FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

HABILITATION THESIS

Martin Branda

Advances in stochastic programming approaches to optimization under uncertainty

Department of Probability and Mathematical Statistics

Mathematics
Probability and Mathematical Statistics

Prague 2018

Contents

Preface

This thesis presents several contributions of the author to the branch of mathematics called stochastic programming or stochastic optimization. Many real life problems lead to optimization problems where some parts need not to be known precisely, thus they are uncertain. Stochastic programming provides methods for dealing with uncertainty when it is represented by random variables with a known or estimated probability distribution.

In chance (probabilistic) constrained problems, the constraints dependent on a random vector with known distribution must be fulfilled with a high prescribed probability and can be violated with the small complementary probability. Such trade-off is useful in many real life applications, e.g., in finance and engineering, but at the same time it leads to problems which are difficult to solve. In Section 1.1, we introduce two solution approaches to the chance constrained problems. The first approach is based on modern tools of variational analysis developed for nonlinear nonconvex problems. It employs the optimality conditions and regularization to construct a strong numerical method, cf. Section 1.1.1. The second approach is based on penalty functions, see Section 1.1.3. We review a recent result on an asymptotic equivalence of the chance constrained problems and the problems with penalty objective with the exact penalty property. The optimality conditions were derived and the regularization was used for similar problems with cardinality constraints where the number of nonzero elements of the decision vector is bounded, see Section 1.1.2. The probabilistic criteria can appear also in scheduling problems. In Section 1.1.4, we deal with the fixed interval scheduling problems with stochastic delays in processing times where the probability that the schedule remains feasible is maximized.

Dealing with the random parts of the stochastic programming problems is highly demanding task. In Section 1.2, we discuss the sample approximation techniques where the original distribution is replaced by a random sample. In particular, we focus on the problems with mixed-integer set of feasible solutions. Section 1.2.1 deals with the chance constrained problems, whereas section 1.2.2 is focused on the problems with several expected value constraints. We will discuss convergence of the sample approximated sets to the original ones under increasing sample size and provide rates of convergence.

Section 1.3 is devoted to a new approach to efficiency analysis of investment opportunities available on financial markets. Data envelopment analysis (DEA) models were generalized to take into account diversification among the considered assets leading to diversification-consistent (DC) DEA. This generalization is done in compliance with the basic principles of financial mathematics, stochastic and multiobjective optimization. Section 1.3.2 is focused on the formulations and basic properties of the DC DEA models. In Section 1.3.3, relations of the DC

DEA models to the stochastic dominance efficiency are investigated.

This thesis gives only a review of the main results and basic literature, thus many explanations, references and all proofs are skipped. We refer the readers to the original papers which are in most cases attached to this thesis and contain all necessary details.

The thesis is primarily based on the papers listed below:

1. Adam, L., Branda, M. (2016). **Nonlinear chance constrained problems: optimality conditions, regularization and solvers.** *Journal of Optimization Theory and Applications* 170 (2): 419–436.
   The paper derives new optimality conditions for the nonlinear chance constrained problems and their relaxations. The relaxed problems can be obtained by relaxing the binary variables which appear in a mixed-integer reformulation of the chance constraints. Relations between the stationary points, local and global minima of the problems are discussed. New numerical methods based on relaxation and regularization are introduced, their convergence is proven and the numerical performance is verified.

2. Branda, M., Bucher, M., Červinka, M., Schwartz, A. (2018). **Convergence of a Scholtes-type Regularization Method for Cardinality-Constrained Optimization Problems with an Application in Sparse Robust Portfolio Optimization.** *Computational Optimization and Applications* 70 (2): 503–530.
   The paper is focused on the optimization problems with cardinality constraints where the number of nonzero elements of the decision vector is bounded. The authors investigate so called Scholtes regularization, for which convergence of the obtained sequence of KKT points to a strong stationary point is proven. In the numerical study, the performance of the regularization and alternative solution approaches is compared on several investment problems with robust risk measures.

3. Branda, M. (2013). **On relations between chance constrained and penalty function problems under discrete distributions.** *Mathematical Methods of Operations Research* 77 (2): 265–277.
   The relations between the chance constrained problems and the problems with penalty objective are discussed. Earlier results on asymptotic equivalence of these problems are generalized in two ways. First, the equivalence is shown under a discrete distribution of the random parts. Then, the exact penalty property which is ensured by a modified calmness condition is employed to verify the equivalence for a finite value of the penalty parameter.

4. Branda, M., Novotný, J., Olstad, A. (2016). **Fixed interval scheduling under uncertainty - a tabu search algorithm for an extended robust coloring formulation.** *Computers & Industrial Engineering* 93: 45–54.
   The paper is focused on fixed interval scheduling problems where the job processing intervals can be delayed. The goal is to obtain a schedule with the highest probability that it remains feasible after observing the delays. This

leads to a difficult stochastic integer programming problem with probabilistic objective. Under particular assumptions, the problem is reformulated as a generalized robust coloring problem. Several solution approaches are discussed.

5. Branda, M. (2012). **Sample approximation technique for mixed-integer stochastic programming problems with several chance constraints.** *Operations Research Letters* 40 (3): 207–211.
The paper deals with the sample approximation technique applied to stochastic programming problems with chance constraints. The results on rates of convergence are extended for problems with mixed-integer bounded sets of feasible solutions and several chance constraints. New estimates on the sample size necessary to get a feasible solution of the original problem using the sample approximation are derived. An application to a vehicle routing problem with stochastic elements is presented.

6. Branda, M. (2014). **Sample approximation technique for mixed-integer stochastic programming problems with expected value constraints.** *Optimization Letters* 8 (3): 861–875.
The paper deals with the theory of sample approximation techniques applied to stochastic programming problems with expected value constraints. The results on the rates of convergence are extended to the problems with a mixed-integer bounded set of feasible solutions and several expected value constraints. Moreover, non-iid sampling and Hölder-calmness of the constraints are enabled. The estimates on the sample size necessary to get a feasible solution of the original problem using the sample approximation are derived. An application to an investment problem with transaction costs is included.

7. Branda, M. (2013). **Diversification-consistent data envelopment analysis with general deviation measures.** *European Journal of Operational Research* 226 (3): 626–635.
The paper introduces new efficiency tests which are based on traditional DEA models and take into account portfolio diversification. The goal is to identify the investment opportunities that perform well without specifying our attitude to risk. General deviation measures are used as the inputs and return measures as the outputs. The choice of the set of investment opportunities is discussed. We compare the optimal values (efficiency scores) of all proposed models leading to the relations between the sets of efficient opportunities. Strength of the new DEA models is then discussed and compared in the empirical part.

8. Branda, M. (2015). **Diversification-consistent data envelopment analysis based on directional-distance measures.** *Omega–International Journal of Management Science* 52: 65–76.
New diversification-consistent DEA models based on directional distance measures are introduced. The models enable to use several risk measures as the inputs and return measures as the outputs, which can take both positive and negative values. It is shown that strongly, semi-strongly or weakly

Pareto–Koopmans efficient investment opportunities can be identified depending on the selected model. Moreover, the optimal solutions correspond to efficient investment opportunities and can be used by investors to revise the inefficient ones. It is proven that under proper choice of the inputs and outputs, the strongest model is able to identify efficient investment opportunities with respect to the second-order stochastic dominance. Moreover, the model can be formulated as a linear programming problem. The proposed DEA models are applied to 48 representative industry portfolios from US stock markets.

9. Branda, M. (2016). **Mean-value at risk portfolio efficiency: approaches based on data envelopment analysis models with negative data and their empirical behaviour.** *4OR–A Quaterly Journal of Operations Research* 14 (1): 77–99.
   The paper deals with the problem of an investor who is using a mean-risk model for accessing efficiency of investment opportunities. The investor employs value at risk on several risk levels at the same time which corresponds to the approach called risk shaping. It is shown that a diversification-consistent extension of the DEA models based on a directional distance measure can be used to identify the Pareto–Koopmans efficient investment opportunities. Reformulations as chance constrained, nonlinear and mixed-integer problems are derived under particular assumptions. In the numerical study, efficiency of US industry representative portfolios is accessed.

10. Branda, M., Kopa, M. (2016). **DEA models equivalent to general N-th order stochastic dominance efficiency tests.** *Operations Research Letters* 44 (2): 285–289.
   The paper introduces data envelopment analysis (DEA) models equivalent to efficiency tests with respect to the N-th order stochastic dominance (NSD). In particular, it is focused on strong and weak variants of convex NSD efficiency and NSD portfolio efficiency. The proposed DEA models are in relation with the strong and weak Pareto–Koopmans efficiencies and employ N-th order lower and co-lower partial moments.

Author's contributions to stochastic programming theory and applications

Although the topics in the following sections may seem unrelated, there are strong connections between the results discussed below. Sample approximation techniques can be used to obtain the discrete distribution which we assume in the parts focused on the optimality conditions and penalization for the chance constrained problems. Financial data envelopment analysis models lead to solving chance constrained problem when Value at Risk measures are used as the inputs. The other DEA models lead to expectation constrained problems for which the sample approximation technique is also discussed.

1.1 Chance constrained programming and related problems

The chance constrained problems (CPP) were introduced by Charnes et al. (1958) and since then many theoretical results, algorithms and applications were proposed. First, we give an overview of recent results concerning CCP. A general approach called sample (empirical) approximation is based on substituting the underlying continuous distribution by a finite sample and on reformulation as a (large) mixed-integer programming problem. The crucial question is the choice of the sample size, which is usually based on the exponential rates of convergence derived, e.g., by Kaňková (1990); Luedtke and Ahmed (2008). We will focus on this topic in Section 1.2 where the estimates for mixed-integer problems are introduced. However, these estimates can be too conservative, cf. Henrion (2013). Recently, Barrera et al. (2016) employed the importance sample technique to solve a chance constrained telecommunications problem with Bernoulli input distributions. Exploiting its structure, they derived conditions to ensure a uniform variance reduction.

For linear constraints and finite discrete distribution, strong results and algorithms based on cutting planes for mixed-integer reformulations are available, cf. Beraldi and Bruni (2010); Luedtke (2014); Luedtke et al. (2010). Recently, Zhao et al. (2017) derived new strong valid inequalities based on intersection of multiple mixing sets for the chance constrained problems with random right-hand side.

Nonlinear programming algorithms were suggested for CCP by Prékopa (2003) and further developed by Dentcheva and Martinez (2012, 2013). Recently, Geletu et al. (2017) proposed a smooth approximation approach employing an inner and an outer analytic approximation of chance constraints leading to two classes of nonlinear programming problems. Xie and Ahmed (2017) introduced quantile cuts which can be obtained as a projection of the mixing inequalities valid for the mixed-integer nonlinear programming reformulation onto the original problem space. Ahmed et al. (2017) proposed new Lagrangian dual problems for CCP and derived dual bounds which are superior to the bounds obtained from the continuous relaxation of a standard mixed-integer programming. They also provided several exact and heuristic algorithms for construction of the new bounds.

A wide class of approaches is based on approximation of the indicator function by a more tractable function. Approximation based on conditional value at risk has been deeply investigated by Rockafellar and Uryasev (2000, 2002); Sun et al. (2014). Similar idea was used by Haneveld and van der Vlerk (2006) who employed the so-called integrated chance constraints. Other methods were based on penalty functions applied to the random constraints Branda (2012c,a, 2013b). Bernstein approximation has been introduced by Nemirovski and Shapiro (2007) for constraints affine in random coefficients and further developed by Ahmed (2014). Recently, algorithmic approaches based on representation using difference of convex (DC) functions appeared in the literature, see Shan et al. (2014); Sun et al. (2014); Wozabal et al. (2010). A second-order cone programming reformulation was obtained by Cheng and Lisser (2012) for problems with linear constraints under normally distributed random coefficient and under independence and copula dependence of the rows. For these linear-Gaussian problems, Henrion and Möller (2012) provided an explicit gradient formula and derived an efficient solution procedure. This formula was generalized for nonlinear CCP by van Ackooij and Henrion (2014) who employed the spherical-radial decomposition of Gaussian random vectors.

1.1.1 Optimality condition for nonlinear chance constrained problems

Adam and Branda (2016a) contributed to deeper understanding and algorithmic approaches for nonlinear chance constrained problems. The investigated problem with single individual chance constraint may be formulated as follows:

$$\begin{aligned}\min_x \ & f(x)\\ \text{s. t.} \ & P(g(x,\xi)\le 0)\ge 1-\varepsilon,\\ & h_j(x)\le 0,\ j=1,\dots,J.\end{aligned} \tag{1.1}$$

Here $x \in \mathbb{R}^n$ is the decision variable, $0 \le \varepsilon < 1$ is a prescribed probabilistic level, $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$ and $h_j : \mathbb{R}^n \to \mathbb{R}$ are functions which are continuously differentiable in variable x and finally $\xi \in \mathbb{R}^d$ is a random vector with known probability distribution P.

We assume that ξ has known discrete distribution and enumerate all possible realizations by $\xi_1, \dots, \xi_S$ and the corresponding positive probabilities by

$p_1, \dots, p_S$. We may then reformulate problem (1.1) into

$$
\begin{aligned}
\min_{x}\ & f(x) \\
\text{s. t.}\ & \sum_{i=1}^{S} p_i \mathbb{I}\left(g(x,\xi^i) \le 0\right) \ge 1-\varepsilon, \\
& h_j(x) \le 0,\ \ j = 1,\dots,J,
\end{aligned}
\tag{1.2}
$$

where $\mathbb{I}$ stands for the characteristic function which equals to 1 if $g(x,\xi^i) \le 0$ and 0 otherwise. Introducing artificial binary variable $y \in \{0,1\}^S$ to deal with $\mathbb{I}$, we obtain the following mixed–integer nonlinear problem

$$
\begin{aligned}
\min_{x,y}\ & f(x) \\
\text{s. t.}\ & \sum_{i=1}^{S} p_i y_i \ge 1-\varepsilon, \\
& y_i \in \{0,1\},\ \ i = 1,\dots,S, \\
& g(x,\xi^i) y_i \le 0,\ \ i = 1,\dots,S, \\
& h_j(x) \le 0,\ \ j = 1,\dots,J.
\end{aligned}
\tag{1.3}
$$

To avoid the problematic constraint $g(x,\xi^i)y_i \le 0$, usually this constraint is replaced by $g(x,\xi^i) \le M(1-y_i)$, where M is a sufficiently large constant, cf. Raike (1970). Since this problem is difficult to tackle by mixed-integer (nonlinear) programming techniques in any of the previous forms, we relax binary constraint $y_i \in \{0,1\}$ into $y_i \in [0,1]$ to obtain nonlinear programming relaxed problem

$$
\begin{aligned}
\min_{x,y}\ & f(x) \\
\text{s. t.}\ & \sum_{i=1}^{S} p_i y_i \ge 1-\varepsilon, \\
& 0 \le y_i \le 1,\ \ i = 1,\dots,S, \\
& g(x,\xi^i) y_i \le 0,\ \ i = 1,\dots,S, \\
& h_j(x) \le 0,\ \ j = 1,\dots,J.
\end{aligned}
\tag{1.4}
$$

Even though problems (1.2) and (1.4) are not equivalent, there are close similarities between them. The global minima are identical, whereas the relaxed problem can contain additional local minima and stationary points, see Adam and Branda (2016a) for details.

To obtain a local nonlinear programming description of the feasibility set, we define the following index sets

$$
\begin{aligned}
I_{00}(x,y) &:= \left\{i : g(x,\xi^i) = 0,\ y_i = 0\right\}, & I_0(x) &:= \left\{i : g(x,\xi^i) = 0\right\}, \\
I_{0+}(x,y) &:= \left\{i : g(x,\xi^i) = 0,\ 0 < y_i < 1\right\}, & J_0(x) &:= \{j : h_j(x) = 0\}, \\
I_{01}(x,y) &:= \left\{i : g(x,\xi^i) = 0,\ y_i = 1\right\}.
\end{aligned}
\tag{1.5}
$$

Further, we define sets which will be used later to derive the optimality conditions

$$\hat{\mathcal{I}}(x) := \left\{ \tilde{I} \subset \{1,\dots,S\} : \sum_{i\in\tilde{I}} p_i \geq 1-\varepsilon, \begin{array}{ll} g(x,\xi^i) & < 0 \implies i \in \tilde{I} \\ g(x,\xi^i) & > 0 \implies i \notin \tilde{I} \end{array} \right\},$$

$$\mathcal{I}(x) := \text{minimal elements of } \hat{\mathcal{I}}(x) \text{ with respect to set inclusion,} \tag{1.6}$$

$$Y(x) := \Big\{ y \in \{0,1\}^S : \exists I \in \mathcal{I}(x) \text{ s.t. } y_i = 1 \text{ if } i \in I;\ y_i = 0 \text{ if } i \notin I \Big\}. \tag{1.7}$$

The purpose of the definition of these sets is that the union of $\{x : \ g(x,\xi^i) \leq 0,\ i \in I\}$ with respect to $I \in \hat{\mathcal{I}}(\bar{x})$ or with respect to $I \in \mathcal{I}(\bar{x})$ has the same shape as the feasible set of problem (1.2) on the neighborhood of $\bar{x}$.

To be able to derive optimality conditions for both the original chance constrained problem (1.2) and the nonlinear relaxed problem (1.4), we impose the following assumption.

Assumption 1.1.1. *Let $\bar{x}$ be a feasible point of problem* (1.2)*. Assume that at least one of the following two conditions is satisfied:*

- *function $g(\cdot,\xi^i)$ and h_j are affine linear.*
- *the following implication is satisfied for all $I \in \mathcal{I}(\bar{x})$*

$$\left.\begin{array}{r} \sum_{i\in I_0(\bar{x})} \lambda_i \nabla_x g(\bar{x},\xi^i) + \sum_{j\in J_0(\bar{x})} \mu_j \nabla h_j(\bar{x}) = 0 \\ \lambda_i = 0,\ i \in I_0(\bar{x}) \setminus I \\ \lambda_i \geq 0,\ i \in I_0(\bar{x}) \cap I \\ \mu_j \geq 0,\ j \in J_0(\bar{x}) \end{array}\right\} \implies \begin{array}{l} \lambda_i = 0,\ i \in I_0(\bar{x}), \\ \mu_j = 0,\ j \in J_0(\bar{x}). \end{array}$$

The second part of the assumption represents a modified Mangasarian-Fromowitz constraints qualification.

For a set $Z \subset \mathbb{R}^n$ and a point $\bar{x} \in Z$, we can define the tangent and Fréchet normal cones to Z at $\bar{x}$ respectively as

$$\begin{aligned} T_Z(\bar{x}) &:= \Big\{ d \in \mathbb{R}^n : \exists d^k \to d\ \exists t^k \searrow 0 : \bar{x} + t^k d^k \in Z \Big\}, \\ \hat{N}_Z(\bar{x}) &:= \{ x^* \in \mathbb{R}^n : \langle x^*, d\rangle \leq 0 \text{ for all } d \in T_Z(\bar{x}) \}. \end{aligned}$$

General necessary optimality condition at $\bar{x} \in Z$ for a problem with differentiable objective function f

$$\min_{x\in Z} f(x)$$

states

$$-\nabla f(\bar{x}) \in \hat{N}_Z(\bar{x}), \tag{1.8}$$

see (Rockafellar and Wets, 1998, Theorem 6.12).

Adam and Branda (2016a) derived necessary optimality condition for the original problem (1.2) and its relaxation (1.4).

Theorem 1.1.1. (Adam and Branda (2016a))
Let Assumption 1.1.1 be satisfied. If $\bar{x}$ is a local minimum of chance constrained

problem (1.2)*, then for every* $I \in \mathcal{I}(\bar{x})$ *there exist multipliers* $\lambda_i,\ i \in I_0(\bar{x})$ *and* $\mu_j,\ j \in J_0(\bar{x})$ *such that*

$$\begin{aligned} \nabla f(\bar{x}) + \sum_{i \in I_0(\bar{x})} \lambda_i \nabla_x g(\bar{x}, \xi^i) + \sum_{j \in J_0(\bar{x})} \mu_j \nabla h_j(\bar{x}) = 0,& \\ \lambda_i = 0,\ i \in I_0(\bar{x}) \setminus I,& \\ \lambda_i \geq 0,\ i \in I_0(\bar{x}) \cap I,& \\ \mu_j \geq 0,\ j \in J_0(\bar{x}).& \end{aligned} \tag{1.9}$$

Theorem 1.1.2. (Adam and Branda (2016a))
Assume that $(\bar{x}, \bar{y})$ *is a local minimum of problem* (1.4) *and that Assumption 1.1.1 is satisfied at* $\bar{x}$*. Then there exist multipliers* $\lambda_i,\ i \in I_0(\bar{x})$ *and* $\mu_j,\ j \in J_0(\bar{x})$ *such that*

$$\begin{aligned} \nabla f(\bar{x}) + \sum_{i \in I_0(\bar{x})} \lambda_i \nabla_x g(\bar{x}, \xi^i) + \sum_{j \in J_0(\bar{x})} \mu_j \nabla h_j(\bar{x}) = 0,& \\ \lambda_i = 0,\ i \in I_{00}(\bar{x}, \bar{y}),& \\ \lambda_i \geq 0,\ i \in I_{01}(\bar{x}, \bar{y}) \cup I_{0+}(\bar{x}, \bar{y}),& \\ \mu_j \geq 0,\ j \in J_0(\bar{x}).& \end{aligned} \tag{1.10}$$

The proofs employ general optimality conditions (1.8). The crucial part is the expression of the tangent and normal cones for the set of feasible solution with our special constraint structure. Moreover, the derivation of the optimality conditions for the relaxed problem uses the decomposition in x and y dimensions and the chain rule for the tangent and normal cones, cf. (Rockafellar and Wets, 1998, Chapter 6).

Since the numerical methods for the relaxed problem are still unstable, Adam and Branda (2016a) introduced an algorithm based on the regularization. This technique enlarges the feasible set and solves the resulting regularized problem while driving the regularization parameter to infinity. Thus, we consider the regularized problem

$$\begin{aligned} \min_{x,y}\ & f(x) \\ \text{s. t.}\ & \sum_{i=1}^{S} p_i y_i \geq 1 - \varepsilon, \\ & 0 \leq y_i \leq 1,\ i = 1, \ldots, S, \\ & y_i \leq \phi_t(g(x, \xi^i)),\ i = 1, \ldots, S, \\ & h_j(x) \leq 0,\ j = 1, \ldots, J, \end{aligned} \tag{1.11}$$

where $\phi_t : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable decreasing function which depends on a parameter $t > 0$ and which satisfies the following properties:

$$\phi_t(0) = 1, \tag{1.12}$$

$$\phi_t(z) > 0, \quad \text{for } z \in \mathbb{R}, \tag{1.13}$$

$$\phi_t(z^t) \to 0, \quad \text{whenever } z^t \overset{t \to \infty}{\to} \bar{z} > 0, \tag{1.14}$$

$$\frac{\phi_t'(z^t)}{\phi_t'(\tilde{z}^t)} \to 0, \quad \text{whenever } \phi_t(z^t) \searrow 0 \text{ and } \phi_t(\tilde{z}^t) \to \bar{z} > 0. \tag{1.15}$$

As an example of such regularizing function, we may consider $\phi_t(z) = e^{-tz}$. Convergence of the stationary points can be obtained.

Theorem 1.1.3. (Adam and Branda (2016a))
Consider $(\bar{x}^t, \bar{y}^t)$ *to be stationary points of problem* (1.11). *Assume that the second part of Assumption 1.1.1 is satisfied at* $\bar{x}$ *and that* $(\bar{x}^t, \bar{y}^t) \to (\bar{x}, \bar{y})$ *as* $t \to \infty$. *Then* $(\bar{x}, \bar{y})$ *is a stationary point of problem* (1.4).

The above results were used by Adam and Branda (2016a) to derive several numerical methods for difficult nonlinear chance constrained problems. Branda (2016a) employed the algorithms to solve an investment problem with variance–skewness objective and a chance constraint imposed on a minimal acceptable portfolio return. Adam et al. (2018) generalized the above results to the joint chance constrained problems and derived a new algorithm based on the Benders decomposition to further improve the numerical performance for a high number of scenarios. They solved a difficult gas network design problem with random demands where the network topology was investigated by Gotzes et al. (2016).

1.1.2 Problems with cardinality constraints

Burdakov et al. (2016); Červinka et al. (2016) employed a similar approach based on optimality condition and regularization for cardinality constrained optimization problems, where the number of nonzero elements of the decision vector is limited. Such limitation appears, e.g., in portfolio theory where the restricted number of assets can stabilize out-of-sample performance of the portfolio and at the same time reduce the transaction costs, cf. Beck and Eldar (2013); Bonami and Lejeune (2009); Gao and Li (2013); Zheng et al. (2014). Inverse modelling in atmosphere in another important area of applications, see Adam and Branda (2016b). We will focus on the most recent contribution proposed by Branda et al. (2018).

For a vector $x \in \mathbb{R}^n$ we denote the support and its cardinality by

$$\mathrm{supp}(x) := \{i = 1, \dots, n \mid x_i \neq 0\} \quad \text{and} \quad \|x\|_0 := |\,\mathrm{supp}(x)|.$$

Let us formulate a general cardinality constrained problem

$$\begin{aligned} \min_x\, & f(x) \\ \text{s. t.}\ \ & \|x\|_0 \leq \kappa, \\ & h_j(x) \leq 0,\ j = 1, \dots, J, \end{aligned} \tag{1.16}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $h_j : \mathbb{R}^n \to \mathbb{R}$ are assumed to be continuously differentiable. As in the chance constrained case, we can use additional binary variables y_i to identify the nonzero elements of the decision vector x. The problem (1.16) can be reformulated using binary decision variables into

$$\begin{aligned} \min_{x,y}\ \ & f(x) \\ \text{s. t.}\ \ & \sum_{i=1}^{n} y_i \geq n - \kappa, \\ & y_i \in \{0, 1\},\ i = 1, \dots, n, \\ & x_i y_i = 0,\ i = 1, \dots, n, \\ & h_j(x) \leq 0,\ j = 1, \dots, J, \end{aligned} \tag{1.17}$$

If x_i is nonzero, then the corresponding y_i must be equal to zero, whereas if x_i is zero, then the corresponding y_i can be equal to one and by the reformulated cardinality constraint $\sum_{i=1}^{n} y_i \geq n - \kappa$ this must happen at least $n - \kappa$ times. However, as even for simple instances of cardinality constrained problems Bienstock (1996) showed the problem to be NP-complete, solving (1.17) even using specialized global solution techniques can be computationally very time demanding. Thus, we instead consider the continuous reformulation of (1.17) suggested by Burdakov et al. (2016) where the binary restrictions are relaxed to continuous ones, i.e., we impose $0 \leq y_i \leq 1$ only. Note that the relaxed problem is closely related to a mathematical program with complementarity constraints (MPCC) due to the "half-complementarity" constraints $y_i \geq 0, x_i y_i = 0$.

Branda et al. (2018) focused on the Scholtes regularization technique where the constraints $x_i y_i = 0$ are replaced by $-t \leq x_i y_i \leq t$ with a positive parameter t. The theoretical analysis as well as the numerical method concern the case when the parameter t decreases to zero. Branda et al. (2018) investigated convergence of the stationary points under the modified Mangasarian-Fromowitz constraint qualification. They showed that the stationary points of the regularized problems converge to a strong-stationary (S-stationary) point, see Červinka et al. (2016) for the definition of stationary points for the cardinality constrained problems. The proposed numerical method was then applied to several robust portfolio optimization problems showing its good performance in comparison with the commercial mixed-integer solver Gurobi and the Kanzow-Schwartz regularization method, see Branda et al. (2018) for details.

1.1.3 Asymptotic equivalence under exact penalization

An asymptotic equivalence of the chance constrained problems and the problems with penalty objective was shown first by Ermoliev et al. (2000) for a particular choice of the penalty function and one chance constraint. The approach was extended to the whole class of penalty functions, cf. Branda and Dupačová (2012), and to problems with several individual and joint chance constraints, see Branda (2012a). However, the assumptions of the theorems restricted the validity to continuous distributions only even though the equivalence was observed also for problems with finite discrete distribution of random parts. This section is focused on results obtained by Branda (2013b) who proposed bounds on optimal values and convergence of optimal solutions under the discrete distribution. Moreover, exact penalization under a modified calmness property is employed to improve the results. In particular, they proposed a modified calmness property which ensures the exact penalization and at the same time enables to derive the asymptotic equivalence. In general, the performance of the penalty method is improved in the case of exact penalization where a finite value of the penalty parameter is sufficient to get a local optimal solution of the original nonlinear programming problem, see, e.g., Bazaraa et al. (2006); Rockafellar and Wets (1998). The most general results can be obtained under the calmness condition (without convexity), see Clarke (1983); Burke (1991a,b).

Let $f(x)$ be a real function on $\mathbb{R}^n$, $g_i(x, \xi), i = 1, \ldots, k$, be real functions on $\mathbb{R}^n \times \mathbb{R}^{n'}$ measurable in ξ for all $x \in X \subseteq \mathbb{R}^n$, and ξ be a random vector on $(\Omega, \mathcal{A}, P)$ with values in $(\Xi, \mathcal{F})$, $\Xi \subseteq \mathbb{R}^{n'}$. The problem with one joint chance constraint can be formulated as follows:

$$\begin{array}{rcl} \varphi_\varepsilon^{CCP} & = & \min_{x\in X} f(x) \\ & \text{s.t.} & \\ & & P\Big(g_1(x,\xi) \le 0, \ldots, g_k(x,\xi) \le 0\Big) \ \ \ge \ \ 1-\varepsilon, \end{array} \tag{1.18}$$

with an optimal solution x_ε^{CCP} and a given level $\varepsilon \in (0,1)$. Let the distribution of random vector ξ be discrete with finite number of realizations ξ^s, $s = 1, \ldots, S$ with known probabilities $0 < p_s \le 1$, $\sum_{s=1}^S p_s = 1$.

Below, we will consider the penalty function $\vartheta : \mathbb{R}^k \to \mathbb{R}_+$ which is continuous nondecreasing in its components, equal to 0 on $\mathbb{R}^k_-$ and positive otherwise. Two special penalty functions are readily available: $\vartheta^{1,o}(u) = \sum_{i=1}^k ([u_i]^+)^o, o > 0$, usually $o = 1$ or $o = 2$, and $\vartheta^2(u) = \max_{1\le i\le k}[u_i]^+$. Both penalty functions can preserve convexity of the random constraints in the decision vector. We denote the penalized constraints

$$\Phi(x,\xi) = \vartheta(g_1(x,\xi), \ldots, g_k(x,\xi)) : \mathbb{R}^n \times \mathbb{R}^{n'} \to \mathbb{R}.$$

Then it holds

$$\begin{array}{rcl} P\Big(g_1(x,\xi) \le 0, \ldots, g_k(x,\xi) \le 0\Big) & \ge & 1-\varepsilon \\ & \Longleftrightarrow & \\ P\Big(\Phi(x,\xi) = 0\Big) & \ge & 1-\varepsilon. \end{array}$$

The expectations of the penalized constraints can be incorporated into the objective function as a penalty term leading to the problem:

$$\varphi_N = \min_{x\in X} \Big[f(x) + N \cdot \mathbb{E}[\Phi(x,\xi)]\Big], \tag{1.19}$$

with N being a positive parameter. We denote x_N an optimal solution of (1.19).

The optimal permanently feasible solution can be defined as the optimal solution of the following problem

$$\begin{array}{l} \min\limits_{x\in X} f(x) \\ \text{s.t.} \\ g_1(x,\xi^s) \le 0, \ldots, g_k(x,\xi^s) \le 0, \ s = 1, \ldots, S, \end{array} \tag{1.20}$$

which is formulated as nonlinear programming problem with inequality constraints. The same solution can be obtained also for the choice $\varepsilon < \min_s p_s$ in problem (1.18).

We will show how the exact penalization can improve the results on asymptotic equivalence of chance constrained and penalty function problems. The result is based on general calmness condition, see Burke (1991a,b). We consider the following perturbed version of the problem (1.20):

$$\begin{array}{l} \min\limits_{x\in X} f(x) \\ \quad \text{s.t.} \\ \quad g_1(x,\xi^s) \le u_{1s}, \ldots, g_k(x,\xi^s) \le u_{ks}, s = 1, \ldots, S. \end{array} \tag{1.21}$$

We define modified l_1-norm for a vector $u \in \mathbb{R}^{kS}$ as

$$\|u\| = \sum_{s=1}^{S} p_s \sum_{i=1}^{k} |u_{is}|,$$

which is necessary for further steps and to show the asymptotic equivalence. We propose the definition of calmness given by (Burke, 1991a, Definition 1).

Definition 1.1.1. *Let x^* be feasible for the unperturbed problem (1.20). Then the problem is said to be calm at x^* if there exist constant $\tilde{N}$ (modulus) and $\epsilon > 0$ (radius) such that for all $(x, u) \in \mathbb{R}^n \times \mathbb{R}^{kS}$ satisfying $x \in B_\epsilon(x^*)$ and $g_i(x, \xi^s) \leq u_{is}$, one has*

$$f(x) + \tilde{N}\,\|u\| \geq f(x^*).$$

Note that then x^* is necessarily a local solution to the unperturbed problem.

Proposition 1.1.4. (Branda (2013b))
Let x^* be feasible for the unperturbed problem, i.e. (1.21) with $u_{ks} = 0$, $i = 1, \ldots, k, s = 1, \ldots, S$. Then the unperturbed problem is calm at x^* with modulus $\tilde{N}$ and radius $\epsilon > 0$ if and only if x^* is a local minimum of the function

$$f(x) + N \sum_{s=1}^{S} p_s \sum_{i=1}^{k} |g_i(x, \xi^s)|_+$$

over $B_\epsilon(x^*)$ for all $N \geq \tilde{N}$.

The proposition follows directly from the basic theorem on exact penalization which was proposed by (Burke, 1991a, Theorem 1.1).

The following theorem focuses on the local minimizers and shows the asymptotic equivalence under the exact penalization. We can preserve the notion of optimal values for local optimal values where X is replaced by $X \cap B_\epsilon(x^*)$ for some local optimal solution x^* of (1.20), see Branda (2013b) for details.

Theorem 1.1.5. (Branda (2013b))
We consider the two problems (1.18) and (1.19) and assume:

(i) $g_i(x', \xi^s) \leq 0,\ i = 1, \ldots, k$ for all $s = 1, \ldots, S$ for at least one $x' \in X$.

(ii) the corresponding unperturbed problem (1.20) is calm at its local optimal solutions modulus $\tilde{N}$ and radius $\epsilon > 0$.

For arbitrary $\gamma \in (0, 1)$, $N > 0$ and $\varepsilon \in (0, 1)$ put

$$\begin{aligned}
\varepsilon(x) &= \sum_{s=1}^{S} p_s \mathbb{I}\left(\Phi(x, \xi^s) > 0\right), \\
\alpha_N(x) &= N \sum_{s=1}^{S} p_s \Phi(x, \xi^s), \\
\beta_\varepsilon(x) &= \frac{1}{\varepsilon^\gamma} \sum_{s=1}^{S} p_s \Phi(x, \xi^s).
\end{aligned}$$

Then for any prescribed $\varepsilon \in (0,1)$ there always exists $N \leq \tilde{N}$ large enough so that minimization (1.19) generates optimal solutions x_N which also satisfy the probabilistic constraint with the given ε.

Moreover, bounds on the local optimal value ψ_ε of (1.27) based on the local optimal value φ_N of (1.19) and vice versa can be constructed:

$$\begin{array}{rcccl} \varphi_{1/\varepsilon^\gamma(x_N)} - \beta_{\varepsilon(x_N)}(x^{CCP}_{\varepsilon(x_N)}) & \leq & \psi_{\varepsilon(N)} & \leq & \varphi_N - \alpha_N(x_N), \\ \psi_{\varepsilon(x_N)} + \alpha_N(x_N) & \leq & \varphi_N & \leq & \psi_{N^{-1/\gamma}} + \beta_{N^{-1/\gamma}}(x^{CCP}_{N^{-1/\gamma}}), \end{array} \tag{1.22}$$

with

$$\lim_{N \to \tilde{N}_-} \alpha_N(x_N) = \lim_{N \to \tilde{N}_-} \varepsilon(x_N) = \lim_{\varepsilon \to \tilde{\varepsilon}_+} \lambda_\varepsilon(x^{CCP}_\varepsilon) = 0,$$

for any sequences of the optimal solutions x_N and x^{CCP}_ε, where $\tilde{\varepsilon} < \min_s p_s$.

1.1.4 Fixed interval scheduling with random delays

Probabilistic functions in optimization problems were also investigated by Branda et al. (2016); Branda and Hájek (2017) who dealt with fixed interval scheduling problems under uncertain processing intervals. The probabilistic functions were employed to express the probability that the obtained schedule remains feasible after observing the random delays. This probability is then maximized in the operational problem or bounded by a prescribed level in the tactical problem. The probability maximization leads to difficult problems in general, see, e.g., Fábián et al. (2018), however we will propose a trackable reformulation for FIS using a generalized robust coloring problem.

Let $\mathcal{C}$ be the set of machines and $\mathcal{J}$ denotes the set of jobs. The starting times s_j are known, but the finishing times are random $f_j(\xi)$. We consider the following random real finishing time which consists of a prescribed completion time f^0_j and a random nonnegative delay $D_j(\xi)$, i.e.

$$f_j(\xi) = f^0_j + D_j(\xi).$$

We assume that the distribution P of the random delays is known with a support Ξ and a probability mass in zero, i.e. $D_j(\xi) = 0$ has a positive probability meaning that the job can be finished in time. We will use the argument ξ to mark that some parameter is random, e.g. $D_j(\xi)$, $f_j(\xi)$. Binary decision variable x_{jc}, $j \in \mathcal{J}$, $c \in \mathcal{C}$ is used to assign job j to machine c, i.e. it is equal to one if the job is assigned to the machine and it is zero otherwise. We use the convention that at time $f_j(\xi)$, when job j finishes, a new job j' with starting time $s_{j'} \geq f_j(\xi)$ can be assigned to the same machine. The following constraints ensure that in each time at most one job is processed by each machine

$$\sum_{j:\ s_j \leq t < f_j(\xi)} x_{jc} \leq 1,\ t \in \mathcal{T}, c \in \mathcal{C}. \tag{1.23}$$

However, the constraints need not be verified at all times, we can consider them only at the known starting times, i.e. $\mathcal{T}$ can be replaced by $\hat{\mathcal{T}} = \{s_1, \ldots, s_{|\mathcal{J}|}\}$, cf. Kroon et al. (1995).

The constraints (1.23) involve random parameters and we would like to find a job assignment with the highest probability that it remains feasible. This leads to the following stochastic programming problem:

$$
\begin{aligned}
\max_x P\Bigg(\sum_{j:\ s_j \le t < f_j(\xi)} x_{jc} &\le 1,\ t \in \hat{\mathcal{T}}, c \in \mathcal{C} \Bigg) \\
\text{s.t.} \sum_{j:\ s_j \le t < f_j^0} x_{jc} &\le 1,\ c \in \mathcal{C},\ t \in \hat{\mathcal{T}}, \\
\sum_{c \in \mathcal{C}} x_{jc} &= 1,\ j \in \mathcal{J}, \\
x_{jc} &\in \{0,1\},\ c \in \mathcal{C}, j \in \mathcal{J}.
\end{aligned}
\tag{1.24}
$$

We maximize the probability that in each time a machine processes one job only, i.e., the probability that the job assignment is feasible with respect to the random delays. The constraints ensure that there is at most one job assigned to a machine at each time with respect to the prescribed job processing times and a job is assigned to exactly one machine. The decision variables x_{jc} are binary.

Below, we review the reformulation of (1.24) as a robust coloring problem proposed by Branda et al. (2016). We work with an interval graph for FIS defined as follows. Let $\mathcal{C}$ be the available machines (colors) and $\mathcal{J}$ denotes the set of jobs which represent the vertices. The starting times s_j and the prescribed finishing times f_j^0 are fixed and known. Then, the set of edges E contains all pairs of jobs $\{j, j'\}$ which processing times overlap for sure, i.e. $s_j \le s_{j'} < f_j^0$. If the distribution of delays has an unbounded support, then the complementary edge set $\overline{E}$ contains in general all pairs of vertices $\{j, j'\}$ such that $f_j^0 \le s_{j'}$, i.e. when the prescribed completion time f_j^0 of job j comes earlier than starting time $s_{j'}$ of job j'. Otherwise, the set can be reduced to include only those pairs of jobs, which processing times can overlap with a positive probability lower than one. Obviously $E \cap \overline{E} = \emptyset$.

We will consider a joint distribution of delays which follows an Archimedean copula. Let $F_j(x) = P(D_j(\xi) \le x)$ be the univariate cumulative distribution functions, and $\psi : [0,1] \longrightarrow [0,\infty]$ be a generator of an Archimedean copula, i.e. a continuous strictly decreasing function satisfying $\psi(1) = 0$, $\lim_{x \to 0_+} \psi(x) = \infty$, see McNeil and Nešlehová (2009), Hering and Stadtmüller (2012) for details. We assume that the joint probability distribution of random delays follows a n-dimensional Archimedean copula function with a generator ψ, i.e.

$$
P(D_1(\xi) \le x_1, \ldots, D_n(\xi) \le x_n) = \psi^{-1}\left(\sum_{j=1}^{n} \psi(F_j(x_j)) \right). \tag{1.25}
$$

Yanez and Ramirez (2003) investigated the robust coloring problem which assigns penalties to the edges connecting the nodes with the same color. Branda et al. (2016) showed that such formulation is not sufficient to solve the problem (1.24). Thus, they proposed a generalized robust coloring reformulation.

Proposition 1.1.6. (Branda et al. (2016))
Let the multivariate distribution of delays be represented by an Archimedean copula with generator ψ. We can formulate a generalized robust coloring problem

where we maximize the schedule reliability as follows

$$
\begin{aligned}
\min_{x,y,z} \sum_{\{j,j'\}\in\overline{E}} q_{jj'} z_{jj'} & & & \\
\text{s.t.} \sum_{c\in\mathcal{C}} x_{jc} &=& 1,\ j\in\mathcal{J}, & \\
x_{jc}+x_{j'c} &\le& 1,\ \{j,j'\}\in E, c\in\mathcal{C}, & \\
x_{jc}+x_{j'c} &\le& 1+y_{jj'},\ \{j,j'\}\in\overline{E}, c\in\mathcal{C}, & \qquad (1.26)\\
y_{jj'} + \sum_{k:\ \{j,k\}\in\overline{E}\ \&\ s_k\ge f^0_{j'}} z_{jk} &\le& 1,\ \{j,j'\}\in\overline{E}, & \\
\sum_{k:\{j,k\}\in\overline{E}} y_{jk} &\le& |\mathcal{J}|\cdot \sum_{k:\{j,k\}\in\overline{E}} z_{jk},\ j\in\mathcal{J}, & \\
x_{jc} &\in& \{0,1\},\ c\in\mathcal{C}, j\in\mathcal{J}, &
\end{aligned}
$$

where the costs are defined as

$$q_{jj'} = \psi(P(D_j(\xi)\le s_{j'} - f^0_j)), \{j,j'\}\in\overline{E}.$$

The probability can be obtained by applying ψ^{-1} to the optimal value.

The binary variables x_{jc} are used to assign job j to machine c, $y_{jj'}$ express that jobs $j, j' \in \mathcal{J}$ share the same color, whereas $z_{jj'}$ are used to identify the successor j' of job j. The objective is to minimize the penalty assigned to successors. The first constraints forbid the same color for vertices connected by hard edges, whereas the second constraints enable the same color for vertices connected by soft edges, where the corresponding variables $y_{jj'}$ have to be equal to one. The next two constraints are used to identify the successor and to assign the right penalties, see Branda et al. (2016) for details. This reformulation was used to solve smaller simulated instances to optimality, however a tabu search algorithm was suggested for solving larger instances. The tabu search is a strong method for combinatorial problems, see, e.g., Michalopoulos et al. (2015) for other application. Branda et al. (2015) proposed a two-stage stochastic programming formulation which aims at maximizing the reward for processing the selected jobs and at the same time minimizing the costs for outsourcing additional machines. Branda and Hájek (2017) observed that a network-based formulation enables us to optimally solve a larger instance by the mixed-integer programming solver IBM Cplex. Moreover, this formulation enabled the consideration of heterogeneous machines and job delays dependent on a selected machine.

1.2 Sample approximation techniques for mixed-integer stochastic problems

Branda (2012b, 2014b) contributed to the theory of sample approximation techniques and elaborated two general classes of stochastic mixed-integer problems. These problems are motivated by many real-life stochastic optimization problems where continuous and integer variables appear together. However, earlier papers dealt with these case separately, cf. de Mello and Bayraksan (2014); Wang and Ahmed (2008). The main goal is to estimate the speed of convergence of the sample approximation to the true problem for the set of feasible solutions, optimal values and optimal solutions. Based on this speed formula, an estimate for the sample size which is necessary to reach a prescribed reliability of the estimates can be derived. Recent research in this area focuses on more advances sampling techniques, e.g., importance sampling, cf. Barrera et al. (2016), and quasi-Monte Carlo, see Heitsch et al. (2016).

1.2.1 Problems with several chance constraints

The following approach was suggested by Branda (2012b). Let $X \subseteq \mathbb{R}^n$, $Y \subseteq \mathbb{Z}^{n'}$ and

$$Z = \{(x, y) \in X \times Y : \ h_1(x, y) \le 0, \dots, h_k(x, y) \le 0\}$$

be the deterministic mixed-integer part of the set of feasible solutions with $h_j(x, y) : \mathbb{R}^n \times \mathbb{R}^{n'} \to \mathbb{R}$. Let ξ be a random vector on the probability space $(\Xi, \mathcal{A}, P)$, $g_{ji}(x, y, \xi), i = 0, \dots, k_j, j = 1, \dots, m$, be real functions on $\mathbb{R}^n \times \mathbb{R}^{n'} \times \mathbb{R}^{n''}$ measurable in ξ for all $x \in X$ and $y \in Y$. We assume that the objective function $f(x, y) : \mathbb{R}^n \times \mathbb{R}^{n'} \to \mathbb{R}$ does not depend on the random vector. Denote $G_j(x, y, \xi) = \max\{g_{j1}(x, y, \xi), \dots, g_{jk_j}(x, y, \xi)\}$. We can define the probability functions

$$\phi_j(x, y) = P(G_j(x, y, \xi) \le 0) = P\Big(\{\xi : \ G_j(x, y, \xi) \le 0\}\Big).$$

Then, the chance constrained problem with several chance constraints can be formulated as follows:

$$\begin{array}{rcl} \min_{(x,y)\in Z} f(x, y), & & \\ \text{s.t.} & & \\ \phi_1(x, y) & \ge & 1 - \varepsilon_1, \\ & \vdots & \\ \phi_m(x, y) & \ge & 1 - \varepsilon_m, \end{array} \tag{1.27}$$

where $\epsilon = (\varepsilon_1, \dots, \varepsilon_m)$, with the levels $\varepsilon_j \in (0, 1)$. We denote by Z_ε the set of feasible solutions. The formulation covers the joint ($k_1 > 1$ and $m = 1$) as well as the separate ($k_j = 1$ and $m > 1$) chance constrained problems as special cases. We will refer to the problem (1.27) as the original problem.

Let $\xi^1, \dots, \xi^S$ be an independent Monte Carlo sample of the random vector ξ. Then, the sample version of the probability function ϕ_j is defined to be

$$\hat{\phi}_j^S(x, y) = S^{-1} \sum_{s=1}^{S} \mathbb{I}\Big(G_j(x, y, \xi^s) \le 0\Big), \tag{1.28}$$

where $\mathbb{I}(\cdot)$ denotes the indicator function which is equal to one if the condition is fulfilled, and to zero otherwise. Finally, the sample version of the chance constrained problem with several chance constraints (1.27) is defined as

$$\begin{array}{rcl} \min_{(x,y)\in Z} f(x,y), & & \\ \text{s.t.} & & \\ \hat{\phi}_1^S(x,y) & \geq & 1-\gamma_1, \\ \vdots & & \\ \hat{\phi}_m^S(x,y) & \geq & 1-\gamma_m, \end{array} \tag{1.29}$$

where the levels $\gamma_j \in (0,1)$ are allowed to be different from the original levels ε_j. We denote Z_γ^S the set of feasible solutions of the sample approximated problem.

It is necessary to consider the constraints which are satisfied strictly, i.e. with some deviation τ:

$$\begin{aligned} Z_{\gamma_j,\tau}^S &= \{(x,y)\in Z : \frac{1}{S}\sum_{s=1}^{S}\mathbb{I}(G_j(x,y,\xi^s)+\tau \leq 0) \geq 1-\gamma_j\}, \\ Z_{\gamma,\tau}^S &= \bigcap_{j=1}^{m} X_{\gamma_j,\tau}^S. \end{aligned}$$

The following theorem by Branda (2012b) provides an exponential rate of convergence for the probability that the relaxed sample-approximated set of feasible solutions is contained in the set of feasible solutions of the original problem. It is assumed that the integral feasibility set is finite, whereas the continuous part is uniformly bounded. Lipschitz continuity of the random constraints is required only with respect to the continuous variables.

Theorem 1.2.1. (Branda (2012b))
Let

1. $\gamma_j < \varepsilon_j$, *i.e., that the levels of the sample approximated problem are more restrictive,*

2. $Y \subseteq \mathbb{R}^{n'}$ *be finite,*

3. $X(y) \subseteq \mathbb{R}^n$ *be uniformly bounded for all* $y \in Y$,
i.e. $D = \sup_{y\in Y}\sup\{\|x-x'\|_\infty : x,x' \in X(y)\}$ *be a finite diameter,*

4. *functions* $G_j(x,y,\xi)$ *be Lipschitz continuous in the real variable* x, *i.e. for arbitrary* $y \in Y$ *and* $\xi \in \Xi$

$$|G_j(x,y,\xi) - G_j(x',y,\xi)| \leq L_j \|x-x'\|_\infty, \ \forall x,x' \in X(y),$$

for some $L_j > 0$.

Then the probability that the relaxed sample-approximated set of feasible solutions is contained in the set of feasible solutions of the original problem increases exponentially fast with increasing sample size, and it holds that

$$\begin{aligned} P(Z_\gamma^{\tau,S} \subseteq Z_\epsilon) &\geq \\ &\geq 1 - m\left\lceil \frac{1}{\lambda_{min}} \right\rceil \left\lceil \frac{2L_{\max}D}{\tau} \right\rceil^n |Y|^2 \exp\Big\{-2S\min_j(\varepsilon_j-\gamma_j-\lambda_j)^2\Big\}, \end{aligned}$$

where $L_{max} = \max_j L_j$ and $\lambda_{min} = \min_j \lambda_j$. It is possible to estimate the sample size S such that the feasible solutions of the relaxed sample-approximated problems are feasible for the original problem with a high probability $1-\delta$, i.e.

$$S \geq \frac{1}{2\min_j(\varepsilon_j - \gamma_j - \lambda_j)^2}\left(\ln\frac{m|Y|^2}{\delta} + \ln\left\lceil\frac{1}{\lambda_{min}}\right\rceil + n\ln\left\lceil\frac{2L_{max}D}{\tau}\right\rceil\right). \quad (1.30)$$

Branda (2012b) applied this approach to a stochastic vehicle routing problem with random traveling times and random demand and derived the estimates of the sample size.

1.2.2 Problems with several expected value constraints

Now, we turn our attention to the problems with several expected value constraints, cf. Branda (2014b). Sample approximation techniques were successfully applied to stochastic programs with expectation type objective, see Dai et al. (2000); Shapiro (2003); de Mello (2008); Xu (2010), and expected value constraints, see Branda (2012c), Wang and Ahmed (2008). However, these results were limited to the finite and real bounded set of feasible solution. In this part, we will extend the results for the problem with several expected value constraints and mixed-integer set of feasible solutions. Moreover, our estimates are valid for particular types of non-iid sampling.

The estimate is based on Cramér's large deviation theory, see Dembo and Zeitouni (1998). It remains valid also for non-iid sampling if the Gärtner-Ellis theorem is used and a condition on the convergence of logarithmic moment generating functions is added, c.f. (Dembo and Zeitouni, 1998, Theorem 2.3.6). This condition is trivially fulfilled for iid samples. Moreover, it can be verified for finite state Markov chains. Instead of the traditional Lipschitz continuity, which was necessary to obtain the rate of convergence for the continuous problems by Wang and Ahmed (2008), we will employ more general calmness condition.

Let $X \subseteq \mathbb{R}^n$, $Y \subseteq \mathbb{Z}^{n'}$ and

$$Z = \{(x,y) \in X \times Y : \; h_1(x,y) \leq 0, \ldots, h_k(x,y) \leq 0\}$$

be the deterministic mixed-integer part of the set of feasible solutions with $h_j(x,y) : \mathbb{R}^n \times \mathbb{R}^{n'} \to \mathbb{R}$. Let ξ be a random vector on the probability space $(\Xi, \mathcal{A}, P)$, $g_j(x,y,\xi), j = 1,\ldots,m$, be real functions on $\mathbb{R}^n \times \mathbb{R}^{n'} \times \mathbb{R}^{n''}$ measurable in ξ for all $x \in X$ and $y \in Y$. We assume that the objective function $f(x,y) : \mathbb{R}^n \times \mathbb{R}^{n'} \to \mathbb{R}$ does not depend on the random vector. We denote the set of feasible solutions by

$$Z_L = \left\{(x,y) \in Z : \; g_j(x,y) := \mathbb{E}[g_j(x,y,\xi)] \leq L_j, \; j = 1,\ldots,m\right\}$$

for some prescribed upper bounds $L_j \in \mathbb{R}$, $L = (L_1,\ldots,L_m)'$. We assume that the levels are chosen in such a way that the set of feasible solutions is nonempty and that the expectations are finite for all $(x,y) \in Z$. Then, the stochastic programming problem with the expected value constraints can be formulated as

$$\min_{(x,y)\in Z_L} f(x,y). \quad (1.31)$$

Let $\xi^1, \dots, \xi^S$ be a Monte-Carlo sample of the underlying distribution of the random vector ξ. We denote the set of feasible solutions of the sample-approximated problem as

$$Z_L^S = \Big\{(x,y) \in Z: \ g_j^S(x,y) := \frac{1}{S}\sum_{s=1}^{S} g_j(x,y,\xi^s) \le L_j, \ j = 1,\dots,m\Big\}.$$

The sample version of the problem with several expected value constraints (1.31) is defined as

$$\min_{(x,y)\in Z_L^S} f(x,y). \tag{1.32}$$

where the levels $L_j \in \mathbb{R}$ are allowed to be different from the original levels.

Let I_y denote the large deviation rate functions, i.e. the Fenchel dual to the logarithm of the finite moment generating function of the difference $g_j(x,y,\xi) - g_j(x,y)$ which is defined as

$$I_{jxy}(\tau) = \sup_{t\in\mathbb{R}} \Big\{t\tau - \Psi_{jxy}(t)\Big\},$$

where

$$\Psi_{jxy}(t) = \ln \mathbb{E}\Big[e^{t(g_j(x,y,\xi)-g_j(x,y))}\Big], \ g_j(x,y) = \mathbb{E}[g_j(x,y,\xi)].$$

The theorem by Branda (2014b) verifies the exponential rate of convergence for the set of feasible solutions. Based on this rate, we can derive a sample estimate to ensure that the original feasibility set is contained in the relaxed sample-approximated feasibility set with a high probability. We will consider non-iid sampling and employ Hölder-calmness (H-calmness) instead of traditional Lipschitz continuity. The assumptions impose finiteness of the integral part and boundeness of the continuous one. H-calmness of the random constraints is required only with respect to the continuous variables. Moreover, random modulus is enabled. The last important assumptions require the convergence of the moment generating functions.

Theorem 1.2.2. (Branda (2014b))
Let

(i) $Y \subseteq \mathbb{R}^{n'}$ *be finite, and* $X \subseteq \mathbb{R}^n$ *be bounded,*
i.e. $D = \sup\{\|x - x'\|_\infty : \ x, x' \in X\}$ *be a finite diameter,*

(ii) $g_j(x,y,\xi)$ *be uniformly H-calm in* $x \in X$ *for each* $y \in Y$*, moduli* $M_j(\xi) > 0$*, and order* $\gamma_j > 0$*:*

$$|g_j(x,y,\xi) - g_j(x',y,\xi)| \le M_j(\xi)\, \|x - x'\|^{\gamma_j}, \ \forall x, x' \in X, \ \forall \xi \in \Xi, \ \forall j,$$

with $M_j = \mathbb{E}[M_j(\xi)] < \infty, \forall j$*,*

(iii) the logarithmic moment generating functions $\Psi_{jxy}(t)$ *of* $g_j(x,y,\xi) - g_j(x,y)$ *be finite around 0 and*

$$\Psi_{jxy}(t) = \lim_{S\to\infty} \frac{\Psi_{jxy}^S(St)}{S},$$

for all $t \in \mathbb{R}$ *and for all* $(x,y) \in Z$*.*

(iv) *the logarithmic moment generating function* $\Psi_{M_j}(t)$ *of* $M_j(\xi) - M_j$ *be finite around 0 and*

$$\Psi_{M_j}(t) = \lim_{S\to\infty} \frac{\Psi^S_{M_j}(St)}{S},$$

for all $t \in \mathbb{R}$.

Then for $\tau_j > 0$ *small,* $\tau = (\tau_1, \dots, \tau_m)$*,*

(a) *the probability that the set of feasible solutions is contained in the relaxed sample-approximated set of feasible solutions increases exponentially with increasing sample size, and it holds*

$$P(Z_L \subseteq Z^S_{L+\tau}) \geq 1 - m\left(1 + |Y|\frac{D^n}{\upsilon^n}\right)\exp\Big\{-Sd(\tau)\Big\},$$

where

$$\begin{aligned}
\sigma^2_{jxy} &= Var[g_j(x,y,\xi) - g_j(x,y)],\\
\sigma^2_{M_j} &= Var[M_j(\xi) - M_j],\\
\upsilon &= \max_j \left(\frac{\tau_j}{4M_j + \tau_j}\right)^{1/\gamma_j},\\
d(\tau) &= \min\left\{\min_{j,x,y} \frac{\tau_j^2}{8\sigma^2_{jxy}}, \min_j \frac{\tau_j^2}{8\sigma^2_{M_j}}\right\},
\end{aligned}$$

and the minimum is taken over $(x,y) \in Z_\upsilon$ *and* $j \in \{1,\dots,m\}$.

(b) *we can get an estimate for the sample size which is necessary to ensure that the original feasibility set is contained in the relaxed sample-approximated feasibility set with a high probability, equal to* $1-\delta$, $\delta \in (0,1)$*:*

$$S \geq \frac{1}{d(\tau)}\left(\ln\frac{m}{\delta} + \ln\left(1 + |Y|\frac{D^n}{\upsilon^n}\right)\right).$$

Branda (2014b) applied the estimates to the investment problems with several Conditional Value at Risk constraints and with the set of feasible portfolio weights including transaction costs modelled using integer variables.

1.3 Diversification-consistent data envelopment analysis models

This section is focused on special class of Data Envelopment Analysis (DEA) models, called diversification-consistent DEA (DC DEA), which are motivated by efficiency analysis of investment opportunities available on financial markets, e.g., assets, mutual funds, portfolios. Compared with the traditional DEA models introduced by Charnes et al. (1978), the envelops of diversification-consistent models are not linear (polyhedral) which is caused by the risk measures that are used as the inputs of the DC DEA models, see below. These new models combine several approaches from mathematical optimization and finance. In particular, several risk measures can be used at the same time leading to an approach called "risk-shaping". It can be shown that the DC DEA models are in compliance with the Pareto–Koopmans efficiency and multiobjective optimization principles. Finally, under particular choice of the inputs and outputs, the DEA efficiency corresponds to the stochastic dominance efficiency which is based on the utility theory.

Since the seminal work of Murthi et al. (1997), DEA models found their place in the efficiency analysis of various investment opportunities. First two moments of the random returns were considered as the inputs and outputs by Briec et al. (2004), and the third moment (skewness) was added by Joro and Na (2006); Briec et al. (2007) and further elaborated by Kerstens et al. (2011). Dokov et al. (2017) investigated mean-variance-skewness-kurtosis efficiency of portfolios. A general class of diversification-consistent models was proposed by Lamb and Tee (2012). However, they considered positive parts of coherent measures as the inputs which can be quite limiting, since the risk measures can take also negative values. Branda (2013a) suggested to use general deviation measures, which are always nonnegative, as the inputs of the models with diversification. Branda and Kopa (2014) provided DEA models equivalent to the second order stochastic dominance tests. In this section, we mainly focus on the DC DEA models based on a directional distance measure proposed by Branda (2015b) who considered coherent risk measures as the inputs and return measures as the outputs. It was even shown, that under particular choice of the inputs and outputs, the proposed models are equivalent to the stochastic dominance tests. Moreover, these models project inefficient investment opportunities to the efficient frontier, which can be used by the investors to revise their inefficient portfolios. The traditional DEA models were used to approximate the DC DEA efficient frontier and to assess performance of portfolios by Liu et al. (2015). Branda (2015a, 2016b) discussed models with Value at Risk measure which does not belong to any above mentioned classes but still is a popular risk measure. These models lead to chance constrained problems which were discussed in Section 1.1. Branda and Kopa (2016) proposed the DEA models which are consistent with the higher order stochastic dominance tests. Branda (2015b), Branda (2016b) performed large numerical experiments where empirical performance of the proposed DC DEA models was investigated.

1.3.1 Preliminaries

Let $\mathcal{X}$ be a set of available investment opportunities. We prefer higher values to lower, i.e. we deal with profits, returns, wealth etc. We will identify each investment opportunity by a random variable which represents its return. We consider n assets and denote by R_i the return of i-th asset, which is a real random variable defined on a probability space. The following choice of the set of investment opportunities will be used:

$$\mathcal{X} = \left\{ \sum_{i=1}^{n} R_i x_i : \sum_{i=1}^{n} x_i = 1, x_i \geq 0 \right\}, \tag{1.33}$$

which enables full diversification of our portfolio across all assets. Other choices of the set are also possible, e.g. with limited number of assets, allowing short sales, borrowing or including proportional and fixed transaction costs, margin requirements, cf. Branda and Kopa (2014); Branda (2014a).

Branda (2015b) considered the models where coherent risk measures are used as the inputs. The coherent risk measures, cf. Artzner et al. (1999), are functionals on $\mathcal{L}_p(\Omega)$ for some $p \in \{1, \dots, \infty\}$, usually $p = 1$ or $p = 2$, that satisfy

(R1) translation equivariance: $\mathcal{R}(X + C) = \mathcal{R}(X) - C$ for all $X \in \mathcal{L}_p(\Omega)$ and constants C,

(R2) positive homogenity: $\mathcal{R}(0) = 0$, and $\mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$ for all $X \in \mathcal{L}_p(\Omega)$ and all $\lambda \geq 0$,

(R3) subadditivity: $\mathcal{R}(X_1 + X_2) \leq \mathcal{R}(X_1) + \mathcal{R}(X_2)$ for all $X_1, X_2 \in \mathcal{L}_p(\Omega)$,

(R4) monotonicity: $\mathcal{R}(X_1) \leq \mathcal{R}(X_2)$ when $X_1 \geq X_2$, $X_1, X_2 \in \mathcal{L}_p(\Omega)$.

The space $\mathcal{L}_p(\Omega)$ of random variables with finite p-th moment is selected to ensure finiteness of a considered measure. Note that the axioms (R2) and (R3) imply convexity of the measures. We can employ several return measures defined as minus coherent risk measures, i.e., $\mathcal{E}(X) = -\mathcal{R}(X)$, which quantify more-likely profitable outcomes by higher real values.

Conditional Value at Risk (CVaR) for a confidence level $\alpha \in (0, 1)$ can be defined as the "Expected value of $(1-\alpha)$*100% worst losses". However, to preserve the above defined coherency for a general distribution of losses, Rockafellar and Uryasev (2002) proposed a definition based on so called $\alpha-$tail distribution. For our purposes, it is more convenient to use the equivalent definition using the minimization formula because it can be used directly in the optimization problems, cf. Rockafellar and Uryasev (2002):

$$CVaR_\alpha(X) = \min_{y \in \mathbb{R}} y + \frac{1}{1-\alpha} \mathbb{E}[\max\{(-X - y), 0\}].$$

The above minimum is attained at any $(1-\alpha)$-th quantile of the random variable $-X$.

1.3.2 Model definition and properties

We employ the traditional definition of DEA efficiency based on the optimal values. It can be shown that the optimal values of the studied models (DEA scores) belong to the interval $(0, 1]$. We can sort the investigated investment opportunities according to these scores, which correspond to relative distances to a set of efficient opportunities. The opportunities with optimal values equal to one are classified as efficient, the other as inefficient.

Branda (2015b) proposed models where several coherent risk measures $\mathcal{R}_k$ are used as the inputs and several return measures $\mathcal{E}_j$ serve as the outputs. Since both return and risk measures can take arbitrary real values, Branda (2015b) generalized the directional distance measures for the DEA models with diversification. First, we discuss the choice of the directions. Let $X_0 \in \mathcal{X}$ be the investigated opportunity. If all input and output data are positive, the directions can be set to

$$e_j(X_0) = \mathcal{E}_j(X_0),\ d_k(X_0) = \mathcal{R}_k(X_0),$$

which corresponds to input-output oriented radial model, see Branda (2013a,c). However, if both the inputs and outputs can take negative and positive values, then the following choices of the directions are reasonable:

$$e_j(X_0) = \max_{X \in \mathcal{X}} \mathcal{E}_j(X) - \mathcal{E}_j(X_0),\ d_k(X_0) = \mathcal{R}_k(X_0) - \min_{X \in \mathcal{X}} \mathcal{R}_k(X), \qquad (1.34)$$

see Portela et al. (2004) for an analogy in the traditional DEA models. These directions are always nonnegative and will be used in our models, although the formulations enable us to employ also other directions. Note that the choice of the directions was investigated by Kerstens et al. (2012) for the mean–variance efficiency models.

We will use $K \geq 1$ risk measures as the inputs and $J \geq 1$ return measures as the outputs. We will focus on the strongest model provided by Branda (2015b) (referred as model 3) which enables us to measure necessary relative improvements to reach the efficient frontier with respect to each input and output separately:

$$\begin{aligned}
\theta_3(X_0) = \min_{\theta_k, \varphi_j, x_i} \frac{1 - \frac{1}{K}\sum_{k=1}^{K} \theta_k}{1 + \frac{1}{J}\sum_{j=1}^{J} \varphi_j} & \\
\text{s.t.} & \qquad (1.35)\\
\mathcal{E}_j\left(\sum_{i=1}^{n} R_i x_i\right) &\geq \mathcal{E}_j(X_0) + \varphi_j \cdot e_j(X_0),\ j = 1, \dots, J,\\
\mathcal{R}_k\left(\sum_{i=1}^{n} R_i x_i\right) &\leq \mathcal{R}_k(X_0) - \theta_k \cdot d_k(X_0),\ k = 1, \dots, K,\\
\sum_{i=1}^{n} x_i &= 1,\ x_i \geq 0,\ \varphi_j \geq 0,\ \theta_k \geq 0.
\end{aligned}$$

For directions (1.34), if $e_j(X_0) = 0$ for some j, it means that the output reaches its maximal value. This necessarily means that no improvement in this output is possible and thus we set $\varphi_j^* = 0$. Similarly, if $d_k(X_0) = 0$ for some k, then $\theta_k^* = 0$.

Definition 1.3.1. *We say that $X_0 \in \mathcal{X}$ is DEA efficient with respect to the set $\mathcal{X}$ if the optimal value of the DEA model (DEA score) is equal to 1. Otherwise, X_0 is inefficient and the optimal value measures the inefficiency.*

We will denote by $\theta_3(X_0)$ the optimal value (DEA score) for a benchmark X_0 and by Ψ_3 the sets of efficient investment opportunities with respect to the set $\mathcal{X}$, i.e.:

$$\Psi_3 \;=\; \{X \in \mathcal{X}:\; \theta_3(X) = 1\}.$$

As it was shown by Branda (2015b), the proposed DEA models identify various efficient points in the production possibility set which can be defined as follows:

$$\mathcal{PPS} = \{(-\mathcal{R}_1(X), \ldots, -\mathcal{R}_K(X), \mathcal{E}_1(X), \ldots, \mathcal{E}_J(X)):\; X \in \mathcal{X}\}.$$

Risk measures are multiplied by -1 to ensure consistency with ordering defined below. This set is not based on any axioms as is often stated for traditional DEA and it contains only those inputs and outputs, which can be obtained using the set of available investment opportunities. Partial ordering known from multiobjective optimization will be closely related to our models through the production possibility set as will be shown later.

Definition 1.3.2. *Let $v, z \in \mathbb{R}^n$. We say that z partially weakly dominates v with respect to an index set $S \subseteq \{1, \ldots, n\}$, denoted by $v \prec_{\mathrm{pw}(S)} z$, if $v_i \leq z_i$ for all $i \in \{1, \ldots, n\}$ and there exists at least one $\tilde{i} \in S$ for which $v_{\tilde{i}} < z_{\tilde{i}}$.*

Definition 1.3.3. *An investment opportunity $X_0 \in \mathcal{X}$ is strongly Pareto–Koopmans efficient if there is no $\tilde{X} \in \mathcal{X}$ and corresponding vector v from the production possibility set for which*

$$(-\mathcal{R}_1(X_0), \ldots, -\mathcal{R}_K(X_0), \mathcal{E}_1(X_0), \ldots, \mathcal{E}_J(X_0)) \prec_{\mathrm{pw}(\{1,\ldots,K+J\})} v$$

We can formulate conditions to fully characterize the set of efficient investment opportunities with respect to the model (1.35).

Proposition 1.3.1. (Branda (2015b))
$X_0 \in \Psi_3$:

- *if and only if X_0 is strongly Pareto–Koopmans efficient, or equivalently*
- *if and only if there is no $X \in \mathcal{X}$ for which $\mathcal{E}_j(X) \geq \mathcal{E}_j(X_0)$ for all j and $\mathcal{R}_k(X) \leq \mathcal{R}_k(X_0)$ for all k with at least one inequality is strict.*

The original formulation (1.35) is a nonconvex optimization problem. However, using the following substitution, we can arrive at a convex problem. If we set $1/t = 1 + \frac{1}{J}\sum_{j=1}^{J} \varphi_j$, $\tilde{x}_i = t x_i$, $\tilde{\theta}_k = t\theta_k$, and $\tilde{\varphi}_j = t\varphi_j$, we can reformulate (1.35) as

$$\begin{aligned}
\theta_3(X_0) = \min_{\tilde{\theta}_k, \tilde{\varphi}_j, t, \tilde{x}_i} \; & t - \frac{1}{K}\sum_{k=1}^{K} \tilde{\theta}_k \\
& \text{s.t.} \\
t + \frac{1}{J}\sum_{j=1}^{J} \tilde{\varphi}_j \;&=\; 1, \\
\mathcal{E}_j\left(\sum_{i=1}^{n} R_i \tilde{x}_i\right) \;&\geq\; t \cdot \mathcal{E}_j(X_0) + \tilde{\varphi}_j \cdot e_j(X_0),\; j = 1, \ldots, J, \\
\mathcal{R}_k\left(\sum_{i=1}^{n} R_i \tilde{x}_i\right) \;&\leq\; t \cdot \mathcal{R}_k(X_0) - \tilde{\theta}_k \cdot d_k(X_0),\; k = 1, \ldots, K, \\
\sum_{i=1}^{n} \tilde{x}_i \;&=\; t,\; \tilde{x}_i \geq 0,\; \tilde{\varphi}_j \geq 0,\; \tilde{\theta}_k \geq 0,\; t \geq 0.
\end{aligned} \tag{1.36}$$

Similar arguments as for the previous model can be used to show that $\theta_3(X_0) \in [0,1]$ and $t=0$ is never optimal, thus the inverse transformation can be used to obtain the optimal solution of (1.35).

1.3.3 Efficiency with respect to the second-order stochastic dominance

In this section, we will compare the diversification-consistent DEA efficiency with the stochastic dominance efficiency. We show that by proper choice of the inputs and outputs we can obtain a DEA model that is able to identify efficient investment opportunities with respect to the second-order stochastic dominance (SSD). Several researchers were trying to obtain such relation, but usually their DEA models provided a necessary condition only, see, e.g., Lozano and Gutiérrez (2008).

Let $F_X(\eta)$ denote the cumulative probability distribution function of an investment opportunity $X \in \mathcal{X}$, i.e.

$$F_X(t) = P(X \leq t).$$

The twice cumulative probability distribution function is defined as:

$$F_X^{(2)}(t) \quad = \quad \int_{-\infty}^{t} F_X(\eta)\,\mathrm{d}\eta,$$

which is well defined for $X \in \mathcal{L}_1(\Omega)$. Then the investment opportunity X dominates $\tilde{X}$ with respect to the second-order stochastic dominance (SSD) if and only if

$$F_X^{(2)}(t) \leq F_{\tilde{X}}^{(2)}(t),\ \forall t \in \mathbb{R}, \tag{1.37}$$

which we will denote $\tilde{X} \preceq_{SSD} X$. The relation is strict, denoted by $\tilde{X} \prec_{SSD} X$, if the inequality in (1.37) is strict for at least one $t \in \mathbb{R}$. An alternative definition is based on concave utility functions, see, e.g., Levy (2006). We say that an investment opportunity $X \in \mathcal{X}$ is SSD efficient if there is no other $\tilde{X} \in \mathcal{X}$ for which $X \prec_{SSD} \tilde{X}$. For our choice of $\mathcal{X}$, it corresponds to SSD portfolio efficiency.

We will consider discretely distributed rates of return with realizations r_{is}, $i = 1,\dots,n$, $s = 1,\dots,S$ and equal probabilities $p_s = 1/S$. Sample approximation technique can be employed to deal with continuous distributions of random returns and to get such discrete distributions, see Section 1.2 or Branda (2014c). For discretely distributed returns with equal probabilities, CVaR of a portfolio can be computed as

$$\mathrm{CVaR}_\alpha^S\left(\sum_{i=1}^{n} R_i x_i\right) = \min_{y\in\mathbb{R}}\left[y + \frac{1}{S(1-\alpha)}\sum_{s=1}^{S}\max\left\{-\sum_{i=1}^{n} x_i r_{is} - y, 0\right\}\right], \tag{1.38}$$

see Rockafellar and Uryasev (2002).

The DEA model (1.35) with $K = S-1$ CVaRs on levels $\alpha_k \in \{1/S,\dots,(S-1)/S\}$ used as the inputs and the expected return as the only output, i.e. $J=1$, can be formulated as

$$
\begin{aligned}
\theta_3^{SSD}(X_0) = \min_{\theta_k,\varphi,x_i} & \frac{1 - \frac{1}{S-1}\sum_{k=1}^{S-1}\theta_k}{1+\varphi} \\
& \text{s.t.} \\
\mathbb{E}\left(\sum_{i=1}^{n} R_i x_i\right) & \geq \mathbb{E}(X_0) + \varphi \cdot e(X_0), \\
\mathrm{CVaR}_{k/S}\left(\sum_{i=1}^{n} R_i x_i\right) & \leq \mathrm{CVaR}_{k/S}(X_0) - \theta_k \cdot d_k(X_0),\ k = 1, \ldots, S-1, \\
\sum_{i=1}^{n} x_i & = 1,\ x_i \geq 0,\ \varphi \geq 0,\ \theta_k \geq 0.
\end{aligned}
\tag{1.39}
$$

with the directions

$$e(X_0) = \max_{X\in\mathcal{X}} \mathbb{E}[X] - \mathbb{E}[X_0],\ d_k(X_0) = \mathrm{CVaR}_{k/S}(X_0) - \min_{X\in\mathcal{X}} \mathrm{CVaR}_{k/S}(X).$$

The DEA model (1.39) can be used by investors as SSD efficiency test, because it allows us to identify SSD efficient investment opportunities, in other words, to test SSD efficiency.

Proposition 1.3.2. (Branda (2015b))
Let the distribution of random returns be discrete. An investment opportunity $X_0 \in \mathcal{X}$ is SSD efficient if and only if it is DEA efficient according to model (1.39), i.e. $\theta_3^{SSD}(X_0) = 1$.

Similar relation to the SSD efficiency was observed by Branda and Kopa (2014) for a radial input-output oriented DEA model. However, their DEA model has no fixed structure for all considered benchmark investment opportunities. Negative risk measures for the benchmark had to be transformed into expectation measures and considered as the outputs. Then, the whole relative improvement in inputs and outputs was not taken into account, hence the scores of inefficient opportunities cannot be compared. This is not a problem for our DEA model (1.39).

As an immediate consequence of previous proposition, we obtain the following important property of our model. As Branda (2015b) demonstrated in the empirical part, the optimal solutions representing efficient portfolios can be used by investors who rebalance their inefficient portfolios to get SSD efficient ones.

Proposition 1.3.3. (Branda (2015b))
Let the distribution of random returns be discrete. Denote by x_i^* the optimal portfolio weights in (1.39). If the benchmark X_0 is inefficient, then the model finds an investment opportunity $X^* = \sum_{i=1}^{n} R_i x_i^*$ which strictly dominates the benchmark, i.e. $X^* \succ_{SSD} X_0$, and at the same time is SSD efficient.

This is not a usual property and is not ensured, e.g., for the SSD efficiency tests proposed by Grechuk (2014), Kuosmanen (2007) and Post (2008). Note that Kuosmanen (2007) considered the basic directional-distance DEA models which are related to the weak Pareto–Koopmans efficiency, whereas we have shown that the strong efficiency is necessary to get a DC DEA model equivalent to SSD efficiency tests.

An investor obtains always a higher score for the SSD dominating investment opportunity compared with the dominated one.

Proposition 1.3.4. (Branda (2015b))
Let the distribution of random returns be discrete. Consider $X_1, X_2 \in \mathcal{X}$. The following implications hold:

- *if $X_1 \preceq_{SSD} X_2$, then $\theta_3^{SSD}(X_1) \leq \theta_3^{SSD}(X_2)$,*
- *if $X_1 \prec_{SSD} X_2$, then $\theta_3^{SSD}(X_1) < \theta_3^{SSD}(X_2)$.*

Branda (2015b) showed that the DC DEA models can be reformulated as convex programming problems. Under our choice of inputs and outputs, the model investigated in this section can be even reformulated as a large linear programming problem, which is very helpful for its solving.

Proposition 1.3.5. (Branda (2015b))
Under the discrete distribution of returns, the model (1.39) can be formulated as a linear programming problem

$$
\begin{aligned}
\theta_3^{SSD}(X_0) = \min_{\tilde{\theta}_k, \tilde{\varphi}, t, \tilde{x}_i, y_k, u_{sk}} \quad & t - \frac{1}{S-1}\sum_{k=1}^{S-1} \tilde{\theta}_k \\
\text{s.t.} \quad & \\
t + \tilde{\varphi} &= 1, \\
\frac{1}{S}\sum_{s=1}^{S}\sum_{i=1}^{n} \tilde{x}_i r_{is} &\geq t \cdot \mathbb{E}[X_0] + \tilde{\varphi} \cdot e(X_0), \\
y_k + \frac{1}{S-k}\sum_{s=1}^{S} u_{sk} &\leq t \cdot \mathrm{CVaR}_{k/S}(X_0) - \tilde{\theta}_k \cdot d_k(X_0),\ k = 1, \ldots, K, \\
u_{sk} &\geq -\sum_{i=1}^{n} \tilde{x}_i r_{is} - y_k,\ s = 1, \ldots, S, k = 1, \ldots, K, \\
\sum_{i=1}^{n} \tilde{x}_i &= t,\ \tilde{x}_i \geq 0,\ \tilde{\varphi} \geq 0,\ \tilde{\theta}_k \geq 0,\ t \geq 0,\ u_{sk} \geq 0.
\end{aligned}
\tag{1.40}
$$

Branda and Kopa (2016) extended the above results and obtained DEA models equivalent to general N-th order stochastic dominance efficiency tests. These models use the lower partial moments and co-lower partial moments as the inputs and outputs and can be also reformulated as linear programming problems.

Bibliography

Adam, L. and Branda, M. (2016a). Nonlinear Chance Constrained Problems: Optimality Conditions, Regularization and Solvers. *Journal of Optimization Theory and Applications*, 170(2):419–436.

Adam, L. and Branda, M. (2016b). Sparse optimization for inverse problems in atmospheric modelling. *Environmental Modelling & Software*, 79:256–266.

Adam, L., Branda, M., Heitsch, H., and Henrion, R. (2018). Solving joint chance constrained problems using regularization and Benders' decomposition. *Optimization Online*, January 2018(Submitted to ANOR):1–21.

Ahmed, S. (2014). Convex relaxations of chance constrained optimization problems. *Optimization Letters*, 8(1):1–12.

Ahmed, S., Luedtke, J., Song, Y., and Xie, W. (2017). Nonanticipative duality, relaxations, and formulations for chance-constrained stochastic programs. *Mathematical Programming*, 162(1):51–81.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

Barrera, J., Homem-de Mello, T., Moreno, E., Pagnoncelli, B. K., and Canessa, G. (2016). Chance-constrained problems and rare events: an importance sampling approach. *Mathematical Programming*, 157(1):153–189.

Bazaraa, M. S., Sherali, H. D., and Shetty, C. M. (2006). *Nonlinear Programming: Theory And Algorithms*. Wiley-Interscience.

Beck, A. and Eldar, Y. C. (2013). Sparsity constrained nonlinear optimization: Optimality conditions and algorithms. *SIAM Journal on Optimization*, 23(3):1480–1509.

Beraldi, P. and Bruni, M. (2010). An exact approach for solving integer problems under probabilistic constraints with random technology matrix. *Annals of Operations Research*, 177(1):127–137.

Bienstock, D. (1996). Computational study of a family of mixed-integer quadratic programming problems. *Mathematical Programming*, 74(2):121–140.

Bonami, P. and Lejeune, M. A. (2009). An exact solution approach for portfolio optimization problems under stochastic and integer constraints. *Operations Research*, 57(3):650–670.

Branda, M. (2012a). Chance constrained problems: penalty reformulation and performance of sample approximation technique. *Kybernetika*, 48(1):105–122.

Branda, M. (2012b). Sample approximation technique for mixed-integer stochastic programming problems with several chance constraints. *Operations Research Letters*, 40(3):207–211.

Branda, M. (2012c). Stochastic programming problems with generalized integrated chance constraints. *Optimization*, 61(8):949–968.

Branda, M. (2013a). Diversification-consistent data envelopment analysis with general deviation measures. *European Journal of Operational Research*, 226(3):626–635.

Branda, M. (2013b). On relations between chance constrained and penalty function problems under discrete distributions. *Mathematical Methods of Operations Research*, 77(2):265–277.

Branda, M. (2013c). Reformulations of input-output oriented DEA tests with diversification. *Operations Research Letters*, 41(5):516–520.

Branda, M. (2014a). Influence of short sales and margin requirements on portfolio efficiency - a DEA-risk approach. In Culik, M, editor, *7th International Scientific Conference on Managing and Modelling of Financial Risks, Ostrava, SEP 08-09, 2014*, pages 97–102. VSB Tech Univ Ostrava.

Branda, M. (2014b). Sample approximation technique for mixed-integer stochastic programming problems with expected value constraints. *Optimization Letters*, 8(3):861–875.

Branda, M. (2014c). Sample approximation techniques for DEA-risk efficiency tests. In Talasova, J and Stoklasa, J and Talasek, T, editor, *32nd International Conference on Mathematical Methods in Economics (MME 2014), Olomouc, SEP 10-12*, pages 55–60.

Branda, M. (2015a). DEA-risk models with Value at Risk inputs. In Martincik, D and Ircingova, J and Janecek, P, editor, *33rd International Conference on Mathematical Methods in Economics (MME 2015), Cheb, SEP 09-11*, pages 68–73.

Branda, M. (2015b). Diversification-consistent data envelopment analysis based on directional-distance measures. *Omega-International Journal of Management Science*, 52:65–76.

Branda, M. (2016a). A chance constrained investment problem with portfolio variance and skewness criteria - solution technique based on the Successive Iterative Regularization. In Kocourek, A and Vavrousek, M, editor, *34th International Conference Mathematical Methods in Economics (MME 2016), Liberec, SEP 06-09*, pages 67–72. Tech Univ Liberec.

Branda, M. (2016b). Mean-value at risk portfolio efficiency: approaches based on data envelopment analysis models with negative data and their empirical behaviour. *4OR-A Quaterly Journal of Operations Research*, 14(1):77–99.

Branda, M., Bucher, M., Červinka, M., and Schwartz, A. (2018). Convergence of a scholtes-type regularization method for cardinality-constrained optimization problems with an application in sparse robust portfolio optimization. *Computational Optimization and Applications*, 70(2):503–530.

Branda, M. and Dupačová, J. (2012). Approximation and contamination bounds for probabilistic programs. *Annals of Operations Research*, 193(1):3–19.

Branda, M. and Hájek, Š. (2017). Flow-based formulations for operational fixed interval scheduling problems with random delays. *Computational Management Science*, 14(1, SI):161–177.

Branda, M. and Kopa, M. (2014). On relations between DEA-risk models and stochastic dominance efficiency tests. *Central European Journal of Operations Research*, 22(1):13–35.

Branda, M. and Kopa, M. (2016). DEA models equivalent to general Nth order stochastic dominance efficiency tests. *Operations Research Letters*, 44(2):285–289.

Branda, M., Novotný, J., and Olstad, A. (2016). Fixed interval scheduling under uncertainty - A tabu search algorithm for an extended robust coloring formulation. *Computers & Industrial Engineering*, 93:45–54.

Branda, M., Novotný, J., Olstad, A., and Popela, P. (2015). A note on fixed interval scheduling with stochastic elements. *Informational Bulletin of the Czech Statistical Society*, 26(4):1–7.

Briec, W., Kerstens, K., and Jokung, O. (2007). Mean-variance-skewness portfolio performance gauging: A general shortage function and dual approach. *Management Science*, 53(1):135–149.

Briec, W., Kerstens, K., and Lesourd, J. B. (2004). Single-period markowitz portfolio selection, performance gauging, and duality: A variation on the luenberger shortage function. *Journal of Optimization Theory and Applications*, 120(1):1–27.

Burdakov, O. P., Kanzow, C., and Schwartz, A. (2016). Mathematical programs with cardinality constraints: Reformulation by complementarity-type conditions and a regularization method. *SIAM Journal on Optimization*, 26(1):397–425.

Burke, J. V. (1991a). Calmness and exact penalization. *SIAM Journal on Control and Optimization*, 29(2):493–497.

Burke, J. V. (1991b). An exact penalization viewpoint of constrained optimization. *SIAM Journal on Control and Optimization*, 29(4):968–998.

Červinka, M., Kanzow, C., and Schwartz, A. (2016). Constraint qualifications and optimality conditions for optimization problems with cardinality constraints. *Mathematical Programming*, 160(1):353–377.

Charnes, A., Cooper, W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6):429 – 444.

Charnes, A., Cooper, W. W., and Symonds, G. H. (1958). Cost horizons and certainty equivalents: An approach to stochastic programming of heating oil. *Management Science*, 4(3):235–263.

Cheng, J. and Lisser, A. (2012). A second-order cone programming approach for linear programs with joint probabilistic constraints. *Operations Research Letters*, 40(5):325–328.

Clarke, F. H. (1983). *Optimization and Nonsmooth Analysis*. SIAM Publications Classics in Applied Mathematics.

Dai, L., Chen, C. H., and Birge, J. R. (2000). Convergence properties of two-stage stochastic programming. *Journal of Optimization Theory and Applications*, 106(3):489–509.

de Mello, T. H. (2008). On rates of convergence for stochastic optimization problems under non-independent and identically distributed sampling. *SIAM Journal on Optimization*, 19(2):524–551.

de Mello, T. H. and Bayraksan, G. (2014). Monte carlo sampling-based methods for stochastic optimization. *Surveys in Operations Research and Management Science*, 19(1):56–85.

Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*. Applications of mathematics. Springer.

Dentcheva, D. and Martinez, G. (2012). Augmented Lagrangian method for probabilistic optimization. *Annals of Operations Research*, 200(1):109–130.

Dentcheva, D. and Martinez, G. (2013). Regularization methods for optimization problems with probabilistic constraints. *Mathematical Programming*, 138(1-2):223–251.

Dokov, S., Morton, D. P., and Popova, I. (2017). Mean-variance-skewness-kurtosis efficiency of portfolios computed via moment-based bounds. In *2017 International Conference on Information Science and Communications Technologies (ICISCT)*, pages 1–5.

Ermoliev, Y., Ermolieva, T., MacDonald, G., and Norkin, V. (2000). Stochastic optimization of insurance portfolios for managing exposure to catastrophic risks. *Annals of Operations Research*, 99(1):207–225.

Fábián, C. I., Csizmás, E., Drenyovszki, R., van Ackooij, W., Vajnai, T., Kovács, L., and Szántai, T. (2018). Probability maximization by inner approximation. *Acta Polytechnica Hungarica*, 15(1):105–125.

Gao, J. and Li, D. (2013). Optimal cardinality constrained portfolio selection. *Operations Research*, 61(3):745–761.

Geletu, A., Hoffmann, A., Kläppel, M., and Li, P. (2017). An inner-outer approximation approach to chance constrained optimization. *SIAM Journal on Optimization*, 27(3):1834–1857.

Gotzes, C., Heitsch, H., Henrion, R., and Schultz, R. (2016). On the quantification of nomination feasibility in stationary gas networks with random load. *Mathematical Methods of Operations Research*, 84(2):427–457.

Grechuk, B. (2014). A simple ssd-efficiency test. *Optimization Letters*, 8(7):2135–2143.

Haneveld, W. and van der Vlerk, M. (2006). Integrated chance constraints: Reduced forms and an algorithm. *Computational Management Science*, 3(4):245–269.

Heitsch, H., Leövey, H., and Römisch, W. (2016). Are quasi-monte carlo algorithms efficient for two-stage stochastic programs? *Computational Optimization and Applications*, 65(3):567–603.

Henrion, R. (2013). A critical note on empirical (Sample Average, Monte Carlo) approximation of solutions to chance constrained programs. In Hömberg, D. and Tröltzsch, F., editors, *System Modeling and Optimization*, volume 391 of *IFIP Advances in Information and Communication Technology*, pages 25–37. Springer Berlin Heidelberg.

Henrion, R. and Möller, A. (2012). A gradient formula for linear chance constraints under Gaussian distribution. *Mathematics of Operations Research*, 37(3):475–488.

Hering, C. and Stadtmüller, U. (2012). Estimating archimedean copulas in high dimensions. *Scandinavian Journal of Statistics*, 39(3):461–479.

Joro, T. and Na, P. (2006). Portfolio performance evaluation in a mean–variance–skewness framework. *European Journal of Operational Research*, 175(1):446 – 461.

Kaňková, V. (1990). On the convergence rate of empirical estimates in chance constrained stochastic programming. *Kybernetika*, 26(6):449–461.

Kerstens, K., Mounir, A., and de Woestyne, I. V. (2011). Geometric representation of the mean-variance-skewness portfolio frontier based upon the shortage function. *European Journal of Operational Research*, 210(1):81 – 94.

Kerstens, K., Mounir, A., and Van de Woestyne, I. (2012). Benchmarking mean-variance portfolios using a shortage function: the choice of direction vector affects rankings! *Journal of the Operational Research Society*, 63(9):1199–1212.

Kroon, L. G., Salomon, M., and Wassenhove, L. N. V. (1995). Exact and approximation algorithms for the operational fixed interval scheduling problem. *European Journal of Operational Research*, 82(1):190–205.

Kuosmanen, T. (2007). Performance measurement and best-practice benchmarking of mutual funds: combining stochastic dominance criteria with data envelopment analysis. *Journal of Productivity Analysis*, 28(1):71–86.

Lamb, J. D. and Tee, K.-H. (2012). Data envelopment analysis models of investment funds. *European Journal of Operational Research*, 216(3):687 – 696.

Levy, H. (2006). *Stochastic Dominance: Investment Decision Making under Uncertainty*. Studies in Risk and Uncertainty. Springer US.

Liu, W., Zhou, Z., Liu, D., and Xiao, H. (2015). Estimation of portfolio efficiency via dea. *Omega*, 52:107 – 118.

Lozano, S. and Gutiérrez, E. (2008). Data envelopment analysis of mutual funds based on second-order stochastic dominance. *European Journal of Operational Research*, 189(1):230 – 244.

Luedtke, J. (2014). A branch-and-cut decomposition algorithm for solving chance-constrained mathematical programs with finite support. *Mathematical Programming*, 146(1-2):219–244.

Luedtke, J. and Ahmed, S. (2008). A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19(2):674–699.

Luedtke, J., Ahmed, S., and Nemhauser, G. (2010). An integer programming approach for linear programs with probabilistic constraints. *Mathematical Programming*, 122(2):247–272.

McNeil, A. J. and Nešlehová, J. (2009). Multivariate archimedean copulas, d-monotone functions and l_1-norm symmetric distributions. *The Annals of Statistics*, 37(5B):3059–3097.

Michalopoulos, D. P., Barnes, J. W., and Morton, D. P. (2015). Prioritized interdiction of nuclear smuggling via tabu search. *Optimization Letters*, 9(8):1477–1494.

Murthi, B., Choi, Y. K., and Desai, P. (1997). Efficiency of mutual funds and portfolio performance measurement: A non-parametric approach. *European Journal of Operational Research*, 98(2):408 – 418.

Nemirovski, A. and Shapiro, A. (2007). Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996.

Portela, M. C. A. S., Thanassoulis, E., and Simpson, G. (2004). Negative data in dea: a directional distance approach applied to bank branches. *Journal of the Operational Research Society*, 55(10):1111–1121.

Post, T. (2008). On the dual test for ssd efficiency: With an application to momentum investment strategies. *European Journal of Operational Research*, 185(3):1564 – 1573.

Prékopa, A. (2003). Probabilistic programming. In Ruszczyński, A. and Shapiro, A., editors, *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*, pages 267–351. Elsevier.

Raike, W. M. (1970). Dissection methods for solutions in chance constrained programming problems under discrete distributions. *Management Science*, 16(11):708–715.

Rockafellar, R. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41.

Rockafellar, R. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471.

Rockafellar, R. T. and Wets, R. J.-B. (1998). *Variational Analysis*. Springer.

Shan, F., Zhang, L., and Xiao, X. (2014). A smoothing function approach to joint chance-constrained programs. *Journal of Optimization Theory and Applications*, 163(1):181–199.

Shapiro, A. (2003). Monte carlo sampling methods. In Ruszczyński, A. and Shapiro, A., editors, *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*, pages 353–425. Elsevier.

Sun, H., Xu, H., and Wang, Y. (2014). Asymptotic analysis of sample average approximation for stochastic optimization problems with joint chance constraints via conditional value at risk and difference of convex functions. *Journal of Optimization Theory and Applications*, 161(1):257–284.

van Ackooij, W. and Henrion, R. (2014). Gradient formulae for nonlinear probabilistic constraints with gaussian and gaussian-like distributions. *SIAM Journal on Optimization*, 24(4):1864–1889.

Wang, W. and Ahmed, S. (2008). Sample average approximation of expected value constrained stochastic programs. *Operations Research Letters*, 36(5):515 – 519.

Wozabal, D., Hochreiter, R., and Pflug, G. C. (2010). A difference of convex formulation of value-at-risk constrained optimization. *Optimization*, 59(3):377–400.

Xie, W. and Ahmed, S. (2017). On quantile cuts and their closure for chance constrained optimization problems. *Mathematical Programming*.

Xu, H. (2010). Uniform exponential convergence of sample average random functions under general sampling with applications in stochastic programming. *Journal of Mathematical Analysis and Applications*, 368(2):692 – 710.

Yanez, J. and Ramirez, J. (2003). The robust coloring problem. *European Journal of Operational Research*, 148(3):546 – 558.

Zhao, M., Huang, K., and Zeng, B. (2017). A polyhedral study on chance constrained program with random right-hand side. *Mathematical Programming*, 166(1):19–64.

Zheng, X., Sun, X., Li, D., and Sun, J. (2014). Successive convex approximations to cardinality-constrained convex programs: a piecewise-linear dc approach. *Computational Optimization and Applications*, 59(1):379–397.

Appendix: Attached papers

1. Adam, L., Branda, M. (2016). Nonlinear chance constrained problems: optimality conditions, regularization and solvers. *Journal of Optimization Theory and Applications* 170 (2): 419–436.
`http://dx.doi.org/10.1007/s10957-016-0943-9`

2. Branda, M., Bucher, M., Červinka, M., Schwartz, A. (2018). Convergence of a Scholtes-type Regularization Method for Cardinality-Constrained Optimization Problems with an Application in Sparse Robust Portfolio Optimization. *Computational Optimization and Applications* 70 (2): 503–530.
`http://dx.doi.org/10.1007/s10589-018-9985-2`

3. Branda, M. (2013). On relations between chance constrained and penalty function problems under discrete distributions. *Mathematical Methods of Operations Research* 77 (2): 265–277.
`http://dx.doi.org/10.1007/s00186-013-0428-7`

4. Branda, M., Novotný, J., Olstad, A. (2016). Fixed interval scheduling under uncertainty - a tabu search algorithm for an extended robust coloring formulation. *Computers & Industrial Engineering* 93: 45–54.
`http://dx.doi.org/10.1016/j.cie.2015.12.021`

5. Branda, M. (2012). Sample approximation technique for mixed-integer stochastic programming problems with several chance constraints. *Operations Research Letters* 40 (3): 207–211.
`http://dx.doi.org/10.1016/j.orl.2012.01.002`

6. Branda, M. (2014). Sample approximation technique for mixed-integer stochastic programming problems with expected value constraints. *Optimization Letters* 8 (3): 861–875.
`http://dx.doi.org/10.1007/s11590-013-0642-5`

7. Branda, M. (2013). Diversification-consistent data envelopment analysis with general deviation measures. *European Journal of Operational Research* 226 (3): 626–635.
`http://dx.doi.org/10.1016/j.ejor.2012.11.007`

8. Branda, M. (2015). Diversification-consistent data envelopment analysis based on directional-distance measures. *Omega–International Journal of Management Science* 52: 65–76.
`http://dx.doi.org/10.1016/j.omega.2014.11.004`

9. Branda, M. (2016). Mean-value at risk portfolio efficiency: approaches based on data envelopment analysis models with negative data and their empirical behaviour. *4OR–A Quaterly Journal of Operations Research* 14 (1): 77–99.
http://dx.doi.org/10.1007/s10288-015-0296-5

10. Branda, M., Kopa, M. (2016). DEA models equivalent to general N-th order stochastic dominance efficiency tests. *Operations Research Letters* 44 (2): 285–289.
http://dx.doi.org/10.1016/j.orl.2016.02.007