On Distributionally Robust Multistage Convex Optimization: Data-driven Models and Performance

Shixuan Zhang[1] and Xu Andy Sun[2]

[1]Wm Michael Barnes '64 Department of Industrial & Systems Engineering, Texas A&M University, shixuan.zhang@tamu.edu
[2]Operations Research and Statistics, Sloan School of Management, Massachusetts Institute of Technology, sunx@mit.edu

Abstract

This paper presents a novel algorithmic study with extensive numerical experiments of distributionally robust multistage convex optimization (DR-MCO). Following the previous work on dual dynamic programming (DDP) algorithmic framework for DR-MCO [48], we focus on data-driven DR-MCO models with Wasserstein ambiguity sets that allow probability measures with infinite supports. These data-driven Wasserstein DR-MCO models have out-of-sample performance guarantees and adjustable in-sample conservatism. Then by exploiting additional concavity or convexity in the uncertain cost functions, we design exact single stage subproblem oracle (SSSO) implementations that ensure the convergence of DDP algorithms. We test the data-driven Wasserstein DR-MCO models against multistage robust convex optimization (MRCO), risk-neutral and risk-averse multistage stochastic convex optimization (MSCO) models on multi-commodity inventory problems and hydro-thermal power planning problems. The results show that our DR-MCO models could outperform MRCO and MSCO models when the data size is small.

Keywords: distributionally robust optimization, multistage convex optimization, dual dynamic programming algorithm, multi-commodity inventory, hydro-thermal power planning problem

1 Introduction

Multistage convex optimization is a decision-making problem where objective functions and constraints are convex and decisions need to be made each time some of the uncertainty information is revealed. As it is often challenging to gain precise knowledge of the probability distributions of the uncertainty, distributionally robust multistage convex optimization (DR-MCO) allows ambiguities in probability distribution and aims to find optimal decisions that minimize the overall expected objective cost with respect to the worst-case distribution. The DR-MCO framework encompasses multistage stochastic convex optimization (MSCO) and multistage robust convex optimization (MRCO) as special cases, and thus have been widely applied in many areas including energy systems, supply chain and inventory planning, portfolio optimization, and finance [43, 6].

Distributionally robust optimization (DRO) has received significant research attention in recent years. Common choices of the ambiguity sets include moment-based ambiguity sets [40, 10, 42, 47], discrepancy or distance-based ambiguity sets [5, 4, 36, 7], and many other ones that are constructed from shape requirements or special properties of the distributions (we refer any interested reader to the review papers [37, 29] for more details). Among all these choices, data-driven Wasserstein ambiguity sets, i.e., Wasserstein distance balls centered at an empirical probability distribution obtained from given uncertainty data, has gained increasing popularity because of the following reasons: (1) measure concentration in Wasserstein distance guarantees with high probability that the decisions or policies obtained from solving the in-sample model provides an upper bound on the out-of-sample mean cost [16, 15]; and (2)

even if we consider general probability distributions in the ambiguity sets, tractable finite-dimensional reformulation or approximation can be derived using strong duality for 1-Wasserstein distance of probability measures on Euclidean spaces [15, 50, 23], or more generally for any p-Wasserstein distance of probability measures on Polish spaces [17]. We remark that these reformulations or approximations are derived for single-stage or two-stage settings. To the best of our knowledge, it remains unclear how to solve DR-MCO problems with Wasserstein ambiguity sets and infinite uncertainty sets.

There are also many works on DRO in the multistage settings. In particular, since taking expectation with respect to a worst-case probability distribution in an ambiguity set gives a coherent risk measure, multistage DRO falls into the category of risk-averse multistage stochastic optimization [43, 35], which dates back to at least [14]. When the underlying uncertainty is stagewise independent, a random nested cutting plane algorithm, which is called stochastic dual dynamic programming (DDP) and is very often used for risk-neutral multistage stochastic linear optimization (MSLO) problems [31], has been extended to risk-averse MSLO with polyhedral risk measures in [21, 45]. DDP algorithms iteratively build under-approximations of the worst-case expected cost-to-go functions in the dynamic programming recursion, which lead to policies that minimize the approximate total cost starting from each stage. To estimate the quality of the obtained policy, deterministic over-approximation is proposed in addition to the under-approximations for risk-averse MSLO problem in [34]. Alternatively, for time-consistent conditional-value-at-risk (CVaR) risk measures, a new sampling scheme based on importance sampling is proposed to achieve tighter estimation of the policy quality [25]. Further exploiting the deterministic over-approximation, a deterministic version of the DDP algorithm is proposed for risk-averse MSLO in [1]. As a variant, robust DDP is proposed for multistage robust linear optimization (MRLO) using similar deterministic over-approximation of cost-to-go functions in [18]. In [32], stochastic DDP algorithm is used to solve distributionally robust multistage linear optimization (DR-MLO) using ambiguity sets defined by a modified χ^2-distance for probability distributions supported on the historical data. In [13], stochastic DDP algorithm is used to solve DR-MLO with Wasserstein ambiguity sets on finitely supported probability distributions with asymptotic convergence and promising out-of-sample performance. We comment that the above solution approaches for risk-averse MSLO or DR-MLO rely either on sample average approximations of the risk measures or on the assumption that all uncertainties have finite support.

As DDP-type algorithms are extensively applied in solving risk-neutral and risk-averse MSLO and DR-MLO, their convergence analysis and complexity study become central questions. The finite time convergence of stochastic DDP algorithms is first proved for MSLO problems using polyhedrality of cost-to-go functions in [33, 41]. Such convergence is similarly proved for deterministic DDP algorithms for MSLO in [2], for MRLO in [18], and for DR-MLO in [1]. An asymptotic convergence of stochastic DDP algorithms is proved using monotone convergence argument in the space of convex cost-to-go function approximations for general risk-neutral multistage stochastic convex optimization (MSCO) in [19] and risk-averse MSCO in [20]. Complexity study of DDP algorithms is established using Lipschitz continuity of the under-approximations of the cost-to-go functions for risk-neutral MSCO in [26, 27], and for risk-neutral multistage mixed-integer nonlinear optimization in [49]. Our recent work [48] adopts an abstract definition of single stage subproblem oracles (SSSO) and proves SSSO-based complexity bounds, hence also the convergence, for two DDP algorithms applied to DR-MCO problems with general uncertainty supports and ambiguity sets.

Following [48], we aim to further study data-driven DR-MCO models and compare their numerical performance to other baselines models. In particular, our paper makes the following contributions to the literature.

1. We prove the out-of-sample performance guarantee using measure concentration results, adjustable in-sample conservatism assuming Lipschitz continuity of the value functions in the uncertainty variables for the data-driven DR-MCO models with Wasserstein ambiguity sets.

2. We discuss implementations of SSSO in the context of Wasserstein ambiguity sets containing infinitely supported probability distributions. Such SSSO allows DDP algorithms to converge with

provable complexity bounds to an ε-global optimal first-stage solution, which could help minimize the influence from the sub-optimality caused by early termination of the algorithms when we compare model performances.

3. We present extensive numerical experiments using multi-commodity inventory problem with either uncertain demands or uncertain prices, and hydro-thermal power planning problems with real-world data, that compares the out-of-sample performances of our DR-MCO models against risk-neutral and CVaR risk-averse MSCO models, as well as the MRCO model. To the best of our knowledge, these are the first numerical experiments in the literature that compare DR-MCO models with MSCO and MRCO models when the uncertainty has infinitely many outcomes.

The rest of the paper is organized as follows. In Section 2, we define the DR-MCO model considered in this paper and show some of its favorable properties. In Section 3, we review DDP algorithms for DR-MCO and study the implementation of SSSO for Wasserstein ambiguity sets. In Section 4, we present numerical experiment results comparing DR-MCO models against other baseline models on two application problems. We provide some concluding remarks in Section 5.

2 Data-driven Model and Its Properties

2.1 Data-driven Model Formulation

In this section, we present a data-driven model for DR-MCO and some of its properties. Let $\mathcal{T} := \{1, \ldots, T\}$ denote the set of stage indices. In each stage $t \in \mathcal{T}$, we use $\mathcal{X}_t \subset \mathbb{R}^{d_t}$ to denote the convex state space and x_t its elements, which is known as the state vector. We denote the set of uncertainties before stage t as $\Xi_t \subseteq \mathbb{R}^{\delta_t}$ and its elements as ξ_t. For simplicity, we use the notation $\mathcal{X}_0 = \{x_0\}$ and $\Xi_1 = \{\xi_1\}$ to denote parameter sets of the given initial state. After each stage $t \in \mathcal{T} \setminus \{T\}$, the uncertainty ξ_t is assumed to be distributed according to an unknown probability measure p_t taken from a subset of the probability measures $\mathcal{P}_t \subset \mathcal{M}^{\mathrm{Prob}}(\Xi_t)$ supported on Ξ_t. The cost in each stage $t \in \mathcal{T}$ is described through a nonnegative, lower semicontinuous local cost function $f_t(x_{t-1}, x_t; \xi_t)$, which is assumed to be convex in x_{t-1} and x_t for every $\xi_t \in \Xi_t$. We allow f_t to take $+\infty$ to model constraints relating the states x_{t-1}, x_t and the uncertainty ξ_t (so that $\mathcal{X}_t$ can be independent of x_{t-1} and ξ_t). The DR-MCO can be written in a nested formulation as follows.

$$\inf_{x_1 \in \mathcal{X}_1} \quad f_1(x_0, x_1; \xi_1) + \sup_{p_2 \in \mathcal{P}_2} \mathbb{E}_{\xi_2 \sim p_2} \left[\inf_{x_2 \in \mathcal{X}_2} f_2(x_1, x_2; \xi_2) + \right. \tag{1}$$

$$+ \sup_{p_3 \in \mathcal{P}_3} \mathbb{E}_{\xi_3 \sim p_3} \left[\inf_{x_3 \in \mathcal{X}_3} f_3(x_2, x_3; \xi_3) + \cdots \right.$$

$$\left. \left. + \sup_{p_T \in \mathcal{P}_T} \mathbb{E}_{\xi_T \sim p_T} \left[\inf_{x_T \in \mathcal{X}_T} f_T(x_{T-1}, x_T; \xi_T) \right] \cdots \right] \right].$$

Here, $\mathbb{E}_{\xi_t \sim p_t}$ is the expectation with respect to variable ξ_t distributed according to the probability measure p_t. We remark that both the uncertainty sets Ξ_t and the ambiguity sets $\mathcal{P}_t$ are independent between stages, which are usually referred to as *stagewise independence*.

Based on stagewise independence in the nested formulation, we can write the following recursion that is equivalent to (1) using the (worst-case expected) cost-to-go functions,

$$\mathcal{Q}_{t-1}(x_{t-1}) := \sup_{p_t \in \mathcal{P}_t} \mathbb{E}_{\xi_t \sim p_t} \left[\inf_{x_t \in \mathcal{X}_t} f_t(x_{t-1}, x_t; \xi_t) + \mathcal{Q}_t(x_t) \right], \tag{2}$$

for each $t \in \mathcal{T}$ and we set by convention $\mathcal{Q}_T(x_T) := 0$ for any $x_T \in \mathcal{X}_T$. To simplify the notation, we also define the following value function for each stage $t \in \mathcal{T}$:

$$Q_t(x_{t-1}; \xi_t) := \inf_{x_t \in \mathcal{X}_t} f_t(x_{t-1}, x_t; \xi_t) + \mathcal{Q}_t(x_t). \tag{3}$$

Using these value functions, we may write the optimal value of the DR-MCO (1) as $Q_1(x_0; \xi_1)$ and further simplify the recursion (2) as

$$\mathcal{Q}_{t-1}(x_{t-1}) = \sup_{p_t \in \mathcal{P}_t} \int_{\Xi_t} Q_t(x_{t-1}; \xi_t) \, \mathrm{d}p_t(\xi_t). \tag{4}$$

While there are many different choices of the ambiguity set $\mathcal{P}_t$ for each stage $t \in \mathcal{T}$ (see e.g., [47]), we focus on the data-driven ambiguity sets constructed as follows. Suppose we have the knowledge of n_t samples $\hat{\xi}_{t,1}, \ldots, \hat{\xi}_{t,n_t}$ of the uncertainty ξ_t. The empirical probability measure is given by $\hat{\nu}_t := \frac{1}{n_t} \sum_{k=1}^{n_t} \delta_{\hat{\xi}_{t,k}}$, where for each $k = 1, \ldots, n_t$, $\delta_{\hat{\xi}_{t,k}}$ is the Dirac probability measure supported at the point $\hat{\xi}_{t,k} \in \Xi_t$, i.e., $\int_{\Xi_t} f \, \mathrm{d}\delta_{\hat{\xi}_{t,k}} = f(\hat{\xi}_{t,k})$ for any compactly supported function f on Ξ_t. Such an empirical probability measure $\hat{\nu}_t$ captures the information from the sample data and is often used to build the sample average approximation for multistage stochastic optimization [43].

Fix any distance function $d_t(\cdot, \cdot)$ on Ξ_t, the Wasserstein distance of order 1 (a.k.a, Kantorovich-Rubinstein distance) is defined as

$$W_t(\mu, \nu) := \inf_{\pi \in \mathcal{M}^{\mathrm{Prob}}(\Xi_t \times \Xi_t)} \left\{ \int_{\Xi_t \times \Xi_t} d_t(\xi^1, \xi^2) \, \mathrm{d}\pi(\xi^1, \xi^2) : P^1_*(\pi) = \mu, \ P^2_*(\pi) = \nu \right\}, \tag{5}$$

for any two probability measures $\mu, \nu \in \mathcal{M}^{\mathrm{Prob}}(\Xi_t)$, where $P^i_*(\pi)$ is the pushforward measure induced by the projection maps $P^i : \Xi_t \times \Xi_t \to \Xi_t$ by sending $P^i(\xi^1, \xi^2) = \xi^i$, for $i = 1$ or 2. That is, the joint probability measure π in (5) has marginal probability measures equal to μ and ν.

It can be shown that W_t is indeed a distance on the space of probability measures $\mathcal{M}^{\mathrm{Prob}}(\Xi_t)$ [46, Definition 6.1] except that it may take the value of $+\infty$. Thus it is natural to restrict our attention to the convex subset of probability measures with finite distance to a Dirac measure on Ξ_t

$$\mathcal{W}_t := \left\{ \mu \in \mathcal{M}^{\mathrm{Prob}}(\Xi_t) : \int_{\Xi_t} d_t(\bar{\xi}, \xi) \, \mathrm{d}\mu(\xi) < +\infty, \text{ for some } \bar{\xi} \in \Xi_t \right\}. \tag{6}$$

Note that any continuous function $g(\xi)$ that satisfies $|g(\xi)| \le C(1 + d_t(\bar{\xi}, \xi))$ for some $C > 0$ and $\bar{\xi} \in \Xi_t$ would be integrable for any probability measure in $\mathcal{W}_t$. Now given any such continuous functions $g_{t,1}, \ldots, g_{t,m_t}$ on Ξ_t and a real vector $\rho_t := (\rho_{t,j})_{t=0}^{m_t} \in \mathbb{R}^{m_t+1}$, we define the Wasserstein ambiguity set $\mathcal{P}_t$ as

$$\mathcal{P}_t := \left\{ p \in \mathcal{W}_t : W_t(p, \hat{\nu}_t) \le \rho_{t,0}, \ \langle g_{t,j}, p \rangle \le \rho_{t,j}, \ j = 1, \ldots, m_t \right\}, \tag{7}$$

The first inequality constraint in the definition (7) bounds the Wasserstein distance of the probability measure $p \in \mathcal{W}_t$ from the empirical measure $\hat{\nu}_t$, while the second set of constraints on p are defined as $\langle g_{t,j}, p \rangle := \int_{\Xi_t} g_{t,j}(\xi) \, \mathrm{d}p(\xi)$, which could be bounds on the moments.

It is well studied that the recursion for data-driven Wasserstein DR-MCO problems (2) can be reformulated as a finite-dimensional minimum-supremum optimization problem, under the assumption of strict feasibility.

Assumption 1. *The empirical probability measure $\hat{\nu}_t$ is a strictly feasible solution of* (7), *i.e. it satisfies* $\langle g_{t,j}, \hat{\nu}_t \rangle = \frac{1}{n_t} \sum_{i=1}^{n_t} g_{t,j}(\hat{\xi}_{t,k}) < \rho_{t,j}$ *for all* $j = 1, \ldots, m_t$.

Theorem 1. *Under Assumption 1, in any stage $t \ge 2$, the expected cost-to-go function (2) can be equivalently rewritten as*

$$\mathcal{Q}_{t-1}(x_{t-1}) = \tag{8}$$

$$\min_{\lambda \ge 0} \left\{ \sum_{j=0}^{m_t} \rho_{t,j} \lambda_j + \frac{1}{n_t} \sum_{k=1}^{n_t} \sup_{\xi_k \in \Xi_t} \left\{ Q_t(x_{t-1}; \xi_k) - \lambda_0 d_{t,k}(\xi_k) - \sum_{j=1}^{m_t} \lambda_j g_{t,j}(\xi_k) \right\} \right\}.$$

We provide the derivation and proof details in Section A. As a corollary, we can prove a special version of the Kantorovich-Rubinstein duality formula [46, Remark 6.5].

Corollary 2. *Under Assumption 1, if the value function $Q_t(x_{t-1}; \xi_t)$ is l_t-Lipschitz continuous in the uncertainty $\xi_t \in \Xi_t$ for any $x_{t-1} \in \mathcal{X}_{t-1}$, then we have*

$$\mathcal{Q}_{t-1}(x_{t-1}) \leq \rho_{t,0} l_t + \frac{1}{n_t} \sum_{k=1}^{n_t} Q_t(x_{t-1}; \hat{\xi}_{t,k}).$$

Proof. Take a feasible solution $\lambda_0 = l_t$ and $\lambda_j = 0$ for $j = 1, \ldots, m_t$ in (8) of Theorem 1. Note that by the Lipschitz continuity assumption, the supremum is attained at $\hat{\xi}_{t,k}$ for each $k = 1, \ldots, n_t$. $\square$

2.2 Out-of-Sample Performance Guarantee

A major motivation for using Wasserstein DR-MCO models is the out-of-sample performance guarantee, which ensures that the decisions evaluated on the true probability distribution would perform no worse than the in-sample training with high probability. To begin with, we say that a probability measure $\mu \in \mathcal{W}_t$ is sub-Gaussian if $\int_{\Xi_t} \exp(C\xi^2) \, d\mu(\xi) < +\infty$ for some constant $C > 0$, or it has finite third moments if $\int_{\Xi_t} \|\xi\|^3 \, d\mu(\xi) < +\infty$. Our discussion is based on the following specialized version of the measure concentration inequality [16, Theorem 2].

Theorem 3. *Fix any probability measure $\nu_t \in \mathcal{M}^{\mathrm{Prob}}(\Xi_t)$ in stage t and let $\hat{\nu}_t$ denote the empirical measure constructed from n_t independent and identically distributed (iid) samples of ν_t. Then for any $\rho_{t,0} > 1$, we have*

$$\mathbb{P}\left(W_t(\nu_t, \hat{\nu}_t) > \rho_{t,0}\right) \leq \begin{cases} C_t \exp\left(-C_t' n_t \rho_{t,0}^{\max\{\delta_t, 2\}}\right), & \text{if } \nu_t \text{ is sub-Gaussian and } \delta_t \neq 2, \quad (9) \\ C_t''(n_t \rho_{t,0}^2)^{-1}, & \text{if } \nu_t \text{ has finite third moments,} \quad (10) \end{cases}$$

for some positive constants $C_t, C_t', C_t'' > 0$ that depend only on ν_t.

The measure concentration bound in (9) becomes slightly more intricate when the dimension of the uncertainty $\delta_t = 2$ (see the details in [16]), so we focus our discussion below on the other cases.

The out-of-sample performance refers to the evaluation of the solutions and policies obtained from solving DR-MCO (1) on the true probability measures ν_t for each $t \in \mathcal{T}$. To be precise, fix any optimal policy given by the DR-MCO (1), i.e., an optimal initial stage state $x_1^* \in \arg\min_{x_1 \in \mathcal{X}_1}\{f_1(x_0, x_1; \xi_1) + \mathcal{Q}_1(x_1)\}$, and a collection of mappings $x_t^* : \mathcal{X}_{t-1} \times \Xi_t \to \mathcal{X}_t$ for $t = 2, \ldots, T$, such that

1. $x_t^*(x_{t-1}, \cdot)$ is Borel measurable for any $x_{t-1} \in \mathcal{X}_{t-1}$;

2. $x_t^*(x_{t-1}, \xi_t) \in \arg\min_{x_t \in \mathcal{X}_t}\{f_t(x_{t-1}, x_t; \xi_t) + \mathcal{Q}_t(x_t)\}$ for any $x_{t-1} \in \mathcal{X}_{t-1}$ and $\xi_t \in \Xi_t$.

We define recursively the costs in the out-of-sample evaluation cost-to-go functions as

$$\mathcal{Q}_{t-1}^{\mathrm{Eval}}(x_{t-1}) := \int_{\Xi_t} \left[f_t(x_{t-1}, x_t^*(x_{t-1}, \xi_t); \xi_t) + \mathcal{Q}_t^{\mathrm{Eval}}(x_t^*(x_{t-1}, \xi_t))\right] d\nu_t(\xi_t), \quad (11)$$

with $\mathcal{Q}_T^{\mathrm{Eval}}(x_T) = 0$ for any $x_T \in \mathcal{X}_T$. Then the out-of-sample evaluation mean cost associated with this policy is defined as $v^{\mathrm{Eval}} := f_1(x_0, x_1^*; \xi_1) + \mathcal{Q}_1^{\mathrm{Eval}}(x_1^*)$. The next theorem provides a lower bound on the probability that the event $v^{\mathrm{Eval}} \leq Q_1(x_0; \xi_1)$ happens, which is often known as the out-of-sample performance guarantee.

Theorem 4. *Fix any probability measure $\nu_t \in \mathcal{W}_t$ and let $\hat{\nu}_t$ denote the empirical measure from n_t iid samples of ν_t for all stages $t \in \mathcal{T}$. Assume that $\langle g_{t,j}, \nu_t \rangle \leq \rho_{t,j}$ for $j = 1, \ldots, m_t$, and $\rho_{t,0} > 1$ for all $t \in \mathcal{T}$. Then for any $\alpha \in (0, 1)$, we have $v^{\mathrm{Eval}} \leq Q_1(x_0; \xi_1)$ with probability at least α if either of the following conditions holds for each $t \in \mathcal{T}$:*

5

1. *the probability measure ν_t is sub-Gaussian, $\delta_t \neq 2$, and*

$$n_t \cdot \rho_{t,0}^{\max\{\delta_t, 2\}} \geq \frac{1}{C_t'} \left[\ln C_t - \ln\left(1 - \alpha^{1/(T-1)}\right) \right],$$

2. *the probability measure ν_t has finite third moments and*

$$n_t \cdot \rho_{t,0}^2 \geq \frac{C_t''}{1 - \alpha^{1/(T-1)}},$$

where C_t, C_t', and C_t'' are the positive constants in Theorem 3, that depend only on ν_t.

Proof. If either of the conditions is satisfied, then it is straightforward to check from Theorem 3 that the probability $\mathbb{P}(W_t(\nu_t, \hat{\nu}_t) > \rho_{t,0}) \leq 1 - \alpha^{1/(T-1)}$. By the assumption on the iid sampling of $\hat{\nu}_t$ and that $\langle g_{t,j}, \nu_t \rangle \leq \rho_{t,j}$ for $j = 1, \ldots, m_t$, the event $E := \{\nu_t \in \mathcal{P}_t \text{ for } t = 2, \ldots, T\}$ has the probability

$$\mathbb{P}(E) = \mathbb{P}\{W_t(\nu_t, \hat{\nu}_t) \leq \rho_{t,0} \text{ for all } t \in \mathcal{T}\} = \prod_{t=2}^{T} \mathbb{P}(W_t(\nu_t, \hat{\nu}_t) \leq \rho_{t,0}) \geq \alpha.$$

To complete the proof, we claim that $\mathcal{Q}_t^{\mathrm{Eval}}(x_t) \leq \mathcal{Q}_t(x_t)$ for all $x_t \in \mathcal{X}_t$ and $t \in \mathcal{T}$ everywhere on E. Note that $\mathcal{Q}_T^{\mathrm{Eval}}(x_T) = \mathcal{Q}_T(x_T) = 0$ for all $x_T \in \mathcal{X}_T$. Now if $\mathcal{Q}_t^{\mathrm{Eval}}(x_t) \leq \mathcal{Q}_t(x_t)$ for all $x_t \in \mathcal{X}_t$, then

$$
\begin{aligned}
&f_t(x_{t-1}, x_t^*(x_{t-1}, \xi_t); \xi_t) + \mathcal{Q}_t^{\mathrm{Eval}}(x_t^*(x_{t-1}, \xi_t)) \\
&\leq f_t(x_{t-1}, x_t^*(x_{t-1}, \xi_t); \xi_t) + \mathcal{Q}_t(x_t^*(x_{t-1}, \xi_t)) \\
&= \min_{x_t \in \mathcal{X}_t} \{f_t(x_{t-1}, x_t; \xi_t) + \mathcal{Q}_t(x_t)\} \\
&= Q_t(x_{t-1}; \xi_t).
\end{aligned}
$$

Therefore, on the event E, we have

$$
\begin{aligned}
\mathcal{Q}_{t-1}^{\mathrm{Eval}}(x_{t-1}) &= \int_{\Xi_t} \left[f_t(x_{t-1}, x_t^*(x_{t-1}, \xi_t); \xi_t) + \mathcal{Q}_t^{\mathrm{Eval}}(x_t^*(x_{t-1}, \xi_t)) \right] \mathrm{d}\nu_t(\xi_t) \\
&\leq \int_{\Xi_t} Q_t(x_{t-1}; \xi_t) \, \mathrm{d}\nu_t(\xi_t) \leq \sup_{p_t \in \mathcal{P}_t} \int_{\Xi_t} Q_t(x_{t-1}; \xi_t) \, \mathrm{d}p_t(\xi_t) = \mathcal{Q}_{t-1}(x_{t-1}).
\end{aligned}
$$

This recursion shows that $\mathcal{Q}_1^{\mathrm{Eval}}(x_1) \leq \mathcal{Q}_1(x_1)$ for all $x_1 \in \mathcal{X}_1$, and thus

$$v^{\mathrm{Eval}} = f_1(x_0, x_1^*; \xi_1) + \mathcal{Q}_1^{\mathrm{Eval}}(x_1^*) \leq \min_{x_1 \in \mathcal{X}_1} \{f_1(x_0, x_1; \xi_1) + \mathcal{Q}_1(x_1)\} = Q_1(x_0; \xi_1).$$

$\square$

While Theorem 4 is a direct consequence of Theorem 3, it shows some interesting aspects of DR-MCO using Wasserstein ambiguity sets. First, to get a certain probabilistic bound for the out-of-sample performance guarantee, we may need to increase the number of samples n_t or the Wasserstein distance bound $\rho_{t,0}$ for a larger number of stages T. Second, for probability measures that are not sub-Gaussian (or more generally those with heavy tails, see [15]), for which we would like to apply the second condition, such increase require the product $n_t \cdot \rho_{t,0}^2$ to grow approximately on the order of $\mathcal{O}(T)$ when α is close to 1. Therefore, it is sometimes useful to take larger values of $\rho_{t,0}$ for out-of-sample performance guarantees, especially when the number of sample n_t is limited, or when the true probability measures ν_t are not sub-Gaussian.

2.3 Adjustable In-Sample Conservatism

Another well studied approach to guarantee out-of-sample performance is the multistage robust convex optimization (MRCO) model [18, 48]. However, MRCO considers only the worst-case outcomes of the uncertainties and thus can be overly conservative. To be precise, we define the nominal MSCO from the empirical measures $\hat{\nu}_t$ without risk aversion by the following recursion

$$\mathcal{Q}_{t-1}^{\mathrm{Nomin}}(x_{t-1}) := \int_{\Xi_t} Q_t^{\mathrm{Nomin}}(x_{t-1}; \xi_t) \, \mathrm{d}\hat{\nu}_t(\xi_t) = \frac{1}{n_t} \sum_{k=1}^{n_t} Q_t^{\mathrm{Nomin}}(x_{t-1}; \hat{\xi}_{t,k}), \tag{12}$$

where

$$Q_t^{\mathrm{Nomin}}(x_{t-1}; \xi_t) := \inf_{x_t \in \mathcal{X}_t} f_t(x_{t-1}, x_t; \xi_t) + \mathcal{Q}_t^{\mathrm{Nomin}}(x_t), \tag{13}$$

and $\mathcal{Q}_T^{\mathrm{Nomin}}(x_T) = 0$ for any $x_T \in \mathcal{X}_T$. Any MRCO that is built directly from data could have a much larger optimal cost than the nominal MSCO with a probability growing with the numbers of samples n_t and stages T, as illustrated by the following example.

Example 1. *Consider an MSCO with local cost functions $f_t(x_{t-1}, x_t; \xi_t) := x_t + \xi_t$ and state spaces $\mathcal{X}_t := [0, 1] \subseteq \mathbb{R}$ for all $t \in \mathcal{T}$. For each $t \geq 2$, the uncertainties are described by a probability measure ν_t on the set $\Xi_t := \mathbb{R}_{\geq 0}$ such that for any $C > 0$, we have $\nu_t(C, +\infty) > 0$. Now to approximate this MSCO, suppose we are given iid samples $\hat{\xi}_{t,1}, \ldots, \hat{\xi}_{t,n_t}$ from the probability measure ν_t in each stage $t \geq 2$. Then any MRCO defined by the following recursion of expected cost-to-go functions*

$$\mathcal{Q}_{t-1}^{\mathrm{Robust}}(x_{t-1}) := \sup_{\xi_t \in \hat{\Xi}_t} \min_{x_t \in \mathcal{X}_t} f_t(x_{t-1}, x_t; \xi_t) + \mathcal{Q}_t^{\mathrm{Robust}}(x_t), \quad \forall \, t \geq 2,$$

where $\hat{\Xi}_t \supseteq \{\hat{\xi}_{t,1}, \ldots, \hat{\xi}_{t,n_t}\}$ is an uncertainty subset constructed from the data, would have its optimal value $Q_1^{\mathrm{Robust}}(x_0; \xi_1) \geq x_0 + \xi_1 + \sum_{t=2}^{T} \hat{\xi}_t^{\max}$, where $\hat{\xi}_t^{\max} := \max\{\hat{\xi}_{t,1}, \ldots, \hat{\xi}_{t,n_t}\}$. The optimal value of the corresponding nominal MSCO is $Q_1^{\mathrm{Nomin}}(x_0; \xi_1) = x_0 + \xi_1 + \sum_{t=2}^{T} \hat{\xi}_t^{\mathrm{mean}}$ where $\hat{\xi}_t^{\mathrm{mean}} := \frac{1}{n_t} \sum_{k=1}^{n_t} \hat{\xi}_{t,k}$. Therefore, for any constant $C > \max_{t=2,\ldots,T} \mathbb{E}[\xi_t]$, we can show that

$$\mathbb{P}\left\{Q_1^{\mathrm{Robust}}(x_0; \xi_1) - Q_1^{\mathrm{Nomin}}(x_0; \xi_1) > C\right\} = 1 - \prod_{t=2}^{T}\left((\nu_t[0, 2C])^{n_t} + \mathbb{P}\left\{\hat{\xi}_t^{\mathrm{mean}} > C\right\}\right),$$

which goes to 1 as $n_t \to \infty$.

In contrast, the difference of the optimal values between the DR-MCO and the nominal MSCO $Q_1(x_0; \xi_1) - Q_1^{\mathrm{Nomin}}(x_0; \xi_1)$ can be bounded by the Wasserstein distances $\rho_{t,0}$, as shown below.

Theorem 5. *Suppose the value function $Q_t(x_{t-1}; \xi_t)$ is l_t-Lipschitz continuous in $\xi_t \in \Xi_t$ for any $x_{t-1} \in \mathcal{X}_{t-1}$, for each stage $t \in \mathcal{T}$. Under Assumption 1, the difference of optimal values between the DR-MCO and the nominal MSCO satisfies*

$$Q_1(x_0; \xi_1) - Q_1^{\mathrm{Nomin}}(x_0; \xi_1) \leq \sum_{t=2}^{T} l_t \rho_{t,0}.$$

Proof. We first observe that if there exists $\varepsilon \geq 0$ such that $\sup_{x_t \in \mathcal{X}_t} |\mathcal{Q}_t(x_t) - \mathcal{Q}_t^{\mathrm{Nomin}}(x_t)| \leq \varepsilon$, then the value functions satisfy $Q_t(x_{t-1}; \xi_t) - Q_t^{\mathrm{Nomin}}(x_{t-1}; \xi_t) \leq \varepsilon$ by the definitions (3) and (13). Now we prove by recursion that $\mathcal{Q}_t(x_t) - \mathcal{Q}_t^{\mathrm{Nomin}}(x_t) \leq \sum_{s>t} l_s \rho_{s,0}$ for any $x_t \in \mathcal{X}_t$, which holds trivially for $t = T$. For

any $t \in \mathcal{T}$, we have

$$
\begin{aligned}
\mathcal{Q}_{t-1}(x_{t-1}) - \mathcal{Q}_{t-1}^{\text{Nomin}}(x_{t-1}) &= \left(\sup_{p_t \in \mathcal{P}_t} \int_{\Xi_t} Q_t(x_{t-1}; \xi_t) \, \mathrm{d}p_t(\xi_t) - \int_{\Xi_t} Q_t(x_{t-1}; \xi_t) \, \mathrm{d}\hat{\nu}_t(\xi_t) \right) \\
&\quad + \int_{\Xi_t} \left(Q_t(x_{t-1}; \xi_t) - Q_t^{\text{Nomin}}(x_{t-1}; \xi_t) \right) \mathrm{d}\hat{\nu}_t(\xi_t) \\
&\leq l_t \rho_{t,0} + \sum_{s>t} l_s \rho_{s,0} = \sum_{s>t-1} l_s \rho_{s,0},
\end{aligned}
$$

where the first part before the inequality is bounded by $l_t \rho_{t,0}$ using Corollary 2, and the second part is bounded by our observation above. This recursion shows that $\mathcal{Q}_1(x_1) - \mathcal{Q}_1^{\text{Nomin}}(x_1) \leq \sum_{t=2}^{T} l_t \rho_{t,0}$, which completes the proof by the same observation. $\qquad\square$

Theorem 5 shows that the conservatism of the DR-MCO can be adjusted linearly with the Wasserstein distance bound $\rho_{t,0}$, assuming the Lipschitz continuity of the value functions in the uncertainties. However, as the out-of-sample performance guarantee in Theorem 4 depends on some unknown constants in Theorem 3, it is not easy to numerically determine the Wasserstein distance bounds $\rho_{t,0}$. We discuss some practical choices of the bounds $\rho_{t,0}$ in Section 4.

3 Dual Dynamic Programming Algorithm

In this section, we first review the recursive cutting plane approximations and the dual dynamic programming (DDP) algorithm. Then we focus on different realizations of the single stage subproblem oracles (SSSO) for DR-MCO with Wasserstein ambiguity sets that would guarantee the convergence of the DDP algorithm.

3.1 Recursive Approximations and Regularization

Recall that for any convex function $\mathcal{Q} : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, an affine function $\mathcal{V} : \mathcal{X} \to \mathbb{R}$ is called a (valid) linear cut if $\mathcal{Q}(x) \geq \mathcal{V}(x)$ for all $x \in \mathcal{X}$. A collection of such valid linear cuts $\{\mathcal{V}^j\}_{1 \leq j \leq i}$ defines a valid under-approximation $\underline{\mathcal{Q}}^i(x) := \max_{1 \leq j \leq i} \mathcal{V}^j(x)$ of $\mathcal{Q}(x)$. Similarly by convexity, given a collection of overestimate values $v^j \geq \mathcal{Q}(x^j)$ for $j = 1, \ldots, i$, we can define a valid over-approximation by the convex envelope $\overline{\mathcal{Q}}^i(x) := \text{conv}_{1 \leq j \leq i}(v^j + \iota_{x^j}(x))$, where $\iota_{x^j}(x) = 0$ when $x = x^j$ and $+\infty$ otherwise, is the convex indicator function centered at x^j. The validness of these approximations $\underline{\mathcal{Q}}(x) \leq \mathcal{Q}(x) \leq \overline{\mathcal{Q}}(x)$ for all $x \in \mathcal{X}$ suggests that we may use them in the place of $\mathcal{Q}$ for recursive updates during a stagewise decomposition algorithm.

To see how the recursive updates may work, let us assume temporarily in this subsection that the worse-case probability measure p_t^* in (4) exists and can be found. Given any under-approximation $\underline{\mathcal{Q}}_t$ for $\mathcal{Q}_t$, for any $\hat{\xi}_t \in \Xi_t$, we can generate a linear cut $V_t(x_{t-1}; \xi_t) := \underline{v}(\hat{\xi}_t) + u(\hat{\xi}_t)^{\mathsf{T}}(x_{t-1} - \tilde{x}_{t-1})$ for the value function $Q_t(x_{t-1}; \hat{\xi}_t)$ with some $\tilde{x}_{t-1} \in \mathcal{X}_{t-1}$, where $\underline{v}(\hat{\xi}_t)$ and $u(\hat{\xi}_t)$ are the optimal value and an optimal dual solution to the following Lagrangian dual of the under-approximation problem that is parametrized by $\hat{\xi}_t$:

$$
\underline{v}(\hat{\xi}_t) := \sup_{u \in \mathbb{R}^{d_{t-1}}} \min \left\{ f_t(z_t, x_t; \hat{\xi}_t) + \underline{\mathcal{Q}}_t(x_t) + u^{\mathsf{T}}(\tilde{x}_{t-1} - z_t) : x_t \in \mathcal{X}_t, z_t \in \mathbb{R}^{d_{t-1}} \right\}, \tag{14}
$$

assuming that $u(\hat{\xi}_t)$ exists. Then we can aggregate these linear cuts $V_t(x_{t-1}; \xi_t)$ into a linear cut $\mathcal{V}_{t-1}(\cdot) := \mathbb{E}_{\xi_t \sim p_t^*} V_t(\cdot; \xi_t)$ for $\mathcal{Q}_{t-1}$, where the expectation is taken componentwise with respect to the probability measure p_t^*. In this way, we can use the under-approximation $\underline{\mathcal{Q}}_t$ for $\mathcal{Q}_t$ to update the under-approximation $\underline{\mathcal{Q}}_{t-1}$ for $\mathcal{Q}_{t-1}$ by the aggregated linear cut $\mathcal{V}_{t-1}$.

Likewise, given any over-approximation $\overline{Q}_t$ for Q_t, for any $\hat{\xi}_t \in \Xi_t$ and $\tilde{x}_{t-1} \in \mathcal{X}_{t-1}$, we can solve the following over-estimation problem

$$\bar{v}(\hat{\xi}_t) := \min \left\{ f_t(\tilde{x}_{t-1}, x_t; \hat{\xi}_t) + \overline{Q}_t(x_t) : x_t \in \mathcal{X}_t \right\}, \tag{15}$$

which gives an overestimate value $\bar{v}(\hat{\xi}_t) \geq Q_t(\tilde{x}_{t-1}; \hat{\xi}_t)$. Again we can aggregate the overestimate value by setting $v_{t-1} := \mathbb{E}_{\xi_t \sim p_t^*} \bar{v}(\xi_t)$, which by definition satisfies $v_{t-1} \geq Q_{t-1}(\tilde{x}_{t-1})$ and thus can be used to update the over-approximation $\overline{Q}_{t-1}$ for Q_{t-1}.

There are, however, some potential issues with this recursive approximation method. First, the supremum in the Lagrangian dual problem (14) may not be attained, which could happen if $\underline{v}(\hat{\xi}_t) = +\infty$ (i.e., $\tilde{x}_{t-1}$ is an infeasible state for $\hat{\xi}_t$) or if any neighborhood of $\tilde{x}_{t-1}$ contains such an infeasible state. In this case, we may fail to generate a linear cut $\mathcal{V}_t(\cdot; \hat{\xi}_t)$. Second, Lipschitz constants of the linear cuts $\mathcal{V}_{t-1}$ may be affected by the under-approximation $\underline{Q}_t$, causing a worse approximation quality. This may happens when $\tilde{x}_{t-1}$ is an extreme point of $\mathcal{X}_{t-1}$. In fact, it is shown in [48] that the Lipschitz constants of $\underline{Q}_t$ could exceed those of Q_t and grow with the total number of stages T. Third, the over-approximation function evaluates to $+\infty$ at any point that is not in the convex hull of previously visited points, which makes the gap $\overline{Q}_t(x) - \underline{Q}_t(x)$ less useful as an estimate of the quality of the current solutions or policies.

To remedy these issues, we consider a technique called *Lipschitzian regularization*. Given regularization factors $M_t > 0$, we define the regularized local cost function as

$$f_t^R(x_{t-1}, x_t; \xi_t) := \inf_{z_t \in \mathbb{R}^{d_{t-1}}} f_t(z_t, x_t; \xi_t) + M_t \| x_{t-1} - z_t \|, \tag{16}$$

and the regularized value function

$$Q_t^R(x_{t-1}; \xi_t) := \min_{x_t \in \mathcal{X}_t} f_t^R(x_{t-1}, x_t; \xi_t) + \mathcal{Q}_t^R(x_t), \tag{17}$$

recursively for $t = T, T-1, \ldots, 2$, where $\mathcal{Q}_t^R$ is the regularized expected cost-to-go function defined as

$$\mathcal{Q}_t^R(x_t) := \sup_{p_{t+1} \in \mathcal{P}_{t+1}} \mathbb{E}_{\xi_{t+1} \sim p_{t+1}} Q_{t+1}^R(x_t; \xi_{t+1}), \tag{18}$$

for $t \leq T-1$, and $\mathcal{Q}_T^R(x_T) \equiv 0$ for any $x_T \in \mathcal{X}_T$. It is then straightforward to check that $Q_t^R(x_{t-1}; \xi_t)$ is uniformly M_t-Lipschitz continuous in x_{t-1} for each $\xi_t \in \Xi_t$, and consequently $\mathcal{Q}_{t-1}^R$ is also M_t-Lipschitz continuous. In the definitions (14) and (15), we can replace accordingly the original cost functions f_t with the regularized cost function f_t^R, which would guarantee the M_t-Lipschitz continuity of the generated cuts and allow us to enhance the over-approximation to be M_t-Lipschitz continuous.

Lipschitzian regularization in general only gives under-approximations of the true value and expected cost-to-go functions. We need the following assumption to preserve the optimality and feasibility of the solutions.

Assumption 2. *For the given regularization factors $M_t > 0$, $t \in \mathcal{T}$, the optimal value of the regularized DR-MCO satisfies*

$$\min_{x_1 \in \mathcal{X}_1} f_1(x_0, x_1; \xi_1) + \mathcal{Q}_1(x_1) = \min_{x_1 \in \mathcal{X}_1} f_1(x_0, x_1; \xi_1) + \mathcal{Q}_1^R(x_1)$$

and the sets of optimal first-stage solutions are the same, i.e., $\arg\min\{f_1(x_0, x_1; \xi_1) + \mathcal{Q}_1^R(x_1) : x_1 \in \mathcal{X}_1\} = \arg\min\{f_1(x_0, x_1; \xi_1) + \mathcal{Q}_1(x_1) : x_1 \in \mathcal{X}_1\}$.

We remark by the following proposition that Assumption 2 can be satisfied in any problem that already have uniformly Lipschitz continuous value function $Q_t(\cdot; \xi_t)$ for all $\xi_t \in \Xi_t$.

Proposition 6 (Proposition 4,[48]). *Suppose each state space $\mathcal{X}_t \subseteq \mathbb{R}^{d_t}$ is full dimensional, i.e., $\mathrm{int}\mathcal{X}_t \neq \varnothing$. Then Assumption 2 holds if for each stage $t \geq 2$, the value function $Q_t(\cdot; \xi_t)$ is M_t-Lipschitz continuous for any $\xi_t \in \Xi_t$.*

In general, we can execute the DDP algorithm even with $M_t = +\infty$, as described below.

3.2 Dual Dynamic Programming Algorithm

DDP algorithms generally refer to the recursive cutting plane algorithms that exploit the stagewise independence structure. We first review the definitions of single stage subproblem oracles (SSSO), which symbolize the subroutines of subproblem solving in each stage with under- and over-approximations of the expected cost-to-go functions [48].

Definition 1 (Initial stage subproblem oracle). *Let $\underline{\mathcal{Q}}_1, \overline{\mathcal{Q}}_1 : \mathcal{X}_1 \to \bar{\mathbb{R}}$ denote two lsc convex functions, representing an under-approximation and an over-approximation of the cost-to-go function $\mathcal{Q}_1^{\mathrm{R}}$ in (18), respectively. Consider the following subproblem for the first stage $t = 1$,*

$$\min_{x_1 \in \mathcal{X}_1} f_1(x_0, x_1; \xi_1) + \underline{\mathcal{Q}}_1(x_1). \tag{I}$$

The initial stage subproblem oracle provides an optimal solution x_1 to (I) and calculates the approximation gap $\gamma_1 := \overline{\mathcal{Q}}_1(x_1) - \underline{\mathcal{Q}}_1(x_1)$ at the solution. We thus define the subproblem oracle formally as a map $\mathscr{O}_1 : (\underline{\mathcal{Q}}_1, \overline{\mathcal{Q}}_1) \mapsto (x_1; \gamma_1)$.

Definition 2 (Noninitial stage subproblem oracle). *Fix any regularization parameter $M_t > 0$. Let $\underline{\mathcal{Q}}_t, \overline{\mathcal{Q}}_t : \mathcal{X}_t \to \bar{\mathbb{R}}$ denote two lsc convex functions, representing an under-approximation and an over-approximation of the cost-to-go function $\mathcal{Q}_t^{\mathrm{R}}$ in (18), respectively, for some stage $t > 1$. Then given a feasible state $x_{t-1} \in \mathcal{X}_{t-1}$, the noninitial stage subproblem oracle provides a feasible state $x_t \in \mathcal{X}_t$, an M_t-Lipschitz continuous linear cut $\mathcal{V}_{t-1}(\cdot)$, and an over-estimate value v_{t-1} such that*
- *they are valid, i.e., $\mathcal{V}_{t-1}(x) \le \mathcal{Q}_{t-1}^{\mathrm{R}}(x)$ for any $x \in \mathcal{X}_{t-1}$ and $v_{t-1} \ge \mathcal{Q}_{t-1}^{\mathrm{R}}(x_{t-1})$;*
- *the gap is controlled, i.e., $v_{t-1} - \mathcal{V}_{t-1}(x_{t-1}) \le \gamma_t := \overline{\mathcal{Q}}_t(x_t) - \underline{\mathcal{Q}}_t(x_t)$.*

We thus define the subproblem oracle formally as a map $\mathscr{O}_t : (x_{t-1}, \underline{\mathcal{Q}}_t, \overline{\mathcal{Q}}_t) \mapsto (\mathcal{V}_{t-1}, v_{t-1}, x_t; \gamma_t)$.

Now we can present the (consecutive) DDP algorithm based on the SSSO. In Algorithm 1, each iteration $i \in \mathbb{N}$ consists of two steps: the noninitial stage step and the initial stage step. In the noninitial stage step, we evaluate the SSSO from stage $t = 2$ to $t = T$ to collect feasible states x_t^i, overestimate values v_{t-1}^i, and a valid linear cut $\mathcal{V}_{t-1}^i$ for updating the approximations $\underline{\mathcal{Q}}_{t-1}^i$ and $\overline{\mathcal{Q}}_{t-1}^i$. Then we evaluate the initial stage SSSO to get an optimal solution x_1^{i+1} and its optimality gap γ_1^{i+1}. We terminate the algorithm when the gap is sufficiently small.

Algorithm 1 is proved to terminate with an ε-optimal solution in finitely many iterations [48]. We include the theorem here for completeness.

Theorem 7. *Suppose that all the state spaces $\mathcal{X}_t$ have the dimensions bounded by d and diameters by D, and let $M := \max\{M_t : t = 1, \dots, T - 1\}$. If for each stage $t \in \mathcal{T}$, the local cost functions are strictly positive $f_t(x_{t-1}, x_t; \xi_t) \ge C$ for all feasible solutions $x_t \in \mathcal{X}_t$, uncertainties $\xi_t \in \Xi_t$ and some constant $C > 0$, then the total number of noninitial stage subproblem oracle evaluations before achieving an α-relative optimal solution x_1^* for Algorithm 1 is bounded by*

$$\texttt{\#Eval} \le 1 + T(T-1)\left(1 + \frac{2MD}{\alpha C}\right)^d.$$

We remark that the DDP algorithm can also be executed in a nonconsecutive way with a similar complexity bound. However, since we need to run the algorithm for a fixed number of iterations to compare different models, we restrict our attention to the consecutive version here. Any interested reader is referred to [48] for more detailed discussion.

In Section 3.1, we have seen how SSSO could be implemented if the worst-case probability measure exists in (4) and can be found. However, it is well known that the worst-case probability measure may not exist (see e.g., [15, Example 2]). Moreover, the integration with respect to some worst-case probability measure could also be numerically challenging. Therefore, we next provide two possible SSSO implementation methods directly using the recursion (8).

Algorithm 1 Dual Dynamic Programming Algorithm

Require: subproblem oracles $\mathcal{O}_t$ for $t \in \mathcal{T}$, optimality gap $\varepsilon > 0$

Ensure: an ε-optimal first stage solution x_1^* to the regularization (18)

1: initialize: $\underline{\mathcal{Q}}_t^0 \leftarrow 0, \overline{\mathcal{Q}}_t^0 \leftarrow +\infty, t \in \mathcal{T} \backslash \{T\}; \underline{\mathcal{Q}}_T^j, \overline{\mathcal{Q}}_T^j \leftarrow 0, j \in \mathbb{N}; i \leftarrow 1$

2: evaluate $(x_1^1; \gamma_1^1) \leftarrow \mathcal{O}_1(\underline{\mathcal{Q}}_1^0, \overline{\mathcal{Q}}_1^0)$

3: set LOWERBOUND $\leftarrow f_1(x_0, x_1^1; \xi_1)$, UPPERBOUND $\leftarrow +\infty$

4: **while** UPPERBOUND $-$ LOWERBOUND $> \varepsilon$ **do**

5:     **for** $t = 2, \ldots, T$ **do**

6:         evaluate $(\mathcal{V}_{t-1}^i, v_{t-1}^i, x_t^i; \gamma_t^i) = \mathcal{O}_t(x_{t-1}^i, \underline{\mathcal{Q}}_t^{i-1}, \overline{\mathcal{Q}}_t^{i-1})$ ▷ noninitial stage step

7:         update $\underline{\mathcal{Q}}_{t-1}^i(x) \leftarrow \max\{\underline{\mathcal{Q}}_{t-1}^{i-1}(x), \mathcal{V}_{t-1}^i(x)\}$

8:         update $\overline{\mathcal{Q}}_{t-1}^i(x) \leftarrow \mathrm{conv}\{\overline{\mathcal{Q}}_{t-1}^{i-1}(x), v_{t-1}^i + M_{t-1}\|x - x_{t-1}^i\|\}$

9:     evaluate $(x_1^{i+1}; \gamma_1^{i+1}) \leftarrow \mathcal{O}_1(\underline{\mathcal{Q}}_1^i, \overline{\mathcal{Q}}_1^i)$ ▷ initial stage step

10:     update LOWERBOUND $\leftarrow f_1(x_0, x_1^{i+1}; \xi_1) + \underline{\mathcal{Q}}_1^i(x_1^{i+1})$

11:     update UPPERBOUND$' \leftarrow f_1(x_0, x_1^{i+1}; \xi_1) + \overline{\mathcal{Q}}_1^i(x_1^{i+1})$

12:     **if** UPPERBOUND$'$ $<$ UPPERBOUND **then**

13:         set $x_1^* \leftarrow x_1^{i+1}$, UPPERBOUND $\leftarrow$ UPPERBOUND$'$

14:     update $i \leftarrow i + 1$

3.3 Subproblem Oracles: Concave Uncertain Cost Functions

If the following assumption holds, we are able to reformulate the recursion (8) into a convex optimization problem.

Assumption 3. *The local cost function $f_t(x_{t-1}, x_t; \xi_t)$ is concave and upper semicontinuous in the uncertainty ξ_t for any $x_{t-1} \in \mathcal{X}_{t-1}$ and $x_t \in \mathcal{X}_t$.*

A direct consequence of Assumption 3 is that the effective domain of the state x_t does not depend on the uncertainty ξ_t, as shown in the following lemma.

Lemma 8. *Under Assumption 3, we have $\mathrm{dom}\, f_t(x_{t-1}, \cdot; \xi_t) = \mathrm{dom}\, f_t(x_{t-1}, \cdot; \xi_t')$ for any $x_{t-1} \in \mathcal{X}_{t-1}$ and $\xi_t, \xi_t' \in \Xi_t$.*

Proof. Assume for contradiction that there exists some $x_{t-1} \in \mathcal{X}_{t-1}$, $x_t \in \mathcal{X}_t$, and $\xi_t, \xi_t' \in \Xi_t$ such that $f_t(x_{t-1}, x_t; \xi_t) < +\infty$ but $f_t(x_{t-1}, x_t; \xi_t') = +\infty$. Then for $c \in (0, 1)$, we have $f_t(x_{t-1}, x_t; (1-c)\xi_t + c\xi_t') = +\infty$ by the concavity and nonnegativity of f_t. It follows from upper semicontinuity that $f_t(x_{t-1}, x_t; \xi_t) \geq \limsup_{c \to 0+} f_t(x_{t-1}, x_t; (1-c)\xi_t + c\xi_t') = +\infty$, which is a contradiction. $\square$

We are now ready to prove the alternative formulation of the recursion (8).

Theorem 9. *Under Assumption 3, if we further assume that the continuous functions $g_{t,j}$ are convex for $j = 1, \ldots, m$ and $d_{t,k}(\xi_{t,k}) = \|\xi_{t,k} - \hat{\xi}_{t,k}\|$ for $k = 1, \ldots, n_t$, then we have*

$$\mathcal{Q}_{t-1}(x_{t-1}) = \min \quad \sum_{j=0}^{m_t} \rho_{t,j}\lambda_{t,j} + \frac{1}{n_t}\sum_{k=1}^{n_t} \left[h_{t,k}(z_t, x_{t,k}, \zeta_{t,k}, \lambda_t) + \mathcal{Q}_t(x_{t,k}) \right]$$

$$\text{s.t.} \quad \left\| \zeta_{t,k} \right\|_* \leq \lambda_{t,0}, \tag{19}$$

$$z_t = x_{t-1},$$

$$\lambda_t \in \mathbb{R}_{\geq 0}^{m_t+1}, x_{t,k} \in \mathcal{X}_t,$$

where for each $k = 1, \ldots, n_t$, $h_{t,k}$ is defined as

$$h_{t,k}(x_{t-1}, x_{t,k}, \zeta_{t,k}, \lambda_t) := \sup_{\xi_{t,k} \in \Xi_t} f_t(x_{t-1}, x_{t,k}; \xi_{t,k}) - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\xi_{t,k}) + \zeta_{t,k}^\mathsf{T}(\xi_{t,k} - \hat{\xi}_{t,k}).$$

Proof. By Lemma 8, for any $x_{t-1} \in \mathcal{X}_{t-1}$, we can define a set $\mathcal{X}_t(x_{t-1}) := \mathrm{dom}\, f_t(x_{t-1}, \cdot; \xi_t) \subseteq \mathcal{X}_t$ which is independent of $\xi_t \in \Xi_t$ and closed by the lower semicontinuity of f_t. Note that the norm function has the dual representation $d_{t,k}(\xi_{t,k}) = \|\xi_{t,k} - \hat{\xi}_{t,k}\| = \max_{\|\zeta\|_* \leq 1} \zeta^\mathsf{T}(\xi_{t,k} - \hat{\xi}_{t,k})$. Thus by the recursion (8), we can write

$$\mathcal{Q}_{t-1}(x_{t-1}) = \min_{\lambda_t \in \mathbb{R}^{m_t+1}_{\geq 0}} \sum_{j=0}^{m_t} \rho_{t,j} \lambda_{t,j} + \frac{1}{n_t} \sum_{k=1}^{n_t} \sup_{\xi_{t,k} \in \Xi_t} \left[-\sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\xi_{t,k}) \right.$$

$$\left. + \min_{x_{t,k}, \zeta_{t,k}} \quad f_t(x_{t-1}, x_{t,k}; \xi_{t,k}) + \mathcal{Q}_t(x_{t,k}) + \zeta_{t,k}^\mathsf{T}(\xi_{t,k} - \hat{\xi}_{t,k}) \right]$$

$$\text{s.t.} \quad \|\zeta_{t,k}\|_* \leq \lambda_{t,0},$$

$$x_{t,k} \in \mathcal{X}_t(x_{t-1}).$$

Now for any fixed x_{t-1} and λ_t, we see that the sets $\{\zeta_{t,k} : \|\zeta_{t,k}\|_* \leq \lambda_{t,0}\}$ and $\mathcal{X}_t(x_{t-1})$ are compact. Moreover, the function inside the supremum of $\xi_{t,k}$ is concave and upper semicontinuous in $\xi_{t,k}$, while convex and lower semicontinuous in $\xi_{t,k}$ and $\zeta_{t,k}$. Thus the result follows by applying Sion's minimax theorem [24]. $\qquad \square$

Remark. The proof remains valid if we replace simultaneously $\mathcal{Q}_t$, $\mathcal{Q}_{t-1}$, and f_t with $\mathcal{Q}_t^\mathrm{R}$, $\mathcal{Q}_{t-1}^\mathrm{R}$, and f_t^R in the theorem. In this case we use $h_{t,k}^\mathrm{R}$ to denote the convex conjugate functions.

We provide a possible implementation for noninitial stage SSSO in Algorithm 2 based on Theorem 9. Its correctness is verified by the following corollary.

Algorithm 2 Single Stage Subproblem Oracle Implementation Under Assumption 3

Require: function h_t, over- and under-approximations $\overline{\mathcal{Q}}_t^i$ and $\underline{\mathcal{Q}}_t^i$, and a state $x_{t-1} \in \mathcal{X}_{t-1}$
Ensure: a linear cut $\mathcal{V}_{t-1}$, an overestimate v_{t-1}, a state x_t, and a gap value γ_t
1: Solve the minimization (19) with $\mathcal{Q}_t$ replaced by $\underline{\mathcal{Q}}_t^i$ and store the optimal value v_{t-1}^*, optimal solutions λ_t^* and $(x_{t,k}^*, \zeta_{t,k}^*)_{k=1}^{n_t}$ and the dual solutions u_t associated with the constraints $z_t = x_{t-1}$
2: **for** $k = 1, \ldots, n_t$ **do**
3: $\quad$ Compute the gap value $\gamma_{t,k} := \overline{\mathcal{Q}}_t^i(x_{t,k}^*) - \underline{\mathcal{Q}}_t^i(x_{t,k}^*)$
4: Set $\mathcal{V}_{t-1}(\cdot) \leftarrow v_{t-1}^* + u_t^\mathsf{T}(\cdot)$
5: Set $v_{t-1} \leftarrow v_{t-1}^* + \frac{1}{n_t} \sum_{k=1}^{n_t} \gamma_{t,k}$
6: Take any $k^* \in \arg\max\{\gamma_{t,k} : k = 1, \ldots, n_t\}$ and set $x_t \leftarrow x_{t,k^*}^*$, $\gamma_t \leftarrow \gamma_{t,k^*}$

Corollary 10. *Under the same assumptions of Theorem 9, the outputs $(\mathcal{V}_{t-1}, v_{t-1}, x_t; \gamma_t)$ of Algorithm 2 satisfy the conditions in Definition 2.*

Proof. To check the validness of v_{t-1}, let $(z_t^*, x_{t,k}^*, \zeta_{t,k}^*, \lambda_t^*)$ denote an optimal solution in the minimization (19) with $\mathcal{Q}_t$ replaced by $\underline{\mathcal{Q}}_t^i$. Then we have

$$v_{t-1} = v_{t-1}^* + \frac{1}{n_t} \sum_{k=1}^{n_t} \gamma_{t,k}$$

$$= \sum_{j=0}^{m_t} \rho_{t,j} \lambda_{t,j}^* + \frac{1}{n_t} \sum_{k=1}^{n_t} [h_{t,k}(z_t^*, x_{t,k}^*, \zeta_{t,k}^*, \lambda_t^*) + \underline{\mathcal{Q}}_t^i(x_{t,k}^*) + \gamma_{t,k}]$$

$$\geq \sum_{j=0}^{m_t} \rho_{t,j} \lambda_{t,j}^* + \frac{1}{n_t} \sum_{k=1}^{n_t} [h_{t,k}(z_t^*, x_{t,k}^*, \zeta_{t,k}^*, \lambda_t^*) + \mathcal{Q}_t(x_{t,k}^*)] \geq \mathcal{Q}_{t-1}(x_{t-1}),$$

the last inequality is due to the feasibility of $(\mathcal{V}_{t-1}, v_{t-1}, x_t; \gamma_t)$ in the minimization (19). For the validness of $\mathcal{V}_{t-1}(\cdot)$, note that the value v_{t-1}^* and the dual solution u_t define a valid linear under-approximation for the function $\mathcal{Q}'_{t-1}(\cdot)$ defined by replacing $\mathcal{Q}_t$ with $\underline{\mathcal{Q}}_t^i$ in the minimization (19). Since clearly $\mathcal{Q}'_{t-1}(x_{t-1}) \leq \mathcal{Q}_{t-1}(x_{t-1})$ for all $x_{t-1} \in \mathcal{X}_{t-1}$, we see that $\mathcal{V}_{t-1}(\cdot)$ is a valid under-approximation for $\mathcal{Q}_{t-1}(\cdot)$. Finally the gap $v_{t-1} - \mathcal{V}_{t-1}(x_{t-1}) = \frac{1}{n_t} \sum_{k=1}^{n_t} \gamma_{t,k} \leq \gamma_t$ is controlled. $\qquad\square$

Theorem 9 and Algorithm 2 would be most useful when the functions $h_{t,k}$ can be written explicitly as minimization problems. We thus spend the rest of this section to derive the form of $h_{t,k}$ in a special yet practically important case, where the local cost function f_t can be written as

$$
f_t(x_{t-1}, x_t; \xi_t) = \min \quad (A_t \xi_t + a_t)^\mathsf{T} y_t \tag{20}
$$
$$
\text{s.t.} \quad (x_{t-1}, y_t, x_t) \in \mathcal{F}_t,
$$

for some compact convex set $\mathcal{F}_t \subseteq \mathcal{X}_{t-1} \times \mathbb{R}^{d'_t} \times \mathcal{X}_t$ in each stage $t \in \mathcal{T}$. It is straightforward to check that f_t in (20) is lower semicontinuous and convex in (x_{t-1}, x_t) for any $\xi_t \in \Xi_t$. To simplify our discussion, we assume that $f_t(\cdot, x_t; \xi_t)$ is M_t-Lipschitz continuous, so by Proposition 6 $\mathcal{Q}_t = \mathcal{Q}_t^\mathrm{R}$ for all $t \in \mathcal{T}$. The problem (20) is a common formulation in the usual MSCO literature, such as [41], where $\mathcal{F}_t$ is supposed to be a polytope.

Proposition 11. *Suppose the local cost function $f_t(x_{t-1}, x_t; \xi_t)$ is given in the form* (20) *and uniformly M_t-Lipschitz continuous in the variable x_{t-1}. Fix any point $\bar{\xi}_t \in \mathrm{int}\Xi_t$ and let $\sigma_t(\zeta) := \sup_{\xi \in \Xi_t} \zeta^\mathsf{T}(\xi - \bar{\xi}_t)$ denote the support function of the set $\Xi_t - \bar{\xi}_t$. If the functions $g_{t,j}(\xi_t) = \xi_t^\mathsf{T} B_{t,j} \xi_t + b_{t,j}^\mathsf{T} \xi_t$ are quadratic with coefficients $B_{t,j} \in \mathcal{S}_{\succeq 0}^{\delta_t}$ and $b_{t,j} \in \mathbb{R}^{\delta_t}$ for $j = 1, \ldots, m_t$, then we can write*

$$
h_{t,k}(x_{t-1}, x_{t,k}, \zeta_{t,k}, \lambda_t) = \min_{y_{t,k}, w_{t,j}, w'_{t,j}, \kappa_j} \quad (A_t \bar{\xi}_t + a_t)^\mathsf{T} y_{t,k} - \sum_{j=1}^{m_t} \lambda_{t,j} [\bar{\xi}_t^\mathsf{T} B_{t,j} \bar{\xi}_t + b_{t,j}^\mathsf{T} \bar{\xi}_t] \tag{21}
$$
$$
+ \zeta_{t,k}^\mathsf{T}(\bar{\xi}_t - \hat{\xi}_{t,k}) + \sum_{j=1}^{m_t} \kappa_{t,j} + \sigma_t(w_{t,0})
$$
$$
\text{s.t.} \quad \sum_{j=0}^{m_t} w_{t,j} = \zeta_{t,k} + A_t^\mathsf{T} y_{t,k} - \sum_{j=1}^{m_t} \lambda_{t,j} [2 B_{t,j} \bar{\xi}_t + b_{t,j}],
$$
$$
\kappa_{t,j} \geq 0, \quad j = 1, \ldots, m,
$$
$$
w_{t,j} \in \mathbb{R}^{\delta_t}, \quad j = 0, \ldots, m,
$$
$$
\kappa_{t,j} + \lambda_{t,j} \geq \|(\kappa_{t,j} - \lambda_{t,j}, U_{t,j} w_{t,j})\|_2, \quad j = 1, \ldots, m_t,
$$
$$
w_{t,j} = B_{t,j} w'_{t,j}, \ w'_{t,j} \in \mathbb{R}^{\delta_t}, \quad j = 1, \ldots, m,
$$
$$
(x_{t-1}, y_{t,k}, x_{t,k}) \in \mathcal{F}_t.
$$

Here, $U_{t,j}$ is a $\delta_t \times \delta_t$ real matrix such that $U_{t,j}^\mathsf{T} U_{t,j}$ is the pseudoinverse of $B_{t,j}$.

Proof. Under the assumptions, we can write the function $h_{t,k}$ as

$$
h_{t,k}(x_{t-1}, x_{t,k}, \zeta_{t,k}, \lambda_t) = \sup_{\xi_t \in \Xi_t} \min_{y_{t,k}} \quad (A_t \xi_t + a_t)^\mathsf{T} y_{t,k} + \zeta_{t,k}^\mathsf{T}(\xi_t - \hat{\xi}_{t,k}) - \sum_{j=1}^{m_t} \lambda_{t,j} (\xi_t^\mathsf{T} B_{t,j} \xi_t + b_{t,j}^\mathsf{T} \xi_t)
$$
$$
\text{s.t.} \quad (x_{t-1}, y_{t,k}, x_{t,k}) \in \mathcal{F}_t.
$$

Note that the objective function in (21) is continuous in both y_t and $\xi_{t,k}$, and the projection of $\mathcal{F}_t$ onto the variables $y_{t,k}$ is compact. Thus by the minimax theorem [24], we can exchange the supremum and

13

minimum operations

$$h_{t,k}(x_{t-1}, x_{t,k}, \zeta_{t,k}, \lambda_t) = \min_{y_{t,k}} \quad (A_t \bar{\xi}_t + a_t)^\mathsf{T} y_{t,k} - \sum_{j=1}^{m_t} \lambda_{t,j} [\bar{\xi}_t^\mathsf{T} B_{t,j} \bar{\xi}_t + b_{t,j}^\mathsf{T} \bar{\xi}_t] + \zeta_{t,k}^\mathsf{T} (\bar{\xi}_t - \hat{\xi}_{t,k})$$

$$+ \sup_{\xi_t \in \mathbb{R}^{\delta_t}} \left\{ \zeta_{t,k}^\mathsf{T} \xi_t - \iota_t(\xi_t) + y_{t,k}^\mathsf{T} A_t \xi_t - \sum_{j=1}^{m_t} \lambda_{t,j} (\xi_t^\mathsf{T} B_{t,j} \xi_t + 2\bar{\xi}_t B_{t,j} \xi_t + b_{t,j}^\mathsf{T} \xi_t) \right\} \quad (22)$$

$$\text{s.t.} \quad (x_{t-1}, y_{t,k}, x_{t,k}) \in \mathcal{F}_t,$$

where ι_t is the convex indicator function of the set $\Xi_t - \bar{\xi}_t$, the convex conjugate of which is the support function σ_t by definition. If we further denote $\varphi_{t,j}(\xi_t; \lambda_{t,j}) := \lambda_{t,j}(\xi_t^\mathsf{T} B_{t,j} \xi_t)$, the supremum can be written using convex conjugacy as

$$\left(\iota_t + \sum_{j=1}^{m_t} \varphi_{t,j}(\cdot; \lambda_{t,j}) \right)^* \left(\zeta_{t,k} + A_t^\mathsf{T} y_{t,k} - \sum_{j=1}^{m_t} \lambda_{t,j} [2B_{t,j} \bar{\xi}_t + b_{t,j}] \right).$$

Note that for each $j = 1, \ldots, m_t$, the parametrized conjugate function $\varphi_{t,j}^*(\cdot; \lambda_{t,j})$ can be written as [39, Example 11.10]

$$\varphi_{t,j}^*(w; \lambda_{t,j}) = \begin{cases} \dfrac{w^\mathsf{T} B_{t,j}^\dagger w}{4\lambda_{t,j}}, & \text{if } w \in \text{range} B_{t,j}, \\ +\infty, & \text{otherwise}, \end{cases}$$

$$= \min \left\{ \kappa_{t,j} \geq 0 : 4\kappa_{t,j} \lambda_{t,j} \geq (U_{t,j} w)^\mathsf{T} (U_{t,j} w), w = B_{t,j} w' \right\}$$

$$= \min \left\{ \kappa_{t,j} \geq 0 : \kappa_{t,j} + \lambda_{t,j} \geq \|(\kappa_{t,j} - \lambda_{t,j}, U_{t,j} w)\|_2, w = B_{t,j} w' \right\},$$

which is nonnegative and second-order conic representable. Here the convention for $\lambda_{t,j} = 0$ is consistent: we have $\varphi_{t,j}^*(0; 0) = 0$ and $\varphi_{t,j}^*(w; 0) = +\infty$ for any $w \neq 0$ because $(U_{t,j} B_{t,j})^\mathsf{T} (U_{t,j} B_{t,j}) = B_{t,j} B_{t,j}^\dagger B_{t,j} = B_{t,j}$, which implies that $U_{t,j} w = U_{t,j} B_{t,j} w' \neq 0$. Now using the formula for convex conjugate of sum of convex functions, we have

$$\left(\iota_t + \sum_{j=1}^{m_t} \varphi_{t,j}(\cdot; \lambda_{t,j}) \right)^* = \text{cl} \left(\sigma_t \square \varphi_{t,1}^*(\cdot; \lambda_{t,1}) \square \cdots \square \varphi_{t,m_t}^*(\cdot; \lambda_{t,m_t}) \right), \quad (23)$$

where $\square$ denotes the infimal convolution (a.k.a. epi-addition) of two convex functions and cl denotes the lower semicontinuous hull of a proper function. Since $\bar{\xi}_t \in \text{int} \Xi_t$, the support function is coercive, i.e., $\lim_{\|w\| \to \infty} \sigma_t(w) = +\infty$. Moreover, each $\varphi_{t,j}^*$ is bounded below as it is nonnegative. Therefore, the closure operation is superficial and the convex conjugate of the sum is indeed lower semicontinuous [3, Proposition 12.14]. The rest of the proof follows from substitution of this convex conjugate expression (23) into the supremum in (22). □

3.4   Subproblem Oracles: Convex Uncertain Cost Functions

We provide another useful reformulation of the recursion (8) in this section, based on the following assumption.

Assumption 4. *The local cost function $f_t(x_{t-1}, x_t; \xi_t)$ is jointly convex in the state variable x_t and the uncertainty ξ_t, for any $x_{t-1} \in \mathcal{X}_{t-1}$. Moreover, the uncertainty set Ξ_t is a pointed polyhedron and the distance function $d_{t,k}(\cdot)$ is polyhedrally representable.*

14

We first mention some direct consequences of Assumption 4. First, the value function $Q_t(x_{t-1}; \xi_t)$ would be a convex function in the uncertainty ξ_t for each state $x_{t-1} \in \mathcal{X}_{t-1}$, although it may not be a jointly convex function. Second, recall that a polyhedron is pointed if it does not contain any lines. Any point in a pointed polyhedron can be written as a convex combination of its extreme points and rays [9, Theorem 3.37]. Now under Assumption 4, we may define a lifted uncertainty set as $\tilde{\Xi}_{t,k} := \{(\zeta, \xi) : \xi \in \Xi_t, \zeta \geq d_{t,k}(\xi)\}$. It is easy to see that $\tilde{\Xi}_{t,k}$ is also a pointed polyhedron. We denote the finite set of extreme points of it as $\text{ext} \, \tilde{\Xi}_{t,k} = \{(\tilde{\zeta}_l, \tilde{\xi}_l)\}_{l \in E_{t,k}}$ where $E_{t,k}$ is the set of indices, and want to show that the maximization in (8) can be taken over the finite set $\{(\tilde{\zeta}_l, \tilde{\xi}_l)\}_{l \in E_{t,k}}$ in two important cases. First, we consider problems with bounded uncertainty sets Ξ_t.

Proposition 12. *Under Assumption 4, if we further assume that Ξ_t is bounded and all functions $g_{t,j}$ are concave for $j = 1, \ldots, m_t$, then the problem (8) can be equivalently reformulated as*

$$\mathcal{Q}_{t-1}(x_{t-1}) = \min_{\lambda_t, \tau_t} \quad \sum_{j=0}^{m_t} \rho_{t,j} \lambda_{t,j} + \frac{1}{n_t} \sum_{k=1}^{n_t} \tau_{t,k} \tag{24}$$

$$\text{s.t.} \quad \lambda_t \geq 0,$$

$$\tau_{t,k} \geq Q_t(x_{t-1}; \tilde{\xi}_l) - \lambda_{t,0} \tilde{\zeta}_l - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\tilde{\xi}_l), \quad \text{for all } l \in E_{t,k} \text{ and } k = 1, \ldots, n_t.$$

Proof. From the definition of lifted uncertainty set $\tilde{\Xi}_t$, we have

$$\sup_{\xi_k \in \Xi_t} \left\{ Q_t(x_{t-1}; \xi_k) - \lambda_{t,0} d_{t,k}(\xi_k) - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\xi_k) \right\}$$

$$= \max_{\xi_k \in \Xi_t, \zeta_k \in \mathbb{R}} \left\{ Q_t(x_{t-1}; \xi_k) - \lambda_{t,0} \zeta_k - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\xi_k) : \zeta_k \geq d_{t,k}(\xi_k) \right\}$$

$$= \max_{(\zeta_k, \xi_k) \in \tilde{\Xi}_{t,k}} \left\{ Q_t(x_{t-1}; \xi_k) - \lambda_{t,0} \zeta_k - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\xi_k) \right\}$$

$$= \max_{l \in E_{t,k}} \left\{ Q_t(x_{t-1}; \tilde{\xi}_l) - \lambda_{t,0} \tilde{\zeta}_l - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\tilde{\xi}_l) \right\}.$$

To see the last equality, note that if Ξ_t is bounded, then the only recession direction of the lifted uncertainty set $\tilde{\Xi}_{t,k}$ is $(1, 0)$. Since $\lambda_{t,0} \geq 0$, any maximum solution (ζ_k^*, ξ_k^*) lies in the convex hull of $\text{ext} \, \tilde{\Xi}_{t,k}$. Now the last equality follows from the convexity of the function $Q_t(x_{t-1}; \xi_k) - \lambda_{t,0} \zeta_k - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\xi_k)$ in terms of ξ_k and ζ_k. Finally, the reformulation is done by replacing the maximum of finitely many functions by its epigraphical representation $\tau_{t,k} \geq Q_t(x_{t-1}; \xi_l) - \lambda_{t,0} \zeta_l - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\xi_l)$ for all $l \in E_{t,k}$ and $k = 1, \ldots, n_t$. $\square$

If the uncertainty sets Ξ_t are unbounded, then in general the supremum in (8) can take $+\infty$ in some unbounded directions of Ξ_t, even when the value function $Q_t(x_{t-1}; \cdot)$ has finite values everywhere. To avoid such situation, we consider the growth rate of the value function $Q_t(x_{t-1}; \cdot)$ defined as

$$r_t(x_{t-1}) := \limsup_{\substack{d_{t,k}(\xi_t) \to \infty, \\ \xi_t \in \Xi_t}} \frac{Q_t(x_{t-1}; \xi_t) - Q_t(x_{t-1}; \hat{\xi}_{t,k})}{d_{t,k}(\xi_t)} \geq 0, \tag{25}$$

for any real-valued $Q_t(x_{t-1}, \cdot)$, where the limit superior is in fact independent of the choice of $k = 1, \ldots, n_t$, and the inequality is due to that $Q_t(x_{t-1}; \cdot)$ is assumed to be lower bounded by 0. Our convention is to set $r_t(x_{t-1}) \equiv 0$ when Ξ_t is bounded. We now consider problems with unbounded uncertainty sets Ξ_t.

15

Proposition 13. *Under Assumption [4], if $Q_t(x_{t-1}; \cdot)$ has finite growth rate $r_t(x_{t-1})$ and all functions $g_{t,j}$ are bounded and concave for $j = 1, \ldots, m_t$, then the problem [8] with any $x_{t-1} \in \mathcal{X}_{t-1}$ such that $\mathcal{Q}_{t-1}(x_{t-1}) < +\infty$ can be equivalently reformulated as*

$$\mathcal{Q}_{t-1}(x_{t-1}) = \min_{\lambda_t, \tau_t} \quad \sum_{j=0}^{m_t} \rho_{t,j} \lambda_{t,j} + \frac{1}{n_t} \sum_{k=1}^{n_t} \tau_{t,k} \tag{26}$$

$$\text{s.t.} \quad \lambda_t \geq 0,$$

$$\lambda_{t,0} \geq r_t(x_{t-1}),$$

$$\tau_{t,k} \geq Q_t(x_{t-1}; \tilde{\xi}_l) - \lambda_{t,0} \tilde{\zeta}_l - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\tilde{\xi}_l), \quad \text{for all } l \in E_{t,k} \text{ and } k = 1, \ldots, n_t.$$

Proof. We claim that the supremum

$$\sup_{\xi_k \in \Xi_t} \left\{ Q_t(x_{t-1}; \xi_k) - \lambda_{t,0} d_{t,k}(\xi_k) - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\xi_k) \right\} < +\infty$$

if and only if $\lambda_{t,0} \geq r_t(x_{t-1})$, for each $k = 1, \ldots, n_t$. Suppose $\lambda_{t,0} < r_t(x_{t-1})$. By definition [25], there exists a sequence $\{\xi_k^{(i)}\}_{i \in \mathbb{N}} \subseteq \Xi_t$ and a constant $\varepsilon > 0$ such that $d_{t,k}(\xi_k^{(i)}) \to \infty$ as $i \to \infty$ and $Q_t(x_{t-1}; \xi_k^{(i)}) \geq Q_t(x_{t-1}; \hat{\xi}_{t,k}) + (\lambda_{t,0} + \varepsilon) d_{t,k}(\xi_k^{(i)})$. Thus $\sup_{i \in \mathbb{N}} \{Q_t(x_{t-1}; \xi_k^{(i)}) - \lambda_{t,0} d_{t,k}(\xi_k^{(i)}) - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\xi_k^{(i)})\} \geq \sup_{i \in \mathbb{N}} \{\varepsilon d_{t,k}(\xi_k^{(i)}) - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\xi_k^{(i)})\} = +\infty$ as $g_{t,j}(\xi_k^{(i)})$ for $j = 1, \ldots, m_t$ are bounded.

Conversely, by definition [25], there exists a constant $\bar{d} > d_{t,k}(\hat{\xi}_{t,k})$ such that $Q_t(x_{t-1}; \xi_k) \leq Q_t(x_{t-1}; \hat{\xi}_{t,k}) + \lambda_{t,0} d_{t,k}(\xi_k)$ for all $\xi_k \in \Xi_t$ with $d_{t,k}(\xi_k) \geq \bar{d}$. Thus we have

$$\sup_{\xi_k \in \Xi_t} \left\{ Q_t(x_{t-1}; \xi_k) - \lambda_{t,0} d_{t,k}(\xi_k) - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\xi_k) \right\}$$

$$\leq \sup_{d_{t,k}(\xi_k) \leq \bar{d}} \left\{ Q_t(x_{t-1}; \xi_k) - \lambda_{t,0} d_{t,k}(\xi_k) \right\} + \sup_{\xi_k \in \Xi_t} \sum_{j=1}^{m_t} \left(-\lambda_{t,j} g_{t,j}(\xi_k) \right)$$

$$= \max_{(\zeta_k, \xi_k) \in \tilde{\Xi}_t(\bar{d})} \left\{ Q_t(x_{t-1}; \xi_k) - \lambda_{t,0} \zeta_k \right\} + \sup_{\xi_k \in \Xi_t} \sum_{j=1}^{m_t} \left(-\lambda_{t,j} g_{t,j}(\xi_k) \right) < +\infty,$$

where $\tilde{\Xi}_t(\bar{d}) := \{(\zeta, \xi) : \xi \in \Xi_t, d_{t,k}(\xi) \leq \bar{d}, \zeta \geq d_{t,k}(\xi)\}$, and the maximum is finite because it is attained on some extreme point $(\bar{\zeta}_k, \bar{\xi}_k) \in \tilde{\Xi}_t(\bar{d})$ by convexity, so $Q_t(x_{t-1}; \bar{\xi}_k) - \lambda_{t,0} \bar{\zeta}_k < +\infty$.

Now from this claim, we see that for any $x_{t-1} \in \mathcal{X}_{t-1}$ such that $\mathcal{Q}_{t-1}(x_{t-1}) < +\infty$, the problem [8] can be formulated equivalently as

$$\mathcal{Q}_{t-1}(x_{t-1}) = \min_{\lambda_t \geq 0} \quad \frac{1}{n_t} \sum_{k=1}^{n_t} \sup_{(\zeta_k, \xi_k) \in \Xi_t} \left\{ Q_t(x_{t-1}; \xi_k) - \lambda_{t,0} \zeta_k - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\xi_k) \right\} + \sum_{j=0}^{m_t} \rho_{t,j} \lambda_{t,j}$$

$$\text{s.t.} \quad \lambda_{t,0} \geq r_t(x_{t-1}).$$

The supremum can be attained in $\tilde{\Xi}'_{t,k} := \text{conv}(\text{ext}\,\tilde{\Xi}_{t,k})$: otherwise there exists a point $(\check{\zeta}_k, \check{\xi}_k) \in \tilde{\Xi}_{t,k} \setminus \tilde{\Xi}'_{t,k}$ and $(\bar{\zeta}_k, \bar{\xi}_k) \in \tilde{\Xi}'_{t,k}$ such that

$$Q_t(x_{t-1}; \check{\xi}_k) - \lambda_{t,0} \check{\zeta}_k - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\check{\xi}_k) > Q_t(x_{t-1}; \bar{\xi}_k) - \lambda_{t,0} \bar{\zeta}_k - \sum_{j=1}^{m_t} \lambda_{t,j} g_{t,j}(\bar{\xi}_k).$$

16

In other words, $(\check{\zeta}_k, \check{\xi}_k) - (\bar{\zeta}_k, \bar{\xi}_k)$ defines a strictly increasing ray of $\tilde{\Xi}_t$, which by convexity implies that the supremum is $+\infty$, a contradiction. Using the convexity again as in the proof of Proposition 12, we conclude that the supremum is indeed attained in $\text{ext}\,\tilde{\Xi}_{t,k}$, and this completes the proof. $\qquad\square$

Proposition 13 reduces to Proposition 12 since the growth rate $r_t(x_{t-1}) = 0$ and any continuous function $g_{t,j}$ over a bounded polyhedron is bounded. The finite growth rate condition is often satisfied, especially when the value function $Q_t(x_{t-1}; \cdot)$ is Lipschitz continuous. However, it is in general difficult to estimate the growth rate (25). Fortunately, in some application problem, we can derive explicitly the growth rate, as discussed in Section 4.3.

Note that the problems (24) and (26) are standard linear optimization problems in the variables λ_t and τ_t. Thus by strong duality, we can write the dual problem as

$$
\mathcal{Q}_{t-1}(x_{t-1}) = \max_{\theta_t, \kappa_{t,k,l} \geq 0} \quad \theta_t r_t(x_{t-1}) + \sum_{k=1}^{n_t} \sum_{l \in E_{t,k}} \kappa_{t,k,l} Q_t(x_{t-1}; \tilde{\xi}_l)
$$

$$
\text{s.t.} \quad \sum_{l \in E_{t,k}} \kappa_{t,k,l} = \frac{1}{n_t}, \quad k = 1, \ldots, n_t,
$$

$$
\theta_t + \sum_{k=1}^{n_t} \sum_{l \in E_{t,k}} \tilde{\zeta}_l \kappa_{t,k,l} \leq \rho_{t,0}, \tag{27}
$$

$$
\sum_{k=1}^{n_t} \sum_{l \in E_{t,k}} g_{t,j}(\tilde{\xi}_l) \kappa_{t,k,l} \leq \rho_{t,j}, \quad j = 1, \ldots, m_t.
$$

Consequently, any feasible dual solutions θ_t and $\kappa_{t,k,l}$ to the dual (27) define a valid under-approximation

$$
\mathcal{Q}_{t-1}(x_{t-1}) \geq \theta_t r_t(x_{t-1}) + \sum_{k=1}^{n_t} \sum_{l \in E_{t,k}} \kappa_{t,k,l} Q_t(x_{t-1}; \tilde{\xi}_l), \quad \forall\, x_{t-1} \in \mathcal{X}_{t-1}. \tag{28}
$$

We now describe an SSSO implementation in Algorithm 3. Its correctness is verified in the following corollary.

Corollary 14. *Suppose that the growth rate function $r_t(\cdot)$ is convex. Under the assumptions of of Proposition 13, the outputs $(\mathcal{V}_{t-1}, v_{t-1}, x_t; \gamma_t)$ of Algorithm 3 satisfy the conditions in Definition 2.*

Proof. The validness of $\mathcal{V}_{t-1}(\cdot)$ follows directly from the inequality (28) and the fact that $\underline{Q}_t(x_{t-1}; \xi_k) \leq Q_t(x_{t-1}; \xi_k)$ for any $x_{t-1} \in \mathcal{X}_{t-1}$ and $\xi_k \in \Xi_t$ by definition. To see the validness of v_{t-1}, note that for any $k = 1, \ldots, n_t$ and $l \in E_{t,k}$, we have

$$
\begin{aligned}
Q_t(x_{t-1}; \tilde{\xi}_l) &\leq \min_{x_t \in \mathcal{X}_t} \left[f_t(x_{t-1}, x_t; \tilde{\xi}_l) + \overline{\mathcal{Q}}_t^i(x_t) \right] \\
&\leq f_t(x_{t-1}, x_{t,k,l}; \tilde{\xi}_l) + \overline{\mathcal{Q}}_t^i(x_{t,k,l}) \\
&\leq f_t(x_{t-1}, x_{t,k,l}; \tilde{\xi}_l) + \underline{\mathcal{Q}}_t^i(x_{t,k,l}) + \gamma_{t,k,l} \\
&\leq \underline{Q}_t(x_{t-1}; \tilde{\xi}_l) + \gamma_{t,k}
\end{aligned}
$$

by the definition of $\gamma_{t,k}$ in Algorithm 3. Thus for any optimal solution λ_t^* to the problem (26) with $Q_t(x_{t-1}; \tilde{\xi}_l)$ replaced by $\underline{Q}_t(x_{t-1}; \tilde{\xi}_l)$, we have

$$
\max_{l \in E_{t,k}} Q_t(x_{t-1}; \tilde{\xi}_l) - \lambda_{t,0}^* \tilde{\zeta}_l - \sum_{j=1}^{m_t} \lambda_{t,j}^* g_{t,j}(\tilde{\xi}_l) \leq \max_{l \in E_{t,k}} \underline{Q}_t(x_{t-1}; \tilde{\xi}_l) - \lambda_{t,0}^* \tilde{\zeta}_l - \sum_{j=1}^{m_t} \lambda_{t,j}^* g_{t,j}(\tilde{\xi}_l) + \gamma_{t,k},
$$

and consequently $v_{t-1} \geq \mathcal{Q}_{t-1}(x_{t-1})$ since λ_t^* is also a feasible solution to the minimization in (26). Finally, the gap is controlled since $v_{t-1} - \mathcal{V}_{t-1}(x_{t-1}) = \frac{1}{n_t} \sum_{k=1}^{n_t} \gamma_{t,k} \leq \gamma_{t,k^*} = \overline{\mathcal{Q}}_t^i(x_t) - \underline{\mathcal{Q}}_t^i(x_t)$. $\qquad\square$

17

Algorithm 3 Single Stage Subproblem Oracle Implementation Under Assumption 4

Require: over- and under-approximations $\overline{\mathcal{Q}}_t^i$ and $\underline{\mathcal{Q}}_t^i$, a state $x_{t-1} \in \mathcal{X}_{t-1}$, growth rate $r_t(x_{t-1})$, and extreme point sets $\text{ext}\,\tilde{\Xi}_{t,k}$ for $k = 1, \ldots, n_t$

Ensure: a linear cut $\mathcal{V}_{t-1}$, an overestimate v_{t-1}, a state x_t, and a gap value γ_t

1: **for** $k = 1, \ldots, n_t$ **do**
2:     **for** $l \in E_{t,k}$ **do**
3:         Evaluate the approximate value function

$$\underline{Q}_t(x_{t-1}; \tilde{\xi}_l) := \min_{x_t \in \mathcal{X}_t} f_t(x_{t-1}, x_t; \tilde{\xi}_l) + \underline{\mathcal{Q}}_t^i(x_t)$$

    with a minimizer stored as $x_{t,k,l}$ and a subgradient vector as $u_{t,k,l} \in \partial \underline{Q}_t(\cdot; \tilde{\xi}_l)$ at x_{t-1}
4:         Calculate $\gamma_{t,k,l} := \overline{\mathcal{Q}}_t^i(x_{t,k,l}) - \underline{\mathcal{Q}}_t^i(x_{t,k,l})$
5: Solve the problem (26) (or (24) if Ξ_t is bounded) with $Q_t(x_{t-1}; \tilde{\xi}_l)$ replaced by $\underline{Q}_t(x_{t-1}; \tilde{\xi}_l)$ and store the optimal value v_{t-1}^* and dual solutions θ_t^*, $\kappa_{t,k,l}^*$ to (27)
6: **for** $k = 1, \ldots, n_t$ **do**
7:     Take any $l^* \in \arg\max \left\{ \overline{\mathcal{Q}}_t^i(x_{t,k,l}) - \underline{\mathcal{Q}}_t^i(x_{t,k,l}) : l \in E_{t,k} \right\}$
8:     Set $\gamma_{t,k} \leftarrow \gamma_{t,k,l^*}$ and $x_{t,k} \leftarrow x_{t,k,l^*}$
9: Take a subgradient $w_t \in \partial r_t(\cdot)$ at x_{t-1}
10: Set $\mathcal{V}_{t-1}(\cdot) \leftarrow v_{t-1}^* + \theta_t^* w_t^\mathsf{T}(\cdot - x_{t-1}) + \sum_{k=1}^{n_t} \sum_{l \in E_t} \kappa_{t,k,l}^* u_{t,k,l}^\mathsf{T}(\cdot - x_{t-1})$
11: Set $v_{t-1} \leftarrow v_{t-1}^* + \frac{1}{n_t} \sum_{k=1}^{n_t} \gamma_{t,k}$
12: Take any $k^* \in \arg\max\{\gamma_{t,k} : k = 1, \ldots, n_t\}$ and set $x_t \leftarrow x_{t,k^*}$, $\gamma_t \leftarrow \gamma_{t,k^*}$

4 Numerical Experiments

In this section, we first introduce baseline models used for comparison against the DR-MCO model (1). Then we present comprehensive numerical studies of two application problems: the multi-commodity inventory problem with either uncertain demands or uncertain prices, and the hydro-thermal power system planning problem with uncertain water inflows.

4.1 Baseline Models and Experiment Settings

For performance comparison, we introduce two types of baseline models in addition to the DR-MCO with Wasserstein ambiguity sets (7). The first baseline model is the simple multistage robust convex optimization (MRCO) model, where we simply consider the worst-case outcome out of the uncertainty set Ξ_t in each stage t. Namely, the cost-to-go functions of the MRCO can be defined recursively as

$$\mathcal{Q}_{t-1}^{\text{Robust}}(x_{t-1}) := \sup_{\xi_t \in \Xi_t} \min_{x_t \in \mathcal{X}_t} f_t(x_{t-1}, x_t; \xi_t) + \mathcal{Q}_t^{\text{Robust}}(x_t), \quad t = T, T-1, \ldots, 2. \tag{29}$$

When the sum $f_t(x_{t-1}, x_t; \xi_t) + \mathcal{Q}_t^{\text{Robust}}(x_t)$ is jointly convex in the state x_t and the uncertainty ξ_t for any given x_{t-1}, then the supremum can be attained at some extreme point of the convex hull of Ξ_t if it is finite. In particular, if we have relatively complete recourse, (i.e., the sum is always finite for any given x_{t-1}), and if Ξ_t is a polytope, (i.e., it is a convex hull of finitely many points), then we can enumerate over the extreme points of Ξ_t to find the supremum, which allows us to solve the simple MRCO by Algorithm 1. In general, if the uncertainty set Ξ_t is unbounded, then the cost-to-go functions of the MRCO model can take $+\infty$ everywhere, so we will only use the baseline MRCO model when we have polytope uncertainty sets Ξ_t.

    The second type of baseline models consists of risk-neutral and risk-averse multistage stochastic convex optimization (MSCO) models. The nominal probability measures in the MSCO models can be either the

empirical measure $\hat{\nu}_t$, or a probability measure associated with the sample average approximation (SAA) of a fitted probability measure, which we denote as $\tilde{\nu}_t = \frac{1}{n'_t} \sum_{k=1}^{n'_t} \delta_{\tilde{\xi}_{t,k}}$. The main difference here is that the outcomes in an SAA probability measure $\tilde{\xi}_{t,1}, \ldots, \tilde{\xi}_{t,n'_t}$ can be different from those given by the empirical measure $\hat{\xi}_{t,1}, \ldots, \hat{\xi}_{t,n_t}$. Moreover, we are able to take $n'_t > n_t$ for a potentially better training effect. To ease the notation, we also allow $\tilde{\nu}_t = \hat{\nu}_t$ to happen when we describe the risk measures in the rest of this section.

For the risk-averse MSCO models, we use the risk measure that is called conditional value-at-risk (CVaR, a.k.a. average value-at-risk or expected shortfall). Its coherence leads to a dual representation [34], that allows the risk-averse MSCO models solved by Algorithm 1 with a straightforward implementation of SSSO. For simplicity, we only introduce the CVaR risk-averse MSCO based on this dual representation, and any interested reader is referred to [43] for the primal definition and the proof of duality.

Given parameters $\alpha \in (0,1)$ and $\beta \in [0,1]$, we define the cost-to-go functions associated with the (α, β)-CVaR risk measures recursively for $t = T, T-1, \ldots, 2$ as

$$\mathcal{Q}_{t-1}^{\mathrm{CVaR}}(x_{t-1}) := \max_{p_t \in \mathcal{P}_t^{\mathrm{CVaR}}} \sum_{k=1}^{n'_t} p_{t,k} \left\{ \min_{x_t \in \mathcal{X}_t} f_t(x_{t-1}, x_t; \tilde{\xi}_{t,k}) + \mathcal{Q}_t^{\mathrm{CVaR}}(x_t) \right\}, \tag{30}$$

where the ambiguity set is defined as

$$\mathcal{P}_t^{\mathrm{CVaR}} := \left\{ p_t = (p_{t,1}, \ldots, p_{t,n'_t}) \in \mathbb{R}^{n'_t} : \begin{array}{l} 0 \leq p_{t,k} \leq \dfrac{\beta}{n_t} + \dfrac{1-\beta}{\alpha n_t}, \\ k = 1, \ldots, n'_t; \displaystyle\sum_{k=1}^{n'_t} p_{t,k} = 1, \end{array} \right\}. \tag{31}$$

Note that when $\beta = 1$, the ambiguity set $\mathcal{P}_t^{\mathrm{CVaR}} = \{(\frac{1}{n'_t}, \ldots, \frac{1}{n'_t})\}$ has only one element corresponding to the SAA probability measure $\tilde{\nu}_t$. Thus the CVaR risk-averse MSCO model (30) reduces to the risk-neutral nominal MSCO in this case. Alternatively, if $\beta = 0$ and $\alpha \leq \frac{1}{n_t}$, then the CVaR risk-averse MSCO model 30 considers only the worst outcome of the SAA probability measure in each stage. We remark that both the simple MRCO model (29) and the CVaR risk-averse MSCO model (30) can be solved by Algorithm 1 since only finitely many outcomes need to be considered in each stage t. More details on the SSSO for these two baseline models can be found in [48].

Our numerical experiments aim to demonstrate two attractive aspects of the DR-MCO models on some application problems: better out-of-sample performance compared to the baseline models, and ability to achieve out-of-sample performance guarantee with reasonable conservatism. For ease of evaluation, we assume that we have the knowledge of the true underlying probability measure $\nu \in \mathcal{M}(\Xi_2 \times \cdots \times \Xi_T)$, and thus the marginal probability measures $\nu_t := P_*^t(\nu)$, where P_*^t is the pushforward of the canonical projection map $P^t : \Xi_2 \times \cdots \Xi_T \to \Xi_t$, for $t = 2, \ldots, T$. Here, we do not restrict our attention to the case that ν is a product of $\nu_2, \ldots, \nu_T$, i.e., the true probability measure is stagewise independent, so our modeling (2) can be used as approximation for problems under general stochastic processes. The experiments are then carried out in the following procedures with a uniform number of data points $n_t = n$ in each stage.

1. Draw n iid samples from ν to form the empirical probability measures $\hat{\nu}_t$;
2. Construct the baseline models and DR-MCO models using $\hat{\nu}_t$;
3. Solve these models using our DDP algorithm (Algorithm 1) to a desired accuracy or within the maximum number of iterations or computation time;
4. Draw N iid sample paths from ν and evaluate the performance profiles (mean, variance, and quantiles) of the models on these sample paths.

In particular, we focus on limited or moderate training sample sizes $n \in \{5, 10, 20, 40\}$, while keeping our sizes of evaluation sample paths to be relatively large ($N = 100,000$). In each independent test run of

our numerical experiment, the training samples used in a smaller-sized test are kept in larger-sized tests, and the evaluation sample paths are held unchanged for all models and sample sizes.

Our algorithms and numerical examples are implemented using Julia 1.6 [8], with Gurobi 9.0 [22] interfaced through the JuMP package (version 0.23) [12]. We use 25 single-core 2.1GHz Intel Xeon processors (24 for the worker processes and 1 for the manager process) with 50 GByte of RAM to allow parallelization of the SSSO (Algorithm 3).

4.2 Multi-commodity Inventory Problems

We consider a multi-commodity inventory problem which is adapted from the ones studied in [18, 48]. Let $\mathcal{J} := \{1, 2, \ldots, J\}$ denote the set of product indices. We first describe the variables in each stage $t \in \mathcal{T}$. We use $x_{t,j}^l$ to denote the variable of inventory level, $y_{t,j}^a$ (resp. $x_{t,j}^b$) to denote the amount of express (resp. standard) order fulfilled in the current (resp. subsequent) stage, and $y_{t,j}^r$ to denote the amount of rejected order of each product $j \in \mathcal{J}$. Let $x_t := (x_{t,1}^l, \ldots, x_{t,K}^l, x_{t,1}^b, \ldots, x_{t,K}^b)$ be the state variable and $y_t := (y_{t,1}^a, \ldots, y_{t,K}^a, y_{t,1}^r, \ldots, y_{t,K}^r)$ be the internal variable for each stage $t \in \mathcal{T}$. The stage t subproblem can be defined through the local cost functions f_t as

$$f_t(x_{t-1}, x_t; \xi_t) := \min_{y_t} \quad C^F + \sum_{j \in \mathcal{J}} \left(C_{t,j}^a y_{t,j}^a + C_{t,j}^b x_{t,j}^b + C_j^r y_{t,j}^r + C_j^H [x_{t,j}^l]_+ + C_j^B [x_{t,j}^l]_- \right) \tag{32}$$

$$\text{s.t.} \quad \sum_{j \in \mathcal{J}} y_{t,j}^a \leq B^c,$$

$$x_{t,j}^l \leq x_{t-1,j}^l + y_{t,j}^a + x_{t-1,j}^b + y_{t,j}^r - D_{t,j}, \forall j \in \mathcal{J},$$

$$y_{t,j}^a \in [0, B_j^a], \ y_{t,j}^r \in [0, D_{t,j}], \ \forall j \in \mathcal{J},$$

$$x_{t,j}^b \in [0, B_j^b], \ x_{t,j}^l \in [B_j^{l,-}, B_j^{l,+}], \ \forall j \in \mathcal{J}.$$

In the definition (32), we use $C_{t,j}^a = C_{t,j}^a(\xi_t)$ (resp. $C_{t,j}^b = C_{t,j}^b(\xi_t)$) to denote the uncertain express (resp. standard) order unit cost, C_j^H (resp. C_j^B) the inventory holding (resp. backlogging) unit cost, C_j^r the penalty on order rejections, $C^F \equiv 1$ a positive fixed cost, B_j^a (resp. B_j^b) the bound for the express (resp. standard) order, and $B_j^{l,-}, B_j^{l,+}$ the bounds on the backlogging and inventory levels, $D_{t,j} = D_{t,j}(\xi_t)$ the uncertain demand for the product j, respectively. The first constraint in (32) is a cumulative bound B^c on the express orders, the second constraint characterizes the change in the inventory level, and the rest are bounds on the decision variables with respect to each product. The notations $[x]_+ := \max\{x, 0\}$ and $[x]_- := -\min\{0, x\}$ are used to denote the positive and negative parts of a real number x. The initial state is given by $x_{0,j}^b = x_{0,j}^l = 0$ for all $j \in \mathcal{J}$. Before we discuss the details of the uncertain parameters $C_{t,j}^a, C_{t,j}^b$ or $D_{t,j}$, we make the following remarks on the definition (32).

First, it is easy to check that if we $C_{t,j}^a, C_{t,j}^b$ (resp. $D_{t,j}$) are deterministic, then Assumption 4 (resp. Assumption 3) is satisfied so we are able to apply the SSSO implementations discussed in Sections 3.4 and 3.3. Second, as the bounds $B_j^{l,-}, B_j^{l,+}$ do not change with t and all orders can be rejected (i.e., $y_{t,j}^r = D_{t,j}$ is feasible for all $j \in \mathcal{J}$), we see that the problem (32) has relatively complete recourse. Third, the Lipschitz constant of the value functions $Q_t(\cdot; \xi_t)$ is uniformly bounded by $\sum_{j \in \mathcal{J}} C_j^r$, so Proposition 6 can be applied here if we set the regularization factors to be sufficiently large $M_t \geq \sum_{j \in \mathcal{J}} C_j^r$. Besides, the Lipschitz continuity guarantees the in-sample adjustable conservatism by Theorem 5. Last, since all state variables are bounded, together with the above observation, we know by Theorem 7 that Algorithm 1 would always converge on our inventory problem (32).

4.2.1 Inventory Problems with Uncertain Demands

First, we consider the inventory problems with uncertain demands, where the goal is to seek a policy with minimum mean inventory cost plus the penalty on order rejections. The uncertain demands are modeled

by the following expression:

$$D_{t,j}(\xi_t) := D_0 \left[1 + \cos\left(\frac{2\pi(t+j)}{\tau} \right) \right] + \bar{D} \cdot \xi_{t,j}, \quad j \in \mathcal{J}, t \in \mathcal{T}. \tag{33}$$

Here, D_0 is a factor and τ is the period for the base demands, and $\bar{D}$ is the bound on the uncertain demands. The uncertain vector $\xi_t \in [0,1]^J$ has its components described as follows: $\xi_{t,1} \sim \text{Uniform}(0,1)$, and for $j = 2, \ldots, J$, we have

$$\xi_{t,j} \mid \xi_{t,j-1} \sim \begin{cases} \text{Uniform}(0, (1 + \xi_{t,j-1})/2), & \text{if } \xi_{t,j-1} \leq \frac{1}{2}, \\ \text{Uniform}(\xi_{t,j-1}/2, 1), & \text{otherwise.} \end{cases} \tag{34}$$

For the experiments, we consider $J = 3$ products and $T = \tau = 5$ stages. The unit prices of each product are deterministically set to $C_{t,j}^a = 5$ and $C_{t,j}^b = 1$ for all $t \in \mathcal{T}$; the inventory and holding costs are $C_j^H = 2$ and $C_j^B = 10$, and the rejection costs are $C_j^r = 100$, for each $j \in \mathcal{J}$. The bounds are set to $B^c = 15$, $B_j^a = 10$, $B_j^b = 20$, $B_j^{l,-} = 10$, and $B_j^{l,+} = 100$ for each $j \in \mathcal{J}$. We pick the uncertainty parameters $D_0 = 5$ and $\bar{D} = 50$. We terminate the DDP algorithm if it reaches 1% relative optimality or 2000 iterations. For Wasserstein ambiguity sets, we only consider the radius constraint (i.e., $m_t = 0$) with radius set to be relative to the following estimation of the distance among data points:

$$\hat{d}_t := \max_{k=1,\ldots,n_t} W_t(\hat{\nu}_t, \delta_{\hat{\xi}_{t,k}}) = \max_{k=1,\ldots,n_t} \frac{1}{n_t} \sum_{k' \neq k} \|\hat{\xi}_{t,k} - \hat{\xi}_{t,k'}\|. \tag{35}$$

For the MSCO models, we directly use the empirical probability measures $\hat{\nu}_t$ for each $t \in \mathcal{T}$. Further, we consider parameters $\alpha \in \{0.01, 0.05, 0.10\}$ and $\beta \in \{0.0, 0.25, 0.50, 0.75\}$ for the CVaR risk-averse MSCO models.

Using the experiment procedure described in Section 4.1, we present the results of our data-driven DR-MCO model with Wasserstein ambiguity sets and the baseline models.

Figure 1 (and Figures 5 and 6 in Section B) displays the out-of-sample cost quantiles of the nominal stochastic model and the DR-MCO models with different Wasserstein radii, constructed from the empirical probability measures $\hat{\nu}_t$. Here we use the log radius $-\infty$ to denote the nominal stochastic model, i.e., $\rho_{t,0} = 0$. From the plot, we see that in small-sample case ($n_t = 5$), the Wasserstein DR-MCO model significantly reduces the top 10% out-of-sample evaluation costs when the radius is set to be $10^{-1.4} - 10^{-1.0}$ of the estimation $\hat{d}_t$. Moreover, the difference between top 10% and bottom 10% of the out-of-sample evaluation costs becomes smaller around the $10^{-0.8} \cdot \hat{d}_t$ even for larger sample sizes. However, the median out-of-sample cost increases with the Wasserstein radius, suggesting that larger Wasserstein radii may lead to overly conservative policies.

To better quantify the trade-off between mean and variance of the out-of-sample evaluation costs, we present Figure 2 (and Figures 7 and 8 in Section B). Here, the lines connect the points representing Wasserstein DR-MCO models from the smallest radius to the largest one. We say one policy dominates another policy if the former has smaller mean and standard deviation than the latter does, and have the following observations. First, in all cases, the policy from the robust model is dominated by some policy from the Wasserstein DR-MCO model. Second, the policies from the nominal and CVaR MSCO models are all dominated by some policy from the DR-MCO model, when the sample size is small ($n_t = 5$). Last, policies from the CVaR MSCO model could penetrate the frontier formed by policies from the DR-MCO model, especially when the sample sizes increase to $n_t = 20$ or 40.

4.2.2 Inventory Problems with Uncertain Prices

Now we discuss the inventory problems with uncertain prices and fixed demands. These problems can be viewed as a simplified model for supply contract problems [28], where the goal is to estimate the total
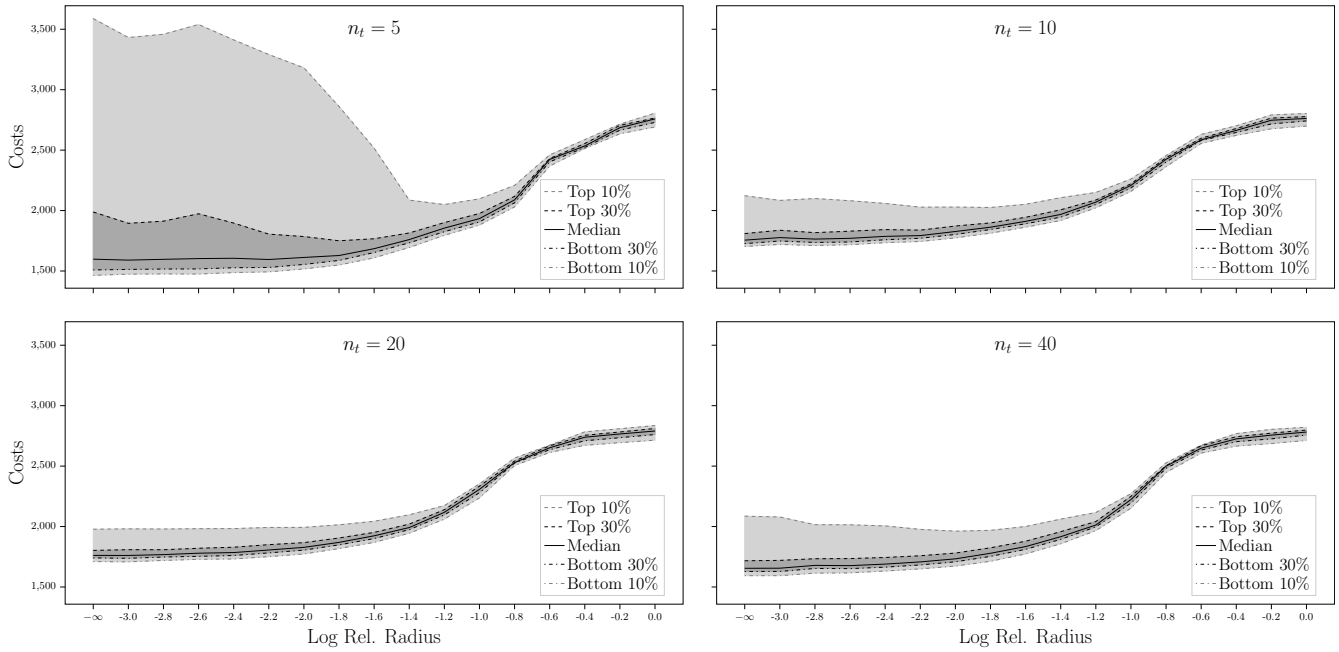
Figure 1: Out-of-sample Cost Quantiles for Different Radii on Multi-commodity Inventory with Uncertain Demands
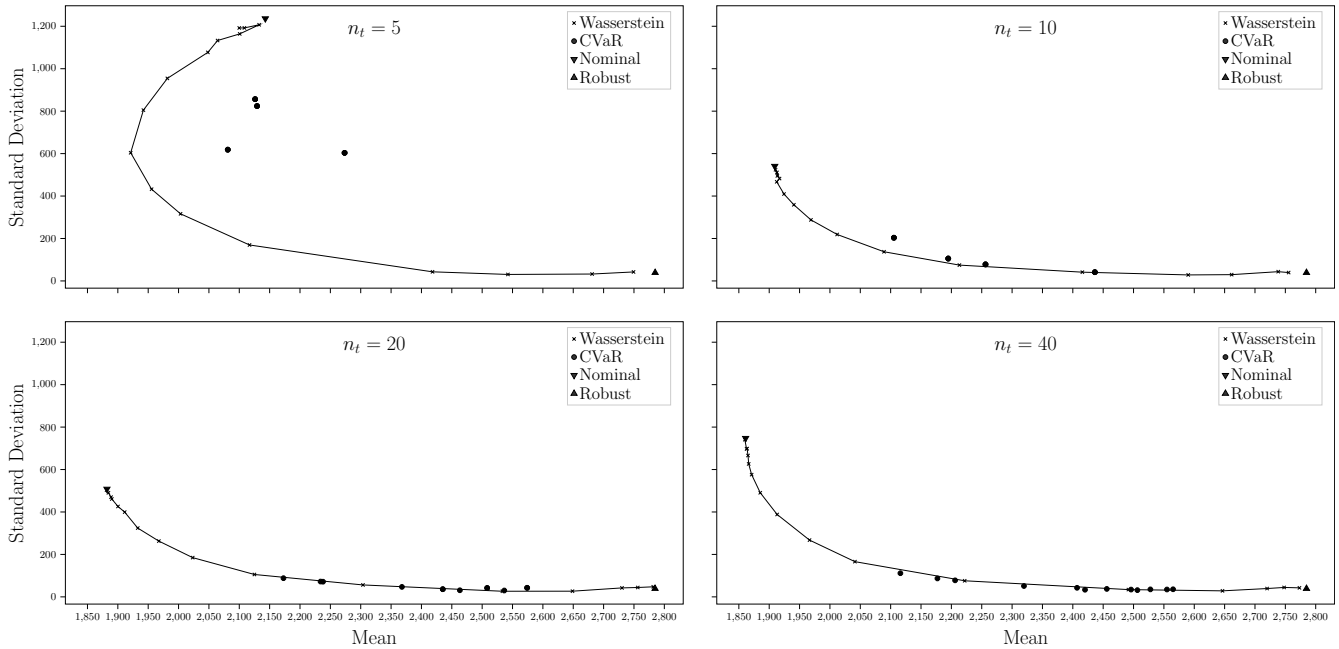

Figure 2: Comparison against Baseline Models on Multi-commodity Inventory with Uncertain Demands

cost of such supply contract and under-estimation would be very undesirable. The uncertain prices are modeled by the following expression:

$$C_{t,j}^b(\xi_t) := \xi_{t,j}, \quad C_{t,j}^a(\xi_t) := C_1 \cdot \xi_{t,j}, \quad j \in \mathcal{J}, t \in \mathcal{T}. \tag{36}$$

Here, C_1 is a factor for express orders. The uncertain vector $\xi_t \in \mathbb{R}_{\geq 0}^J$ follows a truncated multivariate normal distribution:

$$\xi_t := \max\left\{\text{Normal}(\mu_t, \bar{C} \cdot \Sigma_t), \underline{C}\right\}, \quad \mu_t := C_0\left[1 + \sin\left(\frac{2\pi(t+j)}{\tau}\right)\right], \tag{37}$$

where the maximum is taken componentwise, C_0 is a factor for base prices, τ is the period, $\bar{C}$ is the magnitude on the price variation, $\underline{C}$ is the lower bound on the prices, and the covariance matrix Σ_t is randomly generated (by multiplying a uniformly distributed random matrix with its transpose) and normalized to have its maximum eigenvalue equal to 1. The demands are deterministically given by

$$D_{t,j} := D_0\left[1 + \cos\left(\frac{2\pi(t+j)}{\tau}\right)\right] + \bar{D}, \quad j \in \mathcal{J}, t \in \mathcal{T}. \tag{38}$$

For the experiments, we consider $J = 5$ products, $T = 10$ stages, and the period $\tau = 5$. The price uncertainty has parameters $C_0 = 1$, $C_1 = 5$, $\bar{C} = 0.1$, and $\underline{C} = 0.001$. We choose the demand parameters $D_0 = 5$ and $\bar{D} = 10$. The inventory and holding costs are $C_j^H = 1$ and $C_j^B = 10$, and the rejection costs are $C_j^r = 100$, for each $j \in \mathcal{J}$. The bounds are set to $B^c = 15$, $B_j^a = 10$, $B_j^b = 20$, $B_j^{l,-} = 20$, and $B_j^{l,+} = 20$ for each $j \in \mathcal{J}$. The Wasserstein radii in the DR-MCO models are set relatively with respect to $\hat{d}_t$ defined in (35). The baseline MSCO models are constructed in the same way as described in Section 4.2.1. We do not consider the baseline MRCO model here as the uncertainty set Ξ_t is unbounded.
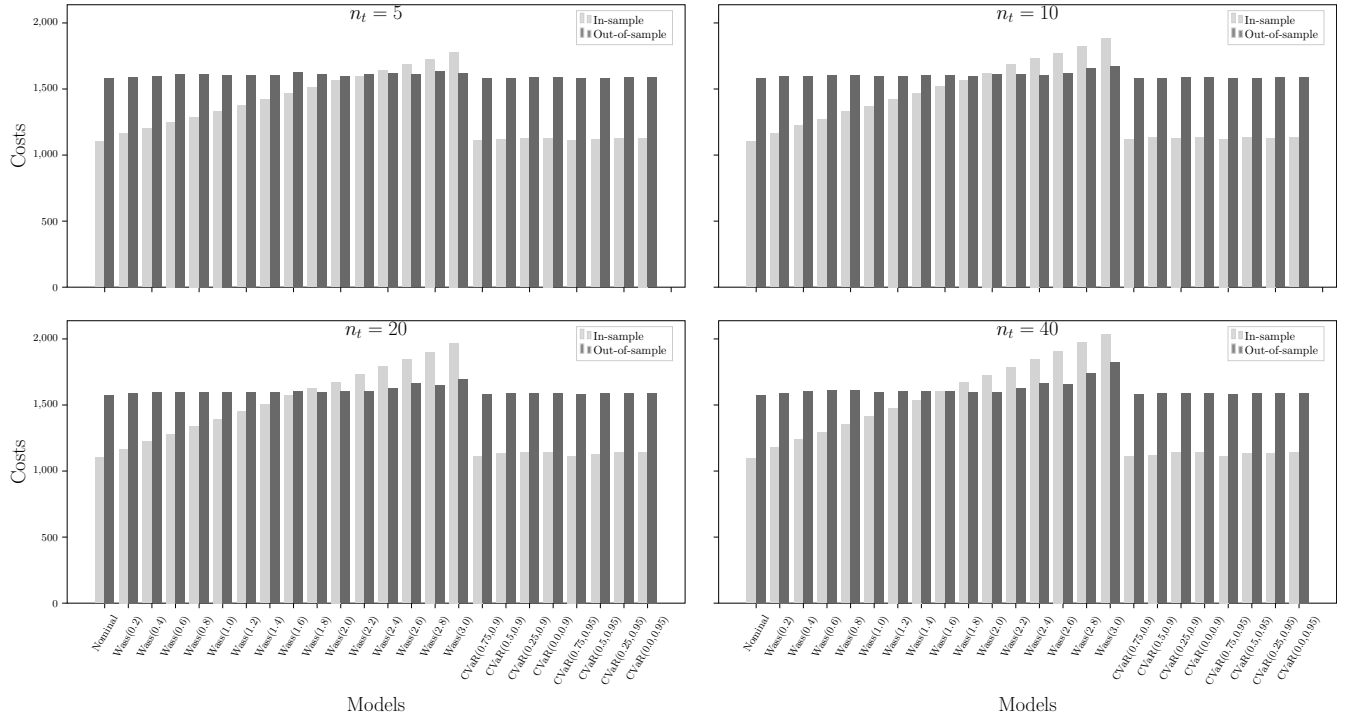

Figure 3: In-sample and Out-of-sample Mean Costs on Multi-commodity Inventory with Uncertain Prices

We plot the in-sample objective costs and out-of-sample mean evaluation costs in Figure 3 (and Figures 9 and 10 in Section B). The label Nominal refers to the nominal risk-neutral MSCO model using

the empirical probability measures $\hat{\nu}_t$; Wass(γ) refers to the Wasserstein DR-MCO model with radius $\rho_{t,0} = \gamma \cdot \hat{d}_t$ in each stage $t \geq 2$; and CVaR(α, β) refers to the CVaR risk-averse MSCO model with parameters α and β. As the uncertainty vectors now have an unbounded support, the robust model is no longer applicable. We see that in all cases, the in-sample objective cost grows linearly with respect to the Wasserstein distance, as predicted by Theorem 5. As the nominal stochastic model inevitably under-estimates the mean evaluation costs, using Wasserstein DR-MCO models with a relative radius $\gamma \in [1.6, 2.4]$ depending on the sample size n_t could achieve the out-of-sample performance guarantee in almost all test cases. Moreover, none of the CVaR risk-averse model in the experiments could achieve similar effect. Thus we believe that the Wasserstein DR-MCO models are particularly more favorable in the context of supply contracts. It is however worth mentioning that we do not observe any improvement of the mean or the variance of evaluation costs from the Wasserstein DR-MCO model over the baseline models.

4.3 Hydro-thermal Power Planning Problem

We next consider the Brazilian interconnected power system planning problem described in [11]. Let $\mathcal{J} = \{1, \ldots, J\}$ denote the indices of four regions in the system, and $\mathcal{L} = \cup_{j \in \mathcal{J}} \mathcal{L}_j$ the indices of thermal power plants, where each of the disjoint subsets $\mathcal{L}_j$ is associated with the region $j \in \mathcal{J}$. We first describe the decision variables in each stage $t \in \mathcal{T}$. We use $x_{t,j}^l$ to denote the stored energy level, $y_{t,j}^h$ to denote the hydro power generation, and $y_{t,j}^s$ to denote the energy spillage, of some region $j \in \mathcal{J}$; and $y_{t,l}^g$ to denote the thermal power generation for some thermal power plant $l \in \mathcal{L}$. For two different regions $j \neq j' \in \mathcal{J}$, we use $y_{t,j,j'}^e$ to denote the energy exchange from region j to region j', and $y_{t,j,j'}^a$ to denote the deficit account for region j in region j'. Let $x_t := (x_{t,1}^l, \ldots, x_{t,J}^l)$ be the state vector of energy levels, y_t the internal decision vector consisting of $y_{t,j}^s$ and $y_{t,j}^h$ for $j \in \mathcal{J}$, $y_{t,l}^g$ for all $l \in \mathcal{L}$, $y_{t,j,j'}^e$ and $y_{t,j,j'}^a$ for any $j \neq j' \in \mathcal{J}$; and $(\xi_{t,1}, \ldots, \xi_{t,J}) \in \mathbb{R}_{\geq 0}^J$ the uncertain vector energy inflows in stage t. We define the DR-MCO by specifying cost functions f_t as

$$f_t(x_{t-1}, x_t; \xi_t) := \min_{y_t} \quad \sum_{j \in \mathcal{J}} \left(C^s y_{t,j}^s + \sum_{l \in \mathcal{L}_j} C_l^g y_{t,l}^g + \sum_{j' \neq j} \left(C_{j,j'}^e y_{t,j,j'}^e + C_{j,j'}^a y_{t,j,j'}^a \right) \right) \tag{39}$$
$$\text{s.t.} \quad x_{t,j}^l + y_{t,j}^h + y_{t,j}^s = x_{t-1,j}^l + \xi_{t,j}, \ \forall j \in \mathcal{J},$$
$$y_{t,j}^h + \sum_{l \in \mathcal{L}_j} y_{t,l}^g + \sum_{j' \neq j} (y_{t,j,j'}^a - y_{t,j,j'}^e + y_{t,j',j}^e) = D_{t,j}, \ \forall j \in \mathcal{J},$$
$$y_{t,l}^g \in [B_l^{g,-}, B_l^{g,+}], \ \forall l \in \mathcal{L},$$
$$x_{t,j}^l \in [0, B_j^l], \ y_{t,j}^h \in [0, B_j^h], \ \forall j \in \mathcal{J},$$
$$y_{t,j,j'}^a \in [0, B_{j,j'}^a], \ y_{t,j,j'}^e \in [0, B_{j,j'}^e], \ \forall j, j' \in \mathcal{J}.$$

Here in the formulation, C^s denotes the unit penalty on energy spillage, C_l^g the unit cost of thermal power generation of plant l, $C_{j,j'}^e$ the unit cost of power exchange from region j to region j', $C_{j,j'}^a$ the unit cost on the energy deficit account for region j in region k', $D_{t,j}$ the deterministic power demand in stage t and region k, B_j^l the bound on the storage level in region j, B_j^h the bound on hydro power generation in region j, $B_l^{g,-}, B_l^{g,+}$ the lower and upper bounds of thermal power generation in plant l, $B_{j,j'}^a$ the bound on the deficit account for region j in region j' such that $\sum_{j' \neq j} B_{j,j'}^a = D_{t,j}$, and $B_{j,j'}^e$ the bound on the energy exchange from region j to region j'. The first constraint in (39) characterizes the change of energy storage levels in each region, the second constraint imposes the power generation-demand balance for each region, and the rest are bounds on the decision variables. The initial state x_0 and uncertainty vector ξ_1 are given by data. In our experiment, we consider $T = 13$ and all other parameters used in this problem can be found in [11].

For the problem (39), we have the following remarks. First, we always have the relatively complete recourse as we allow spillage for extra energy inflows and deficit for energy demands in each region. Then

it is straightforward to check that the Lipschitz constant of f_t in either x_{t-1} or ξ_t can be bounded by the maximum of the deficit cost $C_{j,j'}^a$ and the spillage penalty C^s. Second, as now the uncertainty has an unbounded support $\mathbb{R}_{\geq 0}^J$, we need to estimate the growth rate of the value function $r_t(x_{t-1})$. Note that if for any region the inflow is sufficiently large such that all demands and energy exchanges have met their upper bounds, then the only cost incurred by further increasing the inflow is simply the spillage penalty. Thus we conclude that $r_t(x_{t-1}) = C^s \cdot J$, which is a constant function for all $x_{t-1} \in \mathcal{X}_{t-1}$. Last, it is easy to see that Assumption 4 holds for the problem (39) and that the state variables $\mathcal{X}_t$ is compact, so our SSSO implementation (Algorithm 3) would guarantee the convergence of our DDP algorithm (Algorithm 1).

Now we assume that the true uncertainty can be described by the following logarithmic autoregressive time series:
$$\ln \xi_t - \mu_t = \varphi_t(\ln \xi_{t-1} - \mu_{t-1}) + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{Normal}(0, \Sigma_t), \tag{40}$$
where the logarithm and the product are taken componentwise, and the parameters $\mu_t, \varphi_t \in \mathbb{R}^J$ and $\Sigma_t \in \mathcal{S}_{\succeq 0}^J$ are fit from historical data (see modeling details in [44] and coding details in [11]). Note that (40) is not linear with respect to the uncertainty vectors ξ_t, and consequently it cannot directly be reformulated into a stagewise independent MSCO (or a DR-MCO) [30]. While there are approaches based on Markov chain DDP or linearized version of the model (40), they would require alteration of the DDP algorithm or an increase in the state space dimension. Alternatively, we would like to study the effects of the stagewise independence assumption [13] and the Wasserstein ambiguity sets in our DR-MCO model (2).

We can see that under the assumption on the true uncertainty (40), each uncertainty ξ_t follows a multivariate lognormal distribution. Thus instead of directly using the empirical probability measures $\hat{\nu}_t$ in the MSCO models, we can fit lognormal distributions based on the empirical outcomes $\hat{\xi}_{t,1}, \ldots, \hat{\xi}_{t,n_t}$, from which we further construct the SAA probability measures $\tilde{\nu}_t$ for the MSCO models. Moreover, we can also use this SAA probability measure $\tilde{\nu}_t$ to estimate the Wasserstein distance bound $\rho_{t,0}$ using

$$
\begin{aligned}
\tilde{d}_t := W_t(\hat{\nu}_t, \tilde{\nu}_t) = \min_{\pi_{k,k'} \geq 0} \quad & \sum_{k=1}^{n_t} \sum_{k'=1}^{n_t'} \pi_{k,k'} \cdot d_t(\hat{\xi}_{t,k}, \tilde{\xi}_{t,k'}) \\
\text{s.t.} \quad & \sum_{k'=1}^{n_t'} \pi_{k,k'} = \frac{1}{n_t}, \quad k = 1, \ldots, n_t, \\
& \sum_{k=1}^{n_t} \pi_{k,k'} = \frac{1}{n_t'}, \quad k = 1, \ldots, n_t'.
\end{aligned} \tag{41}
$$

We then set the Wasserstein radius to be $\rho_{t,0} = \gamma \cdot \tilde{d}_t$ for the relative factors $\gamma = 10^{-2.0}, 10^{-1.8}, \ldots, 10^{-0.2}, 1.0$. For the baseline risk-averse MSCO models, we use CVaR parameters $\alpha = 0.1$ and $\beta \in \{0.0, 0.1, \ldots, 0.9, 1.0\}$.

Note that as the SAA resampling step is random, the performance of our DR-MCO models and MSCO models would also be random. In addition, as it is often very challenging to solve the problem (39) to certain optimality gap, we choose to terminate it with a maximum of 1000 iterations, and allow random sampling in the Algorithm 1, in which the noninitial stage step does not strictly follow Definition 2 and guarantees only the validness. The benefit of such random sampling is that empirically the under-approximations (hence the policies) often improve faster especially in the beginning stage of the algorithms. We refer any interested readers to stochastic DDP literature (e.g., [2, 49, 26]) for more details.

Due to the randomness of the models and the algorithm, we repeat each test 3 times with the same empirical probability measures $\hat{\nu}_t$ in the experiment procedure. Then we plot the median values with error bars indicating the maximum and minimum values in Figure 4 (and Figures 11 and 12 in Section B) with increasing values of γ in the Wasserstein DR-MCO models and decreasing values of β in the CVaR risk-averse MSCO models. First, we see that most of the Wasserstein DR-MCO models and the CVaR
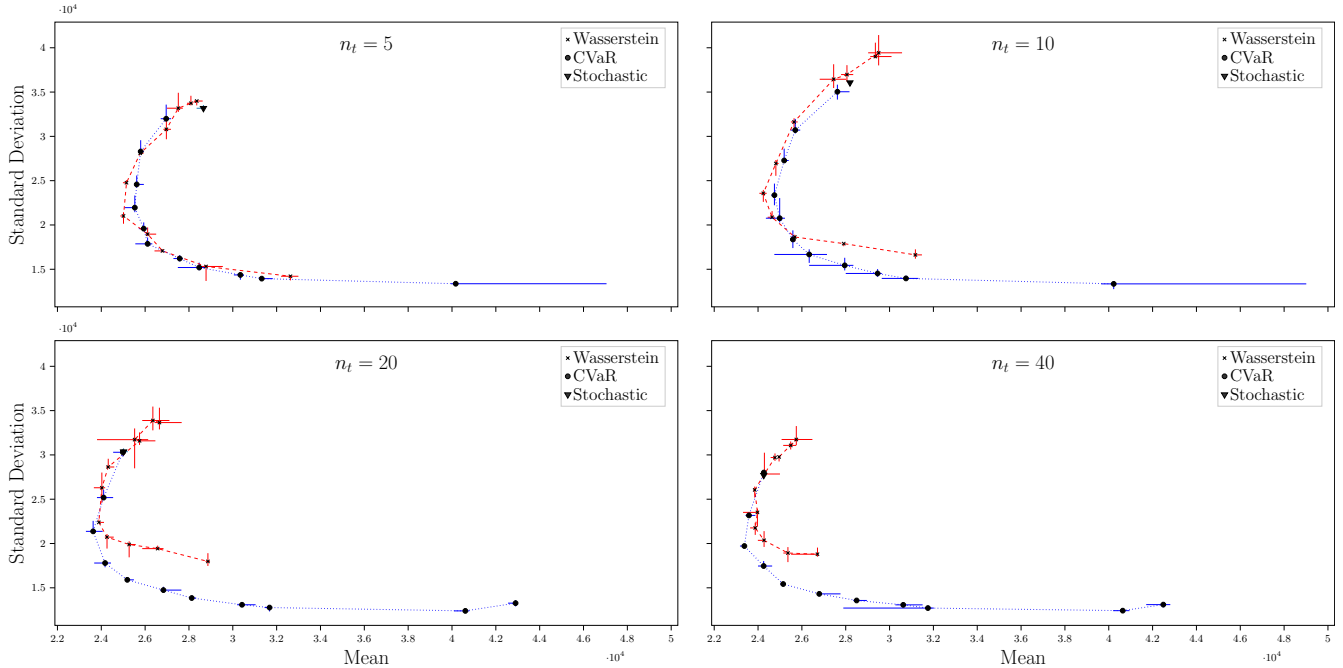
Figure 4: Comparison against Baseline Models on Hydro-thermal Power Planning

risk-averse MSCO models achieve better performances in either the mean or the standard deviation of evaluation costs, compared with the risk-neutral MSCO model. This observation justifies the usage of ambiguity sets or risk aversion when we approximate a stagewise dependent problem with stagewise independent models. Second, when the sample sizes are small ($n_t = 5$ or 10), we see that the Wasserstein DR-MCO could outperform the CVaR risk-averse MSCO models (e.g., for $\gamma \in [10^{-1.0}, 10^{-0.6}]$) in the out-of-sample mean cost. Third, as the sample size grows to $n_t = 20$ or 40, while the policies obtained from Wasserstein DR-MCO models are always dominated by those from the CVaR risk-averse MSCO models, the latter could achieve both lower mean cost and standard deviation. Our conjecture is that for larger sample sizes, the cost-to-go functions of the Wasserstein DR-MCO models have more complicated shapes, thus making it hard to approximate in the limited 1000 iterations. This also suggests that the CVaR risk-averse MSCO models could lead to good out-of-sample performances when an probability distribution fitting is possible and the computation budget is limited.

5 Concluding Remarks

In this work, we study the data-driven DR-MCO models with Wasserstein ambiguity sets. We show that with a sufficiently large Wasserstein radius, such DR-MCO model satisfies out-of-sample performance guarantee with high probability even with limited data sizes. Using convex dual reformulation, we show that the in-sample conservatism is linearly bounded by the radius when the value functions are Lipschitz continuous in the uncertainties. To numerically solve the data-driven DR-MCO models, we design exact SSSO subproblems for the DDP algorithms by exploiting either the concavity of the convexity of the cost functions in terms of the uncertainties. We conduct extensive numerical experiments to compare the performance of our DR-MCO models against MRCO models, risk-neutral and risk-averse MSCO models. On the multi-commodity inventory problems with uncertain demands, we observe that the DR-MCO models are able to provide policies that dominate those baseline model policies in the out-of-sample evaluations when the in-sample data size is small. Moreover, on the inventory problems with uncertain prices, the DR-MCO models could achieve out-of-sample performance guarantee with little compromise of the objective value, which has not been achieved by the baseline models. On the hydro-thermal power

planning problems with uncertain energy inflows, we see that with limited number of iterations, while the policies from the DR-MCO models could achieve better out-of-sample performances than the risk-averse MSCO models for small data sizes, they are dominated by the risk-averse MSCO models for larger data sizes. We hope these numerical experiments could serve as benchmarks for future studies on DR-MCO problems.

References

[1] R Baucke, A Downward, and G Zakeri. A deterministic algorithm for solving stochastic minimax dynamic programmes. *Optimization Online*, 2018.

[2] Regan Baucke, Anthony Downward, and Golbon Zakeri. A deterministic algorithm for solving multistage stochastic programming problems. *Optimization Online*, pages 1–25, 2017.

[3] Heinz H Bauschke, Patrick L Combettes, et al. *Convex analysis and monotone operator theory in Hilbert spaces*, volume 408. Springer, 2011.

[4] Güzin Bayraksan and David K Love. Data-driven stochastic programming using phi-divergences. In *The operations research revolution*, pages 1–19. INFORMS, 2015.

[5] Aharon Ben-Tal, Dick Den Hertog, Anja De Waegenaere, Bertrand Melenberg, and Gijs Rennen. Robust solutions of optimization problems affected by uncertain probabilities. *Management Science*, 59(2):341–357, 2013.

[6] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*, volume 28. Princeton university press, 2009.

[7] Dimitris Bertsimas, Vishal Gupta, and Nathan Kallus. Data-driven robust optimization. *Mathematical Programming*, 167(2):235–292, 2018.

[8] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98, 2017.

[9] Michele Conforti, Gérard Cornuéjols, Giacomo Zambelli, et al. *Integer programming*, volume 271. Springer, 2014.

[10] Erick Delage and Yinyu Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations research*, 58(3):595–612, 2010.

[11] Lingquan Ding, Shabbir Ahmed, and Alexander Shapiro. A python package for multi-stage stochastic programming. *Optimization online*, pages 1–42, 2019.

[12] Iain Dunning, Joey Huchette, and Miles Lubin. Jump: A modeling language for mathematical optimization. *SIAM Review*, 59(2):295–320, 2017.

[13] Daniel Duque and David P Morton. Distributionally robust stochastic dual dynamic programming. *SIAM Journal on Optimization*, 30(4):2841–2865, 2020.

[14] Andreas Eichhorn and Werner Römisch. Polyhedral risk measures in stochastic programming. *SIAM Journal on Optimization*, 16(1):69–95, 2005.

[15] Peyman Mohajerin Esfahani and Daniel Kuhn. Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1):115–166, 2018.

[16] Nicolas Fournier and Arnaud Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3):707–738, 2015.

[17] Rui Gao and Anton Kleywegt. Distributionally robust stochastic optimization with wasserstein distance. *Mathematics of Operations Research*, 2022.

[18] Angelos Georghiou, Angelos Tsoukalas, and Wolfram Wiesemann. Robust dual dynamic programming. *Operations Research*, 67(3):813–830, 2019.

[19] Pierre Girardeau, Vincent Leclere, and Andrew B Philpott. On the convergence of decomposition methods for multistage stochastic convex programs. *Mathematics of Operations Research*, 40(1):130–145, 2015.

[20] Vincent Guigues. Convergence analysis of sampling-based decomposition methods for risk-averse multistage stochastic convex programs. *SIAM Journal on Optimization*, 26(4):2468–2494, 2016.

[21] Vincent Guigues and Werner Römisch. Sampling-based decomposition methods for multistage stochastic programs based on extended polyhedral risk measures. *SIAM Journal on Optimization*, 22(2):286–312, 2012.

[22] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2022.

[23] Grani A Hanasusanto and Daniel Kuhn. Conic programming reformulations of two-stage distributionally robust linear programs over wasserstein balls. *Operations Research*, 66(3):849–869, 2018.

[24] Hidetoshi Komiya. Elementary proof for sion's minimax theorem. *Kodai mathematical journal*, 11(1):5–7, 1988.

[25] Václav Kozmík and David P Morton. Evaluating policies in risk-averse multi-stage stochastic programming. *Mathematical Programming*, 152(1):275–300, 2015.

[26] Guanghui Lan. Complexity of stochastic dual dynamic programming. *Mathematical Programming*, pages 1–38, 2020.

[27] Guanghui Lan. Correction to: Complexity of stochastic dual dynamic programming. *Mathematical Programming*, pages 1–3, 2022.

[28] Chung-Lun Li and Panos Kouvelis. Flexible and risk-sharing supply contracts under price uncertainty. *Management science*, 45(10):1378–1398, 1999.

[29] Fengming Lin, Xiaolei Fang, and Zheming Gao. Distributionally robust optimization: A review on theory and applications. *Numerical Algebra, Control & Optimization*, 12(1):159, 2022.

[30] Nils Löhndorf and Alexander Shapiro. Modeling time-dependent randomness in stochastic dual dynamic programming. *European Journal of Operational Research*, 273(2):650–661, 2019.

[31] Mario VF Pereira and Leontina MVG Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical programming*, 52(1):359–375, 1991.

[32] Andrew B Philpott, Vitor L de Matos, and Lea Kapelevich. Distributionally robust sddp. *Computational Management Science*, 15(3):431–454, 2018.

[33] Andrew B Philpott and Ziming Guan. On the convergence of stochastic dual dynamic programming and related methods. *Operations Research Letters*, 36(4):450–455, 2008.

[34] Andy Philpott, Vitor de Matos, and Erlon Finardi. On solving multistage stochastic programs with coherent risk measures. *Operations Research*, 61(4):957–970, 2013.

[35] Alois Pichler and Alexander Shapiro. Mathematical foundations of distributionally robust multistage optimization. *SIAM Journal on Optimization*, 31(4):3044–3067, 2021.

[36] Hamed Rahimian, Güzin Bayraksan, and Tito Homem-de Mello. Identifying effective scenarios in distributionally robust stochastic programs with total variation distance. *Mathematical Programming*, 173(1):393–430, 2019.

[37] Hamed Rahimian and Sanjay Mehrotra. Distributionally robust optimization: A review. *arXiv preprint arXiv:1908.05659*, 2019.

[38] R Tyrrell Rockafellar. *Convex analysis*. Number 28. Princeton university press, 1970.

[39] R Tyrrell Rockafellar and Roger J-B Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 2009.

[40] Herbert E Scarf. *A min-max solution of an inventory problem*. Rand Corporation Santa Monica, 1957.

[41] Alexander Shapiro. Analysis of stochastic dual dynamic programming method. *European Journal of Operational Research*, 209(1):63–72, 2011.

[42] Alexander Shapiro and Shabbir Ahmed. On a class of minimax stochastic programs. *SIAM Journal on Optimization*, 14(4):1237–1249, 2004.

[43] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczynski. *Lectures on stochastic programming: modeling and theory*. SIAM, 2021.

[44] Alexander Shapiro, Wajdi Tekaya, Joari P da Costa, and Murilo P Soares. Final report for technical cooperation between georgia institute of technology and ons-operador nacional do sistema elétrico. *Georgia Tech ISyE Report*, 2012.

[45] Alexander Shapiro, Wajdi Tekaya, Joari Paulo da Costa, and Murilo Pereira Soares. Risk neutral and risk averse stochastic dual dynamic programming method. *European journal of operational research*, 224(2):375–391, 2013.

[46] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

[47] Wolfram Wiesemann, Daniel Kuhn, and Melvyn Sim. Distributionally robust convex optimization. *Operations Research*, 62(6):1358–1376, 2014.

[48] Shixuan Zhang and Xu Andy Sun. On distributionally robust multistage convex optimization: New algorithms and complexity analysis. *arXiv preprint arXiv:2010.06759*, 2020.

[49] Shixuan Zhang and Xu Andy Sun. Stochastic dual dynamic programming for multistage stochastic mixed-integer nonlinear optimization. *Mathematical Programming*, pages 1–51, 2022.

[50] Chaoyue Zhao and Yongpei Guan. Data-driven risk-averse stochastic optimization with wasserstein metric. *Operations Research Letters*, 46(2):262–267, 2018.

A Finite-dimensional Dual Recursion for DR-MCO

In this section, we briefly review some general strong Lagrangian duality theory and then apply it to derive our finite-dimensional dual recursion for our DR-MCO problems (2).

A.1 Generalized Slater Condition and Lagrangian Duality

Given an $\mathbb{R}$-vector space $\mathcal{M}$, we consider the following optimization problem.

$$v^{\mathrm{P}} := \inf_{\mu \in C} \quad \varphi_0(\mu) \tag{42}$$
$$\text{s.t.} \quad \varphi_j(\mu) \leq 0, \quad j = 1, \ldots, l,$$
$$\varphi_j(\mu) = 0, \quad j = l+1, \ldots, m.$$

Here, $C \subset \mathcal{M}$ is a convex subset, the functions $\varphi_j : \mathcal{M} \to \mathbb{R} \cup \{+\infty\}$ are convex for each $j = 0, 1, \ldots, l$ and $\varphi_j : \mathcal{M} \to \mathbb{R}$ are affine for each $j = l+1, \ldots, m$. Using a vector of multipliers $\lambda \in \mathbb{R}^m$, the Lagrangian dual problem of (42) can be written as

$$v^{\mathrm{D}} := \sup_{\lambda \in \Lambda} \inf_{\mu \in C} \left\{ \varphi_0(\mu) + \sum_{j=1}^{m} \lambda_j \varphi_j(\mu) \right\}, \tag{43}$$

where the admissible set for the multipliers is defined as $\Lambda := \{\lambda \in \mathbb{R}^m : \lambda_j \geq 0, \, \forall\, j = 1, \ldots, l\}$. We want to show the strong duality between (42) and (43), given the following condition.

Definition 3. *We say that the problem* (42) *satisfies the (generalized) Slater condition if the point $\eta = 0$ is in the relative interior of the effective domain of the convex value function associated with the primal problem* (42)

$$v(\eta) := \inf_{\mu \in C} \left\{ \varphi_0(\mu) : \varphi_j(\mu) = \eta_j, \, j = 1, \ldots, l, \text{ and } \varphi_j(\mu) \leq \eta_j, \, j = l+1, \ldots, m \right\}, \quad \eta \in \mathbb{R}^m.$$

Recall that the effective domain of a convex function $v : \mathbb{R}^m \to \mathbb{R} \cup \{\pm\infty\}$ is defined as $\mathrm{dom}\, v := \{\eta \in \mathbb{R}^m : v(\eta) < +\infty\}$, which is clearly a convex set. The affine hull of a convex set $K \subset \mathbb{R}^m$ is defined to be the smallest affine space containing K, and the relative interior of K is the interior of K viewed as a subset of its affine hull (equipped with the subspace topology). By convention, we have $v(\eta) = +\infty$ if there is no $\mu \in C$ such that $\varphi_j(\mu) \leq \eta_j$ for all $j = 1, \ldots, m$.

Proposition 15. *Assuming the Slater condition, the strong duality holds $v^{\mathrm{P}} = v^{\mathrm{D}}$ with an optimal dual solution $\lambda^* \geq 0$ (i.e., the supremum in the dual problem* (43) *is attained).*

Proof. The weak duality $v^{\mathrm{P}} \geq v^{\mathrm{D}}$ holds with a standard argument of exchanging the inf and sup operators, so it suffices to show that $v^{\mathrm{P}} \leq v^{\mathrm{D}}$. If $v^{\mathrm{P}} = -\infty$ then the inequality holds trivially, so we assume that $v^{\mathrm{P}} > -\infty$. Given the Slater condition, the value function $v(\eta)$ of the primal problem (42) must be proper $v(\eta) > -\infty$ for all $\eta \in \mathbb{R}^m$ (ref. Theorem 7.2 in [38]) because $\eta = 0$ is in the relative interior of the effective domain of v and $v(0) > -\infty$. Thus it is also subdifferentiable at the point $\eta = 0$ (ref. Theorem 23.4 in [38]), i.e., there exists a subgradient vector $\lambda^* \in \mathbb{R}^m$ such that $v(\eta) \geq v(0) - (\lambda^*)^{\mathsf{T}}\eta$ for any $\eta \in \mathbb{R}^m$. Here, for each $j = 1, \ldots, l$, the multiplier λ_j^* must be nonnegative since the function $v(\eta)$ is not increasing in the j-th component, so we have $\lambda \in \Lambda$. Since the inequality $v(\eta) + (\lambda^*)^{\mathsf{T}}\eta \geq v(0) = v^{\mathrm{P}}$ holds

for any $\eta \in \mathbb{R}^m$, we have

$$
\begin{aligned}
v^{\mathrm{P}} &\leq \inf_{\eta \in \mathbb{R}^m} \{v(\eta) + (\lambda^*)^{\mathsf{T}} \eta\} \\
&= \inf_{\mu \in C} \inf_{\eta \in \mathbb{R}^m} \left\{ \varphi_0(\mu) + \sum_{j=1}^m \lambda_j^* \eta_j : \varphi_j(\mu) \leq \eta_j, \ j = 1, \dots, l, \ \varphi_j(\mu) = \eta_j, \ j = l+1, \dots, m \right\} \\
&= \inf_{\mu \in C} \left\{ \varphi_0(\mu) + \sum_{j=1}^m \lambda_j^* \varphi_j(\mu) \right\} \leq v^{\mathrm{D}}.
\end{aligned}
$$

The first equality here results from exchanging two infimum operators, while the second one follows by taking $\eta_j = \varphi_j(\mu)$, due to the nonnegativity of λ_j^* for each $j = 1, \dots, l$, and replacing η_j with $\varphi_j(\mu)$ for each $j = l+1, \dots, m$. $\qquad\square$ $\qquad\qquad\square$

The strong Lagrangian duality guaranteed by the Slater condition is useful for many applications because we do not have to specify the topology on the vector space $\mathcal{M}$. The corollary below summarizes a special case where there is no equality constraint, and all the inequality constraints can be strictly satisfied.

Corollary 16. *For problems* (42) *and* (43) *with $l = m$ (no equality constraints), the strong duality holds if there exists a point $\bar{\mu} \in C$ such that $\varphi_j(\bar{\mu}) < 0$, for each $j = 1, \dots, m$.*

Proof. Let $\varepsilon_j := -\varphi_j(\mu) > 0$ for $j = 1, \dots, m$, and $U := \prod_{j=1}^m (-\frac{\varepsilon_j}{2}, \frac{\varepsilon_j}{2}) \subset \mathbb{R}^m$ be an open hyperrectangle. Then for any $\eta \in U$, we have $v(\eta) \leq \varphi_0(\bar{\mu}) < +\infty$. Therefore, we know that the Slater condition holds because $0 \in U \subset \operatorname{dom} v$, and the result follows from Proposition 15. $\qquad\square$ $\qquad\qquad\square$

Assuming that $\mathcal{M}$ is normed and complete, we could have another generalized Slater condition for problems with equality constraints. However, it is in general much harder to check so we do not base our discussion in Section 2 on it.

A.2 Finite-dimensional Dual Recursion

Now we derive the finite-dimensional dual recursion for the DR-MCO (2). Recall that by the definition of Wasserstein distance (5), the constraint $W_t(p, \hat{\nu}_t) \leq \rho_{t,0}$ is ensured if there exists a probability measure on the product space $\pi_t \in \mathcal{M}^{\mathrm{Prob}}(\Xi_t \times \Xi_t)$ with marginal probability measures $P_*^1(\pi_t) = p$ and $P_*^2(\pi_t) = \hat{\nu}_t$, such that

$$
\rho_{t,0} \geq \int_{\Xi_t \times \Xi_t} d_t \, \mathrm{d}\pi_t = \int_{\Xi_t} \int_{\Xi_t} d_t(\xi, \xi') \, \mathrm{d}\pi_t(\xi | \xi') \, \mathrm{d}\hat{\nu}_t(\xi') = \frac{1}{n_t} \sum_{k=1}^{n_t} \int_{\Xi_t} d_{t,k}(\xi) \, \mathrm{d}p_{t,k}(\xi) \tag{44}
$$

where we define $p_{t,k} := \pi_t(\cdot | \hat{\xi}_{t,k}) \in \mathcal{W}_t$ to be the probability measure conditioned on $\{\xi' = \hat{\xi}_{t,k}\}$ and $d_{t,k}(\xi) := d(\xi, \hat{\xi}_{t,k})$ for any $\xi \in \Xi_t$. Then we have $p = \frac{1}{n_t} \sum_{k=1}^{n_t} p_{t,k}$ by the law of total probability. Due to this condition, we define the following parametrized optimization problem for any $\rho > 0$ and any state $x_{t-1} \in \mathcal{X}_{t-1}$

$$
\begin{aligned}
q_t(x_{t-1}; \rho) := \sup_{p_{t,k} \in \mathcal{W}_t} \quad & \frac{1}{n_t} \sum_{k=1}^{n_t} \int_{\Xi_t} Q_t(x_{t-1}; \xi_t) \, \mathrm{d}p_{t,k}(\xi_t) \\
\text{s.t.} \quad & \frac{1}{n_t} \sum_{k=1}^{n_t} \int_{\Xi_t} d_{t,k} \, \mathrm{d}p_{t,k} \leq \rho, \\
& \frac{1}{n_t} \sum_{k=1}^{n_t} \int_{\Xi_t} g_{t,j} \, \mathrm{d}p_{t,k} \leq \rho_{t,j}, \quad j = 1, \dots, m_t.
\end{aligned} \tag{45}
$$

From the discussion above, we see that $\mathcal{Q}_{t-1}(x_{t-1}) \geq q_t(x_{t-1}; \rho_{t,0})$. We next show that the equality holds assuming the strict feasibility of the empirical measure $\hat{\nu}_t$.

Let $\lambda \in \mathbb{R}_{\geq 0}^{m_t+1}$ denote a multiplier vector. Assumption 1 guarantees by Corollary 16 the strong duality of the following Lagrangian dual problem

$$q_t(x_{t-1}; \rho) = \min_{\lambda \geq 0} \left\{ \rho\lambda_0 + \sum_{j=1}^{m_t} \rho_{t,j}\lambda_j + \frac{1}{n_t} \sum_{k=1}^{n_t} \sup_{p_{t,k} \in \mathcal{W}_t} \int_{\Xi_t} \left[Q_t(x_{t-1}; \cdot) - \lambda_0 d_{t,k} - \sum_{j=1}^{m_t} \lambda_j g_{t,j} \right] \mathrm{d}p_{t,k} \right\}$$

$$= \min_{\lambda \geq 0} \left\{ \rho\lambda_0 + \sum_{j=1}^{m_t} \rho_{t,j}\lambda_j + \frac{1}{n_t} \sum_{k=1}^{n_t} \sup_{\xi_k \in \Xi_t} \left\{ Q_t(x_{t-1}; \xi_k) - \lambda_0 d_{t,k}(\xi_k) - \sum_{j=1}^{m_t} \lambda_j g_{t,j}(\xi_k) \right\} \right\}. \quad (46)$$

where the second equality holds because each Dirac measure centered at $\xi \in \Xi_t$ satisfies $\delta_\xi \in \mathcal{W}_t$ and each $p_{t,k} \in \mathcal{W}_t$ is a probability measure. We are now ready to show that $\mathcal{Q}_{t-1}(x_{t-1}) = q_t(x_{t-1}; \rho_{t,0})$, which implies Theorem 1.

Theorem 17. *Under Assumption 1, in any stage $t \geq 2$, we have $q_t(x_{t-1}; \rho)$ is a concave function in ρ for any fixed x_{t-1}. Consequently, the expected cost-to-go function (2) satisfies $\mathcal{Q}_{t-1}(x_{t-1}) = q_t(x_{t-1}; \rho_{t,0})$ and thus can be equivalently rewritten as*

$$\mathcal{Q}_{t-1}(x_{t-1}) = \min_{\lambda \geq 0} \left\{ \sum_{j=0}^{m_t} \rho_{t,j}\lambda_j + \frac{1}{n_t} \sum_{k=1}^{n_t} \sup_{\xi_k \in \Xi_t} \left\{ Q_t(x_{t-1}; \xi_k) - \lambda_0 d_{t,k}(\xi_k) - \sum_{j=1}^{m_t} \lambda_j g_{t,j}(\xi_k) \right\} \right\}.$$

Proof. Let us fix any $x_{t-1} \in \mathcal{X}_{t-1}$. The first assertion on the concavity of $q_t(x_{t-1}; \rho)$ follows directly from (46) since $q_t(x_{t-1}; \cdot)$ is a minimum of affine functions. Moreover, from the definition (45) we see that $q_t(x_{t-1}; \rho) \geq 0$ for any $\rho > 0$ as the measures $p_{t,k} = \delta_{\hat{\xi}_{t,k}}$ satisfy the constraints and $Q_t(x_{t-1}; \xi_t) \geq 0$ by the nonnegativity of the cost functions f_t. If $q_t(x_{t-1}; \rho_{t,0}) = +\infty$, then the equality holds trivially as we already showed that $\mathcal{Q}_{t-1}(x_{t-1}) \geq q_t(x_{t-1}; \rho_{t,0})$. Otherwise, we must have $q_t(x_{t-1}; \rho) < +\infty$ for any $\rho > 0$ due to the concavity. Thus $q_t(x_{t-1}; \cdot)$ is a continuous function on $(0, +\infty)$.

To prove the inequality $\mathcal{Q}_{t-1}(x_{t-1}) \leq q_t(x_{t-1}; \rho_{t,0})$, take any $\varepsilon > 0$. From the definition (5), the constraint $W_t(p, \hat{\nu}_t) \leq \rho_{t,0}$ implies that there exists a probability measure $\pi_t \in \mathcal{M}^{\text{Prob}}(\Xi_t \times \Xi_t)$ with marginal probability measures $P_*^1(\pi_t) = p$ and $P_*^2(\pi_t) = \hat{\nu}_t$ such that $\int d_t \, \mathrm{d}\pi_t \leq \rho_{t,0} + \varepsilon$. In other words, we have $\mathcal{Q}_{t-1}(x_{t-1}) \leq q_t(x_{t-1}; \rho_{t,0} + \varepsilon)$. Now by the continuity of $q_t(x_{t-1}; \cdot)$, we conclude that $\mathcal{Q}_{t-1}(x_{t-1}) \leq \lim_{\varepsilon \to 0+} q_t(x_{t-1}; \rho_{t,0} + \varepsilon) = q_t(x_{t-1}; \rho_{t,0})$. □ □

B Supplemental Numerical Results

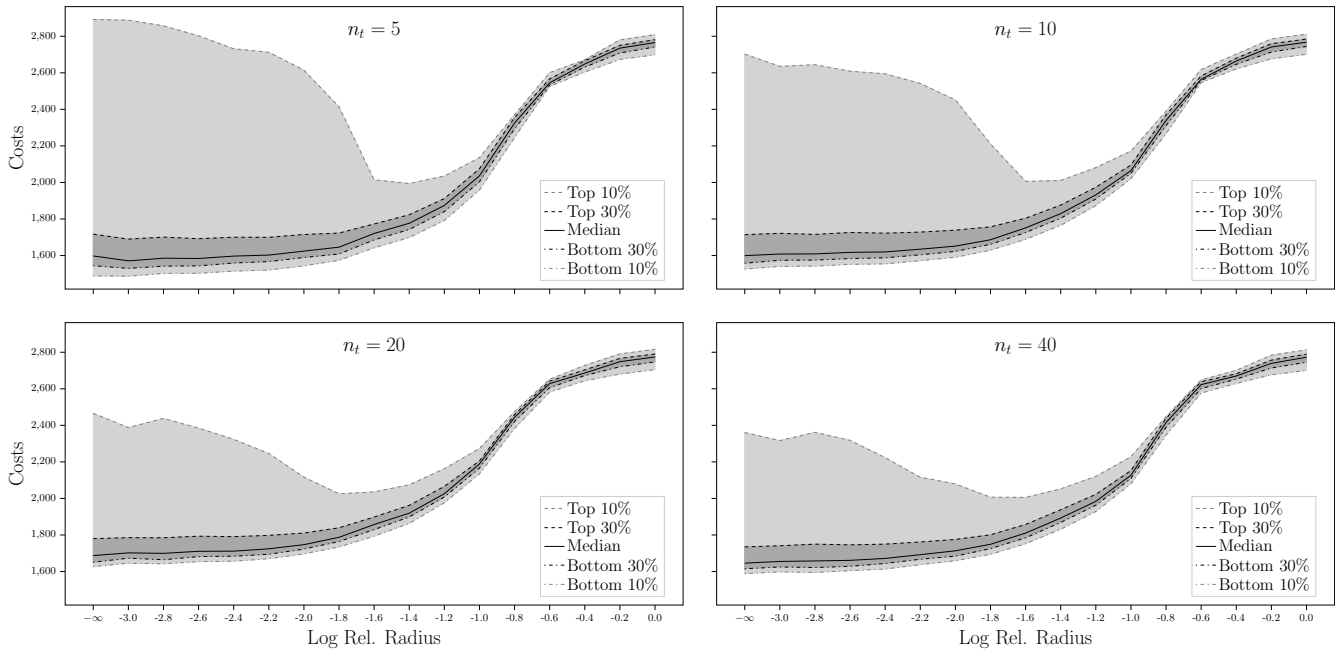In this section, we display supplemental results from our numerical experiments in Section 4.

Figure 5: Out-of-sample Cost Quantiles for Different Radii on Multi-commodity Inventory with Uncertain Demands, Additional Run 1
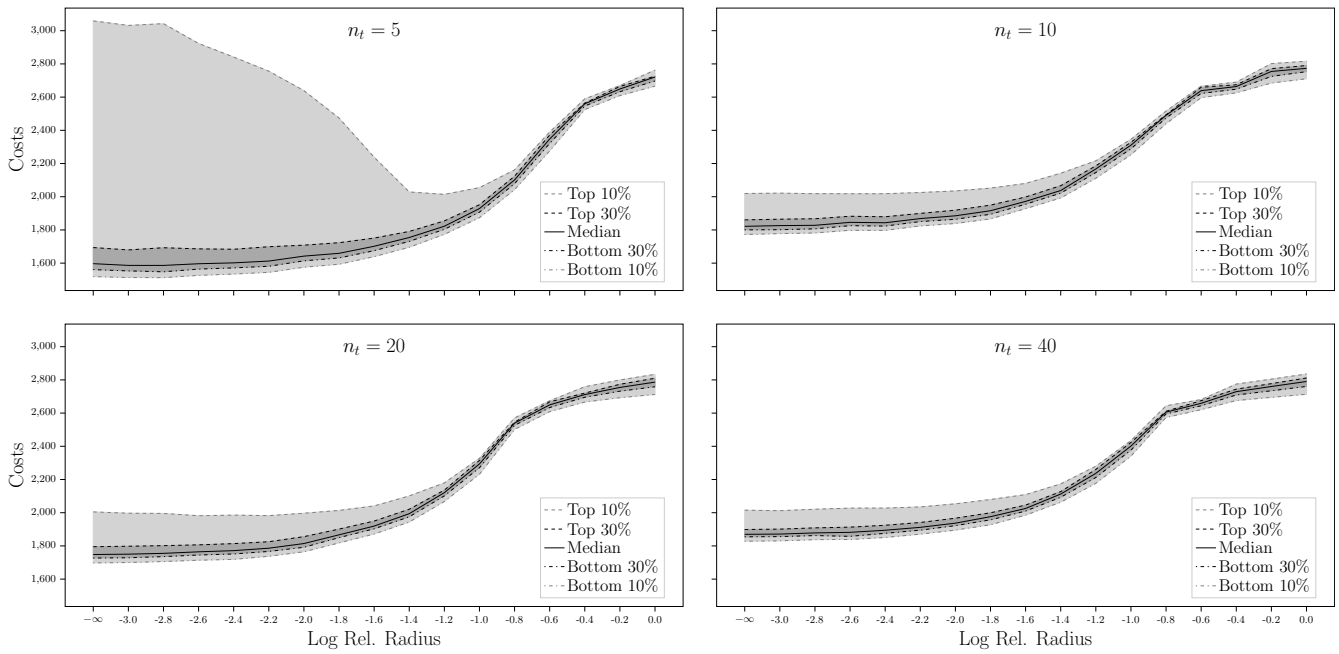

Figure 6: Out-of-sample Cost Quantiles for Different Radii on Multi-commodity Inventory with Uncertain Demands, Additional Run 2
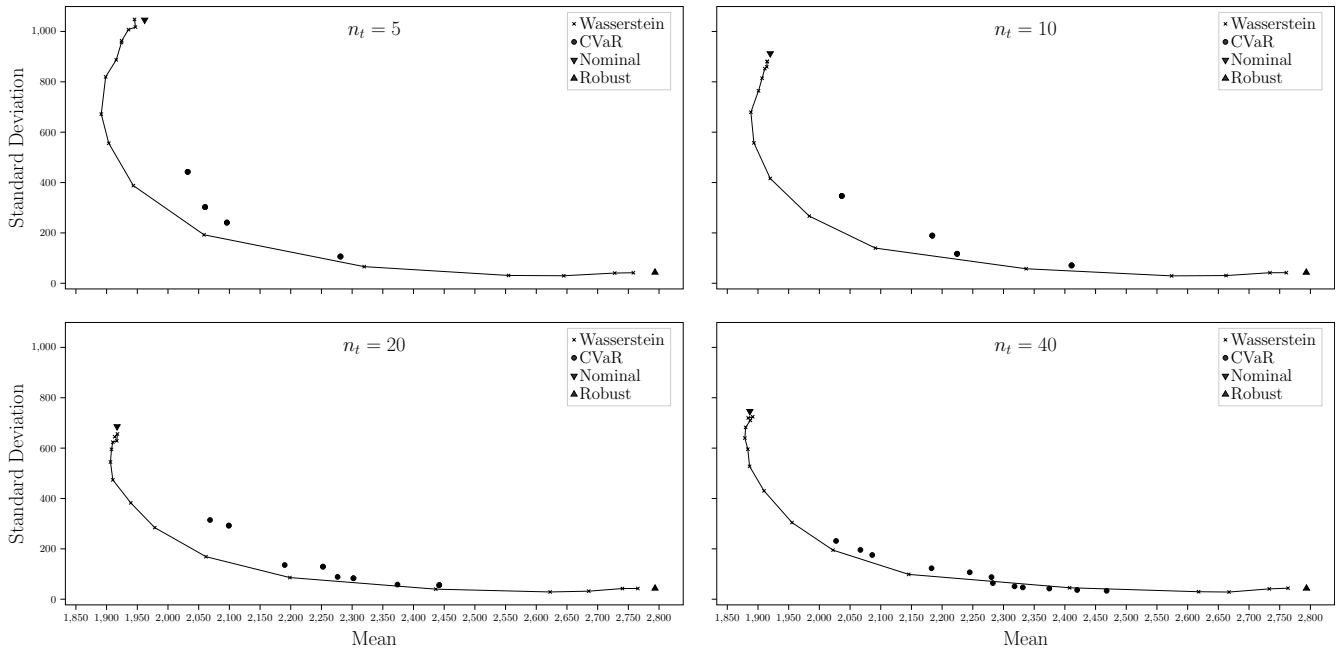
Figure 7: Comparison against Baseline Models on Multi-commodity Inventory with Uncertain Demands, Additional Run 1
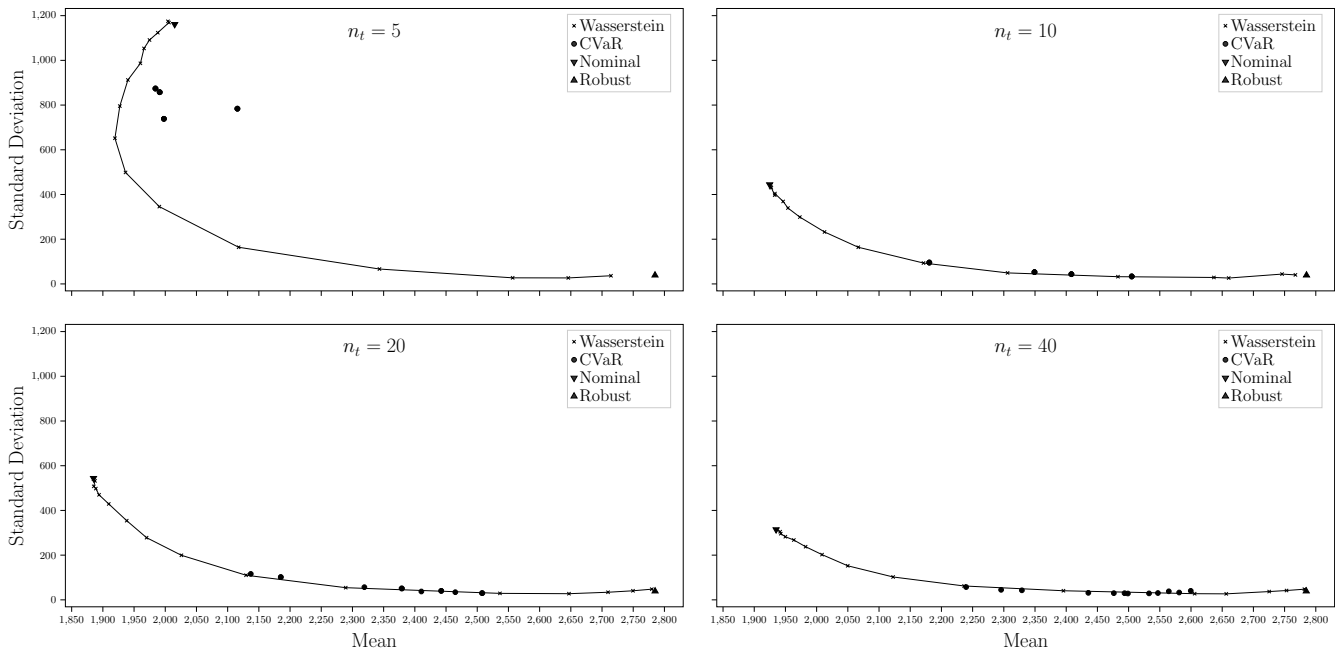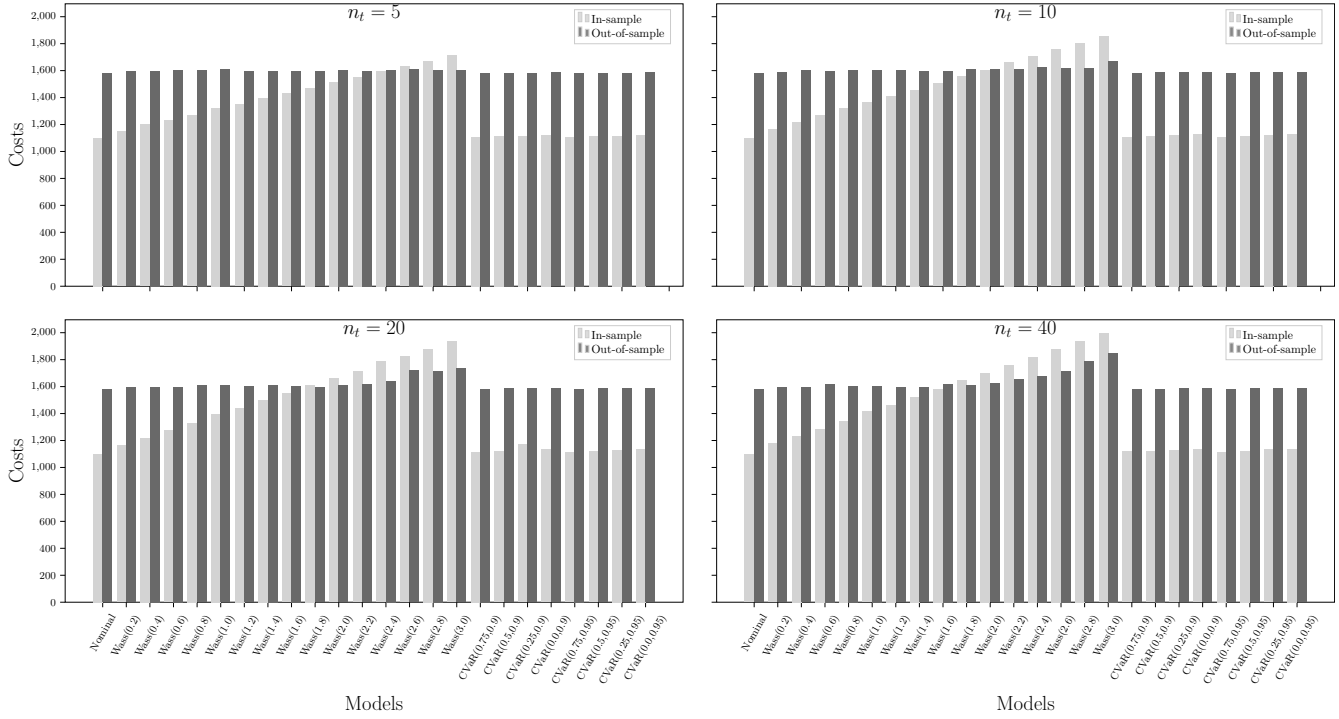

Figure 8: Comparison against Baseline Models on Multi-commodity Inventory with Uncertain Demands, Additional Run 2

Figure 9: In-sample and Out-of-sample Mean Costs on Multi-commodity Inventory with Uncertain Prices, Additional Run 1
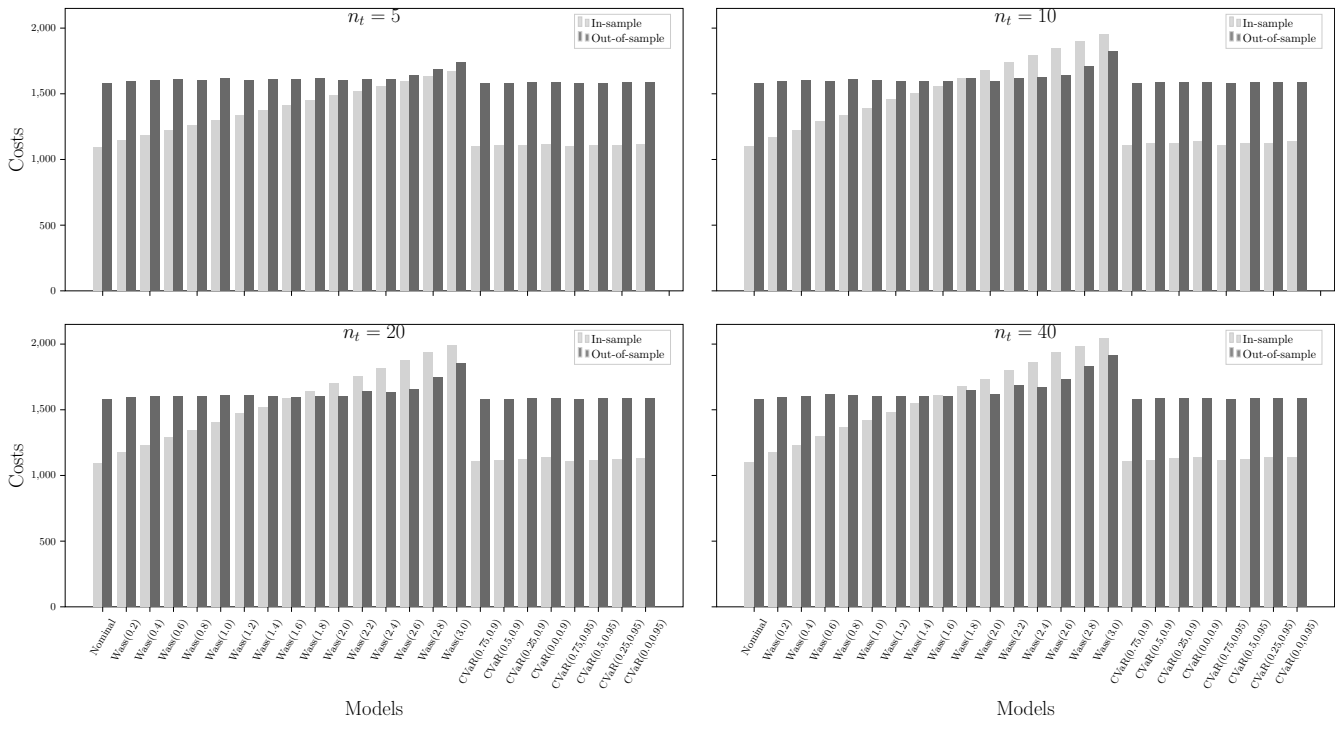

Figure 10: In-sample and Out-of-sample Mean Costs on Multi-commodity Inventory with Uncertain Prices, Additional Run 2
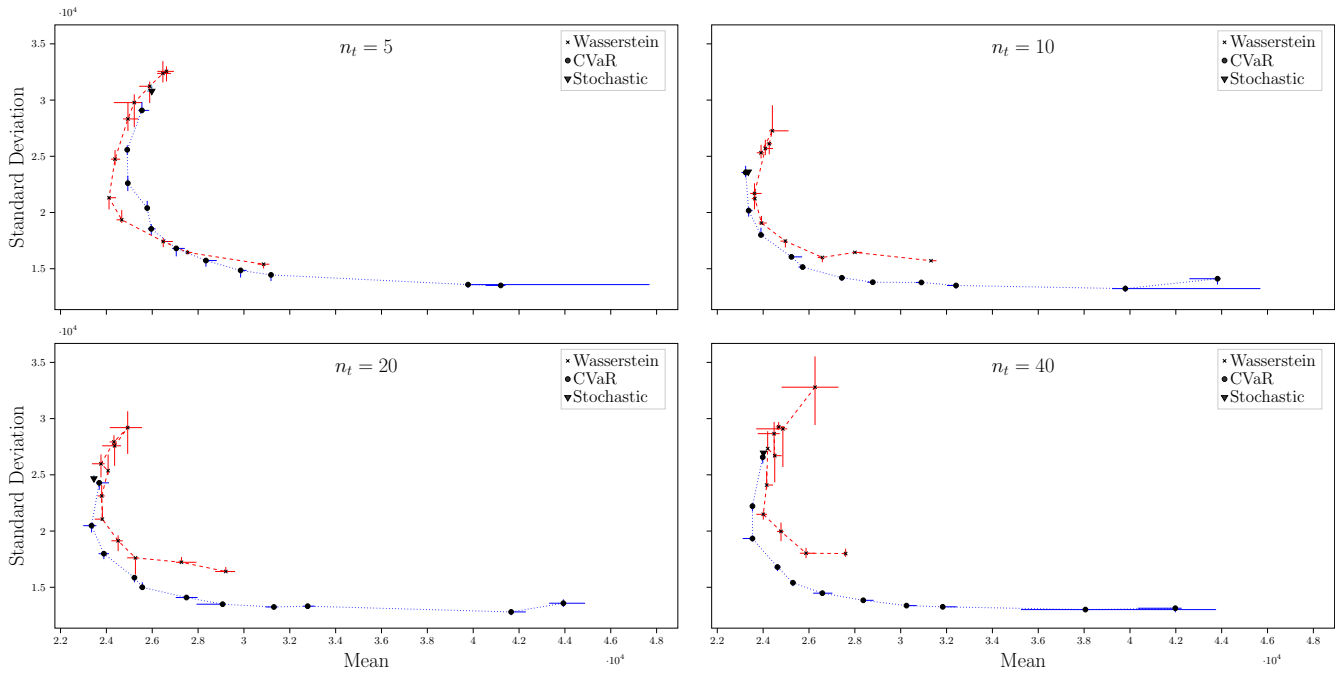
Figure 11: Comparison against Baseline Models on Hydro-thermal Power Planning, Additional Run 1
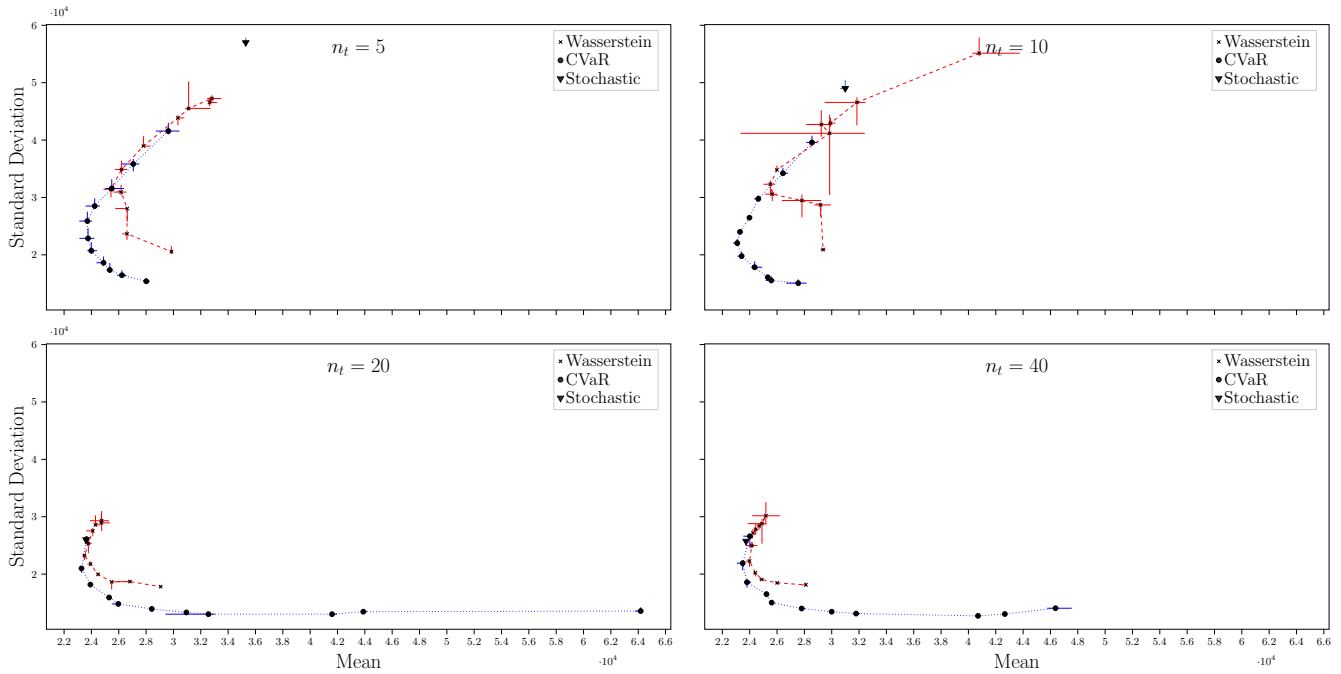

Figure 12: Comparison against Baseline Models on Hydro-thermal Power Planning, Additional Run 2

36