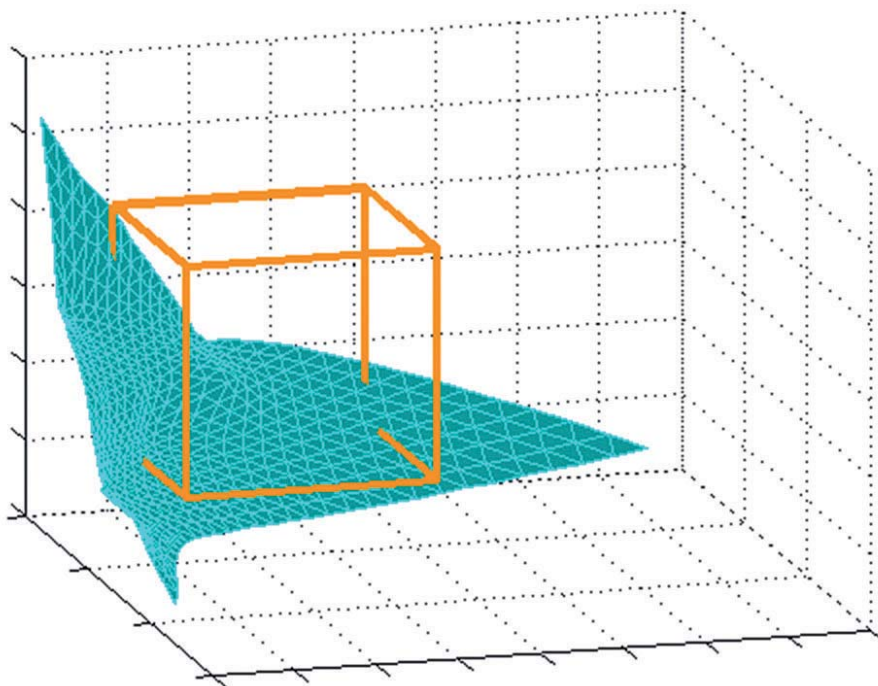Tomi Haanpää

Approximation Method for Computationally Expensive Nonconvex Multiobjective Optimization Problems

Tomi Haanpää

Approximation Method for Computationally Expensive Nonconvex Multiobjective Optimization Problems

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2012

Approximation Method for Computationally Expensive Nonconvex Multiobjective Optimization Problems

Tomi Haanpää

Approximation Method for Computationally Expensive Nonconvex Multiobjective Optimization Problems

UNIVERSITY OF JYVÄSKYLÄ

ABSTRACT

The importance of optimization for decision makers solving real-life engineering problems has increased in recent decades. In the literature, methods created for multiobjective optimization problems are typically directed for problems where objective function values can be calculated fast. If function evaluations are computationally expensive, i.e. time-consuming, then only few methods are available and, usually, they involve assumptions like convex objective functions. To be able to solve more demanding multiobjective optimization problems in future, both the nonconvexity and the computational cost of the problem in question must be considered. In this thesis, an approximation method, Interactive HyperBox Exploration (IHBE), is proposed for computationally expensive nonconvex multiobjective optimization problems.

The IHBE method has been developed so that a decision maker can conveniently investigate the objective space and learn about possibilities and limitations of solutions which are available for the multiobjective optimization problem at hand. In order to lessen the computational expense of the problem, an approximation of the Pareto front is used in the method. The approximation is based on a surrogate function and pre-calculated Pareto optimal objective vectors. By specifying desirable upper and lower bounds for the vectors, the decision maker makes queries about what kind of approximated Pareto optimal objective vectors are available. The method either presents approximated Pareto optimal objective vectors satisfying the bounds or helps the decision maker in setting the bounds in a way that approximated Pareto optimal objective vectors can be found.

Keywords: Multiobjective optimization, computational cost, computational efficiency, Pareto front approximation, surrogate function, interactive decision making, decision maker, psychological convergence, Pareto dominance, Pareto optimality.

Author   Tomi Haanpää
                    Department of Mathematical Information Technology
                    University of Jyväskylä
                    Finland

Supervisors   Professor Kaisa Miettinen
                    Department of Mathematical Information Technology
                    University of Jyväskylä
                    Finland

                    Senior Assistant Jussi Hakanen
                    Department of Mathematical Information Technology
                    University of Jyväskylä
                    Finland

Reviewers   Professor David Rios Insua
                    Department of Statistic and Operations Research
                    Rey Juan Carlos University
                    Spain

                    Professor emeritus Theodor Stewart
                    Department of Statistical Sciences
                    University of Cape Town
                    South Africa

Opponent   Professor emeritus Pekka Korhonen
                    Department of Information and Service Economy
                    Aalto University School of Business
                    Finland

ACKNOWLEDGEMENTS

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

Multiobjective optimization [123] has been created for real-life problems in which several conflicting goals, such as the cost of the product to be minimized and the profit of the product to be maximized, are to be considered simultaneously. Increased computation power has motivated engineers in various real-life application areas to model more and more complex systems mathematically in an attempt to find an optimal (i.e. "the best") solution to a problem for which it is not yet known. In practice, the complexity means that a great number of calculations must be done in order to solve the problem. Typically, for the calculations, simulators are used, and a single simulation can take a considerable amount of time, see e.g. [5, 60]. In this thesis, problems for which the calculations are *time-consuming* are called *computationally expensive (costly) problems*.

Methods created in the seventies, eighties and nineties for multiobjective optimization problems are not necessarily in themselves best suited for today's computationally expensive problems. During the 21st century, multiobjective optimization using evolutionary algorithms [33, 43, 44] has become popular and commonly used, but the applicability of the methods for computationally expensive multiobjective problems still requires further developments. The time required to solve these problems is still a challenge for the methods and algorithms. In the worst case, it can take several days for the methods to solve the problem because of the time-consuming calculations. This may not be acceptable, for example, in radio therapy treatment planning [158]. However, in numerical mathematics and statistics, so-called *surrogate (regression) functions* [28, 53, 160, 181] have been developed for problems in which estimates are required. The estimate can concern, for example, the growth of the population of a country over certain years or even in the future based on some numeric data. This leads to the topic of this thesis.

In this thesis, a new method, Interactive HyperBox Exploration (IHBE), is proposed for computationally expensive multiobjective problems. The method is based on a surrogate function which *approximates* (generates educated guesses of) optimal solutions to the problem at hand. In fact, the use of surrogate functions in multiobjective optimization is introduced in a new and innovative way in

this thesis and opens doors for methods to be used for computationally expensive multiobjective optimization problems. These methods were originally created for computationally inexpensive problems, for which calculation of solutions is fast. To mention some advantages of the IHBE method, *approximated solutions* for computationally expensive multiobjective optimization problems can be calculated in a reasonable time, and numerical details such as approximated trade-off rates for the solutions can be calculated without a need for long computation times.

In what follows, first, in Section 1.1, the mathematical notation and terminology used in this thesis are fixed. Then, the basics of multiobjective optimization, interactive methods and the decision maker are briefly discussed in Section 1.2. In order to beat the time challenge in multiobjective optimization problems, so-called *Pareto front* of the multiobjective optimization problem is approximated by a surrogate function. This is the topic in Section 1.3. Then, surrogate functions are briefly discussed in Section 1.4. Finally, in Section 1.5, the structure and the contributions of this thesis are outlined.

1.1 Mathematical Notation and Terminology Used

In this thesis, we assume that the *Euclidian* vector space $\mathbb{R}^n$ is familiar for the reader. A vector $\mathbf{x} \in \mathbb{R}^n$ is denoted by $(x_1, x_2, \ldots, x_n)^T$, where x_j is called the jth coordinate value of the vector $\mathbf{x}$. If there is an index i for a vector $\mathbf{x} \in \mathbb{R}^n$, then the index is denoted by a superscript $\mathbf{x}^i$.

A *dot product* between vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is a real number

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{i=1}^{n} x_i y_i$$

and a *norm* of $\mathbf{x}$ is a real number $\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$. In addition, vectors $\mathbf{x}$ and $\mathbf{y}$ are said to be *perpendicular* if the dot product $\langle \mathbf{x}, \mathbf{y} \rangle$ is equal to zero.

The following sets have important roles in this thesis. The *positive orthant* of $\mathbb{R}^n$ is the set $\mathbb{R}^n_+ := \{\mathbf{x} \in \mathbb{R}^n : x_j \geq 0 \text{ for all } j = 1, \ldots, n\}$. Respectively, the *negative orthant* of $\mathbb{R}^n$ is the set $\mathbb{R}^n_- := \{\mathbf{x} \in \mathbb{R}^n : x_j \leq 0 \text{ for all } j = 1, \ldots, n\}$. Let $\mathbf{x} \in \mathbb{R}^n$ and $r \geq 0$. Then a set $B(\mathbf{x}, r) := \{\mathbf{x}' : \|\mathbf{x} - \mathbf{x}'\| < r\}$ is called an *open ball*. In what follows, let $A \subset \mathbb{R}^n$, $A \neq \emptyset$. A vector $\mathbf{x} \in \mathbb{R}^n$ is called

- an *interior vector* of A if there exists an open ball $B(\mathbf{x}, r)$, $r > 0$, so that $B(\mathbf{x}, r) \subset A$; the set of interior vectors of A (the interior of A) is denoted by int A,
- an *exterior vector* of A if there exists an open ball $B(\mathbf{x}, r)$, $r > 0$, so that $B(\mathbf{x}, r) \cap A = \emptyset$; the set of exterior vectors of A (the exterior of A) is denoted by ext A,
- a *boundary vector* of A if it is neither an exterior nor an interior vector of A; the set of boundary vectors of A (the boundary of A) is denoted by ∂A,
- a *limit vector* of A if for all $r > 0$ there exists $\mathbf{a}^r \in A$ so that $\mathbf{a}^r \in B(\mathbf{x}, r) \setminus \{\mathbf{x}\}$.

The set A is called

- *open* if int $A = A$,
- *closed* if the complement of A, that is, $A^C := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \notin A\}$, is open,
- *bounded* if there exist $r > 0$ and $\mathbf{x} \in \mathbb{R}^n$ so that $A \subset B(\mathbf{x}, r)$,
- *compact* if A is closed and bounded,
- an *open neighborhood* of $\mathbf{x} \in \mathbb{R}^n$ if A is open and $\mathbf{x} \in A$,
- *convex* if every *closed line segment* $[\mathbf{a}^1, \mathbf{a}^2] := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \lambda \mathbf{a}^1 + (1 - \lambda)\mathbf{a}^2, \lambda \in [0,1]\}$ of A, where $\mathbf{a}^1, \mathbf{a}^2 \in A$, is a subset of A,
- *nonconvex* if it is not convex,
- *disconnected* if there exist nonempty open sets $B, C \subset \mathbb{R}^n$ so that $A \cap B \neq \emptyset$, $A \cap C \neq \emptyset$, $B \cap C \neq \emptyset$ and $A \subset B \cup C$,
- *connected* if it is not disconnected.

For a set $A \subset \mathbb{R}^n$, the following sets are essential. The *closure* of A is the set $\overline{A} := A \cup \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \text{ is a limit vector of } A\}$. The *convex hull* of A is a set $\mathrm{CH}\, A := \bigcap B$, $B \subset \mathbb{R}^n$ is a convex set and $A \subset B$. A subset B of A is called *relatively open*, if there exists an open set $C \subset \mathbb{R}^n$ so that $A \cap C = B$. Similarly, a subset B of A is called *relatively closed*, if there exists a closed set $C \subset \mathbb{R}^n$ so that $A \cap C = B$.

A function $f : \mathbb{R}^n \to \mathbb{R}$ is continuous if for each open set $B \subset \mathbb{R}$ the preimage $f^{\{-1\}}(B) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \in B\}$ is open. Function $f : \mathbb{R}^n \to \mathbb{R}^k$ is said to be differentiable at $\mathbf{a} \in \mathbb{R}^n$ if there exists a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^k$ so that

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{\|f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - L(\mathbf{h})\|}{\|\mathbf{h}\|} = 0.$$

The linear transformation L is denoted by $Df(\mathbf{a})$ and called *derivative* of f at $\mathbf{a}$. If the derivate exists for all $\mathbf{a} \in A$, then function f is called *differentiable* on A.

Let a function f be real valued (i.e. $f : \mathbb{R}^n \to \mathbb{R}$) and $\mathbf{a} \in \mathbb{R}^n$. If the limit

$$\lim_{h \to 0} \frac{f((a_1, \ldots, a_i + h, \ldots, a_n)^{\mathrm{T}}) - f((a_1, \ldots, a_i, \ldots, a_n)^{\mathrm{T}})}{h} =: \frac{\partial f}{\partial a_i}(\mathbf{a})$$

exists, then the limit is called the ith *partial derivative* of f at $\mathbf{a}$. The vector of partial derivatives at $\mathbf{a}$ is called a *gradient* of f at $\mathbf{a}$ and denoted by $\nabla f(\mathbf{a})$. The function f is called *convex* on A if the *epigraph* $\mathrm{epi}\, f := \{(\mathbf{a}, \alpha) \in A \times \mathbb{R} : f(\mathbf{a}) \geq \alpha\}$ is a convex set. Respectively, function f is called *nonconvex* if it is not convex. On the other hand, function f is called *concave* on A, if the *hypograph* $\mathrm{hyp}\, f := \{(\mathbf{a}, \alpha) \in A \times \mathbb{R} : f(\mathbf{a}) \leq \alpha\}$ is a convex set.

1.2 Multiobjective optimization

In this section, basics of multiobjective optimization are discussed. A problem formulation considered in this thesis is given and interactive methods for multiobjective optimization problems involving a decision maker are briefly introduced.

A multiobjective optimization problem is a problem in which several objective function values are optimized simultaneously [123]. More formally, let S be a nonempty set called a *feasible set* in the space $\mathbb{R}^n$, which is called a *decision (design,variable) space*. Let the function f be a vector-valued function defined on S, that is, $f : S \rightarrow \mathbb{R}^k$,

$$f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x}))^{\mathrm{T}},$$

where real valued functions f_i, $i = 1, \ldots, k$, defined on S are called *objective functions*. The space $\mathbb{R}^k$ is called *objective (outcome) space* and vectors in it are called *objective(outcome) vectors*. A vector $f(\mathbf{x})$ is called a *feasible objective vector (outcome)* and the image $f(S)$ is called a *feasible set* in $\mathbb{R}^k$. An objective vector $\mathbf{z} \in \mathbb{R}^k$ is called *infeasible*, if $\mathbf{z} \notin f(S)$.

In this thesis, the MultiObjective Optimization (MOO) problem considered has the form

$$\min_{\mathbf{x} \in S} f(\mathbf{x}),$$

where every continuous (nonconvex) objective function f_i is assumed to be minimized simultaneously over a (nonconvex) set S. A MOO problem is called *nonconvex* if some of the objective functions is nonconvex or S is nonconvex. Respectively, the problem is *convex* if every objective function is convex and S is convex. A unique set of solutions to the MOO problem (no matter whether or not it is convex) is a set consisting of feasible objective vectors satisfying the so-called *Pareto optimality*.

Definition 1 (Pareto optimality). A decision vector $\mathbf{x} \in S$ is called a *Pareto optimal decision vector* if there does not exist another decision vector $\bar{\mathbf{x}} \in S$ so that $f_i(\bar{\mathbf{x}}) \leq f_i(\mathbf{x})$ for all $i = 1, \ldots, k$ and $f(\bar{\mathbf{x}}) \neq f(\mathbf{x})$. Respectively, a vector $f(\mathbf{x})$ is called a *Pareto optimal objective vector*, if the vector $\mathbf{x}$ is a Pareto optimal decision vector. We call the set of Pareto optimal objective vectors on $f(S)$ a *Pareto front*, and we denote it by $\mathrm{PF}(f, S)$.

In Figure 1, we have illustrated a Pareto front $\mathrm{PF}(f, S)$ of a nonconvex MOO problem by black bold curves. Note that the front may contain infinitely many vectors, which are in the mathematical sense equally good.

The *ideal vector* of the Pareto front is

$$\mathbf{ideal} := \left(\inf_{\mathbf{z} \in \mathrm{PF}(f,S)} z_1, \ldots, \inf_{\mathbf{z} \in \mathrm{PF}(f,S)} z_k \right)^{\mathrm{T}}.$$

Respectively, the *nadir vector* of the Pareto front is

$$\mathbf{nadir} := \left(\sup_{\mathbf{z} \in \mathrm{PF}(f,S)} z_1, \ldots, \sup_{\mathbf{z} \in \mathrm{PF}(f,S)} z_k \right)^{\mathrm{T}}.$$

For vector-wise comparisons in subsets of the objective space, a concept of dominance is essential and in this thesis we consider the following:
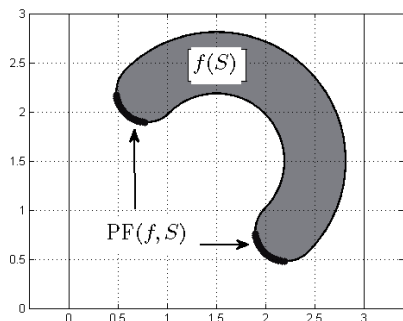
FIGURE 1   A Pareto front $\mathrm{PF}(f, S)$ illustrated in $\mathbb{R}^2$.

Definition 2 (Pareto dominance)**.** Let Z be a nonempty subset of the objective space. A vector $\mathbf{z} \in Z$ is said to be *nondominated* in the set Z if there does not exist another vector $\bar{\mathbf{z}} \in Z$ so that $\bar{z}_i \leq z_i$ for all $i = 1, \ldots, k$ and $\bar{\mathbf{z}} \neq \mathbf{z}$. Otherwise, the vector $\mathbf{z}$ is said to be *dominated* (by vector $\bar{\mathbf{z}}$, or vector $\bar{\mathbf{z}}$ dominates vector $\mathbf{z}$) in the set Z. We denote the set of nondominated vectors of Z by Z^*. In addition, we say that the set Z is *nondominated* if $Z = Z^*$.

In this thesis, *solving* the MOO problem means that the entire Pareto front is calculated. However, in practical (engineering type of) MOO problems, there is no need to calculate the entire Pareto front. In other words, a vector in the Pareto front considered to be implemented in practice is called a *compromise solution (preferred solution)* to the MOO problem. In this thesis, the process of selecting the compromise solution in one way or another is called a *solution process*. To calculate a compromise solution, several methods have been proposed in the literature of Multiple Criteria Decision Making (MCDM). To mention the most well-known of these, achievement (scalarizing) function methods [188], the ε-constraint method [30, 59] and the weighted sum method [52, 198] are often used in the literature. However, in this thesis, the search for a compromise solution to the MOO problem with the IHBE method is done *interactively* involving a human *decision maker*. Here, the word interactively means that the IHBE method is classified to the fourth class of the methods presented in [123] (1. no-preference methods, 2. a posteriori methods, 3. a priori methods, 4. interactive methods). The decision maker is assumed to be an expert on the application area related to the MOO problem. In other words, the decision maker knows and understands the meanings of objective function values of the problem in practice.

In the literature, well-known interactive methods include GUESS [24], NIMBUS [124, 126, 129], STEM [15] and STOM [135, 136] (for method comparisons, see [9, 23, 25, 36, 82, 103, 125, 183]). To put it briefly, the idea of an interactive method is that the decision maker selects the compromise solution based on his or her *preferences*, which may evolve during the process. Typically, the compromise solution will be found by taking the following steps iteratively [131]):

(1) present a starting Pareto optimal objective vector to the decision maker,

(2) ask for preference information from the decision maker,

(3) calculate one or more new objective vectors according to the preferences and present it or them with additional information, if available, to the decision maker,

(4) ask the decision maker to select the best one so far, and stop if the decision maker is satisfied. Otherwise, go to step (2).

Based on the four basic steps introduced above, the solution process with interactive methods can be divided into two phases [131]: a learning phase and a decision phase, which can also be used iteratively. In this thesis, the IHBE method is focused to the learning phase, meaning that it is especially created for supporting and helping the decision maker in learning about the problem and Pareto optimal objective vectors in relation to the preferences the decision maker may have. Then, based on the facts learned, the decision maker may be able to make a final decision about the compromise solution to the MOO problem while being confident about the decision. In this thesis, the construction of the IHBE method is to be done from the decision maker's point of view in order to create a method that most decision makers would like to use. This means that the method should be easy, understandable and useful in searching for desirable compromise solutions to MOO problems.

It is the decision maker, who uses the IHBE method for searching for the compromise solution. In order to make this possible in practice, an *analyst* (a human or a computer program, see e.g. [123]) is also required and is responsible for the mathematical side of the solution process with the IHBE method.

Now, we have introduced the basics of the multiobjective optimization that are required in the IHBE method. For further details of multiobjective optimization, we refer to [20, 46, 123, 165, 176, 194] and references therein.

1.3 Approximating Pareto Front

In this section, the topic is an approximation to be used in the IHBE method. First, we discuss why some MOO problems are computationally expensive and then we point out the main concern, which may arise when a decision maker is searching for a compromise solution to such a problem. Then, we briefly introduce an idea for an approximation found in the literature that we follow in some sense in this thesis. After this, we set up the goals for the approximation as research objectives. Finally, other approximation methods are briefly discussed.

Through the thesis, it is assumed (unless stated otherwise) that the MOO problem is computationally expensive, meaning that, for any $\mathbf{x} \in S$, it is time-consuming to evaluate $f(\mathbf{x})$. In other words, it may take at least from minutes to hours or even days in the worst case to evaluate a single objective vector $f(\mathbf{x})$. The MOO problem can be computationally expensive if e.g. simulators are used instead of explicitly known objective functions f_i or if the objective functions f_i are solutions of partial differential equations that can be occasionally very diffi-

cult to be solved, see e.g. [5, 6, 60, 61, 62, 68]. Note that to calculate a single Pareto optimal objective vector, several hundreds (maybe thousands) of objective function values must be typically calculated, which introduces waiting times for the decision maker in the solution process of the MOO problem. This point out the main concern in searching compromise solutions to computationally expensive MOO problems: the solution process, due to its time-consuming, may stop without a proper compromise solution, as indicated in [104] (the decision maker can be bored of waiting and then stop the process too early). In order to avoid these long computation times, approximations created for Pareto fronts can be used. The idea in the IHBE method is to mainly calculate *approximated Pareto optimal objective vectors*, instead of true Pareto optimal objective vectors.

In this thesis, we follow the general idea of approximating the Pareto front by a hypersurface as presented, for example, in [118], where the authors propose to use a regression function $f_r : \mathbb{R}^k \to \mathbb{R}$ to create the hypersurface. In their case, the hypersurface is the preimage $f_r^{\{-1\}}(0)$, meaning that an objective vector $\mathbf{z} \in \mathbb{R}^k$ is considered to be an approximated Pareto optimal objective vector if $f_r(\mathbf{z})$ is equal to zero. On the other hand, if $f_r(\mathbf{z}) \neq 0$, then the vector $\mathbf{z}$ is either dominated or infeasible. The advantage of this method can be attributed to the simplicity of evaluating whether or not an objective vector is (can be) a Pareto optimal objective vector. In addition, according to [16, 118, 131], it is easier for the decision maker to consider only the objective vectors in the learning phase. However, the hypersurface can be constructed in different ways, and in this thesis a different approach is proposed. The main technical difference in the approximation proposed in this thesis compared to the approximation in [118] is the characterization of the approximated Pareto front. As said, in [118], the approximation is the preimage $f_r^{\{-1\}}(0)$. In this thesis, a function f_a is constructed so that the image of a specific simple set Δ_{k-1} (to be defined later) is the approximated Pareto front. In other words, the approximated Pareto front is the image (set) $f_a(\Delta_{k-1})$.

From a practical point of view, the mathematical structure of the function f_a allows us to construct computationally efficient methods to be used, for example, in identifying dominated and nondominated vectors in the approximation (the methods are developed and discussed in later chapters). Then, for example, a validation of the approximation can be done computationally efficiently. The methods are simple and they can be applied after modifications to the approximation presented in [118], but the computational costs of these methods may increase. Note that the approximation presented in [118] does not in itself determine whether or not the vectors in the approximation are nondominated.

Another focused topic in this thesis is the actual use of the approximation. The approximation is constructed so that it can be suited to the learning phase in the solution process of the MOO problem without increasing computational costs. The use of the approximation in the IHBE method is discussed from the decision maker's point of view, because, as noted in [65, page 33], this aspect has not been widely considered in constructing approximations for Pareto fronts. In other words, the challenges the decision maker may face during the solution

process of a nonconvex MOO problem and how the approximation can help him or her in such cases are discussed in this thesis. We demonstrate that nonconvex MOO problems can be much more complicated for the decision maker than convex MOO problems.

To get an idea of the approximation proposed in this thesis, we present a simple example, in which we have a continuous nonconvex MOO problem with three objective functions. Let us have a set of pre-calculated Pareto optimal objective vectors to the MOO problem, and let them be located in the objective space as shown in Figure 2 (black dots). Because the MOO problem is continuous, then it is possible that there exist other Pareto optimal objective vectors between the pre-calculated Pareto optimal objective vectors. Then a hypersurface created to approximate the Pareto front (as the set of the vectors between the pre-calculated Pareto optimal objective vectors) can be, for example, the dark surface presented in Figure 3.
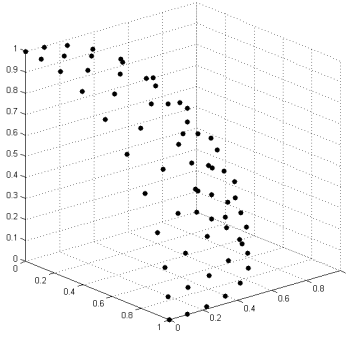
FIGURE 2   A set of Pareto optimal objective vectors in $\mathbb{R}^3$.
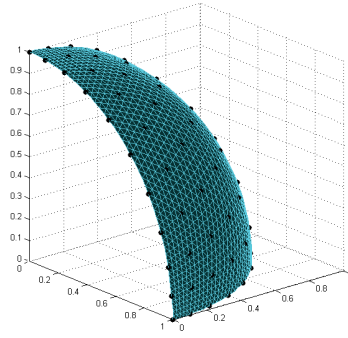
FIGURE 3   A surface including the set of Pareto optimal objective vectors in $\mathbb{R}^3$.

As said, the IHBE method is based on the approximation. Therefore, we list certain goals as research objectives for the approximation. In later chapters, the following goals are referred to by symbols of the form $\mathcal{G}X$, where X indicates the goal number X.

In addition to the formulation of the MOO problem, the approximation must be suitable for the following goals.

$\mathcal{G}1$ The MOO problem considered can consist of more than two objective functions.

$\mathcal{G}2$ The MOO problem considered can be nonconvex.

$\mathcal{G}3$ The MOO problem considered can be computationally expensive.

$\mathcal{G}4$ Only a set consisting of a finite number of known Pareto optimal objective vectors to the MOO problem is available in generating the approximation.

$\mathcal{G}5$ The approximation can be generated computationally efficiently.

$\mathcal{G}6$ The approximation can be validated computationally efficiently.

$\mathcal{G}$7 Generating vectors on the approximation can be done computationally efficiently.

$\mathcal{G}$8 An external interface for specifying the decision maker's preferences can be constructed for the approximation.

$\mathcal{G}$9 The approximation provides additional information for the decision maker in the solution process of the MOO problem.

The first four goals $\mathcal{G}$1-$\mathcal{G}$4 guarantee that the approximation is as general as possible for various continuous MOO problems. Note that although we have assumed the MOO problem to be nonconvex, the approximation can be used also for convex MOO problems. However, the goal $\mathcal{G}$4 can be considered as a minimum requirement to be known for the MOO problem before constructing an approximation for the Pareto front. Then, goals $\mathcal{G}$5-$\mathcal{G}$7 mean in practice that the usage of the approximation is computationally efficient covering the construction and validation of the approximation, and the actual use of it in generating vectors on the approximated Pareto front to be presented for the decision maker in the solution process of the MOO problem. The last two goals, $\mathcal{G}$8-$\mathcal{G}$9, mean that the approximation is more than a theoretical approximation. In other words, it can be used in the IHBE method for learning about the problem. The contents of the goal $\mathcal{G}$9 are clarified later in this thesis.

Next, we briefly review other approximations proposed for the Pareto fronts in the MCDM literature. The simplest ones produce a set consisting of finite number of Pareto optimal objective vectors [7, 33, 42, 43, 121, 166] while in methods such as *'sandwich'* or polyhedral approximations [48, 50, 95, 96, 133, 153] piecewise linear sets (faces) are generated for computationally expensive convex MOO problems. Approximation methods created originally for nonconvex MOO problems are introduced, for example, in [13, 66, 67, 113, 197], in which the idea is to approximate only the set of Pareto optimal objective vectors. Other methods used in the literature approximate the objective functions by surrogate functions (metamodels) [99, 191] or achievement scalarizing functions [97]. Another way is to express one of the coordinates of the Pareto optimal objective vectors by a function of the other coordinates [56, 170], that is, to represent the Pareto front by a graph of a function. By using the surrogate functions, it is possible to have a connection to the decision space, meaning that it can be known approximately how an outcome can be implemented in practice. Note that most of the approximations constructed in the objective space lose the connection to the decision space, that is, for an approximated Pareto optimal objective vector the corresponding decision vector values are not known. Approximations constructed in the decision space also exist and the idea is to approximate the set of Pareto optimal decision vectors, see e.g. [110, 114, 200]. Overviews and discussions of over 50 different approximations can be found e.g. in [65, 159].

In this thesis, comparisons to other approximation methods are not presented for the following reasons. Most of the other approximations generate a discrete set of Pareto optimal objective vectors. The approximation proposed in this thesis requires a discrete set of Pareto optimal objective vectors, which means

that some of the approximation methods are, in fact, a pre-step for the approximation. In other words, there is no sense to compare the approximation to these ones. On the other hand, a comparison between an approximation in [65] and the approximation proposed in this thesis should be presented and discussed. However, for a fair comparison, an independent group of researchers who would like to use the approximations, but who are not developers of the approximations, should make the comparisons and evaluations. Thus the properties measured in the tests would not be biased by the developers, who might be tempted to select indicators, that high-light the positive sides of their approximation and denigrate the other approximations. However, we are honest in this thesis and we discuss openly the possible drawbacks of the approximation used in the IHBE method. Nevertheless, making comparisons between the approximations can be challenging, because the approximation proposed in this thesis is more suited for learning about the problem. This means that it is more than an approximation created only to generate approximated Pareto optimal objective vectors.

Even though the approximation proposed in this thesis is as general as possible for many kinds of continuous MOO problems, it must be clarified that approximations tailored for certain problem types can be much better for their niche. However, the mathematical structure of the approximation proposed in this thesis can provide new aspects for the Pareto fronts of MOO problems for which different kinds of approximations have already been tailored. For example, the approximation can be used in identifying computationally whether or not the Pareto front of a nonconvex MOO problem is disconnected. These aspects are pointed later, when the approximation is plugged to the IHBE method.

1.4   On Surrogate Functions

As said, we construct an approximation for the Pareto front. The approximation is based on the function f_a in which we use a *surrogate function*. Next, we introduce what surrogate functions are.

Surrogate functions (also known as metamodels or response surfaces) are typically real-valued functions defined on $\mathbb{R}^n$, and they are used, for example, in replacing computationally expensive functions. Surrogate functions include radial basis functions [28], support vector regressions [181] and Kriging models as in [160]. Vector-valued surrogate functions include also neural networks [69], but in this thesis, only real-valued surrogate functions are considered.

This section has been organized as follows. First, in Subsection 1.4.1, we present, in a general level, what surrogate functions are, how they are typically used and how they are constructed. Then, in Subsections 1.4.2 and 1.4.3, radial basis functions and Kriging functions are respectively discussed. Finally, in Subsection 1.4.4, we briefly discuss regression functions.

1.4.1 Surrogate Functions Generally

A surrogate function $f_{surrogate}$ typically replaces some other function which can be computationally expensive or a black box function of a simulator. Basically, the idea is that the surrogate function approximates (interpolates, models) some other function or data.

To replace or model a function $g : \mathbb{R}^n \to \mathbb{R}$ or data by a surrogate function $f_{surrogate}$, some pre-calculated data $\mathcal{D}$ consisting of m pairs $(\mathbf{x}^i, y^i) \in \mathbb{R}^n \times \mathbb{R}$, $i = 1, \ldots, m$, must be available. In the previous, vectors $\mathbf{x}^i$ are called *input (design) vectors* and the corresponding real values y^i $(=: g(\mathbf{x}^i))$ are called *output (target) values*. Then the actual replacement can be done by *fitting $f_{surrogate}$* to data $\mathcal{D}$ so that

$$f_{surrogate}(\mathbf{x}^i) = y^i + \varepsilon(\mathbf{x}^i),$$

where $\varepsilon(\cdot)$ is an *error estimation function*.

Depending on the data $\mathcal{D}$, surrogate functions can be divided into two categories: *interpolative* and *regression* functions. If for all $i = 1, \ldots, m$ we have $f_{surrogate}(\mathbf{x}^i) = y^i$, then we say that the function $f_{surrogate}$ is interpolative (the function interpolates). Otherwise, we say that the function $f_{surrogate}$ is a regression function. On the left hand side in Figure 4, we have illustrated a radial basis function (black curve) interpolating a given data $\mathcal{D}$ (gray dots). On the right hand side, a support vector regression has been fitted to *noisy* data. In what follows, we assume that the data used in constructing the approximation does not contain noise. This actually follows from the goal $\mathcal{G}4$ set on page 20. In other words, we do not have to consider regression functions for noisy data. However, with certain selections, regression functions might interpolate. Therefore, they are discussed here because we need a surrogate function that interpolates the data.
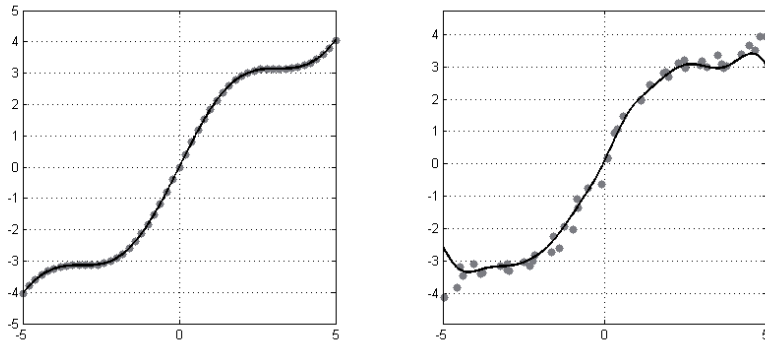
FIGURE 4 On the left hand side, a radial basis function has been fitted to data, which is not noisy. If the data is noisy, then a regression function can be used, as on the right hand side, where a support vector regression has been fitted to the data.

A common feature in all surrogate functions is their computational efficiency in evaluating function values, that is, the wall-clock time for an evalua-

tion of a function value $f_{surrogate}(\mathbf{x})$ is significantly short. This means that the use of surrogate functions in the approximation is justified if we consider the goal $\mathcal{G}7$, meaning that generating vectors in the approximation should not be time-consuming. However, to enjoy of the fast calculations by surrogate functions, they must be fitted to the data.

Typically, fitting a surrogate function to the data means that values for some unknown constants or parameters included to the function must be selected or optimized. However, if we consider the goal $\mathcal{G}5$ for generating the approximation, then surrogate functions such as kriging functions may be out of the question unless some fixed values are selected for the parameters. Otherwise, the parameter values must be optimized which can be a time-consuming task. With fixed values, a statistical model used in kriging may not be the best one, i.e. it is not the optimal one to be used in approximating (interpolating) the data. However, a radial basis function can be fitted to the data by solving only a system of linear equations, but it may also have parameters to be optimized in certain basis functions.

In selecting a surrogate function to be used in practice, we refer to [10, 17, 55, 69, 77, 81, 85, 93, 99, 139, 144, 157, 171, 174, 182], where comparisons and suggestions are presented. It is obvious that different kinds of approximations can be generated by different kinds of surrogate functions. However, in this thesis, we do not suggest which surrogate function the analyst should select for the approximation. What we recommend is that the selection must be based on validations introduced later in this thesis. With the help of the validations, the analyst might find a proper surrogate function for approximating the Pareto front of the MOO problem.

In general, it must be clarified that surrogate functions are fitted to the data $\mathcal{D}$, which means that the outcomes of the surrogate functions adapt the data. In other words, surrogate functions model the data and not necessarily the natural phenomena or the physical object behind the data. Occasionally, an approximation created for Pareto front approximates in some sense the vectors in the front, but the vectors in the approximation may not be Pareto optimal objective vectors or even nondominated. It is possible that some of them are only objective vectors in the objective space and nothing else. This means that the approximation does not necessarily approximate the state of a nondominated vector, but the approximation as a set has in some sense the same "shape" of the set of Pareto optimal objective vectors. However, we believe that to find at least one "shape" or a sufficiently good approximation for the Pareto front of a computationally expensive nonconvex MOO problem is essential for the decision maker, who then can have estimations of feasible Pareto optimal objective vectors without spending too much time in calculating the true ones.

Next, we introduce some well-known surrogate functions the analyst can use in constructing an approximation in the IHBE method. We also briefly comment on the challenges the analyst may face in fitting the surrogate function.

1.4.2   On Radial Basis Functions

In this subsection, we present the basics of *Radial Basis Functions* (RBF) [28]. According to [181, page 387], radial basis functions are originally developed from the *method of potential functions*. To mention some examples of the wide range of the application area of RBFs, we might point out a design of a passenger train [18], intrusion detection [73], an explosively-formed oil well penetrator [119], characterization of porous materials [169], solving differential equations [149] and forecasting financial time series [180]. In optimization, RBFs have been applied even in optimization algorithms created for computationally expensive objective functions, see e.g. [58]. Different variations include *Adaptive* RBF [74, 75], *Extended* RBF [134] and *Constrained Optimization using Response Surfaces* RBF [150, 151, 152]. In addition, RBFs have been used, for example, in evolutionary single-objective optimization [27, 45, 100, 155, 164].

In the literature, an RBF surrogate function for data $\mathcal{D}$ consisting of m pairs $(\mathbf{x}^i, y^i)$ is typically introduced as follows.

$$f_{RBF}(\mathbf{x}) = \sum_{i=1}^{m} \lambda_i \phi(||\mathbf{x} - \mathbf{x}^i||), \tag{1}$$

where

- $||\cdot||$ is, for example, the *Euclidian* norm on the space $\mathbb{R}^n$ and vectors $\mathbf{x}^i \in \mathbb{R}^n$ are called *centers*,
- coefficients $\lambda_i \in \mathbb{R}$ are solutions of problem

$$\Phi\lambda = Y, \tag{2}$$

where

$$\lambda = \begin{bmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_k \end{bmatrix}^{\mathrm{T}}$$

$$\Phi = \begin{bmatrix} \phi(||\mathbf{x}^1 - \mathbf{x}^1||) & \phi(||\mathbf{x}^1 - \mathbf{x}^2||) & \cdots & \phi(||\mathbf{x}^1 - \mathbf{x}^m||) \\ \phi(||\mathbf{x}^2 - \mathbf{x}^1||) & \phi(||\mathbf{x}^2 - \mathbf{x}^2||) & \cdots & \phi(||\mathbf{x}^2 - \mathbf{x}^m||) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(||\mathbf{x}^m - \mathbf{x}^1||) & \phi(||\mathbf{x}^m - \mathbf{x}^2||) & \cdots & \phi(||\mathbf{x}^m - \mathbf{x}^m||) \end{bmatrix}$$

$$Y = \begin{bmatrix} y^1 & y^2 & \cdots & y^m \end{bmatrix}^{\mathrm{T}}.$$

- mapping ϕ is a given *basis function* from $\mathbb{R}$ into $\mathbb{R}$, see typical choices in Table 1.

In other words, an RBF surrogate function is a linear combination of a basis function, which only depends on the radius to the centers. Therefore, the name radial basis function. The basic construction is very simple and for each $(\mathbf{x}^i, y^i) \in \mathcal{D}$ we have

$$f_{RBF}(\mathbf{x}^i) = y^i.$$

However, depending on the basis function used and vectors $\mathbf{x}^i$, the matrix Φ in (2) can be singular. This means that the analyst cannot fit the surrogate function

TABLE 1 Typical choices in the literature for ϕ.

name	$\phi(r)$	
polyharmonic spline	$r^d,$	$d = 1, 3, 5, \ldots$
thin plate spline	$r^d \ln r,$	$d = 2, 4, 6, \ldots$
multiquadric	$\sqrt{r^2 + \varepsilon^2},$	$\varepsilon > 0$
Gaussian	$e^{-\varepsilon r^2},$	$\varepsilon > 0$

for the approximation. In order to avoid singular matrices, a polynomial can be included to the RBF form (1), see e.g. [58]:

$$f_{RBF}(\mathbf{x}) = \sum_{i=1}^{m} \lambda_i \phi(||\mathbf{x} - \mathbf{x}^i||) + p(\mathbf{x}). \tag{3}$$

In order to fit an RBF containing a polynomial to $\mathcal{D}$, let us assume for simplicity that the polynomial in (3) is a first degree polynomial, $p(\mathbf{x}) = a^T \mathbf{x} + b$. Then, coefficients λ_i, a_j and b ($i = 1, 2, \ldots m$, $j = 1, 2, \ldots, n$) can be solved from a system of equations

$$\begin{bmatrix} \Phi & P \\ P^T & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda \\ c \end{bmatrix} = \begin{bmatrix} Y \\ 0 \end{bmatrix}, \tag{4}$$

where Φ is a $m \times m$ matrix with $\Phi_{ij} = \phi(||\mathbf{x}^i - \mathbf{x}^j||)$ and

$$P = \begin{bmatrix} \mathbf{x}^{1T} & 1 \\ \mathbf{x}^{2T} & 1 \\ \vdots & \vdots \\ \mathbf{x}^{kT} & 1 \end{bmatrix}, \quad \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{bmatrix}, \quad c = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix}.$$

For certain basis functions, the degree of the polynomial must be selected in a proper way to get the matrix of the system of linear equations nonsingular in (3). Proper degrees are presented, for example, in [58].

1.4.3 On Kriging Functions

A kriging function is a stochastic function originally developed for efficient mining [40]. In the field of global optimization, kriging functions are applied, for example, in an algorithm called EGO (Efficient Global Optimization) [87] and in its variants [79, 97, 98, 108, 147, 148]. Through the optimization algorithms, kriging functions have been applied in various applications: circuit optimization [11], metal-forming [78], aerodynamic design [84], production of liquid ammonia [144], oil and gas production [148] and in finite element methods [161, 162]. A free implementation can be found e.g. in [141].

To get started with kriging functions, let us first assume that we have a *random mapping* (a random process or a stochastic process)

$$Y : \mathbb{R}^n \rightarrow \mathbb{R} \tag{5}$$

and let y^i be *realization*s of Y ($Y(\mathbf{x}^i) = y^i$ for all $i = 1, 2, \ldots, m$). Then a *kriging* function[1] to (5) based on m data pairs $(\mathbf{x}^i, y^i)$ is a function $f_K : \mathbb{R}^n \to \mathbb{R}$ so that

$$f_K(\mathbf{x}) = \sum_{i=1}^{k} K_{\mathbf{x}^i}^* Y(\mathbf{x}^i),$$

where coefficients $K_{\mathbf{x}^i}^* \in \mathbb{R}$ are solutions of a minimization problem of a variance subject to an expectation value condition:

$$
\begin{aligned}
&\min_{K_{\mathbf{x}^i}} \text{Var}(Y_K - Y(\mathbf{x})), \\
&\text{s.t. } E(Y_K) = E(Y(\mathbf{x})),
\end{aligned}
\tag{6}
$$

where

$$Y_K = \sum_{i=1}^{m} K_{\mathbf{x}^i} Y(\mathbf{x}^i).$$

In practice, it has been typically assumed that the stochastic process Y in (5) has the form

$$Y(\mathbf{x}) = g(\mathbf{x}) + Z(\mathbf{x}), \tag{7}$$

where

- g is a mapping from $\mathbb{R}^n$ into $\mathbb{R}$ consisting of l ($l < m$) functions $g_i : \mathbb{R}^n \to \mathbb{R}$ that were selected beforehand and which are called *global model*s (*universal model*s),

$$g(\mathbf{x}) = \sum_{i=1}^{l} \lambda_i g_i(\mathbf{x}),$$

  where $\lambda_i \in \mathbb{R}$ are unknown constants to be solved or optimized,
- Z is a random process from $\mathbb{R}^n$ into $\mathbb{R}$ with a mean zero and a covariance

$$\text{Cov}(Z(\mathbf{x}), Z(\mathbf{y})) = \sigma_Z^2 R(\mathbf{x}, \mathbf{y}),$$

  where σ_Z^2 is the variance and $R(\mathbf{x}, \mathbf{y})$ is the correlation of Z. Usually, correlation R has been selected to be

$$R(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{n} e^{-\theta_i |x_i - y_i|^2}$$

  and parameters θ_i are suggested to be optimized through the Maximum Likelihood Estimation (MLE) [41, 160].

Then, according to [160], the kriging function to (7) is

$$f_K(\mathbf{x}) = g_{\mathbf{x}}^{\mathrm{T}} \hat{\lambda} + r_{\mathbf{x}}^{\mathrm{T}} R_s^{-1} (Y_s - G_s \hat{\lambda}), \tag{8}$$

---

[1]   In the literature, a kriging function is also known as the *Best Linear Unbiased Predictor* (BLUP) or the *Best Linear Unbiased Estimator* (BLUE).

where

$$\hat{\lambda} = (G_s^\mathrm{T} R_s^{-1} G_s)^{-1} G_s^\mathrm{T} R_s^{-1} Y_s,$$

$$Y_s = \begin{bmatrix} Y_D(\mathbf{x}^1) & Y_D(\mathbf{x}^2) & \cdots & Y_D(\mathbf{x}^m) \end{bmatrix}^\mathrm{T},$$

$$g_\mathbf{x} = \begin{bmatrix} g_1(\mathbf{x}) & g_2(\mathbf{x}) & \cdots & g_l(\mathbf{x}) \end{bmatrix}^\mathrm{T},$$

$$G_s = \begin{bmatrix} g_1(\mathbf{x}^1) & g_2(\mathbf{x}^1) & \cdots & g_l(\mathbf{x}^1) \\ g_1(\mathbf{x}^2) & g_2(\mathbf{x}^2) & \cdots & g_l(\mathbf{x}^2) \\ \vdots & \vdots & \ddots & \vdots \\ g_1(\mathbf{x}^m) & g_2(\mathbf{x}^m) & \cdots & g_l(\mathbf{x}^m) \end{bmatrix},$$

$$r_\mathbf{x} = \begin{bmatrix} R(\mathbf{x}^1,\mathbf{x}) & R(\mathbf{x}^2,\mathbf{x}) & \cdots & R(\mathbf{x}^m,\mathbf{x}) \end{bmatrix}^\mathrm{T},$$

$$R_s = \begin{bmatrix} R(\mathbf{x}^1,\mathbf{x}^1) & R(\mathbf{x}^1,\mathbf{x}^2) & \cdots & R(\mathbf{x}^1,\mathbf{x}^m) \\ R(\mathbf{x}^2,\mathbf{x}^1) & R(\mathbf{x}^2,\mathbf{x}^2) & \cdots & R(\mathbf{x}^2,\mathbf{x}^m) \\ \vdots & \vdots & \ddots & \vdots \\ R(\mathbf{x}^m,\mathbf{x}^1) & R(\mathbf{x}^m,\mathbf{x}^2) & \cdots & R(\mathbf{x}^m,\mathbf{x}^m) \end{bmatrix}.$$

It is easy to show that $f_K(\mathbf{x}^i) = y^i$ for all $i = 1,\ldots,m$, see e.g. [86]. In addition, an error estimation for a value $f_K(\mathbf{x})$ can be calculated as follows

$$\mathrm{MSE}(f_K(\mathbf{x})) = \hat{\sigma}_Z^2 \left(1 - \begin{bmatrix} g_\mathbf{x}^\mathrm{T} & r_\mathbf{x}^\mathrm{T} \end{bmatrix} \begin{bmatrix} 0 & G_s^\mathrm{T} \\ G_s & R_s \end{bmatrix}^{-1} \begin{bmatrix} g_\mathbf{x} \\ r_\mathbf{x} \end{bmatrix} \right),$$

where

$$\hat{\sigma}_Z^2 = \frac{1}{k}(Y_s - G_s\hat{\lambda})^\mathrm{T} R_s^{-1}(Y_s - G_s\hat{\lambda})$$

is an estimation through the MLE, see e.g. [160]. Note that the minimization problem (6) means, in fact, solving a *Mean Square Error* problem [160].

In practice, the analyst may have to use some time in solving the optimal θ_i if he or she has decided to use a kriging function in the approximation. Otherwise, the fitting process of the kriging function is computationally efficient.

1.4.4 Briefly on Regression Functions

As said, the analyst may not have $f_{surrogate}(\mathbf{x}^i) = y^i$ for all $i = 1,\ldots,m$, if he or she has used a regression function for the surrogate function. However, by properly choosing parameters in the regression function, it can also interpolate the data in $\mathcal{D}$. Therefore, we present here some of the commonly known regression functions, which have this property.

Regression functions have a long history in constructing a function to represent data. Already in the fifties, polynomial-based regression functions were used in modeling real life problems [19] and, nowadays for example, in producing liquid ammonia [144] or in aero dynamical problems [76, 171].

Typically, regression functions are considered under the *Response Surface Methods* (RSM), and the most widely used response function has the form of a

linear regression,

$$f_{lr}(\mathbf{x}) = \sum_{i=1}^{l} \lambda_i g_i(\mathbf{x}) + \varepsilon(\mathbf{x}), \tag{9}$$

where $\varepsilon(\cdot)$ is an error estimation, coefficients $\lambda_i \in \mathbb{R}$ are unknown and to be selected properly, and functions $g_i : \mathbb{R}^n \to \mathbb{R}$ are given.

To fit a regression function to data $\mathcal{D}$, a system of linear equations derived from (9) can be written in a matrix notation as follows:

$$\mathbf{y} := \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^k \end{bmatrix} = \begin{bmatrix} g_1(\mathbf{x}^1) & g_2(\mathbf{x}^1) & \cdots & g_l(\mathbf{x}^1) \\ g_1(\mathbf{x}^2) & g_2(\mathbf{x}^2) & \cdots & g_l(\mathbf{x}^2) \\ \vdots & \vdots & \ddots & \vdots \\ g_1(\mathbf{x}^k) & g_2(\mathbf{x}^k) & \cdots & g_l(\mathbf{x}^k) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_l \end{bmatrix} =: \mathbf{g}\lambda.$$

Then the unknown coefficients λ_i can be calculated by using the formula

$$\lambda = (\mathbf{g}^\mathsf{T}\mathbf{g})^{-1}\mathbf{g}^\mathsf{T}\mathbf{y},$$

which is commonly known as a solution of a *linear least square problem*, see e.g. [142]. In order to interpolate the data, the analyst can select l in (9) to be equal to m, the number of pairs $(\mathbf{x}^i, y^i)$ in $\mathcal{D}$. Normally, $l < m$.

Recently, *Support Vector Regression* (SVR) (a variant of a *Support Vector Machine* (SVM) [37, 168, 173, 181]) has become more popular as a regression function. It has been used, for example, in tourism demand forecasting [32], face detection [109], image processing [140], software efforts [143], flood forecasting [195], gene expression [186], traffic data analysis [192], protein predictions [199] and in solving optimization problems [137, 138, 196, 197]. A free implementation can be found in [29]. Basically, a support vector regression takes the form of

$$f_{SVR}(\mathbf{x}) = \sum_{i=1}^{l} \lambda_i K(\mathbf{x}, \mathbf{x}^i) + b,$$

where

- the function K from $\mathbb{R}^n \times \mathbb{R}^n$ into $\mathbb{R}$ is called a *kernel*, and is, for example, some of the kernels presented in Table 2.
- coefficients λ_i and b are solutions of an optimization problem. See, for example, [168, 181] for several optimization problem formulations and computationally efficient solving techniques for them.
- number l is the number of so-called *support vectors*.

Depending on the kernel function and the method used for calculating coefficients λ_i and b, the analyst can have an interpolative support vector regression, if the number of support vectors is increased or some other parameters in calculating coefficients λ_i and b are selected properly.

TABLE 2   Typical choices in the literature for the kernel.

name	$K(\mathbf{x}, \mathbf{x}^i)$
polynomial	$(\langle \mathbf{x}, \mathbf{x}^i \rangle + 1)^d,$ $\quad d = 1, 2, 3, \ldots$
radial basis	$\exp(-\frac{1}{2\sigma^2} \|\mathbf{x} - \mathbf{x}^i\|^2),$ $\quad \sigma^2 > 0$
multilayer perceptron	$\tanh(\beta_0 \langle \mathbf{x}, \mathbf{x}^i \rangle + \beta_1)$
a sum of kernels	$\sum_{j=1}^{m} K_j(\mathbf{x}, \mathbf{x}^i)$
a function that holds *Mercer*'s condition [120]	
Coefficients d, σ^2, β_0 and β_1 are typically given.	

1.5   Outline of This Thesis

The main contribution of this thesis is the IHBE method (Interactive HyperBox Exploring). The method is designed for supporting the decision maker in learning about the MOO problem and his or her preferences in relation what is available. Then, the decision maker may found a compromise solution to a computationally expensive nonconvex MOO problem and be confident about it. During the construction of the method, three major aspects are considered: theoretical aspects, computational aspects and the decision maker's aspects. However, the contribution of this thesis is divided into four different steps as follows.

In order to find a compromise solution for the computationally expensive MOO problem, the first step is to create an approximation for the Pareto front. The approximation is introduced in Chapter 2. The approximation must be theoretically valid for computationally expensive nonconvex MOO problems, but from the practical point of view, the actual use of the approximation must be computationally efficient because it is designed for practical problems. Contribution: a fast method for evaluations of approximated Pareto optimal objective vectors.

Since the approximation models the Pareto front, then the second part of this thesis is the IHBE method based on the approximation. The IHBE method is introduced in Chapter 3, and it consists of well-known interactive methods. In this way, the IHBE method is supposed to be easy to use by the decision maker, and informative in helping and supporting him or her through the solution process. In other words, the decision maker's perspective is considered in the construction of the IHBE method without forgetting the computational aspects in calculating approximated Pareto optimal objective vectors that obey, in some sense, the decision maker's preferences. Contribution: an easy method for the decision maker to investigate approximated Pareto optimal objective vectors.

The third part is presented in Chapters 4 and 5, where methods created for validating the approximation are discussed. The methods are modifications of other methods or theories used in MCDM. The new contributions of the methods are based on their computationally efficient characterizations of whether or not a vector in the approximation is dominated. The computational methods are based on mathematically proved theorems. Contribution: fast methods for validating

the approximation.

The fourth and the last part of this thesis is an interpretation of dominated vectors in the approximation. The interpretation is illustrated in Chapter 6. From the decision maker's point of view, the nonconvexity of the MOO problem may cause some challenging situations, which require special tools to be used in the solution process. Contribution: methods for analyzing the approximation.

Finally, all the methods used in the IHBE method are summarized in Chapter 7, and concluding remarks are presented in Chapter 8.

2 APPROXIMATING THE PARETO FRONT BY A SURROGATE FUNCTION

In this chapter, we propose a mathematical model for approximating the Pareto front of the MOO problem. While constructing the approximation, we focus on goals $\mathcal{G}1$-$\mathcal{G}5$ and $\mathcal{G}7$ presented on page 20. In other words, everything related to the approximation must be computationally efficient. In addition, only a finite set of Pareto optimal objective vectors is assumed to be available in the construction. Note that the dimension of the vectors can be more than two. And finally, the MOO problem can be computationally expensive and even nonconvex.

This chapter has been organized as follows. First, in Subsection 2.1, a basis for the mathematical model is set and then, in Subsection 2.2, the actual approximation is introduced. It uses a surrogate function for approximating the Pareto front. After this, a concluding discussion and remarks for the approximation are presented in Subsection 2.3.

2.1 Setting up a Basis for the Approximation

In this section, we construct a basis for the approximation. It is based on $m \geq 2$ distinct pre-calculated Pareto optimal objective vectors. The vectors are denoted by $\mathbf{z}^1, \mathbf{z}^2, \ldots, \mathbf{z}^m$, and the set consisting of them is denoted by Z_m. To put it briefly, the basis contains three different elements: a set Δ_{k-1}, vectors $\mathbf{z}^{s_i}$ and scalars λ^i. The set Δ_{k-1} is a domain space for a surrogate function, which is fitted to data consisting of vectors $\mathbf{z}^{s_i}$ and scalars λ^i. In this section, we introduce only the basis, but in the next section, we show how the basis is used in approximating the Pareto front.

To get started with the basis, we first define the set Δ_{k-1}.

Definition 3. A set

$$\Delta_{k-1} := \{\mathbf{z} \in \mathbb{R}^k : \exists \lambda_i^\Delta \geq 0 \text{ so that } \mathbf{z} = \sum_{i=1}^k \lambda_i^\Delta \mathbf{e}^i \text{ and } \sum_{i=1}^k \lambda_i^\Delta = 1\},$$

where $\mathbf{e}^i$ is the i:th standard basis vector of the space $\mathbb{R}^k$, is called a $(k-1)$-*dimensional simplex*. If the dimension $k-1$ is obvious from the context, then it is called a *simplex*.

Geometrically, a $(k-1)$-dimensional simplex is a convex hull of the standard basis vectors of $\mathbb{R}^k$. In space $\mathbb{R}^2$, the simplex Δ_1 is a closed interval from $(1,0)^{\mathrm{T}}$ to $(0,1)^{\mathrm{T}}$ as illustrated in Figure 5 (black line).

For later use and calculations, we define two useful vectors in the objective space denoting the extreme vectors of Z_m

$$\mathbf{ideal}^{Z_m} = (\min_{i=1,\dots,m} z^i_1, \dots, \min_{i=1,\dots,m} z^i_k)^{\mathrm{T}} \in \mathbb{R}^k \tag{10}$$

and

$$\mathbf{nadir}^{Z_m} = (\max_{i=1,\dots,m} z^i_1, \dots, \max_{i=1,\dots,m} z^i_k)^{\mathrm{T}} \in \mathbb{R}^k. \tag{11}$$

An example of a two-dimensional case of $\mathbf{ideal}^{Z_m}$ and the corresponding set Z_m is illustrated in Figure 5 (black dots).
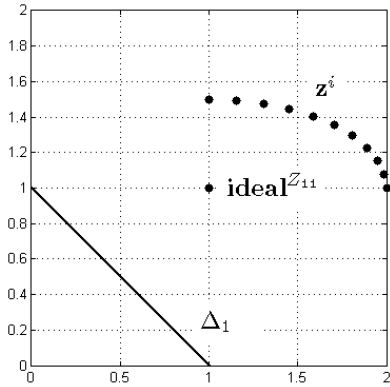
FIGURE 5 Set Δ_1 and vector $\mathbf{ideal}^{Z_{11}}$ based on Z_{11}.
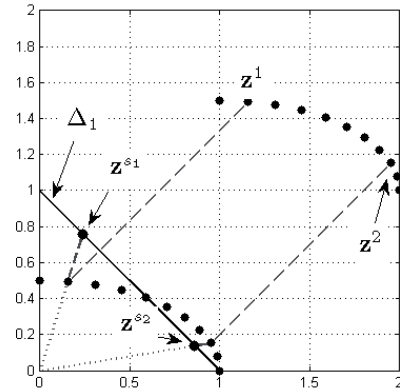
FIGURE 6 Mapping vectors of Z_{11} into a simplex.

Remark 4. Note that $\mathbf{ideal}^{Z_m}$ dominates every vector in Z_m and if one of the coordinates of $\mathbf{ideal}^{Z_m}$ is increased, then there is a vector in Z_m so that it is not anymore dominated by $\mathbf{ideal}^{Z_m}$. In other words, for all $\mathbf{z}^i \in Z_m$ and for all $j = 1, \dots, k$ we have

$$z^i_j - \mathrm{ideal}^{Z_m}_j \geq 0$$

and for some index j the inequality is strict, because Z_m contains at least two different Pareto optimal objective vectors. Thus, for later use, we have an essential property:

$$\sum_{j=1}^{k} z^i_j - \mathrm{ideal}^{Z_m}_j > 0. \tag{12}$$

Let a function $f_{Z_m}^p : (\mathbf{ideal}^{Z_m} + \mathbb{R}_+^k) \setminus \mathbf{ideal}^{Z_m} \to \Delta_{k-1}$ be defined as follows:

$$f_{Z_m}^p(\mathbf{z}) = \frac{\mathbf{z} - \mathbf{ideal}^{Z_m}}{\sum_{j=1}^k z_j - \mathbf{ideal}_j^{Z_m}}. \tag{13}$$

Next we prove that the function $f_{Z_m}^p$ is well-defined.

Lemma 5. For each vector $\mathbf{z} \in (\mathbf{ideal}^{Z_m} + \mathbb{R}_+^k) \setminus \mathbf{ideal}^{Z_m}$, we have $f_{Z_m}^p(\mathbf{z}) \in \Delta_{k-1}$.

Proof. Let $\mathbf{z} \in (\mathbf{ideal}^{Z_m} + \mathbb{R}_+^k) \setminus \mathbf{ideal}^{Z_m}$. To show that $f_{Z_m}^p(\mathbf{z}) \in \Delta_{k-1}$, we have to find coefficients $\lambda_j^{\mathbf{z}} \geq 0$ so that

$$f_{Z_m}^p(\mathbf{z}) = \sum_{j=1}^k \lambda_j^{\mathbf{z}} \mathbf{e}^j \text{ and } \sum_{j=1}^k \lambda_j^{\mathbf{z}} = 1.$$

By choosing

$$\lambda_j^{\mathbf{z}} = \frac{z_j - \mathbf{ideal}_j^{Z_m}}{\sum_{l=1}^k z_l - \mathbf{ideal}_l^{Z_m}} \geq 0, \quad j = 1, \ldots, k,$$

where the last inequality follows from Remark 4, we have $\sum_{j=1}^k \lambda_j^{\mathbf{z}} = 1$ and

$$\sum_{j=1}^k \lambda_j^{\mathbf{z}} \mathbf{e}^j = \sum_{j=1}^k \frac{z_j - \mathbf{ideal}_j^{Z_m}}{\sum_{l=1}^k z_l - \mathbf{ideal}_l^{Z_m}} \mathbf{e}^j = \frac{\mathbf{z} - \mathbf{ideal}^{Z_m}}{\sum_{l=1}^k z_l - \mathbf{ideal}_l^{Z_m}} = f_{Z_m}^p(\mathbf{z}).$$

$\square$

Geometrically, the function $f_{Z_m}^p$ first shifts all vectors in the set $(\mathbf{ideal}^{Z_m} + \mathbb{R}_+^k) \setminus \mathbf{ideal}^{Z_m}$ by $-\mathbf{ideal}^{Z_m}$. Then the vector-wise ideal, as in (10), for all the vectors $\mathbf{z}^i \in Z_m$ shifted is the origin. After this, all the vectors shifted are mapped into the simplex by multiplying by scalars $1/(\sum_{j=1}^k z_j - \mathbf{ideal}_j^{Z_m})$, respectively. The shift and the multiplication are illustrated in Figure 6 as gray dashed lines for the vectors $\mathbf{z}^1, \mathbf{z}^2 \in Z_m$.

For the basis to the approximation, we calculate representative vectors $\mathbf{z}^{s_i} \in \Delta_{k-1}$ for each $\mathbf{z}^i \in Z_m$ by using the function $f_{Z_m}^p$. In other words,

$$\mathbf{z}^{s_i} := f_{Z_m}^p(\mathbf{z}^i), \quad i = 1, \ldots, m. \tag{14}$$

We denote the set of the vectors $\mathbf{z}^{s_i}$ by Z_m^s.

Next, we show how to calculate $\mathbf{z}^i$ from $\mathbf{z}^{s_i}$. Then the calculation is used for the scalar values λ^i for the basis of the approximation.

Lemma 6. There exist unique scalars $\lambda^i > 0$ for each $\mathbf{z}^i \in Z_m$, $i = 1, \ldots, m$, so that

$$\mathbf{z}^i = \mathbf{ideal}^{Z_m} + \lambda^i \mathbf{z}^{s_i}, \quad i = 1, \ldots, m.$$

Proof. We divide the proof into two steps. First, we show the existence and then the uniqueness.

1. By choosing

$$\lambda^i = \sum_{j=1}^{k} z_j^i - \text{ideal}_j^{Z_m} > 0,$$

we have

$$\mathbf{ideal}^{Z_m} + \lambda^i \mathbf{z}^{s_i} = \mathbf{ideal}^{Z_m} + \left(\sum_{j=1}^{k} z_j^i - \text{ideal}_j^{Z_m} \right) \frac{\mathbf{z}^i - \mathbf{ideal}^{Z_m}}{\sum_{j=1}^{k} z_j^i - \text{ideal}_j^{Z_m}}$$
$$= \mathbf{ideal}^{Z_m} + \mathbf{z}^i - \mathbf{ideal}^{Z_m} = \mathbf{z}^i.$$

2. Let us assume that there exist two different scalars λ^{i_1} and λ^{i_2} so that

$$\mathbf{ideal}^{Z_m} + \lambda^{i_1} \mathbf{z}^{s_i} = \mathbf{z}^i = \mathbf{ideal}^{Z_m} + \lambda^{i_2} \mathbf{z}^{s_i}.$$

Then,

$$\lambda^{i_1} \mathbf{z}^{s_i} = \lambda^{i_2} \mathbf{z}^{s_i} \Leftrightarrow (\lambda^{i_1} - \lambda^{i_2}) \mathbf{z}^{s_i} = \mathbf{0}.$$

Since $\mathbf{z}^{s_i} \in \Delta_{k-1}$ and $\mathbf{0} \notin \Delta_{k-1}$, then $\lambda^{i_1} = \lambda^{i_2}$. In other words, the scalar λ^i is unique.

□

Lemmas 5 and 6 show that connections between $\mathbf{z}^i$ and $\mathbf{z}^{s_i}$ exist, i.e., we can calculate one from the other. Note it is possible that for two arbitrary vectors $\hat{\mathbf{z}}, \bar{\mathbf{z}} \in \mathbb{R}^k$ of the form $\hat{z}_i, \bar{z}_i \geq \text{ideal}_i^{Z_m}$, $i = 1, \ldots, k$, we might have $\hat{\mathbf{z}}^s := f_{Z_m}^p(\hat{\mathbf{z}}) = f_{Z_m}^p(\bar{\mathbf{z}}) =: \bar{\mathbf{z}}^s$. This is the case in Figure 7, where vectors $\hat{\mathbf{z}}^s, \bar{\mathbf{z}}^s, \hat{\mathbf{z}}, \bar{\mathbf{z}}$ are depicted by black dots.
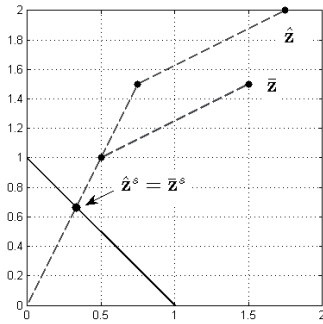
FIGURE 7   Two different vectors $\hat{\mathbf{z}}, \bar{\mathbf{z}}$ so that they are mapped by $f_{Z_m}^p$ to the same vector in the simplex Δ_1.

Remark 7. The reason why cases as in Figure 7 exist is obvious. The function $f_{Z_m}^p$ is not injective - it is surjective, because $f_{Z_m}^p(\mathbf{ideal}^{Z_m} + \Delta_{k-1}) = \Delta_{k-1}$. However, it is not possible that there exists a vector in Δ_{k-1} so that it is a representative one for two different vectors $\mathbf{z}^i, \mathbf{z}^j \in Z_m$. This is guaranteed by the assumption of Pareto optimality for the vectors in Z_m, as we will see next.

Theorem 8. Let $\mathbf{z}^i, \mathbf{z}^j \in Z_m$. Then $\mathbf{z}^i \neq \mathbf{z}^j$ if and only if $\mathbf{z}^{s_i} \neq \mathbf{z}^{s_j}$.

Proof.

"$\Rightarrow$" Let us assume that there exist two different Pareto optimal objective vectors $\mathbf{z}^i, \mathbf{z}^j \in Z_m$ so that $\mathbf{z}^{s_i} = \mathbf{z}^{s_j}$. According to Lemma 6, we have unique scalars $\lambda^i, \lambda^j \geq 0$ so that $\mathbf{z}^i = \mathbf{ideal}^{Z_m} + \lambda^i \mathbf{z}^{s_i}$ and $\mathbf{z}^j = \mathbf{ideal}^{Z_m} + \lambda^j \mathbf{z}^{s_j}$. Since $\mathbf{z}^i \neq \mathbf{z}^j$, then $\lambda^i \neq \lambda^j$. We can assume that $\lambda^i < \lambda^j$ without loss of generality. Now for all $l = 1, \ldots, k$, we have

$$z_l^i = \mathrm{ideal}_l^{Z_m} + \lambda^i z_l^{s_i} < \mathrm{ideal}_l^{Z_m} + \lambda^j z_l^{s_j} = z_l^j,$$

which means that $\mathbf{z}^i$ dominates $\mathbf{z}^j$, and that is a contradiction.

"$\Leftarrow$" This follows directly from the definition of the function $f_{Z_m}^p$ via a contradiction.

$\square$

Now that the elements for a basis for the approximation have been presented, we are ready to use them (the simplex Δ_{k-1}, vectors $\mathbf{z}^{s_i}$ and scalars λ^i calculated in Lemma 6) in constructing an approximation for the Pareto front.

2.2 Pareto Front Approximation with Surrogate Models

In Section 2.1, a basis for the Pareto front approximation was created: the set Δ_{k-1}, vectors $\mathbf{z}^{s_i}$ and scalars λ^i were introduced. In this section, we illustrate how they are used in approximating the Pareto front.

We have assumed that pre-calculated Pareto optimal objective vectors to the MOO problem are available in the set Z_m. Then by using the function $f_{Z_m}^p$, defined in (13), we map every vector in Z_m into Δ_{k-1}. As a result, we obtain a unique set of vectors in Δ_{k-1} (Lemma 5 and Theorem 8). Then in Lemma 6, we proved that there exist unique scalars λ^i for each $\mathbf{z}^i \in Z_m$ so that $\mathbf{z}^i = \mathbf{ideal}^{Z_m} + \lambda^i \mathbf{z}^{s_i}$. Next, we generalize this to be a function for approximating the Pareto front.

Definition 9. We *model* the Pareto front by a function $f_a : \Delta_{k-1} \to \mathbb{R}^k$,

$$f_a(\mathbf{z}) = \mathbf{ideal}^{Z_m} + f_s(\mathbf{z})\mathbf{z}, \tag{15}$$

where f_s is a function from Δ_{k-1} into $\mathbb{R}$. The image $f_a(\Delta_{k-1})$ is referred to as an *approximated Pareto front*.

In Definition 9, the function f_s can be any surrogate function presented in Section 1.4, but it must satisfy the following condition:

Definition 10. Function f_s in (15) must be selected so that for the data consisting of pairs $(\mathbf{z}^{s_i}, \lambda^i)$ we have

$$f_s(\mathbf{z}^{s_i}) = \lambda^i, \quad i = 1, \ldots, m, \tag{16}$$

where $\lambda^i = \sum_{j=1}^{k} z_j^i - \text{ideal}_j^{Z_m}$ is a scalar as calculated in Lemma 6. In addition, the function f_s must be positive on Δ_{k-1} in order to avoid having vectors in the approximation that dominate the vector $\textbf{ideal}^{Z_m}$ and the vectors in Z_m.

Now, it can be seen how the basis created in Subsection 2.1 is used in approximating the Pareto front. The set Δ_{k-1} is the domain space for functions f_a and f_s. Then, the vectors $\textbf{z}^{s_i}$ and the scalars λ^i are used in fitting f_s to the data $\mathcal{D}$ consisting of m pairs $(\textbf{z}^{s_i}, \lambda^i)$. Note that condition (16) is important because it forces the function f_s to interpolate the data. Then the function f_a interpolates data consisting of m pairs $(\textbf{z}^{s_i}, \textbf{z}^i)$, that is, $f_a(\textbf{z}^{s_i}) = \textbf{z}^i$ for all $i = 1, \ldots, m$. This means that the approximation contains at least the vectors in the set Z_m, which contains Pareto optimal objective vectors.

A geometrical illustration of the approximated Pareto front is presented in Figure 8, where an RBF model with a polyharmonic spline of a third degree has been used for f_s. On the right hand side, the approximated Pareto front based on vectors in Z_{21} (black dots on the left hand side of Figure 8) has been depicted as a gray dashed curve. As discussed in Subsection 1.4.1, RBFs interpolate the data, and thus, the analyst can be sure about the condition (16) that it holds when an RBF is used for f_s. Basically, the function f_s approximates scalar values for vectors on the simplex so that the vectors multiplied by the scalars end up to approximate the Pareto front when they are shifted by $\textbf{ideal}^{Z_m}$. In other words, the function f_s captures the shape of the Pareto front in terms of vectors $\textbf{z}^{s_i}$ and scalars λ^i. However, to put the shape into a right position in the objective space, a shift by $\textbf{ideal}^{Z_m}$ is needed.
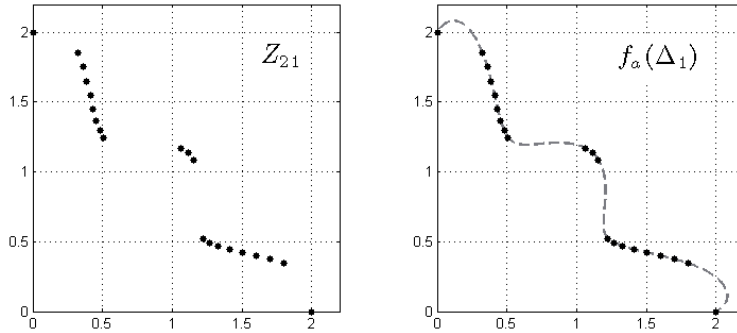
FIGURE 8   Pareto front approximation $f_a(\Delta_1)$.

The image (set) $f_a(\Delta_{k-1})$ is a subset of the objective space. Likewise is the Pareto front, but vectors in the Pareto front are nondominated. The formulation of the function f_a is simple, but some of the vectors in $f_a(\Delta_{k-1})$ can be dominated (the interpretation of dominated vectors in the approximation is presented in Chapter 6). However, the nondominated vectors in $f_a(\Delta_{k-1})$ are the main core

of the approximation we are constructing. Let us first present a definition and then illustrate this geometrically.

Definition 11. Let $\mathbf{z} \in f_a(\Delta_{k-1})$. If the vector $\mathbf{z}$ belongs to $f_a(\Delta_{k-1})^*$ (i.e. it is nondominated in $f_a(\Delta_{k-1})$), then the vector $\mathbf{z}$ is said to be *an approximated Pareto optimal objective vector*.

Definition 11 means in practice that the set $f_a(\Delta_{k-1})$ is approximating the set of Pareto optimal objective vectors of the MOO problem, but only vectors in $f_a(\Delta_{k-1})^*$ are interpreted as approximated Pareto optimal objective vectors. By Definition 11, we particularly consider the approximation to be suitable for nonconvex MOO problems. We demonstrate this as follows. Let us have a non-convex MOO problem, which has been solved, i.e. the entire Pareto front has been calculated. Let the Pareto front be as on the left hand side in Figure 9 (black curves). On the right hand side, we have depicted the set $f_a(\Delta_{k-1})$ as a gray dashed curve, and on top of it bold gray curves illustrate the set $f_a(\Delta_{k-1})^*$. We have also included the Pareto front on top of $f_a(\Delta_{k-1})^*$ as black curves. As it can be seen, nondominated vectors in $f_a(\Delta_{k-1})^*$ approximate (represent) the Pareto front. The dominated vectors in the approximation can be interpreted in a special way, but we return to them in Chapter 6.
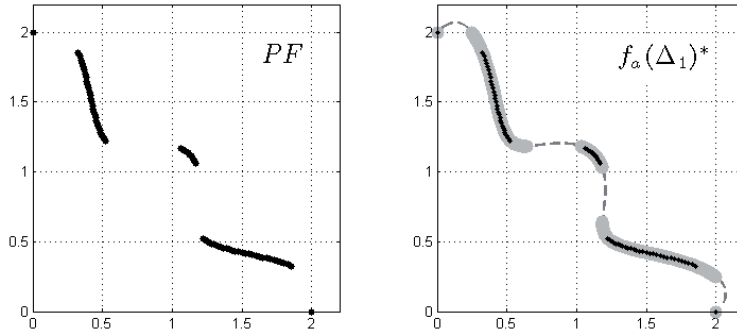
FIGURE 9   A pareto front approximation $f_a(\Delta_1)$ and the nondominated subset $f_a(\Delta_1)^*$.

As seen in the above example, the approximation may contain dominated vectors. Now, a justified question is the following. Can dominated vectors in the approximation be avoided? The answer is "yes" and "no", depending on the MOO problem. Usually, in nonconvex MOO problems, the approximation contains dominated vectors because of the mathematical structure of the function f_a. We illustrate this as follows. In theory, a Pareto front can be, for example, a set of consisting of vectors

$$\{\lambda(0,2)^{\mathrm{T}} + (1-\lambda)(0.5,1.5)^{\mathrm{T}} : \lambda \in (0,1]\} \cup \{\lambda(0.5,0.5)^{\mathrm{T}} + (1-\lambda)(1,0)^{\mathrm{T}} : \lambda \in [0,1]\}.$$

Now, for this set it is not possible to construct a function $f_a : \Delta_1 \to \mathbb{R}^2$ so that $f_a(\Delta_1) = f_a(\Delta_1)^*$ and $\mathrm{PF}(f, S) \subset f_a(\Delta_1)$. This can be seen in the following. In Figure 10, we have illustrated the set $\mathrm{PF}(f, S)$. Let vector $\mathbf{z}$ be $(0.4, 0.6)^\mathrm{T} \in \Delta_1$ (a black dot in Figure 10).
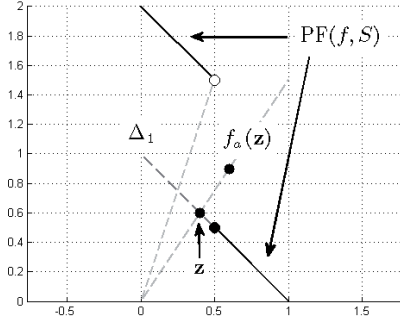
FIGURE 10   A Pareto front that cannot be modeled by the approximation so that the approximation does not contain dominated vectors.

Now, if the coordinate value $f_a(\mathbf{z})_1$ is

1. less than 0.5, then the vector $f_a(\mathbf{z})$ dominates the vector $((f_a(\mathbf{z})_1 + 0.5)/2, 2 - (f_a(\mathbf{z})_1 + 0.5)/2)^\mathrm{T} \in \mathrm{PF}(f, S)$, or
2. greater or equal to 0.5, then the vector $(0.5, 0.5)^\mathrm{T} \in \mathrm{PF}(f, S)$ dominates the vector $f_a(\mathbf{z})$.

In other words, the nonconvexity of the MOO problem has forced the approximation to contain dominated vectors.

Definition 11 for approximated Pareto optimal objective vectors may seem to be a nice theoretical definition, and it might be useless in practice because it should be possible to identify whether or not a vector $f_a(\mathbf{z})$ is nondominated in the set $f_a(\Delta_{k-1})$. However, the situation is not hopeless, because the mathematical structure of the function f_a sets a good basis for such an analysis. It turns out that computationally efficient methods can be created to identify nondominated and dominated vectors in the approximation. The methods are discussed in Chapters 4 and 5.

The main idea of the function f_a is to provide a simple access to calculate vectors on the approximation. This can be done as follows (we present the steps next and after that we discuss the details of the steps):

Method 12.

(1) Select vector(s) in the simplex Δ_{k-1}.
(2) Use function f_a to map the selected vector(s) into the approximation.

Method 12 is a computationally efficient method to be used in calculating vectors on the approximation. Note that the function f_s is a surrogate function, which by

a definition is a computationally efficient function, and the other operations used in the function f_a are not computationally expensive to be evaluated. However, in the solution process of a MOO problem, selecting the vectors on the simplex may be unintuitive for the decision maker. Therefore, in Section 3.2, a method is proposed for the selection. The method generates vectors on the simplex so that they correspond to the decision maker's preferences. In other words, the decision maker does not have to select them in the simplex by himself or herself.

As said, the simple structure of the function f_a can be used in identifying dominated vectors in the approximation. We believe that the analyst might be interested in knowing the error estimations (confidence levels) for vectors in the approximation. However, an error estimation for vectors in the approximation is not necessarily that straightforward in practice because we are approximating a set by a set. On the other hand, if error estimations for the function f_s are available and calculated, then they can be included into f_a as follows:

$$\tilde{f}_a(\mathbf{z}) = \mathbf{ideal}^{Z_m} + (f_s(\mathbf{z}) \pm \varepsilon_{\mathbf{z}})\mathbf{z}, \tag{17}$$

where $\varepsilon_{\mathbf{z}}$ is the error estimation for f_s at $\mathbf{z}$. However, the above error estimation does not calculate, for example, any probability for an objective vector to be non-dominated in the set $f_a(\Delta_{k-1})$. It only reflects the error estimations calculated for the function f_s into the approximation. Other challenges in error estimations are discussed in Subsection 3.1.2.1.

To conclude this section, a mathematical model (function f_a) has been proposed for approximating the set of Pareto optimal objective vectors of the MOO problem. After having constructed the function f_a, the set $f_a(\Delta_{k-1})^*$ contains vectors, which are interpreted as approximated Pareto optimal objective vectors to the MOO problem.

2.3 Discussion and Conclusions on the New Approximation

In this section, we summarize the main features of the approximation. In order to make cross-references for later use, we use notation $\mathcal{AM}\mathrm{f}X$ to indicate feature X of the $\mathcal{A}$pproximation $\mathcal{M}$odel consisting of the functions f_a, f_s and the set Δ_{k-1}.

$\mathcal{AM}\mathrm{f}1$ Let us start with the function f_a defined in (15). It contains a shift by $\mathbf{ideal}^{Z_m}$ and multiplication of a scalar value $f_s(\mathbf{z})$ and vector $\mathbf{z}$, where $\mathbf{z} \in \Delta_{k-1}$. These two operations are not computationally expensive. This means that we have reached the seventh goal $\mathcal{G}7$, set on page 20 for the approximation. In other words, generating vectors on the approximated Pareto front can be done computationally efficiently with the approximation. However, for some vectors in Δ_{k-1}, the corresponding vectors in the approximation can be dominated.

$\mathcal{AM}\mathrm{f}2$ The second feature is the condition (16) for the function f_s in Definition 10. Condition $f_s(\mathbf{z}^{s_i}) = \lambda^i$ for all $i = 1, \ldots, m$, and Theorem 8 guarantee

that $f_a(\mathbf{z}^{s_i}) = \mathbf{z}^i$ for all $i = 1, \ldots, m$. Therefore, the set Z_m is a subset of $f_a(\Delta_{k-1})$, meaning that the approximation contains true Pareto optimal objective vectors.

Calculation of data pairs $(\mathbf{z}^{s_i}, \lambda^i)$ does not include computationally expensive operations. However, if the function f_s requires some parameters to be optimized like in kriging models [160], then some computation time is required. On the other hand, if a radial basis function with a polyharmonic spline of a third degree [28] has been selected, then only a system of linear equations must be solved. Thus, depending on the surrogate function selected and its fitting to the data, it can be said that goal $\mathcal{G}5$ has been reached partially.

As mentioned in Section 1.4, we have several different types of surrogate functions which interpolate the data. Therefore, we do not have to construct such a function by ourselves. However, the great variety of surrogate functions may cause a so-called *selection problem* for the analyst: which surrogate function should be selected as the function f_s? This is not a trivial problem to be solved, but the analyst might find a proper one after validating different surrogate functions by methods introduced in later chapters. In other words, we propose to use special methods created for validations in selecting a proper surrogate function.

$\mathcal{AM}$f3   The third feature is more theoretical. It is easy to show (here we leave it for the reader because it follows immediately from Definition 9, Lemmas 5, 6 and Theorem 8) that every Pareto optimal objective vector $\mathbf{z}^p$ satisfying $z_i^p \geq \mathrm{ideal}_i^{Z_m}$, for all $i = 1, \ldots, k$, can be modeled by f_a. So, from the theoretical point of view, the function f_a is justified in modeling the Pareto front. However, to have $f_a(\Delta_{k-1})^* \subset \mathrm{PF}(f, S)$ in practice is extremely rare unless something concrete of the true front, for example the front is a hyperplane, is known.

A possible drawback related to the Pareto optimal objective vectors of a MOO problem and the function f_a based on vectors in Z_m is the following. A Pareto optimal objective vector $\mathbf{z}^p$ satisfying $z_i^p < \mathrm{ideal}_i^{Z_m}$ for some $i = 1, \ldots, k$ cannot be obtained by f_a because f_a is restricted to Δ_{k-1}. Such vectors might exist if Z_m does not represent the true Pareto front well enough.

$\mathcal{AM}$f4   The fourth feature is related to the simplex Δ_{k-1}. It is a convex hull defined by the standard basis vectors of $\mathbb{R}^k$, meaning that it is easy to construct also from the computational point of view. However, a simple structure for a set is not the main purpose for the simplex. Instead, no matter what the MOO problem is, the domain spaces for functions f_a and f_s are always the same (naturally, the number k of objective functions affects to the "dimension" of the simplex). Thus, the goal $\mathcal{G}1$ has been reached. In other words, it does not matter for the approximation what the number of the objective functions of the MOO problem is.

$\mathcal{AM}$f5   The fifth and the last feature is probably the most important feature for the approximation. We have not assumed anything of the MOO problem

when constructing the function f_a. We have only assumed that vectors in the set Z_m are distinct Pareto optimal objective vectors. The MOO problem behind the vectors can be nonconvex or computationally expensive! Thus, goals $\mathcal{G}2$, $\mathcal{G}3$ and $\mathcal{G}4$ have been reached.

The approximation proposed in this thesis is rather simple from the mathematical point of view. Next, we demonstrate how it can be used efficiently in the solution process of a MOO problem.

3 THE NEW APPROXIMATION IN THE SOLUTION PROCESS OF MULTIOBJECTIVE OPTIMIZATION PROBLEMS

In this chapter, we introduce and discuss in detail the IHBE method (Interactive HyperBox Exploration), which is created for the solution process of the MOO problem. The method is based on the approximation constructed in Chapter 2. By the approximation, we enable an easy way for the decision maker to learn what kind of approximated Pareto optimal objective vectors are available without having long computation times even in computationally expensive problems. In this chapter, we focus more to the actual use of the approximation in the solution process of the MOO problem, because, as noted in [65, page 33], only few papers introduce an approximation for interactive methods so that the decision maker's point of view has been considered in the use of the approximation.

To clarify terminology, we define two working environments for the solution process of the MOO problem: an *approximation environment* and a *Pareto front environment*. Generally speaking, the approximation environment means that the decision maker is using an approximation of the Pareto front. In the Pareto front environment, the decision maker deals directly with the true Pareto front without using any approximation. In what follows, sometimes the word approximation is in brackets, meaning that the sentence is meant to be understood in the Pareto front environment, but the same message in the sentence can also be considered in the approximation environment.

The importance of learning in decision making has been pointed out, for example, in [14, 131]. Through the learning, the decision maker is dealing with a constructive process, where he or she is building a conviction of what is possible. This is extremely important in the solution process, because if it happens that a preferred solution is infeasible, then the decision maker must learn about his or her preferences in relation to what is available. It will then be obvious that the decision maker has to vary his or her preferences in order to find a (an approximated) Pareto optimal objective vector to the MOO problem. The decision maker may even consider a totally different (approximated) Pareto optimal objective

vector compared to the preferred one at the beginning of the solution process. This would mean that he or she has changed his or her opinion regarding which objective function value is the most important. Therefore, the decision maker should not aim at a mathematical convergence when searching for the compromise solution during the solution process. Instead, the decision maker should aim at a psychological convergence, which means to learn about the importance of his or her preferences in relation to what kind of (approximated) Pareto optimal objective vectors are available. As a result of the psychological convergence, the decision maker might find a (an approximated) Pareto optimal objective vector, which satisfies him or her as a compromise solution to the MOO problem. The satisfaction[1] means in this thesis that the decision maker is convinced that no other Pareto optimal objective vector is better than the compromise solution. To be convinced about the compromise solution, learning about what kind of (approximated) Pareto optimal objective vectors are available is naturally needed. Otherwise, the decision maker cannot be sure why some Pareto optimal objective vector could be the compromise solution and no other is better than that particular one. For this purpose, we discuss a process called *a creation of an overall understanding of available (approximated) Pareto optimal objective vectors* in the solution process of the MOO problem.

To clarify the purpose and significance of the creation of the overall understanding of available (approximated) Pareto optimal objective vectors, we first assume that the decision maker has presented clearly at the beginning of the solution process of the MOO problem what kind of objective vector(s) he or she is preferring as the compromise solution(s) to the problem. Then we assume that a Pareto optimal objective vector(s) corresponding to the decision maker's preferences has (have) been calculated. Now, the final assumption is that the decision maker is disappointed with the Pareto optimal objective vector calculated, that is, the vector does not satisfy the decision maker. After getting over the setback, we believe that the decision maker will start making compromises and reconsidering what is really important in relation to the problem. Then, in our opinion, the only way to find a new preferred Pareto optimal objective vector is by investigating the available vectors in the (approximated) Pareto front and after this to decide what is the compromise solution to the problem based on the facts of the available vectors. Based on this, the importance of the creation of an overall understanding of available (approximated) Pareto optimal objective vectors is obvious in the solution process of the MOO problem, because the decision maker may have to decide which one would be the compromise solution in a set consisting of "not necessarily good vectors". Naturally, the situation is totally different if the decision maker can get what he or she is looking for. Then, there is not necessarily any need to learn about the preferences and the problem. However, if the decision maker can get what he or she is looking for, then he or she may still wonder or be curious to know whether or not there is something better available after all.

---

[1] The definition of a psychological convergence in this thesis is less strict compared to the definition given, for example, in [89] where trade-off rates are also included in the definition.

We have assumed that the MOO problem is computationally expensive. To beat the time challenge in calculating vectors for the creation of an overall understanding of available (approximated) Pareto optimal objective vectors, a method called *HyperBox Exploration* (HBE) is proposed in this chapter. It is based on elements of well-known interactive methods proposed in the MCDM literature, and the approximation constructed in Chapter 2. The main idea in HBE is that the decision maker asks questions in a form of a query, and then a mathematical tool (which will be introduced in Section 3.2) answers the query. By simplifying, the question in the query is the following.

I am interested in these kinds of (approximated) Pareto optimal objective vectors.
Do those vectors exist?

By the question, we focus on the creation of an overall understanding of available (approximated) Pareto optimal objective vectors by offering a simple way for the decision maker to start the exploration in the objective space. In what follows, the HBE method is constructed so that in a case, where the decision maker cannot have (approximated) Pareto optimal objective vectors he or she is interested in, the method provides an explanation why such vectors do not exist. In this way, the HBE method may support the decision maker in the learning phase because then the decision maker may realize how his or her preferences are related to the (approximated) Pareto front. At the beginning of the solution process, the decision maker might have optimistic or pessimistic expectations for the compromise solution and then he or she must be educated about the preferences. In addition, during the learning phase, the decision maker may proportion his or her expectations to what is available. On the other hand, the situation is totally different if the decision maker gets what he or she is looking for. Then the decision maker does not necessarily have to reconsider the preferences or present new ones in order to find (approximated) Pareto optimal objective vectors. However, the question in the query is over-simplified and, therefore, some conversions must be done. In what follows, we present the conversions as steps of HBE, and then we briefly discuss the steps.

Method 13 (HBE)**.**

(1) The decision maker defines his or her preferences as aspiration and reservation levels.

(2) The decision maker defines a number for vectors to be generated on the approximation.

(3) Generate vectors on the approximation so that they correspond to the aspiration and reservation levels.

(4) Present the findings (at least nondominated vectors generated in the previous step if they exist) to the decision maker.

In the first step, the decision maker presents his or her preferences in the form of aspiration and reservation levels. Then the computer can "understand" which

part of the objective space should be searched. Aspiration and reservation levels are described and discussed in more detail in Subsection 3.1.1.1. The number asked in the second step of HBE is essential in several ways. The use of the approximation created in this thesis is based on the number of sample vectors related to the decision maker's preferences. When that number is small, the decision maker does not receive too many vectors for making, for example, vector-wise comparisons at the end of the solution process after having found interesting (approximated) Pareto optimal objective vectors. However, if the number is large, the decision maker can get more detailed information of the approximated Pareto front, because a larger set of vectors represents the approximated Pareto front more accurately. A detailed discussion about the role of the number of these vectors is given in Subsection 3.1.1.2. In the third step of HBE, a mathematical model based on the approximation constructed in Chapter 2 is used to generate vectors on the approximation so that they, in some sense, correspond to the aspiration and reservation levels. The mathematical model is introduced in Section 3.2. After this, in the fourth and the last step of HBE, nondominated vectors, i.e. approximated Pareto optimal objective vectors, corresponding to the aspiration and reservation levels (if they exist) are presented to the decision maker. Additional findings related to the vectors are discussed in Section 3.3. Basically, the fourth step consists of several methods introduced in later sections and chapters in this thesis.

It must be clarified that HBE is not in itself an interactive method for the solution process of the MOO problem, because it does not allow the decision maker to redefine new aspiration and reservation levels for an overall understanding of available approximated Pareto optimal objective vectors. Redefining the levels is important, because in practice the decision maker might first check whether certain types of (approximated) Pareto optimal objective vectors are available and, if not found, look for other types for a "backup plan". Keeping this in mind, the IHBE method (Interactive HyperBox Exploration) is presented next. After this we briefly discuss the steps of IHBE.

Method 14 (IHBE)**.**

   (1) Initialize the set Z_m by calculating Pareto optimal objective vectors for the original computationally expensive MOO problem.

   (2) Generate the function f_a for the Pareto front approximation.

   (3) Validate the approximation to make it sufficiently good for the solution process of the MOO problem. If the approximation is not sufficiently good, then generate a different approximation, for example, by replacing the function f_s by another.

   (4) Collect information of the regions in the objective space the decision maker is interested in by using the HBE method.

   (5) If the decision maker needs some new information from the approximation, then go to step (4).

   (6) Calculate (if it can be done in a reasonable time) and present true Pareto optimal objective vector(s) corresponding to the approximated one(s) the

decision maker is interested in. Otherwise, stop the solution process temporarily and continue with the decision maker when the vectors are calculated and ready to be analyzed by the decision maker.

(7) If the decision maker feels that the psychological convergence still requires more (approximated) Pareto optimal objective vectors, then add the calculated true Pareto optimal objective vector(s), which does(do) not yet exist in Z_m into Z_m and go to step (2). Otherwise, stop.

Generally speaking, in steps (1)-(3) initializations for the approximation are done by the analyst. The content of step (3) will be clarified later in Chapters 4 and 5, in which we show how to do the validations computationally. After the validations, the approximation should not contain vectors, which, for example, dominate vectors in the set Z_m. A *sufficiently good approximation* is an informal definition, but to present a formal definition for a good approximation can be difficult because of the nature of nonconvex MOO problems (the Pareto front can have the most pathological "shape" a human can imagine). In cases of invalid approximations, there is no single rule what to do in order to obtain a valid approximation. However, the simplest rule is perhaps to replace the function f_s by a different one and then do the validation again. In this way, the analyst might find a proper surrogate function as discussed in Section 1.4. On the other hand, validation methods proposed in later chapters provide information which tells in which sense the approximation is invalid. It may happen that the approximation is invalid only in a very particular small region in the objective space, and therefore there is no need to generate a new one. This would mean that the approximation is sufficiently valid. Therefore, we encourage the use of common sense in validating the approximation instead of just repeating the steps of IHBE.

During the initializations, the decision maker is not needed, but, in steps (4)-(7), the decision maker is gaining overall understanding of available (approximated) Pareto optimal objective vectors while aiming at psychological convergence. To be more exact, in steps (4) and (5), the decision maker is operating in the approximation environment, meaning that he or she is dealing with the approximation. In steps (6) and (7), the environment is the Pareto front environment. In step (6), *reasonable time* cannot be exactly specified, but if the decision maker has to wait, for example, longer than is needed to drink a cup of coffee or tea, then we can start to talk about overrunning reasonable times. Step (7) is the final step of IHBE. If the IHBE method stops at that step, then the compromise solution to the MOO problem is the Pareto optimal objective vector for which the psychological convergence has been obtained, i.e., there does not exists a better vector for the vector[2].

What comes to the computation efficiency of the IHBE and HBE methods, we have not yet touched these issues. In the HBE method, the only time-consuming step might be the third one in which vectors on the approximation are calculated

---

[2] It is possible that the decision maker might not have found a compromise solution to the problem. While learning, he or she may have noticed that, for example, the formulation of the MOO problem is not a correct one, that is, an important objective or constraint has not been included in the formulation of the problem.

corresponding to the aspiration and reservation levels. However, a mathematical model introduced later in Section 3.2 is generated so that the calculation is computationally efficient. Therefore, the HBE method can be considered as a computationally efficient method in the approximation environment. The computational cost of the IHBE method is an issue that is a bit tricky. Basically, most of the steps of IHBE (i.e. (4)-(7)) in which the decision maker is involved with are dealt with in the approximation environment. However, step (6) is naturally computationally costly, because at least one Pareto optimal objective vector to the MOO problem must be calculated. This is something that cannot be avoided in practice. Therefore, the IHBE method is computationally costly, but in practice the decision maker knows that the problem is computationally expensive and because of this (we believe) he or she might try to avoid unnecessary calculations in step (6). From this point of view, the IHBE method can be considered as a computationally efficient method to be used in the solution process of the MOO problem.

The IHBE method combines learning and aspiration and reservation levels into one packet. However, the combination is not a new idea in the MCDM literature. In methods like AIM (Aspiration-level Interactive Method) [111], GUESS [24], LBS (Light Beam Search) [83] or reference point methods [131], the idea is to tell something reasonable of the preferences (in the form of aspiration and reservation levels) for the decision maker for the learning purposes. According to [111], the idea of learning something of the preferences was presented already in the fifties [117]. From the solution process point of view, learning trough the aspiration and reservation levels seems to be understandable for most of the decision makers, see e.g. [111, 167, 185, 188, 189]. What comes to the psychological convergence, studies like [23, 25, 36] support the idea of utilizing the free-search in creating an overall understanding of available (approximated) Pareto optimal objective vectors. Thus, the IHBE and HBE methods can be considered for the solution process of MOO problems.

The rest of this chapter has been organized as follows. In Section 3.1, we discuss the assumptions the decision maker must satisfy and deal with when using the IHBE and HBE methods. In Section 3.2, we propose a computational method to be used in step (3) of HBE. Through the vectors calculated in step (3) of HBE, several additional findings can be presented to the decision maker. These findings are considered in Section 3.3. Finally, Section 3.4 summarizes this chapter.

3.1 Decision Maker and the IHBE and HBE Methods

In this section, the actual use of the IHBE and HBE methods from the decision maker's viewpoint is discussed. As said, the IHBE method is an interactive method, and therefore it may contain some assumptions or features that the decision maker might see as problems in the solution process of the MOO

problem. These aspects must be discussed because interactive methods have also been criticized in the literature of MCDM. For example, the mathematics used in the method do not necessarily match the decision making, or how the decision making is used or should be used in an interactive method, see e.g. [26, 34, 35, 51, 71, 91].

This section has been organized as follows. Assumptions set on the decision maker when using IHBE and HBE are discussed in Subsection 3.1.1. The IHBE method connects together both approximation and Pareto front environments, meaning that switching from the approximation to a true Pareto front may not be that simple for the decision maker. These concerns are discussed in Subsection 3.1.2.

3.1.1 Assumptions on Decision Maker

The IHBE and HBE methods are rather simple, and it can be stated that the HBE method is the core of the IHBE because it is the method which communicates with the decision maker. Therefore, some assumptions must be set on the decision maker so that he or she would be able to use and receive support for the creation of an overall understanding of available (approximated) Pareto optimal objective vectors from the IHBE and HBE methods.

When using IHBE and HBE the decision maker is assumed to

(1) be able to express his or her preferences as aspiration and reservation levels,

(2) be able to set the number of sample vectors to be generated from the approximation,

(3) be willing to explore and learn about the objective space and (approximated) Pareto optimal objective vectors within it.

To clarify, assumptions (1) and (2) are needed in HBE and assumption (3) in IHBE. Next, we further discuss these assumptions, which we have organized in separate subsections so that in Subsection 3.1.1.X we discuss assumption (X).

3.1.1.1 Aspiration and Reservation Levels

The HBE method assumes in step (1) that the decision maker is able to define his or her preferences as aspiration and reservation levels for each objective function. Aspiration and reservation levels mean in practice that the decision maker defines the lower and upper bounds for each objective function. An aspiration level is a real value, which the decision maker would like to obtain. A reservation level is a real value, which is still admissible and should at least be achieved. In addition, every value between the aspiration and reservation level is admissible for the decision maker.

In a state-of-the-art article [131], aspiration and reservation levels are discussed and included in the category of *reference point methods* in the solution pro-

cess of MOO problems. In addition, the authors have presented the fundamentals of reference point methods also from the learning point of view. The aspects pointed out are based on the intuition of the decision maker [189] and sovereignty of the decision maker [188]. It is easier for the decision maker to handle the objective function values directly instead specifying preferences in some other ways, as e.g. weights, because objective function values are connected to the problem in a natural way. This means that the levels are more meaningful and understandable as numbers for the decision maker. Therefore, the use of aspiration and reservation levels is justified in HBE. In addition, empirical studies also support the justification, see e.g. [111, 167].

In HBE, we do not use reference point methods as they have been proposed in the literature, because the MOO problem considered in this thesis has been assumed to be computationally expensive. Typically, in a reference point method, a maximization problem (or minimization, depending on formulation) of an *achievement (scalarizing) function*[3] based on the objective functions f_i must be solved. This means that a lot of function evaluations must be done, which is time-consuming in this case. To avoid unnecessary objective functions evaluations, we propose in Section 3.2 a computationally efficient method instead of achievement functions to be used in HBE.

In principle, we are not criticizing reference point methods and achievement functions, but we believe that these methods are not the best suited for the solution process of computationally expensive MOO problems. However, criticism about and problems in using aspiration and reservation levels in the solution process can be found mainly in two categories in the literature of MCDM. In the first category, the concern is about how the decision maker can set the levels without seeing the ranges of the objective functions (see e.g. [104, 177]). The second category is related to problems in setting some parameters (e.g. weights) in optimization problems related to certain scalarization functions, see e.g. [139, 177]. The concerns in the first category are valid also in HBE and, therefore, they must be discussed. However, the issues in the second category do not exist in HBE because we do not use achievement functions or methods like goal programming[4] in HBE. Goal programming is mentioned here because, as said in [177, 179] (and noticed already in [101]), the reference point method can be regarded as a method of goal programming, or as a generalization of goal programming with certain order-consistent scalarization functions [190, page 237].

In the first category, the generation of objective function ranges in the set of Pareto optimal objective vectors forms a problem. To have the exact ranges of Pareto optimal objective vectors and in order to have the upper bounds, global optimization problems must be solved. Note that calculating the nadir vector is not a trivial problem to be solved [123, page 16]. Instead of solving the optimization problems in HBE, the ranges are generated by the minimal and maximal coordinate-wise values of the pre-calculated Pareto optimal objective vectors in Z_m. This is not necessarily the best way to generate the ranges, but it is a start as

---

[3]   Different achievement scalarizing function formulations can be found e.g. in [127, 187, 190].

[4]   The term goal programming was originally proposed in [31].

proposed in [50]. However, we want to remind that the approximation $f_a(\Delta_{k-1})$ for the Pareto front can only approximate Pareto optimal objective vectors for which the ith coordinate value can be greater than or equal to the corresponding minimum value of the coordinates of the vectors in Z_m (see $\mathcal{AM}f3$). In other words, the lower bounds are justified, because there is at least one known Pareto optimal objective vector for which at least one of the coordinate values is equal to the lower bounds. Setting the upper bounds is not that critical in HBE as long as they are at least the coordinate-wise maxima of the vectors in Z_m. Otherwise in the worst case, it might be possible that (approximated) Pareto optimal objective vectors for which the objective function values are between the aspiration and reservation levels at the same time do not exist, no matter what the levels are as long as they are in the ranges. This has been demonstrated in Figure 11, where the selected upper bounds (gray lines) are too small.
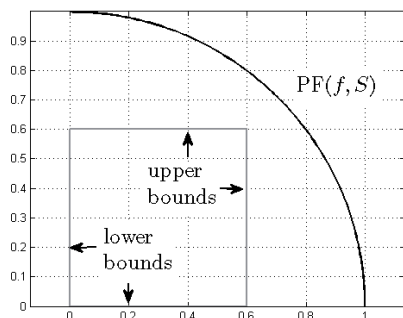
FIGURE 11   If the upper bounds selected for the ranges are too small, then it is possible that (approximated) Pareto optimal objective vectors do not exist inside the ranges.

Another concern in the first category related to the aspiration and reservation levels can be attributed to the *trial-error* problem [104, 177, 183]: how to set or find feasible levels before seeing any (approximated) Pareto optimal objective vectors even when the ranges are presented for the decision maker? In practice, the trial-error problem means that the decision maker tries some levels and then sees what happens. If the levels tried are not feasible, that is, no (approximated) Pareto optimal objective vector is found between the levels, then an error would occur. Otherwise, there is no error and the decision maker has found what he or she was looking for. In a case of an error, the decision maker must try some other levels and then see what happens. In the HBE method, it cannot be said beforehand what the feasible levels are for the decision maker because one cannot know beforehand what part of the objective space the decision maker wants to explore and study for the creation of an overall understanding of available (approximated) Pareto optimal objective vectors. However, HBE presents to the decision maker approximated Pareto optimal objective vectors (if they exists) corresponding in some sense to the aspiration and reservation levels, see Section 3.2. Thus, the problem of the trial-error has been fixed only partially in the HBE method.

However, we believe that the decision maker is able to set new aspiration and reservation levels based on the information presented to him or her.

To summarize, aspiration and reservation levels are selected to be used in HBE because it has been pointed out that these levels supports the learning and are easy to use as numbers in the interaction between a computer software and the decision maker.

3.1.1.2 The Number of Sample Vectors in the HBE Method

The second assumption set on the decision maker is related to the ability to define the number of sample vectors to be generated on the approximation. The use of the approximation is based on these sample vectors corresponding in some sense to the aspiration and reservation levels provided by the decision maker. In this section, we discuss that number from the decision maker's point of view. To put it briefly, the greater the number is, then more detailed information of the approximated Pareto front can be obtained, that is, different kinds of approximated Pareto optimal objective vectors. However, with a small number, vector-wise comparisons are easier for the decision maker. Next, we investigate briefly these two options starting from great numbers and then moving towards small(er) ones.

At the beginning of the creation of an overall understanding of available (approximated) Pareto optimal objective vectors, the decision maker may first set aspiration and reservation levels close to the bounds of the ranges for the objective functions. Then he or she may select a large number for vectors to be generated on the approximation because by a large number of vectors a larger set of different (approximated) Pareto optimal objective vectors can be obtained. There is an analogy with people: ten persons do not necessarily represent the whole nation of 100 000 people. In this way, the decision maker may receive a good distribution of different kinds of approximated Pareto optimal vectors to get an idea of what is available. However, when dealing with numbers, the capacity of processing information in a human mind must be considered.

In [132], Miller studied *the magical number seven, plus or minus two* in processing information. The special significance of that number is based on several different studies by other researches. Their studies point out that a human starts to make mistakes in judgments if there are more than 7 ± 2 alternatives[5] to remember or choose from. In addition, it is cognitively valid for the decision maker if the amount of information is low [107]. Therefore to choose, for example, from 1000 different (approximated) Pareto optimal objective vectors may highly exceed the decision maker's information processing capacity. However, before processing information, i.e. making vector-wise comparisons between (approximated) Pareto optimal objective vectors, the decision maker must have an idea of different kinds of vectors. There is no sense in starting to make vector-wise comparisons if better vectors, which would make the previous vector-wise com-

---

[5] Depending on studies, the number of alternatives may vary. However, typically, the number is from three to nine [39, 72, 116].

parisons meaningless, can be found. In other words, we believe that the decision maker makes judgments on superiority of vectors or vector-wise comparisons after having found interesting regions (or approximated Pareto optimal objective vectors) in the objective space. To find interesting regions from a large set of vectors, methods like filtering data as in [8] or visualizations [105, 112] can be used.

The number of vectors generated on the approximation is a technical detail in HBE, but in IHBE it has a significant role in creating an overall understanding of available approximated Pareto optimal objective vectors. However, setting a strict limit for the number is not that clear. In MCDM, many researchers believe that the number of vectors to be generated is a result of an individual dynamic process, see for example [122, 201]. It is also said that the fixed set of alternatives is the third of the ten myths in MCDM. Although, the fixed number of vectors to be generated is not directly related to the fixed set of alternatives, we do not see a difference between them. Therefore, the number of vectors to be generated should not be fixed by the method (or the analyst or some other method or algorithm) because dynamic changes must be possible while using IHBE and HBE. It is thus assumed that the decision maker must be able to define the number by himself or herself in HBE. We believe that he or she is always able to do that after seeing the vectors and other findings discussed in Section 3.3.1 in step (4) of HBE.

To summarize, balancing between the accuracy and vector-wise comparisons was briefly discussed from the perspective of creating of an overall understanding of available (approximated) Pareto optimal objective vector.

3.1.1.3 Willingness to Learn

The third and the last assumption set on the decision maker is the willingness to learn what kind of (approximated) Pareto optimal objective vectors are available. Compared to the two previous assumptions for the HBE method, this one is needed in steps (4) and (7) of IHBE. In the creation of an overall understanding of available (approximated) Pareto optimal objective vectors, we believe that the decision maker must be interested in learning about the problem. Therefore, we have to assume this when the decision maker is using the IHBE method. However, we would like to categorize IHBE into the interactive methods based on how it is used. If the decision maker is willing to use an interactive method for the solution process, then he or she is willing to learn about the problem [123, 131, 178]. Thus, if the decision maker is willing to use the IHBE method, then the assumption of willingness to learn holds.

Next, we briefly discuss how learning about the problem has been treated in the MCDM literature. In that discussion, we investigate both IHBE and HBE, and how they behave in relation to these issues according to the literature.

In a state-of-the-art article on learning [14], the researchers in the field of MCDM consider several aspects. In their article, several interesting questions are presented as titles of sections:

1. "What Does a Decision Maker Learn?",
2. "How Do We Know if a Decision Maker Has Learned?",

3. "How Does a Decision Maker Learn?" and

4. "How do We Investigate Learning?".

To summarize the answers given in [14], a decision maker learns what is feasible and desirable. Learning can be indicated if the decision maker can explain his or her decisions and choices. To enable learning, the decision maker must be motivated to learn, and the interaction between a computer program and a human must be easy and understandable for the decision maker. In addition, factors such as losses rather than gains may have influence on the decision maker and his or her motivation to learn. According to [14], investigation of the learning is challenging, because learning is a process of an individual.

To put it briefly, in the solution process of the MOO problem, the psychological convergence requires that the decision maker is aware of different kinds of (approximated) Pareto optimal objective vectors, and after having reached the convergence he or she knows why the compromise solution is really the compromise solution, i.e. no other Pareto optimal objective vector is better than that. This means that the first two questions above are considered in the IHBE method through psychological convergence. However, it must be clarified although that the IHBE and HBE methods are tools to be used when aiming at psychological convergence, but they do not necessarily guarantee the convergence in every case with all decision makers.

The third question is partially covered by the third assumption of the willingness to learn, but keeping up the motivation to learn is highly related to the decision maker himself or herself and to a possible user interface for the IHBE method. In the literature of MCDM, the importance of user interfaces in learning was recognized already in the eighties and nineties [101, 102, 104] and nowadays even in the applications running on the Internet [184]. As mentioned in [14], the user interface affects the motivation to learn and thus the ability to understand available (approximated) Pareto optimal objective vectors. To support and help the decision maker in this understanding, a graphical user interface and different types of visualizations [104, 105, 112, 145] for the (approximated) Pareto optimal objective vectors must be available, if needed. However, the appearance of these graphical user interfaces is out of the scope in this thesis. Therefore, the IHBE method partially provides an answer to the third question.

The fourth question is a tricky one, if we consider how the IHBE and HBE methods handle or investigate the learning of available (approximated) Pareto optimal objective vectors or overall understanding of the vectors. Handling here means that either IHBE or HBE maintain the process of learning and, based on that, are able to guide the decision maker to find new (approximated) Pareto optimal objective vectors, which are significantly different from the already studied or preferred ones. The challenges in investigating individual learning may arise from many aspects, for example, from those presented in [104]:

– "The DM wants to control the search process. Also he/she wants that support is available when needed."

– "The DM does not always want to have maximal support."

– "The DM does not always recognize the necessity of support."

The IHBE and HBE methods do not provide tools or methods to recognize when or how much support (in learning) is needed. Therefore, one can state that IHBE is not meant for learning. This can be true, and it must be admitted that a tool or a mechanism or a method created for the Pareto optimal objective vectors the decision maker has learned and investigated would be an essential element in IHBE. Then it would be possible to guide the decision maker to alternatives that should be considered at least when aiming at psychological convergence. However, creation or construction of such a mechanism can be difficult because to know beforehand the alternatives, which should be considered at least, is perhaps impossible in arbitrary MOO problems. Therefore, the decision maker controls his or her individual learning in the IHBE method. In other words, we rely on the decision maker's expertise on the topic of the MOO problem and ability to be interested in studying and learning about new possible (approximate) Pareto optimal objective vectors. It has been said that interactive methods are for creating new possible alternatives [91].

In guiding the decision maker to new areas in the objective space, machine learning can be used and so has been done in MCDM. There are positive results, see e.g. [57, 64] and references therein, on machine learning created for learning the preference structure of the decision maker. Typically, a so-called *utility function* and its updating based on the decision maker's new preferences are used for this purpose. The benefit of such an approach is clear. The utility function can be used in guiding the decision maker to study and consider some other alternative (approximated) Pareto optimal objective vectors corresponding to his or her preferences. However, we have doubts on these methods and on how they detect mistakes committed by the decision maker in judging, for example, which one of two alternatives he or she prefer. These doubts may be due to examples such as those presented e.g. in [107]. On the other hand, in practice, the decision maker might think that

"This alternative is miserable, but might satisfy me if it turns out that there are no better ones.",

which means that the decision maker is not able to decide about his or her preference without seeing other possible alternatives. To have a more concrete example, let us consider a decision maker who is unemployed. He or she has been lucky to receive a dream job offer in a neighboring city, but there is one problem. There is no public transportation between the city and the house of the decision maker, and the job will start the following week. Therefore, the decision maker has to buy a car. He or she may have to consider older ones, because a brand new car may be out of reach financially. In other words, an old rusty car might be fine for the decision maker if no better one is available within the budget. Based on these examples, we (in some sense) believe that machine learning is not necessarily suitable in the creation of the overall understanding of available (approximated) Pareto optimal objective vectors. Machine learning is (in our opinion) more suitable for problems, where mathematical convergence matters more, i.e. where a

method or an algorithm decides what is the best for the decision maker rather than psychological convergence which allows the decision maker to decide what is finally the best for him or her.

Because it is difficult to investigate the learning about the problem, then it can be asked whether or not the IHBE method helps the decision maker in aiming at the psychological convergence. Because no graphical user interface has yet been created, then we have to leave the empirical studies of IHBE use in practice to the future. However, we believe that the basis of IHBE is good, because it is based on aspiration and reservation levels as discussed in Subsection 3.1.1.1. In addition, the free-search in learning is supported by studies [23, 25, 36].

Actual willingness to learn has been briefly discussed in this subsection. Generally, it is very difficult to say how and when the decision maker should be supported in learning because the learning process is always individual. However, the importance of learning is essential in the solution process of the MOO problem because the compromise solution is a result of the learned facts no matter how much of time and patience might have been expended.

3.1.2 Possible Challenges in the IHBE Method for the Decision Maker

The IHBE method is very simple. However, IHBE operates in both approximation and Pareto front environments. As noted in [65, page 33], the use of an approximation in the solution process of MOO problems has not been investigated much. This becomes evident as follows. In articles dealing with method comparisons [9, 22, 23, 25, 36, 83, 125, 130, 183], approximations have not been used. In articles such as [50, 90, 95, 153, 159, 170], approximations are used in the solution process of the MOO problem, but there is no indication about how confident the decision maker was with the approximation used in a method. However, in the above papers where approximations were used, the convexity of the MOO problem plays a great and important role, which may mean that the approximations used were rather accurate and close to the Pareto front. This means that the approximated Pareto front can be almost equal to the true front and, therefore, it makes no difference to the decision maker whether he or she is dealing with an approximation or not. Let us remind ourselves that the IHBE and HBE methods are suited for nonconvex MOO problems, which are not that straightforward compared to convex MOO problems, as we point out in this subsection.

In this subsection, we address several elements in the IHBE method that the decision maker may consider difficult or problematic in practice when using an approximation in the solution process. To clarify, the elements discussed in this subsection are not only problems in the IHBE method: other methods like PAINT [65] and Pareto navigator [50] may suffer from these problems.

When using IHBE, the decision maker must understand that

(1) the psychological convergence is created in both the approximation environment and the Pareto front environment,

(2) the interpretation of error estimations can be challenging for him or her,

(3) he or she can decide himself or herself whether the solutions presented in step (4) of IHBE guide him or her or not and

(4) the values of decision variables are not available in the approximation environment.

Each concern listed above is discussed in separate subsections as follows. In Subsection 3.1.2.X concern (X) is discussed.

3.1.2.1 Psychological Convergence in the Approximation Environment

As said, the HBE method operates in the approximation environment. In other words, the approximated Pareto optimal objective vectors presented in step (4) of HBE are based on educated guesses of the Pareto front. Therefore, the decision maker must understand that the information for psychological convergence in steps (4) and (5) of IHBE is based on educated guesses. Occasionally, those educated guesses may not present accurate information of the true Pareto front. This is the price the decision maker must pay (or risk to take) when using approximations in the solution process of computationally expensive MOO problems.

Operating in the approximation environment with the HBE method can be questioned when the focus is on the psychological convergence. In addition, it might be thought that the psychological convergence in computationally expensive nonconvex MOO problems is pointless because of time-consuming calculations of Pareto optimal objective vectors. This means that only few vectors can be calculated in practice. In other words, the decision maker may not have an overall understanding of available Pareto optimal objective vectors for the decision about the compromise solution because those few vectors may not represent all possible and interesting Pareto optimal objective vectors. We do not know whether or not the psychological convergence should be considered in the approximation environment but we believe that investigating approximated ones is a better option than doing nothing or having long waiting times for calculating single Pareto optimal objective vectors for computationally nonconvex MOO problems. However, dealing with the approximated Pareto optimal objective vectors of a nonconvex MOO problem is not necessarily that straightforward and we illustrate this as follows.

Because the MOO problem can be nonconvex, then in a pathological example the set Z_m can be the Pareto front and no other Pareto optimal objective vector exists. In addition, there can be a hyperplane $\mathcal{H}$ so that all the vectors in Z_m may lie on it. Such a case has been illustrated in Figure 12 (vectors in Z_3 are black dots and the hyperplane $\mathcal{H}$ is the black dashed line). Now, the hyperplane can be a very good educated guess to approximate the front, as can be seen in Figure 12. In Figure 13, we have selected on purpose the image Z to be a black curve of two arcs of circles so that the Pareto optimal objective vectors are the vectors in Z_3. As it can be seen, every vector in $\mathcal{H} \setminus Z_3$ is nondominated in $\mathcal{H}$, and, therefore, they are considered by Definition 11 as approximated Pareto optimal objective vectors. However, one cannot know beforehand how far, for example, vector

$(0.5, 1.5)^{\mathrm{T}} \in \mathcal{H}$ is from the true Pareto front. In other words, in the worst case, the closest true Pareto optimal objective vectors can be too far away from $(0.5, 1.5)^{\mathrm{T}}$, which means that some of the coordinate values of the vector are not desirable for the decision maker. Then the approximation might be providing "too good promises" of the Pareto front by vector $(0.5, 1.5)^{\mathrm{T}}$ for the decision maker, even though it is a nondominated vector in the approximation.
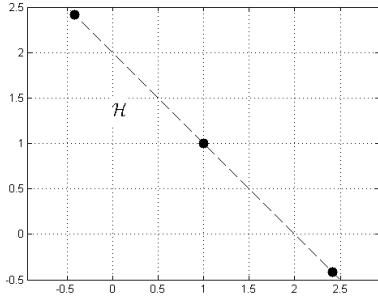
FIGURE 12   A set Z_3 included into a hyperplane.
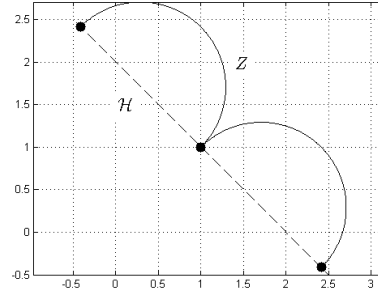
FIGURE 13   A Pareto front of a continuous nonconvex MOO problem in which only three Pareto optimal objective vectors exist.

Because approximated Pareto optimal objective vectors are only guesses, then the decision maker, in the worst case, may feel uncertain about the available approximated Pareto optimal objective vectors. To remove that uncertainty, true Pareto optimal objective vectors which correspond to the approximated ones that the decision maker is interested in are calculated in step (6) of IHBE. The only way to be sure here is by calculating Pareto optimal objective vectors. However, to convince the decision maker that some approximated Pareto optimal objective vectors are reliable, that is, that they are rather close to the Pareto front of the MOO problem, error estimations might be suggested as suitable means for the purpose. However, the error estimations can also be challenging in practice and, especially, in nonconvex MOO problems.

3.1.2.2   Challenges in Interpretations of Error Estimations in the Approximation Methods

The approximation $f_a(\Delta_{k-1})$ constructed in Chapter 2 highly depends on the surrogate function f_s. In addition, if error estimations for function f_s are available and calculated, then error estimations for the vectors in the approximation can be calculated, as in (17). Typically, error estimations for functions are calculated to get an idea of the reliability of a surrogate function in approximating some other function or phenomenon. In other words, by error estimations the decision maker may estimate how reliable an approximated Pareto optimal objective vector is. In IHBE, this reflects immediately to step (6), in which the decision maker decides

about the approximated Pareto optimal objective vectors for which corresponding true Pareto optimal objective vectors should be calculated. In this subsection, we briefly discuss the use of error estimation in step (6) of IHBE that might be occasionally challenging for the decision maker.

Let us consider the case in Figure 12 from the viewpoints of error estimation and step (6) of IHBE. The error estimations of vectors on the approximation can be very close to zero or even zero because the hyperplane describes the data in Z_m completely. In other words, the decision maker would be able to understand that the approximated Pareto optimal objective vectors represent or describe the Pareto front completely and that such Pareto optimal objective vectors corresponding to the approximated ones exist. However, in a pathological case, Pareto optimal objective vectors "very close" to approximated ones may not exist (Figure 13). This reflects immediately to step (6) of IHBE. How can the decision maker know for which approximated Pareto optimal objective vectors the corresponding true Pareto optimal objective vectors should be calculated? Note in this case that the error estimation can be misleading. A wise choice for the decision maker would be to make a list of interesting approximated Pareto optimal objective vectors and have the corresponding true ones calculated in one session as allowed by step (6) of IHBE. However, the decision maker must accept that the calculation may take some time because the MOO problem is computationally expensive.

Generally speaking, it is not clear how the error estimations in cases of nonconvex MOO problems should be handled in the solution processes of MOO problems. Typically, in convex MOO problems, approximated Pareto optimal objective vectors are either Pareto optimal objective vectors or vectors dominated by some previously unknown Pareto optimal objective vectors. This means that it does not necessarily matter what the error estimations are because the approximated Pareto optimal objective vectors are not going to be infeasible. In nonconvex problems, the approximated Pareto optimal objective vectors can also be infeasible objective vectors as seen in Figure 13. In addition, it is not clear how the error estimations may affect the decision maker, that is, how uncomfortable the decision-making may feel to him or her if error estimations are constantly being presented.

To summarize this and the previous subsection: balancing between the approximation and Pareto front environments might be challenging for the decision maker when creating the overall understanding of available (approximated) Pareto optimal objective vectors. This is because the computational cost of the MOO problem may force the decision maker to deal most of the time with approximated Pareto optimal objective vectors which are educated guesses. Uncertainty might increase or the trust to the approximation decreases if the corresponding true Pareto optimal objective vectors do not match with the approximated ones. This may happen in pathological cases, and it is not clear whether or not such situations take place when dealing with practical problems.

3.1.2.3 Full Power to Control

In step (5) of IHBE, a new region in the objective space defined by the new aspiration and reservation levels given by the decision maker is searched. However, giving new levels can be occasionally challenging for the decision maker in the solution process of the MOO problem. In the IHBE method when creating the overall understanding of available (approximated) Pareto optimal objective vectors, the decision maker decides himself or herself which regions of the objective space must be investigated in aiming at psychological convergence. This is something the decision maker must understand when using the IHBE method. Next, the full power to control the method is considered.

In the literature of MCDM, it is recommended for interactive methods that the decision maker control the method [104]. We follow this recommendation in the IHBE method. In other words, the IHBE or the HBE method do not inform the decision maker what to do in step (5) of IHBE, that is, whether to continue to step (4) or move forward to step (6). This so-called "no-guidance" can sometimes be a problem for the decision maker as reported in [23]. The decision maker may face a situation where he or she does not know what to do next or which Pareto optimal objective vector would be the compromise solution. If the previous one is the case, then one may wonder whether or not it matters for the decision maker which Pareto optimal objective vector is implemented in practice. In addition, is then an interactive method really needed in the solution process of the MOO problem with the decision maker? The question is in our opinion very deep and philosophical, and we believe that there is no single answer for it. The decision maker can become frustrated in searching the compromise solution, or be confused about the massive amount of different kinds of (approximated) Pareto optimal objective vectors, not being able to see much difference between them. It is difficult to say how often the decision maker is faced this dilemma or whether such decision makers are in a minority because the free-search has been noted to be rather successful in practice.

We believe that positive results reported in [25, 36, 111, 167, 185, 188, 189] of the AIM, GUESS and other methods that are based on aspiration and reservation levels used in the solution process of a MOO problem are due to the full power to control the method (even though it can be occasionally problematic for the decision maker). In addition, we believe that the full control and through that free-search in the objective space makes it easier for the decision maker to achieve an overall understanding of available (approximated) Pareto optimal objective vectors. In this way, the decision maker can easily find (or create or investigate) new possible (approximated) Pareto optimal objective vectors to the MOO problem. In addition, we believe that it is also easier for the decision maker to reconsider already rejected alternatives. It has been noted, for example, in [23, 104] that the decision maker may sometimes want to reconsider them. Let us remind ourselves that the decision maker may think that

"This alternative is miserable, but might satisfy me if it turns out that there are no better ones."

This illustrates the goodness of having a full controlling power over the IHBE method because "returning" to investigate old vectors is not a problem. Consequently, there is no reason to ask in steps (4) and (7) of IHBE whether the alternative(s) presented for the decision maker satisfy the decision maker, who cannot necessarily affirm that beforehand.

Full power to control the IHBE method has thus been discussed in this subsection. This full power enables the decision maker freely choose the regions in the objective space where to investigate (approximated) Pareto optimal objective vectors.

3.1.2.4 Decision Variables in the Approximation Environment

In practical MOO problems, the values of decision variables may be of importance to the decision of the compromise solution to be implemented for the MOO problem. For example, these variables can correspond to the amount of radiation in radio therapy treatment planning [133, 158], their values are describing how the compromise solution can be produced or implemented in practice.

The approximation $f_a(\Delta_{k-1})$ created in Section 2.2 and used in HBE operates only in the objective space. The approximation is based on the Pareto optimal objective vectors in the set Z_m and the simplex Δ_{k-1}. At this step, the connection to the decision space has been lost. The approximation does not use the information of the decision variables and, therefore, the decision maker cannot have that information while using HBE. This is something what the decision maker must understand. However, in step (6) of IHBE the connection to the decision space is opened again for the decision maker. When calculating a true Pareto optimal objective vector corresponding to the approximated one, the values of decision variables are calculated and available for the true ones.

3.1.3 Summary of Section 3.1

In Section 3.1 we have discussed the assumptions the decision maker must satisfy in order to use the IHBE and HBE methods. Generally, he or she must be willing to learn what kinds of (approximated) Pareto optimal objective vectors are available for aiming at the psychological convergence. From the technical point of view, while learning, the decision maker must be able to define the number of approximated Pareto optimal objective vectors to be generated on the approximation related to the preferences given by him or her in the form of aspiration and reservation levels.

Several aspects of the use of the IHBE and the HBE methods were discussed. The decision maker must understand that the information created for the psychological convergence is generated in both the approximation environment and the Pareto front environment. This means that, for example, the decision variables are not constantly available for the decision maker. On the other hand, the decision maker is allowed to enjoy a full power to control the solution process with the IHBE method.

3.2   A Method to be Used in Step (3) of HBE

The HBE method is rather simple. It consist of four steps, where in the first two steps the decision maker is asked to present preferences in the form of aspiration and reservation levels, and specify the number of sample vectors to be generated on the approximation. Then, in the last two steps, the vectors corresponding in some sense to the aspiration and reservation levels are calculated and presented for the decision maker. So, basically, HBE answers the question "Do those vectors exist?" formulated in the query

"I am interested in these kinds of (approximated) Pareto optimal objective vectors.
Do those vectors exist?"

In other words, if some of the vectors satisfy the preferences, then the decision maker can have such vectors, i.e. they exist. Then the answer is naturally "Yes, they exist". Otherwise, the answer would be "No, they do not exist", which would mean that the decision maker cannot have such (approximated) Pareto optimal objective vectors.

In this section, we propose a method for HBE's step (3), where vectors on the approximation related in some sense to the decision maker's preferences are calculated. Based on the vectors calculated, the HBE method can then answer whether or not such (approximated) Pareto optimal objective vectors exist in practice. However, in this section, we only introduce the way to calculate vectors on the approximation so that they correspond to the decision maker's preferences. Note that when the HBE method was introduced, we mentioned that it would inform the decision maker why such (approximated) Pareto optimal objective vectors do not exist if that happened to be the case. This is the topic of the next section where we introduce the use of the calculated vectors on the approximation to explain why the decision maker cannot have (approximated) Pareto optimal objective vectors asked in the query.

In order to calculate the vectors on the approximation that correspond to the preferences, first, a geometrical illustration for the correspondence between the vectors and preferences is presented in Subsection 3.2.1. The rest of this section has been organized as follows. To calculate vectors which correspond in some sense to the preferences, a geometrical sketch of the calculation is presented in Subsection 3.2.2. Then, space sampling techniques used in the calculation are discussed in Subsection 3.2.3. Subsection 3.2.4 provides the mathematical proofs for the geometrical presentations earlier. Finally, in Subsection 3.2.5, there are the concluding remarks.

3.2.1   Aspiration and Reservation Levels as a Hyperbox and a Motivation for the Method to be used in Step (3) of HBE

Geometrically, aspiration and reservation levels define a *hyperbox* in the objective space. This is illustrated in Figure 14, where on the left-hand side the levels for

a three-objective problem are presented on a line representing the corresponding objective function values. On the right-hand side of Figure 14, the hyperbox defined by the levels is illustrated by a gray box in the three-dimensional objective space.
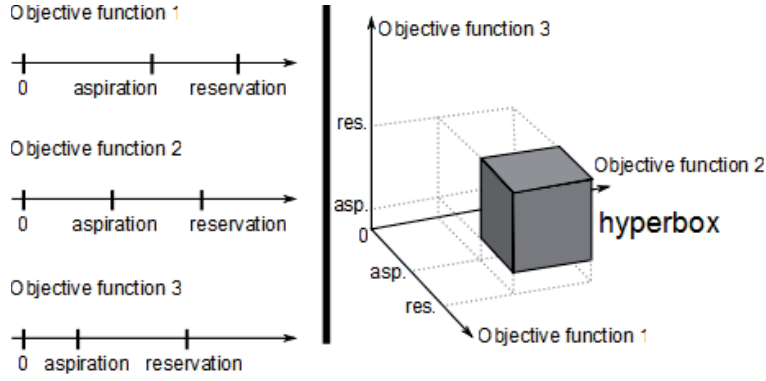
FIGURE 14 Aspiration and reservation levels for each objective function and the corresponding hyperbox in the objective space.

During the solution process of the MOO problem, the decision maker may define the levels in infinitely many ways. Figure 15 demonstrates three different hyperboxes (gray boxes) in the objective space related to the approximated Pareto front. One of them intersects the approximated Pareto front and, to be more precise, the set of nondominated vectors of the approximation. This means that there exist vectors which belong to both the approximation and the hyperbox. Thus the HBE method can answer the question in the query with "Yes, they do exist!". On the other hand, the other two hyperboxes do not intersect the approximation, which means that the answer might be, for example, "No, they do not exist!". In such a case, the vectors in the approximation corresponding to the aspiration and reservation levels should be calculated because the decision maker should be informed about where the existing approximated Pareto front is in a relation to the levels, in order to help him or her to find approximated Pareto optimal objective vectors. In addition, the vectors calculated provide an explanation for the decision maker why he or she cannot have such (approximated) Pareto optimal objective vectors asked in the query. The method for step (3) of HBE should thus be flexible enough to be able to deal with different kinds of hyperboxes, no matter where they are in the objective space.

When calculating vectors on the approximation that in some sense correspond to the aspiration and reservation levels, we propose to select them from the following set.

Definition 15. We say that vector $\mathbf{z} \in f_a(\Delta_{k-1})$ *corresponds to the aspiration and reservation levels* if there exists a vector $\mathbf{z}^l \in \mathbf{ideal}^{Z_m} + \mathbb{R}_+^k$ and a multiplier $r > 0$ so that the vector $\mathbf{z}^l$ belongs to the hyperbox (the coordinates of it are between the levels) and

$$r(z_i^l - \mathrm{ideal}_i^{Z_m}) = z_i - \mathrm{ideal}_i^{Z_m}, \quad i = 1, \ldots, k.$$
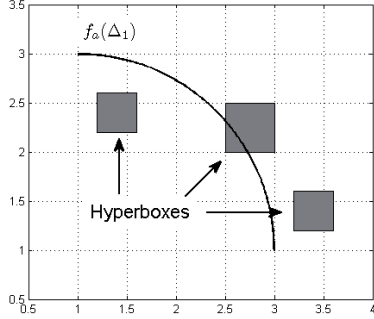
FIGURE 15   Hyperboxes can be located in several regions in the objective space, and therefore the calculation of the vectors corresponding in some sense to the aspiration and reservation levels must be done in a general way.

The set of vectors $\mathbf{z} \in f_a(\Delta_{k-1})$ corresponding to the aspiration and reservation levels is denoted by $f_a(\Delta_{k-1})^{res}_{asp}$, and it is said that set $f_a(\Delta_{k-1})^{res}_{asp}$ *corresponds to the aspiration and reservation levels*.

Geometrically, Definition 15 means that the vector $\mathbf{z}$ in Figure 16 (black dot) corresponds to the aspiration and reservation levels (gray box), because for the vector $\mathbf{z}^l$, which belongs to the hyperbox, can be found $r > 0$ so that $r(z^l_i - \text{ideal}^{Z_m}_i) = z_i - \text{ideal}^{Z_m}_i$, $i = 1, \ldots, k$. Actually, the vector $\mathbf{z}$ corresponds to the aspiration and reservation levels, if a line going through the vectors $\mathbf{ideal}^{Z_m}$ and $\mathbf{z}$ intersects the hyperbox. In Figure 17, we have illustrated several sets (bold gray curves on a black curve representing $f_a(\Delta_1)$) corresponding to different aspiration and reservation levels (the gray boxes are the same as the boxes in Figure 15).
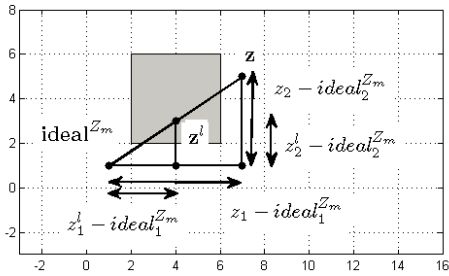
FIGURE 16   Vector $\mathbf{z}$ corresponds to the aspiration and reservation levels.
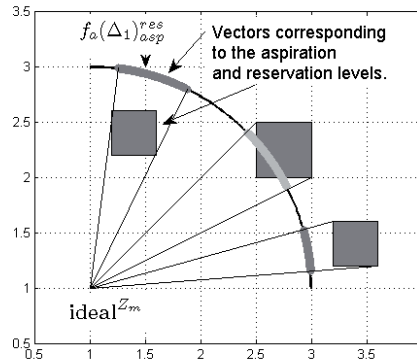
FIGURE 17   Several hyperboxes and corresponding subsets of the approximated Pareto front.

As can be seen in Figure 17, there exist nondominated vectors in the ap-

proximated Pareto front (light gray curve) that correspond to the aspiration and reservation levels (in this case, the nondominance property can be identified by methods proposed in later chapters). On the other hand, the two other sets (dark gray curves) corresponding to the aspiration and reservation levels do not intersect the hyperbox defined by the levels. In other words, by selecting a vector in the set $f_a(\Delta_{k-1})^{res}_{asp}$, the HBE method can, indeed, tell whether or not (approximated) Pareto optimal objective vectors preferred by the decision maker exist. However, it must be clarified that the method does not generate the whole set $f_a(\Delta_{k-1})^{res}_{asp}$ but a discrete subset of it. As can be seen in Figure 17, there are vectors in the light gray curve that do not belong to the corresponding hyperbox. Then, the discrete subset generated by the method proposed in this section can also contain vectors, which do not belong to the hyperbox. However, they can be easily removed from the subset, if needed. On the other hand, the set $f_a(\Delta_{k-1})^{res}_{asp}$ can contain vectors for which some of the objective function values are better than the corresponding aspiration levels and the other function values being between the corresponding levels. This is the case in Figure 17 for the vectors in the light gray set outside the hyperbox. This can be essential information for the decision maker. However, the correspondence in Definition 15 needs some further clarification.

The correspondence in Definition 15 means that the relative importances of objective functions have been considered in defining the set $f_a(\Delta_{k-1})^{res}_{asp}$. In other words, if the objective function f_i is important and f_j is less important for the decision maker, then he or she sets the aspiration and reservation levels so that for the function f_i the levels are as small as possible when compared to the levels for f_j. The vectors in the approximation corresponding to the aspiration and reservation levels are then selected so that the relative importances of objective functions are considered through values $r(z_i^l - \text{ideal}_i^{Z_m})$, $z_i - \text{ideal}_i^{Z_m}$ and conditions $r(z_i^l - \text{ideal}_i^{Z_m}) = z_i - \text{ideal}_i^{Z_m}$, $i = 1, \dots, k$. Thus, it can be stated that the set $f_a(\Delta_{k-1})^{res}_{asp}$ corresponds to the aspiration and reservation levels because the vector $\mathbf{z}^l$ satisfies the levels. Therefore for the hyperbox pointed by an arrow in Figure 17 the set $f_a(\Delta_{k-1})^{res}_{asp}$ consists only of vectors for which the first objective function value is close to $\text{ideal}_1^{Z_m}$, because the aspiration and reservation levels to the first objective function have been selected close to $\text{ideal}_1^{Z_m}$.

Now, we are ready to present a proposal for a method to be used in step (3) of HBE in calculating vectors on the approximation that correspond to the aspiration and reservation levels defined by the decision maker.

Method 16.

(1) Sample vectors within the aspiration and reservation levels (given in step (1) of HBE) so that the sample size is the number given in step (2) of HBE.

(2) Map each of the sample vectors into the approximation so that each one of the mapped sample vectors corresponds to a sample vector as in Definition 15. (That is, a mapped sample vector is the vector $\mathbf{z}$, and the corresponding sample vector is the vector $\mathbf{z}^l$ in Definition 15.)

Before introducing the explicit formulas of step (2) in Method 16, we present the mapping in a geometrical way. Then we briefly discuss the tools, such as space sampling techniques, required in the realization for sampling. After this, we confirm that the formulas used in step (2) generate vectors on the approximation so that the vectors really correspond to the aspiration and reservation levels.

3.2.2 Generating Geometrically Vectors on the Approximation Corresponding to the Aspiration and Reservation Levels

To answer the question "Do they exist?" in the query of the HBE method, we start an investigation to find out whether or not (approximated) Pareto optimal objective vectors exist in the hyperbox, i.e. whether the hyperbox and the set $f_a(\Delta_{k-1})^*$ intersect. In step (2) of HBE, the decision maker is asked to state how many vectors he or she would like to have on the approximation to be generated. Let that number be n_h. Then, the hyperbox is sampled by this number of sample vectors. The sample vectors can be selected randomly in the simplest case, but more advanced space sampling (hyperbox sampling) techniques are presented in the next subsection.

Let the sample vectors be denoted by

$$\mathbf{z}^{h_i}, \quad i = 1, \ldots, n_h. \tag{18}$$

Figure 18 illustrates sample vectors $\mathbf{z}^{h_i}$ as black dots and a hyperbox as black lines, indicating that we have passed step (1) of Method 16.
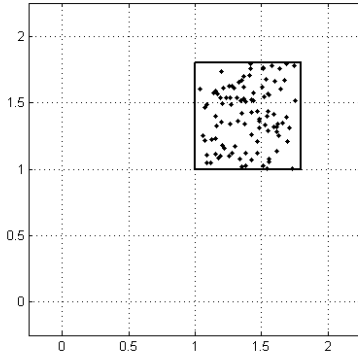
FIGURE 18 Randomly selected sample vectors $\mathbf{z}^{h_i}$ in a hyperbox in the objective space.
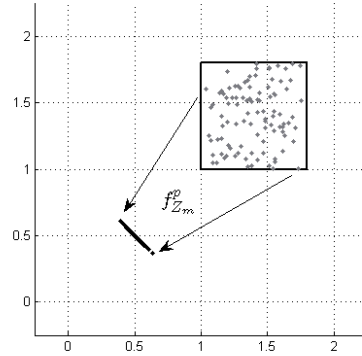
FIGURE 19 Sample vectors $\mathbf{z}^{h_i}$ mapped into the simplex Δ_1.

Once the sample vectors have been generated, they are mapped into the simplex Δ_{k-1} used in the construction of the approximation. The mapping for this purpose is the same function $f_{Z_m}^p$ that is used to map Pareto optimal objective vectors in Z_m into the simplex Δ_{k-1}. Now, let the mapped sample vectors on the
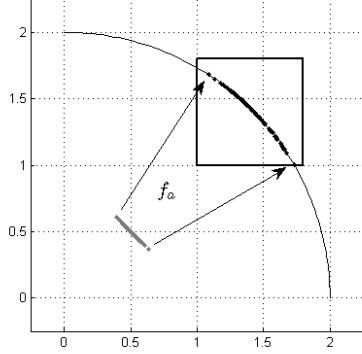
FIGURE 20 Vectors $\mathbf{z}^{sh_i}$ mapped by function f_a into the approximated Pareto front.
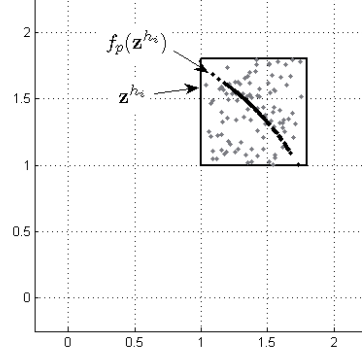
FIGURE 21 Sample vectors $\mathbf{z}^{h_i}$ mapped into $f_a(\Delta_1)$.

simplex be denoted by

$$\mathbf{z}^{sh_i} := f_{Z_m}^p(\mathbf{z}^{h_i}), \quad i = 1, \ldots, n_h. \tag{19}$$

In Figure 19, we have depicted vectors $\mathbf{z}^{sh_i}$ as black dots. Now, because $\mathbf{z}^{sh_i} \in \Delta_{k-1}$, $i = 1, \ldots, n_h$, we naturally use function f_a to map them into the approximation[6]. This is the case in Figure 20. In fact, we have used Method 12 for the sample vectors in the simplex. Thus, a function for generating vectors on the approximation in relation to the aspiration and reservation levels is the function

$$f_p := f_a \circ f_{Z_m}^p \tag{20}$$

and the sample vectors originally sample in the hyperbox. In other words, we have passed step (2) of Method 16, and with the help of the vectors $f_p(\mathbf{z}^{h_i})$ we can answer the question "Do those vectors exist?", as already discussed in the previous subsection.

In Figure 21, where we have mapped vectors $\mathbf{z}^{h_i}$ by function f_p (the vectors $f_p(\mathbf{z}^{h_i})$ are depicted as black dots, and the gray dots being vectors $\mathbf{z}^{h_i}$). As can be seen, the approximated Pareto front related to the aspiration and reservation levels has been found by vectors $f_p(\mathbf{z}^{h_i})$ quite well. In fact, we have proposed a computational method to be used in investigating whether two sets containing infinitely many vectors intersect. This is not a trivial problem to be solved. A good coverage of the approximated Pareto front in Figure 21 was obtained because the sample vectors $\mathbf{z}^{h_i}$ were initially covering the hyperbox well: that is, they are initially everywhere in the hyperbox. In Figure 22, we have not sampled the hyperbox well enough, and because of that we have only partially obtained vectors in the approximated Pareto front related to the aspiration and reservation levels. Here, the Pareto front is the same as in Figures 18- 21.

---

6   In the case of Figure 21, the approximated Pareto front is an arc of an origin-centered circle with radius 2 in the positive orthant.
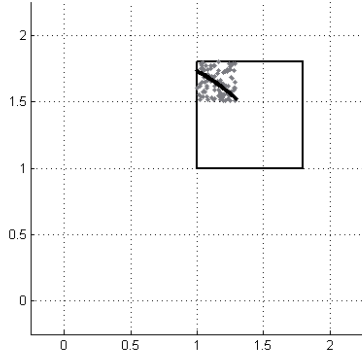
FIGURE 22 If the sample vectors are not distributed everywhere in the hyperbox, then the possible approximated Pareto front can be found partially.
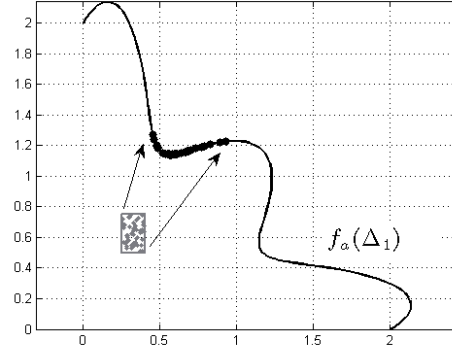
FIGURE 23 Mapped sample vectors do not belong to the hyperbox.

The reader may wonder that the "end-parts" of the approximated Pareto front (parts of the Pareto front close to the edges of the hyperbox) in Figure 20 are not necessarily well represented by vectors $f_p(\mathbf{z}^{h_i})$. In MOO problems consisting of two objective functions, visualization as in Figure 20 can be done. Then the decision maker "sees" such behavior. However, let us consider a MOO problem consisting of ten objective functions and a hyperbox, which intersects the set $f_a(\Delta_9)^*$. In this particular case, it is not necessarily that simple to "see" whether such "end-parts" exist because there can be arbitrary many of them. Note that a discrete set may not always represent a subset of a hypersurface well enough. On the other hand, we do not know whether or not the decision maker is interested in knowing how well the vectors represent the intersection. Also, it is not always that simple to find computationally the intersection of the hyperbox and the set $f_a(\Delta_{k-1})^*$. However, if it happens that the decision maker is really interested in "seeing" whether or not such "end-parts" exist, then he or she can make a new query targeting a region he or she is suspecting to be an "end-part" by specifying new aspiration and reservation levels there. Note that if vectors $f_p(\mathbf{z}^{h_i})$ do not exist in a certain region in the hyperbox, then it does not necessarily mean that approximated Pareto optimal objective vectors do not exist in that particular region. In Figure 22, a resampling (a new query) has been done in one of the "end-parts" and, as it can be seen, approximated Pareto optimal objective vectors have been found. To cover the intersection well enough at first hand, there must be a sufficient large number of sample vectors $\mathbf{z}^{h_i}$ initially covering the hyperbox, but to say how many vectors must be sampled naturally depends on the approximated Pareto front. If there is a lot of variety in the approximated Pareto front, then more sample vectors are probably needed.

A technical advantage of Method 16 is the computational efficiency of the

function f_p in generating vectors on the approximation in relation to the aspiration and reservation levels ($\mathcal{AM}$f1 and the function $f_{Z_m}^p$ do not contain computationally expensive operations). Note that Method 16 does not contain optimization problems to be solved likewise the reference point methods [131] discussed in Subsection 3.1.1.1 require. However, it must be mentioned that vectors $f_p(\mathbf{z}^{h_i})$ may not necessarily be uniformly distributed on the set $f_a(\Delta_{k-1})$ limited to the aspiration and reservation levels. This can be seen, for example, in Figure 21, where vectors $f_p(\mathbf{z}^{h_i})$ have been clustered to the middle part of the approximation. In order to remove similar $f_p(\mathbf{z}^{h_i})$ in the set of the sample vectors by methods created for clusters, we refer to [12]. However, a clear drawback in the method is the property to generate vectors which are sometimes dominated in the approximated Pareto front. This can be seen in Figure 23. However, in this case the dominated vectors can be removed, and in general cases most of dominated vectors can be detected, see Chapters 4 and 5. We continue with the case presented in Figure 23 in Subsection 3.3.3, where we discuss how the sample vectors are used in supporting and helping the decision maker in the solution process.

Next, we study in more detail the whole process of generating vectors $f_p(\mathbf{z}^{h_i})$. We start by the vectors $\mathbf{z}^{h_i}$ defined in (18), and then we continue with the function f_p defined in (20) to make sure that it maps vectors $\mathbf{z}^{h_i}$ into the approximation so that the mapped ones correspond to the aspiration and reservation levels as defined earlier.

3.2.3 On Space Sampling Techniques

With the aspiration and reservation levels in step (1) of HBE, a hyperbox has been defined in the objective space. Then the question is: do the hyperbox and the approximation intersect each other, i.e does there exist at least one vector, which belongs to both sets? In the geometrical illustration of Method 16, a space sampling technique was used at the beginning to generate sample vectors. Note, that vectors $\mathbf{z}^{h_i}$ were selected randomly for simplicity and this technique is a simple space sampling technique. However, sampling the hyperbox by random vectors may not be the best technique for high-dimensional objective spaces.

Sampling the hyperbox by random vectors may cause some unexpected situations, for example, the vectors do not cover the hyperbox well enough. This can be shown by a simple probability calculation. Let us assume that we have a ten-dimensional hyperbox $[0, 10]^{10}$. Then the probability that a uniformly distributed random vector $\mathbf{v}^r$ belongs to $[1, 9]^{10}$ (= the "interior" of the hyperbox) is

$$p(\mathbf{v}^r \in [1, 9]^{10}) = \frac{8^{10}}{10^{10}} \approx 0.1074.$$

This means that approximately every tenth random vector is in the "interior" of the hyperbox, and it must be mentioned that the margin from the "interior" to the boundary of $[0, 10]^{10}$ is quite small. In other words, most of the random vectors in $[0, 10]^{10}$ belong to the "margin", that is, a set $[0, 10]^{10} \setminus [1, 9]^{10}$ (note that the margin as a set contains most of the hyperbox in the sense of a volume) of the

hyperbox[7].

If space sampling by random vectors does not produce a good distribution in the hyperbox, then a regular grid in the hyperbox will sample the hyperbox uniformly. However, the number of grid points increases exponentially when the dimension increases. To illustrate this, let us mention that there are $2^{10} = 1024$ corner vectors and $11^{10} = 25\,937\,424\,601$ vectors of form $(i_1, i_2, \ldots, i_{10})^T$, where i_j is integer from 0 to 10, in the ten-dimensional hyperbox discussed above.

Now by the simple illustrations by numbers, it can be stated that space sampling can be challenging in high-dimensional objective spaces. However, quasi-random sequences like Halton [63], Hammersley [88] or Sobol [21] can be used and are recommended to be used in sampling the hyperbox as vectors $\mathbf{z}^{h_i}$. Basically, they combine the randomness of random vectors and put the vectors in some sense regularly in the objective space so that in small regions there do not exist several similar vectors as it is possible with random vectors. In other words, these vectors combine the randomness and a regular grid into one packet meaning that the sample vectors are initially covering the hyperbox well. Further discussion on the sequences can be found, for example, in [115] and references therein.

Once we have generated the sample vectors, we have passed the first step of Method 16. Then it is time to consider the second step.

3.2.4 Function f_p Related to the Aspiration and Reservation Levels

Function f_p defined in (20) is rather simple. It uses the same functions defined in Subsection 2.2 and therefore it is computationally efficient. However, the property to generate vectors on the approximation $f_a(\Delta_{k-1})$ corresponding to the aspiration and reservation levels requires further considerations. In other words, we must show that the mapped sample vectors belong to $f_a(\Delta_{k-1})_{asp}^{res}$.

Let us start from the simplest case.

Theorem 17. Let $\mathbf{z}^h \in f_a(\Delta_{k-1})$. Then $f_p(\mathbf{z}^h) = \mathbf{z}^h$.

Proof. Because $\mathbf{z}^h \in f_a(\Delta_{k-1})$, then there is a vector $\mathbf{z}^{sh} \in \Delta_{k-1}$ so that $f_a(\mathbf{z}^{sh}) = \mathbf{z}^h$. In other words, there exists a scalar $\lambda_{sh} > 0$ so that $\mathbf{ideal}^{Z_m} + \lambda_{sh}\mathbf{z}^{sh} = \mathbf{z}^h$. Now,

$$f_{Z_m}^p(\mathbf{z}^h) = f_{Z_m}^p(\mathbf{ideal}^{Z_m} + \lambda_{sh}\mathbf{z}^{sh}) = \frac{\lambda_{sh}\mathbf{z}^{sh}}{\sum_{i=1}^k \lambda_{sh} z_i^{sh}} = \frac{\lambda_{sh}\mathbf{z}^{sh}}{\lambda_{sh}\sum_{i=1}^k z_i^{sh}}$$
$$= \mathbf{z}^{sh}.$$

Thus,

$$f_p(\mathbf{z}^h) = f_a \circ f_{Z_m}^p(\mathbf{z}^h) = f_a(\mathbf{z}^{sh}) = \mathbf{z}^h.$$

$\square$

---

[7] This case is opposite to the situation in Figure 21, where most of the mapped sample vectors lie on the middle part of the approximated front related to the levels. In a higher-dimensional space, most of the mapped random vectors may not cluster to the "interior", because the volume for the "margin" is much larger.

To see that $f_p(\mathbf{z}^h) \in f_a(\Delta_{k-1})^{res}_{asp}$, select $r = 1$ in Definition 15. In other words, approximated Pareto optimal objective vectors satisfying the aspiration and reservation levels belong to $f_a(\Delta_{k-1})^{res}_{asp}$. For later use, Theorem 17 provides a useful result.

Corollary 18. If the function $f^p_{Z_m}$ has been limited on $f_a(\Delta_{k-1})$, then the function f_a is the inverse function of $f^p_{Z_m}$.

Proof. The result follows immediately from Theorem 17. □

To simplify the notation, let $\mathcal{HB}$ denote the hyperbox defined by the aspiration and reservation levels and

$$\mathcal{HBA} := f_a(\Delta_{k-1}) \cap \mathcal{HB},$$
$$\mathcal{HBA}^* := f_a(\Delta_{k-1})^* \cap \mathcal{HB}.$$

In other words, if $\mathcal{HBA} \neq \varnothing$, then every vector in $\mathcal{HBA}$ is a fixed point of function f_p, that is, $f_p(\mathbf{z}) = \mathbf{z}$. However, a sample vector $\mathbf{z}^h$ does not, usually, belong to $f_a(\Delta_{k-1})$, meaning that the situation in Theorem 17 is rare in practice. In other words, it is quite rare that a sample vector belongs initially to the approximated Pareto front.

In cases where the hyperbox does not intersect the set $f_a(\Delta_{k-1})$ (or the sample vectors do not belong to the approximation), we have to show that there exists a multiplier $r > 0$ for a sample vector and the corresponding mapped vector so that they correspond to each other as required in Definition 15. However, what joins vectors $\mathbf{z}^h$, $\mathbf{ideal}^{Z_m} + \mathbf{z}^{sh}$ and $f_p(\mathbf{z}^h)$ together is a line and this line is the key element in finding the multiplier r. In Figure 24, vectors $\mathbf{z}^h$, $\mathbf{ideal}^{Z_m} + \mathbf{z}^{sh}$ and $f_p(\mathbf{z}^h)$ are depicted as black dots. The line to which the vectors belong is the black dashed line.
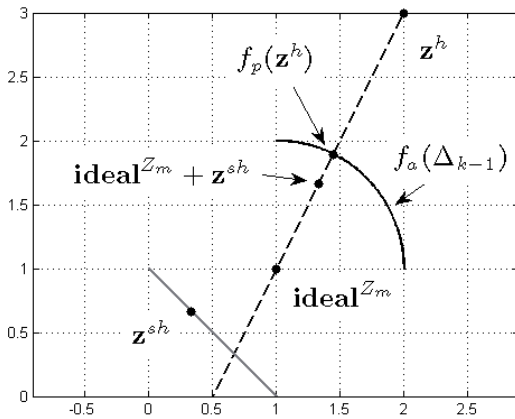
FIGURE 24   A line joining vectors $\mathbf{z}^h$, $\mathbf{ideal}^{Z_m} + \mathbf{z}^{sh}$ and $f_p(\mathbf{z}^h)$ together in the objective space.

In other words, it is clear that there are scalars $\lambda_1 > 0$ and $\lambda_2 > 0$ so that

$$\mathbf{z}^h = \mathbf{ideal}^{Z_m} + \lambda_1 \mathbf{z}^{sh} \quad \text{(definition of function } f_{Z_m}^p)$$

$$f_p(\mathbf{z}^h) = \mathbf{ideal}^{Z_m} + \lambda_2 \mathbf{z}^{sh} \quad \text{(definition of function } f_a).$$

Then

$$\mathbf{z}^{sh} = \frac{1}{\lambda_2}(f_p(\mathbf{z}^h) - \mathbf{ideal}^{Z_m}),$$

$$\mathbf{z}^h = \mathbf{ideal}^{Z_m} + \frac{\lambda_1}{\lambda_2}(f_p(\mathbf{z}^h) - \mathbf{ideal}^{Z_m}),$$

which means that for each $i = 1, \ldots, k$, we have

$$z_i^h - \text{ideal}_i^{Z_m} = \frac{\lambda_1}{\lambda_2}(f_p(\mathbf{z}^h)_i - \text{ideal}_i^{Z_m}) =: r(f_p(\mathbf{z}^h)_i - \text{ideal}_i^{Z_m}).$$

This means that $f_p(\mathbf{z}^h) \in f_a(\Delta_{k-1})_{asp}^{res}$, and we have proved that the function f_p in Method 16 produces vectors corresponding to the aspiration and reservation levels in the sense of Definition 15.

3.2.5 Summary of Section 3.2

In Section 3.2, we have presented a method to be used in step (3) of HBE for the calculation of vectors on the approximation that correspond to the aspiration and reservation levels defined by the decision maker. Technically, first, a hyperbox defined by the aspiration and reservation levels is sampled by sample vectors $\mathbf{z}^{h_i}$, which can be selected randomly or as Halton, Hammersley or Sobol sequences. Then the vectors $\mathbf{z}^{h_i}$ are mapped into the approximation by the function f_p, which is a computationally efficient function. Thus, the method (Method 16) consist of the generation of sample vectors $\mathbf{z}^{h_i}$ and the use of function f_p. This means in practice that the decision maker does not have long waiting times in calculation of approximated Pareto optimal objective vectors corresponding to his or her aspiration and reservation levels in the solution process of the MOO problem even when the problem is computationally expensive. Method 16 completes the HBE method to be used in practice, in the IHBE method. Next, the topic would be the interpretation of the sample vectors generated in step 3 of HBE.

3.3 Additional Information for the Decision Maker

In Section 3.1 we discussed the assumptions the decision maker must satisfy in order to use the IHBE and HBE methods. In Section 3.2, we proposed a method to be used in HBE's step (3) in which vectors on the approximation corresponding to the aspiration and reservation levels defined by the decision maker are calculated. In other words, the IHBE and HBE methods are ready for action in practice. However, we have not yet considered how the HBE method could explain why

the decision maker cannot have (approximated) Pareto optimal objective vectors if those vectors asked in the query do not exist. Now, sample vectors $\mathbf{z}^{h_i}$ and the corresponding vectors $f_p(\mathbf{z}^{h_i})$ in the approximation calculated are used for this purpose. They are used for characterizing a type of the hyperbox, and with the help of the type of the hyperbox, the HBE method can explain why those vectors do not exist.

In naming the type of the hyperbox, we investigate the aspiration and reservation levels. It is possible that the decision maker cannot have the (approximated) Pareto optimal objective vectors he or she is interested in because the levels are either optimistic or pessimistic. By relating the optimism or pessimism of the levels, the HBE method can explain in an intuitive way for the decision maker why he or she cannot have those (approximated) Pareto optimal objective vectors he or she is interested in. When the sample vectors and the type of the hyperbox are presented to the decision maker, he or she might be able to understand why he or she cannot have such vectors, and then he or she can set more realistic levels in order to find (approximated) Pareto optimal objective vectors. This can be essential at the beginning of the solution process in educating the decision maker to set feasible levels as discussed the error-trial problem in Subsection 3.1.1.1. In addition, a lot of other information that can help the decision maker to learn about the approximated Pareto front are calculated for the sample vectors. Thus, in this section, we consider additional findings which are based on the sample vectors calculated in step (3) of HBE.

In what follows, we have organized the discussion into separate subsections. The contents of the subsections depend on the hyperbox and its location in the objective space in relation to the approximated Pareto front. We start with an *optimistic hyperbox* in Subsection 3.3.1. Briefly, an optimistic hyperbox means that the decision maker has set too optimistic values for each objective function to be achieved simultaneously, i.e. objective vectors satisfying the levels are infeasible. In Subsection 3.3.2, the topic a *pessimistic hyperbox*, an opposite hyperbox of an optimistic hyperbox. In the case of a pessimistic hyperbox, the decision maker can improve all the objective function values related to the aspiration levels at the same time. Thus improving some objective function value does not mean that some other objective function value must get worse. However, in cases of optimistic and pessimistic hyperboxes, it is shown that the subsets of nondominated vectors in the approximated Pareto front do not intersect the hyperbox. Then in Subsection 3.3.3, other types of hyperboxes, which intersect the approximated Pareto front, are discussed. Finally, in Subsection 3.3.4, additional numerical information such as approximated trade-off rates is discussed.

To clarify the notation in the following subsections, let the vector $\mathbf{z}^a$ be the vector of aspiration levels given in step (1) of HBE, that is, z_i^a is the ith aspiration level. Similarly, let the vector $\mathbf{z}^r$ be the vector of reservation levels.

3.3.1 Optimistic Hyperbox

Let us start the discussion of an optimistic hyperbox with a geometrical illustration in a two-dimensional objective space and then continue with general cases in higher-dimensional objective spaces. In Figure 25, the approximated Pareto front has been depicted by a black curve and an optimistic hyperbox by a gray box (note that in this case $f_a(\Delta_1) = f_a(\Delta_1)^*$). As can be seen in Figure 25, it is not possible to have approximated Pareto optimal objective vectors for which all objective function values are between the aspiration and reservation levels[8].
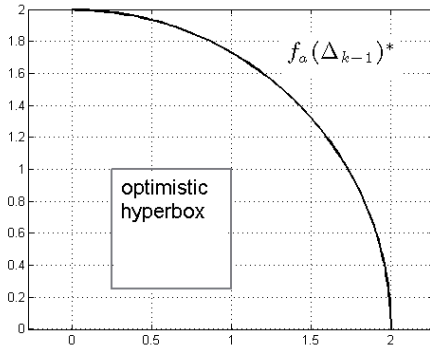
FIGURE 25   An optimistic hyperbox in the objective space.

More formally, an optimistic hyperbox is defined as follows.

Definition 19. A hyperbox defined by the aspiration and reservation levels in step (1) of HBE is called an *optimistic hyperbox*, if $\mathbf{z}^r$ dominates some vector $\mathbf{z} \in f_a(\Delta_{k-1})^*$.

First of all, let us prove that if a hyperbox is optimistic, then the decision maker cannot have an approximated Pareto optimal objective vector so that every coordinate value of the vector is between the corresponding aspiration and reservation levels.

Theorem 20. Let a hyperbox $\mathcal{HB}$ be optimistic. Then $\mathcal{HB} \cap f_a(\Delta_{k-1}) = \varnothing$ (i.e. $\mathcal{HBA} = \varnothing$).

Proof. On the contrary, let $\tilde{\mathbf{z}} \in \mathcal{HBA}$. Since the hyperbox is optimistic, then the vector $\mathbf{z}^r$ dominates some vector $\bar{\mathbf{z}} \in f_a(\Delta_{k-1})^*$. In other words, the vector $\tilde{\mathbf{z}}$ dominates the vector $\bar{\mathbf{z}}$, which is a contradiction. Thus, an optimistic hyperbox does not intersect the set $f_a(\Delta_{k-1})$. □

Theorem 20 proves that an optimistic hyperbox such as in Figure 25 does not contain approximated Pareto optimal objective vectors. In addition, Theorem 20

---

8   An optimistic hyperbox does not mean that the objective function values between the aspiration and reservation levels do not exist separately for each objective.

and Definition 19 clarify that the term optimistic can be used for an optimistic hyperbox. In other words, the decision maker can now be informed in step (4) of HBE about the infeasible region in the objective space he or she is investigating. Basically, it is enough to tell the decision maker that the region defined by the levels is infeasible. This explains why he or she cannot have (approximated) Pareto optimal objective vectors in the query. Thus, the HBE method can explain to the decision maker that the aspiration and reservation levels are optimistic.

The definition presented for an optimistic hyperbox in Definition 19 is obvious, but in practice it can be quite difficult to identify computationally whether a hyperbox is optimistic or not because a vector in $f_a(\Delta_{k-1})^*$ dominated by $\mathbf{z}^r$ should be found. However, in step (3) of HBE, sample vectors $\mathbf{z}^{h_i}$ in the hyperbox and corresponding vectors $f_p(\mathbf{z}^{h_i})$ in the approximation are calculated. We use them now to investigate whether or not a hyperbox is optimistic.

Method 21. If the three following conditions hold,

(1) every sample vector $\mathbf{z}^{h_i}$ dominates the corresponding $f_p(\mathbf{z}^{h_i})$,
(2) vectors $f_p(\mathbf{z}^{h_i})$ do not belong to the hyperbox, and
(3) the vector $\mathbf{z}^r$ dominates some of the vectors $f_p(\mathbf{z}^{h_i})$,

then the hyperbox can be optimistic. If the vector $f_p(\mathbf{z}^{h_i})$ dominated by the vector of reservation levels belongs to $f_a(\Delta_{k-1})^*$, then the hyperbox is optimistic.

We have demonstrated Method 21 in characterizing whether a hyperbox is optimistic or not in Figure 26. As can be seen, every vector $\mathbf{z}^{h_i}$ dominates the corresponding $f_p(\mathbf{z}^{h_i})$ and none of the vectors $f_p(\mathbf{z}^{h_i})$ belong to the hyperbox. In addition, there exists a nondominated vector $f_p(\mathbf{z}^{h_i})$ in the set $f_a(\Delta_1)$ so that the vector of reservation levels dominates it. Thus, the hyperbox is optimistic.

To identify whether or not a vector $f_p(\mathbf{z}^{h_i})$ is nondominated in $f_a(\Delta_{k-1})$, methods introduced in Chapters 4 and 5 are required. Note that to conclude the optimism of the hyperbox by using only the vector of reservation levels and the corresponding vector in the approximation is not enough. The corresponding vector can be dominated in $f_a(\Delta_{k-1})$ as it is in Figure 26 (the gray dots connected by dashed line). Therefore, the whole set of sample vectors $\mathbf{z}^{h_i}$ is proposed to be used in identifying the optimism.

Method 21 is rather simple. However, its simplicity means that the method may occasionally identify a hyperbox to be optimistic that is not. This is the case in Figure 27, where dominated vectors in the approximated Pareto front cannot be used for a correct identification. Thus, more studies and developments are required in the future for identifying computationally whether a hyperbox is optimistic or not in nonconvex MOO problems. However, Method 21 is a computationally efficient way to characterize the optimism by using already calculated vectors in step (3) of HBE.

So far, an optimistic hyperbox defined by the aspiration and reservation levels given in step (1) of HBE has been discussed. Briefly, an optimistic hyperbox lies between the vector $\mathbf{ideal}^{Z_m}$ and the set $f_a(\Delta_{k-1})$ meaning that the decision maker has set too optimistic levels for approximated Pareto optimal objective
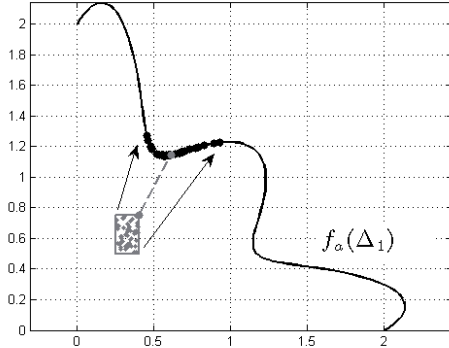
FIGURE 26   An example to demonstrate how to identify an optimistic hyperbox.
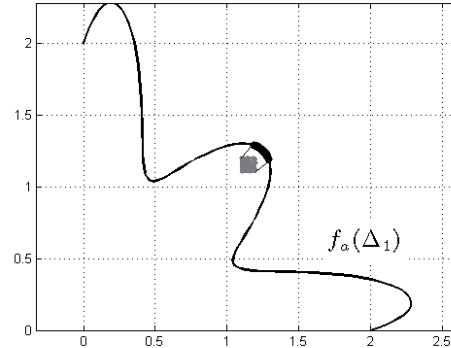
FIGURE 27   A hyperbox which might be identified as optimistic.

vectors. In practice, the decision maker is investigating an infeasible region in the objective space. This may happen at the beginning of the solution process of the MOO problem because the decision maker might not have a clear understanding of available (approximated) Pareto optimal objective vectors. Informing the decision maker about the optimism of the preferences could be essential, especially, when creating an overall understanding of available (approximated) Pareto optimal objective vectors. To identify an optimistic hyperbox, a very simple computational method has been proposed. By Method 21, the HBE method can give a clear explanation for the decision maker about why he or she cannot have (approximated) Pareto optimal objective vectors if the aspiration and reservation levels were optimistic.

3.3.2 Pessimistic Hyperbox

Occasionally at the beginning of the solution process of a MOO problem, the decision maker may set pessimistic values for the aspiration and reservation levels. In such a case, every objective function value can be improved at the same time without impairing any of them. In other words, the decision maker is investigating a region in the objective space that is an "opposite" of an optimistic hyperbox. Informing the decision maker about the pessimism of the levels can be essential in the creation of an overall understanding of available (approximated) Pareto optimal objective vectors. In other words, pessimism explains why the decision maker cannot have (approximated) Pareto optimal objective vectors asked in the query.

Theoretically, a pessimistic hyperbox can be defined as follows.

Definition 22. A hyperbox defined by the aspiration and reservation levels in step (1) of HBE is called a *pessimistic hyperbox*, if there exists a vector $\mathbf{z} \in f_a(\Delta_{k-1})^*$ so that it dominates vector $\mathbf{z}^a$.

As with an optimistic hyperbox, we obtain a similar result for a pessimistic hyperbox. In other words, a pessimistic hyperbox does not intersect the set $f_a(\Delta_{k-1})^*$.

Theorem 23. If a hyperbox is pessimistic, then $\mathcal{HB} \cap f_a(\Delta_{k-1})^* = \varnothing$ (i.e. $\mathcal{HBA}^* = \varnothing$).

Proof. On the contrary, let us assume that $\tilde{\mathbf{z}} \in \mathcal{HBA}^*$. Since the hyperbox is pessimistic, then there is a vector $\bar{\mathbf{z}} \in f_a(\Delta_{k-1})^*$ so that it dominates the vector of the aspiration levels. In other words, vector $\bar{\mathbf{z}}$ dominates $\tilde{\mathbf{z}}$, which is a contradiction. Thus, a pessimistic hyperbox does not intersect the set of nondominated vectors of $f_a(\Delta_{k-1})$. □

Theorem 23 means in practice that a pessimistic hyperbox is located "beyond" the nondominated vectors of the approximated Pareto front from the perspective of the vector $\mathbf{ideal}^{Z_m}$ as depicted in Figure 28. Therefore, labeling the hyperbox pessimistic is justified.
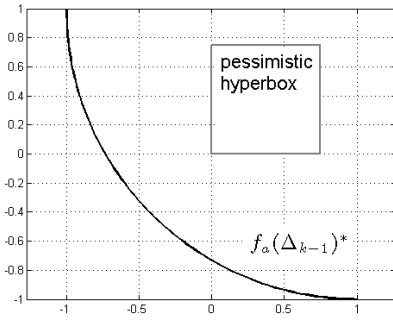
FIGURE 28   A pessimistic hyperbox in the objective space.

To identify whether a hyperbox is pessimistic or not, the characterization can be done as follows.

Method 24. If some of the following holds

(1) $f_p(\mathbf{z}^a)$ dominates the vector of aspiration levels $\mathbf{z}^a$, or
(2) some of the sample vectors $f_p(\mathbf{z}^{h_i})$ dominates vector $\mathbf{z}^a$,

then the hyperbox is pessimistic.

We have illustrated Method 24 in Figure 29 in order to identify the pessimism of a hyperbox. As can be seen, there are vectors which are dominating the vector of aspiration levels, i.e. the hyperbox is pessimistic. However, Method 24 might not find all the pessimistic hyperboxes, and therefore, more developments must take place in future studies. In Figure 30, we have illustrated a pessimistic hyperbox, which cannot be identified by Method 24. However, the proposed technique is computationally efficient and is based on pre-calculated vectors in identifying the pessimism.
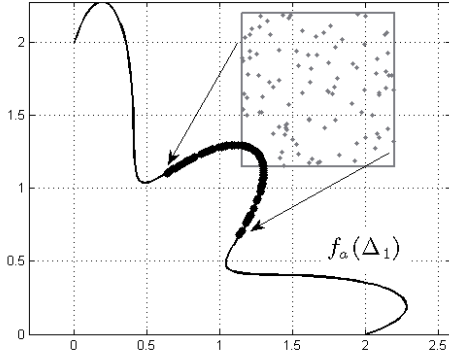
FIGURE 29 An example to demonstrate how to identify a pessimistic hyperbox.
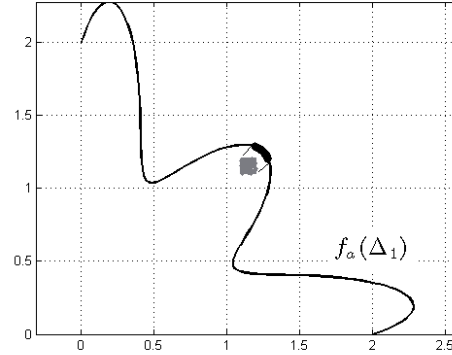
FIGURE 30 A pessimistic hyperbox in an objective space.

To conclude this subsection, a pessimistic hyperbox has been discussed. In cases of pessimistic hyperboxes, there exists at least one vector in $f_a(\Delta_{k-1})^*$, which dominates the vector of aspiration levels. This means that the decision maker can improve each objective function value without sacrificing any of them. In creating the overall understanding of available approximated Pareto optimal objective vectors, this additional information can be useful for the decision maker in step (4) of HBE. We have also proposed a simple method for identifying pessimistic hyperboxes. In other words, the HBE method can explain in such a case for the decision maker why he or she cannot have the (approximated) Pareto optimal objective vectors he or she is interested in.

3.3.3 Neither Optimistic nor Pessimistic Hyperbox

If the hyperbox and the set $f_a(\Delta_{k-1})^*$ intersect, then the decision maker has found approximated Pareto optimal objective vectors between the aspiration and reservation levels. However, we do not label such hyperbox because it is clear that the decision maker has found a region of the approximated Pareto front he or she is interested in. It is up to the decision maker now whether he or she wants investigate the vectors any further. For that eventuality, the next subsection covers additional numerical information about the vectors.

In step (6) of IHBE, calculations are made for the Pareto optimal objective vectors corresponding to approximated ones the decision maker is interested in. Occasionally, the true ones may not be close to the approximated ones, as discussed in Subsection 3.1.2.1. In such cases, the decision maker may start to wonder about the reason for it. This is one of the topics discussed in this subsection. We also set up a specific case for which additional information must be created and presented to the decision maker in step (4) of HBE. The information can be essential in interpreting "strange looking" results in nonconvex MOO problems.

Basically, this special information concerns about a special hyperbox which is neither optimistic nor pessimistic. Note that there are the hyperboxes which are neither optimistic nor pessimistic, just like the one described in the previous paragraph.

We have assumed that the MOO problem is nonconvex. This means that the set of approximated Pareto optimal objective vectors can be depicted as in Figure 31 (black lines). A hyperbox defined by the aspiration and reservation levels can be located in the objective space, as it is presented in the same figure. In this kind of situation, the decision maker should be informed about the region in the objective space where Pareto optimal objective vectors do not exist and where the hyperbox is neither optimistic nor pessimistic. Note that the hyperbox could contain (approximated) Pareto optimal objective vectors because it is, in a way, in the middle of the (approximated) Pareto front (the vectors could be, for example, on the gray dashed line). Example presented in Figure 31 introduces a special case. However, to present an exact definition for the hyperbox in this case, philosophical issues must first be considered. Therefore, a more detailed discussion will follow in Chapter 6.
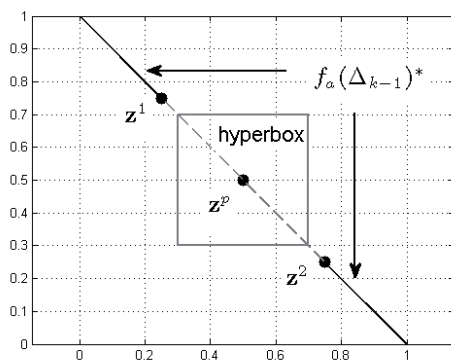
FIGURE 31   A hyperbox located in an objective space so that it does not intersect the set of approximated Pareto optimal objective vectors and where it is neither optimistic nor pessimistic.
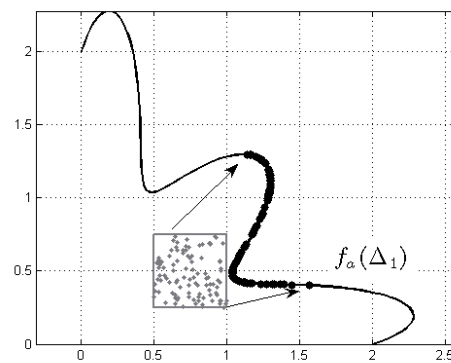
FIGURE 32   A hyperbox in an objective space, which is neither optimistic nor pessimistic nor intersecting the approximated Pareto front.

The case in Figure 31 introduces also a challenge, which might occur in step (6) of IHBE, and we can demonstrate it as follows. Let us assume that the approximation also contains the gray dashed line in Figure 31, but let the Pareto front be the two black lines (i.e. the set $f_a(\Delta_1)^*$). It may happen that the decision maker is interested in the approximated Pareto optimal objective vector $\mathbf{z}^p$ in step (6) of IHBE. He or she might be interested in knowing what is in some sense a closest true Pareto optimal objective vector to $\mathbf{z}^p$. If, for example, reference point methods [131] are used in the step, then the decision maker may receive the vector

$\mathbf{z}^1$. Even though the first objective function value at $\mathbf{z}^1$ is better than the corresponding value at $\mathbf{z}^p$, the decision maker may not be happy with $\mathbf{z}^1$ because the second objective function value is too high. By slightly changing the vector $\mathbf{z}^p$, the decision maker may receive the vector $\mathbf{z}^2$, for which he or she may have corresponding feelings. However, after having seen the vectors $\mathbf{z}^1$ and $\mathbf{z}^2$, the decision maker may start to wonder why he or she cannot get anything else between the vectors and why the second vector is totally different when compared to the first vector. The decision maker should now be informed clearly about the region in the objective space where Pareto optimal objective vectors do not exist and where the region is neither optimistic nor pessimistic. This would enable the decision maker to understand why, for example, a Pareto optimal objective vector generated by the reference point method may be a way off from an approximated one.

The cases discussed above are rather theoretical, and the first one, requiring philosophical discussions, can be handled by the approximation, see Chapter 6. However, the second case described in the previous paragraph is here an unsolved problem (this problem might be difficult one as already discussed in Subsection 3.1.2.2). The value of having been informed about the situation in the second case might be significant for the decision maker as it enables him or her to continue the investigation in other regions in the objective space instead of the original one. We do not yet know how this information could be generated by computationally efficient methods. Note that even though reference point methods can be used but in this case they could be computationally extremely costly.

In this subsection, two different types of hyperboxes have been discussed (the one which intersects the set $f_a(\Delta_{k-1})^*$ and the one which does not but is neither optimistic nor pessimistic). Here the idea is not to introduce all possible variations of them - even as a collection of hyperboxes they do not necessarily cover all possible hyperboxes in the objective space. In Figure 32, we have illustrated a hyperbox, which is not one of the labeled and discussed hyperbox in this section. Cases seen in Figure 32 may occur especially in nonconvex MOO problems, but creating straightforward rules in naming such hyperboxes can be difficult. However, as an explanation for the decision maker about why he or she cannot have (approximated) Pareto optimal objective vector asked in the query, the HBE method can inform him or her that that the sample vectors do not belong to the hyperbox and the sample vectors generated originally in the hyperbox dominate the mapped ones in the approximation. That is, the decision maker might have set aspiration and reservation levels too optimistic.

To conclude this subsection, we can say that the nonconvexity of the MOO problem introduces special cases in which creating an explanation for the decision maker can be challenging. Some of the challenges can be solved computationally, but some of them are still open for good ideas.

3.3.4 Additional Numerical Information for the Decision Maker

Hyperboxes discussed in this section provide additional information for the decision maker about how the hyperbox is located in the objective space in relation to the approximated Pareto front. Naturally, this information is not necessarily the only information the decision maker is interested in, especially once he or she has found the most promising and interesting regions of the (approximated) Pareto front. In other words, the decision maker may want to have some more specific information for the (approximated) Pareto optimal objective vectors in the interesting regions.

In this subsection, we assume that the sample vectors generated in the third step of the HBE method satisfy the decision maker's aspiration and reservation levels. In general, the additional numerical information can be calculated for the sample vectors even though they do not necessarily satisfy the preferences.

The following numerical information can be included into HBE in order to help the decision maker reach the psychological convergence after having achieved in some sense, an overall understanding of available (approximated) Pareto optimal objective vectors:

1. The closest vectors $\mathbf{z}^j \in Z_m$ corresponding to nondominated vectors $f_p(\mathbf{z}^{h_i})$ and the corresponding decision variable values for each $\mathbf{z}^j$.
2. A local approximated trade-off rates[9] for the sample vectors on the approximation.
3. A certainty number (defined later) for the sample vectors on the approximation (the number describes how reliable the vector is when it is considered as a Pareto optimal objective vector).

With the numerical information, the decision maker can focus the investigation to the details of the (approximated) Pareto optimal objective vectors by doing, for example, vector-wise comparisons. The first piece of numerical information is obvious, but the second one can divide opinions in the field of MCDM. Methods like ISWT [30], GDF [54], GRIST [193] and Z-W [202] are based on trade-off rates, but the use of trade-off rates can also be difficult for decision makers, see [25, 36, 92, 107, 177]. However, we want to include the approximated trade-off rates into the HBE method because then it is possible to investigate variations on the approximated Pareto front. With the help of trade-off rates, the decision maker can obtain information about approximated improvements for some objective function values if some other is impaired by one unit. However, this is local information based on tangent hyperplanes, meaning that the approximated improvement based on the hyperplane is a very rough one and does not necessarily approximate the reality well. For example, the derivative value of function $f(x) = x^2$ at the origin is zero, but this does not mean that no changes would take place in the function values when moving away from the origin. We present the calculation of approximated trade-off rates in Chapter 4.

---

[9] Here, we discuss local *approximated* trade-off rates, because the rates are based on the approximation. Detailed discussion of trade-off rates can be found e.g. in [123, 131]

The third piece of numerical information can also divide opinions. As discussed in Subsection 3.1.2.1, error estimations are educated guesses and they may not represent the truth. However, they are convenient for the decision maker in establishing the reliability of the approximated Pareto optimal objective vector in question. A proposal for such a number in a general case is presented in Chapter 5.

Let us clarify that the additional numerical information listed above is not necessarily the only information that can be presented in step (4) of HBE. In various steps of IHBE and HBE, data and numerical information have been calculated, and, in this thesis, the aim is to use that efficiently for several purposes. The listed numerical information consists of by-products of some other calculations. For example, the methods for identifying the type of a hyperbox (optimism, pessimism) are by-products of vectors calculated in step (3) of HBE. In this way, the IHBE and HBE methods aim at being as computationally efficient as possible, without destroying the capability to be informative for the decision maker in the solution process. However, there are some limitations to the numerical information that is to be presented to the decision maker.

Additional numerical information for aiming at psychological convergence in the solution process can be both positive and negative. If too much information is presented, then the decision maker may become saddled with a cognitive load [107]. On the other hand, if too little information is presented, then the decision maker may become misled as pointed out in [104]. Thus, although the literature of MCDM knows about the possible drawbacks of extremes, it seems that pinpointing the limit for a reasonable amount of information is still incomplete in that field of research. However, if we assume that the decision maker gets a sufficient amount of information, then the actual presentation of additional numeric information remains as the challenge for the user interface between a human decision maker and a computer software. A handful set of various kinds of nice visualizations [105, 112] and interfaces for selecting the information to be visualized should be made available. As said earlier, future studies will provide a graphical interface and deal with the outcomes of HBE and IHBE.

In the IHBE and HBE methods, the decisions about the amount of information to be presented are made as follows. Because it is not clear how much information should be presented at a time, then the decision maker decides what information is presented. In this way, the decision maker can proportion the information to his or her capacity to handle information. This, in turn, is based on his or her knowledge about the (approximated) Pareto front and available (approximated) Pareto optimal objective vectors. After having found interesting regions in the (approximated) Pareto front, the decision maker may first want to examine some basic information about, for example, the so-far closest known true Pareto optimal objective vector to an approximated one. As the understanding about the (approximated) Pareto optimal objective vectors increases, then the decision maker may increase the amount of information in order to see the differences between the vectors and be able to select a compromise solution. In other words, it is the decision maker's responsibility to choose the information to be

presented in the IHBE and HBE methods. In this way, the decision maker enjoys the power of control over the methods, but problems discussed in the previous paragraph may still remain.

In this subsection, additional numerical information to be presented in step (4) of HBE for the sample vectors the decision maker is interested in was discussed. By calculating, for example, approximated trade-off rates, the decision maker may get more details of the approximated Pareto front for aiming at psychological convergence. In general, numerical information has been proposed to be used in the final steps in the creation of an overall understanding of available (approximated) Pareto optimal objective vectors, but if the decision maker wants that information to be available at the beginning, it is possible in the HBE method.

3.3.5 Summary of Section 3.3

For the decision maker analyzing the MOO problem at hand, we have proposed the use hyperboxes defined by the aspiration and reservation levels given in step (1) of HBE as information to be presented at the beginning of the creation of an overall understanding of available (approximated) Pareto optimal objective vectors. After having created some overall understanding, the decision maker may want to have more detailed information for the vectors. Such information can consist of, for example, values of decision variables of the closest known true Pareto optimal objective vectors or approximated trade-off rates.

3.4 Summary of Chapter 3

Chapter 3 has been organized so that the reader can have an idea of the approximation created in Chapter 2 and its computationally efficient use in the solution process of computationally expensive continuous nonconvex MOO problems. In Chapter 3, we have discussed how the two last goals ($\mathcal{G}8$ and $\mathcal{G}9$) set on page 20 can be achieved with the approximation. In other words, we have shown how well-known and empirically studied methods used in the literature of MCDM can be applied and to be compatible with the approximation. In addition, we have discussed what kind of information the approximation can provide for the decision maker. Next, we summarize the main topics discussed in Chapter 3.

In the solution process of the MOO problem, psychological convergence was selected as the main goal. For this convergence, the creation of an overall understanding of available (approximated) Pareto optimal objective vectors has been discussed. To achieve that understanding, first, a method called Hyper-Box Exploring was proposed. It is a method created for investigating whether or not the decision maker will be able to obtain the approximated Pareto optimal objective vectors he or she is interested in. A special feature of the method is its ability to present an explanation about why the decision maker cannot have those approximated Pareto optimal objective vectors in case they do not exists. Then

a method called Interactive HyperBox Exploring uses the HBE method interactively in order to create an overall understanding of available (approximated) Pareto optimal objective vectors.

In the solution process, when using HBE and IHBE, the decision maker must satisfy several assumptions. Briefly, the decision maker should be able to express his or her preferences in a form of aspiration and reservation levels. Then the decision maker should give the number of sample vectors he or she would like to have generated on the approximation. The third and the last assumption is the willingness to learn what kinds of (approximated) Pareto optimal objective vectors are available.

A computationally efficient method was constructed for the step (3) of HBE, where sample vectors in the approximation related to the decision maker preferences are calculated. Technically speaking, space sampling techniques and functions used in Chapter 2 are the basis of that method. As a bonus, the method provides by-products based in the sample vectors that can be used in helping the decision maker to achieve psychological convergence in the solution process of the MOO problem.

4 NONDOMINANCE AND NORMAL VECTORS

The approximation created in Chapter 2 is a simple and computationally efficient model to approximate Pareto fronts. In Chapter 3, the IHBE method was proposed for the solution process. In step (3) of IHBE, the approximation is validated by the analyst in order to guarantee that it does not contain vectors which, for example, dominate the Pareto optimal objective vectors in Z_m. If the approximation has not been validated before the decision maker starts to use it in the solution process, then it is possible that the approximation may provide unrealistic visions of possible compromise solutions that cannot be implemented or produced in practice.

The main goal in this chapter is to introduce two methods for validating the approximation. Basically, the idea is to find computational techniques to identify whether or not a vector in the approximation is nondominated. The methods for validations can be based on these findings. In this chapter, normal vectors on the approximation and their connections to dominated and nondominated vectors in the approximation are investigated. A brief geometrical definition of a normal vector would state that it is a vector that belongs to space $\mathbb{R}^k$, and it is *perpendicular* to a *tangent hyperplane* at $\mathbf{z} \in f_a(\Delta_{k-1})$ to the approximated Pareto front. In this thesis, a normal vector at $\mathbf{z}$ is depicted by a black arrow, and in figures it has typically been shifted to $\mathbf{z}$.

We can demonstrate, geometrically, the connection between normal vectors and nondominated or dominated vectors in the approximation as follows. Figure 33 illustrates how normal vectors can be used in identifying whether a vector on the approximation is either dominated or nondominated. As can be seen, a normal vector (black arrow) in Figure 33 shifted to $\mathbf{z}^1 \in f_a(\Delta_1)$ belongs to the interior of negative orthant $\mathbb{R}^2_-$. In addition, the vector $\mathbf{z}^1$ is (locally) nondominated in $f_a(\Delta_1)$ as can be seen in the figure. On the other hand, a normal vector (black arrow) at $\mathbf{z}^2$ does not belong to the negative orthant and the vector $\mathbf{z}^2$ is dominated in $f_a(\Delta_1)$ as it can be seen in Figure 33. In this chapter, these findings are generalized with mathematical proofs to higher-dimensional objective spaces. Then the results are extended to also cover subsets of the approximation instead of single vectors in the approximation.
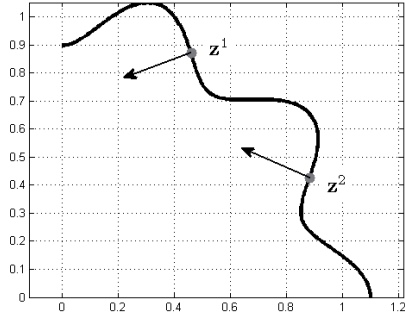
FIGURE 33   A simple example to illustrate how normal vectors can be used in investigating whether a vector in $f_a(\Delta_{k-1})$ is dominated or nondominated. Here vector $\mathbf{z}^1$ is locally nondominated and vector $\mathbf{z}^2$ is a dominated.

The results obtained with the normal vectors are significant as they enable one to say whether a vector in the approximation is nondominated, i.e., an approximated Pareto optimal objective vector. It is obvious that these results are used in validating the approximation. In the solution process, normal vectors are also calculated for the approximated Pareto optimal objective vectors the decision maker is investigating. Then approximated trade-off rates discussed in Section 3.3 can be calculated efficiently with the help of the normal vectors without solving any optimization problems. This means that the decision maker can obtain useful information of the approximated Pareto optimal objective vectors in step (4) of IHBE without having long computation times that the use of the by-products of normal vectors would entail.

Normal vectors are investigated in this chapter as follows. In Section 4.1, we establish a mathematical basis for calculating normal vectors, and then we present the main results related to the Pareto dominance. After this, in Section 4.2, we consider the results from the viewpoints of IHBE and HBE methods, that is, how to implement the results in a computationally efficient way for the methods. Finally, in Section 4.3, we present a brief summary.

4.1   Normal Vectors in Characterizing Pareto Dominance

In this section, we investigate normal vectors on the approximated Pareto front and their connection to the Pareto dominance. In the MCDM literature, such studies have already been done, for example, in [128, 163]. It turns out that the connection in question exists. However, in what follows, we do not assume the MOO problem to be convex or certain constraint conditions to hold at a Pareto optimal decision vector. This means that the results presented in this thesis are more general compared to the results presented in [49, 70, 128, 163].

To get started, we define the basis for calculating normal vectors as follows.

Definition 25. *Relative interior* of Δ_{k-1} is a set denoted by

$$\Delta_{k-1}^{\circ} := \operatorname{rel\,int}(\Delta_{k-1}) = \{\mathbf{v} \in \Delta_{k-1} : v_i > 0 \text{ for all } i = 1, \ldots, k\}.$$

Correspondingly, set $f_a(\Delta_{k-1}^{\circ})$ is called a *relative interior* of $f_a(\Delta_{k-1})$.

In order to calculate normal vectors, we have to assume that the function f_s in the function f_a is continuously differentiable on Δ_{k-1}° (for example, a radial basis function with polyharmonic spline of a third degree is such surrogate function). Then a *normal vector* at $f_a(\mathbf{z})$, $\mathbf{z} \in \Delta_{k-1}^{\circ}$, is a vector

$$f_a(\mathbf{z})^{\perp} := \begin{bmatrix} \frac{\partial f_s}{\partial z_1}(\mathbf{z}) & \cdots & \frac{\partial f_s}{\partial z_k}(\mathbf{z}) \end{bmatrix} \begin{bmatrix} g_1^1(\mathbf{z}) & \cdots & g_k^1(\mathbf{z}) \\ \vdots & \ddots & \vdots \\ g_1^k(\mathbf{z}) & \cdots & g_k^k(\mathbf{z}) \end{bmatrix} - \mathbf{1}^{\mathrm{T}}, \qquad (21)$$

where

$$g_j^i(\mathbf{z}) = \begin{cases} \frac{1-z_i}{f_s(\mathbf{z})}, & i = j \\ -\frac{z_i}{f_s(\mathbf{z})}, & i \neq j \end{cases}.$$

If for all $f_a(\mathbf{z})$, $\mathbf{z} \in \Delta_{k-1}^{\circ}$, there exists a normal vector, then the set $f_a(\Delta_{k-1}^{\circ})$ is called *smooth*.

Remark 26. Typically, normal vectors are defined through a local affine approximation to a manifold, see e.g. [175], and that is the case also for the function formulation (21) derived in Appendix 1.

In addition, the calculation of normal vectors by function 21 is also valid for nonconvex MOO problems. In other words, it does not matter whether or not the MOO problem is convex. This means that normal vectors illustrated in Figure 33 are calculated by using function (21).

Normal vectors define several sets in the objective space $\mathbb{R}^k$. Let $\mathbf{z} \in \Delta_{k-1}^{\circ}$. Then a set

$$\mathcal{T}_{\mathbf{z}} := \{\tilde{\mathbf{z}} \in \mathbb{R}^k : \left\langle f_a(\mathbf{z})^{\perp}, \tilde{\mathbf{z}} \right\rangle = 0\}$$

is called a *tangent space*. A *tangent hyperplane* at $f_a(\mathbf{z})$ is a set

$$\mathcal{T}_{f_a(\mathbf{z})} := \{\tilde{\mathbf{z}} \in \mathbb{R}^k : \left\langle f_a(\mathbf{z})^{\perp}, \tilde{\mathbf{z}} - f_a(\mathbf{z}) \right\rangle = 0\}.$$

In Figure 34, we have illustrated the normal vector $f_a(\mathbf{z})^{\perp}$ at $f_a(\mathbf{z})$ by a black arrow. The gray dashed line in the figure is the tangent hyperplane $\mathcal{T}_{f_a(\mathbf{z})}$. As it can be seen, the normal vector is perpendicular to the tangent hyperplane (note that the tangent hyperplane is a shifted tangent space, which is actually a vector space, see [175]). The example seen in Figure 34 also introduces a smooth set $f_a(\Delta_{k-1}^{\circ})$. In other words, geometrically, a smooth set $f_a(\Delta_{k-1}^{\circ})$ does not contain shapes such as "spikes" or "angles" or "corners" as in Figure 35 at $f_a(\mathbf{z})$.

The tangent hyperplane $\mathcal{T}_{f_a(\mathbf{z})}$ is essential in many cases because it introduces a local affine approximation to $f_a(\Delta_{k-1}^{\circ})$ at $f_a(\mathbf{z}) \in f_a(\Delta_{k-1}^{\circ})$. According
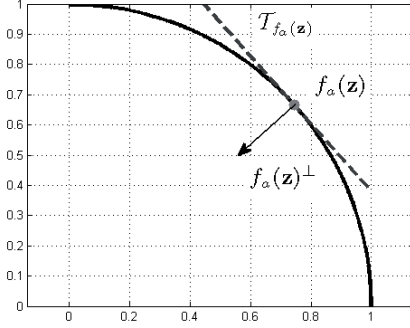
FIGURE 34   A normal vector $f_a(\mathbf{z})^\perp$ and the corresponding tangent hyperplane $\mathcal{T}_{f_a(\mathbf{z})}$ at $f_a(\mathbf{z})$ in the objective space $\mathbb{R}^2$.
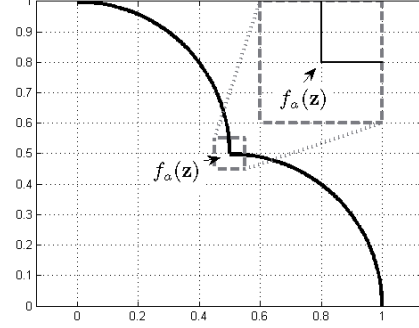
FIGURE 35   Shapes like "spikes", "angles" or "corners" do not exist in a smooth set $f_a(\Delta_{k-1}^\circ)$.

to Theorems 5-1, 5-2 and Problems 2-29, 5-13 in [175], we have an open neighborhood W of $f_a(\mathbf{z})$ and a differentiable function $g : \mathcal{T}_{f_a(\mathbf{z})} \cap W \to \mathbb{R}$ so that $\tilde{\mathbf{z}} + g(\tilde{\mathbf{z}})(f_a(\mathbf{z})^\perp)^{\mathrm{T}} \in f_a(\Delta_{k-1}^\circ)$, where $\tilde{\mathbf{z}} \in \mathcal{T}_{f_a(\mathbf{z})} \cap W$, see Figure 36. Then, for example,

$$\lim_{t \to 0} \frac{t\mathbf{z}^d + g(t\mathbf{z}^d)(f_a(\mathbf{z})^\perp)^{\mathrm{T}} - f_a(\mathbf{z})}{\left\| t\mathbf{z}^d + g(t\mathbf{z}^d)(f_a(\mathbf{z})^\perp)^{\mathrm{T}} - f_a(\mathbf{z}) \right\|} = \frac{\mathbf{z}^d}{\|\mathbf{z}^d\|}, \quad \mathbf{z}^d \in \mathcal{T}_{f_a(\mathbf{z})} \cap W, \text{ and}$$

$$\lim_{\bar{\mathbf{z}} \to f_a(\mathbf{z})} \left\langle \frac{f_a(\mathbf{z})^\perp}{\|f_a(\mathbf{z})^\perp\|}, \frac{\bar{\mathbf{z}} - f_a(\mathbf{z})}{\|\bar{\mathbf{z}} - f_a(\mathbf{z})\|} \right\rangle = 0, \quad \bar{\mathbf{z}} \in f_a(\Delta_{k-1}^\circ) \cap W,$$

(22)

where the first limit describes a directional derivative, and the second limit says that normalized vectors of vector sequences converging to $f_a(\mathbf{z})$ are becoming perpendicular to the normalized normal vector $f_a(\mathbf{z})^\perp$.

We are now ready to start investigating relations between normal vectors and Pareto dominance. First, we show a condition which is characteristic of a situation where a vector in the approximated Pareto front is dominated by some other vector in the approximation.

Theorem 27. Let $f_a(\Delta_{k-1}^\circ)$ be smooth and vector $\mathbf{z} \in f_a(\Delta_{k-1}^\circ)$. If $\mathbf{z}^\perp \notin \mathbb{R}_+^k \cup \mathbb{R}_-^k$, then $\mathbf{z} \notin f_a(\Delta_{k-1}^\circ)^*$.

Proof. Because $\mathbf{z}^\perp \notin \mathbb{R}_+^k \cup \mathbb{R}_-^k$ we have non-empty index sets I_+ and I_- so that $z_j^\perp < 0$ when $j \in I_-$ and $z_j^\perp > 0$ when $j \in I_+$. Let $r_\mathbf{z}$ be selected so that $B(\mathbf{z}, r_\mathbf{z}) \subset \mathbf{ideal}^{Z_m} + \mathrm{int}\,\mathbb{R}_+^k$. Now, let vector $\mathbf{z}^d \in \mathcal{T}_\mathbf{z}$ be selected so that $z_j^d < 0$, for all $j = 1, \ldots, k$, and $\mathbf{z} + \mathbf{z}^d \in B(\mathbf{z}, r_\mathbf{z})$ (the existence of the vector $\mathbf{z}^d$ follows from $\mathcal{T}_\mathbf{z}$ and non-empty index sets I_+, I_-).

Now, because $f_a(\Delta_{k-1}^\circ)$ is smooth at $\mathbf{z}$, then according to (22) for each $|t| < r$, a sufficiently small $r > 0$, there exists a scalar λ_t so that $\mathbf{z} + t\mathbf{z}^d + \lambda_t \mathbf{z}^\perp \in f_a(\Delta_{k-1}^\circ)$
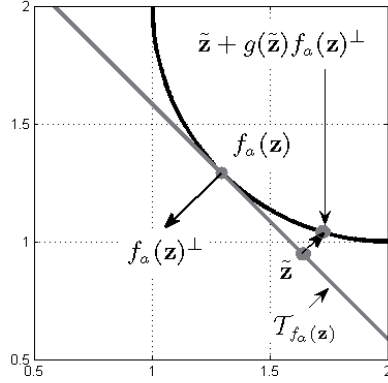
FIGURE 36   A smooth approximated Pareto front can be locally represented from the tangent hyperplane by a differentiable function g.

and

$$\lim_{t \searrow 0} \frac{\mathbf{z} + t\mathbf{z}^d + \lambda_t(\mathbf{z}^\perp)^{\mathrm{T}} - \mathbf{z}}{\left\| \mathbf{z} + t\mathbf{z}^d + \lambda_t(\mathbf{z}^\perp)^{\mathrm{T}} - \mathbf{z} \right\|} = \frac{\mathbf{z}^d}{\left\| \mathbf{z}^d \right\|}.$$

Because $z_j^d < 0$ for all $j = 1, \ldots, k$, then there exists $\tilde{t} \in (0, r)$ so that

$$\mathbf{z}_j + \tilde{t}\mathbf{z}_j^d + \lambda_{\tilde{t}}\mathbf{z}_j^\perp < z_j \text{ for all } j = 1, \ldots, k.$$

In other words, the vector $\mathbf{z} + \tilde{t}\mathbf{z}^d + \lambda_{\tilde{t}}\mathbf{z}^\perp \in f_a(\Delta_{k-1})$ dominates the vector $\mathbf{z}$, which means that $\mathbf{z}$ does not belong to $f_a(\Delta_{k-1}^\circ)^*$. $\qquad\square$

Theorem 27 proves that the vector $\mathbf{z}^2$ introduced in Figure 33 is dominated, but it does not say what or where the dominating vector is. However, we immediately obtain a more important result, which starts to show how the Pareto dominance is related to the normal vectors.

Theorem 28. Let $f_a(\Delta_{k-1}^\circ)$ be smooth and $\mathbf{z} \in f_a(\Delta_{k-1}^\circ)$. If $\mathbf{z} \in f_a(\Delta_{k-1}^\circ)^*$, then $\mathbf{z}^\perp \in \mathbb{R}_+^k \cup \mathbb{R}_-^k$.

Proof. If the argument had not been true, then we would have a contraction to Theorem 27. $\qquad\square$

Theorems 27 and 28 present a necessary condition for a nondominated vector $\mathbf{z} \in f_a(\Delta_{k-1}^\circ)$. In other words, to be a nondominated vector in a smooth approximated Pareto front, the normal vector must belong to $\mathbb{R}_+^k \cup \mathbb{R}_-^k$. However, the converses of Theorems 27 and 28 do not hold. In other words, if some vector $\tilde{\mathbf{z}}$ dominates the vector $\mathbf{z}$, then $\mathbf{z}^\perp$ can belong to the negative orthant. We have presented a case in Figure 37, where for a vector $\mathbf{z} \in f_a(\Delta_1) \setminus f_a(\Delta_1)^*$ the normal vector is $(-1, -1)^{\mathrm{T}}$. This means that the converse of Theorem 27 does not hold. To show that the converse of Theorem 28 does not hold, see the case illustrated

in Figure 38, where the normal vector for a dominated vector $\mathbf{z} = (1, 1/4)^\mathrm{T}$ is $(-1, 0)^\mathrm{T} \in \mathbb{R}_+^2 \cup \mathbb{R}_-^2$. However, the examples discussed in this paragraph motivate us to study the Pareto dominance locally.
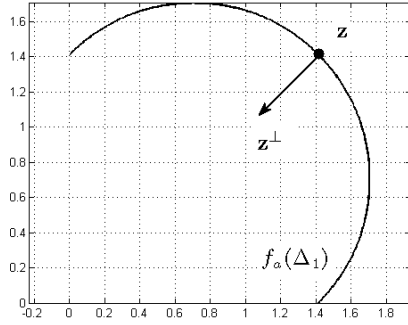
FIGURE 37 A dominated vector $\mathbf{z}$ so that the corresponding normal vector is $(-1, -1)^\mathrm{T}$.
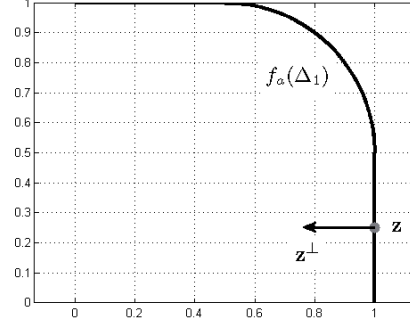
FIGURE 38 Condition $\mathbf{z}^\perp \in \mathbb{R}_+^k \cup \mathbb{R}_-^k$ does not guarantee Pareto dominance.

Normal vectors are based on local affine approximations [175] (tangent hyperplane $\mathcal{T}_{f_a(\mathbf{z})}$). Therefore, being a nondominated vector in the whole set $f_a(\Delta_{k-1}^\circ)$ cannot be concluded for sure. However, the local affine approximations and normal vectors through them establish a connection between vectors in $f_a(\Delta_{k-1}^\circ)$ and Pareto dominance locally as we see next.

Definition 29. Let $f_a(\Delta_{k-1}^\circ)$ be smooth. The vector $\mathbf{z} \in f_a(\Delta_{k-1}^\circ)$ is said to be *locally nondominated*, if there exists a radius $r > 0$ so that none of the vectors in $f_a(\Delta_{k-1}^\circ) \cap B(\mathbf{z}, r)$ dominate $\mathbf{z}$.

Typically, in the MCDM literature e.g. [123], local nondominance has been defined in the decision space and not in the objective space as we have done. Next, we show that locally nondominated vectors can be identified by the normal vectors, that is, we prove that the vector $\mathbf{z}^1$ in Figure 33 is locally nondominated.

Theorem 30. Let $f_a(\Delta_{k-1}^\circ)$ be smooth and $\mathbf{z} \in f_a(\Delta_{k-1}^\circ)$. If $\mathbf{z}^\perp$ belongs to $\mathrm{int}(\mathbb{R}_+^k \cup \mathbb{R}_-^k)$, then $\mathbf{z}$ is locally nondominated.

Proof. On the contrary, let us assume that for each $r_i = 1/i$, $i = 1, 2, \ldots$, we have a vector $\bar{\mathbf{z}}^i \in B(\mathbf{z}, r_i) \cap f_a(\Delta_{k-1}^\circ)$ so that it dominates the vector $\mathbf{z}$. We can assume that a sequence $(\bar{\mathbf{z}}^i)$ has been selected so that $\bar{\mathbf{z}}^i \neq \bar{\mathbf{z}}^j$ when $i \neq j$. From the smoothness of $f_a(\Delta_{k-1}^\circ)$, we obtain

$$\lim_{i \to \infty} \left\langle \frac{\bar{\mathbf{z}}^i - \mathbf{z}}{\|\bar{\mathbf{z}}^i - \mathbf{z}\|}, \frac{\mathbf{z}^\perp}{\|\mathbf{z}^\perp\|} \right\rangle = 0. \tag{23}$$

Let a subsequence $(\bar{\mathbf{z}}^{i_j})$ be selected so that for some index l and for all i_j, we have

$$\frac{|\bar{z}_l^{i_j} - z_l|}{\|\bar{\mathbf{z}}^{i_j} - \mathbf{z}\|} > \frac{1}{\sqrt{k}}.$$

(Note that such a sequence exists because otherwise for all i we would have $\left\|\frac{\bar{\mathbf{z}}^i - z}{\|\bar{\mathbf{z}}^i - z\|}\right\| < 1$, which cannot be true.) Now, let

$$0 < \varepsilon < \frac{1}{\sqrt{k}} \min_{i=1,\ldots,k} \frac{|\mathbf{z}_i^\perp|}{\|\mathbf{z}^\perp\|}.$$

Because (23) holds, then there exists an index $i_{\bar{j}}$ so that

$$\left| \left\langle \frac{\bar{\mathbf{z}}^{i_{\bar{j}}} - \mathbf{z}}{\left\|\bar{\mathbf{z}}^{i_{\bar{j}}} - \mathbf{z}\right\|}, \frac{\mathbf{z}^\perp}{\|\mathbf{z}^\perp\|} \right\rangle \right| < \varepsilon. \tag{24}$$

Now, we have two cases, which we have to study separately.

1. If $z_i^\perp < 0$ for all $i = 1, \ldots, k$, then some of the terms

$$\frac{(\bar{z}_m^{i_{\bar{j}}} - z_m)z_m^\perp}{\|\bar{\mathbf{z}}^{i_{\bar{j}}} - \mathbf{z}\|\,\|\mathbf{z}^\perp\|}, \quad m \neq l, \tag{25}$$

   in (24) must be negative because for the index l we have

$$\frac{(\bar{z}_l^{i_{\bar{j}}} - z_l)z_l^\perp}{\|\bar{\mathbf{z}}^{i_{\bar{j}}} - \mathbf{z}\|\,\|\mathbf{z}^\perp\|} = \left| \frac{(\bar{z}_l^{i_{\bar{j}}} - z_l)z_l^\perp}{\|\bar{\mathbf{z}}^{i_{\bar{j}}} - \mathbf{z}\|\,\|\mathbf{z}^\perp\|} \right| > \left| \frac{1}{\sqrt{k}} \frac{z_l^\perp}{\|\mathbf{z}^\perp\|} \right| > \varepsilon.$$

   Now, for some index $m \neq l$ only one of the numbers $\bar{z}_m^{i_{\bar{j}}} - z_m$ or $z_m^\perp$ can be positive and the other negative to obtain a negative term in (25). However, we know that for each $m \neq l$ the two above numbers are negative because the vector $\bar{\mathbf{z}}^i$ dominates $\mathbf{z}$ and the normal vector belongs to the negative orthant. In other words, we have met a problem, which means that this case cannot hold.

2. If $z_i^\perp > 0$ for all $i = 1, \ldots, k$, then some of the terms

$$\frac{(\bar{z}_m^{i_{\bar{j}}} - z_m)z_m^\perp}{\|\bar{\mathbf{z}}^{i_{\bar{j}}} - \mathbf{z}\|\,\|\mathbf{z}^\perp\|}, \quad m \neq l, \tag{26}$$

   in (24) must be positive, because for the index l we have

$$\frac{(\bar{z}_l^{i_{\bar{j}}} - z_l)z_l^\perp}{\|\bar{\mathbf{z}}^{i_{\bar{j}}} - \mathbf{z}\|\,\|\mathbf{z}^\perp\|} < -\frac{1}{\sqrt{k}} \frac{z_l^\perp}{\|\mathbf{z}^\perp\|} < -\varepsilon.$$

   However, to have a positive term in (26), the term $\bar{z}_m^{i_{\bar{j}}} - z_m$ should be positive for some m. This cannot be true because $\bar{\mathbf{z}}^{i_{\bar{j}}}$ dominates $\mathbf{z}$. This is a contradiction.

Because we have ended up with impossible situations in the above two cases we have a contradiction. Thus, the vector $\mathbf{z}$ is locally nondominated. $\qquad\square$

With minor changes in the proof of Theorem 30, we obtain the following.

Theorem 31. Let $f_a(\Delta_{k-1}^\circ)$ be smooth and $\mathbf{z} \in f_a(\Delta_{k-1}^\circ)$. If $\mathbf{z}^\perp$ belongs to $\mathrm{int}(\mathbb{R}_+^k \cup \mathbb{R}_-^k)$, then $\mathbf{z}$ does not dominate any of the vectors in $f_a(\Delta_{k-1}^\circ) \cap B(\mathbf{z}, r)$.

Proof. The proof is very similar to the proof of Theorem 30. □

Geometrically, the result obtained in Theorem 30 can be illustrated as follows. In Figure 39, there exist radii r_1 and r_2 so that vectors $\mathbf{z}^1, \mathbf{z}^2 \in f_a(\Delta_{k-1}^\circ)$, for which the normal vectors belong to the negative orthant, are locally nondominated. Note that the vector $\mathbf{z}^1$ dominates $\mathbf{z}^2$, which means that the vector $\mathbf{z}^2$ is only locally nondominated. However, Theorem 30 shows only an existence for radii r_1 and r_2 and, therefore, it cannot be said generally how small (large) the radii r_1 and r_2 must be (can be) for the local nondominance.
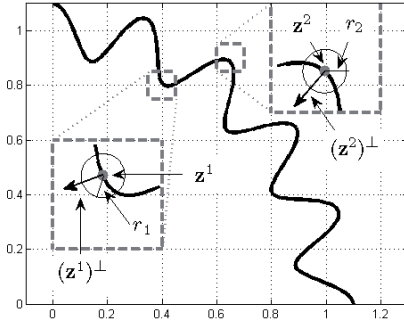
FIGURE 39 Two locally nondominated vectors on the approximated Pareto front.
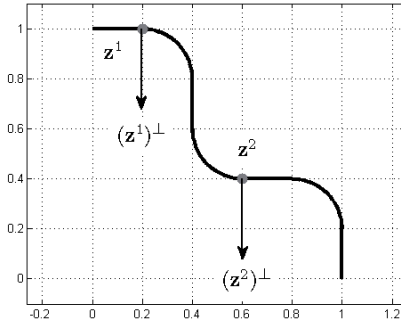
FIGURE 40 An example to illustrate that condition $\mathbf{z}^\perp \in \mathrm{int}(\mathbb{R}_+^k \cup \mathbb{R}_-^k)$ cannot be relaxed in Theorem 30.

The condition $\mathbf{z}^\perp \in \mathrm{int}(\mathbb{R}_+^k \cup \mathbb{R}_-^k)$ is essential in Theorem 30 and it cannot be relaxed. It is possible that for a vector $\mathbf{z}$ dominated by some other vector in $f_a(\Delta_{k-1}^\circ)$ the normal vector belongs to $\partial(\mathbb{R}_+^k \cup \mathbb{R}_-^k)$. In Figure 40, it does not matter how small a radius r_1 for the vector $\mathbf{z}^1$ is because the vector $(z_1^1 - r_1/2, 1)^\mathrm{T}$ dominates it. However, it is possible that for a locally nondominated vector $\mathbf{z}$, the normal vector belongs to $\partial(\mathbb{R}_+^k \cup \mathbb{R}_-^k)$. This is the case at $\mathbf{z}^2$ in Figure 40. In other words, the converse of Theorem 30 does not hold generally.

Vectors $\mathbf{z}^1$ and $\mathbf{z}^2$ depicted in Figure 40 can also be used to show the importance of having the condition $\mathbf{z}^\perp \in \mathrm{int}(\mathbb{R}_+^k \cup \mathbb{R}_-^k)$ in Theorem 31. In other words, no matter how small a radius r_2 is, the vector $\mathbf{z}^2$ dominates the vector $(z_1^2 + r_2/2, 0.4)^\mathrm{T}$. On the other hand, the vector $\mathbf{z}^1$ does not dominate any of the vectors in $f_a(\Delta_{k-1}^\circ) \cap B(\mathbf{z}^1, r_1)$ for a sufficiently small r_1 but $(\mathbf{z}^1)^\perp$ belongs to $\partial(\mathbb{R}_+^k \cup \mathbb{R}_-^k)$, which means that the converse of Theorem 31 does not generally hold.

In Theorems 30 and 31, the continuous differentiability of the function f_s has not been used explicitly. However, in the derivation of formula (21), it has been

assumed that the function f_s is continuously differentiable, see Appendix 1. The continuous differentiability of the function f_s means in practice that all normal vectors on a relative open neighborhood of $\mathbf{z} \in \Delta^\circ_{k-1}$, for which the normal vector $f_a(\mathbf{z})^\perp$ belongs to $\mathrm{int}(\mathbb{R}^k_+ \cup \mathbb{R}^k_-)$, belong to $\mathrm{int}(\mathbb{R}^k_+ \cup \mathbb{R}^k_-)$ as we show next.

Theorem 32. Let $f_a(\Delta^\circ_{k-1})$ be smooth and $\mathbf{z} \in \Delta^\circ_{k-1}$. If $f_a(\mathbf{z})^\perp \in \mathrm{int}(\mathbb{R}^k_+ \cup \mathbb{R}^k_-)$, then there exists a relatively open neighborhood W of $\mathbf{z}$ on Δ°_{k-1} so that for every $\mathbf{z}' \in W$ we have $f_a(\mathbf{z}')^\perp \in \mathrm{int}(\mathbb{R}^k_+ \cup \mathbb{R}^k_-)$, and $f_a(W) = f_a(W)^*$.

Proof. We first prove the existence of a relatively open neighborhood W. The continuous differentiability of function f_s means that the function defined in (21) used for calculating normal vectors is continuous. Then, for a sufficient small $\varepsilon > 0$ there exists $\delta > 0$ so that for all $\mathbf{z}' \in B(\mathbf{z}, \delta)$ we have for the corresponding normal vectors, $\|f_a(\mathbf{z}')^\perp - f_a(\mathbf{z})^\perp\| < \varepsilon$ meaning that $f_a(\mathbf{z}')^\perp \in \mathrm{int}(\mathbb{R}^k_+ \cup \mathbb{R}^k_-)$. In other words, there exists a relatively open neighborhood W of $\mathbf{z}$ on Δ°_{k-1}, which can be assumed, without loss of generality, to be also convex. Next, we prove that the sets $f_a(W)$ and $f_a(W)^*$ are the same.

  Inclusion "$\supset$" is obvious. To prove the other inclusion, let us suppose by a contradiction that we have $\mathbf{z}, \bar{\mathbf{z}} \in W$ so that $f_a(\bar{\mathbf{z}})$ dominates $f_a(\mathbf{z})$. Let curve $f_c : [0, 1] \to \mathbb{R}^k$ be defined as follows: $\mathbf{z}^t := f_c(t) = f_a(t\mathbf{z} + (1 - t)\bar{\mathbf{z}})$. Let $\mathcal{L}$ be a set consisting of vectors $\{\hat{\mathbf{z}} \in \mathbb{R}^k : \hat{\mathbf{z}} = \lambda(f_a(\bar{\mathbf{z}}) - f_a(\mathbf{z})) + f_a(\mathbf{z}), \lambda \in \mathbb{R}\}$. Now, let the function $d_{\mathcal{L}} : [0, 1] \to \mathbb{R}$ be defined as follows:

$$d_{\mathcal{L}}(t) = \inf_{\mathbf{v} \in \mathcal{L}} \|\mathbf{v} - f_c(t)\|.$$

Note that the function $d_{\mathcal{L}}$ is continuous. Then for all $t \in [0, 1]$, there exists a vector $\mathbf{v}^t \in \mathcal{L}$ so that $\|\mathbf{v}^t - f_c(t)\| = d_{\mathcal{L}}(t)$, because $[0, 1]$ is a compact set. Similarly, there exists $t^* \in [0, 1]$ so that

$$d_{\mathcal{L}}(t^*) = \sup_{t \in [0,1]} d_{\mathcal{L}}(t).$$

Now, because $f_a(\Delta^\circ_{k-1})$ is smooth, then so is f_c ([175, Theorems 5-1, 5-2 and Problem 2-9]), meaning that

$$\left\langle (\mathbf{z}^{t^*})^\perp, f_a(\mathbf{z}) - f_a(\bar{\mathbf{z}}) \right\rangle = 0. \tag{27}$$

Because $f_a(\bar{\mathbf{z}})$ dominates $f_a(\mathbf{z})$, then $f_a(\mathbf{z})_i - f_a(\bar{\mathbf{z}})_i \geq 0$ for all $i = 1, \ldots, k$, and for some i the inequality is strict. Without loss of generality, we can assume that $(z^{t^*}_i)^\perp < 0$ for all $i = 1, \ldots, k$. Then

$$\left\langle (\mathbf{z}^{t^*})^\perp, f_a(\mathbf{z}) - f_a(\bar{\mathbf{z}}) \right\rangle = \sum_{i=1}^{k} \underbrace{(z^{t^*})^\perp_i}_{<0} \underbrace{(f_a(\mathbf{z})_i - f_a(\bar{\mathbf{z}})_i)}_{\geq 0} < 0,$$

which contradicts with (27). In other words, there do not exist vectors $\bar{\mathbf{z}}$ and $\mathbf{z}$ in W so that $f_a(\bar{\mathbf{z}})$ dominates $f_a(\mathbf{z})$. Thus $f_a(W) \subset f_a(W)^*$. $\square$

Now, we immediately obtain the following corollary.

Corollary 33. Let $f_a(\Delta^\circ_{k-1})$ be smooth. If for all $\mathbf{z} \in \Delta^\circ_{k-1}$ normal vectors $f_a(\mathbf{z})^\perp$ belong to $\mathrm{int}(\mathbb{R}^k_+ \cup \mathbb{R}^k_-)$, then $f_a(\Delta^\circ_{k-1}) = f_a(\Delta^\circ_{k-1})^*$.

Proof. The result follows immediately from Theorem 32. □

Before we demonstrate the significance of Corollary 33 in practice, differences between Theorems 30, 31 and 32 must be clarified because a careful reader may have noticed that Theorems 30 and 31 are corollaries of Theorem 32. However, Theorems 30 and 31 do not necessarily require the function f_s to be continuously differentiable. The proofs of Theorems 30 and 31 require that a differentiable function g, as in (22), at $f_a(\mathbf{z})$ exists. Note that such a function g may exist for a differentiable function f_s. This means that the theorems can be proved by only using the tangent hyperplane and the normal vector, as it has been done in this thesis. The continuous differentiability of function f_s guarantees that a function g exists at any vector $f_a(\mathbf{z})$ in the relative interior of $f_a(\Delta_{k-1})$. In other words, Theorems 30 and 31 are valid for fewer assumptions set on the function f_s. On the other hand, relaxing the continuous differentiability of the function f_s in Theorem 32 to be only differentiable allows us to create examples in which for every radius $r > 0$ a pair of vectors can be found so that one dominates the other. Such a case has been sketched in Appendix 2. Another difference between the theorems can be pointed to the proof of Theorem 32. Note that the theorem could be proved by using Theorem 30. However, in this thesis, we have proved Theorem 32 without using Theorem 30.

Corollary 33 is essential in analyzing whether or not the whole set $f_a(\Delta^\circ_{k-1})$ is nondominated. Let us illustrate this as follows.

Example 34. Let us consider an arc of an origin-centered circle with radius 1 in $\mathbb{R}^2_+$ as an approximated Pareto front. Then

$$f_s(\mathbf{z}) = \frac{1}{\sqrt{z_1^2 + z_2^2}}, \quad \frac{\partial f_s}{\partial z_1}(\mathbf{z}) = -\frac{z_1}{(z_1^2 + z_2^2)^{3/2}}, \quad \frac{\partial f_s}{\partial z_2}(\mathbf{z}) = -\frac{z_2}{(z_1^2 + z_2^2)^{3/2}},$$

meaning that the function f_s is continuously differentiable. Now, at any $\mathbf{z} = (z_1, z_2) \in \Delta^\circ_1$ we have

$$f_a(\mathbf{z})^\perp = -\begin{bmatrix} \frac{z_1}{(z_1^2+z_2^2)^{3/2}} & \frac{z_2}{(z_1^2+z_2^2)^{3/2}} \end{bmatrix} \begin{bmatrix} \frac{1-z_1}{f_s(\mathbf{z})} & -\frac{z_1}{f_s(\mathbf{z})} \\ -\frac{z_2}{f_s(\mathbf{z})} & \frac{1-z_2}{f_s(\mathbf{z})} \end{bmatrix} - \mathbf{1}^\mathrm{T}$$

$$= -\begin{bmatrix} \frac{z_1-z_1^2-z_2^2}{z_1^2+z_2^2} & \frac{z_2-z_2^2-z_1^2}{z_1^2+z_2^2} \end{bmatrix} - \mathbf{1}^\mathrm{T} = -\begin{bmatrix} \frac{z_1}{z_1^2+z_2^2} & \frac{z_2}{z_1^2+z_2^2} \end{bmatrix} < \mathbf{0}^\mathrm{T}.$$

In other words, for each $\mathbf{z} \in \Delta^\circ_1$ the normal vector belongs to $\mathrm{int}(\mathbb{R}^2_+ \cup \mathbb{R}^2_-)$. Thus, $f_a(\Delta^\circ_1) = f_a(\Delta^\circ_1)^*$.

Example 34 can be easily modified to show that a set consisting of vectors $\{\mathbf{z} \in \mathbb{R}^k : \|\mathbf{z}\| = 1, z_i > 0 \text{ for all } i = 1, \dots, k\}$ is nondominated. However, the set considered in Example 34 can be proved easily to be nondominated, without using normal vectors. Next, we present a nontrivial example.

Example 35.

Let

$$f_s(z_1, z_2) = 1 + \frac{1}{6\pi}\sin(6\pi z_1)$$

and $\mathbf{ideal}^{Z_m} = \mathbf{0}$.
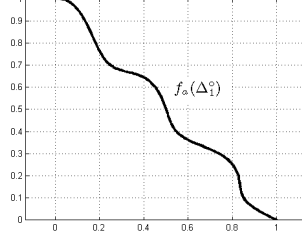
Then $f_a(\Delta_1^\circ)$ is nondominated because for an arbitrary $\mathbf{z} = (z_1, z_2) \in \Delta_1^\circ$, we have

$$f_a(\mathbf{z})^\perp = \begin{bmatrix} \cos(6\pi z_1) & 0 \end{bmatrix} \begin{bmatrix} \frac{1-z_1}{f_s(\mathbf{z})} & -\frac{z_1}{f_s(\mathbf{z})} \\ -\frac{z_2}{f_s(\mathbf{z})} & \frac{1-z_2}{f_s(\mathbf{z})} \end{bmatrix} - \mathbf{1}^T$$

$$= \begin{bmatrix} \frac{\cos(6\pi z_1)(1-z_1)}{1+\frac{1}{6\pi}\sin(6\pi z_1)} & -\frac{\cos(6\pi z_1)(z_1)}{1+\frac{1}{6\pi}\sin(6\pi z_1)} \end{bmatrix} - \mathbf{1}^T \leq \begin{bmatrix} \frac{1-z_1}{1+z_1} & -\frac{z_1}{1+z_1} \end{bmatrix} - \mathbf{1}^T < \mathbf{0}^T.$$

Example 35 demonstrates how Corollary 33 characterizes an approximated Pareto front to be nondominated. Note that the approximated Pareto front can be visualized very easily because the objective space is only two-dimensional. However, the front visualized may inform that the relative interior of the approximated Pareto front would be nondominated. As we know, visualization does not prove anything, but Corollary 33 confirms the set seen in the visualization as nondominated. In other words, what we see is true! Note that to show that a set containing infinitely many vectors is nondominated is not a trivial problem to be solved by using some other methods than normal vectors.

To conclude this subsection, normal vectors on the relative interior of the approximated Pareto front have been studied. It has been proved that Pareto dominance is highly related to the normal vectors. As a main result of this section, we have proved that if normal vectors for all vectors on the relative interior of the approximated Pareto front belong to the interior of the union of negative and positive orthants then the relative interior of the approximated Pareto front is nondominated.

4.2 Normal Vectors Before and During the Solution Process

Normal vectors have been pointed out to be useful in the investigation whether a vector in the approximation is nondominated i.e. an approximated Pareto optimal objective vector. This is an essential feature for the IHBE and HBE methods because the approximation can now be validated, as we show in this section. In addition, approximated trade-off rates at vectors on the approximation can be calculated through normal vectors. In other words, the normal vectors can be used in step (3) of IHBE to validate the approximation, and in step (4) of HBE to present additional numerical information to the decision maker.

The implementation of the calculation of normal vectors on the approxi-

mation can be done efficiently because we have the derived formula (21). The essential attribute of (21) is its validity for nonconvex MOO problems. The actual methods to be used in the steps of IHBE are introduced as follows. In Subsection 4.2.1, we discuss the validation based on normal vectors. Then, in Subsection 4.2.2, approximated trade-off rates are calculated with the help of the normal vectors.

4.2.1 Validating the Approximation by Normal Vectors

Before the decision maker starts the actual solution process with the IHBE method, the approximation must be validated by the analyst. In this way, it is possible to avoid situations where some of the vectors in the approximation may dominate Pareto optimal objective vectors in Z_m. Such situation may occur if, for example, the function f_s has been selected to be a polynomial. As known, a polynomial may occasionally oscillate [94] which means that the approximated Pareto front resemble the one illustrated in Figure 41 (black curve).
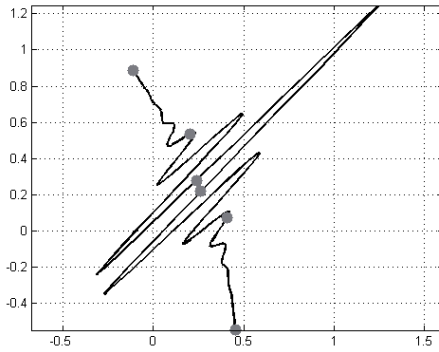
FIGURE 41   In validating the approximated Pareto front, normal vectors can be used at vectors in Z_6 in order to avoid these kinds of approximations.

If we calculate normal vectors at $\mathbf{z} \in Z_6$ of Figure 41 then most of the normal vectors at vectors in Z_6 do not belong to $\mathbb{R}_+^k \cup \mathbb{R}_-^k$. According to Theorem 27 such vectors are dominated in the approximation. This indicates that the approximation is not valid because some of the known Pareto optimal objective vectors are dominated (for a valid approximation, vectors in Z_m should be nondominated). In practice, this means that the function f_s in f_a should be replaced by some other surrogate function in order to remove vectors in the approximation that dominate Pareto optimal objective vectors in Z_m. To clarify, the normal vectors do not characterize dominated vectors caused by the oscillation of the function f_s, but they can indicate such behavior in the approximated Pareto front. Since normal vectors can be used when investigating whether vectors in Z_m are nondominated in the approximation, a validation method for the approximation is, for example, the following:

Method 36.

(1) Calculate normal vectors at vectors in Z_m.
(2) Validate the approximation as follows:

    (2.1) If every normal vector belongs to $\mathbb{R}^k_+ \cup \mathbb{R}^k_-$ then the approximation is valid.

    (2.2) If some of the normal vectors do not belong to $\mathbb{R}^k_+ \cup \mathbb{R}^k_-$ then the approximation is not valid and present the dominated and nondominated vectors if needed.

Method 36 is rather simple, and the validation in it is based on Theorems 27 and 28. However, the method only validates vectors in Z_m. For a larger set of vectors, we propose the following:

Method 37.

(1) Generate sample vectors on Δ°_{k-1} and calculate normal vectors at the corresponding vectors in the approximation.
(2) Validate the approximation as follows:

    (2.1) If every normal vector belongs to $\mathbb{R}^k_+ \cup \mathbb{R}^k_-$, then the approximation may be a nondominated set.

    (2.2) If some of the normal vectors do not belong to $\mathbb{R}^k_+ \cup \mathbb{R}^k_-$, then the approximation contains dominated vectors and present them, if needed.

Method 37 is clearly a computational counterpart to Corollary 33. In other words, Method 37 validates computationally whether or not the set $f_a(\Delta^\circ_{k-1})$ is nondominated. However, the validation is based on sample vectors on the simplex, and, to be sure of nondominance, a large set of sample vectors must be generated. On the other hand, this is not a problem because generating vectors on the simplex is not a problem (see, $\mathcal{AM}f4$) and carrying out the calculation of normal vectors can be considered a computationally efficient operation. Note that if the partial derivatives are not available or are computationally costly, then they can be approximated efficiently by finite differences because the function f_s has been assumed to be computationally efficient.

    When it happens that some of the sample vectors in the approximation are dominated (detected by Method 37), the interpretation of these vectors is discussed in Chapter 6. However, if Method 36 indicates dominated vectors in Z_m, then there are several possible actions that can be performed to fix the problem. For example:

1. Replace function f_s by some other surrogate function.
2. Use function f_s in the approximation, but if the decision maker is investigating a region close to a dominated vector ($\in Z_m$), then inform the decision maker about the dominated vector and invalid behavior of the approximation related to the Pareto dominance between the vectors in the approximation.

3. Add some dummy vectors to the fitting data of the function f_s so that normal vectors at Z_m will belong to $\mathbb{R}_+^k \cup \mathbb{R}_-^k$ after having fitted the function f_s again to the data.

To make it clear, the validations discussed in this subsection are meant to be done by the analyst. Methods proposed in this subsection are meant to be used side by side in validating the approximation. In later chapters, we propose other methods for the validation purposes. Therefore, the information returned by the methods must be proportioned to the feature in the approximation the method is validating. For example, if Method 37 indicates that some of the vectors are dominated then this does not mean that the function f_s must be replaced by some other because the nonconvexity of the MOO problem may have forced the approximation to contain dominated vectors, as discussed in Section 2.2. Later, in Chapter 6, we discuss how these dominated vectors can be interpreted in the solution process of the MOO problem.

4.2.2 Approximated Trade-off Rates and Normal Vectors

In step (3) of HBE, sample vectors on the approximated Pareto front in relation to the decision maker's preferences are calculated. At the end of the solution process of the MOO problem, when an interesting region containing a possible compromise solution has been located, the decision maker may be interested in detailed information for the sample vectors. For example, he or she may have found certain interesting vectors, but some fine-tuning of the objective function values is required. Here, the fine-tuning means that by changing slightly some objective function values, a better (approximated) Pareto optimal objective vector may be found. In other words, the decision maker may be interested in knowing how much the value of objective function f_i must be approximately impaired if the value of the objective function f_j is to be improved by one unit while the other objective function values remain unchanged. Alternatively, how much the value of the objective function f_i will approximately improve if the value of the objective function f_j is impaired by one unit and the other objective function values remain unchanged[1]. In the literature, this is known as calculating a (partial) trade-off rate [49, 123, 131]. In other words, the trade-off rate λ_{ij} approximates the impairment/improvement between the values of the objective functions f_i and f_j, and, typically in the literature, it has been defined at a feasible decision vector $\mathbf{x}$ with a feasible direction $\mathbf{d} \in \mathbb{R}^n$ (there exists an $\alpha' > 0$ so that for all $\alpha < \alpha'$ the vector $\mathbf{x} + \alpha\mathbf{d}$ belongs to S) as follows:

$$\lambda_{ij} := \lim_{\alpha \to 0^+} \frac{f_i(\mathbf{x} + \alpha\mathbf{d}) - f_i(\mathbf{x})}{f_j(\mathbf{x} + \alpha\mathbf{d}) - f_j(\mathbf{x})}. \tag{28}$$

---

[1] Occasionally, improving/impairing the value of the objective function f_i by one unit may be too much. The impairment/improvement for the value of the objective function f_j can also be calculated for any $0 < \varepsilon < 1$ improvement/impairment for the value of the objective function f_j.

Definition (28) for the trade-off rate is as general as possible because the trade-off rate can be calculated for any feasible decision vector and feasible direction. Figure 42 illustrates geometrically the trade-off rate, and it can be said that the trade-off rate defines a tangent line (a tangent space) at $f(\mathbf{x})$ to the values of the objective functions f_i and f_j at the vector $\mathbf{x}$ to the direction $\mathbf{d}$.
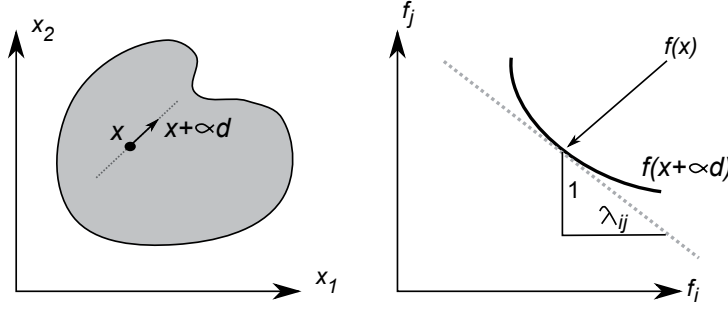
FIGURE 42   A trade-off rate λ_{ij} is the slope between the values of the objective functions f_i and f_j at $\mathbf{x} \in S$ to a feasible direction $\mathbf{d}$.

In order to connect the trade-off rate λ_{ij} to the explanation given above for the impairments/improvements of objective function values at vectors in the Pareto front that the decision maker is interested in, a specific direction $\mathbf{d}$ in the decision space must be determined. For this purpose, the trade-off rate must be connected to the Pareto front in a special way. Note that the trade-off rate λ_{ij} may not always represent the impairment between the coordinate values i and j of the objective vectors in the Pareto front no matter which direction $\mathbf{d}$ has been selected. That is, the tangent line at $f(\mathbf{x})$ is not necessarily a tangent line for the Pareto front. This may happen, especially when the MOO problem is nonconvex. This is actually the case at $\mathbf{x}$ in Figure 43, where $S = [0,1]$ and subsets A, B, C have been mapped to the objective space by a continuously differentiable function $f : [0,1] \rightarrow \mathbb{R}^2$ so that the objective vectors in the set $f(A)$ are dominated by the vectors in the sets $f(B)$ and $f(C)$, which are subsets of the Pareto front $\mathrm{PF}(f, [0,1])$ in this case. In addition, $f(x) \in \mathrm{PF}(f, [0,1])$. Note that the curves can be connected together by the function f, but for the simplicity we have not illustrated the other parts of the image in the figure.

We say that the trade-off rate λ_{ij} at $\mathbf{x}$ to a direction $\mathbf{d}$ is in *relation to the Pareto front* if there exists a continuously differentiable injective function $\bar{c} : [-1,1] \rightarrow \mathbb{R}^k$ so that $\bar{c}(0) = f(\mathbf{x})$, $\bar{c}([-1,1]) \subset \mathrm{PF}(f, S)$, $\bar{c}(t)_l = f_l(\mathbf{x})$ for all $t \in [-1,1]$, $l = 1, \ldots, k, l \neq i, j$ and λ_{ij} is equal to the ratio of derivatives

$$\frac{\bar{c}_i'(0)}{\bar{c}_j'(0)}.$$

In other words, if a trade-off rate is in relation to the Pareto front, then the slopes (i.e. the rates) are the same, as it can be seen in Figure 44. Note that the trade-off rate λ_{ij} is not in relation to the Pareto front in the case presented in Figure 43

Decision space   Objective space
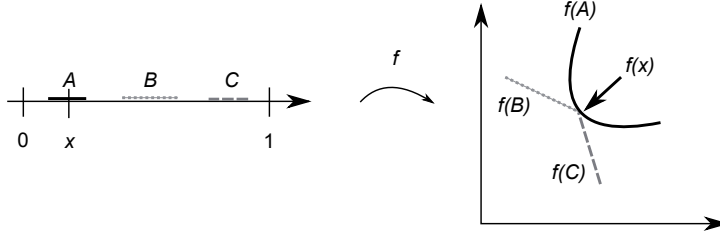
FIGURE 43   For nonconvex MOO problems, the trade-off rate λ_{ij} does not necessarily express the impairments/improvements for the vectors in the Pareto front.

because the tangent hyperplane at $f(\mathbf{x})$ is not a tangent hyperplane at $f(\mathbf{x})$ for the Pareto front.
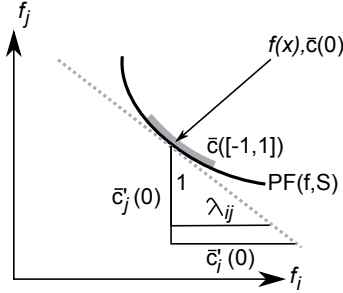
FIGURE 44   The trade-off rate λ_{ij} at $\mathbf{x}$ to some direction $\mathbf{d}$ is in relation to the Pareto front, if a hyperplane defined by the trade-off rate is a tangent hyperplane for the Pareto front at $f(\mathbf{x})$.

Now, if a direction $\mathbf{d}$ is based on derivatives of a continuously differentiable function $c : [-1, 1] \to \mathbb{R}^n$ for which $c(0) = \mathbf{x}$, $c'(0) \neq \mathbf{0}$, the function $f \circ c : [-1, 1] \to \mathbb{R}^k$ is injective and continuously differentiable, $f \circ c([-1, 1]) \subset$ PF(f, S), $f \circ c(t)_l = f_l(\mathbf{x})$ for all $t \in [-1, 1]$ and $l = 1, \ldots, k, l \neq i, j$, as follows

$$\mathbf{d} = (c'_1(0), \ldots, c'_n(0))^{\mathrm{T}},$$

then the trade-off rate λ_{ij} is in relation to the Pareto front. We have illustrated the above in Figure 45. However, constructing such a smooth curve c can be challenging in practice. Therefore, in the literature, trade-off rates considered in our terminology in a relation to the Pareto front are calculated in a different way.

According to [131], it has been proved under certain regularity and second order conditions in [30] that if the KKT coefficients [106] of a minimization problem (also known as the ε-constraint method [30, 59])

$$\begin{aligned}
\min \quad & f_i(\mathbf{x}) \\
\text{s.t.} \quad & f_j(\mathbf{x}) \leq \varepsilon_j \quad \text{for all } j = 1, \ldots, k, \ j \neq i, \\
& \mathbf{x} \in S,
\end{aligned}$$
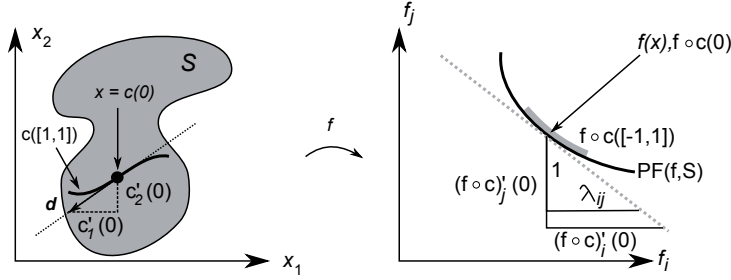
FIGURE 45 A Pareto optimal decision vector **x** and a direction **d** so that the trade-off rate is in a relation to the Pareto front.

where ε_j are upper bounds for the objectives, are positive, then by multiplying by -1 the KKT coefficient of the jth constraint is equal to the trade-off rate λ_{ij} in relation to the Pareto front. It can also be proved that the trade-off rates $\lambda_{i1}, \dots, \lambda_{i(i-1)}, \lambda_{i(i+1)}, \dots, \lambda_{ik}$ in relation to the Pareto front define a tangent space at a vector $f(\mathbf{x})$ in a smooth Pareto front, see e.g. [131]. This can be seen, for example, in Figure 44, where the gray dashed line is the shifted tangent space. The tangent space shifted to $f(\mathbf{x})$ is a tangent hyperplane similar to the tangent hyperplane $\mathcal{T}_{f_a(\mathbf{z})}$ at $f_a(\mathbf{z}) \approx f(\mathbf{x})$ for some previously unknown Pareto optimal decision vector **x**. Therefore, the tangent space $\mathcal{T}_{\mathbf{z}}$ corresponding to the hyperplane $\mathcal{T}_{f_a(\mathbf{z})}$ can be considered as an approximated tangent space for the Pareto front at the vector $f(\mathbf{x})$.

Because the trade-off rates define the tangent space at a vector in the Pareto front and the tangent space $\mathcal{T}_{\mathbf{z}}$ can be considered to approximate it, then the normal vector $f_a(\mathbf{z})^{\perp}$ to the tangent space $\mathcal{T}_{\mathbf{z}}$ can be used for calculating approximated trade-off rates in relation to the Pareto front at $f_a(\mathbf{z})$. Calculating trade-off rates in relation to the Pareto front through normal vectors is not a new idea, because in the MCDM literature such studies are presented, for example, in [70, 128, 163]. However, the calculations are based on KKT coefficients, which can be obtained by solving a system of nonlinear equations. This means that the decision maker may have to deal with long waiting times due to the computationally expensive calculations of objective functions in solving the system of equations. In the HBE method, approximated trade-off rates in relation to the Pareto front can be calculated without solving the KKT coefficients, as follows.

An *approximated trade-off rate* in relation to the Pareto front at $f_a(\mathbf{z})$ in the approximated Pareto front can be calculated by using the normal vector $f_a(\mathbf{z})^{\perp}$. Here, the approximated trade-off rate is denoted by a function value of the function $f_{ij}^{tor} : \Delta_{k-1}^{\circ} \to \mathbb{R}$,

$$f_{ij}^{tor}(\mathbf{z}) = -\frac{f_a(\mathbf{z})_j^{\perp}}{f_a(\mathbf{z})_i^{\perp}}, \tag{29}$$

which approximates a trade-off rate λ_{ij} in relation to the Pareto front at some unknown Pareto optimal decision vector **x** for which $f(\mathbf{x}) \approx f_a(\mathbf{z})$. Note that in this chapter the approximated Pareto front has been assumed to be smooth,

and, consequently normal vectors on it can be calculated. In order to see that the function $f_{ij}^{tor}(\mathbf{z})$ calculates approximated trade-off rates in relation to the Pareto front, we have depicted the trade-off rate $-\lambda_{ij}$ by r in Figure 46, and r can be solved as follows:
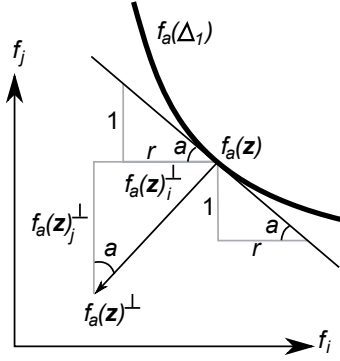
FIGURE 46   The normal vector $f_a(\mathbf{z})^\perp$ can be used in calculating approximated trade-off rates, which are in relation to the approximated Pareto front.

For the angle a in Figure 46, we have

$$\tan(a) = \frac{1}{r} \quad \text{and} \quad \tan(a) = \frac{f_a(\mathbf{z})_i^\perp}{f_a(\mathbf{z})_j^\perp}.$$

In other words,

$$r = \frac{f_a(\mathbf{z})_j^\perp}{f_a(\mathbf{z})_i^\perp}$$

and, thus, $\lambda_{ij} \approx f_{ij}^{tor}(\mathbf{z})$.

Now, a method for approximated trade-off rates to be used in the steps (3-4) of HBE, where sample vectors on the approximation are calculated in relation to the decision maker's preferences, can be formulated as follows:

Method 38.

(1) Calculate approximated trade-off rates at vectors $f_p(\mathbf{z}^{h_i})$ by the formula given in (29), where the vector $f_p(\mathbf{z}^{h_i})$ is the vector in the approximation corresponding to a sample vector $\mathbf{z}^{h_i}$ satisfying the aspiration and reservation levels given by the decision maker.

(2) If the decision maker wants to see the approximated trade-off rates calculated, then present the rates.

Method 38 is rather simple and computationally efficient for calculating the approximated trade-off rates at the vectors in the approximated Pareto front. However, to clarify the first step, the function f_{ij}^t must be used at vectors $\mathbf{z}^{sh_i}$, which are the vectors in Δ_{k-1}° corresponding to $\mathbf{z}^{h_i}$ as defined in Subsection 3.2.2. Note that the function f_{ij}^{tor} has been defined on Δ_{k-1}°.

The approximated impairment/improvement in (29) is based on the local affine approximation $\mathcal{T}_{f_a(\mathbf{z})}$ to the approximated Pareto front. This means that the impairment/improvement is a very rough estimation as is the actual trade-off rate itself. However, if the decision maker is willing to improve/impair only $0 < \varepsilon < 1$ instead of one unit, then the value $\varepsilon f_{ij}^{tor}(\mathbf{z})$ is valid for the approximated impairment/improvement.

The significance of formula (29) compared with the traditional ways of calculating trade-off rates is explained by the use of formula (21). Typically, in the literature, see e.g. [49, 131, 163, 193], trade-off rates are calculated through KKT coefficients [106], which are solutions of a system of nonlinear equations. However, to solve the equations, in the worst case a time-consuming step must be taken. In addition, the MOO problem must be convex in order to use the coefficients. In other words, calculating approximated trade-off rates without the KKT coefficients by formula (21) is significant in three ways:

1. there is no need to solve a system of nonlinear equations,
2. the MOO problem can be nonconvex and computationally expensive and
3. there is no need to introduce a direction for the trade-off rate, which is in a relation to the approximate Pareto front. (Note that the Pareto front can be nonsmooth, even though the objective functions are continuously differentiable e.g. for linear problems. In addition, there does not necessarily exist a direction for a vector $\mathbf{x}$ so that a trade-off rate will be in a relation to the Pareto front, as in the case presented in Figure 43.)

In practice, the above means that the decision maker does not have to wait in order to see the approximated trade-off rates in the HBE method. In this way, the decision maker can make local fine-tunings[2] around interesting approximated Pareto optimal objective vectors without time-consuming calculations.

4.3 Summary of Chapter 4

The simplicity of the approximation introduced in Chapter 2 has been a key element in implementing the approximation to be used in the HBE and IHBE methods. In this chapter, the simplicity has provided an easy access to calculate normal vectors on the approximated Pareto front. Therefore, calculations of (approximated) trade-off rates can be achieved very fast for supporting the decision maker in psychological convergence. In addition, normal vectors can be used in characterizing (non)dominated vectors, i.e. that normal vectors can be used in validating the approximation. Next, we summarize the main results of Chapter 4.

Theorem

---

[2] The statement about the fine-tuning can be supported by the studies and notes presented in [25].

27 The main point in Theorem 27 is to prove that if the normal vector $f_a(\mathbf{z})^\perp$ at $f_a(\mathbf{z}) \in f_a(\Delta_{k-1}^\circ)$ does not belong to $\mathbb{R}_+^k \cup \mathbb{R}_-^k$, then the vector $f_a(\mathbf{z})$ is dominated in $f_a(\Delta_{k-1}^\circ)$. In other words, the normal vector describes when a vector in the approximated Pareto front is dominated, and this is essential in validating the approximation.

28 If we assume that $f_a(\mathbf{z}) \in f_a(\Delta_{k-1}^\circ) \cap f_a(\Delta_{k-1})^*$, then the normal vector $f_a(\mathbf{z})^\perp$ belongs to $\mathbb{R}_+^k \cup \mathbb{R}_-^k$. This is the case proved in Theorem 28. In other words, for an approximated Pareto optimal objective vector, the normal vector must belong to the negative or positive orthant.

30 In this theorem, it is shown that a vector $\mathbf{z} \in f_a(\Delta_{k-1}^\circ)$, for which the normal vector belongs to $\mathrm{int}(\mathbb{R}_+^k \cup \mathbb{R}_-^k)$, is locally nondominated. Locally means in this case that there exists a radius $r_\mathbf{z} > 0$ so that none of the vectors in $B(\mathbf{z}, r_\mathbf{z}) \cap f_a(\Delta_{k-1}^\circ)$ dominate $\mathbf{z}$.

31 If for a vector $\mathbf{z} \in f_a(\Delta_{k-1}^\circ)$ the normal vector belongs to $\mathrm{int}(\mathbb{R}_+^k \cup \mathbb{R}_-^k)$ then there exists a radius $r_\mathbf{z} > 0$ so that vector $\mathbf{z}$ does not dominate any of the vectors in $B(\mathbf{z}, r_\mathbf{z}) \cap f_a(\Delta_{k-1}^\circ)$. This was proved in this theorem.

32 This theorem can be used to show that a set on the approximation is nondominated. In other words, if every normal vector on a relatively open convex neighborhood of $\mathbf{z} \in \Delta_{k-1}^\circ$ belongs to $\mathrm{int}(\mathbb{R}_+^k \cup \mathbb{R}_-^k)$ then $f_a(W) \subset f_a(W)^*$.

It is obvious that the mathematical model introduced in Chapter 2 is flexible because of its simplicity. As seen in this chapter, its simplicity provides an easy access to the calculation of normal vectors as a function on Δ_{k-1}° (formula (21)). Therefore, it is worthwhile to name several features for the approximation, as it has been done already on page 40.

$\mathcal{AM}$f6   In feature $\mathcal{AM}$f4, we discussed the simplex Δ_{k-1}. It is the domain space for the function f_a. Now, its relative interior serves as a domain space for the normal vectors. In other words, when calculating a vector on the approximation and normal vector at it, both can be calculated at the same vector $\mathbf{z} \in \Delta_{k-1}^\circ$.

$\mathcal{AM}$f7   Normal vectors provide information about whether or not a vector in the approximation is nondominated. However, to calculate normal vectors, we have to assume that the function f_s in f_a is continuously differentiable, and this is not a problem because such surrogate functions exist, e.g. RBFs. Note that it has been typically assumed in the literature of MCDM that the objective functions f_i are continuously differentiable i.e. the partial derivative functions of them must be available. This means that, in this thesis, the assumptions about the problem are on the same level as in the literature. However, the next feature clarifies the main significant difference when compared with the literature.

$\mathcal{AM}$f8   The most important feature of normal vectors introduced in this thesis is their simple calculation. We do not have to solve for example the KKT coefficients, as it has been done in the MCDM literature. In other words, if

the partial derivative functions of f_s have been calculated, then calculations of normal vectors can be performed computationally efficiently. And best of all, the normal vectors can be calculated even if the MOO problem is nonconvex.

$\mathcal{AM}$f9   Because the calculation of normal vectors is based on formula (21), the information concluded or derived from the normal vectors related to the Pareto dominance can be easily applied to the solution process of the MOO problem. This is a clear advantage because there is no need to use external solvers for calculating the KKT coefficients.

$\mathcal{AM}$f10   In Section 3.3, we introduced approximated trade-off rates as additional numerical information to be presented to the decision maker in step (4) of HBE. The rates can be obtained computationally efficiently through the normal vectors. In addition, normal vectors have been used in simple validation methods for the approximation. In other words, normal vectors can be used in several different ways in the HBE and IHBE methods strengthening the mathematical model by validation methods. This means that goals $\mathcal{G}$6 and $\mathcal{G}$9 set on page 20 for the approximation are considered in this thesis.

5 NONDOMINANCE AND FUNCTION f_s

In Chapter 4, we have studied normal vectors in characterizing whether vectors in the approximated Pareto front are nondominated or dominated. The characterization is naturally used in validating the approximation before the decision maker starts to use it in the solution process of the MOO problem. However, this characterization must be improved, and we demonstrate the reason for the improvement as follows. In Figure 47, vectors in the set Z_9 have been depicted as gray dots, and we have selected two vectors in the approximated Pareto front $\mathbf{z}^1$ and $\mathbf{z}^2$ (black dots) for which the normal vectors belong to the negative orthant, meaning that they are locally nondominated (Theorem 30). However, there are vectors in Z_9 which dominate the vector $\mathbf{z}^1$. In addition, the vector $\mathbf{z}^2$ dominates at least one of the Pareto optimal objective vectors in Z_9. In other words, vectors like $\mathbf{z}^2$ question the approximation to be used in the solution process of the MOO problem. Thus, we have geometrical evidence that improvements to the characterization of dominated vectors based on normal vectors are needed in validating the approximation.

In this chapter, we introduce functions $f_l, f_u : \Delta_{k-1}^{\circ} \to \mathbb{R}$ as a lower function and an upper function, respectively, for function f_s. That is, if we select a vector $\mathbf{z} \in \Delta_{k-1}^{\circ}$, then we can calculate how much we can multiply the vector $\mathbf{z}$ by a scalar value $f_s(\mathbf{z})$ without dominating any of the vectors in Z_m by $f_a(\mathbf{z})$. In addition, we can calculate how much the vector $\mathbf{z}$ can be multiplied without dominating $f_a(\mathbf{z})$ by any of the vectors in Z_m. In other words, we prove that, if a scalar value $f_s(\mathbf{z})$ is too small ($f_s(\mathbf{z}) < f_l(\mathbf{z})$), then the vector $f_a(\mathbf{z})$ dominates some of the Pareto optimal objective vectors in Z_m. On the other hand, if a scalar value $f_s(\mathbf{z})$ is too large ($f_u(\mathbf{z}) < f_s(\mathbf{z})$), then some of the vectors in Z_m dominate $f_a(\mathbf{z})$. In addition, if $f_l(\mathbf{z}) < f_s(\mathbf{z}) < f_u(\mathbf{z})$, then the vector $f_a(\mathbf{z})$ does not dominate any of the vectors in Z_m and none of the vectors in Z_m dominate $f_a(\mathbf{z})$.

Geometrically, functions f_l and f_u are defined so that by replacing the function f_s in the function f_a by the function f_l, we have a Pareto front approximation, which is illustrated in Figure 48 as a black dashed curve. In addition, by replacing f_s by the function f_u, we have a Pareto front approximation, which is illustrated as a black dotted curve. The sets (curves) presented in Figure 48 follow
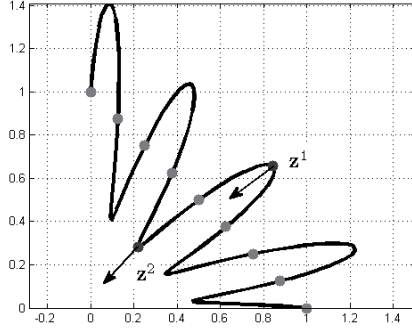
FIGURE 47 Two vectors on the approximated Pareto front so that they either dominate some of the vectors in Z_9 or some of the vectors in Z_9 dominate them.
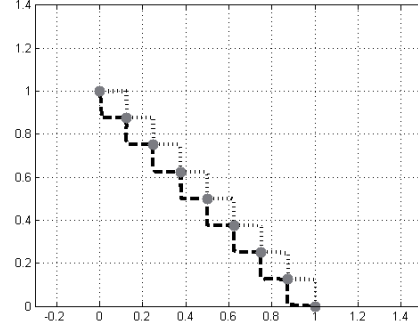
FIGURE 48 Lower and upper limits for function f_s so that the vector $\mathbf{z}$ in the approximated Pareto front does not dominate any of the vectors in Z_9 and any of the vectors in Z_9 do not dominate vector $\mathbf{z}$.

in some sense the ideas presented in [47, 146]. However, the calculations of the sets (curves) in Figure 48 are based on the mathematical structure of the function f_a. In addition, the black dashed curve (bdc) is not a lower bound set as defined in [47, 146] for Z_m because it does not satisfy the condition $bdc \subset (bdc + \mathbb{R}^k_+)^*$ proposed in [47, 146] for a lower bound set. Similarly, the black dotted curve (btc) is not an upper bound set for Z_m because it does not satisfy the condition $btc \subset (btc + \mathbb{R}^k_+)^*$. Even though the sets (curves) presented in Figure 48 do not satisfy the conditions of bound sets, the function value comparisons discussed in the previous paragraph are more essential for our purposes in validating the approximation. Let us demonstrate this as follows.

If we compare Figures 47 and 48, in which the set Z_9 is the same (gray dots), then we can see that vectors $\mathbf{z}^1$ and $\mathbf{z}^2$ do not belong to the boxes defined by the functions f_l and f_u in Figure 48. Here, belonging to the boxes means that $f_l(\mathbf{z}) < f_s(\mathbf{z}) < f_u(\mathbf{z})$. To see that the vectors do not belong to the boxes, let $\tilde{\mathbf{z}}^1, \tilde{\mathbf{z}}^2 \in \Delta_1^\circ$ so that $f_a(\tilde{\mathbf{z}}^1) = \mathbf{z}^1$ and $f_a(\tilde{\mathbf{z}}^2) = \mathbf{z}^2$. It turns out that $f_s(\tilde{\mathbf{z}}^2) < f_l(\tilde{\mathbf{z}}^2)$ and $f_u(\tilde{\mathbf{z}}^1) < f_s(\tilde{\mathbf{z}}^1)$ meaning that $\mathbf{z}^2$ dominates some of the vectors in Z_9, and some of the vectors in Z_9 dominate vector $\mathbf{z}^1$. These findings of the functions f_l, f_u and f_s are the elements in characterizing dominated vectors in this chapter. However, the characterization of dominated vectors is not the only use for the functions.

The functions f_l and f_u define the bounds for the function f_s. Let us assume that the vector $f_a(\mathbf{z})$ is an approximated Pareto optimal objective vector (i.e. condition $f_l(\mathbf{z}) < f_s(\mathbf{z}) < f_u(\mathbf{z})$ holds, and the normal vector belongs to the negative orthant (or the positive orthant)). The decision maker may now wonder how reliable the vector is, that is, how certain he or she can be about the vector, if the vector is considered as a Pareto optimal objective vector to the MOO prob-

lem. Now, the number $f_u(\mathbf{z}) - f_l(\mathbf{z})$ indicates how much the value $f_s(\mathbf{z})$ can vary without dominating any of the vectors in Z_m or vice versa. Then, if $f_u(\mathbf{z}) - f_l(\mathbf{z})$ is small, then it is rather certain that the vector $f_a(\mathbf{z})$ can be a Pareto optimal objective vector compared to a case in which the number $f_u(\mathbf{z}) - f_l(\mathbf{z})$ is very large. To calculate such a *certainty number* for the reliability of the vector as being a Pareto optimal objective vector, a function $f_c : \Delta_{k-1}^{\circ} \to [0,1]$ is proposed in this chapter.

The results of this chapter are essential for the IHBE and HBE methods because we can both validate the approximated Pareto front and also provide additional numerical information for the decision maker about the approximated Pareto optimal objective vectors. In applying the results to the IHBE and HBE methods, computationally efficient calculations are used, i.e. the use of the functions f_l and f_u during the solution process of the MOO problem does not introduce waiting times for the decision maker.

To get started, explicit forms of the functions f_l and f_u are presented in Section 5.1 with studies of characterizations of dominated vectors. In Section 5.2, we combine the result of the previous section to the HBE method, introducing the certainty calculation. Finally, in Section 5.3, we conclude this chapter.

5.1   Scalar values $f_s(\mathbf{z})$ in Characterizations of Dominated Vectors

This chapter concerns the functions f_l and f_u from Δ_{k-1}° into $\mathbb{R}$ and the explicit forms of them are the following.

Definition 39. Functions f_l and f_u are defined as follows:

$$f_l(\mathbf{z}) = \max_{i=1,\dots,m} \min_{j=1,\dots,k} \frac{z_j^i - \mathrm{ideal}_j^{Z_m}}{z_j}, \quad f_u(\mathbf{z}) = \min_{i=1,\dots,m} \max_{j=1,\dots,k} \frac{z_j^i - \mathrm{ideal}_j^{Z_m}}{z_j},$$

where $\mathbf{z}^i \in Z_m$, $i = 1, \dots, m$.

The explicit forms of functions f_l and f_u are rather simple, and the construction of them is presented in Appendix 3. Next, we prove that functions f_l and f_u characterize dominated vectors in the approximation.

Theorem 40. Let $\mathbf{z} \in \Delta_{k-1}^{\circ}$.

1. Then $f_l(\mathbf{z}) \leq f_u(\mathbf{z})$.
2. Functions f_l and f_u are continuous at $\mathbf{z}$.
3. If $f_s(\mathbf{z}) < f_l(\mathbf{z})$, then $f_a(\mathbf{z})$ dominates at least one vector $\mathbf{z}^l \in Z_m$.
4. If $f_u(\mathbf{z}) < f_s(\mathbf{z})$, then some vector $\mathbf{z}^u \in Z_m$ dominates $f_a(\mathbf{z})$.
5. If $f_l(\mathbf{z}) < f_s(\mathbf{z})$, then $f_a(\mathbf{z})$ does not dominate any of the vectors in Z_m.
6. If $f_s(\mathbf{z}) < f_u(\mathbf{z})$, then none of the vectors in Z_m dominate $f_a(\mathbf{z})$.
7. If $f_l(\mathbf{z}) < f_s(\mathbf{z}) < f_u(\mathbf{z})$, then none of the vectors in Z_m dominate $f_a(\mathbf{z})$ and $f_a(\mathbf{z})$ does not dominate any of the vectors in Z_m.

Proof. We split the proof into its corresponding parts.

1. Let us assume by a contradiction that there is $\mathbf{z} \in \Delta_{k-1}^{\circ}$ so that $f_l(\mathbf{z}) > f_u(\mathbf{z})$. Let i_l and i_u be the indices for which we have

$$f_l(\mathbf{z}) = \min_j \frac{z_j^{i_l} - \text{ideal}_j^{Z_m}}{z_j} \quad \text{and} \quad f_u(\mathbf{z}) = \max_j \frac{z_j^{i_u} - \text{ideal}_j^{Z_m}}{z_j}.$$

Now, for all $j = 1, \ldots, k$ we have

$$z_j^{i_l} - \text{ideal}_j^{Z_m} = \frac{z_j^{i_l} - \text{ideal}_j^{Z_m}}{z_j} z_j \geq f_l(\mathbf{z})z_j > f_u(\mathbf{z})z_j \geq \frac{z_j^{i_u} - \text{ideal}_j^{Z_m}}{z_j} z_j$$

$$= z_j^{i_u} - \text{ideal}_j^{Z_m}.$$

In other words, we have $z_j^{i_l} > z_j^{i_u}$ for all $j = 1, \ldots, k$. This means that the vector $\mathbf{z}^{i_u}$ dominates the vector $\mathbf{z}^{i_l}$, which is a contradiction.

2. Because

$$\max\{a_1, a_2\} = \frac{a_1 + a_2}{2} + \frac{|a_1 - a_2|}{2}, \quad \min\{a_1, a_2\} = \frac{a_1 + a_2}{2} - \frac{|a_1 - a_2|}{2},$$

then functions f_l and f_u can be written in a form of a sum of continuous functions.

3. Let $\mathbf{z} \in \Delta_{k-1}^{\circ}$ and let i^* be the index for which we obtain

$$\min_j \frac{z_j^{i^*} - \text{ideal}_j^{Z_m}}{z_j} = \max_i \min_j \frac{z_j^{i} - \text{ideal}_j^{Z_m}}{z_j}.$$

Then

$$f_a(\mathbf{z})_l = \text{ideal}_l^{Z_m} + f_s(\mathbf{z})z_l < \text{ideal}_l^{Z_m} + \max_i \min_j \frac{z_j^{i} - \text{ideal}_j^{Z_m}}{z_j} z_l$$

$$= \text{ideal}_l^{Z_m} + \min_j \frac{z_j^{i^*} - \text{ideal}_j^{Z_m}}{z_j} z_l \leq \text{ideal}_l^{Z_m} + \frac{z_l^{i^*} - \text{ideal}_l^{Z_m}}{z_l} z_l = z_l^{i^*}.$$

In other words, inequality $f_a(\mathbf{z})_l < z_l^{i^*}$ holds for all $l = 1, \ldots, k$, which means that $f_a(\mathbf{z})$ dominates $\mathbf{z}^{i^*}$.

4. The proof is similar to the one above.

5. Let us assume by contradiction that $f_a(\mathbf{z})$ dominates $\tilde{\mathbf{z}} \in Z_m$. It is clear that the vector $\textbf{ideal}^{Z_m} + f_l(\mathbf{z})\mathbf{z}$ dominates $f_a(\mathbf{z})$. Then

$$\text{ideal}_l^{Z_m} + f_l(\mathbf{z})z_l < \tilde{z}_l \quad \text{for all} \quad l = 1, \ldots, k. \tag{30}$$

By calculating, we obtain

$$\text{ideal}_l^{Z_m} + f_l(\mathbf{z})z_l = \text{ideal}_l^{Z_m} + \max_i \min_j \frac{z_j^{i} - \text{ideal}_j^{Z_m}}{z_j} z_l$$

$$\geq \text{ideal}_l^{Z_m} + \min_j \frac{\tilde{z}_j - \text{ideal}_j^{Z_m}}{z_j} z_l \tag{31}$$

for all l. By combining (30) and (31), inequalities

$$\tilde{z}_l - \text{ideal}_l^{Z_m} > \min_j \frac{\tilde{z}_j - \text{ideal}_j^{Z_m}}{z_j} z_l$$

holds for all $l = 1, \ldots, k$ and especially for the index j^* for which we obtain $\min_j \frac{\tilde{z}_j - \text{ideal}_j^{Z_m}}{z_j}$. Therefore $\tilde{z}_{j^*} - \text{ideal}_{j^*}^{Z_m} > \tilde{z}_{j^*} - \text{ideal}_{j^*}^{Z_m}$, which is a contradiction. Thus, the vector $f_a(\mathbf{z})$ does not dominate any of the vectors in Z_m.

6. The proof is similar to the one above.
7. Follows from the two previous cases.

$\square$

One may wonder why in Theorem 40 only strict inequalities are considered. If we relax the strict inequalities to also cover equalities in 3.-7., then we cannot conclude for sure, for example, whether $f_a(\mathbf{z})$ dominates some vector in Z_m. We demonstrate this as follows. Let vectors in Z_3 be as presented in Figure 49 (gray dots) and let $\mathbf{z}', \mathbf{z}'' \in \{\tilde{\mathbf{z}} \in \mathbb{R}^k : \tilde{\mathbf{z}} = \mathbf{0} + f_l(\mathbf{z})\mathbf{z}, \mathbf{z} \in \Delta_1^\circ\}$ (black dots). For simplicity, let $\mathbf{z}^{s'} = f_{Z_m}^p(\mathbf{z}')$ and $\mathbf{z}^{s''} = f_{Z_m}^p(\mathbf{z}'')$. Now, we have $f_s(\mathbf{z}^{s'}) = f_l(\mathbf{z}^{s'})$ and $f_s(\mathbf{z}^{s''}) = f_l(\mathbf{z}^{s''})$, which means that $\mathbf{z}'$ dominates $\mathbf{z}^2 \in Z_3$. However, the vector $\mathbf{z}'' (= \mathbf{z}^2)$ does not dominate any of the vectors in Z_3 because a vector in the objective space cannot be dominated by itself. In other words, the characterization of dominated vectors in Theorem 40 is not clear because of the equalities.
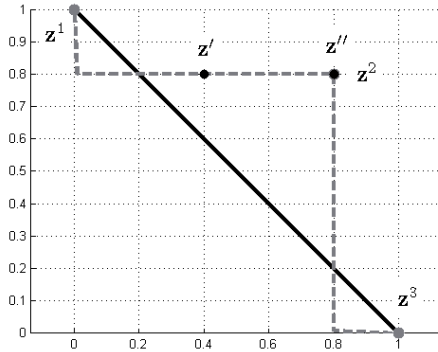
FIGURE 49   Strict inequality conditions in Theorem 40 are essential.

To allow equalities in Theorem 40, several special cases must be considered. After that, we modify Theorem 40 to include the equalities.

As seen in Figure 49, vectors $\mathbf{z}''$ and $\mathbf{z}^2$ are the same. Note that the condition $f_l(\mathbf{z}) = f_s(\mathbf{z})$ for a vector $\mathbf{z} \in \Delta_{k-1}^\circ$ does not guarantee that $f_a(\mathbf{z}) \in Z_m$. In Figure 49, we have $f_l(\mathbf{z}^{s'}) = f_s(\mathbf{z}^{s'})$, but $f_a(\mathbf{z}^{s'}) \notin Z_m$. However, if we also have $f_s(\mathbf{z}) = f_u(\mathbf{z})$, then $f_a(\mathbf{z}) \in Z_m$ as we show next.

Proposition 41. Let $\mathbf{z} \in \Delta_{k-1}^{\circ}$. Then $f_l(\mathbf{z}) = f_s(\mathbf{z}) = f_u(\mathbf{z})$ if and only if $f_a(\mathbf{z}) \in Z_m$.

Proof.

"$\Rightarrow$" Let i_l and i_u be the indices for which we have

$$f_l(\mathbf{z}) = \min_j \frac{z_j^{i_l} - \text{ideal}_j^{Z_m}}{z_j} \quad \text{and} \quad f_u(\mathbf{z}) = \max_j \frac{z_j^{i_u} - \text{ideal}_j^{Z_m}}{z_j}.$$

Now, for all $j = 1, \ldots, k$

$$z_j^{i_l} - \text{ideal}_j^{Z_m} = \frac{z_j^{i_l} - \text{ideal}_j^{Z_m}}{z_j} z_j \geq f_l(\mathbf{z}) z_j = f_u(\mathbf{z}) z_j \geq \frac{z_j^{i_u} - \text{ideal}_j^{Z_m}}{z_j} z_j \quad (32)$$

$$= z_j^{i_u} - \text{ideal}_j^{Z_m},$$

$$\Rightarrow z_j^{i_l} \geq z_j^{i_u}. \quad (33)$$

Because every vector in Z_m is Pareto optimal, then any of the inequalities cannot be strict in (33). Otherwise, the vector $\mathbf{z}^{i_u}$ would dominate the vector $\mathbf{z}^{i_l}$. In other words, the vectors $\mathbf{z}^{i_u}$ and $\mathbf{z}^{i_l}$ are the same. This means that the indices i_l and i_u are the same. Thus

$$f_a(\mathbf{z}) = \textbf{ideal}^{Z_m} + f_s(\mathbf{z})\mathbf{z} = \textbf{ideal}^{Z_m} + f_l(\mathbf{z})\mathbf{z} = \mathbf{z}^{i_l} \in Z_m,$$

where the last equality follows from (32).

"$\Leftarrow$" If we have $f_s(\mathbf{z}) < f_l(\mathbf{z})$, then by Theorem 40 the vector $f_a(\mathbf{z})$ dominates some of the vectors in Z_m, which cannot be possible because every vector in Z_m is Pareto optimal. Therefore, we must have $f_l(\mathbf{z}) \leq f_s(\mathbf{z})$.

If we have $f_l(\mathbf{z}) < f_s(\mathbf{z})$, then for all $j = 1, \ldots, k$ we have

$$f_a(\mathbf{z})_j - \text{ideal}_j^{Z_m} = f_s(\mathbf{z})z_j > f_l(\mathbf{z})z_j = \max_i \min_l \frac{z_l^i - \text{ideal}_l^{Z_m}}{z_l} z_j$$

$$\geq \min_l \frac{f_a(\mathbf{z})_l - \text{ideal}_l^{Z_m}}{z_l} z_j. \quad (34)$$

Let l^* be the index for which we have

$$\frac{f_a(\mathbf{z})_{l^*} - \text{ideal}_{l^*}^{Z_m}}{z_{l^*}} z_j = \min_l \frac{f_a(\mathbf{z})_l - \text{ideal}_l^{Z_m}}{z_l} z_j.$$

Since (34) holds for all $j = 1, \ldots, k$, also for l^*, then

$$f_a(\mathbf{z})_{l^*} - \text{ideal}_{l^*}^{Z_m} > \frac{f_a(\mathbf{z})_{l^*} - \text{ideal}_{l^*}^{Z_m}}{z_{l^*}} z_{l^*} = f_a(\mathbf{z})_{l^*} - \text{ideal}_{l^*}^{Z_m},$$

which is a contradiction. Thus $f_s(\mathbf{z}) = f_l(\mathbf{z})$.

The proof for $f_s(\mathbf{z}) = f_u(\mathbf{z})$ is very similar to the one above.

$\square$

Now, Proposition 41 completes Theorem 40 in the following way:

Theorem 42. Let $\mathbf{z} \in \Delta^{\circ}_{k-1}$.

1. We have $f_s(\mathbf{z}) \leq f_l(\mathbf{z})$ and $f_s(\mathbf{z}) < f_u(\mathbf{z})$ if and only if $f_a(\mathbf{z})$ dominates some vector in Z_m.
2. In addition, $f_u(\mathbf{z}) \leq f_s(\mathbf{z})$ and $f_l(\mathbf{z}) < f_s(\mathbf{z})$ if and only if some vector in Z_m dominates $f_a(\mathbf{z})$.
3. The vector $f_a(\mathbf{z})$ does not dominate any of the vectors in Z_m, and none of the vectors in Z_m dominate $f_a(\mathbf{z})$ if and only if either $f_l(\mathbf{z}) < f_s(\mathbf{z}) < f_u(\mathbf{z})$ or $f_l(\mathbf{z}) = f_s(\mathbf{z}) = f_u(\mathbf{z})$ is true.

Proof.

1. "$\Rightarrow$" If we have $f_s(\mathbf{z}) < f_l(\mathbf{z})$, then Theorem 40 guarantees that we have a vector in Z_m, which is dominated by $f_a(\mathbf{z})$. Let us next have $f_s(\mathbf{z}) = f_l(\mathbf{z})$. Then $f_a(\mathbf{z})$ cannot belong to Z_m. Otherwise, we would have $f_s(\mathbf{z}) = f_u(\mathbf{z})$ by Proposition 41, and that would be a contradiction with the assumption $f_s(\mathbf{z}) < f_u(\mathbf{z})$.

   Let i_l be the index for which we have

$$f_l(\mathbf{z}) = \min_j \frac{z_j^{i_l} - \text{ideal}_j^{Z_m}}{z_j}.$$

   Now, for all $j = 1, \ldots, k$ we have

$$z_j^{i_l} - \text{ideal}_j^{Z_m} = \frac{z_j^{i_l} - \text{ideal}_j^{Z_m}}{z_j} z_j \geq f_l(\mathbf{z})z_j = f_s(\mathbf{z})z_j$$

$$\Rightarrow$$

$$z_j^{i_l} \geq \text{ideal}_j^{Z_m} + f_s(\mathbf{z})z_j = f_a(\mathbf{z})_j. \tag{35}$$

   Since $f_a(\mathbf{z}) \notin Z_m$, then some of the inequalities in (35) must be strict. Thus, the vector $f_a(\mathbf{z})$ dominates the vector $\mathbf{z}^{i_l}$.

   "$\Leftarrow$" Let us assume, on the contrary, that the statement does not hold for some $\mathbf{z} \in \Delta^{\circ}_{k-1}$. In other words,

$$\neg \left[\left[f_s(\mathbf{z}) \leq f_l(\mathbf{z}) \right] \wedge \left[f_s(\mathbf{z}) < f_u(\mathbf{z}) \right] \right] =$$
$$\neg \left[f_s(\mathbf{z}) \leq f_l(\mathbf{z}) \right] \vee \neg \left[f_s(\mathbf{z}) < f_u(\mathbf{z}) \right] =$$
$$\left[f_s(\mathbf{z}) > f_l(\mathbf{z}) \right] \vee \left[f_s(\mathbf{z}) \geq f_u(\mathbf{z}) \right]$$

   is true. Thus, at least one of the following must be true
   (a) $f_s(\mathbf{z}) > f_l(\mathbf{z})$ and $f_s(\mathbf{z}) \geq f_u(\mathbf{z})$,
   (b) $f_s(\mathbf{z}) > f_l(\mathbf{z})$ and $f_s(\mathbf{z}) < f_u(\mathbf{z})$,
   (c) $f_s(\mathbf{z}) \leq f_l(\mathbf{z})$ and $f_s(\mathbf{z}) \geq f_u(\mathbf{z})$.

If (a) is true, then $f_a(\mathbf{z})$ does not dominate any of the vectors in Z_m by Theorem 40. This is a contradiction, and therefore (a) cannot be true. Similarly, alternative (b) cannot be true. Case (c) would be true if $f_l(\mathbf{z}) = f_s(\mathbf{z}) = f_u(\mathbf{z})$, but then $f_a(\mathbf{z}) \in Z_m$ by Proposition 41 and therefore it does not dominate any of the vectors in Z_m. Thus (c) is not true. Because neither (a) nor (b) nor (c) is true, then we have a contradiction.

2. The proof is similar to 1.

3. "$\Leftarrow$" Follows immediately from Theorem 40 and Proposition 41.

"$\Rightarrow$" Let us assume, on the contrary, that for some $\mathbf{z} \in \Delta_{k-1}^\circ$ none of the conditions $f_l(\mathbf{z}) < f_s(\mathbf{z}) < f_u(\mathbf{z})$ and $f_l(\mathbf{z}) = f_s(\mathbf{z}) = f_u(\mathbf{z})$ is true. In other words,

$$\neg\left[\left[(f_l(\mathbf{z}) < f_s(\mathbf{z})) \wedge (f_s(\mathbf{z}) < f_u(\mathbf{z}))\right] \vee \left[(f_l(\mathbf{z}) = f_s(\mathbf{z})) \wedge (f_s(\mathbf{z}) = f_u(\mathbf{z}))\right]\right] =$$
$$\neg\left[(f_l(\mathbf{z}) < f_s(\mathbf{z})) \wedge (f_s(\mathbf{z}) < f_u(\mathbf{z}))\right] \wedge \neg\left[(f_l(\mathbf{z}) = f_s(\mathbf{z})) \wedge (f_s(\mathbf{z}) = f_u(\mathbf{z}))\right] =$$
$$\left[\neg(f_l(\mathbf{z}) < f_s(\mathbf{z})) \vee \neg(f_s(\mathbf{z}) < f_u(\mathbf{z}))\right] \wedge \left[\neg(f_l(\mathbf{z}) = f_s(\mathbf{z})) \vee \neg(f_s(\mathbf{z}) = f_u(\mathbf{z}))\right] =$$
$$\left[(f_l(\mathbf{z}) \geq f_s(\mathbf{z})) \vee (f_s(\mathbf{z}) \geq f_u(\mathbf{z}))\right] \wedge \left[(f_l(\mathbf{z}) \neq f_s(\mathbf{z})) \vee (f_s(\mathbf{z}) \neq f_u(\mathbf{z}))\right]$$

is true. This means that at least one of the following conditions

(a) $f_s(\mathbf{z}) \leq f_l(\mathbf{z})$ and $f_l(\mathbf{z}) \neq f_s(\mathbf{z})$
(b) $f_s(\mathbf{z}) \leq f_l(\mathbf{z})$ and $f_s(\mathbf{z}) \neq f_u(\mathbf{z})$
(c) $f_u(\mathbf{z}) \leq f_s(\mathbf{z})$ and $f_l(\mathbf{z}) \neq f_s(\mathbf{z})$
(d) $f_u(\mathbf{z}) \leq f_s(\mathbf{z})$ and $f_s(\mathbf{z}) \neq f_u(\mathbf{z})$
(e) $f_s(\mathbf{z}) \leq f_l(\mathbf{z})$ and $f_u(\mathbf{z}) \leq f_s(\mathbf{z})$ and $f_l(\mathbf{z}) \neq f_s(\mathbf{z})$
(f) $f_s(\mathbf{z}) \leq f_l(\mathbf{z})$ and $f_u(\mathbf{z}) \leq f_s(\mathbf{z})$ and $f_s(\mathbf{z}) \neq f_u(\mathbf{z})$
(g) $f_s(\mathbf{z}) \leq f_l(\mathbf{z})$ and $f_l(\mathbf{z}) \neq f_s(\mathbf{z})$ and $f_s(\mathbf{z}) \neq f_u(\mathbf{z})$
(h) $f_u(\mathbf{z}) \leq f_s(\mathbf{z})$ and $f_l(\mathbf{z}) \neq f_s(\mathbf{z})$ and $f_s(\mathbf{z}) \neq f_u(\mathbf{z})$
(i) $f_s(\mathbf{z}) \leq f_l(\mathbf{z})$ and $f_u(\mathbf{z}) \leq f_s(\mathbf{z})$ and $f_l(\mathbf{z}) \neq f_s(\mathbf{z})$ and $f_s(\mathbf{z}) \neq f_u(\mathbf{z})$

must be true.

If (a) is true, then $f_s(\mathbf{z}) < f_l(\mathbf{z})$ and Theorem 40 guarantees that $f_a(\mathbf{z})$ dominates at least one vector in Z_m. This is contrary to the assumption that $f_a(\mathbf{z})$ does not dominate any of the vectors in Z_m. Thus (a) is not true.

If (b) is true, then we have two possible alternatives:

– $f_s(\mathbf{z}) < f_l(\mathbf{z})$ and $f_s(\mathbf{z}) \neq f_u(\mathbf{z})$, which cannot be true by Theorem 40,
– $f_s(\mathbf{z}) = f_l(\mathbf{z})$ and $f_l(\mathbf{z}) < f_u(\mathbf{z})$, which cannot be true by 1.

Thus (b) is not true.

If (c) is true, then we have two possible alternatives:

– $f_u(\mathbf{z}) < f_s(\mathbf{z})$ and $f_s(\mathbf{z}) \neq f_u(\mathbf{z})$, which cannot be true by Theorem 40,
– $f_u(\mathbf{z}) = f_s(\mathbf{z})$ and $f_l(\mathbf{z}) < f_u(\mathbf{z})$, which cannot be true by 1.

Thus (c) is not true.

If (d) is true, then $f_u(\mathbf{z}) < f_s(\mathbf{z})$ and Theorem 40 guarantee that some vector in Z_m dominates $f_a(\mathbf{z})$. This is against the assumption that none of the vectors in Z_m dominates $f_a(\mathbf{z})$. Thus (d) is not true.

None of the cases (e)-(i) is true, because otherwise some of the cases (a)-(d) would be true. In other words, none of the possible cases (a) - (i) is true. This is a contradiction.

$\square$

Now, we are ready to discuss how functions f_l and f_u are used in the IHBE and HBE methods.

5.2 Functions f_l, f_u and f_s in the IHBE and HBE Methods

Functions f_l, f_u and f_s can be used in investigating whether or not a vector $f_a(\mathbf{z})$ dominates some of the vectors in Z_m or vice versa. Implementations of the results of theorems presented in the previous section for the IHBE and HBE methods can be done easily because only values $f_s(\mathbf{z})$, $f_l(\mathbf{z})$ and $f_u(\mathbf{z})$ must be calculated and each of them is not a computationally expensive operation to be performed. Then making comparisons, i.e. $f_s(\mathbf{z}) < f_l(\mathbf{z})$, is not a problem. In other words, functions f_l and f_u can be implemented and used without high computational costs.

In this section, we update Method 37 to cover also dominated vectors identified by theorems found for functions f_l, f_u and f_s. This is done in Subsection 5.2.1. In other words, we strengthen the validation, which is based on the normal vectors, to make it more accurate. In Subsection 5.2.2, we propose a method to be used in steps (3-4) of HBE in estimating how reliable or certain an approximated Pareto optimal objective vector is.

5.2.1 Validating the Approximation by Functions f_l, f_u and f_s

Before the decision maker starts the solution process of the MOO problem, the approximation must be validated in order to avoid cases where e.g. an approximated Pareto optimal objective vector dominates some of the vectors in Z_m. As shown in Theorem 40, if $f_s(\mathbf{z}) < f_l(\mathbf{z})$, then $f_a(\mathbf{z})$ dominates some of the vectors in Z_m, and this motivates us to update Method 37 as follows.

Method 43 ("Method 37.1").

(1) Generate sample vectors $\mathbf{z}^{s_i}$ on Δ_{k-1}° and calculate both the normal vectors and function values f_l, f_u and f_s at them.

(2) Validate the approximation as follows.

(2.1) If every normal vector belongs to $\mathbb{R}_+^k \cup \mathbb{R}_-^k$ and for all sample vectors the condition $f_l(\mathbf{z}^{s_i}) < f_s(\mathbf{z}^{s_i}) < f_u(\mathbf{z}^{s_i})$ holds, then the approximation may be a nondominated set.

(2.2) If for some of the sample vectors the condition $f_s(\mathbf{z}^{s_i}) < f_l(\mathbf{z}^{s_i})$ holds, then the approximation contains vectors which dominate Pareto optimal objective vectors to the MOO problem.

(2.3) If some of the normal vectors do not belong to $\mathbb{R}^k_+ \cup \mathbb{R}^k_-$ or for some of the sample vectors condition $f_u(\mathbf{z}^{s_i}) < f_s(\mathbf{z}^{s_i})$ holds, then the approximation contains dominated vectors.

Method 43 is still simple after the update. However, one may wonder why condition $f_u(\mathbf{z}^{s_i}) < f_s(\mathbf{z}^{s_i})$ is not in step (2.2). As shown, Theorem 40 guarantees that at least one of the vectors in Z_m dominates vector $f_a(\mathbf{z}^{s_i})$, if $f_u(\mathbf{z}^{s_i}) < f_s(\mathbf{z}^{s_i})$. The reason why it has not been included in Method 43 can be illustrated as follows. As already discussed in Section 2.2, the nonconvexity of the MOO problem may force the approximation to contain dominated vectors. In Figure 50, we have a case in which $f_u(\mathbf{z}) < f_s(\mathbf{z})$ at $\mathbf{z} \in \Delta_1^\circ$. This means that the vector $f_a(\mathbf{z})$ is dominated by Theorem 40. However, vectors $(0,1)^T$, $(0.2,0.8)^T$, $(0.79,0.79)^T$, $(0.8,0.2)^T$ and $(1,0)^T$ (gray dots) in the set Z_5 define a very narrow margin $[f_l(\mathbf{z}), f_u(\mathbf{z})]$ for the scalar $f_s(\mathbf{z})$ at $\mathbf{z} \in \Delta_1$. Note that the margin can be arbitrarily small if we replace vector $(0.79,0.79)^T$ by vector $(0.8 - \varepsilon, 0.8 - \varepsilon)^T$, $\varepsilon > 0$. In other words, staying between the ranges $[f_l(\mathbf{z}), f_u(\mathbf{z})]$ can be sometimes difficult in practice with the surrogate functions. That is, the black curve representing the approximated Pareto front is not inside the boxes defined by the functions f_l and f_u. However, if it turns out that for some vector $\mathbf{z} \in \Delta_{k-1}^\circ$ the difference $f_s(\mathbf{z}) - f_u(\mathbf{z})$ is extremely large, then we have an indication that the function f_s may not behave properly in its relation to the Pareto dominance, i.e. it can oscillate as in cases illustrated in Figures 41 and 47. An update for Method 43 can be done here as well, but it is not clear how much the difference must be. Therefore, we have left out the condition $f_u(\mathbf{z}^{s_i}) < f_s(\mathbf{z}^{s_i})$ in step (2.2).
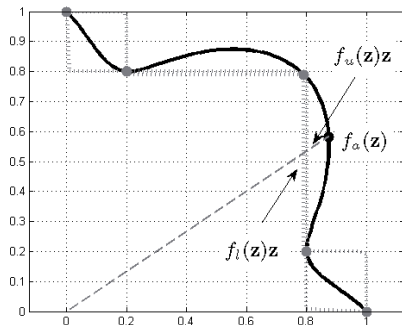
FIGURE 50   Condition $f_u(\mathbf{z}) < f_s(\mathbf{z})$ is not necessarily that critical in investigating dominated vectors in $f_a(\Delta_{k-1}^\circ)$.

When using functions f_l and f_u in validating the approximation, it must be clarified that the validation is based on vectors in Z_m and a vector $\mathbf{z} \in \Delta_{k-1}^\circ$. In other words, the validation is done in a vector-wise sense in Theorems 40 and 42,

and not as a local neighborhood validation likewise the validation based on normal vectors. However, a validation based on functions f_l and f_u can be established on an open neighborhood, as we see next.

Theorem 44. Let $\mathbf{z} \in \Delta^{\circ}_{k-1}$. If the function f_s is continuous, then there exists a radius $r > 0$ so that for all $\mathbf{z}' \in B(\mathbf{z}, r) \cap \Delta^{\circ}_{k-1}$ we have $f_s(\mathbf{z}') < f_l(\mathbf{z}')$ or $f_u(\mathbf{z}') < f_s(\mathbf{z}')$ if $f_s(\mathbf{z}) < f_l(\mathbf{z})$ or $f_u(\mathbf{z}) < f_s(\mathbf{z})$ respectively.

Proof. Because the function f_l is continuous (Theorem 40), then for a sufficiently small $0 < \varepsilon < (f_l(\mathbf{z}) - f_s(\mathbf{z}))/3$, we have a radius $r_{\mathbf{z}} > 0$ so that

$$|f_s(\mathbf{z}) - f_s(\mathbf{z}')| < \varepsilon, |f_l(\mathbf{z}) - f_l(\mathbf{z}')| < \varepsilon$$

for all $\mathbf{z}' \in \Delta^{\circ}_{k-1}$, $\|\mathbf{z}' - \mathbf{z}\| < r_{\mathbf{z}}$. Now,

$$\begin{aligned}
|f_l(\mathbf{z}) - f_s(\mathbf{z})| &= |f_l(\mathbf{z}) - f_l(\mathbf{z}') + f_l(\mathbf{z}') - f_s(\mathbf{z}') + f_s(\mathbf{z}') - f_s(\mathbf{z})| \\
&\leq |f_l(\mathbf{z}) - f_l(\mathbf{z}')| + |f_l(\mathbf{z}') - f_s(\mathbf{z}')| + |f_s(\mathbf{z}') - f_s(\mathbf{z})| \\
&\leq 2\varepsilon + |f_l(\mathbf{z}') - f_s(\mathbf{z}')| < \frac{2}{3}|f_l(\mathbf{z}) - f_s(\mathbf{z})| + |f_l(\mathbf{z}') - f_s(\mathbf{z}')| \\
&\Leftrightarrow \\
0 &< \frac{1}{3}|f_l(\mathbf{z}) - f_s(\mathbf{z})| < |f_l(\mathbf{z}') - f_s(\mathbf{z}')| \\
&\Leftrightarrow \\
\varepsilon &< |f_l(\mathbf{z}') - f_s(\mathbf{z}')|.
\end{aligned}$$

In other words, for all $\mathbf{z}' \in \Delta^{\circ}_{k-1}$, $\|\mathbf{z}' - \mathbf{z}\| < r_{\mathbf{z}}$, we have $f_s(\mathbf{z}') < f_l(\mathbf{z}')$ or $f_l(\mathbf{z}') < f_s(\mathbf{z}')$. However, if the last one holds then $f_l(\mathbf{z}') < f_s(\mathbf{z})$, because $|f_s(\mathbf{z}) - f_s(\mathbf{z}')| < \varepsilon$ and $f_s(\mathbf{z}') - f_l(\mathbf{z}') > \varepsilon$, which means that

$$|f_l(\mathbf{z}) - f_s(\mathbf{z})| < |f_l(\mathbf{z}) - f_l(\mathbf{z}')| < \varepsilon < |f_l(\mathbf{z}) - f_s(\mathbf{z})|/3.$$

Thus, for all $\mathbf{z}' \in \Delta^{\circ}_{k-1}$, $\|\mathbf{z}' - \mathbf{z}\| < r_{\mathbf{z}}$, we have $f_s(\mathbf{z}') < f_l(\mathbf{z}')$.

The proof for the function f_u is similar. $\square$

Theorem 44 extends Theorem 40 to an open neighborhood of $\mathbf{z} \in \Delta^{\circ}_{k-1}$. Thus, for example, on an relatively open neighborhood of $\mathbf{z} \in \Delta^{\circ}_{k-1}$ every vector dominates some vector in Z_m if $f_s(\mathbf{z}) < f_l(\mathbf{z})$. This increases the value of the functions f_l and f_u in the validations. However, Theorem 44 does not present any clue about how large the open set can be. In other words, we can now find dominated vectors in the approximation more accurately. This leads to a question presented in the next paragraph.

In Method 43, dominated vectors in the approximation are detected. Now one can ask whether all dominated vectors can be detected. The answer is, unfortunately, no. In nonconvex MOO problems, it is possible that a vector $f_a(\mathbf{z})$ is dominated in the approximated Pareto front even though conditions $f_l(\mathbf{z}) < f_s(\mathbf{z}) < f_u(\mathbf{z})$ and the normal vector belong to the negative orthant hold. We have illustrated such a case in Figure 51 at $f_a(\mathbf{z}^3)$. This means that more attention

must be paid in future studies to cases like the vector $f_a(\mathbf{z}^3)$ because neither normal vectors nor functions f_l and f_u indicate that the vector is dominated in the approximated Pareto front. However, methods proposed in this thesis for detecting dominated vectors complete each other. In other words, if a normal vector does not indicate a dominated vector, then the function f_l may indicate that it is dominated, see vector $f_a(\mathbf{z}^1)$ in Figure 51. On the other hand, functions f_l or f_u may not indicate the dominance, but then normal vector indicates something else, see vector $f_a(\mathbf{z}^2)$. Thus, both of them are justified for characterizing dominated vectors in the approximated Pareto front (note that it can be possible that both of them detect a dominated vector, that is, the normal vector does not belong to $\mathbb{R}^k_+ \cup \mathbb{R}^k_-$ and $f_u(\mathbf{z}) < f_s(\mathbf{z})$). This is essential for the validation and for a later use in Chapter 6.
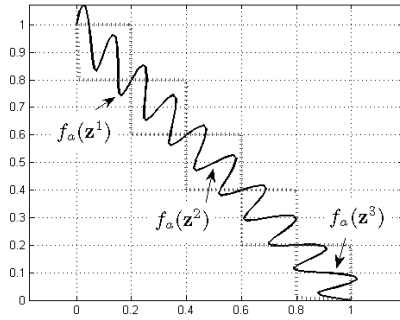
FIGURE 51   Dominated vectors which can be detected by using the validation methods proposed in this thesis. However, they do not cover all the cases.

Remark 45. We believe that it would be quite unlikely (or impossible) to be able to present in one thesis a complete collection of computational methods which detect all dominated vectors correctly. To see this, let us assume that we have a collection of computational methods which detect both nondominated and dominated vectors correctly in a set which can contain infinitely many vectors. Then it can be said that we have a computational method (a combination of the methods) which characterizes the definitions of Pareto optimality and Pareto dominance. Does such a computational method exist? We do not know for sure, but we believe that it is not possible to construct such computationally methods, which work perfectly with all arbitrary nonconvex MOO problems. On the other hand, we do not underestimate the results obtained by normal vectors and functions f_l and f_u. They are pointing out that such computational methods can be created, and in practice they can provide good results in detecting nondominated and dominated vectors.

To conclude this subsection, validating the approximation by functions f_l, f_u and f_s was discussed. It turned out that a computationally efficient method can be formulated for validating the approximation. In this subsection, the validation of

the approximation means finding out whether or not the approximation contains vectors which dominate some of the vectors in Z_m.

5.2.2 Reliability of Vectors in the Approximated Pareto Front

In this subsection, we propose a function $f_c : \Delta_{k-1}^\circ \rightarrow [0,1]$ estimating how reliable an approximated Pareto optimal objective vector is if it is considered as a Pareto optimal objective vector to the MOO problem. By the function value, the HBE method can provide an estimation of the reliability in step (4) for a sample vector in the approximation that has been generated based on the decision maker's preferences. Another use for the function could be in step (6) of IHBE, where the decision maker can point the approximated Pareto optimal objective vectors for which the corresponding true Pareto optimal objective vectors should be calculated. With the help of the value of the function f_c, the decision maker may estimate for which vectors there is no need to calculate a true Pareto optimal objective vector. Let us illustrate this as follows. If the *certainty number* $f_c(\mathbf{z})$ is close to one, then the vector $f_a(\mathbf{z})$ can be interpreted to be rather reliable (certain), that is, it could be a Pareto optimal objective vector to the MOO problem. In other words, there is not necessarily a need to calculate a true Pareto optimal objective vector to the approximated one[1]. On the other hand, if $f_c(\mathbf{z})$ is close to zero, then the vector $f_a(\mathbf{z})$ is *uncertain*, i.e. we cannot be sure whether or not the vector $f_a(\mathbf{z})$ can be considered as a Pareto optimal objective vector. This means that a true Pareto optimal objective vector corresponding to the approximated one should be calculated. Note that calculating the true ones requires time because the MOO problem is computationally expensive. In other words, with the help of the certainty number, the decision maker may save computation time by rejecting calculations of certain (reliable) approximated vectors. However, we want to remind that the function f_c is only an educated guess and situations discussed in Subsection 3.1.2.2 may take place with it.

To put it briefly, we propose the value $f_c(\mathbf{z})$ to be calculated as follows:

$$f_c(\mathbf{z}) := 1 - \min\left\{ \frac{f_u(\mathbf{z}) - f_l(\mathbf{z})}{\sum_{i=1}^k \text{nadir}_i^{Z_m} - \text{ideal}_i^{Z_m}}, 1 \right\}.$$

Geometrically, the numerator $f_u(\mathbf{z}) - f_l(\mathbf{z})$ measures the range of scalar values $f_s(\mathbf{z})$ such that the vector $f_a(\mathbf{z})$ does not dominate any of the vectors in the set Z_m and vice versa. The denominator $\sum_{i=1}^k \text{nadir}_i^{Z_m} - \text{ideal}_i^{Z_m}$ is in some sense a scaling factor in order to get the value to the range $[0,1]$. The scaling factor represent in some sense a maximum of differences $f_u(\mathbf{z}) - f_l(\mathbf{z})$ that can be possible in practice. This can be seen as follows. Let Z_3 consists of three vectors: $(1,0,0)^T$, $(0,1,0)^T$ and $(0,0,1)^T$. Then the function value $f_u((1/3,1/3,1/3)^T) - f_l((1/3,1/3,1/3)^T)$ is $\sum_{i=1}^3 \text{nadir}_i^{Z_m} - \text{ideal}_i^{Z_m}$ and that is the maximum difference, which can be found in this case. However, as demonstrated in Appendix 3,

---

[1]   If the decision maker is interested in decision variable values, then a corresponding true Pareto optimal objective vector must be calculated.

the function f_u can have arbitrarily large values, meaning that the number $f_u(\mathbf{z}) - f_l(\mathbf{z})$ can be arbitrarily large. In other words, the difference $f_u(\mathbf{z}) - f_l(\mathbf{z})$ can be more than the corresponding $\sum_{i=1}^{k} \text{nadir}_i^{Z_m} - \text{ideal}_i^{Z_m}$, which means that the difference has, in some sense, crossed a theoretical maximum for the difference. Therefore, the function f_c contains the minimum-operator in order to keep the certainty number in the range of $[0, 1]$.

Next, we demonstrate the values of the function f_c. Let us assume that for all vectors $\mathbf{z}$ in Δ_{k-1}° we have $f_l(\mathbf{z}) < f_s(\mathbf{z}) < f_u(\mathbf{z})$. In Figure 52, we have plotted function values $f_c(\mathbf{z})$, where Z_5 consists of vectors $(2,0)^T$, $(0.4,1.6)^T$, $(1.0,1.2)^T$, $(1.6,0.4)^T$ and $(0,2)^T$. The objective space $\mathbb{R}^2$ is the xy-plane in Figure 52 and the function values $f_c(\mathbf{z})$ on the simplex Δ_1 are plotted in parallel to z-axis. As it can be seen, if the function value $f_c(\mathbf{z})$ is close to 1, then the vector $f_a(\mathbf{z})$ can be interpreted as a Pareto optimal objective vector because the vector $f_a(\mathbf{z})$ is close to a known Pareto optimal objective vector. However, the value $f_c(\mathbf{z})$ should not be confused with a probability to be a Pareto optimal objective vector because it has nothing to do with probabilities.
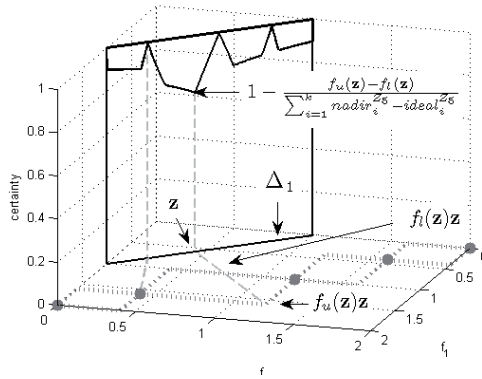
FIGURE 52   Function values $f_c(\mathbf{z})$ on simplex Δ_1. Note that the certainty value 1 is obtained at vectors on the relative interior of the simplex that map to Z_m by function f_a.

In implementing the function f_c to be used in the HBE method, one has to calculate values $f_l(\mathbf{z})$ and $f_u(\mathbf{z})$. Then one must calculate the difference and normalize it by vectors $\mathbf{ideal}^{Z_m}$ and $\mathbf{nadir}^{Z_m}$ and check whether the result is smaller than 1. Finally, one must calculate the difference to 1. In other words, to calculate function values $f_c(\mathbf{z})$ in the HBE method is not a computationally expensive operation. This means that the decision maker does not have to wait during the calculation. Therefore, we propose to use the following method in steps (3-4) in HBE.

Method 46.

(1) Calculate in step (3) of HBE certainty numbers as function values f_c at vectors $f_p(\mathbf{z}^{h_i})$, where the vector $f_p(\mathbf{z}^{h_i})$ is the vector in the approximation

corresponding to the sample vector $\mathbf{z}^{h_i}$ satisfying the aspiration and reservation levels given by the decision maker.

(2) If the decision maker wants to see the certainty numbers in step (4) of HBE, then present the calculated numbers.

Method 46 is rather simple computationally efficient method to be used in HBE. However, to clarify the first step, the function f_c must be applied to vectors $\mathbf{z}^{sh_i}$, which are the vectors on Δ_{k-1}° corresponding to $\mathbf{z}^{h_i}$.

Because of the simplicity, Method 46 has several drawbacks. First, the function f_c has been defined on Δ_{k-1}°. This means that estimations on the relative boundary of Δ_{k-1} cannot be calculated by the function f_c. In addition, the certainty numbers may not be close to one for vectors in the relative interior of $f_a(\Delta_{k-1})$ that are close to vectors, which belong to Z_m and the relative boundary of $f_a(\Delta_{k-1})$, see the values in Figure 52 at vectors close to vectors $(1,0)^T$ and $(0,1)^T$ in the simplex. However, it is possible in higher-dimensional objective spaces that for vectors in $f_a(\Delta_{k-1}^\circ)$ close to the relative boundary of $f_a(\Delta_{k-1}^\circ)$ there does not exist a vector in Z_m so that it is close to the vectors, see Appendix 3. Another drawback in the function f_c is that it does not take into account the value $f_s(\mathbf{z})$. In other words, the certainty value $f_c(\mathbf{z})$ is the same for all $f_s(\mathbf{z}) \in (f_l(\mathbf{z}), f_u(\mathbf{z}))$. This means that further investigations related to the function f_c must be done in the future studies. However, its simplicity for computationally expensive nonconvex MOO problems defends its position in the IHBE and HBE methods.

In this subsection, a computationally efficient function f_c was proposed for estimating how reliable an approximated Pareto optimal objective vector is, if it is interpreted as a Pareto optimal objective vector to the MOO problem. Then a method based on the function f_c and sample vectors generated on the approximated Pareto front that somehow corresponds to the decision maker's preferences were proposed for the HBE method.

5.3 Summary of Chapter 5

Chapter 5 has introduced functions f_l and f_u to the function f_s as a lower function and an upper function, respectively. They can be used, for example, in validating the approximation or in calculating how reliable an approximated Pareto optimal objective vector is. Next, we briefly summarize the main theorems of this chapter.

Theorem

40 In this theorem, it was shown how the functions f_l and f_u can be used in characterizing dominated vectors in the approximation. First, it was proved that $f_l(\mathbf{z}) \leq f_u(\mathbf{z})$ for all $\mathbf{z} \in \Delta_{k-1}^\circ$. Then dominance relations between vectors in Z_m and $f_a(\mathbf{z})$ were considered. If $f_s(\mathbf{z}) < f_l(\mathbf{z})$, then $f_a(\mathbf{z})$ dominates at least one of the vectors in Z_m. In addition, if $f_u(\mathbf{z}) < f_s(\mathbf{z})$, then at least

one of the vectors in Z_m dominates $f_a(\mathbf{z})$. However, if $f_l(\mathbf{z}) < f_s(\mathbf{z}) < f_u(\mathbf{z})$, then $f_a(\mathbf{z})$ does not dominate any of the vectors in Z_m and none of the vectors in Z_m dominates $f_a(\mathbf{z})$.

Proposition

41 Theorem 40 does not directly say when $f_a(\mathbf{z})$ belongs into Z_m. In Proposition 41, it has been shown that $f_l(\mathbf{z}) = f_s(\mathbf{z}) = f_u(\mathbf{z})$ if and only if $f_a(\mathbf{z}) \in Z_m$.

Theorem

42 Proposition 41 completes Theorem 40 to contain equalities instead of strict inequalities which were studied in this theorem.

44 For validation purposes, Theorem 44 introduces a condition through the continuity of the function f_s when dominated vectors on an open neighborhood of $\mathbf{z} \in \Delta_{k-1}^{\circ}$ are studied. It turned out that if f_s is continuous, then there exists a radius $r > 0$ so that $f_s(\mathbf{z}') < f_l(\mathbf{z}')$ or $f_u(\mathbf{z}') < f_s(\mathbf{z}')$, if $f_s(\mathbf{z}) < f_l(\mathbf{z})$ or $f_u(\mathbf{z}) < f_s(\mathbf{z})$, respectively, for all $\mathbf{z}' \in B(\mathbf{z}, r) \cap \Delta_{k-1}^{\circ}$, $\mathbf{z} \in \Delta_{k-1}^{\circ}$. This means that on an open neighborhood of $\mathbf{z}$ the corresponding vectors in the approximation are either dominating some of the vectors in Z_m or vice versa.

From the approximation model $\mathcal{AM}$'s point of view, several features must be mentioned.

$\mathcal{AM}$f11   Functions f_l and f_u can be considered as computationally efficient functions. They do not contain any computationally expensive operation to be evaluated.

      The simplex Δ_{k-1}° has also a significant role with functions f_l and f_u. Both functions are defined on Δ_{k-1}° and this is one of the reason why the explicit expressions of functions f_l and f_u are simple.

$\mathcal{AM}$f12   The surrogate function f_s has been fitted to data consisting of pairs $(\mathbf{z}^{s_i}, \lambda^i)$ in Definition 10. The data does not contain in itself any information of the Pareto dominance relation between objective vectors. As discussed in Subsection 5.2.1, functions f_l, f_u and f_s can be used in validating the approximation. Especially, through the function f_l, it can be detected, when a vector in the approximation dominates at least one of the vectors in Z_m. In other words, in some sense, we have provided answers to the goal $\mathcal{G}6$ set on page 20 to the approximation, where the approximation must be validated computationally efficiently.

$\mathcal{AM}$f13   In supporting the decision maker in the solution process of the MOO problem, the function f_c was proposed in this chapter. It is computationally efficient, and it calculates a number between [0,1] for each $\mathbf{z} \in \Delta_{k-1}^{\circ}$ to indicate how reliable the corresponding approximated Pareto optimal objective vector is. The number can be used in estimating whether the approximated Pareto optimal objective vector could be a Pareto optimal objective vector

to the MOO problem. This means that the decision maker can get significant information of the Pareto front. Thus we have answered the goal $\mathcal{G}9$ on page 20.

We have proposed in this chapter computational methods for identifying dominated vectors in the solution process produced by the HBE method. Next, we discuss how the dominated vectors are proposed to be interpreted in the solution process.

6 INTERPRETATION OF DOMINATED VECTORS AND SOME TOPOLOGICAL PROPERTIES OF THE APPROXIMATION

Normal vectors studied in Chapter 4 characterize dominated and nondominated vectors together with the functions f_l and f_u introduced in Chapter 5. In the IHBE method, they are used in validating the approximation. After the validation, the approximation should not, for example, contain vectors which dominate Pareto optimal objective vectors in the set Z_m. In other words, after having validated the approximation, it does not present, to the decision maker, unrealistic (infeasible) objective vectors, as it may happen for an invalid approximation. However, the nonconvexity of the MOO problem may introduce challenging situations discussed already in Subsection 3.3.3. In this chapter, we focus to the case, where philosophical discussion is required.

Let us assume exceptionally in this chapter that a Pareto front is available for visualizations and only for visualizations. Otherwise, it would be difficult to illustrate geometrically the cases discussed in this chapter. However, the assumption of having the exact Pareto front available is only for visualization, the theories presented in this chapter are not based on available Pareto fronts (the case is reintroduced in this chapter).

The IHBE method proposed in Chapter 3 allows the decision maker to freely study the objective space by the HBE method. The search is based on aspiration and reservation levels, which define a hyperbox in the objective space. Let us now assume that the decision maker is dealing with the IHBE method in the Pareto front environment. That is, the decision maker deals with the true Pareto front instead of an approximated one (having assumed that the front is available). We have assumed that the MOO problem can be nonconvex. The Pareto front could then be the dark surface illustrated in Figure 53. In informal terms, the Pareto front in the figure contains a *hole*. The hole is in the middle of the Pareto front, and through it we can partially see the coordinate planes of the objective space behind the Pareto front. In other words, the hole is in some sense a *missing piece of the Pareto front*. However, because the decision maker may search freely

124

the objective space with the HBE method, then a hyperbox (a gray dashed box) defined the by aspiration and reservation levels given by the decision maker can be located in the hole as we have demonstrated in Figure 54, where the Pareto front is the same as in Figure 53 but from a different perspective. This leads to the main concern of this chapter.
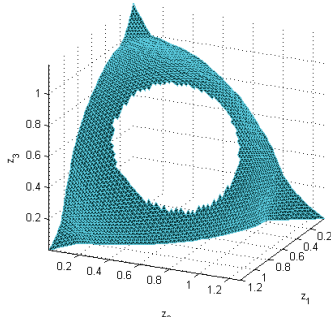
FIGURE 53   A Pareto front in objective space $\mathbb{R}^3$ with an obvious hole.
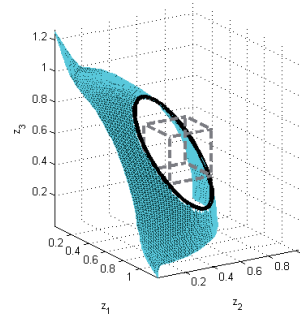
FIGURE 54   Occasionally, the decision maker may search for Pareto optimal objective vectors in the hole.

It is obvious that the decision maker cannot have Pareto optimal objective vectors he or she prefers because the hyperbox is in the hole (the Pareto front and the hyperbox as sets do not intersect). Now, the main concern is what to do if the hyperbox is in the hole. How to inform the decision maker about the situation? How to guide him or her away from the hole? What should the HBE method present for the decision maker in such a situation? The challenges in answering the questions emerge in the following paragraphs.

In the MCDM literature, Pareto fronts containing holes[1] are not that widely studied, especially from the decision maker's point of view. We believe that the reason for this is that many interactive methods created for nonconvex MOO problems present a Pareto optimal objective vector in some sense close to the decision maker's preferences. Then it would not matter whether or not there is a hole in the Pareto front. However, it is typically an algorithm that decides what is in some sense close to the preferences. In our opinion, this is not simple in practice. Also, selecting a close one by an algorithm can be counterproductive for the decision maker. Let us illustrate this as follows.

In Figure 54, we have added a black circle, which illustrates in some sense closest Pareto optimal objective vectors to the hyperbox. As it can be seen in the figure, there are several types of Pareto optimal objective vectors, which are close to the hyperbox and which we have classified in the following way:

---

[1]   We use the word hole because there is not yet a standard word for the concept in the literature. Words such as a disconnected set [43], hole [80, 154], branch [114] or non-convex or discontinuous frontier [191] are occasionally used.

Class 1. the first objective function value is better than the corresponding aspiration level,

Class 2. the second objective function value is better than the corresponding aspiration level,

Class 3. the third objective function value is better than the corresponding aspiration level,

Class 4. the first and the second objective function values are better than the corresponding aspiration levels,

Class 5. the first and the third objective function values are better than the corresponding aspiration levels and

Class 6. the second and the third objective function values are better than the corresponding aspiration levels.

In other words, if an algorithm presents a Pareto optimal objective vector to the decision maker, then the algorithm probably selects one vector, which belongs to some of the classes introduced above. This is the main concern discussed in this chapter because the decision maker may have different opinions about the vector selected by the algorithm. For example, the algorithm may have presented a Pareto optimal objective vector, which belongs to the first class, although the decision maker might have liked (or was hoping/preferring) to have a vector, which belongs to the sixth class because the second and the third objective functions can be more important for the decision maker. Because the algorithm has presented a Pareto optimal objective vector, which belongs to the first class, then the decision maker may think that there does not exist Pareto optimal objective vectors, which belong to the sixth class. Note that it is possible in this case to have Pareto optimal objective vectors in the sixth class. In other words, the algorithm can in the worst case mislead the decision maker and this is something we do not want to happen with the IHBE and HBE methods.

In order to avoid misleading the decision maker, presenting vectors in each class may sound like a good idea. Nevertheless, the number of classes increases exponentially when the number of objective functions increases. This means that, in a MOO problem consisting of four objective functions, the decision maker may receive up to $2^4 - 1 = 15$ different types of Pareto optimal objective vectors close to the hyperbox. This could exceed the decision maker's capability to process information as discussed in Subsection 3.1.1.2. Another way to avoid providing misleading information is to visualize the hole as in Figure 54. However, to "see" a hole in the Pareto front of a MOO problem consisting of four objective functions is not necessarily that straightforward because visualization is no longer as simple as in Figures 53 and 54. Note that with the help of the visualization, the decision maker can "see" alternative Pareto optimal objective vectors close to the hyperbox.

We have seen that "getting out of the hole or seeing a hole", that is, presenting Pareto optimal objective vectors close to the hyperbox in a hole can be challenging. Note that "getting out of the hole" is important, because a compromise solution should be found and such a vector does not exist in a hole in the

Pareto front. However, if the analyst introduces the concept of holes in the Pareto front for the decision maker, then we believe that the decision maker receives more support for and understanding about the (approximated) Pareto front for the creation of an overall understanding of available (approximated) Pareto optimal objective vectors. The concept of holes in the Pareto front then explains why the decision maker cannot have the (approximated) Pareto optimal objective vectors he or she is asking for if it turned out that the hyperbox is in a hole. In addition, the decision maker may realize that it is possible to find different kinds of Pareto optimal objective vectors which in some sense are close to the preferences instead of presenting a single vector selected by some algorithm. In other words, the concept of holes in the Pareto front must be clear for the decision maker. Then he or she can benefit of that information in the solution process. The decision maker understands that the hyperbox is in the middle of the Pareto front, where no Pareto optimal objective vectors exist but close to the hyperbox different kinds of Pareto optimal objective vectors can be found. In addition, depending on the decision maker's interest to mathematical details about the MOO problem, the analyst can clarify to the decision maker that the Pareto front may contain holes because the MOO problem is nonconvex. It must be noted that in continuous convex MOO problems there are no holes in the Pareto front, and this is guaranteed by the convexity of the set S and convexities of the objective functions.

In the IHBE and HBE methods, a hyperbox in a hole is understood in the following way. As illustrated, there can be different kinds of Pareto optimal objective vectors close to a hyperbox in a hole in the Pareto front. In addition, it can be challenging to present all of them to the decision maker. Therefore, (approximated) Pareto optimal objective vectors close to the hyperbox in a hole are not calculated in the IHBE and HBE methods. Let us remind ourselves that the IHBE method is based on the HBE method and the HBE method is created for answering a query in a form of a question:

I am interested in these kinds of (approximated) Pareto optimal objective vectors. Do those vectors exist?

If a hyperbox defined by the aspiration and reservation levels is in a hole, then, by using the concept of holes in the Pareto front, the HBE method can provide an explanation about why the decision maker cannot have (approximated) Pareto optimal objective vectors. For example, the explanation can be "They do not exist, because the particular region you are interested in is in an approximated hole in the Pareto front" (an approximated hole in the Pareto front will be clarified in the next section, and the answer is naturally possible only after the concept of holes has been computationally implemented later in this chapter).

The answer by the HBE method is honest and explains why the decision maker did not receive any (approximated) Pareto optimal objective vectors. In other words, it is not forcing the decision maker to some direction defined by some algorithm. It will be up to the decision maker to determine which part of the objective space to investigate next. In this way, the IHBE and HBE methods

can be directed to the regions that are of interest to the decision maker. This is the advantage of having a free-search -feature in the creation of an overall understanding of available (approximated) Pareto optimal objective vectors, as discussed in Subsection 3.1.2.3. However, in case the decision maker has no idea where to continue in the objective space, the vectors in the set Z_m can be presented to him or her as proposed in Subsection 3.3.4.

The result of the characterization of approximated holes in the Pareto front is significant in the creation of an overall understanding of available (approximated) Pareto optimal objective vectors, i.e. the decision maker can be informed about the region, where (approximated) Pareto optimal objective vectors do not exist and where the preferences are neither optimistic nor pessimistic. The decision maker can thus save time for some other regions where (approximated) Pareto optimal objective vectors may exist.

The rest of this chapter has been organized as follows. First, we geometrically illustrate the way approximated holes in the Pareto front can be identified by the approximation created in Chapter 2. The dominated vectors in the approximated Pareto front are the key elements in the characterization. This takes place in Section 6.1, where approximated holes are defined geometrically. Then, a formal definition for approximated holes is defined mathematically in Section 6.2. After this, the characterization of approximated holes in practice is discussed in Section 6.3. Thus we introduce computational counterparts for the mathematical and theoretical characterizations for the approximated holes. Section 6.4 slightly differs by its contents from the other sections of this chapter. We discuss how the analyst can use topological properties of the approximation in validating the approximation. Finally, this chapter is summarized in Section 6.5.

6.1 Characterization of Holes in the Pareto Front Using the Approximation

In nonconvex MOO problems, Pareto fronts can be disconnected, that is, they can consist of separate sets as in Figure 55 (black curves). In Subsection 3.3.3, we have briefly discussed a case in which the hyperbox defined by the aspiration and reservation levels can be located in the objective space, as it has been depicted in Figure 55 (gray box). The decision maker may wonder why no approximated Pareto optimal objective vectors inside the hyperbox can be obtained while very close to the hyperbox they can be found if the aspiration and reservation levels are slightly changed. However, if a visualization (similar to that of Figure 55) of the Pareto front was available for the decision maker, the decision maker could immediately realize why he or she cannot get Pareto optimal objective vectors inside the hyperbox. He or she can see that the hyperbox is in some sense in the middle of the Pareto front, where Pareto optimal objective vectors do not exist (the hyperbox is between the Pareto optimal objective vectors $\mathbf{z}^1$ and $\mathbf{z}^2$ in the figure, and between them no other Pareto optimal objective vectors exist).

Informally stated, the hyperbox is in a *hole* in the Pareto front.
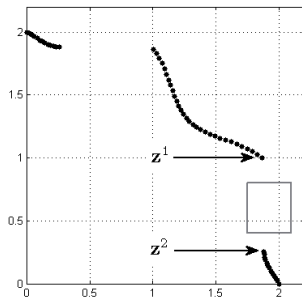
FIGURE 55 In nonconvex MOO prob-
lems, the Pareto front can
consist of several separated
sets and, occasionally, the
decision maker may study
a region, defined by aspira-
tion and reservation levels,
where   (approximated)
Pareto optimal objective
vectors do not exist.
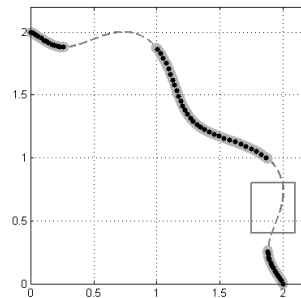
FIGURE 56 The decision maker is sup-
ported by the information
that he or she is investi-
gating a hole which can be
characterized by dominated
vectors in the approxima-
tion.

To detect computationally that the hyperbox is in a hole can be done by us-
ing the approximation. We demonstrate this as follows. In Figure 56, we have
added the approximation in the figure (gray dashed/bold curves, where bold
curves represent nondominated vectors in the approximation and dashed repre-
sent dominated vectors). Note that the nonconvexity of the MOO problem has
forced the approximation, which is a "nice" smooth curve, into a "shape", which
contains dominated vectors. As shown by Figure 56, the approximated Pareto
front intersects the hyperbox, and the vectors in the intersection are dominated
in the approximation. Note that the dominated vectors and especially the ones
in the intersection can be detected to be dominated by the methods introduced in
Chapters 4 and 5. Now, the dominated vectors are used in Method 16 to indicate
that the hyperbox is in an *approximated hole in the Pareto front*. We use the term an
approximated hole because the identification of the hole is based on information
the approximation is providing with the help of the dominated vectors. Note that
the dominated vectors in the hyperbox can be found by using Method 16 created
for calculating vectors on the approximated Pareto front.

Indeed, dominated vectors in the approximation can be used for indicating
approximated holes in the Pareto front. However, here we have only a geomet-
rical representation illustrating how holes in the Pareto front can be detected. To
do this computationally, we first have to set up a mathematical basis for a concept
of a hole in the Pareto front. For that, we have to determine what a hole actually
is.

FIGURE 57   A clear hole in the wall.

FIGURE 58   A piece of the wall is missing.

In this thesis, a hole is defined mathematically based on our example about *"hole in the wall"*. In Figure 57, we have a wall (the gray rectangle) with a hole (the white disc). Intuitively, a hole can be considered a piece that is missing from the wall. Through that hole we can see behind the wall (the white paper in this case). To understand better what characterizes a hole, let us consider the wall in Figure 58. There we have a piece missing from the wall. We believe that most of the people would not probably consider the missing piece a hole because the wall does not surround it as it does in Figure 57. In other words and semi-formally, to consider an object $\mathcal{O}_1$ to be a hole in an object $\mathcal{O}_2$ (a part of a plane) requires that

1. elements in $\mathcal{O}_2$ do not belong to $\mathcal{O}_1$ and elements in $\mathcal{O}_1$ do not belong to $\mathcal{O}_2$,
2. it must be possible to go from one side of $\mathcal{O}_2$ to the other through $\mathcal{O}_1$ and
3. the object $\mathcal{O}_2$ must surround $\mathcal{O}_1$ everywhere.

The term *surrounding* means that there exists a clear *border*, which separates the object $\mathcal{O}_1$ from the object $\mathcal{O}_2$ and no matter which point and direction in the object $\mathcal{O}_1$ have been selected, a point in the border will be hit eventually when moving from the point selected to the direction selected.

The semi-formal definition above based on the objects $\mathcal{O}_1$ and $\mathcal{O}_2$ for a hole is the basis for an approximated hole in the Pareto front. However, to define an approximated hole directly in the Pareto front can be challenging because the Pareto front can be a twisted or a bended hypersurface making it difficult to say which of the vectors in the objective space actually belong to the hole. However, the above semi-formal definition for approximated holes in a plane (the wall may be considered a two-dimensional plane) can be interpreted to mean that vectors in $\mathcal{O}_1$ inside the object $\mathcal{O}_2$ are hole vectors. This is an important observation because then the "hole in the wall" example can be adapted to the simplex Δ_{k-1}, which is one of the elements in the basis of the approximation. We claim that the "hole in the wall" example is an intuitive demonstration of a hole in real life. This intuition can now be adapted to the approximated Pareto front, as we will demonstrate in the next section.

6.2 On Dominated Subsets of the Set $f_a(\Delta_{k-1})$ and Approximated Holes in the Pareto Front

To start the discussion on approximated holes in the Pareto front, we propose the following set as a basis.

Definition 47. A vector $\mathbf{z} \in \Delta_{k-1}$ is called a *hole vector* , if

$$\mathbf{z} \in \operatorname{rel int} \left(\mathrm{CH}(\Delta_{k-1}^*) \cap (\Delta_{k-1}^*)^C \right) =: \Delta_{k-1}^h,$$

where Δ_{k-1}^* is the set of vectors for which the corresponding vector in the approximation is nondominated. The vector $f_a(\mathbf{z})$ is called a hole vector if $\mathbf{z}$ is a hole vector.

The vector $\mathbf{z} \in \Delta_{k-1}$ is called a *boundary hole vector* if $f_a(\mathbf{z})$ is not a hole vector, but

$$\mathbf{z} \in \overline{\mathrm{CH}(\Delta_{k-1}^*) \cap (\Delta_{k-1}^*)^C}.$$

The set of boundary hole vectors of Δ_{k-1}^* is denoted by $\Delta_{k-1}^{\partial h}$. In addition, the vector $f_a(\mathbf{z})$ is called a boundary hole vector if $\mathbf{z}$ is a boundary hole vector. If $\Delta_{k-1}^h \cup \Delta_{k-1}^{\partial h} = \varnothing$ and f_a is continuous on Δ_{k-1}^*, then $f_a(\Delta_{k-1}^*)$ $(= f_a(\Delta_{k-1})^*)$ is said to be *continuous*.

Definition 47 for an approximated hole in the Pareto front may seem cryptic at first. However, it follows the analogy of the "*hole in the wall*" example, and we demonstrate this as follows. Let us consider the set Δ_{k-1}^* as the wall (the light gray part on the left hand side in Figure 59). As we can see, there is a hole in Δ_{k-1}^* caused by the dominated vectors in the approximation (the white faces of triangles in the middle picture). Now, the (separated) set(s) of dominated vectors in the approximation indicate an approximated hole in the Pareto front as can be seen on the right hand side in Figure 59. In other words, if the dominated vectors in the approximation are removed, then we have a hole in the approximation. However, the dominated vectors in the approximation are not removed in practice. Nevertheless, if a hyperbox contains such vectors, then the hyperbox is in an approximated hole in the Pareto front, as discussed in the previous section. Note that this example is in a three-dimensional objective space. Definition 47 is valid for any k-dimensional objective space ($k \geq 2$).

The border, which was introduced in the "*hole in the wall*" example and which separates the hole from the approximated Pareto front is, naturally, the set $\Delta_{k-1}^{\partial h}$ and, correspondingly in the approximated Pareto front, the set $f_a(\Delta_{k-1}^{\partial h})$. Note that a boundary hole vector in $f_a(\Delta_{k-1})$ can be nondominated or dominated. Next, we prove that the set $\Delta_{k-1}^{\partial h}$ surrounds holes, i.e. (separated) set(s) Δ_{k-1}^h. In other words, we show that no matter which hole vector $\mathbf{z}^h$ and direction $\mathbf{z}^d \in \{\mathbf{z}' : \mathbf{z}' \neq \mathbf{0}, \mathbf{z}' = \mathbf{z}'' - (1/k, \dots, 1/k), \mathbf{z}'' \in \Delta_{k-1}\}$ have been selected, then we will eventually end up at the boundary $\Delta_{k-1}^{\partial h}$ when we start to move from $\mathbf{z}^h$ to the direction $\mathbf{z}^d$.
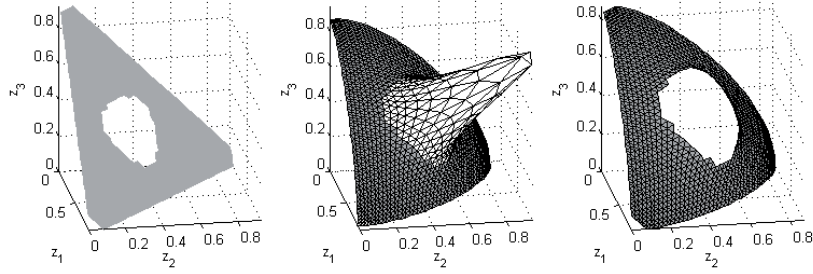
FIGURE 59 A "wall" on the simplex representing the nondominated vectors in the approximation. The dominated vectors in the approximation are subsequently interpreted as an approximated hole in the Pareto front.

Theorem 48. The set $\Delta_{k-1}^{\partial h}$ surrounds the set Δ_{k-1}^{h}, i.e. for all vectors $\mathbf{z}^h \in \Delta_{k-1}^{\partial h}$ and directions $\mathbf{z}^d$ in the simplex, there exists a scalar $\lambda_d > 0$ so that $\mathbf{z}^h + \lambda_d \mathbf{z}^d \in \Delta_{k-1}^{\partial h}$ and the vector $\mathbf{z}^h + \lambda \mathbf{z}^d$ is a hole vector for all $\lambda < \lambda_d$.

Proof. It is clear that a radius $r > 0$ exists so that every vector on line $\mathbf{z}^h + r\mathbf{z}^d$ is a hole vector, because $\mathbf{z}^h$ belongs initially to a relatively open set of the simplex. Now, for some $\lambda_d > 0$ the vector $\mathbf{z}^h + \lambda_d \mathbf{z}^d$ belongs to $\Delta_{k-1}^{\partial h}$, and this can be seen as follows. Because the set Δ_{k-1} is compact, then there is a maximal $\lambda_m > 0$ so that for some i the coordinate value $(\mathbf{z}^h + \lambda_m \mathbf{z}^d)_i$ is equal to zero and the vector $\mathbf{z}^h + \lambda_m \mathbf{z}^d$ belongs to Δ_{k-1}. If for all $\lambda_d < \lambda_m$ the vector $\mathbf{z}^h + \lambda_d \mathbf{z}^d$ does not belong to $\Delta_{k-1}^{\partial h}$, then the vector $\mathbf{z}^h + \lambda_m \mathbf{z}^d$ does because it belongs to the relative boundary of the set rel int Δ_{k-1}. $\qquad\square$

In other words, Theorem 48 proves the semi-formal "surrounding" requirement between objects $\mathcal{O}_1$ and $\mathcal{O}_2$. Next, we present illustrative examples of holes, and after that we consider the theoretical conditions which characterize hole vectors.

Example 49. Hole vectors are defined on the simplex, therefore in the following examples we only illustrate sets on the simplex.

1. Let
$$\Delta_2^* = \{\mathbf{z} \in \Delta_2 : \text{ there exists } i \text{ so that } z_i < 0.1\}.$$

In Figure 60, the set Δ_2^* is depicted as the black part on the simplex. The set Δ_2^h has been colored dark gray, and the set $\Delta_2^{\partial h}$ consists of three dashed white lines. Now, sets Δ_2^h and $\Delta_2^{\partial h}$ have an intuitive connection to the "*hole in the wall*" example. Note that the set Δ_2^* also surrounds the hole. However, it is possible that the set Δ_2^* does not surround the hole. We demonstrate this next.

2. Let Δ_2^* be consisting of vectors of the form $0 \leq z_3 \leq 0.5$ and $0.4 \leq z_1 \leq 0.5$ or $0.4 \leq z_2 \leq 0.5$. In Figure 61, the set Δ_2^* is represented by the black part

on the simplex. The set Δ_2^h has been colored dark gray, and the set $\Delta_2^{\partial h}$ is represented by the dashed light gray lines. Note that Δ_2^* does not surround the set Δ_2^h, but $\Delta_2^{\partial h}$ does.

3. If Δ_{k-1}^* is convex, then sets Δ_{k-1}^h and $\Delta_{k-1}^{\partial h}$ are empty. In other words, the set $f_a(\Delta_{k-1}^*)$ is continuous if Δ_{k-1}^* is convex and the function f_s is continuous. Thus, for example, the set

$$\{\mathbf{z} : \|\mathbf{z}\| = 1, z_i \geq 0\},$$

considered as $f_a(\Delta_{k-1}^*)$, is continuous, which is trivial.

4. Let Δ_2^* consist of only vectors of rational coordinate values. Then the set Δ_2^h is empty, but $\Delta_2^{\partial h} = \Delta_2$.
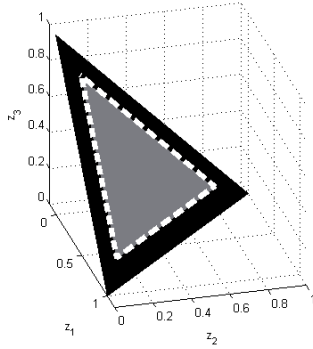
FIGURE 60   A simple example of the simplex for the *a hole in the wall* example.
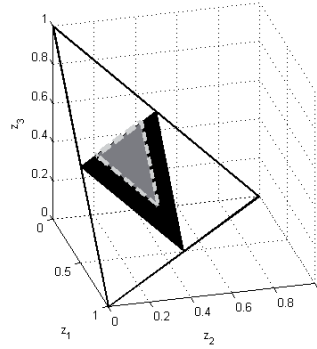
FIGURE 61   Set $\Delta_2^{\partial h}$ artificially separates the hole vectors from the vectors for which the corresponding vectors in the approximation are nondominated.

Example 49.2. demonstrates Definition 47 in an interesting way. It shows that the set of hole vectors is only defined "inside" the convex hull of Δ_{k-1}^* (the dark gray part is "inside" the black part, i.e. the set Δ_{k-1}^*) and not "outside" the convex hull of Δ_{k-1}^*, see Figures 60 and 61. In other words, Definition 47 does not characterize vectors close to the corners of the simplex (the standard unit vectors) in Example 49.2. as hole vectors, as seen in Figure 61. The corresponding vectors in the approximation are dominated, but they are not hole vectors because they are not "surrounded" by nondominated vectors. This is a consequence of having the set $\mathrm{CH}(\Delta_{k-1}^*) \cap (\Delta_{k-1}^*)^C$ in Definition 47. On the other hand, in Figure 58, we came up to a conclusion that the missing piece in the wall is not a hole, because the wall does not surround it. However, one can see that the missing piece, i.e. the hole in the approximated Pareto front in Example 49.2., corresponds to this missing piece in the wall. This is a challenging situation for us because it is difficult to say whether or not the set Δ_{k-1}^h in Example 49.2. is a hole "geometrically"

or a by-product or a drawback of taking the convex hull of the set Δ_{k-1}^* in Definition 47. However, there is a clear explanation for it being considered as a hole. In Example 49.2., the situation is a natural consequence of taking the convex hull of Δ_{k-1}^*. Note that the nondominated vectors in the set Δ_{k-1}^* "surround" the hole vectors because the set Δ_{k-1}^h is a subset of the convex hull of the set Δ_{k-1}^*. In other words, the set Δ_{k-1}^h is "inside" the convex hull of the set Δ_{k-1}^*, and the set Δ_{k-1}^h does not contain vectors, which do not belong to the convex hull of the set Δ_{k-1}^*. Therefore, the hole in Example 49.2. can, indeed, be interpreted as a hole.

So far, we have geometrically demonstrated holes in the simplex and the corresponding vectors in the approximation. Next, we show how the characterization of hole vectors can be done theoretically based on the theories already presented in this thesis.

Theorem 50. Let $\mathbf{z} \in \operatorname{rel\,int} \operatorname{CH}(\Delta_{k-1}^*)$ and let the function f_s be continuously differentiable. If the normal vector $f_a(\mathbf{z})^\perp$ does not belong to $\mathbb{R}_-^k \cup \mathbb{R}_+^k$ or $f_u(\mathbf{z}) < f_s(\mathbf{z})$, then $\mathbf{z}$ is a hole vector.

Proof. To show that $\mathbf{z}$ is a hole vector, we should find a radius $r_{\mathbf{z}} > 0$ so that for each $\mathbf{z}' \in \operatorname{rel\,int} \operatorname{CH}(\Delta_{k-1}^*)$, $\|\mathbf{z}' - \mathbf{z}\| < r_{\mathbf{z}}$, vector $f_a(\mathbf{z}')$ is dominated in $f_a(\Delta_{k-1})$. Theorem 44 guarantees that such a radius exists if we have $f_u(\mathbf{z}) < f_s(\mathbf{z})$. A radius for the condition of normal vector follows similarly from the proof of Theorem 32 applied to Theorem 27. $\qquad\square$

As said, Theorem 50 is a theoretical way to identify hole vectors. However, it is a basis for a computational method to be used in the HBE method when investigating whether or not a hyperbox is located in an approximated hole in the Pareto front.

6.3 Implementing the Concept of a Hole

Calculating normal vectors $f_a(\mathbf{z})^\perp$ and function values $f_u(\mathbf{z})$ can be done computationally efficiently. As proved in Theorem 50, they can be used in characterizing hole vectors. Naturally, Theorem 50 would be the one to be implemented for the HBE method. However, there are two challenges in Theorem 50 from a practical point of view. Firstly, the set Δ_{k-1}^* should be known. Secondly, the vector $\mathbf{z}$ should belong to the relative interior of the set $\operatorname{CH}(\Delta_{k-1}^*) \cap (\Delta_{k-1}^*)^{\mathsf{C}}$. Once the first challenge has been solved, the second one can be easily checked.

Instead of using $\operatorname{CH}\Delta_{k-1}^*$ in the HBE method, we propose to use $\operatorname{CH} Z_m^s$ ($Z_m^s = f_{Z_m}^p(Z_m)$) and the function $f_{Z_m}^p$ was used to map vectors in Z_m into the simplex). Let us assume that the approximated Pareto front $f_a(\Delta_{k-1})$ has been validated, i.e. vectors in Z_m are nondominated in $f_a(\Delta_{k-1})$. Then, $Z_m^s \subset \Delta_{k-1}^*$, which means that $\operatorname{rel\,int} \operatorname{CH} Z_m^s \subset \operatorname{rel\,int} \operatorname{CH} \Delta_{k-1}^*$. Now, if we replace $\operatorname{rel\,int} \operatorname{CH} \Delta_{k-1}^*$ by $\operatorname{rel\,int} \operatorname{CH} Z_m^s$, then the vector $\mathbf{z}$ in Theorem 50 is a hole vector (small changes must be done in the proof of Theorem 50). However, $\operatorname{rel\,int} \operatorname{CH} Z_m^s$ covers less

than rel int CH Δ_{k-1}^* on the simplex, so we are not necessarily able to characterize as many hole vectors as Theorem 50 does.

The reason for the use of CH Z_m^s is obvious, because Z_m is finite. Then, to calculate whether or not $\mathbf{z}$ belongs to CH Z_m^s is much simpler, as we show next.

Let us consider the following linear minimization problem at $\mathbf{z} \in \Delta_{k-1}$.

$$
\begin{aligned}
\min_{\alpha_1,\ldots,\alpha_m,t} \quad & t \\
\text{subject to} \quad & \sum_{i=1}^{m} \alpha_i z_j^{s_i} - z_j \leq t, \quad j = 1,\ldots,k, \\
& z_j - \sum_{i=1}^{m} \alpha_i z_j^{s_i} \leq t, \quad j = 1,\ldots,k, \\
& \sum_{i=1}^{m} \alpha_i = 1, \quad \alpha_i \geq 0, \quad i = 1,\ldots,m.
\end{aligned}
\tag{36}
$$

Let t^* be the minimum value of the minimization problem (36) satisfying all the constraints. Then we have

$$
\left| \sum_{i=1}^{m} \alpha_i z_j^{s_i} - z_j \right| \leq t^*, \quad j = 1,\ldots,k.
$$

In addition, if $t^* = 0$ and all the constraints hold, then we have

$$
\mathbf{z} = \sum_{i=1}^{m} \alpha_i \mathbf{z}^{s_i}, \quad \sum_{i=1}^{m} \alpha_i = 1, \quad \alpha_i \geq 0, \quad i = 1,\ldots,m,
$$

which means that the vector $\mathbf{z}$ is a convex combination of vectors in Z_m^s. In other words, if $t^* = 0$ and the constraints hold in (36), then $\mathbf{z} \in \mathrm{CH}\, Z_m^s$. On the other hand, if $t^* \neq 0$ and all the constraints hold, then $\mathbf{z} \notin \mathrm{CH}\, Z_m^s$, because there exists always at least one j so that $\left| \sum_{i=1}^{m} \alpha_i z_j^{s_i} - z_j \right| > 0$ no matter what α_i:s are as far as conditions $\sum_{i=1}^{m} \alpha_i = 1$, $\alpha_i \geq 0$, $i = 1,\ldots,m$, hold. This means that we cannot represent $\mathbf{z}$ as a convex combination of vectors $\mathbf{z}^{s_i}$. Thus, the minimal t to the linear optimization problem (36) determines whether or not $\mathbf{z}$ belongs to CH Z_m^s. This means that we have solved the first challenge related to the implementation of Theorem 50 in the HBE method.

To solve the second challenge, we should be able to identify whether $\mathbf{z}$ belongs to rel int CH Z_m^s or not. Note that the linear minimization problem (36) does not characterize this because it is possible that the vector in $\overline{\mathrm{CH}\, Z_m^s}$ is a boundary hole vector. However, to identify that a vector belongs to the relative interior, we propose the following procedure for the HBE method.

Let $\varepsilon > 0$ be sufficiently small. If each of the vectors

$$
\varepsilon(\mathbf{e}^i - (1/k,\ldots,1/k)^{\mathrm{T}}) + \mathbf{z}, \quad i = 1,\ldots,k
\tag{37}
$$

belongs to CH Z_m^s, then $\mathbf{z} \in \mathrm{rel\,int}\,\mathrm{CH}\, Z_m^s$, because the relative interior of the convex hull of the vectors is an open neighborhood of $\mathbf{z}$. We have geometrically
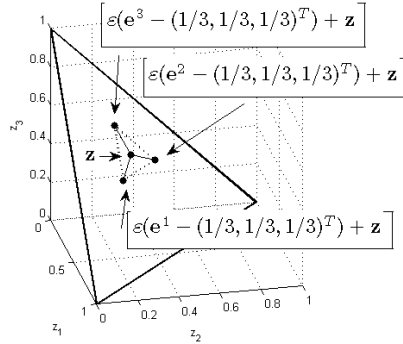
FIGURE 62　A simple method to characterize a vector as an interior vector.

sketched this in Figure 62. Thus, we have a proposal to the second problem related to the implementation of hole vectors for the HBE method.

　　Now, we are ready to propose a method to be used in the HBE method to characterize hole vectors. Then the method can be used in investigating whether a hyperbox is located in an approximate hole in the Pareto front.

Method 51. Let $\mathbf{z}$ be a dominated vector in the approximation identified by the normal vector or functions f_u and f_s.

(1) Solve the linear minimization problem (36) at $\mathbf{z}$.

　(1.1) If solution t^* to the problem is positive, then the vector $\mathbf{z}$ is not a hole vector by the approximation.

(2) Investigate whether all vectors in (37) with a sufficiently small $\varepsilon > 0$ are hole vectors (check first that they are dominated, and then use the linear optimization problem (36) for checking that they belong to the convex hull of CH Z_m^s).

　(2.1) If all of them belong to CH Z_m^s, then the vector $\mathbf{z}$ is a hole vector.
　(2.2) If some of them do not belong to CH Z_m^s, then the vector $\mathbf{z}$ is not necessarily a hole vector by the approximation.

　　Method 51 requires an external solver for the linear minimization problem (36). However, problem (36) does not contain any special requirements. Therefore, any commercial solver such as IBM ILOG CPLEX [3], Optimization toolbox™ [4] or noncommercial ones like GLPK [1] or GNU Octave [2] can be used. However, what comes to the computational cost of Method 51, it can be said that the cost does not increase that significantly, even though problem (36) must be solved $k + 1$ times for a single vector. Note that the problem is a linear optimization problem, which can be considered fast solvable.

　　In Subsections 3.3.1 and 3.3.2, we have proposed simple methods to be used in identifying optimistic and pessimistic hyperboxes by using the sample vectors $f_p(\mathbf{z}^{h_i})$ calculated in step (3) of HBE (the sample vectors $f_p(\mathbf{z}^{h_i})$ in the approximation correspond to the aspiration and reservation levels given by the decision

maker). Then the identification of whether or not a hyperbox is in an approximated hole in the Pareto front can be done similarly.

Method 52.

   (1) If every sample vector $f_p(\mathbf{z}^{h_i})$ is a hole vector identified by Method 51 and some of these vectors belong to the hyperbox defined by the aspiration and reservation levels, then the decision maker may be informed about the hyperbox which is in a hole.

   (2) If some of the vectors $f_p(\mathbf{z}^{h_i})$ are hole vectors identified by Method 51 and some of them belong to the hyperbox defined by the aspiration and reservation levels, then inform the decision maker about the hyperbox which is partially in a hole.

Next, we briefly discuss Method 52 and how it can be used in the HBE method. Let us remind that the HBE method provides an answer to whether or not the decision maker can have the (approximated) Pareto optimal objective vector he or she is interested in. If it happens that the hyperbox is in an approximated hole in the Pareto front, then Method 52 explains the situation for the decision maker. This is the first case in Method 52. Now, we need to clarify that Method 52 investigates whether or not a hyperbox is in an approximated hole in the Pareto front. The method does not investigate whether or not the hyperbox contains nondominated vectors in the approximation. In other words, the second case in Method 52 informs the decision maker about the hyperbox which is partially in an approximated hole in the Pareto front. It may be possible that the hyperbox contains both nondominated and hole vectors, i.e. that the hyperbox is around the border of an approximated hole in the Pareto front. Note that the HBE method presents at least the nondominated vectors to the decision maker, but now his or her preferences can be clarified by informing about the approximated hole in the Pareto front close to the nondominated vectors.

Method 52 for the approximated holes in the Pareto front is rather simple, but it does not necessarily characterize all holes, which means that further investigations must be carried out in future research. An undetected hole may exist, because Method 52 is based on normal vectors and the condition $f_u(\mathbf{z}) < f_s(\mathbf{z})$. As known, these conditions do not detect every dominated vector in the approximated Pareto front. However, for the future studies, we have illustrated in Figures 63 and 64 several different cases in which detection of hole vectors must be improved in order to find more details of the approximated hole in the Pareto front. In the figures, we have used the following coloring theme. Vectors in Z_m are depicted as black dots, and the approximated Pareto front is a black curve. Dominated vectors characterized by normal vectors are bold dark gray curves on the black curves. The bold light gray curves represent dominated vectors detected by the condition $f_u(\mathbf{z}) < f_s(\mathbf{z})$. Dotted black lines illustrate the upper bounds to the function f_s, i.e. they are the approximated Pareto fronts defined by the function f_u as discussed in the introduction of Chapter 5. As it can be seen, some of the dominated vectors in the approximation do not belong to the bold

curves as they should because they are dominated vectors in the approximation and "inside" the convex hull of nondominated vectors.
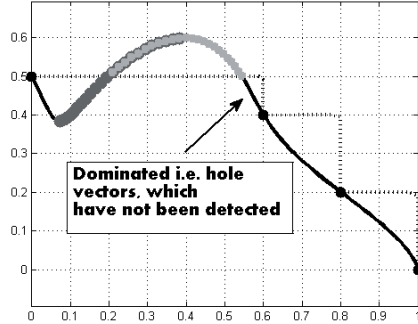
FIGURE 63 Normal vectors and condition $f_u(\mathbf{z}) < f_s(\mathbf{z})$ characterize hole vectors. In future studies, more attention must be paid to investigate whether or not the hole is larger than the methods used indicate.
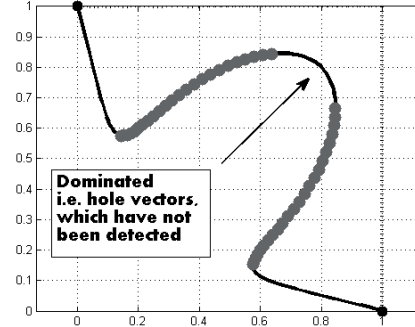
FIGURE 64 Occasionally, it may happen that there are dominated vectors between hole vectors even though the methods in the HBE method indicate something else. In future studies, this must be investigated in more detail.

To conclude this section, we have proposed Method 52 for the HBE method in order to explain for the decision maker the reason of having not received (approximated) Pareto optimal objective vectors in a special case. It indicates that the decision maker is investigating a region in the objective space that is interpreted as an approximated hole in the Pareto front. There, naturally, are no (approximated) Pareto optimal objective vectors. After having this information, the decision maker can use his or her valuable time for other regions in the objective space, where (approximated) Pareto optimal objective vectors may exist.

6.4   Convexity and Connectedness of Sets $f_a(\Delta_{k-1}^*)$ and Δ_{k-1}^*

A convex hull of the set Δ_{k-1}^* was used in defining hole vectors. However, we have not studied convexity of the set Δ_{k-1}^* and, especially, convexity of the function f_s in this thesis. In this section, we investigate these aspects briefly. In this way, the analyst can obtain useful information of the approximation and its nature in the objective space in relation to the MOO problem. This information can be used in selecting and validating the function f_s in the function f_a. Note that the selection of the approximation through the function f_s can also be based on topological features discussed in this section.

In this section, the studies are separated as follows. First, in Subsection 6.4.1, we investigate convexity and concavity of the function f_s. Then, in Subsection 6.4.2,

connectedness of sets Δ^*_{k-1} and $f_a(\Delta^*_{k-1})$ is briefly discussed. Finally, in Subsection 6.4.3, the results are discussed from the practical point of view.

6.4.1 On Convexity and Concavity of Function f_s

Here, we focus on the function f_s because the function f_a is a vector-valued function for which defining convexity is not necessarily that straightforward. In addition, a convex set $f_a(\Delta^*_{k-1})$ is not interesting for further investigations regarding its topological features because trivially a convex nondominated set in the objective space is a convex subset of a hyperplane. Instead, here we investigate a specific set we call a *relative epigraph* in which the function f_s has an essential role to play. Then, we can consider "convex" functions f_a on Δ^*_{k-1}.

Definition 53. Set $f_a(\Delta^*_{k-1})$ is said to be *relatively convex*, if the relative epigraph

$$\mathrm{epi}(f_a, \Delta^*_{k-1}) := \{\mathbf{z}' \in \mathbb{R}^k : \quad \mathbf{z}' = \mathbf{ideal}^{Z_m} + \alpha\mathbf{z}, \quad 0 < f_s(\mathbf{z}) \leq \alpha, \mathbf{z} \in \Delta^*_{k-1}\}$$

is convex. Correspondingly, the set $f_a(\Delta^*_{k-1})$ is said to be *relatively concave*, if a set called a *relative hypograph*

$$\mathrm{hyp}(f_a, \Delta^*_{k-1}) := \{\mathbf{z}' \in \mathbb{R}^k : \quad \mathbf{z}' = \mathbf{ideal}^{Z_m} + \alpha\mathbf{z}, \quad 0 < \alpha \leq f_s(\mathbf{z}), \mathbf{z} \in \Delta^*_{k-1}\}$$

is convex.

Geometrically, sets $\mathrm{epi}(f_a, \Delta^*_{k-1})$ and $\mathrm{hyp}(f_a, \Delta^*_{k-1})$ lie on different sides of the set $f_a(\Delta^*_{k-1})$ as demonstrated in Figure 65, where the function f_s is a constant, $f_s = 1.2$ at every vector on Δ_{k-1}. The relative epigraph $\mathrm{epi}(f_a, \Delta^*_{k-1})$ is the light gray area and the relative hypograph $\mathrm{hyp}(f_a, \Delta^*_{k-1})$ is the dark gray area in the figure. In this example, both sets $\mathrm{epi}(f_a, \Delta^*_{k-1})$ and $\mathrm{hyp}(f_a, \Delta^*_{k-1})$ are convex, but usually not at the same time or none of them is convex.
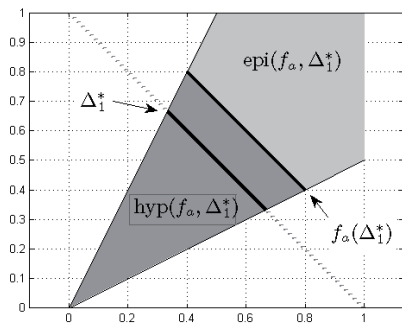
FIGURE 65   Sets $\mathrm{epi}(f_a, \Delta^*_{k-1})$ and $\mathrm{hyp}(f_a, \Delta^*_{k-1})$ in a simple example, where $f_a(\mathbf{z}) = 1.2\mathbf{z}$.

Definition 53 connects the function f_s and the relative convexity of f_a. In this way, we are able to consider convexity of functions as it has been done in [156]. However, first we have to prove the following:

Theorem 54. If $f_a(\Delta_{k-1}^*)$ is relatively convex, then the epigraph

$$\mathrm{epi}(f_s, \Delta_{k-1}^*) := \{(\mathbf{z}, \alpha) \in \Delta_{k-1}^* \times \mathbb{R} : \alpha \geq f_s(\mathbf{z})\}$$

is convex.

Proof. Let $(\mathbf{z}', \alpha'), (\mathbf{z}'', \alpha'') \in \mathrm{epi}(f_s, \Delta_{k-1}^*), 0 < \lambda < 1$ and $\mathbf{z}^\lambda := \lambda \mathbf{z}' + (1-\lambda)\mathbf{z}''$.
Let

$$t = \frac{\lambda \alpha''}{(1-\lambda)\alpha' + \lambda \alpha''} \quad (0 < t < 1).$$

Then

$$\lambda = \frac{t\alpha'}{t\alpha' + (1-t)\alpha''}, \quad (1-\lambda) = \frac{(1-t)\alpha''}{t\alpha' + (1-t)\alpha''}.$$

Let

$$\tilde{\mathbf{z}}^\lambda := t(\mathbf{ideal}^{Z_m} + \alpha'\mathbf{z}') + (1-t)(\mathbf{ideal}^{Z_m} + \alpha''\mathbf{z}'')$$
$$= \mathbf{ideal}^{Z_m} + t\alpha'\mathbf{z}' + (1-t)\alpha''\mathbf{z}''.$$

Because $\mathbf{z}', \mathbf{z}'' \in \Delta_{k-1}^*$, then

$$\sum_{i=1}^{k} z_i' = 1 = \sum_{i=1}^{k} z_i''$$

and, thus

$$f_{Z_m}^p(\tilde{\mathbf{z}}^\lambda) = \frac{t\alpha'\mathbf{z}' + (1-t)\alpha''\mathbf{z}''}{t\alpha' + (1-t)\alpha''} = \frac{t\alpha'}{t\alpha' + (1-t)\alpha''}\mathbf{z}' + \frac{(1-t)\alpha''}{t\alpha' + (1-t)\alpha''}\mathbf{z}''$$
$$= \lambda\mathbf{z}' + (1-\lambda)\mathbf{z}'' = \mathbf{z}^\lambda.$$

Clearly, vectors $\mathbf{ideal}^{Z_m} + \alpha'\mathbf{z}'$ and $\mathbf{ideal}^{Z_m} + \alpha''\mathbf{z}''$ belong to $\mathrm{epi}(f_a, \Delta_{k-1}^*)$ and because $\mathrm{epi}(f_a, \Delta_{k-1}^*)$ is convex, then so does the vector $\tilde{\mathbf{z}}^\lambda$. In addition, we have $t\alpha' + (1-t)\alpha'' \geq f_s(\mathbf{z}^\lambda)$ because $\mathrm{epi}(f_a, \Delta_{k-1}^*)$ is convex.
    Next, we show that $\lambda\alpha' + (1-\lambda)\alpha'' \geq t\alpha' + (1-t)\alpha''$.

$$0 \leq (\alpha' - \alpha'')^2$$
$$\Leftrightarrow$$
$$0 \leq (\lambda - \lambda^2)(\alpha'^2 + \alpha''^2) + 2(\lambda^2 - \lambda)\alpha'\alpha''$$
$$= (\lambda - \lambda^2)(\alpha'^2 + \alpha''^2) + (\lambda^2 + (1-\lambda)^2 - 1)\alpha'\alpha''$$
$$\Leftrightarrow$$
$$\alpha'\alpha'' \leq \lambda\alpha'(1-\lambda)\alpha' + \lambda\alpha'\lambda\alpha'' + (1-\lambda)\alpha''(1-\lambda)\alpha' + \lambda\alpha''(1-\lambda)\alpha''$$
$$= (\lambda\alpha' + (1-\lambda)\alpha'')((1-\lambda)\alpha' + \lambda\alpha'')$$
$$\Leftrightarrow$$
$$\lambda\alpha' + (1-\lambda)\alpha'' \geq \frac{\alpha'\alpha''}{(1-\lambda)\alpha' + \lambda\alpha''} = \frac{\lambda\alpha''\alpha'}{(1-\lambda)\alpha' + \lambda\alpha''} + \frac{(1-\lambda)\alpha'\alpha''}{(1-\lambda)\alpha' + \lambda\alpha''}$$
$$= t\alpha' + (1-t)\alpha''.$$

Therefore, $\lambda\alpha' + (1-\lambda)\alpha'' \geq f_s(\mathbf{z}^\lambda) = f_s(\lambda\mathbf{z}' + (1-\lambda)\mathbf{z}'')$, which means that

$$(\lambda\mathbf{z}' + (1-\lambda)\mathbf{z}'', \lambda\alpha' + (1-\lambda)\alpha'') \in \mathrm{epi}(f_s, \Delta^*_{k-1}).$$

Thus, the relative epigraph $\mathrm{epi}(f_s, \Delta^*_{k-1})$ is convex. $\qquad\square$

Practically, Theorem 54 means that, if $f_a(\Delta^*_{k-1})$ is relatively convex, then f_s is a convex function on Δ^*_{k-1}. However, the converse of Theorem 54 does not hold in general. In other words, if an epigraph $\mathrm{epi}(f_s, \Delta^*_{k-1})$ is convex (the function f_s is convex), then the relative epigraph $\mathrm{epi}(f_a, \Delta^*_{k-1})$ is not necessarily convex. This can be seen in the following:

Let $\mathbf{ideal}^{Z_m} = \mathbf{0}$ and f_s on Δ_1 be defined as follows:

$$f_s(\mathbf{z}) = 2 - z_1.$$

In other words, for $0 < \lambda < 1$ and $\mathbf{z}', \mathbf{z}'' \in \Delta_1$, we have

$$f_s(\lambda\mathbf{z}' + (1-\lambda)\mathbf{z}'') = 2 - (\lambda z_1' + (1-\lambda)z_1'') = \lambda(2 - z_1') + (1-\lambda)(2 - z_1'')$$
$$= \lambda f_s(\mathbf{z}') + (1-\lambda)f_s(\mathbf{z}''),$$

which means that f_s is a convex function [156]. Let

$$\mathbf{z}' := (0,1)^\mathrm{T}, \quad \mathbf{z}'' := (1,0)^\mathrm{T},$$
$$\tilde{\mathbf{z}}' := \mathbf{ideal}^{Z_m} + f_s(\mathbf{z}')\mathbf{z}' = \mathbf{0} + 2\mathbf{z}' = (0,2)^\mathrm{T},$$
$$\tilde{\mathbf{z}}'' := \mathbf{ideal}^{Z_m} + f_s(\mathbf{z}'')\mathbf{z}'' = \mathbf{0} + 1\mathbf{z}'' = (1,0)^\mathrm{T},$$
$$\tilde{\mathbf{z}}^{1/2} := \frac{1}{2}\tilde{\mathbf{z}}' + (1 - \frac{1}{2})\tilde{\mathbf{z}}'' = (1/2, 1)^\mathrm{T}.$$

Then

$$\mathbf{z}^{1/2} := f^p_{Z_m}(\tilde{\mathbf{z}}^{1/2}) = (1/3, 2/3)^\mathrm{T} \quad \left(= \frac{\tilde{\mathbf{z}}^{1/2} - \mathbf{ideal}^{Z_m}}{\sum_{i=1}^2 \tilde{z}_i^{1/2} - \mathrm{ideal}_i^{Z_m}} \right)$$

and furthermore

$$f_s(\mathbf{z}^{1/2}) = 2 - 1/3 = 5/3 > 3/2 = (1/2 + 1) = \sum_{i=1}^2 \tilde{z}_i^{1/2} - \mathrm{ideal}_i^{Z_m},$$

which means that the vector $\tilde{\mathbf{z}}^{1/2}$ does not belong to $\mathrm{epi}(f_a, \Delta^*_{k-1})$. Thus, $\mathrm{epi}(f_a, \Delta^*_{k-1})$ is not convex, and f_a is not relatively convex. However, if we assume the function f_s to be concave, then the situation is different, as we show next.

Theorem 55. If the hypograph

$$\mathrm{hyp}(f_s, \Delta^*_{k-1}) := \{(\mathbf{z}, \alpha) \in \Delta^*_{k-1} \times \mathbb{R} : \alpha \leq f_s(\mathbf{z})\}$$

is convex, then $f_a(\Delta^*_{k-1})$ is relatively concave.

Proof. Let $\mathbf{z}', \mathbf{z}'' \in \text{hyp}(f_a, \Delta^*_{k-1})$ and $0 < \lambda < 1$. In other words, we have vectors $\tilde{\mathbf{z}}', \tilde{\mathbf{z}}'' \in \Delta^*_{k-1}$ and scalars $0 < \alpha' \leq f_s(\tilde{\mathbf{z}}')$ and $0 < \alpha'' \leq f_s(\tilde{\mathbf{z}}'')$ so that $\mathbf{z}' = \mathbf{ideal}^{Z_m} + \alpha'\tilde{\mathbf{z}}'$ and $\mathbf{z}'' = \mathbf{ideal}^{Z_m} + \alpha''\tilde{\mathbf{z}}''$. Now,

$$\lambda\mathbf{z}' + (1 - \lambda)\mathbf{z}'' = \lambda(\mathbf{ideal}^{Z_m} + \alpha'\tilde{\mathbf{z}}') + (1 - \lambda)(\mathbf{ideal}^{Z_m} + \alpha''\tilde{\mathbf{z}}'')$$
$$= \mathbf{ideal}^{Z_m} + \alpha'\lambda\tilde{\mathbf{z}}' + \alpha''(1 - \lambda)\tilde{\mathbf{z}}''.$$

Let

$$t = \frac{\lambda\alpha'}{\lambda\alpha' + (1 - \lambda)\alpha''} \quad (0 < t < 1).$$

Because $\tilde{\mathbf{z}}', \tilde{\mathbf{z}}'' \in \Delta^*_{k-1}$, then

$$\sum_{i=1}^{k} \tilde{z}'_i = 1 = \sum_{i=1}^{k} \tilde{z}''_i$$

and, thus, the vector

$$\tilde{\mathbf{z}}^\lambda := f^p_{Z_m}(\mathbf{ideal}^{Z_m} + \alpha'\lambda\tilde{\mathbf{z}}' + \alpha''(1 - \lambda)\tilde{\mathbf{z}}'') = \frac{\lambda\alpha'\tilde{\mathbf{z}}' + (1 - \lambda)\alpha''\tilde{\mathbf{z}}''}{\lambda\alpha' + (1 - \lambda)\alpha''}$$
$$= \frac{\lambda\alpha'\tilde{\mathbf{z}}'}{\lambda\alpha' + (1 - \lambda)\alpha''} + \frac{(1 - \lambda)\alpha''\tilde{\mathbf{z}}''}{\lambda\alpha' + (1 - \lambda)\alpha''} = t\tilde{\mathbf{z}}' + (1 - t)\tilde{\mathbf{z}}''$$

is a convex combination of vectors $\tilde{\mathbf{z}}'$ and $\tilde{\mathbf{z}}''$. Because we have assumed f_s to be concave on Δ^*_{k-1}, then $t\alpha' + (1 - t)\alpha'' \leq f_s(\tilde{\mathbf{z}}^\lambda)$.

Next, we show that $\lambda\alpha' + (1 - \lambda)\alpha'' \leq t\alpha' + (1 - t)\alpha''$.

$$0 \geq -(\alpha' - \alpha'')^2$$
$$\Leftrightarrow$$
$$0 \geq (\lambda^2 - \lambda)\alpha'^2 + (\lambda^2 - \lambda)\alpha''^2 - 2(\lambda^2 - \lambda)\alpha'\alpha''$$
$$= (\lambda^2 - \lambda)\alpha'^2 + ((1 - \lambda)^2 - (1 - \lambda))\alpha''^2 + 2(\lambda - \lambda^2)\alpha'\alpha''$$
$$\Leftrightarrow$$
$$\lambda\alpha'^2 + (1 - \lambda)\alpha''^2 \geq \lambda^2\alpha'^2 + \lambda\alpha'(1 - \lambda)\alpha'' + (1 - \lambda)\alpha''\lambda\alpha' + (1 - \lambda)^2\alpha''^2$$
$$= (\lambda\alpha' + (1 - \lambda)\alpha'')(\lambda\alpha' + (1 - \lambda)\alpha'')$$
$$\Leftrightarrow$$
$$\lambda\alpha' + (1 - \lambda)\alpha'' \leq \frac{\lambda\alpha'^2 + (1 - \lambda)\alpha''^2}{\lambda\alpha' + (1 - \lambda)\alpha''} = \frac{\lambda\alpha'^2}{\lambda\alpha' + (1 - \lambda)\alpha''} + \frac{(1 - \lambda)\alpha''^2}{\lambda\alpha' + (1 - \lambda)\alpha''}$$
$$= t\alpha' + (1 - t)\alpha''.$$

Therefore $\lambda\alpha' + (1 - \lambda)\alpha'' \leq f_s(\tilde{\mathbf{z}}^\lambda)$ (note that $\mathbf{ideal}^{Z_m} + (\lambda\alpha' + (1 - \lambda)\alpha'')\tilde{\mathbf{z}}^\lambda = \lambda\mathbf{z}' + (1 - \lambda)\mathbf{z}''$). Thus, the relative hypograph $\text{hyp}(f_a, \Delta^*_{k-1})$ is convex and $f_a(\Delta^*_{k-1})$ is relatively concave. $\square$

In practice, Theorem 55 means that if the function f_s is concave on Δ^*_{k-1}, then $f_a(\Delta^*_{k-1})$ is relatively concave. However, the converse of Theorem 55 does not necessarily hold. This can be seen in the following example:

Let $\mathbf{ideal}^{Z_m} = \mathbf{0}$ and $\mathrm{hyp}(f_a, \Delta^*_{k-1})$ be the convex hull of vectors $(0,0)^{\mathrm{T}}$, $(0,2)^{\mathrm{T}}$, $(1,0)^{\mathrm{T}}$. Clearly, the set $\mathrm{hyp}(f_a, \Delta^*_{k-1})$ is convex. Let

$$\mathbf{z}' := ((0,1)^{\mathrm{T}}, 2), \quad \mathbf{z}'' := ((1,0)^{\mathrm{T}}, 1),$$

$$\mathbf{z}^{1/2} := \frac{1}{2}\mathbf{z}' + (1 - \frac{1}{2})\mathbf{z}'' = ((1/2, 1/2)^{\mathrm{T}}, 3/2),$$

$$\tilde{\mathbf{z}}' := (0,2)^{\mathrm{T}} = \mathbf{0} + 2(0,1)^{\mathrm{T}} \quad (= \mathbf{ideal}^{Z_m} + f_s((0,1)^{\mathrm{T}})(0,1)^{\mathrm{T}}),$$

$$\tilde{\mathbf{z}}'' := (1,0)^{\mathrm{T}} = \mathbf{0} + 1(1,0)^{\mathrm{T}} \quad (= \mathbf{ideal}^{Z_m} + f_s((1,0)^{\mathrm{T}})(1,0)^{\mathrm{T}}).$$

Now, the vector $(2/3, 2/3)^{\mathrm{T}} = 1/3\tilde{\mathbf{z}}' + (1 - 1/3)\tilde{\mathbf{z}}''$ is a convex combination of the vectors $\tilde{\mathbf{z}}'$ and $\tilde{\mathbf{z}}''$. Note that

$$(2/3, 2/3)^{\mathrm{T}} = \mathbf{0} + 4/3(1/2, 1/2)^{\mathrm{T}} \quad (= \mathbf{ideal}^{Z_m} + f_s((1/2, 1/2)^{\mathrm{T}})(1/2, 1/2)^{\mathrm{T}}).$$

However, $f_s((1/2, 1/2)^{\mathrm{T}}) = 4/3 < 3/2$, which means that the pair $(\mathbf{z}^{1/2}, 3/2)$ does not belong to $\mathrm{hyp}(f_s, \Delta^*_{k-1})$. Thus, $\mathrm{hyp}(f_s, \Delta^*_{k-1})$ is not convex.

Let us next summarize the cases of convexity and concavity as follows: If the function f_s is

- convex on Δ^*_{k-1}, then the set $f_a(\Delta^*_{k-1})$ is not necessarily relatively convex,
- concave on Δ^*_{k-1}, then the set $f_a(\Delta^*_{k-1})$ is relatively concave.

However, if the set $f_a(\Delta^*_{k-1})$ is

- relatively convex, then the function f_s is convex on Δ^*_{k-1},
- relatively concave, then the function f_s is not necessarily concave on Δ^*_{k-1}.

To use more familiar sets, we first prove the following connection to the *Edgeworth-Pareto hull* of $f_a(\Delta^*_{k-1})$, that is, the set $f_a(\Delta^*_{k-1}) + \mathbb{R}^k_+$.

Theorem 56. The set $f_a(\Delta^*_{k-1})$ is relatively convex if and only if the set $f_a(\Delta^*_{k-1}) + \mathbb{R}^k_+$ is convex.

Proof.

"$\Rightarrow$" Let $\mathbf{z}', \mathbf{z}'' \in f_a(\Delta^*_{k-1}) + \mathbb{R}^k_+$ and $0 < \lambda < 1$. In other words, we have vectors $\tilde{\mathbf{z}}', \tilde{\mathbf{z}}'' \in f_a(\Delta^*_{k-1})$ and $\bar{\mathbf{z}}', \bar{\mathbf{z}}'' \in \mathbb{R}^k_+$ so that $\mathbf{z}' = \tilde{\mathbf{z}}' + \bar{\mathbf{z}}'$ and $\mathbf{z}'' = \tilde{\mathbf{z}}'' + \bar{\mathbf{z}}''$. Therefore,

$$\lambda\mathbf{z}' + (1 - \lambda)\mathbf{z}'' = \lambda(\tilde{\mathbf{z}}' + \bar{\mathbf{z}}') + (1 - \lambda)(\tilde{\mathbf{z}}'' + \bar{\mathbf{z}}'')$$
$$= \lambda\tilde{\mathbf{z}}' + (1 - \lambda)\tilde{\mathbf{z}}'' + \lambda\bar{\mathbf{z}}' + (1 - \lambda)\bar{\mathbf{z}}''.$$

Because $\mathbb{R}^k_+$ is convex, then $\lambda(\bar{\mathbf{z}}') + (1 - \lambda)\bar{\mathbf{z}}''$ naturally belongs to $\mathbb{R}^n_+$. In addition, because $f_a(\Delta^*_{k-1})$ is relatively convex, then we have $\hat{\mathbf{z}}^\lambda \in \Delta^*_{k-1}$ and $s_\lambda \geq f_s(\hat{\mathbf{z}}^\lambda)$ so that

$$\mathbf{ideal}^{Z_m} + s_\lambda\hat{\mathbf{z}}^\lambda = \lambda\tilde{\mathbf{z}}' + (1 - \lambda)\tilde{\mathbf{z}}''.$$

Let
$$\mathbf{z}^d := \mathbf{ideal}^{Z_m} + s_\lambda \hat{\mathbf{z}}^\lambda - f_a(\hat{\mathbf{z}}^\lambda) = (s_\lambda - f_s(\hat{\mathbf{z}}^\lambda))\hat{\mathbf{z}}^\lambda \quad (\in \mathbb{R}^n_+).$$

Now,

$$f_a(\hat{\mathbf{z}}^\lambda) + \underbrace{\mathbf{z}^d + \lambda\bar{\mathbf{z}}' + (1-\lambda)\bar{\mathbf{z}}''}_{\in \mathbb{R}^k_+} = \mathbf{ideal}^{Z_m} + s_\lambda \hat{\mathbf{z}}^\lambda + \lambda\bar{\mathbf{z}}' + (1-\lambda)\bar{\mathbf{z}}''$$

$$= \lambda\tilde{\mathbf{z}}' + (1-\lambda)\tilde{\mathbf{z}}'' + \lambda\bar{\mathbf{z}}' + (1-\lambda)\bar{\mathbf{z}}''$$
$$= \lambda\mathbf{z}' + (1-\lambda)\mathbf{z}''.$$

In other words, $\lambda\mathbf{z}' + (1-\lambda)\mathbf{z}''$ belongs to $f_a(\Delta^*_{k-1}) + \mathbb{R}^n_+$ and, thus, $f_a(\Delta^*_{k-1}) + \mathbb{R}^k_+$ is convex.

"$\Leftarrow$" Let $\mathbf{z}', \mathbf{z}'' \in \mathrm{epi}(f_a, \Delta^*_{k-1})$ and $0 < \lambda < 1$. In other words, we have vectors $\tilde{\mathbf{z}}', \tilde{\mathbf{z}}'' \in \Delta^*_{k-1}$ and $\alpha' \geq f_s(\tilde{\mathbf{z}}')$, $\alpha'' \geq f_s(\tilde{\mathbf{z}}'')$ so that $\mathbf{z}' = \mathbf{ideal}^{Z_m} + \alpha'\tilde{\mathbf{z}}'$ and $\mathbf{z}'' = \mathbf{ideal}^{Z_m} + \alpha''\tilde{\mathbf{z}}''$.

Let
$$\tilde{\mathbf{z}}^\lambda := f^p_{Z_m}(\lambda\mathbf{z}' + (1-\lambda)\mathbf{z}'').$$

Clearly $\tilde{\mathbf{z}}^\lambda$ is a convex combination of vectors $\tilde{\mathbf{z}}'$ and $\tilde{\mathbf{z}}''$ and it belongs to Δ^*_{k-1} because $f_a(\Delta^*_{k-1}) + \mathbb{R}^k_+$ is convex.

Let $\alpha_\lambda > 0$ be selected so that $\lambda\mathbf{z}' + (1-\lambda)\mathbf{z}'' = \mathbf{ideal}^{Z_m} + \alpha_\lambda\tilde{\mathbf{z}}^\lambda$. Now, because $f_a(\Delta^*_{k-1}) + \mathbb{R}^k_+$ is convex, then vectors $\mathbf{z}'$, $\mathbf{z}''$ and $\lambda\mathbf{z}' + (1-\lambda)\mathbf{z}''$ belong to $f_a(\Delta^*_{k-1}) + \mathbb{R}^k_+$. Therefore, $f_s(\tilde{\mathbf{z}}^\lambda) \leq \alpha_\lambda$, which means that $\lambda\mathbf{z}' + (1-\lambda)\mathbf{z}''$ belongs to $\mathrm{epi}(f_a, \Delta^*_{k-1})$. Thus $\mathrm{epi}(f_a, \Delta^*_{k-1})$ is convex and $f_a(\Delta^*_{k-1})$ is relatively convex.

$\square$

Now, we immediately obtain the following result.

Theorem 57. If the Edgeworth-Pareto hull of $f_a(\Delta^*_{k-1})$ (i.e. $f_a(\Delta^*_{k-1}) + \mathbb{R}^k_+$) is convex, then the function f_s is convex on Δ^*_{k-1}.

Proof. This follows directly from Theorems 54 and 56. $\square$

Correspondingly, we have a similar result for concavity.

Theorem 58. The set $f_a(\Delta^*_{k-1})$ is relatively concave if and only if the set $f_a(\Delta^*_{k-1}) + \mathbb{R}^k_-$ is convex.

Proof. The proof is similar to that of Theorem 56. $\square$

To conclude this subsection, relative convexity or concavity of the set $f_a(\Delta^*_{k-1})$ characterizes the shape of the set of approximated Pareto optimal objective vectors. Perhaps the most interesting finding is that the convexity of the function f_s does not guarantee that the set $f_a(\Delta^*_{k-1})$ is relatively convex, i.e. the Edgeworth-Pareto hull of the nondominated vectors in the approximation is convex. However, the results obtained in this subsection and the theorems can be useful for the analyst, who can use them as follows:

For example, if the MOO problem is convex, then the Edgeworth-Pareto hull of the Pareto front is convex. The analyst can now use Theorem 57 to investigate the approximation, that is, to study how the function f_s behaves on the nondominated vectors in the simplex. If it turns out that the function f_s is not convex on that particular set, then the analyst immediately recognizes that the function f_s in the function f_a is not necessarily a proper one. Had the approximation been a proper one, then the function f_s would be a convex function on Δ_{k-1}^* by Theorem 58. On the other hand, if it turns out that the function f_s is concave, then such a function cannot be used for a convex MOO problem because the set $f_a(\Delta_{k-1}^*)$ is relatively concave. However, the analyst must be aware that for a nonconvex MOO problem the Pareto front can be relatively convex, according to the terminology defined in this chapter for the approximation. Then, in investigating the shape of the approximated Pareto front of a nonconvex MOO problem, the analyst must be careful in drawing right conclusions.

6.4.2 On Connected Sets Δ_{k-1}^* and $f_a(\Delta_{k-1}^*)$

Relative convexity of a set $f_a(\Delta_{k-1}^*)$ means that a set Δ_{k-1}^* is convex. In other words, the set Δ_{k-1}^* is connected. In this subsection, we briefly investigate the connectedness of the sets Δ_{k-1}^* and $f_a(\Delta_{k-1}^*)$.

From the theoretical point of view, the connectedness can be interesting in describing the shape of the approximated Pareto front. One might think that the connectedness can be used in validating the approximation. This is possible but can be very challenging at the same time.

Let us first demonstrate that the relative convexity of the set $f_a(\Delta_{k-1}^*)$ does not necessarily mean that the set $f_a(\Delta_{k-1}^*)$ is connected.

Example 59. Let $\mathbf{ideal}^{Z_m} = \mathbf{0}$ and

$$f_s(\mathbf{z}) = \begin{cases} 2, & \mathbf{z} = (1,0) \text{ or } \mathbf{z} = (0,1), \\ 1, & \text{otherwise.} \end{cases}$$

Then $\Delta_1^* = \Delta_1$ (connected) and $f_a(\Delta_1^*)$ is relatively convex, but it is not connected, see Figure 66.

In other words, the analyst cannot necessarily conclude whether or not the Pareto front is connected in validating the approximation. Note that for a convex MOO problem the Pareto front can be the set $f_a(\Delta_1)$ in Example 59 (a convex function on $[0,1]$ is not necessarily continuous). This introduces challenges in using the connectedness of the set $f_a(\Delta_{k-1})$ in validating the approximation in relation to the topological properties of the MOO problem because even the convexity of the MOO problem does not necessarily mean that the Pareto front is connected. However, for a continuous convex MOO problem, the Pareto front is connected trivially.

Example 59 is in several ways interesting. First of all, the set $f_a(\Delta_{k-1}^*)$ is relatively convex, but it is not connected. Secondly, let us assume that the Pareto front
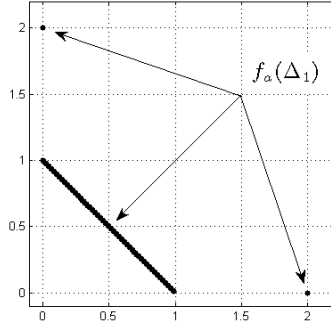
FIGURE 66   Convexity of set Δ_{k-1}^{*} does not mean that set $f_a(\Delta_{k-1}^{*})$ is connected even if function f_s is convex.
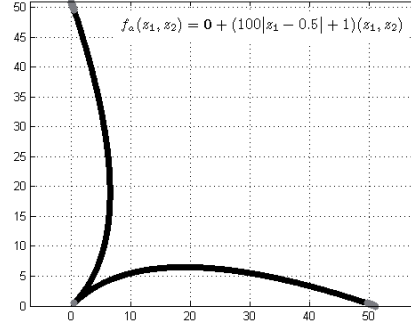
FIGURE 67   A convex function f_s can be justified for a nonconvex MOO problem.

of a MOO problem is $f_a(\Delta_1)$. Then one may wonder whether there exist holes in the Pareto front. Note that neither normal vectors nor the condition $f_u(\mathbf{z}) < f_s(\mathbf{z})$ characterize holes, because they are defined on Δ_{k-1}°. In other words, it can be stated that the Pareto front contains holes, but methods proposed in this thesis do not characterize them. However, from a practical point of view, Example 59 is a pathological example because in practice a Pareto optimal objective vector can be found arbitrarily close to the vector $(0,1)^{\mathrm{T}}$. The decision maker might now select a Pareto optimal objective vector close to the vector $(0,1)^{\mathrm{T}}$ instead of the vector $(0,2)^{\mathrm{T}}$, as we are minimizing the values of objective functions.

Example 59 demonstrates that it cannot be said for sure whether the set $f_a(\Delta_{k-1}^{*})$ is connected, when the set Δ_{k-1}^{*} is connected. In other words and in general, if $f_a(\Delta_{k-1}^{*})$ is connected, then it is connected. However, if f_s is continuous, then the set $f_a(\Delta_{k-1}^{*})$ is connected if and only if Δ_{k-1}^{*} is connected. For simplicity, we only prove the case where $\Delta_{k-1}^{*} \subset \Delta_{k-1}^{\circ}$.

Theorem 60. Let the function f_s be continuous and $\Delta_{k-1}^{*} \subset \Delta_{k-1}^{\circ}$. Then $f_a(\Delta_{k-1}^{*})$ is connected if and only if Δ_{k-1}^{*} is connected.

Proof.

"$\Rightarrow$" Because the function $f_{Z_m}^{p}$ limited to $f_a(\Delta_{k-1})$ is the inverse function of f_s and it is continuous, then the claim is clear.
"$\Leftarrow$" Trivially.

$\square$

This completes our brief investigation of the connectedness of sets Δ_{k-1}^{*} and $f_a(\Delta_{k-1}^{*})$. Next, we show how the analyst can use the convexity, concavity and connectedness in selecting or validating the approximation.

6.4.3 Convexity, Concavity and Connectedness from the Practical Point of View

It is obvious that topological features such as convexity and connectedness describe in a theoretical way the shape of the approximation, and the analyst can use the theoretical result discussed in this section in validating the approximation. In other words, the analyst can check whether or not the approximation is justified for the MOO problem the decision maker is investigating by studying topological properties of the problem and the approximation.

In these kinds of investigations, the analyst must be careful, and we can illustrate this as follows: Let us assume that the MOO problem is nonconvex. The analyst may think that for nonconvex MOO problems there is no reason to use a convex function f_s. Note that the convexity of the function f_s does not mean that it should not be used for a nonconvex MOO problem. This can be seen as follows: Let the function f_s be defined on Δ_1 so that $f_s(\mathbf{z}) = 100|z_1 - 0.5| + 1$ (a convex function). In Figure 67, we have depicted the set $f_a(\Delta_1)$ (black curves representing dominated vectors and gray curves representing nondominated vectors). As it can be seen, the convex function f_s has generated a Pareto front approximation that is suitable for a nonconvex MOO problem. Thus, a convex function f_s can be used to approximate the Pareto front of a nonconvex MOO problem. Another interesting feature in this case is that the convexity of the function f_s does not guarantee that the set Δ_{k-1}^* is convex. In other words, the set Δ_{k-1}^* is not connected. As demonstrated in the previous subsection, connectedness can be challenging in validating the approximation. However, if the analyst can identify the set Δ_{k-1}^* as connected and the function f_s as continuous, then he or she can conclude that the set of nondominated vectors in the approximation is connected i.e. the set does not consist of at least two separated sets by Theorem 60. This information can be useful if the analyst is investigating whether or not an approximated hole in the Pareto front separates the front into several separated sets.

The above examples demonstrate that the analyst must be aware about the topological properties such as convexity of the set S or convexity of the function f_s and about how these can be used in selecting the approximation or investigating some other topological properties related to the sets that indicate what kind of set the Pareto front may be. Such studies can occasional be challenging for the analyst, but simple rules can be created to help in this matter. For example, if a concave function f_s is used for a convex MOO problem, then the approximation may not be a sufficiently good one. This is guaranteed by Theorem 55 and by the fact that the Edgeworth-Pareto hull of a convex MOO problem is convex. Alternatively, for a valid approximation created for a continuous convex MOO problem, the set Δ_{k-1}^* should be connected.

6.5 Summary of Chapter 6

Let us now summarize the main topics of this chapter. Briefly, we have discussed the concept of approximated holes in the Pareto front. Holes characterize regions in the objective space, where (approximated) Pareto optimal objective vectors could exist, but they do not exist because of the nonconvexity of the MOO problem. In this way, the decision maker can be informed in step (4) of HBE about the region under investigation. The decision maker can now use time for other regions in the objective space, where (approximated) Pareto optimal objective vectors may exist.

Topological properties were also studied briefly in this chapter. We showed, for example, that for a convex function f_s, the Edgeworth-Pareto hull of an approximated Pareto front is not necessarily convex. With the help of the topological information, the analyst can investigate whether or not the approximation is valid for the MOO problem from a topological point of view.

Let us next summarize the theorems studied in this chapter.
Theorem

50 In Definition 47, hole vectors were defined. However, the definition does not provide a clue about how to characterize hole vectors computationally. On the other hand, it introduces a theoretical way to do the characterization by using the results of Chapters 4 and 5. In other words, if the function f_s is continuously differentiable and a vector $\mathbf{z} \in \mathrm{rel\,int\,CH}\,\Delta_{k-1}^*$ is dominated (characterized by the normal vector or the condition $f_u(\mathbf{z}) < f_s(\mathbf{z})$) and all vectors in a relative open neighborhood of $\mathbf{z}$ are dominated, then the vector $f_a(\mathbf{z})$ is a hole vector. This was proved in this theorem.

54 In defining hole vectors, convexity of the set $\mathrm{CH}\,\Delta_{k-1}^*$ was considered. One may wonder whether the convexity of the function f_s means anything related to the approximated Pareto optimal objective vectors. It turns out in Theorem 54 that if the set $f_a(\Delta_{k-1}^*)$ is relatively convex, then the function f_s is convex on Δ_{k-1}^*. However, the converse does not hold generally.

55 In this theorem, we studied concavity instead of convexity because that one was done in Theorem 54. It turned out that, if the function f_s is concave on Δ_{k-1}^*, then $f_a(\Delta_{k-1}^*)$ is relatively concave. However, the converse is not generally true.

56 In the literature, an Edgeworth-Pareto hull is a commonly used concept [112, 113], and in Theorem 56 we connected this to Definition 53. We showed that it does not matter whether we use the Edgeworth-Pareto hull or Definition 53. However, Definition 53 combines the convexity of the epigraph and hypograph of the function f_s to the relative convexity and concavity of the set $f_a(\Delta_{k-1}^*)$.

58 This theorem is similar to Theorem 56, but it is based on relative concavity.

60 Finally, we briefly investigated connectedness of the sets Δ_{k-1}^* and $f_a(\Delta_{k-1})^*$.

We proved that if the function f_s is continuous, then Δ_{k-1}^* is connected if and only if the set $f_a(\Delta_{k-1})^*$ is connected.

Next, we summarize the results related to the approximation method $\mathcal{AM}$.

$\mathcal{AM}$f14   It is obvious that the hole vectors introduced in this chapter are listed here. They can be used in characterizing regions in the objective space where (approximated) Pareto optimal objective vectors may exist but do not. This is essential information for the decision maker in the creation of an overall understanding of available (approximated) Pareto optimal objective vectors because he or she can save time for other regions. This means that the hole vectors are related to the goal $\mathcal{G}9$ set on page 20.

However, the characterization can only be done in the relative interior of the simplex and not on the relative boundary of it. Technically, the characterization of hole vectors is based on results created in Chapters 4 and 5. From the practical and computational implementation point of view, linear optimization problems must be solved, but they do not introduce significant computational costs to the HBE method.

$\mathcal{AM}$f15   The relative convexity or concavity of the set $f_a(\Delta_{k-1}^*)$ and the convexity or concavity of the function f_s on Δ_{k-1}^* can be essential or interesting from a theoretical point of view. They characterize the shape of the set of the approximated Pareto optimal objective vectors. This means that the analyst can use these features in selecting the approximation to the MOO problem, as discussed in Subsection 6.4.3. In other words, we have considered the validation of the approximation, i.e. the goal $\mathcal{G}6$, in this chapter.

In this chapter, we have considered how some of the dominated vectors in the approximation can be interpreted in the solution process of the MOO problem. It turned out that they do characterize approximated holes in the Pareto front, and this can help the decision maker in the creation of an overall understanding of available (approximated) Pareto optimal objective vectors. However, this is one of the characterizations for a specific topic in helping and supporting the decision maker in the solution process. Next, we summarize all the methods proposed in this thesis, i.e. how they can be used in the solution process and how they may help the decision maker in learning about the MOO problem.

7 INTERACTIVE HYPERBOX EXPLORATION

In this thesis, an approximation for a Pareto front of a computationally expensive multiobjective optimization problem has been proposed in Chapter 2. Then, in Chapter 3, we proposed an interactive method called Interactive HyperBox Exploration, which uses the approximation. In that chapter, we also proposed several methods to be used during the solution process of the MOO problem in order to help and support the decision maker in finding a compromise solution to the MOO problem. Chapters 4 and 5 were mainly focused on the validation methods created for the approximation. In addition, we proposed several methods for the HBE method employed by in the IHBE method for analyzing the vectors in the approximated Pareto front. Finally, in Chapter 6, we discussed the way dominated vectors in the approximation can be interpreted in the solution process. For the interpretations, we proposed a couple of methods to be used in connection with the HBE method. Overall, we have proposed several methods for different purposes for the IHBE and HBE methods. Thus, it is reasonable to summarize what information the IHBE and HBE methods provide for the decision maker in the solution process of the MOO problem at hand.

To describe the background of the IHBE and HBE methods briefly, we developed the IHBE method for supporting and helping the decision maker in the learning phase of the solution process. For learning about the problem at hand, we proposed a process which we described as the creation of an overall understanding of available (approximated) Pareto optimal objective vectors. The goal for the process is to help the decision maker to aim at the psychological convergence in the solution process. For this purpose, the IHBE and HBE methods allow the decision maker to freely search the objective space and (approximated) Pareto optimal objective vectors within the space. In order to find approximated Pareto optimal objective vectors that the decision maker is interested in, the HBE method answers the following question formulated in a query:

I am interested in these kinds of (approximated) Pareto optimal objective vectors.
Do those vectors exist?

To answer the question, the HBE method generates sample vectors on the approx-

imation so that they correspond to the decision maker's preferences, which are asked in the form of aspiration and reservation levels. Then these sample vectors are used for answering the question (the answers are summarized later in this chapter). In this way, the decision maker can learn about the problem and about what kinds of (approximated) Pareto optimal objective vectors exist in relation to his or her preferences. Based on these facts learned, the decision maker may be confident about a compromise solution finally selected to the MOO problem.

Next, we summarize the steps of the IHBE method, and then we focus on the information that the HBE method provides for the decision maker. In the IHBE method, there are seven steps, and the first three steps consist of generating the approximation. The approximation is based on a surrogate function and, in order to find a proper surrogate function to be used in the approximation, we have proposed validation methods for the analyst. Basically, the approximation is based on pre-calculated Pareto optimal objective vectors to the MOO problem. The validation methods proposed in Chapters 4 and 5 provide information for the analyst about whether the approximation is valid or not in relation to the pre-calculated Pareto optimal objective vectors. The validity is based on the pre-calculated Pareto optimal objective vectors and on whether or not some vectors in the approximation dominate them.

In the fourth step of IHBE, the decision maker starts to use the approximation in the solution process. In other words, the decision maker communicates with the HBE method in order to learn about what kinds of (approximated) Pareto optimal objective vectors are available. Then, in the last three steps of IHBE, the decision maker can update the approximation by asking the method to calculate new true Pareto optimal objective vectors to the MOO problem. In addition, the decision maker can decide either to continue with the HBE method in searching them or stop the IHBE method.

The answer provided by the HBE method to the decision maker concerning specific (approximated) Pareto optimal objective vectors is based on sample vectors in the approximation, which correspond to the decision maker's preferences. With the help of the sample vectors, the HBE method can also provide additional information about the Pareto front. The decision maker can select the information he or she wants to see in the answer. In the simplest case, the HBE method simply answer "Yes, they exist!" or "No, they do not exist!". The answer naturally depends on whether or not the sample vectors in the approximation satisfy the decision maker's preferences.

When the answer to the question is "Yes, they exist!", the HBE method naturally present the sample vectors satisfying the aspiration and reservation levels. Next, we summarize what kind of information the HBE method can provide for the sample vectors also.

Because the HBE method is created for learning, then the decision maker can select what kind of additional information is wanted to find the sample vectors that satisfy the preferences. The decision maker could also point out the sample vectors for which some more information is needed. In this way, the HBE method adapts to the decision maker's capability to handle information, that is,

it does not present too much information by default to the decision maker. In other words, the HBE method allows the decision maker to control cognitive load. However, it is the decision maker's responsibility to tell the HBE method what information to present. The information can consist, for example, of the closest true Pareto optimal objective vectors to the sample vectors, of approximated trade-off rates at the sample vectors, or of certainty numbers for the sample vectors. These all are considered in this thesis. In this way, the decision maker may learn where the so-far known true Pareto optimal objective vectors are in relation to the sample vectors. In addition, the decision maker may learn what kinds of approximated Pareto optimal objective vectors may be found around a small neighborhood of the sample vectors.

Next, we discuss a case where the decision maker cannot have approximated Pareto optimal objective vectors satisfying the preferences, i.e. the answer to the question is "No, they do not exist!". We divide the related discussion into two paragraphs. In the first one, we consider a case where the decision maker has started the solution process. In the second one, we discuss how to continue in such a case, that is, how the HBE method supports or helps the decision maker to continue, i.e. to redefine new levels.

When the answer to the question is negative, the HBE method provides (if possible) an explanation concerning why such approximated Pareto optimal objective vectors do not exist. In our opinion, it is very important to explain that for the decision maker at the beginning of the solution process. The decision maker can then adjust his or her probably unrealistic preferences and concentrate on the matching ones. To make an explanation possible, the HBE method investigates the aspiration and reservation levels given by the decision maker. These can be, for example, optimistic or pessimistic. On the other hand, the hyperbox defined by the levels can be in an approximated hole in the Pareto front. In order to identify such situations, we have developed specific methods for each of the cases for the HBE method, and the investigations are based on the sample vectors. The explanation based on the sample vectors can be, for example: "The (approximated) Pareto optimal objective vectors you are interested in do not exist, because your preferences are optimistic.", "... pessimistic.", or "... , because the Pareto front contains an approximated hole in the region you are interested in." Through these explanations, we believe that the decision maker can learn about his or her preferences in relation what kind of (approximated) Pareto optimal objective vectors exist. This can be essential at the beginning of the solution process.

It is obvious that informing the decision maker about optimistic or pessimistic levels is not necessarily enough to get to know about the approximated Pareto front, especially when the answer to the query is negative. Therefore, the HBE method presents, if the decision maker so wants, the sample vectors, which somehow correspond to the aspiration and reservation levels. In this way, the HBE method helps the decision maker to find the (approximated) Pareto front in some sense related to the preferences. In addition, the decision maker may learn how to set new levels by the sample vectors presented. In other words, the

decision maker may set new levels, i.e. to make a new query, so that the sample vectors satisfy the new levels. Then, the decision maker can obtain and learn detailed information around the old sample vectors by the new sample vectors in the new query. Note that the new ones may represent in more details some certain region in the approximated Pareto front represented earlier partially by the old sample vectors. In addition, the HBE method can present to the decision maker all the numerical information mentioned earlier for the sample vectors.

As pointed out earlier, the HBE method aims at explaining why the decision maker cannot have Pareto optimal objective vectors satisfying the preferences. However, the named and discussed hyperboxes in this thesis do not cover all possible hyperboxes defined by the aspiration and reservation levels. It is possible that the decision maker might not have preferred Pareto optimal objective vectors and the levels, staying neither optimistic nor pessimistic. In such a case, the HBE method does not state why the decision maker cannot have such (approximated) Pareto optimal objective vectors because this information cannot always be derived from the sample vectors. However, the HBE method can present the sample vectors corresponding to the preferences and analyze them, i.e. show which ones are, for example, dominated, nondominated or hole vectors by the methods proposed in this thesis. If the decision maker is interested in numerical information mentioned earlier for the sample vectors, then these can also be presented by the method.

As stated earlier, the HBE method provides information for the decision maker in the learning phase of the solution process. The information, naturally, depends on the time the decision maker has used for the solution process. At the beginning of the process, the information might consists only of few items, but after having learned about the approximated Pareto front, the decision maker may work with more detailed information, which is generated for the approximated Pareto optimal objective vectors presented in the answer for the question in the query. There are still some future challenges in developing a graphical user interface dedicated for the purpose for the IHBE method. However, we are highly motivated to continue developing the IHBE method, not least because the IHBE method is based on methods, which were introduced in the MCDM literature and have been developed for the communications between a human decision maker and a computer software. These methods have received positive feedback from the users, i.e. the real life decision makers. The approximation created in this thesis has allowed us to use the methods also for computationally expensive nonconvex MOO problems. The calculations can now be as fast as they are for computationally inexpensive MOO problems.

We have now summarized the IHBE and the HBE methods. In Figure 68, we have listed the methods discussed in this chapter, and we have organized them into the steps of the IHBE method, where they are used. Where it comes to the research objectives set on page 20, we can say that most of the goals have been clearly reached. For example, it does not matter whether or not the MOO problem consist of more than two objective functions, the objective functions being nonconvex and computationally expensive, only a finite number of Pareto optimal

objective vectors are used for generating the approximation and generating the vectors on the approximation that somehow satisfy the decision maker's preferences can be done computationally efficiently, meaning that the goals $\mathcal{G}1$-$\mathcal{G}4$ and $\mathcal{G}7$ are clear. In addition, the goal $\mathcal{G}8$ has been reached because there exists a way for the decision maker to specify his or her preferences for the IHBE method. However, the goals $\mathcal{G}5$ and $\mathcal{G}6$ have been only partially reached: depending on the surrogate function used in the function f_a, the function may contain some parameters to be optimized that may take some time, and the validations based on normal vectors and the functions f_l, f_u and f_s may not identify all the dominated vectors in the approximation. However, the actual generation of the approximation and its validations can be performed computationally efficiently, but the above details clarify why these goals are not fully reached. The goal $\mathcal{G}9$ is a bit tricky to reach because the IHBE and HBE methods provide information for the decision maker, but we do not know yet how significant and essential it is for the decision maker in the solution process. Therefore, we say that the goal has been partially reached in this thesis and is open for future studies.

Interactive Hyperbox Exploration

Initialization (generating the approximation), steps 1-3

- Investigate dominated vectors for validations by Methods 36, 37 and 43.
- Investigate topological properties discussed in Section 6.4.

HBE (generating vectors on the approximation), step 4

- **Generating the sample vectors on the approximated Pareto front for the answer by Method 16**.
  - Investigate nondominated and dominated sample vectors by using Theorems 28, 29, 30, 40 and 42.
  - Investigate hole vectors by Method 51.
- **Additional information for the answer:**
  - Closest true Pareto optimal objective vectors and corresponding decision variable values discussed in Subsection 3.3.4.
  - Approximated trade-off rates for the sample vectors by Method 38.
  - Certainty numbers for the sample vectors by Method 46.
  - Calculating error estimations by using function (17).
- **For the explanation:**
  - Investigate optimistic/pessimistic hyperboxes by Methods 21 and 24.
  - Investigate an approximated hole in the Pareto front by Method 52.
  - Investigate whether the sample vectors dominate each other as discussed at the end of Subsection 3.3.3.

Updating (the approximation), steps 5-7

- Certainty numbers for the interesting approximated Pareto optimal objective vectors by Method 46.
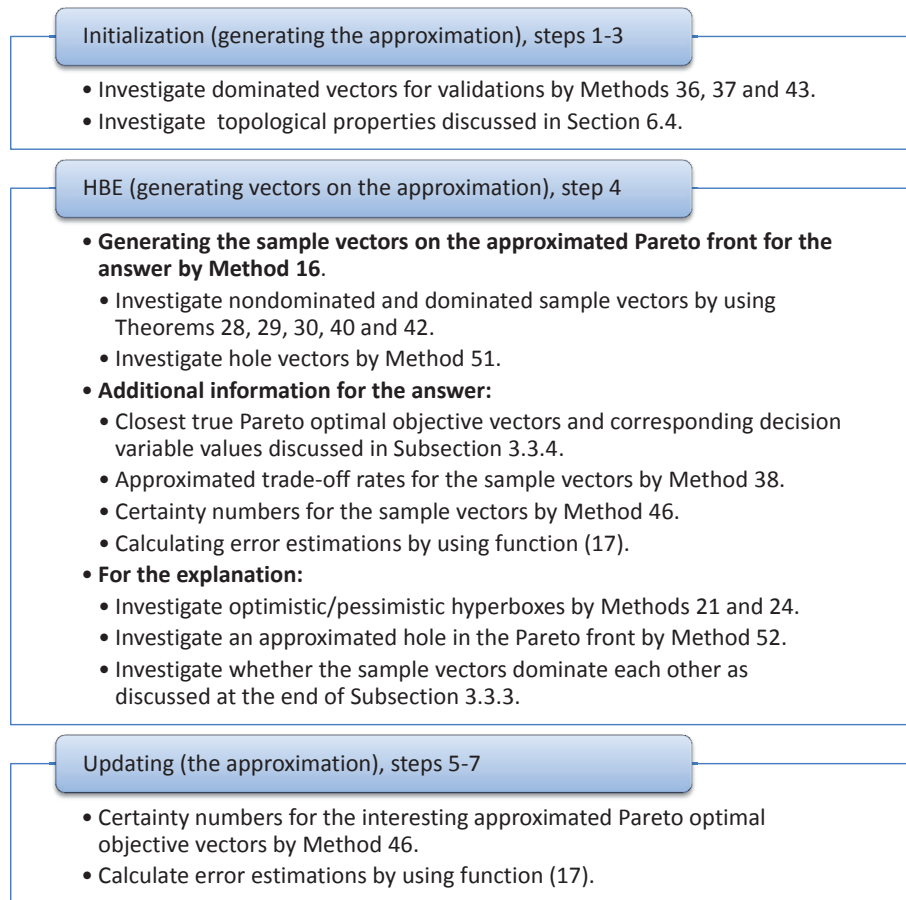- Calculate error estimations by using function (17).

FIGURE 68   Methods and theorems discussed and developed in this thesis that can be used in the steps of IHBE.

8 CONCLUSIONS AND FURTHER STUDIES

In this thesis, a method called Interactive HyperBox Exploration (IHBE) has been proposed for computationally expensive nonconvex multiobjective optimization problems. The method was summarized in the previous chapter. Three aspects were considered during the construction of the IHBE method: theoretical, computational, and the decision maker's point of views.

To clarify the theoretical aspect, an approximation was created to approximate the Pareto front of the MOO problem. Then it was proved mathematically that the approximation is suitable for computationally expensive convex/nonconvex MOO problems through a surrogate function used in the approximation. In addition, the approximation can be validated before the decision maker starts any interaction with it. The validations were based on simple formulas which were proved mathematically. For example, characterization of nondominated or dominated vectors by normal vectors is based on mathematical theories. In addition, the concept of holes in the Pareto front is rather mathematical. However, in this thesis, computational counterparts are presented for the theoretical ones, and this leads to the next aspect.

The computational aspect is as important feature in the IHBE method as the mathematical theories introduced in this thesis. As said, mathematical theories and theorems are used for different purposes in the approximation. However, from the practical point of view, all the theories and theorems should have a computational counterpart if they are considered to be used in the solution process of a MOO problem. As showed in this thesis, computational methods based on the theories and theorems were found, but some of the methods are rather simple and requires improvements in future studies. On the other hand, the most important feature of the methods is their computational efficiency. This means that during the solution process of a MOO problem, the decision maker does not have to wait in order to obtain, for example, vectors in the approximation related to his or her preferences. In other words, computational efficiency has been considered in this thesis, even though the problem can be computationally expensive.

In this thesis, the most important aspect is the communication between the IHBE method considered as a computer software and a human decision maker.

This aspect has been taken into account here and discussed, in detail, from a human decision maker's point of view. This thesis does not develop new communication methods, instead well-known MCDM methods/techniques and empirically proved and studied methods such as aspiration and reservation levels were used for the communication. For this reason, the IHBE method should suit for most decision makers. In addition, the HBE method allows the decision maker to decide what information should be presented in different phases of the solution process. More formally, the information provided by the IHBE method is focused to the learning phase in the solution process. This allows the decision maker to learn about his or her preferences in relation to what kinds of approximated Pareto optimal objective vectors are available. In this way, the IHBE method has been aimed at supporting and helping the decision maker to find a compromise solution based on the available (approximated) Pareto optimal objective vectors.

The approximation used in the IHBE method is rather simple, but innovative. As pointed out in this thesis, the simplicity of the approximation is remarkable, for example, in calculating approximated trade-off rates through the normal vectors or constructing bounds for nondominated vectors in the approximation. In addition, the computational efficiency is another reason that makes these properties essential and significant. For example, to calculate normal vectors, typically KKT-coefficients that are solutions of a system of nonlinear equations are calculated in the literature. There is no need with the help of the approximation to calculate the KKT-coefficients. This means that the approximation has opened new doors for methods originally developed only for computationally inexpensive problems and can now be used in computationally expensive MOO problems.

There is a large potential use for the IHBE method and the approximation in various MOO problems because it does not matter whether the problem is convex (expensive) or not. This is why the generality of the method and the approximation have been the targets of this thesis. However, this thesis does not present any real-life examples in which the IHBE method has been used leaving it for future studies. This thesis does not consider a graphical interface for the IHBE method. This is why the thesis does not contain any usability analysis or opinions from real-life decision makers about the method or how it may help one in searching for a compromise solution to a MOO problem under investigation. However, a preliminary software exists (not discussed in this thesis) to prove that the approximation can be constructed in practice and used in the solution process. Thus, a natural future sequel related to this thesis is the software and its completion for testing. In addition, the current simple methods for investigating, for example, whether or not a hyperbox is optimistic or pessimistic must be improved in future studies in order to make the methods more accurate.

YHTEENVETO (FINNISH SUMMARY)

Tässä väitöskirjassa, suomenkieliseltä nimeltään "Menetelmä laskennallisesti vaativille monitavoiteoptimointitehtäville", esitetään interaktiivinen menetelmä IHBE (Interactive HyperBox Exploration). Käytännön tarve menetelmälle tulee korostumaan lähitulevaisuudessa, koska laskennallisesti raskaat simulaattorit yleistyvät teollisuudessa yhä monimutkaisempien kokonaisuuksien matemaattisessa mallinnuksessa. Käytännössä tämä monimutkaisuus ilmenee kasvavana laskenta-aikana, jonka vuoksi tehtävää kutsutaan laskennalliseksi vaativaksi/raskaaksi. Tyypillisesti simulaattorit eivät sisällä optimointiin ja parhaan kompromissiratkaisun valintaan eri tavoitteiden välillä liittyviä työkaluja. Kun jokin sopiva monitavoiteoptimointimenetelmä on valittu tuottamaan kompromissiratkaisuja simulaattorin rinnalle, on IHBE-menetelmä suora jatke näille. Se tarjoaa joustavan tavan tarkastella simulaattorin ja monitavoiteoptimointimenetelmän tuottamia tuloksia. Tällöin päätöksentekijä (tehtäväalan asiantuntija) voi löytää parhaan kompromissiratkaisun tehtävän laskennallisesta vaativuudesta huolimatta.

IHBE-menetelmän kehittämisessä on erityistä huomiota kiinnitetty kolmeen tutkimuskohteeseen: menetelmän teoreettiseen soveltuvuuteen epäkonvekseille jatkuville monitavoiteoptimointitehtäville, menetelmän laskennalliseen tehokkuuteen sekä päätöksentekijän ja menetelmän vuorovaikutukseen. Monimutkaiset matemaattiset mallit simulaattoreiden osana muodostavat tyypillisesti monitavoiteoptimoinnissa laskennallisesti vaativia epäkonvekseja jatkuvia tehtäviä. Käytännössä tämä tarkoittaa, että yhdenkin ratkaisun tuottaminen saattaa kestää minuuteista tunteihin mikä saattaa muodostua ongelmaksi päätöksentekijälle, koska hän voi turhautua odottamiseen. IHBE-menetelmässä tämä ongelma ratkaistiin kehittämällä erityinen approksimaatio, joka matemaattisesti osoitettiin laskennallisesti tehokkaaksi ja soveltuvaksi kyseisille tehtäville. Approksimaatiota käyttäen tarvittavat laskenta-ajat IHBE-menetelmässä jäävät sekunteihin, jolloin päätöksentekijän ei enää tarvitse odottaa häntä kiinnostavien approksimoitujen ratkaisujen tuottamista.

Päätöksentekijän ja IHBE-menetelmän vuorovaikutus on oleellisessa asemassa parhaan kompromissiratkaisun valintaprosessissa. IHBE-menetelmän on kyettävä ymmärtämään minkälaista ratkaisua päätöksentekijä on etsimässä. Toisaalta, IHBE-menetelmän antamat tulokset saattavat vaikuttaa parhaan kompromissiratkaisun valintaan. Hyvän vuorovaikutuksen saamiseksi ja mahdollistamiseksi, IHBE-menetelmä koostuu kirjallisuudessa empiirisesti hyviksi todetuista menetelmistä. Pienillä muutoksilla menetelmät saatiin yhteensopiviksi aiemmin mainitun approksimaation kanssa. Muunnoksissa kuitenkin huomioitiin ettei menetelmien perusluonne muutu ja ettei menetelmien tarvitsemat laskenta-ajat kasva kohtuuttomasti.

Työssä kehitetty approksimaatio on yksinkertainen, jonka vuoksi sitä voidaan helposti soveltaa haastaviin monitavoiteoptimointitehtäviin, mikä tekee siitä käytännössä arvokkaan. Approksimaation avulla on mahdollista tuoda päätöksenteon tukemiseen oleellisia havainnollistuksia kuten esimerkiksi vaihtosuhteet,

mikä on aiemmin ollut haasteellista ellei jopa mahdotonta laskennallisesti vaativien tehtävien kohdalla. Tämän lisäksi yksinkertaisuus on avannut uusia innovatiivisia tarkastelutapoja tehtävän kompromissiratkaisujen tulkintoihin, minkä vuoksi jatkotutkimus on perusteltua syvällisempien ominaisuuksien esiin tuomiseen niin käytännössä kuin teoriassa.

REFERENCES

[1] GLPK (GNU Linear Programming Kit). Last accessed: July 16, 2012, http://www.gnu.org/software/glpk/.

[2] GNU Octave. Last accessed: July 16, 2012, http://www.gnu.org/software/octave/.

[3] IBM ILOG CPLEX optimization studio home page. Last accessed: July 16, 2012, http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/.

[4] The MathWorks Inc. Optimization toolbox™. Last accessed: July 16, 2012, http://www.mathworks.se/products/optimization/.

[5] T. Aittokoski. *On Challenges of Simulation-Based Global and Multiobjective Optimization*. PhD thesis, University of Jyväskylä, 2009.

[6] T. Aittokoski and K. Miettinen. Cost effective simulation-based multiobjective optimization in the performance of an internal combustion engine. *Engineering Optimization*, 40(7):593–612, 2008.

[7] T. Aittokoski and K. Miettinen. Efficient evolutionary approach to approximate the Pareto-optimal set in multiobjective optimization, UPS-EMOA. *Optimization Methods and Software*, 25(6):841–858, 2010.

[8] T. Aittokoski and S. Tarkkanen. User preference extraction using dynamic query sliders in conjunction with UPS-EMO algorithm. *ArXiv e-prints*, 2011. http://arxiv.org/abs/1110.5992v1.

[9] Y. Aksoy, T. W. Butler, and E. D. Minor. Comparative studies in interactive multiple objective mathematical programming. *European Journal of Operational Research*, 89(2):408–422, 1996.

[10] N. M. Alexandrov, J. E. Dennis, R. M. Lewis, and V. Torczon. A trust-region framework for managing the use of approximation models in optimization. *Structural and Multidisciplinary Optimization*, 15:16–23, 1998.

[11] R. Aslett, R. J. Buck, S. G. Duvall, J. Sacks, and W. J. Welch. Circuit optimization via sequential computer experiments: Design of an output buffery. *Journal of the Royal Statistical Society: Series C: Applied Statistics*, 47(1):31–48, 1998.

[12] S. Äyrämö. *Knowledge Mining and Using Robust Clustering*. PhD thesis, University of Jyväskylä, 2006.

[13] H. Aytug and S. Sayin. Using support vector machines to learn the efficient set in multiple objective discrete optimization. *European Journal of Operational Research*, 193(2):510–519, 2009.

[14] V. Belton, J. Branke, P. Eskelinen, S. Greco, J. Molina, F. Ruiz, and R. Slowiński. Interactive multiobjective optimization from a learning perspective. In J. Branke, K. Deb, K. Miettinen, and R. Słowiński, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, pages 213–243. Springer-Verlag, Berlin, Heidelberg, 2008.

[15] R. Benayoun, J. de Montgolfier, J. Tergny, and O. Laritchev. Linear programming with multiple objective functions: Step method (STEM). *Mathematical Programming*, 1(1):366–375, 1971.

[16] H. P. Benson and S. Sayin. Towards finding global representations of the efficient set in multiple objective mathematical programming. *Naval Research Logistics*, 44(1):47–67, 1997.

[17] W. E. Biles, J. P. C. Kleijnen, W. C. M. van Beers, and I. van Nieuwenhuyse. Kriging metamodeling in constrained simulation optimization: An explorative study. In S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, editors, *WSC '07: Proceedings of the 39th conference on Winter simulation*, pages 355–362. IEEE Press, 2007.

[18] M. Björkman and K. Holmström. Global optimization of costly nonconvex functions using radial basis functions. *Optimization and Engineering*, 1(4):373–397, 2000.

[19] G. E. P. Box and K. B. Wilson. On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 13(1):1–45, 1951.

[20] J. Branke, K. Deb, K. Miettinen, and R. Słowiński, editors. *Multiobjective Optimization: Interactive and Evolutionary Approaches*. Springer, Berlin, Heidelberg, 2008.

[21] P. Bratley and B. L. Fox. Algorithm 659: Implementing Sobol's quasirandom sequence generator. *ACM Transactions on Mathematical Software*, 14(1):88–100, 1988.

[22] J. Buchanan and L. Gardiner. A comparison of two reference point methods in multiple objective mathematical programming. *European Journal of Operational Research*, 149(1):17–34, 2003.

[23] J. T. Buchanan. An experimental evaluation of interactive MCDM methods and the decision making process. *Operational Research Society*, 45(9):1050–1059, 1994.

[24] J. T. Buchanan. A naive approach for solving MCDM problems: the GUESS method. *Journal of Operational Research Society*, 48(2):202–206, 1997.

[25] J. T. Buchanan and J. L. Corner. The effects of anchoring in interactive MCDM solution methods. *Computers and Operational Research*, 24(10):907–918, 1997.

[26] J. T. Buchanan, E. J. Henig, and M. I. Henig. Objectivity and subjectivity in the decision making process. *Annals of Operations Research*, 80:333–345, 1998.

[27] O. Buchtala, M. Klimek, and B. Sick. Evolutionary optimization of radial basis function classifiers for data mining applications. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 35(5):928–947, 2005.

[28] M. Buhmann. *Radial Basis Functions: Theory and Implementations*. Cambridge University Press, 2003.

[29] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[30] V. Chankong and Y. Y. Haimes. *Multiobjective Decision Making. Theory and Methodology*. North-Holland, New York, 1983.

[31] A. Charnes and W. W. Cooper. *Management Models and Industrial Applications of Linear Programming*. Wiley, New York, 1961.

[32] K.-Y. Chen and C.-H. Wang. Support vector regression with genetic algorithms in forecasting tourism demand. *Tourism Management*, 28(1):215–226, 2007.

[33] C. A. Coello Coello, G. B. Lamont, and D. A. Van Veldhuizen. *Evolutionary Algorithms for Solving Multi-Objective Problems (Genetic and Evolutionary Computation)*. Springer-Verlag New York, Secaucus, NJ, USA, 2006.

[34] J. Corner, J. Buchanan, and M. Henig. Dynamic decision problem structuring. *Journal of Multi-Criteria Decision Analysis*, 10(3):129–141, 2001.

[35] J. L. Corner and P. D. Corner. Characteristics of decisions in decision analysis practice. *The Journal of the Operational Research Society*, 46(3):304–314, 1995.

[36] J.L. Corner and J. T. Buchanan. Capturing decision maker preference: Experimental comparison of decision analysis and MCDM techniques. *Operational Research*, 98(98):85–97, 1997.

[37] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

[38] R. Courant and F. John. *Introduction to Calculus and Analysis I*. Springer-Verlag, New York, 1989.

[39] N. Cowan. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1):87–114, 2001.

[40] N. Cressie. The origins of kriging. *Mathematical Geology*, 22(3):239–252, 1990.

162

[41] C. Currin, T. Mitchell, M. Morris, and D. Ylvisaker. A Bayesian approach to the design and analysis of computer experiments. Technical Report ORNL-6498, OAK Ridge National Laboratory, 1988.

[42] I. Das and J. Dennis. Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems. *SIAM Journal on Optimization*, 8(3):631–657, 1998.

[43] K. Deb. *Multi-Objective Optimization using Evolutionary Algorithms*. John Wiley & Sons, 2001.

[44] K. Deb. Introduction to evolutionary multiobjective optimization. In J. Branke, K. Deb, K. Miettinen, and R. Slowinski, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, pages 59–96. Springer Berlin, Heidelberg, 2008.

[45] H. Ding, Y. Xiao, and J. Yue. Adaptive training of radial basis function networks using particle swarm optimization algorithm. In L. Wang, K. Chen, and Y.S. Ong, editors, *Advances in Natural Computation*, pages 119–128. Springer, Berlin Heidelberg, 2005.

[46] M. Ehrgott. *Multicriteria Optimization*. Springer, Berlin, 2005.

[47] M. Ehrgott and X. Gandibleux. Bound sets for biobjective combinatorial optimization problems. *Computers & Operations Research*, 34(9):2674–2694, 2007.

[48] M. Ehrgott, L. Shao, and A. Schöbel. An approximation algorithm for convex multi-objective programming problems. *Journal of Global Optimization*, 50:397–416, 2011.

[49] P. Eskelinen and K. Miettinen. Trade-off analysis approach for interactive nonlinear multiobjective optimization. *OR Spectrum*, 34(4):803–816, 2012.

[50] P. Eskelinen, K. Miettinen, K. Klamroth, and J. Hakanen. Pareto navigator for interactive nonlinear multiobjective optimization. *OR Spectrum*, 32(1):211–227, 2010.

[51] S. French. Interactive multi-objective programming: Its aims, applications and demands. *Journal of the Operational Research Society*, 35(9):827–834, 1984.

[52] S. Gass and T. Saaty. The computational algorithm for the parametric objective function. *Naval Research Logistics Quarterly*, 2(1–2):39–45, 1955.

[53] A. Gelman and J. Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, 2006.

[54] A. M. Geoffrion, J. S. Dyer, and A. Feinberg. An interactive approach for multi-criterion optimization, with an application to the operation of an academic department. *Management Science*, 19(4):357–368, 1972.

[55] A. A. Giunta and L. T. Watson. A comparison of approximation modeling techniques: Polynomial versus interpolating models. Technical Report NASA-aiaa-98-4758, NASA Langley Technical Report Server, 1998.

[56] T. Goel, R. Vaidyanathan, R. T. Haftka, W. Shyy, N. V. Queipo, and K. Tucker. Response surface approximation of Pareto optimal front in multi-objective optimization. *Computer Methods in Applied Mechanics and Engineering*, 196(4-6):879–893, 2007.

[57] S. Greco, B. Matarazzo, and R. Slowiński. Dominance-based rough set approach to interactive multiobjective optimization. In J. Branke, K. Deb, K. Miettinen, and R. Slowiński, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, pages 121–155. Springer-Verlag, Berlin, Heidelberg, 2008.

[58] H.-M. Gutmann. A radial basis function method for global optimization. *Journal of Global Optimization*, 19(3):201–227, 2001.

[59] Y. Y. Haimes, L. S. Lasdon, and D. A. Wismer. On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Transactions on Systems, Man and Cybernetics*, 1(3):296–297, 1971.

[60] J. Hakanen. *On Potential of Interactive Multiobjective Optimization in Chemical Process Design*. PhD thesis, University of Jyväskylä, 2006.

[61] J. Hakanen, K. Miettinen, and K. Sahlstedt. Wastewater treatment: New insight provided by interactive multiobjective optimization. *Decision Support Systems*, 51(2):328–337, 2011.

[62] J Hakanen, Kawajiri Y., K. Miettinen, and L.T. Biegler. Interactive multiobjective optimization for simulated moving bed processes. *Control and Cybernetics*, 36(2):282–320, 2007.

[63] J. H. Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik*, 2:84–90, 1960.

[64] T. Hanne. Meta decision problems in multiple criteria decision making. In T. Gal, T. J. Stewart, and T. Hanne, editors, *Multicriteria Decision Making: Advances in MCDM Models, Algorithms, Theory, and Applications*, pages 6–1–6–25. Kluwer Academic Publishers, 1999.

[65] M. Hartikainen. *Approximation Through Interpolation in Nonconvex Multiobjective Optimization*. PhD thesis, University of Jyväskylä, 2011.

[66] M. Hartikainen, K. Miettinen, and M. Wiecek. Constructing a Pareto front approximation for decision making. *Mathematical Methods of Operations Research*, 73:209–234, 2011.

[67] M. Hartikainen, K. Miettinen, and M. Wiecek. PAINT: Pareto front interpolation for nonlinear multiobjective optimization. *Computational Optimization and Applications*, 52(3):845–867, 2012.

[68] M. Hartikainen and V. Ojalehto. Demonstrating the applicability of PAINT to computationally expensive real-life multiobjective optimization. *ArXiv e-prints*, 2011. http://arxiv.org/pdf/1109.3411v1.pdf.

[69] S. Haykin. *Neural Networks and Learning Machines*. Pearson Education, 3 edition, 2009.

[70] J. C. Henig and J. T. Buchanan. Tradeoff directions in multiobjective optimization problems. *Mathematical Programming*, 78(3):357–374, 1997.

[71] M. I. Henig and J. T. Buchanan. Solving MCDM problems: Process concepts. *Journal of Multi-Criteria Decision Analysis*, 5(1):3–21, 1996.

[72] S. Herbert. How big is a chunk? *Science*, 183:482–488, 1957.

[73] A. Hofmann and B. Sick. Evolutionary optimization of radial basis function networks for intrusion detection. *2003. Proceedings of the International Joint Conference on Neural Networks*, 1:415–420, 2003.

[74] K. Holmström. An adaptive radial basis algorithm (ARBF) for expensive black-box global optimization. *Journal of Global Optimization*, 41(3):447–464, 2008.

[75] K. Holmström, N.-H. Quttineh, and M. M. Edvall. An adaptive radial basis algorithm (ARBF) for expensive black-box mixed-integer constrained global optimization. *Optimization and Engineering*, 9(4):311–339, 2008.

[76] S. Hosder, L. T. Watson, B. Grossman, W. H. Mason, H. Kim, R. T. Haftka, and S. E. Cox. Polynomial response surface approximations for the multidisciplinary design optimization of a high speed civil transport. *Optimization and Engineering*, 2(4):431–452, 2001.

[77] W. Hu, L. Enying, G.Y. Li, and Z.H. Zhong. Development of metamodeling based optimization system for high nonlinear engineering problems. *Advances in Engineering Software*, 39(8):629–645, 2008.

[78] D. Huang, T. T. Allen, W. I. Notz, and R. A. Miller. Sequential kriging optimization using multiple-fidelity evaluations. *Structural and Multidisciplinary Optimization*, 32(4):369–382, 2006.

[79] D. Huang, T. T. Allen, W. I. Notz, and N. Zeng. Global optimization of stochastic black-box systems via sequential kriging meta-models. *Journal of Global Optimization*, 34(0):441–466, 2006.

[80] E. Hughes. Many objective optimisation: Direct objective boundary identification. In G. Rudolph, T. Jansen, S. Lucas, C. Poloni, and N. Beume, editors, *Parallel Problem Solving from Nature - PPSN X*, volume 5199, pages 733–742. Springer Berlin, Heidelberg, 2008.

[81] M. F. Hussain, R. R. Barton, and S. B. Joshi. Metamodeling: Radial basis functions, versus polynomials. *European Journal of Operational Research*, 138(1):142–154, 2002.

[82] P. Iz and L. Krajewski. Comparative evaluation of three interactive multiobjective programming techniques as group decision support tools. *INFOR*, 30(4):349–363, 1992.

[83] A. Jaszkiewicz and R. Slowiński. The 'light beam search' approach - an overview of methodology and applications. *European Journal of Operational Research*, 113(2):300–314, 1999.

[84] S. Jeong and S. Obayashi. Efficient global optimization (EGO) for multiobjective problem and data mining. In *The 2005 IEEE Congress on Evolutionary Computation*, volume 3, pages 2138–21453, 2005.

[85] R. Jin, W. Chen, and T. W. Simpson. Comparative studies of metamodelling techniques under multiple modelling criteria. *Structural and Multidisciplinary Optimization*, 23(1):1–13, 2001.

[86] D. R. Jones. A taxonomy of global optimization methods based on response surfaces. *Journal of Global Optimization*, 21(4):345–383, 2001.

[87] D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):445–492, 1998.

[88] J. R. Kalagnanam and U. M. Diwekar. An efficient sampling technique for off-line quality control. *Technometrics*, 39(3):308–319, 1997.

[89] I. Kaliszewski and W. Michalowski. Searching for psychologically stable solutions of multiple criteria decision problems. *European Journal of Operational Research*, 118(3):549–562, 1999.

[90] E. Karasakal and M. Köksalan. Generating a representative subset of the nondominated frontier in multiple criteria decision making. *Operations Research*, 57(1):187–199, 2009.

[91] E. Kasanen, H. Wallenius, J. Wallenius, and S. Zionts. A study of high-level managerial decision processes, with implications for MCDM research. *European Journal of Operational Research*, 120(3):496–510, 2000.

[92] R. L. Keeney. Common mistakes in making value trade-offs. *Operations Research*, 50(6):935–945, 2002.

[93] M. C. Kennedy and A. O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001.

[94] D. Kincaid and W. Cheney. *Numerical Analysis: Mathematics of Scientific Computing*. Brooks/Cole Publishing Co., third edition, 2002.

[95] K. Klamroth and K. Miettinen. Integrating approximations and interactive decision making in multicriteria optimization. *Operations Research*, 56(1):222–234, 2008.

[96] K. Klamroth, J. Tind, and M. Wiecek. Unbiased approximation in multicriteria optimization. *Mathematical Methods of Operations Research*, 56:413–437, 2002.

[97] J. Knowles. ParEGO: a hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Transactions on Evolutionary Computation*, 10(1):50–66, 2006.

[98] J. Knowles and E. J. Hughes. Multiobjective optimization on a budget of 250 evaluations. In C. A. Coello Coello et. al., editor, *Evolutionary Multi-Criterion Optimization*, volume 3410, pages 176–190. Springer, Berlin, Heidelberg, 2005.

[99] J. Knowles and H. Nakayama. Meta-modeling in multiobjective optimization. In J. Branke, K. Deb, K. Miettinen, and R. Slowiński, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, pages 245–284. Springer-Verlag, Berlin, Heidelberg, 2008.

[100] N. Kondo, T. Hatanaka, and K. Uosaki. Pattern classification by evolutionary RBF networks ensemble based on multi-objective optimization. In *International Joint Conference on Neural Networks 2006. IJCNN '06.*, pages 2919–2925, 2006.

[101] P. Korhonen and J. Laakso. Solving generalized goal programming problems using a visual interactive approach. *European Journal of Operational Research*, 26(3):355–363, 1986.

[102] P. Korhonen and J. Laakso. A visual interactive method for solving the multiple criteria problem. *European Journal of Operational Research*, 24(2):277–287, 1986.

[103] P. Korhonen and J. Wallenius. Observations regarding choice behaviour in interactive multiple criteria decision-making environments: an experimental investigation. In Lewandowski A. and Stanchev I., editors, *Methodology and Software for interactive decision support*, volume 337, pages 163–170. Springer-Verlag, 1989.

[104] P. Korhonen and J. Wallenius. Behavioral issues in MCDM: Neglected research questions. In J. Clímaco, editor, *Multicriteria Analysis*, pages 412–422. Springer-Verlag, Berlin, Heidelberg, 1997.

[105] P. Korhonen and J. Wallenius. Visualization in the multiple objective decision-making framework. In J. Branke, K. Deb, K. Miettinen, and R. Slowiński, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, pages 195–212. Springer-Verlag, Berlin, Heidelberg, 2008.

[106] H. W. Kuhn and A. W. Tucker. Nonlinear programming. In Neyman J., editor, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 481–492. University of California Press, Berkeley, 1951.

[107] O. I. Larichev. Cognitive validity in decision-aiding techniques. *Journal of Multi-Criteria Decision Analysis*, 1(3):127–138, 1992.

[108] S. J. Leary, A. Bhaskar, and A. J. Keane. A constraint mapping approach to the structural optimization of an expensive model using surrogates. *Optimization and Engineering*, 2(4):385–398, 2001.

[109] Y. Li, S. Gong, and H. Liddell. Support vector regression and classification based multi-view face detection and recognition. In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pages 300–305, 2000.

[110] I. Loshchilov, M. Schoenauer, and M. Sebag. A Pareto-compliant surrogate approach for multiobjective optimization. In *Proceedings of the 12th annual conference companion on Genetic and evolutionary computation*, GECCO '10, pages 1979–1982, New York, NY, USA, 2010. ACM.

[111] V. Lotfi, T. Stewart, and S. Zionts. An aspiration-level interactive model for multiple criteria decision making. *Computers & Operations Research*, 19(7):671–681, 1992.

[112] A. V. Lotov and K. Miettinen. Visualizing the Pareto frontier. In J. Branke, K. Deb, K. Miettinen, and R. Slowiński, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, pages 213–243. Springer-Verlag Berlin, Heidelberg, 2008.

[113] A.V. Lotov, V.A. Bushenkov, and G.K. Kamenev. *Interactive Decision Maps*. Kluwer Academic Publishers, 2004.

[114] A. Lovison. Singular continuation: Generating piecewise linear approximations to Pareto sets via global analysis. *SIAM Journal on Optimization*, 21(2):463–490, 2011.

[115] H. Maaranen. *On Heuristic Hybrid Methods and Structured Point Sets in Global Continuous Optimization*. PhD thesis, University of Jyväskylä, 2004.

[116] N. K. Malhotra. Information load and consumer decision making. *Journal of Consumer Research*, 8(4):419–430, 1982.

[117] J. March and H. A. Simon. *Organizations*. Wiley, New York, 1958.

[118] J. Martin, C. Bielza, and D. Rios Insua. Approximating nondominated sets in continuous multiobjective optimization problems. *Naval Research Logistics*, 52(5):469–480, 2005.

[119] D. B. McDonald, W. J. Grantham, W. L. Tabor, and M. J. Murphy. Global and local optimization using radial basis function response surface models. *Applied Mathematical Modelling*, 31(10):2095–2110, 2007.

[120] J. Mercer. Functions of positive and negative type, and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society of London. Series A*, 209:415–446, 1909.

[121] A. Messac, A. Ismail-Yahaya, and C.A. Mattson. The normalized normal constraint method for generating the Pareto frontier. *Structural and Multidisciplinary Optimization*, 25:86–98, 2003.

[122] W. Michalowski. MCDM at the crossroads. In J. Clímaco, editor, *Multicriteria Analysis*, pages 579–584. Springer-Verlag Berlin, Heidelberg, 1997.

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

[124] K. Miettinen and M.M. Mäkelä. Interactive bundle-based method for nondifferentiable multiobjeective optimization: NIMBUS. *Optimization*, 34(3):231–246, 1995.

[125] K. Miettinen and M.M. Mäkelä. Comparative evaluation of some interactive reference point-based methods for multi-objective optimisation. *Journal of the Operational Research Society*, 50(9):949–959, 1999.

[126] K. Miettinen and M.M. Mäkelä. Interactive multiobjective optimization system WWW-NIMBUS on the internet. *Computers & Operations Research*, 27(7-8):709–723, 2000.

[127] K. Miettinen and M.M. Mäkelä. On scalarization functions in multiobjective optimization. *OR Spectrum*, 24(2):193–213, 2002.

[128] K. Miettinen and M.M. Mäkelä. On trade-off directions in nonconvex multiobjective optimization. *Mathematical Programming*, 92(1):141–151, 2002.

[129] K. Miettinen and M.M. Mäkelä. Synchronous approach in interactive multiobjective optimization. *European Journal of Operational Research*, 170(3):909–922, 2006.

[130] K. Miettinen, M.M. Mäkelä, and K. Kaario. Experiments with classification-based scalarizing functions in interactive multiobjective optimization. *European Journal of Operational Research*, 175(2):931–947, 2006.

[131] K. Miettinen, F. Ruiz, and A. P. Wierzbicki. Introduction to multiobjective optimization: Interactive approaches. In J. Branke, K. Deb, K. Miettinen, and R. Slowiński, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, pages 27–57. Springer-Verlag, Berlin, Heidelberg, 2008.

[132] G. A. Miller. The magical number seven, plus or minus two some limits on our capacity for processing information. *Psychological Review*, 63(2):343–352, 1955.

[133] M. Monz. *Pareto Navigation - Algorithmic Foundation of Interactive Multicirteria IMRT Planning*. PhD thesis, University of Kaiserslautern, 2006.

[134] A. A. Mullur and A. Messac. Metamodeling using extended radial basis functions: a comparative approach. *Engineering with Computers*, 21(3):203–217, 2006.

[135] H. Nakayama. Aspiration level approach to interactive multi-objective programming and its applications. In P. M. Pardalos, Y. Siskos, and C. Zopounidis, editors, *Advances in Multicriteria Analysis*, pages 147–174. Kluwer Academic Publishers, Dordrecht, 1995.

[136] H. Nakayama and Y. Sawaragi. Satisficing trade-off method for multiobjective programming. In M. Grauer and A. P. Wierzbicki, editors, *Interactive Decision Analysis*, pages 113–122. Springer, Heidelberg, 1984.

[137] H. Nakayama and Y. Yun. Support vector regression based on goal programming and multi-objective programming. In *International Joint Conference on Neural Networks, IJCNN '06*, pages 1156–1161, 2006.

[138] H. Nakayama, Y. Yun, T. Asada, and M. Yoon. A family of support vector machines using MOP/GP. In *MCDM, Whistler, B. C., Canada, August 6-11*, 2004.

[139] H. Nakayama, Y. Yun, and M. Yoon. *Sequential Approximate Multiobjective Optimization Using Computational Intelligence*. Springer, Berlin, Heidelberg, 2009.

[140] K. S. Ni and T. Q. Nguyen. Image superresolution using support vector regression. *IEEE Transactions on Image Processing*, 16(6):1596–1610, 2007.

[141] H. B. Nielsen, S. N. Lophaven, and J. Søndergaard. DACE - a matlab kriging toolbox, 2002. Software available at http://www2.imm.dtu.dk/~hbn/dace/.

[142] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, 1999.

[143] A. L. I. Oliveira. Estimation of software project effort with support vector regression. *Neurocomputing*, 69(13-15):1749–1753, 2006.

[144] K. Palmer and M. Realff. Optimization and validation of steady-state flow-sheet simulation metamodels. *Chemical Engineering Research and Design*, 80(7):773–782, 2002.

[145] S. Poles, M. Vassileva, and D. Sasaki. Multiobjective optimization software. In J. Branke, K. Deb, K. Miettinen, and R. Slowiński, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, pages 213–243. Springer-Verlag, Berlin, Heidelberg, 2008.

[146] A. Przybylski, X. Gandibleux, and M. Ehrgott. A two phase method for multi-objective integer programming and its application to the assignment problem with three objectives. *Discrete Optimization*, 7(3):149–165, 2010.

[147] N. V. Queipo, S. Pintos, N. Rincón, N. Contreras, and J. Colmenares. Surrogate modeling-based optimization for the integration of static and dynamic data into a reservoir description. *Journal of Petroleum Science and Engineering*, 35(3–4):167–181, 2002.

[148] N. V. Queipo, A. J. Verde, J. Canelón, and S. Pintos. Efficient global optimization for hydraulic fracturing treatment design. *Journal of Petroleum Science and Engineering*, 35(3–4):151–166, 2002.

[149] P. A. Ramachandran and K. Balakrishnan. Radial basis functions as approximate particular solutions: review of recent progress. *Engineering Analysis with Boundary Elements*, 24(7-8):575–582, 2000.

[150] R. G. Regis and C. A. Shoemaker. Constrained global optimization of expensive black box functions using radial basis functions. *Journal of Global Optimization*, 31(1):153–171, 2005.

[151] R. G. Regis and C. A. Shoemaker. Improved strategies for radial basis function methods for global optimization. *Journal of Global Optimization*, 37(1):113–135, 2007.

[152] R. G. Regis and C. A. Shoemaker. Parallel radial basis function methods for the global optimization of expensive functions. *European Journal of Operational Research*, 182(2):514–535, 2007.

[153] G. Rennen, E. R. van Dam, and D. den Hertog. Enhancement of sandwich algorithms for approximating higher-dimensional convex Pareto sets. *INFORMS journal on Computing*, 23(4):493–517, 2011.

[154] E. Rigoni. Hole functions problem. Technical report, Esteco 2004-002, 2004.

[155] A. J. Rivera, I. Rojas, J. Ortega, and M. J. del Jesus. A new hybrid methodology for cooperative-coevolutionary optimization of radial basis function networks. *Soft Computing*, 11(7):655–668, 2007.

[156] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, New Jersey, 1970.

[157] J. F. Rodriguez, J. E. Renaud, B. A. Wujek, and R. V. Tappeta. Trust region model management in multidisciplinary design optimization. *Journal of Computational and Applied Mathematics*, 124(1-2):139–154, 2000.

[158] H. Ruotsalainen, K. Miettinen, J.-E. Palmgren, and T. Lahtinen. Interactive multiobjective optimization for anatomy-based three-dimensional HDR brachytherapy. *Physics in Medicine and Biology*, 55(16):4703–4719, 2010.

[159] S. Ruzika and M. M. Wiecek. Approximation methods in multiobjective programming. *Journal of Optimization Theory and Applications*, 126(3):473–501, 2005.

[160] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn. Design and analysis of computer experiments. *Statistical Science*, 4(4):409–423, 1989.

[161] S. Sakata, F. Ashida, and M. Zako. Structural optimization using kriging approximation. *Computer Methods in Applied Mechanics and Engineering*, 192(7-8):923–939, 2003.

[162] S. Sakata, F. Ashida, and M. Zako. Approximate structural optimization using kriging method and digital modeling technique considering noise in sampling data. *Computers and Structures*, 86(13-14):1477–1485, 2008.

[163] M. Sakawa and H. Yano. Trade-off rates in the hyperplane method for multiobjective optimization problems. *European Journal of Operational Research*, 44(1):105–118, 1990.

[164] L.V. Santana-Quintero, V.A. Serrano-Hernandez, C.A. Coello Coello, A.G. Hernandez-Diaz, and J. Molina. Use of radial basis functions and rough sets for evolutionary multi-objective optimization. In *IEEE Symposium on Computational Intelligence in Multicriteria Decision Making*, pages 107–114, 2007.

[165] Y. Sawaragi, H. Nakayama, and T. Tanino. *Theory of Multiobjective Optimization*. Academic Press, 1985.

[166] S. Sayin. A procedure to find discrete representations of the efficient set with specified coverage errors. *Operations Research*, 51(3):427–436, 2003.

[167] R. Scheubrein and S. Zionts. A problem structuring front end for a multiple criteria decision support system. *Computers & Operations Research*, 33(1):18–31, 2006.

[168] B. Schölkopf and A. J. Smola. *Learning with Kernels*. Massachusetts Institute of Technology, 2002.

[169] A. Shahsavand and A. Ahmadpour. Application of optimal RBF neural networks for optimization and characterization of porous materials. *Computers and Chemical Engineering*, 29(10):2134–2143, 2005.

[170] A. Y. D. Siem, D. den Hertog, and A. L. Hoffmann. The effect of transformations on the approximation of univariate (convex) functions with applications to Pareto curves. *European Journal of Operational Research*, 189(2):347–362, 2008.

[171] T. W. Simpson, T. M. Mauery, J. J. Korte, and F. Mistree. Comparison of response surface and kriging models for multidisciplinary design optimization. Technical Report 98–4755, AIAA, 1998.

[172] D.R. Smart. *Fixed Point Theorems*. Cambridge University Press, New York, 1975.

[173] A. J. Smola and B. Schölkopf. A tutorial on support vector regression. *Journal of Statistics and Computing*, 14(3):199–222, 2004.

[174] A. Sóbester, S. J. Leary, and A. J. Keane. A parallel updating scheme for approximating and optimizing high fidelity computer simulations. *Structural and Multidisciplinary Optimization*, 27(5):371–383, 7 2004.

[175] M. Spivak. *Calculus on Manifolds: a Modern Approach to Classical Theorems of Advanced Calculus*. Westview press, 1965. twenty-seventh printing, 1998.

[176] R. E. Steuer. *Multiple Criteria Optimization: Theory, Computation and Application*. John Wiley, New York, 1986.

[177] T. J. Stewart. A critical survey on the status of multiple criteria decision making theory and practice. *Omega*, 20(5/6):569–586, 1992.

[178] T. J. Stewart. Concepts of interactive progamming. In T. Gal, T. J. Stewart, and T. Hanne, editors, *Multicriteria Decision Making: Advances in MCDM Models, Algorithms, Theory, and Applications*, pages 10–1–10–28. Kluwer Academic Publishers, 1999.

[179] T. J. Stewart. Goal programming and cognitive biases in decision-making. *Operational Research Society*, 56(10):1166–1175, 2005.

[180] Y. F. Sun, Y. C. Liang, W. L. Zhang, H. P. Lee, W. Z. Lin, and L. J. Cao. Optimal partition algorithm of the RBF neural network and its application to financial time series forecasting. *Neural Computing and Applications*, 14(1):36–44, 2005.

[181] V. N. Vapnik. *Statistical Learning Theory*. Wiley-Interscience, 1998.

[182] I. Voutchkov and A.J. Keane. Multiobjective optimization using surrogates. In *Proceedings of the 7th International Conference on Adaptive Computing in Design and Manufacture*, pages 167–175. The M.C.Escher Company, 2006.

[183] J. Wallenius. Comparative evaluation of some interactive approaches to multicriterion optimization. *Management Science*, 21(12):1387–1396, 1975.

[184] J. Wallenius, J. S. Dyer, P. C. Fishburn, R. E. Steuer, S. Zionts, and K. Deb. Multiple criteria decision making, multiattribute utility theory: Recent accomplishments and what lies ahead. *Management Science*, 54(7):1336–1349, 2008.

[185] J. Wang and S. Zionts. The aspiration level interactive method (AIM) reconsidered: Robustness of solutions. *European Journal of Operational Research*, 175(2):948–958, 2006.

[186] X. Wang, A. Li, Z. Jiang, and H. Feng. Missing value estimation for DNA microarray gene expression data by support vector regression imputation and orthogonal coding scheme. *BMC Bioinformatics*, 7(1):32, 2006.

[187] A. P. Wierzbicki. The use of reference objectives in multiobjective optimization. In G. Fandel and T. Gal, editors, *Multiple Criteria Decision Making: Theory and Application*, pages 431–445. Springer, Berlin, 1980.

[188] A. P. Wierzbicki. On the completeness and constructiveness of parametric characterizations to vector optimization problems. *Journal of Multi-Criteria Decision Analysis*, 8(2):73–87, 1986.

[189] A. P. Wierzbicki. On the role of intuition in decision making and some ways of multicriteria aid of intuition. *Journal of Multi-Criteria Decision Analysis*, 6(2):65–76, 1997.

[190] A. P. Wierzbicki. Multi-objective and reference point optimization tools. In A. P. Wierzbicki, M. Makowski, and J. Wessels, editors, *Model-Based Decision Support Methodology with Environmental Applications*. Kluwer Academic Publishers, 2000.

[191] B. Wilson, D. Cappelleri, T.W. Simpson, and M. Frecker. Efficient Pareto frontier exploration using surrogate approximations. *Optimization and Engineering*, 2:31–50, 2001.

[192] C.-H. Wu, J.-M. Ho, and D. T. Lee. Travel-time prediction with support vector regression. *Intelligent Transportation Systems*, 5(4):276–281, 2004.

[193] J. Yang and D. Li. Normal vector identification and interactive tradeoff analysis using minimax formulation in multiobjective optimization. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 32(3):305–319, 2002.

[194] P-L. Yu. *Multiple-Criteria Decision Making*. Mathematical Concepts and methods in Science and Engineering, Plenum Press New York, 1985.

[195] P.-S. Yu, S.-T. Chen, and I.-F. Chang. Support vector regression for real-time flood stage forecasting. *Journal of Hydrology*, 328(3-4):704–716, 2006.

174

[196] Y. Yun, M. Yoon, and H. Nakayama. Multi-objective optimization based on meta-modeling by using support vector regression. *Optimization and Engineering*, 10:167–181, 2009.

[197] Y. B. Yun, H. Nakayama, and M. Yoon. Sequential approximation method in multi-objective optimization using aspiration level approach. *SICE Transactions of the society of Instrument and Control Engineers*, 43(8):672–678, 2007.

[198] L. Zadeh. Optimality and non-scalar-valued performance criteria. *IEEE Transactions on Automatic Control*, 8(1):59–60, 1963.

[199] Y. Zheng and H. Bixing. Prediction of protein accessible surface areas by support vector regression. *Proteins: Structure, Function, and Bioinformatics*, 57(3):558–564, 2004.

[200] A. Zhou, Q. Zhang, and Y. Jin. Approximating the set of Pareto-optimal solutions in both the decision and objective spaces by an estimation of distribution algorithm. *IEEE Transactions on Evolutionary Computation*, 13(5):1167–1189, 2009.

[201] S. Zionts. Some thoughts on MCDM: Myths and ideas. In J. Clímaco, editor, *Multicriteria Analysis*, pages 602–607. Springer-Verlag, Berlin, Heidelberg, 1997.

[202] S. Zionts and J. Wallenius. An interactive programming method for solving the multiple criteria problem. *Management Science*, 22(6):652–663, 1976.

APPENDIX 1   DERIVATION OF NORMAL VECTORS

In Section 4.1, we introduced formula (21) for calculating normal vectors on the approximated Pareto front. In this appendix, we derive formula (21) and we follow the procedure presented in [175, Problem 5-13 on page 121].

In order to calculate normal vectors explicitly on $f_a(\Delta_{k-1})$, we first have to construct a function $g : \mathbf{ideal}^{Z_m} + \text{int}\,\mathbb{R}_+^k \to \mathbb{R}$ so that the preimage of $0 \in \mathbb{R}$ is $f_a(\Delta_{k-1})$ ($= M$, the notation used in [175]). Here, we define a function g so that

$$g(\mathbf{z}) = f_s(f_{Z_m}^p(\mathbf{z})) - \sum_{i=1}^{k} z_i - \text{ideal}_i^{Z_m}. \tag{38}$$

Next, we show that $g^{\{-1\}}(0) = f_a(\Delta_{k-1})$.

Theorem 61. Let $\mathbf{z} \in \mathbf{ideal}^{Z_m} + \text{int}\,\mathbb{R}_+^k$. Then $g(\mathbf{z}) = 0$ if and only if $\mathbf{z} \in f_a(\Delta_{k-1})$.

Proof. Let $\mathbf{z}^s = f_{Z_m}^p(\mathbf{z})$.

"$\Rightarrow$" Because $g(\mathbf{z}) = 0$, then

$$f_s(\mathbf{z}^s) = \sum_{i=1}^{k} z_i - \text{ideal}_i^{Z_m}.$$

It is clear that $\mathbf{z}^s \in \Delta_{k-1}$. Now,

$$f_a(\mathbf{z}^s) = \mathbf{ideal}^{Z_m} + f_s(\mathbf{z}^s)\mathbf{z}^s = \mathbf{ideal}^{Z_m} + \left(\sum_{i=1}^{k} z_i - \text{ideal}_i^{Z_m}\right) \frac{\mathbf{z} - \mathbf{ideal}^{Z_m}}{\sum_{j=1}^{k} z_j - \text{ideal}_j^{Z_m}}$$
$$= \mathbf{ideal}^{Z_m} + \mathbf{z} - \mathbf{ideal}^{Z_m} = \mathbf{z}.$$

In other words, $\mathbf{z} \in f_a(\Delta_{k-1})$.

"$\Leftarrow$" Because $\mathbf{z} \in f_a(\Delta_{k-1})$, then

$$\mathbf{z} = \mathbf{ideal}^{Z_m} + f_s(\mathbf{z}^s)\mathbf{z}^s = \mathbf{ideal}^{Z_m} + f_s(\mathbf{z}^s) \frac{\mathbf{z} - \mathbf{ideal}^{Z_m}}{\sum_{j=1}^{k} z_j - \text{ideal}_j^{Z_m}}.$$

On the other hand,

$$\mathbf{z} = \mathbf{ideal}^{Z_m} + \mathbf{z} - \mathbf{ideal}^{Z_m} = \mathbf{ideal}^{Z_m} + \left(\sum_{i=1}^{k} z_i - \text{ideal}_i^{Z_m}\right) \frac{\mathbf{z} - \mathbf{ideal}^{Z_m}}{\sum_{j=1}^{k} z_j - \text{ideal}_j^{Z_m}}.$$

In other words,

$$\sum_{i=1}^{k} z_i - \text{ideal}_i^{Z_m} = f_s(\mathbf{z}^s).$$

Thus $g(\mathbf{z}) = 0$.

$\square$

Now, according to [175, Problem 5-13 on page 121], the normal vector $f_a(\mathbf{z})^\perp$ is the gradient vector of g at $\tilde{\mathbf{z}} = f_a(\mathbf{z})$:

$$f_a(\mathbf{z})^\perp = \nabla g(f_a(\mathbf{z})) = \nabla g(\tilde{\mathbf{z}})$$

$$= \left[\frac{\partial f_s}{\partial \tilde{z}_1}(f^p_{Z_m}(\tilde{\mathbf{z}})) \quad \cdots \quad \frac{\partial f_s}{\partial \tilde{z}_k}(f^p_{Z_m}(\tilde{\mathbf{z}})) \right] \begin{bmatrix} \frac{\partial f^{p_1}_{Z_m}}{\partial \tilde{z}_1}(\tilde{\mathbf{z}}) & \cdots & \frac{\partial f^{p_1}_{Z_m}}{\partial \tilde{z}_k}(\tilde{\mathbf{z}}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f^{p_k}_{Z_m}}{\partial \tilde{z}_1}(\tilde{\mathbf{z}}) & \cdots & \frac{\partial f^{p_k}_{Z_m}}{\partial \tilde{z}_k}(\tilde{\mathbf{z}}) \end{bmatrix} - \mathbf{1}^T, \quad (39)$$

where

$$(f^{p_1}_{Z_m}(\tilde{\mathbf{z}}), \ldots, f^{p_k}_{Z_m}(\tilde{\mathbf{z}}))^T := f^p_{Z_m}(\tilde{\mathbf{z}}),$$

$$\frac{\partial f^{p_i}_{Z_m}}{\partial \tilde{z}_j}(\tilde{\mathbf{z}}) = \begin{cases} \dfrac{\sum_{\substack{l=1 \\ l \neq i}}^{k} \tilde{z}_l - \text{ideal}^{Z_m}_l}{\left(\sum_{l=1}^{k} \tilde{z}_l - \text{ideal}^{Z_m}_l\right)^2}, & i = j \\ -\dfrac{\tilde{z}_i - \text{ideal}^{Z_m}_i}{\left(\sum_{l=1}^{k} \tilde{z}_l - \text{ideal}^{Z_m}_l\right)^2}, & i \neq j \end{cases}.$$

As shown in Theorem 61, $\sum_{i=1}^{k} \tilde{z}_i - \text{ideal}^{Z_m}_i = f_s(f^p_{Z_m}(\tilde{\mathbf{z}}))$ and $\tilde{z}_i = \text{ideal}^{Z_m}_i + f_s(f^p_{Z_m}(\tilde{\mathbf{z}})) f^{p_i}_{Z_m}(\tilde{\mathbf{z}})$ for all $i = 1, \ldots, k$. Then partial derivatives of functions $f^{p_i}_{Z_m}$ in (39) can now be rewritten as follows

$$\frac{\partial f^{p_i}_{Z_m}}{\partial \tilde{z}_j}(\tilde{\mathbf{z}}) = \begin{cases} \dfrac{1 - f^{p_i}_{Z_m}(\tilde{\mathbf{z}})}{f_s(f^p_{Z_m}(\tilde{\mathbf{z}}))}, & i = j \\ -\dfrac{f^{p_i}_{Z_m}(\tilde{\mathbf{z}})}{f_s(f^p_{Z_m}(\tilde{\mathbf{z}}))}, & i \neq j \end{cases}.$$

Thus,

$$f_a(\mathbf{z})^\perp = \left[\frac{\partial f_s}{\partial \tilde{z}_1}(\mathbf{z}) \quad \cdots \quad \frac{\partial f_s}{\partial z_k}(\mathbf{z}) \right] \begin{bmatrix} g^1_1(\mathbf{z}) & \cdots & g^1_k(\mathbf{z}) \\ \vdots & \ddots & \vdots \\ g^k_1(\mathbf{z}) & \cdots & g^k_k(\mathbf{z}) \end{bmatrix} - \mathbf{1}^T,$$

where

$$g^i_j(\mathbf{z}) = \begin{cases} \dfrac{1 - z_i}{f_s(\mathbf{z})}, & i = j \\ -\dfrac{z_i}{f_s(\mathbf{z})}, & i \neq j \end{cases}.$$

In other words, we have derived formula (21).

APPENDIX 2   LOCAL VS. GLOBAL NONDOMINANCE

To show that for a locally nondominated vector $\mathbf{z}$, the set $\{\tilde{\mathbf{z}} \in f_a(\Delta_{k-1}^\circ) : \|\mathbf{z} - \tilde{\mathbf{z}}\| < r\}$, where r is a radius for local nondominance, can contain dominated vectors, we present the following sketch.

Let $\mathbf{ideal}^{Z_m} = \mathbf{0}$ for simplicity. To get started, we construct two sequences of vectors. Let $f_{s'}(\mathbf{z}) = (z_1 - 0.5)^4 + 1$, $f_{a'}(\mathbf{z}) = \mathbf{ideal}^{Z_m} + f_{s'}(\mathbf{z})\mathbf{z}$ and

$$\bar{\mathbf{z}}^1 = (0.25, 0.75)^\mathrm{T},$$

$$\hat{\mathbf{z}}^1 = \frac{f_{a'}\left(\bar{\mathbf{z}}^1\right) + (f_{a'}\left(\bar{\mathbf{z}}^1\right)_1, \bar{z}_2^1)^\mathrm{T}}{2},$$

$$\bar{\mathbf{z}}^n = (\hat{z}_1^{n-1}, 1 - \hat{z}_1^{n-1})^\mathrm{T}, n = 2, 3, \ldots,$$

$$\hat{\mathbf{z}}^n = \frac{f_{a'}\left(\bar{\mathbf{z}}^n\right) + (f_{a'}\left(\bar{\mathbf{z}}^n\right)_1, \bar{z}_2^n)^\mathrm{T}}{2}, n = 2, 3, \ldots.$$

Geometrically, every $\bar{\mathbf{z}}^n$ has been selected from a closed line $[(1,0)^\mathrm{T}, (0,1)^\mathrm{T}]$, and the vector $\hat{\mathbf{z}}^n$ is the mean vector of vectors $f_{a'}(\bar{\mathbf{z}}^n)$ and $(f_{a'}(\bar{\mathbf{z}}^n)_1, \bar{z}_2^n)^\mathrm{T}$. It is clear that the vector $\bar{\mathbf{z}}^n$ dominates $\hat{\mathbf{z}}^n$ for each $n = 1, 2, \ldots$. This can be seen in Figure 69.
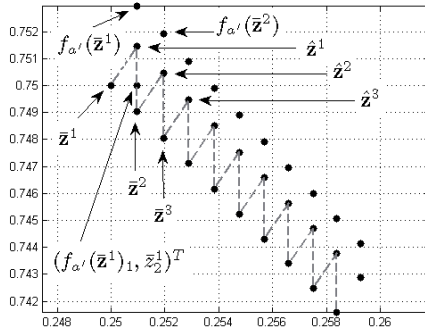

FIGURE 69   Sequences $(\bar{\mathbf{z}}^n)$ and $(\hat{\mathbf{z}}^n)$ in the objective space $\mathbb{R}^2$.
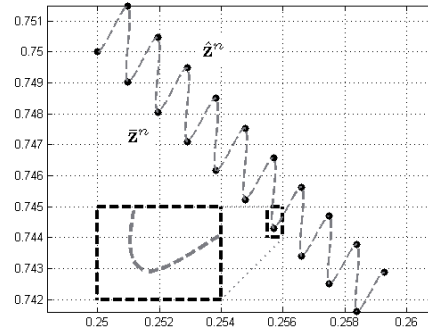
FIGURE 70   A sine or cosine function fitted to the sequences $(\bar{\mathbf{z}}^n)$ and $(\hat{\mathbf{z}}^n)$.

Limits

$$\lim_{n\to\infty} \hat{\mathbf{z}}^n = (0.5, 0.5)^\mathrm{T} = \lim_{n\to\infty} \bar{\mathbf{z}}^n \tag{40}$$

exist because the function $g : [0.25, 0.5] \to [0.25, 0.5]$, $g(t) = ((t - 0.5)^4 + 1)t$ is shrinking, that is, $|g(t_1) - g(t_2)| < |t_1 - t_2|$ for all $t_1, t_2 \in [0.25, 0.5]$, see [38, page 178]. Then the function g has a unique fixed point [172], that is, there is only one $t' \in [0.25, 0.5]$ so that $g(t') = t'$. In this case, $t' = 0.5$.

Now, the function f_s is constructed so that $\bar{\mathbf{z}}^n, \hat{\mathbf{z}}^n \in f_a(\Delta_1^\circ)$ for all $n = 1, 2, \ldots$ and the normal vector $((0.5, 0.5)^\mathrm{T})^\perp \in \mathbb{R}_+^2 \cup \mathbb{R}_-^2$. Let the function f_s be a sine or

a cosine function so that the lower peaks of the function f_a hit $\bar{\mathbf{z}}^n$ and the upper peaks $\hat{\mathbf{z}}^n$, when $z_1 < 0.5$. This can be done by defining separate scalings for each interval defined by $\bar{\mathbf{z}}^n$ and $f_{Z_m}^p(\hat{\mathbf{z}}^n)$ on Δ_1. The rough idea of this is presented in Figure 70, where f_a is depicted as a gray dashed curve. We set $f_s(\mathbf{z}) = 1$ when $z_1 \geq 0.5$.

It can be proved (it is not difficult, but very technical) that f_a is differentiable (but not continuously) and a normal vector at $(0.5, 0.5)^T$ is $(-1, -1)^T \in \mathbb{R}_+^2 \cup \mathbb{R}_-^2$. However, for every $n = 1, 2, \ldots$, we have a pair of vectors in

$$B\big((0.5, 0.5)^T, r_{(0.5,0.5)^T}/n\big) \cap f_a(\Delta_1^\circ)$$

so that one dominates the other.

APPENDIX 3 DERIVATION OF FUNCTIONS f_l AND f_u

The explicit forms of functions f_l and f_u in Definition 39 are rather simple, but one may wonder how they have been found. Geometrically, the function f_l has been constructed as follows.

Let $\mathbf{z} \in \Delta_{k-1}^\circ$ and $\mathbf{z}^i \in Z_m$. It is obvious that $f_a(\mathbf{z})$ dominates the vector $\mathbf{z}^i \in Z_m$ if for all $j = 1, \ldots, k$ we have

$$\text{ideal}_j^{Z_m} + f_s(\mathbf{z})z_j < z_j^i \quad \Leftrightarrow \quad f_s(\mathbf{z}) < \frac{z_j^i - \text{ideal}_j^{Z_m}}{z_j},$$

see Figure 71. Then by choosing

$$s_\mathbf{z}^i = \min_{j=1,\ldots,k} \frac{z_j^i - \text{ideal}_j^{Z_m}}{z_j},$$

we have $\text{ideal}_j^{Z_m} + s_\mathbf{z}^i z_j \leq z_j^i$ for all $j = 1, \ldots, k$. If we select $s_\mathbf{z}$ to be the maximum of scalar values $s_\mathbf{z}^i$, $i = 1, \ldots, m$, then there is at least one vector $\mathbf{z}' \in Z_m$ so that

$$\text{ideal}_j^{Z_m} + s_\mathbf{z} z_j \leq z_j' \quad \text{for all } j = 1, \ldots, k.$$

Now, by setting

$$f_l(\mathbf{z}) := s_\mathbf{z} = \max_{i=1,\ldots,m} \min_{j=1,\ldots,k} \frac{z_j^i - \text{ideal}_j^{Z_m}}{z_j},$$

we have derived the function f_l.
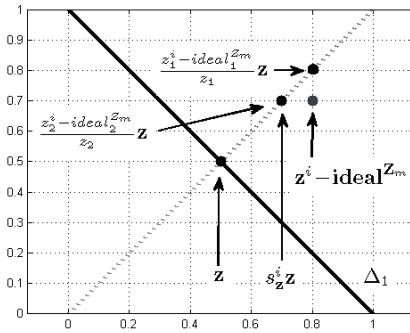
FIGURE 71   Derivation of function f_l.
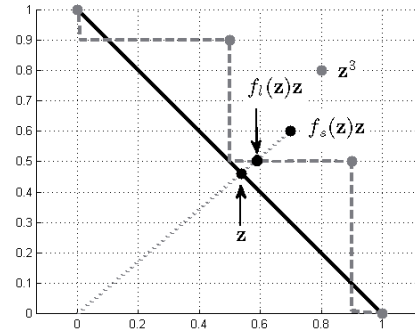
FIGURE 72   Taking the maximum in the formulation of function f_l is essential.

Derivation of the function f_u is similar to the derivation of the function f_l. However, there are several aspects to be considered. First of all, taking the maximum in the function f_l is essential. Otherwise, it would be possible that the vector $f_a(\mathbf{z})$ dominates some of the vectors from Z_m even though $f_l(\mathbf{z}) < f_s(\mathbf{z})$.

This would be against Theorem 40, claim (4). We are illustrating this in Figure 72, where $\mathbf{ideal}^{Z_m} = \mathbf{0}$ and $s_{\mathbf{z}}^3$ has not been considered in calculating scalar value $f_l(\mathbf{z})$. As it can be seen, we have $f_l(\mathbf{z}) < f_s(\mathbf{z})$, but the vector $f_a(\mathbf{z})$ dominates the vector $\mathbf{z}^3 \in Z_5$.

In a case, where the MOO problem consists of two objective functions only, the set $\{\mathbf{z}' \in \mathbb{R}^3 : \mathbf{z}' = \mathbf{ideal}^{Z_m} + f_u(\mathbf{z})\mathbf{z}, \mathbf{z} \in \Delta_2^\circ\}$ is always bounded as seen in Figure 48. However, in higher-dimensional objective spaces, this may not always happen. Let us demonstrate this as follows. Let $Z_3 = \{(0.5, 0.5, 0)^T =: \mathbf{z}^1, (0.5, 0, 0.5)^T =: \mathbf{z}^2, (0, 0.5, 0.5)^T =: \mathbf{z}^3\}$. In Figure 73, vectors $\mathbf{z}^1$, $\mathbf{z}^2$ and $\mathbf{z}^3$ are depicted as black dots on a gray simplex. The set $\{\mathbf{z}' \in \mathbb{R}^3 : \mathbf{z}' = \mathbf{ideal}^{Z_m} + f_u(\mathbf{z})\mathbf{z}, \mathbf{z} \in \Delta_2^\circ\}$ is illustrated by grayscaled planes. Now, each of the planes continues to the infinity and, therefore, the set is not bounded because none of the vectors in Z_3 dominates the vector $(\hat{z}, 0.25, 0.25)^T$, $\hat{z} > 0.5$. In other words, the number $\hat{z}$ can be arbitrary large and still the vector $(\hat{z}, 0.25, 0.25)^T$ is not dominated by $\mathbf{z}^1$, $\mathbf{z}^2$ and $\mathbf{z}^3$. Note that the vector $(\hat{z}, 0.25, 0.25)^T$ can be expressed by the function f_a.
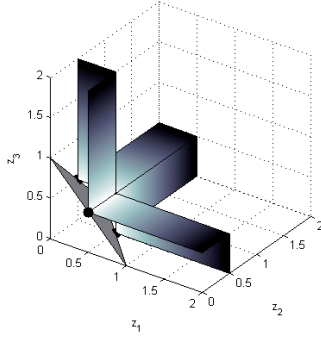
FIGURE 73   Unbounded upper bounds for function f_s.

The set $\{\mathbf{z}' \in \mathbb{R}^k : \mathbf{z}' = \mathbf{ideal}^{Z_m} + f_l(\mathbf{z})\mathbf{z}, \mathbf{z} \in \Delta_{k-1}^\circ\}$ is always bounded, no matter how large k is. The boundedness is guaranteed by Δ_{k-1}° and by the formulation of the function f_l. Note that the set can be a singleton. This can happen, for example if $Z_2 = \{(0, 0, 1)^T, (0, 1, 0)^T\}$.

JYVÄSKYLÄ STUDIES IN COMPUTING

1 ROPPONEN, JANNE, Software risk management - foundations, principles and empirical findings. 273 p. Yhteenveto 1 p. 1999.

2 KUZMIN, DMITRI, Numerical simulation of reactive bubbly flows. 110 p. Yhteenveto 1 p. 1999.

3 KARSTEN, HELENA, Weaving tapestry: collaborative information technology and organisational change. 266 p. Yhteenveto 3 p. 2000.

4 KOSKINEN, JUSSI, Automated transient hypertext support for software maintenance. 98 p. (250 p.) Yhteenveto 1 p. 2000.

5 RISTANIEMI, TAPANI, Synchronization and blind signal processing in CDMA systems. - Synkronointi ja sokea signaalinkäsittely CDMA järjestelmässä. 112 p. Yhteenveto 1 p. 2000.

6 LAITINEN, MIKA, Mathematical modelling of conductive-radiative heat transfer. 20 p. (108 p.) Yhteenveto 1 p. 2000.

7 KOSKINEN, MINNA, Process metamodelling. Conceptual foundations and application. 213 p. Yhteenveto 1 p. 2000.

8 SMOLIANSKI, ANTON, Numerical modeling of two-fluid interfacial flows. 109 p. Yhteenveto 1 p. 2001.

9 NAHAR, NAZMUN, Information technology supported technology transfer process. A multi-site case study of high-tech enterprises. 377 p. Yhteenveto 3 p. 2001.

10 FOMIN, VLADISLAV V., The process of standard making. The case of cellular mobile telephony. - Standardin kehittämisen prosessi. Tapaustutkimus solukkoverkkoon perustuvasta matkapuhelintekniikasta. 107 p. (208 p.) Yhteenveto 1 p. 2001.

11 PÄIVÄRINTA, TERO, A genre-based approach to developing electronic document management in the organization. 190 p. Yhteenveto 1 p. 2001.

12 HÄKKINEN, ERKKI, Design, implementation and evaluation of neural data analysis environment. 229 p. Yhteenveto 1 p. 2001.

13 HIRVONEN, KULLERVO, Towards better employment using adaptive control of labour costs of an enterprise. 118 p. Yhteenveto 4 p. 2001.

14 MAJAVA, KIRSI, Optimization-based techniques for image restoration. 27 p. (142 p.) Yhteenveto 1 p. 2001.

15 SAARINEN, KARI, Near infra-red measurement based control system for thermo-mechanical refiners. 84 p. (186 p.) Yhteenveto 1 p. 2001.

16 FORSELL, MARKO, Improving component reuse in software development. 169 p. Yhteenveto 1 p. 2002.

17 VIRTANEN, PAULI, Neuro-fuzzy expert systems in financial and control engineering. 245 p. Yhteenveto 1 p. 2002.

18 KOVALAINEN, MIKKO, Computer mediated organizational memory for process control. Moving CSCW research from an idea to a product. 57 p. (146 p.) Yhteenveto 4 p. 2002.

19 HÄMÄLÄINEN, TIMO, Broadband network quality of service and pricing. 140 p. Yhteenveto 1 p. 2002.

20 MARTIKAINEN, JANNE, Efficient solvers for discretized elliptic vector-valued problems. 25 p. (109 p.) Yhteenveto 1 p. 2002.

21 MURSU, ANJA, Information systems development in developing countries. Risk management and sustainability analysis in Nigerian software companies. 296 p. Yhteenveto 3 p. 2002.

22 SELEZNYOV, ALEXANDR, An anomaly intrusion detection system based on intelligent user recognition. 186 p. Yhteenveto 3 p. 2002.

23 LENSU, ANSSI, Computationally intelligent methods for qualitative data analysis. 57 p. (180 p.) Yhteenveto 1 p. 2002.

24 RYABOV, VLADIMIR, Handling imperfect temporal relations. 75 p. (145 p.) Yhteenveto 2 p. 2002.

25 TSYMBAL, ALEXEY, Dynamic integration of data mining methods in knowledge discovery systems. 69 p. (170 p.) Yhteenveto 2 p. 2002.

26 AKIMOV, VLADIMIR, Domain decomposition methods for the problems with boundary layers. 30 p. (84 p.). Yhteenveto 1 p. 2002.

27 SEYUKOVA-RIVKIND, LUDMILA, Mathematical and numerical analysis of boundary value problems for fluid flow. 30 p. (126 p.) Yhteenveto 1 p. 2002.

28 HÄMÄLÄINEN, SEPPO, WCDMA Radio network performance. 235 p. Yhteenveto 2 p. 2003.

29 PEKKOLA, SAMULI, Multiple media in group work. Emphasising individual users in distributed and real-time CSCW systems. 210 p. Yhteenveto 2 p. 2003.

30 MARKKULA, JOUNI, Geographic personal data, its privacy protection and prospects in a location-based service environment. 109 p. Yhteenveto 2 p. 2003.

31 HONKARANTA, ANNE, From genres to content analysis. Experiences from four case organizations. 90 p. (154 p.) Yhteenveto 1 p. 2003.

32 RAITAMÄKI, JOUNI, An approach to linguistic pattern recognition using fuzzy systems. 169 p. Yhteenveto 1 p. 2003.

33 SAALASTI, SAMI, Neural networks for heart rate time series analysis. 192 p. Yhteenveto 5 p. 2003.

34 NIEMELÄ, MARKETTA, Visual search in graphical interfaces: a user psychological approach. 61 p. (148 p.) Yhteenveto 1 p. 2003.

35 YOU, YU, Situation Awareness on the world wide web. 171 p. Yhteenveto 2 p. 2004.

36 TAATILA, VESA, The concept of organizational competence – A foundational analysis. - Perusteanalyysi organisaation kompetenssin käsitteestä. 111 p. Yhteenveto 2 p. 2004.

37 LYYTIKÄINEN, VIRPI, Contextual and structural metadata in enterprise document management. - Konteksti- ja rakennemetatieto organisaation dokumenttien hallinnassa. 73 p. (143 p.) Yhteenveto 1 p. 2004.

38 KAARIO, KIMMO, Resource allocation and load balancing mechanisms for providing quality of service in the Internet. 171 p. Yhteenveto 1 p. 2004.

39 ZHANG, ZHEYING, Model component reuse. Conceptual foundations and application in the metamodeling-based systems analysis and design environment. 76 p. (214 p.) Yhteenveto 1 p. 2004.

40 HAARALA, MARJO, Large-scale nonsmooth optimization variable metric bundle method with limited memory. 107 p. Yhteenveto 1 p. 2004.

41 KALVINE, VIKTOR, Scattering and point spectra for elliptic systems in domains with cylindrical ends. 82 p. 2004.

42 DEMENTIEVA, MARIA, Regularization in multistage cooperative games. 78 p. 2004.

43 MAARANEN, HEIKKI, On heuristic hybrid methods and structured point sets in global continuous optimization. 42 p. (168 p.) Yhteenveto 1 p. 2004.

44 FROLOV, MAXIM, Reliable control over approximation errors by functional type a posteriori estimates. 39 p. (112 p.) 2004.

45 ZHANG, JIAN, Qos- and revenue-aware resource allocation mechanisms in multiclass IP networks. 85 p. (224 p.) 2004.

46 KUJALA, JANNE, On computation in statistical models with a psychophysical application. 40 p. (104 p.) 2004.,

47 SOLBAKOV, VIATCHESLAV, Application of mathematical modeling for water environment problems. 66 p. (118 p.) 2004.

48 HIRVONEN, ARI P., Enterprise architecture planning in practice. The Perspectives of information and communication technology service provider and end-user. 44 p. (135 p.) Yhteenveto 2 p. 2005.

49 VARTIAINEN, TERO, Moral conflicts in a project course in information systems education. 320 p. Yhteenveto 1p. 2005.

50 HUOTARI, JOUNI, Integrating graphical information system models with visualization techniques. - Graafisten tietojärjestelmäkuvausten integrointi visualisointitekniikoilla. 56 p. (157 p.) Yhteenveto 1p. 2005.

51 WALLENIUS, EERO R., Control and management of multi-access wireless networks. 91 p. (192 p.) Yhteenveto 3 p. 2005.

52 LEPPÄNEN, MAURI, An ontological framework and a methodical skeleton for method engineering – A contextual approach. 702 p. Yhteenveto 2 p. 2005.

53 MATYUKEVICH, SERGEY, The nonstationary Maxwell system in domains with edges and conical points. 131 p. Yhteenveto 1 p. 2005.

54 SAYENKO, ALEXANDER, Adaptive scheduling for the QoS supported networks. 120 p. (217 p.) 2005.

55 KURJENNIEMI, JANNE, A study of TD-CDMA and WCDMA radio network enhancements. 144 p. (230 p.) Yhteenveto 1 p. 2005.

56 PECHENIZKIY, MYKOLA, Feature extraction for supervised learning in knowledge discovery systems. 86 p. (174 p.) Yhteenveto 2 p. 2005.

57 IKONEN, SAMULI, Efficient numerical methods for pricing American options. 43 p. (155 p.) Yhteenveto 1 p. 2005.

58 KÄRKKÄINEN, KARI, Shape sensitivity analysis for numerical solution of free boundary problems. 83 p. (119 p.) Yhteenveto 1 p. 2005.

59 HELFENSTEIN, SACHA, Transfer. Review, reconstruction, and resolution. 114 p. (206 p.) Yhteenveto 2 p. 2005.

60 NEVALA, KALEVI, Content-based design engineering thinking. In the search for approach. 64 p. (126 p.) Yhteenveto 1 p. 2005.

61 KATASONOV, ARTEM, Dependability aspects in the development and provision of location-based services. 157 p. Yhteenveto 1 p. 2006.

62 SARKKINEN, JARMO, Design as discourse: Representation, representational practice, and social practice. 86 p. (189 p.) Yhteenveto 1 p. 2006.

63 ÄYRÄMÖ, SAMI, Knowledge mining using robust clustering. 296 p. Yhteenveto 1 p. 2006.

64 IFINEDO, PRINCELY EMILI, Enterprise resource planning systems success assessment: An integrative framework. 133 p. (366 p.) Yhteenveto 3 p. 2006.

65 VIINIKAINEN, ARI, Quality of service and pricingin future multiple service class networks. 61 p. (196 p.) Yhteenveto 1 p. 2006.

66 WU, RUI, Methods for space-time parameter estimation in DS-CDMA arrays. 73 p. (121 p.) 2006.

67 PARKKOLA, HANNA, Designing ICT for mothers. User psychological approach. – Tieto- ja viestintätekniikoiden suunnittelu äideille. Käyttäjäpsykologinen näkökulma. 77 p. (173 p.) Yhteenveto 3 p. 2006.

68 HAKANEN, JUSSI, On potential of interactive multiobjective optimization in chemical process design. 75 p. (160 p.) Yhteenveto 2 p. 2006.

69 PUTTONEN, JANI, Mobility management in wireless networks. 112 p. (215 p.) Yhteenveto 1 p. 2006.

70 LUOSTARINEN, KARI, Resource , management methods for QoS supported networks. 60 p. (131 p.) 2006.

71 TURCHYN, PAVLO, Adaptive meshes in computer graphics and model-based simulation. 27 p. (79 p.) Yhteenveto 1 p.

72 ZHOVTOBRYUKH, DMYTRO, Context-aware web service composition. 290 p. Yhteenveto 2 p. 2006.

73 KOHVAKKO, NATALIYA, Context modeling and utilization in heterogeneous networks. 154 p. Yhteenveto 1 p. 2006.

74 MAZHELIS, OLEKSIY, Masquerader detection in mobile context based on behaviour and environment monitoring. 74 p. (179 p). Yhteenveto 1 p. 2007.

75 SILTANEN, JARMO, Quality of service and dynamic scheduling for traffic engineering in next generation networks. 88 p. (155 p.) 2007.

76 KUUVA, SARI, Content-based approach to experiencing visual art. - Sisältöperustainen lähestymistapa visuaalisen taiteen kokemiseen. 203 p. Yhteenveto 3 p. 2007.

77 RUOHONEN, TONI, Improving the operation of an emergency department by using a simulation model. 164 p. 2007.

78 NAUMENKO, ANTON, Semantics-based access control in business networks. 72 p. (215 p.) Yhteenveto 1 p. 2007.

79 WAHLSTEDT, ARI, Stakeholders' conceptions of learning in learning management systems development. - Osallistujien käsitykset oppimisesta oppimisympäristöjen kehittämisessä. 83 p. (130 p.) Yhteenveto 1 p. 2007.

80 ALANEN, OLLI, Quality of service for triple play services in heterogeneous networks. 88 p. (180 p.) Yhteenveto 1 p. 2007.

81 NERI, FERRANTE, Fitness diversity adaptation in memetic algorithms. 80 p. (185 p.) Yhteenveto 1 p. 2007.

82 KURHINEN, JANI, Information delivery in mobile peer-to-peer networks. 46 p. (106 p.) Yhteenveto 1 p. 2007.

83 KILPELÄINEN, TURO, Genre and ontology based business information architecture framework (GOBIAF). 74 p. (153 p.) Yhteenveto 1 p. 2007.

84 YEVSEYEVA, IRYNA, Solving classification problems with multicriteria decision aiding approaches. 182 p. Yhteenveto 1 p. 2007.

85 KANNISTO, ISTO, Optimized pricing, QoS and segmentation of managed ICT services. 45 p. (111 p.) Yhteenveto 1 p. 2007.

86 GORSHKOVA, ELENA, A posteriori error estimates and adaptive methods for incompressible viscous flow problems. 72 p. (129 p.) Yhteenveto 1 p. 2007.

87 LEGRAND, STEVE, Use of background real-world knowledge in ontologies for word sense disambiguation in the semantic web. 73 p. (144 p.) Yhteenveto 1 p. 2008.

88 HÄMÄLÄINEN, NIINA, Evaluation and measurement in enterprise and software architecture management. - Arviointi ja mittaaminen kokonais- ja ohjelmistoarkkitehtuurien hallinnassa. 91 p. (175 p.) Yhteenveto 1 p. 2008.

89 OJALA, ARTO, Internationalization of software firms: Finnish small and medium-sized software firms in Japan. 57 p. (180 p.) Yhteenveto 2 p. 2008.

90 LAITILA, ERKKI, Symbolic Analysis and Atomistic Model as a Basis for a Program Comprehension Methodology. 321 p. Yhteenveto 3 p. 2008.

91 NIHTILÄ, TIMO, Performance of Advanced Transmission and Reception Algorithms for High Speed Downlink Packet Access. 93 p. (186 p.) Yhteenveto 1 p. 2008.

92 SETÄMAA-KÄRKKÄINEN, ANNE, Network connection selection-solving a new multiobjective optimization problem. 52 p. (111p.) Yhteenveto 1 p. 2008.

93 PULKKINEN, MIRJA, Enterprise architecture as a collaboration tool. Discursive process for enterprise architecture management, planning and development. 130 p. (215 p.) Yhteenveto 2 p. 2008.

94 PAVLOVA, YULIA, Multistage coalition formation game of a self-enforcing international environmental agreement. 127 p. Yhteenveto 1 p. 2008.

95 NOUSIAINEN, TUULA, Children's involvement in the design of game-based learning environments. 297 p. Yhteenveto 2 p. 2008.

96 KUZNETSOV, NIKOLAY V., Stability and oscillations of dynamical systems. Theory and applications. 116 p. Yhteenveto 1 p. 2008.

97 KHRIYENKO, OLEKSIY, Adaptive semantic Web based environment for web resources. 193 p. Yhteenveto 1 p. 2008.

98 TIRRONEN, VILLE, Global optimization using memetic differential evolution with applications to low level machine vision. 98 p. (248 p.) Yhteenveto 1 p. 2008.

99 VALKONEN, TUOMO, Diff-convex combinations of Euclidean distances: A search for optima. 148 p. Yhteenveto 1 p. 2008.

100 SARAFANOV, OLEG, Asymptotic theory of resonant tunneling in quantum waveguides of variable cross-section. 69 p. Yhteenveto 1 p. 2008.

101 POZHARSKIY, ALEXEY, On the electron and phonon transport in locally periodical waveguides. 81 p. Yhteenveto 1 p. 2008.

102 AITTOKOSKI, TIMO, On challenges of simulation-based global and multiobjective optimization. 80 p. (204 p.) Yhteenveto 1 p. 2009.

103 YALAHO, ANICET, Managing offshore outsourcing of software development using the ICT-supported unified process model: A cross-case analysis. 91 p. (307 p.) Yhteenveto 4 p. 2009.

104 KOLLANUS, SAMI, Tarkastuskäytänteiden kehittäminen ohjelmistoja tuottavissa organisaatioissa. - Improvement of inspection practices in software organizations. 179 p. Summary 4 p. 2009.

105 LEIKAS, JAANA, Life-Based Design. 'Form of life' as a foundation for ICT design for older adults. - Elämälähtöinen suunnittelu. Elämänmuoto ikääntyville tarkoitettujen ICT tuotteiden ja palvelujen suunnittelun lähtökohtana. 218 p. (318 p.) Yhteenveto 4 p. 2009.

106 VASILYEVA, EKATERINA, Tailoring of feedback in web-based learning systems: Certitude-based assessment with online multiple choice questions. 124 p. (184 p.) Yhteenveto 2 p. 2009.

107 KUDRYASHOVA, ELENA V., Cycles in continuous and discrete dynamical systems. Computations, computer assisted proofs, and computer experiments. 79 p. (152 p.) Yhteenveto 1 p. 2009.

108 BLACKLEDGE, JONATHAN, Electromagnetic scattering and inverse scattering solutions for the analysis and processing of digital signals and images. 297 p. Yhteenveto 1 p. 2009.

109 IVANNIKOV, ANDRIY, Extraction of event-related potentials from electroencephalography data. - Herätepotentiaalien laskennallinen eristäminen EEG-havaintoaineistosta. 108 p. (150 p.) Yhteenveto 1 p. 2009.

110 KALYAKIN, IGOR, Extraction of mismatch negativity from electroencephalography data. - Poikkeavuusnegatiivisuuden erottaminen EEG-signaalista. 47 p. (156 p.) Yhteenveto 1 p. 2010.

111 HEIKKILÄ, MARIKKA, Coordination of complex operations over organisational boundaries. 265 p. Yhteenveto 3 p. 2010.

112 FEKETE, GÁBOR, Network interface management in mobile and multihomed nodes. 94 p. (175 p.) Yhteenveto 1 p. 2010.

113 KUJALA, TUOMO, Capacity, workload and mental contents - Exploring the foundations of driver distraction. 146 p. (253 p.) Yhteenveto 2 p. 2010.

114 LUGANO, GIUSEPPE, Digital community design - Exploring the role of mobile social software in the process of digital convergence. 253 p. (316 p.) Yhteenveto 4 p. 2010.

115 KAMPYLIS, PANAGIOTIS, Fostering creative thinking. The role of primary teachers. - Luovaa ajattelua kehittämässä. Alakoulun opettajien rooli. 136 p. (268 p.) Yhteenveto 2 p. 2010.

116 TOIVANEN, JUKKA, Shape optimization utilizing consistent sensitivities. - Muodon optimointi käyttäen konsistentteja herkkyyksiä. 55 p. (130p.) Yhteenveto 1 p. 2010.

117 MATTILA, KEIJO, Implementation techniques for the lattice Boltzmann method. - Virtausdynamiikan tietokonesimulaatioita Hila-Boltzmann -menetelmällä: implementointi ja reunaehdot. 177 p. (233 p.) Yhteenveto 1 p. 2010.

118 CONG, FENGYU, Evaluation and extraction of mismatch negativity through exploiting temporal, spectral, time-frequency, and spatial features. - Poikkeavuusnegatiivisuuden (MMN) erottaminen aivosähkönauhoituksista käyttäen ajallisia, spektraalisia, aikataajuus- ja tilapiirteitä. 57 p. (173 p.) Yhteenveto 1 p. 2010.

119 LIU, SHENGHUA, Interacting with intelligent agents. Key issues in agent-based decision support system design. 90 p. (143 p.) Yhteenveto 2 p. 2010.

120 AIRAKSINEN, TUOMAS, Numerical methods for acoustics and noise control. - Laskennallisia menetelmiä akustisiin ongelmiin ja melunvaimennukseen. 58 p. (133 p.) Yhteenveto 2 p. 2010.

121 WEBER, MATTHIEU, Parallel global optimization Structuring populations in differential evolution. - Rinnakkainen globaalioptimointi. Populaation rakenteen määrittäminen differentiaalievoluutiossa. 70 p. (185 p.) Yhteenveto 2 p. 2010.

122 VÄÄRÄMÄKI, TAPIO, Next generation networks, mobility management and appliances in intelligent transport systems. - Seuraavan sukupolven tietoverkot, liikkuvuuden hallinta ja sovellutukset älykkäässä liikenteessä. 50 p. (111 p.) Yhteenveto 1 p. 2010.

123 VIUKARI, LEENA, Tieto- ja viestintätekniikkavälitteisen palvelun kehittämisen kolme diskurssia. - Three discourses for an ICT-service development . 304 p. Summary 5 p. 2010.

124 PUURTINEN, TUOMAS, Numerical simulation of low temperature thermal conductance of corrugated nanofibers. - Poimutettujen nanokuitujen lämmönjohtavuuden numeerinen simulointi matalissa lämpötiloissa . 114 p. Yhteenveto 1 p. 2010.

125 HILTUNEN, LEENA, Enhancing web course design using action research . - Verkkoopetuksen suunnittelun kehittäminen toimintatutkimuksen keinoin . 192 p. Yhteenveto 2 p. 2010.

126 AHO, KARI, Enhancing system level performance of third generation cellular networks through VoIP and MBMS services. 121 p. (221 p.). Yhteenveto 2 p. 2010.

127 HÄKKINEN, MARKKU, Why alarms fail. A cognitive explanatory model. 102 p. (210 p.). Yhteenveto 1 p. 2010.

128 PENNANEN, ANSSI, A graph-based multigrid with applications. - Graafipohjainen monihilamenetelmä sovelluksineen. 52 p. (128 p.). Yhteenveto 2 p. 2010.

129 AHLGREN, RIIKKA, Software patterns, organizational learning and software process improvement. 70 p. (137 p.). Yhteenveto 1 p. 2011.

130 NIKITIN, SERGIY, Dynamic aspects of industrial middleware architectures 52 p. (114 p.). Yhteenveto 1 p. 2011.

131 SINDHYA, KARTHIK, Hybrid Evolutionary Multi-Objective Optimization with Enhanced Convergence and Diversity. 64 p. (160 p.). Yhteenveto 1 p. 2011.

132 MALI, OLLI, Analysis of errors caused by incomplete knowledge of material data in mathematical models of elastic media. 111 p. Yhteenveto 2 p. 2011.

133 MÖNKÖLÄ, SANNA, Numerical Simulation of Fluid-Structure Interaction Between Acoustic and Elastic Waves. 136 p. Yhteenveto 2 p. 2011.

134 PURANEN, TUUKKA, Metaheuristics Meet Meta-models. A Modeling Language and a Product Line Architecture for Route Optimization Systems. 270 p. Yhteenveto 1 p. 2011.

135 MÄKELÄ, JUKKA, Mobility Management in Heterogeneous IP-networks. 86 p. (145 p.) Yhteenveto 1 p. 2011.

136 SAVOLAINEN, PAULA, Why do software development projects fail? Emphasising the supplier's perspective and the project start-up. 81 p. (167 p.) Yhteenveto 2 p. 2011.

137 KUZNETSOVA, OLGA, Lyapunov quantities and limit cycles in two-dimensional dynamical systems: analytical methods, symbolic computation and visualization. 80 p. (121 p.) Yhteenveto 1 p. 2011.

138 KOZLOV, DENIS, The quality of open source software and its relation to the maintenance process. 125 p. (202 p.) Yhteenveto 1 p. 2011.

139 IACCA, GIOVANNI, Memory-saving optimization algorithms for systems with limited hardware. 100 p. (236 p.) Yhteenveto 1 p. 2011.

140 ISOMÖTTÖNEN, VILLE, Theorizing a one-semester real customer student software project course. 189 p. Yhteenveto 1 p. 2011.

141 HARTIKAINEN, MARKUS, Approximation through interpolation in nonconvex multiobjective optimization. 74 p. (164 p.) Yhteenveto 1 p. 2011.

142 MININNO, ERNESTO, Advanced optimization algorithms for applications in control engineering. 72 p. (149 p.) Yhteenveto 1 p. 2011.

143 TYKHOMYROV, VITALIY, Mitigating the amount of overhead arising from the control signaling of the IEEE 802.16 OFDMa System. 52 p. (138 p.) Yhteenveto 1 p. 2011.

144 MAKSIMAINEN, JOHANNA, Aspects of values in human-technology interaction design—a content-based view to values. - Ihmisen ja teknologian vuorovaikutussuunnittelun arvoulottuvuudet—sisältöperustainen lähestymistapa arvoihin. 111 p. (197 p.) Yhteenveto 2 p. 2011.

145 JUUTINEN, SANNA, Emotional obstacles of e-learning. 97 p. (181 p.) Yhteenveto 3 p. 2011.

146 TUOVINEN, TERO, Analysis of stability of axially moving orthotropic membranes and plates with a linear non-homogeneous tension profile. 104 p. Yhteenveto 1 p. 2011.

147 HILGARTH, BERND, The systemic cognition of e-Learning success in internationally operating organizations. - Kokonaisvaltainen käsitys e-oppimisen menestyksestä kansainvälisissä organisaatioisssa. 100 p. (181 p.) Yhteenveto 1 p. 2011.

148 JERONEN, JUHA, On the mechanical stability and out-of-plane dynamics of a travelling panel sub-merged in axially flowing ideal fluid. A study into paper production in mathematical terms. - Ideaali virtaukseen upotetun, aksiaalisesti liikkuvan paneelin mekaanisesta stabiilisuudesta ja dynamiikasta. Tutkimus paperintuotannosta matemaattisin käsittein. 243 p. Yhteenveto 3 p. 2011.

149 FINNE, AUVO, Tanzanit - Towards a comprehensive quality meta-model for information systems: Case studies of information system quality modelling in East Africa. 209 p. Yhteenveto 2 p. 2011.

150 KANKAANPÄÄ, IRJA, IT Artefact Renewal: Triggers, Timing and Benefits. 79 p. (164 p.) Yhteenveto 1 p. 2011.

151 KOTILAINEN, NIKO, Methods and Applications for Peer-to-Peer Networking. 46 p. (133 p.) Yhteenveto 1 p. 2011.

152 SKRYPNYK, IRYNA, Unstable feature relevance in classification tasks. - Epävakaiden ominaisuuksien merkitys luokittelutehtävissä. 232 p. Yhteenveto 1 p. 2011.

153 ZAIDENBERG, NEZER JACOB, Applications of virtualization in systems design. 297 p. Yhteenveto 1 p. 2012.

154 MARTIKAINEN, HENRIK, PHY and MAC Layer Performance Optimization of the IEEE 802.16 System. 80 p. (150 p.) Yhteenveto 1 p. 2012.

155 LESKINEN, JYRI, Distributed multi-objective optimization methods for shape design using evolutionary algorithms and game strategies. 86 p. (151 p.) Yhteenveto 1 p. 2012.

156 KUUSIO, ARI, Tietokannan hallintajärjestelmäportfolion hallinta ja sen arviointi suurissa asiakasorganisaatioissa. - DBMS portfolio management and assessment in large customer organizations. 354 p. Summary 3 p. 2012.

157 HAANPÄÄ, TOMI, Approximation method for computationally expensive nonconvex multiobjective optimization problems. 180 p. Yhteenveto 2 p. 2012.