



# Robust multiobjective optimization & applications in portfolio optimization



Jörg Fliege<sup>a,\*</sup>, Ralf Werner<sup>b</sup>

<sup>a</sup> CORMSIS, University of Southampton, UK

<sup>b</sup> Universität Augsburg, Institut für Mathematik, Germany

## ARTICLE INFO

### Article history:

Available online 26 October 2013

### Keywords:

Robustification

Uncertainty

Markowitz

Multiobjective

## ABSTRACT

Motivated by Markowitz portfolio optimization problems under uncertainty in the problem data, we consider general convex parametric multiobjective optimization problems under data uncertainty. For the first time, this uncertainty is treated by a *robust multiobjective formulation* in the gist of Ben-Tal and Nemirovski. For this novel formulation, we investigate its relationship to the original multiobjective formulation as well as to its scalarizations. Further, we provide a characterization of the location of the *robust Pareto frontier* with respect to the corresponding original Pareto frontier and show that standard techniques from multiobjective optimization can be employed to characterize this robust efficient frontier. We illustrate our results based on a standard mean–variance problem.

Crown Copyright © 2013 Published by Elsevier B.V. All rights reserved.

## 1. Motivation and overview

Prompted by the well-known strong data dependency of mean–variance optimization, we investigate how to treat multiobjective optimization problems with uncertain parameters. The framework of multiobjective optimization, in which several conflicting objectives have to be minimized simultaneously, can be seen as the natural setting for portfolio optimization problems, as such problems invariably have to deal with the conflicting notions of revenue and risk. However, and equally naturally, these problems also have to deal with uncertainty: problem data (like expected future return and covariances of random variables) are not known precisely, and only estimates are available. We are thus facing multiobjective optimization problems with uncertain parameters.

In this paper, we want to follow the idea of the *robust counterpart approach* where an entire set of possible parameter realizations – called *uncertainty set* – is used for the optimization, but no assumptions about the distribution of the unknown parameters is needed, in contrast to many other robustification approaches. In the context of portfolio optimization, several authors have considered instances varying from theoretical to practical settings, amongst others Goldfarb and Iyengar (2003), Tütüncü and Koenig (2004), Ceria and Stubbs (2006), Meucci (2005), Lutgens (2004) or Schöttle and Werner (2006). Most of these approaches have in common that the uncertainty sets are chosen based on statistical reasoning, but other approaches are followed as well. All these approaches are based on just one

specific choice of the risk-aversion parameter, i.e. only robustification of a particular instance of risk–return trade-off is considered.

In contrast to these approaches, we propose a different and completely novel approach. Instead of following Markowitz (1952) and referring to some scalar portfolio optimization problem, we start with the multiobjective formulation of the mean–variance portfolio problem in the gist of Kuhn and Tucker (1951). Our main aim is then to robustify the complete efficient frontier to obtain a *robust efficient frontier*. In this context, our main contributions are as follows:

- We introduce for the first time a robust counterpart to a multiobjective programming problem in the style of Ben-Tal and Nemirovski (1998, 1999).
- We investigate the relationship between the robust Pareto frontier and the original Pareto frontier and show that the robust frontier lies between the original nominal efficient frontier and some corresponding easy-to-determine upper bound.
- We demonstrate that robust efficient frontiers can be found by standard methods of robust and multiobjective programming under commonly made assumptions on the uncertainty.
- We pay particular attention to the case of portfolio optimization and show that the resulting robust multiobjective counterpart of the mean–variance portfolio optimization problem can be treated in a numerically efficient manner.

We want to emphasize that, although motivated by portfolio optimization, our methodology is general enough to be applied to any convex parametric multiobjective optimization problem under data uncertainty. As such, it might be especially useful for optimization problems arising in engineering.

\* Corresponding author. Tel.: +44 23 8059 8453.

E-mail addresses: [J.Fliege@soton.ac.uk](mailto:J.Fliege@soton.ac.uk) (J. Fliege), [ralf.werner@math.uni-augsburg.de](mailto:ralf.werner@math.uni-augsburg.de) (R. Werner).

The rest of the paper is organized as follows: Section 2 gives an overview of the necessary machinery from multiobjective optimization (Section 2.2), robust optimization (Section 2.3) and (robust) portfolio optimization (Section 2.4) that we need further on. Readers already familiar with these topics can easily skip this part of the presentation. Section 3 then contains the main results of this paper: we provide a proper notion of a robustified multiobjective problem and show that it leads, via scalarization, to families of robust single-objective problems that can also be derived in a different way. We also provide an alternative motivation for the robust counterpart and characterize the location of the robust efficient frontier with respect to the original frontier. Section 4 considers an illustrative numerical example from portfolio optimization that provides further insight into the proposed robust counterpart. Finally, we conclude in Section 5.

## 2. Introduction to multiobjective and robust optimization

### 2.1. Notation

The following notation is used throughout. We use  $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$  and  $\mathbb{R}_{++} = \{x \in \mathbb{R} : x > 0\}$  for the set of non-negative resp. strictly positive real numbers.  $\mathcal{K}(M)$  denotes the space of all non-empty, convex and compact subsets of a given non-empty and convex set  $M \subset \mathbb{R}^n$ . Further,  $B_\varepsilon(x)$  denotes the closed  $\varepsilon$ -ball around some vector  $x \in \mathbb{R}^n$ ,  $\|x\|$  denotes the 2-norm for  $x \in \mathbb{R}^n$  and  $\|\Sigma\|$  denotes the Frobenius norm for  $\Sigma \in \mathbb{S}^n$ , where  $\mathbb{S}^n$  is the space of all symmetric  $n \times n$  matrices and  $\mathbb{S}_+^n$  is the cone of positive semidefinite matrices.

### 2.2. Multiobjective optimization

In this section we provide a brief introduction of the concepts in multiobjective optimization that are used in this paper. We follow closely the exposition given by Fliege and Vicente (2006).

In multiobjective optimization, several functions

$$f_1, \dots, f_p : \mathbb{R}^n \rightarrow \mathbb{R}$$

with  $p > 1$  have to be minimized simultaneously over a set of feasible points characterized by a (convex) compact set  $X \in \mathcal{K}(\mathbb{R}^n)$ . The general problem can be conveniently stated in the form

$$\begin{aligned} \text{efmin} \quad & f : \mathbb{R}^n \rightarrow \mathbb{R}^p \\ \text{s.t.} \quad & x \in X, \end{aligned} \quad (M)$$

where  $f = (f_1, \dots, f_p)^\top$ , and the exact meaning of “efmin” still has to be specified. We will do so in what follows.

**Remark 2.1.** Multicriteria optimization is the ideal setting to analyse portfolio optimization problems in the sense of Markowitz. If we work in a financial market with  $n$  risky assets and  $x \in \mathbb{R}^n$  is a portfolio vector ( $\sum_{i=1}^n x_i = 1$ ), we can simply set  $p = 2$ , let, say,  $f_1(x) = s(x) = x^\top \Sigma x$  be the risk function for some covariance matrix  $\Sigma \in \mathbb{S}_+^n$  and let  $f_2(x) = -m(x) = -\mu^\top x$  be the return function for some vector of expected returns  $\mu \in \mathbb{R}^n$ .

The reason for the formulation given here is that there is no standard total order for the image space  $\mathbb{R}^p$ . In contrast to this situation, in the classical single-objective case one always uses the standard total order defined by  $x < y : \Leftrightarrow y - x \in \mathbb{R}_{++}$  ( $x, y \in \mathbb{R}$ ). Nevertheless, the idea of specifying an order by using a specific set defining it can be conveniently employed in multiobjective optimization, as the following discussion will show.

If an arbitrary order relation  $\prec$  on  $\mathbb{R}^p$  and a set  $A \subseteq \mathbb{R}^p$  are given, the vector  $a \in \mathbb{R}^p$  is called *minimal* or a *minimizer* w.r.t.  $\prec$  in  $A$  if

$a \in A$  and  $a \preceq b$  for all  $b \in A$ . (Here,  $\preceq$  is the reflexive hull of  $\prec$ , i.e.,  $a \preceq b$  if and only if  $a = b$  or  $a \prec b$ .) Minimal points usually do not exist, one reason being that it is seldom the case that  $\preceq$  is a total order. A weaker concept, the concept of *domination* is therefore needed. A point  $a$  *dominates* a point  $b$ , if  $a \prec b$  and  $a \neq b$  holds. A point  $a$  is *nondominated* in  $A$ , if  $a \in A$  and there does not exist a point  $c \in A$  with  $c \prec a$  and  $c \neq a$ . This approach raises the question about which of the many orders in  $\mathbb{R}^p$  one should choose when solving multicriteria problems.

Let  $K \subseteq \mathbb{R}^p$  be an arbitrary set. Define  $A := f(X)$  and the order

$$x <_K y : \Leftrightarrow y - x \in K. \quad (2.1)$$

For such an order relation, define further

$$\text{eff}_K(A) := \{a \in A \mid a \text{ nondominated in } A \text{ w.r.t. } <_K\},$$

the set of all *nondominated* or *efficient* points of the set  $A$ . It is this notion of efficiency or optimality that we will use when we speak of solutions of (M), and the operator  $\text{efmin}$  in (M) is understood to search for such efficient points: the set of solutions of the problem (M) is the preimage of all nondominated points of the set  $A = f(X)$  with respect to the order  $<_K$ .

The next theorem is well known, see, e.g. Göpfert and Nehse (1990) or Fliege and Vicente (2006).

**Theorem 2.2.** Let  $K \subseteq \mathbb{R}^p$  be a set and let  $<_K$  be the binary relation defined by  $K$  as in (2.1). Then, the following statements hold:

1. If  $0 \in K$  then  $<_K$  is reflexive.
2. If  $K + K \subseteq K$  then  $<_K$  is transitive.
3. If  $K$  is a cone containing no lines, i.e.,  $K \cap -K \subseteq \{0\}$ , then  $<_K$  is anti-symmetric. (In this case, the set  $K$  is also called pointed.)
4. The order  $<_K$  is total if and only if  $K \cup -K = \mathbb{R}^p$ .
5. The set  $K$  is closed if and only if the relation  $<_K$  is “continuous at 0” in the following sense. For all  $a \in \mathbb{R}^p$  and all sequences  $(a^{(i)})_{i \in \mathbb{N}}$  in  $\mathbb{R}^p$  with  $\lim_{i \rightarrow +\infty} a^{(i)} = a$  and  $0 <_K a^{(i)}$  for all  $i \in \mathbb{N}$  it follows that  $0 <_K a$  holds.

Note that  $K + K \subseteq K$  holds if  $K$  is a convex cone. According to the theorem above, practitioners prefer to choose a closed convex cone  $K$  with  $0 \in K$  which contains no lines to define the partial order  $<_K$ . (Note that the lexicographic order in  $\mathbb{R}^p$  is defined by a cone which is not closed.) Moreover, in our context the space  $\mathbb{R}^p$  will be the image space of functions to be minimized. As a consequence, it is important for numerical reasons to have scale-invariance of the induced order. This means that if  $x <_K y$  and  $\lambda > 0$  then  $\lambda x <_K \lambda y$ , a property which holds if and only if the set  $K$  is a cone.

Using a fixed set  $K$  to define an order relation as in (2.1) has one additional advantage. For an arbitrary relation  $\prec$ , the sets

$$C(a) := \{b \in \mathbb{R}^p \mid a \prec b\} - a \quad (2.2)$$

are constant if there exists a set  $K$  such that  $\prec = <_K$  holds. Indeed, if  $\prec = <_K$  then  $C(a) = \{b \in \mathbb{R}^p \mid a \prec b\} = \{b \in \mathbb{R}^p \mid a <_K b\} = \{b \in \mathbb{R}^p \mid b - a \in K\} = a + K$ . This means that  $\prec$  is translation-invariant, i.e.,  $x + z \prec y + z$  for all  $z$  if and only if  $x \prec y$ .

To summarize the discussion above: we need to choose a convex cone  $K \subseteq \mathbb{R}^p$  with  $0 \in K$  in order to define an order  $<_K$ . Other attributes of  $K$  that can be used to our advantage are closedness, pointedness, and  $K \cup -K = \mathbb{R}^p$  but, as pointed out before, we cannot have all of these at the same time.

Often,  $K$  is chosen to be the positive orthant,  $K = \mathbb{R}_+^p$ , which gives exactly the standard definition of order in multicriteria optimization and especially in the specific setting of mean–variance optimization.

We will now consider linear forms from  $\text{int}(K^*)$ , where  $K^*$  is the dual cone of  $K$  defined by  $K^* = \{\lambda \mid \forall a \in K : \lambda^\top a \geq 0\}$ . It turns out that in the case of convex cones and sets, these linear forms can

be used to characterize almost all efficient points of a multiobjective optimization problem, as we will see below. More precisely, we need the linear forms from the *quasi-interior* of  $K^*$ , i.e. from the set

$$K^+ := \{\lambda \in \mathbb{R}^p \mid \forall a \in K \setminus \{0\} : \lambda^\top a > 0\}.$$

With this, it turns out that basically “all” efficient points can be found by minimizing functions of the form  $y \mapsto \lambda^\top y$  over  $A$ , as the following theorem shows.

**Theorem 2.3.** *Let  $K, A \subseteq \mathbb{R}^p$  and define*

$$P_K(A) := \bigcup_{\lambda \in K^+} \arg \min \{\lambda^\top a \mid a \in A\}.$$

*Then, the following statements hold:*

1. *Let  $K$  be a convex cone with  $0 \in K$  and  $\{0\} \neq K \neq \mathbb{R}^p$ . Then*

$$P_K(A) \subseteq \text{eff}_K(A).$$

2. *Let  $K$  be a closed convex cone with  $0 \in K$  such that  $K$  contains no lines. Let  $A$  be closed and convex. Then*

$$\text{eff}_K(A) \subseteq \text{cl}(P_K(A)).$$

3. *Let  $K = \mathbb{R}_+^p$  and  $A$  polyhedral. Then*

$$\text{eff}_K(A) = P_K(A).$$

Different proofs can be found in various textbooks and original articles, see e.g. Göpfert and Nehse (1990, Section 2.22), Sawaragi, Nakayama, and Tanino (1985, p. 74), or Durier (1988). The first proof of part 2 is due to Arrow, Barankin, and Blackwell (1953). The set  $P_K(A)$  is often called the set of *properly efficient points* and also allows for a different characterization, see Geoffrion (1968). Also, note that other, nonlinear  $K$ -monotone functions have only recently attracted some attention, mainly for numerical reasons. See Fliege (2004) for more details.

We especially note that, under the assumptions listed in part 1 and 2 of the theorem above, we have that

$$P_K(A) \subseteq \text{eff}_K(A) \subseteq \text{cl}(P_K(A)) \quad (2.3)$$

holds. In other words, the sets  $\text{eff}_K(A)$  and  $P_K(A)$  differ at most by some boundary points from  $\text{cl}(P_K(A)) \setminus P_K(A)$  and are therefore, in a numerical setting with limited precision floating point arithmetic, indistinguishable from each other.

Our motivation was to use  $A = f(X)$ . So, in the convex case, by varying  $\lambda$  over “all” points  $\lambda \in K^+$  we can generate “all” nondominant points in  $f(X)$  w.r.t.  $<_K$  and their corresponding preimages by solving the problems

$$\begin{aligned} \min \quad & \lambda^\top f(x) \\ \text{s.t.} \quad & x \in X. \end{aligned} \quad (P(\lambda))$$

Some boundary points (more precisely, a subset of measure zero) are left out in the convex case, but, as mentioned above, this is a detail rarely of importance in applications. Such a methodology, in which a multiobjective problem is replaced by several standard single-objective problems is called *scalarization*, and the scalarization technique described above is the classical *weighted-sum scalarization* henceforth denoted by WSS. Alternative numerical methods for treating multiobjective problems have only recently gained some attention, see Fliege and Svaiter (2000) and Fliege, Grana Drummond, and Svaiter (2009).

We thus call the multiobjective problem  $(M)$  *equivalent* to the family of parameterized single-objective problems  $P(\lambda)$  if and only if (2.3) holds. Theorem 2.3 gives sufficient conditions for this equivalence.

Applying these results to the case  $K = \mathbb{R}_+^p$ , we readily arrive at  $K^+ = \text{int}(\mathbb{R}_+^p) = (\mathbb{R}_{++})^p$  as the set of linear forms to consider, and as the scalar problems are equivalent under scaling, we see that only problems of the form

$$\begin{aligned} \min \quad & \sum_{i=1}^p \lambda_i f_i(x) \\ \text{s.t.} \quad & x \in X. \end{aligned} \quad (2.4)$$

need to be considered for  $\lambda \in \mathbb{R}^p$  with  $\lambda_i > 0$  ( $i = 1, \dots, p$ ) and  $\sum_{i=1}^p \lambda_i = 1$ .

**Remark 2.4.** Considering the case of portfolio optimization in the Markowitz formulation, we see that we have, as mentioned before in Remark 2.1,  $p = 2$ , and it is also straightforward to see that all assumptions of item 1 and item 2 of Theorem 2.3 hold. As such, we arrive at (2.3), and we see that we can generate almost all efficient portfolios by solving all problems of the form

$$\begin{aligned} \min \quad & \lambda_1 s(x) - \lambda_2 m(x) = \lambda_1 x^\top \Sigma x - \lambda_2 \mu^\top x \\ \text{s.t.} \quad & x \in X. \end{aligned} \quad (2.5)$$

for all  $\lambda_1, \lambda_2 > 0$  with  $\lambda_1 + \lambda_2 = 1$ . It is now also clear which boundary points we are missing from the set of efficient points when employing this strategy: the minimum-variance portfolio as well as the maximum return portfolios. But these can readily be generated by considering (2.5) also for the cases  $\lambda_1 = 1, \lambda_2 = 0$  and  $\lambda_1 = 0, \lambda_2 = 1$ .

While solving problems of the form  $P(\lambda)$  appears to be an attractive method, especially as it is easy to interpret the meaning of the “weight vectors”  $\lambda$ , the problems  $P(\lambda)$  are often extremely sensitive to changes in  $\lambda$ . Thus, other methods have also been developed, among them the so-called  $\varepsilon$ -constraint scalarization, denoted by ECS in what follows. For this, consider an arbitrary  $k \in \{1, \dots, p\}$ , a parameter vector  $\varepsilon \in \mathbb{R}^p$ , and define the scalar optimization problem

$$\begin{aligned} \min \quad & f_k(x) \\ \text{s.t.} \quad & f_i(x) \leq \varepsilon_i \quad \forall i \in \{1, \dots, p\} \setminus \{k\}, \\ & x \in X. \end{aligned} \quad (P(k, \varepsilon))$$

(Note that, formally, the problem  $P(k, \varepsilon)$  does not depend on the parameter  $\varepsilon_k$ .) The interconnections between solutions of  $P(k, \varepsilon)$  and the given multiobjective problem are described in the following theorem.

**Theorem 2.5.** *Let  $k \in \{1, \dots, p\}$  be fixed and define*

$$S(k) := \bigcup_{\varepsilon \in \mathbb{R}^p} \{f(x) \in X \mid x \text{ solves } P(k, \varepsilon)\}$$

*as well as  $A := f(X)$ . Let  $K = \mathbb{R}_+^p$ . Then, the following holds.*

1. *We have*

$$\text{eff}_K(A) \subseteq S(k) \subseteq \text{eff}_{\text{int}(K)}(A)$$

2. *Let  $p = 2$ . Then,*

$$\text{eff}_K(A) = S(k). \quad (2.6)$$

A proof can be found, e.g., in Eichfelder (2008, Corollaries 2.28 and 2.29).

In a similar vein to the above, we can thus call the multiobjective problem  $(M)$  *equivalent* to the family of parameterized single-objective problems  $P(k, \varepsilon)$ , if and only if (2.6) holds. This is clearly the case for the considered problem of Markowitz portfolio optimization.

### 2.3. Robust optimization

Over the last decade the idea to deal with uncertainty in optimization by robust optimization has gained more and more interest among both academics and practitioners. An extensive and very recent introduction to the topic can be found in the book by Ben-Tal, El Ghaoui, and Nemirovski (2009) or in the review by Bertsimas, Brown, and Caramanis (2011), which contain large lists with references demonstrating the popularity and broadness of this approach. The robust optimization approach, especially the concept of the *robust counterpart*, was introduced by Ben-Tal, El Ghaoui and Nemirovski.<sup>1</sup> The robust counterpart takes into account uncertainty by replacing uncertain functions, i.e. functions depending on some uncertain parameter, by a conservative over-estimator, the worst-case value of the function on the whole uncertainty set. Thus, if applied to uncertain constraints, it guarantees feasibility under all (reasonable) realizations of the parameters, and if applied to the objective function, it guarantees a worst-case performance in all circumstances.

To proceed, let us consider the following convex parametric optimization problem

$$\begin{aligned} \min_{x \in X} \quad & f(x, u) \\ \text{s.t.} \quad & g(x, u) \leq 0, \end{aligned} \quad (P_u)$$

For the purpose of the subsequent results, a few mild assumptions on  $(P_u)$  are necessary which are in essence identical to the assumptions made in the seminal paper (Ben-Tal & Nemirovski, 1998) by Ben-Tal and Nemirovski:

- Let us assume throughout that  $X \in \mathcal{K}(\mathbb{R}^n)$  and let  $u \in \bar{U}$  with  $\bar{U} \in \mathcal{K}(\mathbb{R}^d)$ . Here,  $u$  denotes the uncertain data within a *global uncertainty set*.
- For convenience, we assume that both the objective  $f: X \times \bar{U} \rightarrow \mathbb{R}$  and all constraints  $g: X \times \bar{U} \rightarrow \mathbb{R}^m$  are continuously differentiable on  $X \times \bar{U}$ .
- To obtain a convex problem  $(P_u)$  let us assume that  $f$  and all constraints  $g = (g_1, \dots, g_m)^\top$  are convex in  $x$  for all  $u \in \bar{U}$ .

In the above setting, it is assumed that  $X$  subsumes all *certain* constraints, whereas all *uncertain* constraints explicitly depending on  $u$  are handled by the inequality  $g(x, u) \leq 0$ . Of course, the whole setup can be generalized to conic problems by replacing the constraint  $g(x, u) \leq 0$  with a convex conic constraint  $g(x, u) \in -K$ .

**Remark 2.6.** In general, it is reasonable to assume that uncertainty enters in a convex way, as it is expected that the worst uncertain parameter for a given point  $x$  lies on the boundary of  $U$ , which is the case for convex uncertainty. Taking into account a different notation for the constraints, convex uncertainty was assumed in Ben-Tal and Nemirovski (1998). Actually, in most practical applications, and especially in the situation of portfolio optimization described below, uncertainty even enters in a linear way. Nevertheless, we do not impose any assumption on the structure of uncertainty in the following.

Analogously to the definition of  $(P^*)$  in Ben-Tal and Nemirovski (1998), let us define the *robust counterpart* of a family of optimization problems  $(P_u)_{u \in U}$  for some uncertainty set  $U \in \mathcal{K}(\bar{U})$  as

$$\begin{aligned} \min_{x \in X} \quad & \max_{u \in U} f(x, u) \\ \text{s.t.} \quad & g(x, u) \leq 0 \quad \forall u \in U. \end{aligned} \quad (RP_u)$$

To avoid pathological situations in the investigations later on, we require that there exists one point  $\bar{x} \in X$  which is feasible for all  $u \in \bar{U}$ :  $g(\bar{x}, u) \leq 0$  for all  $u \in \bar{U}$ .

The robust counterpart approach as such represents a worst-case approach by considering the worst-case value of the objective over the whole set  $U$  of possible parameter realizations. Further, the solution has to be feasible for all realizations of the uncertain parameter within the uncertainty set  $U$ . Thus, the feasibility of the solution is immunized against perturbations of the parameter – which is the obvious benefit of robustification.

**Definition 2.7.** Let us introduce the *robust counterparts*  $f_U^{RC}$  of the function  $f$  (also called *robustified version* of  $f$ ) as

$$f_U^{RC}(x) := \max_{u \in U} f(x, u)$$

and similarly for  $g$  in a component-wise fashion

$$g_U^{RC}(x) := \begin{pmatrix} \max_{u \in U} g_1(x, u) \\ \dots \\ \max_{u \in U} g_m(x, u) \end{pmatrix} = \begin{pmatrix} g_{1,U}^{RC}(x) \\ \dots \\ g_{m,U}^{RC}(x) \end{pmatrix}$$

Then we immediately obtain for all  $x \in X$

$$f(x, u) \leq f_U^{RC}(x) \quad \forall u \in U$$

and

$$g(x, u) \leq g_U^{RC}(x) \quad \forall u \in U$$

where the second inequality is understood in the usual component-wise manner, i.e. using  $\leq_{\mathbb{R}^m}$ . This means that the robustified versions are convex overestimators of the original functions and thus lead to more conservative, i.e. robust, solutions. Based on the robust counterparts of  $f$  and  $g$ , the robust counterpart of  $(P_u)$  becomes

$$\begin{aligned} \min_{x \in X} \quad & f_U^{RC}(x) \\ \text{s.t.} \quad & g_U^{RC}(x) \leq 0. \end{aligned}$$

**Remark 2.8.** Under the given assumptions, the robust counterpart remains a convex problem, which is especially important for its numerical solution. However, due to the max operator, the once smooth problem now might become non-smooth, but still sub-differentiable. Depending on the specific choice of the uncertainty set  $U$  and the structure how uncertainty enters the functions  $f$  and  $g$ , analytical expressions for the robustified versions are available for a large class of functions. This means that under suitable assumptions, the robust counterpart can be solved with similarly efficient methods as the original nominal problem, which is for instance the case for most practical examples of robust mean-variance optimization. If no analytical reformulation is available the robust counterpart can still be solved rather efficiently based on numerical methods from semi-infinite programming, cf. Werner (2008).

It can be noted that in the special case  $U = \{u\}$  the robust counterpart coincides with the original problem instance  $(P_u)$ . In most applications, a suitably chosen neighborhood  $U(\hat{u})$ , e.g.  $U(\hat{u}) = B_\epsilon(\hat{u})$ , around some initially given nominal parameter  $\hat{u}$  is used as uncertainty set. For more details on the choice of the uncertainty sets in the context of mean-variance optimization, let us refer to Section 2.4.

**Remark 2.9.** The robust counterpart is a rather elegant approach as it allows to express both design and (obviously) parameter uncertainties in a unified manner. For instance, *design uncertainty*, i.e. small deviations in the implementation of the optimal solution, can be considered via the uncertainty

<sup>1</sup> More precisely, the robust counterpart approach was independently introduced by Ben-Tal and Nemirovski (1998, 1999, 2002) and by El Ghaoui and Lebret (1997, 1998).



$$g(x + u) \leq 0 \quad \forall u \in B_\delta(0)$$

instead of an original constraint  $g(x) \leq 0$ . This then guarantees that the optimal solution remains feasible after slight perturbations.

#### 2.4. Robust mean–variance optimization

Our main motivation for the consideration of multiobjective problems under data uncertainty stems from mean–variance portfolio optimization as laid out by Markowitz (1952, 1959). The basic idea is that a portfolio is solely characterized by the two quantities *risk* (mostly measured in terms of the variance or volatility) and *expected return*. Since an investor is seeking for an allocation with low risk and high expected return, a trade-off between these two conflicting aims has to be made.

Although the mean–variance framework is probably still the most popular approach for asset allocation, there is a large inherent risk in this framework, evolving from using possibly incorrect parameter values in the computation of the portfolio allocation. Markowitz (1952) already acknowledged that it is an open question how to derive the proper input parameters for the mean–variance model. Nowadays, this is often referred to as *estimation risk* in the academic literature, as usually the value that is finally used for the solution of the portfolio allocation problem is just an estimate (mostly based on a data sample) for the unknown true parameter. Due to this estimation risk, it is by now well-known that the concept of mean–variance efficient frontier has two major weaknesses, namely:

- It has been demonstrated (see e.g. Barry, 1974; Jorion, 1992 or Best & Grauer, 1991) that the optimal solution of the mean–variance problem strongly depends on the particular values of the expected returns and their covariances and that especially the expected returns have a significant influence<sup>2</sup> on the optimal asset allocation.
- In addition, Michaud (1998) pointed out that especially the right end of the efficient frontier (the *maximum return portfolio*) promises unreliable high expected returns for somewhat too low volatilities. This is typically due to over-optimistic optimization based on error-prone estimates, resulting in large deviations between ex-ante and ex-post returns and volatilities.

As estimation risk represents a popular and heavily researched topic, there are various approaches to account for estimation risk, especially in the context of portfolio optimization, see for example the early considerations by Jobson and Korkie (1981) or Best and Grauer (1991). To improve the allocation, many authors propose to use parameter estimators other than the classical maximum likelihood as these seem to be more robust and thus reduce the influence of estimation risk on the optimal portfolios. Bawa, Brown, and Klein (1979), Jobson and Korkie (1981) and also Efron and Morris (1975) use shrinkage or Stein-type estimators, while others (e.g., Jorion (1986)) suggest Bayesian estimators which combine a traditional parameter estimate with external prior information. More recently, robust estimators in more general terms are considered in Perret-Gentil and Victoria-Feser (2004), by Kempf and Memmel (2006) in the context of minimum variance portfolios and most recently by DeMiguel and Nogales (2009). A detailed comparison of the Bayesian approach to the popular

Black–Litterman approach together with its robustification can be found in Schöttle, Werner, and Zagst (2010).

Alternative types of optimization models that address the problem of estimation risk are stochastic programming and chance-constrained optimization which incorporate distributional assumptions and uncertainty about parameters into the optimization problem in terms of probabilistic constraint formulations, cf. Dupacova (1999) or Erdogan and Iyengar (2006) and the references therein, or sample-average multiobjective optimization, see Fliege and Xu (2011). Other approaches formulate the problem over a discrete set of alternatives (portfolios) and use tools from multicriteria decision aid (MCDA), see, e. g., Zopounidis and Doumpos (1999) and their review (Zopounidis & Doumpos, 1999).

To formalize the approach, let us consider a financial market with  $n$  risky assets defined on a suitable probability space in a single period setting. Since it is well-known that the mean–variance approach is consistent with utility theory under the assumption of multivariate elliptical distributions, see e.g. Ingersoll (1987), we assume that the linear one period return  $R \in \mathbb{R}^n$  is multivariate elliptically distributed with parameters  $\mu, \Sigma$  and characteristic generator  $\psi$ , i.e.

$$R \sim \mathcal{E}(\mu, \Sigma, \psi).$$

Further, we assume that  $R$  possesses a density and has finite second moments, and that  $2\psi'(0) = -1$ , which means

$$\mathbb{E}[R] = \mu \quad \text{and} \quad \text{cov}[R] = -2\psi'(0) \Sigma = \Sigma.$$

For more details on elliptical distributions, let us refer to Fang and Zhang (1990) or Fang, Kotz, and Ng (1990). Finally, let us assume that the set of feasible portfolios  $X$  is non-empty, convex and compact and that portfolio weights add up to one, i.e.

$$X \in \mathcal{K}(\{x \in \mathbb{R}^n | x^\top \mathbf{1} = 1\}).$$

Thus, all additional constraints on the feasible portfolios which do not depend on  $\mu$  and  $\Sigma$  are subsumed in  $X$ . Compactness of  $X$  is not a very strong restriction, since it is usually fulfilled in practical settings, for example if some short-selling constraints (i.e.  $x \geq 0$ ) are present.

If asset returns  $R$  are distributed with mean  $\mu$  and covariance matrix  $\Sigma$ , the expected return of a portfolio  $x \in X$  equals  $m(x) := \mathbb{E}[R^\top x] = \mu^\top x$  and the corresponding risk figure in the mean–variance framework is given by the variance of the portfolio return  $s(x) := \text{Var}[R^\top x] = x^\top \Sigma x$ . Translating the original work of Markowitz in our context, Markowitz suggests to solve the portfolio allocation problem, i.e. the computation of the whole efficient frontier of the following multiobjective optimization problem

$$\text{efmin}_{x \in X} \begin{pmatrix} s(x) \\ -m(x) \end{pmatrix} = \text{efmin}_{x \in X} \begin{pmatrix} x^\top \Sigma x \\ -\mu^\top x \end{pmatrix} \quad (\text{MV})$$

by the following procedure: start with the minimum variance portfolio  $x_{MV}^*$  which is obtained as the unique solution of

$$\min_{x \in X} s(x) = \min_{x \in X} x^\top \Sigma x$$

and then increase the required expected return level from the expected return of the minimum variance portfolio

$$\mu_{\min} = m(x_{MV}^*) = \mu^\top x_{MV}^*$$

to the expected return of the maximum return portfolio  $x_{MR}^*$  given by

$$\mu_{\max} = \max_{x \in X} m(x) = \max_{x \in X} \mu^\top x,$$

computing the corresponding efficient portfolios for each required expected return level. In modern language this means that Markowitz proposed to solve the multiobjective problem by the

<sup>2</sup> This rather general statement is true in most practical situations, see e.g. Chopra and Ziemba (1993), but of course it depends – among others factors – on the particular choice of the risk aversion parameter. It further depends on the ratio of the number of assets to the size of the historical data sample, as shown recently by Kan and Zhou (2007).

$\varepsilon$ -constraint method (see Section 2.2 for more details), i.e. he proposed to trace the efficient frontier by solving the following family of convex optimization problems for varying target levels of return  $\mu_T \in [\mu_{\min}, \mu_{\max}]$ ,

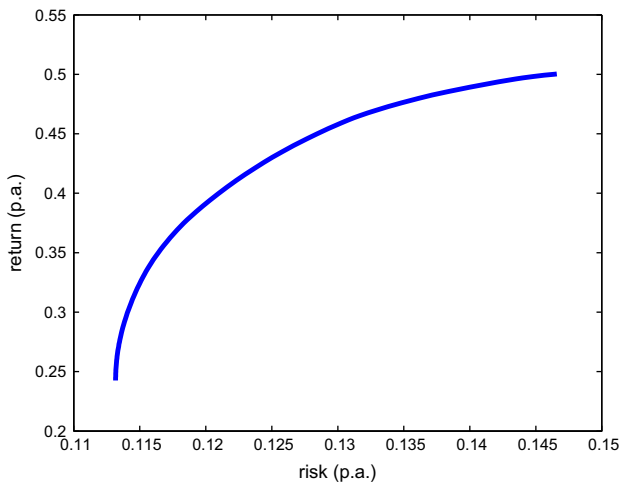
$$\begin{aligned} \min_{x \in X} \quad & x^\top \Sigma x \\ \text{s.t.} \quad & \mu_T - \mu^\top x \leq 0, \end{aligned} \quad (\text{MV}_{\mu_T})$$

In Fig. 1 we have illustrated such a resulting efficient frontier in the upper part together with the according portfolio allocations  $x^*(\mu_T)$  in the lower part, cf. Section 4 for more details, where  $\mu_T$  ranges from 6.3% to 10%.

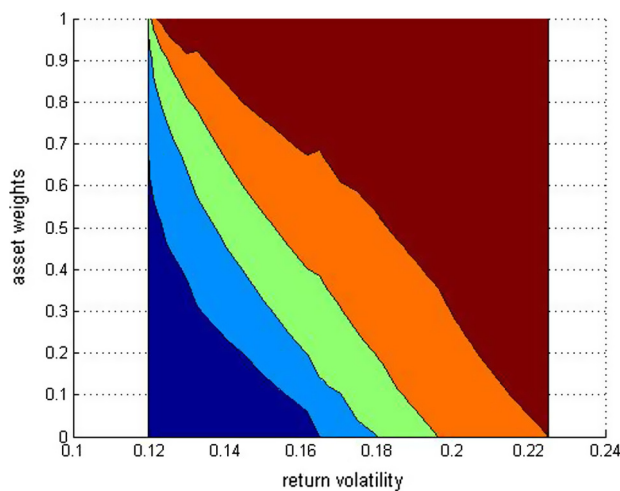
As we have seen in Section 2.2, one could alternatively solve the following family of optimization problems:

$$\min_{x \in X} (1 - \lambda)x^\top \Sigma x + \lambda(-\mu^\top x) \quad (\text{MV}_\lambda)$$

for  $0 \leq \lambda \leq 1$ . In Section 2.2 more details on the relationship between the original multiobjective problem, its  $\varepsilon$ -constraint formulation and its weighted-sum scalarization were already given.



(a) efficient frontier



(b) portfolio allocations

**Fig. 1.** Illustration of an efficient frontier and its associated optimal portfolio allocations. The percentage allocation of assets 1–6 in the optimal allocation  $x^*(\mu_T)$  have been illustrated in different colors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Remark 2.10.** In the specific situation outlined above the fraction  $\lambda(1 - \lambda)$  has two nice and well-known interpretations:

- It can be interpreted as the Lagrange multiplier to the corresponding constraint in problem  $(\text{MV}_{\mu_T})$  for suitably chosen return level  $\mu_T$ .
- It can also be interpreted as the risk-aversion factor from the corresponding logarithmic utility function.

As we have already pointed out, the parameter  $u := (\mu, \Sigma) \in \mathbb{R}^n \times \mathbb{S}_+^n$  is not known exactly but subject to uncertainty. Hence some appropriate method to deal with the uncertainty in problem  $(\text{MV})$  is critical. However, so far, uncertainty has not been considered directly in the formulation  $(\text{MV})$ , but only in its scalarized version  $(\text{MV}_{\mu_T})$  and  $(\text{MV}_\lambda)$ . For this purpose we have to choose a joint uncertainty set  $U$  for the uncertain data  $u = (\mu, \Sigma)$ . For simplicity of the exposition that follows, we choose an ellipsoid around a nominal point  $(\hat{\mu}, \hat{\Sigma})$  of size  $\delta$ :

$$U_\delta(\hat{\mu}, \hat{\Sigma}) := \{(\mu, \Sigma) \in \mathbb{R}^n \times \mathbb{S}_+^n : \|\mu - \hat{\mu}\| + c\|\Sigma - \hat{\Sigma}\| \leq \delta\}. \quad (2.7)$$

Of course, more sophisticated uncertainty sets can be chosen which lead to other robust counterparts. For a more detailed discussion on the choice of the uncertainty set let us refer to Schöttle and Werner (2006, 2009). Based on the joint uncertainty set  $U_\delta(\hat{\mu}, \hat{\Sigma})$  we obtain the robust counterpart of  $(\text{MV}_{\mu_T})$  as

$$\begin{aligned} \min_{x \in X} \quad & \max_{(\mu, \Sigma) \in U_\delta(\hat{\mu}, \hat{\Sigma})} x^\top \Sigma x \\ \text{s.t.} \quad & \max_{(\mu, \Sigma) \in U_\delta(\hat{\mu}, \hat{\Sigma})} \mu_T - \mu^\top x \leq 0, \end{aligned} \quad (\text{RMV}_{\mu_T})$$

and the robust counterpart of  $(\text{MV}_\lambda)$  as

$$\min_{x \in X} \max_{(\mu, \Sigma) \in U_\delta(\hat{\mu}, \hat{\Sigma})} (1 - \lambda)x^\top \Sigma x + \lambda(-\mu^\top x) \quad (\text{RMV}_\lambda)$$

The problem family  $(\text{RMV}_\lambda)$  has been considered in great detail in Schöttle and Werner (2009). There, it has been shown that the resulting robust frontier (i.e. the efficient portfolios) is actually a 'shortened' version of the original efficient frontier if the joint uncertainty set  $U$  is chosen in a statistically meaningful way. However, the family of robust problems  $(\text{RMV}_{\mu_T})$  has not yet been investigated in this manner. Therefore, let us consider  $(\text{RMV}_{\mu_T})$  more closely: First, it can be observed that now the required expected return level  $\mu_T^r$  has to be chosen from the interval  $[\mu_{\min}^r, \mu_{\max}^r]$  where the robust minimum variance portfolio  $x_{\text{RMV}}^*$  is obtained as the unique solution of

$$\min_{x \in X} \max_{(\mu, \Sigma) \in U_\delta(\hat{\mu}, \hat{\Sigma})} x^\top \Sigma x$$

and taking into account signs (i.e. switching from max to min),

$$\mu_{\min}^r = \min_{(\mu, \Sigma) \in U_\delta(\hat{\mu}, \hat{\Sigma})} \mu^\top x_{\text{RMV}}^*$$

to the robust expected return of the robust maximum return portfolio

$$\mu_{\max}^r = \max_{x \in X} \min_{(\mu, \Sigma) \in U_\delta(\hat{\mu}, \hat{\Sigma})} \mu^\top x.$$

Second, using results on the  $\varepsilon$ -constraint method of convex multiobjective problems, see Theorem 2.5, it can be easily shown that the family of problems  $(\text{RMV}_{\mu_T})$  is equivalent to the multiobjective problem

$$\text{efmin}_{x \in X} \begin{pmatrix} \max_{(\mu, \Sigma) \in U_\delta(\hat{\mu}, \hat{\Sigma})} x^\top \Sigma x \\ \max_{(\mu, \Sigma) \in U_\delta(\hat{\mu}, \hat{\Sigma})} -\mu^\top x \end{pmatrix} \quad (\text{GRMV})$$

which in turn is – by weighted-sum scalarization – equivalent to the family of problems

$$\min_{x \in X} (1 - \lambda) \max_{(\mu, \Sigma) \in U_\delta(\hat{\mu}, \hat{\Sigma})} x^\top \Sigma x + \lambda \max_{(\mu, \Sigma) \in U_\delta(\hat{\mu}, \hat{\Sigma})} -\mu^\top x. \quad (\text{GRMV}_\lambda)$$

As the families of optimization problems  $(\text{GRMV}_\lambda)$  and  $(\text{RMV}_\lambda)$  do not have the same objective function, it cannot be expected that they possess the same solutions nor that they yield the same efficient frontier. Further, it is not clear yet if one formulation is advantageous over the other, and for what reasons.

### 3. Robust multiobjective optimization

As we have already motivated, alongside the single-objective problem with uncertainty  $(P_u)$  we also need to consider a more general setting, where now the objective  $f$  becomes multi-dimensional, i.e.  $f: X \times \bar{U} \rightarrow \mathbb{R}^p$  with  $p \geq 2$ , and thus  $(P_u)$  becomes a *multiobjective optimization problem*. In this framework we are then looking for efficient points with respect to a partial ordering induced by an ordering cone  $K \subset \mathbb{R}^p$ . In the following, we only consider the standard cone  $K = \mathbb{R}_+^p$ .

#### 3.1. Robustification of multiobjective problems

Instead of the single-objective problem  $(P_u)$  let us consider its multiobjective extension

$$\begin{aligned} \text{efmin}_{x \in X} \quad & f(x, u) \\ \text{s.t.} \quad & g(x, u) \leq 0, \end{aligned} \quad (\text{MOP}_u)$$

Subsequently, we denote the efficient frontier of  $(\text{MOP}_u)$  as  $\text{eff}_K(\text{MOP}_u)$  and its corresponding preimage, i.e. the set of corresponding optimal solutions, as  $\text{eff}_K^x(\text{MOP}_u)$ . In the following, we usually do not distinguish between efficiency of points  $f(x)$  in the image space or the corresponding preimage  $x \in X$  if it is clear from the context which space is referred to.

**Remark 3.1.** Markowitz called a *portfolio allocation*  $x \in X$  *efficient* if its image under the following mapping

$$f: x \mapsto \begin{pmatrix} x^\top \Sigma x \\ -\mu^\top x \end{pmatrix}$$

is efficient according to the above definition with respect to the partial ordering induced by the cone  $K = \mathbb{R}_+^2$ .

We have already discussed that there are two popular scalarization methods to calculate the whole efficient frontier  $\text{eff}_K(\text{MOP}_u)$ :

- the  $\varepsilon$ -constraint scalarization method (ECS), and
- the weighted-sum scalarization (WSS).

As we have demonstrated in [Theorems 2.3 and 2.5](#), both scalarization methods yield the same efficient frontier (modulo closure) as  $(\text{MOP}_u)$  under reasonably mild assumptions. Unfortunately, this is no longer true in the robust multiobjective case:

**Example 3.2.** Let us consider the following two-dimensional linear multiobjective problem with linear uncertainty with feasible set  $X = [-1, 1]$  and uncertainty set  $U = [-1, 1]$ :

$$\text{efmin}_{x \in [-1, 1]} \begin{pmatrix} ux \\ -ux \end{pmatrix}$$

Let us consider the two approaches as introduced above:

1. Repeating the above steps, i.e. first scalarization via the ECS method, then robustification, and finally re-transformation back to a multiobjective formulation leads to

$$\text{efmin}_{x \in [-1, 1]} \begin{pmatrix} |x| \\ |x| \end{pmatrix}$$

as  $\max_{u \in [-1, 1]} ux = |x|$ . The only optimal solution to this multiobjective problem is obviously  $x^* = 0$  which yields the robust efficient frontier  $\{(0, 0)^\top\}$ .

2. However, if instead of ECS we use WSS and then robustify the scalarization obtained by the WSS method, we obtain

$$\min_{x \in [-1, 1]} \max_{u \in [-1, 1]} \lambda ux + (1 - \lambda)(-ux)$$

which is, as easy calculations show, the same as

$$\min_{x \in [-1, 1]} |x| |2\lambda - 1|.$$

For  $\lambda \neq \frac{1}{2}$  only  $x^* = 0$  is optimal, but for  $\lambda = \frac{1}{2}$  all  $x \in [-1, 1]$  are actually optimal.

Thus, both approaches lead, as proposed, to completely different solution sets, even in a very simple linear setting.

In the previous section we have encountered a similar situation, where the transformation of the multiobjective portfolio optimization problem to a family of scalarized problems is not consistent with robustification, i.e. the operations *robustification* and *scalarization* do not commute.

To understand the reason for this surprising behavior let us consider robustification and scalarization for  $(\text{MOP}_u)$  in its general form. The ECS method allows to equivalently rewrite  $(\text{MOP}_u)$  as

$$\begin{aligned} \min_{x \in X} \quad & f_1(x, u) \\ \text{s.t.} \quad & f_i(x, u) \leq \varepsilon_i \quad \text{for } i = 2, \dots, p, \\ & g_j(x, u) \leq 0 \quad \text{for } j = 1, \dots, m. \end{aligned} \quad (\text{MOP}_{u, \varepsilon})$$

with suitably chosen  $\varepsilon_i$ . Note that the same parameter  $u$  is used in all  $f_i$  and in all  $g_j$  in  $(\text{MOP}_{u, \varepsilon})$ . Let us call  $(\text{MOP}_{u, \varepsilon})$  a *traditional instance* of a parametric optimization problem. The robust counterpart to this problem becomes

$$\begin{aligned} \min_{x \in X} \quad & \max_{u \in U} f_1(x, u) \\ \text{s.t.} \quad & \max_{u \in U} f_i(x, u) \leq \varepsilon_i \quad \text{for } i = 2, \dots, p, \\ & \max_{u \in U} g_j(x, u) \leq 0 \quad \text{for } j = 1, \dots, m. \end{aligned} \quad (\text{RMOP}_{U, \varepsilon})$$

Now, let us consider what we like to call a *generalized instance* of the original problem, where now each function  $f_i$  and constraint  $g_j$  gets its own copy of  $u$ :

$$\begin{aligned} \min_{x \in X} \quad & f_1(x, u_1) \\ \text{s.t.} \quad & f_i(x, u_i) \leq \varepsilon_i \quad \text{for } i = 2, \dots, p, \\ & g_j(x, u_{p+j}) \leq 0 \quad \text{for } j = 1, \dots, m. \end{aligned} \quad (\text{GMOP}_{U, \varepsilon})$$

with  $u_i \in U, i = 1, \dots, p + m$ , or equivalently  $(u_1, \dots, u_{p+m}) \in U^{p+m}$ . Interestingly, the robust counterpart of  $(\text{GMOP}_{U, \varepsilon})$  is the same as the robust counterpart of  $(\text{MOP}_{u, \varepsilon})$  which is  $(\text{RMOP}_{U, \varepsilon})$ . This means that after robustification we cannot distinguish any longer if all  $f_i$  shared the same parameter  $u$  or if they all had an individual parameter  $u_i$  from the same uncertainty set. (This observation was already made by [Ben-Tal & Nemirovski \(1998\)](#).) We have already seen in the previous example that due to this feature of robustification the operations scalarization and robustification cannot commute in general.

**Remark 3.3.** If we now equivalently re-transform the robust counterpart of the two  $\varepsilon$ -constraint families back to a multiobjective problem, we obtain in *both* cases:

$$\text{efmin}_{\substack{x \in X \\ g_U^{\text{RC}}(x) \leq 0}} \begin{pmatrix} \max_{u \in U} f_1(x, u) \\ \dots \\ \max_{u \in U} f_p(x, u) \end{pmatrix}.$$

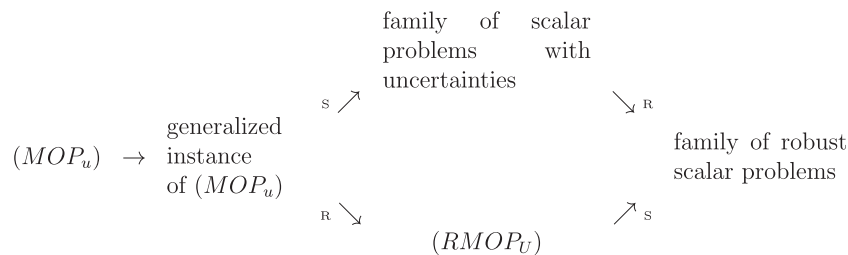
This observation leads us to a canonical definition of a robust counterpart of a multiobjective problem:

**Definition 3.4.** The *multiobjective robust counterpart* for the multiobjective problem with uncertainties  $(\text{MOP}_u)$  is defined as

$$\text{efmin}_{x \in X} f_U^{\text{RC}}(x) \quad (\text{RMOP}_U)$$

$$\text{s.t. } g_U^{\text{RC}}(x) \leq 0$$

with



$$f_U^{\text{RC}}(x) := \begin{pmatrix} \max_{u \in U} f_1(x, u) \\ \dots \\ \max_{u \in U} f_p(x, u) \end{pmatrix} \quad \text{and} \quad g_U^{\text{RC}}(x) := \begin{pmatrix} \max_{u \in U} g_1(x, u) \\ \dots \\ \max_{u \in U} g_m(x, u) \end{pmatrix}$$

i.e. each component of the objective function and the constraints is replaced by its robust counterpart. The efficient frontier of  $(\text{RMOP}_U)$  is referred to as *robust efficient frontier* and denoted by  $\text{eff}_K(\text{RMOP}_U)$  in the following.

We note that for  $p = 1$  the definition of the multiobjective robust counterpart coincides with the definition of the usual robust counterpart and hence provides a proper generalization of this concept to multiobjective optimization. Further, the robustification of the  $p$ -dimensional objective is now the very same as the robustification of the  $m$ -dimensional constraint.

**Remark 3.5.** Using this definition of a robust counterpart, we have already seen that if  $(\text{MOP}_u)$  is given as a generalized instance, robustification and  $\varepsilon$ -scalarization commute, i.e. it holds:

$$\text{ECS}(\text{robustification}(\text{MOP}_u)) = \text{robustification}(\text{ECS}(\text{MOP}_u))$$

Let us now consider the case of WSS and see if the above result also holds in this case. The WSS of  $(\text{RMOP}_U)$  is given by

$$\min_{\substack{x \in X \\ g_U^{\text{RC}}(x) \leq 0}} \sum_{i=1}^p \lambda_i \max_{u \in U} f_i(x, u)$$

with  $\sum_{i=1}^p \lambda_i = 1, \lambda_i > 0$ . Obviously, this is equivalent to

$$\min_{\substack{x \in X \\ g_U^{\text{RC}}(x) \leq 0}} \max_{(u_1, \dots, u_p) \in U^p} \sum_{i=1}^p \lambda_i f_i(x, u_i), \quad (\text{RMOP}_{U, \lambda})$$

but quite different to the robust counterpart of the weighted-sum scalarization of  $(\text{MOP}_u)$ :

$$\min_{\substack{x \in X \\ g_U^{\text{RC}}(x) \leq 0}} \max_{u \in U} \sum_{i=1}^p \lambda_i f_i(x, u).$$

We can summarize the above discussion in the following proposition:

**Proposition 3.6.** Consider the multiobjective problem with uncertainties  $(\text{MOP}_u)$ . Then we have the following.

1. Robustifying  $(\text{MOP}_u)$  and then applying ECS leads to the same family of scalar problems as applying ECS first to  $(\text{MOP}_u)$  and the robustifying the result:  $\text{ECS}(\text{robustification}(\text{MOP}_u)) = \text{robustification}(\text{ECS}(\text{MOP}_u))$ .
2. Robustifying  $(\text{MOP}_u)$  and then applying WSS leads to the same family of scalar problems as applying WSS first to  $(\text{MOP}_u)$  and the robustifying the result:  $\text{WSS}(\text{robustification}(\text{MOP}_u)) = \text{robustification}(\text{WSS}(\text{MOP}_u))$ .

We can visualize the results of this proposition by the commutative diagram

where arrows labeled with “R” denote a robustification operation and arrows labeled with “S” denote a scalarization. Of course, the actual type of scalar problems depends on the scalarization strategy in use.

**Remark 3.7.** The above proposition is the key observation to the numerical solution of  $(\text{RMOP}_U)$ . It states that  $(\text{RMOP}_U)$  can be solved via the numerical solution of a family of ordinary scalar robust optimization problems. Therefore, it is open to a large variety of methods designed to solve robust counterpart problems, see e.g. Ben-Tal et al. (2009).

**Remark 3.8.** For the following let us also introduce the notion of the *robust image* of the efficient frontier  $\text{eff}_K(\text{MOP}_u)$  as the image of  $\text{eff}_K^x(\text{MOP}_u) \cap \{x \in X | g_U^{\text{RC}}(x) \leq 0\}$  under  $f_U^{\text{RC}}$ , abbreviated as  $\text{eff}_K^{\text{RC}}(\text{MOP}_u)$ . We immediately note the following properties:

- $\text{eff}_K^{\text{RC}}(\text{MOP}_u)$  is in general not an efficient frontier itself.
- For the special case of the robust counterpart we have  $\text{eff}_K(\text{RMOP}_U) = \text{eff}_K^{\text{RC}}(\text{RMOP}_U)$ .

### 3.2. An alternative motivation for the robust counterpart

On first sight it might seem that the construction of the generalized robust counterpart is quite artificial and based on the very specific  $\varepsilon$ -constraint method. However, in the following we will provide an alternative motivation for the generalized robust counterpart by considerations which are independent to the  $\varepsilon$ -constraint method and motivated by a pure multiobjective programming point of view. For this purpose, let us rewrite the original problem



$$\begin{aligned} \min_{x \in X} \quad & f(x, u) \\ \text{s.t.} \quad & g(x, u) \leq 0, \end{aligned} \quad (\text{P}_u)$$

as an optimization with linear objective function that does not depend on  $u$  by introducing an auxiliary variable  $\alpha \in \mathbb{R}$

$$\begin{aligned} \min_{\substack{x \in X \\ \alpha \in \mathbb{R}}} \quad & \alpha \\ \text{s.t.} \quad & f(x, u) - \alpha \leq 0, \\ & g(x, u) \leq 0. \end{aligned} \quad (\text{Q}_u)$$

The robust counterpart of  $(\text{Q}_u)$  is given by

$$\begin{aligned} \min_{\substack{x \in X \\ \alpha \in \mathbb{R}}} \quad & \alpha(\text{RQ}_u) \\ \text{s.t.} \quad & \max_{u \in U} f(x, u) - \alpha \leq 0, \\ & \max_{u \in U} g_i(x, u) \leq 0, \text{ for } i = 1, \dots, m. \end{aligned}$$

which is obviously equivalent to the robust counterpart  $(\text{RP}_u)$  of  $(\text{P}_u)$ .

Now, the same methodology can be applied to the multiobjective problem  $(\text{MOP}_u)$ , which then becomes

$$\begin{aligned} \text{efmin}_{\substack{x \in X \\ \alpha \in \mathbb{R}^p}} \quad & \alpha(\text{MOQ}_u) \\ \text{s.t.} \quad & f(x, u) - \alpha \leq_K 0, \\ & g(x, u) \leq 0. \end{aligned} \quad (\text{MOQ}_u)$$

after introduction of an auxiliary variable  $\alpha \in \mathbb{R}^p$ . Note that the only difference to  $(\text{Q}_u)$  is due to the  $p$ -dimensional conic<sup>3</sup> inequality  $f(x, u) \leq_K \alpha$  instead of the ordinary one-dimensional inequality  $f(x, u) \leq \alpha$  before. As we now deal with a conic inequality, its robustified version

$$f(x, u) - \alpha \leq_K 0 \quad \forall u \in U$$

cannot be rewritten in a similar manner as before. To proceed, let us fix  $x$  and let us consider the problem in  $\alpha$  only for the time being:

$$\begin{aligned} \text{efmin}_{\alpha \in \mathbb{R}^p} \quad & \alpha \\ \text{s.t.} \quad & f(x, u) \leq_K \alpha \quad \forall u \in U. \end{aligned}$$

Interestingly, for the standard cone  $K = \mathbb{R}_+^p$  this problem has a unique solution  $\alpha^*$ , given by

$$\alpha^* = \begin{pmatrix} \max_{u \in U} f_1(x, u) \\ \dots \\ \max_{u \in U} f_p(x, u) \end{pmatrix} = f_U^{\text{RC}}(x).$$

This immediately yields that the multiobjective robust counterpart of  $(\text{MOQ}_u)$  is identical to  $(\text{RMOP}_u)$ . Thus, the multiobjective robust counterpart seems to be the canonical generalization of the concept of a robust counterpart to multiobjective optimization. In the following theorem we summarize the insights gained from the above.

**Theorem 3.9.** *Let  $g(x, u) = g(x)$ , i.e. let the feasible set of  $(\text{MOP}_u)$  not be subject to uncertainty, and suppose  $K = \mathbb{R}_+^p$ . Then the robust efficient frontier is "bracketed" between the nominal efficient frontier and its robust image in the following sense:*

1. For all  $u \in U$  and all  $x_r^* \in \text{eff}_K^x(\text{RMOP}_u)$  there exists an  $\hat{x} \in \text{eff}_K^x(\text{MOP}_u)$  such that
 
$$f(\hat{x}, u) \leq f^{\text{RC}}(x_r^*).$$

2. For all  $u \in U$  and all  $\hat{x} \in \text{eff}_K^x(\text{MOP}_u)$  there exists an  $x_r^* \in \text{eff}_K^x(\text{RMOP}_u)$  such that
 
$$f^{\text{RC}}(x_r^*) \leq f^{\text{RC}}(\hat{x}).$$

**Proof.**

1. From the definition of  $f^{\text{RC}}$  we have  $f(x_r^*, u) \leq f^{\text{RC}}(x_r^*)$ . If  $x_r^*$  is already efficient for  $u$ , i.e. if  $x_r^* \in \text{eff}_K^x(\text{MOP}_u)$ , the result follows immediately. Otherwise, if  $x_r^*$  is not efficient for  $(\text{MOP}_u)$ , we can find an  $\hat{x} \in \text{eff}_K^x(\text{MOP}_u)$  which dominates  $f(x_r^*, u)$ , i.e.  $f(\hat{x}, u) \leq f(x_r^*, u)$ , which proves the first statement.
2. If  $\hat{x}$  is already robust efficient, i.e. if  $\hat{x} \in \text{eff}_K^x(\text{RMOP}_u)$ , the result follows immediately. Otherwise  $f^{\text{RC}}(\hat{x})$  is not robust efficient, i.e. there exists an  $x_r^* \in \text{eff}_K^x(\text{RMOP}_u)$  which dominates  $\hat{x}$ , i.e.  $f^{\text{RC}}(x_r^*) \leq f^{\text{RC}}(\hat{x})$ .  $\square$

This bracketing of the robust efficient frontier between the nominal efficient frontier and its robust image is illustrated in Fig. 3. We note that the robust efficient frontier and the robust image of the original nominal frontier might shrink and move to the top right significantly. Further, as we have seen from the proof of Part 1 of Theorem 3.9, the statement will remain true even if the feasible set is also subject to uncertainty. However, in case of uncertain feasibility set, Part 2 of Theorem 3.9 will no longer hold true for all  $\hat{x} \in \text{eff}_K^x(\text{MOP}_u)$  but only for those which are also feasible in  $(\text{RMOP}_u)$ . Therefore, the following corollary applies:

**Corollary 3.10.** *Let  $K = \mathbb{R}_+^p$ . Then the robust efficient frontier is "bracketed" between the nominal efficient frontier and its robust image, i.e. it holds:*

1. For all  $u \in U$  and all  $x_r^* \in \text{eff}_K^x(\text{RMOP}_u)$  there exists an  $\hat{x} \in \text{eff}_K^x(\text{MOP}_u)$  such that
 
$$f(\hat{x}, u) \leq f^{\text{RC}}(x_r^*).$$
2. For all  $u \in U$  and all  $\hat{x} \in \text{eff}_K^x(\text{MOP}_u)$  with  $g_U^{\text{RC}}(\hat{x}) \leq 0$  (i.e.  $f_U^{\text{RC}}(\hat{x}) \in \text{eff}_K^{\text{RC}}(\text{MOP}_u)$ ) there exists an  $x_r^* \in \text{eff}_K^x(\text{RMOP}_u)$  such that
 
$$f^{\text{RC}}(x_r^*) \leq f^{\text{RC}}(\hat{x}).$$

#### 4. An illustrative example

For an illustration of the above concepts, we have chosen a simple example from mean–variance optimization. Although the following results can also be obtained with a more realistic example and sophisticated choices of uncertainty sets, the basic idea can be similarly well illustrated in such a simplified way. If not mentioned otherwise, it is easy to show that all conclusions remain valid in more general cases as well.

Let us consider  $n = 6$  assets with nominal parameters, i.e. expected returns and volatilities,

$$\hat{\mu} = \begin{pmatrix} 0.05 \\ 0.06 \\ 0.07 \\ 0.08 \\ 0.09 \\ 0.10 \end{pmatrix}, \quad \hat{\sigma} = \begin{pmatrix} 0.100 \\ 0.125 \\ 0.150 \\ 0.175 \\ 0.200 \\ 0.225 \end{pmatrix}$$

and pairwise return correlations equal to  $\hat{\rho} = 0.60$ . Further, we let  $X = \{x \in \mathbb{R}_+^n | x^T \mathbf{1} = 1\}$ .

The resulting efficient frontier, e.g. obtained by the numerically solving problems from the weighted-sum scalarization, has already been illustrated in Fig. 1.

<sup>3</sup> We emphasize the meaning of the inequality  $\leq_K$  by explicitly mentioning the cone  $K$  to ease understanding, but actually this is nothing else than the ordinary inequality  $\leq$  which has been used for the constraints  $g$  throughout.

Taking into account uncertainty in expected returns and in the covariance matrix, we use the previously introduced uncertainty set

$$U_\delta(\hat{\mu}, \hat{\Sigma}) = \{(\mu, \Sigma) \in \mathbb{R}^n \times \mathbb{S}_+^n \mid \|\mu - \hat{\mu}\| + c\|\Sigma - \hat{\Sigma}\| \leq \delta\}$$

with  $\delta = 0.0075$  and  $c = 2$ . This leads to the corresponding robust multiobjective mean–variance formulation

$$\operatorname{efmin}_{x \in X} \begin{pmatrix} s^{RC}(x) \\ -m^{RC}(x) \end{pmatrix} = \operatorname{efmin}_{x \in X} \begin{pmatrix} x^\top \hat{\Sigma} x + \frac{\delta}{c} \|x\|^2 \\ -\hat{\mu}^\top x + \delta \|x\| \end{pmatrix}.$$

Here, based on the specific choice of the uncertainty set (2.7), the robustified versions (i.e. the robust counterparts) of  $s(x) = x^\top \Sigma x$  and  $-m(x) = -\mu^\top x$  can be analytically obtained as

$$s^{RC}(x) = \max_{(\mu, \Sigma) \in U_\delta(\hat{\mu}, \hat{\Sigma})} x^\top \Sigma x = x^\top \left( \hat{\Sigma} + \frac{\delta}{c} I \right) x = x^\top \hat{\Sigma} x + \frac{\delta}{c} \|x\|^2$$

and

$$-m^{RC}(x) = \max_{(\mu, \Sigma) \in U_\delta(\hat{\mu}, \hat{\Sigma})} -\mu^\top x = -\hat{\mu}^\top x + \delta \|x\|,$$

see Schöttle and Werner (2009) for more details on the corresponding calculations. Hence, scalarizations of the robust counterpart can be solved in a numerically efficient manner by standard optimization software.

**Remark 4.1.** As already mentioned before, this is one of the cases where the robustified versions can be obtained in analytical form due to the specific interplay of the choice of the uncertainty set and the linear dependence on the uncertain parameters. In the given special situation it can be observed that robustification acts as a penalty for missing diversification as the norm of a portfolio can be interpreted as a measure of its diversification. The better the diversification, the lower the penalty due to robustification. Therefore, the penalty term is directly linked to the sensitivity of the functions with respect to the uncertain parameters, the less sensitive, the less the penalty.

Although both observations are very specific to the example at hand, similar interpretations of robustification still hold true, see e.g. Schöttle and Werner (2006) for a practical discussion.

Fig. 2 illustrates for three specific portfolios (the minimum variance portfolio obtained with parameter  $\lambda = 0$ ; some average efficient portfolio corresponding to  $\lambda = 0.25$ ; and the maximum return portfolio with  $\lambda = 1$ ) how the return and volatility estimate of the resulting optimal portfolio (obtained by the solution of the corresponding optimization problem with nominal data  $\hat{\mu}$  and  $\hat{\Sigma}$ )

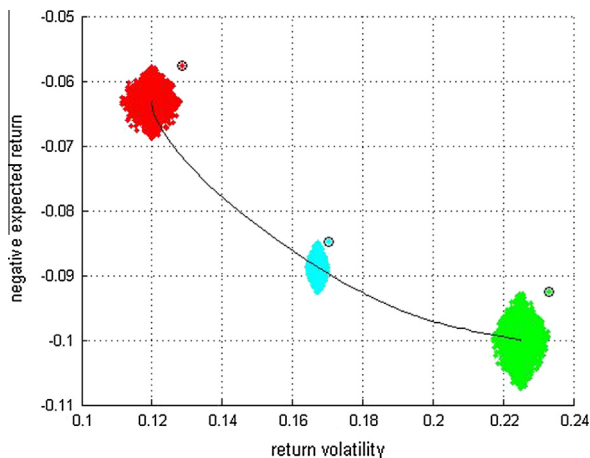


Fig. 2. Illustration of uncertainty for selected efficient portfolios.

fluctuate. The colored sets represent the images of the uncertainty set  $U_\delta(\hat{\mu}, \hat{\Sigma})$  under the two objective functions  $\sqrt{s(x)}$  and  $-m(x)$ . Additionally, the three corresponding elements in the robust image, i.e. the images of the portfolios under the robustified objective are illustrated (as small circles). According to Theorem 3.9 these are always located to the top right of the image of the uncertainty set and to the top right of the original element of the efficient frontier.

The location of the robust efficient frontier (in red) in comparison to the original nominal frontier (in black) is illustrated in Fig. 3.

**Remark 4.2.** Interestingly, the robust efficient frontier passes through the corresponding robust images of the average and the maximum return portfolio, but not through the robust image of the minimum variance portfolio. This immediately implies that the minimum variance portfolio cannot be robust efficient. This example shows that in general a nominal efficient solution does not necessarily need to be robust efficient as well.

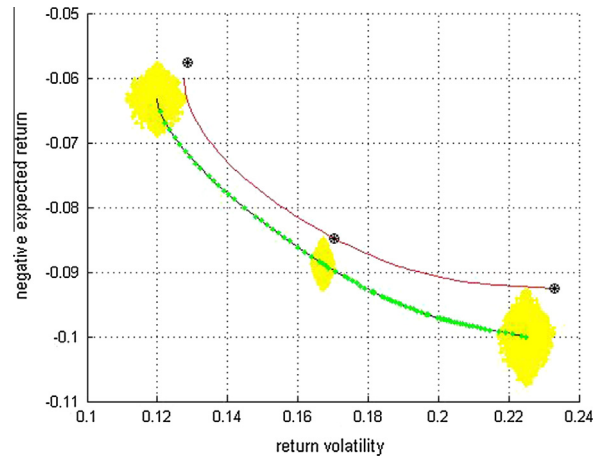


Fig. 3. Illustration of the robust efficient frontier (red), compared to the original nominal frontier (black) and the image of the robust efficient portfolios under the nominal parameter (green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

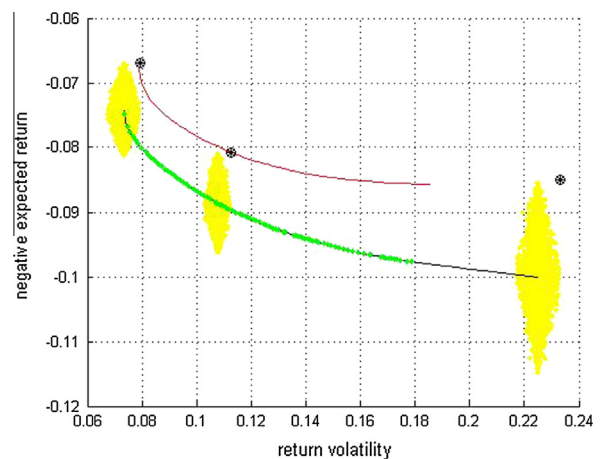


Fig. 4. Illustration of the robust efficient frontier (red), compared to the original nominal frontier (black) and the image of the robust efficient portfolios under the nominal parameter (green) based on altered nominal data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

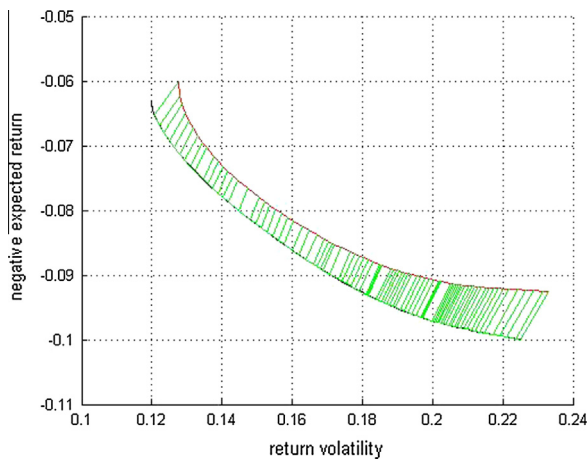


Fig. 5. Illustration of the efficiency loss of the robust efficient frontier compared to the original nominal frontier.

However, this observation depends on the specific problem data. Setting the correlation  $\hat{\rho} = 0$ ,  $\delta = 0.015$  and  $c = 4$ , a different picture can be obtained, cf. Fig. 4. Now the minimum variance portfolio itself is already robust. However, more risky portfolios further down the efficient frontier are no longer robust; the robust efficient frontier gets significantly shortened compared to the original frontier. Due to the specific parameter choice, this example falls into the framework of Schöttle and Werner (2009) where this result has been obtained analytically.

For practical purposes, it is helpful to link the robust efficient frontier to the nominal performance and to visualize this in comparison to the nominal efficient frontier. Figures like Fig. 5 help the decision maker in finding an appropriate solution, as the trade-off between optimality and robustness is easily obtainable from the figure. In this representation, each element of the robust frontier is linked to its corresponding nominal element by a line segment, which gives an idea about the efficiency lost due to uncertainty. The length of the line segment represents the loss of efficiency due to optimizing a robustified objective instead of the true objective. Such an illustration, interpretation and decomposition of the efficiency loss due to uncertainty and due to optimizing a robustified objective is easily obtained in our multiobjective framework. We have thus extended the scalar case by taking into account additional information that would not have been available beforehand.

## 5. Conclusion

Motivated by the well-known strong data dependency of mean-variance optimization, we have investigated how to treat convex multiobjective optimization problems with uncertain parameters. We have adopted the robust counterpart concept of Ben-Tal and Nemirovski and generalized it to the multiobjective setting. It has been demonstrated that this generalization seems to be the most natural concept in this context. It has further been emphasized that scalarization and robustification only commute if generalized instances are considered – which nevertheless lead to the same (multiobjective) robust counterpart. This is also the key observation to the numerical solution of the robust multiobjective problem which boils down to the solution of a family of ordinary scalar robust problems. It remains an open question for future research how to characterize the differences which arise if one would start with traditional problem instances instead. In addition we have demonstrated that the robust efficient frontier can be easily bracketed by the original nominal frontier and its robust image.

Based on an example from portfolio theory it was shown that this allows for a better interpretation of the efficiency loss due to robustification.

## References

- Arrow, K. J., Barankin, E. W., & Blackwell, D. (1953). Admissible points of convex sets. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the theory of games* (pp. 481–492). Princeton, New Jersey: Princeton University Press.
- Barry, C. B. (1974). Portfolio analysis under uncertain means, variances and covariances. *Journal of Finance*, 29(2), 512–522.
- Bawa, V. S., Brown, S. J., & Klein, R. W. (1979). Optimal portfolio choice under uncertainty: A Bayesian approach. In: V.S. Bawa, S.J. Brown, R.W. Klein (Eds.), *Estimation risk and optimal portfolio choice*.
- Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). *Robust optimization*. Princeton: Princeton University Press.
- Ben-Tal, A., & Nemirovski, A. (1998). Robust convex optimization. *Mathematics of Operations Research*, 23(4), 769–805.
- Ben-Tal, A., & Nemirovski, A. (1999). Robust solutions of uncertain linear programs. *Operations Research Letters*, 25, 1–13.
- Ben-Tal, A., & Nemirovski, A. (2002). Robust optimization—methodology and applications. *Mathematical Programming*, 92(3), 453–480.
- Bertsimas, D., Brown, D. B., & Caramanis, C. (2011). Theory and applications of robust optimization. *SIAM Review*, 53(3), 464–501.
- Best, M. J., & Grauer, R. R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results. *The Review of Financial Studies*, 4(2), 315–342.
- Ceria, S., & Stubbs, R. A. (2006). Incorporating estimation errors into portfolio selection: Robust portfolio construction. *Journal of Asset Management*, 7(2), 109–127.
- Chopra, V. K., & Ziemba, W. T. (1993). The effect of errors in means, variances and covariances on optimal portfolio choice. *Journal of Portfolio Management*, 19(2), 6–11.
- DeMiguel, V., & Nogales, F. (2009). Portfolio selection with robust estimation. *Operations Research*, 57(3), 560–577.
- Dupacova, J. (1999). Portfolio optimization via stochastic programming: Methods of output analysis. *Mathematical Methods of Operations Research*, 50(2), 245–270.
- Durier, Roland (1988). Weighting factor results in vector optimization. *Journal of Optimization Theory and Applications*, 58(3), 411–430.
- Efron, B., & Morris, C. (1975). Data analysis using Stein's estimator and its generalizations. *Journal of the American Statistical Association*, 70, 311–319.
- Eichfelder, Gabriele (2008). *Adaptive scalarization methods in multiobjective optimization*. Berlin: Springer.
- El Ghaoui, L., & Lebret, H. (1997). Robust solutions to least-square problem with uncertain data matrices. *SIAM Journal of Matrix Analysis and Application*, 18, 1035–1064.
- El Ghaoui, L., Oustry, F., & Lebret, H. (1998). Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9, 33–52.
- Erdogan, E., & Iyengar, G. (2006). Ambiguous chance constrained problems and robust optimization. *Mathematical Programming, Series B*, 107(1), 37–61.
- Fang, K.-T., Kotz, S., & Ng, K.-W. (1990). *Symmetric multivariate and related distributions*. London: Chapman and Hall.
- Fang, K.-T., & Zhang, Y.-T. (1990). *Generalized multivariate analysis*. Beijing: Science Press.
- Fliege, Jörg (2004). Gap-free computation of Pareto-points by quadratic scalarizations. *Mathematical Methods of Operations Research*, 59, 69–89.
- Fliege, Jörg, Grana Drummond, L. M., & Svaiter, Benar Fux (2009). Newton's method for multiobjective optimization. *SIAM Journal on Optimization*, 20, 602–626.
- Fliege, Jörg, & Svaiter, Benar Fux (2000). Steepest descent methods for multicriteria optimization. *Mathematical Methods of Operations Research*, 51, 479–494.
- Fliege, Jörg, & Vicente, Luis (2006). Multicriteria approach to bilevel optimization. *Journal of Optimization Theory and Applications*, 131, 209–225.
- Fliege, Jörg, & Xu, Huifu (2011). Stochastic multiobjective optimization: Sample average approximation and applications. *Journal of Optimization Theory and Applications*, 151, 135–162.
- Geoffrion, Arthur M. (1968). Proper efficiency and the theory of vector maximization. *Journal of Optimization Theory and Applications*, 22, 618–630.
- Goldfarb, D., & Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1), 1–38.
- Göpfert Alfred & Nehse Reinhard. (1990). *Vektoroptimierung—Theorie, Verfahren und Anwendungen*, Mathematisch-Naturwissenschaftliche Bibliothek (Vol. 74). Leipzig: BSB B.G. Teubner Verlagsgesellschaft.
- Ingersoll, J.-E. (1987). *Theory of financial decision making*. Maryland: Rowman & Littlefield.
- Jobson, J. D., & Korkie, B. (1981). Estimation for Markowitz efficient portfolios. *Journal of the American Statistical Association*, 75, 544–554.
- Jobson, J. D., & Korkie, B. (1981). Putting Markowitz theory to work. *Journal of Portfolio Management*, 7(4), 70–74.
- Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21(3), 279–292.
- Jorion, P. (1992). Portfolio optimization in practice. *Financial Analysts Journal*, 48(1), 68–74.
- Kan, R., & Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, 42(3), 621–656.

- Kempf, A., & Memmel, C. (2006). Estimating the global minimum variance portfolio. *Schmalenbach Business Review*, 58(4), 332–348.
- Kuhn, H.W. & Tucker, A.W. (1951). Nonlinear programming. In J. Neyman, editor, *Proceedings of the second berkeley symposium on mathematical statistics and probability* (pp. 481–492). University of California Press; Berkeley.
- Lutgens, F. (2004). *Robust portfolio optimization*. PhD Thesis, Maastricht University.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1), 77–91.
- Markowitz, H. M. (1959). *Portfolio selection—efficient diversification of investments*. New York: John Wiley & Sons.
- Meucci, A. (2005). *Risk and asset allocation*. Berlin: Springer.
- Michaud, R. O. (1998). *Efficient asset management*. Boston: Harvard Business School Press.
- Perret-Gentil, C. & Victoria-Feser, M.-P. (2004). *Robust mean–variance portfolio selection*. FAME Research Paper No. 140.
- Sawaragi, Yoshikazu, Nakayama, Hirotaka, & Tanino, Tetsuzo (1985). *Theory of multiobjective optimization. Mathematics in science and engineering* (Vol. 176). Academic Press.
- Schöttle, K., & Werner, R. (2006). Towards reliable efficient frontiers. *Journal of Asset Management*, 7(2), 128–141.
- Schöttle, K., & Werner, R. (2009). Robustness properties of mean–variance portfolios. *Optimization*, 58(6), 641–663.
- Schöttle, K., Werner, R., & Zagst, R. (2010). Comparison and robustification of Bayes and Black–Litterman models. *Mathematical Methods of Operations Research*, 71(3), 453–476.
- Tütüncü, R. H., & Koenig, M. (2004). Robust asset allocation. *Annals of Operations Research*, 132(1–4), 157–187.
- Werner, R. (2008). Cascading: an adjusted exchange method for robust conic programming. *Central European Journal of Operations Research*, 16(2), 179–189.
- Zopounidis, C., & Doumpos, M. (1999). Multi-criteria decision aid in financial decision making: methodologies and literature review. *Journal of Multi-Criteria Decision Analysis*, 11(4–5), 167–186.
- Zopounidis, C., & Doumpos, M. (1999). A multicriteria decision aid methodology for sorting decision problems: The case of financial distress. *Computational Economics*, 14(3), 197–218.