Two-stage and Lagrangian Dual Decision Rules for Multistage Adaptive Robust Optimization

Maryam Daryalal
Department of Decision Sciences, HEC Montréal, Montréal, Québec H3T 2A7, Canada,
maryam.daryalal@hec.ca

Ayşe N. Arslan
Univ. Bordeaux, CNRS, INRIA, Bordeaux INP, IMB, UMR 5251, F-33400 Talence, France,
Centre Inria de l'Universite de Bordeaux, F-33405 Talence, France,
ayse-nur.arslan@inria.fr

Merve Bodur
School of Mathematics, University of Edinburgh, Edinburgh EH9 3FD, UK,
merve.bodur@ed.ac.uk

In this work, we design primal and dual bounding methods for multistage adaptive robust optimization (MSARO) problems motivated by two decision rules rooted in the stochastic programming literature. From the primal perspective, this is achieved by applying decision rules that restrict the functional forms of only a certain subset of decision variables resulting in an approximation of MSARO as a two-stage adjustable robust optimization problem. We leverage the two-stage robust optimization literature in the solution of this approximation. From the dual perspective, decision rules are applied to the Lagrangian multipliers of a Lagrangian dual of MSARO, resulting in a two-stage stochastic optimization problem. As the quality of the resulting dual bound depends on the distribution chosen when developing the dual formulation, we define a distribution optimization problem with the aim of optimizing the obtained bound and develop solution methods tailored to the nature of the recourse variables. Our framework is general-purpose and does not require strong assumptions such as a stage-wise independent uncertainty set, and can consider integer recourse variables. Computational experiments on newsvendor, location-transportation, and capital budgeting problems show that our bounds yield considerably smaller optimality gaps compared to the existing methods.

Key words: Optimization under uncertainty, Robust optimization, Decision rules

1. Introduction

Many practical planning, design and operational problems involve making decisions under uncertainty at consecutive stages, where the decisions in one stage affect the decisions of the future stages. In such sequential decision-making problems, first-stage (*here-and-now*) decisions are the ones that are immediately implementable. Subsequent recourse (*wait-and-see*) decisions depend on the state of the system, which is a result of previous decisions and observations of the uncertain parameters. A solution is then an adaptable *policy* or *decision rule* that takes the previous decisions and history of uncertainty realizations as an input, and returns a new implementable decision. The dynamics of a sequential decision-making problem is depicted in Figure 1.

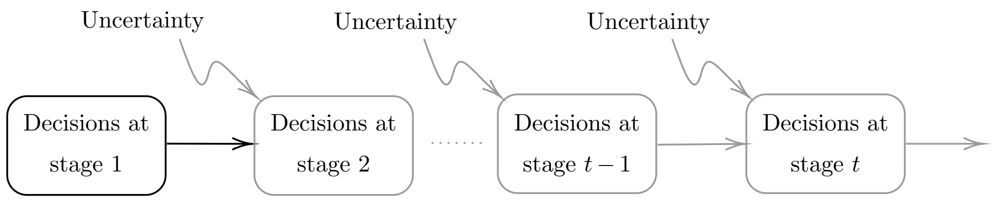

Figure 1 Sequential decision-making under uncertainty

There are several modeling frameworks for sequential decision-making problems under uncertainty. When the probability distribution governing the uncertain parameters is known, these problems may be addressed by the *multistage stochastic programming* (MSP) paradigm, with the goal of optimizing some statistical performance measure over the planning horizon. There is an extensive body of research on MSP problems of various structures, with a rich literature on problems with continuous decision variables. However, Bertsimas and Thiele (2006) point out that implemented solutions may perform poorly if the probability distribution used in the MSP model is different than the *true* distribution, even if both distributions share the same first and second moments. To mitigate this effect, *distributionally robust optimization* (DRO) models are proposed for making decisions that are based on a family of probability distributions, often defined by using historical data (Goh and Sim 2010). These models aim to hedge against tuning decisions to a perceived distribution. While the DRO framework has received significant attention (Mohajerin Esfahani and Kuhn (2018), Cheramin et al. (2022)) from the research community and some recent studies have proposed tractable solution methods for linear DRO problems under certain conditions (Philpott et al. 2018, Bertsimas et al. 2019), they remain largely challenging to solve especially in the multi-stage setting.

Multistage adaptive/adjustable robust optimization (MSARO) is another framework for modeling sequential decision-making problems under uncertainty that does not require any knowledge about the probability distribution governing the uncertain parameters. This framework is also adapted to contexts where the underlying uncertainty is not stochastic in nature, for instance, in the case of adversarial participants. In MSARO, the uncertainty is represented as belonging to a pre-structured (often compact) set, called the *uncertainty set*, and the decisions are optimized with respect to the worst-case outcome in this set.

In this paper, we focus on the MSARO framework. Throughout, we use $[a] := \{1, 2, \dots, a\}$ and $[a, b] := \{a, a+1, \dots, b\}$ for positive integers a and b (with $a \le b$), and $(\cdot)^\top$ for the transpose operator. For a problem with T decision-making stages, we denote with $\Xi^T \subseteq \mathbb{R}^{\ell^T}$ the uncertainty set which governs the set of uncertain parameters $(\xi_1, \xi_2, \dots, \xi_T)$, where ξ_t denotes the vector of parameters associated with stage t, and $\xi_1 = 1$, by convention. For ease of presentation, we also define the sequence of uncertain parameter vectors up to stage t along with their (projected) support as $\xi^t := (\xi_1, \dots, \xi_t) \in \Xi^t := \mathrm{proj}_{\xi^t}(\Xi^T) \subseteq \mathbb{R}^{\ell^t}$. We assume Ξ^t is compact for all $t \in [T]$. Then, we study the following general MSARO problem:

$$\min_{x_1 \in X_1(\xi_1)} c_1(\xi_1)^\top x_1 + \sup_{\xi_2 : (\xi^1, \xi_2) \in \Xi^2} \min_{\substack{x_2 \in X_2(\xi^2): \\ A_2(\xi^2) x_2 + B_2(\xi^2) x_1 \le b_2(\xi^2)}} c_2(\xi^2)^\top x_2 + \cdots \tag{1}$$

$$\cdots + \sup_{\xi_t : (\xi^{t-1}, \xi_t) \in \Xi^t} \min_{\substack{x_t \in X_t(\xi^t): \\ A_t(\xi^t) x_t + B_t(\xi^t) x_{t-1} \le b_t(\xi^t)}} c_t(\xi^t)^\top x_t + \cdots + \sup_{\xi_T : (\xi^{T-1}, \xi_T) \in \Xi^T} \min_{\substack{x_T \in X_T(\xi^T): \\ A_T(\xi^T) x_T + B_T(\xi^T) x_{T-1} \le b_T(\xi^T)}} c_T(\xi^T)^\top x_T$$

where $X_t(\xi^t) := \left\{ x_t \in \mathbb{R}^{n_t - n_t^{\mathrm{i}}} \times \mathbb{Z}^{n_t^{\mathrm{i}}} : \ D_t(\xi^t) x_t \le d_t(\xi^t) \right\}$ and $c_t : \mathbb{R}^{\ell^t} \to \mathbb{R}^{n_t}, b_t : \mathbb{R}^{\ell^t} \to \mathbb{R}^{m_t^{\mathrm{S}}}, d_t : \mathbb{R}^{\ell^t} \to \mathbb{R}^{m_t^{\mathrm{R}}}, A_t : \mathbb{R}^{\ell^t} \to \mathbb{R}^{m_t^{\mathrm{S}} \times n_t}, B_t : \mathbb{R}^{\ell^t} \to \mathbb{R}^{m_t^{\mathrm{S}} \times n_{t-1}}, D_t : \mathbb{R}^{\ell^t} \to \mathbb{R}^{m_t^{\mathrm{R}} \times n_t}$ for $t \in [T]$. The main output of model (1) is the first-stage deterministic (here-and-now) decisions x_1 which minimize the worst-case objective value over T stages taking into account the sequential uncertainty realizations and optimal wait-and-see decisions. In the sequential framework, at each stage $t \in T$, the worst-case realization ξ_t that is consistent with the history of the realizations up to stage $t-1$, ξ^{t-1}, is revealed. The vector ξ_t combined with the history ξ^{t-1} yields the history of realizations up to stage t, that is, $\xi^t = (\xi^{t-1}, \xi_t)$ from the support Ξ^t. This determines the parameters of the stage-t minimization problem, from which the optimal wait-and-see decision vector x_t is obtained. As such, the wait-and-see decisions, $x_2, \dots, x_T$, also known as recourse decisions, are adapted to the history of the uncertain parameter realizations up to their decision-making stage, $\xi^2, \dots, \xi^T$, respectively. We remark that, by definitions of X_t and Ξ^t, we allow for the possibility of mixed-integer wait-and-see decisions and dependence between the uncertain parameters of different stages. In the following, we assume, for the data, that all uncertain vectors and matrices are affine functions of the associated uncertain parameters ξ^t, as the majority of the literature mentioned makes this assumption; further assumptions will be specified explicitly when necessary. When the set $X_t(\xi^t), t \in [2, T]$ does not have integrality restrictions ($n_t^{\mathrm{i}} = 0$), MSARO problem has continuous recourse, otherwise, it has (mixed-)integer recourse.

MSARO problems are highly challenging to solve. Indeed, as has been recently proven by Goerigk et al. (2024), they are, in general, harder than NP-hard problems, lying at the higher levels of the polynomial hierarchy, and their complexity increases with the number of decision stages. Specifically, T-stage MSARO problems with certain uncertainty set structures are $\Sigma^{\mathrm{p}}_{2T-1}$-hard. However, some sub-classes and special cases of MSARO problems are theoretically and/or computationally more tractable, depending on the number of stages, the structure of the uncertainty set, and the nature of recourse decisions. Among these, the most well-known is *static* robust optimization, which considers that all decisions are here-and-now. For a considerable number of problem structures, e.g., when the uncertainty set is a polyhedron or an ellipsoid, it is possible to derive a monolithic reformulation of the static robust optimization problem through a compact reformulation of the adversarial problem, usually relying on duality techniques (Ben-Tal and Nemirovski 1999, Ben-Tal et al. 2009, Bertsimas et al. 2011a, 2015). However, this paradigm cannot capture the flexibility offered by the possibility of adapting some decisions to the realization of uncertainty, thus often producing overly conservative decisions. As such, there has been significant research effort on developing exact and approximate solution methods for two-stage adjustable robust optimization (2ARO) problems, i.e., MSARO with $T = 2$ (Ben-Tal et al. 2004, Zeng and Zhao 2013, Postek and Hertog 2016, Subramanyam et al. 2020). In developing these methods, the presence of discrete recourse variables, known as (mixed-)integer recourse, poses additional challenges in ensuring exact or high-quality solutions within a reasonable computational effort compared to the continuous recourse case.

On the other hand, scientific progress on general MSARO has been much more limited. Given the aforementioned theoretical complexity of these problems, the focus of most existing studies is the approximate solution of these problems. Approximations proposed for MSARO mostly rely on reducing the multi-stage problem to a static problem, with a view to leverage the tractability of these problems. While these approximations can produce feasible solutions for MSARO problems, they can lead to a significant degradation in solution quality. Furthermore, some of these methods are quite restrictive being only applicable to special classes such as MSARO with continuous recourse.

To address these limitations, this paper aims to develop approximations for general MSARO problems of form (1). More specifically, inspired by the recent developments in the MSP literature, namely the two-stage linear decision rules (Bodur and Luedtke 2018), we propose applying decision rule approximations to only a certain subset of decision variables, resulting in an approximation of MSARO problems as 2ARO problems. In so doing we have three motivations: (i) the strength of the added flexibility in adapting recourse decisions to uncertainty in 2ARO compared to static robust optimization, (ii) the significantly reduced theoretical complexity of 2ARO compared to MSARO, and (iii) the progress made in computationally viable exact and approximate solution methodologies for 2ARO. As a result of point (iii), our proposed framework is capable of considering a large variety of MSARO classes, most notably the mixed-integer recourse case, and will directly benefit from the developments in the highly active 2ARO literature in the future, e.g., the incorporation of machine learning for computational enhancements (Julien et al. 2022, Dumouchelle et al. 2023).

While the aforementioned ideas are aimed at providing feasible solutions for MSARO problems, an important question arises as to the quality of the obtained solution. In order to evaluate the quality of a feasible policy, one could use a dual bound on the optimal value of the MSARO problem. Unfortunately, obtaining dual bounds for MSARO problems is a largely unexplored topic in the literature, especially in the case of mixed-integer recourse. To fill this gap, we propose to develop dual approximations for MSARO. In particular, we propose a Lagrangian dual for the MSARO problem and apply decision rules to the Lagrangian multipliers, leveraging ideas rooted in the MSP literature, namely, Lagrangian dual decision rules (Daryalal et al. 2024). In deriving a Lagrangian dual of the MSARO problem, we assign a probability distribution with the support as the uncertainty set and use the assigned distribution in dualizing a subset of constraints along with their Lagrangian multipliers. As a result, we obtain a dual approximation of MSARO in the form of a two-stage stochastic optimization problem, which can be solved with the help of state-of-the-art methods for two-stage stochastic problems. Since the quality of the resulting dual bound depends on the assigned distribution used while developing the dual formulation, we define a distribution optimization problem with the aim of identifying the strongest such dual bound. We develop appropriate solution methods tailored to the nature of the recourse variables for the resulting distribution optimization problem.

The contributions of our work are summarized as follows:

- We develop a solution framework for MSARO problems that returns adaptable policies as well as a dual bound measuring the quality of these policies. We do this by employing two-stage and Lagrangian dual decision rules, leading to novel techniques that reveal new theoretical and practical avenues.
- We propose to approximate MSARO problems via 2ARO problems to leverage existing solution methodologies and future developments for the latter in designing adaptable policies for the former. To this end, we present two-stage decision rules for MSARO, the first adaptation of a generalization of two-stage linear decision rules from the MSP literature in robust optimization, which can be applied to a broad range of problems. We employ, for an illustration of our approach, a tailored constraint-and-column generation algorithm to solve the resulting 2ARO approximation. The optimal solution of this approximation not only provides a primal policy but can also contribute to the calculation of a dual bound through identification of critical realizations in the uncertainty set.
- With a similar motivation, we derive a dual approximation of MSARO in the form of a two-stage stochastic optimization problem, which we show to be a strong dual in certain cases.

Moreover, in order to obtain the strongest possible such dual bound, we study the numerical solution of a distribution optimization problem. More specifically, we apply decision rules to dual variables, and design a cutting-plane algorithm to solve the obtained restricted dual problem. We also show that in the special case of continuous recourse, the restricted dual problem can be reformulated as a monolithic bilinear program. Additionally, we present an alternative decomposable dual problem which offers an improved numerical performance. These novel techniques contribute to the scarce literature for obtaining dual bounds for MSARO problems with mixed-integer recourse.

- We evaluate the performance of our solution framework over multistage versions of three classical problems from the MSARO literature: (*i*) the newsvendor problem, (*ii*) the location-transportation problem, and (*iii*) the capital budgeting problem. Each of these problem classes is suitable for a different solution method developed in this work and our analysis over various instances attests to the quality of the returned primal and dual bounds.

The remainder of the paper is organized as follows. In Section 2 we review the literature relevant to our work. In Section 3 we introduce two-stage decision rules for obtaining primal adaptable policies. In Section 4 we present our approach to deriving dual bounds. This is followed by numerical experiments in Section 5 and concluding remarks. We remark that all proofs are deferred to Appendix B.

2. Literature Review

Figure 2 presents a summary of existing solution methods for obtaining exact/approximate solutions and dual bounds for MSARO, with methods developed specifically for 2ARO separately categorized. In the following, we briefly discuss each method and the specific problem structure it can address.

Exact solution methods are scarce in the MSARO literature and the existing studies mostly focus on 2ARO. For 2ARO problems with fixed recourse and finite or polyhedral uncertainty set, Zeng and Zhao (2013) developed a constraint-and-column generation algorithm. For the same type of problems restricted to continuous recourse, Bertsimas et al. (2012) designed a Benders decomposition-type algorithm and applied it to a unit commitment problem, whereas Georghiou et al. (2020) proposed a convergent method based on enumeration of the extreme points of the uncertainty set combined with affine decision rules to provide gradually improving primal and dual bounds. For 2ARO problems with continuous fixed recourse, Zhen et al. (2018) used Fourier-Motzkin elimination iteratively to remove the second-stage decisions, eventually forming an equivalent static robust optimization problem. This computationally expensive approach is also extended to multi-stage problems. In the case of mixed-binary recourse and only objective uncertainty, Arslan and

Figure 2 Solution methods for MSARO

Detienne (2022) proposed an exact method based on a Dantzig-Wolfe reformulation of the recourse problem based on a technical assumption on the structure of the linking constraints. Similarly, using Dantzig-Wolfe reformulation, for a subclass of 2ARO problems with fixed and mixed-integer recourse, block diagonal recourse matrix and a finite uncertainty set, Hashemi Doulabi et al. (2021) derived a static formulation which is amenable to Benders decomposition. For continuous MSARO problems with a stage-wise rectangular uncertainty set, Georghiou et al. (2019) developed robust dual dynamic programming (RDDP) and proved finite/asymptomatic convergence for various problem sub-classes. RDDP is an adaptation of the stochastic dual dynamic programming algorithm from the MSP literature (Pereira and Pinto 1991) to MSARO.

Approximate solution methods are more common in the MSARO literature, with the central idea of restricting adaptable/adjustable decisions to follow a certain functional form, known as *decision rules*. Ben-Tal et al. (2004) proposed the first decision rule for MSARO problems with *continuous recourse*, LDRs, where recourse decisions are expressed as affine functions of uncertain parameters where the parameters of this function are to be optimized. The resulting LDR-restricted problem being a static optimization problem, it can be reformulated as a linear optimization problem in certain cases. Nonlinear decision rules were also proposed, such as deflected and segregated affine (Chen et al. 2008), extended affine (Chen and Zhang 2009), piecewise affine (Goh and Sim 2010), truncated linear (See and Sim 2010), piecewise affine with exponentially many pieces (Ben-Tal et al. 2020), quadratic (Xu and Hanasusanto 2021) and polynomial (Bertsimas et al. 2011b) decision rules. However, the resulting reformulations when using non-linear decision rules are often nonlinear, e.g, semidefinite or copositive programs. For a comprehensive list of nonlinear decision rules, interested reader may refer to the survey by Yanıkoğlu et al. (2019).

In the case of *mixed-integer recourse*, LDRs and most of its aforementioned extensions lead to non-adjustable decisions for the integer variables. Thus, alternative approaches have been proposed, with the key idea of (implicitly or explicitly) partitioning the uncertainty set and determining a constant recourse solution corresponding to each subset. A popular approach for 2ARO problems uses the notion of finite adaptability, first introduced by Bertsimas and Caramanis (2010). In finite or K-adaptability, the decision-maker a priori commits to K recourse decisions (while making the first-stage decisions), and then chooses among them after observing the uncertainty realization which leads to an implicit K-partition of the uncertainty set. While Bertsimas and Caramanis (2010) presented an exact formulation for the 2-adaptability problem, for the general K-adaptability case, Hanasusanto et al. (2015) proposed a monolithic formulation for problems with binary recourse, and Subramanyam et al. (2020) developed a branch-and-bound algorithm for problems with mixed-integer recourse. For MSARO problems on the other hand, explicit uncertainty set partitioning is considered in an iterative heuristic framework, with the aim of obtaining a sequence of improving approximations (Bertsimas and Dunning 2016, Postek and Hertog 2016, Romeijnders and Postek 2020). Lastly, for MSARO problems with pure-binary recourse, Bertsimas and Georghiou (2018) introduced binary decision rules, whereas for the mixed-binary recourse case, Bertsimas and Georghiou (2015) implicitly designed piecewise linear/constant decision rules.

While the aforementioned primal approximations can be shown to be exact in some special cases (Bertsimas et al. 2010, Bertsimas and Goyal 2012, Iancu et al. 2013, Hanasusanto et al. 2015, Zhen et al. 2018), in general they do not provide optimal solutions. In order to assess the quality of their feasible solutions, dual bounds can be used. To this end, Kuhn et al. (2011) presented the idea of deriving a dual problem for MSARO with only continuous variables and applying LDRs on the dual variables. Since their approach was originally derived for stochastic programs its application to MSAROs requires assigning a probability distribution to the uncertainty set. The impact of the chosen distribution on the quality of the obtained dual bound was observed by Kuhn et al. (2011), as such a distribution optimization problem was mentioned. This problem was later formalized by Hadjiyiannis et al. (2011) for a 2ARO with continuous variables and shown to be of the same theoretical difficulty as the original problem. Hadjiyiannis et al. (2011) proposed to solve instead a 2ARO problem for a finite set of scenarios from the uncertainty set, selected based on a primal decision rule restriction, to reach a dual bound. This procedure can also be extended to obtain dual bounds for general MSAROs. Finally, for special cases of 2ARO problems with continuous recourse, Georghiou et al. (2020) proposed a framework to derive progressive dual bounds, by considering the linear programming dual of their primal extreme point reformulation.

3. Primal Bounding

The MSARO problem given in a nested form in (1) can be reformulated as a monolithic optimization problem by explicitly introducing the functional form of the decision variables, $x_t(\xi^t): \Xi^t \to \mathbb{R}^{n_t - n_t^{\mathrm{i}}} \times \mathbb{Z}^{n_t^{\mathrm{i}}}$ for all $t \in [T]$, along with a deterministic variable, z, representing the worst-case objective value:

$$
\begin{aligned}
\nu^\star := \min \quad & z && && \text{(2a)}\\
\text{s.t.} \quad & \sum_{t\in[T]} c_t(\xi^t)^\top x_t(\xi^t) \le z && \xi^T \in \Xi^T && \text{(2b)}\\
& A_t(\xi^t) x_t(\xi^t) + B_t(\xi^t) x_{t-1}(\xi^{t-1}) \le b_t(\xi^t) && t \in [2,T],\ \xi^t \in \Xi^t && \text{(2c)}\\
& x_t(\xi^t) \in X_t(\xi^t) && t \in [T],\ \xi^t \in \Xi^t. && \text{(2d)}
\end{aligned}
$$

Together with constraints (2b), the objective function (2a) minimizes the worst outcome. Constraints (2c) and (2d) are *state* and *recourse* constraints, respectively: while the former link different stages, the latter are local restrictions for a specific stage.

Throughout the paper, we make the following assumptions:

ASSUMPTION 1. *$X_1(\xi^1)$ is non-empty.*

ASSUMPTION 2. *The problem has relatively complete recourse, i.e., for all $t \in [2,T]$, $\xi^t \in \Xi^t$ and a history of feasible decisions made up to t, $\{x_{t'}(\xi^{t'})\}_{t'\in[t-1]}$, there always exists a feasible decision at stage t, $x_t(\xi^t)$.*

ASSUMPTION 3. *For $t \in [T]$ and $\xi^t \in \Xi^t$, the maximum diameter of $X_t(\xi^t)$ is finite, i.e., the feasibility sets are bounded.*

We note that combined with the compactness assumption of the uncertainty set and its projections, these assumptions imply that the studied MSARO problem has a finite optimal objective value.

To derive feasible policies to the MSARO problem, it is quite common in the literature to restrict *all* the decisions $x_t(\xi^t)$ to follow a simple functional form, such as an affine or piecewise constant decision rule. By breaking the temporal dependencies between stages, this approach approximates problem (2) with a static robust optimization problem. Our goal in this section is to employ a new paradigm where a specific *subset* of the decision variables $x_t(\xi^t)$ are enforced to follow a structured decision rule, leading to a restriction in the form of a 2ARO problem. We introduce this approach, two-stage decision rules for MSARO, in Section 3.1, then present its specific instantiation and possible solution methodologies in Sections 3.2 and 3.3.

3.1. Two-stage Decision Rules

In a similar manner to constraints (2c) and (2d), we partition the decision variables $x_t(\xi^t), t \in [T]$ into $x_t^{\mathsf{s}}(\xi^t) \in \mathbb{R}^{q_t}$ and $x_t^{\mathsf{r}}(\xi^t) \in \mathbb{R}^{p_t}$, *state* and *recourse* variables, as those that appear in the state constraints of subsequent stages and the others, respectively. We have $q_t + p_t = n_t$, with integrality restrictions on the variables, if any, embedded in the set $X_t(\xi^t)$. Then the MSARO (2) can be written more explicitly as follows:

$$\nu^\star = \min z \tag{3a}$$

$$\text{s.t.} \sum_{t\in[T]} c_t^{\mathsf{s}}(\xi^t)^\top x_t^{\mathsf{s}}(\xi^t) + c_t^{\mathsf{r}}(\xi^t)^\top x_t^{\mathsf{r}}(\xi^t) \le z \qquad \xi^T \in \Xi^T \tag{3b}$$

$$A_t^{\mathsf{s}}(\xi^t) x_t^{\mathsf{s}}(\xi^t) + B_t^{\mathsf{s}}(\xi^t) x_{t-1}^{\mathsf{s}}(\xi^{t-1}) + A_t^{\mathsf{r}}(\xi^t) x_t^{\mathsf{r}}(\xi^t) \le b_t(\xi^t) \qquad t \in [2,T],\ \xi^t \in \Xi^t \tag{3c}$$

$$(x_t^{\mathsf{s}}(\xi^t), x_t^{\mathsf{r}}(\xi^t)) \in X_t(\xi^t) \qquad t \in [T],\ \xi^t \in \Xi^t \tag{3d}$$

where $c_t^{\mathsf{s}}(\xi^t), A_t^{\mathsf{s}}(\xi^t), B_t^{\mathsf{s}}(\xi^t), D_t^{\mathsf{s}}(\xi^t)$ are sub-arrays/sub-matrices of $c_t(\xi^t), A_t(\xi^t), B_t(\xi^t), D_t(\xi^t)$ associated with the state variables with appropriate dimensions, while $c_t^{\mathsf{r}}(\xi^t), A_t^{\mathsf{r}}(\xi^t), D_t^{\mathsf{r}}(\xi^t)$ have the same role for the recourse variables, and $X_t(\xi^t) = \big\{x_t^{\mathsf{s}} \in \mathbb{R}^{q_t - q_t^{\mathsf{i}}} \times \mathbb{Z}^{q_t^{\mathsf{i}}}, x_t^{\mathsf{r}} \in \mathbb{R}^{p_t - p_t^{\mathsf{i}}} \times \mathbb{Z}^{p_t^{\mathsf{i}}} :\ D_t^{\mathsf{s}}(\xi^t) x_t^{\mathsf{s}} + D_t^{\mathsf{r}}(\xi^t) x_t^{\mathsf{r}} \le d_t(\xi^t)\big\}$. For notational convenience, we drop the parametrization for the first-stage variables as well as their feasible set and the objective vector, *i.e.*, use $x_1 = (x_1^{\mathsf{s}}, x_1^{\mathsf{r}}), X_1$, and $c_1 = (c_1^{\mathsf{s}}, c_1^{\mathsf{r}})$.

For $t \in [2,T]$, let $x_t^{\mathsf{s}}(\xi^t)$ be approximated by a decision rule, i.e., $x_t^{\mathsf{s}}(\xi^t) = \Theta_t(\xi^t, \beta_t)$, where $\Theta_t : \mathbb{R}^{\ell^t} \times \mathbb{R}^{K_t} \to \mathbb{R}^{q_t}$ represents the rule, and $\beta_t \in \mathbb{R}^{K_t}$ represents its vector of design parameters. By substituting this rule in problem (3), we obtain an approximation that can be reformulated as:

$$\nu^{2\mathrm{S}} := \min\ c_1^\top x_1 + \max_{\xi^T \in \Xi^T} \min_{x^{\mathsf{r}} \in \mathcal{X}(x_1^{\mathsf{s}}, \beta, \xi^T)} \sum_{t\in[2,T]} c_t^{\mathsf{s}}(\xi^t)^\top \Theta_t(\xi^t, \beta_t) + c_t^{\mathsf{r}}(\xi^t)^\top x_t^{\mathsf{r}} \tag{4a}$$

$$\text{s.t.} \quad x_1 \in X_1 \tag{4b}$$

$$\beta_t \in \mathbb{R}^{K_t} \qquad t \in [2,T], \tag{4c}$$

where:

$$\begin{aligned}
\mathcal{X}(x_1^{\mathsf{s}}, \beta, \xi^T) := \Big\{ \big(x_t^{\mathsf{r}}\big)_{t\in[2,T]} \in \mathbb{R}^{p_2} \times \mathbb{R}^{p_3} \times \cdots \times \mathbb{R}^{p_T} :& \\
A_t^{\mathsf{r}}(\xi^t) x_t^{\mathsf{r}} \le b_t(\xi^t) - \Big(A_t^{\mathsf{s}}(\xi^t)\Theta_t(\xi^t, \beta_t) + B_t^{\mathsf{s}}(\xi^t) x_1^{\mathsf{s}}\Big) & \qquad t = 2 \\
A_t^{\mathsf{r}}(\xi^t) x_t^{\mathsf{r}} \le b_t(\xi^t) - \Big(A_t^{\mathsf{s}}(\xi^t)\Theta_t(\xi^t, \beta_t) + B_t^{\mathsf{s}}(\xi^t)\Theta_{t-1}(\xi^{t-1}, \beta_{t-1})\Big) & \qquad t \in [3,T] \\
(\Theta_t(\xi^t, \beta_t), x_t^{\mathsf{r}}) \in X_t(\xi^t) & \qquad t \in [2,T] \Big\}.
\end{aligned}$$

Note that the decision rules are solely applied to the state variables, whereas the recourse variables remain fully adjustable to the uncertain parameters (see Figure 3). Problem (4) is a 2ARO since the temporal dependency between stages is removed thanks to the application of the two-stage

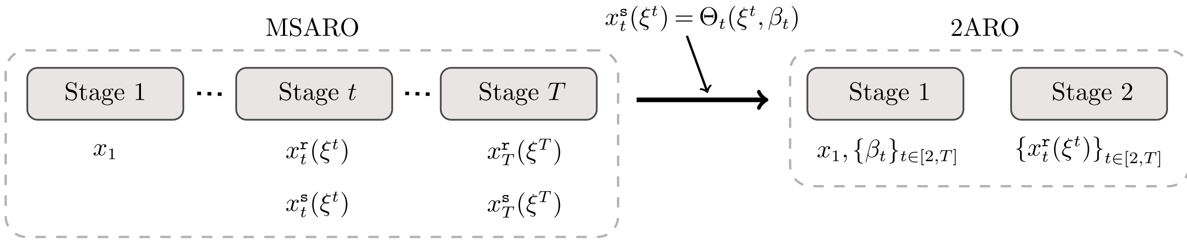

Figure 3 Two-stage decision rules

decision rules (see Appendix A for a detailed proof). We remark that the relatively complete recourse assumption stated in Assumption 2 for the MSARO problem, does not guarantee that the 2ARO model has relatively complete recourse, but it can be ensured, for instance, by following the techniques mentioned in (Bodur and Luedtke 2018).

A practical result of such an approximation is that the resulting problem can be solved using the existing solution methods for 2ARO. In the following sections, we present two possible choices for the decision rule $\Theta_t(\xi^t, \beta_t)$, respectively applicable to continuous and integer state variables. Together, they permit the approximation of an MSARO problem with mixed-integer state variables via a 2ARO model. We remark that the nature of the recourse variables does not impact the reduction of an MSARO to a 2ARO, but plays an important role in the choice of an appropriate solution method for the resulting 2ARO model. In what follows, we illustrate the application of these decision rules and the algorithmic solution of ensuing 2ARO models. For ease of exposition, we present MSARO problems with only continuous and only integer state variables separately.

3.2. Two-stage Linear Decision Rules for MSAROs with Continuous State Variables

If the state variables are continuous, we can approximate them via a decision rule with an affine form. For $t \in [2, T]$, letting $\Phi_t(\xi^t) = \big(\Phi_{t1}(\xi^t), \ldots, \Phi_{tK_t}(\xi^t)\big) : \mathbb{R}^{\ell^t} \to \mathbb{R}^{q_t \times K_t}$ be a vector of chosen basis functions, the *two-stage LDR* is enforced by using

$$\Theta_t(\xi^t, \beta_t) = \Phi_t(\xi^t)\beta_t \tag{5}$$

in (4) where we use a compact matrix representation[1] for notational convenience. The resulting 2ARO is written as:

$$\nu^{\text{2S-LDR}} := \min \ c_1^\top x_1 + \max_{\xi^T \in \Xi^T} \min_{x^{\mathtt{r}} \in \mathcal{X}(x_1^{\mathtt{s}}, \beta, \xi^T)} \sum_{t\in[2,T]} c_t^{\mathtt{s}}(\xi^t)^\top \Phi_t(\xi^t)\beta_t + c_t^{\mathtt{r}}(\xi^t)^\top x_t^{\mathtt{r}} \tag{6a}$$

[1] This representation is obtained, without loss of generality, by concatenating individual LDR restrictions applied to each state variable $x_{ti}^{\mathtt{s}}(\xi^t)$ for $i \in [q_t]$. For example, consider an instance where there are two state variables at stage $t = 2$ and the history consists of two components $\xi^2 = (\xi_1, \xi_2) \in \mathbb{R}^2$ (with the convention that $\xi_1 = 1$). Then the decisions rules $x_{21}(\xi^2) = \beta_{21}^1\xi_1 + \beta_{22}^1\xi_2$ and $x_{22}(\xi^2) = \beta_{21}^2\xi_1 + \beta_{22}^2\xi_2$ can be represented in the more compact matrix form with $K_2 = 4$ using the concatenated decision vector $\beta \in \mathbb{R}^4$ and the basis function matrix $\Phi_2(\xi^2) = \begin{bmatrix} \xi_1 & \xi_2 & 0 & 0 \\ 0 & 0 & \xi_1 & \xi_2 \end{bmatrix}$.

$$\text{s.t.} \quad x_1 \in X_1 \tag{6b}$$

$$\beta_t \in \mathbb{R}^{K_t} \qquad t \in [2,T]. \tag{6c}$$

As reviewed in Section 2, the 2ARO problem (6) can either be solved by means of approximation (e.g., K-adaptability (Subramanyam et al. 2020), uncertainty set partitioning (Postek and Hertog 2016, Bertsimas and Dunning 2016), Neur2RO (Dumouchelle et al. 2023)) or exactly, most notably via the commonly used constraint-and-column generation (C&CG) method, initially proposed by Zeng and Zhao (2013), which we detail next.

The C&CG method draws on the fact that not all realizations in Ξ^T contribute to the worst-case objective value. It then strives to identify *necessary realizations* by starting from a smaller uncertainty set and gradually expanding it. This leads to the generation of new columns and constraints, respectively corresponding to recourse variables and second-stage constraints for the newly identified realization. More specifically, consider a relaxation of problem (6) where instead of the uncertainty set Ξ^T, a potentially empty subset $\hat{\Xi} \subseteq \Xi^T$ is used to obtain the following master problem, which we denote by $\mathcal{MP}(\hat{\Xi})$:

$$\min \quad c_1^\top x_1 + \eta \tag{7a}$$

$$\text{s.t.} \quad \eta \geq \sum_{t\in[2,T]} \left(c_t^{\mathtt{s}}(\xi^t)^\top \Phi_t(\xi^t)\beta_t + c_t^{\mathtt{r}}(\xi^t)^\top x_{t,\xi^T}^{\mathtt{r}} \right) \qquad \xi^T \in \hat{\Xi}^{\text{Feas}} \tag{7b}$$

$$A_t^{\mathtt{r}}(\xi^t) x_{t,\xi^T}^{\mathtt{r}} + A_t^{\mathtt{s}}(\xi^t)\Phi_t(\xi^t)\beta_t + B_t^{\mathtt{s}}(\xi^t) x_1^{\mathtt{s}} \leq b_t(\xi^t) \qquad t = 2, \xi^T \in \hat{\Xi} \tag{7c}$$

$$A_t^{\mathtt{r}}(\xi^t) x_{t,\xi^T}^{\mathtt{r}} + A_t^{\mathtt{s}}(\xi^t)\Phi_t(\xi^t)\beta_t + B_t^{\mathtt{s}}(\xi^t)\Phi_{t-1}(\xi^{t-1})\beta_{t-1} \leq b_t(\xi^t) \qquad t \in [3,T], \xi^T \in \hat{\Xi} \tag{7d}$$

$$(\Phi_t(\xi^t)\beta_t, x_{t,\xi^T}^{\mathtt{r}}) \in X_t(\xi^t) \qquad t \in [2,T], \xi^T \in \hat{\Xi} \tag{7e}$$

$$x_1 \in X_1, \quad \eta \in \mathbb{R} \tag{7f}$$

$$\beta_t \in \mathbb{R}^{K_t} \qquad t \in [2,T], \tag{7g}$$

where $\hat{\Xi}^{\text{Feas}} \subseteq \hat{\Xi}$ is the set of identified necessary realizations ξ^T for which there exists a feasible first-stage solution $\hat{\beta}$ such that the feasibility space $\mathcal{X}(x_1^{\mathtt{s}}, \hat{\beta}, \xi^T)$ is nonempty. If the recourse variables are all continuous (i.e., $p_t^{\mathtt{i}} = 0, t \in [2,T]$), then (7) is a linear program, otherwise it is a mixed-integer linear program. We remark that if $\hat{\Xi}^{\text{Feas}}$ is empty at initialization, a valid lower bound on the η variable can be added to the model to avoid unboundedness. If $\mathcal{MP}(\hat{\Xi})$ is infeasible for any $\hat{\Xi} \subseteq \Xi^T$, the 2ARO model (6) is proven to be infeasible. Next, we consider the cases where $\mathcal{MP}(\hat{\Xi})$ is feasible and bounded.

The optimal objective value of $\mathcal{MP}(\hat{\Xi})$ is a lower bound on $\nu^{\text{2S-LDR}}$, the optimal objective value of (6). To obtain an exact solution to problem (6), we may need to gradually expand $\hat{\Xi}$ with necessary realizations (and consequently the set of recourse variable copies $x_{t,\xi^T}^{\mathtt{r}}$). At convergence, solving $\mathcal{MP}(\hat{\Xi})$ should return an optimal solution $(\hat{x}_1, \hat{\beta}, \hat{\eta})$ such that $\hat{\eta}$ accurately measures the worst-case second-stage cost over the complete uncertainty set Ξ^T (or conclude infeasibility of (6)). To check whether this convergence criterion is satisfied, we solve the adversarial problem for a given first-stage solution $(\hat{\beta}, \hat{x}_1^{\mathtt{s}})$, which results in the following subproblem:

$$\mathcal{SP}(\hat{\beta}, \hat{x}_1^{\mathtt{s}}) := \max_{\xi^T \in \Xi^T} \left\{ \sum_{t \in [2,T]} c_t^{\mathtt{s}}(\xi^t)^\top \Phi_t(\xi^t) \hat{\beta}_t + \min_{x^{\mathtt{r}} \in \mathcal{X}(\hat{x}_1^{\mathtt{s}}, \hat{\beta}, \xi^T)} \sum_{t \in [2,T]} c_t^{\mathtt{r}}(\xi^t)^\top x_t^{\mathtt{r}} \right\}. \tag{8}$$

If $\hat{\eta} = \mathcal{SP}(\hat{\beta}, \hat{x}_1^{\mathtt{s}})$, then $\hat{\eta}$ exactly measures the worst-case cost of the second-stage problem and $(\hat{x}_1, \hat{\beta}, \hat{\eta})$ is an optimal solution to problem (6), i.e., $\nu^{\text{2S-LDR}} = c_1^\top \hat{x}_1 + \hat{\eta}$. Otherwise, let $\hat{\xi}^T$ be the optimal solution of subproblem (8) if it is feasible, or a scenario such that $\mathcal{X}(\hat{x}_1^{\mathtt{s}}, \hat{\beta}, \hat{\xi}^T)$ is an empty set. We create new variables $x_{t,\hat{\xi}^T}^{\mathtt{r}}, t \in [2,T]$ and update the set of necessary realizations with $\hat{\Xi} = \hat{\Xi} \cup \{\hat{\xi}^T\}$, and accordingly update (7c)-(7e). If subproblem (8) is feasible, then we update $\hat{\Xi}^{\text{Feas}} = \hat{\Xi}^{\text{Feas}} \cup \{\hat{\xi}^T\}$ and (7b) as well. In this case, constraints (7b)-(7e) make up the optimality cuts. If subproblem (8) is infeasible, then (7c)-(7e) act as feasibility cuts. We repeat solving the master problem and the subproblem in a cutting-plane fashion and add the appropriate cuts until $\hat{\Xi}$ includes all the necessary realizations and $\hat{\eta} = \mathcal{SP}(\hat{\beta}, \hat{x}_1^{\mathtt{s}})$, or $\mathcal{MP}(\hat{\Xi})$ becomes infeasible.

In general, subproblem (8) can be numerically challenging to solve especially if the inner minimization problem contains integer variables and/or non-linearities in ξ^T. However, there are certain practical cases in which (8) can be reformulated as a mixed-integer linear program and then directly be given to an optimization solver. In the context of our subproblem, one notable example of this is provided in Remark 1. Further, extending this case to allow integer recourse variables ($p_t^{\mathtt{i}} \neq 0, t \in [2,T]$), while keeping the other assumptions in Remark 1, Zhao and Zeng (2012) proposed a nested constraint-and-column generation algorithm. In addition, we note that it is not necessary to solve the subproblem exactly at each iteration; the solution process can be stopped as soon as a violated cut for the master problem is identified. Recent studies have also considered the use of neural networks in order to approximate the inner minimization problem (Dumouchelle et al. 2023), especially to handle integer recourse variables and non-linearities.

Remark 1. Consider an MSARO where the uncertainty set is a polytope, the basis functions $\Phi_t(\xi^t)$ are chosen to be affine in ξ^t for all $t \in [2,T]$, all the recourse variables are continuous (i.e., $p_t^{\mathtt{i}} = 0, t \in [2,T]$) and we have fixed parameters associated with the state variables, i.e., $c_t^{\mathtt{s}}(\xi^t) = c_t^{\mathtt{s}}, A_t^{\mathtt{s}}(\xi^t) = A_t^{\mathtt{s}}, B_t^{\mathtt{s}}(\xi^t) = B_t^{\mathtt{s}}, D_t^{\mathtt{s}}(\xi^t) = D_t^{\mathtt{s}}$ for all $t \in [2,T]$ and $\xi^t \in \Xi^t$. In this case, using linear programming duality or the KKT optimality conditions (under a further relatively complete recourse assumption for the 2ARO problem (6)) yields a monolithic bilinear continuous optimization problem, which can further be linearized using big-M constraints to obtain a mixed-integer linear program (see Appendix C.1 for details).

We further remark that the presence of decision-rule design variables β may make the master problem $\mathcal{MP}(\hat{\Xi})$ numerically challenging to solve, compared to traditional use cases of C&CG. In that regard, recently proposed enhancement ideas can be employed, such as the inexact C&CG algorithm proposed by Tsang et al. (2023). In their framework, master problems are solved to a given relative optimality gap which gradually reduces to zero over the course of the algorithm. Additionally, it is possible to design the basis functions $\Phi_t(\cdot)$ to use a small information basis rather than the entire history ξ^t so that the number of design variables is reduced. This would be especially helpful for problems with larger number of decision stages.

Another consideration concerning the exact solution of (6) using the C&CG algorithm is its convergence. Indeed, the C&CG may not have finite convergence in general. If Ξ^T is a finite set, then the finite convergence is straightforward, otherwise more conditions are needed to ensure this property. Zeng and Zhao (2013) proved finite convergence, considering a (bounded) polyhedral uncertainty set, for problems with only right-hand-side uncertainty represented as an affine function of uncertain parameters. However, in problem (6), basis functions $\Phi_t(\xi^t)$ appear as coefficients of β_t, both in the objective function (6a) (in the term $c_t^{\mathsf{s}}(\xi^t)^\top \Phi_t(\xi^t)$) and the second-stage constraints (terms $A_t^{\mathsf{s}}(\xi^t)\Phi_t(\xi^t), B_t^{\mathsf{s}}(\xi^t)\Phi_{t-1}(\xi^{t-1})$ and $D_t^{\mathsf{s}}(\xi^t)\Phi_t(\xi^t)$). The following proposition provides a sufficient condition for finite convergence of the C&CG algorithm (proof is given in Appendix B) in this more general context.

Proposition 1. *Consider an MSARO with only right-hand-side uncertainty, continuous recourse, and (bounded) polyhedral uncertainty set. If the basis functions $\Phi_t(\xi^t)$ are chosen to be affine in ξ^t for all $t \in [2,T]$, the C&CG algorithm converges to $\nu^{2S\text{-}LDR}$ in a finite number of iterations.*

In case the conditions of Proposition 1 are not satisfied, the C&CG algorithm still converges but asymptotically to an optimal solution of problem (6) if it is feasible (since it is bounded under the boundedness assumption imposed on the original MSARO problem).

Lastly, we note that in the case of continuous recourse, a linear decision rule can be applied to the recourse decision variables as well resulting in the LDR approach proposed by Ben-Tal et al. (2004). Since recourse variables $x_t^{\mathsf{r}}(\xi^t)$ in (6) are fully adjustable, it immediately follows that

$$\nu^\star \le \nu^{\text{2S-LDR}} \le \nu^{\text{LDR}},$$

where ν^{LDR} refers to the bound obtained from the commonly used LDR approach.

3.3. Two-stage Piecewise-constant Decision Rules for MSAROs with Integer State Variables

In this section, we study the application of two-stage decision rules in another special case of MSAROs, where each state variable $x_{ti}^{\mathsf{s}}(\xi^t) \in \mathbb{Z}, i \in [q_t], t \in [2,T]$ is a bounded integer with a given domain $[\underline{\kappa}_{ti}, \overline{\kappa}_{ti}]$ where boundedness follows from Assumption 3. We enforce the *two-stage piecewise-constant decision rule* (PCDR) by using $x^{\mathsf{s}}_{ti}(\xi^t) = \Theta_{ti}(\xi^t, \beta^i_t)$ where

$$\Theta_{ti}(\xi^t, \beta^i_t) = \begin{cases} \underline{\kappa}_{ti} & \Upsilon_{ti}(\xi^t, \beta^i_t) \in \mathcal{K}^i_{t1} \\ \underline{\kappa}_{ti} + 1 & \Upsilon_{ti}(\xi^t, \beta^i_t) \in \mathcal{K}^i_{t2} \\ \vdots & \vdots \\ \overline{\kappa}_{ti} & \Upsilon_{ti}(\xi^t, \beta^i_t) \in \mathcal{K}^i_{tJ_i}, \end{cases} \tag{9}$$

where $\mathcal{K}^i_{tj} \subset [-1,1], j \in [J_i]$ are disjoint sets with $\bigcup_{j\in[J_i]} \mathcal{K}^i_{tj} = [-1,1]$, and $\Upsilon_{ti}(\xi^t, \beta^i_t) : \mathbb{R}^{\ell^t} \times \mathbb{R}^{K_t} \to [-1,1]$ are functions defining the policy for $t \in [2,T]$ and $i \in [q_t]$. Semantically, the PCDR partitions the interval $[-1,1]$ into subsets, and then assigns an integer value to each partition.

A special case of PCDRs can be defined by restricting the form of all mappings $\Upsilon_{ti}(\xi^t, \beta^i_t)$ to be a linear function of the decision rule design variables β. In the following, we show that such a decision rule results in a model that is structurally very similar to (6), thus it is amenable to the C&CG method. Let, without loss of generality, $\mathcal{K}^i_{t1} = [a^i_{t1}, b^i_{t1}]$, and $\mathcal{K}^i_{tj} = (a^i_{tj}, b^i_{tj}], j \in [J_i] \setminus \{1\}$ be intervals with $a^i_{t1} = -1$, $b^i_{t,J_i} = 1$, and $a^i_{tj} = b^i_{t,j-1}$ for $j \in [2, J_i]$. Let further, for $t \in [2,T]$, $\Upsilon_{ti}(\xi^t, \beta^i_t) = \hat{\Upsilon}_{ti}(\xi^t)^\top \beta^i_t$ be chosen as an affine function of basis functions $\hat{\Upsilon}_{ti}(\xi^t)$. This implies, in particular, that $-1 \le \hat{\Upsilon}_{ti}(\xi^t)^\top \beta^i_t \le 1$ which is enforced through robust constraints. Then problem (4) becomes:

$$\begin{aligned}
\nu^{\text{2S-PCDR}} := \min\ & c_1^\top x_1 + \mathcal{SP}^{\text{PCDR}}(\beta, x^{\mathsf{s}}_1) && && \text{(10a)}\\
\text{s.t. }\ & x_1 \in X_1 && && \text{(10b)}\\
& \beta_t \in \mathbb{R}^{K_t} && t \in [2,T] && \text{(10c)}\\
& -1 \le \hat{\Upsilon}_{ti}(\xi^t)^\top \beta^i_t \le 1 && t \in [2,T], i \in [q_t], \xi^T \in \Xi^T && \text{(10d)}
\end{aligned}$$

where:

$$\begin{aligned}
\mathcal{SP}^{\text{PCDR}}(\beta, x^{\mathsf{s}}_1) := \max_{\xi^T \in \Xi^T}\ & \sum_{t\in[2,T]} c^{\mathsf{s}}_t(\xi^t)^\top x^{\mathsf{s}}_t + \min_{x^{\mathsf{r}} \in \mathcal{X}(x^{\mathsf{s}}_1, \beta, \xi^T)} \sum_{t\in[2,T]} c^{\mathsf{r}}_t(\xi^t)^\top x^{\mathsf{r}} && && \text{(11a)}\\
\text{s.t. }\ & \sum_{j\in[J_i]} (\underline{\kappa}_{ti} + j - 1) \upsilon_{tij} = x^{\mathsf{s}}_{ti} && t \in [2,T], i \in [q_t] && \text{(11b)}\\
& \sum_{j\in[J_i]} \omega_{tij} = \hat{\Upsilon}_{ti}(\xi^t)^\top \beta^i_t && t \in [2,T], i \in [q_t] && \text{(11c)}\\
& (a^i_{tj} + \epsilon_j) \upsilon_{tij} \le \omega_{tij} \le b^i_{tj} \upsilon_{tij} && t \in [2,T], i \in [q_t], j \in [J_i] && \text{(11d)}\\
& \sum_{j\in[J_i]} \upsilon_{tij} = 1 && t \in [2,T], i \in [q_t] && \text{(11e)}\\
& \upsilon_{tij} \in \{0,1\} && t \in [2,T], i \in [q_t], j \in [J_i]. && \text{(11f)}
\end{aligned}$$

with $\epsilon_1 = 0$. The PCDR is modeled using the auxiliary variables $\upsilon_{tij}, \omega_{tij}$ and constraints (11b)-(11f). Variables υ determine in which interval the quantity $\hat{\Upsilon}_{ti}(\xi^t)^\top \beta^i_t$ falls, in accordance with the variables ω, and the integer values assigned to the state variables. Here, ϵ_j is added to the lower bound on ω_{tij} to ensure that partitions $\mathcal{K}_{tj}$ are disjoint. We remark that, it is possible to choose $\epsilon_j = 0$ for $j \in [J_i]$, in which case the intervals would intersect at their boundaries. In this case, whenever the quantity $\hat{\Upsilon}_{ti}(\xi^t)^\top \beta_t^i$ is a boundary point, the model allows assigning either one of the corresponding integer values to the associated state variable. The objective function then dictates that the solution leading to the worst objective value is chosen.

Similar to Remark 1, in certain special cases, we can reformulate problem (11) as a monolithic mixed-integer linear program.

Remark 2. Consider an MSARO where the uncertainty set is a polytope, the basis functions $\hat{\Upsilon}_{ti}(\xi^t)$ are chosen to be affine in ξ^t for all $t \in [2,T]$ and $i \in [q_t]$, all the recourse variables are continuous, and we have fixed parameters associated with the state variables, i.e., $c_t^{\mathsf{s}}(\xi^t) = c_t^{\mathsf{s}}, A_t^{\mathsf{s}}(\xi^t) = A_t^{\mathsf{s}}, B_t^{\mathsf{s}}(\xi^t) = B_t^{\mathsf{s}}, D_t^{\mathsf{s}}(\xi^t) = D_t^{\mathsf{s}}$ for all $t \in [2,T]$ and $\xi^t \in \Xi^t$. In this case, using linear programming duality yields a monolithic bilinear continuous optimization problem whose objective function involves products between variables x_{it}^{s} and linear programming dual variables of the inner minimization problem. Thanks to constraints (11b), variables x_{it}^{s} can be substituted for a weighted sum of binary variables υ, as such these bilinear terms can be linearized using big-M constraints to obtain a mixed-integer linear program.

Further, in applying the C&CG method to model (10), robust constraints (10d) will appear in the master problem. If the uncertainty set is a polytope and the basis functions $\hat{\Upsilon}_{ti}(\xi^t)$ are chosen to be affine in ξ^t for all $t \in [2,T]$ and $i \in [q_t]$ then these semi-infinite constraints can be reformulated as a finite set of linear constraints using classical robust optimization techniques based on linear programming duality. We also note that the arguments presented in Section 3.2 imply similarly that the C&CG algorithm converges asymptotically to the optimal solution of $\nu^{\text{2S-PCDR}}$ if it is feasible.

Lastly, the methods presented in Sections 3.2 and 3.3 can be combined to address MSAROs with mixed-integer state variables.

Remark 3. For an MSARO with mixed-integer state variables, the application of linear and piecewise constant decision rules, given by equations (5) and (9), to the continuous and integer state variables, respectively, yields a 2ARO approximation. The resulting model is presented in detail in Appendix C.2. This model is similarly amenable to the C&CG method for exact solution but can also benefit from other 2ARO solution methods from the literature.

4. Dual Bounding

In this section, we introduce a new dual problem that provides a lower bound for MSARO problems with mixed-integer recourse. Due to the existence of integer variables, we rely on Lagrangian duality techniques, where we create a Lagrangian relaxation and optimize over the Lagrangian dual multipliers. As the MSARO involves constraints corresponding to every realization of uncertainty, Lagrangian multipliers are functions of uncertainty, as such they are usually high (possibly infinite) dimensional. To overcome the difficulty in their optimization, we propose to apply decision rule restrictions to Lagrangian multipliers, leveraging ideas rooted in the MSP literature (Kuhn et al. 2011, Daryalal et al. 2024). In deriving a Lagrangian dual of the MSARO problem, we *choose* a probability distribution with the support as the uncertainty set, and use the associated density function to scale the constraints to be dualized, resulting in expectation terms in the objective function of the relaxation. Accordingly, we obtain a dual approximation of MSARO in the form of a two-stage stochastic program. This probability distribution-based approach has several benefits, most notably the possibility of leveraging state-of-the-art stochastic programming techniques to solve the dual problem. However, the quality of the resulting dual bound depends on the probability distribution used while developing the dual formulation, as previously observed by Kuhn et al. (2011) for MSAROs with continuous recourse. With the aim of identifying the strongest such dual bound, we formally pose a distribution optimization problem (akin to what was developed in (Hadjiyiannis et al. 2011)) and develop appropriate solution methods (tailored to the nature of the recourse variables) for the resulting distribution optimization problem. To the best of our knowledge, numerical solution of such a bounding problem and the quality of the obtained bounds have not been studied before.

In what follows, in Section 4.1, we introduce the nonanticipative reformulation of the MSARO problem and its Lagrangian dual. In Section 4.2, we present the restricted Lagrangian dual problem and define the associated distribution optimization problem for which we develop solution methods in Section 4.3. Lastly, in Section 4.4, we propose an alternative dual problem, which can be weaker in terms of the quality of the obtained bound but has computational advantages thanks to its decomposable structure.

4.1. Nonanticipative Dual of the MSARO

The nonanticipative (NA) dual is based on a reformulation of the MSARO problem where we create a copy of decision variables for every stage and every realization, and explicitly enforce nonanticipativity constraints. To this end, we introduce the copy variables $y(\xi^T) = (y_1(\xi^T), \dots, y_T(\xi^T))$ for all $\xi^T \in \Xi^T$ as perfect information variables depending on the entire realization $\xi^T = (\xi_1, \dots, \xi_T)$.

We denote the decision variables by y instead of x used in previous sections to emphasize the fact that they are perfect information variables. We also note that we do not need to distinguish state and recourse variables in this section, thus vectors y involve all the decisions. We can then obtain the NA reformulation of the MSARO problem (2) as:

$$
\begin{aligned}
\min\ & z && && \text{(12a)}\\
\text{s.t.}\ & \sum_{t\in[T]} c_t(\xi^t)^\top y_t(\xi^T) \le z && \xi^T \in \Xi^T && \text{(12b)}\\
& A_t(\xi^t)y_t(\xi^T) + B_t(\xi^t)y_{t-1}(\xi^T) \le b_t(\xi^t) && t\in[2,T],\ \xi^T\in\Xi^T && \text{(12c)}\\
& D_t(\xi^t)y_t(\xi^T) \le d_t(\xi^t) && t\in[T],\ \xi^T\in\Xi^T && \text{(12d)}\\
& y_t(\xi^T) = y_t(\xi'^T) && t\in[T],\ \xi^T,\xi'^T\in\Xi^T \text{ with } \xi^t=\xi'^t && \text{(12e)}\\
& y_t(\xi^T) \in \mathbb{R}^{n_t-n_t^{\mathrm{i}}}\times\mathbb{Z}^{n_t^{\mathrm{i}}} && t\in[T],\ \xi^T\in\Xi^T. && \text{(12f)}
\end{aligned}
$$

Constraints (12e) are *nonanticipativity constraints* which ensure that at stage t for every partial realization of ξ^T, the decisions made are consistent (i.e., the decisions made in all realizations sharing the history ξ^t are the same).

Remark 4. The nonanticipativity constraints are redundant for stage T, however, we include them in model (12) for notational convenience. In our implementation for the numerical results presented in Section 5, we exclude those redundant constraints.

In deriving a Lagrangian relaxation, we will first scale the constraints to be relaxed. As the constraints correspond to uncertainty realizations, we *choose* the scaling factors in such a way that they induce a probability measure over the support Ξ^T. To this end, we let $\mathbb{P}$ denote this probability measure such that $\mathbb{P}(\Xi^T)=1$, which we interchangeably refer to as the probability distribution. We define $p^{\mathbb{P}}:\Xi^T\to\mathbb{R}^+$ as the associated density function. Lastly, we let $\mathcal{P}^> := \{\mathbb{P} \mid p^{\mathbb{P}}(\xi^T)>0, \xi^T\in\Xi^T\}$, i.e., every $\mathbb{P}\in\mathcal{P}^>$ has a density function assigning a strictly positive value to all $\xi^T\in\Xi^T$.

Before introducing the NA dual derived from (12), we present the following lemma, which we use to reformulate the nonanticipativity constraints (proof in Appendix B).

Lemma 1. *For any $\mathbb{P}\in\mathcal{P}^>$, constraints (12e) are equivalent to the following:*

$$
y_t(\xi^T) = \mathbb{E}_{\xi'^T\sim\mathbb{P}}\left[y_t(\xi'^T) \;\middle|\; \xi'^t=\xi^t \right], \quad t\in[T],\ \xi^T\in\Xi^T. \tag{13}
$$

One advantage of this reformulation is the reduction in the number of constraints and in turn in the number of dual multipliers to be introduced. Let, to this end, $\lambda_t(\cdot):\mathbb{R}^{\ell^t}\to\mathbb{R}^{n_t}$ for $t\in[T]$, where $\mathbb{E}_{\xi^T\sim\mathbb{P}}[\lambda_t(\xi^T)]<+\infty$, be the dual functionals to be used in relaxing the nonanticipativity constraints (13). Let further the feasibility space for $\xi^T\in\Xi^T$ be $Y(\xi^T):=\big\{\big(z,y_1(\xi^T),\dots,y_T(\xi^T)\big):$

(12b) – (12d) and (12f)$\}$. Then, after scaling (13) with the probability densities associated with the distribution $\mathbb{P}$, we obtain the following NA Lagrangian relaxation problem:

$$\mathcal{L}^{\text{NA}}_{\text{LR}}(\mathbb{P},\lambda_1(\cdot),\dots,\lambda_T(\cdot)) := \min\ z + \sum_{t\in[T]} \mathbb{E}_{\xi^T\sim\mathbb{P}}\Big[\lambda_t(\xi^T)^\top\Big(y_t(\xi^T) - \mathbb{E}_{\xi'^T\sim\mathbb{P}}\big[y_t(\xi'^T)|\xi'^t=\xi^t\big]\Big)\Big] \tag{14a}$$

$$\text{s.t. } \big(z, y_1(\xi^T),\dots,y_T(\xi^T)\big) \in Y(\xi^T) \qquad \xi^T\in\Xi^T. \tag{14b}$$

whose optimal objective value yields a lower bound for the original MSARO problem for any $\mathbb{P}\in\mathcal{P}^>$. The NA Lagrangian dual problem aims to find the best bound among all such Lagrangian relaxation bounds:

$$\mathcal{L}^{\text{NA}}(\mathbb{P}) := \max_{\lambda_1(\cdot),\dots,\lambda_T(\cdot)} \quad \mathcal{L}^{\text{NA}}_{\text{LR}}(\mathbb{P},\lambda_1(\cdot),\dots,\lambda_T(\cdot)). \tag{15}$$

The following proposition shows that regardless of the choice of $\mathbb{P}$, $\mathcal{L}^{\text{NA}}(\mathbb{P})$ is an exact dual bound for MSARO problems with *continuous recourse* (proof in Appendix B).

Proposition 2. *Let $\mathbb{P}$ be any probability measure in $\mathcal{P}^>$. For MSARO problems with* continuous recourse*, (15) is a strong dual of (2), i.e., $\mathcal{L}^{\text{NA}}(\mathbb{P}) = \nu^\star$.*

We remark that our construction of the dual postulates that we multiply the constraints with a density function whose support is Ξ^T. Therefore, the strictly positive density property of $\mathcal{P}^>$ is necessary for the exactness of our formulation. Proposition 2 suggests that the solution of problem (15) gives an exact dual bound for (2) (hence for (12)) if all decision variables are continuous. Although this bound is not necessarily exact in the case of mixed-integer recourse, the potential of leveraging the literature of multistage stochastic programming in achieving a dual bound for MSARO is quite appealing.

The objective function of the NA Lagrangian relaxation problem (14) contains (conditional) expectations of decision variables $y_t(\xi^T)$, which is computationally challenging. Because of our initial assumption that $X_t(\xi^t)$ are bounded (Assumption 3) we have, by letting $\tilde{Y}(\xi^T) := \text{proj}_{y_1(\cdot),\dots,y_T(\cdot)} Y(\xi^T)$, that $\mathbb{E}_{\xi^T\sim\mathbb{P}}[\text{diam}(\tilde{Y}(\xi^T))] < +\infty$. Since we further have that $\mathbb{E}_{\xi^T\sim\mathbb{P}}[\lambda_t(\xi^T)] < +\infty$, we can apply Lemma 1 of (Daryalal et al. 2024) and replace the expectation term for t in the objective function (14a) with:

$$\mathbb{E}_{\xi^T\sim\mathbb{P}}\Big[\Big(\lambda_t(\xi^T) - \mathbb{E}_{\xi'^T\sim\mathbb{P}}\big[\lambda_t(\xi'^T)|\xi'^t=\xi^t\big]\Big)^\top y_t(\xi^T)\Big].$$

For given $\lambda_t(\xi^T)$, this exchange allows us to compute the coefficients of $y_t(\xi^T)$ in the Lagrangian relaxation problem. Still, the optimal form of the dual functionals $\lambda_t(\cdot)$ need to be determined, making the problem (15) computationally intractable. In the next section, we restrict these Lagrangian multipliers to follow LDRs and obtain a restricted dual problem with decision variables of smaller (finite) dimension. Furthermore, this new dual problem is amenable to well-known solution techniques from the literature of two-stage stochastic programming which are designed to approximately solve a problem with expectation in the objective function.

4.2. Lagrangian Dual Decision Rules

We restrict the NA Lagrangian dual problem (15) for a given $\mathbb{P} \in \mathcal{P}^>$ by enforcing LDRs on the Lagrangian multipliers, referred to as Lagrangian dual decision rules (LDDRs). For a set of pre-determined basis functions $\Psi_t : \Xi^T \to \mathbb{R}^{n_t \times K_t}$ and LDR decision variables $\alpha_t \in \mathbb{R}^{K_t}$, we restrict the form of $\lambda_t(\xi^T)$ at stage $t \in [T]$ as follows:

$$\lambda_t(\xi^T) = \Psi_t(\xi^T)\alpha_t,$$

which gives us a restricted NA Lagrangian dual problem with respect to $\mathbb{P}$:

$$\mathcal{L}_R^{\mathrm{NA}}(\mathbb{P}) := \max_{\alpha_1,\ldots,\alpha_T} \quad \mathcal{L}_{\mathrm{LR}}^{\mathrm{NA}}(\mathbb{P}, \Psi_1(\xi^T)\alpha_1, \ldots, \Psi_T(\xi^T)\alpha_T). \tag{16}$$

Since problem (16) is a restriction of (15), we have $\mathcal{L}_R^{\mathrm{NA}}(\mathbb{P}) \leq \mathcal{L}^{\mathrm{NA}}(\mathbb{P})$.

Using Lemma 2 in (Daryalal et al. 2024), the primal characterization of $\mathcal{L}_R^{\mathrm{NA}}(\mathbb{P})$ is:

$$\min \quad z \tag{17a}$$

$$\text{s.t.} \quad \big(z, y_1(\xi^T), \ldots, y_T(\xi^T)\big) \in \operatorname{conv}(Y(\xi^T)) \qquad \xi^T \in \Xi^T \tag{17b}$$

$$\mathbb{E}_{\xi^T \sim \mathbb{P}}\left[\Psi_t(\xi^T)^\top \Big(y_t(\xi^T) - \mathbb{E}_{\xi'^T \sim \mathbb{P}}\big[y_t(\xi'^T) \mid \xi'^t = \xi^t\big]\Big)\right] = \mathbf{0} \qquad t \in [T] \tag{17c}$$

For a given $\mathbb{P}$, comparing (17) to the primal characterization of the (unrestricted) NA Lagrangian dual problem (15) (provided in (42) of Appendix B), it is clear that, the former is a relaxation of the latter since constraints (17c) are an aggregation of their counterpart (42c). Consequently, unlike $\mathcal{L}^{\mathrm{NA}}(\mathbb{P})$, even for MSARO with continuous recourse, $\mathcal{L}_R^{\mathrm{NA}}(\mathbb{P})$ is not necessarily a strong bound. Furthermore, due to constraints (17c), the strength of the restricted NA Lagrangian dual bound depends on the choice of probability measure $\mathbb{P}$. A similar observation was made by Kuhn et al. (2011) and Hadjiyiannis et al. (2011) concerning their dual bound for MSARO problems *with only continuous variables*. This observation motivates us to optimize over the probability distribution $\mathbb{P}$ and LDR variables α to find the best such dual bound. As such, we propose to solve a distribution optimization (DO) problem over the set of probability distributions $\mathcal{P}^>$, defined as follows:

$$\nu_R^{\text{NA-DO}} := \sup_{\mathbb{P} \in \mathcal{P}^>} \quad \mathcal{L}_R^{\mathrm{NA}}(\mathbb{P}) \tag{18}$$

where, as before, $\mathcal{P}^> = \{\mathbb{P} \mid p^{\mathbb{P}}(\xi^T) > 0, \xi^T \in \Xi^T\}$.

In linear and mixed-integer programming, strict inequalities such as the ones required for $\mathbb{P} \in \mathcal{P}^>$ (the strictly positive density property for all $\xi^T \in \Xi^T$) often cause numerical and theoretical difficulties, thus are not desirable. To avoid these inequalities, in the following discussions, we modify the DO model to improve its numerical behaviour. Denote by $\mathcal{P}^{\geq}$ a superset of $\mathcal{P}^{>}$ that also admits distributions that allow $p(\xi^T) = 0$ for some $\xi^T \in \Xi^T$. Consider the problem:

$$\bar{\nu}_R^{\text{NA-DO}} := \max_{\mathbb{P} \in \mathcal{P}^{\geq}} \mathcal{L}_R^{\text{NA}}(\mathbb{P}). \tag{19}$$

Because (19) is a relaxation of (18), it yields an upper bound for $\nu_R^{\text{NA-DO}}$. Thus, it does not immediately follow that such a bound is a valid lower bound for the optimal value of the MSARO problem, $\nu^\star$. The following proposition shows that (19) indeed leads to a valid dual (lower) bound (see Appendix B for proof).

Proposition 3. *$\bar{\nu}_R^{\text{NA-DO}}$ is a lower bound for $\nu^\star$.*

Hereafter, we refer to (19) or its equivalent explicit form

$$\bar{\nu}_R^{\text{NA-DO}} = \max_{\mathbb{P}, \alpha} \quad \mathcal{Q}(\mathbb{P}, \alpha) \tag{20a}$$

$$\text{s.t.} \quad \alpha_t \in \mathbb{R}^{K_t} \quad t \in [T] \tag{20b}$$

$$\mathbb{P} \in \mathcal{P}^{\geq}, \tag{20c}$$

as the DO problem, where

$$\mathcal{Q}(\mathbb{P}, \alpha) := \min \quad z + \sum_{t \in [T]} \mathbb{E}_{\xi^T \sim \mathbb{P}} \left[\left(\left(\Psi_t(\xi^T) - \mathbb{E}_{\xi'^T \sim \mathbb{P}} \left[\Psi_t(\xi'^T) \middle| \xi'^t = \xi^t \right] \right) \alpha_t \right)^\top y_t(\xi^T) \right] \tag{21a}$$

$$\text{s.t.} \quad \left(z, y_1(\xi^T), \ldots, y_T(\xi^T)\right) \in Y(\xi^T) \qquad \xi^T \in \Xi^T. \tag{21b}$$

Problem (21) is a *two-stage stochastic program* (2SP) and can benefit from its rich literature. In the next section, we build on well-known stochastic programming techniques to design a decomposition method to solve the DO problem (20).

4.3. Solving the DO Problem

There are two main challenges associated with the solution of the DO problem: the expectation terms in the objective of (21a) and the max-min structure in (20).

If the uncertainty set of the MSARO problem, Ξ^T, is not discrete, objective function (21a) includes the expectation of a nonsmooth concave function. Further, even when Ξ^T is discrete calculating the expectation term exactly can be prohibitive from a computational point of view. The literature of two-stage stochastic programming addresses such a difficulty by means of sampling-based approaches that replace the expectation in the objective function with the average of a sample drawn from the underlying distribution and has favourable theoretical convergence results (see e.g., Shapiro et al. (2009)). We follow the sample average approximation (SAA) approach in overcoming the first challenge and show in Section 4.3.1 that it leads to a valid dual bound for the MSARO problem.

With regards to the second challenge, we propose a cutting plane algorithm that iteratively constructs improving approximations of $\mathcal{Q}(\mathbb{P},\alpha)$ through its supporting hyperplanes obtained by solving the SAA approximation of (21). We present this general algorithm in Section 4.3.2 and propose an alternative monolithic formulation in the special case of MSARO with continuous recourse in Section 4.3.3. Figure 4 summarizes the methods presented in this section for solving the DO model.

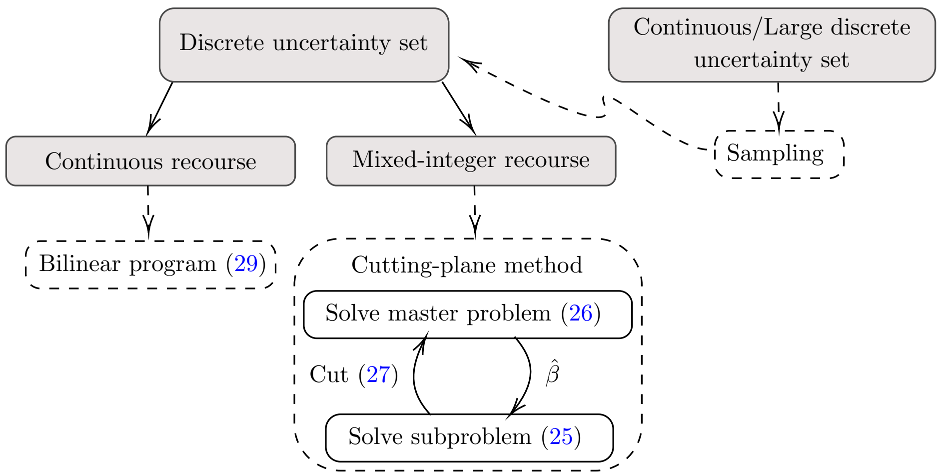

Figure 4 Solution methods for the DO problem

4.3.1. Sample average approximation (SAA) Let $\Omega \subseteq \Xi^T$ be a finite subset of the uncertainty set Ξ^T. We define the set of probability measures $\mathcal{P}^{\geq}_{\Omega}$ such that $\mathbb{P}_{\Omega} \in \mathcal{P}^{\geq}_{\Omega}$ implies that $\mathbb{P}_{\Omega}(\Omega) = 1$ and the associated density function has value zero for any realization not in Ω:

$$p^{\mathbb{P}_{\Omega}}(\xi^T) = 0,\ \xi^T \in \Xi^T \setminus \Omega,\ \text{and}\ p^{\mathbb{P}_{\Omega}}(\xi^T) \geq 0,\ \xi^T \in \Omega.$$

Since $\mathcal{P}^{\geq}_{\Omega} \subseteq \mathcal{P}^{\geq}$ and $\bar{\nu}^{\text{NA-DO}}_{R}$ is obtained by maximizing $\mathcal{L}^{\text{NA}}_{R}(\mathbb{P})$ over $\mathbb{P} \in \mathcal{P}^{\geq}$, we have, for any $\mathbb{P}_{\Omega} \in \mathcal{P}^{\geq}_{\Omega}$, that $\mathcal{L}^{\text{NA}}_{R}(\mathbb{P}_{\Omega}) \leq \bar{\nu}^{\text{NA-DO}}_{R} \leq \nu^{\star}$. It then follows that

$$\max_{\mathbb{P}_{\Omega} \in \mathcal{P}_{\Omega}} \mathcal{L}^{\text{NA}}_{R}(\mathbb{P}_{\Omega}) \leq \bar{\nu}^{\text{NA-DO}}_{R} \leq \nu^{\star}. \tag{22}$$

We finally have that

$$\begin{aligned} \max_{\mathbb{P}_{\Omega} \in \mathcal{P}^{\geq}_{\Omega}} \mathcal{L}^{\text{NA}}_{R}(\mathbb{P}_{\Omega}) = \max_{\mathbb{P}_{\Omega},\alpha} \quad & \mathcal{Q}(\mathbb{P}_{\Omega},\alpha) \\ \text{s.t.} \quad & \alpha_t \in \mathbb{R}^{K_t} \quad t \in [T] \\ & \mathbb{P}_{\Omega} \in \mathcal{P}^{\geq}_{\Omega}, \end{aligned} \tag{23}$$

where for any $\mathbb{P}_\Omega \in \mathcal{P}_\Omega^{\geq}$ the expectation terms in the objective function (21a), used in calculating $\mathcal{Q}(\mathbb{P}_\Omega, \alpha)$, are replaced by their sample average over Ω.

In the remainder of this section, we omit the notation Ω and write our models with a finite discrete uncertainty set Ξ^T which can either be the full uncertainty set of the MSARO problem or a set of realizations sampled from it.

4.3.2. MSARO with mixed-integer recourse In a discrete uncertainty set with realizations $\xi^T \in \Xi^T$, $\mathcal{P}^{\geq}$ can be modeled by a set of (in)equalities, that is $\mathcal{P}^{\geq} = \{\rho \in \mathbb{R}_+^{|\Xi^T|} \mid \mathbf{1}^\top \rho = 1\}$, where a vector $\rho^{\mathbb{P}} \in \mathcal{P}^{\geq}$ characterizes the probability measure $\mathbb{P}$ such that for all $\xi^T \in \Xi^T$, $\rho^{\mathbb{P}}_{\xi^T}$ is the probability of realization ξ^T with respect to $\mathbb{P}$. Define $\rho^{\mathbb{P}}_{|\xi^t} := \sum_{\substack{\xi'^T \in \Xi^T: \\ \xi'^t = \xi^t}} \rho^{\mathbb{P}}_{\xi'^T}$ as the sum of the probabilities of all realizations sharing the same history ξ^t up to stage t and let $\rho^{\mathbb{P}}_{\xi'^T|\xi^t}$ be the conditional probability of realization $\xi'^T \in \Xi^T$ given history ξ^t. More precisely, given ξ^t, for $\xi'^T \in \Xi^T$ with $\xi'^t = \xi^t$ we have that $\rho^{\mathbb{P}}_{\xi'^T|\xi^t} = \frac{\rho^{\mathbb{P}}_{\xi'^T}}{\rho^{\mathbb{P}}_{|\xi^t}}$ if $\rho^{\mathbb{P}}_{|\xi^t} > 0$ and that $\rho^{\mathbb{P}}_{\xi'^T|\xi^t} = 0$ otherwise. Then, the expectation term in the objective function (21a) can be written as:

$$\sum_{t\in[T]} \sum_{\xi^T \in \Xi^T} \rho^{\mathbb{P}}_{\xi^T} \bigg(\Big(\Psi_t(\xi^T) - \sum_{\substack{\xi'^T \in \Xi^T: \\ \xi'^t = \xi^t}} \rho^{\mathbb{P}}_{\xi'^T|\xi^t} \Psi_t(\xi'^T) \Big) \alpha_t \bigg)^\top y_t(\xi^T).$$

Let $\Psi_{tk}(\xi^T)$ be the k^{th} column of the matrix $\Psi_t(\xi^T)$. Then $\mathcal{Q}(\mathbb{P}, \alpha)$ can be expressed as:

$$\begin{aligned} \min \quad & z + \sum_{t\in[T]} \sum_{\xi^T \in \Xi^T} \rho^{\mathbb{P}}_{\xi^T} \bigg(\sum_{k\in[K_t]} \Big(\Psi_{tk}(\xi^T) - \sum_{\substack{\xi'^T \in \Xi^T: \\ \xi'^t = \xi^t}} \rho^{\mathbb{P}}_{\xi'^T|\xi^t} \Psi_{tk}(\xi'^T) \Big) \alpha_{tk} \bigg)^\top y_t(\xi^T) \\ \text{s.t.} \quad & \big(z, y_1(\xi^T), \ldots, y_T(\xi^T)\big) \in Y(\xi^T) \quad \xi^T \in \Xi^T. \end{aligned}$$

Let $\gamma_{tk\xi^T} := \rho^{\mathbb{P}}_{\xi^T} \alpha_{tk}$ and define $\beta_{t\xi^T} := \sum_{k\in[K_t]} \Big(\gamma_{tk\xi^T} \Psi_{tk}(\xi^T) - \sum_{\substack{\xi'^T \in \Xi^T: \\ \xi'^t = \xi^t}} \rho^{\mathbb{P}}_{\xi'^T|\xi^t} \gamma_{tk\xi^T} \Psi_{tk}(\xi'^T) \Big)$ as the coefficient vector of variables $y_t(\xi^T)$. Then with a change of variables in the DO problem (20) we have:

$$\bar{\nu}_R^{\text{NA-DO}} = \max_{\rho,\alpha,\gamma,\beta} \mathcal{Q}(\beta) \tag{24a}$$

$$\text{s.t.} \sum_{\xi^T \in \Xi^T} \rho_{\xi^T} = 1 \tag{24b}$$

$$\gamma_{tk\xi^T} = \rho_{\xi^T} \alpha_{tk} \qquad t \in [T], k \in [K_t], \xi^T \in \Xi^T \tag{24c}$$

$$\sum_{\substack{\xi'^T \in \Xi^T: \\ \xi'^t = \xi^t}} \rho^{\mathbb{P}}_{\xi'^T} \beta_{t\xi^T} - \sum_{k\in[K_t]} \Big(\sum_{\substack{\xi'^T \in \Xi^T: \\ \xi'^t = \xi^t}} \rho^{\mathbb{P}}_{\xi'^T} \gamma_{tk\xi^T} \Psi_{tk}(\xi^T) - \sum_{\substack{\xi'^T \in \Xi^T: \\ \xi'^t = \xi^t}} \rho_{\xi^T} \gamma_{tk\xi^T} \Psi_{tk}(\xi'^T) \Big) = \mathbf{0} \qquad t \in [T], \xi^T \in \Xi^T \tag{24d}$$

$$-\boldsymbol{M} \rho_{\xi^T} \leq \beta_{t\xi^T} \leq \boldsymbol{M} \rho_{\xi^T} \qquad t \in [T], \xi^T \in \Xi^T \tag{24e}$$

$$\rho \in \mathbb{R}_+^{|\Xi^T|} \tag{24f}$$

$$\alpha_t \in \mathbb{R}^{K_t},\ \gamma_t \in \mathbb{R}^{K_t \times |\Xi^T|},\ \beta_t \in \mathbb{R}^{|\Xi^T| \times n_t} \qquad t \in [T] \tag{24g}$$

where

$$\mathcal{Q}(\beta) = \min\Big\{ z + \sum_{t\in[T]} \sum_{\xi^T \in \Xi^T} \beta_{t\xi^T}^\top y_t(\xi^T) \ \Big|\ \big(z, y_1(\xi^T), \ldots, y_T(\xi^T)\big) \in Y(\xi^T),\ \xi^T \in \Xi^T \Big\}, \tag{25}$$

and $\boldsymbol{M} \geq \mathbf{0}$ is a vector of sufficiently large numbers.

In model (24), constraints (24b) along with the nonnegativity restrictions imposed on variables ρ induce a probability distribution over the realizations Ξ^T with some realizations potentially assigned zero probability. For given ξ^T, if its associated probability ρ_{ξ^T} is zero, then constraints (24c) and (24e) imply, respectively, that all variables γ and β indexed by ξ^T are zero, as such constraints (24d) trivially hold. Otherwise, since probability $\rho_{\xi^T} > 0$, we have that $\sum_{\substack{\xi'^T \in \Xi^T: \\ \xi'^t = \xi^t}} \rho^{\mathbb{P}}_{\xi'^T} > 0$, as such constraints (24d) correctly impose the definition of β by dividing all terms by $\sum_{\substack{\xi'^T \in \Xi^T: \\ \xi'^t = \xi^t}} \rho^{\mathbb{P}}_{\xi'^T}$ and plugging in the definition of γ from constraints (24c). Then given coefficients β, (25) evaluates the expected value of the optimal decisions $y(\cdot)$.

Remark 5. The choice of values for the vector $\boldsymbol{M}$ impacts the quality of the bound obtained from model (24). In particular, in the limit case where $\boldsymbol{M} = \mathbf{0}$, one obtains the perfect information bound, that is, all nonanticipativity constraints are relaxed from the NA reformulation of the MSARO problem. Otherwise, for larger values of $\boldsymbol{M}$, the optimal value of (24) is lower bounded by the perfect information bound since choosing $\beta = \mathbf{0}$ is always feasible. One can therefore expect to obtain a better bound from (24).

Model (24) can be solved via a cutting-plane method in which $\mathcal{Q}(\beta)$ is approximated by a set of linear inequalities. At each iteration, we solve the following bilinear program as the master problem:

$$\max_{\rho,\alpha,\gamma,\beta} \Big\{ \eta \ \Big|\ (24b) - (24g),\ (\eta, \beta) \in \mathcal{H} \Big\}, \tag{26}$$

where η is an auxiliary variable representing $\mathcal{Q}(\beta)$, and $\mathcal{H}$ is a set described by optimality cuts approximating $\mathcal{Q}(\beta)$. Note that, as β only parameterizes the objective function of $\mathcal{Q}(\beta)$, i.e., it does not impact the feasibility space, there is no need for feasibility cuts. With $(\hat{\eta}, \hat{\beta})$ returned from solving the master problem (26), we solve the subproblem (25) to compute $\mathcal{Q}(\hat{\beta})$, resulting in $\hat{y}_t(\xi^T)$ as the optimal solution. If $\hat{\eta} \leq \mathcal{Q}(\hat{\beta})$, we have found the optimal solution of the DO problem. Otherwise we add the following optimality cut to the master problem:

$$\eta \leq \mathcal{Q}(\hat{\beta}) + \sum_{t\in[T]} \sum_{\xi^T \in \Xi^T} \Big(\beta_{t\xi^T} - \hat{\beta}_{t\xi^T} \Big)^\top \hat{y}_t(\xi^T). \tag{27}$$

This procedure continues until no more optimality cuts are found. The objective function of the subproblem, $\mathcal{Q}(\beta)$, is a concave function in β (pointwise minimum of linear functions with respect to β). Each cut (27) is a hyperplane approximating the subproblem from above. The cutting-plane method iteratively finds improving approximations of $\mathcal{Q}(\beta)$. We note that it is not necessary to execute the cutting-plane procedure until convergence in order to obtain a valid dual bound. Indeed at each iteration of the algorithm the value $\mathcal{Q}(\hat{\beta})$ provides a valid dual bound for $\nu^\star$.

Remark 6. In our implementation for the numerical results presented in Section 5, rather than creating the copies of first-stage variables $y_1(\cdot)$ and relaxing their nonanticipativity constraints, we keep them as static variables in (25), same as variable z.

4.3.3. MSARO with continuous recourse As a special case, we study MSARO with continuous recourse. Our goal here is to use this particular structure and derive a monolithic formulation as an alternative to the cutting-plane algorithm, to leverage off-the-shelf solvers. Denote by u_{ξ^T}, $v_{t\xi^T}$ and $w_{t\xi^T}$, the dual variables associated with the set of constraints described by $Y(\xi^T)$ for given $\xi^T \in \Xi^T$ (corresponding to (12b), (12c) and (12d), respectively). The linear programming dual of the inner minimization problem (25), i.e., the subproblem of the cutting-plane algorithm, is:

$$
\begin{aligned}
\mathcal{Q}^D(\beta) = \max \quad & \sum_{t\in[T]} \sum_{\xi^T\in\Xi^T} b_t(\xi^T)^\top v_{t\xi^T} + \sum_{t\in[T]} \sum_{\xi^T\in\Xi^T} d_t(\xi^t)^\top w_{t\xi^T} && \text{(28a)}\\
\text{s.t.} \quad & \sum_{\xi^T\in\Xi^T} u_{\xi^T} = 1 && \text{(28b)}\\
& -c_T(\xi^T)u_{\xi^T} + A_T(\xi^T)^\top v_{T\xi^T} + D^\top_{T\xi^T} w_{T\xi^T} - \beta_{T\xi^T} = \mathbf{0} \quad \xi^T \in \Xi^T && \text{(28c)}\\
& -c_t(\xi^t)u_{\xi^T} + A_t(\xi^t)^\top v_{t\xi^T} + D_t(\xi^t)^\top w_{t\xi^T} + \\
& \qquad B_{t+1}(\xi^{t+1})^\top v_{t+1,\xi^T} - \beta_{t\xi^T} = \mathbf{0} \quad t\in[T-1], \xi^T \in \Xi^T && \text{(28d)}\\
& u_{\xi^T} \geq 0 \quad \xi^T \in \Xi^T && \text{(28e)}\\
& v_{t\xi^T}, w_{t\xi^T} \leq \mathbf{0} \quad t\in[T], \xi^T \in \Xi^T. && \text{(28f)}
\end{aligned}
$$

Merging the two maximization problems in (24), we get the monolithic *bilinear* program:

$$
\begin{aligned}
\bar{\nu}_R^{\text{NA-DO}} = \max \quad & \sum_{t\in[T]} b_t(\xi^T)^\top v_{t\xi^T} + \sum_{t\in[T]} \sum_{\xi^T\in\Xi^T} d_t(\xi^t)^\top w_{t\xi^T} && \text{(29a)}\\
\text{s.t.} \quad & (24b) - (24g) && \text{(29b)}\\
& (28b) - (28f). && \text{(29c)}
\end{aligned}
$$

There is a large body of research on solution methods for bilinear problems that can be used in solving model (29). Furthermore, many optimization solvers, such as MOSEK (ApS 2022) and

Gurobi (Gurobi Optimization, LLC 2022), offer off-the-shelf alternatives to solve problems of type (29). Further, (29) can be solved heuristically to obtain a valid dual bound. For instance, one could alternate between optimizing over variables ρ and (α, β, γ) (given that all bilinear terms involve variables ρ) to converge towards a local optimal solution.

4.4. Restricted Decomposable NA Dual

In solving problem (24) using the cutting-plane method, we frequently optimize (25) to compute $\mathcal{Q}(\beta)$. This can become computationally demanding when there is a large number of realizations. Further, we cannot decompose $\mathcal{Q}(\beta)$ by realizations, as they are linked through the z variables. In this section, we present an alternative NA reformulation of the MSARO problem that can yield, through its associated DO problem, a potentially weaker bound than the one provided by (24). However, this alternative reformulation offers a computational advantage since in the framework of the cutting-plane method it leads to decomposable subproblems when calculating $\mathcal{Q}(\beta)$. To this end, in addition to the decision variable copies $y_t(\xi^T)$, we introduce copy variables $z(\xi^T)$ and explicitly enforce them to be equal via nonanticipativity constraints. For an assigned probability measure $\mathbb{P} \in \mathcal{P}^>$, the alternative NA reformulation of the MSARO problem (2) is:

$$\begin{aligned}
\min \quad & \mathbb{E}_{\xi^T \sim \mathbb{P}} \left[z(\xi^T) \right] && && (30a)\\
\text{s.t.} \quad & \sum_{t \in [T]} c_t(\xi^t)^\top y_t(\xi^T) \le z(\xi^T) && \xi^T \in \Xi^T && (30b)\\
& z(\xi^T) = \mathbb{E}_{\xi'^T \sim \mathbb{P}} \left[z(\xi'^T) \right] && \xi^T \in \Xi^T && (30c)\\
& (12c), (12d), (12f), (13). && && (30d)
\end{aligned}$$

Together, (30b) and (30c) capture the semantics of the worst-case outcome, which is minimized in the objective function (30a). We remark that constraints (30c) are obtained from individual nonanticipativity constraints similarly to the derivation of constraints (13) provided in Lemma 1.

Relaxation of the nonanticipativity constraints (13) and (30c) with the Lagrangian multipliers $\lambda_t^y(\cdot)$ and $\lambda^z(\cdot)$ such that $\mathbb{E}_{\xi^T \sim \mathbb{P}}[\lambda_t^y(\xi^T)] < +\infty$ and $\mathbb{E}_{\xi^T \sim \mathbb{P}}[\lambda_t^z(\xi^T)] < +\infty$, leads to the decomposable NA Lagrangian dual problem $\mathcal{L}^{\text{DNA}}(\mathbb{P}) = \max_{\lambda^y(\cdot), \lambda^z(\cdot)} \mathcal{L}_{\text{LR}}^{\text{DNA}}(\mathbb{P}, \lambda_1^y(\cdot), \dots, \lambda_T^y(\cdot), \lambda^z(\cdot))$, where

$$\begin{aligned}
\mathcal{L}_{\text{LR}}^{\text{DNA}}(\mathbb{P}, \lambda_1^y(\cdot), \dots, \lambda_T^y(\cdot), \lambda^z(\cdot)) = \min \quad & \mathbb{E}_{\xi^T \sim \mathbb{P}} \left[\left(1 + \lambda^z(\xi^T) - \mathbb{E}_{\xi'^T \sim \mathbb{P}} \left[\lambda^z(\xi'^T) \right] \right) z(\xi^T) \right] + \\
& \sum_{t \in [T]} \mathbb{E}_{\xi^T \sim \mathbb{P}} \left[\left(\lambda_t^y(\xi^T) - \mathbb{E}_{\xi'^T \sim \mathbb{P}} \left[\lambda_t^y(\xi'^T) | \xi'^t = \xi^t \right] \right)^\top y_t(\xi^T) \right] \\
\text{s.t.} \quad & \left(z(\xi^T), y_1(\xi^T), \dots, y_T(\xi^T) \right) \in Y(\xi^T) \qquad \xi^T \in \Xi^T,
\end{aligned}$$

with $Y(\xi^T)$ the scenario feasibility space described by constraints (30b), (12c), (12d), and (12f). We note that, in the first expectation term in the objective function the inner expectation is not conditional, since constraints (30c) (and accordingly their associated dual functions) are not defined at every stage $t \in [T]$ as decision variable $z(\xi^T) \in \mathbb{R}$ captures the cost of an entire realization ξ^T. Since in $\mathcal{L}_{\mathrm{LR}}^{\mathrm{DNA}}(\cdot)$, the objective function, constraints and variables are decomposable in ξ^T they can be optimized individually.

In deriving the decomposable NA Lagrangian relaxation problem $\mathcal{L}_{\mathrm{LR}}^{\mathrm{DNA}}(\cdot)$ we apply Lemma 1 of (Daryalal et al. 2024) to obtain both expectation terms. We remark that the lemma requires the condition $\mathbb{E}_{\xi^T \sim \mathbb{P}}[\mathrm{diam}(\mathrm{proj}_{z(\cdot)} Y(\xi^T))] < +\infty$ which is not naturally satisfied. However, as a result of Assumptions 1-3 and the compactness of the uncertainty set Ξ^T the optimal value of the MSARO problem is bounded. As such the functionals $z(\cdot)$ can be artificially bounded without changing the optimal value of $\mathcal{L}_{\mathrm{LR}}^{\mathrm{DNA}}(\cdot)$.

After substituting the decision rules $\lambda_t^y(\xi^T) = \Psi_t(\xi^T)\alpha_t^y, t \in [T]$ and $\lambda^z(\xi^T) = \Psi_T(\xi^T)\alpha^z$ in $\mathcal{L}^{\mathrm{DNA}}(\mathbb{P})$, where $\alpha^z \in \mathbb{R}^{K_T}$, and merging with the optimization over the probability distributions $\mathbb{P} \in \mathcal{P}^{\geq}$ the decomposable DO problem is:

$$\bar{\nu}_R^{\text{DNA-DO}} := \max_{\mathbb{P} \in \mathcal{P}^{\geq}, \alpha_1^y, \ldots, \alpha_T^y, \alpha^z} \mathcal{L}_{\mathrm{LR}}^{\mathrm{DNA}}(\mathbb{P}, \Psi_1(\xi^T)\alpha_1^y, \ldots, \Psi_T(\xi^T)\alpha_T^y, \Psi_T(\xi^T)\alpha^z).$$

This DO problem can be solved using the same methods developed in the previous sections for the non-decomposable DO problem. Due to the relaxation of the nonanticipativity constraints on $z(\cdot)$ variables, for given $\mathbb{P}$ and $\lambda^y(\cdot)$, subproblem $\mathcal{L}_{\mathrm{LR}}^{\mathrm{DNA}}(\cdot)$ is a relaxation of $\mathcal{L}_{\mathrm{LR}}^{\mathrm{NA}}(\cdot)$. Therefore $\bar{\nu}_R^{\text{DNA-DO}} \leq \bar{\nu}_R^{\text{NA-DO}}$, i.e., $\bar{\nu}_R^{\text{NA-DO}}$ is a potentially stronger bound. However, the fact that $\mathcal{L}_{\mathrm{LR}}^{\mathrm{DNA}}(\cdot)$ is decomposable is highly desirable. Particularly for continuous or large discrete uncertainty sets where we rely on sampling, and the quality of the bound varies based on the selected sample. Since the decomposable model can afford samples of larger sizes it can potentially yield better bounds compared to the bound obtained from non-decomposable model over a smaller sample. We explore this trade-off in our numerical section.

5. Numerical Experiments

We evaluate the performance of the proposed bounding framework over multistage versions of three classical decision-making problems under uncertainty: (*i*) the newsvendor problem, (*ii*) the location-transportation problem, and (*iii*) the capital budgeting problem. Depending on their structure, each problem is solved by using the appropriate models and methods described in Sections 3 and 4, illustrating the applicability of the developed concepts to a large array of problem classes.

5.1. Benchmarks and Implementation Details

To assess the quality of the primal and dual bounds, we measure the relative distance of the bounds from the true optimal value when an exact solution of MSARO is available (in small-size instances).

Otherwise, we report the optimality gap between the bounds obtained from the proposed methods, and compare it against a gap from traditional bounding methods if one exists. In problems with continuous recourse, we consider LDRs (i.e., their application to all decision variables) as the benchmark for the primal decision rules. On the dual side, we use the perfect information (PI) bound (denoted by ν^{PI}) for comparison, which often can be conveniently evaluated for a general MSARO problem, as well as the bound obtained by solving model (2) using only the binding realizations identified from the primal decision rule solution. The PI bound corresponds to the optimal objective value of the MSARO problem reformulated as in (12) without the nonanticipativity constraints (12e), i.e., it finds the cost of every realization in the uncertainty set individually, and then selects the one with the worst-case cost.

The algorithms are implemented in Python and use the Gurobi Optimizer 9.5.1 (Gurobi Optimization, LLC 2022) as the mixed-integer/bilinear programming solver. The computational experiments are carried out on the Niagara supercomputer servers (Loken et al. 2010, Ponce et al. 2019). For instances with discrete uncertainty sets, the programs have a time limit of 1 hour. We report the valid lower/upper bound at the point of termination. For instances with continuous uncertainty sets, this time limit is extended to 10 hours. As a common design choice for the basis functions of the LDRs, we use the uncertain parameters themselves, i.e., the standard basis functions. Any implementation nuances and enhancements used for improving the performance of the algorithms are discussed for each problem class in a dedicated section, along with the characteristics of the studied instances.

5.2. Robust Newsvendor Problem

In this section, we extend the two-stage newsvendor problem studied in (Xu and Hanasusanto 2021) to the multistage setting. In this problem, a decision-maker (the newsvendor) needs to order from a set of items to be sold (only) at the next decision stage, with the objective of maximizing the worst-case profit over the planning horizon. Let $d_{it}(\xi^t)$ be the uncertain demand of item $i \in [I]$ at stage $t \in [2,T]$, c_i and s_i the purchase and shortage costs of item i, respectively, and r_i its sale price. To meet the customers' demands of stage t, at stage $t-1$ the decision-maker decides on the amounts to be ordered from each item, such that the total spending over the T stages does not exceed a predetermined budget of B. Denote by $x_{it}(\xi^t)$ the decision variable for the amount of item i ordered at stage $t \in [T-1]$. The multistage multi-item budgeted newsvendor problem is:

$$\max \quad z \tag{32a}$$

$$\text{s.t.} \quad z \leq \sum_{i\in[I]} \sum_{t\in[2,T]} y_{it}(\xi^t) \qquad \xi^T \in \Xi^T \tag{32b}$$

$$y_{it}(\xi^t) \leq \big(r_i - c_i\big)x_{i,t-1}(\xi^{t-1}) - r_i\big(x_{i,t-1}(\xi^{t-1}) - d_i(\xi^t)\big) \qquad i \in [I], t \in [2,T], \xi^t \in \Xi^t \tag{32c}$$

$$y_{it}(\xi^t) \leq \big(r_i - c_i\big)x_{i,t-1}(\xi^{t-1}) - s_i\big(d_i(\xi^t) - x_{i,t-1}(\xi^{t-1})\big) \qquad i \in [I], t \in [2,T], \xi^t \in \Xi^t \tag{32d}$$

$$\sum_{i\in[I]} \sum_{t\in[T-1]} x_{it}(\xi^t) \leq B \qquad \xi^{T-1} \in \Xi^{T-1} \tag{32e}$$

$$x_t(\xi^t) \in \mathbb{R}^I_+ \qquad t \in [T-1], \xi^t \in \Xi^t. \tag{32f}$$

where auxiliary variable $y_{it}(\xi^t)$ captures the profit from item i at stage t, by means of (32c)- (32d). Constraints (32e) impose a budget of B over the order amounts throughout the planning horizon. The objective function is the worst-case profit of the newsvendor, modeled via (32a) and (32b).

5.2.1. Problem Instances Our instance generation loosely follows the procedure described in (Ardestani-Jaafari and Delage 2021) for the two-stage robust newsvendor problem. Parameters r_i, s_i and c_i are drawn uniformly from the intervals $[140,160]$, $[80,90]$ and $[50,70]$, respectively. We consider a discrete uncertainty set modeled as a stagewise-dependent scenario tree with branching factor BR (i.e., every node of the tree prior to the leaves has BR many child nodes). Demand realizations $d_{it}(\xi^t), i \in [I], t \in [2,T]$ at a child node are drawn uniformly from $[\mu_{it} - \sigma_{it}, \mu_{it} + \sigma_{it}]$, where μ_{it} and σ_{it} are uniformly drawn from the intervals $[20,40]$ and $[10,20]$. We have generated 26 small-size instances with $T \in [3,5]$, $I \in [2,5]$, $B \in \{100,150,200,250,300\}$, and $\text{BR} \in \{2,3,4,5,10\}$, such that the number of realizations $|\Xi^T| = \text{BR}^{T-1}$ is less than 150. Additionally, we have generated 18 large-size instances with $T \in [4,8]$. For $T = 4$, the number of items I lies in the set $[3,5]$, with a budget $B \in \{200,300\}$ and $\text{BR} \in \{10,15,20\}$. For $T \in [5,8]$, our instances have $I \in \{3,4\}$ items, budget of $B \in \{300,400,500,600\}$, and $\text{BR} \in [3,6]$, restricted to the cases with $|\Xi^T| \leq 3000$.

5.2.2. Quality of the Bounds The small-size instances are easily optimized by solving model (12) over all realizations in the uncertainty set. In our case, this computation takes less than 3 seconds. Using these optimal values, we can examine the quality of a primal/dual bound by measuring its relative distance to the optimal objective value $\nu^\star$. For this problem, all primal and dual problems are solved by the extensive form (i.e., monolithic) of their respective models. Figures 5 and 6 present the gap between the bound and the optimal value of the exact solution, defined as $100\big(\frac{\nu^\star - \nu^{(\cdot)}}{\nu^\star}\big)$ and $100\big(\frac{\nu^{(\cdot)} - \nu^\star}{\nu^\star}\big)$ for primal and dual bounds, respectively, and presented as a percentage (detailed results along with solution times are given in Appendix D.1). In each figure, the solid bars depict the performance of the newly proposed bounds, while the hatched bars represent the benchmarks. The results show that $\nu^{\text{2S-LDR}}$ and $\bar{\nu}_R^{\text{NA-DO}}$ outperform the benchmark bounds by orders of magnitude. More precisely, $\nu^{\text{2S-LDR}}$ on average achieves 84% improvement over ν^{LDR}, with reductions in relative distance ranging from 64% to 98%. Interestingly, the quality of the bound $\nu^{\text{2S-LDR}}$ remains rather stable with changes in the number of items I and budget B, compared to the drastic changes of ν^{LDR} with variations in the inputs. The notable performance of

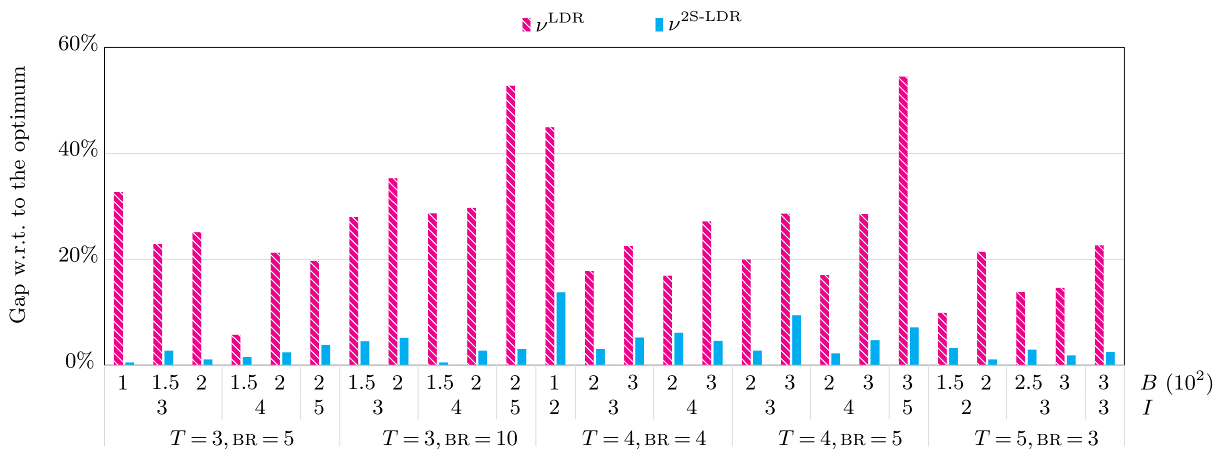

Figure 5 Quality of the primal bounds from LDRs and two-stage LDRs for small-size newsvendor instances

the two-stage LDRs for the newsvendor problem contrasted with the LDRs can be explained by the nature of the recourse variables. In model (32), $y_{it}(\xi^t)$ determines the profit of a given realization at stage t for item i, which for the newsvendor problem is by definition nonlinear. In fact, $y_{it}(\xi^t)$ is an auxiliary variable, defined to linearize the following net profit at stage t from item i:

$$r_i \min\left\{x_{i,t-1}(\xi^{t-1}), d_{it}(\xi^t)\right\} - c_i x_{i,t-1}(\xi^{t-1}) - s_i \max\left\{d_{it}(\xi^t) - x_{i,t-1}(\xi^{t-1}), 0\right\}.$$

Therefore, for the newsvendor problem LDRs *always* return suboptimal decisions as they restrict the form of the nonlinear profit function to be affine, while two-stage LDRs allow them to take any form, giving them an immediate advantage over LDRs.

From the dual perspective, $\bar{\nu}_R^{\mathrm{NA\text{-}DO}}$ achieves an average improvement of 55% compared to ν^{PI}, and in 9 instances fully closes the gap. From the results of Figure 6, a common observation is that LDDRs return a bound of higher quality for smaller values of the ratio $\frac{B}{(T-1)\times I}$, which is an

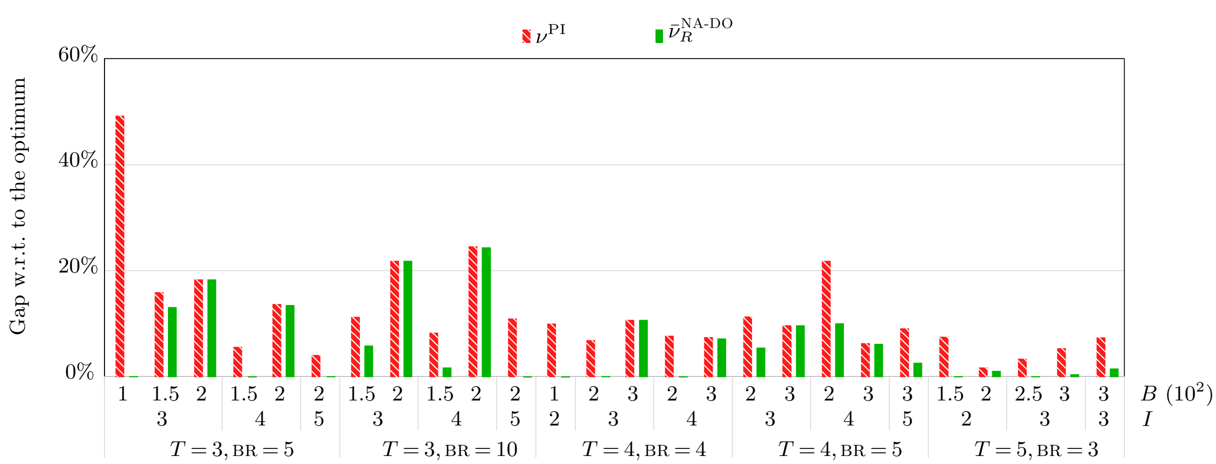

Figure 6 Quality of the dual bounds from the PI and LDDRs for small-size newsvendor instances

estimate of the average budget available per product at every stage. This trend suggests that the restricted NA dual bound performs better with tighter budget, when there is a higher dependency between stages, in which case the importance of intermediate decisions becomes more pronounced.

5.2.3. Optimality Gap For instances with larger number of realizations, we compare the optimality gap from the benchmark methods ($\text{OPT}^{\mathtt{T}}$) with the gap obtained by applying the newly proposed methods, namely two-stage LDRs and LDDRs ($\text{OPT}^{\mathtt{N}}$):

$$\text{OPT}^{\mathtt{T}} = 100\Big(\frac{\nu^{\text{PI}} - \nu^{\text{LDR}}}{\nu^{\text{LDR}}}\Big), \quad \text{OPT}^{\mathtt{N}} = 100\Big(\frac{\bar{\nu}_R^{\text{NA-DO}} - \nu^{\text{2S-LDR}}}{\nu^{\text{2S-LDR}}}\Big).$$

Table 1 presents the optimality gaps from the benchmark and proposed models (whose running times are provided in Appendix D.1). Achieving an average gap reduction of 81.6%, there is a

Table 1 Optimality gaps for larger instances of the newsvendor problem

Instance	T	BR	$\lvert\Xi^T\rvert$	I	B	Primal Bounds ν^{LDR}	Primal Bounds $\nu^{\text{2S-LDR}}$	Dual Bounds $\bar{\nu}_R^{\text{NA-DO}}$	Dual Bounds ν^{PI}	Optimality Gap $\text{OPT}^{\mathtt{N}}$	Optimality Gap $\text{OPT}^{\mathtt{T}}$	Gap reduction
1	4	10	1000	3	200	5648.8	8142.9	9002.8	9353.0	10.6%	65.6%	83.9%
2				3	300	9143.5	13687.6	17853.0	17853.0	30.4%	95.3%	68.1%
3				4	200	-104.4	440.0	642.4	919.0	46.0%	980.4%	95.3%
4				4	300	11029.5	15854.7	18432.0	18432.0	16.3%	67.1%	75.8%
5				5	300	724.9	6125.5	6614.0	7368.0	8.0%	916.5%	99.1%
6	4	15	3375	3	200	5111.3	7072.0	8222.0	8222.0	16.3%	60.9%	73.3%
7	4	20	8000	3	300	9030.9	13040.7	17615.0	17615.0	35.1%	95.1%	63.1%
8	5	5	625	3	300	11134.5	15053.4	16216.4	16680.0	7.7%	49.8%	84.5%
9				4	300	3651.5	9494.8	10403.1	10628.0	9.6%	191.1%	95.0%
10	5	6	1296	3	400	15013.4	20271.7	25157.0	25157.0	24.1%	67.6%	64.3%
11	6	4	1024	3	400	15492.3	21457.9	28163.0	28163.0	31.2%	81.8%	61.8%
12				4	400	7124.3	14805.2	15473.0	15473.0	4.5%	117.2%	96.2%
13				4	500	14445.4	24887.7	32973.0	32973.0	32.5%	128.3%	74.7%
14	7	3	729	3	300	24.3	1495.9	2150.0	2383.0	43.7%	9692.2%	99.5%
15				3	400	12994.7	17774.1	19554.5	19983.0	10.0%	53.8%	81.4%
16				4	400	3267.5	4303.1	5118.1	5427.0	18.9%	66.1%	71.3%
17	8	3	2187	3	500	16149.1	25321.4	30259.0	30259.0	19.5%	87.4%	77.7%
18				4	600	16892.3	27608.1	28747.0	28747.0	4.1%	70.2%	94.1%

considerable value in using the two-stage LDRs and LDDRs in devising policies for the multistage newsvendor problem. The optimality gaps $\text{OPT}^{\mathtt{N}}$ range from 4% to 46%. There are some interesting cases such as instance 3 where LDR policies can even lead to a profit loss in the worst case. In the majority of the instances, both two-stage LDRs and LDDRs contribute to the improvement of the gap, although the primal side clearly has the larger impact. For instance, in five instances, the dual bound $\bar{\nu}_R^{\text{NA-DO}}$ and the benchmark ν^{PI} are the same. In some instances, such as instances 8, 12 and 18, the PI bound might already be strong enough so that using LDDRs does not make a tangible difference in strengthening it.

Lastly, for the primal bounds, we observe that the solution times for obtaining the traditional and two-stage LDR bounds are comparable. For the dual bounds, the times for obtaining the PI bound are quite small whereas for our dual problem the solution time significantly increases with problem size with several instances not being solved to optimality within the 1-hour time limit.

5.3. Robust Location-Transportation Problem

The two-stage robust location-transportation problem studied by Zeng and Zhao (2013) is as follows. Given a set of I potential facilities with building cost f_i and unit capacity cost $a_i, i \in [I]$, we have to meet the uncertain demand of a set of customers J with unit transportation cost $c_{ij}, i \in [I], j \in [J]$. The goal is to decide which facilities to open and their initial capacities, such that the worst-case total cost of facility deployment and future transportation is minimized. Letting $d_{jt}(\xi^t)$ be the demand of customer j at stage t, we define the MSARO extension of the problem:

$$
\begin{aligned}
\min \quad & z && && \text{(33a)}\\
\text{s.t.} \quad & z \geq \sum_{i\in[I]}(f_i y_i + a_i s_{i1}) + \sum_{t\in[2,T]}\sum_{i\in[I]}\Big(a_i s_{it}(\xi^t) + \sum_{j\in[J]} c_{ij} x_{ijt}(\xi^t)\Big) && \xi^T \in \Xi^T && \text{(33b)}\\
& s_{i1} \leq K_i y_i && i \in [I] && \text{(33c)}\\
& s_{it}(\xi^t) = s_{i1} - \sum_{j\in[J]} x_{ijt}(\xi^t) && i \in [I], t = 2, \xi^t \in \Xi^t && \text{(33d)}\\
& s_{it}(\xi^t) = s_{i,t-1}(\xi^{t-1}) - \sum_{j\in[J]} x_{ijt}(\xi^t) && i \in [I], t \in [3,T], \xi^t \in \Xi^t && \text{(33e)}\\
& \sum_{i\in[I]} x_{ijt}(\xi^t) \geq d_{jt}(\xi^t) && j \in [J], t \in [2,T], \xi^t \in \Xi^t && \text{(33f)}\\
& y \in \{0,1\}^I, \; s_1 \in \mathbb{R}^I_+ && && \text{(33g)}\\
& s_t(\xi^t) \in \mathbb{R}^I_+, \; x_t(\xi^t) \in \mathbb{R}^{I\times J}_+ && t \in [2,T], \xi^t \in \Xi^t, && \text{(33h)}
\end{aligned}
$$

where y_i is a binary variable equal to 1 if facility i is built, s_{i1} determines the initial capacity of facility i, $s_{it}(\xi^t)$ is the state variable calculating the remaining capacity of facility i at stage t, and $x_{ijt}(\xi^t)$ is the amount of goods transported from facility i to customer j at stage t. Objective function (33a) together with constraint (33b) measures the worst-case cost. Constraints (33c) bound the initial capacity of the facilities, whereas constraints (33d)-(33e) are the state equations calculating the remaining capacities. Constraints (33f) ensure that the customer demands are met.

5.3.1. Problem Instances Our instances are generated using the instance parameters described in (Zeng and Zhao 2013) for the two-stage problem. For number of stages $T \in [3,5]$, we have five combinations for the number of facilities and customers, $(I,J) \in \{(5,5),(5,7),(5,10),(10,10),(20,20)\}$. Fixed installation, unit capacity, and unit transportation costs are drawn from $f_i \in [100, 1000], a_i \in [10, 100], c_{ij} \in [1, 1000]$, respectively. Maximal capacity is set to $K_i = 2 \times 10^4$ based on preliminary experiments to make sure the instances are feasible. Customer demands are random parameters with support $\left[\mu_{jt}, (1+\alpha^d)\mu_{jt}\right]$, where $\mu_{jt} \in [10, 500]$ and $\alpha^d \in \{0.1, 0.3, 0.5\}$ are given, with α^d a parameter controlling the variation among demand realizations of customer j. For this problem, we have generated 32 instances with $\alpha^d = 0.5$ over small scenario trees to compare the bounds with the optimal objective value. In addition, we have generated 83 instances with demands $d_{jt} = \mu_{jt} + \xi_{jt}\sigma_{jt},\ j \in [J],\ t \in [2, T]$, where ξ belongs to the following budgeted uncertainty set:

$$\Xi^T = \left\{ \xi \in \mathbb{R}_+^{J \times T-1} \;\middle|\; \xi_{jt} \in [0,1],\ j \in [J],\ t \in [2,T],\ \sum_{t \in [2,T]} \sum_{j \in [J]} \xi_{jt} \le \Gamma \right\}.$$

Parameter Γ correlates the demands of all customers and stages together, which results in stagewise (temporal) dependence between the decision stages. We use $\Gamma = \alpha^u I$, with $\alpha^u \in \{0.1, 0.4, 0.7, 1\}$.

5.3.2. Scenario-Tree Instances In solving the primal (2S-LDR) and exact models of the scenario-tree instances, we have used their respective extensive forms, while for computing our dual bound we implemented the cutting-plane method described in Section 4.3.

Table 2 presents the gap between the bound and the optimal value of the exact solution, $100\big(\frac{\nu^{(\cdot)} - \nu^\star}{\nu^\star}\big)$ and $100\big(\frac{\nu^\star - \nu^{(\cdot)}}{\nu^\star}\big)$ for the primal and dual problems, respectively. Results show that, among the group of instances with similar characteristics, as the branching factor or the number of stages increases, the ν^{LDR} bound gets noticeably worse. In contrast, $\nu^{\text{2S-LDR}}$ stays very close to the

Table 2 Quality of the bounds for the scenario-tree instances of the location-transportation problem

T	I	J	BR	$\lvert\Xi^T\rvert$	ν^{LDR}	$\nu^{\text{2S-LDR}}$	$\bar{\nu}_R^{\text{NA-DO}}$	ν^{PI}
3	5	10	3	9	2.9%	0.0%	1.4%	4.7%
			4	16	0.8%	0.0%	3.3%	3.4%
			5	25	4.6%	0.0%	2.5%	13.2%
			6	36	5.6%	0.0%	17.1%	27.5%
			7	49	8.2%	0.0%	9.4%	9.4%
			8	64	8.8%	0.0%	4.8%	14.6%
			9	81	6.5%	0.0%	6.1%	6.1%
			10	100	6.6%	0.0%	8.0%	10.2%
3	10	10	3	9	2.3%	0.0%	2.5%	13.2%
			4	16	2.7%	0.1%	1.2%	8.4%
			5	25	6.1%	0.0%	12.0%	19.3%
			6	36	9.8%	0.0%	10.4%	22.6%
			7	49	13.7%	0.1%	6.7%	14.1%
			8	64	13.9%	0.3%	5.5%	5.5%
			9	81	12.2%	0.2%	15.5%	15.5%
			10	100	12.3%	0.2%	10.4%	19.0%
3	20	20	3	9	2.5%	0.7%	9.5%	11.3%
			4	16	3.9%	0.5%	7.0%	7.0%
			5	25	5.8%	0.5%	28.4%	28.4%
			6	36	6.6%	0.5%	13.0%	19.2%
			7	49	9.1%	0.5%	13.3%	31.3%
4	5	10	3	27	8.0%	0.0%	2.6%	8.0%
			4	64	6.8%	0.0%	2.0%	2.0%
			5	125	13.5%	0.1%	5.0%	5.0%
4	10	10	3	27	11.9%	0.1%	17.3%	17.3%
			4	64	13.3%	0.0%	10.2%	21.0%
			5	125	13.1%	0.4%	10.0%	10.0%
4	20	20	3	27	9.0%	0.9%	11.7%	18.1%
			4	64	11.9%	0.5%	9.8%	21.7%
5	5	10	3	81	14.7%	1.0%	4.4%	6.3%
			4	256	13.9%	0.5%	11.7%	19.1%
5	10	10	3	81	16.8%	0.7%	8.2%	8.2%

optimal value, with an average relative distance of 0.2% among all the instances (compared to 8.7% for ν^{LDR}). We do not observe the same trend for the dual bounds, and their relative distance to the optimal value fluctuates even between two instances that only differ in the number of realizations. Nevertheless, $\bar{\nu}_R^{\text{NA-DO}}$ considerably outperforms PI, with an average improvement of 36.2%.

5.3.3. Budgeted-Uncertainty Instances For the instances with the budgeted uncertainty set, we use, as benchmarks, ν^{LDR} as an upper bound and $\nu^{\Omega(\text{LDR})}$ as a lower bound, obtained by solving problem (33) using only the binding realizations identified from the benchmark primal solution. We calculate $\nu^{\text{2S-LDR}}$ using the C&CG method described in Section 3.2 (detailed models are given in Appendix C.3.1). Similar to the scenario-tree instances, we use the cutting-plane method to obtain the dual bound $\bar{\nu}_R^{\text{NA-DO}}$. To do so, we use a sample of size at least $50(T-2)$, which includes both the binding realizations from the two-stage LDR solution and additional randomly generated realizations from the uncertainty set. For both algorithms, we stop when the optimality gap of the method falls below 5% or we reach the 10-hour time limit. Table 3 presents the optimality gaps between the benchmark bounds and the proposed bounds, respectively, for the 83 instances considered. Detailed results for each instance are provided in Appendix D.2. Our methods return an average optimality gap of 17.3% across all instances, compared to an average gap of 27% from the benchmarks. Given the strength of the two-stage LDR bound observed in the scenario-tree instances, and its resilience to increases in the size of the tree, it is likely that the dual bounds are further from the optimal value.

Figure 7 illustrates the improvement achieved from $\bar{\nu}_R^{\text{NA-DO}}$ over the benchmark dual bound, $\nu^{\Omega(\text{LDR})}$. Each bar in the figure represents an individual instance. Instances are grouped into blocks based on their shared parameters T, (I, J), and α^d, and within each block, instances are arranged in ascending order of their α^u values. Across all instances, $\bar{\nu}_R^{\text{NA-DO}}$ demonstrates an average improvement of 7%. This figure also highlights that improved identification of critical realizations, guided by the two-stage LDR solution, and the subsequent solution of the discretized relaxation of the problem using these realizations, independently contribute to an average 3.6% enhancement of the dual bound. A key consideration here is that the value of $\bar{\nu}_R^{\text{NA-DO}}$ depends on the sample used for its computation. An insufficiently large sample can lead to a poor bound. However, because of the bilinear form of the cutting-plane master problem, we were not able to solve the model with large samples using off-the-shelf commercial solvers. By employing specialized algorithms developed for bilinear problems, it could be possible to increase the size of the sample and improve the $\bar{\nu}_R^{\text{NA-DO}}$ bound, which we leave for future research.

Analyzing the performance of the C&CG and cutting-plane algorithms in solving our instances provides further insight into the quality of the bounds. Solution times, number of iterations, and

Table 3 Optimality gaps for the budgeted-uncertainty instances, calculated using both benchmark and proposed bounding methods. For this sampling instance, OPT^{T} is defined as the gap between ν^{LDR} and $\nu^{\Omega(\text{LDR})}$. As before, OPT^{N} represents the gap between $\nu^{\text{2S-LDR}}$ and $\bar{\nu}_R^{\text{NA-DO}}$.

(T,I,J)	α^d	α^u	OPT^{T}	OPT^{N}
(3,10,10)	0.1	0.1	9.0%	7.4%
		0.4	20.8%	12.4%
		0.7	32.0%	22.5%
		1	30.6%	22.2%
	0.3	0.1	15.3%	12.0%
		0.4	27.0%	16.7%
		0.7	30.3%	21.5%
		1	29.1%	21.3%
	0.5	0.1	17.8%	12.9%
		0.4	28.3%	15.1%
		0.7	32.7%	28.4%
		1	26.8%	7.4%
(3,10,15)	0.1	0.1	32.8%	24.4%
		0.4	31.5%	27.1%
		0.7	12.1%	7.2%
		1	31.5%	17.3%
	0.3	0.1	12.8%	8.3%
		0.4	35.3%	25.6%
		0.7	17.2%	5.7%
		1	26.2%	13.6%
	0.5	0.1	32.4%	28.2%
		0.4	25.8%	15.3%
		0.7	29.4%	18.2%
		1	29.7%	20.7%
(4,5,5)	0.1	0.1	14.3%	6.0%
		0.4	33.2%	25.7%
		0.7	27.7%	10.8%
		1	12.7%	5.6%
	0.3	0.1	13.5%	7.8%
		0.4	25.3%	18.1%
		0.7	23.6%	13.8%
		1	26.2%	9.6%
	0.5	0.1	31.8%	27.1%
		0.4	22.6%	14.1%
		0.7	29.8%	19.6%
		1	29.8%	5.4%
(4,5,7)	0.1	0.1	15.7%	13.6%
		0.4	24.0%	19.4%
		1	38.3%	24.5%
	0.3	0.1	23.2%	19.3%
		0.4	30.8%	21.6%
		0.7	10.2%	6.4%
		1	30.7%	18.7%
	0.5	0.1	24.2%	21.5%
		0.4	39.5%	27.1%
		0.7	32.7%	17.9%
		1	41.0%	23.3%

(T,I,J)	α^d	α^u	OPT^{T}	OPT^{N}
(4,5,10)	0.1	0.1	22.4%	16.8%
		0.4	32.1%	28.1%
		0.7	15.2%	10.2%
		1	31.2%	19.0%
	0.3	0.1	28.5%	20.8%
		0.4	37.7%	30.6%
		0.7	42.9%	27.6%
		1	24.3%	14.2%
	0.5	0.1	33.0%	27.7%
		0.4	22.2%	11.4%
		0.7	27.7%	17.9%
		1	33.1%	23.5%
(4,10,10)	0.1	0.1	19.4%	17.5%
		0.4	32.1%	26.6%
		0.7	38.7%	32.7%
		1	21.7%	10.8%
	0.3	0.1	11.1%	6.6%
		0.4	19.8%	16.5%
		0.7	22.6%	17.4%
		1	38.1%	28.0%
	0.5	0.1	24.9%	12.4%
		0.4	23.3%	11.0%
		0.7	31.4%	18.2%
		1	31.2%	25.7%
(5,5,10)	0.1	0.1	17.2%	13.7%
		0.4	24.5%	20.6%
		0.7	34.2%	21.8%
		1	20.3%	8.6%
	0.3	0.1	28.9%	25.6%
		0.4	32.5%	19.3%
		0.7	39.1%	31.9%
		1	37.5%	29.4%
	0.5	0.1	36.6%	31.8%
		0.4	39.8%	26.6%
		0.7	39.2%	17.6%
		1	41.6%	27.1%

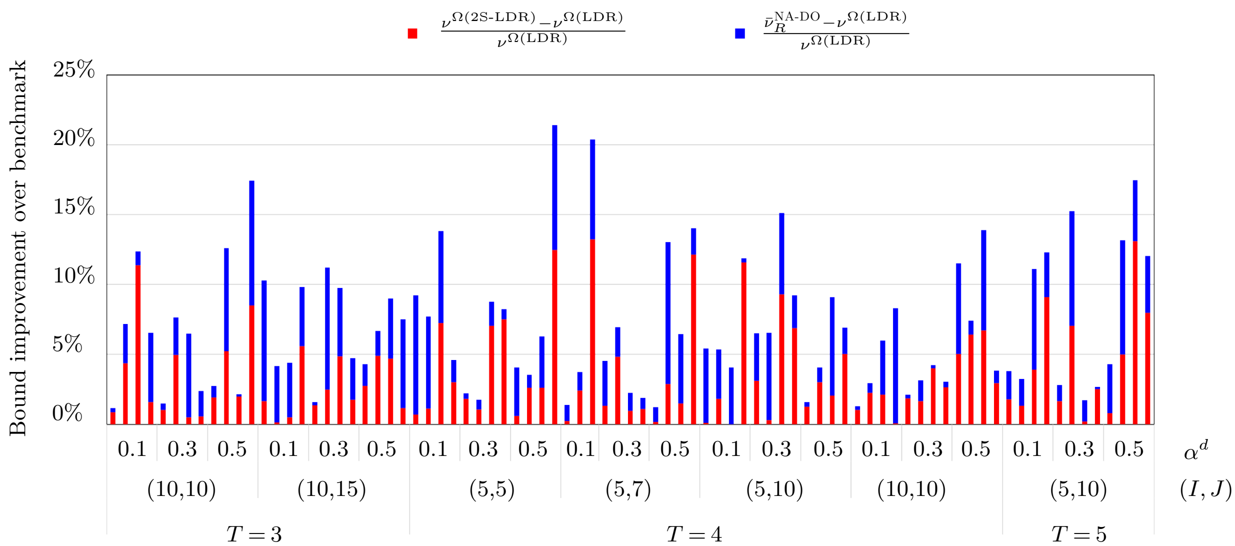

Figure 7 Improvement over the benchmark bound $\nu^{\Omega(\text{LDR})}$ in the the budgeted-uncertainty instances. Here, $\nu^{\Omega(\text{LDR})}$ and $\nu^{\Omega(\text{2S-LDR})}$ are defined as the bounds obtained by solving problem (33) using only the binding realizations of ν^{LDR} and $\nu^{\text{2S-LDR}}$, respectively.

final optimality gaps at the time limit (namely the termination gap) are provided in Appendix D.2. Our findings show that while our methods do require extra computational effort, they yield stronger bounds as a result. Another key aspect is that a large termination gap can negatively impact the quality of the bounds. Our results show that in all instances, the C&CG stops at a solution of the 2ARO model with less than 5% termination gap. In fact, the method proves to be quite powerful in detecting the significant realizations for the 2ARO approximation, such that in 69 out of 83 instances it achieves this gap after only two iterations. On the other hand, in 39 instances the cutting-plane method is not able to reach the optimality gap of 5% within the time limit of 10 hours albeit having less than 10% optimality gap in all instances with one exception of 13.4%. Consequently, for calculating $\text{OPT}^{\mathtt{N}}$, we use the best lower bound on $\bar{\nu}_R^{\text{NA-DO}}$ obtained at the end of 10 hours. This can further contribute to an increased optimality gap. This suggests that, in addition to having a difficult nonlinear master problem, the cutting-plane method itself requires algorithmic enhancements, such as the design of stronger cuts.

5.4. Robust Capital Budgeting Problem

In the capital budgeting problem, a company wants to invest in a subset of I projects with uncertain cost and profit, subject to an initial budget of B that can be increased by getting a loan. A variant of the two-stage problem is studied by Subramanyam et al. (2020). In the following, we formulate the multistage capital budgeting problem as an MSARO. Over a planning horizon of T stages, let $x_{it}(\xi^t)$ be a binary decision variable taking the value of 1 if the company decides to invest in the project $i \in [I]$ at stage $t \in [T]$, with a cost of $c_{it}(\xi^t)$ and profit of $r_{it}(\xi^t)$. Further, let $L_t(\xi^t)$ be a continuous decision variable determining the amount of loan the company decides to get at stage $t \in [T]$ with a unit cost of $c_{\mathrm{L}}\mu^{t-1}, \mu > 1$. The MSARO model is as follows:

$$
\begin{aligned}
\max \quad & z && && \text{(34a)}\\
\text{s.t.} \quad & z \le \sum_{t\in[T]}\sum_{i\in[I]} r_{it}(\xi^t)x_{it}(\xi^t) - \sum_{t\in[T]} c_{\mathrm{L}}\mu^{t-1}L_t(\xi^t) && \xi^T \in \Xi^T && \text{(34b)}\\
& B_t(\xi^t) - B_{t-1}(\xi^{t-1}) + C_{t-1}(\xi^{t-1}) - L_t(\xi^t) = 0 && t \in [T], \xi^t \in \Xi^t && \text{(34c)}\\
& \sum_{i\in[I]} c_{it}(\xi^t)x_{it}(\xi^t) = C_t(\xi^t) && t \in [T], \xi^t \in \Xi^t && \text{(34d)}\\
& B_t(\xi^t) - C_t(\xi^t) \ge 0 && t \in [T], \xi^t \in \Xi^t && \text{(34e)}\\
& x_t(\xi^t) \in \{0,1\}^I, L_t(\xi^t) \in \mathbb{R}_+ && t \in [T], \xi^t \in \Xi^t, && \text{(34f)}
\end{aligned}
$$

where $B_t(\xi^t)$ is the amount of available funds at stage $t \in [T]$, determined by constraints (34c), while $C_t(\xi^t)$ is the expenditure calculated through constraints (34d), with initial values of $B_0(\boldsymbol{\xi}^0) = B$ and $C_0(\boldsymbol{\xi}^0) = 0$. Constraints (34e) bound the expenditure amount by available funds. The objective is to maximize the worst-case profit over the planning horizon, measured by constraints (34b).

5.4.1. Problem Instances Our instance generation follows the procedure of Subramanyam et al. (2020) for the two-stage problem. To incorporate the dynamic nature of multistage capital budgeting, project costs and profits are modeled as affine functions of evolving risk factors $\boldsymbol{\xi}_t, t \in [2,T]$:

$$
c_{it}(\boldsymbol{\xi}_t) = c_{it}^0\left(1 + \boldsymbol{\Phi}_{it}^\top \boldsymbol{\xi}_t/2\right), \qquad r_{it}(\boldsymbol{\xi}_t) = r_{it}^0\left(1 + \boldsymbol{\Psi}_{it}^\top \boldsymbol{\xi}_t/2\right),
$$

where c_{it}^0 and r_{it}^0 represent the baseline cost and profit, respectively, assuming all risk factors are held at their neutral value of zero; $\boldsymbol{\Phi}_{it}$ and $\boldsymbol{\Psi}_{it}$ are factor loading vectors governing the sensitivity of cost and profit to deviations from the neutral state; and c_{i1} and r_{i1} are set to the nominal values. We consider four risk factors at each stage $t \in [2,T]$ such that $\boldsymbol{\xi}_t \in [-1,1]^4$. Thus, $\boldsymbol{\Phi}_{it} \in \mathbb{R}^4$ and $\boldsymbol{\Psi}_{it} \in \mathbb{R}^4$, quantify the influence of each of the four risk factors on the project's financial outcome. When sampling from $\mathbb{R}^4$, we ensure that $\boldsymbol{\Phi}_{it}^\top \mathbf{e} = \boldsymbol{\Psi}_{it}^\top \mathbf{e} = 1$ for all $i \in \mathcal{I}$ and $t \in [2,T]$, where $\mathbf{e}$ is a vector of ones.

For $T \in \{3,4,5\}$ and $I \in \{5,8,10\}$, the nominal cost vector $c_{it}^0, t \in [2,T], i \in [I]$ is drawn uniformly from $[0,10]^I$, and nominal profits are set as $r_{it}^0 = \frac{c_{it}^0}{5}$. Loan purchase cost is $c_L\mu^{t-1} = 0.12(1.2)^{t-1}$ per unit of loan. For each combination of T and I we consider different levels of initial budget which impacts the dependence between stages.

5.4.2. Optimality Gap Due to presence of binary variables $x_t(\xi^t)$, with the existing methods in the literature of 2ARO we cannot calculate $\nu^{\text{2S-LDR}}$ exactly. Therefore, we solve an approximation of it using the K-adaptability method of Subramanyam et al. (2020) with $K=2$ and use the obtained bound ν^{K} in measuring the optimality gap. On the dual side, we study three options: (i) ν^{Ω}, the upper bound from solving the model (34) with a sample $|\Omega| = 250$, (ii) the NA bound $\bar{\nu}_R^{\text{NA-DO}}$ with the same sample Ω, (iii) the decomposable NA bound $\bar{\nu}_R^{\text{DNA-DO}}$ with a sample of size 500, which includes the sample Ω. For each option, when the model is not solved to optimality within the time limit, the best valid bound is used in the calculations. Note that, if we can solve both options (i) and (ii) to optimality, we should expect a better bound from option (i). Figure 8 presents the optimality gaps of the capital budgeting instances between the bound ν^{K} and the three choices of upper bound (detailed results are given in Appendix D.3). Results show that, even though the bound ν^{Ω} should theoretically be at least as good as the bound $\bar{\nu}_R^{\text{NA-DO}}$, on average the NA bound returns a better upper bound within the same time limit. This is a testament to the difficulty of the multistage problem even when it is solved for a discrete set of realizations. The decomposable NA bound $\bar{\nu}_R^{\text{DNA-DO}}$ further improves the results by using a larger sample which is viable because of its superior computational performance. Note that, the best gaps in Figure 8, ranging between 7% to 33%, are obtained from approximations over approximations on both primal and dual side. Accordingly, these rather large gaps can be attributed to both bounding methods.

General Integer Recourse Variables. To assess our algorithms' performance with general integer recourse, we conducted a set of experiments where loan amounts, $L_t(\xi^t)$, were restricted to integer values. Figure 9 in Appendix D.4 presents the same analysis as in Figure 8 for the capital budgeting problem, but with the added constraint of $L_t(\xi^t) \in \mathbb{Z}_+$. Interestingly, requiring integer loan amounts did not significantly alter the optimal investment decisions compared to the continuous case. This explains the visual similarity between the two figures, with both exhibiting similar patterns in optimal decisions. While there are slight differences in objective function values, the overall investment strategies remain largely unaffected by the integrality constraints.

It is important to note that although the advantage of our dual methods appears to diminish with an increasing number of stages, even the initial gap relies on the strength of our primal solution, ν^{K}. Therefore, these figures demonstrate our ability to achieve further improvements beyond the initial strong primal bound.

Appendix D.3 provides a detailed analysis of the computational requirements for bounding the capital budgeting problem, considering both fractional and integral loans. For the primal side, we present the solution times of our approach, as no alternative primal method is available for comparison. On the dual side, we compare the solution times of the three previously discussed bounding

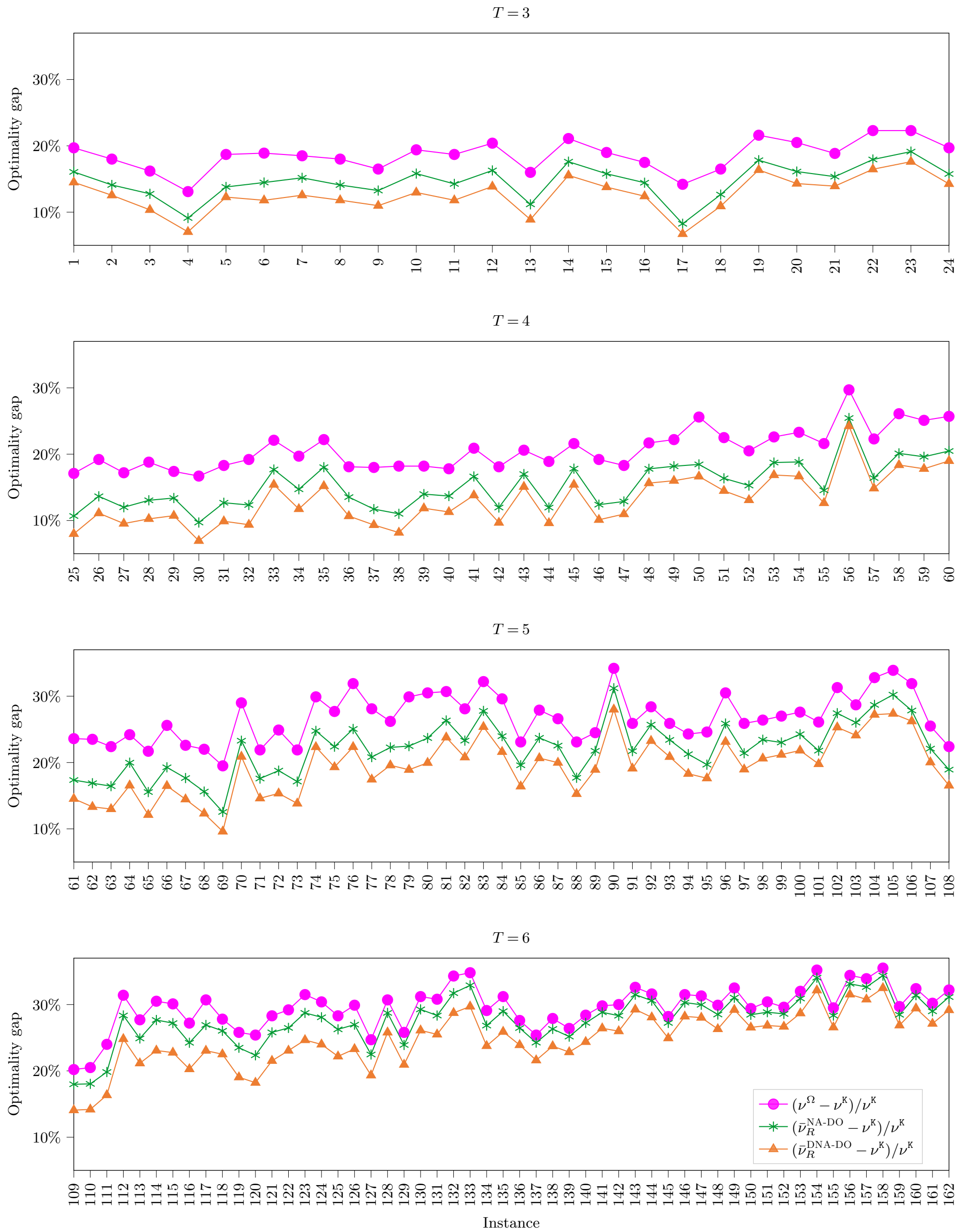

Figure 8 Optimality gap improvements for capital budgeting problems when using LDDR-based methods. In the legend, ν^{Ω} denotes the bound obtained by solving model (34) over a sample set.

approaches. The solution times demonstrate that the extra computational effort associated with the decomposable NA model yields a demonstrably improved dual bound.

6. Conclusion

Robust optimization models are built on a different premise than stochastic programming in the sense that they do not assume any knowledge about the probability distribution, focusing instead on optimizing the worst-case outcomes. In this paper, we study general MSAROs for which we develop primal and dual bounding methods by adapting two decision rule approximations from the stochastic programming literature (namely two-stage LDRs and LDDRs). These approximations allow us to reduce MSARO to a 2ARO from the primal side and a two-stage stochastic programming problem from the dual side. As such, the resulting approximations drastically reduce the theoretical complexity of the studied problems. Since our dual bounds are dependent on the choice of a probability distribution while deriving the dual model, we propose to solve a distribution optimization problem to obtain a stronger bound. We develop various solution methods for our proposed bounding problems where we also leverage existing methods from both the robust optimization and the stochastic programming literature. Our extensive numerical study demonstrates that our methods considerably improve both primal and dual bounds compared to the commonly used approaches in the literature. Our work opens the door to the direct application of existing two-stage robust optimization and stochastic programming algorithms and other future algorithmic developments in these areas to MSAROs. For instance, as the algorithms such as C&CG and K-adaptability for 2ARO improve, our models can be solved more efficiently.

We believe that our work can initiate additional methodological and numerical developments. Both from the primal and the dual side the question of solving the problems we pose in a more numerically efficient manner definitely merits more attention. Further methodological work may also explore the synergies between primal and dual decision rules. Finally, the following directions can be the subject of future research: exploiting problem structure in order to approximate MSAROs with numerically more favorable, e.g, decomposable models, identifying special cases of MSAROs in which the proposed approximations can be proven to be exact, and application of similar approaches in related fields such as distributionally robust optimization.

Acknowledgments

This work was supported by Natural Sciences and Engineering Research Council of Canada [Grants RGPIN-2018-04984 and RGPIN-2024-05908] and Agence Nationale de la Recherche of France [Grant ANR-22-CE048-0018]. Computations were performed on the Niagara supercomputer at the SciNet HPC Consortium. SciNet is funded by: the Canada Foundation for Innovation; the Government of Ontario; Ontario Research Fund - Research Excellence; and the University of Toronto.

References

ApS M (2022) *MOSEK modeling cookbook*.

Ardestani-Jaafari A, Delage E (2021) Linearized robust counterparts of two-stage robust optimization problems with applications in operations management. *INFORMS Journal on Computing* 33(3):1138–1161.

Arslan AN, Detienne B (2022) Decomposition-based approaches for a class of two-stage robust binary optimization problems. *INFORMS Journal on Computing* 34(2):857–871.

Ayoub J, Poss M (2016) Decomposition for adjustable robust linear optimization subject to uncertainty polytope. *Computational Management Science* 13(2):219–239.

Ben-Tal A, El Ghaoui L, Nemirovski A (2009) Robust optimization. *Robust optimization* (Princeton university press).

Ben-Tal A, El Housni O, Goyal V (2020) A tractable approach for designing piecewise affine policies in two-stage adjustable robust optimization. *Mathematical Programming* 182(1):57–102.

Ben-Tal A, Goryashko A, Guslitzer E, Nemirovski A (2004) Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* 99(2):351–376.

Ben-Tal A, Nemirovski A (1999) Robust solutions of uncertain linear programs. *Operations Research Letters* 25(1):1–13.

Bertsimas D, Brown DB, Caramanis C (2011a) Theory and applications of robust optimization. *SIAM review* 53(3):464–501.

Bertsimas D, Caramanis C (2010) Finite adaptability in multistage linear optimization. *IEEE Transactions on Automatic Control* 55(12):2751–2766.

Bertsimas D, Dunning I (2016) Multistage robust mixed-integer optimization with adaptive partitions. *Operations Research* 64(4):980–998.

Bertsimas D, Georghiou A (2015) Design of near optimal decision rules in multistage adaptive mixed-integer optimization. *Operations Research* 63(3):610–627.

Bertsimas D, Georghiou A (2018) Binary decision rules for multistage adaptive mixed-integer optimization. *Mathematical Programming* 167(2):395–433.

Bertsimas D, Goyal V (2012) On the power and limitations of affine policies in two-stage adaptive optimization. *Mathematical Programming* 134(2):491–531.

Bertsimas D, Goyal V, Lu BY (2015) A tight characterization of the performance of static solutions in two-stage adjustable robust linear optimization. *Mathematical Programming* 150(2):281–319.

Bertsimas D, Iancu DA, Parrilo PA (2010) Optimality of affine policies in multistage robust optimization. *Mathematics of Operations Research* 35(2):363–394.

Bertsimas D, Iancu DA, Parrilo PA (2011b) A hierarchy of near-optimal policies for multistage adaptive optimization. *IEEE Transactions on Automatic Control* 56(12):2809–2824.

Bertsimas D, Litvinov E, Sun XA, Zhao J, Zheng T (2012) Adaptive robust optimization for the security constrained unit commitment problem. *IEEE Transactions on Power Systems* 28(1):52–63.

Bertsimas D, Sim M, Zhang M (2019) Adaptive distributionally robust optimization. *Management Science* 65(2):604–618.

Bertsimas D, Thiele A (2006) A robust optimization approach to inventory theory. *Operations Research* 54(1):150–168.

Bodur M, Luedtke JR (2018) Two-stage linear decision rules for multi-stage stochastic programming. *Mathematical Programming* 1–34.

Chen X, Sim M, Sun P, Zhang J (2008) A linear decision-based approximation approach to stochastic programming. *Operations Research* 56(2):344–357.

Chen X, Zhang Y (2009) Uncertain linear programs: Extended affinely adjustable robust counterparts. *Operations Research* 57(6):1469–1482.

Cheramin M, Cheng J, Jiang R, Pan K (2022) Computationally efficient approximations for distributionally robust optimization under moment and wasserstein ambiguity. *INFORMS Journal on Computing* 34(3):1768–1794.

Daryalal M, Bodur M, Luedtke JR (2024) Lagrangian dual decision rules for multistage stochastic mixed-integer programming. *Operations Research* 72(2):717–737.

Dumouchelle J, Julien E, Kurtz J, Khalil EB (2023) Neur2RO: Neural two-stage robust optimization. *arXiv preprint arXiv:2310.04345* .

Geoffrion AM (1974) *Lagrangean relaxation for integer programming*, volume 2, 82–114 (Berlin, Heidelberg: Springer Berlin Heidelberg).

Georghiou A, Tsoukalas A, Wiesemann W (2019) Robust dual dynamic programming. *Operations Research* 67(3):813–830.

Georghiou A, Tsoukalas A, Wiesemann W (2020) A primal–dual lifting scheme for two-stage robust optimization. *Operations Research* 68(2):572–590.

Goerigk M, Lendl S, Wulf L (2024) On the complexity of robust multi-stage problems with discrete recourse. *Discrete Applied Mathematics* 343:355–370.

Goh J, Sim M (2010) Distributionally robust optimization and its tractable approximations. *Operations Research* 58(4-part-1):902–917.

Gurobi Optimization, LLC (2022) Gurobi Optimizer Reference Manual. URL https://www.gurobi.com.

Hadjiyiannis MJ, Goulart PJ, Kuhn D (2011) A scenario approach for estimating the suboptimality of linear decision rules in two-stage robust optimization. *50th IEEE Conference on Decision and Control and European Control Conference*, 7386–7391 (IEEE).

Hanasusanto GA, Kuhn D, Wiesemann W (2015) K-adaptability in two-stage robust binary programming. *Operations Research* 63(4):877–891.

Hashemi Doulabi H, Jaillet P, Pesant G, Rousseau LM (2021) Exploiting the structure of two-stage robust optimization models with exponential scenarios. *INFORMS Journal on Computing* 33(1):143–162.

Hendrix EM, Boglárka G, et al. (2010) *Introduction to nonlinear and global optimization*, volume 37 (Springer).

Iancu DA, Sharma M, Sviridenko M (2013) Supermodularity and affine policies in dynamic robust optimization. *Operations Research* 61(4):941–956.

Julien E, Postek K, Birbil Şİ (2022) Machine learning for K-adaptability in two-stage robust optimization. *arXiv preprint arXiv:2210.11152* .

Kuhn D, Wiesemann W, Georghiou A (2011) Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming* 130(1):177–209.

Loken C, Gruner D, Groer L, Peltier R, Bunn N, Craig M, Henriques T, Dempsey J, Yu CH, Chen J, et al. (2010) SciNet: lessons learned from building a power-efficient top-20 system and data centre. *Journal of Physics-Conference Series*, volume 256, 012026.

Mohajerin Esfahani P, Kuhn D (2018) Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming* 171(1-2):115–166, URL http://dx.doi.org/10.1007/s10107-017-1172-1.

Pereira MV, Pinto LM (1991) Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming* 52(1-3):359–375.

Philpott AB, de Matos VL, Kapelevich L (2018) Distributionally robust SDDP. *Computational Management Science* 15(3):431–454.

Ponce M, van Zon R, Northrup S, Gruner D, Chen J, Ertinaz F, Fedoseev A, Groer L, Mao F, Mundim BC, et al. (2019) Deploying a top-100 supercomputer for large parallel workloads: The Niagara supercomputer. *Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (learning)*, 1–8 (Association for Computing Machinery, New York, NY, United States).

Postek K, Hertog Dd (2016) Multistage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set. *INFORMS Journal on Computing* 28(3):553–574.

Romeijnders W, Postek K (2020) Piecewise constant decision rules via branch-and-bound based scenario detection for integer adjustable robust optimization. *INFORMS Journal on Computing* .

See CT, Sim M (2010) Robust approximation to multiperiod inventory management. *Operations Research* 58(3):583–594.

Shapiro A, Dentcheva D, Ruszczyński A (2009) *Lectures on Stochastic Programming: Modeling and Theory* (SIAM).

Subramanyam A, Gounaris CE, Wiesemann W (2020) K-adaptability in two-stage mixed-integer robust optimization. *Mathematical Programming Computation* 12(2):193–224.

Tsang MY, Shehadeh KS, Curtis FE (2023) An inexact column-and-constraint generation method to solve two-stage robust optimization problems 51(1):92–98.

Wolsey LA, Nemhauser GL (1999) *Integer and combinatorial optimization*, volume 55 (John Wiley & Sons).

Xu G, Hanasusanto GA (2021) Improved decision rule approximations for multi-stage robust optimization via copositive programming. *arXiv preprint arXiv:1808.06231* .

Yanıkoğlu İ, Gorissen BL, den Hertog D (2019) A survey of adjustable robust optimization. *European Journal of Operational Research* 277(3):799–813.

Zeng B, Zhao L (2013) Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operations Research Letters* 41(5):457–461.

Zhao L, Zeng B (2012) An exact algorithm for two-stage robust optimization with mixed integer recourse problems.

Zhen J, Den Hertog D, Sim M (2018) Adjustable robust optimization via Fourier–Motzkin elimination. *Operations Research* 66(4):1086–1100.

Appendices

A. Obtaining 2ARO Model via Two-stage Decision Rules

In this section, we detail the transition from the MSARO model (3) to the 2ARO model (4) as a result of applying two-stage decision rules proposed in Section 3.1. Substituting the state variables in the MSARO problem with the decision rules $x_t^{\mathtt{s}}(\xi^t) = \Theta_t(\xi^t, \beta_t), t \in [2,T]$, we obtain the following model:

$$
\begin{aligned}
\min\ & c_1^\top x_1 + z^{\text{rest}} && && \text{(35a)}\\
\text{s.t. } & x_1 \in X_1 && && \text{(35b)}\\
& \beta_t \in \mathbb{R}^{K_t} && t \in [2,T] && \text{(35c)}\\
& z^{\text{rest}} \ge \sum_{t\in[2,T]} c_t^{\mathtt{s}}(\xi^t)^\top \Theta_t(\xi^t,\beta_t) + \sum_{t\in[2,T]} c_t^{\mathtt{r}}(\xi^t)^\top x_t^{\mathtt{r}}(\xi^t) && \xi^T \in \Xi^T && \text{(35d)}\\
& A_t^{\mathtt{r}}(\xi^t) x_t^{\mathtt{r}}(\xi^t) \le b_t(\xi^t) - \Big(A_t^{\mathtt{s}}(\xi^t)\Theta_t(\xi^t,\beta_t) + B_t^{\mathtt{s}}(\xi^t) x_1^{\mathtt{s}}\Big) && t=2, \xi^t \in \Xi^t && \text{(35e)}\\
& A_t^{\mathtt{r}}(\xi^t) x_t^{\mathtt{r}}(\xi^t) \le b_t(\xi^t) - \Big(A_t^{\mathtt{s}}(\xi^t)\Theta_t(\xi^t,\beta_t) + B_t^{\mathtt{s}}(\xi^t)\Theta_{t-1}(\xi^{t-1},\beta_{t-1})\Big) && t \in [3,T], \xi^t \in \Xi^t && \text{(35f)}\\
& (\Theta_t(\xi^t,\beta_t), x_t^{\mathtt{r}}(\xi^t)) \in X_t(\xi^t) && t \in [2,T], \xi^t \in \Xi^t && \text{(35g)}
\end{aligned}
$$

Given the first-stage decisions x_1 and β, observe that the feasible set of the recourse variables, defined by (35e)-(35g) decomposes by stage t and history ξ^t (as both the decision variables, $x_t^{\mathtt{r}}(\xi^t)$, and the constraints, (35e)-(35g), are separately defined for each stage and history, i.e., there is no link between them). For convenience, let us represent this feasible space in a decomposed form via $x_t^{\mathtt{r}}(\xi^t) \in \mathcal{X}_t^{\mathtt{r}}(x_1^{\mathtt{s}}, \beta, \xi^t), t \in [2,T], \xi^t \in \Xi^t$, and write (35) more compactly as follows:

$$
\begin{aligned}
\min_{x_1 \in X_1, \beta}\ & c_1^\top x_1 + z^{\text{rest}} && && \text{(36a)}\\
\text{s.t. } & z^{\text{rest}} \ge \sum_{t\in[2,T]} c_t^{\mathtt{s}}(\xi^t)^\top \Theta_t(\xi^t,\beta_t) + \sum_{t\in[2,T]} c_t^{\mathtt{r}}(\xi^t)^\top x_t^{\mathtt{r}}(\xi^t) && \xi^T \in \Xi^T && \text{(36b)}\\
& x_t^{\mathtt{r}}(\xi^t) \in \mathcal{X}_t^{\mathtt{r}}(x_1^{\mathtt{s}}, \beta, \xi^t) && t \in [2,T], \xi^t \in \Xi^t && \text{(36c)}
\end{aligned}
$$

In order to minimize z^{rest}, the second summation term in (36b) should be minimized over the recourse decisions $x_t^{\mathtt{r}}(\xi^t)$. Since this term is additively separable over stages $t \in [2,T]$ and their associated history $\xi^t \in \Xi^t$, the recourse decisions of each stage can be optimized separately, yielding the following equivalent model:

$$
\begin{aligned}
\min_{x_1 \in X_1, \beta}\ & c_1^\top x_1 + z^{\text{rest}} && && \text{(37a)}\\
\text{s.t. } & z^{\text{rest}} \ge \sum_{t\in[2,T]} c_t^{\mathtt{s}}(\xi^t)^\top \Theta_t(\xi^t,\beta_t) + \sum_{t\in[2,T]} \min_{x_t^{\mathtt{r}}(\xi^t) \in \mathcal{X}_t^{\mathtt{r}}(x_1^{\mathtt{s}},\beta,\xi^t)} c_t^{\mathtt{r}}(\xi^t)^\top x_t^{\mathtt{r}}(\xi^t) && \xi^T \in \Xi^T && \text{(37b)}
\end{aligned}
$$

Inspecting the minimization problems in (37b), we observe that the objective function coefficients and the feasible set are parametrized by the history, as such changing the parametrization of the recourse decisions from the history ξ^t, to a full uncertainty realization ξ^T should not change the optimal objective value. Thus, instead, we can consider the following equivalent model:

$$\min_{x_1 \in X_1, \beta} \; c_1^\top x_1 + z^{\text{rest}} \tag{38a}$$

$$\text{s.t. } z^{\text{rest}} \geq \sum_{t \in [2,T]} c_t^{\mathsf{s}}(\xi^t)^\top \Theta_t(\xi^t, \beta_t) + \sum_{t \in [2,T]} \min_{x_t^{\mathsf{r}}(\xi^T) \in \mathcal{X}_t^{\mathsf{r}}(x_1^{\mathsf{s}}, \beta, \xi^t)} c_t^{\mathsf{r}}(\xi^t)^\top x_t^{\mathsf{r}}(\xi^T) \qquad \xi^T \in \Xi^T \tag{38b}$$

More formally, we can show that given an optimal solution to the inner minimization problem in (38b), $\hat{x}_t^{\mathsf{r}} : \xi^T \to \mathbb{R}^{p_t}$, we can construct a feasible solution $\tilde{x}_t^{\mathsf{r}} : \xi^t \to \mathbb{R}^{p_t}$ to the inner minimization problem in (37b) which attains the same objective value. Since the minimization problem in (38b) is a relaxation of that in (37b), having more flexible decision variables, the aforementioned construction is sufficient to conclude the proof of our equivalence claim. Given stage $t \in [2,T]$ and history $\breve{\xi}^t \in \Xi^t$, we let $\tilde{x}_t^{\mathsf{r}}(\breve{\xi}^t) := \hat{x}_t^{\mathsf{r}}(\mathring{\xi}^T)$ where $\mathring{\xi}^T$ is an arbitrarily selected element from the set of realizations with the same history $\{\xi^T \in \Xi^T : \xi^t = \breve{\xi}^t\}$. The feasibility of the constructed policy is straightforward since the two inner minimization problems have the same feasibility set, defined by $\mathcal{X}_t^{\mathsf{r}}(x_1^{\mathsf{s}}, \beta, \xi^t), t \in [2,T]$. Next, we observe that $c_t^{\mathsf{r}}(\xi^t)^\top \hat{x}_t^{\mathsf{r}}(\xi^T)$ is the same for all realizations in $\{\xi^T \in \Xi^T : \xi^t = \breve{\xi}^t\}$ since the minimization problem in (38b) has the same feasible set and objective coefficient vector for any of those realizations and $\hat{x}_t^{\mathsf{r}}(\cdot)$ is chosen to be an optimal policy. Lastly, by construction, the policy $\tilde{x}_t^{\mathsf{r}}(\cdot)$ achieves the same objective value.

As there is no link between the different stage optimization problems in (38b), we can swap the summation and minimization operators and optimize over the recourse variables associated with all the stages together given an uncertainty realization:

$$\min_{x_1 \in X_1, \beta} \; c_1^\top x_1 + z^{\text{rest}} \tag{39a}$$

$$\text{s.t. } z^{\text{rest}} \geq \sum_{t \in [2,T]} c_t^{\mathsf{s}}(\xi^t)^\top \Theta_t(\xi^t, \beta_t) + \min_{(x_t^{\mathsf{r}}(\xi^T))_{t \in [2,T]} \in \mathcal{X}(x_1^{\mathsf{s}}, \beta, \xi^T)} \sum_{t \in [2,T]} c_t^{\mathsf{r}}(\xi^t)^\top x_t^{\mathsf{r}}(\xi^T) \quad \xi^T \in \Xi^T \tag{39b}$$

where $\mathcal{X}(x_1^{\mathsf{s}}, \beta, \xi^T) = \bigtimes_{t \in [2,T]} \mathcal{X}_t^{\mathsf{r}}(x_1^{\mathsf{s}}, \beta, \xi^t)$. Since z^{rest} is equal to the maximum of the right-hand side of (38b) over the realizations $\xi^T \in \Xi^T$ at an optimal solution, we can turn this problem into the following nested formulation:

$$\min \;\; c_1^\top x_1 + \max_{\xi^T \in \Xi^T} \; \min_{x^{\mathsf{r}} \in \mathcal{X}(x_1^{\mathsf{s}}, \beta, \xi^T)} \sum_{t \in [2,T]} c_t^{\mathsf{s}}(\xi^t)^\top \Theta_t(\xi^t, \beta_t) + c_t^{\mathsf{r}}(\xi^t)^\top x_t^{\mathsf{r}} \tag{40a}$$

where the parametrization of the recourse variables is omitted since ξ^T is given as an input to the inner minimization problem.

B. Proofs

In this section, we present the proofs of the three propositions and one lemma mentioned in the body of the paper, for which we also restate the claims for convenience.

PROPOSITION 1. *Consider an MSARO with only right-hand-side uncertainty, continuous recourse, and (bounded) polyhedral uncertainty set. If the basis functions* $\Phi_t(\xi^t)$ *are chosen to be affine in* ξ^t *for all* $t\in[2,T]$*, the C&CG algorithm converges to* $\nu^{2S\text{-}LDR}$ *in a finite number of iterations.*

Proof Denote by $f(x_1^{\mathsf{s}},\beta,\xi^T)$, the objective function of the inner minimization in the two-stage problem (6). Then, we can rewrite problem (6) as:

$$\nu^{\text{2S-LDR}}=\min\left\{c_1^\top x_1+\max_{\xi^T\in\Xi^T} f(x_1^{\mathsf{s}},\beta,\xi^T)\mid x_1\in X_1,\ \beta_t\in\mathbb{R}^{K_t},\ t\in[2,T]\right\}.$$

First, we show that $f(x_1^{\mathsf{s}},\beta,\xi^T)$ is convex in ξ^T for given $x_1=(x_1^{\mathsf{s}},x_1^{\mathsf{r}})\in X_1$ and β. For $\hat{\xi}^T,\tilde{\xi}^T\in\Xi^T$ and $\lambda\in[0,1]$, the following (in)equalities hold:

$$\begin{aligned}
&\lambda f(x_1^{\mathsf{s}},\beta,\hat{\xi}^T)+(1-\lambda)f(x_1^{\mathsf{s}},\beta,\tilde{\xi}^T) &&= \\
&\sum_{t\in[2,T]} {c_t^{\mathsf{s}}}^\top\big(\lambda\Phi_t(\hat{\xi}^T)+(1-\lambda)\Phi_t(\tilde{\xi}^t)\big)\beta_t+ \\
&\qquad\lambda\Big(\min_{x^{\mathsf{r}}\in\mathcal{X}(x_1^{\mathsf{s}},\beta,\hat{\xi}^T)}\sum_{t\in[2,T]}{c_t^{\mathsf{r}}}^\top x_t^{\mathsf{r}}\Big)+(1-\lambda)\Big(\min_{x^{\mathsf{r}}\in\mathcal{X}(x_1^{\mathsf{s}},\beta,\tilde{\xi}^T)}\sum_{t\in[2,T]}{c_t^{\mathsf{r}}}^\top x_t^{\mathsf{r}}\Big) &&= \qquad (41a)\\
&\sum_{t\in[2,T]} {c_t^{\mathsf{s}}}^\top\Phi_t(\lambda\hat{\xi}^t+(1-\lambda)\tilde{\xi}^t)\beta_t+ \\
&\qquad\lambda\Big(\min_{x^{\mathsf{r}}\in\mathcal{X}(x_1^{\mathsf{s}},\beta,\hat{\xi}^T)}\sum_{t\in[2,T]}{c_t^{\mathsf{r}}}^\top x_t^{\mathsf{r}}\Big)+(1-\lambda)\Big(\min_{x^{\mathsf{r}}\in\mathcal{X}(x_1^{\mathsf{s}},\beta,\tilde{\xi}^T)}\sum_{t\in[2,T]}{c_t^{\mathsf{r}}}^\top x_t^{\mathsf{r}}\Big) &&\geq \qquad (41b)\\
&\sum_{t\in[2,T]} {c_t^{\mathsf{s}}}^\top\Phi_t(\lambda\hat{\xi}^t+(1-\lambda)\tilde{\xi}^t)\beta_t+\min_{x^{\mathsf{r}}\in\mathcal{X}(x_1^{\mathsf{s}},\beta,\lambda\hat{\xi}^T+(1-\lambda)\tilde{\xi}^T)}\sum_{t\in[2,T]}{c_t^{\mathsf{r}}}^\top x_t^{\mathsf{r}} &&= f(x_1^{\mathsf{s}},\beta,\lambda\hat{\xi}^T+(1-\lambda)\tilde{\xi}^T).
\end{aligned}$$

Equality (41a) holds because $\Phi_t(\xi^t)$ is affine in ξ^t. Inequality (41b) follows from the convexity of the optimal value of the inner minimization problem as a function of ξ^T, since uncertainty appears only on the right-hand-sides of constraints and recourse variables are continuous by assumption.

Now, let Ξ^{Ext} be the set of extreme points of Ξ^T. In maximization of a convex function over a compact polyhedral set, there is an optimal solution that is an extreme point (Hendrix et al. 2010). Then, the two-stage problem becomes:

$$\nu^{\text{2S-LDR}}=\min\left\{c_1^\top x_1+\eta\mid \eta\geq f(x_1^{\mathsf{s}},\xi^T,\beta),\ \xi^T\in\Xi^{\text{Ext}},\ x_1\in X_1,\ \beta_t\in\mathbb{R}^{K_t},\ t\in[2,T]\right\}.$$

C&CG is then the process of gradually adding constraints for each extreme point. Because $|\Xi^{\text{Ext}}|<+\infty$, iterating over all extreme points takes finitely many steps, which concludes the proof. □

Lemma 1. *For any* $\mathbb{P} \in \mathcal{P}^{>}$*, constraints* (12e) *are equivalent to the following:*

$$y_t(\xi^T) = \mathbb{E}_{\xi'^T \sim \mathbb{P}}\left[y_t(\xi'^T) \;\middle|\; \xi'^t = \xi^t \right], \quad t \in [T],\ \xi^T \in \Xi^T. \tag{13}$$

Proof For a fixed t and ξ^T, we start by multiplying both sides of (12e) by the density $p^{\mathbb{P}}(\xi'^T)$ for every ξ'^T (including ξ^T) that shares the same history with ξ^T up to t:

$$p^{\mathbb{P}}(\xi'^T) y_t(\xi^T) = p^{\mathbb{P}}(\xi'^T) y_t(\xi'^T) \qquad t \in [T],\ \xi^T, \xi'^T \in \Xi^T \text{ with } \xi^t = \xi'^t.$$

Since $p^{\mathbb{P}}(\xi'^T) > 0$ for every realization ξ'^T under any probability distribution $\mathbb{P} \in \mathcal{P}^{>}$, the feasible set of model (12) remains the same. Then, we integrate (or sum if Ξ^T is discrete) the scaled constraints over all such ξ'^T realizations to obtain

$$y_t(\xi^T) \int\limits_{\xi'^T : \xi'^t = \xi^t} p^{\mathbb{P}}(\xi'^T) d\xi'^T = \int\limits_{\xi'^T : \xi'^t = \xi^t} p^{\mathbb{P}}(\xi'^T) y_t(\xi'^T) d\xi'^T \qquad t \in [T],\ \xi^T \in \Xi^T.$$

Let $\delta := \int\limits_{\xi'^T : \xi'^t = \xi^t} p^{\mathbb{P}}(\xi'^T) d\xi'^T$ where $\delta > 0$ since $\xi^T \in \{\xi'^T : \xi'^t = \xi^t\}$ with $p^{\mathbb{P}}(\xi^T) > 0$. Then, we obtain (13) via

$$y_t(\xi^T) = \int\limits_{\xi'^T : \xi'^t = \xi^t} \frac{p^{\mathbb{P}}(\xi'^T)}{\delta} y_t(\xi'^T) d\xi'^T = \mathbb{E}_{\xi'^T \sim \mathbb{P}}\left[y_t(\xi'^T) \;\middle|\; \xi'^t = \xi^t \right] \qquad t \in [T],\ \xi^T \in \Xi^T.$$

Since constraints (13) are obtained as an aggregation of the original constraints (12e), they are valid for model (12). Further, they imply the nonanticipativity constraints (12e) since for any given t along with ξ^T and $\hat{\xi}^T$ such that $\xi^t = \hat{\xi}^t$, the right-hand side of (13) is the same, i.e., $\mathbb{E}_{\xi'^T \sim \mathbb{P}}\left[y_t(\xi'^T) \;\middle|\; \xi'^t = \xi^t \right] = \mathbb{E}_{\xi'^T \sim \mathbb{P}}\left[y_t(\xi'^T) \;\middle|\; \xi'^t = \hat{\xi}^t \right]$, enforcing $y_t(\xi^T) = y_t(\hat{\xi}^T)$. □

Proposition 2. *Let* $\mathbb{P}$ *be any probability measure in* $\mathcal{P}^{>}$*. For MSARO problems with* continuous recourse*,* (15) *is a strong dual of* (2)*, i.e.,* $\mathcal{L}^{\text{NA}}(\mathbb{P}) = \nu^\star$*.*

Proof The strength of a Lagrangian dual problem can be studied by its primal characterization, derived by Geoffrion (1974) for a mixed-integer linear optimization problem. Then, using standard Lagrangian duality theory (see, for instance, Wolsey and Nemhauser (1999)), the primal characterization of (15) is:

$$\min \ z \tag{42a}$$

$$\text{s.t.} \ \left(z, y_1(\xi^T), \ldots, y_T(\xi^T)\right) \in \operatorname{conv}(Y(\xi^T)) \qquad \xi^T \in \Xi^T \tag{42b}$$

$$y_t(\xi^T) = \mathbb{E}_{\xi'^T \sim \mathbb{P}}\left[y_t(\xi'^T) \;\middle|\; \xi'^t = \xi^t \right] \qquad t \in [T], \xi^T \in \Xi^T. \tag{42c}$$

The result follows from the equivalence of (13) and (12e), as well as the fact that for MSARO problems with continuous recourse, we have that $\operatorname{conv}(Y(\xi^T)) = Y(\xi^T)$ for all $\xi^T \in \Xi^T$. □

PROPOSITION 3. *$\bar{\nu}_R^{\text{NA-DO}}$ is a lower bound for $\nu^\star$.*

Proof We have shown that $\mathcal{L}_R^{\text{NA}}(\mathbb{P}) \le \nu^\star$ for any $\mathbb{P} \in \mathcal{P}^>$. Now we claim that, for any $\hat{\mathbb{P}} \in \mathcal{P}^\ge \setminus \mathcal{P}^> = \{\mathsf{P} \in \mathcal{P}^\ge \mid \exists \xi^T \in \Xi^T: \ p^{\mathsf{P}}(\xi^T) = 0\}$ where $p^{\hat{\mathbb{P}}}: \Xi^T \to \mathbb{R}_+$ is the density function of $\hat{\mathbb{P}}$, the inequality $\mathcal{L}_R^{\text{NA}}(\hat{\mathbb{P}}) \le \nu^\star$ also holds. Given $\hat{\mathbb{P}}$, let $\hat{\Xi}^T := \{\xi^T \in \Xi^T : p^{\hat{\mathbb{P}}}(\xi^T) > 0\}$. Using $\hat{\Xi}^T$, we create a relaxation of (12), the NA reformulation of the MSARO problem, as

$$
\begin{aligned}
\hat{\nu}(\hat{\mathbb{P}}) := \min \ & z && && \text{(43a)}\\
\text{s.t.} \ & \sum_{t\in[T]} c_t(\xi^t)^\top y_t(\xi^T) \le z && \xi^T \in \Xi^T && \text{(43b)}\\
& A_t(\xi^t) y_t(\xi^T) + B_t(\xi^t) y_{t-1}(\xi^T) \le b_t(\xi^t) && t \in [2,T], \ \xi^T \in \Xi^T && \text{(43c)}\\
& D_t(\xi^t) y_t(\xi^T) \le d_t(\xi^t) && t \in [T], \ \xi^T \in \Xi^T && \text{(43d)}\\
& y_t(\xi^T) = y_t(\xi'^T) && t \in [T], \ \xi^T \in \hat{\Xi}^T, \xi'^T \in \hat{\Xi}^T \text{ with } \xi^t = \xi'^t && \text{(43e)}\\
& y_t(\xi^T) \in \mathbb{R}^{n_t - n_t^{\mathrm{i}}} \times \mathbb{Z}^{n_t^{\mathrm{i}}} && t \in [T], \ \xi^T \in \Xi^T && \text{(43f)}
\end{aligned}
$$

where the NA constraints are only imposed for the pairs of realizations in $\hat{\Xi}^T \subset \Xi^T$. Therefore, we have $\hat{\nu}(\hat{\mathbb{P}}) \le \nu^*$.

Now that $\hat{\mathbb{P}}$ assigns a positive density to all the realizations from $\hat{\Xi}^T$, via Lemma 1, the NA constraints in (43e) can be equivalently reformulated as

$$
y_t(\xi^T) = \mathbb{E}_{\xi'^T \sim \hat{\mathbb{P}}}\left[y_t(\xi'^T) \;\middle|\; \xi'^t = \xi^t \right], \quad t \in [T], \ \xi^T \in \hat{\Xi}^T.
$$

Subsequently, relaxation of these reformulated NA constraints and construction of the Lagrangian dual problem (15) with respect to $\hat{\mathbb{P}}$ yields a relaxation of a relaxation of a minimization problem, where for each fixed $\lambda_t(\cdot)$ we have $\mathcal{L}_{\text{LR}}^{\text{NA}}(\hat{\mathbb{P}}, \lambda_1(\cdot). \dots, \lambda_T(\cdot)) \le \hat{\nu}(\hat{\mathbb{P}})$, and consequently $\mathcal{L}^{\text{NA}}(\hat{\mathbb{P}}) \le \hat{\nu}(\hat{\mathbb{P}})$. Furthermore, restricting the Lagrangian duals to follow LDRs, we obtain $\mathcal{L}_R^{\text{NA}}(\hat{\mathbb{P}}) \le \mathcal{L}^{\text{NA}}(\hat{\mathbb{P}})$.

Since we showed $\mathcal{L}_R^{\text{NA}}(\mathbb{P}) \le \nu^*$ for all $\mathbb{P} \in \mathcal{P}^\ge$, we have $\bar{\nu}_R^{\text{NA-DO}} := \max_{\mathbb{P}\in\mathcal{P}^\ge} \ \mathcal{L}_R^{\text{NA}}(\mathbb{P}) \le \nu^*$, which completes the proof. □

C. Detailed Models

C.1. Monolithic Form of Model (8)

Consider the case mentioned in Remark 1, namely an MSARO where the uncertainty set is a polytope, the basis functions $\Phi_t(\xi^t)$ are chosen to be affine in ξ^t for all $t \in [2,T]$, all the recourse variables are continuous, we have fixed parameters associated with the state variables, and the 2ARO problem (6) has relatively complete recourse. We next detail how a monolithic mixed-integer linear programming formulation of the inner minimization problem (8) can be derived as in (Ayoub and Poss 2016, Zeng and Zhao 2013).

Let $\pi^{\mathtt{S}}$ and $\pi^{\mathtt{R}}$ be the linear programming dual variables associated with the state and recourse constraints in $\mathcal{X}(\hat{x}_1^{\mathtt{s}}, \hat{\beta}, \xi^T)$, in the inner minimization problem of (8), respectively. Then, using KKT conditions, subproblem (8) can be modelled as follows:

$$\max \sum_{t\in[2,T]} c_t^{\mathtt{s}\top}\Phi_t(\xi^t)\hat{\beta}_t + c_t^{\mathtt{r}\top}x_t^{\mathtt{r}} \tag{44a}$$

$$\text{s.t. } A_t^{\mathtt{r}}x_t^{\mathtt{r}} + A_t^{\mathtt{s}}\Phi_t(\xi^t)\hat{\beta}_t + B_t^{\mathtt{s}}\hat{x}_1^{\mathtt{s}} - b_t(\xi^t) \le 0 \qquad t = 2 \tag{44b}$$

$$A_t^{\mathtt{r}}x_t^{\mathtt{r}} + A_t^{\mathtt{s}}\Phi_t(\xi^t)\hat{\beta}_t + B_t^{\mathtt{s}}\Phi_{t-1}(\xi^{t-1})\hat{\beta}_{t-1} - b_t(\xi^t) \le 0 \qquad t\in[3,T] \tag{44c}$$

$$D_t^{\mathtt{s}}\Phi_t(\xi^t)\hat{\beta}_t + D_t^{\mathtt{r}}x_t^{\mathtt{r}} - d_t(\xi^t) \le 0 \qquad t\in[2,T] \tag{44d}$$

$$A_t^{\mathtt{r}\top}\pi_t^{\mathtt{S}} + D_t^{\mathtt{r}\top}\pi_t^{\mathtt{R}} = c_t^{\mathtt{r}} \qquad t\in[2,T] \tag{44e}$$

$$\Big(A_t^{\mathtt{r}}x_t^{\mathtt{r}} + A_t^{\mathtt{s}}\Phi_t(\xi^t)\hat{\beta}_t + B_t^{\mathtt{s}}\hat{x}_1^{\mathtt{s}} - b_t(\xi^t)\Big)^\top\pi_t^{\mathtt{S}} = 0 \qquad t = 2 \tag{44f}$$

$$\Big(A_t^{\mathtt{r}}x_t^{\mathtt{r}} + A_t^{\mathtt{s}}\Phi_t(\xi^t)\hat{\beta}_t + B_t^{\mathtt{s}}\Phi_{t-1}(\xi^{t-1})\hat{\beta}_{t-1} - b_t(\xi^t)\Big)^\top\pi_t^{\mathtt{S}} = 0 \qquad t\in[3,T] \tag{44g}$$

$$\Big(D_t^{\mathtt{s}}\Phi_t(\xi^t)\hat{\beta}_t + D_t^{\mathtt{r}}x_t^{\mathtt{r}} - d_t(\xi^t)\Big)^\top\pi_t^{\mathtt{R}} = 0 \qquad t\in[2,T] \tag{44h}$$

$$x_t^{\mathtt{r}} \in \mathbb{R}^{p_t}, \quad \pi_t^{\mathtt{S}} \in \mathbb{R}_-^{m_t^{\mathtt{S}}}, \quad \pi_t^{\mathtt{R}} \in \mathbb{R}_-^{m_t^{\mathtt{R}}} \qquad t\in[2,T] \tag{44i}$$

$$\xi^T \in \Xi^T. \tag{44j}$$

Inequalities (44b)-(44d) are the primal feasibility constraints at $\hat{\beta}_t$, (44e) the dual feasibility constraints, and (44f)-(44h) are the complementary slackness constraints. The latter include bilinear terms, as basis functions $\Phi_t(\xi^t)$ are functions of ξ^T which are decision variables in (44). They can be linearized with the addition of binary decision variables via the so-called big-M constraints. Since the basis functions are chosen to be affine, the resulting model is a mixed-integer linear program. A detailed example of building a mixed-integer linear model for a multistage location-allocation problem is provided in Appendix C.3.

C.2. Two-stage Decision Rules for MSAROs with Mixed-integer State Variables

As mentioned in Remark 3, for MSAROs with mixed-integer state variables, linear and piecewise constant decision rules can be combined to obtain a 2ARO approximation. To this end, consider the partition of the index set of the state variables into sets $\mathcal{I}_t^{\mathtt{i}}$ and $\mathcal{I}_t^{\mathtt{c}}$ for integer and continuous variables, respectively, i.e., $\mathcal{I}_t^{\mathtt{i}} \cup \mathcal{I}_t^{\mathtt{c}} = [q_t], \mathcal{I}_t^{\mathtt{i}} \cap \mathcal{I}_t^{\mathtt{c}} = \emptyset, |\mathcal{I}_t^{\mathtt{i}}| = q_t^{\mathtt{i}}$. Similarly, let $x_t^{\mathtt{s}} = (x_t^{\mathtt{s},\mathtt{i}}, x_t^{\mathtt{s},\mathtt{c}})$ and $\beta = (\beta^{\mathtt{i}}, \beta^{\mathtt{c}})$ be the vectors of state variables and decision rule design variables with sub-vectors corresponding to the integer and continuous state variables. Then, the application of LDRs (5) to the continuous state variables and PCDRs (9) to the integer state variables yields the following 2ARO model:

$$\nu^{\text{2S-LDR-PCDR}} := \min\ c_1^\top x_1 + \mathcal{SP}^{\text{2S-LDR-PCDR}}(\beta, x_1^{\mathtt{s}}) \tag{45a}$$

$$
\begin{aligned}
\text{s.t. } & x_1 \in X_1 && && (45b)\\
& \beta_t^{\mathtt{c}} \in \mathbb{R}^{K_t^{\mathtt{c}}} && t \in [2,T] && (45c)\\
& \beta_t^{\mathtt{i}} \in \mathbb{R}^{K_t^{\mathtt{i}}} && t \in [2,T] && (45d)\\
& -1 \le \hat{\Upsilon}_{ti}(\xi^t)^\top \beta_{ti}^{\mathtt{i}} \le 1 && t \in [2,T], i \in \mathcal{I}_t^{\mathtt{i}}, \xi^T \in \Xi^T && (45e)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{SP}^{\text{2S-LDR-PCDR}}(\beta, x_1^{\mathtt{s}}) := \max_{\xi^T \in \Xi^T} \quad & \sum_{t \in [2,T]} c_t^{\mathtt{s}}(\xi^t)^\top x_t^{\mathtt{s}} + \min_{x^{\mathtt{r}} \in \mathcal{X}(x^{\mathtt{s}}, \xi^T)} \sum_{t \in [2,T]} c_t^{\mathtt{r}}(\xi^t)^\top x^{\mathtt{r}} && && (46a)\\
\text{s.t.} \quad & \Phi_t(\xi^t)\beta_t^{\mathtt{c}} = x_t^{\mathtt{s},\mathtt{c}} && t \in [2,T] && (46b)\\
& \sum_{j \in [J_i]} (\underline{\kappa}_{ti} + j - 1)\upsilon_{tij} = x_{ti}^{\mathtt{s},\mathtt{i}} && t \in [2,T], i \in \mathcal{I}_t^{\mathtt{i}} && (46c)\\
& \sum_{j \in [J_i]} \omega_{tij} = \hat{\Upsilon}_{ti}(\xi^t)^\top \beta_t^i && t \in [2,T], i \in \mathcal{I}_t^{\mathtt{i}} && (46d)\\
& (a_{tj}^i + \epsilon_j)\upsilon_{tij} \le \omega_{tij} \le b_{tj}^i \upsilon_{tij} && t \in [2,T], i \in \mathcal{I}_t^{\mathtt{i}}, j \in [J_i] && (46e)\\
& \sum_{j \in [J_i]} \upsilon_{tij} = 1 && t \in [2,T], i \in \mathcal{I}_t^{\mathtt{i}} && (46f)\\
& \upsilon_{tij} \in \{0,1\} && t \in [2,T], i \in \mathcal{I}_t^{\mathtt{i}}, j \in [J_i]. && (46g)
\end{aligned}
$$

Given the state variables and an uncertainty realization, the recourse feasible set is defined as

$$
\begin{aligned}
\mathcal{X}(x^{\mathtt{s}}, \xi^T) := \Big\{ & \big(x_t^{\mathtt{r}}\big)_{t \in [2,T]} \in \mathbb{R}^{p_2} \times \mathbb{R}^{p_3} \times \cdots \times \mathbb{R}^{p_T} : \\
& A_t^{\mathtt{r}}(\xi^t)x_t^{\mathtt{r}} \le b_t(\xi^t) - \Big(A_t^{\mathtt{s}}(\xi^t)x_t^{\mathtt{s}} + B_t^{\mathtt{s}}(\xi^t)x_1^{\mathtt{s}}\Big) && t = 2\\
& A_t^{\mathtt{r}}(\xi^t)x_t^{\mathtt{r}} \le b_t(\xi^t) - \Big(A_t^{\mathtt{s}}(\xi^t)x_t^{\mathtt{s}} + B_t^{\mathtt{s}}(\xi^t)x_{t-1}^{\mathtt{s}}\Big) && t \in [3,T]\\
& (x_t^{\mathtt{s}}, x_t^{\mathtt{r}}) \in X_t(\xi^t) && t \in [2,T] \Big\}.
\end{aligned}
$$

Lastly, we note that in the special case where the assumptions of Remarks 1 and 2 are satisfied, the subproblem (46) can be similarly reformulated as a monolithic mixed-integer linear program.

C.3. Models for the Location-Transportation

C.3.1. Column-and-constraint Generation with Two-stage Linear Decision Rules Applying LDRs on the state variables of model (33), $s_{it}(\xi^t) = \beta_{it}^0 + \sum_{t' \in [2,t]} \sum_{j \in [J]} d_{jt'}(\xi^t)\beta_{it}^{jt'}, i \in [I], t \in [2,T]$, results in the following 2ARO problem in monolithic form:

$$
\begin{aligned}
\min \quad & z\\
\text{s.t.} \quad & z \ge \sum_{i \in [I]} f_i y_i + \sum_{i \in [I]} a_i s_{i1} + \sum_{t \in [2,T]} \sum_{i \in [I]} a_i \big(\beta_{it}^0 + \sum_{t' \in [2,t]} \sum_{j \in [J]} d_{jt'}(\xi'^t)\beta_{it}^{jt'}\big) +
\end{aligned}
$$

$$\sum_{t\in[2,T]}\sum_{i\in[I]}\sum_{j\in[J]} c_{ij}x_{ijt}(\xi^t) \qquad \xi^T\in\Xi^T$$

$$s_{i1}\le K_i y_i \qquad i\in[I]$$

$$\beta^0_{i2}+\sum_{j\in[J]} d_{j2}(\xi^2)\beta^{j2}_{i2} = s_{i1}-\sum_{j\in[J]} x_{ij2}(\xi^2) \qquad i\in[I],\xi^2\in\Xi^2$$

$$\beta^0_{it}+\sum_{t'\in[2,t]}\sum_{j\in[J]} d_{jt'}(\xi^t)\beta^{jt'}_{it} = \beta^0_{i,t-1}+\sum_{t'\in[2,t-1]}\sum_{j\in[J]} d_{jt'}(\xi^{t-1})\beta^{jt'}_{i,t-1}-\sum_{j\in[J]} x_{ijt}(\xi^t) \qquad i\in[I],t\in[3,T],\xi^t\in\Xi^t$$

$$\sum_{i\in[I]} x_{ijt}(\xi^t)\ge d_{jt}(\xi^t) \qquad j\in[J],t\in[2,T],\xi^t\in\Xi^t$$

$$\beta^0_{it}+\sum_{t'\in[2,t]}\sum_{j\in[J]} d_{jt'}(\xi^t)\beta^{jt'}_{it}\ge 0 \qquad i\in[I],t\in[2,T],\xi^t\in\Xi^t$$

$$y\in\{0,1\}^I,\ z,x\ge 0.$$

For better clarity, let us re-write the above formulation in the more common min-max-min form of 2ARO:

$$\begin{aligned}
\min\ & \sum_{i\in[I]} f_i y_i+\sum_{i\in[I]} a_i\big(s_{i1}+\sum_{t\in[2,T]}\beta^0_{it}\big) + \max_{\xi^T\in\Xi^T}\ \min \sum_{t'\in[2,t]} a_i d_{jt'}(\xi^t)\beta^{jt'}_{it}+\sum_{t\in[2,T]}\sum_{i\in[I]}\sum_{j\in[J]} c_{ij}x_{ijt}\\
\text{s.t. } & s_{i1}\le K_i y_i \quad i\in[I] \qquad\qquad \text{s.t. } \sum_{j\in[J]} x_{ij2}=s_{i1}-\beta^0_{i2}-\sum_{j\in[J]} d_{j2}(\xi^2)\beta^{j2}_{i2} \quad i\in[I]\\
& \beta^0_{it}+\sum_{t'\in[2,t]}\sum_{j\in[J]} d_{jt'}(\xi^t)\beta^{jt'}_{it}\ge 0 \quad i\in[I],t\in[2,T],\xi^t\in\Xi^t \qquad \sum_{j\in[J]} x_{ijt}=\beta^0_{i,t-1}-\beta^0_{it}-\sum_{j\in[J]}\sum_{t'\in[2,t]} d_{jt'}(\xi^t)\beta^{jt'}_{it}\\
& y\in\{0,1\}^I. \qquad\qquad +\sum_{j\in[J]}\sum_{t'\in[2,t-1]} d_{jt'}(\xi^{t-1})\beta^{jt'}_{i,t-1} \quad i\in[I],t\in[3,T]\\
& \qquad\qquad \sum_{i\in[I]} x_{ijt}\ge d_{jt}(\xi^t) \quad j\in[J],t\in[2,T]\\
& \qquad\qquad x\ge 0
\end{aligned}$$

Note that we have strengthened the outer minimization problem by adding the non-negativity constraints from the inner minimization problem as robust constraints. This constraint can be rewritten as follows:

$$\beta_{it}^0 + \sum_{t' \in [2,t]} \sum_{j \in [J]} d_{jt'}(\xi^t) \beta_{it}^{jt'} \geq 0 \qquad i \in [I], t \in [2,T], \xi^t \in \Xi^t \rightarrow$$

$$\beta_{it}^0 + \sum_{t' \in [2,t]} \sum_{j \in [J]} \mu_{jt'} \beta_{it}^{jt'} + \min \Big\{ \sum_{t' \in [2,t]} \sum_{j \in [J]} \xi_{jt'} \sigma_{jt'} \beta_{it}^{jt'} : \sum_{t' \in [2,T]} \sum_{j \in [J]} \xi_{jt'} \leq \Gamma, \; 0 \leq \xi_{jt'} \leq 1, j \in [J], t' \in [2,T] \Big\} \geq 0 \qquad i \in [I], t \in [2,T] \rightarrow$$

$$\beta_{it}^0 + \sum_{t' \in [2,t]} \sum_{j \in [J]} \mu_{jt'} \beta_{it}^{jt'} + \max \Big\{ u\Gamma + \sum_{j \in [J]} \sum_{t' \in [2,T]} \omega_{jt'} : u + \omega_{jt'} \leq \sigma_{jt'} \beta_{it}^{jt'}, j \in [J], t' \in [2,t], \; u \leq 0, \omega_{jt'} \leq 0, j \in [J], t' \in [2,T] \Big\} \geq 0 \qquad i \in [I], t \in [2,T] \rightarrow$$

$$\begin{cases} \beta_{it}^0 + \sum_{t' \in [2,t]} \sum_{j \in [J]} \mu_{jt'} \beta_{it}^{jt'} + u\Gamma + \sum_{j \in [J]} \sum_{t' \in [2,T]} \omega_{jt'} \geq 0 & i \in [I], t \in [2,T] \\ u + \omega_{jt'} \leq \sigma_{jt'} \beta_{it}^{jt'} & i \in [I], j \in [J], t \in [2,T], t' \in [2,t] \\ \omega_{jt'} \leq 0 & i \in [I], j \in [J], t \in [2,T], t' \in [2,T] \\ u \leq 0 \end{cases}$$

In order to establish relatively complete recourse, we add the following constraints to the outer minimization problem:

$$\sum_{i \in [I]} \Big[s_{i1} - \beta_{i2}^0 - \sum_{j \in [J]} d_{j2}(\xi^2) \beta_{i2}^{j2} \Big] \geq \sum_{j \in [J]} d_{j2}(\xi^2) \qquad \xi^2 \in \Xi^2$$

$$\sum_{i \in [I]} \Big[\beta_{i,t-1}^0 - \beta_{it}^0 - \sum_{j \in [J]} \sum_{t' \in [2,t]} d_{jt'}(\xi^t) \beta_{it}^{jt'} + \sum_{j \in [J]} \sum_{t' \in [2,t-1]} d_{jt'}(\xi^{t-1}) \beta_{i,t-1}^{jt'} \Big] \geq \sum_{j \in [J]} d_{jt}(\xi^t) \quad t \in [3,T], \xi^t \in \Xi^t,$$

which can be reformulated as follows:

$$\sum_{i \in [I]} \Big[s_{i1} - \beta_{i2}^0 \Big] \geq \max_{\xi^2 \in \Xi^2} \Big\{ \sum_{j \in [J]} d_{j2}(\xi^2) \Big[1 + \sum_{i \in [I]} \beta_{i2}^{j2} \Big] \Big\}$$

$$\sum_{i \in [I]} \Big[\beta_{i,t-1}^0 - \beta_{it}^0 \Big] \geq \max_{\xi^t \in \Xi^t} \Big\{ \sum_{i \in [I]} \Big[\sum_{j \in [J]} \sum_{t' \in [2,t]} d_{jt'}(\xi^t) \beta_{it}^{jt'} - \sum_{j \in [J]} \sum_{t' \in [2,t-1]} d_{jt'}(\xi^{t-1}) \beta_{i,t-1}^{jt'} \Big] + \sum_{j \in [J]} d_{jt}(\xi^t) \Big\} \quad t \in [3,T]$$

The linearization process is similar to the one previously discussed for non-negativity constraints. Denote by π^1, π^2, π^3 the dual variables associated with three constraint sets of the inner minimization problem. After taking its linear programming dual and the adding the KKT conditions, we get the following as the subproblem of the column-and-constraint generation:

$$\max \quad \sum_{t \in [2,T]} \sum_{i \in [I]} \sum_{j \in [J]} \sum_{t' \in [2,t]} a_i d_{jt'} \beta_{it}^{jt'} + \sum_{t \in [2,T]} \sum_{i \in [I]} \sum_{j \in [J]} c_{ij} x_{ijt}$$

$$
\begin{aligned}
\text{s.t.} \quad & \sum_{j\in[J]} x_{ij2} = s_{i1} - \beta^0_{i2} - \sum_{j\in[J]} d_{j2}\beta^{j2}_{i2} && i\in[I]\\
& \sum_{j\in[J]} x_{ijt} = \beta^0_{i,t-1} - \beta^0_{it} - \sum_{j\in[J]}\sum_{t'\in[2,t]} d_{jt'}\beta^{jt'}_{it} + \sum_{j\in[J]}\sum_{t'\in[2,t-1]} d_{jt'}\beta^{jt'}_{i,t-1} && i\in[I], t\in[3,T]\\
& \sum_{i\in[I]} x_{ijt} \ge d_{jt} && j\in[J], t\in[2,T]\\
& \pi^1_i + \pi^3_{j2} \le c_{ij} && i\in[I], j\in[J]\\
& \pi^2_{it} + \pi^3_{jt} \le c_{ij} && i\in[I], j\in[J], t\in[3,T]\\
& \Big(d_{jt} - \sum_{i\in[I]} x_{ijt}\Big)\pi^3_{jt} = 0 \xrightarrow{\text{linearization}} \\
& \qquad \pi^3_{jt} \le M(1-\ell^R_{jt}), \quad \sum_{i\in[I]} x_{ijt} - d_{jt} \le M\ell^R_{jt} && j\in[J], t\in[2,T]\\
& \big(\pi^1_i + \pi^3_{j2} - c_{ij}\big)x_{ij2} = 0 \xrightarrow{\text{linearization}} \\
& \qquad x_{ij2} \le M(1-\ell^D_{ij2}), \quad c_{ij} - \pi^1_i - \pi^3_{j2} \le M\ell^D_{ij2} && i\in[I], j\in[J]\\
& \big(\pi^2_{it} + \pi^3_{jt} - c_{ij}\big)x_{ijt} = 0 \xrightarrow{\text{linearization}} \\
& \qquad x_{ijt} \le M(1-\ell^D_{ijt}), \quad c_{ij} - \pi^2_{it} - \pi^3_{jt} \le M\ell^D_{ijt} && i\in[I], j\in[J], t\in[3,T]\\
& d_{jt} = \mu_{jt} + \xi_{jt}\sigma_{jt} && j\in[J], t\in[2,T]\\
& \sum_{t\in[2,T]}\sum_{j\in[J]} \xi_{jt} \le \Gamma \\
& x_{ijt} \ge 0,\ \ell^D_{ijt}\in\{0,1\} && i\in[I], j\in[J], t\in[2,T]\\
& 0 \le \xi_{jt} \le 1,\ 0 \le \pi^3_{jt},\ \ell^R_{jt}\in\{0,1\} && j\in[J], t\in[2,T].
\end{aligned}
$$

C.3.2. Restricted NA Dual The $\mathcal{Q}(\beta^s, \beta^x)$ function for the location-transportation problem:

$$
\begin{aligned}
& \mathcal{Q}(\beta^s,\beta^x) = \\
& \quad \min \quad z + \sum_{t\in[2,T]}\sum_{i\in[I]} \beta^s_{it}(\xi^T)s_{it}(\xi^T) + \sum_{t\in[2,T]}\sum_{i\in[I]}\sum_{j\in[J]} \beta^x_{ijt}(\xi^T)x_{ijt}(\xi^T)\\
& \quad \text{s.t.} \quad z \ge \sum_{i\in[I]} y_i + \sum_{t\in[T]}\sum_{i\in[I]} a_i s_{it}(\xi^T) + \sum_{t\in[2,T]}\sum_{i\in[I]}\sum_{j\in[J]} c_{ij}x_{ijt}(\xi^T) && \xi^T\in\Xi^T\\
& \qquad\quad s_{i1} \le K_i y_i, && i\in[I]\\
& \qquad\quad s_{it}(\xi^T) = s_{i,t-1}(\xi^{T-1}) - \sum_{j\in[J]} x_{ijt}(\xi^T) && i\in[I], t\in[2,T], \xi^T\in\Xi^T\\
& \qquad\quad \sum_{i\in[I]} x_{ijt}(\xi^T) \ge d_{jt}(\xi^t) && i\in[I], t\in[2,T], \xi^T\in\Xi^T\\
& \qquad\quad y\in\{0,1\}^I, z\in\mathbb{R}^{I\times T}_+, x\in\mathbb{R}^{I\times J\times(T-1)}_+.
\end{aligned}
$$

For this mixed-integer subproblem, the cutting-plane method presented in Section 4.3 is applicable.

D. Detailed Results

D.1. Newsvendor Problem

Table 4 Detailed results for the newsvendor problem

T	BR	$\lvert\Xi^T\rvert$	I	B	ν^{LDR}	$\nu^{\text{2S-LDR}}$	$\nu^{\star}$	$\bar{\nu}_R^{\text{NA-DO}}$	ν^{PI}
3	5	25	3	100	656.6	970.2	975.3	975.3	1457.0
			3	150	6668.5	8410.4	8647.2	9802.6	10036.0
			3	200	8662.0	11446.2	11569.7	13709.0	13709.0
			4	150	5154.8	5384.4	5466.8	5466.8	5781.0
			4	200	10049.4	12452.0	12759.9	14501.9	14525.0
			5	200	5944.7	7120.6	7402.4	7402.4	7713.0
3	10	100	3	150	5752.1	7620.9	7984.6	8590.3	8895.0
			3	200	6279.6	9199.0	9700.9	11833.0	11833.0
			4	150	3502.7	4886.5	4910.2	5110.1	5323.0
			4	200	7922.9	10961.1	11273.4	14036.5	14060.0
			5	200	2865.5	5877.7	6063.1	6063.1	6737.0
4	4	64	2	100	669	1047	1214	1214	1337
			3	200	7807	9198	9491	9492	10157
			3	300	13421	16414	17317	19105	19188
			4	200	2034	2297	2447	2447	2640
			4	300	13645	17860	18721	20139	20139
4	5	125	3	200	6643.8	8070.1	8295.6	8785.8	9249.0
			3	300	11854.3	15035.4	16606.8	18240.0	18240.0
			4	200	1443.7	1700.8	1739.8	1917.7	2122.0
			4	300	13186.4	17579.4	18450.8	19618.0	19635.0
			5	300	5142.0	10494.4	11299.5	11804.9	12343.0
4	10	1000	3	200	5648.8	8142.9		9353.0	9353.0
			3	300	9143.5	13687.6		17853.0	17853.0
			4	200	-104.4	440.0		642.4	919.0
			4	300	11029.5	15854.7		18508.0	18432.0
			5	300	724.9	6125.5		6740.1	7368.0
4	15	3375	3	200	5111.3	7072.0		7646.6	8222.0
			3	300	9030.9	13040.7		17615.0	17615.0
5	3	81	2	150	2490.9	2673.5	2763.4	2763.3	2975.4
			2	200	8888.4	11192.5	11313.7	11525.6	11525.6
			3	250	7809.0	8800.3	9065.4	9104.8	9382.3
			3	300	14702.3	16898.8	17214.3	17308.7	18156.1
5	4	256	3	300	12367.7	15591.0	15989.0	16511.1	17192.6
5	5	625	3	300	11134.5	15053.4		16191.6	16680.0
			4	300	3651.5	9494.8		10628.0	10628.0
5	6	1296	3	400	15013.4	20271.7		25157.0	25157.0
6	4	1024	3	400	15492.3	21457.9		28163.0	28163.0
			4	400	7124.3	14805.2		15400.6	15473.0
			4	500	14445.4	24887.7		32973.0	32973.0
7	3	729	3	300	24.3	1495.9		2137.3	2383.0
			3	400	12994.7	17774.1		19554.5	19983.0
			4	400	3267.5	4303.1		5388.1	5427.0
8	3	2187	3	500	16149.1	25321.4		30259.0	30259.0
			4	600	16892.3	27608.1		28717.7	28747.0

Table 5 Running times for the larger-size newsvendor problem instances

T	BR	$\lvert\Xi^T\rvert$	I	B	Time (s)			
					ν^{LDR}	$\nu^{\text{2S-LDR}}$	$\bar{\nu}_R^{\text{NA-DO}}$	PI
4	10	1000	3	200	38	31	89	1
			3	300	50	31	20	1
			4	200	47	28	637	1
			4	300	58	30	2712	1
			5	300	59	33	3088	2
	15	3375	3	200	158	140	558	3
	20	8000	3	300	726	675	2313	7
5	5	625	3	300	16	15	1423	1
			4	300	22	19	$>1h$	1
	6	1296	3	400	54	54	64	1
6	4	1024	3	400	57	53	99	1
			4	400	70	67	12	2
			4	500	74	69	3405	2
7	3	729	3	300	37	38	280	1
			3	400	37	36	$>1h$	1
			4	400	50	50	$>1h$	1
8	3	2187	3	500	364	335	408	4
			4	600	462	464	$>1h$	5

D.2. Location-Transportation Problem

Table 6 Detailed results for the location-transportation problem over a budgeted uncertainty set

(T,I,J,α^d)	α^u	$\nu^{\text{2S-LDR}}$	$\bar{\nu}_R^{\text{NA-DO}}$
$(3,10,10,0.1)$	0.1	738969.9	684460.5
	0.4	738967.8	647069.1
	0.7	738967.0	573052.3
	1	738969.3	574580.2
$(3,10,10,0.3)$	0.1	1657362.2	1457789.5
	0.4	1637485.0	1364705.6
	0.7	1535360.3	1205974.5
	1	1554797.6	1223643.1
$(3,10,10,0.5)$	0.1	1641883.1	1430838.7
	0.4	1605346.2	1363127.9
	0.7	1646129.3	1178537.7
	1	1592763.4	1474506.6
$(3,10,15,0.1)$	0.1	1133251.5	857198.3
	0.4	1133251.5	825580.3
	0.7	1157049.6	1073536.8
	1	1156385.8	956570.9
$(3,10,10,0.3)$	0.1	1157948.4	1062129.6
	0.4	1157948.4	861882.4
	0.7	1157948.4	1092108.1
	1	1157948.4	1000064.7
$(3,10,10,0.5)$	0.1	1157948.4	831435.5
	0.4	1157948.4	981328.1
	0.7	1156079.7	946148.4
	1	1116748.8	885303.1
$(4,5,5,0.1)$	0.1	770943.7	724852.8
	0.4	770943.7	573157.6
	0.7	770943.7	687320.0
	1	770943.7	727980.6
$(4,5,5,0.3)$	0.1	770943.7	710778.0
	0.4	770943.7	631020.7
	0.7	770943.7	664643.9
	1	770943.7	696703.2
$(4,5,5,0.5)$	0.1	770943.7	562210.5
	0.4	770943.7	662298.1
	0.7	770943.7	620073.7
	1	770943.7	729544.4
$(4,5,7,0.1)$	0.1	1052392.2	908927.4
	0.4	1052392.2	847845.0
	1	1052392.2	794593.6
$(4,5,7,0.3)$	0.1	1052392.2	848889.1
	0.4	1052392.2	824873.8
	0.7	1052392.2	984627.8
	1	1052392.7	855154.4
$(4,5,7,0.5)$	0.1	1052392.2	826648.8
	0.4	1052393.0	767446.4
	0.7	1052392.4	863507.3
	1	1052392.9	806706.2
$(4,5,10,0.1)$	0.1	2414351.2	2008281.8
	0.4	2414351.2	1736011.9
	0.7	2414351.2	2168851.3
	1	2414351.2	1954758.7
$(4,5,10,0.3)$	0.1	2414351.2	1912871.0
	0.4	2414351.2	1675507.4
	0.7	2414351.2	1747647.4
	1	2414351.5	2071113.7
$(4,5,10,0.5)$	0.1	2414351.1	1745320.2
	0.4	2342846.9	2076574.6
	0.7	2386394.3	1959412.9
	1	2386394.3	1824441.4
$(4,10,10,0.1)$	0.1	2502997.3	2064097.4
	0.4	2508960.6	1842514.1
	0.7	2508960.5	1688369.1
	1	2508960.8	2237510.8
$(4,10,10,0.3)$	0.1	2344765.1	2189919.0
	0.4	2508960.6	2095408.1
	0.7	2508961.0	2071323.3
	1	2401193.3	1728094.5
$(4,10,10,0.5)$	0.1	2432431.8	2130923.4
	0.4	2474962.8	2203068.7
	0.7	2347843.6	1921134.8
	1	2347843.6	1743911.8
$(5,5,10,0.1)$	0.1	3562840.8	3075701.1
	0.4	3595989.4	2855492.7
	0.7	3563956.9	2788612.9
	1	3563956.4	3256824.0
$(5,5,10,0.3)$	0.1	3623768.3	2694805.6
	0.4	3501499.4	2826482.5
	0.7	3649157.4	2483339.2
	1	3470084.5	2450441.3
$(5,5,10,0.5)$	0.1	3630095.7	2474291.5
	0.4	3573873.7	2622461.9
	0.7	3610736.5	2973602.4
	1	3489038.0	2543037.8

Table 7 Algorithmic details for the location-transportation problem over a budgeted uncertainty set

(T,I,J,α^d)	α^u	ν^{LDR}	$\nu^{\text{2S-LDR}}$, C&CG			$\nu^{\Omega(\text{LDR})}$	$\nu^{\Omega(\text{2S-LDR})}$	$\bar{\nu}_R^{\text{NA-DO}}$		
		Time (s)	Time (s)	#iterations	Gap	Time (s)	Time (s)	Time (s)	#iterations	Gap
$(3,10,10,0.1)$	0.1	0.5	168.3	2	0.0%	2.4	1.4	1376.9	617	4.9%
	0.4	6.5	59.6	2	0.0%	9.2	11.5	1113.5	809	2.6%
	0.7	10.8	87.0	2	0.0%	9.3	9.8	1088.4	726	4.3%
	1	1.3	141.8	2	0.0%	16.7	10.9	$>10h$	1315	7.9%
$(3,10,10,0.3)$	0.1	14.6	867.2	12	4.8%	2.2	0.1	$>10h$	1013	8.3%
	0.4	4.7	51.9	2	4.9%	8.1	16.5	1082.4	242	4.1%
	0.7	12.6	66.1	2	1.2%	3.1	1.6	$>10h$	730	7.1%
	1	8.4	71.6	2	4.9%	2.6	0.9	588.0	984	3.0%
$(3,10,10,0.5)$	0.1	12.9	124.6	2	4.7%	7.9	10.5	$>10h$	271	7.0%
	0.4	16.3	111.8	3	4.9%	1.0	0.7	$>10h$	665	9.1%
	0.7	1.8	83.2	3	4.3%	1.7	0.7	1132.4	1064	3.2%
	1	25.2	96.5	3	3.9%	2.1	2.2	$>10h$	1405	4.1%
$(3,10,15,0.1)$	0.1	10.2	150.1	3	1.2%	2.1	1.3	2634.8	258	2.9%
	0.4	3.5	112.6	2	1.2%	2.3	1.1	$>10h$	337	6.9%
	0.7	29.3	96.6	2	0.0%	1.7	2.0	$>10h$	432	5.5%
	1	8.3	29.3	2	0.0%	2.2	2.7	$>10h$	824	9.9%
$(3,10,10,0.3)$	0.1	2.5	169.7	2	1.1%	3.0	2.4	4193.8	771	2.1%
	0.4	13.4	95.6	2	1.2%	3.8	0.4	4706.4	974	4.9%
	0.7	13.0	94.6	2	1.1%	2.3	1.2	5043.6	475	2.0%
	1	10.2	119.9	2	1.1%	1.6	2.0	5015.7	1187	4.4%
$(3,10,10,0.5)$	0.1	5.7	97.1	6	1.3%	2.0	1.9	3757.0	1253	1.1%
	0.4	2.8	155.3	2	1.2%	3.0	1.5	2902.3	803	4.5%
	0.7	0.7	112.0	2	0.0%	8.5	1.6	$>10h$	884	9.7%
	1	2.3	89.9	2	1.2%	2.3	4.5	3898.8	1120	4.6%
$(4,5,5,0.1)$	0.1	0.2	0.8	2	0.0%	7.3	16.5	1198.6	951	4.8%
	0.4	0.3	1.2	2	0.0%	7.1	21.8	$>10h$	1483	7.2%
	0.7	0.6	1.1	2	0.0%	2.5	5.4	1558.8	629	2.1%
	1	0.5	1.2	2	0.0%	1.4	2.6	$>10h$	1267	7.3%
$(4,5,5,0.3)$	0.1	0.7	1.3	2	0.0%	5.6	0.9	1649.9	1399	0.6%
	0.4	0.4	0.9	2	0.0%	8.2	3.0	$>10h$	569	9.0%
	0.7	0.3	0.4	2	0.0%	3.5	1.1	905.6	1016	3.3%
	1	0.9	0.4	2	0.0%	1.5	3.3	$>10h$	548	8.4%
$(4,5,5,0.5)$	0.1	0.4	0.9	2	0.0%	2.2	0.0	$>10h$	550	7.0%
	0.4	0.5	1.3	2	0.0%	9.5	14.1	$>10h$	284	7.0%
	0.7	0.2	1.4	2	0.0%	3.9	2.0	1064.2	920	2.6%
	1	0.2	1.2	2	0.0%	3.9	2.5	$>10h$	912	5.6%
$(4,5,7,0.1)$	0.1	4.7	77.2	2	0.0%	1.4	0.9	1207.0	332	4.8%
	0.4	9.7	133.8	2	0.0%	4.8	10.1	$>10h$	433	8.7%
	1	1.1	159.8	2	0.0%	2.7	5.5	$>10h$	1073	6.8%
$(4,5,7,0.3)$	0.1	2.4	123.2	2	0.0%	2.9	4.6	886.7	864	4.9%
	0.4	6.6	197.4	2	0.0%	4.9	2.1	1364.0	1361	4.4%
	0.7	5.3	179.7	2	0.0%	2.2	2.1	$>10h$	787	6.9%
	1	3.2	156.8	2	0.0%	3.7	0.8	$>10h$	966	9.3%
$(4,5,7,0.5)$	0.1	6.2	121.9	2	0.0%	1.5	1.2	1047.0	1296	4.5%
	0.4	3.9	101.3	2	0.0%	10.3	14.3	988.4	392	4.7%
	0.7	25.5	135.8	2	0.0%	2.2	6.6	970.9	562	4.8%
	1	4.6	204.7	2	0.0%	4.1	3.1	$>10h$	591	8.9%
$(4,5,10,0.1)$	0.1	4.8	83.7	2	1.7%	3.0	7.2	$>10h$	1508	9.4%
	0.4	13.6	88.8	2	1.6%	1.1	1.0	$>10h$	723	6.4%
	0.7	6.8	102.2	2	1.6%	2.1	5.0	4502.0	479	4.9%
	1	6.4	118.4	2	1.8%	12.6	4.2	$>10h$	566	7.4%
$(4,5,10,0.3)$	0.1	17.9	199.3	2	1.8%	3.8	8.2	7782.2	1059	4.2%
	0.4	25.3	103.5	2	1.8%	6.7	6.3	12134.1	667	2.6%
	0.7	13.8	50.2	2	1.7%	9.0	10.3	11770.6	586	4.5%
	1	8.6	107.1	2	1.7%	2.0	3.8	9143.7	421	2.2%
$(4,5,10,0.5)$	0.1	10.4	189.7	2	1.7%	1.2	0.6	$>10h$	366	9.0%
	0.4	6.2	38.2	2	1.8%	5.7	5.5	$>10h$	1429	9.3%
	0.7	4.0	72.6	2	1.8%	3.1	3.1	12449.6	617	4.9%
	1	1.0	41.1	2	1.8%	17.4	40.7	11130.6	739	4.3%

Table 7 Algorithmic details for the location-transportation problem over a budgeted uncertainty set (continued)

(T, I, J, α^d)	α^u	ν^{LDR}	$\nu^{\text{2S-LDR}}$, C&CG			$\nu^{\Omega(\text{LDR})}$	$\nu^{\Omega(\text{2S-LDR})}$	$\bar{\nu}_R^{\text{NA-DO}}$		
		Time (s)	Time (s)	#iterations	Gap	Time (s)	Time (s)	Time (s)	#iterations	Gap
(4, 10, 10, 0.1)	0.1	10.7	13318.8	1	4.1%	21.0	43.0	$> 10h$	1880	5.2%
	0.4	18.9	13171.6	1	4.8%	14.8	9.2	$> 10h$	2154	8.7%
	0.7	33.7	4419.5	2	4.6%	18.0	49.5	$> 10h$	776	8.2%
	1	3.4	7159.0	2	4.8%	14.6	6.9	14098.5	1275	4.1%
(4, 10, 10, 0.3)	0.1	35.8	10763.5	1	2.4%	4.6	9.1	9702.9	885	4.8%
	0.4	12.7	7688.3	2	4.8%	9.3	8.2	$> 10h$	755	8.6%
	0.7	24.5	12108.4	1	4.6%	8.5	5.2	$> 10h$	1343	13.4%
	1	33.8	7580.7	2	3.2%	8.8	1.2	13045.5	862	4.9%
(4, 10, 10, 0.5)	0.1	23.1	8361.0	2	2.4%	36.8	7.7	10972.8	765	1.5%
	0.4	17.5	9441.9	2	4.4%	36.9	69.7	$> 10h$	791	5.7%
	0.7	16.8	5638.9	2	2.3%	38.5	50.6	14347.7	1511	4.4%
	1	26.7	3497.6	2	2.4%	29.5	20.6	$> 10h$	696	8.7%
(5, 5, 10, 0.1)	0.1	2.3	372.6	4	1.6%	34.3	17.0	$> 10h$	1439	7.5%
	0.4	5.8	989.5	4	2.8%	28.3	38.4	$> 10h$	267	8.2%
	0.7	5.4	657.6	2	1.9%	21.4	31.2	15167.1	605	4.6%
	1	3.6	784.6	3	1.7%	31.3	52.0	$> 10h$	904	7.0%
(5, 5, 10, 0.3)	0.1	1.8	883.5	2	2.8%	39.3	84.0	$> 10h$	789	5.2%
	0.4	5.5	607.0	2	1.9%	24.5	23.9	16175.2	1053	4.8%
	0.7	4.5	414.9	10	2.8%	27.4	33.6	13106.8	762	4.8%
	1	3.5	959.7	2	1.3%	19.5	32.5	14376.6	1016	2.0%
(5, 5, 10, 0.5)	0.1	5.0	1009.4	2	1.8%	35.5	94.5	$> 10h$	418	5.6%
	0.4	6.2	723.7	2	3.0%	38.5	52.0	13067.1	1329	4.5%
	0.7	7.9	383.6	2	2.5%	39.5	53.6	15368.2	911	4.0%
	1	9.9	724.3	2	1.6%	39.5	66.5	14150.1	521	4.7%

D.3. Capital Budgeting with Loan

Table 8 Results for the capital budgeting problem with unrestricted loans

Instance	(T, I)	B	ν^{κ}	UB			Optimality Gap (%)		
				ν^{Ω}	$\bar{\nu}_R^{\text{NA-DO}}$	$\bar{\nu}_R^{\text{DNA-DO}}$	$\left(\frac{\nu^{\Omega}-\nu^{\kappa}}{\nu^{\kappa}}\right)$	$\left(\frac{\bar{\nu}_R^{\text{NA-DO}}-\nu^{\kappa}}{\nu^{\kappa}}\right)$	$\left(\frac{\bar{\nu}_R^{\text{DNA-DO}}-\nu^{\kappa}}{\nu^{\kappa}}\right)$
1	(3,5)	0	1.8	2.2	2.1	2.1	19.7%	16.1%	14.5%
2		50	6.7	7.9	7.6	7.5	18.0%	14.1%	12.6%
3		100	7.6	8.8	8.5	8.4	16.2%	12.8%	10.4%
4		150	7.8	8.8	8.5	8.3	13.1%	9.1%	7.0%
5	(3,10)	0	6.7	7.9	7.6	7.5	18.7%	13.8%	12.3%
6		50	11.7	13.9	13.4	13.1	18.9%	14.5%	11.8%
7		100	16.7	19.7	19.2	18.8	18.5%	15.2%	12.6%
8		150	16.7	19.7	19.1	18.7	18.0%	14.1%	11.8%
9	(3,15)	0	8.9	10.4	10.1	9.9	16.5%	13.3%	11.0%
10		50	13.7	16.4	15.9	15.5	19.4%	15.8%	13.0%
11		100	18.9	22.4	21.6	21.1	18.7%	14.3%	11.8%
12		150	21.6	26.0	25.1	24.6	20.4%	16.3%	13.9%
13	(3,20)	0	11.9	13.8	13.2	13.0	16.0%	11.2%	8.9%
14		50	16.4	19.8	19.2	18.9	21.1%	17.6%	15.6%
15		100	21.7	25.8	25.1	24.7	19.0%	15.8%	13.8%
16		150	27.1	31.8	31.0	30.4	17.5%	14.4%	12.4%
17	(3,25)	0	17.7	20.2	19.1	18.8	14.2%	8.3%	6.7%
18		50	22.5	26.2	25.3	24.9	16.5%	12.7%	10.9%
19		100	26.5	32.2	31.2	30.8	21.6%	17.8%	16.4%
20		150	31.7	38.2	36.8	36.2	20.5%	16.1%	14.3%
21	(3,30)	0	22.2	26.4	25.6	25.3	18.9%	15.4%	13.9%
22		50	26.5	32.4	31.2	30.8	22.3%	17.9%	16.5%
23		100	31.4	38.4	37.4	36.9	22.3%	19.1%	17.6%
24		150	37.1	44.4	42.9	42.4	19.7%	15.8%	14.3%

Table 8 Results for the capital budgeting problem with unrestricted loans (continued)

Instance	(T, I)	B	ν^{K}	UB			Optimality Gap (%)		
				ν^{Ω}	$\bar{\nu}_R^{\text{NA-DO}}$	$\bar{\nu}_R^{\text{DNA-DO}}$	$\left(\frac{\nu^{\Omega}-\nu^{\text{K}}}{\nu^{\text{K}}}\right)$	$\left(\frac{\bar{\nu}_R^{\text{NA-DO}}-\nu^{\text{K}}}{\nu^{\text{K}}}\right)$	$\left(\frac{\bar{\nu}_R^{\text{DNA-DO}}-\nu^{\text{K}}}{\nu^{\text{K}}}\right)$
25	(4,5)	0	3.3	3.8	3.6	3.5	17.1%	10.7%	8.0%
26		50	8.3	9.9	9.4	9.2	19.2%	13.7%	11.1%
27		100	8.7	10.2	9.7	9.5	17.2%	12.0%	9.5%
28		150	7.6	9.1	8.6	8.4	18.8%	13.1%	10.3%
29		200	8.7	10.2	9.9	9.6	17.4%	13.4%	10.7%
30		250	8.7	10.2	9.6	9.3	16.7%	9.7%	6.9%
31	(4,10)	0	8.4	9.9	9.5	9.2	18.3%	12.7%	9.9%
32		50	13.4	16.0	15.1	14.7	19.2%	12.3%	9.4%
33		100	18.0	22.0	21.2	20.8	22.1%	17.7%	15.4%
34		150	20.8	25.0	23.9	23.3	19.7%	14.7%	11.7%
35		200	20.4	25.0	24.1	23.5	22.2%	18.0%	15.2%
36		250	21.1	25.0	24.0	23.4	18.1%	13.5%	10.7%
37	(4,15)	0	12.9	15.3	14.5	14.2	18.0%	11.7%	9.3%
38		50	18.0	21.3	20.0	19.5	18.2%	11.0%	8.2%
39		100	23.1	27.3	26.3	25.8	18.2%	14.0%	11.9%
40		150	28.3	33.3	32.1	31.4	17.8%	13.7%	11.3%
41		200	31.6	38.2	36.9	36.0	20.9%	16.6%	13.8%
42		250	32.3	38.2	36.2	35.5	18.1%	12.0%	9.7%
43	(4,20)	0	16.0	19.3	18.8	18.4	20.6%	17.0%	15.1%
44		50	21.3	25.3	23.9	23.4	18.9%	12.0%	9.6%
45		100	25.8	31.3	30.4	29.7	21.6%	17.8%	15.4%
46		150	31.3	37.3	35.2	34.5	19.2%	12.4%	10.1%
47		200	36.6	43.3	41.3	40.6	18.3%	12.9%	11.0%
48		250	39.7	48.3	46.8	45.9	21.7%	17.8%	15.6%
49	(4,25)	0	22.7	27.7	26.8	26.3	22.2%	18.2%	16.0%
50		50	26.8	33.7	31.8	31.3	25.6%	18.5%	16.7%
51		100	32.4	39.7	37.7	37.1	22.5%	16.4%	14.5%
52		150	37.9	45.7	43.7	42.9	20.5%	15.3%	13.1%
53		200	42.2	51.7	50.1	49.3	22.6%	18.7%	16.9%
54		250	46.8	57.7	55.6	54.6	23.3%	18.8%	16.7%
55	(4,30)	0	23.0	27.9	26.3	25.9	21.6%	14.6%	12.6%
56		50	26.1	33.9	32.8	32.5	29.7%	25.5%	24.2%
57		100	32.6	39.9	38.0	37.5	22.3%	16.4%	14.9%
58		150	36.4	45.9	43.7	43.1	26.1%	20.1%	18.4%
59		200	41.5	51.9	49.6	48.9	25.1%	19.6%	17.8%
60		250	46.1	57.9	55.5	54.8	25.7%	20.5%	19.0%
61	(5,5)	0	4.9	6.1	5.8	5.6	23.6%	17.4%	14.5%
62		50	10.5	13.0	12.3	12.0	23.5%	16.9%	13.3%
63		100	14.4	17.6	16.7	16.2	22.4%	16.4%	13.0%
64		150	14.1	17.6	17.0	16.5	24.2%	20.0%	16.6%
65		200	14.4	17.6	16.7	16.2	21.7%	15.6%	12.1%
66		250	14.0	17.6	16.7	16.3	25.6%	19.3%	16.5%
67		300	14.3	17.6	16.9	16.4	22.6%	17.7%	14.5%
68		350	14.4	17.6	16.6	16.2	22.0%	15.6%	12.3%
69	(5,10)	0	10.0	12.0	11.3	11.0	19.5%	12.6%	9.6%
70		50	14.0	18.0	17.2	16.9	29.0%	23.3%	20.9%
71		100	19.7	24.0	23.1	22.6	21.9%	17.6%	14.6%
72		150	24.0	30.0	28.5	27.7	24.9%	18.8%	15.4%
73		200	24.6	30.0	28.8	28.0	21.9%	17.1%	13.8%
74		250	23.1	30.0	28.8	28.2	29.9%	24.8%	22.3%
75		300	23.5	30.0	28.7	28.0	27.7%	22.4%	19.3%
76		350	22.7	30.0	28.4	27.8	31.9%	25.1%	22.4%

Table 8 Results for the capital budgeting problem with unrestricted loans (continued)

Instance	(T, I)	B	ν^{K}	UB			Optimality Gap (%)		
				ν^{Ω}	$\bar{\nu}_R^{\text{NA-DO}}$	$\bar{\nu}_R^{\text{DNA-DO}}$	$\left(\frac{\nu^{\Omega}-\nu^{\mathrm{K}}}{\nu^{\mathrm{K}}}\right)$	$\left(\frac{\bar{\nu}_R^{\text{NA-DO}}-\nu^{\mathrm{K}}}{\nu^{\mathrm{K}}}\right)$	$\left(\frac{\bar{\nu}_R^{\text{DNA-DO}}-\nu^{\mathrm{K}}}{\nu^{\mathrm{K}}}\right)$
77	(5,15)	0	12.7	16.2	15.3	14.9	28.1%	20.8%	17.5%
78		50	17.6	22.2	21.5	21.1	26.2%	22.3%	19.6%
79		100	21.7	28.2	26.6	25.8	29.9%	22.5%	18.9%
80		150	26.2	34.2	32.4	31.5	30.5%	23.7%	20.0%
81		200	30.8	40.2	38.9	38.1	30.7%	26.3%	23.8%
82		250	31.6	40.5	39.0	38.2	28.1%	23.3%	20.8%
83		300	30.7	40.5	39.2	38.4	32.2%	27.8%	25.4%
84		350	31.3	40.5	38.8	38.0	29.6%	24.0%	21.6%
85	(5,20)	0	19.1	23.6	22.9	22.3	23.1%	19.6%	16.4%
86		50	23.1	29.6	28.6	27.9	27.9%	23.7%	20.7%
87		100	28.1	35.6	34.4	33.7	26.6%	22.6%	20.0%
88		150	33.8	41.6	39.7	38.9	23.1%	17.7%	15.3%
89		200	38.2	47.6	46.5	45.4	24.5%	21.8%	18.9%
90		250	39.9	53.6	52.4	51.1	34.2%	31.2%	28.0%
91		300	46.8	58.9	57.0	55.7	25.9%	21.7%	19.1%
92		350	45.9	58.9	57.7	56.6	28.4%	25.7%	23.3%
93	(5,25)	0	23.5	29.6	29.0	28.4	25.9%	23.4%	20.9%
94		50	28.6	35.6	34.7	33.9	24.3%	21.3%	18.3%
95		100	33.4	41.6	39.9	39.3	24.6%	19.7%	17.6%
96		150	36.5	47.6	45.9	44.9	30.5%	25.9%	23.1%
97		200	42.6	53.6	51.7	50.6	25.9%	21.4%	19.0%
98		250	47.1	59.6	58.2	56.9	26.4%	23.4%	20.6%
99		300	51.6	65.6	63.5	62.6	27.0%	23.0%	21.2%
100		350	56.1	71.6	69.7	68.3	27.6%	24.3%	21.8%
101	(5,30)	0	25.6	32.2	31.1	30.6	26.1%	21.8%	19.8%
102		50	29.1	38.2	37.1	36.5	31.3%	27.4%	25.3%
103		100	34.4	44.2	43.3	42.6	28.7%	26.0%	24.1%
104		150	37.8	50.2	48.7	48.1	32.8%	28.7%	27.2%
105		200	42.0	56.2	54.7	53.5	33.9%	30.2%	27.4%
106		250	47.2	62.2	60.3	59.5	31.9%	27.8%	26.2%
107		300	54.4	68.2	66.4	65.3	25.5%	22.1%	20.0%
108		350	60.6	74.2	72.1	70.7	22.4%	18.9%	16.5%
109	(6,5)	0	5.6	6.8	6.6	6.4	20.2%	18.0%	14.1%
110		50	10.6	12.8	12.5	12.1	20.5%	18.0%	14.2%
111		100	13.6	16.9	16.3	15.8	24.0%	19.8%	16.4%
112		150	12.8	16.9	16.5	16.0	31.4%	28.4%	24.8%
113		200	13.2	16.9	16.5	16.0	27.7%	24.9%	21.2%
114		250	12.9	16.9	16.5	15.9	30.5%	27.7%	23.1%
115		300	13.0	16.9	16.5	15.9	30.1%	27.2%	22.8%
116		350	13.3	16.9	16.5	16.0	27.2%	24.2%	20.3%
117		400	12.9	16.9	16.4	15.9	30.7%	26.9%	23.1%
118	(6,10)	0	10.9	14.0	13.8	13.4	27.8%	26.0%	22.5%
119		50	15.9	20.0	19.6	18.9	25.8%	23.5%	19.0%
120		100	20.7	26.0	25.4	24.5	25.4%	22.4%	18.2%
121		150	24.9	32.0	31.4	30.3	28.3%	25.8%	21.5%
122		200	27.1	35.0	34.2	33.3	29.2%	26.5%	23.1%
123		250	26.6	35.0	34.2	33.2	31.5%	28.7%	24.7%
124		300	26.8	35.0	34.4	33.3	30.4%	28.1%	24.0%
125		350	27.3	35.0	34.4	33.3	28.3%	26.3%	22.2%
126		400	26.9	35.0	34.2	33.2	29.9%	27.0%	23.3%

Table 8 Results for the capital budgeting problem with unrestricted loans (continued)

Instance	(T, I)	B	ν^{K}	UB			Optimality Gap (%)		
				ν^{Ω}	$\bar{\nu}_R^{\text{NA-DO}}$	$\bar{\nu}_R^{\text{DNA-DO}}$	$\left(\frac{\nu^{\Omega}-\nu^{\mathrm{K}}}{\nu^{\mathrm{K}}}\right)$	$\left(\frac{\bar{\nu}_R^{\text{NA-DO}}-\nu^{\mathrm{K}}}{\nu^{\mathrm{K}}}\right)$	$\left(\frac{\bar{\nu}_R^{\text{DNA-DO}}-\nu^{\mathrm{K}}}{\nu^{\mathrm{K}}}\right)$
127		0	15.9	19.9	19.5	19.0	24.7%	22.5%	19.3%
128		50	19.8	25.9	25.5	24.9	30.7%	28.7%	25.8%
129		100	25.3	31.9	31.4	30.6	25.8%	23.9%	20.9%
130		150	28.9	37.9	37.3	36.4	31.2%	29.3%	26.1%
131	(6,15)	200	33.5	43.9	43.1	42.1	30.8%	28.3%	25.5%
132		250	37.0	49.7	48.7	47.6	34.3%	31.7%	28.7%
133		300	36.9	49.7	49.0	47.8	34.8%	32.9%	29.7%
134		350	38.5	49.7	48.8	47.6	29.1%	26.8%	23.8%
135		400	37.9	49.7	48.9	47.7	31.2%	29.0%	25.9%
136		0	22.4	28.6	28.3	27.8	27.6%	26.5%	23.9%
137		50	27.6	34.6	34.3	33.5	25.4%	24.3%	21.6%
138		100	31.7	40.6	40.1	39.3	27.9%	26.3%	23.7%
139		150	36.9	46.6	46.1	45.3	26.4%	25.2%	22.8%
140	(6,20)	200	41.0	52.6	52.1	50.9	28.4%	27.2%	24.4%
141		250	45.1	58.6	58.2	57.1	29.8%	28.8%	26.4%
142		300	49.7	64.6	63.8	62.6	30.0%	28.3%	26.0%
143		350	53.2	70.6	70.0	68.8	32.6%	31.5%	29.3%
144		400	54.3	71.5	70.9	69.6	31.6%	30.6%	28.1%
145		0	29.0	37.2	36.9	36.2	28.2%	27.2%	24.9%
146		50	32.8	43.2	42.8	42.1	31.5%	30.3%	28.2%
147		100	37.5	49.2	48.7	48.0	31.3%	30.0%	28.0%
148		150	42.5	55.2	54.6	53.7	29.9%	28.5%	26.3%
149	(6,25)	200	46.2	61.2	60.5	59.7	32.5%	31.0%	29.2%
150		250	51.9	67.2	66.7	65.7	29.4%	28.5%	26.6%
151		300	56.1	73.2	72.3	71.2	30.4%	28.9%	26.8%
152		350	61.1	79.2	78.6	77.4	29.6%	28.6%	26.7%
153		400	64.5	85.2	84.5	83.0	32.0%	30.9%	28.7%
154		0	31.5	42.6	42.2	41.6	35.2%	34.1%	32.2%
155		50	37.5	48.6	48.1	47.5	29.5%	28.3%	26.6%
156		100	40.6	54.6	54.1	53.4	34.4%	33.1%	31.5%
157		150	45.2	60.6	60.0	59.2	33.9%	32.6%	30.8%
158	(6,30)	200	49.1	66.6	66.0	65.1	35.5%	34.4%	32.5%
159		250	56.0	72.6	71.9	71.0	29.7%	28.5%	26.9%
160		300	59.4	78.6	78.0	76.8	32.4%	31.4%	29.4%
161		350	65.0	84.6	83.8	82.6	30.2%	29.0%	27.1%
162		400	68.5	90.6	89.8	88.5	32.2%	31.1%	29.2%

D.4. Capital Budgeting with Integer Loan

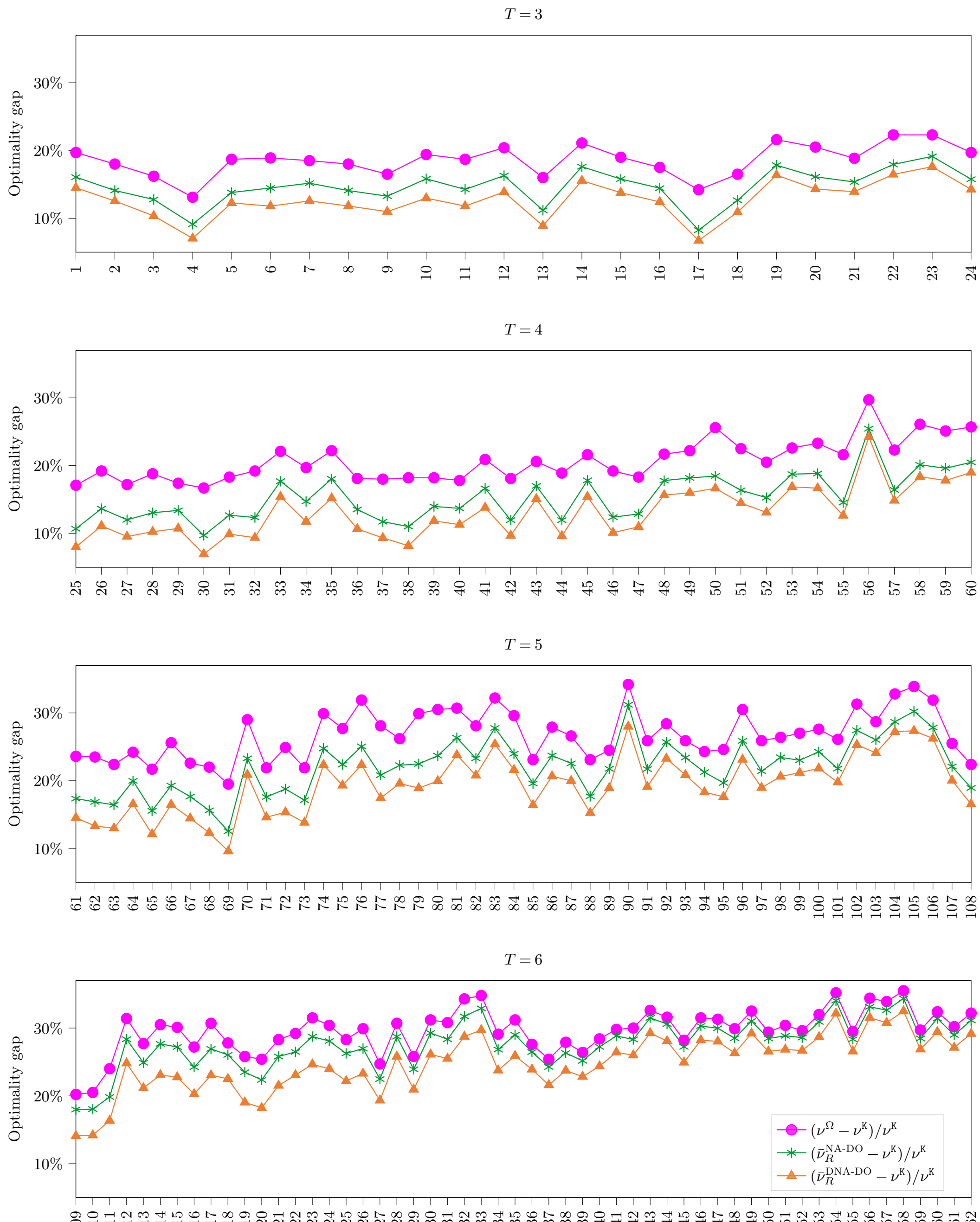

Figure 9 Optimality gap improvements for capital budgeting instances with general integer loan decisions.

Table 9 Solution times of different bounding methods in seconds for the capital budgeting problem

Instance	(T,I)	B	Fractional loan ν^{κ}	Fractional loan ν^{Ω}	Fractional loan $\bar{\nu}_R^{\mathrm{NA}}$	Fractional loan $\bar{\nu}_R^{\mathrm{DNA}}$	Integral loan ν^{κ}	Integral loan ν^{Ω}	Integral loan $\bar{\nu}_R^{\mathrm{NA}}$	Integral loan $\bar{\nu}_R^{\mathrm{DNA}}$
1	(3,5)	0	27.0	63.0	56.1	62.1	67.2	72.7	112.3	80.8
2		50	16.2	85.2	83.8	74.9	52.6	99.3	678.9	142.2
3		100	25.8	60.8	55.5	52.2	43.5	95.6	260.8	287.1
4		150	25.2	56.2	55.3	35.5	64.0	59.3	259.9	156.1
5	(3,10)	0	35.4	547.4	547.4	521.2	111.3	840.0	1133.1	948.6
6		50	40.5	384.5	254.5	376.1	77.5	771.1	413.1	557.9
7		100	34.5	340.5	392.5	366.2	96.5	373.2	478.1	432.2
8		150	44.4	514.4	657.4	471.2	107.7	1215.7	1447.6	604.0
9	(3,15)	0	90.9	1502.9	1178.9	1390.9	218.5	2029.6	3812.6	3569.0
10		50	90.0	1169.0	1007.0	1061.0	201.8	1467.1	2698.8	2868.9
11		100	96.3	1438.3	1168.3	1556.3	169.1	2261.3	2553.9	5394.1
12		150	94.8	1535.8	1247.8	1355.8	200.5	1974.3	3441.4	3313.6
13	(3,20)	0	611.4	3138.4	1896.4	2568.4	689.5	4305.1	3531.1	5606.8
14		50	653.1	3809.1	2498.1	3005.1	772.2	5003.3	3962.0	7371.5
15		100	667.2	3057.2	2643.2	4295.2	743.4	3777.1	5651.2	12202.7
16		150	609.6	3539.6	2918.6	2557.6	765.9	4162.9	5034.6	5319.8
17	(3,25)	0	4292.4	8689.4	4392.9	5521.4	4561.3	9868.9	7116.6	10188.8
18		50	3762.0	6794.0	3465.3	5684.5	3953.5	7354.5	4636.6	11134.0
19		100	3759.0	6921.0	3830.0	4160.5	4054.2	8311.5	6707.6	6271.3
20		150	4012.8	7998.8	4716.5	4608.3	4336.0	8751.4	6883.0	8159.8
21	(3,30)	0	1808.1	5929.1	3754.7	8207.1	2188.2	6894.2	6115.2	20660.0
22		50	2083.8	5512.8	2817.2	8133.8	2394.6	6036.5	3818.2	15524.7
23		100	2027.1	5422.1	3152.7	9087.1	2307.5	5763.8	4756.4	23493.2
24		150	1850.7	6203.7	4069.8	7591.7	2225.7	7524.8	5719.4	15765.4
25	(4,5)	0	89.2	1012.2	626.8	1185.2	173.3	3395.6	686.6	1483.1
26		50	97.2	1141.2	592.8	1236.2	190.1	1536.5	665.1	1478.5
27		100	106.0	1189.0	624.7	1259.0	173.2	2337.4	706.7	1488.1
28		150	80.8	849.8	470.5	932.8	170.7	1740.1	516.6	1071.5
29		200	85.2	838.2	390.8	952.2	160.4	3339.3	423.1	1060.8
30		250	107.2	1039.2	596.8	1173.2	202.7	2826.5	665.2	1332.0
31	(4,10)	0	590.0	2420.0	1277.3	3150.0	713.9	3186.4	1544.7	5161.8
32		50	593.2	2247.2	1050.1	3178.2	720.7	5845.5	1295.2	4602.0
33		100	580.8	2189.8	1049.2	2941.8	779.5	12174.6	1236.0	4424.5
34		150	470.0	2230.0	1449.3	3485.0	636.5	2278.7	1820.4	5099.7
35		200	417.2	2208.2	1285.5	2833.2	602.3	5185.9	1572.6	4482.1
36		250	431.6	2198.6	1353.7	3231.6	583.6	10838.1	1576.6	4862.5
37	(4,15)	0	1372.8	5153.8	3131.9	5853.8	1679.6	7835.0	4739.6	11098.8
38		50	1370.4	4828.4	2509.6	6153.4	1724.7	8235.8	3476.6	12602.2
39		100	1268.0	5458.0	3486.7	7861.0	1525.9	15446.3	5541.5	21591.5
40		150	1507.6	5316.6	2784.4	7977.6	1750.2	17220.9	3909.3	21449.1
41		200	1426.4	5243.4	3039.6	6315.4	1613.6	12527.3	4895.8	11460.3
42		250	1589.2	5702.2	2990.8	8046.2	1911.2	19692.0	4633.7	21338.5
43	(4,20)	0	4676.4	10535.4	5935.6	17371.4	5228.1	20173.7	9827.4	$>10h$
44		50	4988.0	10909.0	5992.7	16613.0	5402.2	14597.9	7704.6	$>10h$
45		100	4942.8	10226.8	5601.9	17168.8	5356.3	20126.0	9011.5	$>10h$
46		150	4111.2	9830.2	5337.5	13249.2	4642.0	27734.9	7855.0	$>10h$
47		200	4633.6	9934.6	6239.1	16356.6	5176.0	16374.7	9543.7	$>10h$
48		250	4344.4	9147.4	5522.3	14321.4	4749.8	18371.0	7313.3	$>10h$
49	(4,25)	0	8113.2	16000.2	10434.8	25348.2	8769.9	$>10h$	13120.7	$>10h$
50		50	7384.0	14978.0	9753.3	21113.0	7679.0	$>10h$	12535.0	$>10h$
51		100	7619.2	14613.2	9046.1	20856.2	7865.7	$>10h$	12054.9	$>10h$
52		150	7864.4	15145.4	9942.3	22623.4	8270.7	$>10h$	12352.3	$>10h$
53		200	7440.4	14736.4	8354.9	25575.4	7731.2	$>10h$	9985.8	$>10h$
54		250	7986.4	15734.4	8246.9	27221.4	8265.6	$>10h$	10005.2	$>10h$

Table 9 Solution times of different bounding methods in seconds for the capital budgeting problem (continued)

Instance	(T,I)	B	Fractional loan ν^{κ}	Fractional loan ν^{Ω}	Fractional loan $\bar{\nu}_R^{\mathrm{NA}}$	Fractional loan $\bar{\nu}_R^{\mathrm{DNA}}$	Integral loan ν^{κ}	Integral loan ν^{Ω}	Integral loan $\bar{\nu}_R^{\mathrm{NA}}$	Integral loan $\bar{\nu}_R^{\mathrm{DNA}}$
55	(4,30)	0	13898.8	20608.8	12197.9	$>10h$	14146.4	$>10h$	15337.6	$>10h$
56		50	13221.6	20082.6	10940.4	33232.6	13470.7	$>10h$	14137.2	$>10h$
57		100	12854.0	20045.0	13182.0	$>10h$	13435.0	$>10h$	19973.4	$>10h$
58		150	13757.6	20898.6	13569.7	$>10h$	14320.1	$>10h$	19708.7	$>10h$
59		200	12260.8	19245.8	10654.5	$>10h$	12773.9	$>10h$	14818.3	$>10h$
60		250	13656.4	20471.4	12559.6	$>10h$	14180.5	$>10h$	18532.9	$>10h$
61	(5,5)	0	1188.0	4539.0	2692.7	6018.0	1302.5	6041.9	2993.6	7706.0
62		50	1219.0	4453.0	2802.0	5922.0	1297.0	4489.5	3164.9	6700.7
63		100	1469.0	4346.0	2864.0	5529.0	1545.1	12850.6	3207.0	6774.4
64		150	1315.0	4093.0	2428.7	5442.0	1443.7	7996.8	2647.2	6284.1
65		200	1099.0	4026.0	2384.0	5303.0	1247.0	12676.3	2573.5	6182.0
66		250	1194.0	4438.0	2825.3	6075.0	1294.4	6630.7	3110.0	7217.1
67		300	812.5	3603.5	2135.7	4741.5	923.0	12057.9	2253.7	5600.9
68		350	1376.5	5019.5	3046.3	6457.5	1484.3	19850.1	3340.3	7406.8
69	(5,10)	0	4275.5	10708.5	9053.2	14882.5	4512.4	14699.0	11601.7	20366.7
70		50	3354.0	11245.0	9677.3	13469.0	3547.8	13864.4	11116.0	16596.1
71		100	3764.5	12332.5	10597.7	14540.5	3959.8	20538.2	12756.1	18039.9
72		150	3769.0	10903.0	10864.4	13991.0	3991.5	22267.6	13768.9	19160.7
73		200	3507.5	10891.5	10037.7	13358.5	3679.3	21045.9	12135.6	18143.1
74		250	4005.5	12899.5	11522.3	16746.5	4244.5	17293.1	13665.4	20972.2
75		300	3507.0	10998.0	13136.7	13005.0	3724.3	25089.7	18343.2	17303.2
76		350	3803.0	9881.0	8232.7	12413.0	4006.9	21655.3	9519.8	17092.7
77	(5,15)	0	7936.0	22038.0	14625.3	28044.0	8368.5	$>10h$	16465.2	$>10h$
78		50	8516.5	22334.5	14089.7	29802.5	8893.6	$>10h$	15924.1	$>10h$
79		100	9214.0	21936.0	14357.3	29063.0	9757.7	$>10h$	15971.1	$>10h$
80		150	9150.0	22127.0	13484.7	28673.0	9541.0	$>10h$	14723.9	$>10h$
81		200	8487.5	21143.5	13829.0	27331.5	8850.1	$>10h$	15523.1	$>10h$
82		250	7664.5	21657.5	13638.3	28499.5	8033.2	$>10h$	14898.5	$>10h$
83		300	8442.5	22287.5	14258.3	27426.5	8785.5	$>10h$	15829.6	$>10h$
84		350	9361.5	19241.5	12027.7	27496.5	9747.9	$>10h$	13072.9	$>10h$
85	(5,20)	0	12474.0	31538.0	24198.3	$>10h$	13006.9	$>10h$	28183.8	$>10h$
86		50	13804.0	30383.0	22819.2	$>10h$	14552.9	$>10h$	26801.1	$>10h$
87		100	13531.5	27682.5	21402.1	$>10h$	14425.2	$>10h$	25016.9	$>10h$
88		150	12393.5	30610.5	23529.6	$>10h$	13144.1	$>10h$	28047.3	$>10h$
89		200	13991.5	32146.5	25330.4	$>10h$	14732.0	$>10h$	31559.2	$>10h$
90		250	12486.0	28563.0	22135.8	$>10h$	13412.9	$>10h$	26835.3	$>10h$
91		300	14566.5	32117.5	24472.9	$>10h$	15184.6	$>10h$	28655.3	$>10h$
92		350	14411.0	32416.0	26284.2	$>10h$	15094.5	$>10h$	31906.3	$>10h$
93	(5,25)	0	19659.5	$>10h$	22510.6	$>10h$	20733.1	$>10h$	$>10h$	$>10h$
94		50	19808.5	$>10h$	26676.1	$>10h$	20452.7	$>10h$	$>10h$	$>10h$
95		100	21565.0	$>10h$	27794.4	$>10h$	22605.3	$>10h$	$>10h$	$>10h$
96		150	20473.5	$>10h$	25248.3	$>10h$	21390.7	$>10h$	$>10h$	$>10h$
97		200	20571.5	$>10h$	25523.9	$>10h$	21776.2	$>10h$	$>10h$	$>10h$
98		250	21483.0	$>10h$	24536.7	$>10h$	22207.9	$>10h$	$>10h$	$>10h$
99		300	22307.5	$>10h$	27119.4	$>10h$	23392.4	$>10h$	$>10h$	$>10h$
100		350	18784.5	$>10h$	25705.0	$>10h$	19746.9	$>10h$	$>10h$	$>10h$
101	(5,30)	0	16882.5	$>10h$	21286.1	$>10h$	17358.9	$>10h$	$>10h$	$>10h$
102		50	15126.5	$>10h$	17585.0	$>10h$	15856.9	$>10h$	$>10h$	$>10h$
103		100	17118.0	$>10h$	25825.6	$>10h$	17614.8	$>10h$	$>10h$	$>10h$
104		150	16249.5	$>10h$	20582.8	$>10h$	16861.9	$>10h$	$>10h$	$>10h$
105		200	17442.5	$>10h$	22880.6	$>10h$	18291.0	$>10h$	$>10h$	$>10h$
106		250	16167.0	$>10h$	21657.8	$>10h$	16757.5	$>10h$	$>10h$	$>10h$
107		300	16427.5	$>10h$	19808.3	$>10h$	17289.3	$>10h$	$>10h$	$>10h$
108		350	16505.0	$>10h$	21061.1	$>10h$	17336.6	$>10h$	$>10h$	$>10h$

Table 9 Solution times of different bounding methods in seconds for the capital budgeting problem (continued)

Instance	(T, I)	B	Fractional loan				Integral loan			
			ν^{κ}	ν^{Ω}	$\bar{\nu}_R^{\mathrm{NA}}$	$\bar{\nu}_R^{\mathrm{DNA}}$	ν^{κ}	ν^{Ω}	$\bar{\nu}_R^{\mathrm{NA}}$	$\bar{\nu}_R^{\mathrm{DNA}}$
109	(6,5)	0	3016.8	7432.8	5463.7	9377.8	3240.6	13927.0	6806.2	13651.4
110		50	5783.4	10153.4	7404.2	13493.4	5940.9	11026.0	9714.2	18790.5
111		100	5794.8	10968.8	8336.0	15323.8	6040.9	18911.0	11222.6	25385.0
112		150	4342.2	9224.2	6689.4	11467.2	4508.2	13912.7	8315.9	16565.2
113		200	4884.6	10710.6	7782.0	14405.6	5035.6	18792.8	10012.1	23365.9
114		250	4446.6	8944.6	6575.8	11688.6	4707.9	16686.1	7770.8	17262.4
115		300	5827.2	11153.2	8325.5	14773.2	6062.7	17358.7	11622.5	25515.4
116		350	4345.8	8927.8	6816.8	12029.8	4605.7	16729.0	9140.3	17563.5
117		400	4180.2	9836.2	7160.2	12836.2	4367.5	14595.2	8522.6	19430.3
118	(6,10)	0	10104.6	$> 10h$	10200.6	24872.6	10546.4	27686.4	13701.4	$> 10h$
119		50	8583.0	$> 10h$	9351.0	21351.0	9209.9	22440.2	12711.7	35626.3
120		100	11280.6	$> 10h$	11376.6	22130.6	11747.3	29362.6	16320.9	35307.2
121		150	9476.4	$> 10h$	9860.4	23221.4	9962.1	31165.4	12165.8	$> 10h$
122		200	9439.2	$> 10h$	9919.2	20601.2	9959.5	23400.1	13121.1	$> 10h$
123		250	8892.6	$> 10h$	9084.6	17979.6	9535.3	24697.2	10903.3	26120.8
124		300	9023.4	$> 10h$	9983.4	18604.4	9466.8	30363.2	11968.1	26395.9
125		350	11331.6	$> 10h$	11427.6	27149.6	11807.6	28773.7	13932.5	$> 10h$
126		400	11216.4	$> 10h$	12080.4	28663.4	11832.9	26791.1	14588.3	$> 10h$
127	(6,15)	0	18997.2	$> 10h$	$> 10h$	$> 10h$	19922.6	$> 10h$	$> 10h$	$> 10h$
128		50	18051.6	$> 10h$	$> 10h$	$> 10h$	19040.8	$> 10h$	$> 10h$	$> 10h$
129		100	16756.2	$> 10h$	$> 10h$	$> 10h$	17648.3	$> 10h$	$> 10h$	$> 10h$
130		150	20196.0	$> 10h$	$> 10h$	$> 10h$	21130.6	$> 10h$	$> 10h$	$> 10h$
131		200	19142.4	$> 10h$	$> 10h$	$> 10h$	20092.3	$> 10h$	$> 10h$	$> 10h$
132		250	19387.2	$> 10h$	$> 10h$	$> 10h$	20388.0	$> 10h$	$> 10h$	$> 10h$
133		300	16451.4	$> 10h$	$> 10h$	$> 10h$	17483.9	$> 10h$	$> 10h$	$> 10h$
134		350	16513.2	$> 10h$	$> 10h$	$> 10h$	17490.8	$> 10h$	$> 10h$	$> 10h$
135		400	15548.4	$> 10h$	$> 10h$	$> 10h$	16621.3	$> 10h$	$> 10h$	$> 10h$
136	(6,20)	0	26038.2	$> 10h$	$> 10h$	$> 10h$	28075.0	$> 10h$	$> 10h$	$> 10h$
137		50	25179.0	$> 10h$	$> 10h$	$> 10h$	26638.6	$> 10h$	$> 10h$	$> 10h$
138		100	27861.0	$> 10h$	$> 10h$	$> 10h$	29812.5	$> 10h$	$> 10h$	$> 10h$
139		150	27286.2	$> 10h$	$> 10h$	$> 10h$	29240.4	$> 10h$	$> 10h$	$> 10h$
140		200	26205.0	$> 10h$	$> 10h$	$> 10h$	27575.8	$> 10h$	$> 10h$	$> 10h$
141		250	26002.2	$> 10h$	$> 10h$	$> 10h$	27503.4	$> 10h$	$> 10h$	$> 10h$
142		300	28728.0	$> 10h$	$> 10h$	$> 10h$	30580.6	$> 10h$	$> 10h$	$> 10h$
143		350	24949.8	$> 10h$	$> 10h$	$> 10h$	26952.8	$> 10h$	$> 10h$	$> 10h$
144		400	29950.2	$> 10h$	$> 10h$	$> 10h$	31406.3	$> 10h$	$> 10h$	$> 10h$
145	(6,25)	0	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
146		50	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
147		100	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
148		150	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
149		200	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
150		250	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
151		300	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
152		350	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
153		400	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
154	(6,30)	0	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
155		50	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
156		100	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
157		150	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
158		200	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
159		250	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
160		300	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
161		350	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$
162		400	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$	$> 10h$