

---

# Distributionally Risk-Receptive and Robust Multistage Stochastic Integer Programs and Interdiction Models

Sumin Kang · Manish Bansal

First Version: June 2023; Latest Version: Sep 2024

**Abstract** In this paper, we study distributionally risk-receptive and distributionally robust (or risk-averse) multistage stochastic mixed-integer programs (denoted by DRR- and DRO-MSIPs). We present cutting plane-based and reformulation-based approaches for solving DRR- and DRO-MSIPs without and with decision-dependent uncertainty to optimality. We show that these approaches are finitely convergent with probability one. Furthermore, we introduce generalizations of DRR- and DRO-MSIPs by presenting multistage stochastic disjunctive programs and algorithms for solving them. These frameworks are useful for optimization problems under uncertainty where the focus is on analyzing outcomes based on multiple decision-makers' differing perspectives, such as interdiction problems that are attacker-defender games having non-cooperative players. To assess the performance of the algorithms for DRR- and DRO-MSIPs, we consider instances of distributionally ambiguous multistage maximum flow and facility location interdiction problems that are important in their own right. Based on our computational results, we observe that the cutting plane-based approaches are 2800% and 2410% (on average) faster than the reformulation-based approaches for the foregoing instances with distributional risk-aversion and risk-receptiveness, respectively. Additionally, we conducted out-of-sample tests to showcase the significance of the DRR framework in revealing network vulnerabilities and also in mitigating the impact of data corruption.

---

This research is partially funded by National Science Foundation Grant CMMI-1824897 and 2034503, and Commonwealth Cyber Initiative grants which are gratefully acknowledged.

Sumin Kang · Manish Bansal

Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA 24061, USA

E-mail: suminkang@vt.edu

E-mail: bansal@vt.edu

**Keywords** distributionally risk-receptive optimization · distributionally robust optimization · multistage stochastic programs · disjunctive programs · network interdiction · robust statistics · decision-dependent uncertainty

**Mathematics Subject Classification (2020)** 90C11 · 90C15

## 1 Introduction

Stochastic programming is a framework for modeling optimization problems with uncertain input data, assuming that the probability distribution associated with the uncertain parameters is known. Typically, it employs a probability distribution that approximates the *true* probability distribution (which is often not known in practice) using prior knowledge or historical data. When the complete information of the distribution is not known, distributionally robust optimization provides a framework to a decision-maker for modeling risk-aversion to such *distributional ambiguity* [39]. However, optimization problems often involve multiple decision-makers, and the risk-aversion modeled by distributionally robust optimization may not be sufficient when the focus is on analyzing outcomes from their diverse perspectives.

For an example, consider a two-player Stackelberg game played between an attacker and a defender with distributional ambiguity. The attacker with the aim of degrading the defender’s performance makes their decisions before the defender, thereby influencing the defender’s subsequent decision-making process. These games are often used to analyze vulnerabilities in the defender’s system/network (refer to [10, 14, 40, 24] for specific examples). In this adversarial context, considering the attacker as risk-averse would represent an optimistic view of the defender, downplaying the successfully-attacked scenarios instead. Indeed, this underscores the necessity of a diametrically opposed framework, where the defender evaluates the optimistic view of the attacker as well. We characterize this risk behavior as *distributionally risk-receptiveness* (DRR), in contrast with the conventional notion of distributionally risk-aversion [24]. Despite this juxtaposition, we adhere to the community convention by using the term *distributional robustness (or robust)* (DRO) interchangeably with distributional risk-aversion.

The DRR framework is also applicable when given samples may be contaminated. This optimistic optimization approach first rectifies the contamination and then makes decision. We note that concepts similar to DRR have also been explored to mitigate the impact of corrupted data in data-driven decision-making processes (namely, robust statistics [8, 9, 23] and Rockafellian relaxations [37]; refer to Sect. 2 for details). In the context of the Stackelberg games for disrupting domestic sex trafficking networks [25] and illegal drug supply chain networks [29] where attacker (or interdictor) is the protagonist, the DRR framework allows a risk-receptive interdictor to mitigate the impact of a network user (or defender) who can deceive the interdictor by providing false information about critical arcs of the network for maintaining the maximum flow in the network (refer to Sect. 8.1 for details).

In this paper, we study DRR and DRO models within the framework of multistage stochastic programming. A single-stage DRR model is formulated as a min-min problem. The outer problem seeks a decision that minimizes the expected cost for a best-case probability distribution, selected by the inner problem from a set of possible probability distributions, referred to as *ambiguity set*. Conversely, a single-stage DRO model is formulated as a min-max problem, where the inner problem selects a worst-case probability distribution from an ambiguity set. We extend these models to multistage stochastic integer programs (MSIPs), where a sequence of decisions is made over a time horizon. The objective of each stage is to minimize the sum of the current cost and the expected cost of future stages. We introduce DRR-MSIP and DRO-MSIP models to address distributional ambiguity in MSIPs. Furthermore, we study DRR-MSIPs with decision-dependent uncertainty as decisions often affect uncertainty in many real-world situations such as facility location [4, 42] and machine scheduling [33]. For example, in an agricultural supply chain, a farmer would sign up to transport the harvest to the storage centers, only if there is a nearby storage center that has been assigned to a processing mill with adequate demand to consume the farmer's harvest. This implies that decisions such as selection of storage centers and their assignment to the processing mill also impact the uncertain amount of harvest available at the storage centers and hence, the probability distribution of this uncertain parameter.

Despite the simplicity of the DRR models, they face significant computational challenges in terms of solution approaches. Directly solving a DRR model as a single-level minimization is impractical due to the integral defining the expectation, unless the ambiguity set is specifically chosen. With a finite support, the expectation is expressed as a summation, but this transforms the model without decision-dependent uncertainty into a mixed-integer non-convex bilinear program, which is notoriously difficult to solve. Even without any integrality restrictions and in two-stage setup, the objective function of a DRR model is nonconvex. In a multistage stochastic decision-making setup, the complexity increases further due to factors such as the number of scenarios and stages, which expand the model size. In contrast, some solution approaches for DRO models that utilize approximations are available in the literature under certain conditions on the cost functions and feasible sets [2, 3, 17, 18, 19, 36, 42]. It is worth noting that if the ambiguity sets are singleton, then DRO- and DRR-MSIPs reduce to a (risk-neutral) MSIP.

In this paper, we first address DRR- and DRO-MSIPs without decision-dependent uncertainty by presenting convex lower-approximations of the expected cost-to-go functions for best-case or worst-case distributions. We derive valid cutting planes and utilize reformulation techniques, considering finite or continuous supports. We note that while reformulations for DRR and DRO-MSIPs are derived using standard techniques, they are presented in this paper for comparative purposes as no other approaches for solving DRR-MSIPs are known in the literature. These approximations enable the application of decomposition methods, which effectively reduce complexity in a multistage decision-making setup. Specifically, we introduce customized SDDP (Stochas-

tic Dual Dynamic Programming) algorithms [35] for DRR- and DRO-MSIPs that utilize these approximations. We also generalize these algorithms to solve DRR and DRO multistage stochastic disjunctive linear programs (DRR- and DRO-MSDPs) where the feasible sets are defined by disjunctive linear sets.

DRO problems where the ambiguity sets are dependent on the decisions have been addressed by deriving their tractable reformulations [28,42] and then using state-of-the-art approaches for solving these reformulations. However, a solution approach for DRR (or optimistic) optimization with decision-dependent ambiguity set has not been studied in the literature, to the best of our knowledge. We fill this gap by introducing a reformulation-based approximation for decision-dependent DRR-MSIPs where the ambiguity sets are defined using Wasserstein distance. We also present lower-bound approximation for the associated decision-dependent optimistic “cost-to-go-function” and then utilize it to introduce a cutting plane-based algorithm for DRR-MSIPs with decision-dependent uncertainty.

To demonstrate practical applications of DRR- and DRO-MSIPs, we consider multistage network interdiction problems (NIPs). Specifically, we introduce multistage stochastic models for the maximum flow interdiction problem and the facility location interdiction problem, denoted by MS-MFIP and MS-FLIP, respectively. MS-MFIP has applications in disrupting domestic sex trafficking networks [25] and illegal drug supply chain networks [29], while MS-FLIP is applicable for identifying critical assets in infrastructure systems [13]. Despite these applications, their multistage stochastic variants have not been studied in the literature. To the best of our knowledge, only two-stage variants where each stage corresponds to the decision-making process of a single player have been explored in [14,22,38].

Through our DRO and DRR frameworks, we investigate the significance of varying risk-appetite of the decision-maker in MS-MFIP. Also, we examine an adversarial setting where data is corrupted, and demonstrate that the DRR framework can enhance the out-of-sample performance under such environments. Lastly, we provide numerical results for both MS-MFIP and MS-FLIP, showcasing the computational efficiency of the proposed approaches.

In the remainder of this section, we present formulations of general DRR- and DRO multistage stochastic programs (MSPs) along with the contributions and organization of this paper.

*Notation.* We use  $[d]$  to denote set  $\{1, \dots, d\}$  for any positive integer  $d$ .

### 1.1 Problem Formulation: DRO-MSP and DRR-MSP

We first present the Bellman equation of an (risk-neutral) MSP with a planning horizon of  $T$  stages and then generalize it to the cases with distributional ambiguity, i.e., DRO- and DRR-MSPs. There are two types of decision variables at each stage: *state* decision variables that influence subsequent stages and *local* decision variables specific to the current stage. Let  $x_t$  and  $y_t$  be

the state and local decision vector for stage  $t \in [T]$ , respectively. An MSP is formulated as

$$\min_{(x_1, y_1) \in X_1} \left\{ f_1(x_1, y_1) + \mathbb{E}_{P_2} [Q_2(x_1, \boldsymbol{\omega}_2)] \right\}, \quad (1)$$

where  $\boldsymbol{\omega}_t$  is a random vector that represents uncertain parameters for each stage  $t = 2, \dots, T$ , and the cost-to-go functions are defined as

$$Q_t(x_{t-1}, \boldsymbol{\omega}_t) = \min_{(x_t, y_t) \in X_t(x_{t-1}, \boldsymbol{\omega}_t)} \left\{ f_t(x_t, y_t, \boldsymbol{\omega}_t) + \mathbb{E}_{P_{t+1}} [Q_{t+1}(x_t, \boldsymbol{\omega}_{t+1})] \right\}, \quad (2)$$

for  $t = 2, \dots, T$ , and  $Q_{T+1} = 0$ . Set  $X_1$  denotes the feasible region of the first stage, and set  $X_t(x_{t-1}, \boldsymbol{\omega}_t)$  denotes the feasible region of each stage  $t \in \{2, \dots, T\}$  that depends on a decision of the previous stage  $x_{t-1}$ . The random vector  $\boldsymbol{\omega}_t$  has a sample space  $\Omega_t$ , which is associated with a probability distribution  $P_t$ .

Now, to present DRO- and DRR-MSPs, we consider a set  $\mathcal{P}_t$  of probability distributions as the ambiguity set for  $t \in \{2, \dots, T\}$ . Then, the bellman equation form of a DRO-MSP is given by

$$\min_{(x_1, y_1) \in X_1} \left\{ f_1(x_1, y_1) + \max_{P_2 \in \mathcal{P}_2} \mathbb{E}_{P_2} [Q_2^{RA}(x_1, \boldsymbol{\omega}_2)] \right\}, \quad (3)$$

where

$$Q_t^{RA}(x_{t-1}, \boldsymbol{\omega}_t) = \min_{(x_t, y_t) \in X_t(x_{t-1}, \boldsymbol{\omega}_t)} \left\{ f_t(x_t, y_t, \boldsymbol{\omega}_t) + \max_{P_{t+1} \in \mathcal{P}_{t+1}} \mathbb{E}_{P_{t+1}} [Q_{t+1}^{RA}(x_t, \boldsymbol{\omega}_{t+1})] \right\} \quad (4)$$

for  $t = 2, \dots, T$ , and  $Q_{T+1}^{RA} = 0$ . We refer to the value function

$$Q_{t+1}^{RA}(x_t) := \max_{P_{t+1} \in \mathcal{P}_{t+1}} \mathbb{E}_{P_{t+1}} [Q_{t+1}^{RA}(x_t, \boldsymbol{\omega}_{t+1})]$$

as the *pessimistic expected cost-to-go* function. By minimizing with respect to the probability distribution within the ambiguity set in (3) and (4), instead of maximizing, we can formulate a DRR-MSP as follows:

$$\min_{(x_1, y_1) \in X_1} \left\{ f_1(x_1, y_1) + \min_{P_2 \in \mathcal{P}_2} \mathbb{E}_{P_2} [Q_2^{RR}(x_1, \boldsymbol{\omega}_2)] \right\}, \quad (5)$$

where

$$Q_t^{RR}(x_{t-1}, \boldsymbol{\omega}_t) = \min_{(x_t, y_t) \in X_t(x_{t-1}, \boldsymbol{\omega}_t)} \left\{ f_t(x_t, y_t, \boldsymbol{\omega}_t) + \min_{P_{t+1} \in \mathcal{P}_{t+1}} \mathbb{E}_{P_{t+1}} [Q_{t+1}^{RR}(x_t, \boldsymbol{\omega}_{t+1})] \right\} \quad (6)$$

for  $t = 2, \dots, T$ , and  $Q_{T+1}^{RR} = 0$ . The value function

$$Q_{t+1}^{RR}(x_t) := \min_{P_{t+1} \in \mathcal{P}_{t+1}} \mathbb{E}_{P_{t+1}} [Q_{t+1}^{RR}(x_t, \boldsymbol{\omega}_{t+1})]$$

is referred to as the *optimistic expected cost-to-go* function. For DRR-MSPs with decision-dependent uncertainty, we consider ambiguity sets that depend on decisions, i.e., the optimistic expected cost-to-go functions are defined by

$$Q_{t+1}^{RR}(x_t) = \min_{P_{t+1} \in \mathcal{P}_{t+1}(x_t)} \mathbb{E}_{P_{t+1}} [Q_{t+1}^{RR}(x_t, \boldsymbol{\omega}_{t+1})].$$

When the feasible region  $X_t(x_{t-1}, \boldsymbol{\omega}_t)$  is defined by a set of linear inequalities and disjunctive constraints, problems (3)-(4) and (5)-(6) are referred to as DRO- and DRR-MSDPs, respectively, i.e.,

$$X_t(x_{t-1}, \boldsymbol{\omega}_t) := \left\{ (x_t, y_t) \in \mathbb{R}_+^{d_x} \times \mathbb{R}_+^{d_y} : \bigvee_{h \in H_t} \left( A_t^h(\boldsymbol{\omega}_t)x_t + B_t^h(\boldsymbol{\omega}_t)y_t \geq b_t^h(\boldsymbol{\omega}_t) - C_t^h(\boldsymbol{\omega}_t)x_{t-1} \right) \right\}. \quad (7)$$

Here, notation  $\vee$  is used to denote disjunction (“or” logical operator). The disjunctive constraints generalize integrality constraints on variables. For example, a binary restriction on a variable, i.e.,  $x \in \{0, 1\}$ , is equivalent to a disjunction:  $(x = 0) \vee (x = 1)$ . If the disjunctive constraints in (7) represent linear inequalities with integrality constraints on variables, then the DRO- and DRR-MSDPs reduce to DRO- and DRR-MSIPs. The main challenges encountered in solving these problems arises from the nonconvexity of feasible regions, caused by logical disjunctions or integer variables, as well as the nonlinearity and discontinuity of objective functions. Additionally, in the DRR problems, each stage problem’s objective function is nonconvex even if all variables are continuous.

## 1.2 Contributions and Organization of this Paper

In Sect. 2, we provide a review of previous studies on MSPs with and without distributional ambiguity. The organization of the rest of this paper and our contributions are as follows.

- *DRR-MSIPs with finite support.* In Sect. 3, we present a class of cutting planes to under-approximate the optimistic expected cost-to-go function  $Q_{t+1}^{RR}$  at stage  $t$ . For the sake of computational comparisons, we also reformulate DRR-MSIPs by considering the probability distribution at each stage as a decision vector and applying a McCormick envelope. We show that our SDDP-based algorithms using the cutting plane- and reformulation-based approximations are exact and finitely convergent (Sect. 6). Our computational results in Sect. 8 show that the algorithm using the cutting plane-based approach is 24.1 times faster (on average) than the algorithm using the reformulation approach.

- *DRR-MSIPs with continuous support and decision-dependent ambiguity sets.* In Sect. 4, we study DRR-MSIPs with Wasserstein ambiguity sets where the radii of the Wasserstein balls depend on the decision variables. We derive a dual formulation of the inner minimization at each stage (6), thereby deriving a class of valid cutting planes for the optimistic expected cost-to-go function and a cutting plane-based solution approach.
- *DRO-MSIPs.* In Sect. 5, we present a cutting plane-based approximation for DRO-MSIPs with a general family of ambiguity sets and finite supports. Also, we present a reformulation-based approximation using a dual formulation of DRO-MSIPs with Wasserstein ambiguity sets and continuous/finite support. We observe that the former is 28 times (on average) faster than the latter approach for solving DRO-MSIP with finite support.
- *Out-of-sample Performance.* By conducting out-of-sample tests, we demonstrate the significance of DRR and DRO in the context of interdiction problem, in particular, MS-MFIP. The results in Sect. 8.1.2 show that the DRO framework enhances the robustness of decision policies under uncertainty, and the DRR framework enables the identification of network vulnerabilities by giving more weights to unfavorable scenarios for the network user. Furthermore, in Sect. 8.1.3 we illustrate out-of-sample performance in an adversarial setting, where the sample data are intentionally corrupted (refer to robust statistics in Sect. 2 for details). The results show that the DRR policies behaves robust to data corruption by decreasing the significance of the corrupted data points.
- *DRR-MSDPs and DRO-MSDPs.* In Sect. 7, we introduce algorithms for solving MSDPs, DRR-MSDPs, and DRO-MSDPs by deriving tight extended formulations for parametric disjunctive constraints in each stage. Since MSIPs are special cases of MSDPs, we utilize the foregoing approach for solving MSIPs with(out) distributional ambiguity as well, where a hierarchy of relaxations of the feasible regions is obtained in each iteration.
- *MS-MFIP and MS-FLIP.* As mentioned in Sect. 1, MS-MFIP and MS-FLIP are important interdiction problems in their own right and have not been studied in the literature. We present algorithms to solve these problems and their distributionally ambiguous variants, thereby generalizing results of [13, 14, 22, 29, 25] that study special cases of these problems.

In Sect. 8, we present our computational results and concluding remarks in Sect. 9. For readers' convenience, we list major abbreviations used in this paper in Appendix A and provide all proofs in Appendix C. Throughout the paper, we made the following assumptions:

**Assumption 1** The random vectors are stage-wise independent, i.e.,  $\omega_t$  is independent of  $\omega_{[t-1]} = (\omega_2, \dots, \omega_{t-1})$  for all  $t = 3, \dots, T$ .

We note that the stage-wise independence assumption is required for computationally efficient algorithms, but not deriving convex approximations.

**Assumption 2** The state variables  $x_t$  are binary for all  $t \in [T]$ . The feasible sets  $X_1$  and  $X_t(\cdot, \cdot)$  are defined as mixed-integer sets (and disjunctive sets only in Sect. 7).

*Remark 1* For general integer or discrete state variables  $x_t$  that are bounded, we can obtain their equivalent binary representation by using the binary expansion.

**Assumption 3** Sets  $X_1$  and  $X_t(x_{t-1}, \omega_t)$ , given any  $x_{t-1} \in \{0, 1\}^{d_x}$  and  $\omega_t \in \Omega_t$ , for all  $t = 2, \dots, T$ , are nonempty and compact.

**Assumption 4** The supports  $\Omega_t$ , for all  $t = 2, \dots, T$ , are finite, i.e.,  $|\Omega_t| < \infty$ . Accordingly, let  $p_t^i$  be the probability of scenario  $\omega_t^i$ , i.e.,  $P_t(\omega_t = \omega_t^i)$ , for  $i \in \mathcal{N}_t$ , where  $\mathcal{N}_t := \{1, \dots, N_t\}$  is the index set associated with the support  $\Omega_t$ ,  $t = 2, \dots, T$ .

Throughout the paper, we assume finite supports unless stated otherwise. Specifically, we will relax this assumption and consider continuous supports in Sect. 4 and Remark 2 in Sect. 5.

**Assumption 5** Functions  $f_1(x_1, y_1)$  and  $f_t(x_t, y_t, \omega_t)$ , for  $\omega_t \in \Omega_t$  and  $t = 2, \dots, T$  are linear.

## 2 Literature Review

In this section, we examine studies related to the DRR framework. In addition, we review the literature on solution approaches for multistage stochastic linear programs (MSLPs), MSIPs, and their distributionally robust variants.

### 2.1 Distributionally Risk-Receptive Programs and Robust Statistics

In the literature, the DRR framework has been studied in three main contexts: (a) making decisions robust to outliers, errors, and adversarial corruptions in a data-driven setting [8, 37], (b) analyzing outcomes based on multiple decision-makers' differing perspective and varying risk-appetite of the decision-makers [24], and (c) obtaining the bounds on the true expectations [16, 11].

In the context of robust statistics, Blanchet et al. [8] discuss the connection between the DRR framework and robust statistics, which aim to seek a reliable estimator given samples that may be contaminated. They present the DRR framework as an optimistic optimization approach that first rectifies the contamination and then makes decision. In [9], as a rectifying-optimizing approach, a DRR program with an ambiguity set defined by an optimal transport distance has been studied. Royset et al. [37] also study a similar optimistic approach based on *Rockafellian relaxations*. Jiang and Xie [23] show that the DRR framework with a specific selection of ambiguity sets can recover many robust statistics, such as median and the least trimmed squares.



In another direction, DRR programs are utilized in reinforcement learning [41] to find an optimistic policy and Bayesian statistics [32] to approximate a likelihood. In [21], the authors investigate the DRO and DRR frameworks in comparison to a sample average approximation approach for out-of-sample performance. In particular, they show that by solving both DRO and DRR programs, one of their solutions always outperforms the sample average approximation solutions in out-of-sample test.

In the context of obtaining bounds, Duchi et al. [16] employ the DRR and DRO frameworks to construct a confidence interval for the true optimal objective of a stochastic optimization problem, which can be used to determine the size of the ambiguity sets for a given confidence interval size. Similarly, Cao and Gao [11] address problems involving covariate data where uncertain parameters belong to a specific uncertainty set. They show that solving robust and optimistic optimization problems yields worst-case and best-case rewards, respectively, thus forming a confidence interval for the true reward. Nakao et al. [31] consider a partially observable Markov decision process with distributional ambiguity. They solve a DRR model to obtain an upper bound on the true value function of a DRO partial observable Markov decision process.

To the best of our knowledge, the literature lacks a solution approach for DRR programs involving multistage decision-making with integer variables, which is applicable to interdiction problems (as discussed in Section 1).

## 2.2 Multistage Stochastic and Distributionally Robust Programs

*Multistage Stochastic Linear Programs.* To solve MSLPs, Nested Benders Decomposition (NBD) approximates the cost-to-go function  $Q_{t+1}$  at each stage  $t$  by a piecewise linear convex function using Benders cutting planes, which are constructed by the dual solutions from the problem at the subsequent stage [7]. For MSLPs under the stage-wise independence assumption, SDDP [35] is a NBD-like approach that harnesses scenario sampling to mitigate “curse of dimensionality” of dynamic programming without losing the (almost surely) finite convergence of the NBD algorithm. The iteration complexity of SDDP is also examined in [26].

*Distributionally Robust (or Risk-averse) MSLPs.* Recently, Philpott et al. [36] consider distributionally robust multistage stochastic linear programs (denoted by DRO-MSLPs) where ambiguity sets are constructed based on  $\chi^2$  distance from a reference probability distribution. Their approaches are based on SDDP embedding separation algorithms that compute a worst-case probability distribution for the different reference probability distributions. Note that their problems consider only continuous variables. Duque and Morton [18] present an SDDP-based algorithm for DRO-MSLPs where the ambiguity set is defined using Wasserstein metric. Their approach is based on the dualization of the inner problem for finding a worst-case probability distribution. They present comparison analysis of the results from their algorithm and those from the

modified algorithm of [36] for Wasserstein metric. Park and Bayraksan [34] investigate DRO-MSLPs, where the ambiguity set is defined using  $\phi$ -divergence, for an application to a water allocation problem. They propose a NBD-type algorithm relying on the dual reformulation of the inner problem for finding a worst-case probability distribution.

*Multistage Stochastic Integer Programs.* For solving MSIPs with binary state variables, an extension to SDDP, referred to as SDDiP, has been proposed by Zou et al. [44]. SDDiP uses a new class of cutting planes, constructed based on a Lagrangian relaxation where the strong duality holds for the resulting Lagrangian dual, to approximate the cost-to-go function in each stage. They provide the cut conditions under which SDDiP is finitely convergent. For a more general class of MSIPs, where all decision variables are allowed to be mixed-integer, there are several studies applying scenario-wise decomposition schemes to the deterministic equivalent formulation. For example, Carøe and Schultz [12] present a branch-and-bound algorithm based on a dual decomposition approach applied to the deterministic equivalent formulation. The approach uses a Lagrangian relaxation of *non-anticipativity* constraints which enforce the scenarios that follow the same history up to stage  $t$  to have the same decisions until stage  $t$ . Lulli and Sen [27] propose a branch-and-price algorithm, i.e., a branch-and-bound algorithm with column generation, for the same class of MSIPs. We also note that for multistage stochastic mixed-integer nonlinear programs, Zhang and Sun [43] present three decomposition algorithms—one based on NBD and another based on SDDP—which rely on a regularization of the expected cost-to-go function, i.e.,  $\mathbb{E}_{P_{t+1}}[Q_{t+1}(x_t, \omega_{t+1})]$  at stage  $t$ , and a class of cutting planes called *generalized conjugacy cut* to approximate the function.

*Distributionally Robust MSIPs.* Yu and Shen [42] investigate decision dependent DRO-MSIPs, where state variables are binary and the ambiguity set depends on the state decisions made at the previous stage. They consider three types of ambiguity sets constructed based on the decision-dependent moment information (e.g., mean and variance). They propose mixed-integer linear programming and mixed-integer semidefinite programming reformulations of the problems and solve them using SDDiP. Recently, in the dissertation of Nakao [30], a dual decomposition approach is presented for a DRO-MSIP where variables can be mixed-integer and the ambiguity sets are defined using Wasserstein metric. The approach applies the dual reformulations to the inner problems over Wasserstein ambiguity sets in a consecutive manner for deriving a monolithic-minimization deterministic equivalent formulation of the DRO-MSIP. Then, they use a Lagrangian relaxation of the non-anticipativity constraints in the deterministic equivalent formulation to derive a Lagrangian dual. Bayraksan et al. [6] divide the scenario tree into subgroups to obtain lower bounds and present conditions for selecting the radii of the ambiguity sets of the subgroup problems, thereby leading to another decomposition approach for DRO-MSIPs.

### 3 Convex Approximations for DRR-MSIPs having Finite Supports

In this section, we present two convex approximations of the optimistic expected cost-to-go function  $Q_{t+1}^{RR}(x_t)$  at stage  $t$ . We start by considering affine cuts valid for  $Q_{t+1}^{RR}$ :

$$Q_{t+1}^{RR}(x_t, \omega_{t+1}^i) \geq (\alpha_t^{i,k})^\top x_t + \beta_t^{i,k}, \quad \forall x_t \in \{0, 1\}^{d_x}, \quad (8)$$

and denote them by their coefficients  $(\alpha_t^{i,k}, \beta_t^{i,k})$ , where  $i \in \mathcal{N}_{t+1}$  indexes the support  $\Omega_{t+1}$ , and  $k \in [K_t]$ , with  $K_t \geq 1$  represents the number of cuts. Throughout this section, we suppose these cuts are given for every stage  $t$ . Discussions regarding the cut generating procedure and properties of these cuts are deferred to Sect. 6. Using the cuts, we define the following nonconvex (bilinear) approximating problem:

$$\phi_t^B(x_t) := \min_{\theta_t, P_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i \theta_t^i \quad (9a)$$

$$\text{s.t. } \theta_t^i \geq (\alpha_t^{i,k})^\top x_t + \beta_t^{i,k}, \quad \forall k \in [K_t], i \in \mathcal{N}_{t+1}, \quad (9b)$$

$$P_{t+1} \in \mathcal{P}_{t+1}. \quad (9c)$$

Note that we consider the probabilities  $P_{t+1} = (p_{t+1}^i)_{i \in \mathcal{N}_{t+1}}$  as decision variables. This leads to bilinear terms involving variables  $p_{t+1}^i$  and  $\theta_t^i$  in the objective function. By construction, the function  $\phi_t^B(x_t)$  provides a lower bound for the optimistic expected cost-to-go function  $Q_{t+1}^{RR}(x_t)$ , yet solving this problem directly is not desirable due to the bilinear terms. In the following sections, we derive convex approximations of this problem using a new class of cutting planes and a mixed-integer linear programming reformulation, respectively. For simplicity of exposition, we let  $K_t = 1$  for all  $t \in [T-1]$  and suppress the index  $k$  in notation.

#### 3.1 A cutting plane-based approximation for DRR-MSIP

Consider any  $\hat{x}_t \in \{0, 1\}^{d_x}$ . We define a cutting plane-based approximating function as:

$$\phi_t^C(x_t) := \min \left\{ \phi : \phi \geq \pi_t^\top (x_t - \hat{x}_t) + \gamma_t \right\} \quad (10)$$

where the parameters of the inequalities are given by

$$\pi_{t,j} := \begin{cases} \min_{P_{t+1} \in \mathcal{P}_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i \alpha_{t,j}^i, & \text{if } \hat{x}_{t,j} = 0, \\ \max_{P_{t+1} \in \mathcal{P}_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i \alpha_{t,j}^i, & \text{if } \hat{x}_{t,j} = 1, \end{cases} \quad \text{for } j \in [d_x], \quad (11)$$

and  $\gamma_t := \min_{P_{t+1} \in \mathcal{P}_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i \left( (\alpha_t^i)^\top \hat{x}_t + \beta_t^i \right)$ , for each  $t \in [T-1]$ . Apparently, this function is convex. Also, it provides a lower-bound for  $Q_{t+1}^{RR}$  as shown in Theorem 1.

**Theorem 1** *The function  $\phi_t^C$  provides a lower bound for the optimistic expected cost-to-go function, i.e.,  $\phi_t^C(x_t) \leq \mathcal{Q}_{t+1}^{RR}(x_t)$  for all  $x_t \in \{0, 1\}^{d_x}$  and  $t \in [T - 1]$ .*

An important property of this cut is that it preserves the tightness of the cuts  $(\alpha_t^i, \beta_t^i)$ , i.e., it intersects with  $\mathcal{Q}_{t+1}^{RR}$  at  $\hat{x}_t$  if the cuts  $(\alpha_t^i, \beta_t^i)$  intersect with  $\mathcal{Q}_{t+1}^{RR}(\cdot, \omega_{t+1}^i)$  at  $\hat{x}_t$  for all  $i \in \mathcal{N}_{t+1}$ . This property plays a key role in showing the finite convergence of our algorithm presented in Sect. 6. Additionally, it is important to note that the cuts (10) are different from Benders cuts, which rely on the duality results of problems. Since the DRR problem (6) is nonconvex even when involving continuous variables, convex combination of Benders cuts cannot be applied to approximate the optimistic expected cost-to-go function  $\mathcal{Q}_{t+1}^{RR}(x_t)$  as it can be done for risk-neutral and pessimistic expected cost-to-go-function.

### 3.2 A reformulation-based approximation for DRR-MSIP

We present a reformulation-based approximation for DRR-MSIP by treating the probability distribution as variables and using McCormick envelopes for bilinear terms. Though straightforward, it remains the only approach using existing techniques to solve this problem and therefore, we include it for comparative purpose. More specifically, it works as follows.

For each constraint in (9a), we multiply  $p_{t+1}^i$  to both sides of inequalities for each  $i \in \mathcal{N}_{t+1}$ , and replace  $p_{t+1}^i x_t$  with a decision vector  $\eta_t^i$  in the right-hand side of the resulting inequalities. This yields the following system of inequalities:

$$p_{t+1}^i \theta_t^i \geq (\alpha_t^i)^\top \eta_t^i + \beta_t^i p_{t+1}^i, \quad \forall i \in \mathcal{N}_{t+1}, \quad (12a)$$

$$\eta_t^i \leq x_t, \quad \eta_t^i \leq p_{t+1}^i, \quad \eta_t^i \geq p_{t+1}^i + x_t - 1, \quad \eta_t^i \geq 0, \quad \forall i \in \mathcal{N}_{t+1}. \quad (12b)$$

Notice that a system of inequalities (12b) ensure that a feasible  $\eta_t^i$  equals to  $p_{t+1}^i x_t$ , given any  $x_t \in \{0, 1\}^{d_x}$  and  $(p_{t+1}^i)_{i \in \mathcal{N}_{t+1}} \in [0, 1]^{N_{t+1}}$ . We introduce an additional variable  $\bar{\theta}_t^i$  to replace  $p_{t+1}^i \theta_t^i$ . This yields an approximating function of  $\mathcal{Q}_{t+1}^{RR}$  for each  $t \in [T - 1]$  as follows:

$$\phi_t^R(x_t) := \min_{P_{t+1}, \bar{\theta}, \eta_t} \left\{ \sum_{i \in \mathcal{N}_{t+1}} \bar{\theta}_i : \bar{\theta}_i \geq (\alpha_t^i)^\top \eta_t^i + \beta_t^i p_{t+1}^i, \quad \forall i \in \mathcal{N}_{t+1}, \quad (13a) \right.$$

$$\left. (12b), P_{t+1} \in \mathcal{P}_{t+1} \right\} \quad (13b)$$

Since  $\theta_t^i$  is not restricted, we can readily show that the equivalence between (12a) and constraint (13a). Therefore, the function  $\phi_t^R$  equals to the function  $\phi_t^B$ , which is a lower bound for  $\mathcal{Q}_{t+1}^{RR}(x_t)$ . When the ambiguity set  $\mathcal{P}_{t+1}$  is defined by a polytope (e.g., Wasserstein ambiguity set with a finite support), the problem (13) is a linear program. Consequently,  $\phi_t^R$  becomes piecewise linear and convex by linear programming duality.

#### 4 Convex Approximation for DRR-MSIPs having Continuous Supports and Decision-dependent Ambiguity Sets

In this section, we investigate DRR-MSIPs, where for each stage  $t$  the support  $\Omega_t$  is continuous. Suppose that a finite set of data  $\Omega_t := \{\omega_t^1, \dots, \omega_t^{N_t}\}$  is available. Consider an empirical distribution  $\bar{P}_t = \frac{1}{N_t} \sum_{i \in \mathcal{N}_t} \delta_{\omega_t^i}$ , where  $\delta_{\omega_t^i}$  is the Dirac delta function centered at  $\omega_t^i$  for  $i \in \mathcal{N}_t$ . We now present a dual-based convex approximation of the optimistic expected cost-to-go function  $Q_{t+1}^{RR}(x_t)$  at stage  $t$ . This approximation can be applied to a further generalized case where ambiguity sets are decision-dependent. In particular, we focus on decision-dependent ambiguity sets defined using Wasserstein metric as follows:

$$\mathcal{P}_t(x_{t-1}) := \left\{ P_t \in \mathcal{M}(\Omega_t) : \mathcal{W}(P_t, \bar{P}_t) \leq \epsilon_t(x_{t-1}) \right\} \quad (14)$$

for  $t = 2, \dots, T$  [28], where  $\mathcal{M}(\Omega_t)$  is a set of all probability distributions supported on  $\Omega_t$ , and  $\mathcal{W}(P_t, \bar{P}_t)$  is the 1-Wasserstein distance defined as

$$\mathcal{W}(P_t, \bar{P}_t) := \inf_{P \in \mathcal{P}(\Omega_t \times \Omega_t)} \left\{ \mathbb{E}_\pi \left[ \|\omega_t - \omega'_t\| \right] : P(\omega_t) = P_t, P(\omega'_t) = \bar{P}_t \right\}. \quad (15)$$

Here,  $\mathcal{P}(\Omega_t \times \Omega_t)$  is the set of all joint probability distributions supported on  $\Omega_t \times \Omega_t$ . The marginal distributions of  $\omega_t$  and  $\omega'_t$  are denoted by  $P(\omega_t)$  and  $P(\omega'_t)$ , respectively, and  $\|\cdot\|$  is an arbitrary norm. For a given  $\epsilon_t(x_{t-1}) > 0$ , each ambiguity set  $\mathcal{P}_t(x_{t-1})$ ,  $t \in \{2, \dots, T\}$ , is a Wasserstein ball containing all probability distributions within a certain radius  $\epsilon_t(x_{t-1})$  from the empirical probability distribution. For the ease of exposition in this section, we use  $x_t$  to represent all decision variables at each stage  $t$ .

**Proposition 1 (Strong duality)** *For the ambiguity set  $\mathcal{P}_t(x_{t-1})$  defined as (14), the optimistic expected cost-to-go function can be reformulated as*

$$\begin{aligned} Q_t^{RR}(x_{t-1}) &= \min_{P_t \in \mathcal{P}_t(x_{t-1})} \mathbb{E}[Q_t^{RR}(x_{t-1}, \omega_t)] \\ &= \max_{\rho_t \geq 0} \left\{ -\epsilon_t(x_{t-1})\rho_t \right. \\ &\quad \left. + \sum_{i \in \mathcal{N}_t} \frac{1}{N_t} \min_{\omega_t \in \Omega_t} \left\{ \rho_t \|\omega_t - \omega_t^i\| + Q_t^{RR}(x_{t-1}, \omega_t) \right\} \right\}. \end{aligned} \quad (16)$$

Using Proposition 1, we derive an approximation for  $Q_t^{RR}$  in Theorem 2 under the following assumptions.

**Assumption 6** The support  $\Omega_t$  for each stage  $t$  is a bounded hyperrectangle, i.e.,  $\Omega_t = [l_t, u_t]^{d_\omega}$ .

**Assumption 7** Random vector  $\omega \in \mathbb{R}^{d_\omega}$  is only associated with the right-hand side of the constraints defining  $X_t(x_{t-1}, \omega_t)$ . Let  $\mathcal{X}_t \subseteq \mathbb{R}^{d_x}$  be the set

defined by the integrality constraints for  $x_t$ . Then the feasible sets can be written as

$$X_t(x_{t-1}, \omega_t) = \left\{ x_t \in \mathcal{X}_t : A_t x_t \geq \omega_t - C_t x_{t-1} \right\}. \quad (17)$$

**Assumption 8** The decision-dependent radius  $\epsilon_t(x_{t-1}) : \mathcal{X}_{t-1} \rightarrow \mathbb{R}_+$  is an affine function.

**Theorem 2** Given valid cuts' coefficients  $\{(\pi_t^k, \gamma_t^k)\}_{k \in [K_t]}$  for  $\mathcal{Q}_{t+1}^{RR}(x_t)$ ,

(a) the optimistic expected cost-to-go function  $\mathcal{Q}_t^{RR}(x_{t-1})$  is lower-approximated as follows:

$$\begin{aligned} \mathcal{Q}_t^{RR}(x_{t-1}) \geq & -\epsilon_t(x_{t-1})\rho_t + \sum_{i \in \mathcal{N}_t} \frac{1}{N_t} \left( (-\lambda_t^i)^\top C_t x_{t-1} + (\mu_t^i)^\top l_t \right. \\ & \left. - (\nu_t^i)^\top u_t + (\lambda_t^i - \mu_t^i + \nu_t^i)^\top \omega_t^i + \sum_{k \in [K_t]} \zeta_{tk}^i \gamma_t^k \right), \end{aligned} \quad (18)$$

such that parameters  $(\rho_t, \{\lambda_t^i, \mu_t^i, \nu_t^i, \zeta_t^i\}_{i \in \mathcal{N}_t})$  satisfy the following system of equations and inequalities (19):

$$A_t^\top \lambda_t^i - \sum_{k \in [K_t]} \zeta_{tk}^i \pi_t^k = c_t, \quad \forall i \in \mathcal{N}_t, \quad (19a)$$

$$\|\lambda_t^i + \mu_t^i - \nu_t^i\|_* \leq \rho_t, \quad \forall i \in \mathcal{N}_t, \quad (19b)$$

$$\sum_{k \in [K_t]} \zeta_{tk}^i = 1, \quad \forall i \in \mathcal{N}_t \quad (19c)$$

$$\rho_t \geq 0, \quad (\lambda_t^i, \mu_t^i, \nu_t^i, \zeta_t^i) \geq 0, \quad \forall i \in \mathcal{N}_t, \quad (19d)$$

where  $\|\cdot\|_*$  is the dual norm of  $\|\cdot\|$ .

(b) the strongest lower bound approximation (18) is obtained by solving the following problem:

$$\begin{aligned} \max \left\{ & -\epsilon_t(x_{t-1})\rho_t + \frac{1}{N_t} \sum_{i \in \mathcal{N}_t} \left( (-\lambda_t^i)^\top C_t x_{t-1} + (\mu_t^i)^\top l_t \right. \right. \\ & \left. \left. - (\nu_t^i)^\top u_t + (\lambda_t^i - \mu_t^i + \nu_t^i)^\top \omega_t^i + \sum_{k \in [K_t]} \zeta_{tk}^i \gamma_t^k \right) : (19a)-(19d) \right\}. \end{aligned} \quad (20)$$

Problem (20) is a cut-generating problem, where any feasible solution yields a cut in the form of (18). With the choice of specific norms, problem (20) becomes computationally manageable. For instance, if  $\|\cdot\|$  represents the  $l_1$  norm, the dual norm is the  $l_\infty$  norm, allowing the linearization of constraints (19b) and transformation of problem (20) into a linear program. Also, if  $\|\cdot\|$  is the  $l_2$  norm, its dual norm is also the  $l_2$  norm, resulting in problem (20) becoming a second-order conic program.

Using the cuts generated as above, we derive a cutting plane-based approximating function for  $Q_t^{RR}$  as in the form of (10). It is important to note that the cuts for next stage ( $t + 1$ ) in Theorem 2 are naturally given through our algorithmic procedure (discussed in Sect. 6). Our algorithm produces cuts in a descending order of stages, from  $T$  down to 1, and thus there always exists cuts for stage ( $t + 1$ ) when generating a cut at stage  $t$ .

## 5 Convex Approximations for DRA-MSIPs

In this section, we present two convex approximations of the pessimistic expected cost-to-go functions  $Q_{t+1}^{RA}(x_t)$  for stage  $t$ . These approximations are constructed using cutting planes derived by a separation approach and a reformulation derived by utilizing strong duality, respectively.

### 5.1 A cutting plane-based approximation for DRA-MSIP

We assume that the supports are finite (Assumption 4), and the valid cuts defined by  $(\alpha_t^i, \beta_t^i)$  for  $Q_{t+1}^{RA}(\cdot, \omega_{t+1}^i)$  for  $t \in [T - 1]$  and  $i \in \mathcal{N}_{t+1}$  are available. Given cuts  $(\alpha_t^i, \beta_t^i)$ , for  $i \in \mathcal{N}_{t+1}$ , and solution  $\hat{x}_t$ , we identify a worst-case probability distribution by solving the following problem, referred to as *distribution separation problem* at stage  $t \in [T - 1]$ :

$$\max_{P_{t+1} \in \mathcal{P}_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i \left( (\alpha_t^i)^\top \hat{x}_t + \beta_t^i \right). \quad (21)$$

Let  $\hat{P}_{t+1} = (\hat{p}_{t+1}^1, \dots, \hat{p}_{t+1}^{N_{t+1}})$  be an optimal solution to the distribution separation problem (21). We define a cutting plane-based approximating function for  $t \in [T - 1]$  as

$$\phi_t^S(x_t) := \min \left\{ \phi : \phi \geq \pi_t^\top x_t + \gamma_t \right\}, \quad (22)$$

where  $\pi_t = \sum_{i \in \mathcal{N}_{t+1}} \hat{p}_{t+1}^i \alpha_t^i$ , and  $\gamma_t = \sum_{i \in \mathcal{N}_{t+1}} \hat{p}_{t+1}^i \beta_t^i$ . Clearly, the function  $\phi_t^S$  under-approximates  $Q_{t+1}^{RA}$ , since for solution  $\bar{x}_t \neq \hat{x}_t$  it holds that  $Q_{t+1}^{RA}(\bar{x}_t) \geq \max_{P_{t+1} \in \mathcal{P}_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i ((\alpha_t^i)^\top \bar{x}_t + \beta_t^i) \geq \phi_t^S(\bar{x}_t)$ . Note that there are various types of ambiguity sets for which we can solve problem (21) and compute the coefficients (11) using a finite-time algorithm; e.g., ambiguity sets constructed using Wasserstein metric [20], moment information [42], total variation distance [5] and  $\chi^2$  distance [36], where their supports are finite.

### 5.2 A reformulation-based approximation for DRA-MSIP

We present a reformulation-based approximation that relies on the dualization of the inner maximization problems in (3) and (4), i.e.,  $\max\{\mathbb{E}_{P_{t+1}}[Q_{t+1}^{RA}(\cdot, \cdot)]:$

$P_{t+1} \in \mathcal{P}_{t+1}$  for  $t \in [T-1]$ . We note that similar reformulation-based approaches are studied in recent papers for DRA-MSLPs with  $\phi$ -divergence-based ambiguity sets [34], DRA-MSLPs with Wasserstein ambiguity sets [18], and DRA-MSIPs with moment-based ambiguity sets [42]. In the following, we demonstrate the DRA-SDDP-R algorithm for DRA-MSIPs with Wasserstein ambiguity sets, but we note that this approach can be applied to DRA-MSIPs with a general family of ambiguity sets for which dual formulations of DRA models are available.

The Wasserstein ambiguity set with finite support is defined for  $t = 2, \dots, T$  as

$$\mathcal{P}_t = \left\{ P_t \in \mathbb{R}_+^{N_t} : \sum_{i \in \mathcal{N}_t} p_t^i = 1, \sum_{j \in \mathcal{N}_t} v_{ij} = p_t^i, i \in \mathcal{N}_t, \sum_{i \in \mathcal{N}_t} v_{ij} = \bar{p}_t^j, j \in \mathcal{N}_t, \sum_{i \neq j \in \mathcal{N}_t} \|\omega_t^i - \omega_t^j\| v_{ij} \leq \epsilon_t, v_{ij} \geq 0, \forall i, j \in \mathcal{N}_t \right\},$$

where  $\{\bar{p}_t^i\}_{i \in \mathcal{N}_t}$  is a reference probability distribution on  $\Omega_t$  for  $t \in \{2, \dots, T\}$ . The dual of the maximization in (3) and (4) for  $t \in [T-1]$  is given as follows:

$$\min_{\rho_t \geq 0} \left\{ \epsilon_{t+1} \rho_t + \sum_{i \in \mathcal{N}_{t+1}} \bar{p}_{t+1}^i \nu_t^i : \nu_t^i + \left\| \omega_{t+1}^i - \omega_{t+1}^j \right\| \rho_t \geq Q_{t+1}^{RA}(x_t, \omega_{t+1}^j), \forall i, j \in \mathcal{N}_{t+1} \right\}.$$

Here, the strong duality holds by Theorem 1 in Gao and Kleywegt [20]. Then, given cuts  $(\alpha_t^i, \beta_t^i), i \in \mathcal{N}_{t+1}$ , we define a reformulation-based approximating function for  $t \in [T-1]$  as

$$\phi_t^D(x_t) := \min \left\{ \epsilon_{t+1} \rho_t + \sum_{i \in \mathcal{N}_{t+1}} \bar{p}_{t+1}^i \nu_t^i : \rho_t \geq 0, \right. \quad (23a)$$

$$\left. \nu_t^i + \left\| \omega_{t+1}^i - \omega_{t+1}^j \right\| \rho_t \geq (\alpha_t^i)^\top x_t + \beta_t^i, \forall i, j \in \mathcal{N}_{t+1} \right\}. \quad (23b)$$

*Remark 2* The convex approximation described in this section can be applied to DRA-MSIPs having continuous supports by employing additional algorithmic techniques. More specifically, the dual of the inner maximization in (3) and (4) with continuous support is given by the following semi-infinite program:

$$\min_{\rho_t \geq 0} \left\{ \epsilon_{t+1} \rho_t + \sum_{i \in \mathcal{N}_{t+1}} \bar{p}_{t+1}^i \nu_t^i : \nu_t^i + \left\| \omega_{t+1}^i - \omega_{t+1} \right\| \rho_t \geq Q_{t+1}^{RA}(x_t, \omega_{t+1}), \forall i \in \mathcal{N}_{t+1}, \omega_{t+1} \in \Omega_{t+1} \right\}. \quad (24)$$

Given  $\omega_{t+1}^j \in \Omega_{t+1}$  and cuts  $(\alpha_t^j, \beta_t^j)$  valid for  $Q_{t+1}^{RA}(\cdot, \omega_{t+1}^j)$ , for some  $j$ , we obtain valid cutting planes in the form of (23b) for this semi-infinite program.



These constraints provide an outer approximation of the feasible region of the semi-infinite program. Consequently, the optimal value within the outer approximation is a lower bound on the value of (24). We iteratively improve the outer approximation through the following steps: (i) Solve the outer approximation of (24) with a finite number of cutting planes; (ii) Given its solution  $(\bar{x}_t, \bar{\rho}_t)$ , solve the separation problem  $\max_{\omega_{t+1} \in \Omega_{t+1}} \{Q_{t+1}^{RA}(\bar{x}_t, \omega_{t+1}) - \|\omega_{t+1}^i - \omega_{t+1}\| \bar{\rho}_t\}$ ; (iii) Add a cutting plane in the form of (23b) associated with a scenario identified by solving the separation problem to the outer approximation of (24); (iv) Repeat until the outer approximation is as tight as a predetermined level. These iterative steps refine the outer approximation, making it increasingly accurate in representing the feasible region of the semi-infinite program, and thereby solve the DRA-MSIP with continuous support.

## 6 SDDP-based Algorithms for DRR-MSP and DRO-MSP

### 6.1 Distributionally ambiguous SDDP

We present a customized SDDP algorithm, referred to as *distributionally ambiguous SDDP* (DA-SDDP), for both DRO-MSIPs and DRR-MSIPs, where the convex approximations discussed in Sects. 3 and 5 are utilized. The DA-SDDP approximates the pessimistic and optimistic expected cost-to-go functions  $Q_{t+1}^{RA}$  and  $Q_{t+1}^{RR}$  for each  $t \in [T-1]$ . Notably, DA-SDDP shares the key sampling ideas of SDDP for approximating the expectation functions, but the crux of this algorithm is in deriving convex approximations and iterative refinement methods for them. Recall that SDDP solves the linear programs and use dual solutions to derive a valid cut (or approximations) for MSLPs. However, this is not applicable for DRR programs.

---

#### Algorithm 1 Distributionally Ambiguous SDDP

---

- 1: **Initialize**  $l \leftarrow 1$ ;  $x_0 \leftarrow$  initial state;  $\omega_1 \leftarrow$  data at the first stage;  $\Omega_1 := \{\omega_1\}$ ;  $K_t^l \leftarrow 0$  for  $t \in [T-1]$ ;
  - 2: **while** (satisfying none of stopping conditions) **do**
  - 3:   Sample a scenario path  $\xi^t \in \Xi := \Omega_1 \times \dots \times \Omega_T$
  - 4:   **for**  $t \in [T]$  **do** ▷ Forward Step
  - 5:     Solve subproblem  $\mathbb{P}_t^l(x_{t-1}^l, \xi_t^l)$  and obtain  $(x_t^l, y_t^l)$  and  $\hat{Q}_t^l(x_{t-1}^l, \xi_t^l)$
  - 6:   **for**  $t = T, \dots, 2$  **do** ▷ Backward Step
  - 7:     **for**  $i \in \mathcal{N}_t$  **do**
  - 8:       Solve relaxation  $\tilde{\mathbb{P}}_t^l(x_{t-1}^l, \omega_t^i)$  and obtain cut  $(\alpha_{t-1}^{i,l}, \beta_{t-1}^{i,l})$
  - 9:       Refine approximating function  $\phi_{t-1}^l$  by using cuts  $(\alpha_{t-1}^{i,l}, \beta_{t-1}^{i,l}), i \in \mathcal{N}_t$
  - 10:      $K_{t-1}^l \leftarrow K_{t-1}^l + 1$
  - 11:   Solve subproblem  $\mathbb{P}_1^l(x_0^l, \omega_1)$  and obtain the bound  $LB$
  - 12:    $K_t^{l+1} \leftarrow K_t^l$  for  $t = 1, \dots, T-1$ ;  $l \leftarrow l + 1$
  - 13: **return** Subproblems  $\{\mathbb{P}_t^l\}_{t \in [T-1]}$ ,  $LB$
-

A pseudocode of DA-SDDP is given in Algorithm 1. The algorithm is initialized with a predetermined initial state of the model, denoted by  $x_0$ , the input data for the first stage, denoted by  $\omega_1$ , and the singleton set  $\Omega_1 := \{\omega_1\}$ , which are introduced to simplify the notation later. We also set the iteration counter  $l$  to 1, and the number of cuts  $K_t^l$  for  $t \in [T - 1]$  at iteration  $l$  to 0. At iteration  $l$ , the algorithm samples a scenario path  $\xi^l = (\xi_1^l, \dots, \xi_T^l)$  from  $\Xi := \Omega_1 \times \dots \times \Omega_T$  (Line 3). For the finite convergence of the algorithm, we assume that the scenario path sampling is conducted with replacement. Note that this can be readily extended to the sampling of multiple scenario paths per iteration. The remainder of the iteration consists of a forward step (Lines 4 and 5) and a backward step (Lines 6-11).

*Forward Step.* For each stage  $t \in [T]$ , DA-SDDP solves the following approximation of Problem (4) (or (6)), which we refer to as *subproblem* and denote by  $\mathbb{P}_t^l(x_{t-1}^l, \xi_t^l)$  (Line 5):

$$\hat{Q}_t^l(x_{t-1}^l, \xi_t^l) := \min_{(x_t, y_t) \in X_t(x_{t-1}^l, \xi_t^l)} \left\{ f_t(x_t, y_t, \xi_t^l) + \phi_t^l(x_t) \right\}, \quad t \in [T], \quad (25)$$

where  $\phi_T^l(\cdot) = 0$ , and  $x_{t-1}^l, t = 2, \dots, T$ , is an optimal stage- $(t-1)$  solution. To simplify notation, we let  $x_0^l := x_0$  and  $X_1(x_0^l, \xi_1^l) := X_1$ . Function  $\phi_t^l(x_t)$ , for each  $t \in [T - 1]$ , is a convex function that is constructed by  $K_t^l$  cuts, and it serves as an under-approximation of the pessimistic and optimistic expected cost-to-go functions— $Q_{t+1}^{RA}$  and  $Q_{t+1}^{RR}$ —while solving DRO-MSIP and DRR-MSIP, respectively. The approximating functions presented in Sect. 3 and 5 are viable substitutes for the function  $\phi_t^l$  when they are defined using the set of cuts  $\{(\alpha_t^{i,k}, \beta_t^{i,k})\}_{k \in [K_t^l]}$  available at iteration  $l$  and stage  $t$ .

*Backward Step.* For each  $t = T, \dots, 2$ , the algorithm solves relaxations of the subproblems, denoted by  $\hat{\mathbb{P}}_t^l(x_{t-1}^l, \omega_t^i)$ , and compute affine cuts  $(\alpha_{t-1}^{i,l}, \beta_{t-1}^{i,l})$  for  $i \in \mathcal{N}_t$  (Line 8) using the information obtained by solving the relaxations. These cuts provide a lower-bounding approximation of the value function  $\hat{Q}_t^l$  such that

$$\hat{Q}_t^l(x_{t-1}, \omega_t^i) \geq (\alpha_{t-1}^{i,l})^\top x_{t-1} + \beta_{t-1}^{i,l}, \quad \forall x_{t-1} \in \{0, 1\}^{d_x}, i \in \mathcal{N}_t, \quad (26a)$$

$$\hat{Q}_t^l(x_{t-1}^l, \omega_t^i) = (\alpha_{t-1}^{i,l})^\top x_{t-1}^l + \beta_{t-1}^{i,l}. \quad (26b)$$

It should be noted that the cut's validity is defined for  $\hat{Q}_t^l$ , yet this condition is sufficient to construct a lower-approximation for the optimistic or pessimistic expected cost-to-go function, since  $\hat{Q}_t^l$  is always lower-bounding the exact value function  $Q_t^{RR}$  or  $Q_t^{RA}$ , respectively. Using the cuts  $\{(\alpha_{t-1}^{i,l}, \beta_{t-1}^{i,l})\}_{i \in \mathcal{N}_t}$ , the algorithm tightens the approximating function  $\phi_{t-1}^l(x_{t-1})$  and improves the lower bound (Line 9). In Line 11, the algorithm computes the lower bound on the overall optimal objective value by solving the subproblem associated with the first stage.

The algorithm repeats these iterations with the forward and backward steps until one of predetermined stopping conditions is satisfied. These conditions can include a maximum number of iterations, a limit on elapsed time, or convergence of the lower bound.

There are various ways of generating a cut  $(\alpha_{t-1}^{i,l}, \beta_{t-1}^{i,l}), i \in \mathcal{N}_t$ , which is a supporting hyperplane of the epigraph of  $\hat{Q}_t^l(x_{t-1}, \omega_t^i)$ , intersecting at  $x_{t-1} = x_{t-1}^l$ , i.e., a cut satisfying both (26a) and (26b). For example, by solving the subproblem to optimality, we can obtain an integer optimality cut, given a lower bound  $L$  for the value function  $\hat{Q}_t^l$ , in the following form:

$$\hat{Q}_t^l(x_{t-1}, \omega_t^i) \geq (\hat{q}_t^l - L) \left( \sum_{i \in [d_x]} 2x_{t-1,i}^l x_{t-1,i} - x_{t-1,i} - x_{t-1,i}^l \right) + \hat{q}_t^l,$$

where  $\hat{q}_t^l := \hat{Q}_t^l(x_{t-1}^l, \omega_t^i)$ . As another example, consider a Benders cut obtained by solving a linear programming relaxation of the subproblem when solving a DRO-MSIP (or DRR-MSIP). This cut satisfies (26a), though not necessarily (26b). However, we can derive a mixed-binary linear programming reformulation of the subproblem by adding binary variables replacing integer variables, use the hierarchy of relaxations (discussed in Sect. 7.2) to solve the reformulation to optimality, and obtain a Benders cut that satisfies both (26a) and (26b).

## 6.2 Finite Convergence

Now, we show the finite convergence of DA-SDDP equipped with the convex approximations presented earlier. For DRR-MSIPs, we use DRR-SDDP-C and DRR-SDDP-R to denote variants of DA-SDDP with the cutting plane-based approximation  $\phi_t^C$  (10) and the reformulation-based approximation  $\phi_t^R$  (13), respectively. We define a *policy* by a collection of functions  $\{\bar{x}_t(\xi_{[t]}), \bar{y}_t(\xi_{[t]})\}_{t \in [T]}$ , where  $\xi_{[t]} = (\xi_1, \dots, \xi_t)$ , which serves as a decision rule given any scenario path  $(\xi_1, \dots, \xi_T)$ . A policy is *optimal* for a DRR-MSIP if  $(\bar{x}_t(\xi_{[t]}), \bar{y}_t(\xi_{[t]}))$  is optimal to the  $t$ -th stage problem (6) ((5) for  $t = 1$ ) for  $t \in [T]$  and all  $\xi \in \Xi$ .

**Theorem 3** *The forward step of the DRR-SDDP-C algorithm defines an optimal policy for a DRR-MSIP in a finite number of iterations of its while loop with probability one. Furthermore, each iteration of the while loop is executed in a finite time if there exists a finite-time algorithm for computing the cut coefficients (11).*

*Remark 3* When the ambiguity sets are singletons, the finite convergence result presented in Theorem 3 applies to MSIPs, where the finite convergence result previously established by Zou et al. [44]. While their proof utilizes assumptions regarding the validity, tightness, and finiteness of cuts, the proof in this paper with binary state variables demonstrates that the convergence holds even without the finiteness assumption on cuts.

**Theorem 4** *The forward step of the DRR-SDDP-R algorithm defines an optimal policy to a DRR-MSIP in a finite number of iterations of its while loop with probability one. Furthermore, each iteration of the while loop is executed in a finite time if the ambiguity set at every stage is defined by a polytope or a mixed-binary linear set.*

Similarly, for DRO-MSIPs, we refer to the variants of DA-SDDP as DRO-SDDP-C and DRO-SDDP-R, utilizing the cutting plane-based approximation  $\phi_t^S$  (22) and the reformulation-based approximation  $\phi_t^D$  (23), respectively.

**Theorem 5** *The forward step of the DRO-SDDP-C algorithm provides an optimal policy for DRO-MSIP in a finite number of iterations of its while loop with probability one. Furthermore, each iteration of the while loop is executed in a finite time if there exists a finite-time algorithm for solving the distribution separation problem (21).*

## 7 Extensions to Multistage Stochastic Disjunctive Programs with Distributional Ambiguity

In this section, we present extensions of the DA-SDDP algorithms to DRR- and DRO-MSDPs defined in Section 1.1, under the following assumption: For  $h \in H_t$ , set  $\{(x_t, y_t) \in \mathbb{R}_+^{d_x} \times \mathbb{R}_+^{d_y} : A_t^h(\omega_t)x_t + B_t^h(\omega_t)y_t \geq b_t^h(\omega_t) - C_t^h(\omega_t)x_{t-1}\}$  is nonempty and compact for any  $x_{t-1} \in \{0, 1\}^{d_x}$  and  $\omega_t \in \Omega_t$ . Also, its constraints include  $x_t \geq 0$  and  $x_t \leq 1$ .

### 7.1 DA-SDDP algorithms for DRR- and DRO-MSDPs

Let us consider a set of cuts,  $\{(\pi_t^k, \gamma_t^k)\}_{k \in [K_t^l]}$ , for iteration  $l$  and stage  $t$ , constructing an approximation of the pessimistic or optimistic expected cost-to-go function. Then, with the same definition of  $x_0$  and  $\omega_1$  as in Algorithm 1, the subproblem at iteration  $l$  is given by

$$\hat{Q}_t^l(x_{t-1}, \omega_t) = \min_{(x_t, y_t) \in X_t(x_{t-1}, \omega_t)} \left\{ f_t(x_t, y_t, \omega_t) + \phi_t : \right. \\ \left. \phi_t \geq (\pi_t^k)^\top x_t + \gamma_t^k, \quad k \in [K_t^l] \right\} \quad (27)$$

for  $t \in [T]$  and  $\omega_t \in \Omega_t$ , where  $\phi_T = K_T^l = 0$  and  $X_t(x_{t-1}, \omega_t)$  is defined by a disjunctive set (7). We define the feasible set of the foregoing subproblem by

$$\mathcal{D}_t^l(x_{t-1}, \omega_t) := \left\{ (x_t, y_t, \phi_t) \in \mathbb{R}_+^{d_x} \times \mathbb{R}_+^{d_y} \times \mathbb{R}_+ : \right. \\ \left. \bigvee_{h \in H_t} \left( \phi_t - (\pi_t^k)^\top x_t \geq \gamma_t^k, \quad k \in [K_t^l], \right. \right. \\ \left. \left. A_t^h(\omega_t)x_t + B_t^h(\omega_t)y_t \geq b_t^h(\omega_t) - C_t^h(\omega_t)x_{t-1} \right) \right\}, \quad (28)$$

and derive the convex hull of  $\mathcal{D}_t^l(x_{t-1}, \omega_t)$  in the following proposition.

**Proposition 2** For any  $x_{t-1} \in \{0, 1\}^{d_x}$  and  $\omega_t \in \Omega_t$ , the convex hull of the set  $\mathcal{D}_t^l(x_{t-1}, \omega_t)$ ,  $t \in [T]$ , is equivalent to the projection of polyhedral set  $\tilde{\mathcal{D}}_t^l(x_{t-1}, \omega_t)$  onto the  $(x_t, y_t, \phi_t)$ -space where  $\tilde{\mathcal{D}}_t^l(x_{t-1}, \omega_t)$  is given by

$$\left\{ \begin{array}{l} \sum_{h \in H_t} \zeta_{t,0}^h = 1, \sum_{h \in H_t} \zeta_{t,1}^h - x_t = 0, \\ \sum_{h \in H_t} \zeta_{t,2}^h - y_t = 0, \sum_{h \in H_t} \zeta_{t,3}^h = x_{t-1}, \sum_{h \in H_t} \zeta_{t,4}^h - \phi_t = 0, \end{array} \right.$$

$$\begin{aligned} A_t^h(\omega_t)\zeta_{t,1}^h + B_t^h(\omega_t)\zeta_{t,2}^h + C_t^h(\omega_t)\zeta_{t,3}^h - b_t^h(\omega_t)\zeta_{t,0}^h &\geq 0, \quad h \in H_t, \\ \zeta_{t,4}^h - (\pi_t^k)^\top \zeta_{t,1}^h - \gamma_t^k \zeta_{t,0}^h &\geq 0, \quad h \in H_t, k \in K_t^l, \\ x_t \in \mathbb{R}_+^{d_x}, y_t \in \mathbb{R}_+^{d_y}, \phi_t \in \mathbb{R}_+, \\ \zeta_{t,0}^h \in \mathbb{R}_+, \zeta_{t,1}^h \in \mathbb{R}_+^{d_x}, \zeta_{t,2}^h \in \mathbb{R}_+^{d_y}, \zeta_{t,3}^h \in \mathbb{R}_+^{d_x}, \zeta_{t,4}^h \in \mathbb{R}_+, \quad h \in H_t \end{aligned} \left. \vphantom{\begin{aligned}} \right\}.$$

Using Proposition 2, we also derive an extension of DA-SDDP for DRR- and DRO-MSDPs, namely DA-SDDP-DP. Its pseudocode is provided in Appendix D. DA-SDDP-DP shares a similar structure to DA-SDDP, but note that it involves distinct subproblems and a special subroutine for adding cuts to the subproblems. In particular, it solves the linear programming equivalents of subproblems (27), derived using Proposition 2 and referred to as *LP-subproblems*:

$$\min \left\{ f_t(x_t, y_t, \omega_t) + \phi_t : (x_t, y_t, \phi_t) \in \text{Proj}_{x_t, y_t, \phi_t}(\tilde{\mathcal{D}}_t^l(x_{t-1}, \omega_t)) \right\}, \quad (29)$$

for  $t \in [T]$  and  $\phi_T = 0$ . We note that the DA-SDDP-DP algorithms for DRR- and DRO-MSDPs have the finite convergence if cuts  $(\pi_{t-1}^l, \gamma_{t-1}^l)$  are obtained as in the cutting plane-based algorithms for DRR- and DRO-MSIPs, respectively. The comprehensive description of the algorithm is provided in Appendix D.

## 7.2 Application of Proposition 2 for solving DRR- and DRO-MSIPs using Hierarchical Relaxations

In this section, we present a hierarchy of relaxations ranging from linear relaxation to tight extended formulations for each stage to solve DRR- and DRO-MSIPs by applying Proposition 2. For the ease of exposition, let  $y_t$  be continuous. Then, the subproblem (25) of DRR- and DRO-MSIPs for iteration

$l$  and stage  $t$  can be rewritten as

$$\min \left\{ f_t(x_t, y_t, \xi_t^l) + \phi_t^l : \left( x_{t,j} = 0 \vee x_{t,j} = 1, j = 1, \dots, d_x \right) \right. \\ \left. \bigwedge \left( \phi_t - (\pi_t^k)^\top x_t \geq \gamma_t^k, k \in [K_t^l], \right. \right. \\ \left. \left. A_t(\xi_t^l)x_t + B_t(\xi_t^l)y_t \geq b_t(\xi_t^l) - C_t(\xi_t^l)x_{t-1} \right) \right\}. \quad (30)$$

We use  $\mathcal{D}_t^l(x_{t-1}, \xi_t^l)$  and  $\mathcal{D}_t^{l,LP}(x_{t-1}, \xi_t^l)$  to denote the feasible region of the sub-problem (30) and its linear programming relaxation, respectively. A relaxation of  $\mathcal{D}_t^l(x_{t-1}, \xi_t^l)$  can be defined as follows:  $\mathcal{D}_t^{l,s}(x_{t-1}, \xi_t^l) := \mathcal{D}_t^{l,LP}(x_{t-1}, \xi_t^l) \cap \{x_{t,j} = 0 \vee x_{t,j} = 1, j \in [s]\}$ , for  $s \in [d_x]$ . It is easy to see that the set  $\mathcal{D}_t^{l,s}(\cdot, \cdot)$  for  $s = d_x$  is equivalent to the original set  $\mathcal{D}_t^l(\cdot, \cdot)$ . Moreover,  $\mathcal{D}_t^{l,LP}(\cdot, \cdot) \supseteq \mathcal{D}_t^{l,1}(\cdot, \cdot) \supseteq \dots \supseteq \mathcal{D}_t^{l,d_x}(\cdot, \cdot) = \mathcal{D}_t^l(\cdot, \cdot)$ , and thus  $\text{conv}(\mathcal{D}_t^{l,1}(\cdot, \cdot)) \supseteq \dots \supseteq \text{conv}(\mathcal{D}_t^{l,d_x}(\cdot, \cdot)) = \text{conv}(\mathcal{D}_t^l(\cdot, \cdot))$ . This provides a hierarchy of relaxations of the feasible region  $\mathcal{D}_t^l(\cdot, \cdot)$  of DRR- and DRO-MSIPs. The tight extended formulation of the convex hull of the relaxations can be obtained using Proposition 2.

**Proposition 3** *The convex hull of the set  $\mathcal{D}_t^{l,s}(x_{t-1}, \omega_t)$  is the projection of the following set onto the  $(x_t, y_t, \phi_t)$ -space for any  $x_{t-1} \in \{0, 1\}^{d_x}$  and  $\omega_t \in \Omega_t$ :*

$$\left\{ \begin{array}{l} \sum_{h \in [|\mathcal{J}_t^s|]} \zeta_{t,0}^h = 1, \quad \sum_{h \in [|\mathcal{J}_t^s|]} \zeta_{t,1}^h - x_t = 0, \\ \sum_{h \in [|\mathcal{J}_t^s|]} \zeta_{t,2}^h - y_t = 0, \quad \sum_{h \in [|\mathcal{J}_t^s|]} \zeta_{t,3}^h = x_{t-1}, \quad \sum_{h \in [|\mathcal{J}_t^s|]} \zeta_{t,4}^h - \phi_t = 0, \\ A_t(\omega_t)\zeta_{t,1}^h + B_t(\omega_t)\zeta_{t,2}^h + C_t(\omega_t)\zeta_{t,3}^h \geq b_t(\omega_t), \quad h \in [|\mathcal{J}_t^s|], \\ \zeta_{t,1,j}^h = 0, \quad j \in J_1^h, \quad h \in [|\mathcal{J}_t^s|], \\ \zeta_{t,1,j}^h = \zeta_{t,0}^h, \quad j \in J_2^h, \quad h \in [|\mathcal{J}_t^s|], \\ \zeta_{t,4}^h - (\pi_t^k)^\top \zeta_{t,1}^h - \gamma_t^k \zeta_{t,0}^h \geq 0, \quad h \in [|\mathcal{J}_t^s|], k \in [K_t^l], \\ x_t \in \mathbb{R}_+^{d_x}, y_t \in \mathbb{R}_+^{d_y}, \phi_t \in \mathbb{R}_+, \\ \zeta_{t,0}^h \in \mathbb{R}_+, \zeta_{t,1}^h \in \mathbb{R}_+^{d_x}, \zeta_{t,2}^h \in \mathbb{R}_+^{d_y}, \zeta_{t,3}^h \in \mathbb{R}_+^{d_x}, \zeta_{t,4}^h \in \mathbb{R}_+, \quad h \in [|\mathcal{J}_t^s|] \end{array} \right\}, \quad (31)$$

where  $\mathcal{J}_t^s := \{(J_1^h, J_2^h) : h \in [|\mathcal{J}_t^s|]\}$ ,  $s \in [d_x], t \in [T]$ , be a set of all pairs of disjoint sets  $(J_1, J_2)$  such that  $J_1, J_2 \subseteq [d_x], J_1 \cap J_2 = \emptyset$ , and  $|J_1 \cup J_2| = s$ .

Proposition 3 provides the following relaxation that can be used to generate cuts in Line 8 of Algorithm 1:  $\tilde{Q}_t^{l,s}(x_{t-1}^l, \xi_t^l) = \min \left\{ f_t(x_t, y_t, \xi_t^l) + \right.$

$\phi_t^l(x_t) : (x_t, y_t) \in \text{Proj}_{x_t, y_t}(\mathcal{D}_t^{l,s}(x_{t-1}^l, \xi_t^l))\}$ , where  $\tilde{Q}_t^{l,1}(x_{t-1}^l, \xi_t^l) \leq \dots \leq \hat{Q}_t^{l,d_x}(x_{t-1}^l, \xi_t^l) = \hat{Q}_t^l(x_{t-1}^l, \xi_t^l)$ . For  $s = d_x$ , a Benders cut obtained by solving this relaxation is a supporting hyperplane satisfying (26b). For the smaller values of  $s$ , the Benders cut does not necessarily support the value function  $\hat{Q}_t^l(\cdot)$ , but the relaxations are computationally easier to solve. The value of  $s$  can be adjusted either before or during the execution of the algorithm to address this trade-off between the computational effort and the effectiveness of the cuts.

*Remark 4* The above relaxations can be readily extended to the case where  $y_t$  are mixed-binary variables. Furthermore, if  $y_t$  are mixed-integer, a hierarchy of relaxations can be derived by employing the binary expansion as discussed in Remark 1.

## 8 Computational Tests

In this section, we present computational results for the DRR-SDDP-C, DRR-SDDP-R, DRO-SDDP-C, and DRO-SDDP-R algorithms. Recall that these are specific implementations of DA-SDDP, utilizing the approximations (10), (13), (22), and (23), respectively. We apply these algorithms to solve instances of multistage maximum flow and facility location interdiction problems with distributional ambiguity. All the algorithms are implemented in Julia 1.8 where subproblems, coefficient-computing problems (11), and distribution separation problem (21) are solved using Gurobi 9.5 with an optimality tolerance of  $10^{-4}$ . We also integrate our implementation of the inner functionalities of our DA-SDDP algorithms with SDDP.jl [15] package to be consistent with the research community, e.g., [36], thereby making it convenient for future computational and applied users of these algorithms. We conducted all tests on a machine equipped with an Intel Core i7 processor (3.8 GHz), utilizing a single thread, and 32 GB RAM.

In our implementations, all the algorithms generate a strengthened Benders cut and an integer optimality cut alternately during the backward steps. For details on the strengthened Benders cuts, we refer readers to Zou et al. [44]. For ambiguity set, we consider Wasserstein ambiguity set with the  $l_1$  norm throughout all test instances. Note that in this case the distribution separation problem (21) is a linear program and the subproblem (25) in the DRR-SDDP-R algorithm is a mixed-binary linear program.

### 8.1 MS-MFIP with Distributional Ambiguity

Throughout this section, we consider DRR and DRO variants of the MS-MFIP formulation that is provided in Appendix B. In the subsequent sections, we present the results of a comparative analysis of the algorithms and demonstrate the significance of DRR and DRO for MS-MFIP. In addition, we extend the

analysis to the case involving data corruption where the defender can deceive the interdictor by providing false information about critical arcs in a network.

### 8.1.1 Instance generation and computational results

Networks are randomly generated, following the method presented in Cormican et al. [14]. First, we place all nodes, excluding the source and sink nodes, in a grid pattern. Next, we establish connections between the leftmost and rightmost nodes in the grid to the source and sink nodes, respectively, using non-interdictable arcs with infinite capacity. Then, every pair of adjacent nodes in the grid is connected by an arc. Horizontal arcs are oriented from left to right, and vertical arcs, connecting the leftmost or rightmost nodes, are oriented from up to down. The orientations of the remaining arcs are randomly chosen. To avoid trivial solutions, e.g., removing all horizontal arcs in the same column, we set 80 percent of all arcs to be interdictable. Following the above procedure, we generate two distinct networks with different sizes. For the first network and the second network, we sample realizations of the random capacity of each arc uniformly distributed on  $[30, 60]$  and  $[20, 90]$ , respectively, to construct the support  $\Omega_t$  of size  $|\Omega|$  for each stage  $t$ . We set the Wasserstein ball size parameter  $\epsilon$  to 30. The interdiction budget for each stage is set to one to avoid a trivial solution where an interdiction solution in an early stage completely separates the source and sink nodes, leaving no arcs to remove in later stages.

Table 1 summarizes the details of the test instances. Each row corresponds to a single instance that is labeled accordingly under the *Instance* column. The naming convention NI- $i$ - $T$ - $|\Omega|$  is used for an instance with the  $i$ th network among the two aforementioned networks,  $T$  stages, and  $|\Omega|$  realizations per stage. The labels  $|N| \times |A|$  and #Scenario denote the number of nodes and arcs of the network and the total number of scenario paths, respectively. For termination conditions, we specify a time limit of 3 hours for all algorithms. Also, there is an early-termination condition based on the convergence of lower bound. In particular, the algorithm is stopped if the lower bound fails to improve for 100 consecutive iterations.

Table 1: Details of DRR- and DRA-MFIP instances

Instance	$ N  \times  A $	$T$	$ \Omega $	#Scenario
NI-1-3-5	37 x 73	3	5	25
NI-1-3-10			10	100
NI-1-3-15			15	225
NI-1-4-5		4	5	125
NI-1-4-10			10	1000
NI-1-4-15			15	3375
NI-1-5-5		5	5	625
NI-1-5-10			10	10000
NI-1-5-15			15	50625
NI-1-6-5		6	5	3125
NI-1-6-10			10	100000
NI-1-6-15			15	759375
NI-2-3-5	52 x 106	3	5	25
NI-2-3-10			10	100
NI-2-3-15			15	225
NI-2-4-5		4	5	125
NI-2-4-10			10	1000
NI-2-4-15			15	3375



Table 2: Performance comparison of algorithms for DRR- and DRA-MFIP instances

Instance	DRA-SDDP-C		DRA-SDDP-R		DRR-SDDP-C		DRR-SDDP-R	
	LBound	Time (s)	LBound	Time (s)	LBound	Time (s)	LBound	Time (s)
NI-1-3-5	537.08	39.1	537.08	224.1	528.10	41.1	528.10	585.0
NI-1-3-10	535.00	50.3	535.00	1653.2	527.05	60.3	527.05	2172.9
NI-1-3-15	534.76	105.0	534.76	10800+	528.11	103.1	528.11	6774.8
NI-1-4-5	655.58	169.1	655.58	4794.1	644.76	173.8	644.76	10800+
NI-1-4-10	670.44	547.2	662.12	10800+	655.45	453.2	600.68	10800+
NI-1-4-15	657.02	401.6	641.21	10800+	643.02	525.3	590.32	10800+
NI-1-5-5	735.84	1019.2	729.85	10800+	727.05	1455.2	645.57	10800+
NI-1-5-10	718.36	1443.9	676.38	10800+	701.54	1958.1	583.27	10800+
NI-1-5-15	792.14	704.2	774.19	10800+	769.31	746.9	650.32	10800+
NI-1-6-5	737.51	3058.8	705.42	10800+	725.69	3491.4	631.55	10800+
NI-1-6-10	762.90	5066.5	672.96	10800+	744.84	6853.9	593.40	10800+
NI-1-6-15	748.97	10800+	594.48	10800+	715.37	10800+	576.73	10800+
NI-2-3-5	1018.51	54.2	1018.51	115.3	1013.16	61.7	1013.16	906.5
NI-2-3-10	892.61	73.9	892.61	994.0	884.66	96.8	884.66	3543.9
NI-2-3-15	903.28	148.8	903.28	263.7	891.10	171.7	891.10	10244.7
NI-2-4-5	1181.45	410.0	1181.45	9978.1	1173.41	436.3	1136.69	10800+
NI-2-4-10	1096.20	1426.5	1062.48	10800+	1086.77	1574.4	981.57	10800+
NI-2-4-15	1105.07	3387.2	1027.92	10800+	1089.51	3553.0	824.64	10800+

In Table 2, we report lower bounds and solution times in seconds for each algorithm, labeled LBound and Time (s), respectively. We may also obtain upper bounds computed statistically through a sampling approach using the policy after termination. However, to focus on the computational performance of the algorithms in terms of convergence while running, we will only report lower bounds and solution times here. The results indicate that the DRO-SDDP-C algorithm provides better lower bounds and solution times than the DRO-SDDP-R algorithm for all instances. On average, the DRO-SDDP-C algorithm converges 17.2 times faster than the DRO-SDDP-R algorithm. This performance advantage increases to 26.4 times (on average) for eight instances where both algorithms provide the same lower bounds. The results show that the DRO-SDDP-R algorithm’s performance is more susceptible to  $T$  and  $|\Omega|$  than the DRO-SDDP-C algorithm. For example, as  $|\Omega|$  increases from 5 to 15 for NI-1-3-5, the DRO-SDDP-R algorithm’s solution time increases by 48.2 times, while the DRO-SDDP-C algorithm’s solution time increases by 2.7 times. Similarly, when  $T$  increases to 5 and 6 from 3 for NI-1-3-5, the DRO-SDDP-R algorithm fails to solve the 5-stage and 6-stage instances—NI-1-5-5 and NI-1-6-5—within the time limit, whereas the DRO-SDDP-C algorithm solves all instances within the time limit. This is mainly because the DRO-SDDP-R algorithm adds a significantly larger number of cuts ( $|\Omega|^2$  cuts for every subproblem solved) for each iteration, which increases the solution times for subproblems. Regarding DRR-MFIP, the DRR-SDDP-C algorithm outperforms the DRR-SDDP-R algorithm for all test instances. In terms of solution time, the DRR-SDDP-C algorithm converges, on average, 22.3 times faster

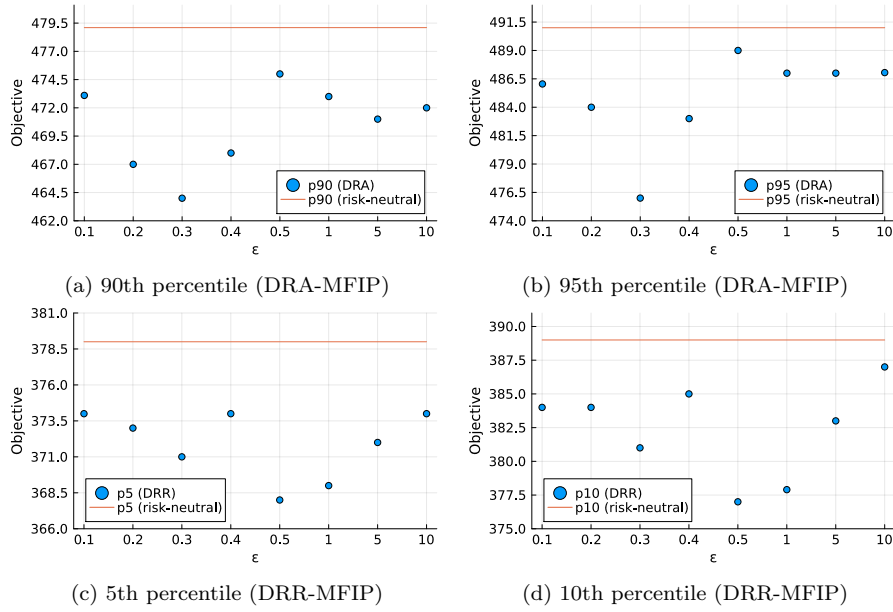


Fig. 1: Plots of percentiles from out-of-sample tests on DRA-MFIP (90th and 95th percentiles) and DRR-MFIP (5th and 10th percentiles)

than the DRR-SDDP-R algorithm, and this advantage increases to 41.3 times for seven instances where both the algorithms provide the same bounds. This is because the DRR-SDDP-R algorithm solves larger subproblems that arise from (13b), incorporated by the linearization and the ambiguity set, respectively.

### 8.1.2 Impact of DRO and DRR on MS-MFIP

To demonstrate the impact of DRO and DRR on MS-MFIP, we present the results from out-of-sample tests, which are conducted as follows. We first sample realizations of the capacity over the stages. Then, the DRO-MFIP and DRR-MFIP instances obtained over this sample ( $\{\Omega_t : t \in [T]\}$ ) are solved using the algorithms. The resulting subproblems along with the cuts generated by the algorithms define a policy for MS-MFIP, i.e., the decision rule that selects the set of arcs to remove given a realization of the capacity. We simulate the DRR and DRO policies using scenario paths of the capacity that are sampled independently from the realizations used in solving the problems.

Throughout the out-of-sample tests in this section, we consider the problem with 4 stages, 30 realizations per stage, 3000 independently-sampled scenario paths, and a network with 30 nodes and 60 arcs. All capacity realizations are sampled from a truncated normal distribution with a mean of 30, a standard deviation of 5, and values constrained to the interval  $[10, 50]$ . The policies are

generated for the set of different Wasserstein ball size parameter  $\epsilon$  belonging to  $\{0, 0.1, 0.2, 0.3, 0.4, 0.5, 1, 5, 10\}$ . We present the results from the simulations in Figure 1: Figures 1a and 1b show the 90th and 95th percentiles of the objective values obtained by the DRO policies, and Figures 1c and 1d show the 5th and 10th percentiles of the objective values obtained by the DRR policies. Each figure contains the result from the policy generated by SDDiP with  $\epsilon = 0$ , i.e., the risk-neutral policy, as an orange horizontal line for comparison.

The DRO policies give the smaller 90th and 95th percentiles compared to the risk-neutral policy. As the ball size  $\epsilon$  increases from 0.1 to 0.3, the corresponding policies yield the lower 90th and 95th percentiles, indicating the better out-of-sample performance. As the ball size  $\epsilon$  increases further, to 0.5, 1, 5, and 10, the overall performance of the policies drops, yet their 90th and 95th percentiles remain smaller than those of the risk-neutral policy. This demonstrates that the DRO policies achieve the robustness of interdiction solutions by incorporating conservatism over unfavorable realizations in  $\Omega_t, t \in [T]$ . On the other hand, the DRR policies give the smaller 5th and 10th percentiles compared to the risk-neutral policy. This indicates that the DRR policies yield more effective interdiction solutions than the risk-neutral policy for certain scenario paths and reflect a more pessimistic perspective of the network user on the network performance. Consequently, this pessimistic view can identify the network vulnerabilities that are unnoticed by the risk-neutral policy. The level of pessimism increases as  $\epsilon$  increases from 0.1 to 0.5, except for  $\epsilon = 0.4$ . However, as it continues to increase up to  $\epsilon = 10$ , the pessimistic view on the performance diminishes.

### 8.1.3 Significance of DRR framework for MS-MFIP with data corruption

In this test setup, we assume the network user has a capability of deceiving the interdictor by providing false information about critical arcs, i.e., most essential arcs to maintain the maximum flow. Specifically, the network user introduces uncertainty about the success of interdiction on the critical arcs before the interdictor makes a decision. In this adversarial setting, we consider a risk-receptive interdictor, i.e., an optimistic decision-maker, to mitigate the impact of the false information. The significance of DRR is demonstrated through out-of-sample tests we conduct as follows.

To simplify analysis, we focus on a two-stage case of MS-MFIP, where the interdiction decision is made only in the first stage, and the flow decision is made only in the second stage. That is, the interdictor's decision is made based on the corrupted sample data of the random vector in the second stage. Let  $G = (N, A)$  represent the network. For random data generation, we generate samples of the capacity of each arc  $a \in A$  from a normal distribution with mean  $\mu_a$  and standard deviation  $\sigma_a = \mu_a/4$ , where  $\mu_a$  is drawn from a discrete uniform distribution in the interval  $[20, 40]$ . The network user manipulates these sample data by solving deterministic MS-MFIP on  $G$  with true mean values, identifying the critical arcs in  $A$ , and introducing fake data where interdiction on these arcs fails, replacing an  $\bar{\alpha}$  proportion of the sample data.

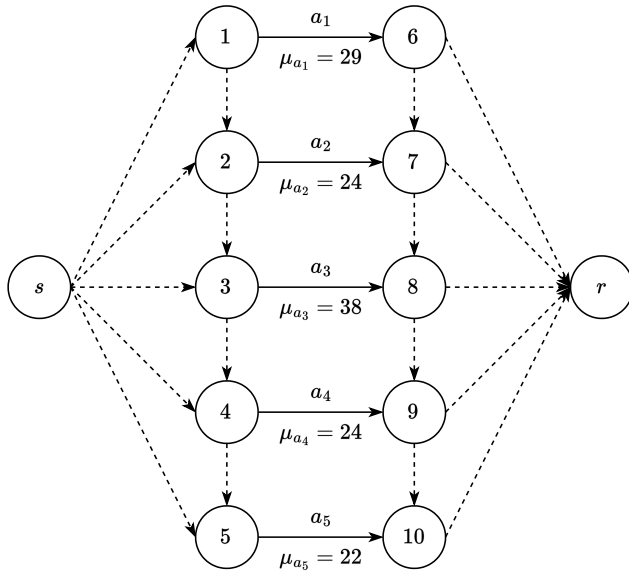


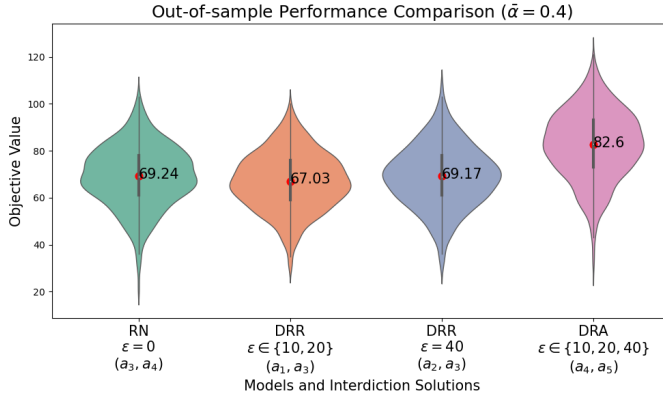
Fig. 2: A network in a  $5 \times 2$  grid shape.

The parameter  $\bar{\alpha} \in [0, 1]$  denotes the contamination level. Throughout all tests, we fix the number of scenarios to 30 for consistency.

We first consider a simple network depicted in Figure 2. The solid lines represent interdictible arcs, and the dotted lines represent non-interdictible arcs. The label on each arc represents its name and mean of capacity. In this network, the most critical arc is  $a_3$ , followed by  $a_1$  as the second most critical. The third most critical arc is either  $a_2$  or  $a_4$ , as they share the same mean value. We assume the network user manipulates the data associated with arcs  $a_1, a_2$ , and  $a_3$ . Now we assess the impact of the corrupted data and DRR on optimal interdiction solutions by solving the DRR variant of MS-MFIP with Wasserstein ambiguity set. We also solve the DRO variant for comparative purposes. The contamination level  $\bar{\alpha}$  varies in  $\{0.2, 0.4, 0.6\}$ , and the Wasserstein ball size parameter  $\epsilon$  varies in  $\{0, 10, 20, 40\}$ . The interdiction budget  $b$  is set to 2. In Table 3, we present the optimal interdiction solutions for these models. When the interdictor is risk-neutral, i.e.,  $\epsilon = 0$ , solutions  $(a_3, a_4)$  and  $(a_4, a_5)$  are chosen. This demonstrates that the fake uncertainty of interdiction success motivates the interdictor to deviate from the optimal solution  $(a_1, a_3)$  with the true mean values by choosing arcs  $a_4$  and  $a_5$ , where the success of interdiction is guaranteed across all scenarios. The risk-averse interdictor prioritizes the most conservative solution  $(a_4, a_5)$ , illustrating how DRO could be susceptible to data corruption in adversarial settings. In contrast, the risk-receptive interdictor includes the critical arcs  $a_1, a_2$ , and  $a_3$  in their solutions in almost all cases. For  $\bar{\alpha} = 0.6$ , the optimal solution diverges from these critical arcs when  $\epsilon = 10$ , selecting  $(a_4, a_5)$ , but it reverts to interdicting the critical arcs as  $\epsilon$  in-

Table 3: Optimal interdiction solutions for the DRR and DRO models with  $\bar{\alpha} \in \{0.2, 0.4, 0.6\}$  and  $\epsilon \in \{0, 10, 20, 40\}$ .

$\bar{\alpha}$	$\epsilon$						
	Risk-neutral	DRR			DRO		
	0	10	20	40	10	20	40
0.2	$(a_3, a_4)$	$(a_1, a_3)$	$(a_1, a_3)$	$(a_2, a_3)$	$(a_3, a_4)$	$(a_4, a_5)$	$(a_4, a_5)$
0.4	$(a_3, a_4)$	$(a_1, a_3)$	$(a_1, a_3)$	$(a_2, a_3)$	$(a_4, a_5)$	$(a_4, a_5)$	$(a_4, a_5)$
0.6	$(a_4, a_5)$	$(a_4, a_5)$	$(a_1, a_3)$	$(a_1, a_3)$	$(a_4, a_5)$	$(a_4, a_5)$	$(a_4, a_5)$

Fig. 3: Comparison of interdiction solutions from different models in out-of-sample test. Network has a  $5 \times 2$  grid shape, and contamination level  $\bar{\alpha}$  is 0.4.

creases to 20 and 30. Next, we evaluate the out-of-sample performance of these interdiction solutions. We randomly sample 1000 scenario paths from the clean distributions and solve the second-stage problem using the solutions provided in Table 3 for  $\bar{\alpha} = 0.4$ . The plot of objective values for the sample paths is illustrated in Figure 3, with average values denoted by red marks along with their respective numbers. Since we are evaluating the interdictor's solutions, lower objective values indicate better performance. Comparing average performances, the DRR interdiction solutions,  $(a_1, a_3)$  and  $(a_2, a_3)$ , outperform the other solutions. The DRO interdiction solution  $(a_4, a_5)$  shows the most inferior out-of-sample performance among the four solutions with an average objective value that is 23.3 percent higher than that of the DRR interdiction solution  $(a_1, a_3)$ .

To assess the significance of DRR in a more complicated instance, we perform additional tests with the network in a  $6 \times 3$  grid shape and the interdiction budget of 5. We set  $\bar{\alpha} = 0.4$ , obtain the DRR and DRO solutions for  $\epsilon \in \{0, 10, 20, 40\}$ , and conduct the out-of-sample test with these solutions. The results are plotted in Figure 4. Consistent with our observations in the simpler instance, the out-of-sample performance of the DRR solutions im-

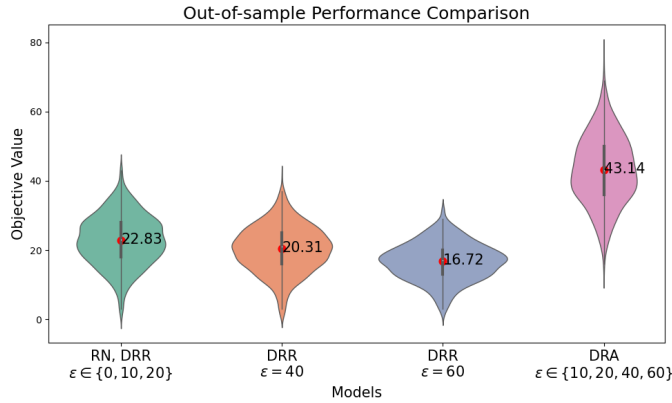


Fig. 4: Comparison of interdiction solutions from different models in out-of-sample test. Network has a  $6 \times 3$  grid shape, interdiction budget  $b$  is 5, and contamination level  $\bar{\alpha}$  is 0.4.

proves as  $\epsilon$  increases. In contrast, the DRO model produces identical solutions for all values of  $\epsilon$ , yielding the poorest average performance among the four solutions.

## 8.2 Multistage facility location interdiction problem with distributional ambiguity

We evaluate the computational performance of the algorithms for MS-FLIP with distributional ambiguity, referred to as DRR-FLIP and DRO-FLIP. The detailed formulation of MS-FLIP is provided in Appendix B.

### 8.2.1 Computational results

To generate instances, we use the following method that is similar to the one used in [42]. We first randomly sample  $(L+M)$  points from a  $100 \times 100$  grid and place demand points and facilities. For each demand point  $l \in [L]$ , we randomly choose  $\mu_l$  from a uniform discrete distribution  $[20, 40]$ . Then, to construct the support  $\Omega_t$  of size  $|\Omega|$  for each stage  $t$ , we randomly sample  $|\Omega|$  realizations of the random demand for  $l \in [L]$  from a truncated normal distribution, where the mean is  $\mu_l$ , the standard deviation is  $\sigma_l = \mu_l/4$ , and the truncation interval is  $[1, 60]$ . For all test instances, we set the Wasserstein ball size  $\epsilon$  to 10 and set the interdiction budget for each stage to one, i.e.,  $r_t = 1, t \in [T]$ . The details of the test instances are given in Table 4. Each instance, denoted by  $LI-i-T-|\Omega|$ , involves the  $i$ th network out of two randomly generated networks,  $T$  stages, and  $|\Omega|$  realizations per stage. For each row of the table, the labels  $L \times M$  and #Scenario denote the number of demand points and facilities and the number

of total scenario paths, respectively. The termination conditions are identical to those used for the DRR- and DRO-MFIP instances.

Table 5 presents the upper bounds and solution times in seconds obtained by each algorithm under the label UBound and Time (s). The numbers in each row of the table correspond to the result for a single instance. Note that smaller bounds are better since they are upper bounds. For DRO-FLIP, the DRO-SDDP-C algorithm provides the upper bounds better than the DRO-SDDP-R algorithm for all instances. Also, on average, the DRO-SDDP-C algorithm is 38.9 times faster than the DRO-SDDP-R algorithm, and this advantage increases to 48.4 times for 11 instances where both the algorithms produce the same bounds. The performance of the DRO-SDDP-R algorithm is comparatively susceptible to the number of realizations

per stage. For example, the DRO-SDDP-R algorithm takes 458.7 seconds to solve LI-2-3-10, but it fails to converge within the time limit when solving LI-2-3-50. The DRO-SDDP-C algorithm, however, converges for the both instances within the time limit. When comparing the results for DRR-FLIP, the DRR-SDDP-C algorithm provides better upper bounds than the DRR-SDDP-R algorithm for all the instances. In terms of solution time, the DRR-SDDP-C algorithm is, on average, 25.8 times faster than the DRR-SDDP-R algorithm for all instances, and 26.7 times faster for 13 instances where both the algorithms produce the same bounds. As discussed in the previous section, this shows that the both DRO-SDDP-R and DRR-SDDP-R algorithms require more time to solve due to the larger subproblems resulting from the reformulation techniques.

## 9 Conclusion

We studied multistage stochastic integer and disjunctive programs under distributional ambiguity, considering the distributional risk-receptiveness and robustness in a decision making process. For distributionally risk-receptive multistage stochastic integer programs (DRR-MSIPs) without and with decision dependent uncertainty, we presented new classes of cutting planes and reformulation-based approaches to derive convex approximations of the op-

Table 4: Details of DRR- and DRA-FLIP instances

Instance	$L \times M$	$T$	$ \Omega $	#Scenario
LI-1-3-10	10 x 20	3	10	100
LI-1-3-20			20	400
LI-1-3-50			50	2500
LI-1-4-10		4	10	1000
LI-1-4-20			20	8000
LI-1-4-50			50	125000
LI-1-5-10		5	10	10000
LI-1-5-20			20	160000
LI-1-5-50			50	6250000
LI-2-3-10	15 x 30	3	10	100
LI-2-3-20			20	400
LI-2-3-50			50	2500
LI-2-4-10		4	10	1000
LI-2-4-20			20	8000
LI-2-4-50			50	125000
LI-2-5-10		5	10	10000
LI-2-5-20			20	160000
LI-2-5-50			50	6250000

Table 5: Performance comparison of algorithms for DRR- and DRA-FLIP instances

Instance	DRA-SDDP-C		DRA-SDDP-R		DRR-SDDP-C		DRR-SDDP-R	
	UBound	Time (s)	UBound	Time (s)	UBound	Time (s)	UBound	Time (s)
LI-1-3-10	431.29	29.1	431.29	169.7	437.28	28.4	437.28	109.1
LI-1-3-20	327.32	49.8	327.32	1761.0	333.39	41.3	333.39	630.6
LI-1-3-50	384.83	75.8	384.83	10800+	392.50	74.5	392.50	3373.3
LI-1-4-10	489.80	53.2	489.80	1170.2	498.44	48.6	498.44	622.1
LI-1-4-20	479.23	68.5	479.23	6511.9	491.03	78.7	491.03	1963.0
LI-1-4-50	651.45	105.6	651.45	10800+	668.28	106.8	668.28	10800+
LI-1-5-10	654.12	73.6	654.12	1744.3	670.83	74.5	670.83	1501.3
LI-1-5-20	774.27	89.5	774.27	4853.6	792.99	80.3	792.99	4635.5
LI-1-5-50	929.85	156.2	931.32	10800+	955.35	162.9	978.63	10800+
LI-2-3-10	452.23	74.1	452.23	458.7	456.76	78.3	456.76	333.4
LI-2-3-20	463.56	119.4	463.56	1894.1	468.56	112.7	468.56	772.8
LI-2-3-50	489.47	459.4	518.37	10800+	495.10	453.1	495.10	10800+
LI-2-4-10	500.09	369.0	500.09	10800+	506.22	378.5	506.22	4815.8
LI-2-4-20	508.09	476.5	531.57	10800+	514.78	444.9	514.78	8376.8
LI-2-4-50	556.50	761.3	721.03	10800+	564.69	692.1	1194.59	10800+
LI-2-5-10	599.37	534.6	661.32	10800+	605.32	550.7	1009.53	10800+
LI-2-5-20	816.66	879.1	904.14	10800+	828.21	1026.6	1339.16	10800+
LI-2-5-50	756.16	1732.8	949.12	10800+	767.99	1927.1	1756.89	10800+

timistic expected cost-to-go functions. For distributionally robust multistage stochastic integer programs (DRO-MSIPs), we presented a cutting plane-based and reformulation-based approximations of the pessimistic expected cost-to-go functions. We developed algorithms using these approximations and provided their finite convergence analysis. Furthermore, we extended the algorithms for distributionally risk-receptive and distributionally robust multistage stochastic disjunctive programs (DRR- and DRO-MSDPs) and then presented applications of them to solve DRR- and DRO-MSIPs using a hierarchy of relaxations. We compared the algorithms for DRR- and DRO-MSIPs by solving multistage stochastic network interdiction problems under distributional ambiguity that are multi-round attacker-defender games and have not been addressed in the literature. The computational results show that the algorithms using the cutting plane-based approximations outperform the algorithms using the reformulation-based approximations by 26.1 times, on average, in terms of solution time until convergence. In addition, we conducted out-of-sample tests, and their results demonstrate that the DRA policies provide robust decision rules for uncertainty, while the DRR policies may reveal the network vulnerabilities that are overlooked by risk-neutral policies for uncertainty. In an adversarial setting, we showcased that the DRR policies can be used to mitigate the impact of data corruption.

*Data Availability Statement* The instances used for computational studies in this paper are available in “Multistage-Interdiction-Problems” folder at <https://github.com/Bansal-ORGroup/>.



## References

1. Balas, E.: Disjunctive programming. Springer (2018)
2. Bansal, M., Huang, K.L., Mehrotra, S.: Decomposition algorithms for two-stage distributionally robust mixed binary programs. *SIAM Journal on Optimization* **28**(3), 2360–2383 (2018)
3. Bansal, M., Zhang, Y.: Scenario-based cuts for structured two-stage stochastic and distributionally robust  $p$ -order conic mixed integer programs. *Journal of Global Optimization* **81**(2), 391–433 (2021)
4. Basciftci, B., Ahmed, S., Shen, S.: Distributionally robust facility location problem under decision-dependent stochastic demand. *European Journal of Operational Research* **292**(2), 548–561 (2021)
5. Bayraksan, G., Love, D.K.: Data-driven stochastic programming using phi-divergences, pp. 1–19. *INFORMS TutORials in Operations Research* (2015)
6. Bayraksan, G., Maggioni, F., Faccini, D., Yang, M.: Bounds for multistage mixed-integer distributionally robust optimization. *SIAM Journal on Optimization* **34**(1), 682–717 (2024)
7. Birge, J.R.: Decomposition and partitioning methods for multistage stochastic linear programs. *Operations Research* **33**(5), 989–1007 (1985)
8. Blanchet, J., Li, J., Lin, S., Zhang, X.: Distributionally robust optimization and robust statistics. *arXiv preprint arXiv:2401.14655* (2024)
9. Blanchet, J., Li, J., Pelger, M., Zanotti, G.: Automatic outlier rectification via optimal transport. *arXiv preprint arXiv:2403.14067* (2024)
10. Brown, G., Carlyle, M., Salmerón, J., Wood, K.: Defending critical infrastructure. *Interfaces* **36**(6), 530–544 (2006)
11. Cao, J., Gao, R.: Contextual decision-making under parametric uncertainty and data-driven optimistic optimization. *Optimization Online preprint:2021/10/8634* (2021)
12. Carøe, C.C., Schultz, R.: Dual decomposition in stochastic integer programming. *Operations Research Letters* **24**(1), 37–45 (1999)
13. Church, R.L., Scaparra, M.P., Middleton, R.S.: Identifying critical infrastructure: The median and covering facility interdiction problems. *Annals of the Association of American Geographers* **94**(3), 491–502 (2004)
14. Cormican, K.J., Morton, D.P., Wood, R.K.: Stochastic network interdiction. *Operations Research* **46**(2), 184–197 (1998)
15. Dowson, O., Kapelevich, L.: Sddp. jl: a julia package for stochastic dual dynamic programming. *INFORMS Journal on Computing* **33**(1), 27–33 (2020)
16. Duchi, J.C., Glynn, P.W., Namkoong, H.: Statistics of robust optimization: A generalized empirical likelihood approach. *Mathematics of Operations Research* **46**(3), 946–969 (2021)
17. Duque, D., Mehrotra, S., Morton, D.P.: Distributionally robust two-stage stochastic programming. *SIAM Journal on Optimization* **32**(3), 1499–1522 (2022)
18. Duque, D., Morton, D.P.: Distributionally robust stochastic dual dynamic programming. *SIAM Journal on Optimization* **30**(4), 2841–2865 (2020)
19. Gangammanavar, H., Bansal, M.: Stochastic decomposition method for two-stage distributionally robust linear optimization. *SIAM Journal on Optimization* **32**(3), 1901–1930 (2022)
20. Gao, R., Kleywegt, A.: Distributionally robust stochastic optimization with wasserstein distance. *Mathematics of Operations Research* **48**(2), 603–655 (2023)
21. Gotoh, J.y., Kim, M.J., Lim, A.E.: A data-driven approach to beating saa out of sample. *Operations Research* (2023)
22. Janjarassuk, U., Linderoth, J.: Reformulation and sampling to solve a stochastic network interdiction problem. *Networks* **52**(3), 120–132 (2008)
23. Jiang, N., Xie, W.: Distributionally favorable optimization: A framework for data-driven decision-making with endogenous outliers. *SIAM Journal on Optimization* **34**(1), 419–458 (2024)
24. Kang, S., Bansal, M.: Distributionally risk-receptive and risk-averse network interdiction problems with general ambiguity set. *Networks* **81**(1), 3–22 (2023)

25. Kosmas, D., Sharkey, T.C., Mitchell, J.E., Maass, K.L., Martin, L.: Multi-period max flow network interdiction with restructuring for disrupting domestic sex trafficking networks. *Annals of Operations Research* pp. 1–64 (2022)
26. Lan, G.: Complexity of stochastic dual dynamic programming. *Mathematical Programming* **191**(2), 717–754 (2022)
27. Lulli, G., Sen, S.: A branch-and-price algorithm for multistage stochastic integer programming with application to stochastic batch-sizing problems. *Management Science* **50**(6), 786–796 (2004)
28. Luo, F., Mehrotra, S.: Distributionally robust optimization with decision dependent ambiguity sets. *Optimization Letters* **14**(8), 2565–2594 (2020)
29. Malaviya, A., Rainwater, C., Sharkey, T.: Multi-period network interdiction problems with applications to city-level drug enforcement. *IIE Transactions* **44**(5), 368–380 (2012)
30. Nakao, H.: Distributionally robust optimization in sequential decision making. Ph.d. thesis, University of Michigan, Ann Arbor (2021)
31. Nakao, H., Jiang, R., Shen, S.: Distributionally robust partially observable markov decision process with moment-based ambiguity. *SIAM Journal on Optimization* **31**(1), 461–488 (2021)
32. Nguyen, V.A., Shafieezadeh-Abadeh, S., Yue, M.C., Kuhn, D., Wiesemann, W.: Optimistic distributionally robust optimization for nonparametric likelihood approximation. In: *Advances in Neural Information Processing Systems*, vol. 32. Curran Associates, Inc. (2019)
33. Noyan, N., Rudolf, G., Lejeune, M.: Distributionally robust optimization under a decision-dependent ambiguity set with applications to machine scheduling and humanitarian logistics. *INFORMS Journal on Computing* **34**(2), 729–751 (2022)
34. Park, J., Bayraksan, G.: A multistage distributionally robust optimization approach to water allocation under climate uncertainty. *European Journal of Operational Research* **306**(2), 849–871 (2023)
35. Pereira, M.V., Pinto, L.M.: Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming* **52**(1), 359–375 (1991)
36. Philpott, A.B., de Matos, V.L., Kapelevich, L.: Distributionally robust SDDP. *Computational Management Science* **15**(3), 431–454 (2018)
37. Royset, J.O., Chen, L.L., Eckstrand, E.: Rockafellian relaxation in optimization under uncertainty: Asymptotically exact formulations. *arXiv preprint arXiv:2204.04762* (2023)
38. Sadana, U., Delage, E.: The value of randomized strategies in distributionally robust risk-averse network interdiction problems. *INFORMS Journal on Computing* **35**(1), 216–232 (2023)
39. Scarf, H.: A min-max solution of an inventory problem, pp. 201–209. Stanford University Press (1958)
40. Smith, J.C., Song, Y.: A survey of network interdiction models and algorithms. *European Journal of Operational Research* **283**(3), 797–811 (2020)
41. Song, J., Zhao, C.: Optimistic distributionally robust policy optimization. *arXiv preprint arXiv:2006.07815* (2020)
42. Yu, X., Shen, S.: Multistage distributionally robust mixed-integer programming with decision-dependent moment-based ambiguity sets. *Mathematical Programming* **196**(1), 1025–1064 (2022)
43. Zhang, S., Sun, X.A.: Stochastic dual dynamic programming for multistage stochastic mixed-integer nonlinear optimization. *Mathematical Programming* **196**(1), 935–985 (2022)
44. Zou, J., Ahmed, S., Sun, X.A.: Stochastic dual dynamic integer programming. *Mathematical Programming* **175**(1), 461–502 (2019)

## Appendix A List of Major Abbreviations

- *DRO/DRR*: Distributionally robust (or robustness)/Distributionally risk-receptive (or risk-receptiveness).
- *MSLP/MSIP/MSDP*: Multistage stochastic linear/integer/disjunctive program.
- *SDDP*: Stochastic dual dynamic programming.
- *DA-SDDP*: Our customized SDDP that addresses distributional ambiguity.
- *DRR-SDDP-C*: DA-SDDP for DRR-MSIPs with the cutting plane-based approximation (10).
- *DRR-SDDP-R*: DA-SDDP for DRR-MSIPs with the reformulation-based approximation (13).
- *DRO-SDDP-C*: DA-SDDP for DRO-MSIPs with the cutting plane-based approximation (22).
- *DRO-SDDP-R*: DA-SDDP for DRO-MSIPs with the reformulation-based approximation (23).
- *MS-MFIP*: Multistage stochastic maximum flow interdiction problem
- *MS-FLIP*: Multistage stochastic facility location interdiction problem

## Appendix B Formulations of the Network Interdiction Problems

We provide detailed formulations of the multistage NIPs used in our computational tests. We focus on the risk-neutral formulations, as their DRO and DRR variants can be readily derived from these by appropriately maximizing or minimizing with respect to the probability distribution within an ambiguity, as in (3) and (5).

### B.1 Formulation of MS-MFIP

Consider a directed and capacitated network, denoted by  $G = (N, A)$  where  $N$  is a set of nodes and  $A$  is a set of directed arcs of the network. The interdictor's objective is to minimize the total flow from the source node  $s$  to the sink node  $r$  of the network  $G$  by interdicting a subset of arcs in  $A$ . In contrast, the network user's objective is to maximize the total flow given the interdicted network. At each stage, both the interdictor and the network user make their decisions as follows: The interdictor removes a set of arcs from the network  $G$  given an interdiction budget, and the network user finds a maximum flow after observing the interdictor's decision. It is assumed that after an arc is interdicted, the network user cannot use it till the end of the time horizon.

Let  $x_t$  and  $y_t$  be the interdiction decision vector and the flow decision vector, respectively, for stage  $t \in [T]$ . An interdiction decision  $x_{t,a}$ , for each arc  $a \in A$ , is binary, i.e.,  $x_{t,a} = 1$ , if an interdiction occurs on  $a \in A$ , and  $x_{t,a} = 0$ , otherwise. Each arc  $a \in A$  is associated with the interdiction cost, denoted by  $f_{t,a}$ . The interdiction budget is denoted by  $b_t$ , for each stage  $t \in [T]$ . We

assume that the capacity of arc  $a \in A$  is uncertain and denoted by  $c_{t,a}(\omega_t)$ . We denote the set of the outgoing arcs and the set of the incoming arcs of node  $n \in N$  by  $\delta^+(n)$  and  $\delta^-(n)$ , respectively. For the brevity, we assume there exists a dummy arc from  $r$  to  $s$  in  $A$  associated with infinite capacity and interdiction cost. Then, the bellman equation form of MS-MFIP is given by (1) and (2) where for each  $t \in [T]$ ,

$$Q_t(x_{t-1}, \omega_t) := \min \psi_t(x_t, \omega_t) + \mathbb{E}_{P_{t+1}} [Q_{t+1}(x_t, \omega_{t+1})] \quad (32a)$$

$$\text{s.t. } x_t \geq x_{t-1} \quad (32b)$$

$$\sum_{a \in A} f_{t,a} x_{t,a} \leq b_t + \sum_{a \in A} f_{t,a} x_{t-1,a} \quad (32c)$$

$$x_t \in \{0, 1\}^{|A|}, \quad (32d)$$

and

$$\psi_t(x_t, \omega_t) := \max_{y_{t,a} \geq 0} \left\{ y_{t,(r,s)} : \sum_{a \in \delta^+(n)} y_{t,a} - \sum_{a \in \delta^-(n)} y_{t,a} = 0, \quad \forall n \in N, \quad (33a) \right.$$

$$\left. y_{t,a} \leq c_{t,a}(\omega_t)(1 - x_{t,a}), \quad \forall a \in A \right\}. \quad (33b)$$

In interdictor's problem (32), constraint (32b) ensures that the impact of interdiction on an arc remains till the end of the time horizon, and constraint (32c) restricts the total interdiction cost within the given budget  $b_t$ . Given an interdiction solution  $x_t$ , the objective function of the interdictor's problem at stage  $t \in [T]$ ,  $\psi_t(x_t, \omega_t)$ , is a value function that provides a maximum flow over the interdicted network. The function  $\psi_t(x_t, \omega_t)$  is computed by solving the network user's problem (33) where decision variable  $y_{t,a}$  represents a flow on arc  $a$  in  $A$ , constraints (33a) enforce the flow balance on nodes in  $N$ , and constraints (33b) restrict the capacity of arcs in  $A$ . Notice that because of constraints (33b),  $y_{t,a}$  is restricted to be zero if the interdiction occurs on arc  $a \in A$ , i.e., if  $x_{t,a} = 1$ .

To solve each stage's problem during our tests, we first dualize the network user's problem (33) and then reformulate the problem (32) into a single-level minimization mixed-binary program.

## B.2 Formulation of MS-FLIP

The single-stage (deterministic) FLIP, also known as the  $r$ -interdiction median problem, is introduced by Church et al. [13]. The objective of the single-stage FLIP is to find a subset of  $r$  facilities whose removal maximizes the network user's objective of minimizing the total weighted distance. MS-FLIP extends this problem to a multistage stochastic setting, where interdiction decisions are made sequentially over multiple time periods under demand uncertainty.

Let  $L$  and  $M$  be the number of demand points and facilities, respectively. At stage  $t \in [T]$ , let  $x_{tm} \in \{0, 1\}$  be an interdiction decision variable, which

equals 1, if the interdiction occurs on facility  $m \in [M]$ , or equals 0, otherwise. Variable  $y_{tlm} \in \{0, 1\}$  denotes an assignment decision that represents whether demand point  $l$  is assigned to facility  $m$ . We denote the random weighted distance between  $l$  and  $m$  by  $c_{tlm}(\omega_t) = a_{tl}(\omega_t)d_{lm}$ , where  $d_{lm} > 0$  is the Euclidean distance between  $l$  and  $m$ , and  $a_{tl}(\omega_t)$  is an uncertain demand at point  $l \in [L]$ . The formulation of MS-FLIP is given by (1) and (2) with

$$Q_t(x_{t-1}, \omega_t) := \max \sum_{l \in [L], m \in [M]} c_{tlm}(\omega_t) y_{tlm} + \mathbb{E}_{P_{t+1}}[Q_{t+1}(x_t, \omega_{t+1})] \quad (34a)$$

$$\text{s.t. } \sum_{m \in [M]} y_{tlm} = 1, \quad \forall l \in [L], \quad (34b)$$

$$x_t \geq x_{t-1}, \quad (34c)$$

$$\sum_{m \in [M]} (x_{tm} - x_{t-1,m}) = r_t, \quad (34d)$$

$$\sum_{n \in S_{lm}} y_{tln} \leq x_{tm}, \quad \forall l \in [L], m \in [M], \quad (34e)$$

$$x_t \in \{0, 1\}^M, \quad y_t \in \{0, 1\}^{L \times M}. \quad (34f)$$

Here  $S_{lm} := \{n \in [M] : d_{ln} > d_{lm}\}$  is the set of facilities that are farther than facility  $m \in [M]$  is from demand point  $l \in [L]$ . The first term of the objective function (34a) represents the total weighted distance of the assignment of demand points to non-interdicted facilities. Constraints (34b) enforce each demand point to be assigned to a facility. Constraint (34c) ensures that the facilities interdicted from the previous stages remain interdicted for the current stage. Constraint (34d) ensures that the total number of interdictions occurred at the current stage equals to the budget  $r_t$ . Constraint (34e), for each  $l \in [L]$  and  $m \in [M]$ , prevents the demand point  $l$  from being assigned to facilities farther than the facility  $m$ , unless the facility  $m$  is interdicted. It should be noted that problem (34) is a single-level maximization problem, but not the max-min form of typical interdiction problems, because of the assumption that each facility has enough capacity to cover all demand values and the demand point is always assigned to the closest facility through constraints (34e).

## Appendix C Proofs

### C.1 Proof of Theorem 1

*Proof* For any  $k \in [K_t^l]$  and  $x_t \in \{0, 1\}^{d_x}$ , it is satisfied that

$$\begin{aligned}
\min_{P_{t+1} \in \mathcal{P}_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i \hat{Q}_{t+1}^l(x_t, \omega_{t+1}^i) &\geq \min_{P_{t+1} \in \mathcal{P}_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i \left( (\alpha_t^{i,k})^\top x_t + \beta_t^{i,k} \right) \\
&= \min_{P_{t+1} \in \mathcal{P}_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i \left( (\alpha_t^{i,k})^\top (x_t - \hat{x}_t) + (\alpha_t^{i,k})^\top \hat{x}_t + \beta_t^{i,k} \right) \\
&\geq \gamma_t^k + \min_{P_{t+1} \in \mathcal{P}_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i (\alpha_t^{i,k})^\top (x_t - x_t^k) \\
&\geq \gamma_t^k + \sum_{j \in [d_x]} \min_{P_{t+1} \in \mathcal{P}_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i \alpha_{t,j}^{i,k} (x_{t,j} - x_{t,j}^k).
\end{aligned}$$

In the last inequality, if  $x_{t,j}^k = 0$ , then  $(x_{t,j} - x_{t,j}^k) \in \{0, 1\}$ . It follows that we can fix the coefficient of the term  $(x_{t,j} - x_{t,j}^k)$  to

$$\pi_{t,j}^k = \min_{P_{t+1} \in \mathcal{P}_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i \alpha_{t,j}^{i,k}.$$

If  $x_{t,j}^k = 1$ , then  $(x_{t,j} - x_{t,j}^k) \in \{0, -1\}$  and we can fix the coefficient to

$$\pi_{t,j}^k = \max_{P_{t+1} \in \mathcal{P}_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i \alpha_{t,j}^{i,k}.$$

Fixing the coefficients for all  $j \in d_x$  and  $k \in [K_t^l]$ , we have

$$\min_{P_{t+1} \in \mathcal{P}_{t+1}} \sum_{i \in \mathcal{N}_{t+1}} p_{t+1}^i \hat{Q}_{t+1}^l(x_t, \omega_{t+1}^i) \geq (\pi_t^k)^\top (x_t - x_t^k) + \gamma_t^k, \quad \forall k \in [K_t^l]. \quad (35)$$

Function  $\phi_t^{l,C}(x_t)$  is constructed using the affine functions in the right-hand side of (35), and thus it follows that  $\phi_t^{l,C}(x_t) \leq \mathcal{Q}_{t+1}^{RR}(x_t)$ .  $\square$

### C.2 Proof of Proposition 1

*Proof* Function  $\mathcal{Q}_t^{RR}(x_{t-1}) = -\max_{P_t \in \mathcal{P}_t(x_{t-1})} \mathbb{E}_{P_t}[-Q_t^{RR}(x_{t-1}, \omega_t)]$ . By applying Theorem 1 in [20] to the maximization problem, the dual formulation is given by

$$\begin{aligned}
\mathcal{Q}_t^{RR}(x_{t-1}) &= - \min_{\rho_t \geq 0, v_t^i} \epsilon_t(x_{t-1}) \rho_t + \sum_{i \in \mathcal{N}_t} \bar{p}_t^i v_t^i \\
&\text{s.t. } \rho_t \|\omega_t - \omega_t^i\| + v_t^i \geq -Q_t^{RR}(x_{t-1}, \omega_t), \quad \forall \omega_t \in \Omega_t, i \in \mathcal{N}_t.
\end{aligned}$$

where  $\bar{p}_t^i = 1/N_t$ . Consequently, we obtain

$$\begin{aligned} & \mathcal{Q}_t^{RR}(x_{t-1}) \\ &= -\min_{\rho_t \geq 0} \left\{ \epsilon_t(x_{t-1})\rho_t + \sum_{i \in \mathcal{N}_t} \bar{p}_t^i \max_{\omega_t \in \Omega_t} \left\{ -\rho_t \|\omega_t - \omega_t^i\| - Q_t^{RR}(x_{t-1}, \omega_t) \right\} \right\} \\ &= \max_{\rho_t \geq 0} \left\{ -\epsilon_t(x_{t-1})\rho_t + \sum_{i \in \mathcal{N}_t} \bar{p}_t^i \min_{\omega_t \in \Omega_t} \left\{ \rho_t \|\omega_t - \omega_t^i\| + Q_t^{RR}(x_{t-1}, \omega_t) \right\} \right\}. \quad \square \end{aligned}$$

### C.3 Proof of Theorem 2

*Proof* By Proposition 1, we have

$$\mathcal{Q}_t^{RR}(x_{t-1}) = \max_{\rho_t \geq 0} \left\{ -\epsilon_t(x_{t-1})\rho_t + \sum_{i \in \mathcal{N}_t} \bar{p}_t^i V_t^i(x_{t-1}, \rho_t) \right\}, \quad (36)$$

where  $\bar{p}_t^i = 1/N_t$  and

$$V_t^i(x_{t-1}, \rho_t) := \min_{\omega_t \in \Omega_t} \left\{ \rho_t \|\omega_t - \omega_t^i\| + Q_t^{RR}(x_{t-1}, \omega_t) \right\} \text{ for } i \in \mathcal{N}_t. \quad (37)$$

Using the valid cuts with coefficients  $\{(\pi_t^k, \gamma_t^k)\}_{k \in [K_t]}$  for  $Q_t^{RR}(x_t)$ , we obtain a lower-bounding approximation  $\underline{V}_t^i$  of  $V_t^i$  under assumptions 3 and 5, where

$$\underline{V}_t^i(x_{t-1}, \rho_t) := \min_{\omega_t, x_t} \rho_t \|\omega_t - \omega_t^i\| + c_t^\top x_t + \theta_t \quad (38a)$$

$$\text{s.t. } A_t x_t - \omega_t \geq -C_t x_{t-1} \quad (38b)$$

$$\omega_t \in \Omega_t = [l_t, u_t]^{d_\omega} \quad (38c)$$

$$\theta_t \geq (\pi_t^k)^\top x_t + \gamma_t^k, \quad \forall k \in [K_t] \quad (38d)$$

$$x_t \in \mathcal{X}_t, \omega_t \in \mathbb{R}^{d_\omega}, \theta_t \in \mathbb{R}. \quad (38e)$$

Then, a Lagrangian relaxation of (38) without integrality restrictions on  $x_t$ , where  $\lambda_t^i, (\mu_t^i, \nu_t^i), \zeta_{tk}^i$  are dual multipliers associated with constraints (38b), (38c), (38d), respectively, is given by

$$\begin{aligned} & \min_{(x_t, \omega_t, \theta_t) \in \mathbb{R}^{d_x + d_\omega + 1}} -(\lambda_t^i)^\top C_t x_{t-1} + (\mu_t^i)^\top l_t - (\nu_t^i)^\top u_t + \rho_t \|\omega_t - \omega_t^i\| \\ & \quad + (\lambda_t^i - \mu_t^i + \nu_t^i)^\top \omega_t + \left( c_t - \lambda_t^i A_t + \sum_{k \in [K_t]} \pi_t^k \right)^\top x_t \quad (39) \\ & \quad + \left( 1 - \sum_{k \in [K_t]} \zeta_{tk}^i \right) \theta_t + \sum_{k \in [K_t]} \gamma_t^k \zeta_{tk}^i. \end{aligned}$$

Next, we prove that the Lagrangian dual is equivalent to

$$\begin{aligned} \max \quad & -(\lambda_t^i)^\top C_t x_{t-1} + (\mu_t^i)^\top l_t - (\nu_t^i)^\top u_t \\ & + (\lambda_t^i - \mu_t^i + \nu_t^i)^\top \omega_t^i + \sum_{k \in [K_t]} \gamma_t^k \zeta_{tk} \end{aligned} \quad (40a)$$

$$\text{s.t. } A_t^\top \lambda_t^i - \sum_{k \in [K_t]} \pi_t^k \zeta_{tk} = c_t \quad (40b)$$

$$\|\lambda_t^i - \mu_t^i + \nu_t^i\|_* \leq \rho_t \quad (40c)$$

$$\sum_{k \in [K_t]} \zeta_{tk}^i = 1 \quad (40d)$$

$$(\lambda_t^i, \mu_t^i, \nu_t^i, \zeta_t^i) \geq 0. \quad (40e)$$

As a consequence, we can substitute  $V_t^i$  for all scenarios  $i \in \mathcal{N}_t$  in (36) with the value functions of these Lagrangian duals and obtain problem (20). Since the Lagrangian duals are maximization problems, any feasible solutions of them yield a valid cut in the form of (18), and tightest cuts are attained by optimal solutions.

Constraints (40b) and (40d) are straightforwardly derived to ensure  $(c_t - \lambda A_t + \sum_{k \in [K_t]} \pi_t^k) = 0$ , and  $(1 - \sum_{k \in [K_t]} \zeta_k) = 0$ , thereby ensuring that (39) is not unbounded because of unrestricted  $x_t$  and  $\theta_t$ . Likewise, as shown below, constraint (40c) ensures that

$$\tau := \inf_{\omega_t \in \mathbb{R}^{d_\omega}} \{ \rho_t \|\omega_t - \omega_t^i\| + (\lambda_t^i - \mu_t^i + \nu_t^i)^\top \omega_t \} > -\infty.$$

For the ease of exposition, we use  $(\lambda, \mu, \nu)$  instead of  $(\lambda_t^i, \mu_t^i, \nu_t^i)$  in the following claim.

*Claim* Let  $\|\cdot\|_*$  be the dual norm of  $\|\cdot\|$ . Given  $\rho_t \geq 0$ , we have  $\tau > -\infty$  if and only if  $(\rho_t - \|\lambda - \mu + \nu\|_*) \geq 0$ . Also, if  $(\rho_t - \|\lambda - \mu + \nu\|_*) \geq 0$ , then  $\tau = (\lambda - \mu + \nu)^\top \omega_t^i$ .

*Proof* By letting  $\omega_t' = \omega_t^i - \omega_t$ , the infimum  $\tau$  can be rewritten as follows.

$$\tau = (\lambda - \mu + \nu)^\top \omega_t^i + \inf_{\omega_t' \in \mathbb{R}^{d_\omega}} \left\{ \rho_t \|\omega_t'\| - (\lambda - \mu + \nu)^\top \omega_t' \right\}. \quad (41a)$$

Let  $\tau^0 := \inf_{\omega_t' \in \mathbb{R}^{d_\omega}} \{ \rho_t \|\omega_t'\| - (\lambda - \mu + \nu)^\top \omega_t' \}$ . If  $(\rho_t - \|\lambda - \mu + \nu\|_*) \geq 0$ , then we have

$$\tau^0 \geq \inf_{\omega_t' \in \mathbb{R}^{d_\omega}} \left\{ \|\omega_t'\| \cdot (\rho_t - \|\lambda - \mu + \nu\|_*) \right\} = 0.$$

The inequality holds by Hölder's inequality. The infimum  $\tau^0 = 0$  since  $\omega_t' = 0$  is feasible. Therefore,  $\tau = (\lambda - \mu + \nu)^\top \omega_t^i$ .

Now we show the “only-if” direction using proof by contradiction. Suppose that  $\tau > -\infty$  and  $(\rho_t - \|\lambda - \mu + \nu\|_*) < 0$ . This implies that  $\tau^0$  is also bounded.



By definition of dual norm, i.e.,  $\|\lambda - \mu + \nu\|_* = \max_{\omega_t \in \mathbb{R}^{d_\omega}} \{\omega_t^\top (\lambda - \mu + \nu) : \|\omega_t\| \leq 1\}$ , there exists  $\bar{\omega}_t \in \mathbb{R}^{d_\omega}$  such that

$$\rho_t < \bar{\omega}_t^\top (\lambda - \mu + \nu) \text{ and } \|\bar{\omega}_t\| \leq 1.$$

Thus,  $\rho_t \|\bar{\omega}_t\| - \bar{\omega}_t^\top (\lambda - \mu + \nu) < 0$ . Considering scalar  $r \rightarrow \infty$ , we have  $\rho_t \|r\bar{\omega}_t\| - (r\bar{\omega}_t)^\top (\lambda - \mu + \nu) \rightarrow -\infty$ . By taking  $\omega'_t = r\bar{\omega}_t$ , it follows that  $\tau^0 = -\infty$ , and  $\tau = -\infty$ .  $\square$

#### C.4 Proof of Theorem 3

*Proof* We first show that the finiteness of the while loop. Let  $\{\bar{x}_t^l(\xi_{[t]}), \bar{y}_t^l(\xi_{[t]}) : t \in [T]\}$  be a policy that is defined by the forward step at iteration  $l$  of the DRR-SDDP-C algorithm. Then, to achieve the optimality of the policy, it suffices to show that  $\phi_t^{l,C}(\bar{x}_t^l(\xi_{[t]})) = \mathcal{Q}_{t+1}^{RR}(\bar{x}_t^l(\xi_{[t]}))$  for  $t \in [T-1]$  and all  $\xi \in \Xi$ .

Let  $\mathcal{L}$  denote the set of iterations where the policy is non-optimal. For each  $t \in [T-1]$ , let  $l_t \in \mathcal{L}$  be the largest iteration such that  $\phi_t^{l_t,C}(\bar{x}_t^{l_t}(\xi_{[t]})) < \mathcal{Q}_{t+1}^{RR}(\bar{x}_t^{l_t}(\xi_{[t]}))$  for some  $\xi \in \Xi$ . Also, given iteration  $m$ , let  $\bar{X}_t^m, t \in [T-1]$ , be the set of  $\bar{x}_t^l(\xi_{[t]})$  for all future iterations  $l \in \mathcal{L} \setminus [m]$  and  $\xi \in \Xi$ , i.e.,  $\bar{X}_t^m = \{\bar{x}_t^l(\xi_{[t]}) : \xi \in \Xi, l > m, l \in \mathcal{L}\}$ .

We first show that  $l_{T-1}$  is finite with probability one. At the forward step of iteration  $l$ , if we observe that  $\phi_{T-1}^{l,C}(\bar{x}) < \mathcal{Q}_T^{RR}(\bar{x})$  for any  $\bar{x} \in \bar{X}_{T-1}^1$ , then for all future iterations  $m > l$  we have

$$\phi_{T-1}^{m,C}(\bar{x}) \geq \min_{P_T \in \mathcal{P}_T} \sum_{i \in \mathcal{N}_T} p_T^i \hat{Q}_T^l(\bar{x}, \omega_T^i) = \mathcal{Q}_T^{RR}(\bar{x}). \quad (42)$$

The above holds since the cut added at iteration  $l$  (in the form of constraint (10)) satisfies both (26a) and (26b), and by definition  $\hat{Q}_T^l$  is equivalent to  $\mathcal{Q}_T^{RR}$ . It follows that  $\phi_{T-1}^{m,C}(\bar{x}) = \mathcal{Q}_T^{RR}(\bar{x})$  for any  $m > l$ . Since  $|\bar{X}_{T-1}^1| < \infty$  and every  $\xi \in \Xi$  has a positive probability to be sampled, it holds with probability one that  $\phi_{T-1}^{l,C}(\bar{x}) = \mathcal{Q}_T^{RR}(\bar{x})$  for all  $\bar{x} \in \bar{X}_{T-1}^1$  after finitely many iterations. This shows that  $l_{T-1} < \infty$  with probability one.

Next, we show that  $l_{T-2}$  is also finite with probability one. Suppose at iteration  $l \geq l_{T-1}$  we observe that  $\phi_{T-2}^{l,C}(\bar{x}) < \mathcal{Q}_{T-1}^{RR}(\bar{x})$  for any  $\bar{x} \in \bar{X}_{T-2}^{l_{T-1}}$ . Then, we have  $\phi_{T-2}^{m,C}(\bar{x}) = \mathcal{Q}_{T-1}^{RR}(\bar{x})$  for all future iterations  $m > l$  since

$$\begin{aligned} \phi_{T-2}^{m,C}(\bar{x}) &\geq \min_{P_{T-1} \in \mathcal{P}_{T-1}} \sum_{i \in \mathcal{N}_{T-1}} p_{T-1}^i \hat{Q}_{T-1}^l(\bar{x}, \omega_{T-1}^i) \\ &= \min_{P_{T-1} \in \mathcal{P}_{T-1}} \sum_{i \in \mathcal{N}_{T-1}} p_{T-1}^i \mathcal{Q}_{T-1}^{RR}(\bar{x}, \omega_{T-1}^i) = \mathcal{Q}_{T-1}^{RR}(\bar{x}). \end{aligned} \quad (43)$$

The first equality holds since  $\phi_{T-1}^{m,C}(\bar{x}) = \mathcal{Q}_T^{RR}(\bar{x})$  for  $m > l_{T-1}$ . So, we have  $\phi_{T-2}^{m,C}(\bar{x}) = \mathcal{Q}_{T-1}^{RR}(\bar{x})$  for any future iteration  $m > l$ . Since  $|\bar{X}_{T-2}^{l_{T-1}}| < \infty$  and

every  $\xi \in \Xi$  has a positive probability, there exists a finite iteration  $l$  such that  $\phi_{T-2}^{l,C}(\bar{x}) = \mathcal{Q}_{T-1}^{RR}(\bar{x})$  for all  $\bar{x} \in \bar{X}_{T-2}^{l,T-1}$  with probability one, which implies  $l_{T-2} < \infty$  with probability one. Similarly, we can prove by induction that  $l_t$ , for all  $t \in [T-1]$ , are finite. This proves that  $|\mathcal{L}| < \infty$  with probability one.

Now, we show that each iteration terminates in finite time. This is followed by the facts that the subproblems are bounded and thus solvable in a finite time using a branch-and-cut algorithm and Line 9 is executed in a finite time by assumption.  $\square$

### C.5 Proof of Theorem 4

*Proof* By following a similar argument to the proof of Theorem 3, we can show that the number of the while-loop iterations required to define an optimal policy is finite with probability one and each while-loop iteration is executed in a finite time.  $\square$

### C.6 Proof of Theorem 5

*Proof* Let  $\{\bar{x}_t^l(\xi_{[t]}), \bar{y}_t^l(\xi_{[t]})\}_{t \in [T]}$  be a policy that is defined by the forward step at iteration  $l$ . To show its optimality, it suffices to show  $\phi_t^{l,S}(\bar{x}_t^l(\xi_{[t]})) = \mathcal{Q}_{t+1}^{RA}(\bar{x}_t^l(\xi_{[t]}))$  for  $t \in [T-1]$  and all  $\xi \in \Xi$ . We can prove the statement by following a similar argument to the proof of Theorem 3.  $\square$

### C.7 Proof of Proposition 2

*Proof* Let  $z_t = (x_t, y_t, \phi_t)$ . For any  $\bar{x}_{t-1} \in \{0, 1\}^{d_x}$ , stage  $t$ , and iteration  $l$ , the convex hull of  $\mathcal{D}_t^l(\bar{x}_{t-1}, \omega_t)$  is equivalent to the convex hull of  $\mathcal{F}_t^l(\omega_t) \cap \mathcal{E}(\bar{x}_{t-1})$  where

$$\mathcal{F}_t^l(\omega_t) := \left\{ (x_{t-1}, z_t) \in \mathbb{R}_+^{d_x} \times \mathbb{R}_+^{d_x+d_y+1} : \right. \\ \left. \bigvee_{h \in H_t} \left( \phi_t - (\pi_t^k)^\top x_t \geq \gamma_t^k, k \in [K_t^l], \right. \right. \\ \left. \left. A_t^h(\omega_t)x_t + B_t^h(\omega_t)y_t + C_t^h(\omega_t)x_{t-1} \geq b_t^h(\omega_t) \right) \right\}, \quad (44)$$

and  $\mathcal{E}(\bar{x}_{t-1}) := \{(x_{t-1}, z_t) : x_{t-1} = \bar{x}_{t-1}\}$ .

We claim that  $\text{conv}(\mathcal{F}_t^l(\omega_t) \cap \mathcal{E}(\bar{x}_{t-1})) = \text{conv}(\mathcal{F}_t^l(\omega_t)) \cap \mathcal{E}(\bar{x}_{t-1})$ . Clearly,  $\text{conv}(\mathcal{F}_t^l(\omega_t) \cap \mathcal{E}(\bar{x}_{t-1})) \subseteq \text{conv}(\mathcal{F}_t^l(\omega_t)) \cap \mathcal{E}(\bar{x}_{t-1})$ .

To show that  $\text{conv}(\mathcal{F}_t^l(\omega_t) \cap \mathcal{E}(\bar{x}_{t-1})) \supseteq \text{conv}(\mathcal{F}_t^l(\omega_t)) \cap \mathcal{E}(\bar{x}_{t-1})$ , pick any point  $(\bar{x}_{t-1}, \bar{z}_t) \in \text{conv}(\mathcal{F}_t^l(\omega_t)) \cap \mathcal{E}(\bar{x}_{t-1})$ . Then, there exist  $(x_{t-1}^j, z_t^j) \in \mathcal{F}_t^l(\omega_t)$  and  $\lambda^j \in (0, 1], j = 1, \dots, J$  such that  $\sum_{j \in [J]} \lambda^j = 1, \sum_{j \in [J]} \lambda^j x_{t-1}^j =$

$\bar{x}_{t-1}, \sum_{j \in [J]} \lambda^j z_t^j = \bar{z}_t$ . Since  $\bar{x}_{t-1}$  is binary and  $x_{t-1}^j$  belongs to  $[0, 1]^{d_x}$ , this implies that  $x_{t-1}^j = \bar{x}_{t-1}$  for all  $j \in [J]$ . Consequently,  $(x_{t-1}^j, z_{t-1}^j) \in \mathcal{F}_t^l(\omega_t) \cap \mathcal{E}(\bar{x}_{t-1}), j \in [J]$ , and  $(\bar{x}_{t-1}, \bar{z}_t) \in \text{conv}(\mathcal{F}_t^l(\omega_t) \cap \mathcal{E}(\bar{x}_{t-1}))$ . This completes the proof of the claim.

To obtain  $\text{conv}(\mathcal{F}_t^l(\omega_t))$ , we first use Theorem 2.1 in [1] and derive a tight extended formulation of  $\mathcal{F}_t^l(\omega_t)$ , which is given by

$$\widehat{\mathcal{F}}_t^l(\omega_t) := \left\{ \begin{array}{l} \sum_{h \in H_t} \zeta_{t,0}^h = 1, \sum_{h \in H_t} \zeta_{t,1}^h - x_t = 0, \sum_{h \in H_t} \zeta_{t,2}^h - y_t = 0, \\ \sum_{h \in H_t} \zeta_{t,3}^h - x_{t-1} = 0, \sum_{h \in H_t} \zeta_{t,4}^h - \phi_t = 0, \\ A_t^h(\omega_t) \zeta_{t,1}^h + B_t^h(\omega_t) \zeta_{t,2}^h + C_t^h(\omega_t) \zeta_{t,3}^h - b_t^h(\omega_t) \zeta_{t,0}^h \geq 0, \quad h \in H_t, \\ \zeta_{t,4}^h - (\pi_t^k)^\top \zeta_{t,1}^h - \gamma_t^k \zeta_{t,0}^h \geq 0, \quad h \in H_t, k \in K_t^l, \\ x_t \in \mathbb{R}_+^{d_x}, y_t \in \mathbb{R}_+^{d_y}, x_{t-1} \in \mathbb{R}_+^{d_x}, \phi_t \in \mathbb{R}_+, \\ \zeta_{t,0}^h \in \mathbb{R}_+, \zeta_{t,1}^h \in \mathbb{R}_+^{d_x}, \zeta_{t,2}^h \in \mathbb{R}_+^{d_y}, \zeta_{t,3}^h \in \mathbb{R}_+^{d_x}, \zeta_{t,4}^h \in \mathbb{R}_+, h \in H_t \end{array} \right\}. \quad (45)$$

Since  $\mathcal{F}_t^l(\omega_t)$  is unbounded, the projection of the above formulation (45) onto the  $(x_{t-1}, z_t)$ -space is the closed convex hull of  $\mathcal{F}_t^l(\omega_t)$ . Consider  $|H_t|$  polyhedra defined by disjunctive constraints for  $h \in H_t$  in (44). They are nonempty and have identical recession cones. This implies that  $\text{conv}(\mathcal{F}_t^l(\omega_t))$  is a polyhedron, and thus the closed convex hull of  $\mathcal{F}_t^l(\omega_t)$  is equivalent to  $\text{conv}(\mathcal{F}_t^l(\omega_t))$ . Hence, using the tight extended formulation (45), the convex hull of  $\mathcal{D}_t^l(\bar{x}_{t-1}, \omega_t)$  is given by  $\text{Proj}_{x_{t-1}, z_t}(\widehat{\mathcal{F}}_t^l(\omega_t) \cap \mathcal{E}(\bar{x}_{t-1}))$ , which is equivalent to the projection of the set  $\tilde{\mathcal{D}}_t^l(\bar{x}_{t-1}, \omega_t)$  onto the  $z_t$ -space.  $\square$

### C.8 Proof of Proposition 3

*Proof* For  $s \in [d_x]$ , the set  $\mathcal{D}_t^{l,s}(x_{t-1}, \omega_t), t \in [T]$ , is given by the following disjunctive constraints:

$$\bigvee_{(J_1, J_2) \in \mathcal{J}_t^s} \left( x_{t,j} = 0, j \in J_1, x_{t,j} = 1, j \in J_2, \phi_t - (\pi_t^k)^\top x_t \geq \gamma_t^k, k \in [K_t^l], \right. \\ \left. A_t(\omega_t)x_t + B_t(\omega_t)y_t \geq b_t(\omega_t) - C_t(\omega_t)x_{t-1} \right), \quad (46)$$

where  $\mathcal{J}_t^s = \{(J_1, J_2) : J_1, J_2 \subseteq [d_x], J_1 \cap J_2 = \emptyset, |J_1 \cup J_2| = s\}$ . By applying Proposition 2 to this disjunctive set, we obtain the tight extended formulation (31).  $\square$

## Appendix D Pseudocode of DA-SDDP-DP Algorithm

---

### Algorithm 2 DA-SDDP-DP

---

```

1: Initialize  $l \leftarrow 1$ ;  $x_0 \leftarrow$  initial state;  $\omega_1 \leftarrow$  data at the first stage;  $\Omega_1 := \{\omega_1\}$ ;  $K_t^l \leftarrow 0$ 
   for  $t = 1, \dots, T - 1$ ;
2: while (satisfying none of stopping conditions) do
3:   Sample a scenario path  $\xi^l \in \Xi := \Omega_1 \times \dots \times \Omega_T$ 
4:   for  $t \in [T]$  do ▷ Forward Step
5:     Solve  $t$ -stage LP-subproblem (27) given  $x_{t-1} = x_{t-1}^l$  and  $\omega_t = \xi_t^l$ 
6:   for  $t = T, \dots, 2$  do ▷ Backward Step
7:     for  $i \in \mathcal{N}_t$  do
8:       Solve  $t$ -stage LP-subproblem (27) given  $x_{t-1} = x_{t-1}^l$  and  $\omega_t = \omega_t^i$  and obtain
         cut  $(\sigma_{t-1,1}^{i,l}, \sigma_{t-1,0}^{i,l})$ 
9:       Add cuts  $(\pi_{t-1}^l, \gamma_{t-1}^l)$  to  $(t - 1)$ -stage LP-subproblem (27) by using cuts
          $(\sigma_{t-1,1}^{i,l}, \sigma_{t-1,0}^{i,l}), i \in \mathcal{N}_t$ 
10:       $K_{t-1}^l \leftarrow K_{t-1}^l + 1$ 
11:     Solve LP-subproblem (27) for  $t = 1$  to obtain the bound  $LB$ 
12:      $K_t^{l+1} \leftarrow K_t^l$  for  $t = 1, \dots, T - 1$ 
13:      $l \leftarrow l + 1$ 
14: return LP-subproblems,  $LB$ 

```

---

In this algorithm, let  $x_t^l$  be an optimal solution obtained by solving the LP-subproblem (29) during a forward step (Line 5) for iteration  $l$  and stage  $t$ . In a backward step (Line 8), the algorithm solves the LP-subproblem (29) given  $x_t^l$  and  $\omega_t^i$  to obtain a cut in the form of Benders cut,  $(\sigma_{t-1,1}^{i,l}, \sigma_{t-1,0}^{i,l})$ , where  $\sigma_{t-1,1}^{i,l}$  and  $\sigma_{t-1,0}^{i,l}$  are the optimal dual multipliers, associated with the constraints  $\sum_{h \in H_t} \zeta_{t,0}^h = 1$  and  $\sum_{h \in H_t} \zeta_{t,3}^h = x_{t-1}^l$ , respectively. These cuts  $(\sigma_{t-1,1}^{i,l}, \sigma_{t-1,0}^{i,l})$  for  $i \in \mathcal{N}_t$  are used to derive a cut  $(\pi_{t-1}^l, \gamma_{t-1}^l)$ ; the cut  $(\pi_{t-1}^l, \gamma_{t-1}^l)$  takes the form of (10), if we solve a DRR-MSDP, and it takes the form of (22), if we solve a DRO-MSDP. Then, its copies for  $h \in H_t$  are added to the  $(t - 1)$ -stage LP-subproblem (Line 9). In Line 11, it computes the lower bound by solving the first-stage LP-subproblem. DA-SDDP-DP repeats this procedure until a stopping condition is met. Both the DRR- and DRO-SDDP-DP algorithms, defined as Algorithm 2 with cut  $(\pi_{t-1}^l, \gamma_{t-1}^l)$  obtained as above for DRR-MSDPs and DRO-MSDPs, respectively, have the finite convergence that can be proved using Theorems 3 and 5, respectively.

*Remark 5* In the implementation of Algorithm 2, we can establish the LP-subproblems once and reuse them in each iteration, without the need for repeated construction, by adding constraints  $\zeta_{t,4}^h - (\pi_t^k)^\top \zeta_{t,1}^h - \gamma_t^k \zeta_{t,0}^h$  to the LP-subproblems as needed (in Line 9).