j,\ j \in J_u \right\}$$

for an index set $J_u \subseteq \{1, \ldots, n\}$ and $u_j > 0$ for all $j \in J_u$.

Proof. 1. Follows from Proposition 4.10 and Proposition 4.14.

2. Follows from item 1, as the LICQ is satisfied for all $x \in S$.

□

Corollary 4.18. Let B be a complementary index set with $|B| = r$ and let $\mathscr{A}_B = (\mathcal{I}, \mathcal{J})$ be the corresponding active set. B is a complementary basis if and only if the matrix

$$\begin{bmatrix} I_{\mathcal{I}\bullet} \\ A_{\mathcal{J}\bullet} \end{bmatrix} \tag{4.22}$$

has full rank.

Proof. Follows from Proposition 4.14 and Proposition 4.11. □

4.2.2 Weight Cells

The dimension and other properties of the cells $\Lambda^B(B)$ for more general parametric linear complementarity problems have been investigated by Adelgren [2]. In this section we will discuss properties of weight cells $\Lambda^B(B)$ of the parametric linear complementarity problem (pLCP).
The following extension of Proposition 4.10 was shown by Adelgren [1]:

Proposition 4.19. [1] Let B be a complementary basis of (pLCP). Exactly one of the following statements holds:

1. $\det(M_B(\lambda)) > 0$ for all $\lambda \in \Lambda^B(B)$ or
2. $\det(M_B(\lambda)) < 0$ for all $\lambda \in \Lambda^B(B)$.

Proof. The proof can be found with Proposition 5.14 in Adelgren [1]. □

Definition 4.20. Let B be a complementary basis of (pLCP). Then the functions

$$(q_B(\lambda))_k : \mathbb{R}^m \to \mathbb{R}$$

are called **basic value functions** for $k = 1, \dots, r$.
The hypersurfaces that define the boundary segments of $\Lambda^B(B)$ in Λ are denoted in the following way:

$$\mathfrak{h}_B^k := \{\lambda \in \mathbb{R}^m \ : \ (q_B(\lambda))_k = 0\}$$

Proposition 4.21. [1] Let B be an efficient complementary basis of (pLCP). Then the hypersurfaces $\mathfrak{h}_B^k$ are semi-algebraic sets for all $k = 1, \dots, r$, i.e. $\mathfrak{h}_B^k$ is defined by polynomial equations for $k = 1, \dots, r$. Additionally, the weight cell $\Lambda^B(B)$ is a semi-algebraic set, i.e. $\Lambda^B(B)$ is defined by polynomial equations and inequalities.

Proof. We follow the proof of Proposition 5.14 in [1]. Since B is a complementary basis we can write $q_B(\lambda)$ in the following way:

$$q_B(\lambda) = (M_B(\lambda))^{-1} q(\lambda) = \frac{\mathrm{Adj}(M_B(\lambda))}{\det(M_B(\lambda))} q(\lambda) \tag{4.23}$$

Where $\mathrm{Adj}(M_B(\lambda))$ refers to the adjoint matrix of $M_B(\lambda)$. Notice that $\mathrm{Adj}(M_B(\lambda))q(\lambda)$ is polynomial in λ. Thus,

$$\Lambda^B(B) = \{\lambda \in \Lambda \ : \ \mathrm{sign}(\det(M_B(\lambda)))\mathrm{Adj}(M_B(\lambda))q(\lambda) \geqq 0\}\,.$$

□

Corollary 4.22. The weight cells $\Lambda^B(B)$ are closed and bounded sets.

Proof. Follows from Proposition 4.21 as the polynomials that describe $\Lambda^B(B)$ are continuous over $\mathbb{R}^m$. □

Theorem 4.23. [13, 1] Let B_{eff} be the set of efficient complementary bases of (pLCP).

$$\Lambda = \bigcup_{B \in B_{\text{eff}}} \Lambda^B(B).$$

Furthermore, the set Λ can be decomposed into cells $\Lambda^B(B)$, i.e.

$$\mathrm{relint}\Lambda^B(B) \cap \mathrm{relint}\Lambda^B(B') = \emptyset$$

for any $B, B' \in B_{\text{eff}}$, $B \neq B'$.

Proof. For every $\lambda \in \Lambda$ we can solve the weighted sum scalarization problem (WQP) using the criss-cross method (Algorithm 2.1) and compute an efficient complementary bases.

Now, assume that there exists $\lambda \in \Lambda^B(B) \cap \Lambda^B(B')$ for some $B, B' \in B_{\text{eff}}$, $B \neq B'$. Recall that the solution $\bar{x} := \bar{x}(\lambda)$ of the weighted sum problem (WQP) is unique for every $\lambda \in \Lambda$. Let $\mathscr{A}_B = (\mathcal{I}, \mathcal{J})$ and $\mathscr{A}_{B'} = (\mathcal{I}', \mathcal{J}')$ be the regular active sets corresponding to B and B', respectively. Notice that the efficient active sets are not identical $\mathscr{A}_B \neq \mathscr{A}_{B'}$.
The optimal solution $\bar{x}$ of the weighted sum problem (WQP) has to satisfy

$$\bar{x}_{\mathcal{I} \cup \mathcal{I}'} = 0 \text{ and } A_{\mathcal{J} \cup \mathcal{J}' \bullet} \bar{x} = b_{\mathcal{J} \cup \mathcal{J}'}.$$

Consider $i \in \mathcal{I}' \setminus \mathcal{I}$. The unique solution $(\bar{x}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}})$ of the active set system wrt. $\mathscr{A}_B$ must satisfy the equation $\bar{x}_i = 0$. Thus $\lambda \in \text{bd } \Lambda^A(\mathscr{A}_B)$. Similarly if $j \in \mathcal{J}' \setminus \mathcal{J}$ then $\bar{x}$ satisfies $A_{j \bullet} \bar{x} = b_j$ for $\mathscr{A}_B$, showing that $\lambda \in \text{bd } \Lambda^A(\mathscr{A}_B)$. By interchanging the roles of $\mathcal{I}$ and $\mathcal{I}'$ or $\mathcal{J}$ and $\mathcal{J}'$ we see that:

$$\mathrm{relint}\Lambda^A(\mathscr{A}_B) \cap \mathrm{relint}\Lambda^A(\mathscr{A}_{B'}) = \emptyset.$$

□

4.2.3 Decomposition of the Efficient Set

Using the decomposition of the weight space by active sets of complementary bases we can also decompose the efficient set. Recall that $\bar{x}(\lambda)$ is the unique optimal solution of the weighted sum scalarization problem for $\lambda \in \Lambda$.

Definition 4.24. Let B be an efficient complementary basis of (pLCP) and let $\bar{x}(\lambda)$ be the optimal solution of (WQP) for $\lambda \in \Lambda$. Then we define the corresponding subset of the efficient points by

$$\bar{X}(B) := \left\{ \bar{x}(\lambda) \ : \ \lambda \in \Lambda^B(B) \right\}.$$

Proposition 4.25. [38] Let B_{eff} be the set of efficient complementary bases (pLCP) and let X_E be the efficient set of (MQP). Then the following statement holds:

$$X_E = \bigcup_{B \in B_{\text{eff}}} \bar{X}(B)$$

Proof. Follows from Theorem 4.23, Theorem 2.25 and Theorem 2.26 □

Example 4.26. Consider the following biobjective problem with two variables:

$$\begin{array}{ll} \min & f_i(x) = \frac{1}{2}x^T Q^i x + (c^i)^T x, \quad i = 1,2 \\ \text{s.t.} & Ax \geqq b, \; x \geqq 0 \end{array}$$

with

$$Q^1 = \begin{pmatrix} 4 & -3 \\ -3 & 5 \end{pmatrix},\; Q^2 = I_2,\; c^1 = \begin{pmatrix} 3.5 \\ -9.5 \end{pmatrix},\; c^2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix},\; A = \begin{pmatrix} 4 & -1 \\ -8 & 1 \end{pmatrix},\; b = \begin{pmatrix} 2 \\ -6 \end{pmatrix}$$

After substituting $\lambda_2 = 1 - \lambda_1$ we can write the parametric linear complementarity problem as

$$\begin{pmatrix} 1 & 0 & 0 & 0 & (-3\lambda_1 - 1) & 3\lambda_1 & 4 & -8 \\ 0 & 1 & 0 & 0 & 3\lambda_1 & (-4\lambda_1 - 1) & -1 & 1 \\ 0 & 0 & 1 & 0 & -4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 8 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ s_1 \\ s_2 \\ x_1 \\ x_2 \\ \pi_1 \\ \pi_2 \end{pmatrix} = \begin{pmatrix} \frac{9\lambda_1}{2} - 1 \\ -\frac{19\lambda_1}{2} \\ -2 \\ 6 \end{pmatrix} \tag{4.24}$$

with $x_i y_i = 0$ for $i = 1, \dots, n$ and $s_j \pi_j = 0$ for $j = 1, \dots, p$ and $y, s, x, \pi \geqq 0$.
For $\lambda = (0,1)^T$ the feasible basis of the linear complementarity problem (4.24) is $B_1 = (s_1, \pi_2, x_1, x_2)$ with the corresponding efficient point $\bar{x}((0,1)) = \frac{1}{65}(49,2)^T$. For B_1 the basic value is given by:

$$\begin{pmatrix} s_1 \\ \pi_2 \\ x_1 \\ x_2 \end{pmatrix} = q_{B_1}(\lambda) = \begin{pmatrix} 0 & (-3\lambda_1 - 1) & 3\lambda_1 \\ 0 & 3\lambda_1 & (-4\lambda_1 - 1) \\ 1 & -4 & 1 \\ 0 & 8 & -1 \end{pmatrix}^{-1} \begin{pmatrix} \frac{9\lambda_1}{2} - 1 \\ -\frac{19\lambda_1}{2} \\ -2 \\ 6 \end{pmatrix} = \frac{1}{211\lambda_1 + 65} \begin{pmatrix} 64 - 138\lambda_1 \\ 51\lambda_1^2 - \frac{117}{2}\lambda_1 + 2 \\ \frac{491}{2}\lambda_1 + 49 \\ 698\lambda_1 + 2 \end{pmatrix}$$

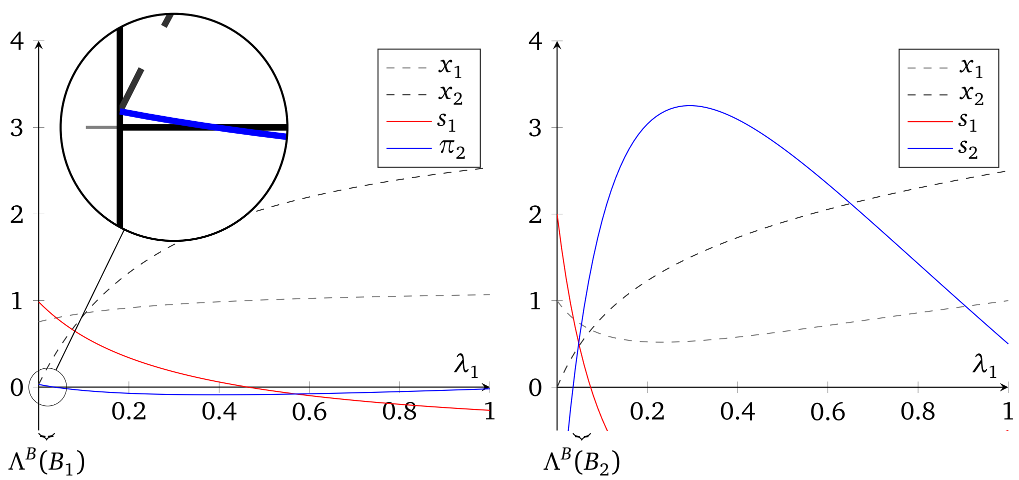

Figure 4.1: Basic values for B_1 (left) and B_2 (right) in Example 4.26.

The graphs of basic values as a function of λ_1 for B_1 are shown in the left part of Figure 4.1. The basic value of π_2 vanishes at approximately $\lambda_1 = 0.0353$. The complementary basis B_1 remains a feasible basis for the parametric linear complementarity problem for all $\lambda_1 \in [0, 0.0353]$, i.e. $\Lambda^B(B_1) = [0, 0.0353]$.
Another efficient complementary basis can be found by a diagonal pivot of π_2. The new basis $B_2 = (s_1, s_2, x_1, x_2)$ is feasible for $\lambda_1 = 0.0353$. In an analogous manner the basic values as a function of B_2 can be computed. The graphs of these functions are shown on the right part of Figure 4.1.
The graphs of the basic functions of two additional efficient bases $B_3 = (x_1, x_2, \pi_1, s_2)$ and $B_4 = (x_1, x_2, \pi_1, \pi_2)$ are illustrated in Figure 4.2. We can make the following observations:

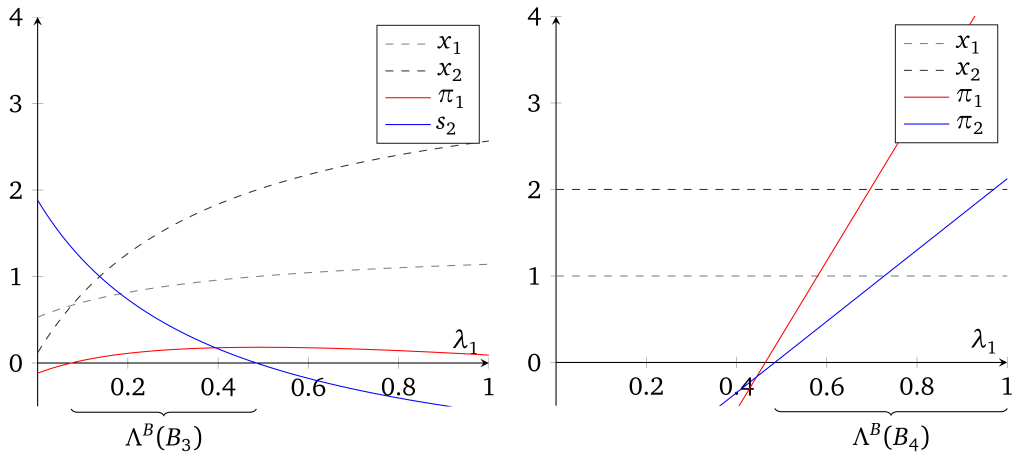

Figure 4.2: Basic values for B_3 (left) and B_4 (right) in Example 4.26.

- The bases B_1 and B_4 are connected by a pivot of the complementary variable pair s_1/π_1 but B_1 and B_4 are not adjacent, i.e. $\Lambda^B(B_1) \cap \Lambda^B(B_2) = \emptyset$.
- Basis B_4 represents an active set with 2 active constraints and $\bar{X}(B_4)$ consists only of one point. The basic values are linear in λ_1 for this efficient basis. This property motivates Proposition 4.27.

The weight space decomposition of (4.24) can now be used to analytically describe the efficient set which is illustrated in Figure 4.3. Additionally, the curves

$$C^l = \left\{(x_1, x_2) \ : \ x_1 = (q_{B_l}(\lambda))_3 \text{ and } x_2 = (q_{B_l}(\lambda))_4,\ \lambda \in \Lambda\right\}$$

are shown in dotted lines for $l = 1, 2, 3$ to visualize the connection to active sets. For example C^2 is the efficient set of the unconstrained problem $\text{vmin}_{x \in \mathbb{R}^n} f(x)$ and a segment of C^2 is a part of the efficient set of (4.24).

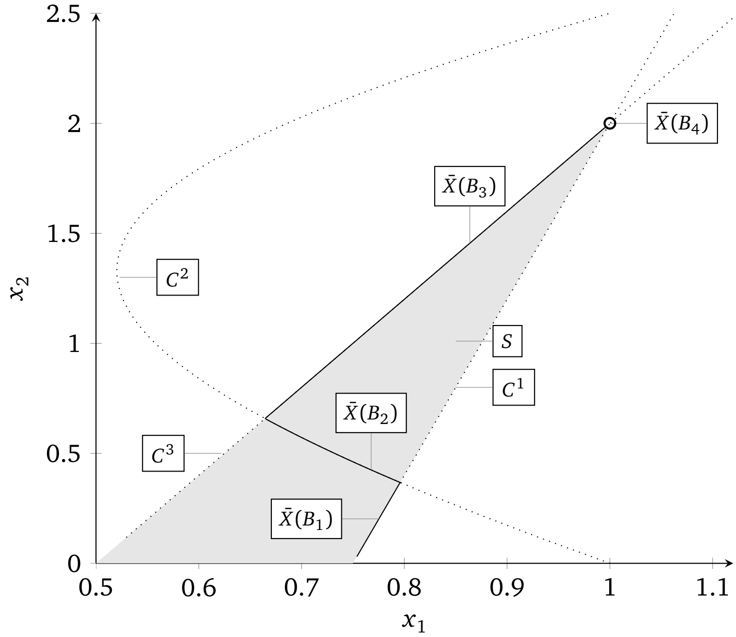

Figure 4.3: Decomposition of the Efficient Set of Example 4.26

Proposition 4.27. Let B be an efficient complementary basis of (MQP) and let $\mathscr{A}_B = (\mathcal{I}, \mathcal{J})$ be the corresponding active set.
If $\dim \bar{X}(B) = 0$ and $|\mathcal{I}| + |\mathcal{J}| = n$ then $\Lambda^B(B)$ is a convex polyhedron.

Proof. Let $\tilde{x} \in S$ be such that $\bar{X}(B) = \{\tilde{x}\}$. Since B is a complementary basis the efficient active set $\mathscr{A}_B = (\mathcal{I}, \mathcal{J}))$ is regular and $\Lambda^B(B) = \Lambda^A(\mathscr{A}_B)$ (see Proposition 4.16). Consider the active set system:

$$\begin{bmatrix} Q(\lambda) & -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T \\ -I_{\mathcal{I}\bullet} & 0 & 0 \\ -A_{\mathcal{J}\bullet} & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \\ -b_{\mathcal{J}} \end{pmatrix} \tag{4.25}$$

Since only $\tilde{x}$ solves the systems of equations (4.25) we can insert $\tilde{x}$ into (4.25) and reduce it to

$$\begin{bmatrix} -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T \end{bmatrix} \begin{pmatrix} y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \end{pmatrix} = -c(\lambda) - Q(\lambda)\tilde{x}. \tag{4.26}$$

The matrix $\begin{bmatrix} -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T \end{bmatrix}$ is (by assumption) square and regular (otherwise B would not be a basis of (pLCP)). Thus, we can formulate an explicit representation of the dual variables $y_{\mathcal{I}}$ and $\pi_{\mathcal{J}}$ for every $\lambda \in \Lambda$.

$$\begin{pmatrix} y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \end{pmatrix} = \begin{bmatrix} -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T \end{bmatrix}^{-1} (-c(\lambda) - Q(\lambda)\tilde{x}) \tag{4.27}$$

The right-hand-side and thus the entries of the solutions $y_{\mathcal{I}}$ and $\pi_{\mathcal{J}}$ of (4.26) consists of polynomials of degree 1 wrt. λ. Hence, the non-negativity constraints $y_{\mathcal{I}} \geqq 0$ and $\pi_{\mathcal{J}} \geqq 0$ are linear inequality constraints in λ, i.e. the set

$$\left\{ \lambda \in \Lambda \;:\; \begin{bmatrix} (I_{\mathcal{I}\bullet})^T & (A_{\mathcal{J}\bullet})^T \end{bmatrix}^{-1} (-c(\lambda) - Q(\lambda)\tilde{x}) \geqq 0 \right\}$$

is a convex polyhedron.

□

An example for efficient complementary basis with a few polyhedral weight cells is given in Example 4.47.

Example 4.28. Consider the following convex multiobjective quadratic problem (MQP):

$$\begin{array}{ll} \text{vmin} & f_i(x) = \frac{1}{2}x^T Q^i x + (c^i)^T x, \;\; i = 1, \dots, m \\ \text{s.t.} & Ax \geqq b, \;\; x \geqq 0 \end{array}$$

with

- $n = 2$, $m = 3$, $p = 1$
- $Q^1 = Q^3 = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $Q^2 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$
- $c^1 = \begin{pmatrix} 0 \\ -1.8 \end{pmatrix}, c^2 = \begin{pmatrix} -3.6 \\ 0 \end{pmatrix}, c^3 = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$
- $A = \begin{bmatrix} 1 & 1 \end{bmatrix}$, $b = 2$

The lexicographic minima of the unconstrained problem $\text{vmin}_{x \in \mathbb{R}^n} f(x)$ are

$$x^1_{\text{lex}} = \begin{pmatrix} 0 \\ 1.8 \end{pmatrix}, \; x^2_{\text{lex}} = \begin{pmatrix} 1.8 \\ 0 \end{pmatrix}, \; x^3_{\text{lex}} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

The weight space decomposition is illustrated in the right part of Figure 4.4. There are two efficient complementary bases $B_1 = (x_1, x_2, \pi_1)$ (shown in green) and $B_2 = (x_1, x_2, s_1)$ (shown in blue). The corresponding points in the efficient set are shown in the left part of Figure 4.4.
We can observe that the curve $(q_{B_1}(\lambda))_3 = 0$ (shown as a dashed line in the right part of Figure 4.4) has two unconnected segments common with weight space Λ. The weight cell $\Lambda^B(B_1)$ is not connected.

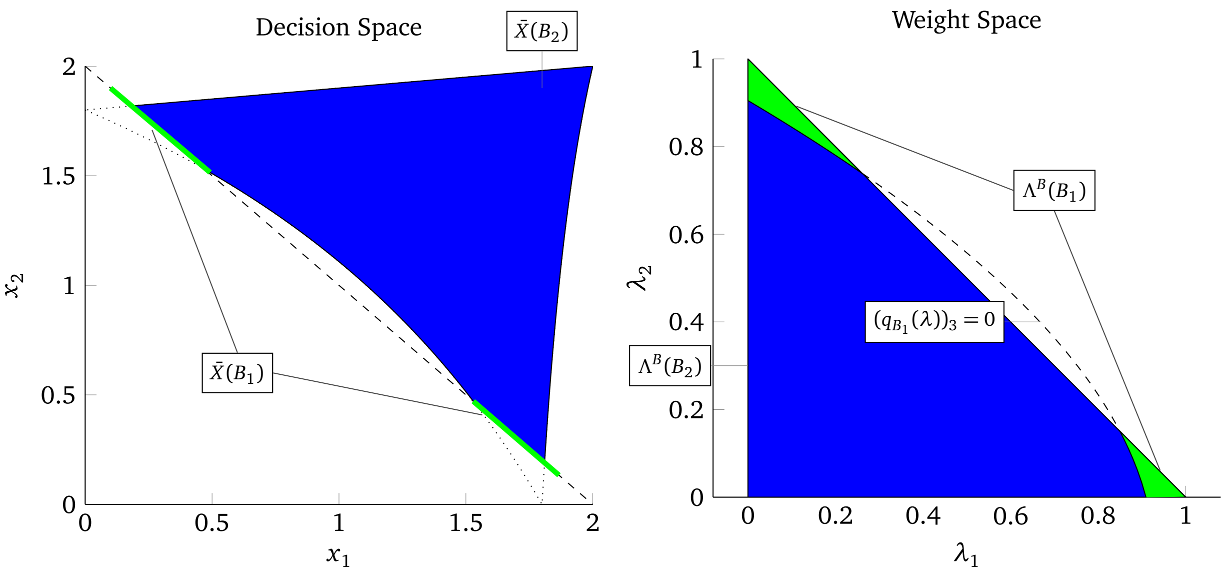

Figure 4.4: Decision Space (left) and Weight Space (right) of Example 4.28.

Observation 4.29. $\Lambda^B(B)$ can be unconnected.

4.3 Determination of Efficient Bases

Recall the definition of the weight cells $\Lambda^B(B)$ for a given efficient complementary basis B:

$$\Lambda^B(B) := \{\lambda \in \Lambda \;:\; q_B(\lambda) \geqq 0\}$$

From the definition of $\Lambda^B(B)$ be can observe that the boundary of $\Lambda^B(B)$ consists of $\lambda \in \Lambda$ for which there exists $k \in \{1, \dots, r\}$ such that $(q_B(\lambda))_k = 0$ or there exists some $i \in \{1, \dots, m\}$ with $\lambda_i = 0$. Only in the former case another efficient complementary basis B' can be found with $\Lambda^B(B) \cap \Lambda^B(B') \neq \emptyset$.
Goh and Yang [38] proposed an approach where for each $k \in \{1, \dots, r\}$ the hypersurfaces $\mathfrak{h}_B^k := \{\lambda \in \mathbb{R}^m \;:\; (q_B(\lambda))_k = 0\}$ that define the boundary of $\Lambda^B(B)$ are computed using symbolic calculations. This approach is only valid if all efficient complementary bases are connected by diagonal pivots, i.e. every efficient active set is regular.
The approach was extended by Adelgren [1] to cover

- a broader class of problems: parametric linear complementarity problems
- multiobjective quadratic optimization problems for which not every efficient active set is regular
- only the determination of efficient complementary bases B with $\dim \Lambda^B(B) \geq m - 2$.

The approach by Adelgren is reviewed in Section 4.3.3. In the following we discuss an approach based on the approaches by Goh and Yang [38] and Adelgren [1] to compute all efficient complementary bases.

Recall the definitions of diagonal and exchange pivots from Definition 2.14. First, we show that it is enough to only consider diagonal and exchange pivots for the determination of all efficient complementary bases:

Proposition 4.30. [1, 13] The set of efficient complementary bases of (pLCP) is connected by diagonal and exchange pivots.

Proof. Proposition 5.54 from [1] (based on Lemma 3.8 from [13]) shows that if two efficient complementary bases B and B' share an $(m-2)$-dimensional boundary segment then $|B \cap B'| \geq r - 2$, which is equivalent to either $B' = \text{diag}(B, k)$ for some $k \in \{1, \dots, r\}$ or $B' = \text{exch}(B, k, l)$ for some $k, l \in \{1, \dots, r\}$ with $l \neq k$. □

Goh and Yang [38] considered a method where the boundary segments of two weight cells are determined using symbolic calculations. The following similar approach only considers individual points on the boundary of $\Lambda^B(B)$ and does not require any symbolic calculations.
First, we discuss sufficient and necessary criteria for diagonal and exchange pivots leading to efficient complementary bases.

Proposition 4.31. Let B be an efficient complementary basis of (pLCP) and let $k \in \{1, \dots, r\}$.
If $B' = \text{diag}(B, k)$ is a basis and there exists $\lambda \in \Lambda^B(B)$ such that $(q_B(\lambda))_k = 0$ then B' is an efficient complementary basis of (MQP).

Proof. Let λ be a weight such that $\lambda \in \Lambda^B(B)$ and $(q_B(\lambda))_k = 0$. Let $T(\lambda)$ be the dictionary wrt. B and λ. From Proposition 2.20 we know that $T_{kk}(\lambda) \neq 0$ for λ because B' is a basis.
Recall the definition of the working tableau $[T(\lambda)\; I_r\; q_B(\lambda)]$ from Section 2.1.2.
The basic value $q_{B'}(\lambda)$ wrt. the new basis B' can be computed from $q_B(\lambda)$ by a principal pivot of the working tableau:

$$(q_{B'}(\lambda))_l = (q_B(\lambda))_l - \frac{T_{lk}(\lambda)}{T_{kk}(\lambda)} \underbrace{(q_B(\lambda))_k}_{=0} = (q_B(\lambda))_l$$

for $l = 1, \ldots, r$. Since $\lambda \in \Lambda^B(B)$ we know that $q_B(\lambda) \geqq 0$ showing that $q_{B'}(\lambda) \geqq 0$ and $\Lambda^B(B') \neq \emptyset$.

□

A similar result can be shown for exchange pivots:

Proposition 4.32. Let B be an efficient complementary basis of (pLCP) and let $k, l \in \{1, \ldots, r\}$ with $k \neq l$.
If $B' = \text{exch}(B, k, l)$ is a basis and there exists $\lambda \in \Lambda^B(B)$ such that $(q_B(\lambda))_k = 0$ and $(q_B(\lambda))_l = 0$ then B' is an efficient complementary basis of (MQP).

Proof. Let λ be a weight such that $\lambda \in \Lambda^B(B)$ and $(q_B(\lambda))_k = (q_B(\lambda))_l = 0$. Let $T(\lambda)$ be the dictionary wrt. B and λ. Consider the following two cases:

- There exists a complementary basis B'' such that $B'' = \text{diag}(B, k)$ and $B' = \text{diag}(B'', l)$ or $B'' = \text{diag}(B, l)$ and $B' = \text{diag}(B'', k)$.
  In either case we can apply Proposition 4.31 twice, once for the pivot of the basic variable $z_k \in B$ and again for $z_l \in B'$.
- Otherwise, let T be the dictionary wrt. B and let $\lambda \in \Lambda^B(B)$ be given with $(q_B(\lambda))_k = 0$ and $(q_B(\lambda))_l = 0$. Then according to Theorem 2.21
  - $T_{kk} = 0$ and $T_{lk} < 0$ or
  - $T_{ll} = 0$ and $T_{kl} < 0$.

  Since both cases are symmetrical wrt. k and l we only consider the first case. The basic values for B' can be computed by principal pivots of the working tableau $[T \;\; I_r \;\; q_B(\lambda)]$. The computation of the new basic value after an exchange pivot can be found in Section 2.1.2.

$$q_{B'} = \begin{pmatrix} (q_B(\lambda))_1 - \frac{T_{1l}}{T_{kl}}(q_B(\lambda))_k - \frac{T_{1k}}{T_{lk}}\left((q_B(\lambda))_l - \frac{T_{ll}}{T_{kl}}(q_B(\lambda))_k\right) \\ \vdots \\ \frac{(q_B(\lambda))_k}{T_{kl}} \\ \vdots \\ \frac{1}{T_{lk}}\left((q_B(\lambda))_l - \frac{T_{ll}}{T_{kl}}(q_B(\lambda))_k\right) \\ \vdots \\ (q_B(\lambda))_r - \frac{T_{rl}}{T_{kl}}(q_B(\lambda))_k - \frac{T_{rk}}{T_{lk}}\left((q_B(\lambda))_l - \frac{T_{ll}}{T_{kl}}(q_B(\lambda))_k\right) \end{pmatrix}$$

  Note that since $(q_B(\lambda))_k = (q_B(\lambda))_l = 0$ it follows that $q_{B'}(\lambda) = q_B(\lambda)$. Hence, $q_{B'}(\lambda) \geqq 0$ and B' is an efficient complementary basis.

□

Proposition 4.31 and 4.32 are sufficient criteria to check whether a given basis is efficient. The following two propositions show that if $\Lambda^B(B) \cap \Lambda^B(B') \neq \emptyset$ then a weight $\lambda \in \Lambda^B(B)$ can be found that satisfies the assumptions of Proposition 4.31 or 4.32.

Proposition 4.33. Let B and B' be efficient complementary bases of (pLCP) with $B' = \text{diag}(B, k)$ for some $k \in \{1, \ldots, r\}$. If $\Lambda^B(B) \cap \Lambda^B(B') \neq \emptyset$ then there exists $\lambda \in \Lambda^B(B)$ such that $(q_B(\lambda))_k = 0$.

Proof. Let $\mathscr{A}_B = (\mathcal{I}, \mathcal{J})$ and $\mathscr{A}_{B'} = (\mathcal{I}', \mathcal{J}')$ be the active sets corresponding to B and B', respectively. Then exactly one of the following situations is possible:

- $k \leq p$, $\mathcal{I}' = \mathcal{I}$ and $\mathcal{J}' = \mathcal{J} \cup \{k\}$
- $k \leq p$, $\mathcal{I}' = \mathcal{I}$ and $\mathcal{J} = \mathcal{J}' \cup \{k\}$
- $k > p$, $\mathcal{I} = \mathcal{I}' \cup \{k-p\}$ and $\mathcal{J}' = \mathcal{J}$
- $k > p$, $\mathcal{I}' = \mathcal{I} \cup \{k-p\}$ and $\mathcal{J}' = \mathcal{J}$

We will only consider the first case as all other cases can be handled analogously. Consider the active systems (4.14) for $\lambda \in \Lambda^B(B) \cap \Lambda^B(B')$:

$$\begin{bmatrix} Q(\lambda) & -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T \\ -I_{\mathcal{I}\bullet} & 0 & 0 \\ -A_{\mathcal{J}\bullet} & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \\ -b_{\mathcal{J}} \end{pmatrix} \tag{4.28}$$

and

$$\begin{bmatrix} Q(\lambda) & -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T & -(A_{k\bullet})^T \\ -I_{\mathcal{I}\bullet} & 0 & 0 & 0 \\ -A_{\mathcal{J}\bullet} & 0 & 0 & 0 \\ -A_{k\bullet} & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \\ \pi_k \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \\ -b_{\mathcal{J}} \\ -b_k \end{pmatrix} \tag{4.29}$$

Let $(\bar{x}(\lambda), \bar{y}_{\mathcal{I}}(\lambda), \bar{\pi}_{\mathcal{J}}(\lambda))$ be a solution of (4.3). Then $(\bar{x}(\lambda), \bar{y}_{\mathcal{I}}(\lambda), \bar{\pi}_{\mathcal{J}}(\lambda), 0)$ is a solution of (4.3) and since (4.3) has a unique solution for every $\lambda \in \Lambda$ it is also unique. The variables π_k and s_k are nonbasic in B and B', respectively. Hence, it follows that $0 = \pi_k = (q_B(\lambda))_k = (q_{B'}(\lambda))_k = s_k$ at $\lambda \in \Lambda^B(B) \cap \Lambda^B(B')$.

□

Proposition 4.34. Let B and B' be efficient complementary bases of (pLCP) with $B' = \text{exch}(B, k, l)$ for some $k, l \in \{1, \dots, r\}$ with $k \neq l$. If $\Lambda^B(B) \cap \Lambda^B(B') \neq \emptyset$ then there exists $\lambda \in \Lambda^B(B)$ such that $(q_B(\lambda))_k = 0$ and $(q_B(\lambda))_l = 0$.

Proof. The proof is analogous to the proof of Proposition 4.33. □

To determine whether a complementary index set is a basis we can use the dictionary (see Section 2.1.2, particularly Proposition 2.20 and Theorem 2.21). Alternatively, Proposition 4.14 can be used to check if a given complementary index set is a complementary basis by checking if the corresponding active set system has a unique solution for all $\lambda \in \Lambda$.

Our goal is to compute all efficient complementary bases of (pLCP). In general, the set of efficient bases is not connected by diagonal pivots alone. However, we don't have to consider every possible exchange pivot: If B is a complementary basis and, for example, $B' = \text{diag}(B, k)$ and $B'' = \text{exch}(B, k, l)$ are complementary bases for some $k, l \in \{1, \dots, r\}$ with $k \neq l$ then B'' can be obtained form B' by a diagonal pivot, i.e. $B'' = \text{diag}(B', l)$. This pivot can be considered when all adjacent bases of B' are computed. Thus, we only need to consider exchange pivots with $k, l \in \{1, \dots, r\}$ if $B' = \text{diag}(B, k)$ is not a basis.
Let B be an efficient complementary basis of (pLCP) and $k \in \{1, \dots, r\}$ such that $B' = \text{diag}(B, k)$ is a complementary basis. Consider the following nonlinear optimization problem:

$$\Psi_d(B, k) = \begin{cases} \min\limits_{\lambda \in \mathbb{R}^m} & (q_B(\lambda))_k \\ \text{s.t.} & q_B(\lambda) \geqq 0 \\ & \lambda \in \Lambda \end{cases} \tag{4.30}$$

Procedure 4.1: Compute Possible Adjacent Bases

```
DPivots = ∅, EPivots = ∅;
for k = 1,...,r do
    if B' = diag(B,k) is a complementary basis then
        DPivots = DPivots ∪ {k};
    else
        for l = 1,...,r, l ≠ k do
            if B' = exch(B,k,l) is a complementary basis and exch(B,l,k) ∉ EPivots then
                EPivots = EPivots ∪ {(k,l)};
Output: DPivots, EPivots
```

Proposition 4.35. Let B be an efficient complementary basis and let $k \in \{1,\dots,r\}$ such that $B' = \text{diag}(B,k)$ is a basis. If $\Psi_d(B,k) = 0$ then B' is an efficient complementary basis.

Proof. Follows directly from Proposition 4.31. □

Remark 4.36. Note that the converse of Proposition 4.35 is not true in general. More precisely, B and B' may be efficient complementary bases but $\Lambda^B(B) \cap \Lambda^B(B') = \emptyset$. One occurrence for this can be seen in Example 4.26 where B_4 can be obtained from B_1 via a diagonal pivot and both B_1 and B_4 are efficient complementary bases but $\Lambda^B(B_1) \cap \Lambda^B(B_4) = \emptyset$.

For exchange pivots we can formulate a similar result using the following optimization problem:

$$\Psi_e(B,k,l) = \begin{cases} \min\limits_{\lambda \in \mathbb{R}^m} & (q_B(\lambda))_k + (q_B(\lambda))_l \\ \text{s.t.} & q_B(\lambda) \geqq 0 \\ & \lambda \in \Lambda \end{cases} \tag{4.31}$$

Proposition 4.37. Let B be an efficient complementary basis and let $k,l \in \{1,\dots,r\}$ with $k \neq l$ and $B' = \text{exch}(B,k,l)$ being a basis. If $\Psi_e(B,k) = 0$ then B' is an efficient complementary basis.

Proof. Follows directly from Proposition 4.32. □

For numerical reasons it is preferable to avoid the explicit enumeration of the inverse of the matrix $M_B(\lambda)$. In fact, we observed during numerical experiments that the nonlinear optimization solver of MATLAB, fmincon, evaluated the objective function of (4.30) for $\lambda \notin \Lambda$ for which the regularity of $M_B(\lambda)$ is not guaranteed. If $M_B(\lambda)$ was not regular fmincon stopped with an error message as the objective function of (4.30) could not be evaluated.
To avoid this situation the following equivalent optimization problems are used:

$$\Psi_d(B,k) = \begin{cases} \min\limits_{\lambda \in \mathbb{R}^m,\ v \in \mathbb{R}^r} & v_k \\ \text{s.t.} & M_B(\lambda)v = q(\lambda) \\ & v \geqq 0,\ \lambda \in \Lambda \end{cases} \tag{4.32}$$

$$\Psi_e(B,k,l) = \begin{cases} \min\limits_{\lambda \in \mathbb{R}^m,\ v \in \mathbb{R}^r} & v_k + v_l \\ \text{s.t.} & M_B(\lambda)v = q(\lambda) \\ & v \geqq 0,\ \lambda \in \Lambda \end{cases} \tag{4.33}$$

Notice that any $\lambda \in \Lambda^B(B)$ is feasible for (4.30) and (4.31). Similarly, $(\lambda, q_B(\lambda))$ is feasible for (4.32) and (4.33) for any $\lambda \in \Lambda^B(B)$.

4.3.1 The Algorithm

We now have all the ingredients to formulate a strategy to compute all efficient complementary bases of (MQP).
During the execution of Algorithm 4.2 two sets of complementary bases F and E are maintained. E consists of all efficient bases for which all possible adjacent bases have been checked for efficiency. F consists of all efficient bases for which adjacent bases have to be checked.
In the initialization of Algorithm 4.2 an initial efficient complementary basis of (pLCP) is computed. Since we assume that the feasible set is nonempty and the objective functions $f_i(x)$ of (MQP) are strictly convex for all $i = 1, \dots, m$ we know that there exists at least one efficient complementary basis. The time complexity of Algorithm 4.2 is discussed in Section 4.3.1.1.

Algorithm 4.2: Adjacency Search

```
Determine an initial efficient complementary basis B_0 by solving the linear complementarity
  problem for some λ ∈ Λ, for example λ = (1/m, ... 1/m), with Algorithm 2.1.
Set F = {B_0} and E = ∅
while F ≠ ∅ do
    Choose B ∈ F
    F = F \ {B}
    E = E ∪ {B}
    Compute DPivots und EPivots with Procedure 4.1.
    for each k ∈ DPivots do
        B' = diag(B, k)
        if B' ∉ E ∩ F then
            if Ψ_d(B, k) = 0 then
                F = F ∪ {B'}
    for each (k, l) ∈ EPivots do
        B' = exch(B, k, l)
        if B' ∉ E ∩ F then
            if Ψ_e(B, k, l) = 0 then
                F = F ∪ {B'}
Output: Set of efficient complementary bases E
```

4.3.1.1 Worst-case Time Complexity

Several steps in Algorithm 4.2 have exponential worst-case time complexity. First, in order to compute the initial efficient complementary basis a linear complementarity problem has to be solved. The criss-cross method (see Section 2.1.3) has exponential worst-case time complexity [14]. The practical performance of the criss-cross method is acceptable for the size of problems we consider here. However, there are polynomial-time algorithms available, for example the algorithm proposed by Kojima et. al. [54] where an approximate solution of a linear complementarity problem (LCP) is computed that satisfies

$$u_k v_k \leq \varepsilon$$

for a small $\varepsilon > 0$. In order to find a feasible complementary basis from this solution a basis identification algorithm has to be applied which can be done in polynomial time [5].
Another important step is the computation of $\Psi_d(B, k)$ and $\Psi_e(B, k, l)$. Since (4.32) and (4.33) are nonconvex problems with possibly unconnected feasible sets, they are in general NP-hard problems [63].

In each iteration at most r diagonal pivots and $\frac{r^2-r}{2}$ exchange pivots have to be considered [13]. A particular complementary basis might be considered up to $r+\frac{r^2-r}{2}$ times as a possible adjacent efficient bases unless it is found to be an efficient complementary basis in a previous iteration.
Additionally, the number of efficient complementary bases can be exponential in the number of complementary variable pairs $r = n + p$. To see this, consider the following instance of (MQP):

Example 4.38. Let $n \geq 2$, $m \geq 2$, $p = 0$ and $c^1 = 0$ be given and consider the following problem:

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{vmin}} & f_i(x) = \frac{1}{2}x^T Q^i x + (c^i)^T x \quad i = 1,\ldots.m \\ \text{s.t.} & \qquad x \geqq 0 \end{array} \tag{4.34}$$

Notice that the active set system

$$\begin{bmatrix} Q(\lambda) & -(I_{\mathcal{J}\bullet})^T \\ -I_{\mathcal{J}\bullet} & 0 \end{bmatrix} \begin{pmatrix} x \\ y_{\mathcal{J}} \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \end{pmatrix} \tag{4.35}$$

has a unique solution for all $\lambda \in \Lambda$ and for all active sets of (4.34) since $I_{\mathcal{J}\bullet}$ has full rank for all subsets $\mathcal{J}$ of $\{1,\ldots,n\}$.
The global minimum of f_1 over $\mathbb{R}^n$ is the vector 0 and 0 is also feasible for (4.34). Thus, $(0,\ldots,0)^T$ is an efficient solution of (4.34). Notice that $x = 0$ and $y_{\mathcal{J}} = 0$ solve (4.35) for $\lambda = (1,0,\ldots,0)^T$ for every $\mathcal{J} \subseteq \{1,\ldots,n\}$. So all possible active sets and the corresponding complementary bases of (4.34) are efficient, and (4.35) yields

$$1 + \sum_{k=1}^{n} \binom{n}{k}$$

efficient complementary bases because every subset of $\{1,\ldots,n\}$ yields an efficient active set of (4.34).

4.3.1.2 Notes on the Implementation

Algorithm 4.2 and the criss-cross method (Algorithm 2.1) were implemented in MATLAB 2018a. The nonlinear optimization problems (4.32) or (4.33) were solved with the included nonlinear optimization solver, fmincon, using the sqp method with default settings. The gradients of the nonlinear equality constraints of (4.32) and (4.33) were explicitly computed and passed as an argument to fmincon. The sets E and F are stored as lists. More efficient data structures could be used here but in the numerical experiments it was observed that the bulk of the computational time (more that 90%) is used for the computation of $\Psi_d(B,k)$ and $\Psi_e(B,k,l)$.
For the first efficient complementary basis B_0 computed as the solution of (pLCP) for the initial weight $\lambda \in \Lambda$ we know a weight in $\Lambda^B(B_0)$. If an adjacent basis B' of B_0 is found then an optimal solution $\lambda' \in \Lambda^B(B)$ of (4.32) or (4.33) is also in $\Lambda^B(B')$ and can be used as an initial feasible solution for (4.32) or (4.33) when the adjacent bases of B' are considered.

4.3.2 Computational Experiment

The time required to compute all efficient complementary solutions depends on the dimensions of the problem (both m, n and $r = n + p$), the number of efficient complementary bases and the number of adjacent efficient complementary bases of each efficient complementary basis which indicates the number of pivots to be tested in each iteration.
The following numerical experiments aim to visualize the effect of the dimensions m, n, p on the time required to run Algorithm 4.2.
It should be noted at this point that the dimensions of the parametric linear complementarity problem (pLCP) only yield an upper bound for the number of efficient bases [13]. A given problem may only have one efficient complementary bases, even if the number of variables is large. For example, if the minimum of every objective function is feasible, then the multiobjective quadratic programming problem may only have one efficient complementary basis.

4.3.2.1 The Test Instance

There are collections of test instances available in the literature for singleobjective quadratic programming problems, for example the collection by Maros and Meszaros [61].

To test the performance of Algorithm 4.2 a set of random problem instances were computed with 2 to 6 objectives and 2 to 8 variables. For each combination of number of variables and objectives 10 distinct problems were generated.

The objective matrices Q^i are random symmetric positive definite matrices computed in the following way for $i = 1, \ldots, m$: Each objective matrix Q^i is computed by two regular matrices U and D with $Q^i = U^T DU$. The entries of matrix U are uniformly generated random integers from the interval $[-5, 5]$. D is a diagonal matrix with uniformly generated diagonal entries random integers from the interval $[1, 5]$.
The feasible sets are identical for each n and consist of the box $[0, 5] \times \cdots \times [0, 5]$ of the corresponding dimension. The linear independence constraint qualification is guaranteed to be satisfied for each feasible point. Thus, only diagonal pivots are considered for solving these problems.
The vectors c^i are computed in the following way for $i = 1, \ldots, m$:
For each objective a random integer vectors x^i with entries in the interval $[-1, 6]$ is computed. c^i is then set to $c^i = -Q^i x^i$. The resulting objective functions $f_i(x) = \frac{1}{2} x^T Q^i x + (c^i)^T x$ have their minimum in x^i for $i = 1, \ldots, m$
This choice results in problems where most of the lexicographic minima are infeasible.
The efficient complementary bases of the test problems were computed using an implementation of Algorithm 4.2 in MATLAB 2018a on an Intel i5-2400 processor with 3.2 GHz and 8 GB of RAM. The average number of efficient bases, average total time, average number of calls to the solver fmincon and average time per efficient bases are shown in Tables 4.1, 4.2, 4.3 and 4.4, respectively.

Table 4.1 shows that the number of efficient bases increases with the number of objectives and the number of variables. Due to the construction of the problems this is to be expected as the minima of the objective functions were placed mostly outside of the feasible set.

$m \backslash n$	2	3	4	5	6	7	8
2	2.2	3.3	3.8	6.1	5.3	6.3	8.5
3	4.6	5.5	7.7	11.6	18.1	18.7	21.4
4	4.5	8.5	11.2	14.4	23.2	31.2	57.2
5	5.2	10.5	16.5	25.9	40.8	56.6	73.9
6	6.6	11.4	22.1	30.8	62.1	88.6	115.6

Table 4.1: Average Number of Efficient Bases computed by Algorithm 4.2

The average time it took to compute the efficient complementary bases (cp. Table 4.2) is closely linked to the amount of solver calls shown in Table 4.3.

$m\backslash n$	2	3	4	5	6	7	8
2	0.045923	0.109	0.21662	0.41219	0.53495	1.0445	1.1453
3	0.063387	0.19475	0.55568	1.4651	2.7838	4.1779	5.1032
4	0.068963	0.27616	0.74217	2.0782	4.3808	8.3114	16.2376
5	0.069432	0.29767	1.1201	3.2825	7.1066	15.5916	26.1433
6	0.073817	0.37863	1.2704	3.3338	9.0906	22.253	38.3013

Table 4.2: Average Time (in seconds)

The number of adjacent bases increases exponentially with the number of complementary variables of (pLCP). Hence, it is to be expected that the number of solver calls grows faster with the number of variables than with the number of objectives.

$m\backslash n$	2	3	4	5	6	7	8
2	6.1	13.9	21.1	43	49.3	66.8	97.7
3	8.3	19.6	39	71.9	136.8	182.9	244
4	8.8	24.7	52.3	91.3	174.6	285.7	595.8
5	9.5	26.1	64.1	141.6	272.7	481.1	736.8
6	9.5	30.5	77.4	161.1	383.9	680.7	1079

Table 4.3: Average Number of Solver Calls

The average time per bases (cp. Table 4.4) increases faster with the number of objectives than with the number of variables.

$m\backslash n$	2	3	4	5	6	7	8
2	0.021868	0.033967	0.064266	0.080295	0.10287	0.20877	0.16251
3	0.014522	0.037155	0.077106	0.13496	0.19398	0.24732	0.26287
4	0.016266	0.035465	0.068675	0.14622	0.19361	0.27282	0.28902
5	0.013776	0.03072	0.072058	0.12913	0.18848	0.28149	0.36559
6	0.011999	0.035796	0.069126	0.11792	0.15443	0.25656	0.34586

Table 4.4: Average Time (in seconds) per Basis

4.3.3 Determination of Boundary Segments

Algorithm 4.2 did not consider which of the hypersurfaces

$$\mathfrak{h}_B^k := \{\lambda \in \mathbb{R}^m \ : \ (q_B(\lambda))_k = 0\}$$

for some $k \in \{1, \ldots, r\}$ actually is a $(m-2)$-dimensional boundary segment of $\Lambda^B(B)$ for some $B \in B_{\text{eff}}$.

In this section we will discuss the two-phase approach introduced by Adelgren [1] for parametric linear complementarity problems. The first phase consists of the determination of a complementary basis with a full dimensional, i.e. in the case of (pLCP) $(m-1)$-dimensional, weight cell. In the second phase all complementary basis with weight cells of dimension $(m-1)$ and $(m-2)$ are computed by identifying the $(m-2)$-dimensional boundaries of the weight cells.
For the complete representation of the efficient set of (pLCP) it is enough to only consider efficient complementary bases with $\dim(\Lambda^B(B)) = m-1$ [23].
Recall that the dimension of the parameter space considered here is $\dim(\Lambda) = m-1$. In order for the results from Adelgren [1] to apply directly we will assume that the matrix $\Delta q \in \mathbb{R}^{r \times m}$ defined such that

$$q(\lambda) = q_0 + \Delta q \lambda$$

has full rank. In the case that Δq does not have full rank Adelgren and Wiecek [2] provide a scheme to consider this case as well.
A smaller dimensional parametric linear complementarity problem is considered, from which the efficient complemenatry bases and the weight cells of the original problem can be computed.

Definition 4.39. [1] The **parametric complementary cone** of an efficient complementary basis B is defined as

$$C(B, \lambda) := \left\{a \in \mathbb{R}^r \ : \ M_B^{-1} a \geqq 0\right\}.$$

Using the parametric complementary cone of B we can see that [1]:

$$\lambda \in \Lambda^B(B) \Leftrightarrow q_B(\lambda) \in C(B, \lambda)$$

Adelgren [1] showed the following extension of Proposition 4.30:

Proposition 4.40. [1] Let B and B' be efficient complementary bases of (pLCP) with $\dim(\Lambda^B(B)) = \dim(\Lambda^B(B')) = m-1$ and $\dim(\Lambda^B(B) \cap \Lambda^B(B'))) = m-2$. Then there exists a sequence of efficient complementary bases $(B_j)_j$, $j = 1, \ldots, J$ with $\dim(\Lambda^B(B')) \geq m-2$ and $\dim(C(B_j, \lambda) \cap C(B_{j+1}, \lambda)) = m-2$ for $j = 1, \ldots, J-1$ and $B_1 = B$ and $B_J = B'$.

One of the consequences of Proposition 4.40 is that in order to compute all efficient bases B of (pLCP) with a full dimensional weight cell $\Lambda^B(B)$ it is enough to consider efficient complementary bases B with $\dim(\Lambda^B(B)) \geq m-2$.
Let B be an efficient complementary basis of (pLCP) with $\dim(\Lambda^B(B) = m-1)$. Recall the definition of the hypersurfaces $\mathfrak{h}_B^k$ for $k \in \{1, \ldots, r\}$:

$$\mathfrak{h}_B^k := \{\lambda \in \mathbb{R}^m \ : \ (q_B(\lambda))_k = 0\}$$

We will now consider which of the hypersurfaces $\mathfrak{h}_B^k$ for $k = 1, \ldots, r$ can not be $(m-2)$-dimensional boundary segments of $\Lambda^B(B)$. We can see that under the following conditions that $\mathfrak{h}_B^k$ cannot be a $(m-2)$-dimensional boundary segments of $\Lambda^B(B)$ for some $k \in \{1, \ldots, r\}$ [1]:

- If $(q_B(\lambda))_k$ is constantly zero on $\Lambda^B(B)$, then $\mathfrak{h}_B^k$ does not form a boundary segment of $\Lambda^B(B)$.
  We denote the corresponding set of indices by

$$\mathcal{Z}_B := \{k \in \{1, \ldots, r\} \ : \ (q_B(\lambda))_k = 0 \ \forall \lambda \in \Lambda\}.$$

- If $\mathfrak{h}_B^k$ for some $k \in \{1,\dots,r\}$ does not intersect $\Lambda^B(B)$, then $\mathfrak{h}_B^k$ does not form a boundary segment of $\Lambda^B(B)$.

  We denote the corresponding set of indices by

$$\mathcal{R}_B := \left\{k \in \{1,\dots,r\} \;:\; \mathfrak{h}_B^k \cap \Lambda^B(B) = \emptyset\right\}.$$

- Additionally, we do not consider hypersurfaces $\mathfrak{h}_B^k$ for which the intersection with $\Lambda^B(B)$ is contained in the intersection of $\Lambda^B(B)$ with another hypersurface, say hypersurface $\mathfrak{h}_B^l$.

  We denote the corresponding set of indices for $k = 1,\dots,r$ by

$$\mathcal{H}_B^k := \left\{l \in \{1,\dots,r\} \setminus (\{k\} \cup \mathcal{R}_B) \;:\; \left(\mathfrak{h}_B^k \cap \Lambda^B(B)\right) \subseteq \left(\mathfrak{h}_B^l \cap \Lambda^B(B)\right)\right\}.$$

The set $\mathcal{Z}_B$ can be determined easily [1], for example, by maximizing $(q_B(\lambda))_k$ over $\Lambda^B(B)$ for $k = 1,\dots,m$, or by using symbolic calculations. Adelgren [1] provides conditions to determine the sets $\mathcal{H}_B^k$ and $\mathcal{R}_B$. Consider the following nonlinear optimization problem for some $k, l \in \{1,\dots,r\}$, $k \neq l$:

$$\Psi_{\mathcal{H}}(B,k,l) := \begin{cases} \max\limits_{\lambda \in \mathbb{R}^m,\ \alpha \in \mathbb{R}} & \alpha \\ \text{s.t.} & (q_B(\lambda))_j \geq 0 \quad \forall j \in \{1,\dots,r\} \setminus (\{k,l\} \cup \mathcal{Z}_B) \\ & (q_B(\lambda))_k = 0 \\ & (q_B(\lambda))_l \geq \alpha \\ & \lambda \in \Lambda \end{cases} \tag{4.36}$$

Proposition 4.41. [1] Let B be an efficient complementary basis of (pLCP) with $\dim(\Lambda^B(B)) \geq m-2$ and $k, l \in \{1,\dots,r\}$ with $l \neq k$. Then $l \in \mathcal{H}_B^k$ if and only if the optimal objective value of the nonlinear optimization problem (4.36) is zero.

The nonlinear optimization problem (4.36) can also be used to compute the set $\mathcal{R}_B$:

Proposition 4.42. [1] Let B be an efficient complementary basis of (pLCP) with $\dim(\Lambda^B(B)) \geq m-2$ and $k \in 1,\dots,r$ with $l \neq k$. Then $k \in \mathcal{R}_B$ if and only if there exists $l \in \{1,\dots,r\} \setminus \{k\}$ such that (4.36) is infeasible or has a strictly negative objective value.

Now, the sets $\mathcal{Z}_B$, $\mathcal{R}_B$ and $\mathcal{H}_B^k$ for $k = 1,\dots,r$ can be determined for a given complementary basis B. To determine if an basis B has a full-dimensional weight cell $\Lambda^B(B)$ consider the following nonlinear optimization problem:

$$\Psi_{\text{full}}(B) := \begin{cases} \max\limits_{\lambda \in \mathbb{R}^m,\ \alpha \in \mathbb{R}} & \alpha \\ \text{s.t.} & (q_B(\lambda))_k \geq \alpha \quad \forall k \in \{1,\dots,r\} \setminus \mathcal{Z}_B \\ & \lambda \in \Lambda \end{cases} \tag{4.37}$$

Proposition 4.43. [1] Let B be an efficient complementary basis of (pLCP). Then $\dim(\Lambda^B(B)) = m-1$ if and only if $|\mathcal{Z}_B| \leq r-(m-1)$ and (4.37) has a strictly positive optimal objective value.

Notice that Proposition 4.43 does not provide a statement about the boundary segments of $\Lambda^B(B)$. Now, consider the following optimization problem for $k \in \{1,\dots,r\}$:

$$\Psi_{\text{bd}}(B,k) := \begin{cases} \max\limits_{\lambda \in \mathbb{R}^m,\ \alpha \in \mathbb{R}} & \alpha \\ \text{s.t.} & (q_B(\lambda))_l \geq \alpha \quad \forall l \in \{1,\dots,r\} \setminus \left(\{k\} \cup \mathcal{Z}_B \cup \mathcal{H}_B^k\right) \\ & (q_B(\lambda))_k = 0 \\ & \lambda \in \Lambda \end{cases} \tag{4.38}$$

Proposition 4.44. [1] Let B be an efficient complementary basis of (pLCP). If (4.38) has a strictly positive optimal objective value for some $k \in \{1,\dots,r\} \setminus (\mathcal{Z}_B \cup \mathcal{R}_B)$ then $\mathfrak{h}_B^k$ forms a $(m-2)$-dimensional boundary of $\Lambda^B(B)$.

Proposition 4.44 can also be extended for exchange pivots. Consider the following nonlinear optimization problem:

$$\Psi_{\text{exbd}}(B,k,l,B') := \begin{cases} \max\limits_{\lambda\in\mathbb{R}^m,\ \alpha\in\mathbb{R}} & \alpha \\ \text{s.t.} & (q_B(\lambda))_j \geq \alpha \quad \forall j \in \{1,\dots,r\} \setminus \left(\{k\} \cup \mathcal{Z}_B \cup \mathcal{H}_B^k\right) \\ & (q_{B'}(\lambda))_i \geq \alpha \quad \forall i \in \{1,\dots,r\} \setminus \left(\{l\} \cup \mathcal{Z}_{B'} \cup \mathcal{H}_{B'}^l\right) \\ & (q_B(\lambda))_k = 0 \\ & \lambda \in \Lambda \end{cases} \tag{4.39}$$

Proposition 4.45. [1] Let B be an efficient complementary basis of (pLCP) with $\dim(\Lambda^B(B)) = m-1$ and let $\mathfrak{h}_B^k$ be a $(m-2)$-dimensional boundary of $\Lambda^B(B)$. For any complementary basis B' with $B' \neq B$ and $|B \cap B'| \geq r-2$ $\Lambda^B(B)$ and $\Lambda^B(B')$ are adjacent along $\mathfrak{h}_B^k$ (i.e. $\Lambda^B(B) \cap \Lambda^B(B') \subseteq \mathfrak{h}_B^k$) if and only if one of the following conditions hold:

1. $B' = \text{diag}(B,k)$ and B' is a complementary basis.
2. $B' = \text{exch}(B,k,l)$ for some $l \in \{1,\dots,r\} \setminus \{k\}$, B' is a complementary basis and (4.39) has a strictly positive optimal objective value.

Similar results for efficient complementary bases B with $\dim(\Lambda^B(B)) = m-2$ can also be found in [1]. Given a efficient complementary basis B of (pLCP) with a full-dimensional weight cell $\Lambda^B(B)$ we can now follow the following steps to compute all $(m-2)$-dimensional boundary segments of $\Lambda^B(B)$:

1. Determine the sets $\mathcal{Z}_B$, $\mathcal{R}_B$ and $\mathcal{H}_B^k$ for $k = 1,\dots,r$.
2. Consider the set $\mathscr{K}_0 = \{1,\dots,r\} \setminus (\mathcal{Z}_B \cup \mathcal{R}_B \cup \left\{k \in \{1,\dots,r\} : \mathcal{H}_B^k \neq \emptyset\right\})$. For each $k \in \mathscr{K}_0$ determine if whether $\mathfrak{h}_B^k$ forms a $(m-2)$-dimensional boundary of $\Lambda^B(B)$ using Proposition 4.44 or Proposition 4.45.

Notice that multiple nonlinear (in general) nonconvex optimization problems have to be solved in order to determine the sets $\mathcal{Z}_B$, $\mathcal{R}_B$ and $\mathcal{H}_B^k$ and then to determine the $(m-2)$-dimensional boundary segments of $\Lambda^B(B)$. This is a considerably higher effort for each efficient complementary basis compared to Algorithm 4.2. Algorithm 4.2 solves only one nonlinear optimization problem for each possible adjacent basis.
For multiobjective optimization problems where every or most weight cell are full dimensional Algorithm 4.2 solves a considerably lower number of nonlinear optimization problems. One such instance is given in Example 4.47.
However, for problems with many lower dimensional weight cells $\Lambda^B(B)$ the approach by Adelgren [1] has some advantages. No efficient complementary bases B with $\dim(\Lambda^B(B)) < m-2$ are computed. The following example is a degenerate case and is an instance where the approach by Adelgren [1] has a huge advantage over Algorithm 4.2:

Example 4.46. Consider the following problem similar to Example 4.38:

$$\begin{array}{lll} \underset{x\in\mathbb{R}^n}{\text{vmin}} & f_i(x) = \frac{1}{2}x^T I_n x + (c^i)^T x & i = 1,\dots .m \\ \text{s.t.} & x \geqq 0 & \end{array} \tag{4.40}$$

with $c^1 = 0$ and $c^i = (1,\dots,1)^T$ for $i = 2,\dots,m$. Let $\mathscr{A} = \mathcal{J}$ be an active set of (4.40). Then the systems of equations

$$\begin{bmatrix} I_n & -(I_{\mathcal{J}\bullet})^T \\ -I_{\mathcal{J}\bullet} & 0 \end{bmatrix} \begin{pmatrix} x \\ y_{\mathcal{J}} \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \end{pmatrix} \tag{4.41}$$

is solved by x and $y_{\mathcal{J}}$ with

$$x_j = \begin{cases} 0 & \text{if } j \in \mathcal{J} \\ -\sum_{i=2}^m \lambda_i & \text{otherwise} \end{cases} \quad \text{for } j = 1,\dots,n, \quad \text{and } y_j = \sum_{i=2}^m \lambda_i \text{ for } j \in \mathcal{J}$$

Notice that for every $\mathcal{J} \subseteq \{1,\dots,n\}$ except for $\mathcal{J} = \emptyset$ that (4.41) has no solution $x \geqq 0$, $y_{\mathcal{J}} \geqq 0$ for any $\lambda \in \Lambda$ with $\lambda_i > 0$ for any $2 \leq i \leq m$. The efficient complementary basis corresponding to the active set $\mathscr{A} = \mathcal{J}$ with $\mathcal{J} = \emptyset$ is given by $B = (x_1,\dots,x_n)$. B is the only efficient complementary basis of (4.40) with an $(m-1)$-dimensional weight cell $\Lambda^B(B)$.

Example 4.46 illustrates a case where Algorithm 4.2 computes an exponential number of efficient complementary bases whereas the approach by Adelgren [1] would only compute one efficient complementary basis.
The following example has only $(m-1)$-dimensional weight cells. Thus, both approaches would compute the same efficient complementary bases.

Example 4.47. Consider the following triobjective convex quadratic programming problem:

$$\begin{array}{ll} \min & f_1(x) = \frac{1}{2}x^T \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} x + \begin{pmatrix} -40 \\ -24 \end{pmatrix}^T x \\ \min & f_2(x) = \frac{1}{2}x^T \begin{pmatrix} 8 & 2 \\ 2 & 16 \end{pmatrix} x + \begin{pmatrix} -34 \\ -24 \end{pmatrix}^T x \\ \min & f_3(x) = \frac{1}{2}x^T \begin{pmatrix} 6 & 3 \\ 3 & 12 \end{pmatrix} x + \begin{pmatrix} -27 \\ -3 \end{pmatrix}^T x \\ \text{s.t.} & \begin{pmatrix} -1 & -1 \\ 0 & -1 \\ -1 & 1 \end{pmatrix} x \geqq \begin{pmatrix} -9 \\ -4 \\ -5 \end{pmatrix} \\ & x \geqq 0 \end{array} \tag{4.42}$$

After substituting $\lambda_3 = 1 - \lambda_1 - \lambda_2$ we can write the parametric linear complementarity problem of (4.42) as

$$\begin{gathered} \begin{pmatrix} I_2 & 0 & -Q(\lambda) & A^T \\ 0 & I_3 & -A & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} c(\lambda) \\ -b \end{pmatrix} \\ u_k v_k = 0 \ \forall k = 1,\dots,5 \\ u \geqq 0,\ v \geqq 0 \end{gathered} \tag{4.43}$$

with $u = (y_1,\ y_2,\ s_1,\ s_2,\ s_3)^T$, $v = (x_1,\ x_2,\ \pi_1,\ \pi_2,\ \pi_3)^T$,

$$Q(\lambda) = \begin{pmatrix} 6 & 3 \\ 3 & 12 \end{pmatrix} + \lambda_1 \begin{pmatrix} -2 & -2 \\ -2 & -10 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 & -1 \\ -1 & 4 \end{pmatrix} \text{ and } c(\lambda) = \begin{pmatrix} -27 \\ -3 \end{pmatrix} + \lambda_1 \begin{pmatrix} -13 \\ -21 \end{pmatrix} + \lambda_2 \begin{pmatrix} -7 \\ -21 \end{pmatrix}.$$

Consider the complementary basis of (4.43) $B_1 = \{s_1,\ s_2,\ s_3,\ x_1,\ x_2\}$ and the corresponding matrix

$$M_{B_1}(\lambda) = \begin{pmatrix} 0 & 0 & 0 & 2\lambda_1 - 2\lambda_2 - 6 & 2\lambda_1 + \lambda_2 - 3 \\ 0 & 0 & 0 & 2\lambda_1 + \lambda_2 - 3 & 10\lambda_1 - 4\lambda_2 - 12 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & -1 \end{pmatrix}.$$

We can use symbolic calculations to compute the vector of basic values for the basis B:

$$q_{B_1}(\lambda) = \frac{1}{\det(M_{B_1}(\lambda))} \begin{pmatrix} 248\lambda_1^2 - 360\lambda_1\lambda_2 - 612\lambda_1 - 35\lambda_2^2 + 216\lambda_2 + 315 \\ 80\lambda_1^2 - 155\lambda_1\lambda_2 - 423\lambda_1 - 21\lambda_2^2 + 78\lambda_2 + 315 \\ 152\lambda_1^2 - 178\lambda_1\lambda_2 - 54\lambda_1 + 35\lambda_2^2 + 276\lambda_2 - 63 \\ -88\lambda_1^2 + 45\lambda_1\lambda_2 - 171\lambda_1 + 49\lambda_2^2 + 132\lambda_2 + 315 \\ -16\lambda_1^2 + 27\lambda_1\lambda_2 + 135\lambda_1 + 49\lambda_2^2 + 138\lambda_2 - 63 \end{pmatrix} \tag{4.44}$$

where $\det(M_{B_1}(\lambda)) = 16\lambda_1^2 - 32\lambda_1\lambda_2 - 72\lambda_1 + 7\lambda_2^2 + 54\lambda_2 + 63$.
The hypersurfaces $\mathfrak{h}_{B_1}^k$ that bound $\Lambda^B(B_1)$ for $k = 1, \ldots, 5$ are depicted in Figure 4.5.

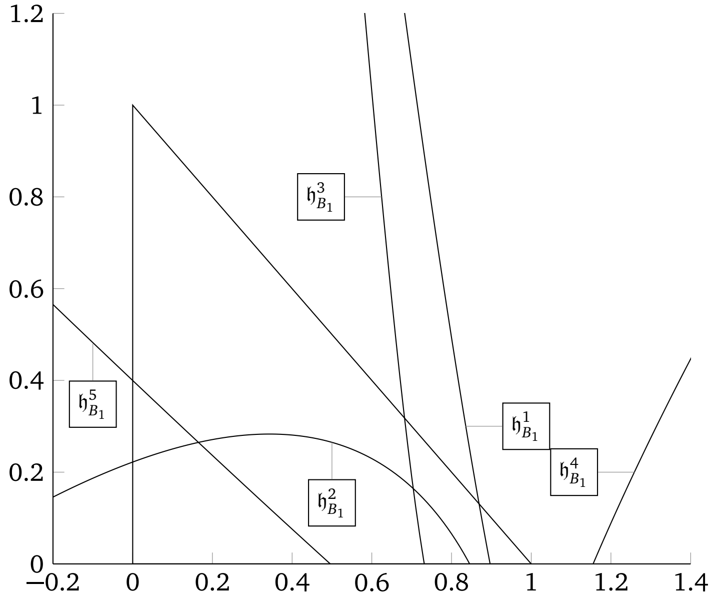

Figure 4.5: Hypersurfaces wrt. B_1

We can observe that the hypersurfaces $\mathfrak{h}_{B_1}^2$, $\mathfrak{h}_{B_1}^3$ and $\mathfrak{h}_{B_1}^5$ bound $\Lambda^B(B_1)$. Furthermore, $\mathfrak{h}_{B_1}^1$ and $\mathfrak{h}_{B_1}^4$ do not intersect $\Lambda^B(B_1)$.
The efficient complementary bases and the corresponding efficient active sets can be found in Table 4.5. The decomposition of the weight space and the efficient set are depicted in Figure 4.6.

Efficient complementary basis	Efficient Active Set	Color
$B_1 = (s_1, s_2, s_3, x_1, x_2)$	$(\emptyset, \emptyset)$	blue
$B_2 = (\pi_1, s_2, s_3, x_1, x_2)$	$(\emptyset, \{1\})$	magenta
$B_3 = (s_1, s_2, \pi_3, x_1, x_2)$	$(\emptyset, \{3\})$	green
$B_4 = (s_1, s_2, s_3, x_1, y_2)$	$(\{2\}, \emptyset)$	red
$B_5 = (\pi_1, s_2, \pi_3, x_1, x_2)$	$(\emptyset, \{1, 3\})$	cyan
$B_6 = (s_1, s_2, \pi_3, x_1, y_2)$	$(\{2\}, \{3\})$	yellow

Table 4.5: Efficient complementary bases of(4.43)

We will revisit this example in Chapter 5.

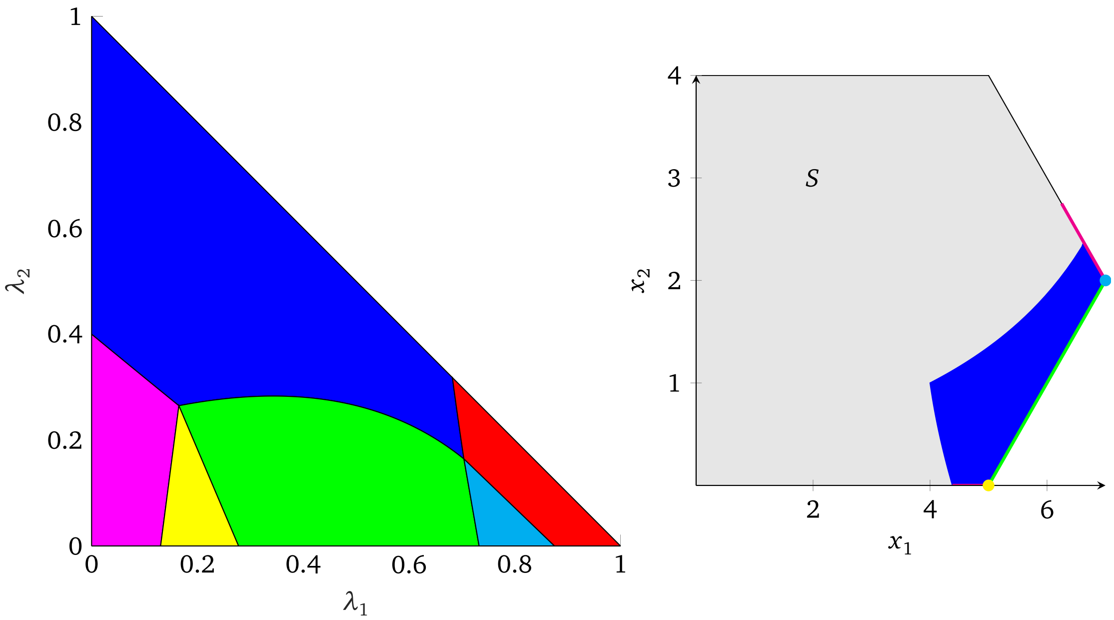

Figure 4.6: Decomposition of the weight space (left) and efficient set (right) in Example 4.47.

In conclusion, one should decide which approach to choose depending on the properties of the problem at hand. Of course, these properties are hard to check in advance.
It should be noted at this point that the boundary-determination approach by Adelgren [1] can be applied to a broader class of parametric linear complementarity problems.

4.4 Multiobjective Convex Quadratic Programming Problems in General Form

So far we have considered strictly convex multiobjective quadratic programming problems in canonical form with linear inequality constraints and nonnegative variables. In this section we will also allow unbounded variables and linear equality constraints. Many of the concepts discussed in Section 4.2 will be revisited here.

Consider a strictly convex multiobjective quadratic programming problem in general form:

$$\begin{aligned} \operatorname*{vmin}_{x\in\mathbb{R}^n} \quad & f_i(x) = \tfrac{1}{2}x^TQ^ix + (c^i)^Tx \quad i = 1,\dots,m \\ \text{s.t.} \quad & Ax \geqq b,\ Hx = h,\ x_{\mathcal{I}_+} \geqq 0 \end{aligned} \tag{gMQP}$$

with $\mathcal{I}_+ \subseteq \{1,\dots,n\}$, $A \in \mathbb{R}^{p\times n}$, $b \in \mathbb{R}^p$, $H \in \mathbb{R}^{q\times n}$, $h \in \mathbb{R}^q$ and $Q^i \in \mathbb{R}^{n\times n}$ symmetric positive definite for $i = 1,\dots,m$.

Additionally, we assume that the matrix of equality constraints H has full rank and that the feasible set $S := \left\{x \in \mathbb{R}^n \,:\, Ax \geqq b,\ Hx = h,\ x_{\mathcal{I}_+} \geqq 0\right\}$ is not empty.

The weighted sum problem of (gMQP) for $\lambda \in \Lambda$ is given by

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & \sum_{i=1}^m \lambda_i \left(\frac{1}{2}x^TQ^ix + (c^i)^Tx\right) \\ \text{s.t.} \quad & Ax \geqq b,\ Hx = h,\ x_{\mathcal{I}_+} \geqq 0 \end{aligned} \tag{gWQP}$$

In the same way as Theorem 4.3 we can formulate the following optimality condition:

Theorem 4.48 (Optimality Conditions for (gMQP)). [23] Let $x \in S$ be a regular feasible solution of (gMQP). Then x is efficient for (gMQP) if and only if there exists $\lambda \in \Lambda$, $\pi \in \mathbb{R}^p$, $y \in \mathbb{R}^{|\mathcal{I}_+|}$ and $\mu \in \mathbb{R}^q$ such that

$$\begin{aligned} Q(\lambda)x + c(\lambda) - A^T\pi - H^T\mu - I_{\bullet\mathcal{I}_+}y &= 0 \qquad (4.45) \\ Ax - b \geqq 0,\ x_{\mathcal{I}_+} \geqq 0,\ Hx &= h \\ \pi \geqq 0,\ y &\geqq 0 \\ x_iy_i &= 0\ \forall i \in \mathcal{I}_+ \\ (A_{j\bullet}x - b_j)\pi_j &= 0\ \forall j = 1,\dots,p \end{aligned}$$

Proof. Follows from Corollary 2.7 with Theorem 2.25 and Theorem 2.26 using the fact that the objective functions $f_i(x)$ are strictly convex for $i = 1,\dots,m$. □

Using additional variables $s = Ax - b$ we can write the system of equations (4.45) as a parametric system of equations and inequalities similar to a mixed linear complementarity problem [14]:

$$\begin{aligned} \underbrace{\begin{bmatrix} I_{\bullet\mathcal{I}_+} & 0 & -Q(\lambda) & A^T & H^T \\ 0 & I_p & -A & 0 & 0 \\ 0 & 0 & -H & 0 & 0 \end{bmatrix}}_{=:M(\lambda)} \begin{pmatrix} y \\ s \\ x \\ \pi \\ \mu \end{pmatrix} &= \underbrace{\begin{pmatrix} c(\lambda) \\ -b \\ -h \end{pmatrix}}_{=:q(\lambda)} \\ s \geqq 0,\ y \geqq 0,\ \pi \geqq 0,\ x_{\mathcal{I}_+} &\geqq 0 \\ x \in \mathbb{R}^n,\ \mu &\in \mathbb{R}^q \\ s_j\pi_j &= 0\ \forall j = 1,\dots,p \\ y_ix_i &= 0\ \forall i \in \mathcal{I}_+ \end{aligned} \tag{pmLCP}$$

The set of complementary variables Z_{comp} of (pmLCP) is given by

$$Z_{\text{comp}} = \{x_i \,:\, i \in \mathcal{I}_+\} \cup \{y_i \,:\, i \in \mathcal{I}_+\} \cup \{s_j \,:\, j = 1,\dots,p\} \cup \{\pi_j \,:\, j = 1,\dots,p\}$$

and the set of free variables Z_{free} of (pmLCP) is given by

$$Z_{\text{free}} = \{x_i \,:\, i \notin \mathcal{I}_+\} \cup \{\mu_j \,:\, j = 1,\dots,q\}.$$

The mixed linear complementarity problem (pmLCP) has in total $r := n + 2p + p_+ + q$ variables with $r_+ := p_+ + p$ complementary variable pairs and $r_0 = q + n - p_+$ free variables.

The following definition of a basis of (pmLCP) is slightly different from the definition usually found in the literature (see for example [14]) to emphasize the role of complementary variables. The free variables $x_{\bar{\mathcal{I}}_+}$ and μ have no constraints apart from the linear equality constraint in (pmLCP). Thus we include the free variables in every basis and only consider exchanging complementary variables:

Let $B \subset Z_{\text{comp}}$ be a complementary index set and let $\lambda \in \Lambda$. Let $M'_B(\lambda) \in \mathbb{R}^{r \times r_+}$ be the matrix of columns of $M(\lambda)$ corresponding to variables $B \subset Z_{\text{comp}}$ and let $M_{\text{free}}(\lambda) \in \mathbb{R}^{r \times r_0}$ be the matrix consisting of the columns of $M(\lambda)$ corresponding to free variables Z_{free}. Consider the following square matrix

$$M_B(\lambda) = \begin{bmatrix} M'_B(\lambda) & M_{\text{free}}(\lambda) \end{bmatrix}. \tag{4.46}$$

Definition 4.49. A complementary set of variables B is called a basis of (pmLCP) for $\lambda \in \Lambda$ if the matrix $M_B(\lambda)$ as defined in equation (4.46) is regular.

The basic values of a complementary basis B of (pmLCP) for $\lambda \in \Lambda$ are given by

$$q_B(\lambda) := (M_B(\lambda))^{-1} q(\lambda). \tag{4.47}$$

Note that the first r_+ entries of $q_B(\lambda)$ correspond to complementary variables where nonnegativity is required in (pmLCP).

Definition 4.50. Let B be a complementary basis of (pmLCP).

1. B is called feasible for $\lambda \in \Lambda$ if
$$q_B(\lambda)_k \geq 0 \;\forall k = 1,\dots,r_+.$$
2. B is called an efficient complementary basis of (pmLCP) if there exists $\lambda \in \Lambda$ with
$$q_B(\lambda)_k \geq 0 \;\forall k = 1,\dots,r_+.$$
3. The set
$$\Lambda^B(B) := \{\lambda \in \Lambda \,:\, q_B(\lambda)_k \geq 0 \;\forall k = 1,\dots,r_+\}$$
is called the weight cell of B.

Definition 4.51. The functions $(q_B(\lambda))_k : \mathbb{R}^m \to \mathbb{R}$, $k = 1,\dots,r_+$, are called **basic value functions** of (pmLCP).

The criss-cross method (Algorithm 2.1) can be applied to the mixed linear complementarity problem (pmLCP) for a fixed weight $\lambda \in \Lambda$ by keeping the free variables $x_i,\ i \notin I_+$ and μ in every basis and only performing pivots with the complementary variable pairs.
The concept of efficient active sets can also be extended from multiobjective quadratic programming problems in canonical form (MQP) to multiobjective quadratic programming problems in general (gMQP) in the following way:

Definition 4.52. A pair of index sets $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ with $\mathcal{I} \subseteq \mathcal{I}_+$ is called efficient active set of (gMQP), if there exists $\lambda \in \Lambda$ such that there exists a solution $(\bar{x}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}, \bar{\mu}) \in \mathbb{R}^{n \times |\mathcal{I}| \times |\mathcal{J}| \times q}$ of the active set system

$$\begin{bmatrix} Q(\lambda) & -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T & -H^T \\ -I_{\mathcal{I}\bullet} & 0 & 0 & 0 \\ -A_{\mathcal{J}\bullet} & 0 & 0 & 0 \\ -H & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \\ \mu \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \\ -b_{\mathcal{J}} \\ -h \end{pmatrix} \tag{4.48}$$

that satisfies $\bar{x} \in S$ and $(\bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}) \geqq 0$.
The weight cell corresponding to an efficient active set is defined as

$$\Lambda^A(\mathscr{A}) = \left\{\lambda \in \Lambda \ :\ \exists (\bar{x}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}, \mu) \in \mathbb{R}^{n \times |\mathcal{I}| \times |\mathcal{J}| \times q} \text{ solving (4.48) for } \lambda \text{ with } \bar{x} \in S \text{ and } (\bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}) \geqq 0\right\}.$$

The set of efficient active sets of (gMQP) is denoted by $\mathscr{A}_{\text{eff}}$.

We will refer to (4.48) as the active set system for an active set $\mathscr{A} = (\mathcal{I}, \mathcal{J})$.

Definition 4.53. An active set $\mathscr{A}$ of (gMQP) is called regular if the system of equations (4.48) has a unique solution for all $\lambda \in \Lambda$.

The following statements can be shown in an analogous way as for (MQP).

Proposition 4.54. [1] $\Lambda^B(B)$ is a semi-algebraic set for every efficient complementary basis B of (pmLCP).

Proof. The proof is analogous to the proof of Proposition 4.21. □

Proposition 4.55. If $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ is a regular efficient active set of (gMQP) then $B_{\mathscr{A}}$ is an efficient complementary basis of (pmLCP) and $\Lambda^A(\mathscr{A}) = \Lambda^B(B_{\mathscr{A}})$.

Proof. The proof is almost identical to the proof of Proposition 4.14 using the assumption that the matrix of equality constraints H is regular. □

Proposition 4.56. If B is an efficient complementary basis of (pmLCP), then $\mathscr{A}_B$ is is a regular efficient active set of (gMQP) and $\Lambda^B(B) = \Lambda^A(\mathscr{A}_B)$.

Proof. The proof is almost identical to the proof of Proposition 4.16 using the assumption that the matrix of equality constraints H is regular. □

Proposition 4.57. [38] Let B_{eff} be the set of efficient complementary bases of (gMQP) and let $\bar{x}(\lambda)$ be the optimal solution of (gWQP) for $\lambda \in \Lambda$. Let $\bar{X}(B)$ be defined as

$$\bar{X}(B) := \left\{\bar{x}(\lambda) \;:\; \lambda \in \Lambda^B(B)\right\}.$$

Then the following statement holds:

$$\bigcup_{B \in B_{\text{eff}}} \bar{X}(B) = X_E$$

Proof. A feasible complementary basis of (pmLCP) can be computed for every $\lambda \in \Lambda$. Thus,

$$\bigcup_{B \in B_{\text{eff}}} \Lambda^B(B) = \Lambda.$$

Applying Theorem 2.26 and 2.25 shows that every efficient point is an optimal solution of (gWQP) for some $\lambda \in \Lambda$. □

4.5 Regularization of Positive Semidefinite Objective Matrices

Throughout this chapter we have assumed that all objective matrices Q^i for $i = 1, \dots, m$ are symmetric positive definite and thus $Q(\lambda)$ is symmetric positive definite for every $\lambda \in \Lambda$. This assumption limits the applicability of the theory discussed in Chapter 4 to applications where the objective matrices may only guaranteed to be symmetric positive semidefinite.
In this section we will discuss possible difficulties that are caused by singular objective matrices as well as conditions under which a weight space decomposition is well defined and can be obtained.
Consider a multiobjective quadratic programming problem in general form:

$$\begin{array}{ll} \operatorname*{vmin}_{x \in \mathbb{R}^n} & x^T Q^i x + (c^i)^T x, \ i = 1, \dots, m \\ \text{s.t.} & Ax \geqq b, \ Hx = h, \ x_{\mathcal{I}_+} \geqq 0 \end{array} \tag{4.49}$$

with positive semidefinite matrices Q^i for $i = 1, \dots, m$.
The corresponding weighted sum problem of (4.49) is given by:

$$\begin{array}{ll} \operatorname*{vmin}_{x \in \mathbb{R}^n} & x^T Q(\lambda) x + c(\lambda)^T x \\ \text{s.t.} & Ax \geqq b, \ Hx = h, \ x_{\mathcal{I}_+} \geqq 0 \end{array} \tag{4.50}$$

The weighted sum problem (4.50) may be unbounded for some $\bar{\lambda} \in \Lambda$. In this case there exists at least one $i \in \{1, \dots, m\}$ such that the singleobjective problem

$$\begin{array}{ll} \operatorname*{vmin}_{x \in \mathbb{R}^n} & x^T Q^i x + (c^i)^T x \\ \text{s.t.} & Ax \geqq b, \ Hx = h, \ x_{\mathcal{I}_+} \geqq 0 \end{array} \tag{4.51}$$

is unbounded. In this case (pmLCP) is infeasible for $\bar{\lambda}$ and no active set can be assigned to the weight $\bar{\lambda}$.
Even if the weighted sum problem (4.50) is bounded for every $\lambda \in \Lambda$, the solution of (4.50) is not unique in general.
Since (4.50) does in general not have a unique solution for $\lambda \in \Lambda$, we can in general not define a weight space decomposition with non-overlapping weight cells. However, a special case of (4.49) is discussed in Section 4.6 where only one objective is strictly convex and the other objectives are affine functions .

4.5.1 Regularization of Objective Matrices

The regularization of convex optimization problems is a topic of ongoing research, see for example Friedlander and Tseng [33]. The following approach for regularization was discussed in the literature [38]. We will only briefly discuss how a multiobjective convex quadratic problem can be regularized and how the weight space decomposition is affected by the modification.
Let $\varepsilon > 0$ be a (possibly small) number. Consider the following modified weighted sum problem

$$\begin{array}{ll} \operatorname{vmin}_{x \in \mathbb{R}^n} & x^T (Q(\lambda) + \varepsilon I_n) x + c(\lambda)^T x \\ \text{s.t.} & Ax \geqq 0, \ Hx = h, \ x_{\mathcal{I}_+} \geqq 0 \end{array} \tag{4.52}$$

First we show that the objective matrix $Q(\lambda) + \varepsilon I_n$ is positive definite for all $\lambda \in \Lambda$:

Proposition 4.58 (Raleigh-Ritz)**.** [16] Let $Q \in \mathbb{R}^{n \times n}$. Then the smallest eigenvalue $e_{\min}$ of Q is given by

$$e_{\min} = \min_{x \in \mathbb{R}^n} \left\{ x^T Q x \ : \ \|x\|_2 = 1 \right\}.$$

Proposition 4.59. Let $Q \in \mathbb{R}^{n\times n}$ be positive semidefinite. Then the matrix $Q + \varepsilon I_n$ is positive definite for all $\varepsilon > 0$.

Proof. The smallest eigenvalue of $Q + \varepsilon I_n$ can be determined with Proposition 4.58:

$$\begin{aligned} e_{\min} &= \min_{x\in\mathbb{R}^n} \left\{ x^T(Q+\varepsilon I_n)x \;:\; \|x\|_2 = 1 \right\} \\ &= \min_{x\in\mathbb{R}^n} \left\{ \underbrace{x^T Q x}_{\geq 0} + \varepsilon \underbrace{x^T x}_{=1} \;:\; \|x\|_2 = 1 \right\} \\ &\geq \varepsilon > 0 \end{aligned}$$

which shows that $Q + \varepsilon I_n$ is positive definite. □

We can now modify the objective functions such that the weighted sum problem has a unique solution for every $\lambda \in \Lambda$ by adding the regularization term εI_n to every objective matrix. In the following the influence of this regularization is discussed:
Let $\bar{x}$ be an optimal solution of (4.49) for one $\lambda \in \Lambda$ with a corresponding active set $\mathscr{A} = (\mathcal{I}, \mathcal{J})$. Then there exists $\bar{y}_{\mathcal{I}} \geqq 0$, $\bar{\pi}_{\mathcal{J}} \geqq 0$ and $\bar{\mu} \in \mathbb{R}^q$ such that $(\bar{x}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}, \bar{\mu})$ is a solution of

$$\begin{bmatrix} Q(\lambda) & -I_{\mathcal{I}\bullet}^T & -A_{\mathcal{J}\bullet}^T & -H^T \\ -I_{\mathcal{I}\bullet}^T & 0 & 0 & 0 \\ -A_{\mathcal{J}\bullet} & 0 & 0 & 0 \\ -H & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \\ \mu \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \\ -b_{\mathcal{J}} \\ -h \end{pmatrix}. \tag{4.53}$$

Consider now the system of equations of the modified problem (4.52) for the same active set $\mathscr{A} = (\mathcal{I}, \mathcal{J})$:

$$\begin{bmatrix} Q(\lambda)+\varepsilon I_n & -I_{\mathcal{I}\bullet}^T & -A_{\mathcal{J}\bullet}^T & -H^T \\ -I_{\mathcal{I}\bullet}^T & 0 & 0 & 0 \\ -A_{\mathcal{J}\bullet} & 0 & 0 & 0 \\ -H & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \\ \mu \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \\ -b_{\mathcal{J}} \\ -h \end{pmatrix} \tag{4.54}$$

After inserting $(\bar{x}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}, \bar{\mu})$ into (4.54) we see that

$$\begin{aligned} &(Q(\lambda)+\varepsilon I_n)\bar{x} - (I_{\mathcal{I}\bullet})^T \bar{y}_{\mathcal{I}} - (A_{\mathcal{J}\bullet})^T \bar{\pi}_{\mathcal{J}} - H^T \mu + c(\lambda) \\ =&\underbrace{Q(\lambda)\bar{x} - (I_{\mathcal{I}\bullet})^T \bar{y}_{\mathcal{I}} - (A_{\mathcal{J}\bullet})^T \bar{\pi}_{\mathcal{J}} - H^T \mu}_{=-c(\lambda)} + c(\lambda) + \varepsilon \bar{x} \\ =&\varepsilon \bar{x} \xrightarrow{\varepsilon \to 0} 0 \end{aligned}$$

This shows that if the weighted sum problem (4.50) has a unique solution then any solution of the active set system (4.53) is an approximate solution for the active set system of the modified problem (4.52). Consequently, the weight space decomposition of (4.49) can be approximated by that of (4.52).

4.6 Multiobjective Mixed Linear and Convex Quadratic Problems

In this section we discuss a particular case of multiobjective convex quadratic programming problems where only one objective is a convex quadratic function and the other objectives are linear. Optimization problems of this type have been considered, for example, by Markowitz [59], Hirschberger et. al. [46], Columbano et. al. [13] and Adelgren and Wiecek [2].
We motivate this section with the following biobjective quadratic programming problem in the field of portfolio optimization proposed by Markowitz [59]: Given n possible assets with an expected return of μ_j $j = 1,\dots,n$ and the positive definite covariance matrix $Q \in \mathbb{R}^n$. The portfolio selection problem according to Markowitz [59] consists of two objectives and a budget constraint. The first objective is to minimize the risk and the second objective is to maximize the expected return. The portfolio selection problem can be formulated in the following way:

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{vmin}} \quad & \begin{pmatrix} x^TQx \\ -\mu^T x \end{pmatrix} \\ \text{s.t.} \quad & \sum_{j=1}^{n} x_j = 1 \\ & 0 \leq x_j \leq 1 \ \forall j = 1,\dots,n \end{aligned} \tag{4.55}$$

An interesting property shown by Markowitz [59] for the biobjective case and Hirschberger et. al. [46] in the triobjective case is that the weight cells for the particular portfolio selection problems discussed in the papers are connected intervals or convex polyhedrons, respectively.

We have assumed that every objective matrix Q^i is symmetric positive definite for multiobjective convex quadratic programming problems in general form (gMQP) which specifically excludes any linear objective functions. In this section we will investigate multiobjective convex programming problems with one strictly convex quadratic and $(m-1)$ linear objective functions.

Consider a multiobjective convex programming problem with one quadratic and $(m-1)$ linear objectives in general form:

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{vmin}} \quad & \begin{pmatrix} \frac{1}{2}x^TQ^1x + (c^1)^Tx \\ (c^2)^Tx \\ \vdots \\ (c^m)^Tx \end{pmatrix} \\ \text{s.t.} \quad & Ax \geqq b,\ Hx = h,\ x_{\mathcal{I}_+} \geqq 0 \end{aligned} \tag{MMLQP}$$

with $\mathcal{I}_+ \subseteq \{1,\dots,n\}$, $p_+ = |\mathcal{I}_+|$, $A \in \mathbb{R}^{p\times n}$, $b \in \mathbb{R}^p$, $H \in \mathbb{R}^{q\times n}$, $h \in \mathbb{R}^q$. Let $Q^1 \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix. We also assume that the feasible set

$$S := \left\{x \in \mathbb{R}^n \ : \ Ax \geqq b,\ Hx = h,\ x_{\mathcal{I}_+} \geqq 0\right\}$$

is not empty and that the matrix H is regular. Additionally, we assume that all singleobjective problems

$$\min_{x\in S} f_i(x) \tag{4.56}$$

are bounded for $i = 1,\dots,m$.
The weighted sum problem of (MMLQP) is given by

$$\begin{aligned} \min \quad & \lambda_1\left(\frac{1}{2}x^TQ^1x + (c^1)^Tx\right) + \sum_{i=2}^{m}\lambda_i(c^i)^Tx \\ \text{s.t.} \quad & Ax \geqq b,\ Hx = h,\ x_{\mathcal{I}_+} \geqq 0 \end{aligned} \tag{4.57}$$

Notice that the optimal solution of (4.57) for some $\lambda \in \Lambda$ with $\lambda_1 = 0$ is in general neither unique nor efficient. The corresponding linear programming problem is not strictly convex and in general the weighted sum problem (4.57) may have more than one optimal solution for $\lambda \in \Lambda$ with $\lambda_1 = 0$. However, we can still compute an efficient solution using a reoptimization routine [23]:

Let $\lambda_1 = 0$ and let $X_{\text{opt}}(\lambda)$ be the set of optimal solutions of (4.57). Since (4.57) is a linear programming problem $X_{\text{opt}}(\lambda)$ is a convex polyhedron. Consider the following quadratic optimization problem:

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f_1(x) \\ \text{s.t.} \quad & x \in S \\ & x \in X_{\text{opt}}(\lambda) \end{aligned} \tag{4.58}$$

Proposition 4.60. The optimal solution of (4.58) is efficient for (MMLQP).

Proof. The optimal solution x^* of (4.58) is unique since f_1 is strictly convex and the feasible set of (4.58) is convex. Assume there that exists $x' \in S$ such that $f(x')$ dominates $f(x^*)$. Then x' is also feasible for (4.58) with $f_1(x') \leq f_1(x^*)$ which contradicts the fact that x^* is the unique optimal solution of (4.58). □

Proposition 4.61. [44] $\Lambda^A(\mathscr{A})$ is a polyhedron for all efficient active sets $\mathscr{A} \in \mathscr{A}_{\text{eff}}$ of (MMLQP).

Proof. $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ is an efficient active set if there exists $\lambda \in \Lambda$ such that a solution $(\bar{x}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}}, \bar{\mu})$ of

$$\begin{bmatrix} \lambda_1 Q^1 & -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T & -H^T \\ -I_{\mathcal{I}\bullet} & 0 & 0 & 0 \\ -A_{\mathcal{J}\bullet} & 0 & 0 & 0 \\ -H & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \\ \mu \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \\ -b_{\mathcal{J}} \\ -h \end{pmatrix} \tag{4.59}$$

satisfies $\bar{x} \in S$, $\bar{y}_{\mathcal{I}} \geqq 0$ and $\bar{\pi}_{\mathcal{J}} \geqq 0$.
Recall the definition of the $\mathcal{W}$-parameterization of the weight space from Chapter 2:

$$\mathcal{W} = \{w \in \mathbb{R}^m \;:\; w \geqq 0,\; w_1 = 1\}$$

Consider the following system of equations for parameters $w \in \mathcal{W}$:

$$\begin{bmatrix} Q^1 & -(I_{\mathcal{I}\bullet})^T & -(A_{\mathcal{J}\bullet})^T & -H^T \\ -I_{\mathcal{I}\bullet} & 0 & 0 & 0 \\ -A_{\mathcal{J}\bullet} & 0 & 0 & 0 \\ -H & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \\ \mu \end{pmatrix} = \begin{pmatrix} -c^1 - \sum_{i=2}^m w_i c^i \\ 0 \\ -b_{\mathcal{J}} \\ -h \end{pmatrix} \tag{4.60}$$

Hirschnbergeret. al. [44] showed the following:

$$\mathcal{W}(\mathscr{A}) := \left\{ w \in \mathcal{W} \;:\; \text{there exists a solution } (\bar{x}', \bar{y}'_{\mathcal{I}}, \bar{\pi}'_{\mathcal{J}}, \mu') \text{ of (4.60) with } \bar{x}' \in S,\; \bar{y}'_{\mathcal{I}} \geqq 0 \text{ and } \bar{\pi}'_{\mathcal{J}} \geqq 0 \right\}$$

is a polyhedron since w is only found on the right hand side of (4.60).
Let $w \in \mathcal{W}(\mathscr{A})$ be given. Consider the following weight $\gamma(w) \in \Lambda$:

$$\gamma(w) := \frac{w}{1 + w_2 + \cdots + w_m}$$

First, we will show that $\gamma(w) \in \Lambda^A(\mathscr{A})$ for all $w \in \mathcal{W}(\mathscr{A})$. From $w \in \mathcal{W}(\mathscr{A})$ we know that there exists $\bar{x}' \in S$, $\bar{y}'_{\mathcal{I}} \geqq 0$, $\bar{\pi}'_{\mathcal{J}} \geqq 0$ and $\mu' \in \mathbb{R}^q$ solving (4.60). Moreover, any solution $(\bar{x}', \bar{y}_{\mathcal{I}}, \bar{\pi}'_{\mathcal{J}}, \bar{\mu}')$ of (4.60) satisfies:

$$Q^1 \bar{x}' - (I_{\mathcal{I}\bullet})^T \bar{y}' - (A_{\mathcal{J}\bullet})^T \bar{\pi}' - H^T \bar{\mu}' = -c^1 - \sum_{i=2}^m w_i c^i. \tag{4.61}$$

Then

$$\left(\bar{x}', \frac{\bar{y}'}{1+w_2+\ldots,w_m}, \frac{\bar{\pi}'}{1+w_2+\ldots,w_m}, \frac{\bar{\mu}'}{1+w_2+\ldots,w_m}\right)^T$$

is a solution of (4.59) for $\lambda = \gamma(w)$. This shows that

$$\gamma(w) \in \Lambda^A(\mathscr{A})\ \forall w \in \mathcal{W}(\mathscr{A}).$$

Since $\mathcal{W}(\mathscr{A})$ is a polyhedron we can describe $\mathcal{W}(\mathscr{A})$ by extreme points and extreme rays. Let E be the set of extreme points of $\mathcal{W}(\mathscr{A})$ and let R be the set of extreme rays of $\mathcal{W}(\mathscr{A})$. Let $\mathfrak{r} \in R$ and $w \in E$. Note that $\mathfrak{r}_1 = 0$ must hold for all $\mathfrak{r} \in R$. For $t \geq 0$ consider the vector $\gamma(w + t\mathfrak{r})$:

$$\gamma(w+t\mathfrak{r}) = \frac{w+t\mathfrak{r}}{1+(w_2+t\mathfrak{r}_2)+\ldots,(w_m+t\mathfrak{r}_m)}$$

Taking the limit we see that

$$\lim_{t\to\infty} \gamma(w+t\mathfrak{r}) = \frac{\mathfrak{r}}{\mathfrak{r}_2+\cdots+\mathfrak{r}_m} \in \Lambda$$

Since $\gamma(w+t\mathfrak{r}) \in \mathcal{W}(\mathscr{A})$ for all $t \geq 0$ we see that $\frac{\mathfrak{r}}{\mathfrak{r}_2+\cdots+\mathfrak{r}_m} \in \Lambda^A(\mathscr{A})$. Now consider the set

$$\Gamma := \mathrm{conv}\left(\{\gamma(\mathrm{w})\ :\ \mathrm{w} \in \mathrm{E}\} \cup \left\{\frac{\mathfrak{r}}{\mathfrak{r}_2+\ldots\mathfrak{r}_\mathrm{m}}\ :\ \mathfrak{r} \in \mathrm{R}\right\}\right).$$

So far we have shown that $\Gamma \subseteq \Lambda^A(\mathscr{A})$. Now, consider $\lambda \in \Lambda^A(\mathscr{A}) \cap \{\lambda \in \Lambda\ :\ \lambda_1 > 0\}$. Similar to the first part of this proof it is easy to see that

$$\frac{\lambda}{\lambda_1} \in \mathcal{W}(\mathscr{A}).$$

Also notice that

$$\gamma\left(\frac{\lambda}{\lambda_1}\right) = \frac{\frac{\lambda}{\lambda_1}}{1+\frac{\lambda_2}{\lambda_1}+\cdots+\frac{\lambda_m}{\lambda_1}} = \frac{\lambda}{\lambda_1+\cdots+\lambda_m} = \lambda$$

which shows that

$$\Lambda^A(\mathscr{A}) \cap \{\lambda \in \Lambda\ :\ \lambda_1 > 0\} \subseteq \Gamma.$$

$\Lambda^A(\mathscr{A})$ and Γ are closed sets. Thus

$$\mathrm{cl}(\Lambda^\mathrm{A}(\mathscr{A}) \cap \{\lambda \in \Lambda\ :\ \lambda_1 > 0\}) = \Lambda^\mathrm{A}(\mathscr{A}) \subseteq \mathrm{cl}(\Gamma) = \Gamma$$

which shows that

$$\Lambda^A(\mathscr{A}) = \Gamma.$$

□

4.6.1 Computation of Polyhedral Weight Cells

There are multiple efficient methods available in the literature to compute the weight space decomposition of (MMLQP), for example, a parametric procedure by Hirschberger et. al. [46] and the solution approaches for parametric linear complementarity problems by Columbano et. al. [13] and Adelgren and Wiecek [2].

4.7 Multiobjective Convex Quadratic Problems with Identical Objective Matrices

In this section we will discuss a special case of multiobjective convex quadratic optimization that is similar to the one discussed in Section 4.6.
Consider a multiobjective quadratic programming problem in general form

$$\begin{array}{ll} \operatorname*{vmin}_{x\in\mathbb{R}^n} & f_i(x) = \frac{1}{2}x^TQ^ix + (c^i)^Tx, \;\; i = 1,\dots,m \\ \text{s.t.} & Ax \geqq b, \; Hx = h, \; x_{\mathcal{I}_+} \geqq 0 \end{array} \tag{SMQP}$$

with $I_+ \subseteq \{1,\dots,n\}$, $p_+ = |\mathcal{I}_+|$, $A \in \mathbb{R}^{p\times n}$, $b \in \mathbb{R}^p$, $H \in \mathbb{R}^{q\times n}$, $h \in \mathbb{R}^q$ and $Q^i \in \mathbb{R}^{n\times n}$ symmetric positive definite for $i = 1,\dots,m$ and let there exist $\alpha \in \mathbb{R}^m$ with $\alpha > 0$ and $\alpha_i = 1$ such that

$$Q^i = \alpha_i Q^1 \text{ for } i = 2,\dots,m.$$

The feasible set is defined as

$$S = \left\{x \in \mathbb{R}^n \;:\; Ax \geqq b, \; Hx = h, \; x_{\mathcal{I}_+} \geqq 0\right\}$$

and we assume that $S \neq \emptyset$.

Proposition 4.62. $\Lambda^B(B)$ is a convex polyhedron for every efficient complementary basis $B \in B_{\text{eff}}$ of (SMQP).

Proof. Consider the following objective functions

$$\tilde{f}_i(x) = \begin{cases} \frac{1}{2}x^TQ^1x + (c^1)^Tx & \text{for i=1} \\ \frac{1}{2}x^TQ^1x + \frac{1}{\alpha_i}(c^i)^Tx & \text{for i=2,...,m} \end{cases}$$

and the following multiobjective quadratic optimization problem

$$\begin{array}{ll} \text{vmin} & \tilde{f}_i(x) = \frac{1}{2}x^TQ^1x + (\tilde{c}^i)^Tx, \;\; i = 1,\dots,m \\ \text{s.t.} & Ax \geqq b, \; Hx = h, \; x_{I_+} \geqq 0 \end{array} \tag{4.62}$$

with $\tilde{c}^i = \frac{1}{\alpha_i}(c^i)$ for $i = 1,\dots,m$. The multiobjective optimization problems (SMQP) and (4.62) have identical efficient sets, but the weighted sum problems of (SMQP) and (4.62) may have different optimal solutions for the same weights $\lambda \in \Lambda$.
The weighted sum problem of (4.62) for $\lambda \in \Lambda$ is given by

$$\begin{array}{ll} \operatorname*{vmin}_{x\in\mathbb{R}^n} & \sum_{i=1}^m \lambda_i \tilde{f}_i(x) = x^TQ^1x + \sum_{i=1}^m \frac{\lambda_i}{\alpha_i}(c^i)^Tx \\ \text{s.t.} & Ax \geqq b, \; Hx = h, \; x_{I_+} \geqq 0 \end{array} \tag{4.63}$$

with the objective matrix Q^1. Consider the parametric mixed linear complementary system of the KKT conditions of (4.63):

$$\begin{bmatrix} I_{\bullet\mathcal{I}_+} & 0 & -Q^1 & A^T & H^T \\ 0 & I_p & -A & 0 & 0 \\ 0 & 0 & -H & 0 & 0 \end{bmatrix} \begin{pmatrix} y \\ s \\ x \\ \pi \\ \mu \end{pmatrix} = \begin{pmatrix} c(\lambda) \\ -b \\ -h \end{pmatrix} \tag{4.64}$$

$$\begin{aligned} s \geqq 0, \; y \geqq 0, \; \pi \geqq 0, \; x_{\mathcal{I}_+} \geqq 0 \\ x \in \mathbb{R}^n, \; \mu \in \mathbb{R}^q \\ s_j\pi_j = 0 \;\forall j = 1,\dots,p \\ y_ix_i = 0 \;\forall i \in \mathcal{I}_+ \end{aligned}$$

Observe that the parameters λ are only present in the right-hand-side of the parametric mixed linear complementary system (4.64). Let B be an efficient complementary basis of (4.64). Recall the definition of the weight cell $\Lambda^B(B)$

$$\Lambda^B(B) = \{\lambda \in \Lambda \ : \ q_B(\lambda) \geqq 0\}$$

where for (4.64) the basic value vector $q_B(\lambda)$ is given by

$$q_B(\lambda) = M_B^{-1} q(\lambda) \tag{4.65}$$

where M_B does not contain any parameter. Then $\Lambda^B(B)$ is described by linear equality and inequality constraints. Hence, $\Lambda^B(B)$ is a convex polyhedron for all $B \in B_{\text{eff}}$. □

Objective functions based on the Euclidean norm can sometimes be written as quadratic functions. Such functions have, for example applications in location analysis, which is the subject of Section 4.10.

4.8 Multiobjective Convex Quadratic Programming Problems with Diagonal Objective Matrices and Box Constraints

Consider a multiobjective strictly convex quadratic problem with box constraints:

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{vmin}} & \frac{1}{2}x^T Q^i x + (c^i)^T x \;\; i = 1,\dots,m \\ \text{s.t.} & \boldsymbol{l} \leqq x \leqq \boldsymbol{u} \end{array} \tag{4.66}$$

with symmetric positive definite diagonal objective matrices $Q^i \in \mathbb{R}^{n\times n}$ and vectors $c^i \in \mathbb{R}^n$ for $i = 1,\dots,m$ and lower and upper bounds $\boldsymbol{l}$ and $\boldsymbol{u}$ satisfying $\boldsymbol{l}_j \in \mathbb{R}\cup\{-\infty\}$ and $\boldsymbol{u}_j \in \mathbb{R}\cup\{\infty\}$ for $j = 1,\dots,n$.

Let $\mathcal{U}$ be the index set of variables x_j with an upper bound, i.e.

$$\mathcal{U} = \left\{ j \in \{1,\dots,n\} \;:\; \boldsymbol{u}_j < \infty \right\}.$$

The problem (4.66) can be transformed such that all lower bounds $\boldsymbol{l}_j$ are either 0 or $-\infty$ by setting $x'_j = x_j - \boldsymbol{l}_j$ for $j \in \{1,\dots,n\} : l_j > -\infty$. For this reason we only consider problems where the lower bounds $\boldsymbol{l}_j$ are either $-\infty$ or 0 for $j = 1,\dots,n$.

Now, consider the following notation:

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{vmin}} & \frac{1}{2}x^T Q^i x + (c^i)^T x \;\; i = 1,\dots,m \\ \text{s.t.} & x_{\mathcal{J}_+} \geqq 0,\; x_{\mathcal{U}} \leqq \boldsymbol{u}_{\mathcal{U}} \end{array} \tag{MDQP}$$

with symmetric positive definite diagonal objective matrices $Q^i \in \mathbb{R}^{n\times n}$ and vectors $c^i\mathbb{R}^n$ for $i = 1,\dots,m$ and with index sets $\mathcal{J}_+ \subset \{1,\dots,n\}$ and $\mathcal{U} \subset \{1,\dots,n\}$. Additionally, we assume that if $j \in \mathcal{J}_+ \cap \mathcal{U}$, then $\boldsymbol{u}_j > 0$.

The weighted sum problem of (MDQP) is given by

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{vmin}} & \frac{1}{2}x^T Q(\lambda) x + c(\lambda)^T x \;\; i = 1,\dots,m \\ \text{s.t.} & x_{\mathcal{J}_+} \geqq 0,\; x_{\mathcal{U}} \leqq \boldsymbol{u}_{\mathcal{U}} \end{array} \tag{WDQP}$$

Proposition 4.63. [23] The feasible point $x \in S$ is an efficient solution of (MDQP) if and only if there exist y and $\pi \in \mathbb{R}^n$ such that:

$$Q(\lambda)x + c(\lambda) - (I_{\mathcal{J}_+\bullet})^T y + (I_{\mathcal{U}\bullet})^T \pi = 0 \tag{4.67}$$

$$0 \leqq x \leqq \boldsymbol{u}$$

$$y \geqq 0,\; \pi \geqq 0$$

$$x_j y_j = 0 \;\forall j \in \mathcal{J}_+ \tag{4.68}$$

$$(\bar{x}_j - \boldsymbol{u}_j)\pi_j = 0 \;\forall j \in \mathcal{U} \tag{4.69}$$

Proof. Follows from Corollary 2.7 with Theorem 2.25 and Theorem 2.26 using the fact that the objective functions $f_i(x)$ are strictly convex for $i = 1,\dots,m$. □

The solutions of (WDQP) for a given weight $\lambda \in \Lambda$ can be computed in $O(n+m)$ time:

Proposition 4.64. Let $\lambda \in \Lambda$ and let $\hat{x}$ be defined as

$$\hat{x} = -(Q(\lambda))^{-1} c(\lambda).$$

Then the optimal solution $\bar{x}(\lambda)$ of the weighted sum scalarization problem (WDQP) is given by:

$$\bar{x}(\lambda)_j = \begin{cases} 0 & \text{if } j \in \mathcal{I}_+ \text{ and } \hat{x}_j < 0 \\ \boldsymbol{u}_j & \text{if } j \in \mathcal{U} \text{ and } \hat{x}_j > \boldsymbol{u}_j \\ \hat{x}_j & \text{otherwise} \end{cases} \tag{4.70}$$

for $j = 1, \dots, n$.

Proof. Note that $\hat{x} = -(Q(\lambda))^{-1} c(\lambda)$ can be computed in $O(m+n)$ time as m additions of n entries is necessary to compute $Q(\lambda)$. Computing the inverse can be done in $O(n)$ time as $Q(\lambda)$ is a diagonal matrix for all $\lambda \in \Lambda$. The vector $\hat{x}$ is then computed by the multiplication of a diagonal matrix with a vector, which can be done in $O(n)$ time. Thus, the computation of $\hat{x}$ takes $O(n+m)$ time.

Let $\bar{x} = \bar{x}(\lambda)$ as defined in (4.70) for $\lambda \in \Lambda$. It is easy to see that $\bar{x}$ is feasible for (MDQP). Let

$$\begin{aligned} Q(\lambda)\bar{x} + c(\lambda) - (I_{\mathcal{I}_+\bullet})^T y + (I_{\mathcal{U}\bullet})^T \pi = 0 \\ y \geqq 0,\ \pi \geqq 0 \\ x_j y_j = 0\ \forall j \in \mathcal{I}_+ \\ (\bar{x}_j - \boldsymbol{u}_j)\pi_j = 0\ \forall j \in \mathcal{U} \end{aligned}$$

We partition the variable according to the bonding constraints. Let $\mathcal{I} \subset \mathcal{I}_+$ be given such that $\hat{x}_{\mathcal{I}} = 0$ and let $\mathcal{J} \subset \mathcal{U}$ be given such that $\hat{x}_{\mathcal{J}} = \boldsymbol{u}_{\mathcal{J}}$. Let the complements of $\mathcal{I}$ and $\mathcal{J}$ be denoted by $\bar{\mathcal{I}}$ and $\bar{\mathcal{J}}$.
We will now show that $\bar{x}(\lambda)$ as defined in (4.70) is an efficient solution of (MDQP) and optimal solution of (WDQP) for $\lambda \in \Lambda$ using Proposition 4.64.
Notice that since $Q(\lambda)$ is a diagonal matrix each variable x_j is only contained in the j-th row of equation (4.67) for $j = 1, \dots, n$:

$$Q(\lambda)\bar{x} + c(\lambda) - (I_{\mathcal{I}_+\bullet})^T y + (I_{\mathcal{U}\bullet})^T \pi = 0 \tag{4.71}$$

Notice that the complementarity conditions (4.68) and (4.69) demand that $y_{\bar{\mathcal{I}}} = 0$ and $\pi_{\bar{\mathcal{J}}} = 0$.

- For a variable x_j with $j \in \{1, \dots, n\} \setminus (\mathcal{I} \cup \mathcal{J})$ the corresponding line of equation (4.71) is given by

  $$\underbrace{Q_{jj}(\lambda)}_{>0} x_j + (c(\lambda))_j = 0$$

  which is equivalent to

  $$x_j = -\frac{1}{Q_{jj}(\lambda)}(c(\lambda))_j = \hat{x}_j.$$

- For $j \in \mathcal{I}$ first consider that

  $$\hat{x}_j = -\underbrace{\frac{1}{Q(\lambda)_{jj}}}_{>0} c(\lambda)_j < 0 \Leftrightarrow c(\lambda)_j > 0. \tag{4.72}$$

  Now consider the corresponding row of (4.71):

  $$Q(\lambda)_{jj} \underbrace{\bar{x}_j}_{=0} + c(\lambda)_j = y_j \Leftrightarrow c(\lambda)_j = y_j$$

  Using (4.72) we see that $y_j = c(\lambda)_j > 0$.

- Similarly, for $j \in \mathcal{J}$ we see that

$$\hat{x}_j = -\underbrace{\frac{1}{Q(\lambda)_{jj}}}_{>0} c(\lambda)_j > \boldsymbol{u}_j \Leftrightarrow -c(\lambda)_j > Q(\lambda)_{jj}\boldsymbol{u}_j, \tag{4.73}$$

$$Q(\lambda)_{jj}\bar{x}_j + c(\lambda)_j + \pi_j = 0 \text{ and}$$

$$\pi_j = -Q(\lambda)_{jj}\underbrace{\bar{x}_j}_{=\boldsymbol{u}_j} - c(\lambda)_j = -Q(\lambda)_{jj}\boldsymbol{u}_j - \underbrace{c(\lambda)_j}_{<Q(\lambda)_{jj}\boldsymbol{u}_j} > 0.$$

So in summary we found $y_{\mathcal{I}} \geq 0$ and$\pi_{\mathcal{J}} \geq 0$ such that $\bar{x}$ is a KKT point of (WDQP) for $\lambda \in \Lambda$. Using Proposition 4.64 we see that $\bar{x}$ is efficient for (MDQP).

□

4.8.1 Efficient Active Sets

Corollary 4.65. Every efficient active set of (MDQP) is regular.

Proof. The LICQ are satisfied at all feasible points. Thus, the conditions of Corollary 4.17 are satisfied.

□

In the following we will use a different notation: Let $\Delta Q \in \mathbb{R}^{n\times m}$ be the matrix comprising all diagonal elements of the matrices Q^i $i = 1,\dots,m$, such that $\Delta Q_{ji} = Q^i_{jj}$ for $i = 1,\dots,m$ and $j = 1,\dots,n$. Similarly, let $\Delta c \in \mathbb{R}^{n\times m}$ be such that the i-th row of Δc is equal to c^i for $i = 1\dots,m$. With this notation we can write matrix vector multiplications of the form $Q(\lambda)x$ as $(\Delta Qx)\lambda$ for $x \in \mathbb{R}^n$ and $\lambda \in \Lambda$.
Let $\Delta Q_{\mathcal{J}} * u_{\mathcal{J}}$ be defined as a kind of component-wise product of $\Delta Q_{\mathcal{J}}$ and $u_{\mathcal{J}}$:

$$(\Delta Q_{\mathcal{J}} * u_{\mathcal{J}})_{ij} := (\Delta Q_{\mathcal{J}})_{ij}(u_{\mathcal{J}})_i \tag{4.74}$$

Now we consider the active sets of (MDQP). Note that every active set can be thought of as a participation of the index set $\{1,\dots,n\}$ into three sets: Variables at their upper bound, variables at their lower bound and variable between both bounds. In terms of efficient active set the first two sets are denoted by $\mathcal{I} \subseteq \mathcal{I}_+$ and $\mathcal{J} \subseteq \mathcal{U}$.
Let the index set of the remaining variables be defined as

$$\mathcal{K} = \{1,\dots,n\} \setminus (\mathcal{I} \cup \mathcal{J}). \tag{4.75}$$

Theorem 4.66. Let $\mathscr{A} = (\mathcal{J},\mathcal{I})$ be an efficient active set of (MDQP). Then the weight cell $\Lambda(\mathscr{A})$ is a convex polyhedron.

Proof. Recall from Definition 4.52 that $\mathscr{A}$ is an efficient active set of (MDQP) if and only if there exists $\lambda \in \Lambda$ such that there exists a solution $(\bar{x}, \bar{y}_{\mathcal{I}}, \bar{\pi}_{\mathcal{J}})$ of

$$\begin{bmatrix} Q(\lambda) & -(I_{\mathcal{I}\bullet})^T & (A_{\mathcal{J}\bullet})^T \\ -I_{\mathcal{I}\bullet} & 0 & 0 \\ A_{\mathcal{J}\bullet} & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y_{\mathcal{I}} \\ \pi_{\mathcal{J}} \end{pmatrix} = \begin{pmatrix} -c(\lambda) \\ 0 \\ \boldsymbol{u}_{\mathcal{J}} \end{pmatrix} \tag{4.76}$$

satisfying $\bar{x} \in S$, $\bar{y}_{\mathcal{I}} \geqq 0$ and $\bar{\pi}_{\mathcal{J}} \geqq 0$.

Given the active set $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ and $\mathcal{K}$ defined as in (4.75) we can partition the first set of equations in (4.76) into three parts corresponding to $\mathcal{I}$, $\mathcal{J}$ and $\mathcal{K}$.
Let $J \subset 1,\dots,n$. $c(\lambda)_J$ denotes the vector with columns $c_j,\ j \in J$. Similarly, Q_{JJ} denotes submatrix of Q containing only entries of rows and columns j of Q in J.
Then we have that

$$Q(\lambda)x - (I_{\mathcal{I}\bullet})^T y_{\mathcal{I}} + (A_{\mathcal{J}\bullet})^T \pi_{\mathcal{J}} = -c(\lambda)$$

is equivalent to

$$Q_{\mathcal{I}\mathcal{I}} x_{\mathcal{I}} - y_{\mathcal{I}} = -c(\lambda)_{\mathcal{I}} \tag{4.77}$$
$$Q_{\mathcal{J}\mathcal{J}} x_{\mathcal{J}} + \pi_{\mathcal{J}} = -c(\lambda)_{\mathcal{J}} \tag{4.78}$$
$$Q_{\mathcal{K}\mathcal{K}} x_{\mathcal{K}} = -c(\lambda)_{\mathcal{K}} \tag{4.79}$$

- First consider equation (4.77):

$$Q_{\mathcal{I}\mathcal{I}} \underbrace{x_{\mathcal{I}}}_{=0} - y_{\mathcal{I}} = -c(\lambda)_{\mathcal{I}} \Leftrightarrow y_{\mathcal{J}} = c(\lambda)_{\mathcal{I}}$$

  The dual variables $y_{\mathcal{I}}$ have to be nonnegative:

$$y_{\mathcal{J}} = c(\lambda)_{\mathcal{I}} \geqq 0 \Leftrightarrow \Delta c_{\mathcal{I}} \lambda \geqq 0 \tag{4.80}$$

- Similarly for equation (4.78) we see that:

$$Q_{\mathcal{J}\mathcal{J}} \underbrace{x_{\mathcal{J}}}_{=\boldsymbol{u}_{\mathcal{J}}} + \pi_{\mathcal{J}} = -c(\lambda)_{\mathcal{J}} \Leftrightarrow \pi_{\mathcal{J}} = -Q_{\mathcal{J}\mathcal{J}} \boldsymbol{u}_{\mathcal{J}} - c(\lambda)_{\mathcal{J}}.$$

  The dual variables $\pi_{\mathcal{J}}$ have to be nonnegative:

$$\pi_{\mathcal{J}} = -Q_{\mathcal{J}\mathcal{J}} \boldsymbol{u}_{\mathcal{J}} - c(\lambda)_{\mathcal{J}} \geqq 0 \Leftrightarrow -(\Delta Q_{\mathcal{J}} * \boldsymbol{u}_{\mathcal{J}} + \Delta c_{\mathcal{J}})\lambda \geqq 0 \tag{4.81}$$

  Where $\Delta Q_{\mathcal{J}} * u_{\mathcal{J}}$ is defined as in (4.74).

- Now, consider equation (4.79):

$$Q_{\mathcal{K}\mathcal{K}} x_{\mathcal{K}} = -c(\lambda)_{\mathcal{K}} \Leftrightarrow x_{\mathcal{K}} - Q^{-1}_{\mathcal{K}\mathcal{K}} c(\lambda)_{\mathcal{K}}$$

  In order to be feasible $x_{\mathcal{K}}$ has to satisfy the constraints of (MDQP), i.e. $x_{\mathcal{K}\cap\mathcal{I}_+} \geqq 0$ and $x_{\mathcal{K}\cap\mathcal{U}} \leqq \boldsymbol{u}_{\mathcal{K}\cap\mathcal{U}}$. Notice that $Q^{-1}_{\mathcal{K}\cap\mathcal{I}_+,\mathcal{K}\cap\mathcal{I}_+} a$ is a rational function with constant numerator wrt. λ and positive denominator for every $a \in \mathbb{R}^{|\mathcal{K}\cap\mathcal{I}_+|}$. Thus,

$$\begin{aligned} x_{\mathcal{K}\cap\mathcal{I}_+} \geqq 0 &\Leftrightarrow -Q^{-1}_{\mathcal{K}\cap\mathcal{I}_+,\mathcal{K}\cap\mathcal{I}_+} c(\lambda)_{\mathcal{K}\cap\mathcal{I}_+} \geqq 0 \\ &\Leftrightarrow -c(\lambda)_{\mathcal{K}\cap\mathcal{I}_+} \geqq 0 \\ &\Leftrightarrow -\Delta c_{\mathcal{K}\cap\mathcal{I}_+} \lambda \geqq 0 \end{aligned} \tag{4.82}$$

  Now consider the upper bounds:

$$\begin{aligned} x_{\mathcal{K}\cap\mathcal{U}} \leqq \boldsymbol{u}_{\mathcal{K}\cap\mathcal{U}} &\Leftrightarrow -Q^{-1}_{\mathcal{K}\cap\mathcal{I}_+,\mathcal{K}\cap\mathcal{I}_+} c(\lambda)_{\mathcal{K}\cap\mathcal{I}_+} \leqq \boldsymbol{u}_{\mathcal{K}\cap\mathcal{U}} \\ &\Leftrightarrow (\Delta c_{\mathcal{K}\cap\mathcal{U}} + \Delta Q_{\mathcal{K}\cap\mathcal{U}} * \boldsymbol{u}_{\mathcal{K}\cap\mathcal{U}})\lambda \geqq 0 \end{aligned} \tag{4.83}$$

The inequalities (4.80), (4.81), (4.82) and (4.83) have to hold for every $\lambda \in \Lambda^A(\mathscr{A})$. Hence, the weight cell is given by:

$$\Lambda^A(\mathscr{A}) = \left\{ \lambda \in \Lambda \ : \ \begin{bmatrix} \Delta c_{\mathcal{I}} \\ -(\Delta Q_{\mathcal{J}} * \boldsymbol{u}_{\mathcal{J}} + \Delta c_{\mathcal{J}}) \\ -\Delta c_{\mathcal{K}\cap\mathcal{I}_+} \\ \Delta Q_{\mathcal{K}\cap\mathcal{U}} * \boldsymbol{u}_{\mathcal{K}\cap\mathcal{U}} + \Delta c_{\mathcal{K}\cap\mathcal{U}} \end{bmatrix} \lambda \geqq 0 \right\}$$

This shows that $\Lambda^A(\mathscr{A})$ is a convex polyhedron.

□

4.8.2 Arrangement of Hyperplanes

Let $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ be an active set and let $\mathcal{K}$ be defined as in (4.75), i.e. $\mathcal{K} = \{1, \dots, n\} \setminus (\mathcal{I} \cup \mathcal{J})$. The corresponding weight cell $\Lambda^A(\mathscr{A})$ is given by

$$\Lambda^A(\mathscr{A}) = \left\{ \lambda \in \Lambda \ : \ \begin{bmatrix} \Delta c_{\mathcal{I}} \\ -(\Delta Q_{\mathcal{J}} * \boldsymbol{u}_{\mathcal{J}} + \Delta c_{\mathcal{J}}) \\ -\Delta c_{\mathcal{K}\cap\mathcal{I}_+} \\ \Delta Q_{\mathcal{K}\cap\mathcal{U}} * \boldsymbol{u}_{\mathcal{K}\cap\mathcal{U}} + \Delta c_{\mathcal{K}\cap\mathcal{U}} \end{bmatrix} \lambda \geqq 0 \right\} \tag{4.84}$$

Let $j \in \mathcal{I}_+$ be given. Consider the following hyperplane in $\mathbb{R}^m$:

$$\boldsymbol{h}_j := \left\{ \lambda \in \mathbb{R}^m \ : \ \Delta c_j^T \lambda = 0 \right\}$$

and the half-spaces in $\mathbb{R}^m$:

$$\boldsymbol{h}_j^+ := \left\{ \lambda \in \mathbb{R}^m \ : \ \Delta c_j^T \lambda \geq 0 \right\} \text{ and } \boldsymbol{h}_j^- := \left\{ \lambda \in \mathbb{R}^m \ : \ \Delta c_j^T \lambda \leq 0 \right\}.$$

Let $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ be an efficient active set of (MDQP). If $j \in \mathcal{I}$ then $\Lambda^A(\mathscr{A}) \subseteq \boldsymbol{h}_j^+$ and if $j \notin \mathcal{I}$ then $\Lambda^A(\mathscr{A}) \subseteq \boldsymbol{h}_j^-$. So every efficient active set of (MDQP) is either a subset of $\boldsymbol{h}_j^+$ or $\boldsymbol{h}_j^-$ for all $j \in \mathcal{I}_+$.
Similarly, consider $j \in \mathcal{U}$ and the hyperplane $\boldsymbol{g}_j$ in $\mathbb{R}^m$

$$\boldsymbol{g}_j := \left\{ \lambda \in \mathbb{R}^m \ : \ (\Delta Q_j * \boldsymbol{u}_j + \Delta c_j)\lambda = 0 \right\}.$$

Also consider the corresponding half-spaces

$$\boldsymbol{g}_j^+ := \left\{ \lambda \in \mathbb{R}^m \ : \ (\Delta Q_j * \boldsymbol{u}_j + \Delta c_j)^T \lambda \geq 0 \right\} \text{ and } \boldsymbol{g}_j^- := \left\{ \lambda \in \mathbb{R}^m \ : \ (\Delta Q_j * \boldsymbol{u}_j + \Delta c_j)^T \lambda \leq 0 \right\}.$$

Notice that every efficient set is either a subset of $\boldsymbol{g}_j^+$ or $\boldsymbol{g}_j^-$ for every $j \in \mathcal{U}$.
Hence, the hyperplanes $\boldsymbol{h}_j$, $j \in \mathcal{I}_+$ and $\boldsymbol{g}_j$, $j \in \mathcal{U}$, decompose the weight space Λ into the weight cells $\Lambda^A(\mathscr{A})$. Such a decomposition is called an arrangement of hyperplanes [22].

Proposition 4.67. The weight space of (MDQP) is decomposed by an arrangement of at most $2n$ hyperplanes.

Schulze et. al. [79] showed that the weight space decomposition of the following multiobjective unconstrained combinatorial optimization problem is an arrangement of hyperplanes:

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & \sum_{j=1}^{n} P_{ij} x_j \qquad i = 1, \dots, m \\ & x_j \in \{0, 1\} \quad j = 1, \dots, n \end{aligned} \tag{MUCO}$$

with a cost matrix $P \in \mathbb{Z}^{m\times n}$. The weight space decomposition of (MUCO) given by an arrangement of the following hyperplanes [79]:

$$\hat{h}_j = \left\{\lambda \in \mathbb{R}^m \ : \ (P_{j\bullet})^T \lambda = 0\right\},\ j = 1,\dots,n.$$

The linear programming relaxation of (MUCO) was also discussed in Schulze et. al. [79] and shown to have the same weight space decomposition as (MUCO).

By using this result we can consider a multiobjective linear programming problem with a weight space decomposition that has the same arrangement of hyperplanes as (MDQP):

Theorem 4.68. The weight space decomposition of the following multiobjective linear programming problem is the same arrangement of hyperplanes as for the weight space decomposition of (MDQP):

$$\begin{array}{ll} \text{vmin} & \begin{bmatrix} \Delta c_{\mathcal{I}_+} \\ \Delta Q_{\mathcal{U}} * \boldsymbol{u}_{\mathcal{U}} + \Delta c_{\mathcal{U}} \end{bmatrix} (w, w')^T \\ \text{s.t.} & 0 \leqq w \leqq 1 \\ & 0 \leqq w' \leqq 1 \end{array} \tag{4.85}$$

Let (w, w') be an efficient basic solution of (4.85). Then the weight cell of (**??**) $\Lambda(w, w')$ corresponds to $\Lambda^A(\mathscr{A})$ for $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ with $\mathcal{I} = \left\{j \in \mathcal{I}_+ \ : \ w_j = 0\right\}$ and $\mathcal{J} = \left\{j \in \mathcal{U} \ : \ w'_j = 0\right\}$.
The fact that the weight space of (MDQP) is an arrangement of at most $2n$ hyperplanes yields an upper bound to the number of efficient active sets of (MDQP):

Lemma 4.69. [79, 12] The number of efficient active sets of (MDQP) is bounded by

$$\sum_{i=1}^{m} \binom{2n}{i}.$$

Proof. Buck [12] showed that the number of cells of an arrangement of $2n$ hyperplanes in $\mathbb{R}^{m-1}$ is bounded by

$$\sum_{i=1}^{m} \binom{2n}{i}.$$

□

Example 4.70. Consider the following instance of (MDQP) with $n = 8$ variables from the unit box and $m = 3$ objectives given by ΔQ and Δc:

$$\Delta Q = \begin{pmatrix} 3 & 1 & 2 \\ 4 & 1 & 1 \\ 2 & 1 & 4 \\ 2 & 3 & 3 \\ 4 & 3 & 4 \\ 3 & 2 & 1 \\ 3 & 2 & 3 \\ 2 & 3 & 3 \end{pmatrix}, \ \Delta c = \begin{pmatrix} 2 & 9 & -9 \\ 8 & -10 & 6 \\ -8 & -7 & 10 \\ 5 & -4 & -2 \\ 10 & 6 & 7 \\ -7 & -5 & 3 \\ 0 & -10 & 5 \\ -2 & -7 & -2 \end{pmatrix}$$

The weight space decomposition was computed using the multiobjective linear programming problem (4.85) and the multiobjective linear programming solver bensolve [58] and can be seen in Figure 4.7.

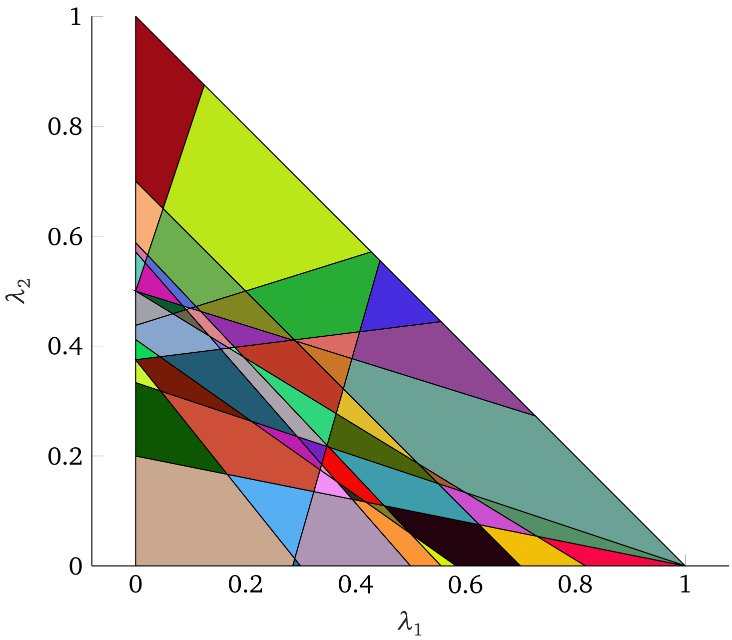

Figure 4.7: Weight Space Decomposition for Example 4.70

4.9 The e-constraint Scalarization for Multiobjective Convex Quadratic Programming Problems

In this section we will discuss a useful connection between the weight space decomposition of (gMQP) and another scalarization method: the e-constraint scalarization.

Recall the definition of a multiobjective convex quadratic programming problem in general form:

$$\begin{array}{lll} \operatorname*{vmin}_{x\in\mathbb{R}^n} & f_i(x) = \frac{1}{2}x^T Q^i x + (c^i)^T x & i = 1,\dots,m \\ \text{s.t.} & Ax \geqq b,\ Hx = h,\ x_{\mathcal{J}_+} \geqq 0 & \end{array} \tag{gMQP}$$

with a nonempty feasible set

$$S = \left\{x \in \mathbb{R}^n \ : \ Ax \geqq b,\ Hx = h,\ x_{\mathcal{J}_+} \geqq 0\right\} \neq \emptyset.$$

Consider a multiobjective strictly convex quadratic programming problem (gMQP). Let $\mathfrak{e} \in \mathbb{R}^{m-1}$ be a real vector. Then the e-constraint problem of (gMQP) is given by:

$$\begin{array}{lll} \min_{x\in\mathbb{R}^n} & f_m(x) & \\ \text{s.t.} & f_i(x) \leq \mathfrak{e}_i & \forall\ i = 1,\dots,m-1 \\ & Ax \geqq b,\ Hx = h,\ x_{\mathcal{J}_+}, \geqq 0 & \end{array} \tag{4.86}$$

Any of the objective function $f_i(x)$ of (gMQP) can be chosen as the objective function of (4.86). In comparison to the weighted sum scalarization not all parameter $\mathfrak{e} \in \mathbb{R}^{m-1}$ lead to a nonempty feasible set.

Proposition 4.71. Let $\mathfrak{e} \in \mathbb{R}^{m-1}$ be a set of parameters. The e-constraint scalarization problem (4.86) is either infeasible or has a unique solution.

Proof. The level sets $\mathcal{L}(f_i, \mathfrak{e}_i)$ are convex compact sets in $\mathbb{R}^n$ for every $i = 1, \dots, m-1$ and S is a convex polyhedron. Hence, the feasible set of (4.86) is convex and compact. The objective function $f_m(x)$ is continuous and attains it's minimum on the feasible set of (4.86).
Additionally, since $f_m(x)$ is strictly convex and the feasible set of (4.86) is convex the optimal solution of (4.86) is unique. □

We summirze a classical result [23] about the e-constraint scalarization:

Proposition 4.72. [23] A feasible point $x \in S$ is an efficient solution of (gMQP) if and only if x is an optimal solution of (4.86) for $\mathfrak{e} = (f_1(x), \dots, f_{m-1}(x))$.

Proof.

- First, consider an efficient solution $\bar{x} \in S$ of (gMQP). The e-constraint scalarization problem (4.86) is then feasible for $\mathfrak{e} = (f_1(\bar{x}), \dots, f_m(\bar{x}))$ and has a unique optimal solution $\tilde{x} \in S$. Assume that $\tilde{x} \neq \bar{x}$. Since is a unique optimal solution of (4.86) it follows that $f_m(\tilde{x}) < f_m(\bar{x})$ which implies that $f(\tilde{x})$ weakly dominates $f(\bar{x})$. Thus $\bar{x}$ is not efficient.
- Now let $x \in S$ be the optimal solution of (4.86) for $\mathfrak{e} = (f_1(x), \dots, f_{m-1}(x))$. Assume there exists $x' \in S$ such that $f(x')$ dominates $f(x)$. Notice that x' is then feasible for (4.86) with $\mathfrak{e} = (f_1(x), \dots, f_{m-1}(x))$. x' is also an optimal solution of (4.86) since $f_m(x') \leq f_m(x)$ which contradicts the uniqueness of optimal solutions of (4.86).

□

The e-constraint scalarization problem (4.86) has a compact convex feasible set and a strictly convex objective function. For this reason the KKT conditions are necessary and sufficient for optimal solutions of (4.86) under appropriate regularity assumptions. We can show a result analogous to Theorem 4.48 for the weighted sum problem:

Proposition 4.73. [23] Let $\bar{x}$ be a regular point of (4.86) wrt. the constraints of (4.86). $\bar{x}$ is an optimal solution of (4.86) with parameters $\mathfrak{e} \in \mathbb{R}^{m-1}$ if and only if there exists $w \in \mathbb{R}^{m-1}$, $\pi \in \mathbb{R}^p$, $y \in \mathbb{R}^{p_+}$ and $\mu \in \mathbb{R}^q$ such that

$$\left(\sum_{i=1}^{m-1} w_i(Q^i x + c^i)\right) + Q^m x + c^m - A^T\pi - H^T\mu - I_{\bullet\mathcal{J}_+} y = 0 \tag{4.87}$$

$$\begin{aligned} Ax - b \geqq 0,\ x_{\mathcal{J}_+} \geqq 0,\ Hx = h \\ f_i(x) - \mathfrak{e}_i \leq 0\ \forall i = 1, \dots, m-1 \\ \pi \geqq 0,\ y \geqq 0,\ w \geqq 0 \\ x_i y_i = 0\ \forall i \in \mathcal{J}_+ \\ (A_{j\bullet}x - b_j)\pi_j = 0\ \forall j = 1, \dots, p \\ (f_i(x) - \mathfrak{e}_i)w_i = 0\ \forall i = 1, \dots, m-1 \end{aligned}$$

4.9.1 Parameter Space Decomposition for e-constraint Scalarization Problems

The following definition of a minimal complete parameter set is an extension of the definition by Hansen [42]:

Definition 4.74. A set $\mathfrak{E} \subseteq \mathbb{R}^{m-1}$ is called **minimal complete representation** of the efficient set of (gMQP) if for all $x \in X_E$ there exists exactly one $\mathfrak{e} \in \mathfrak{E}$ such that x is the unique optimal solution of the e-constraint scalarization problem (4.86) for parameters $\mathfrak{e}$.

For example, the set

$$\mathfrak{E} = \{(f_1(x), \ldots, f_{m-1}(x)) \;:\; x \in X_E)\} \tag{4.88}$$

is a minimal complete representation if all objective functions are strictly convex. Note that for every $\mathfrak{e} \in \mathfrak{E}$ the optimal solution x of (4.86) satisfies $f_i(x) - \mathfrak{e}_i = 0$ for all $i = 1, \ldots, m-1$.
Similar to the decomposition of the weight space Λ we can define a decomposition of the parameter space for the e-constraint scalarization:

Definition 4.75. Let $\mathscr{A} = (\mathfrak{I}, \mathfrak{J})$ be an efficient active set of (gMQP). Let $\bar{x}(\lambda)$ be the optimal solution of the weighted sum problem of (gMQP) for $\lambda \in \Lambda$. Then the e-constraint parameter set corresponding to $\mathscr{A}$ is defined as:

$$\mathfrak{E}(\mathscr{A}) := \left\{(f_1(\bar{x}(\lambda)), \ldots, f_{m-1}(\bar{x}(\lambda))) \;:\; \lambda \in \Lambda^A(\mathscr{A})\right\}$$

Proposition 4.76. The set

$$\bigcup_{\mathscr{A} \in \mathscr{A}_{\text{eff}}} \mathfrak{E}(\mathscr{A})$$

is a complete representation of the efficient set of (4.86).

Proof. From Proposition 4.57 we know that

$$X_E = \bigcup_{\mathscr{A} \in \mathscr{A}_{\text{eff}}} \bar{X}(\mathscr{A}).$$

and with Proposition 4.72 it follows that

$$\bigcup_{\mathscr{A} \in \mathscr{A}_{\text{eff}}} \mathfrak{E}(\mathscr{A})$$

is a complete representation of X_E. □

The e-constraint parameter space decomposition for Example4.70 is shown in Figure 4.8.

Corollary 4.77. If $\lambda \in \Lambda^A(\mathscr{A}) \cap \Lambda^A(\mathscr{A}')$ then $\mathfrak{e} = (f_1(\bar{x}(\lambda)), \ldots, f_{m-1}(\bar{x}(\lambda)))$ satisfies $\mathfrak{e} \in \mathfrak{E}(\mathscr{A}) \cap \mathfrak{E}(\mathscr{A}')$.

Corollary 4.77 can be used to show that the parameter space decomposition for the e-constraint scalarization problem and the weighted sum scalarization problems share some properties:

Corollary 4.78. Let $\mathscr{A}_{\text{eff}}$ be the set of efficient active sets of (gMQP). Then the following statements hold:

1. $\operatorname{int} \mathfrak{E}(\mathscr{A}) \cap \operatorname{int} \mathfrak{E}(\mathscr{A}') = \emptyset$ for all regular active sets $\mathscr{A} \neq \mathscr{A}'$
2. $\mathfrak{E}(\mathscr{A})$ is a connected set if and only if $\Lambda^A(\mathscr{A})$ is connected.

Proof.

1. Follows directly from Corollary 4.77
2. Follows from the fact that the efficient set is connected and that all objective functions are continuous.

□

In this section we have seen that the weight space decomposition of (gMQP) can be used to compute a parameter space decomposition for the e-constraint problem as well. We will apply this result in a particular context that arises in a problem from the field of location theory in Section 4.10.

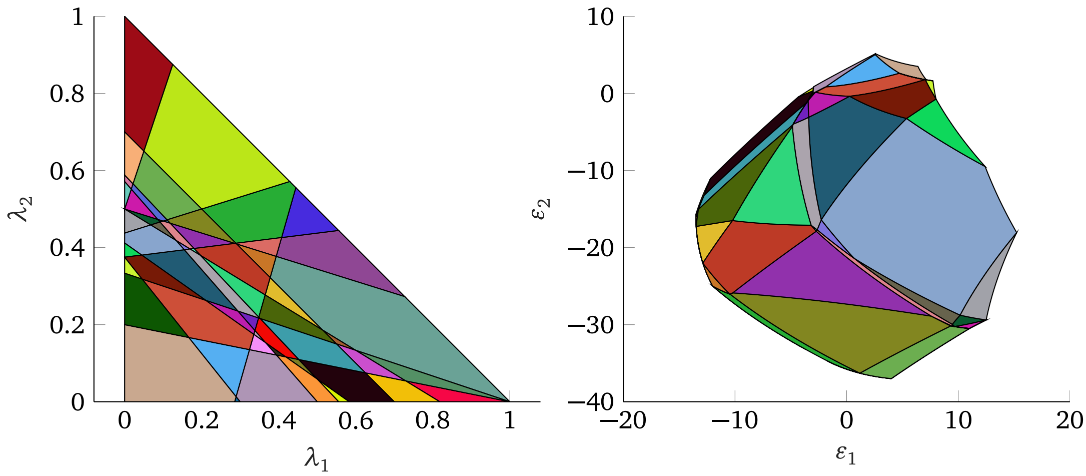

Figure 4.8: Weight Space Decomposition and e-constraint Parameter Decomposition for Example 4.70

4.10 Applications in Location Analysis

In this section, we consider an application of the results from Section 4.7 to a problem from the field of location theory. For a general introduction to location theory we refer to the books by Nickel and Puerto [64], Laporte et. al. [55] and Drezner and Hamacher [21].
The aim of this section is to provide a decomposition of the e-constraint parameter space for multiobjective constrained location problems with l_2^2 and l_2 norms, respectively.

Consider, for example, the placement of a new warehouse from which a number of customers have to be supplied. We are interested in a location for the new warehouse that minimizes the distance from the warehouse to the customers simultaneously. In general, there does not exist a location that minimizes the distance to all customers. Thus, a decision about the placement of the new warehouse has to take into account that the objectives, i.e. minimizing the distances to each customer, are conflicting objectives. Multiobjective location problems have been studied in the literature, see for example Wendell and Hurter [81], Juel and Love [49] and Pelegrin and Fernandez [69].
For this section we consider a multiobjective multidimensional location problem: Given m existing locations $a^i \in \mathbb{R}^n$, $i = 1, \dots, m$, we want to find a new location $x \in \mathbb{R}^n$ such that the distances between x and all existing locations are minimized simultaneously. For the distance measures we consider the l_2^2 and l_2 norms. Consider the following multiobjective optimization problems:

$$\operatorname*{vmin}_{x\in\mathbb{R}^n} \quad f_i(x) = \|x - a^i\|_2^2 \quad i = 1, \dots, m \tag{4.89}$$

$$\operatorname*{vmin}_{x\in\mathbb{R}^n} \quad g_i(x) = \|x - a^i\|_2 \quad i = 1, \dots, m \tag{4.90}$$

Francis and Cabot [32] show the following result:

Proposition 4.79. [32] The efficient sets of (4.89) and (4.90) are given by $\operatorname{conv}\left(\left\{\mathrm{a}^1, \dots, \mathrm{a}^\mathrm{m}\right\}\right)$.

We will now consider the constrained case. Let $S \subseteq \mathbb{R}^n$ be given by linear equality and inequality constraints as defined for (gMQP), i.e.

$$S = \left\{x \in \mathbb{R}^n \ : \ Ax \geqq 0, x_{\mathcal{I}_+} \geqq 0, Hx = h\right\} \neq \emptyset.$$

Consider the constrained multiobjective location problem with l_2^2 norm:

$$\begin{array}{lll} \operatorname*{vmin}_{x\in\mathbb{R}^n} & f_i(x) = \|x-a^i\|_2^2 & i=1,\dots,m \\ \text{s.t.} & x \in S & \end{array} \tag{4.91}$$

Notice that the objective functions of (4.89) are quadratic functions for each $i = 1,\dots,m$, since

$$f_i(x) = \|x-a^i\|_2^2 = (x-a^i)^T(x-a^i) = x^T I_n x - 2(a^i)^T x + (a^i)^T a^i \quad \text{for } i=1,\dots,m.$$

Thus, we can apply the results from Section 4.4 and Section 4.1.1 to (4.91). In particular, the weight cells $\Lambda(\mathscr{A})$ are convex polyhedra for all efficient active sets $\mathscr{A}$ of (4.91) as shown in Proposition 4.62. For applications it may be more relevant to consider the Euclidean norm and the following multiobjective optimization problem:

$$\begin{array}{lll} \operatorname*{vmin}_{x\in\mathbb{R}^n} & g_i(x) = \|x-a^i\|_2 & i=1,\dots,m \\ \text{s.t.} & x \in S & \end{array} \tag{4.92}$$

In the following we will now extend Proposition 4.79 to the constrained case:
Recall the definition of level sets $\mathcal{L}(f,y,S)$ and level curves $\mathcal{L}_=(f,y,S)$ from Definition 2.29 for $y \in \mathbb{R}^m$:

$$\mathcal{L}(f,y,S) = \{x \in S\ : f_i(x) \le y_i\ \forall i=1,\dots,m\}$$
$$\mathcal{L}_=(f,y,S) = \{x \in S\ : f_i(x) = y_i\ \forall i=1,\dots,m\}$$

Notice that for the functions f and g as defined in (4.91) and (4.92) we can observe the following for $y \in \mathbb{R}$, $y \ge 0$:

$$\begin{aligned} \mathcal{L}(f_i,y,S) &= \left\{x \in S\ :\ \|x-a^i\|_2^2 \le y\right\} \\ &= \left\{x \in S\ :\ \|x-a^i\|_2 \le \sqrt{y}\right\} \\ &= \mathcal{L}(g_i,\sqrt{y},S)\ \forall i=1\dots,m \end{aligned} \tag{4.93}$$

$$\begin{aligned} \mathcal{L}(g_i,y,S) &= \left\{x \in S\ :\ \|x-a^i\|_2 \le y\right\} \\ &= \left\{x \in S\ :\ \|x-a^i\|_2^2 \le y^2\right\} \\ &= \mathcal{L}(f_i,y^2,S)\ \forall i=1\dots,m \end{aligned}$$

Similarly, we can show the same for the level curves for every $i = 1,\dots,m$:

$$\mathcal{L}_=(f_i,y,S) = \mathcal{L}_=^S(g_i,\sqrt{y},S)$$
$$\mathcal{L}_=(g_i,y,S) = \mathcal{L}_=^S(f_i,y^2,S)$$

In Theorem 2.30 it was shown that efficient solutions can be characterized by level sets and level curves. For the efficient solutions of (4.91) and (4.92) we show the following result:

Proposition 4.80. The efficient sets of (4.91) and (4.92) are identical.

Proof. Let $\bar{x} \in S$ be a feasible point with $f(\bar{x}) = y$. Notice that the function values of $f(x)$ and $g(x)$ are always nonnegative. Let the vector $\tilde{y}$ be defined by

$$\tilde{y}_i = \sqrt{y_i}, \quad i=1,\dots,m.$$

Then according to Theorem 2.30

$$\begin{aligned} \bar{x} \text{ is efficient for (4.91)} &\Leftrightarrow \mathcal{L}(f,y,S) = \mathcal{L}_=(f,y,S) \\ &\Leftrightarrow \mathcal{L}(g,\tilde{y},S) = \mathcal{L}_=(g,\tilde{y},S) \\ &\Leftrightarrow \bar{x} \text{ is efficient for (4.92)} \end{aligned}$$

□

The optimal solution of the weighted sum scalarization of (4.92) can, in general, not be determined explicitly [65]. But we can consider the e-constraint scalarization of (4.91) and (4.92) for $\mathfrak{e} \in \mathbb{R}^{m-1}$:

$$\begin{aligned} \min_{x \in S} \quad & \|x - a^m\|_2^2 \\ \text{s.t.} \quad & \|x - a^i\|_2^2 \leq \mathfrak{e}_i \quad i = 1, \ldots, m-1 \end{aligned} \tag{4.94}$$

$$\begin{aligned} \min_{x \in S} \quad & \|x - a^m\|_2 \\ \text{s.t.} \quad & \|x - a^i\|_2 \leq \mathfrak{e}_i \quad i = 1, \ldots, m-1 \end{aligned} \tag{4.95}$$

Notice that the feasible sets of (4.94) and (4.95) are given by

$$\bigcap_{i=1}^{m-1} \mathcal{L}(f_i, \mathfrak{e}_i, S) \text{ and } \bigcap_{i=1}^{m-1} \mathcal{L}(g_i, \mathfrak{e}_i, S),$$

respectively.
In particular, using (4.93) we can see that the feasible set of (4.94) for $\mathfrak{e} \in \mathbb{R}^{m-1}$, $\mathfrak{e} \geqq 0$, is identical to the feasible set of (4.95) with parameter $\mathfrak{e}'$ defined as

$$\mathfrak{e}'_i = \sqrt{\mathfrak{e}_i} \; i = 1, \ldots, m.$$

In addition, the global minima of the objective functions of (4.94) and (4.95) are attained at $x = a^m$ (in the unconstrained case).
Let $\mathfrak{E}(\mathscr{A})$ be a cell in the e-constraint parameter space of the e-constraint scalarization problem (4.94) as defined in Section 4.9.1. Then the corresponding cell in the parameter space of (4.95) is given by

$$\mathfrak{E}'(\mathscr{A}) = \left\{ \mathfrak{e}' \in \mathbb{R}^{m-1} \; : \; \mathfrak{e} \in \mathfrak{E}(\mathscr{A}), \; \mathfrak{e}'_i = \sqrt{\mathfrak{e}_i}, \; i = 1, \ldots, m-1 \right\}. \tag{4.96}$$

Hence, the e-constraint parameter space decomposition for (4.92) can be computed from (4.91) and vice-versa.

4.11 Conclusion

In Chapter 4 we have reviewed a parametric solution approach for multiobjective convex quadratic optimization problems using the weight space decomposition by efficient complementary bases of the parametric linear complementarity problem (pLCP). We have shown that there exists a one-to-one correspondence between efficient complementary bases and efficient active sets.
We proposed an algorithm for the determination of all efficient complementary bases without symbolic computations. We considered a generalization of multiobjective convex problems in canonical form to the general form.
Three special cases were discussed for which the weight cells are convex polyhedra. Multiobjective convex quadratic problems with diagonal objective matrices and lower and upper bounds were discussed and it was shown that the weight space decomposition of such problems is an arrangement of hyperplanes. This provides a polynomial bound on the number of efficient complementary bases (for a fixed m). Furthermore, we have found a multiobjective linear programming problem with a weight space decomposition that has the same arrangement of hyperplanes.
Furthermore, we considered the parameter space decomposition for the e-constraint scalarization that can be computed using the weight space decomposition. An application of this relationship was used to provide a method for the computation of the e-constraint parameter space decomposition for multiobjective location problems with l_2 and l_2^2 norms.

For multiobjective convex quadratic optimization problems with more than 3 objectives it is difficult to use the analytic description of the efficient set provided by the weight space decomposition, as the weight cells are in general $m-1$-dimensional semi-algebraic sets [1].
An approach to approximate the weight cells is discussed in Chapter 5 using multiobjective convex piecewise-linear optimization problems.
An interesting question is whether the weight space decomposition by active sets can be generalized to other multiobjective convex optimization problems, such as multiobjective convex polynomial optimization problems.

Chapter 5

Approximation of Multiobjective Convex Optimization Problems by Multiobjective Piecewise-Linear Problems

Many multiobjective optimization problems have a large number of efficient solutions - in the case of continuous optimization the efficient set can be infinite.

From a theoretical or technical standpoint it may be difficult to compute a complete description of the efficient set, even if an analytical description is available. We observed in Chapter 4 that the efficient set of multiobjective convex programming problems can be described analytically using efficient active sets and parametric optimization. However, we also observed that, apart from some special cases, computing the analytical representation is difficult and in large-dimensional cases impractical.

Furthermore, for real-world problems the aim is often to find an efficient solution that satisfies the preferences of the decision maker. For this task it may be sufficient to provide the decision maker with an approximation of the efficient set or the nondominated set, respectively.

For these reasons a variety of approaches have been provided in the literature that aim to compute a representation of the efficient set or the nondominated set, or both, for different types of multiobjective problems [74].

One category of such approaches are point approximations for which different quality measures are discussed in the literature, such as the Hausdorff distance between the nondominated set and the representation set (referred to as coverage) and uniformity [77, 78]. Point approximations can be computed using different techniques, for example dichotomic search [23, 71, 72] or evolutionary algorithms [10, 83].

A number of approaches construct a piecewise-linear inner or outer approximation of the nondominated set of convex multiobjective optimization problems [7, 25, 26, 52, 53]. We refer to [74] for a detailed survey.

In this chapter, we will consider an approach proposed by Oberdieck and Pistikopoulos [67] for approximating the weight space decomposition of multiobjective convex quadratic optimization problems. In this approach the objective functions are approximated by piecewise-linear functions. A weight space decomposition of the multiobjective piecewise-linear problem can then be computed.

In Section 5.1 the weight space decomposition for multiobjective convex continuous optimization problems is introduced.

In Section 5.2 multiobjective convex piecewise-linear programming problems and the corresponding formulation as a multiobjective linear programming problem are reviewed. Additionally, the weight space decomposition for multiobjective convex piecewise-linear programming problems is introduced.

The properties of the weight space decomposition for different types of multiobjective convex optimization problems are summarized in Section 5.3 using results from Chapter 4 and the literature [23]. In Section 5.4 the outer approximation approach by Oberdieck and Pistikopoulos [67] is reviewed. We discuss the convergence properties using results from the field of approximation of convex compact sets by polyhedrons.

In Section 5.5 a we construct an approximation of the weight space decomposition of a triobjective convex quadratic optimization problem and discuss the result.

5.1 A Weight Space Decomposition for Multiobjective Convex Optimization Problems

In this section we will first discuss an extension of the weight space decomposition by active sets for multiobjective convex quadratic optimization problems.
Consider a multiobjective convex programming problem with linear equality and inequality constraints:

$$\begin{array}{lll} \underset{x\in\mathbb{R}^n}{\text{vmin}} & f_i(x) & i=1,\dots,m \\ \text{s.t.} & Ax \geqq b,\ x_{\mathcal{I}_+} \geqq 0, Hx = h & \end{array} \tag{MCP}$$

with convex objective functions $f_i(x)$ for $i = 1,\dots,n$, $A \in \mathbb{R}^{p\times n}$, $b \in \mathbb{R}^p$, $H \in \mathbb{R}^{q\times n}$, $h \in \mathbb{R}^q$ and index set $\mathcal{I}_+ \subseteq \{1,\dots,n\}$ of nonnegative variables.
The feasible set is denoted by

$$S = \left\{x \in \mathbb{R}^n \ :\ Ax \geqq b,\ Hx = h,\ x_{\mathcal{I}_+} \geqq 0\right\}.$$

Additionally, we assume that the objective functions $f_i(x)$, $i = 1,\dots,m$, are continuous, but not necessarily differentiable, on S.
The weighted sum scalarization problem of (MCP) for $\lambda \in \Lambda$ is given by

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\min} & \sum_{i=1}^m \lambda_i f_i(x) \\ \text{s.t.} & Ax \geqq b,\ x_{\mathcal{I}_+} \geqq 0, Hx = h \end{array} \tag{WCP}$$

The set of optimal solutions of the weighted sum scalarization problem (WCP) is denoted by $X_{\text{opt}}(\lambda)$ for a fixed $\lambda \in \Lambda$.
In Definition 4.52 the efficient active sets for multiobjective convex quadratic optimization problems (gMQP) are defined using the KKT conditions. Since we do not assume that the objective functions of (MCP) $f_i(x)$ are differentiable for all $i = 1,\dots,m$ we consider here the following definition of active sets:

Definition 5.1. Let X_E be the efficient set of (MCP). The **active set** $\mathscr{A}(x) = (\mathcal{I}(x), \mathcal{J}(x))$ with $\mathcal{I} \subset \mathcal{I}_+$ of a feasible point $x \in S$ is defined as:

$$\mathcal{I}(x) = \left\{j \in \mathcal{I}_+ \ :\ x_j = 0\right\} \text{ and } \mathcal{J}(x) = \left\{j \in \{1,\dots,p\} \ :\ A_j x = b_j\right\}. \tag{5.1}$$

An active set $\bar{\mathscr{A}} = (\bar{\mathcal{I}}, \bar{\mathcal{J}})$ is called efficient active set of (MCP) if there exists $x \in X_E$ such that $\mathscr{A}(x) = \bar{\mathscr{A}}$.
The set of efficient active sets is defined as:

$$\mathscr{A}_{\text{eff}} = \{\mathscr{A}(x) \ :\ x \in X_E\} \tag{5.2}$$

For a given efficient active set $\bar{\mathscr{A}}$ we can define the weights $\lambda \in \Lambda$ such that there exists an optimal solution $x \in S$ of the weighted sum scalarization problem (WCP) with $\mathscr{A}(x) = \bar{\mathscr{A}}$:

$$\Lambda^C(\bar{\mathscr{A}}) := \left\{\lambda \in \Lambda \ :\ \exists x \in X_{\text{opt}}(\lambda) \cap X_E \ \text{ such that } \ \mathscr{A}(x) = \bar{\mathscr{A}}\right\} \tag{5.3}$$

In general, it is difficult to determine the set $\Lambda^C(\mathscr{A})$ for multiobjective convex problems for which no analytical or closed-formula solution of the weighted sum scalarization problem is available.
Two classes of multiobjective convex optimization problems have been discussed in this dissertation for which the weight cells can be determined analytically: Multiobjective linear programming problems, which were introduced in Section 2.2.3, and multiobjective convex quadratic optimization problems, which were discussed in Chapter 4.

In Section 5.2 we will discuss another class of multiobjective convex optimization problems for which a decomposition of the weight space can be computed explicitly.

5.2 Multiobjective Convex Piecewise-Linear Programming Problems

Singleobjective convex piecewise-linear optimization problems arise in the field of location theory, in particular when considering norms with polyhedral unit balls, see, for example, Hamacher and Nickel [41].

Multiobjective convex piecewise-linear optimization problems will play an important role in this chapter. We will shortly discuss how the efficient active sets and a decomposition of the weight space can be computed for multiobjective convex piecewise-linear programming problems using multiobjective linear programming techniques.

For a more detailed investigation see Fang et. al. [28] and in particular for the structure of the efficient set Nickel and Wiecek [65].

Consider a multiobjective convex piecewise-linear optimization problem:

$$\begin{array}{lll} \operatorname*{vmin}_{x\in\mathbb{R}^n} & f_i(x) := \max\limits_{k=1,\dots,K^i} (\boldsymbol{G}^i_{k\bullet})^T x + \boldsymbol{g}^i_k & i=1,\dots,m \\ \text{s.t.} & Ax \geqq b,\ x_{\mathcal{I}_+} \geqq 0, Hx = h & \end{array} \tag{5.4}$$

with matrices $\boldsymbol{G}^i \in \mathbb{R}^{K^i \times n}$ and vectors $\boldsymbol{g}^i \in \mathbb{R}^{K^i}$ for $K^i > 0$, $i = 1,\dots,m$. The linear constraints are given by a matrix $A \in \mathbb{R}^{p\times n}$, vector $b \in \mathbb{R}^p$, matrix $H \in \mathbb{R}^{q\times n}$ and vector $h \in \mathbb{R}^q$. Each objective function is given by the maximum of K^i affine-linear functions $(\boldsymbol{G}^i_{k\bullet})^T x + \boldsymbol{g}^i_k$ for each $i = 1,\dots,m$.

Notice that the objective functions $f_i(x)$, $i = 1,\dots,m$, are in general not differentiable over $\mathbb{R}^n$. However, (5.4) can be reformulated as a multiobjective linear optimization problem [28]:

$$\begin{array}{lll} \operatorname*{vmin}_{x\in\mathbb{R}^n,\ \alpha\in\mathbb{R}^m} & \alpha_i & i=1,\dots,m \\ \text{s.t.} & (\boldsymbol{G}^i_{k\bullet})^T x + \boldsymbol{g}^i_k \leq \alpha_i & \forall k=1,\dots,K^i\ \forall i=1,\dots,m \\ & Ax \geqq b,\ x_{\mathcal{I}_+} \geqq 0, Hx = h & \end{array} \tag{5.5}$$

Let $\boldsymbol{J} = \{1,\dots,n\} \setminus \mathcal{I}_+$ be the index set of unconstrained variables of (5.4). The multiobjective linear programming problem (5.5) can then be transformed into standard form [40] by adding slack variables s for the linear inequality constraints of (5.4) and a^i_k for $k = 1,\dots,K^i$ and $i = 1,\dots,m$ and by splitting the unconstrained variables $x_{\boldsymbol{J}}$ and α:

$$\begin{array}{lll} \text{vmin} & \alpha^+_i - \alpha^-_i & i=1,\dots,m \\ \text{s.t.} & (\boldsymbol{G}^i_{k\bullet})^T x - (\boldsymbol{G}^i_{k\boldsymbol{J}})^T z_{\boldsymbol{J}} + a^i_k - \alpha^+_i + \alpha^-_i = -\boldsymbol{g}^i_k & \forall k=1,\dots,K^i\ \forall i=1,\dots,m \\ & Ax - A_{\bullet\boldsymbol{J}} z_{\boldsymbol{J}} - s = b,\ Hx - H_{\bullet\boldsymbol{J}} z_{\boldsymbol{J}} = h & \\ & x \geqq 0,\ z \geqq 0,\ s \geqq 0,\ \alpha^+ \geqq 0,\ \alpha^- \geqq 0, & \\ & a^i_k \geq 0\ \forall k=1,\dots,K^i\ \forall i=1,\dots,m & \end{array} \tag{5.6}$$

Recall the definition of efficient basic solutions of multiobjective linear programming problems from Definition 2.33. Let $\mathcal{E}_E$ be the set of efficient basic solutions of (5.6). Then for each $\tilde{x} \in \mathcal{E}_E$, a corresponding cell in Λ can be defined in the following way (see Definition 2.34):

$$\Lambda(\tilde{x}) = \left\{\lambda \in \Lambda\ :\lambda^T C\tilde{x} \leq \lambda^T C\tilde{x}'\ \forall \tilde{x}' \in \tilde{S}\right\},$$

where $\tilde{S}$ and C are the feasible set and the objective matrix of (5.6), respectively.

The weight space decomposition for (5.6) can be computed using a variety of methods, for example the multiobjective simplex method [23], Benson's method [25, 58] or dichotomic search [72].

Let $\tilde{x} \in \mathcal{E}_E$ be an efficient basic solution of (5.6) and let $\bar{x}$ be the corresponding solution of (5.4), i.e. if $\tilde{x} = (x, z, s, \alpha^+, \alpha^-, a)$ then $\bar{x}$ is given by

$$\bar{x}_i = \begin{cases} x_j & \text{if } j \in \mathcal{I}_+ \\ x_j - z_j & \text{otherwise} \end{cases}.$$

Then $\bar{x}$ is an efficient solution of (5.4) and the weight cell $\Lambda(\tilde{x})$ can be associated with $\mathscr{A}(\bar{x})$.

Example 5.2. Consider the following multiobjective convex piecewise-linear optimization problem:

$$\begin{aligned}
\operatorname*{vmin}_{x\in\mathbb{R}^2} \quad & \max\left(\begin{pmatrix}-8 & 2\\ 9 & -9\\ 10 & -6\end{pmatrix}x+\begin{pmatrix}-3\\ 7\\ -10\end{pmatrix}\right)\\
& \max\left(\begin{pmatrix}-10 & 5\\ -7 & 3\\ 3 & -1\end{pmatrix}x+\begin{pmatrix}-3\\ 7\\ -10\end{pmatrix}\right)\\
& \max\left(\begin{pmatrix}-7 & -3\\ 4 & 3\\ -7 & 6\end{pmatrix}x+\begin{pmatrix}-9\\ 9\\ 6\end{pmatrix}\right)\\
\text{s.t.} \quad & -x_1-x_2\geq 2,\ x\geqq 0
\end{aligned} \tag{5.7}$$

The set of efficient basic solutions of the linear programming formulation of (5.7) can be computed , for example, using the multiobjective linear solver bensolve [58]. The efficient basic solutions of (5.7) are given by:

$$X_E=\left\{\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}0\\0.9091\end{pmatrix},\begin{pmatrix}0\\0.6667\end{pmatrix},\begin{pmatrix}0.4286\\1.5714\end{pmatrix},\begin{pmatrix}0.3902\\1.5122\end{pmatrix},\begin{pmatrix}0.4211\\1.5789\end{pmatrix}\right\}$$

The efficient active sets and the corresponding efficient basic solutions of (5.7) are given in Table 5.1.

Active Set	Color in Fig. 5.1	Efficient basic solutions
$\mathscr{A}_1=(\{1,2\},\emptyset)$	turquoise	$\begin{pmatrix}0\\0\end{pmatrix}$
$\mathscr{A}_2=(\{1\},\emptyset)$	red	$\begin{pmatrix}0\\0.9091\end{pmatrix},\begin{pmatrix}0\\0.6667\end{pmatrix}$
$\mathscr{A}_3=(\emptyset,\{1\})$	green	$\begin{pmatrix}0.4286\\1.5714\end{pmatrix},\begin{pmatrix}0.4211\\1.5789\end{pmatrix}$
$\mathscr{A}_4=(\emptyset,\emptyset)$	blue	$\begin{pmatrix}0.3902\\1.5122\end{pmatrix}$

Table 5.1: Efficient Active Sets in Example 5.2.

The weight space decomposition for the multiobjective linear programming problem can be seen in Figure 5.1. The color of each cell corresponds to the active set of the corresponding efficient basic solution.

As each cell corresponds to a particular efficient basic solution of (5.7) we can assign each weight cell to a particular efficient active set of (5.7).

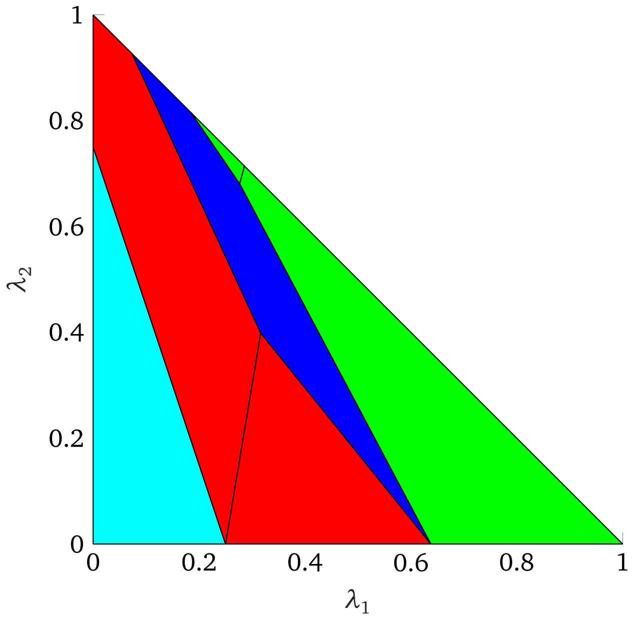

Figure 5.1: Decomposition of the weight space for Example 5.2.

5.3 Comparison of the Weight Space Decomposition by Active Sets for Multiobjective Convex Optimization Problems

Table 5.2 summarizes the properties of the weight space decomposition by efficient active sets (and efficient complementary bases in the case of multiobjective quadratic programming problems) for different types of multiobjective convex problems discussed in this thesis.

Problem	Constraints	Type of cells	References	Literature
MLP	linear	convex polyhedron	Section 2.2.3	Isermann [48]
MDQP	bounds	arrangement of hyperplanes	Section 4.8	
MMLQP	linear	convex polyhedron	Section 4.6	Markowitz [59], Hirschberger et. al. [46]
MSQP	linear	convex polyhedron	Section 4.7	
MPLP	linear	union of convex polyhedrons	Section 5.2	
MQP	linear	semi-algebratic sets	Sections 4.2 and 4.4	Adelgren [1]
MCP	linear		Section 5.1	

Table 5.2: Types of weight cells for different convex multiobjective optimization problems with linear constraints.

Consider approximating the weight space decomposition of a multiobjective convex quadratic programming problem by constructing a multiobjective convex optimization problem for which an explicit computation of the weight space decomposition by active sets is possible. In the case of multiobjective linear programming problems and the special cases of multiobjective convex quadratic programming problems discussed in Sections 4.6, 4.7 and 4.8 we know that the weight space is decomposed into convex polyhedrons by efficient active sets.
However, the weight cells $\Lambda^C(\mathscr{A})$ for multiobjective convex optimization problems are in general not convex, as can be seen in Example 4.47 for multiobjective convex quadratic problems, and in Example 5.2 for multiobjective convex piecewise-linear optimization problems. But the weight space decomposition of multiobjective convex piecewise-linear optimization problems consists of unions of convex polyhedrons that can be nonconvex. Thus, from the selection of multiobjective convex optimization problems discussed here, multiobjective convex piecewise-linear optimization problems seem to be a good candidate for an approximation procedure.

In the following section, we discuss an approach by Oberdieck and Pistikopoulos [67] that approximate the weight space decomposition of a multiobjective convex quadratic optimization problem using a multiobjective convex piecewise-linear optimization problem. One step in this approach is the approximation of each nonlinear convex objective function by a convex piecewise-linear objective function.

5.4 An Outer Approximation Algorithm for Convex Functions

The approximation of functions by polyhedral structures is discussed in the literature in different contexts, for example in branch-and-cut algorithms [6, 56] and approximations of the nondominated set for multiobjective convex optimization problems [53, 80].

In this section we will consider an approximation algorithm that was used by Oberdieck and Pistikopoulos [67] to approximate convex quadratic functions. The procedure is similar to the outer approximation discussed by Bertsekas and Yu [6] for convex functions.

Let $X \subseteq \mathbb{R}^n$ be a compact polyhedron over which a continuously differentiable convex function $f : X \to \mathbb{R}$ is to be approximated. The aim of this section is to construct a piecewise-linear function $\mathcal{T}$ that is an outer approximation of f over X satisfying

$$\mathcal{T}(x) \leq f(x)\ \forall x \in X \ \text{ and } |f(x) - \mathcal{T}(x)| \leq \varepsilon\ \forall x \in X$$

for some $\varepsilon > 0$.
Since f is convex and differentiable we know that for any $z \in X$ that the following inequality holds:

$$f(x) \geq f(z) + \nabla f(z)^T (x - z)\ \forall x \in X$$

Using a set of finitely many points $Z \subseteq X$ an outer approximation of f can be defined by

$$\mathcal{T}(x, Z) = \max_{z \in Z} \big(f(z) + \nabla f(z)^T (x - z) \big)\ \forall x \in \mathbb{R}^n.$$

It is easy to see that $\mathcal{T}(x, Z)$ is a convex piece-wise linear function and an outer approximation of f for every set of points $Z \subseteq X$, $Z \neq \emptyset$. Given an approximating point set $Z \subset X$ the approximation error at $x \in X$ is defined as

$$\eta(x, Z) := f(x) - \mathcal{T}(x, Z)$$

and the approximation error for the whole set X is defined as

$$\eta(X, Z) := \max_{x \in X} \eta(x, Z).$$

Oberdieck and Pistikopoulos [67] formulated a minimax optimization problem to compute the approximation error $\eta(X, Z)$ in the case where f is a strictly convex quadratic function:

$$\eta(X, Z) = - \min_{x \in X} \Big(\max_{z \in Z} \big(f(z) + \nabla f(z)^T (x - z) \big) - f(x) \Big) \tag{5.8}$$

The minimax problem (5.8) is in general nonconvex and not differentiable. Algorithm 5.1 was proposed by Oberdieck and Pistikopoulos [67] to compute a set of approximating points Z^k such that $\eta(X, Z) \leq \eta^*$ for a desired approximation error $\eta^* > 0$.

Algorithm 5.1: Approximation Algorithm for Convex Functions [67]

Input: Convex function f, compact set $X \subseteq \mathbb{R}^n$, desired approximation error $\eta^* > 0$, initial set of approximating points Z^0

Set $k := 0$.

Compute $\eta_0 = \eta(X, Z^0)$ and let z^k be the corresponding optimal solution of (5.8).

while $\eta_k > \eta^*$ **do**

- $Z^{k+1} := Z^k \cup \{z^k\}$.
- Compute η_{k+1} and z^{k+1} such that $\eta_{k+1} = \eta(X, Z^{k+1}) = f(z^{k+1}) - \mathcal{T}(x, Z^{k+1})$.
- Set $k := k + 1$.

Output: Set of approximating points Z^k

In the upcoming sections we want to investigate the following questions about Algorithm 5.1:

- What are the convergence properties of Algorithm5.1?
- How can the new approximating point be computed in each iteration?

In order to answer these questions we consider a related field of research that considers the approximation of convex bodies by convex polyhedra.

5.4.1 Approximation of Convex Bodies by Polyhedra

A summary of results about the approximation of convex sets by polyhedrons can be found in Bronstein [11]. Kamenev [50, 51] investigated the convergence properties of a class of algorithms for constructing approximations of convex bodies by convex polyhedrons, called Hausdorff algorithms or Hausdorff schemes [50].

Definition 5.3. [8] The epigraph of $f : \mathbb{R}^n \to \mathbb{R}$ over $X \subseteq \mathbb{R}^n$ is defined as the set

$$\mathrm{Epi} f = \{(x, y) \in X \times \mathbb{R} \; : \; f(x) \leq y\}.$$

The epigraph Epi$f$ of a convex function f over a convex domain is also convex [73]. However, the epigraph is unbounded. Let X be a closed convex polyhedron. Let y^* be an upper bound for the largest function value of f over X, for example

$$y_{\max} = \max_{x \in X} f(x).$$

Then the following set $\mathcal{C}$ is convex and compact:

$$\mathcal{C} = \{(x, y) \in X \times (-\infty, y^*] \; : \; f(x) \leq y\}$$

We call $\mathcal{C}$ the bounded epigraph of f and $y_{\max}$ the **cut-off level**. Given a set of approximating points Z we can compute the bounded epigraph of $\mathcal{T}(x, Z)$:

$$\begin{aligned} P(Z) &= \{(x, y) \in X \times (-\infty, y_{\max}] \; : \; \mathcal{T}(x, Z) \leq y\} \\ &= \left\{(x, y) \in X \times (-\infty, y_{\max}] \; : \; \max_{z \in Z} \left(f(z) + \nabla f(z)^T (x - z)\right) \leq y\right\} \\ &= \left\{(x, y) \in X \times (-\infty, y_{\max}] \; : \; f(z) + \nabla f(z)^T (x - z) \leq y \;\; \forall z \in Z\right\} \end{aligned}$$

Notice that $P(Z)$ is a bounded convex polyhedron and $\mathcal{C} \subseteq P(Z)$ for all non empty approximation point sets $Z \subseteq X$. Hence, $P(Z)$ is a polyhedral outer approximation of $\mathcal{C}$. Algorithm 5.2 is a formulation of Algorithm 5.1 in the context of approximating $\mathcal{C}$ by a sequence of polyhedrons $P^k = P(Z^k)$ with approximating point sets Z^k.

Algorithm 5.2: Outer Approximation Algorithm for Compact Convex Sets Using the Approximation Error

Input: Compact convex set $\mathcal{C}$, initial approximating point set Z^0, desired error bound $\eta^* > 0$
Compute $P^0 = P(Z^0)$.
Determine $p^0 = (x^0, y^0) \in P^0$ such that $\eta(P^0) = f(x^0) - y^0$.
Set $\eta_0 := f(x^0) - y^0$ and $k := 0$.
while $\eta_k > \eta^*$ **do**
  $Z^{k+1} = Z^k \cup \{x^k\}$.
  $P^{k+1} = P(Z^{k+1})$.
  Determine $p^{k+1} = (x^{k+1}, y^{k+1}) \in P^{k+1}$ such that $\eta(P^{k+1}) = f(x^{k+1}) - y^{k+1}$.
  Set $\eta_{k+1} := f(x^{k+1}) - y^{k+1}$ and $k = k + 1$.

In order to analyze the convergence properties of Algorithm 5.2 we consider a different approximation error:

Definition 5.4. [39]

1. The (Euclidean) distance between two points $x, y \in \mathbb{R}^n$ is defined as:
$$d(x,y) = \|x-y\|_2$$
2. The (Euclidean) distance between a point $x \in \mathbb{R}^n$ and a set $M \subseteq \mathbb{R}^n$ is defined as:
$$d(x,M) = \min_{y \in M} d(x,y)$$
3. The Hausdorff distance of two sets $C_1, C_2 \subseteq \mathbb{R}^n$ is defined as:
$$d_H(C_1,C_2) = \max\left\{\sup_{x\in C_1} d(x,C_2),\ \sup_{y \in C_2} d(y,C_1)\right\}$$
4. The approximation error for $p = (x_p, y_p) \in P(Z) \subset \mathbb{R}^n$ for some set of approximating points Z is denoted by
$$\eta(p) = f(x_p) - y_p.$$

Kamenev [50] considered an approximation procedure that minimizes the Hausdorff distance between a convex compact set $\mathcal{C}$ and an outer approximation polyhedron. Algorithm 5.3 is the method formulated for the approximation of the bounded epigraph $\mathcal{C}$.

Algorithm 5.3: Outer Approximation Algorithm for Compact Convex Sets Using Hausdorff Distances [50]

Input: Compact convex set $\mathcal{C}$, initial approxmating point set Z^0, desired error bound $\delta^* > 0$
Compute $P^0 = P(Z^0)$.
Determine $p^0 \in P^0$ and $c^0 = (x^0, y^0) \in \mathcal{C}$ such that $d_H(P^0, \mathcal{C}) = d(p^0, c^0)$.
Set $\delta_0 := d(p^0, c^0)$ and $k := 0$.
while $\delta_k > \delta^*$ **do**
  $Z^{k+1} = Z^k \cup \{x^k\}$.
  $P^{k+1} = P(Z^{k+1})$.
  Determine $p^{k+1} \in P^{k+1}$ and $c^{k+1} = (x^{k+1}, y^{k+1}) \in \mathcal{C}$ such that $d_H(P^{k+1}, \mathcal{C}) = d(p^{k+1}, c^{k+1})$.
  Set $\delta_{k+1} := d(p^{k+1}, c^{k+1})$ and $k = k+1$.

To compute the approximation error and to show a convergence result we consider the points where the approximation error is attained.

It is a well-known fact that the Hausdorff-distance between a convex compact set and an enclosing polyhedron is attained at an extreme point of the polyhedron [39]:

Proposition 5.5. [6, 39] For every $w \in P$ there exists an extreme point $v \in P$ such that $d(v, \mathcal{C}) \geq d(w, \mathcal{C})$.

This result can, for example, be shown using linear programming theory.
The following proposition is similar to the proof of Theorem 5.2.2 in Bard [4] and shows a similar result to Proposition 5.5 for the approximation error η.

Proposition 5.6. [4] For every $w \in P$ there exists an extreme point $v \in P$ such that $\eta(v) \geq \eta(w)$.

Proof. [4] Let $v^1, \ldots, v^r$ be the extreme points of P with $v^i = (x^i, y^i)$ for $i = 1, \ldots, r$. Then, for every $\hat{v}' \in P$ there exists $\alpha \in \mathbb{R}^r$, $\sum_{i=1}^r \alpha_i = 1$, $\alpha \geqq 0$ such that

$$\hat{v} = \sum_{i=1}^{r} \alpha_i v^i.$$

Using Jensen's inequality:

$$\begin{aligned}
\eta(\hat{v}) &= f(\hat{x}) - \hat{y} \\
&= f\left(\sum_{i=1}^{r} \alpha_i x^i\right) - \sum_{i=1}^{r} \alpha_i y^i \\
&\leq \sum_{i=1}^{r} \alpha_i f(x^i) - \sum_{i=1}^{r} \alpha_i y^i \\
&= \sum_{i=1}^{r} \alpha_i \left(f(x^i) - y^i\right) \\
&= \sum_{i=1}^{r} \alpha_i \eta(v^i)
\end{aligned}$$

Which shows that

$$\eta(\hat{v}) \leq \sum_{i=1}^{r} \alpha_i \eta(v^i) \tag{5.9}$$

Now assume that $\eta(\hat{v}) > \eta(v^i)$ for $i = 1, \ldots, r$. Then

$$\sum_{i=1}^{r} \alpha_i \eta(v^i) < \sum_{i=1}^{r} \alpha_i \eta(\hat{v}) = \eta(\hat{v})$$

which contradicts (5.9). □

In order to analyze the convergence properties of Algorithm 5.2 we show that the sequence of polyhedra P^k, $k \in \mathbb{N}$, computed by Algorithm 5.2 is a so-called Hausdorff sequence:

Definition 5.7. [51] Let $\mathcal{C}$ be a convex compact set. A sequence of polytopes $(P^k)_k$ is called a **Hausdorff sequence** if there exists a constant $\gamma > 0$ such that

$$d_H(P^k, P^{k+1}) \geq \gamma d_H(P^k, \mathcal{C}) \ \forall k = 1, \ldots$$

Notice that Algorithm 5.3 computed Hausdorff sequences with $\gamma = 1$ [50].

Proposition 5.8. Let $(P^k)_k$ be a sequence constructed by Algorithm 5.2. Then there exists $\beta > 0$ such that

$$\eta_k \leq \beta d_H(P^k, \mathcal{C}), \ \forall k = 1, \ldots.$$

Proof. Let $p = (x_p, y_p)$ be an extreme point of P^k. Let $c = (x_c, y_c) \in C$ be such that $d(p, \mathcal{C}) = d(p, c)$. The situation is illustrated in Figure 5.2. Consider the point $c' = (x_p, f(x_p))$.

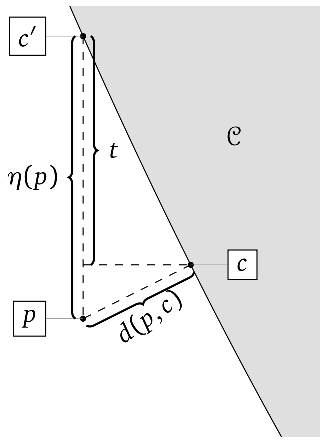

Figure 5.2: Bounding the approximation error at an extreme point of P^k.

Consider the line segment $t := f(x_p) - f(x_c)$. Since f is convex the following holds:

$$f(x_c) \geq f(x_p) + \nabla f(x_p)^T (x_c - x_p) \Leftrightarrow \underbrace{f(x_p) - f(x_c)}_{=t} \leq \nabla f(x_p)^T (x_p - x_c) \tag{5.10}$$

By assumption, $||\nabla f(x)||$ is continuous over X and X is a compact set. Hence, $||\nabla f(x)||$ attains it's maximum in X. This leads to a bound on t:

$$t \leq \nabla f(x_p)^T (x_p - x_c) \leq ||\nabla f(x_p)|| \underbrace{||x_p - x_c||}_{\leq d(p,c)} \leq \max_{x \in X} ||\nabla f(x)|| \cdot d(p, c)$$

Now, consider the approximation error for p:

$$\eta(p) \leq t + \underbrace{|y_c - y_p|}_{\leq d(p,c)} \leq \left(1 + \max_{x \in X} ||\nabla f(x)||\right) \cdot d(p, c)$$

Let $p \in P$ be an extreme point such that $\eta_k = \eta(p)$. Using the fact that $d(p, \mathcal{C}) \leq d_H(P^k, \mathcal{C})$ for all $p \in P^k$ we see that

$$\eta_k \leq \left(1 + \max_{x \in X} ||\nabla f(x)||\right) d(p, \mathcal{C}) \leq \underbrace{\left(1 + \max_{x \in X} ||\nabla f(x)||\right)}_{=:\beta} d_H(P^k, \mathcal{C}).$$

Notice that β is independent of k. □

Proposition 5.9. Let $(P^k)_k$ be a sequence constructed by Algorithm 5.2. Then there exists $\gamma > 0$ such that

$$d_H(P^k, P^{k+1}) \geq \gamma d_H(P^k, \mathcal{C}) \; \forall k = 1, \dots$$

Proof. Let $p = (x_p, y_p) \in P^k$ be given such that $\eta_k = f(x_p) - y_p$ and $Z^{k+1} = Z^k \cup \{x_p\}$. Let $c \in \mathcal{C}$ be the point $c = (x_p, f(x_p))$. Notice that the polyhedron P^{k+1} is given by:

$$P^{k+1} = P^k \cap \left\{(x, y) \in \mathbb{R}^{n+1} \;:\; y \geq f(x_p) + \nabla f(x_p)^T(x - x_p)\right\} \tag{5.11}$$

Also obverse that c is on the graph of $\mathcal{T}(\bullet, Z^{k+1})$ and $\mathcal{T}(x_p, Z^{k+1}) = f(x_p)$.
Now, let $q = (x_q, y_q) \in P^{k+1}$ such that $d(p, P^{k+1}) = d(p, q)$. The situation is illustrated in Figure 5.3.

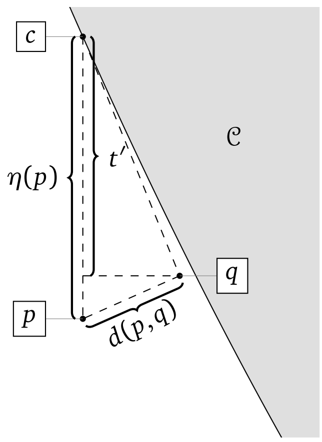

Figure 5.3: Bounding of $d_H(P^k, P^{k+1})$.

We can observe the following:

$$\eta(p) = f(x_p) - y_p = \underbrace{f(x_p) - y_q}_{=:t'} + \underbrace{y_q - y_p}_{\leq d(p,q)} \tag{5.12}$$

Consider the distance t' and notice that c and q are points on the graph of $\mathcal{T}(\bullet, Z^{k+1})$. Thus, it holds that

$$t' = f(x_p) - y_q = \mathcal{T}(x_p, Z^{k+1}) - \mathcal{T}(x_q, Z^{k+1}).$$

Note that the subgradients of $\mathcal{T}(x, Z^{k+1})$ are bounded by gradients of f on Z^{k+1} [6]:

$$\partial\mathcal{T}(x, Z^{k+1}) = \left\{v \in \mathbb{R}^n \;:\; \mathcal{T}(x, Z^{k+1}) - \mathcal{T}(x', Z^{k+1}) \geq v^T(x - x') \; \forall x \in \mathbb{R}^n\right\} \subseteq \operatorname{conv}(\left\{\nabla \mathrm{f(x)} \;:\; \mathrm{x} \in \mathrm{Z}^{k+1}\right\})$$

Thus, all subgradients of $\mathcal{T}(x, Z^{k+1})$ are bounded for all $k \in \mathbb{N}$ and

$$\max_{v \in \partial\mathcal{T}(x, Z^{k+1})} \|v\| \leq \max_{x \in X} \|\nabla f(x)\|. \tag{5.13}$$

Using (5.13) we can see that

$$t' = \mathcal{T}(x_p, Z^{k+1}) - \mathcal{T}(x_q, Z^{k+1}) \leq \left(\max_{x \in X} \|\nabla f(x)\|\right) d(p, q). \tag{5.14}$$

Using (5.12) and (5.14) we can observe that

$$\eta(p) \leq t' + d(p, q) \leq \left(1 + \max_{x \in X} \|\nabla f(x)\|\right) d(p, q). \tag{5.15}$$

Since $d(p,q) = d(p, P^{k+1})$ we also know that $d(p,q) \leq d_H(P^k, P^{k+1})$. Applying this and the fact that $\eta_k \geq d_H(P^k, \mathcal{C})$ to (5.15) we observe that

$$d_H(P^k, \mathcal{C}) \leq \eta_k = \eta(p) \leq \left(1 + \max_{x \in X} \|\nabla f(x)\|\right) d(p,q) \leq \left(1 + \max_{x \in X} \|\nabla f(x)\|\right) d_H(P^k, P^{k+1}).$$

This is equivalent to the following bound for $d_H(P^k, P^{k+1})$ which holds for every $k \in \mathbb{N}$:

$$d_H(P^k, P^{k+1}) \geq \left(1 + \max_{x \in X} \|\nabla f(x)\|\right)^{-1} d_H(P^k, \mathcal{C})$$

□

Proposition 5.9 shows that indeed the sequence computed by Algorithm 5.2 is a Hausdorff sequence. We can now apply one of the main results from Kamenev [50, 51]:

Theorem 5.10. [50] Let $(P^k)_k$ be a sequence constructed by Algorithm 5.2 or by Algorithm 5.3, respectively. Then the following statements hold:

1. If f is twice continuously differentiable, then

$$d_H(C, P^k) \leq O(k^{\frac{2}{-n}}) \tag{5.16}$$

2. If f is once continuously differentiable, then

$$d_H(C, P^k) \leq O(k^{\frac{1}{-n}}) \tag{5.17}$$

We have now established that Algorithm 5.2 and, by extension, Algorithm 5.1 do converge. Using Proposition 5.8 we also know the following:

Proposition 5.11. Let $(P^k)_k$ be a sequence constructed by Algorithm 5.2.

1. If f is twice continuously differentiable, then

$$\eta_k \leq O(k^{\frac{2}{-n}}) \tag{5.18}$$

2. If f is once continuously differentiable, then

$$\eta_k \leq O(k^{\frac{1}{-n}}) \tag{5.19}$$

Proof. The result follows from Theorem 5.10 and Proposition 5.8. □

5.4.2 Implementation of the Approximation Scheme

In each iteration of Algorithm 5.2 a point $p^{k+1} \in P^k$ has to be computed where the approximation error $\eta(p^{k+1})$ is maximal. This could, for example, be done by computing the optimal solution of the minimax problem (5.8). However, as (5.8) is in general not convex, global optimization routines have to be used to ensure that the correct point is found. The convergence proof (Proposition 5.11) relies on the fact that, at least after a finite number of iterations, the approximation error decreases.
In order to ensure that the global optimum of (5.8) is found global optimization techniques can be used. For further details consider, for example, Pardalos and Romeijn [68].
Proposition 5.6 shows that the maximum approximation error is attained at an extreme point of P^k. Hence, a reliable approach is to compute all extreme points of P^k and compute the approximation error at each point. Recall that $P^k = P(Z^k)$ is given by:

$$P(Z^k) = \left\{(x, y) \in X \times (-\infty, y_{\max}] \;:\; f(z) + \nabla f(z)^T(x - z) \leq y \;\; \forall z \in Z^k\right\}$$

where $y_{\max}$ is the cut-off level for the bounded epigraph. Notice that $P(Z^k)$ is defined by a set of linear inequalities.
Computing all extreme points of a polyhedron from a description by linear inequalities, also referred to as vertex enumeration, is a common subject in the field of computational geometry [76]. One method to compute all extreme points is the double description method [35].

Example 5.12. Consider the first objective function $f_1(x)$ from Example 4.47:

$$f_1(x) = \frac{1}{2}x^T \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} x + \begin{pmatrix} -40 \\ -24 \end{pmatrix}^T x$$

We use Algorithm 5.2 to compute an approximation of $f_1(x)$ over the set

$$X = \{x \in \mathbb{R}^n \ : \ 0.1039 \leq x_1 \leq 15.8961, \ -3.1668 \leq x_2 \leq 19.1668\}.$$

The graph of f_1 over X is depicted in Figure 5.4. For selected iterations the approximation error η_k is shown in Table 5.3. A selection of approximating polyhedra P^k is shown in Figure 5.5.
To enumerate the extreme points of P^k in each iteration k we used the ccdmex implementation of the double desciption method from the MPT-toolbox [43]. The following results were obtained on a machine with an Intel Core i5 processor with 3.2 GHz and 8 GB of ram.

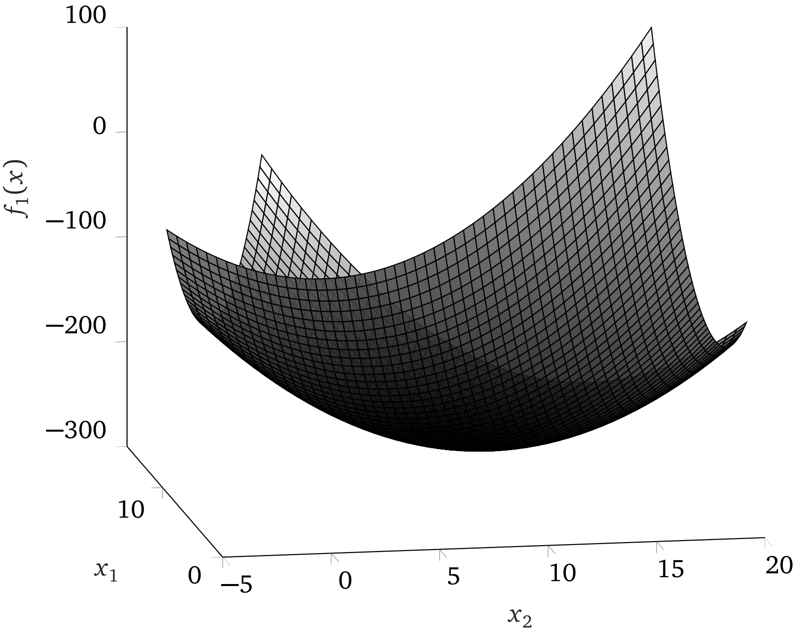

Figure 5.4: Graph of the function f_1 in Example 5.12

The sequence of approximation error behaves as expected, considering the approximation results given in Proposition 5.11. As the number of vertecies of P^k increases in each iteration

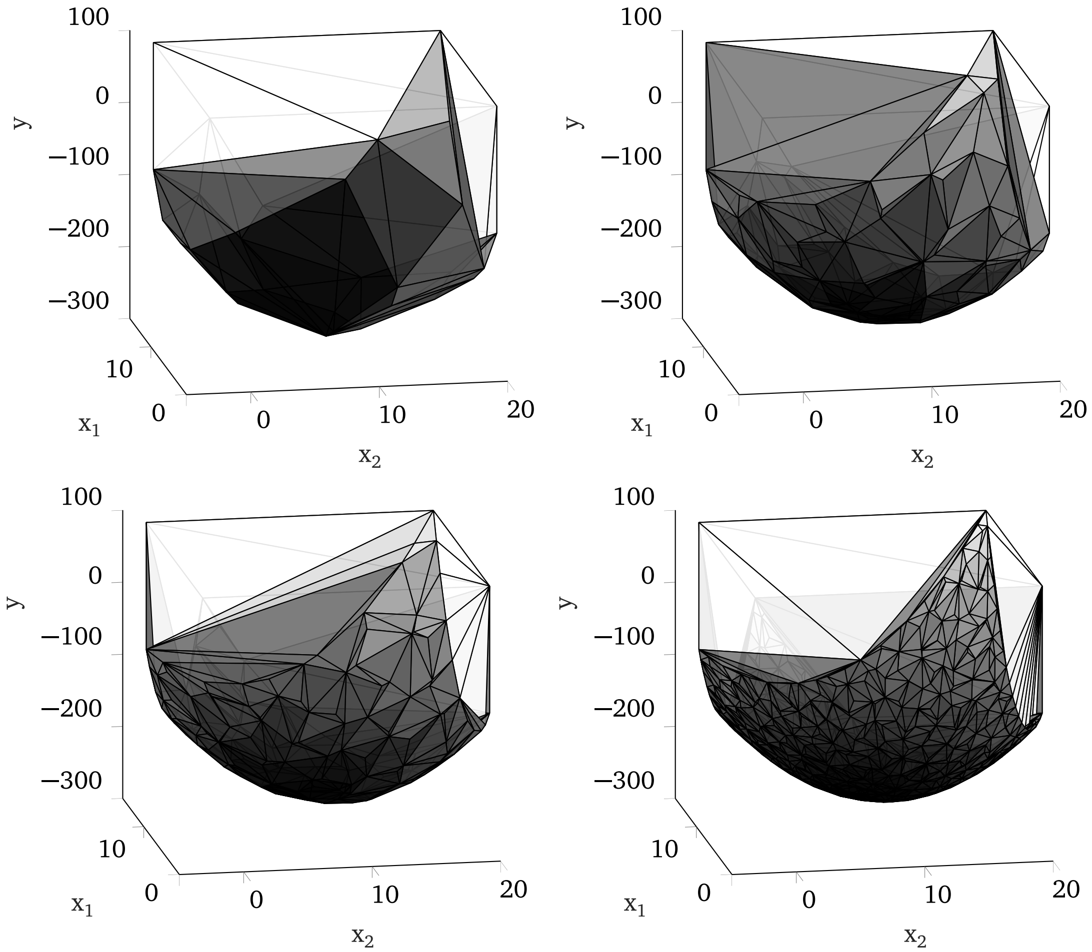

Figure 5.5: Approximating polyhedra of Example 5.12 after 10 (top left), 40 (top right), 80 (bottom left), 300 (bottom right) iterations

Iterations	Approximation Error	Time (in seconds)
10	40.7177	0.008
20	17.5871	0.021
30	12.9486	0.041
40	8.3292	0.068
50	6.8500	0.103
60	5.8895	0.145
70	5.0592	0.197
80	4.4458	0.258
90	3.9916	0.328
100	3.4737	0.410
200	1.6626	2.031
300	1.0967	5.777
400	0.8164	12.745
500	0.6443	23.997
1000	0.3224	190.939
2000	0.1604	1660.136

Table 5.3: Sequence of approximation errors in Example 5.12.

5.5 Approximation of the Weight Space Decomposition

In this section we will discuss the application of the approximation procedure reviewed in Section 5.4 to the multiobjective problems considered in this thesis. In the approximation scheme proposed by Oberdiek and Pistikopoulos [67] all objective functions are approximated by piecewise-linear functions in order to compute an approximation of the weight space decomposition.

Consider a multiobjective convex programming problem as introduced in Section 5.1.

$$\begin{array}{lll} \underset{x\in\mathbb{R}^n}{\text{vmin}} & f_i(x) & i=1,\dots,m \\ \text{s.t.} & Ax \geqq b,\ x_{\mathcal{I}_+} \geqq 0, Hx = h & \end{array} \tag{5.20}$$

with strictly-convex objective functions $f_i(x)$ for $i=1,\dots,n$, $A \in \mathbb{R}^{p\times n}$, $b \in \mathbb{R}^p$, $H \in \mathbb{R}^{q\times n}$, $h \in \mathbb{R}^q$ and index set of nonnegative variables $\mathcal{I}_+ \subseteq \{1,\dots,n\}$. In addition we assume that each objective function f_i is continuously differentiable and bounded on S for all $i=1,\dots,m$.

We will now suggest the approximation scheme for approximating the weight space decomposition by active sets of (5.20). We are most interested in a tight approximation of the objective functions over the efficient set of (5.20). For this reason we choose the sets X^i over which the function $f_i(x)$ is to be approximated such that

$$\mathcal{L}(f_i, y_i^N) \subseteq X^i\ \forall i=1,\dots,m,$$

where y^N is the nadir point of (5.20). This can be done, for example, by computing a bounding box of $\mathcal{L}(f_i, y_i^N)$. If the nadir point of (5.20) is not available a upper bound of the nadir point can be used instead. The extreme points of X^i can then be used as the initial set of approximation points.
Using Algorithm 5.2 we can then compute the approximating sets Z^i for each objective function f_i, $i=1,\dots,m$. The outer approximation of (5.20) problem is then given by:

$$\begin{array}{lll} \underset{x\in\mathbb{R}^n}{\text{vmin}} & \mathcal{T}_i(x, Z^i) & i=1,\dots,m \\ \text{s.t.} & Ax \geqq b,\ x_{\mathcal{I}_+} \geqq 0, Hx = h & \end{array} \tag{5.21}$$

Consider that the approximation error for each objective is smaller than η^*.

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\min} & \sum_{i=1}^m \lambda_i \mathcal{T}_i(x, Z^i) \\ \text{s.t.} & Ax \geqq b,\ x_{\mathcal{I}_+} \geqq 0, Hx = h \end{array} \tag{5.22}$$

Let $\lambda \in \Lambda$ be given. Consider the difference between the weighted sum of the original functions $f_i(x)$, $i = 1, \dots, m$ and the piecewise-linear functions $\mathcal{T}_i(x, Z^i)$, $i = 1, \dots, m$:

Proposition 5.13. Let $\lambda \in \Lambda$ be given and let Z^i, $i = 1, \dots, m$, be the sets of approximating points computed by Algorithm 5.2 with approximation error η^* for each objective function $f_i(x)$, $i = 1, \dots, m$. Then $\sum_{i=1}^m \lambda_i \mathcal{T}_i(x, Z^i)$ is an outer approximation function of $\sum_{i=1}^m \lambda_i f_i(x)$ and

$$\sum_{i=1}^{m} \lambda_i f_i(x) - \sum_{i=1}^{m} \lambda_i \mathcal{T}_i(x, Z^i) \leq \eta^*.$$

Proof. It is easy to see for every $\lambda \in \Lambda$ that

$$\sum_{i=1}^{m} \lambda_i \underbrace{f_i(x)}_{\geq \mathcal{T}_i(x,Z)} \geq \sum_{i=1}^{m} \lambda_i \mathcal{T}_i(x, Z).$$

Now consider the difference between the weighted sums:

$$\begin{aligned} \sum_{i=1}^{m} \lambda_i f_i(x) - \sum_{i=1}^{m} \lambda_i \mathcal{T}_i(x, Z) &= \sum_{i=1}^{m} \lambda_i \left(\underbrace{f_i(x) - \mathcal{T}_i(x, Z^i)}_{\leq \eta^*} \right) \\ &\leq \sum_{i=1}^{m} \lambda_i \eta^* = \eta^* \end{aligned}$$

□

To assess the quality of the approximation of the weight space decomposition of (5.20) by the weight space decomposition of (5.21) we consider the following:

Definition 5.14. Let $\tilde{X}_E$ be the efficient set of (5.21) and let $\tilde{X}_{\text{opt}}(\lambda)$ be the set of optimal solutions of the weighted sum problem (5.22) of the outer approximation (5.21) for $\lambda \in \Lambda$.
Let $\varepsilon \geq 0$ and let an active set $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ be given. Let $S_\varepsilon(\mathscr{A}) \subseteq S$ be given by:

$$S_\varepsilon(\mathscr{A}) = \left\{ x \in S \ : \ x_{\mathcal{I}} \leqq \varepsilon, \, A_{\mathcal{J}\bullet} x - b_{\mathcal{J}} \leqq \varepsilon \right\}$$

Then we define the following set in $\mathbb{R}^m$:

$$\Omega_\varepsilon(\mathscr{A}) := \left\{ \lambda \in \Lambda : \tilde{X}_E \cap \tilde{X}_{\text{opt}}(\lambda) \cap S_\varepsilon(\mathscr{A}) \neq \emptyset \right\}.$$

The sets $\Omega_\varepsilon(\mathscr{A})$ can be understood as approximations of the weight cells $\Lambda^A(\mathscr{A})$ for a given active set $\mathscr{A}$.

Proposition 5.15. Let $\bar{x}(\lambda)$ be the optimal solution of the weighted sum problem of (5.20) for $\lambda \in \Lambda$. Let $\tilde{X}_E$ be the efficient set of (5.21) and let $\tilde{X}_{\text{opt}}(\lambda)$ be the set of optimal solutions of the weighted sum problem (5.22) for $\lambda \in \Lambda$ of the outer approximation (5.21).
Let Z^i be approximating point sets for $i = 1, \dots, m$ such that there exists $\delta > 0$ such that

$$d(\bar{x}(\lambda), \tilde{X}_{\text{opt}}(\lambda) \cap \tilde{X}_E) \leq \delta \ \forall \lambda \in \Lambda.$$

Then there exists $\zeta > 0$ such that

$$\Lambda^C(\mathscr{A}) \subseteq \Omega_\varepsilon(\mathscr{A})$$

for any efficient active set $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ of (5.20) and $\varepsilon = \zeta\delta$.

Proof. Let $\mathscr{A} = (\mathcal{I}, \mathcal{J})$ be an efficient active set of (5.20) and let $\lambda \in \Lambda^C(\mathscr{A})$ be given. Let $\bar{x} = \bar{x}(\lambda)$ be the optimal solution of the weighted sum problem of (5.20) for λ.
By assumption, there exits $\hat{x} \in \tilde{X}_{\text{opt}}(\lambda) \cap \tilde{X}_E$ such that $d(\bar{x}, \hat{x}) = \|\bar{x} - \hat{x}\| \leq \delta$.
Consider the active linear inequality constraints:

$$\begin{aligned}\left\|A_{j\bullet}\hat{x} - b_j\right\| &= \left\|A_{j\bullet}\hat{x} - A_{j\bullet}\bar{x} + A_{j\bullet}\bar{x} - b_j\right\| \\ &\leq \left\|A_{j\bullet}(\hat{x} - \bar{x})\right\| + \underbrace{\left\|A_{j\bullet}\bar{x} - b_j\right\|}_{=0} \\ &\leq \left\|A_{j\bullet}\right\| \underbrace{\|\hat{x} - \bar{x}\|}_{\leq \delta} \\ &\leq \left\|A_{j\bullet}\right\| \delta \ \forall j \in \mathcal{J}\end{aligned}$$

Furthermore, consider the active nonnegativity constraints:

$$\begin{aligned}\|\hat{x}_{\mathcal{J}}\| &= \|\hat{x}_{\mathcal{J}} - \bar{x}_{\mathcal{J}} + \bar{x}_{\mathcal{J}}\| \\ &\leq \|\hat{x}_{\mathcal{J}} - \bar{x}_{\mathcal{J}}\| + \underbrace{\|\bar{x}_{\mathcal{J}}\|}_{=0} \\ &\leq \|\hat{x}_{\mathcal{J}} - \bar{x}_{\mathcal{J}}\| \\ &\leq \delta\end{aligned}$$

By choosing the constant ζ independently of the active set.

$$\zeta = \max\left\{1, \|A_{1\bullet}\|, \dots, \left\|A_{p\bullet}\right\|\right\}$$

we see that

$$\left\|A_{j\bullet}\hat{x} - b_j\right\| \leq \left\|A_{j\bullet}\right\| \delta \leq \zeta\delta \ \forall j \in \mathcal{J}$$

and

$$\|\hat{x}_{\mathcal{J}}\| \leq \delta \leq \zeta\delta.$$

This shows that $\hat{x} \in S_\varepsilon(\mathscr{A})$ for $\varepsilon = \zeta\delta$. Since $\hat{x}$ was also in $\tilde{X}_{\text{opt}}(\lambda) \cap \tilde{X}_E$ we see that $\lambda \in \Omega_\varepsilon(\mathscr{A})$. □

Unfortunately, the assumptions of Proposition 5.15 are quite strong, and it is not clear if Algorithm 5.2 computes an outer approximation that satisfies these assumptions for a sufficiently small error bound η^*. For this reason we formulate the following conjecture:

Conjecture 5.16. Let $\delta > 0$ be given. Then there exists $\eta^* > 0$ such that Algorithm 5.2 computes approximating points sets Z^i with approximation errors smaller than η^* for each $i = 1, \dots, m$ such that the efficient set $\tilde{X}_E$ of the outer approximation (5.21) and the set optimal solutions $\tilde{X}_{\text{opt}}(\lambda)$ of the weighted sum problem (5.22) for $\lambda \in \Lambda$ satisfy

$$d(\bar{x}(\lambda), \tilde{X}_{\text{opt}}(\lambda) \cap \tilde{X}_E) \leq \delta \ \forall \lambda \in \Lambda$$

where $\bar{x}(\lambda)$ is the optimal solution of the weighted sum problem of (5.20) for $\lambda \in \Lambda$.

In order to proof a proposition similar to Conjecture 5.16 more assumption may be necessary. Additionally, other approximation error for Algorithm 5.2 may be considered.
Bertsekas and Yu [6] provide a framework for the approximation of singleobjective optimization problems and a convergence result that is similar to the assumption of Proposition 5.15. It may be possible to extend this result to the weighted sum problem.

We consider a triobjective strictly convex ooptimization problem to visualize the approximation procedure:

Example 5.17. Consider the triobjective convex quadratic programming problem from Example 4.47:

$$\begin{aligned}
\min \quad & f_1(x) = \tfrac{1}{2}x^T \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} x + \begin{pmatrix} -40 \\ -24 \end{pmatrix}^T x \\
\min \quad & f_2(x) = \tfrac{1}{2}x^T \begin{pmatrix} 8 & 2 \\ 2 & 16 \end{pmatrix} x + \begin{pmatrix} -34 \\ -24 \end{pmatrix}^T x \\
\min \quad & f_3(x) = \tfrac{1}{2}x^T \begin{pmatrix} 6 & 3 \\ 3 & 12 \end{pmatrix} x + \begin{pmatrix} -27 \\ -3 \end{pmatrix}^T x \\
\text{s.t.} \quad & \begin{pmatrix} -1 & -1 \\ 0 & -1 \\ -1 & 1 \end{pmatrix} x \geqq \begin{pmatrix} -9 \\ -4 \\ -5 \end{pmatrix} \\
& x \geqq 0
\end{aligned} \tag{5.23}$$

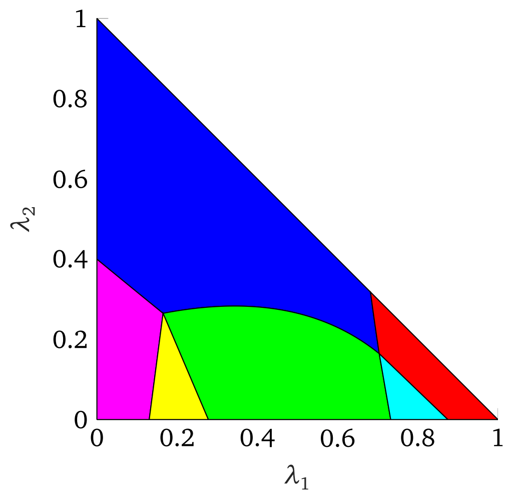

Figure 5.6: Weight Space Decomposition of (5.23).

The weight space decomposition of (5.23) is depicted in Figure 5.6. In a first step we approximate the nadir point of (5.23) by the evaluation of objective function values on a dense grid of sample points.

$$y^N = \begin{pmatrix} -146,89 \\ -37,375 \\ 27,125 \end{pmatrix}$$

Using the nonlinear optimization solver fmincon we compute a bounding box X^i of the set $\mathcal{L}(f_i, y_i^N)$ for each $i = 1, \dots, m$.

$$\begin{aligned}
X^1 &= \{x \in \mathbb{R}^n \ : \ 0.10 \leq x_1 \leq 15.89, 3.17 \leq x_2 \leq 19.17\} \\
X^2 &= \{x \in \mathbb{R}^n \ : \ 0.68 \leq x_1 \leq 7.317, 1.34 \leq x_2 \leq 3.35\} \\
X^3 &= \{x \in \mathbb{R}^n \ : \ -0.96 \leq x_1 \leq 10.95, 5.21 \leq x_2 \leq 3.21\}
\end{aligned}$$

The maximal function values of f_1, f_2 and f_3 over X^1, X^2 and X^3 are approximately 81.57, 23.55 and 222.11, respectively.
Algorithm 5.2 was applied to each objective function $f_i(x)$,$i = 1, 2, 3$, for 2000 iterations.
The approximation error and the computation time of Algorithm 5.2 can be found in Table 5.4 and Table 5.5, respectively. The approximation of the weight space decomposition is shown in Figures 5.7 and 5.8.
After about 200 iterations we can observe that the all efficient active sets of (5.23) are present in the weight space decomposition of the approximation problem (cp. Figures 5.6 and 5.7).

Iteration	10	100	200	300	400	500	1000	2000
f_1	40.718	3.474	1.663	1.097	0.816	0.644	0.322	0.160
f_2	12.984	1.279	0.617	0.408	0.307	0.245	0.121	0,058
f_3	34.752	3.196	1.427	0.952	0.707	0.540	0.275	0.137

Table 5.4: Approximation error

Iteration	10	100	200	300	400	500	1000	2000
f_1	0.008	0.410	2.031	5.777	12.745	23.997	190.939	1660.136
f_2	0.007	0.409	2.018	5.746	12.690	23.965	189.124	1623.037
f_3	0.007	0.402	1.987	5.633	12.355	23.208	183.360	1579.276

Table 5.5: Time in seconds

In Figure 5.8 we can observe that the weight cells of the multiobjective piecewise-linear problem are devided into smaller cells with increasing number of approximating points. The boundaries between the weight cells of the approximation problem resemble the boundaries of the weight cells of (5.23). However, consider the final approximation after 2000 iterations depicted in the bottom right of Figure 5.8. Even though the approximation error relatively small, for example, with $\eta_{2000} = 0.16$ for the first objective function with a cut-off level of 81.57 the weight space decomposition still differs significantly from the weight space decomposition of (5.23).

We can observe that some of the weight cells of the outer approximation problem (5.21) are rather long cells. This can be attributed to the fact that the original problem has only 2 variables. The magenta, green and red cells correspond to active sets for which the corresponding efficient points lie on lines in $\mathbb{R}^2$ at the boundary of the efficient set (the decomposition of the efficient set can be seen in Figure 4.6). The yellow and cyan cells correspond to a single points, respectively.
In conclusion, we see that the approximation approach does indeed provide an approximation of the weight space decomposition of (5.23) after a relatively small number of iterations. However, for larger number of iterations the approximation becomes only slightly more accurate.

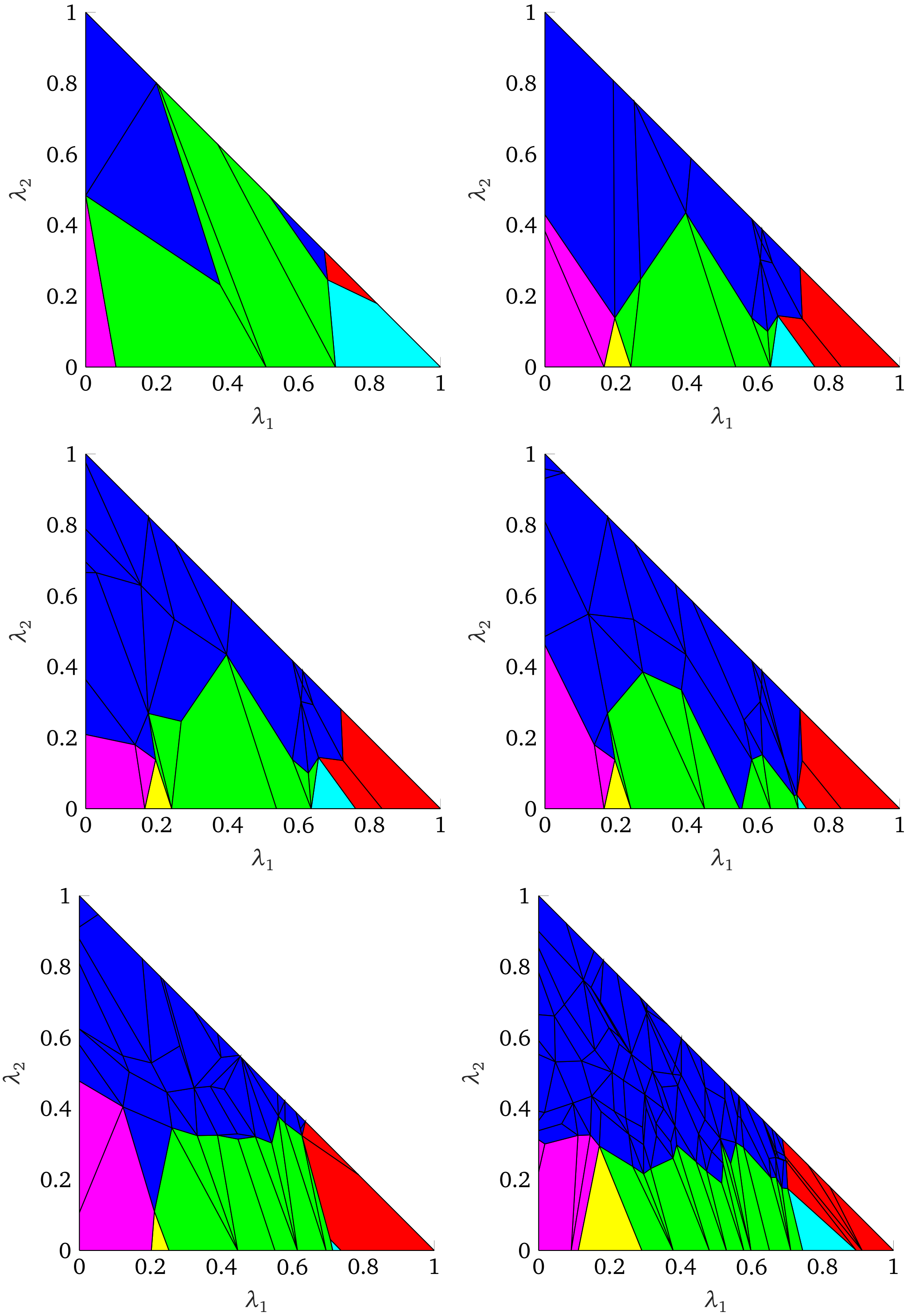

Figure 5.7: Approximation of the Weight Space Decomposition after 10, 20, 30, 40, 80 and 200 Iterations (from top left to bottom right).

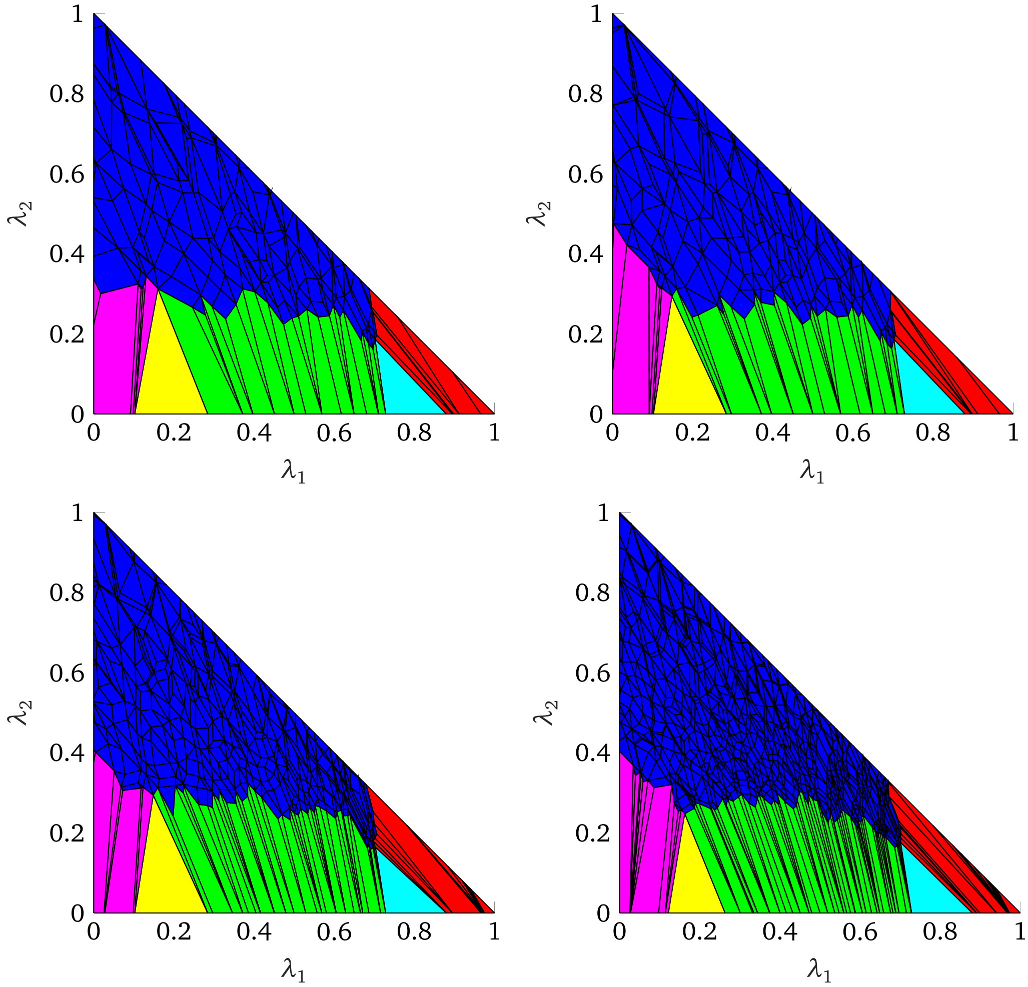

Figure 5.8: Approximation of the Weight Space Decomposition after 400, 500, 1000 and 2000 iterations (from top left to bottom right.

5.6 Conclusion

In this chapter we have defined a weight space decomposition for convex multiobjective optimization problems. We have shown that the weight space of multiobjective convex piecewise-linear optimization problems can be computed using a multiobjective linear programming problem. The corresponding weight cells are unions of convex polyhedra and can be computed using known methods for multiobjective linear programming.
An approximation procedure proposed by Oberdieck and Pistikopoulos [67] was reviewed and using results from the field of approximation of convex bodies by polyhedra [50] a convergence rate was shown.
Finally, using a concrete example we have seen that the approximation procedure by Oberdieck and Pistikopoulos [67] does produce a good approximation.

One interesting question for future research is the analysis of a measure of appropriation error in the weight space and the relationship with the approximation error used in Algorithm 5.2.
Furthermore, it is interesting to consider other types of multiobjective convex optimization problems as approximation problems. The main premise of the approximation scheme discussed in Section 5.5 was to compute an approximation of the weight space of a given multiobjective convex quadratic optimization problem by computing the weight space decomposition of a multiobjective convex piecewise-linear optimization problem. Methods for approximating the weight space directly may provide better results. An interesting approach was proposed by Ruzika and Halffmann [75] using a point approximation of the weight space.

Chapter 6

Conclusion

In this thesis several multiobjective optimization problems and solution techniques have been considered.
In Chapter 3 we reviewed a class of multiobjective descent algorithms that can be interpreted as an extension of singleobjective descent methods. We introduced the class of weight-based descent methods, in particular the compromise descent method.

In Chapter 4 we investigated properties of multiobjective convex optimization problems. In particular, we formulated a weight space decomposition by efficient complementary bases of the parametric linear complementarity system that arises from the KKT conditions of the weighted sum problem and showed that this approach is equivalent to the concept of efficient active sets. We have suggested an algorithm to compute all efficient complementary bases.
Furthermore, we have discussed three particular cases for which the weight cells of multiobjective convex optimization problems are convex polyhedra. For multiobjective convex optimization problems with diagonal objective matrices and lower and upper bounds a stronger result was shown: In this case the weight space is an arrangement of hyperplanes and the number of efficient active set can be bounded by a polynomial in the number of variables and objectives.
We have also considered a parameter space decomposition for the e-constraint scalarization and an application to a problem from the field of location analysis.

In Chapter 5 we defined a weight space decomposition for general convex multiobjective problems and applied this definition to multiobjective convex piecewise-linear optimization problems. We reviewed an approach for the approximation of the weight space decomposition for multiobjective convex optimization problems using an outer approximation by convex piecewise-linear functions of the objective functions. A convergence result was shown for this algorithm.
An interesting question for future research is the generalization of the weight space decomposition to other multiobjective convex problems.

Bibliography

[1] Nathan Adelgren. *Solution Techniques for Classes of Biobjective and Parametric Programs*. PhD thesis, Clemson University, 2016.

[2] Nathan Adelgren and Margaret M. Wiecek. A two-phase algorithm for the multiparametric linear complementarity problem. *European Journal of Operational Research*, 254(3):715 – 738, 2016.

[3] Arif A. Akkeles, László Balogh, and Tibor Illés. New variants of the criss-cross method for linearly constrained convex quadratic programming. *European Journal of Operational Research*, 157(1):74–86, 2004.

[4] Jonathan F. Bard. *Practical Bilevel Optimization*. Springer US, 1998.

[5] Arjan B. Berkelaar, Benjamin Jansen, Kees Roos, and Tamás Terlaky. Basis- and partition identification for quadratic programming and linear complementarity problems. *Mathematical Programming*, 86(2):261–282, 1999.

[6] Dimitri P. Bertsekas and Huizhen Yu. A unifying polyhedral approximation framework for convex optimization. *SIAM Journal on Optimization*, 21(1):333–360, 2011.

[7] Rasmus Bokrantz and Anders Forsgren. An algorithm for approximating convex pareto surfaces based on dual techniques. *INFORMS Journal on Computing*, 25(2):377–393, 2013.

[8] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[9] Sebastiaan Breedveld. *Towards Automated Treatment Planning in Radiotherapy: A Mathematical Approach to Automated and Integrated Multi-Criterial Optimization of Beam Angles and IMRT Fluence Profiles*. PhD thesis, E, March 2013.

[10] Dimo Brockhoff and Eckart Zitzler. Are all objectives necessary? On dimensionality reduction in evolutionary multiobjective optimization. In Thomas Philip Runarsson, Hans-Georg Beyer, Edmund Burke, Juan J. Merelo-Guervós, L. Darrell Whitley, and Xin Yao, editors, *Parallel Problem Solving from Nature - PPSN IX: 9th International Conference, Reykjavik, Iceland, September 9-13, 2006, Proceedings*, pages 533–542. Springer Berlin Heidelberg, 2006.

[11] E. M. Bronstein. Approximation of convex sets by polytopes. *Journal of Mathematical Sciences*, 153(6):727–762, 2008.

[12] R. C. Buck. Partition of space. *The American Mathematical Monthly*, 50(9):541–544, 1943.

[13] Sebastiano Columbano, Komei Fukuda, and Colin N Jones. An output-sensitive algorithm for multi-parametric LCPs with sufficient matrices. *Polyhedral Computation*, 48:73, 2009.

[14] R. Cottle, J. Pang, and R. Stone. *The Linear Complementarity Problem*. Society for Industrial and Applied Mathematics, 2009.

[15] R.W. Cottle, J.-S. Pang, and V. Venkateswaran. Sufficient matrices and the linear complementarity problem. *Linear Algebra and its Applications*, 114-115:231 – 249, 1989.

[16] Etienne de Klerk. *Aspects of Semidefinite Programming*. Springer US, 2002.

[17] Jean-Antoine Désidéri. MGDA II: A direct method for calculating a descent direction common to several criteria. Research Report RR-7922, INRIA, April 2012.

[18] Jean-Antoine Désidéri. Mgda variants for multi-objective optimization. Research Report RR-8068, INRIA, September 2012.

[19] Jean-Antoine Désidéri. Multiple-gradient descent algorithm for multiobjective optimization. In Josef Eberhardsteiner, editor, *European Congress on Computational Methods in Applied Sciences and Engineering*, Vienne, Austria, 2012. Institute for Mechanics of Materials and Structures.

[20] Jean-Antoine Désidéri. Multiple-gradient descent algorithm (mgda) for multiobjective optimization. *Comptes Rendus Mathematique*, 350(5-6):313 – 318, 2012.

[21] Zvi Drezner and Horst W Hamacher. *Facility Location: Applications and Theory*. Springer Science & Business Media, 2001.

[22] Herbert Edelsbrunner. *Algorithms in Combinatorial Geometry*, volume 10. Springer Science & Business Media, 2012.

[23] Matthias Ehrgott. *Multicriteria Optimization*. Springer Science & Business Media, 2006.

[24] Matthias Ehrgott, Horst W. Hamacher, Kathrin Klamroth, Stefan Nickel, Anita Schöbel, and Margaret M. Wiecek. Equivalence of balance points and pareto solutions in multiple-objective programming. *Journal of Optimization Theory and Applications*, 92(1):209–212, 1997.

[25] Matthias Ehrgott, Andreas Löhne, and Lizhen Shao. A dual variant of benson's outer approximation algorithm for multiple objective linear programming. *Journal of Global Optimization*, 52(4):757–778, 2012.

[26] Matthias Ehrgott, Lizhen Shao, and Anita Schöbel. An approximation algorithm for convex multi-objective programming problems. *Journal of Global Optimization*, 50(3):397–416, 2011.

[27] Matthias Ehrgott and Dagmar Tenfelde-Podehl. Computation of ideal and nadir values and implications for their use in MCDM methods. *European Journal of Operational Research*, 151(1):119–139, nov 2003.

[28] Ya Ping Fang, Kaiwen Meng, and Xiao Qi Yang. Piecewise linear multicriteria programs: The continuous case and its discontinuous generalization. *Operations Research*, 60(2):398–409, 2012.

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

[30] Jörg Fliege and Benar Fux Svaiter. Steepest descent methods for multicriteria optimization. *Mathematical Methods of Operations Research*, 51(3):479–494, 2000.

[31] Jörg Fliege and A. Ismael F. Vaz. A method for constrained multiobjective optimization based on sqp techniques. *SIAM Journal on Optimization*, 26(4):2091–2119, 2016.

[32] Richard L. Francis and A. Victor Cabot. Properties of a multifacility location problem involving euclidian distances. *Naval Research Logistics Quarterly*, 19(2):335–353, jun 1972.

[33] M. Friedlander and P. Tseng. Exact regularization of convex programs. *SIAM Journal on Optimization*, 18(4):1326–1350, 2008.

[34] Ellen H. Fukuda and Luis Mauricio Grana Drummond. A survey on multiobjective descent methods. *Pesquisa Operacional*, 34:585 – 620, 12 2014.

[35] Komei Fukuda and Alain Prodon. Double description method revisited. In Michel Deza, Reinhardt Euler, and Ioannis Manoussakis, editors, *Combinatorics and Computer Science*, pages 91–111. Springer Berlin Heidelberg, 1996.

[36] Carl Geiger and Christian Kanzow. *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben*. Springer-Verlag, 2013.

[37] Carl Geiger and Christian Kanzow. *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer-Verlag, 2013.

[38] C.J. Goh and X.Q. Yang. Analytic efficient solution set for multi-criteria quadratic programs. *European Journal of Operational Research*, 92(1):166 – 181, 1996.

[39] Peter M. Gruber. *Approximation of convex bodies*, pages 131–162. Birkhäuser Basel, Basel, 1983.

[40] Horst W Hamacher and Kathrin Klamroth. *Lineare Optimierung und Netzwerkoptimierung*. Springer, 2006.

[41] Horst. W. Hamacher and Stefan Nickel. Multicriteria planar location problems. *European Journal of Operational Research*, 94(1):66 – 86, 1996.

[42] Pierre Hansen. Bicriterion path problems. In Günter Fandel and Tomas Gal, editors, *Multiple Criteria Decision Making Theory and Application*, pages 109–127, Berlin, Heidelberg, 1980. Springer Berlin Heidelberg.

[43] M. Herceg, M. Kvasnica, C.N. Jones, and M. Morari. Multi-Parametric Toolbox 3.0. In *Proc. of the European Control Conference*, pages 502–510, Zürich, Switzerland, July 17–19 2013. `http://control.ee.ethz.ch/~mpt`.

[44] Markus Hirschberger, Yue Qi, and Ralph E. Steuer. Quadratic parametric programming for portfolio selection with random problem generation and computational experience department of banking and finance. unpublished, 2004.

[45] Markus Hirschberger, Yue Qi, and Ralph E. Steuer. Large-scale {MV} efficient frontier computation via a procedure of parametric quadratic programming. *European Journal of Operational Research*, 204(3):581 – 588, 2010.

[46] Markus Hirschberger, Ralph E. Steuer, Sebastian Utz, Maximilian Wimmer, and Yue Qi. Computing the nondominated surface in tri-criterion portfolio selection. *Operations Research*, 61(1):169–183, 2013.

[47] S. Huband, P. Hingston, L. Barone, and L. While. A review of multiobjective test problems and a scalable test problem toolkit. *IEEE Transactions on Evolutionary Computation*, 10(5):477–506, Oct 2006.

[48] Heinz Isermann. Proper efficiency and the linear vector maximum problem. *Operations Research*, 22(1):189–191, 1974.

[49] Henrik Juel and Robert F. Love. Hull properties in location problems. *European Journal of Operational Research*, 12(3):262 – 265, 1983.

[50] George Kamenev. A class of adaptive algorithms for approximating convex bodies by polyhedra. *Computational Mathematics and Mathematical Physics*, 32, 01 1992.

[51] George Kamenev. Analysis of an algorithm for approximating convex bodies. *Computational Mathematics and Mathematical Physics*, 34:521–528, 04 1994.

[52] Kathrin Klamroth and Tind Jørgen. Constrained optimization using multiple objective programming. *Journal of Global Optimization*, 37(3):325–355, jul 2006.

[53] Kathrin Klamroth, Jørgen Tind, and Margaret M. Wiecek. Unbiased approximation in multicriteria optimization. *Mathematical Methods of Operations Research*, 56(3):413–437, 2003.

[54] Masakazu Kojima, Shinji Mizuno, and Akiko Yoshise. A polynomial-time algorithm for a class of linear complementarity problems. *Mathematical Programming*, 44(1):1–26, May 1989.

[55] Gilbert Laporte, Stefan Nickel, and Francisco Saldanha da Gama, editors. *Location Science*. Springer International Publishing, 2015.

[56] Rafael Lazimy. Interactive polyhedral outer approximation (ipoa) strategy for general multiobjective optimization problems. *Annals of Operations Research*, 210(1):73–99, Nov 2013.

[57] C. E. Lemke. A method of solution for quadratic programs. *Management Science*, 8(4):442–453, 1962.

[58] Andreas Löhne and Benjamin Weißing. The vector linear program solver bensolve – notes on theoretical background. *European Journal of Operational Research*, 260(3):807–813, aug 2017.

[59] Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

[60] Harry Markowitz. The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly*, 3(1-2):111–133, 1956.

[61] Istvan Maros and Csaba Meszaros. A repository of convex quadratic programming problems. *Optimization Methods and Software (OMS)*, 11, 12:671–681, December 1999.

[62] Kaisa Miettinen. *Nonlinear Multiobjective Optimization*, volume 12. Springer Science & Business Media, 2012.

[63] Katta G. Murty and Santosh N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39(2):117–129, jun 1987.

[64] Stefan Nickel and Justo Puerto. *Location Theory*. Springer-Verlag, 2005.

[65] Stefan Nickel and Margaret M Wiecek. Multiple objective programming with piecewise linear functions. *Journal of Multicriteria Decision Analysis*, 8(6):322, 1999.

[66] Jorge Nocedal and Stephen Wright. *Numerical Optimization*, volume 35 of *Springer Series in Operations Research*. Springer, second edition edition, 2006.

[67] Richard Oberdieck and Efstratios N. Pistikopoulos. Multi-objective optimization with convex quadratic cost functions: A multi-parametric programming approach. *Computers & Chemical Engineering*, 85:36 – 39, 2016.

[68] Panos M Pardalos and H Edwin Romeijn. *Handbook of Global Optimization*, volume 2. Springer Science & Business Media, 2013.

[69] B. Pelegrin and F. R. Fernandez. Determination of efficient points in multiple-objective location problems. *Naval Research Logistics*, 35(6):697–705, dec 1988.

[70] A. Povalej. Quasi-newton method for multiobjective optimization. *Journal of Computational and Applied Mathematics*, 255:765 – 777, 2014.

[71] Anthony Przybylski, Xavier Gandibleux, and Matthias Ehrgott. A recursive algorithm for finding all nondominated extreme points in the outcome set of a multiobjective integer programme. *INFORMS Journal on Computing*, 22(3):371–386, 2010.

[72] Anthony Przybylski, Kathrin Klamroth, and Renaud Lacour. A simple and efficient dichotomic search algorithm for multi-objective integer linear programmes. unpublished as of july 2019, 2017.

[73] R Tyrrell Rockafellar. *Convex Analysis, volume 28 of Princeton Mathematics Series*. Princeton University Press, 1970.

[74] S. Ruzika and M. M. Wiecek. Approximation methods in multiobjective programming. *Journal of Optimization Theory and Applications*, 126(3):473–501, sep 2005.

[75] Stefan Ruzika and Pascal Halffmann. Can we approximate a weight set decomposition? Recent Advances in Multi-Objective Optimization, 2018. talk.

[76] Jörg-Rüdiger Sack and Jorge Urrutia. *Handbook of Computational Geometry*. Elsevier, 1999.

[77] Serpil Sayın. Measuring the quality of discrete representations of efficient sets in multiple objective mathematical programming. *Mathematical Programming*, 87(3):543–560, 2000.

[78] Serpil Sayin. A procedure to find discrete representations of the efficient set with specified coverage errors. *Operations Research*, 51(3):427–436, 2003.

[79] Britta Schulze, Kathrin Klamroth, and Michael Stiglmayr. Multi-objective unconstrained combinatorial optimization: a polynomial bound on the number of extreme supported solutions. *Journal of Global Optimization*, 74(3):495–522, 2019.

[80] Mohit Tawarmalani and Nikolaos V. Sahinidis. A polyhedral branch-and-cut approach to global optimization. *Mathematical Programming*, 103(2):225–249, 2005.

[81] Richard E. Wendell and Arthur P. Hurter. Location theory, dominance, and convexity. *Operations Research*, 21(1):314–320, 1973.

[82] Elizabeth Lai Sum Wong. *Active-set Methods for Quadratic Programming*. PhD thesis, University of California, San Diego, 2011.

[83] Eckart Zitzler, Kalyanmoy Deb, and Lothar Thiele. Comparison of multiobjective evolutionary algorithms: Empirical results. *Evolutionary computation*, 8(2):173–195, 2000.