

## Stability of multistage stochastic programs incorporating polyhedral risk measures

Andreas Eichhorn\* and Werner Römisch\*

(Received 00 Month 200x; In final form 00 Month 200x)

We analyse stability aspects of linear multistage stochastic programs with polyhedral risk measures in the objective. In particular, we consider sensitivity of the optimal value with respect to perturbations of the underlying stochastic input process. An existing stability result for multistage stochastic programs with expectation objective is carried forward to the case of polyhedral risk-averse objectives. Beside  $L_r$ -distances these results also involve filtration distances of the perturbations of the stochastic process. We discuss additional requirements for the polyhedral risk measures such that the problem dependent filtration distances can be bounded by problem independent ones. Stability and such bounds are the basis for scenario tree approximation techniques used in practical problem solving.

### 1 Multistage stochastic programming and stability

Multistage stochastic programs are a model the situation of a decision maker faced with a finite number of timesteps  $t = 1, \dots, T$  at each of which he/she observes some random outcomes  $\xi_t$  and has to make an (optimal) decision  $x_t$  based on the exact knowledge of the past ( $\xi_1, \dots, \xi_t$  and  $x_1, \dots, x_{t-1}$ ) and on statistical information about the future ( $\xi_{t+1}, \dots, \xi_T$ ); cf., e.g., [21]. The random outcomes may enter both, the objective as well as the constraints for the decisions. The presence of statistical information is expressed by assuming  $\xi = (\xi_1, \dots, \xi_T)$  to be a (multivariate) stochastic process on some fixed probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Note that it is assumed that the stochastic process is a pure input parameter and, hence, does not depend on the decisions.

In the following, we assume  $\xi_1, \dots, \xi_T \in L_r(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$  with some numbers  $r \in [1, \infty]$  and  $d \in \mathbb{N}$ . We set  $\xi^t := (\xi_1, \dots, \xi_t)$  and we introduce the  $\sigma$ -fields  $\mathcal{F}_t := \sigma(\xi^t)$  for  $t = 1, \dots, T$ . We assume that that  $\xi_1$  is deterministic and that  $\sigma(\xi) = \mathcal{F}$ . Thus we have the following *filtration*:  $\mathcal{F}_1 = \{\emptyset, \Omega\} \subseteq \mathcal{F}_2 \subseteq \dots \subseteq$

---

\*Humboldt University Berlin, Department of Mathematics, 10099 Berlin, Germany. Email: romisch@math.hu-berlin.de, eichhorn@math.hu-berlin.de, URL [www.math.hu-berlin.de/~romisch](http://www.math.hu-berlin.de/~romisch)

$\mathcal{F}_T = \mathcal{F}$ . We will consider *linear* multistage stochastic programs of the form

$$\min \left\{ \mathbb{E} \left[ \sum_{t=1}^T \langle b_t(\xi_t), x_t \rangle \right] \left| \begin{array}{l} x_t \in L_{r'}(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^{m_t}) \quad (t = 1, \dots, T), \\ x_t \in X_t \text{ a.s.} \quad (t = 1, \dots, T), \\ \sum_{\tau=0}^{t-1} A_{t,\tau}(\xi_t) x_{t-\tau} = h_t(\xi_t) \text{ a.s.} \quad (t = 2, \dots, T) \end{array} \right. \right\} \quad (1)$$

with some numbers  $m_t, n_t \in \mathbb{N}$  and  $r' \in [1, \infty]$ , polyhedral sets  $X_t \subseteq \mathbb{R}^{m_t}$ , matrices  $A_{t,\tau} \in \mathbb{R}^{n_t \times m_{t-\tau}}$ , and vectors  $h_t \in \mathbb{R}^{n_t}$  and  $b_t \in \mathbb{R}^{m_t}$ . We assume that  $A_{t,\tau}$ ,  $h_t$ , and  $b_t$  depend affinely linearly on  $\xi_t$  ( $t = 1, \dots, T$ ). The matrices  $A_{t,0}$  are called the recourse matrices ( $t = 2, \dots, T$ ) and for  $\tau > 0$  the matrices  $A_{t,\tau}$  are called technology matrices. The vectors  $b_t$  can be interpreted as cost factors. Note that optimality of the stochastic costs  $\langle b_t(\xi_t), x_t \rangle$  is evaluated in terms of expectation.

For various reasons it is of interest to analyse stability properties of stochastic programs with respect to perturbations of the underlying stochastic input process  $\xi = (\xi_1, \dots, \xi_T)$ . In particular, quantitative stability results have a significant impact on methods for approximating  $\xi$  suitably by finite scenario trees. For the special case  $T = 2$ , a lot is known for different types of stochastic programs; see [16, 20] for a broad exposition and [9] for applications to scenario approximation. This case is much easier to handle since the information structure is fix:  $\mathcal{F}_1 = \{\emptyset, \Omega\}$ ,  $\mathcal{F}_2 = \mathcal{F}$ . For  $T > 2$ , the situation is much more challenging; only few approaches can be found in literature. In [10] a stability result for the optimal values of (1) was stated introducing a so-called filtration distance. Scenario tree approximation methods based on this stability result have been presented in [8].

In many applications it is of interest to consider risk measures alternatively to the expectation functional  $\mathbb{E}$  in the objective of (1). Typically, risk measures are inherently non-linear. Since the existing stability results rest to some extent on the linearity of the objective, it seems very difficult to carry them forward to problems with objectives incorporating arbitrary risk measures. However, in [3] the class of polyhedral risk measures was introduced containing ordinary risk measures such as CVaR / AVaR as well as multiperiod risk measures. As it will be demonstrated in section 2, these risk measures, though non-linear, behave particularly suitable in the objective of (1) since polyhedral risk measures themselves are defined as optimal values of certain stochastic programs. In section 3 of the present paper, we will proof stability theorems similar to that from [10] which apply to the situation obtained by incorporating a polyhedral risk measure into (1). These stability results consist of local Lipschitz type estimates involving  $L_r(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^s)$  norm distances (where  $s = Td$ ) as well as filtration distances. The filtration distances depend on the solution behaviour of the particular problems, but in [8] it has already been discussed how to bound or estimate them by problem independent metrics in the context of

scenario tree approximation. However, these estimates are valid only if the sets of  $\varepsilon$ -optimal solutions are uniformly bounded. Hence, in section 4, conditions for this boundedness will be analysed and it will be seen that all risk measure instances from [3] don't cause any problems with respect to these conditions. Hence, we conclude in section 5 that polyhedral risk measures are a meaningful tool for risk aversion in multistage stochastic programming.

## 2 Risk measures in the objective

### 2.1 Measures of risk

Basically, risk measures  $\rho$  are (extended) real-valued functionals on some space  $\mathcal{Z}$  of random variables on the measurable space  $(\Omega, \mathcal{F})$  (e.g.,  $\mathcal{Z} = L_p(\Omega, \mathcal{F}, \mathbb{P})$  with  $p \geq 1$  or even  $p = 0$ ) or random processes (e.g.,  $\mathcal{Z} = \times_{j=1}^J L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P})$  with timesteps  $1 \leq t_1 < t_2 < \dots < t_J \leq T$ ), i.e.,  $\rho : \mathcal{Z} \rightarrow \overline{\mathbb{R}}, z \mapsto \rho(z)$ . Typically,  $\rho$  is essentially nonlinear. The number  $\rho(z)$  is intended to represent the chance of ending up with undesired realizations  $z(\omega)$  of  $z$  or to represent the degree of uncertainty (spread) associated with  $z$ . In any case, if there is a choice among different  $z \in \mathcal{Z}$ , one is interested to find a  $z$  such that the value  $\rho(z)$  is rather low, i.e., one may want to *minimise*  $\rho(z)$  over a subset of  $\mathcal{Z}$ .

Consider the one-period case, i.e.,  $\mathcal{Z} = L_p(\Omega, \mathcal{F}, \mathbb{P})$ . We assume in the following, that for  $z \in \mathcal{Z}$  higher outcomes  $z(\omega)$  are preferred to lower ones, e.g.,  $z = -\sum_{t=1}^T \langle b_t(\xi_t), x_t \rangle$ . Classical functionals in this context are, e.g., the *variance* [11] ( $p \geq 2$  required) or the *Value-at-Risk* at level  $\alpha \in (0, 1)$  [6, Chapter 4.4] given by  $\text{VaR}_\alpha(z) = -\bar{q}_\alpha(z)$  with  $\bar{q}_\alpha(z) = \inf\{a \in \mathbb{R} : \mathbb{P}(z \leq a) > \alpha\}$  denoting the upper alpha quantile. Note that both of these functionals are known to have certain drawbacks in particular when being used for optimization. Other well known risk functionals are semideviations [15], expected utility, shortfall risk [6, Chapter 4.6], etc. It may also be desirable to minimise a mixture  $\gamma \cdot \rho(z) - (1 - \gamma) \cdot \mathbb{E}[z]$  of a risk measure and the expectation with some number  $\gamma \in [0, 1]$  (mean-risk models, cf. [11, 15, 23]). Important work on axiomatic characterisations of risk measures has been reported in [1] and [6, Chapter 4.1], but also [15] contains considerations in that direction.

For the case that discrete time random processes  $z = (z_{t_1}, \dots, z_{t_J})$  are to be evaluated, multiperiod risk measures have to be used [2, 7, 12–14, 17]; see also [22, sections 11–13]. In this case, axiomatic characterisation turned out to be more controversial and, too, fewer instances are suggested in literature. If a multiperiod risk measure shall be minimised within a multistage stochastic programming framework such as (1), the risk measure does not necessarily need to take all timesteps  $t = 1, \dots, T$  into account but may be restricted to a subset  $t_1, \dots, t_J$  of timesteps. Hence, for the multiperiod case we will consider

$$\mathcal{Z} = \times_{j=1}^J L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P}).$$

## 2.2 Polyhedral risk measures

For the purpose of being minimised in a (multistage) stochastic program, polyhedral risk measures have been introduced in [3] and applied to electricity models in [4, 5]. Risk measures from this class are defined as optimal values of certain simple-structured stochastic minimisation problems. Consider the multiperiod case with some (fixed) timesteps  $1 = t_0 < t_1 < \dots < t_J = T$  and  $\mathcal{Z} = \times_{j=1}^J L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P})$ . Then a functional  $\rho$  is called a (multiperiod) polyhedral risk measure if it has the following form:

$$\rho(z) = \inf \left\{ \mathbb{E} \left[ \sum_{j=0}^J \langle c_j, y_j \rangle \right] \left| \begin{array}{l} y_j \in L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P}; \mathbb{R}^{k_j}) \quad (j = 0, \dots, J), \\ y_j \in Y_j \text{ a.s.} \quad (j = 0, \dots, J), \\ \sum_{\tau=0}^j \langle w_{j,\tau}, y_{j-\tau} \rangle = z_{t_j} \text{ a.s.} \quad (j = 1, \dots, J) \end{array} \right. \right\} \quad (2)$$

with some numbers  $k_j \in \mathbb{N}$  and vectors  $c_j \in \mathbb{R}^{k_j}$  ( $j = 0, \dots, J$ ),  $w_{j,\tau} \in \mathbb{R}^{k_{j-\tau}}$ , ( $j = 1, \dots, J$ ,  $\tau = 0, \dots, j$ ), a polyhedral set  $Y_0 \subseteq \mathbb{R}^{k_0}$ , and polyhedral cones  $Y_j \subseteq \mathbb{R}^{k_j}$  ( $j = 1, \dots, J$ ). Typically, when using this type of functional in the objective of a multistage stochastic program (cf. (1)), one has  $z_{t_j} = -\sum_{t=1}^{t_j} \langle b_t(\xi_t), x_t \rangle$  for  $z = (z_{t_1}, \dots, z_{t_J}) \in \mathcal{Z}$ . Note that the case that it is intended to minimise a combination like  $\gamma \cdot \rho(z_{t_1}, \dots, z_{t_J}) - (1 - \gamma) \cdot \mathbb{E}[z_T]$  is fully included in this framework, since such a mixture can be expressed by modifying the vectors  $c_j$  in (2) suitably [3, Remark 2.3]. For  $\mathcal{Z} = L_p(\Omega, \mathcal{F}, \mathbb{P})$ , i.e., for the one-period case, the definition is accordingly ( $J = 1$  and  $\mathcal{F}_{t_1} = \mathcal{F}$ ).

One reason why polyhedral risk measures are particularly suitable for being minimised is as follows. For a stochastic program of the form (1) with a polyhedral risk measure in the objective

$$\min \left\{ \rho(z_{t_1}, \dots, z_{t_J}) \left| \begin{array}{l} x_t \in L_{r'}(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^{m_t}), \quad x_t \in X_t \text{ a.s.} \quad (t = 1, \dots, T), \\ \sum_{\tau=0}^{t-1} A_{t,\tau}(\xi_t) x_{t-\tau} = h_t(\xi_t) \text{ a.s.} \quad (t = 2, \dots, T) \\ z_t = z_t(\xi^t, x^t) := -\sum_{\tau=1}^t \langle b_\tau(\xi_\tau), x_\tau \rangle \quad (t = 1, \dots, T) \end{array} \right. \right\} \quad (3)$$

there is an obvious equivalence to

$$\min \left\{ \mathbb{E} \left[ \sum_{j=0}^J \langle c_j, y_j \rangle \right] \left| \begin{array}{l} x_t \in L_{r'}(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^{m_t}), \quad x_t \in X_t \text{ a.s.} \quad (t = 1, \dots, T), \\ y_j \in L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P}; \mathbb{R}^{k_j}), \quad y_j \in Y_j \text{ a.s.} \quad (j = 0, \dots, J), \\ \sum_{\tau=0}^{t-1} A_{t,\tau}(\xi_t) x_{t-\tau} = h_t(\xi_t) \text{ a.s.} \quad (t = 2, \dots, T), \\ \sum_{\tau=0}^j \langle w_{j,\tau}, y_{j-\tau} \rangle + \sum_{\tau=1}^{t_j} \langle b_\tau(\xi_\tau), x_\tau \rangle = 0 \text{ a.s.} \\ (j = 1, \dots, J) \end{array} \right. \right\} \quad (4)$$

by inserting the definition (2). The equivalence is basically in terms of optimal values rather than in terms of solution sets [3, Proposition 4.1]. The resulting problem (4) is almost of the form (1) (but the matrices  $A_{t,\tau}$  then depend on  $\xi^t$  rather than  $\xi_t$  only). This equivalence is, e.g., useful for algorithmic approaches.

**Example 2.1** For  $\mathcal{Z} = L_p(\Omega, \mathcal{F}, \mathbb{P})$ , i.e., for the one-period case, the *Conditional* or *Average Value-at-Risk* at level  $\alpha \in (0, 1)$  ( $\text{CVaR}_\alpha$  or  $\text{AVaR}_\alpha$ , cf. [18, 23] and [6, Chapter 4.4]) is given by

$$\text{AVaR}_\alpha(z) := \frac{1}{\alpha} \int_0^\alpha \text{VaR}_{\bar{\alpha}}(z) d\bar{\alpha} = \inf_{y_0 \in \mathbb{R}} \left\{ y_0 + \frac{1}{\alpha} \mathbb{E} \left[ (y_0 + z)^- \right] \right\} \quad (5)$$

where the second representation on the right is due to [18]. By introducing variables for positive and negative parts of  $y + z$ , respectively,  $\text{AVaR}_\alpha$  can be rewritten in the form (2) with  $J = 1$ ,  $k_0 = 1$ ,  $k_1 = 2$ ,  $c_0 = 1$ ,  $c_1 = (0, \frac{1}{\alpha})$ ,  $w_{1,0} = (1, -1)$ ,  $w_{1,1} = -1$ ,  $Y_0 = \mathbb{R}$ , and  $Y_1 = \mathbb{R}_+^2$ , and, hence, is a polyhedral risk measure. Moreover,  $\text{AVaR}_\alpha$  is known to be a convex measure of risk in the sense of [6], a coherent risk measure in the sense of [1], and it is consistent with 2nd order stochastic dominance [15].

**Example 2.2** Consider the *expected regret* or *expected loss* defined by

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

with some fixed target  $\beta \in \mathbb{R}$ . This functional, too, can be written in the form (2) with  $J = 1$ ,  $k_0 = 1$ ,  $k_1 = 2$ ,  $c_0 = 0$ ,  $c_1 = (0, 1)$ ,  $w_{1,0} = (1, -1)$ ,  $w_{1,1} = 1$ ,  $Y_0 = \{\beta\}$ , and  $Y_1 = \mathbb{R}_+ \times \mathbb{R}_+$ .

**Example 2.3** For the multiperiod case  $J > 1$ , not many instances of risk measures are known. In [3], four different multiperiod risk measures, i.e., four possible choices for  $Y_j$ ,  $c_j$ , and  $w_{j,\tau}$ , have been suggested, cf. Table 1. These instances  $\rho_1$ ,  $\rho_2$ ,  $\rho_3$ , and  $\rho_4$  of (2) can be understood as multiperiod extensions of the Average Value-at-Risk. As a start,  $\rho_1$  is just an average of AVaRs applied to different time stages, whereas  $\rho_2$  is deduced herefrom by interchanging minimisation and summation:

$$\begin{aligned} \rho_1(z) &= \frac{1}{J} \sum_{j=1}^J \inf_{y \in \mathbb{R}} \left\{ y + \frac{1}{\alpha_j} \mathbb{E} \left[ (z_{t_j} + y)^- \right] \right\} \\ \rho_2(z) &= \inf_{y \in \mathbb{R}} \left\{ y + \frac{1}{J} \sum_{j=1}^J \frac{1}{\alpha_j} \mathbb{E} \left[ (z_{t_j} + y)^- \right] \right\} \end{aligned}$$

The instances  $\rho_3$  and  $\rho_4$  are such that the information structure of the value process  $z$  has a definite impact. In particular,  $\rho_4$  can be understood as the

Table 1. Multiperiod polyhedral risk measures, i.e., instances of (2) from [3]. The notation  $y_{j,k}$  refers to  $k$ -th component of  $y_j \in \mathbb{R}^{k_j}$  for  $k = 1, \dots, k_j$  and  $j = 0, \dots, J$ .

No.	primal representation (2)	
$\rho_1$	$\inf \left\{ \frac{1}{J} \sum_{j=1}^J \left( y_{0,j} + \frac{1}{\alpha_j} \mathbb{E}[y_{j,2}] \right) \right\}$	$\left. \begin{array}{l} y_0 \in \mathbb{R}^J, y_j \in L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P}; \mathbb{R}^2) (j = 1, \dots, J), \\ y_j \in \mathbb{R}_+ \times \mathbb{R}_+ \text{ a.s. } (j = 1, \dots, J), \\ y_{j,1} - y_{j,2} = z_{t_j} + y_{0,j} \text{ a.s. } (j = 1, \dots, J) \end{array} \right\}$
$\rho_2$	$\inf \left\{ y_{0,1} + \frac{1}{J} \sum_{j=1}^J \frac{1}{\alpha_j} \mathbb{E}[y_{j,2}] \right\}$	$\left. \begin{array}{l} y_0 \in \mathbb{R}, y_j \in L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P}; \mathbb{R}^2) (j = 1, \dots, J), \\ y_j \in \mathbb{R}_+ \times \mathbb{R}_+ \text{ a.s. } (j = 1, \dots, J), \\ y_{j,1} - y_{j,2} = z_{t_j} + y_{0,1} \text{ a.s. } (j = 1, \dots, J) \end{array} \right\}$
$\rho_3$	$\inf \left\{ y_{0,1} + \frac{1}{J} \sum_{j=1}^J \frac{1}{\alpha_j} \mathbb{E}[y_{j,2}] \right\}$	$\left. \begin{array}{l} y_0 \in \mathbb{R} \times \{0\}, y_j \in L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P}; \mathbb{R}^2) (j = 1, \dots, J), \\ y_j \in \mathbb{R}_+ \times \mathbb{R}_+ \text{ a.s. } (j = 1, \dots, J), \\ y_{j,1} - y_{j,2} = z_{t_j} + y_{0,1} + y_{j-1,2} \text{ a.s. } (j = 1, \dots, J) \end{array} \right\}$
$\rho_4$	$\inf \left\{ \frac{1}{J} \left( y_{0,1} + \sum_{j=1}^J \frac{1}{\alpha_j} \mathbb{E}[y_{j,2}] \right) \right\}$	$\left. \begin{array}{l} y_0 \in \mathbb{R}, y_j \in L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P}; \mathbb{R}^2) (j = 1, \dots, J), \\ y_j \in \mathbb{R} \times \mathbb{R}_+ \text{ a.s. } (j = 1, \dots, J-1), \\ y_J \in \mathbb{R}_+ \times \mathbb{R}_+ \text{ a.s.}, \\ y_{j,1} - y_{j,2} = z_{t_j} + y_{j-1,1} \text{ a.s. } (j = 1, \dots, J) \end{array} \right\}$

multiperiod extension of AVaR according to [17]. In [3] it is shown that each of these four risk measures is multiperiod coherent in the sense of [2].

**Example 2.4** The multiperiod risk measure suggested in [12–14], is based on the concept of the *value of perfect information* (cf., e.g., [21, Chapter 1.2.5]). The risk measure  $\mathcal{R}$  is defined as a difference of two functionals assessing the utility of a financial income stream  $z = (z_{t_1}, \dots, z_{t_J})$  with one functional being derived from the other by relaxing the information constraints, i.e., by assuming that the actual values of all future incomes are perfectly known from the beginning (clairvoyance). The difference  $\mathcal{R}(z)$  is supposed to measure the financial value of being clairvoyant. The utility functional (including the information constraints) is denoted by  $\rho_5$ , it is given<sup>1</sup> by the optimal value of a simple multistage model using given constants  $q_j$  (shortfall cost factors),  $s_j$  (surplus utility factors), and  $d$  (discount factor):

$$\rho_5(z) = \inf \left\{ \begin{array}{l} -s_0 y_{0,1} + \mathbb{E} \left[ \sum_{j=1}^{J-1} (-s_j y_{j,1} + q_j y_{j,3}) - d y_{J,2} + q_J y_{J,3} \right] : \\ y_j \in L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P}; \mathbb{R}^3) (j = 0, \dots, J), y_{0,2} = y_{0,3} = y_{J,1} = 0, \\ y_{j,2} \geq 0 \text{ a.s.}, y_{j,3} \geq 0 \text{ a.s. } (j = 1, \dots, J), \\ y_{j,2} - y_{j,3} = y_{j-1,2} + z_{t_j} - y_{j-1,1} \text{ a.s. } (j = 1, \dots, J) \end{array} \right\} \quad (6)$$

For economic and mathematical consistency, the constants have to satisfy the relations  $d < s_{J-1} < \dots < s_1 < s_0$  and  $s_{j-1} < q_j$  for  $j = 1, \dots, J$ . The functional (6) is of the form (2) with  $k_j = 3$  ( $j \geq 0$ ),  $Y_0 = \mathbb{R} \times \{0\} \times \{0\}$ ,  $Y_j = \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$  ( $1 \leq j < J$ ),  $Y_J = \{0\} \times \mathbb{R}_+ \times \mathbb{R}_+$ ,  $w_{1,1} = (1, 0, 0)$ ,

<sup>1</sup>We adapted the notation of [13, 14] to the notation of polyhedral risk measures (2). To this end, the original identifiers  $t$ ,  $a_t$ ,  $K_t$ ,  $M_t$ ,  $I_t$ , and  $c_t$  from [13, 14] have been replaced by  $j$ ,  $y_{j-1,1}$ ,  $y_{j,2}$ ,  $y_{j,3}$ ,  $z_{t_j}$ , and  $s_{j-1}$ , respectively.

$w_{j,0} = (0, 1, -1)$  ( $j \geq 1$ ),  $w_{j,1} = (1, -1, 0)$  ( $j > 1$ ), further  $w_{j,\tau} = 0$  for  $\tau > 1$ ,  $c_0 = (-s_0, 0, 0)$ ,  $c_j = (-s_j, 0, q_j)$  ( $1 \leq j < J$ ), and  $c_J = (0, -d, q_J)$ . Hence, it is a polyhedral risk measure. It is a coherent risk measure in the sense of [2] if  $s_1 = 1$  (cf. [3]). The assumption of being clairvoyant with respect to the income process is expressed by  $y_j \in L_p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^3)$  instead of  $y_j \in L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P}; \mathbb{R}^3)$ . This relaxation simplifies the utility functional drastically (cf. [13, 14]) such that the overall risk measure, i.e., the difference between  $\rho_5$  and its clairvoyance modification, is given by

$$\mathcal{R}(z_{t_1}, \dots, z_{t_J}) = \rho_5(z_{t_1}, \dots, z_{t_J}) + \sum_{j=1}^J s_{j-1} \mathbb{E}[z_{t_j}]$$

which is always nonnegative. Note that it is intended to minimise this functional when the  $z_{t_j}$  values represent incomes, e.g.,  $z_{t_j} = -\sum_{t=t_{j-1}+1}^{t_j} \langle b_t(\xi_t), x_t \rangle$  instead of  $z_{t_j} = -\sum_{t=1}^{t_j} \langle b_t(\xi_t), x_t \rangle$ .

*Remark 1* In [3], dual representations for (2) have been derived. For these results it is required that the following conditions for  $Y_j$ ,  $c_j$ , and  $w_{j,\tau}$  hold:

- *complete recourse*:  $\langle w_{j,0}, Y_j \rangle = \mathbb{R}$  ( $j = 1, \dots, J$ ),
- *dual feasibility*:  $\bigcap_{j=0}^J \mathcal{D}_{\rho,j} \neq \emptyset$   
with  $\mathcal{D}_{\rho,j} := \{u \in \mathbb{R}^J : c_j + \sum_{\nu=\max\{1,j\}}^J u_\nu w_{\nu,\nu-j} \in -Y_j^*\}$ .

By using the latter notation, the dual representation of (2) reads

$$\rho(z) = \sup \left\{ \begin{array}{l} \inf_{y_0 \in Y_0} \langle c_0 + \sum_{\nu=1}^J \mathbb{E}[\lambda_\nu] w_{\nu,\nu}, y_0 \rangle \\ -\mathbb{E} \left[ \sum_{j=1}^J \lambda_j z_{t_j} \right] \end{array} \middle| \begin{array}{l} \lambda \in \times_{j=1}^J L_{p'}(\Omega, \mathcal{F}_{t_j}, \mathbb{P}), \\ \mathbb{E}[\lambda | \mathcal{F}_{t_j}] \in \mathcal{D}_{\rho,j} \text{ a.s.} \\ (j = 1, \dots, J) \end{array} \right\} \quad (7)$$

with  $p' \in [1, \infty]$  such that  $\frac{1}{p} + \frac{1}{p'} = 1$ , or

$$\rho(z) = \sup \left\{ -\mathbb{E} \left[ \sum_{j=1}^J \lambda_j z_{t_j} \right] \middle| \begin{array}{l} \lambda \in \times_{j=1}^J L_{p'}(\Omega, \mathcal{F}_{t_j}, \mathbb{P}), \\ \mathbb{E}[\lambda | \mathcal{F}_{t_j}] \in \mathcal{D}_{\rho,j} \text{ a.s.} (j = 0, \dots, J) \end{array} \right\} \quad (8)$$

for the case that  $Y_0$  is a cone. Moreover, it has been shown in [3] that, if complete recourse and dual feasibility hold, the polyhedral risk measure  $\rho$  is finite, continuous, and convex on  $\mathcal{Z}$ . Further, a criterion for (multiperiod) coherence (cf. [1, 2]) has been stated based on the dual representation (8).

Table 2. Feasible sets of the dual representations (7) for the exemplary polyhedral risk measures.

risk.	$\mathcal{D}_{\rho,0}$	$\mathcal{D}_{\rho,j} (j = 1, \dots, J-1)$	$\mathcal{D}_{\rho,J}$
AVaR $_{\alpha}$	$\{1\}$		$[0, \frac{1}{\alpha}]$
EL $_{\beta}$	$\text{sign}(\beta) \cdot \mathbb{R}_+$		$[0, 1]$
$\rho_1$	$\{(\frac{1}{J}, \dots, \frac{1}{J})\}$	$\{u \in \mathbb{R}^J : 0 \leq u_j, u_j \leq \frac{1}{J\alpha_j}\}$	$\{u \in \mathbb{R}^J : u_J \in [0, \frac{1}{J\alpha_J}]\}$
$\rho_2$	$\{u \in \mathbb{R}^J : \sum u_j = 1\}$	$\{u \in \mathbb{R}^J : 0 \leq u_j, u_j \leq \frac{1}{J\alpha_j}\}$	$\{u \in \mathbb{R}^J : u_J \in [0, \frac{1}{J\alpha_J}]\}$
$\rho_3$	$\{u \in \mathbb{R}^J : \sum u_j = 1\}$	$\{u \in \mathbb{R}^J : 0 \leq u_j, u_j + u_{j+1} \leq \frac{1}{J\alpha_j}\}$	$\{u \in \mathbb{R}^J : u_J \in [0, \frac{1}{J\alpha_J}]\}$
$\rho_4$	$\{u \in \mathbb{R}^J : u_1 = \frac{1}{J}\}$	$\{u \in \mathbb{R}^J : u_j = u_{j+1}, u_j \leq \frac{1}{J\alpha_j}\}$	$\{u \in \mathbb{R}^J : u_J \in [0, \frac{1}{J\alpha_J}]\}$

### 3 Stability of multistage stochastic programs

We consider a multistage stochastic program of the form (3) with a polyhedral risk measure  $\rho$  of the form (2) in the objective and study the stability behaviour of its optimal value with respect to perturbations of the stochastic input process  $\xi = (\xi_1, \dots, \xi_T)$ . One possible approach for this analysis would be to analyse the equivalent problem (4) which is similar to problem (1). However, it has turned out to be more fruitful to pursue the other approach, namely to analyse sensitivity of  $\rho$  and then to use these results to analyse problem (3) directly.

For the sensitivity analysis of  $\rho$  resp. (2) with regard to perturbations of  $\xi$  in (3), observe that  $\rho$  does not only depend on  $z = (z_{t_1}, \dots, z_{t_J})$  but also depends on  $\xi$  via the  $\sigma$ -fields  $\mathcal{F}_t = \sigma(\xi^t)$ . Moreover, perturbations of  $\xi$  in (3) may cause variations of  $x$  and, since  $z = z(\xi, x)$  in (3), variations of  $z$  in (2). Therefore we will use notations like  $\rho(z, \xi)$  instead of just  $\rho(z)$  from now on. Further, we introduce the notations

$$\mathbb{Z} := \{(z, \xi) : \xi \in L_r(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^s), z \in \times_{j=1}^J L_p(\Omega, \sigma(\xi^{t_j}), \mathbb{P})\}$$

for pairs of processes such that  $z$  is adapted to  $\xi$ . For  $(z, \xi) \in \mathbb{Z}$  we set

$$\mathcal{Y}_{\rho}(z, \xi) := \left\{ y \in \times_{j=0}^J L_p(\Omega, \sigma(\xi^{t_j}), \mathbb{P}; \mathbb{R}^{k_j}) \left| \begin{array}{l} y_j \in Y_j \text{ a.s. } (j = 0, \dots, J), \\ \sum_{\tau=0}^{j-1} \langle w_{j,\tau}, y_{j-\tau} \rangle = z_{t_j} \text{ a.s.} \\ (j = 1, \dots, J) \end{array} \right. \right\} \tag{9}$$

for the feasible set with  $y = (y_0, y_1, \dots, y_J)$ . Accordingly we set

$$F_{\rho}(y) := \mathbb{E} \left[ \sum_{j=0}^J \langle c_j, y_j \rangle \right] \tag{10}$$

for the objective. With these notations formula (2) can be written in the following short form:  $\rho(z, \xi) = \inf\{F_{\rho}(y) : y \in \mathcal{Y}_{\rho}(z, \xi)\}$ . For a given level

$\varepsilon \geq 0$  the sets

$$l_{\rho,\varepsilon}(z, \xi) := \{y \in \mathcal{Y}_\rho(z, \xi) : F_\rho(y) \leq \rho(z, \xi) + \varepsilon\} \quad (11)$$

are called the *level sets*. For  $\varepsilon > 0$  these level sets are nonempty. For  $\varepsilon = 0$ ,  $l_{\rho,0}(z, \xi) =: S_\rho(z, \xi)$  is called the *solution set*.

**PROPOSITION 3.1** *Consider a multiperiod polyhedral risk measure  $\rho$  of the form (2) on  $L_p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^J)$  satisfying complete recourse and dual feasibility. Then there exists a constant  $K_\rho > 0$  such that*

$$|\rho(z, \xi) - \rho(\tilde{z}, \tilde{\xi})| \leq K_\rho \left( \|z - \tilde{z}\|_p + D_\rho((z, \xi), (\tilde{z}, \tilde{\xi})) \right) \quad (12)$$

for  $(z, \xi), (\tilde{z}, \tilde{\xi}) \in \mathbb{Z}$ . Here,  $D_\rho$  denotes the filtration distance for  $\rho$  given by

$$\begin{aligned} D_\rho((z, \xi), (\tilde{z}, \tilde{\xi})) &:= \sup_{\varepsilon > 0} D_{\rho,\varepsilon}((z, \xi), (\tilde{z}, \tilde{\xi})) \\ D_{\rho,\varepsilon}((z, \xi), (\tilde{z}, \tilde{\xi})) &:= \\ \inf \left\{ \sum_{j=1}^{J-1} \max \left\{ \|\bar{y}_j - \mathbb{E}[\bar{y}_j | \sigma(\tilde{\xi}^{t_j})]\|_p, \|\tilde{y}_j - \mathbb{E}[\tilde{y}_j | \sigma(\xi^{t_j})]\|_p \right\} \right. & \left. \begin{array}{l} \bar{y} \in l_{\rho,\varepsilon}(z, \xi), \\ \tilde{y} \in l_{\rho,\varepsilon}(\tilde{z}, \tilde{\xi}) \end{array} \right\} \end{aligned}$$

*Proof* Let  $\varepsilon > 0$ ,  $(z, \xi), (\tilde{z}, \tilde{\xi}) \in \mathbb{Z}$ , and  $\bar{y} = (\bar{y}_0, \bar{y}_1, \dots, \bar{y}_J) \in l_{\rho,\varepsilon}(z, \xi)$ . In the following, an element  $\tilde{y} = (\tilde{y}_0, \tilde{y}_1, \dots, \tilde{y}_J) \in \mathcal{Y}_\rho(\tilde{z}, \tilde{\xi})$  is recursively constructed such that its distance to  $\mathbb{E}[\bar{y}_j | \sigma(\tilde{\xi}^{t_j})]$  is small in some sense. To this end, consider the set-valued mappings (multifunctions)

$$\begin{aligned} M_j : \mathbb{R} &\rightrightarrows \mathbb{R}^{k_j} \\ u &\mapsto M_j(u) := \{y_j \in Y_j : \langle w_{j,0}, y_j \rangle = u\} \end{aligned}$$

for  $j = 1, \dots, J$ . Note that each  $M_j$  has polyhedral graph and, hence, is Lipschitz continuous with respect to the Hausdorff distance [19, Example 9.35] with some modulus  $l_j$ , thus,

$$\inf_{y \in M_j(u)} |\hat{y} - y| \leq l_j |\hat{u} - u| \quad (13)$$

for all (non-random)  $\hat{u}, u \in \mathbb{R}$  and  $\hat{y} \in M_j(\hat{u})$ . Here,  $|\cdot|$  denotes the Euclidian norm in  $\mathbb{R}^{k_j}$ . Now, the random element  $\tilde{y}$  is constructed as follows: For  $j = 0$ , we set  $\tilde{y}_0 := \bar{y}_0$ . For  $j > 0$ , consider the random elements

$$\bar{u}_j(\cdot) := z_{t_j}(\cdot) - \sum_{\tau=1}^j \langle w_{j,\tau}, \bar{y}_{j-\tau}(\cdot) \rangle \quad \tilde{u}_j(\cdot) := \tilde{z}_{t_j}(\cdot) - \sum_{\tau=1}^j \langle w_{j,\tau}, \tilde{y}_{j-\tau}(\cdot) \rangle$$

as well as the following set-valued mappings:

$$M_j^1 : \Omega \rightrightarrows \mathbb{R}^{k_j} \qquad M_j^2 : \Omega \rightrightarrows \mathbb{R}^{k_j}$$

$$\omega \mapsto M_j(\tilde{u}_j(\omega)) \qquad \omega \mapsto \arg \min_{y \in M_j^1(\omega)} |\mathbb{E}[\tilde{y}_j | \sigma(\tilde{\xi}^{t_j})](\omega) - y|$$

Obviously,  $M_j^1(\omega)$  is closed, convex, and, due to the complete recourse assumption, non-empty for every  $\omega \in \Omega$ .  $M_j^2(\omega)$  is non-empty for  $\omega \in \Omega$  because the distance function  $|\mathbb{E}[\tilde{y}_j | \sigma(\tilde{\xi}^{t_j})](\omega) - \cdot|$  is coercive. Further, since  $\tilde{u}_j$  is measurable with respect to  $\sigma(\tilde{\xi}^{t_{j-1}})$ ,  $M_j^1$  and  $M_j^2$  are measurable with respect to  $\sigma(\tilde{\xi}^{t_j})$ ; cf., e.g., [19, Theorem 14.36] and [19, Theorem 14.37]. The latter theorem also guarantees the existence of a  $\sigma(\tilde{\xi}^{t_j})$ -measurable function  $\tilde{y}_j$  with  $\tilde{y}_j(\omega) \in M_j^2(\omega)$  for  $\omega \in \Omega$ . Now, using (13) with  $\hat{y} = \mathbb{E}[\tilde{y}_j | \sigma(\tilde{\xi}^{t_j})](\omega)$ ,  $\hat{u} = \mathbb{E}[\tilde{u}_j | \sigma(\tilde{\xi}^{t_j})](\omega)$ ,  $u = \tilde{u}_j(\omega)$ , and  $y = \tilde{y}_j(\omega)$  (note that  $\tilde{y}_j$  was chosen as a pointwise minimiser) yields the estimate

$$\begin{aligned} & |\mathbb{E}[\tilde{y}_j | \sigma(\tilde{\xi}^{t_j})] - \tilde{y}_j| \\ & \leq l_j |\mathbb{E}[\tilde{u}_j | \sigma(\tilde{\xi}^{t_j})] - \tilde{u}_j| \\ & = l_j |\mathbb{E}[z_{t_j} - \sum_{\tau=1}^j \langle w_{j,\tau}, \tilde{y}_{j-\tau} \rangle | \sigma(\tilde{\xi}^{t_j})] - \tilde{z}_{t_j} + \sum_{\tau=1}^j \langle w_{j,\tau}, \tilde{y}_{j-\tau} \rangle| \\ & \leq l_j \left( |\mathbb{E}[z_{t_j} - \tilde{z}_{t_j} | \sigma(\tilde{\xi}^{t_j})]| + \sum_{\tau=1}^j |w_{j,\tau}| |\mathbb{E}[\tilde{y}_{j-\tau} | \sigma(\tilde{\xi}^{t_j})] - \tilde{y}_{j-\tau}| \right) \\ & \leq l_j \left( \mathbb{E}[|z_{t_j} - \tilde{z}_{t_j}| | \sigma(\tilde{\xi}^{t_j})] \right. \\ & \quad \left. + \sum_{\tau=1}^j |w_{j,\tau}| |\mathbb{E}[\tilde{y}_{j-\tau} | \sigma(\tilde{\xi}^{t_{j-\tau}})] - \tilde{y}_{j-\tau}| \right. \\ & \quad \left. + \sum_{\tau=1}^j |w_{j,\tau}| |\mathbb{E}[\tilde{y}_{j-\tau} | \sigma(\tilde{\xi}^{t_{j-\tau}})] - \mathbb{E}[\tilde{y}_{j-\tau} | \sigma(\tilde{\xi}^{t_j})]| \right) \end{aligned}$$

pointwise on  $\Omega$  for  $j = 1, \dots, J$ . Note that Jensen's inequality is used for the first term of the final estimate. Putting these estimates together recursively (recall that  $\bar{y}_0 = \tilde{y}_0$ ) yields

$$\begin{aligned} |\mathbb{E}[\bar{y}_j | \sigma(\tilde{\xi}^{t_j})] - \tilde{y}_j| & \leq \sum_{i=1}^j K_{j,i} \mathbb{E}[|z_{t_i} - \tilde{z}_{t_i}| | \sigma(\tilde{\xi}^{t_i})] \\ & \quad + \sum_{i=1}^j \sum_{\tau=1}^i C_{j,i,\tau} |\mathbb{E}[\bar{y}_{i-\tau} | \sigma(\tilde{\xi}^{t_{i-\tau}})] - \mathbb{E}[\bar{y}_{i-\tau} | \sigma(\tilde{\xi}^{t_i})]| \end{aligned}$$

with some positive constants  $K_{j,i}$  and  $C_{j,i,\tau}$ . Hence, since  $\bar{y} \in l_{\rho,\varepsilon}(z, \xi)$  and

$\tilde{y} \in \mathcal{Y}_\rho(\tilde{z}, \tilde{\xi})$ , we have

$$\begin{aligned}
 & \rho(\tilde{z}, \tilde{\xi}) - \rho(z, \xi) \\
 & \leq \mathbb{E} \left[ \sum_{j=0}^J \langle c_j, \tilde{y}_j \rangle \right] - \mathbb{E} \left[ \sum_{j=0}^J \langle c_j, \bar{y}_j \rangle \right] + \varepsilon \\
 & = \sum_{j=1}^J \mathbb{E} \left[ \langle c_j, \tilde{y}_j - \mathbb{E}[\tilde{y}_j | \sigma(\tilde{\xi}^{t_j})] \rangle \right] + \varepsilon \\
 & \leq \sum_{j=1}^J |c_j| \mathbb{E} \left[ |\tilde{y}_j - \mathbb{E}[\tilde{y}_j | \sigma(\tilde{\xi}^{t_j})]| \right] + \varepsilon \\
 & \leq \sum_{j=1}^J \left( K_j \mathbb{E} \left[ |z_{t_j} - \tilde{z}_{t_j}| \right] + \sum_{\tau=1}^j C_{j,\tau} \mathbb{E} \left[ |\mathbb{E}[\tilde{y}_{j-\tau} | \sigma(\tilde{\xi}^{t_j-\tau})] - \bar{y}_{j-\tau}| \right] \right) + \varepsilon \\
 & = \sum_{j=1}^J K_j \mathbb{E} \left[ |z_{t_j} - \tilde{z}_{t_j}| \right] + \sum_{j=1}^{J-1} C_j \mathbb{E} \left[ |\mathbb{E}[\tilde{y}_j | \sigma(\tilde{\xi}^{t_j})] - \bar{y}_j| \right] + \varepsilon \\
 & \leq C \left( \|z - \tilde{z}\|_p + \sum_{j=1}^{J-1} \|\mathbb{E}[\tilde{y}_j | \sigma(\tilde{\xi}^{t_j})] - \bar{y}_j\|_p \right) + \varepsilon
 \end{aligned}$$

with some other positive constants  $K_j$ ,  $C_{j,\tau}$ ,  $C_j$ , and  $C$ . Note that the terms in the final line of the previous display do not depend on  $\tilde{y}$  which has been constructed dependent on an arbitrary  $\bar{y} \in l_{\rho,\varepsilon}(z, \xi)$ . Thus, the roles of  $(z, \xi)$  and  $(\tilde{z}, \tilde{\xi})$  can be changed, i.e., for arbitrary  $\tilde{y} \in l_{\rho,\varepsilon}(\tilde{z}, \tilde{\xi})$  it holds that

$$\rho(z, \xi) - \rho(\tilde{z}, \tilde{\xi}) \leq \hat{C} \left( \|z - \tilde{z}\|_p + \sum_{j=1}^{J-1} \|\mathbb{E}[\tilde{y}_j | \sigma(\xi^{t_j})] - \tilde{y}_j\|_p \right) + \varepsilon$$

with some positive constant  $\hat{C}$ . With  $K_\rho := \max\{C, \hat{C}\}$  it follows that

$$\begin{aligned}
 & |\rho(z, \xi) - \rho(\tilde{z}, \tilde{\xi})| \\
 & \leq K_\rho \left( \|z - \tilde{z}\|_p + \sum_{j=1}^{J-1} \max \left\{ \|\bar{y}_j - \mathbb{E}[\tilde{y}_j | \sigma(\tilde{\xi}^{t_j})]\|_p, \|\tilde{y}_j - \mathbb{E}[\tilde{y}_j | \sigma(\xi^{t_j})]\|_p \right\} \right) + \varepsilon
 \end{aligned}$$

for arbitrary  $\bar{y} \in l_{\rho,\varepsilon}(z, \xi)$  and  $\tilde{y} \in l_{\rho,\varepsilon}(\tilde{z}, \tilde{\xi})$ . Hence, we can pass to the infimum arriving at

$$\begin{aligned}
 |\rho(z, \xi) - \rho(\tilde{z}, \tilde{\xi})| & \leq K_\rho \left( \|z - \tilde{z}\|_p + D_{\rho,\varepsilon}(z, \xi, \tilde{z}, \tilde{\xi}) \right) + \varepsilon \\
 & \leq K_\rho \left( \|z - \tilde{z}\|_p + D_\rho(z, \xi, \tilde{z}, \tilde{\xi}) \right) + \varepsilon
 \end{aligned}$$

and because  $\varepsilon$  was chosen arbitrarily the assertion follows.  $\square$

Next, we make use of the latter result for the analysis of the risk-averse stochastic program (3). To this end, we introduce similar notations as for  $\rho$  that stress the dependence on  $\xi$ :

$$F(\xi, x) := \rho(z(\xi, x))$$

for the objective with

$$z(\xi, x) := (z_{t_1}(\xi, x), \dots, z_{t_J}(\xi, x)) \quad z_t(\xi, x) = z_t(\xi^t, x^t) := -\sum_{\tau=1}^t \langle b_\tau(\xi_\tau), x_\tau \rangle$$

and

$$\mathcal{X}(\xi) := \left\{ x \in \times_{t=1}^T L_{r'}(\Omega, \sigma(\xi^t), \mathbb{P}; \mathbb{R}^{m_t}) \mid \begin{array}{l} x_1 \in X_1, \\ x_t \in \mathcal{X}_t(x^{t-1}, \xi^t) \text{ a.s. } (t = 2, \dots, T) \end{array} \right\}$$

for the constraint set with

$$\mathcal{X}_t(x^{t-1}, \xi^t) := \{x_t \in X_t : \sum_{\tau=0}^{t-1} A_{t,\tau} x_{t-\tau} = h_t(\xi_t)\}$$

for  $t = 2, \dots, T$ . Then the model (3) can be written in the following short form:

$$\min\{F(\xi, x) : x \in \mathcal{X}(\xi)\} \tag{14}$$

and with  $v(\xi) := \inf\{F(\xi, x) : x \in \mathcal{X}(\xi)\}$  we denote the optimal value of (14). For any  $\varepsilon \geq 0$  let

$$l_\varepsilon(F(\xi, \cdot)) := \{x \in \mathcal{X}(\xi) : F(\xi, x) \leq v(\xi) + \varepsilon\}$$

denote its  $\varepsilon$ -level set. For the integrability numbers  $r, r', p \geq 1$ , we will set  $r$  and  $r'$  in dependence of the class of problem (3) by the assignment

$$r := \begin{cases} \in [p, \infty) \text{ arbitrarily,} & \text{if only costs or right-hand sides are random} \\ 2p & \text{, if only costs and right-hand sides are random} \\ pT & \text{, if all technology matrices are random} \end{cases}$$

$$r' := \begin{cases} \frac{pr}{r-p} & \text{, if only costs are random} \\ r & \text{, if only right-hand sides are random} \\ r = 2p & \text{, if costs and right-hand sides are random} \\ \infty & \text{, if all technology matrices are random} \end{cases} \tag{15}$$

which implies  $r \geq p$  and  $r' \geq p$ . We will consider the following conditions the optimization model (3):

- (A1)  $\xi \in L_r(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^s)$
- (A2) There exists a  $\delta_1 > 0$  such that for any  $\tilde{\xi} \in L_r(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^s)$  with  $\|\tilde{\xi} - \xi\|_r \leq \delta_1$ , any  $t = 2, \dots, T$  and any  $x_1 \in X_1$ ,  $x_\tau \in L_{r'}(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^{m_\tau})$  with  $x_\tau \in \mathcal{X}_\tau(x^{\tau-1}, \tilde{\xi}_\tau)$ ,  $\tau = 2, \dots, t-1$ , the  $t$ -th feasibility set  $\mathcal{X}_t(x^{t-1}, \tilde{\xi}_t)$  is nonempty (*relatively complete recourse locally around  $\xi$* ).
- (A3) The optimal values  $v(\tilde{\xi})$  of (14) with input  $\tilde{\xi}$  are finite for all  $\tilde{\xi}$  in a neighbourhood of  $\xi$  and the objective function  $F$  is *level-bounded locally uniformly*

at  $\xi$ , i.e., for some  $\varepsilon_0 > 0$  there exists a  $\delta_2 > 0$  and a bounded subset  $B$  of  $L_{r'}(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^m)$  such that  $v(\xi) \in \mathbb{R}$  and  $\emptyset \neq l_{\varepsilon_0}(F(\xi, \cdot)) \subseteq B$  for all  $\tilde{\xi} \in L_r(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^s)$  with  $\|\tilde{\xi} - \xi\|_r \leq \delta_2$ .

- (A4) The recourse matrices  $A_{t,0}(\xi_t)$  are fixed, i.e., they do not depend on  $\xi_t$  ( $t = 1, \dots, T$ ).

**THEOREM 3.2** *For the multistage stochastic program (3) respectively (14), let  $p \in [1, \infty)$  and  $r$  and  $r'$  be defined by (15) and assume that the multiperiod polyhedral risk measure  $\rho$  on  $L_p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^J)$  of the form (2) satisfies complete recourse and dual feasibility. Furthermore, let (A1)–(A4) be satisfied and  $X_1$  be bounded. Then there exists positive constants  $K$ ,  $\varepsilon_0$  and  $\delta$  such that the estimate*

$$|v(\xi) - v(\tilde{\xi})| \leq K \left( \|\xi - \tilde{\xi}\|_r + D_{f,\rho}(\xi, \tilde{\xi}) \right) \quad (16)$$

holds for all random elements  $\tilde{\xi} \in L_r(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^s)$  with  $\|\tilde{\xi} - \xi\|_r \leq \delta$ . Here, the filtration distance  $D_{f,\rho}(\xi, \tilde{\xi})$  is given by

$$D_{f,\rho}(\xi, \tilde{\xi}) := \sup_{\varepsilon \in (0, \varepsilon_0]} D_{f,\rho,\varepsilon}$$

$$D_{f,\rho,\varepsilon} := \inf \left\{ \begin{array}{l} \sum_{t=2}^{T-1} \max\{\|\mathbb{E}[x_t | \sigma(\xi^t)] - x_t\|_{r'}, \|\mathbb{E}[\tilde{x}_t | \sigma(\xi^t)] - \tilde{x}_t\|_{r'}\} \\ + \sum_{j=1}^{J-1} \max\{\|\mathbb{E}[y_j | \sigma(\xi^{t_j})] - y_j\|_p, \|\mathbb{E}[\tilde{y}_j | \sigma(\xi^{t_j})] - \tilde{y}_j\|_p\} \end{array} \right\}$$

where the infimum is taken with respect to all  $x \in l_\varepsilon(F(\xi, \cdot))$ ,  $\tilde{x} \in l_\varepsilon(F(\tilde{\xi}, \cdot))$ ,  $y \in l_{\rho,\varepsilon}(z(\xi, x), \xi)$ , and  $\tilde{y} \in l_{\rho,\varepsilon}(z(\tilde{\xi}, \tilde{x}), \tilde{\xi})$ .

*Proof* For the sake of clarity and without loss of generality we restrict the following presentation to the case that  $A_{t,\tau} = 0$  for  $\tau \geq 2$ . Since [10, Theorem 2.1]) deals with the same problem but with expectation objective, we will here use from the proof of [10, Theorem 2.1]) some formulas of which the derivation does not depend on the objective.

Let  $\varepsilon_0$ ,  $\delta_1$ , and  $\delta_2$  be selected as in (A2) and (A3) and set  $\delta := \min\{\delta_1, \delta_2\} > 0$ . Let  $\varepsilon \in (0, \varepsilon_0]$  and  $\tilde{\xi} \in L_r(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^s)$  be such that  $\|\tilde{\xi} - \xi\|_r < \delta$ . First, recall from the proof of Proposition 3.1 that there exists a positive constant  $K_\rho$  such that

$$\begin{aligned} \rho(\tilde{z}, \tilde{\xi}) - \rho(z, \xi) &\leq K_\rho \left( \|z - \tilde{z}\|_p + \sum_{j=1}^{J-1} \|\mathbb{E}[\tilde{y}_j | \sigma(\tilde{\xi}^{t_j})] - \tilde{y}_j\|_p \right) + \varepsilon \\ \rho(z, \xi) - \rho(\tilde{z}, \tilde{\xi}) &\leq K_\rho \left( \|z - \tilde{z}\|_p + \sum_{j=1}^{J-1} \|\mathbb{E}[\tilde{y}_j | \sigma(\xi^{t_j})] - \tilde{y}_j\|_p \right) + \varepsilon \end{aligned} \quad (17)$$

holds for all  $\tilde{y} \in l_{\rho,\varepsilon}(z, \xi)$  and  $\tilde{y} \in l_{\rho,\varepsilon}(\tilde{z}, \tilde{\xi})$  and all pairs  $(z, \xi)$  and  $(\tilde{z}, \tilde{\xi})$  in  $\mathbb{Z}$ . Now, let  $\tilde{x} \in l_\varepsilon(F(\xi, \cdot))$ . In the following, we construct  $\tilde{x} \in \mathcal{X}(\tilde{\xi})$  in same manner as in the proof of [10, Theorem 2.1] (similarly to  $\tilde{y}$  in the proof of

Proposition 3.1) such that  $\bar{x}_1 = \tilde{x}_1$  and the estimate<sup>1</sup>

$$\begin{aligned} |\mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)] - \tilde{x}_t| &\leq l_t \left( \sum_{\tau=2}^t \max\{1, |\tilde{\xi}^t|^{t-\tau}\} \mathbb{E}[|\xi_\tau - \tilde{\xi}_\tau| | \sigma(\tilde{\xi}^\tau)] \right. \\ &\quad \left. + \sum_{\tau=2}^{t-1} \max\{1, |\tilde{\xi}^t|^{t-\tau}\} \mathbb{E}[|\bar{x}_\tau - \mathbb{E}[\bar{x}_\tau | \sigma(\tilde{\xi}^\tau)]| | \sigma(\tilde{\xi}^{\tau+1})] \right) \end{aligned} \quad (18)$$

holds with some positive constants  $l_t$  for  $t = 2, \dots, T$ . Note that the first sum on the right-hand side disappears if only costs are random and that both max-terms vanish if the technology matrices are not random.

Now, because  $\bar{x} \in l_\varepsilon(F(\xi, \cdot))$  and  $\tilde{x} \in \mathcal{X}(\tilde{\xi})$ , we have that for the optimal values it holds that for any  $\bar{y} \in l_{\rho, \varepsilon}(z(\xi, \bar{x}), \xi)$

$$\begin{aligned} v(\tilde{\xi}) - v(\xi) &\leq \rho(z(\tilde{\xi}, \tilde{x}), \tilde{\xi}) - \rho(z(\xi, \bar{x}), \xi) + \varepsilon \\ &\leq K_\rho \left( \|z(\xi, \bar{x}) - z(\tilde{\xi}, \tilde{x})\|_p + \sum_{j=1}^{J-1} \|\mathbb{E}[\bar{y}_j | \sigma(\tilde{\xi}^{t_j})] - \bar{y}_j\|_p \right) + 2\varepsilon \end{aligned} \quad (19)$$

where (17) is used for the second estimate.

Next, we derive an estimate for  $\|z(\xi, \bar{x}) - z(\tilde{\xi}, \tilde{x})\|_p$  by making use of (18).

With  $\hat{x}_t := \mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)]$  and  $\hat{x} = (\hat{x}_1, \dots, \hat{x}_T)$  we have

$$\begin{aligned} \|z(\xi, \bar{x}) - z(\tilde{\xi}, \tilde{x})\|_p &\leq \|z(\xi, \bar{x}) - z(\tilde{\xi}, \bar{x})\|_p + \|z(\tilde{\xi}, \bar{x}) - z(\tilde{\xi}, \hat{x})\|_p + \|z(\tilde{\xi}, \hat{x}) - z(\tilde{\xi}, \tilde{x})\|_p \end{aligned} \quad (20)$$

and for the first summand we obtain

$$\begin{aligned} \|z(\xi, \bar{x}) - z(\tilde{\xi}, \bar{x})\|_p &= \left( \mathbb{E} \left[ \sum_{j=1}^J |z_{t_j}(\xi, \bar{x}) - z_{t_j}(\tilde{\xi}, \bar{x})|^p \right] \right)^{\frac{1}{p}} \\ &= \left( \sum_{j=1}^J \mathbb{E} \left[ \left| \sum_{t=1}^{t_j} \langle b_t(\xi_t) - b_t(\tilde{\xi}_t), \bar{x}_t \rangle \right|^p \right] \right)^{\frac{1}{p}} \\ &\leq \sum_{j=1}^J \left( \mathbb{E} \left[ \left| \sum_{t=1}^{t_j} \langle b_t(\xi_t) - b_t(\tilde{\xi}_t), \bar{x}_t \rangle \right|^p \right] \right)^{\frac{1}{p}} \\ &\leq \sum_{j=1}^J \sum_{t=1}^{t_j} \left( \mathbb{E} \left[ |\langle b_t(\xi_t) - b_t(\tilde{\xi}_t), \bar{x}_t \rangle|^p \right] \right)^{\frac{1}{p}} \\ &\leq J \sum_{t=1}^T \left( \mathbb{E} \left[ |\langle b_t(\xi_t) - b_t(\tilde{\xi}_t), \bar{x}_t \rangle|^p \right] \right)^{\frac{1}{p}} \\ &\leq J \sum_{t=1}^T \left( \mathbb{E} \left[ |b_t(\xi_t) - b_t(\tilde{\xi}_t)|^p |\bar{x}_t|^p \right] \right)^{\frac{1}{p}} \\ &\leq J \sum_{t=1}^T \|b_t(\xi_t) - b_t(\tilde{\xi}_t)\|_r \|\bar{x}_t\|_{r'} \end{aligned}$$

---

<sup>1</sup>In the proof of [10, Theorem 2.1] the term  $(1 + |\bar{x}_{\tau-1}|)$  occurs additionally in the first conditional expectation on the right-hand side of (18) if the technology matrices are not random. However, in this case, due to (A3), we have that  $B$  is bounded in  $L_\infty$ , hence, since  $\bar{x} \in l_\varepsilon(F(\xi, \cdot)) \subseteq B$ ,  $|\bar{x}_{\tau-1}|$  can be estimated by  $\|\bar{x}_{\tau-1}\|_\infty$  and we assume the latter norm to be integrated in the constant  $l_t$ .

where Minkowski's inequality in  $L_p(\Omega, \mathcal{F}, \mathbb{P})$  as well as the Cauchy-Schwarz inequality in  $\mathbb{R}^{m_t}$  have been used. For the final estimate, a generalised version of Hölder's inequality has been used which is valid for  $\frac{1}{r} + \frac{1}{r'} = \frac{1}{p}$  (the case of stochastic cost and deterministic technology matrices) as well as for  $p \leq r < r' = \infty$  (the case of stochastic technology matrices). For the case that only right-hand sides are random, this estimate is also valid, because then the deterministic<sup>1</sup> cost factors  $b_t$  can be moved outside the expectation and Lyapunov's inequality yields the same result. Now, since  $\bar{x} \in B$ ,  $B$  is  $L_{r'}$ -bounded and  $b_t(\cdot)$  is affine linear, it holds that that

$$\|z(\xi, \bar{x}) - z(\tilde{\xi}, \bar{x})\|_p \leq C_1 \|\xi - \tilde{\xi}\|_r$$

with some positive constant  $C_1$  depending on  $B$  and  $b_t$  ( $t = 1, \dots, T$ ). For the second and the third summand in (20) we conclude analogously:

$$\begin{aligned} \|z(\tilde{\xi}, \bar{x}) - z(\tilde{\xi}, \hat{x})\|_p &\leq J \sum_{t=1}^T \left( \mathbb{E} \left[ |b_t(\tilde{\xi}_t)|^p |\bar{x}_t - \mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)]|^p \right] \right)^{\frac{1}{p}} \\ &\leq J \sum_{t=1}^T \|b_t(\tilde{\xi}_t)\|_r \|\bar{x}_t - \mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)]\|_{r'} \\ \|z(\tilde{\xi}, \hat{x}) - z(\tilde{\xi}, \tilde{x})\|_p &\leq J \sum_{t=1}^T \left( \mathbb{E} \left[ |b_t(\tilde{\xi}_t)|^p |\mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)] - \tilde{x}_t|^p \right] \right)^{\frac{1}{p}} \\ &\leq J \sum_{t=1}^T \|b_t(\tilde{\xi}_t)\|_r \|\mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)] - \tilde{x}_t\|_{r'} \end{aligned}$$

where we have resubstituted  $\hat{x}_t = \mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)]$ . Since  $\xi \in L_r(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^s)$  and  $\|\xi - \tilde{\xi}\|_r \leq \delta$ , it holds that

$$\begin{aligned} \|z(\tilde{\xi}, \bar{x}) - z(\tilde{\xi}, \hat{x})\|_p &\leq C_2 \sum_{t=1}^T \|\bar{x}_t - \mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)]\|_{r'} \\ \|z(\tilde{\xi}, \hat{x}) - z(\tilde{\xi}, \tilde{x})\|_p &\leq C_2 \sum_{t=1}^T \|\mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)] - \tilde{x}_t\|_{r'} \end{aligned}$$

with some positive constant  $C_2$  depending on  $\xi$ ,  $\delta$ , and  $b_t$  ( $t = 1, \dots, T$ ). Now, the latter estimate will be continued by inserting (18).

First, we consider the situation that only cost are random and  $r' < \infty$ . We use Minkowski's and Jensen's inequality and arrive at

$$\begin{aligned} \|z(\tilde{\xi}, \hat{x}) - z(\tilde{\xi}, \tilde{x})\|_p &\leq C_2 \sum_{t=1}^T \left( \mathbb{E} \left[ |\mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)] - \tilde{x}_t|^{r'} \right] \right)^{\frac{1}{r'}} \\ &\leq C_2 \sum_{t=1}^T l_t \left( \mathbb{E} \left[ \left| \sum_{\tau=2}^{t-1} \mathbb{E}[\bar{x}_\tau - \mathbb{E}[\bar{x}_\tau | \sigma(\tilde{\xi}^\tau)] | \sigma(\tilde{\xi}^{\tau+1})] \right|^{r'} \right] \right)^{\frac{1}{r'}} \\ &\leq C_3 \sum_{t=1}^T \left( \mathbb{E} \left[ |\bar{x}_t - \mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)]|^{r'} \right] \right)^{\frac{1}{r'}} \end{aligned}$$

---

<sup>1</sup>Of course, if  $b_t$  are non-random, both sides of the above estimate are zero anyway, but the same argument will be used again below where this is not the case.

with some positive constant  $C_3$ .

Next, we consider the situation that right-hand sides are random but technology matrices are non-random. Then we have  $r = r' < \infty$  and analogously we obtain

$$\begin{aligned} \|z(\tilde{\xi}, \hat{x}) - z(\tilde{\xi}, \tilde{x})\|_p &\leq C_2 \sum_{t=1}^T \left( \mathbb{E} \left[ \left| \mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)] - \tilde{x}_t \right|^{r'} \right] \right)^{\frac{1}{r'}} \\ &\leq C_2 \sum_{t=1}^T l_t \left( \mathbb{E} \left[ \left| \sum_{\tau=2}^t \mathbb{E}[\xi_\tau - \tilde{\xi}_\tau | \sigma(\tilde{\xi}^\tau)] \right. \right. \right. \\ &\quad \left. \left. \left. + \sum_{\tau=2}^{t-1} \mathbb{E} \left[ \left| \bar{x}_\tau - \mathbb{E}[\bar{x}_\tau | \sigma(\tilde{\xi}^\tau)] \right| \left| \sigma(\tilde{\xi}^{\tau+1}) \right| \right] \right| \right] \right)^{\frac{1}{r'}} \\ &\leq C_4 \left( \|\xi - \tilde{\xi}\|_r + \sum_{t=1}^T \left( \mathbb{E} \left[ \left| \bar{x}_t - \mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)] \right|^{r'} \right] \right)^{\frac{1}{r'}} \right) \end{aligned}$$

with some constant  $C_4$ .

Finally, we consider the case that the technology matrices are random and  $r = Tp < r' = \infty$ . Then, however, we need to start at the point *before* Hölder's inequality was applied and obtain

$$\begin{aligned} \|z(\tilde{\xi}, \hat{x}) - z(\tilde{\xi}, \tilde{x})\|_p &\leq J \sum_{t=1}^T \left( \mathbb{E} \left[ \left| b_t(\tilde{\xi}_t)^p \left| \mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)] - \tilde{x}_t \right|^p \right] \right)^{\frac{1}{p}} \\ &\leq J \sum_{t=1}^T l_t \left( \mathbb{E} \left[ \left| b_t(\tilde{\xi}_t)^p \left| \sum_{\tau=2}^t \max\{1, |\tilde{\xi}^t|^{t-\tau}\} \mathbb{E}[\xi_\tau - \tilde{\xi}_\tau | \sigma(\tilde{\xi}^\tau)] \right. \right. \right. \\ &\quad \left. \left. \left. + \sum_{\tau=2}^{t-1} \max\{1, |\tilde{\xi}^t|^{t-\tau}\} \mathbb{E} \left[ \left| \bar{x}_\tau - \mathbb{E}[\bar{x}_\tau | \sigma(\tilde{\xi}^\tau)] \right| \left| \sigma(\tilde{\xi}^{\tau+1}) \right| \right]^p \right] \right)^{\frac{1}{p}} \\ &\leq C_5 \left( \|\xi - \tilde{\xi}\|_r + \sum_{t=1}^T \|\bar{x}_t - \mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)]\|_\infty \right) \end{aligned}$$

with a constant  $C_5$  depending on  $b_t(\cdot)$ ,  $\|\xi\|_r$ , and  $\delta^r$ .

Hence, in all cases we can bound each of the three summands on the right-hand of (20) suitably, i.e., in each case there is a constant  $C$  such that

$$\|z(\xi, \bar{x}) - z(\tilde{\xi}, \tilde{x})\|_p \leq C \left( \|\xi - \tilde{\xi}\|_r + \sum_{t=1}^T \|\bar{x}_t - \mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)]\|_{r'} \right)$$

holds for each  $\bar{x} \in l_\varepsilon(F(\xi, \cdot))$  (and  $\tilde{x}$  constructed appropriately). Hence, we can continue (19) as follows:

$$\begin{aligned} v(\tilde{\xi}) - v(\xi) &\leq \\ &\bar{K} \left( \|\xi - \tilde{\xi}\|_r + \sum_{t=2}^{T-1} \|\bar{x}_t - \mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)]\|_{r'} + \sum_{j=1}^{J-1} \|\bar{y}_j - \mathbb{E}[\bar{y}_j | \sigma(\tilde{\xi}^{t_j})]\|_p \right) + 2\varepsilon \end{aligned} \tag{21}$$

with some positive constant  $\bar{K}$ . The estimate is valid for any  $\bar{x} \in l_\varepsilon(F(\xi, \cdot))$  and any  $\bar{y} \in l_{\rho, \varepsilon}(z(\xi, \bar{x}), \xi)$  and does no longer depend on  $\tilde{x}$ . Changing the role

of  $\xi$  and  $\tilde{\xi}$  yields another constant  $\tilde{K}$  such that

$$v(\xi) - v(\tilde{\xi}) \leq \tilde{K} \left( \|\xi - \tilde{\xi}\|_r + \sum_{t=2}^{T-1} \|\mathbb{E}[\tilde{x}_t | \sigma(\xi^t)] - \tilde{x}_t\|_{r'} + \sum_{j=1}^{J-1} \|\mathbb{E}[\tilde{y}_j | \sigma(\xi^{t_j})] - \tilde{y}_j\|_p \right) + 2\varepsilon \quad (22)$$

for any  $\tilde{x} \in l_\varepsilon(F(\tilde{\xi}, \cdot))$  and  $\tilde{y} \in l_{\rho, \varepsilon}(z(\tilde{\xi}, \tilde{x}), \tilde{\xi})$ . We note that the second and third summands in (21) and (22) are bounded by

$$\begin{aligned} & \sum_{t=2}^{T-1} \max\{\|\mathbb{E}[\bar{x}_t | \sigma(\tilde{\xi}^t)] - \bar{x}_t\|_{r'}, \|\mathbb{E}[\tilde{x}_t | \sigma(\xi^t)] - \tilde{x}_t\|_{r'}\} \\ & \sum_{j=1}^{J-1} \max\{\|\mathbb{E}[\bar{y}_j | \sigma(\tilde{\xi}^{t_j})] - \bar{y}_j\|_p, \|\mathbb{E}[\tilde{y}_j | \sigma(\xi^{t_j})] - \tilde{y}_j\|_p\} \end{aligned}$$

and this leads directly to

$$|v(\xi) - v(\tilde{\xi})| \leq K \left( \|\xi - \tilde{\xi}\|_r + D_{f, \rho}(\xi, \tilde{\xi}) \right) + 2\varepsilon$$

with  $K := \max\{\bar{K}, \tilde{K}\}$ . Finally, it remains to take the the infimum of the right-hand side with respect to  $\varepsilon > 0$  and the proof is complete.  $\square$

*Remark 1* The filtration distance  $D_{f, \rho}$  depends on the  $\varepsilon$ -level-sets, i.e., on the solution behaviour of the problem, which is typically unknown in practice. The question arises, whether it can be estimated by objects that are easier to compute. In particular, for making use of Theorem 3.2 for scenario tree approximation of  $\xi$  this question becomes important. For the scenario tree generation procedure described in [8], such an upper bound of  $D_{f, \rho}$  is derived. Namely, if (A3) is satisfied and the set

$$\bigcup_{\tilde{x} \in l_\varepsilon(F(\tilde{\xi}, \cdot)), \|\tilde{\xi} - \xi\|_r \leq \delta} l_{\rho, \varepsilon}(z(\tilde{\xi}, \tilde{x}), \tilde{\xi}) \quad (23)$$

is bounded in  $L_p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^J)$  for some  $\varepsilon > 0$ , then there exists a constant  $C > 0$  such that

$$\begin{aligned} D_{f, \rho}(\xi, \tilde{\xi}) & \leq C \left( \sup_{\|x\|_{r'} \leq 1} \sum_{t=2}^{T-1} \|\mathbb{E}[x_t | \sigma(\xi^t)] - \mathbb{E}[x_t | \sigma(\tilde{\xi}^t)]\|_{r'} \right. \\ & \quad \left. + \sup_{\|y\|_p \leq 1} \sum_{j=1}^{J-1} \|\mathbb{E}[y_j | \sigma(\xi^{t_j})] - \mathbb{E}[y_j | \sigma(\tilde{\xi}^{t_j})]\|_p \right) \\ & \leq \hat{C} \sup_{\|x\|_{r'} \leq 1} \sum_{t=2}^{T-1} \|\mathbb{E}[x_t | \sigma(\xi^t)] - \mathbb{E}[x_t | \sigma(\tilde{\xi}^t)]\|_{r'} \end{aligned}$$

where the second estimate holds with some other constant  $\hat{C}$  because  $r' \geq p$  and  $J \leq T - 1$ . The upper bound for  $D_{f, \rho}(\xi, \tilde{\xi})$  represents a distance measure

for the filtrations of  $\xi$  and its perturbation  $\tilde{\xi}$  and does not depend on the particular problem.

The level sets  $l_\varepsilon(F(\xi, \cdot))$  and  $l_\varepsilon(F(\tilde{\xi}, \cdot))$  are bounded in  $L_{r'}(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^m)$  due to condition (A3) (e.g., if the sets  $X_t$  are bounded for  $t = 1, \dots, T$ ). However, the corresponding level sets  $l_{\rho, \varepsilon}(z(\xi, x), \xi)$  and  $l_{\rho, \varepsilon}(z(\tilde{\xi}, \tilde{x}), \tilde{\xi})$  of the polyhedral risk measure and, hence, (23) may be unbounded in  $L_p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^J)$ , since the sets  $Y_j$  are assumed to be cones, i.e., unbounded. By definition of the elements  $z(\tilde{\xi}, \tilde{x})$  in  $L_p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^J)$ , the pairs  $(z(\tilde{\xi}, \tilde{x}), \tilde{\xi})$  in (23) vary in a bounded subset of  $\mathbb{Z}$  if (A3) is satisfied. Hence, it remains to clarify the question, under what conditions the level sets of the polyhedral risk measures are uniformly bounded over bounded subsets of  $\mathbb{Z}$ .

#### 4 Level-sets of polyhedral risk measures

As just motivated in the above remark, it is of interest for the stability analysis to know, whether the sets of  $\varepsilon$ -optimal solutions is bounded uniformly on bounded subsets  $Z \subseteq \mathbb{Z}$ . However, the following example shows that, for  $p > 1$ , the level sets, even for a single element  $(z, \xi) \in \mathbb{Z}$ , are typically unbounded.

**Example 4.1** Consider the Average Value-at-Risk at level  $\alpha \in (0, 1)$  ( $\text{AVaR}_\alpha$ , cf. Example 2.1) and let  $z \in L_p(\Omega, \mathcal{F}, \mathbb{P})$  with some  $p \in [1, \infty]$ . Due to the results in [18] it is known that the solution set of (2) is given by

$$S_{\text{AVaR}_\alpha}(z) = \{(y_0, (z + y_0)^+, (z + y_0)^-) : y_0 \in [-\bar{q}_\alpha(z), -q_\alpha(z)]\}$$

with  $\bar{q}_\alpha(z) = \inf\{a \in \mathbb{R} : \mathbb{P}(z \leq a) > \alpha\}$  and  $q_\alpha(z) = \inf\{a \in \mathbb{R} : \mathbb{P}(z \leq a) \geq \alpha\}$  denoting the upper and lower quantile of the distribution of  $z$ , respectively. Hence, since the interval  $[-\bar{q}_\alpha(z), -q_\alpha(z)]$  is always compact, the solution set  $S_{\text{AVaR}_\alpha}(z)$  is bounded in  $L_p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^3)$ .

However, things are different for the level sets  $l_{\text{AVaR}_\alpha, \varepsilon}(z)$  for  $\varepsilon > 0$ . Suppose the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  is rich enough such that there exist  $A_n \in \mathcal{F}$  such that  $\mathbb{P}(A_n) = \frac{1}{n}$ . Consider

$$y^{(n)}(z) := \left( -\bar{q}_\alpha, (z - \bar{q}_\alpha(z))^+ + \varepsilon n \mathbf{1}_{A_n}, (z - \bar{q}_\alpha(z))^- + \varepsilon n \mathbf{1}_{A_n} \right)$$

for  $n \in \mathbb{N}$ . Obviously  $y^{(n)}(z) \in \mathcal{Y}_{\text{AVaR}_\alpha}(z)$ , i.e.,  $y^{(n)}(z)$  is feasible, and  $F_{\text{AVaR}_\alpha}(y^{(n)}(z)) = \rho(z) + \varepsilon$ , i.e.,  $y^{(n)}(z) \in l_{\text{AVaR}_\alpha, \varepsilon}(z)$ . But even if we assume  $z \in L_\infty(\Omega, \mathcal{F}, \mathbb{P})$  we have that  $\|y^{(n)}(z)\|_p \sim n^{1-\frac{1}{p}} \rightarrow \infty$  for  $p \in (1, \infty]$ , i.e., the level set  $l_{\text{AVaR}_\alpha, \varepsilon}(z)$  for a single random variable  $z$  is unbounded in  $L_p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^3)$  for  $p > 1$ . Thus, for the boundedness of the  $\text{AVaR}_\alpha$  level sets,

there is only hope for  $p = 1$ . It will be seen below that  $l_{\text{AVaR}_{\alpha,\varepsilon}}(z)$  is bounded in  $L_1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^3)$  indeed, actually in a uniform manner.

Since the multiperiod polyhedral risk measures (cf. section 2.2) from [3] boil down to AVaR if it is set  $J = 1$ , and, hence, their level sets are unbounded in  $L_p$  if  $p > 1$ , we will assume  $p = 1$  from now on (and accordingly  $p' = \infty$ ). In the following, a simple criterion will be derived which guarantees the sort of uniform  $L_1$  boundedness of the level sets  $l_\rho(z, \xi)$  as it is required in Remark 1 in section 3. This criterion, though appearing to be very specific, applies for most of the polyhedral risk measures  $\rho$  introduced so far. Here, the extended real-valued function  $\Phi_\rho$ , called the *value function* given by

$$\Phi_\rho(y_0, z, \xi) := \inf_{y_1, \dots, y_J} \{F_\rho(y_0, y_1, \dots, y_J) : (y_0, y_1, \dots, y_J) \in \mathcal{Y}_\rho(z, \xi)\}$$

will be used. Observe that  $\rho(z, \xi) = \inf_{y_0 \in Y_0} \Phi_\rho(y_0, z, \xi)$ . The letter  $\pi_j$  will denote the projection to the  $j$ th component.

**PROPOSITION 4.2** *Let  $\rho$  be a functional of the form (2) satisfying complete recourse and dual feasibility and assume*

- (i)  $k_j = 2$ ,  $\langle c_j, Y_j \rangle \subseteq \mathbb{R}_+$  for  $j = 1, \dots, J$ ,
- (ii) the vectors  $c_j$  and  $w_{j,0}$  are linearly independent for  $j = 1, \dots, J$ ,
- (iii)  $\pi_j \left( \bigcap_{\nu=j}^J \mathcal{D}_{\rho,\nu} \right)$  is bounded in  $\mathbb{R}$  for  $j = 1, \dots, J$ , and
- (iv)  $Y_0$  is bounded, or alternatively
- (iv')  $k_0 = 1$ ,  $c_0 > 0$ , and  $\inf \left\{ \sum_{j=1}^J u_j w_{j,j} : u \in \bigcap_{j=1}^J \mathcal{D}_{\rho,j} \right\} < -c_0$ .

Let  $Z \subseteq \mathbb{Z}$  such that the projection  $\pi_1(Z)$  to the  $z$  component is bounded in  $L_1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^J)$ . Then the union over all level sets  $\bigcup_{(z,\xi) \in Z} l_{\rho,\varepsilon}(z, \xi)$  is bounded in  $L_1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{\sum k_j})$  for  $\varepsilon > 0$ .

*Proof* First of all, consider the numbers  $M_Z := \sup\{\|z\|_1 : (z, \xi) \in Z\}$ ,  $M_D := \sup\{\|u\|_\infty : u \in \bigcap_{j=1}^J \mathcal{D}_{\rho,j}\}$ , and  $M_\rho := \sup\{|\rho(z)| : (z, \xi) \in Z\}$ . Observe that  $M_Z < \infty$  according to the assumptions about  $Z$  and that  $M_D < \infty$  due to assumption (iii). First, we show that also  $M_\rho < \infty$ . To this end, consider the dual representation (7) and note that due to assumption (iii) the feasible set

$$\Lambda_\rho(\xi) := \left\{ \lambda \in \times_{j=1}^J L_\infty(\Omega, \sigma(\xi^{t_j}), \mathbb{P}) : \mathbb{E}[\lambda | \xi^{t_j}] \in \mathcal{D}_{\rho,j} \text{ a.s. } (j = 1, \dots, J) \right\}$$

is bounded in  $L_\infty(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^J)$  with a bound  $M_\lambda$  not depending on  $\xi$ . Hence,

(7) yields the following estimates:

$$\begin{aligned} \rho(z, \xi) &\leq \sup_{\lambda \in \Lambda_\rho(\xi)} \inf_{y_0 \in Y_0} \left\langle c_0 + \sum_{\nu=1}^J \mathbb{E}[\lambda_\nu] w_{\nu, \nu}, y_0 \right\rangle + M_Z M_\lambda \\ \rho(z, \xi) &\geq \sup_{\lambda \in \Lambda_\rho(\xi)} \inf_{y_0 \in Y_0} \left\langle c_0 + \sum_{\nu=1}^J \mathbb{E}[\lambda_\nu] w_{\nu, \nu}, y_0 \right\rangle - M_Z M_\lambda \end{aligned}$$

and since

$$\begin{aligned} &\sup_{\lambda \in \Lambda_\rho(\xi)} \inf_{y_0 \in Y_0} \left\langle c_0 + \sum_{\nu=1}^J \mathbb{E}[\lambda_\nu] w_{\nu, \nu}, y_0 \right\rangle \\ &= \sup_{u \in \bigcap_{\nu=1}^J \mathcal{D}_{\rho, \nu}} \inf_{y_0 \in Y_0} \left\langle c_0 + \sum_{\nu=1}^J u_\nu w_{\nu, \nu}, y_0 \right\rangle \end{aligned}$$

it becomes clear that this number, which does not depend on  $(z, \xi)$ , must be finite (otherwise  $\rho(z, \xi)$  would be infinite). Hence,  $M_\rho$  is finite indeed.

Now, let  $\varepsilon > 0$ . We prove boundedness of the level sets for each component  $j = 0, 1, \dots, J$  successively. For  $j = 0$ , we show that, if  $Y_0$  is unbounded, the value function  $\Phi_\rho(y_0, z, \xi)$  grows to infinity uniformly on  $Z$  as  $|y_0| \rightarrow \infty$ . For  $y_0 \rightarrow +\infty$  this is obvious since  $\Phi_\rho(y_0, z, \xi) \geq c_0 y_0$  due to assumption (i) and  $c_0 y_0 \rightarrow \infty$  due to assumption (iv'). For  $y_0 < 0$ , we obtain the following estimate by making use of [19, Theorem 14.60] and LP duality [19, Example 11.43]:

$$\begin{aligned} &\Phi_\rho(y_0, z, \xi) \\ &= c_0 y_0 + \inf \left\{ \mathbb{E} \left[ \sum_{j=1}^J \langle c_j, y_j \rangle \right] \left| \begin{array}{l} y \in \times_{j=1}^J L_1(\Omega, \sigma(\xi^{t_j}), \mathbb{P}; \mathbb{R}^{k_j}), \\ y \in \times_{j=1}^J Y_j \text{ a.s.}, \\ \sum_{\tau=0}^{j-1} \langle w_{j, \tau}, y_{j-\tau} \rangle = z_{t_j} - w_{j, j} y_0 \text{ a.s.} \end{array} \right. \right\} \\ &\geq c_0 y_0 + \inf \left\{ \mathbb{E} \left[ \sum_{j=1}^J \langle c_j, y_j \rangle \right] \left| \begin{array}{l} y \in \times_{j=1}^J L_1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{k_j}), \\ y \in \times_{j=1}^J Y_j \text{ a.s.}, \\ \sum_{\tau=0}^{j-1} \langle w_{j, \tau}, y_{j-\tau} \rangle = z_{t_j} - w_{j, j} y_0 \text{ a.s.} \end{array} \right. \right\} \\ &= c_0 y_0 + \mathbb{E} \left[ \inf \left\{ \sum_{j=1}^J \langle c_j, y_j \rangle \left| \begin{array}{l} y \in \times_{j=1}^J Y_j, \\ \sum_{\tau=0}^{j-1} \langle w_{j, \tau}, y_{j-\tau} \rangle = z_{t_j} - w_{j, j} y_0 \end{array} \right. \right\} \right] \\ &= c_0 y_0 + \mathbb{E} \left[ \sup \left\{ \sum_{j=1}^J u_j (w_{j, j} y_0 - z_{t_j}) \mid u = (u_1, \dots, u_J) \in \bigcap_{j=1}^J \mathcal{D}_{\rho, j} \right\} \right] \\ &\geq c_0 y_0 + y_0 \inf \left\{ \sum_{j=1}^J u_j w_{j, j} \mid u \in \bigcap_{j=1}^J \mathcal{D}_{\rho, j} \right\} - M_Z M_D, \end{aligned}$$

thus  $\Phi_\rho(y_0, z, \xi) \rightarrow +\infty$  as  $y_0 \rightarrow -\infty$  due to assumption (iv'). Hence, there is a real number  $M_0$  such that for all  $(z, \xi) \in Z$  and for all  $\hat{y} = (\hat{y}_0, \dots, \hat{y}_J) \in l_{\rho, \varepsilon}(z, \xi)$  it holds that  $|\hat{y}_0| \leq M_0$ . Now, for  $j = 1$  it holds due to assumption

Table 3. Details to verify condition (iv') of Proposition 4.2 for some polyhedral risk measures when  $\alpha_j$  is set to  $\alpha < 1$  for  $j = 1, \dots, J$ .

riskm.	$\bigcap_{j=1}^J \mathcal{D}_{\rho,j}$	$c_0$	$w_{1,1}$	$w_{j,j}$ ( $j > 1$ )	$\inf_{u \in \mathcal{D}_{\rho,j}} \sum u_j w_{j,j}$
AVaR $_{\alpha}$	$[0, \frac{1}{\alpha}]$	1	-1		$-\frac{1}{\alpha}$
$\rho_2$	$\times_{j=1}^J [0, \frac{1}{\alpha}]$	1	-1	-1	$-\frac{1}{\alpha}$
$\rho_3$	$\{u \in \mathbb{R}_+^J : u_J \leq \frac{1}{J\alpha}, u_j + u_{j+1} \leq \frac{1}{J\alpha}\}$	1	-1	-1	$-\frac{1}{2\alpha}$
$\rho_4$	$\{u \in \mathbb{R}_+^J : u_1 = \dots = u_J \leq \frac{1}{J\alpha}\}$	$\frac{1}{J}$	-1	0	$-\frac{1}{J\alpha}$

(i) that

$$\begin{aligned} \|\langle c_1, \hat{y}_1 \rangle\|_1 &= \mathbb{E}[\langle c_1, \hat{y}_1 \rangle] \leq \sum_{j=1}^J \mathbb{E}[\langle c_j, \hat{y}_j \rangle] = F_{\rho}(\hat{y}) - \langle c_0, \hat{y}_0 \rangle \\ &\leq \rho(z_{t_1}, \dots, z_{t_J}) + \varepsilon + |c_0| \cdot M_0 \leq M_{\rho} + \varepsilon + |c_0| M_0 \\ \|\langle w_{1,0}, \hat{y}_1 \rangle\|_1 &= \|z_{t_1} - \langle w_{1,1}, \hat{y}_0 \rangle\|_1 \\ &\leq \|z_{t_1}\|_1 + |w_{1,1}| M_0 \leq M_Z + |w_{1,1}| M_0, \end{aligned}$$

i.e.,  $(c_1, w_{1,0})' \hat{y}_1$  is bounded in  $L_1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^2)$  by a number that does not depend on  $(\xi, z)$ . The  $2 \times 2$  matrix  $(c_1, w_{1,0})$  is regular due to assumption (ii), hence,  $\hat{y}_1$  is  $L_1$  bounded. By induction we conclude analogously for  $j > 1$ .  $\square$

This proposition applies directly to the exemplary polyhedral risk measures  $EL_{\beta}$ , AVaR $_{\alpha}$ ,  $\rho_2$ ,  $\rho_3$ , and  $\rho_4$  as far as, say,  $\alpha_j = \alpha < 0.5$ , cf. Table 2 and 3. Moreover, uniform level boundedness of risk measure  $\rho_1$  is guaranteed, too, since its level sets can be understood as a Cartesian product of level sets of AVaR $_{\alpha_j}$ .

**Example 4.3** Of course, since Proposition 4.2 appears rather technical and all the examples from section 2.2 have bounded level sets in  $L_1$ , the question arises, whether there exist any polyhedral risk measures on  $L_1$  satisfying complete recourse and dual feasibility that have unbounded level sets. The answer can be given directly: Consider AVaR $_1$ , i.e., the Average Value-at-Risk (cf. Example 2.1) at level  $\alpha = 1$  (note that, typically,  $\alpha < 1$  is assumed), which still satisfies complete recourse and dual feasibility but not condition (iv) or (iv') of Proposition 4.2. For  $z \equiv 0$ , formula (5) reveals that the  $y_0$  component of the solution set is given by  $\pi_0(S_{AVaR_1}(0, \xi)) = \mathbb{R}^-$ , i.e., it is unbounded in  $\mathbb{R}$ . Hence,  $S_{AVaR_1}(0, \xi)$  and thus  $l_{AVaR_1, \varepsilon}(0, \xi)$  are unbounded in  $L_1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^3)$ . We conclude that complete recourse and dual feasibility are *not* sufficient conditions for bounded level sets.

**Example 4.4** Regrettably, Proposition 4.2 does not apply for the value of perfect information based risk measure (6) from [13, 14], cf. example 2.4, since  $k_j \neq 2$ , i.e., condition (i) is not satisfied. However, it has been observed in [14]

that the risk measure decomposes into functionals for each time period:

$$\rho_5(z) = \sum_{j=1}^J \mathbb{E} \left[ -s_j z_{t_j} + (s_{j-1} - s_j) \text{AVaR}_{\alpha_j}(z_{t_j} | \mathcal{F}_{t_{j-1}}) \right]$$

where  $\alpha_j = (s_{j-1} - s_j)/(q_j - s_j)$ . This decomposition will simplify the analysis of the level sets drastically. In [14] it has been derived via the dual representation (8) of  $\rho_5$ , but it can be deduced directly from (6) by making use of the dynamic constraints  $z_{t_j} = y_{j,2} - y_{j,3} - y_{j-1,2} + y_{j-1,1}$ :

$$\begin{aligned} & \rho_5(z) + \sum_{j=1}^J s_j \mathbb{E} [z_{t_j}] \\ = \inf & \left\{ \begin{array}{l} -s_0 y_{0,1} + \mathbb{E} \left[ \sum_{j=1}^J (-s_j y_{j,1} + q_j y_{j,3}) - d y_{j,2} \right] \\ + \mathbb{E} \left[ \sum_{j=1}^J s_j (y_{j,2} - y_{j,3} - y_{j-1,2} + y_{j-1,1}) \right] : \\ y_j \in L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P}; \mathbb{R}^3) \ (j = 0, \dots, J), \ y_{0,2} = y_{0,3} = y_{J,1} = 0, \\ y_{j,2} \geq 0 \text{ a.s.}, \ y_{j,3} \geq 0 \text{ a.s.}, \ y_{j,2} - y_{j,3} = y_{j-1,2} + z_{t_j} - y_{j-1,1} \text{ a.s.} \\ (j = 1, \dots, J) \end{array} \right\} \\ = \inf & \left\{ \begin{array}{l} -y_{0,1}(s_0 - s_1) \\ + \mathbb{E} \left[ \sum_{j=1}^{J-1} (s_j - s_{j+1})(y_{j,2} - y_{j,1}) + \sum_{j=1}^J (q_j - s_j) y_{j,3} \right] : \\ y_j \in L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P}; \mathbb{R}^3) \ (j = 0, \dots, J), \ y_{0,2} = y_{0,3} = y_{J,1} = 0, \\ y_{j,2} \geq 0 \text{ a.s.}, \ y_{j,3} \geq 0 \text{ a.s.}, \ y_{j,2} - y_{j,3} = y_{j-1,2} + z_{t_j} - y_{j-1,1} \text{ a.s.} \\ (j = 1, \dots, J) \end{array} \right\} \end{aligned}$$

where it is set  $s_J := d$  for convenience. Substituting  $\tilde{y}_{j,1} := y_{j,2} - y_{j,1}$  yields immediately

$$\begin{aligned} \rho_5(z) &= \sum_{j=1}^J ((s_{j-1} - s_j) \rho_{5,j}(z_{t_j}) - s_j \mathbb{E}[z_{t_j}]) \text{ resp.} \\ \mathcal{R}(z) &= \sum_{j=1}^J (s_{j-1} - s_j) (\rho_{5,j}(z_{t_j}) + \mathbb{E}[z_{t_j}]) \end{aligned}$$

with

$$\rho_{5,j}(z_{t_j}) = \inf \left\{ \mathbb{E} \left[ \tilde{y}_{j-1,1} + \frac{q_j - s_j}{s_{j-1} - s_j} y_{j,3} \right] \left| \begin{array}{l} \tilde{y}_{j-1,1} \in L_p(\Omega, \mathcal{F}_{t_{j-1}}, \mathbb{P}), \\ y_{j,2}, y_{j,3} \in L_p(\Omega, \mathcal{F}_{t_j}, \mathbb{P}), \\ y_{j,2} \geq 0 \text{ a.s.}, \ y_{j,3} \geq 0 \text{ a.s.}, \\ y_{j,2} - y_{j,3} = z_{t_j} + \tilde{y}_{j-1,1} \text{ a.s.} \end{array} \right. \right\} \quad (24)$$

for  $j = 1, \dots, J$ . Interchanging minimisation and integration can give the above interpretation  $\rho_{5,j}(z_{t_j}) = \mathbb{E}[\text{AVaR}_{\alpha_j}(z_{t_j} | \mathcal{F}_{t_{j-1}})]$  from [14].

**PROPOSITION 4.5** *Let  $Z \subseteq \mathbb{Z}$  such that the projection  $\pi_1(Z)$  to the  $z$  component is bounded in  $L_1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^J)$ . Then, for the risk measure  $\rho_5$  in (6), it holds that the union over all  $\varepsilon$ -level sets  $\bigcup_{(z, \xi) \in Z} l_{\rho_5, \varepsilon}(z, \xi)$  is bounded in  $L_1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{3(J+1)})$  for  $\varepsilon > 0$ .*

*Proof* We show that for  $\varepsilon > 0$  and for each  $j = 1, \dots, J$  the union  $\bigcup_{(z, \xi) \in Z} l_{\rho_{5,j}, \varepsilon}(z_{t_j}, \xi)$  of all  $\varepsilon$ -level sets of  $\rho_{5,j}(z_{t_j}, \xi)$ , cf. (24), is bounded in  $L_1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^3)$ . To this end, we first note that the number  $M_{\rho_{5,j}} := \sup\{|\rho_{5,j}(z, \xi)| : (z, \xi) \in Z\}$  is finite. This can easily be seen by considering the dual of (24) given by

$$\rho_{5,j}(z, \xi) = \sup \left\{ -\mathbb{E}[\lambda_j z_{t_j}] \left| \begin{array}{l} \lambda_j \in L_p(\Omega, \sigma(\xi^{t_j}), \mathbb{P}), \\ 0 \leq \lambda_j \leq \frac{q_j - s_j}{s_{j-1} - s_j} \text{ a.s., } \mathbb{E}[\lambda_j | \xi^{t_{j-1}}] = 1 \text{ a.s.} \end{array} \right. \right\}.$$

Now, let  $(z^{(n)}, \xi^{(n)}) \in Z$  and  $y^{(n)} = (\tilde{y}_{j-1,1}^{(n)}, y_{j,2}^{(n)}, y_{j,3}^{(n)}) \in l_{\rho_{5,j}, \varepsilon}(z_{t_j}^{(n)}, \xi^{(n)})$  for  $n \in \mathbb{N}$ . Suppose there is a subsequence  $(y^{(n_k)})$  such that  $\|(\tilde{y}_{j-1,1}^{(n_k)})^-\|_1 \rightarrow \infty$ . In this case, the following estimate for the objective of (24) would hold:

$$\begin{aligned} F_{\rho_{5,j}}(y^{(n_k)}) &= \mathbb{E} \left[ \tilde{y}_{j-1,1}^{(n_k)} + \frac{1}{\alpha_j} y_{j,3}^{(n_k)} \right] \geq \mathbb{E} \left[ \mathbf{1}_{\{\tilde{y}_{j-1,1}^{(n_k)} \leq 0\}} \left( \tilde{y}_{j-1,1}^{(n_k)} + \frac{1}{\alpha_j} y_{j,3}^{(n_k)} \right) \right] \\ &= \mathbb{E} \left[ \mathbf{1}_{\{\tilde{y}_{j-1,1}^{(n_k)} \leq 0\}} \left( \tilde{y}_{j-1,1}^{(n_k)} + \frac{1}{\alpha_j} (y_{j,2}^{(n_k)} - \tilde{y}_{j-1,1}^{(n_k)} - z^{(n_k)}) \right) \right] \\ &= \left( \frac{1}{\alpha_j} - 1 \right) \|(\tilde{y}_{j-1,1}^{(n_k)})^-\|_1 + \frac{1}{\alpha_j} \mathbb{E} \left[ \mathbf{1}_{\{\tilde{y}_{j-1,1}^{(n_k)} \leq 0\}} \left( y_{j,2}^{(n_k)} - z^{(n_k)} \right) \right] \rightarrow \infty. \end{aligned}$$

The convergence to infinity holds because  $1/\alpha_j > 1$ ,  $y_{j,2}^{(n_k)} \geq 0$ , and because the sequence  $(z^{(n)})$  is  $L_1$ -bounded. However,  $F_{\rho_{5,j}}(y^{(n_k)}) \rightarrow \infty$  is a contradiction to  $y^{(n)} \in l_{\rho_{5,j}, \varepsilon}(z^{(n)}, \xi^{(n)})$  since the sequence  $(\rho_{5,j}(z^{(n)}, \xi^{(n)}))$  is bounded due to  $M_{\rho_{5,j}} < \infty$ . Hence, the sequence  $((\tilde{y}_{j-1,1}^{(n)})^-)$  is  $L_1$ -bounded. Suppose there is a subsequence  $(y^{(n_k)})$  such that  $\|(\tilde{y}_{j-1,1}^{(n_k)})^+\|_1 \rightarrow \infty$ . Obviously, this would also imply  $F_{\rho_{5,j}}(y^{(n_k)}) \rightarrow \infty$  since  $y_{j,3}^{(n_k)} \geq 0$  and thus cause a contradiction. Hence, the sequence  $(\tilde{y}_{j-1,1}^{(n)})$  is  $L_1$ -bounded. Also the existence of subsequences  $(y^{(n_k)})$  such that  $\|y_{j,2}^{(n_k)}\|_1 \rightarrow \infty$  or  $\|y_{j,3}^{(n_k)}\|_1 \rightarrow \infty$  would cause a contradiction in the same manner. Hence, the overall sequence  $(y^{(n)})$  is  $L_1$ -bounded. That is, the union over all level sets of  $\rho_{5,j}$  is indeed bounded in  $L_1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^3)$ . Finally, note that this boundedness for  $\rho_{5,j}$  implies uniform boundedness of the  $\varepsilon$ -level sets of (6) for  $\rho_5$  because the substitution  $(y_1, y_2) \mapsto (y_2 - y_1, y_2)$  in example 4.4 is bijective in  $\mathbb{R} \times \mathbb{R}_+$ .  $\square$

## 5 Conclusion for stability and scenario tree approximation

In [3], the class of polyhedral risk measures has been suggested. As discussed in section 2.2, replacing the expectation in (1) by a (multiperiod) polyhedral risk measure yields problem (3). The equivalent problem (4) of (3) has an expect-

tation objective and is of a similar form (but not the same) as (1) with additional stochastic decision variables  $y_j$  and additional constraints. The stability theorem from [10], however, does not hold. Here, we provided an equivalent stability theorem (Theorem 3.2) for problem (3) based on sensitivity analysis for polyhedral risk measures (Proposition 3.1).

Stability according to Theorem 3.2 involves so-called filtration distances which are based on the sets of  $\varepsilon$ -optimal solutions (level sets) of the underlying problem. In order to make use of Theorem 3.2 in the context of scenario tree approximation, it turns out to be necessary to have these level sets bounded; cf. Remark 1 in section 3, see also [8]. However, though in many application the original decision variables can be assumed to be bounded from the outset, the additional  $y_j$  variables arising from the polyhedral risk measures are inherently unbounded in terms of feasibility. For this reason, criteria for the boundedness of the  $y_j$  components of the level sets are derived in section 4; in particular, it has been stated that boundedness is guaranteed for all the instances of the class of polyhedral risk measures from [3, 14] if the integrability number  $p$  of the risk measure arguments is set to 1.

As in [10], Theorem 3.2 makes several restrictions for the integrability number  $r$  of the stochastic input process  $\xi$ . At the first glance there seem to be more degrees of freedom for  $r$  than in [10] since, theoretically,  $p$  may be chosen arbitrarily. But, as mentioned above, in the context of scenario approximation  $p = 1$  is the only choice. Then, however, the situation is the same as in [10].

To conclude, by means of the present paper the results from [8, 10] apply to problem (3) where  $\mathbb{E}$  is replaced by a polyhedral risk measure from [3]. In particular, the same scenario approximation techniques can be used as soon as the criteria for the boundedness of the level sets for the polyhedral risk measure are satisfied.

## Acknowledgments

This work was supported by the DFG Research Center MATHEON *Mathematics for Key Technologies* in Berlin (<http://www.matheon.de>).

## References

- [1] Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D., 1999, Coherent measures of risk. *Mathematical Finance*, **9**, 203–228.
- [2] Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., and Ku, H., 2006, Coherent multiperiod risk adjusted values and Bellman’s principle. *Annals of Operations Research*, to appear.
- [3] Eichhorn, A., and Römisch, W., 2005, Polyhedral risk measures in stochastic programming. *SIAM Journal on Optimization*, **16**, 69–95.
- [4] Eichhorn, A., and Römisch, W., 2006, Mean-risk optimization models for electricity portfolio management. *Proceedings of the 9th International Conference on Probabilistic Methods Applied to Power Systems* (PMAPS 2006), Stockholm, Sweden.

- [5] Eichhorn, A., Römisch, W., and Wegner, I., 2005, Mean-risk optimization of electricity portfolios using multiperiod polyhedral risk measures. *IEEE St. Petersburg PowerTech Proceedings*, St. Petersburg, Russia.
- [6] Föllmer, H., and Schied, A., 2002, *Stochastic Finance: An Introduction in Discrete Time* (Berlin: Walter de Gruyter).
- [7] Frittelli, M., and Scandolo, G., 2006 Risk measures and capital requirements for processes. *Mathematical Finance*, **16**, 589–612.
- [8] Heitsch, H., and Römisch, W., 2006, Scenario tree modelling for multistage stochastic programs, Preprint 324, DFG Research Center MATHEON "Mathematics for key technologies", and submitted.
- [9] Heitsch, H., and Römisch, W., 2006, A note on scenario reduction for two-stage stochastic programs, Preprint 335, DFG Research Center MATHEON "Mathematics for key technologies", and submitted to *Operations Research Letters*.
- [10] Heitsch, H., Römisch, W., and Strugarek, C., 2006, Stability of multistage stochastic programs. *SIAM Journal on Optimization*, **17**, 511–525.
- [11] Markowitz, H., 1952, Portfolio selection. *The Journal of Finance*, **7**, 77–91.
- [12] Pflug, G. Ch., 2004, The value of perfect information as a risk measure. In K. Marti, Y. Ermoliev, and G. Ch. Pflug (Eds), *Dynamic Stochastic Optimization*, (Berlin: Springer), pp. 275–291.
- [13] Pflug, G. Ch. and Ruszczyński, A., 2004, Risk measures for income processes. In: G. Szegő (Ed), *Risk Measures for the 21st Century*, (Chichester, UK: Wiley), pp. 249–269.
- [14] Pflug, G. Ch. and Ruszczyński, A., 2005, Measuring Risk for Income Streams. *Computational Optimization and Applications*, **32**, 161–178.
- [15] Ogryczak, W., and Ruszczyński, A., 2001 On consistency of stochastic dominance and mean-semideviation models. *Mathematical Programming*, **89**, 217–232.
- [16] Rachev, S. T., and Römisch, W., 2002, Quantitative stability in stochastic programming: The method of probability metrics. *Mathematics of Operations Research*, **27**, 792–818.
- [17] Riedel, F., 2004, Dynamic coherent risk measures. *Stochastic Processes and their Applications*, **112**, 185–200.
- [18] Rockafellar, R. T., and Uryasev, S., 2002, Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, **26**, 1443–1471.
- [19] Rockafellar, R. T., and Wets, R. J-B, 1998, *Variational Analysis* (1st edn, corr. 2nd printing 2002) (Berlin: Springer).
- [20] Römisch, W., 2002, Stability of stochastic programming problems. Chapter 8 in [21], pp. 483–554.
- [21] Ruszczyński, A., and Shapiro A. (Eds), 2003, *Stochastic Programming* (Amsterdam: Elsevier).
- [22] Ruszczyński, A., and Shapiro A., 2006, Optimization of risk measures, In: G. Calafiore and F. Dabbene (Eds) *Probabilistic and Randomized Methods for Design under Uncertainty* (London: Springer), pp. 119–157.
- [23] Schultz, R., and Tiedemann, S., 2005 Conditional Value-at-Risk in stochastic programs with mixed-integer recourse. *Mathematical Programming*, **105**, 365–386.