A scenario decomposition algorithm for stochastic programming problems with a class of downside risk measures

Maciej Rysz[1] Alexander Vinel[1] Pavlo A. Krokhmal[1]* Eduardo L. Pasiliao[2]

[1]Department of Mechanical and Industrial Engineering,
The University of Iowa, 3131 Seamans Center, Iowa City, IA, 52242

[2]Air Force Research Lab, 101 West Eglin Blvd, Eglin AFB, FL 32542

Abstract

We present an efficient scenario decomposition algorithm for solving large-scale convex stochastic programming problems that involve a particular class of downside risk measures. The considered risk functionals encompass coherent and convex measures of risk that can be represented as an infimal convolution of a convex certainty equivalent, and include well-known measures, such as conditional value-at-risk, as special cases. The resulting structure of the feasible set is then exploited via iterative solving of relaxed problems, and it is shown that the number of iterations is bounded by a parameter that depends on the problem size. The computational performance of the developed scenario decomposition method is illustrated on portfolio optimization problems involving two families of nonlinear measures of risk, the higher-moment coherent risk measures and log-exponential convex risk measures. It is demonstrated that for large-scale nonlinear problems the proposed approach can provide up to an order of magnitude of improvement in computational time in comparison to state-of-the-art solvers, such as CPLEX, Gurobi, and MOSEK.

Keywords: Stochastic optimization, risk measures, utility theory, certainty equivalent, scenario decomposition, higher moment coherent risk measures, log-exponential convex risk measures.

1 Introduction and Motivation

Quantification of uncertainties and risk via axiomatically defined statistical functionals, such as the *coherent measures of risk* of Artzner et al. (1999), has become a widely accepted practice in stochastic optimization and decision making under uncertainty (Shapiro et al., 2009; Krokhmal et al., 2011; Uryasev and Rockafellar, 2013). Many of such risk measures admit effective utilization in "scenario-based" formulations of stochastic programming models, i.e., the stochastic optimization problems where the random parameters are assumed to have a known distribution over a finite support that is commonly

*Corresponding author, `krokhmal@engineering.uiowa.edu`

called the *scenario set*. A typical instance of such a problem can be written as

$$\min_{\mathbf{x}\in C} \ \rho(X(\mathbf{x},\omega)), \tag{1}$$

where ρ is the risk measure, $X(\mathbf{x},\omega)$ represents a stochastic loss or cost function dependent on the decision vector $\mathbf{x} \in C \subset \mathbb{R}^n$ and a random event ω from the finite set $\Omega = \{\omega_1, \ldots, \omega_N\}$. In many practical applications accurate approximations of uncertainties may, however, require very large scenario sets ($N \gg 1$), thus potentially leading to substantial computational difficulties.

In this work, we propose an efficient algorithm for solving large-scale stochastic optimization problems involving a class of "downside", or "tail" risk measures that are constructed via *certainty equivalents*, a well known concept in the utility theory. The presented scenario decomposition algorithm exploits the special structure of the feasible set induced by the respective risk measures as well as the properties common to the considered class of risk functionals. As an illustrative example of the general approach, we consider stochastic optimization problems with higher-moment coherent risk measures (HMCR), which quantify risk via higher moments of cost or loss distributions (Krokhmal, 2007), making them advantageous in the presence of "heavy-tailed" uncertainty. We also apply the proposed method to problems with log-exponential convex risk (LogExpCR) measures (Vinel and Krokhmal, 2015).

Perhaps, the most frequently implemented risk measure in problems of type (1) is the well known Conditional Value-at-Risk (CVaR) (Rockafellar and Uryasev, 2000, 2002). When X is piecewise linear in $\mathbf{x}$ and set C is polyhedral, formulation (1) with CVaR objective or constraints reduces to a linear programming (LP) problem. Several recent studies addressed the solution efficiency of LPs with CVaR objectives or constraints for cases when the number of scenarios is large. Lim, Sherali, and Uryasev (2010) noted that (1) in this case may be viewed as a nondifferentiable optimization problem and implemented a two-phase solution approach to solve large-scale instances. In the first phase, they exploit descent-based optimization techniques to circumvent nondifferentiable points by perturbing the solution to differentiable solutions within their "relative neighborhood". The second phase employs a deflecting subgradient search direction with a step size established by an adequate target value. They further extended this approach with a third phase that resorts to the simplex algorithm after achieving convergence by employing an advanced crash-basis dependent on solutions obtained from the first two phases.

Künzi-Bay and Mayer (2006) developed a solution technique for problem (1), with measure ρ chosen as the CVaR, that utilized a specialized L-shaped method after reformulating it as a two-stage stochastic programming problem. However, Subramanian and Huang (2008) noted that the problem structure does not naturally conform to the characteristics of a two-stage stochastic program and introduced a polyhedral reformulation of the CVaR constraint with a statistics based CVaR estimator to solve a closely related version of the problem. In a followup study (Subramanian and Huang, 2009), they retained Value-at-Risk (VaR) and CVaR as unknown variables in the CVaR constraints, enabling a more efficient decomposition algorithm, as opposed to Klein Haneveld and van der Vlerk (2006), where the problem was solved as a canonical integrated chance constraint problem with preceding estimates of VaR. Espinoza and Moreno (2012) proposed a solution method for problems (1) with CVaR measures that entailed generation of aggregated scenario constraints to form smaller relaxation problems whose optimal outcomes were then used to directly evaluate the respective upper bound on the objective of the original problem.

In what follows, we develop a general scenario decomposition solution framework for solving stochastic optimization problems with certainty equivalent-based risk measures by utilizing principles related to those in Espinoza and Moreno (2012). The rest of the paper is organized as follows: A class of certainty equivalent-based risk measures that are in the focus of this study and their implementation in

mathematical programming problems are discussed in Section 2. In Section 3 we propose the scenario decomposition algorithm for stochastic programming problems with structure that is induced by the risk measures described in Section 2. Lastly, experimental studies on portfolio optimization problems with large-scale data sets that demonstrate the effectiveness of the developed technique are presented in Section 4 .

2 A Class of Downside Risk Measures Based on Certainty Equivalents

In this section we describe a class of risk measures that encompasses some popular instances in risk management literature. A general solution algorithm that utilizes special properties of this class of measures will be presented in the sequel. Specifically, this algorithm applies to the so-called *coherent* and *convex* measures of risk that can be represented as an infimal convolution of certainty equivalent of some utility function. Below we recall the definitions of coherent and convex risk measures and describe the representation that motivated the present development.

In general, a risk measure $\rho(X)$ over a random outcome (specifically, a cost or a loss) X from probability space $(\Omega, \mathcal{F}, \mathsf{P})$ is defined as a lower semi-continuous (l.s.c.) mapping $\rho : \mathcal{X} \mapsto \mathbb{R}$, with $\mathcal{X}$ being the space of bounded $\mathcal{F}$-measurable functions $X : \Omega \mapsto \mathbb{R}$. In order to avoid an excessively technical discussion, we will implicitly assume that $\mathcal{X}$ is endowed with the properties necessary in the given context (e.g., integrability, and so on). Additional properties of ρ are introduced to make the corresponding risk measure well-suited for a specific application area.

Artzner et al. (1999) and Delbaen (2002) proposed the following four axioms as the desirable characteristics that a "good", or *coherent* measure of risk should possess:

(A1) *monotonicity:* $\rho(X) \leq \rho(Y)$ *for all* $X, Y \in \mathcal{X}$ *such that* $X \leq Y$*;*

(A2) *convexity:* $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$ *for all* $X, Y \in \mathcal{X}$ *and* $0 \leq \lambda \leq 1$*;*

(A3) *positive homogeneity:* $\rho(\lambda X) = \lambda\rho(X)$ *for all* $X \in \mathcal{X}$ *and* $\lambda > 0$*;*

(A4) *translation invariance:* $\rho(X + a) = \rho(X) + a$ *for all* $X \in \mathcal{X}$ *and* $a \in \mathbb{R}$.

The following interpretations may be given to the above axioms: Axiom (A1) ensures that smaller losses lead to lower risk. From the risk management point of view, the convexity axiom (A2) promotes risk reduction via diversification; it is also of fundamental importance in the optimization context. The positive homogeneity property (A3) postulates that scaling losses by a positive factor scales risk correspondingly. Axiom (A4) allows for eliminating risk of an uncertain cost/loss profile X by adding a deterministic hedge, $\rho(X - \rho(X)) = 0$.

Since being proposed in Artzner et al. (1999) and Delbaen (2002), the axiomatic approach to defining risk measures has been widely adopted in literature, and a number of risk functionals tailored to particular preferences emerged thereafter (see, e.g., Krokhmal et al., 2011; Uryasev and Rockafellar, 2013). In particular, it has been argued that the positive homogeneity property (A3) may be omitted in many situations; the corresponding risk measures that satisfy axioms (A1), (A2), and (A4) are called *convex measures of risk* (Ruszczyński and Shapiro, 2006).

Our interest in these two classes of risk measures stems from the following *infimal convolution* representation that facilitates their use in mathematical programming problems.

Theorem 1 (Krokhmal, 2007; Vinel and Krokhmal, 2014a) *Function $\rho(X)$ is a proper coherent (resp., convex) measure of risk if and only if it can be represented by the following infimal convolution of a l.s.c. function $\phi : \mathcal{X} \mapsto \mathbb{R}$ such that $\phi(0) = 0$, $\phi(\eta) > \eta$ for all real $\eta \neq 0$, and which satisfies (A1)–(A3) (resp., (A1)–(A2)):*

$$\rho(X) = \inf_{\eta} \ \eta + \phi(X - \eta). \tag{2}$$

Moreover, the infimum in (2) *is attained for all X, so* $\inf_\eta$ *may be replaced by* $\min_{\eta\in\mathbb{R}}$.

Representation (2) can be used for construction of coherent (convex) risk measures through an appropriate choice of function ϕ. The present work concerns risk measures of type (2) that can directly incorporate decision maker's risk preferences as given by the utility theory of von Neumann and Morgenstern (1944). This is desirable in view of the well-known fact (see, e.g., Schied and Follmer, 2002) that risk preferences expressed by coherent/convex measures of risk are generally not compatible with rational risk-averse preferences (i.e., those defined by a non-decreasing concave utility function u).

Given that we operate with stochastic cost/loss variables, let $v(t) = -u(-t)$ be the utility function adapted to loss variable X, or *deutility* function that quantifies dissatisfaction with cost or loss X. Then, $\mathrm{CE}(X) = v^{-1}\big(\mathsf{E}v(X)\big)$ represents the certainty equivalent (CE) of loss X, i.e., such a deterministic loss that a rational decision maker with deutility function v would be indifferent between $\mathrm{CE}(X)$ and stochastic loss profile X. The following argument can be used to construct risk measures of the form (2) that employ rational utility maximizer's preferences via certainty equivalents (Vinel and Krokhmal, 2015, see also Ben-Tal and Teboulle, 2007). Consider a decision maker who faces an uncertain future loss X, but who can allocate an amount η of resources now to cover the future loss. It will cost $v^{-1}\mathsf{E}v(X-\eta)_+$ to cover the remaining losses $(X-\eta)_+$, where $t_+ = \max\{0, t\}$ and an operator-like notation is used for v, i.e., $v^{-1}\mathsf{E}v(X-\eta)_+ = v^{-1}(\mathsf{E}v((X-\eta)_+))$. The total cost can then be optimized with an appropriate choice of η, such that the risk $\rho(X)$ of a future loss X reduces to

$$\rho(X) = \min_{\eta} \ \ \eta + \frac{1}{1-\alpha} v^{-1}\mathsf{E}v(X-\eta)_+, \quad \alpha \in (0,1), \tag{3}$$

where $(1-\alpha)^{-1} > 1$ is a penalty factor (a detailed discussion of representation (3) and related aspects is presented in Vinel and Krokhmal, 2014a).

Notably, expressing $\phi(X)$ in (2) via certainty equivalents necessarily requires that $(\cdot)_+$ appears in (3) in order for $\phi(X)$ to conform to the conditions of Theorem 1 (Vinel and Krokhmal, 2014a). The conditions on v that guarantee convexity of $\mathrm{CE}(X) = v^{-1}\mathsf{E}v(X)$, and, correspondingly, of $\phi(X)$, can be found, for example, in Ben-Tal and Teboulle (2007): v should be three times continuously differentiable, and $v'(t)/v''(t)$ be convex. In what follows, we implicitly assume that $\phi(X) = (1-\alpha)^{-1}v^{-1}\mathsf{E}v(X_+)$ is convex and satisfies the conditions of Theorem 1:

(U1) *Function $v(t)$ is continuously differentiable, increasing, convex, and, moreover, such that $v(0) = 0$ and the certainty equivalent $v^{-1}\mathsf{E}v(X)$ is convex in X.*

A key property of risk measures (3) is isotonicity with respect to second order stochastic dominance (SSD), provided that deutility function v is convex and nondecreasing:

(A5) *SSD isotonicity: $\rho(X) \leq \rho(Y)$ for all $X, Y \in \mathcal{X}$ such that $(-X) \succeq_{\mathrm{SSD}} (-Y)$.*

Recall that payoff profile Y_1 dominates Y_2 with respect to SSD, $Y_1 \succeq_{\mathrm{SSD}} Y_2$, if and only if $\mathsf{E}u(Y_1) \geq \mathsf{E}u(Y_2)$ holds for all non-decreasing concave utility functions u, or, in other words, if every rational risk-averse decision maker prefers Y_1 over Y_2. In this regard, (A5) implies that risk measures (3) "inherit" the

risk preferences given by the utility u (equivalently, v). It is important to note that coherent and convex measures of risk are generally not SSD-isotonic (Krokhmal et al., 2011).

Another common property of risk measures (3) is that they are "tail" risk measures in the sense that the tail $\{X : X \geq \eta^*_\alpha(X)\}$ of the loss distribution is used to quantify risk, where the location of the "tail cutoff" point $\eta^*_\alpha(X)$, which is a minimizer in (3), can be adjusted according to risk preferences via the parameter α (see Krokhmal, 2007; Vinel and Krokhmal, 2014a).

Several practical and interesting risk measure families can be obtained from (3) by selecting a specific deutility function v. If $v(t) = t$, then (3) defines the well-known Conditional Value-at-Risk measure (Rockafellar and Uryasev, 2002, 2000):

$$\mathsf{CVaR}_\alpha(X) = \min_\eta \ \eta + (1-\alpha)^{-1}\mathsf{E}(X-\eta)_+, \quad \alpha \in (0,1). \tag{4}$$

If $v(t) = t^p$ for $t \geq 0$ and $p > 1$, then representation (3) yields a two-parametric family of *higher-moment coherent risk measures* (HMCR) (Krokhmal, 2007):

$$\mathsf{HMCR}_{p,\alpha}(X) = \min_\eta \ \eta + (1-\alpha)^{-1}\|(X-\eta)_+\|_p, \quad \alpha \in (0,1), \quad p \geq 1, \tag{5}$$

where $\|X\|_p = \big(\mathsf{E}|X|^p\big)^{1/p}$. If $v(t) = \lambda^t - 1$, $\lambda > 1$, then one obtains the family of log-exponential convex measures of risk (Vinel and Krokhmal, 2015):

$$\mathsf{LogExpCR}_{\lambda,\alpha}(X) = \min_\eta \ \eta + (1-\alpha)^{-1}\log_\lambda \mathsf{E}\lambda^{(X-\eta)_+}, \quad \alpha \in (0,1), \quad \lambda > 1. \tag{6}$$

Unlike the CVaR and HMCR measures that are coherent, the LogExpCR measure is *convex* but not coherent as it does not satisfy the positive homogeneity axiom (A3).

Perhaps one of the most widely used coherent measures of risk is defined by (4), which represents, roughly speaking, the conditional expectation of losses that may occur in the $(1-\alpha)\cdot 100\%$ of worst realizations of X. Clearly, CVaR measure is a special case of (5) when $p = 1$, $\mathsf{HMCR}_{1,\alpha}(X) = \mathsf{CVaR}_\alpha(X)$. When $p > 1$, HMCR measures quantify risk via higher tail moments $\|(X-\eta)_+\|_p$, and have been shown to be better suited for applications that involve heavy tailed loss distributions (Krokhmal, 2007). Likewise, the LogExpCR family (6) is designed for dealing with heavy-tailed distributions; moreover, in addition to being SSD-isotonic, LogExpCR measures are isotonic with respect to stochastic dominance of arbitrary order (kSD), see Vinel and Krokhmal (2015).

Next we discuss the implementation of the risk measures discussed above in mathematical programming problems.

2.1 Implementation in Stochastic Programming

Assume that loss X is a function of the decision variable $\mathbf{x}$, $X = X(\mathbf{x}, \omega)$, where $\omega \in \Omega$. Then, for a compact and convex feasible set $C \subset \mathbb{R}^n$, consider a stochastic programming problem with a risk constraint in the form

$$\min\big\{g(\mathbf{x}) : \rho\big(X(\mathbf{x},\omega)\big) \leq h(\mathbf{x}),\ \mathbf{x} \in C\big\}. \tag{7}$$

Theorem 2 *Consider problem* (7) *where set $C \subset \mathbb{R}^n$ is compact and convex, and functions $g(\mathbf{x})$ and $h(\mathbf{x})$ are convex and concave on C, respectively. If, further, the cost or loss function $X(\mathbf{x},\omega)$ is convex in*

$\mathbf{x}$, *and* ρ *is a coherent or convex measure of risk with representation* (2)*, then problem* (7) *is equivalent to*

$$\min\left\{g(\mathbf{x}) : \eta + \phi\big(X(\mathbf{x},\omega) - \eta\big) \le h(\mathbf{x}),\ (\mathbf{x},\eta) \in C \times \mathbb{R}\right\}, \tag{8}$$

in the sense that (7) *and* (8) *achieve minima at the same values of the decision variable* $\mathbf{x}$ *and their optimal objective values coincide. Further, if the risk constraint in* (7) *is binding at optimality,* $(\mathbf{x}^*, \eta^*)$ *achieves the minimum of* (8) *if and only if* $\mathbf{x}^*$ *is an optimal solution of* (7) *and*

$$\eta^* \in \arg\min_\eta\ \eta + \phi\big(X(\mathbf{x}^*,\omega) - \eta\big).$$

Proof: See Krokhmal (2007). □

Remark 1 Note that the risk minimization problem

$$\min\left\{\rho\big(X(\mathbf{x},\omega)\big) : \mathbf{x} \in C\right\} \tag{9}$$

is obtained from (7) by introduction of a dummy variable x_{n+1} and letting $g(\mathbf{x}) = h(\mathbf{x}) = x_{n+1}$.

Let function ϕ in (8) have the form $\phi(X) = (1-\alpha)^{-1} v^{-1} \mathsf{E} v(X_+)$. Given a discrete set of scenarios $\{\omega_1, \ldots, \omega_N\} = \Omega$ that induce cost or loss outcomes $X(\mathbf{x},\omega_1), \ldots, X(\mathbf{x},\omega_N)$ for any given decision vector $\mathbf{x}$, it is easy to see that the risk constraint in (8) can be represented by the following set of inequalities:

$$\eta + (1-\alpha)^{-1} w_0 \le h(\mathbf{x}), \tag{10a}$$

$$w_0 \ge v^{-1}\Big(\sum_{j\in\mathcal{N}} \pi_j v(w_j)\Big), \tag{10b}$$

$$w_j \ge X(\mathbf{x},\omega_j) - \eta, \quad j \in \mathcal{N}, \tag{10c}$$

$$w_j \ge 0, \quad j \in \mathcal{N}, \tag{10d}$$

where $\mathcal{N}$ denotes the set of scenario indices, $\mathcal{N} = \{1, \ldots, N\}$, and $\pi_j = \mathsf{P}(\omega_j) > 0$ represent the corresponding scenario probabilities that satisfy $\pi_1 + \cdots + \pi_N = 1$.

In the above discussion it was shown that several types of risk measures emerge from different choices of the deutility function v. Here we note that the corresponding representations of constraint (10b) in the context of HMCR and LogExpCR measures lead to sufficiently "nice", i.e., convex, mathematical programming models. For HMCR measures inequality (10b) becomes

$$w_0 \ge \Big(\sum\nolimits_{j\in\mathcal{N}} \pi_j w_j^p\Big)^{1/p}, \tag{11}$$

which is equivalent to a standard p-order cone under affine scaling. Noteworthy instances of (11) for which readily available mathematical programming solution methods exist include $p = 1, 2$. In the particular case of $p = 1$, which corresponds to CVaR, the problem reduces to a linear programming (LP) model. For instances when $p = 2$, a second-order cone programming (SOCP) model that is efficiently solvable using long-step self-dual interior point methods transpires. However, no similarly efficient solution methods exist for solving p-order conic constrained problems when $p \in (1,2) \cup (2,\infty)$ due to the fact that the p-cone is not self-dual in this case. Additional discussion and computational

considerations for such instances are given in Section 4.1. Lastly, the following exponential inequality corresponds to constraint (10b) when ρ is a LogExpCR measure:

$$w_0 \geq \ln \sum\nolimits_{j \in \mathcal{N}} \pi_j e^{w_j}, \tag{12}$$

which is also convex and allows for the resulting optimization problem to be solved using appropriate (e.g., interior point) methods.

3 Scenario Decomposition Algorithm

Large-scale stochastic optimization models with CVaR measure (4) and the corresponding solution algorithms have received considerable attention in the literature. In this section we propose an efficient scenario decomposition algorithm for solving large-scale mathematical programming problems that use certainty equivalent-based risk measures (3), which contain CVaR as a special case.

The algorithm relies on solving a series of relaxation problems containing linear combinations of scenario-based constraints that are systematically decomposed until an optimal solution of the original problem is found or the problem is proven to be infeasible. Naturally, the core assumption behind such a scheme is that sequential solutions of smaller relaxation problems can be achieved within shorter computation times. By virtue of Section 2, when the distribution of loss function $X(\mathbf{x}, \omega)$ has a finite support (scenario set) $\Omega = \{\omega_1, \ldots, \omega_N\}$ with probabilities $\mathsf{P}(\omega_j) = \pi_j > 0$, the stochastic programming problem with risk constraint (8) admits the form

$$\begin{aligned}
\min \quad & g(\mathbf{x}) && \text{(13a)}\\
\text{s. t.} \quad & \mathbf{x} \in C, && \text{(13b)}\\
& \eta + (1-\alpha)^{-1} w_0 \leq h(\mathbf{x}), && \text{(13c)}\\
& w_0 \geq v^{-1}\Big(\sum_{j \in \mathcal{N}} \pi_j v(w_j) \Big), && \text{(13d)}\\
& w_j \geq X(\mathbf{x}, \omega_j) - \eta, \quad j \in \mathcal{N}, && \text{(13e)}\\
& w_j \geq 0, \quad j \in \mathcal{N}, && \text{(13f)}
\end{aligned}$$

where $\mathcal{N} = \{1, \ldots, N\}$. If we assume that function $g(\mathbf{x})$ and feasible set C are "nice" in the sense that problem $\min\{g(\mathbf{x}) : \mathbf{x} \in C\}$ admits efficient solution methods, then formulation (13) may present challenges that are two-fold. First, constraint (13d) may need a specialized solution approach, especially in the case of large N. Similarly, when N is large, computational difficulties may be associated with handling the large number of constraints (13e)–(13f). In this work we present an iterative procedure for dealing with a large number of scenario-based inequalities (13e)–(13f).

Since the original problem (13) with many constraints of the form (13e)–(13f) may be hard solve, a relaxation of (13) can be constructed by aggregating some of the scenario constraints. Let $\{\mathcal{S}_k : k \in \mathcal{K}\}$ denote a *partition* of the set $\mathcal{N}$ of scenario indices (which we will simply call scenario set), i.e.,

$$\bigcup_{k \in \mathcal{K}} \mathcal{S}_k = \mathcal{N}, \quad \mathcal{S}_i \cap \mathcal{S}_j = \emptyset \quad \text{for all} \quad i, j \in \mathcal{K},\ i \neq j.$$

The aggregation of scenario constraints by adding inequalities (13e) within sets $\mathcal{S}_k$ produces the following *master problem*:

$$\begin{aligned}
\min \quad & g(\mathbf{x}) && \text{(14a)}\\
\text{s. t.} \quad & \mathbf{x} \in C, && \text{(14b)}\\
& \eta + (1-\alpha)^{-1} w_0 \le h(\mathbf{x}), && \text{(14c)}\\
& w_0 \ge v^{-1}\Big(\sum_{j\in\mathcal{N}} \pi_j v(w_j)\Big), && \text{(14d)}\\
& \sum_{j\in\mathcal{S}_k} w_j \ge \sum_{j\in\mathcal{S}_k} X(\mathbf{x},\omega_j) - |\mathcal{S}_k|\eta, \quad k \in \mathcal{K}, && \text{(14e)}\\
& w_j \ge 0, \quad j \in \mathcal{N}. && \text{(14f)}
\end{aligned}$$

Clearly, any feasible solution of (13) is also feasible for (14), and the optimal value of (14) represents a lower bound for that of (13). Since the relaxed problem contains fewer scenario-based constraints (14e), it is potentially easier to solve. It would then be of interest to determine the conditions under which an optimal solution of (14) is also optimal for the original problem (13). Assuming that $\mathbf{x}^*$ is an optimal solution of (14), consider the problem

$$\begin{aligned}
\min \quad & \eta + (1-\alpha)^{-1} w_0 && \text{(15a)}\\
\text{s. t.} \quad & w_0 \ge v^{-1}\Big(\sum_{j\in\mathcal{N}} \pi_j v(w_j)\Big), && \text{(15b)}\\
& w_j \ge X(\mathbf{x}^*,\omega_j) - \eta, \quad j \in \mathcal{N}, && \text{(15c)}\\
& w_j \ge 0, \quad j \in \mathcal{N}. && \text{(15d)}
\end{aligned}$$

Proposition 1 *Consider problem* (13) *and its relaxation* (14) *obtained by aggregating scenario constraints* (13e) *over sets* $\mathcal{S}_k$, $k \in \mathcal{K}$, *that form a partition of* $\mathcal{N} = \{1,\dots,N\}$. *Assuming that* (13) *is feasible, consider problem* (15) *where* $\mathbf{x}^*$ *is an optimal solution of relaxation* (14). *Let* $(\eta^{**}, \mathbf{w}^{**})$ *be an optimal solution of* (15). *If the optimal value of* (15) *satisfies condition*

$$\eta^{**} + (1-\alpha)^{-1} w_0^{**} \le h(\mathbf{x}^*), \tag{16}$$

then $(\mathbf{x}^*, \eta^{**}, \mathbf{w}^{**})$ *is an optimal solution of the original problem* (13).

Proof: Let $\mathbf{x}^\circ$ be an optimal solution of (13). Obviously, one has $g(\mathbf{x}^*) \le g(\mathbf{x}^\circ)$. The statement of the proposition then follows immediately by observing that inequality (16) guarantees the triple $(\mathbf{x}^*, \eta^{**}, \mathbf{w}^{**})$ to be feasible for problem (13). □

The statement of Proposition 1 allows one to solve the original problem (13) by constructing an appropriate partition of $\mathcal{N}$ and solving the corresponding master problem (14). Below we outline an iterative procedure that accomplishes this goal.

Step 0: The algorithm is initialized by including all scenarios in a single partition, $\mathcal{K} = \{0\}$, $\mathcal{S}_0 = \{1,\dots,N\}$.

Step 1: For a current partition $\{\mathcal{S}_k : k \in \mathcal{K}\}$, solve the master problem (14). If (14) is infeasible, then the original problem (13) is infeasible as well, and the algorithm terminates. Otherwise, let $\mathbf{x}^*$ be an optimal solution of the master (14).

Step 2: Given a solution $\mathbf{x}^*$ of the master, solve problem (15), and let $(\eta^{**}, \mathbf{w}^{**})$ denote the corresponding optimal solution. If condition (16) is satisfied, the algorithm terminates with $(\mathbf{x}^*, \eta^{**}, \mathbf{w}^{**})$ being an optimal solution of (13) due to Proposition 1. If, however, condition (16) is violated,

$$\eta^{**} + (1-\alpha)^{-1} w_0^{**} > h(\mathbf{x}^*),$$

then the algorithm proceeds to Step 3 to update the partition.

Step 3: Determine the set of scenario-based constraints in (15) that, for a given solution of the master $\mathbf{x}^*$, are binding at optimality:

$$\mathcal{J} = \{j \in \mathcal{N} : w_j^{**} = X(\mathbf{x}^*, \omega_j) - \eta^{**} > 0\} \tag{17}$$

Then, the elements of $\mathcal{J}$ are removed from the existing sets $\mathcal{S}_k$:

$$\mathcal{S}_k = \mathcal{S}_k \setminus \mathcal{J}, \quad k \in \mathcal{K},$$

and added to the partition as single-element sets:

$$\{\mathcal{S}_0, \ldots, \mathcal{S}_K\} \cup \{S_{K+1}, \ldots, \mathcal{S}_{K+|\mathcal{J}|}\}, \text{ where } \mathcal{S}_{K+i} = \{j_i\} \text{ for each } j_i \in \mathcal{J}, \ i = 1, \ldots, |\mathcal{J}|,$$

and the algorithm proceeds to Step 1.

Theorem 3 *Assume that in problem* (13) *functions* $g(\mathbf{x})$ *and* $X(\mathbf{x}, \omega)$ *are convex in* $\mathbf{x}$*,* $h(\mathbf{x})$ *is concave in* $\mathbf{x}$*,* v *satisfies assumption (U1), and the set* C *is convex and compact. Then, the described scenario decomposition algorithm either finds an optimal solution of problem* (13) *or declares its infeasibility after at most* N *iterations.*

Proof: Let us show that during an iteration of the algorithm the size of the partition of the set $\mathcal{N}$ of scenarios increases by at least one.

Let $\{\mathcal{S}_k : k \in \mathcal{K}\}$ be the current partition of $\mathcal{N}$, $(\mathbf{x}^*, \eta^*, \mathbf{w}^*)$ be the corresponding optimal solution of (14), and $(\eta^{**}, \mathbf{w}^{**})$ be an optimal solution of (15) for the given $\mathbf{x}^*$, such that the stopping condition (16) is not satisfied,

$$\eta^{**} + (1-\alpha)^{-1} w_0^{**} > h(\mathbf{x}^*). \tag{18}$$

Let $\bar{\mathcal{S}}^*$ denote the set of constraints (15c) that are binding at optimality,

$$\bar{\mathcal{S}}^* = \{j : w_j^{**} = X(\mathbf{x}^*, \omega_j) - \eta^{**} > 0, \ j \in \mathcal{N}\}.$$

Next, consider a problem obtained from (15) with a given $\mathbf{x}^*$ by aggregating the constraints (15c) that are non-binding at optimality:

$$\min \quad \eta + (1-\alpha)^{-1} w_0 \tag{19a}$$

$$\text{s. t.} \quad w_0 \geq v^{-1}\Big(\sum_{j \in S_0} \pi_j v(w_j)\Big), \tag{19b}$$

$$w_j \geq X(\mathbf{x}^*, \omega_j) - \eta, \quad j \in \bar{\mathcal{S}}^*, \tag{19c}$$

$$\sum_{j \in \mathcal{S}^*} w_j \geq \sum_{j \in \mathcal{S}^*} X(\mathbf{x}^*, \omega_j) - |\mathcal{S}^*|\eta, \tag{19d}$$

$$w_j \geq 0, \quad j \in \mathcal{N}, \tag{19e}$$

where $\mathcal{S}^* = \mathcal{N} \setminus \bar{\mathcal{S}}^*$. Obviously, an optimal solution $(\eta^{**}, \mathbf{w}^{**})$ of (15) will also be optimal for (19).

Next, observe that at any stage of the algorithm, the partition $\{\mathcal{S}_k : k \in \mathcal{K}\}$ is such that there exists at most one set with $|\mathcal{S}_k| > 1$, namely set $\mathcal{S}_0$, and the rest of the sets in the partition satisfy $|\mathcal{S}_k| = 1$, $k \neq 0$. Let us denote

$$\bar{\mathcal{S}}_0 = \mathcal{N} \setminus \mathcal{S}_0 = \bigcup_{k \in \mathcal{K} \setminus \{0\}} \mathcal{S}_k.$$

Assume that $\bar{\mathcal{S}}^* \subseteq \bar{\mathcal{S}}_0$. By rewriting the master problem (14) as

$$\begin{aligned}
\min \quad & g(\mathbf{x}) && \text{(20a)}\\
\text{s. t.} \quad & \mathbf{x} \in C, && \text{(20b)}\\
& \eta + (1-\alpha)^{-1} w_0 \leq h(\mathbf{x}), && \text{(20c)}\\
& w_0 \geq v^{-1}\Big(\sum_{j \in \mathcal{N}} \pi_j v(w_j)\Big), && \text{(20d)}\\
& w_j \geq X(\mathbf{x}, \omega_j) - \eta, \quad j \in \bar{\mathcal{S}}_0, && \text{(20e)}\\
& \sum_{j \in \mathcal{S}_0} w_j \geq \sum_{j \in \mathcal{S}_0} X(\mathbf{x}, \omega_j) - |\mathcal{S}_0|\eta, && \text{(20f)}\\
& w_j \geq 0, \quad j \in \mathcal{N}, && \text{(20g)}
\end{aligned}$$

we observe that the components $\eta^*, \mathbf{w}^*$ of its optimal solution are feasible for (19). Indeed, from (20e) one has that

$$w_j^* \geq X(\mathbf{x}^*, \omega_j) - \eta^*, \quad j \in \bar{\mathcal{S}}^*,$$

which satisfies (19c), and also

$$w_j^* \geq X(\mathbf{x}^*, \omega_j) - \eta^*, \quad j \in \bar{\mathcal{S}}_0 \setminus \bar{\mathcal{S}}^* = \mathcal{S}^* \setminus \mathcal{S}_0.$$

Adding the last inequalities yields

$$\sum_{j \in \mathcal{S}^* \setminus \mathcal{S}_0} w_j^* \geq \sum_{j \in \mathcal{S}^* \setminus \mathcal{S}_0} X(\mathbf{x}^*, \omega_j) - |\mathcal{S}^* \setminus \mathcal{S}_0|\, \eta^*,$$

which can then be aggregated with (20f) to produce

$$\sum_{j \in \mathcal{S}^*} w_j^* \geq \sum_{j \in \mathcal{S}^*} X(\mathbf{x}^*, \omega_j) - |\mathcal{S}^*|\eta^*,$$

verifying the feasibility of $(\eta^*, \mathbf{w}^*)$ for (19). Since (20c) has to hold for $(\mathbf{x}^*, \eta^*, \mathbf{w}^*)$, we obtain that

$$\eta^{**} + (1-\alpha)^{-1} w^{**} \leq \eta^* + (1-\alpha)^{-1} w^* \leq h(\mathbf{x}^*),$$

which furnishes a contradiction with (18). Therefore, one has to have $\bar{\mathcal{S}}_0 \subset \bar{\mathcal{S}}^*$ for (18) to hold, meaning that at least one additional scenario from $\bar{\mathcal{S}}^*$ will be added to the partition during Step 3 of the algorithm. It is easy to see that the number of iterations cannot exceed the number N of scenarios. □

Remark 2 The fact that the proposed scenario decomposition method terminates within at most N iterations represents an important advantage over several existing cutting-plane methods that were developed in the literature for problems involving Conditional Value-at-Risk measure (Künzi-Bay and Mayer, 2006), integrated chance constraints (Klein Haneveld and van der Vlerk, 2006), and SSD constraints (Roman et al., 2006). In the mentioned works, the cutting-plane algorithms utilized supporting hyperplane representations for scenario constraints, which were themselves exponential in the size N of scenario sets. Although finite convergence of the cutting plane techniques was guaranteed by the polyhedral structure of the scenario constraints (in the case when $X(\mathbf{x}, \omega)$ is linear in $\mathbf{x}$), no estimate for the sufficient number of iterations was provided. A level-type regularization of cutting plane method for problems with SSD constraints, which allows for an estimate of the number of cuts due to Lemaréchal et al. (1995), is discussed in Fábián et al. (2011).

3.1 An Efficient Solution Method for Sub-Problem (15)

Although formulation (15) may be solved using appropriate mathematical programming techniques, an efficient alternative solution method can be employed by noting that (15) is equivalent to

$$\min \ \eta + \frac{1}{1-\alpha} v^{-1}\left(\sum_{j \in \mathcal{N}} \pi_j \, v\big(X(\mathbf{x}^*, \omega_j) - \eta\big)_+ \right), \tag{21}$$

which is a mathematical programming implementation of representation (3) under a finite scenario model where realizations $X(\mathbf{x}^*, \omega_j)$ represent scenario losses corresponding to an optimal decision $\mathbf{x}^*$ in the master problem (14). An optimal value of η in (15) and (21) can be computed directly using its properties dictated by representation (3).

Namely, let $X_j = X(\mathbf{x}^*, \omega_j)$ represent the optimal loss in scenario j for problem (14), and let $X_{(m)}$ be the m-th smallest outcome among $X_1, \ldots, X_N$, such that

$$X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(N)}.$$

The following proposition enables evaluation of η^{**} as a "cutoff" point within the tail of the loss distribution.

Proposition 2 *Given a function $v(\cdot)$ that satisfies (U1) and an $\alpha \in (0, 1)$, a sufficient condition for η^{**} to be an optimal solution in problems* (21) *and* (15) *has the form*

$$\frac{\sum_{j: X_j > \eta^{**}} \pi_j \, v'(X_j - \eta^{**})}{v'(v^{-1}(\sum_{j \in \mathcal{N}} \pi_j \, v(X - \eta^{**})_+))} + \alpha - 1 = 0, \tag{22}$$

where v' denotes the derivative of v.

Proof: The underlying assumption (U1) on v entails that $\phi(X) = (1-\alpha)^{-1} v^{-1} \mathsf{E} v(X)$ is convex, whence the objective function of (21)

$$\Phi_X(\eta) = \eta + \phi(X - \eta) = \eta + \frac{1}{1-\alpha} v^{-1}\Big(\sum\nolimits_{j \in \mathcal{N}} \pi_j \, v(X_j - \eta)_+ \Big) \tag{23}$$

is convex on $\mathbb{R}$. Moreover, the condition $\phi(\eta) > \eta$ for $\eta \neq 0$ of Theorem 1 guarantees that the set of minimizers of $\Phi_X(\eta)$ is compact and convex in $\mathbb{R}$. Indeed, it is easy to see that $\Phi_X(\eta) = \eta$ for $\eta \geq X_{(N)}$ and $\Phi_X(\eta) \sim -\frac{\alpha\eta}{1-\alpha}$ for $\eta \ll -1$.

Now, consider the left derivative of $\Phi_X(\eta)$ at a given point $\eta = \eta^{**}$:

$$
\begin{aligned}
&-(1-\alpha) + (1-\alpha)\frac{d^-}{d\eta}\Phi_X(\eta)\Big|_{\eta=\eta^{**}} = \frac{d^-}{d\eta}\Big\{v^{-1}\Big(\textstyle\sum_{j\in\mathcal{N}} \pi_j v(X_j - \eta)_+\Big)\Big\}\Big|_{\eta=\eta^{**}} \\
&= \lim_{\epsilon\to 0+} \frac{1}{-\epsilon}\Big\{v^{-1}\Big(\sum_{j:X_j\geq\eta^{**}} \pi_j v(X_j - \eta^{**} + \epsilon)\Big) - v^{-1}\Big(\sum_{j:X_j\geq\eta^{**}} \pi_j v(X_j - \eta^{**})\Big)\Big\} \\
&= \frac{d^-}{d\eta}\Big\{v^{-1}\Big(\sum_{j:X_j\geq\eta^{**}} \pi_j v(X_j - \eta)\Big)\Big\}\Big|_{\eta=\eta^{**}} = \frac{d}{d\eta}\Big\{v^{-1}\Big(\sum_{j:X_j\geq\eta^{**}} \pi_j v(X_j - \eta)\Big)\Big\}\Big|_{\eta=\eta^{**}},
\end{aligned}
$$

where the last equality follows from the continuous differentiability of function $v^{-1}\big(\sum_{j:X_j\geq\eta^{**}} \pi_j v(X_j - \eta^{**})\big)$ at the point η^{**} due to the assumed properties of v. Analogously, the right derivative of $\Phi_X(\eta)$ at $\eta = \eta^{**}$ equals to

$$
\frac{d^+}{d\eta}\Phi_X(\eta)\Big|_{\eta=\eta^{**}} = 1 + \frac{1}{1-\alpha}\frac{d}{d\eta}\Big\{v^{-1}\Big(\sum_{j:X_j>\eta^{**}} \pi_j v(X_j - \eta)\Big)\Big\}\Big|_{\eta=\eta^{**}},
$$

where the strict inequality in summation is due to fact that $v(X_j - \eta^{**} - \epsilon)_+ = 0$ for all $\epsilon > 0$ if $\eta^{**} \leq X_j$.

Observe that $\Phi_X(\eta)$ may only be non-differentiable at points $\eta = X_j$. Indeed, for any $\eta^{**} \neq X_j$, $j \in \mathcal{N}$, the obtained expressions for left and right derivatives become equivalent, and equation (22) is obtained from the first order optimality conditions by computing the derivatives of the functions in braces and noting that $\sum_{j:X_j\geq\eta^{**}} \pi_j v(X_j - \eta^{**}) = \sum_{j:X_j>\eta^{**}} \pi_j v(X_j - \eta^{**}) = \sum_{j\in\mathcal{N}} \pi_j v(X_j - \eta^{**})_+$. $\square$

Recall that the presented above scenario decomposition algorithm uses the subproblem (15) for determining an optimal value of η^{**}, as well as for identifying (during Step 3) the set $\mathcal{J}$ of scenarios that are binding at optimality, i.e., for which $X(\mathbf{x}^*, \omega_j) - \eta^{**} > 0$. This can be accomplished with the help of the derived optimality condition (22) as follows.

Step (i) Compute values $X_j = X(\mathbf{x}^*, \omega_j)$, where $\mathbf{x}^*$ is an optimal solution of (14), and sort them in ascending order: $X_{(1)} \leq \ldots \leq X_{(N)}$.

Step (ii) For $m = N, N-1, \ldots, 1$, compute values T_m as

$$
\begin{aligned}
&T_N = 1 - \alpha, \\
&T_m = 1 - \alpha - \frac{\sum_{j=m+1}^{N} \pi_j v'\big(X_{(j)} - X_{(m)}\big)}{v'\Big(v^{-1}\Big(\sum_{j=m+1}^{N} \pi_j v\big(X_{(j)} - X_{(m)}\big)\Big)\Big)}, \quad m = N-1, \ldots, 1,
\end{aligned}
\tag{24}
$$

until m^* is found such that

$$
T_{m^*} \leq 0, \;\; T_{m^*+1} > 0. \tag{25}
$$

Step (iii) If $T_{m^*} = 0$, then the solution η^{**} of (15), (21) is equal to $X_{(m^*)}$. Otherwise, η^{**} satisfies

$$
\eta^{**} \in \big(X_{(m^*)}, X_{(m^*+1)}\big],
$$

and its value can be found by using an appropriate numerical procedure, such as Newton's method. The set $\mathcal{J}$ in (17) is then obtained as

$$\mathcal{J} = \{j : X_j = X_{(k)},\ k = m^* + 1, \ldots, N\}.$$

Proposition 3 *Given an optimal solution* $\mathbf{x}^*$ *of the master problem* (14)*, the algorithm described in steps (i)–(iii) yields an optimal value* η^{**} *in* (15)*,* (21) *and the set* $\mathcal{J}$ *to be used during steps 2 and 3 of the scenario decomposition algorithm.*

Proof: First, observe that an optimal solution η^{**} of (15) and (21) satisfies $\eta^{**} \leq X_{(N)}$. Indeed, assume to the contrary that $\eta^{**} = X_{(N)} + \epsilon$ for some $\epsilon > 0$. The optimal value of (15) and (3) is then equal to $X_{(N)} + \epsilon$, and can be improved by selecting, e.g., $\epsilon = \epsilon/2$.

Next, observe that quantities T_m are equal, up to a factor $1-\alpha$, to the right derivatives of function $\Phi_X(\eta)$ (23) at $\eta = X_{(m)}$, i.e., $T_m = (1-\alpha)\frac{d^+}{d\eta}\Phi_X(\eta)\big|_{\eta=X_{(m)}}$. The value of $T_N = 1-\alpha$ follows directly from the fact that $\Phi_X(\eta) = \eta$ for $\eta \geq X_{(N)}$. Then, if strict inequalities in (25) hold, two cases are possible. Namely, an optimal η^{**} is located inside the interval $\big(X_{(m^*)}, X_{(m^*+1)}\big)$ if $\frac{d^-}{d\eta}\Phi_X(X_{(m^*+1)}) > 0$. Alternatively, $\eta^{**} = X_{(m^*+1)}$ if $\frac{d^-}{d\eta}\Phi_X(X_{(m^*+1)}) \leq 0$. Thus, we have the second statement of step (iii).

If $T_{m^*} = 0$ in (25), observe that necessarily $\frac{d^-}{d\eta}\Phi_X(X_{m^*}) \leq 0$ since the left derivative of Φ_X at $X_{(m)}$ differs from the expression (24) by an extra summand $\pi_m v'(0)$ in the numerator. If $v'(0) = 0$ then $\frac{d^-}{d\eta}\Phi_X(X_{m^*}) = \frac{d^+}{d\eta}\Phi_X(X_{m^*}) = 0$ and $\eta^{**} = X_{(m^*)}$ is a minimum due to Proposition 2. If $v'(0) > 0$ then $\frac{d^-}{d\eta}\Phi_X(X_{m^*}) < 0$ and $\eta^{**} = X_{(m^*)}$ is again either a unique minimizer, or represents the left endpoint of the set of minimizers. This validates the first claim of step (iii).

Once the value of η^{**} is obtained during step (iii), the set $\mathcal{J}$ in (17) is constructed as the set of scenario indices corresponding to $X_{(m^*+1)}, X_{(m^*+2)}, \ldots, X_{(N)}$.

Note that it is not necessary to prove that there always exists $m^* \in \{1, \ldots, N-1\}$ such that $T_{m^*} \leq 0$ and $T_{m^*+1} > 0$. If indeed it were to happen that $T_m > 0$ for all $m = 1, \ldots, N$, this would imply that set $\mathcal{J}$ must contain all scenarios, i.e., $\mathcal{J} = \mathcal{N}$, making the exact value of η^{**} irrelevant in this case, since the original problem (13) would have to be solved at the next iteration of the scenario decomposition algorithm. □

Remark 3 We conclude this section by noting that the presented scenario decomposition approach is applicable, with appropriate modifications, to more general forms of downside risk measures $\rho(X) = \min_\eta\{\eta + \phi((X-\eta)_+)\}$. The focus of our discussion on the case when function ϕ has the form of a certainty equivalent, $\phi(X) = v^{-1}\mathsf{E}v(X_+)$, is dictated mainly by the fact that the resulting constraint (13d) encompasses a number of interesting and practically relevant special cases, such as second-order cone, p-order cone, and log-exponential constraints.

4 Computational Experiments: Portfolio Optimization with HMCR and LogExpCR Measures

Portfolio optimization problems are commonly used as an experimental platform in risk management and stochastic optimization. In this section we illustrate the computational performance of the proposed

scenario decomposition algorithm on a portfolio optimization problem, where the investment risk is quantified using HMCR or LogExpCR measures.

A standard formulation of portfolio optimization problem entails determining the vector of portfolio weights $\mathbf{x} = (x_1, \ldots, x_n)^\top$ of n assets so as to minimize the risk while maintaining a prescribed level of expected return. We adopt the traditional definition of portfolio losses X as negative portfolio returns, $X(\mathbf{x}, \omega) = -\mathbf{r}(\omega)^\top \mathbf{x}$, where $\mathbf{r}(\omega) = (r_1(\omega), \ldots, r_n(\omega))^\top$ are random returns of the assets. Then, the portfolio selection model takes the general form

$$\begin{aligned} \min \quad & \rho\big(-\mathbf{r}(\omega)^\top \mathbf{x}\big) && \text{(26a)} \\ \text{s. t.} \quad & \mathbf{1}^\top \mathbf{x} = 1, && \text{(26b)} \\ & \mathsf{E}\big[\mathbf{r}(\omega)^\top \mathbf{x}\big] \geq \bar{r}, && \text{(26c)} \\ & \mathbf{x} \geq \mathbf{0}, && \text{(26d)} \end{aligned}$$

where $\mathbf{1} = (1, \ldots, 1)^\top$, equality (26b) represents the budget constraint, (26b) ensures a minimum expected portfolio return level, $\bar{r}$, and (26d) corresponds to no-short-selling constraints.

The distribution of the random vector $\mathbf{r}(\omega)$ of assets' returns is given by a finite set of N equiprobable scenarios $\mathbf{r}_j = \mathbf{r}(\omega_j) = (r_{1j}, \ldots, r_{nj})^\top$,

$$\pi_j = \mathsf{P}\big\{\mathbf{r} = (r_{1j}, \ldots, r_{nj})^\top\big\} = 1/N, \quad j \in \mathcal{N} \equiv \{1, \ldots, N\}. \tag{27}$$

4.1 Portfolio Optimization with Higher Moment Coherent Risk Measures

In the case when risk measure ρ in (26) is selected as a higher moment coherent risk measure, $\rho(X) = \mathsf{HMCR}_{p,\alpha}(X)$, the portfolio optimization problem (26) can be written in a stochastic programming form that is consistent with the general formulation (13) as

$$\begin{aligned} \min \quad & \eta + (1-\alpha)^{-1} w_0 && \text{(28a)} \\ \text{s. t.} \quad & w_0 \geq \|(w_1, \ldots, w_N)\|_p, && \text{(28b)} \\ & \pi_j^{-1/p} w_j \geq -\mathbf{r}_j^\top \mathbf{x} - \eta, \quad j \in \mathcal{N}, && \text{(28c)} \\ & \mathbf{x} \in C, \quad \mathbf{w} \geq \mathbf{0}, && \text{(28d)} \end{aligned}$$

where C represents a polyhedral set comprising the expected return, budget, and no-short-selling constraints on the vector of portfolio weights $\mathbf{x}$:

$$C = \Big\{\mathbf{x} \in \mathbb{R}^n \ : \ \sum\nolimits_{j \in \mathcal{N}} \pi_j \mathbf{r}_j^\top \mathbf{x} \geq \bar{r}, \ \ \mathbf{1}^\top \mathbf{x} = 1, \ \ \mathbf{x} \geq \mathbf{0}\Big\}. \tag{29}$$

Due to the presence of p-order cone constraint (28b), formulation (28) constitutes a p-order cone programming problem (pOCP).

Solution methods for problem (28) are dictated by the specific value of parameter p in (28b). As has been mentioned, in the case of $p = 1$ formulation (28) reduces to a LP problem that corresponds to a choice of risk measure as the CVaR, a case that has received a considerable attention in the literature. In view of this, of particular interest are nonlinear instances of problem (28), which correspond to values of the parameter $p \in (1, +\infty)$.

Below we consider instances of (28) with $p = 2$ and $p = 3$. In the case of $p = 2$, problem (28) can be solved using SOCP self-dual interior point methods. In the case of $p = 3$ and, generally, $p \in (1, 2) \cup (2, \infty)$, the p-cone (28b) is not self-dual, and we employ two techniques for solving (28) and the corresponding master problem (14): (i) a SOCP-based approach that relies on the fact that for a rational p, a p-order cone can be equivalently represented via a sequence of second order cones, and (ii) an LP-based approach that allows for obtaining exact solutions of pOCP problems via cutting-plane methods.

Detailed discussions of the respective formulations of problems (28) are provided below. Throughout this section, we use abbreviations in brackets to denote the different formulations of the "complete" versions of (28) (i.e., with complete set of scenario constraints (28c)). For each "complete" formulation, we also consider the corresponding scenario decomposition approach, indicated by suffix "SD". Within the scenario decomposition approach, we present formulations of the master problem (denoted by subscript "MP"); the respective subproblems are then constructed accordingly. For example, the SOCP version of the complete problem (28) with $p = 2$ is denoted [SOCP], while the same problem solved by scenario decomposition is referred to as [SOCP-SD], with the master problem being denoted as [SOCP-SD]$_{\text{MP}}$ (see below).

4.1.1 SOCP Formulation in $p = 2$ Case.

In case when $p = 2$, formulation (28) constitutes a standard SOCP problem that can be solved using a number of available SOCP solvers, such as CPLEX, MOSEK, GUROBI, etc. In order to solve it using the scenario decomposition algorithm presented in Section 3, the master problem (14) is formulated with respect to the original problem (28) with $p = 2$ as follows:

$$\begin{aligned} \min \quad & \eta + (1-\alpha)^{-1} w_0 \\ \text{s. t.} \quad & w_0 \geq \|(w_1, \ldots, w_N)\|_2, \\ & \sum_{j \in \mathcal{S}_k} \frac{\pi_j^{1/2}}{\pi^{(k)}} w_j \geq \left(\sum_{j \in \mathcal{S}_k} \frac{\pi_j}{\pi^{(k)}} \mathbf{r}_j^\top \right) \mathbf{x} - \eta, \quad k \in \mathcal{K}, \\ & \mathbf{w} \geq \mathbf{0}, \;\; \mathbf{x} \in C. \end{aligned} \qquad \text{[SOCP-SD]}_{\text{MP}}$$

Note that in the case of $\mathsf{HMCR}_{2,\alpha}$ measure, the function $v(t) = t^2$ is positive homogeneous of degree two, which allows for eliminating the scenario probabilities π_j from constraint (14d) and representing the latter in the form of a second order cone in the full formulation (28) and in the master problem [SOCP-SD]$_{\text{MP}}$. This affects constraints (14d), which then can be written in the form of the second constraint in [SOCP-SD]$_{\text{MP}}$. The subproblem (15) is reformulated accordingly.

4.1.2 SOCP Reformulation of p-Order Cone Program.

One of the possible approaches for solving the pOCP problem (28) with $p = 3$ involves reformulating the p-cone constraint (28b) via a set of quadratic cone constraints. Such an exact reformulation is possible when the parameter p has a rational value, $p = q/s$. Then, a (q/s)-order cone constraint in the positive orthant $\mathbb{R}_+^{N+1}$

$$\left\{ \mathbf{w} \geq \mathbf{0} \;:\; w_0 \geq (w_1^{q/s} + \ldots + w_N^{q/s})^{s/q} \right\} \tag{30}$$

may equivalently be represented as the following set in $\mathbb{R}_+^{N+1} \times \mathbb{R}_+^N$:

$$\left\{\mathbf{w}, \mathbf{u} \geq \mathbf{0} \ : \ w_0 \geq \|\mathbf{u}\|_1, \ \ w_j^q \leq u_j^s w_0^{q-s}, \ \ j \in \mathcal{N}\right\}. \tag{31}$$

Each of the N nonlinear inequalities in (31) can in turn be represented as a sequence of three-dimensional rotated second-order cones of the form $\xi_0^2 \leq \xi_1 \xi_2$, resulting in a SOCP reformulation of the rational-order cone (30) (Nesterov and Nemirovski, 1994; Alizadeh and Goldfarb, 2003; Krokhmal and Soberanis, 2010). Such a representation, however, is not unique and in general may comprise a varying number of rotated second order cones for a given $p = q/s$. In this case study we use the technique of Morenko et al. (2013), which allows for representing rational order p-cones with $p = q/s$ in ρ^{N+1} via $N\lceil \log_2 q \rceil$ second order cones. Namely, in the case of $p = 3$, when $q = 3$, $s = 1$, the 3-order cone (30) can equivalently be replaced with $\lceil \log_2 3 \rceil N = 2N$ quadratic cones

$$\left\{\mathbf{w}, \mathbf{u}, \mathbf{v} \geq \mathbf{0} \ : \ w_0 \geq \|\mathbf{u}\|_1, \ \ w_j^2 \leq w_0 v_j, \ \ v_j^2 \leq w_j u_j, \ \ j \in \mathcal{N}\right\}. \tag{32}$$

In accordance with the above, a p-order cone inequality in $\mathbb{R}^{N+1}$ can be represented by a set of 3D second order cone constraints and a linear inequality when p is a positive rational number. Thus, the [SpOCP] problem (28) takes the following form:

$$\begin{aligned}
\min \quad & \eta + (1-\alpha)^{-1} w_0 \\
\text{s. t.} \quad & w_0 \geq \|\mathbf{u}\|_1, \\
& w_j^2 \leq w_0 v_j, \ v_j^2 \leq w_j u_j, \quad j \in \mathcal{N}, \\
& \pi_j^{-1/p} w_j \geq -\mathbf{r}_j^\top \mathbf{x} - \eta, \quad j \in \mathcal{N}, \\
& \mathbf{x} \in C, \ \mathbf{w}, \mathbf{v}, \mathbf{u} \geq \mathbf{0}.
\end{aligned} \tag{[SpOCP]}$$

The corresponding master problem sub-problem $[\text{SpOCP-SD}]_{\text{MP}}$ in the scenario decomposition-based method is constructed by replacing constraints of the form (28c) in the last problem as follows:

$$\begin{aligned}
\min \quad & \eta + (1-\alpha)^{-1} w_0 \\
\text{s. t.} \quad & w_0 \geq \|\mathbf{u}\|_1, \\
& w_j^2 \leq w_0 v_j, \ v_j^2 \leq w_j u_j, \quad j \in \mathcal{N}, \\
& \sum_{j \in \mathcal{S}_k} \frac{\pi_j^{1-1/p}}{\pi^{(k)}} w_j \geq \left(\sum_{j \in \mathcal{S}_k} \frac{\pi_j}{\pi^{(k)}} \mathbf{r}_j^\top \right) \mathbf{x} - \eta, \quad k \in \mathcal{K}, \\
& \mathbf{x} \in C, \ \mathbf{w}, \mathbf{v}, \mathbf{u} \geq \mathbf{0}.
\end{aligned} \tag{$[\text{SpOCP-SD}]_{\text{MP}}$}$$

4.1.3 An Exact Solution Method for pOCP Programs Based on Polyhedral Approximations.

Computational methods for solving p-order cone programming problems that are based on polyhedral approximations (Krokhmal and Soberanis, 2010; Vinel and Krokhmal, 2014b) represent an alternative to interior-point approaches, and can be beneficial in situations when a pOCP problem needs to be solved repeatedly, with small variations in problem data or problem structure.

Thus, in addition to the SOCP-based approaches for solving the pOCP problem (28) discussed above, we also employ an exact polyhedral-based approach with $O(\varepsilon^{-1})$ iteration complexity that was proposed in

Vinel and Krokhmal (2014b). It consists in reformulating the p-order cone $w_0 \geq \|(w_1, \ldots, w_N)\|_p$ via a set of three-dimensional p-cones

$$w_0 = w_{2N-1}, \quad w_{N+j} \geq \|(w_{2j-1}, w_{2j})\|_p, \quad j = 1, \ldots, N-1, \tag{33}$$

and then iteratively building outer polyhedral approximations of the 3D p-cones until the solution of desired accuracy $\varepsilon > 0$ is obtained,

$$\|(w_1, \ldots, w_N)\|_p \leq (1+\varepsilon) w_0.$$

In the context of the lifted representation (33), the above ε-relaxation of p-cone inequality translates into $N-1$ corresponding approximation inequalities for 3D p-cones:

$$\|(w^*_{2j-1}, w^*_{2j})\|_p \leq (1+\epsilon) w^*_{N+j}, \quad j = 1, \ldots, N-1, \tag{34}$$

where $\epsilon = (1+\varepsilon)^{1/\lceil \log_2 N \rceil} - 1$. Then, for a given $\varepsilon > 0$, an ε-approximate solution of pOCP portfolio optimization problem (28) is obtained by iteratively solving the linear programming problem

$$\begin{aligned}
\min \quad & \eta + (1-\alpha)^{-1} w_0 \\
\text{s. t.} \quad & w_0 = w_{2N-1}, \\
& w_{N+j} \geq \alpha_p(\theta_{k_j}) w_{2j-1} + \beta_p(\theta_{k_j}) w_{2j}, \quad \theta_{k_j} \in \Theta_j, \quad j = 1, \ldots, N-1, \\
& \pi_j^{-1/p} w_j \geq -\mathbf{r}_j^\top \mathbf{x} - \eta, \quad j \in \mathcal{N}, \\
& \mathbf{x} \in C, \ \mathbf{w} \geq \mathbf{0},
\end{aligned} \tag{LpOCP}$$

where coefficients α_p and β_p are defined as

$$\alpha_p(\theta) = \frac{\cos^{p-1}\theta}{(\cos^p\theta + \sin^p\theta)^{1-\frac{1}{p}}}, \quad \beta_p(\theta) = \frac{\sin^{p-1}\theta}{(\cos^p\theta + \sin^p\theta)^{1-\frac{1}{p}}}.$$

If, for a given solution $\mathbf{w}^* = (w^*_0, \ldots, w^*_{2N-1})$ of [LpOCP], the approximation condition (34) is not satisfied for some $j = 1, \ldots, N-1$,

$$\|(w^*_{2j-1}, w^*_{2j})\|_p > (1+\epsilon) w^*_{N+j}, \tag{35}$$

then a cut of the form

$$w_{N+j} \geq \alpha_p(\theta^*_j) w_{2j-1} + \beta_p(\theta^*_j) w_{2j}, \quad \theta^*_j = \arctan \frac{w^*_{2j}}{w^*_{2j-1}}, \tag{36}$$

is added to [LpOCP]. The process is initialized with $\Theta_j = \{\theta_1\}$, $\theta_1 = \pi/4$, $j = 1, \ldots, N-1$, and continues until no violations of condition (35) are found. In Vinel and Krokhmal (2014b) it was shown that this cutting-plane procedure generates an ε-approximate solution to pOCP problem (28) within $O(\varepsilon^{-1})$ iterations.

The described cutting plane scheme can be employed to solve the master problem corresponding to the pOCP problem (28). Namely, the cutting-plane formulation of this master problem is obtained by

replacing the p-cone constraint (28b) with cutting planes similarly to [LpOCP], and the set of N scenario constraints (28c) with the aggregated constraints (compare to [SpOCP-SD]$_{\text{MP}}$):

$$\begin{aligned}
\min \quad & \eta + (1-\alpha)^{-1} t \\
\text{s. t.} \quad & w_0 = w_{2N-1}, \\
& w_{N+j} \geq \alpha_p(\theta_{k_j}) w_{2j-1} + \beta_p(\theta_{k_j}) w_{2j}, \quad \theta_{k_j} \in \Theta_j, \quad j = 1, \ldots, N-1, \\
& \sum_{j \in \mathcal{S}_k} \frac{\pi_j^{1-1/p}}{\pi^{(k)}} w_j \geq \left(\sum_{j \in \mathcal{S}_k} \frac{\pi_j}{\pi^{(k)}} \mathbf{r}_j^\top \right) \mathbf{x} - \eta, \quad k \in \mathcal{K}, \\
& \mathbf{x} \in C, \ \mathbf{w} \geq \mathbf{0}.
\end{aligned} \tag{[LpOCP-SD]$_{\text{LB}}$}$$

4.2 Portfolio Optimization with Log Exponential Convex Risk Measures

In order to demonstrate the applicability of the proposed method when solving problems with measures of risk other than the HMCR class, we examine an analogous experimental framework for instances when $\rho(X) = \mathsf{LogExpCR}_{e,\alpha}(X)$. The portfolio optimization problem (26) may then be written as

$$\begin{aligned}
\min \quad & \eta + (1-\alpha)^{-1} w_0 \\
\text{s. t.} \quad & w_0 \geq \ln \sum_{j \in \mathcal{N}} \pi_j e^{w_j}, \\
& w_j \geq -\mathbf{r}_j^\top \mathbf{x} - \eta, \quad j \in \mathcal{N}, \\
& \mathbf{x} \in C, \ \mathbf{w} \geq \mathbf{0}.
\end{aligned} \tag{[LogExpCP]}$$

Note that in contrast to pOCP and SOCP problems discussed in the preceding subsections, the above formulation is not a conic program. Since it involves a convex log-exponential constraint, we call this problem a log-exponential convex programming problem (LogExpCP) that can be solved with interior point methods.

The corresponding master problem for the scenario decomposition algorithm is obtained from [LogExpCP] by aggregating the scenario constraints in accordance to (14):

$$\begin{aligned}
\min \quad & \eta + (1-\alpha)^{-1} w_0 \\
\text{s. t.} \quad & w_0 \geq \ln \sum_{j \in \mathcal{N}} \pi_j e^{w_j}, \\
& \sum_{j \in \mathcal{S}_k} w_j \geq -\sum_{j \in \mathcal{S}_k} \mathbf{r}_j^\top \mathbf{x} - |\mathcal{S}_k| \eta, \quad k \in \mathcal{K}, \\
& \mathbf{x} \in C, \ \mathbf{w} \geq \mathbf{0}.
\end{aligned} \tag{[LogExpCP-SD]$_{\text{MP}}$}$$

In the next section we examine the computational performances within each implementation class of problem (28).

4.3 Computational Results

The portfolio optimization problems described in Section 4.1 and 4.2 were implemented in C++ using callable libraries of three solvers, CPLEX 12.5, GUROBI 5.02, and MOSEK 6. Computations ran on

a six-core 2.30GHz PC with 128GB RAM in 64-bit Windows environment. In the context of benchmarking, each adopted formulation was tested against its scenario decomposition-based implementation. Moreover, it was of particular interest to examine the performance of the scenario decomposition algorithm using various risk measure configurations, thus, the following problem settings were solved: problems [SOCP]-[SOCP-SD] with risk measure as defined by (5) for $p = 2$; problems [SpOCP]-[SpOCP-SD] and [LpOCP]-[LpOCP-SD] with measure (5) for $p = 3$; and problems [LogExpCP]-[LogExpCP-SD] with risk measure (6). The value of parameter α in the employed risk measures was fixed at $\alpha = 0.9$ throughout.

The scenario data in our numerical experiments was generated as follows. First, a set of n stocks ($n = 50$, 100, 200) was selected at random from the S&P500 index. Then, a covariance matrix of daily returns as well as the expected returns were estimated for the specific set of n stocks using historical prices from January 1, 2006 to January 1, 2012. Finally, the desired number N of scenarios, ranging from 1,000 to 100,000, have been generated as N independent and identically distributed samples from a multivariate normal distribution with the obtained mean and covariance matrix.

On account of precision arithmetic errors associated with the numerical solvers, we introduced a tolerance level $\epsilon > 0$ to specify the permissible gap in the stopping criterion (16):

$$\eta^{**} + (1-\alpha)^{-1} w_0^{**} \leq h(\mathbf{x}^*) + \epsilon. \tag{37}$$

Specifically, the value $\epsilon = 10^{-5}$ was was chosen to match the reduced cost of the simplex method in CPLEX and GUROBI. In a similar manner, we adjust (24) around m^* for precision errors as

$$T_{m^*+1}(p) - \epsilon < 0 \quad \text{and} \quad T_{m^*}(p) + \epsilon > 0.$$

Empirical observations suggest the accumulation of numerical errors is exacerbated by the use of fractional values of scenarios in assets returns, r_{ij}. To alleviate the numerical accuracy issues, the data in respective problem instances of the scenario decomposition algorithm were appropriately scaled.

The results of our numerical experiments are summarized in Tables 1 – 5. Unless stated otherwise, the reported running time values are averaged over 20 instances. Table 1 presents the computational times observed during solving the full formulation, [SOCP], of problem (28) with HMCR measure and $p = 2$, and solving the same problem using the scenario decomposition algorithm, [SOCP-SD], with the three solvers, CPLEX, GUROBI, and MOSEK. Observe that the scenario decomposition method performs better for all instances and solvers, with the exception of the largest three scenario instances when using GUROBI with $n = 50$ assets. However, this trend is tampered as the number of assets increases.

Table 2 reports the running times observed during solving of the second-order cone reformulation of the pOCP version of problem (28) with $p = 3$, in the full formulation ([SpOCP]) and via the scenario decomposition algorithm ([SpOCP-SD]). The obtained results indicate that, although the scenario decomposition algorithm is slower on smaller problem instances, it outperforms direct solution methods as the numbers of scenarios N and assets n in the problem increase. Due to observed numerical instabilities, the CPLEX solver was not considered for this particular experiment.

Next, the same problem is solved using using the polyhedral approximation cutting-plane method described in Section 4.1. Table 3 shows the running times achieved by all three solvers for problems [LpOCP] and [LpOCP-SD] with $p = 3$. In this case, the scenario decomposition method resulted in order-of-magnitude improvements, which can be attributed to the "warm-start" capabilities of CPLEX and GUROBI's simplex solvers. Consistent with these conclusions is also the fact that the simplex-based solvers of CPLEX and GUROBI yield improved solution times on the full problem formulation

n	N	CPLEX [SOCP]	CPLEX [SOCP-SD]	GUROBI [SOCP]	GUROBI [SOCP-SD]	MOSEK [SOCP]	MOSEK [SOCP-SD]
50	1000	1.00	0.46	0.62	0.45	0.26	0.15
	2500	3.03	0.51	1.88	1.07	0.60	0.36
	5000	6.58	0.55	3.81	2.78	1.24	0.72
	10000	13.72	1.35	9.56	7.89	2.56	1.61
	25000	31.03	3.53	32.40	34.04	7.33	5.18
	50000	60.62	9.05	101.09	117.24	17.64	12.43
	100000	137.14	25.25	327.95	449.78	36.78	33.02
100	1000	2.46	0.86	1.73	0.42	0.61	0.18
	2500	6.14	0.99	4.87	1.17	1.50	0.47
	5000	13.69	1.10	11.13	3.55	3.25	1.15
	10000	27.06	2.21	21.94	9.63	6.69	3.03
	25000	72.95	8.85	71.34	37.48	20.41	6.88
	50000	157.25	20.88	185.56	129.37	44.01	16.61
	100000	319.90	58.29	464.12	467.35	79.75	41.58
200	1000	6.87	2.19	5.60	0.58	6.68	0.29
	2500	17.48	2.10	15.36	1.37	4.49	0.73
	5000	34.93	2.98	33.96	4.15	9.36	1.92
	10000	76.13	5.03	63.67	16.50	19.54	5.51
	25000	206.29	24.16	196.45	54.00	53.89	29.15
	50000	447.85	55.93	438.40	152.76	112.47	28.85
	100000	950.17	112.60	998.86	539.46	234.68	61.98

Table 1: Average computation times (in seconds) obtained by solving problems [SOCP] and [SOCP-SD] for $p = 2$ using CPLEX, GUROBI and MOSEK. All running times are averaged over 20 instances.

comparing to the SOCP-based reformulation [SpOCP], where barrier solvers were invoked. The discrepancy between [LpOCP] and [LpOCP-SD] solution times is especially prominent for MOSEK, but in this case it appears that MOSEK's interior-point LP solver was much less effective at solving the [LpOCP] formulation using the cutting plane method.

Finally, Table 4 displays the running times for the discussed implementation of problems [LogExpCR] and [LogExpCP-SD]. Of the three solvers considered in this case study, only MOSEK was capable of handling problems with constraints that involve sums of univariate exponential functions. Again, the scenario decomposition-based solution method appears to be preferable in comparison to solving the full formulation. Note, however, that computational times were not averaged over 20 instances in this case due to numerical difficulties associated with the solver for many instances of [LogExpCP].

It is also of interest to comment on the number of scenarios that had to be generated during the scenario decomposition procedure in order to yield an optimal solution. Table 5 lists the corresponding average number of scenarios partitioned for each problem type over all instances. Although these numbers may slightly differ among the three solvers, we only present results for MOSEK as it was the only solver used to solve all the problem in Sections 4.1 and 4.2. Observe that far fewer scenarios are required relative to the total set size N. In fact, as a percentage of the total number of scenarios, the number of scenarios that were generated during the algorithm in order to achieve optimality was between 0.7% and 11% of the total scenario set size.

n	N	GUROBI [SpOCP]	GUROBI [SpCOP-SD]	MOSEK [SpOCP]	MOSEK [SpCOP-SD]
50	1000	2.58	2.73	0.18	0.63
	2500	10.63	6.61	0.49	0.96
	5000	32.01	19.27	1.06	1.70
	10000	87.27	41.34	2.31	3.49
	25000	198.56	92.39	7.14	6.70
	50000	455.63	540.09	16.36	13.70
	100000	1217.96	2080.34	35.33	30.29
100	1000	7.16	3.14	0.30	0.75
	2500	29.47	8.44	0.85	1.37
	5000	90.25	19.74	1.88	2.32
	10000	277.72	44.31	4.52	3.91
	25000	642.63	92.11	12.66	8.66
	50000	1365.37	1716.37	28.64	15.10
	100000	—	—	65.48	28.29
200	1000	17.86	3.87	0.69	1.01
	2500	78.28	8.65	1.90	1.56
	5000	276.89	22.40	4.41	2.47
	10000	799.65	49.02	9.88	4.84
	25000	2118.11	107.14	29.99	9.60
	50000	—	—	64.52	17.41
	100000	—	—	139.87	34.99

Table 2: Average computation times (in seconds) obtained by solving problems [SpCOP] and [SpCOP-SD] for $p = 3$ using GUROBI and MOSEK. All running times are averaged over 20 instances and symbol "—" indicates that the time limit of 3600 seconds was exceeded.

5 Conclusions

In this work, we propose an efficient algorithm for solving large-scale convex stochastic programming problems that involve a class of risk functionals in the form of infimal convolutions of certainty equivalents. We exploit the property induced by such risk functionals that a significant portion of scenarios is not required to obtain an optimal solution. The developed scenario decomposition technique is contingent on the identification and separation of "non-redundant" scenarios by solving a series of smaller relaxation problems. It is shown that the number of iterations of the algorithm is bounded by the number of scenarios in the problem. Numerical experiments with portfolio optimization problems based on simulated return data following the covariance structure of randomly chosen S&P500 stocks demonstrate that significant reductions in solution times may be achieved by employing the proposed algorithm. Particularly, performance improvements were observed for the large-scale instances when using HMCR measures with $p = 2, 3$, and LogExpCR measures.

Acknowledgements This work was supported in part by the AFOSR grant FA9550-12-1-0142 and the U.S. Department of Air Force grant FA8651-12-2-0010. In addition, support by the AFRL Mathematical Modeling and Optimization Institute is gratefully acknowledged.

n	N	CPLEX [LpOCP]	CPLEX [LpOCP-SD]	GUROBI [LpOCP]	GUROBI [LpOCP-SD]	MOSEK [LpOCP]	MOSEK [LpOCP-SD]
50	1000	0.27	0.12	0.22	0.59	0.82	0.46
	2500	1.65	0.24	0.74	0.83	4.26	0.66
	5000	6.81	0.46	2.31	1.54	15.08	1.46
	10000	19.20	1.42	7.73	3.86	60.66	3.75
	25000	31.93	3.93	56.52	13.74	381.67	11.34
	50000	179.49	16.07	117.72	36.51	1412.81	25.47
	100000	903.36	62.79	474.68	112.72	—	54.45
100	1000	0.37	0.13	0.23	0.61	2.94	0.65
	2500	2.22	0.28	0.86	0.98	7.11	1.06
	5000	8.58	0.79	2.82	1.76	32.20	1.95
	10000	28.71	2.18	9.28	4.13	122.75	4.99
	25000	45.37	4.99	35.11	13.13	1138.99	15.34
	50000	200.12	18.80	122.21	39.78	2753.54	34.17
	100000	3336.26	82.79	1316.29	138.74	—	80.15
200	1000	0.61	0.20	0.33	0.89	15.68	1.06
	2500	3.13	0.44	1.30	1.17	20.64	1.37
	5000	13.25	1.01	3.72	2.11	70.49	2.97
	10000	47.97	3.31	13.20	4.72	322.36	8.12
	25000	195.28	6.98	94.45	14.77	2418.52	26.91
	50000	936.60	27.20	665.61	45.43	—	53.62
	100000	—	114.08	3301.44	160.92	—	123.89

Table 3: Average computation times (in seconds) obtained by solving problems [LpOCP] and [LpOCP-SD] for $p = 3$ using CPLEX, GUROBI and MOSEK. All running times are averaged over 20 instances and symbol "—" indicates that the time limit of 3600 seconds was exceeded.

References

Alizadeh, F. and Goldfarb, D. (2003) "Second-order cone programming," *Mathematical Programming*, **95** (1), 3–51.

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

Ben-Tal, A. and Teboulle, M. (2007) "An Old-New Concept of Convex Risk Measures: An Optimized Certainty Equivalent," *Mathematical Finance*, **17** (3), 449–476.

Delbaen, F. (2002) "Coherent risk measures on general probability spaces," in: K. Sandmann and P. J. Schönbucher (Eds.) "Advances in Finance and Stochastics: Essays in Honour of Dieter Sondermann," 1–37, Springer.

Espinoza, D. and Moreno, E. (2012) "Fast sample average approximation for minimizing Conditional-Value-at-Risk," *Preprint Paper*.

Fábián, C. I., Mitra, G., Roman, D., and Zverovich, V. (2011) "An enhanced model for portfolio choice with SSD criteria: a constructive approach," *Quantitative Finance*, **11**, 1525–1534.

Klein Haneveld, W. K. and van der Vlerk, M. H. (2006) "Integrated chance constraints: reduced forms and an algorithm," *Computational Management Science*, **3**, 245–269.

n	N	MOSEK [LogExpCP]	MOSEK [LogExpCP-SD]	Instances Solved
50	1000	0.61	0.27	12
	2500	0.97	0.58	14
	5000	1.89	1.18	12
	10000	4.88	2.57	9
	25000	14.99	7.94	12
	50000	26.65	18.76	15
	100000	65.45	61.48	17
100	1000	0.57	0.25	17
	2500	1.65	0.53	16
	5000	3.69	1.14	10
	10000	9.18	2.53	15
	25000	24.61	13.83	13
	50000	50.66	39.72	19
	100000	148.54	59.02	16
200	1000	5.25	0.37	19
	2500	4.22	0.75	17
	5000	9.53	1.39	18
	10000	21.17	2.63	17
	25000	62.03	7.59	17
	50000	145.89	16.47	18
	100000	333.73	43.56	19

Table 4: Average computation times (in seconds) obtained by solving a specified number of instances for problems [LogExpCP] and [LogExpCP-SD] using MOSEK solver.

Krokhmal, P. (2007) "Higher Moment Coherent Risk Measures," *Quantitative Finance*, **7** (4), 373–387.

Krokhmal, P. and Soberanis, P. (2010) "Risk optimization with p-order conic constraints: A linear programming approach," *European Journal of Operational Research*, **301** (3), 653–671.

Krokhmal, P., Zabarankin, M., and Uryasev, S. (2011) "Modeling and optimization of risk," *Surveys in Operations Research and Management Science*, **16** (2), 49–66.

Künzi-Bay, A. and Mayer, J. (2006) "Computational aspects of minimizing conditional value-at-risk," *Computational Management Science*, **3** (1), 3–27.

Lemaréchal, C., Nemirovskii, A., and Nesterov, Y. (1995) "New variants of bundle methods," *Mathematical Programming*, **69**, 111–147.

Lim, C., Sherali, H. D., and Uryasev, S. (2010) "Portfolio optimization by minimizing conditional value-at-risk via nondifferentiable optimization," *Computational Optimization and Applications*, **46** (3), 391–415.

Morenko, Y., Vinel, A., Yu, Z., and Krokhmal, P. (2013) "On p-cone linear discrimination," *European Journal of Operational Research*, **231**, 784–789.

Nesterov, Y. E. and Nemirovski, A. (1994) *Interior Point Polynomial Algorithms in Convex Programming*, volume 13 of *Studies in Applied Mathematics*, SIAM, Philadelphia, PA.

n	N	MOSEK [SOCP-SD]	[SpOCP-SD]	[LpOCP-SD]	[LogExpCP-SD]
50	1000	80.3	24.8	21.3	61.8
	2500	180.8	47.8	47.0	77.8
	5000	349.3	80.3	79.0	104.6
	10000	711.6	133.4	128.3	154.3
	25000	1834.9	232.0	318.3	178.2
	50000	3582.1	445.4	675.0	841.7
	100000	6945.1	774.1	1346.4	1447.5
100	1000	87.2	32.0	27.0	81.4
	2500	191.2	73.6	74.1	107.8
	5000	367.6	107.4	102.4	192.2
	10000	711.1	148.9	156.9	229.7
	25000	1808.6	278.1	348.6	1869.1
	50000	3802.9	457.8	729.7	2418.6
	100000	7323.3	831.3	1395.8	923.4
200	1000	108.2	39.5	36.4	100.7
	2500	201.7	72.7	73.0	154.5
	5000	395.6	116.3	119.6	198.1
	10000	744.0	184.9	171.2	304.6
	25000	1805.5	308.3	347.0	464.2
	50000	3607.8	512.2	697.6	788.1
	100000	7198.9	865.0	1384.3	1153.5

Table 5: Average number of partitioned scenarios from solving the scenario decomposition-based problems listed in Section 4.1 and 4.2.

Rockafellar, R. T. and Uryasev, S. (2000) "Optimization of Conditional Value-at-Risk," *Journal of Risk*, **2**, 21–41.

Rockafellar, R. T. and Uryasev, S. (2002) "Conditional Value-at-Risk for General Loss Distributions," *Journal of Banking and Finance*, **26** (7), 1443–1471.

Roman, D., Darby-Dowman, K., and Mitra, G. (2006) "Portfolio construction based on stochastic dominance and target return distributions," *Mathematical Programming*, **108**, 541–569.

Ruszczyński, A. and Shapiro, A. (2006) "Optimization of Convex Risk Functions," *Mathematics of Operations Research*, **31** (3), 433–452.

Schied, A., A.ed and Follmer, H. (2002) "Robust preferences and convex measures of risk," in: K. Sandmann and P. J. Schönbucher (Eds.) "Advances in Finance and Stochastics: Essays in Honour of Dieter Sondermann," 39–56, Springer.

Shapiro, A., Dentcheva, D., and Ruszczyński (2009) *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, Philadelphia.

Subramanian, D. and Huang, P. (2008) "A Novel Algorithm for Stochastic Linear Programs with Conditional-value-at-risk (CVaR) Constraints," *IBM Research Report, RC24752*.

Subramanian, D. and Huang, P. (2009) "An Efficient Decomposition Algorithm for Static, Stochastic, Linear and Mixed-Integer Linear Programs with Conditional-Value-at-Risk Constraints," *IBM Research Report, RC24752*.

Uryasev, S. and Rockafellar, R. T. (2013) "The fundamental risk quadrangle in risk management, optimization and statistical estimation," *Surveys in Operations Research and Management Science*, **18**, 33–53.

Vinel, A. and Krokhmal, P. (2014a) "On valid inequalities for mixed integer p-order cone programming," *Journal of Optimization Theory and Applications*, **160** (2), 439–456.

Vinel, A. and Krokhmal, P. (2014b) "Polyhedral approximations in p-order cone programming," *Optimization Methods & Software*, **29**, 1210–1237.

Vinel, A. and Krokhmal, P. (2015) "Certainty equivalent measures of risk," *Annals of Operations Research*, DOI:10.1007/s10479–015–1801–0.

von Neumann, J. and Morgenstern, O. (1944) *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, 1953rd edition.