Mathematical Programming manuscript No.
(will be inserted by the editor)

A Stability Result for Linear Markov Decision Processes

David Wozabal · Adriana Kiszka

Received: date / Accepted: date

Abstract In this paper, we propose a semi-metric for Markov processes that allows to bound optimal values of linear Markov Decision Processes (MDPs). Similar to existing notions of distance for general stochastic processes our distance is based on transportation metrics. Apart from the specialization to MDPs, our contribution is to make the distance problem specific, i.e., explicitly dependent on the data of the problem whose objective value we want to bound. As a result, we are able to consider problems with randomness in the constraints as well as in the objective function and therefore relax an assumption in the extant literature. We derive several properties of the proposed semi-metrics and demonstrate its use in a stylized numerical example.

Keywords Stochastic Optimization · Wasserstein Distance · Scenario Lattices

Mathematics Subject Classification (2010) 90C15 · 90C31 · 90C40 · 60J05

1 Introduction

Stochastic optimization is concerned with the solution of optimization problems that involve random quantities as data. Consequently, the decisions $x(\xi)$ depend on values of a random process ξ, making stochastic optimization a

David Wozabal
School of Management, Technical University of Munich, Arcisstraße 21, 80333 Munich, Germany
E-mail: david.wozabal@tum.de

Adriana Kiszka
School of Management, Technical University of Munich, Arcisstraße 21, 80333 Munich, Germany
E-mail: adriana.kiszka@tum.de

problem in function spaces. Mirroring the situation in deterministic optimization, only few stochastic optimization problems lend itself to analytical treatment and allow for closed form solutions. Consequently, in the following, we focus on discrete time problems that are solved with numerical methods.

The theory of stochastic optimization as well as the development of solution methods made great advances in the last decades. In particular, there exists a sound theory for two-stage stochastic optimization problems, i.e., problems with only one decision stage in the future (see [3] for an overview). Two-stage stochastic optimization is nowadays routinely applied by many researchers and industry practitioners alike. The resulting methods are based on discrete representations of the, possible continuous, source of randomness in the form of a finite set of samples or scenarios. This can either be achieved by sample average approaches (see [35] for an introduction) or by explicitly choosing representative scenarios. In this paper, we will focus on the latter.

Despite the abovementioned successes, it became clear quite early that the computational complexity of stochastic optimization does not scale well in the number of random variables involved in the problem formulation. More specifically, it has been shown that stochastic optimization problems exhibit non-polynomial increase in complexity, as the number of random variables increases [11]. The problem underlying these difficulties is the numerical evaluation of high dimensional integrals, which in turn is related to the problem of optimal discretization of probability distributions.

The situation is even more complicated for multi-stage problems, where we deal with random processes resulting in additional random variables in each additional stage and the issue of finding discretizations for conditional distributions. Consequently, it was observed in [32, 33] that solving multi-stage stochastic optimization problems is often practically intractable.

Notwithstanding these problems, there is a rich literature on multi-stage stochastic optimization. The majority of authors use scenario trees as representation of discrete stochastic processes (see the left panel in Figure 1 for a depiction). In a scenario tree, nodes represent possible states of the world and are assigned to a point in time. All nodes that represent states for the same point in time are usually depicted at the same level of the tree. Possible transitions between nodes in consecutive stages are depicted by probability weighted arcs connecting the nodes. Consequently, the collection of transition probabilities between a node and the nodes of the next stage connected by arcs describes the distribution of the random process conditional on that node. Note that the resulting graph structure is required to be a tree, implying that every node is allowed to have exactly one predecessor on the previous stage.

There are various ways to construct scenario trees for multi-stage stochastic programs (see [7, 19] for surveys). In [16, 17] a recursive application of moment matching is presented. The approach is easy to understand and apply, but suffers from an exponential explosion of nodes in the resulting trees as the number of stages increases and offers no theoretical insight regarding

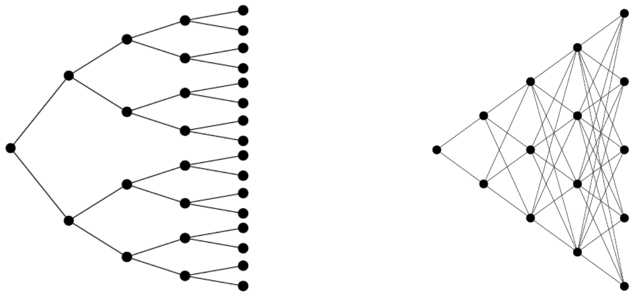


Fig. 1 A scenario tree with 31 nodes representing 16 scenarios on the left and a scenario lattice with 15 nodes representing 120 scenarios on the right. The transition probabilities on the arcs are not depicted to keep the picture legible.

discretization error made when replacing the original process by the resulting tree.

[23, 24] propose a method for the construction of scenario trees that is based on integration quadratures and ensures that the approximated problems based on scenario trees epi-converge to the true infinite dimensional problem yielding convergence in optimal value as well as in optimal decisions. However, the results are asymptotic in nature, i.e., the approximation scheme doesn't offer finite sample guarantees for any given discrete approximation.

Another approach is based on the principle of bound-based constructions [5, 9, 10, 20]. The idea is to construct two discrete stochastic programs which provide upper and lower bounds on the optimal value of the original problem.

The results in this paper extend a stream of literature that uses probability metrics to define notions of distance for stochastic processes and allows inference about the accuracy of approximating trees [8, 12, 13, 26, 27, 28]. [8, 12] consider a distance between discrete stochastic processes and assume that both processes are defined on the same probability space. This assumption is relaxed in [27, 28] who develop a nested distance between value-and-information structures which can be applied to continuous processes. [14, 15] prove stability results using the sum of a L_r-distance and a filtration distance to bound objective values of a certain class of stochastic optimization problems.

Scenario trees are discrete approximations of general processes and therefore lend themselves to the construction of a general theory of stochastic optimization. However, the requirement that every node only has one predecessor, makes it hard to construct scenario trees with many stages that model the conditional distributions well, i.e., ensure every node has a sufficient number of successors, and at the same time avoid an exponential growth of the number of nodes.

A possible way out of this dilemma is to restrict the type of the stochastic optimization problem to MDPs where the random processes in the problem formulation are Markov processes [21] or, even more common, independent [25]. In this setting the *history* of the random variables and the decisions is condensed in the state variable of the problem and there is no need to *remember* the whole history of the randomness and the decisions. This paves the way for leaner discretizations, which we call a *scenario lattices* in this paper. A scenario lattice consists of the same building blocks as a scenario tree but relaxes the requirement that every node has only one predecessor and therefore solves the problem of the exponential explosion of the number of nodes as the number of stages grows (see the right panel in Figure 1). In the same way that a scenario tree is a natural representation for a general discrete stochastic process, a scenario lattice is a natural represention of a discrete Markov process.

Even though the abovementioned problem class is quite popular there are practically no theoretical results on how to construct optimal scenario lattices for MDPs. An exception is [1, 2] who design an algorithm for the construction of scenario lattices for Brownian motions based on ideas of optimal quantization.

With this paper we contribute to the development of a theory for discrete approximations of Markov processes to be used in MDPs. In particular, we propose a class of problem specific semi-distances for Markov processes and show that the objective values of a certain class of linear MDPs is Lipschitz continuous with respect to these distances. This lays the foundations for constructing scenario lattices approximating general Markov processes that in turn can be used to formulate approximating optimization problems. In particular, the results in this paper, can be used to control the error that results from replacing a MDP that is formulated using a complex (possibly continuous) Markov process by another simpler MDP using a compact scenario lattice instead of the original process. Furthermore, we discuss a LP formulation of our distance for discrete Markov processes, i.e., lattices, and demonstrate the use of our results in a multi-stage version of the well known news-vendor problem.

Our approach is inspired by [27] who work on optimal scenario trees and general stochastic optimization problems. In contrast to [27] our approach is specialized to linear MDPs, which results in tighter bounds for this problem class and additionally allows for problems where the randomness does not only affect the objective function but also the feasible set. The latter makes it necessary to adopt a different technique of proof based on stability results for linear programs rather than the idea of transporting solutions from one problem to the other. To the best of our knowledge, we are the first to propose stability results based on transportation distances that encompass problems where the feasible set depends on randomness via inequality constraints: [8, 12, 27, 28] require the feasible set to be independent of randomness, while in [14, 15] the constraints involving random parameters are required to be equality constraints.

This paper is structured as follows: In Section 2, we introduce some notation and the problem setup. In Section 3, we define the problem dependent lattice distance and establish some of its key properties. Section 4 contains the main results of the paper which allow to connect the lattice distance to optimal values of linear MDPs, while Section 5 is devoted to the case of discretely supported processes representable by lattices and a numerical example. Section 6 concludes the paper.

2 Problem Description

We consider a class of discrete time, finite horizon, linear MDPs, i.e., stochastic dynamic programming problems depending on a Markov process. The time periods in our problem are indexed by $t \in \mathbb{T} = \{0, 1, \dots, T\}$, where $t = 0$ represents the deterministic start state of the problem. We partition the state space in an *environmental state* ξ and a *resource state* S. The former is governed by a (possibly inhomogeneous) Markov process $\xi = (\xi_0, \xi_1, \dots, \xi_T)$, $\xi_t : \Omega_t \to \mathbb{R}^{n_t}$ which is assumed to be independent of the decisions. Examples are prices, demand for a product, or weather related variables such as temperature. The resource state S_t on the other hand describes the part of the state space that is influenced by the decision maker. Examples include inventory levels, states of machinery, and contractual obligations.

We equip the probability space Ω_t with the σ-algebra $\sigma_t = \sigma(\xi_t)$ and define the path space $\Omega = \Omega_1 \times \dots \times \Omega_T$ and the corresponding σ-algebra $\mathcal{F} = \sigma_1 \otimes \dots \otimes \sigma_T$. Note that we base our σ-algebras only on the random variables ξ_t and not on the whole history of random variables until t as it is usually done when working with scenario trees. Consequently, $\sigma_1, \dots, \sigma_T$ is not a filtration.

Furthermore, we define the paths for which the event $A \in \sigma_t$ happens as

$$A_t^\Omega := \Omega_1 \times \dots \times \Omega_{t-1} \times A \times \Omega_{t+1} \times \dots \times \Omega_T.$$

and the corresponding σ-algebra as

$$\sigma_t^\Omega = \{\Omega_1 \times \dots \times \Omega_{t-1} \times A \times \Omega_{t+1} \times \dots \times \Omega_T : A \in \sigma_t\} = \{A_t^\Omega : A \in \sigma_t\}.$$

We will additionally require

$$\sigma_{t-1,t}^\Omega = \{A_{t-1}^\Omega \cap B_t^\Omega : A \in \sigma_{t-1}, B \in \sigma_t\}$$

linking information from two consecutive stages.

The distribution of ξ is described by a sequence of Markov kernels and we write $P_t^{\omega_{t-1}}$ for the distribution of ξ_t given $\omega_{t-1} \in \Omega_{t-1}$. The kernel as a function from Ω_{t-1} to the set of probability measures on Ω_t is σ_{t-1}-measurable. For a given sequence of Markov kernels, we define the distribution on Ω as

$$P(A) := \int_{\Omega_1} \dots \int_{\Omega_T} \mathbb{1}_A(\omega) P_T^{\omega_{T-1}}(d\omega_T) \dots P_1^{\omega_0}(d\omega_1)$$

for every $A \in \mathcal{F}$ and $\omega = (\omega_1, \dots, \omega_T)$.

We consider MDPs that can be written as

$$V_0(S_0, \xi_0) = \begin{cases} \max_x \mathbb{E}\left(\sum_{t=0}^T c_t(\xi_t)^\top x_t\right) \\ \text{s.t.} \quad (x_t, S_{t+1}) \in \mathcal{X}_t(S_t, \xi_t), \ \forall t = 0, 1, \dots, T \end{cases} \tag{1}$$

with feasible sets

$$\mathcal{X}_t(S_t,\xi_t) = \left\{ (x_t, S_{t+1}) : \begin{array}{l} A_{1,t}x_t \leq b_{1,t}(\xi_t) + C_{1,t}S_t \\ A_{2,t}x_t = S_{t+1} \\ A_{2,t}x_t \leq b_{2,t+1} \\ x_t, S_{t+1} \geq 0, \end{array} \right\}. \tag{2}$$

In particular, we assume that for planning in stage t, the decision maker knows S_t, i.e., the system's resource state at the beginning of the period as well as ξ_t, i.e., the realization of the Markov process in period t and given this information the feasible set for the decision x_t as well as the definition of S_{t+1} can be expressed using linear inequality constraints. The decisions x_t are auxiliary decision variables in stage t that are not part of the resource state.

Remark 1 Usually we would expect a state transition equation of the form $S_{t+1} = S_t + Ax_t$. However, due to the requirement to make our distance independent of the resource state, we formulate the state transition using x_t. More specifically, we assign S_t to variables in x_t in the first constraint. The state transition is subsequently modelled in the equality constraint using those variables instead of S_t. Alternatively, we could assign S_{t+1} to variables in x_t in the equality constraint and then formulate the state transition using the first inequality constraint. We refer to the example in Section 5 for an illustration of this principle.

Because of its recursive structure, problem (1) can be equivalently written in terms of its dynamic programming equations using value functions, i.e.,

$$V_t(S_t,\xi_t) = \begin{cases} \max\limits_{x_t} c_t(\xi_t)^\top x_t + \mathbb{E}\left(V_{t+1}(S_{t+1},\xi_{t+1})|\xi_t\right) \\ \text{s. t.} \quad (x_t, S_{t+1}) \in \mathcal{X}_t(S_t,\xi_t) \end{cases}, \quad \forall t \in \mathbb{T} \tag{3}$$

and $V_{T+1} \equiv 0$ or, more generally, a known piecewise linear concave function. Since ξ is a Markov process and V_t only depends on the current state of the system, the problem is a MDP [31].

If we are dealing with discrete Markov processes, the expectations of the value functions V_t, which are concave functions of the resource state, can be written as a minimum of finitely many affine functions. We formalize this well known fact in the following lemma whose proof can be found for example in [21, 29, 34].

Lemma 1 *If ξ is finitely supported, then, for every realization of ξ_t, $S_{t+1} \mapsto \mathbb{E}\left(V_{t+1}(S_{t+1},\xi_{t+1})|\xi_t\right)$ is a concave, polyhedral function. In particular, there is a set $\mathcal{I}_t(\xi_t)$ and coefficients (β_i^0, β_i) such that*

$$\mathbb{E}\left(V_{t+1}(S_{t+1},\xi_{t+1})|\xi_t\right) = \min_{i\in\mathcal{I}_t(\xi_t)} \beta_i^0 + \beta_i^\top S_{t+1}.$$

3 A Distance for Markov Processes

In order to introduce the concept of a distance between Markov processes, we first recall the Wasserstein or Kantorovich distance for distributions [18, 36]. Loosely speaking, the Wasserstein distance is defined as the total cost of passing from a given distribution to a desired one by *moving* probability masses accordingly.

Definition 1 Let $\xi : (\Omega, \mathcal{A}) \to \mathbb{R}^n$ and $\tilde{\xi} : (\tilde{\Omega}, \tilde{\mathcal{A}}) \to \mathbb{R}^n$ be two random variables with distributions P and $\tilde{P}$, respectively. The Wasserstein distance of order r $(r \geq 1)$ between ξ and $\tilde{\xi}$ is defined as

$$W_r(\xi, \tilde{\xi}) = \begin{cases} \inf\limits_{\pi} \left(\int\limits_{\Omega \times \tilde{\Omega}} ||\xi(\omega) - \tilde{\xi}(\tilde{\omega})||_r^r \, \pi\,(d\omega, d\tilde{\omega}) \right)^{\frac{1}{r}} \\ \text{s.t. } \pi(A \times \tilde{\Omega}) = P(A), \quad \forall A \in \mathcal{A} \\ \qquad \pi(\Omega \times B) = \tilde{P}(B), \quad \forall B \in \tilde{\mathcal{A}}, \end{cases} \tag{4}$$

where the infimum runs over all probability measures π on $(\Omega \times \tilde{\Omega}, \mathcal{A} \otimes \tilde{\mathcal{A}})$.

Remark 2 The above problem is bounded and an optimal transportation measure π exists, due to weak-compactness of the set of transportation plans (see [36], Lemma 4.4). Furthermore, according to the famous Kantorovich-Rubinstein Theorem, for $r = 1$, the dual of (4) can be written as the following maximization problem

$$W_1(\xi, \tilde{\xi}) = \begin{cases} \sup\limits_{f} \left(\int f dP - \int f d\tilde{P} \right) \\ \text{s.t. } \operatorname{Lip}(f) \leq 1, \end{cases}$$

where $\operatorname{Lip}(f)$ is the Lipschitz constant of f.

Clearly, for a two-stage stochastic optimization problem

$$v(P) = \begin{cases} \inf\limits_{x} \; f(x) + \mathbb{E}_P(Q(x, \xi)) \\ \text{s.t. } x \in \mathcal{X} \end{cases}, \quad Q(x, \xi) = \begin{cases} \inf\limits_{y} \; g(y, \xi) \\ \text{s.t. } y \in \mathcal{Y}(x) \end{cases}$$

with

$$|Q(x, \xi) - Q(x, \xi')| \leq L \; ||\xi - \xi'||_1, \tag{5}$$

we have

$$v(P) - v(\tilde{P}) \leq \mathbb{E}_P(Q(\tilde{x}^*, \xi)) - \mathbb{E}_{\tilde{P}}(Q(\tilde{x}^*, \tilde{\xi})) \leq L \; W_1(\xi, \tilde{\xi})$$

where $\tilde{x}^*$ is the optimal solution for $v(\tilde{P})$. By symmetry it follows that

$$|v(P) - v(\tilde{P})| \leq L \; W_1(\xi, \tilde{\xi}),$$

i.e., that the objective value of the two-stage stochastic program is Lipschitz continuous with respect to W_1, as long as the cost-to-go function Q is Lipschitz in ξ. This was first recognized by [26].

[27, 28] generalize these ideas to a multi-stage setting using the notion of nested distributions which correspond to *generalized* scenario trees. Based on a modified transportation problem and an assumption similar to the uniform Lipschitz property in (5), they obtain a distance with respect to which the objective value of a general multi-stage problem is Hölder continuous.

We aim for a similar result for scenario lattices and problems of the form (1). Additionally, we relax one major assumption in the abovementioned approaches, namely that randomness enters the problem only in the objective function. Observe, that the argument above hinges on the fact that the set $\mathcal{Y}$ does not depend on ξ. The same restriction applies to the results on multi-period problems in [27, 28].

We begin with analysing the following simple deterministic linear optimization problem, which is of a similar structure as (3), with the last inequality constraint and the second term in the objective function modeling the piecewise linear value function (see Lemma 1)

$$\max_{x\in\mathbb{R}^n, y\in\mathbb{R}^m, z\in\mathbb{R}^k} \left\{ c_1^\top x + c_2^\top y : \begin{array}{l} A_1 x \le b_1 \\ A_2 x = z \\ A_2 x \le b_2 \\ A_3 y \le b_3 + C_3 z \\ x, z \ge 0 \end{array} \right\}. \tag{6}$$

We first prove the following result which is motivated by Hoffmann's Lemma and in particular its discussion in [35], Theorem 7.11 and Theorem 7.12. For what follows, we adopt the notational convention that the addition of a vector $x = (x_1, \dots, x_n)$ and a scalar $y \in \mathbb{R}$ is to be interpreted pointwise, i.e., results in the vector $(x_1 + y, \dots, x_n + y)$ and similarly inequalities of the form $x \le y$ are interpreted pointwise as well.

Lemma 2 *Let $V(b_1)$ be the optimal value of problem* (6) *in dependence on the parameter b_1. There exists a positive constant $\gamma(A_1, A_2)$, depending only on A_1, A_2, such that for any b_1, b_1' for which* (6) *is feasible*

$$|V(b_1) - V(b_1')| \le \gamma(A_1, A_2)||c_1||_\infty ||b_1 - b_1'||_1. \tag{7}$$

Proof We start by rewriting (6) as

$$\max_{t\in\mathbb{R}, x\in\mathbb{R}^n, y\in\mathbb{R}^m, z\in\mathbb{R}^k} \left\{ t : \begin{array}{l} t - \tilde{c}_1^\top x - \tilde{c}_2^\top y \le 0 \\ A_1 x \le b_1 \\ A_2 x = z \\ A_2 x \le b_2 \\ A_3 y \le b_3 + C_3 z \\ x, z \ge 0 \end{array} \right\} \tag{8}$$

where $\tilde{c}_1 = \frac{c_1}{||c_1||_\infty}$ and $\tilde{c}_2 = \frac{c_2}{||c_1||_\infty}$.

Denote by $\mathcal{M}(b_1)$ the set of feasible points of problem (8) and consider a point $\alpha = (x, y, z, t) \in \mathcal{M}(b_1)$. Note that for any $a \in \mathbb{R}^n$, $||a||_1 = \sup_{||u||_\infty \le 1} u^\top a$ and

$$\begin{aligned}\operatorname{dist}((x,y,z,t),\mathcal{M}(b_1')) &= \inf_{\alpha'\in\mathcal{M}(b_1')} ||\alpha-\alpha'||_1 \\ &= \inf_{\alpha'\in\mathcal{M}(b')} \sup_{||u||_\infty\le 1} u^\top(\alpha-\alpha') \\ &= \sup_{||u||_\infty\le 1} \inf_{\alpha'\in\mathcal{M}(b_1')} u^\top(\alpha-\alpha').\end{aligned}$$

By changing variables defining $w = (w_1, w_2, w_3, w_4) = \alpha - \alpha'$ and using linear optimization duality with $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7)$, we have

$$\begin{aligned}\inf_{\alpha'\in\mathcal{M}(b_1')} u^\top(\alpha-\alpha') = \inf_{w\in\tilde{\mathcal{M}}(b_1')} u^\top w = \sup_{\lambda\in\tilde{\mathcal{M}}^*(u)} &\lambda_1^\top(t-\tilde{c}_1^\top x-\tilde{c}_2^\top y) \\ &+\lambda_2^\top(A_1x-b_1')+\lambda_4^\top(A_2x-b_2)+\lambda_5^\top(A_3y-b_3-C_3z) \\ &+\lambda_6^\top(-x)+\lambda_7^\top(-z),\end{aligned}$$

where

$$\tilde{\mathcal{M}}(b_1') = \left\{ w : \begin{array}{l} t-\tilde{c}_1^\top x-\tilde{c}_2^\top y \le w_4-\tilde{c}_1^\top w_1-\tilde{c}_2^\top w_2, \\ A_1x-b_1' \le A_1w_1, \\ A_2w_1 = w_3, \\ A_2x-b_2 \le A_2w_1, \\ A_3y-b_3-C_3z \le A_3w_2-C_3w_3, \\ -x \le -w_1, -z \le -w_3 \end{array} \right\}$$

and

$$\tilde{\mathcal{M}}^*(u) = \left\{ \lambda : \begin{array}{l} -\tilde{c}_1\lambda_1+A_1^\top\lambda_2+A_2^\top\lambda_3+A_2^\top\lambda_4-\mathbb{I}_n\lambda_6 = u_1, \\ -\tilde{c}_2\lambda_1+A_3^\top\lambda_5 = u_2 \\ \lambda_3-C_3^\top\lambda_5-\lambda_7 = u_3, \\ \lambda_1 = u_4, \\ \lambda_1,\lambda_2,\lambda_4,\lambda_5,\lambda_6,\lambda_7 \ge 0 \end{array} \right\}.$$

Consequently we obtain that

$$\begin{aligned}\operatorname{dist}(\alpha,\mathcal{M}(b_1')) = \sup_{||u||_\infty\le 1,\lambda\in\tilde{\mathcal{M}}^*(u)} &\lambda_1^\top(t-\tilde{c}_1^\top x-\tilde{c}_2^\top y)+\lambda_2^\top(A_1x-b_1') \\ &+\lambda_4^\top(A_2x-b_2)+\lambda_5^\top(A_3y-b_3-C_3z) \\ &+\lambda_6^\top(-x)+\lambda_7^\top(-z).\end{aligned} \tag{9}$$

The right-hand side of (9) has a finite optimal value (since the left-hand side of (9) is finite) and, hence, has an optimal solution $(\hat{u}, \hat{\lambda})$. It follows that

$$\begin{aligned}\operatorname{dist}(\alpha,\mathcal{M}(b_1')) &= \hat{\lambda}_1^\top(t-\tilde{c}_1^\top x-\tilde{c}_2^\top y)+\hat{\lambda}_2^\top(A_1x-b_1') \\ &\quad+\hat{\lambda}_4^\top(A_2x-b_2)+\hat{\lambda}_5^\top(A_3y-b_3-C_3z) \\ &\quad+\hat{\lambda}_6^\top(-x)+\hat{\lambda}_7^\top(-z).\end{aligned}$$

Since $\alpha \in \mathcal{M}(b_1)$, we have

$$\begin{aligned}\operatorname{dist}(\alpha, \mathcal{M}(b_1')) &\leq \hat{\lambda}_2^\top (A_1 x - b_1') = \hat{\lambda}_2^\top (A_1 x - b_1) + \hat{\lambda}_2^\top (b_1 - b_1') \\ &\leq \hat{\lambda}_2^\top (b_1 - b_1') \leq ||\hat{\lambda}_2||_\infty ||b_1 - b_1'||_1.\end{aligned}$$

To find a bound for $||\hat{\lambda}_2||_\infty$, we analyze the extreme points of the feasible set

$$P := \left\{ \lambda : \begin{array}{l} ||-\tilde{c}_1\lambda_1 + A_1^\top \lambda_2 + A_2^\top \lambda_3 + A_2^\top \lambda_4 - \lambda_6||_\infty \leq 1 \\ ||-\tilde{c}_2\lambda_1 + A_3^\top \lambda_5||_\infty \leq 1 \\ ||\lambda_3 - C_3^\top \lambda_5 - \lambda_7||_\infty \leq 1 \\ ||\lambda_1||_\infty \leq 1 \\ \lambda_1, \lambda_2, \lambda_4, \lambda_5, \lambda_6, \lambda_7 \geq 0. \end{array} \right\}$$

Since we know that $||\tilde{c}_1||_\infty \leq 1$ and $|\lambda_1| \leq 1$, we can replace the constraint

$$||-\tilde{c}_1\lambda_1 + A_1^\top \lambda_2 + A_2^\top \lambda_3 + A_2^\top \lambda_4 - \lambda_6||_\infty \leq 1$$

with the constraint

$$||A_1^\top \lambda_2 + A_2^\top \lambda_3 + A_2^\top \lambda_4 - \lambda_6||_\infty \leq 2$$

and remove the other constraints to increase the feasible set of problem (9), and hence increase its optimal value. Moreover, since we are only interested in $\hat{\lambda}_2$, we can ignore variables which don't influence $\hat{\lambda}_2$. Consequently,

$$\max_{\lambda \in P} ||\lambda_2||_\infty \leq \max_{\lambda \in P'} ||\lambda_2||_\infty$$

with

$$P' := \left\{ (\lambda_2, \lambda_3, \lambda_6) : \begin{array}{l} ||A_1^\top \lambda_2 + A_2^\top \lambda_3 - \lambda_6||_\infty \leq 2 \\ \lambda_2, \lambda_6 \geq 0 \end{array} \right\}.$$

The polyhedral set P' depends only on A_1, A_2, and has a finite number of extreme points. Hence, $||\hat{\lambda}_2||_\infty$ can be bounded by a constant $\gamma(A_1, A_2)$ which depends on A_1, A_2, and

$$\operatorname{dist}(\alpha, \mathcal{M}(b_1')) \leq ||\hat{\lambda}_2||_\infty ||b_1 - b_1'|| \leq \gamma(A_1, A_2)||b_1 - b_1'||_1. \tag{10}$$

Assume that $\alpha = (x, y, z, t)$ is the optimal solution of problem (8) and $||c_1||_\infty t = V(b_1)$. Let further $\alpha' \in \mathcal{M}(b_1')$ be a point minimizing the distance $\operatorname{dist}(\alpha, \mathcal{M}(b_1'))$. Then (10) implies

$$|t - t'| \leq \gamma(A_1, A_2)||b_1 - b_1'||_1$$

and we obtain

$$\begin{aligned}V(b_1) - V(b_1') &\leq V(b_1) - ||c_1||_\infty t' = ||c_1||_\infty t - ||c_1||_\infty t' \\ &\leq ||c_1||_\infty \gamma(A_1, A_2)||b_1 - b_1'||_1.\end{aligned}$$

Analogously, we get

$$V(b_1') - V(b_1) \leq ||c_1||_\infty \gamma(A_1, A_2)||b_1 - b_1'||_1,$$

and finally (7). □

The above lemma allows to bound the variation in the objective as a function of the right hand side data of the problem (6). Next we will prove a similar result to bound the objective value when the objective coefficient c_1 changes.

Lemma 3 *Let $V(c_1)$ be the optimal value of the problem (6) in dependence on the objective value coefficient c_1. There exists a positive constant $\phi(A_1, b_1, A_2, b_2)$, depending on A_1, b_1, A_2, b_2, such that*

$$|V(c_1) - V(\tilde{c}_1)| \leq \phi(A_1, b_1, A_2, b_2)||c_1 - \tilde{c}_1||_1.$$

Proof Let (x^*, y^*, z^*) be an optimal solution to $V(c_1)$, then we have

$$\begin{aligned} V(c_1) &= c_1^\top x^* + c_2^\top y^* = c_1^\top x^* + c_2^\top y^* - \tilde{c}_1^\top x^* + \tilde{c}_1^\top x^* = \\ &= (c_1 - \tilde{c}_1)^\top x^* + \tilde{c}_1^\top x^* + c_2^\top y^* \leq ||c_1 - \tilde{c}_1||_1 ||x^*||_\infty + V(\tilde{c}_1) \end{aligned}$$

and by symmetry this implies

$$|V(c_1) - V(\tilde{c}_1)| \leq \max(||x^*||_\infty, ||\tilde{x}^*||_\infty)||c_1 - \tilde{c}_1||_1$$

for an optimal solution $(\tilde{x}^*, \tilde{y}^*, \tilde{z}^*)$ to $V(\tilde{c}_1)$. Notice that the set of feasible points is invariant with respect to the parameter c_1. Hence, x^* and $\tilde{x}^*$ are extreme points of the same polyhedral set

$$P := \{x \in \mathbb{R}^n : A_1 x \leq b_1, A_2 x \leq b_2, A_2 x \geq 0, x \geq 0\}.$$

P depends on A_1, b_1, A_2, b_2, and has a finite number of vertices. Therefore $||x^*||_\infty$ and $||\tilde{x}^*||_\infty$ can be bounded by a constant $\phi(A_1, b_1, A_2, b_2)$ and we get

$$|V(c_1) - V(\tilde{c}_1)| \leq \phi(A_1, b_1, A_2, b_2)||c_1 - \tilde{c}_1||_1$$

which finishes the proof. □

Remark 3 Applying the above lemma to the problem (3), $b_{1,t}(\xi_t) + C_{1,t}S_t$ corresponds to the second parameter of ϕ. Since we would like to avoid a dependence of our distance on the resource state, we note that ϕ is increasing with respect to this parameter and replace $b_{1,t}(\xi_t) + C_{1,t}S_t$ by $b_{1,t}(\xi_t) + C_{1,t}^+ b_{2,t}$ where $C_{1,t}^+ = (\max(c_{ij}, 0))_{i,j}$ and c_{ij} are the entries in the matrix $C_{1,t}$. In this way, we make the bound slightly looser and at the same time independent of S_t.

Equipped with the above results, we define a transportation distance between two Markov processes. The distance is defined for *a given problem* of the form (1), i.e., we do not propose one distance but a whole family of problem specific distances, which differ in the matrices and vectors used to define the constants γ and ϕ in Lemma 2 and Lemma 3. To avoid cluttered notation, we write $\gamma_t = \gamma(A_{1,t}, A_{2,t})$ and $\phi_t(\xi_t) = \phi(A_{1,t}, b_{1,t}(\xi_t) + C_{1,t}^+ b_{2,t}, A_{2,t}, b_{2,t+1})$. Furthermore, we omit the explicit dependence of ξ on ω wherever no confusion can arise, i.e., write ξ instead of $\xi(\omega)$.

Definition 2 Let ξ and $\tilde{\xi}$ be two Markov processes defined on probability spaces Ω and $\tilde{\Omega}$, respectively, and P and $\tilde{P}$ corresponding measures on Ω and $\tilde{\Omega}$. We define a transportation distance for the problem (1) as the following optimization problem taking the infimum over all probability measures π defined on $\mathcal{F} \otimes \tilde{\mathcal{F}}$

$$D_L(\xi, \tilde{\xi}) = \begin{cases} \inf_{\pi} \int\limits_{\Omega \times \tilde{\Omega}} d\left(\xi(\omega), \tilde{\xi}(\tilde{\omega})\right) \pi(d\omega, d\tilde{\omega}) \\ \text{s.t. } \pi_t^{\omega_{t-1}, \tilde{\omega}_{t-1}}(A \times \tilde{\Omega}_t) = P_t^{\omega_{t-1}}(A), \\ \quad\;\; \pi_t^{\omega_{t-1}, \tilde{\omega}_{t-1}}(\Omega_t \times \tilde{A}) = \tilde{P}_t^{\tilde{\omega}_{t-1}}(\tilde{A}), \end{cases} \tag{11}$$

where the constraints hold for all $t \in \mathbb{T} \backslash \{0\}$, for all $(\omega_{t-1}, \tilde{\omega}_{t-1}) \in \Omega_{t-1} \times \tilde{\Omega}_{t-1}$, as well as all $A \times \tilde{A} \in \sigma_t \otimes \tilde{\sigma}_t$ with $\omega = (\omega_1, \ldots, \omega_T)$,

$$d(\xi, \tilde{\xi}) = \sum_{t=0}^{T} \min \left\{ d_t(\xi_t, \tilde{\xi}_t), d_t(\tilde{\xi}_t, \xi_t) \right\}, \tag{12}$$

and

$$d_t(\xi_t, \tilde{\xi}_t) = \gamma_t \|c_t(\xi_t)\|_\infty \|b_{1,t}(\xi_t) - b_{1,t}(\tilde{\xi}_t)\|_1 + \phi_t(\tilde{\xi}_t) \|c_t\,(\xi_t) - c_t(\tilde{\xi}_t)\|_1. \tag{13}$$

Note that the objective function in (11) is defined in terms of the unconditional transport plan π between the joint distributions P and $\tilde{P}$ while the constraints rely on the corresponding disintegration in the form of Markov kernels $\pi_t^{\omega_{t-1}, \tilde{\omega}_{t-1}}$, which are guaranteed to exist [30] and relate to π via

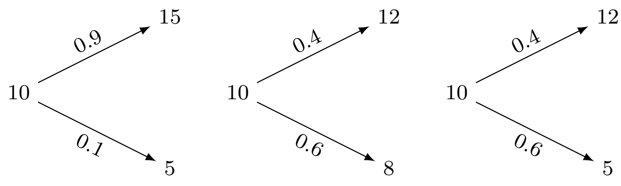
$$\pi(A \times \tilde{A}) = \int\limits_{\Omega_1 \times \tilde{\Omega}_1} \ldots \int\limits_{\Omega_T \times \tilde{\Omega}_T} \mathbb{1}_{A \times \tilde{A}}(\omega, \tilde{\omega})\, \pi_T^{\omega_{T-1}, \tilde{\omega}_{T-1}}(d\omega_T, d\tilde{\omega}_T) \ldots \pi_1(d\omega_1, d\tilde{\omega}_1).$$

Remark 4 Analogously to the Remark 2 and [27, 28] the infimum in the above definition is attained due to weak-compactness of the set of transportation plans.

Next we prove some properties of the lattice distance. We will start by showing that D_L is a *semi-metric*, i.e., that it is non-negative and symmetric. Example 1 demonstrates that it actually does not fulfill the triangle inequality.

Proposition 1 *D_L is a semi-metric.*

Proof From the non-negativity of the norms and the constants ϕ_t and γ_t, we obtain that $D_L \geq 0$. Clearly, $d(\xi, \tilde{\xi}) = d(\tilde{\xi}, \xi)$. If π^* is the optimal transportation plan for $D_L(\xi, \tilde{\xi})$, then $\tilde{\pi}^*(\tilde{\omega}, \omega) = \pi^*(\omega, \tilde{\omega})$ is the optimal transportation plan for $D_L(\tilde{\xi}, \xi)$. Therefore we have $D_L(\xi, \tilde{\xi}) = D_L(\tilde{\xi}, \xi)$. □



(a) Process ξ (b) Process $\hat{\xi}$ (c) Process $\tilde{\xi}$

Fig. 2 The three processes used in Example 1 to show that the trianlge inequality of D_L does not hold.

Example 1 In the following example we demonstrate that the triangle inequality does not hold in general for D_L. To that end, consider a simple two-stage problem with the objective function in period t defined by

$$c(\xi_t)^\top x_t = (\xi_t, 0)x_t$$

where ξ_t is a one-dimensional random variable. The constraints in the form of (2) are described by

$$A_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad b_1(\xi_t) = \begin{pmatrix} 0 \\ \xi_t \end{pmatrix}, \quad C_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0, 1 \end{pmatrix}, \quad b_2 = 20.$$

Considering the three random processes presented in Figure 2 and using definition (11), we obtain the following values of the lattice distance between every pair of processes

$$D_L(\xi, \hat{\xi}) = 147.8, \quad D_L(\hat{\xi}, \tilde{\xi}) = 54, \quad D_L(\xi, \tilde{\xi}) = 202.8.$$

We refer to Section 5 for a detailed description on how to calculate the distances. Hence, we have

$$D_L(\xi, \hat{\xi}) + D_L(\hat{\xi}, \tilde{\xi}) = 147.8 + 54 = 201.8 < 202.8 = D_L(\xi, \tilde{\xi})$$

confirming that the triangle inequality does not hold.

Next we show that there is always at least one feasible transport plan between any two Markov processes, i.e., there are no processes with infinite distance.

Proposition 2 *The defining optimization problem of D_L is always feasible. In particular, the product measure $\pi := P \otimes \tilde{P}$ is always part of the feasible set.*

Proof Let $A \in \sigma_{t+1}$ and $B \in \tilde{\sigma}_{t+1}$ for given t and $C \in \mathcal{F}$ and $D \in \tilde{\mathcal{F}}$, then we have

$$\int\limits_{C \times D} P_{t+1}^{\omega_t}(A) \cdot \tilde{P}_{t+1}^{\tilde{\omega}_t}(B) \pi(d\omega, d\tilde{\omega}) = \int\limits_{C} P_{t+1}^{\omega_t}(A) P(d\omega) \cdot \int\limits_{D} \tilde{P}_{t+1}^{\tilde{\omega}_t}(A) \tilde{P}(d\tilde{\omega})$$

$$
\begin{aligned}
&= P\left(C \cap A^{\Omega}_{t+1}\right) \cdot \tilde{P}\left(D \cap B^{\tilde{\Omega}}_{t+1}\right) \\
&= \pi\left(\left(C \cap A^{\Omega}_{t+1}\right) \times \left(D \cap B^{\tilde{\Omega}}_{t+1}\right)\right) \\
&= \pi\left((C \times D) \cap \left(A^{\Omega}_{t+1} \times B^{\tilde{\Omega}}_{t+1}\right)\right) \\
&= \int\limits_{C \times D} \pi^{\omega_t,\tilde{\omega}_t}_{t+1}(A \times B)\pi(d\omega, d\tilde{\omega})
\end{aligned}
$$

where the first equality follows from the properties of the product measure. Since the sets A, B, C, and D were arbitrary and $P^{\omega_t}_{t+1}(A) \cdot \tilde{P}^{\tilde{\omega}_t}_{t+1}(B)$ as well as $\pi^{\omega_t,\tilde{\omega}_t}_{t+1}(A \times B)$ are $\sigma_t \otimes \tilde{\sigma}_t$ measurable, it follows that they coincide π-almost everywhere, i.e.,

$$P^{\omega_t}_{t+1}(A) \cdot \tilde{P}^{\tilde{\omega}_t}_{t+1}(B) = \pi^{\omega_t,\tilde{\omega}_t}_{t+1}(A \times B).$$

For the particular choices $A = \Omega_{t+1}$ or $B = \tilde{\Omega}_{t+1}$, we get the conditions in problem (11). □

Similar to the classical duality result for the Wasserstein distance as well as the derivations in [27], we can give a dual characterization of D_L. To this end, we denote the set of $\mathcal{F} \otimes \tilde{\mathcal{F}}$-measurable real valued functions as $\mathcal{L}^0(\Omega \times \tilde{\Omega}, \mathcal{F} \otimes \tilde{\mathcal{F}})$ and define projections

$$\mathrm{Pr}_{t-1} : \mathcal{L}^0(\Omega \times \tilde{\Omega}, \mathcal{F} \otimes \tilde{\mathcal{F}}) \to \mathcal{L}^0(\Omega \times \tilde{\Omega}, \sigma^{\Omega}_{t-1} \otimes \tilde{\mathcal{F}})$$

for functions of the type $\mathbb{1}_{A \times \tilde{A}}$ with $A \in \mathcal{F}$ and $\tilde{A} \in \tilde{\mathcal{F}}$ as

$$\mathrm{Pr}_{t-1}(\mathbb{1}_{A \times \tilde{A}}) = \mathbb{E}_P(\mathbb{1}_A | \sigma^{\Omega}_{t-1}) \circ i \cdot \mathbb{1}_{\tilde{A}} \circ \tilde{i}$$

where $i : \Omega \times \tilde{\Omega} \to \Omega$, $\tilde{i} : \Omega \times \tilde{\Omega} \to \tilde{\Omega}$ are natural projections. Since the functions $\mathbb{1}_{A \times \tilde{A}}$ form a basis of $\mathcal{L}^0(\Omega \times \tilde{\Omega}, \mathcal{F} \otimes \tilde{\mathcal{F}})$, we can extend Pr_{t-1} to a well defined operator on the whole space. We define the projection $\widetilde{\mathrm{Pr}}_{t-1}$ analogously.

Proposition 3 *The dual representation of* (11) *can be written as*

$$
\begin{aligned}
\sup_{M,\mu_t,\tilde{\mu}_t} \quad & M \\
\text{s.t.} \quad & M + \sum_{s=1}^{T} \mu_s + \sum_{s=1}^{T} \mu_s \leq d(\xi, \tilde{\xi}) \\
& \mu_t \in \mathcal{L}^0(\Omega \times \tilde{\Omega}, \sigma^{\Omega}_{t-1,t} \otimes \tilde{\sigma}^{\tilde{\Omega}}_{t-1}), \ \forall t \in \mathbb{T}\backslash\{0\} \\
& \tilde{\mu}_t \in \mathcal{L}^0(\Omega \times \tilde{\Omega}, \sigma^{\Omega}_{t-1} \otimes \tilde{\sigma}^{\tilde{\Omega}}_{t-1,t}), \ \forall t \in \mathbb{T}\backslash\{0\} \\
& \mathrm{Pr}_{t-1}(\mu_t) = 0, \ \widetilde{\mathrm{Pr}}_{t-1}(\tilde{\mu}_t) = 0
\end{aligned}
$$

The proof of the above proposition can be found in Appendix A.

4 Bounding Linear Markov Decision Problems

We start by showing the following theorem demonstrating that distances between any Markov processes can be approximated to an arbitrary precision by distances where one of the processes is replaced by a discrete approximation. This result is of interest by itself and will also be required in the proof of Theorem 2.

Theorem 1 *Let $\theta_t(\xi_t, \tilde{\xi}_t) = \min\{d_t(\xi_t, \tilde{\xi}_t), d_t(\tilde{\xi}_t, \xi_t)\}$ and π be transportation plan that minimizes $D_L(\xi, \tilde{\xi})$ for two given processes ξ and $\tilde{\xi}$. If for all $1 \leq t \leq T$, $\theta_t(\xi_t, \tilde{\xi}_t) \in \mathcal{L}^p(\Omega_t \times \tilde{\Omega}_t, \pi_t)$ for some $p > 1$ and there is a $x_{0t} \in \mathbb{R}^{n_t}$ such that*

$$\int_{\Omega \times \tilde{\Omega}} \theta_t(\xi_t, x_{0t})\, \nu(d\omega, d\tilde{\omega}) < \infty$$

for all feasible transportation plans ν, then there is a sequence of discrete approximations $\tilde{\xi}^k$ such that $D_L(\xi, \tilde{\xi}^k) \xrightarrow{k \to \infty} D_L(\xi, \tilde{\xi})$.

Corollary 1 *Every Markov process ξ can be approximated arbitrarily well in terms of D_L by a discrete process.*

Proof Use ξ instead of $\tilde{\xi}$ in Theorem 1. □

Since θ_t is continuous, it is uniformly continuous on $B_k := B(0,k) \times B(0,k)$, where $B(0,k)$ is the ball around 0 in $\mathbb{R}^{n_t}$. To prove Theorem 1, we define the a discrete approximation in the following way: for each k define a discrete random variable $\tilde{\xi}^k_t : \tilde{\Omega}_t \to \mathbb{R}^{n_t}$ with atoms $\tilde{\xi}^k_{t,m}$ and $E^k_{t,m} = \{\tilde{\omega}_t \in \tilde{\Omega}_t : \tilde{\xi}^k_t(\tilde{\omega}_t) = \tilde{\xi}^k_{t,m}\}$ such that

$$|\theta_t\left(\xi_t(\omega_t), \tilde{\xi}_t(\tilde{\omega}_t)\right) - \theta_t\left(\xi_t(\omega_t), \tilde{\xi}^k_t(\tilde{\omega}_t)\right)| \leq k^{-1}, \quad \forall \omega_t\ \forall \tilde{\omega}_t : \tilde{\xi}_t(\tilde{\omega}_t) \in B(0,k)$$

and $\tilde{\xi}^k_t(\tilde{\omega}_t) = x_{0t}$ for all $\tilde{\omega}_t$ such that $\tilde{\xi}_t(\tilde{\omega}_t) \notin B(0,k)$. Furthermore, define corresponding Markov kernels as

$$\tilde{P}^{\tilde{\xi}^k_{t-1,m}}_{t,k}(\tilde{\xi}^k_{t,j}) = \int_{E^k_{t-1,m}} \tilde{P}^{\tilde{\omega}_{t-1}}_t(E^k_{t,j})\, \tilde{P}_{t-1}(d\tilde{\omega}_{t-1})$$

and the functions

$$f^t_k(\nu) = \int_{\Omega_t \times \tilde{\Omega}_t} \theta_t(\xi_t, \tilde{\xi}^k_t)\, \nu_t(d\omega_t, d\tilde{\omega}_t), \quad f^t_0(\nu) = \int_{\Omega_t \times \tilde{\Omega}_t} \theta_t(\xi_t, \tilde{\xi}_t)\, \nu_t(d\omega_t, d\tilde{\omega}_t).$$

for $\nu_t \in \mathcal{L}^q(\Omega_t \times \tilde{\Omega}_t, \pi_t)$ with $q^{-1} + p^{-1} = 1$ the unconditional distributions of the transportation plan ν in stage t.

Lemma 4 *f^t_k epi-converges to f^t_0 for all t, i.e., $f^t_k \xrightarrow{epi} f^t_0$.*

Proof Define

$$f^t_{kn}(\nu) = \int\limits_{\Omega_t \times \tilde{\Omega}_t} \theta_t(\xi_t, \tilde{\xi}^k_t) \mathbb{1}_{B_n}(\xi_t, \tilde{\xi}_t)\ \nu_t(d\omega_t, d\tilde{\omega}_t)$$

$$f^t_{0n}(\nu) = \int\limits_{\Omega_t \times \tilde{\Omega}_t} \theta_t(\xi_t, \tilde{\xi}_t) \mathbb{1}_{B_n}(\xi_t, \tilde{\xi}_t)\ \nu_t(d\omega_t, d\tilde{\omega}_t).$$

Fix $\epsilon > 0$. By assumption, there is a compact set $K \subset \mathbb{R}^{n_t} \times \mathbb{R}^{n_t}$ such that

$$\int\limits_{\Omega_t \times \tilde{\Omega}_t} \theta_t(\xi_t, x_{0t}) \mathbb{1}_{K^c}(\xi_t, \tilde{\xi}_t)\ \nu_t(d\omega_t, d\tilde{\omega}_t) < \epsilon,$$

$$\int\limits_{\Omega_t \times \tilde{\Omega}_t} \theta_t(\xi_t, \tilde{\xi}_t) \mathbb{1}_{K^c}(\xi_t, \tilde{\xi}_t)\ \nu_t(d\omega_t, d\tilde{\omega}_t) < \epsilon.$$

Now choose $k \in \mathbb{N}$ such that $K \subseteq B_k$ and $k > \epsilon^{-1}$ and note that

$$\begin{aligned} |f^t_{kn}(\nu) - f^t_{0n}(\nu)| \leq & \int\limits_{\Omega_t \times \tilde{\Omega}_t} |\theta_t(\xi_t, \tilde{\xi}^k_t) - \theta_t(\xi_t, \tilde{\xi}_t)| \mathbb{1}_{B_k}(\xi_t, \tilde{\xi}_t)\ \nu_t(d\omega_t, d\tilde{\omega}_t) \\ & + \int\limits_{\Omega_t \times \tilde{\Omega}_t} \theta_t(\xi_t, \tilde{\xi}_t) \mathbb{1}_{B_n \setminus B_k}(\xi_t, \tilde{\xi}_t)\ \nu_t(d\omega_t, d\tilde{\omega}_t) \\ & + \int\limits_{\Omega_t \times \tilde{\Omega}_t} \theta_t(\xi_t, x_{0t}) \mathbb{1}_{B_n \setminus B_k}(\xi_t, \tilde{\xi}_t)\ \nu_t(d\omega_t, d\tilde{\omega}_t) \leq 3\epsilon, \end{aligned}$$

i.e., $f^t_{kn} \to f^t_{0n}$ uniformly for all n. Note further that

$$f^t_0 = \lim_n f^t_{0n} = \lim_n \lim_k f^t_{kn} = \lim_k \lim_n f^t_{kn} = \lim_k f^t_k$$

where the two limits can be exchanged because of the uniform convergence shown above and the first equality follows by the monotone convergence theorem.

$\mathcal{L}^p(\Omega_t \times \tilde{\Omega}_t, \pi_t)$ is reflexive and therefore the weak topology is barrelled (see [22], Theorem 23.22). Since $f^t_k \to f^t_0$ weakly, the set $\{f^t_k, f^t_0\}$ is weakly bounded and therefore weakly equi-continuous by the uniform boundedness principle (see [4], Theorem III.2.1). Since $\{f^t_{kn} : n \in \mathbb{N}_0\}$ is equi-continuous, it is equi–lower semi-continuous and $f^t_k \xrightarrow{epi} f^t_0$ (see [6], Theorem 2.18). □

Proof (Theorem 1) Because of the epi-convergence proved in Lemma 4, we have

$$D_L(\xi, \tilde{\xi}^k) = \min_{\nu \in \Gamma} \sum_{t=0}^{T} f^t_k \to \min_{\nu \in \Gamma} \sum_{t=0}^{T} f^t_0 = D_L(\xi, \tilde{\xi}).$$

Note that the feasible set Γ can be w.l.o.g. assumed to be the feasible set of $D_L(\xi, \tilde{\xi})$, since for every feasible transportation plan for $D_L(\xi, \tilde{\xi}^k)$ there exists a plan that is feasible for $D_L(\xi, \tilde{\xi})$ yielding the same objective. □

Next we prove the main result establishing that the optimal value of the stochastic optimization problem associated to D_L is Lipschitz with respect to D_L. We first note the following useful lemma.

Lemma 5 *For a measurable function $f : \Omega_t \to \mathbb{R}$ and measures P_t, $\tilde{P}_t$, π_t that fulfill the conditions in* (11)*, we have*

$$\mathbb{E}_{\pi_t}(f \circ i) = \mathbb{E}_{P_t}(f).$$

Proof The result clearly holds for functions $f = \mathbb{1}_A$ with $A \in \Omega_t$ and therefore, by the usual argument, also for general measurable functions. □

Theorem 2 *Let ξ and $\tilde{\xi}$ be Markov decision processes for which the conditions of Theorem 1 hold, then*

$$|V_0(S_0, \xi_0) - \tilde{V}_0(S_0, \tilde{\xi}_0)| \leq D_L(\xi, \tilde{\xi}).$$

Proof We first assume that $\tilde{\xi}$ is finitely supported. Defining

$$\begin{aligned}
\delta_t^1(\xi_t, \tilde{\xi}_t) &= \gamma(A_{1,t}, A_{2,t})||c_t(\xi_t)||_\infty ||b_{1,t}(\xi_t) - b_{1,t}(\tilde{\xi}_t)||_1 \\
\delta_t^2(\xi_t, \tilde{\xi}_t) &= \phi(A_{1,t}, b_{1,t}(\tilde{\xi}_t) + C_{1,t}^{+} b_{2,t}, A_{2,t}, b_{2,t+1})||c_t(\xi_t) - c_t(\tilde{\xi}_t)||_1
\end{aligned}$$

as well as $\delta_t(\xi_t, \tilde{\xi}_t) = \delta_t^1(\xi_t, \tilde{\xi}_t) + \delta_t^2(\xi_t, \tilde{\xi}_t)$, we note that

$$\begin{aligned}
V_T(S_T, \xi_T) &= \max \left\{ c_T(\xi_T)^\top x_T : (x_T, S_{T+1}) \in \mathcal{X}_T(S_T, \xi_T) \right\} \\
&\geq \max \left\{ c_T(\xi_T)^\top x_T : (x_T, S_{T+1}) \in \mathcal{X}_T(S_T, \tilde{\xi}_T) \right\} - \delta_T^1(\xi_T, \tilde{\xi}_T) \\
&\geq \max \left\{ c_T(\tilde{\xi}_T)^\top x_T : (x_T, S_{T+1}) \in \mathcal{X}_T(S_T, \tilde{\xi}_T) \right\} - \delta_T(\xi_T, \tilde{\xi}_T) \\
&= \tilde{V}_T(S_T, \tilde{\xi}_T) - \delta_T(\xi_T, \tilde{\xi}_T),
\end{aligned} \tag{14}$$

where the two inequalities follow from Lemma 2, Lemma 3 and Remark 3. Exchanging the order of steps in which Lemma 2 and Lemma 3 are applied yields

$$\tilde{V}_T(S_T, \tilde{\xi}_T) - \delta_T(\tilde{\xi}_T, \xi_T) \leq V_T(S_T, \xi_T)$$

and exchanging the roles of V_T and $\tilde{V}_T$ finally results in

$$|\tilde{V}_T(S_T, \tilde{\xi}_T) - V_T(S_T, \xi_T)| \leq \min \left\{ \delta_T(\xi_T, \tilde{\xi}_T), \delta_T(\tilde{\xi}_T, \xi_T) \right\} =: \Delta_T(\xi_T, \tilde{\xi}_T).$$

Proceeding to the next stage, we have

$$\begin{aligned}
V_{T-1}(S_{T-1}, \xi_{T-1}) &= \begin{cases} \max\ c_{T-1}(\xi_{T-1})^\top x_{T-1} + \mathbb{E}_{P_T}\left[V_T(S_T, \xi_T)|\xi_{T-1}\right] \\ \text{s.t.}\ \ (x_{T-1}, S_T) \in \mathcal{X}_{T-1}(S_{T-1}, \xi_{T-1}) \end{cases} \\
&= \begin{cases} \max\ c_{T-1}(\xi_{T-1})^\top x_{T-1} + \mathbb{E}_{\pi_T}\left[V_T(S_T, \xi_T) \circ i|(\xi_{T-1}, \tilde{\xi}_{T-1})\right] \\ \text{s.t.}\ \ (x_{T-1}, S_T) \in \mathcal{X}_{T-1}(S_{T-1}, \xi_{T-1}) \end{cases} \\
&\geq \begin{cases} \max\ c_{T-1}(\xi_{T-1})^\top x_{T-1} + \mathbb{E}_{\pi_T}\left[\tilde{V}_T(S_T, \tilde{\xi}_T) \circ \tilde{i} - \Delta_T|(\xi_{T-1}, \tilde{\xi}_{T-1})\right] \\ \text{s.t.}\ \ (x_{T-1}, S_T) \in \mathcal{X}_{T-1}(S_{T-1}, \xi_{T-1}) \end{cases}
\end{aligned}$$

$$= \begin{cases} \max\ c_{T-1}(\xi_{T-1})^\top x_{T-1} + \mathbb{E}_{\tilde{P}_T}\left[\tilde{V}_T(S_T,\tilde{\xi}_T)|\tilde{\xi}_{T-1}\right] - \mathbb{E}_{\pi_T}\left[\Delta_T|(\xi_{T-1},\tilde{\xi}_{T-1})\right] \\ \text{s.t.}\ \ (x_{T-1},S_T) \in \mathcal{X}_{T-1}(S_{T-1},\xi_{T-1}) \end{cases}$$

$$= \begin{cases} \max\ c_{T-1}(\xi_{T-1})^\top x_{T-1} + \tilde{\gamma} - \mathbb{E}_{\pi_T}\left[\Delta_T|(\xi_{T-1},\tilde{\xi}_{T-1})\right] \\ \text{s.t.}\ \ (x_{T-1},S_T) \in \mathcal{X}_{T-1}(S_{T-1},\xi_{T-1}) \\ \qquad \tilde{\gamma} \le \tilde{\beta}_0 + \tilde{\beta}_1 S_T \end{cases}$$

$$\ge \begin{cases} \max\ c_{T-1}(\xi_{T-1})^\top x_{T-1} + \tilde{\gamma} \\ \text{s.t.}\ \ (x_{T-1},S_T) \in \mathcal{X}_{T-1}(S_{T-1},\tilde{\xi}_{T-1}) \\ \qquad \tilde{\gamma} \le \tilde{\beta}_0 + \tilde{\beta}_1 S_T \end{cases}$$

$$- \mathbb{E}_{\pi_T}\left[\Delta_T|(\xi_{T-1},\tilde{\xi}_{T-1})\right] - \delta^1_{T-1}(\xi_{T-1},\tilde{\xi}_{T-1})$$

$$\ge \begin{cases} \max\ c_{T-1}(\tilde{\xi}_{T-1})^\top x_{T-1} + \tilde{\gamma} \\ \text{s.t.}\ \ (x_{T-1},S_T) \in \mathcal{X}_{T-1}(S_{T-1},\tilde{\xi}_{T-1}) \\ \qquad \tilde{\gamma} \le \tilde{\beta}_0 + \tilde{\beta}_1 S_T \end{cases}$$

$$- \mathbb{E}_{\pi_T}\left[\Delta_T|(\xi_{T-1},\tilde{\xi}_{T-1})\right] - \delta_{T-1}(\xi_{T-1},\tilde{\xi}_{T-1})$$

$$= \tilde{V}_{T-1}(S_{T-1},\tilde{\xi}_{T-1}) - \mathbb{E}_{\pi_T}\left[\Delta_T|(\xi_{T-1},\tilde{\xi}_{T-1})\right] - \delta_{T-1}(\xi_{T-1},\tilde{\xi}_{T-1})$$

where the second equality follows by Lemma 5, the first inequality from (14), the following equality again from Lemma 5 and the subsequent inequalities follows from Lemma 2 and Lemma 3. As in the derivation of (14), we can exchange the order in which Lemma 2 and Lemma 3 are applied to get the above inequality with $\delta_{T-1}(\xi_{T-1},\tilde{\xi}_{T-1})$ replaced by $\delta_{T-1}(\tilde{\xi}_{T-1},\xi_{T-1})$. Exchanging the roles of V_t and $\tilde{V}_t$ we obtain

$$|\tilde{V}_{T-1}(S_{T-1},\tilde{\xi}_{T-1}) - V_{T-1}(S_{T-1},\xi_{T-1})| \le \mathbb{E}_{\pi_T}\left[\Delta_T(\xi_T,\tilde{\xi}_T)|(\xi_{T-1},\tilde{\xi}_{T-1})\right] + \Delta_{T-1}(\xi_{T-1},\tilde{\xi}_{T-1}).$$

Proceeding by backward induction, we arrive at

$$|\tilde{V}_0(S_0,\tilde{\xi}_0) - V_0(S_0,\xi_0)| \le D_L(\xi,\tilde{\xi}).$$

For the general case construct discrete approximations $\tilde{P}_k$ to $\tilde{P}$ and note that

$$\begin{aligned} |\tilde{V}_0(S_0,\tilde{\xi}_0) - V_0(S_0,\xi_0)| &\le |\tilde{V}^k_0(S_0,\tilde{\xi}^k_0) - V_0(S_0,\xi_0)| + |\tilde{V}^k_0(S_0,\tilde{\xi}^k_0) - \tilde{V}_0(S_0,\tilde{\xi}_0)| \\ &\le D_L(\xi,\tilde{\xi}^k) + D_L(\tilde{\xi}^k,\tilde{\xi}) \to D_L(\xi,\tilde{\xi}) \end{aligned}$$

by Theorem 1 and the special case proven above. □

Remark 5 Every MDP can equivalently be viewed as a stochastic optimization problem. Assuming no randomness in the constraints, [27] provide stability results that are similar to our results for this general problem class. Hence, a comparison is of interest.

In [27] it proven that for a general objective function $h : \mathbb{X} \times \Omega \to \mathbb{R}$

$$\left|\min_{x\in\mathbb{X}} \mathbb{E}(h(x,\xi)) - \min_{x\in\mathbb{X}} \mathbb{E}(h(x,\tilde{\xi}))\right| \le L\ D_T(\xi,\tilde{\xi})$$

assuming that there is an L such that

$$|h(x,\xi) - h(x,\tilde{\xi})| \leq L\ ||\xi - \tilde{\xi}||_1, \quad \forall x \in \mathbb{X}, \quad \forall (\omega, \tilde{\omega}) \in \Omega \times \tilde{\Omega}.$$

Defining $\mathcal{G}_t = \sigma(\xi_1, \ldots, \xi_t)$, $\tilde{\mathcal{G}}_t = \sigma(\tilde{\xi}_1, \ldots, \tilde{\xi}_t)$ as the σ-algebras generated by the history of the processes, a distance D_T for arbitrary stochastic processes can be defined as

$$D_T(\xi, \tilde{\xi}) = \begin{cases} \inf\limits_{\pi} \displaystyle\int_{\Omega \times \tilde{\Omega}} ||\xi - \tilde{\xi}||_1\ \pi(d\omega, d\tilde{\omega}) \\ \text{s.t. } \pi(A \times \tilde{\Omega} | \mathcal{G}_t \otimes \tilde{\mathcal{G}}_t) = P(A|\mathcal{G}_t), \quad \forall A \in \mathcal{G}_T \\ \quad\ \ \pi(\Omega \times \tilde{A} | \mathcal{G}_t \otimes \tilde{\mathcal{G}}_t) = \tilde{P}(\tilde{A}|\tilde{\mathcal{G}}_t), \quad \forall \tilde{A} \in \tilde{\mathcal{G}}_T. \end{cases} \tag{15}$$

In our case $h(x,\xi) = \sum\limits_{t=0}^{T} c_t(\xi_t)^\top x_t$ and L can be calculated as $L = \max\limits_t L_{c_t} \phi_t$ assuming that the functions c_t are Lipschitz with constants L_{c_t}. Note that ϕ_t is deterministic in the case of deterministic feasible set.

It is easy to see that for two Markov processes the permissible transportation plans π for D_T and for D_L are equivalent. Assume that π^* is optimal transportation plan for D_T, then we have

$$\begin{aligned} D_L(\xi, \tilde{\xi}) &\leq \int\limits_{\Omega \times \tilde{\Omega}} \sum_{t=0}^{T} \phi_t \|c_t\,(\xi_t) - c_t(\tilde{\xi}_t)\|_1\ \pi^*(d\omega, d\tilde{\omega}) \\ &\leq \int\limits_{\Omega \times \tilde{\Omega}} \sum_{t=0}^{T} \phi_t L_{c_t} \|\xi_t - \tilde{\xi}_t\|_1\ \pi^*(d\omega, d\tilde{\omega}) \\ &\leq L \int\limits_{\Omega \times \tilde{\Omega}} \|\xi - \tilde{\xi}\|_1\ \pi^*(d\omega, d\tilde{\omega}) = L\ D_T(\xi, \tilde{\xi}). \end{aligned}$$

Hence, we have shown that our bound is tighter than D_T for problems where both bounds are applicable, i.e., linear MDPs with deterministic feasible set $\mathbb{X}$.

5 Implementation for Finite Lattices

In this section, we focus on the computation of D_L for two finitely supported Markov processes. We show that, similar to the case of the classical Wasserstein distance and [27], the distance can be computed by solving a linear optimization problem to find the optimal transport plan π.

To that end, we represent the discrete Markov processes ξ and $\tilde{\xi}$ by lattices. At every stage $t \in \mathbb{T}$ we define the probability spaces

$$\Omega_t = \{i : 1 \leq i \leq N_t\},\ \tilde{\Omega}_t = \left\{\tilde{i} : 1 \leq \tilde{i} \leq M_t\right\}$$

where N_t and M_t are the number of atoms of P_t and $\tilde{P}_t$ respectively. The conditional transition from a given state i at time $(t-1)$ to a state j at time t is described by a conditional probability $P_t^i(j)$ and $\tilde{P}_t^i(j)$, respectively.

To compute the lattice distance for discrete Markov processes with image measures P and $\tilde{P}$, we note that

$$\begin{aligned} & \pi_t^{i,r}(\{j\} \times \tilde{\Omega}_t) = P_t^i(j) \\ \Leftrightarrow & \sum_{\omega:\omega_{t-1}=i,\omega_t=j} \sum_{\tilde{\omega}:\tilde{\omega}_{t-1}=r} \pi(\omega,\tilde{\omega}) = P_t^i(j) \sum_{\omega:\omega_{t-1}=i} \sum_{\tilde{\omega}:\tilde{\omega}_{t-1}=r} \pi(\omega,\tilde{\omega}) \end{aligned}$$

and formulate (11) as a linear program over all probability measures π on $\Omega \times \tilde{\Omega}$ as

$$D_L(\xi,\tilde{\xi}) = \begin{cases} \min\limits_{\pi} \sum\limits_{\omega,\tilde{\omega}} ||\xi(\omega) - \tilde{\xi}(\tilde{\omega})||_1 \; \pi(\omega,\tilde{\omega}) \\ \text{s.t.} \;\; \pi_t^{i,r}(\{j\} \times \tilde{\Omega}_t) = P_t^i(j) \\ \qquad \pi_t^{i,r}(\Omega_t \times \{s\}) = \tilde{P}_t^r(s) \end{cases} \tag{16}$$

where the constraints hold for all $1 \leq i \leq N_{t-1}, 1 \leq j \leq N_t, 1 \leq r \leq M_{t-1}$ and $1 \leq s \leq M_t$.

As a demonstration, we consider a multi-stage extension of the classical newsboy problem – the problem of a flowergirl selling flowers, facing a random demand and a random sales price with the possibility to store excess flowers for the next period. The problem has $(T+1)$ stages, with stage $t=0$ being the deterministic start state. In every stage t, we start with the inventory level S_t limited by the storage capacity $\bar{S}_t$. Due to the perishable nature of flowers, l% of the flowers cannot be sold on the next day. After the demand ξ_t^1 and the price ξ_t^2 become known in stage t, the flowergirl places an order x_t^1 for flowers to be delivered from a wholesaler for a price p on the next day. The order in stage t has to be placed without knowing the random demand ξ_{t+1}^1. On the next day the flowers can be sold at a market price ξ_{t+1}^2 not known on day t. If the available quantity exceeds the demand, the flowergirl can add the excess to her inventory for sale in $(t+2)$. The flowergirl starts in period $t=0$ without any stock and no demand, i.e., $S_0 = 0$ and $\xi_0^1 = 0$.

The decisions in every stage consist of the number of flowers to order for the next stage x_t^1, the number of flowers to sell x_t^2, and the inventory level of the next day x_t^3. Furthermore, we define the state variable $S_{t+1} = x_t^3$. Note that, as described in Remark 1, the environmental state variable S_{t+1} is duplicated as an auxiliary variable so as to make the feasible set fit (2).

The storage equation consequently is

$$x_t^3 = (1-l) \cdot (S_t - x_t^2) + x_t^1, \quad \forall t = 0, \ldots, T.$$

The sales decisions are constrained by the random demand as well as the storage level, i.e.,

$$x_t^2 \leq \min(\xi_t^1, S_t), \quad \forall t = 0, \ldots, T, \quad a.s.$$

Furthermore, we impose the following constraints

$$x_t^3 = S_{t+1}, \quad 0 \leq x_t^3 \leq \bar{S}_{t+1}, \quad x_t^1,\ x_t^2,\ x_t^3 \geq 0, \quad \forall t = 0, \dots, T.$$

The flowergirl maximizes her expected profit, which is given by

$$\mathbb{E}\left(\sum_{t=0}^{T} \xi_t^2 x_t^2 - p x_t^1\right).$$

For our numerical example, we consider the three-stage version of the problem, i.e., the problem with $T = 2$, $l = 0.1$, $p = 5$ and the vector of storage capacities $\bar{S} = \left(\bar{S}_0, \bar{S}_1, \bar{S}_2, \bar{S}_3\right) = (0, 11, 9, 0)$. To rewrite the problem to the form in (1), we define $c(\xi_t) = (-p, \xi_t^2, 0)^\top$ and the vectors and matrices appearing in the constraints as

$$A_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ -1 & -(1-l) & 1 \\ 1 & 1-l & -1 \end{pmatrix}, \quad b_1(\xi_t) = \begin{pmatrix} \xi_t^1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad C_1 = \begin{pmatrix} 0 \\ 1 \\ 1-l \\ -(1-l) \end{pmatrix},$$
$$A_2 = \begin{pmatrix} 0 \ 0 \ 1 \end{pmatrix}, \quad b_{2,t} = \begin{pmatrix} \bar{S}_t \end{pmatrix}.$$

The flowergirl problem can then be formulated as follows

$$V_t(S_t, \xi_t) = \begin{cases} \max\limits_{x_t} \ c(\xi_t)^\top x_t + \mathbb{E}\left(V_{t+1}(S_{t+1}, \xi_{t+1}) | \xi_t\right) \\ \text{s.t.} \ \ A_1 x_t \leq b_1(\xi_t) + C_1 S_t \\ \qquad A_2 x_t = S_{t+1} \\ \qquad A_2 x_t \leq b_{2,t+1} \\ \qquad S_{t+1}, x_t \geq 0. \end{cases}$$

We consider the two Markov processes ξ and $\tilde{\xi}$ presented in Figure 3a and Figure 3b with transition probabilities

$$P_1 = \begin{pmatrix} 0.8 & 0.1 & 0.1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 0.15 & 0.10 & 0.10 & 0.65 \\ 0.60 & 0.10 & 0.15 & 0.15 \\ 0.50 & 0.10 & 0.30 & 0.10 \end{pmatrix},$$
$$\tilde{P}_1 = \begin{pmatrix} 0.1 & 0.9 \end{pmatrix}, \quad \tilde{P}_2 = \begin{pmatrix} 0.20 & 0.15 & 0.65 \\ 0.20 & 0.25 & 0.55 \end{pmatrix}.$$

To bound the difference in the solutions, we calculate $D_L(\xi, \tilde{\xi})$. We find that $\lambda_1 = (3.8, 0, 2)$ maximizes $||\cdot||_\infty$ over the extreme points of the set

$$P = \{(\lambda_1, \lambda_2, \lambda_3) : ||A_1^\top \lambda_1 + A_2^\top \lambda_2 - \lambda_3||_\infty \leq 2, \lambda_1, \lambda_3 \geq 0\},$$

resulting in $\gamma = 3.8$. The constant $\phi_t(\xi_t) = \phi\left(A_1, b_1(\xi_t) + C_1^+ b_{2,t}, A_2, b_{2,t+1}\right)$ can be found by maximizing $||x||_\infty$ over the extreme points x of the set

$$\hat{P}(\xi_t) = \{x : A_1 x \leq b_1(\xi_t) + C_1^+ b_{2,t}, A_2 x \leq b_{2,t+1}, A_2 x \geq 0, x \geq 0\}.$$

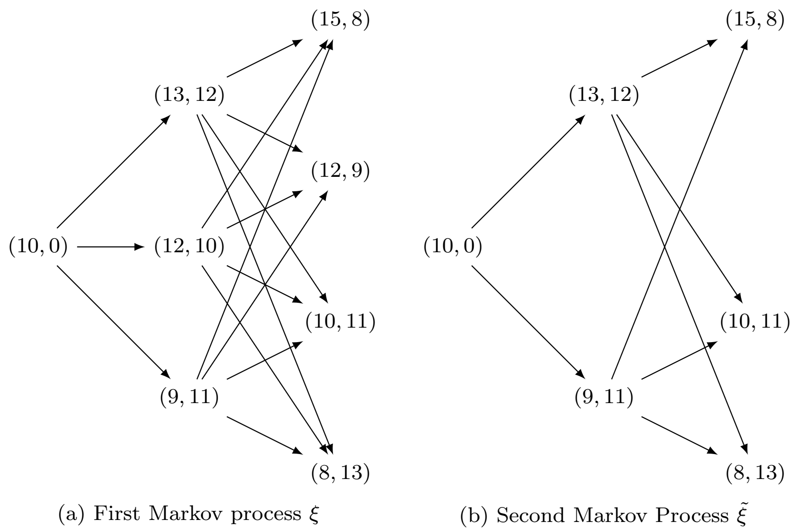


Fig. 3 Depiction of the two Markov processes used for the numerical calculation of the flowergirl example.

Having calculated the constants, we proceed by computing $d(\xi, \tilde{\xi})$ using (12) and (13). We can then determine the joint distribution π that minimizes the distance between processes by solving (16).

The resulting optimal transportation plan yields a distance of $D_L(\xi, \tilde{\xi}) = 75.99$. The optimal value of our problem for ξ is equal to 123.83 and for $\tilde{\xi}$ the optimal value equals 87.44 resulting in a difference of 36.39.

Lastly, we compare the performance of D_L to the performance of the nested distance defined in [27, 28]. For this it is necessary to simplify the problem so as to make the constraints independent of the randomness. To achieve this, we fix the demand at each stage. In particular, we assume that the demand is equal to 0, 11, and 9 in the stages $t = 0$, 1, and 2, respectively. For this simplified setup, we obtain $D_L(\xi, \tilde{\xi}) = 34.42$ and $D_T(\xi, \tilde{\xi}) = 38.28$ demonstrating that for our problem D_L provides tighter bounds than D_T (see also Remark 5).

6 Conclusions

MDPs strike a good balance between the complexity of the required discrete approximations of the randomness and the expressiveness of the corresponding problem class. In particular, since scenario lattices offer leaner discretization structures than scenario trees, the unfavorable computational properties of general stochastic optimization problems can be, in part, mitigated in MDPs as has been demonstrated in the literature.

In this paper, we define a family of problem dependent semi-distances for linear MDPs that can be used to bound objective values. We also show that every Markov process can be approximated to arbitrary precision in terms of the defined distances. Therefore the concepts in this paper can be used to find arbitrary precise discrete approximation of complicated MDPs, possibly with continuous state space.

Furthermore, we contribute to the literature on transportation distances by an approach that is capable of dealing with randomness in the constraints. This necessitates a different technique of proof, since the transport of solutions between problems becomes impossible in this framework. We therefore base our results on stability results for linear programs.

Hence, we contribute to the development of a theory of optimal scenario reduction for linear MDPs – a topic that received very little attention till now, although the corresponding problem class is quite popular.

In this paper we laid the foundations for a theory driven method to generate scenario lattices. Further research is required to find computationally efficient ways to do so and to evaluate the outcomes on real world problems.

References

1. V. Bally and P. Pagès. Error analysis of the quantization algorithm for obstacle problems. *Stochastic Processes & Their Applications*, 106(1):1–40, 2003.
2. V. Bally and P. Pagès. A quantization algorithm for solving discrete time multi-dimensional optimal stopping problems. *Bernoulli*, 9(6):1003–1049, 2003.
3. J. R. Birge and F. Louveaux. *Introduction to stochastic programming.* Springer Series in Operations Research and Financial Engineering, 2000.
4. N. Bourbaki, H. G. Eggleston, and S. Madan. *Topological Vector Spaces.* Éléments de mathématique. Springer-Verlag, 1987.
5. M. S. Casey and S. Sen. The scenario generation algorithm for multistage stochastic linear programming. *Mathematics of Operations Research*, 30:615–631, 2005.
6. S. Dolecki, G. Salinetti, and R. J. B. Wets. Convergence of functions: Equi-semicontinuity. *Transactions of the American Mathematical Society*, 276(1):409–429, 1983.
7. J. Dupacová, G. Consigli, and S. W. Wallace. Scenarios for multistage stochastic programs. *Annals of Operations Research*, 100:25–53, 2000.
8. J. Dupacová, N. Gröwe-Kuska, and W. Römisch. Scenario reduction in stochastic programming an approach using probability metrics. *Mathematical Programming, Series B*, 95(3):493–511, 2003.
9. N. C. P. Edirisinghe. Bound-based approximations in multistage stochastic programming: Nonanticipativity aggregation. *Annals of Operations Research*, 85:103–127, 1996.
10. K. Frauendorfer. Barycentric scenario trees in convex multistage stochastic programming. *Mathematical Programming*, 75:277–293, 1996.

11. G. A. Hanasusanto, D. Kuhn, and W. Wiesemann. A comment on "computational complexity of stochastic programming problems". *Mathematical Programming*, 159(1):557–569, 2016.
12. H. Heitsch and W. Römisch. Scenario reduction algorithms in stochastic programming. *Computational Optimization and Applications*, 24(2-3):187–206, 2003.
13. H. Heitsch and W. Römisch. Scenario tree modeling for multistage stochastic programs. *Mathematical Programming*, 118:371–406, 2009.
14. H. Heitsch and W. Römisch. Stability and scenario trees for multistage stochastic programs. *International series in Operations Research & Management Science*, 150:139–164, 2011.
15. H. Heitsch, W. Römisch, and C. Strugarek. Stability of multistage stochastic programs. *SIAM Journal on Optimization*, 17:511–525, 2006.
16. K. Høyland, M. Kaut, and S. W. Wallace. A heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 24(2):169–185, 2003.
17. K. Høyland and S. W. Wallace. Generating Scenario Decision Trees for Multistage Problems. *Management Science*, 47(2):295–307, 2001.
18. L. Kantorovich. On the translocation of masses. *C.R. Acad. Sci. URSS*, 1(37):199–201, 1942.
19. M. Kaut and S. W. Wallace. Evaluation of scenario generation methods for stochastic programming. *Pacific Journal of Optimization*, 3(2):257–271, 2007.
20. D. Kuhn. *Generalized Bounds for Convex Multistage Stochastic Programs.* Springer-Verlag Berlin Heidelberg, 2005.
21. N. Löhndorf, D. Wozabal, and S. Minner. Optimizing trading decisions for hydro storage systems using approximate dual dynamic programming. *Operations Research*, 61(4):810–823, 2013.
22. R. Meise, D. Vogt, and M. S. Ramanujan. *Introduction to Functional Analysis.* Clarendon Press, 1997.
23. T. Pennanen. Epi-convergent discretizations of multistage stochastic programs. *Mathematics of Operations Research*, 30(1):245–256, 2005.
24. T. Pennanen. Epi-convergent discretizations of multistage stochastic programs via integration quadratures. *Mathematical Programming*, 116(1-2):461–479, 2009.
25. M. V. F. Pereira and L. M. V. G. Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52(2):359–375, 1991.
26. G. Ch. Pflug. Scenario tree generation for multiperiod financial optimization by optimal discretization. *Mathematical Programming, Series B*, 89(2):251–271, 2001.
27. G. Ch. Pflug and A. Pichler. A distance for multistage stochastic optimization models. *SIAM Journal on Optimization*, 22(1):1–23, 2012.
28. G. Ch. Pflug and A. Pichler. *Multistage Stochastic Optimization.* Springer Series in Operations Research and Financial Engineering, 2014.

29. A. B. Philpott and Z. Guan. On the convergence of stochastic dual dynamic programming and related methods. *Operations Research Letters*, 36(4):450 – 455, 2008.
30. D. Pollard. *A User's Guide to Measure Theoretic Probability*. Cambridge Series in Statistical and Probabilistic Mathematics, 2002.
31. M. L. Puterman. *Markov decision processes*. John Wiley & Sons, 1994.
32. A. Shapiro. On complexity of multistage stochastic programs. *Operations Research Letters*, 34(1):1–8, 2006.
33. A. Shapiro. Computational complexity of stochastic programming: Monte carlo sampling approach. *Proceedings of the International Congress of Mathematicians*, pages 2979–2995, 2010.
34. A. Shapiro. Analysis of stochastic dual dynamic programming method. *European Journal of Operational Research*, 209(1):63 – 72, 2011.
35. A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming*. MPS-SIAM Series on Optimization 9, 2009.
36. C. Villani. Topics in optimal transportation. *American Mathematical Society*, vol. 58 of Graduate Studies in Mathematics, 2003.

A Proof of Proposition 3

To prove Proposition 3, we require the following lemma, providing an equivalent characterizations of the constraints in (11).

Lemma 6 *The measure π satisfies the conditions*

$$\pi_t^{\omega_{t-1},\tilde{\omega}_{t-1}}(A \times \tilde{\Omega}_t) = P_t^{\omega_{t-1}}(A) \tag{17}$$

for every $A \in \sigma_t$ and $(\omega_{t-1}, \tilde{\omega}_{t-1}) \in \Omega_{t-1} \times \tilde{\Omega}_{t-1}$ if and only if

$$\mathbb{E}_\pi(\lambda) = \mathbb{E}_\pi(\mathrm{Pr}_{t-1}(\lambda))$$

holds for all integrable and $\sigma^{\Omega}_{t-1,t} \otimes \tilde{\sigma}^{\tilde{\Omega}}_{t-1}$-measurable functions λ.

Proof Clearly it is enough to consider the basis of functions of the following type

$$\begin{aligned}\lambda &= \mathbb{1}_{(\Omega_0 \times \ldots \times A \times B \times \ldots \times \Omega_T) \times (\tilde{\Omega}_0 \times \ldots \times \tilde{A} \times \ldots \times \tilde{\Omega}_T))} \\ &= \mathbb{1}_{(\Omega_0 \times \ldots \times A \times \ldots \times \Omega_T) \times (\tilde{\Omega}_0 \times \ldots \times \tilde{A} \times \ldots \times \tilde{\Omega}_T))} \mathbb{1}_{(\Omega_0 \times \ldots \times B \times \ldots \times \Omega_T) \times \tilde{\Omega}}\end{aligned}$$

where $A \in \sigma_{t-1}$, $B \in \sigma_t$, and $\tilde{A} \in \tilde{\sigma}_{t-1}$. Because of (17), we have

$$\begin{aligned}\mathbb{E}_\pi(\lambda) &= \int\limits_{\Omega_0 \times \tilde{\Omega}_0} \cdots \int\limits_{\Omega_{t-1} \times \tilde{\Omega}_{t-1}} \mathbb{1}_{A \times \tilde{A}} \int\limits_{\Omega_t \times \tilde{\Omega}_t} \mathbb{1}_{B \times \tilde{\Omega}_t} d\pi_t^{\omega_{t-1},\tilde{\omega}_{t-1}} d\pi_{t-1}^{\omega_{t-2},\tilde{\omega}_{t-2}} \ldots d\pi_0 \\ &= \int\limits_{\Omega_0 \times \tilde{\Omega}_0} \cdots \int\limits_{\Omega_{t-1} \times \tilde{\Omega}_{t-1}} \mathbb{1}_{A \times \tilde{A}} \int\limits_{\Omega_t} \mathbb{1}_B dP_t^{\omega_{t-1}} d\pi_{t-1}^{\omega_{t-2},\tilde{\omega}_{t-2}} \ldots d\pi_0 \\ &= \int\limits_{\Omega_0 \times \tilde{\Omega}_0} \cdots \int\limits_{\Omega_{t-1} \times \tilde{\Omega}_{t-1}} \mathbb{1}_A \int\limits_{\Omega_t} \mathbb{1}_B dP_t^{\omega_{t-1}} \circ i \cdot \mathbb{1}_{\tilde{A}} \circ \tilde{i} d\pi_{t-1}^{\omega_{t-2},\tilde{\omega}_{t-2}} \ldots d\pi_0\end{aligned}$$

$$= \int_{\Omega_0 \times \tilde\Omega_0} \dots \int_{\Omega_{t-1} \times \tilde\Omega_{t-1}} \mathrm{Pr}_{t-1}(\lambda) d\pi_{t-1}^{\omega_{t-2},\tilde\omega_{t-2}} \dots d\pi_0 = \mathbb{E}_\pi(\mathrm{Pr}_{t-1}(\lambda))$$

To prove the converse implication we need to show

$$\pi_t^{\omega_{t-1},\tilde\omega_{t-1}}(A \times \tilde\Omega_t) = P_t^{\omega_{t-1}}(A)$$

for every $A \in \sigma_t$ and $(\omega_{t-1}, \tilde\omega_{t-1}) \in \Omega_{t-1} \times \tilde\Omega_{t-1}$. As both, the left hand side and the right hand side are $\sigma_{t-1}^{\Omega} \otimes \tilde\sigma_{t-1}^{\tilde\Omega}$-measurable densities it is sufficient to show that

$$\int_{C \times D} P_t(A|\sigma_{t-1}^{\Omega}) \circ i d\pi = \int_{C \times D} \pi_t(A \times \tilde\Omega_t | \sigma_{t-1}^{\Omega} \otimes \tilde\sigma_{t-1}^{\tilde\Omega}) d\pi \tag{18}$$

for all sets $C \in \sigma_{t-1}^{\Omega}$ and $D \in \tilde\sigma_{t-1}^{\tilde\Omega}$. To this end observe first that

$$\begin{aligned}
&\int_{C \times D} \pi_t(A \times \tilde\Omega_t | \sigma_{t-1}^{\Omega} \otimes \tilde\sigma_{t-1}^{\tilde\Omega}) d\pi = \mathbb{E}_\pi(\mathbb{1}_{C\times D} \pi(A_t^{\Omega} \times \tilde\Omega | \sigma_{t-1}^{\Omega} \otimes \tilde\sigma_{t-1}^{\tilde\Omega}))\\
&= \mathbb{E}_\pi(\mathbb{1}_{C\times D} \mathbb{E}_\pi(\mathbb{1}_{A_t^{\Omega} \times \tilde\Omega} | \sigma_{t-1}^{\Omega} \otimes \tilde\sigma_{t-1}^{\tilde\Omega})) = \mathbb{E}_\pi(\mathbb{E}_\pi(\mathbb{1}_{A_t^{\Omega} \times \tilde\Omega} \mathbb{1}_{C\times D} | \sigma_{t-1}^{\Omega} \otimes \tilde\sigma_{t-1}^{\tilde\Omega}))\\
&= \mathbb{E}_\pi(\mathbb{1}_{(A_t^{\Omega} \times \tilde\Omega)\cap(C\times D)}) = \mathbb{E}_\pi(\mathbb{1}_{(A_t^{\Omega}\cap C)\times D}) = \mathbb{E}_\pi(\mathbb{1}_{A_t^{\Omega}\cap C} \circ i \cdot \mathbb{1}_D \circ \tilde i)
\end{aligned}$$

secondly

$$\begin{aligned}
&\int_{C \times D} P_t(A|\sigma_{t-1}^{\Omega}) \circ i d\pi = \mathbb{E}_\pi(\mathbb{1}_{C\times D} P(A_t^{\Omega} | \sigma_{t-1}^{\Omega}) \circ i)\\
&= \mathbb{E}_\pi(\mathbb{1}_{C\times D} \mathbb{E}_P(\mathbb{1}_{A_t^{\Omega}} | \sigma_{t-1}^{\Omega}) \circ i) = \mathbb{E}_\pi(\mathbb{1}_C \circ i \cdot \mathbb{1}_D \circ \tilde i \cdot \mathbb{E}_P(\mathbb{1}_{A_t^{\Omega}} | \sigma_{t-1}^{\Omega}) \circ i)\\
&= \mathbb{E}_\pi(\mathbb{E}_P(\mathbb{1}_{A_t^{\Omega}} \mathbb{1}_C | \sigma_{t-1}^{\Omega}) \circ i \cdot \mathbb{1}_D \circ \tilde i) = \mathbb{E}_\pi(\mathbb{E}_P(\mathbb{1}_{A_t^{\Omega}\cap C} | \sigma_{t-1}^{\Omega}) \circ i \cdot \mathbb{1}_D \circ \tilde i)\\
&= \mathbb{E}_\pi(Pr_{t-1}(\mathbb{1}_{A_t^{\Omega}\cap C} \circ i \cdot \mathbb{1}_D \circ \tilde i))
\end{aligned}$$

and by the assumption (18) holds. The quantities coincide and we conclude that desired assertion holds true. □

Remark 6 In definition (11) optimization over probability measures can be equivalently replaced by optimization over positive measures, because if constraints are satisfied, then π is automatically probability measure.

Proof Using Proposition (6) to encode the primal conditions (11) in the Lagrangian, the primal problem (11) rewrites

$$\begin{aligned}
\inf_{\pi \geq 0} \sup_{M, \lambda_t, \tilde\lambda_t} &\int_{\Omega \times \tilde\Omega} d(\xi, \tilde\xi) d\pi - M(\mathbb{E}_\pi(\mathbb{1}_{\Omega\times\tilde\Omega}) - 1)\\
&- \sum_{s=1}^{T} (\mathbb{E}_\pi(\lambda_s) - \mathbb{E}_\pi(Pr_{s-1}(\lambda_s))) - \sum_{s=1}^{T} \Big(\mathbb{E}_\pi(\tilde\lambda_s) - \mathbb{E}_\pi(\tilde{Pr}_{s-1}(\tilde\lambda_s))\Big)
\end{aligned}$$

where the infimum is among all positive measures π (not only probability measures) and the supremum is among numbers M, $\sigma_{t-1,t}^{\Omega} \otimes \tilde\sigma_{t-1}^{\tilde\Omega}$-measurable functions λ_t and $\sigma_{t-1}^{\Omega} \otimes \tilde\sigma_{t-1,t}^{\tilde\Omega}$-measurbale functions $\tilde\lambda_t$. According to Sion's min-max Theorem this saddle point has the same objective value as

$$\sup_{M, \lambda_t, \tilde\lambda_t} \inf_{\pi \geq 0} M + \mathbb{E}_\pi \left(d(\xi, \tilde\xi) - M\mathbb{1}_{\Omega\times\tilde\Omega} - \sum_{s=1}^{T} (\lambda_s - Pr_{s-1}(\lambda_s)) - \sum_{s=1}^{T} \Big(\tilde\lambda_s - \tilde{Pr}_{s-1}(\tilde\lambda_s)\Big) \right). \tag{19}$$

Notice that the infimum is equal to $-\infty$ unless integrand is positive for every measure $\pi \geq 0$ which means that

$$M + \sum_{s=1}^{T} (\lambda_s - Pr_{s-1}(\lambda_s)) + \sum_{s=1}^{T} \left(\tilde{\lambda}_s - \tilde{Pr}_{s-1}(\tilde{\lambda}_s)\right) \leq d(\xi, \tilde{\xi})$$

has to hold for every $(\omega, \tilde{\omega}) \in \Omega \times \tilde{\Omega}$. For a positive integrand the infimum over all expectations in (19) is 0. We thus get rid of the primal variable π and obtain

$$\begin{array}{ll} \sup\limits_{M,\lambda_t,\tilde{\lambda}_t} & M \\ s.t. & M + \sum\limits_{s=1}^{T} (\lambda_s - Pr_{s-1}(\lambda_s)) + \sum\limits_{s=1}^{T} \left(\tilde{\lambda}_s - \tilde{Pr}_{s-1}(\tilde{\lambda}_s)\right) \leq d(\xi, \tilde{\xi}) \\ & \lambda_t \sim (\Omega \times \tilde{\Omega}, \sigma^{\Omega}_{t-1,t} \otimes \tilde{\sigma}^{\tilde{\Omega}}_{t-1}), \tilde{\lambda}_t \sim (\Omega \times \tilde{\Omega}, \sigma^{\Omega}_{t-1} \otimes \tilde{\sigma}^{\tilde{\Omega}}_{t-1,t}), \ \forall t \in \mathbb{T} \backslash \{0\} \end{array}$$

Now just define $\mu_t = \lambda_t - Pr_{t-1}(\lambda_t)$ and $\tilde{\mu}_t = \tilde{\lambda}_t - \tilde{Pr}_{t-1}(\tilde{\lambda}_t)$ to rewrite our problem in the following way

$$\begin{array}{ll} \sup\limits_{M,\mu_t,\tilde{\mu}_t} & M \\ s.t. & M + \sum\limits_{s=1}^{T} \mu_s + \sum\limits_{s=1}^{T} \mu_s \leq d(\xi, \tilde{\xi}) \\ & \mu_t \sim (\Omega \times \tilde{\Omega}, \sigma^{\Omega}_{t-1,t} \otimes \tilde{\sigma}^{\tilde{\Omega}}_{t-1}), \ \forall t \in \mathbb{T} \backslash \{0\} \\ & \tilde{\mu}_t \sim (\Omega \times \tilde{\Omega}, \sigma^{\Omega}_{t-1} \otimes \tilde{\sigma}^{\tilde{\Omega}}_{t-1,t}), \ \forall t \in \mathbb{T} \backslash \{0\} \\ & Pr_{t-1}(\mu_t) = 0, \tilde{Pr}_{t-1}(\tilde{\mu}_t) = 0 \end{array}$$

which is the desired assertion. □