SIAM J. Optim.
Vol. 24, No. 3, pp. 993–1010

STOCHASTIC PROGRAMMING OF TIME-CONSISTENT EXTENSIONS OF AVaR*

M. DENSING†

Abstract. We discuss multiperiod stochastic programming formulations of time-consistent extensions of average value-at-risk (AVaR); AVaR measures the risk of a random financial value. Multiperiod risk measures that are recursively defined over time are known to be time consistent. For a multiperiod extension of AVaR for stochastic value processes, we reformulate the recursion as a linear stochastic program, such that the extension can be applied in multiperiod mean-risk optimization. In the special case of risk measurement for a final random value at a time horizon, we give a lower bound in terms of AVaR.

Key words. multistage stochastic linear programming, average value-at-risk, conditional value-at-risk, time consistency, mean-risk optimization

AMS subject classifications. 90C15, 90C90

DOI. 10.1137/130905046

1. Introduction.

Let an uncertain financial loss be given by a random variable. A coherent risk measure is a functional that maps the random variable to a real number and that satisfies the coherency properties as defined in the seminal work of Artzner et al. [3]. Risk measures have the sign-convention that losses are positive. For multiple periods, considering losses as negative (and profit as positive) is more convenient [4, Rem. 2]. Consequently, our convention is to use sign-reversed risk measures, which are called *risk-adjusted values* [4] or *acceptability functionals* [19]. For example, a widely applied coherent risk-adjusted value is average value-at-risk (AVaR) [19]. (The corresponding risk measure is conditional value-at-risk (CVaR) [18, 22].) AVaR can be calculated by a linear optimization problem, which allows AVaR to be incorporated in linear mean-risk optimization problems [5, 14, 17, 22].

Risk-adjusted values can be extended to multiple time periods (or to continuous time); for an introduction to multiperiod and continuous-time risk measurement, see [1, 19, 25]. Multiperiod risk measurement can be for a stochastic process or for only a final random variable at a time horizon. With multiple periods, the risk can also be measured from the viewpoint of a specific state and time, that is, the risk-adjusted value is evaluated conditionally on a specific state and time.

Risk-adjusted values can be used as acceptance criteria in decision problems. Such a criterion may be required to be time consistent: For example, if a final random value has a higher conditionally evaluated risk-adjusted value than a second final random variable in all possible states at a future time, then the (unconditional) risk-adjusted value as of today for the first random variable should be higher than for the second random variable, too. Unfortunately, a conditionally evaluated AVaR for a final value is generally not time consistent. An example is in Figure 1 for a binary three-stage scenario tree. (See also the example on a larger tree in [4, Chap. 5.3].) At the second stage, the conditionally evaluated AVaR of the final random variable X is always

*Received by the editors January 7, 2013; accepted for publication (in revised form) April 30, 2014; published electronically July 10, 2014. Parts of this work were developed at the Institute for Operations Research, University of Zurich, Switzerland.
http://www.siam.org/journals/siopt/24-3/90504.html
†Energy Economics Group, Paul Scherrer Institute, 5232 Villigen PSI, Switzerland (martin.densing@psi.ch).

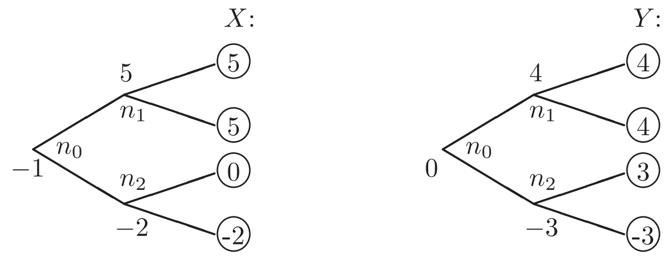

FIG. 1. *Violation of time consistency by conditionally evaluated AVaR on a scenario tree with equiprobable terminal nodes (circles). In this example, AVaR is the arithmetic mean of values below or equal to the* 50% *quantile. Random variables* X *and* Y *take values on the terminal nodes (the tree is duplicated for readability). For example, AVaR of* X *has in node* n_0 *the value* $(-2+0)/2 = -1$*, and in node* n_2 *the value is the lower of* -2 *and* 0*. Today, i.e., in* n_0*, the AVaR of* Y *is higher:* $0 > -1$*; later the AVaR of* X *is higher:* $5 > 4$ *in* n_1*, and* $-2 > -3$ *in* n_2*.*

higher than for the random variable Y, whereas at the first stage, the AVaR of X is lower than for Y. (See also the general Definitions 3.1 and 4.2 of (conditional) AVaR.)

In discrete time and with a finite time horizon, time consistency is ensured by a recursive definition (nested over time) of the risk-adjusted value [4]. Recursive extensions of AVaR have different names in the literature and are in most cases defined as a risk-adjusted value of a final random variable: *nested AVaR* [19], *dynamic AVaR* [6], *composed AVaR* [7], *dynamically consistent TVaR* [23], and *conditional risk mappings* [24].

Our definition of a recursive risk-adjusted value is for stochastic processes and is based on [4]; a risk-adjusted value of final random variables is a special case where intermediate values are neglected. Intermediate values are important for risk measurement if they have some effect; for example, the intermediate value distribution may be used for an intermediate regulatory rating or it may influence intermediate cash flows.

In section 2, we define a risk-adjusted value for stochastic processes by a recursive formula. We show that the recursively defined risk-adjusted value is coherent and time consistent. Then, we choose a particular risk-adjusted value to extend AVaR to multiple periods (section 3). The recursive definition of the extension is reformulated as a stochastic linear program; the proof uses a lemma on strong duality for conditional AVaR, which may be of its own interest (section 4). The linear program formulation allows us to incorporate the risk-adjusted value in mean-risk optimization problems (section 5). For final random variables, the risk-adjusted value is shown to have a lower bound by AVaR, and we provide a dual formulation, which is similar to that of AVaR (section 6). Note that some preliminary work was presented in [10].

Applications of multiperiod risk measurement are relatively scarce. An optimization of a time-consistent risk-adjusted value for a dynamic portfolio of commodities is considered in [13]; a multiperiod mean-risk energy-production problem with so-called polyhedral risk measures is in [11]; a hydrothermal scheduling problem with time consistency is considered in [20]; and a pension fund management problem is in [15].

The risk-adjusted value of a stochastic process as proposed in this paper is applied for an energy-production problem in [9]. Specifically, the proposed extension of AVaR is used in a mean-risk model for the optimal operation of a pump-storage hydropower plant; the intermediate financial value is the sum of realized cash flows and of the expected value of remaining water in the storage reservoir, the time horizon is several months within a year, and the model has monthly time steps. The obtained numerical solutions indicate that the proposed multiperiod risk-adjusted value can be used to reduce the probability of low intermediate values, while low final values are also reduced to an extent that is comparable to alternative model runs where the risk was measured by AVaR only on final values. Furthermore, the increase in computational running time compared to AVaR was found to be negligible. Because the proposed risk measurement is coherent and time consistent, we can ensure these desirable properties for the decision maker of the plant, though we note that there are multiperiod extensions of AVaR that relax or replace time consistency by other suitable properties [11, 16, 23].

The probability space is $(\Omega, \mathcal{F}, \mathbb{P})$. The scenario space Ω is finite. The finiteness allows us to ignore measure-theoretic subtleties; the necessary changes for an extension to infinite spaces are discussed in the conclusion. To keep notation simple, we use general denotations of probability theory and avoid more lengthly formulations with realizations of random variables. A random variable is a measurable function $X\colon \Omega \to \mathbb{R}$, and the expected value under probability measure $\mathbb{P}$ is $\mathbb{E}[X]$. Equalities involving random variables hold almost surely (i.e., on every event for Ω finite). Note that for general Ω, the (pointwise) infimum in subsequent expressions must be generally replaced by the essential infimum.

2. Recursively defined risk-adjusted values. In this section, we revisit the recursive definition of a risk-adjusted value process given in [4]. The properties of coherency and of time consistency are shown with fewer assumptions.

Let an uncertain financial value be represented by a bounded random variable $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$. A *coherent risk-adjusted value* $\pi[X]$ can be given by

$$\pi[X] = \inf_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\mathbb{Q}}[X], \tag{2.1}$$

where $\mathbb{E}_{\mathbb{Q}}[\cdot]$ is the expected value with respect to probability measure $\mathbb{Q}$, and $\mathcal{P}$ is a set of probability measures on $(\Omega, \mathcal{F})$ [4, sect. 1]. The values of the measure $\mathbb{Q}$ are interpreted as *test probabilities*, such that $\pi[X]$ is the worst test result.

In a multiperiod setting, let $t = 0, \ldots, T$ be the time steps. The gain of information is given by a filtration of σ-algebras $(\mathcal{F}_t)_{t=0,\ldots,T}$, with $\mathcal{F}_t \subseteq \mathcal{F}_{t+1}$ for $t = 0, \ldots, T-1$, $\mathcal{F}_0 = \{\emptyset, \Omega\}$, and $\mathcal{F}_T = \mathcal{F}$. A sequence of uncertain financial values is represented by a stochastic process $(X) := (X_t)_{t=0,\ldots,T}$ that is adapted to the filtration; for example, X_t is the sum of realized cash flows until t and of an assessment of the market value of future cash flows. In a multiperiod setting, the risk of the stochastic process is measured at different states and time, such that the risk-adjusted values over time are also a stochastic process, which is denoted by $(R_t^{(X)})_{t=0,\ldots,T}$.

The risk-adjusted value process is assumed (i) to be smaller than the value process and (ii) not to decrease over time for all feasible test probabilities because information increases. Choosing the process with largest values under (i) and (ii) leads to the following definition (see also [4, sect. 4.1]).

DEFINITION 2.1 (risk-adjusted process). *Let $\mathcal{P}$ be a set of probability measures on the space $(\Omega, \mathcal{F})$, and let (X) be a bounded stochastic process adapted to the filtration $(\mathcal{F}_t)_{t=0,\ldots,T}$. The* risk-adjusted process $(R_t^{(X)})_{t=0,\ldots,T}$ of (X) *is*

$$R_t^{(X)} = \begin{cases} X_T & a.s.,\ t = T, \\ \min\Big(X_t,\ \inf_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\mathbb{Q}}[R_{t+1}^{(X)}|\mathcal{F}_t]\Big) & a.s.,\ t = 0, \ldots, T-1. \end{cases} \tag{2.2}$$

$R_t^{(X)}$ *is called* risk adjusted value at a time t. *A* risk-adjusted process *(without mentioning (X)) is the sequence of functionals $(R_t^{(\cdot)})_{t=0,\dots,T}$ which map a process (X) to $R_t^{(X)}$ for all t.*

In our finite setting, the inf-operation in (2.2) preserves $\mathcal{F}_t$-measurability. Hence, the risk-adjusted process is adapted to the filtration $(\mathcal{F}_t)_{t=0,\dots,T}$. In particular, $R_0^{(X)}$ is measurable on the trivial σ-algebra $\mathcal{F}_0$ and therefore deterministic. If we would allow for infinite Ω, then the (pointwise) infimum in (2.2) must be replaced by the essential infimum to preserve measurability [12, sect. A.4].

Next, we show that a risk-adjusted process given by Definition 2.1 has the convenient properties of coherency and time consistency. The properties follow without further assumptions on the set $\mathcal{P}$; this may not be obvious from the proofs in [4].

We use the notation $(\lambda X + Y) = (\lambda X_0 + Y_0, \dots, \lambda X_T + Y_T)$ and $(X + \lambda) = (X_0 + \lambda, \dots, X_T + \lambda)$, where $\lambda \in \mathbb{R}$ and $(X) = (X_t)_{t=0,\dots,T}$, $(Y) = (Y_t)_{t=0,\dots,T}$ are stochastic processes.

PROPOSITION 2.2 (multiperiod coherency). *Let (X) and (Y) be adapted stochastic processes, and let a risk-adjusted process $(R_t^{(\cdot)})_{t=0,\dots,T}$ be given by Definition 2.1. Then*

(i) $R_0^{(X+Y)} \geq R_0^{(X)} + R_0^{(Y)}$,
(ii) $R_0^{(\lambda X)} = \lambda R_0^{(X)}$ *for all* $\lambda \geq 0$,
(iii) $X_0 \leq Y_0, \dots, X_T \leq Y_T$ *a.s.* $\implies R_0^{(X)} \leq R_0^{(Y)}$,
(iv) $R_0^{(X+\lambda)} = R_0^{(X)} + \lambda$ *for all* $\lambda \in \mathbb{R}$.

Proof. Only (i) is not obvious. We use induction. At final time T, by (2.2) (in case $t = T$) we have

$$R_T^{(X+Y)} = X_T + Y_T \geq R_T^{(X)} + R_T^{(Y)} \quad \text{a.s.}$$

Let $t \in \{1, \dots, T-1\}$ and suppose that $R_t^{(X+Y)} \geq R_t^{(X)} + R_t^{(Y)}$ a.s. Applying (2.2) (in case $t < T$) we obtain (with abbreviation $\inf_{\mathbb{Q}} = \inf_{\mathbb{Q}\in\mathcal{P}}$)

$$\begin{aligned}
R_{t-1}^{(X+Y)} &\overset{(2.2)}{=} \min\Big(X_{t-1} + Y_{t-1}, \inf_{\mathbb{Q}} \mathbb{E}_{\mathbb{Q}}\big[R_t^{(X+Y)} \big| \mathcal{F}_{t-1}\big]\Big) \\
&\overset{(*)}{\geq} \min\Big(X_{t-1} + Y_{t-1}, \inf_{\mathbb{Q}} \mathbb{E}_{\mathbb{Q}}\big[R_t^{(X)} + R_t^{(Y)} \big| \mathcal{F}_{t-1}\big]\Big) \\
&\overset{(**)}{\geq} \min\Big(X_{t-1} + Y_{t-1}, \inf_{\mathbb{Q}} \mathbb{E}_{\mathbb{Q}}\big[R_t^{(X)} \big| \mathcal{F}_{t-1}\big] + \inf_{\mathbb{Q}} \mathbb{E}_{\mathbb{Q}}\big[R_t^{(Y)} \big| \mathcal{F}_{t-1}\big]\Big) \\
&\overset{(***)}{\geq} \min\big(X_{t-1}, \inf_{\mathbb{Q}} \mathbb{E}_{\mathbb{Q}}\big[R_t^{(X)} \big| \mathcal{F}_{t-1}\big]\big) + \min\big(Y_{t-1}, \inf_{\mathbb{Q}} \mathbb{E}_{\mathbb{Q}}\big[R_t^{(Y)} \big| \mathcal{F}_{t-1}\big]\big) \\
&\overset{(2.2)}{=} R_{t-1}^{(X)} + R_{t-1}^{(Y)} \quad \text{a.s.,}
\end{aligned}$$

where $(*)$ follows by assumption and by the monotonicity of expected values; $(**)$ follows because separate minimization gives smaller values; and $(***)$ holds because

$$\min(a + b, c + d) \geq \min\big(a + \min(b, d), c + \min(b, d)\big) = \min(a, c) + \min(b, d)$$

for all a, b, c, $d \in \mathbb{R}$. $\square$

The following definition of time consistency is based on the one in [4].

Definition 2.3 (time consistency for processes). *Let* $(F_t^{(\cdot)})_{t=0,\dots,T}$ *be a sequence of functionals that map a stochastic process* (Z) *to an* $\mathcal{F}_t$*-measurable random variable* $F_t^{(Z)}$ *for all* t*. The sequence* $(F_t^{(\cdot)})_{t=0,\dots,T}$ *is* time consistent *if for all stochastic adapted processes* (X) *and* (Y) *such that there exists* $t \in \{1,\dots,T\}$ *with*

$$F_t^{(X)} \geq F_t^{(Y)} \text{ a.s. and } X_s = Y_s \text{ a.s. for all } s = 0,\dots,t-1,$$

it holds that $F_0^{(X)} \geq F_0^{(Y)}$.

The definition of time consistency roughly says that if (X) and (Y) are ordered at time t in the sense that $F_t^{(X)} \geq F_t^{(Y)}$ a.s., then the direction of the order is the same today at zero time. A comparison at t in terms of risk should not depend on realized values prior to t; hence, the same order is required in Definition 2.3 only for processes that are equal up to t.

Proposition 2.4. *A risk-adjusted process* $(R_t^{(\cdot)})_{t=0,\dots,T}$ *given by Definition* 2.1 *is time consistent.*

The proof follows easily from the following lemma. We use alternatively the more explicit notation $R_t^{(X_0,\dots,X_T)} = R_t^{(X)}$.

Lemma 2.5. *Let* $(X) = (X_0,\dots,X_T)$ *be an adapted process. A risk-adjusted process given by Definition* 2.1 *fulfills*

$$R_0^{(X_0,\dots,X_T)} = R_0^{(X_0,\dots,X_{t-1},R_t^{(X)},\dots,R_t^{(X)})} \text{ a.s. for } t = 1,\dots,T. \tag{2.3}$$

Note that the process $(X_0,\dots,X_{t-1},R_t^{(X)},\dots,R_t^{(X)})$ in (2.3) is adapted, because $R_t^{(X)}$ is $\mathcal{F}_s$-measurable for $s = t,\dots,T$.

Proof. Equalities hold a.s. Let $t \in \{1,\dots,T\}$. We use induction from final time T to t and a second induction from t to 0. At final time T, it follows from (2.2) that

$$R_T^{(X_0,\dots,X_{t-1},R_t^{(X)},\dots,R_t^{(X)})} = R_t^{(X)}. \tag{2.4}$$

Let $s \in \{t+1,\dots,T\}$. Suppose

$$R_s^{(X_0,\dots,X_{t-1},R_t^{(X)},\dots,R_t^{(X)})} = R_t^{(X)}. \tag{2.5}$$

The induction step is

$$R_{s-1}^{(X_0,\dots,X_{t-1},R_t^{(X)},\dots,R_t^{(X)})} = \min\Big(R_t^{(X)}, \inf_{\mathbb{Q}} \mathbb{E}_{\mathbb{Q}}[R_t^{(X)}|\mathcal{F}_{s-1}]\Big) = R_t^{(X)}, \tag{2.6}$$

where the first equality is by (2.2) and by assumption (2.5), and the second equality follows from $\mathbb{E}_{\mathbb{Q}}[R_t^{(X)}|\mathcal{F}_{s-1}] = R_t^{(X)}$ since $R_t^{(X)}$ is $\mathcal{F}_{s-1}$-measurable for $s = t+1,\dots,T$. Induction with base case (2.4) yields

$$R_t^{(X_0,\dots,X_{t-1},R_t^{(X)},\dots,R_t^{(X)})} = R_t^{(X)}. \tag{2.7}$$

For the second induction, let $s \in \{1, \ldots, t\}$. Suppose

$$R_s^{(X_0,\ldots,X_{t-1},R_t^{(X)},\ldots,R_t^{(X)})} = R_s^{(X)}. \tag{2.8}$$

The induction step is

(2.9)
$$\begin{aligned} R_{s-1}^{(X_0,\ldots,X_{t-1},R_t^{(X)},\ldots,R_t^{(X)})} &\overset{(2.2)}{=} \min\Big(X_{s-1}, \inf_{\mathbb{Q}} \mathbb{E}_{\mathbb{Q}}[R_s^{(X_0,\ldots,X_{t-1},R_t^{(X)},\ldots,R_t^{(X)})}|\mathcal{F}_{s-1}]\Big) \\ &\overset{(2.8)}{=} \min\Big(X_{s-1}, \inf_{\mathbb{Q}} \mathbb{E}_{\mathbb{Q}}[R_s^{(X)}|\mathcal{F}_{s-1}]\Big) \overset{(2.2)}{=} R_{s-1}^{(X)}. \end{aligned}$$

Induction with base case (2.7) yields $R_0^{(X_0,\ldots,X_{t-1},R_t^{(X)},\ldots,R_t^{(X)})} = R_0^{(X)}$. ∎

Proof (Proposition 2.4). Let $t \in \{1, \ldots, T\}$. Let $(X) = (X_s)_{s=0,\ldots,T}$ and $(Y) = (Y_s)_{s=0,\ldots,T}$ be adapted processes such that

$$R_t^{(X)} \geq R_t^{(Y)} \text{ a.s.} \quad \text{and } X_s = Y_s \text{ a.s. for all } s = 0, \ldots, t-1.$$

The monotonicity property (iii) in Lemma 2.2 implies

$$R_0^{(X_0,\ldots,X_{t-1},R_t^{(X)},\ldots,R_t^{(X)})} \geq R_0^{(Y_0,\ldots,Y_{t-1},R_t^{(Y)},\ldots,R_t^{(Y)})},$$

which is $R_0^{(X)} \geq R_0^{(Y)}$ by (2.3) of Lemma 2.5. ∎

3. Multiperiod AVaR-set of probability measures. A coherent risk-adjusted value is defined by its set $\mathcal{P}$ of probability measures. (See (2.1) and Definition 2.1 in the previous section for single- and multiperiod risk-adjusted values, respectively.) The recursive definition of a multiperiod risk-adjusted value can be simplified to a linear stochastic program for suitably chosen $\mathcal{P}$ (see section 4). In this section, we define such a suitable set $\mathcal{P}$ that corresponds to a stepwise conditional valuation by AVaR.

Let $\alpha \in (0,1)$. The set of probability measures for the coherent risk-adjusted value AVaR can be defined as follows. A probability measure $\mathbb{Q}$ that is feasible fulfills $\mathbb{Q}[A] \leq \frac{1}{\alpha}\mathbb{P}[A]$ for all events $A \in \mathcal{F}$; α is usually small in applications (e.g., 5%). The upper bound implies that a feasible $\mathbb{Q}$ is absolutely continuous to $\mathbb{P}$; hence, the Radon–Nikodym probability density $d\mathbb{Q}/d\mathbb{P}$ exists, which allows us to define the set in terms of densities.

Definition 3.1 (AVaR). *Let $\alpha \in (0,1)$ and X be a bounded random variable. The* risk-adjusted value AVaR of X at level α *is*

$$\mathrm{AVaR}^\alpha[X] = \inf_{\mathbb{Q} \in \mathcal{Q}^\alpha} \mathbb{E}_{\mathbb{Q}}[X], \tag{3.1}$$

where the AVaR-set $\mathcal{Q}^\alpha$ *of probabilty measures at level α is*

$$\mathcal{Q}^\alpha = \left\{ \mathbb{Q} : \frac{d\mathbb{Q}}{d\mathbb{P}} \leq \frac{1}{\alpha} \ a.s. \right\}. \tag{3.2}$$

The definition corresponds to that in [19] and that of Tail-VaR in [4]. A (sign-reversed) risk measure AVaR is also defined in [12]. The original, related notion is the risk measure CVaR [21].

The AVaR-set $\mathcal{Q}^\alpha$ is extended to multiple periods as follows. First, we give the general definition, then we give a more explicit definition on a scenario tree.

Definition 3.2 (multiperiod AVaR-set). *Let* $\alpha \in (0,1)$. *The* multiperiod AVaR-set $\mathcal{P}^\alpha$ *of probability measures is*

$$\mathcal{P}^\alpha = \left\{ \mathbb{Q} : \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \middle| \mathcal{F}_t \right] \le \frac{1}{\alpha}\mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \middle| \mathcal{F}_{t-1} \right] \ a.s.,\ t = 1, \dots, T \right\}, \tag{3.3}$$

where we assume that the density $d\mathbb{Q}/d\mathbb{P}$ *exists, that is, a feasible* $\mathbb{Q}$ *is absolutely continuous to* $\mathbb{P}$.

If $T = 1$, then $\mathcal{P}^\alpha$ is the AVaR-set $\mathcal{Q}^\alpha$. For simplicity, the level α in (3.3) is constant over time. (The level could be generally an adapted process and most of the following propositions would still hold after minor adjustments.) Such a generalized set is defined in [6, sect. 2.3.1] (denoted there by $\mathcal{Q}$). The set $\mathcal{P}^\alpha$ corresponds also to [19, eq. 3.51] (denoted there by $\mathcal{W}$ as a set of densities).

On a finite probability space, the filtration $(\mathcal{F}_t)_{t=0,\dots,T}$ corresponds to a scenario tree: Given a time t, each atom $\hat{n}$ of $\mathcal{F}_t$ is identified with a node $n \in \mathcal{N}_t$, where $\mathcal{N}_t$ is the set of nodes of the tree at t. (More precisely, $\hat{n}$ is an atom of the probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$.) The set Ω is identified with the root node n_0. Let $\mathbb{Q}$ be probability measure on $(\Omega, \mathcal{F})$. The probability of a node n is defined as the probability $\mathbb{Q}[\hat{n}]$. Let n now be a nonterminal node with immediate successor (children) nodes $m_1, \dots, m_{k(n)}$, where $k(n)$ is the number of children. In case $\mathbb{Q}[\hat{n}] > 0$ we define $\mathbf{q}_n = (q_{nm_1}, q_{nm_2}, \dots, q_{nm_{k(n)}})^\top \in \mathbb{R}^{k(n)}$ to be the (single-period) transition probabilities from n to the children. Note that $\boldsymbol{e}_n^\top \boldsymbol{q}_n = 1$ for $\boldsymbol{e}_n = (1, \dots, 1)^\top$. In case $\mathbb{Q}[\hat{n}] = 0$ we define arbitrarily $\mathbf{q}_n = \boldsymbol{e}_n / k(n)$. On the scenario tree, we *identify* a probability measure $\mathbb{Q}$ with the sequence $(\boldsymbol{q}_n)_{n \in \mathcal{N}_t, t=0,\dots,T-1}$.

Definition 3.3 (multiperiod AVaR-set on scenario tree). *On a scenario tree, the* multiperiod AVaR-set $\mathcal{P}^\alpha$ *of probability measures at level* $\alpha \in (0,1)$ *is given by the set of sequences of single-period transition probabilities*

$$\mathcal{P}^\alpha = \left\{ (\boldsymbol{q}_n)_{n \in \mathcal{N}_t, t=0,\dots,T-1} : 0 \le \boldsymbol{q}_n \le \frac{1}{\alpha}\boldsymbol{p}_n,\ \boldsymbol{e}_n^\top \boldsymbol{q}_n = 1 \right\}, \tag{3.4}$$

where $\mathbf{p}_n$ *is the vector of transition probabilities from* n *to its children according to probability measure* $\mathbb{P}$.

In other words, we consider in each nonterminal node n a probability space. A scenario in this space corresponds to a child node, and the probability of the scenario is the transition probability from n to the child node. In each space, we choose the AVaR-set of probability measures at level α (Definition 3.1).

To see that the set $\mathcal{P}^\alpha$ in (3.4) is indeed equivalent to (3.3) we consider for a node $n_t \in \mathcal{N}_t$ the probabilities $\mathbb{Q}[\hat{n}_t] = q_{n_0 n_1} q_{n_1 n_2} \cdots q_{n_{t-1} n_t}$ and $\mathbb{P}[\hat{n}_t] = p_{n_0 n_1} p_{n_1 n_2} \cdots p_{n_{t-1} n_t}$, where $(n_0, n_1, n_2, \dots, n_t)$ is the path from root node n_0 to n_t. The conditional expectation of the density in node n_t is

$$\mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \middle| \mathcal{F}_t \right]\bigg|_{\hat{n}_t} = \frac{\mathbb{Q}[\hat{n}_t]}{\mathbb{P}[\hat{n}_t]}, \tag{3.5}$$

where we assume $\mathbb{P}[\hat{n}_t] > 0$ without loss of generality (else we reduce the scenario tree by cutting the node n_t which has zero probability in measure $\mathbb{P}$ and hence also in $\mathbb{Q}$). Substituting (3.5) in (3.2) we obtain (3.4).

The set $\mathcal{P}^\alpha$ (3.4) has the property of *stability* as follows. Consider two elements $\mathbb{Q}$ and $\mathbb{Q}'$ of $\mathcal{P}^\alpha$. If we replace in a nonterminal node n the vector $\mathbf{q}_n$ of $\mathbb{Q}$ by the corresponding vector $\mathbf{q}'_n$ of $\mathbb{Q}'$, then we get another measure $\mathbb{Q}''$. Because every concatenation of feasible vectors is feasible in $\mathcal{P}^\alpha$, the measure $\mathbb{Q}''$ is again in $\mathcal{P}^\alpha$. Details on the stability property can be found in [4, 8].

4. Stochastic linear programming. For an arbitrary set $\mathcal{P}$ of probability measures, we have seen that a risk-adjusted process given by Definition 2.1 is coherent and time consistent. Specifically, we can choose for $\mathcal{P}$ the multivariate AVaR-set $\mathcal{P}^\alpha$ at level α (Definitions 3.2 and 3.4). In this section, we show that for this coherent and time-consistent extension of AVaR the risk-adjusted value at time zero can be obtained by a stochastic program as follows.

PROPOSITION 4.1. *Let (X) be an adapted bounded stochastic process, and let $\alpha \in (0,1)$. The risk-adjusted value $R_0^{(X)}$ (Definition 2.1) with multiperiod AVaR-set (Definition 3.2) of probability measures, $\mathcal{P} = \mathcal{P}^\alpha$, is the optimal objective value of the stochastic linear optimization problem*

$$\left.\begin{aligned} &\max R_0 \quad \textit{s.t.} && \\ &\quad R_t \le X_t, && t = 0,\dots,T \\ &\quad R_t \le Q_t - \frac{1}{\alpha}\mathbb{E}[Z_{t+1}|\mathcal{F}_t], && t = 0,\dots,T-1 \\ &\quad Z_t \ge Q_{t-1} - R_t, && t = 1,\dots,T \\ &\quad Z_t \ge 0, && t = 1,\dots,T \end{aligned}\right\}, \tag{4.1}$$

where R_t, Q_t, and Z_t are $\mathcal{F}_t$-measurable random variables for all t, and inequalities hold a.s. Furthermore, $R_t \le R_t^{(X)}$ a.s. for all feasible R_t and all t.

For the proof of Proposition 4.1 we need a duality result of conditional AVaR. A conditionally evaluated AVaR was already informally considered in Figure 1. The definition of conditional AVaR is as follows. (See also similar definitions in [19, Exp. 2.53(b)] and [4, Def. 5.5].)

DEFINITION 4.2 (conditional AVaR). *Let $t \in \{0,\dots,T\}$. The* conditional AVaR *at level $\alpha \in (0,1)$ for a bounded random variable X at time t is pointwise*

$$\left.\begin{aligned} \mathrm{AVaR}_t^\alpha[X] = \inf_H\ &\mathbb{E}[HX|\mathcal{F}_t] \quad \textit{s.t.} \\ &\mathbb{E}[H|\mathcal{F}_t] = 1 \quad \textit{a.s.} \\ &0 \le H \le \frac{1}{\alpha} \quad \textit{a.s.} \end{aligned}\right\}, \tag{4.2}$$

where H is an integrable random variable.

In a finite setting on a scenario tree, "pointwise" means on each node at t; in an infinite setting, the infimum would need to be replaced by the essential infimum. See also the remark below Definition 2.1.

A related pointwise maximization on $\mathcal{F}_t$ for $t = 0,\dots,T$ is

$$\sup_Q \left(Q - \frac{1}{\alpha}\mathbb{E}[(Q-X)^+|\mathcal{F}_t] \right), \tag{4.3}$$

where Q is an $\mathcal{F}_t$-measurable random variable (the pointwise values are real numbers), and $(\cdot)^+ = \max(\cdot, 0)$. In the unconditional case of $t = 0$ it is well-known that $\mathrm{AVaR}_0^\alpha[X]$, which is $\mathrm{AVaR}^\alpha[X]$ in (3.1), equals the optimal objective value of (4.3) [2, 12, 19, 22].

LEMMA 4.3 (optimality for conditional AVaR). *Let $t \in \{0, \ldots, T\}$ and $\alpha \in (0, 1)$. We have the following:*

(i) *Problems* (4.2) *and* (4.3) *have the same optimal objective value (pointwise, i.e., on every event in $\mathcal{F}_t$).*

(ii) *An optimal solution of* (4.2) *is given by*

$$\hat{H} = \begin{cases} \dfrac{1}{\alpha} 1_{\{X<Q_\alpha\}} + \dfrac{1 - \frac{1}{\alpha}\mathbb{P}[X < Q_\alpha|\mathcal{F}_t]}{\mathbb{P}[X = Q_\alpha|\mathcal{F}_t]} 1_{\{X=Q_\alpha\}} & \textit{if } \mathbb{P}[X = Q_\alpha|\mathcal{F}_t] > 0, \\[2ex] \dfrac{1}{\alpha} 1_{\{X \le Q_\alpha\}} & \textit{if } \mathbb{P}[X = Q_\alpha|\mathcal{F}_t] = 0, \end{cases}$$

where $1_{\{.\}} \colon \Omega \to \{0, 1\}$ is the indicator function, equations hold a.s., and Q_α is an $\mathcal{F}_t$-measurable random variable that satisfies

$$\mathbb{P}[X < Q_\alpha|\mathcal{F}_t] \le \alpha \le \mathbb{P}[X \le Q_\alpha|\mathcal{F}_t] \qquad \textit{a.s.} \tag{4.4}$$

(iii) *An optimal solution of* (4.3) *is given by every Q_α that satisfies* (4.4).

The random variable Q_α in Lemma 4.3 is called a *conditional α-quantile of X*.

Proof. We show optimality of $\hat{H}$ and Q_α in three steps.

Step 1. We show that the objective value of problem (4.2) is an upper bound for that of problem (4.3) for all feasible solutions. With H and Q feasible, we obtain

$$\begin{aligned} \mathbb{E}[HX|\mathcal{F}_t] &\overset{(*)}{\ge} \mathbb{E}[H(Q + \min(X - Q, 0))|\mathcal{F}_t] = \mathbb{E}[H(Q - (Q - X)^+)|\mathcal{F}_t] \\ &\overset{(**)}{=} QE[H|\mathcal{F}_t] - \mathbb{E}[H(Q - X)^+|\mathcal{F}_t] = Q - \mathbb{E}[H(Q - X)^+|\mathcal{F}_t] \\ &\overset{(***)}{\ge} Q - \mathbb{E}\left[\frac{1}{\alpha}(Q - X)^+ \,\middle|\, \mathcal{F}_t\right] \ \textit{a.s.}, \end{aligned}$$

where $(*)$ follows from $H \ge 0$ a.s., $(**)$ follows from the $\mathcal{F}_t$-measurability of Q, and $(***)$ follows from $H \le 1/\alpha$ a.s.

Step 2. We show that the objective value of (4.2) with $\hat{H}$ equals that of (4.3) with Q_α. On the event $\mathbb{P}[X = Q_\alpha|\mathcal{F}_t] = 0$, the objective value of (4.2) is

$$\begin{aligned} \mathbb{E}[\hat{H}X|\mathcal{F}_t] &= \mathbb{E}\left[\frac{1}{\alpha}1_{\{X \le Q_\alpha\}}X \,\middle|\, \mathcal{F}_t\right] \\ &= \frac{1}{\alpha}\mathbb{E}[X1_{\{X\le Q_\alpha\}} - Q_\alpha 1_{\{X \le Q_\alpha\}}|\mathcal{F}_t] + \frac{1}{\alpha}\mathbb{E}[Q_\alpha 1_{\{X\le Q_\alpha\}}|\mathcal{F}_t] \\ &= -\frac{1}{\alpha}\mathbb{E}[(Q_\alpha - X)^+|\mathcal{F}_t] + \frac{1}{\alpha}Q_\alpha\mathbb{P}[X \le Q_\alpha|\mathcal{F}_t] \\ &= Q_\alpha - \frac{1}{\alpha}\mathbb{E}[(Q_\alpha - X)^+|\mathcal{F}_t] \quad \text{a.s.} \end{aligned}$$

On the complementary event $\mathbb{P}[X = Q_\alpha|\mathcal{F}_t] > 0$, the objective value is

$$
\begin{aligned}
\mathbb{E}[\hat{H}X|\mathcal{F}_t] &= \mathbb{E}\left[\left(\frac{1}{\alpha}1_{\{X<Q_\alpha\}} + \frac{1-\frac{1}{\alpha}\mathbb{P}[X<Q_\alpha|\mathcal{F}_t]}{\mathbb{P}[X=Q_\alpha|\mathcal{F}_t]}1_{\{X=Q_\alpha\}}\right)X \,\middle|\, \mathcal{F}_t\right] \\
&= \frac{1}{\alpha}\mathbb{E}[X1_{\{X<Q_\alpha\}}|\mathcal{F}_t] + \frac{1-\frac{1}{\alpha}\mathbb{P}[X<Q_\alpha|\mathcal{F}_t]}{\mathbb{P}[X=Q_\alpha|\mathcal{F}_t]}\mathbb{E}[X1_{\{X=Q_\alpha\}}|\mathcal{F}_t] \\
&= \frac{1}{\alpha}\mathbb{E}[X1_{\{X<Q_\alpha\}} - Q_\alpha 1_{\{X<Q_\alpha\}}|\mathcal{F}_t] + \frac{1}{\alpha}\mathbb{E}[Q_\alpha 1_{\{X<Q_\alpha\}}|\mathcal{F}_t] \\
&\quad + Q_\alpha\left(1 - \frac{1}{\alpha}\mathbb{P}[X<Q_\alpha|\mathcal{F}_t]\right) \\
&= -\frac{1}{\alpha}\mathbb{E}[(Q_\alpha - X)^+|\mathcal{F}_t] + \frac{1}{\alpha}Q_\alpha\mathbb{P}[X<Q_\alpha|\mathcal{F}_t] + Q_\alpha\left(1-\frac{1}{\alpha}\mathbb{P}[X<Q_\alpha|\mathcal{F}_t]\right) \\
&= Q_\alpha - \frac{1}{\alpha}\mathbb{E}[(Q_\alpha - X)^+|\mathcal{F}_t] \quad \text{a.s.}
\end{aligned}
$$

Hence, $\hat{H}$ and Q_α yield the same objective values for (4.2) and (4.3).

Step 3. Finally, we have to check that $\hat{H}$ and Q_α are feasible. Q_α is trivially feasible because it is $\mathcal{F}_t$-measurable by definition; we can check the feasibility of $\hat{H}$ by inserting $\hat{H}$ into the constraints of (4.2) on each of the two events as in the previous step and by observing that Q_α satisfies (4.4). $\square$

Proof (Proposition 4.1). The proof has three steps. Equalities hold a.s.

Step 1. We reformulate the minimization of the conditional expectation in (2.2) as a maximization. Let $(R_t^{(X)})_{t=0,\dots,T}$ be a risk-adjusted process with a multiperiod AVaR-set of probability measures, $\mathcal{P} = \mathcal{P}^\alpha$. For $t = 0,\dots,T-1$, we have

$$
\begin{aligned}
\inf_{\mathbb{Q}\in\mathcal{P}^\alpha}\mathbb{E}_\mathbb{Q}[R_{t+1}^{(X)}|\mathcal{F}_t] &= \begin{cases} \inf_{\mathbb{Q}\in\mathcal{P}^\alpha}\mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\left(\mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_t\right]\right)^{-1}R_{t+1}^{(X)} \;\middle|\; \mathcal{F}_t\right] & \text{if } \mathbb{E}[\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_t] > 0, \\ 0 & \text{if } \mathbb{E}[\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_t] = 0 \end{cases} \\
&= \inf\left\{\mathbb{E}[HR_{t+1}^{(X)}|F_t] : H \in L^1(\Omega,\mathcal{F},\mathbb{P}),\ \mathbb{E}[H|F_t] = 1,\ 0 \le H \le \frac{1}{\alpha}\right\} \\
&= \max_Q\left(Q - \frac{1}{\alpha}\mathbb{E}\left[(Q - R_{t+1}^{(X)})^+ \;\middle|\; \mathcal{F}_t\right]\right),
\end{aligned} \tag{4.5}
$$

where the first equality follows from properties of conditional expectation; on the event where $\mathbb{E}[d\mathbb{Q}/d\mathbb{P}|\mathcal{F}_t] = 0$, we have also $d\mathbb{Q}/d\mathbb{P} = 0$, and therefore $\mathbb{E}_\mathbb{Q}[Y|\mathcal{F}_t] = \mathbb{E}[Y d\mathbb{Q}/d\mathbb{P}|\mathcal{F}_t] = 0$ on that event for a random variable Y. For the second equality in (4.5) we used the substitution $H = d\mathbb{Q}/d\mathbb{P}(\mathbb{E}[d\mathbb{Q}/d\mathbb{P}|\mathcal{F}_t])^{-1}$ on the event $\mathbb{E}[d\mathbb{Q}/d\mathbb{P}|\mathcal{F}_t] > 0$ (else we set $H = 0$), and we used Definition 3.3 of $\mathcal{P}^\alpha$. The last equality in (4.5) follows from Lemma 4.3.

Step 2. We show that $R_0^{(X)}$ is the optimal objective value of

$$
\left.\begin{aligned}
z = \max_{(R_t)}\ & R_0 \quad \text{s.t.} & \\
& R_t \le X_t, & t = 0,\dots,T \\
& R_t \le \max_{Q_t}\Big(Q_t - \frac{1}{\alpha}\mathbb{E}[(Q_t - R_{t+1})^+|\mathcal{F}_t]\Big), & t = 0,\dots,T-1
\end{aligned}\right\}, \tag{4.6}
$$

where R_t is $\mathcal{F}_t$-measurable for all t, and the maximization over the $\mathcal{F}_t$-measurable Q_t is understood pointwise on $\mathcal{F}_t$ for all t. The random variables $(R_t^{(X)})_{t=0,\dots,T}$ are feasible in problem (4.6) because they satisfy (2.2) and we have (4.5). Hence, $R_0^{(X)} \le z$. The reverse inequality is shown by induction as follows. Let $(R_t)_{t=0,\dots,T}$ be feasible in (4.6). At final time, $R_T \le X_T \overset{(2.2)}{=} R_T^{(X)}$. Let $t \in \{0,\dots,T-1\}$. Suppose $R_{t+1} \le R_{t+1}^{(X)}$. The induction step is

$$
\begin{aligned}
R_t &\overset{(4.6)}{\le} \min\left\{X_t,\ \max_{Q_t}\left(Q_t - \frac{1}{\alpha}\mathbb{E}[(Q_t - R_{t+1})^+|\mathcal{F}_t]\right)\right\} \\
&\overset{(4.5)}{=} \min\big\{X_t,\ \min_{\mathbb{Q}\in\mathcal{P}^\alpha} \mathbb{E}_{\mathbb{Q}}[R_{t+1}|\mathcal{F}_t]\big\} \\
&\le \min\big\{X_t,\ \min_{\mathbb{Q}\in\mathcal{P}^\alpha} \mathbb{E}_{\mathbb{Q}}[R_{t+1}^{(X)}|\mathcal{F}_t]\big\} \\
&\overset{(2.2)}{=} R_t^{(X)}.
\end{aligned}
\tag{4.7}
$$

By induction $R_0 \le R_0^{(X)}$, hence $z \le R_0^{(X)}$. In combination with the reverse inequality from above we get $R_0^{(X)} = z$.

Step 3. Finally, we show that problem (4.6) is equivalent to problem (4.1). Suppose $(\hat{R}_t)_{t=0,\dots,T}$ is feasible in (4.6) with corresponding optimal solution $(\hat{Q}_t)_{t=0,\dots,T-1}$ of the inner maximization. We can immediately check that $(\hat{R}_t)_{t=0,\dots,T}$, $(\hat{Q}_t)_{t=0,\dots,T-1}$, and $(\hat{Z}_t)_{t=1,\dots,T}$ given by $\hat{Z}_t = (\hat{Q}_{t-1} - \hat{R}_t)^+$ is feasible in (4.1).

On the other hand, suppose $(\hat{R}_t)_{t=0,\dots,T}$, $(\hat{Q}_t)_{t=0,\dots,T-1}$, $(\hat{Z}_t)_{t=0,\dots,T-1}$ are feasible in (4.1). Then for all $t \in \{0,\dots,T-1\}$,

$$
\begin{aligned}
\max_Q\left(Q - \frac{1}{\alpha}\mathbb{E}\big[(Q - \hat{R}_{t+1})^+\big|\mathcal{F}_t\big]\right) &\ge \hat{Q}_t - \frac{1}{\alpha}\mathbb{E}\big[(\hat{Q}_t - \hat{R}_{t+1})^+\big|\mathcal{F}_t\big] \\
&\overset{(4.1)}{\ge} \hat{Q}_t - \frac{1}{\alpha}\mathbb{E}\big[\hat{Z}_{t+1}\big|\mathcal{F}_t\big] \overset{(4.1)}{\ge} \hat{R}_t.
\end{aligned}
$$

Hence, $(\hat{R}_t)_{t=0,\dots,T}$ is feasible in (4.6). In combination with the opposite statement from above, it follows that problem (4.6) is equivalent to problem (4.1), and the optimal objective values are the same. Furthermore, to prove the last statement of Proposition 4.1 we observe that for every feasible R_t in (4.1) we have $R_t \le R_t^{(X)}$ for all t by using (4.7). ∎

5. Mean-risk optimization.

In the previous section, we considered a linear reformulation of a multiperiod extension of the risk-adjusted value AVaR; speficially, we considered a risk-adjusted process with a multiperiod AVaR-set of probability measures. In this section, we show how we can incorporate the extension in a multiperiod mean-risk optimization problem. Furthermore, we provide the explicit formulation on a scenario tree, which is needed for applications.

In decision making under uncertainty, risk-adjusted values can be used to decide whether a random variable or a stochastic process is acceptable in terms of risk (e.g., see [4, sect. 1]). A stochastic process $(X) = (X_t)_{t=0,\dots,T}$ is accepted if the risk-adjusted value of today is above a threshold, that is, $R_0^{(X)} \ge \rho \in \mathbb{R}$. We can assume ρ to be zero because it follows from the translation equivariance (iv) of Lemma 2.2 that for the translated process $\hat{X}_t = X_t - \rho$, $t = 0, \dots, T$, the acceptance is $R_0^{(\hat{X})} \geq 0$.

We consider multiperiod mean-risk optimization problems of form

$$\left.\begin{aligned} &\sup_{(X)} \mathbb{E}[g(X_0, \dots, X_T)] \quad \text{s.t.} \\ &\qquad R_0^{(X)} \geq 0 \\ &\qquad (X) \in \mathcal{X} \end{aligned}\right\}, \tag{5.1}$$

where $g\colon \mathbb{R}^{T+1} \to \mathbb{R}$ is a measurable function, and the set $\mathcal{X}$ represents the problem-specific constraints. We assume that g and $\mathcal{X}$ are such that the supremum is finite. We can reformulate (5.1) with a linearized constraint on risk as follows.

PROPOSITION 5.1 (mean-risk optimization). *Let the risk-adjusted process be given by Definition* 2.1 *with multiperiod AVaR-set $\mathcal{P}^\alpha$ of probability measures (Definition* 3.2*) at level $\alpha \in (0, 1)$. Then an optimal solution of the mean-risk problem* (5.1) *is given by the optimal variables $(X) = (X_t)_{t=0,\dots,T}$ of the stochastic optimization problem*

$$\left.\begin{aligned} &\sup \mathbb{E}[g(X_0, \dots, X_T)] \quad \textit{s.t.} \\ &\qquad R_0 \geq 0 \\ &\qquad R_t \leq X_t, && t = 0, \dots, T \\ &\qquad R_t \leq Q_t - \frac{1}{\alpha}\mathbb{E}[Z_{t+1}|\mathcal{F}_t], && t = 0, \dots, T-1 \\ &\qquad Z_t \geq Q_{t-1} - R_t, && t = 1, \dots, T \\ &\qquad Z_t \geq 0, && t = 1, \dots, T \\ &\qquad (X_t)_{t=0,\dots,T} \in \mathcal{X} \end{aligned}\right\}, \tag{5.2}$$

where X_t, R_t, Q_t, and Z_t are $\mathcal{F}_t$-measurable random variables for all t, and inequalities hold a.s.

Proof. We show that problems (5.1) and (5.2) are equivalent. Suppose $(\hat{X})$ is feasible in (5.1), and let $(\hat{R}) = \hat{R}_{t=0,\dots,T}$, $(\hat{Q}) = \hat{Q}_{t=0,\dots,T-1}$, $(\hat{Z}) = \hat{Z}_{t=1,\dots,T}$ be optimal in (4.1). It can be easily checked that $(\hat{X})$, $(\hat{R})$, $(\hat{Q})$ are also feasible in (5.2). On the other hand, suppose $(\hat{X})$, $(\hat{R})$, $(\hat{Q})$, $(\hat{Z})$ are feasible in (5.2). By (4.1), we get $R_0^{(\hat{X})} \geq \hat{R}_0 \geq 0$; that is, $(\hat{X})$ is feasible in (5.1).

Hence, problems (5.1) and (5.2) are equivalent, and the optimal objective values are the same. ∎

If $T = 1$, then the mean-risk optimization (5.2) corresponds to the optimization with risk measure CVaR in [22, Thm. 16] (by using different sign conventions and in their special case of $l = 1$).

In the following, the mean-risk optimization problem (5.2) is formulated on a scenario tree, which is the formulation needed for numerical applications. Some of the notation on a scenario tree was introduced above Definition 3.4. On the finite probability space, let $x_{tn} \in \mathbb{R}$ denote the value of an $\mathcal{F}_t$-measurable random variable X_t on the atom $\hat{n} \in \mathcal{F}_t$ at time t, where $\hat{n}$ corresponds to node $n \in \mathcal{N}_t$. The transition probability from a nonterminal node n to node m is denoted by p_{nm}. The value of a conditional expectation on a nonterminal node $n \in \mathcal{N}_t$ is

$$\mathbb{E}[X_{t+1}|\mathcal{F}_t]\Big|_{\hat{n}} = \sum_{m \in \mathcal{N}_{t+1}} p_{nm} x_{(t+1)m},$$

where we can sum over every node at time $t+1$ because the transition probabilities from n to only its successor nodes are nonzero. With the notation from above, problem (5.2) can be formulated as the linear program

$$\begin{aligned}
\sup \sum_{n_T \in \mathcal{N}_T} & p_{n_0 n_T} g(x_{0n_0}, \ldots, x_{Tn_T}) \quad \text{s.t.} && \\
& r_{0n_0} \geq 0, && \\
& r_{tn} \leq x_{tn}, && t = 0, \ldots, T, n \in \mathcal{N}_t, \\
& r_{tn} \leq q_{tn} + \frac{1}{\alpha} \sum_{m \in \mathcal{N}_{t+1}} p_{nm} z_{(t+1)m}, && t = 0, \ldots, T-1, n \in \mathcal{N}_t, \\
& z_{tn} \geq q_{(t-1)n^-} - r_{tn}, && t = 1, \ldots, T, n \in \mathcal{N}_t, \\
& z_{tn} \geq 0, && t = 1, \ldots, T, n \in \mathcal{N}_t, \\
& (x_{tn})_{t=0,\ldots,T} \in \hat{\mathcal{X}}, && n \in \mathcal{N}_t,
\end{aligned}$$

where n_0 is the root node of the scenario tree, n^- is the parent node of n, and the set $\hat{\mathcal{X}}$ is defined for the realizations of $(X) \in \mathcal{X}$.

In our finite setting, if α is sufficiently close to zero, then the risk-adjusted value $\mathrm{AVaR}^\alpha[X]$ (3.1) attains the minimal value of the random variable X. Similarly, the risk-adjusted value $R_0^{(X)}$ with multiperiod AVaR-set $\mathcal{P} = \mathcal{P}^\alpha$ attains for sufficiently small α the minimal value of $(X) = (X_t)_{t=0,\ldots,T}$; this can be seen from the recursive definition (2.2). To avoid such worst-case values, it is sufficient that all transition probabilities over (single) time steps are strictly smaller than α. Accordingly, all nonterminal nodes of the scenario tree have sufficiently many children, which leads to numerically demanding problem sizes.

6. Risk-adjusted final values. In the previous sections, we discussed risk-adjusted values for a stochastic process $X_0, \ldots, X_T$ of financial values. In certain applications, the intermediate values $X_0, \ldots, X_{T-1}$ are not a concern for the decision maker; for example, the decision maker is concerned only by the payoff at final time T of a contractually locked-in value. The risk-adjusted value of a final value X_T can be considered as the one of a stochastic process that has sufficiently large intermediate values. For example, if we assume $X_t = \|X_T\|_\infty$ for $t = 0, \ldots, T-1$, then $X_0, \ldots, X_{T-1}$ are not relevant in the recursive calculation (2.2) of the risk-adjusted process; see also [4, Definition 5.1]. Accordingly, the stochastic programming formulation (4.1) for risk-adjusted values that have a multiperiod AVaR-set $\mathcal{P}^\alpha$ of probability measures can be simplified by removing the constraints $R_t \leq X_t$ for $t = 0, \ldots, T-1$. Equivalently, a coherent risk-adjusted final value can be defined in the form of problem (2.1) as follows.

DEFINITION 6.1 (risk-adjusted final value). *Let X be a bounded $\mathcal{F}_T$-measurable random variable, and $\alpha \in (0,1)$. The* risk-adjusted value of (final) X with multiperiod AVaR-set $\mathcal{P}^\alpha$ *is the optimal objective value of*

$$\inf_{\mathbb{Q} \in \mathcal{P}^\alpha} \mathbb{E}_{\mathbb{Q}}[X]. \tag{6.1}$$

By using Definition 3.2 of $\mathcal{P}^\alpha$ and substituting $H = d\mathbb{Q}/d\mathbb{P}$, we can write problem (6.1) more explicitly as

$$\left.\begin{aligned} &\inf_H \mathbb{E}[HX] \quad \text{s.t.} \\ &\quad \mathbb{E}[H] = 1, \quad H \geq 0 \quad \text{a.s.} \\ &\quad \mathbb{E}[H|\mathcal{F}_{t+1}] \leq \frac{1}{\alpha}\mathbb{E}[H|\mathcal{F}_t] \quad \text{a.s.}, \quad t = 0, \dots, T-1, \qquad (*) \end{aligned}\right\}, \tag{6.2}$$

where H is an $\mathcal{F}_T$-measurable integrable random variable.

Problem (6.1) corresponds in the literature to *dynamic AVaR* [6, sect. 2.3.1], *composed AVaR* [7, sect. 5], *conditional risk mappings* [24, Defs. 5.8, 6.18], *dynamically consistent TVaR* [23, Def. 8.2], and *nested-AVaR* (in form of the equivalent problem (6.6) below) [19, Ex. 3.34, Prop. 3.35].

The risk-adjusted final value (6.1) is bounded from below by AVaR as follows.

PROPOSITION 6.2. *Let X be a bounded random variable, and let $\mathcal{P}^\alpha$ be the multiperiod AVaR-set (Definition* 3.2*) at level $\alpha \in (0,1)$. Then*

$$\mathrm{AVaR}^{\alpha^T}[X] \leq \inf_{Y\in\mathcal{Y}} \inf_{\mathbb{Q}\in\mathcal{P}^\alpha} \mathbb{E}_{\mathbb{Q}}[Y], \tag{6.3}$$

where $\mathcal{Y}$ is the set of random variables having the same distribution law as X, and AVaR is given by Definition 3.1.

By the bound (6.3), we may choose the level α for multiperiod risk-adjusted values in the order of the Tth power of the level for single-period AVaR. This multiplicative effect is also observed in [23, Exp. 8.8]. It would be desirable that multiperiod risk-adjusted values are bounded also from above by AVaR. Unfortunately, at least in continuous time, a time-consistent risk-adjusted value that is more conservative than AVaR is basically the value in the worst-case scenario [8, Thm. 9]. However, upper bounds exist in terms of other multiperiod risk-adjusted values [19, Prop. 3.36]. For a different conditional multiperiod risk measure, an upper bound in terms of AVaR is given in [26].

Proof. Equalities hold a.s. We show that problem (6.2) is a relaxation of problem (4.2) in the unconditional case $t = 0$. By building the product of the left-hand sides of the constraints $(*)$ in (6.2) and also the product of the right-hand sides, we can form a constraint on the products:

$$\prod_{t=0}^{T-1} \mathbb{E}[H|\mathcal{F}_{t+1}] = H \prod_{t=0}^{T-1} \mathbb{E}[H|\mathcal{F}_t] \leq \frac{1}{\alpha^T} \prod_{t=0}^{T-1} \mathbb{E}[H|\mathcal{F}_t]. \tag{6.4}$$

Equation (6.4) is satisfied for every feasible H in problem (6.2). Indeed, let H be a feasible $\mathcal{F}_T$-measurable random variable. The equality in (6.4) is satisfied because $\mathbb{E}[H|\mathcal{F}_T] = H$ and $\mathbb{E}[H|\mathcal{F}_0] = 1$, and the inequality is implied by the constraints $(*)$ since all terms are nonnegative. By dividing the inequality in (6.4) on the event where the product is nonzero, we obtain

$$H \leq \frac{1}{\alpha^T} \qquad \text{if } \prod_{t=0}^{T-1} \mathbb{E}[H|\mathcal{F}_t] > 0.$$

The complementary event is where $\prod_{t=0}^{T-1} \mathbb{E}[H|\mathcal{F}_t] = 0$. Note that for nonnegative H if $\mathbb{E}[H|\mathcal{F}_t] = 0$ on an event, then $\mathbb{E}[H|\mathcal{F}_{t+1}] = 0$ since $\mathbb{E}[H|\mathcal{F}_t] = \mathbb{E}[\mathbb{E}[H|\mathcal{F}_{t+1}]|\mathcal{F}_t]$. Hence, on the complementary event it holds that $\mathbb{E}[H|\mathcal{F}_T] = 0$, which implies $H = 0$. Therefore, on every event, the constraints $(*)$ imply $H \leq 1/\alpha^T$ a.s., which is a constraint in problem (4.2) of AVaR at level α^T. The other constraints in (4.2) and problem (6.2) are the same. Hence, $\mathrm{AVaR}^{\alpha^T}[X] \leq \inf_{\mathbb{Q}\in\mathcal{P}^\alpha} \mathbb{E}_{\mathbb{Q}}[X]$. Furthermore, only the distribution of X enters the calculation of AVaR because X is inside an expected value under $\mathbb{P}$ in formulation (4.3) ($t = 0$). Consequently, we can take the infimum on the right-hand side over all random variables having the same distribution. ∎

As a final result, we show that the dual of problem (6.1) has a similar form as the well-known dual problem (4.3) of AVaR. For this result, we make explicit use of our general assumption that Ω is finite.

Proposition 6.3. *Let X be a bounded random variable, and let $\mathcal{P}^\alpha$ be the multiperiod AVaR-set (Definition 3.2) at level $\alpha \in (0,1)$. Suppose $X = \sum_{t=1}^T Y_t$, where Y_t is an $\mathcal{F}_t$-measurable random variable for all t. Then $\min_{\mathbb{Q}\in\mathcal{P}^\alpha} \mathbb{E}_{\mathbb{Q}}[X]$ is equal to the optimal objective value of*

$$\left.\begin{aligned} &\max\left(q - \frac{1}{\alpha}\mathbb{E}[Z_1]\right) \quad s.t. \\ &Z_t = \begin{cases} \left(A_t + \frac{1}{\alpha}\mathbb{E}[Z_{t+1}|\mathcal{F}_t] - Y_t\right)^+ & a.s.,\ t = 1,\dots,T \\ 0 & a.s.,\ t = T+1 \end{cases} \\ &q = \sum_{t=1}^T A_t \quad a.s. \end{aligned}\right\}, \tag{6.5}$$

where $q \in \mathbb{R}$, A_t, Z_t are $\mathcal{F}_t$-measurable random variables for all t, and $(\cdot)^+ = \max(\cdot, 0)$.

An interpretation of problem (6.5) is as follows. A deterministic financial amount q is committed in advance. The amount q must be matched by a sum of successive amounts, $(A_t)_{t=1,\dots,T}$, that can depend on the paths in the scenario tree and that are drawn from an income stream $(Y_t)_{t=1,\dots,T}$. Because q is committed in advance, a shortfall is insured. The insurance is proportional to the expected total shortfall Z_1, which is recursively obtained from the shortfalls on each stage.

Proof. Because Ω is finite, strong linear duality holds; it remains to show that (6.5) can be formulated as a linear dual problem to (6.2). Equalities hold a.s. Using the substitution $H_t = \mathbb{E}[H|\mathcal{F}_t]$ for all $t = 0,\dots,T$, and $X = \sum_{t=1}^T Y_t$, we can write problem (6.2) equivalently as

$$\left.\begin{aligned} &\min_{(H_t)} \sum_{t=1}^T \mathbb{E}[H_t Y_t] \quad \text{s.t.} \\ &\mathbb{E}[H_{t+1}|\mathcal{F}_t] = H_t, \quad t = 0,\dots,T-1, && (M_t) \\ &H_0 = 1 \\ &H_t \geq 0, \qquad t = 0,\dots,T \\ &H_{t+1} \leq \frac{1}{\alpha} H_t, \qquad t = 0,\dots,T-1, && (Z_{t+1}) \end{aligned}\right\}, \tag{6.6}$$

where H_t is an $\mathcal{F}_t$-measurable random variable for all t, and we have indicated the $\mathcal{F}_t$-measurable random variables M_t and $Z_t \geq 0$ as Lagrange multipliers for the calculation of the dual. The Lagrangian is

$$\begin{aligned} L &= \sum_{t=1}^{T} \mathbb{E}[H_t Y_t] + \sum_{t=0}^{T-1} \mathbb{E}\left[Z_{t+1}\left(H_{t+1} - \frac{1}{\alpha} H_t\right)\right] + \sum_{t=0}^{T-1} \mathbb{E}\big[M_t\big(\mathbb{E}[H_{t+1}|\mathcal{F}_t] - H_t\big)\big] \\ &= \mathbb{E}\big[H_T\big(Y_T + Z_T + M_{T-1}\big)\big] + \sum_{t=1}^{T-1} \mathbb{E}\left[H_t\left(Y_t + Z_t - \frac{1}{\alpha}\mathbb{E}[Z_{t+1}|\mathcal{F}_t] + M_{t-1} - M_t\right)\right] \\ &\quad - \mathbb{E}\left[H_0\left(M_0 + \frac{1}{\alpha} Z_1\right)\right]. \end{aligned}$$

By restricting the minimization of L to finite value, we obtain the constraints of the dual problem:

$$\inf_{H_t \geq 0} L > -\infty \iff \begin{cases} Y_T + Z_T + M_{T-1} \geq 0, \\ Y_t + Z_t - \dfrac{1}{\alpha}\mathbb{E}[Z_{t+1}|\mathcal{F}_t] - M_t + M_{t-1} \geq 0, \quad t = 1, \ldots, T-1. \end{cases}$$

We use the substitution $A_t = M_t - M_{t-1}$ for $t = 1, \ldots, T$ by defining $M_T = 0$; the substitution back is $M_t = -\sum_{s=t+1}^{T} A_s$. Hence, $\sum_{t=1}^{T} A_t = -M_0 \in \mathbb{R}$. By setting $q = -M_0$ in L and observing that $H_0 = 1$, we obtain the dual problem of (6.6) as

$$\begin{aligned} &\max_{q,(A_t),(Z_t)} q - \frac{1}{\alpha}\mathbb{E}[Z_1] \quad \text{s.t.} \\ &\qquad \left.\begin{aligned} &Z_t \geq 0, && t = 1, \ldots, T \\ &Z_t \geq A_t + \frac{1}{\alpha}\mathbb{E}[Z_{t+1}|\mathcal{F}_t] - Y_t, && t = 1, \ldots, T-1 \end{aligned}\right\} (*) \\ &\qquad Z_T \geq A_T - Y_T, \\ &\qquad q = \sum_{t=1}^{T} A_t. \end{aligned}$$

Let a feasible q and feasible $(A_t)_{t=1,\ldots,T}$ be fixed. Note that only Z_1 appears in the objective function (and not $Z_2, \ldots, Z_T$); an optimal Z_1 can be chosen as low as feasibility allows in every node of the scenario tree at $t = 1$. Z_1 is bounded from below by the constraints $(*)$ for $t = 1$. Hence, an optimal Z_2 can be lowered—without causing a decrease in objective value—as long as Z_2 is feasible. A feasible Z_2 is in turn constrained by $(*)$ for $t = 2$. Iteratively, it follows that at each time t and each node, one of the two inequalities in $(*)$ can be chosen to be tight in an optimal solution. Therefore, we can restrict the feasible set to that of problem (6.5). ∎

7. Conclusion. We considered an extension of AVaR for evaluating the risk of a sequence of financial values, and as a special case for evaluating the risk of a final value at a time horizon. The provided linear formulation allows us to incorporate the extension into multiperiod mean-risk optimization problems, and the upper bound for the risk of final values gives a hint about how to choose the level α of the multiperiod AVaR-set of probability measures in applications.

The considered extension of AVaR is applied in [9], where we consider a mean-risk model for the optimal operation of a pump-storage hydropower plant (for details, see the introduction). The obtained numerical solution shows that the risk of intermediate values could be captured without detrimental effects on final values. The properties of coherency and of time consistency ensure that the decision making in such models is free of possible pathologies if these properties are not always fulfilled.

To avoid technical details and to make the results more comprehensible, we assumed a finite probability space. Indeed, we did not refer substantially to the finiteness assumptions in the proofs (apart from Proposition 6.3, where we use strong linear duality on finite spaces). We already indicated how to change notation when allowing for infinite spaces; e.g., see the remark below Definition 2.1 on the essential infimum. Moreover, by using general notation of probability theory and providing also notational counterparts using realizations of random variables, we hope that this paper helps to improve the link between the theoretical literature on risk measurement and applications.

Acknowledgment. I thank two referees for helping to improve the paper considerably. I thank the research staff of the Energy Economics Group at the Paul Scherrer Institute and my supervisors P. Kall, H.-J. Lüthi, and J. Mayer.

REFERENCES

[1] B. Acciaio and I. Penner, *Dynamic risk measures*, in Advanced Mathematical Methods for Finance, G. D. Nunno and B. Øksendal, eds., Springer, Berlin, 2011, pp. 1–34.

[2] C. Acerbi and D. Tasche, *On the coherence of expected shortfall*, J. Banking Finance, 26 (2002), pp. 1487–1503.

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

[4] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, and H. Ku, *Coherent multiperiod risk adjusted values and Bellman's principle*, Ann. Oper. Res., 152 (2007), pp. 5–22.

[5] D. Bertsimas, G. J. Lauprete, and A. Samarov, *Shortfall as a risk measure: properties, optimization and applications*, J. Econom. Dynam. Control, 28 (2004), pp. 1353–1381.

[6] P. Cheridito and M. Kupper, *Compostion of time-consistent dynamic monetary risk measures in discrete time*, Int. J. Theor. Appl. Finance, 14 (2011), pp. 137–162.

[7] P. Cheridito and M. Stadje, *Time-inconsistency of VaR and time-consistent alternatives*, Finance, 6 (2009), pp. 40–46.

[8] F. Delbaen, *The Structure of m-Stable Sets and in Particular of the Set of Risk Neutral Measures*, Lecture Notes in Math. 1874, Springer, New York, 2006, pp. 215–258.

[9] M. Densing, *Price-driven hydropower dispatch under uncertainty*, in Handbook of Risk Management in Energy Production and Trading, R. M. Kovacevic, G. Ch. Pflug, and M. T. Vespucci, eds., Internat. Ser. Oper. Res. Management Sci. 199, Springer, New York, 2013.

[10] M. Densing and J. Mayer, *Multiperiod stochastic optimization problems with time-consistent risk constraints*, in Operations Research Proceedings 2011, D. Klatte, H.-J. Lüthi, and K. Schmedders, eds., Springer, Berlin, 2012, pp. 521–526.

[11] A. Eichhorn, H. Heitsch, and W. Römisch, *Stochastic optimization of electricity portfolios: Scenario tree modeling and risk management*, in Handbook of Power Systems II, S. Rebennack, P. M. Pardalos, M. V. F. Pereira, and N. A. Iliadis, eds., Springer, Berlin, 2010, pp. 405–432.

[12] H. Föllmer and A. Schied, *Stochastic Finance*, de Gruyter Stud. Math., de Gruyter, Berlin, 2002.

[13] H. Geman and S. Ohana, *Time-consistency in managing a commodity portfolio: A dynamic risk measure approach*, J. Banking Finance, 32 (2008), pp. 1991–2005.

[14] P. Kall and J. Mayer, *Stochastic Linear Programming: Models, Theory, and Computation*, 2nd ed., Internat. Ser. Oper. Res. Management Sci. 156, Springer, New York, 2011.

[15] S. Kilianová and G. Ch. Pflug, *Optimal pension fund management under multi-period risk minimization*, Ann. Oper. Res., 166 (2009), pp. 261–270.

[16] R. Kovacevic and G. Ch. Pflug, *Time consistency and information monotonicity of multiperiod acceptability functionals*, in Advanced Financial Modelling, H. Albrecher,

W. J. Runggaldier, and W. Schachermayer, eds., Radon Ser. Comput. Appl. Math. 8, de Gruyter, Berlin, 2009, pp. 347–370.

[17] A. Künzi-Bay and J. Mayer, *Computational aspects of minimizing conditional value-at-risk*, Comput. Management Sci., 3 (2006), pp. 3–27.

[18] G. Ch. Pflug, *Some remarks on the value-at-risk and the conditional value-at-risk*, in Pobabilistic Constrained Optimization, S. P. Uryasev, ed., Nonconvex Optim. Appl. 49, Kluwer Academic Publishers, Dordrecht, 2000, pp. 272–281.

[19] G. Ch. Pflug and W. Römisch, *Modeling, Measuring, and Managing Risk*, World Scientific, Hackensack, NJ, 2007.

[20] A. B. Philpott and V. L. de Matos, *Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion*, Eur. J. Oper. Res., 218 (2012), pp. 470–483.

[21] R. T. Rockafellar and S. Uryasev, *Optimization of conditional value-at-risk*, J. Risk, 2 (2000), pp. 21–41.

[22] R. T. Rockafellar and S. Uryasev, *Conditional value-at-risk for general loss distributions*, J. Banking Finance, 26 (2002), pp. 1443–1471.

[23] B. Roorda and J. M. Schumacher, *Time consistency conditions for acceptability measures, with an application to tail value at risk*, Insurance Math. Econom., 40 (2007), pp. 209–230.

[24] A. Ruszczyński and A. Shapiro, *Conditional risk mappings*, Math. Oper. Res., 31 (2006), pp. 544–561.

[25] G. Szegö, ed., *Risk Measures for the 21st Century*, Wiley, New York, 2004.

[26] L. Xin and A. Shapiro, *Bounds for nested law invariant coherent risk measures*, Oper. Res. Lett., 40 (2012), pp. 431–435.