SIAM J. OPTIM.
Vol. 16, No. 1, pp. 69–95

POLYHEDRAL RISK MEASURES IN STOCHASTIC PROGRAMMING*

ANDREAS EICHHORN† AND WERNER RÖMISCH†

Abstract. We consider stochastic programs with risk measures in the objective and study stability properties as well as decomposition structures. Thereby we place emphasis on dynamic models, i.e., multistage stochastic programs with multiperiod risk measures. In this context, we define the class of polyhedral risk measures such that stochastic programs with risk measures taken from this class have favorable properties. Polyhedral risk measures are defined as optimal values of certain linear stochastic programs where the arguments of the risk measure appear on the right-hand side of the dynamic constraints. Dual representations for polyhedral risk measures are derived and used to deduce criteria for convexity and coherence. As examples of polyhedral risk measures we propose multiperiod extensions of the Conditional-Value-at-Risk.

Key words. stochastic programming, convex risk measure, coherent, polyhedral, mean-risk, quantitative stability, probability metrics, dual decomposition

AMS subject classifications. 90C15, 91B30

DOI. 10.1137/040605217

1. Introduction. Stochastic programs are essentially known to minimize, maximize, or bound expected values. From a theoretical point of view they easily offer the possibility to minimize or bound risk functionals since they rest upon stochastic models. This idea goes back to [14]. However, in practice it may happen that incorporating risk measures in stochastic programs makes them much harder to solve, especially if integer decisions are included. In addition, other favorable properties like stability with respect to approximations or duality results may get lost. In this paper considerations are made about the question as to how risk measures should be designed so that stochastic programs incorporating them show similar properties as stochastic programs based on expected values only. As a result, the class of polyhedral risk measures is introduced.

Of course, when analyzing risk measures with respect to their practicability for stochastic programs, one has to determine first of all what is understood by the expression *risk measure* and what properties are required from the viewpoint of economic considerations. Here, a (one-period) risk measure ρ will be understood as a functional from some set of real random variables to the real numbers. Random variables will be denoted by the letter z, they will represent uncertain (usually monetary) values for which larger outcomes are preferred to lower ones. The value $\rho(z)$ gives information about the riskiness of z, i.e., a high value $\rho(z)$ indicates a high danger of reaching low values.

Risk measures are broadly discussed in financial mathematics. For one-period risk measures, i.e., for risk measures that depend on one random variable only, there is a relatively high degree of agreement among the community about the desirable

*Received by the editors March 15, 2004; accepted for publication (in revised form) September 13, 2004; published electronically August 17, 2005. This research was supported by the DFG Research Center MATHEON *Mathematics for Key Technologies* at Berlin and by a grant of EDF Electricité de France.

http://www.siam.org/journals/siopt/16-1/60521.html

†Humboldt-University Berlin, Institute of Mathematics, 10099 Berlin, Germany (eichhorn@math.hu-berlin.de, romisch@math.hu-berlin.de, www.math.hu-berlin.de/~romisch).

properties. Possibly the most important work in this context is the axiomatic characterization of coherent risk measures [1], where the risk $\rho(z)$ is understood as the minimal amount of additional (risk-free) capital that is required to make the position z acceptable. Several generalizations of this paper followed, e.g., [6, 13, 10, 28]; see also Chapter 4 in the monograph [11]. Further desirable properties, namely, the consistency of risk measures with stochastic dominance rules, were suggested in [15, 17, 18, 19]. In addition, there are papers dealing with specific risk measures, e.g., [27, 20, 38]; see also the volumes [7, 41]. Recently, a theory for convex optimization of convex risk measures has been developed in [35].

Currently, generalizations of one-period risk measures to different dynamic settings are discussed in the literature. Such generalizations become necessary when information is revealed gradually with the passing of time and a sequence of random variables $z_1, \ldots, z_T$ is to be assessed with respect to its riskiness. In the literature, the settings as well as the postulated properties for risk functionals differ more than in the one-period case. Generally speaking, there are two classes of settings depending on whether liquidity risk over a time period is considered or intermediate monitoring by supervisors is to be anticipated. In the latter case an entire risk measure process $\rho_1, \ldots, \rho_T$ is defined; see [25, 42] and also [3, 2]. The more important case from the viewpoint of optimization is the case where one has one real number $\rho(z_1, \ldots, z_T)$ that represents the risk of the entire process (multiperiod risk). Such concepts are presented in [22, 36, 21] and again in [3, 2]. As in the one-period case, the number $\rho(z_1, \ldots, z_T)$ can be understood as minimal capital requirement for the overall time period so that the strategy corresponding to $z_1, \ldots, z_T$ is acceptable.

In the present paper, we consider (mixed-integer) multistage stochastic programs of the form

$$\min \left\{ \mathbb{E}\left[\sum_{t=1}^{T} \langle b_t(\xi_t), x_t\rangle\right] \;\middle|\; \begin{array}{l} x_t \text{ is } \mathcal{F}_t\text{-measurable}, \\ \sum_{\tau=0}^{t-1} A_{t,\tau}(\xi_t) x_{t-\tau} = h_t(\xi_t) \text{ a.s.}, \\ x_t \in X_t \text{ a.s. } (t = 1, \ldots, T) \end{array} \right\} \tag{1.1}$$

as starting point, where $(\xi_t)_{t=1}^T$ is a stochastic process and $\mathcal{F}_t = \sigma(\xi_1, \ldots, \xi_t)$, the sets X_t are closed and have polyhedral convex hulls, $b_t(\cdot)$ are cost coefficients, $h_t(\cdot)$ are right-hand sides, and $A_{t,\tau}(\cdot)$, $\tau = 0, \ldots, t-1$, are matrices having appropriate dimensions and possibly depending on ξ_t for $t = 1, \ldots, T$.

Much is known for expectation-based stochastic programs, e.g., on optimality and duality, decomposition methods, and statistical approximations and stability (cf. [34]). Most of these results are essentially based on the fact that $\mathbb{E}$ is a linear operator. As will be seen below in section 2, risk measures are usually by no means linear. Hence, if we change from expectation to a risk measure in (1.1), many known results will no longer be valid. Nevertheless, there are results about incorporating certain risk functionals into (stochastic) optimization problems, e.g., [38, 35, 37]. In particular, the Conditional-Value-at-Risk turns out to behave very opportunely in stochastic programs because it allows a reformulation of the risk aversive problem as an expectation-based problem with additional variables (cf. [27, 20, 40]).

However, from an economic point of view not every risk measure is suitable for any application. In particular, for multistage stochastic programs it may become necessary to incorporate multiperiod risk measures, i.e., to minimize $\rho(z_1, \ldots, z_T)$ with $z_t = -\sum_{\tau=1}^{t} \langle b_\tau(\xi_\tau), x_\tau\rangle$ denoting the intermediate values. Hence, it would be convenient to have an entire class of risk measures at hand such that every risk measure from this class behaves opportunely in stochastic programs.

Such a class will be introduced in section 2 for the one-period case, namely the class of polyhedral risk measures. Conditions implying that polyhedral risk measures are coherent and consistent with second order stochastic dominance are provided. In section 3 this class will be extended to the multiperiod case. Briefly, polyhedral risk measures are defined as optimal values of certain simple linear stochastic programs. In section 4 it will be shown that, indeed, several properties of expectation-based stochastic programs remain valid for stochastic programs with polyhedral risk measures as objectives. This is due to the fact that a problem of the form (1.1) with $\mathbb{E}$ replaced by a polyhedral risk measure ρ can easily be transformed into a stochastic program with additional variables and an objective consisting of the expectation of a linear function. In particular, we present stability results for two-stage stochastic programs with polyhedral risk measures and show that dual decomposition structures are maintained.

2. Polyhedral risk measures. We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and the linear space of real random variables $L_p(\Omega, \mathcal{F}, \mathbb{P})$ with some $p \in [1, \infty]$. According to [10, 11] a functional $\rho : L_p(\Omega, \mathcal{F}, \mathbb{P}) \to \bar{\mathbb{R}}$ is called a *risk measure* if it satisfies the following two conditions for all $z, \tilde{z} \in L_p(\Omega, \mathcal{F}, \mathbb{P})$:

(i) If $z \leq \tilde{z}$ a.s., then $\rho(z) \geq \rho(\tilde{z})$ *(monotonicity).*

(ii) For each $r \in \mathbb{R}$ we have $\rho(z + r) = \rho(z) - r$ *(translation invariance).*

A risk measure ρ is called *convex* if it satisfies the condition

$$\rho(\mu z + (1-\mu)\tilde{z}) \leq \mu\rho(z) + (1-\mu)\rho(\tilde{z})$$

for all $z, \tilde{z} \in L_p(\Omega, \mathcal{F}, \mathbb{P})$ and $\mu \in [0, 1]$. A convex risk measure is called *coherent* if it is *positively homogeneous*, i.e., $\rho(\mu z) = \mu\rho(z)$ for all $\mu \geq 0$ and $z \in L_p(\Omega, \mathcal{F}, \mathbb{P})$. There is a number of representation theorems for convex and especially for coherent risk measures in the literature emerging from convex duality. Next, we cite one of these representations adapted to our needs. Therefore, we set

$$\mathcal{D} := \{f \in L_1(\Omega, \mathcal{F}, \mathbb{P}) : f \geq 0 \text{ a.s.}, \mathbb{E}[f] = 1\},$$

the set of all density functions for $(\Omega, \mathcal{F}, \mathbb{P})$.

THEOREM 2.1. *Let $\rho : L_p(\Omega, \mathcal{F}, \mathbb{P}) \to \bar{\mathbb{R}}$ with $p \in [1, \infty]$. Assume that ρ is lower semicontinuous. Then ρ is a coherent risk measure if and only if the following condition holds:*

$$\exists\, \mathcal{P}_\rho \subseteq \mathcal{D} \text{ convex} : \rho(z) = \sup_{f \in \mathcal{P}_\rho} \mathbb{E}[-zf] \;\; \forall\, z \in L_p(\Omega, \mathcal{F}, \mathbb{P}).$$

Proof. "⇒" is stated in [35, Corollary 1] and "⇐" is easily seen by checking the four properties of the definition above; see also [11, 6, 28]. □

Now we are ready to define the class of polyhedral risk measures.

DEFINITION 2.2. *A risk measure ρ on $L_p(\Omega, \mathcal{F}, \mathbb{P})$ with some $p \in [1, \infty]$ will be called* polyhedral *if there exist $k_1, k_2 \in \mathbb{N}$, $c_1, w_1 \in \mathbb{R}^{k_1}$, $c_2, w_2 \in \mathbb{R}^{k_2}$, a nonempty polyhedral set $Y_1 \subseteq \mathbb{R}^{k_1}$, and a polyhedral cone $Y_2 \subseteq \mathbb{R}^{k_2}$ such that*

$$\rho(z) = \inf \left\{ \langle c_1, y_1 \rangle + \mathbb{E}[\langle c_2, y_2 \rangle] \;\middle|\; \begin{array}{l} y_1 \in Y_1, \\ y_2 \in L_p(\Omega, \mathcal{F}, \mathbb{P}),\ y_2 \in Y_2 \ a.s., \\ \langle w_1, y_1 \rangle + \langle w_2, y_2 \rangle = z \ a.s. \end{array} \right\} \tag{2.1}$$

for every $z \in L_p(\Omega, \mathcal{F}, \mathbb{P})$. Here, $\mathbb{E}$ denotes the expectation on $(\Omega, \mathcal{F}, \mathbb{P})$ and $\langle \cdot, \cdot \rangle$ a scalar product on $\mathbb{R}^{k_1}$ or $\mathbb{R}^{k_2}$.

Hence, expressed in the language of stochastic programming, a polyhedral risk measure is given as the optimal value of a certain two-stage stochastic program with random right-hand side. We use the term *polyhedral* because, for $\#\Omega < \infty$, the space $L_p(\Omega, \mathcal{F}, \mathbb{P})$ can be identified with $\mathbb{R}^{\#\Omega}$ and in this case a risk measure defined by (2.1) is indeed a polyhedral function on $\mathbb{R}^{\#\Omega}$.

Remark 2.3. Of course, the negative expectation is a polyhedral risk measure. Moreover, a convex combination of (negative) expectation and a polyhedral risk measure is again a polyhedral risk measure: Let $\mu \in [0,1]$ and ρ be a polyhedral risk measure with dimensions k_t, vectors c_t and w_t ($t = 1,2$), and polyhedral set/cone Y_1 / Y_2. Then the risk measure $\hat{\rho} := \mu\rho - (1-\mu)\mathbb{E}$ is polyhedral with the same dimensions k_t and the same sets Y_t and vectors $\hat{w}_1 := w_1$, $\hat{w}_2 := w_2$, $\hat{c}_1 := \mu c_1 - (1-\mu)w_1$, and $\hat{c}_2 := \mu c_2 - (1-\mu)w_2$. Thus, so-called *mean-risk models*, where expectation and risk are optimized simultaneously, do not need to be considered separately.

Next, we derive dual representations for (2.1). To this end, we do not need to assume that ρ is a risk measure in the sense of [10, 11], i.e., that it is monotone and translation invariant. We conclude in our first result that ρ is a convex functional. To state this result, we use the notation[1]

$$D_{\rho,t} := \{u \in \mathbb{R} : c_t + uw_t \in -Y_t^*\} \quad (t = 1,2)$$

for the so-called *dual feasible sets*.

THEOREM 2.4. *Let ρ be a functional of the form* (2.1) *on $L_p(\Omega, \mathcal{F}, \mathbb{P})$ with some $p \in [1, \infty)$. Assume*

(i) *complete recourse: $\langle w_2, Y_2 \rangle = \mathbb{R}$,*

(ii) *dual feasibility: $D_{\rho,1} \cap D_{\rho,2} \neq \emptyset$.*

Then ρ is finite, convex, and continuous. Further, the representation

$$\rho(z) = \inf_{y_1 \in Y_1} \left\{ \langle c_1, y_1 \rangle + \mathbb{E}\left[\max_{\ell=1,2} u_\ell \left(\langle w_1, y_1 \rangle - z \right) \right] \right\} \tag{2.2}$$

holds with two real numbers u_1 and u_2 that are the endpoints of $D_{\rho,2}$ which is a compact interval in $\mathbb{R}$. Furthermore, with $\frac{1}{p} + \frac{1}{p'} = 1$, ρ admits the dual representation

$$\rho(z) = \sup \left\{ -\mathbb{E}[\lambda z] + \inf_{y_1 \in Y_1} \langle c_1 + \mathbb{E}[\lambda] w_1, y_1 \rangle \;\middle|\; \begin{array}{l} \lambda \in L_{p'}(\Omega, \mathcal{F}, \mathbb{P}), \\ c_2 + \lambda w_2 \in -Y_2^* \ \textit{a.s.} \end{array} \right\}. \tag{2.3}$$

In particular, if Y_1 is a cone, then ρ is positively homogeneous and (2.3) *becomes*

$$\rho(z) = \sup \left\{ -\mathbb{E}[\lambda z] \;\middle|\; \begin{array}{l} \lambda \in L_{p'}(\Omega, \mathcal{F}, \mathbb{P}), \\ c_1 + \mathbb{E}[\lambda] w_1 \in -Y_1^*, \ c_2 + \lambda w_2 \in -Y_2^* \ \textit{a.s.} \end{array} \right\}. \tag{2.4}$$

Proof. Finiteness, convexity, continuity, and the representations (2.3) and (2.4) will be proved in a more general framework in section 3, Theorem 3.9. Representation (2.2) follows from LP duality applied to the second stage program. (Note that due to [29, Theorem 14.60] the minimization for the second stage can be carried out pointwise on Ω.) Namely, it holds for each $y_1 \in Y_1$ and each $z \in \mathbb{R}$ that

$$\begin{aligned} &\min \{\langle c_2, y_2 \rangle : y_2 \in Y_2, \ \langle w_1, y_1 \rangle + \langle w_2, y_2 \rangle = z\} \\ &= \max \{u (\langle w_1, y_1 \rangle - z) : c_2 + uw_2 \in -Y_2^*\}. \end{aligned}$$

[1]Thereby Y_t^* denotes the *polar cone* of Y_t. For a nonempty set Y the polar cone Y^* is defined by $Y^* = \{y^* : \langle y, y^* \rangle \leq 0 \ \forall y \in Y\}$.

Due to complete recourse and dual feasibility the feasible sets of both problems are nonempty and the joint optimal value is finite for each $y_1 \in Y_1$ and each $z \in \mathbb{R}$. Since the expression $\langle w_1, y_1\rangle - z$ can reach any real number and the feasible set of the right problem $D_{\rho,2}$ does not depend on y_1 and z, it is clear that the latter is bounded, i.e., it is a compact interval in $\mathbb{R}$. Of course, the maximum is attained for u being an endpoint of $D_{\rho,2}$. ∎

If a functional ρ on $L_p(\Omega, \mathcal{F}, \mathbb{P})$ is defined by formula (2.1), the question arises for which choice of c_t, w_t, and Y_t $(t = 1, 2)$ this functional is a (convex) risk measure in the sense of [10, 11]. Formula (2.4) provides a sufficient criterion for a functional of the form (2.1) to be a coherent risk measure in case Y_1 is a cone.

COROLLARY 2.5. *Let ρ be a functional on $L_p(\Omega, \mathcal{F}, \mathbb{P})$ of the form* (2.1) *with Y_1 being a polyhedral cone and $1 \le p < \infty$. Let the conditions of Theorem* 2.4 *be satisfied (complete recourse, dual feasibility) and assume that*

$$\Lambda_\rho := \left\{ \lambda \in L_{p'}(\Omega, \mathcal{F}, \mathbb{P}) \;\middle|\; \begin{array}{l} c_1 + \mathbb{E}[\lambda]\, w_1 \in -Y_1^*, \\ c_2 + \lambda w_2 \in -Y_2^* \;\; a.s. \end{array} \right\} \subseteq \mathcal{D}. \tag{2.5}$$

Then ρ is a coherent risk measure.

Proof. The proof follows immediately from Theorems 2.1 and 2.4 with $\mathcal{P}_\rho := \Lambda_\rho$ since, of course, continuity implies lower semicontinuity. ∎

The following result provides a sufficient criterion for a functional of the form (2.1) to be a convex risk measure in case Y_1 is *not* a cone.

PROPOSITION 2.6. *Let ρ be a functional on $L_p(\Omega, \mathcal{F}, \mathbb{P})$ of the form* (2.1) *with $p \in [1, \infty)$. Assume that complete recourse and dual feasibility hold and that $D_{\rho,2} \subseteq \mathbb{R}_+$ and let c_1, w_1, and Y_1 be of the form $c_1 = (\hat{c}_1, 1)$, $w_1 = (\hat{w}_1, -1)$, and $Y_1 = \hat{Y}_1 \times \mathbb{R}$, where $\hat{w}_1, \hat{c}_1 \in \mathbb{R}^{k_1-1}$, and $\hat{Y}_1 \subseteq \mathbb{R}^{k_1-1}$. Then ρ is a (polyhedral) convex risk measure.*

Proof. Finiteness and convexity of ρ follow from Theorem 2.4. The monotonicity property (i) follows from the representation (2.2) and the fact that u_1 and u_2 are nonnegative. Indeed, let $z, \tilde{z} \in L_p(\Omega, \mathcal{F}, \mathbb{P})$ be such that $z \le \tilde{z}$ a.s.; then we have $\mathbb{E}[\max_{\ell=1,2} u_\ell(\langle w_1, y_1\rangle - z)] \ge \mathbb{E}[\max_{\ell=1,2} u_\ell(\langle w_1, y_1\rangle - \tilde{z})]$ for every $y_1 \in Y_1$. The translation invariance condition (ii) follows by setting $y_1 = (\hat{y}_1, \bar{y}_1)$, $\tilde{y}_1 := \bar{y}_1 + r \in \mathbb{R}$ as a consequence of the identity

$$\begin{aligned}
&\rho(z + r) \\
&= \inf \Big\{ \langle \hat{c}_1, \hat{y}_1\rangle + \bar{y}_1 + \mathbb{E}\left[\max\nolimits_{\ell=1,2} u_\ell\left(\langle \hat{w}_1, \hat{y}_1\rangle - \bar{y}_1 - (z + r)\right)\right] : \hat{y}_1 \in \hat{Y}_1,\ \bar{y}_1 \in \mathbb{R} \Big\} \\
&= \inf \Big\{ \langle \hat{c}_1, \hat{y}_1\rangle + \tilde{y}_1 + \mathbb{E}\left[\max\nolimits_{\ell=1,2} u_\ell\left(\langle \hat{w}_1, \hat{y}_1\rangle - \tilde{y}_1 - z\right)\right] : \hat{y}_1 \in \hat{Y}_1,\ \tilde{y}_1 \in \mathbb{R} \Big\} - r \\
&= \rho(z) - r
\end{aligned}$$

for each $r \in \mathbb{R}$ and $z \in L_p(\Omega, \mathcal{F}, \mathbb{P})$. ∎

The assumptions of Proposition 2.6 guarantee even a stronger type of monotonicity than imposed earlier for risk measures. Such stronger monotonicity properties are based on so-called *integral stochastic orders* or *stochastic dominance rules* (see [15] for a recent survey). For real random variables z and $\tilde{z}$ in $L_1(\Omega, \mathcal{F}, \mathbb{P})$, stochastic dominance rules are defined by classes $\mathcal{F}$ of measurable real-valued functions on $\mathbb{R}$. A stochastic dominance rule is defined by

$$z \preceq_{\mathcal{F}} \tilde{z} \quad \text{if} \quad \mathbb{E}[f(z)] \le \mathbb{E}[f(\tilde{z})]$$

for each $f \in \mathcal{F}$ such that the expectations exist. Important special cases are the class of $\mathcal{F}_{nd}$ of nondecreasing functions and the class $\mathcal{F}_{ndc}$ of nondecreasing concave functions. In these cases the rules are also called *first order stochastic dominance* and *second order stochastic dominance* and denoted by $\preceq_{FSD}$ and $\preceq_{SSD}$, respectively. Clearly, $z \preceq_{FSD} \tilde{z}$ implies $z \preceq_{SSD} \tilde{z}$. The relation $z \preceq_{FSD} \tilde{z}$ is equivalent to $\mathbb{P}(z \leq t) \geq \mathbb{P}(\tilde{z} \leq t)$ for each $t \in \mathbb{R}$. Furthermore, $z \preceq_{SSD} \tilde{z}$ is equivalent to the condition $\mathbb{E}[\min\{z,t\}] \leq \mathbb{E}[\min\{\tilde{z},t\}]$ for each $t \in \mathbb{R}$ (cf. [15, section 8]). Other equivalent characterizations of $z \preceq_{SSD} \tilde{z}$ are $\int_{-\infty}^{\eta} \mathbb{P}(z \leq t)dt \geq \int_{-\infty}^{\eta} \mathbb{P}(\tilde{z} \leq t)dt$ for each $t \in \mathbb{R}$ (cf. [17, 18]) and $\int_0^p q_\alpha(z)d\alpha \leq \int_0^p q_\alpha(\tilde{z})d\alpha$ for each $p \in (0,1]$ (cf. [19]) with $q_\alpha(z) = \inf\{r \in \mathbb{R} : \mathbb{P}(z \leq r) \geq \alpha\}$ denoting the (lower) α-quantile of the random variable z.

In [19, 17, 18] the consistency of risk measures ρ with certain stochastic dominance rules $\preceq_{\mathcal{F}}$ is studied. In particular, it is said that ρ is *consistent with second order stochastic dominance* if $z \preceq_{SSD} \tilde{z}$ implies $\rho(z) \geq \rho(\tilde{z})$.

PROPOSITION 2.7. *Let ρ be a functional on $L_p(\Omega, \mathcal{F}, \mathbb{P})$ of the form* (2.1) *with $p \in [1, \infty)$. Assume that complete recourse and dual feasibility hold and that $D_{\rho,2} \subseteq \mathbb{R}_+$. Then ρ is consistent with second order stochastic dominance.*

Proof. Due to Theorem 2.4 the representation (2.2) holds with $u_1, u_2 \in \mathbb{R}_+$. Define for $y_1 \in Y_1$ the real-valued function g_{y_1} given by

$$g_{y_1}(t) := \langle c_1, y_1 \rangle + \max_{\ell=1,2} u_\ell \left(\langle w_1, y_1 \rangle - t \right)$$

for $t \in \mathbb{R}$. Note that g_{y_1} is convex and, because of $u_1, u_2 \geq 0$, nonincreasing. Let $z \preceq_{SSD} \tilde{z}$. Then $\mathbb{E}[-g_{y_1}(z)] \leq \mathbb{E}[-g_{y_1}(\tilde{z})]$ for all $y_1 \in Y_1$ and, thus, $\rho(z) = \inf_{y_1 \in Y_1} \mathbb{E}[g_{y_1}(z)] \geq \inf_{y_1 \in Y_1} \mathbb{E}[g_{y_1}(\tilde{z})] = \rho(\tilde{z})$. ∎

Remark 2.8. For a risk measure ρ on $L_p(\Omega, \mathcal{F}, \mathbb{P})$ the *acceptance set* $\mathcal{A}_\rho$ is defined by $\mathcal{A}_\rho = \{z \in L_p(\Omega, \mathcal{F}, \mathbb{P}) : \rho(z) \leq 0\}$ [3, 11]; let the conditions of Theorem 2.4 be satisfied. Then, since ρ is a convex functional, $\mathcal{A}_\rho$ is a convex set. If, in addition, Y_1 is a cone, then $\mathcal{A}_\rho$ is a convex cone. Regarding (2.5) it is obvious that

$$\mathcal{A}_\rho = \{z \in L_p(\Omega, \mathcal{F}, \mathbb{P}) \mid \forall\, \lambda \in \Lambda_\rho : \mathbb{E}\left[\lambda z\right] \geq 0\} = -\Lambda_\rho^*$$

in this case. Of course, if $\Omega = \{\omega_1, \ldots, \omega_S\}$, then Λ_ρ is a polyhedron in $\mathbb{R}^S$, thus $\mathcal{A}_\rho = -\Lambda_\rho^*$ is a polyhedral cone.

For stability analysis of stochastic programs (cf. section 4.1), it is important to know whether first stage solution sets are bounded or not. For a polyhedral risk measure ρ satisfying complete recourse and dual feasibility, the first stage solution set $S(\rho(z)) \subseteq Y_1$ can be written according to the dual representation (2.2) as

$$S(\rho(z)) := \{y_1 \in Y_1 : \langle c_1, y_1 \rangle + \mathbb{E}\left[\max\nolimits_{\ell=1,2} u_\ell \left(\langle w_1, y_1 \rangle - z \right)\right] = \rho(z)\}. \tag{2.6}$$

The following proposition provides a sufficient criterion for the boundedness of $S(\rho(z))$ for a large class of polyhedral risk measures.

PROPOSITION 2.9. *Let ρ be a functional on $L_p(\Omega, \mathcal{F}, \mathbb{P})$ of the form* (2.1) *with $p \in [1, \infty)$. Let the conditions of Theorem* 2.4 *be satisfied (complete recourse, dual feasibility) and assume that $S(\rho(0))$ is a nonempty, bounded subset in $\mathbb{R}^{k_1}$. Then $S(\rho(z))$ is nonempty, convex, and compact for any $z \in L_p(\Omega, \mathcal{F}, \mathbb{P})$.*

Proof. Clearly, Theorem 2.4 implies convexity and closedness of $S(\rho(z))$. It remains to be seen whether $S(\rho(z))$ is nonempty and bounded. The polyhedral set Y_1 can be represented in the form $Y_1 = P_1 + C_1$, where P_1 is a bounded polyhedron and C_1 a polyhedral cone (e.g., [29, Corollary 3.53]). Let $0 \neq d_1 \in C_1$ (hence, $\mu d_1 \in C_1$ for any $\mu \geq 0$) and $g_{d_1}(0) = \langle c_1, d_1 \rangle + \max_{\ell=1,2} u_\ell \langle w_1, d_1 \rangle$. Next we show $g_{d_1}(0) > 0$. Suppose $g_{d_1}(0) < 0$ and let $p_1 \in P_1$, $\mu > 0$. Then $p_1 + \mu d_1 \in Y_1$ and we obtain

$\rho(0) \leq g_{p_1}(0) + \mu g_{d_1}(0)$. This contradicts to the finiteness of ρ since $\mu > 0$ may be chosen arbitrarily large. If $g_{d_1}(0) = 0$, the set $S(\rho(0))$ would contain the unbounded subset $\{\bar{y}_1 + \mu d_1 : \mu \geq 0\}$ for some $\bar{y}_1 \in S(\rho(0))$. Now, let $z \in L_p(\Omega, \mathcal{F}, \mathbb{P})$ and let $(y_{1,n})$ be a sequence with $y_{1,n} = p_{1,n} + d_{1,n} \in Y_1$, $p_{1,n} \in P_1$, $d_{1,n} \in C_1$, and

$$\langle c_1, y_{1,n}\rangle + \mathbb{E}\left[\max\nolimits_{\ell=1,2} u_\ell\left(\langle w_1, y_{1,n}\rangle - z\right)\right] \to \rho(z).$$

Since P_1 is bounded, we may assume without loss of generality that $(p_{1,n})$ is convergent to some $\bar{p}_1 \in P_1$. Suppose that $(y_{1,n})$ is unbounded. Then we may assume without loss of generality that $\|d_{1,n}\| \to \infty$ and $\frac{d_{1,n}}{\|d_{1,n}\|} \to \bar{d}_1 \in C_1$. It follows that

$$\rho(z) = \lim_{n\to\infty}\left(\langle c_1, y_{1,n}\rangle + \mathbb{E}\left[\max\nolimits_{\ell=1,2} u_\ell\left(\langle w_1, y_{1,n}\rangle - z\right)\right]\right) = \lim_{n\to\infty} \|d_{1,n}\|\alpha_n$$

with $\alpha_n := \langle c_1, \frac{y_{1,n}}{\|d_{1,n}\|}\rangle + \mathbb{E}[\max_{\ell=1,2} u_\ell(\langle w_1, \frac{y_{1,n}}{\|d_{1,n}\|}\rangle - \frac{z}{\|d_{1,n}\|})]$. Obviously, it holds that $\alpha_n \to g_{\bar{d}_1}(0) > 0$, hence $\rho(z) = \lim_{n\to\infty} \|y_{1,n}\|\alpha_n = \infty$. This is a contradiction. It follows that each minimizing sequence $(y_{1,n})$ in Y_1 is always bounded. This implies both existence of a solution and boundedness of the solution set $S(\rho(z))$. ∎

Example 2.10. We consider the *Conditional-* or *Average-Value-at-Risk* at level $\alpha \in (0,1)$ ($CVaR_\alpha$ or $AVaR_\alpha$) defined by

$$CVaR_\alpha(z) := \frac{1}{\alpha}\int_0^\alpha VaR_\gamma(z)d\gamma = \inf_{r\in\mathbb{R}}\left\{r + \frac{1}{\alpha}\mathbb{E}\left[(r+z)^-\right]\right\}, \tag{2.7}$$

where $VaR_\gamma(z) := \inf\{r \in \mathbb{R} : \mathbb{P}(z+r < 0) \leq \gamma\} = -\bar{q}_\gamma(z)$ is the Value-at-Risk at level $\gamma \in (0,1)$ (see [11, section 4.4] and [27]) and $a^- = \max\{0, -a\}$ denotes the negative part of a real number a. The number $\bar{q}_\gamma(z)$ is also called the upper γ-quantile of z. Introducing variables for positive and negative parts of the infimum representation in (2.7), respectively, leads to

$$CVaR_\alpha(z) = \inf\left\{ y_1 + \frac{1}{\alpha}\mathbb{E}\left[y_2^{(2)}\right] \;\middle|\; \begin{array}{l} y_1 \in \mathbb{R},\ y_2 \in L_1(\Omega, \mathcal{F}, \mathbb{P}), \\ y_2 \in \mathbb{R}_+ \times \mathbb{R}_+ \text{ a.s.}, \\ y_2^{(1)} - y_2^{(2)} = z + y_1 \text{ a.s.} \end{array}\right\}. \tag{2.8}$$

Thus, $CVaR_\alpha$ is of the form (2.1) by setting $k_1 = 1$, $k_2 = 2$, $w_1 = -1$, $c_1 = 1$, $c_2 = (0, \frac{1}{\alpha})$, $w_2 = (1, -1)$, $Y_1 = \mathbb{R}$, and $Y_2 = \mathbb{R}^2_+$, and, hence, is a polyhedral risk measure. Moreover, $\langle w_2, Y_2\rangle = \mathbb{R}$, $D_{\rho,1} = D_{\rho,1} \cap D_{\rho,2} = \{1\}$, and $D_{\rho,2} = [0, \frac{1}{\alpha}] \subseteq \mathbb{R}_+$, thus the dual representation (2.4) holds and $CVaR_\alpha$ is consistent with second order stochastic dominance. The representation (2.2) holds with $u_1 = 0$ and $u_2 = \frac{1}{\alpha}$. The condition $c_2 + \lambda w_2 \in -Y_2^*$ in the dual representation (2.4) is equivalent to $\lambda \in [0, \frac{1}{\alpha}]$. Hence, (2.4) becomes

$$CVaR_\alpha(z) = \sup\left\{-\mathbb{E}\left[\lambda z\right] : \lambda \in L_{p'}(\Omega, \mathcal{F}, \mathbb{P}),\ \lambda \in \left[0, \tfrac{1}{\alpha}\right] \text{ a.s.}, \mathbb{E}\left[\lambda\right] = 1\right\} \tag{2.9}$$

for each $z \in L_p(\Omega, \mathcal{F}, \mathbb{P})$, $1 \leq p < \infty$. Corollary 2.5 applies thus, $CVaR$ is a coherent risk measure, too. Similar results have already been shown in [28, 19]. Furthermore, it is shown in [27] that the set $\{r \in \mathbb{R} : CVaR_\alpha(z) = r + \frac{1}{\alpha}\mathbb{E}[(r+z)^-]\}$ of first stage solutions is just the interval $[-\bar{q}_\alpha(z), -q_\alpha(z)]$, i.e., the set of all negative α-quantiles of z. Indeed, Proposition 2.9 is inspired by the latter result.

Example 2.11. Consider the *expected regret* or *expected loss* defined by

$$EL(z) = \mathbb{E}\left[(z-\gamma)^-\right]$$

with some fixed target $\gamma \in \mathbb{R}$. This functional, too, can be written in the form (2.1) with $k_1 = 1$, $k_2 = 2$, $w_1 = 1$, $c_1 = 0$, $c_2 = (0, 1)$, $w_2 = (1, -1)$, $Y_1 = \{\gamma\}$, $Y_2 = \mathbb{R}_+ \times \mathbb{R}_+$. Note that, actually, Y_1 is *not* a cone here. Further, $\langle w_2, Y_2\rangle = \mathbb{R}$, $D_{\rho,1} \cap D_{\rho,2} \neq \emptyset$, and $D_{\rho,2} = [0, 1] \subseteq \mathbb{R}_+$, thus the dual representations (2.2) and (2.3) hold and ρ is consistent with second order stochastic dominance. However, ρ is not translation invariant, i.e., not a risk measure in the sense of [10, 11]. Nevertheless, it is used as a risk measure in some applications.

Example 2.12. The utilization of deviation and semideviation measures in stochastic optimization goes back to [14] and is further discussed, e.g., in [17, 18, 19, 28]. For $k \geq 1$ deviation and semideviation are defined by

$$D_k(z) := \left(\mathbb{E}\left[|z - \mathbb{E}[z]|^k\right]\right)^{1/k} \qquad SD_k(z) := \left(\mathbb{E}\left[\left((z - \mathbb{E}[z])^-\right)^k\right]\right)^{1/k},$$

respectively. They are closely related to coherent risk measures (cf. [28]), $-\mathbb{E} + \beta \cdot D_k$ and $-\mathbb{E}+\beta\cdot SD_k$ with $\beta \geq 0$ are translation invariant in the sense of [10, 11] and, hence, candidates for coherent risk measure. However, they are *not* within the framework of polyhedral risk measures, even $SD_1 = \frac{1}{2}D_1$ cannot be written in the form (2.1). But, if we change from expectation $\mathbb{E}[z]$ to the median $q_{\frac{1}{2}}(z)$, then we obtain the median-deviation which is a special case of the so-called dispersion measures at level $\alpha \in (0, 1)$ given by

$$d_\alpha(z) := \mathbb{E}\left[\alpha(z - q_\alpha)^+ + (1 - \alpha)(z - q_\alpha)^-\right] \qquad d_{\frac{1}{2}}(z) = \tfrac{1}{2}\mathbb{E}\left[\left|z - q_{\frac{1}{2}}(z)\right|\right]$$

(cf. [19, 40]). These functionals are polyhedral with $k_1 = 1$, $k_2 = 2$, $c_1 = 0$, $c_2 = (\alpha, 1-\alpha)$, $w_1 = 1$, $w_2 = (1, -1)$, $Y_1 = \mathbb{R}$, and $Y_1 = \mathbb{R}_+ \times \mathbb{R}_+$. Again, $\rho := -\mathbb{E}+\beta\cdot d_\alpha$ is a candidate for a coherent risk measure. According to Remark 2.3 also ρ is polyhedral with $c_1 = -1$, $c_2 = (\alpha\beta-1, (1-\alpha)\beta+1)$, and w_t and Y_t as above. Hence, $D_{\rho,1} = \{1\}$, $D_{\rho,2} = [1-\alpha\beta, 1+(1-\alpha)\beta]$, and $\Lambda_\rho = \{\lambda : \mathbb{E}[\lambda] = 1, \lambda \in [1-\alpha\beta, 1+(1-\alpha)\beta] \text{ a.s.}\}$, i.e., ρ is coherent and second order stochastic dominance consistent if $\beta \leq \frac{1}{\alpha}$ (see also [19]). However, the latter representation reveals that $\rho = -(1 - \alpha\beta)\mathbb{E} + \alpha\beta \cdot CVaR_\alpha$, i.e., quantile dispersion and Conditional-Value-at-Risk is basically the same thing.

3. Multiperiod risk. When random variables $z_1, \ldots, z_T$ with $z_t \in L_p(\Omega, \mathcal{F}_t, \mathbb{P})$, $p \geq 1$, are considered and the available information is revealed with the passing of time, it may become necessary to use multiperiod risk measures (see [3, 2, 22, 25, 42, 36]). We assume that a filtration of σ-fields $\mathcal{F}_t$, $t = 1, \ldots, T$, is given, i.e., $\mathcal{F}_t \subseteq \mathcal{F}_{t+1} \subseteq \mathcal{F}$, and that $\mathcal{F}_1 = \{\emptyset, \Omega\}$, i.e., that z_1 is always deterministic. We will now generalize the concepts of the previous section to this multiperiod framework.

Remark 3.1. When dealing with multiperiod risk measures one has to determine whether the random variables represent (potentially financial) *incomes* or *payments* as, e.g., in [22, 36, 42], or if they have to be understood in a cumulative sense, i.e., as a *wealth* or *value process* as in [3, 2]. Of course, the one can easily be transformed into the other: If Z_t is an income, then one can consider accumulation $z_t = Z_1 + \cdots + Z_t$, and if z_t is an accumulated value, then the income is given by $Z_t = z_t - z_{t-1}$. Throughout this paper we consider $z = (z_1, \ldots, z_T)$ to be a value process.

We give the definition of coherence in the multiperiod case as introduced[2] in [3, 2].

[2]In [3, 2] the definition is slightly different since another framework was considered: The first time stage (i.e., the deterministic stage) was denoted by index 0. Here, the formulation is adapted to our framework with index 1 for the deterministic time stage (i.e., $\mathcal{F}_1 = \{\emptyset, \Omega\}$).

DEFINITION 3.2. *A functional ρ on $\times_{t=1}^{T} L_p(\Omega, \mathcal{F}_t, \mathbb{P})$ is called a* multiperiod coherent risk measure *if the following:*

(i) *if $z_t \le \tilde{z}_t$ a.s., $t = 1, \dots, T$, then $\rho(z_1, \dots, z_T) \ge \rho(\tilde{z}_1, \dots, \tilde{z}_T)$* (monotonicity)*;*

(ii) *for each $r \in \mathbb{R}$ we have $\rho(z_1+r, \dots, z_T+r) = \rho(z)-r$* (translation invariance)*;*

(iii) *$\rho(\mu z_1 + (1-\mu)\tilde{z}_1, \dots, \mu z_T + (1-\mu)\tilde{z}_T) \le \mu\rho(z_1, \dots, z_T) + (1-\mu)\rho(\tilde{z}_1, \dots, \tilde{z}_1)$ for $\mu \in [0, 1]$* (convexity)*;*

(iv) *for $\mu \ge 0$ we have $\rho(\mu z_1, \dots, \mu z_T) = \mu\rho(z_1, \dots, z_T)$* (positive homogeneity).

Remark 3.3. How translation invariance is to be defined in the multiperiod case is still subject to discussion in the ongoing research in financial mathematics. Different suggestions were made, e.g., in [36, 25, 42] such that nonrandom amounts can be shifted in time by means of credits. However, from the viewpoint of capital requirement and optimization it appears reasonable to keep with [3, 2].

Example 3.4. In [3, Example 3] it was shown that $\rho(z) := -\mathbb{E}[\min\{z_1, \dots, z_T\}]$ with $z = (z_1, \dots, z_T)$ is a multiperiod coherent risk measure on $\times_{t=1}^{T} L_\infty(\Omega, \mathcal{F}_t, \mathbb{P})$.

Remark 3.5. Let ρ_t be (one-period) coherent risk measures on $L_p(\Omega, \mathcal{F}_t, \mathbb{P})$, $t = 1, \dots, T$. Let further $\emptyset \ne S \subseteq \{1, \dots, T\}$. Then $\rho(z) := \max_{t \in S} \rho_t(z_t)$ is multiperiod coherent. Let $\mu_t \in \mathbb{R}_+$, $t = 1, \dots, T$, with $\sum_{t=1}^{T} \mu_t = 1$. Then also $\rho(z) := \sum_{t=1}^{T} \mu_t \rho_t(z_t)$ is a multiperiod coherent risk measure. This can easily be verified by checking the four properties of Definition 3.2.

As shown in [3, 2], the representation result for (one-period) risk measures (Theorem 2.1) can be carried over to the multiperiod case. Therefore, the set of densities $\mathcal{D}$ is extended such that the integrals of the time steps sum up to one,

$$\mathcal{D}_T := \Big\{ f \in \times_{t=1}^{T} L_1(\Omega, \mathcal{F}_t, \mathbb{P}) : f_t \ge 0 \text{ a.s. } (t = 1, \dots, T),\ \textstyle\sum_{t=1}^{T} \mathbb{E}\,[f_t] = 1 \Big\}.$$

THEOREM 3.6. *Let $\rho : \times_{t=1}^{T} L_p(\Omega, \mathcal{F}_t, \mathbb{P}) \to \bar{\mathbb{R}}$ and assume that ρ is lower semicontinuous. Then ρ is a multiperiod coherent risk measure if and only if the following condition holds:*

$$\exists\, \mathcal{P}_\rho \subseteq \mathcal{D}_T \text{ convex} : \rho(z) = \sup\Big\{ \textstyle\sum_{t=1}^{T} \mathbb{E}\,[-z_t f_t] : f \in \mathcal{P}_\rho \Big\}. \tag{3.1}$$

Proof. We follow the ideas of [3, 2], but in reverse order. Obviously, ρ is coherent if and only if the corresponding one-period risk measure ρ' on $L_p(\Omega', \mathcal{F}', \mathbb{P}')$ is coherent in the usual sense, where $(\Omega', \mathcal{F}', \mathbb{P}')$ and ρ' are defined as follows:

$$\begin{aligned}
\Omega' &:= \Omega \times \{1, \dots, T\} \\
\mathcal{F}' &:= \Big\{ \textstyle\bigcup_{t=1}^{T} (A_t \times \{t\}) : A_t \in \mathcal{F}_t \Big\} \\
\mathbb{P}' \Big(\textstyle\bigcup_{t=1}^{T} (A_t \times \{t\}) \Big) &:= \tfrac{1}{T} \textstyle\sum_{t=1}^{T} \mathbb{P}(A_t) \\
\rho'(z') &:= \rho\,(z(z'))
\end{aligned}$$

and $z(z')$ is defined by $z(z')(\omega) := (z'(\omega, 1), z'(\omega, 2), \dots, z'(\omega, T))$. Theorem 2.1 says that there exists a convex set of density functions $\mathcal{P}'_\rho \subseteq \mathcal{D}$ such that, for $z \in \times_{t=1}^{T} L_p(\Omega, \mathcal{F}_t, \mathbb{P})$,

$$\rho(z) = \rho'(z'(z)) = \sup\left\{ \mathbb{E}'\,[-z'f'] : f' \in \mathcal{P}'_\rho \right\}$$

with $z'(z)(\omega, t) := z_t(\omega)$. Note that also the conditions from Definition 3.2 are equivalent to those from Theorem 2.1 for $(\Omega', \mathcal{F}', \mathbb{P}')$ and that lower semicontinuity of ρ is equivalent to lower semicontinuity of ρ'. By setting

$$\mathcal{P}_\rho := \left\{ f = \left(\tfrac{1}{T} f'(.,1), \tfrac{1}{T} f'(.,2), \ldots, \tfrac{1}{T} f'(.,T)\right) : f' \in \mathcal{P}'_\rho \right\},$$

the assertion follows. $\square$

Now we are ready to extend Definition 2.2 to the multiperiod case.

DEFINITION 3.7. *A multiperiod risk measure ρ on $\times_{t=1}^T L_p(\Omega, \mathcal{F}_t, \mathbb{P})$ with $p \in [1, \infty]$ is called* multiperiod polyhedral *if there are $k_t \in \mathbb{N}$, $c_t \in \mathbb{R}^{k_t}$, $t = 1, \ldots, T$, $w_{t\tau} \in \mathbb{R}^{k_{t-\tau}}$, $t = 1, \ldots, T$, $\tau = 0, \ldots, t-1$, a polyhedral set $Y_1 \subseteq \mathbb{R}^{k_1}$, and polyhedral cones $Y_t \subseteq \mathbb{R}^{k_t}$, $t = 2, \ldots, T$, such that*

$$\rho(z) = \inf \left\{ \mathbb{E}\left[\sum_{t=1}^T \langle c_t, y_t \rangle\right] \;\middle|\; \begin{array}{ll} y_t \in L_p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^{k_t}), & \\ y_t \in Y_t \;\; a.s., & (t = 1, \ldots, T) \\ \sum_{\tau=0}^{t-1} \langle w_{t,\tau}, y_{t-\tau} \rangle = z_t \;\; a.s. & \end{array} \right\}. \tag{3.2}$$

Remark 3.8. The reader might wonder why, for $T = 2$, this definition does not precisely coincide with the Definition 2.2 for the one-period case. This is due to the fact that, in the literature, the risk of a process $z_1, \ldots, z_T$ is allowed to depend also on z_1 although this value is constant, i.e., deterministic (see [3, 2, 25]), whereas one-period risk depends on one scalar random variable only. Nevertheless, the one-period case can be regarded as a special case of Definition 3.7 because for $T = 2$ the parameters Y_1, c_1, and $w_{1,0}$ can easily be chosen such that z_1 does not contribute to the optimal value of (3.2).

THEOREM 3.9. *Let ρ be a functional of the form* (3.2) *on $\times_{t=1}^T L_p(\Omega, \mathcal{F}_t, \mathbb{P})$ with $p \in [1, \infty)$. Assume*

(i) *complete recourse: $\langle w_{t,0}, Y_t \rangle = \mathbb{R}$ $(t = 1, \ldots, T)$,*

(ii) *dual feasibility: $\{u \in \mathbb{R}^T : c_t + \sum_{\nu=t}^T u_\nu w_{\nu,\nu-t} \in -Y_t^* \; (t = 1, \ldots, T)\} \neq \emptyset$.*

Then ρ is finite, convex, and continuous on $\times_{t=1}^T L_p(\Omega, \mathcal{F}_t, \mathbb{P})$ and with $\frac{1}{p} + \frac{1}{p'} = 1$ the following dual representation holds:

$$\rho(z) = \sup \left\{ \begin{array}{l} \inf\limits_{y_1 \in Y_1} \left\langle c_1 + \sum_{\nu=1}^T \mathbb{E}[\lambda_\nu] \, w_{\nu,\nu-1}, y_1 \right\rangle \\ - \mathbb{E}\left[\sum_{t=1}^T \lambda_t z_t\right] \end{array} \;\middle|\; \begin{array}{l} \lambda_t \in L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}) \; (t = 1, \ldots, T), \\ c_t + \sum_{\nu=t}^T \mathbb{E}[\lambda_\nu | \mathcal{F}_t] \, w_{\nu,\nu-t} \in -Y_t^* \\ a.s. \; (t = 2, \ldots, T) \end{array} \right\}. \tag{3.3}$$

If, in addition, Y_1 is a polyhedral cone, then ρ is positively homogeneous and (3.3) *simplifies to*

$$\rho(z) = \sup \left\{ -\mathbb{E}\left[\sum_{t=1}^T \lambda_t z_t\right] \;\middle|\; \begin{array}{l} \lambda_t \in L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}), \\ c_t + \sum\limits_{\nu=t}^T \mathbb{E}[\lambda_\nu | \mathcal{F}_t] \, w_{\nu,\nu-t} \in -Y_t^* \;\; a.s. \\ (t = 1, \ldots, T) \end{array} \right\}. \tag{3.4}$$

Proof. We use results on conjugate duality (see [26] and [5, section 2.5.1]). Consider the Banach spaces and their duals

$$E := \times_{t=1}^T L_p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^{k_t}) \qquad E^* = \times_{t=1}^T L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^{k_t})$$

$$Z := \times_{t=1}^T L_p(\Omega, \mathcal{F}_t, \mathbb{P}) \qquad Z^* = \times_{t=1}^T L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P})$$

with bilinear forms $\langle e, e^*\rangle_{E/E^*} = \sum_{t=1}^T \mathbb{E}[\langle e_t, e_t^*\rangle_{\mathbb{R}^{k_t}}]$ and $\langle z, z^*\rangle_{Z/Z^*} = \sum_{t=1}^T \mathbb{E}[z_t z_t^*]$, respectively. Due to the complete recourse assumption it holds that $\rho(z) < \infty$ for all $z = (z_1, \dots, z_T) \in E$. Define $Y := \{y \in E : y_t(\omega) \in Y_t \ (t = 1, \dots, T) \text{ for a.e. } \omega \in \Omega\}$, $K = \sum_{t=1}^T k_t$ and

$$\begin{aligned} &\varphi : E \times Z \to \bar{\mathbb{R}} \\ &\qquad (y, z) \mapsto \varphi(y, z) := \langle y, c\rangle_{E/E^*} + \delta_Y(y) + \delta_{\{0\}}(Wy - z) \end{aligned}$$

with δ denoting the indicator function (taking values 0 and $+\infty$ only) and with

$$c = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_T \end{pmatrix} \in \mathbb{R}^K \qquad W = \begin{pmatrix} w'_{1,0} & 0 & 0 & \cdots & 0 \\ w'_{2,1} & w'_{2,0} & 0 & \cdots & 0 \\ w'_{3,2} & w'_{3,1} & w'_{3,0} & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ w'_{T,T-1} & w'_{T,T-2} & w'_{T,T-3} & \cdots & w'_{T,0} \end{pmatrix} \in \mathbb{R}^{T \times K}.$$

Note that φ is proper, lower semicontinuous, and convex since Y is convex. With these notations Definition 3.7 reads $\rho(z) = \inf_{y \in E} \varphi(y, z)$ and due to [5, Proposition 2.143] ρ is convex. The (conjugate) dual problem according to [5] is given by

$$\rho^*(z) = \sup\left\{\langle z, z^*\rangle_{Z/Z^*} - \varphi^*(0, z^*) : z^* \in Z^*\right\} \tag{3.5}$$

in which the conjugate φ^* is given by

$$\begin{aligned} \varphi^*(y^*, z^*) &= \sup\left\{\langle y, y^*\rangle_{E/E^*} + \langle z, z^*\rangle_{Z/Z^*} - \varphi(y, z) : y \in E,\, z \in Z\right\} \\ &= \sup\left\{\langle y, y^* - c\rangle_{E/E^*} + \langle z, z^*\rangle_{Z/Z^*} : y \in Y,\, z = Wy \text{ a.s.}\right\} \\ &= \sup\left\{\langle y, y^* - c\rangle_{E/E^*} + \langle Wy, z^*\rangle_{Z/Z^*} : y \in Y\right\} \\ &= \sup\left\{\langle y, y^* - c + W^* z^*\rangle_{E/E^*} : y \in Y\right\} \end{aligned}$$

with $W^* : Z^* \to E^*$ denoting the adjoint operator implicitly defined by the relation $\langle Wy, z^*\rangle_{Z/Z^*} = \langle y, W^* z^*\rangle_{E/E^*}$ for $y \in E$, $z^* \in Z^*$. Thereby, the matrix W is understood as mapping from E to Z. For the adjoint operator W^* it holds that

$$\begin{aligned} \langle y, W^* z^*\rangle_{E/E^*} &= \langle Wy, z^*\rangle_{Z/Z^*} = \textstyle\sum_{t=1}^T \mathbb{E}\left[z_t^* \sum_{\tau=0}^{t-1} \langle w_{t,\tau}, y_{t-\tau}\rangle_{\mathbb{R}^{k_{t-\tau}}}\right] \\ &= \mathbb{E}\left[\textstyle\sum_{t=1}^T \sum_{\tau=0}^{t-1} \langle z_t^* w_{t,\tau}, y_{t-\tau}\rangle_{\mathbb{R}^{k_{t-\tau}}}\right] \\ &= \mathbb{E}\left[\textstyle\sum_{s=1}^T \sum_{\nu=s}^{T} \langle z_\nu^* w_{\nu,\nu-s}, y_s\rangle_{\mathbb{R}^{k_s}}\right] \\ &= \textstyle\sum_{s=1}^T \mathbb{E}\left[\left\langle \sum_{\nu=s}^T z_\nu^* w_{\nu,\nu-s}, y_s\right\rangle_{\mathbb{R}^{k_s}}\right] \\ &= \textstyle\sum_{s=1}^T \mathbb{E}\left[\left\langle \sum_{\nu=s}^T \mathbb{E}\left[z_\nu^* | \mathcal{F}_s\right] w_{\nu,\nu-s}, y_s\right\rangle_{\mathbb{R}^{k_s}}\right], \end{aligned}$$

hence $W^* z^* = (\sum_{\nu=1}^T \mathbb{E}[z_\nu^*] w_{\nu,\nu-1}, \sum_{\nu=2}^T \mathbb{E}[z_\nu^* | \mathcal{F}_2] w_{\nu,\nu-2}, \dots, z_T^* w_{T,0}) \in E^*$. Utilizing the fact that Y_t are cones for $t = 2, \dots, T$ results in

$$\begin{aligned}
\rho^*(z) &= \sup\left\{\langle z, z^*\rangle_{Z/Z^*} - \sup\left\{\langle y, W^*z^* - c\rangle_{E/E^*} : y \in Y\right\} : z^* \in Z^*\right\}\\
&= \sup\left\{\langle z, z^*\rangle_{Z/Z^*} + \inf\left\{\langle y, c - W^*z^*\rangle_{E/E^*} : y \in Y\right\} : z^* \in Z^*\right\}\\
&= \sup\left\{\begin{array}{l} \langle z, z^*\rangle_{Z/Z^*} + \\ \inf\limits_{y_1\in Y_1} \left\langle y_1, c_1 - \sum_{t=1}^T \mathbb{E}\left[z_t^*\right] w_{t,t-1}\right\rangle \end{array} \left| \begin{array}{l} z^* \in Z^*, \\ c_t - \sum_{\nu=t}^T \mathbb{E}\left[z_\nu^* | \mathcal{F}_t\right] w_{\nu,\nu-t} \in -Y_t^* \\ \text{a.s. } (t = 2, \ldots, T) \end{array}\right.\right\}
\end{aligned}$$

and this is exactly (3.3) with $\lambda = -z^*$. Weak duality holds (cf. [5, section 2.5.1]), i.e., $\rho^*(z) \le \rho(z)$, and dual feasibility ensures $\rho^*(z) > -\infty$, hence

$$-\infty < \rho^*(z) \le \rho(z) < +\infty \quad \forall\, z \in Z.$$

Now, [5, Proposition 2.152] provides that $\rho(z)$ is continuous. In turn, for this case [5, Theorem 2.151] guarantees strong duality, i.e., $\rho^*(z) = \rho(z)$. ◻

As for the one-period case, we define the set of dual multipliers by

$$\Lambda_\rho := \left\{\lambda \in \times_{t=1}^T L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}) \left| \begin{array}{l} c_t + \sum_{\nu=t}^T \mathbb{E}\left[\lambda_\nu | \mathcal{F}_t\right] w_{\nu,\nu-t} \in -Y_t^* \text{ a.s.} \\ (t = 1, \ldots, T) \end{array}\right.\right\}. \tag{3.6}$$

Again, comparing the dual representations (3.1) and (3.4) provides a criterion for a polyhedral functional to be a multiperiod coherent risk measure.

COROLLARY 3.10. *Let ρ be a functional on $\times_{t=1}^T L_p(\Omega, \mathcal{F}_t, \mathbb{P})$ of the form* (3.2) *with Y_1 being a polyhedral cone. Let the conditions of Theorem* 3.9 *be satisfied (complete recourse, dual feasibility) and assume $\Lambda_\rho \subseteq \mathcal{D}_T$. Then ρ is a multiperiod coherent risk measure.*

Proof. Analogously to Corollary 2.5, the assertion here is an immediate consequence of Theorems 3.6 and 3.9 since $\mathcal{P}_\rho := \Lambda_\rho$ does the job. ◻

Example 3.11. A straightforward approach to incorporate risk in terms of the Conditional-Value-at-Risk at all time stages consists in considering a weighted sum

$$\rho_1(z) := \sum_{t=2}^T \gamma_t CVaR_{\alpha_t}(z_t)$$

with some weights $\gamma_t \ge 0$ (e.g., $\gamma_t = \frac{1}{T-1}$) and some confidence levels $\alpha_2, \alpha_3, \ldots, \alpha_T \in (0,1)$. Note that

$$\begin{aligned}
\rho_1(z) &= \sum_{t=2}^T \gamma_t \inf_{r_t\in\mathbb{R}} \left\{r_t + \tfrac{1}{\alpha_t}\mathbb{E}\left[(z_t + r_t)^-\right]\right\}\\
&= \inf_{(r_2,\ldots,r_T)\in\mathbb{R}^{T-1}} \left\{\sum_{t=2}^T \gamma_t \left(r_t + \tfrac{1}{\alpha_t}\mathbb{E}\left[(z_t + r_t)^-\right]\right)\right\}\\
&= \inf\left\{\sum_{t=2}^T \gamma_t \left(y_1^{(t)} + \tfrac{1}{\alpha_t}\mathbb{E}\left[y_t^{(2)}\right]\right) \left| \begin{array}{l} y_1 \in \mathbb{R}^T,\ y_1^{(1)} = z_1, \\ y_t \in L_1(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^2), \\ y_t^{(1)} - y_t^{(2)} = z_t + r_t \text{ a.s.}, \\ y_t \in \mathbb{R}_+ \times \mathbb{R}_+ \text{ a.s. } (t = 2, \ldots, T) \end{array}\right.\right\}
\end{aligned}$$

(set $y_1^{(t)} = r_t$), i.e., ρ_1 is of the form (3.2) with $k_1 = T$, $k_t = 2$ $(t = 2, \ldots, T)$, $c_1 = (0, \gamma_2, \ldots, \gamma_T)$, $c_t = (0, \frac{\gamma_t}{\alpha_t})$ $(t = 2, \ldots, T)$, $w_{1,0} = e_1$, $w_{t,0} = (1, -1)$ $(t = 2, \ldots, T)$, $w_{t,t-1} = -e_t$ $(t = 2, \ldots, T)$, $w_{t,\tau} = 0$ $(\tau = 1, \ldots, t-2,\ t = 3, \ldots, T)$, $Y_1 = \mathbb{R}^T$, $Y_t = \mathbb{R}_+ \times \mathbb{R}_+$ $(t = 2, \ldots, T)$ (with e_t denoting the tth standard basis vector in $\mathbb{R}^T$).

Thus, the risk measure ρ_1 is *multiperiod polyhedral*. Due to Remark 3.5 it is *multiperiod coherent*, too, if $\sum_{t=2}^{T}\gamma_t = 1$. This can also be seen by means of Corollary 3.10. The set of feasible multipliers is given here by

$$\Lambda_{\rho_1} = \left\{ \lambda \in \times_{t=1}^{T} L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}) \;\middle|\; \begin{array}{l} \lambda_1 = 0, \\ 0 \le \lambda_t \le \frac{\gamma_t}{\alpha_t} \text{ a.s. } (t = 2, \ldots, T), \\ \mathbb{E}[\lambda_t] = \gamma_t \end{array} \right\} \tag{3.7}$$

and, of course, $\Lambda_{\rho_1} \subseteq \mathcal{D}_T$. Moreover, the conditions of Theorem 3.9 are satisfied, i.e., complete recourse and dual feasibility hold (take $u = (0, \gamma_2, \ldots, \gamma_T)$).

Next we present more involved examples, which extend the Conditional-Value-at-Risk to the multiperiod situation. The characteristic thing about $CVaR$ is that, in the dual representation, the density functions, i.e., the Lagrangian multipliers are bounded pointwise from above (cf. Example 2.10). This idea will be found somehow in all of the following examples.

Example 3.12. In this example, we define a multiperiod coherent risk measure where not every time step contributes with a fixed weight. When looking at the dual representation (3.3) and at Corollary 3.10, it becomes obvious that each of the dual constraints $c_t + \sum_{\nu=t}^{T}\mathbb{E}[\lambda_\nu|\mathcal{F}_t]w_{\nu,\nu-t} \in -Y_t^*$ has to imply $\lambda_t \ge 0$ for $t = 1, \ldots, T$. A natural candidate for implying $\sum_{\nu=1}^{T}\mathbb{E}[\lambda_\nu] = 1$ is the corresponding constraint for $t = 1$, which reads $c_1 + \sum_{\nu=1}^{T}\mathbb{E}[\lambda_\nu]w_{\nu,\nu-1} \in -Y_1^*$.

Now, setting $k_t = 2$ $(t = 1, \ldots, T)$, $c_1 = (1, 0)$, $c_t = (0, \beta_t)$ with some $\beta_t > 0$ $(t = 2, \ldots, T)$ such that $\sum_{t=2}^{T}\beta_t \ge 1$, $w_{1,0} = (0, 1)$, $w_{t,0} = (1, -1)$ $(t = 1, \ldots, T)$, $w_{t,t-1} = (-1, 0)$ $(t = 2, \ldots, T)$, and $w_{t,\tau} = 0$ $(\tau = 1, \ldots, t-2,\ t = 3, \ldots, T)$, $Y_1 = \mathbb{R} \times \mathbb{R}$, $Y_t = \mathbb{R}_+ \times \mathbb{R}_+$ $(t = 2, \ldots, T)$ leads to

$$\begin{array}{lcll} c_1 + \sum_{\nu=1}^{T}\mathbb{E}[\lambda_\nu]\, w_{\nu,\nu-1} \in -Y_1^* & \iff & \lambda_1 = 0 \quad \text{and} & \sum_{\nu=1}^{T}\mathbb{E}[\lambda_\nu] = 1, \\ c_t + \sum_{\nu=t}^{T}\mathbb{E}[\lambda_\nu|\mathcal{F}_t]\, w_{\nu,\nu-t} \in -Y_t^* & \iff & 0 \le \lambda_t \quad \text{and} & \lambda_t \le \beta_t \ (t = 2, \ldots, T) \end{array}$$

since $Y_1^* = \{0\} \times \{0\}$ and $Y_t^* = \mathbb{R}_- \times \mathbb{R}_-$ $(t = 2, \ldots, T)$. Hence, the dual set Λ_{ρ_2} is of the form

$$\Lambda_{\rho_2} = \left\{ \lambda \in \times_{t=1}^{T} L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}) \;\middle|\; \begin{array}{l} \lambda_1 = 0, \\ 0 \le \lambda_t \le \beta_t \text{ a.s. } (t = 2, \ldots, T), \\ \sum_{t=1}^{T}\mathbb{E}[\lambda_t] = 1 \end{array} \right\}. \tag{3.8}$$

Note that complete recourse and dual feasibility hold. Thus, Corollary 3.10 implies that the functional

$$\rho_2(z) := \inf \left\{ y_1^{(1)} + \sum_{t=2}^{T}\beta_t\mathbb{E}\left[y_t^{(2)}\right] \;\middle|\; \begin{array}{l} y_t \in L_p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^2)\ (t = 1, \ldots, T), \\ y_1 \in \mathbb{R} \times \mathbb{R},\ y_t \in \mathbb{R}_+ \times \mathbb{R}_+ \text{ a.s. } (t = 2, \ldots, T), \\ y_1^{(2)} = z_1, \\ y_t^{(1)} - y_t^{(2)} = z_t + y_1^{(1)} \text{ a.s. } (t = 2, \ldots, T) \end{array} \right\}$$

or simply $\rho_2(z) = \inf_{r\in\mathbb{R}}\{r + \sum_{t=2}^{T}\beta_t\mathbb{E}[(z_t + r)^-]\}$ is a *multiperiod polyhedral* and *coherent risk measure*.

The remaining examples present multiperiod polyhedral coherent risk measures that depend on the filtration $\{\mathcal{F}_t\}_{t=1}^{T}$, i.e., on the information flow over time.

Example 3.13. To incorporate the information structure we adapt the previous example in such a manner that successive time steps are associated. We choose everything as before, only the assignment $w_{t,\tau} = 0$ ($\tau = 1, \dots, t-2$, $t = 3, \dots, T$) is replaced by $w_{t,1} = (0, -1)$ ($t = 3, \dots, T$) and $w_{t,\tau} = 0$ ($\tau = 2, \dots, t-2$, $t = 4, \dots, T$). In addition, we set $c_t = (0, \delta_t)$ with $\delta_t > 0$ for $t = 2, \dots, T$. Hence, the dual set Λ_{ρ_3} is of the form

$$\Lambda_{\rho_3} = \left\{ \lambda \in \times_{t=1}^T L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}) \;\middle|\; \begin{array}{l} \lambda_1 = 0,\ \sum_{t=1}^T \mathbb{E}[\lambda_t] = 1, \\ 0 \le \lambda_t,\ \lambda_t + \mathbb{E}[\lambda_{t+1}|\mathcal{F}_t] \le \delta_t \text{ a.s.} \\ (t = 2, \dots, T-1), \\ 0 \le \lambda_T \le \delta_T \text{ a.s.} \end{array} \right\}. \tag{3.9}$$

Again, the complete recourse condition is satisfied and dual feasibility holds if the parameters δ_t are chosen sufficiently large. Altogether, Corollary 3.10 implies that the functional

$$\rho_3(z) := \inf \left\{ y_1^{(1)} + \sum_{t=2}^T \delta_t \mathbb{E}\left[y_t^{(2)}\right] \;\middle|\; \begin{array}{l} y_t \in L_p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^2)\ (t = 1, \dots, T), \\ y_1 \in \mathbb{R} \times \mathbb{R},\ y_t \in \mathbb{R}_+ \times \mathbb{R}_+ \text{ a.s. } (t = 2, \dots, T), \\ y_1^{(2)} = z_1, \\ y_2^{(1)} - y_2^{(2)} = z_2 + y_1^{(1)} \text{ a.s.}, \\ y_t^{(1)} - y_t^{(2)} = z_t + y_1^{(1)} + y_{t-1}^{(2)} \text{ a.s. } (t = 3, \dots, T) \end{array} \right\}$$

is a *polyhedral multiperiod coherent risk measure.*

Example 3.14. In this approach, the concatenation of the time steps is even stronger than in the previous example. We set $k_t = 2$ ($t = 1, \dots, T$), $c_1 = (\frac{1}{T-1}, 0)$, $c_t = (0, \mu_t)$ ($t = 2, \dots, T$) with some numbers $\frac{1}{T-1} < \mu_2 \le \mu_3 \le \dots \le \mu_T$, $w_{1,0} = (0,1)$, $w_{t,0} = (1, -1)$ ($t = 2, \dots, T$), $w_{t,1} = (-1, 0)$ ($t = 2, \dots, T$), $w_{t,\tau} = 0$ for $\tau > 1$, $Y_1 = \mathbb{R} \times \mathbb{R}$, $Y_t = \mathbb{R} \times \mathbb{R}_+$ ($t = 2, \dots, T-1$), $Y_T = \mathbb{R}_+ \times \mathbb{R}_+$.

The dual constraints $c_t + \sum_{\nu=t}^T \mathbb{E}[\lambda_\nu|\mathcal{F}_t] w_{\nu,\nu-t} \in -Y_t^*$ imply that λ has to be a *martingale* with respect to the filtration $(\mathcal{F}_t)_{t=1}^T$. This implies $\mathbb{E}[\lambda_2] = \dots = \mathbb{E}[\lambda_T]$ and $\lambda_t \ge 0$ since $\lambda_T \ge 0$. Together with (3.6) we obtain

$$\Lambda_{\rho_4} = \left\{ \lambda \in \times_{t=1}^T L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}) \;\middle|\; \begin{array}{l} \lambda_1 = 0, \\ 0 \le \lambda_t \le \mu_t \text{ a.s. } (t = 2, \dots, T), \\ \lambda_t = \mathbb{E}\left[\lambda_{t+1}|\mathcal{F}_t\right]\ (t = 2, \dots, T-1), \\ \mathbb{E}\left[\lambda_2\right] = \dots = \mathbb{E}\left[\lambda_T\right] = \frac{1}{T-1} \end{array} \right\}. \tag{3.10}$$

Complete recourse is satisfied and dual feasibility holds since the vector $u \in \mathbb{R}^T$ with $u_1 = 0$ and $u_t = \frac{1}{T-1}$ for $t = 2, \dots, T$ defines a (constant) element of Λ_{ρ_4}. Hence, Corollary 3.10 applies and the resulting functional

$$\rho_4(z) := \inf \left\{ \tfrac{1}{T-1} y_1^{(1)} + \sum_{t=2}^T \mu_t \mathbb{E}\left[y_t^{(2)}\right] \;\middle|\; \begin{array}{l} y_t \in L_p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^2)\ (t = 1, \dots, T), \\ y_1 \in \mathbb{R} \times \mathbb{R},\ y_T \in \mathbb{R}_+ \times \mathbb{R}_+ \text{ a.s.}, \\ y_t \in \mathbb{R} \times \mathbb{R}_+ \text{ a.s. } (t = 2, \dots, T-1), \\ y_1^{(2)} = z_1, \\ y_t^{(1)} - y_t^{(2)} = z_t + y_{t-1}^{(1)} \text{ a.s. } (t = 2, \dots, T) \end{array} \right\}$$

is a polyhedral *multiperiod coherent risk measure.*

Comparing (3.10) for $\mu_t = \frac{1}{(T-1)\alpha}$ with the dual representation of the Conditional-Value-at-Risk (2.9) it turns out that the multiperiod risk measure ρ_4 defined in this way is a kind of canonical extension of the Conditional-Value-at-Risk in terms of [3, sections 4 and 5] and of [25].[3]

The next example is motivated from the viewpoint of the value of information (cf. [21, 22]).

Example 3.15. In [22], the following multiperiod risk measure was suggested. Given some constants $b_T = 0 \leq d \leq b_{T-1} \leq \cdots \leq b_2 \leq b_1$ and $b_{t-1} \leq q_t$ for $t = 2, \ldots, T$, this risk measure is defined[4] on $\times_{t=1}^T L_p(\Omega, \mathcal{F}_t, \mathbb{P})$ by

$$\rho_5(Z) = -\sup \left\{ \begin{array}{l} \mathbb{E}\left[b_1 A_1 + \sum_{t=2}^{T-1} (b_t A_t - q_t M_t) + d K_T - q_T M_T\right] : \\ A_t \in L_p(\Omega, \mathcal{F}_t, \mathbb{P}) \quad (t = 1, \ldots, T), \\ K_t = [K_{t-1} + Z_t - A_{t-1}]^+ \quad (t = 2, \ldots, T), \\ M_t = [K_{t-1} + Z_t - A_{t-1}]^- \quad (t = 2, \ldots, T) \end{array} \right\}$$

with $K_1 := 0$. However, in [22] $Z = (Z_1, \ldots, Z_T)$ is understood as income process with $Z_1 = 0$, thus this definition does not fit in our framework.

Therefore, we rewrite this definition taking the value processes $z = (z_1, \ldots, z_T)$ with $z_1 = Z_1 = 0$, $z_t = \sum_{\tau=1}^T Z_\tau$, i.e., $Z_t = z_t - z_{t-1}$ for $t > 2$. This reformulation leads to the representation (3.2) with $k_t = 3$ $(t = 1, \ldots, T)$, $Y_1 = \mathbb{R} \times \mathbb{R} \times \{0\}$, $Y_t = \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$ $(t = 2, \ldots, T)$, $y_t = (A_t, M_t, K_t)$, $w_{t,0} = (0, -1, 1)$ $(t = 1, \ldots, T)$, $w_{t,\tau} = (1, -1, 0)$ $(\tau = 1, \ldots, t-2,\ t = 3, \ldots, T)$, $w_{t,t-1} = (1, 0, 0)$ $(t = 2, \ldots, T)$, $c_1 = (-b_1, 0, 0)$, $c_t = (-b_t, q_t, 0)$ $(t = 2, \ldots, T-1)$, $c_T = (0, q_T, -d)$.

To understand this reformulation note that $w_{1,0} = (0, -1, 1)$ implies $M_1 = -z_1 = 0$ and that for $t = 2, \ldots, T$ the recursion $K_t - M_t = K_{t-1} + Z_t - A_{t-1}$ with $K_t \geq 0$ and $M_t \geq 0$ must hold. This recursion can be transformed into a recursion of the type of definition of multiperiod polyhedrality

$$z_t = K_t + \sum_{\tau=1}^{t-1} A_\tau - \sum_{\tau=2}^{t} M_\tau \quad (t = 2, \ldots, T)$$

with $K_1 = 0$. Thus, this risk measure fits into the framework of *multiperiod polyhedral* risk measures.

Furthermore, it is *multiperiod coherent* if $b_1 = 1$. This can be shown by means of Corollary 3.10. Note that

$$c_1 + \sum_{\nu=1}^T \mathbb{E}[\lambda_\nu]\, w_{\nu,\nu-1} \in -Y_1^* \iff \sum_{\nu=2}^T \mathbb{E}[\lambda_\nu] = b_1 \text{ and } \lambda_1 = 0$$

and

$$\begin{array}{l} c_t + \sum_{\nu=t}^T \mathbb{E}[\lambda_\nu | \mathcal{F}_t]\, w_{\nu,\nu-t} \in -Y_t^* \ (t = 2, \ldots, T) \iff \\ d \leq \lambda_T \leq q_T,\ 0 \leq \lambda_t \leq q_t - b_t,\ \sum_{\nu=t+1}^T \mathbb{E}[\lambda_\nu | \mathcal{F}_t] = b_t \ (t = 2, \ldots, T-1), \end{array}$$

[3]The framework in these papers assumes that the multiperiod risk measure is determined only by a set of (scalar) density functions $\mathcal{P}_\rho \subseteq L_1(\Omega, \mathcal{F}, \mathbb{P})$ rather than $\mathcal{P}_\rho \subseteq \times_{t=1}^T L_1(\Omega, \mathcal{F}_t, \mathbb{P})$. Then, the risk $\rho(z)$ is given by expressions like $\sup\{-\frac{1}{T} \sum_{t=1}^T \mathbb{E}[f z_t] : f \in \mathcal{P}_\rho\}$ [25] or $\sup\{-\mathbb{E}[f z_\tau] : f \in \mathcal{P}_\rho,$ τ stopping time$\}$ [3]. Indeed, Λ_{ρ_4} is nothing else but the set of densities for the Conditional-Value-at-Risk (2.9) in terms of [25], i.e., all density functions bounded by $\frac{1}{\alpha}$.

[4]In [22], ρ_5 is called a (negative) utility measure rather than a risk measure. Moreover, the first time stage (i.e., the deterministic stage) is denoted by index 0 there. Here, the formulation is adapted to our framework with index 1 for the deterministic time stage (i.e., $\mathcal{F}_1 = \{\emptyset, \Omega\}$). In addition, the notations c_t and a_t were replaced by the definitions $b_t := c_{t+1}$ and $A_t := a_{t+1}$.

thus

$$\Lambda_{\rho_5} = \left\{ \lambda \in \times_{t=1}^{T} L_{p'}(\Omega, \mathcal{F}_t, P) \;\middle|\; \begin{array}{l} \lambda_1 = 0, \\ 0 \le \lambda_t \le q_t - b_t \text{ a.s. } (t = 2, \ldots, T-1), \\ d \le \lambda_T \le q_T \text{ a.s.}, \\ \mathbb{E}\left[\lambda_t | \mathcal{F}_{t-1}\right] = b_{t-1} - b_t \ (t = 2, \ldots, T) \end{array} \right\}.$$

Further, complete recourse is obviously satisfied and dual feasibility holds since the vector $u \in \mathbb{R}^T$ with $u_1 = 0$, $u_T = b_{T-1}$, and $u_t = b_{t-1} - b_t$ for $t = 2, \ldots, T-1$ defines a (constant) element of Λ_{ρ_5}. Furthermore, $\sum_{t=1}^{T} \mathbb{E}[\lambda_t] = b_1$ for $\lambda \in \Lambda_{\rho_5}$, thus the inclusion $\Lambda_{\rho_5} \subseteq \mathcal{D}_T$ holds indeed if $b_1 = 1$.

An interesting specific case appears for $d = 0$, $b_t = \frac{T-t}{T-1}$, and $q_t = b_t + \frac{1}{(T-1)\alpha_t}$ $(t = 1, \ldots, T)$ with $\alpha_t \in (0, 1)$. Then we obtain

$$\Lambda_{\rho_5} = \left\{ \lambda \in \times_{t=1}^{T} L_{p'}(\Omega, \mathcal{F}_t, P) \;\middle|\; \begin{array}{l} \lambda_1 = 0, \quad 0 \le \lambda_t \le \frac{1}{(T-1)\alpha_t} \text{ a.s.}, \\ \mathbb{E}\left[\lambda_t | \mathcal{F}_{t-1}\right] = \frac{1}{T-1} \end{array} \quad (t = 2, \ldots, T) \right\}$$

and the risk measure ρ_5 on $\times_{t=1}^{T} L_p(\Omega, \mathcal{F}_t, P)$ takes the form

$$\rho_5(z) = \tfrac{1}{T-1} \textstyle\sum_{t=2}^{T} \inf \left\{ \mathbb{E}\left[u_{t-1} + \tfrac{1}{\alpha_t}(z_t + u_{t-1})^- \right] \;\middle|\; u_t \in L_p(\Omega, \mathcal{F}_t, \mathbb{P}) \right\}. \tag{3.11}$$

The tth summand can be interpreted as the expectation of the Conditional-Value-at-Risk of z_t conditioned with respect to the σ-field $\mathcal{F}_{t-1}$. Clearly, (3.11) boils down to the one-period $CVaR$ (2.8) for $T = 2$.

Remark 3.16. Of course, it is interesting to compare these examples. To this end, it is useful to consider the dual representations, i.e., the Lagrange multiplier sets Λ_{ρ_j} $(j = 1, \ldots, 5)$. Hence, regarding formulas (3.8), (3.9), and (3.10), it is obvious that for $\beta_t = \delta_t = \mu_t$ it holds that $\Lambda_{\rho_4} \subseteq \Lambda_{\rho_2} \supseteq \Lambda_{\rho_3}$, thus, since

$$\rho_j(z) = \sup \left\{ -\textstyle\sum_{t=1}^{T} \mathbb{E}\left[\lambda_t z_t\right] : \lambda \in \Lambda_{\rho_j} \right\}, \tag{3.12}$$

the relation $\rho_4 \le \rho_2 \ge \rho_3$ is valid. On the other hand, comparing ρ_3 and ρ_4 for the case $\delta_t = 2\mu_t$ leads to $\Lambda_{\rho_4} \subseteq \Lambda_{\rho_3}$, thus $\rho_4 \le \rho_3$. Hence, ρ_3 is more cautious than ρ_4 in this case. Moreover, if we set $\gamma_t = \frac{1}{T-1}$ and $\beta_t = \mu_t = \frac{1}{(T-1)\alpha_t}$, formula (3.7) shows $\Lambda_{\rho_4} \subseteq \Lambda_{\rho_1} \subseteq \Lambda_{\rho_2}$, hence $\rho_4 \le \rho_1 \le \rho_2$. Thus, ρ_2 is the most cautious or most pessimistic of these risk measures.

More precisely, for a fixed random variable z let $\lambda^j = \lambda^j(z) \in \Lambda_{\rho_j}$ be a maximizer for the dual representations (3.12) of ρ_j, respectively. Then, roughly speaking, λ^j is big where z is small in compliance with the respective restrictions. For $j = 1$ and $j = 4$, the weighting of the time steps is fixed in advance since $\mathbb{E}[\lambda_t^j]$ is fixed. For $j = 2$ the weighting of the time steps is variable, hence the available probability mass of λ^2 is concentrated at time steps at which z is low. Thus, ρ_2 is a kind of worst time step risk measure. This might be desirable or not, depending on the application.

Comparing ρ_1 with ρ_4, one sees that in the first case λ_t^1 is big where z_t is small, independent of the other time steps. In the second case, λ^4 is completely determined by λ_T^4 since $\lambda_t^4 = \mathbb{E}[\lambda_T^4 | \mathcal{F}_t]$ because of the martingale property. This means that the maximization (3.12) takes all time steps into account simultaneously, i.e., the maximization occurs along the paths of the treelike information structure given by the filtration $(\mathcal{F}_t)_{t=1}^{T}$. This latter approach seems to be more efficient in case the risk of paths is of interest. Then, ρ_1 may be more pessimistic than necessary. Furthermore, it does not incorporate the information structure of the problem. On the other hand, the martingale property of ρ_4 seems very restrictive.

Comparing ρ_3 and ρ_4 for the case $\delta_t = 2\mu_t$ leads to $\Lambda_{\rho_4} \subseteq \Lambda_{\rho_3}$, thus $\rho_4 \leq \rho_3$. Hence, ρ_3 is more cautious than ρ_4 in this case. Regarding the dual sets for ρ_5, one obtains $\Lambda_{\rho_5} \subseteq \Lambda_{\rho_1}$ for $\gamma_t = b_{t-1} - b_t$ and $\alpha_t = (b_{t-1} - b_t)/(q_t - b_t)$, and $\Lambda_{\rho_5} \subseteq \Lambda_{\rho_3}$ for $\delta_t = q_t - b_{t+1}$. Hence, $\rho_1 \geq \rho_5 \leq \rho_3$, i.e., ρ_5 is less cautious for this choice of the coefficients.

However, cautiousness is not necessarily a desirable property, because in applications one usually has to pay a price for being cautious. Which risk measure to take depends highly on the intention of the application. It seems that ρ_3 may be a good compromise, since the information structure is taken into account and there is no fixed weighting of the time steps. For initial numerical results we refer to [9].

4. Risk measures in stochastic programs. In this section we study the effect of replacing expectation-based objectives of stochastic programming problems by polyhedral risk measures. In particular, we are interested in consequences for structural and stability properties of the resulting models. We assume that randomness occurs as a (possibly multivariate) stochastic data process $(\xi_t)_{t=1}^T$ and set $\mathcal{F}_t = \sigma(\xi_1, \ldots, \xi_t)$, $t = 1, \ldots, T$. We consider multistage stochastic programs of the form

$$\min \left\{ \mathbb{E}\left[\sum_{t=1}^{T} \langle b_t(\xi_t), x_t\rangle\right] \;\middle|\; \begin{array}{l} x_t \in X_t, \\ H_t(x_t) = 0, \\ B_t(\xi_t)x_t \leq d_t(\xi_t), \\ \sum_{\tau=0}^{t-1} A_{t,\tau}(\xi_t)x_{t-\tau} = h(\xi_t) \end{array} \quad (t = 1, \ldots, T) \right\} \tag{4.1}$$

with closed sets X_t having the property that their convex hull is polyhedral, and with cost coefficients $b_t(\cdot)$, right-hand sides $d_t(\cdot)$ and $h_t(\cdot)$, and matrices $A_{t,\tau}(\cdot)$, $\tau = 0, \ldots, t-1$, and $B_t(\cdot)$ all having suitable dimensions and possibly depending affine linearly on ξ_t for $t = 1, \ldots, T$. The constraints consist of four groups, where the first $x_t \in X_t$ models simple fixed constraints, the second $H_t(z) := z - \mathbb{E}[z|\mathcal{F}_t] = 0$ ensures the nonanticipativity of the decisions x_t, and the third and fourth are the coupling and the dynamic constraints, respectively. By $\mathcal{X}(\xi)$ we denote the set of decisions satisfying all constraints of (4.1).

When replacing the expectation of the stochastic overall costs $\sum_{t=1}^T \langle b_t(\xi_t), x_t\rangle$ by some polyhedral multiperiod risk measure ρ applied to the random vector

$$z(x,\xi) := \Big(-\langle b_1(\xi_1), x_1\rangle, -\langle b_1(\xi_1), x_1\rangle - \langle b_2(\xi_2), x_2\rangle, \ldots, -\textstyle\sum_{\tau=1}^T \langle b_\tau(\xi_\tau), x_\tau\rangle\Big)$$

of negative intermediate costs, we arrive at the following risk averse alternative to problem (4.1):

$$\min \left\{\rho\left(z(x,\xi)\right) \mid x \in \mathcal{X}(\xi)\right\}. \tag{4.2}$$

The polyhedral risk measure ρ is defined by the minimization problem

$$\rho(z) = \inf \left\{ \mathbb{E}\left[\textstyle\sum_{t=1}^T \langle c_t, y_t\rangle\right] \;\middle|\; \begin{array}{l} H_t(y_t) = 0,\ y_t \in Y_t, \\ \sum_{\tau=0}^{t-1} \langle w_{t,\tau}, y_{t-\tau}\rangle = z_t \end{array} (t = 1, \ldots, T) \right\}.$$

This gives rise to the question whether (4.2) is equivalent to the optimization model

(4.3)
$$\min \left\{ \mathbb{E}\left[\sum_{t=1}^{T}\langle c_t, y_t\rangle\right] \;\middle|\; \begin{array}{l} x \in \mathcal{X}(\xi), \\ H_t(y_t) = 0,\; y_t \in Y_t \; (t = 1, \ldots, T), \\ \sum_{\tau=0}^{t-1}\langle w_{t,\tau}, y_{t-\tau}\rangle + \sum_{\tau=1}^{t}\langle b_\tau(\xi_\tau), x_\tau\rangle = 0 \; (t = 1, \ldots, T) \end{array} \right\},$$

where the minimization with respect to the original decision x and the variable y defining ρ is carried out simultaneously. Of course, the answer is positive.

PROPOSITION 4.1. *Minimizing* (4.2) *with respect to* x *is equivalent to minimizing* (4.3) *with respect to all pairs* (x, y) *in the following sense: The optimal values of* (4.2) *and* (4.3) *coincide and a pair* (x^*, y^*) *is a solution of* (4.3) *if and only if* x^* *solves* (4.2) *and* y^* *is a solution of the minimization problem defining* $\rho(z(x^*, \xi))$.

Proof. The minimization with respect to all feasible pairs (x, y) of (4.3) can be carried out by minimizing with respect to y and then by minimizing the latter residual with respect to $x \in \mathcal{X}(\xi)$. Hence, the optimal values coincide and, if the pair (x^*, y^*) solves (4.3), its first component x^* is a solution of (4.2) and y^* is a solution of the problem

$$\min \left\{ \mathbb{E}\left[\sum_{t=1}^{T}\langle c_t, y_t\rangle\right] \;\middle|\; \begin{array}{l} H_t(y_t) = 0,\; y_t \in Y_t, \\ \sum_{\tau=0}^{t-1}\langle w_{t,\tau}, y_{t-\tau}\rangle + \sum_{\tau=1}^{t}\langle b_\tau(\xi_\tau), x^*_\tau\rangle = 0 \end{array} \right\}, \tag{4.4}$$

whose optimal value is just $\rho(z(x^*, \xi))$. Conversely, if x^* is a solution of (4.2) and y^* a solution of (4.4), the pair (x^*, y^*) has to be a solution of (4.3). ☐

Thus, minimizing a stochastic program with a polyhedral risk measure in the objective leads to a "traditional" stochastic program with linear expectation-based objective and with additional variables y and constraints, respectively. Both the variables and the constraints are convenient for stochastic programs since the variables are nicely constrained by polyhedral sets (no integer requirements). Thus, if the original expectation-based stochastic program (4.1) has convenient properties, there is good reason to expect that these properties are maintained when using a polyhedral risk measure for risk aversion.

4.1. Stability of stochastic programs. Stability of solutions and optimal values of stochastic programs with respect to the perturbation of the underlying probability measure is an important issue since in applications the true measure $\mathbb{P}$ is usually unknown and has to be approximated by some other measure $\mathbb{Q}$. Such an approximation may be gained by sampling techniques.

In [30] various stability results involving distances $d(\mathbb{P}, \mathbb{Q})$ of probability measures are developed for different types of (mainly) expectation-based stochastic programs. It is shown there that certain ideal probability metrics (see [23] for an exposition) may be associated with classes of stochastic programs. Here, we briefly show that these stability results remain valid for important classes if the expectation is replaced by a polyhedral risk measure. We restrict ourselves to the two-stage case here since stability properties are best understood for such programs. In the context of distances of probability measures it turns out to be useful to assume that $\Omega = \Xi \subseteq \mathbb{R}^n$ and $\mathcal{F} = \mathcal{B}(\Xi)$.

4.1.1. Linear two-stage programs. In [24, Theorem 3.3] and [30] it is shown that two-stage stochastic programs with fixed recourse of the form

$$\min \left\{ \langle b, x_1\rangle + \mathbb{E}_{\mathbb{P}}\left[\langle p(\cdot), x_2(\cdot)\rangle\right] \;\middle|\; \begin{array}{l} W x_2(\xi) = h(\xi) - T(\xi)x_1, \\ x_1 \in X_1,\; x_2(\xi) \in X_2 \end{array} \right\}, \tag{4.5}$$

with X_1 and Ξ being polyhedral sets, X_2 being a polyhedral cone, and $p(\cdot)$, $h(\cdot)$, $T(\cdot)$ being affine linear functions (of $\xi \in \Xi$), are stable[5] at $\mathbb{P}$ with respect to the probability metric ζ_2 given by

$$\zeta_2(\mathbb{P},\mathbb{Q}) = \sup\left\{ |\mathbb{E}_{\mathbb{P}}[F] - \mathbb{E}_{\mathbb{Q}}[F]| \;\middle|\; \begin{array}{l} F : \Xi \to \mathbb{R}, \\ |F(\xi) - F(\xi')| \le \max\{1, \|\xi\|, \|\xi'\|\} \cdot \|\xi - \xi'\| \\ \forall\, \xi, \xi' \in \Xi \end{array} \right\}$$

if the following four conditions hold:

(A1) $\forall\,(x_1,\xi) \in X_1 \times \Xi \;\exists\, x_2 \in X_2 : W x_2 = h(\xi) - T(\xi)x_1$ (relatively complete recourse).

(A2) $\forall\,\xi \in \Xi \;\exists\, z : W'z - p(\xi) \in X_2^*$ (dual feasibility).

(A3) $\mathbb{E}_{\mathbb{P}}\|\xi\|^2 < \infty$ (finite second moments).

(A4) The first stage solution set $S_{\mathbb{E}} \subseteq X_1$ is nonempty and bounded.

The program (4.5) is equivalent to $\min\{\mathbb{E}_{\mathbb{P}}[z(x_1)] : x_1 \in X_1\}$ using the notations $z(x_1) := \langle b, x_1\rangle + \Phi(p(\cdot), h(\cdot) - T(\cdot)x_1)$ and the second stage value function $\Phi(u,t) := \inf\{\langle u, x_2\rangle : x_2 \in X_2,\, Wx_2 = t\}$ (cf. [34, 29, 30]). Hence, the first stage solution set is given by $S_{\mathbb{E}} := \{x_1 \in X_1 : \mathbb{E}[z(x_1)] = v_{\mathbb{E}}\}$ with $v_{\mathbb{E}} := \inf\{\mathbb{E}[z(x_1)] : x_1 \in X_1\}$ denoting the optimal value.

If we exchange from expectation to a (one-period) polyhedral risk measure $\rho = \rho_{\mathbb{P}}$ according to Definition 2.2, we obtain the problem

$$\min\left\{ \rho\left[-\langle b, x_1\rangle - \langle p(.), x_2(.)\rangle\right] \;\middle|\; \begin{array}{l} Wx_2(\xi) = h(\xi) - T(\xi)x_1, \\ x_1 \in X_1,\; x_2(\xi) \in X_2 \end{array} \right\}, \tag{4.6}$$

which is equivalent to $\min\{\rho[-z(x_1)] : x_1 \in X_1\}$ and, too, equivalent to

$$\min\left\{ \begin{array}{l} \langle c_1, y_1\rangle + \\ \mathbb{E}\left[\langle c_2, y_2(.)\rangle\right] \end{array} \;\middle|\; \begin{array}{l} x_1 \in X_1,\; x_2(\xi) \in X_2, \\ y_1 \in Y_1,\; y_2(\xi) \in Y_2, \\ Wx_2(\xi) = h(\xi) - T(\xi)x_1, \\ \langle p(\xi), x_2(\xi)\rangle + \langle w_2, y_2(\xi)\rangle = -\langle b, x_1\rangle - \langle w_1, y_1\rangle \end{array} \right\}. \tag{4.7}$$

The latter program has almost the same structure as (4.5) with
$\hat{x}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$, $\hat{x}_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$, $\hat{h}(\xi) = \begin{pmatrix} h(\xi) \\ 0 \end{pmatrix}$, $\hat{b} = \begin{pmatrix} 0 \\ c_1 \end{pmatrix}$, $\hat{p} = \begin{pmatrix} 0 \\ c_2 \end{pmatrix}$,
$\hat{W}(\xi) = \begin{pmatrix} W & 0 \\ p(\xi)' & w_2' \end{pmatrix}$, $\hat{T}(\xi) = \begin{pmatrix} T(\xi) & 0 \\ b & w_1' \end{pmatrix}$, $\hat{X}_1 = X_1 \times Y_1$, $\hat{X}_2 = X_2 \times Y_2$,
but now the recourse matrix $\hat{W}$ is random while the cost coefficient $\hat{p}$ is nonrandom.

Moreover, if we also impose complete recourse and dual feasibility for the polyhedral risk measure ρ in the sense of section 2, i.e., (i) $\langle w_2, Y_2\rangle = \mathbb{R}$ and (ii) $D_{\rho,1} \cap D_{\rho,2} \neq \emptyset$, $D_{\rho,2} \subseteq \mathbb{R}_+$, then we can conclude both relatively complete recourse and dual feasibility for the risk aversive alternative (4.7):

(A1) Relatively complete recourse:
Let $(x_1, y_1, \xi) \in X_1 \times Y_1 \times \Xi$; then $\exists\, x_2 \in X_2 : Wx_2 = h(\xi) - T(\xi)x_1$ and $y_2 \in Y_2$ can be chosen such that $\langle w_2, y_2\rangle + \langle p(\xi), x_2\rangle = -\langle b, x_1\rangle - \langle w_1, y_1\rangle$ because $\langle w_2, Y_2\rangle = \mathbb{R}$, thus $\hat{W}(\xi)\hat{x}_2 = \hat{h}(\xi) - \hat{T}(\xi)\hat{x}_1$.

[5] We do not give a precise definition of stability here; see [30] for this. Briefly, stability means that optimal values and (first stage) solution sets behave (quantitatively) continuous at the original measure $\mathbb{P}$ with respect to a distance $d(\mathbb{P},\mathbb{Q})$.

(A2) Dual feasibility:
Let $\xi \in \Xi$. Choose $v \in D_{\rho,2} = \{u \in \mathbb{R} : -(c_2 + uw_2) \in Y_2^*\} \subseteq \mathbb{R}_+$ and z such that $W'z + p(\xi) \in X_2^*$, set $\hat{z} = (vz', -v)'$; then one obtains

$$\hat{W}(\xi)'\hat{z} - \hat{p} = \begin{pmatrix} v(W'z - p(\xi)) \\ -vw_2 - c_2 \end{pmatrix} \in X_2^* \times Y_2^* = \hat{X}_2^*,$$

by making use of the fact that X_2 is a cone.

Since the randomness enters only the last row of $\hat{W}(\xi)$ except for the coefficient in the main diagonal, the stability results from [32] for the random recourse situation with only lower diagonal randomness apply. The model (4.7) with nonrandom costs, however, is again stable with respect to the same metric ζ_2 as for (4.5) if the (first stage) solution set $\bar{S} \subseteq X_1 \times Y_1$ of (4.7) is nonempty and bounded.

PROPOSITION 4.2. *Let ρ be a polyhedral risk measure on $L_1(\Omega, \mathcal{F}, \mathbb{P})$ of the form* (2.1). *Assume that the conditions* (i) $\langle w_2, Y_2 \rangle = \mathbb{R}$ *and* (ii) $D_{\rho,1} \cap D_{\rho,2} \neq \emptyset$, $D_{\rho,2} \subseteq \mathbb{R}_+$, *are satisfied and that the set $S(\rho(0))$ (see* (2.6)*) is nonempty and bounded. Then the set $\bar{S} \subseteq X_1 \times Y_1$ is nonempty and bounded if the solution set $S_\rho := \{x_1 \in X_1 : \rho[-z(x_1)] = \inf_{\hat{x}_1 \in X_1} \rho[-z(\hat{x}_1)]\}$ of* (4.6) *is nonempty and bounded. Hence, the stochastic program* (4.7) *is stable at $\mathbb{P}$ with respect to the metric ζ_2 if the conditions* (A1)–(A3) *are valid and S_ρ is nonempty and bounded.*

Proof. Proposition 4.1 implies that the set $\bar{S}$ is nonempty and bounded if S_ρ is nonempty and bounded and the subset $\bigcup_{x_1 \in S_\rho} S(\rho[-z(x_1)])$ of Y_1 is bounded. Here, $S(\rho(z))$ is defined by (2.6) and is nonempty and bounded due to Proposition 2.9. Clearly, nothing has to be shown if Y_1 is bounded. Now, let Y_1 be unbounded. Suppose $\bigcup_{x_1 \in S_\rho} S(\rho[-z(x_1)])$ is unbounded. Then there exist sequences $(y_{1,n})$ and $(x_{1,n})$ such that $x_{1,n} \in S_\rho$, $y_{1,n} \in S(\rho[-z(x_{1,n})])$ and $\|y_{1,n}\| \geq n$ for $n \in \mathbb{N}$. Because S_ρ is compact, we may assume without loss of generality that $x_{1,n} \to x_{1,0} \in S_\rho$. Since Φ is Lipschitz in t (cf. [43]) we have $z(x_{1,n}) \to z(x_{1,0})$ in $L_1(\Xi)$. Hence, the sequence of probability distributions of $z(x_{1,n})$ converges to the distribution of $z(x_{1,0})$ with respect to the Fortet–Mourier metric ζ_1 (cf. [23, section 5.1]). Now, the set $S(\rho[-z(x_{1,0})])$ is nonempty and bounded. Therefore, the stability result [30, Corollary 25] for two-stage stochastic programs with random right-hand side implies that there must exist an index $n_0 \in \mathbb{N}$ such that for $n \geq n_0$ the sets $S(\rho[-z(x_{1,n})])$ are contained in a fixed bounded neighborhood of $S(\rho[-z(x_{1,0})])$. This contradicts $\|y_{1,n}\| \geq n$, thus $\bigcup_{x_1 \in S_\rho} S(\rho[-z(x_1)])$ must be bounded. □

4.1.2. Linear mixed-integer two-stage programs. In [30, Theorem 35], it is shown that programs of the form

(4.8)
$$\min \left\{ \mathbb{E}_{\mathbb{P}}\left[\langle b, x_1 \rangle + \langle p, x_2(.) \rangle + \langle \bar{p}, \bar{x}_2(.) \rangle\right] \;\middle|\; \begin{array}{l} x_1 \in X_1, \\ x_2(\xi) \in X_2 \cap \mathbb{Z}^m,\ \bar{x}_2(\xi) \in \bar{X}_2, \\ Wx_2(\xi) + \bar{W}\bar{x}_2(\xi) = h(\xi) - T(\xi)x_1 \end{array} \right\}$$

with a closed Euclidean set X_1, a polyhedral set Ξ, and polyhedral cones X_2 and $\bar{X}_2$ are stable with respect to the probability metric ζ_{1,ph_k} with some $k \in \mathbb{N}$ if the following conditions are satisfied:

(B1) $\forall\, (x_1, \xi) \in X_1 \times \Xi\ \exists\, x_2 \in X_2 \cap \mathbb{Z}^m,\ \bar{x}_2 \in \bar{X}_2 : Wx_2 + \bar{W}\bar{x}_2 = h(\xi) - T(\xi)x_1$ (relatively complete recourse).

(B2) $\exists\, z \in \mathbb{R}^r : W'z + p \in X_2^*$ and $\bar{W}'z + \bar{p} \in \bar{X}_2^*$ (dual feasibility).

(B3) $\mathbb{E}_{\mathbb{P}}\|\xi\| < \infty$ (finite first moments).
(B4) W and $\bar{W}$ have rational coefficients only (rational recourse).
(B5) The first stage solution set $S_{\mathbb{E}} \subseteq X_1$ is nonempty and bounded.

The metric ζ_{1,ph_k} is given by

$$\zeta_{1,ph_k}(\mathbb{P},\mathbb{Q}) = \sup\left\{ |\mathbb{E}_{\mathbb{P}}[\chi_B \cdot F] - \mathbb{E}_{\mathbb{Q}}[\chi_B \cdot F]| \;\middle|\; \begin{array}{l} B \in \mathcal{B}_{ph_k}(\Xi),\ F : \Xi \to \mathbb{R} \\ |F(\xi) - F(\xi')| \le \|\xi - \xi'\| \\ \forall\, \xi, \xi' \in \Xi \end{array} \right\},$$

where $\mathcal{B}_{ph_k}(\Xi)$ is the set of polyhedra contained in Ξ with at most k faces and χ denotes the characteristic function, i.e., $\chi_B(\xi) = 1$ if $\xi \in B$ and $= 0$ otherwise.

If we exchange in (4.8) from expectation to a polyhedral risk measure ρ we obtain the problem $\min\{\rho[-z(x_1)] : x_1 \in X_1\}$ with $z(x_1) := \langle b, x_1\rangle + \Phi(h(\cdot) - T(\cdot)x_1)$ and $\Phi(t) := \inf\{\langle p, x_2\rangle + \langle \bar{p}, \bar{x}_2\rangle : x_2 \in X_2 \cap \mathbb{Z}^m,\ \bar{x}_2 \in \bar{X}_2,\ Wx_2 + \bar{W}\bar{x}_2 = t\}$. This problem is equivalent to

$$\min \left\{ \begin{array}{l} \langle c_1, y_1\rangle + \\ \mathbb{E}\left[\langle c_2, y_2(.)\rangle\right] \end{array} \;\middle|\; \begin{array}{l} x_1 \in X_1,\ x_2(\xi) \in X_2 \cap \mathbb{Z}^m,\ \bar{x}_2(\xi) \in \bar{X}_2, \\ y_1 \in Y_1,\ y_2(\xi) \in Y_2, \\ Wx_2(\xi) + \bar{W}\bar{x}_2(\xi) = h(\xi) - T(\xi)x_1, \\ \langle w_2, y_2(\xi)\rangle + \langle p, x_2(\xi)\rangle + \langle \bar{p}, \bar{x}_2(\xi)\rangle = \\ -\langle b, x_1\rangle - \langle w_1, y_1\rangle \end{array} \right\}. \tag{4.9}$$

The latter program has the same structure as (4.8) with
$\hat{x}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$, $\hat{x}_2 = x_2$, $\hat{\bar{x}}_2 = \begin{pmatrix} \bar{x}_2 \\ y_2 \end{pmatrix}$, $\hat{X}_1 = X_1 \times Y_1$, $\hat{X}_2 = X_2$, $\hat{\bar{X}}_2 = \bar{X}_2 \times Y_2$,
$\hat{W} = \begin{pmatrix} W & 0 \\ p' & w_2' \end{pmatrix}$, $\hat{\bar{W}} = \begin{pmatrix} \bar{W} \\ \bar{p}' \end{pmatrix}$, $\hat{T}(\xi) = \begin{pmatrix} T(\xi) & 0 \\ b' & w_1' \end{pmatrix}$, $\hat{h}(\xi) = \begin{pmatrix} h(\xi) \\ 0 \end{pmatrix}$,
$\hat{b} = \begin{pmatrix} 0 \\ c_1 \end{pmatrix}$, $\hat{p} = \begin{pmatrix} 0 \\ c_2 \end{pmatrix}$, $\hat{\bar{p}} = 0$.

As in the previous paragraph, this combined program satisfies relatively complete recourse and dual feasibility if both (4.8) and ρ do so. To have the condition (B4) satisfied, one has to impose additionally that also p, $\bar{p}$, and w_2 have only rational coefficients. Then, however, the same stability (with respect to the metric ζ_{1,ph_k}) as for the original program is guaranteed if the (first stage) solution set $\bar{S} \subseteq X_1 \times Y_1$ of (4.9) is nonempty and bounded. Unfortunately, we cannot conclude as in Proposition 4.2 in the mixed-integer case since Φ is no longer continuous. However, a quantitative stability result is available for the expected loss and the Conditional-Value-at-Risk.

PROPOSITION 4.3. *Let ρ denote the expected loss or the Conditional-Value-at-Risk (see section* 2*). Then the first stage solution set $\bar{S} \subseteq X_1 \times Y_1$ of* (4.9) *is nonempty and bounded if the set $S_\rho := \{x_1 \in X_1 : \rho[-z(x_1)] = \inf_{\hat{x}_1 \in X_1} \rho[-z(\hat{x}_1)]\}$ is nonempty and bounded. Hence, the stochastic program* (4.9) *is stable at $\mathbb{P}$ with respect to ζ_{1,ph_k} if the conditions* (B1)–(B3), (B4)$'$ *W, $\bar{W}$, p, and $\bar{p}$ have rational coefficients only and are satisfied and S_ρ is nonempty and bounded.*

Proof. As in Proposition 4.2, boundedness of $\bar{S}$ is guaranteed if both the set S_ρ of X_1-solutions of (4.9) is nonempty and bounded and the subset $\bigcup_{x_1 \in S_\rho} S(\rho[-z(x_1)])$ of Y_1 is bounded, too. Clearly, the latter set is bounded if Y_1 is bounded which is the case for the expected loss. For the Conditional-Value-at-Risk, we argue as follows. The set of random variables $\{z(x_1) : x_1 \in S_\rho\}$ is bounded in $L_1(\Xi)$ since S_ρ is bounded and the estimate $|\Phi(t) - \Phi(\tilde{t})| \le a\|t - \tilde{t}\| + b$ holds for the second stage function Φ with some positive coefficients a and b (e.g., [30, Lemma 33]). This implies boundedness of the set of their probability distributions $\{D(z(x_1)) : x_1 \in S_\rho\}$ with respect to the Fortet–Mourier metric ζ_1. For real random variables z, $\hat{z}$ and their distributions $D(z)$, $D(\hat{z})$ the metric ζ_1 has the explicit representation (cf. [23, section 5.4])

$$\zeta_1(D(z), D(\hat{z})) = \int_{-\infty}^{\infty} |\mathbb{P}(z \le t) - \mathbb{P}(\hat{z} \le t)| dt.$$

For the $CVaR_\alpha$ we know that for any random variable z the first stage solution set is given by the interval of negative quantiles $S(\rho(z)) = [-\bar{q}_\alpha(z), -q_\alpha(z)]$ (cf. Example 2.10). Fix $\hat{x}_1 \in S_\rho$ and set $\hat{z} := z(\hat{x}_1)$. Let $\Psi_j : \mathbb{R}_+ \to \mathbb{R}_+$ $(j = 1, 2)$ be defined by

$$\Psi_1(r) := \int_{q_\alpha(\hat{z})-r}^{q_\alpha(\hat{z})} (\alpha - \mathbb{P}(\hat{z} \le t))\, dt \qquad \Psi_2(r) := \int_{\bar{q}_\alpha(\hat{z})}^{\bar{q}_\alpha(\hat{z})+r} (\mathbb{P}(\hat{z} \le t) - \alpha)\, dt.$$

Note that the functions Ψ_j $(j = 1, 2)$ are *strictly* increasing. Let z be a random variable. We show that the distances $|q_\alpha(\hat{z}) - q_\alpha(z)|$ and $|\bar{q}_\alpha(\hat{z}) - \bar{q}_\alpha(z)|$ are bounded in terms of $\zeta_1(D(z), D(\hat{z}))$. In case $q_\alpha(z) < q_\alpha(\hat{z})$ it holds that

$$\begin{aligned} \zeta_1(D(z), D(\hat{z})) &= \int_{-\infty}^{\infty} |\mathbb{P}(z \le t) - \mathbb{P}(\hat{z} \le t)| dt &&\ge \int_{q_\alpha(z)}^{q_\alpha(\hat{z})} |\mathbb{P}(z \le t) - \mathbb{P}(\hat{z} \le t)| dt \\ &= \int_{q_\alpha(z)}^{q_\alpha(\hat{z})} (\mathbb{P}(z \le t) - \mathbb{P}(\hat{z} \le t))\, dt &&\ge \int_{q_\alpha(z)}^{q_\alpha(\hat{z})} (\alpha - \mathbb{P}(\hat{z} \le t))\, dt \\ &= \Psi_1(q_\alpha(\hat{z}) - q_\alpha(z)), \end{aligned}$$

hence $|q_\alpha(\hat{z}) - q_\alpha(z)| \le \Psi_1^{-1}(\zeta_1(D(z), D(\hat{z})))$. In case $\bar{q}_\alpha(z) > \bar{q}_\alpha(\hat{z})$ we get

$$\begin{aligned} \zeta_1(D(z), D(\hat{z})) &\ge \int_{\bar{q}_\alpha(\hat{z})}^{\bar{q}_\alpha(z)} |\mathbb{P}(z \le t) - \mathbb{P}(\hat{z} \le t)| dt &&= \int_{\bar{q}_\alpha(\hat{z})}^{\bar{q}_\alpha(z)} (\mathbb{P}(\hat{z} \le t) - \mathbb{P}(z \le t))\, dt \\ &\ge \int_{\bar{q}_\alpha(\hat{z})}^{\bar{q}_\alpha(z)} (\mathbb{P}(\hat{z} \le t) - \alpha)\, dt &&= \Psi_2(\bar{q}_\alpha(z) - \bar{q}_\alpha(\hat{z})), \end{aligned}$$

hence $|\bar{q}_\alpha(z) - \bar{q}_\alpha(\hat{z})| \le \Psi_2^{-1}(\zeta_1(D(z), D(\hat{z})))$. ∎

After this paper was submitted, the authors attention was called to the recent paper [39]. It contains a stability result for the Conditional-Value-at-Risk in mixed-integer two-stage stochastic programs, which is similar to the preceding proposition but proved without relying on Proposition 4.1.

4.2. Lagrangian relaxation and decomposition. We consider again the multistage stochastic program (4.1) and its risk averse alternative (4.2), which, according to Proposition 4.1, is of the form

$$\min \left\{ \mathbb{E}\left[\sum_{t=1}^{T} \langle c_t, y_t \rangle\right] \;\middle|\; \begin{array}{l} x_t \in X_t,\ y_t \in Y_t, \\ H_t(x_t) = 0,\ H_t(y_t) = 0, \\ B_t(\xi_t) x_t \le d_t(\xi_t), \\ \sum_{\tau=0}^{t-1} A_{t,\tau}(\xi_t) x_{t-\tau} = h(\xi_t), \\ \sum_{\tau=0}^{t-1} (\langle w_{t,\tau}, y_{t-\tau} \rangle + \langle b_{\tau+1}(\xi_{\tau+1}), x_{\tau+1} \rangle) = 0 \end{array} \right\}. \tag{4.10}$$

Obviously, (4.10) has a similar structure as (4.1) but additionally with T vector valued random variables and T dynamic (equality) constraints. Thus, decomposition methods that work for (4.1) are likely to work similarly for (4.10), too. We exemplify this here by two important dual decomposition methods.

4.2.1. Scenario decomposition. When solving problems like (4.1) or (4.10) one usually has to approximate $\mathbb{P}$ or, equivalently, ξ by a finite number of scenarios (more precisely: by a finite scenario tree). This can be expressed by $\infty > \#\Omega =: S$ and one can assume without loss of generality $\Omega = \{\xi^1, \dots, \xi^S\}$ and $\mathcal{F} = \wp(\Omega)$. Then the problem is no longer infinite-dimensional and can be solved by standard mixed-integer linear programming techniques, but it is very large scale in most cases. Thus, specialized decomposition techniques are of great interest (cf. [8, 33, 31, 37, 34]).

Scenario decomposition means Lagrange-dualizing the nonanticipativity constraints of (4.10) and solving the dual scenario-wise. Setting $m_t := \dim x_t$ we obtain the dual problem

$$\max \left\{ D(\lambda_1, \lambda_2) : \lambda_{1t} \in L_1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{m_t}),\ \lambda_{2t} \in L_1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{k_t}) \right\},$$

where the dual function $D(\lambda_1, \lambda_2)$ is given by

$$D(\lambda_1, \lambda_2) = \min \left\{ L(\lambda_1, \lambda_2, x, y) \;\middle|\; \begin{array}{l} x_t \in X_t,\ y_t \in Y_t, \\ B_t(\xi_t) x_t \le d_t(\xi_t), \\ \sum_{\tau=0}^{t-1} A_{t,\tau}(\xi_t) x_{t-\tau} = h(\xi_t), \\ \sum_{\tau=0}^{t-1} (\langle w_{t,\tau}, y_{t-\tau} \rangle + \langle b_{\tau+1}(\xi_{\tau+1}), x_{\tau+1} \rangle) = 0 \end{array} \right\}$$

with $L(\lambda_1, \lambda_2, x, y) := \mathbb{E}[\sum_{t=1}^T (\langle c_t, y_t \rangle + \langle \lambda_{1t}, H_t(x_t) \rangle + \langle \lambda_{2t}, H_t(y_t) \rangle)]$ denoting the Lagrangian. Solving this problem is an iterative process: $D(\lambda_1, \lambda_2)$ has to be computed for a fixed pair (λ_1, λ_2) and then (λ_1, λ_2) has to be updated via subgradient-type methods and so on. If the sets X_t are nonconvex, this procedure only leads to lower bounds of the optimal value of (4.1) and suitable globalization techniques based on these lower bounds have to be used in addition.

Because both the restrictions and the Lagrangian are separable with respect to scenarios for a fixed pair (λ_1, λ_2), the calculation of the dual function can be carried out scenario-wise, i.e., $D(\lambda_1, \lambda_2) = \sum_{s=1}^S \mathbb{P}(\{\xi^s\}) D^s(\lambda_1, \lambda_2)$. To derive the separability of the Lagrangian the identities $\mathbb{E}[\langle \lambda_{1t}, H_t(x_t) \rangle] = \mathbb{E}[\langle H_t(\lambda_{1t}), x_t \rangle]$ and $\mathbb{E}[\langle \lambda_{2t}, H_t(y_t) \rangle] = \mathbb{E}[\langle H_t(\lambda_{2t}), y_t \rangle]$ were used.

Hence, instead of one problem with $S \cdot \sum_{t=1}^T (m_t + k_t)$ variables one only has to solve S subproblems each with $\sum_{t=1}^T (m_t + k_t)$ variables to update the multipliers. In comparison with the (dualized form of the) purely expectation-based problem (4.1) one has T additional equality constraints and $\sum_{t=1}^T k_t$ additional variables in each subproblem. Note that the dimensions k_t of y_t are typically small compared to the dimensions m_t of x_t.

4.2.2. Geographical decomposition. In many practical applications the stochastic program (4.1) shows the following kind of block separability $x_i = (x_{i1}, \dots, x_{iT})$, $i = 1, \dots, I$, of components of x:

$$\min \left\{ \mathbb{E}\left[\sum_{i=1}^I \sum_{t=1}^T \langle b_{it}(\xi_t), x_{it} \rangle \right] \;\middle|\; \begin{array}{l} x_{it} \in X_{it}, \\ H_t(x_{it}) = 0, \\ \sum_{i=1}^I B_{it}(\xi_t) x_{it} \le d_t(\xi_t), \\ \sum_{\tau=0}^{t-1} A_{it,\tau}(\xi_t) x_{i,t-\tau} = h_{it}(\xi_t) \end{array} \right\}. \tag{4.11}$$

Hence, the I blocks of x are only coupled by the sum in the third constraint in (4.11). For such programs, *Lagrange relaxation of coupling constraints*, also known as *geographical* or *component decomposition*, may lead to efficient algorithms for computing lower bounds (cf. [8, 31]).

By exchanging from $\mathbb{E}$ to a multiperiod polyhedral risk measure this property is maintained, but an additional block consisting of the y_t variables and T *additional (dynamic) coupling constraints* appear,

$$\min\left\{\mathbb{E}\left[\sum_{t=1}^{T}\langle c_t, y_t\rangle\right] \;\middle|\; \begin{array}{l} x_{it}\in X_{it},\ y_t\in Y_t,\\ H_t(x_{it})=0,\ H_t(y_t)=0,\\ \sum_{i=1}^{I} B_{it}(\xi_t)x_{it}\le d_t(\xi_t),\\ \sum_{\tau=0}^{t-1} A_{it,\tau}(\xi_t)x_{i,t-\tau}=h_{it}(\xi_t),\\ \sum_{\tau=0}^{t-1}(\langle w_{t,\tau}, y_{t-\tau}\rangle+\sum_{i=1}^{I}\langle b_{i,\tau+1}(\xi_t), x_{i,\tau+1}\rangle)=0 \end{array}\right\}. \tag{4.12}$$

Here, Lagrange relaxation of coupling constraints means to assign $\mathcal{F}_t$-measurable Lagrange multipliers λ_{1t} and λ_{2t} to the third *and* fifth constraint in (4.12), respectively, and to arrive at the dual problem

$$\max\left\{D(\lambda_1,\lambda_2) : \lambda_{1t}\in L_{p'}(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{R}_+^{n_t}),\ \lambda_{2t}\in L_{p'}(\Omega,\mathcal{F}_t,\mathbb{P})\right\}.$$

The dual function $D(\lambda_1,\lambda_2)$ is given by

$$D(\lambda_1,\lambda_2)=\min\left\{L(\lambda_1,\lambda_2,x_1,\dots,x_I,y) \;\middle|\; \begin{array}{l} x_{it}\in X_{it},\ y_t\in Y_t,\\ H_t(x_{it})=0,\ H_t(y_t)=0,\\ \sum_{\tau=0}^{t-1} A_{it,\tau}(\xi_t)x_{i,t-\tau}=h_{it}(\xi_t) \end{array}\right\}$$

and the Lagrangian $L(\lambda_1,\lambda_2,x_1,\dots,x_I,y)$ is defined by

$$\begin{aligned} &L(\lambda_1,\lambda_2,x_1,\dots,x_I,y)\\ =&\mathbb{E}\Big[\sum_{t=1}^{T}\Big(\langle c_t,y_t\rangle+\Big\langle\lambda_{1t},\sum_{i=1}^{I}B_{it}(\xi_t)x_{it}-d_t(\xi_t)\Big\rangle\\ &\qquad+\lambda_{2t}\sum_{\tau=0}^{t-1}\Big(\langle w_{t,\tau},y_{t-\tau}\rangle+\sum_{i=1}^{I}\langle b_{i,\tau+1}(\xi_{\tau+1}),x_{i,\tau+1}\rangle\Big)\Big)\Big]. \end{aligned}$$

By rearranging with respect to blocks in the objective, the dual function D decomposes into $I+1$ minimization subproblems and is then of the form

$$D(\lambda_1,\lambda_2)=\sum_{i=1}^{I}D_i(\lambda_1,\lambda_2)+D_R(\lambda_2)-\mathbb{E}\left[\sum_{t=1}^{T}\langle\lambda_{1t},d_t(\xi_t)\rangle\right].$$

The functions D_i correspond to I geographical subproblems

$$\begin{aligned} &D_i(\lambda_1,\lambda_2)\\ =&\min\left\{\mathbb{E}\left[\sum_{t=1}^{T}\Big\langle B_{it}(\xi_t)'\lambda_{1t}+b_{it}(\xi_t)\sum_{\tau=t}^{T}\lambda_{2\tau},x_{it}\Big\rangle\right] \;\middle|\; \begin{array}{l} x_{it}\in X_{it},\\ H_t(x_{it})=0,\\ \sum_{\tau=0}^{t-1}A_{it,\tau}(\xi_t)x_{i,t-\tau}=h_{it}(\xi_t) \end{array}\right\} \end{aligned}$$

and D_R corresponds to the risk subproblem

$$D_R(\lambda_2)=\min\left\{\mathbb{E}\left[\sum_{t=1}^{T}\Big\langle c_t+\sum_{\tau=t}^{T}\lambda_{2\tau}w_{\tau,\tau-t},y_t\Big\rangle\right] \;\middle|\; \begin{array}{l} y_t\in Y_t,\\ H_t(y_t)=0 \end{array}\right\}.$$

Compared to the (dualized form of the) purely expectation-based problem (4.11), the subproblems for the x_i-blocks have the same structure, therefore the same solution methods can be applied. The only change consists in the additional factors $\sum_{\tau=t}^{T} \lambda_{2\tau}$ of $b_{it}(\xi_t)$ in the objective. If Y_1 is a cone, the subproblem for the additional y-block represents a cone constrained linear stochastic program and can be solved explicitly, namely, it holds

$$D_R(\lambda_2) = \begin{cases} 0 & \text{if } -\left(c_t + \sum_{\tau=t}^{T} \mathbb{E}[\lambda_{2\tau}|\mathcal{F}_t] w_{\tau,\tau-t}\right) \in Y_t^* \ (t = 1, \dots, T), \\ -\infty & \text{otherwise.} \end{cases}$$

Hence, the dual problem reads

$$\max \left\{ \sum_{i=1}^{I} D_i(\lambda_1, \lambda_2) - \mathbb{E}\left[\sum_{t=1}^{T} \langle \lambda_{1t}, d_t(\xi_t) \rangle \right] \;\middle|\; \begin{array}{l} \lambda_{1t} \in L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}_+^{n_t}), \\ \lambda_{2t} \in L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}), \\ c_t + \sum_{\tau=t}^{T} \mathbb{E}[\lambda_{2\tau}|\mathcal{F}_t] w_{\tau,\tau-t} \in -Y_t^* \\ (t = 1, \dots, T) \end{array} \right\}$$

and the whole Lagrangian decomposition strategy has the same favorable features for the risk averse model (4.12) as for the expectation-based one (4.11). For example, the known Lagrangian relaxation based algorithms for electricity portfolio optimization (e.g., [4, 12, 16]) apply to risk aversive models after some modifications.

5. Conclusions. We have introduced the class of polyhedral risk measures. Polyhedral risk measures are defined as optimal values of certain linear stochastic programs with recourse where the arguments appear on the right-hand sides of the dynamic constraints. By means of convex duality, criteria for coherence and second order stochastic dominance consistency have been deduced. For the one-period case it has been shown that well-known risk measures are contained in this class: Conditional-Value-at-Risk / quantile dispersion, and expected loss. For the multiperiod case, five polyhedral (coherent) risk measures were suggested.

Stochastic programs with a polyhedral risk measure as objective (or, alternatively, with an objective consisting of a linear combination of an expectation and a polyhedral risk measure) can be easily transformed into expectation-based stochastic programs. This observation has been used to demonstrate that important dual decomposition techniques known for certain expectation-based stochastic programs can be applied to stochastic programs with polyhedral risk measures after some modicifactions. The same is true for stability properties of stochastic programs.

Hence, for large scale problems possibly including integer variables polyhedral risk measures are a reasonable and flexible means to control risk while keeping the problems tractable.

Acknowledgments. The authors wish to thank the referees as well as the associate editor for valuable remarks and suggestions. Further thanks are due to Stephan Tiedemann (University of Duisburg-Essen) for helpful discussions.

REFERENCES

[1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, *Coherent measures of risk*, Math. Finance, 9 (1999), pp. 203–228.
[2] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, and H. Ku, *Coherent Multiperiod Risk Measurement*, Working Paper, 2002, downloadable from www.math.ethz.ch/~delbaen.
[3] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, and H. Ku, *Coherent Multiperiod Risk Adjusted Values and Bellman's Principle*, Working Paper, 2004, downloadable from www.math.ethz.ch/~delbaen.

[4] L. Bacaud, C. Lemaréchal, A. Renaud, and C. Sagastizábal, *Bundle methods in stochastic optimal power management: A disaggregated approach using preconditioners*, Comput. Optim. Appl., 20 (2001), pp. 227–244.
[5] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer Ser. Oper. Res., Springer-Verlag, New York, 2000.
[6] F. Delbaen, *Coherent risk measures on general probability spaces*, in Advances in Finance and Stochastics, K. Sandmann and P. Schönbucher, eds., Springer, Berlin, 2002, pp. 1–37.
[7] M. A. H. Dempster, ed., *Risk Management: Value at Risk and Beyond*, Cambridge University Press, Cambridge, UK, 2002.
[8] D. Dentcheva and W. Römisch, *Duality gaps in nonconvex stochastic optimization*, Math. Program., Ser. A, 101 (2004), pp. 515–535.
[9] A. Eichhorn, W. Römisch, and I. Wegner, *Polyhedral risk measures in electricity portfolio optimization*, PAMM Proc. Appl. Math. Mech., 4 (2004), pp. 7–10.
[10] H. Föllmer and A. Schied, *Convex risk measures and trading constraints*, Finance Stoch., 6 (2002), pp. 429–447.
[11] H. Föllmer and A. Schied, *Stochastic Finance: An Introduction in Discrete Time*, Walter de Gruyter, Berlin, 2002.
[12] N. Gröwe-Kuska, K. C. Kiwiel, M. P. Nowak, W. Römisch, and I. Wegner, *Power management in a hydro-thermal system under uncertainty by Lagrangian relaxation*, in Decision Making under Uncertainty: Energy and Power, C. Greengard and A. Ruszczyński, eds., IMA Vol. Math. Appl. 128, Springer, New York, 2002, pp. 39–70.
[13] S. Kusuoka, *On law invariant coherent risk measures*, in Advances in Mathematical Economics, Vol. 3, S. Kusuoka et al., eds., Springer, Tokyo, 2001, pp. 83–95.
[14] H. Markowitz, *Portfolio selection*, J. Finance, 7 (1952), pp. 77–91.
[15] A. Müller and D. Stoyan, *Comparison Methods for Stochastic Models and Risks*, Wiley, Chichester, UK, 2002.
[16] M. P. Nowak and W. Römisch, *Stochastic Lagrangian relaxation applied to power scheduling in a hydro-thermal system under uncertainty*, Ann. Oper. Res., 100 (2000), pp. 251–272.
[17] W. Ogryczak and A. Ruszczyński, *From stochastic dominance to mean-risk models: Semideviations as risk measures*, European J. Oper. Res., 116 (1999), pp. 33–50.
[18] W. Ogryczak and A. Ruszczyński, *On consistency of stochastic dominance and mean-semideviation models*, Math. Program., Ser. B, 89 (2001), pp. 217–232.
[19] W. Ogryczak and A. Ruszczyński, *Dual stochastic dominance and related mean-risk models*, SIAM J. Optim., 13 (2002), pp. 60–78.
[20] G. C. Pflug, *Some remarks on the value-at-risk and the conditional value-at-risk*, in Probabilistic Constrained Optimization, S. P. Uryasev, ed., Kluwer, Dordrecht, The Netherlands, 2000, pp. 272–281.
[21] G. C. Pflug, *The value of perfect information as a risk measure*, in Dynamic Stochastic Optimization, K. Marti, Y. Ermoliev, and G. C. Pflug, eds., Lecture Notes in Econom. and Math. Systems 532, Springer, Berlin, 2004, pp. 275–291.
[22] G. C. Pflug and A. Ruszczyński, *A risk measure for income processes*, in Risk Measures for the 21st Century, G. Szegö, ed., Wiley, Chichester, UK, 2004, pp. 249–269.
[23] S. T. Rachev, *Probability Metrics and the Stability of Stochastic Models*, Wiley, Chichester, UK, 1991.
[24] S. T. Rachev and W. Römisch, *Quantitative stability in stochastic programming: The method of probability metrics*, Math. Oper. Res., 27 (2002), pp. 792–818.
[25] F. Riedel, *Dynamic coherent risk measures*, Stochastic Process. Appl., 112 (2004), pp. 185–200.
[26] R. T. Rockafellar, *Conjugate Duality and Optimization*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 16, SIAM, Philadelphia, 1974.
[27] R. T. Rockafellar and S. Uryasev, *Conditional value-at-risk for general loss distributions*, Journal of Banking & Finance, 26 (2002), pp. 1443–1471.
[28] R. T. Rockafellar, S. Uryasev, and M. Zarabankin, *Deviation Measures in Risk Analysis and Optimization*, Research Report 2002-7, Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL, 2002.
[29] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
[30] W. Römisch, *Stability of stochastic programming problems*, in Stochastic Programming, A. Ruszczyński and A. Shapiro, eds., Handbooks Oper. Res. Management Sci. 10, Elsevier, Amsterdam, 2003, pp. 483–554.
[31] W. Römisch and R. Schultz, *Multistage stochastic integer programs: An introduction*, in Online Optimization of Large Scale Systems, M. Grötschel, S. O. Krumke, and J. Rambau, eds., Springer, Berlin, 2001, pp. 581–622.
[32] W. Römisch and R. J.-B. Wets, *Stability of ε-Approximate Solutions to Convex Stochastic*

Programs, in preparation.

[33] A. Ruszczyński, *Decomposition methods in stochastic programming*, Math. Program., Ser. B, 79 (1997), pp. 333–353.

[34] A. Ruszczyński and A. Shapiro, eds., *Stochastic Programming*, Handbooks Oper. Res. Management Sci. 10, Elsevier, Amsterdam, 2003.

[35] A. Ruszczyński and A. Shapiro, *Optimization of convex risk functions*, Stochastic Programming E-Print Series, 6 (2004), paper 8, downloadable from www.speps.info.

[36] G. Scandolo, *Risk Measures for Processes and Capital Requirement*, Working Paper, Universitá degli Studi di Milano, Milan, Italy, 2003.

[37] R. Schultz, *Stochastic programming with integer variables*, Math. Program., Ser. B, 97 (2003), pp. 285–309.

[38] R. Schultz and S. Tiedemann, *Risk aversion via excess probabilities in stochastic programs with mixed-integer recourse*, SIAM J. Optim., 14 (2003), pp. 115–138.

[39] R. Schultz and S. Tiedemann, *Conditional value-at-risk in stochastic programs with mixed-integer recourse*, Stochastic Programming E-Print Series, 6 (2004), paper 20, downloadable from www.speps.info.

[40] A. Shapiro and S. Ahmed, *On a class of minimax stochastic programs*, SIAM J. Optim., 14 (2004), pp. 1237–1249.

[41] G. P. Szegö, ed., *Statistical and Computational Problems in Risk Management: VaR and Beyond VaR*, special issue of Journal of Banking & Finance, 26 (7) (2002).

[42] S. Weber, *Distribution-Invariant Dynamic Risk Measures*, Working Paper, Humboldt-University, Berlin, Germany, 2003.

[43] R. J.-B. Wets, *Stochastic programs with fixed recourse: The equivalent deterministic program*, SIAM Rev., 16 (1974), pp. 309–339.