On Risk-Averse Stochastic Semidefinite Programs with Continuous Recourse

Matthias Claus · Rüdiger Schultz ·
Kai Spürkel · Tobias Wollenberg

Abstract The vast majority of the literature on stochastic semidefinite programs (stochastic SDPs) with recourse is concerned with risk-neutral models. In this paper, we introduce mean-risk models for stochastic SDPs and study structural properties as convexity and (Lipschitz) continuity. Special emphasis is placed on stability with respect to changes of the underlying probability distribution. Perturbations of the true distribution may arise from incomplete information or working with (finite discrete) approximations for the sake of computational efficiency. We discuss extended formulations for stochastic SDPs under finite discrete distributions, which turn out to be deterministic (mixed-integer) SDPs that are (almost) block-structured for many popular risk measures.

1 Introduction

Stochastic semidefinite programs with recourse were first considered by Ariyawansa and Zhu in [1], where, for finite discrete distributions, the authors reformulate the risk-neutral stochastic SDP as a block-structured deterministic SDP and discuss an application to the stochastic version of the minimum-volume covering ellipsoid problem (cf. [22], [24]). In [27], the same authors give a multitude of other applications, including problems in geometry, location aided routing, RC circuit design and structural optimization.

M. Claus · R. Schultz · K. Spürkel · T. Wollenberg
University Duisburg-Essen
Thea-Leymann-Strae 9
D-45127 Essen
Tel.: +49 201 183 6887
E-mail: matthias.claus@uni-due.de

Some approaches to the algorithmic treatment of risk neutral programs with linear recourse carry over to expectation based stochastic SDPs. Extending the results of Zhao (cf. [25]), Mehrotra and Özevin derive a polynomial logarithmic barrier algorithm employing Bender's decomposition (cf. [13]). Using the volumetric barrier of Vaidya (cf. [23]), Ariyawansa and Zhu construct algorithms of similar complexity in [2]. Furthermore, in [10], Jin, Ariyawansa and Zhu propose homogeneous self-dual algorithms with complexities comparable to the ones of the methods mentioned before. Motivated by an application in multi-antenna wireless networks, Gaujal and Mertikopoulos establish a stochastic approximation algorithm in [9].

Chance constrained SDP models have been introduced by Ariyawansa and Zhu in [26, Chapter 3], where an application to the stochastic minimum-volume covering ellipsoid problem is considered. A different approach towards risk-aversion is taken by Schultz and Wollenberg, who consider stochastic mixed-integer semidefinite programs arising from unit commitment problems in AC transmission systems. Based on Lagrangian relaxation of the nonanticipativity constraint, a decomposition algorithm for minimizing a weighted sum of the expectation and the probability of exceeding a certain threshold is proposed in [19].

The present work extends the models of [19] and [2] by considering more general risk measures. Instead of focussing on a certain application, we discuss structural properties as convexity and (Lipschitz) continuity of the resulting objective functions. Consequences for quantitative stability of the stochastic SDP models under perturbations of the underlying distribution are pointed out. Such perturbations may arise from incomplete information about the distribution or the choice to work with a simpler (possibly finite discrete) approximation for reasons of computational efficiency.

Furthermore, we establish sufficient conditions for differentiabiliy in the risk neutral setting. Finally, for finite discrete distributions, we establish equivalent SDPs for various risk measures and give indications on how to exploit their special structure for numerical treatment.

2 Two-Stage Stochastic SDPs with Continuous Recourse

Let $\mathcal{S}_+^k$ denote the cone of symmetric positive semidefinite matrices in $\mathbb{R}^{k \times k}$. The componentwise Frobenius product of $A = (a_1, \ldots a_s)^\top \in (\mathcal{S}_+^k)^l$ and $x \in \mathcal{S}_+^k$ is defined as $A \bullet x := (\mathrm{tr}(a_1 x), \ldots, \mathrm{tr}(a_s x))^\top \in \mathbb{R}^s$. Furthermore, the Frobenius norm on $\mathcal{S}_+^k$ is given by $\|x\| := \sqrt{x \bullet x}$.

We shall consider the parametric SDP

$$(\mathbf{P}(z)) \qquad \min_{x,y}\{c \bullet x + q \bullet y \mid T \bullet x + W \bullet y = z,\ x \in X,\ y \in \mathcal{S}_+^m\},$$

where $z \in \mathbb{R}^s$ enters as a parameter. The data is comprised of $c \in \mathcal{S}_+^n$, $q \in \mathcal{S}_+^m$, $T \in (\mathcal{S}_+^n)^s$, $W \in (\mathcal{S}_+^m)^s$ and a nonempty, closed, convex set $X \subseteq \mathcal{S}_+^n$. The set X is usually given as a spectrahedron, i.e. the intersection of the solution sets of a finite number of affine matrix inequalities with the cone of positive semidefinite matrices.

Let $z = Z(\omega)$ be the realization of a random vector $Z : \Omega \to \mathbb{R}^s$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A two-stage stochastic SDP arises from $(\mathbf{P}(z))$ if the decision x

has to be taken without knowledge of the particular realization $Z(\omega)$, while y can be chosen after observing the previously unknown parameter. In this setting, the optimal decision y is governed by the recourse problem

$$\min_{y}\{q \bullet y \mid W \bullet y = Z(\omega) - T \bullet x, \; y \in \mathcal{S}_+^m\}. \tag{1}$$

Let $\varphi : \mathbb{R}^s \to \overline{\mathbb{R}}$ denote the optimal value function of (1) with respect to the right-hand side of the system of matrix equations in its constraints, i.e.

$$\varphi(t) := \min_{y}\{q \bullet y \mid W \bullet y = t, \; y \in \mathcal{S}_+^m\}.$$

Introducing the function $f : \mathcal{S}_+^n \times \mathbb{R}^s \to \overline{\mathbb{R}}$, $f(x,z) := c \bullet x + \varphi(z - T \bullet x)$ we may rewrite $(\mathbf{P}(Z(\cdot))$ as

$$\min_{x}\{f(x, Z(\cdot)) \mid x \in X\}. \tag{2}$$

Due to the assumed interplay between decision and observation, problem (2) is not well-defined without further modelling choices. For any x, $f(x, Z(\cdot))$ belongs to the space $L^0(\Omega, \mathcal{F}, \mathbb{P})$ of extended real-valued random variables on the underlying probability space. We thus may fix any functional $\mathcal{R} : \mathcal{X} \to \overline{\mathbb{R}}$ satisfying

$$\{f(x, Z(\cdot)) \mid x \in X\} \subseteq \mathcal{X} \subseteq L^0(\Omega, \mathcal{F}, \mathbb{P})$$

and consider the optimization problem

$$\min_{x}\{Q_{\mathcal{R}}(x) \mid x \in X\}, \tag{3}$$

where the mapping $Q_{\mathcal{R}} : \mathcal{S}_+^n \to \overline{\mathbb{R}}$ is given by $Q_{\mathcal{R}}(x) = \mathcal{R}[f(x, Z(\cdot))]$.

We shall work with the following assumptions:

A1 (**Complete recourse**) $W \bullet \mathcal{S}_+^m = \mathbb{R}^s$.

A2 (**Strict dual feasibility**) There is some $u \in \mathbb{R}^s$ such that $q - W^\top u$ is positive definite.

Similar, yet more restrictive assumptions are also made in [13].

Lemma 1 *Assume A2, then A1 holds if and only if* $M_D := \{u \in \mathbb{R}^s \mid q - W^\top u \in \mathcal{S}_+^m\}$ *is compact.*

Proof M_D is closed due to the closedness of S_+^m. Suppose that M_D is unbounded, i.e. that there exists a sequence $\{u_k\}_{k \in \mathbb{N}} \subseteq M_D$ with $\lim_{k \to \infty} \|u_k\| = \infty$. Define $v_k := u_k / \|u_k\|$, then $\|v_k\| = 1$ holds for all $k \in \mathbb{N}$. Therefore, the sequence $\{v_k\}_{k \in \mathbb{N}}$ can be assumed to converge to some $v \neq 0$ without loss of generality. By $u_k \in M_D$ we have $q - W^\top u_k \in \mathcal{S}_+^m$ for all $k \in \mathbb{N}$. Thus,

$$-W^\top v = \lim_{k \to \infty} -W^\top v_k = \lim_{k \to \infty} \frac{1}{\|u_k\|}\left(q - W^\top u_k\right) \in \mathcal{S}_+^m.$$

Now select any $u_0 \in M_D$. Then $u_0 + \alpha v \in M_D$ holds for any $\alpha \geq 0$ and we have

$$\lim_{\alpha \to \infty} v^\top(u_0 + \alpha v) = \lim_{\alpha \to \infty} v^\top u_0 + \alpha \|v\|^2 = \infty,$$

verifying $\sup\{v^\top u \mid q - W^\top u \in \mathcal{S}_+^m\} = \infty$. By duality, the set $\{y \in \mathcal{S}_+^m \mid W \bullet y = v\}$ has to be empty, which contradicts A1.

Let M_D be compact, then once again by duality for arbitrary $t \in \mathbb{R}^s$, there exists $u \in M_D$ with $\min\{q \bullet y \mid W \bullet y = t, \; y \in \mathcal{S}_+^m\} = t^\top u$, which implies $t \in W \bullet \mathcal{S}_+^m$ and thus A1. $\qquad \square$

The lemma above shows that $\sup\{t^\top u \mid q - W^\top u \in \mathcal{S}_+^m\}$ is attained for any $t \in \mathbb{R}^s$ whenever A1 and A2 hold true.

Lemma 2 *Assume A1 and A2, then φ is finite, convex and Lipschitz continuous on $\mathbb{R}^s$.*

Proof Due to A1 and A2, strong duality holds true for the SDP defining φ. We thus have

$$\varphi(t) = \max_u \{t^\top u \mid u \in M_D\} \ \ \forall t \in \mathbb{R}^s.$$

As M_D is nonempty and compact by Lemma 1, φ is finite on $\mathbb{R}^s$.

Furthermore, for arbitrary $\lambda \in [0,1]$ and $t_1, t_2 \in \mathbb{R}^s$, strong duality implies

$$\begin{aligned}
\varphi(\lambda t_1 + (1-\lambda)t_2) &= \max_{u \in M_D} (\lambda t_1 + (1-\lambda)t_2)^T u \\
&\leq \lambda \max_{u \in M_D} t_1^T u + (1-\lambda) \max_{u \in M_D} t_2^T u \\
&= \lambda \varphi(t_1) + (1-\lambda)\varphi(t_2),
\end{aligned}$$

which proves the asserted convexity of φ.

To establish Lipschitz continuity, let $t_1, t_2 \in \mathbb{R}^s$ be arbitrary and fixed. Then by strong duality and the compactness of M_D, there exists $u_1, u_2 \in M_D$ such that $\varphi(t_1) = t_1^\top u_1$ and $\varphi(t_2) = t_2^\top u_2$. By $t_1^\top u_1 \geq t_1^\top u_2$ and $t_2^\top u_2 \geq t_2^\top u_1$ we have

$$-\|u_2\| \cdot \|t_1 - t_2\| \leq t_1^\top u_2 - t_2^\top u_2 \leq \varphi(t_1) - \varphi(t_2) \leq t_1^\top u_1 - t_2^\top u_1 \leq \|u_1\| \cdot \|t_1 - t_2\|$$

and thus $|\varphi(t_1) - \varphi(t_2)| \leq \max\{\|u_1\|, \|u_2\|\}\|t_1 - t_2\|$. Set $L_\varphi := \max_{u \in M_D} \|u\| < \infty$, then

$$|\varphi(t_1) - \varphi(t_2)| \leq L_\varphi \cdot \|t_1 - t_2\|$$

holds for all $t_1, t_2 \in \mathbb{R}^s$, which completes the proof.   □

Remark 1 Under assumptions A1 and A2, φ is finite and convex, which implies directional differentiability by [16, Theorem 25.4]. Furthermore, the subdifferential of φ is convex, compact and admits the representation

$$\partial\varphi(t) = \text{Argmax}\{u^\top t \mid u \in M_D\}.$$

By [16, Theorem 25.1], φ is differentiable at t if and only if $\partial\varphi(t)$ is a singleton. In that case, we have $\partial\varphi(t) = \{\nabla\varphi(t)\}$.

Remark 2 In two-stage stochastic linear programming, the counterpart of φ is the optimal value function of a linear program:

$$\varphi_l : \mathbb{R}^s \to \overline{\mathbb{R}}, \ \ \varphi_l(t) := \min\{q_l^\top y_l \mid W_l y_l = t, \ y_l \in \mathbb{R}_+^m\}$$

with $q_l \in \mathbb{R}^m$ and $W_l \in \mathbb{R}^{s \times m}$. By linear programming theory, φ_l is finite on $\mathbb{R}^s$ iff $W_l(\mathbb{R}_+^m) = \mathbb{R}^s$ and $M_{D_l} = \{u \in \mathbb{R}^s \mid W_l^\top u \leq q\} \neq \emptyset$. In this situation, φ_l admits the representation

$$\varphi_l(t) = \max_{j=1,\dots,N} d_j^\top t,$$

where $d_1, \dots, d_N$ denote the vertices of the polytope M_{D_l}. In particular, φ_l is piecewise linear, convex and Lipschitz continuous.

The following example shows that the assumptions A1 and $M_D \neq \emptyset$ are not sufficient to ensure that the optimal value in the problem defining $\varphi(t)$ is attained for all $t \in \mathbb{R}^s$.

Example 1 For $t \in \mathbb{R}$, consider the SDP

$$\min\left\{ \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \bullet y \mid \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \bullet y = t,\ y \in \mathcal{S}_+^2 \right\}. \tag{4}$$

For any $t \in \mathbb{R}$ we have

$$\begin{bmatrix} |t|+1 & t \\ t & |t|+1 \end{bmatrix} \in \operatorname{int} \mathcal{S}_+^2 \ \text{ and } \ \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \bullet \begin{bmatrix} |t|+1 & t \\ t & |t|+1 \end{bmatrix} = t.$$

Consequently, A1 is fulfilled. Moreover, we have

$$M_D = \left\{ u \in \mathbb{R} \mid \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \cdot u \in \mathcal{S}_+^2 \right\} = \{0\}. \tag{5}$$

As (4) is strictly feasible for any right-hand side $t \in \mathbb{R}^s$, strong duality holds and (5) implies that the infimum of (4) is zero. Furthermore, for any $t \in \mathbb{R} \setminus \{0\}$ we have

$$\begin{bmatrix} y_{11} & \frac{t}{2} \\ \frac{t}{2} & y_{22} \end{bmatrix} \in \mathcal{S}_+^2 \ \Leftrightarrow \ y_{11} > 0,\ y_{22} > 0,\ y_{11}y_{22} - \left(\frac{t}{2}\right)^2 \geq 0,$$

which yields the lower bound $y_{11} \geq t^2/(4y_{22}) > 0$ for any y that is feasible for (4). Consequently, the optimal value in (4) is not attained if $t \neq 0$.

3 Structure of Risk-Averse Stochastic SDPs

Let us now return to problem (3) and consider various choices of $\mathcal{R}$. To ensure finiteness, we shall work with moment conditions on the Borel probability measure $\mathbb{P} \circ Z^{-1}$ induced by the underlying random vector $Z(\cdot)$. Let $\mathcal{P}(\mathbb{R}^s)$ denote the space of all Borel probability measures on $\mathbb{R}^s$ and

$$\mathcal{M}_s^p := \{ \mu \in \mathcal{P}(\mathbb{R}^s) \mid \int_{\mathbb{R}^s} \|t\|^p\ \mu(dt) < \infty \}$$

be the subspace of measures having finite moments of order $p \geq 1$.

Lemma 3 *Assume A1, A2 and $\mathbb{P} \circ Z^{-1} \in \mathcal{M}_s^1$. Then $f(x, Z(\cdot)) \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ for all $x \in \mathcal{S}_+^n$ and the mapping $F : \mathcal{S}_+^m \to L^1(\Omega, \mathcal{F}, \mathbb{P})$, $F(x) := f(x, Z(\cdot))$ is convex and Lipschitz continuous with constant $\|c\| + L_\varphi \cdot \|T\|$.*

Proof For any $x \in \mathcal{S}_+^n$ we have

$$\begin{aligned} \|F(x)\|_{L^1} &= \int_{\mathbb{R}^s} |c \bullet x + \varphi(z - T \bullet x)|\ (\mathbb{P} \circ Z^{-1})(dz) \\ &\leq |c \bullet x| + |\varphi(0)| + \int_{\mathbb{R}^s} |\varphi(z - T \bullet x) - \varphi(0)|\ (\mathbb{P} \circ Z^{-1})(dz) \\ &\leq |c \bullet x| + |\varphi(0)| + L_\varphi \|T \bullet x\| + L_\varphi \int_{\mathbb{R}^s} \|z\|\ (\mathbb{P} \circ Z^{-1})(dz) < \infty \end{aligned}$$

by Lemma 2.

For any $x_1, x_2 \in \mathcal{S}_+^n$, $\lambda \in [0, 1]$ and $z \in \mathbb{R}^s$, the convexity of φ yields

$$f(\lambda x_1 + (1 - \lambda)x_2, z) \leq \lambda f(x_1, z) + (1 - \lambda)f(x_2, z)$$

and thus in particular $F(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda F(x_1) + (1 - \lambda)F(x_2)$ with respect to the $\mathbb{P}$-almost sure partial order, proving the asserted convexity of F.

Finally,

$$\|F(x_1) - F(x_2)\|_{L^1} = \int_{\mathbb{R}^s} |c \bullet (x_1 - x_2) + \varphi(z - T \bullet x_1) - \varphi(z - T \bullet x_2)| \, (\mathbb{P} \circ Z^{-1})(dz)$$

$$\leq \|c\| \cdot \|x_1 - x_2\| + L_\varphi \cdot \|T\| \cdot \|x_1 - x_2\|$$

holds for all $x_1, x_2 \in \mathcal{S}_+^n$. □

Definition 1 A mapping $\mathcal{R} : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ defined on some linear subspace $\mathcal{X}$ of $L^0(\Omega, \mathcal{F}, \mathbb{P})$ containing the constants is called a **convex risk measure** if the following conditions are fulfilled:

1. **(Convexity)** For any $Z_1, Z_2 \in \mathcal{X}$ and $\lambda \in [0, 1]$ we have

$$\mathcal{R}[\lambda Z_1 + (1 - \lambda)Z_2] \leq \lambda \mathcal{R}[Z_1] + (1 - \lambda)\mathcal{R}[Z_2].$$

2. **(Monotonicity)** $\mathcal{R}[Z_1] \leq \mathcal{R}[Z_2]$ for all $Z_1, Z_2 \in \mathcal{X}$ satisfying $Z_1 \leq Z_2$ with respect to the $\mathbb{P}$-almost sure partial order.

3. **(Translation equivariance)** $\mathcal{R}[Z_1 + z_2] = \mathcal{R}[Z_1] + z_2$ for all $Z_1 \in \mathcal{X}$ and $z_2 \in \mathbb{R}$.

A convex risk measure $\mathcal{R}$ is **coherent** if the following holds true:

4. **(Positive homogeneity)** $\mathcal{R}[z_2 Z_1] = z_2 \cdot \mathcal{R}[Z_1]$ for all $Z_1 \in \mathcal{X}$ and $z_2 \in [0, \infty)$.

Definition 2 A mapping $\mathcal{R} : L^0(\Omega, \mathcal{F}, \mathbb{P}) \supseteq \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ is called **law-invariant** if for all $Z_1, Z_2 \in L^0(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{P} \circ Z_1^{-1} = \mathbb{P} \circ Z_2^{-1}$ we have $\mathcal{R}[Z_1] = \mathcal{R}[Z_2]$.

We shall give some examples of risk-measures frequently used in stochastic programming as listed in [17], pp. 447-448, and [21]. Later we will give extensive formulations of discrete mean-risk SDPs based on these risk-measures:

(i) The expectation $\mathbb{E} : L^1(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ is a law-invariant coherent risk-measure.

(ii) The expected excess over threshold $\eta \in \mathbb{R}$ (as used in [18]) is the mapping $\mathbb{EE}_\eta : L^1(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ defined by

$$\mathbb{EE}_\eta[Y] = \int_\Omega \max\{Y(\omega) - \eta, 0\} \, \mathbb{P}(d\omega).$$

This is a non-decreasing, convex and law-invariant risk measure, but in general not translation-equivariant.

(iii) The conditional value-at-risk at level $\alpha \in (0, 1)$

$$\mathbb{CV@R}_\alpha : L^1(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}, \ \mathbb{CV@R}_\alpha[Y] = \min_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{1 - \alpha} \mathbb{EE}_\eta(Y) \right\} \qquad (6)$$

is law-invariant and coherent (cf. [14]).

(iv) The value-at-risk at level $\alpha \in (0, 1)$

$$\mathbb{V@R}_\alpha : L^0(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}, \ \mathbb{V@R}_\alpha[Y] = \inf\{\, t \mid \mathbb{P}(Z(\omega) \leq t) \geq \alpha \,\}$$

is nondecreasing, law-invariant, translation-equivariant and positively homogenous, but in general non-convex.

(v) The upper semi-deviation of order p is the mapping $\mathbb{Mad}_p^+ : L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ defined by

$$\mathbb{Mad}_p^+[Y] = \Big(\int \max\{0, Y(\omega) - \mathbb{E}_\mathbb{P}[Z]\}^p \, \mathbb{P}(\mathrm{d}\omega) \Big)^{\frac{1}{p}}.$$

For $\rho \in [0, 1]$ this gives rise to the law-invariant and coherent risk measure $\mathbb{E} + \rho \, \mathbb{Mad}_p$ (cf. [21], p. 276).

Proposition 1 *Assume A1 and A2, let $\mathcal{X}$ be a convex subset of $L^0(\Omega, \mathcal{F}, \mathbb{P})$ that contains $F(\mathcal{S}_+^n)$ and fix a convex and nondecreasing mapping $\mathcal{R} : \mathcal{X} \to \mathbb{R}$. Then $Q_\mathcal{R}$ is finite and convex on $\mathcal{S}_+^n$. In particular, problem (3) is convex.*

Proof Finiteness of $Q_\mathcal{R}$ follows directly from the finiteness of $\mathcal{R}$. Furthermore, for any $x_1, x_2 \in \mathcal{S}_+^n$ and $\lambda \in [0, 1]$ we have

$$\begin{aligned} Q_\mathcal{R}(\lambda x_1 + (1-\lambda)x_2) &= \mathcal{R}[F(\lambda x_1 + (1-\lambda)x_2)] \\ &\leq \mathcal{R}[\lambda F(x_1) + (1-\lambda)F(x_2)] \\ &\leq \lambda \mathcal{R}[F(x_1)] + (1-\lambda)\mathcal{R}[F(x_2)]. \end{aligned}$$

The first inequality above holds due to the monotonicity of $\mathcal{R}$ and the convexity of F (by Lemma 3), while the second one is justified by the convexity of $\mathcal{R}$. $\quad\square$

Proposition 2 *Assume A1, A2 and that the support of $\mathbb{P} \circ Z^{-1}$ is bounded. Furthermore, let $\mathcal{R} : L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R} \cup \{\infty\}$ be a coherent risk measure and assume that there is some $Y \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathcal{R}[Y] < \infty$. Then $Q_\mathcal{R}$ is finite and Lipschitz continuous with constant $\|c\| + L_\varphi \cdot \|T\|$ on $\mathcal{S}_+^n$.*

Proof $\mathcal{R}$ is finite and Lipschitz continuous with constant 1 with respect to the L^∞-norm on by $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ by [8, Lemma 4.3].

For any $x \in \mathcal{S}_+^n$, the mapping $f(x, \cdot)$ is continuous by Lemma 2, which implies

$$\sup_{z \in \mathrm{supp}(\mathbb{P} \circ Z^{-1})} |f(x, z)| < \infty.$$

Thus, $F(\mathcal{S}_+^n) \subseteq L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, which implies the asserted finiteness of $Q_\mathcal{R}$.

Furthermore, for any $x_1, x_2 \in \mathcal{S}_+^n$, we have

$$\begin{aligned} |Q_\mathcal{R}(x_1) - Q_\mathcal{R}(x_2)| &= |\mathcal{R}[F(x_1)] - \mathcal{R}[F(x_2)]| \\ &\leq \|F(x_1) - F(x_2)\|_{L^\infty} \\ &\leq \|F(x_1) - F(x_2)\|_{L^1} \\ &\leq (\|c\| + L_\varphi \cdot \|T\|) \cdot \|x_1 - x_2\| \end{aligned}$$

by Lemma 3. $\quad\square$

If the support of $\mathbb{P} \circ Z^{-1}$ is unbounded, $F(\mathcal{S}^n_+)$ may fail to be a subset of $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. While Lipschitz continuity with respect to any L^p-norm with $p < \infty$ does not hold for general coherent risk measures, the Conditional Value-at-Risk CVaR_α is known to be Lipschitz continuous with respect to the L^1-norm with constant $\frac{1}{1-\alpha}$ (cf. [15, Corollary 3.7]). Using the Kusuoka representation (cf. [12]), this allows to replace the boundedness of the support of $\mathbb{P} \circ Z^{-1}$ with a less restrictive assumption on the moments of $\mathbb{P} \circ Z^{-1}$ for special classes of risk measures.

Definition 3 Random variables Z_1 and Z_2 are called **comonotonic** if (Z_1, Z_2) is distributionally equivalent to $(F_{Z_1}^{-1}(U), F_{Z_2}^{-1}(U))$ where U is uniformly distributed on $[0, 1]$.

A coherent risk measure $\mathcal{R} : \mathcal{X} \to \mathbb{R}$ is said to be **comonotonic** if for any two comonotonic random variables $Z_1, Z_2 \in \mathcal{X}$ we have $\mathcal{R}(Z_1 + Z_2) = \mathcal{R}(Z_1) + \mathcal{R}(Z_2)$.

For a discussion of comonotonicity we refer to [6] and [7]. A proof of the following result is given in [20, Theorem 2]:

Theorem 1 *A law-invariant coherent risk measure $\mathcal{R} : L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ with $p \in [1, \infty)$ is comonotonic if and only if there exists probability measure ν on $[0, 1)$ such that*

$$\mathcal{R}(Y) = \int_0^1 \mathbb{CV@R}_\alpha(Y) \ \nu(d\alpha) \tag{7}$$

holds for all $Y \in L^p(\Omega, \mathcal{F}, \mathbb{P})$. Furthermore, the measure ν in representation (7) is defined uniquely.

Example 2 Using δ_{α_0} to denote the Dirac measure at $\alpha_0 \in [0, 1)$

$$\mathbb{CV@R}_{\alpha_0}(Y) = \int_0^1 \mathbb{CV@R}_\alpha(Y) \ \delta_{\alpha_0}(d\alpha)$$

and, in particular,

$$\mathbb{E}[Y] = \int_0^1 \mathbb{CV@R}_\alpha(Y) \ \delta_0(d\alpha)$$

hold for all $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$.

Proposition 3 *Let $\mathcal{R} : L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ with $p \in [1, \infty)$ be a law-invariant, comonotonic coherent risk measure. Assume A1, A2, $\mathbb{P} \circ Z^{-1} \in \mathcal{M}^p_s$ and*

$$L_\nu := \int_0^1 \frac{1}{1-\alpha} \ \nu(d\alpha) < \infty,$$

where ν denotes the uniquely defined probability measure form representation (7). Then $Q_\mathcal{R}$ is Lipschitz continuous with constant $L_\nu \cdot (\|c\| + L_\varphi \cdot \|T\|)$ on $\mathcal{S}^n_+$.

Proof For any $x_1, x_2 \in \mathcal{S}^n_+$, we have

$$\begin{aligned}
|Q_\mathcal{R}(x_1) - Q_\mathcal{R}(x_2)| &\leq \int_0^1 |\mathrm{CVaR}_\alpha(F((x_1)) - \mathrm{CVaR}_\alpha(F((x_2))| \ \nu(d\alpha) \\
&\leq \int_0^1 \frac{1}{1-\alpha} \cdot \|F((x_1)) - F((x_2))\|_{L^1} \ \nu(d\alpha) \\
&\leq \int_0^1 \frac{1}{1-\alpha} \cdot (\|c\| + L_\varphi \cdot \|T\|) \cdot \|x_1 - x_2\| \ \nu(d\alpha) \\
&= L_\nu \cdot (\|c\| + L_\varphi \cdot \|T\|) \cdot \|x_1 - x_2\|.
\end{aligned}$$

The second inequality above holds due to [15, Corollary 3.7], while the third one is justified by Lemma 3. □

We shall now study the dependence of $Q_{\mathcal{R}}$ on the underlying probability measure $\mathbb{P} \circ Z^1$. This is motivated by the fact that in applications the true probability distribution of the random parameter may be unknown. In such situations, one may work with an approximation if the optimal value function and the optimal solution set mapping of (3) are at least semicontinuous with respect to changes of the underlying distribution.

Let $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ be an atomless probability space, i.e. assume that for any $A \in \mathcal{F}_0$ with $\mathbb{P}_0(A) > 0$ there exists some $B \subsetneq A$ with $B \in \mathcal{F}_0$ and $\mathbb{P}_0(B) > 0$, and fix any $p \geq 1$. Then for any $\nu \in \mathcal{M}_p^1$ there exists some $Z_\nu \in L^p(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ such that $\mathbb{P}_0 \circ Z_\nu^{-1}$. Thus, given any law-invariant mapping $\mathcal{R}_0 : L^p(\Omega_0, \mathcal{F}_0, \mathbb{P}_0) \to \mathbb{R}$, the function

$$\Theta_{\mathcal{R}_0} : \mathcal{M}_p^1 \to \mathbb{R}, \quad \Theta_{\mathcal{R}_0}[\nu] := \mathcal{R}_0[Z_\nu]$$

is well-defined. Furthermore, we can construct a mapping $\mathcal{R}_{\mathcal{R}_0} : L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ by setting $\mathcal{R}_{\mathcal{R}_0}[Z_1] := \Theta_{\mathcal{R}_0}[\mathbb{P} \circ Z_1^{-1}]$. To ease the notation, we shall assume that $(\Omega, \mathcal{F}, \mathbb{P})$ itself is atomless. Given any law-invariant mapping $\mathcal{R} : L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$, we shall consider the function

$$\mathcal{Q}_{\mathcal{R}} : \mathcal{S}_+^n \times \mathcal{M}_s^p \to \mathbb{R}, \quad \mathcal{Q}_{\mathcal{R}}(x, \mu) := \Theta_{\mathcal{R}}[\mu \circ f(x, \cdot)^{-1}].$$

For the following analysis, we equip the space $\mathcal{P}(\mathbb{R}^s)$ with the topology of weak convergence, where a sequence $\{\mu_k\}_{k \in \mathbb{N}} \subseteq \mathcal{P}(\mathbb{R}^s)$ converges to some $\mu \in \mathcal{P}(\mathbb{R}^s)$, written $\mu_k \overset{w}{\to} \mu$ if and only if

$$\int_{\mathbb{R}^s} h(t) \; \mu_k(dt) \to \int_{\mathbb{R}^s} h(t) \; \mu(dt)$$

holds for any bounded and continuous function $h : \mathbb{R}^s \to \mathbb{R}$. It is well known that even for linear recourse one cannot expect weak continuity of $\mathcal{Q}_{\mathcal{R}}$ on the entire space $\mathcal{S}_+^n \times \mathcal{M}_s^p$. Along the lines of [5], we shall thus restrict the analysis to appropriate subspaces.

Definition 4 A set $\mathcal{M} \subseteq \mathcal{M}_s^p$ is called **locally uniformly $\|\cdot\|^p$-integrating** if for any $\mu \in \mathcal{M}$ and any $\epsilon > 0$ there exists some open neighborhood $\mathcal{N}$ of μ with respect to the topology of weak convergence such that

$$\lim_{a \to \infty} \sup_{\nu \in \mathcal{N} \cap \mathcal{M}} \int_{\mathbb{R}^s} \mathbb{1}_{(a, \infty)}(\|t\|^p) \cdot \|t\|^p \; \nu(dt) \leq \epsilon.$$

Example 3 (a) For any $K, \epsilon > 0$ and $p \geq 1$, the set

$$U(\epsilon, K) := \{\nu \in \mathcal{M}_s^p : \int_{\mathbb{R}^s} \|t\|^{1+\epsilon} \; \nu(dt) \leq K\}$$

of measures having uniformly bounded moments of order $1 + \epsilon$ is locally uniformly $\|\cdot\|^p$-integrating (cf. [4, Lemma 2.69]).

(b) For any $p \geq 1$ and compact set $\Xi \subset \mathbb{R}^s$, the set

$$\{\nu \in \mathcal{M}_s^p : \int_{\Xi} 1 \; \nu(dt) = 1\}$$

of measures with support in Ξ is locally uniformly $\|\cdot\|^p$-integrating by [11, Lemma 5.1].

(c) Any singleton $\{\mu\} \subseteq \mathcal{M}_s^p$ is locally uniformly $\|\cdot\|^p$-integrating for any $p \geq 1$ by [11, Lemma 5.2].

Theorem 2 *Let $\mathcal{R} : L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ with $p \geq 1$ be law-invariant, convex and nondecreasing. Assume A1 and A2 and let $\mathcal{M} \subseteq \mathcal{M}_s^p$ be locally uniformly $\|\cdot\|^p$-integrating. Then the following statements hold true:*

1. *The restriction of $\mathcal{Q}_\mathcal{R}$ to the set $\mathcal{S}_+^n \times \mathcal{M}$ is continuous with respect to the product topology of the the standard topology on $\mathcal{S}_+^n$ and the relative topology of weak convergence on $\mathcal{M}$.*

2. *The optimal value function*

$$\phi : \mathcal{M} \to \overline{\mathbb{R}}, \;\; \phi(\mu) := \min_x \{\mathcal{Q}_\mathcal{R}(x, \mu) \mid x \in X\}$$

   is weakly upper semicontinuous.

Additionally assume that X is compact. Then

3. *ϕ is weakly continuous.*
4. *The optimal solution set mapping*

$$\Phi : \mathcal{M} \rightrightarrows \mathcal{S}_+^n, \;\; \Phi(\mu) := \mathrm{Argmin}_x \{\mathcal{Q}_\mathcal{R}(x, \mu) \mid x \in X\}$$

   is weakly upper semicontinuous in the sense of Berge, i.e. for any $\mu_0 \in \mathcal{M}$ and any open set $\mathcal{O} \subseteq \mathcal{S}_+^n$ with $\Phi(\mu_0) \subseteq \mathcal{O}$ there exists a weakly open neighborhood $\mathcal{N}$ of μ_0 such that $\Phi(\mu) \subseteq \mathcal{O}$ for all $\mu \in \mathcal{N} \cap \mathcal{M}$. Furthermore, $\Phi(\mu)$ is nonempty and compact for any $\mu \in \mathcal{M}$.

Proof Invoking Lemma 2, the result follows from [5, Corollary 2]. $\qquad\square$

Corollary 1 *Let $\mathcal{R} : L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ with $p \geq 1$ be law-invariant, convex and nondecreasing and assume A1 and A2. Then $Q_\mathcal{R}$ is continuous.*

Proof By part (c) of Example 3 we may apply the first part of Theorem 2 to $\mathcal{M} = \{\mathbb{P} \circ Z^{-1}\}$. The asserted continuity follows from $Q_\mathcal{R}(x) = \mathcal{Q}_\mathcal{R}(x, \mathbb{P} \circ Z^{-1})$ for any $x \in \mathcal{S}_+^n$. $\qquad\square$

We shall now turn our attention to questions of differentiability, but confine the analysis to the risk neutral model.

Lemma 4 *Assume A1, A2 and $\mathbb{P} \circ Z^{-1} \in \mathcal{M}_s^1$, then the functional $Q_\mathbb{E} : \mathcal{S}_+^n \to \mathbb{R}$, $Q_\mathbb{E}(x) := \mathbb{E}[F(x)]$ is directionally differentiable and*

$$Q_\mathbb{E}'(x; v) := \int_{\mathbb{R}^s} \varphi'(z - T \bullet x; v) \, (\mathbb{P} \circ Z^{-1})(dz)$$

holds for all $x, v \in \mathcal{S}_+^n$.

Proof $Q_\mathbb{E}$ is finite valued by Lemma 3, convex by Proposition 1 and thus directionally differentiable (cf. [16, Theorem 25.4]). Furthermore, $\varphi'(\cdot - Tx; v)$ is a pointwise limit of measurable functions and thus measurable for any $x, v \in \mathcal{S}_+^n$. The asserted representation of the directional derivative is justified by Lemma 2 and [3, Proposition 2.1]. $\qquad\square$

Sufficient conditions for differentiability $Q_\mathbb{E}$ can be obtained using the same arguments as for linear recourse (cf. [21]).

Lemma 5 *Assume A1, A2 and $\mathbb{P} \circ Z^{-1} \in \mathcal{M}_s^1$ and let $x_0 \in \mathcal{S}_+^n$ be such that*

$$\operatorname{Argmax}\{u^\top (z - T \bullet x_0) \mid u \in M_D\}$$

is a singleton for $(\mathbb{P} \circ Z^{-1})$-almost all $z \in \mathbb{R}^s$. Then $Q_\mathbb{E}$ is differentiable at x_0.

Proof For $(\mathbb{P} \circ Z^{-1})$-almost all $z \in \mathbb{R}^s$, $h_z : \mathcal{S}_+^n \to \mathbb{R}$, $h_z(x) = c \bullet x + \varphi(z - T \bullet x)$ is differentiable with measurable derivative

$$h_z'(x) = c + -T^\top \cdot \operatorname{Argmax}\{u^\top (z - T \bullet x_0) \mid u \in M_D\}.$$

Consider the functions $g_z : \mathcal{S}_+^n \to \mathbb{R}$ defined by

$$g_z(x) := \frac{h_z(x) - h_z(x_0) - h_z'(x_0)^\top (x - x_0)}{\|x - x_0\|},$$

then $\lim_{x \to x_0} g_z(x) = 0$ holds for $(\mathbb{P} \circ Z^{-1})$-almost all $z \in \mathbb{R}^s$. Furthermore, Lemma 2 implies $\|g_z(x)\| \leq 2(L_\varphi \|T\| + \|c\|)$ for all $x \in \mathcal{S}_+^n$ and $z \in \mathbb{R}^s$. Hence, by Lebesgue's dominated convergence theorem, we have

$$\lim_{x \to x_0} \frac{Q_\mathbb{E}(x) - Q_\mathbb{E}(x_0) - \int_{\mathbb{R}^s} h_z'(x_0)^\top (x - x_0) \, (\mathbb{P} \circ Z^{-1})(dz)}{\|x - x_0\|}$$

$$= \lim_{x \to x_0} \int_{\mathbb{R}^s} g_z(x) \, (\mathbb{P} \circ Z^{-1})(dz) = \int_{\mathbb{R}^s} \lim_{x \to x_0} g_z(x) \, (\mathbb{P} \circ Z^{-1})(dz) = 0.$$

Consequently, $Q_\mathbb{E}$ is differentiable at x_0 and $Q_\mathbb{E}'(x_0) = \int_{\mathbb{R}^s} h_z'(x_0) \, (\mathbb{P} \circ Z^{-1})(dz)$. $\square$

Corollary 2 *Assume A1, A2 and that $\mathbb{P} \circ Z^{-1} \in \mathcal{M}_s^1$ is absolutely continuous with respect to the Lebesgue measure. Then $Q_\mathbb{E}$ is continuously differentiable on $\mathcal{S}_+^n$.*

Proof Let $N_\varphi \subset \mathbb{R}^s$ denote the set of points of nondifferentiability of φ. By [16, Theorem 25.5],
$$N_x := \{z \in \mathbb{R}^s \mid z - T \bullet x \in N_\varphi\}$$
is a null set with respect to the Lebesgue measure for any $x \in \mathcal{S}_+^n$, which implies $(\mathbb{P} \circ Z^{-1})[N_x] = 0$. Consequently, $Q_\mathbb{E}$ is differentiable on $\mathcal{S}_+^n$. Continuity of the derivative follows from [16, Theorem 25.5] and the convexity of $Q_\mathbb{E}$. $\square$

Remark 3 Assuming A1, A2 and $\mathbb{P} \circ Z^{-1} \in \mathcal{M}_s^1$, the subdifferential of $Q_\mathbb{E}$ admits the representation

$$\partial Q_\mathbb{E}(x) = c + \int_\mathbb{R}^s \partial_x \varphi(z - T \bullet x) \, (\mathbb{P} \circ Z^{-1})(dz)$$

$$= \left\{ c + \int_{\mathbb{R}^s} \rho(z) \, (\mathbb{P} \circ Z^{-1})(dz) \mid \rho : \mathbb{R}^s \to \mathcal{S}_+^n \text{ measurable}, \ \rho(z) \in \partial_x \varphi(z - T \bullet x) \text{ a.s.} \right\}.$$

Furhter details are given in [3].

Corollary 3 *Assume A2 and that the underlying random variable Z follows a finite discrete distribution with realizations $z_1, \ldots, z_S \in \mathbb{R}^s$ and respective probabilities $\pi_1, \ldots, \pi_S > 0$. Furthermore, assume that $\{y \in \mathcal{S}_+^m \mid W \bullet y = z_i - T \bullet x\}$ is nonempty for any $i \in \{1, \ldots, S\}$ and $x \in \mathcal{S}_+^n$. Then*

$$\partial Q_{\mathbb{E}}(x) = c + \sum_{i=1}^{s} \pi_i \cdot \partial_x \varphi(z_i - T \bullet x)$$

$$= c + \sum_{i=1}^{s} -\pi_i \cdot T^\top \cdot \mathrm{Argmax}\{u^\top(z_i - T \bullet x) \mid u \in M_D\}$$

holds for any $x \in \mathcal{S}_+^n$.

Proof The result follows directly from [16, Theorem 23.8]. $\qquad\square$

4 Extensive Formulations for Finite Discrete Distributions

Throughout this section, we shall assume A1, A2 and that the underlying random variable Z follows a finite discrete distribution with realizations $z_1, \ldots, z_S \in \mathbb{R}^s$ and respective probabilities $\pi_1, \ldots, \pi_S > 0$. Furthermore, we denote the index set $\{1, \ldots, S\}$ by $\mathcal{I}_S$.

It is well known that in the risk neutral setting, the stochastic SDP admits a reformulation as a block-structured SDP (cf. [1], [13]):

Proposition 4 *The risk neutral stochastic SDP*

$$\min\{Q_{\mathbb{E}}(x) \mid x \in X\} \tag{8}$$

is equivalent to the SDP

$$\min_{x, y_1, \ldots, y_S} \left\{ c \bullet x + \sum_{i=1}^{S} \pi_i q \bullet y_i \mid T \bullet x + W \bullet y_i = z_i \ \forall i \in \mathcal{I}_S, \right. \tag{9}$$

$$\left. x \in X, \ y_i \in \mathcal{S}_+^m \ \forall i \in \mathcal{I}_S \right\},$$

in the sense that the infimal values of the problems coincide. Furthermore, x is an optimal solution for (8) if and only if there exist v and $y_1, \ldots, y_S$ such that $(x, v, y_1, \ldots, y_S)$ is an optimal solution for (9).

Proof By definition of φ,

$$Q_{\mathbb{E}}(x) = c \bullet x + \sum_{i=1}^{S} \pi_i \varphi(z_i - T \bullet x) \leq c \bullet x + \sum_{i=1}^{S} \pi_i q \bullet y_i \tag{10}$$

holds for any $x \in X$, $y_1, \ldots, y_S \in \mathcal{S}_+^m$ satisfying $T \bullet x + W \bullet y_i = z_i$ for all $i \in \mathcal{I}_S$. Thus, the infimal value of (8) is less or equal to the infimal value of (9). Furhtermore, (10) is satisfied as equality if and only if

$$y_i \in \mathrm{Argmin}\{q \bullet y \mid T \bullet x + W \bullet y = z_i, \ y \in \mathcal{S}_+^m\}$$

holds for all $i \in \mathcal{I}_S$. The optimal solution set above is nonempty by strong duality, which holds due to A1 and A2. $\qquad\square$

We continue with extensive formulations of the SDP (3) for mean-risk models based on the risk measures immediately following Definition 2. In this context, ρ shall always be a nonnegative, predefined parameter indicating risk-aversion in the optimization.

Proposition 5

$$\min \left\{ Q_{\mathbb{E}+\rho\,\mathbb{E}\mathbb{E}_\eta}(x) \mid x \in X \right\}, \tag{11}$$

with $\eta \in \mathbb{R}$ as a given parameter, can be equivalently restated as

$$\min_{\substack{x,v_1,\ldots,v_S, \\ y_1,\ldots,y_S}} \left\{ c \bullet x + \sum_{i=1}^{S} \pi_i q \bullet y_i + \rho \sum_{i=1}^{S} \pi_i v_i \mid T \bullet x + W \bullet y_i = z_i \; \forall i \in \mathcal{I}_S, \right. \tag{12}$$
$$v \geq 0, \; v_i \geq c \bullet x + q \bullet y_i - \eta \; \forall i \in \mathcal{I}_S,$$
$$\left. x \in X, \; y_i \in \mathcal{S}_+^m \; \forall i \in \mathcal{I}_S \right\}.$$

Proof As the objective function of (12) is increasing with respect to v, any optimal solution $(x, v_1, \ldots, v_S, y_1, \ldots, y_S)$ satisfies $v_i = \max\{c \bullet x + q \bullet y_i - \eta, 0\}$ for all $i \in \mathcal{I}_S$. The asserted equivalence of (11) and (12) then follows as in the proof of Proposition 4. □

Proposition 6

$$\min \left\{ Q_{\mathbb{E}+\rho\,\mathbb{C}\mathbb{V}@R_\alpha}(x) \mid x \in X \right\}$$

can be equivalently restated as

$$\min_{\substack{x,v_1,\ldots,v_S, \\ y_1,\ldots,y_S,\eta}} \left\{ c \bullet x + \sum_{i=1}^{S} \pi_i q \bullet y_i + \rho\,\eta + \frac{\rho}{1-\alpha} \sum_{i=1}^{S} \pi_i v_i \mid \right. \tag{13}$$
$$T \bullet x + W \bullet y_i = z_i \; \forall i \in \mathcal{I}_S,$$
$$v \geq 0, \; v_i \geq c \bullet x + q \bullet y_i - \eta \; \forall i \in \mathcal{I}_S,$$
$$\left. \eta \in \mathbb{R}, \; x \in X, \; y_i \in \mathcal{S}_+^m \; \forall i \in \mathcal{I}_S \right\}.$$

Proof This follows directly from the variational representation of $\mathbb{C}\mathbb{V}@R$ in (6). The expected-excess can be pushed into the restrictions by the same trick as in Proposition 5. □

As in in the risk-neutral case, problems (12) and (13) exhibit a block structure, i.e. there is no coupling constraint involving variables associated with different scenarios. This allows for a direct adaptation of the decomposition algorithms established for the expectation based model.

Proposition 7 *Consider the problem*

$$\min \left\{ Q_{\mathbb{E}+\rho\,\mathbb{V}@R_\alpha}(x) \mid x \in X \right\}$$

with compact set X. This problem can be equivalently restated as the following SDP with binary variables

$$
\min_{\substack{x,v_1,\ldots,v_S,\\ y_1,\ldots,y_S,\\ \delta_1,\ldots,\delta_S,\eta}} \left\{ (1+\rho)\, c \bullet x + \sum_{i=1}^{S} \pi_i q \bullet y_i + \rho\,\eta \mid \right. \tag{14}
$$

$$
T \bullet x + W \bullet y_i = z_i \ \forall i \in \mathcal{I}_S,
$$

$$
\sum_{i=1}^{S} \delta_i\,\pi_i \geq \alpha,
$$

$$
\eta - q \bullet y_i \geq (1-\delta_i)\,M \ \forall i \in \mathcal{I}_S
$$

$$
\left. \eta \in \mathbb{R},\ x \in X,\ \delta_i \in \{0,1\},\ y_i \in \mathcal{S}_+^m \ \forall i \in \mathcal{I}_S \right\}
$$

if $M \in \mathbb{R}$ is chosen sufficiently big.

Proof As in the preceding propositions introduce a dummy variable η to push $\mathbb{V}@R[\varphi(z - T \bullet x)]$ into the restrictions as $\eta \geq \mathbb{V}@R[\varphi(z - T \bullet x)]$ and minimize over η. Note that $\eta \geq \mathbb{V}@R[\varphi(z - T \bullet x)]$ is equivalent to

$$
\mu(\varphi(z - T \bullet x) \leq \eta) \geq \alpha. \tag{15}
$$

As for given $x \in X$ feasible points to the second stage problem corresponding to realization z_i are denoted as y_i, (15) can be rewritten as

$$
\sum_{i \in \mathcal{I}_S \,:\, q \bullet y_i \leq \eta} \pi_i \geq \alpha.
$$

This conditional summation can in turn be cast into inequalities with binary variables $\delta_i,\ i \in \mathcal{I}_S$,

$$
\eta - q \bullet y_i \geq (1-\delta_i)\,M,\ i \in \mathcal{I}_S
$$

$$
\sum_{i \in \mathcal{I}_S} \delta_i\,\pi_i \geq \alpha
$$

if M is chosen such that $\eta - q \bullet y_i < M$ for all feasible y_i and all η close to $\mathbb{V}@R[\varphi(z_i - T \bullet x)]$. Since $-q \bullet y_i \leq -\varphi(z_i - T \bullet x)$ the existence of M follows from compactness of X, as $\max_{x \in X} \varphi(z_i - T \bullet x) < \infty$ for all $i \in \mathcal{I}_S$.   □

Unlike the previous models, (14) does not decompose scenariowise due to the coupling constraint $\sum_{i=1}^{S} \delta_i\,\pi_i \geq \alpha$, which involves variables from all scenarios. Furthermore, it has an additional binary variable for each scenario. Problems of a similar structure have been considered in the context of minimizing a weighted sum of the expectation and the probability of exceeding a fixed threshold in [19], where Lagrangian relaxation of the coupling constraint enables an approach based on Bender's decomposition. This direction seems also very promising for the algorithmic treatment of (14).

Proposition 8

$$\min \left\{ Q_{\mathbb{E}+\rho\, \mathbb{M}ad_p^+}(x) \mid x \in X \right\},$$

can be equivalently restated as

$$\min_{\substack{x,v_1,\ldots,v_S, \\ y_1,\ldots,y_S}} \left\{ c \bullet x + \sum_{i=1}^{S} \pi_i q \bullet y_i + \rho \left(\sum_{i=1}^{S} \pi_i v_i^p \right)^{\frac{1}{p}} \mid T \bullet x + W \bullet y_i = z_i \; \forall i \in \mathcal{I}_S, \right.$$

$$v \geq 0, \; v_i \geq c \bullet x + q \bullet y_i - \sum_{j=1}^{S} \pi_j \, q \bullet y_j \; \forall i \in \mathcal{I}_S,$$

$$\left. x \in X, \; y_i \in \mathcal{S}_+^m \; \forall i \in \mathcal{I}_S \right\}.$$

Proof Analogous to Proposition 5. □

Unlike (14), the equivalent SDP in Proposition 8 contains an individual coupling constraint for each scenario. While Lagrangian relaxation still is possible, it remains to be examined whether this approach is sensible form a computational point of view.

References

1. K. A. Ariyawansa, Y. Zhu, *Stochastic semidefinite programming: a new paradigm for stochastic optimization*, 4OR, 4(3), pp. 239-253 (2006)
2. K. A. Ariyawansa, Y. Zhu, *A class of polynomial volumetric barrier decomposition algorithms for stochastic semidefinite programming*, Mathematics of Computation, 80, no. 275, pp.1639-1661 (2011)
3. D. P. Bertsekas, *Stochastic optimization problems with nondifferentiable cost functionals*, Journal of Optimization Theory and Applications, 12, pp. 218-231 (1973)
4. M. Claus, *Advancing stability analysis of mean-risk stochastic programs : bilevel and two-stage models*, PhD thesis, University of Duisburg-Essen (2016)
5. M. Claus, V. Krätschmer and R. Schultz, *Weak continuity of risk functionals with applications to stochastic programming*, SIAM Journal on Optimization, 27(1), pp. 91-108 (2017)
6. J. Dhaene, M. Denuit, M. J. Goovaerts, R. Kaas, D. Vyncke, *The concept of comonotonicity in actuarial science and finance: theory*, Insurance: Math. Econom., 31, pp. 3-33 (2002)
7. J. Dhaene, S. Vanduffel, M. J. Goovaerts, R. Kaas, Q. Tang, D. Vyncke, *Risk Measures and Comonotonicity: A Review*, Stochastic Models, 22, pp. 573-606 (2006)
8. H. Föllmer, A. Schied, *Stochastic Finance: An Introduction in Discrete Time*, 2nd ed., de Gruyter Stud. Math. 27, de Gruyter, Berlin (2004)
9. B. Gaujal, P. Mertikopoulos, *A stochastic approximation algorithms for stochastic semidefinite programming*, Probability in the Engineering and Informational Sciences, 30, pp. 431454 (2016)
10. S. Jin, K. A. Ariyawansa, Y. Zhu, *Homogeneous Self-dual Algorithms for Stochastic Semidefinite Programming*, Journal of Optimization Theory and Applications, 155(3), pp.1073-1083 (2012)
11. V. Krätschmer, A. Schied, H. Zähle, *Domains of weak continuity of statistical functionals with a view on robust statistics*, Journal of Multivariate Analysis, 158, pp. 1-19 (2017)
12. S. Kusuoka, *On law-invariant coherent risk measures*, Advances in Mathematical Economics (eds. S. Kusuoka, T. Maruyama), Vol. 3, Springer, Tokyo, pp. 83-95 (2001)
13. S. Mehrotra, M. G. Özevin, *Decomposition-based interior point methods for two-stage stochastic semidefinite programming*, SIAM J. Optim., 18(1), pp. 206-222 (2007)
14. G. C. Pflug, *Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk*, in Probabilistic Constrained Opimization - Methodology and Applications (ed. S. P. Uryasev), Kluwer Academic Publishers, Dordrecht, pp. 272-281 (2000).

15. A. Pichler, *A quantitative comparison of risk measures*, Ann. Oper. Res., 254, pp. 251-275 (2017)
16. R. T. Rockafellar, *Convex Analysis*, Princeton University Press (1970)
17. A. Ruszczynski, A. Shapiro, *Stochastic Programming*, 1st ed., Elsevier Science, Amsterdam (2003)
18. R. Schultz, S. Tiedemann, *Conditional Value-at-Risk in Stochastic Programs with Mixed-Integer Recourse*, Mathematical Programming, Series B, 105 (2-3), pp. 365-386 (2006)
19. R. Schultz, T. Wollenberg, *Unit commitment under uncertainty in AC transmission systems via risk averse semidefinite stochastic Programs*, RAIRO Oper. Res., 51(2), pp. 391-416 (2017)
20. A. Shapiro, *On Kusuoka Representation of Law Invariant Risk Measures*, Math. of Oper. Res., 38(1), pp. 142-152 (2013)
21. A. Shapiro, D. Dentcheva, A. Ruszczynski, *Lectures on Stochastic Programming*, MPS-SIAM, Philadelphia (2009)
22. P. Sun, R. M. Freund, *Computation of minimum-cost covering ellipsoids*, Oper. Res., 52(5), pp. 690-706 (2004).
23. P. M. Vaidya, *A new algorithm for minimizing convex functions over a convex set*, Math. Program., Ser. A, 73, pp. 291-341 (1996).
24. L. Vandenberghe, S. Boyd, *Semidefinite programming*, SIAM Rev., 38, pp. 49-95 (1996).
25. G. Zhao, *A log-barrier method with Benders decomposition for solving two-stage stochastic linear programs*, Math. Program., Ser. A, 90, pp. 507-536 (2001).
26. Y. Zhu, *Semidefinite Programming under Uncertainy*, PhD thesis, Washington State University (2006).
27. Y. Zhu, K. A. Ariyawansa, *A preliminary set of applications leading to stochastic semidefinite programs and chance-constrained semidefinite programs*, Applied Mathematical Modelling, 35, pp. 2425-2442 (2011).