



Stage- t scenario dominance for risk-averse multi-stage stochastic mixed-integer programs

İ. Esra Büyüктаhtakın¹

Accepted: 27 October 2021 / Published online: 1 December 2021
© The Author(s) 2022

Abstract

This paper presents a new and general approach, named “Stage- t Scenario Dominance,” to solve the risk-averse multi-stage stochastic mixed-integer programs (M-SMIPs). Given a monotonic objective function, our method derives a partial ordering of scenarios by pairwise comparing the realization of uncertain parameters at each time stage under each scenario. Specifically, we derive bounds and implications from the “Stage- t Scenario Dominance” by using the partial ordering of scenarios and solving a subset of individual scenario sub-problems up to stage t . Using these inferences, we generate new cutting planes to tackle the computational difficulty of risk-averse M-SMIPs. We also derive results on the minimum number of scenario-dominance relations generated. We demonstrate the use of this methodology on a stochastic version of the mean-conditional value-at-risk (CVaR) dynamic knapsack problem. Our computational experiments address those instances that have uncertainty, which correspond to the objective, left-hand side, and right-hand side parameters. Computational results show that our “scenario dominance”-based method can reduce the solution time for mean-risk, stochastic, multi-stage, and multi-dimensional knapsack problems with both integer and continuous variables, whose structure is similar to the mean-risk M-SMIPs, with varying risk characteristics by one-to-two orders of magnitude for varying numbers of random variables in each stage. Computational results also demonstrate that strong dominance cuts perform well for those instances with ten random variables in each stage, and ninety random variables in total. The proposed scenario dominance framework is general and can be applied to a wide range of risk-averse and risk-neutral M-SMIP problems.

Keywords Stage- t scenario dominance · Partial ordering of scenarios · Time-consistent mean-risk multi-stage stochastic mixed-integer programs · Risk-averse · Conditional value-at-risk (CVaR) · Stochastic dynamic knapsack problem · Cutting planes · Bounds

✉ İ. Esra Büyüктаhtakın
esratoy@njit.edu

¹ Department of Mechanical and Industrial Engineering, New Jersey Institute of Technology, Newark, NJ, USA

1 Introduction

Risk is a fundamental issue arising in finance, insurance, project management, and many other areas. The analysis of risk in stochastic programming has recently been quite popular and raised many research questions, from model formulation to algorithmic approaches to tackle the problem difficulty, see, for instance, Schultz (2005), Schultz and Tiedemann (2006), Eichhorn and Römisch (2005), Ruszczyński (2013), and Yin and Büyüktaktakin (2021), among others. This paper addresses the computational difficulty of solving multi-stage stochastic mixed-integer programming problems (M-SMIPs) involving risk-averse objectives.

Stochastic mixed-integer programs can be formulated as an extensive form of a mixed-integer program (MIP), where a discrete stochastic process is represented by a finite number of realizations, i.e., scenarios. The size of the MIP grows exponentially in the number of decision stages and scenarios, leading to large-scale optimization formulations. Incorporating the risk-aversion criterion in addition to the expectation in the objective function further complicates the problem.

In this paper, we propose a new and general method of “stage- t scenario dominance” to effectively solve the risk-averse M-SMIPs. Our main idea is to derive cutting planes and bounds using the dominance characteristics of scenarios defined by a partial ordering and the solution of a new scenario sub-problem, which includes all the non-anticipativity and original constraints. Specifically, we solve the problem for a number of specific scenario sub-problems and use this information and the partial ordering of scenarios to generate cutting planes so that the general problem can be solved to optimality or close-to-optimality by a cut-and-branch algorithm. We demonstrate the computational benefit from using the proposed approaches on the risk-averse stochastic dynamic knapsack problem.

1.1 Literature review and paper contributions

Solution algorithms proposed for risk-averse multi-stage stochastic programs with integer variables are quite limited due to the non-convexity arisen by integrality. Most approaches proposed for risk-averse multi-stage stochastic optimization problems are extensions of the solution techniques proposed for their risk-neutral counterparts. Schultz and Tiedemann (2006) develop a Lagrangian decomposition algorithm that is based on the relaxation of non-anticipativity constraints to solve the scenario-based formulation of two-stage mixed-integer stochastic programming involving the conditional value-at-risk (CVaR). Heinze and Schultz (2008) later extended the scenario-decomposition algorithm of Schultz and Tiedemann (2006) combined with a branch and bound scheme to the case of multi-stage stochastic integer programs with CVaR risk objective. Yu et al. (2018) employ the Stochastic Dual Dynamic Integer Programming (SDDiP) approach and Lagrangian cuts of Zou et al. (2019) for solving the risk-averse multi-stage stochastic facility location problem using expected conditional risk measures (ECRM). Dentcheva and Martinez (2012) introduce Second-order Stochastic Dominance (SSD) constraints and explore the mathematical properties in two-stage stochastic problems. Escudero et al. (2016) study the first- and second-order stochastic dominance constraints induced by mixed-integer recourse in risk-averse multi-stage stochastic programs. Our scenario dominance approach is fundamentally different than the stochastic dominance constraints introduced in the literature (Müller and Stoyan 2002) in that we pairwise compare the scenarios based on the problem parameters and derive implications using the solution of scenario sub-problems, while in stochastic dominance random variables are compared based on their probability distribution function.

There has also been a stream of research studying bounds and approximation for risk-averse multi-stage stochastic mixed-integer programming problems (Maggioni and Pflug 2019; Mahmutođulları et al. 2018; Sandıkçı and Özaltn 2017; Bakir et al. 2020; Maggioni et al. 2014; Maggioni and Pflug 2016; Sandıkçı et al. 2013). Our focus in this paper is different from those by providing cutting plane algorithms based on the scenario dominance concept to tackle the computational difficulty of solving risk-averse multi-stage stochastic programs. Particularly, our scenario dominance-based cutting planes provide bounds on a single-scenario objective rather than the full objective function as considered in previous studies on bounds and approximations [e.g., see Maggioni and Pflug (2019) and Mahmutođulları et al. (2018)]. Furthermore, our cuts are driven based on a new scenario sub-problem, which is different from the classical scenario sub-problem that have been widely used in prior studies [see, e.g., Madansky (1960), Maggioni et al. (2016)].

Ruszczyński (2002) reformulates the two-stage stochastic programming problems with probabilistic constraints as large-scale mixed integer programming problems with knapsack constraints. The author uses the partial ordering of the scenarios to define precedence relationships between the variables of the knapsack problem and then defines the induced knapsack covers (Park and Park 1997; Boyd 1993) for the precedence-constrained knapsack polyhedra. Our definition of scenario dominance is similar to the ordering of scenarios that was discussed in Ruszczyński (2002); but the way we use the scenario ordering concept is fundamentally different than the approach in Ruszczyński (2002) because we compare the single scenario objectives of two scenarios based on a newly-defined scenario sub-problem to derive cuts for the risk-averse M-SMIPs, while Ruszczyński (2002) uses the ordering of scenarios to define induced knapsack covers. Furthermore, different than former work, our study, for the first time, defines the notion of “stage- t scenario dominance” and utilizes it for solving the risk-averse M-SMIPs more efficiently with tractable computational times.

Due to the limited number of available solution algorithms for the risk-averse multi-stage stochastic programs with integer variables, the main contribution of this paper is to provide a new approach named “scenario dominance” to solving risk-averse M-SMIPs. Specifically, we provide the following contributions: (i) presenting a new approach, named as “stage- t scenario dominance” that compares scenarios based on the realization of uncertain parameters under each scenario up to a certain time period t , derives a partial ordering of scenarios, and studies the implications of this ordering; (ii) defining a new scenario sub-problem, which involves all constraints of the original problem, including the non-anticipativity constraints, as opposed to the classical scenario sub-problem, which ignores the non-anticipativity relations among scenarios, as commonly used in the literature (Birge and Louveaux 2011; Lulli and Sen 2004); (iii) deriving and proving the scenario-based formulation for the mean-CVaR multi-stage stochastic mixed-integer program; (iv) demonstrating the use of “scenario dominance” theory to derive bounds and cutting planes for the risk-averse mean-CVaR M-SMIPs based on the solution of a new scenario sub-problem; (v) deriving results on the minimum number of stage- t scenario-dominance relations in multi-stage scenario trees, and thus cutting planes generated; (vi) presenting the computational benefit using the proposed “scenario dominance” cuts and bounds to solve the risk-averse dynamic stochastic knapsack problem by conducting computational experiments; and (vii) showing that “scenario dominance” cuts can significantly reduce the solution time for large-scale mean-risk stochastic multi-stage multi-dimensional knapsack problems with both integer and continuous variables, whose structure is similar to the mean-risk M-SMIPs. The proposed framework is general and can be applied to a wide range of risk-averse M-SMIP problems.

1.2 Risk-averse multi-stage stochastic mixed-integer programs

1.2.1 Risk-neutral multi-stage stochastic MIPs

We consider a finite number of stages $t \in \mathcal{T} := \{1, \dots, T\}$. Let ξ_1 be known, or deterministic, and (ξ_2, \dots, ξ_T) denotes a sequence of uncertain parameters where $\xi_{[t]} := \{\xi_1, \xi_2, \dots, \xi_t\}$ is the information observed by stage t . Let n_t, q_t, m_t , and k_t represent the number of decision variables, the number of integer variables, the number of constraints, and the number of random variables, respectively, at time $t \in \mathcal{T}$. Let \mathbb{R} denote the set of real numbers. We assume that ξ_2, \dots, ξ_T is a discrete-time stochastic process with a finite probability space (Ξ_t, F_t, P_t) , where the support, $\Xi_t \in \mathbb{R}^{k_t}$, represents a finite sample space for $t = 2, \dots, T$, and $\Xi = \Xi_T, F_t$ is the set of all events defined on Ξ_t , and P_t is the corresponding probability distribution. Let \mathbb{R}_+ and \mathbb{Z}_+ denote the set of positive real numbers and positive integers, respectively. Let $x := (x_1, x_2, \dots, x_T)$ represent the decision vector with $x_t \in \mathbb{R}_+^{n_t - q_t} \times \mathbb{Z}_+^{q_t}$. The sequence of decisions and the realizations of uncertain parameters are represented as $x_1, \xi_2, x_2(x_1, \xi_{[2]}), \xi_3, x_3(x_1, x_2, \xi_{[3]}), \dots, x_T(x_1, \dots, x_{T-1}, \xi_{[T]})$.

We define the multi-stage stochastic mixed-integer program (M-SMIP) to be of the nested form below:

$$\min_{x_1} \left\{ c_1^\top x_1 + \mathbb{E}_{\xi_2} [Q_2(x_1, \xi_{[2]})] : A_1 x_1 = b_1, x_1 \in \mathbb{R}_+^{n_1 - q_1} \times \mathbb{Z}_+^{q_1} \right\} \quad (1)$$

where $Q_t(\bullet, \bullet)$ is defined recursively for each $t = 2, \dots, T$, as

$$Q_t(x_{t-1}, \xi_{[t]}) = \min_{x_t} \{ c_t^\top(\xi_{[t]}) x_t + \mathbb{E}_{\xi_{t+1}|\xi_{[t]}} [Q_{t+1}(x_t, \xi_{[t+1]})] \} : \\ A_t(\xi_{[t]}) x_{t-1} + H_t(\xi_{[t-1]}) x_t = b_t(\xi_{[t]}), x_t \in \mathbb{R}_+^{n_t - q_t} \times \mathbb{Z}_+^{q_t},$$

where $c_1 \in \mathbb{R}^{n_1}$, $A_1 \in \mathbb{R}^{m_1 \times n_1}$, and $b_1 \in \mathbb{R}^{m_1}$ are known; and the uncertain parameters that realize as the stochastic process ξ evolves are given by $c_t(\xi_{[t]}) \in \mathbb{R}^{n_t}$, $A_t(\xi_{[t]}) \in \mathbb{R}^{m_t \times n_t}$, $H_t(\xi_{[t]}) \in \mathbb{R}^{m_t \times n_t}$, and $b_t(\xi_{[t]}) \in \mathbb{R}^{m_t}$; and $\mathbb{E}_{\xi_{t+1}|\xi_{[t]}}[\bullet]$ is the expectation with respect to ξ_{t+1} conditioned on the realization $\xi_{[t]}$, for $t = 2, \dots, T$. Note that for $T = 2$, the M-SMIP (1) is a two-stage SMIP.

Numerical optimization of multi-stage stochastic programs usually requires discrete probability distributions since otherwise, the multi-variate integrals involved become intractable (Heinze and Schultz 2008). So we consider a finite number of realizations of a discrete stochastic process $\xi = (\xi_1, \dots, \xi_T)$, which has a finite support $\Xi = \{\xi^1, \dots, \xi^{\bar{\Omega}}\}$ such that $|\Xi| = \bar{\Omega}$ for some positive integer $\bar{\Omega}$. Let ξ_1^ω be equal to the known parameter ξ_1 for each $\omega \in \Omega := \{1, 2, \dots, \bar{\Omega}\}$. A realization of a sequence of random parameters in stages $1, \dots, T$, $\xi^\omega = (\xi_1^\omega, \xi_2^\omega, \xi_3^\omega, \dots, \xi_T^\omega) \in \Xi$, is then referred as a scenario indexed by $\omega \in \Omega$. Each scenario $\xi^\omega \in \Xi$ is associated with probability p^ω , which is computed as the product of the conditional probabilities of all nodes that belong to the scenario path ξ^ω , such that $\sum_{\omega \in \Omega} p^\omega = 1$. The realizations of a scenario ξ^ω up to stage t is denoted by $\xi_{[t]}^\omega := \{\xi_1^\omega, \xi_2^\omega, \xi_3^\omega, \dots, \xi_t^\omega\}$.

When there is a finite number of scenarios, the M-SMIP (1) can be equivalently formulated as a large-scale deterministic program by replicating the decision variable in stage t , x_t , for each scenario $\xi^\omega \in \Xi$. Denoting x_t^ω as the decision variable for stage $t = 1, \dots, T$, the M-SMIP given in (1) can be equivalently formulated as the extensive form of the multi-stage stochastic MIP (E-MSMIP) (2a):

$$\min_{x_1^\omega, \dots, x_T^\omega} \sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T}} c_t^{\omega \top} x_t^\omega \quad (2a)$$

$$\text{s.t. } A_1 x_1^\omega = b_1, \quad \omega \in \Omega, \quad (2b)$$

$$A_t^\omega x_{t-1}^\omega + H_t^\omega x_t^\omega = b_t^\omega, \quad \forall t \in \mathcal{T} \setminus \{1\}, \quad \omega \in \Omega, \quad (2c)$$

$$x_t^\omega = x_t^{\omega'} \quad \forall t \in \mathcal{T}, \quad \omega, \omega' \in \Omega \quad \text{s.t.} \quad \xi_{[t]}^\omega = \xi_{[t]}^{\omega'}, \quad (2d)$$

$$x_t^\omega \in \mathbb{R}_+^{n_t - q_t} \times \mathbb{Z}_+^{q_t} \quad \forall t \in \mathcal{T}, \quad \omega \in \Omega, \quad (2e)$$

where the random parameters become known with $c_t^\omega \in \mathbb{R}^{n_t}$, $b_t^\omega \in \mathbb{R}^{m_t}$, $A_t^\omega \in \mathbb{R}^{m_t \times n_t}$, and $H_t^\omega \in \mathbb{R}^{m_t \times n_t}$ under scenario ω . The non-anticipativity constraints (2d) ensure that x_t^ω and $x_t^{\omega'}$ must attain the same values as long as the paths corresponding to scenarios ξ^ω and $\xi^{\omega'}$ coincide up to stage t (i.e., $\xi_{[t]}^\omega = \xi_{[t]}^{\omega'}$). Those non-anticipativity constraints ensure the real-time implementability of solutions.

1.2.2 Time consistency in multi-stage stochastic MIPs

The classical stochastic programming approach given in E-MSMIP (2a) may perform poorly if certain realizations of the stochastic parameters lead to an excessively high cost to accept. Thus, a mean-risk optimization formulation, which considers a risk-averse measure in addition to the expectation in the objective function, performs better than the expectation objective when the decision-making is at high-stake. In this paper, we focus on an asymmetric risk measure called Conditional Value-at-Risk (CVaR), which has been widely used as a risk measure in decision-making problems under uncertainty because it is a coherent measure of risk (see, for instance, see Rockafellar et al. (2000), Pflug (2000), and Ogryczak and Ruszczyński (2002) for more detailed discussions).

The incorporation of risk measures, such as CVaR, into the two-stage risk-averse stochastic program is straightforward (Ahmed 2006; Schultz and Tiedemann 2006; Schultz 2005; Fábíán 2008). However, the definition of risk aversion is not apparent in the multi-stage setting. Risk measures can be applied at every stage additively or to the complete scenario path, only defining risk at the end of the planning horizon or can be represented in a nested form. One of the most important considerations in modeling the risk-averse multi-stage stochastic programs is time consistency. Risk-neutral or risk-averse two-stage stochastic programs are time consistent, while time consistency for multi-stage risk-averse stochastic programs is not guaranteed. Based on the definition of Homem-de Mello and Pagnoncelli (2016), time consistency means that if you solve a multi-stage stochastic program today and find solutions for each node of a tree, you should find the same solutions if you re-solve the problem tomorrow given what was observed and decided today. Various researchers have studied how to address the time consistency issue in multi-stage risk-averse stochastic programs. For example, Shapiro et al. (2013) propose a nested conditional risk measure for multi-stage optimization problems and prove its time consistency. However, their nested conditional risk measure is formulated as Bellman equations (Bellman 1957; Büyüktaktın 2011), which cannot be solved in explicit form. Homem-de Mello and Pagnoncelli (2016) address this drawback by proposing a class of expected conditional risk measures (ECRMs) that is proven to be time consistent.

1.2.3 Mean-CVaR multi-stage stochastic MIP

Next, we analyze the incorporation of the multi-period CVaR in the objective function of the E-MSMIP (2a) and present the mean-CVaR multi-stage SMIP formulation, similar to the ECRM approach of Homem-de Mello and Pagnoncelli (2016), by casting the resulting risk-averse problem with ECRM as an extensive form that includes some additional variables

and risk linearization constraints. Our definition of the multi-period CVaR below is also consistent with other multi-period nested CVaR models in the literature (see, e.g., Heinze and Schultz 2008; Dupačová and Kozmík 2015; Alonso-Ayuso et al. 2018).

Definition 1.1 [*Multi-period CVaR* (Eichhorn and Römisch 2005; Heinze and Schultz 2008)] The Conditional Value-at-Risk of a random variable $c_t(\xi_{[t]})^\top x_t$ at confidence level $\alpha_t \in [0, 1)$ at each time stage $t \in \mathcal{T} \setminus \{1\}$ can be expressed as the optimal value of the following optimization problem:

$$CVaR_{\alpha_t} \left[c_t(\xi_{[t]})^\top x_t | \xi_{[t-1]} \right] = \min_{\eta_t \in \mathbb{R}} \left\{ \eta_t + \frac{1}{1 - \alpha_t} \mathbb{E} \left[\left(c_t(\xi_{[t]})^\top x_t - \eta_t \right)_+ | \xi_{[t-1]} \right] \right\}. \quad (3)$$

Proposition 1 Consider the following mean-CVaR problem:

$$\min_{x_t \in X_t} \left\{ \sum_{t \in \mathcal{T}} \mathbb{E} \left[c_t(\xi_{[t]})^\top x_t | \xi_{[t-1]} \right] + \lambda \sum_{t \in \mathcal{T} \setminus \{1\}} CVaR_{\alpha_t} \left[c_t(\xi_{[t]})^\top x_t | \xi_{[t-1]} \right] \right\}, \quad (4)$$

where $X_t(x_{[t-1]}, \xi_{[t]})$ denotes the feasibility set at stage t with decisions up to stage $t - 1$ ($x_{[t-1]}$) and the realization of the uncertain parameters up to stage t ($\xi_{[t]}$), and $\lambda \in \mathbb{R}_+$ is a trade-off coefficient that represents the relative weight of the risk objective ($CVaR_{\alpha_t}[\bullet]$) with respect to the expectation function ($\mathbb{E}[\bullet]$). Given the above notations and definitions, we can equivalently formulate the mean-CVaR problem (4), as the following extensive mean-CVaR multi-stage stochastic MIP (CVaR M-SMIP) problem (5a):

$$(P) \quad \min_{\substack{x_1^\omega, \dots, x_T^\omega, \\ \eta_1^\omega, \dots, \eta_T^\omega, \\ v_1^\omega, \dots, v_T^\omega}} \sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T}} c_t^{\omega \top} x_t^\omega + \lambda \sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right) \quad (5a)$$

$$s.t. \quad A_1 x_1^\omega = b_1, \quad \omega \in \Omega, \quad (5b)$$

$$A_t^\omega x_{t-1}^\omega + H_t^\omega x_t^\omega = b_t^\omega \quad \forall t \in \mathcal{T} \setminus \{1\}, \quad \omega \in \Omega, \quad (5c)$$

$$v_t^\omega \geq c_t^{\omega \top} x_t^\omega - \eta_t^\omega, \quad \forall t \in \mathcal{T} \setminus \{1\}, \quad \omega \in \Omega, \quad (5d)$$

$$x_t^\omega = x_t^{\omega'}, \quad v_t^\omega = v_t^{\omega'}, \quad \forall t \in \mathcal{T}, \quad \omega, \omega' \in \Omega \quad s.t. \quad \xi_{[t]}^\omega = \xi_{[t]}^{\omega'}, \quad (5e)$$

$$x_t^\omega \in \mathbb{R}_+^{n_t - q_t} \times \mathbb{Z}_+^{q_t}, \quad \forall t \in \mathcal{T}, \quad \omega \in \Omega, \quad (5f)$$

$$v_t^\omega \geq 0, \quad \eta_t^\omega \in \mathbb{R}, \quad \forall t \in \mathcal{T}, \quad \omega \in \Omega. \quad (5g)$$

The proof of Proposition 1 is given in Appendix A.

2 Scenario-dominance

In this section, we introduce a new concept of “scenario dominance” and “dominance sets,” considering the partial order relations of scenarios for the CVaR M-SMIP problem (5a). The partial ordering of scenarios is defined by comparing the scenario realizations of uncertain parameters. Using the partial ordering of scenarios and its implications based on the solution of a scenario sub-problem, we derive new bounds and cutting planes to improve the computational solvability of mean-CVaR M-SMIP formulation (5a).

2.1 Defining the scenario-dominance

Similar to the partial ordering of elements of a set, we define the partial ordering of scenarios in the scenario set Ξ . Because a scenario is defined as the realizations of a random variable over multiple stages $t \in \mathcal{T}$, the partial ordering of the scenarios is obtained by the point-wise comparison of scenario realizations at each stage t . Thus, in order to obtain a partial ordering of two scenarios $\xi^\omega \in \Xi$ and $\xi^{\omega'} \in \Xi$, the realization of scenario $\xi^\omega \in \Xi$ at each time period $t \in \mathcal{T}$, $(\xi_1^\omega, \xi_2^\omega, \xi_3^\omega, \dots, \xi_T^\omega)$ will be pairwise compared with the realization of scenario $\xi^{\omega'} \in \Xi$ for each t , $(\xi_1^{\omega'}, \xi_2^{\omega'}, \xi_3^{\omega'}, \dots, \xi_T^{\omega'})$. Note that $\xi_1^\omega = \xi_1^{\omega'}$ because ξ_1 is deterministic. This comparison leads to the following definition for a partial order relation of the scenarios.

Definition 2.1 *Scenario dominance* Define a scenario realization at time $t \in \mathcal{T}$ as ξ_t^ω , and let ξ_t^ω be comprised of random elements $(c_t^\omega, b_t^\omega, A_t^\omega, H_t^\omega)$ for $\omega \in \Omega$. Given any two scenarios ξ^ω and $\xi^{\omega'}$, scenario ξ^ω dominates scenario $\xi^{\omega'}$, denoted by $\xi^{\omega'} \leq \xi^\omega$, with respect to the problem P (5a) if

$$(p^\omega \geq p^{\omega'}) \wedge (c_t^\omega \geq c_t^{\omega'}) \wedge (b_t^\omega \geq b_t^{\omega'}) \wedge (A_t^\omega \leq A_t^{\omega'}) \wedge (H_t^\omega \leq H_t^{\omega'}) \quad \forall t \in \mathcal{T},$$

where \wedge denotes conjunction or ‘and’ operator, and the objective function (5a) portion for a single scenario index $\hat{\omega} = \omega, \omega'$ $(p^{\hat{\omega}} \sum_{t \in \mathcal{T}} c_t^{\hat{\omega}\top} x_t^{\hat{\omega}} + \lambda p^{\hat{\omega}} \sum_{t \in \mathcal{T} \setminus \{1\}} (\eta_t^{\hat{\omega}} + \frac{1}{1-\alpha_t} v_t^{\hat{\omega}}))$ is a non-decreasing function of $(x^{\hat{\omega}}, \eta^{\hat{\omega}}, v^{\hat{\omega}})$.

If, for example, $c_t^{\hat{\omega}} := (c_{1,t}^{\hat{\omega}}, \dots, c_{n_t,t}^{\hat{\omega}})$ and $A_t^{\hat{\omega}} := (A_{1,1,t}^{\hat{\omega}}, \dots, A_{m_t,n_t,t}^{\hat{\omega}})$, then the scenario domination is determined based on an element-wise comparison of the realizations of $c_t^{\hat{\omega}}$ and $A_t^{\hat{\omega}}$ for $\hat{\omega} = \omega, \omega'$, such as:

$$c_{j,t}^\omega \geq c_{j,t}^{\omega'} \quad \forall j = 1, \dots, n_t, \quad \forall t = 1, \dots, T,$$

$$A_{i,j,t}^\omega \leq A_{i,j,t}^{\omega'} \quad \forall i = 1, \dots, m_t, \quad \forall j = 1, \dots, n_t, \quad \forall t = 1, \dots, T.$$

We now generalize the ‘‘Scenario Dominance’’ definition above to the ‘‘Stage- t Scenario Dominance’’ where we compare the realization of scenario $\xi^\omega \in \Xi$ up to stage $t \in \mathcal{T}$, $(\xi_1^\omega, \xi_2^\omega, \xi_3^\omega, \dots, \xi_t^\omega)$ with the realization of scenario $\xi^{\omega'} \in \Xi$ up to stage t , $(\xi_1^{\omega'}, \xi_2^{\omega'}, \xi_3^{\omega'}, \dots, \xi_t^{\omega'})$ to derive the stage- t dominance.

Definition 2.2 *Stage- t Scenario dominance* Given two scenarios ξ^ω and $\xi^{\omega'}$, scenario ξ^ω stage- t dominates scenario $\xi^{\omega'}$ for $t \in \mathcal{T}$, denoted by $\xi_t^{\omega'} \leq \xi_t^\omega$, with respect to the problem P (5a) if

$$(p^\omega \geq p^{\omega'}) \wedge (c_j^\omega \geq c_j^{\omega'}) \wedge (b_j^\omega \geq b_j^{\omega'}) \wedge (A_j^\omega \leq A_j^{\omega'}) \wedge (H_j^\omega \leq H_j^{\omega'})$$

for $j = 1, \dots, t$,

and the objective function (5a) portion for a single scenario index $\hat{\omega} = \omega, \omega'$ $(p^{\hat{\omega}} \sum_{j=1}^t c_j^{\hat{\omega}\top} x_j^{\hat{\omega}} + \lambda p^{\hat{\omega}} \sum_{j=2}^t (\eta_j^{\hat{\omega}} + \frac{1}{1-\alpha_j} v_j^{\hat{\omega}}))$ is a non-decreasing function of $(x^{\hat{\omega}}, \eta^{\hat{\omega}}, v^{\hat{\omega}})$.

Using the definition of the ‘‘scenario dominance’’ above, we define the scenario-dominance sets as follows:

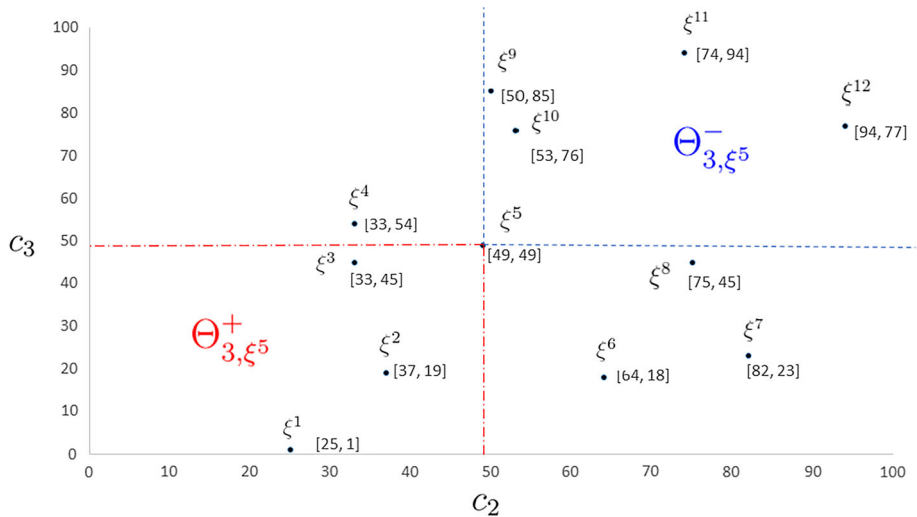


Fig. 1 Example of 12 scenario realizations of the c_t variable in two stages, where each black point represents a scenario $\xi^i := [c_2^i, c_3^i]$ for $i = 1, \dots, 12$. The Θ_{3,ξ^5}^+ represents the index set of scenarios that are stage-3 dominated by ξ^5 , Θ_{3,ξ^5}^- represents the index set of scenarios that stage-3 dominate ξ^5 , and N_{3,ξ^5} represents scenarios neither stage-3 dominate nor are stage-3 dominated by ξ^5

Definition 2.3 Stage- t Scenario-dominance sets Θ_{t,ξ^ω}^+ is the index set of scenarios, which are stage- t dominated by scenario $\xi^\omega \in \Xi$, Θ_{t,ξ^ω}^- is the index set of scenarios which stage- t dominate scenario $\xi^\omega \in \Xi$, and N_{t,ξ^ω} is the index set of scenarios, which neither stage- t dominate nor are stage- t dominated by $\xi^\omega \in \Xi$ for $t \in \mathcal{T}$ as defined below:

$$\begin{aligned} \Theta_{t,\xi^\omega}^+ &:= \left\{ \omega' \in \Omega, \omega' \neq \omega : \xi_t^{\omega'} \preceq \xi_t^\omega \right\}, \\ \Theta_{t,\xi^\omega}^- &:= \left\{ \omega' \in \Omega : \xi_t^\omega \preceq \xi_t^{\omega'} \right\}, \\ N_{t,\xi^\omega} &:= \left\{ \omega' \in \Omega : \xi_t^\omega \not\preceq \xi_t^{\omega'} \text{ and } \xi_t^{\omega'} \not\preceq \xi_t^\omega \right\}. \end{aligned}$$

Example 1 Figure 1 presents an example realization of the objective cost parameter, c_t , in two stages resulting in 12 different scenario realizations. To simplify the presentation, we omit c_1 in a scenario description and scenario-dominance comparisons since it takes the same value under all scenario realizations. Thus, here a scenario ξ^i represents a realization of c_t in stages 2 and 3, $[c_2^i, c_3^i]$. In the example, the scenario ξ^5 realization is [49, 49]. The index set of scenarios, which are stage-3 dominated by ξ^5 is defined as $\Theta_{3,\xi^5}^+ = \{1, 2, 3\}$, while the index set of scenarios, which stage-3 dominate scenario ξ^5 is given by $\Theta_{3,\xi^5}^- = \{5, 9, 10, 11, 12\}$. All other scenarios that neither dominate nor are dominated by ξ^5 is given by $N_{\xi^5} = \{4, 6, 7, 8\}$.

2.2 The number of scenario-dominance relations

The number of scenario-dominance relations depends on a number of factors, including the structure of the scenario tree, how the random variables are generated, and the number of

scenarios. Different methods have been proposed for generating a multi-period scenario tree (Heitsch and Römisch 2009; Pflug and Pichler 2015; Dupačová et al. 2000). Here we focus on the two most common types of multi-stage scenario trees based on the type of random variable generation, where random variables are completely independent and random variables are jointly distributed but stage-wise independent.

The number of scenario dominance cutting planes described in Sect. 2.4 relies on the number of scenario-dominance relations. Thus, next, we discuss the relations of two scenarios that share the same path up to a certain time period and exploit this result to compute the minimum number of scenario-dominance relations based on how the random variables are generated in a multi-stage scenario tree.

Proposition 2 *Let $\xi^\omega \in \Xi$ and $\xi^{\omega'} \in \Xi$ be two scenarios such that $\xi_{[t']}^\omega = \xi_{[t']}^{\omega'}$ for some $t' \in \mathcal{T}$. Then scenarios ξ^ω and $\xi^{\omega'}$ stage- t dominate each other, i.e., $\xi_t^{\omega'} \preceq \xi_t^\omega$ and $\xi_t^\omega \preceq \xi_t^{\omega'}$ for $t = 1, \dots, t'$.*

Proof It follows from non-anticipativity constraints (5e) that are ensured in a multi-stage stochastic program and the definition of the stage- t scenario dominance. \square

By Proposition 2, the scenarios that share the same path up to some stage t dominate each other in a scenario tree where there is at least one random variable at each stage. Thus, the minimum number of scenario-dominance relations could be defined by considering those dominance relations between scenarios that share a common history, motivating the following propositions.

Proposition 3 *Let (ξ_2, \dots, ξ_T) be a sequence of independently identically distributed (i.i.d) random vectors of k_t dimensions each. Each univariate component of a vector ξ_t is jointly distributed. In other words, components of ξ_t are positively or negatively correlated. In a scenario tree with l branches at each node in each stage, the minimum total number of scenario-dominance relations, W , is equal to:*

$$W = \sum_{t=1}^T l^{(2T-t-1)}. \tag{6}$$

Proof In this scenario tree, we have l^{T-1} scenarios. At stage $t' = 1$, we have 1 node, and thus l^{T-1}/l^0 scenarios share the same node at stage 1. At stage $t' = 2$, we have l nodes, and thus l^{T-1}/l^1 scenarios share the same node at stage 2. At stage $t' = 3$, we have l^2 nodes, and thus l^{T-1}/l^2 scenarios share the same node at stage 3. Thus, at stage $t' = t$, we have l^{t-1} nodes, and thus $l^{T-1}/l^{(t-1)}$ scenarios share the same node at stage t . Based on Proposition 2, for those scenarios that share the same path up to stage t dominate each other, there are $(l^{T-1}/l^{t-1})^2$ domination relations. Since l^{t-1} nodes are shared at stage t , there are at least a total of $l^{t-1} (l^{T-1}/l^{t-1})^2$ stage- t scenario-dominance relations. This results in a total number of $\sum_{t=1}^T l^{(2T-t-1)}$ scenario-dominance relations over all stages at the minimum. \square

Proposition 4 *Let (ξ_2, \dots, ξ_T) be a sequence of independently identically distributed (i.i.d) random vectors of $k_t = k$ dimensions each. Each univariate component of a vector ξ_t is independently distributed. In a scenario tree with l branches at each node in each stage, the minimum total number of scenario-dominance relations, W , is equal to:*

$$W = \sum_{t=1}^T l^{(2T-t-1)k}. \tag{7}$$

Proof In this scenario tree, we have $l^{(T-1)k}$ scenarios. At stage $t' = t$, we have $l^{(t-1)k}$ nodes, and thus $l^{(T-1)k_t} / l^{(t-1)k}$ scenarios share the same node at stage t , resulting in a total of $l^{(t-1)k} (l^{(T-1)k} / l^{(t-1)k})^2$ domination relations at stage t . This results in at least a total number of $\sum_{t=1}^T l^{(2T-t-1)k}$ scenario-dominance relations over all stages. \square

Example 2 Consider a CVaR-SMKP instance in which $T = 4$, $l = 2$, and $k_t = 2$ for $t = 1, \dots, T$. In the case where each univariate component of the random vector ξ_t are jointly distributed, we have $\bar{\Omega} = 8$ scenarios and a minimum number of 64, 32, 16, and 8 scenario-dominance relations in stages one, two, three, and four, respectively, resulting in a total of at least $W = 120$ scenario-dominance relations. In the case where each univariate component of the random vector ξ_t is independently distributed, we have $\bar{\Omega} = 64$ scenarios and a total of at least $W = 5440$ scenario-dominance relations at the minimum.

2.3 Scenario sub-problem and bounds

In this section, we define a new scenario sub-problem, which optimizes variables related to a subset of scenarios and assigns feasible solutions for all the other variables. This scenario sub-problem helps us derive new lower and upper bounds on the mean-CVaR multi-stage stochastic MIP problem. As a main contribution of the paper, using the solution of the scenario sub-problem and the scenario-dominance relations described in Sect. 2.1, we derive new cutting planes to chop off non-optimal points and find solutions faster to the highly-complex mean-CVaR multi-stage stochastic MIPs. To introduce the definitions and propositions below, let us define the feasible set for the original problem. The set of feasible solutions of the CVaR M-SMIP (5a) is denoted by \mathcal{X} as given below:

$$\mathcal{X} := \{(x, \eta, v) \in \mathbb{R}_+^{(n_1-q_1)\bar{\Omega}} \times \mathbb{Z}_+^{q_1\bar{\Omega}} \times \dots \times \mathbb{R}_+^{(n_T-q_T)\bar{\Omega}} \times \mathbb{Z}_+^{q_T\bar{\Omega}} \times \mathbb{R}^{T\bar{\Omega}} \times \mathbb{R}_+^{T\bar{\Omega}} : \text{Constraints (5b) – (5g)}\}.$$

Definition 2.4 (*Scenario group sub-problem*) Let $S \subseteq \Omega$. The scenario group sub-problem involving each $\omega \in S$, (P^S) , is formulated as follows:

$$Z^S := \min_{(x, \eta, v) \in \mathcal{X}} \sum_{\omega \in S} \left(p^\omega \sum_{t \in T} c_t^\omega \top x_t^\omega + \lambda p^\omega \sum_{t \in T \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right) \right). \tag{8}$$

Note that the problem (P^S) (8) is equivalent to the original problem P (5a), if $S = \Omega$.

Single scenario sub-problem The scenario group sub-problem for the set consisting of the single scenario, ξ^ω , is denoted as the scenario- ξ^ω problem $(P^{(\omega)})$ and formulated as follows:

$$Z^{(\omega)} := \min_{(x, \eta, v) \in \mathcal{X}} p^\omega \sum_{t \in T} c_t^\omega \top x_t^\omega + \lambda p^\omega \sum_{t \in T \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right). \tag{9}$$

Remark 1 The scenario- ξ^ω problem $(P^{(\omega)})$ is an MIP, whose feasible region comprises all the variables and all the constraints of the original problem (5a) while its objective function includes only one specific scenario with index ω . Consequently, the $P^{(\omega)}$ leads to a complete solution (x, η, v) , which is feasible for (5b)–(5g) but optimizes only for one scenario ω .

In the following proposition, we assume that the objective function for each scenario is non-negative, i.e., $f^\omega(x^\omega, \eta^\omega, v^\omega) = p^\omega \sum_{t \in T} c_t^\omega \top x_t^\omega + \lambda p^\omega \sum_{t \in T \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right) \geq 0$,

implying that $c_t^\omega \geq 0$ for each $\omega \in \Omega$ and $t \in \mathcal{T}$. Note that this assumption is not necessary when $S = \Omega$, since $Z^\Omega = Z$.

Proposition 5 *Let Z be the optimal objective value of the original problem P (5a). Then, P^S is a relaxation of P , that is $Z \geq Z^S \forall S \subseteq \Omega$.*

Proof Define a new variable $H \in \mathbb{R}$. The following problem is equivalent to the original problem P (5a):

$$Z := \min_{(x,\eta,v,H) \in \tilde{\mathcal{X}}} H,$$

where

$$\tilde{\mathcal{X}} = \mathcal{X} \cap \left\{ (x, \eta, v, H) : H \geq \sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T}} c_t^{\omega \top} x_t^\omega + \lambda \sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right) \right\}.$$

Similarly, the scenario group sub-problem P^S (8) can be equivalently written as in the following problem:

$$Z^S := \min_{(x,\eta,v,H) \in \hat{\mathcal{X}}} H,$$

where $\hat{\mathcal{X}} = \mathcal{X} \cap \left\{ (x, \eta, v, H) : H \geq \sum_{\omega \in S} p^\omega \sum_{t \in \mathcal{T}} c_t^{\omega \top} x_t^\omega + \lambda \sum_{\omega \in S} p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right) \right\}$. Clearly, $\tilde{\mathcal{X}} \subseteq \hat{\mathcal{X}}$, which proves the proposition. \square

Definition 2.5 *Stage- t single scenario sub-problem* The single scenario problem, including the variables and constraints up to stage t , for $t \in \mathcal{T}$, is denoted as the stage- t scenario- ξ^ω problem ($P_t^{\{\omega\}}$) and formulated as follows:

$$Z_t^{\{\omega\}} := \min_{(x,\eta,v) \in \mathcal{X}} p^\omega \sum_{j=1}^t c_j^{\omega \top} x_j^\omega + \lambda p^\omega \sum_{j=2}^t \left(\eta_j^\omega + \frac{1}{1 - \alpha_j} v_j^\omega \right). \tag{10}$$

Definition 2.6 *(Relaxed scenario group sub-problem)* The relaxed scenario group sub-problem involving each $\omega \in S$, (P_R^S), is formulated as follows:

$$Z_R^S := \min_{(x,\eta,v) \in \mathcal{X}^S} \sum_{\omega \in S} \left(p^\omega \sum_{t \in \mathcal{T}} c_t^{\omega \top} x_t^\omega + \lambda p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right) \right), \tag{11}$$

where

$$\mathcal{X}^S := \{(x, \eta, v) \in \mathbb{R}_+^{(n_1 - q_1)|S|} \times \mathbb{Z}_+^{q_1|S|} \times \dots \times \mathbb{R}_+^{(n_T - q_T)|S|} \times \mathbb{Z}_+^{q_T|S|} \times \mathbb{R}^{T|S|} \times \mathbb{R}_+^{T|S|} : \text{Constraints (5b)–(5g) only for } \omega \in S\}.$$

Proposition 6 P_R^S is a relaxation of P^S ; that is $Z^S \geq Z_R^S \forall S \subseteq \Omega$.

Remark 2 The $P^{\{\omega\}}$ is different than the conventional single scenario sub-problem ($P_R^{\{\omega\}}$), which has been widely studied in the literature [see, e.g., Madansky (1960), Ahmed (2013), and Sandıkçı and Özaltın (2017)] because in the $P_R^{\{\omega\}}$ the objective function and each constraint of the problem (5a) are defined only for a specific scenario with scenario index ω and non-anticipativity constraints (5e) are relaxed, while all variables and constraints are kept in the $P^{\{\omega\}}$. Clearly, the $P_R^{\{\omega\}}$ is a relaxation of $P^{\{\omega\}}$ as stated in Proposition 6.

Using the solution of the single-scenario sub-problems, we provide lower and upper bounds, which could then be used to strengthen the linear programming relaxation of the original problem, as demonstrated under the computational results under Sect. 4.5.

2.3.1 Lower bound

Definition 2.7 $\Gamma : \mathbb{R}^m \rightarrow \mathbb{R}$ is superadditive on $D \subseteq \mathbb{R}^m$ if $\Gamma(a + b) \geq \Gamma(a) + \Gamma(b)$ for all $a, b \in D$ such that $a + b \in D$.

Proposition 7 $Z^{\{\omega\}}$ is superadditive on $S \subseteq \Omega$. That is for $\omega, \omega' \in S$,

$$Z^{\{\omega, \omega'\}} \geq Z^{\{\omega\}} + Z^{\{\omega'\}}. \tag{12}$$

Proof Let $(\hat{x}, \hat{\eta}, \hat{v})$, $(\bar{x}, \bar{\eta}, \bar{v})$, and $(\check{x}, \check{\eta}, \check{v})$ be the optimal solutions of the scenario- ξ^ω , scenario- $\xi^{\omega'}$, and scenario- $\{\xi^\omega, \xi^{\omega'}\}$ sub-problems, respectively. Then $Z^{\{\omega, \omega'\}} = p^\omega \sum_{t \in T} c_t^{\omega \top} \check{x}_t^\omega + \lambda p^\omega \sum_{t \in T \setminus \{1\}} \left(\check{\eta}_t^\omega + \frac{1}{1-\alpha_t} \check{v}_t^\omega \right) + p^{\omega'} \sum_{t \in T} c_t^{\omega' \top} \check{x}_t^{\omega'} + \lambda p^{\omega'} \sum_{t \in T \setminus \{1\}} \left(\check{\eta}_t^{\omega'} + \frac{1}{1-\alpha_t} \check{v}_t^{\omega'} \right) \geq p^\omega \sum_{t \in T} c_t^{\omega \top} \hat{x}_t^\omega + \lambda p^\omega \sum_{t \in T \setminus \{1\}} \left(\hat{\eta}_t^\omega + \frac{1}{1-\alpha_t} \hat{v}_t^\omega \right) + p^{\omega'} \sum_{t \in T} c_t^{\omega' \top} \bar{x}_t^{\omega'} + \lambda p^{\omega'} \sum_{t \in T \setminus \{1\}} \left(\bar{\eta}_t^{\omega'} + \frac{1}{1-\alpha_t} \bar{v}_t^{\omega'} \right) = Z^{\{\omega\}} + Z^{\{\omega'\}}$ because $p^\omega \sum_{t \in T} c_t^{\omega \top} \check{x}_t^\omega + \lambda p^\omega \sum_{t \in T \setminus \{1\}} \left(\check{\eta}_t^\omega + \frac{1}{1-\alpha_t} \check{v}_t^\omega \right) \geq p^\omega \sum_{t \in T} c_t^{\omega \top} \hat{x}_t^\omega + \lambda p^\omega \sum_{t \in T \setminus \{1\}} \left(\hat{\eta}_t^\omega + \frac{1}{1-\alpha_t} \hat{v}_t^\omega \right)$ and $p^{\omega'} \sum_{t \in T} c_t^{\omega' \top} \check{x}_t^{\omega'} + \lambda p^{\omega'} \sum_{t \in T \setminus \{1\}} \left(\check{\eta}_t^{\omega'} + \frac{1}{1-\alpha_t} \check{v}_t^{\omega'} \right) \geq p^{\omega'} \sum_{t \in T} c_t^{\omega' \top} \bar{x}_t^{\omega'} + \lambda p^{\omega'} \sum_{t \in T \setminus \{1\}} \left(\bar{\eta}_t^{\omega'} + \frac{1}{1-\alpha_t} \bar{v}_t^{\omega'} \right)$ by the optimality of solutions $(\hat{x}, \hat{\eta}, \hat{v})$ and $(\bar{x}, \bar{\eta}, \bar{v})$ to the scenario- ξ^ω and scenario- $\xi^{\omega'}$ problems, respectively. \square

Proposition 8 (Lower bound) Suppose that the objective function for each scenario is non-negative. Then Z^S and $\sum_{\omega \in S} Z^{\{\omega\}}$ provides a lower bound on the optimal objective function value of the original problem P (5a), Z , as in the following inequality:

$$Z \geq Z^S \geq \sum_{\omega \in S} Z^{\{\omega\}}. \tag{13}$$

Proof The first inequality follows by Proposition 5, while the second inequality is the generalization of the superadditivity result in Proposition 7. \square

Remark 3 The lower bound (13) is tightest when $S = \Omega$. In this case, the non-negativity assumption for each scenario’s objective function value is not required.

2.3.2 Upper bound

Proposition 9 (Upper bound) Let \hat{x}^ω be optimal solutions to the scenario- ξ^ω problem $P^{\{\omega\}}$; and let \bar{Z}^ω be the objective function value of P at solution \hat{x}^ω . Then,

$$Z \leq \min_{\omega \in \Omega} \bar{Z}^\omega. \tag{14}$$

2.4 Scenario dominance cuts

In this section, we derive new cutting planes based on the “scenario dominance” concept defined in Sect. 2.1. We first define Lemmas 1 and 2 to motivate the main result on the scenario dominance cutting planes (Theorem 1). Then we present a corollary (Corollary 1.1) based on Theorem 1 to generalize the scenario dominance cuts to multiple time stages and further improve their computational benefits.

In this section, we present results regarding the general uncertainty, that is, the uncertainty in the right-hand side b_t^ω , left-hand side (A_t^ω, H_t^ω) , and the objective coefficients c_t^ω . For notational brevity, we denote $P^{\{\omega\}}$ as P^ω and $Z^{\{\omega\}}$ as Z^ω for the rest of the paper. We give the following two definitions that are commonly used in the results presented in the rest of the paper.

Definition 2.8 Let $(\hat{x}^\omega, \hat{\eta}^\omega, \hat{v}^\omega)$ be a feasible solution for the scenario- ξ^ω problem (9) and Z^ω be the corresponding objective function value, i.e., $Z^\omega := p^\omega \sum_{t \in \mathcal{T}} c_t^{\omega \top} \hat{x}_t^\omega + \lambda p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\hat{\eta}_t^\omega + \frac{1}{1 - \alpha_t} \hat{v}_t^\omega \right)$. Note that Z^ω is a scalar value.

Lemma 1 below states that the portion of the optimal objective function value for the original problem (5a) that corresponds to the scenario- ξ^ω is bounded below by the optimal objective value of the scenario- ξ^ω problem.

Lemma 1 Let (x^*, η^*, v^*) be the optimal solution to the original problem (5a) and $(\hat{x}^\omega, \hat{\eta}^\omega, \hat{v}^\omega)$ be the optimal solution of the scenario- ξ^ω problem (9). Then, the portion of the optimal solution for the original problem (5a) that corresponds to scenario- ξ^ω , $(x^{\omega*}, \eta^{\omega*}, v^{\omega*})$, satisfies the following inequality:

$$p^\omega \sum_{t \in \mathcal{T}} c_t^{\omega \top} x_t^{\omega*} + \lambda p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\eta_t^{\omega*} + \frac{1}{1 - \alpha_t} v_t^{\omega*} \right) \geq Z^\omega \quad \forall \omega \in \Omega.$$

Proof This is a proof by contradiction. For some $\omega \in \Omega$ and any $(x^{\omega*}, \eta^{\omega*}, v^{\omega*})$ assume that

$$\begin{aligned} & p^\omega \sum_{t \in \mathcal{T}} c_t^{\omega \top} x_t^{\omega*} + \lambda p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\eta_t^{\omega*} + \frac{1}{1 - \alpha_t} v_t^{\omega*} \right) \\ & < p^\omega \sum_{t \in \mathcal{T}} c_t^{\omega \top} \hat{x}_t^\omega + \lambda p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\hat{\eta}_t^\omega + \frac{1}{1 - \alpha_t} \hat{v}_t^\omega \right). \end{aligned} \tag{15}$$

Then define a new solution \hat{x} to P as follows:

$$\hat{x}_t = x_t^{\omega*}, \hat{\eta}_t = \eta_t^{\omega*}, \hat{v}_t = v_t^{\omega*} \quad \forall t \in \mathcal{T}. \tag{16}$$

By (16), the new solution $(\hat{x}_t, \hat{\eta}_t, \hat{v}_t)$ is feasible and reduces the value of Z^ω by (16), contradicting to the optimality of $(\hat{x}^\omega, \hat{\eta}^\omega, \hat{v}^\omega)$ for the scenario- ξ^ω problem. \square

Lemma 2 below defines the relations of the optimal objective function value of the sub-problems that correspond to two scenarios that have a dominance relation.

Lemma 2 Let ξ^ω be a scenario such that $\omega \in \Theta_{T, \xi^{\omega'}}^- \subseteq \Omega$ for a given scenario $\xi^{\omega'}$. Let $(\hat{x}^\omega, \hat{\eta}^\omega, \hat{v}^\omega)$ and $(\hat{x}^{\omega'}, \hat{\eta}^{\omega'}, \hat{v}^{\omega'})$ be the optimal solution of the scenario- ξ^ω and the scenario- $\xi^{\omega'}$ problems, Z^ω and $Z^{\omega'}$ be the corresponding objective values, respectively. Then the following inequality holds:

$$Z^\omega \geq Z^{\omega'} \quad \forall \omega \in \Theta_{T, \xi^{\omega'}}^-.$$

Proof Add the following equalities to both scenario- ξ^ω and scenario- $\xi^{\omega'}$ problems, P^ω and $P^{\omega'}$:

$$x_t^\omega = x_t^{\omega'}, \eta_t^\omega = \eta_t^{\omega'}, v_t^\omega = v_t^{\omega'}, \quad \forall t \in \mathcal{T}. \tag{17}$$

Define the new problems as \tilde{P}^ω and $\tilde{P}^{\omega'}$, respectively. Define the following sets of scenarios:

$$\Omega^1 = \left\{ \hat{\omega} \in \Omega \setminus \{ \omega, \omega' \} : \exists t \in \mathcal{T} \text{ such that } \xi_{[t]}^{\hat{\omega}} = \xi_{[t]}^{\omega'} \text{ for some } t > 1 \right\}.$$

$$\Omega^2 = \left\{ \check{\omega} \in \Omega \setminus \{ \Omega^1 \cup \{ \omega \} \cup \{ \omega' \} \} \right\}.$$

Note that the set Ω^1 only includes scenarios $\hat{\omega} \neq \{ \omega, \omega' \}$ that share a common history with scenario ω' up to stage $t > 1$ and $\Omega = \Omega^1 \cup \Omega^2 \cup \{ \omega \} \cup \{ \omega' \}$.

Define a feasible solution $(\tilde{x}, \tilde{\eta}, \tilde{v})$ for \tilde{P}^ω and $\tilde{P}^{\omega'}$ as follows:

$$\tilde{x}_t^\omega = \ddot{x}_t^\omega, \tilde{\eta}_t^\omega = \ddot{\eta}_t^\omega, \tilde{v}_t^\omega = \ddot{v}_t^\omega \quad \forall t \in \mathcal{T}; \tag{18}$$

$$\tilde{x}_t^{\omega'} = \ddot{x}_t^{\omega'}, \tilde{\eta}_t^{\omega'} = \ddot{\eta}_t^{\omega'}, \tilde{v}_t^{\omega'} = \ddot{v}_t^{\omega'} \quad \forall t \in \mathcal{T}; \tag{19}$$

$$\tilde{x}_t^{\check{\omega}} = \ddot{x}_t^{\check{\omega}}, \tilde{\eta}_t^{\check{\omega}} = \ddot{\eta}_t^{\check{\omega}}, \tilde{v}_t^{\check{\omega}} = \ddot{v}_t^{\check{\omega}} \quad \forall \check{\omega} \in \Omega^2, \quad \forall t \in \mathcal{T}; \tag{20}$$

$$\tilde{x}_t^{\hat{\omega}} = \ddot{x}_t^{\hat{\omega}}, \tilde{\eta}_t^{\hat{\omega}} = \ddot{\eta}_t^{\hat{\omega}}, \tilde{v}_t^{\hat{\omega}} = \ddot{v}_t^{\hat{\omega}} \quad \forall \hat{\omega} \in \Omega^1 \text{ s.t. } \xi_{[t]}^{\hat{\omega}} = \xi_{[t]}^{\omega'} \text{ for } t = 1, \dots, t', \quad \forall t, t' \in \mathcal{T}, \tag{21}$$

where scenario $\check{\omega} \in \Omega$ satisfies the following conditions for a given $\hat{\omega} \in \Omega^1$ such that $\xi_{[t]}^{\hat{\omega}} = \xi_{[t]}^{\omega'}$ for $t = 1, \dots, t'$: $\xi_{[t]}^{\check{\omega}} = \xi_{[t]}^{\omega'}$ for $t = 1, \dots, t'$ and $\xi_{[t]}^{\check{\omega}} = \xi_{[t]}^{\hat{\omega}}$ for $t = t' + 1, \dots, T$.

The solution defined in (18)–(21) is feasible for both $\tilde{P}^{\omega'}$ and \tilde{P}^ω because it satisfies constraints (5b)–(5d), as well as the non-anticipativity constraints (5e). Define the optimal objective value of \tilde{P}^ω as \tilde{Z}^ω , and the new optimal value for $\tilde{P}^{\omega'}$ as $\tilde{Z}^{\omega'}$. The solution $(\tilde{x}, \tilde{\eta}, \tilde{v})$ defined in (18)–(21) is optimal for \tilde{P}^ω because, for each feasible $(x^\omega, \eta^\omega, v^\omega)$, we have $\tilde{Z}^\omega \geq Z^\omega$ and $\tilde{Z}^\omega = Z^\omega$. Because we have $\tilde{x}^\omega = \tilde{x}^{\omega'}$, $\tilde{\eta}^\omega = \tilde{\eta}^{\omega'}$, and $\tilde{v}^\omega = \tilde{v}^{\omega'}$ and $c_t^\omega \geq c_t^{\omega'}$,

$$\tilde{Z}^\omega \geq \tilde{Z}^{\omega'}. \tag{22}$$

Since $b_t^\omega \geq b_t^{\omega'}$, $A_t^\omega \leq A_t^{\omega'}$ and $H_t^\omega \leq H_t^{\omega'}$, and $c_t^\omega \geq c_t^{\omega'}$, a feasible solution $(\tilde{x}^{\omega'}, \tilde{\eta}^{\omega'}, \tilde{v}^{\omega'})$ has to satisfy constraints related to the scenario- ξ^ω by the addition of constraints (17), and thus we have

$$\tilde{Z}^{\omega'} \geq Z^{\omega'}. \tag{23}$$

Because $\tilde{x}^\omega = \tilde{x}^{\omega'}$, $\tilde{\eta}^\omega = \tilde{\eta}^{\omega'}$, and $\tilde{v}^\omega = \tilde{v}^{\omega'}$, we have

$$\tilde{Z}^\omega = Z^\omega. \tag{24}$$

Furthermore, we have

$$Z^\omega = \tilde{Z}^\omega \geq \tilde{Z}^{\omega'} \geq Z^{\omega'}. \tag{25}$$

The first and second equalities in (25) hold by (24) and (22), respectively, while the last inequality holds by (23). □

Using those two lemmas above, we show the main theorem below on the scenario dominance cut that provides a lower bound on the portion of the optimal objective value for the original problem (5a) that corresponds to the scenario- ξ^ω problem. This lower bound is obtained by solving the single-scenario problems that are dominated by the scenario ξ^ω .

Theorem 1 *Scenario dominance cuts* Assume that uncertainty in problem (5a) is observed in the objective function c_t^ω , in the right-hand side coefficient b_t^ω , and the left-hand side coefficient (A_t^ω, H_t^ω) . Let ξ^ω and $\xi^{\omega'}$ be two scenarios such that $\xi^{\omega'} \preceq \xi^\omega$, $\omega, \omega' \in \Omega$ and $\omega \neq \omega'$, i.e., $p^\omega \geq p^{\omega'}$, $b_t^\omega \geq b_t^{\omega'}$, $A_t^\omega \leq A_t^{\omega'}$ and $H_t^\omega \leq H_t^{\omega'}$, and $c_t^\omega \geq c_t^{\omega'}$ for all $t \in \mathcal{T}$. Then, the portion of the optimal solution for the original problem (5a) that corresponds to scenario- ξ^ω , $(x^{\omega*}, \eta^{\omega*}, v^{\omega*})$, satisfies the following inequality:

$$p^\omega \sum_{t \in \mathcal{T}} c_t^{\omega \top} x_t^\omega + \lambda p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right) \geq Z^{\omega'} \quad \forall \omega \in \Theta_{T, \xi^{\omega'}}^- \tag{26}$$

Proof The first inequality below follows from Lemma 1, while the second follows from Lemma 2.

$$p^\omega \sum_{t \in \mathcal{T}} c_t^{\omega \top} x_t^{\omega*} + \lambda p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\eta_t^{\omega*} + \frac{1}{1 - \alpha_t} v_t^{\omega*} \right) \geq Z^\omega \geq Z^{\omega'}$$

□

In the corollary following Theorem 1, we generalize the cuts to multiple stages by solving each stage- t problem for $t \in \mathcal{T}$.

Definition 2.9 Let $(\dot{x}_t^\omega, \dot{\eta}_t^\omega, \dot{v}_t^\omega)$ be the optimal solution for the stage- t scenario- ξ^ω problem (10) and Z_t^ω be the corresponding optimal objective value such that

$$Z_t^\omega = p^\omega \sum_{j=1}^t c_j^{\omega \top} \dot{x}_j^\omega + \lambda p^\omega \sum_{j=2}^t \left(\dot{\eta}_j^\omega + \frac{1}{1 - \alpha_j} \dot{v}_j^\omega \right)$$

Corollary 1.1 *Stage- t Scenario dominance cuts* Let ξ^ω and $\xi^{\omega'}$ be two scenarios $l, \omega \in \Omega$ such that $\xi_t^{\omega'} \preceq \xi_t^\omega$ for $t \in \hat{\mathcal{T}} = \{1, \dots, \hat{t}\} \subseteq \mathcal{T}$, where $\hat{t} \in \mathcal{T}$. Then, the portion of the optimal solution for the original problem (5a) that corresponds to scenario- ξ^ω up to stage t , $(x_t^{\omega*}, \eta_t^{\omega*}, v_t^{\omega*})$, satisfies the following inequality:

$$p^\omega \sum_{j=1}^t c_j^{\omega \top} x_j^\omega + \lambda p^\omega \sum_{j=2}^t \left(\eta_j^\omega + \frac{1}{1 - \alpha_j} v_j^\omega \right) \geq Z_t^{\omega'} \quad \forall \omega \in \Theta_{t, \xi^{\omega'}}^-, t \in \hat{\mathcal{T}} \subseteq \mathcal{T} \tag{27}$$

Proof The proof follows from the application of Theorem 1 to each stage $t \in \mathcal{T}$. □

Remark 4 The stage- t scenario dominance cuts (27) are generalizations of the scenario dominance cuts (26). Both (26) and (27) are satisfied by all optimal solutions to the original problem (5a); however they are not valid inequalities for the feasible region of the problem (5a), i.e., they both may cut-off a non-optimal feasible solution, while preserving the optimal solution. This type of inequalities are not uncommon (see, e.g., Bender’s optimality cuts Benders (2005), supervalid inequalities of Israeli and Wood (2002), and the pre-processing cuts presented in Kibis et al. (2020)), and they are useful because they help to shrink the feasible region while preserving the optimal solution.

Remark 5 For problem instances, P , where it is easy to find a feasible solution but hard to find the optimal one, we expect that the single scenario sub-problem, P^ω , could be solved much faster than P because the P^ω is a relaxation of P , as shown in Proposition 5. Furthermore, in the P^ω , we optimize the value of the variables related to a single scenario ω , while we only need to assign a feasible solution to all other variables in order to satisfy the constraint set,

\mathcal{X} . This observation is also justified by the computational results for the dynamic knapsack problems in Sect. 4. On the other hand, if finding a feasible solution is hard, the single scenario sub-problems could take a long time to solve. In such cases, we suggest solving the relaxed single scenario sub-problem P_R^ω rather than the P^ω . Because the P_R^ω is merely a relaxation of the P^ω , including only the constraints and variables of a single scenario, the lower bound (13), scenario dominance cuts (26), and stage- t scenario dominance cuts (27) would still hold when P_R^ω is used to generate those inequalities.

3 Risk-averse multi-stage stochastic knapsack

We consider a class of stochastic mean-CVaR multi-stage multi-dimensional mixed 0-1 knapsack problems (SMKPs) as a benchmark set for our computations. The classical deterministic, single-dimensional, and risk-neutral version of the problem is known as a maximization binary knapsack problem (KP): given a set of items with profits and sizes, and the capacity value of a knapsack, the objective is to maximize the total profit of selected items in the knapsack while satisfying the capacity constraint (Kellerer et al. 2004). Another widely used variant is the minimization knapsack problem (Güntzer and Jungnickel 2000; Han and Makino 2009; Csirik 1991), that is, given a set of items associated with costs and sizes, our goal is to minimize the total cost of selected items while ensuring that the knapsack is covered, i.e., the sum of the items' sizes is larger than or equal to a constant. Güntzer and Jungnickel (2000) consider a minimization equivalent to the maximization KP, where binary selection variables are negated in the formulation. A generalization of the KP is the multi-dimensional knapsack (MKP), which involves multiple constraints, each corresponding to a knapsack, whose capacity should be satisfied. The structure of the MKP is quite similar to the general integer programs with multiple constraints.

When the costs of items are not fixed and are instead given by a probability distribution, the problem is broadly referred to as the stochastic knapsack problem (Morton and Wood 1998; Dean et al. 2008; Angulo et al. 2016). Angulo et al. (2016) and Guan et al. (2009) present cutting plane approaches to improve the solvability of the stochastic dynamic knapsack problem. In the computational experiments, we tackle solving the general stochastic multi-stage multi-dimensional knapsack problem (SMKP) with integer and binary variables and a mean-risk objective.

3.1 Mean-CVaR multi-stage stochastic knapsack formulation

We adapt the scenario-dominance cuts presented in Sect. 2.4 to solve a class of mean-CVaR stochastic multi-dimensional mixed 0-1 knapsack problems (CVaR-SMKPs) and demonstrate their computational benefit on this class of multi-stage stochastic MIPs. CVaR-SMKP represents the general mean-CVaR multi-stage stochastic mixed-integer programs with binary and continuous variables in the first and subsequent stages.

To formulate the CVaR-SMKP, we define the following notation:

Indices:

t : index for time period, where $t \in \mathcal{T} = \{1, \dots, T\}$, and T is the final time period.

i : index for item, where $i \in \mathcal{I} = \{1, \dots, I\}$, and I is the maximum number of items.

ω : index for scenario, where $\omega \in \Omega = \{1, \dots, \bar{\Omega}\}$, and $\bar{\Omega}$ is the maximum number of scenarios.

Variables:

- $x_{it}^\omega \in \mathbb{B}$: 0 – 1 binary variable to select an item i in period t under scenario $\omega \in \Omega$, where \mathbb{B} is the set of binary integers.
- $z_t^\omega \in \mathbb{R}_+$: continuous variable in period t under scenario $\omega \in \Omega$.
- $y_t^\omega \in \mathbb{B}$: 0 – 1 binary variable that relates to the variable x_{it}^ω in period t under scenario $\omega \in \Omega$.
- $\eta_t^\omega \in \mathbb{R}$: VaR variable in period t under scenario $\omega \in \Omega$.
- $v_t^\omega \in \mathbb{R}_+$: continuous variable that represents the excess cost over the VaR variable in period t under scenario $\omega \in \Omega$.

All parameters defined below are real numbers, e.g., belong to the field of \mathbb{R} :
 Stochastic parameters:

- q_t^ω : cost coefficient associated with variable y_t^ω in period t under scenario $\omega \in \Omega$.
- r_t^ω : coefficient corresponding to the continuous variable z_t^ω in period t under scenario $\omega \in \Omega$.
- b_t^ω : minimum total value of all selected items in period t under scenario $\omega \in \Omega$.

Deterministic parameters:

- c_{it} : cost coefficient associated with variable x_{it}^ω for item i in period t .
- d_t : cost coefficient associated with variable z_t^ω in period t .
- a_{it} : value of item i in period t .
- κ_{it} : size of item i in period t .
- w_t : coefficient corresponding to the continuous variable y_t^ω in period t .
- h_t : minimum total size of all selected items in period t .

All objective coefficients are assumed to be non-negative. Using the above notation, the CVaR-SMKP (28a) can be formulated as follows.

$$\min_{\substack{x_{i1}^\omega, \dots, x_{iT}^\omega, \\ z_1^\omega, \dots, z_T^\omega, \\ y_1^\omega, \dots, y_T^\omega, \\ \eta_1^\omega, \dots, \eta_T^\omega, \\ v_1^\omega, \dots, v_T^\omega}} \sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T}} \left(\sum_{i \in \mathcal{I}} c_{it} x_{it}^\omega + d_t z_t^\omega + q_t^\omega y_t^\omega \right) + \lambda \sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right) \tag{28a}$$

$$\text{s.t. } \sum_{j=1}^t \sum_{i \in \mathcal{I}} a_{ij} x_{ij}^\omega + r_t^\omega z_t^\omega \geq b_t^\omega \quad \forall t \in \mathcal{T}, \omega \in \Omega, \tag{28b}$$

$$\sum_{i \in \mathcal{I}} \kappa_{it} x_{it}^\omega + w_t y_t^\omega \geq h_t \quad \forall t \in \mathcal{T}, \omega \in \Omega, \tag{28c}$$

$$v_t^\omega \geq \sum_{i \in \mathcal{I}} c_{it} x_{it}^\omega + d_t z_t^\omega + q_t^\omega y_t^\omega - \eta_t^\omega \quad \forall t \in \mathcal{T}, \omega \in \Omega, \tag{28d}$$

$$x_{it}^\omega = x_{it}^{\omega'} \quad \forall t \in \mathcal{T}, i \in \mathcal{I}, \omega, \omega' \in \Omega \text{ s.t. } \xi_{[t]}^\omega = \xi_{[t]}^{\omega'}, \tag{28e}$$

$$z_t^\omega = z_t^{\omega'}, y_t^\omega = y_t^{\omega'} \quad \forall t \in \mathcal{T}, \omega, \omega' \in \Omega \text{ s.t. } \xi_{[t]}^\omega = \xi_{[t]}^{\omega'}, \tag{28f}$$

$$\eta_t^\omega = \eta_t^{\omega'}, v_t^\omega = v_t^{\omega'} \quad \forall t \in \mathcal{T}, \omega, \omega' \in \Omega \text{ s.t. } \xi_{[t]}^\omega = \xi_{[t]}^{\omega'}, \tag{28g}$$

$$x_{it}^\omega, y_t^\omega \in \{0, 1\} \quad \forall t \in \mathcal{T}, i \in \mathcal{I}, \omega \in \Omega, \tag{28h}$$

$$z_t^\omega, v_t^\omega \geq 0 \quad \forall t \in \mathcal{T}, \omega \in \Omega. \tag{28i}$$

The mean-risk function (28a) minimizes the expected sum of knapsack costs and the weighted VaR and the weighted positive expected excess cost over the VaR over all items $i \in \mathcal{I}$, periods $t \in \mathcal{T}$, and scenarios $\omega \in \Omega$. Constraints (28b) and (28c) define knapsack-related constraints, where constraint (28b) [(28c)] ensures that the sum of the value [size] of selected items should be larger than a given constant. Constraints (28d) define VaR and the excess cost under each scenario over the VaR variable η_t^ω . Constraints (28e)–(28g) represent the nonanticipativity constraints for the x , z , y , η , and v variables, respectively. Finally, (28h) represent binary integer restrictions on x and y variables, while constraints (28i) represent lower bounds on the z , η , and v variables, respectively. Note that CVaR-SMKP formulated above includes uncertainty in the objective function parameter q_t^ω , the left-hand side parameter r_t^ω , and the right-hand side parameter b_t^ω .

3.2 Bounds

Proposition 10 *Let $S \subseteq \Omega$ be a subset of scenarios. For a given $\omega' \in S$, let $(\dot{x}^{\omega'}, \dot{z}^{\omega'}, \dot{y}^{\omega'}, \dot{\eta}^{\omega'}, \dot{v}^{\omega'})$ and $Z^{\omega'}$ be the optimal solution and the corresponding objective function value of the scenario- $\xi^{\omega'}$ problem, respectively. Then an optimal solution to the CVaR-SMKP (28a), $(x^*, z^*, y^*, \eta^*, v^*)$, satisfies the following inequality:*

$$\sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T}} \left(\sum_{i \in \mathcal{I}} c_{it} x_{it}^\omega + d_t z_t^\omega + q_t^\omega y_t^\omega \right) + \lambda \sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right) \geq \sum_{\omega' \in S} Z^{\omega'}. \tag{29}$$

Proof In the CVaR-SMKP (28a), we assume that all objective coefficients are non-negative, implying that the specific scenario objective function values are also non-negative. Then, the result follows immediately by applying inequality (13) in Proposition 8 for the CVaR-SMKP. \square

Proposition 11 *Let $S \subseteq \Omega$ be a subset of scenarios. For a given $\omega' \in S$, let $\bar{Z}^{\omega'}$ be the objective function value of (28a) at the optimal solution of the scenario- $\xi^{\omega'}$ problem, $(\dot{x}^{\omega'}, \dot{z}^{\omega'}, \dot{y}^{\omega'}, \dot{\eta}^{\omega'}, \dot{v}^{\omega'})$. Then an optimal solution to the CVaR-SMKP (28a), $(x^*, z^*, y^*, \eta^*, v^*)$, satisfies the following inequality:*

$$\sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T}} \left(\sum_{i \in \mathcal{I}} c_{it} x_{it}^\omega + d_t z_t^\omega + q_t^\omega y_t^\omega \right) + \lambda \sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right) \leq \min_{\omega' \in S} \bar{Z}^{\omega'}. \tag{30}$$

Proof The result follows immediately by applying inequality (14) in Proposition 9 for the CVaR-SMKP. \square

Example 3 Consider a CVaR-SMKP instance in which $T = 4$ and $I = 2$. This instance has $\bar{\Omega} = 8$ scenarios, and the probability of each scenario is set as $1/8$. The risk parameter λ is set as 1 and $\alpha_t = 0.95$ for each $t \in \mathcal{T}$. We assume that all parameter values are set at 0 at period $t = 1$. The data pertaining the instance is: $c_{1t} = (59, 35, 67)$, $c_{2t} = (23, 16, 14)$, $d_t = (47, 3, 67)$, $a_{1t} = (75, 95, 4)$, $a_{2t} = (82, 56, 26)$, $\kappa_{1t} = (98, 9, 20)$, $\kappa_{2t} = (41, 1, 60)$, $w_t = (84, 97, 57)$, $h_t = (167.2, 80.2, 102.7)$ for $t = \{2, 3, 4\}$. The parameter values at $t = 1$ are assumed to be known and are set as zero in this example. The values of the uncertain parameters $(q_t^\omega, r_t^\omega, b_t^\omega)$ are presented in Table 1.

Table 1 Values of the uncertain parameters $(q_t^\omega, r_t^\omega, b_t^\omega)$ for $\omega = \{1, \dots, 8\}$ and $t = \{2, 3, 4\}$

ω	q_t^ω	r_t^ω	b_t^ω
1	[33, 49, 31]	[37, 47, 29]	[99.8, 198.0, 236.3]
2	[33, 49, 74]	[37, 47, 61]	[99.8, 198.0, 260.3]
3	[33, 62, 31]	[37, 85, 29]	[99.8, 226.5, 236.3]
4	[33, 62, 74]	[37, 85, 61]	[99.8, 226.5, 260.3]
5	[85, 49, 31]	[97, 47, 29]	[144.8, 198.0, 236.3]
6	[85, 49, 74]	[97, 47, 61]	[144.8, 198.0, 260.3]
7	[85, 62, 31]	[97, 85, 29]	[144.8, 226.5, 236.3]
8	[85, 62, 74]	[97, 85, 61]	[144.8, 226.5, 260.3]

Now, let $S = \{1, 2, 3, 4\}$. Then $Z^1 = 69.9$, $Z^2 = 80.6$, $Z^3 = 72.39$, and $Z^4 = 83.14$. For S the lower-bound inequality (29) is:

$$\sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T}} \left(\sum_{i \in \mathcal{I}} c_{it} x_{it}^\omega + d_t z_t^\omega + q_t^\omega y_t^\omega \right) + \lambda \sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right) \geq 306.04. \tag{31}$$

The minimum upper-bound is attained when solving the scenario problem with $\omega = 4$, and thus the upper-bound inequality (30) is:

$$\sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T}} \left(\sum_{i \in \mathcal{I}} c_{it} x_{it}^\omega + d_t z_t^\omega + q_t^\omega y_t^\omega \right) + \lambda \sum_{\omega \in \Omega} p^\omega \sum_{t \in \mathcal{T} \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right) \leq 694.86. \tag{32}$$

The optimal objective value for this instance is 614.08. For $S = \{1, 2, 3, 4, 5, 6, 7, 8\}$, the lower bound of the inequality (31) can be improved to 614.08, while the upper bound in the inequality (32) stays the same.

3.3 Stage- t scenario-dominance cuts

Theorem 2 Let $\xi^{\omega'}$ be a scenario with $\omega' \in S \subseteq \Omega$. For each $t \in \mathcal{T}$, let $(\dot{x}_t^{\omega'}, \dot{z}_t^{\omega'}, \dot{y}_t^{\omega'}, \dot{\eta}_t^{\omega'}, \dot{v}_t^{\omega'})$ represent a vector of optimal solutions to the stage- t scenario- $\xi^{\omega'}$ problem (10) and define the corresponding optimal objective value as:

$$Z_t^{\omega'} := p^{\omega'} \sum_{j=1}^t \left(\sum_{i \in \mathcal{I}} c_{ij} \dot{x}_{ij}^{\omega'} + d_j \dot{z}_j^{\omega'} + q_j^{\omega'} \dot{y}_j^{\omega'} \right) + \lambda p^{\omega'} \sum_{j=2}^t \left(\dot{\eta}_j^{\omega'} + \frac{1}{1 - \alpha_j} \dot{v}_j^{\omega'} \right).$$

Then, the portion of the optimal solution up to stage t to the CVaR-SMKP (28a) corresponding to scenario ξ^ω such that $\xi_t^{\omega'} \leq \xi_t^\omega$ for $t \in \hat{\mathcal{T}} \subseteq \mathcal{T}$, $(x_t^{\omega*}, z_t^{\omega*}, y_t^{\omega*}, \eta_t^{\omega*}, v_t^{\omega*})$, is satisfied by the following inequality:

$$\begin{aligned} & p^\omega \sum_{j=1}^t \left(\sum_{i \in \mathcal{I}} c_{ij} x_{ij}^\omega + d_j z_j^\omega + q_j^\omega y_j^\omega \right) + \lambda p^\omega \sum_{j=2}^t \left(\eta_j^\omega + \frac{1}{1 - \alpha_j} v_j^\omega \right) \\ & \geq Z_t^{\omega'} \quad \forall \omega \in \Theta_{t, \xi^{\omega'}}^-, \quad t \in \hat{\mathcal{T}}. \end{aligned} \tag{33}$$

Proof The result follows immediately by applying the inequality (27) in Corollary 1.1 for the mean-CVaR stochastic multi-stage knapsack problem (28a). \square

Corollary 2.1 *The inequality (33) can be strengthened as in the following inequality:*

$$\begin{aligned}
 & p^\omega \sum_{j=1}^t \left(\sum_{i \in \mathcal{I}} c_{ij} x_{ij}^\omega + d_j z_j^\omega + q_j^\omega y_j^\omega \right) + \lambda p^\omega \sum_{j=2}^t \left(\eta_j^\omega + \frac{1}{1 - \alpha_j} v_j^\omega \right) \\
 & \geq Z_t^{\omega'} \quad \forall \omega \in \Theta_{t, \xi^{\omega'}}^-, t \in \hat{T}.
 \end{aligned} \tag{34}$$

Proof Because the solution $(x^{\omega^*}, z^{\omega^*}, y^{\omega^*}, \eta^{\omega^*}, v^{\omega^*})$ is a part of a feasible solution to the full problem defined in (28a), $\forall \omega, \omega' \in \Omega$ such that $\xi_t^{\omega'} \preceq \xi_t^\omega$ we have

$$\begin{aligned}
 \bar{c} & := p^{\omega'} \sum_{j=1}^t \left(\sum_{i \in \mathcal{I}} c_{ij} x_{ij}^{\omega^*} + d_j z_j^{\omega^*} + q_j^{\omega'} y_j^{\omega^*} \right) + \lambda p^{\omega'} \sum_{j=2}^t \left(\eta_j^{\omega^*} + \frac{1}{1 - \alpha_j} v_j^{\omega^*} \right) \geq \mathcal{B}. \\
 & := p^{\omega'} \sum_{j=1}^t \left(\sum_{i \in \mathcal{I}} c_{ij} \dot{x}_{ij}^{\omega'} + d_j \dot{z}_j^{\omega'} + q_j^{\omega'} \dot{y}_j^{\omega'} \right) + \lambda p^{\omega'} \sum_{j=2}^t \left(\dot{\eta}_j^{\omega'} + \frac{1}{1 - \alpha_j} \dot{v}_j^{\omega'} \right).
 \end{aligned} \tag{35}$$

But $q_t^\omega \geq q_t^{\omega'} \quad \forall t \in \hat{T}$ implies that

$$\begin{aligned}
 \mathcal{A} & := p^\omega \sum_{j=1}^t \left(\sum_{i \in \mathcal{I}} c_{ij} x_{ij}^{\omega^*} + d_j z_j^{\omega^*} + q_j^\omega y_j^{\omega^*} \right) + \lambda p^\omega \sum_{j=2}^t \left(\eta_j^{\omega^*} + \frac{1}{1 - \alpha_j} v_j^{\omega^*} \right) \\
 & \geq \mathcal{C} := p^\omega \sum_{j=1}^t \left(\sum_{i \in \mathcal{I}} c_{ij} x_{ij}^{\omega^*} + d_j z_j^{\omega^*} + q_j^{\omega'} y_j^{\omega^*} \right) + \lambda p^\omega \sum_{j=2}^t \left(\eta_j^{\omega^*} + \frac{1}{1 - \alpha_j} v_j^{\omega^*} \right).
 \end{aligned} \tag{36}$$

Because we have $p^\omega \geq p^{\omega'}$, we have $\mathcal{C} \geq \bar{c}$, and thus the following inequalities hold:

$$\mathcal{A} \geq \mathcal{C} \geq \bar{c} \geq \mathcal{B}. \tag{37}$$

Because the inequality (33) implies that $\mathcal{A} \geq \mathcal{B}$, by the results above in (37) the inequality (34), which implies $\mathcal{C} \geq \mathcal{B}$, strengthens the inequality (33). \square

Example 3 (cont). Again, consider the CVaR-SMKP instance defined in Example 3. The index set of scenarios that dominate ξ^3 is defined as $\Theta_{2, \xi^3}^- = \{1, 2, 3, 4\}$, $\Theta_{3, \xi^3}^- = \{3, 4\}$, $\Theta_{4, \xi^3}^- = \{3\}$. Solving the stage- t scenario- ξ^3 problem for $t = 2, 3, 4$, we have $Z_2^3 = 31.62$, $Z_3^3 = 59.89$, and $Z_4^3 = 72.39$.

For $\omega = 4 \in \Theta_{t, \xi^3}^-$, the inequalities corresponding to (34) for $t = 2, 3$ can be, respectively, written as:

$$\begin{aligned}
 & p^4 \left(\sum_{i \in \mathcal{I}} c_{i2} x_{i2}^4 + d_2 z_2^4 + q_2^3 y_2^4 \right) + \lambda p^4 \left(\eta_2^4 + \frac{1}{1 - \alpha_2} v_2^4 \right) \geq 31.62. \\
 & p^4 \sum_{t \in \{2, 3\}} \left(\sum_{i \in \mathcal{I}} c_{it} x_{it}^4 + d_t z_t^4 + q_t^3 y_t^4 \right) + \lambda p^4 \sum_{t \in \{2, 3\}} \left(\eta_t^4 + \frac{1}{1 - \alpha_t} v_t^4 \right) \geq 59.89.
 \end{aligned}$$

Because $\Theta_{4, \xi^3}^- = \{3\}$, in period $t = 4$, we have only one inequality (34) as given below:

$$p^3 \sum_{t \in \{2,3,4\}} \left(\sum_{i \in \mathcal{I}} c_{it} x_{it}^3 + d_t z_t^3 + q_t^3 y_t^3 \right) + \lambda p^3 \sum_{t \in \{2,3,4\}} \left(\eta_t^3 + \frac{1}{1 - \alpha_t} v_t^3 \right) \geq 72.39.$$

The optimal objective function value to CVaR-SMKP (28a) is 614.08, and its Linear Programming (LP) relaxation is 507.60. By adding inequalities (29), (30), and (34) for $S = \{1, 2, 3, 4\}$, the optimal relaxation solution is cut-off, and the LP relaxation objective of (28a) increases to 560.61, closing the optimality gap by 50%.

3.4 Strong scenario-dominance cuts

Proposition 12 *Let $\xi^{\omega'}$ and $\xi^{\omega''}$ be two scenarios that share a common history up to time period $1, \dots, \tau \in \mathcal{T}$. For each $t \in \mathcal{T}$, where $t \geq \tau$, let $(\dot{x}_t^{\omega'}, \dot{z}_t^{\omega'}, \dot{y}_t^{\omega'}, \dot{\eta}_t^{\omega'}, \dot{v}_t^{\omega'})$ represent a vector of optimal solutions to the stage- t scenario- $\xi^{\omega''}$ problem and define the objective value corresponding to the stage- t scenario- $\xi^{\omega'}$ problem at the solution $(\dot{x}_t^{\omega'}, \dot{z}_t^{\omega'}, \dot{y}_t^{\omega'}, \dot{\eta}_t^{\omega'}, \dot{v}_t^{\omega'})$ as:*

$$\hat{Z}_t^{\omega'} := p^{\omega'} \sum_{j=1}^t \left(\sum_{i \in \mathcal{I}} c_{ij} \dot{x}_{ij}^{\omega''} + d_j \dot{z}_j^{\omega''} + q_j^{\omega'} \dot{y}_j^{\omega''} \right) + \lambda p^{\omega'} \sum_{j=2}^t \left(\dot{\eta}_j^{\omega''} + \frac{1}{1 - \alpha_j} \dot{v}_j^{\omega''} \right).$$

Then, there exists a feasible solution to the CVaR-SMKP (28a) corresponding to scenario- ξ^{ω} problem such that $\xi_t^{\omega'} \leq \xi_t^{\omega}$ for $t \in \hat{\mathcal{T}} \subseteq \mathcal{T}$, $(\bar{x}_t^{\omega}, \bar{z}_t^{\omega}, \bar{y}_t^{\omega}, \bar{\eta}_t^{\omega}, \bar{v}_t^{\omega})$, that is satisfied by the following inequality:

$$p^{\omega} \sum_{j=1}^t \left(\sum_{i \in \mathcal{I}} c_{ij} x_{ij}^{\omega} + d_j z_j^{\omega} + q_j^{\omega} y_j^{\omega} \right) + \lambda p^{\omega} \sum_{j=2}^t \left(\eta_j^{\omega} + \frac{1}{1 - \alpha_t} v_j^{\omega} \right) \geq \hat{Z}_t^{\omega'} \quad \forall \omega \in \Theta_{t, \xi^{\omega'}}^-, \quad t \in \hat{\mathcal{T}}. \tag{38}$$

Remark 6 It is easy to see that there exists a feasible solution to the CVaR-SMKP (28a) that is satisfied by the inequalities (38) because those inequalities provide a lower bound on the x -, z -, y -, η -, and v -variables. Using the inequalities (38), we often get the optimal solutions, but sometimes they may cut-off the optimal solution. This is because $\hat{Z}_t^{\omega'}$ is not optimal but a close-to-optimal objective value to the stage- t scenario- $\xi^{\omega'}$ problem. Therefore, one needs to carefully implement the strong scenario dominance cuts. In Sect. 4, we investigate the trade-off between the quality of the best solution obtained by applying those inequalities versus the computational speed-up obtained with them. We, specifically, present the objective value gap defined by the percentage deviation between the best objective value found by CPLEX for the original problem and the best objective value obtained after applying the cuts (38) to analyze the impact of adding these aggressive inequalities that may cut-off the optimal solution. On the other hand, the stage- t scenario dominance inequalities (34) do not cut-off the optimal solution and reduce the optimality gap considerably, as shown in Section 4. However, the inequalities (38) may be preferable to the inequalities (34) if the user prefers a computational speed-up.

Example 3 (cont). Again, consider the CVaR-SMKP instance defined in Example 3. Consider two similar scenarios $\xi^{\omega'}$ and $\xi^{\omega''}$, where $\omega' = 2$ and $\omega'' = 1$. Solving the stage- t scenario- ξ^1 problem for $t = 2$ and $t = 3$, and using the solution in the objective function of the scenario- ξ^2 problem, we have $\hat{Z}_2^2 = 31.69$ and $\hat{Z}_3^2 = 57.44$, respectively.

For $\omega = 1 \in \Theta_{t,\xi_2}^-$, the inequalities corresponding to (38) for $t = 2, 3$ can be, respectively, written as:

$$p^1 \left(\sum_{i \in \mathcal{I}} c_{i2} x_{i2}^1 + d_2 z_2^1 + q_2^1 y_2^1 \right) + \lambda p^1 \left(\eta_2^1 + \frac{1}{1 - \alpha_2} v_2^1 \right) \geq 31.69,$$

$$p^1 \sum_{t \in \{2,3\}} \left(\sum_{i \in \mathcal{I}} c_{it} x_{it}^1 + d_t z_t^1 + q_t^1 y_t^1 \right) + \lambda p^1 \sum_{t \in \{2,3\}} \left(\eta_t^1 + \frac{1}{1 - \alpha_t} v_t^1 \right) \geq 57.44.$$

4 Computational experiments

4.1 Implementation specifications

In our computational experiments, we demonstrate the effectiveness of the proposed bounds and cuts driven by “scenario dominance” in solving risk-averse multi-stage stochastic knapsack instances with different problem characteristics. In particular, we solve a variety of CVaR-SMKP instances with the following approaches:

- *cpx*: direct solution of the model (28a) by CPLEX using its default settings.
- *bc*: inequalities (29) and (30).
- *scut*: inequalities (34).
- *sdcut*: the *scut* + inequalities (38).

The implementation of the cuts defined above involves appending them to the CVaR-SMKP formulations as CPLEX user cuts, and then solving the resulting model using CPLEX 12.71 with its default settings. The best feasible solution found within 50 CPU seconds (s) is fed into the optimizer before we run the model with the cuts in order to benefit from useful CPLEX cuts. The 50s spent a priori to adding **sdcut** is also included in the cut generation time. We used a formula of $\frac{2^{T-2}}{T-1}$ to determine the number of scenarios to be used for cut generation for solving the CVaR-SMKP. To drive this formula, we performed a preliminary experimental analysis on the fraction of the scenarios that are used for cut generation. After using various fractions, we observe that selecting the half of the scenarios divided by $T - 1$ strikes a good balance in terms of the cut generation time (which includes the solution of all selected scenario sub-problems) and the time benefit obtained from implementing the cuts. However, the number of scenarios to be used for cut generation could be set any number from 1 to Ω . We generated both *scut* and *sdcut* for periods $t = \{T/2, \dots, T\} \in \hat{T}$. The effectiveness of the proposed cuts is evaluated in terms of reducing the solution time, the integrality gap, the number of nodes solved in the branch and bound tree as well as the number of instances solved without memory problems relative to *cpx*, which is the direct solution of the model by CPLEX.

All computational testing was conducted on a desktop computer running with Intel i7 CPU and 64.0 GB of memory and using CPLEX 12.7.1. A time limit of 7200 CPU seconds was imposed for all test instances. We describe our test instance generation scheme for CVaR-SMKPs in Sect. 4.2, and present the detailed results in Sect. 4.3.

4.2 Instance generation and test problems

For each problem class, we randomly generate multiple instances using different seeds from a uniform distribution. All instances have the parameter T that defines the number of stages

in the stochastic scenario tree with $l = 2$ branches emanating from each node. Thus the resulting scenario tree has a size of 2^{T-1} scenarios. All CVaR-SMKP instances generated can be found at the Systems Optimization and Data Analytics Laboratory Mean-CVaR Stochastic Multi-Dimensional Mixed 0-1 Knapsack Test Problem Library (SODAL CVaRSMKP-LIB) (Büyüktaktakın 2021).

Knapsack instances We generate test instances for CVaR-SMKP in a similar fashion to the two-stage binary knapsack instances of Angulo et al. (2016). Different than the study of Angulo et al. (2016), which solved a two-stage risk-neutral version of the multi-item knapsack problem with a maximum of 120 scenarios and 55 constraints, in our study, we consider a larger, risk-averse version of the problem with up to $T = 10$ stages, 512 scenarios, and 67,072 constraints. Also, rather than using only binary variables as in the formulation of Angulo et al. (2016), we consider a mix of binary and continuous variables, further complicating the multi-item knapsack problem considered.

We assume that all parameter values at period $t = 1$ are set at zero since uncertainty starts at stage two. The parameters c_{it} , κ_{it} , d_t , a_{it} , and w_t are independent and identically distributed (i.i.d.) sampled from the uniform distribution over $\{1, \dots, 100\}$, e.g., $U[1, R]$ where $R = 100$ and $\kappa_{it} \in U[1, R]$ for $t = 2, \dots, T$. We have considered uncertainty in the objective function parameter q_t^ω , the left-hand side parameter r_t^ω , and the right-hand side parameter b_t^ω . For the 2-branch problem, all uncertain parameters have two levels: low (L) and high (H), where $(q_t^\omega, r_t^\omega) \in U[1, R/2]$ for the low level and $(q_t^\omega, r_t^\omega) \in U[R/2 + 1, R]$ for the high level and $b_t^\omega = \frac{3}{4} \left(\sum_{j=1}^t \sum_{i \in \mathcal{I}} a_{ij} + r_t^\omega \right)$ for $t = 2, \dots, T$. A scenario ξ^ω stage- t dominates a scenario $\xi^{\omega'}$ with respect to the problem (28a) if $(p^\omega \geq p^{\omega'})$, $(q_j^\omega \geq q_j^{\omega'})$, $(r_j^\omega \leq r_j^{\omega'})$, and $(b_j^\omega \geq b_j^{\omega'})$ hold for each $j = 2, \dots, t$ with $t \in \mathcal{T} \setminus \{1\}$.

We set $h_t = \frac{3}{4} (\sum_{i \in \mathcal{I}} \kappa_{it} + w_t)$. By employing this scheme, we randomly generate instances for various number of stages (T) and the number of items (I). Ten random feasible instances for each stage and item combination $(T, I) \in \{(5, 120), (6, 50), (7, 40), (8, 20), (9, 15), (10, 10)\}$ are generated, resulting in a total number of 60 instances. As we focus on large scale problems of the form (5a), we assume $p^\omega = 1/\bar{\Omega}$, where $\bar{\Omega}$ is the maximum number of scenarios, for the rest of the paper.

Those instances generated represent large-scale and hard CVaR-SMKP formulations with both the number of binary and continuous variables varying from 9656 to 71,140, and the number of constraints varying from 6096 to 67,072. The size and the difficulty of those instances increase with T and I .

4.3 Results for risk-averse multi-dimensional knapsack instances

In Tables 2, 3, 4, 5 and 6, we report results for a combination of columns defined as follows.

- T : number of stages (periods).
- I : number of items.
- exp : solution approach.
- $time$: CPU seconds required to solve the problem, excluding inequality generation time.
- $Ctime$: CPU seconds required to generate the cuts, including the solution of the single scenario sub-problems.
- $Ttime$: total CPU seconds required to solve the problem, including inequality generation time.

- $Tfac$: time factor improvement by $sdcut$ over cpx ($Tfac = \text{Time}^{cpx} / \text{Time}^{sdcut}$), where Time^{cpx} is the solution time ($time$) by cpx and Time^{sdcut} is the total solution time ($Ttime$) by bc , sdc , or $sdcut$.
- cut : number of cuts added to CPLEX as user inequalities.
- $node$ (100K): number of nodes explored in the branch and bound tree in 100,000s.
- gap^1 : final optimality gap by cpx at the end of the time limit.
- gap^2 : percentage deviation between the best objective value found by cpx (obj^{cpx}) and the best objective value obtained by the considered cutting-plane method at the end of the time limit, e.g., $gap^2 = 100 \times (obj^{sdcut} / obj^{cpx} - 1)$, where obj^{sdcut} defines the best objective value obtained by $sdcut$.
- $IGap$: percentage integrality gap of the formulation before inequalities are added ($InitGap = 100 \times (\text{bestobj} - \text{relaxlb}) / \text{bestobj}$), where relaxlb and bestobj are objective function values of the initial LP relaxation and the best feasible solution by cpx , respectively.
- $RGap$: percentage integrality gap of the formulation after inequalities are added ($RootGap = 100 \times (\text{bestobj} - \text{rootlb}) / \text{bestobj}$), where rootlb is the objective function value of the initial LP relaxation after $sdcut$ and bc are added.
- $GapImp$: percentage improvement in the integrality gap at the root node ($GapImp = 100 \times (1 - RGap / IGap)$).
- $unsol$: the number of instances unsolved to optimality within the given time limit.

In all tables, the values under the time-related metrics ($time$, $Ctime$, $Ttime$, and $Tfac$) are rounded to the closest integer to improve the clarity of the presentation.

4.4 Results on bounds and Sdc

Table 2 presents the initial gap ($IGap$), cut , $Ctime$, the root gap ($RGap$), and the optimality gap improvement ($GapImp$) over cpx using the bc and $sdcut$ together ($bc+sdcut$), where we solve both 2^{T-2} and 2^{T-1} number of scenario sub-problems for only one time period $t = T$. Here, we did not include the $sdcut$ because the strong dominance cuts (38) are rather aggressive and may potentially cut-off the optimal solutions. The $IGap$, $RGap$, and $GapImp$ columns are presented in Table 2. Each row of Table 2 presents results for an average of ten instances, while the *Average* row represents results for an average of 60 instances.

In Table 2, the percent initial gap shows an increasing trend as T increases, reaching 5.65% for instances with $T = 10$ and $I = 10$. From the table, the gap improvement due to $bc+sdcut$ also shows an increasing trend as T increases. The average integrality gap of 2.37% is reduced to 1.25% and 0.10% by the implementation of $bc+sdcut$ when 2^{T-2} and 2^{T-1} scenario sub-problems are solved, respectively. When cut generation is varied from solving 2^{T-2} to solving 2^{T-1} scenario sub-problems, the number of cuts generated as well as the cut generation time nearly double, as shown in Table 2. On the other hand, the average gap improvement due to $sdcut$ and bc is 46.4%, when 2^{T-2} scenarios are solved, while it increases to 94.3% when the number of scenario sub-problems solved increases to 2^{T-1} . In other words, the gap improvement due to the $bc + sdcut$ nearly doubles when the number of scenario sub-problems solved is doubled.

Table 5 in Appendix B presents results that compare the performance of cpx , bc , $sdcut$, and $sdcut$ instance-by-instance for the CVaR-SMKP test set with $T = 6$, $I = 50$, where $\lambda = 1$ and $\alpha_t = 0.95$. Each row of Table 5 represents an instance as specified under the *Instance* column, except the *Average* row block, which gives the average results over ten instances with $T = 6$ and $I = 50$. Here, bc and $sdcut$ are implemented by solving 2^{T-1} number of scenario sub-problems while $sdcut$ is generated using 2^{T-2} scenario sub-problems.

Table 2 Optimality gap results due to *bc* and *scut* for CVaR-SMKP instances over 5–10 stages

(T, I)	cpx IGap (%)	bc + scut (2^{T-2})				bc + scut (2^{T-1})			
		Cut	Ctime	RGap (%)	GapImp (%)	Cut	Ctime	RGap (%)	GapImp (%)
(5, 120)	0.19	10	7	0.10	45.7	18	16	0.02	89.7
(6, 50)	0.71	18	10	0.40	44.5	34	17	0.05	93.6
(7, 40)	0.93	34	22	0.52	46.5	66	42	0.05	94.9
(8, 20)	3.31	66	47	1.74	47.5	130	76	0.07	97.6
(9, 15)	3.41	130	106	1.84	46.1	258	239	0.21	93.6
(10, 10)	5.65	258	292	2.92	48.2	514	758	0.19	96.7
Average	2.37	86	80	1.25	46.4	170	191	0.10	94.3

In terms of the solution time performance, *scut* performs the best by reducing the *cpx* solution time from two hours to only 68 CPU seconds. The *bc* and *scut* do not help much reducing the solution time, except the instance number 4, where two hours of *cpx* solution time is reduced to 180 and 203 CPU seconds, respectively. The final MIP optimality gap of 0.08% by *cpx* is improved to 0.05 and 0.04% by *bc* and *scut*, respectively. The main benefit of *bc* and *scut* is observed in reducing the integrality gap by 93.2% on average, as shown under the *GapImp* column. Since the *bc* and *scut* would provide the highest benefit to close the gap from the linear programming relaxation of the problem, we recommend using both *bc* and *scut* to close the integrality gap in the state-of-the-art solvers, such as CPLEX, rather than adding those inequalities as cuts and solving the augmented problem for computational speed-up.

4.5 Results on Sdcut

Table 3 summarizes results using *scut* for instances with a variety of stages ($T = 5 - 10$) and items ($I = 10 - 120$) where $\lambda = 1$ and $\alpha_t = 0.95$. Each row of Table 3 presents results for an average of ten instances, while *Average (Total)* row reports the overall average (total) over sixty instances.

Here, we do not include the *bc* because the initial experiments with the lower and upper bounding inequalities (29) and (30), shown in Section 4.5, do not provide any solution speed-up over the *scut*, despite they close the initial integrality gap substantially. The difficulty of the instances presented in Table 3 is a factor of the time stage T as well as the number of items I . In all implementations, *scut* significantly reduces the solution time by *cpx*. For example, the solution time of 7287 is reduced to 68 CPU seconds for instances with $T = 6$ and $I = 50$, and for the $T = 10$ and $I = 10$ instances, the solution time of 5780 is reduced to 227 CPU seconds due to the *scut*. In particular, the average solution time is reduced from 6,559 to 111 seconds, where 98 seconds are spent for cut generation, including the solution of the scenario sub-problems. When the total computation time is compared, a total of 39,354 solution time of the *cpx* is reduced to 669 CPU seconds by the *scut* to solve all instances. While *scut* improves the solution time of the $T = 5$ and $I = 120$ instances with a factor of 110, it improves the solution speed of the $T = 10$ and $I = 10$ instances with a factor of 26.

Table 3 Experiments for CVaR-SMKP instances with $\lambda = 1$ and $\alpha_t = 0.95$ over 5–10 stages

(T, I)	Exp	Time	Ctime	Ttime	Tfac	Cut	Node (100K)	Gap ¹ (%)	Gap ² (%)	Unsol
(5, 120)	<i>cpx</i>	6534	0	6534		0	14.2	0.02		9
	<i>sdcut</i>	6	53	60	110	21	0.2		0.01	0
(6, 50)	<i>cpx</i>	7287	0	7287		0	5.8	0.08		10
	<i>sdcut</i>	12	56	68	108	35	0.2		0.02	0
(7, 40)	<i>cpx</i>	6053	0	6,053		0	3.6	0.07		9
	<i>sdcut</i>	15	75	90	67	135	0.0		0.02	0
(8, 20)	<i>cpx</i>	7207	0	7207		0	6.3	0.07		8
	<i>sdcut</i>	41	70	111	65	255	0.3		0.00	0
(9, 15)	<i>cpx</i>	6492	0	6493		0	4.2	0.12		9
	<i>sdcut</i>	1	112	114	57	961	0.1		0.05	0
(10, 10)	<i>cpx</i>	5774	0	5780		0	3.5	0.15		8
	<i>sdcut</i>	1	220	227	25	1705	0.1		0.08	0
Average	<i>cpx</i>	6558	0	6559		0	6.3	0.08		9
	<i>sdcut</i>	13	98	111	72	519	0.1		0.03	0
Total	<i>cpx</i>	39,348	0	39,354	0	0	37.6	0.59	0	53
	<i>sdcut</i>	76	586	669	432	3112	22.2		0.19	0

As shown in the overall averages in Table 3, *sdcut* improve results by a factor of 72, when all of the time factor improvement due to *sdcut* is averaged. The solution time reduction due to the *sdcut* is also consistent with the reduction in the number of nodes searched in the branch and bound tree. As the time period is increased from $T = 5$ to $T = 10$, the number of scenarios increases from 16 to 512. Because $\frac{2^{T-2}}{T-1} * \frac{T-1}{2}$ scenarios are solved as single scenario sub-problems, the time required for the cut generation (Ctime) shows an increasing trend as T increases. The average gap that defines the percentage deviation between *obj* obtained by *cpx* and *obj* by *sdcut* is 0.03%. This means that when all instances from $T = 5$ to $T = 10$ are averaged, the objective value provided by *sdcut* is 0.03% higher than the objective value provided by *cpx*. Out of 60 instances, *cpx* could not solve 53 of them to optimality within the 2-h of time limit, as shown under the *unsol* of Table 3, while the implementation with *sdcut* solves all of the 60 instances with only 0.03% deviation from *cpx*, within the 2-hour of time limit.

The *Total* row shows that a total time of 39,354 CPU seconds for solving 60 problems is reduced to only 669 CPU seconds (less than 11 minutes) by using *sdcut*, resulting in a time factor improvement of 432 for solving all instances presented in this table. Varying risk parameters does not change the performance of *sdcut* much (see detailed results for varying risk parameters in Table 6 and the related discussion under Appendix C).

4.6 Generalization to multiple random variables

We perform additional experiments to demonstrate the number of scenario-dominance relations and its impact on the performance of the *sdcut* as we vary the number of random variables. For those experiments, we have generated additional instances where we re-defined

the parameter c_{it} as an uncertain parameter c_{it}^ω , which represents the cost coefficient associated with variable x_{it}^ω for item i in period t under scenario ω , and assumed that all other parameters are deterministic. The random parameters c_{it}^ω are stage-wise independent but are jointly distributed over i for each $t \in \mathcal{T}$. The c_{it}^ω could be positively or negatively correlated for each item i . A scenario ξ^ω stage- t dominates a scenario $\xi^{\omega'}$ with respect to the problem (28a) if $(p^\omega \geq p^{\omega'})$ and $(c_{it}^\omega \geq c_{it}^{\omega'})$ hold for each $i \in \mathcal{I}$ and $j = 2, \dots, t$ with $t \in \mathcal{T} \setminus \{1\}$. Since the case where each univariate component of the random vector is independently distributed could explode the size of the scenario tree, here we generated random variables as stage-wise independent but jointly distributed at each stage. However, results provide insights into the former case as well.

In Table 4, we compare the performance of *sdcut* with respect to *cpX* for a varying number of random variables in the scenario tree for instances with $T = 10$ and $I = 10$. In this table, \hat{I} represents the number of randomly generated items in each stage t , RV shows the total number of random variables generated, $Sd(t = 2)$ and $Sd(t = T)$ represent the number of scenario-dominance relations at stages two and T , and $Sd(total)$ presents the total number of scenario-dominance relations over all time periods for each RV . For example, instances with $\hat{I} = 0$ and $RV = 0$ are deterministic, while $\hat{I} = 3$ represents instances with three items randomly generated in each stage and thus that have a total of $\hat{I}(T - 1) = 27$ random variables since the values of the random parameters at stage 1 are assumed to be known.

Each row of Table 4 presents results for an average of ten instances. Since random variables are stage-wise independent but are dependent in each stage, the number of scenarios for each instance with a varying number of random variables is 512.

According to the average solution times in Table 4, *sdcut* consistently improves the solution times of *cpX* for instances with an increasing number of random variables by an overall average factor of 25, reducing the average solution time of 5,716 seconds to 231 seconds. As the number of random variables increases, the number of second-stage and last-stage scenario-dominance relations reduces. Note that the second-stage dominance relations are much larger than the last-stage dominance relations. When we reach 81 random variables, we also hit the minimum number scenario relations at stages two and T , and in total, as shown by Proposition 3. For the considered instances, having the minimum number scenario relations does not impact the performance of *sdcut* with respect to *cpX*. For example, when we have $\hat{I} = 10$ and $RV = 90$, we have the least total number of scenario-dominance relations ($W = 261, 632$), and 1705 cuts are generated under *sdcut*. While *mipgap* tends to increase as the number of random variables increase, the performance of *sdcut* is not impacted by an increased number of random variables. The results imply that when we increase the number of random variables in each stage, the algorithm still generates plenty of cutting planes due to an ample number of stage- t dominance relations.

5 Concluding remarks

In this paper, we introduced a new concept called “scenario dominance” and “stage- t scenario dominance” to effectively solve the general risk-averse multi-stage stochastic mixed-integer programs. This method derives the implications of one scenario on another by comparing the problem coefficients under each scenario and the optimal objective function values of the individual scenario sub-problems. We derive results on the minimum number of stage- t scenario-dominance relations in a multi-stage scenario tree. By studying those implications and the partial ordering of scenarios, we derive new bounds and cutting planes to tackle risk-

Table 4 Results for instances with $T = 10$, $I = 10$, and a varying number of random variables RV

Exp	\hat{I}	RV	Sd ($t = 2$)	Sd ($t = T$)	Sd (total)	Cut	Ttime	Gap ¹ (%)	Gap ² (%)
<i>cpx</i>	0	0	262, 144	262, 144	2, 359, 296	0	1460	0.03	
<i>sdcut</i>						140, 800	216		0.00
<i>cpx</i>	1	9	196, 608	19, 683	727, 383	0	7223	0.09	
<i>sdcut</i>						17, 319	242		0.03
<i>cpx</i>	2	18	157, 286	3, 889	399, 414	0	5394	0.07	
<i>sdcut</i>						4867	226		0.03
<i>cpx</i>	3	27	150, 733	1563	328, 533	0	5321	0.09	
<i>sdcut</i>						3225	232		0.05
<i>cpx</i>	4	36	131, 072	742	274, 202	0	5843	0.13	
<i>sdcut</i>						2196	237		0.07
<i>cpx</i>	5	45	137, 626	710	277, 011	0	5787	0.09	
<i>sdcut</i>						2142	231		0.04
<i>cpx</i>	6	54	137, 626	538	274, 714	0	6510	0.20	
<i>sdcut</i>						1876	231		0.14
<i>cpx</i>	7	63	137, 626	563	275, 507	0	6500	0.17	
<i>sdcut</i>						1779	230		0.12
<i>cpx</i>	8	72	137, 626	627	281, 997	0	5781	0.13	
<i>sdcut</i>						2046	227		0.11
<i>cpx</i>	9	81	131, 072	512	261, 632	0	5842	0.11	
<i>sdcut</i>						1705	232		0.05
<i>cpx</i>	10	90	131, 072	512	261, 632	0	7216	0.19	
<i>sdcut</i>						1705	236		0.11
<i>cpx</i>	Ave	45	155, 499	26, 499	520, 120	0	5716	0.12	
<i>sdcut</i>						16, 333	231		0.07

averse SMIPs. We perform extensive computational experiments using the proposed cuts on the mean-CVaR stochastic dynamic knapsack problem. Our computations demonstrate that the proposed methodology reduces the solution time of the large-scale, risk-averse multi-stage stochastic mixed-integer optimization problems with varying risk characteristics by one-to-two orders of magnitude. Results also demonstrate that strong dominance cuts perform well for those instances with ten random variables in each stage and ninety random variables in total. They also have the potential to tackle a larger number of random variables because one can generate plenty of stage- t scenario-dominance relations even under the worst-case scenario.

This study opens several exciting research directions worth investigating in the future. For example, “scenario dominance” approach could be investigated for other risk measures, such as the Expected Excess of a given target and the Excess Probability (Schultz and Tiedemann 2003). In this paper, we only used a small subset of the full scenario set to generate the scenario dominance cuts. A future experimental study could analyze the impact of using different fractions and subsets of scenarios and various scenario selection schemes on the quality of the bounds as well as the cuts proposed. Decomposition methods involving Lagrangian relaxation (CarøE and Schultz 1999; Nowak and Römisich 2000), column generation (Singh et al.

2009; Sen 2005), L-shaped decomposition (Angulo et al. 2016; Laporte and Louveaux 1993), progressive hedging (Watson and Woodruff 2011; Gade et al. 2016), scenario-decomposition (Birge 1985; Sandıkçı and Özaltın 2017), and stochastic dual dynamic integer programming (Zou et al. 2019) have been quite useful in solving various classes of multi-stage stochastic integer programs. Thus, another important future research direction is to integrate the proposed “scenario dominance” cut-and-bound generation scheme within a decomposition framework for solving mean-risk M-SMIPs.

In this paper, we focus on the type of stochastic multi-dimensional knapsack instances up to $T = 10$ stages, 10 random variables in each stage, and $I = 120$ items. The considered CVaR-SMKP deterministic counterparts, which involve the expected values of the uncertain parameters, are also computationally challenging. Future research could investigate a variant of CVaR-SMKP instances, which have easier deterministic formulations but a larger number of scenarios than considered in this work, and thus remain computationally hard to solve. For example, a future extension of this work could consider $T = 5$ -stage $I = 1$ -item CVaR-SMKPs with $\hat{I} = 10$ independent random variables and 10 branches in each stage, with 10,000 scenarios. To tackle an instance like this, a parallel distributed platform would be more suitable for solving the individual scenario problems, reducing the dominance cut generation time, and thus improving the overall solution performance. In such a case, based on Proposition 4, we would have a minimum of over more than a hundred million scenario dominance relations, leading to an enormous number of scenario dominance cuts. Thus another important direction for research is to determine the most useful subset and the number of dominance cuts, which would not excessively increase the size of the problem formulation.

The concept of “scenario dominance” as defined in this paper is fundamentally different than the “stochastic dominance” relation defined in the literature (Müller and Stoyan 2002; Luedtke 2008) because two random variables are compared based on their probability distribution function in stochastic dominance, while in our scenario dominance approach we pairwise compare scenarios based on the realization of uncertain parameters at each time stage. Specifically, scenario dominance derives relations among scenarios by comparing possible outcomes of a single random variable or potential joint outcomes of more than one random variable. Yet, an interesting future direction would be to investigate the theoretical relations between the scenario dominance and the first (second)-order stochastic dominance.

The proposed approach in this paper does not require convexity or linearity, and thus can be extended to study the non-linear case and various types of risk-averse multi-stage stochastic MIPs. The generality of our method would require careful implementation of the proposed cuts for each specific problem, such as risk-averse multi-stage stochastic financial investment planning, capacity allocation, and network optimization problems.

Our approach also has some limitations. For example, problems with several uncorrelated random variables may exhibit less scenario-dominance relations and fewer dominance cuts than considered in this study. Therefore, an important future direction is to investigate further the boundaries of the scenario dominance cuts by exploring the breadth of problems that could benefit from our approach.

Acknowledgements The author gratefully acknowledges the partial support of the National Science Foundation CAREER Award co-funded by the CBET/ENG Environmental Sustainability program and the Division of Mathematical Sciences in MPS/NSF Under CMMI-1100765 and Grant # CBET-1554018. We thank the Academic Research and Computing Systems (ARCS) center of NJIT for their support with the NJIT ARCS Kong cluster, which is used to perform our computational testing. We thank three anonymous referees, the associate editor, and the editor, whose remarks helped to improve the content and clarity of our exposition.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendices

A Proof of Proposition 1

Proof By the multi-period CVaR formulation (3) in Definition 1.1 the following formulation holds:

$$\sum_{t \in T \setminus \{1\}} \min_{\eta_t \in \mathbb{R}} \left\{ \eta_t + \frac{1}{1 - \alpha_t} \mathbb{E} \left[\left(c_t(\xi_{[t]})^\top x_t - \eta_t \right)_+ \mid \xi_{[t-1]} \right] \right\}. \tag{39}$$

The formulation (39) can be rewritten as follows:

$$\min_{(\eta_2, \dots, \eta_T) \in \mathbb{R}^{T-1}} \left\{ \sum_{t \in T \setminus \{1\}} \left(\eta_t + \frac{1}{1 - \alpha_t} \mathbb{E} \left[\left(c_t(\xi_{[t]})^\top x_t - \eta_t \right)_+ \mid \xi_{[t-1]} \right] \right) \right\}. \tag{40}$$

In order to linearize the $(\bullet)_+$ term in (40), we define a non-negative variable $v_t \in \mathbb{R}_+$ such that $v_t = (c_t(\xi_{[t]})^\top x_t - \eta_t)_+$ and replace the nonlinear term $(\bullet)_+$ with the following two types of constraints for each $t = 2, \dots, T$:

$$v_t \geq 0, \tag{41}$$

$$v_t \geq c_t(\xi_{[t]})^\top x_t - \eta_t. \tag{42}$$

Replicating the v_t and η_t variables and the constraints (41) and (42) for each scenario $\omega \in \Omega$, the equation (40) can be cast as the following equation:

$$\begin{aligned} & \min_{\substack{x_1^\omega, \dots, x_T^\omega, \\ \eta_1^\omega, \dots, \eta_T^\omega, \\ v_1^\omega, \dots, v_T^\omega}} \left\{ \sum_{\omega \in \Omega} p^\omega \sum_{t \in T \setminus \{1\}} \left(\eta_t^\omega + \frac{1}{1 - \alpha_t} v_t^\omega \right) : v_t^\omega \right. \\ & \geq c_t^\omega{}^\top x_t^\omega - \eta_t^\omega \quad t \in T \setminus \{1\}, \forall \omega \in \Omega; \quad v_t^\omega \geq 0, \eta_t^\omega, \\ & \left. x_t^\omega \in \mathbb{R}_+^{n_t - q_t} \times \mathbb{Z}_+^{q_t}, \quad \forall t \in T, \forall \omega \in \Omega \right\}. \tag{43} \end{aligned}$$

For each stage t , η_t^ω is a “stage- $(t - 1)$ variable”, representing the value-at-risk, VaR_{α_t} , and v_t^ω , is an auxiliary “stage- t variable,” to represent the stage- t excess cost, exceeding η_t^ω . Similar to the x_t^ω -variables in E-MSMIP (2a), we define non-anticipativity constraints on the v_t^ω variables to ensure the time consistency and thus implementability of solutions as follows:

$$v_t^\omega = v_t^{\omega'}, \quad \forall t \in T, \omega, \omega' \in \Omega \text{ s.t. } \xi_{[t]}^\omega = \xi_{[t]}^{\omega'}. \tag{44}$$

Integrating the objective function and constraints of (43) along with corresponding non-anticipativity constraints (44) into the E-MSMIP (2a) leads to the mean-CVaR multi-stage stochastic MIP (CVaR-M-SMIP) problem (5a). □

B Results for bounds and cuts

Table 5 presents results that compare the performance of *cpx*, *bc*, *scut*, and *sdcut* instance-by-instance for the CVaR-SMKP test set with $T = 6$, $I = 50$, where $\lambda = 1$ and $\alpha_t = 0.95$.

Table 5 Experiments for CVaR-SMKP instances with $T = 6$ and $I = 50$

Instance	Exp	Cut	Ctime	Ttime	Tfac	Gap ¹	Gap ²	GapImp
1	<i>cpx</i>	0	0	7307		0.12		
	<i>bc</i>	2	144	7402	1	0.05		93.0
	<i>sdc</i>	224	114	7326	1	0.05		93.0
	<i>sdcut</i>	35	63	85	86		0.05	
2	<i>cpx</i>	0	0	7337		0.08		
	<i>bc</i>	2	112	7440	1	0.04		94.3
	<i>sdc</i>	224	115	7342	1	0.04		94.3
	<i>sdcut</i>	35	55	60	122		0.01	
3	<i>cpx</i>	0	0	7323		0.09		
	<i>bc</i>	2	133	7444	1	0.08		88.9
	<i>sdc</i>	224	123	7326	1	0.07		88.9
	<i>sdcut</i>	35	54	55	134		0.01	
4	<i>cpx</i>	0	0	7251		0.03		
	<i>bc</i>	2	139	180	40	0.01		98.4
	<i>sdc</i>	224	138	203	36	0.01		98.4
	<i>sdcut</i>	35	53	53	136		0.00	
5	<i>cpx</i>	0	0	7241		0.07		
	<i>bc</i>	2	92	7365	1	0.05		89.8
	<i>sdc</i>	224	129	7340	1	0.05		89.8
	<i>sdcut</i>	35	53	66	110		0.01	
6	<i>cpx</i>	0	0	7309		0.11		
	<i>bc</i>	2	146	7451	1	0.07		95.3
	<i>sdc</i>	224	100	7397	1	0.04		95.3
	<i>sdcut</i>	35	56	57	129		0.02	
7	<i>cpx</i>	0	0	7,290		0.10		
	<i>bc</i>	2	139	7357	1	0.07		95.1
	<i>sdc</i>	224	137	7340	1	0.06		95.1
	<i>sdcut</i>	34	58	68	107		0.01	
8	<i>cpx</i>	0	0	7269		0.06		
	<i>bc</i>	2	94	7376	1	0.06		89.3
	<i>sdc</i>	224	130	7466	1	0.03		89.3
	<i>sdcut</i>	35	54	56	130		0.00	
9	<i>cpx</i>	0	0	7241		0.08		
	<i>bc</i>	2	115	7393	1	0.06		92.0
	<i>sdc</i>	224	86	7311	1	0.05		92.0
	<i>sdcut</i>	34	53	65	111		0.00	

Table 5 continued

Instance	Exp	Cut	Ctime	Ttime	Tfac	Gap ¹	Gap ²	GapImp
10	<i>cpx</i>	0	0	7305		0.05		
	<i>bc</i>	2	141	7355	1	0.05		96.4
	<i>sdcut</i>	224	115	7330	1	0.03		96.4
	<i>sdcut</i>	35	57	109	67		0.04	
Average	<i>cpx</i>	0	0	7287		0.08		
	<i>bc</i>	2	125	6676	5	0.05		93.2
	<i>sdcut</i>	224	119	6638	4	0.04		93.2
	<i>sdcut</i>	35	56	68	108		0.02	

C Results for varying risk parameters

Table 6 summarizes results for $(T, I) = (5, 120)$ with varying values of risk parameters λ and α_t . In particular, we consider 15 combinations of λ ranging from 0.001 to 1000 and $\alpha_t \in \{0.7, 0.95\}$. From each combination, three instances are generated with $(T, I) = (5, 120)$, resulting in 45 instances that differ in their λ and α_t values, thus each row of Table 6 presents an average of the results for three instances. The *Ttime (cpx)* and *Ttime (sdcut)* columns of Table 6 gives the CPU times of the CPLEX performance for solving the CVaR-SMKP instances without cuts, and with *sdcut*, respectively, *Ctime* gives the *sdcut* generation time, *gap*¹ presents the *cpx* MIP gap, and *gap*² gives the percentage deviation between the best objective value found by *cpx* and *sdcut*.

According to the total and average solution times, the solution times of *cpx* and *sdcut* are similar for varying λ and α_t values. Thus we do not observe a significant change in instance difficulty and computational times with respect to the λ and α parameters. For all cases with varying λ and α_t values, the *sdcut* drastically improves the performance with respect to *cpx* with a factor over 110 either when averaged or in total. Out of 45 cases, 45 of them cannot be solved to optimality by *cpx* within the 2-h limit, while all instances are solved within 2 minutes with the *sdcut* with an average optimality gap of 0.01%.

Table 6 Experiments for instances with $T = 5$ and $I = 120$ for various λ and α_t values

λ	$\alpha_t = 0.7$				$\alpha_t = 0.95$			
	Ttime (cpx)	Ttime (sdcut)	Gap ¹ (%)	Gap ² (%)	Ttime (cpx)	Ttime (sdcut)	Gap ¹ (%)	Gap ² (%)
0.001	7253	61	0.03	0.00	7255	65	0.03	0.00
0.1	7254	60	0.04	0.00	7251	62	0.03	0.01
1	7283	69	0.03	0.01	7283	67	0.01	0.01
10	7249	65	0.04	0.01	7249	70	0.03	0.01
1000	7271	69	0.03	0.01	7271	68	0.03	0.01
Ave	7262	65	0.03	0.01	7262	67	0.03	0.01
Total	36,310	323	0.17	0.03	36,310	333	0.13	0.03

References

- Ahmed, S. (2006). Convexity and decomposition of mean-risk stochastic programs. *Mathematical Programming*, 106(3), 433–446.
- Ahmed, S. (2013). A scenario decomposition algorithm for 0–1 stochastic programs. *Operations Research Letters*, 41(6), 565–569.
- Alonso-Ayuso, A., Escudero, L. F., Guignard, M., & Weintraub, A. (2018). Risk management for forestry planning under uncertainty in demand and prices. *European Journal of Operational Research*, 267(3), 1051–1074.
- Angulo, G., Ahmed, S., & Dey, S. S. (2016). Improving the integer L-shaped method. *INFORMS Journal on Computing*, 28(3), 483–499.
- Bakir, I., Boland, N., Dandurand, B., & Erera, A. (2020). Sampling scenario set partition dual bounds for multistage stochastic programs. *INFORMS Journal on Computing*, 32(1), 145–163.
- Bellman, R. (1957). A Markovian decision process. *Journal of mathematics and mechanics*, 6(5), 679–684.
- Benders, J. F. (2005). Partitioning procedures for solving mixed-variables programming problems. *Computational Management Science*, 2(1), 3–19.
- Birge, J. R. (1985). Aggregation bounds in stochastic linear programming. *Mathematical Programming*, 31(1), 25–41.
- Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming*. Springer.
- Boyd, E. A. (1993). Polyhedral results for the precedence-constrained knapsack problem. *Discrete Applied Mathematics*, 41(3), 185–201.
- Büyükahtakin, İ. E. (2011). *Dynamic programming via linear programming*. Wiley Encyclopedia of Operations Research and Management Science. Wiley.
- Büyükahtakin, İ. E. (2021). SODAL CVaRSMKP-LIB: Systems optimization and data analytics laboratory mean-CVaR stochastic multi-dimensional mixed 0–1 knapsack test problem library. <https://web.njit.edu/~esratoy/Research.html>
- CarøE, C. C., & Schultz, R. (1999). Dual decomposition in stochastic integer programming. *Operations Research Letters*, 24(1–2), 37–45.
- Csirik, J. (1991). Heuristics for the 0–1 min-knapsack problem. *Acta Cybernetica*, 10(1–2), 15–20.
- Dean, B. C., Goemans, M. X., & Vondrák, J. (2008). Approximating the stochastic knapsack problem: The benefit of adaptivity. *Mathematics of Operations Research*, 33(4), 945–964.
- Dentcheva, D., & Martinez, G. (2012). Two-stage stochastic optimization problems with stochastic ordering constraints on the recourse. *European Journal of Operational Research*, 219(1), 1–8.
- Dupačová, J., Consigli, G., & Wallace, S. W. (2000). Scenarios for multistage stochastic programs. *Annals of operations research*, 100(1), 25–53.
- Dupačová, J., & Kozmík, V. (2015). Structure of risk-averse multistage stochastic programs. *OR spectrum*, 37(3), 559–582.
- Eichhorn, A., & Römisch, W. (2005). Polyhedral risk measures in stochastic programming. *SIAM Journal on Optimization*, 16(1), 69–95.
- Escudero, L. F., Garín, M. A., Merino, M., & Pérez, G. (2016). On time stochastic dominance induced by mixed integer-linear recourse in multistage stochastic programs. *European Journal of Operational Research*, 249(1), 164–176.
- Fábián, C. I. (2008). Handling CVaR objectives and constraints in two-stage stochastic models. *European Journal of Operational Research*, 191(3), 888–911.
- Gade, D., Hackebeil, G., Ryan, S. M., Watson, J.-P., Wets, R.J.-B., & Woodruff, D. L. (2016). Obtaining lower bounds from the progressive hedging algorithm for stochastic mixed-integer programs. *Mathematical Programming*, 157(1), 47–67.
- Guan, Y., Ahmed, S., & Nemhauser, G. L. (2009). Cutting planes for multistage stochastic integer programs. *Operations Research*, 57(2), 287–298.
- Güntzer, M. M., & Jungnickel, D. (2000). Approximate minimization algorithms for the 0/1 knapsack and subset-sum problem. *Operations Research Letters*, 26(2), 55–66.
- Han, X., & K. Makino (2009). Online minimization knapsack problem. In *International workshop on approximation and online algorithms* (pp. 182–193). Springer.
- Heinze, T., & Schultz, R. (2008). A branch-and-bound method for multistage stochastic integer programs with risk objectives. *Optimization*, 57(2), 277–293.
- Heitsch, H., & Römisch, W. (2009). Scenario tree reduction for multistage stochastic programs. *Computational Management Science*, 6(2), 117–133.
- Homem-de Mello, T., & Pagnoncelli, B. K. (2016). Risk aversion in multistage stochastic programming: A modeling and algorithmic perspective. *European Journal of Operational Research*, 249(1), 188–199.

- Israeli, E. & Wood, R. K. (2002). Shortest-path network interdiction. *Networks: An International Journal* 40(2), 97–111.
- Kellerer, H., Pferschy, U., & Pisinger, D. (2004). Multidimensional knapsack problems. In *Knapsack problems* (pp. 235–283). Springer.
- Kibis, E., Büyüktaktın, İ. E., Haight, R.G., Akhundov, N., Knight, K., & Flower, C. (2020). A new multi-stage stochastic programming model and cutting planes for the optimal surveillance and control of emerald ash borer in cities. *Forthcoming in INFORMS Journal on Computing*.
- Laporte, G., & Louveaux, F. V. (1993). The integer L-shaped method for stochastic integer programs with complete recourse. *Operations Research Letters*, 13(3), 133–142.
- Luedtke, J. (2008). New formulations for optimization under stochastic dominance constraints. *SIAM Journal on Optimization*, 19(3), 1433–1450.
- Lulli, G., & Sen, S. (2004). A branch-and-price algorithm for multistage stochastic integer programming with application to stochastic batch-sizing problems. *Management Science*, 50(6), 786–796.
- Madansky, A. (1960). Inequalities for stochastic linear programming problems. *Management Science*, 6(2), 197–204.
- Maggioni, F., Allevi, E., & Bertocchi, M. (2014). Bounds in multistage linear stochastic programming. *Journal of Optimization Theory and Applications*, 163(1), 200–229.
- Maggioni, F., Allevi, E., & Bertocchi, M. (2016). Monotonic bounds in multistage 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.
- Maggioni, F., & Pflug, G. C. (2019). Guaranteed bounds for general nondiscrete multistage risk-averse stochastic optimization programs. *SIAM Journal on Optimization*, 29(1), 454–483.
- Mahmutoğulları, A. İ., Çavuş, Ö., & Aktürk, M. S. (2018). Bounds on risk-averse mixed-integer multi-stage stochastic programming problems with mean-CVaR. *European Journal of Operational Research*, 266(2), 595–608.
- Morton, D. P. & Wood, R. K. (1998). On a stochastic knapsack problem and generalizations. In *Advances in computational and stochastic optimization, logic programming, and heuristic search* (pp. 149–168). Springer.
- Müller, A., & Stoyan, D. (2002). *Comparison methods for stochastic models and risks* (Vol. 389). Wiley.
- Nowak, M. P., & Römisich, W. (2000). Stochastic Lagrangian relaxation applied to power scheduling in a hydro-thermal system under uncertainty. *Annals of Operations Research*, 100(1–4), 251–272.
- Ogryczak, W., & Ruszczyński, A. (2002). Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization*, 13(1), 60–78.
- Park, K., & Park, S. (1997). Lifting cover inequalities for the precedence-constrained knapsack problem. *Discrete Applied Mathematics*, 72(3), 219–241.
- Pflug, G. C. (2000). Some remarks on the value-at-risk and the conditional value-at-risk. In *Probabilistic constrained optimization* (pp. 272–281). Springer.
- Pflug, G. C., & Pichler, A. (2015). Dynamic generation of scenario trees. *Computational Optimization and Applications*, 62(3), 641–668.
- Rockafellar, R. T., Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2, 21–42.
- Ruszczyński, A. (2002). Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra. *Mathematical Programming*, 93(2), 195–215.
- Ruszczyński, A. (2013). Advances in risk-averse optimization. In *Theory Driven by Influential Applications* (pp. 168–190). INFORMS.
- Sandıkçı, B., Kong, N., & Schaefer, A. J. (2013). A hierarchy of bounds for stochastic mixed-integer programs. *Mathematical Programming*, 138(1), 253–272.
- Sandıkçı, B., & Özalın, O. (2017). A scalable bounding method for multistage stochastic programs. *SIAM Journal on Optimization*, 27(3), 1772–1800.
- Schultz, R. (2005). Risk aversion in two-stage stochastic integer programming. *Stochastic Programming*, 150, 165–187.
- Schultz, R., & Tiedemann, S. (2003). Risk aversion via excess probabilities in stochastic programs with mixed-integer recourse. *SIAM Journal on Optimization*, 14(1), 115–138.
- Schultz, R., & Tiedemann, S. (2006). Conditional value-at-risk in stochastic programs with mixed-integer recourse. *Mathematical Programming*, 105(2–3), 365–386.
- Sen, S. (2005). Algorithms for stochastic mixed-integer programming models. *Handbooks in Operations Research and Management Science*, 12, 515–558.
- 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.

- Singh, K. J., Philpott, A. B., & Wood, R. K. (2009). Dantzig-Wolfe decomposition for solving multistage stochastic capacity-planning problems. *Operations Research*, *57*(5), 1271–1286.
- Watson, J.-P., & Woodruff, D. L. (2011). Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. *Computational Management Science*, *8*(4), 355–370.
- Yin, X., Büyüktaktın, İ. E. (2021). Risk-averse multi-stage stochastic programming to optimizing vaccine allocation and treatment logistics for effective epidemic response. *IIEE Transactions on Healthcare Systems Engineering*, 1–52 (just-accepted).
- Yu, X., Shen, S., & Ahmed, S. (2021). On the Value of Multistage Stochastic Facility Location with Risk Aversion. arXiv preprint [arXiv:2105.11005](https://arxiv.org/abs/2105.11005).
- Zou, J., Ahmed, S., & Sun, X. A. (2019). Stochastic dual dynamic integer programming. *Mathematical Programming*, *175*(1–2), 461–502.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.