Journal of Machine Learning Research 23 (2022) 1-52 Submitted 5/21; Revised 9/21; Published 2/22

Optimal Transport for Stationary Markov Chains via Policy Iteration

Kevin O'Connor KOCONN@LIVE.UNC.EDU
Department of Statistics and Operations Research
University of North Carolina, Chapel Hill
Chapel Hill, NC 27599, USA

Kevin McGoff KMCGOFF1@UNCC.EDU
Department of Mathematics and Statistics
University of North Carolina, Charlotte
Charlotte, NC 28223, USA

Andrew B. Nobel NOBEL@EMAIL.UNC.EDU
Department of Statistics and Operations Research
University of North Carolina, Chapel Hill
Chapel Hill, NC 27599, USA

Editor: Animashree Anandkumar

Abstract

We study the optimal transport problem for pairs of stationary finite-state Markov chains, with an emphasis on the computation of optimal transition couplings. Transition couplings are a constrained family of transport plans that capture the dynamics of Markov chains. Solutions of the optimal transition coupling (OTC) problem correspond to alignments of the two chains that minimize long-term average cost. We establish a connection between the OTC problem and Markov decision processes, and show that solutions of the OTC problem can be obtained via an adaptation of policy iteration. For settings with large state spaces, we develop a fast approximate algorithm based on an entropy-regularized version of the OTC problem, and provide bounds on its per-iteration complexity. We establish a stability result for both the regularized and unregularized algorithms, from which a statistical consistency result follows as a corollary. We validate our theoretical results empirically through a simulation study, demonstrating that the approximate algorithm exhibits faster overall runtime with low error. Finally, we extend the setting and application of our methods to hidden Markov models, and illustrate the potential use of the proposed algorithms in practice with an application to computer-generated music.

Keywords: optimal transport, Markov chains, Markov decision processes, stationary processes, entropic regularization

1. Introduction

The application and computation of optimal transport (OT) plans has recently received a great deal of attention within the machine learning community. Applications of optimal transport in machine learning include generative modeling (Arjovsky et al., 2017; Deshpande et al., 2018; Genevay et al., 2018; Kolouri et al., 2018; Salimans et al., 2018) and supervised learning (Frogner et al., 2015; Janati et al., 2019; Luise et al., 2018). In this paper, we study the optimal transport (OT) problem in the case where the objects of interest are stationary Markov chains or processes possessing hidden Markov structure. The problem of interest to us is distinct from traditional applications of coupling to Markov chains, e.g., to establish convergence to a stationary distribution. Our interest is in the computation of optimal transport plans for Markov chains that explicitly account for both stationarity and Markovian structure. In particular, we develop algorithms for computing solutions to a Markov-constrained form of the OT problem. The algorithms leverage recent advances in computational OT as well as techniques from Markov decision processes.

The principled extension of computational OT techniques to classes of distributions that possess additional structure, such as martingales or dependent processes, is an important direction of research. Indeed, some variations of constrained OT have been considered in recent work (Beiglböck et al., 2013; Zaev, 2015; Forrow et al., 2019; Moulos, 2021; Backhoff et al., forthcoming), and several recent applications of OT have focused on dependent observations (Schiebinger et al., 2019; Xu et al., 2018). Extensions of OT to dependent processes open the door to new applications in climate science, finance, epidemiology and other fields, where it is common for observations to possess temporal or spatial structure. The OT problem that we consider is tailored to the alignment and comparison of Markov chains and hidden Markov models (HMMs). As an illustration, we describe in Section 7 an application of the proposed techniques to the analysis of computer-generated music.

The primary contributions of this paper are as follows:

- We formulate a constrained version of the OT problem for stationary Markov chains, referred to as the optimal transition coupling (OTC) problem. The OTC problem aims to align the two chains of interest so as to minimize long-term average cost while preserving Markovity and stationarity.

- We detail an extension of the OTC problem to HMMs. In particular, we describe how one may couple a pair of HMMs via a coupling of their hidden chains using a cost that is derived from the OT cost between their emission distributions.

- We establish a useful connection between the OTC problem and Markov decision processes (MDPs) that provides a means of computing optimal solutions in an efficient manner. Leveraging this connection, we arrive at an algorithm combining policy iteration (Howard, 1960) with OT solvers that we refer to as `ExactOTC` (Algorithm 1). We state in Theorem 7 that if the two Markov chains of interest are irreducible, then `ExactOTC` converges to a solution of the OTC problem in a finite number of iterations.

- We introduce an entropically-constrained OTC problem and an associated regularized algorithm, referred to as `EntropicOTC` (Algorithm 2), that exhibits improved computational efficiency in theory and in practice. In Theorems 9 and 12, we establish upper bounds on the computational complexity of this algorithm, demonstrating that the runtime of each iteration is nearly-linear in the dimension of the couplings under study. This dependence is comparable to the state-of-the-art for computational OT.

- We prove a stability result for the OTC problem, stated formally in Theorem 13. Consistency of the plug-in estimate of the optimal transition coupling and its expected cost follows as a corollary (see Corollary 14).

The rest of the paper is organized as follows: We begin by providing some background on optimal transport and define the OTC problem in Section 2. In Section 3, we detail our extension of the OTC problem to HMMs. In Section 4, we establish the connection between the OTC problem and MDPs and state our result regarding `ExactOTC` for obtaining optimal transition couplings. A faster, regularized algorithm `EntropicOTC` for computing optimal transition couplings is described in Section 5. In Section 6 we present our result regarding the stability of the OTC problem and the statistical consistency of optimal transition couplings computed from data. In Section 7 we describe a simulation study and an application of our algorithms to computer-generated music. We close with a discussion of our results in Section 8. Proofs for all stated results may be found in Section 9.

Notation. Let $\mathbb{R}_+$ be the non-negative reals and $\Delta_n = \{u \in \mathbb{R}^n_+ | \sum_{i=1}^n u_i = 1\}$ denote the probability simplex in $\mathbb{R}^n$. Given a metric space $\mathcal{U}$, let $\mathcal{M}(\mathcal{U})$ denote the set of Borel probability measures on $\mathcal{U}$. For a vector $u \in \mathbb{R}^n$, let $\|u\|_\infty = \max_i |u_i|$ and $\|u\|_1 = \sum_i |u_i|$. Occasionally we will treat matrices in $\mathbb{R}^{n\times n}$ as vectors in $\mathbb{R}^{n^2}$.

2. The Optimal Transition Coupling Problem

The optimal transport problem is defined in terms of couplings and a cost function. Let $\mathcal{U}$ and $\mathcal{V}$ be metric spaces. Given probability measures $\mu \in \mathcal{M}(\mathcal{U})$ and $\nu \in \mathcal{M}(\mathcal{V})$, a *coupling* of μ and ν is a probability measure $\pi \in \mathcal{M}(\mathcal{U}\times\mathcal{V})$ such that $\pi(A\times\mathcal{V}) = \mu(A)$ and $\pi(\mathcal{U}\times B) = \nu(B)$ for every measurable $A \subset \mathcal{U}$ and $B \subset \mathcal{V}$. Let $\Pi(\mu,\nu)$ be the set of couplings of μ and ν. Let $c : \mathcal{U}\times\mathcal{V} \to \mathbb{R}$ be a cost function. We interpret $c(u, v)$ as the cost of transporting one unit of a quantity from $u \in \mathcal{U}$ to $v \in \mathcal{V}$, or vice versa. The optimal transport problem associated with μ, ν, and c is the program

$$\begin{aligned} \text{minimize} \quad & \int c \, d\pi \\ \text{subject to} \quad & \pi \in \Pi(\mu,\nu). \end{aligned} \tag{1}$$

As formulated, the problem (1) makes no particular assumptions about the structure of the measures μ and ν. In most existing applications, μ and ν represent the distribution of static quantities such as 3-dimensional point clouds, images of handwritten digits, social networks, or measurements of gene expression. However, in other application areas, μ and ν may represent dynamic quantities that vary with time or some other index. For example, μ and ν might be distributions of words in a block of text, the heart rate or blood pressure of a patient over a period of observation, or the daily high temperatures at two different locations over some period of time. In such cases, one may wish to constrain the types of couplings under consideration to ensure that they reflect the structure of the underlying distributions.

As a natural first step toward computational OT for dependent processes, we consider the case where μ and ν represent stationary Markov chains $X = (X_0, X_1, ...)$ and $Y = (Y_0, Y_1, ...)$ with values in finite sets $\mathcal{X}$ and $\mathcal{Y}$, respectively. Markov chains are a natural choice: their simple dependence structure is conducive to computation, and they can be studied in terms of transition matrices. Without loss of generality, assume that $\mathcal{X}$ and $\mathcal{Y}$ both contain d points. Let $P, Q \in [0, 1]^{d\times d}$ be the transition matrices, and let $p, q \in \Delta_d$ be the corresponding stationary distributions, of the chains X and Y, respectively. For a brief overview of the necessary background on Markov chains, we refer the reader to Section 9.1. For a more in-depth review of Markov chain theory, we refer the reader to Levin and Peres (2017). The extension of the OTC problem to hidden Markov models, detailed in Section 3, enables us to apply our approach to non-Markovian processes with long-range dependence and Polish alphabets.

Remark 1 *The optimal transport problem traces its roots back to the physical transportation of goods. In particular, the optimal coupling offers a means of stochastically matching a supply of some goods to their demand so as to minimize the expected cost of transporting the goods. In his book on the topic, Villani (Villani, 2008) offers an example of transporting loaves of bread between bakeries and cafés to build intuition for the optimal transport problem:*

> *Consider a large number of bakeries, producing loaves, that should be transported each morning to cafés where consumers will eat them. The amount of bread that can be produced at each bakery, and the amount that will be consumed at each café are known in advance, and can be modeled as probability measures ... on a certain space ... (equipped with the natural metric such that the distance between two points is the shortest path joining them). The problem is to find in practice where each unit of bread should go, in such a way as to minimize the total transport cost.*

In our setting, the collections of bakeries and cafés correspond to the finite sets $\mathcal{X}$ and $\mathcal{Y}$. However, unlike the static problem described by Villani, we consider a dynamic problem in which the number of loaves produced and consumed at the bakeries and cafés evolves over time. Indeed, we suppose that the amounts produced and consumed are determined by the distributions of stationary Markov chains X and Y. As we now have dependence over time to consider, the new problem is to synchronize the supply with the demand so as to minimize the total cost of transportation over the long term while still ensuring that the bakery and cafe owners are satisfied. To make things easier for the delivery driver, one might agree to consider only transport plans that do not change over time (stationary) and under which the deliveries tomorrow only depend on the deliveries today (Markov).

In principle, one may apply the standard optimal transport problem in the Markov setting by taking $\mathcal{U} = \mathcal{X}$, $\mathcal{V} = \mathcal{Y}$ and identifying an optimal coupling of the stationary distributions p and q. However, this marginal approach does not capture the dependence structure of the chains X and Y. Consider, for example, the case when $\mathcal{X} = \mathcal{Y} = \{0, 1\}$ with single-letter cost $c(x, y) = \delta(x \neq y)$, and

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \left[\begin{array}{cc} 1/2 & 1/2 \\ 1/2 & 1/2 \end{array} \right] \end{array} \quad \text{and} \quad Q = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right] \end{array}.$$

Note that the process X corresponding to P is iid, while the process Y corresponding to Q is deterministic (after conditioning on the initial symbol Y_0). Nevertheless, under a marginal analysis, the optimal transport distance between X and Y is zero since their stationary distributions p and q each coincide with the $(1/2, 1/2)$ measure. *In general, optimal coupling* *of stationary distributions yields a joint distribution on the product $\mathcal{X} \times \mathcal{Y}$, but it does not provide a means of generating a joint process having X and Y as marginals.* We seek a variation of (1) that captures and preserves the stochastic structure, namely stationarity and Markovity, of the processes X and Y.

As an alternative to a marginal analysis, one may consider instead the full measures $\mathbb{P} \in \mathcal{M}(\mathcal{X}^{\mathbb{N}})$ and $\mathbb{Q} \in \mathcal{M}(\mathcal{Y}^{\mathbb{N}})$ of the processes X and Y. Formally, $\mathbb{P}$ is the unique probability measure on $\mathcal{X}^{\mathbb{N}}$ such that for any cylinder set $[a_i^j] := \{(x_0, x_1, ...) \in \mathcal{X}^{\mathbb{N}} : x_k = a_k, i \leq k \leq j\}$,

$$\mathbb{P}([a_i^j]) := p(a_i) \prod_{k=i}^{j-1} P(a_k, a_{k+1}).$$

The measure $\mathbb{Q}$ is defined similarly in terms of q and Q. By definition, the measures $\mathbb{P}$ and $\mathbb{Q}$ are stationary, and Markovian. However, a coupling of $\mathbb{P}$ and $\mathbb{Q}$ on the joint sequence space $\mathcal{X}^{\mathbb{N}} \times \mathcal{Y}^{\mathbb{N}}$ need not be stationary or Markovian. To illustrate, let X' and Y' be iid Bernoulli($1/2, 1/2$) processes, independent of each other, defined on the same probability space. For $i \geq 0$ let $\tilde{X}_i = X'_i$, and let $\tilde{Y}_i = X'_i$ if i is a power of 2 and $\tilde{Y}_i = Y'_i$ otherwise. It is easy to see that the joint process $(\tilde{X}, \tilde{Y}) = (\tilde{X}_0, \tilde{Y}_0), (\tilde{X}_1, \tilde{Y}_1), \ldots$ is a coupling of X' and Y', but it is neither stationary nor Markovian. For further examples and discussion of non-Markovian couplings of Markov processes, see Ellis (1976, 1978, 1980b,a).

A joint process $(\tilde{X}, \tilde{Y})$ arising from a non-stationary or non-Markovian coupling of $\mathbb{P}$ and $\mathbb{Q}$ has a very different stochastic structure than the processes X and Y themselves, and will be difficult to work with computationally. Thus we wish to exclude such couplings from the feasible set of an optimal transport problem. An obvious fix is to consider the family $\Pi_{\mathrm{M}}(\mathbb{P}, \mathbb{Q})$, defined as the set of couplings $\mathbb{P}$ and $\mathbb{Q}$ that are stationary and Markovian. Viewed as processes, elements of $\Pi_{\mathrm{M}}(\mathbb{P}, \mathbb{Q})$ correspond to joint processes $(\tilde{X}, \tilde{Y})$ that are stationary, Markov, and satisfy $\tilde{X} \sim X$ and $\tilde{Y} \sim Y$. While this is a natural choice, the optimal transport cost associated with Π_{M} may violate the triangle inequality, even when the underlying cost function c is itself a metric, see Ellis (1976, 1978). Moreover, the family $\Pi_{\mathrm{M}}(\mathbb{P}, \mathbb{Q})$ is not characterized by a simple set of constraints (Boyle and Petersen, 2009). Motivated by the need for ready interpretation and tractable computation, we consider the set of stationary Markov chains on $\mathcal{X} \times \mathcal{Y}$ whose transition distributions are couplings of those of X and Y. A formal definition is given below. The resulting set of couplings, called transition couplings, is characterized by a simple set of linear constraints involving P and Q, and one may show (see Appendix A) that the resulting OT cost does satisfy the triangle inequality as long as the underlying cost c does.

In order to reduce notation when considering vectors and matrices indexed by elements of $\mathcal{X} \times \mathcal{Y}$, we will indicate only the cardinality of the index set and adopt an indexing convention whereby a vector $u \in \mathbb{R}^{d^2}$ is indexed as $u(x, y)$ and a matrix $R \in [0, 1]^{d^2 \times d^2}$ is indexed as $R((x, y), (x', y'))$ for $(x, y), (x', y') \in \mathcal{X} \times \mathcal{Y}$. Note also that vectors of the form $R((x, y), \cdot)$ will be regarded as row vectors.

Definition 2 *Let P and Q be transition matrices on finite state spaces $\mathcal{X}$ and $\mathcal{Y}$, respectively. A transition matrix $R \in [0, 1]^{d^2 \times d^2}$ is a **transition coupling** of P and Q if for every paired-state $(x, y) \in \mathcal{X} \times \mathcal{Y}$, the distribution $R((x, y), \cdot)$ is a coupling of the distributions $P(x, \cdot)$ and $Q(y, \cdot)$, formally $R((x, y), \cdot) \in \Pi(P(x, \cdot), Q(y, \cdot))$. Let $\Pi_{TC}(P, Q)$ denote the set of all transition couplings of P and Q.*

Standard results in Markov chain theory ensure that each transition coupling $R \in \Pi_{\mathrm{TC}}(P, Q)$ admits at least one stationary distribution $r \in \Delta_{d^2}$. Using r and R, one may construct a stationary Markov chain $(\tilde{X}, \tilde{Y}) = \{(\tilde{X}_i, \tilde{Y}_i)\}_{i \geq 0}$ taking values in $\mathcal{X} \times \mathcal{Y}$. We will also refer to couplings constructed in this way as transition couplings, as stated in the following definition.

Definition 3 *Let X and Y be stationary Markov chains with transition matrices P and Q on the finite state spaces $\mathcal{X}$ and $\mathcal{Y}$, respectively. A stationary Markov chain $(\tilde{X}, \tilde{Y}) = \{(\tilde{X}_i, \tilde{Y}_i)\}_{i \geq 0}$ taking values in $\mathcal{X} \times \mathcal{Y}$ with transition matrix $R \in [0,1]^{d^2 \times d^2}$ is a* ***transition coupling*** *of X and Y if $(\tilde{X}, \tilde{Y})$ is a coupling of X and Y and $R \in \Pi_{TC}(P, Q)$.*

Each transition coupling of X and Y may be associated with a process measure $\pi \in \mathcal{M}_s(\mathcal{X}^{\mathbb{N}} \times \mathcal{Y}^{\mathbb{N}})$; let $\Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$ denote the set of all such measures induced by transition couplings of X and Y. As the notation suggests, one may readily show that the process measure π induced by a transition coupling of X and Y is itself a coupling of the process measures $\mathbb{P}$ and $\mathbb{Q}$ associated with X and Y, respectively. As all elements of $\Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$ are also stationary and Markovian, it follows that $\Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q}) \subset \Pi_{\mathrm{M}}(\mathbb{P}, \mathbb{Q})$.

The couplings defined in Definition 3 are sometimes referred to as "Markovian couplings" in the literature (Levin and Peres, 2017), and they have been used, for example, to study diffusions (Banerjee and Kendall, 2018, 2016, 2017). We refer to such couplings as "transition couplings" in order to distinguish them from elements of $\Pi_{\mathrm{M}}(\mathbb{P}, \mathbb{Q})$. Note that $\Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q}) \neq \emptyset$ since it contains the independent coupling, namely, the stationary Markov chain on $\mathcal{X} \times \mathcal{Y}$ with transition matrix $P \otimes Q((x, y), (x', y')) = P(x, x')\, Q(y, y')$ for all (x, y) and (x', y'). The independent coupling corresponds to a paired chain $(\tilde{X}, \tilde{Y}) = \{(\tilde{X}_i, \tilde{Y}_i)\}_{i \geq 0}$ where $\tilde{X}$ and $\tilde{Y}$ are equal in distribution to X and Y, respectively, and evolve independently of one another.

A key advantage of considering $\Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$ over $\Pi_{\mathrm{M}}(\mathbb{P}, \mathbb{Q})$ is that the constraints defining $\Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$ are linear and thus computationally tractable (the constraints defining $\Pi_{\mathrm{M}}(\mathbb{P}, \mathbb{Q})$ are not). As we prove in Proposition 4 below, the set $\Pi_{\mathrm{TC}}(P, Q)$ of transition matrices actually characterizes the set $\Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$ of transition couplings if X and Y are irreducible. Stated differently, the condition $R \in \Pi_{\mathrm{TC}}(P, Q)$ is sufficient to ensure that a chain $(\tilde{X}, \tilde{Y})$ with transition matrix R is a transition coupling of X and Y. On the other hand, if X or Y is reducible, a stationary Markov chain with a transition matrix in $\Pi_{\mathrm{TC}}(P, Q)$ need not be a coupling of X and Y as the stationary distributions of P and Q are not unique. This follows from the fact that a transition coupling of reducible chains may admit as marginals any of the chains with transition matrices P or Q. So in order to solve the OTC problem by optimizing over $\Pi_{\mathrm{TC}}(P, Q)$ instead of $\Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$, we must be careful to avoid this situation. Proposition 4 ensures that this cannot occur if X and Y are irreducible.

Proposition 4 *Let X and Y be irreducible stationary Markov chains with transition matrices P and Q, respectively. Then any stationary Markov chain with a transition matrix contained in $\Pi_{TC}(P, Q)$ is a transition coupling of X and Y.*

As a result of Proposition 4, we may avoid working explicitly with transition couplings of X and Y and work instead with the set of matrices $\Pi_{\mathrm{TC}}(P, Q)$.

Letting $c : \mathcal{X}^{\mathbb{N}} \times \mathcal{Y}^{\mathbb{N}} \to \mathbb{R}$ be a cost function defined on sample sequences of X and Y, we define the *optimal transition coupling (OTC) problem* for X and Y with cost c to be the program

$$\begin{aligned} &\text{minimize} \quad \int c\, d\pi \\ &\text{subject to} \quad \pi \in \Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q}). \end{aligned} \tag{2}$$

The minimum in (2), referred to as the OTC cost, assesses the degree to which the two chains may be "synced up" with respect to c. Any solution to (2) describes the joint distribution of the synchronized chains. Moreover, as a consequence of the pointwise ergodic theorem, any optimal transition coupling $\pi \in \Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$ in Problem (2) is also optimal with respect to the averaged cost $((x_0, x_1, ...), (y_0, y_1, ...)) \mapsto \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} c((x_i, x_{i+1}, ...), (y_i, y_{i+1}, ...))$. In this sense, the quality of an alignment (equivalently, transition coupling) of the two chains X and Y is assessed based on its long-term average cost.

In the remainder of the paper we assume that c is a single-letter cost, i.e., $c((x_0, x_1, ...), (y_0, y_1, ...)) = \tilde{c}(x_0, y_0)$ for some cost function $\tilde{c} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$. In most of what follows we identify c and $\tilde{c}$, regarding c as a function on $\mathcal{X} \times \mathcal{Y}$ and writing $c(x_0, y_0)$ when no confusion will arise. The consideration of single-letter costs is motivated by our focus on computation and reflects existing work on computational OT, where a cost or metric is defined *a priori* on static observations. Single letter costs have also been the focus of previous work on optimal transport problems for stationary processes (Ornstein, 1973; Gray et al., 1975). Our arguments may be easily adapted to the case when the cost depends on a finite number of coordinates. In particular, any k-letter cost $c : \mathcal{X}^k \times \mathcal{Y}^k \to \mathbb{R}_+$ may be regarded as single-letter for the chains $\tilde{X} = (X_0^{k-1}, X_1^k, ...)$ and $\tilde{Y} = (Y_0^{k-1}, Y_1^k, ...)$ on $\mathcal{X}^k$ and $\mathcal{Y}^k$, respectively. For single-letter costs, we show in Appendix A that optimal transition couplings exist, and that the OTC cost satisfies the triangle inequality whenever c does. Note that c is necessarily bounded, as $\mathcal{X}$ and $\mathcal{Y}$ are finite. Moreover, there is no loss in generality in assuming that c is non-negative since our results also hold after adding a constant to c.

A primary contribution of this paper, and the focus of Sections 4 and 5, is the development of efficient algorithms for computing solutions to the OTC problem (2). Note that this problem involves the minimization of a linear objective over the non-convex set $\Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$, which makes it difficult to find a solution with off-the-shelf methods. Proposition 4 shows that one may optimize instead over the convex polyhedron $\Pi_{\mathrm{TC}}(P, Q)$: informally, the program (2) can be reformulated as minimizing $\mathbb{E}c(\tilde{X}_0, \tilde{Y}_0)$ over $R \in \Pi_{\mathrm{TC}}(P, Q)$, where $(\tilde{X}, \tilde{Y})$ is a stationary Markov chain generated by R. However, this reformulation has a non-convex objective, so some care is needed in order to obtain global solutions.

2.1 Related Work

Stationary couplings of stationary processes, known as *joinings*, were first studied by Furstenberg (1967). Distances between processes based on joinings have been proposed in the ergodic theory literature (Gray et al., 1975; Ornstein, 1973; McGoff et al., forthcoming; McGoff and Nobel, 2020, 2021; O'Connor et al., 2021a), but they have been explored primarily as a theoretical tool: no tractable algorithms have been proposed for computing such distances exactly. In the context of Markov chains, coupling methods have been widely used as a tool to establish rates of convergence (see for instance Griffeath 1976 or Lindvall 2002). Examples of optimal Markovian couplings of Markov processes are studied by Ellis

(1976, 1978, 1980a,b). Another line of work has explored total variation-type distances for models with Markovian structure. For example, Chen and Kiefer (2014), and Kiefer (2018) develop algorithms for and consider the computability of the total variation distance between hidden Markov models and labeled Markov chains. Similarly, Daca et al. (2016) study the inestimability of the total variation distance between Markov chains. More recent work has proposed direct adaptations of the optimal transport problem for processes with Markovian structure. Moulos (2021) study the bicausal optimal transport problem for Markov chains and its connection to Markov decision processes. Unlike the OTC problem, in the bicausal transport problem, couplings are not required to be stationary or Markov themselves. O'Connor et al. (2021b) apply the OTC problem and the tools presented in this paper to the comparison and alignment of graphs. We also remark that the optimal transition coupling problem appears in an unpublished manuscript of Aldous and Diaconis (2009).

Some existing work (Song et al., 2016; Zhang, 2000) studies a modified form of the OTC problem that we refer to as the *1-step* transition coupling problem. In the 1-step transition coupling problem the expected cost is measured with respect to the 1-step transition probabilities rather than the stationary distribution of the transition coupling. In particular, a transition coupling $R \in \Pi_{\mathrm{TC}}(P, Q)$ is 1-step optimal if for every $(x, y) \in \mathcal{X} \times \mathcal{Y}$,

$$R((x,y),\cdot) \in \operatorname*{argmin}_{r \in \Pi(P(x,\cdot),Q(y,\cdot))} \sum_{(x',y')} r(x',y')\, c(x',y').$$

Loosely, one can view the OTC problem (2) as an infinite-step version of the 1-step OTC problem, wherein a transition coupling is chosen that minimizes the expected cost averaged over an infinite number of steps. The 1-step transition coupling problem is used by Song et al. (2016) to assess the distance between Markov decision processes. In another direction, Zhang (2000) show that solutions to the 1-step transition coupling problem exist for Markov processes on Polish state spaces and lower semicontinuous cost functions. While the 1-step problem is computationally convenient, in some situations it will yield poor alignments of the two chains of interest. We provide an example to illustrate this in Appendix C, showing that the 1-step approach can yield a transition coupling with arbitrarily high expected cost over time.

Other work has considered modifications of standard computational OT techniques for time series that do not necessarily possess Markovian structure. Cazelles et al. (2020) study the Wasserstein-Fourier distance, which is the Wasserstein distance between normalized power spectral densities, while Muskulus and Verduyn-Lunel (2011) suggest using the optimal transport cost between the k-block empirical measures constructed from observed samples. For general observed sequences, Su and Hua (2018) consider only couplings that do not disturb the ordering of the two sequences too much, as quantified by the inverse difference moment. Another line of work (Cohen et al., 2021; Cuturi and Blondel, 2017; Janati et al., 2020) has explored distances between time series based on dynamic time warping (DTW). Similar in spirit to OT, the DTW problem seeks an alignment of observations in two time series that respects the ordering of the respective observations and minimizes a total cost. In contrast to these approaches, we seek a more direct modification of the optimal transport problem itself that best captures the Markovian dynamics.

Entropic regularization in OT traces its roots back to traffic modeling techniques in transportation theory (Wilson, 1969). Cuturi (2013) showed how one may solve the entropy-regularized OT problem via a matrix scaling algorithm proposed by Sinkhorn (1967). Owing to the increased computational efficiency of matrix multiplication over linear programming, Cuturi's result placed entropic OT as an efficient alternative to standard OT in high-dimensional (large d) scenarios. Altschuler et al. (2017) provided further analysis of Sinkhorn's algorithm, showing that for appropriate choice of regularization coefficient and number of iterations, it yields an approximation of the unregularized OT cost in near-linear time. More recent work (Dvurechensky et al., 2018; Lin et al., 2019; Guo et al., 2020) has considered alternative algorithms based on stochastic gradient decent for solving entropy-regularized OT problems. An entropy rate-regularized optimal joining problem and its statistical properties are studied O'Connor et al. (2021a). We remark that an upper bound on the entropy of each transition distribution of a Markov chain (as considered in this paper) implies an upper bound on the entropy rate of the chain.

3. Extension of OTC to Hidden Markov Models

Markov models are often employed as components of more complex models for sequential observations. Hidden Markov models (HMMs) are a widely used variant of the Markov model in which observations are modeled as conditionally independent random emissions arising from a latent Markov chain. HMMs have been applied successfully to a variety of problems including speech recognition (Bahl et al., 1986; Varga and Moore, 1990), text segmentation (Yamron et al., 1998), and modeling disease progression (Williams et al., 2020). For a detailed overview, we refer the reader to the text of Zucchini et al. (2017).

Formally, a HMM may be characterized by a pair (X, ϕ) where $X = (X_0, X_1, ...)$ is an unobserved Markov chain taking values in a finite set $\mathcal{X}$, and a function $\phi : \mathcal{X} \to \mathcal{M}(\mathcal{U})$ that maps each state $x \in \mathcal{X}$ to a distribution on a fixed observation space $\mathcal{U}$. The pair (X, ϕ) gives rise to a stationary process $U = (U_0, U_1, ...)$ where $U_0, U_1, \ldots \in \mathcal{U}$ are conditionally independent given X with $U_i \sim \phi(X_i)$ for $i \geq 0$. Note that the process U may exhibit long-range dependence. In this way, HMMs provide a simple means of modeling sequences with more complex dependence structures.

The OTC problem may be extended to processes with hidden Markov structure as follows. Let (X, ϕ) and (Y, ψ) be a pair of HMMs with observation spaces $\mathcal{U}$ and $\mathcal{V}$, respectively, and let $c : \mathcal{U} \times \mathcal{V} \to \mathbb{R}_+$. Note that the cost c is specified on the observed spaces $\mathcal{U}$ and $\mathcal{V}$ rather than the state spaces of the unobserved Markov chains X and Y. However, one may extend c to a cost on $\mathcal{X} \times \mathcal{Y}$ by optimally coupling the emission distributions $\phi(x)$ and $\psi(y)$ for every pair $(x, y) \in \mathcal{X} \times \mathcal{Y}$. In more detail, let $\theta : \mathcal{X} \times \mathcal{Y} \to \mathcal{M}(\mathcal{U} \times \mathcal{V})$ and $c' : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$ be defined by

$$\theta(x, y) \in \operatorname*{argmin}_{\pi \in \Pi(\phi(x), \psi(y))} \int c \, d\pi \qquad \text{and} \qquad c'(x, y) = \min_{\pi \in \Pi(\phi(x), \psi(y))} \int c \, d\pi.$$

In other words, we define the functions $\theta : \mathcal{X} \times \mathcal{Y} \to \mathcal{M}(\mathcal{U} \times \mathcal{V})$ and $c' : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$ such that for every $(x, y) \in \mathcal{X} \times \mathcal{Y}$, $\theta(x, y)$ is an optimal coupling and $c'(x, y)$ is the OT cost of the emission distributions $\phi(x)$ and $\psi(y)$ with respect to c. One may then find an optimal transition coupling (X', Y') of X and Y with respect to c' as in problem (2). The expected cost of this transition coupling corresponds to a cost between the HMMs (X, ϕ) and (Y, ψ) taking the original cost c into account. Moreover, the pair $((X', Y'), \theta)$ defines an optimal joint HMM of (X, ϕ) and (Y, ψ) from which samples in $\mathcal{U} \times \mathcal{V}$ may be drawn.

Leveraging the intuition from the standard OTC problem, the optimal transition coupling $((X', Y'), \theta)$ may be thought of as an alignment of the two HMMs (X, ϕ) and (Y, ψ) with respect to c. In this way, we may apply the OTC problem to any processes that can be embedded as or are well-approximated by HMMs. Before proceeding, we remark that Chen et al. (2019) also proposes an OT problem for HMMs based on coupling the emission distributions of the two HMMs of interest. However, the latent Markov chains of either HMM are coupled using standard OT after a registration step. Our approach captures the Markovity of the latent sequences more directly and allows one to generate new samples from the coupled HMM.

4. Computing Optimal Transition Couplings

In this section, we turn our attention toward our primary goal of developing tractable algorithms for solving the OTC problem (2). As discussed in Section 2, the OTC problem is a non-convex, constrained optimization problem and thus there is little hope of obtaining global solutions via generic optimization algorithms. Adopting a more tailored approach, we draw a connection between the OTC problem and Markov decision processes (MDP). Having established this connection, we may leverage the wealth of algorithms for obtaining global solutions to MDPs to solve the OTC problem. As we will show, the framework of policy iteration naturally lends itself to our problem and leads to a computationally tractable algorithm combining standard MDP techniques with OT solvers.

4.1 Connection to Markov Decision Processes

A Markov decision process is characterized by a 4-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, c')$ consisting of a state space $\mathcal{S}$, an action space $\mathcal{A} = \bigcup_s \mathcal{A}_s$ where $\mathcal{A}_s$ is the set of allowable actions in state s, a set of transition distributions $\mathcal{P} = \{p(\cdot|s, a) : s \in \mathcal{S}, a \in \mathcal{A}\}$ on $\mathcal{S}$, and a cost function $c' : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$. At each time step the process occupies a state $s \in \mathcal{S}$ and an agent chooses an action $a \in \mathcal{A}_s$; the process incurs a cost $c'(s, a)$ and then moves to a new state according to the distribution $p(\cdot|s, a)$. Informally, the goal of the agent is to choose actions to minimize her average cost. The behavior of an agent is described by a family $\gamma = \{\gamma_s(\cdot) : s \in \mathcal{S}\}$ of distributions $\gamma_s(\cdot) \in \mathcal{M}(\mathcal{A}_s)$ on the set of admissible actions, which is known as a *policy*. An agent following policy γ chooses her next action according to $\gamma_s(\cdot)$ whenever the system is in state s, independently of her previous actions.

It is easy to see that, in conjunction with the transition distributions $\mathcal{P}$, every policy γ induces a collection of Markov chains on the state space $\mathcal{S}$ indexed by initial states $s \in \mathcal{S}$. In the average-cost MDP problem the goal is to identify a policy for which the induced Markov chain minimizes the limiting average cost, namely a policy γ minimizing

$$\bar{c}_\gamma(s) := \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_\gamma \left[c'(s_t, a_t) \middle| s_0 = s \right], \tag{3}$$

for each $s \in \mathcal{S}$. Note that the expectation in (3) is taken with respect to the Markov chain induced by γ. In general, the limiting average cost $\bar{c}_\gamma(s)$ will depend on the initial state s, but if γ induces an ergodic chain then the average cost will be constant. If all policies induce ergodic Markov chains, the MDP is referred to as "unichain"; otherwise the MDP is classified as "multichain". We refer the reader to Puterman (2005) for more details on MDPs.

The OTC problem (2) may readily be recast as an MDP. In detail, let the state space $\mathcal{S} = \mathcal{X} \times \mathcal{Y}$, and let $s = (x, y)$ denote an element of $\mathcal{S}$. Define the set of admissible actions in state s to be the corresponding set of row couplings $\mathcal{A}_s = \Pi(P(x,\cdot), Q(y,\cdot))$. For each state s and action $r_s \in \mathcal{A}_s$ define the transition distribution $p(\cdot|s, r_s) := r_s(\cdot)$, and the cost function $c'(s, r_s) = c(s) = c(x, y)$. Note that c' is independent of the action r_s. We refer to this MDP as TC-MDP.

Any policy γ for TC-MDP specifies distributions over $\Pi(P(x,\cdot), Q(y,\cdot))$ for each $(x, y) \in \mathcal{X} \times \mathcal{Y}$ and thus corresponds to a single distribution over $\Pi_{\text{TC}}(P, Q)$ that governs the random actions of the agent. In TC-MDP it suffices to consider only deterministic policies γ, namely policies such that for each state $s = (x, y)$ the distribution $\gamma_s(\cdot)$ is a point mass at unique element of $\mathcal{A}_s = \Pi(P(x,\cdot), Q(y,\cdot))$.

Proposition 5 *Let γ be a policy for TC-MDP. Then there exists a deterministic policy $\tilde{\gamma}$ such that $\bar{c}_\gamma(s) = \bar{c}_{\tilde{\gamma}}(s)$ for every $s \in \mathcal{S}$.*

Thus optimization over $\Pi_{\text{TC}}(P, Q)$ is equivalent to optimization over deterministic policies. Importantly, *a deterministic policy corresponds to a fixed transition coupling matrix $R \in \Pi_{\text{TC}}(P, Q)$*. Going forward, we refer to $R \in \Pi_{\text{TC}}(P, Q)$ directly instead of the equivalent deterministic policy $\tilde{\gamma}$ in our notation. We note that, even when X and Y are ergodic, the same may not be true of the stationary Markov chain induced by a transition coupling matrix $R \in \Pi_{\text{TC}}(P, Q)$ (see Appendix B). Specifically, a single element of $\Pi_{\text{TC}}(P, Q)$ may have multiple stationary distributions and thus give rise to multiple stationary Markov chains depending on the initial state $s \in \mathcal{S}$. Thus TC-MDP is classified as multichain. Finally, we may formalize the relationship between the OTC problem and TC-MDP.

Proposition 6 *If X and Y are irreducible, then any $R \in \Pi(P, Q)$ that is an optimal policy for TC-MDP corresponds to an optimal coupling $\pi_R \in \Pi_{TC}(\mathbb{P}, \mathbb{Q})$ with expected cost $\min_{s \in \mathcal{S}} \bar{c}_R(s)$.*

4.2 Policy Iteration

Now that we have shown that the OTC problem can be viewed as an MDP, we can leverage existing algorithms for MDPs to obtain solutions. To this end, we propose to adapt the framework of policy iteration (Howard, 1960). To facilitate our discussion, in what follows, we regard the cost function c and limiting average cost $\bar{c}_R$ as vectors in $\mathbb{R}_+^{d^2}$. For each $R \in \Pi_{\text{TC}}(P, Q)$, standard results (Puterman, 2005) guarantee that the limit $\overline{R} := \lim_{T\to\infty} T^{-1} \sum_{t=0}^{T-1} R^t$ exists. When R is aperiodic and irreducible, the Perron-Frobenius theorem ensures that $\overline{R} = \lim_{T\to\infty} R^T$ and the rows of $\overline{R}$ are equal to the stationary distributions of R.

In policy iteration, one repeatedly evaluates and improves policies. In the context of TC-MDP, for a given transition coupling matrix $R \in \Pi_{\text{TC}}(P, Q)$ the evaluation step computes the average cost (*gain*) vector $g = \overline{R}\, c$ and the total extra cost (*bias*) vector $h = \sum_{t=0}^{\infty} R^t(c - g)$. In practice, g and h may be obtained by solving a linear system of equations rather than evaluating infinite sums (see Algorithm 1a) (Puterman, 2005). The improvement step selects a new transition coupling matrix R' that minimizes $R'\, g$ or, if no improvement is possible, $R'\, h$ in an element-wise fashion (see Algorithm 1b). In more detail, we may select a transition coupling R' such that for each (x, y) the corresponding row $r = R'((x, y), \cdot)$ minimizes rg (or rh) over couplings $r \in \Pi(P(x, \cdot), Q(y, \cdot))$. To denote the element-wise argmin, we write $\text{elem-argmin}_{R \in \Pi_{\text{TC}}(P,Q)}\, R\, g$ (or $R\, h$). The improved matrix R' is obtained by solving d^2 OT problems with marginals $P(x, \cdot)$ and $Q(y, \cdot)$ and cost g (or h). This special feature of TC-MDP enables us to find improved transition coupling matrices in a computationally efficient manner despite working with an infinite action space. Once a fixed point in the evaluation and improvement process is reached, the procedure terminates. The resulting algorithm will be referred to as `ExactOTC` (see Algorithm 1). We initialize Algorithm 1 to the independent transition coupling $P \otimes Q$, defined in Section 2.

Algorithm 1: `ExactOTC`

```
R_0 ← P ⊗ Q, n ← 0
while not converged do
    /* transition coupling evaluation */
    (g_n, h_n) ← ExactTCE(R_n)
    /* transition coupling improvement */
    R_{n+1} ← ExactTCI(g_n, h_n, R_n, Π_TC(P, Q))
    n ← n + 1
return R_n
```

Algorithm 1a: `ExactTCE`

input : R

Solve for (g, h, w) such that

$$\begin{bmatrix} I - R & 0 & 0 \\ I & I - R & 0 \\ 0 & I & I - R \end{bmatrix} \begin{bmatrix} g \\ h \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ c \\ 0 \end{bmatrix}$$

return (g, h)

Algorithm 1b: `ExactTCI`

```
input : g, h, R_0, Π
/* element-wise argmin */
R' ← elem-argmin_{R∈Π} Rg
if R'g = R_0 g then
    R' ← elem-argmin_{R∈Π} Rh
    if R'h = R_0 h then
        return R_0
    else
        return R'
else
    return R'
```

For finite state and action spaces, policy iteration is known to yield an optimal policy for the average-cost MDP in a finite number of steps (Puterman, 2005). While policy iteration may fail to converge for general compact action spaces (Dekker, 1987; Schweitzer, 1985; Puterman, 2005), as is the case for TC-MDP, we may exploit the polyhedral structure of $\Pi_{\text{TC}}(P, Q)$ to establish the following convergence result.

Theorem 7 *Algorithm 1 converges to a solution (g^*, h^*, R^*) of TC-MDP in a finite number of iterations. Moreover, if X and Y are irreducible, R^* is the transition matrix of an optimal transition coupling of X and Y.*

Recall from Proposition 6 that an optimal solution to TC-MDP necessarily yields an optimal solution to (1). Thus Theorem 7 ensures that a solution to the OTC problem can be obtained from Algorithm 1 in a finite number of iterations. A proof of this result can be found in Section 9.3.3.

Remark 8 *One may in principle adapt other MDP algorithms to solve the OTC problem. However, the standard alternatives to policy iteration either do not admit a computationally tractable implementation (e.g. linear programming) or are not as conducive to a convergence analysis (e.g. value iteration). We choose policy iteration because it balances both of these features, admitting a practical implementation while also enabling a theoretical convergence analysis. We acknowledge that OTC solvers based on policy iteration may not be preferable in every scenario and leave a detailed exploration of other MDP algorithms for the OTC problem to future work.*

5. Fast Approximate Policy Iteration

The simplicity of Algorithm 1 in conjunction with the theoretical guarantee of Theorem 7 make it an appealing method for solving the OTC problem when the cardinality d of the state spaces of X and Y is small. However, each call to Algorithm 1a involves solving a system of $3d^2$ linear equations, requiring a total of $\mathcal{O}(d^6)$ operations. Furthermore, each call to Algorithm 1b entails solving d^2 linear programs each with $\mathcal{O}(d)$ constraints, which can be accomplished in a total time of $\mathcal{O}(d^5 \log d)$. We note that a similar dependence on the dimension of each coupling is observed in exact OT algorithms, such as the network simplex algorithm in (Peyré and Cuturi, 2019). For even moderate values of d, this may be too slow for practical use.

To alleviate the poor scaling with the dimension of the couplings in the standard OT problem, one may use entropic regularization, whereby a negative entropy term is added to the OT objective. Cuturi (2013) showed that solutions to the entropy-regularized OT problem may be obtained efficiently via Sinkhorn's algorithm (Sinkhorn, 1967). More recently, Altschuler et al. (2017) proved that Sinkhorn's algorithm yields an approximation of the OT cost with error bounded by ε in near-linear time with respect to the dimension of the couplings under consideration. Subsequent work (Dvurechensky et al., 2018; Lin et al., 2019; Guo et al., 2020) has proposed and studied alternative algorithms for approximating the optimal transport cost, each with runtime scaling at least linearly with the dimension of the couplings in the problem. One might hope that a similar dependence on the size of the elements of $\Pi_{\text{TC}}(P, Q)$ may be achievable for the OTC problem by employing regularization.

In this section, we extend entropic regularization techniques to the OTC problem. This extension leads to an approximate algorithm that runs in $\tilde{\mathcal{O}}(d^4)$ time per iteration, where $\tilde{\mathcal{O}}(\cdot)$ omits non-leading poly-logarithmic factors. This complexity is nearly-linear in the dimension d^4 of the transition couplings. We first propose a truncation-based approximation of the `ExactTCE` transition coupling evaluation algorithm, which we call `ApproxTCE`. When the transition coupling to be evaluated satisfies a simple regularity condition, we show that one can obtain approximations of the gain and bias from `ApproxTCE` with error bounded by ε in $\tilde{\mathcal{O}}(d^4 \log \varepsilon^{-1})$ time.

Mirroring the derivation of entropic OT, we then propose an entropy-regularized approximation of the `ExactTCI` transition coupling improvement algorithm, called `EntropicTCI`. We perform a new analysis of the Sinkhorn algorithm (described in Section 5.3) that is tailored to transition coupling improvement to show that `EntropicTCI` yields an improved transition coupling with error bounded by ε in $\tilde{\mathcal{O}}(d^4 \varepsilon^{-4})$ time. Combining these two algorithms, we obtain the `EntropicOTC` algorithm, which runs in $\tilde{\mathcal{O}}(d^4 \varepsilon^{-4})$ time per iteration. We provide empirical support for these theoretical results through a simulation study in Section 7. We find that the improved efficiency at each iteration of `EntropicOTC` leads to a much faster runtime in practice as compared to `ExactOTC`. Our experiments also show that `EntropicOTC` yields an expected cost that closely approximates the unregularized OTC cost.

5.1 Constrained Optimal Transition Coupling Problem

We begin by defining a constrained set of transition couplings. Let $\mathcal{K}(\cdot\|\cdot)$ be the Kullback-Leibler (KL) divergence defined for $u, v \in \Delta_{d^2}$ by $\mathcal{K}(u\|v) = \sum_s u(s) \log(u(s)/v(s))$ with the convention that $0 \log(0/0) = 0$ and $\mathcal{K}(u\|v) = +\infty$ if $u(s) > 0$ and $v(s) = 0$ for some index s. For every $\eta > 0$ and $(x, y) \in \mathcal{X} \times \mathcal{Y}$, define the set

$$\Pi_\eta(P(x,\cdot), Q(y,\cdot)) = \left\{ r \in \Pi(P(x,\cdot), Q(y,\cdot)) : \mathcal{K}\big(r \| P \otimes Q((x,y),\cdot)\big) \leq \eta \right\},$$

and the subset of transition coupling matrices

$$\Pi_{\mathrm{TC}}^\eta(P, Q) = \{R \in \Pi_{\mathrm{TC}}(P, Q) : R((x,y),\cdot) \in \Pi_\eta(P(x,\cdot), Q(y,\cdot)), \ \forall (x,y) \in \mathcal{X} \times \mathcal{Y}\}.$$

Elements of $\Pi_{\mathrm{TC}}^\eta(P, Q)$ have rows that are close in KL-divergence to the rows of the independent transition coupling $P \otimes Q$. When P and Q are aperiodic and irreducible, the same is true of $P \otimes Q$. Fix $\eta > 0$ and let $\Pi_{\mathrm{TC}}^\eta(\mathbb{P}, \mathbb{Q})$ be the set of transition couplings with transition matrices in $\Pi_{\mathrm{TC}}^\eta(P, Q)$. The *entropic OTC problem* is

$$\begin{aligned} &\text{minimize} \quad && \int c \, d\pi \\ &\text{subject to} \quad && \pi \in \Pi_{\mathrm{TC}}^\eta(\mathbb{P}, \mathbb{Q}). \end{aligned} \tag{4}$$

For completeness, we establish in Appendix A that a solution to (4) exists. As the divergence $\mathcal{K}(r \| P \otimes Q(s,\cdot))$ is bounded for $r \in \Pi(P(x,\cdot), Q(y,\cdot))$ and $s = (x,y) \in \mathcal{X} \times \mathcal{Y}$ (Cuturi, 2013), the program (4) coincides with the unconstrained OTC problem for sufficiently large η. Finally, note that (4) corresponds to an MDP in the same way that (2) does but with a constrained set of policies. In the rest of the section, we develop computationally efficient alternatives to Algorithms 1a and 1b for this constrained MDP.

5.2 Fast Approximate Transition Coupling Evaluation

Next, we propose a fast approximation of Algorithm 1a. Recall from our previous discussion that the gain vector g corresponding to any aperiodic and irreducible $R \in \Pi_{\mathrm{TC}}(P, Q)$ is constant and thus may be written as $g = g_0 \mathbb{1}$ for a scalar g_0. Fixing such an $R \in \Pi_{\mathrm{TC}}(P, Q)$ and $L, T \geq 1$, we approximate the gain g by averaging the cost over L steps of the Markov chain corresponding to R from each possible starting point in $\mathcal{X} \times \mathcal{Y}$. Moreover, we approximate the bias h by summing the total extra cost over T steps with respect to the approximate gain $\tilde{g}$. Formally, let $\tilde{g} := (d^{-2}(R^L c)^\top \mathbb{1})\mathbb{1}$ and $\tilde{h} := \sum_{t=0}^{T} R^t(c - \tilde{g})$. The resulting algorithm, which we refer to as `ApproxTCE`, is detailed in Algorithm 2a.

Algorithm 2a: `ApproxTCE`

input : R, L, T
$\tilde{g} \leftarrow (d^{-2}(R^L c)^\top \mathbb{1})\mathbb{1}$
$\tilde{h} \leftarrow \sum_{t=0}^{T} R^t(c - \tilde{g})$
return $(\tilde{g}, \tilde{h})$

The approximations $\tilde{g}$ and $\tilde{h}$ can be computed in $\mathcal{O}(Ld^4)$ and $\mathcal{O}(Td^4)$ time, respectively. Since g and h are equal to the limits of $\tilde{g}$ and $\tilde{h}$ as $L, T \to \infty$, we expect that larger L and T will yield better approximations. One must ensure that the L and T that are required for a good approximation do not grow too quickly with d. We show that this is the case in Theorem 9 below. We will say that a transition matrix $R \in [0,1]^{d^2 \times d^2}$ with stationary distribution $\lambda \in \Delta_{d^2}$ is mixing with coefficients $M \in \mathbb{R}_+$ and $\alpha \in [0,1)$ if for every $t \in \mathbb{N}$, $\max_{s \in \mathcal{X} \times \mathcal{Y}} \|R^t(s, \cdot) - \lambda\|_1 \leq M\alpha^t$. Recall that R is mixing whenever it is aperiodic and irreducible.

Theorem 9 *Let $R \in \Pi_{TC}(P,Q)$ be aperiodic and irreducible with mixing coefficients $M \in \mathbb{R}_+$ and $\alpha \in [0,1)$ and gain and bias vectors $g \in \mathbb{R}^{d^2}$ and $h \in \mathbb{R}^{d^2}$, respectively. Then for any $\varepsilon > 0$, there exist $L, T \in \mathbb{N}$ such that* `ApproxTCE`(R, L, T) *yields $(\tilde{g}, \tilde{h})$ satisfying $\|\tilde{g} - g\|_\infty \leq \varepsilon$ and $\|\tilde{h} - h\|_1 \leq \varepsilon$ in $\tilde{\mathcal{O}}\left(\frac{d^4}{\log \alpha^{-1}} \log\left(\frac{M}{\varepsilon(1-\alpha)}\right)\right)$ time.*

In particular, `ApproxTCE` does approximate `ExactTCE` in time scaling like $\tilde{\mathcal{O}}(d^4)$. Explicit choices of L and T are given in the proof of Theorem 9, which may be found in Section 9.4.

Remark 10 *In practice, values of L and T satisfying the conclusion of Theorem 9 are unknown. In our experiments, we found that running Algorithm 2a with large, fixed values of L and T yields a high approximation accuracy while still running significantly more quickly than Algorithm 1a. Alternatively, L and T may be chosen adaptively by computing vectors $\tilde{g}$ and $\tilde{h}$ iteratively for larger and larger values of L and T until some convergence criterion is satisfied or L and T hit some prespecified thresholds. For example, letting $\tilde{g}^L$ and $\tilde{h}^T$ be the iterates of this procedure, one may iterate until $\|\tilde{g}^L - \tilde{g}^{L-1}\|_\infty < \varepsilon$ and $\|\tilde{h}^T - \tilde{h}^{T-1}\|_\infty < \varepsilon$. This approach achieves the same worst-case complexity as Algorithm 2a but allows for time-savings when the chain R mixes quickly.*

For a set $\mathcal{U} \subset \mathbb{R}^n$, let $B_\varepsilon(u) \subset \mathbb{R}^n$ be the open ball of radius $\varepsilon > 0$ centered at $u \in \mathcal{U}$, and let $\text{aff}(\mathcal{U})$ denote the affine hull, defined as $\text{aff}(\mathcal{U}) = \{\sum_{i=1}^k \alpha_i u_i : k \in \mathbb{N}, u_1, ..., u_k \in \mathcal{U}, \sum_{i=1}^k \alpha_i = 1\}$. Let $\text{ri}(\cdot)$ denote the relative interior, defined as $\text{ri}(\mathcal{U}) = \{u \in \mathcal{U} : \exists \varepsilon > 0 \text{ s.t. } B_\varepsilon(u) \cap \text{aff}(\mathcal{U}) \subset \mathcal{U}\}$.

Proposition 11 *If P and Q are aperiodic and irreducible then every $R \in \text{ri}(\Pi_{TC}(P,Q))$ is also aperiodic and irreducible, and thus mixing.*

As a consequence of Proposition 11, we need only verify that $R \in \mathrm{ri}(\Pi_{\mathrm{TC}}(P,Q))$ to ensure that Theorem 9 holds and that we may perform fast transition coupling evaluation via `ApproxTCE`. As we show in Theorem 12, this condition is naturally guaranteed when employing entropic OT techniques for speeding up the transition coupling improvement step.

5.3 Entropic Transition Coupling Improvement

Next we describe a means of speeding up Algorithm 1b. For the MDP corresponding to the entropic OTC problem, exact policy improvement can be performed by calling `ExactTCI` with $\Pi = \Pi^{\eta}_{\mathrm{TC}}(P,Q)$. However, no computation time is saved by doing this. Instead, we settle for an algorithm that yields approximately improved transition couplings with better computational efficiency. To find such an approximation, we reconsider the linear optimization problems that comprise the transition coupling improvement step. Namely, for each $(x,y) \in \mathcal{X} \times \mathcal{Y}$,

$$\begin{aligned} \text{minimize} \quad & \langle r, h \rangle \\ \text{subject to} \quad & r \in \Pi_{\eta}(P(x,\cdot), Q(y,\cdot)). \end{aligned} \tag{5}$$

By standard arguments, (5) is equivalent to

$$\begin{aligned} \text{minimize} \quad & \langle r, h \rangle + \frac{1}{\xi} \sum_{s'} r(s') \log r(s') \\ \text{subject to} \quad & r \in \Pi(P(x,\cdot), Q(y,\cdot)), \end{aligned} \tag{6}$$

for some $\xi \in [0,\infty]$ depending on (x,y), η and h. The reformulation (6) suggests that one use computational techniques for entropic OT in the place of linear programming to perform transition coupling improvement for the constrained OTC problem. In particular, we use the `ApproxOT` algorithm of (Altschuler et al., 2017), detailed in Appendix 9.4. Using `ApproxOT` instead of solving (6) exactly, we obtain the `EntropicTCI` algorithm detailed in Algorithm 2b.

Algorithm 2b: `EntropicTCI`

input : h, ξ, ε
for $(x,y) \in \mathcal{X} \times \mathcal{Y}$ **do**
  $R(s,\cdot) \leftarrow$ `ApproxOT`$(P(x,\cdot)^{\top}, Q(y,\cdot)^{\top}, h, \xi, \varepsilon)$
return R

To provide further intuition for Algorithm 2b, it is helpful to consider the constrained OTC problem from an alternate perspective. For a probability measure $r \in \Delta_{d^2}$, let $H(r) = -\sum_s r(s) \log r(s)$ be its entropy. Then by duality theory, the constrained OTC problem (4) may be written as the finite-dimensional optimization problem

$$\begin{aligned} \text{minimize} \quad & \langle c, \lambda \rangle - \sum_s \frac{1}{\xi(s)} H(R(s,\cdot)) \\ \text{subject to} \quad & R \in \Pi_{\mathrm{TC}}(P,Q) \\ & \lambda R = \lambda \\ & \langle \mathbb{1}, \lambda \rangle = 1, \end{aligned} \tag{7}$$

for some $\xi \in [0,\infty]^{d^2}$. In order to solve the problem above, we study its Lagrangian. Let $\alpha, \beta \in \mathbb{R}^{d^3}$, $\gamma \in \mathbb{R}^{d^2}$, and $\delta \in \mathbb{R}$ be Lagrange multipliers. The Lagrangian may be written as

$$\begin{aligned}\mathcal{L}(R,\lambda,\alpha,\beta,\gamma,\delta) = \langle c, \lambda\rangle - \sum_{x,y} \frac{1}{\xi(x,y)} H(R((x,y),\cdot)) \\ + \sum_{x,y,x'} \alpha(x,y,x')\left(\sum_{y'} R((x,y),(x',y')) - P(x,x')\right) \\ + \sum_{x,y,y'} \beta(x,y,y')\left(\sum_{x'} R((x,y),(x',y')) - Q(y,y')\right) \\ + \sum_{x',y'} \gamma(x',y')\left(\sum_{x,y} \lambda(x,y) R((x,y),(x',y')) - \lambda(x',y')\right) \\ + \delta\left(\sum_{x,y}\lambda(x,y) - 1\right).\end{aligned}$$

Taking the partial derivative of $\mathcal{L}$ with respect to $R((x,y),(x',y'))$ and setting it equal to zero, we find that

$$R(s,(x',y')) = \exp\left\{-\xi(s)\alpha(s,x') - \frac{1}{2}\right\}\exp\left\{-\xi(s)\lambda(s)\gamma(x',y')\right\}\exp\left\{-\xi(s)\beta(s,y') - \frac{1}{2}\right\},$$

where we have used $s = (x,y)$ to reduce notation. When viewed as a $d \times d$ matrix, $R((x,y),\cdot)$ can be written as UKV where U and V are both non-negative diagonal matrices. Note that when $\xi(x,y) < \infty$, this implies that R is aperiodic and irreducible since R lies in the relative interior of $\Pi_{\mathrm{TC}}(P,Q)$ (see Theorem 12).

A similar matrix form appears in the analysis of Cuturi (2013). An important difference is the matrix $K \in \mathbb{R}_+^{d\times d}$, which satisfies

$$K(x',y') = \exp\left\{-\xi(x,y)\lambda(x,y)\gamma(x',y')\right\}.$$

In Cuturi (2013), one finds that $K = e^{-\xi C}$ where C is the cost matrix. To better understand this difference, it is helpful to look at the partial derivative of the Lagrangian with respect to $\lambda(x,y)$. Evaluating this partial derivative and setting it equal to zero, we find

$$\gamma(x,y) = c(x,y) + \sum_{x',y'} R((x,y),(x',y'))\gamma(x',y') + \delta.$$

Absorbing the scalar δ into c to obtain an augmented cost $\tilde{c} = c + \delta$, we have

$$\gamma(x,y) = \tilde{c}(x,y) + \sum_{x',y'} R((x,y),(x',y'))\gamma(x',y').$$

Letting g be the gain of the policy R with respect to $\tilde{c}$, we recognize that the equation above is the Bellman recursion for the bias of R with respect to the cost $\tilde{c} + g$. As R is aperiodic and irreducible, g is a constant vector. Moreover, as the bias is invariant under constant shifts in cost, γ is exactly the bias h that appears in `EntropicTCI`. Returning to the form of $R((x,y),\cdot)$ established earlier, we find that

$$R((x,y),\cdot) = U \exp\{-\xi(x,y)\lambda(x,y)h\}\, V = U \exp\{-\tilde{\xi}(x,y)h\} V,$$

for non-negative diagonal matrices U and V and the constant $\tilde{\xi}(x,y) := \xi(x,y)\lambda(x,y)$. In this way, the bias h plays the role of the cost matrix C of Cuturi (2013).

In order to solve the program (7), one must grapple with the interdependence between the bias h and the policy R. A natural approach for doing so is to consider an alternating optimization algorithm in which one repeatedly solves for the bias h from a given policy R, then solves for a new policy R given the bias h. Indeed, this is the procedure one follows in `ExactOTC`. In practice, given a policy R, one approximately computes the bias h (`ApproxTCE`) in order to save time. Given a bias vector h, one solves for a new policy R by performing Sinkhorn iterations with the bias h as a cost matrix for each $R((x,y),\cdot)$ (`EntropicTCI`).

It was shown in Altschuler et al. (2017) that `ApproxOT` yields an approximation of the OT cost in near-linear time with respect to the size of the couplings of interest. However, in order to control the approximation error of `EntropicTCI`, we rely on a different analysis showing that one can obtain an approximation of the entropic optimal coupling in near-linear time (see Lemma 18) . To the best of our knowledge, this result does not exist in the literature, so we provide a proof in Section 9.4. Using this result, we show the complexity bound below.

Theorem 12 *Let P and Q be aperiodic and irreducible, $h \in \mathbb{R}^{d^2}$, $\xi > 0$, and $\varepsilon > 0$. Then* `EntropicTCI`(h, ξ, ε) *returns $\hat{R} \in \mathrm{ri}(\Pi_{TC}(P,Q))$ with $\max_s \|\hat{R}(s,\cdot) - R^*(s,\cdot)\|_1 \leq \varepsilon$ for some $R^* \in \mathrm{argmin}_{R' \in \Pi_{TC}(P,Q)} R'h - {}^1\!/\!_\xi H(R')$ in $\tilde{\mathcal{O}}(d^4\varepsilon^{-4})$ time.*

To summarize, this result states that `EntropicTCI` yields an approximately improved transition coupling in $\tilde{\mathcal{O}}(d^4)$ time rather than $\tilde{\mathcal{O}}(d^5)$ as previously discussed. In practice, further speedups are possible by utilizing the fact that the d^2 entropic OT problems to be solved are decoupled and thus may be computed in parallel.

5.4 `EntropicOTC`

Finally, using Algorithms 2a and 2b, we define the `EntropicOTC` algorithm, detailed in Algorithm 2. Essentially, `EntropicOTC` is defined by replacing `ExactTCE` and `ExactTCI` by the efficient alternatives, `ApproxTCE` and `EntropicTCI`. As stated in Theorem 12, `EntropicTCI` returns transition couplings in the relative interior of $\Pi_{\mathrm{TC}}(P,Q)$, so the iterates of `EntropicOTC` are not restricted to the finite set of extreme points of $\Pi_{\mathrm{TC}}(P,Q)$. Thus, convergence for Algorithm 2 must be assessed differently than in Algorithm 1. In our simulations we found that the element-wise inequality $\tilde{g}_{n+1} \geq \tilde{g}_n$ works well as an indicator of convergence.

Algorithm 2: `EntropicOTC`

```
input : L, T, ξ, ε
n ← 0
while n = 0 or g̃_{n+1} < g̃_n do
    /* transition coupling evaluation */
    (g̃_n, h̃_n) ← ApproxTCE(R_n, L, T)
    /* transition coupling improvement */
    R_{n+1} ← EntropicTCI(h̃_n, ξ, ε)
    n ← n + 1
return R_{n+1}
```

6. Consistency

The computational and theoretical results presented above assume that one has complete knowledge of the transition matrices P and Q of the Markov chains X and Y under study. In practice, one may not have direct access to P and Q, but may instead have estimates $\hat{P}_n$ and $\hat{Q}_n$ derived from n observations of the chains X and Y. In the simplest case, $\hat{P}_n$ and $\hat{Q}_n$ may be obtained from the observed relative frequencies of each transition between states. In Theorem 13 below, we show that the cost and solution sets of the standard and regularized optimal transition coupling problems possess natural stability properties with respect to the marginal transition matrices. As a corollary, we obtain a consistency result for the OTC problem applied to the estimates $\hat{P}_n$ and $\hat{Q}_n$.

Recall that we use Δ_d to denote the probability simplex in $\mathbb{R}^d$ and note that the set of $d \times d$-dimensional transition matrices may be written as Δ_d^d. Likewise, the set of $d^2 \times d^2$-dimensional transition matrices may be written as $\Delta_{d^2}^{d^2}$. Note that we endow the sets of $d \times d$- and $d^2 \times d^2$-dimensional transition matrices with the topologies they inherit as subsets of $\mathbb{R}^{d\times d}$ and $\mathbb{R}^{d^2\times d^2}$, respectively, and adopt the same convention for the set $\Delta_{d^2} \times \Delta_{d^2}^{d^2}$. Now, we may reformulate Problems (2) and (4) as follows:

$$\begin{aligned} \text{minimize} \quad & \langle c, \lambda \rangle \\ \text{subject to} \quad & R \in \Pi(P, Q) \\ & \lambda R = \lambda \\ & \lambda \in \Delta_{d^2}. \end{aligned} \tag{I}$$

$$\begin{aligned} \text{minimize} \quad & \langle c, \lambda \rangle \\ \text{subject to} \quad & R \in \Pi_\eta(P, Q) \\ & \lambda R = \lambda \\ & \lambda \in \Delta_{d^2}. \end{aligned} \tag{II}$$

Let $\rho(P, Q)$ and $\rho_\eta(P, Q)$ denote the optimal values of Problems (I) and (II), respectively, and let $\Phi^*(P, Q)$ and $\Phi_\eta^*(P, Q)$ denote the associated sets of optimal solutions $(\lambda, R) \in \Delta_{d^2} \times \Delta_{d^2}^{d^2}$ to Problems (I) and (II), respectively. For metric spaces $\mathcal{U}$ and $\mathcal{Z}$, we will say that a function $F : \mathcal{U} \to 2^{\mathcal{Z}}$ is upper semicontinuous at a point $u_0 \in \mathcal{U}$ if for any neighborhood V of $F(u_0)$, there exists a neighborhood U of u_0 such that $F(u) \subset V$ for every $u \in U$.

Theorem 13 *Let $P, Q \in \Delta_d^d$ be irreducible transition matrices. Then the following hold:*

- *$\rho(\cdot, \cdot)$ is continuous and $\Phi^*(\cdot, \cdot)$ is upper semicontinuous at (P, Q)*
- *For any $\eta > 0$, $\rho_\eta(\cdot, \cdot)$ is continuous and $\Phi_\eta^*(\cdot, \cdot)$ is upper semicontinuous at (P, Q)*

Theorem 13 states that the optimal values and optimal solution sets of the OTC and entropic OTC problems are stable in the marginal transition matrices P and Q. We may use this result to prove a consistency result for either problem when applied to estimates $\hat{P}_n$ and $\hat{Q}_n$ derived from data. In stating the following result, we make use of the following definition: For a sequence of sets $\{A_n\}_{n\geq 0}$ in a topological space $\mathcal{A}$, let $\limsup_{n\to\infty} A_n = \bigcap_{n=0}^{\infty} \mathrm{cl}\,(\bigcup_{m=n}^{\infty} A_m)$, where $\mathrm{cl}(\cdot)$ denotes the closure with respect to topology of $\mathcal{A}$. Note that the presence of $\mathrm{cl}(\cdot)$ in our definition of limit superior of a sequence of sets differs from that commonly used in probability but is consistent with the definition appearing, for example, in Rockafellar and Wets (2009).

Corollary 14 *Let $X = \{X_i\}_{i\geq 0}$ and $Y = \{Y_i\}_{i\geq 0}$ be stationary, ergodic processes taking values in $\mathcal{X}$ and $\mathcal{Y}$, and defined on a common Borel probability space. Suppose further that X and Y have marginal, one-step transition matrices P and Q, respectively. Let $\hat{P}_n$ and $\hat{Q}_n$ be the one-step transition matrices estimated via relative frequencies from the sequences $X_0, ..., X_{n-1}$ and $Y_0, ..., Y_{n-1}$. Then with probability one, the following hold:*

- $\rho(\hat{P}_n, \hat{Q}_n) \to \rho(P, Q)$ *and* $\limsup_{n\to\infty} \Phi^*(\hat{P}_n, \hat{Q}_n) \subseteq \Phi^*(P, Q)$
- *For any $\eta > 0$,* $\rho_\eta(\hat{P}_n, \hat{Q}_n) \to \rho_\eta(P, Q)$ *and* $\limsup_{n\to\infty} \Phi^*_\eta(\hat{P}_n, \hat{Q}_n) \subseteq \Phi^*_\eta(P, Q)$

Corollary 14 allows us to apply the computational tools described above to real data in a principled manner. In particular, when the marginal transition matrices P and Q are unknown, we may use $\hat{P}_n$ and $\hat{Q}_n$ as proxies in the OTC problem to estimate the set of optimal transition couplings and their expected cost when n is large. Note that we do not require the generating processes themselves to be Markov: they need only be stationary and ergodic, so that the estimates $\hat{P}_n$ and $\hat{Q}_n$ converge to the true one-step transition matrices P and Q as n tends to infinity.

7. Experiments

In this section, we validate the proposed algorithms empirically by applying them to stationary Markov chains derived from both synthetic and real data. We begin by comparing the runtime of the proposed algorithms and approximation error of `EntropicOTC` via a simulation study. Subsequently, we illustrate the potential use of the OTC problem in practice through an application to computer-generated music.

We remark that an application of the OTC problem to graphs is studied in O'Connor et al. (2021b). In particular, a weighted graph may be associated with a stationary Markov chain by means of a simple random walk on its nodes with transition probabilities proportional to its edge weights. Leveraging this perspective, we propose to perform OT on the graphs of interest by applying the OTC problem to their associated Markov chains. In the aforementioned work, we demonstrate that this approach performs on par with state-of-the-art graph OT methods in a variety of graph comparison and alignment tasks on real and synthetic data.

`Matlab` implementations of `ExactOTC` and `EntropicOTC` as well as code for reproducing the experimental results to follow are available at `https://github.com/oconnor-kevin/OTC`. For `ApproxOT` and related OT algorithms, we used the implementation found at `https://github.com/JasonAltschuler/OptimalTransportNIPS17`.

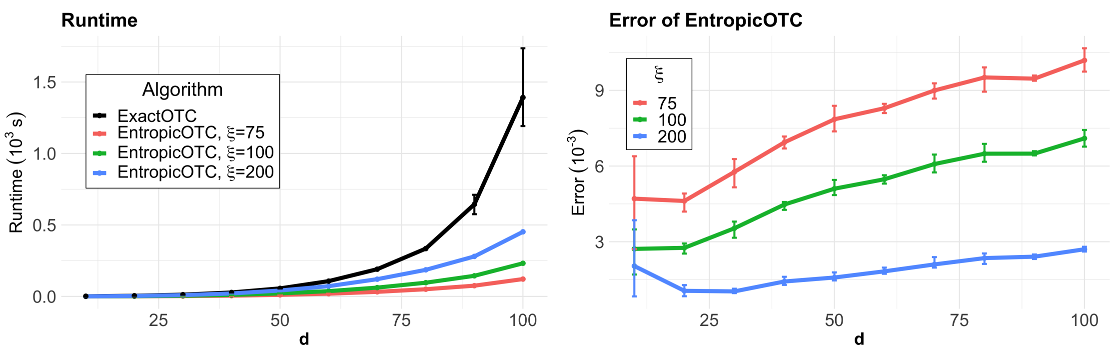


Figure 1: A comparison of total runtimes between `ExactOTC` and `EntropicOTC` and approximation errors of `EntropicOTC` for a range of d and ξ via simulation. Error bars show the minimum and maximum values observed over five simulations. Note that the error bars for the runtimes of `EntropicOTC` are not visible because little variation in runtime was observed over the simulations performed. Runtime is reported in units of 10^3 seconds while error is reported in units of 10^{-3} relative to the maximum value of the cost function c.

7.1 Simulation Study

In order to validate the use of Algorithm 2 as a fast alternative to Algorithm 1, we performed a simulation study to compare their runtimes and the error of the entropic OTC cost as an approximation of the OTC cost. For each choice of the marginal state space size $d \in \{10, 20, ..., 100\}$, we perform five simulations, obtaining estimates of the runtimes and approximation error in each. In each simulation, we generate transition matrices $P \in [0,1]^{d\times d}$ and $Q \in [0,1]^{d\times d}$ and a cost matrix $c \in \mathbb{R}_+^{d\times d}$ by drawing each element of the matrix of interest independently from a standard normal distribution and then applying an appropriate normalization to the matrix. In the case of the transition matrices, we apply a softmax normalization with weight 0.1 to each row of P and Q:

$$P(x,x') \mapsto \frac{e^{0.1P(x,x')}}{\sum_{\tilde{x}} e^{0.1P(x,\tilde{x})}}, \qquad Q(y,y') \mapsto \frac{e^{0.1Q(y,y')}}{\sum_{\tilde{y}} e^{0.1Q(y,\tilde{y})}}.$$

For the cost matrix, we apply an absolute value element-wise so that $c \in \mathbb{R}_+^{d\times d}$ and then divide each element by the maximum element in the matrix so that $\|c\|_\infty = 1$. After generating both the transition matrices and cost matrix, we run both `ExactOTC` and `EntropicOTC` for each $\xi \in \{75, 100, 200\}$ until convergence. In all runs of `EntropicOTC`, we choose L and T adaptively as described in Remark 10 with tolerance (ε) equal to 10^{-12} and upper bounds of 100 and 1000, respectively. For each choice of $\xi \in \{75, 100, 200\}$, we use 50, 100, and 200 Sinkhorn iterations, respectively. Runtimes of `ExactOTC` and `EntropicOTC` in a given iteration are measured from the start to convergence and thus correspond to *total* runtime rather than the runtime of individual iterations. The approximation error of `EntropicOTC` in a given iteration is measured by taking the absolute difference between the expected cost returned by `EntropicOTC` and that returned by `ExactOTC`. Note that after randomization, the cost function c is scaled to $\|c\|_\infty = 1$ and the error is reported on that scale.

The results of the simulation study are shown in Figure 1. The error bars in either plot denote the maximum and minimum values observed for each choice of parameters over the five repeated simulations. In our simulations, we found that the time savings in each iteration of `EntropicOTC` resulted in substantial time savings over the entire runtime of the algorithm without substantial loss of accuracy. For example, when $d = 100$ and $\xi = 100$, we observed that `EntropicOTC` yielded a time savings of roughly 80% compared to `ExactOTC`. Moreover, weakening the regularization by increasing ξ reduces the error of `EntropicOTC` with little additional runtime. This supports our theoretical findings, indicating that `EntropicOTC` is a good alternative to `ExactOTC` when d is large.

We remark that running this simulation experiment for values of d beyond 100 caused some of the matrices in the computing environment to exceed the available memory on the machine we were using. This occurs when one tries to represent a transition coupling as a $d^2 \times d^2$-dimensional array in memory. It is likely that such barriers may be overcome in some cases by making use of more sophisticated techniques for working with large, sparse matrices.

7.2 Application to Computer-Generated Music

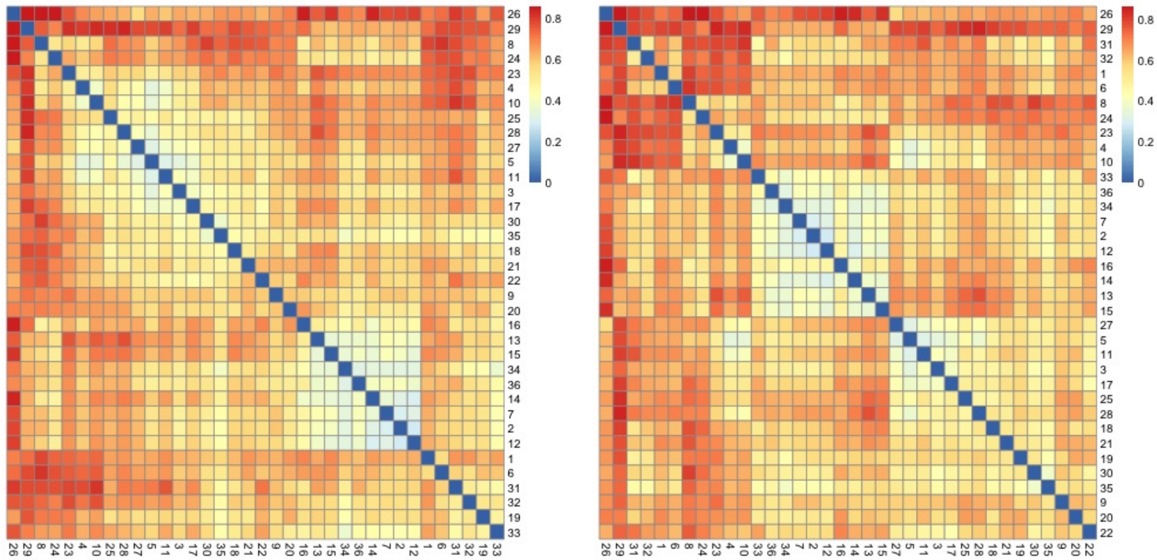
Next we illustrate the OTC problem in practice through an application to aligning and comparing computer-generated music. HMMs and other state-space models have been explored as a tool for modeling musical arrangements Ames (1989); Liu and Selfridge-Field (2002); Weiland et al. (2005); Allan and Williams (2005); Pikrakis et al. (2006); Ren et al. (2010); Bell (2011); Yanchenko and Mukherjee (2017); Das et al. (2018). In this line of work, sequences of notes are commonly modeled as a stationary processes with latent Markovian structure. As described in Section 2, the OTC problem easily extends to this setting, allowing one to apply OT methods to analyzing generative models for music. We apply the computational tools developed above to two tasks: comparing pieces based on the sequences of notes they contain and generating paired sequences of notes based on existing pieces.

We analyzed a data set of 36 pieces of classical music from 3 different classical composers (Bach, Beethoven and Mozart) downloaded from https://www.mfiles.co.uk/classical-midi.htm. The pieces considered along with the composer, musical key, and reference number between 1 and 36 may be found in Table 1. For each piece, a 3-layer HMMs with 5 hidden states was trained using the code provided in Yanchenko and Mukherjee (2017). We refer the reader to Yanchenko and Mukherjee (2017) and Oliver et al. (2004) for details on layered HMMs but note that once a layered HMM is trained it may be recast as a standard HMM and thus the extension of OTC to HMMs described in Section 3 still applies. We considered two different cost functions between notes. The first cost function equal to 0 if the two notes are equal or some number of octaves (intervals of 12 semitones) apart, and 1 otherwise. The second cost function is 0 when the first cost function is 0, 1 when the two notes are 5 or 7 semitones apart (perfect consonance), 2 when the two notes are 4 or 9 semitones apart (imperfect consonance) and 10 otherwise. This tiered cost function incorporates a preference for unison over perfect consonance, perfect consonance over imperfect consonance, and imperfect consonance over dissonance.

In the first task, we computed the OTC cost for every pair of pieces, obtaining a pairwise cost matrix. Note that when running `EntropicOTC`, we use $L = 100$, $T = 1000$, $\xi = 50$,

	Composer	Piece	Key
1	Bach	Toccata and Fugue	D minor
2	Bach	Book 1, Fugue 2	C minor
3	Bach	Book 1, Fugue 10	E minor
4	Bach	Book 1, Fugue 14	F# minor
5	Bach	Book 1, Fugue 24	B minor
6	Bach	Book 1, Prelude 1	C major
7	Bach	Book 1, Prelude 2	C minor
8	Bach	Book 1, Prelude 3	C# major
9	Bach	Book 1, Prelude 6	D minor
10	Bach	Book 1, Prelude 14	F# minor
11	Bach	Book 1, Prelude 24	B minor
12	Bach	Book 2, Fugue 2	C minor
13	Bach	Book 2, Fugue 7	D# major
14	Bach	Book 2, Prelude 2	C minor
15	Bach	Book 2, Prelude 7	D# major
16	Bach	Book 2, Prelude 12	F minor
17	Bach	Bourrée in E minor	E minor
18	Bach	2 Part Invention, No. 13	A minor
19	Bach	2 Part Invention, No. 4	D minor
20	Bach	Prelude in C major	C major
21	Beethoven	Für Elise	A minor
22	Beethoven	Minuet in G	G major
23	Beethoven	Moonlight Sonata, Movement 1	C# minor
24	Beethoven	Sonata Pathétique, Movement 2	C minor
25	Beethoven	Symphony No. 7, Movement 2	A minor
26	Beethoven	Symphony No. 9, Movement 4	D minor
27	Beethoven	Violin Sonata 1, Movement 1	D major
28	Mozart	Piano Sonata No. 11, Movement 3	A major
29	Mozart	Horn Concerto 4, Movement 3	D# major
30	Mozart	Minuet and Trio, K.1	G major
31	Mozart	Minuet in F major, K.2	F major
32	Mozart	Österreichische Bundeshymne	D# major
33	Mozart	Piano Concerto No. 21, Movement 2	C major
34	Mozart	Piano Sonata No. 13, Movement 1	A# major
35	Mozart	Piano Sonata No. 16	C major
36	Mozart	Symphony No. 40, Movement 1	G minor

Table 1: Pieces considered in the application of OTC to computer-generated music.



(a) `ExactOTC` (b) `EntropicOTC`

Figure 2: Heatmap of costs for all pairs of pieces as computed by `ExactOTC` and `EntropicOTC`. Lower cost (indicated by blue) indicates a better correspondence between the two pieces. The list of pieces and composers considered may be found in Table 1.

and 20 Sinkhorn iterations. The cost matrices obtained using `ExactOTC` and `EntropicOTC` are both depicted in Figure 2. The correspondence between rows and columns of the two heatmaps and the musical pieces considered can be found in Table 1. We remark that pieces in the same key tended to have lower OTC cost. For example, Bach's Fugue 2 from Book 1 (2 in Figure 2) and Fugue 2 from Book 2 (12 in Figure 2), both in C minor, had the lowest OTC and entropic OTC costs among all pairs considered. We observe that pairwise costs obtained by either algorithm only differ by 8×10^{-3} on average. In other words, `EntropicOTC` approximates the result of `ExactOTC` with high accuracy.

In the second task, we explored the samples generated from the optimal transition coupling of each pair of fitted HMMs. The optimal transition coupling maximizes the probability of generating consonant pairs of notes while preserving the distributions of the two sequences. This results in sequences that sound harmonious together more frequently. In Figure 3, we provide a paired sequence drawn from the output of `ExactOTC` applied to pieces from Bach and Beethoven. Note that no dissonant pairs of notes were sampled in this sequence. Audio files for this sequence and sequences drawn from other pairings may be found in the accompanying supplemental materials.

8. Discussion

In this paper, we introduced an optimal transport problem for stationary Markov chains that takes the Markovian dynamics into account called the optimal transition coupling (OTC) problem. Intuitively, the OTC problem aims to synchronize the Markov chains of interest so as to minimize long-term average cost. We demonstrated how this problem may

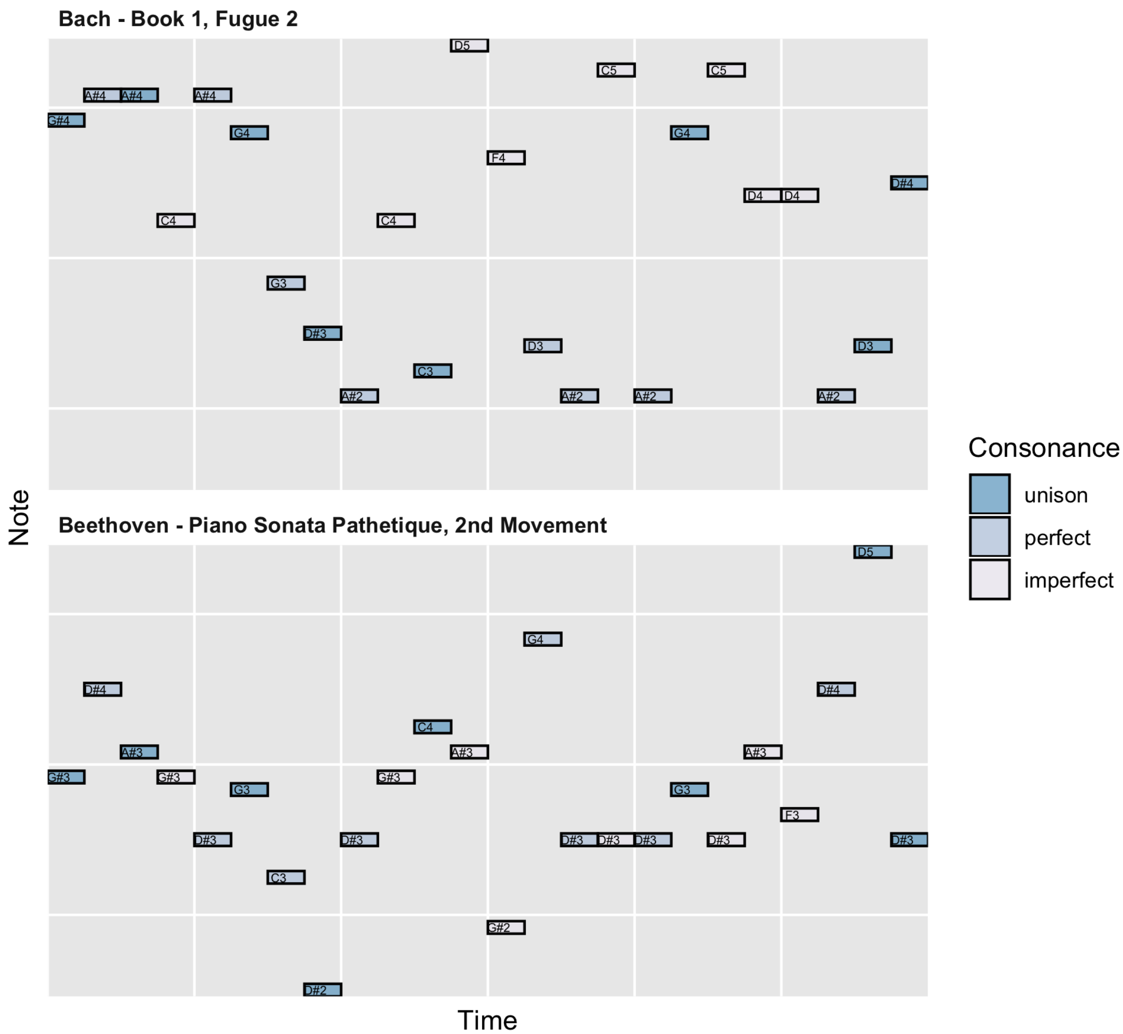


Figure 3: An illustration of samples drawn from an optimal transition coupling of Bach's Book 1, Fugue 2 and Beethoven's Sonata Pathétique, Movement 2, both in C minor. The color of each sampled note denotes its consonance with the other note played at the same time.

be easily extended to formulate an OT problem for HMMs. In the interest of computation, we recast this problem as a Markov decision process and leveraged this connection to prove that solutions can be obtained via an adaptation of the policy iteration algorithm, referred to as `ExactOTC`. Mirroring the development of entropic OT in Cuturi (2013), we also proposed an entropic OTC problem and an associated approximate algorithm, `EntropicOTC`, which scales better with dimension. For cases when the marginal Markov chains must be estimated from data, we showed that the plug-in estimates for either problem are consistent. We showed empirically that `EntropicOTC` approximates the OTC cost with high accuracy and substantially faster runtime than `ExactOTC` in large state space regimes. Finally, we illustrated the use of the OTC problem and the proposed algorithms in practice via an application to computer-generated music.

Future work may consider extending the ideas of the OTC problem to processes with more flexible structure such as Gibbs processes or dynamical linear models. We expect that the extension of our work to processes with richer temporal structure will present interesting computational challenges. Alternatively, future work may explore further applications of the OTC problem in practice. Our approach to analyzing computer-generated music may be easily transferred to any data that may be modeled by an HMM. HMMs and other sequence models with hidden Markov structure are commonly used in a variety of fields including genomics, speech recognition, protein folding, and natural language processing.

9. Proofs

In what follows, we detail the proofs of our results.

9.1 Preliminaries

We begin by introducing some additional notation, covering some preliminaries on Markov chains, and remarking on some technical aspects relating to our results.

9.1.1 Additional Notation

We adopt the following additional notation: For a finite set $\mathcal{U} \subset \mathbb{R}$, we define $\min_{>0} \mathcal{U} = \min\{u \in \mathcal{U} : u > 0\}$. We define the inner product $\langle \cdot, \cdot \rangle$ for matrices $U, V \in \mathbb{R}^{n\times n}$ by $\langle U, V \rangle := \sum_{i,j} U_{ij} V_{ij}$. All vector and matrix equations and inequalities should be understood to hold element-wise. For $i \leq j$, we let $u_i^j = (u_i, ..., u_j)$ and we will denote infinite sequences by boldface, lowercase letters such as $\mathbf{u} = (u_0, u_1, ...)$. For a collection of sets $\mathcal{U}_s \subset \mathbb{R}^{d^2}$ indexed by $s \in \mathcal{X} \times \mathcal{Y}$, we define $\bigotimes_s \mathcal{U}_s$ to be the set of matrices $U \in \mathbb{R}^{d^2 \times d^2}$ such that for every $s \in \mathcal{X} \times \mathcal{Y}$, $U(s, \cdot) \in \mathcal{U}_s$. In particular, we write $\Pi_{\text{TC}}(P, Q) = \bigotimes_{(x,y)} \Pi(P(x, \cdot), Q(y, \cdot))$.

9.1.2 Preliminaries on Markov Chains

For a finite metric space $\mathcal{U}$, we say that a measure $\mu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}})$ is *Markov* or *corresponds to a Markov chain taking values in* $\mathcal{U}$ if for any cylinder set $[u_0 \cdots u_k] \subset \mathcal{U}^{\mathbb{N}}$, $\mu([u_0 \cdots u_k])/\mu([u_0 \cdots u_{k-1}]) = \mu([u_{k-1} u_k])/\mu([u_{k-1}])$, where we let $0/0 = 0$. We say that μ is *stationary* if $\mu = \mu \circ \sigma^{-1}$, where $\sigma : \mathcal{U}^{\mathbb{N}} \to \mathcal{U}^{\mathbb{N}}$ is the left-shift map defined such that for any $\mathbf{u} \in \mathcal{U}^{\mathbb{N}}$, $\sigma(\mathbf{u})_i = u_{i+1}$. When $\mathcal{U}$ has cardinality $n \geq 1$, we define the transition matrix $U \in \mathbb{R}^{n\times n}$ of μ such that for every $u_{k-1}, u_k \in \mathcal{U}$, $U(u_{k-1}, u_k) = \mu([u_{k-1} u_k])/\mu([u_{k-1}])$. If μ is also stationary, its stationary distribution $\lambda_U \in \Delta_n$ is defined such that $\lambda_U(u) = \mu([u])$ for any $u \in \mathcal{U}$. We say that μ or U is *irreducible* if for every $u, u' \in \mathcal{U}$, there exists $k \geq 1$, possibly depending on u and u', such that $U^k(u, u') > 0$. We call μ or U *aperiodic* if $\gcd\{k \geq 1 : U^t(u, u') > 0\} = 1$ for every $u, u' \in \mathcal{U}$. Note that if μ is irreducible, its stationary distribution λ_U is unique. Furthermore, if μ is also aperiodic, there exists $M < \infty$ and $\alpha \in (0, 1)$ such that for any $t \geq 1$, $\max_u \|U^t(u, \cdot) - \lambda_U\|_1 \leq M\alpha^t$. For more details on basic Markov chain theory, we refer the reader to Levin and Peres (2017).

9.1.3 Technical Considerations

We endow the finite set $\mathcal{X}\times\mathcal{Y}$ with the discrete topology and $\mathcal{X}^{\mathbb{N}}\times\mathcal{Y}^{\mathbb{N}}$ with the corresponding product topology. For each $(x,y)\in\mathcal{X}\times\mathcal{Y}$ and $\eta>0$, we endow both $\Pi(P(x,\cdot),Q(y,\cdot))$ and $\Pi_\eta(P(x,\cdot),Q(y,\cdot))$ with the subspace topology inherited from the Euclidean topology on $\mathbb{R}^{d^2}$. Similarly, we endow $\Pi_{\text{TC}}(P,Q)$ and $\Pi^{\eta}_{\text{TC}}(P,Q)$ with the subspace topologies inherited from the Euclidean topology on $\mathbb{R}^{d^2\times d^2}$. Unless stated otherwise, continuity of any function will be understood to mean with respect to the corresponding topology above.

9.2 Proofs from Section 2

Proposition 4 *Let X and Y be irreducible stationary Markov chains with transition matrices P and Q, respectively. Then any stationary Markov chain with a transition matrix contained in $\Pi_{\text{TC}}(P,Q)$ is a transition coupling of X and Y.*

Proof Let $\pi\in\mathcal{M}((\mathcal{X}\times\mathcal{Y})^{\mathbb{N}})$ be the distribution of a stationary Markov chain with transition matrix $R\in\Pi_{\text{TC}}(P,Q)$ and stationary distribution $r\in\Delta_{d^2}$. Furthermore, let $r_{\mathcal{X}}$ and $r_{\mathcal{Y}}\in\Delta_d$ be the $\mathcal{X}$ and $\mathcal{Y}$ marginals of r, respectively. For a metric space $\mathcal{U}$ and a probability measure $\mu\in\mathcal{M}(\mathcal{U}^{\mathbb{N}})$, we define $\mu_k\in\mathcal{M}(\mathcal{U}^k)$ as the k-dimensional marginal distribution of μ. Formally, for any cylinder set $[a_0^{k-1}]=\{\mathbf{u}\in\mathcal{U}^{\mathbb{N}}:u_j=a_j,0\le j\le k-1\}$, $\mu_k(a_0^{k-1}):=\mu([a_0^{k-1}])$.

We wish to show that $\pi\in\Pi_{\text{TC}}(\mathbb{P},\mathbb{Q})$. Since π corresponds to a stationary Markov chain and $R\in\Pi_{\text{TC}}(P,Q)$ by assumption, it suffices to show that $\pi\in\Pi(\mathbb{P},\mathbb{Q})$. We will do this by showing that $\pi_k\in\Pi(\mathbb{P}_k,\mathbb{Q}_k)$ for every $k\ge1$. Starting with $k=1$, for any $y\in\mathcal{Y}$,

$$\begin{aligned}
r_{\mathcal{Y}}(y) &= \sum_x r(x,y)\\
&= \sum_x\sum_{x',y'} r(x',y')R((x',y'),(x,y))\\
&= \sum_{x',y'} r(x',y')\sum_x R((x',y'),(x,y))\\
&= \sum_{x',y'} r(x',y')Q(y',y)\\
&= \sum_{y'} r_{\mathcal{Y}}(y')Q(y',y).
\end{aligned}$$

We have proven that $r_{\mathcal{Y}}$ is invariant with respect to Q. Since Q is irreducible, the stationary distribution q of Q is unique. Thus, $r_{\mathcal{Y}}=q$. A similar argument will show that $r_{\mathcal{X}}=p$. Thus, $r\in\Pi(p,q)$ and therefore, $\pi_1\in\Pi(\mathbb{P}_1,\mathbb{Q}_1)$.

Now suppose that $\pi_k\in\Pi(\mathbb{P}_k,\mathbb{Q}_k)$ for some $k\ge1$. Fixing $y_0^k\in\mathcal{Y}^{k+1}$, it follows that

$$\begin{aligned}
\sum_{x_0^k}\pi_{k+1}(x_0^k,y_0^k) &= \sum_{x_0^k}\pi_k(x_0^{k-1},y_0^{k-1})R((x_{k-1},y_{k-1}),(x_k,y_k))\\
&= \sum_{x_0^{k-1}}\pi_k(x_0^{k-1},y_0^{k-1})Q(y_{k-1},y_k)
\end{aligned}$$

$$= \mathbb{Q}_k(y_0^{k-1}) Q(y_{k-1}, y_k)$$
$$= \mathbb{Q}_{k+1}(y_0^k).$$

Again the proof for the other marginal is identical. So we find that $\pi_{k+1} \in \Pi(\mathbb{P}_{k+1}, \mathbb{Q}_{k+1})$ and since $k \geq 1$ was arbitrary, we conclude that $\pi \in \Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$. ■

9.3 Proofs from Section 4

Next we prove our results from Section 4 regarding the correspondence between TC-MDP and the OTC problem and the convergence of `ExactOTC`.

9.3.1 Existence of a Deterministic Policy

Proposition 5 *Let γ be a policy for TC-MDP. Then there exists a deterministic policy $\tilde{\gamma}$ such that $\bar{c}_\gamma(s) = \bar{c}_{\tilde{\gamma}}(s)$ for every $s \in \mathcal{S}$.*

Proof Before proving the result, it will be helpful to fix some additional notation. Let $\gamma = \{\gamma_s(\cdot) : s \in \mathcal{X} \times \mathcal{Y}\}$ be a policy for TC-MDP. Recall that for each $s = (x, y)$, $\gamma_s(\cdot)$ describes a distribution on $\mathcal{A}_s = \Pi(P(x, \cdot), Q(y, \cdot))$. Define the deterministic policy $\tilde{\gamma} = \{\tilde{\gamma}_s(\cdot) : s \in \mathcal{X} \times \mathcal{Y}\}$ such that for every s, $\tilde{\gamma}_s(\cdot)$ assigns probability one to

$$\tilde{r}_s := \int_{\mathcal{A}_s} r_s \gamma_s(dr_s).$$

Here, $\tilde{r}_s$ is the expected action taken by the agent while occupying a state s and following the policy γ. Note that $\tilde{r}_s \in \mathcal{A}_s$ due to the convexity of $\mathcal{A}_s$. As such, we may collect the row vectors $\{\tilde{r}_s : s \in \mathcal{X} \times \mathcal{Y}\}$ into a single transition matrix $\tilde{R} \in \Pi_{\mathrm{TC}}(P, Q)$ where $\tilde{R}(s, \cdot) = \tilde{r}_s(\cdot)$ for every $s \in \mathcal{X} \times \mathcal{Y}$. In what follows, let $\mathrm{Prob}_\gamma(\cdot|s_0)$ and $\mathrm{Prob}_{\tilde{\gamma}}(\cdot|s_0) \in \mathcal{M}(\{\mathcal{A} \times (\mathcal{X} \times \mathcal{Y})\}^{\mathbb{N}})$ be the probability measures corresponding to the action-state processes with initial state s_0 induced by γ and $\tilde{\gamma}$, respectively. In particular,

$$\mathrm{Prob}_\gamma(dr_{s_0}, s_1, ..., dr_{s_{t-1}}, s_t|s_0) = \gamma_{s_0}(dr_{s_0}) r_{s_0}(s_1) \cdots \gamma_{s_{t-1}}(dr_{s_{t-1}}) r_{s_{t-1}}(s_t)$$

and the analogous statement holds for $\mathrm{Prob}_{\tilde{\gamma}}(\cdot|s_0)$. In the case of $\tilde{\gamma}$, one may also show that $\mathrm{Prob}_{\tilde{\gamma}}(s_t|s_0) = \tilde{R}^t(s_0, s_t)$. Finally, let $\mathbb{E}_\gamma[\cdot|s_0]$ and $\mathbb{E}_{\tilde{\gamma}}[\cdot|s_0]$ denote expectation with respect to $\mathrm{Prob}_\gamma(\cdot|s_0)$ and $\mathrm{Prob}_{\tilde{\gamma}}(\cdot|s_0)$, respectively.

Now, we can prove the result. For any $s_0 \in \mathcal{X} \times \mathcal{Y}$ and $t \geq 1$,

$$\begin{aligned}
\mathbb{E}_\gamma\left[c(s_t)|s_0\right] &= \sum_{s_t} c(s_t) \,\mathrm{Prob}_\gamma(s_t|s_0) \\
&= \sum_{s_t} c(s_t) \int_{\mathcal{A}_{s_0}} \sum_{s_1} \cdots \int_{\mathcal{A}_{s_{t-1}}} \mathrm{Prob}_\gamma(dr_{s_0}, s_1, ..., dr_{s_{t-1}}, s_t|s_0) \\
&= \sum_{s_t} c(s_t) \int_{\mathcal{A}_{s_0}} \sum_{s_1} \cdots \int_{\mathcal{A}_{s_{t-1}}} \gamma_{s_0}(dr_{s_0})\, r_{s_0}(s_1) \cdots \gamma_{s_{t-1}}(dr_{s_{t-1}})\, r_{s_{t-1}}(s_t) \\
&= \sum_{s_1^t} c(s_t) \int_{\mathcal{A}_{s_0}} \cdots \int_{\mathcal{A}_{s_{t-1}}} \gamma_{s_0}(dr_s)\, r_{s_0}(s_1) \cdots \gamma_{s_{t-1}}(dr_{s_{t-1}})\, r_{s_{t-1}}(s_t)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{s_1^t} c(s_t)\tilde{r}_{s_0}(s_1)\cdots\tilde{r}_{s_{t-1}}(s_t) \\
&= \sum_{s_1^t} c(s_t)\tilde{R}(s_0,s_1)\cdots\tilde{R}(s_{t-1},s_t) \\
&= \sum_{s_t} c(s_t)\tilde{R}^t(s_0,s_t) \\
&= \sum_{s_t} c(s_t)\,\mathrm{Prob}_{\tilde{\gamma}}(s_t|s_0) \\
&= \mathbb{E}_{\tilde{\gamma}}\left[c(s_t)|s_0\right].
\end{aligned}
$$

Thus, for every $s \in \mathcal{X}\times\mathcal{Y}$,

$$
\bar{c}_\gamma(s) = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_\gamma\left[c(s_t)|s_0=s\right] = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{\tilde{\gamma}}\left[c(s_t)|s_0=s\right] = \bar{c}_{\tilde{\gamma}}(s).
$$

■

9.3.2 Correspondence Between TC-MDP and the OTC Problem

Next, we prove Proposition 6 showing that optimal solutions to TC-MDP necessarily provide optimal solutions to the OTC problem. We rely on the basic idea of recurrent classes of states for finite-state Markov chains. For details on recurrence for Markov chains, we refer the reader to Levin and Peres (2017). For any $R \in \Pi_{\mathrm{TC}}(P,Q)$, let $\Lambda(R) := \{\lambda \in \mathcal{M}(\mathcal{X}) : \lambda R = \lambda\}$ denote the set of stationary distributions for R and let $\bigsqcup$ denote a disjoint union. Before proving the proposition, we require a lemma stating that for a given transition coupling matrix $R \in \Pi_{\mathrm{TC}}(P,Q)$, the stationary distribution of R that incurs the least expected cost may be chosen to be the unique stationary distribution of one of R's recurrent classes.

Lemma 15 *Let $R \in \Pi_{TC}(P,Q)$ and let $\mathcal{S}_r$ be the set of states belonging to some recurrent class of R. Moreover, for every $s \in \mathcal{S}_r$, let $\lambda_{R,s} \in \Lambda(R)$ denote the stationary distribution of R corresponding to the recurrent class in which s lies. Then $\lambda_{R,s}$ is uniquely defined and*

$$
\min_{s\in\mathcal{S}_r}\langle c, \lambda_{R,s}\rangle = \min_{\lambda\in\Lambda(R)}\langle c,\lambda\rangle.
$$

Proof The uniqueness of $\lambda_{R,s}$ follows from the fact that the chain obtained by restricting R to the recurrent class of s is necessarily irreducible. Now suppose that R has m recurrent classes $\{S_r^i\}_{i=1}^m$ and thus $\mathcal{S}_r = \bigsqcup_{i=1}^m S_r^i$. Then by (Puterman, 2005, Theorem A.5), there exist m linearly independent stationary distributions of R. Note that necessarily, the unique stationary distributions $\{\lambda_i\}_{i=1}^m$ corresponding to the m recurrent classes of R are linearly independent and constitute such a choice. Moreover, it is straightforward to show that $\Lambda(R)$ is equal to the convex hull of $\{\lambda_i\}_{i=1}^m$ and is thus compact. Then since minima of linear functions over a compact, convex set occur at the extreme points of the feasible set,

$$
\min_{s\in\mathcal{S}_r}\langle c,\lambda_{R,s}\rangle = \min_{i=1,\ldots,m}\langle c,\lambda_i\rangle = \min_{\lambda\in\Lambda(R)}\langle c,\lambda\rangle.
$$

■

Proposition 6 *If X and Y are irreducible, then any $R \in \Pi(P,Q)$ that is an optimal policy for TC-MDP corresponds to an optimal coupling $\pi_R \in \Pi_{TC}(\mathbb{P},\mathbb{Q})$ with expected cost $\min_{s\in\mathcal{S}} \bar{c}_R(s)$.*

Proof For every $R \in \Pi_{\text{TC}}(P,Q)$ and $s \in \mathcal{S}$, let $\lambda_{R,s} \in \Lambda(R)$ be the stationary distribution of R defined by

$$\lambda_{R,s} := \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} R^t(s,\cdot).$$

Note that $\lambda_{R,s}$ is well-defined by (Puterman, 2005, Theorem A.5). Moreover, we will use $\mathcal{S}_r(R)$ to refer to the set of all states in $\mathcal{S}$ that belong to a recurrent class of R. Since the space $\mathcal{S}$ is finite, $\mathcal{S}_r(R)$ is necessarily non-empty for every $R \in \Pi_{\text{TC}}(P,Q)$. Finally, note that whenever $s \in \mathcal{S}_r(R)$, $\lambda_{R,s}$ is the unique stationary distribution of R associated with the recurrent class in which s lies.

Now let $R^* \in \Pi_{\text{TC}}(P,Q)$ be optimal for TC-MDP. We will construct a transition coupling $\pi_* \in \Pi_{\text{TC}}(\mathbb{P},\mathbb{Q})$ from R^* that is optimal in the OTC problem. Note that by definition, $\bar{c}_R(s) = \langle c, \lambda_{R,s}\rangle$. Then by the optimality of R^* in TC-MDP, $\langle c, \lambda_{R^*,s}\rangle = \min_{R\in\Pi_{\text{TC}}(P,Q)}\langle c, \lambda_{R,s}\rangle$ for every $s \in \mathcal{S}$. So by Lemma 15,

$$\min_{s\in\mathcal{S}_r}\langle c, \lambda_{R^*,s}\rangle = \min_{s\in\mathcal{S}}\langle c, \lambda_{R^*,s}\rangle = \min_{R\in\Pi_{\text{TC}}(P,Q)} \min_{s\in\mathcal{S}}\langle c, \lambda_{R,s}\rangle = \min_{R\in\Pi_{\text{TC}}(P,Q)} \min_{\lambda\in\Lambda(R)}\langle c, \lambda\rangle. \tag{8}$$

Let $s^* \in \operatorname{argmin}_{s\in\mathcal{S}_r}\langle c, \lambda_{R^*,s}\rangle$ and define $\pi_* \in \Pi_{\text{TC}}(\mathbb{P},\mathbb{Q})$ to be the transition coupling with transition matrix R^* and stationary distribution λ_{s^*}. Then by (8),

$$\int c\, d\pi_* = \langle c, \lambda_{R^*,s^*}\rangle = \min_{R\in\Pi_{\text{TC}}(P,Q)} \min_{\lambda\in\Lambda(R)}\langle c, \lambda\rangle.$$

But at this point, we recognize that the quantity on the right is exactly the OTC cost. To see this, note that by Proposition 4 every $\pi \in \Pi_{\text{TC}}(\mathbb{P},\mathbb{Q})$ is uniquely characterized by a transition matrix $R \in \Pi_{\text{TC}}(P,Q)$ and a stationary distribution $\lambda \in \Lambda(R)$, and $\int c\, d\pi = \langle c, \lambda\rangle$. Thus

$$\int c\, d\pi_* = \min_{R\in\Pi_{\text{TC}}(P,Q)} \min_{\lambda\in\Lambda(R)}\langle c, \lambda\rangle = \min_{\pi\in\Pi_{\text{TC}}(\mathbb{P},\mathbb{Q})} \int c\, d\pi,$$

and we conclude that π_* is optimal for the OTC problem. Finally, by construction, $\int c\, d\pi_* = \min_{s\in\mathcal{S}}\langle c, \lambda_{R^*,s}\rangle = \min_{s\in\mathcal{S}} \bar{c}_{R^*}(s)$. ■

9.3.3 Convergence of ExactOTC

Next, we prove the convergence of Algorithm 1 to a solution of TC-MDP. For any polyhedron $\mathcal{P} \in \mathbb{R}^{n\times n}$, let $\mathcal{E}(\mathcal{P})$ denote the extreme points of $\mathcal{P}$. Recall that if $\mathcal{P}$ is bounded, a linear function on $\mathcal{P}$ achieves its minimum on $\mathcal{E}(\mathcal{P})$ (Bertsimas and Tsitsiklis, 1997). Note that for every $(x,y) \in \mathcal{X}\times\mathcal{Y}$, since $\Pi(P(x,\cdot),Q(y,\cdot))$ is a bounded subset of $\mathbb{R}^{d^2}$ defined by a finite set of linear equality and inequality constraints, it is a bounded polyhedron.

Theorem 7 *Algorithm 1 converges to a solution (g^*, h^*, R^*) of TC-MDP in a finite number of iterations. Moreover, if X and Y are irreducible, R^* is the transition matrix of an optimal transition coupling of X and Y.*

Proof We will first show that Algorithm 1 converges to some (g^*, h^*, R^*) and then argue that this is a solution to TC-MDP. Recall that for every $s = (x, y)$, $\mathcal{A}_s = \Pi(P(x, \cdot), Q(y, \cdot))$ and $\mathcal{A} = \bigcup_s \mathcal{A}_s$. In this proof, it is most convenient to consider the concatenatation of the state-action spaces instead of the union $\bigcup_s \mathcal{A}_s$. Abusing notation, we let $\mathcal{A} = \bigotimes_s \mathcal{A}_s$ for the remainder of the proof. Furthermore, let $\mathcal{A}'_s = \mathcal{E}(\mathcal{A}_s)$ be the set of extreme points of $\mathcal{A}_s$. As $\mathcal{A}_s$ is a bounded polyhedron, $\mathcal{A}'_s$ is finite. For every $n \geq 1$, let (g_n, h_n, R_n) be the n'th iterate of Algorithm 1. Since the rows of R_n are solutions of the linear programs in Algorithm 1b,$R_n(s, \cdot) \in \mathcal{E}(\mathcal{A}'_s)$ for every s. Thus the iterates of Algorithm 1 are the same as the iterates of the policy iteration algorithm for the restricted MDP $(\mathcal{X} \times \mathcal{Y}, \bigcup_s \mathcal{A}'_s, \{p(\cdot|s, a)\}, c)$ constructed by restricting the state-action spaces $\mathcal{A}_s$ of TC-MDP to $\mathcal{A}'_s$ for each s. Since $\mathcal{A}'_s$ is finite for every s, standard results (Puterman, 2005, Theorem 9.2.3) ensure that the iterates $\{(g_n, h_n, R_n)\}$ of Algorithm 1 will converge to a solution (g^*, h^*, R^*) in a finite number of iterations. Thus, we need only show that any stationary point of Algorithm 1 is necessarily a solution to TC-MDP.

Let (g^*, h^*, R^*) be a stationary point of Algorithm 1. Then $R^* = \texttt{ExactTCI}(g^*, h^*, R^*, \bigotimes_s \mathcal{A}'_s)$ and consequently, $R^*(s, \cdot) \in \operatorname{argmin}_{r \in \mathcal{A}'_s} rh^*$ for every s. Since $\mathcal{A}_s$ is a bounded polyhedron, $\min_{r \in \mathcal{A}_s} rh^* = \min_{r \in \mathcal{A}'_s} rh^*$ and we find that $R^*(s, \cdot) \in \operatorname{argmin}_{r \in \mathcal{A}_s} rh^*$. Since $\mathcal{A} = \bigotimes_s \mathcal{A}_s$, we may write $R^* \in \operatorname{argmin}_{R \in \mathcal{A}} Rh^*$ where the minimum is understood to be element-wise. Using the assumption that (g^*, h^*, R^*) is a stationary point of Algorithm 1 again, $(g^*, h^*) = \texttt{ExactTCE}(R^*)$. It follows that

$$g^* + h^* = R^* h^* + c. \tag{9}$$

Since $R^* \in \operatorname{argmin}_{R \in \mathcal{A}} Rh^*$, we obtain

$$g^* + h^* = \min_{R \in \mathcal{A}} Rh^* + c.$$

Then by (Puterman, 2005, Theorem 9.1.2 (c)), g^* is the optimal expected cost for TC-MDP. Moreover, by (9) and (Puterman, 2005, Theorem 8.2.6 (b)), $g^* = \overline{R}^* c = \bar{c}_{R^*}$, where we remind the reader that $\overline{R}^* = \lim_{T \to \infty} 1/T \sum_{t=0}^{T-1} R^{*t}$. Thus R^* has optimal expected cost among policies for TC-MDP and we conclude that (g^*, h^*, R^*) is a solution to TC-MDP.

If X and Y are irreducible, then by Proposition 4, every transition coupling matrix in $\Pi_{\mathrm{TC}}(P, Q)$ induces a transition coupling in $\Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$. Since R^* has minimal expected cost over all elements of $\Pi_{\mathrm{TC}}(P, Q)$, it attains the minimum in Problem (2) and is thus an optimal transition coupling. ■

9.4 Proofs from Section 5

Next we prove our results from Section 5 regarding the complexities of `ApproxTCE` and `EntropicTCI`.

9.4.1 Complexity of Approximate Transition Coupling Evaluation

Theorem 9 *Let $R \in \Pi_{TC}(P, Q)$ be aperiodic and irreducible with mixing coefficients $M \in \mathbb{R}_+$ and $\alpha \in [0, 1)$ and gain and bias vectors $g \in \mathbb{R}^{d^2}$ and $h \in \mathbb{R}^{d^2}$, respectively. Then for any $\varepsilon > 0$, there exist $L, T \in \mathbb{N}$ such that* `ApproxTCE`(R, L, T) *yields $(\tilde{g}, \tilde{h})$ satisfying $\|\tilde{g} - g\|_\infty \leq \varepsilon$ and $\|\tilde{h} - h\|_1 \leq \varepsilon$ in $\tilde{\mathcal{O}}\left(\frac{d^4}{\log \alpha^{-1}} \log\left(\frac{M}{\varepsilon(1-\alpha)}\right)\right)$ time.*

Proof Briefly, we remind the reader that $g = \overline{R}c$ and $h = \sum_{t=0}^{\infty} R^t(c - g)$, and that for integers $L, T \geq 1$ to be chosen later,

$$\tilde{g} = \langle 1/d^2 R^L c, \mathbb{1} \rangle \mathbb{1} \quad \text{and} \quad \tilde{h} = \sum_{t=0}^{T} R^t(c - \tilde{g}).$$

Note that the expression for $\tilde{g}$ may also be written as

$$\tilde{g} = \left(\frac{1}{d^2} \sum_s R^L(s, \cdot)c\right) \mathbb{1}.$$

We begin by studying the approximation error for $\tilde{h}$ by first considering the intermediate quantity $h' := \sum_{t=0}^{T} R^t(c - g)$. By the triangle inequality,

$$\|\tilde{h} - h\|_1 \leq \|\tilde{h} - h'\|_1 + \|h' - h\|_1, \tag{10}$$

so it suffices to control the two terms on the right hand side. Using Hölder's inequality, it follows that

$$\begin{aligned}
\|\tilde{h} - h'\|_1 &= \left\| \sum_{t=0}^{T} R^t(\tilde{g} - g) \right\|_1 \\
&\leq \sum_{t=0}^{T} \left\| R^t(\tilde{g} - g) \right\|_1 \\
&\leq d^2 \sum_{t=0}^{T} \max_s \left| R^t(s, \cdot)(\tilde{g} - g) \right| \\
&\overset{(*)}{\leq} d^2 \sum_{t=0}^{T} \|\tilde{g} - g\|_\infty \\
&= (T + 1)d^2 \|\tilde{g} - g\|_\infty,
\end{aligned}$$

where (*) uses the fact that $\|R^t(s, \cdot)\|_1 = 1$ for every $t \geq 1$ and $s \in \mathcal{X} \times \mathcal{Y}$. Next we wish to bound $\|h' - h\|_1$. Since $R^t\overline{R} = \overline{R}$ for any $t \geq 1$, we may write h and h' as

$$h = \sum_{t=0}^{\infty} (R^t - \overline{R})c \quad \text{and} \quad h' = \sum_{t=0}^{T} (R^t - \overline{R})c.$$

Moreover, since R is aperiodic and irreducible, the Perron-Frobenius theorem implies that $\overline{R}(s,\cdot) = \lambda_R$ for every $s \in \mathcal{X} \times \mathcal{Y}$, where $\lambda_R \in \Delta_{d^2}$ is the unique stationary distribution of R. Now by Hölder's inequality and the mixing assumption on R,

$$\begin{aligned}
\|h' - h\|_1 &= \left\| \sum_{t=T+1}^{\infty} (R^t - \overline{R})c \right\|_1 \\
&\leq \sum_{t=T+1}^{\infty} \|(R^t - \overline{R})c\|_1 \\
&\leq d^2 \sum_{t=T+1}^{\infty} \max_s \left|(R^t(s,\cdot) - \lambda_R)c\right| \\
&\leq \|c\|_\infty d^2 \sum_{t=T+1}^{\infty} \max_s \left\|R^t(s,\cdot) - \lambda_R\right\|_1 \\
&\leq \|c\|_\infty d^2 \sum_{t=T+1}^{\infty} M\alpha^t \\
&= M\|c\|_\infty \frac{\alpha^{T+1}}{1-\alpha} d^2.
\end{aligned}$$

Thus by (10),

$$\|\tilde{h} - h\|_1 \leq (T+1)\|\tilde{g} - g\|_\infty d^2 + M\|c\|_\infty \frac{\alpha^{T+1}}{1-\alpha} d^2. \tag{11}$$

So in order to bound $\|\tilde{h} - h\|_1$, we require a bound on $\|\tilde{g} - g\|_\infty$. Using the fact that $\tilde{g}$ and g are constant vectors, Hölder's inequality and the mixing assumption on R,

$$\begin{aligned}
\|\tilde{g} - g\|_\infty &= \left\| \left(\frac{1}{d^2} \sum_s R^L(s,\cdot)c \right) \mathbb{1} - \overline{R}c \right\|_\infty \\
&= \left| \frac{1}{d^2} \sum_s R^L(s,\cdot)c - \lambda_R c \right| \\
&\leq \frac{1}{d^2} \sum_s \left|(R^L(s,\cdot) - \lambda_R)c\right| \\
&\leq \frac{1}{d^2} \sum_s \|c\|_\infty \|R^L(s,\cdot) - \lambda_R\|_1 \\
&\leq \frac{1}{d^2} \sum_s M\alpha^L \|c\|_\infty \\
&\leq M\alpha^L \|c\|_\infty.
\end{aligned}$$

Plugging this into (11),

$$\|\tilde{h} - h\|_1 \leq M\alpha^L \|c\|_\infty (T+1) d^2 + M\|c\|_\infty \frac{\alpha^{T+1}}{1-\alpha} d^2.$$

Then choosing

$$T + 1 \geq \frac{1}{\log \alpha^{-1}} \log \left(\frac{2M \|c\|_\infty d^2 \varepsilon^{-1}}{(1-\alpha)} \right) = \tilde{\mathcal{O}} \left(\frac{1}{\log \alpha^{-1}} \log \left(\frac{M}{\varepsilon(1-\alpha)} \right) \right) \tag{12}$$

and

$$L \geq \frac{\log \left(2(T+1) M \|c\|_\infty d^2 \varepsilon^{-1} \right)}{\log \alpha^{-1}} = \tilde{\mathcal{O}} \left(\frac{1}{\log \alpha^{-1}} \log \left(\frac{M}{\varepsilon} \right) \right), \tag{13}$$

we obtain $\|\tilde{h} - h\|_1 \leq \varepsilon$. Note that for this choice of L, $\|\tilde{g} - g\|_\infty \leq \varepsilon/2(T+1)$. Since $T + 1 \geq 1$, this implies that $\|\tilde{g} - g\|_\infty \leq \varepsilon$. So the error for $\tilde{g}$ is controlled at the desired level as well.

Now consider the cost of computing $\tilde{g}$ and $\tilde{h}$. Computing $\tilde{g}$ requires L multiplications of a vector in $\mathbb{R}^{d^2}$ by $R \in \mathbb{R}^{d^2 \times d^2}$, which takes $\mathcal{O}(Ld^4)$ time, followed by an inner product with $\mathbb{1} \in \mathbb{R}^{d^2}$, multiplication with $\mathbb{1} \in \mathbb{R}^{d^2}$ and multiplication by $1/d^2$, each in $\mathcal{O}(d^2)$ time. This requires $\mathcal{O}(Ld^4) + \mathcal{O}(d^2) + \mathcal{O}(d^2) + \mathcal{O}(d^2) = \mathcal{O}(Ld^4)$ time. Letting L be the minimum integer satisfying (13), this takes time

$$\mathcal{O}(Ld^4) = \tilde{\mathcal{O}} \left(\frac{d^4}{\log \alpha^{-1}} \log \left(\frac{M}{\varepsilon} \right) \right).$$

On the other hand, given $\tilde{g}$, computing $\tilde{h}$ requires computing $c - \tilde{g} \in \mathbb{R}^{d^2}$ in $\mathcal{O}(d^2)$ operations then multiplying by $R \in \mathbb{R}^{d^2 \times d^2}$ $T+1$ times in $\mathcal{O}(Td^4)$ time. Finally, the sum may also be evaluated in $\mathcal{O}(Td^4)$, requiring a total time of $\mathcal{O}(d^2) + \mathcal{O}(Td^4) + \mathcal{O}(Td^4) = \mathcal{O}(Td^4)$. Letting T be the minimum integer satisfying (12), this takes time

$$\mathcal{O}(Td^4) = \tilde{\mathcal{O}} \left(\frac{d^4}{\log \alpha^{-1}} \log \left(\frac{M}{\varepsilon(1-\alpha)} \right) \right). \tag{14}$$

In total, we find that `ApproxTCE`(R, L, T) takes time

$$\tilde{\mathcal{O}} \left(\frac{d^4}{\log \alpha^{-1}} \log \left(\frac{M}{\varepsilon} \right) \right) + \tilde{\mathcal{O}} \left(\frac{d^4}{\log \alpha^{-1}} \log \left(\frac{M}{\varepsilon(1-\alpha)} \right) \right) = \tilde{\mathcal{O}} \left(\frac{d^4}{\log \alpha^{-1}} \log \left(\frac{M}{\varepsilon(1-\alpha)} \right) \right).$$

■

9.4.2 Aperiodicity and Irreducibility of Elements of $\text{ri}(\Pi_{\text{TC}}(P,Q))$

Next we prove Proposition 11 regarding the aperiodicity and irreducibility of elements of $\text{ri}(\Pi_{\text{TC}}(P,Q))$. We begin with two elementary lemmas about the independent transition coupling.

Lemma 16 *For any $k \geq 1$, $(P \otimes Q)^k = P^k \otimes Q^k$.*

Proof The result clearly holds for $k = 1$, so assume that it holds for some $k \geq 1$. For any (x, y), $(x', y') \in \mathcal{X} \times \mathcal{Y}$, we can show

$$(P \otimes Q)^{k+1}((x,y),(x',y')) = \sum_{\tilde{x},\tilde{y}} (P \otimes Q)^k((x,y),(\tilde{x},\tilde{y}))\, P \otimes Q((\tilde{x},\tilde{y}),(x',y'))$$

$$
\begin{aligned}
&= \sum_{\tilde{x},\tilde{y}} P^k(x,\tilde{x})\, Q^k(y,\tilde{y})\, P(\tilde{x},x')\, Q(\tilde{y},y') \\
&= \sum_{\tilde{x}} P^k(x,\tilde{x})\, P(\tilde{x},x') \sum_{\tilde{y}} Q^k(y,\tilde{y})\, Q(\tilde{y},y') \\
&= P^{k+1}(x,x')\, Q^{k+1}(y,y') \\
&= P^{k+1} \otimes Q^{k+1}((x,y),(x',y')).
\end{aligned}
$$

By induction, the lemma is proven. ■

Lemma 17 *If P and Q are aperiodic and irreducible, then the independent transition coupling $P \otimes Q$ is aperiodic and irreducible.*

Proof Since P and Q are aperiodic and irreducible, there exist $\ell_0, m_0 \geq 1$ such that for any $\ell \geq \ell_0$ and $m \geq m_0$, $P^\ell > 0$ and $Q^m > 0$ (Levin and Peres, 2017, Proposition 1.7). Defining $k_0 := \ell_0 \vee m_0$, for every $k \geq k_0$, $P^k, Q^k > 0$. By Lemma 16, it follows that $(P \otimes Q)^k = P^k \otimes Q^k > 0$ for all $k \geq k_0$. Thus $P \otimes Q$ is irreducible. Furthermore, for every $s \in \mathcal{X} \times \mathcal{Y}$, $\gcd\{k \geq 1 : (P \otimes Q)^k(s,s) > 0\} = \gcd\{..., k_0, k_0+1, ...\} = 1$ and we conclude that $P \otimes Q$ is also aperiodic. ■

Next we prove Proposition 11. Recall that for a set $\mathcal{U} \subset \mathbb{R}^n$, $B_\varepsilon(u) \subset \mathbb{R}^n$ denotes the open ball of radius $\varepsilon > 0$ centered at $u \in \mathcal{U}$, $\mathrm{aff}(\mathcal{U})$ denotes the affine hull, defined as $\mathrm{aff}(\mathcal{U}) = \{\sum_{i=1}^k \alpha_i u_i : k \in \mathbb{N}, u_1, ..., u_k \in \mathcal{U}, \sum_{i=1}^k \alpha_i = 1\}$, and $\mathrm{ri}(\mathcal{U})$ denotes the relative interior, defined as $\mathrm{ri}(\mathcal{U}) = \{u \in \mathcal{U} : \exists \varepsilon > 0 \text{ s.t. } B_\varepsilon(u) \cap \mathrm{aff}(\mathcal{U}) \subset \mathcal{U}\}$.

Proposition 11 *If P and Q are aperiodic and irreducible then every $R \in \mathrm{ri}(\Pi_{TC}(P,Q))$ is also aperiodic and irreducible, and thus mixing.*

Proof First we establish that $P \otimes Q(s,s') > 0$ implies that $R(s,s') > 0$ for every $s, s' \in \mathcal{X} \times \mathcal{Y}$. Suppose for the sake of contradiction that there exist $s, s' \in \mathcal{X} \times \mathcal{Y}$ such that $P \otimes Q(s,s') > 0$ and $R(s,s') = 0$. By definition, there is some $\varepsilon > 0$ such that $B_\varepsilon(R) \cap \mathrm{aff}(\Pi_{\mathrm{TC}}(P,Q)) \subset \Pi_{\mathrm{TC}}(P,Q)$. Defining $R' = R + \frac{\varepsilon}{2} d$ where $d = (R - P \otimes Q)/\|R - P \otimes Q\|_2$, one may verify that $R' \in B_\varepsilon(R) \cap \mathrm{aff}(\Pi_{\mathrm{TC}}(P,Q))$. Thus by the choice of R, we have $R' \in \Pi_{\mathrm{TC}}(P,Q)$. However, our assumptions imply that $R'(s,s') < 0$, a contradiction. This proves the preliminary claim.

By nature of the fact that $R((x,y),\cdot) \in \Pi(P(x,\cdot), Q(y,\cdot))$, one may easily establish that the reverse implication holds: $R(s,s') > 0$ implies that $P \otimes Q(s,s') > 0$ for every $s, s' \in \mathcal{X} \times \mathcal{Y}$. As such, one may find a positive constant $a > 0$ such that $aP \otimes Q \leq R$ where the inequality is understood to hold element-wise. Now, by Lemma 17, $P \otimes Q$ is aperiodic and irreducible. Thus there exists $k \geq 1$ such that $(P \otimes Q)^k > 0$. Thus, $R^k \geq a^k (P \otimes Q)^k > 0$ and it follows that R is aperiodic and irreducible as well. The mixing property of R follows from (Levin and Peres, 2017, Theorem 4.9). ■

9.4.3 Complexity of Entropic Transition Coupling Improvement

Next we aim to prove Theorem 12, showing that `EntropicTCI` returns an improved transition coupling with error bounded by $\varepsilon > 0$ in $\tilde{\mathcal{O}}(d^4\varepsilon^{-4})$ time. Recall that `EntropicTCI` improves policies by solving d^2 entropy-regularized OT transport problems, calling the `ApproxOT` algorithm (Altschuler et al., 2017) for each problem. Before we can prove Theorem 12, we must analyze the computational complexity of `ApproxOT`. In the following discussion as well as Lemma 18, we find it most convenient to adopt the notation of Altschuler et al. (2017). Thus, we fix two probability vectors $r \in \Delta_m$ and $c \in \Delta_n$, a non-negative cost matrix $C \in \mathbb{R}_+^{m\times n}$, a regularization parameter $\xi > 0$, and an error tolerance $\varepsilon > 0$. For vectors in $\mathbb{R}^m$ or $\mathbb{R}^n$ and matrices in $\mathbb{R}^{m\times n}$, we temporarily drop the double-indexing convention, using subscripts instead to denote elements (i.e. u_i and X_{ij}). Finally, for a coupling $X \in \Pi(r, c)$, let $H(X) = -\sum_{ij} X_{ij} \log X_{ij}$ be the Shannon entropy.

Recall that the entropic OT problem is defined as,

$$\begin{aligned} \text{minimize} \quad & \langle X, C\rangle - \frac{1}{\xi} H(X) \\ \text{subject to} \quad & X \in \Pi(r, c). \end{aligned} \tag{15}$$

Cuturi (2013) showed that solutions to (15) have a computationally convenient form. Namely, if $X_\xi^* \in \Pi(r, c)$ is the solution to (15), then it is unique and can be written as $X_\xi^* = \text{diag}(e^{u^*})K\,\text{diag}(e^{v^*})$ for some $u^* \in \mathbb{R}^m$ and $v^* \in \mathbb{R}^n$, where $K = e^{-\xi C}$. As a result, (15) can be formulated as a matrix scaling problem and solved using Sinkhorn's algorithm (Sinkhorn, 1967).

Altschuler et al. (2017) introduced the `ApproxOT` algorithm (Algorithm 3), which combines Sinkhorn's algorithm with a rounding step to obtain an approximate solution to the OT problem. In particular, `ApproxOT` runs `Sinkhorn` (Algorithm 4) to obtain a coupling of the form $X' = \text{diag}(e^{u'})K\,\text{diag}(e^{v'}) \in \Pi(r', c')$, where $\|r - r'\|_1 + \|c - c'\|_1 \leq \varepsilon$, then applies `Round` (Algorithm 5) to X' to obtain $\hat{X} \in \Pi(r, c)$. `ApproxOT` was originally intended for approximating the OT cost, but we use it to approximate the regularized optimal coupling $X_\xi^* \in \Pi(r, c)$. In particular, we wish to show that for appropriate choice of parameters, `ApproxOT` yields a coupling $\hat{X} \in \Pi(r, c)$ such that $\|\hat{X} - X_\xi^*\|_1 \leq \varepsilon$ in $\tilde{\mathcal{O}}(mn\varepsilon^{-4})$ time. To the best of our knowledge, this result has not appeared in the literature. So we state and prove it in Lemma 18.

Note that `ApproxOT` was originally defined for fully-supported marginal probability vectors $(r, c > 0)$. However, this will not always be the case in Algorithm 2b. In particular, transition couplings may be sparse, even when P and Q are strictly positive. Thus we add an extra step to `ApproxOT` that subsets the quantities of interest to their positive entries. For an index set $\mathcal{I}$ and a vector / matrix A we let $A_{\mathcal{I}}$ denote the subvector / matrix that retains only elements with indices contained in $\mathcal{I}$.

Algorithm 3: `ApproxOT`

result : optimal coupling
input : $r, c, C, \xi, \varepsilon$
`/* subset to positive elements */`
$\mathcal{R} \leftarrow \{i : r_i > 0\}$, $\mathcal{C} \leftarrow \{j : c_j > 0\}$
$\mathcal{S} \leftarrow \mathcal{R} \times \mathcal{C}$, $\tilde{r} \leftarrow r_{\mathcal{R}}$, $c \leftarrow c_{\mathcal{C}}$
`/* set parameters */`
$J \leftarrow 4 \log n \|C_{\mathcal{S}}\|_\infty / \varepsilon - \log \min_{ij}\{\tilde{r}_i, \tilde{c}_j\}$
$\varepsilon' \leftarrow \varepsilon^2 / 8J$
$K \leftarrow \exp(-\xi C_{\mathcal{S}})$
`/* approximate Sinkhorn projection */`
$X' \leftarrow \texttt{Sinkhorn}(K, \tilde{r}, \tilde{c}, \varepsilon')$
`/* round to feasible coupling */`
$X' \leftarrow \texttt{Round}(X', \Pi(\tilde{r}, \tilde{c}))$
`/* replace zeroes */`
$\hat{X} \leftarrow 0_{d \times d}$, $\hat{X}_{\mathcal{S}} \leftarrow X'$
return $\hat{X}$

Algorithm 4: `Sinkhorn`

result : approximate Sinkhorn projection
input : K, r, c, ε'
$k \leftarrow 0$
$X_0 \leftarrow K / \|K\|_1$, $u^0 \leftarrow 0$, $v^0 \leftarrow 0$
while $\|X_k \mathbb{1} - r\|_1 + \|X_k^\top \mathbb{1} - c\|_1 > \varepsilon'$ **do**
  $k \leftarrow k + 1$
  if k *odd* **then**
    $r^k \leftarrow X_k \mathbb{1}$
    $u_i \leftarrow \log(r_i / r_i^k)$ for $i \in [n]$
    $u^k \leftarrow u^{k-1} + u$, $v^k \leftarrow v^{k-1}$
  else
    $c^k \leftarrow X_k^\top \mathbb{1}$
    $v_j \leftarrow \log(c_j / c_j^k)$ for $j \in [n]$
    $v^k \leftarrow v^{k-1} + v$, $u^k \leftarrow u^{k-1}$
  $X_k \leftarrow \mathrm{diag}(e^{u^k}) K \,\mathrm{diag}(e^{v^k})$
return X_k

Algorithm 5: `Round`

result : feasible coupling
input : $F, \Pi(r,c)$
$r' \leftarrow F\mathbb{1}$
$X \leftarrow \operatorname{diag}(x)$ with $x_i = r_i/r'_i \wedge 1$
$F' \leftarrow XF$
$c' \leftarrow (F')^\top \mathbb{1}$
$Y \leftarrow \operatorname{diag}(y)$ with $y_j = c_j/c'_j \wedge 1$
$F'' \leftarrow F'Y$
$r'' \leftarrow F''\mathbb{1}$, $c'' \leftarrow (F'')^\top \mathbb{1}$
$\operatorname{err}_r \leftarrow r - r''$, $\operatorname{err}_c \leftarrow c - c''$
return $F'' + \operatorname{err}_r \operatorname{err}_c^\top / \| \operatorname{err}_r \|_1$

Lemma 18 *Let $r \in \Delta_m$ and $c \in \Delta_n$ have all positive entries, $C \in \mathbb{R}_+^{m\times n}$, $\xi > 0$ and $\varepsilon \in (0,1)$. Then* `ApproxOT`(r,c,C,ξ,ε) *returns a coupling $\hat{X} \in \Pi(r,c)$ such that $\|\hat{X} - X_\xi^*\|_1 \leq \varepsilon$, where $X_\xi^* \in \operatorname{argmin}_{X\in\Pi(r,c)} \langle X, C\rangle - 1/\xi H(X)$, in time $\tilde{\mathcal{O}}(mn\varepsilon^{-4}\xi\|C\|_\infty(\xi^2\|C\|_\infty^2 + (\log b^{-1})^2))$ where $b = \min_{ij}\{r_i, c_j\}$.*

Proof Let $\varepsilon' > 0$, $K = e^{-\xi C}$, $X' \in \Delta_{m\times n}$ be the output of `Sinkhorn`(K,r,c,ε') and $\hat{X} \in \Pi(r,c)$ be the output of `Round`$(X', \Pi(r,c))$. By the triangle inequality,

$$\|\hat{X} - X_\xi^*\|_1 \leq \|\hat{X} - X'\|_1 + \|X' - X_\xi^*\|_1. \tag{16}$$

We will first describe how to control the second term on the right hand side. By Pinsker's inequality, $\|X' - X_\xi^*\|_1^2 \leq 2\mathcal{K}(X_\xi^* \| X')$, so it suffices to bound the KL-divergence between the two couplings. From Lemma 2 of Cuturi (2013) that $X_\xi^* = \operatorname{diag}(e^{u^*})K\operatorname{diag}(e^{v^*})$ for some $u^* \in \mathbb{R}^m$, $v^* \in \mathbb{R}^n$, and $K = e^{-\xi C}$. By construction we also have $X' = \operatorname{diag}(e^{u'})K\operatorname{diag}(e^{v'})$ for some $u' \in \mathbb{R}^m$ and $v' \in \mathbb{R}^n$. Now rewriting the KL-divergence,

$$\begin{aligned}
\mathcal{K}(X_\xi^* \| X') &= \sum_{ij} X_{\xi,ij}^* \log X_{\xi,ij}^* - \sum_{ij} X_{\xi,ij}^* \log X'_{ij} \\
&= \sum_{ij} X_{\xi,ij}^* \left(u_i^* + v_j^* - \xi C_{ij}\right) - \sum_{ij} X_{\xi,ij}^* \left(u'_i + v'_j - \xi C_{ij}\right) \\
&= \sum_{ij} X_{\xi,ij}^* (u_i^* - u'_i) + \sum_{ij} X_{\xi,ij}^* (v_j^* - v'_j) \\
&= \sum_i (u_i^* - u'_i) \sum_j X_{\xi,ij}^* + \sum_j (v_j^* - v'_j) \sum_i X_{\xi,ij}^* \\
&= \sum_i (u_i^* - u'_i) r_i + \sum_j (v_j^* - v'_i) c_j \\
&= \langle u^* - u', r\rangle + \langle v^* - v', c\rangle.
\end{aligned}$$

Writing $\psi(u,v) = \langle \mathbb{1}, \operatorname{diag}(e^u)K\operatorname{diag}(e^v)\mathbb{1}\rangle - \langle u, r\rangle - \langle v, c\rangle$ for the objective of the dual entropic OT problem (Dvurechensky et al., 2018), we immediately see that

$$\tilde{\psi}(u', v') := \psi(u', v') - \psi(u^*, v^*) = \langle u^* - u', r\rangle + \langle v^* - v', c\rangle.$$

Now let r' and c' be the row and column marginals of X', respectively. Using the two previous displays and applying the upper bound from (Dvurechensky et al., 2018, Lemma 2), we obtain

$$\mathcal{K}(X^*_\xi \| X') = \tilde{\psi}(u, v) \leq J\left(\|r' - r\|_1 + \|c' - c\|_1\right),$$

where $J = \xi\|C\|_\infty - \log \min_{ij}\{r_i, c_j\}$. For ease of notation, we will let $b := \min_{ij}\{r_i, c_j\}$. Now by (Altschuler et al., 2017, Theorem 2) and the fact that each iteration of `Sinkhorn` takes $\mathcal{O}(mn)$ time, `Sinkhorn`(K, r, c, ε') returns a coupling with $X' \in \Pi(r', c')$ satisfying $\|r' - r\|_1 + \|c' - c\|_1 \leq \varepsilon'$ in $\mathcal{O}(mn(\varepsilon')^{-2}\log(s/\ell))$ time where $s = \sum_{ij} K_{ij}$ and $\ell = \min_{ij} K_{ij}$. As C is non-negative, $s = \sum_{ij} e^{-\xi C_{ij}} \leq \sum_{ij} 1 = mn$. Furthermore, $\ell = e^{-\xi\|C\|_\infty}$ so we get a total runtime of $\mathcal{O}(mn(\varepsilon')^{-2}(\log mn + \xi\|C\|_\infty)) = \tilde{\mathcal{O}}(mn(\varepsilon')^{-2}\xi\|C\|_\infty)$. Now choosing $\varepsilon' = \varepsilon^2/8J$, we have

$$\|X' - X^*_\xi\|_1 \leq \sqrt{2J(\|r' - r\|_1 + \|c' - c\|_1)} \leq \sqrt{2J\varepsilon'} = \sqrt{2J\varepsilon^2/8J} = \varepsilon/2.$$

Since $\varepsilon' = \varepsilon^2/8J$, the runtime becomes

$$\begin{aligned}\tilde{\mathcal{O}}(mn(\varepsilon')^{-2}\xi\|C\|_\infty) &= \tilde{\mathcal{O}}(mn(\varepsilon^2/8J)^{-2}\xi\|C\|_\infty) \\ &= \tilde{\mathcal{O}}(mn\varepsilon^{-4}\xi\|C\|_\infty J^2) \\ &= \tilde{\mathcal{O}}(mn\varepsilon^{-4}\xi\|C\|_\infty(\xi\|C\|_\infty - \log b)^2) \\ &= \tilde{\mathcal{O}}(mn\varepsilon^{-4}\xi\|C\|_\infty(\xi^2\|C\|_\infty^2 + (\log b^{-1})^2)).\end{aligned}$$

Now we must bound $\|\hat{X} - X'\|_1$. By (Altschuler et al., 2017, Lemma 7), Algorithm 5 returns $\hat{X}$ satisfying

$$\|\hat{X} - X'\|_1 \leq 2(\|r' - r\|_1 + \|c' - c\|_1),$$

in $\mathcal{O}(mn)$ time. So it suffices to check that $\|r' - r\|_1 + \|c' - c\|_1 \leq \varepsilon' = \varepsilon^2/8J$ is enough to guarantee that $\|\hat{X} - X'\|_1 \leq \varepsilon/2$. This will follow immediately from $\|\hat{X} - X'\|_1 \leq 2\varepsilon' = \varepsilon^2/4J \leq \varepsilon/2J$ if we can establish that $J \geq 1$. To see this, first note that $b = \min_{i,j}\{r_i, c_j\} \leq 1/(m \vee n)$. This implies that $-\log b \geq \log(m \vee n)$ and since $\xi > 0$,

$$J = \xi\|C\|_\infty - \log b \geq -\log b \geq \log(m \vee n) \geq 1,$$

assuming that $m \vee n > 2$. If $m \vee n = 2$, then one can check that letting $\varepsilon' = \varepsilon^2 \log 2/8J$ is enough to obtain the desired bounds without affecting the computational complexity. Thus by (16), we obtain $\|\hat{X} - X^*_\xi\|_1 \leq \varepsilon$ in time $\tilde{\mathcal{O}}(mn\varepsilon^{-4}\xi\|C\|_\infty(\xi^2\|C\|_\infty^2 + (\log b^{-1})^2) + mn) = \tilde{\mathcal{O}}(mn\varepsilon^{-4}\xi\|C\|_\infty(\xi^2\|C\|_\infty^2 + (\log b^{-1})^2))$. ■

Now we can proceed to the proof of Theorem 12.

Theorem 12 *Let P and Q be aperiodic and irreducible, $h \in \mathbb{R}^{d^2}$, $\xi > 0$, and $\varepsilon > 0$. Then* `EntropicTCI`(h, ξ, ε) *returns $\hat{R} \in \mathrm{ri}(\Pi_{TC}(P, Q))$ with $\max_s \|\hat{R}(s, \cdot) - R^*(s, \cdot)\|_1 \leq \varepsilon$ for some $R^* \in \mathrm{argmin}_{R' \in \Pi_{TC}(P,Q)} R'h - {}^1/_\xi H(R')$ in $\tilde{\mathcal{O}}(d^4\varepsilon^{-4})$ time.*

Proof Without loss of generality, we may assume that h is non-negative. Otherwise, one can consider the modified bias $h + \|h\|_\infty\mathbb{1}$. Since we are interested in optimal couplings with respect to h rather than expected cost and $\|h + \|h\|_\infty \mathbb{1}\|_\infty = \mathcal{O}(\|h\|_\infty)$, this has no effect on the output of `ApproxOT` or the computational complexity. Now, in order to analyze the complexity of `EntropicTCI`, we must first analyze the complexity of `ApproxOT`. Fix $s = (x, y) \in \mathcal{X} \times \mathcal{Y}$ and, after removing points outside of the supports of $P(x, \cdot)$ and $Q(y, \cdot)$, consider the entropic OT problem for marginal probability measures $P(x, \cdot)$ and $Q(y, \cdot)$ and cost h,

$$\begin{aligned} \text{minimize} \quad & \langle r, h \rangle - \frac{1}{\xi} H(r) \\ \text{subject to} \quad & r \in \Pi(P(x, \cdot), Q(y, \cdot)). \end{aligned} \tag{17}$$

Then by (Cuturi, 2013, Lemma 2), there exists a unique solution $r_s^* \in \Pi(P(x, \cdot), Q(y, \cdot))$ to problem (17). Furthermore by Lemma 18, $\texttt{ApproxOT}(P(x, \cdot)^\top, Q(y, \cdot)^\top, h, \xi, \varepsilon)$ returns $\hat{r}_s \in \Pi(P(x, \cdot), Q(y, \cdot))$ such that $\|\hat{r}_s - r_s^*\|_1 \leq \varepsilon$ in $\tilde{\mathcal{O}}(d^2 \varepsilon^{-4})$ time. One may also verify using arguments in Altschuler et al. (2017) that $\hat{r}_s \in \text{ri}(\Pi(P(x, \cdot), Q(y, \cdot)))$.

Now we may analyze the error and computational complexity of $\texttt{EntropicTCI}(h, \xi, \varepsilon)$. Calling $\texttt{ApproxOT}(P(x, \cdot)^\top, Q(y, \cdot)^\top, h, \xi, \varepsilon)$ for every $s = (x, y) \in \mathcal{X} \times \mathcal{Y}$, we obtain $\hat{R} \in \Pi_{\text{TC}}(P, Q)$, where $\hat{R}(s, \cdot) = \hat{r}_s(\cdot)$, in $d^2 \tilde{\mathcal{O}}(d^2 \varepsilon^{-4}) = \tilde{\mathcal{O}}(d^4 \varepsilon^{-4})$ time. Note that since the relative interior commutes with cartesian products of convex sets, $\hat{R} \in \text{ri}(\Pi_{\text{TC}}(P, Q))$. Then defining $R^* \in \Pi_{\text{TC}}(P, Q)$ such that $R^*(s, \cdot) = r_s^*(\cdot)$, we have

$$\max_s \|\hat{R}(s, \cdot) - R^*(s, \cdot)\|_1 = \max_s \|\hat{r}_s - r_s^*\|_1 \leq \varepsilon,$$

by construction. This concludes the proof. ■

9.5 Proofs from Section 6

Our proof of Theorem 13 relies on a well-known result regarding the stability of certain optimization problems. Before stating this result, fix spaces $\mathcal{Z}$ and $\mathcal{U}$ corresponding to the set of possible solutions and set of parameters for the optimization problem of interest, respectively. Now consider the following problem.

$$\begin{aligned} \text{minimize} \quad & f(z, u) \\ \text{subject to} \quad & z \in \Phi(u). \end{aligned} \tag{18}$$

Note that $f(\cdot, u) : \mathcal{Z} \to \mathbb{R}$ describes the objective to be minimized and $\Phi(u) \subset \mathcal{Z}$ represents the feasible set of Problem (18), both indexed by a parameter $u \in \mathcal{U}$. We will call a set $\mathcal{V} \subset \mathcal{Z}$ a neighborhood of a subset $\mathcal{W} \subset \mathcal{Z}$ if $\mathcal{W} \subset \text{int}\, \mathcal{V}$. Neighborhoods in $\mathcal{U}$ will be defined similarly. Recall that a multifunction $F : \mathcal{U} \to 2^{\mathcal{Z}}$ is upper semicontinuous at a point $u_0 \in \mathcal{U}$ if for any neighborhood $\mathcal{V}_{\mathcal{Z}}$ of the set $F(u_0)$, there exists a neighborhood $\mathcal{V}_{\mathcal{U}}$ of u_0 such that for every $u \in \mathcal{V}_{\mathcal{U}}$, $F(u) \subset \mathcal{V}_{\mathcal{Z}}$.

Theorem 19 (Bonnans and Shapiro (2013), Proposition 4.4) *Let u_0 be a given point in the parameter space $\mathcal{U}$. Suppose that (i) the function $f(z, u)$ is continuous on $\mathcal{Z} \times \mathcal{U}$, (ii) the graph of the multifunction $\Phi(\cdot)$ is a closed subset of $\mathcal{U} \times \mathcal{Z}$, (iii) there exists $\alpha \in \mathbb{R}$*

and a compact set $C \subset \mathcal{Z}$ such that for every u in a neighborhood of u_0, the level set $\{z \in \Phi(u) : f(z,u) \leq \alpha\}$ is nonempty and contained in C, (iv) for any neighborhood $\mathcal{V}_{\mathcal{Z}}$ of the set $\operatorname{argmin}_{z\in\Phi(u_0)} f(z,u_0)$ there exists a neighborhood $\mathcal{V}_U$ of u_0 such that $\mathcal{V}_{\mathcal{Z}} \cap \Phi(u) \neq \emptyset$ for all $u \in \mathcal{V}_{\mathcal{U}}$. Then the optimal value function $u \mapsto \min_{z\in\Phi(u)} f(z,u)$ is continuous at $u = u_0$ and the multifunction $u \mapsto \operatorname{argmin}_{z\in\Phi(u)} f(z,u)$ is upper semicontinuous at u_0.

Both Problems (I) and (II) may be recast in the form of Problem (18). Let

$$\mathcal{Z} = \Big\{(\lambda, R) \in \Delta_{d^2} \times \Delta_{d^2}^{d^2} : R \in \Pi_{\mathrm{TC}}(P,Q) \text{ for some } P,Q \in \Delta_d^d, \lambda R = \lambda\Big\}$$

and $\mathcal{U} = \Delta_d^d \times \Delta_d^d$ be the set of all valid pairs of transition matrices in $\mathbb{R}^{d\times d}$. It is straightforward to verify that $\mathcal{Z}$ and $\mathcal{U}$ are in fact compact subsets of $\mathbb{R}^{d^2} \times \mathbb{R}^{d^2\times d^2}$ and $\mathbb{R}^{d\times d} \times \mathbb{R}^{d\times d}$, respectively. The objective function $f(\cdot)$ is identified with the map $(\lambda, R) \mapsto \langle c, \lambda\rangle$ and does not depend on the parameter $u = (P,Q)$. We will refer to the constraint functions for Problems (I) and (II) by $\Phi : \mathcal{U} \to 2^{\mathcal{Z}}$ and $\Phi_\eta : \mathcal{U} \to 2^{\mathcal{Z}}$, and their optimal solution functions by $\Phi^* : \mathcal{U} \to 2^{\mathcal{Z}}$ and $\Phi^*_\eta : \mathcal{U} \to 2^{\mathcal{Z}}$, respectively.

Theorem 13 *Let $P,Q \in \Delta_d^d$ be irreducible transition matrices. Then the following hold:*

- *$\rho(\cdot,\cdot)$ is continuous and $\Phi^*(\cdot,\cdot)$ is upper semicontinuous at (P,Q)*
- *For any $\eta > 0$, $\rho_\eta(\cdot,\cdot)$ is continuous and $\Phi^*_\eta(\cdot,\cdot)$ is upper semicontinuous at (P,Q)*

Proof We will prove the result for Problem (2) as the proof for Problem (4) is similar. As the two problems are equivalent, it suffices to check the conditions of Theorem 19 for Problem (I) at the point $u_0 = (P,Q) \in \mathcal{U}$. First, (i) is vacuously true since the objective $f(\cdot)$ does not depend on u. Next, we will show that the graph of $\Phi(\cdot)$ is a closed subset of $\mathcal{U} \times \mathcal{Z}$. Fix a sequence $\{(P_n, Q_n, \lambda_n, R_n)\}_{n\geq1} \subset \operatorname{graph}\Phi(\cdot)$. As a subset of the compact set $\Delta_d^d \times \Delta_d^d \times \Delta_{d^2} \times \Delta_{d^2}^{d^2}$, it has a subsequence, which we also label as $\{(P_n, Q_n, \lambda_n, R_n)\}_{n\geq1}$ converging to some $(P', Q', \lambda', R') \in \Delta_d^d \times \Delta_d^d \times \Delta_{d^2} \times \Delta_{d^2}^{d^2}$. Taking limits of the linear equations $R_n \in \Pi_{\mathrm{TC}}(P_n, Q_n)$ and $\lambda_n R_n = \lambda_n$, we conclude that $R' \in \Pi_{\mathrm{TC}}(P', Q')$ and $\lambda' R' = \lambda'$. Thus $(P', Q', \lambda', R') \in \operatorname{graph}\Phi(\cdot)$ and (ii) holds. To show that (iii) is satisfied, note that one may let $\alpha = \|c\|_\infty$ and use the fact that the entire set $\mathcal{Z}$ is compact. Finally, we will show that (iv) is satisfied. Let $\mathcal{V}_{\mathcal{Z}} \subset \mathcal{Z}$ be a neighborhood of $\operatorname{argmin}_{z\in\Phi(u_0)} f(z,u_0)$. Then define the neighborhood $\mathcal{V}_{\mathcal{U}}$ of $u_0 = (P,Q)$ as

$$\mathcal{V}_{\mathcal{U}} := \{(P,Q) \in \Delta_d^d \times \Delta_d^d : R \in \Pi_{\mathrm{TC}}(P,Q) \text{ for some } (\lambda, R) \in \mathcal{V}_{\mathcal{Z}}\}.$$

Note that $\mathcal{V}_{\mathcal{U}}$ is nonempty by the non-emptiness of $\mathcal{V}_{\mathcal{Z}}$ and the definition of $\mathcal{Z}$. Moreover, $\mathcal{V}_{\mathcal{Z}} \cap \Phi(u) \neq \emptyset$ for all $u \in \mathcal{V}_{\mathcal{U}}$ by construction. Thus all the conditions of Theorem 19 are satisfied and the desired convergence holds. ■

Acknowledgments

The authors would like to thank Quoc Tran-Dinh for helpful discussions and Jason Altschuler for contributions to the proof of Lemma 18. K.O. and A.B.N. were supported in part by NIH Grant R01 HG009125-01 and NSF Grant DMS-1613072. K.M. was supported in part by NSF Grant DMS-1847144. K.M. and A.B.N. were supported in part by NSF Grant DMS-1613261.

Appendix A. Properties of the OTC Problems

In this appendix, we prove that solutions to the OTC and constrained OTC problems exist via continuity and compactness arguments and establish the triangle inequality for the unconstrained problem. For a metric space $\mathcal{U}$ and a sequence of Borel probability measures $\{\mu^n\} \subset \mathcal{M}(\mathcal{U})$, we say that μ^n *converges weakly to* $\mu \in \mathcal{M}(\mathcal{U})$, denoted by $\mu^n \Rightarrow \mu$, if for every continuous and bounded function $f : \mathcal{U} \to \mathbb{R}$, $\int f \, d\mu^n \to \int f \, d\mu$. A set $\Pi \subset \mathcal{M}(\mathcal{U})$ is said to be *weakly compact* if every sequence in Π contains a subsequence converging weakly to an element of Π. Π is said to be *tight* if for every $\varepsilon > 0$, there exists a compact set $K \subset \mathcal{U}$ such that $\mu(K) > 1 - \varepsilon$ for every $\mu \in \Pi$. Tightness and relative compactness are related by Prohorov's theorem which states that if $\mathcal{U}$ is a separable metric space, $\Pi \subset \mathcal{M}(\mathcal{U})$ is tight if and only if its closure is relatively compact. Note that $\mathcal{X}^{\mathbb{N}} \times \mathcal{Y}^{\mathbb{N}}$ is complete and separable when equipped with the metric

$$d((\mathbf{x}^1, \mathbf{y}^1), (\mathbf{x}^2, \mathbf{y}^2)) = \sum_{k=0}^{\infty} 2^{-k} \delta((x_k^1, y_k^1) \neq (x_k^2, y_k^2)).$$

Finally, we remark that since $\tilde{c} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$ is continuous and bounded, $c(\mathbf{x}, \mathbf{y}) = \tilde{c}(x_0, y_0)$ is as well.

A.1 Existence of Solutions to the OTC Problem

Lemma 20 $\Pi_{TC}(\mathbb{P}, \mathbb{Q})$ *is weakly compact.*

Proof It is well-known that the set of stationary couplings (joinings) of two stationary processes is weakly compact (see the arguments in McGoff and Nobel 2021). As a subset of this set, $\Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$ is tight. Thus by Prohorov's theorem, we need only prove that $\Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$ is closed. Take a sequence $\{\pi^n\} \subset \Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$ such that $\pi^n \Rightarrow \pi$ for some joining π of $\mathbb{P}$ and $\mathbb{Q}$. Since the state space $\mathcal{X} \times \mathcal{Y}$ of the Markov chains of interest is finite, the conditions of Theorem 1 of Karr (1975) are satisfied and we have that π is Markov. Thus it suffices to check that the transition matrix R of π satisfies the transition coupling condition. Let R_n denote the transition matrix of π^n. The weak convergence $\pi^n \Rightarrow \pi$ implies that $R_n(s, s') \to R(s, s')$ for every $s, s' \in \mathcal{X} \times \mathcal{Y}$. Then for any $(x, y) \in \mathcal{X} \times \mathcal{Y}$ and $y' \in \mathcal{Y}$,

$$\sum_{x'} R_n((x, y), (x', y')) \to \sum_{x'} R((x, y), (x', y')).$$

But as $R_n \in \Pi_{\mathrm{TC}}(P, Q)$, $\sum_{x'} R_n((x, y), (x', y')) = Q(y, y')$ and it follows that $\sum_{x'} R((x, y), (x', y')) = Q(y, y')$. Employing a similar argument to the other marginal of R, one may show that in fact $R \in \Pi_{\mathrm{TC}}(P, Q)$. Therefore, $\pi \in \Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$ and we conclude that $\Pi_{\mathrm{TC}}(\mathbb{P}, \mathbb{Q})$ is weakly compact. ∎

Proposition 21 *The OTC problem* (2) *has a solution. Moreover, this solution may be assumed to have a transition matrix that admits a unique stationary distribution.*

Proof Let $\{\pi^n\} \subset \Pi_{\text{TC}}(\mathbb{P}, \mathbb{Q})$ be a sequence such that

$$\int c\, d\pi^n \to \inf_{\pi \in \Pi_{\text{TC}}(\mathbb{P},\mathbb{Q})} \int c\, d\pi.$$

By Lemma 20, $\Pi_{\text{TC}}(\mathbb{P}, \mathbb{Q})$ is weakly compact. Thus, there exists a subsequence $\{\pi^{n_k}\}$ such that $\pi^{n_k} \Rightarrow \pi^*$ for some $\pi^* \in \Pi_{\text{TC}}(\mathbb{P}, \mathbb{Q})$. Since c is continuous and bounded,

$$\int c\, d\pi^* = \lim_{k\to\infty} \int c\, d\pi^{n_k} = \inf_{\pi \in \Pi_{\text{TC}}(\mathbb{P},\mathbb{Q})} \int c\, d\pi.$$

Thus π^* is an optimal solution for Problem (2).

In order to establish the second half of the proposition, we will construct a second transition coupling of $\mathbb{P}$ and $\mathbb{Q}$ from π^* that is also optimal and satisfies the conclusion of the proposition. Let R be the transition matrix of π^*. By Lemma 15, $\int c\, d\pi^* = \langle c, \lambda \rangle$ where λ is the unique stationary distribution of one of the communicating classes of R. Let $\mathcal{S} \subset \mathcal{X} \times \mathcal{Y}$ denote the set of states in this class and define a new transition matrix $\tilde{R}$ on $\mathcal{X} \times \mathcal{Y}$ by

$$\tilde{R}((x,y),(x',y')) = \begin{cases} R((x,y),(x',y')) & (x,y) \in \mathcal{S} \\ P(x,x')Q(y,y') & \text{otherwise} \end{cases}.$$

In other words, $\tilde{R}$ corresponds to the Markov chain that transitions according to R when the chain occupies a state in $\mathcal{S}$ and independently otherwise. One may verify that λ is the unique stationary distribution of $\tilde{R}$ and thus the stationary Markov process $\tilde{\pi}$ corresponding to $\tilde{R}$ has expected cost equal to that of π^*. Moreover, it is clear from the definition of $\tilde{R}$ that $\tilde{\pi} \in \Pi_{\text{TC}}(\mathbb{P}, \mathbb{Q})$ so $\tilde{\pi}$ is in fact an optimal transition coupling of $\mathbb{P}$ and $\mathbb{Q}$ satisfying the second conclusion of the proposition. ■

A.2 Existence of Solutions to the Constrained OTC Problem

Lemma 22 *For any* $\eta \geq 0$, $\Pi^\eta_{TC}(\mathbb{P}, \mathbb{Q})$ *is weakly compact.*

Proof The proof is similar to that of Lemma 20. Let $\{\pi^n\} \subset \Pi^\eta_{\text{TC}}(\mathbb{P}, \mathbb{Q})$ be a sequence such that $\pi^n \Rightarrow \pi$ for some joining π of $\mathbb{P}$ and $\mathbb{Q}$. By Lemma 20, $\Pi_{\text{TC}}(\mathbb{P}, \mathbb{Q})$ is weakly compact so $\pi \in \Pi_{\text{TC}}(\mathbb{P}, \mathbb{Q})$. Letting R be the transition matrix of π, we need only show that $R \in \Pi^\eta_{\text{TC}}(P, Q)$. Letting R_n be the transition matrix of π^n, the weak convergence $\pi^n \Rightarrow \pi$ implies that $R_n \to R$. Using the weak lower semicontinuity of the KL-divergence, for every $s \in \mathcal{X} \times \mathcal{Y}$,

$$\mathcal{K}(R(s,\cdot) \| P \otimes Q(s,\cdot)) \leq \liminf_{n\to\infty} \mathcal{K}(R_n(s,\cdot) \| P \otimes Q(s,\cdot)) \leq \eta.$$

Therefore, $R \in \Pi_\eta(P, Q)$ and we find that $\pi \in \Pi^\eta_{\text{TC}}(\mathbb{P}, \mathbb{Q})$. Thus, we conclude that $\Pi^\eta_{\text{TC}}(\mathbb{P}, \mathbb{Q})$ is weakly compact. ■

Proposition 23 *For any $\eta > 0$, the constrained OTC problem (4) has a solution.*

The proof of Proposition 23 is similar to that of Proposition 21 with an application of Lemma 22 rather than Lemma 20.

A.3 Triangle Inequality

Next we prove that the optimal transition coupling cost satisfies the triangle inequality when the cost does. In the rest of the section, we fix irreducible stationary Markov processes $\mathbb{P}_1$, $\mathbb{P}_2$, $\mathbb{P}_3 \in \mathcal{M}(\mathcal{X}^{\mathbb{N}})$ with transition matrices P_1, P_2, and P_3, respectively. Let $\Pi_{\mathrm{TC}}(\mathbb{P}_1, \mathbb{P}_2, \mathbb{P}_3)$ denote the set of three-way transition couplings of $\mathbb{P}_1$, $\mathbb{P}_2$, and $\mathbb{P}_3$, defined in the obvious way.

Lemma 24 (Gluing Lemma) *Let $\pi_{12} \in \Pi_{TC}(\mathbb{P}_1, \mathbb{P}_2)$ and $\pi_{23} \in \Pi_{TC}(\mathbb{P}_2, \mathbb{P}_3)$ have transition matrices R_{12} and R_{23} with unique stationary distributions. Then there exists $\pi \in \Pi_{TC}(\mathbb{P}_1, \mathbb{P}_2, \mathbb{P}_3)$ admitting π_{12} and π_{23} as marginals.*

Proof Let R be the transition matrix on $\mathcal{X} \times \mathcal{X} \times \mathcal{X}$ such that for any $x, y, z, x', y', z' \in \mathcal{X}$,

$$R((x, y, z), (x', y', z')) = \frac{R_{12}((x, y), (x', y'))R_{23}((y, z), (y', z'))}{P_2(y, y')},$$

where we let $0/0 = 0$. Let $\lambda \in \mathcal{M}(\mathcal{X} \times \mathcal{X} \times \mathcal{X})$ be a stationary distribution of R and $\pi \in \mathcal{M}(\mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}})$ be the stationary process measure corresponding to the Markov process constructed from λ and R. It is straightforward to check that in fact $R \in \Pi_{\mathrm{TC}}(P_1, P_2, P_3)$ and thus $\pi \in \Pi_{\mathrm{TC}}(\mathbb{P}_1, \mathbb{P}_2, \mathbb{P}_3)$. In order to see that π admits π_{12} and π_{23} as marginals, one may use the work above to show that $\lambda_{12} \in \mathcal{M}(\mathcal{X} \times \mathcal{X})$ and $\lambda_{23} \in \mathcal{M}(\mathcal{X} \times \mathcal{X})$, defined by

$$\lambda_{12}(x, y) = \sum_z \lambda(x, y, z) \quad \text{and} \quad \lambda_{23}(y, z) = \sum_x \lambda(x, y, z),$$

are stationary for R_{12} and R_{23}, respectively. Since the stationary distributions of R_{12} and R_{23} are assumed to be unique, this implies that λ_{12} and λ_{23} are in fact the stationary distributions of π_{12} and π_{23}. Thus for any $n \geq 1$ and $x_1^n, y_1^n \in \mathcal{X}^n$,

$$\begin{aligned}
&\sum_{z_1^n} \pi([x_1^n, y_1^n, z_1^n]) \\
&\quad = \sum_{z_1^n} \lambda(x_1, y_1, z_1)R((x_1, y_1, z_1), (x_2, y_2, z_2)) \cdots R((x_{n-1}, y_{n-1}, z_{n-1}), (x_n, y_n, z_n)) \\
&\quad = \sum_{z_1^{n-1}} \lambda(x_1, y_1, z_1)R((x_1, y_1, z_1), (x_2, y_2, z_2)) \cdots R_{12}((x_{n-1}, y_{n-1}), (x_n, y_n)) \\
&\quad \vdots \\
&\quad = \lambda_{12}(x_1, y_1)R_{12}((x_1, y_1), (x_2, y_2)) \cdots R_{12}((x_{n-1}, y_{n-1}), (x_n, y_n)) \\
&\quad = \pi_{12}([x_1^n, y_1^n]).
\end{aligned}$$

In the same way, one may show that $\sum_{x_1^n} \pi([x_1^n, y_1^n, z_1^n]) = \pi_{23}([y_1^n, z_1^n])$ for every $y_1^n, z_1^n \in \mathcal{X}^n$. Thus, we conclude that π admits π_{12} and π_{23} as marginals. ■

Proposition 25 (Triangle Inequality) *Let $c : \mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}} \to \mathbb{R}_+$ be a single-letter cost that satisfies the triangle inequality. Then the OTC problem satisfies*

$$\min_{\pi \in \Pi_{TC}(\mathbb{P}_1, \mathbb{P}_3)} \int c\, d\pi \leq \min_{\pi \in \Pi_{TC}(\mathbb{P}_1, \mathbb{P}_2)} \int c\, d\pi + \min_{\pi \in \Pi_{TC}(\mathbb{P}_2, \mathbb{P}_3)} \int c\, d\pi. \tag{19}$$

Proof By Proposition 21, there exist $\pi_{12} \in \Pi_{\mathrm{TC}}(\mathbb{P}_1, \mathbb{P}_2)$ and $\pi_{23} \in \Pi_{\mathrm{TC}}(\mathbb{P}_2, \mathbb{P}_3)$ that are optimal in the two problems on the right hand side of (19) and have transition matrices with unique stationary distributions. Then by Lemma 24, there exists $\pi \in \Pi_{\mathrm{TC}}(\mathbb{P}_1, \mathbb{P}_2, \mathbb{P}_3)$ that admits π_{12} and π_{23} as marginals. Define the measure $\pi_{13} \in \mathcal{M}(\mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}})$ by $\pi_{13}(A_1 \times A_3) = \pi(A_1 \times \mathcal{X}^{\mathbb{N}} \times A_3)$ for every A_1, $A_3 \subset \mathcal{X}^{\mathbb{N}}$. Then clearly π_{13} is a joining of $\mathbb{P}_1$ and $\mathbb{P}_3$ with a one-step transition matrix that satisfies the transition coupling condition with respect to P_1 and P_3. However, π_{13} need not be Markov and thus need not be a transition coupling of $\mathbb{P}_1$ and $\mathbb{P}_3$. Using the stationary distribution and one-step transition distribution of π_{13}, one may construct a Markov process with an identical stationary distribution and one-step transition distribution, denoted by $\tilde{\pi}_{13} \in \mathcal{M}(\mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}})$. Then one has $\tilde{\pi}_{13} \in \Pi_{\mathrm{TC}}(\mathbb{P}_1, \mathbb{P}_3)$ and it follows that

$$\begin{aligned}
\min_{\pi \in \Pi_{\mathrm{TC}}(\mathbb{P}_1, \mathbb{P}_3)} \int c\, d\pi &\leq \int_{\mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}}} c(\mathbf{x}, \mathbf{z})\, d\tilde{\pi}_{13}(\mathbf{x}, \mathbf{z}) \\
&= \int_{\mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}}} c(\mathbf{x}, \mathbf{z})\, d\pi_{13}(\mathbf{x}, \mathbf{z}) \qquad (20) \\
&= \int_{\mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}}} c(\mathbf{x}, \mathbf{z})\, d\pi(\mathbf{x}, \mathbf{y}, \mathbf{z}) \\
&\leq \int_{\mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}}} (c(\mathbf{x}, \mathbf{y}) + c(\mathbf{y}, \mathbf{z}))\, d\pi(\mathbf{x}, \mathbf{y}, \mathbf{z}) \\
&= \int_{\mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}}} c(\mathbf{x}, \mathbf{y})\, d\pi_{12}(\mathbf{x}, \mathbf{y}) + \int_{\mathcal{X}^{\mathbb{N}} \times \mathcal{X}^{\mathbb{N}}} c(\mathbf{y}, \mathbf{z})\, d\pi_{23}(\mathbf{y}, \mathbf{z}) \\
&= \min_{\pi \in \Pi_{\mathrm{TC}}(\mathbb{P}_1, \mathbb{P}_2)} \int c\, d\pi + \min_{\pi \in \Pi_{\mathrm{TC}}(\mathbb{P}_2, \mathbb{P}_3)} \int c\, d\pi,
\end{aligned}$$

where in (20) we use the single-letter property of c and the fact that $\tilde{\pi}_{13}$ and π_{13} have identical stationary distributions. ■

Appendix B. Reducible Transition Coupling of Irreducible Chains

In this appendix, we provide an example showing that a transition coupling of two irreducible transition matrices is not necessarily irreducible. Let

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 0.25 & 0.25 & 0.50 \\ 0.25 & 0.25 & 0.50 \\ 0.25 & 0.25 & 0.50 \end{bmatrix} \end{array} \quad \text{and} \quad Q = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 0.25 & 0.25 & 0.50 \\ 0.25 & 0.25 & 0.50 \\ 0.50 & 0.25 & 0.25 \end{bmatrix} \end{array}.$$

Both P and Q are clearly irreducible, but the transition coupling

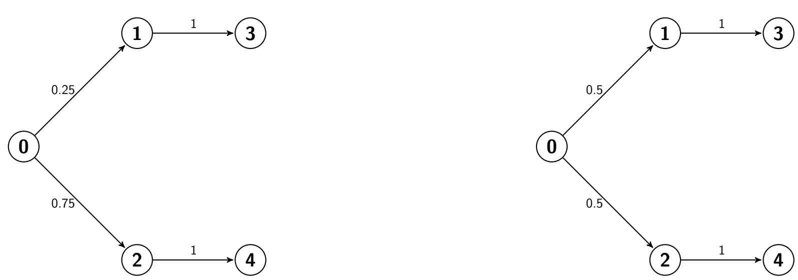
$$R = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (0,2) \\ (1,0) \\ (1,1) \\ (1,2) \\ (2,0) \\ (2,1) \\ (2,2) \end{array} \begin{array}{c} \begin{array}{ccccccccc} (0,0) & (0,1) & (0,2) & (1,0) & (1,1) & (1,2) & (2,0) & (2,1) & (2,2) \end{array} \\ \begin{bmatrix} 0 & 0.25 & 0 & 0.25 & 0 & 0 & 0 & 0 & 0.50 \\ 0 & 0 & 0.25 & 0 & 0 & 0.25 & 0.25 & 0.25 & 0 \\ 0 & 0 & 0.25 & 0.25 & 0 & 0 & 0.25 & 0.25 & 0 \\ 0.25 & 0 & 0 & 0 & 0 & 0.25 & 0 & 0.25 & 0.25 \\ 0 & 0 & 0.25 & 0.25 & 0 & 0 & 0 & 0.25 & 0.25 \\ 0 & 0.25 & 0 & 0 & 0 & 0.25 & 0.50 & 0 & 0 \\ 0 & 0.25 & 0 & 0.25 & 0 & 0 & 0 & 0 & 0.50 \\ 0.25 & 0 & 0 & 0 & 0 & 0.25 & 0 & 0.25 & 0.25 \\ 0 & 0.25 & 0 & 0 & 0 & 0.25 & 0.50 & 0 & 0 \end{bmatrix} \end{array}$$

is reducible. While we do not provide an example here, we remark that transition coupling matrices of aperiodic and irreducible transition matrices may also have multiple recurrent classes.

Appendix C. Comparison to 1-step Optimal Transition Coupling

In this appendix, we demonstrate how the 1-step transition coupling problem described in Section 2 prioritizes expected cost in the next step over long-term average cost as the OTC problem does.

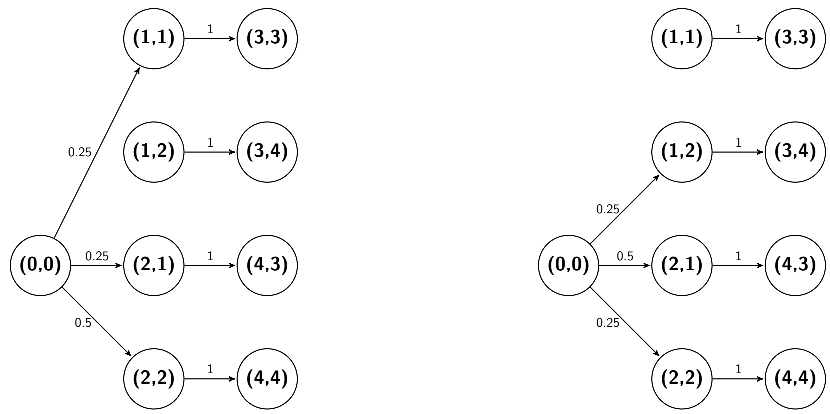
Example 1 *Consider stationary Markov chains X and Y with transition distributions defined by the graphs in Figure 4. In order to find an OTC of X and Y, we must specify a cost for every pair of states $(x, y) \in \mathcal{X} \times \mathcal{Y}$. Let states $(0,0)$, $(1,2)$, $(2,1)$, $(2,2)$, and $(3,3)$ have cost 0, states $(1,1)$, $(3,4)$ and $(4,3)$ have cost 1, state $(4,4)$ have cost 9, and let all other states have a cost sufficiently large 1-step OTC and OTC do not assign them positive probability.*



(a) X transition probabilities (b) Y transition probabilities

Figure 4: Marginal stationary Markov chains. Both chains return to state 0 from states 3 and 4 with probability one.

The transition distributions of the OTC and 1-step OTC are largely the same except for the transitions from $(0,0)$ to $(1,1)$, $(1,2)$, $(2,1)$ and $(2,2)$ (see Figure 5 for an illustration). In particular, since the OTC chooses the transitions to minimize expected cost over the complete



(a) 1-step OTC (expected cost of 5/3) (b) OTC (expected cost of 1)

Figure 5: An example where the 1-step OTC has sub-optimal expected cost. Both chains return to state $(0,0)$ from states $(3,3)$, $(3,4)$, $(4,3)$, and $(4,4)$ with probability one. Note that Figures 5a and 5b omit the edges that are the same between the two transition couplings.

trajectory of the chain, it assigns lower probability to the transition $(0,0) \to (2,2)$ *in order to avoid the costly state* $(4,4)$. *On the other hand, the 1-step OTC does not incorporate this information in deciding how to transition from* $(0,0)$ *and assigns a higher probability to the transition* $(0,0) \to (2,2)$. *As a result, the expected cost of the 1-step OTC is* $5/3$ *compared to an expected cost of* 1 *for the OTC. In fact, by increasing the cost of the state* $(4,4)$, *one can make the difference between the 1-step OTC and OTC costs arbitrarily large. The lower expected cost indicates that the OTC constitutes a better alignment of* X *and* Y *as compared to the 1-step OTC.*

References

David J Aldous and Persi Diaconis. https://www.stat.berkeley.edu/ aldous/unpub/persi.pdf. 2009.

Moray Allan and Christopher KI Williams. Harmonising chorales by probabilistic inference. *Advances in Neural Information Processing Systems*, 17:25–32, 2005.

Jason Altschuler, Jonathan Weed, and Philippe Rigollet. Near-linear time approximation algorithms for optimal transport via Sinkhorn iteration. In *Advances in Neural Information Processing Systems*, pages 1964–1974, 2017.

Charles Ames. The Markov process as a compositional model: A survey and tutorial. *Leonardo*, 22(2):175–187, 1989.

Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *International Conference on Machine Learning*, pages 214–223, 2017.

Julio Backhoff, Daniel Bartl, Mathias Beiglböck, and Johannes Wiesel. Estimating processes in adapted Wasserstein distance. *The Annals of Applied Probability*, forthcoming.

Lalit Bahl, Peter Brown, Peter De Souza, and Robert Mercer. Maximum mutual information estimation of hidden Markov model parameters for speech recognition. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 11, pages 49–52, 1986.

Sayan Banerjee and Wilfrid S Kendall. Coupling the Kolmogorov diffusion: maximality and efficiency considerations. *Advances in Applied Probability*, 48(A):15–35, 2016.

Sayan Banerjee and Wilfrid S Kendall. Rigidity for Markovian maximal couplings of elliptic diffusions. *Probability Theory and Related Fields*, 168(1-2):55–112, 2017.

Sayan Banerjee and Wilfrid S Kendall. Coupling polynomial Stratonovich integrals: the two-dimensional Brownian case. *Electronic Journal of Probability*, 23, 2018.

Mathias Beiglböck, Pierre Henry-Labordère, and Friedrich Penkner. Model-independent bounds for option prices—a mass transport approach. *Finance and Stochastics*, 17(3): 477–501, 2013.

Chip Bell. Algorithmic music composition using dynamic Markov chains and genetic algorithms. *Journal of Computing Sciences in Colleges*, 27(2):99–107, 2011.

Dimitris Bertsimas and John N Tsitsiklis. *Introduction to Linear Optimization*, volume 6. Athena Scientific Belmont, MA, 1997.

J Frédéric Bonnans and Alexander Shapiro. *Perturbation Analysis of Optimization Problems*. Springer Science & Business Media, 2013.

Mike Boyle and Karl Petersen. Hidden Markov processes in the context of symbolic dynamics. *arXiv preprint arXiv:0907.1858*, 2009.

Elsa Cazelles, Arnaud Robert, and Felipe Tobar. The Wasserstein-Fourier distance for stationary time series. *IEEE Transactions on Signal Processing*, 2020.

Taolue Chen and Stefan Kiefer. On the total variation distance of labelled Markov chains. In *Joint Meeting of the Twenty-Third EACSL Annual Conference on Computer Science Logic (CSL) and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–10, 2014.

Yukun Chen, Jianbo Ye, and Jia Li. Aggregated Wasserstein distance and state registration for hidden Markov models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(9):2133–2147, 2019.

Samuel Cohen, Giulia Luise, Alexander Terenin, Brandon Amos, and Marc Deisenroth. Aligning time series on incomparable spaces. In *International Conference on Artificial Intelligence and Statistics*, pages 1036–1044, 2021.

Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, pages 2292–2300, 2013.

Marco Cuturi and Mathieu Blondel. Soft-dtw: a differentiable loss function for time-series. In *International Conference on Machine Learning*, pages 894–903, 2017.

Przemyslaw Daca, Thomas A Henzinger, Jan Kretínskỳ, and Tatjana Petrov. Linear distances between Markov chains. In *International Conference on Concurrency Theory*, 2016.

Orchisama Das, Blair Kaneshiro, and Tom Collins. Analyzing and classifying guitarists from rock guitar solo tablature. In *Sound and Music Computing Conference*, 2018.

Rommert Dekker. Counter examples for compact action Markov decision chains with average reward criteria. *Stochastic Models*, 3(3):357–368, 1987.

Ishan Deshpande, Ziyu Zhang, and Alexander Schwing. Generative modeling using the sliced Wasserstein distance. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3483–3491, 2018.

Pavel Dvurechensky, Alexander Gasnikov, and Alexey Kroshnin. Computational optimal transport: Complexity by accelerated gradient descent is better than by Sinkhorn's algorithm. In *International Conference on Machine Learning*, pages 1367–1376, 2018.

Martin Ellis. The $\overline{d}$-distance between two Markov processes cannot always be attained by a Markov joining. *Israel Journal of Mathematics*, 24(3-4):269–273, 1976.

Martin Ellis. Distances between two-state Markov processes attainable by Markov joinings. *Transactions of the American Mathematical Society*, 241:129–153, 1978.

Martin Ellis. Conditions for attaining $\bar{d}$ by a Markovian joining. *The Annals of Probability*, 8(3):431–440, 1980a.

Martin Ellis. On Kamae's conjecture concerning the $\overline{d}$-distance between two-state Markov processes. *The Annals of Probability*, pages 372–376, 1980b.

Aden Forrow, Jan-Christian Hütter, Mor Nitzan, Philippe Rigollet, Geoffrey Schiebinger, and Jonathan Weed. Statistical optimal transport via factored couplings. In *International Conference on Artificial Intelligence and Statistics*, pages 2454–2465, 2019.

Charlie Frogner, Chiyuan Zhang, Hossein Mobahi, Mauricio Araya, and Tomaso Poggio. Learning with a Wasserstein loss. In *Advances in Neural Information Processing Systems*, pages 2053–2061, 2015.

Harry Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Theory of Computing Systems*, 1(1):1–49, 1967.

Aude Genevay, Gabriel Peyré, and Marco Cuturi. Learning generative models with Sinkhorn divergences. In *International Conference on Artificial Intelligence and Statistics*, pages 1608–1617, 2018.

Robert Gray, David Neuhoff, and Paul Shields. A generalization of Ornstein's $\overline{d}$-distance with applications to information theory. *The Annals of Probability*, pages 315–328, 1975.

David Scott Griffeath. *Coupling Methods for Markov Processes*. Cornell University, January, 1976.

Wenshuo Guo, Nhat Ho, and Michael Jordan. Fast algorithms for computational optimal transport and Wasserstein barycenter. In *International Conference on Artificial Intelligence and Statistics*, pages 2088–2097, 2020.

Ronald Howard. *Dynamic Programming and Markov Processes*. John Wiley, 1960.

Hicham Janati, Marco Cuturi, and Alexandre Gramfort. Wasserstein regularization for sparse multi-task regression. In *International Conference on Artificial Intelligence and Statistics*, volume 89, 2019.

Hicham Janati, Marco Cuturi, and Alexandre Gramfort. Spatio-temporal alignments: Optimal transport through space and time. In *International Conference on Artificial Intelligence and Statistics*, pages 1695–1704, 2020.

Alan F Karr. Weak convergence of a sequence of Markov chains. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 33(1):41–48, 1975.

Stefan Kiefer. On computing the total variation distance of hidden Markov models. In *International Colloquium on Automata, Languages, and Programming*, 2018.

Soheil Kolouri, Phillip E Pope, Charles E Martin, and Gustavo K Rohde. Sliced Wasserstein auto-encoders. In *International Conference on Learning Representations*, 2018.

David A Levin and Yuval Peres. *Markov Chains and Mixing Times*, volume 107. American Mathematical Soc., 2017.

Tianyi Lin, Nhat Ho, and Michael Jordan. On efficient optimal transport: An analysis of greedy and accelerated mirror descent algorithms. In *International Conference on Machine Learning*, pages 3982–3991, 2019.

Torgny Lindvall. *Lectures on the Coupling Method*. Courier Corporation, 2002.

Yi-Wen Liu and Eleanor Selfridge-Field. Modeling music as Markov chains: Composer identification, 2002.

Giulia Luise, Alessandro Rudi, Massimiliano Pontil, and Carlo Ciliberto. Differential properties of Sinkhorn approximation for learning with Wasserstein distance. In *International Conference on Neural Information Processing Systems*, pages 5864–5874, 2018.

Kevin McGoff and Andrew B. Nobel. Empirical risk minimization and complexity of dynamical models. *The Annals of Statistics*, 48(4):2031–2054, 2020.

Kevin McGoff and Andrew B. Nobel. Empirical risk minimization for dynamical systems and stationary processes. *Information and Inference: A Journal of the IMA*, 2021.

Kevin McGoff, Sayan Mukherjee, and Andrew B. Nobel. Gibbs posterior convergence and the thermodynamic formalism. *The Annals of Applied Probability*, forthcoming.

Vrettos Moulos. Bicausal optimal transport for Markov chains via dynamic programming. In *IEEE International Symposium on Information Theory*, pages 1688–1693, 2021.

Michael Muskulus and Sjoerd Verduyn-Lunel. Wasserstein distances in the analysis of time series and dynamical systems. *Physica D: Nonlinear Phenomena*, 240(1):45–58, 2011.

Kevin O'Connor, Kevin McGoff, and Andrew B. Nobel. Estimation of stationary optimal transport plans. *arXiv preprint arXiv:2107.11858*, 2021a.

Kevin O'Connor, Bongsoo Yi, Kevin McGoff, and Andrew B. Nobel. Graph optimal transport with transition couplings of random walks. *arXiv preprint arXiv:2106.07106*, 2021b.

Nuria Oliver, Ashutosh Garg, and Eric Horvitz. Layered representations for learning and inferring office activity from multiple sensory channels. *Computer Vision and Image Understanding*, 96(2):163–180, 2004.

Donald S Ornstein. An application of ergodic theory to probability theory. *The Annals of Probability*, 1(1):43–58, 1973.

Gabriel Peyré and Marco Cuturi. Computational optimal transport. *Foundations and Trends in Machine Learning*, 11(5-6):355–607, 2019.

Aggelos Pikrakis, Sergios Theodoridis, and Dimitris Kamarotos. Classification of musical patterns using variable duration hidden Markov models. *IEEE Transactions on Audio, Speech, and Language Processing*, 14(5):1795–1807, 2006.

Martin Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* John Wiley & Sons Inc., 2005.

Lu Ren, David Dunson, Scott Lindroth, and Lawrence Carin. Dynamic nonparametric Bayesian models for analysis of music. *Journal of the American Statistical Association*, 105(490):458–472, 2010.

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

Tim Salimans, Dimitris Metaxas, Han Zhang, and Alec Radford. Improving GANs using optimal transport. In *International Conference on Learning Representations*, 2018.

Geoffrey Schiebinger, Jian Shu, Marcin Tabaka, Brian Cleary, Vidya Subramanian, Aryeh Solomon, Joshua Gould, Siyan Liu, Stacie Lin, Peter Berube, et al. Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell*, 176(4):928–943, 2019.

Paul Schweitzer. On undiscounted Markovian decision processes with compact action spaces. *RAIRO-Operations Research*, 19(1):71–86, 1985.

Richard Sinkhorn. Diagonal equivalence to matrices with prescribed row and column sums. *The American Mathematical Monthly*, 74(4):402–405, 1967.

Jinhua Song, Yang Gao, Hao Wang, and Bo An. Measuring the distance between finite Markov decision processes. In *International Conference on Autonomous Agents & Multiagent Systems*, pages 468–476, 2016.

Bing Su and Gang Hua. Order-preserving optimal transport for distances between sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(12):2961–2974, 2018.

A_P Varga and RK Moore. Hidden Markov model decomposition of speech and noise. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 845–848, 1990.

Cédric Villani. *Optimal Transport: Old and New*, volume 338. Springer Science & Business Media, 2008.

Michele Weiland, Alan Smaill, and Peter Nelson. Learning musical pitch structures with hierarchical hidden Markov models. *Journees d'Informatique Musical*, 2005.

Jonathan P Williams, Curtis B Storlie, Terry M Therneau, Clifford R Jack Jr, and Jan Hannig. A Bayesian approach to multistate hidden Markov models: application to dementia progression. *Journal of the American Statistical Association*, 115(529):16–31, 2020.

Alan Geoffrey Wilson. The use of entropy maximising models, in the theory of trip distribution, mode split and route split. *Journal of Transport Economics and Policy*, pages 108–126, 1969.

Hongteng Xu, Wenlin Wang, Wei Liu, and Lawrence Carin. Distilled Wasserstein learning for word embedding and topic modeling. In *Advances in Neural Information Processing Systems*, pages 1716–1725, 2018.

Jonathan P Yamron, Ira Carp, Larry Gillick, Steve Lowe, and Paul van Mulbregt. A hidden Markov model approach to text segmentation and event tracking. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 1, pages 333–336, 1998.

Anna K Yanchenko and Sayan Mukherjee. Classical music composition using state space models. *arXiv preprint arXiv:1708.03822*, 2017.

Danila A Zaev. On the Monge–Kantorovich problem with additional linear constraints. *Mathematical Notes*, 98(5-6):725–741, 2015.

Shaoyi Zhang. Existence and application of optimal Markovian coupling with respect to non-negative lower semi-continuous functions. *Acta Mathematica Sinica*, 16(2):261–270, 2000.

Walter Zucchini, Iain L MacDonald, and Roland Langrock. *Hidden Markov Models for Time Series: An Introduction Using R*. CRC press, 2017.