Data-driven Multistage Distributionally Robust Optimization with Nested Distance

Rohit Arora and Rui Gao
Department of Information, Risk and Operations Management, University of Texas at Austin, Austin, TX 78705
arorarohit@utexas.edu, rui.gao@mccombs.utexas.edu

We study multistage distributionally robust linear optimization, where the uncertainty set is a ball of distributions defined through the nested distance (Pflug and Pichler 2012) centered at a scenario tree. This choice of uncertainty set, as opposed to alternatives like the Wasserstein distance between stochastic processes, takes into account information evolution, making it hedge against a plausible family of data processes. Our contributions are two-fold. First, we develop a recursive reformulation to evaluate the worst-case risk of a given policy and related it to the conditional risk mapping with single-period Wasserstein distance. Second, under the stagewise independence assumption, we derive dynamic programming reformulations for finding the optimal robust policy and identify tractable cases when the uncertainty appears in the objective or the right-hand side.

Key words: Multistage stochastic programming, Nested distance, Wasserstein distance, Time consistency

1. Introduction

Distributionally Robust Optimization (DRO) is an emerging paradigm for data-driven decision-making, offering robust solutions that account for data uncertainty. For static problems, much significant progress has been made recently in terms of computation (Delage and Ye 2010, Goh and Sim 2010, Ben-Tal et al. 2013, Wiesemann et al. 2014, Mohajerin Esfahani and Kuhn 2018), regularization (Lam 2016, Duchi and Namkoong 2019, Shafieezadeh-Abadeh et al. 2019, Gao et al. 2022), and statistical guarantees (Lam 2019, Duchi et al. 2021, Blanchet et al. 2019, Gao 2022), etc. However, the landscape for sequential problems remains challenging, with results being both limited and less satisfactory.

In multistage problems, data processes are often depicted using scenario trees, which are constructed based on historical data. Approaches for constructing scenario trees include Monte Carlo sampling techniques such as conditional sampling (Shapiro 2003a,b), which encompasses stagewise independent sampling as a significant case, as well as scenario generation approaches (Dupačová et al. 2000, Høyland and Wallace 2001, Dupačová et al. 2003, Rom, Henrion and Römisch 2022). The scenario trees provide a discrete approximation of the true underlying stochastic process and are used to formulate decision-making problems along the sample paths within the tree. However, it is crucial to recognize that the policies developed based on these scenario trees may not be well-defined for unseen sample paths. In many cases, heuristic policies do not come with optimality guarantees (Ben-Tal et al. 2009, Note and Remarks 14.1). Consequently, when dealing with multistage problems, data scarcity becomes a significant challenge (Shapiro and Nemirovski 2005), highlighting the need for a distributionally robust formulation that can generalize to unseen scenarios. The challenges in multistage DRO stem from both modeling and computational aspects.

1.1. Modeling Challenges

The literature has yet to establish a consensus regarding the formulation of multistage DRO problems. Seemingly natural extensions of the single-stage formulation can result in distinct frameworks, leading to ongoing discussions and research in this area (Pichler and Shapiro 2021, Shapiro and Pichler 2022).

One straightforward formulation, referred to as the *multistage-static formulation*, is a direct extention of the standard single-stage formulation:

$$\inf_{\boldsymbol{x}\in\mathcal{X}} \sup_{\mathbb{P}\in\mathfrak{M}} \mathbb{E}_{\boldsymbol{\xi}_{[T]}\sim\mathbb{P}}\left[\sum_{t=1}^{T} c_t(\boldsymbol{x}_t,\boldsymbol{\xi}_t)\right]. \tag{1}$$

Here $c_t(\boldsymbol{x}_t,\boldsymbol{\xi}_t)$ denotes the per-stage cost associated with a T-stage sample path $\boldsymbol{\xi}_{[T]} := (\xi_1,\boldsymbol{\xi}_2,\ldots,\boldsymbol{\xi}_T)\in\Xi_1\times\cdots\times\Xi_T$ under a policy $\boldsymbol{x}=(x_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T)$[1]. The formulation hedges against an uncertainty set $\mathfrak{M}$ of T-stage stochastic processes and optimizes the worst-case expected cumulative cost over the set $\mathcal{X}$ of policies satisfying some feasibility and non-anticipativity constraints $x_1\in\mathcal{X}_1$, $\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1})$, $t=2,\ldots,T$. Typically, the uncertainty set is constructed based on summary statistics of the stochastic process, such as support and moment information (Bertsimas et al. 2019, Xin and Goldberg 2022), or based on statistical distance such as relative entropy (Hansen and Sargent 2001), Wasserstein distance (Bertsimas et al. 2022, Sturt 2023), and nested distance (Analui and Pflug 2014, Glanzer et al. 2019).

The multistage-static formulation (1) is conceptually simple and offers a clear interpretation as a method to mitigate uncertainty in the data process. However, the coupling of decisions and uncertainties over time in the multistage-static objective is not explicitly adjusted for dynamics of the decision process (Pichler and Shapiro 2021), making it challenging to use dynamic programming recursions and potentially raising concerns about time inconsistency (Iancu et al. 2015), a concept criticized in decision theory for its violation of rational behavior.

An alternative formulation, referred to as the *multistage-dynamic formulation*, has been devised to facilitate dynamic programming recursion and is routinely employed in computational studies. In this formulation, the cost-to-go function takes on a recursive form

$$Q_t(\boldsymbol{x}_{t-1},\boldsymbol{\xi}_{[t]}) = \inf_{\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1})}\left\{c_t(\boldsymbol{x}_t,\boldsymbol{\xi}_t) + \sup_{\mathbb{P}_{t+1}\in\mathfrak{M}_{t+1}} \mathbb{E}_{\mathbb{P}_{t+1}}\left[Q_{t+1}(\boldsymbol{x}_t,\boldsymbol{\xi}_{[t+1]})\right]\right\}, \quad t\in[T], \tag{2}$$

and $Q_{T+1}(\cdot,\cdot)\equiv 0$. Here the uncertainty set $\mathfrak{M}_{t+1}$ can be defined through composite distributionally robust functionals (Shapiro 2016, Pichler and Shapiro 2021), as well as conditional distributionally robust functionals (Shapiro and Pichler 2022) (also known as conditional risk mappings (Ruszczyński and Shapiro 2006)) based on Average Value-at-Risk (AVaR), entropic risk measure ϕ-divergence (Klabjan et al. 2013, Park and Bayraksan 2020, Rahimian et al. 2021), and Wasserstein distance (Shapiro and Pichler 2022). Of particular interest is the stagewise independent setting where these two functionals are equivalent. Common choices of $\mathfrak{M}_{t+1}$ include moment-based sets (Shapiro and Xin 2020, Xin and Goldberg 2021, Yu and Shen 2020) and sets based on statistical distances such as χ^2-divergence (Philpott et al. 2018), L_∞-norm (Huang et al. 2017) and Wasserstein distance (Duque and Morton 2020, Zhang and Sun 2020, 2022).

In comparison to the multistage-static formulation (1), the multistage-dynamic formulation (2) is generally more computationally friendly. Nevertheless, it should be noted that if not appropriately specified, it can be overly conservative and lack interpretability. For instance, the composition of single-period AVaR, also known as iterated conditional tail expectation, takes tail risks of quantities that are already tail risks. This multi-period risk measure does not offer the same straightforward interpretation as AVaR (Shapiro 2012) and can potentially result in overly conservative risk assessments (Iancu et al. 2015). Another example is seen in the formulation with composite distributionally robust functionals, where the worst-case T-stage distribution depends on historical realizations and thus can vary among realizations at each stage. This raises concerns about the pessimism of the resulting policy.

In light of the discussion above, the following question is natural and important, yet remains largely open:

[1] We use bold font for random variables and regular font for deterministic values like constants or elements in the sample space. In line with the convention in stochastic programming literature, we consider the first-stage data $\boldsymbol{\xi}_1$ as deterministic. Therefore, we do not differentiate between x_1 and $\boldsymbol{x}_1$, or between ξ_1 and $\boldsymbol{\xi}_1$. We simply set Ξ_1 as ξ_1.

Q1 : Is there a modeling choice which can be easily interpreted from the viewpoint of the multistage-static formulation (1)*, while simultaneously allowing for an equivalent decomposition into the multistage-dynamic formulation* (2) *with interpretable single-period uncertainty sets?*

Indeed, if the multistage-static formulation (1) can be equivalently represented in a multistage-dynamic form (2), it opens up the possibility of solving (1) through dynamic programming. This would result in a time-consistent robust optimal policy. Conversely, if the multistage-dynamic formulation (2) can be equivalently transformed into a multistage-static form (1) with a natural choice of $\mathfrak{M}$, it would enhance the interpretability of the composite risk measure. Additionally, this transformation could help address the issue of conservativeness by allowing us to work with the more interpretable multistage-static counterpart.

We will affirmatively address Question Q1 by showing that the multistage-static formulation (1) with nested distance is indeed equivalent to the multistage-dynamic formulation (2) with single-period Wasserstein distance. Moreover, the nested distance uncertainty set offers several appealing advantages from a modeling standpoint: (i) It fully utilizes the entire distributional information, distinguishing it from moment-based and risk-measure-based sets that rely only on partial data information like moments and tails. (ii) It provides effective protection against data perturbations that extend beyond the support of the empirical scenario tree. This distinguishes it from divergence-based sets, which impose strict restrictions on the support of relevant distributions (Bayraksan and Love 2015). (iii) It defines a plausible family of stochastic processes, allowing for non-anticipative perturbations of scenario paths. This stands in contrast to the Wasserstein distance, which permits perturbations dependent on future information. Further elaboration on this topic can be found in Section 2.2.

1.2. Computational Challenge

Multistage stochastic programming, even without distributional robustness, is known to be computationally challenging due to the curse of horizon (Shapiro and Nemirovski 2005). Nevertheless, significant progress has been made in recent years for solving linear and convex scenario-based multistage stochastic programming problems by employing cutting-plane approximations of the cost-to-go (value) functions derived from dynamic programming equations (Lan and Shapiro 2023), such as Stochastic Dual Dynamic Programming (SDDP) (Birge 1985, Pereira and Pinto 1991). Since multistage DRO with nested distance reduces to its non-robust counterpart when the radius becomes zero, we hope to leverage these advancements to enhance the computational tractability of the robust counterpart.

The computational challenges in solving (1) with nested distance stem from both the inner robust risk evaluation and the outer policy optimization.

First, the inner robust risk evaluation involves an infinite dimensional optimization over probability distributions. To make this problem more tractable, a common approach is to reformulate it into a finite-dimensional problem using duality and conditioning. However, the nested distance uncertainty set becomes non-convex as soon as $T \geq 3$ (Pflug and Pichler 2014, Section 7.3). The current methodology, as detailed in Pflug and Pichler (2014, Section 7.3.3), exploits a successive programming strategy that iteratively discretizes stochastic processes. Unfortunately, this approach encounters scalability issues, even in stagewise independent settings.

Second, the outer policy optimization involves an infinite-dimensional optimization over policies. Unlike its non-robust scenario approximation counterpart, which only specify policy values on a finite number of sample paths, solving (1) entails decisions at each stage that are functions of all possible realizations of distributions in the uncertainty set. Moreover, previous research (Hanasusanto and Kuhn 2018, Xie 2020) has shown that obtaining the first-stage deterministic decision is generally NP-hard, even in the case of $T = 2$ where (1) with nested distance reduces to a two-stage Wasserstein DRO problem. On the other hand, there have been positive tractability results in specific scenarios of two-stage

Wasserstein DRO. Hanasusanto and Kuhn (2018) provided co-positive program reformulations for 2-Wasserstein DRO with complete recourse and linear program reformulations for 1-Wasserstein DRO with sufficiently expensive recourse. Xie (2020) provides sufficient conditions for ∞-Wasserstein DRO under which the problem admits tractable convex program reformulations.

Given these advances, we would like to answer the following question:

Q2 : *Can we identify conditions under which the formulation* (1) *with nested distance uncertainty set become computationally tractable, at least as tractable as its non-robust counterpart?*

1.3. Our Contributions

We study data-driven multistage distributionally robust linear program (1), where the uncertainty set $\mathfrak{M}$ consists of all stochastic processes within certain nested distance from a scenario tree. Our contributions are as follows.

(I) We derive dynamic programming reformulations for (1), providing a positive answer to Question Q1.

In Section 3.1, we develop dynamic programming reformulation to evaluate the worst-case risk for a fixed policy, in spite of the non-convexity of the nested distance ball. For ∞-nested distance, the inner maximization in (1) can be equivalently decomposed into dynamic programs defined via single-period ∞-Wasserstein balls centered at the nominal conditional distribution. For p-nested distance ($p \in [1, \infty)$), a soft-penalty counterpart of the inner maximization in (1) can be equivalently decomposed into dynamic programs defined via single-period p-Wasserstein penalty centered at the nominal conditional distribution.

Furthermore, in Section 3.2 we demonstrate that, under the stagewise independence assumption, the multistage-static formulation (1) with nested distance is equivalent to the multistage-dynamic formulation (2) with single-period Wasserstein distance. To the best of our knowledge, this is the first non-degenerate uncertainty set that reconciles both static and dynamic formulations for generic multistage linear programs.

(II) We address Question Q2 by deriving computationally tractable dynamic reformulations under objective / right-hand side uncertainty in Section 4.

For objective uncertainty (Section 4.1), the reformulation can be interpreted as norm-regularized scenario approximation problem. It penalizes large norms on the decision variables at every stage, either with a hard constraint ($p = 1$) or a soft penalty $p \in (1, \infty]$.

For right-hand side uncertainty (Section 4.2), the reformulation encourages large norms on the dual variables at every stage. This leads to a solution that regularizes extreme perturbations of the right-hand side when $p = \infty$, and a solution that coincides with the non-robust counterpart formulation when $p = 1$. These results extend the insights of static/two-stage Wasserstein DRO (Mohajerin Esfahani and Kuhn 2018, Shafieezadeh-Abadeh et al. 2019, Gao et al. 2022, Hanasusanto and Kuhn 2018, Xie 2020, Duque et al. 2022) to the multistage setting in a non-trivial manner.

Furthermore, the optimal robust policy is well-defined for every possible sample path of the distributions in the uncertainty set, extending beyond those included in the nominal scenario tree (Corollary 1). In addition, it may not be unique (Proposition 1).

(III) We apply our results to a portfolio selection problem and develop an SDDP algorithm to solve it. The out-of-sample performance of the optimal robust policy, when compared with the non-robust policy, demonstrates the superiority of the robust approach.

1.4. Related Literature

The closest work to ours comes from the monograph Pflug and Pichler (2014, Section 7.3) (see also Analui and Pflug (2014)). Our results differ from theirs in several major ways, which also considers multistage distributionally robust optimization with nested distance. First, their solution approach is based on successive programming that are not scalable and they do not pursue a dynamic programing equivalent reformulation; whereas we develop a general minimax nested dynamic programming formulation and identify computationally tractable cases. Second, they assume the space of sample paths is finite due to computational reasons; whereas we allow it to be a general space while still maintain computational efficiency. Third, they do not discuss the time consistency of the formulation (1); whereas our result clarifies the issue of time consistency as well as statistical consistency. Fourth, both their work and part of our analysis entails a minimax theorem that is proved based on certain convex relaxation whose (approximately) worst-case distribution is contained in the original nested distance ball, however, our convex relaxation is based on relaxing the non-anticipativity constraints, which is different from their construction of convex hull based on compounding finite trees. Below, we review some relevant literature other than those above-mentioned works.

On nested distance. Nested distance was first introduced and studied in Pflug (2010), Pflug and Pichler (2012) in stochastic programming literature. Since then, it has been applied to scenario generation/reduction and approximation of multistage stochastic programming (Pflug and Pichler 2015, Maggioni and Pflug 2016, Kovacevic and Pichler 2015, Chen and Yan 2018, Horejšová et al. 2020). Its statistical properties (Pflug and Pichler 2016, Glanzer et al. 2019, Veraguas et al. 2020) and computational properties (Cabral and da Costa 2017, Pichler and Weinhardt 2022) have also been investigated thoroughly. The idea of imposing non-anticipativity constraints on the transport plan can be traced back to the Yamada-Watanabe criterion for stochastic differential equations (Yamada and Watanabe 1971)as well as the causal transportation in continuous time (Lassalle 2018) and in discrete time (Backhoff et al. 2017). In the optimal transport and mathematical finance literature, the nested distance is also called bi-causal transport distance or adapted Wasserstein distance (Backhoff-Veraguas et al. 2020, Backhoff et al. 2022). Incorporating nested distance in multistage distributionally robust optimization was first considered in Analui and Pflug (2014), Pflug and Pichler (2014) and then in Glanzer et al. (2019) for a pricing problem. Unlike the algorithmic approach in these works, our algorithm is more computationally friendly, and enjoy similar tractability as the SAA counterpart. A recent paper (Yang et al. 2022) studies DRO with causal transport distance, which can be viewed as a convex relaxation of our problem in three stages.

On time consistency. Various concepts of time consistency have been discussed in economics literature (Strotz 1955, Hansen and Sargent 2001, Epstein and Schneider 2003, Etner et al. 2012), in mathematical finance literature (Wang 1999, Föllmer and Schied 2002, Artzner et al. 2007, Roorda and Schumacher 2007, Cheridito and Kupper 2009) and in robust control literature (Iyengar 2005, Nilim and El Ghaoui 2005, Wiesemann et al. 2013). Our notion of time consistency is more aligned with the stochastic programming literature (Ruszczyński and Shapiro 2006, Shapiro 2012, 2016, Shapiro and Xin 2020, Xin and Goldberg 2021, Pichler et al. 2021). In general, the multistage-static formulation (1) does not have a time-consistent optimal policy due to a lack of dynamic programming representation (Pichler and Shapiro 2021). In fact, the only known examples of $\mathfrak{M}$ that lead to a tractable dynamic nested risk measure representation (2) are two degenerate ones (Shapiro et al. 2021, Remark 33): singleton (corresponding to the conditional expectation or risk neutral) and entirety (corresponding to max-risk measures, or most risky). For certain classes of problems in inventory control, it has been shown that the multistage-static formulation (1) with some moment-based uncertainty sets has a time-consistent optimal policy under certain conditions (Shapiro 2012, Xin and Goldberg 2021, 2022). In contrast, our result on time consistency holds for generic multistage linear programs.

On computation of multistage DRO with transport distance based sets. There are several computational works directly solving the multistage-dynamic formulation (2) instead of working with (1). For formulation (2) with 1-Wasserstein set, Duque and Morton (2020) proposed a stochastic dual dynamic programming algorithm which restricts the support of scenarios on a pre-specified finite set including the empirical support and additional "tail" scenarios. Yet, it is not entirely clear on how to choose these extreme scenarios a priori. This issue is mitigated in a recent work by Zhang and Sun (2022). We consider (1) with nested distance, aiming at tackling the its non-convexity as observed in Pflug and Pichler (2014). We would like to emphasize that, in this paper, we are content with a reformulation that is as tractable as multistage SAA, although the latter alone has many computational issues (Shapiro et al. 2021, Chapter 5.8). These issues are beyond the scope of this paper.

The rest of the paper is organized as follows. In Section 2, we introduce the distributionally robust linear multistage program, and provide a quick overview on the nested distance used for constructing the ambiguity set. In Section 3, we first develop a general dynamic programming reformulation of the multistage-static problem, and then discuss its implications on statistical consistency and time consistency. In Section 4, we specialize our result to three cases that admit tractable reformulations. In Section 5, we apply our result to the portfolio selection problem and develop a stochastic dual dynamic programming algorithm and out-of-sample testing procedure. We conclude the paper in Section 6. All proofs are deferred to the Appendices.

2. Multistage Distributionally Robust Optimization with Nested Distance

In this section, we discuss the nested distance via examples and present the distributionally robust formulation.

2.1. Multistage Stochastic Programming and its Distributionally Robust Counterpart

Consider a T-stage stochastic linear optimization problem

$$\begin{aligned} \min_{x_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T} \quad & \mathbb{E}_{\mathbb{P}}\left[c_1^\top x_1 + \boldsymbol{c}_2^\top \boldsymbol{x}_2 + \cdots + \boldsymbol{c}_T^\top \boldsymbol{x}_T\right], \\ s.t. \quad & A_1 x_1 = b_1,\ x_1 \geq 0, \\ & \boldsymbol{B}_t \boldsymbol{x}_{t-1} + \boldsymbol{A}_t \boldsymbol{x}_t = \boldsymbol{b}_t,\ \boldsymbol{x}_t \geq 0, \quad t = 2, \ldots, T, \end{aligned}$$

where $\boldsymbol{\xi}_t := (\boldsymbol{A}_t, \boldsymbol{B}_t, \boldsymbol{c}_t, \boldsymbol{b}_t) \in \Xi_t \subset \mathbb{R}^{d_t}$, $t \in [T] := \{1, \ldots, T\}$, are data vectors and matrices, some or all of which may be random. For $t \in [T]$, we denote by $\boldsymbol{\xi}_{[t]} := (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_t) \in \Xi_{[t]} := \Xi_1 \times \cdots \Xi_t$ the history of the data process up to time t. Let $\mathcal{P}(\Xi_{[t]})$ be the set of probability distributions on $\Xi_{[t]}$, $t \in [T]$. For any distribution $\mathbb{P} \in \mathcal{P}(\Xi_{[T]})$ of a stochastic process $\boldsymbol{\xi}_{[T]}$, we denote by $\mathbb{P}_{[t]}$ the marginal distribution of $\boldsymbol{\xi}_{[t]}$ under $\mathbb{P}$, and by $\mathbb{P}_{t|\boldsymbol{\xi}_{[t-1]}}$ the conditional distribution of $\boldsymbol{\xi}_t$ given the history $\boldsymbol{\xi}_{[t-1]}$. The minimization is performed over the set of non-anticipative policies $\boldsymbol{x}_t = \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})$, each of which is measurable with respect to $\sigma(\boldsymbol{\xi}_{[t]})$, the σ-algebra induced by $\boldsymbol{\xi}_{[t]}$. To ease notations, we denote the feasible regions

$$\begin{aligned} \mathcal{X}_1 &:= \{x_1 \geq 0 :\ A_1 x_1 = b_1\}, \\ \mathcal{X}_t(x_{t-1}, \boldsymbol{\xi}_t) &:= \{x_t \geq 0 :\ \boldsymbol{A}_t x_t = \boldsymbol{b}_t - \boldsymbol{B}_t x_{t-1}\}, \quad t = 2, \ldots, T. \end{aligned}$$

We will always assume $\mathcal{X}_1 \neq \emptyset$. The multistage problem above admits a dynamic programming formulation

$$Q_t(x_{t-1}, \boldsymbol{\xi}_{[t]}) = \min_{x_t \in \mathcal{X}_t(x_{t-1}, \boldsymbol{\xi}_t)} \left\{ c_t^\top x_t + \mathbb{E}_{\mathbb{P}_{t+1|\boldsymbol{\xi}_{[t]}}}[Q_{t+1}(x_t, \boldsymbol{\xi}_{[t+1]})] \right\}, \quad t \in [T],$$

with $Q_{T+1} \equiv 0$ and $\mathcal{X}_1(x_0, \xi_1) \equiv \mathcal{X}_1$. We represent the set of non-anticipative and feasible policy as

$$\mathcal{X} := \left\{ \left(x_1, \boldsymbol{x}_2(\cdot), \ldots, \boldsymbol{x}_T(\cdot)\right) :\ \boldsymbol{x}_t \in \mathcal{X}_t(\boldsymbol{x}_{t-1}),\ t \in [T] \right\}.$$

Here the shorthand notation $\boldsymbol{x}_t \in \mathcal{X}_t(\boldsymbol{x}_{t-1})$ is interpreted as $\boldsymbol{x}_t(\xi_{[t]}) \in \mathcal{X}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}), \xi_t)$ for all $\xi_{[t]} \in \Xi_{[t]}$.

Quite often in practice, the data-generating distribution of the random process $\boldsymbol{\xi}_{[T]}$ is not known exactly. A common approach is to replace the underlying data-generating distribution with a scenario tree $\widehat{\mathbb{P}}$, which is typically constructed using conditional sampling or scenario reduction. Let us denote by $\widehat{\boldsymbol{\xi}}_{[T]}$ the stochastic process with a finitely-supported distribution $\widehat{\mathbb{P}}$ and by $\widehat{\Xi}_t$ the support of $\widehat{\boldsymbol{\xi}}_t$, $t \in [T]$. To account for the distributional uncertainty, we consider the following multistage distributionally robust optimization

$$\inf_{\boldsymbol{x}\in\mathcal{X}} \sup_{\mathbb{P}\in\mathfrak{M}} \mathbb{E}_{\mathbb{P}}\left[\sum_{t\in[T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\right], \tag{$\mathrm{P_{static}}$}$$

and the distributional uncertainty set $\mathfrak{M}$ specifies a set of T-stage distributions to hedge against. In particular, we consider the following uncertainty set

$$\mathfrak{M} := \left\{\mathbb{P} \in \mathcal{P}(\Xi_{[T]}) : \mathsf{D}_p(\widehat{\mathbb{P}}, \mathbb{P}) \le \vartheta\right\}, \tag{3}$$

where $\vartheta > 0$ is the radius of the uncertainty set, and D_p is the *p-nested distance* proposed by Pflug (2010), Pflug and Pichler (2012), which takes account of the information evolution in the multistage problem, as will be elaborated on in Section 2.2. We will also consider a soft robust formulation when $p \in [1, \infty)$:

$$\inf_{\boldsymbol{x}\in\mathcal{X}} \sup_{\mathbb{P}\in\mathcal{P}(\Xi_{[T]})} \left\{\mathbb{E}_{\mathbb{P}}\left[\sum_{t\in[T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\right] - \lambda \mathsf{D}_p^p(\widehat{\mathbb{P}}, \mathbb{P})\right\}. \tag{$\mathrm{P_{static}}$-soft}$$

2.2. Nested Distance

The nested distance calculates the minimum cost needed to transport probability mass from $\widehat{\mathbb{P}}$ to $\mathbb{P}$ among a set of non-anticipative transport plans. Similar to the Wasserstein distance, it is based on an optimal transport problem. But in addition to the marginal constraints on the transport plan as in the definition of Wasserstein distance, it also requires that transport plan should be non-anticipative with respect to the filtration $\sigma(\widehat{\boldsymbol{\xi}}_{[t]}) \otimes \sigma(\boldsymbol{\xi}_{[t]})$.

Let $\mathsf{d}(\cdot,\cdot)$ be a metric on $\Xi_{[T]}$. For any $\widehat{\mathbb{P}}, \mathbb{P} \in \mathcal{P}(\Xi_{[T]})$, we denote by $\Gamma(\widehat{\mathbb{P}}, \mathbb{P})$ the set of joint distributions on $\Xi_{[T]}^2$ with marginals $\widehat{\mathbb{P}}$ and $\mathbb{P}$. For a joint distribution $\gamma \in \Gamma(\widehat{\mathbb{P}}, \mathbb{P})$, we use $\gamma_{\widehat{\xi}_t|(\widehat{\xi}_{[t-1]},\xi_{[t-1]})}$ to denote the conditional distribution of $\widehat{\boldsymbol{\xi}}_t$ given $(\widehat{\boldsymbol{\xi}}_{[t-1]}, \boldsymbol{\xi}_{[t-1]})$ under γ. Recall that $\widehat{\mathbb{P}}_{t|\widehat{\xi}_{[t-1]}}$ denotes the conditional distribution of $\widehat{\boldsymbol{\xi}}_t$ given $\widehat{\boldsymbol{\xi}}_{[t-1]}$ under $\widehat{\mathbb{P}}$.

Definition 1 (Nested Distance). Define the set of transport plans

$$\Gamma_{bc}(\widehat{\mathbb{P}}, \mathbb{P}) = \left\{\gamma \in \Gamma(\widehat{\mathbb{P}}, \mathbb{P}) : \gamma_{\widehat{\xi}_{t+1}|(\widehat{\xi}_{[t]},\xi_{[t]})} = \widehat{\mathbb{P}}_{t+1|\widehat{\xi}_{[t]}},\ \gamma_{\xi_{t+1}|(\widehat{\xi}_{[t]},\xi_{[t]})} = \mathbb{P}_{t+1|\xi_{[t]}},\ \forall t = 1, \ldots, T-1\right\}. \tag{4}$$

The nested distance $\mathsf{D}_p(\widehat{\mathbb{P}}, \mathbb{P})$ between $\widehat{\mathbb{P}}$ and $\mathbb{P}$ is defined as

$$\mathsf{D}_p(\widehat{\mathbb{P}}, \mathbb{P}) := \begin{cases} \left(\inf_{\gamma\in\Gamma_{bc}(\widehat{\mathbb{P}},\mathbb{P})} \mathbb{E}_{(\widehat{\xi}_{[T]},\xi_{[T]})\sim\gamma}\left[\mathsf{d}(\widehat{\boldsymbol{\xi}}_{[T]}, \boldsymbol{\xi}_{[T]})^p\right]\right)^{1/p}, & p \in [1, \infty), \\ \inf_{\gamma\in\Gamma_{bc}(\widehat{\mathbb{P}},\mathbb{P})} \ \gamma\text{-}\operatorname{ess\,sup}_{(\widehat{\xi}_{[T]},\xi_{[T]})\in\Xi^2} \ \mathsf{d}(\widehat{\boldsymbol{\xi}}_{[T]}, \boldsymbol{\xi}_{[T]}), & p = \infty. \end{cases} \tag{5}$$

◇

The non-anticipativity constraints can be equivalently stated as follows. Under the joint distribution γ,

$$\boldsymbol{\xi}_{[t]} \perp \widehat{\boldsymbol{\xi}}_{t+1} \mid \widehat{\boldsymbol{\xi}}_{[t]}, \tag{6a}$$

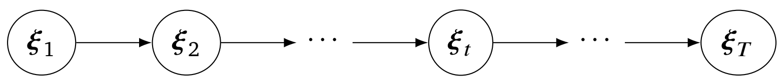
$$\widehat{\boldsymbol{\xi}}_{[t]} \perp \boldsymbol{\xi}_{t+1} \mid \boldsymbol{\xi}_{[t]}, \tag{6b}$$

namely, $\boldsymbol{\xi}_{[t]}$ and $\widehat{\boldsymbol{\xi}}_{t+1}$ are *conditionally independent* given $\widehat{\boldsymbol{\xi}}_{[t]}$; and $\widehat{\boldsymbol{\xi}}_{[t]}$ and $\boldsymbol{\xi}_{t+1}$ are conditionally independent given $\boldsymbol{\xi}_{[t]}$. Suppose there exists a transport map $\mathbb{T} = (\mathbb{T}_1, \ldots, \mathbb{T}_T)$ from $\widehat{\mathbb{P}}$ to $\mathbb{P}$, then (6a) implies that where $\widehat{\boldsymbol{\xi}}_{[t]}$ is transported (i.e., $\boldsymbol{\xi}_{[t]} = \mathbb{T}_{[t]}(\widehat{\boldsymbol{\xi}}_{[T]})$) is independent of the future $\widehat{\boldsymbol{\xi}}_{t+1}$. Thereby, $\mathbb{T}$ satisfies (6a) if and only if it is of the form $\mathbb{T}(\widehat{\xi}_{[T]}) = (\mathbb{T}_1(\widehat{\xi}_1), \mathbb{T}_2(\widehat{\xi}_{[2]}), \ldots, \mathbb{T}_T(\widehat{\xi}_{[T]}))$, $\forall \widehat{\xi}_{[T]} \in \widehat{\Xi}$. Similarly, the condition (6b) indicates that where $\boldsymbol{\xi}_{[t]}$ is transported should not dependent on future information $\boldsymbol{\xi}_{t+1}$. Thereby, if $\mathbb{T}_t$, $t = 1, \ldots, T$, are invertible, then $\mathbb{T}$ satisfies (6b) as well (Backhoff et al. 2017). The equivalent definition (6) provides a convient way to check whether a transport plan is causal or not.

In the literature, a transport plan from $\widehat{\mathbb{P}}$ to $\mathbb{P}$ is termed *causal* if it satisfies the non-anticipativity constraint (6a), and is termed *bi-causal* if it satisfies both non-anticipativity constraints (6). When only the first set of constraints in (4) is imposed, the resulting distance is called *causal transport distance* (Backhoff et al. 2017), denoted as $\mathsf{C}_p(\widehat{\mathbb{P}}, \mathbb{P})$. If we replace Γ_{bc} by Γ, then (5) becomes the defining expression for the Wasserstein distance. Note that $\Gamma_{bc}(\widehat{\mathbb{P}}, \mathbb{P})$ is always a non-empty subset of $\Gamma(\widehat{\mathbb{P}}, \mathbb{P})$, containing at least the independent transport plan, namely the product distribution with marginals $\widehat{\mathbb{P}}$ and $\mathbb{P}$.

Let us illustrate these concepts with the following examples.

Example 1. In many applications, the data process $\boldsymbol{\xi}_{[T]}$ often adheres to a causal relationship, as represented by the following causal diagram



This relationship can be observed in various scenarios, such as the demand process for a product or the return rate process of financial assets. The data uncertainty of $\boldsymbol{\xi}_t$ can arise directly from errors like sampling or measurement inaccuracies in $\boldsymbol{\xi}_t$. Alternatively, it might be indirectly influenced by errors in statistical modeling or data processing propagated from historical data $\boldsymbol{\xi}_{[t-1]}$. Consequently, it is logical to consider data perturbations that exhibit historical dependencies. However, such perturbations should not be dependent on future uncertainties, as these are typically unknown. This rationale provides a justification for the non-anticipativity constraints in Definition 1. ♣

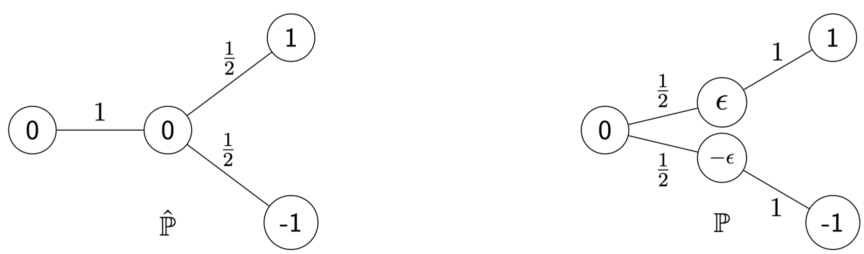


Figure 1 The p-Wasserstein distance between $\widehat{\mathbb{P}}$ and $\mathbb{P}$ is ϵ whereas the nested distance $\mathsf{D}_1(\widehat{\mathbb{P}}, \mathbb{P}) = 1 + \epsilon$.

Example 2. Consider two stochastic processes $\widehat{\mathbb{P}}$ and $\mathbb{P}$ represented by the two scenario trees plotted in Figure 1. The two processes $\widehat{\mathbb{P}}$ and $\mathbb{P}$ have different evolution of information. For $\widehat{\mathbb{P}}$, conditional on observing $\widehat{\boldsymbol{\xi}}_2 = 0$, $\widehat{\boldsymbol{\xi}}_3$ takes ± 1 with equal probability; while for $\mathbb{P}$ with any $\epsilon > 0$, conditional on observing $\boldsymbol{\xi}_2$, the value of $\boldsymbol{\xi}_3$ is certain: $\mathbb{P}_{\xi_3 = 1 | \xi_2 = \epsilon} = \mathbb{P}_{\widehat{\xi}_3 = -1 | \widehat{\xi}_2 = -\epsilon} = 1$.

Suppose $\mathsf{d}(\widehat{\xi}_{[3]}, \xi_{[3]}) = \|\widehat{\xi}_{[3]} - \xi_{[3]}\|_p$. Then we have $\mathcal{W}_p(\widehat{\mathbb{P}}, \mathbb{P}) = \epsilon$ for all $p \in [1, \infty]$. On the other hand, a causal transport plan γ from $\widehat{\mathbb{P}}$ to $\mathbb{P}$ should satisfy $\mathbb{P}_{\widehat{\xi}_3 | \widehat{\xi}_2 = 0} = \frac{1}{2} = \gamma_{\widehat{\xi}_3 | \widehat{\xi}_2 = 0, \xi_2}$. Hence the only feasible transport plan is the independent product distribution $\gamma(\widehat{\boldsymbol{\xi}}_{[3]}, \boldsymbol{\xi}_{[3]}) = \frac{1}{4}$ for all pairs of $\widehat{\boldsymbol{\xi}}_{[3]}, \boldsymbol{\xi}_{[3]}$, which is in fact bi-causal. Thus we have

$$\mathsf{D}_p(\widehat{\mathbb{P}}, \mathbb{P}) = 2^{-1/p}(\epsilon^p + (2 + \epsilon)^p)^{1/p} \geq 1 + \epsilon/2.$$

As such, the nested-distance ball centered at $\widehat{\mathbb{P}}$ with radius 1 would not contain $\mathbb{P}$ with any $\epsilon > 0$.

When forming an uncertainty set $\mathfrak{M}$, it makes sense to consider stochastic processes that share a similar filtration as the nominal data process. If we define $\mathfrak{M}$ based on a Wasserstein ball centered at $\widehat{\mathbb{P}}$, then it would always contain a scenario tree whose structure is similar to $\mathbb{P}$ but significantly different from $\widehat{\mathbb{P}}$. In contrast, a nested distance ball can rule out this undesirable situation. ♣

EXAMPLE 3. In Figure 2 we plot two three-stage scenario trees, with labels along the edges indicating the conditional probabilities of realizing a scenario given their parent nodes. Specifically, define the sample paths

$$\widehat{\xi}^1_{[3]} = (\hat{a}, \hat{b}, \hat{d}),\ \widehat{\xi}^2_{[3]} = (\hat{a}, \hat{b}, \hat{e}),\ \widehat{\xi}^3_{[3]} = (\hat{a}, \hat{c}, \hat{f}),$$
$$\xi^1_{[3]} = (a, b, d),\ \xi^2_{[3]} = (a, c, e),\ \xi^3_{[3]} = (a, c, f),$$

where a sample path is represented by a triple of nodes. Then the two trees represent probability distributions on the three sample paths

$$\widehat{\mathbb{P}} = \frac{1}{6}\boldsymbol{\delta}_{\widehat{\xi}^1_{[3]}} + \frac{1}{6}\boldsymbol{\delta}_{\widehat{\xi}^2_{[3]}} + \frac{2}{3}\boldsymbol{\delta}_{\widehat{\xi}^3_{[3]}}, \quad \mathbb{P} = \frac{1}{2}\boldsymbol{\delta}_{\xi^1_{[3]}} + \frac{1}{4}\boldsymbol{\delta}_{\xi^2_{[3]}} + \frac{1}{4}\boldsymbol{\delta}_{\xi^3_{[3]}},$$

where $\boldsymbol{\delta}_\xi$ indicates a Dirac mass at a sample path ξ.

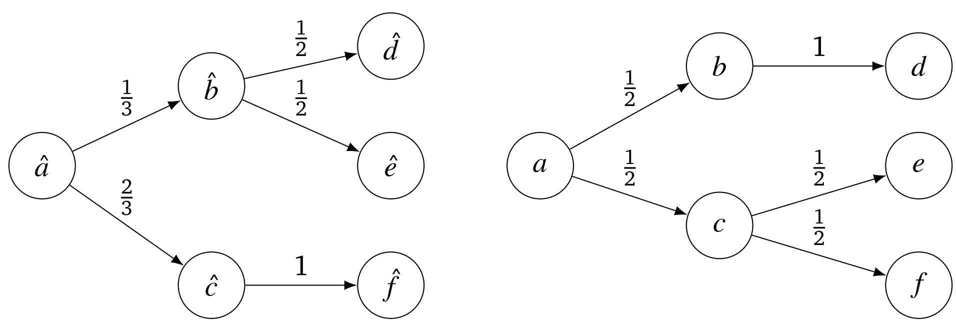


Figure 2 Two three-stage scenario trees

Consider the following three transport plans between the two trees, represented by a joint distribution with marginals $\widehat{\mathbb{P}}$ and $\mathbb{P}$:

$$\gamma = \begin{pmatrix} 1/6 & 0 & 0 \\ 0 & 1/6 & 0 \\ 1/3 & 1/12 & 1/4 \end{pmatrix}, \quad \gamma_c = \begin{pmatrix} 1/24 & 1/8 & 0 \\ 1/24 & 1/8 & 0 \\ 5/12 & 0 & 1/4 \end{pmatrix}, \quad \gamma_{bc} = \begin{pmatrix} 1/24 & 1/16 & 1/16 \\ 1/24 & 1/16 & 1/16 \\ 5/12 & 1/8 & 1/8 \end{pmatrix}.$$

In each matrix, its element in the i-th row and the j-th column represents the probability mass transported from the path $\widehat{\xi}^i_{[3]}$ to the path $\xi^j_{[3]}$, $i, j = 1, 2, 3$. We have the following observations.

(I) The transport plan γ is not causal, as both causal constraints in (6) are violated. Indeed, $\widehat{\boldsymbol{\xi}}_2 = \widehat{b}$ is transported to b if $\widehat{\boldsymbol{\xi}}_3 = \hat{d}$ and to c if $\widehat{\boldsymbol{\xi}}_3 = \hat{e}$. This means that the future value of $\widehat{\boldsymbol{\xi}}_3$ affects the value of $\boldsymbol{\xi}_2$. On the other hand, $\boldsymbol{\xi}_2 = c$ is transported to $\hat{c}$ if $\boldsymbol{\xi}_3 = f$ and is split to $\hat{c}$ and $\hat{b}$ if $\boldsymbol{\xi}_3 = e$.

(II) The transport plan γ_c is causal from $\widehat{\mathbb{P}}$ to $\mathbb{P}$, but not from $\mathbb{P}$ to $\widehat{\mathbb{P}}$. Indeed, regardless of the value of $\widehat{\boldsymbol{\xi}}_3$, $\hat{b}$ splits $1/24$ probabilty mass to b and $1/8$ probability mass to c, and $\hat{c}$ splits $5/12$ probability mass to b and $1/4$ probability mass to c. On the other hand, (6b) is violated, as $\boldsymbol{\xi}_2 = c$ is transported to $\hat{b}$ if $\boldsymbol{\xi}_3 = e$ and to $\hat{c}$ if $\boldsymbol{\xi}_3 = f$.

(III) The transport plan γ_{bc} is bi-causal. Indeed, regardless of the value of $\widehat{\boldsymbol{\xi}}_3$, $\hat{b}$ splits $1/24$ probabilty mass to b and $1/8$ probability mass to c, and $\hat{c}$ splits $5/12$ probability mass to b and $1/4$ probability mass to c. Furthermore, regardless of the value of $\boldsymbol{\xi}_3$, b splits $1/12$ probabilty mass to $\hat{b}$ and $5/12$ probability mass to c, and c splits $1/8$ probability mass to $\hat{b}$ and $1/8$ probability mass to $\hat{c}$. ♣

In the rest of this paper, we will assume Ξ_t is a non-empty subset of some normed space, $t \in [T]$, and set

$$\mathsf{d}(\widehat{\xi}_{[T]}, \xi_{[T]}) = \begin{cases} \left(\sum_{t\in[T]} \|\widehat{\xi}_t - \xi_t\|^p \right)^{1/p}, & p \in [1,\infty), \\ \max_{t\in[T]} \|\widehat{\xi}_t - \xi_t\|, & p = \infty, \end{cases} \quad \forall\, \widehat{\xi}_{[T]}, \xi_{[T]} \in \Xi_{[T]},$$

where $\|\cdot\|$ is some norm defined on the corresponding spaces, and we use overloaded notations

$$\mathsf{d}(\widehat{\xi}_{[t]}, \xi_{[t]}) = \sum_{s\in[t]} \|\widehat{\xi}_s - \xi_s\|, \quad \forall\, \widehat{\xi}_{[t]}, \xi_{[t]} \in \Xi_{[t]},\ t \in [T].$$

Note that the norm in each stage can be chosen differently, but we omit such dependence for the ease of notation.

3. Dynamic Programming Reformulations

In this section, we develop dynamic programming reformulations for ($\mathsf{P}_{\text{static}}$). In Section 3.1, we focus on the inner maximization of ($\mathsf{P}_{\text{static}}$), which evaluates the worst-case risk of a fixed policy, and we discuss the outer minimization over policies in Section 3.2.

3.1. Robust Risk Evaluation

We first consider $p = \infty$. The following theorem provides a dynamic programming equivalent reformulation for the inner supremum in ($\mathsf{P}_{\text{static}}$).

THEOREM 1. *Let $p = \infty$. Then for any continuous policy $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$, it holds that*

$$\begin{aligned} &\sup_{\mathbb{P}\in\mathfrak{M}} \mathbb{E}_{\mathbb{P}}\left[\sum_{t\in[T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\right] \\ &= c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\left[\sup_{\xi_2\in\Xi_2: \|\xi_2-\widehat{\xi}_2\|\le\vartheta} \left\{ c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_{3|\widehat{\xi}_2}}\left[\sup_{\xi_3\in\Xi_3: \|\xi_3-\widehat{\xi}_3\|\le\vartheta} \left\{ c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots \right.\right.\right.\right. \\ &\qquad\qquad \left.\left.\left.\left. + \mathbb{E}_{\widehat{\mathbb{P}}_{T|\widehat{\xi}_{[T-1]}}}\left[\sup_{\xi_T\in\Xi_T: \|\xi_T-\widehat{\xi}_T\|\le\vartheta} c_T^\top \boldsymbol{x}_T(\xi_{[T]})\right] \cdots \right\}\right]\right\}\right]. \end{aligned} \tag{7}$$

Theorem 1 shows that ($\mathsf{P}_{\text{static}}$) admits a nested form. Note that (7) can also be written in the form of dynamic programming. Set $V_{T+1}^{\widehat{\xi}_{[T+1]}} \equiv 0$ and for $t \in [T]$, set

$$V_t^{\widehat{\xi}_{[t]}}(\boldsymbol{\xi}_{[t]}) := \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) + \mathbb{E}_{\widehat{\boldsymbol{\xi}}_{t+1}\sim\widehat{\mathbb{P}}_{t+1|\widehat{\xi}_{[t]}}}\left[\sup_{\xi_{t+1}\in\Xi_{t+1}: \|\xi_{t+1}-\widehat{\boldsymbol{\xi}}_{t+1}\|\le\vartheta} V_{t+1}^{\widehat{\boldsymbol{\xi}}_{[t+1]}}(\boldsymbol{\xi}_{[t]}, \xi_{t+1})\right]. \tag{8}$$

Then (7) implies that

$$\sup_{\mathbb{P}\in\mathfrak{M}} \mathbb{E}_{\mathbb{P}}\left[\sum_{t\in[T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\right] = V_1^{\widehat{\xi}_1}(\xi_1) =: V_1.$$

The risk-to-go function $V_t^{\widehat{\xi}_{[t]}}(\cdot)$ is defined with respect to the nominal realization $\widehat{\boldsymbol{\xi}}_{[t]}$. It assesses the current cost along with the next-stage risk relative to the nominal conditional distribution $\widehat{\mathbb{P}}_{t+1|\widehat{\xi}_{[t]}}$. More specifically, the risk $V_t^{\widehat{\xi}_{[t]}}(\boldsymbol{\xi}_{[t]})$ at stage t is broken down into two components: (i) the current-stage cost, $c_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})$; (ii) the risk-to-go, which evaluates the worst-case value of the risk function $V_{t+1}^{\widehat{\xi}_{[t+1]}}(\boldsymbol{\xi}_{[t]}, \cdot)$ in a ϑ-neighborhood of $\widehat{\boldsymbol{\xi}}_{t+1}$, and then averaged over the nominal conditional distribution

$\widehat{\boldsymbol{\xi}}_{t+1} \sim \widehat{\mathbb{P}}_{t|\widehat{\boldsymbol{\xi}}_{[t]}}$. If $\vartheta = 0$, then (8) reduces to the standard Bellman recursion. When $T = 2$, the two-stage DRO with ∞-nested distance becomes the two-stage DRO with ∞-Wasserstein distance, and the result in Theorem 1 is consistent with Xie (2020, Theorem 2).

The proof idea of Theorem 1 can be summarized as follows. Due to the non-convexity of the nested distance uncertainty set, directly dualizing the inner supremum in ($\mathsf{P}_{\text{static}}$) is not a viable approach. To overcome this challenge, we first consider a convex relaxation of ($\mathsf{P}_{\text{static}}$), replacing the nested distance uncertainty set with the causal distance uncertainty set $\mathfrak{M}^{\mathsf{C}} := \{\mathbb{P} \in \mathcal{P}(\Xi) : \mathsf{C}_\infty(\widehat{\mathbb{P}}, \mathbb{P}) \le \vartheta\}$. Using induction and the tower property of conditional expectation, we are able to derive a dynamic programming reformulation for the relaxed problem. Next, we show that this convex relaxation is, in fact, tight. By modifying the worst-case distribution within the causal distance ball, it is shown that there exists a distribution whose nested distance to $\widehat{\mathbb{P}}$ is approximately equal to the causal distance, and moreover, it yields an objective value that is approximately equal to worst-case risk over the causal distance ball. Thereby, the expression holds for the nested distance ball as well. A complete proof can be found in Appendix EC.1.1.

Next, we consider $p \in [1, \infty)$. We have the following result, whose proof can be found in Appendix EC.1.2.

THEOREM 2. *Let $p \in [1, \infty)$. Then for any continuous policy $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)$, it holds that*

$$
\begin{aligned}
&\sup_{\mathbb{P}\in\mathcal{P}(\Xi)} \left\{ \mathbb{E}_{\mathbb{P}}\Big[\sum_{t\in[T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) \Big] - \lambda \mathsf{D}_p^p(\widehat{\mathbb{P}}, \mathbb{P}) \right\} \\
&= c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\left[\sup_{\xi_2\in\Xi_2} \left\{ c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_{3|\widehat{\xi}_2}}\left[\sup_{\xi_3\in\Xi_3} \left\{ c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots \right.\right.\right.\right. \\
&\qquad \left.\left.\left.\left. + \mathbb{E}_{\widehat{\mathbb{P}}_{T|\widehat{\xi}_{[T-1]}}}\left[\sup_{\xi_T\in\Xi_T} \left\{ c_T^\top \boldsymbol{x}_T(\xi_{[T]}) - \lambda\|\xi_T - \widehat{\boldsymbol{\xi}}_T\|^p \right\} \right] \cdots - \lambda\|\xi_3 - \widehat{\boldsymbol{\xi}}_3\|^p \right\} \right] - \lambda\|\xi_2 - \widehat{\boldsymbol{\xi}}_2\|^p \right\} \right].
\end{aligned} \tag{9}
$$

Moreover, it holds that

$$
\begin{aligned}
&\sup_{\mathbb{P}\in\mathfrak{M}} \mathbb{E}_{\mathbb{P}}\Big[\sum_{t\in[T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) \Big] \\
&= \min_{\lambda\ge 0} \left\{ \lambda\vartheta^p + c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\left[\sup_{\xi_2\in\Xi_2} \left\{ c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_{3|\widehat{\xi}_2}}\left[\sup_{\xi_3\in\Xi_3} \left\{ c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots \right.\right.\right.\right.\right. \\
&\qquad \left.\left.\left.\left.\left. + \mathbb{E}_{\widehat{\mathbb{P}}_{T|\widehat{\xi}_{[T-1]}}}\left[\sup_{\xi_T\in\Xi_T} \left\{ c_T^\top \boldsymbol{x}_T(\xi_{[T]}) - \lambda\|\xi_T - \widehat{\boldsymbol{\xi}}_T\|^p \right\} \right] \cdots - \lambda\|\xi_3 - \widehat{\boldsymbol{\xi}}_3\|^p \right\} \right] - \lambda\|\xi_2 - \widehat{\boldsymbol{\xi}}_2\|^p \right\} \right] \right\}.
\end{aligned}
$$

Theorem 2 establishes the nested reformulations of both (1) and its soft variant. The reformulation (9) of the soft problem is often convenient to work with. Indeed, set $V_{T+1}^{\widehat{\boldsymbol{\xi}}_{[T+1]}} \equiv 0$ and for $t \in [T]$, set

$$
V_t^{\widehat{\boldsymbol{\xi}}_{[t]}}(\boldsymbol{\xi}_{[t]}) := \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) + \mathbb{E}_{\widehat{\boldsymbol{\xi}}_{t+1}\sim\widehat{\mathbb{P}}_{t+1|\widehat{\boldsymbol{\xi}}_{[t]}}}\left[\sup_{\xi_{t+1}\in\Xi_{t+1}} \left\{ V_{t+1}^{\widehat{\boldsymbol{\xi}}_{[t+1]}}(\boldsymbol{\xi}_{[t]}, \xi_{t+1}) - \lambda\|\xi_{t+1} - \widehat{\boldsymbol{\xi}}_{t+1}\|^p \right\} \right]. \tag{10}
$$

Then (9) implies that

$$
\sup_{\mathbb{P}\in\mathcal{P}(\Xi)} \left\{ \mathbb{E}_{\mathbb{P}}\Big[\sum_{t\in[T]} \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) \Big] - \lambda \mathsf{D}_p^p(\widehat{\mathbb{P}}, \mathbb{P}) \right\} = V_1^{\widehat{\boldsymbol{\xi}}_1}(\xi_1) =: V_1.
$$

Remark 1 (Connection with Coherent Conditional Risk Mappings). The dynamic recursive form (8) is conceptually related to the conditional risk mapping introduced in Ruszczyński and Shapiro (2006). Specifically, a coherent conditional risk mapping can be represented as

$$\rho_{t|\xi_{[t-1]}}[Z(\boldsymbol{\xi}_{[t]})] = \sup_{\mathbb{P}_t \in \mathfrak{M}_t} \mathbb{E}_{\xi_t \sim \mathbb{P}_t}[Z(\boldsymbol{\xi}_{[t-1]}, \boldsymbol{\xi}_t)],$$

where Z is any measurable function on $\Xi_{[t]}$, and a convex conditional risk mapping can be represented as

$$\rho_{t|\xi_{[t-1]}}[Z(\boldsymbol{\xi}_{[t]})] = \sup_{\mathbb{P}_t \in \mathcal{P}(\Xi_t)} \left\{ \mathbb{E}_{\xi_t \sim \mathbb{P}_t}[Z(\boldsymbol{\xi}_{[t-1]}, \boldsymbol{\xi}_t)] - \lambda J_t(\mathbb{P}_t) \right\},$$

where J is a convex functional on $\mathcal{P}(\Xi_t)$. In our setting, let us define a conditional risk mapping $\rho^{\widehat{\boldsymbol{\xi}}_{[t-1]}}_{t|\xi_{[t-1]}}$ via

$$\begin{aligned} &\rho^{\widehat{\boldsymbol{\xi}}_{[t-1]}}_{t|\xi_{[t-1]}}[Z^{\widehat{\boldsymbol{\xi}}_{[t-1]}}(\boldsymbol{\xi}_{[t]})] \\ &= \begin{cases} \sup_{\gamma_t \in \Gamma(\widehat{\mathbb{P}}_{t|\widehat{\xi}_{[t-1]}}, \cdot)} \left\{ \mathbb{E}_{(\widehat{\xi}_t, \xi_t) \sim \gamma_t}\left[Z^{\widehat{\boldsymbol{\xi}}_{[t]}}(\boldsymbol{\xi}_{[t]})\right] : \ \gamma_t\text{-}\operatorname*{ess\,sup}_{(\xi_t, \widehat{\xi}_t) \in \Xi_t^2} \|\boldsymbol{\xi}_t - \widehat{\boldsymbol{\xi}}_t\| \le \vartheta \right\}, & p = \infty, \\ \sup_{\gamma_t \in \Gamma(\widehat{\mathbb{P}}_{t|\widehat{\xi}_{[t-1]}}, \cdot)} \left\{ \mathbb{E}_{(\widehat{\xi}_t, \xi_t) \sim \gamma_t}\left[Z^{\widehat{\boldsymbol{\xi}}_{[t]}}(\boldsymbol{\xi}_{[t]})\right] - \lambda \mathbb{E}_{(\widehat{\xi}_t, \xi_t) \sim \gamma_t}\left[\|\boldsymbol{\xi}_t - \widehat{\boldsymbol{\xi}}_t\|^p\right] \right\}, & p \in [1, \infty), \end{cases} \end{aligned} \tag{11}$$

where $\Gamma(\widehat{\mathbb{P}}_{t|\widehat{\xi}_{[t-1]}}, \cdot)$ is the set of joint distributions on Ξ_t^2 whose first marginal distribution is $\widehat{\mathbb{P}}_{t|\widehat{\xi}_{[t-1]}}$. Note that the uncertainty set in the case of $p = \infty$ is equivalent to a single-period ∞-Wasserstein ball

$$\mathfrak{M}_t^{\widehat{\boldsymbol{\xi}}_{[t-1]}} = \left\{ \mathbb{P}_t \in \mathcal{P}(\Xi_t) : \mathcal{W}_\infty(\widehat{\mathbb{P}}_{t|\widehat{\xi}_{[t-1]}}, \mathbb{P}_t) \le \vartheta \right\},$$

centered at the nominal conditional distribution $\widehat{\mathbb{P}}_{t+1|\widehat{\xi}_{[t]}}$ with radius ϑ, and the penalty term in the case of $p \in [1, \infty)$ is equivalent to a single-period p-Wasserstein distance penalty

$$J_t(\mathbb{P}_t) = \mathcal{W}_p^p(\widehat{\mathbb{P}}_{t|\widehat{\xi}_{[t-1]}}, \mathbb{P}_t).$$

With this definition, (7) and (9) can be rewritten as

$$V_t^{\widehat{\boldsymbol{\xi}}_{[t]}}(\boldsymbol{\xi}_{[t]}) = \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) + \rho^{\widehat{\boldsymbol{\xi}}_{[t]}}_{t+1|\xi_{[t]}}\left[V_{t+1}^{\widehat{\boldsymbol{\xi}}_{[t+1]}}(\boldsymbol{\xi}_{[t+1]})\right].$$

It is important to note that in the original definition of a conditional risk mapping, $\rho_{t|\xi_{[t-1]}}$ considers only the filtration generated by $\boldsymbol{\xi}_{[t]}$. However, in our case, it extends to involve the filtration generated by the nominal stochastic process $\widehat{\boldsymbol{\xi}}_{[T]}$. ♣

Remark 2 (Connection with Wasserstein DRO). Suppose $\widehat{\mathbb{P}}$ is a scenario fan, that is, $\widehat{\mathbb{P}}_{t+1|\widehat{\xi}_{[t]}}$ is a Dirac measure for all $t = 1, \dots, T-1$. Then (7) and (9) become respectively

$$\sup_{\mathbb{P} \in \mathfrak{M}} \mathbb{E}_{\mathbb{P}}\left[\sum_{t \in [T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\right] = c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\left[\sup_{\substack{\xi_t \in \Xi_t : \|\xi_t - \widehat{\xi}_t\| \le \vartheta \\ t = 2, \dots, T}} \sum_{t=2}^{T} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\right]$$

and

$$\sup_{\mathbb{P} \in \mathcal{P}(\Xi)} \left\{ \mathbb{E}_{\mathbb{P}}\Big[\sum_{t \in [T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\Big] - \lambda \mathrm{D}_p^p(\widehat{\mathbb{P}}, \mathbb{P}) \right\} = c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\left[\sup_{\substack{\xi_t \in \Xi_t \\ t = 2, \dots, T}} \left\{ \sum_{t=2}^{T} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) - \lambda \sum_{t=2}^{T} \|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p \right\}\right].$$

Notably, they coincide with the dual formulation of Wasserstein DRO.

Now suppose the underlying stochastic process has a density, then $\widehat{\mathbb{P}}$ is a scenario fan with probability one. As a result, the statistical consistency of the worst-case risk under nested distance follows easily from that of Wasserstein DRO (Kuhn et al. 2019, Gao et al. 2022). This is in contrast to the statistical inconsistency of the nested distance itself (Pflug and Pichler 2016). ♣

3.2. Time-Consistent Policy Optimization under Stagewise Independence

In numerical studies on multistage DRO, the following problem is often considered

$$\min_{x_1\in\mathcal{X}_1} c_1^\top x_1 + \min_{\boldsymbol{x}_2\in\mathcal{X}_2(x_1)} \rho_2\Bigg[c_2^\top \boldsymbol{x}_2(\boldsymbol{\xi}_2) + \min_{\boldsymbol{x}_3\in\mathcal{X}_3(\boldsymbol{x}_2)} \rho_{3|\boldsymbol{\xi}_{[2]}}^{\widehat{\boldsymbol{\xi}}_{[2]}}\Big[c_3^\top \boldsymbol{x}_3(\boldsymbol{\xi}_{[3]}) + \cdots + \min_{\boldsymbol{x}_T\in\mathcal{X}_T(\boldsymbol{x}_{T-1})} \rho_{T|\boldsymbol{\xi}_{[T-1]}}^{\widehat{\boldsymbol{\xi}}_{[T-1]}}\big[c_T^\top \boldsymbol{x}_T(\boldsymbol{\xi}_{[T]})\big]\cdots\Big]\Bigg]. \tag{12}$$

In general, this is just a lower bound of the original problem

$$\min_{\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1}),\forall t\in[T]} c_1^\top x_1 + \rho_2\Bigg[c_2^\top \boldsymbol{x}_2(\boldsymbol{\xi}_2) + \rho_{3|\boldsymbol{\xi}_{[2]}}^{\widehat{\boldsymbol{\xi}}_{[2]}}\Big[c_3^\top \boldsymbol{x}_3(\boldsymbol{\xi}_{[3]}) + \cdots + \rho_{T|\boldsymbol{\xi}_{[T-1]}}^{\widehat{\boldsymbol{\xi}}_{[T-1]}}\big[c_T^\top \boldsymbol{x}_T(\boldsymbol{\xi}_{[T]})\big]\cdots\Big]\Bigg] \tag{13}$$

as it involves exchange of minimization and expectation. Nevertheless, the two problem are equivalent when the nominal scenario tree $\widehat{\mathbb{P}}$ is stagewise independent. In this case, we set $\rho_t = \rho_{t|\boldsymbol{\xi}_{[t-1]}}^{\widehat{\boldsymbol{\xi}}_{[t-1]}}$.

ASSUMPTION 1 **(Stagewise independence)**. *The nominal scenario tree $\widehat{\mathbb{P}}$ is stagewise independent, namely, $\widehat{\mathbb{P}}_{t|\widehat{\boldsymbol{\xi}}_{[t-1]}} = \widehat{\mathbb{P}}_t$ for all $t\in[T]$.*

In the remainder of the paper, we make the following assumption (c.f. Shapiro et al. (2021, Definition 3.1)).

ASSUMPTION 2 **(Relatively complete recourse)**. *For every $t=2,\dots,T$ and every sequence of feasible decisions $(\boldsymbol{x}_1,\dots,\boldsymbol{x}_{t-1})$, the set $\mathcal{X}_t(\boldsymbol{x}_{t-1},\xi_t)$ is non-empty for every $\xi_t\in\Xi_t$.*

Without this assumption, the worst-case risk is always infinite for $p\in[1,\infty)$. For $p=\infty$, the condition above can be relaxed by replacing Ξ_t with $\cup_{\widehat{\xi}_t\in\operatorname{supp}\widehat{\mathbb{P}}_t}\{\xi_t\in\Xi_t : \|\xi_t-\widehat{\boldsymbol{\xi}}_t\|\le\vartheta\}$.

Below we establish the dynamic programming reformulations of (1). The proof is based on the interchangeability principle and can be found in Appendix EC.2.

Define the dynamic programming formulation

$$\begin{aligned}
&\mathcal{Q}_{T+1} := 0,\\
&\mathcal{Q}_t(x_{t-1},\boldsymbol{\xi}_{[t-1]}) := \min_{\boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\in\mathcal{X}_t(x_{t-1},\cdot)} \rho_t\Big[\boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) + \mathcal{Q}_{t+1}\big(\boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}),\boldsymbol{\xi}_{[t]}\big)\Big],\\
&\mathcal{Q}_1 := \min_{x_1\in\mathcal{X}_1}\left\{c_1^\top x_1 + \mathcal{Q}_2(x_1)\right\}.
\end{aligned} \tag{$\mathsf{P}_{\text{dynamic}}$}$$

THEOREM 3. *Suppose Assumptions 1 and 2 hold. Then the optimal value of the dynamic program ($\mathsf{P}_{\text{dynamic}}$) equals the optimal value of ($\mathsf{P}_{\text{static}}$) when $p=\infty$ and the optimal value of ($\mathsf{P}_{\text{static}}$-soft) when $p\in[1,\infty)$. Moreover, set $Q_{T+1}\equiv 0$ and*

$$Q_t(x_{t-1},\widehat{\boldsymbol{\xi}}_t) := \begin{cases}\sup_{\xi_t\in\Xi_t:\|\xi_t-\widehat{\xi}_t\|\le\vartheta} \min_{x_t\in\mathcal{X}_t(x_{t-1},\xi_t)}\left\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t,\widehat{\boldsymbol{\xi}}_{t+1})\big]\right\}, & p=\infty,\\ \sup_{\xi_t\in\Xi_t} \min_{x_t\in\mathcal{X}_t(x_{t-1},\xi_t)}\left\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t,\widehat{\boldsymbol{\xi}}_{t+1})\big] - \lambda\|\xi_t-\widehat{\boldsymbol{\xi}}_t\|^p\right\}, & p\in[1,\infty),\end{cases} \tag{14}$$

for $t=2,\dots,T$. Then it holds that $\mathbb{E}_{\widehat{\mathbb{P}}_t}[Q_t(x_{t-1},\widehat{\boldsymbol{\xi}}_t)] = \mathcal{Q}_t(x_{t-1},\boldsymbol{\xi}_{[t-1]})$.

Theorem 3 shows that expected risk-to-go function $\mathcal{Q}_t$ defined in ($\mathsf{P}_{\text{dynamic}}$), which involves functional optimization over $\boldsymbol{x}_t(\boldsymbol{\xi}_{[t-1]},\cdot)$, can be evaluated through the expectation of the risk-to-go function Q_t, which involves only a finite-dimensional problem. Note that due to the stagewise independence Assumption 1, the risk-to-go function $Q_t(x_{t-1},\cdot)$ depends only on the current-stage uncertainty and the expected risk-to-go function $\mathcal{Q}_t(x_{t-1},\cdot)$ is a constant function. Thus, from now on we will omit the second argument of $\mathcal{Q}_t$ and denote it as $\mathcal{Q}_t(x_{t-1})$. Also note that the supremum over ξ_t may not always be tractable. Nevertheless, in Section 4, we will explore cases where tractable solutions are possible.

By expanding the recursion using (14), ($\mathsf{P}_{\mathsf{dynamic}}$) is equivalent to

$$\min_{x_1\in\mathcal{X}_1}\left\{c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\left[\sup_{\|\xi_2-\widehat{\xi}_2\|\le\vartheta}\min_{x_2\in\mathcal{X}_2(x_1,\xi_2)}\left\{c_2^\top x_2+\mathbb{E}_{\widehat{\mathbb{P}}_3}\left[\sup_{\|\xi_3-\widehat{\xi}_3\|\le\vartheta}\min_{x_3\in\mathcal{X}_3(x_2,\xi_3)}\left\{c_3^\top x_3+\right.\right.\right.\right.\right.$$
$$\left.\left.\left.\left.\left.\cdots+\mathbb{E}_{\widehat{\mathbb{P}}_T}\left[\sup_{\|\xi_T-\widehat{\xi}_T\|\le\vartheta}\min_{x_T\in\mathcal{X}_T(x_{T-1},\xi_T)}c_T^\top x_T\right]\cdots\right\}\right]\right\}\right]\right\}$$

when $p=\infty$, and

$$\min_{\lambda\ge0,x_1\in\mathcal{X}_1}\left\{\lambda\vartheta^p+c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\left[\sup_{\xi_2\in\Xi_2}\min_{x_2\in\mathcal{X}_2(x_1)}\left\{c_2^\top \boldsymbol{x}_2(\xi_2)+\mathbb{E}_{\widehat{\mathbb{P}}_3}\left[\sup_{\xi_3\in\Xi_3}\min_{x_3\in\mathcal{X}_3(x_2)}\left\{c_3^\top \boldsymbol{x}_3(\xi_{[3]})+\right.\right.\right.\right.\right.$$
$$\left.\left.\left.\left.\left.\cdots+\mathbb{E}_{\widehat{\mathbb{P}}_T}\left[\sup_{\xi_T\in\Xi_T}\min_{x_T\in\mathcal{X}_T(x_{T-1},\xi_T)}c_T^\top \boldsymbol{x}_T(\xi_{[T]})-\lambda\|\xi_T-\widehat{\boldsymbol{\xi}}_T\|^p\right]\cdots\right\}-\lambda\|\xi_3-\widehat{\boldsymbol{\xi}}_3\|^p\right]-\lambda\|\xi_2-\widehat{\boldsymbol{\xi}}_2\|^p\right\}\right]\right\}$$

when $p\in[1,\infty)$. Using duality for Wasserstein DRO, they are equivalent to

$$\min_{x_1\in\mathcal{X}_1}\left\{c_1^\top x_1+\sup_{\mathbb{P}_2\in\mathfrak{M}_2}\mathbb{E}_{\mathbb{P}_2}\left[\min_{x_2\in\mathcal{X}_2(x_1,\xi_2)}\left\{c_2^\top x_2+\sup_{\mathbb{P}_3\in\mathfrak{M}_3}\mathbb{E}_{\mathbb{P}_3}\left[\min_{x_3\in\mathcal{X}_3(x_2,\xi_3)}\left\{c_3^\top x_3+\cdots+\right.\right.\right.\right.\right.$$
$$\left.\left.\left.\left.\left.\sup_{\mathbb{P}_T\in\mathfrak{M}_T}\mathbb{E}_{\mathbb{P}_T}\left[\min_{x_T\in\mathcal{X}_T(x_{T-1},\xi_T)}c_T^\top x_T\right]\cdots\right\}\right]\right\}\right]\right\},$$

and

$$\min_{\lambda\ge0,x_1\in\mathcal{X}_1}\left\{\lambda\vartheta^p+c_1^\top x_1+\sup_{\mathbb{P}_2\in\mathcal{P}(\Xi_2)}\mathbb{E}_{\mathbb{P}_2}\left[\min_{x_2\in\mathcal{X}_2(x_1)}\left\{c_2^\top \boldsymbol{x}_2(\xi_2)+\sup_{\mathbb{P}_3}\mathbb{E}_{\mathbb{P}_3\in\mathcal{P}(\Xi_3)}\left[\min_{x_3\in\mathcal{X}_3(x_2)}\left\{c_3^\top \boldsymbol{x}_3(\xi_{[3]})+\cdots+\right.\right.\right.\right.\right.$$
$$\left.\left.\left.\left.\left.\sup_{\mathbb{P}_T}\mathbb{E}_{\mathbb{P}_T\in\mathcal{P}(\Xi_T)}\left[\min_{x_T\in\mathcal{X}_T(x_{T-1},\xi_T)}c_T^\top \boldsymbol{x}_T(\xi_{[T]})\right]-\lambda\mathcal{W}_p^p(\widehat{\mathbb{P}}_T,\mathbb{P}_T)\cdots\right\}\right]-\lambda\mathcal{W}_p^p(\widehat{\mathbb{P}}_3,\mathbb{P}_3)\right\}\right]-\lambda\mathcal{W}_p^p(\widehat{\mathbb{P}}_2,\mathbb{P}_2)\right\}$$

respectively. These formulations have been considered in numerical studies (Duque and Morton 2020, Zhang and Sun 2020, 2022). These works, however, either assume a finite sample space, or do not render a policy with provable optimality guarantees.

4. Tractable Policy Optimization

In this section, we aim to provide tractable formulations for obtaining a policy for ($\mathsf{P}_{\mathsf{dynamic}}$). As a corollary of Theorem 3, we have the following result.

Corollary 1. *Recall the Q-function defined in* (14). *An optimal policy* $(\boldsymbol{x}_1^\star,\ldots,\boldsymbol{x}_T^\star)$ *of* ($\mathsf{P}_{\mathsf{dynamic}}$) *is given recursively by*

$$\begin{aligned}\boldsymbol{x}_1^\star&\in\operatorname*{arg\,min}_{x_1\in\mathcal{X}_1}\left\{c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}[\mathcal{Q}_2(x_1)]\right\},\\ \boldsymbol{x}_t^\star(\xi_{[t]})&\in\operatorname*{arg\,min}_{x_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1}^\star(\xi_{[t-1]}),\xi_t)}\left\{c_t^\top x_t+\mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[\mathcal{Q}_{t+1}(x_t,\widehat{\boldsymbol{\xi}}_{t+1})\right]\right\},\quad \xi_{[t]}\in\Xi_{[t]},t=2,\ldots,T.\end{aligned}\tag{15}$$

The optimal policy (15) is defined on the entire sample space, marking a significant difference from its non-robust counterpart, which yields a solution defined solely on the sample paths within the nominal scenario tree.

Below, we identify cases where the risk-to-go function Q_t, as defined in (14), and thus ($\mathrm{P}_{\mathrm{dynamic}}$), can be computed efficiently. Note that when $T = 2$, ($\mathrm{P}_{\mathrm{dynamic}}$) reduces to two-stage Wasserstein DRO, for which exact tractable reformulations have been established when $p \in \{1, \infty\}$ and the uncertainty appears in the objective or right-hand side (Mohajerin Esfahani and Kuhn 2018, Hanasusanto and Kuhn 2018, Xie 2020). Below, we will show that these results can be extended to multistage problems with essentially the same assumptions.

4.1. Objective Uncertainty Only

We first consider problems with objective uncertainty only, in which case we identify $\boldsymbol{\xi}_t$ with $\boldsymbol{c}_t$, and the constraint set $\mathcal{X}_t(x_{t-1}, \xi_t) = \mathcal{X}_t(x_{t-1}) = \{x_t \geq 0 : A_t x_t = b_t - B_t x_{t-1}\}$ is deterministic once x_{t-1} is given. The following theorem shows an equivalent reformulation of ($\mathrm{P}_{\mathrm{dynamic}}$).

Corollary 2. *Suppose $p = \infty$, $\Xi_t = (\mathbb{R}^{d_t}, \|\cdot\|)$. Then Q_t, $t = 2, \ldots, T$, defined in (14) can be computed as*

$$Q_t(x_{t-1}, \widehat{\boldsymbol{c}}_t) = \begin{cases} \min_{x_t \in \mathcal{X}_t(x_{t-1}), \|x_t\|_* \leq \lambda} \left\{ \widehat{\boldsymbol{c}}_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{c}}_{t+1}) \right] \right\}, & p = 1, \\ \min_{x_t \in \mathcal{X}_t(x_{t-1})} \left\{ \widehat{\boldsymbol{c}}_t^\top x_t + (1 - 1/p)(\frac{1}{p\lambda})^{\frac{1}{p-1}} \|x_t\|_*^{\frac{p}{p-1}} + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{c}}_{t+1}) \right] \right\}, & p \in (1, \infty), \\ \min_{x_t \in \mathcal{X}_t(x_{t-1})} \left\{ \widehat{\boldsymbol{c}}_t^\top x_t + \vartheta \|x_t\|_* + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{c}}_{t+1}) \right] \right\}, & p = \infty. \end{cases}$$

Corollary 2 shows that when the uncertainty appears only in the objective, ($\mathrm{P}_{\mathrm{dynamic}}$) is equivalent to a scenario approximation problem with norm regularization on each $\boldsymbol{x}_t$. The regularization is a hard constraint when $p = 1$ and a soft penalty when $p \in (1, \infty]$. The proof is based on the straightforward application of Hölder's and Young's inequalities; see EC.3.1 for details.

We remark that the optimal robust policy may not be unique, as demonstrated in the following result.

Proposition 1. *Set*

$$\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}) \in \begin{cases} \arg\min_{x_t \in \mathcal{X}_t(\widehat{\boldsymbol{x}}_{t-1}(\widehat{\boldsymbol{c}}_{[t-1]})), \|x_t\|_* \leq \lambda} \left\{ \widehat{\boldsymbol{c}}_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{c}}_{t+1}) \right] \right\}, & p = 1, \\ \arg\min_{x_t \in \mathcal{X}_t(\widehat{\boldsymbol{x}}_{t-1}(\widehat{\boldsymbol{c}}_{[t-1]}))} \left\{ \widehat{\boldsymbol{c}}_t^\top x_t + (1 - \frac{1}{p})(\frac{1}{p\lambda})^{\frac{1}{p-1}} \|x_t\|_*^{\frac{p}{p-1}} + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{c}}_{t+1}) \right] \right\}, & p \in (1, \infty), \\ \arg\min_{x_t \in \mathcal{X}_t(\widehat{\boldsymbol{x}}_{t-1}(\widehat{\boldsymbol{c}}_{[t-1]}))} \left\{ \widehat{\boldsymbol{c}}_t^\top x_t + \vartheta \|x_t\|_* + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{c}}_{t+1}) \right] \right\}, & p = \infty, \end{cases}$$

and set recursively $\widehat{\boldsymbol{c}}_1^c := c_1$ and

$$\widehat{\boldsymbol{c}}_t^c := \begin{cases} \displaystyle\arg\min_{\widehat{\boldsymbol{c}}_t \in \operatorname{supp} \widehat{\mathbb{P}}_t} \left\{ \widehat{\boldsymbol{c}}_t^\top \widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}\big(\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t), \widehat{\boldsymbol{c}}_{t+1}\big) \right] + \lambda \|\widehat{\boldsymbol{c}}_t - c_t\|^p \right\}, & p \in [1, \infty), \\ \displaystyle\arg\min_{\widehat{\boldsymbol{c}}_t \in \operatorname{supp} \widehat{\mathbb{P}}_t : \|\widehat{\boldsymbol{c}}_t - c_t\| \leq \vartheta} \widehat{\boldsymbol{c}}_t^\top \widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t) + \vartheta \|\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t)\|_* + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}\big(\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t), \widehat{\boldsymbol{c}}_{t+1}\big) \right], & p = \infty. \end{cases}$$

Then the policy $\bar{\boldsymbol{x}}_t(c_{[t]}) := \widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}^c)$ is also optimal for ($\mathrm{P}_{\mathrm{dynamic}}$).

The policy $\bar{\boldsymbol{x}} = (\bar{\boldsymbol{x}}_1, \ldots, \bar{\boldsymbol{x}}_T)$ is defined for every sample path in $\Xi_{[T]}$ when $p \in [1, \infty)$ and in $\{c_{[T]} \in \Xi_{[T]} : \|c_{[T]} - \widehat{\boldsymbol{c}}_{[T]}\|_\infty \leq \vartheta\}$ when $p = \infty$. The sample path $\widehat{\boldsymbol{c}}_{[T]}^c$ represents the best, with regard to the norm-regularized cost-to-go, in-sample path within a ϑ-neighborhood of $c_{[T]}$ when $p = \infty$, or within a λ-soft neighborhood of $c_{[T]}$ when $p \in [1, \infty)$. Notably, the computation of the policy $\bar{\boldsymbol{x}}$ requires knowledge of only the optimal robust policy values on sample paths from the nominal scenario tree, contrasting with the policy $\boldsymbol{x}^\star$ defined in (15) that requires the entire cost-to-go function $Q_t(\cdot, \widehat{\boldsymbol{c}}_t)$. Such computational advantage enhances the appeal of the policy $\bar{\boldsymbol{x}}$. Its optimality is obtained by verifying that the worst-case risk of $\bar{\boldsymbol{x}}_t$ does not exceed the risk-to-go as defined by Q_t. For the detailed proof, please refer to EC.3.1.

4.2. Right-hand Side Uncertainty Only

Next, we consider problems with right-hand side uncertainty only. To ease the presentation, we consider either $\boldsymbol{\xi}_t = \boldsymbol{b}_t$ or $\boldsymbol{\xi}_t = \boldsymbol{B}_t$. The following theorem provides an equivalent reformulation of ($\mathsf{P}_{\mathsf{dynamic}}$) for $p = \infty$.

Corollary 3. *Suppose $p=\infty$, $\boldsymbol{\xi}_t = \boldsymbol{b}_t \in \Xi_t = (\mathbb{R}^{d_t}, \|\cdot\|)$ or $\boldsymbol{\xi}_t = \boldsymbol{B}_t \in \Xi_t = (\mathbb{R}^{d_t \times m_t}, \|\cdot\|)$. Set $\psi_t(x_t) := c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t, \widehat{\boldsymbol{\xi}}_{t+1})\big] + \mathbb{I}\{x_t \geq 0\}$. Then Q_t, $t = 2, \ldots, T$, defined in (14) can be computed as*

$$\begin{aligned} Q_t(x_{t-1}, \widehat{\boldsymbol{b}}_t) &= \max_{y_t \in \mathbb{R}^{d_t}} \left\{ (\widehat{\boldsymbol{b}}_t - B_t x_{t-1})^\top y_t + \vartheta \|y_t\|_* - \psi_t^*(A_t^\top y_t) \right\}, \qquad \text{if } \boldsymbol{\xi}_t = \boldsymbol{b}_t, \\ Q_t(x_{t-1}, \widehat{\boldsymbol{B}}_t) &= \max_{y_t \in \mathbb{R}^{d_t}} \left\{ (b_t - \widehat{\boldsymbol{B}}_t x_{t-1})^\top y_t + \vartheta \|y_t x_{t-1}^\top\|_* - \psi_t^*(A_t^\top y_t) \right\}, \quad \text{if } \boldsymbol{\xi}_t = \boldsymbol{B}_t. \end{aligned} \tag{16}$$

When $\|\cdot\| = \|\cdot\|_1$, it holds that

$$Q_t(x_{t-1}, \widehat{\boldsymbol{b}}_t) = \max_{j \in [d_t], \delta \in \{1,-1\}} \min_{x_t \geq 0} \left\{ c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t, \widehat{\boldsymbol{b}}_{t+1})\big] : A_t x_t = \widehat{\boldsymbol{b}}_t - B_t x_{t-1} + \vartheta \delta e_j \right\},$$

where e_j is the j-th unit vector. When $\|\cdot\| = \|\cdot\|_{\mathrm{op}}$, where $\|B\|_{\mathrm{op}} = \sup_{\|v\| \leq 1} \|Bv\|_1$, it holds that

$$Q_t(x_{t-1}, \widehat{\boldsymbol{B}}_t) = \max_{j \in [d_t], \delta \in \{1,-1\}} \min_{x_t \geq 0} \left\{ c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t, \widehat{\boldsymbol{B}}_{t+1})\big] : A_t x_t = b_t - \widehat{\boldsymbol{B}}_t x_{t-1} + \vartheta \delta \|x_{t-1}\| e_j \right\}.$$

The risk-to-go function Q_t encourages a large norm of the dual variable y_t. It also penalizes a large norm on the primal variable x_t when the uncertainty is present in $\boldsymbol{B}_t$. Note that solving (16) involves maximizing a convex norm function, which can be hard in general (Xie 2020, Proposition 6). Nevertheless, it becomes tractable when the dual norm entails an inf-norm, which can be represented as component-wise maximum absolute value; see EC.3.3. The resulting reformulation of Q_t involves solving $2d_t$ problems in total, each of which perturbs the constraints of the scenario approximation problem by a unit vector with a magnitude proportional to ϑ. When $T = 2$, this is consistent with Xie (2020, Theorem 3). We will apply this result to a dynamic portfolio selection problem in Section 5.

Corollary 4. *Suppose $\Xi_t = (\mathbb{R}^{d_t}, \|\cdot\|)$ and $p = 1$. Set $\mathcal{Y}_{T+1} \equiv \{0\}$, and for $t \in [T]$, define recursively*

$$\mathcal{S}_t := \left\{ y_t \in \mathbb{R}^{d_t} : \exists \boldsymbol{y}_{t+1} \in \mathcal{Y}_{t+1} \text{ s.t. } A_t^\top y_t + B_{t+1}^\top \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}[\boldsymbol{y}_{t+1}(\widehat{\boldsymbol{\xi}}_{t+1})] \leq c_t \right\},$$

where $\mathcal{Y}_t$ is the space of functions from $\widehat{\Xi}_t$ to $\mathcal{S}_t$. Set $\widehat{Q}_{T+1}(x_t, \widehat{\boldsymbol{\xi}}_{[t+1]}) \equiv 0$ and

$$\widehat{Q}_t(x_{t-1}, \widehat{\boldsymbol{\xi}}_t) = \min_{x_t \in \mathcal{X}_t(x_{t-1}, \widehat{\boldsymbol{\xi}}_t)} \left\{ c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[\widehat{Q}_{t+1}(x_t, \widehat{\boldsymbol{\xi}}_{t+1})\right] \right\}, \quad t = 2, \ldots, T.$$

Then

$$\begin{aligned} Q_t(x_{t-1}, \widehat{\boldsymbol{b}}_t) &= \widehat{Q}_t(x_{t-1}, \widehat{\boldsymbol{b}}_t) + \infty \cdot \mathbf{1}\left\{ \lambda < \max_{s=t,\ldots,T} \max_{y \in \mathcal{S}_s} \|y\|_* \right\}, \qquad \text{if } \boldsymbol{\xi}_t = \boldsymbol{b}_t, \\ Q_t(x_{t-1}, \widehat{\boldsymbol{B}}_t) &= \widehat{Q}_t(x_{t-1}, \widehat{\boldsymbol{B}}_t) + \infty \cdot \mathbf{1}\left\{ \lambda < \max_{s=t,\ldots,T} \max_{y \in \mathcal{S}_s} \|x_t y^\top\|_* \right\}, \quad \text{if } \boldsymbol{\xi}_t = \boldsymbol{B}_t, \end{aligned}$$

and the optimal value of ($\mathsf{P}_{\text{dynamic}}$) *equals*

$$\begin{cases} \min_{x_1 \in \mathcal{X}_1} \left\{ c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}[\widehat{Q}_2(x_1, \widehat{\boldsymbol{b}}_2)] \right\} + \vartheta \cdot \max_{t=2,\dots,T} \max_{y \in \mathcal{S}_t} \|y\|_*, & \text{if } \boldsymbol{\xi}_t = \boldsymbol{b}_t, \\ \min_{x_1 \in \mathcal{X}_1} \left\{ c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}[\widehat{Q}_2(x_1, \widehat{\boldsymbol{B}}_2)] \right\} + \vartheta \cdot \max_{t=2,\dots,T} \max_{y \in \mathcal{S}_t} \|x_t y^\top\|_*, & \text{if } \boldsymbol{\xi}_t = \boldsymbol{B}_t. \end{cases} \tag{17}$$

Note that $\widehat{Q}_t$ is the cost-to-go function for the (non-robust) scenario approximation problem. The first term of (17) is the optimal value of the scenario approximation problem, while the second term of (17) is a linear function of ϑ, whose value is independent of the policy $\boldsymbol{x}$. Consequently, (17) share the same optimal policy values as the scenario approximation problem on each sample path from the nominal scenario tree. We would like to emphasize that the optimal robust policy for ($\mathsf{P}_{\text{dynamic}}$), as in (15), is defined for all sample paths in $\Xi_{[T]}$, whereas the optimal policy for the scenario approximation problem is only defined for the sample paths within the scenario tree. In this respect, the robust formulation induces a safe way to extend the optimal solution to the scenario approximation problem across the entire sample space, and justifies the heuristic policy in the literature (Shapiro et al. 2012, Keutchayan et al. 2017, Zhang and Sun 2022).

Corollary 4 generalizes the results for two-stage Wasserstein DRO (Hanasusanto and Kuhn 2018, Duque et al. 2022). Below we give an example in the context of the multi-stage newsvendor problem. A similar observation was made in Mohajerin Esfahani and Kuhn (2018) for the static newsvendor problem.

Example 4 (Newsvendor). Consider a multistage distributionally robust newsvendor model. Let $\boldsymbol{x}_t$ be the inventory level after having ordered in stage t but before the demand $\boldsymbol{\xi}_t$ in that stage is realized. Let c_t, c_t^b and c_t^h be the ordering, back-order penalty and holding costs per unit in stage t, respectively. The multistage newsvendor is given by

$$\begin{aligned} \min_{\boldsymbol{x}_t, t \in [T]} \quad & \sup_{\mathbb{P} \in \mathfrak{M}} \mathbb{E}_{\mathbb{P}} \left[\sum_{t \in [T]} c_t(\boldsymbol{x}_t - \boldsymbol{x}_{t-1} + \boldsymbol{\xi}_{t-1}) + c_t^h(\boldsymbol{x}_{t-1} - \boldsymbol{\xi}_{t-1})_+ + c_t^b(\boldsymbol{\xi}_{t-1} - \boldsymbol{x}_{t-1})_+ \right] \\ s.t. \quad & \boldsymbol{x}_t \ge \boldsymbol{x}_{t-1} - \boldsymbol{\xi}_{t-1}, \\ & \boldsymbol{x}_t \ge 0. \end{aligned}$$

With additional auxiliary variables z_t, we can rewrite it as a multistage linear program with right-hand uncertainty

$$\begin{aligned} \min_{\boldsymbol{x}_t, z_t, t \in [T]} \quad & \sup_{\mathbb{P} \in \mathfrak{M}} \mathbb{E}_{\mathbb{P}} [c_1 \boldsymbol{x}_1 + z_2 + \cdots + z_T] \\ s.t. \quad & -\boldsymbol{x}_t + \boldsymbol{x}_{t-1} \le \boldsymbol{\xi}_{t-1}, \\ & c_t \boldsymbol{x}_t - z_t + (c_t^h - c_t)\boldsymbol{x}_{t-1} \le (c_t^h - c_t)\boldsymbol{\xi}_{t-1}, \\ & c_t \boldsymbol{x}_t - z_t - (c_t + c_t^b)\boldsymbol{x}_{t-1} \le -(c_t + c_t^b)\boldsymbol{\xi}_{t-1}, \\ & \boldsymbol{x}_t \ge 0. \end{aligned}$$

Then Corollary 4 indicates that the Wasserstein robust solution and the non-robust solution coincide on sample paths within the nominal scenario tree. ♣

5. Application in Dynamic Portfolio Selection

5.1. Problem Formulation

We consider a portfolio selection problem of an investor who seeks to minimize the dis-utility of the terminal wealth. Given some initial wealth W_1, she invests in n assets. The monetary value of all n investments are represented using a vector, $x_t \in \mathbb{R}^n$. The return rate at time period t is modeled by a random variable $\boldsymbol{\xi}_t \in \mathbb{R}_t$. At each stage before the terminal, she may re-balance her wealth by taking long-only positions across the n investments. Suppose the investor's terminal dis-utility function is given by $U(W_T) := \max(-\alpha_0 - \beta_0 W_T, -\alpha_1 - \beta_1 W_T)$, where $\alpha_0, \alpha_1, \beta_0, \beta_1$ are used to encode the investor's preferences. We can write the robust portfolio selection problem as

$$\begin{aligned} \min_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{T-1}\geq 0, \boldsymbol{x}_T} \; & \max_{\mathbb{P}\in\mathfrak{M}} \mathbb{E}_{\mathbb{P}}[U(\boldsymbol{x}_T)] \\ s.t. \quad & \mathbf{1}^\top \boldsymbol{x}_1 = W_1, \\ & \mathbf{1}^\top \boldsymbol{x}_t = \boldsymbol{\xi}_t^\top \boldsymbol{x}_{t-1}, \quad t = 2,\ldots,T-1, \\ & \boldsymbol{x}_T = \boldsymbol{\xi}_T^\top \boldsymbol{x}_{T-1}. \end{aligned} \tag{18}$$

Using Corollary 3, we obtain the following reformulation, whose proof is given in EC.4.1.

Corollary 5. *Using the setup in Corollary 3, the dynamic programming reformulation of* (18) *is given by*

$$\begin{aligned} Q_T(x_{T-1}, \widehat{\boldsymbol{\xi}}_{[T]}) &:= \max_{\substack{z\in\{0,1\} \\ \delta\in\{1,-1\}}} \max\Big(-\alpha_0 - \beta_0 \widehat{\boldsymbol{\xi}}_T^\top x_{T-1} + \vartheta(1-z)\delta\|x_{T-1}\|, \; -\alpha_1 - \beta_1 \widehat{\boldsymbol{\xi}}_T^\top x_{T-1} + \vartheta z\delta\|x_{T-1}\| \Big), \\ Q_t(x_{t-1}, \widehat{\boldsymbol{\xi}}_{[t]}) &:= \max_{\delta\in\{1,-1\}} \min_{\substack{\mathbf{1}^\top x_t = \widehat{\boldsymbol{\xi}}_t^\top x_{t-1} + \vartheta\delta\|x_{t-1}\|, \\ x_t\geq 0}} \mathbb{E}_{\widehat{\mathbb{P}}_{t+1|\widehat{\boldsymbol{\xi}}_{[t]}}} \Big[Q_{t+1}(x_t, \widehat{\boldsymbol{\xi}}_{[t+1]}) \Big], \; t = 2,\ldots,T-1, \\ Q_1 &:= \min_{\substack{\mathbf{1}^\top x_1 = W_1 \\ x_1\geq 0}} \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{\xi}_2}} \Big[Q_2(x_1, \widehat{\boldsymbol{\xi}}_2) \Big]. \end{aligned}$$

5.2. Experiment Setup

Let $\alpha_0 = 0$, $\alpha_1 = (1-a)W_1$, $\beta_0 = 1$, $\beta_1 = a$, which corresponds to the dis-utility

$$U(W_T) := \begin{cases} -W_T, & W_T \leq W_1, \\ -W_1 - a(W_T - W_1), & W_T > W_1. \end{cases}$$

Here $0 < a < 1$ encodes the investor's preference for gains, which is set to be 0.5 in our experiments. Suppose the investor starts with an initial wealth of $W_1 = 10000$ units, and can invest the wealth across the following 5 assets, iShares MSCI Emerging Markets ETF (EEM), iShares 20+ Year Treasury Bond ETF (TLT), Schwab US TIPS ETF (SCHP), SPDR S&P Oil & Gas Equipment & Services ETF (XES), and ProShares UltraShort Financials ETF (SKF). We simulate the monthly asset returns using a log-normal distribution with mean and covariance estimated using adjusted closing prices from January 1, 2018 to June 30, 2021; see Appendix EC.4.2 for estimation results. We assume stage-wise independence, for the convenience of out-of-sample test.

We describe our out-of-sample testing procedure as follows. The training dataset is a T-stage scenario tree, where at stage $t = 2,\ldots,T$ there are $\widehat{N}_t$ independent scenarios. The testing dataset is another T-stage scenario tree, independent from the training tree, where at stage $t = 2,\ldots,T$ there are N_t independent scenarios. The reformulated problem in Corollary 5 can be viewed as a regularized SAA problem, so we solve it using a modified SDDP algorithm; see Algorithm 1 in Appendix EC.4.3 for details. Since the SDDP algorithm does not provide a policy but only the first-stage decision, we need to resolve the remaining subproblems to obtain subsequent robust decisions.

To describe our out-of-sample testing procedure, consider a 3-stage problem as an example, illustrated in Figure 3. First, we use the entire training tree $\{(\widehat{\boldsymbol{\xi}}_t^1, \ldots, \widehat{\boldsymbol{\xi}}_t^{N_t})\}_{t=2,\ldots,T}$ to obtain a first-stage robust decision x_1^{rob} and evaluate the first-stage cost. Next, at stage 2, we observe a sample $\xi_2^{i_2}$ from the second-stage scenarios in the testing tree $(\xi_2^1, \ldots, \xi_2^{N_2})$, and use all scenarios in the remaining stages

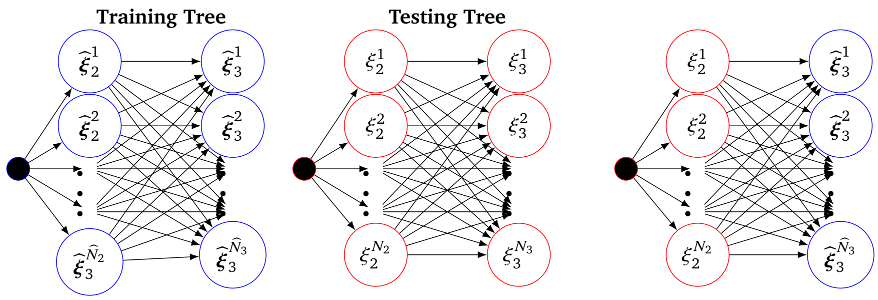


Figure 3 Illustration of the out-of-sample testing procedure for a 3-stage problem. In the last tree, the second-stage decision is computed using the training scenarios in the third stage after observing a testing scenario in the second stage.

(i.e., stage 3 in this setting) from the training tree $(\widehat{\boldsymbol{\xi}}_3^1,\dots,\widehat{\boldsymbol{\xi}}_3^{\widehat{N}_3})$ to solve for the second-stage robust decision x_2^{rob} and evaluate its second-stage cost. At stage 3, we draw a sample $\xi_3^{i_3}$ from the third-stage scenarios in the testing tree $(\xi_3^1,\dots,\xi_2^{N_3})$, and use all scenarios in the remaining stages from the training tree (i.e., none in this setting) to solve for the third-stage robust decision x_3^{rob} and evaluate its third-stage cost. After going through all stages, we sum up the per-stage costs in all stages, which gives a realization of the out-of-sample cost associate with the testing sample path $(\xi_2^{i_2},\xi_3^{i_3})$. To get an estimate of the expected out-of-sample cost, we following the procedure above to sample M testing paths and average over them. We refer to Algorithm 2 in Appendix EC.4.3 for a pseudocode for general T-stage problems.

5.3. Numerical Results

In our experiments, we set $T\in\{3,4,5\}$, $\widehat{N}_2=\cdots=\widehat{N}_T\in\{2,5,10\}$, $N_2=\cdots=N_T=30$, $M=25$. We are interested in how the out-of-sample performance depend on different parameter values by varying $\vartheta\in\{0.1,0.3,0.5\}$. The benchmark is chosen as the sample average approximation counterpart of our robust formulation. We repeat the above out-of-sample testing procedure 30 times, each of which has an independent instance, and we report the resulting boxplots in Figure 4.

The left column of Figure 4 shows the out-of-sample expected utility of the optimal SAA solution and the optimal robust solution for different choices of T and $\widehat{N}_t$. We have the following observations:

(I) As $\widehat{N}_t$ increases, both the SAA solution and robust solutions achieve a higher out-of-sample expected utility. This makes sense because a larger sample size yields a more faithful representation of the underlying stochastic process.

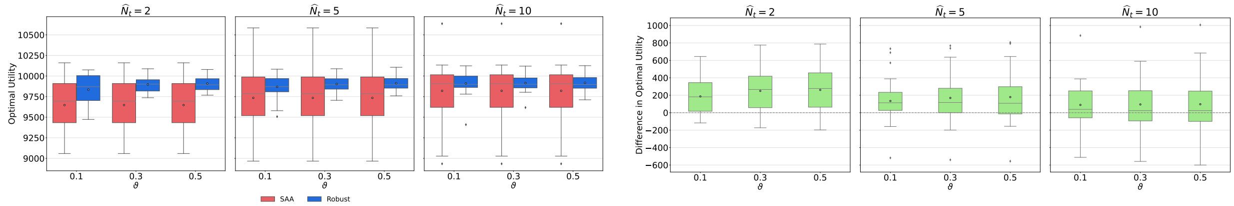
(II) The average out-of-sample performance (as indicated by the circles) of the robust solution is consistently better than that of the SAA solution, and the variability of the out-of-sample performance (as indicated by the box length) of the robust solutions is consistently smaller than that of the SAA solution. This shows the practical importance of having a robust formulation to achieve better out-of-sample performance.

(III) When $\widehat{N}_t=2,5$, a large radius $\vartheta=0.5$ has the best out-of-sample performance; whereas when $\widehat{N}_t=10$, a large radius does not have clear advantage anymore.

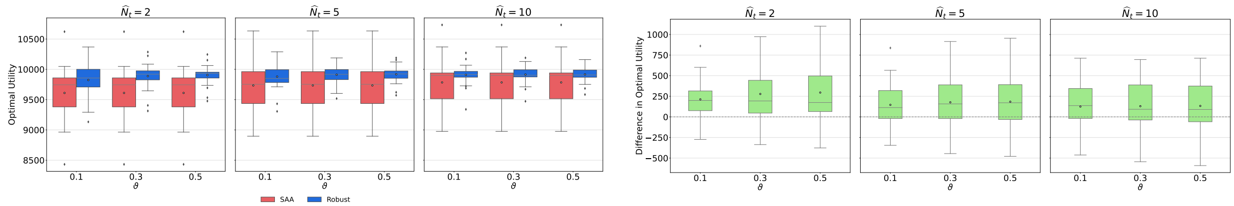
These observations are consistent with our intuition and hold for all choices of T.

To further investigate the impact of the sample size and the radius on the out-of-sample performance, on the right column of Figure 4, we plot the instance-wise difference between SAA and robust solutions. A positive value means the robust solution performs better than the SAA solution out-of-sample for a particular instance. We have the following observations:

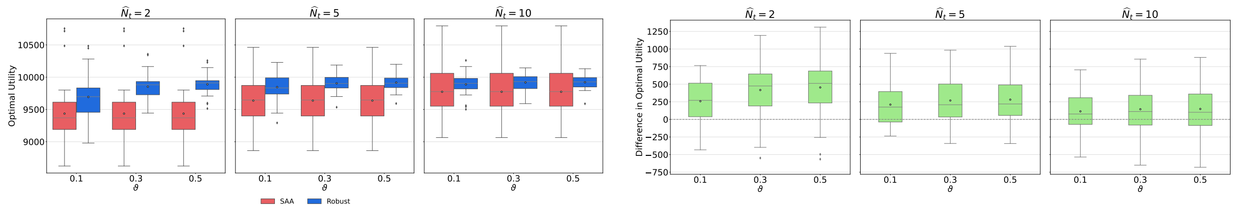
(I) The performance of the robust solution has a clear advantage over the SAA solution when the sample size $\widehat{N}_t$ is small, and the advantage diminishes as the sample size becomes larger.



(a) $T = 3$



(b) $T = 4$



(c) $T = 5$

Figure 4 Out-of-sample expected optimal utility yielding from robust and SAA formulations (left column) and their differences (right column)

(II) The best radius decreases as the sample size $\widehat{N}_t$ increases and increases as T increases. This makes sense because the distributionally uncertainty reduces when more sample paths are observed and amplifies when there are more stages.

These observations are also consistent with our intuition and validate the robust approach. In summary, the numerical results demonstrate the clear advantage of our robust formulation as compared to the sample average approximation.

6. Concluding Remarks

In this paper, we develop reformulations for distributionally robust optimization with nested distance. These reformulations unveil equivalence between static and dynamic formulations of multistage distributionally robust problem and can be viewed as sample average approximation with norm regularization. For the future work, it is interesting to study multistage problems with general convex objective and constraints, as well as the finite-sample performance guarantees of the robust solution.

References

Bita Analui and Georg Ch Pflug. On distributionally robust multiperiod stochastic optimization. *Computational Management Science*, 11(3):197–220, 2014.

Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, David Heath, and Hyejin Ku. Coherent multiperiod risk adjusted values and bellman's principle. *Annals of Operations Research*, 152(1):5–22, 2007.

Julio Backhoff, Mathias Beiglbock, Yiqing Lin, and Anastasiia Zalashko. Causal transport in discrete time and applications. *SIAM Journal on Optimization*, 27(4):2528–2562, 2017.

Julio Backhoff, Daniel Bartl, Mathias Beiglböck, and Johannes Wiesel. Estimating processes in adapted wasserstein distance. *The Annals of Applied Probability*, 32(1):529–550, 2022.

Julio Backhoff-Veraguas, Daniel Bartl, Mathias Beiglböck, and Manu Eder. Adapted wasserstein distances and stability in mathematical finance. *Finance and Stochastics*, 24(3):601–632, 2020.

Güzin Bayraksan and David K Love. Data-driven stochastic programming using phi-divergences. In *The Operations Research Revolution*, pages 1–19. INFORMS, 2015.

Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*. Princeton University Press, 2009.

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.

Dimitris Bertsimas, Melvyn Sim, and Meilin Zhang. Adaptive distributionally robust optimization. *Management Science*, 65(2):604–618, 2019.

Dimitris Bertsimas, Shimrit Shtern, and Bradley Sturt. A data-driven approach to multi-stage stochastic linear optimization. *Management Science (Forthcoming)*, 2022.

John R. Birge. Decomposition and Partitioning Methods for Multistage Stochastic Linear Programs. *Operations Research*, 33(5):989–1007, 1985.

Jose Blanchet, Yang Kang, and Karthyek Murthy. Robust Wasserstein profile inference and applications to machine learning. *Journal of Applied Probability*, 56(3):830–857, 2019.

Filipe Goulart Cabral and Bernardo Freitas Paulo da Costa. Nested distance for stagewise-independent processes. *arXiv preprint arXiv:1711.10633*, 2017.

Zhiping Chen and Zhe Yan. Scenario tree reduction methods through clustering nodes. *Computers & Chemical Engineering*, 109:96–111, 2018.

Patrick Cheridito and Michael Kupper. Recursiveness of indifference prices and translation-invariant preferences. *Mathematics and Financial Economics*, 2(3):173–188, 2009.

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

John C Duchi and Hongseok Namkoong. Variance-based regularization with convex objectives. *Journal of Machine Learning Research*, 20(68):1–55, 2019.

John C Duchi, Peter W Glynn, and Hongseok Namkoong. Statistics of robust optimization: A generalized empirical likelihood approach. *Mathematics of Operations Research*, (forthcoming), 2021.

J. Dupačová, N. Gröwe-Kuska, and W. Römisch. Scenario reduction in stochastic programming an approach using probability metrics. *Mathematical Programming, Series B*, 95:493–511, 3 2003.

Jitka Dupačová, Giorgio Consigli, and Stein W. Wallace. Scenarios for multistage stochastic programs. *Ann. Oper. Res.*, 100:25–53, 2000.

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

Daniel Duque, Sanjay Mehrotra, and David P. Morton. Distributionally robust two-stage stochastic programming. *SIAM Journal on Optimization*, 32(3):1499–1522, 2022.

Larry G Epstein and Martin Schneider. Recursive multiple-priors. *Journal of Economic Theory*, 113(1):1–31, 2003.

Johanna Etner, Meglena Jeleva, and Jean-Marc Tallon. Decision theory under ambiguity. *Journal of Economic Surveys*, 26(2):234–270, 2012.

Hans Föllmer and Alexander Schied. Convex measures of risk and trading constraints. *Finance and stochastics*, 6 (4):429–447, 2002.

Rui Gao. Finite-sample guarantees for Wasserstein distributionally robust optimization: Breaking the curse of dimensionality. *Operations Research (Forthcoming)*, 2022.

Rui Gao, Xi Chen, and Anton J Kleywegt. Wasserstein distributionally robust optimization and variation regularization. *Operations Research*, 2022.

Martin Glanzer, Georg Ch Pflug, and Alois Pichler. Incorporating statistical model error into the calculation of acceptability prices of contingent claims. *Mathematical Programming*, 174(1):499–524, 2019.

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

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.

LarsPeter Hansen and Thomas J Sargent. Robust control and model uncertainty. *American Economic Review*, 91 (2):60–66, 2001.

R. Henrion and W. Römisch. Problem-based optimal scenario generation and reduction in stochastic programming. *Mathematical Programming*, 191:183–205, 1 2022.

Markéta Horejšová, Sebastiano Vitali, Miloš Kopa, and Vittorio Moriggia. Evaluation of scenario reduction algorithms with nested distance. *Computational Management Science*, 17(2):241–275, 2020.

Jianqiu Huang, Kezhuo Zhou, and Yongpei Guan. A study of distributionally robust multistage stochastic optimization. *arXiv preprint arXiv:1708.07930*, 2017.

Kjetil Høyland and Stein W. Wallace. Generating scenario trees for multi-stage decision problems. *Manage. Sci.*, 47:295–307, 2001.

Dan A Iancu, Marek Petrik, and Dharmashankar Subramanian. Tight approximations of dynamic risk measures. *Mathematics of Operations Research*, 40(3):655–682, 2015.

Garud N Iyengar. Robust dynamic programming. *Mathematics of Operations Research*, 30(2):257–280, 2005.

Julien Keutchayan, Michel Gendreau, and Antoine Saucier. Quality evaluation of scenario-tree generation methods for solving stochastic programming problems. *Computational Management Science*, 14:333–365, 7 2017.

Diego Klabjan, David Simchi-Levi, and Miao Song. Robust stochastic lot-sizing by means of histograms. *Production and Operations Management*, 22(3):691–710, 2013.

Raimund M Kovacevic and Alois Pichler. Tree approximation for discrete time stochastic processes: a process distance approach. *Annals of operations research*, 235(1):395–421, 2015.

Daniel Kuhn, Peyman Mohajerin Esfahani, Viet Anh Nguyen, and Soroosh Shafieezadeh-Abadeh. Wasserstein distributionally robust optimization: Theory and applications in machine learning. In *Operations Research & Management Science in the Age of Analytics*, pages 130–166. INFORMS, 2019.

Henry Lam. Robust sensitivity analysis for stochastic systems. *Mathematics of Operations Research*, 41(4): 1248–1275, 2016.

Henry Lam. Recovering best statistical guarantees via the empirical divergence-based distributionally robust optimization. *Operations Research*, 67(4):1090–1105, 2019.

Guanghui Lan and Alexander Shapiro. Numerical Methods for Convex Multistage Stochastic Optimization, March 2023.

Rémi Lassalle. Causal transport plans and their monge–kantorovich problems. *Stochastic Analysis and Applications*, 36(3):452–484, 2018.

Francesca Maggioni and Georg Ch Pflug. Bounds and approximations for multistage stochastic programs. *SIAM Journal on Optimization*, 26(1):831–855, 2016.

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.

Arnab Nilim and Laurent El Ghaoui. Robust control of markov decision processes with uncertain transition matrices. *Operations Research*, 53(5):780–798, 2005.

Jangho Park and Guzin Bayraksan. A multistage distributionally robust optimization approach to water allocation under climate uncertainty. *arXiv preprint arXiv:2005.07811*, 2020.

M. V. F. Pereira and L. M. V. G. Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52(1):359–375, May 1991.

G Ch Pflug. Version-independence and nested distributions in multistage stochastic optimization. *SIAM Journal on Optimization*, 20(3):1406–1420, 2010.

Georg Ch Pflug and Alois Pichler. A distance for multistage stochastic optimization models. *SIAM Journal on Optimization*, 22(1):1–23, 2012.

Georg Ch Pflug and Alois Pichler. *Multistage stochastic optimization*, volume 1104. Springer, 2014.

Georg Ch Pflug and Alois Pichler. Dynamic generation of scenario trees. *Computational Optimization and Applications*, 62:641–668, 2015.

Georg Ch. Pflug and Alois Pichler. From Empirical Observations to Tree Models for Stochastic Optimization: Convergence Properties. *SIAM Journal on Optimization*, 26(3):1715–1740, January 2016.

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

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

Alois Pichler and Michael Weinhardt. The nested sinkhorn divergence to learn the nested distance. *Computational Management Science*, 19(2):269–293, 2022.

Alois Pichler, Rui Peng Liu, and Alexander Shapiro. Risk-averse stochastic programming: Time consistency and optimal stopping. *Operations Research*, 2021.

Hamed Rahimian, Guzin Bayraksan, and Tito Homem-de Mello. Effective scenarios in multistage distributionally robust optimization with a focus on total variation distance. *arXiv preprint arXiv:2109.06791*, 2021.

Berend Roorda and Johannes M Schumacher. Time consistency conditions for acceptability measures, with an application to tail value at risk. *Insurance: Mathematics and Economics*, 40(2):209–230, 2007.

Andrzej Ruszczyński and Alexander Shapiro. Conditional risk mappings. *Mathematics of operations research*, 31 (3):544–561, 2006.

Soroosh Shafieezadeh-Abadeh, Daniel Kuhn, and Peyman Mohajerin Esfahani. Regularization via mass transportation. *Journal of Machine Learning Research*, 20(103):1–68, 2019.

Alexander Shapiro. Inference of statistical bounds for multistage stochastic programming problems. *Mathematical Methods of Operations Research*, 58:57–68, 9 2003a.

Alexander Shapiro. Monte carlo sampling methods. *Handbooks in Operations Research and Management Science*, 10:353–425, 1 2003b.

Alexander Shapiro. Time consistency of dynamic risk measures. *Operations Research Letters*, 40(6):436–439, 2012.

Alexander Shapiro. Rectangular sets of probability measures. *Operations Research*, 64(2):528–541, 2016.

Alexander Shapiro and Arkadi Nemirovski. On complexity of stochastic programming problems. In *Continuous optimization*, pages 111–146. Springer, 2005.

Alexander Shapiro and Alois Pichler. Conditional distributionally robust functionals. *Available on Optimization Online*, 2022.

Alexander Shapiro and Linwei Xin. Time inconsistency of optimal policies of distributionally robust inventory models. *Operations Research*, 68(5):1576–1584, 2020.

Alexander Shapiro, Wajdi Tekaya, Joari Paulo Da Costa, and Murilo Pereira Soares. Final report for technical cooperation between georgia institute of technology and ons-operador nacional do sistema elétrico. 2012.

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

Robert Henry Strotz. Myopia and inconsistency in dynamic utility maximization. *The review of economic studies*, 23(3):165–180, 1955.

Bradley Sturt. A nonparametric algorithm for optimal stopping based on robust optimization. *Operations Research*, 5 2023.

Dávid Terjék. Virtual adversarial lipschitz regularization. *arXiv preprint arXiv:1907.05681*, 2019.

Julio Backhoff Veraguas, Mathias Beiglböck, Manu Eder, and Alois Pichler. Fundamental properties of process distances. *Stochastic Processes and their Applications*, 130(9):5575–5591, 2020.

Tan Wang. A class of dynamic risk measures. *University of British Columbia*, 21:58, 1999.

Wolfram Wiesemann, Daniel Kuhn, and Berç Rustem. Robust markov decision processes. *Mathematics of Operations Research*, 38(1):153–183, 2013.

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

Weijun Xie. Tractable reformulations of two-stage distributionally robust linear programs over the type-∞ Wasserstein ball. *Operations Research Letters*, 48(4):513–523, July 2020.

Linwei Xin and David A Goldberg. Time (in) consistency of multistage distributionally robust inventory models with moment constraints. *European Journal of Operational Research*, 289(3):1127–1141, 2021.

Linwei Xin and David Alan Goldberg. Distributionally robust inventory control when demand is a martingale. *Mathematics of Operations Research*, 2022.

Toshio Yamada and Shinzo Watanabe. On the uniqueness of solutions of stochastic differential equations. *Journal of Mathematics of Kyoto University*, 11(1):155–167, 1971.

Jincheng Yang, Luhao Zhang, Ningyuan Chen, Rui Gao, and Ming Hu. Decision-making with side information: A causal transport robust approach. *Available on Optimization Online*, 2022.

Xian Yu and Siqian Shen. Multistage distributionally robust mixed-integer programming with decision-dependent moment-based ambiguity sets. *Mathematical Programming*, pages 1–40, 2020.

Luhao Zhang, Jincheng Yang, and Rui Gao. A simple duality proof for wasserstein distributionally robust optimization. *arXiv preprint arXiv:2205.00362*, 2022.

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

Shixuan Zhang and Xu Andy Sun. On distributionally robust multistage convex optimization: Data-driven models and performance. *arXiv preprint arXiv:2210.08433*, 2022.

Supplementary Material

EC.1. Proofs for Section 3.1

We consider a relaxation

$$\sup_{\mathbb{P}\in\mathfrak{M}^{\mathsf{C}}} \mathbb{E}_{\mathbb{P}}\left[Z^{\pi}(\boldsymbol{\xi}_{[T]})\right], \tag{EC.1}$$

where the nested distance uncertainty set (3) is relaxed to the causal transport distance uncertainty set

$$\mathfrak{M}^{\mathsf{C}} = \left\{\mathbb{P}\in\mathcal{P}(\Xi):\ \mathsf{C}_p(\widehat{\mathbb{P}},\mathbb{P})\le\vartheta\right\}.$$

EC.1.1. Proof of Theorem 1

We have the following result for the relaxed problem (EC.1).

PROPOSITION EC.1. *Let $p=\infty$. Then for any feasible policy $(\boldsymbol{x}_1,\dots,\boldsymbol{x}_T)$, it holds that*

$$\begin{aligned}&\sup_{\mathbb{P}\in\mathfrak{M}^{\mathsf{C}}} \mathbb{E}_{\mathbb{P}}\left[\sum_{s\in[T]}\boldsymbol{c}_s^\top\boldsymbol{x}_s(\boldsymbol{\xi}_{[s]})\right]\\&=\sup_{\gamma_{[t]}\in\Gamma^{\mathsf{C}}_{[t]}}\mathbb{E}_{\gamma_{[t]}}\left[\sum_{s\in[t]}\boldsymbol{c}_s^\top\boldsymbol{x}_s(\boldsymbol{\xi}_{[s]})+\mathbb{E}_{\widehat{\mathbb{P}}_{t+1|\widehat{\boldsymbol{\xi}}_{[t]}}}\left[\sup_{\xi_{t+1}\in\Xi_{t+1}:\|\xi_{t+1}-\widehat{\boldsymbol{\xi}}_{t+1}\|\le\vartheta}V_{t+1|\widehat{\boldsymbol{\xi}}_{[t+1]}}(\boldsymbol{\xi}_{[t]},\xi_{t+1})\right]\right].\end{aligned}$$

Proof of Proposition EC.1. We prove by induction. By definition, the case of $t=T$ holds trivially. Suppose we have proved for the case of t, where $t=2,\dots,T$, now we prove the result also holds for $t-1$. Denote by $\Gamma^{\mathsf{C}}_{[t]}$ the set of all causal couplings on $\Xi_{[t]}$ between $\widehat{\mathbb{P}}_{[t]}$ and the distributions in $\mathfrak{M}^{\mathsf{C}}_{[t]} := \{\mathbb{P}_{[t]}\in\mathcal{P}(\Xi_{[t]}):\ \mathsf{C}_\infty(\widehat{\mathbb{P}}_{[t]},\mathbb{P}_{[t]})\le\vartheta\}$. According to the induction hypothesis, we have that

$$\begin{aligned}V &:= \sup_{\mathbb{P}\in\mathfrak{M}^{\mathsf{C}}} \mathbb{E}_{\mathbb{P}}\left[\sum_{s\in[T]}\boldsymbol{c}_s^\top\boldsymbol{x}_s(\boldsymbol{\xi}_{[s]})\right]\\&=\sup_{\gamma_{[t]}\in\Gamma^{\mathsf{C}}_{[t]}}\mathbb{E}_{\gamma_{[t]}}\left[\sum_{s\in[t]}\boldsymbol{c}_s^\top\boldsymbol{x}_s(\boldsymbol{\xi}_{[s]})+\mathbb{E}_{\widehat{\mathbb{P}}_{t+1|\widehat{\boldsymbol{\xi}}_{[t]}}}\left[\sup_{\xi_{t+1}\in\Xi_{t+1}:\|\xi_{t+1}-\widehat{\boldsymbol{\xi}}_{t+1}\|\le\vartheta}V_{t+1|\widehat{\boldsymbol{\xi}}_{[t+1]}}(\boldsymbol{\xi}_{[t]},\xi_{t+1})\right]\right]\\&=: \sup_{\gamma_{[t]}\in\Gamma^{\mathsf{C}}_{[t]}}\mathbb{E}_{\gamma_{[t]}}\left[\sum_{s\in[t]}\boldsymbol{c}_s^\top\boldsymbol{x}_s(\boldsymbol{\xi}_{[s]})+\bar{V}_{t+1|\widehat{\boldsymbol{\xi}}_{[t]}}(\boldsymbol{\xi}_{[t]})\right].\end{aligned}$$

For any causal transport plan $\gamma_{[t]}$ between $\widehat{\mathbb{P}}_{[t]}$ and $\mathbb{P}_{[t]}$, by the tower property of conditional expectation, it holds that

$$\begin{aligned}
&\mathbb{E}_{\gamma_{[t]}}\left[\sum_{s\in[t]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) + \bar{V}_{t+1|\widehat{\boldsymbol{\xi}}_{[t]}}(\boldsymbol{\xi}_{[t]})\right] \\
&= \mathbb{E}_{\gamma_{[t-1]}}\left[\sum_{s\in[t-1]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]})\right] + \mathbb{E}_{\gamma_{[t]}}\left[\boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) + \bar{V}_{t+1|\widehat{\boldsymbol{\xi}}_{[t]}}(\boldsymbol{\xi}_{[t]})\right] \\
&= \mathbb{E}_{\gamma_{[t-1]}}\left[\sum_{s\in[t-1]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]})\right] + \mathbb{E}_{\gamma_{[t-1]}}\left[\mathbb{E}_{\gamma_{\widehat{\boldsymbol{\xi}}_t|(\widehat{\boldsymbol{\xi}}_{[t-1]},\boldsymbol{\xi}_{[t-1]})}}\left[\mathbb{E}_{\gamma_{\boldsymbol{\xi}_t|(\widehat{\boldsymbol{\xi}}_{[t]},\boldsymbol{\xi}_{[t-1]})}}\left[\boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) + \bar{V}_{t+1|\widehat{\boldsymbol{\xi}}_{[t]}}(\boldsymbol{\xi}_{[t]})\right]\right]\right] \\
&= \mathbb{E}_{\gamma_{[t-1]}}\left[\sum_{s\in[t-1]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]})\right] + \mathbb{E}_{\gamma_{[t-1]}}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{t|\widehat{\boldsymbol{\xi}}_{[t-1]}}}\left[\mathbb{E}_{\gamma_{\boldsymbol{\xi}_t|(\widehat{\boldsymbol{\xi}}_{[t]},\boldsymbol{\xi}_{[t-1]})}}\left[\boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) + \bar{V}_{t+1|\widehat{\boldsymbol{\xi}}_{[t]}}(\boldsymbol{\xi}_{[t]})\right]\right]\right],
\end{aligned} \tag{EC.2}$$

where, in the last equations above, we have used the property of the causal transport plan $\gamma_{\widehat{\boldsymbol{\xi}}_t|(\widehat{\boldsymbol{\xi}}_{[t-1]},\boldsymbol{\xi}_{[t-1]})} = \widehat{\mathbb{P}}_{\widehat{\boldsymbol{\xi}}_t|\widehat{\boldsymbol{\xi}}_{[t-1]}}$. Note that given the distribution $\widehat{\mathbb{P}}_{[t]}$, the causal transport plan $\gamma_{[t]}$ is determined completely by $\gamma_{[t-1]}$ and $\gamma_{\boldsymbol{\xi}_t|(\widehat{\boldsymbol{\xi}}_{[t]},\boldsymbol{\xi}_{[t-1]})}$, and that by definition, the constraint

$$\gamma_{[t]}\text{-}\operatorname*{ess\,sup}_{\widehat{\boldsymbol{\xi}}_{[t]},\boldsymbol{\xi}_{[t]}\in\Xi_{[t]}} \max_{s\in[t]} \|\widehat{\boldsymbol{\xi}}_s - \boldsymbol{\xi}_s\| \le \vartheta$$

is equivalent to

$$\gamma_{[t-1]}\text{-}\operatorname*{ess\,sup}_{\widehat{\boldsymbol{\xi}}_{[t-1]},\boldsymbol{\xi}_{[t-1]}\in\Xi_{[t-1]}} \max_{s\in[t-1]} \|\widehat{\boldsymbol{\xi}}_s - \boldsymbol{\xi}_s\| \le \vartheta$$

and

$$\gamma_{\boldsymbol{\xi}_t|(\widehat{\boldsymbol{\xi}}_{[t]},\boldsymbol{\xi}_{[t-1]})}\text{-}\operatorname*{ess\,sup}_{\boldsymbol{\xi}_t\in\Xi_t} \|\widehat{\boldsymbol{\xi}}_t - \boldsymbol{\xi}_t\| \le \vartheta, \quad \forall \widehat{\boldsymbol{\xi}}_{[t]} \in \operatorname{supp}\widehat{\mathbb{P}}_{[t]}, \boldsymbol{\xi}_{[t-1]} \in \operatorname{supp}\mathbb{P}_{[t]}.$$

Thereby maximizing over $\gamma_{[t]} \in \Gamma^{\mathsf{C}}_{[t]}$ is equivalent to maximizing over $\gamma_{[t-1]} \in \Gamma^{\mathsf{C}}_{[t-1]}$ and $\gamma_{\boldsymbol{\xi}_t|(\widehat{\boldsymbol{\xi}}_{[t]},\boldsymbol{\xi}_{[t-1]})} \in \{\mathbb{P}_t \in \mathcal{P}(\Xi_t) : \mathbb{P}_t\text{-}\operatorname*{ess\,sup}_{\xi_t\in\Xi_t}\|\xi_t - \widehat{\boldsymbol{\xi}}_t\| \le \vartheta\} =: \mathfrak{M}_t$. Observe that

$$\begin{aligned}
&\sup_{\gamma_{\boldsymbol{\xi}_t|(\widehat{\boldsymbol{\xi}}_{[t]},\boldsymbol{\xi}_{[t-1]})}\in\mathfrak{M}_t} \mathbb{E}_{\gamma_{\boldsymbol{\xi}_t|(\widehat{\boldsymbol{\xi}}_{[t]},\boldsymbol{\xi}_{[t-1]})}}\left[\boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) + \bar{V}_{t+1|\widehat{\boldsymbol{\xi}}_{[t]}}(\boldsymbol{\xi}_{[t]})\right] \\
= &\sup_{\xi_t\in\Xi_t:\|\xi_t-\widehat{\boldsymbol{\xi}}_t\|\le\vartheta} \left\{c_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t-1]},\xi_t) + \bar{V}_{t+1|\widehat{\boldsymbol{\xi}}_{[t]}}(\boldsymbol{\xi}_{[t-1]},\xi_t)\right\} \\
= &\sup_{\xi_t\in\Xi_t:\|\xi_t-\widehat{\boldsymbol{\xi}}_t\|\le\vartheta} V_{t|\widehat{\boldsymbol{\xi}}_{[t]}}(\boldsymbol{\xi}_{[t-1]},\xi_t).
\end{aligned}$$

Hence, together with (EC.2), it follows that

$$V = \sup_{\gamma_{[t-1]}\in\Gamma^{\mathsf{C}}_{[t-1]}} \mathbb{E}_{\gamma_{[t-1]}}\left[\sum_{s\in[t-1]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t|\widehat{\boldsymbol{\xi}}_{[t-1]}}}\left[\sup_{\xi_t\in\Xi_t:\|\xi_t-\widehat{\boldsymbol{\xi}}_t\|\le\vartheta} V_{t|\widehat{\boldsymbol{\xi}}_{[t]}}(\boldsymbol{\xi}_{[t-1]},\xi_t)\right]\right],$$

which completes the induction. □

Proof of Theorem 1. We prove the theorem by showing that for every policy $(\boldsymbol{x}_1(\boldsymbol{\xi}_1), \cdots, \boldsymbol{x}_T(\boldsymbol{\xi}_{[T]})) \in \Pi$, it holds that

$$\sup_{\mathbb{P}\in\mathfrak{M}} \mathbb{E}_{\mathbb{P}}\left[\sum_{t\in[T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\right] = \sup_{\mathbb{P}\in\mathfrak{M}^{\mathsf{C}}} \mathbb{E}_{\mathbb{P}}\left[\sum_{t\in[T]} \boldsymbol{c}_s^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\right].$$

Since $\mathfrak{M} \subset \mathfrak{M}^{\mathsf{C}}$, the left-hand side is less than or equal to the right-hand side. It remains to prove the other direction. Applying Proposition EC.1 to a special case where $\mathcal{X}_t(x_{t-1}, \boldsymbol{\xi}_t) = \boldsymbol{x}_t(\boldsymbol{\xi}_t)$, the right-hand side above equals

$$\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{\xi}_2}}\left[\sup_{\xi_2 \in \Xi_2 : \|\widehat{\xi}_2 - \xi_2\| \le \vartheta} \left\{ c_2^\top x_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{\xi}_3|\widehat{\xi}_2}}\left[\sup_{\xi_3 \in \Xi_3 : \|\widehat{\xi}_3 - \xi_3\| \le \vartheta} \left\{ c_3^\top x_3(\xi_3) + \cdots + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{\xi}_T|\widehat{\xi}_{[T-1]}}}\left[\sup_{\xi_T \in \Xi_T : \|\widehat{\xi}_T - \xi_T\| \le \vartheta} c_T^\top x_T(\xi_T)\right]\right\}\right]\right\}\right].$$

It follows that there exists a worst-case distribution $\tilde{\mathbb{P}}$ of the form

$$\tilde{\mathbb{P}} = \mathbb{T}_\# \widehat{\mathbb{P}}, \text{ where } \mathbb{T}(\widehat{\xi}_{[T]}) = (\mathbb{T}_1(\widehat{\xi}_1), \mathbb{T}_2(\widehat{\xi}_{[2]}), \ldots, \mathbb{T}_T(\widehat{\xi}_{[T]})),\ \forall \widehat{\xi}_{[T]} \in \Xi_{[T]},$$

where

$$\mathbb{T}_t(\widehat{\xi}_{[t]}) \in \underset{\xi_t \in \Xi_t : \|\widehat{\xi}_t - \xi_t\| \le \vartheta}{\arg\max} \left\{ c_t^\top x_t(\xi_t) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1|\widehat{\xi}_{[t]}}}[Q_{t+1}(x_t(\xi_t) \mid \widehat{\boldsymbol{\xi}}_{[t+1]})] \right\}.$$

Since $\tilde{\mathbb{P}}$ is finitely supported, for every $\epsilon > 0$, we can define a perturbed transport map $\mathbb{T}^\epsilon = (\mathbb{T}_1^\epsilon, \ldots, \mathbb{T}_T^\epsilon)$, where $\mathbb{T}_t^\epsilon : \Xi_{[t]} \to \Xi_t$, such that $\mathbb{T}_t^\epsilon$ is bijective and $\|\mathbb{T}_t^\epsilon - \mathbb{T}_t\|_\infty \le \epsilon$, $t \in [T]$. For example, if $\mathbb{T}_t(\xi_{[t]}) = \mathbb{T}_t(\xi'_{[t]})$ for $\xi_{[t]} \ne \xi'_{[t]}$, we can slightly perturb the value of $\mathbb{T}_t(\xi'_{[t]})$ to make them different. This makes $\mathbb{T}^\epsilon$ a bi-causal transport map as $(\mathbb{T}^\epsilon)^{-1} := ((\mathbb{T}_1^\epsilon)^{-1}, \ldots, (\mathbb{T}_T^\epsilon)^{-1})$ is causal. Define

$$\mathbb{P}^\epsilon := (1 - \tfrac{\epsilon}{\vartheta+\epsilon})\mathbb{T}^\epsilon_\# \widehat{\mathbb{P}} + \tfrac{\epsilon}{\vartheta+\epsilon}\widehat{\mathbb{P}}.$$

It follows that

$$\begin{aligned}
\mathsf{D}_\infty(\widehat{\mathbb{P}}, \mathbb{P}^\epsilon) &\le (1 - \tfrac{\epsilon}{\vartheta+\epsilon}) \sup_{\widehat{\xi}_{[T]} \in \operatorname{supp} \widehat{\mathbb{P}}} \|\widehat{\xi}_{[T]} - \mathbb{T}^\epsilon(\widehat{\xi}_{[T]})\| \\
&\le (1 - \tfrac{\epsilon}{\vartheta+\epsilon}) \sup_{\widehat{\xi}_{[T]} \in \operatorname{supp} \widehat{\mathbb{P}}} (\|\widehat{\xi}_{[T]} - \mathbb{T}(\widehat{\xi}_{[T]})\| + \epsilon) \\
&\le (1 - \tfrac{\epsilon}{\vartheta+\epsilon})(\mathsf{C}_\infty(\widehat{\mathbb{P}}, \tilde{\mathbb{P}}) + \epsilon) \\
&\le \vartheta.
\end{aligned}$$

Hence $\mathbb{P}^\epsilon$ is a feasible distribution. Define a compact set

$$\widehat{\Xi}_\vartheta := \bigcup_{\widehat{\xi} \in \operatorname{supp} \widehat{\mathbb{P}}} \operatorname{cl}\{\xi \in \Xi :\ \|\xi - \widehat{\xi}\| \le \vartheta\},$$

where $\operatorname{cl}(\cdot)$ denotes the closure of a set. Then by definition of $\mathfrak{M}$ and ∞-Wasserstein distance, we have $\operatorname{supp} \mathbb{P} \subset \widehat{\Xi}_\vartheta$ for every $\mathbb{P} \in \mathfrak{M}$. It follows that

$$\begin{aligned}
&\sup_{\mathbb{P} \in \mathfrak{M}} \mathbb{E}_{\mathbb{P}}\left[\sum_{t \in [T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\right] \\
&\ge \mathbb{E}_{\mathbb{P}^\epsilon}\left[\sum_{t \in [T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\right] \\
&= (1 - \tfrac{\epsilon}{\vartheta+\epsilon})\mathbb{E}_{\widehat{\mathbb{P}}}\left[\sum_{t \in [T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\mathbb{T}^\epsilon(\widehat{\boldsymbol{\xi}}_{[t]}))\right] + \tfrac{\epsilon}{\vartheta+\epsilon}\mathbb{E}_{\widehat{\mathbb{P}}}\left[\sum_{t \in [T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\right] \\
&\ge \mathbb{E}_{\tilde{\mathbb{P}}}\left[\sum_{t \in [T]} \boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]})\right], \text{ as } \epsilon \to 0,
\end{aligned}$$

where the last step holds because $\boldsymbol{x}$ is continuous. □

EC.1.2. Proof of Theorem 2

PROPOSITION EC.2. *Let $p \in [1,\infty)$. Then for any feasible policy $(\boldsymbol{x}_1,\dots,\boldsymbol{x}_T)$, it holds that*

$$\begin{aligned}&\sup_{\mathbb{P}\in\mathfrak{M}^{\mathsf{C}}} \mathbb{E}_{\mathbb{P}}\left[\sum_{s\in[T]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]})\right]\\ &= \sup_{\gamma_{[t]}\in\Gamma^{\mathsf{C}}_{[t]}} \mathbb{E}_{\gamma_{[t]}}\left[\sum_{s\in[t]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1|\widehat{\xi}_{[t]}}}\left[\sup_{\xi_{t+1}\in\Xi_{t+1}}\left\{V_{t+1|\widehat{\xi}_{[t+1]}}(\boldsymbol{\xi}_{[t]},\xi_{t+1}) - \lambda\|\xi_{t+1}-\widehat{\boldsymbol{\xi}}_{t+1}\|\right\}\right]\right].\end{aligned}$$

Proof of Proposition EC.2. We prove by induction. By definition, the case of $t=T$ holds trivially. Suppose we have proved for the case of t, where $t=2,\dots,T$, now we prove the result also holds for $t-1$. Denote by $\Gamma^{\mathsf{C}}_{[t]}$ the set of all causal coupling on $\Xi_{[t]}$ between $\widehat{\mathbb{P}}_{[t]}$ and any distribution on $\mathcal{P}(\Xi_{[t]})$. According to the induction hypothesis, we have that

$$\begin{aligned}V &:= \sup_{\mathbb{P}\in\mathfrak{M}^{\mathsf{C}}} \mathbb{E}_{\mathbb{P}}\left[\sum_{s\in[T]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]})\right]\\ &= \sup_{\gamma_{[t]}\in\Gamma^{\mathsf{C}}_{[t]}} \mathbb{E}_{\gamma_{[t]}}\left[\sum_{s\in[t]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) - \lambda\sum_{s\in[t]}\|\boldsymbol{\xi}_s-\widehat{\boldsymbol{\xi}}_s\|^p + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1|\widehat{\xi}_{[t]}}}\left[\sup_{\xi_{t+1}\in\Xi_{t+1}}\left\{V_{t+1|\widehat{\xi}_{[t+1]}}(\boldsymbol{\xi}_{[t]},\xi_{t+1}) - \lambda\|\xi_{t+1}-\widehat{\boldsymbol{\xi}}_{t+1}\|^p\right\}\right]\right]\\ &=: \sup_{\gamma_{[t]}\in\Gamma^{\mathsf{C}}_{[t]}} \mathbb{E}_{\gamma_{[t]}}\left[\sum_{s\in[t]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) - \lambda\sum_{s\in[t]}\|\boldsymbol{\xi}_s-\widehat{\boldsymbol{\xi}}_s\|^p + \bar{V}_{t+1|\widehat{\xi}_{[t]}}(\boldsymbol{\xi}_{[t]})\right]\end{aligned}$$

For any causal transport plan $\gamma_{[t]}$ between $\widehat{\mathbb{P}}_{[t+1]}$ and $\mathbb{P}_{[t+1]}$, by the tower property of conditional expectation, it holds that

$$\begin{aligned}&\mathbb{E}_{\gamma_{[t]}}\left[\sum_{s\in[t]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) - \lambda\sum_{s\in[t]}\|\boldsymbol{\xi}_s-\widehat{\boldsymbol{\xi}}_s\|^p + \bar{V}_{t+1|\widehat{\xi}_{[t]}}(\boldsymbol{\xi}_{[t]})\right]\\ &= \mathbb{E}_{\gamma_{[t-1]}}\left[\sum_{s\in[t-1]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) - \lambda\sum_{s\in[t-1]}\|\boldsymbol{\xi}_s-\widehat{\boldsymbol{\xi}}_s\|^p\right] + \mathbb{E}_{\gamma_{[t]}}\left[\boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) - \lambda\|\boldsymbol{\xi}_t-\widehat{\boldsymbol{\xi}}_t\|^p + \bar{V}_{t+1|\widehat{\xi}_{[t]}}(\boldsymbol{\xi}_{[t]})\right]\\ &= \mathbb{E}_{\gamma_{[t-1]}}\left[\sum_{s\in[t-1]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) - \lambda\sum_{s\in[t-1]}\|\boldsymbol{\xi}_s-\widehat{\boldsymbol{\xi}}_s\|^p\right]\\ &\quad + \mathbb{E}_{\gamma_{[t-1]}}\left[\mathbb{E}_{\gamma_{\widehat{\xi}_t|(\widehat{\xi}_{[t-1]},\xi_{[t-1]})}}\left[\mathbb{E}_{\gamma_{\xi_t|(\widehat{\xi}_{[t]},\xi_{[t-1]})}}\left[\boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) - \lambda\|\boldsymbol{\xi}_t-\widehat{\boldsymbol{\xi}}_t\|^p + \bar{V}_{t+1|\widehat{\xi}_{[t]}}(\boldsymbol{\xi}_{[t]})\right]\right]\right]\\ &= \mathbb{E}_{\gamma_{[t-1]}}\left[\sum_{s\in[t-1]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) - \lambda\sum_{s\in[t-1]}\|\boldsymbol{\xi}_s-\widehat{\boldsymbol{\xi}}_s\|^p\right]\\ &\quad + \mathbb{E}_{\gamma_{[t-1]}}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{t|\widehat{\xi}_{[t-1]}}}\left[\mathbb{E}_{\gamma_{\xi_t|(\widehat{\xi}_{[t]},\xi_{[t-1]})}}\left[\boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) - \lambda\|\boldsymbol{\xi}_t-\widehat{\boldsymbol{\xi}}_t\|^p + \bar{V}_{t+1|\widehat{\xi}_{[t]}}(\boldsymbol{\xi}_{[t]})\right]\right]\right],\end{aligned} \tag{EC.3}$$

where, in the last equations above, we have used the property of the causal transport plan $\gamma_{\widehat{\xi}_t|(\widehat{\xi}_{[t-1]},\xi_{[t-1]})} = \widehat{\mathbb{P}}_{\widehat{\xi}_t|\widehat{\xi}_{[t-1]}}$. Note that given the distribution $\widehat{\mathbb{P}}_{[t]}$, the joint distribution $\gamma_{[t]}$ is determined completely by $\gamma_{[t-1]}$ and $\gamma_{\xi_t|(\widehat{\xi}_{[t]},\xi_{[t-1]})}$. Thereby maximizing over $\gamma_{[t]}\in\Gamma^{\mathsf{C}}_{[t]}$ is equivalent to maximizing over $\gamma_{[t-1]} \in \Gamma^{\mathsf{C}}_{[t-1]}$ and $\gamma_{\xi_t | (\widehat{\xi}_{[t]}, \xi_{[t-1]})} \in \mathcal{P}(\Xi_t)$. Observe that

$$\begin{aligned}
&\sup_{\gamma_{\xi_t | (\widehat{\xi}_{[t]}, \xi_{[t-1]})} \in \mathcal{P}(\Xi_t)} \mathbb{E}_{\gamma_{\xi_t | (\widehat{\xi}_{[t]}, \xi_{[t-1]})}} \left[\boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) - \lambda \|\boldsymbol{\xi}_t - \widehat{\boldsymbol{\xi}}_t\|^p + \bar{V}_{t+1|\widehat{\xi}_{[t]}}(\boldsymbol{\xi}_{[t]}) \right] \\
&= \sup_{\xi_t \in \Xi_t} \left\{ c_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t-1]}, \xi_t) + \bar{V}_{t+1|\widehat{\xi}_{[t]}}(\boldsymbol{\xi}_{[t-1]}, \xi_t) - \lambda \|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p \right\} \\
&= \sup_{\xi_t \in \Xi_t} \left\{ V_{t|\widehat{\xi}_{[t]}}(\boldsymbol{\xi}_{[t-1]}, \xi_t) - \lambda \|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p \right\}.
\end{aligned}$$

Hence, together with (EC.2), it follows that

$$V = \sup_{\gamma_{[t-1]} \in \Gamma^{\mathsf{C}}_{[t-1]}} \mathbb{E}_{\gamma_{[t-1]}} \left[\sum_{s \in [t-1]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t|\widehat{\xi}_{[t-1]}}} \left[\sup_{\xi_t \in \Xi_t} \left\{ V_{t|\widehat{\xi}_{[t]}}(\boldsymbol{\xi}_{[t-1]}, \xi_t) - \lambda \|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p \right\} \right] \right],$$

which completes the induction. □

Proof of Theorem 2. We prove the theorem by showing that for every $(\boldsymbol{x}_1(\boldsymbol{\xi}_1), \cdots, \boldsymbol{x}_T(\boldsymbol{\xi}_{[T]}))$, it holds that

$$\sup_{\mathbb{P} \in \mathfrak{M}} \mathbb{E}_{\mathbb{P}} \left[\sum_{s \in [T]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) \right] = \sup_{\mathbb{P} \in \mathfrak{M}^{\mathsf{C}}} \mathbb{E}_{\mathbb{P}} \left[\sum_{s \in [T]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) \right].$$

Since $\mathfrak{M} \subset \mathfrak{M}^{\mathsf{C}}$, the left-hand side is less than or equal to the right-hand side. It remains to prove the other direction. Using the characterization of the extreme points of the set of causal transport plans from $\widehat{\mathbb{P}}$ (Veraguas et al. 2020), there exists a worst-case distribution of the form

$$\tilde{\mathbb{P}} = (1-q)\overline{\mathbb{T}}_\# \widehat{\mathbb{P}} + q \underline{\mathbb{T}}_\# \widehat{\mathbb{P}},$$

where $q \in [0,1]$ and $\overline{\mathbb{T}}$ and $\underline{\mathbb{T}}$ are causal transport maps. Since $\tilde{\mathbb{P}}$ is finitely supported, for every $\epsilon > 0$, we can define perturbed transport maps $\overline{\mathbb{T}}^\epsilon$ and $\underline{\mathbb{T}}^\epsilon$ such that both $\overline{\mathbb{T}}^\epsilon$ and $\underline{\mathbb{T}}^\epsilon$ are bijective and their images has no overlap. It follows that both $\overline{\mathbb{T}}^\epsilon$ and $\underline{\mathbb{T}}^\epsilon$ are bi-causal transport maps. Moreover, their mixture $\mathbb{P}^\epsilon = (1-q)\overline{\mathbb{T}}^\epsilon_\# \widehat{\mathbb{P}} + q \underline{\mathbb{T}}^\epsilon_\# \widehat{\mathbb{P}}$ is a bi-causal transport plan, since by construction, there is a causal transport map from $\mathbb{P}^\epsilon$ to $\widehat{\mathbb{P}}$, defined as

$$\mathbb{T}^\dagger(\xi) := \begin{cases} \left((\overline{\mathbb{T}}^\epsilon_1)^{-1}(\xi_1), \ldots, (\overline{\mathbb{T}}^\epsilon_T)^{-1}(\xi_{[T]})\right), & \text{if } \xi \in \operatorname{supp} \overline{\mathbb{T}}^\epsilon_\# \widehat{\mathbb{P}}, \\ \left((\underline{\mathbb{T}}^\epsilon_1)^{-1}(\xi_1), \ldots, (\underline{\mathbb{T}}^\epsilon_T)^{-1}(\xi_{[T]})\right), & \text{if } \xi \in \operatorname{supp} \underline{\mathbb{T}}^\epsilon_\# \widehat{\mathbb{P}}. \end{cases}$$

Let

$$\tilde{\mathbb{P}}^\epsilon := (1 - \tfrac{\epsilon}{\vartheta+\epsilon})\mathbb{P}^\epsilon + \tfrac{\epsilon}{\vartheta+\epsilon}\widehat{\mathbb{P}}.$$

It follows that

$$\mathsf{D}_p(\widehat{\mathbb{P}}, \tilde{\mathbb{P}}^\epsilon) \le (1 - \tfrac{\epsilon}{\vartheta+\epsilon})\mathsf{D}_p(\widehat{\mathbb{P}}, \mathbb{P}^\epsilon) \le (1 - \tfrac{\epsilon}{\vartheta+\epsilon})(\mathsf{D}_p(\widehat{\mathbb{P}}, \tilde{\mathbb{P}}) + \epsilon) \le \vartheta.$$

Hence $\mathbb{P}^\epsilon$ is a feasible distribution. Moreover,

$$\begin{aligned}
&\sup_{\mathbb{P} \in \mathfrak{M}} \mathbb{E}_{\mathbb{P}} \left[\sum_{s \in [T]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) \right] \\
&\ge \mathbb{E}_{\tilde{\mathbb{P}}^\epsilon} \left[\sum_{s \in [T]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) \right] \\
&= (1 - \tfrac{\epsilon}{\vartheta+\epsilon})\mathbb{E}_{\mathbb{P}^\epsilon} \left[\sum_{s \in [T]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) \right] + \tfrac{\epsilon}{\vartheta+\epsilon}\mathbb{E}_{\widehat{\mathbb{P}}} \left[\sum_{s \in [T]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) \right] \\
&= (1 - \tfrac{\epsilon}{\vartheta+\epsilon}) \sup_{\mathbb{P} \in \mathfrak{M}^{\mathsf{C}}} \mathbb{E}_{\mathbb{P}} \left[\sum_{s \in [T]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) \right] + \tfrac{\epsilon}{\vartheta+\epsilon}\mathbb{E}_{\widehat{\mathbb{P}}} \left[\sum_{s \in [T]} \boldsymbol{c}_s^\top \boldsymbol{x}_s(\boldsymbol{\xi}_{[s]}) \right].
\end{aligned}$$

Letting $\epsilon \to 0$ yields the desired result. □

EC.2. Proof of Theorem 3

EC.2.1. $p=\infty$

Using Theorem 1, problem (1) is equivalent to

$$\min_{\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1}),\forall t\in[T]}\Bigg\{c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\|\xi_2-\widehat{\xi}_2\|\le\vartheta}\Big\{c_2^\top\boldsymbol{x}_2(\xi_2)+\mathbb{E}_{\widehat{\mathbb{P}}_{3|\widehat{\xi}_2}}\Big[\sup_{\|\xi_3-\widehat{\xi}_3\|\le\vartheta}\Big\{c_3^\top\boldsymbol{x}_3(\xi_{[3]})+ \cdots+\mathbb{E}_{\widehat{\mathbb{P}}_{T|\widehat{\xi}_{[T-1]}}}\Big[\sup_{\|\xi_T-\widehat{\xi}_T\|\le\vartheta}c_T^\top\boldsymbol{x}_T(\xi_{[T]})\Big]\cdots\Big\}\Big]\Big\}\Bigg]\Bigg\}.$$

Using Assumption 1, this becomes

$$\min_{\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1}),\forall t\in[T]}\Bigg\{c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\|\xi_2-\widehat{\xi}_2\|\le\vartheta}\Big\{c_2^\top\boldsymbol{x}_2(\xi_2)+\mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\|\xi_3-\widehat{\xi}_3\|\le\vartheta}\Big\{c_3^\top\boldsymbol{x}_3(\xi_{[3]})+ \cdots+\mathbb{E}_{\widehat{\mathbb{P}}_T}\Big[\sup_{\|\xi_T-\widehat{\xi}_T\|\le\vartheta}c_T^\top\boldsymbol{x}_T(\xi_{[T]})\Big]\cdots\Big\}\Big]\Big\}\Bigg]\Bigg\}.$$

Let us show recursively that for fixed $\boldsymbol{x}_{[t-1]}$,

$$\begin{aligned}
V_t(\boldsymbol{x}_{[t-1]}) &:= \inf_{\boldsymbol{x}_s\in\mathcal{X}_s(\boldsymbol{x}_{s-1}),s=t,\dots T}\Bigg\{c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\|\xi_2-\widehat{\xi}_2\|\le\vartheta}\Big\{c_2^\top\boldsymbol{x}_2(\xi_2)+\mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\|\xi_3-\widehat{\xi}_3\|\le\vartheta}\Big\{c_3^\top\boldsymbol{x}_3(\xi_{[3]})+ \\
&\qquad\cdots+\mathbb{E}_{\widehat{\mathbb{P}}_T}\Big[\sup_{\|\xi_T-\widehat{\xi}_T\|\le\vartheta}c_T^\top\boldsymbol{x}_T(\xi_{[T]})\Big]\cdots\Big\}\Big]\Big\}\Bigg]\Bigg\} \\
&= c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\|\xi_2-\widehat{\xi}_2\|\le\vartheta}\Big\{c_2^\top\boldsymbol{x}_2(\xi_2)+\mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\|\xi_3-\widehat{\xi}_3\|\le\vartheta}\Big\{c_3^\top\boldsymbol{x}_3(\xi_{[3]})+\cdots+ \\
&\qquad\mathbb{E}_{\widehat{\mathbb{P}}_{t-1}}\Big[\sup_{\|\xi_{t-1}-\widehat{\xi}_{t-1}\|\le\vartheta}\big\{c_{t-1}^\top\boldsymbol{x}_{t-1}(\xi_{[t-1]})+\mathcal{Q}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}),\xi_{[t-1]})\big\}\Big]\cdots\Big\}\Big]\Big\}\Bigg] \\
&= c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\|\xi_2-\widehat{\xi}_2\|\le\vartheta}\Big\{c_2^\top\boldsymbol{x}_2(\xi_2)+\mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\|\xi_3-\widehat{\xi}_3\|\le\vartheta}\Big\{c_3^\top\boldsymbol{x}_3(\xi_{[3]})+\cdots+ \\
&\qquad\mathbb{E}_{\widehat{\mathbb{P}}_{t-1}}\Big[\sup_{\|\xi_{t-1}-\widehat{\xi}_{t-1}\|\le\vartheta}\big\{c_{t-1}^\top\boldsymbol{x}_{t-1}(\xi_{[t-1]})+\mathbb{E}_{\widehat{\mathbb{P}}_t}\big[\mathcal{Q}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}),\widehat{\boldsymbol{\xi}}_t)\big]\big\}\Big]\cdots\Big\}\Big]\Big\}\Bigg].
\end{aligned} \tag{EC.4}$$

The case of $t=T+1$ holds trivially. Suppose (EC.4) holds for some $t+1$, $t=2,\dots,T$, now we prove for t. Using the induction hypothesis, it holds that

$$\begin{aligned}
V_t(\boldsymbol{x}_{[t-1]}) &= \inf_{\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1})}\Bigg\{c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\|\xi_2-\widehat{\xi}_2\|\le\vartheta}\Big\{c_2^\top\boldsymbol{x}_2(\xi_2)+\mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\|\xi_3-\widehat{\xi}_3\|\le\vartheta}\Big\{c_3^\top\boldsymbol{x}_3(\xi_{[3]})+\cdots+ \\
&\qquad\mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\|\xi_t-\widehat{\xi}_t\|\le\vartheta}\big\{c_t^\top\boldsymbol{x}_t(\xi_{[t]})+\mathcal{Q}_{t+1}(\boldsymbol{x}_t(\xi_{[t]}),\xi_{[t]})\big\}\Big]\cdots\Big\}\Big]\Big\}\Bigg]\Bigg\} \\
&= \inf_{\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1})}\Bigg\{c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\|\xi_2-\widehat{\xi}_2\|\le\vartheta}\Big\{c_2^\top\boldsymbol{x}_2(\xi_2)+\mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\|\xi_3-\widehat{\xi}_3\|\le\vartheta}\Big\{c_3^\top\boldsymbol{x}_3(\xi_{[3]})+\cdots+ \\
&\qquad\mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\|\xi_t-\widehat{\xi}_t\|\le\vartheta}\big\{c_t^\top\boldsymbol{x}_t(\xi_{[t]})+\mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[\mathcal{Q}_{t+1}(\boldsymbol{x}_t(\xi_{[t]}),\widehat{\boldsymbol{\xi}}_{t+1})\big]\big\}\Big]\cdots\Big\}\Big]\Big\}\Bigg]\Bigg\}.
\end{aligned}$$

Exchanging $\inf_{\boldsymbol{x}_t}$ and $\mathbb{E}_{\widehat{\mathbb{P}}_s}[\sup_{\xi_s}]$, $s=2,\ldots,t-1$, it follows that

$$\begin{aligned}
&V_t(\boldsymbol{x}_{[t-1]})\\
&\geq c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\|\xi_2-\widehat{\boldsymbol{\xi}}_2\|\leq\vartheta}\Big\{c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\|\xi_3-\widehat{\boldsymbol{\xi}}_3\|\leq\vartheta}\Big\{c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots +\\
&\qquad \inf_{\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1})}\mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\|\xi_t-\widehat{\boldsymbol{\xi}}_t\|\leq\vartheta}\{c_t^\top \boldsymbol{x}_t(\xi_{[t]}) + \mathcal{Q}_{t+1}(\boldsymbol{x}_t(\xi_{[t]}),\xi_{[t]})\}\Big]\cdots\Big\}\Big]\Big\}\Bigg]\\
&\geq c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\|\xi_2-\widehat{\boldsymbol{\xi}}_2\|\leq\vartheta}\Big\{c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\|\xi_3-\widehat{\boldsymbol{\xi}}_3\|\leq\vartheta}\Big\{c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots +\\
&\qquad \inf_{\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1})}\sup_{\mathbb{P}_t\in\mathcal{P}(\Xi_t):\mathcal{W}_\infty(\widehat{\mathbb{P}}_t,\mathbb{P}_t)\leq\vartheta}\mathbb{E}_{\mathbb{P}_t}\Big[\boldsymbol{c}_t^\top \boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}) + \mathcal{Q}_{t+1}(\boldsymbol{x}_t(\boldsymbol{\xi}_{[t]}),\boldsymbol{\xi}_{[t]})\Big]\cdots\Big\}\Big]\Big\}\Bigg]\\
&= c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\|\xi_2-\widehat{\boldsymbol{\xi}}_2\|\leq\vartheta}\Big\{c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\|\xi_3-\widehat{\boldsymbol{\xi}}_3\|\leq\vartheta}\Big\{c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots + \mathcal{Q}_t(\boldsymbol{x}_{t-1}(\boldsymbol{\xi}_{[t-1]}),\boldsymbol{\xi}_{[t-1]})\cdots\Big\}\Big]\Big\}\Bigg].
\end{aligned} \tag{EC.5}$$

Define $g_t(x_t,\xi_{[t]}):\mathbb{R}^{d_{x_t}}\times\Xi_{[t]}\to\mathbb{R}\cup\{+\infty\}$ as

$$g_t(x_t,\xi_{[t]}) := \begin{cases} c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[\mathcal{Q}_{t+1}(x_t,\widehat{\boldsymbol{\xi}}_{t+1})\big], & \text{if } x_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}),\xi_t),\\ +\infty, & \text{otherwise.}\end{cases}$$

Then g is random lower semi-continuous. Define $\mathcal{R}_{[t]}:\mathcal{Z}_{[t]}\to\mathbb{R}\cup\{\infty\}$ as

$$\begin{aligned}\mathcal{R}_{[t]}(\nu) := c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\|\xi_2-\widehat{\boldsymbol{\xi}}_2\|\leq\vartheta}\Big\{c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\|\xi_3-\widehat{\boldsymbol{\xi}}_3\|\leq\vartheta}\Big\{c_3^\top \boldsymbol{x}_3(\xi_{[3]}) +\\
\cdots + \mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\|\xi_t-\widehat{\boldsymbol{\xi}}_t\|\leq\vartheta}\nu(\xi_{[t]})\Big]\cdots\Big\}\Big]\Big\}\Bigg].\end{aligned}$$

By definition, $\mathcal{R}_{[t]}$ is monotone and continuous with respect to the L^∞-norm. Let $\mathfrak{X}_t$ be the space of measurable functions from $\Xi_{[t]}$ to X_t. It follows that

$$
\begin{aligned}
&\inf_{\chi_t\in\mathfrak{X}_t} \mathcal{R}_{[t]}(g_t(\chi_t(\cdot),\cdot))\\
&=\inf_{\chi_t\in\mathfrak{X}_t}\left\{c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\left[\sup_{\|\xi_2-\widehat{\xi}_2\|\le\vartheta}\left\{c_2^\top \boldsymbol{x}_2(\xi_2)+\mathbb{E}_{\widehat{\mathbb{P}}_3}\left[\sup_{\|\xi_3-\widehat{\xi}_3\|\le\vartheta}\left\{c_3^\top\boldsymbol{x}_3(\xi_{[3]})+\cdots+\right.\right.\right.\right.\right.\\
&\qquad\left.\left.\left.\left.\left.\mathbb{E}_{\widehat{\mathbb{P}}_t}\left[\sup_{\|\xi_t-\widehat{\xi}_t\|\le\vartheta} g_t(\chi_t(\xi_{[t]}),\xi_{[t]})\right]\cdots\right\}\right]\right\}\right]\right\}\\
&=\inf_{\substack{\chi_t(\xi_{[t]})\in\mathcal{X}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}),\xi_t)\\ \forall \xi_{[t]}}}\left\{c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\left[\sup_{\|\xi_2-\widehat{\xi}_2\|\le\vartheta}\left\{c_2^\top \boldsymbol{x}_2(\xi_2)+\mathbb{E}_{\widehat{\mathbb{P}}_3}\left[\sup_{\|\xi_3-\widehat{\xi}_3\|\le\vartheta}\left\{c_3^\top\boldsymbol{x}_3(\xi_{[3]})+\cdots+\right.\right.\right.\right.\right.\\
&\qquad\left.\left.\left.\left.\left.\mathbb{E}_{\widehat{\mathbb{P}}_t}\left[\sup_{\|\xi_t-\widehat{\xi}_t\|\le\vartheta}\left\{c_t^\top\chi_t(\xi_{[t]})+\mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[Q_{t+1}(\chi_t(\xi_{[t]}),\widehat{\boldsymbol{\xi}}_{t+1})\right]\cdots\right\}\right]\right\}\right]\right\}\right]\right\}\\
&=\inf_{\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1})}\left\{c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\left[\sup_{\|\xi_2-\widehat{\xi}_2\|\le\vartheta}\left\{c_2^\top \boldsymbol{x}_2(\xi_2)+\mathbb{E}_{\widehat{\mathbb{P}}_3}\left[\sup_{\|\xi_3-\widehat{\xi}_3\|\le\vartheta}\left\{c_3^\top\boldsymbol{x}_3(\xi_{[3]})+\cdots+\right.\right.\right.\right.\right.\\
&\qquad\left.\left.\left.\left.\left.\mathbb{E}_{\widehat{\mathbb{P}}_t}\left[\sup_{\|\xi_t-\widehat{\xi}_t\|\le\vartheta}\left\{c_t^\top\boldsymbol{x}_t(\xi_{[t]})+\mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[Q_{t+1}(\boldsymbol{x}_t(\xi_{[t]}),\widehat{\boldsymbol{\xi}}_{t+1})\right]\right\}\right]\cdots\right\}\right]\right\}\right]\right\}\\
&=V_t(\boldsymbol{x}_{[t-1]}).
\end{aligned}
$$

Define $\nu\in\mathcal{Z}_{[t]}$ as

$$
\nu_t(\xi_{[t]}):=\inf_{x_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}),\xi_t)} g_t(x_t,\xi_{[t]}),\quad \xi_{[t]}\in\Xi_{[t]}.
$$

By (Shapiro et al. 2021, Theorem 9.110), we have that

$$
\begin{aligned}
&V_t(\boldsymbol{x}_{[t-1]})\\
&=\inf_{\chi_t\in\mathfrak{X}_t}\mathcal{R}(g_t(\chi_t(\cdot),\cdot))\\
&=\mathcal{R}(\nu_t)\\
&=c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\left[\sup_{\|\xi_2-\widehat{\xi}_2\|\le\vartheta}\left\{c_2^\top \boldsymbol{x}_2(\xi_2)+\mathbb{E}_{\widehat{\mathbb{P}}_3}\left[\sup_{\|\xi_3-\widehat{\xi}_3\|\le\vartheta}\left\{c_3^\top\boldsymbol{x}_3(\xi_{[3]})+\right.\right.\right.\right.\\
&\qquad\left.\left.\left.\left.\cdots+\mathbb{E}_{\widehat{\mathbb{P}}_t}\left[\sup_{\|\xi_t-\widehat{\xi}_t\|\le\vartheta}\inf_{x_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}),\xi_t)}\left\{c_t^\top\boldsymbol{x}_t(\xi_{[t]})+\mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[Q_{t+1}(x_t,\widehat{\boldsymbol{\xi}}_{t+1})\right]\right\}\right]\cdots\right\}\right]\right\}\right]\\
&=c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}\left[\sup_{\|\xi_2-\widehat{\xi}_2\|\le\vartheta}\left\{c_2^\top \boldsymbol{x}_2(\xi_2)+\mathbb{E}_{\widehat{\mathbb{P}}_3}\left[\sup_{\|\xi_3-\widehat{\xi}_3\|\le\vartheta}\left\{c_3^\top\boldsymbol{x}_3(\xi_{[3]})+\cdots+\mathbb{E}_{\widehat{\mathbb{P}}_t}\left[Q_t(\boldsymbol{x}_{t-1},\widehat{\boldsymbol{\xi}}_t)\right]\cdots\right\}\right]\right\}\right].
\end{aligned}
$$

This completes the induction. Therefore, using (EC.4) we have shown that the optimal value of (1) is

$$
\min_{x_1\in\mathcal{X}_1}\left\{c_1^\top x_1+\mathbb{E}_{\widehat{\mathbb{P}}_2}[Q_2(x_1,\widehat{\boldsymbol{\xi}}_2)]\right\}.
$$

In addition, comparing with (EC.5), this shows that

$$
\mathbb{E}_{\widehat{\mathbb{P}}_t}\left[Q_t(x_{t-1},\widehat{\boldsymbol{\xi}}_t)\right]\ge \mathcal{Q}_t(x_{t-1},\xi_{[t-1]}).
$$

Next, we prove by induction that

$$\mathbb{E}_{\widehat{\mathbb{P}}_t}\left[Q_t(x_{t-1},\widehat{\boldsymbol{\xi}}_t)\right] = \mathcal{Q}_t(x_{t-1},\xi_{[t-1]}).$$

The base case $t = T+1$ trivially holds. Suppose we have shown the case for some $t+1$, $t = 2,\dots,T$. It suffices to prove $\mathbb{E}_{\widehat{\mathbb{P}}_t}\left[Q_t(x_{t-1},\widehat{\boldsymbol{\xi}}_t)\right] \le \mathcal{Q}_t(x_{t-1},\xi_{t-1})$. By definition of Q_t, we have that

$$\begin{aligned}
&\mathbb{E}_{\widehat{\mathbb{P}}_t}\left[Q_t(x_{t-1},\widehat{\boldsymbol{\xi}}_t)\right]\\
&= \mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\|\xi_t-\widehat{\boldsymbol{\xi}}_t\|\le\vartheta}\ \min_{x_t\in\mathcal{X}_t(x_{t-1},\xi_t)}\left\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[Q_{t+1}(x_t,\widehat{\boldsymbol{\xi}}_{t+1})\right]\right\}\Big]\\
&= \sup_{\mathbb{P}_t\in\mathcal{P}(\Xi_t):\mathcal{W}_\infty(\widehat{\mathbb{P}}_t,\mathbb{P}_t)\le\vartheta} \mathbb{E}_{\mathbb{P}_t}\Big[\min_{x_t\in\mathcal{X}_t(x_{t-1},\xi_t)}\left\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[Q_{t+1}(x_t,\widehat{\boldsymbol{\xi}}_{t+1})\right]\right\}\Big]\\
&= \sup_{\mathbb{P}_t\in\mathcal{P}(\Xi_t):\mathcal{W}_\infty(\widehat{\mathbb{P}}_t,\mathbb{P}_t)\le\vartheta}\ \min_{\boldsymbol{x}_t(\xi_{[t-1]},\cdot)\in\mathcal{X}_t(x_{t-1},\cdot)} \mathbb{E}_{\mathbb{P}_t}\left[c_t^\top \boldsymbol{x}_t(\xi_{[t-1]},\boldsymbol{\xi}_t) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[Q_{t+1}\big(\boldsymbol{x}_t(\xi_{[t-1]},\boldsymbol{\xi}_t),\widehat{\boldsymbol{\xi}}_{t+1}\big)\right]\right]\\
&\le \min_{\boldsymbol{x}_t(\xi_{[t-1]},\cdot)\in\mathcal{X}_t(x_{t-1},\cdot)}\ \sup_{\mathbb{P}_t\in\mathcal{P}(\Xi_t):\mathcal{W}_\infty(\widehat{\mathbb{P}}_t,\mathbb{P}_t)\le\vartheta} \mathbb{E}_{\mathbb{P}_t}\left[c_t^\top \boldsymbol{x}_t(\xi_{[t-1]},\boldsymbol{\xi}_t) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[Q_{t+1}\big(\boldsymbol{x}_t(\xi_{[t-1]},\boldsymbol{\xi}_t),\widehat{\boldsymbol{\xi}}_{t+1}\big)\right]\right]\\
&= \mathcal{Q}_t(x_{t-1},\xi_{[t-1]}).
\end{aligned}$$

where the second equality follows from the Wasserstein DRO reformulation, and the third equality follows from the interchangeability principle (Shapiro et al. 2021, Theorem 9.108). This completes the induction. In addition, observe that

$$\begin{aligned}
&\sup_{\|\xi_t-\widehat{\boldsymbol{\xi}}_t\|\le\vartheta}\left\{c_t^\top \boldsymbol{x}_t^\star(\xi_{[t]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[Q_{t+1}(\boldsymbol{x}_t^\star(\xi_{[t]}),\widehat{\boldsymbol{\xi}}_{t+1})\right]\right\}\\
&= \sup_{\|\xi_t-\widehat{\boldsymbol{\xi}}_t\|\le\vartheta}\ \min_{x_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1}^\star(\xi_{[t-1]}),\xi_t)}\left\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[Q_{t+1}(x_t,\widehat{\boldsymbol{\xi}}_{t+1})\right]\right\}\\
&= Q_t(\boldsymbol{x}_{t-1}^\star(\xi_{[t-1]}),\widehat{\boldsymbol{\xi}}_t).
\end{aligned}$$

Moreover, the optimal value of the inner minimization problem is continuous in $\xi_{[t]}$. Hence the worst-case risk over the nested distance ball equals that over the causal transport distance ball. This verifies the optimality of $(\boldsymbol{x}_1^\star,\dots,\boldsymbol{x}_T^\star)$. □

EC.2.2. $p\in[1,\infty)$

Using Theorem 2, the soft penalty version of problem (1) is equivalent to

$$\begin{aligned}
\min_{\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1}),\forall t\in[T]}&\Bigg\{\lambda\vartheta^p + c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2\in\Xi_2}\Bigg\{c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_{3|\widehat{\xi}_2}}\Bigg[\sup_{\xi_3\in\Xi_3}\Bigg\{c_3^\top \boldsymbol{x}_3(\xi_{[3]}) +\\
&\cdots + \mathbb{E}_{\widehat{\mathbb{P}}_{T|\widehat{\xi}_{[T-1]}}}\Big[\sup_{\xi_T\in\Xi_T}\Big\{c_T^\top \boldsymbol{x}_T(\xi_{[T]}) - \lambda\|\xi_T-\widehat{\boldsymbol{\xi}}_T\|^p\Big\}\Big]\cdots - \lambda\|\xi_3-\widehat{\boldsymbol{\xi}}_3\|^p\Bigg\}\Bigg] - \lambda\|\xi_2-\widehat{\boldsymbol{\xi}}_2\|^p\Bigg\}\Bigg]\Bigg\}.
\end{aligned}$$

Using Assumption 1, this becomes

$$\begin{aligned}
\min_{\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1}),\forall t\in[T]}&\Bigg\{\lambda\vartheta^p + c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2\in\Xi_2}\Bigg\{c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Bigg[\sup_{\xi_3\in\Xi_3}\Bigg\{c_3^\top \boldsymbol{x}_3(\xi_{[3]}) +\\
&\cdots + \mathbb{E}_{\widehat{\mathbb{P}}_T}\Big[\sup_{\xi_T\in\Xi_T}\Big\{c_T^\top \boldsymbol{x}_T(\xi_{[T]}) - \lambda\|\xi_T-\widehat{\boldsymbol{\xi}}_T\|^p\Big\}\Big]\cdots - \lambda\|\xi_3-\widehat{\boldsymbol{\xi}}_3\|^p\Bigg\}\Bigg] - \lambda\|\xi_2-\widehat{\boldsymbol{\xi}}_2\|^p\Bigg\}\Bigg]\Bigg\}.
\end{aligned}$$

Let us show recursively that for fixed λ and $\boldsymbol{x}_{[t-1]}$,

$$
\begin{aligned}
&V_t(\boldsymbol{x}_{[t-1]}) \\
&:= \inf_{\boldsymbol{x}_s \in \mathcal{X}_s(\boldsymbol{x}_{s-1}), s=t,\dots T} \Bigg\{ c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2 \in \Xi_2} \Bigg\{ c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Bigg[\sup_{\xi_3 \in \Xi_3} \Big\{ c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \\
&\qquad \cdots + \mathbb{E}_{\widehat{\mathbb{P}}_T}\Big[\sup_{\xi_T \in \Xi_T} \Big\{ c_T^\top \boldsymbol{x}_T(\xi_{[T]}) - \lambda \|\xi_T - \widehat{\boldsymbol{\xi}}_T\|^p \Big\} \Big] \cdots - \lambda \|\xi_3 - \widehat{\boldsymbol{\xi}}_3\|^p \Big\} \Bigg] - \lambda \|\xi_2 - \widehat{\boldsymbol{\xi}}_2\|^p \Bigg\} \Bigg] \Bigg\} \\
&= c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2 \in \Xi_2} \Bigg\{ c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Bigg[\sup_{\xi_3 \in \Xi_3} \Big\{ c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots + \mathbb{E}_{\widehat{\mathbb{P}}_{t-1}}\Big[\sup_{\xi_{t-1} \in \Xi_{t-1}} \big\{ c_{t-1}^\top \boldsymbol{x}_{t-1}(\xi_{[t-1]}) + \\
&\qquad \mathcal{Q}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}), \xi_{[t-1]}) - \lambda \|\xi_{t-1} - \widehat{\boldsymbol{\xi}}_{t-1}\|^p \big\} \Big] \cdots - \lambda \|\xi_3 - \widehat{\boldsymbol{\xi}}_3\|^p \Big\} \Bigg] - \lambda \|\xi_2 - \widehat{\boldsymbol{\xi}}_2\|^p \Bigg\} \Bigg]. \\
&= c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2 \in \Xi_2} \Bigg\{ c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Bigg[\sup_{\xi_3 \in \Xi_3} \Big\{ c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots + \mathbb{E}_{\widehat{\mathbb{P}}_{t-1}}\Big[\sup_{\xi_{t-1} \in \Xi_{t-1}} \big\{ c_{t-1}^\top \boldsymbol{x}_{t-1}(\xi_{[t-1]}) + \\
&\qquad \mathbb{E}_{\widehat{\mathbb{P}}_t}\big[\mathcal{Q}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}), \widehat{\boldsymbol{\xi}}_t) \big] - \lambda \|\xi_{t-1} - \widehat{\boldsymbol{\xi}}_{t-1}\|^p \big\} \Big] \cdots - \lambda \|\xi_3 - \widehat{\boldsymbol{\xi}}_3\|^p \Big\} \Bigg] - \lambda \|\xi_2 - \widehat{\boldsymbol{\xi}}_2\|^p \Bigg\} \Bigg].
\end{aligned}
\tag{EC.6}
$$

The case of $t = T+1$ holds trivially. Suppose (EC.6) holds for some $t+1$, $t = 2, \dots, T$, now we prove for t. Using the induction hypothesis, it holds that

$$
\begin{aligned}
&V_t(\boldsymbol{x}_{[t-1]}) \\
&= \inf_{\boldsymbol{x}_t \in \mathcal{X}_t(\boldsymbol{x}_{t-1})} \Bigg\{ c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2 \in \Xi_2} \Bigg\{ c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Bigg[\sup_{\xi_3 \in \Xi_3} \Big\{ c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots + \\
&\quad \mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\xi_t \in \Xi_t} \big\{ c_t^\top \boldsymbol{x}_t(\xi_{[t]}) + \mathcal{Q}_{t+1}(\boldsymbol{x}_t(\xi_{[t]}), \xi_{[t]}) - \lambda \|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p \big\} \Big] \cdots - \lambda \|\xi_3 - \widehat{\boldsymbol{\xi}}_3\|^p \Big\} \Bigg] - \lambda \|\xi_2 - \widehat{\boldsymbol{\xi}}_2\|^p \Bigg\} \Bigg] \Bigg\} \\
&= \inf_{\boldsymbol{x}_t \in \mathcal{X}_t(\boldsymbol{x}_{t-1})} \Bigg\{ c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2 \in \Xi_2} \Bigg\{ c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Bigg[\sup_{\xi_3 \in \Xi_3} \Big\{ c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots + \\
&\quad \mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\xi_t \in \Xi_t} \big\{ c_t^\top \boldsymbol{x}_t(\xi_{[t]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[\mathcal{Q}_{t+1}(\boldsymbol{x}_t(\xi_{[t]}), \widehat{\boldsymbol{\xi}}_{t+1}) \big] - \lambda \|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p \big\} \Big] \cdots - \lambda \|\xi_3 - \widehat{\boldsymbol{\xi}}_3\|^p \Big\} \Bigg] - \lambda \|\xi_2 - \widehat{\boldsymbol{\xi}}_2\|^p \Bigg\} \Bigg] \Bigg\}.
\end{aligned}
$$

Exchanging $\inf_{\boldsymbol{x}_t}$ and $\mathbb{E}_{\widehat{\mathbb{P}}_s}[\sup_{\xi_s}]$, $s=2,\ldots,t-1$, we obtain that

$$
\begin{aligned}
&V_t(\boldsymbol{x}_{[t-1]})\\
&\geq c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2\in\Xi_2}\Big\{c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\xi_3\in\Xi_3}\Big\{c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots + \inf_{\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1})}\\
&\qquad \mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\xi_t\in\Xi_t}\big\{c_t^\top \boldsymbol{x}_t(\xi_{[t]}) + \mathcal{Q}_{t+1}(\boldsymbol{x}_t(\xi_{[t]}),\xi_{[t]}) - \lambda\|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p\big\}\Big]\cdots - \lambda\|\xi_3-\widehat{\boldsymbol{\xi}}_3\|^p\Big\}\Big] - \lambda\|\xi_2-\widehat{\boldsymbol{\xi}}_2\|^p\Big\}\Bigg]\\
&= c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2\in\Xi_2}\Big\{c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\xi_3\in\Xi_3}\Big\{c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots + \inf_{\boldsymbol{x}_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1})}\sup_{\mathbb{P}_t\in\mathcal{P}(\Xi_t)}\big\{\mathbb{E}_{\mathbb{P}_t}\big[\boldsymbol{c}_t^\top \boldsymbol{x}_t(\xi_{[t-1]},\boldsymbol{\xi}_t)\\
&\qquad + \mathcal{Q}_{t+1}\big(\boldsymbol{x}_t(\xi_{[t-1]},\boldsymbol{\xi}_t),(\xi_{[t-1]},\boldsymbol{\xi}_t)\big)\big] - \lambda\mathcal{W}_p^p(\widehat{\mathbb{P}}_t,\mathbb{P}_t)\big\}\cdots - \lambda\|\xi_3-\widehat{\boldsymbol{\xi}}_3\|^p\Big\}\Big] - \lambda\|\xi_2-\widehat{\boldsymbol{\xi}}_2\|^p\Big\}\Bigg].\\
&= c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2\in\Xi_2}\Big\{c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\xi_3\in\Xi_3}\Big\{c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots + \mathcal{Q}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}),\xi_{t-1})\\
&\qquad \cdots - \lambda\|\xi_3-\widehat{\boldsymbol{\xi}}_3\|^p\Big\}\Big] - \lambda\|\xi_2-\widehat{\boldsymbol{\xi}}_2\|^p\Big\}\Bigg].
\end{aligned}
\tag{EC.7}
$$

Define $g_t(x_t,\xi_{[t]}):\mathbb{R}^{d_{x_t}}\times\Xi_{[t]}\to\mathbb{R}\cup\{+\infty\}$ as

$$
g_t(x_t,\xi_{[t]}) := \begin{cases} c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[\mathcal{Q}_{t+1}(x_t,\widehat{\boldsymbol{\xi}}_{t+1})\big] - \lambda\|\xi_t-\widehat{\boldsymbol{\xi}}_t\|^p, & \text{if } x_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}),\xi_t),\\ +\infty, & \text{otherwise.}\end{cases}
$$

Then g is random lower semi-continuous. Define $\mathcal{R}_{[t]}:\mathcal{Z}_{[t]}\to\mathbb{R}\cup\{\infty\}$ as

$$
\begin{aligned}
\mathcal{R}_{[t]}(\nu) := c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[&\sup_{\xi_2\in\Xi_2}\Big\{c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\xi_3\in\Xi_3}\Big\{c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots +\\
&\mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\xi_t\in\Xi_t}\big\{\nu(\xi_{[t]}) - \lambda\|\xi_t-\widehat{\boldsymbol{\xi}}_t\|^p\big\}\Big]\cdots - \lambda\|\xi_3-\widehat{\boldsymbol{\xi}}_3\|^p\Big\}\Big] - \lambda\|\xi_2-\widehat{\boldsymbol{\xi}}_2\|^p\Big\}\Bigg].
\end{aligned}
$$

By definition, $\mathcal{R}_{[t]}$ is monotone and continuous with respect to the L^∞-norm. Let $\mathfrak{X}_t$ be the space of measurable functions from $\Xi_{[t]}$ to X_t. It follows that

$$\begin{aligned}
&\inf_{\chi_t \in \mathfrak{X}_t} \mathcal{R}_{[t]}(g_t(\chi_t(\cdot),\cdot)) \\
&= \inf_{\chi_t \in \mathfrak{X}_t} \Bigg\{ c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2 \in \Xi_2} \Big\{ c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\xi_3 \in \Xi_3} \Big\{ c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots + \\
&\qquad \mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\xi_t \in \Xi_t} \{ g_t(\chi_t(\xi_{[t]}), \xi_{[t]}) - \lambda \|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p \} \Big] \cdots - \lambda \|\xi_3 - \widehat{\boldsymbol{\xi}}_3\|^p \Big\} \Big] - \lambda \|\xi_2 - \widehat{\boldsymbol{\xi}}_2\|^p \Big\} \Bigg] \Bigg\} \\
&= \inf_{\substack{\chi_t(\xi_{[t]}) \in \mathcal{X}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}), \xi_t) \\ \forall \xi_{[t]}}} \Bigg\{ c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2 \in \Xi_2} \Big\{ c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\xi_3 \in \Xi_3} \Big\{ c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots + \mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\xi_t \in \Xi_t} \{ \\
&\qquad c_t^\top \chi_t(\xi_{[t]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}[Q_{t+1}(\chi_t(\xi_{[t]}), \widehat{\boldsymbol{\xi}}_{t+1})] - \lambda \|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p \} \Big] \cdots - \lambda \|\xi_3 - \widehat{\boldsymbol{\xi}}_3\|^p \Big\} \Big] - \lambda \|\xi_2 - \widehat{\boldsymbol{\xi}}_2\|^p \Big\} \Bigg] \Bigg\} \\
&= \inf_{\boldsymbol{x}_t \in \mathcal{X}_t(\boldsymbol{x}_{t-1})} \Bigg\{ c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2 \in \Xi_2} \Big\{ c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\xi_3 \in \Xi_3} \Big\{ c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots + \mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\xi_t \in \Xi_t} \{ \\
&\qquad c_t^\top \boldsymbol{x}_t(\xi_{[t]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}[Q_{t+1}(\boldsymbol{x}_t(\xi_{[t]}), \widehat{\boldsymbol{\xi}}_{t+1})] - \lambda \|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p \} \Big] \cdots - \lambda \|\xi_3 - \widehat{\boldsymbol{\xi}}_3\|^p \Big\} \Big] - \lambda \|\xi_2 - \widehat{\boldsymbol{\xi}}_2\|^p \Big\} \Bigg] \Bigg\} \\
&= V_t(\boldsymbol{x}_{[t-1]}).
\end{aligned}$$

Define $\nu \in \mathcal{Z}_{[t]}$ as

$$\nu_t(\xi_{[t]}) := \inf_{x_t \in \mathcal{X}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}), \xi_t)} g_t(x_t, \xi_{[t]}), \quad \xi_{[t]} \in \Xi_{[t]}.$$

By (Shapiro et al. 2021, Theorem 9.110), we have that

$$\begin{aligned}
&V_t(\boldsymbol{x}_{[t-1]}) \\
&= \inf_{\chi_t \in \mathfrak{X}_t} \mathcal{R}_{[t]}(g_t(\chi_t(\cdot),\cdot)) \\
&= \mathcal{R}_{[t]}(\nu_t) \\
&= c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2 \in \Xi_2} \Big\{ c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\xi_3 \in \Xi_3} \Big\{ c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots + \mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\xi_t \in \Xi_t} \inf_{x_t \in \mathcal{X}_t(\boldsymbol{x}_{t-1}(\xi_{[t-1]}), \xi_t)} \{ \\
&\qquad c_t^\top \boldsymbol{x}_t(\xi_{[t]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}[Q_{t+1}(x_t, \widehat{\boldsymbol{\xi}}_{t+1})] - \lambda \|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p \} \Big] \cdots - \lambda \|\xi_3 - \widehat{\boldsymbol{\xi}}_3\|^p \Big\} \Big] - \lambda \|\xi_2 - \widehat{\boldsymbol{\xi}}_2\|^p \Big\} \Bigg] \\
&= c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}\Bigg[\sup_{\xi_2 \in \Xi_2} \Big\{ c_2^\top \boldsymbol{x}_2(\xi_2) + \mathbb{E}_{\widehat{\mathbb{P}}_3}\Big[\sup_{\xi_3 \in \Xi_3} \Big\{ c_3^\top \boldsymbol{x}_3(\xi_{[3]}) + \cdots + \mathbb{E}_{\widehat{\mathbb{P}}_t}[Q_t(\boldsymbol{x}_{t-1}, \widehat{\boldsymbol{\xi}}_t)] \cdots \\
&\qquad - \lambda \|\xi_3 - \widehat{\boldsymbol{\xi}}_3\|^p \Big\} \Big] - \lambda \|\xi_2 - \widehat{\boldsymbol{\xi}}_2\|^p \Big\} \Bigg].
\end{aligned}$$

This completes the induction. Therefore, using (EC.4) we have shown that the optimal value of (1) is

$$\min_{x_1 \in \mathcal{X}_1} \left\{ \lambda \vartheta^p + c_1^\top x_1 + \mathbb{E}_{\widehat{\mathbb{P}}_2}[Q_2(x_1, \widehat{\boldsymbol{\xi}}_2)] \right\}.$$

In addition, comparing with (EC.7), this shows that

$$\mathbb{E}_{\widehat{\mathbb{P}}_t}\left[Q_t(x_{t-1}, \widehat{\boldsymbol{\xi}}_t)\right] \geq \mathcal{Q}_t(x_{t-1}, \xi_{[t-1]}).$$

Next, we prove by induction that the above inequality holds as equality. The base case $t = T+1$ trivially holds. Suppose we have shown the case for some $t+1$, $t = 2,\dots,T$. It suffices to prove $\mathbb{E}_{\widehat{\mathbb{P}}_t}\big[Q_t(x_{t-1},\widehat{\boldsymbol{\xi}}_t)\big] \le \mathcal{Q}_t(x_{t-1},\xi_{t-1})$. By definition of Q_t, we have that

$$
\begin{aligned}
&\mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[Q_t(x_{t-1},\widehat{\boldsymbol{\xi}}_t)\Big]\\
&= \mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[\sup_{\xi_t\in\Xi_t}\min_{x_t\in\mathcal{X}_t(x_{t-1},\xi_t)}\Big\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t,\widehat{\boldsymbol{\xi}}_{t+1})\big] - \lambda\|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p\Big\}\Big]\\
&= \sup_{\mathbb{P}_t\in\mathcal{P}(\Xi_t)}\Big\{\mathbb{E}_{\mathbb{P}_t}\Big[\min_{x_t\in\mathcal{X}_t(x_{t-1},\xi_t)}\big\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t,\widehat{\boldsymbol{\xi}}_{t+1})\big]\big\}\Big] - \lambda\mathcal{W}_p^p(\widehat{\mathbb{P}}_t,\mathbb{P}_t)\Big\}\\
&= \sup_{\mathbb{P}_t\in\mathcal{P}(\Xi_t)}\min_{\boldsymbol{x}_t(\xi_{[t-1]},\cdot)\in\mathcal{X}_t(x_{t-1},\cdot)}\Big\{\mathbb{E}_{\mathbb{P}_t}\Big[c_t^\top \boldsymbol{x}_t(\xi_{[t-1]},\boldsymbol{\xi}_t) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}\big(\boldsymbol{x}_t(\xi_{[t-1]},\boldsymbol{\xi}_t),\widehat{\boldsymbol{\xi}}_{t+1}\big)\big]\Big] - \lambda\mathcal{W}_p^p(\widehat{\mathbb{P}}_t,\mathbb{P}_t)\Big\},
\end{aligned}
$$

where the second equality follows from the Wasserstein DRO reformulation (Zhang et al. 2022), and the third equality follows from the interchangeability principle (Shapiro et al. 2021, Theorem 9.108). Exchanging sup and min, we obtain

$$
\begin{aligned}
&\mathbb{E}_{\widehat{\mathbb{P}}_t}\Big[Q_t(x_{t-1},\widehat{\boldsymbol{\xi}}_t)\Big]\\
&\le \min_{\boldsymbol{x}_t(\xi_{[t-1]},\cdot)\in\mathcal{X}_t(x_{t-1},\cdot)}\sup_{\mathbb{P}_t\in\mathcal{P}(\Xi_t)}\Big\{\mathbb{E}_{\mathbb{P}_t}\Big[c_t^\top \boldsymbol{x}_t(\xi_{[t-1]},\boldsymbol{\xi}_t) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}\big(\boldsymbol{x}_t(\xi_{[t-1]},\boldsymbol{\xi}_t),\widehat{\boldsymbol{\xi}}_{t+1}\big)\big]\Big] - \lambda\mathcal{W}_p^p(\widehat{\mathbb{P}}_t,\mathbb{P}_t)\Big\}\\
&= \mathcal{Q}_t(x_{t-1},\xi_{[t-1]}).
\end{aligned}
$$

This completes the induction. In addition, observe that

$$
\begin{aligned}
&\sup_{\xi_t\in\Xi_t}\Big\{c_t^\top \boldsymbol{x}_t^\star(\xi_{[t]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(\boldsymbol{x}_t^\star(\xi_{[t]}),\widehat{\boldsymbol{\xi}}_{t+1})\big] - \lambda\|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p\Big\}\\
&= \sup_{\xi_t\in\Xi_t}\min_{x_t\in\mathcal{X}_t(\boldsymbol{x}_{t-1}^\star(\xi_{[t-1]}),\xi_t)}\Big\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t,\widehat{\boldsymbol{\xi}}_{t+1})\big] - \lambda\|\xi_t - \widehat{\boldsymbol{\xi}}_t\|^p\Big\}\\
&= Q_t(\boldsymbol{x}_{t-1}^\star(\xi_{[t-1]}),\widehat{\boldsymbol{\xi}}_t).
\end{aligned}
$$

Moreover, the optimal value of the inner minimization problem is continuous in $\xi_{[t]}$. Hence the worst-case risk with respect to the nested distance equals that with respect to the causal transport distance. This verifies the optimality of $(\boldsymbol{x}_1^\star,\dots,\boldsymbol{x}_T^\star)$. □

EC.3. Proofs for Section 4

EC.3.1. Proof of Corollary 2

Applying Theorem 3, we have that for $p=\infty$,

$$
\begin{aligned}
Q_t(x_{t-1},\widehat{\boldsymbol{c}}_t) &= \sup_{\|c_t-\widehat{\boldsymbol{c}}_t\|\le\vartheta}\min_{x_t\in\mathcal{X}_t(x_{t-1})}\Big\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t,\widehat{\boldsymbol{c}}_{t+1})\big]\Big\}\\
&= \min_{x_t\in\mathcal{X}_t(x_{t-1})}\sup_{\|c_t-\widehat{\boldsymbol{c}}_t\|\le\vartheta}\Big\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t,\widehat{\boldsymbol{c}}_{t+1})\big]\Big\}\\
&= \min_{x_t\in\mathcal{X}_t(x_{t-1})}\Big\{\widehat{\boldsymbol{c}}_t^\top x_t + \vartheta\|x_t\|_* + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t,\widehat{\boldsymbol{c}}_{t+1})\big]\Big\}.
\end{aligned}
$$

For $p=1$,

$$
\begin{aligned}
Q_t(x_{t-1},\widehat{\boldsymbol{c}}_t) &:= \sup_{c_t\in\Xi_t}\min_{x_t\in\mathcal{X}_t(x_{t-1})}\Big\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t,\widehat{\boldsymbol{c}}_{t+1})\big] - \lambda\|c_t - \widehat{\boldsymbol{c}}_t\|\Big\}\\
&= \min_{x_t\in\mathcal{X}_t(x_{t-1})}\sup_{c_t\in\Xi_t}\Big\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t,\widehat{\boldsymbol{c}}_{t+1})\big] - \lambda\|c_t - \widehat{\boldsymbol{c}}_t\|\Big\}\\
&= \min_{x_t\in\mathcal{X}_t(x_{t-1}),\|x_t\|_*\le\lambda}\Big\{\widehat{\boldsymbol{c}}_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t,\widehat{\boldsymbol{c}}_{t+1})\big]\Big\}.
\end{aligned}
$$

For $p \in (1,\infty)$,

$$\begin{aligned} Q_t(x_{t-1}, \widehat{\boldsymbol{c}}_t) &:= \sup_{c_t \in \Xi_t} \min_{x_t \in \mathcal{X}_t(x_{t-1})} \left\{ c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{c}}_{t+1}) \right] - \lambda \|c_t - \widehat{\boldsymbol{c}}_t\|^p \right\} \\ &= \min_{x_t \in \mathcal{X}_t(x_{t-1})} \sup_{c_t \in \Xi_t} \left\{ c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{c}}_{t+1}) \right] - \lambda \|c_t - \widehat{\boldsymbol{c}}_t\|^p \right\} \\ &= \min_{x_t \in \mathcal{X}_t(x_{t-1})} \left\{ \widehat{\boldsymbol{c}}_t^\top x_t + (1-1/p)(\tfrac{1}{p\lambda})^{\frac{1}{p-1}} \|x_t\|_*^{\frac{p}{p-1}} + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{c}}_{t+1}) \right] \right\}. \end{aligned}$$

□

EC.3.2. Proof of Proposition 1

We first prove for the case of $p = \infty$. Consider

$$\sup_{\|c_t - \widehat{\boldsymbol{c}}_t\| \le \vartheta} \left\{ c_t^\top \bar{\boldsymbol{x}}_t(c_{[t]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(\bar{\boldsymbol{x}}_t(c_{[t]}), \widehat{\boldsymbol{c}}_{t+1}) \right] \right\}.$$

By definition of $\bar{\boldsymbol{x}}_t$, we have

$$c_t^\top \bar{\boldsymbol{x}}_t(c_{[t]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}\big(\bar{\boldsymbol{x}}_t(c_{[t]}), \widehat{\boldsymbol{c}}_{[t]}\big) \right] = c_t^\top \widehat{\boldsymbol{x}}_t\big(\widehat{\boldsymbol{c}}_{[t]}^c\big) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}\big(\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}^c), \widehat{\boldsymbol{c}}_{t+1}\big) \right].$$

Decompose the right-hand side as

$$\begin{aligned} & c_t^\top \widehat{\boldsymbol{x}}_t\big(\widehat{\boldsymbol{c}}_{[t]}^c\big) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}\big(\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}^c), \widehat{\boldsymbol{c}}_{t+1}\big) \right] \\ &= (\widehat{\boldsymbol{c}}_t^c)^\top \widehat{\boldsymbol{x}}_t\big(\widehat{\boldsymbol{c}}_{[t]}^c\big) + (c_t - \widehat{\boldsymbol{c}}_t^c)^\top \widehat{\boldsymbol{x}}_t\big(\widehat{\boldsymbol{c}}_{[t]}^c\big) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}\big(\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}^c), \widehat{\boldsymbol{c}}_{t+1}\big) \right] \\ &\le (\widehat{\boldsymbol{c}}_t^c)^\top \widehat{\boldsymbol{x}}_t\big(\widehat{\boldsymbol{c}}_{[t]}^c\big) + \vartheta \|\widehat{\boldsymbol{x}}_t\big(\widehat{\boldsymbol{c}}_{[t]}^c\big)\|_* + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}\big(\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}^c), \widehat{\boldsymbol{c}}_{t+1}\big) \right]. \end{aligned}$$

By definition of $\widehat{\boldsymbol{c}}_{[t]}^c$, the last expression is upper bounded by

$$\widehat{\boldsymbol{c}}_t^\top \widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t) + \vartheta \|\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t)\|_* + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t), \widehat{\boldsymbol{c}}_{t+1}) \right] = Q_t(\widehat{\boldsymbol{x}}_{t-1}(\widehat{\boldsymbol{c}}_{[t-1]}^c), \widehat{\boldsymbol{c}}_t).$$

Therefore, we have shown that

$$\sup_{\|c_t - \widehat{\boldsymbol{c}}_t\| \le \vartheta} \left\{ c_t^\top \bar{\boldsymbol{x}}_t(c_{[t]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(\bar{\boldsymbol{x}}_t(c_{[t]}), \widehat{\boldsymbol{c}}_{t+1}) \right] \right\} \le Q_t(\bar{\boldsymbol{x}}_{t-1}(c_{[t-1]}), \widehat{\boldsymbol{c}}_t).$$

Taking the expectation over $\widehat{\boldsymbol{c}}_t \sim \widehat{\mathbb{P}}_t$ and using Theorem 3, we obtain that

$$\mathbb{E}_{\widehat{\mathbb{P}}_t} \left[\sup_{\|c_t - \widehat{\boldsymbol{c}}_t\| \le \vartheta} \left\{ c_t^\top \bar{\boldsymbol{x}}_t(c_{[t]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(\bar{\boldsymbol{x}}_t(c_{[t]}), \widehat{\boldsymbol{c}}_{t+1}) \right] \right\} \right] \le \mathcal{Q}_t(\bar{\boldsymbol{x}}_{t-1}(c_{[t-1]})).$$

Hence we have

$$\bar{\boldsymbol{x}}_t(c_{[t-1]}, \cdot) \in \operatorname*{arg\,min}_{\boldsymbol{x}_t(c_{[t-1]}, \cdot) \in \mathcal{X}_t(\bar{\boldsymbol{x}}_{t-1}(c_{[t-1]}), \cdot)} \rho_t \left[\boldsymbol{c}_t^\top \boldsymbol{x}_t(c_{[t-1]}, \boldsymbol{c}_t) + \mathcal{Q}_{t+1}\big(\boldsymbol{x}_t(c_{[t-1]}, \boldsymbol{c}_t)\big) \right],$$

which shows the optimality of the policy $(\bar{\boldsymbol{x}}_1, \ldots, \bar{\boldsymbol{x}}_T)$.

Next, we prove for the case of $p \in [1,\infty)$. Consider

$$\sup_{c_t \in \Xi_t} \left\{ c_t^\top \bar{\boldsymbol{x}}_t(c_{[t]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(\bar{\boldsymbol{x}}_t(c_{[t]}), \widehat{\boldsymbol{c}}_{t+1}) \right] - \lambda \|c_t - \widehat{\boldsymbol{c}}_t\|^p \right\}.$$

Using the expression for $\bar{\boldsymbol{x}}_t$, we obtain

$$c_t^\top \bar{\boldsymbol{x}}_t(c_{[t]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}\big(\bar{\boldsymbol{x}}_t(c_{[t]}), \widehat{\boldsymbol{c}}_{[t]}\big) \right] = c_t^\top \widehat{\boldsymbol{x}}_t\big(\widehat{\boldsymbol{c}}_{[t]}^c\big) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}\big(\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}^c), \widehat{\boldsymbol{c}}_{t+1}\big) \right].$$

Decompose the right-hand side as

$$\begin{aligned}
&c_t^\top \widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_t^c) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}^c), \widehat{\boldsymbol{c}}_{t+1})\big] \\
&= (\widehat{\boldsymbol{c}}_{[t]}^c)^\top \widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}^c) + (c_t - \widehat{\boldsymbol{c}}_t^c)^\top \widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}^c) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}^c), \widehat{\boldsymbol{c}}_{t+1})\big] \\
&\le (\widehat{\boldsymbol{c}}_t^c)^\top \widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}^c) + \frac{1}{p}\big((p\lambda)^{1/p}\|c_t - \widehat{\boldsymbol{c}}_t^c\|\big)^p + (1-\frac{1}{p})\big((p\lambda)^{-1/p}\|\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}^c)\|_*\big)^{\frac{p}{p-1}} + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t]}^c), \widehat{\boldsymbol{c}}_{t+1})\big],
\end{aligned}$$

where the last inequality follows from the Young's inequality. By definition of $\widehat{\boldsymbol{c}}_{[t]}^c$, the last expression is upper bounded by

$$\widehat{\boldsymbol{c}}_t^\top \widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t) + \lambda\|c_t - \widehat{\boldsymbol{c}}_t\|^p + (1-\frac{1}{p})\big((p\lambda)^{-1/p}\|\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t)\|_*\big)^{\frac{p}{p-1}} + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t), \widehat{\boldsymbol{c}}_{t+1})\big].$$

Thereby we have

$$\begin{aligned}
&\sup_{c_t\in\Xi_t}\Big\{c_t^\top \bar{\boldsymbol{x}}_t(c_{[t]}) + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[\widehat{Q}_{t+1}(\bar{\boldsymbol{x}}_t(c_{[t]}), \widehat{\boldsymbol{c}}_{t+1})\big] - \lambda\|c_t - \widehat{\boldsymbol{c}}_t\|^p\Big\} \\
&\le \widehat{\boldsymbol{c}}_t^\top \widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t) + (1-\frac{1}{p})\big((p\lambda)^{-1/p}\|\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t)\|_*\big)^{\frac{p}{p-1}} + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(\widehat{\boldsymbol{x}}_t(\widehat{\boldsymbol{c}}_{[t-1]}^c, \widehat{\boldsymbol{c}}_t), \widehat{\boldsymbol{c}}_{t+1})\big]. \\
&= Q_t(\widehat{\boldsymbol{x}}_{t-1}(\widehat{\boldsymbol{c}}_{[t-1]}^c), \widehat{\boldsymbol{c}}_t) \\
&= Q_t(\bar{\boldsymbol{x}}_{t-1}(c_{[t-1]}), \widehat{\boldsymbol{c}}_t),
\end{aligned}$$

which shows the optimality of the policy $(\bar{\boldsymbol{x}}_1, \ldots, \bar{\boldsymbol{x}}_T)$. □

EC.3.3. Proof of Corollary 3

We first compute

$$Q_t(x_{t-1}, \widehat{\boldsymbol{b}}_t) := \sup_{\|b_t - \widehat{\boldsymbol{b}}_t\|\le\vartheta} \min_{x_t\in\mathcal{X}_t(x_{t-1}, b_t)} \Big\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t, \widehat{\boldsymbol{b}}_{t+1})\big]\Big\},$$

Thanks to (Shapiro et al. 2021, Section 3.2.1), the objective $c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t, \widehat{\boldsymbol{b}}_{t+1})\big]$ is convex in x_t. Using convex programming duality we obtain that

$$\begin{aligned}
&\min_{x_t\in\mathcal{X}_t(x_{t-1}, b_t)} \Big\{c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t, \widehat{\boldsymbol{b}}_{t+1})\big]\Big\} \\
&= \max_{y_t\in\mathbb{R}^{d_t}} \Big\{y_t^\top(b_t - B_t x_{t-1}) + \min_{x_t\ge 0}\Big\{(c_t - A_t^\top y_t)^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t, \widehat{\boldsymbol{b}}_{t+1})\big]\Big\}\Big\}.
\end{aligned}$$

It follows that

$$\begin{aligned}
&Q_t(x_{t-1}, \widehat{\boldsymbol{b}}_t) \\
&= \sup_{\|b_t - \widehat{\boldsymbol{b}}_t\|\le\vartheta} \max_{y_t\in\mathbb{R}^{d_t}} \Big\{y_t^\top(b_t - B_t x_{t-1}) + \min_{x_t\ge 0}\Big\{(c_t - A_t^\top y_t)^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t, \widehat{\boldsymbol{b}}_{t+1})\big]\Big\}\Big\} \\
&= \max_{y_t\in\mathbb{R}^{d_t}} \sup_{\|b_t - \widehat{\boldsymbol{b}}_t\|\le\vartheta} \Big\{y_t^\top(b_t - B_t x_{t-1}) + \min_{x_t\ge 0}\Big\{(c_t - A_t^\top y_t)^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t, \widehat{\boldsymbol{b}}_{t+1})\big]\Big\}\Big\} \\
&= \max_{y_t\in\mathbb{R}^{d_t}} \Big\{y_t^\top(\widehat{\boldsymbol{b}}_t - B_t x_{t-1}) + \vartheta\|y_t\|_* + \min_{x_t\ge 0}\Big\{(c_t - A_t^\top y_t)^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[Q_{t+1}(x_t, \widehat{\boldsymbol{b}}_{t+1})\big]\Big\}\Big\} \\
&= \max_{y_t\in\mathbb{R}^{d_t}} \Big\{(\widehat{\boldsymbol{b}}_t - B_t x_{t-1})^\top y_t + \vartheta\|y_t\|_* - \psi_t^*(A_t^\top y_t)\Big\}.
\end{aligned}$$

The case of uncertainty in $\boldsymbol{B}_t$ can be dealt in a similar way. We first compute

$$Q_t(x_{t-1}, \widehat{\boldsymbol{B}}_t) := \sup_{\|B_t - \widehat{\boldsymbol{B}}_t\| \leq \vartheta} \min_{x_t \in \mathcal{X}_t(x_{t-1}, B_t)} \left\{ c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{B}}_{t+1}) \right] \right\},$$

Thanks to (Shapiro et al. 2021, Section 3.2.1), the objective $c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{b}}_{t+1}) \right]$ is convex in x_t. Using convex programming duality we obtain that

$$\begin{aligned}
&\min_{x_t \in \mathcal{X}_t(x_{t-1}, B_t)} \left\{ c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{B}}_{t+1}) \right] \right\} \\
&= \max_{y_t \in \mathbb{R}^{d_t}} \left\{ y_t^\top (b_t - B_t x_{t-1}) + \min_{x_t \geq 0} \left\{ (c_t - A_t^\top y_t)^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{B}}_{t+1}) \right] \right\} \right\}.
\end{aligned}$$

It follows that

$$\begin{aligned}
&Q_t(x_{t-1}, \widehat{\boldsymbol{B}}_{[t]}) \\
&= \sup_{\|B_t - \widehat{\boldsymbol{B}}_t\| \leq \vartheta} \max_{y_t \in \mathbb{R}^{d_t}} \left\{ y_t^\top (b_t - B_t x_{t-1}) + \min_{x_t \geq 0} \left\{ (c_t - A_t^\top y_t)^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{B}}_{t+1}) \right] \right\} \right\} \\
&= \max_{y_t \in \mathbb{R}^{d_t}} \sup_{\|B_t - \widehat{\boldsymbol{B}}_t\| \leq \vartheta} \left\{ y_t^\top (b_t - B_t x_{t-1}) + \min_{x_t \geq 0} \left\{ (c_t - A_t^\top y_t)^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[Q_{t+1}(x_t, \widehat{\boldsymbol{B}}_{t+1}) \right] \right\} \right\} \\
&= \max_{y_t \in \mathbb{R}^{d_t}} \left\{ y_t^\top (b_t - \widehat{\boldsymbol{B}}_t x_{t-1}) + \vartheta \|x_{t-1} y_t^\top\|_* + \min_{x_t \geq 0} \left\{ (c_t - A_t^\top y_t)^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[\widehat{Q}_{t+1}(x_t, \widehat{\boldsymbol{b}}_{t+1}) \right] \right\} \right\} \\
&= \max_{y_t \in \mathbb{R}^{d_t}} \left\{ (b_t - \widehat{\boldsymbol{B}}_t x_{t-1})^\top y_t + \vartheta \|x_{t-1} y_t^\top\|_* - \psi_t^*(A_t^\top y_t) \right\}.
\end{aligned}$$

When $\|\cdot\| = \|\cdot\|_{\text{op}}$, it holds that

$$\max_{\|\Delta\|_{\text{op}} \leq \vartheta} y_t^\top \Delta x_{t-1} = \vartheta \|y_t\|_\infty \|x_{t-1}\| = \max_{j \in [m_t], \delta \in \{1, -1\}} \vartheta \delta y_t^\top e_j \|x_{t-1}\|.$$

Indeed, using Hölder's inequality and the definition of the operator norm, it holds that

$$y^\top \Delta x \leq \|y\|_\infty \|\Delta x\|_1 \leq \|y\|_\infty \|\Delta\|_{\text{op}} \|x\|.$$

Moreover, the inequality holds as equality at $\tilde{\Delta} = \vartheta \tilde{y} \tilde{x}^\top$, where $\tilde{x}$, $\tilde{y}$ are such that $\tilde{x}^\top x = \|x\|$, $\|\tilde{x}\|_* = 1$ and $\tilde{y}^\top y = \|y\|_\infty$, $\|\tilde{y}\|_1 = 1$. In fact, we verify that

$$\|\tilde{\Delta}\|_{\text{op}} = \sup_{\|v\| \leq 1} \|\vartheta \tilde{y} \tilde{x}^\top v\|_1 = \vartheta \sup_{\|v\| \leq 1} \|\tilde{y}\|_1 |\tilde{x}^\top v| = \vartheta,$$

and

$$y^\top \tilde{\Delta} x = \vartheta y^\top \tilde{y} \tilde{x}^\top x = \vartheta \|y\|_\infty \|x\|.$$

□

EC.3.4. Proof of Corollary 4

Define a function $\ell_t : \Xi_t \to \mathbb{R}$ as

$$\ell_t(b_t) := \max_{y_t \in \mathbb{R}^{d_t}} \left\{ y_t^\top (b_t - B_t x_{t-1}) - \widehat{\psi}_t^*(A_t^\top y_t) \right\}.$$

We prove by induction that

$$Q_t(x_{t-1}, \widehat{\boldsymbol{b}}_t) = \min_{x_t \in \mathcal{X}_t(x_{t-1}, b_t)} \left\{ c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[\widehat{Q}_{t+1}(x_t, \widehat{\boldsymbol{b}}_{t+1}) \right] \right\} + \infty \cdot \mathbf{1} \left\{ \lambda < \max_{s=t,\ldots,T} \|\ell_t\|_{\text{Lip}} \right\}.$$

Thanks to (Shapiro et al. 2021, Section 3.2.1), the objective $c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\big[\widehat{Q}_{t+1}(x_t, \widehat{\boldsymbol{c}}_{t+1})\big]$ is convex in x_t. Using convex programming duality we obtain that

$$
\begin{aligned}
&\min_{x_t \in \mathcal{X}_t(x_{t-1}, b_t)} \left\{ c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[\widehat{Q}_{t+1}\big(x_t, \widehat{\boldsymbol{b}}_{t+1}\big)\right]\right\} \\
&= \max_{y_t \in \mathbb{R}^{d_t}} \left\{ y_t^\top (b_t - B_t x_{t-1}) + \min_{x_t \geq 0}\left\{ (c_t - A_t^\top y_t)^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[\widehat{Q}_{t+1}\big(x_t, \widehat{\boldsymbol{b}}_{t+1}\big)\right]\right\}\right\}.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&Q_t(x_{t-1}, \widehat{\boldsymbol{b}}_t) \\
&= \sup_{b_t \in \Xi_t} \left\{ \min_{x_t \in \mathcal{X}_t(x_{t-1}, b_t)} \left\{ c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[Q_{t+1}(x_t, \widehat{\boldsymbol{b}}_{t+1})\right]\right\} - \lambda \|b_t - \widehat{\boldsymbol{b}}_t\| \right\} + \infty \cdot \mathbf{1}\left\{\lambda < \max_{s=t+1,\ldots,T} \|\ell_s\|_{\mathrm{Lip}}\right\} \\
&= \sup_{b_t \in \Xi_t} \max_{y_t \in \mathbb{R}^{d_t}} \left\{ y_t^\top (b_t - B_t x_{t-1}) + \min_{x_t \geq 0}\left\{ (c_t - A_t^\top y_t)^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[\widehat{Q}_{t+1}\big(x_t, \widehat{\boldsymbol{b}}_{t+1}\big)\right]\right\} - \lambda \|b_t - \widehat{\boldsymbol{b}}_t\| \right\} \\
&\qquad + \infty \cdot \mathbf{1}\left\{\lambda < \max_{s=t+1,\ldots,T} \|\ell_s\|_{\mathrm{Lip}}\right\} \\
&= \max_{y_t \in \mathbb{R}^{d_t}} \sup_{b_t \in \Xi_t} \left\{ y_t^\top (b_t - B_t x_{t-1}) + \min_{x_t \geq 0}\left\{ (c_t - A_t^\top y_t)^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[\widehat{Q}_{t+1}\big(x_t, \widehat{\boldsymbol{b}}_{t+1}\big)\right]\right\} - \lambda \|b_t - \widehat{\boldsymbol{b}}_t\| \right\} \\
&\qquad + \infty \cdot \mathbf{1}\left\{\lambda < \max_{s=t+1,\ldots,T} \|\ell_s\|_{\mathrm{Lip}}\right\} \\
&= \max_{y_t \in \mathbb{R}^{d_t}} \sup_{b_t \in \Xi_t} \left\{ y_t^\top (b_t - B_t x_{t-1}) - \widehat{\psi}_t^*(A_t^\top y_t) - \lambda \|b_t - \widehat{\boldsymbol{b}}_t\| \right\} + \infty \cdot \mathbf{1}\left\{\lambda < \max_{s=t+1,\ldots,T} \|\ell_s\|_{\mathrm{Lip}}\right\}.
\end{aligned}
$$

Observe that the function ℓ_t is convex and Lipschitz. It follows that

$$
\begin{aligned}
&\sup_{b_t \in \Xi_t} \left\{ y_t^\top (b_t - B_t x_{t-1}) - \widehat{\psi}_t^*(A_t^\top y_t) - \lambda \|b_t - \widehat{\boldsymbol{b}}_t\| \right\} \\
&= \begin{cases} \sup_{b_t \in \Xi_t} \left\{ y_t^\top (b_t - B_t x_{t-1}) - \widehat{\psi}_t^*(A_t^\top y_t) - \lambda \|b_t - \widehat{\boldsymbol{b}}_t\| \right\}, & \text{if } \lambda \geq \|\ell_t\|_{\mathrm{Lip}}, \\ \infty, & \text{otherwise.} \end{cases}
\end{aligned}
$$

Hence

$$
\begin{aligned}
Q_t(x_{t-1}, \widehat{\boldsymbol{b}}_t) &= \max_{y_t \in \mathbb{R}^{d_t}} \left\{ (\widehat{\boldsymbol{b}}_t - B_t x_{t-1})^\top y_t - \widehat{\psi}_t^*(A_t^\top y_t) \right\} + \infty \cdot \mathbf{1}\left\{\lambda < \max_{s=t,\ldots,T} \|\ell_s\|_{\mathrm{Lip}}\right\} \\
&= \min_{x_t \in \mathcal{X}_t(x_{t-1}, \widehat{\boldsymbol{b}}_t)} \left\{ c_t^\top x_t + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}\left[\widehat{Q}_{t+1}(x_t, \widehat{\boldsymbol{b}}_{t+1})\right]\right\} + \infty \cdot \mathbf{1}\left\{\lambda < \max_{s=t,\ldots,T} \|\ell_s\|_{\mathrm{Lip}}\right\}.
\end{aligned}
$$

Furthermore, we have $\lambda^\star = \max_{s=2,\ldots,T} \|\ell_s\|_{\mathrm{Lip}}$.

It remains to compute

$$
\|\ell_t\|_{\mathrm{Lip}} = \sup_{y \in \mathrm{dom}\, \widehat{\psi}_t^*(A_t^\top \cdot)} \|y\|_*. \tag{EC.8}
$$

When $t = T$, we have

$$
\widehat{\psi}_T^*(A_T^\top y_T) = \max_{x \geq 0} \left\{ y_T^\top A_T x - c_T^\top x \right\},
$$

which is zero when $A_T^\top y_T \le c_T$ and infinite otherwise. Thus dom $\widehat{\psi}_T^*(A_T^\top \cdot) = \{y \in \mathbb{R}^{d_T} : A_T^\top y \le c_T\}$ and dom $\widehat{\psi}_T^*(A_T^\top \cdot) = \mathcal{S}_T$. Now suppose (EC.8) holds for some $t+1$, let us prove the case of t, where $t \in [T-1]$. By definition of $\widehat{\psi}_t^*$ and the induction hypothesis, we have

$$
\begin{aligned}
&\widehat{\psi}_t^*(A_t^\top y_t)\\
&= \max_{x_t \ge 0} \left\{ y_t^\top A_t x_t - c_t^\top x_t - \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[\max_{y_{t+1} \in \mathbb{R}^{d_{t+1}}} \left\{ (\widehat{\boldsymbol{b}}_{t+1} - B_{t+1} x_t)^\top y_{t+1} - \widehat{\psi}_{t+1}^*(A_{t+1}^\top y_{t+1}) \right\} \right] \right\}\\
&= \max_{x_t \ge 0} \left\{ y_t^\top A_t x_t - c_t^\top x_t - \max_{\boldsymbol{y}_{t+1} \in \mathcal{Y}_{t+1}} \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[(\widehat{\boldsymbol{b}}_{t+1} - B_{t+1} x_t)^\top \boldsymbol{y}_{t+1}(\widehat{\boldsymbol{b}}_{t+1}) - \widehat{\psi}_{t+1}^*(A_{t+1}^\top \boldsymbol{y}_{t+1}(\widehat{\boldsymbol{b}}_{t+1})) \right] \right\}\\
&= \max_{x_t \ge 0} \min_{\boldsymbol{y}_{t+1} \in \mathcal{Y}_{t+1}} \left\{ y_t^\top A_t x_t - c_t^\top x_t - \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[(\widehat{\boldsymbol{b}}_{t+1} - B_{t+1} x_t)^\top \boldsymbol{y}_{t+1}(\widehat{\boldsymbol{b}}_{t+1}) - \widehat{\psi}_{t+1}^*(A_{t+1}^\top \boldsymbol{y}_{t+1}(\widehat{\boldsymbol{b}}_{t+1})) \right] \right\}\\
&= \min_{\boldsymbol{y}_{t+1} \in \mathcal{Y}_{t+1}} \left\{ \max_{x_t \ge 0} \left\{ y_t^\top A_t x_t - c_t^\top x_t - \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[(\widehat{\boldsymbol{b}}_{t+1} - B_{t+1} x_t)^\top \boldsymbol{y}_{t+1}(\widehat{\boldsymbol{b}}_{t+1}) \right] \right\} + \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[\widehat{\psi}_{t+1}^*(A_{t+1}^\top \boldsymbol{y}_{t+1}(\widehat{\boldsymbol{b}}_{t+1})) \right] \right\},
\end{aligned}
$$

where the third equality follows from the strong duality of the polyhedral problem (Shapiro et al. 2021, Chapter 3.2). Observe that

$$
\begin{aligned}
&\max_{x_t \ge 0} \left\{ y_t^\top A_t x_t - c_t^\top x_t - \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[(\widehat{\boldsymbol{b}}_{t+1} - B_{t+1} x_t)^\top \boldsymbol{y}_{t+1}(\widehat{\boldsymbol{b}}_{t+1}) \right] \right\}\\
&= -\mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[\widehat{\boldsymbol{b}}_{t+1}^\top \boldsymbol{y}_{t+1}(\widehat{\boldsymbol{b}}_{t+1}) \right] + \max_{x_t \ge 0} \left\{ (A_t^\top y_t - c_t)^\top x_t - \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[-x_t^\top B_{t+1}^\top \boldsymbol{y}_{t+1}(\widehat{\boldsymbol{b}}_{t+1}) \right] \right\}\\
&= \begin{cases} -\mathbb{E}_{\widehat{\mathbb{P}}_{t+1}} \left[\widehat{\boldsymbol{b}}_{t+1}^\top \boldsymbol{y}_{t+1}(\widehat{\boldsymbol{b}}_{t+1}) \right], & \text{if } A_t^\top y_t + B_{t+1}^\top \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}[\boldsymbol{y}_{t+1}(\widehat{\boldsymbol{b}}_{t+1})] \le c_t,\\ \infty, & \text{otherwise.} \end{cases}
\end{aligned}
$$

Hence, $\widehat{\psi}_t^*(A_t^\top y_t)$ is finite if there exists $\boldsymbol{y}_{t+1} \in \mathcal{Y}_{t+1}$ such that $A_t^\top y_t + B_{t+1}^\top \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}[\boldsymbol{y}_{t+1}(\widehat{\boldsymbol{b}}_{t+1})] \le c_t$. Thus we have shown that

$$
\operatorname{dom} \widehat{\psi}_t^*(A_t^\top y_t) = \left\{ y_t \in \mathbb{R}^{d_t} : \exists \boldsymbol{y}_{t+1} \in \mathcal{Y}_{t+1} \ s.t.\ A_t^\top y_t + B_{t+1}^\top \mathbb{E}_{\widehat{\mathbb{P}}_{t+1}}[\boldsymbol{y}_{t+1}(\widehat{\boldsymbol{b}}_{[t+1]})] \le c_t \right\}.
$$

which completes the induction for (EC.8). □

EC.4. Additional Details for Section 5

EC.4.1. Proof of Corollary 5

Introducing variables $\boldsymbol{x}_T^\pm, s_1, s_2 \ge 0$ we rewrite the problem as

$$
\begin{aligned}
\min_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{T-1} \ge 0, \boldsymbol{x}_T^\pm, W_T} \ & \max_{\mathbb{P} \in \mathfrak{M}} \mathbb{E}[\boldsymbol{x}_T^+ - \boldsymbol{x}_T^-]\\
s.t.\ \ & \mathbf{1}^\top \boldsymbol{x}_1 = W_1,\\
& \mathbf{1}^\top \boldsymbol{x}_{t-1} = \boldsymbol{\xi}_{t-1}^\top \boldsymbol{x}_{t-2}, \quad t = 2, \ldots, T,\\
& W_T = \boldsymbol{\xi}_T^\top \boldsymbol{x}_{T-1},\\
& \boldsymbol{x}_T^+ - \boldsymbol{x}_T^- - s_1 = -\alpha_0 - \beta_0 \boldsymbol{\xi}_T^\top x_{T-1},\\
& \boldsymbol{x}_T^+ - \boldsymbol{x}_T^- - s_2 = -\alpha_1 - \beta_1 \boldsymbol{\xi}_T^\top x_{T-1}.
\end{aligned}
$$

Then the result follows from substituting the following parameter values in Corollary 3:

$$
\begin{aligned}
&A_1 = \mathbf{1}^\top,\ B_1 = 0,\ b_1 = W_1,\ c_1 = \mathbf{0},\\
&A_t = \mathbf{1}^\top,\ B_t = -\boldsymbol{\xi}_t^\top,\ b_t = 0,\ c_t = \mathbf{0},\\
&A_T = \begin{pmatrix} 1 & -1 & -1 & 0 \\ 1 & -1 & 0 & -1 \end{pmatrix}, B_T = (\beta_0 \boldsymbol{\xi}_T^\top, \beta_1 \boldsymbol{\xi}_T^\top)^\top,\ b_T = (-\alpha_0, -\alpha_1)^\top,\ c_T = (1, -1, 0, 0)^\top.
\end{aligned}
$$

□

EC.4.2. Data

The estimated mean vector and covariance matrix are as the following.

$$\widehat{\mu} = \begin{array}{c} \begin{array}{ccccc} \text{EEM} & \text{TLT} & \text{SCHP} & \text{XES} & \text{SKF} \end{array} \\ \begin{bmatrix} 0.005142 & 0.006166 & 0.004729 & -0.005437 & -0.026391 \end{bmatrix} \end{array}$$

$$\widehat{\Sigma} = \begin{array}{c} \begin{array}{ccccc} \text{EEM} & \text{TLT} & \text{SCHP} & \text{XES} & \text{SKF} \end{array} \\ \begin{bmatrix} 0.002862 & -0.000634 & 0.000211 & 0.007217 & -0.003882 \\ -0.000634 & 0.001495 & 0.000231 & -0.002881 & 0.001658 \\ 0.000211 & 0.000231 & 0.000111 & 0.000426 & -0.000244 \\ 0.007217 & -0.002881 & 0.000426 & 0.030046 & -0.013416 \\ -0.003882 & 0.001658 & -0.000244 & -0.013416 & 0.010840 \end{bmatrix} \end{array} \begin{array}{c} \text{EEM} \\ \text{TLT} \\ \text{SCHP} \\ \text{XES} \\ \text{SKF} \end{array}$$

EC.4.3. Algorithms

Algorithm 1 SDDP Algorithm for Robust Reformulation with Uncertainty in $\boldsymbol{B}_t$

Initialize: $\{\mathfrak{Q}_t\}_{t\in[T]}$ (initial lower approximation), $\bar{x}_0^k = 0$

while not converge **do**

  Sample K scenario paths $\{\widehat{\boldsymbol{B}}^k_{[T]}\}_{k=1}^K$

  for $t = 1, \dots, T-1$ **do** ▹ *Forward pass*

    for $k = 1, \dots, K$ **do**

      for $(\delta, j) \in \{1, -1\} \times [m_t]$ **do**

        $(x_{tk}^{\delta j}, v_{tk}^{\delta j}) \leftarrow \min_{x_t \geq 0} \left\{ c_t x_t + \mathfrak{Q}_{t+1}(x_t) : \; A_t x_t + \widehat{\boldsymbol{B}}_t^k \bar{x}_{t-1}^k = b_t + \vartheta\delta \|\bar{x}_{t-1}^k\| e_j \right\}$ ▹ *Optimal solution and optimal value*

      end for

      $(\delta^*, j^*) \leftarrow \arg\max_{(\delta, j)} v_{tk}^{\delta j}$; $\bar{x}_t^k \leftarrow x_{tk}^{\delta^* j^*}$

    end for

  end for

  for $t = T, \dots, 2$ **do** ▹ *Backward pass*

    for $k = 1, \dots, K$ **do**

      for $i = 1, \dots, \widehat{N}_t$ **do**

        for $(\delta, j) \in \{1, -1\} \times [m_t]$ **do**

          $(v_{ti}^{\delta j}(\bar{x}_{t-1}^k); y_{tki}^{\delta j}) \leftarrow \min_{x_t \geq 0} \left\{ c_t x_t + \mathfrak{Q}_{t+1}(x_t) : \; A_t x_t + \widehat{\boldsymbol{B}}_t^i \bar{x}_{t-1}^k = b_t + \vartheta\delta \|\bar{x}_{t-1}^k\| e_j \right\}$ ▹ *Optimal value and optimal dual solution*

        end for

        $(\delta^\star, j^\star) \leftarrow \arg\max_{(\delta, j)} v_{ti}^{\delta j}$; $\tilde{Q}_{ti}(\bar{x}_{t-1}^k) \leftarrow v_{ti}^{\delta^\star j^\star}$; $\pi_{ti}^k \leftarrow y_{tik}^{\delta^\star j^\star}$

      end for

      $\tilde{\mathfrak{Q}}_t(\bar{x}_{t-1}^k) \leftarrow \frac{1}{N_t} \sum_{i=1}^{N_t} \tilde{Q}_{ti}(\bar{x}_{t-1}^k)$; $\tilde{g}_t^k \leftarrow \frac{1}{N_t} \sum_{i=1}^{N_t} \vartheta \delta^\star {\pi_{ti}^k}^\top e_{j^\star} \nabla \|\bar{x}_{t-1}^k\| - (\widehat{\boldsymbol{B}}_t^i)^\top \pi_{ti}^k$

      $\mathfrak{Q}_t(\cdot) \leftarrow \max \left(\mathfrak{Q}_t(\cdot), \; \tilde{\mathfrak{Q}}_t(\bar{x}_{t-1}^k) + {\tilde{g}_t^k}^\top (\cdot - \bar{x}_{t-1}^k) \right)$

    end for

  end for

end while

Algorithm 2 Out-of-sample Test

Input: A training tree $(\widehat{\boldsymbol{\xi}}_1, \{\widehat{\boldsymbol{\xi}}_2^{i_2}\}_{i_2=1}^{\widehat{N}_2}, \ldots, \{\widehat{\boldsymbol{\xi}}_1^{i_T}\}_{i_T=1}^{\widehat{N}_T})$, a testing tree $(\boldsymbol{\xi}_1, \{\boldsymbol{\xi}_2^{i_2}\}_{i_2=1}^{N_1}, \ldots, \{\boldsymbol{\xi}_1^{i_T}\}_{i_T=1}^{N_T})$, number of testing paths M

Output: Average out-of-sample value V_{Avg}

```
for m = 1 : M do
    V_m ← 0
    Sample a path (ξ_1, ξ_2^{i_2}, ..., ξ_T^{i_T}) from the testing tree
    for t = 1 : T do
        Solve for the optimal decision x_t using x_{t-1} and the training sub-tree ({ξ̂_t^{i_t}}_{i_t=1}^{N̂_t}, ..., {ξ̂_1^{i_T}}_{i_T=1}^{N̂_T})
        Observe a testing scenario ξ_t^{i_t}
        Evaluate the per-stage out-of-sample cost C_t at stage t using x_t and ξ_t^{i_t}
        V_m ← V_m + C_t
    end for
end for
V_Avg ← (1/M) Σ_{m=1}^{M} V_m
```