Stochastic Multi-Objective Optimization: a Survey on Non-Scalarizing Methods *

Walter J. Gutjahr †
Department of Statistics and Operations Research
University of Vienna, Austria

Alois Pichler ‡
Norwegian University of Science and Technology, Norway

Abstract. Currently, stochastic optimization on the one hand and multi-objective optimization on the other hand are rich and well-established special fields of Operations Research. Much less developed, however, is their intersection: the analysis of decision problems involving multiple objectives and stochastically represented uncertainty simultaneously. This is amazing, since in economic and managerial applications, the features of multiple decision criteria and uncertainty are very frequently co-occurring. Part of the existing quantitative approaches to deal with problems of this class apply scalarization techniques in order to reduce a given stochastic multi-objective problem to a stochastic single-objective one. The present article gives an overview over a second strand of the recent literature, namely methods that preserve the multi-objective nature of the problem during the computational analysis. We survey publications assuming a risk-neutral decision maker, but also articles addressing the situation where the decision maker is risk-averse. In the second case, modern risk measures play a prominent role, and generalizations of stochastic orders from the univariate to the multivariate case have recently turned out as a promising methodological tool. Modeling questions as well as issues of computational solution are discussed.

Keywords: Stochastic optimization, multi-objective optimization, Pareto optimality, risk measures, multivariate stochastic dominance.

1 Introduction

Among the most influential papers in optimization is George B. Dantzig's article *Linear Programming Under Uncertainty* [16]. This short paper cites contributions from Harry

*Invited Survey Article for *Annals of Operations Research: Special Volume*
†Corresponding author. Universitaetsstrasse 5/9, A-1010 Vienna, Austria. Tel.: +43 1 4277 38632. E-mail: walter.gutjahr@univie.ac.at.
‡The author gratefully acknowledges support of the Norwegian grant 207690/E20.

Markowitz and Richard Bellman, who became famous in different scientific disciplines later. But first and foremost, the paper introduces uncertainty represented by stochastic models into the just created field of Linear Programming. Stochastic considerations thus were at the very beginning of the entire discipline of optimization, and of linear optimization in particular.

About at the same time as optimization under uncertainty based on stochastic models, another important branch of computational optimization emerged: *multi-objective* optimization. Already several decades before, the works by the economists Francis Edgeworth and Vilfredo Pareto had laid the formal foundations for the analysis of (what was called later) multi-objective optimization problems, but it was in the fifties of the 20th century when researchers started to design computer-based algorithms to provide decision makers with concrete solutions to such problems. Koopmans [47], for example, analyzed resource allocation problems based on the notion of "efficient vectors", and Saaty and Gass [65] developed algorithms for the determination of efficient solutions by a parametric weighted-sum approach. (The reader is referred to [46] for a survey on the history of multi-criteria decision analysis).

Stochastic optimization and multi-objective optimization saw a rapid, impressive and extremely fruitful development in the decades since then. Surprisingly enough, however, these two areas evolved to a good part separately from each other, and even today, their intersection is addressed by only a comparably small fraction of publications, although real-life decision making frequently encompasses *both* uncertainty and multiple objectives. The last has been observed in a large range of application fields such as finance [12, 64], energy [9], transportation logistics [38], facility location [27], manufacturing and production planning [40, 70], supply chain management [2, 3], humanitarian logistics [72], telecommunication, health care management [4, 7] or project management [35]. For numerous problems in these and other fields, it is desirable to have optimization models and computational solution techniques that deal with the multi-objective nature *and* with the stochastic features of the problem *simultaneously*.

This survey article addresses methods for modeling and solving stochastic multi-objective optimization problems. Contrary to the focus of a recent related overview by Ben Abdelaziz [11], we concentrate here on methods and models that do *not* (or at least not primarily) use some implicit or explicit form of *scalarization* to reduce the given multi-objective stochastic problem to a single-objective – though still stochastic – variant.[1] Rather than that, the approaches surveyed here sustain the multi-objective nature of the problem during its computational analysis. In the most typical approaches of this kind, a final choice on the "weights" assigned to the objectives (which is of course necessary to come to a decision) is done by the decision maker only at a late stage and can already be based on the insights obtained from the computational analysis. Our survey is not planned to be exhaustive. Instead, we have tried to select a subset of (mostly) methodologically oriented articles giving a flavor of diverse possible ways how to deal with stochastic multiobjective problems.

The organization of the paper is as follows: Section 2 introduces stochastic optimization problems, multi-objective optimization problems, and combined stochastic multi-

[1] Expressed in the terms used in [11], this means that we focus on the "multiobjective method" instead of the "stochastic method".

objective problems. Section 3 addresses methods suitable for applications where the underlying implicit utility function of the decision maker can be assumed as linear; this corresponds to the case of risk-neutral decision making. In Section 4, the more intricate situation of risk-averse decision making is addressed. Section 5, finally, contains some concluding remarks.

2 Problem Description and Background

2.1 Stochastic Optimization Problems

The simplest stochastic optimization problems (SOPs) employ the expected value in the objective, and a decision is only possible *prior* to the observation of some random event influencing the outcome. In financial optimization for example, the investor decides about her or his asset allocation or portfolio decomposition, and at the end of the holding period (s)he will observe the random outcome of this portfolio. This random outcome is represented mathematically by writing

$$f(x, Y),$$

where x is the decision (in our example, the chosen portfolio), and Y is the random variable observed (in our example, the vector of the returns of all assets). For asset allocation, e.g., the investor would pick the linear function $f(x, Y) = x^\top Y$ representing the final return.

In mathematical terms, the expression $f(x, Y)$ is shorthand for the real-valued random variable assigning to each so-called *scenario* $\omega \in \Omega$ the value $f(x, Y(\omega))$; therein, Ω is the sample space[2]. Each scenario ω determines a realization of the random variable $Y(\omega)$ and therefore also of the random variable $f(x, Y(\omega))$.

Since $Y = Y(\omega)$ is completely determined by the scenario ω, we will simply write $f(x, \omega)$ instead of $f(x, Y(\omega))$ from now on.

The easiest stochastic optimization problem, expressing maximization of expected return, then reads

$$\begin{array}{ll} \max & \mathbb{E}\,[f(x,\omega)] \\ \text{s.t.} & x \in \mathcal{X}, \end{array} \tag{1}$$

where $\mathbb{E}$ is the expectation operator and $\mathcal{X}$ the set of feasible solutions (decisions); $\mathcal{X}$ is called the *solution space* or *decision space*.

In the indicated financial optimization context, formulation (1) does not yet lead to meaningful solutions, because it ignores *risk*. To quantify the risk associated with a decision, a convex risk measure ρ is frequently employed (we use here the axioms as stated in [26], although they are adapted by many authors for their particular problem). [3] Then

[2]The sample space Ω is the first component of the probability space $(\Omega, \mathcal{A}, P)$. Note that in stochastic optimization, the probability model given by $(\Omega, \mathcal{A}, P)$ is always assumed as known, and therefore in particular also the distributions of the considered random variables.

[3]The axioms proposed, in the convex setting, are essentially

(i) Monotonicity: $\rho(X) \le \rho(Y)$ provided that $X \ge Y$ almost everywhere, the problem to maximize the expected return, now under risk constraints, can be stated as

$$\begin{array}{ll} \max & \mathbb{E}\left[f(x,\omega)\right] \\ \text{s.t.} & \rho\left(f(x,\omega)\right) \le r,\ x \in \mathcal{X}. \end{array} \tag{2}$$

The decision x is here only allowed provided that the risk of the associated investment, $\rho\left(f(x,\omega)\right)$, does not exceed a given threshold r. Depending on r, the problem is possibly infeasible.

On the other hand, the investor could just as well be interested in obtaining at least an expected return of μ, say, and at the same time intend to minimize the associated risk:

$$\begin{array}{ll} \min & \rho\left(f(x,\omega)\right) \\ \text{s.t.} & \mathbb{E}\left[f(x,\omega)\right] \ge \mu,\ x \in \mathcal{X}. \end{array}$$

As above, depending on μ, this problem is possibly infeasible, and if it is feasible, its solution may differ from that of (2).

From an integrated risk management's perspective (cf. [61]) the problem is often stated in a yet different form, namely in a *scalarized* form as

$$\begin{array}{ll} \max & (1-\gamma)\cdot\mathbb{E}\left[f(x,\omega)\right] - \gamma\cdot\rho\left(f(x,\omega)\right) \\ \text{s.t.} & x \in \mathcal{X}, \end{array} \tag{3}$$

where γ with $0 \le \gamma \le 1$ is a parameter governing the degree to which the investor avoids risk.

Finally, we could also cast the problem in a *bi-objective* formulation by writing

$$\begin{array}{ll} \max & \left(\mathbb{E}\left[f(x,\omega)\right],\ -\rho\left(f(x,\omega)\right)\right) \\ \text{s.t.} & x \in \mathcal{X}. \end{array} \tag{4}$$

This formulation contains a *vector* optimization in its objective, and we shall discuss its possible meaning in Subsection 2.2. For the moment let us confine ourselves to two remarks.

(i) Although (4) is already a multi-objective problem, its two objectives (expected return and risk) have been derived from a *single* quantity of interest, namely (financial) return. In this article we shall go beyond this case by proceeding to the more complex situation where two or more *inherently different* quantities play a role as objectives in the optimization. Some examples include

- cost and health effect in health-care management,
- immediate financial return and strategic positioning in project management,
- delivery cost and customer satisfaction in transportation logistics,

---

(ii) Convexity: $\rho((1-\lambda)Y_0 + \lambda Y_1) \le (1-\lambda)\rho(Y_0) + \lambda\rho(Y_1)$,

(iii) Positive Homogeneity: $\rho(\lambda Y) = \lambda\rho(Y)$ for all $\lambda > 0$ and

(iv) Translation Invariance: $\rho(Y + c) = \rho(Y) - c$.

- cost and reliability in safety-critical technologies, or
- return, market share and emission reduction in production economics.

It is clear, on the other hand, that models (and solution techniques) for this more complex case will be related to those for the case outlined above, because the second case extends the first.

(ii) Formulations of the type (3) *scalarize* the multi-objective problem by combining the objectives to a weighted mean or some other weighted aggregation with pre-specified weights. This issue will be discussed in more detail in Subsection 2.2.

Several other types of formulations for SOPs are possible. Among them, *chance constraints* (cf. Subsection 2.3) have to be mentioned. Moreover, *multistage* formulations are common. Therein, a first-stage decision x_0 is made prior to the observation of any random event. After a partial observation of the random outcome, a subsequent decision x_1 is made, etc. Only before or after the last decision, the observation of the random parameters is revealed completely. We shall shortly address models of this kind in Subsections 3.2.5 and 4.1.4.

A challenging problem of outstanding practical relevance involves *stochastic orders*. As it will be explained in detail in Section 4, a stochastic order $\succeq$ is a dominance relation (partial order) between random outcomes with known distributions, usually (but not necessarily) defined based on expected utilities. In the literature stochastic orders are typically applied in constraints (cf. [18]). Let us illustrate this concept again within the context of financial portfolio optimization by assuming that a fund manager intends to outperform a certain benchmark, the Dow Jones Industrial Average, say. Then (s)he can consider the following problem:

$$\begin{array}{ll} \max & \mathbb{E}\,[f(x,\omega)] \\ \text{s.t.} & f(x,\omega) \succeq \text{Dow Jones.} \end{array} \tag{5}$$

In this formulation, only decisions that are at least as favorable as the Dow Jones (also with respect to risk!) are allowed, and within this set, the expected return is maximized.

The financial engineering examples above build on *continuous* sets of solutions (decisions). In several application areas, problems with a *discrete* feasible set are more typical. Then, one faces a *stochastic integer problem*. A simple example is the Probabilistic Traveling Salesperson Problem (PTSP), where a given set of customers has to be arranged in a fixed order, described by a list x, on the assumption that on a considered day, each customer requires a visit only with a certain probability. An actual tour is constructed from x by visiting the customers in the order given by the list, simply skipping those customers who do not require a visit this day, and returning then to the depot. The objective is to determine the permutation x minimizing the expected length of the actual tour. Other examples are stochastic scheduling problems or 0-1 project portfolio selection problems. Often, also stochastic *mixed-integer* problems occur.

2.2 Multi-Objective Optimization Problems

The generic form of a multi-objective optimization problem (MOOP) is

$$\max\ (f_1(x), \ldots, f_m(x)) \quad \text{s.t.} \quad x \in \mathcal{X}, \tag{6}$$

where $\mathcal{X}$ is again the solution space, and $f_j : \mathcal{X} \to \mathbb{R}$ $(j = 1, \ldots, m)$ are the m objective functions $(m \geq 2)$. From (6), it is not yet clear what is meant by maximizing the vector-valued function $f = (f_1, \ldots, f_m)$. There are several alternatives between which the decision maker (DM) can choose, see, e.g., [21, 22]. Leaving lexicographic optimization aside, which does not give each objective the same priority but considers a rank list between objectives instead, most of the mentioned alternatives fall into one of two classes, namely scalarizing methods on the one hand and Pareto-type methods on the other hand.

Scalarizing methods reduce the multi-objective problem to a single-objective problem by aggregating the objectives in some way or other, e.g., by a weighted average or by a weighted Chebyshev distance. Although these methods considerably facilitate the computational solution of the MOOP, they have the disadvantage that for their application, the decision maker is forced to define the particular way of aggregation (including the weights assigned to the objectives) in advance. Basically, this means that (s)he has to specify her or his utility function in quantitative terms before the computational analysis can start. Very frequently, decision makers are not willing or not able to do that. For this reason, we leave scalarizing methods outside the scope of this survey and focus on methods that do *not* require the decision makers to assign weights to their objectives as a prerequisite for the computational decision analysis.

The standard (non-scalarizing) concept of multi-objective optimization is that of Pareto optimality (cf. [15]):

Definition 1. A solution $x \in \mathcal{X}$ is called *efficient* with respect to objectives $f_1, \ldots, f_m$ if it is not dominated by any feasible $y \in \mathcal{X}$ $(y \neq x)$, that is to say if there is no $y \in \mathcal{X}$ such that (i) $f_j(y) \geq f_j(x)$ for all $j = 1, \ldots, m$, and (ii) $f_j(y) > f_j(x)$ for at least one j. The set of efficient solutions is called the *efficient set.* The image point $(f_1(x), \ldots, f_m(x))$ of an efficient solution x is called *Pareto-optimal*, and the set of Pareto-optimal points, i.e., the image of the efficient set in the objective space $\mathbb{R}^m$ under $f = (f_1, \ldots, f_m)$, is called the *Pareto frontier.*

By providing the decision maker with the efficient set, the computational analysis enables her or him to choose among a considerably restricted set of final solution candidates instead of the entire solution space $\mathcal{X}$. This can greatly reduce the cognitive burden. Let us mention that it is often sufficient to have only one representative from the pre-image of each Pareto-optimal point z in the solution set, which reduces the efficient set to a candidate set that may be equally suitable (and easier to oversee) from the viewpoint of the DM.

Pareto analysis is well-founded in utility theory by the following observation, for which a proof is easy: Let a vector $(f_1, \ldots, f_m)$ of objective functions on $\mathcal{X}$ be given. If solution x^* maximizes $u(f_1(x), \ldots, f_m(x))$ on $\mathcal{X}$ for a (strictly) increasing utility function $u : \mathbb{R}^m \to \mathbb{R}$, then x^* must be efficient. Conversely, if x^* is efficient, then there exists a nondecreasing utility function $u : \mathbb{R}^m \to \mathbb{R}$ such that x^* maximizes $u(f_1(x), \ldots, f_m(x))$ on $\mathcal{X}$.

2.3 Stochastic Multi-Objective Optimization Problems

Let us now turn to the combined situation of a stochastic multi-objective optimization problem (SMOOP), i.e., an optimization problem that is both an SOP and a MOOP. In particular, we consider problems of the form

$$\max \, (f_1(x,\omega), \ldots, f_m(x,\omega)) \quad \text{s.t.} \quad x \in \mathcal{X}, \tag{7}$$

where ω and $\mathcal{X}$ are as in Subsection 2.1, and the real-valued random variables $f_j(x,\omega)$ are defined on $\mathcal{X} \times \Omega$ ($j = 1, \ldots, m$). It is assumed that the joint distribution of the random variables $f_j(x,\omega)$ is known.

In this survey we focus on the special case where the original problem only contains *deterministic* feasibility constraints[4], i.e., the definition of $\mathcal{X}$ is assumed not to contain ω. In particular, we do not consider *chance constraints* of the form $Pr\,(\{\omega : x \in \mathcal{X}(\omega)\}) \geq \alpha$. For the treatment of chance constraints within a stochastic multi-objective optimization context, the reader is referred to [11].

It is clear that by (7) the SMOOP is not yet fully defined. It has to be specified in which sense the vector maximization is to be interpreted, and what is done to remove the ambiguity of the values of the objectives caused by the random influence ω. An elegant general answer to both questions can be given by means of the concept of *multivariate stochastic dominance*, which will be outlined in Subsection 4.2. For the moment we consider two simple and intuitively appealing alternative ways of specifying the meaning of (7). In the literature (see [10, 13] and the recent survey [11]), these two ways have been called the *multiobjective method* and the *stochastic method*. The two approaches differ by the order in which the given SMOOP is reduced to simpler problems: The multi-objective method starts by reducing the SMOOP to a MOOP, which is then solved by techniques of multi-objective optimization. The stochastic method, on the other hand, starts by reducing the SMOOP to a SOP, to which techniques of stochastic optimization are applied.

(i) The *multi-objective method* defines for each objective function f_j a vector

$$\left(\mathcal{F}_j^{(1)}(f_j(x,\omega)), \ldots, \mathcal{F}_j^{(r_j)}(f_j(x,\omega))\right) \tag{8}$$

of one or more functionals $\mathcal{F}_j^{(s)}$ of the random variable $f_j(x,\cdot)$, which can be the expectation, the variance, a quantile, some risk measure, or some other summary statistics of the random variable, and specifies (7) as the deterministic MOOP

$$\max \, \left(Z_1^{(1)}(x), \ldots, Z_1^{(r_1)}(x), \ldots\ldots, Z_m^{(1)}(x), \ldots, Z_m^{(r_m)}\right) \quad \text{s.t.} \quad x \in \mathcal{X} \tag{9}$$

with $Z_j^{(s)}(x) = \mathcal{F}_j^{(s)}(f_j(x,\omega))$ ($s = 1, \ldots, r_j$, $j = 1, \ldots, m$). The MOOP (9) is solved in the sense of the computation of the set of efficient solutions (see Definition 1).

[4] In Subsections 4.1.3 and 4.2.3, also problem formulations encompassing stochastic dominance constraints will be discussed. It is clear that these constraints cannot be considered anymore as "deterministic". This case will nevertheless be included here, since stochastic dominance constraints are not so much externally defined conditions restricting the feasible set, but rather constitute an alternative way to express the preferences of a decision maker. We admit that the distinction is not sharp.

The multi-objective approach can be considered as a generalization of traditional *mean-risk models* from the case $m = 1$ of only one objective to that of several objectives.

Because the information on the entire joint distribution of the objectives is reduced to the consideration of a finite set of summary statistics, the multi-objective method may over-simplify the given SMOOP. In particular, the dependency between the random objectives is not taken into account (cf. [13]). This issue will be discussed below in more detail.

(ii) The *stochastic method* combines the random objectives $f_1(x, \omega), \ldots, f_m(x, \omega)$ by some aggregation function $u : \mathbb{R}^m \to \mathbb{R}$ and specifies (7) as the single-objective SOP

$$\max\ u\left(f_1(x, \omega), \ldots, f_m(x, \omega)\right) \quad \text{s.t.} \quad x \in \mathcal{X}. \tag{10}$$

The stochastic method takes account of the dependency between the stochastic objectives, but it has the obvious disadvantage that the aggregation function u, which should represent the utility function of the decision maker to make the optimization model meaningful, has to be known before the computational analysis can start. Since we restrict ourselves here to non-scalarizing methods, we shall not consider the stochastic method in this survey.

3 Risk-Neutral Decision Making

3.1 Expected Value Efficient Solutions

In this section we investigate the simplest special variant of the multi-objective method, namely that where as functionals $\mathcal{F}_j^{(s)}(f_j(x, \omega))$ only expectations $\mathbb{E}[f_j(x, \omega)]$ are taken. This leads to the so-called *expected value efficient solutions* (cf. [11]), i.e., the efficient solutions of the m-objective problem

$$\max\ (\mathbb{E}[f_1(x, \omega)], \ldots, \mathbb{E}[f_m(x, \omega)]) \quad \text{s.t.} \quad x \in \mathcal{X}. \tag{11}$$

It would be desirable to justify the computation of expected value efficient solutions by using the sound framework of expected utility theory, which we will shortly recall in Subsection 4.1. Unfortunately, however, the remark at the end of Subsection 2.2 concerning the connection between efficient solutions and utility maximizers does not generalize to the stochastic situation. It is, for example, not necessarily the case that for an efficient solution x of (11), there exists a nondecreasing utility function u such that x is a solution of

$$\max\ \mathbb{E}[u(f_1(x, \omega), \ldots, f_m(x, \omega))] \quad \text{s.t.} \quad x \in \mathcal{X}. \tag{12}$$

This is a simple consequence of the fact that in general, the expectation operator cannot be interchanged with the utility function u in (12). Only if interchanging is possible, the observation at the end of Subsection 2.2 extends to the stochastic situation.

It is obviously allowed to switch $\mathbb{E}$ and u if the utility function u is *linear*. Let

$$u\left(f_1, \ldots, f_m\right) = \sum_{j=1}^{m} c_j f_j. \tag{13}$$

with $c_j > 0$ for all j. Such a utility reflects the attitude of a *risk-neutral* decision maker. The relation between risk-neutrality and linear utility functions is well-known; let us illustrate it by considering so-called *mean-preserving spreads* of the objective function values.[5] A distribution P' on $\mathbb{R}$ is a mean-preserving spread of a distribution P if for some X' distributed according to P', the random variable X' can be represented as $X' = X + Z$ with X distributed according to P and $\mathbb{E}(Z \mid X = x) = 0$ for all x. A decision maker is regarded as risk-neutral if (s)he is indifferent with respect to any mean-preserving spread of her/his gain(s). It is obvious that the expected value of the utility u given by (13) is insensitive with respect to mean-preserving spreads of the random variables $f_j = f_j(x, \omega)$.

By the consideration above, each solution x^* maximizing an expected utility of the form (13) must occur in the efficient set of (11).

The range of utility functions for which expectation and utility are allowed to be switched can still be extended by including the possibility that some of the objective functions are actually deterministic: Let $f_1(x), \ldots, f_k(x)$ and $f_{k+1}(x, \omega), \ldots, f_m(x, \omega)$ be the deterministic and the (properly) stochastic objective functions, respectively. Then it is easy to see (cf. [32]) that also for utilities of the form

$$u\,(f_1, \ldots, f_m) = \varphi(f_1, \ldots, f_k) + \sum_{j=k+1}^{m} \psi_j(f_1, \ldots, f_k)\, f_j, \tag{14}$$

with increasing functions φ and ψ_j $(j = k+1, \ldots, m)$, each expected utility maximizer is an efficient solution of (11). Evidently, utilities of this form reflect a risk-neutral attitude as well.

A mix of deterministic and stochastic objective functions does not only occur frequently in applications, it is also interesting from the viewpoint of computational solution because contrary to the purely stochastic case with linear utility (13), it can lead to so-called *unsupported* solutions. In multi-objective optimization, an efficient solution is called *supported* if it can be obtained by optimizing a suitable weighted sum of objectives, and unsupported otherwise. For example, in the case $u(f_1, f_2) = \varphi(f_1) + f_2$ for deterministic f_1, unknown increasing φ and stochastic f_2, the expected utility maximizers are the expected value efficient solutions, but some of them (including desirable "compromise solutions") can be unsupported. In the microeconomics literature, utility functions of the form $u(f_1, f_2) = \varphi(f_1) + f_2$ are called *quasilinear*.

[5] An alternative way to judge risk-neutrality is to apply the Arrow-Pratt risk aversion measure defined by $-u''(\cdot)/u'(\cdot)$ to each random variable f_j. Evidently, the values vanish for the utility u given by (13). Using a more general observation (Theorem 1 in [31]), it can be shown that on the additional constraint that u is nondecreasing and concave, (13) is the *only* form u can have for a decision maker who is risk-neutral toward all properly stochastic objective functions. This does not hold anymore if the restriction to concave utilities is dropped.

3.2 Computational Issues

3.2.1 Finite Probability Distributions and Scenario Generation

To make a SMOOP computationally tractable it is convenient to have finite probability distributions of the form

$$P = \sum_{i=1}^{N} p_i \, \delta_{\omega_i} \tag{15}$$

available. The symbol δ_{ω_i} denotes the point mass (Dirac measure) at ω_i, and the number $p_i = P(\{\omega_i\})$ is the probability that the value ω_i is attained ($i = 1, \ldots, N$, $\sum_{i=1}^{N} p_i = 1$). In this situation the expected value, as well as many risk measures reduce to finite sums, and the optimization problem under consideration often reduces to a linear program. This is particularly the case if P is an empirical probability measure, i.e., a measure obtained by *random sampling* from a basic distribution. It is clear that in this situation the results of the optimization problem have random character themselves, they depend on the randomly observed *samples*. The respective statistical behavior was studied, e.g., in [58, 69, 68], the statistical behavior of risk measures based on empirical observations was investigated in [62].

Quite generally, if the true probability measure is approximated by a finite measure (15), additional considerations are necessary to quantify the quality of the approximating problem: the final decisions x in the original and in the approximating space should be *close* to each other, and so should the corresponding objectives. This is the subject of results of robustness, ambiguity, or stability of stochastic programs. To pick a single example from the wide literature we refer to [39].

As elements of the sample space Ω, the points $\omega_1, \ldots \omega_N$ in (15) are again called *scenarios* (cf. Subsection 2.1). For huge finite distributions it is often computationally advantageous to reduce the number of scenarios to a reasonable amount. Related techniques are required for finding good finite probability distributions to replace an originally *continuous* distribution. This process is called *scenario generation* in the literature. Various approaches are used for scenario generation in practice. Although most of them are based on heuristics, a rigorous mathematical approach is elaborated in [57] and [59], and for multistage stochastic approximation details are elaborated in [60].

3.2.2 Application to the Stochastic Multi-Objective Case

The conceptually simplest way to (approximately) determine the expected value efficient solutions (11) relies on a fixed set $S = \{\omega_1, \ldots, \omega_N\}$ of scenarios drawn *in advance* to approximate the true probability model by a finite probability model, as described in Subsection 3.2.1. If the original probability distribution is used for sampling, the probabilities p_i are equal, i.e., $p_i = 1/N$ ($i = 1, \ldots, N$). In the multi-objective optimization model, each objective $\mathbb{E}[f_j(x, \omega)]$ is replaced then by the corresponding *sample average estimate*

$$\frac{1}{N} \sum_{\nu=1}^{N} f_j(x, \omega_\nu), \tag{16}$$

which is an unbiased consistent estimator of $\mathbb{E}(f_j(x,\omega))$. This results in a deterministic MOOP which can be solved by a standard solution technique for MOOPs. In the linear case, e.g., an algorithm for linear MOOPs (see [21]) can be used. In the nonlinear case, one can apply the well-known *Epsilon-Constraint Algorithm* to obtain an approximation to the Pareto frontier, and in the discrete finite case, the *Adaptive Epsilon-Constraint Algorithm* [48] can be applied as a general solution technique.

3.2.3 Variable Samples

The techniques mentioned in Subsection 3.2.2 can be computationally demanding. The (Adaptive) Epsilon-Constraint Algorithm, for example, requires the solution of a possibly large series of single-objective optimization problem instances. In the case of large sample size N, each of these instances can be computationally very expensive itself. For an NP-hard integer programming problem, say, the growth of the number of needed variables as a function of the sample size N can result in an exponential growth of the computation time in N. Typically, this makes only a comparably small number N of scenarios computationally feasible.

On the other hand, approximating a complex stochastic model by only few random scenarios is often inappropriate, especially in the case where events with low probability and high impact can occur. If, for example, hardware and/or software of an aircraft is to be optimized, one has to consider many possible random events during a flight. In aviation, some failures occur with a rate of only 10^{-8} per hour, but can have fatal consequences nevertheless. In such situations, a stochastic optimization procedure that is only based on a fixed sample of, say, about hundred scenarios, runs the risk of missing relevant scenarios. (This holds even if the variance of the estimation has been reduced by rare-event simulation methods. For the sake of brevity, we do not address variance reduction techniques in this article.)

In the context of single-objective optimization, the drawback of limited samples has been taken account of by the development of *variable-sample* methods (see [41]) which allow a much larger number of scenarios to be considered during the optimization process. A variable-sample technique for discrete stochastic *multi-objective* problems has been developed under the name *Adaptive Pareto Sampling* (APS) in [29]. It performs an iterative refinement of an approximation to the Pareto frontier by (i) drawing, in each iteration, new samples containing only a moderate number of scenarios, (ii) solving the corresponding subproblems (with the help of a suitable MOO algorithm) to obtain new candidate solutions, and (iii) improving the accuracy of all objective function evaluations by re-sampling (similarly as in the *Stochastic Branch-and-Bound* technique by Norkin et al. [52]). APS can be shown to converge to the exact Pareto frontier under mild conditions [29, 35], and it can be demonstrated [33] that the computation time for solving the *stochastic* MOOP exceeds that for solving the deterministic MOO subproblem only by a factor of order $O(n^3)$, which is modest compared to the required effort for solving the (typically NP-hard) deterministic subproblem.

3.2.4 Metaheuristic Techniques

Especially in the case where a SMOOP involves integrality constraints such that an NP-hard discrete SMOOP can result, it may be necessary to resort to metaheuristic solution techniques (for metaheuristics in stochastic combinatorial optimization, see the survey [30]). Most metaheuristics for SMOOPs are based on diverse variants of evolutionary algorithms. Let us give a few examples. Hughes [45], Teich [71], and Eskandari et al. [24] adapt a multi-objective genetic algorithm, the multi-objective evolutionary algorithms SPEA and the nondominated sorting genetic algorithm NSGA-II, respectively, to the stochastic situation. Ding et al. [20] propose a multi-objective genetic algorithm as well, combined with a simulation procedure. In [28], multi-objective variants of Ant Colony Optimization and of Simulated Annealing are generalized to the stochastic case and compared to each other. Amodeo et al. [3] combine a discrete-event simulation procedure with the multi-objective algorithms SPEA-II and NSGA-II, and compare with a multi-objective Particle Swarm Optimization algorithm. Eskandari and Geiger [23] report on computational experiments with an extension of the approach in [45]. Syberfeldt et al. [70] use a multi-objective evolutionary algorithm supported by an artificial neural network, combined with a simulation routine.

The *quality indicator* technique proposed by Basseur and Zitzler [8] and by Liefhooge et al. [49, 50] reduces the multi-objective problem to a single-objective one, but not by a scalarization of the objective functions, but by shifting the optimization process from the level of single solutions to that of sets of solutions intended as approximations of the efficient set.

Also the APS algorithm mentioned in Subsection 3.2.3 can be reshaped to a SMOO metaheuristic by solving the deterministic MOO subproblem with the help of a MOO metaheuristic instead of an exact MOO solution algorithm.

3.2.5 Two-Stage Stochastic Multi-Objective Optimization

In practice, decision making in a multi-objective and stochastic context frequently encompasses a *series* of decisions, between which some of the uncertain parameters become gradually known. This can be cast in the classical framework of *multistage* stochastic optimization, but now complicated by the presence of multiple objective. Already on the modeling level (not to mention the computational solution), such a situation poses difficult challenges. To state things more precisely, let us first consider the single-objective case. We restrict ourselves here to the special case of two decision stages.

The term "two-stage stochastic program" has become accepted in the literature for a stochastic optimization problem in which a part of the objective function is itself a result of an optimization problem. Typically, a two-stage stochastic problem models a situation where at a time, when the realizations of certain random variables are not known yet, a first decision has to be made. Later, after their realizations have been observed, a second decision is made which can adapt to the current state. A two-stage stochastic program is given as

$$\begin{array}{ll} \max & f(x) + \mathbb{E}[g(x,\omega)] \\ \text{s.t.} & x \in \mathcal{X}, \end{array} \tag{17}$$

where the function g is called *recourse function*. The recourse function is evaluated in an additional stage, the second stage. In the linear case, g is defined by the result of the program

$$g(x,\omega) = \begin{array}{ll} \max & q(\omega)^\top y \\ \text{s.t.} & T(\omega)x + Wy = h(\omega) \\ & y \geq 0. \end{array} \tag{18}$$

The decision x in (17) is called *here-and-now* or 1^{st} stage decision, whereas y in (18), which depends on the realization ω, is called *wait-and-see*, or 2^{nd} stage decision.

Special solution techniques have been developed to solve problem (17), most of them based on the *deterministic equivalent*, assuming that the probability space is discrete (see (15)). The deterministic equivalent is a formulation considering simultaneously problem (17) together with N versions of problem (18), one for each scenario ω_i. The resulting problem, which is typically large scale, has a particular form (L-shape), which is exploited by dedicated solution algorithms (such as the L-shaped method, a variant of the Bender's decomposition adapted to this particular problem). It is evident that many optimization problems can be formulated as two-stage stochastic programs and indeed, a large variety of applications is known in the Operations Research literature.

In spite of its importance, the *multi-objective* case is treated in only few publications. The majority of them (for examples, see [27, 14, 72]) assumes a bi-objective situation where one of the two objectives only depends on the 1^{st} stage decision, whereas the other objective depends on the decisions in both stages. This facilitates the analysis, because within a solution framework of, say, the (Adaptive) Epsilon-Constraint Algorithm, constraints can be generated with respect to the first objective, which produces a series of ordinary (single-objective) two-stage stochastic programs to be solved as auxiliary subproblems. For the case of more than one objective depending on both the 1^{st} stage and the 2^{nd} stage decision, the mentioned technique is not applicable anymore. Papers addressing this more intricate case are [1, 63].

4 Risk-Averse Decision Making

Whereas the preferences of a risk-neutral decision maker are represented by linear utility functions, a standard approach to a formal representation of risk aversion is the assumption of a nonlinear utility function. If the dependence of the utility on the outcome of some payoff or objective function value is described by a concave function, then an expected-utility maximizer is no longer indifferent with respect to a mean-preserving spread of her or his payoff, but tends to prefer random distributions with small variation over such with large variation, given that the expected values are the same. In the nonlinear case, contrary to the case of linear utility functions, the formal treatment of the expected-utility optimization problem becomes nontrivial already on a conceptual level, not to mention the level of computational solution. In particular, in a multi-objective situation, the *multi-objective method* shows distinct limitations in the presence of nonlinear utilities, as has been pointed out by other authors (see, e.g., Caballero et al. [13] or Ben Abdelaziz [11]).

A different concept, which has turned out to be very helpful when dealing with SMOOPs under nonlinear utilities, is that of *multivariate stochastic dominance*. Univariate stochastic dominance concepts and, related to them, so-called *mean-risk models* have a long tradition in the treatment of single-objective stochastic optimization problems under risk aversion. Recently, a growing number of articles studies the generalization of these techniques to the multivariate case, i.e., the case of multi-objective stochastic optimization.

We shall start with a short survey on univariate stochastic dominance and mean-risk models in Subsection 4.1, and turn then to the multivariate case in Subsection 4.2.

4.1 Univariate Stochastic Orders

A fundamental concept in economics is that of expected utility maximization: a random outcome X is preferred over another random outcome Y, if the corresponding expected utility values satisfy

$$\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)] \tag{19}$$

for an *a priori* chosen utility function $u : \mathbb{R} \to \mathbb{R}$. In the (frequent) case where the decision maker does not specify a concrete utility function, it is necessary to express the preference of X over Y not just by a single utility u as in (19), but with respect to an entire class of utility functions:

$$X \succeq_{\mathscr{U}} Y \;\Leftrightarrow\; \mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)] \text{ for all } u \in \mathscr{U} \text{ such that the expectations exist.}$$

The relation $\succeq_{\mathscr{U}}$ introduces a partial order – a so-called stochastic order relation – on a set of random variables on a common probability space. (The assumption that the considered expectations exist will also be needed in the following; for the sake of brevity, we will not mention it explicitly anymore. Alternatively, its validity can be achieved by imposing additional regularity constraints on the elements of $\mathscr{U}$.)

Many classes of stochastic order relations have been considered in the literature. Important examples include

- $\mathscr{U}_{(1)} = \{u : \mathbb{R} \to \mathbb{R} : u \text{ nondecreasing}\}$ and
- $\mathscr{U}_{(2)} = \{u : \mathbb{R} \to \mathbb{R} : u \text{ nondecreasing and concave}\}$.

Each utility function of a rational decision maker can be supposed to belong to the class $\mathscr{U}_{(1)}$, and each utility function of a rational and not risk-seeking decision maker can be supposed to belong even to the smaller class $\mathscr{U}_{(2)}$, considering that risk aversion corresponds to a concave utility function.[6] The stochastic orders $\succeq_{(1)}=\succeq_{\mathscr{U}_{(1)}}$ and $\succeq_{(2)}=\succeq_{\mathscr{U}_{(2)}}$ are called the *first-order stochastic dominance relation* and the *second-order stochastic dominance relation*, respectively.

For a comprehensive analysis of stochastic order relations we refer to the books [51] and [67].

[6] Let us also mention the s-*convex* class $\mathscr{U}_s$ given by $\mathscr{U}_s = \{u : \mathbb{R} \to \mathbb{R} : (-1)^s u^{(s)} \geq 0\}$, where $u^{(s)}$ is the s-th derivative of u. This class contains the function $x \mapsto x^k$ as well as the function $x \mapsto -x^k$ for every $k < s$. Notably, the relation $X \succeq_{\mathscr{U}_s} Y$ thus ensures that the moments of order less than s coincide: $\mathbb{E}[X^k] = \mathbb{E}[Y^k]$ for all $k < s$. The special case $\mathscr{U}_2$ of the 2-convex class is closely related to the class $\mathscr{U}_{(2)}$ defined above, but note that it has an equality constraint for expected values instead of an inequality constraint.

For a given stochastic order $\succeq$ (e.g., an order $\succeq_{\mathscr{U}}$ derived from some class $\mathscr{U}$), we say that X *stochastically dominates* Y, written $X \succ Y$, if $X \succeq Y$ and $Y \not\succeq X$. This relation can be extended from random vectors to solutions of an optimization problem in a natural way by defining that a solution x stochastically dominates solution y with respect to $\succeq$ and f if $f(x,\omega) \succ f(y,\omega)$.

The preference $\succeq_{\mathscr{U}}$ represents in fact multiple criteria, typically a continuum of criteria.[7] Therefore, one frequently has $X \not\succeq_{\mathscr{U}} Y$ and $X \not\succeq_{\mathscr{U}} Y$. For this reason, we cannot expect to be able to solve the optimization problem

$$\begin{array}{ll} \max & f(x,\omega) \\ \text{s.t.} & x \in \mathcal{X} \end{array} \tag{20}$$

in the sense of finding an $x \in \mathcal{X}$ such that $f(x,\omega) \succeq_{\mathscr{U}} f(y,\omega)$ for all $y \in \mathcal{X}$. However, by a solution to (20), we can understand the set of stochastically nondominated solutions $x \in \mathcal{X}$ with respect to $\succeq$ and f, that is to say, the set of all solutions $x \in \mathcal{X}$ for which there exists no solution $y \in \mathcal{X}$ such that $f(y,\omega) \succ f(x,\omega)$.

It is seen that already in the case of only one objective function $f = f_1$, the given stochastic single-objective problem implicitly splits to a multi-objective problem, usually even with an infinite number of objectives. So-called *mean-risk models* aim at reducing the dimensionality of this multi-objective problems to only two objectives, the expected value and some measure related to the spread of the distribution, which can be (but needs not to be) a risk measure in the sense of [6]. We shall recall some especially important risk measures in Subsection 4.1.1 and will then give an outline of mean-risk models in Subsection 4.1.2.

4.1.1 Risk Measures

An important driver of stochastic optimization was – and is – the problem of asset allocation. This problem already exposes the dilemma of an investor to maximize the profit, but to keep the portfolio's risk, at the same time, as small as possible. When dealing with the asset allocation problem, Harry Markowitz employed the *variance* to quantify the risk. This has been criticized, because the variance treats positive and negative deviations (high profits and losses) in the same way, though high profits are welcome but losses are not. For this reason, other measures have been developed to reflect the asymmetry of upwards and downwards deviations of the objective function from the viewpoint of the DM. Widely accepted nowadays are risk measures as introduced in the pioneering paper [6].

A still frequently used, but theoretically not fully satisfying measure of risk is the *Value-at-Risk* (V@R). We shall shortly recall the V@R, because the theoretically much more attractive Average Value-at-Risk described next builds on the V@R. The Value-at-Risk of the random outcome X at level α, written $\mathsf{V@R}_\alpha(X)$, is simply the negative α-quantile of the random variable X, i.e.,

$$\mathsf{V@R}_\alpha(X) = -F_X^{-1}(\alpha)$$

[7] For example, one criterion for each $u \in \mathscr{U}$, but also other representations of the criteria can be chosen, as it will be seen later. Under some special circumstances, some of these criteria may coincide in a way making the problem finite-dimensional even for infinite $\mathscr{U}$, but in general, this does not hold.

with F_X denoting the distribution function of X. The $\mathsf{V@R}$ does *not* satisfy all axioms proposed in [5].

An outstanding risk measure is the *Average Value-at-Risk*. The Average Value-at-Risk $\mathsf{AV@R}_\alpha(X)$ of the outcome X at level α $(0 < \alpha \leq 1)$ is defined as

$$\mathsf{AV@R}_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha \mathsf{V@R}_p(X)\mathrm{d}p$$

and can be represented equivalently as

$$\mathsf{AV@R}_\alpha(X) = \min_{q\in\mathbb{R}} \left(\frac{1}{\alpha}\mathbb{E}(q-X)_+ - q\right)$$

with $Y_+ = \max(Y, 0)$. The definition can be extended to cover the boundary value $\alpha = 0$. The first equation introduces the Average Value-at-Risk as the average of all Values-at-Risk $\mathsf{V@R}_p(X)$ for levels p below the given threshold α. The second equation is a representation that is particularly useful in computational optimization, it expresses the $\mathsf{AV@R}$ as a minimum of an optimization problem. Analogous representations for other risk measures – if they exist at all – are typically significantly more complicated.

The Average Value-at-Risk is sometimes also called *Conditional Value-at-Risk*, *Expected Shortfall*, *Tail Value-at-Risk*, *Conditional Tail Expectation* (in an insurance context), or recently also *Super-Quantile*. The name "Conditional Value-at-Risk" derives from the representation

$$\mathsf{AV@R}_\alpha(Y) = \mathbb{E}\left[-Y \,|\, Y \leq -\mathsf{V@R}_\alpha(Y)\right]$$

interpreting $\mathsf{AV@R}_\alpha$ as the expected loss in the worst $\alpha \cdot 100$ percent of all possible scenarios. For $\alpha = 1$, we have $\mathsf{AV@R}_1(X) = -\mathbb{E}[X]$, so the expectation is a special case of the $\mathsf{AV@R}$.

The importance of the $\mathsf{AV@R}$ becomes also evident from Kusuoka's theorem which states that *every* general, positively homogeneous and law invariant risk measure ρ can be represented as

$$\rho(Y) = \sup_{\mu\in\mathscr{M}}\int_0^1 \mathsf{AV@R}_\alpha(Y)\mu(\mathrm{d}\alpha),$$

where $\mathscr{M}$ is a set of probability measures on $[0, 1]$.

4.1.2 Mean-Risk Models

Mean-risk models are among the oldest attempts in stochastic optimization to address risk aversion. Markowitz' famous formulation of the portfolio optimization problem is already a mean-risk model, although (as stated above) measuring risk by variance has distinct drawbacks.

The basic idea of a mean-risk model is to replace the problem of computing all stochastically nondominated solutions $x \in \mathcal{X}$ with respect to some stochastic order $\succeq$ and objective function f by the solution of a bi-objective problem of the form (4). Therein, ρ is some measure related to risk, which can be a risk measure in the strict sense of the axioms in [26] or a measure quantifying risk in a wider sense. As it can be seen, in formulation (4), the solution x dominates (is preferred to) the solution y if

$$\mathbb{E}[f(x,\omega)] \geq \mathbb{E}[f(y,\omega)] \text{ and } \rho(f(x,\omega)) \leq \rho(f(y,\omega)),$$

with one of the inequalities being strict.

Except of the already mentioned variance, possible candidates for $\rho(X)$ are (among others):

- the absolute semi-deviation,

$$\bar{\delta}(X) = \mathbb{E}[\max(\mathbb{E}[X] - X, 0)],$$

- the standard semi-deviation,

$$\delta(X) = \left(\mathbb{E}[(\max(\mathbb{E}[X] - X, 0))^2]\right)^{1/2}, \text{ or}$$

- the Average Value-at-Risk,

$$\mathsf{AV@R}_\alpha(X) \quad \text{for some fixed } \alpha \in [0,1].$$

Computing the set of the efficient solutions of the bi-objective problem (4) instead of the set of stochastically nondominated solutions raises the question how the two sets of solutions are related to each other. Considering the dimensionality of the multi-objective problems, it is clear that except in the risk-neutral cases discussed in Section 3, the set of stochastically nondominated solutions will typically be "larger" than the set of efficient solutions. However, can we at least be sure that each efficient solution of the bi-objective problem (4) is stochastically nondominated?

Definition 2. The mean-risk model (4) is called *consistent* with the stochastic order $\succeq$ if for each $x \in \mathcal{X}$ and $y \in \mathcal{X}$,

$$f(x,\omega) \succeq f(y,\omega) \;\Rightarrow\; \mathbb{E}[f(x,\omega)] \geq \mathbb{E}[f(y,\omega)] \text{ and } \rho(f(x,\omega)) \leq \rho(f(y,\omega)). \tag{21}$$

It is easy to see that if (4) is consistent with $\succeq$, then an efficient solution x of (4) cannot be stochastically dominated with respect to $\succeq$ by any $y \in \mathcal{X}$, except for solutions y with the same image point as x under the objectives of the mean-risk model.[8]

It is well-known that Markowitz' mean-*variance* model is *not* consistent with second-order stochastic dominance relation $\succeq_{(2)}$, so it is not perfectly suited as a decision aid for a rational, risk-averse decision maker.

The mean-risk model using the Average Value-at-Risk at some level α *is* consistent with the second order stochastic dominance relation. Indeed, it holds by [55, Theorem 3.2] that

$$X \succeq_{(2)} Y \text{ if and only if } \mathsf{AV@R}_p(X) \leq \mathsf{AV@R}_p(Y) \text{ for all } 0 \leq p \leq 1. \tag{22}$$

[8]The last part of the sentence says that an efficient solution x needs not necessarily to be stochastically nondominated itself, but it can only be stochastically dominated by a solution y with $\mathbb{E}[f(x,\omega)] = \mathbb{E}[f(y,\omega)]$ and $\rho(f(x,\omega)) = \rho(f(y,\omega))$. Therefore, if the pre-image of a Pareto-optimal point of the mean-risk model (4) is finite, it contains at least one stochastically nondominated solution.

From the special cases $p = 1$ and $p = \alpha$, we conclude that (21) is satisfied. Hence the Average Value-at-Risk can be employed in accordance with the second order stochastic dominance relation.

If the absolute semideviation or the standard semideviation is used for measuring risk in a mean-risk model, we do not immediately obtain consistency with $\succeq_{(2)}$. However, as shown in [54], Propositions 7 and 8, consistent mean-risk models are obtained by subtracting the expectation: For $\rho(X) = \bar{\delta}(X) - \mathbb{E}[X]$ as well as for $\rho(X) = \delta(X) - \mathbb{E}[X]$, the mean-risk model (4) is consistent with $\succeq_{(2)}$.

We refer to [56] for further examples of consistent mean-risk models.

4.1.3 Stochastic Dominance Constraints

It has been mentioned above that the determination of all stochastically nondominated solutions of $\max_{x \in \mathcal{X}} f(x, \omega)$ cannot be interpreted anymore as a Pareto analysis problem with a *finite* number of objectives. Rather than that, we may consider it as a Pareto analysis problem with a *continuum* of objective functions. In the case of the order $\succeq_{(2)}$, for example, by (22), all values $\mathsf{AV@R}_\alpha$ for $\alpha \in [0, 1]$ can be considered as objectives. In mean-risk models, *two* objectives are selected. Some articles (as [64]) go beyond that by taking *three* objectives into account, the expected value and two different risk measures. However, the possibilities of increasing the number of objectives are limited: a larger number of objectives increases the computational difficulty and extends the dimensionality of the Pareto frontier, making it less intuitive to interpret.

An alternative is to shift the preferences concerning risk rather to the constraints of the given optimization problem than having it expressed within the objectives. This is done by *stochastic dominance constraints.* Using stochastic dominance constraints, we get problem formulations analogous to (5):

$$\begin{array}{ll} \max & \mathbb{E}\left[f(x, \omega)\right] \\ \text{s.t.} & f(x, \omega) \succeq Y, \end{array} \tag{23}$$

where Y is some reference benchmark chosen by the DM, for example, $Y = f(y, \omega)$ with some known (acceptable) solution $y \in \mathcal{X}$. Only those solutions that are not worse than the benchmark Y with respect to the stochastic order $\succeq$ are admitted to the optimization.

Stochastic dominance constraints have been introduced in [18], and also some mathematical properties of the corresponding optimization problems have been analyzed there. Among other papers, the recent article [25] continues the analysis, adding also computational results.

4.1.4 Two-Stage Problems

In a risk-neutral context, we have outlined and discussed two-stage stochastic optimization problems in Subsection 3.2.5. Let us now turn to the risk-averse case, focusing on the single-objective situation. Risk can be incorporated in problem (17) by different means. Involving risk measures is again an attractive option. We point out a recently developed variant of the classical two-stage stochastic program in which the 1^{st} stage problem is of the following form:

$$\begin{array}{ll} \max & f(x) + \mathbb{E}[g(x,Y)] \\ \text{s.t.} & g(x,Y) \succeq Z,\ x \in \mathcal{X}. \end{array}$$

This formulation has been introduced and discussed in [17]. As it is seen, it uses a stochastic dominance constraint applied to the recourse function. The authors investigate 2^{nd} stage problems of the type (18) and deal with the stochastic order $\succeq_{(2)}$. After presenting two new characterizations of this order, they develop decomposition methods (a multi-cut decomposition method and a method using quantile functions) for the computational solution of the overall problem. Also convergence proofs are provided.

4.2 Multivariate Stochastic Orders

In order to proceed from SOPs to SMOOPs, let us now turn to *multivariate* stochastic orders, i.e., stochastic order relations defined on sets of *vectors* of random variables. There are several concepts generalizing the notion of a stochastic order to the multivariate case (cf. Müller and Stoyan [51], Chapter 3). Below we recall two ways to define a multivariate order based on a given univariate order. Let $\mathbf{X}$ and $\mathbf{Y}$ be two random variables with values in $\mathbb{R}^m$, and let $\mathscr{U}$ be a set of utility functions $u : \mathbb{R} \to \mathbb{R}$. We define the

- *component order*: $\mathbf{X} \succeq_{\mathscr{U}}^{comp} \mathbf{Y}$ if $X_j \succeq_{\mathscr{U}} Y_j$ for each $j = 1, \ldots, m$, and the
- *positive linear order*: $\mathbf{X} \succeq_{\mathscr{U}}^{lin} \mathbf{Y}$ if $\mathbf{a}^\top \mathbf{X} \succeq_{\mathscr{U}} \mathbf{a}^\top \mathbf{Y}$ for all $\mathbf{a} \geq \mathbf{0}$.

Whereas the component order refers to the case where the m components of the random variables can be considered as independent, the positive linear order refers to a situation where components (in a multi-objective optimization context: the outcomes of the objective functions) can be *traded against each other*, but at some *a priori* unknown prices $a_j \geq 0$. After combining the objective functions to a weighted sum with the prices as weights, some (in general nonlinear) single-dimensional utility function is applied to them.

Most frequently, the constructions above are applied to the special case where $\mathscr{U}$ is the set $\mathscr{U}_{(2)}$ of nondecreasing concave utility functions. This gives the *componentwise second-order dominance relation* $\succeq_{(2)}^{comp}$ and the *nondecreasing positive linear concave order* $\succeq_{(2)}^{lin}$, respectively.

Multivariate orders can also be defined based directly on sets $\mathscr{U}^{[m]}$ of multi-dimensional utility functions $u : \mathbb{R}^m \to \mathbb{R}$, setting $\mathbf{X} \succeq_{\mathscr{U}^{[m]}} \mathbf{Y}$ if $\mathbb{E}[u(\mathbf{X})] \geq \mathbb{E}[u(\mathbf{Y})]$ for all $u \in \mathscr{U}^{[m]}$. If we take, for example, for $\mathscr{U}^{[m]}$ the set $\mathscr{U}_{(2)}^{[m]}$ of nondecreasing concave utility functions $u : \mathbb{R}^m \to \mathbb{R}$, we get the multivariate version of the *nondecreasing concave order*, which will be denoted again by $\succeq_{(2)}$. Also the component order and the positive linear order can be expressed in this framework: It is easy to see that if one chooses $\mathscr{U}^{[m]}$ as the set of utility functions $u : \mathbb{R}^m \to \mathbb{R}$ of the form

$$u(\mathbf{X}) = u_1(X_1) + \ldots + u_m(X_m),\ u_j \in \mathscr{U}\ (j = 1, \ldots, m), \tag{24}$$

the component order $\succeq_{\mathscr{U}}^{comp}$ results, and if one chooses $\mathscr{U}^{[m]}$ as the set of utility functions $u : \mathbb{R}^m \to \mathbb{R}$ of the form

$$u(\mathbf{X}) = \bar{u}(\mathbf{a}^\top \mathbf{X}),\ \bar{u} \in \mathscr{U},\ \mathbf{a} \geq \mathbf{0}, \tag{25}$$

the positive linear order $\succeq_{\mathscr{U}}^{lin}$ is obtained.

In [51, 4], the following relations are derived:

$$\mathbf{X} \succeq_{(2)} \mathbf{Y} \;\Rightarrow\; \mathbf{X} \succeq_{(2)}^{lin} \mathbf{Y} \;\Rightarrow\; \mathbf{X} \succeq_{(2)}^{comp} \mathbf{Y}, \tag{26}$$

and it is shown that the inclusions caused by these implications are strict.

Using a multivariate stochastic order $\succeq$, we can provide a general specification of the (under-specified) SMOOP (7). For a given vector $f = (f_1, \ldots, f_m)$ of objective functions, let us extend a given stochastic order relation $\succeq$ from random vectors to solutions in a quite analogous way as we did it in the case $m = 1$: We say that a solution x *stochastically dominates* solution y with respect to $\succeq$ and f iff

$$(f_1(x,\omega), \ldots, f_m(x,\omega)) \succ (f_1(y,\omega), \ldots, f_m(y,\omega)) \tag{27}$$

with $\mathbf{X} \succ \mathbf{Y}$ if $\mathbf{X} \succeq \mathbf{Y}$ and $\mathbf{Y} \not\succeq \mathbf{X}$. A solution $x \in \mathcal{X}$ is called *stochastically nondominated* if there is no solution $y \in \mathcal{X}$ dominating x. By the solution of (7), we understand then the set of stochastically nondominated solutions $x \in \mathcal{X}$ with respect to $\succeq$ and f. [9]

The relation between the solution of (7) in the sense above, using some specific multivariate order $\succeq$, and the solution of (7) by the *multi-objective method*, using some specific set of functionals $\mathcal{F}_j^{(s)}$ (cf. (8)), is again a topic of consistency results.

4.2.1 The Component Order

If $\succeq$ is chosen as the component order $\succeq_{\mathscr{U}}^{comp}$ for some $\mathscr{U}$, then (27) can be decomposed to

$$f_j(x,\omega) \succeq_{\mathscr{U}} f_j(y,\omega) \quad (j = 1, \ldots, m)$$

with at least one dominance being strict: $f_j(x,\omega) \succ_{\mathscr{U}} f_j(y,\omega)$. For the special case where $\mathscr{U}$ consists of linear utility functions, this leads again back to the expected value efficient solution, since then $f_j(x,\omega) \succeq_{\mathscr{U}} f_j(y,\omega)$ exactly if $\mathbb{E}[f_j(x,\omega)] \geq \mathbb{E}[f_j(y,\omega)]$. In the case of nonlinear utilities, say for $\mathscr{U} = \mathscr{U}_{(2)}$, the decomposition still holds, but the expected values have to be extended by risk measures, as in the univariate situation. Whereas the so-called *expected value standard deviation efficient solution* (cf. [13]), which uses the standard deviations as risk measures, is in general not consistent with the component order induced by $\mathscr{U}_{(2)}$ (because it is already not consistent in the boundary case $m = 1$), a consistent vector of summary statistics would be obtained by extending each of the m expectations by, say, an AV@R value.

4.2.2 Partial Risk Neutrality

Frequently, a decision maker is risk-neutral in some of her or his objectives while risk-averse in other objectives. For example, a budget limit may induce risk aversion with respect to costs (as the cost approaches or even exceeds the limit, the utility decreases faster than linearly), whereas with respect to a second (non-monetary) objective, there may be no reason for risk aversion. For a twice differentiable utility function $u : \mathbb{R}^m \to \mathbb{R}$,

[9] Analogously as in the case mentioned in Subsection 2.2, it is often sufficient to have only one representative from the pre-image of each stochastically nondominated random vector in the solution set.

this means that $\partial^2 u \,/\, \partial f_j^2 = 0$ for some j corresponding to the objective toward which the decision maker is risk-neutral. Let us assume risk aversion holds toward objectives $f_1, \ldots, f_k$, and risk neutrality holds toward objectives $f_{k+1}, \ldots, f_m$. Note that even in the case of risk neutrality toward all objectives, the utility function needs not to be linear; an example is the function $u(f_1, f_2) = f_1 \cdot f_2$.

It can be shown [31, 34] that at least in the case where as the basic set $\mathscr{U}^{[m]}$ of utilities, the set $\mathscr{U}_{(2)}^{[m]}$ is taken, a restriction of this set to those utilities for which risk neutrality toward $f_{k+1}, \ldots, f_m$ holds allows it to "split off" the objectives $f_{k+1}, \ldots, f_m$ in the form of expectations, such that risk measures have only to be introduced for the objectives $f_1, \ldots, f_k$. Although this is what one would expect, the result is nontrivial insofar as it hinges on the properties of $\mathscr{U}_{(2)}^{[m]}$: it does not hold anymore for the larger class $\mathscr{U}_{(1)}^{[m]}$ of nondecreasing utility functions and not even for the class of nondecreasing, componentwise concave utility functions.

4.2.3 Stochastic Dominance Constraints in the Multivariate Case

As mentioned above, already in the single-objective case $m = 1$, the determination of all stochastically nondominated solutions of $\max_{x \in \mathcal{X}} f(x, \omega)$ can be considered as an optimization problem with an infinite number of objectives. This holds *a fortiori* for the case $m > 1$. In the last case, even if we restrict ourselves to a finite selection of summary statistics $\mathcal{F}_j^{(s)}$, the number of required statistics can become too large to make the determination of the corresponding Pareto frontier a convenient approach. (Note that already for four objectives, no natural graphical representation of the Pareto frontier exists, and since the number of efficient solutions rapidly grows with the number of objectives, also computational issues may become a problem.)

This difficulty is aggravated if we pass from the component order $\succeq_{(2)}^{comp}$ to the "richer" orders $\succeq_{(2)}^{lin}$ or even $\succeq_{(2)}$, as by (26), proceeding in this way leads to a growth of the number of stochastically nondominated solutions. Thus, formulations adopting stochastic dominance constraints, as explained in Subsection 4.1.3, are especially attractive in the multi-objective case: By proceeding to "richer" stochastic orders, the number of solutions is then not increased but reduced.

The multivariate stochastic-dominance-constrained problem is given as

$$\max \; g(x) \quad \text{s.t.} \quad f(x, \omega) \succeq \mathbf{Y},\; x \in \mathcal{X}, \tag{28}$$

where $\succeq$ is a multivariate stochastic order, and $\mathbf{Y}$ is a reference benchmark. The objective $g(x)$ of (28) can also be an expected value. For instance, one might choose $g(x) = \mathbb{E}[f_j(x, \omega)]$ with some selected objective function f_j.

In [19], Dentcheva and Ruszczyński extend their pioneering work [18] on stochastic dominance constraints to the multivariate case. The objective function $g(x)$ is taken as an expectation $\mathbb{E}[H(x, \omega)]$, where $x \mapsto H(x, \omega)$ is assumed to be concave in x for almost all ω, as well as $x \mapsto f(x, \omega)$. In the constraint, the positive linear order $\succeq_{(2)}^{lin}$ is applied. A generator set for this order is derived. Under a technical condition concerning uniform dominance, it is shown that the original problem allows an equivalent dual formulation, and that no duality gap occurs. The Lagrange multipliers can be identified as elements of the indicated generator set. The paper also turns to the nonconvex case. In this case, at least necessary optimality conditions can be given. This article lays the mathematical fundamentals for several related later papers on optimization through multivariate stochastic dominance constraints.

Homem-de-Mello and Mehrotra [42] confine themselves to the case where both the function $g(x)$ and the function $f(x,\omega)$ in (28) are linear, i.e., $g(x) = d^\top x$ and $f(x,\omega) = A(\omega) \cdot x$ with a random $[m \times n]$ matrix $A(\omega)$ and $x \in \mathcal{X} \subseteq \mathbb{R}^n$. For this case, the authors develop a cutting surface technique for solving (28) computationally. As in [19], the stochastic dominance constraints are based on the positive linear order $\succeq_{(2)}^{lin}$. However, the authors generalize the definition of $\succeq_{(2)}^{lin}$ by a restriction of the weight vector $\mathbf{a}$ to a polyhedron $\mathcal{P}$: The random vector $\mathbf{X}$ is said to dominate the random vector $\mathbf{Y}$ in *polyhedral second order*, if $\mathbf{a}^\top\mathbf{X} \succeq_{(2)} \mathbf{a}^\top\mathbf{Y}$ for all $\mathbf{a} \in \mathcal{P}$. It is seen that the possible "prices" for the outcomes of the objectives can now be restricted by linear constraints. For the purpose of computational solution, the authors use the canonical generator set of the set $\mathscr{U}_{(2)}$ and show (by means of the decomposition theorem for polyhedra and by reformulating the problem as a concave minimization problem with additional variables) that a *finite* number of constraints is sufficient to give an equivalent representation of the (infinitely many) inequalities contained in the stochastic dominance constraint. This allows a re-formulation of (28) (in the present linear context) as a linear program (LP). Duality results are presented, and a cut-generation algorithm for solving the LP in an efficient manner by gradually adding constraints to relaxations is proposed.

An interesting application of the approach developed in [42] to a budget allocation problem in homeland security is given in [43]. The authors use sample average approximation (SAA, cf. Subsection 3.2.2) and show that the model asymptotically converges as the sample size tends to infinity. Moreover, statistical lower and upper bounds are derived.

Haskell et al. [37] work with relaxations of the order $\succeq_{(2)}$ by considering parametrized subfamilies $\mathscr{U}^{[m]} = \{u(\cdot;\xi) \,|\, \xi \in \Xi\}$ of the set $\mathscr{U}_{(2)}^{[m]}$ of nondecreasing concave utilities. Their main contribution consists in applying the idea of *regret minimization* to this situation. They define to each utility u a *benchmark* $y(u) = \max_{x \in \mathcal{X}} \mathbb{E}[u(f(x,\omega))]$, where $f(x,\omega) = (f_1(x,\omega), \dots, f_m(x,\omega))$ with $x \mapsto f(x,\omega)$ being concave for almost all ω, and define the (negative) *regret* as

$$\min_{\xi \in \Xi} \{\mathbb{E}[u(f(x,\omega);\xi)] - y(u(\cdot;\xi))\} \tag{29}$$

which is always a nonpositive value. The regret minimization problem reads then

$$\max_{x \in \mathcal{X}} \min_{\xi \in \Xi} \{\mathbb{E}[u(f(x,\omega);\xi)] - y(u(\cdot;\xi))\}\,. \tag{30}$$

Applying Lagrangian duality, this problem can be re-formulated as a convex semi-infinite programming problem, and strong duality results can be shown. After that, Haskell et al. extend their results to the case where a reference solution is used and show that the regret objective can be interpreted as a penalty function for corresponding stochastic dominance constraints.

In [36], compared to [37], more general sets $\mathscr{U}^{[m]}$ of nondecreasing concave utility functions (still forming compact subsets of $\mathscr{U}_{(2)}^{[m]}$) are considered as a basis for stochastic dominance constraints. The higher degree of generality requires the introduction of certain perturbation terms in the stochastic dominance constraints as a technical mean to make Slater conditions satisfied. If this is done, Lagrange multipliers of the constrained problem turn out as nondecreasing concave functions. Moreover, [36] investigates SAA in the addressed framework and extends the asymptotic consistency results for SAA given in [44] to the more general situation considered here. Finally, it is shown that by a coupling approach using a suitable joint distribution of the random variables $f(x, \omega)$ and the reference random variable $\mathbf{Y}$ in the stochastic dominance constraint (instead of assuming them as independent), computational advantages can be gained.

Noyan and Rudolf [53] consider the problem (28) for special variants of the linear order $\succeq^{lin}$. The variants are derived from an Average Value-at-Risk $\mathsf{AV@R}_\alpha$. As in [42], the weight vectors $\mathbf{a}$ used for the linear combination $\mathbf{a}^\top \mathbf{X}$ are restricted to a polyhedron $\mathcal{P}$. Noyan and Rudolf define the relation of AV@R-preferability by saying that $\mathbf{X}$ is AV@R-preferable[10] to $\mathbf{Y}$ at confidence level α with respect to $\mathcal{P}$, iff

$$\mathsf{AV@R}_\alpha(\mathbf{a}^\top \mathbf{X}) \geq \mathsf{AV@R}_\alpha(\mathbf{a}^\top \mathbf{Y}) \quad \text{for all } \mathbf{a} \in \mathcal{P}. \tag{31}$$

Let us denote this order by $\succeq^{lin,\mathcal{P}}_{\mathsf{AV@R}_\alpha}$. Observe that $\mathbf{a}^\top \mathbf{X} \succeq_{(2)} \mathbf{a}^\top \mathbf{Y}$ implies $\mathsf{AV@R}_\alpha(\mathbf{a}^\top \mathbf{X}) \geq \mathsf{AV@R}_\alpha(\mathbf{a}^\top \mathbf{Y})$ for all α, so AV@R-preferability relaxes the stochastic dominance in polyhedral second order introduced in [42, 43]. On the other hand, the continuum of AV@R-constraints for all confidence levels $\alpha \in [0, 1]$ is equivalent to the constraint with respect to $\succeq_{(2)}$, so by several AV@R-constraints with different α, the latter can be approximated. In their computational framework, the authors use LP formulations for the computation of $\mathsf{AV@R}_\alpha$ in the case of a finite scenario model, and they show that in the last-mentioned case, it is sufficient to consider a finite subset of weight vectors $\mathbf{a} \in \mathcal{P}$. For the LP formulations of the problem (28) with $\succeq$ replaced by $\succeq^{lin,\mathcal{P}}_{\mathsf{AV@R}_\alpha}$ and f and g linear, the authors provide duality results. Moreover, they extend their theory from AV@R to more general coherent risk measures. A cut-generation algorithm for the computational solution of the problem is presented, and it is shown that it terminates after a finite number of iterations. Illustrative examples (one of which is the homeland security budget allocation problem from [43]) show the computational feasibility of the approach.

Instead of relying on positive linear $\succeq^{lin}$ or, more generally, polyhedral second order $\succeq^{lin,\mathcal{P}}$, Armbruster and Luedtke [4] use the ordinary second-order stochastic dominance $\succeq_{(2)}$ in the constraints of (28). (Results for first-order stochastic dominance are given in [4] as well; we skip their description here.) This approach has the advantage that the dominance constraint is easily interpretable: only those solutions are taken into consideration that every rational, risk-averse decision maker will prefer to the reference benchmark $\mathbf{Y}$. Of course, since $\succeq_{(2)}$ is even stronger than $\succeq^{lin}_{(2)}$, the number of feasible solutions may be severely limited in this way. The authors argue that this drawback can be overcome by using a random outcome of reduced quality compared to the benchmark. (Also in other papers, e.g. [53], it has been suggested to relax the dominance over the benchmark by a tolerance parameter.) Armbruster and Luedtke show that their approach has computational advantages as well, allowing a compact LP formulation for the constraint. The LP formulation is derived through a characterization of second-order stochastic dominance

[10]Actually, the authors use the term CVaR instead of AV@R.

by Strassen's theorem which says that $X \succeq_{(2)} Y$ holds if and only if there exist random vectors X' and Y' with the same distributions as X and Y, respectively, such that X' is larger or equal to the conditional expectation of Y', given X'. The authors test their approach both on the homeland security budget allocation problem from [43]) and on an interesting medical application, a multi-objective radiation treatment planing problem under uncertainty, where the aim is to optimize the radiation dose targeted at a tumor while keeping the dose to critical organs in the neighborhood as small as possible.

5 Conclusions and Future Research

Stochastic optimization is a well established quantitative methodology for decision problems with *a priori* unknown parameters. Such problems are ubiquitous in today's economic environment and in numerous managerial decision contexts that involve planning for an uncertain future.

Many of these problems also encompass decision making with more than one objective. Traditionally, multi-objective optimization approaches have been used to tackle with situations where multiple criteria play a role.

We have surveyed articles dealing with optimization problems that are both stochastic and multi-objective. Our attention has been focused on approaches that do not eliminate the multi-objective aspect of the problem by scalarization before the computational analysis (this would require the decision maker to weigh her or his objectives), but rather sustain the multi-objective character during the analysis. We have seen that for the context of risk-neutral decisions, standard modeling concepts exist, though the computational solution frequently remains challenging. In the case of risk-averse decision making, already the modeling part becomes highly nontrivial. No silver bullet to manage problems of the latter kind has been found yet, but there are some very promising recent approaches based on risk measures and multivariate stochastic orders to deal with them.

There is vast room for future research in the intersection of stochastic and multi-objective optimization. Let us give a few examples. First, whereas a good part of the current literature focuses on qualitatively different objectives (as cost vs. health effects or return vs. environmental aims), a multi-objective perspective is already introduced into a single-objective stochastic optimization application by the consideration of different time horizons. In the simplest case, this generates a bi-objective tradeoff between a short-term (operative) and a long-term (strategic) perspective. As an example, cf. Ahmed et al. [2]. However, also several successive time horizons may be envisaged, which raises the interesting question whether a sequence of objectives with an inherent linear order structure can be treated by specifically tailored methods.

Second, state-of-the-art multi-objective optimization offers a considerable amount of theoretical results and of solution algorithms, starting with the well-developed methodology of multi-objective linear programming (MOLP) and proceeding to areas as multi-objective combinatorial or nonlinear optimization. It is widely unknown which of these results and procedures generalize (in which way) to the stochastic case.

Third, the inclusion of risk-averse problem variants in stochastic multi-objective optimization is still rather new. This is especially true for discrete problem variants. In the single-objective context, some authors have started to extend methods of risk-averse optimization (originally developed in financial engineering) from the continuous case to integer programming (see, e.g., Schultz [66]), but despite a strong demand in diverse areas, an elaborated methodology of multi-objective discrete optimization under risk does not yet seem to exist.

Fourth, as a consequence of the high complexity of SMOOPs, many real-world applications require suitable heuristic solution techniques. Whereas, e.g., multi-objective evolutionary algorithms for deterministic problems are already well-developed and are frequently used in application-oriented publications, there is still much to do to extend these techniques to the stochastic situation and to gain experience with their application to practical problems.

Finally, the field of *multistage* stochastic multi-objective optimization (which has to be distinguished from basically *static* decision making with short-term and long-term objectives) is still widely unexplored, even on the modeling level. A considerable amount of conceptual work, as well as the development of efficient solution algorithms and their test on applied problems, will still be needed to cope successfully with multistage SMOOPs.

References

[1] M. Abbas and F. Bellahcene. Cutting plane method for multiple objective stochastic integer linear programming. *European J. of Operational Research*, 168:967–984, 2006.

[2] S. Ahmed, P. Schütz, and A. Tomasgard. Supply chain design under uncertainty using sample average approximation and dual decomposition. *European Journal of Operational Research*, 199:409–419, 2009.

[3] L. Amodeo, C. Prins, and D. Sanchez. Comparison of metaheuristic approaches for multi-objective simulation-based optimization in supply chain inventory management. In *Applications of Evolutionary Computing*, Lecture Notes in Computer Science. Springer, 2009.

[4] B. Armbruster and J.B. Luedtke. Models and formulations for multivariate dominance constrained stochastic programs. Technical report, 2011.

[5] P. Artzner, F. Delbaen, J. Eber, and D. Heath. Coherent Measures of Risk. *Mathematical Finance*, 9:203–228, 1999.

[6] P. Artzner, F. Delbaen, and D. Heath. Thinking coherently. *Risk*, 10:68–71, November 1997.

[7] F.F. Baesler and J.A. Sepúlveda. Multi-objective simulation optimization for a cancer treatment center. In *Proc. Winter Simulation Conference 2001*, pages 1405–1411, 2001.

[8] M. Basseur and E. Zitzler. Handling uncertainty in indicator-based multiobjective optimization. *International Journal of Computational Intelligence Research*, 2:255–272, 2006.

[9] S.K. Bath, J.S. Dhillon, and D.P. Kothari. Stochastic multiobjective generation dispatch. *Electric Power Components and Systems*, 32:1083–1103, 2004.

[10] F. Ben Abdelaziz. *L'efficacité en Programmation Multi-objectifs Stochastique.* PhD thesis, Université de Laval, Québec, 1992.

[11] F. Ben Abdelaziz. Solution approaches for the multiobjective stochastic programming. *European Journal of Operations Research*, 216:1–16, 2012.

[12] F. Ben Abdelaziz, B. Aouni, and R. El Fayedh. Multiobjective stochastic programming for portfolio selection. *European Journal of Operations Research*, 177:1811–1823, 2007.

[13] R. Caballero, E. Cerda, M.M. Muños, and L. Rey. Stochastic approach versus multiobjective approach for obtaining efficient solutions in stochastic multiobjective programming problems. *European Journal of Operational Research*, 158:633–648, 2004.

[14] Y. Cardona-Valdés, A. Álvarez, and D. Ozdemir. A bi-objective supply chain design problem with uncertainty. *Transportation Research Part C*, 19:821–832, 2011.

[15] A. Chinchuluun and P.M Pardalos. A survey of recent developments in multiobjective optimization. *Annals of Operations Research*, 154:29–50, 2007.

[16] G.B. Dantzig. Linear programming under uncertainty. *Management Science*, pages 197–206, 1955.

[17] D. Dentcheva and G. Martinez. Two-stage stochastic optimization problems with stochastic ordering constraints on the recourse. *European Journal of Operational Research*, 219(1):1–8, 2012.

[18] D. Dentcheva and A. Ruszczyński. Optimization with stochastic dominance constraints. *SIAM Journal on Optimization*, 14(2):548–566, 2003.

[19] D. Dentcheva and A. Ruszczyński. Optimization with multivariate stochastic dominance constraints. *Mathematical Programming (Ser. B)*, 117:111–127, 2009.

[20] H. Ding, L. Benyucef, and X. Xie. A simulation-based multi-objective genetic algorithm approach for networked enterprises optimization. *Engineering Applications of Artificial Intelligence*, 19:609–623, 2006.

[21] M. Ehrgott. *Multicriteria Optimization.* Springer, Berlin, 2000.

[22] M. Ehrgott and X. Gandibleux. A survey and annotated bibliography of multiobjective combinatorial optimization. *OR Spectrum*, 22:425–460, 2000.

[23] H. Eskandari and C. D Geiger. Evolutionary multiobjective optimization in noisy problem environments. *Journal of Heuristics*, 15:559–595, 2009.

[24] H. Eskandari, L. Rabelo, and M Mollaghasemi. Multiobjective simulation optimization using an enhanced genetic algorithm. In *Proceedings of the 37th conference on Winter simulation, WSC '05. Orlando, Florida*, pages 833–841, 2005.

[25] C.I. Fábián, G. Mitra, and D. Roman. Processing second order stochastic dominance models using cutting plane representations. *Mathematical Programming*, 130(1):33–37.

[26] H. Föllmer and A. Schied. *Stochastic Finance – An Introduction in Discrete Time*. de Gruyter Studies in Mathematics 27. de Gruyter, 2004.

[27] M. Fonseca, A. García-Sánchez, M. Ortega-Mier, and F. Saldanha da Gama. A stochastic bi-objective location model for strategic reverse logistics. *TOP*, 18:158–184, 2010.

[28] W.J. Gutjahr. Two metaheuristics for multiobjective stochastic combinatorial optimization. In *Proc. SAGA 2005 (Stochastic Algorithms: Foundations and Applications)*, pages 116–125. Springer, 2005.

[29] W.J. Gutjahr. A provably convergent heuristic for stochastic bicriteria integer programming. *Journal of Heuristics*, 15:227–258, 2009.

[30] W.J. Gutjahr. Recent trends in metaheuristics for stochastic combinatorial optimization. *Central European Journal of Computer Science*, 1:58–66, 2011.

[31] W.J. Gutjahr. Multi-objective stochastic optimization under partial risk neutrality. Technical report, Univ. of Vienna, 2012.

[32] W.J. Gutjahr. A note on the utility-theoretic justification of the pareto approach in stochastic multi-objective combinatorial optimization. In L. Zadnik Stirn, J. Zerovnik, J. Povh, S. Drobne, and A. Lisec, editors, *Proc. 11th Internat. Symposium on Operational Research in Slovenia (SOR '11), Dolenjske Toplice, Slovenia, Sept. 28–30*, 2012.

[33] W.J. Gutjahr. Runtime analysis of an evolutionary algorithm for stochastic multi-objective combinatorial optimization. *Evolutionary Computation*, 20:395–421, 2012.

[34] W.J. Gutjahr. A three-objective optimization approach to cost effectiveness analysis under uncertainty. In *Proc. OR2012 (International Annual Conference of the German Operations Research Society 2012)*, 2012. Accepted for publication.

[35] W.J. Gutjahr and P. Reiter. Bi-objective project portfolio selection and staff assignment under uncertainty. *Optimization Methods and Software*, 59:417–445, 2010.

[36] W.B. Haskell, J.G. Shanthikumar, and Z.M. Shen. Aspects of increasing concave stochastic order constrained optimization. Technical report, 2012.

[37] W.B. Haskell, J.G. Shanthikumar, and Z.M. Shen. Optimization with stochastic objectives. Technical report, 2012.

[38] H.A. Hassan-Pour, M. Mosadegh-Khah, and R. Tavakkoli-Moghaddam. Solving a multi-objective multi-depot stochastic location-routing problem by a hybrid simulated annealing algorithm. In *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, pages 1045–1054, 2009.

[39] H. Heitsch, W. Römisch, and C. Strugarek. Stability of multistage stochastic programs. *SIAM J. Optimization*, 17:511–525, 2006.

[40] F. Hnaien, X. Delorme, and A. Dolgui. Multi-objective optimization for inventory control in two-level assembly systems under uncertainty of lead times. *Computers and Operations Research*, 20:1835–1843, 2010.

[41] T. Homem-de Mello. Variable-sample methods for stochastic optimization. *ACM Transactions on Modeling and Computer Simulation*, 13:108–133, 2003.

[42] T. Homem-de Mello and S. Mehrotra. A cutting surface method for uncertain linear programs with polyhedral stochastic dominance constraints. *SIAM Journal on Optimization*, 20:1250–1273, 2009.

[43] J. Hu, T. Homem-de Mello, and S. Mehrotra. Risk-adjusted budget allocation models with application in homeland security. *IIE Transactions*, 43:819–839, 2011.

[44] J. Hu, T. Homem-de Mello, and S. Mehrotra. Sample average approximation of stochastic dominance constrained programs. *Mathematical Programming*, 133:171–201, 2012.

[45] E.J. Hughes. Evolutionary multi-objective ranking with uncertainty and noise. In *Proc. EMO '01 (Evolutionary Multicriterion Optimization)*, pages 329–343. Springer, 2001.

[46] M. Köksalan, J. Wallenius, and S. Zionts. *Multiple Criteria Decision Making: From Early History to the 21st Century*. World Scientific Publishing: Singapore, 2011.

[47] T.C. Koopmans. An analysis of production as an efficient combination of activities. In *Activity Analysis of Production and Allocation*. Wiley, 1951.

[48] M. Laumanns, L. Thiele, and E. Zitzler. An efficient, adaptive parameter variation scheme for metaheuristics based on the epsilon-constraint method. *European Journal of Operational Research*, 169:932–942, 2006.

[49] A. Liefooghe, M. Basseur, L. Jourdan, and E.-G. Talbi. Combinatorial optimization of stochastic multi-objective problems: an application to the flow-shop scheduling problem. In *Evolutionary Multi-Criterion Optimization*, Lecture Notes in Computer Science 4403, pages 457–471. Springer, 2007.

[50] A. Liefooghe, M. Basseur, L. Jourdan, and E.-G. Talbi. ParadisEO-MOEO: a framework for evolutionary multi-objective optimization. In *Evolutionary Multi-Criterion Optimization*, Lecture Notes in Computer Science 4403, pages 386–400. Springer, 2007.

[51] A. Müller and D. Stoyan. *Comparison methods for stochastic models and risks*. Wiley series in probability and statistics. Wiley, Chichester, 2002.

[52] V.I. Norkin, Y.M. Ermoliev, and A. Ruszczyński. On optimal allocation of indivisibles under uncertainty. *Operations Research*, 46:381–395, 1998.

[53] N. Noyan and G. Rudolf. Optimization with multivariate conditional value-at-risk constraints. Technical report, 2012.

[54] W. Ogryczak and A. Ruszczyński. From stochastic dominance to mean-risk models: Semideviations as risk measures. *European Journal of Operational Research*, 116:33–50, 1999.

[55] W. Ogryczak and A. Ruszczyński. On consistency of stochastic dominance and mean–semideviation models. *Math. Program., Ser. B*, 89:217–232, 2001.

[56] W. Ogryczak and A. Ruszczyński. Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization*, 13(1):60–78, 2002.

[57] G.Ch. Pflug. Scenario tree generation for multiperiod financial optimization by optimal discretization. *Mathematical Programming*, 89:251–271, 2001.

[58] G.Ch. Pflug. Stochastic optimization and statistical inference. In Andrzej Ruszczyński and Alexander Shapiro, editors, *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*, chapter 7, pages 427–482. Elsevier, 2003.

[59] G.Ch. Pflug and A. Pichler. *Approximations for Probability Distributions and Stochastic Optimization Problems*, volume 163 of *International Series in Operations Research & Management Science*, chapter 15, pages 343–387. Springer New York, 2011.

[60] G.Ch. Pflug and A. Pichler. A distance for multistage stochastic optimization models. *SIAM Journal on Optimization*, 22:1–23, 2012.

[61] G.Ch. Pflug and W. Römisch. *Modeling, Measuring and Managing Risk.* World Scientific, River Edge, NJ, 2007.

[62] G.Ch. Pflug and N. Wozabal. Asymptotic distribution of law-invariant risk functionals. *Finance and Stochastics*, 14:397–418, 2010.

[63] S. Rath, M. Gendreau, and W.J. Gutjahr. Bi-objective stochastic programming models for determining depot locations in disaster relief operations planning. Technical report, 2012.

[64] D. Roman, K. Darby-Dowman, and G. Mitra. Mean-risk models using two risk measures: a multi-objective approach. *Quantitative Finance*, 7:443–458, 2007.

[65] Th. Saaty and S. Gass. Parametric objective function (part i). *Operations Research*, pages 316–319, 1954.

[66] R. Schultz. Risk aversion in two-stage stochastic integer programming. In *Stochastic Programming*, International Series in Operations Research and Management Science 150, pages 165–187. Springer New York, 2011.

[67] M. Shaked and J. G. Shanthikumar. *Stochastic Order*. Springer Series in Statistics. Springer, 2007.

[68] A. Shapiro, A. Ruszczyński, and D. Dentcheva. *Lectures on Stochastic Programming*. MPS-SIAM Series on Optimization 9, 2009.

[69] A. Shapiro and H. Xu. Uniform laws of large numbers for set-valued mappings and subdifferentials of random functions. *Mathematical Analysis and Applications*, 325:1390–1399, 2007.

[70] A. Syberfeldt, A. Ng, R.I. John, and Ph. Moore. Multi-objective evolutionary simulation-optimisation of a real-world manufacturing problem. *Robotics and Computer-Integrated Manufacturing*, 25:926–931, 2009.

[71] J. Teich. Pareto-front exploration with uncertain objectives. In *Proc. EMO '01 (Evolutionary Multicriterion Optimization)*, pages 314–328. Springer, 2001.

[72] F. Tricoire, A. Graf, and W.J. Gutjahr. The bi-objective stochastic covering tour problem. *Computers an Operations Research*, 39:1582–1592, 2012.