



Stochastics and Statistics

Bounds on risk-averse mixed-integer multi-stage stochastic programming problems with mean-CVaR



Ali İrfan Mahmutogulları, Özlem Çavuş*, M. Selim Aktürk

Department of Industrial Engineering, Bilkent University, Ankara 06800, Turkey

ARTICLE INFO

Article history:

Received 2 July 2016

Accepted 17 October 2017

Available online 13 November 2017

Keywords:

Stochastic programming

Mixed-integer multi-stage stochastic programming

Dynamic measures of risk

CVaR

Bounding

ABSTRACT

Risk-averse mixed-integer multi-stage stochastic programming forms a class of extremely challenging problems since the problem size grows exponentially with the number of stages, the problem is non-convex due to integrality restrictions, and the objective function is nonlinear in general. We propose a scenario tree decomposition approach, namely group subproblem approach, to obtain bounds for such problems with an objective of dynamic mean conditional value-at-risk (mean-CVaR). Our approach does not require any special problem structure such as convexity and linearity, therefore it can be applied to a wide range of problems. We obtain lower bounds by using different convolution of mean-CVaR risk measures and different scenario partition strategies. The upper bounds are obtained through the use of optimal solutions of group subproblems. Using these lower and upper bounds, we propose a solution algorithm for risk-averse mixed-integer multi-stage stochastic problems with mean-CVaR risk measures. We test the performance of the proposed algorithm on a multi-stage stochastic lot sizing problem and compare different choices of lower bounds and partition strategies. Comparison of the proposed algorithm to a commercial solver revealed that, on the average, the proposed algorithm yields 1.13% stronger bounds. The commercial solver requires additional running time more than a factor of five, on the average, to reach the same optimality gap obtained by the proposed algorithm.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

In risk-averse stochastic optimization problems, risk measures are used to assess the risk involved in the decisions made. Due to the structural properties of risk measures, risk-averse models are more challenging than their risk-neutral counterparts. The multi-stage risk-averse stochastic models are even more complicated due to their dynamic nature and excessive amount of decision variables. Both the risk-neutral and risk-averse multi-stage stochastic problems are non-convex when some of the decision variables are required to be integer valued. Therefore, the solution methods suggested for convex multi-stage stochastic problems cannot be used to solve these problems.

In this study, we consider risk-averse mixed-integer multi-stage stochastic problems with an objective function of dynamic mean conditional value-at-risk (mean-CVaR). Both CVaR and mean-CVaR are coherent measures of risk that have been used in the literature extensively (see, Rockafellar & Uryasev, 2002). Coherent measures of risk and their axiomatic properties are introduced in the pio-

neering paper by Artzner, Delbaen, Eber, and Heath (1999). Later, the theory of coherent risk measures is extended by Ruszczyński and Shapiro (2006a, 2006b), and references therein.

In a multi-stage decision horizon, risk involved in a stream of random outcomes is considered. Therefore, dynamic coherent risk measures are introduced to quantify the risk in multi-stage models (see, Artzner, Delbaen, Eber, Heath, and Ku, 2007; Kovacevic and Pflug, 2009; Pflug and Römisch, 2007; Ruszczyński and Shapiro, 2006a; 2006b, and references therein).

For the multi-stage stochastic optimization problems with dynamic measures of risk, some exact solution techniques are suggested under the assumption that the decision variables are continuous. These techniques, such as stochastic dual dynamic programming (SDDP), which is first suggested by Pereira and Pinto (1991) for risk-neutral problems and then extended to risk-averse problems by Shapiro (2011), Shapiro, Tekaya, da Costa, and Soares (2013), Philpott, de Matos, and Finardi (2013), and Lagrangian relaxation of nonanticipativity constraints suggested by Collado, Papp, and Ruszczyński (2012) rely on the convex structure of the problem, therefore, they cannot be used to find an exact solution when some of the decision variables are integer valued. On the other hand, these methods can be used to obtain lower bounds on the optimal value of multi-stage stochastic integer problems.

* Corresponding author.

E-mail address: ozlem.cavus@bilkent.edu.tr (Ö. Çavuş).

Bonnans, Cen, and Christel (2012) propose an extension of SDDP method for the risk-neutral problems with integer variables by relaxing the integrality requirements in the backward steps of the algorithm. Later, Bruno, Ahmed, Shapiro, and Street (2016) extend this approach to risk-averse integer problems. Zou, Ahmed, and Sun (2016) consider SDDP method to solve risk-neutral multi-stage mixed-integer problems with binary state variables. They prove that SDDP method provides an exact solution to the problem in finite number of iterations when the cuts satisfy some sufficient conditions. Similarly, Schultz (2003) uses Lagrangian relaxation of nonanticipativity constraints to obtain lower bounds within a branch-and-bound procedure for risk-neutral multi-stage problems with integer variables. However, these approaches rely on some restrictive assumptions. SDDP method requires stagewise independence of random process and the branch-and-bound procedure requires complete recourse assumptions. Therefore, they cannot be applicable to a wide range of problems.

A recent stream of research proposes an alternative way of obtaining bounds for mixed-integer multi-stage stochastic problems via a scenario tree decomposition. In that approach, the sample space is partitioned into subspaces called as groups, and the problem is solved for the scenarios in a group instead of the original sample space. These smaller problems are called as group subproblems. Sandikçi, Kong, and Schaefer (2013) propose a group subproblem approach for risk-neutral mixed-integer two-stage stochastic problems. They show that the expected value of the optimal values of group subproblems gives a lower bound on the optimal value of the original problem. Later, this approach is extended to the risk-neutral multi-stage problems by Sandikçi and Özalp (2014), Zenarosa, Prokopyev, and Schaefer (2014), and Maggioni, Allevi, and Bertocchi (2016). Recently, Maggioni and Pflug (2016) apply group subproblem approach to risk-averse mixed-integer multi-stage stochastic problems where the objective is a concave utility function applied to the total cost over the planning horizon. Although, group subproblems include less number of scenarios than the original problem, the length of the decision horizon in group subproblems and the original problem is the same. Therefore, one may argue that scalability is a drawback of this approach when the decision horizon is too long.

In this study, we propose a scenario tree decomposition algorithm for risk-averse mixed-integer multi-stage stochastic problems with a dynamic objective function defined via mean-CVaR. The suggested algorithm is based on group subproblem approach and is used to find lower and upper bounds on the optimal value of the problem. We propose infinitely many valid lower bounds on mean-CVaR risk measure that can be used within the frame of the algorithm. We also investigate the effect of scenario partitioning strategies on the quality of the different lower bounds by considering different partitioning strategies based on the structure of the scenario tree and disparateness of scenario realizations.

As outlined earlier, our approach does not require any special structural property such as convexity and linearity of the feasible set. Moreover, it does not require complete recourse or stagewise independence assumptions, therefore, it can be applied to a wide range of problems. We conduct computational experiments on a multi-stage lot sizing problem by considering different choices of bounds and scenario tree partitions. The experiments reveal that the obtained bounds are tight and require reasonable CPU times. Our approach yields 1.13% stronger bounds than solving the problem with IBM ILOG CPLEX. On the other hand, CPLEX requires more than 5.45 times of CPU time to obtain the same optimality gaps of our approach.

The organization of the paper is as follows: In Section 2, we present problem definition and some preliminaries. Section 3 includes our main results on obtaining different lower bounds for mean-CVaR via a scenario grouping approach. We consider the

application of these lower bounds to a risk-averse mixed-integer multi-stage stochastic problem with a dynamic objective function defined via mean-CVaR. We also suggest a method to obtain an upper bound. The computational study conducted on a multi-stage lot sizing problem and related discussions are presented in Section 4. Section 5 is devoted to concluding remarks and future research directions.

2. Risk-averse mixed-integer multi-stage stochastic problems with dynamic mean-CVaR objective

We consider a multi-stage discrete decision horizon where the decisions at stage $t \in \{1, \dots, T\}$ are made based on the available information up to that stage. Let Ω be a finite sample space and $\{0, \emptyset\} = \mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_T = \mathcal{F}$ be a filtration, that is, an ordered set of sigma algebras on Ω , representing gradually increasing information through stages. We use ξ_t and x_t to denote the vector of problem parameters and decisions at stage $t \in \{1, \dots, T\}$, respectively. For each $t \in \{1, 2, \dots, T\}$, ξ_t and x_t are \mathcal{F}_t -measurable. At first stage, the vector of problem parameters ξ_1 and decisions x_1 are deterministic, since $\mathcal{F}_1 = \{0, \emptyset\}$. At stage $t \in \{2, \dots, T\}$, some or all problem parameters are random.

An element ω of Ω is called as a scenario. A scenario $\omega \in \Omega$ corresponds to a realization of a sequence of random parameters $\xi_2(\omega), \dots, \xi_T(\omega)$ in stages 2, ..., T.

Our main interest is a risk-averse mixed-integer multi-stage stochastic problem with an objective of dynamic risk measure $\varrho_1, \tau(\cdot)$ over the horizon 1, ..., T. The problem can be defined as:

$$\min_{x \in \mathcal{X}} \varrho_{1,T}(f_1(x_1), f_2(x_2, \xi_2), \dots, f_T(x_T, \xi_T)), \quad (1)$$

where $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2(x_1, \xi_2) \times \dots \times \mathcal{X}_T(x_{T-1}, \xi_T)$ is the abstract representation of possibly nonlinear feasibility set. Let \mathbb{R} and \mathbb{Z} denote the set of real numbers and integers, respectively. $\mathcal{X}_1 \subseteq \mathbb{R}^{n_1} \times \mathbb{Z}^{m_1}$ is a mixed-integer deterministic set and, for $t \in \{2, \dots, T\}$, $\mathcal{X}_t : \mathbb{R}^{n_{t-1}} \times \mathbb{Z}^{m_{t-1}} \times \Omega \Rightarrow \mathbb{R}^{n_t} \times \mathbb{Z}^{m_t}$ are \mathcal{F}_t -measurable mixed-integer point-to-set mappings. The cost in the first stage is deterministic and represented by a possibly nonlinear, real-valued function $f_1 : \mathbb{R}^{n_1} \times \mathbb{Z}^{m_1} \rightarrow \mathbb{R}$. The cost functions $f_t : \mathbb{R}^{n_t} \times \mathbb{Z}^{m_t} \times \Omega \rightarrow \mathbb{R}$, $t \in \{2, \dots, T\}$ are \mathcal{F}_t -measurable, real-valued, and may be nonlinear.

Classical solution methods such as SDDP and Lagrangian relaxation of nonanticipativity constraints cannot be used to solve problem (1) due to integrality restrictions of some decision variables. Therefore, our focus is to obtain bounds on (1) where the objective function $\varrho_1, \tau(\cdot)$ is a dynamic risk measure defined via mean-CVaR.

Now, we present some necessary concepts and notation on coherent, conditional, and dynamic risk measures to exploit the structure of problem (1).

2.1. Coherent measures of risk

Let $\mathcal{Z} := \mathcal{L}_\infty(\Omega, \mathcal{F}, P)$ be the space of bounded and \mathcal{F} -measurable random variables with respect to sample space Ω and probability distribution P . Let $Z, W \in \mathcal{Z}$ represent uncertain outcomes for which lower realizations are preferable. Also, let Z_ω be the value that the random variable Z takes under scenario $\omega \in \Omega$. As defined in Artzner et al. (1999), a function $\rho : \mathcal{Z} \rightarrow \mathbb{R}$ is called a coherent measure of risk if it satisfies:

- (A1) Convexity: $\rho(\alpha Z + (1 - \alpha)W) \leq \alpha \rho(Z) + (1 - \alpha)\rho(W)$ for all $Z, W \in \mathcal{Z}$ and $\alpha \in [0, 1]$,
- (A2) Monotonicity: $Z \succeq W$ implies $\rho(Z) \geq \rho(W)$ for all $Z, W \in \mathcal{Z}$,
- (A3) Translational Equivariance: $\rho(Z + t) = \rho(Z) + t$ for all $t \in \mathbb{R}$ and $Z \in \mathcal{Z}$,
- (A4) Positive Homogeneity: $\rho(tZ) = t\rho(Z)$ for all $t > 0$ and $Z \in \mathcal{Z}$,

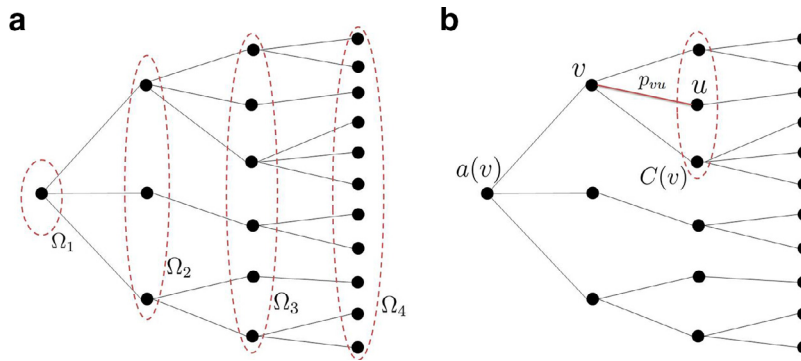


Fig. 1. An example of four-stage scenario tree. (a) Ω_1 , Ω_2 , Ω_3 and Ω_4 are the set of nodes at stages 1, 2, 3 and 4, respectively. (b) $C(v)$ is the set of children nodes of node v , $a(v)$ is the ancestor node of node v and p_{vu} is the conditional probability of node u given v .

where $Z \succeq W$ indicates pointwise partial ordering such that $Z_\omega \geq W_\omega$ for a.e. $\omega \in \Omega$.

We assume that cardinality of Ω is finite and \mathcal{F} is the set of all events defined on Ω . Then, the probability of a scenario $\omega \in \Omega$ can be specified as $p_\omega > 0$. In this case, elements of both \mathcal{Z} and its dual space \mathcal{Z}^* can be represented as elements of $\mathbb{R}^{|\Omega|}$ since both \mathcal{Z} and \mathcal{Z}^* are isomorphic to $\mathbb{R}^{|\Omega|}$.

Let μ_ω be the value that $\mu \in \mathcal{Z}^*$ takes under scenario $\omega \in \Omega$. For $Z \in \mathcal{Z}$ and $\mu \in \mathcal{Z}^*$, the scalar product $\langle \cdot, \cdot \rangle$ is defined as

$$\langle \mu, Z \rangle := \sum_{\omega \in \Omega} p_\omega \mu_\omega Z_\omega.$$

The following fact is known as dual representation of coherent measures of risk (see, [Ruszczyński and Shapiro, 2006b](#), for example): if $\rho(\cdot)$ is a coherent measure of risk, then, under some assumptions, for every random variable $Z \in \mathcal{Z}$,

$$\rho(Z) = \max_{\mu \in \mathcal{A}} \langle \mu, Z \rangle, \quad (2)$$

where $\mathcal{A} \subseteq \mathcal{Z}^*$ is a compact and convex set. We call this set as the dual set of the risk measure $\rho(\cdot)$. A coherent measure of risk can be characterized via its dual set. The reader is referred to [Ruszczyński and Shapiro \(2006b\)](#) for a detailed discussion on the dual representation of coherent measures of risk.

2.2. Conditional and dynamic risk measures

When a multi-stage stochastic process is considered, all realizations of the process form a scenario tree in the finite distribution case. In this section, we follow the notation used by [Collado et al. \(2012\)](#) to represent the scenario tree. Let Ω_t be the set of nodes at stage $t \in \{1, \dots, T\}$. At stage $t = 1$, there is only one node, called as root node and it is represented by v_1 . The nodes at stages $t \in \{2, \dots, T\}$ represent elementary events in \mathcal{F}_t , that is $\mathcal{F}_t = \sigma(\Omega_t)$, a sigma algebra on Ω_t .

The set Ω_T corresponds to all possible scenarios, that is $\Omega_T = \Omega$. Each node $v \in \Omega_t$, $t \in \{2, \dots, T\}$ has a unique ancestor at stage $t - 1$ and this ancestor node is called as $a(v)$. Also, each node $v \in \Omega_t$, $t \in \{1, \dots, T - 1\}$ has a set of children nodes $C(v)$ such that $C(v) = \{u \in \Omega_{t+1} : a(u) = v\}$. The probability measure P can be specified by conditional probabilities

$$p_{vu} := P[u|v], v \in \Omega_t, u \in C(v), t \in \{1, \dots, T - 1\},$$

and probability of a scenario $\omega \in \Omega_T$ can be computed as

$$p_\omega = p_{v_1 v_2} p_{v_2 v_3} \dots p_{v_{T-1} \omega},$$

where $v_1, v_2, \dots, v_{T-1}, \omega$ is the unique path from root node v_1 to node ω .

The notation mentioned above is depicted in [Fig. 1](#) for a four-stage scenario tree.

For a multi-stage decision horizon with stages $t \in \{1, \dots, T\}$, let $\mathcal{Z}_t := \mathcal{L}_\infty(\Omega, \mathcal{F}_t, P)$. The mapping $\rho_{\mathcal{F}_{t+1}|\mathcal{F}_t} : \mathcal{Z}_{t+1} \rightarrow \mathcal{Z}_t$ is called as one-step conditional risk measure if it satisfies axioms (A1)–(A4) for corresponding spaces \mathcal{Z}_t and \mathcal{Z}_{t+1} for all $t \in \{1, 2, \dots, T - 1\}$.

The risk involved in a sequence of random variables $Z_t \in \mathcal{Z}_t$, $t \in \{1, \dots, T\}$ adopted to the filtration \mathcal{F}_t , $t \in \{1, \dots, T\}$ can be evaluated by a time consistent dynamic measure of risk $\varrho_{1,T}(\cdot)$, that is,

$$\varrho_{1,T}(Z_1, Z_2, \dots, Z_T) = Z_1 + \rho_{\mathcal{F}_2|\mathcal{F}_1}(Z_2 + \rho_{\mathcal{F}_3|\mathcal{F}_2}(Z_3 + \dots + \rho_{\mathcal{F}_T|\mathcal{F}_{T-1}}(Z_T) \dots)). \quad (3)$$

The structure (3) is presented in [Ruszczyński and Shapiro \(2006a\)](#). Later, [Ruszczyński \(2010\)](#) shows that the representation (3) can be constructed using monotonicity of conditional risk measures and the concept of time consistency. A time consistent dynamic risk measure $\varrho_{1,T}(\cdot)$ is not, in general, law invariant, even in the case it is a composition of law invariant conditional risk measures (see, [Shapiro, 2012](#)). The reader is referred to [Ruszczyński \(2010\)](#) and [Eckstein, Eskandani, and Fan \(2016\)](#) for the definition of time consistency and law invariance.

[Collado et al. \(2012\)](#) show that the dual representation of coherent risk measures can be extended to dynamic measures of risk. If $\varrho_{1,T}(\cdot)$ is a dynamic risk measure given as in (3), then for every sequence of random variables $\{Z_t \in \mathcal{Z}_t\}_{t=1}^T$,

$$\varrho_{1,T}(Z_1, Z_2, \dots, Z_T) = \max_{q_T \in \mathcal{Q}_T} \langle q_T, Z_1 + Z_2 + \dots + Z_T \rangle, \quad (4)$$

where

$$\mathcal{Q}_T = \mathcal{A}_{T-1} \circ \dots \circ \mathcal{A}_2 \circ \mathcal{A}_1, \quad (5)$$

and \mathcal{A}_t , $t \in \{2, \dots, T\}$ is a convex and compact set used in the dual representation of $\rho_{\mathcal{F}_{t+1}|\mathcal{F}_t}(\cdot)$. The operator “ \circ ” defines convolution of probability measures, that is,

$$(\mu_t \circ q_t)(u) = q_t(a(u)) \mu_t(a(u), u), \forall u \in \Omega_{t+1},$$

and

$$\mathcal{A}_t \circ \mathcal{Q}_t = \{\mu_t \circ q_t : q_t \in \mathcal{Q}_t, \mu_t \in \mathcal{A}_t\},$$

for all $t \in \{1, 2, \dots, T - 1\}$. Recall that $a(u)$ is the ancestor node of u .

In this study, we use conditional mean-CVaR as one-step conditional risk measure. Therefore, the next section is devoted to the definition of mean-CVaR.

2.3. CVaR and mean-CVaR

An important and extensively used example of coherent measures of risk is Conditional Value-at-Risk (CVaR). CVaR of $Z \in \mathcal{Z}$ at level $\alpha \in [0, 1)$ is defined as (see, [Rockafellar & Uryasev, 2002](#))

$$\text{CVaR}_\alpha(Z) := \inf_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{1-\alpha} \mathbb{E}[(Z - \eta)_+] \right\}, \quad (6)$$

where $(a)_+ := \max\{a, 0\}$ for $a \in \mathbb{R}$.

Given a level parameter $\alpha \in [0, 1)$ and a weight parameter $\epsilon_1 \in [0, 1]$, mean-CVaR of $Z \in \mathcal{Z}$ is defined as

$$\rho(Z) := (1 - \epsilon_1) \mathbb{E}[Z] + \epsilon_1 \text{CVaR}_\alpha(Z). \quad (7)$$

As seen in (7), despite CVaR, mean-CVaR risk measure conveys the expected value information of a random variable, as well. As α or ϵ_1 increase, the decision-maker gets more risk-averse.

The expression in (7) can equivalently be represented as following linear program for finite probability spaces.

$$\begin{aligned} \rho(Z) = \underset{\vartheta, \eta}{\text{minimize}} \quad & (1 - \epsilon_1) \sum_{\omega \in \Omega} p_\omega Z_\omega + \epsilon_1 \left(\eta + \frac{1}{1-\alpha} \sum_{\omega \in \Omega} p_\omega \vartheta_\omega \right) \\ \text{subject to} \quad & \vartheta_\omega \geq Z_\omega - \eta, \quad \forall \omega \in \Omega \\ & \vartheta_\omega \geq 0, \quad \forall \omega \in \Omega. \end{aligned}$$

When the sample space is finite, the dual representation (2) holds for mean-CVaR with the set \mathcal{A} represented as (see, Ruszczyński & Shapiro, 2006b):

$$\mathcal{A} = \{\mu \in \mathcal{Z}^* : 1 - \epsilon_1 \leq \mu_\omega \leq 1 + \epsilon_2, \forall \omega \in \Omega \text{ and } \mathbb{E}[\mu] = 1\}, \quad (8)$$

where

$$\epsilon_2 := \frac{\alpha}{1-\alpha} \epsilon_1 \geq 0,$$

and $\mathbb{E}[\mu] = \sum_{\omega \in \Omega} p_\omega \mu_\omega$.

For any $Z_{t+1} \in \mathcal{Z}_{t+1}$, the one-step conditional mean-CVaR risk measure $\rho_{\mathcal{F}_{t+1}|\mathcal{F}_t}(Z_{t+1})$ with parameters $\alpha_t \in [0, 1)$ and $\epsilon_{1t} \in [0, 1]$ and its dual set \mathcal{A}_t are defined similar to (7) and (8). However, in (6), the infimum is over $\eta_t \in \mathcal{Z}_t$ and the expectation operators in (6)–(8) are replaced with conditional expectations with respect to \mathcal{F}_t .

For the remainder of the paper, we will focus on mean-CVaR risk measure. Hence, we will use $\rho(\cdot)$ to refer to mean-CVaR and $\rho_{\mathcal{F}_{t+1}|\mathcal{F}_t}(\cdot)$, $t \in \{1, 2, \dots, T-1\}$ to refer to one-step conditional mean-CVaR.

3. Bounds

The main motivation of this section is to propose lower and upper bounds for problem (1) with an objective of dynamic mean-CVaR. Therefore, in Section 3.1, we first propose a continuum of time consistent lower bounds for mean-CVaR risk measure by scenario grouping. Some possible lower bounds are presented in Section 3.2. The application of these time consistent bounds to a risk-averse mixed-integer multi-stage stochastic problems with an objective of (3) is presented in Section 3.3. Extension of the proposed lower bounds to other dynamic mean-CVaR risk measures is discussed in Section 3.4. In Section 3.5, we propose a method for obtaining an upper bound to the problem. The proposed algorithm benefits from these results and yields lower and upper bounds for problem (1).

3.1. Lower bounds for mean-CVaR risk measure

Let $\rho(\cdot)$ be a mean-CVaR risk measure with dual set \mathcal{A} . We would like to construct another coherent risk measure $\tilde{\rho}(\cdot)$ which provides a time consistent lower bound for $\rho(\cdot)$. The risk measure $\tilde{\rho}(\cdot)$, or equivalently its dual set $\tilde{\mathcal{A}}$, can be constructed in different ways. When the cardinality of the sample space is large, due to computational concerns, one may think of dealing with subsets of sample space separately and then obtaining a lower bound for $\rho(\cdot)$. For such construction, we need the definition of scenario groups and partition.

A subset of scenarios $S \subseteq \Omega$ is called as a group. Let $S = \{S_j\}_{j=1}^J$ be a collection of groups that forms a partition of Ω , that is, $\bigcup_{j=1}^J S_j = \Omega$ and $S_j \cap S_{j'} = \emptyset$ for all $j, j' \in \{1, 2, \dots, J\}$ such that $j \neq j'$. Note that the groups may not be necessarily disjoint (see, Sandıkçı & Özalpın, 2014), i.e. $S_j \cap S_{j'} \neq \emptyset$, but for the ease of representation, we partition the sample space into disjoint groups. Let \mathcal{G} be a σ -algebra generated by partition S where each group $S_j \in S$, $j \in \{1, 2, \dots, J\}$ corresponds to an elementary event j of \mathcal{G} . The probability of an elementary event j is $p_j = \sum_{\omega \in S_j} p_\omega$ which is the total probability of scenarios in S_j . We also define the adjusted probability of each scenario ω as $p_{j\omega} = p_\omega/p_j$ for all $\omega \in S_j$ and $j \in \{1, 2, \dots, J\}$. Note that, \mathcal{G} is a sub σ -algebra of \mathcal{F} .

Once a partition of the sample space Ω is given, one way to construct $\tilde{\rho}(\cdot)$ is to define it as a convolution of a coherent risk measure $\tilde{\rho}_{\mathcal{G}} : \mathcal{L}_\infty(\Omega, \mathcal{G}, P) \rightarrow \mathbb{R}$ with dual set $\tilde{\mathcal{A}}_{\mathcal{G}}$ and a one-step conditional risk measure $\tilde{\rho}_{\mathcal{F}|\mathcal{G}} : \mathcal{Z} \rightarrow \mathcal{L}_\infty(\Omega, \mathcal{G}, P)$ with dual set $\tilde{\mathcal{A}}_{\mathcal{F}|\mathcal{G}}$. That is, $\tilde{\rho}(\cdot) = (\tilde{\rho}_{\mathcal{G}} \circ \tilde{\rho}_{\mathcal{F}|\mathcal{G}})(\cdot)$, and its dual set is the convolution of the sets $\tilde{\mathcal{A}}_{\mathcal{G}}$ and $\tilde{\mathcal{A}}_{\mathcal{F}|\mathcal{G}}$ such that $\tilde{\mathcal{A}} = \tilde{\mathcal{A}}_{\mathcal{F}|\mathcal{G}} \circ \tilde{\mathcal{A}}_{\mathcal{G}}$.

Note that, $\tilde{\rho}_{\mathcal{F}|\mathcal{G}}(\cdot)$ can be represented in terms of $\rho_{S_j}(\cdot)$, $j \in \{1, 2, \dots, J\}$, that is, $[\tilde{\rho}_{\mathcal{F}|\mathcal{G}}(\cdot)]_j = \rho_{S_j}(\cdot)$ (see, Miller and Ruszczyński, 2011, for example) where $\rho_{S_j} : \mathcal{L}_\infty(\Omega, \sigma(S_j), P) \rightarrow \mathbb{R}$ is a coherent risk measure and $\sigma(S_j)$ is the σ -algebra on S_j . Fig. 2 depicts aforementioned notation for a given partition of a scenario tree with five scenarios.

For mean-CVaR, $\tilde{\rho}(\cdot)$ or equivalently its dual set $\tilde{\mathcal{A}}$, can be explicitly stated. Let parameters of $\tilde{\rho}_{\mathcal{G}}$ be $\alpha^1 \in [0, 1)$, $\epsilon_1^1 \in [0, 1]$, and $\epsilon_2^1 = \frac{\alpha^1}{1-\alpha^1} \epsilon_1^1$, and parameters of $\tilde{\rho}_{\mathcal{F}|\mathcal{G}}$ be $\alpha^2 \in [0, 1)$, $\epsilon_1^2 \in [0, 1]$ and $\epsilon_2^2 = \frac{\alpha^2}{1-\alpha^2} \epsilon_1^2$. Consider the convolution $\tilde{\rho} = \tilde{\rho}_{\mathcal{G}} \circ \tilde{\rho}_{\mathcal{F}|\mathcal{G}} : \mathcal{Z} \rightarrow \mathbb{R}$ and its dual set

$$\begin{aligned} \tilde{\mathcal{A}} &= \tilde{\mathcal{A}}_{\mathcal{F}|\mathcal{G}} \circ \tilde{\mathcal{A}}_{\mathcal{G}} = \{\mu \in \mathcal{Z}^* : \mu = \mu^1 \circ \mu^2, \mu_1 \in \tilde{\mathcal{A}}_{\mathcal{G}}, \mu_2 \in \tilde{\mathcal{A}}_{\mathcal{F}|\mathcal{G}}\} \\ &= \{\mu \in \mathcal{Z}^* : \mu = \mu^1 \circ \mu^2, 1 - \epsilon_1^1 \leq \mu_j^1 \leq 1 + \epsilon_2^1, \forall j \in 1, 2, \dots, J \\ &\text{and } \mathbb{E}[\mu^1] = 1, 1 - \epsilon_1^2 \leq \mu_\omega^2 \leq 1 + \epsilon_2^2, \forall \omega \in \Omega \text{ and } \mathbb{E}[\mu^2|\mathcal{G}] = \mathbf{1}\}, \end{aligned} \quad (9)$$

where $\mathbb{E}[\mu^1] = \sum_{j \in \{1, \dots, J\}} p_j \mu_j^1$, $[\mathbb{E}[\mu^2|\mathcal{G}]]_j = \sum_{\omega \in S_j} p_{j\omega} \mu_\omega^2$ for $j \in \{1, \dots, J\}$, and $\mathbf{1}$ is a \mathcal{G} -measurable random variable that takes value of one in all realizations. Construction of the set $\tilde{\mathcal{A}}$ for the example in Fig. 2 can be seen in Appendix A.

Now, we are ready to prove that a lower bound for mean-CVaR risk measure $\rho(\cdot)$ can be obtained by $\tilde{\rho}(\cdot) = (\tilde{\rho}_{\mathcal{G}} \circ \tilde{\rho}_{\mathcal{F}|\mathcal{G}})(\cdot)$.

Proposition 1. Let $\rho(\cdot)$ be a mean-CVaR risk measure with parameters $\alpha \in [0, 1)$, $\epsilon_1 \in [0, 1]$, $\epsilon_2 = \frac{\alpha}{1-\alpha} \epsilon_1 \geq 0$, and dual set \mathcal{A} . Also let $\tilde{\rho}(\cdot) = (\tilde{\rho}_{\mathcal{G}} \circ \tilde{\rho}_{\mathcal{F}|\mathcal{G}})(\cdot)$ where $\tilde{\rho}_{\mathcal{G}}$ is a mean-CVaR risk measure with parameters $\alpha^1 \in [0, 1)$, $\epsilon_1^1 \in [0, 1]$, $\epsilon_2^1 = \frac{\alpha^1}{1-\alpha^1} \epsilon_1^1$, and dual set $\tilde{\mathcal{A}}_{\mathcal{G}}$; and $\tilde{\rho}_{\mathcal{F}|\mathcal{G}}$ is a one-step conditional mean-CVaR risk measure with parameters $\alpha^2 \in [0, 1)$, $\epsilon_1^2 \in [0, 1]$, $\epsilon_2^2 = \frac{\alpha^2}{1-\alpha^2} \epsilon_1^2$, and dual set $\tilde{\mathcal{A}}_{\mathcal{F}|\mathcal{G}}$. Then, $\tilde{\rho}(Z) \leq \rho(Z)$ for all $Z \in \mathcal{Z}$ if

$$1 - \epsilon_1 \leq (1 - \epsilon_1^1)(1 - \epsilon_2^1) \text{ and } \left(1 + \frac{\alpha^1}{1-\alpha^1} \epsilon_1^1\right) \left(1 + \frac{\alpha^2}{1-\alpha^2} \epsilon_1^2\right) \leq 1 + \frac{\alpha}{1-\alpha} \epsilon_1. \quad (10)$$

Proof. Let $\mu \in \tilde{\mathcal{A}} = \tilde{\mathcal{A}}_{\mathcal{F}|\mathcal{G}} \circ \tilde{\mathcal{A}}_{\mathcal{G}}$. Then, from (9), there exist $\mu^1 \in \tilde{\mathcal{A}}_{\mathcal{G}}$ and $\mu^2 \in \tilde{\mathcal{A}}_{\mathcal{F}|\mathcal{G}}$ such that $\mu = \mu^1 \circ \mu^2$ with $\mathbb{E}[\mu^1] = 1$ and $\mathbb{E}[\mu^2|\mathcal{G}] = \mathbf{1}$. Properties of conditional expectation implies that $\mathbb{E}[\mu] = \mathbb{E}[\mathbb{E}[\mu|\mathcal{G}]] = \mathbb{E}[\mathbb{E}[\mu^1 \circ \mu^2|\mathcal{G}]] = \mathbb{E}[\mu^1 \circ \mathbb{E}[\mu^2|\mathcal{G}]] = \mathbb{E}[\mu^1 \circ \mathbf{1}] = \mathbb{E}[\mu^1] = 1$.

From the definition of ϵ_2, ϵ_1^1 and ϵ_2^2 , second part of (10) implies $(1 + \epsilon_2^1)(1 + \epsilon_2^2) \leq 1 + \epsilon_2$. Moreover, by (9), $(1 - \epsilon_1^1)(1 - \epsilon_2^1) \leq \mu_\omega \leq (1 + \epsilon_2^1)(1 + \epsilon_2^2)$ for all $\omega \in \Omega$. If $1 - \epsilon_1 \leq (1 - \epsilon_1^1)(1 - \epsilon_2^1)$

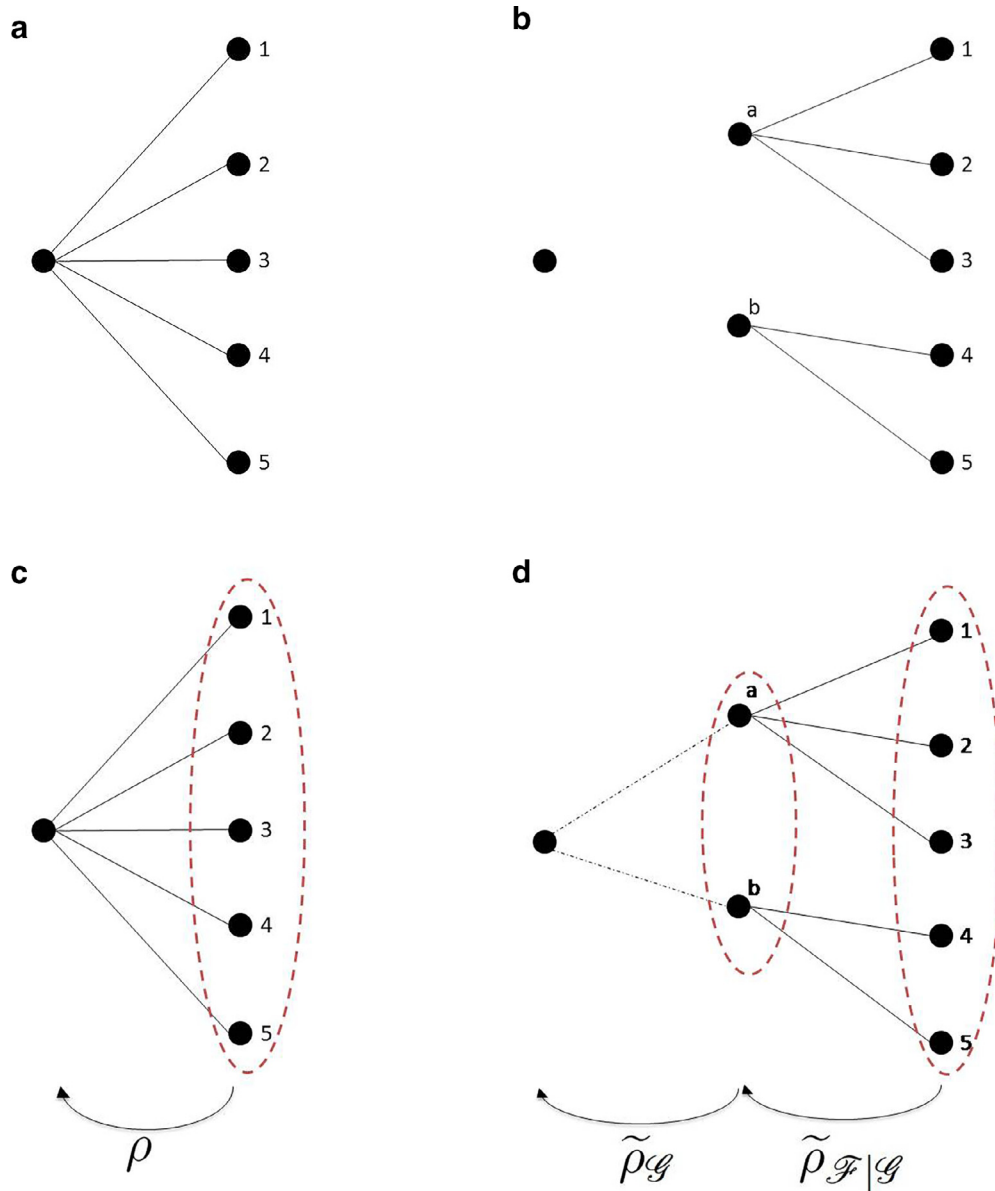


Fig. 2. (a) An example partition for a two-stage scenario tree: There are five scenarios 1,2,3,4, and 5 with probabilities p_1, p_2, p_3, p_4 , and p_5 , respectively. (b) $S = \{S_a, S_b\}$ is a partition of Ω where $S_a = \{1, 2, 3\}$ and $S_b = \{4, 5\}$. Nodes a and b correspond to groups S_a and S_b with probabilities $p_a = p_1 + p_2 + p_3$ and $p_b = p_4 + p_5$, respectively. (c) $\rho : \mathcal{Z} \rightarrow \mathbb{R}$ is the original risk measure. (d) \mathcal{G} is a sub σ -algebra of \mathcal{F} . $\tilde{\rho}_{\mathcal{G}} : \mathcal{L}_{\infty}(\Omega, \mathcal{G}, P) \rightarrow \mathbb{R}$ is a coherent risk measure and $\tilde{\rho}_{\mathcal{F}|\mathcal{G}} : \mathcal{Z} \rightarrow \mathcal{L}_{\infty}(\Omega, \mathcal{G}, P)$ is a one-step conditional risk measure that can be represented via $\rho_{S_a} : \mathcal{L}_{\infty}(\Omega, \sigma(S_a), P) \rightarrow \mathbb{R}$ and $\rho_{S_b} : \mathcal{L}_{\infty}(\Omega, \sigma(S_b), P) \rightarrow \mathbb{R}$ as $[\tilde{\rho}_{\mathcal{F}|\mathcal{G}}(\cdot)]_a = \rho_{S_a}(\cdot)$ and $[\tilde{\rho}_{\mathcal{F}|\mathcal{G}}(\cdot)]_b = \rho_{S_b}(\cdot)$.

$(1 - \epsilon_1^2)$ and $(1 + \epsilon_1^2)(1 + \epsilon_2^2) \leq 1 + \epsilon_2$, then $1 - \epsilon_1 \leq \mu_{\omega} \leq 1 + \epsilon_2$, for all $\omega \in \Omega$ which implies, $\mu \in \mathcal{A}$. Since μ is arbitrary, $\tilde{\mathcal{A}} \subseteq \mathcal{A}$.

For any $Z \in \mathcal{Z}$, let $\tilde{\rho}(Z) = \max_{\mu \in \tilde{\mathcal{A}}} \langle \mu, Z \rangle$ and $\mu^* \in \arg \max_{\mu \in \tilde{\mathcal{A}}} \langle \mu, Z \rangle$. If $\tilde{\mathcal{A}} \subseteq \mathcal{A}$, then $\mu^* \in \mathcal{A}$ and $\tilde{\rho}(Z) = \langle \mu^*, Z \rangle \leq \max_{\mu \in \mathcal{A}} \langle \mu, Z \rangle = \rho(Z)$. Since Z is arbitrary, $\tilde{\rho}(Z) \leq \rho(Z)$ for all $Z \in \mathcal{Z}$. \square

Proposition 1 partially extends Theorem 8 and Corollary 6 of Iancu, Petrik, and Subramanian (2015) to mean-CVaR risk measure. It implies that, under conditions (10), $\tilde{\rho}(\cdot) = (\tilde{\rho}_{\mathcal{G}} \circ \tilde{\rho}_{\mathcal{F}|\mathcal{G}})(\cdot)$ is a valid lower bound for $\rho(\cdot)$ for any partition S of Ω . If $\rho(\cdot)$ is a conditional mean-CVaR risk measure, Proposition 1 still applies. In this case, the expectations in the proof are replaced with corresponding conditional expectations.

3.2. Possible lower bounds

We have shown that a lower bound for $\rho(\cdot)$ can be obtained by convolutions of mean-CVaR risk measures whose parameters

satisfy condition (10). Due to Proposition 1, we can generate infinitely many lower bounds. Under the settings on Proposition 1, Table 1 presents some special cases of parameters of $\tilde{\rho}_{\mathcal{G}}(\cdot)$ and $\tilde{\rho}_{\mathcal{F}|\mathcal{G}}(\cdot)$ such that they can be used to obtain a lower bound for $\rho(\cdot)$.

Bounds $\rho_{\mathcal{G}} \circ \mathbb{E}_{\mathcal{F}|\mathcal{G}}$ and $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$ represent the extreme cases where either $\tilde{\rho}_{\mathcal{G}}(\cdot)$ or $\tilde{\rho}_{\mathcal{F}|\mathcal{G}}(\cdot)$ is the expectation operator. Bound $\rho_{\mathcal{G}}^s \circ \rho_{\mathcal{F}|\mathcal{G}}^s$ is an intermediate case where both $\tilde{\rho}_{\mathcal{G}}(\cdot)$ and $\tilde{\rho}_{\mathcal{F}|\mathcal{G}}(\cdot)$ have the same parameters, that is, $\alpha^1 = \alpha^2$, $\epsilon_1^1 = \epsilon_1^2$ and $\epsilon_2^1 = \epsilon_2^2$. Under these conditions, in order to construct the largest set $\tilde{\mathcal{A}}$, the inequalities in (10) are forced to hold at equality.

An interesting question is whether one of the possible lower bounds presented above is always preferable among others. Following example reveals that $\rho_{\mathcal{G}}^s \circ \rho_{\mathcal{F}|\mathcal{G}}^s$ is not necessarily the tightest bound among others.

Example 1. Consider a random variable Z with sample space $\Omega = \{\omega_i\}_{i=1}^4$. All four realizations have equal probabilities, that is, $p_{\omega_i} =$

Table 1Possible choices of $\tilde{\rho}_{\mathcal{G}}(\cdot)$ and $\tilde{\rho}_{\mathcal{F}|\mathcal{G}}(\cdot)$ that can be used to obtain lower bound on mean-CVaR risk measure $\rho(\cdot)$.

| $\tilde{\rho}_{\mathcal{G}} \circ \tilde{\rho}_{\mathcal{F} \mathcal{G}}$ | Parameters of $\tilde{\rho}_{\mathcal{G}}$ | | | Parameters of $\tilde{\rho}_{\mathcal{F} \mathcal{G}}$ | | |
|---|--|-----------------------------|---|--|-----------------------------|---|
| | ϵ_1^1 | ϵ_2^1 | α^1 | ϵ_1^2 | ϵ_2^2 | α^2 |
| $\rho_{\mathcal{G}} \circ \mathbb{E}_{\mathcal{F} \mathcal{G}}$ | ϵ_1 | ϵ_2 | α | 0 | 0 | 0 |
| $\rho_{\mathcal{G}}^s \circ \rho_{\mathcal{F} \mathcal{G}}^s$ | $1 - \sqrt{1 - \epsilon_1}$ | $\sqrt{1 + \epsilon_2} - 1$ | $\frac{\sqrt{1 + \epsilon_2} - 1}{\sqrt{1 + \epsilon_2} - \sqrt{1 - \epsilon_1}}$ | $1 - \sqrt{1 - \epsilon_1}$ | $\sqrt{1 + \epsilon_2} - 1$ | $\frac{\sqrt{1 + \epsilon_2} - 1}{\sqrt{1 + \epsilon_2} - \sqrt{1 - \epsilon_1}}$ |
| $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F} \mathcal{G}}$ | 0 | 0 | 0 | ϵ_1 | ϵ_2 | α |

Table 2

Values of different lower bounds (LB's) for Example 1.

| LB Choice | \mathcal{S} | \mathcal{S}' |
|---|---------------|----------------|
| $\rho_{\mathcal{G}} \circ \mathbb{E}_{\mathcal{F} \mathcal{G}}$ | 3.5 | 2.5 |
| $\rho_{\mathcal{G}}^s \circ \rho_{\mathcal{F} \mathcal{G}}^s$ | 3.12 | 3 |
| $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F} \mathcal{G}}$ | 3 | 3.5 |

1/4 for all $i \in \{1, 2, 3, 4\}$. The value that Z takes under scenario ω_i is i , that is, $Z_{\omega_i} = i$ for $i \in \{1, 2, 3, 4\}$.

Let $\epsilon_1 = 1$ and $\alpha = 0.5$, then (7) reduces to CVaR value at $\alpha = 0.5$ and then $\rho(Z) = 3.5$.

Two different partitions of scenarios are $\mathcal{S} = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\}$ and $\mathcal{S}' = \{\{\omega_1, \omega_4\}, \{\omega_2, \omega_3\}\}$. Values of the three bounds for partitions \mathcal{S} and \mathcal{S}' are given in Table 2.

As seen in Table 2, the tightest bounds for partitions \mathcal{S} and \mathcal{S}' are bounds $\rho_{\mathcal{G}} \circ \mathbb{E}_{\mathcal{F}|\mathcal{G}}$ and $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$, respectively. Another observation is the fact that $\rho_{\mathcal{G}}^s \circ \rho_{\mathcal{F}|\mathcal{G}}^s$ is not necessarily the tightest bound among others. In Example 1, although either $\rho_{\mathcal{G}} \circ \mathbb{E}_{\mathcal{F}|\mathcal{G}}$ or $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$ can be the tightest bound among others under different scenario partitions, the computational experiments in Section 4 reveal that $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$ is the most promising lower bound choice.

Although Shapiro, Dentcheva, and Ruszczyński (2009) show that, under some assumptions, the lower bound $\rho_{\mathcal{G}} \circ \mathbb{E}_{\mathcal{F}|\mathcal{G}}$ can be extended to any coherent risk measures, the other bounds provided in Table 1 may not be applicable for all coherent risk measures. Example 2 reveals that $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$ is not necessarily a valid lower bound for an arbitrary coherent risk measure.

Example 2. Consider a random variable Z that takes values $Z_{\omega_1} = 100$, $Z_{\omega_2} = 0$, $Z_{\omega_3} = 1$ and $Z_{\omega_4} = 500$ with probabilities 0.3, 0.2, 0.4 and 0.1, respectively. We use the first-order mean semi-deviation as a risk measure, that is:

$$\rho(Z) = \mathbb{E}[Z] + \kappa \mathbb{E}[(Z - \mathbb{E}[Z])_+], \quad \kappa \in [0, 1]. \quad (11)$$

Let $\rho_{\mathcal{F}|\mathcal{G}}(\cdot)$ be the one-step conditional first-order mean semi-deviation with the same κ value as in (11). Set $\kappa = 0.5$. For partition $\mathcal{S} = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\}$, $\rho(Z) = 104.32$ but $(\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}})(Z) = 106.36$.

Therefore, $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$ is not necessarily a valid lower bound for all coherent risk measures.

3.3. Lower bound for optimization problem

In this section, we extend the lower bound proposed in Proposition 1 to a risk-averse mixed-integer multi-stage stochastic problem with an objective of dynamic mean-CVaR risk measure. Using the structure presented in (3), the problem (1) can be written as

$$(\mathbf{P}) \min_{x_1 \in \mathcal{X}_1} f_1(x_1) + \rho(Q(x_1, \xi)), \quad (12)$$

where

$$Q(x_1, \xi) = \min_{x_t \in \mathcal{X}_t, t \in \{2, \dots, T\}} Q_{2,T}(f_2(x_2, \xi_2), \dots, f_T(x_T, \xi_T)), \quad (13)$$

$\xi = \{\xi_t\}_{t=2}^T$, $\rho(\cdot)$ is a mean-CVaR risk measure with parameters $\alpha \in [0, 1]$ and $\epsilon_1 \in [0, 1]$, and $Q_{2,T}(\cdot)$ is a dynamic mean-CVaR. Let x_1^* and z^* be an optimal first stage solution and the optimal value of (P), respectively.

Recall the partition $\mathcal{S} = \{S_j\}_{j=1}^J$ of Ω and sigma algebra \mathcal{G} induced by this partition. Then, the j th group subproblem is just problem (P) with sample space S_j and adjusted probabilities $p_{j\omega}$, $\omega \in S_j$. Additionally, the risk measure $\rho(\cdot)$ in (12) is replaced by $\rho_{S_j}(\cdot)$. For $j \in \{1, 2, \dots, J\}$, let z^j be the optimal value of j th group subproblem. Also let Z_{LB} be a \mathcal{G} -measurable random variable that takes value of z^j with probability $p_j = \sum_{\omega \in S_j} p_{j\omega}$.

In Theorem 1, we show that a lower bound for risk-averse mixed-integer multi-stage stochastic problem (P) can be obtained by using optimal values of group subproblems.

Theorem 1. Let $\tilde{\rho}_{\mathcal{G}} : \mathcal{L}_{\infty}(\Omega, \mathcal{G}, P) \rightarrow \mathbb{R}$ be a mean-CVaR risk measure with parameters $\alpha^1 \in [0, 1]$ and $\epsilon_1^1 \in [0, 1]$; and $\tilde{\rho}_{\mathcal{F}|\mathcal{G}} : \mathcal{L}_{\infty}(\Omega, \mathcal{F}, P) \rightarrow \mathcal{L}_{\infty}(\Omega, \mathcal{G}, P)$ be a conditional mean-CVaR risk measure with parameters $\alpha^2 \in [0, 1]$ and $\epsilon_1^2 \in [0, 1]$ satisfying $1 - \epsilon_1 \leq (1 - \epsilon_1^1)(1 - \epsilon_1^2)$ and $(1 + \frac{\alpha^1}{1 - \alpha^1} \epsilon_1^1)(1 + \frac{\alpha^2}{1 - \alpha^2} \epsilon_1^2) \leq 1 + \frac{\alpha}{1 - \alpha} \epsilon_1$. Then, $z^* \geq \tilde{\rho}_{\mathcal{G}}(Z_{LB})$.

Proof. Recall that x_1^* is an optimal first stage solution of (P). Note that, it is a feasible first stage solution for each group subproblem. By optimality of each group subproblem, we have

$$f_1(x_1^*) + \rho_{S_j}(Q(x_1^*, \xi)) \geq z^j, \quad \forall j \in \{1, \dots, J\}$$

and

$$f_1(x_1^*) + \tilde{\rho}_{\mathcal{F}|\mathcal{G}}(Q(x_1^*, \xi)) \geq Z_{LB}. \quad (14)$$

The values on the both sides of inequality (14) are \mathcal{G} -measurable. Since, $\rho_{\mathcal{G}}(\cdot)$ is a coherent risk measure and it satisfies monotonicity axiom (A2), we get

$$\tilde{\rho}_{\mathcal{G}}(f_1(x_1^*) + \tilde{\rho}_{\mathcal{F}|\mathcal{G}}(Q(x_1^*, \xi))) \geq \tilde{\rho}_{\mathcal{G}}(Z_{LB}). \quad (15)$$

Note that, $f_1(x_1^*)$ is an \mathcal{F} -measurable cost. Since \mathcal{G} is a sub σ -algebra of \mathcal{F} , $f_1(x_1^*)$ is \mathcal{G} -measurable, as well. Applying translational equivariance axiom (A3) to the left hand side of (15), we get

$$\tilde{\rho}_{\mathcal{G}}(\tilde{\rho}_{\mathcal{F}|\mathcal{G}}(f_1(x_1^*) + Q(x_1^*, \xi))) \geq \tilde{\rho}_{\mathcal{G}}(Z_{LB}). \quad (16)$$

Since conditions in (10) are satisfied, we can apply Proposition 1 to the left hand side of inequality (16) and obtain:

$$\rho(f_1(x_1^*) + Q(x_1^*, \xi)) \geq \tilde{\rho}_{\mathcal{G}}(Z_{LB}).$$

Finally, using translational equivariance axiom (A3), we get

$$z^* = f_1(x_1^*) + \rho(Q(x_1^*, \xi)) \geq \tilde{\rho}_{\mathcal{G}}(Z_{LB}).$$

□

Theorem 1 implies that a lower bound on the optimal value of (P) can be obtained by solving group subproblems and then applying $\tilde{\rho}_{\mathcal{G}}(\cdot)$ to the optimal values of these group subproblems. Since group subproblems include smaller number of scenarios compared to the original problem, they are computationally less challenging. Moreover, applying $\tilde{\rho}_{\mathcal{G}}(\cdot)$ to the optimal values of group subproblems requires negligible computational effort, since it is only the

calculation of value of a risk measure $\tilde{\rho}_{\mathcal{G}}(\cdot)$ for a given random cost.

Although the nested structure presented in (3) is widely used in the literature, there are other risk measures that can be used to evaluate the risk of a sequence of random variables. We show that our approach can also be applied to the risk-averse mixed-integer multi-stage stochastic problems with different dynamic extensions of mean-CVaR.

3.4. Extension to other dynamic measures of risk

Some examples of dynamic risk measures apart from the nested structure in (3) are multiperiod mean-CVaR and sum of mean-CVaR (see, Pflug and Römisch, 2007; Eichhorn and Römisch, 2005, respectively). For a sequence of random variables $Z_t \in \mathcal{Z}_t, t \in \{1, \dots, T\}$ adopted to the filtration $\mathcal{F}_t, t \in \{1, \dots, T\}$, multiperiod mean-CVaR is defined as

$$\rho^{multi}(\{Z_t\}_{t=2}^T) = \sum_{t=2}^T \lambda_t \mathbb{E}[\rho_{\mathcal{F}_t|\mathcal{F}_{t-1}}(Z_t)], \quad (17)$$

and sum of mean-CVaR is represented as

$$\rho^{sum}(\{Z_t\}_{t=2}^T) = \sum_{t=2}^T \lambda_t \rho_t(Z_t), \quad (18)$$

with $\sum_{t=2}^T \lambda_t = 1, \lambda_t \geq 0$ for $t \in \{2, 3, \dots, T\}$.

Our approach is also applicable for the case where the risk measure is applied to whole scenario cost as a time inconsistent objective function, that is,

$$\rho^{whole}(\{Z_t\}_{t=1}^T) = \rho(Z_1 + Z_2 + \dots + Z_T). \quad (19)$$

Although the risk measure (19) can be applied to a sequence of random variables, it is not a dynamic measure of risk.

The risk measure defined in (17) is a time consistent dynamic measure of risk whereas the risk measures (18) and (19) are not time consistent.

In the following three propositions, we show that a lower bound for these three risk measures can be obtained by scenario grouping. Therefore, our approach is still valid for Problem (P) with an objective of one of these risk measures.

Consider an arbitrary sequence of random variables $Z_t \in \mathcal{Z}_t, t \in \{1, \dots, T\}$ adopted to the filtration $\mathcal{F}_t, t \in \{1, \dots, T\}$. To avoid notational ambiguity, expectation operators and risk measures are given without reference sigma algebras.

Proposition 2. For a multiperiod mean-CVaR risk measure $\rho^{multi}(\cdot)$ as defined in (17), $\mathbb{E} \circ \rho^{multi}(\cdot)$ is a valid lower bound.

Proof. If multiperiod mean-CVaR risk measure (17) is applied to the sequence $Z_t \in \mathcal{Z}_t, t \in \{1, \dots, T\}$, then

$$\rho^{multi}(\{Z_t\}_{t=2}^T) = \sum_{t=2}^T \lambda_t \mathbb{E}[\rho_{\mathcal{F}_t|\mathcal{F}_{t-1}}(Z_t)].$$

Since $\rho_{\mathcal{F}_t|\mathcal{F}_{t-1}}(\cdot)$ is a conditional mean-CVaR risk measure, the lower bound $\mathbb{E} \circ \rho_{\mathcal{F}_t|\mathcal{F}_{t-1}}(\cdot)$ applies for $t \in \{2, 3, \dots, T\}$. Then,

$$\rho^{multi}(\{Z_t\}_{t=2}^T) \geq \sum_{t=2}^T \lambda_t \mathbb{E}[\mathbb{E}[\rho_{\mathcal{F}_t|\mathcal{F}_{t-1}}(Z_t)]].$$

Since expectation is a linear operator, we get

$$\rho^{multi}(\{Z_t\}_{t=2}^T) \geq \mathbb{E} \left[\sum_{t=2}^T \lambda_t \mathbb{E}[\rho_{\mathcal{F}_t|\mathcal{F}_{t-1}}(Z_t)] \right],$$

or equivalently,

$$\rho^{multi}(\{Z_t\}_{t=2}^T) \geq \mathbb{E}[\rho^{multi}(\{Z_t\}_{t=2}^T)].$$

Since the sequence $Z_t \in \mathcal{Z}_t, t \in \{1, \dots, T\}$ is arbitrary, the desired result follows. \square

Proposition 3. For a sum of mean-CVaR risk measure $\rho^{sum}(\cdot)$ as defined in (18), $\mathbb{E} \circ \rho^{sum}(\cdot)$ is a valid lower bound.

Proof. If sum of mean-CVaR risk measure (18) is applied to the sequence $Z_t \in \mathcal{Z}_t, t \in \{1, \dots, T\}$, then

$$\rho^{sum}(\{Z_t\}_{t=2}^T) = \sum_{t=2}^T \lambda_t \rho_t(Z_t).$$

Similarly, $\mathbb{E} \circ \rho_t(\cdot)$ applies for $t \in \{2, 3, \dots, T\}$. Then,

$$\rho^{sum}(\{Z_t\}_{t=2}^T) \geq \sum_{t=2}^T \lambda_t \mathbb{E}[\rho_t(Z_t)],$$

and

$$\rho^{sum}(\{Z_t\}_{t=2}^T) \geq \mathbb{E} \left[\sum_{t=2}^T \lambda_t \rho_t(Z_t) \right],$$

or equivalently,

$$\rho^{sum}(\{Z_t\}_{t=2}^T) \geq \mathbb{E}[\rho^{sum}(\{Z_t\}_{t=2}^T)].$$

Since the sequence $Z_t \in \mathcal{Z}_t, t \in \{1, \dots, T\}$ is arbitrary, the desired result follows. \square

Proposition 4. For the risk measure $\rho^{whole}(\cdot)$ as defined in (19), $\tilde{\rho}_{\mathcal{G}} \circ \tilde{\rho}_{\mathcal{F}|\mathcal{G}}(\cdot)$ is a valid lower bound if parameters of $\tilde{\rho}_{\mathcal{G}}(\cdot)$ and $\tilde{\rho}_{\mathcal{F}|\mathcal{G}}(\cdot)$ satisfy conditions in (10).

Proof. Follows from Proposition 1. \square

As shown above, our proposed lower bound is quite general and can be applied to other dynamic mean-CVaR measures.

3.5. Upper bound for optimization problem

Obtaining an upper bound, or equivalently finding a feasible solution of a minimization problem, is crucial for the instances where an optimal solution is not available. A good quality feasible solution gives the decision maker an action to be taken and measures the quality of obtained lower bound when an optimal solution is not available.

An upper bound for the optimal value of (P) can be obtained by using optimal solutions of group subproblems. Once j th group subproblem is solved, an optimal solution of it, namely x^j , is obtained. Let UB_j be the optimal value of (P) where (some of) the variables appearing in j th group subproblem are set to x^j . We call this problem as restricted problem. Since some of the problem variables are fixed, solving the restricted problem is easier than the original one and the resulting scenario tree can become decomposable.

If the restricted problem does not have a feasible solution, then corresponding upper bound UB_j is set to infinity. The best available upper bound UB is obtained by taking minimum of UB_j values over all $j \in \{1, \dots, J\}$, that is,

$$UB = \min_{j \in \{1, \dots, J\}} UB_j. \quad (20)$$

In Algorithm 1, we present how group subproblem approach can be used to obtain lower and upper bounds for the multi-stage risk-averse mixed-integer problem (P) with dynamic mean-CVaR objective. The algorithm can be easily adopted to the other risk measures given in Section 3.4.

Algorithm 1 Lower and upper bounds for (P).

Input: A risk-averse mixed-integer multi-stage stochastic problem (P) and a partition $S = \{S_j\}_{j=1}^J$ of sample space Ω .

Initialize: $LB \leftarrow -\infty$ and $UB \leftarrow +\infty$

Lower Bounding:

for all $j \in \{1, 2, \dots, J\}$ **do**

Solve the j th group subproblem.

$x^j \leftarrow$ an optimal solution of j th group subproblem

$z^j \leftarrow$ optimal value of j th group subproblem

end for

Let Z_{LB} be a random variable that takes value z^j with probability

$$p_j = \sum_{\omega \in S_j} p_\omega$$

$$LB \leftarrow \tilde{\rho}_{\mathcal{G}}(Z_{LB})$$

Upper Bounding:

for all $j \in \{1, 2, \dots, J\}$ **do**

$UB_j \leftarrow$ the optimal value of the original problem with the additional constraint where (some of) the variables appearing in j th group subproblem are set to x^j .

end for

$$UB \leftarrow \min_{j \in \{1, 2, \dots, J\}} UB_j$$

Return: LB and UB

4. Computational experiments

In this section, we conduct our numerical experiments on a multi-stage lot sizing problem studied in Guan, Ahmed, and Nemhauser (2009). All computational experiments are performed on an Intel(R) Core(TM) i7-4790 CPU@3.60 gigahertz computer with 8.00 gigabyte of RAM with Java 1.8.0.31 and IBM ILOG CPLEX 12.6. We first introduce risk-averse multi-stage lot sizing problem (RAMLSP) with dynamic mean-CVaR defined in (3). Then, we compare the results obtained via usage of different scenario partition strategies and lower bound choices. We also compare the proposed algorithm and CPLEX in terms of solution quality and required CPU time.

4.1. Risk-averse multi-stage lot sizing problem with mean-CVaR

The objective of RAMLSP is to minimize the dynamic mean-CVaR risk measure over T periods subject to demand satisfaction and capacity constraints. RAMLSP-T-r represents an RAMLSP instance with T stages in which random components can take r different values at each stage. Therefore, total number of scenarios in an RAMLSP-T-r instance is r^{T-1} . We generate random test instances as in Guan et al. (2009). The same setting of the parameters is also used by Sandikçi and Özalp (2014), that is, $h_{tu} \sim U[0, 10]$, $\alpha_{tu} \sim U[3.2, 4.8]E[h]$, $\beta_{tu} \sim U[320, 480]E[h]$, $d_{tu} \sim U[0, 100]$ and $M_{tu} \sim U[40T, 60T]$, where $U[a, b]$ represents uniform distribution between a and b .

Using the scenario tree representation given in Section 2.2, RAMLSP can be stated as follows:

$$\begin{aligned} \text{(RAMLSP)} \quad & \text{minimize} \quad Z_1 + \rho_{\mathcal{F}_2|\mathcal{F}_1}(Z_2 \\ & + \rho_{\mathcal{F}_3|\mathcal{F}_2}(Z_3 + \dots + \rho_{\mathcal{F}_T|\mathcal{F}_{T-1}}(Z_T) \dots)) \end{aligned} \quad (21)$$

subject to

$$Z_{tu} = \alpha_{tu}x_{tu} + \beta_{tu}y_{tu} + h_{tu}s_{tu}, \forall u \in \Omega_t, t \in \{1, \dots, T\} \quad (22)$$

$$s_{(t-1)a(u)} + x_{tu} = d_{tu} + s_{tu}, \quad \forall u \in \Omega_t, t \in \{1, \dots, T\} \quad (23)$$

$$x_{tu} \leq M_{tu}y_{tu}, \quad \forall u \in \Omega_t, t \in \{1, \dots, T\} \quad (24)$$

$$x_{tu}, s_{tu} \geq 0 \text{ and integer, } y_{tu} \in \{0, 1\}, \quad \forall u \in \Omega_t, t \in \{1, \dots, T\} \quad (25)$$

$$s_{0a(v_1)} = 0.$$

Here x_{tu} is the production level, y_{tu} is the setup indicator and s_{tu} is the inventory level variables at node $u \in \Omega_t$ in period $t \in \{1, \dots, T\}$. α_{tu} , β_{tu} , h_{tu} , d_{tu} and M_{tu} denote unit production cost, setup cost, inventory holding cost, demand and production capacity parameters, respectively. Z_1 is the sum of deterministic production, setup and inventory holding costs incurred in the first stage. Similarly, Z_{tu} is the cost incurred at node $u \in \Omega_t$ at stage $t \in \{2, \dots, T\}$. Z_t represents the random variable that takes values of Z_{tu} , $u \in \Omega_t$ with respective probabilities. The objective (21) is the dynamic risk value over the planning horizon. Constraint (22) calculate the cost incurred at each node of the scenario tree. Constraints (23) and (24) are inventory balance and capacity constraints, respectively. Constraint (25) are domain constraints. Unlike Guan et al. (2009) and Sandikçi and Özalp (2014), we assume that production and inventory levels are required to be integer valued. Although this assumption increases the problem complexity, we have a more realistic representation to evaluate the performance of the algorithm. In order to linearize RAMLSP, the linearization of mean-CVaR presented in Section 2.3 is used.

For the computational experiments, we use three different values of weight parameter $\epsilon_1 \in \{0.8, 0.5, 0.3\}$ and level parameter $\alpha \in \{0.9, 0.8, 0.7\}$ of mean-CVaR. Therefore, we have nine different risk-aversion settings.

4.2. Choices of scenario partitions and lower bounds

As seen in Example 1, the value of each lower bound highly depends on chosen scenario partition. We consider four possible scenario partition strategies obtained by grouping the scenarios in different ways, namely *index1*, *index2*, *similar* and *different*. For each strategy, we can also specify the number of scenarios in each group as a function of the number of scenarios $|\Omega|$ and the number of groups J . Let $a \% b$ be the remainder after the division of $a \in \mathbb{R}$ by $b \in \mathbb{R}$, $\lceil \cdot \rceil$ be the ceiling function, and $\lfloor \cdot \rfloor$ be the floor function. Then, each scenario grouping strategy yields a scenario partition that has J groups, where $\lceil |\Omega|/J \rceil$ groups have cardinality $\lceil |\Omega|/J \rceil$ and $J - (\lceil |\Omega|/J \rceil)$ groups have cardinality $\lfloor |\Omega|/J \rfloor$. For example, if $|\Omega| = 32$ and $J = 5$ then the cardinality of two groups will be seven and the other three groups will have cardinality of six.

Partition strategies *index1* and *index2* are based on the structure of scenario tree. In *index1*, the last stage nodes sharing the highest number of common nodes are placed into the same group. On the other hand, *index2* is obtained by placing the last stage nodes sharing the least number of common nodes into the same group.

If a priori information on the cost of each single scenario under an optimal solution is available, the groups can also be obtained with respect to similarity and diversity of individual scenarios. Since this information is not available before solving the original problem, the deterministic version of the original problem a priori can be solved for each scenario separately, and the corresponding single scenario costs can be used to obtain two different scenario partition strategies named as *similar* and *different*. Note that, for both strategies, an additional computational effort is required to obtain single scenario costs.

In strategy *similar*, we assign $\lceil |\Omega|/J \rceil$ scenarios that have the largest single scenario costs to the first group. Then, depending on the cardinality of the second group, $\lceil |\Omega|/J \rceil$ or $\lfloor |\Omega|/J \rfloor$ scenarios that have the second largest single scenario costs are assigned to the second group, and so on.

Table 3
Different scenario partitions $\mathcal{S} = \{S_1, S_2, S_3, S_4\}$ for the example scenario tree in Fig. 3.

| Partition strategy | S_1 | S_2 | S_3 | S_4 |
|--------------------|--|--|---|--|
| <i>index1</i> | $\{\omega_1, \omega_2, \omega_3, \omega_4\}$ | $\{\omega_5, \omega_6, \omega_7, \omega_8\}$ | $\{\omega_9, \omega_{10}, \omega_{11}, \omega_{12}\}$ | $\{\omega_{13}, \omega_{14}, \omega_{15}, \omega_{16}\}$ |
| <i>index2</i> | $\{\omega_1, \omega_5, \omega_9, \omega_{13}\}$ | $\{\omega_2, \omega_6, \omega_{10}, \omega_{14}\}$ | $\{\omega_3, \omega_7, \omega_{11}, \omega_{15}\}$ | $\{\omega_4, \omega_8, \omega_{12}, \omega_{16}\}$ |
| <i>Similar</i> | $\{\omega_9, \omega_4, \omega_{11}, \omega_6\}$ | $\{\omega_7, \omega_{16}, \omega_{13}, \omega_1\}$ | $\{\omega_{10}, \omega_2, \omega_8, \omega_3\}$ | $\{\omega_{12}, \omega_{14}, \omega_{15}, \omega_5\}$ |
| <i>Different</i> | $\{\omega_9, \omega_7, \omega_{10}, \omega_{12}\}$ | $\{\omega_4, \omega_{16}, \omega_2, \omega_{14}\}$ | $\{\omega_{11}, \omega_{13}, \omega_8, \omega_{15}\}$ | $\{\omega_6, \omega_1, \omega_3, \omega_5\}$ |

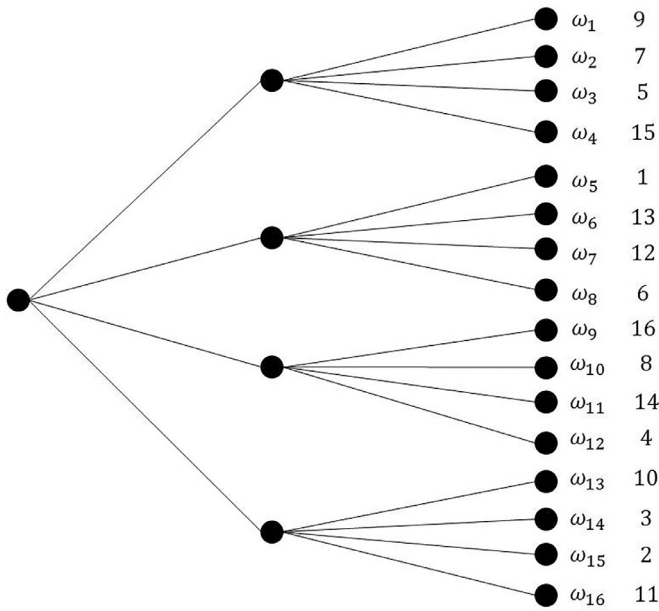


Fig. 3. An example of three-stage scenario tree with 16 scenarios.

In partition strategy *different*, each scenario is assigned to one of J groups by assigning the scenario with the largest single scenario cost to the first group, the scenario with the second largest single scenario cost to the second group, and so on. This assignment process returns to the first group after assigning the first J scenarios and the process restarts. It is ended after all scenarios are placed in a group.

With respect to single scenario cost values, in strategy *similar*, the dispersion within each group is low, however, the dispersion between the groups is high. On the other hand, in partition strategy *different*, the dispersion within each group is high.

Example 3. Fig. 3 depicts the scenario tree for an RAMLSP-3-4 instance where the numbers near the scenarios indicate the cost of each individual scenario. The scenarios can be ordered as $\omega_9, \omega_4, \omega_{11}, \omega_6, \omega_7, \omega_{16}, \omega_{13}, \omega_1, \omega_{10}, \omega_2, \omega_8, \omega_3, \omega_{12}, \omega_{14}, \omega_{15}, \omega_5$ where the individual scenario costs decrease moving through from ω_9 to ω_5 . Table 3 presents different scenario partition strategies for this scenario tree.

In order to observe the quality of bounds obtained by different scenario partition strategies and lower bound choices, the proposed algorithm is applied to five RAMLSP-3-30 instances generated via different random seeds. Total number of scenarios is 900. We consider the number of groups as $J \in \{2, 4, 10\}$, and hence each group subproblem includes 450, 225, and 90 scenarios, for the respective value of J . While obtaining upper bounds, optimal production decisions of group subproblems are fixed in the restricted problems. As noted before, for strategies *similar* and *different*, single scenario costs are required. The CPU time needed to obtain these values are also included in the running time of the algorithm. In order to measure the quality of

lower and upper bounds, an optimality gap information $Gap(\%) = 100((UB - LB)/UB)$ is used. All running times are reported in seconds. The results are presented in Table 4, where the *gap* and *time* values are the average values of five randomly generated instances.

The bolded entries in Table 4 correspond to the smallest optimality gap values among all lower bound choices, partition strategies and number of groups. It can be observed that, the smallest optimality gap values are obtained with strategy *different*, lower bound choice $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$, and $J = 2$. Regarding to the optimality gap, $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$ is the best lower bound choice. In general, for $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$, the partition strategy *different* provides the smallest optimality gap for any J value, and the strategy *similar* is the worst one. This is a consequence of the fact that group subproblems with original dynamic risk measure reflect the risk-aversion behavior of the original problem better when the dispersion within groups is high.

Moreover, the running time of the algorithm decreases as moving through lower bound choices $\rho_{\mathcal{G}} \circ \mathbb{E}_{\mathcal{F}|\mathcal{G}}$, $\rho_{\mathcal{G}}^S \circ \rho_{\mathcal{F}|\mathcal{G}}^S$ and $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$, in general. For example, with partition strategy *different* and $J = 2$, the average running times for lower bound choices $\rho_{\mathcal{G}} \circ \mathbb{E}_{\mathcal{F}|\mathcal{G}}$, $\rho_{\mathcal{G}}^S \circ \rho_{\mathcal{F}|\mathcal{G}}^S$ and $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$ are 57.18, 37.79 and 14.6 seconds, respectively. On the other hand, no partition strategy is preferable among others for all lower bound choices with respect to the running time. The computational experiments summarized in Table 4 reveal that the partition strategy *different* and the lower bound choice $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$ are the most promising choices when the bound quality and the running time are considered.

Although the upper bounds obtained from different partitions are incomparable, a hierarchy of lower bounds can be obtained using refinement chains. Moreover, the lower bound values can be improved by relaxing the requirement that groups should be disjoint. If this requirement is relaxed, some fixed scenarios appear in each group. We call this as scenario fixing. In the next subsection, we will discuss refinement chains, scenario fixing, and their impact on the quality of the lower bound.

4.3. Refinement chains and scenario fixing

Maggioni and Pflug (2016) consider refinement chains and scenario fixing to improve the quality of lower bounds obtained via scenario grouping. We suggest Proposition 5 to construct a refinement chain. We show a relation between two lower bounds obtained via two different special scenario partitioning with the lower bound choice $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$. Let $LB(S^1)$ and $LB(S^2)$ be the lower bounds $\mathbb{E}_{\mathcal{G}^1} \circ \rho_{\mathcal{F}|\mathcal{G}^1}$ and $\mathbb{E}_{\mathcal{G}^2} \circ \rho_{\mathcal{F}|\mathcal{G}^2}$ on the optimal value z^* of (P), obtained by partitions $S^1 = \{S_1^1, \dots, S_J^1\}$ and $S^2 = \{S_1^2, \dots, S_M^2\}$ where \mathcal{G}^1 and \mathcal{G}^2 are the sigma algebras induced by S^1 and S^2 , respectively.

Proposition 5. Let S^1 and S^2 be two different partitions of Ω , where $S_j^1 \cap S_{j'}^1 = \emptyset$ for $j, j' \in \{1, \dots, J\}$, $j \neq j'$ and $S_m^2 \cap S_{m'}^2 = \emptyset$ for $m, m' \in \{1, \dots, M\}$, $m \neq m'$. If for all $S_j^1 \in S^1$, $j \in \{1, \dots, J\}$ there exists $S_k^2 \in S^2$, $k \in \{j_1, \dots, j_{K_j}\}$ such that $S_j^1 = \bigcup_{k \in \{j_1, \dots, j_{K_j}\}} S_k^2$, then $LB(S^1) \geq LB(S^2)$.

Table 4

Average optimality gap and running time values of the proposed algorithm for five different RAMLSP-3-30 instances with different partition and lower bound choices.

| LB choice | α | ϵ_1 | J = 2 | | | | | | | | J = 4 | | | | | | | | J = 10 | | | | | | | | |
|--|--|--------------|---------------|------|---------------|-------|----------------|-------|------------------|-------|---------------|------|---------------|------|----------------|-------|------------------|------|---------------|------|---------------|-------|----------------|-------|------------------|-------|-----|
| | | | <i>index1</i> | | <i>index2</i> | | <i>Similar</i> | | <i>Different</i> | | <i>index1</i> | | <i>index2</i> | | <i>Similar</i> | | <i>Different</i> | | <i>index1</i> | | <i>index2</i> | | <i>Similar</i> | | <i>Different</i> | | |
| | | | Gap (%) | Time | Gap (%) | Time | Gap (%) | Time | Gap (%) | Time | Gap (%) | Time | Gap (%) | Time | Gap (%) | Time | Gap (%) | Time | Gap (%) | Time | Gap (%) | Time | Gap (%) | Time | Gap (%) | Time | |
| $\rho_{ig} \circ \mathbb{E}_{\mathcal{F} g}$ | 0.8 | 0.8 | 10.86 | 7.1 | 12.45 | 16.5 | 9.34 | 11.1 | 12.54 | 13.4 | 9.45 | 2.0 | 13.07 | 6.9 | 7.71 | 2.1 | 13.71 | 6.4 | 4.26 | 7.8 | 18.54 | 9.7 | 5.10 | 6.6 | 16.99 | 6.0 | |
| | | 0.9 | 0.5 | 7.64 | 24.5 | 8.81 | 24.8 | 7.92 | 29.6 | 8.67 | 23.9 | 6.85 | 3.1 | 9.36 | 8.6 | 7.97 | 2.5 | 9.78 | 8.5 | 3.63 | 9.7 | 11.93 | 8.6 | 7.01 | 7.2 | 12.73 | 5.4 |
| | | 0.3 | 5.32 | 18.2 | 5.79 | 52.6 | 6.56 | 95.3 | 5.69 | 48.2 | 4.91 | 4.1 | 6.33 | 13.4 | 8.29 | 3.7 | 6.73 | 11.7 | 3.01 | 11.4 | 8.34 | 7.9 | 8.32 | 7.3 | 8.28 | 7.3 | |
| | 0.8 | 0.8 | 9.61 | 13.2 | 10.82 | 21.9 | 7.16 | 14.2 | 10.97 | 24.0 | 8.14 | 3.3 | 11.69 | 9.6 | 5.32 | 3.7 | 11.58 | 10.1 | 3.46 | 8.2 | 15.55 | 10.3 | 6.89% | 8.9 | 14.43 | 5.5 | |
| | | 0.5 | 6.57 | 22.0 | 7.20 | 55.0 | 6.28 | 116.0 | 7.40 | 31.9 | 5.81 | 4.0 | 7.80 | 11.2 | 6.05 | 3.3 | 8.15 | 11.4 | 3.00 | 9.8 | 9.89 | 8.8 | 7.83 | 7.2 | 10.51 | 5.7 | |
| | | 0.3 | 4.44 | 26.6 | 4.75 | 88.6 | 5.46 | 118.9 | 4.88 | 99.2 | 4.12 | 5.0 | 5.21 | 14.9 | 7.01 | 4.8 | 5.56 | 13.8 | 2.59 | 12.2 | 6.74 | 8.0 | 8.68 | 7.5 | 6.71 | 7.5 | |
| | 0.7 | 0.8 | 8.86 | 21.2 | 10.11 | 74.9 | 5.35 | 24.2 | 10.10 | 44.4 | 7.60 | 4.0 | 10.49 | 12.0 | 5.33 | 4.5 | 10.51 | 12.6 | 3.85 | 7.7 | 13.09 | 9.1 | 7.82 | 7.6 | 12.70 | 7.1 | |
| | | 0.5 | 5.90 | 31.5 | 6.56 | 73.7 | 5.11 | 74.2 | 6.62 | 89.7 | 5.24 | 4.3 | 6.98 | 16.3 | 6.31 | 4.2 | 7.35 | 14.9 | 3.08 | 12.1 | 8.22 | 8.5 | 8.45 | 7.3 | 9.17 | 7.5 | |
| | | 0.3 | 3.95 | 36.4 | 4.27 | 114.0 | 4.73 | 144.6 | 4.29 | 139.9 | 3.70 | 5.6 | 4.50 | 17.0 | 7.08 | 5.4 | 4.80 | 17.2 | 2.48 | 14.5 | 5.50 | 8.5 | 8.91 | 7.8 | 5.81 | 10.6 | |
| | $\rho_{ig}^s \circ \rho_{\mathcal{F} g}^s$ | 0.9 | 0.8 | 4.52 | 7.2 | 5.52 | 14.8 | 7.57 | 8.3 | 5.10 | 9.6 | 4.38 | 1.6 | 5.99 | 6.1 | 9.83 | 1.8 | 6.03 | 6.3 | 4.20 | 6.7 | 8.95 | 4.4 | 14.27 | 6.8 | 8.76 | 5.2 |
| | | | 0.5 | 3.99 | 10.6 | 4.43 | 18.5 | 7.80 | 16.6 | 4.23 | 15.5 | 3.86 | 2.3 | 4.91 | 6.9 | 10.97 | 1.7 | 5.13 | 6.2 | 4.04 | 9.9 | 7.18 | 4.6 | 13.95 | 7.3 | 7.67 | 4.1 |
| | | | 0.3 | 3.08 | 20.4 | 3.14 | 21.5 | 6.57 | 66.6 | 3.02 | 20.1 | 3.04 | 4.0 | 3.56 | 7.4 | 10.47 | 3.2 | 3.79 | 10.6 | 3.38 | 10.5 | 5.31 | 6.0 | 12.83 | 7.4 | 5.18 | 4.3 |
| 0.8 | | 0.8 | 3.67 | 12.5 | 4.45 | 18.9 | 6.16 | 17.3 | 4.40 | 19.9 | 3.66 | 4.1 | 5.06 | 8.5 | 10.85 | 2.2 | 5.06 | 8.9 | 3.35 | 8.5 | 7.58 | 5.5 | 14.81 | 8.7 | 7.41 | 3.8 | |
| | | 0.5 | 3.30 | 12.9 | 3.52 | 22.0 | 6.68 | 34.7 | 3.40 | 59.3 | 3.12 | 3.7 | 4.01 | 10.2 | 9.91 | 2.8 | 4.09 | 9.6 | 3.09 | 9.2 | 5.94 | 5.4 | 13.86 | 7.4 | 5.98 | 4.8 | |
| | | 0.3 | 2.46 | 21.1 | 2.49 | 60.9 | 5.83 | 166.9 | 2.41 | 44.2 | 2.44 | 4.2 | 2.85 | 12.3 | 9.28 | 4.1 | 2.98 | 12.8 | 2.67 | 10.3 | 4.26 | 7.0 | 12.41 | 7.6 | 4.07 | 5.6 | |
| 0.7 | | 0.8 | 3.37 | 21.0 | 4.40 | 19.8 | 5.10 | 22.0 | 4.22 | 28.6 | 3.39 | 4.5 | 4.68 | 11.1 | 10.73 | 4.7 | 4.90 | 8.5 | 3.80 | 7.8 | 6.18 | 5.1 | 15.10 | 7.7 | 6.53 | 3.8 | |
| | | 0.5 | 2.94 | 25.1 | 3.26 | 46.0 | 6.09 | 52.5 | 3.10 | 47.1 | 2.82 | 6.1 | 3.62 | 14.7 | 10.17 | 3.4 | 3.63 | 13.7 | 3.06 | 10.2 | 4.72 | 5.0 | 13.80 | 7.3 | 5.12 | 4.9 | |
| | | 0.3 | 2.30 | 21.9 | 2.19 | 73.9 | 5.57 | 169.8 | 2.10 | 95.8 | 2.26 | 7.1 | 2.49 | 15.2 | 9.50 | 3.5 | 2.57 | 16.4 | 2.39 | 12.4 | 3.36 | 7.0 | 12.46 | 7.9 | 3.50 | 8.8 | |
| $\mathbb{E}_{ig} \circ \rho_{\mathcal{F} g}$ | 0.9 | 0.8 | 1.15 | 6.4 | 0.65 | 5.6 | 11.78 | 6.8 | 0.36 | 5.8 | 1.97 | 2.9 | 1.10 | 3.8 | 20.35 | 1.9 | 1.00 | 3.1 | 6.81 | 9.3 | 3.80 | 4.0 | 25.67 | 6.7 | 4.31 | 3.4 | |
| | | 0.5 | 1.26 | 4.8 | 0.42 | 11.9 | 9.54 | 5.6 | 0.26 | 9.1 | 1.82 | 2.4 | 0.82 | 4.3 | 16.15 | 1.6 | 0.94 | 4.4 | 5.09 | 7.4 | 2.99 | 3.1 | 20.73 | 7.3 | 2.92 | 3.1 | |
| | | 0.3 | 1.16 | 9.2 | 0.43 | 14.4 | 7.67 | 14.7 | 0.29 | 14.1 | 1.51 | 3.1 | 0.76 | 6.8 | 13.31 | 2.0 | 0.90 | 6.8 | 3.88 | 10.7 | 2.41 | 4.1 | 17.00 | 7.4 | 2.00 | 3.3 | |
| | 0.8 | 0.8 | 0.59 | 10.0 | 0.25 | 10.2 | 10.95 | 8.9 | 0.15 | 10.9 | 1.23 | 4.5 | 0.59 | 6.9 | 18.86 | 2.8 | 0.56 | 7.4 | 5.17 | 10.9 | 2.54 | 4.1 | 24.29 | 8.8 | 2.10 | 3.9 | |
| | | 0.5 | 0.64 | 11.3 | 0.27 | 15.3 | 8.74 | 16.8 | 0.12 | 12.3 | 1.07 | 3.2 | 0.59 | 6.2 | 14.79 | 1.9 | 0.52 | 7.0 | 3.78 | 9.2 | 2.08 | 3.6 | 19.49 | 7.3 | 1.89 | 3.2 | |
| | | 0.3 | 0.60 | 16.5 | 0.27 | 17.8 | 7.13 | 58.0 | 0.19 | 18.8 | 1.04 | 3.5 | 0.53 | 7.1 | 12.36 | 2.5 | 0.47 | 8.5 | 2.94 | 10.4 | 1.69 | 4.4 | 15.95 | 7.7 | 1.36 | 4.4 | |
| | 0.7 | 0.8 | 0.74 | 11.2 | 0.26 | 14.9 | 10.70 | 13.0 | 0.23 | 13.2 | 1.04 | 4.3 | 0.49 | 7.6 | 17.62 | 4.5 | 0.45 | 7.8 | 4.09 | 8.0 | 1.72 | 4.2 | 22.96 | 7.8 | 1.77 | 3.7 | |
| | | 0.5 | 0.63 | 16.9 | 0.26 | 14.7 | 8.47 | 22.0 | 0.22 | 18.2 | 0.87 | 3.4 | 0.51 | 7.3 | 14.15 | 3.0 | 0.47 | 8.3 | 3.02 | 9.9 | 1.35 | 3.4 | 18.68 | 7.2 | 1.52 | 3.1 | |
| | | 0.3 | 0.58 | 40.8 | 0.23 | 28.2 | 7.02 | 71.0 | 0.13 | 29.0 | 0.93 | 5.8 | 0.47 | 10.1 | 11.91 | 4.3 | 0.40 | 11.6 | 2.37 | 12.6 | 1.25 | 5.4 | 15.47 | 7.9 | 1.27 | 7.7 | |

Table 5A refinement chain for the scenario tree in Example 3 where the partition strategy is *different*.

| Partition | Groups |
|-----------|---|
| S_1 | Ω |
| S_2 | $\{\omega_9, \omega_{11}, \omega_7, \omega_{13}, \omega_{10}, \omega_8, \omega_{12}, \omega_{15}\}, \{\omega_4, \omega_6, \omega_{16}, \omega_1, \omega_2, \omega_3, \omega_5, \omega_{14}\}$ |
| S_4 | $\{\omega_9, \omega_7, \omega_{10}, \omega_{12}\}, \{\omega_4, \omega_{16}, \omega_2, \omega_{14}\}, \{\omega_{11}, \omega_{13}, \omega_8, \omega_{15}\}, \{\omega_6, \omega_1, \omega_3, \omega_5\}$ |
| S_8 | $\{\omega_9, \omega_{10}\}, \{\omega_4, \omega_2\}, \{\omega_{11}, \omega_8\}, \{\omega_6, \omega_3\}, \{\omega_7, \omega_{12}\}, \{\omega_{16}, \omega_{14}\}, \{\omega_{13}, \omega_{15}\}, \{\omega_1, \omega_5\}$ |
| S_{16} | $\{\omega_9\}, \{\omega_4\}, \{\omega_{11}\}, \{\omega_6\}, \{\omega_7\}, \{\omega_{16}\}, \{\omega_{13}\}, \{\omega_1\}, \{\omega_{10}\}, \{\omega_2\}, \{\omega_8\}, \{\omega_3\}, \{\omega_{12}\}, \{\omega_{14}\}, \{\omega_{15}\}, \{\omega_5\}$ |

Table 6Average lower bound gap and running time for the refinement chain $S_1, S_2, S_4, \dots, S_{128}$ obtained with partition strategy *different* for five different RAMSP-3-32 instances.

| α | 0.9 | | | | | | | | 0.8 | | | | | | | | 0.7 | | | | | | | |
|--------------|-------|---------------|------|--------------|--------|--------------|--------|---------------|------|--------------|--------|--------------|--------|---------------|-------|--------------|--------|--------------|--------|--------|------|--|--|--|
| ϵ_1 | 0.8 | | | | 0.5 | | | | 0.3 | | | | 0.8 | | | | 0.5 | | | | 0.3 | | | |
| Partition | # sce | LB_Gap | Time | LB_Gap | Time | LB_Gap | Time | LB_Gap | Time | LB_Gap | Time | LB_Gap | Time | LB_Gap | Time | LB_Gap | Time | LB_Gap | Time | LB_Gap | Time | | | |
| S_1 | 1024 | – | 81.3 | – | 3285.1 | – | 6503.9 | – | 90.1 | – | 4496.4 | – | 7199.0 | – | 169.3 | – | 3149.7 | – | 7272.5 | | | | | |
| S_2 | 512 | 0.20% | 3.6 | 0.30% | 7.2 | 0.28% | 11.5 | 0.34% | 7.1 | 0.27% | 12.3 | 0.23% | 16.4 | 0.28% | 8.7 | 0.24% | 46.1 | 0.26% | 104.0 | | | | | |
| S_4 | 256 | 0.82% | 1.3 | 0.94% | 1.2 | 0.80% | 2.3 | 0.76% | 1.8 | 0.64% | 2.3 | 0.59% | 3.3 | 0.58% | 1.9 | 0.65% | 3.8 | 0.63% | 5.7 | | | | | |
| S_8 | 128 | 2.60% | 0.4 | 2.46% | 0.46 | 1.99% | 0.6 | 1.93% | 0.8 | 1.75% | 0.6 | 1.47% | 0.9 | 1.66% | 0.6 | 1.59% | 1.0 | 1.44% | 1.5 | | | | | |
| S_{16} | 64 | 4.90% | 0.2 | 4.29% | 0.2 | 3.47% | 0.3 | 4.18% | 0.3 | 3.53% | 0.2 | 2.90% | 0.3 | 3.44% | 0.3 | 3.06% | 0.2 | 2.66% | 0.3 | | | | | |
| S_{32} | 32 | 6.81% | 0.1 | 5.99% | 0.1 | 4.99% | 0.1 | 6.37% | 0.1 | 5.39% | 0.1 | 4.47% | 0.1 | 5.84% | 0.1 | 4.93% | 0.1 | 4.14% | 0.1 | | | | | |
| S_{64} | 16 | 9.10% | 0.0 | 7.84% | 0.0 | 6.58% | 0.0 | 8.68% | 0.0 | 7.25% | 0.0 | 6.03% | 0.0 | 8.36% | 0.0 | 6.93% | 0.0 | 5.78% | 0.0 | | | | | |
| S_{128} | 8 | 11.25% | 0.0 | 9.58% | 0.0 | 8.10% | 0.0 | 10.89% | 0.0 | 9.03% | 0.0 | 7.53% | 0.0 | 10.38% | 0.0 | 8.64% | 0.0 | 7.23% | 0.0 | | | | | |

Proof. For $j \in \{1, \dots, J\}$ and $k \in \{j_1, \dots, j_{K_j}\}$, let z^j and \bar{z}^k be the optimal values of group subproblems defined by groups $S_j^1 \in S^1$ and $S_k^2 \in S^2$, respectively. Also let $p_j = \sum_{\omega \in S_j^1} p_\omega$ and $\bar{p}_k = \sum_{\omega \in S_k^2} p_\omega$.

Since $S_k^2, k \in \{j_1, \dots, j_{K_j}\}$ is a partition of S_j^1 , Theorem 1 implies that

$$z^j \geq \sum_{k \in \{j_1, \dots, j_{K_j}\}} \frac{\bar{p}_k}{p_j} \bar{z}^k, \quad \forall j \in \{1, \dots, J\},$$

where $\frac{\bar{p}_k}{p_j}$ is the total conditional probability of scenarios in $S_k^2, k \in \{j_1, \dots, j_{K_j}\}$ given S_j^1 . Then, we have

$$\sum_{j \in \{1, \dots, J\}} p_j z^j \geq \sum_{j \in \{1, \dots, J\}} p_j \left(\sum_{k \in \{j_1, \dots, j_{K_j}\}} \frac{\bar{p}_k}{p_j} \bar{z}^k \right),$$

or equivalently,

$$\sum_{j \in \{1, \dots, J\}} p_j z^j \geq \sum_{j \in \{1, \dots, J\}} \sum_{k \in \{1, \dots, K_j\}} \bar{p}_k \bar{z}^k.$$

Then, we get

$$\sum_{j \in \{1, \dots, J\}} p_j z^j \geq \sum_{m \in \{1, \dots, M\}} \bar{p}_m \bar{z}^m,$$

where $\bar{p}_m = \sum_{\omega \in S_m^2} p_\omega$ and \bar{z}^m is the optimal value of the group subproblem S_m^2 for $m \in \{1, \dots, M\}$. Hence, $LB(S^1) \geq LB(S^2)$ by definition of $LB(S^1)$ and $LB(S^2)$. \square

A sequence of partitions S^1, S^2, S^3, \dots , for which $LB(S^1) \geq LB(S^2) \geq LB(S^3) \geq \dots$, is called a refinement chain. Table 5 shows a refinement chain for the scenario tree in Example 3, where the partition strategy *different* is used and S_j denotes a partition with J groups. Note that, in partition S_{16} , each group subproblem is a deterministic problem with only one scenario.

We conduct a computational experiment where five different RAMSP-3-32 instances with 1024 scenarios are used with the refinement chain $S_1, S_2, S_4, \dots, S_{128}$ which is obtained with partition strategy *different*. Table 6 presents the number of scenarios in each group (# sce), the average lower bound gap (LB_Gap) and the average running time (Time) for each partition of the refinement chain.

LB_Gap values are calculated as

$$LB_Gap(\%) = 100 \frac{z^* - LB}{z^*},$$

where z^* is the optimal value of our problem.

As expected, the quality of the lower bound obtained by scenario grouping increases when the number of scenarios in each group subproblem increases with a cost of longer running times. Another suggestion of Maggioni and Pflug (2016) is relaxing the disjoint groups assumption to improve lower bound quality. Recall that $\omega \in \Omega$ is a scenario with probability p_ω and J is the number of groups. We can relax disjoint groups assumption by placing ω into $k \in \{2, \dots, J\}$ different groups. In this case, ω is replaced with k identical scenarios each having a probability of p_ω/k and these new scenarios are placed into k different groups.

We conduct another set of computational experiments for five different RAMSP-3-32 instances where eight scenarios with the largest single scenario cost appear at each group and we present the results in Table 7. It is observed that the quality of lower bounds improves when we allow some scenarios to appear in all groups. However, the running times may get larger since the number of scenarios in each group subproblem increases. For example, when $J = 2$, $\alpha = 0.7$, $\epsilon_1 = 0.3$, for the case where eight scenarios are fixed, the LB_Gap decreases by 0.01% but the required time to obtain lower bound increases to 2749.1 from 104 seconds.

4.4. Computational study results for larger number of stages

Up to now, we have shown that the lower bound choice $\mathbb{E}_{\mathcal{G}} \circ \rho_{\mathcal{F}|\mathcal{G}}$, the partition strategy *different*, and considering disjoint groups is the most promising combination among all bound and partition combinations. Therefore, further computational experiments are conducted on the instances with more stages under this setting. We also conduct a set of computational experiments to compare the performance of the proposed algorithm with CPLEX in terms of optimality gap and solution time.

In the upper bounding phase of the proposed algorithm, the restricted problem is solved for each group. When the number of groups J in a partition is large, the upper bounding phase requires long CPU times. Therefore, one may solve the restricted problem for only a subset of groups. Another computational enhancement for the upper bounding phase is running the restricted problems

Table 7
Average lower bound gap and running time for the refinement chain S_1, S_2, S_4, S_8 obtained with partition strategy *different* and eight fixed scenarios for five different RAMLSP-3-32 instances.

| α | | 0.9 | | | | 0.8 | | | | 0.7 | | | |
|--------------|-------|--------------|------|--------------|--------|--------------|--------|--------------|-------|--------------|--------|--------------|--------|
| ϵ_1 | | 0.8 | | 0.5 | | 0.3 | | 0.8 | | 0.5 | | 0.3 | |
| Partition | # sce | LB_Gap | Time | LB_Gap | Time | LB_Gap | Time | LB_Gap | Time | LB_Gap | Time | LB_Gap | Time |
| S_1 | 1024 | – | 81.3 | – | 3285.1 | – | 6503.9 | – | 90.13 | – | 4496.4 | – | 7199.0 |
| S_2 | 516 | 0.11% | 12.8 | 0.23% | 27.6 | 0.24% | 115.9 | 0.22% | 9.9 | 0.23% | 36.7 | 0.23% | 691.3 |
| S_4 | 262 | 0.44% | 1.2 | 0.70% | 2.1 | 0.66% | 3.3 | 0.59% | 1.9 | 0.56% | 3.1 | 0.54% | 4.4 |
| S_8 | 135 | 1.59% | 0.4 | 1.82% | 0.5 | 1.67% | 0.8 | 1.52% | 0.5 | 1.48% | 0.6 | 1.33% | 0.9 |

Table 8
Average optimality gap and running time values of the proposed algorithm for five different RAMLSP-3-30, RAMLSP-4-8, and RAMLSP-5-4 instances with partition strategy *different* and lower bound choice $\mathbb{E}_g \circ \rho_{\mathcal{F}|g}$.

| | | | J = 4 | | J = 8 | | J = 16 | | J = 32 | |
|-------------|-----|-----|---------|-------|---------|-------|---------|------|---------|------|
| α | | | Gap (%) | Time | Gap (%) | Time | Gap (%) | Time | Gap (%) | Time |
| RAMLSP-3-64 | 0.9 | 0.8 | 1.02 | 68.9 | 2.42 | 55.8 | 5.02 | 44.6 | 8.13 | 82.0 |
| | | 0.5 | 0.92 | 132.6 | 2.07 | 66.2 | 4.27 | 60.9 | 6.99 | 86.1 |
| | | 0.3 | 0.74 | 697.4 | 1.57 | 101.4 | 3.33 | 75.6 | 5.80 | 93.1 |
| | 0.8 | 0.8 | 0.71 | 27.5 | 1.67 | 13.0 | 3.82 | 26.3 | 7.10 | 85.2 |
| | | 0.5 | 0.66 | 97.4 | 1.55 | 25.4 | 3.19 | 39.3 | 6.03 | 82.4 |
| | | 0.3 | 0.60 | 376.1 | 1.20 | 49.6 | 2.53 | 55.6 | 4.76 | 86.5 |
| | 0.7 | 0.8 | 0.54 | 51.8 | 1.30 | 14.5 | 2.82 | 37.6 | 5.67 | 82.9 |
| | | 0.5 | 0.56 | 129.2 | 1.12 | 27.4 | 2.37 | 48.2 | 4.74 | 81.9 |
| | | 0.3 | 0.48 | 373.1 | 0.93 | 63.7 | 1.96 | 57.6 | 3.73 | 84.9 |
| RAMLSP-4-8 | 0.9 | 0.8 | 5.92 | 10.4 | 6.05 | 9.0 | 9.62 | 17.6 | 11.23 | 37.5 |
| | | 0.5 | 5.49 | 16.9 | 6.69 | 19.3 | 7.28 | 26.0 | 9.18 | 59.1 |
| | | 0.3 | 4.06 | 21.3 | 5.87 | 27.0 | 6.45 | 37.8 | 7.76 | 66.5 |
| | 0.8 | 0.8 | 3.59 | 13.9 | 5.03 | 11.2 | 6.69 | 18.8 | 9.62 | 50.0 |
| | | 0.5 | 3.64 | 18.6 | 5.53 | 25.5 | 6.35 | 33.2 | 8.26 | 63.5 |
| | | 0.3 | 3.40 | 22.3 | 5.28 | 29.3 | 5.86 | 40.9 | 7.48 | 62.0 |
| | 0.7 | 0.8 | 3.48 | 16.0 | 4.95 | 14.5 | 6.14 | 22.0 | 8.38 | 54.3 |
| | | 0.5 | 3.19 | 21.1 | 5.31 | 24.4 | 5.92 | 33.7 | 7.95 | 61.8 |
| | | 0.3 | 3.12 | 28.3 | 4.86 | 32.2 | 5.54 | 41.1 | 7.09 | 65.1 |
| RAMLSP-5-4 | 0.9 | 0.8 | 6.55 | 8.9 | 10.25 | 4.8 | 13.24 | 11.7 | 16.09 | 29.9 |
| | | 0.5 | 5.53 | 13.7 | 9.09 | 5.6 | 11.69 | 16.0 | 13.27 | 43.2 |
| | | 0.3 | 5.12 | 17.1 | 7.97 | 8.0 | 10.02 | 20.0 | 11.38 | 48.2 |
| | 0.8 | 0.8 | 6.83 | 9.9 | 10.78 | 4.7 | 13.40 | 11.3 | 16.16 | 29.6 |
| | | 0.5 | 5.70 | 13.6 | 9.48 | 6.2 | 11.81 | 14.9 | 13.40 | 43.7 |
| | | 0.3 | 5.28 | 16.4 | 8.27 | 8.6 | 10.08 | 17.4 | 11.53 | 50.1 |
| | 0.7 | 0.8 | 6.04 | 10.3 | 9.92 | 6.0 | 13.04 | 17.0 | 14.58 | 40.5 |
| | | 0.5 | 5.39 | 11.7 | 8.96 | 6.7 | 10.96 | 20.9 | 12.49 | 51.4 |
| | | 0.3 | 4.93 | 15.5 | 7.94 | 9.8 | 9.53 | 22.5 | 11.00 | 56.9 |

with a prespecified time limit and reporting the objective value of current incumbent solution as UB_j . Since, the optimal value of the restricted problem is an upper bound for the original problem, the objective value of any incumbent solution is also a valid upper bound.

We solve RAMLSP-3-64, RAMLSP-4-8, and RAMLSP-5-4 problems with 3, 4, and 5 stages, respectively, and for each risk setting, we generate five instances using different random seeds. The algorithm is applied with lower bound choice $\mathbb{E}_g \circ \rho_{\mathcal{F}|g}$ and the partition strategy *different*, where number of groups, J , takes values of 4, 8, 16, and 32. The number of restricted problems to be solved is $\lceil J/5 \rceil$, which are selected randomly. The time limit for each restricted problem is set to 10 seconds. The results are presented in Table 8.

As seen in Table 8, increasing the number of groups in the partition may not always yield CPU time saving. As J increases, the optimality gap increases, on the other hand, the CPU time may not always decrease. Specifically, when J is increased to 32 from 16, the CPU time increases in all of the instances. As the number of groups J increases, the subproblems get smaller in size. However, the number of group subproblems and the restricted problems to solve increases. Therefore, increasing the number of groups may

not always result in a decrease in the running time of the algorithm.

An interesting question is the comparison of the proposed algorithm with CPLEX in terms of optimality gap and CPU time. To make a fair comparison, we use RAMLSP-3-64 instances where CPLEX is run as long as it reaches to the optimality gap or the CPU time of the proposed algorithm.

When CPLEX is allowed to run with one hour of time limit, it cannot solve none of the instances optimally. Table 9 presents the comparison of the proposed algorithm with CPLEX for $J = 4$.

In Table 9, the column “Gap_CPLEX” corresponds to the optimality gap value reported by CPLEX when it is allowed to run as long as the running time of the proposed algorithm. Moreover, the values in the column “Gap difference” is measured as the difference between CPLEX gap and the gap obtained by the proposed algorithm. When CPLEX is allowed to run as long as the solution time of the proposed algorithm, the algorithm yields 1.13% stronger bounds on the average. For example, when $\alpha = 0.7$ and $\epsilon_1 = 0.8$, our algorithm terminates with an optimality gap of 0.54% within 51.8 seconds. CPLEX stops with an optimality gap of 1.93% within the same time limit, that is, the bounds obtained by our algorithm is 1.39% better than the bounds obtained by CPLEX.

Table 9
Comparison of optimality gaps and running times of the proposed algorithm with CPLEX.

| α | ϵ_1 | Proposed algorithm | | CPLEX | | | |
|----------|--------------|--------------------|-------|---------------|--------------------|------------|-------|
| | | Gap (%) | Time | Gap_CPLEX (%) | GAP difference (%) | Time_CPLEX | Delay |
| 0.9 | 0.8 | 1.02 | 68.9 | 2.76 | 1.74 | 248.3 | 3.60 |
| | 0.5 | 0.92 | 132.6 | 3.47 | 2.55 | 2366.7 | 17.84 |
| | 0.3 | 0.74 | 697.4 | 1.39 | 0.65 | 719.9 | 1.03 |
| 0.8 | 0.8 | 0.71 | 27.5 | 2.04 | 1.33 | 241.9 | 8.79 |
| | 0.5 | 0.66 | 97.4 | 1.48 | 0.82 | 403.9 | 4.15 |
| | 0.3 | 0.60 | 376.1 | 0.89 | 0.29 | 603.0 | 1.60 |
| 0.7 | 0.8 | 0.54 | 51.8 | 2.47 | 1.93 | 246.4 | 4.75 |
| | 0.5 | 0.56 | 129.2 | 1.13 | 0.57 | 444.2 | 3.44 |
| | 0.3 | 0.48 | 373.1 | 0.81 | 0.33 | 1445.6 | 3.87 |

In Table 9, the column “Time_CPLEX” corresponds to the time in seconds that CPLEX takes to reduce its gap to the level of the gap obtained by the proposed algorithm. Also, the values in the column “Delay” is measured as the ratio of Time_CPLEX to the running time of the proposed algorithm (Time). CPLEX requires 5.45 times longer running time to achieve the optimality gap of the proposed algorithm, on the average. For $\alpha = 0.9$ and $\epsilon_1 = 0.5$, CPLEX requires 2366.7 seconds to achieve the optimality gap of the proposed algorithm, that means CPLEX needs to spend more than 17 times of the running time of the proposed algorithm in order to reach this optimality gap. These results show that the proposed algorithm outperforms CPLEX with respect to both optimality gap and running time.

5. Conclusion

In this paper, we propose a group subproblem approach for risk-averse mixed-integer multi-stage stochastic problems with a dynamic risk measure defined by mean-CVaR. To the best of our knowledge, this is the first study where group subproblem approach is applied to a risk-averse problem with an objective of a dynamic risk measure. We show that infinitely many lower bounds on the optimal value of the problem can be obtained by using different convolution of mean-CVaR risk measures. An upper bound is obtained through the use of optimal solutions of group subproblems, as well. The results are tested by a computational study on a multi-stage lot sizing problem. The effect of partition strategies and lower bound choices on the optimality gap of the proposed algorithm is investigated. Possible computational enhancements such as refinement chains and scenario fixing are also considered.

It is revealed that, on the average, the optimality gap of the proposed algorithm is 1.13% stronger than the optimality gap of CPLEX within the same running time. By solving the original problem with CPLEX, the optimality gaps of our algorithm can be achieved with additional running time more than a factor of five.

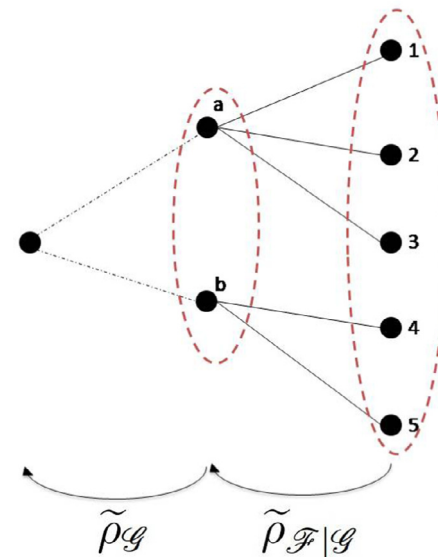
In the lower bounding phase of the proposed algorithm, the group subproblems can be assigned to different threads of a computer and solved in parallel. Similarly, parallel computing can be used to solve the restricted problems in the upper bounding phase. The parallel implementation of the proposed algorithm may decrease the running time significantly, especially for large J values. Therefore, it is considered as a future research direction. Another possible extension of the study is to find better scenario partitioning strategies. Finding optimal grouping strategy is still an interesting research direction.

Acknowledgment

The authors would like to thank the editor and two anonymous reviewers for their comments and suggestions that have improved the manuscript significantly.

Appendix A

Construction of set $\tilde{\mathcal{A}}$ for the example in Fig. 2.



$$\tilde{\mathcal{A}}_{\mathcal{G}} = \left\{ (\mu_a, \mu_b) \in \mathbb{R}^2 : \right.$$

$$1 - \epsilon_1^1 \leq \mu_a \leq 1 + \epsilon_2^1,$$

$$1 - \epsilon_1^1 \leq \mu_b \leq 1 + \epsilon_2^1,$$

$$(p_1 + p_2 + p_3)\mu_a + (p_4 + p_5)\mu_b = 1 \left. \right\}.$$

$$\tilde{\mathcal{A}}_{\mathcal{F}|\mathcal{G}} = \left\{ (\mu_1, \mu_2, \mu_3, \mu_4, \mu_5) \in \mathbb{R}^5 : \right.$$

$$1 - \epsilon_1^2 \leq \mu_1 \leq 1 + \epsilon_2^2,$$

$$1 - \epsilon_1^2 \leq \mu_2 \leq 1 + \epsilon_2^2,$$

$$1 - \epsilon_1^2 \leq \mu_3 \leq 1 + \epsilon_2^2,$$

$$1 - \epsilon_1^2 \leq \mu_4 \leq 1 + \epsilon_2^2,$$

$$1 - \epsilon_1^2 \leq \mu_5 \leq 1 + \epsilon_2^2,$$

$$\frac{p_1}{p_1 + p_2 + p_3} \mu_1 + \frac{p_2}{p_1 + p_2 + p_3} \mu_2 + \frac{p_3}{p_1 + p_2 + p_3} \mu_3 = 1, \\ \frac{p_4}{p_4 + p_5} \mu_4 + \frac{p_5}{p_4 + p_5} \mu_5 = 1 \left. \right\}.$$

$$\tilde{\mathcal{A}} = \tilde{\mathcal{A}}_{\mathcal{F}|\mathcal{G}} \circ \tilde{\mathcal{A}}_{\mathcal{G}} = \left\{ (\mu_a \mu_1, \mu_a \mu_2, \mu_a \mu_3, \mu_b \mu_4, \mu_b \mu_5) \in \mathbb{R}^5 : \right.$$

$$(\mu_a, \mu_b) \in \tilde{\mathcal{A}}_{\mathcal{G}}, (\mu_1, \mu_2, \mu_3, \mu_4, \mu_5) \in \tilde{\mathcal{A}}_{\mathcal{F}|\mathcal{G}} \left. \right\}.$$

References

- Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical finance*, 9(3), 203–228.
- Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., & Ku, H. (2007). Coherent multi-period risk adjusted values and Bellman's principle. *Annals of Operations Research*, 152(1), 5–22.
- Bonnans, J. F., Cen, Z., & Christel, T. (2012). Energy contracts management by stochastic programming techniques. *Annals of Operations Research*, 200(1), 199–222.
- Bruno, S., Ahmed, S., Shapiro, A., & Street, A. (2016). Risk neutral and risk averse approaches to multistage renewable investment planning under uncertainty. *European Journal of Operational Research*, 250(3), 979–989.
- Collado, R. A., Papp, D., & Ruszczyński, A. (2012). Scenario decomposition of risk-averse multistage stochastic programming problems. *Annals of Operations Research*, 200(1), 147–170.
- Eckstein, J., Eskandani, D., & Fan, J. (2016). Multilevel optimization modeling for risk-averse stochastic programming. *INFORMS Journal on Computing*, 28(1), 112–128.
- Eichhorn, A., & Römisch, W. (2005). Polyhedral risk measures in stochastic programming. *SIAM Journal on Optimization*, 16(1), 69–95.
- Guan, Y., Ahmed, S., & Nemhauser, G. L. (2009). Cutting planes for multistage stochastic integer programs. *Operations Research*, 57(2), 287–298.
- Iancu, D. A., Petrik, M., & Subramanian, D. (2015). Tight approximations of dynamic risk measures. *Mathematics of Operations Research*, 40(3), 655–682.
- Kovacevic, R., & Pflug, G. C. (2009). Time consistency and information monotonicity of multiperiod acceptability functionals. *Advanced Financial Modelling*, 8, 347.
- Maggioni, F., Allevi, E., & Bertocchi, M. (2016). Monotonic bounds in multi-stage mixed-integer stochastic programming. *Computational Management Science*, 13(3), 423–457.
- Maggioni, F., & Pflug, G. C. (2016). Bounds and approximations for multistage stochastic programs. *SIAM Journal on Optimization*, 26(1), 831–855.
- Miller, N., & Ruszczyński, A. (2011). Risk-averse two-stage stochastic linear programming: Modeling and decomposition. *Operations Research*, 59(1), 125–132.
- Pereira, M., & Pinto, L. M. (1991). Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52(1–3), 359–375.
- Pflug, G. C., & Römisch, W. (2007). *Modeling, measuring and managing risk*: Vol. 190. World Scientific.
- Philpott, A., de Matos, V., & Finardi, E. (2013). On solving multistage stochastic programs with coherent risk measures. *Operations Research*, 61(4), 957–970.
- Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7), 1443–1471.
- Ruszczynski, A. (2010). Risk-averse dynamic programming for markov decision processes. *Mathematical Programming*, 125(2), 235–261.
- Ruszczynski, A., & Shapiro, A. (2006a). Conditional risk mappings. *Mathematics of Operations Research*, 31(3), 544–561.
- Ruszczynski, A., & Shapiro, A. (2006b). Optimization of convex risk functions. *Mathematics of Operations Research*, 31(3), 433–452.
- Sandıkçı, B., Kong, N., & Schaefer, A. J. (2013). A hierarchy of bounds for stochastic mixed-integer programs. *Mathematical Programming*, 138(1–2), 253–272.
- Sandıkçı, B., & Özaltın, O. Y. (2017). A scalable bounding method for multi-stage stochastic integer programs. *SIAM Journal on Optimization*, 27(3), 1772–1800.
- Schultz, R. (2003). Stochastic programming with integer variables. *Mathematical Programming*, 97(1–2), 285–309.
- Shapiro, A. (2011). Analysis of stochastic dual dynamic programming method. *European Journal of Operational Research*, 209(1), 63–72.
- Shapiro, A. (2012). Time consistency of dynamic risk measures. *Operations Research Letters*, 40(6), 436–439.
- Shapiro, A., Dentcheva, D., & Ruszczyński, A. (2009). *Lectures on stochastic programming: modeling and theory*: Vol. 9. SIAM.
- Shapiro, A., Tekaya, W., da Costa, J. P., & Soares, M. P. (2013). Risk neutral and risk averse stochastic dual dynamic programming method. *European Journal of Operational Research*, 224(2), 375–391.
- Zenarosa, G.L., Prokopyev, O.A., & Schaefer, A.J. (2014). Scenario-tree decomposition: bounds for multistage stochastic mixed-integer programs. Technical Report. Pittsburgh, PA, USA: Technical paper, Department of Industrial Engineering, University of Pittsburgh.
- Zou, J., Ahmed, S., & Sun, X.A. (2016). Nested decomposition of multistage stochastic integer programs with binary state variables. Technical Report. School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA.