Sample-based Optimal Transport and Barycenter Problems

MAX KUANG
New York University, Courant Institute of Mathematical Sciences

AND

ESTEBAN G. TABAK
New York University, Courant Institute of Mathematical Sciences

Abstract

A methodology is developed for the numerical solution to the sample-based optimal transport and Wasserstein barycenter problems. The procedure is based on a characterization of the barycenter and of the McCann interpolants that permits the decomposition of the global problem under consideration into various local problems where the distance among successive distributions is small. These local problems can be formulated in terms of feature functions, and shown to have a unique minimizer that solves a nonlinear system of equations. Both the theoretical underpinnings of the methodology and its practical implementation are developed, and illustrated with synthetic and real data sets.

1 Introduction

The optimal transport problem, as proposed originally by Monge in 1781 [26], addresses the displacement of a pile of soil between two locations with minimal cost. Given the cost $c(\mathbf{x},\mathbf{y})$ of moving a unit mass from point $\mathbf{x}$ to point $\mathbf{y}$, one seeks the map $\mathbf{y}=\mathbf{f}(\mathbf{x})$ that minimizes its integral. After normalizing the two piles so that each has total mass one, they can be regarded as probability measures, and the problem adopts the form

$$\inf_{\mathbf{f}_\sharp\mu=\nu}\int c(\mathbf{x},\mathbf{f}(\mathbf{x}))d\mu(\mathbf{x}), \tag{1.1}$$

where μ and ν are the source and target measures, and $\mathbf{f}_\sharp\mu$ denotes the pushforward measure of μ by the map $\mathbf{f}$.

In the 20th century, Kantorovich [22] relaxed Monge's problem, allowing the movement of soil from one to multiple locations and vice versa. Denoting the mass transported from $\mathbf{x}$ to $\mathbf{y}$ by $\pi(\mathbf{x},\mathbf{y})$, the minimization problem can be rewritten as

$$\inf_{\pi\in\Pi_{\mu,\nu}}\int c(\mathbf{x},\mathbf{y})\pi(\mathbf{x},\mathbf{y})d\mathbf{x}\,d\mathbf{y}, \tag{1.2}$$

Communications on Pure and Applied Mathematics, Vol. 000, 0001–0052 (2000)

where $\Pi_{\mu,\nu}$ is the set of all transfer plans $\pi(x,y)$ satisfying the marginal constraints

$$\int \pi(\mathbf{x},\mathbf{y})d\,\mathbf{y} = \mu(\mathbf{x}),$$
$$\int \pi(\mathbf{x},\mathbf{y})d\,\mathbf{x} = \nu(\mathbf{y}).$$

Since the second half of the 20th century, mathematical properties of the optimal transport solution have been studied extensively. Here we introduce only those developments that are relevant to the methodology of this article. For a comprehensive review, we refer the reader to [45].

The optimal transport with L_p cost function has been particularly well-studied. The induced total transportation cost defines the Wasserstein distance [22],

$$W_p(\mu,\nu) = \left(\inf_{\pi\in\Pi_{\mu,\nu}} \int d(\mathbf{x},\mathbf{y})^p d\pi(\mathbf{x},\mathbf{y})\right)^{\frac{1}{p}}. \tag{1.3}$$

This distance provides a natural metric in the space of probability measures, which is important from a theoretical perspective and also because it arises naturally in many practical problems. A recent development in the field is the Wasserstein barycenter [2], which extends the concept of barycenter to probability measures.

The depth of the mathematical theory of optimal transport and its natural connection to many practical problems has attracted the attention of researchers in various fields. In economics, optimal transport is used in optimal assignment and matching problems (see [15]), for which the Kantorovich dual theorem provides a direct link between optimality and the equilibrium of matching between two populations. In image processing, the optimal transport distance is known as the earth mover's distance [39], used in many applications such as color transfer [34] and image segmentation [30]. Optimal transport is also used in medical research. For instance, it was found in [40] that the Ollivier-Ricci curvature computed through 1-Wasserstein distance is a proxy for robustness of gene expression networks, which can be applied to the characterization of cancer. The Wasserstein distance also appears in machine learning applications [5, 27] as a good measure of the distance between probability distributions.

Computational algorithms have been proposed to solve optimal transport and barycenter problems in a variety of settings. We reference below some recent representatives of the various approaches taken. Benamou and Brenier [6] introduced a computational fluid approach to solve the problem with continuous distributions $\mu_{1,2}$, exploiting the structure of the interpolant of the optimal map to solve the PDE corresponding to the dual optimization problem. Oberman [31] discretized the given continuous distributions and solved the resulting linear programming problem in an adaptive way that exploits the sparse nature of the solution (the fact that the optimal plan has support on a map). In image processing applications, several approaches have been proposed to regularize the discrete linear optimization problem of the earth mover's distance. The Entropy regularization approach [42] adds

an entropy term which leads to efficient algorithms to derive new solutions. It was proposed in [14] to add a graph regularization term to generate more regular solutions. Data-driven formulations take as input not the distributions $\mu_{1,2}$ but sample sets from both. Methodologies proposed include a fluid-flow-like algorithm [43] and an adaptive linear programming approach [9].

In this article, we propose a new family of algorithms to solve sample-based optimal transport and barycenter problems. With finite sample sets as input, the sample-based optimal transport algorithm (SOT) finds the optimal map from one sample set to the other and the sample-based barycenter algorithm (SBC) generates sample points from the barycenter of multiple sample sets. These algorithms solve optimization problems directly defined at the level of the sample sets. Instead of modeling the distributions underlying the data, we model the mapping functions as gradients of feature functions, where the features capture the structure of the sample sets and can be customized by the users. These algorithms, grounded on a necessary and sufficient property that characterizes the barycenter measure and on the McCann interpolants [25], reduce general optimal transport and barycenter problems to "local" optimal transport problems.

This article is organized as follows: section 2 introduces the theoretical formulation of the algorithms. Starting with a key property of barycenter measures and their connection with the McCann interpolants, we introduce a number of "theoretical" algorithms (theoretical because they assume that the distributions defining the problem are given and that a "black box" solver is provided for the simpler problems into which the full problem can be decomposed.) We then develop the concept of L-descending maps to prove the convergence of the proposed algorithms in a general setting. Section 3 develops the practical version of the algorithms by re-formulating the continuous version of optimal transport in a sample-based setting. We propose to model optimal maps using the gradient of feature functions and discuss their selection in practice. Section 4 is devoted to numerical examples. First synthetic examples are used to assess the performance of the new algorithms and draw comparisons with existing ones, and then the sample-based algorithms are applied to the transfer of multimodal distributions and to color transfer and shape transformations. Finally, Section 5 summarizes the work, discusses its limitations and applications, and suggests possible extensions.

2 Solving the theoretical optimal transport and barycenter problems

This section reviews the theoretical formulation of the optimal transport and barycenter problems, introducing formulas to characterize and derive their solutions. Algorithms are proposed as a bridge from the theoretical continuous formulation to the sample-based discrete formulation, which is the topic of section 3.

2.1 Optimal transport and barycenter problems

Let μ_1 and μ_2 be in $P(\mathscr{X})$, the set of Borel probability measures on a Polish space $\mathscr{X}$. Optimal transport asks how to "optimally" move mass from μ_1 to μ_2. An optimal transport problem, requires the following two elements:

Definition 2.1 (Cost Function)**.** The cost function $c : \mathscr{X} \times \mathscr{X} \to \mathbb{R}$ represents the cost of moving a unit mass from location $\mathbf{x}$ to $\mathbf{y}$. For much of this paper, we will focus on the quadratic cost on $\mathbb{R}^d$, $c(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|^2$.

Definition 2.2 (Transfer Plan)**.** A measure $\pi \in P(\mathscr{X} \times \mathscr{X})$ is a *transfer plan* between μ_1 and μ_2 if for any Borel set $E \subset \mathscr{X}$,

$$\pi(E \times \mathscr{X}) = \mu_1(E), \quad \pi(\mathscr{X} \times E) = \mu_2(E). \tag{2.1}$$

We denote by Π_{μ_1,μ_2} the set of all transfer plans between μ_1 and μ_2.

The value of a transfer plan π at $(\mathbf{x},\mathbf{y})$ represents the amount of mass moved from location $\mathbf{x}$ to $\mathbf{y}$, so the total cost $C(\pi)$ of a transfer plan is given by

$$C(\pi) = \int_{\mathscr{X}} c(\mathbf{x},\mathbf{y})\, d\pi(\mathbf{x},\mathbf{y}). \tag{2.2}$$

We call a transfer plan π^* optimal if it minimizes the total transportation cost,

$$\pi^* = \underset{\pi \in \Pi_{\mu_1,\mu_2}}{\operatorname{argmin}} C(\pi). \tag{2.3}$$

This is the **Monge-Kantorovich problem**, for which Kantorovich proved the following duality theorem (theorem 5.10 in [45]):

If $c(x,y)$ is a lower semi-continuous function such that

$$\forall \mathbf{x},\mathbf{y} \in \mathscr{X}, \quad c(\mathbf{x},\mathbf{y}) \geq a(\mathbf{x}) + b(\mathbf{y}), \tag{2.4}$$

for some real-valued upper semicontinuous functions $a \in L^1(\mu_1)$ and $b \in L^1(\mu_2)$, then the following duality principle holds:

(2.5)

$$\min_{\pi \in \Pi(\mu_1,\mu_2)} \int_{\mathscr{X}} c(\mathbf{x},\mathbf{y}) d\pi(\mathbf{x},\mathbf{y}) = \sup_{\substack{(\phi,\psi) \in L^1(\mu_1) \times L^1(\mu_2) \\ \phi + \psi \leq c}} \int_{\mathscr{X}} \phi(\mathbf{x}) d\mu_1(\mathbf{x}) + \int_{\mathscr{X}} \psi(\mathbf{y}) d\mu_2(\mathbf{y}).$$

The two dual functions ϕ and ψ satisfy

$$\phi(\mathbf{x}) = \inf_{\mathbf{y}} \left(c(\mathbf{x},\mathbf{y}) - \psi(\mathbf{y})\right), \quad \psi(\mathbf{y}) = \inf_{\mathbf{x}} \left(c(\mathbf{x},\mathbf{y}) - \phi(\mathbf{x})\right). \tag{2.6}$$

In 1996, [16] investigated optimal transport problems with cost functions of the form

$$c(\mathbf{x},\mathbf{y}) = h(\mathbf{x}-\mathbf{y}), \quad \text{where } h(\mathbf{x}) \text{ is a strictly convex function.} \tag{2.7}$$

Under minor constraints, they proved that the optimal transfer plan is unique and is induced by a map $\mathbf{s}$ that makes μ_1's pushforward measure $\mathbf{s}_\sharp \mu_1$ equals to μ_2. The optimal transfer plan can then be written as

$$\pi = (\mathbf{id} \times \mathbf{s})_\sharp \mu_1, \tag{2.8}$$

where **id** stands for the identity map. Moreover, the optimal map **s** is of the form

$$\mathbf{s}(\mathbf{x}) = \mathbf{x} - \nabla h^*(\nabla\phi(\mathbf{x})), \tag{2.9}$$

where h^* is the Legendre transform of h and ϕ is a c-concave function.

In the special case where $h(\mathbf{x}) = \|\mathbf{x}\|^2$, the statement above adopts the particularly simple form:

> *The optimal transfer plan of the quadratic optimal transport problem in $\mathbb{R}^d$ is induced by a unique optimal map, given by the gradient of a convex function.*

This convex function $\phi(\mathbf{x})$ satisfies the Monge-Ampere equation

$$d\mu_2(\nabla\phi(\mathbf{x}))\det(\nabla^2\phi(\mathbf{x})) = d\mu_1(\mathbf{x}), \tag{2.10}$$

a PDE that can be interpreted as

$$\int_{\mathbb{R}^d} h(\nabla\phi(\mathbf{x}))d\mu_1(\mathbf{x}) = \int_{\mathbb{R}^d} h(\mathbf{y})d\mu_2(\mathbf{y}) \tag{2.11}$$

for all continuous functions h [12].

For the quadratic cost function, the optimal transfer cost $C(\pi)$ is the square of the 2-Wasserstein distance

$$W_2(\mu_1,\mu_2) = \left(\min_{\pi\in\Pi_{\mu_1,\mu_2}} \int_{\mathbb{R}^d} \|\mathbf{x}-\mathbf{y}\|^2 d\pi(\mathbf{x},\mathbf{y})\right)^{\frac{1}{2}}, \tag{2.12}$$

which provides a metric in the space of measures

$$P_2(\mathbb{R}^d) := \left\{\mu \in P(\mathbb{R}^d); \quad \int_{\mathbb{R}^d} \|\mathbf{x}\|^2 d\mu(\mathbf{x}) < +\infty\right\}. \tag{2.13}$$

From now on, unless otherwise specified, we will consider the quadratic cost function and measures in $P_2(\mathscr{X})$, with $\mathscr{X} = \mathbb{R}^d$.

The McCann interpolant is defined as a continuous family of measures $\mu(t)$ ($t \in [0,1]$) between μ_1 and μ_2. If **s** is the optimal map between μ_1 and μ_2, we define the McCann interpolant measure as

$$\mu(t) = [t\,\mathbf{s} + (1-t)\mathbf{id}]_\sharp \mu_1. \tag{2.14}$$

The induced map from ν_s to ν_t ($0 \le s,t \le 1$) is also the optimal map between them. This property, which parallels similar ones for the barycenter of a set of measures, is essential to the algorithms that we will propose.

The barycenter problem is a relatively recent development in the field of optimal transport. Consider first the weighted barycenter of a set of points in $\mathscr{X}$. Given the points $\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_K$ and positive weights $w_1,w_2\cdots,w_K$ ($\sum w_k = 1$), their weighted barycenter $\bar{\mathbf{x}} = \sum_{k=1}^K w_k\mathbf{x}_k$ can be characterized through

$$\bar{\mathbf{x}} = \operatorname*{argmin}_{\mathbf{x}} \sum_{k=1}^K w_k \|\mathbf{x}_k - \mathbf{x}\|^2. \tag{2.15}$$

Moreover, for any $\mathbf{y}$,

$$\sum_{k=1}^{K} w_k \|\mathbf{x}_k - \mathbf{y}\|^2 - \sum_{k=1}^{K} w_k \|\mathbf{x}_k - \bar{\mathbf{x}}\|^2 = \|\mathbf{y} - \bar{\mathbf{x}}\|^2. \tag{2.16}$$

Similarly, the barycenter of a set of measures $\mu_1, \mu_2, \cdots, \mu_K \in P_2(\mathscr{X})$ with weights $w_1, w_2, \cdots, w_K$ is the minimizer of the following problem [2]:

$$\bar{\mu} = \underset{\nu \in P_2(\mathscr{X})}{\operatorname{argmin}} \sum_{k=1}^{K} w_k W_2^2(\mu_k, \nu). \tag{2.17}$$

A duality result similar to the one for the optimal transport problem holds:

$$\min_{\nu \in P_2(\mathscr{X})} \sum_{k=1}^{K} w_k W_2^2(\mu_k, \nu) = \sup_{\sum_{k=1}^{K} \phi_k(\mathbf{x}_k) \le \sum_{k=1}^{K} w_k \|\mathbf{x}_k - \bar{\mathbf{x}}\|^2} \sum_{k=1}^{K} \int_{\mathscr{X}} \phi_k(\mathbf{x}) d\mu_k(\mathbf{x}), \tag{2.18}$$

where $\bar{\mathbf{x}}$ is the barycenter of the K vectors $\mathbf{x}_k$. Still another equivalent form of the barycenter problem is the multidimentional formulation [2, 17]:

$$\min_{\hat{\pi} \in \Pi_{\mu_1, \mu_2, \cdots, \mu_K}} \int_{\mathscr{X}^K} \left(\sum_{k=1}^{K} w_k \|\mathbf{x}_k - \bar{\mathbf{x}}\|^2 \right) d\hat{\pi}(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_K), \tag{2.19}$$

where $\Pi_{\mu_1, \mu_2, \cdots, \mu_K}$ is the set of measures $\hat{\pi}(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_K)$ with kth marginal μ_k. The equivalence of (2.17) and (2.19) is established in [2].

A special case is the two-measure barycenter problem,

$$\min_{\nu \in P_2(\mathscr{X})} w W_2^2(\mu_1, \nu) + (1-w) W_2^2(\mu_2, \nu), \tag{2.20}$$

whose solution coincides with a McCann interpolant measure between μ_1 and μ_2: if $\mu(t)$ is the interpolant measures defined in (2.14), then

$$\mu(1-w) = \underset{\nu \in P_2(\mathscr{X})}{\operatorname{argmin}} \; w W_2^2(\mu_1, \nu) + (1-w) W_2^2(\mu_2, \nu). \tag{2.21}$$

2.2 A sequence of algorithms

This subsection introduces the main theoretical algorithms of this article informally, without proofs.

In order to address the numerical solution to the optimal transport and barycenter problems, we start with a simpler question: assuming that we already have a black box solver for the optimal transport problem, find an efficient algorithm that uses it to solve the barycenter problem.

Given an arbitrary initial measure $\nu \in P_2(\mathscr{X})$, we seek to update it so as to reduce the total cost in the barycenter problem (2.17). Constructing the optimal transfer maps from ν to the μ_k, which we denote by $\mathbf{s}_k$, we have

$$\sum_{k=1}^{K} w_k W_2^2(\nu, \mu_k) = \sum_{k=1}^{K} w_k \int_{\mathscr{X}} \|\mathbf{y} - \mathbf{s}_k(\mathbf{y})\|^2 d\nu(\mathbf{y}) = \int_{\mathscr{X}} \left[\sum_{k=1}^{K} w_k \|\mathbf{y} - \mathbf{s}_k(\mathbf{y})\|^2 \right] d\nu(\mathbf{y}). \tag{2.22}$$

Comparing the integrand of the last expression with the characterization of the pointwise barycenter in (2.15) suggests defining a transformation $\mathbf{f}$ that maps $\mathbf{y}$ to the barycenter of $\mathbf{s}_1(\mathbf{y}), \mathbf{s}_2(\mathbf{y}), \cdots, \mathbf{s}_K(\mathbf{y})$:

$$\mathbf{f}(\mathbf{y}) := \sum_{k=1}^{K} w_k \mathbf{s}_k(\mathbf{y}). \tag{2.23}$$

Pushing forward ν via $\mathbf{f}$, we have the following inequality[1]:

$$\begin{aligned} \sum_{k=1}^{K} w_k W_2^2(\nu, \mu_k) &\geq \int_{\mathscr{X}} \left[\sum_{k=1}^{K} w_k \| \mathbf{f}(\mathbf{y}) - \mathbf{s}_k(\mathbf{y}) \|^2 \right] d\nu(\mathbf{y}) \\ &= \int_{\mathscr{X}} \left[\sum_{k=1}^{K} w_k \| \tilde{\mathbf{y}} - \mathbf{s}_k(\mathbf{f}^{-1}(\tilde{\mathbf{y}})) \|^2 \right] d\mathbf{f}_\sharp \nu(\tilde{\mathbf{y}}) \\ &\geq \sum_{k=1}^{K} w_k W_2^2(\mathbf{f}_\sharp \nu, \mu_k). \end{aligned} \tag{2.24}$$

Applying repeatedly the map $\mathbf{f}$ yields Algorithm 1 [3].

Algorithm 1 Basic Theoretical Barycenter Algorithm (BTB)

(1) Set $\nu = \nu_0$ where $\nu_0 \in P_2(\mathscr{X})$ is an arbitrary initial measure;
(2) Find the optimal maps $\mathbf{s}_k$ from ν to μ_k $(k = 1, 2, \cdots, K)$;
(3) Define $\mathbf{f}$ using (2.23);
(4) Set $\nu = \mathbf{f}_\sharp \nu$ and return to step 2.

Notice that, for the updated measure $\mathbf{f}_\sharp \nu$, the map $\mathbf{s}_k(\mathbf{f}^{-1}(\tilde{\mathbf{y}}))$ between $\mathbf{f}_\sharp \nu$ and μ_k may not be optimal anymore. Hence at every iteration of the BTB algorithm, one needs to recompute the optimal maps between ν and μ_k using the black box optimal transport solver, which can make the complexity of the BTB algorithm quite high.

The same idea can be used to solve a standard optimal transport problem. In the last section, we pointed out the connection between the two-measure barycenter and McCann's interpolant. Applying BTB to a two measure barycenter problem, if the algorithm converges, we will get a McCann Interpolant measure, which also gives us the optimal map $\mathbf{s}_2 \circ \mathbf{s}_1^{-1}$ between μ_1 and μ_2.

It may seem unnecessary to use the BTB algorithm for the optimal transport problem, as this can be solved directly using the black box solver which BBT requires. The algorithm's utility becomes clearer when instead of the barycenter problem we consider the following alternative:

$$\min_{\substack{\nu_0, \nu_1, \cdots, \nu_K \in P_2(\mathscr{X}) \\ \nu_0 = \mu_1, \nu_K = \mu_2}} \sum_{k=1}^{K} w_k W_2^2(\nu_k, \nu_{k-1}). \tag{2.25}$$

[1] Even though $\mathbf{f}$ is assumed to be invertible for this calculation, we prove in section 2.5 a general version for which this assumption is not required.

Since $\nu_0, \nu_1, \cdots, \nu_K$ form a chain between μ_1 and μ_2 and the Wasserstein distance is a metric on $P_2(\mathscr{X})$, Cauchy's inequality yields

$$\left[\sum_{k=1}^{K} w_k W_2^2(\nu_k, \nu_{k-1})\right]\left[\sum_{k=1}^{K} \frac{1}{w_k}\right] \geq \left[\sum_{k=1}^{K} W_2(\nu_k, \nu_{k-1})\right]^2 \geq W_2^2(\mu_1, \mu_2), \tag{2.26}$$

with both equalities attained if and only if $\nu_0, \nu_2, \cdots, \nu_K$ are McCann interpolant measures with the specific time parameters:

$$\lambda_k = \sum_{i=1}^{k} \frac{1}{w_i} \Bigg/ \sum_{i=1}^{K} \frac{1}{w_i}, \tag{2.27}$$

$$\nu_k = \mu(\lambda_k). \tag{2.28}$$

If $\mathbf{s}_k$ is the optimal map from ν_{k-1} to ν_k and $(\nu_0, \nu_1, \cdots, \nu_K)$ is the minimizer of (2.25), the optimal map from μ_1 to μ_2 is

$$\mathbf{s}^c = \mathbf{s}_K \circ \mathbf{s}_{K-1} \circ \cdots \circ \mathbf{s}_1, \tag{2.29}$$

thus leading to the introduction of Algorithm 2 (TOT).

Algorithm 2 Theoretical Optimal Transport Algorithm (TOT)

(1) Let $\nu_0 = \mu_1$ and $\nu_K = \mu_2$. Set ν_k $(k = 1, 2, \cdots, K-1)$ to arbitrary initial measures in $P_2(\mathscr{X})$;
(2) Find the optimal maps $\mathbf{s}_k$ from ν_{k-1} to ν_k $(k = 1, 2, \cdots, K)$;
(3) Define $\mathbf{s}^c$ using (2.29) and the weights λ_k using (2.27);
(4) For $k = 1, 2, \cdots, K-1$, update ν_k to

$$\nu_k = [\lambda_k \mathbf{s}^c + (1-\lambda_k)\mathbf{id}]_\sharp \mu_1 \tag{2.30}$$

and return to step 2.

To find the optimal map from μ_1 to μ_2, the TOT algorithm approximates not only the optimal map itself but also the McCann interpolant measures. Notice that we still need to solve optimal transport problems between ν_{k-1} and ν_k at every iteration. However, by choosing enough interpolant measures, TOT can always make ν_{k-1} and ν_k as close to each other as needed. The solution to these "local" optimal transport problems turns out to be far easier to approximate. Hence TOT provides a feasible way to solve arbitrary optimal transport problems assuming that one can solve local optimal transport problems.

Next we can merge the TOT algorithm above into the barycenter algorithm, using it as a black box solver. Moreover, instead of solving an optimal transport problem at every step, we can simply update the map using one iteration of TOT every step, which gives the following composite theoretical barycenter algorithm 3.

Algorithm 3 Composite Theoretical Barycenter Algorithm (CTB)

(1) Set $\nu = \nu_0$ where $\nu_0 \in P_2(\mathscr{X})$ is an arbitrary initial measure;
(2) Run steps 1,2,3 of the TOT Algorithm 2 once to find a map $\mathbf{s}_k^c$ from ν to μ_k for each $k = 1, 2, \cdots, K$;
(3) Define the map $\mathbf{f}^c$ as:

$$\mathbf{f}^c(\mathbf{y}) = \sum_{k=1}^{K} w_k \mathbf{s}_k^c(\mathbf{y}); \tag{2.31}$$

(4) Update ν and $\mathbf{s}_k^c$:

$$\nu = \mathbf{f}^c_\sharp \nu \tag{2.32}$$

$$\mathbf{s}_k^c = \mathbf{s}_k^c \circ (\mathbf{f}^c)^{-1}; \tag{2.33}$$

(5) Run step 4, 2, 3 of TOT once for each pair (ν, μ_k) to update $\mathbf{s}_k^c$ and return to step 3.

The CTB algorithm 3 solves the barycenter problem under the single assumption that one knows how to solve local optimal transport problems, a problem that will be addressed in sub-section 3.2.

None of the three theoretical algorithms introduced in this section, BTB, TOT and CTB, can be applied to realistic settings directly, since typically in practice one does not know the continuous distributions that define the problem. Instead, the input data often consists of sample sets drawn from these distributions. Yet before discussing how to apply these algorithms in practical scenarios, we prove below their convergence.

2.3 A necessary and sufficient characterization of the barycenter

In the BTB algorithm 1, we perform two basic operations: finding the optimal map between pairs (ν, μ_k) and finding the barycenter of sets of points. The algorithm stops when both of the following conditions are satisfied:

(1) For any $k = 1, 2, \cdots, K$, $\mathbf{s}_k$ is the optimal map from ν to μ_k;
(2) For any $\mathbf{y} \in \mathscr{X}$,

$$\mathbf{y} = \sum_{k=1}^{K} w_k \mathbf{s}_k(\mathbf{y}). \tag{2.34}$$

The next theorem shows these two conditions are both necessary and sufficient for ν to be the barycenter of the μ_k:

Theorem 2.3 ([2])**.** *Consider the barycenter problem* (2.17) *with absolutely continuous measures $\mu_1, \mu_2, \cdots, \mu_K$ and positive weights $w_1, w_2, \cdots, w_K$ ($\sum_{k=1}^K w_k = 1$). A measure ν is the barycenter if and only if for almost all $\mathbf{y} \in \mathrm{supp}(\nu)$,*

$$\mathbf{y} = \sum_{k=1}^{K} w_k \mathbf{s}_k(\mathbf{y}) \tag{2.35}$$

where $\mathbf{s}_k$ is the optimal map from ν to μ_k.

Proof. To start with, the existence of optimal map $\mathbf{s}_k$ is guaranteed by the regularity of ν (ν will be an absolutely continuous measure), which is proved in [2].

For necessity, we use the inequality in (2.24),

$$\sum_{k=1}^{K} w_k W_2^2(\nu,\mu_k) - \sum_{k=1}^{K} w_k W_2^2(\mathbf{f}_\sharp \nu,\mu_k) \geq \int_{\mathscr{X}} \|\mathbf{y}-\mathbf{f}(\mathbf{y})\|^2 d\nu(\mathbf{y}). \tag{2.36}$$

Since ν is the minimizer, it follows that $\mathbf{y}=\mathbf{f}(\mathbf{y})$ ν-almost everywhere.

For sufficiency, let $(\phi_k^*(\mathbf{x}),\psi_k^*(\mathbf{y}))$ be the optimizer of the dual optimal transport problem (2.5) between ν and μ_k. To unify notations, we set the cost function to be $w_k\|\mathbf{x}-\mathbf{y}\|^2$ instead of the usual $\|\mathbf{x}-\mathbf{y}\|^2$. Then the constraints in the dual problems can be written as

$$\phi_k^*(\mathbf{x}) + \psi_k^*(\mathbf{y}) \leq w_k\|\mathbf{x}-\mathbf{y}\|^2. \tag{2.37}$$

Equation (2.9) and related properties imply the following relation:

$$\mathbf{s}_k(\mathbf{y}) = \nabla\left[\frac{\|\mathbf{y}\|^2}{2} - \frac{\psi_k^*(\mathbf{y})}{w_k}\right], \tag{2.38}$$

which combined with (2.35) yields

$$\mathbf{y} = \sum_{k=1}^{K} w_k \mathbf{s}_k(\mathbf{x}) = \sum_{k=1}^{K} w_k \nabla\left[\frac{\|\mathbf{y}\|^2}{2} - \frac{\psi_k^*(\mathbf{y})}{w_k}\right] = \mathbf{y} - \nabla\sum_{k=1}^{K}\psi_k^*(\mathbf{y}). \tag{2.39}$$

This indicates that $\sum_{k=1}^{K}\psi_k^*(\mathbf{y}) = C$ is a constant, which we can set to zero without changing the optimal map. Thus

$$\sum_{k=1}^{K}\int \phi_k^*(\mathbf{x})d\mu_k(\mathbf{x}) = \sum_{k=1}^{K}\left[\int \phi_k^*(\mathbf{x})d\mu_k(\mathbf{x}) + \int \psi_k^*(\mathbf{y})d\nu(\mathbf{y})\right]. \tag{2.40}$$

From the strong duality property of the optimal transport problem,

$$\int \phi_k^*(\mathbf{x})d\mu_k(\mathbf{x}) + \int \psi_k^*(\mathbf{y})d\nu(\mathbf{y}) = w_k W_2^2(\mu_k,\nu), \tag{2.41}$$

which summed up over all k yields

$$\sum_{k=1}^{K}\int \phi_k^*(\mathbf{x})d\mu_k(\mathbf{x}) = \sum_{k=1}^{K} w_k W_2^2(\mu_k,\nu). \tag{2.42}$$

Hence $\{\phi_k^*(\mathbf{x})\}$ and ν satisfy the strong duality of the barycenter problem, implying that ν is the barycenter of the μ_k. □

This theorem is first introduced in [2]. We propose the above proof in the spirit of the barycenter algorithm. The theorem provides an easy check on whether a measure is the barycenter. A direct corollary of this theorem corresponding to the TOT algorithm is

Theorem 2.4. *Consider problem* (2.25) *with absolutely continuous measures* μ_1, μ_2 *measures and positive weights* $w_1, w_2, \cdots, w_K$ ($\sum_{k=1}^K w_k = 1$)*. Absolutely continuous measures* $(\nu_0, \nu_1, \cdots, \nu_K)$ *minimize* (2.25) *if and only if for almost all* $\mathbf{y} \in \text{supp}(\mu_1)$ *and all* $k = 1, 2, \cdots, K$*,*

$$\mathbf{s}_k \circ \cdots \circ \mathbf{s}_2 \circ \mathbf{s}_1(\mathbf{y}) = (1-\lambda_k)\mathbf{y} + \lambda_k \mathbf{s}^c(\mathbf{y}) \tag{2.43}$$

where $\mathbf{s}_k$ *is the optimal map from* ν_{k-1} *to* ν_k *and* $\mathbf{s}^c$ *is defined by* (2.29)*.*

Proof. Necessity follows directly from (2.26): for all inequalities to become equalities, ν_k has to be on the geodesics between μ_1 and μ_2, thus must be absolutely continuous themselves. With the existence of $\mathbf{s}_k$, the condition in the theorem must be satisfied.

On the other hand, it follows from the conditions that, for $k = 1, 2, \cdots, K-1$,

$$\mathbf{s}_k \circ \cdots \circ \mathbf{s}_2 \circ \mathbf{s}_1(\mathbf{y}) = (1 - \frac{\lambda_k}{\lambda_{k+1}})\mathbf{y} + \frac{\lambda_k}{\lambda_{k+1}} \mathbf{s}_{k+1} \circ \cdots \circ \mathbf{s}_2 \circ \mathbf{s}_1(\mathbf{y}). \tag{2.44}$$

Theorem 2.3 for $k = 2, 3, \cdots, K$ implies that $\mathbf{s}_k \circ \cdots \circ \mathbf{s}_2 \circ \mathbf{s}_1$ is the optimal map from μ_1 to ν_k. Specifically, when $k = K$, $\mathbf{s}^c$ is the optimal map between μ_1 and μ_2, thus ν_k is indeed a McCann interpolant measure with time variable λ_k, which proves that $(\nu_0, \nu_1, \cdots, \nu_K)$ is the minimizer of (2.25). □

The above two theorems characterize the barycenter and the McCann interpolant measures. While necessity is quite straightforward, sufficiency is particularly useful, as it provides a way to define the barycenter and the McCann interpolant. To prove the convergence of the algorithms proposed in the prior subsection, we will first prove that they converge to measures that satisfy the above properties. Using these theorems, it is then guaranteed that these limit measures are indeed the barycenter and McCann interpolant measures respectively.

2.4 A general convergence theorem for L-descending maps

To prove the convergence of the proposed algorithms, we will first introduce a set of more general convergence theorems that can be applied to BTB, TOT, CTB and to a broader class of algorithms. We first define the following concept:

Definition 2.5 (L-Descending Map)**.** Suppose $\mathscr{F}$ is a map between measures:

$$\mathscr{F} : P_2(\mathscr{X}) \to P_2(\mathscr{X}) \tag{2.45}$$

and $L : P_2(\mathscr{X}) \to \mathbb{R}$ is a cost function with a lower bound. We call $\mathscr{F}$ an *L-descending map* if it satisfies the following conditions:

(1) $\mathscr{F}$ is a continuous map with respect to the W_2 metric on $P_2(\mathscr{X})$.
(2) For arbitrary $\pi \in P_2(\mathscr{X})$, we have

$$L(\mathscr{F}\pi) \leq L(\pi) \tag{2.46}$$

and equality holds if and only if

$$\mathscr{F}\pi = \pi = \pi^* \tag{2.47}$$

where π^* is a minimizer of L.

Consider an iterative algorithm seeking the minimizer of a cost function L. If each iteration updates the current measure through a map, it is natural to require that the map reduces the value of L. At the same time, a necessary condition for the algorithm to succeed is that when the algorithm can no longer reduce the value of L, the minimizer must have been reached. Combining these two conditions and the continuity of the map, we get exactly the above definition of *L-descending maps*. Next we prove a general convergence theorem.

Theorem 2.6. *Let $\mathscr{F}$ be an L-descending map. We define a sequence $\{\pi^n\}_{n=1,2,...}$ through*

$$\pi^{n+1} = \mathscr{F}\,\pi^n, \quad n = 0, 1, ... \tag{2.48}$$

where $\pi^0 \in P_2(\mathscr{X})$ is an arbitrary initial measure.

If the following two conditions are satisfied:

(1) $\{\pi^n\}_{n=1,2,...}$ is sequentially compact with respect to the W_2 metric,
(2) L is continuous with respect to the W_2 metric and it has a unique minimizer π^*,

then

$$\pi^n \xrightarrow{W_2} \pi^* \tag{2.49}$$

Proof. Because $\{\pi^n\}_{n=1,2,...}$ is sequentially compact, we can extract a subsequence that converges to some measure π':

$$\pi^{n_j} \xrightarrow{W_2} \pi'. \tag{2.50}$$

From the continuity of $\mathscr{F}$ and L, we have that

$$\mathscr{F}\,\pi^{n_j} \xrightarrow{W_2} \mathscr{F}\,\pi', \tag{2.51}$$

and

$$\lim_{j\to+\infty} L(\pi^{n_j}) = L(\pi') \tag{2.52}$$

$$\lim_{j\to+\infty} L(\mathscr{F}\,\pi^{n_j}) = L(\mathscr{F}\,\pi'). \tag{2.53}$$

On the other hand, because $L(\pi^n)$ is a non-increasing sequence,

$$L(\mathscr{F}\,\pi') = \lim_{n\to+\infty} L(\mathscr{F}\,\pi^{n_j}) = \lim_{n\to+\infty} L(\pi^{n_j}) = L(\pi'). \tag{2.54}$$

From the definition of *L-descending maps*, π' must be the unique minimizer of L. The W_2 convergence of the full sequence follows from the fact that any subsequence of $\{\pi^n\}_{n=1,2,...}$ W_2 converges to the same measure π^*. □

2.5 Convergence of the optimal transport and barycenter algorithms

The general convergence theorem for *L-descending maps* helps us prove a series of convergence results, for which we only need to verify the conditions in Theorem 2.6. Consider first the following general algorithm:

Definition 2.7 (General Descending Optimal Transport Algorithm (GDOT))**.** Consider an iterative algorithm that updates a transfer plan $\pi \in \Pi_{\mu_1,\mu_2}$:

$$\pi^{n+1} = \mathscr{F}\,\pi^n, \quad n = 0,1,\cdots \tag{2.55}$$

We call an algorithm a general descending optimal transport algorithm if the associated iteration map $\mathscr{F}$ is an L-descending map in which the cost function $L(\pi)$ is the total transport cost defined in (2.2).

Theorem 2.8. *Assume that the two target measures μ_1 and μ_2 are in $P_2(\mathscr{X})$ and the optimal transfer plan is unique. Then the GDOT algorithm converges to the optimal transfer plan π^* in the W_2 metric.*

Proof. In order to use Theorem 2.6 to prove the W_2 convergence of $\{\pi^n\}$, we verify the two conditions in Theorem 2.6.

To prove that $\{\pi^n\}_{n=1,2,\cdots}$ is sequentially compact with respect to W_2 metric, we prove that Π_{μ_1,μ_2} is tight. First, $\Pi_{\mu_1,\mu_2} \subset P_2(\mathscr{X}^2)$, because

$$\int_{\mathscr{X}^2} (\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) d\pi = \int_{\mathscr{X}} \|\mathbf{x}\|^2 d\mu_1 + \int_{\mathscr{X}} \|\mathbf{y}\|^2 d\mu_2 < \infty. \tag{2.56}$$

Then, for any $\varepsilon > 0$, we can choose compact sets $K_\varepsilon^{\mu_1}$ and $K_\varepsilon^{\mu_2}$ such that

$$\mu_1(\mathscr{X} \setminus K_\varepsilon^{\mu_1}) < \frac{\varepsilon}{2}, \quad \mu_2(\mathscr{X} \setminus K_\varepsilon^{\mu_2}) < \frac{\varepsilon}{2}. \tag{2.57}$$

It follows that, for any $\pi \in \Pi_{\mu_1,\mu_2}$,

$$\pi(\mathscr{X}^2 \setminus K_\varepsilon^{\mu_1} \times K_\varepsilon^{\mu_2}) \le \mu_1(\mathscr{X} \setminus K_\varepsilon^{\mu_1}) + \mu_2(\mathscr{X} \setminus K_\varepsilon^{\mu_2}) < \varepsilon. \tag{2.58}$$

By Prokhorov's theorem [35], $\{\pi^n\}_{n=1,2,\cdots}$ has a subsequence $\{\pi^{n_j}\}_{j=1,2,\cdots}$ that converges weakly to a measure π'. Since the marginal measures of π^{n_j} are the same as those of π''s,

$$\int_{\mathscr{X}^2} (\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) d\pi^{n_j} = \int_{\mathscr{X}} \|\mathbf{x}\|^2 d\mu_1 + \int_{\mathscr{X}} \|\mathbf{y}\|^2 d\mu_2 = \int_{\mathscr{X}^2} (\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) d\pi' \tag{2.59}$$

By Theorem 6.9 of [45], the convergence of quadratic functions indicates that $\pi^{n_j} \xrightarrow{W_2} \pi'$. Thus $\pi^n \in \Pi_{\mu_1,\mu_2}$ indeed has a W_2 convergent subsequence.

For condition 2 of Theorem 2.6, we need to prove that the cost function $C(\pi)$ is continuous with respect to the W_2 metric. Since the quadratic cost can be controlled by quadratic terms:

$$\|\mathbf{x}-\mathbf{y}\|^2 \le 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2), \tag{2.60}$$

if $\pi^n \xrightarrow{W_2} \pi$, by the definition of W_2 convergence, we have that

$$\lim_{n\to\infty} C(\pi^n) = \lim_{n\to\infty} \int_{\mathscr{X}^2} \|\mathbf{x}-\mathbf{y}\|^2 d\pi^n = \int_{\mathscr{X}^2} \|\mathbf{x}-\mathbf{y}\|^2 d\pi = C(\pi). \tag{2.61}$$

Since both conditions are satisfied, Theorem 2.6 shows that $\{\pi^n\}_{n=1,2,\ldots}$ converges to the optimal transfer plan π^*. □

This theorem provides a general framework for proving convergence for optimal transport algorithms. As long as one can show that the map in the algorithm is an *L-descending map* and that the optimal map is unique, convergence is guaranteed. Notice that the theorem does not require the transfer plan to be a map.

The simplest *L-descending map* is the one that maps every transfer plan to the optimal transfer plan. As we will see later, the map in TOT is also an *L-descending map*. Before going into specific cases, we build a general framework for the barycenter problem.

Since the barycenter measure is not a transfer plan, in order to prove convergence to the barycenter measure ν we need to consider the multimarginal measure $\hat{\pi}$. Informally, in every iteration after step 2 of the BTB algorithm, we have a family of optimal transfer plans between the barycenter and the target measures:

$$\pi_k = \text{law}(\mathbf{y}, \mathbf{s}_k(\mathbf{y})), \quad k = 1,2,\cdots,K. \tag{2.62}$$

Combining all the maps together yields the multimarginal measure $\hat{\pi}$:

$$\hat{\pi} = \text{law}(\mathbf{s}_1(\mathbf{y}), \mathbf{s}_2(\mathbf{y}), \cdots, \mathbf{s}_K(\mathbf{y})). \tag{2.63}$$

Instead of thinking of the update of ν at every iteration, we will view the barycenter algorithms as updating the $\hat{\pi}$ measure. Clearly $\hat{\pi}$ stores more information than ν, which can be derived from $\hat{\pi}$ though

$$\nu(E) := \hat{\pi}\left(\{(\mathbf{x}_1,\cdots,\mathbf{x}_K) | \sum_{k=1}^K w_k \mathbf{x}_k \in E\}\right), \text{ for all Borel sets } E. \tag{2.64}$$

However, when we update general transfer plans π_k (not transfer maps), the multimarginal measure is not fully determined by the transfer plans: when we fix a point $\mathbf{y}$ in measure ν, the conditional measure $\pi_k|\mathbf{y}$ in general does not concentrate in one location, as it does when we have a transfer map. Thus for the joint conditional measure $\hat{\pi}|\mathbf{y}$, only its marginals are defined by $\pi_k|\mathbf{y}$. To fully define $\hat{\pi}$, we need to either only consider transfer maps or specify $\hat{\pi}$ for general π_k.

Even if we restrict attention to transfer maps, as in BTB and CTB, we still need the map $\mathbf{f}$ to be invertible in order to fully determine $\hat{\pi}$. In [3], the authors proved the invertibility of $\mathbf{f}$ in BTB in a similar setting, but we found that this property might not hold for the CTB algorithm. Moreover, to build a more general convergence theorem, we would like to consider general transfer plans.

In order to fully define $\hat{\pi}$, we build the conditional measure $\hat{\pi}|\mathbf{y}$ as the product measure of the $\pi_k|\mathbf{y}$:

$$\hat{\pi}|\mathbf{y} = \pi_1|\mathbf{y} \times \pi_2|\mathbf{y} \times \cdots \times \pi_K|\mathbf{y}, \tag{2.65}$$

thus making the conditional marginal measures independent of each other. Then we define a general descending barycenter algorithm:

Definition 2.9 (General Descending Barycenter Algorithm (GDB))**.** Given an arbitrary *L-descending map* $\mathscr{F}$ for the GDOT algorithms, we define the map:

$$\mathscr{G} : \Pi_{\mu_1,\cdots,\mu_2,\cdots,\mu_K} \to \Pi_{\mu_1,\cdots,\mu_2,\cdots,\mu_K} \tag{2.66}$$

and use it to update the current multimarginal measure $\hat{\pi}^n$:

$$\hat{\pi}^{n+1} = \mathscr{G}\,\hat{\pi}^n. \tag{2.67}$$

We call such an algorithm a general descending barycenter algorithm if $\mathscr{G}$ is induced by $\mathscr{F}$ through the following steps:

(1) Derive v^n from $\hat{\pi}^n$ using (2.64);

(2) For $k = 1, 2, \cdots, K$, set the current transfer plan π_k^n to

$$\pi_k^n(E_v \times E_k) := v(E_v) \cdot \hat{\pi}^n\left(\{(\mathbf{x}_1, \cdots, \mathbf{x}_K) | \, \mathbf{x}_k \in E_k\}\right), \text{ for all Borel sets } E_v, E_k; \tag{2.68}$$

(3) Apply the map $\mathscr{F}$ to all π_k^n,

$$\pi_k^{n+1} = \mathscr{F}\,\pi_k^n \tag{2.69}$$

and define the conditional measure $\pi_k^{n+1}|\,\mathbf{y}$;

(4) Set $\hat{\pi}^{n+1}$ as the product measure:

$$\hat{\pi}^{n+1}(E_1 \times E_2 \times \cdots \times E_K) = \int_{\mathscr{X}} \left[\prod_{k=1}^{K} \pi_k^{n+1}|\,\mathbf{y}(E_k)\right] dv(\mathbf{y}). \tag{2.70}$$

The general descending barycenter algorithm is a generalization of the BTB and CTB algorithms that updates transfer plans instead of maps. The optimal transport solver is not specified in GDB: for any *L-descending map* $\mathscr{F}$ for the GDOT algorithm, we can define a corresponding GDB algorithm.

Theorem 2.10. *The general descending barycenter algorithm has the following properties:*

(I) $\mathscr{G}$ *is a L-descending map, where the cost function L is the barycenter cost in the multimarginal formulation* (2.19)*:*

$$L(\hat{\pi}) = \int_{\mathscr{X}^K} \left(\sum_{k=1}^{K} w_k \|\,\mathbf{x}_k - \bar{\mathbf{x}}\|^2\right) d\hat{\pi}(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_K), \quad \hat{\pi} \in \Pi_{\mu_1,\mu_2,\cdots,\mu_K}; \tag{2.71}$$

(II) If the barycenter problem has a unique solution, the GDB algorithm always converges to the minimizer $\hat{\pi}^*$ *of the* $L(\hat{\pi})$ *defined above; the corresponding* v^* *is the barycenter.*

Proof of (I). To prove that $\mathscr{G}$ is an *L-descending map*, we verify the conditions in Definition 2.5.

First, we prove that $\mathscr{G}$ is continuous. Let a sequence of measures $\{\hat{\pi}^n\}_{n=1,2,\cdots}$ converge to some $\hat{\pi}^*$ in the W_2 metric. We would like to show that

$$\mathscr{G}\,\hat{\pi}^n \xrightarrow{W_2} \mathscr{G}\,\hat{\pi}^*. \tag{2.72}$$

We first show that the corresponding ν^n converges to ν^* in W_2. By definition (2.64), $\nu(y)$ can be viewed as $\hat{\pi}$'s marginal measure for $\mathbf{y} = \sum_{k=1}^K w_k \mathbf{x}_k$. For any continuous function ϕ such that $|\phi(\mathbf{y})| \le C(1+\|\mathbf{y}\|^2)$,

$$|\phi(\mathbf{y})| \le C(1+\|\mathbf{y}\|^2) \le C(1+\sum_{k=1}^K w_k \|\mathbf{x}_k\|^2) \le \tilde{C}(1+\sum_{k=1}^K \|\mathbf{x}_k\|^2). \tag{2.73}$$

Since

$$\int_{\mathscr{X}} \phi(\mathbf{y}) d\nu(\mathbf{y}) = \int_{\mathscr{X}^K} \phi(\sum_{k=1}^K w_k \mathbf{x}_k) d\hat{\pi}, \tag{2.74}$$

the W_2 convergence of $\hat{\pi}^n$ yields

$$\lim_{n\to\infty} \int_{\mathscr{X}} \phi(\mathbf{y}) d\nu^n(\mathbf{y}) = \int_{\mathscr{X}} \phi(\mathbf{y}) d\nu^*(\mathbf{y}), \tag{2.75}$$

the W_2 convergence of ν^n. The joint measure of ν^n and $\hat{\pi}^n$ also converges, and so

$$\pi_k^n \xrightarrow{W_2} \pi_k^*. \tag{2.76}$$

Since $\mathscr{F}$ is an *L-descending map*,

$$\mathscr{F}\,\pi_k^n \xrightarrow{W_2} \mathscr{F}\,\pi_k^*. \tag{2.77}$$

Finally, since $\mathscr{G}\,\hat{\pi}^n$ is defined by the product of the $\mathscr{F}\,\pi_k^n$'s conditional measures,

$$\mathscr{G}\,\hat{\pi}^n \xrightarrow{W_2} \mathscr{G}\,\hat{\pi}^*. \tag{2.78}$$

Next we prove that $\mathscr{G}$ reduces the cost function L.

$$
\begin{aligned}
L(\hat{\pi}^n) &= \int_{\mathscr{X}^K} \left[\sum_{k=1}^K w_k \| \mathbf{x}_k - \bar{\mathbf{x}} \|^2 \right] d\hat{\pi}^n(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_K) \\
&= \sum_{k=1}^K w_k \int_{\mathscr{X}^2} \| \mathbf{x}_k - \mathbf{y} \|^2 d\pi_k^n(\mathbf{x}_k, \mathbf{y}) \\
&\geq \sum_{k=1}^K w_k \int_{\mathscr{X}^2} \| \mathbf{x}_k - \mathbf{y} \|^2 d\pi_k^{n+1}(\mathbf{x}_k, \mathbf{y}) \qquad (2.79) \\
&= \sum_{k=1}^K w_k \int_{\mathscr{X}} \int_{\mathscr{X}} \| \mathbf{x}_k - \mathbf{y} \|^2 d\left(\pi_k^{n+1} | \mathbf{y}\right)(\mathbf{x}_k)\, d\nu^n(\mathbf{y}) \\
&= \int_{\mathscr{X}} \int_{\mathscr{X}^K} \left[\sum_{k=1}^K w_k \| \mathbf{x}_k - \mathbf{y} \|^2 \right] d\left(\hat{\pi}^{n+1} | \mathbf{y}\right)(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_K)\, d\nu^n(\mathbf{y}) \\
&\geq \int_{\mathscr{X}} \int_{\mathscr{X}^K} \left[\sum_{k=1}^K w_k \| \mathbf{x}_k - \bar{\mathbf{x}} \|^2 \right] d\left(\hat{\pi}^{n+1} | \mathbf{y}\right)(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_K)\, d\nu^n(\mathbf{y}) \qquad (2.80) \\
&= \int_{\mathscr{X}^K} \left[\sum_{k=1}^K w_k \| \mathbf{x}_k - \bar{\mathbf{x}} \|^2 \right] d\hat{\pi}^{n+1}(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_K) \\
&= L(\hat{\pi}^{n+1}).
\end{aligned}
$$

If $L(\hat{\pi}^n) = L(\hat{\pi}^{n+1})$, from the inequality turned equality in (2.79), we have that

$$
\int_{\mathscr{X}^2} \| \mathbf{x}_k - \mathbf{y} \|^2 d\pi_k^n(\mathbf{x}_k, \mathbf{y}) = \int_{\mathscr{X}^2} \| \mathbf{x}_k - \mathbf{y} \|^2 d\pi_k^{n+1}(\mathbf{x}_k, \mathbf{y}). \qquad (2.81)
$$

Since $\mathscr{F}$ is an *L-descending map*, it follows that π_k^n must be the optimal transfer plan between μ_k and ν^n. The difference between the two sides of the inequality turned equality in (2.80) yields

$$
\int_{\mathscr{X}} \int_{\mathscr{X}^K} \| \bar{\mathbf{x}} - \mathbf{y} \|^2 d\hat{\pi}^{n+1} | \mathbf{y}(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_K) d\nu^n(\mathbf{y}) = 0, \qquad (2.82)
$$

implying that $\bar{\mathbf{x}} = \mathbf{y}$ almost everywhere. So $\nu^n = \nu^{n+1}$ and $\hat{\pi}^n = \hat{\pi}^{n+1}$.

Now using theorem 2.3, we have that this fixed point ν^n must be the barycenter of ν. Moreover $\hat{\pi}^n$ must be the minimizer of $L(\hat{\pi}^n)$. This concludes the proof that $\mathscr{G}$ is an *L-descending map*. □

Proof of (II). Having shown that $\mathscr{G}$ is an *L-descending map*, the proof of the second part follows the same steps as the proof of Theorem 2.8. □

With the above general convergence theorems 2.8, 2.10 for the GDOT and GDB algorithms, the convergence of the specific algorithms proposed in the previous section are simple corollaries: the proofs boil down to showing that the associated maps in the iterative algorithm are *L-descending maps*.

When the BTB, TOT and CTB algorithms were introduced, the transfer plans were limited to transfer maps. As we have discussed, a more accurate and general definition should be made in terms of general transfer plans, which we do here for completeness:

Algorithm 4 Basic Theoretical Barycenter Algorithm (BTB)

(1) Set $\nu = \nu_0$ where $\nu_0 \in P_2(\mathscr{X})$ is an arbitrary initial measure;
(2) Find the optimal transfer plan π_k between ν and μ_k ($k = 1, 2, \cdots, K$);
(3) Define the multimarginal measure $\hat{\pi}$ from π_k;
(4) Derive ν from $\hat{\pi}$ using (2.64) and go to step 2.

Algorithm 5 Theoretical Optimal Transport Algorithm (TOT)

(1) Let $\nu_0 = \mu_1$ and $\nu_K = \mu_2$. Set ν_k ($k = 1, 2, \cdots, K-1$) to arbitrary initial measures in $P_2(\mathscr{X})$;
(2) Find the optimal transfer plan π_k between ν_{k-1} to ν_k ($k = 1, 2, \cdots, K$);
(3) Derive the transfer plan $\hat{\pi}$ between μ_1 and μ_2 from π_k by (2.84) and define λ_k by (2.27);
(4) For $k = 1, 2, \cdots, K-1$, update ν_k to

$$\nu_k(E) := \hat{\pi}\left(\{(\mathbf{x}, \mathbf{y}) | \lambda_k \mathbf{y} + (1-\lambda_k)\mathbf{x} \in E\}\right), \text{ for all Borel sets } E \tag{2.83}$$

and go to step 2.

Algorithm 6 Composite Theoretical Barycenter Algorithm (CTB)

(1) Set $\nu = \nu_0$ where $\nu_0 \in P_2(\mathscr{X})$ is an arbitrary initial measure;
(2) Run step 1,2,3 of TOT once to find a transfer map π_k from ν to μ_k for each $k = 1, 2, \cdots, K$;
(3) Define the multimarginal measure $\hat{\pi}$ from π_k;
(4) Update ν using (2.64) and derive new π_k from $\hat{\pi}$;
(5) Run step 4, 2, 3 of TOT once for each pair (ν, μ_k) to update π_k and go to step 3.

One crucial step in the TOT algorithm is the derivation of the transfer plan $\hat{\pi}$ from the local transfer plans π_k. As for the barycenter multimarginal measure case, with transfer plans more general than maps, we need to specify how to couple the transfer plans together. As in the barycenter case, we couple the conditional measures together independently. Consider the process from $\mathbf{x}$ to $\mathbf{y}$ through $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{K-1}$. For each k, the conditional measure $\pi_k | \mathbf{x}_{k-1}$ can be thought of as a Markov kernel from ν_{k-1} to ν_k. Then one can derive the Markov kernel from μ_1 to μ_2 by compositing all the conditional measures:

$$\hat{\pi}(E_{\mathbf{x}}, E_{\mathbf{y}}) := \int_{E_{\mathbf{x}}} \int_{\mathscr{X}} \cdots \int_{\mathscr{X}} \int_{E_{\mathbf{y}}} \left[\prod_{k=1}^{K} d\pi_k | \mathbf{x}_{k-1}(\mathbf{x}_k) \right] d\mu_1(\mathbf{x}) \tag{2.84}$$

Now we prove the convergence theorems for all the algorithms proposed:

Theorem 2.11 (Convergence of Theoretical Optimal Transport Algorithm)**.** *Consider the quadratic optimal transport problem with measures μ_1 and μ_2. The map $\mathscr{F}_{TOT}$ associated with the TOT algorithm is an L-descending map and TOT always converges to the optimal transfer plan.*

Proof. To show that $\mathscr{F}_{\text{TOT}}$ is an *L-descending map*, we verify all the conditions in definition 2.5.

We first show the continuity of $\mathscr{F}_{\text{TOT}}$. Suppose we have a sequence of transfer plans $\{\pi^n\}_{n=1,2,\cdots}$ converging to π^*. For all k, ν_k^n will also converge to ν_k^* in the W_2 metric. By Theorem 5.20 in [45] and the uniqueness of the optimal transfer plan in our case, we have, for the local transfer plans π_k^n:

$$\pi_k^n \xrightarrow{W_2} \pi_k^*. \tag{2.85}$$

By the construction of $\mathscr{F}_{\text{TOT}}\,\hat{\pi}^n$ from π_k^n (2.84), we have

$$\mathscr{F}_{\text{TOT}}\,\hat{\pi}^n \xrightarrow{W_2} \mathscr{F}_{\text{TOT}}\,\hat{\pi}^*, \tag{2.86}$$

proving continuity. For the descending property of $\mathscr{F}_{\text{TOT}}$, we directly compute the cost function. Defining

$$\mathbf{x}_k = \lambda_k \mathbf{y} + (1-\lambda_k)\mathbf{x}, \tag{2.87}$$

we have that

$$\mathbf{x}_{k-1} - \mathbf{x}_k = (\lambda_k - \lambda_{k-1})(\mathbf{x}-\mathbf{y}), \tag{2.88}$$

and thus

$$\int_{\mathscr{X}^2} \|\mathbf{x}-\mathbf{y}\|^2 d\hat{\pi}^n(\mathbf{x},\mathbf{y}) = \sum_{k=1}^{K} \int_{\mathscr{X}^2} (\lambda_k - \lambda_{k-1}) \|\mathbf{x}-\mathbf{y}\|^2 d\hat{\pi}^n(\mathbf{x},\mathbf{y}) \tag{2.89}$$

$$= \sum_{k=1}^{K} \int_{\mathscr{X}^2} \frac{\|\mathbf{x}_{k-1}-\mathbf{x}_k\|^2}{\lambda_k - \lambda_{k-1}} d\pi_k^n(\mathbf{x}_{k-1},\mathbf{x}_k) \tag{2.90}$$

$$\geq \sum_{k=1}^{K} \int_{\mathscr{X}^2} \frac{\|\mathbf{x}_{k-1}-\mathbf{x}_k\|^2}{\lambda_k - \lambda_{k-1}} d\pi_k^{n+1}(\mathbf{x}_{k-1},\mathbf{x}_k). \tag{2.91}$$

Following the definition of $\hat{\pi}^{n+1}$, we can write the summation of integrals above as one integral over the joint measure $\tilde{\pi}^{n+1}(\mathbf{x},\mathbf{x}_1,\cdots,\mathbf{x}_{K-1},\mathbf{y})$,

(2.92)

$$\begin{aligned}\int_{\mathscr{X}^2}\|\mathbf{x}-\mathbf{y}\|^2 d\hat{\pi}^n(\mathbf{x},\mathbf{y}) &\geq \sum_{k=1}^{K}\int_{\mathscr{X}^2}\frac{\|\mathbf{x}_{k-1}-\mathbf{x}_k\|^2}{\lambda_k-\lambda_{k-1}}d\pi_k^{n+1}(\mathbf{x}_{k-1},\mathbf{x}_k)\\ &= \int_{\mathscr{X}^K}\sum_{k=1}^{K}\frac{\|\mathbf{x}_{k-1}-\mathbf{x}_k\|^2}{\lambda_k-\lambda_{k-1}}d\tilde{\pi}_k^{n+1}(\mathbf{x},\mathbf{x}_1,\cdots,\mathbf{x}_{K-1},\mathbf{y})\\ &\geq \int_{\mathscr{X}^K}\frac{\|\sum_{k=1}^{K}\mathbf{x}_{k-1}-\mathbf{x}_k\|^2}{\sum_{k=1}^{K}\lambda_k-\lambda_{k-1}}d\tilde{\pi}_k^{n+1}(\mathbf{x},\mathbf{x}_1,\cdots,\mathbf{x}_{K-1},\mathbf{y}) \qquad (2.93)\\ &= \int_{\mathscr{X}^K}\|\mathbf{x}-\mathbf{y}\|^2 d\tilde{\pi}_k^{n+1}(\mathbf{x},\mathbf{x}_1,\cdots,\mathbf{x}_{K-1},\mathbf{y})\\ &= \int_{\mathscr{X}^2}\|\mathbf{x}-\mathbf{y}\|^2 d\hat{\pi}^{n+1}(\mathbf{x},\mathbf{y})\end{aligned}$$

This shows the transport cost is not increasing. If it is not decreasing either, we check the inequalities turned into equalities in the above calculation for the conditions that $\hat{\pi}^n$ must satisfy. From (2.92), we have that π_k^n must be the optimal transfer plan between ν_{k-1} and ν_k. From (2.93) and Cauchy's inequality, we have that

$$\frac{\mathbf{x}_{k-1}-\mathbf{x}_k}{\lambda_k-\lambda_{k-1}} = \mathbf{x}-\mathbf{y}, \qquad (2.94)$$

a relation equivalent to (2.44). By Theorem 2.4, we have that $\hat{\pi}^n$ must be the optimal transfer plan between μ_1 and μ_2.

This concludes the proof that $\mathscr{F}_{\text{TOT}}$ is an *L-descending map*. Then Theorem 2.8 shows that the TOT algorithm converges to the optimal transfer plan. □

Theorem 2.12 (Convergence of the Theoretical Barycenter Algorithms)**.** *Consider the barycenter problem* (2.17) *with measures* $\mu_1,\mu_2,...,\mu_K$ *and positive weights* $w_1,w_2,...,w_K$ ($\sum_{k=1}^{K}w_k=1$)*. We have*

(I) The BTB algorithm is a special case of the GDB algorithm with the L-descending map that maps all the transfer plan to the corresponding optimal transfer plan. Thus it converges to the barycenter measure.

(II) The CTB Algorithm is a special case of the GDB algorithm with the L-descending map $\mathscr{F}_{TOT}$*. Thus it converges to the barycenter measure.*

Proof. It follows straightforwardly from the definitions of BTB and CTB that both are special cases of GDB. □

2.6 Conclusions of section 2

This section proposed three theoretical algorithms to solve the optimal transport and barycenter problems. These algorithms perform iterations that reduce the

corresponding cost function. The iterations are *L-descending maps*, so the general convergence theorem Theorem 2.6 provides the basic framework to show their convergence.

For the optimal transport problem, the TOT algorithm approximates the McCann interpolant measures and updates the transfer plans for each of its segments. For the barycenter problem, both the BTB and CTB algorithms alternate between updating the barycenter measure via point-wise barycenter problems and updating the transfer plans between the barycenter and each target measure. A shared feature of these algorithms is that they both alternate between two different operations to reduce the cost function. The sufficiency conditions proved in Theorem 2.3 and Theorem 2.4 guarantee that these algorithms converge to the optimal transfer plan and barycenter measure respectively.

The convergence theorems Theorem 2.8 and Theorem 2.10 show the convergence of a broad family of optimal transport and barycenter algorithms. To apply these theorems, one only needs to verify that the associated map in an algorithm is an *L-descending map*.

Using the TOT and CBC algorithms, one can solve optimal transport and barycenter problems defined by arbitrary $P_2(\mathscr{X})$ measures, assuming that one can solve "local" optimal transport problems. In many cases, these local optimal transport problems are relatively easy to solve. For instance, when two measures are close to each other, the gradient flow [4] offers a good approximation to the optimal transfer plan. Local optimal transport problems also play an important role in the sample-based formulation described below.

3 The sample-based optimal transport and barycenter problems

The prior section discussed optimal transport and barycenter problems when the data is provided as a set of measures. This setting, which connects with the classical theory of optimal transport, enabled us to propose theoretical algorithms. However, the classical setting is not enough for practical data analysis, since more often than not one has access not to the actual conditional distributions of the data but only to samples drawn from these distributions.

The plan of this section is as follows. In the first subsection, we discuss different practical formulations of the optimal transport problem and introduce the sample-based Monge-Ampere equation and the sample-based optimal transport and barycenter problems. In the second subsection, we prove a local solution theorem for the sample-based formulation. Then we adapt all the theoretical algorithms to their sample-based forms using the new local solver in the third subsection. Finally we discuss the selection of the feature functions that the solver requires.

3.1 The sample-based formulation

The starting point of statistics/data analysis is data, which can be modeled as a set of independent realizations of random variables from unknown underlying

distributions. Hence, in order to use optimal transport as a tool for data analysis, one must reformulate the problem in terms of datasets in $\mathbb{R}^d$ instead of known distributions.

Let $\{\mathbf{x}_i\}_{i=1}^{N_x}$ and $\{\mathbf{y}_i\}_{i=1}^{N_y}$ be i.i.d. samples from the unknown distributions μ_1 and μ_2 in $\mathbb{R}^d$, respectively. We would like to define a quadratic optimal map between the two sample sets and develop practical algorithms to find it.

We will focus our attention on optimal maps rather then general transfer plans. On the one hand, the solution to the classical quadratic optimal transport problem is a map. On the other, when transferring a finite sample set, one would expect the result to also be a finite sample set. The simplest such transform consists of moving each sample point to a new sample point, as maps do.

To formulate the optimal transport problem in this setting, we need to give the sample-based version of the following two statements:

(1) A map $\mathbf{f}$ pushes measure μ_1 to μ_2:

$$\mathbf{f}_\sharp \mu_1 = \mu_2 \tag{3.1}$$

(2) A map $\mathbf{f}$ minimizes the transport cost:

$$\int c(\mathbf{x}, \mathbf{f}(\mathbf{x})) d\mu_1(\mathbf{x}) \tag{3.2}$$

One relevant setting is the assignment problem in combinatorial optimization, which seeks a one-to-one mapping between $\{\mathbf{x}_i\}_{i=1}^N$ and $\{\mathbf{y}_i\}_{i=1}^N$ minimizing

$$\sum_{i=1}^{N} \| \mathbf{x}_i - \mathbf{y}_{\sigma(i)} \|^2, \tag{3.3}$$

where σ is a permutation of $\{1, 2, \cdots, N\}$. There are broad applications of the assignment problem [8] and various practical algorithms [29] to solve it. Yet the assignment problem has significant differences with the scenario we are interested in. The most important one is that we do not restrict the images of $\{\mathbf{x}_i\}_{i=1}^N$ to be among the $\{\mathbf{y}_i\}_{i=1}^N$: since the sample points are random variables drawn from some continuous distributions, the range of the map should be the whole support of the continuous distribution, not a finite subset of points. Another limitation of the assignment problem is that it requires both sets to have the same number of elements, which is not a requirement of our problem.

Heuristically, the optimal map between $\{\mathbf{x}_i\}_{i=1}^{N_x}$ and $\{\mathbf{y}_i\}_{i=1}^{N_y}$ should be similar to the classical optimal map between the two measures μ_1 and μ_2. One possible way to proceed would define the sample-based optimal map in two steps: one would obtain density estimates $\tilde{\mu}_1$ and $\tilde{\mu}_2$ from sample the sets $\{\mathbf{x}_i\}_{i=1}^{N_x}$ and $\{\mathbf{y}_i\}_{i=1}^{N_y}$, and then solve the classical optimal transport problem between $\tilde{\mu}_1$ and $\tilde{\mu}_2$.

If one knows which parametric classes are well-suited to approximate μ_1 and μ_2 and the sample size is large enough, one can derive accurate estimators for μ_1 and μ_2. For instance, if one knows that the sample points are drawn from normal

distributions, given enough samples one can obtain good approximations to their means and covariance matrices, and hence to the distributions themselves. Since the optimal transport problem between normal distributions has a closed-form solution [18], one can then define it as the solution to the sample-based problem between $\{\mathbf{x}_i\}_{i=1}^{N_x}$ and $\{\mathbf{y}_i\}_{i=1}^{N_y}$.

However, it is generally difficult both to select a parametric model for the unknown density functions and to do an accurate density estimation with limited sample size. These concerns, common to many statistical problems, make density estimation a not very favorable route, especially for problems in high dimensions. Moreover, even if provided with good density estimates, we would still need to solve the classical optimal transport problem numerically, since there is no closed-form solution for the multidimensional problem between general measures.

A popular formulation of the optimal transport problem is the discrete formulation. Suppose the underlying space $\mathscr{X}$ is discrete and the two distribution functions are $\{\mu_i\}_{i=1}^{N}$ and $\{\rho_j\}_{j=1}^{N}$. The Kantorovich formulation (2.2) can be written as

$$\min_{\substack{\sum_{i=1}^{N}\pi_{ij}=\mu_j \\ \sum_{j=1}^{N}\pi_{ij}=\rho_i}} \sum_{i,j=1}^{N} c_{ij}\pi_{ij}, \tag{3.4}$$

also known in image processing as the earth mover's problem [33]. This fits well imaging applications in which the input (an image) can be naturally represented as a discrete distribution. Yet this is not true for general random sample sets. One disadvantage of the discrete form (3.4) is that its optimal transfer plan is not necessarily a map as in the classical case. On the other hand, to apply the discrete form to our problem, we need to properly discretize the problem first. As discussed in Section 4.2, this is not always a trivial task.

To summarize, properly converting the sample-based problem to either classical problems or to a discrete formulation are nontrivial tasks. In this paper, we will instead formulate the problem directly in terms of sample sets.

To define an optimal map $\mathbf{f}$ directly from the two finite sample sets $\{\mathbf{x}_i\}_{i=1}^{N_x}$ and $\{\mathbf{y}_i\}_{i=1}^{N_y}$, we first need to redefine the constraint $\mathbf{f}_\sharp\mu_1 = \mu_2$ so that it makes sense for samples. The statement that $\mathbf{f}$ pushes μ_1 to μ_2 can be informally paraphrased as $\{\mathbf{f}(\mathbf{x}_i)\}_{i=1}^{N_x}$ and $\{\mathbf{y}_i\}_{i=1}^{N_y}$ are drawn from the same distribution. Since the underlying distribution is unknown and hard to estimate, we would like to develop an equivalence relation that can be directly verified from sample sets.

We propose the following equivalence relation that compares sample sets through feature functions:

Definition 3.1 (sample-based equivalence)**.** We say that a sample set $\{\mathbf{x}_i\}_{i=1}^{N_x}$ is sample-based equivalent to $\{\mathbf{y}_i\}_{i=1}^{N_y}$ with respect to a set of feature functions $\{F_j(\mathbf{x})\}_{j=1}^{M}$:

$\mathbb{R}^d \to \mathbb{R}$ if

$$\frac{1}{N_x}\sum_{i=1}^{N_x} F_j(\mathbf{x}_i) = \frac{1}{N_y}\sum_{j=1}^{N_y} F_j(\mathbf{y}_i), \quad j = 1,2,\cdots,M \tag{3.5}$$

We denote this equivalence relation by

$$\{\mathbf{x}_i\}_{i=1}^{N_x} \sim \{\mathbf{y}_i\}_{i=1}^{N_y} \tag{3.6}$$

Similar ideas have appeared in different contexts to compare measures and sample sets, for instance in the MaxEnt framework of [7] and in the constrained density estimation technique of [24].

One interpretation of the sample-based equivalence arises from the Monge-Ampere equation (2.11). If we replace the test function $h(\mathbf{x})$ in (2.11) by a feature function and substitute the integrals over measures by the corresponding empirical means, we obtain:

$$\frac{1}{N_x}\sum_{i=1}^{N_x} F_j(\nabla\phi(\mathbf{x}_i)) = \frac{1}{N_y}\sum_{i=1}^{N_y} F_j(\mathbf{y}_i), \quad j = 1,2,\cdots,M \tag{3.7}$$

which is equivalent to $\{\nabla\phi(\mathbf{x}_i)\}_{i=1}^{N_x} \sim \{\mathbf{y}_i\}_{i=1}^{N_y}$.

Using the above as constraints on the function $\phi(\mathbf{x})$, the sample-based optimal transport problem can be cast as the following optimization problem:

$$\min_{\{\nabla\phi(\mathbf{x}_i)\}_{i=1}^{N_x} \sim \{\mathbf{y}_i\}_{i=1}^{N_y}} \sum_{i=1}^{N_x} \|\nabla\phi(\mathbf{x}_i) - \mathbf{x}_i\|^2, \tag{3.8}$$

In the above problem we substitute the transportation cost by the empirical cost on the sample set. A closer look at (3.8) reveals that some further specifications are needed. If the function space for ϕ is not further constrained, the values of $\nabla\phi(\mathbf{x}_i)$ for different $\mathbf{x}_i$ are uncorrelated. In addition, for locations $\mathbf{x} \notin \{\mathbf{x}_i\}_{i=1}^{N_x}$, the value of $\nabla\phi(\mathbf{x})$ is not specified or controlled by the optimization problem (3.8). To solve this problem, we constrain ϕ to lie in a finite-dimensional space:

$$\phi(\mathbf{x}) = \frac{\|\mathbf{x}\|^2}{2} + \sum_{j=1}^{M} s_j F_j(\mathbf{x}). \tag{3.9}$$

The $\frac{\|\mathbf{x}\|^2}{2}$ term is included so that when all the s_j are zeros $\phi(\mathbf{x})$ corresponds to the identity map.

To summarize, we define below the sample-based Monge-Ampere equation and the sample-based optimal transport problem:

Definition 3.2 (Sample-based Monge-Ampere Equation)**.** For two sample sets $\{\mathbf{x}_i\}_{i=1}^{N_x}$ and $\{\mathbf{y}_i\}_{i=1}^{N_y}$ and a given set of C^1 feature functions $\{F_j(\mathbf{x})\}_{j=1}^{M}$, we say a function $\phi(\mathbf{x})$ is a solution to the sample-based Monge-Ampere equation if $\phi(\mathbf{x})$ satisfies (3.9) and $\{\nabla\phi(\mathbf{x}_i)\}_{i=1}^{N_x} \sim \{\mathbf{y}_i\}_{i=1}^{N_y}$.

Definition 3.3 (Sample-based Optimal Transport)**.** For two sample sets $\{\mathbf{x}_i\}_{i=1}^{N_x}$ and $\{\mathbf{y}_i\}_{i=1}^{N_y}$ and a given set of feature functions $\{F_j(\mathbf{x})\}_{j=1}^{M}$, the optimal map $\mathbf{f}$ from $\{\mathbf{x}_i\}_{i=1}^{N_x}$ to $\{\mathbf{y}_i\}_{i=1}^{N_y}$ is defined as $\mathbf{f}(\mathbf{x}) = \nabla\phi^*(\mathbf{x})$, where ϕ^* is the solution to the sample-based Monge-Ampere equation that minimizes

$$\sum_{i=1}^{N_x} \|\nabla\phi(\mathbf{x}_i) - \mathbf{x}_i\|^2. \tag{3.10}$$

Expanding ϕ as a linear combination of the feature functions as base functions can be interpreted as applying a finite element method to the classical Monge-Ampere equation (2.10). There is a clear correspondence between the classical and sample-based Monge-Ampere equation: when the two measures μ_1 and μ_2 are known, the classical solution must satisfy (2.11) for all continuous functions h. When we only have finite sets of samples $\{\mathbf{x}_i\}_{i=1}^{N_x}$ and $\{\mathbf{y}_i\}_{i=1}^{N_y}$, the sample-based solution must lie in a prescribed finite dimensional space and satisfy (3.7) for all feature functions.

As in the classical setting, once the sample-based optimal transport problem is defined, we can extend it to the barycenter problem:

Definition 3.4 (Sample-based barycenter problem)**.** Given K sample sets $\{\mathbf{x}_i^k\}_{i=1}^{N_k}$, $k = 1, 2, \cdots, K$, positive weights $w_1, w_2, \cdots, w_K$, $(\sum_{k=1}^{K} w_k = 1)$ and a fixed parameter N_y, we call a sample set $\{\mathbf{y}_i\}_{i=1}^{N_y}$ the barycenter if it minimizes

$$\sum_{k=1}^{K} w_k \sum_{i=1}^{N_y} \|\nabla\phi_k(\mathbf{y}_i) - \mathbf{y}_i\|^2, \tag{3.11}$$

where $\{\nabla\phi_k(\mathbf{y}_i)\}_{i=1}^{N_y} \sim \{\mathbf{x}_i^k\}_{i=1}^{N_k}$ for $k = 1, 2, \cdots, K$.

Instead of a barycenter measure ν, in the sample-based formulation the target is a sample set $\{\mathbf{y}_i\}_{i=1}^{N_y}$, depending on the parameter N_y, the sample size of the barycenter.

If we set $N_y = 1$ and consider the single feature function

$$\mathbf{F}(\mathbf{x}) = \mathbf{x}, \tag{3.12}$$

the solution is given by

$$\mathbf{y}_1 = \sum_{k=1}^{K} w_k \frac{1}{N_k} \sum_{i=1}^{N_k} \mathbf{x}_i^k, \tag{3.13}$$

the weighted average of the barycenters of the given sample sets. With $N_y > 1$ and more general feature functions, the solution captures more detailed structures in the data and becomes harder to solve. We discuss properties of the sample-based formulation in the next subsections.

3.2 The solution of the local sample-based optimal transport problem

As in the classical Monge-Ampere equation, the sample-based version can have multiple solutions: even though there are exactly M unknowns $\{s_j\}_{j=1}^M$ to be solved from the M equations (3.7), the feature functions are typically nonlinear. To derive all the solutions and figure out which gives the smallest transportation cost is computationally intractable. In addition, nonlinearity can lead to the nonexistence of solutions.

However, when the two sample sets are close to each other, we will be able not only to guarantee the existence of a solution but also find the global minimizer of the objective function. We call this type of problems "local sample-based optimal transport problems".

To state the full theorem, we first introduce some notation:

Denote by $\mathbf{F}: \mathbb{R}^d \to \mathbb{R}^M$ the feature function vector:

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \cdots \\ f_M(\mathbf{x}) \end{pmatrix}. \tag{3.14}$$

We define the matrices $\mathbf{A}_i \in \mathbb{R}^{M\times d}, \mathbf{A} \in \mathbb{R}^{M\times dN_x}$, by

$$\mathbf{A}_i = \nabla \mathbf{F}(\mathbf{x}_i), \quad i = 1,2,\cdots,N_x \tag{3.15}$$

and

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_{N_x} \end{pmatrix}. \tag{3.16}$$

Denote the feature values of the two sample sets by column vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^M$:

$$\mathbf{a} = \frac{1}{N_x}\sum_{i=1}^{N_x} \mathbf{F}(\mathbf{x}_i), \quad \mathbf{b} = \frac{1}{N_y}\sum_{i=1}^{N_y} \mathbf{F}(\mathbf{y}_i). \tag{3.17}$$

Inspired by the sample-based Monge-Ampere equation, define $\mathbf{G}: \mathbb{R}^M \to \mathbb{R}^M$:

$$\mathbf{G}(\mathbf{s}) = \frac{1}{N_x}\sum_{i=1}^{N_x} \mathbf{F}(\mathbf{x}_i + \mathbf{s}^T \mathbf{A}_i). \tag{3.18}$$

Theorem 3.5. *Assume that* $\mathbf{A}$ *has full row rank. Then there exist an open set* U *around* $\mathbf{0}$ *and an open set* V *around* $\mathbf{a}$ *such that*

(I) $\mathbf{G}: U \to V$ *gives a bijection between* U *and* V*. Thus for all* $\mathbf{b} \in V$*, the solution to the sample-based Monge-Ampere equation exists and is unique.*

(II) For all $\mathbf{b} \in V$*, we can define a map* ϕ^* *through* (3.9) *with* $\mathbf{s}^* = \mathbf{G}^{-1}(\mathbf{b})$*. Then* $\phi^*(\mathbf{x})$ *gives the global minimum of the sample-based optimal transport problem.*

(III) For a fixed compact set $\mathscr{X} \subset \mathbb{R}^d$*, if all feature functions are* C^2 *functions in* $\mathscr{X}$*,* $\phi^*(\mathbf{x})$ *is convex in* $\mathscr{X}$*.*

Before giving the proof, we interpret the meaning of the theorem. Comparing to the general non-local case, Theorem 3.5 establishes the existence of the sample-based Monge-Ampere solution and states that the local solution is the minimizer of the nonlinear global optimization problem (3.10). In addition, the convexity of the optimal ϕ^* is also guaranteed when the feature functions' Hessian matrices are uniformly bounded, which corresponds to a key property of the solution to the classical quadratic optimal transport problem.

The compactness of set $\mathscr{X}$ in III can be substituted by other conditions. For instance, we can choose $\mathscr{X} = \mathbb{R}^d$ and enforce that all the feature functions' Hessian matrices are uniformly bounded in $\mathbb{R}^d$, from which we can prove the global convexity of the optimal ϕ^*. We ask $\mathscr{X}$ to be compact in the main theorem because the sample sets are finite, so it is reasonable to discuss properties of the optimal map in a compact set.

The first requirement of the theorem is that $\mathbf{A}$ be full row rank. One necessary condition for this is that

$$M < dN_x, \tag{3.19}$$

which loosely states that the number of feature functions should not exceed the number of sample points. If this is not satisfied, we are overfitting and no longer have Theorem 3.5.

The other requirement is that the two sample sets should be close to each other, measured by the feature vectors they yield. The closeness of sample sets coincides with the concept of "local optimal transport" introduced in the theoretical algorithm section. As we will see in the next section, using a solver for the local sample-based Monge-Ampere equation as the black box solver in the algorithms, we can adapt all the theoretical algorithms to solve practical sample-based problems.

The theorem guarantees the existence of an open set V. The size of V represents how close the feature vector $\mathbf{b}$ needs to be to $\mathbf{a}$ for the theorem to hold. An estimate of the size of V can be found by combining the estimates for V_1, V_2 and V_3 in the proof below. The upper bound on the size of V_1 is essentially a bound on the size of the neighborhood on which the function is invertible. Such an estimate can be derived when the Jacobian matrix is Lipschitz –see for instance Theorem 2.9.7 in [19]. For V_2 and V_3, the bounds can be computed from their definitions (3.24) and (3.30). In this article we do not give an explicit bound for the size of V. However we will discuss how to ensure that $\mathbf{b} - \mathbf{a}$ is small enough in practice (Section 3.3). For further discussion, we refer readers to Section 5.

Proof of Theorem 3.5. We would like to use the Inverse Function Theorem to prove I. For the function $\mathbf{G}$ defined in (3.18), we have $\mathbf{G}(\mathbf{0}) = \mathbf{a}$. To apply the inverse function theorem, we verify that the Jacobian matrix at $\mathbf{s} = \mathbf{0}$ is invertible.

$$\frac{\partial \mathbf{G}(\mathbf{s})}{\partial \mathbf{s}} = \frac{1}{N_x} \sum_{i=1}^{N_x} \nabla \mathbf{F}(\mathbf{x}_i + \mathbf{s}^T \mathbf{A}_i) \mathbf{A}_i^T, \tag{3.20}$$

so

$$\frac{\partial \mathbf{G}(\mathbf{0})}{\partial \mathbf{s}} = \frac{1}{N_x}\sum_{i=1}^{N_x} \mathbf{A}_i \mathbf{A}_i^T = \frac{1}{N_x}\mathbf{A}\mathbf{A}^T. \tag{3.21}$$

Since $\mathbf{A}$ is full row rank, the Jacobian matrix is positive definite, thus invertible. By the inverse function theorem, there exists an open set U_1 around $\mathbf{0}$ and an open set V_1 around $\mathbf{a}$ such that $\mathbf{G}$ gives a bijection between U_1 and V_1.

Thus as long as $\mathbf{b} \in V_1$, there exists an $\mathbf{s}^*$ satisfying

$$\mathbf{s}^* = \mathbf{G}^{-1}(\mathbf{b}). \tag{3.22}$$

The corresponding ϕ^* is the local solution for the sample-based Monge-Ampere equation.

For II, we compute the objective function in (3.10):

$$\sum_{i=1}^{N_x} \|\phi^*(\mathbf{x}_i) - \mathbf{x}_i\|^2 = \sum_{i=1}^{N_x} \|\mathbf{s}^T \mathbf{A}_i\|^2 = \mathbf{s}^T \mathbf{A}\mathbf{A}^T \mathbf{s}. \tag{3.23}$$

Since $\mathbf{A}\mathbf{A}^T$ is positive definite, it provides a metric for $\mathbb{R}^M$. Then there exists $\delta > 0$, such that

$$U_2 := \left\{ \mathbf{s} \,\middle|\, \sqrt{\mathbf{s}^T \mathbf{A}\mathbf{A}^T \mathbf{s}} < \delta \right\} \subset U_1. \tag{3.24}$$

Define $V_2 = \mathbf{G}(U_2)$, we have that for any $\mathbf{b} \in V_2$, $\mathbf{s}^* = \mathbf{G}^{-1}(\mathbf{B})$ is the unique point in U_2 that solves $\mathbf{G}(\mathbf{s}) = \mathbf{b}$. Thus $\mathbf{s}^*$ gives the global minimum of the sample-based optimal transport problem.

For III, we first claim that the L_2 norm of all feature functions' Hessian matrices are uniformly bounded on $\mathscr{X}$, i.e. there exists a finite number C, such that for all $\mathbf{x} \in \mathscr{X}$ and j,

$$\|\mathbf{H}F_j(\mathbf{x})\|_2 < C. \tag{3.25}$$

Since $\mathbf{H}F_j(\mathbf{x})$ is a continuous function on the compact set $\mathscr{X}$, and

$$\|\mathbf{H}F_j(\mathbf{x})\|_2 = \max_{\|\mathbf{t}\|=1} \|\mathbf{H}F_j(\mathbf{x})\mathbf{t}\|, \tag{3.26}$$

$\|\mathbf{H}F_j(\mathbf{x})\|_2$ is also a continuous function of $\mathbf{x}$. From the compactness of $\mathscr{X}$, there exists a constant C_j such that

$$\|\mathbf{H}F_j(\mathbf{x})\|_2 < C_j, \forall \mathbf{x} \in \mathscr{X} \tag{3.27}$$

We can than set $C = \max_j C_j$ and establish the uniform boundedness of the Hessian matrices.

Compute the Hessian matrix of ϕ^*,

$$\mathbf{H}\phi^*(\mathbf{x}) = \mathbf{I} + \sum_{j=1}^{M} s_j^* \mathbf{H}F_j(\mathbf{x}) \tag{3.28}$$

Since

$$\| \sum_{j=1}^{M} s_j^* \mathbf{H} F_j(\mathbf{x}) \|_2 \leq \sum_{j=1}^{M} \| s_j^* \mathbf{H} F_j(\mathbf{x}) \|_2 \leq C \| \mathbf{s}^* \|_1, \tag{3.29}$$

if we constrain $\mathbf{s}^*$ such that $C\| \mathbf{s}^* \|_1 < 1$, $\mathbf{H}\phi^*(\mathbf{x})$ is invertible. Since $\| \cdot \|_1$ also gives a metric on $\mathbb{R}^M$, using the same argument we can define U_3 and V_3 that guarantee that ϕ^* is convex on $\mathscr{X}$:

$$U_3 := \left\{ \mathbf{s} \,\middle|\, \| \mathbf{s} \|_1 < \frac{1}{2C} \right\} \subset U_2, \quad V_3 = \mathbf{G}(U_3) \tag{3.30}$$

This concludes the proof of III. □

3.3 Sample-based optimal transport and barycenter algorithms

In this section we adapt the theoretical algorithms introduced in Section 2.2 by using the local solution theorem in Section 3.2 to solve sample-based problems.

Theorem 3.5 guarantees that the local solution to the sample-based Monge-Ampere equation (3.7) is the solution to the sample-based optimal transport problem. In practice, we use a standard iterative nonlinear system solver, such as a trust region method [11] or the Levenberg-Marquardt method [28], to solve the system (3.7). We set the initial point of the numerical algorithm to $\mathbf{s} = \mathbf{0}$, thus the solution given by the solver, if successful, will be the local solution sought.

The theoretical algorithms TOT and CTB introduced in Section 2.2 are based on a black box solver for the local optimal transport problem. This corresponds exactly to the local sample-based problem and the nonlinear system solver. Thus if we use the nonlinear system solver to solve local problems in the theoretical algorithms, we can extend the algorithms to solve sample-based problems.

While the theoretical algorithms can be naturally extended to the sample-based case, the latter requires the specification of feature functions. These can be defined by the user, based on expert knowledge of the nature of the sample sets, which gives great flexibility to the sample-based formulation. We devote the next section to discuss some general ways of selecting feature functions.

To complete the description of the algorithm, we need to further specify initial sets and stopping criteria.

For the initial sets in the SOT algorithm, we need to ensure that $\{\mathbf{x}_i^{k-1}\}_{i=1}^{N_{k-1}}$ and $\{\mathbf{x}_i^k\}_{i=1}^{N_k}$ are close to each other so that Theorem 3.5 holds for the subproblem. We first assign to each point $\mathbf{x}_i$ a random points $\mathbf{z}_i$ in the set $\{\mathbf{y}_i\}_{i=1}^{N_y}$. Then we use step 5 of the SOT algorithm to define all the initial sets. With K large enough, we can guarantee that every pair of sample sets gives close feature vectors. For simplicity, we use equal weights $\frac{1}{K}$.

For the CSB algorithm, the initial set can be chosen arbitrarily. In practice, we choose the set to be identical to one of the sets given.

Algorithm 7 Sample-based Optimal Transport Algorithm (SOT)

(1) Let $\{\mathbf{x}_i^0\}_{i=1}^{N_x} = \{\mathbf{x}_i\}_{i=1}^{N_x}$ and $\{\mathbf{x}_i^K\}_{i=1}^{N_y} = \{\mathbf{y}_i\}_{i=1}^{N_y}$. Set initial sample sets $\{\mathbf{x}_i^k\}_{i=1}^{N_x}$ $(k = 1, 2, \cdots, K-1)$;
(2) Select a set of feature functions $\{F_j(\mathbf{x})\}_{j=1}^{M}$;
(3) Set the current set $\{\mathbf{z}_i\}_{i=1}^{N_x} = \{\mathbf{x}_i^0\}_{i=1}^{N_x}$;
(4) For $k = 1, 2, \cdots, K$, solve the sample-based optimal map $\mathbf{f}$ from $\{\mathbf{z}_i\}_{i=1}^{N_x}$ to $\{\mathbf{x}_i^k\}_{i=1}^{N_x}$ using a standard nonlinear system solver and update $\{\mathbf{z}_i\}_{i=1}^{N_x}$ through

$$\{\mathbf{z}_i\}_{i=1}^{N_x} = \{\mathbf{f}(\mathbf{z}_i)\}_{i=1}^{N_x}; \tag{3.31}$$

(5) For $k = 1, 2, \cdots, K-1$, update $\{\mathbf{x}_i^k\}_{i=1}^{N_x}$ through

$$\{\mathbf{x}_i^k\}_{i=1}^{N_x} = \{\frac{K-k}{K}\mathbf{x}_i^0 + \frac{k}{K}\mathbf{z}_i\}_{i=1}^{N_x} \tag{3.32}$$

and go to step 3.

Algorithm 8 Composite Sample-based Barycenter Algorithm (CSB)

(1) Set the initial sample set $\{\mathbf{y}_i\}_{i=1}^{N_y}$;
(2) Select a set of feature functions $\{F_j(\mathbf{x})\}_{j=1}^{M}$;
(3) Run step 1,3,4 of SOT once for each pair $(\{\mathbf{y}_i\}_{i=1}^{N_y}, \{\mathbf{x}_i^k\}_{i=1}^{N_k})$ to map $\{\mathbf{y}_i\}_{i=1}^{N_y}$ to a new set $\{\mathbf{z}_i^k\}_{i=1}^{N_y}$;
(4) Update $\{\mathbf{y}_i\}_{i=1}^{N_y}$ through

$$\{\mathbf{y}_i\}_{i=1}^{N_y} = \left\{\sum_{k=1}^{K} w_k \mathbf{z}_i^k\right\}_{i=1}^{N_y}. \tag{3.33}$$

(5) Run step 5, 3, 4 of SOT once for each pair $(\{\mathbf{y}_i\}_{i=1}^{N_y}, \{\mathbf{x}_i^k\}_{i=1}^{N_k})$ to update $\{\mathbf{z}_i^k\}_{i=1}^{N_y}$ and go to step 4.

We stop the algorithm when an iteration does not significantly change the map. More specifically, in the SOT algorithm we compare the images of $\{\mathbf{x}_i\}_{i=1}^{N_x}$ before and after one iteration. If

$$\frac{1}{N_x}\sum_{i=1}^{N_x} \| \mathbf{z}_i^{\text{before}} - \mathbf{z}_i^{\text{after}} \|^2 < \varepsilon, \tag{3.34}$$

we stop the algorithm. In the CSB algorithm, we compare the set $\{\mathbf{y}_i\}_{i=1}^{N_x}$ before and after one iteration, if

$$\frac{1}{N_y}\sum_{i=1}^{N_y} \| \mathbf{y}_i^{\text{before}} - \mathbf{y}_i^{\text{after}} \|^2 < \varepsilon, \tag{3.35}$$

we stop the algorithm.

The SOT and CSB algorithms can be directly applied to sample sets. Except for the selection of feature functions, there are only four parameters. For both algorithms, two of the them are for the stopping criteria, one for the local nonlinear system solver and the other the ε introduced above. The third parameter is N_y, the number of sample points sought from the barycenter. Finally, the forth parameter is K, the number of interpolant sample sets in the SOT algorithm (which is also used in the CSB algorithm). K should be large enough so that every subproblem is local so that it has a solution. In practice, we set K to an initial value K_0; if the local problem does not have a solution (a fact informed by the local nonlinear system solver), we double the value of K and rerun the current iteration.

3.4 Selection of feature functions

Both the optimal map and the time complexity of the algorithms for the sample-based optimal transport and barycenter problems are highly dependent on the choice of feature functions. The more feature functions one uses, the richer the structure of the optimal map $\nabla\phi(\mathbf{x})$ (though one must be careful not to overfit the problem) and the slower the algorithm.

Feature functions can be defined by the user, based on expert knowledge of the nature of the sample sets. Feature functions constructed in this way can be very informative and useful. We will discuss an example of problem-specific feature functions in the numerical tests Section 4.3.

In this section we introduce techniques to choose feature functions that are not problem-specific, and attempt to automatically capture the most significant aspects of the sample sets without requiring the user's external insight.

Moments

The simplest feature functions are moments of the sample sets. Choosing for feature functions the components of x:

$$F_j(\mathbf{x}) = \mathbf{x}^{(j)}, \quad j = 1,2,\cdots,d, \tag{3.36}$$

the sample-based Monge-Ampere equation simply matches the means of $\{\nabla\phi(\mathbf{x}_i)\}_{i=1}^{N_x}$ and $\{\mathbf{y}_i\}_{i=1}^{N_y}$. Since $\nabla\mathbf{F}(\mathbf{x}) = \mathbf{I}$, the corresponding function space of the map only contains translations,

$$\nabla\phi(\mathbf{x}) = \mathbf{x} + \mathbf{s}, \tag{3.37}$$

which shift the first sample set so as to match the mean of the second sample set.

Choosing for feature functions the first and second moments,

$$F_j(\mathbf{x}) = \mathbf{x}^{(j)}, \quad j = 1,2,\cdots,d, \tag{3.38}$$

$$F_{\frac{(2n-i+1)i}{2}+j}(\mathbf{x}) = \mathbf{x}^{(i)}\,\mathbf{x}^{(j)}, \quad 1 \le i \le j \le d, \tag{3.39}$$

the sample-based Monge-Ampere equation matches the means and covariance matrices of $\{\nabla\phi(\mathbf{x}_i)\}_{i=1}^{N_x}$ and $\{\mathbf{y}_i\}_{i=1}^{N_y}$. Suppose the means of $\{\mathbf{x}_i\}_{i=1}^{N_x}$ and $\{\mathbf{y}_i\}_{i=1}^{N_y}$ are $\mathbf{m}_1$ and $\mathbf{m}_2$ and the covariance matrices are Σ_1 and Σ_2. The solution of the sample-based Monge-Ampere equation can be written as

$$\nabla\phi(\mathbf{x}) = \mathbf{s}_1 + (\mathbf{x}-\mathbf{m}_1)\mathbf{S}, \tag{3.40}$$

where $\mathbf{s}_1 \in \mathbb{R}^d$ is a row vector, $\mathbf{S} \in \mathbb{R}^{d\times d}$ is a symmetric matrix. From the Monge-Ampere equation, they satisfy

$$\mathbf{s}_1 = \mathbf{m}_2, \tag{3.41}$$

$$\mathbf{S}\Sigma_1\mathbf{S} = \Sigma_2. \tag{3.42}$$

Solving the above yields the optimal linear map:

$$\nabla\phi^*(\mathbf{x}) = \mathbf{m}_2 + (\mathbf{x}-\mathbf{m}_1)\Sigma_1^{-1/2}(\Sigma_1^{1/2}\Sigma_2\Sigma_1^{1/2})^{1/2}\Sigma_1^{-1/2}, \tag{3.43}$$

which agrees with the optimal map solution between multivariate normal distributions [18]. For these moment functions, the sample-based Monge-Ampere equation, because of linearity, has a unique solution.

Since in this case the solution has a closed-form, we do not need to use the SOT algorithm to solve the sample-based optimal transport problem. Instead, we can simply compute the empirical mean and covariance matrices of the two sample sets and apply (3.43).

In the same setting for the sample-based barycenter problem, when the mean and covariance matrix of the kth sample set are $\mathbf{m}_k$ and Σ_k, we can use the BTB algorithm to find the barycenter. Using the closed-form solution (3.43), we have the following adaptation of the BTB algorithm 9.

This algorithm can also be viewed as an algorithm iterating over the mean $\mathbf{m}_y$ and covariance matrix Σ_y. After step 4 of each iteration, we can compute the new $\mathbf{m}_y^{\text{new}}$ and Σ_y^{new} with respect to the old ones,

$$\mathbf{m}_y^{\text{new}} = \sum_{k=1}^{K} w_k\mathbf{m}_k, \tag{3.48}$$

$$\Sigma_y^{\text{new}} = \Sigma_y^{-1/2}\left(\sum_{k=1}^{K} w_k(\Sigma_y^{1/2}\Sigma_k\Sigma_y^{1/2})^{1/2}\right)^2\Sigma_y^{-1/2}. \tag{3.49}$$

The fixed point of this algorithm satisfies

$$\Sigma_y = \sum_{k=1}^{K} w_k(\Sigma_y^{1/2}\Sigma_k\Sigma_y^{1/2})^{1/2}, \tag{3.50}$$

which agrees with the solution of barycenter problems of multiple normal distributions [2]. The iteration formula also agrees with the fixed point iterative methods found in [3].

We can see that using first and second moment functions as feature functions gives us exact solutions of the optimal transport and barycenter problems with

Algorithm 9 Fixed Point Barycenter Algorithm

(1) Set the initial sample set $\{\mathbf{y}_i\}_{i=1}^{N_y}$;

(2) Calculate the empirical mean and covariance matrix of $\{\mathbf{y}_i\}_{i=1}^{N_y}$,

$$\mathbf{m}_y = \frac{1}{N_y}\sum_{i=1}^{N_y}\mathbf{y}_i \tag{3.44}$$

$$\Sigma_y = \frac{1}{N_y}\sum_{i=1}(\mathbf{y}_i-\mathbf{m}_y)^T(\mathbf{y}_i-\mathbf{m}_y); \tag{3.45}$$

(3) Define the optimal maps from $\{\mathbf{y}_i\}_{i=1}^{N_y}$ to $\{\mathbf{x}_i^k\}_{i=1}^{N_k}$ $(k=1,2,\cdots,K)$ using (3.43):

$$\nabla\phi_k(\mathbf{y}) = \mathbf{m}_k + (\mathbf{y}-\mathbf{m}_y)\Sigma_y^{-1/2}(\Sigma_y^{1/2}\Sigma_k\Sigma_y^{1/2})^{1/2}\Sigma_y^{-1/2}; \tag{3.46}$$

(4) Update $\{\mathbf{y}_i\}_{i=1}^{N_y}$ using

$$\mathbf{y}_i = \sum_{k=1}^{K} w_k\nabla\phi_k(\mathbf{y}_i), \quad i=1,2,\cdots,N_y \tag{3.47}$$

and go to step 2.

normal distributions. Moreover, it gives an iterative algorithm to solve the sample-based barycenter problem. The other advantage of the sample-based formulation is that we can use this solution on arbitrary sample sets without the assumption of normal distributions. The result is justified as a specific solution when only the first and second moments are chosen as feature functions.

Kernel functions

Another type of feature functions can be used to characterize the local structure of a sample set. For a given location $\mathbf{z}_0$ and a bandwidth parameter h, we can define a kernel function $F(\mathbf{x})$ such that

$$F(\mathbf{x}) = K\left(\frac{\mathbf{x}-\mathbf{z}_0}{h}\right) \tag{3.51}$$

where $K(\mathbf{x})$ is a nonnegative C^1 function that vanishes as $\mathbf{x}$ tends to infinity. The mean of $F(\mathbf{x})$ over a sample set $\{\mathbf{x}_i\}_{i=1}^{N_x}$ is

$$\frac{1}{N_x}\sum_{i=1}^{N_x}K\left(\frac{\mathbf{x}_i-\mathbf{z}_0}{h}\right), \tag{3.52}$$

which can be interpreted as the value of a kernel density estimator at point $\mathbf{z}_0$. Thus if the feature function set consists of kernel functions at $(\mathbf{z}_0,h_0),(\mathbf{z}_1,h_1),\cdots,(\mathbf{z}_M,h_M)$, the sample-based equivalence 3.1 between two sample sets corresponds to their density functions from kernel density estimation agreeing at the locations $\{\mathbf{z}_j\}_{j=1}^{M}$.

One natural candidate for the kernel function $K(\mathbf{x})$ is the Gaussian kernel

$$K_h(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d}} \exp\left(-\frac{\|\mathbf{x}\|^2}{2h^2}\right), \tag{3.53}$$

which we have used to build feature functions in most of the numerical test in Section 4.

The remaining choice to make is the location/bandwidth set $\{(\mathbf{z}_j, h_j)\}_{j=1}^M$. The simplest strategy is to choose for $\{\mathbf{z}_j\}_{j=1}^M$ a regular grid in sample space. This is shown to work well in applications in Section 4.3 and Section 4.3.

Data-driven approaches can better inform the choice of $\{(\mathbf{z}_j, h_j)\}_{j=1}^M$. For instance, one might want to place more kernels in areas with a higher density of points. One appropriate technique is the mean-shift algorithm [10], which locates local maxima of the probability density underlying data.

Specifically, for all the input sample sets, we choose a set of bandwidth parameters $\{h_j\}$. For each h_j, we use the mean shift algorithm on all the sample sets to locate a set of local maxima $\{\mathbf{z}_j^i\}$. We then construct kernel functions from all pairs of $(\mathbf{z}_j^i, h_j)$ to form a set of feature functions. This scheme is applied in Section 4.2.

Feature functions from feature extraction techniques

Feature extraction techniques seek to extract as much information as possible using a minimum number of functions.

One such technique is principal component analysis. Consider the optimal transport problem between two sample sets in $\mathbb{R}^d$. As mentioned, we can use all the first moment functions to shift one sample set to the other, which requires d feature functions. As an alternative, we can perform PCA on the sample sets and only select the projections on the first few principle directions as feature functions.

A more useful generalization of PCA is kernel PCA [41]. With a kernel function, kernel PCA can capture nonlinear structures in sample sets.

4 Numerical examples

In this section, we apply SOT and CSB algorithms on both synthetic and real examples. We start with optimal transport between normal distributions. We use these examples to assess the performance of the algorithms, since the true solutions are known in closed form.

Then we present a detailed comparison between the sample-based method and classical discrete optimal transport algorithms on a one-dimensional example. We show that when the data consists of i.i.d. sample sets, sample-based optimal transport outperforms the others in both accuracy and computational cost.

Finally we apply sample-based algorithms to more general sample sets, including Gaussian mixture models and shape and color transfer problems. The underlying distributions in these applications include multimodality, complex geometry, three dimensionality and rich structure.

4.1 Optimal transport and barycenter problems of normal distributions

Quadratic optimal transport and barycenter problems with normal distributions have closed-form solutions (Section 3.4). This provides us with good synthetic examples to assess our methods and perform error analysis.

Optimal transport problem between normal distributions

Consider the two-dimensional normal distributions $\mathcal{N}(\mathbf{m}_0, \Sigma_0)$ and $\mathcal{N}(\mathbf{m}_1, \Sigma_1)$ with

$$(\mathbf{m}_0, \mathbf{m}_1) = ((0,0),(5,5)),$$
$$(\Sigma_0, \Sigma_1) = \left(\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \right)$$

Drawing 200 sample points from each distribution, we obtain the datasets shown in Fig. 4.1a.

To solve the corresponding sample-based optimal transport problem, we apply the SOT algorithm 7 with first and second moment functions as feature functions. In the algorithm, we set the number of interpolant sample sets K to 12 and the error limit ε in (3.34) to 10^{-4}.

Fig. 4.1b shows the trajectories of 15 random sample points from the sample set of μ_0. We can see that all the trajectories are straight lines with equal distance between interpolant sample sets, a property of McCann interpolants.

Fig. 4.1c displays the convergence of the SOT algorithm. The error computed is similar to (3.34),

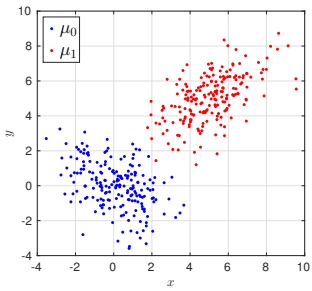
$$\sqrt{\frac{1}{N_x} \sum_{i=1}^{N_x} \| \mathbf{z}_i^j - \mathbf{z}_i^* \|^2}, \tag{4.1}$$

where $\mathbf{z}_i^j$ is the map of $\mathbf{x}_i$ after the jth iteration and $\mathbf{z}_i^*$ is the optimal one. The algorithm converges sublinearly and stops in only 5 iterations.

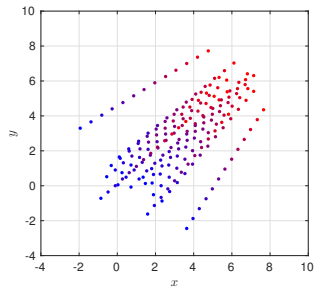
Fig. 4.1d shows 4 interpolant sample sets $\tilde{\mu}_t$ with $t = 0.25, 0.5, 0.75, 1$. We can see the sample sets moving toward μ_1 and gradually changing shape.

Barycenter of normal distributions

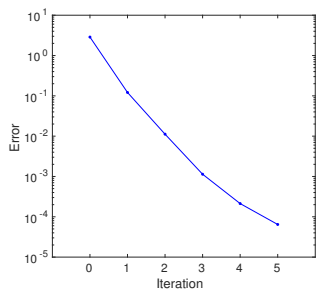
Consider a two-dimensional barycenter problem with five normal distributions (Fig. 4.2). For each of the 400 point sample sets, one can calculate the empirical means and covariance matrices, shown in Table 4.1. With finite sample sizes, sample statistics only provides estimates for the unknown distribution parameters. To analyze the convergence of our algorithms, we will ignore sample errors and directly work with sample statistics. Statistics of the true barycenter measure μ^* in Table 4.1 is calculated using (3.50) with the sample means and covariance matrices.



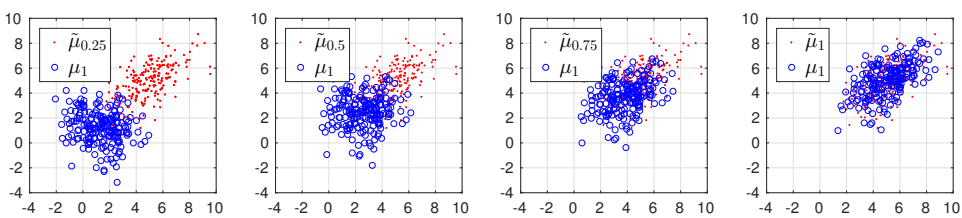
(A) Sample sets



(B) Trajectories of sample points



(C) Empirical L_2 error



(D) Interpolant sets

FIGURE 4.1. Application of the SOT algorithm with parameters $K = 12$, $\varepsilon = 10^{-4}$ to the optimal transport problems between two normal distributions. (a) Sample sets drawn from distributions μ_0 and μ_1 respectively. (b) Trajectories of 15 sample points from μ_0 to μ_1. (c) Empirical L_2 error defined in (4.1) after each iteration of SOT. (d) McCann interpolant sample sets with $t = 0.25, 0.5, 0.75, 1$.

Distribution	$\mathbf{m}_1$	$\mathbf{m}_1$	Σ_{11}	Σ_{22}	Σ_{12}
μ_1	1.22	4.63	0.59	0.25	-0.33
μ_2	4.73	-1.43	0.26	0.11	0.12
μ_3	-1.87	-4.62	0.32	0.57	0.27
μ_4	3.47	-4.83	0.55	0.30	0.15
μ_5	-2.99	-2.75	0.51	0.46	0.09
μ^*	0.91	-1.80	0.38	0.27	0.07

TABLE 4.1. Sample statistics of five normal distributions in the barycenter problem example.

We apply two barycenter algorithms to this problem: the CSB algorithm 8 and the Fixed Point Barycenter Algorithm 9.

There are several differences between these two algorithms. The fixed point algorithm is only applicable to normal distributions, as each iteration uses directly the optimal transport solution for normal distributions (3.43). On the other hand, the CSB algorithm solves general barycenter problems. In each iteration it does not compute full optimal transport solutions but only performs one SOT update for each subproblem.

As mentioned in Section 3.4, the fixed point algorithm can be viewed as a basic sample-based algorithm. If the optimal transport solution is unknown, it runs a full SOT algorithm in every iteration, which typically consists of several SOT updates. This indicates that the fixed point algorithm should generate much better updates per iteration.

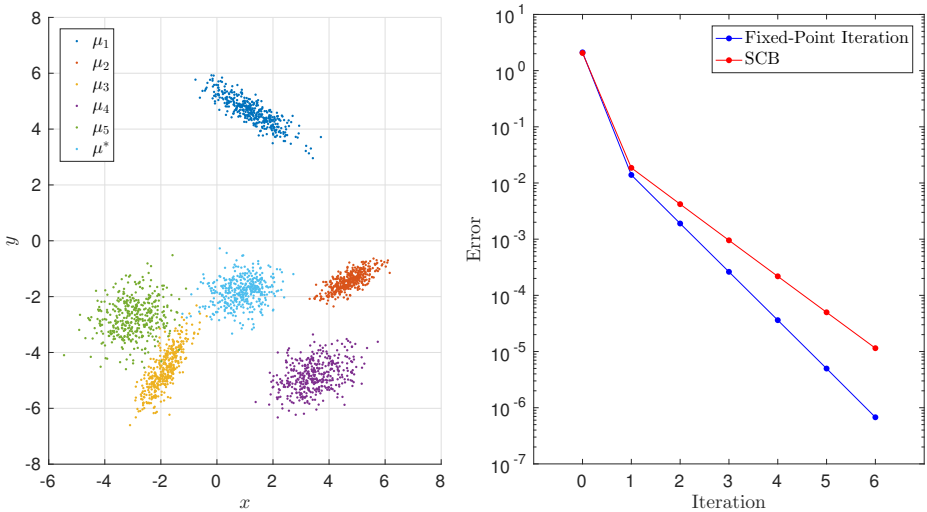
In the experiment, we set $K = 12$ and $\varepsilon = 10^{-4}$ for the CSB algorithm; the feature functions chosen are the first and second moment functions. The resulting barycenter sample set μ^* is shown in Fig. 4.2a.

Fig. 4.2b displays the solution's accuracy after every iteration, measured as the Wasserstein distance between normal distributions. We can see that both algorithms converge linearly and generate accurate solutions in only a few iterations. While the fixed point barycenter algorithm does converge faster due to its additional normal distribution assumption, the CSB algorithm has a comparable convergence rate, even though it only performs one SOT update in each iteration.

4.2 Detailed analysis of a one-dimensional optimal transport problem

In this section, we perform a careful comparison between sample-based optimal transport algorithms and existing ones. Specifically, we compare the discrete linear programming method with the SOT algorithm 7 with meanshift feature functions.

We have so far ignored sample error: with finite sample sets, the true solution between unknown distributions are unavailable. The hope is that the solution derived from sample sets is close to the unknown truth when the sample size is large enough.



(A) Sample sets (B) Trajectories of sample points

FIGURE 4.2. Application of the CSB algorithm with parameters $K = 12$, $\varepsilon = 10^-4$ to a barycenter problem between five normal distributions. (a) Sample sets drawn from the $\mu_i (i = 1, \cdots, 5)$. (b) Errors after each iteration of the CSB algorithm and Fixed Point Barycenter Algorithm. The error is defined as the Wasserstein distance between the true and the estimated distribution.

To measure the quality of an algorithm, we compute the error between the true map $f^*(x)$ of distributions and the map $f(x)$ derived from an instance of finite sample sets:

$$e(f) = \mathbb{E}(f(x) - f^*(x))^2. \tag{4.2}$$

Here the random variable x follows the source distribution. Then we compute the expected error $\mathbb{E}e(f)$ by averaging $e(f)$ over multiple instances of sample sets.

We can see that in order to compute $\mathbb{E}e(f)$, we need to fix both source and target distributions and generate multiple sample set instances as inputs. The other requirement is that we need to be able to compute the true solution between general continuous distributions, which is why we work on one-dimensional distributions.

For meaningful comparison, we create synthetic distributions such that neither method is able to exploit their specific forms. The source distribution chosen is the Gaussian mixture model (Fig. 4.3a):

$$p(x) \sim \frac{1}{2}[N(\frac{2}{9}, \frac{2}{9}) + N(\frac{7}{9}, \frac{1}{9})], \tag{4.3}$$

while a beta distribution is picked as target:

$$q(x) \sim \text{Beta}(2, 5). \tag{4.4}$$

This example requires the algorithms to correctly transfer a symmetric, two-mode distribution into a one-mode, asymmetric one.

We now specify the discrete linear programming method we compare the SOT algorithm with. As mentioned in Section 3.1, in order to use the linear programming method on sample-based problems, we first need to construct histograms from the sample sets to provide as inputs to the linear programming algorithm.

Specifically, for both the source and target sample sets, we divide their corresponding domains into M equal sized bins, denoting the bin centers by s_i and t_j, and the corresponding histogram distributions by p_i and q_j (Figs. 4.3a and 4.3b). Defining the cost function between source bin i and target bin j as $c_{ij} = (s_i - b_j)^2$, we solve the following linear programming problem:

$$\min_{\pi_{ij}\geq 0} \sum_{i,j=1}^{M} c_{ij}\pi_{ij}, \quad \text{s.t.} \sum_{i=1}^{M} \pi_{ij} = q_j, \sum_{j=1}^{M} \pi_{ij} = p_i. \tag{4.5}$$

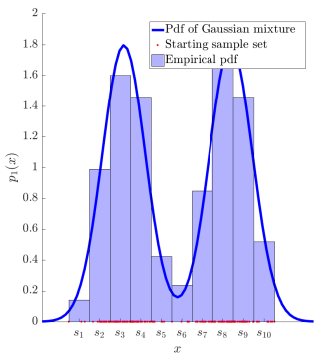
Since the solution π_{ij} of the above optimization does not automatically yield a map, we follow the natural averaging scheme: if sample point x is in the i_0th bin of the source distribution, we define the map f as the conditional expected value of the solution:

$$f(x) = \frac{1}{p_{i_0}} \sum_{j=1}^{M} \pi_{i_0 j} t_j. \tag{4.6}$$

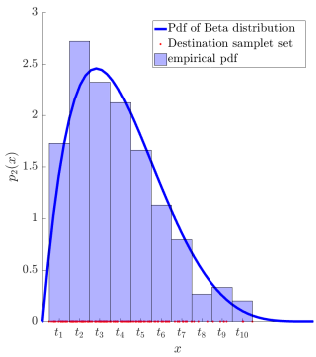
To compare the two algorithms under different sample variances, we choose as sample sizes $50, 100, 200, 400$. For the discrete solver, the main parameter is the number of bins used to obtain the empirical pdf; we use $M = 5, 10, 20, 40, 80, 160$ for all cases. Notice that the number of variables in the corresponding linear programming problem is the square of the bin numbers (i.e. 25 for $M = 5$ and 25600 for $M = 160$.) For the SOT algorithm with mean shift feature functions, we fix the mean shift bandwidths $h = 0.1, 0.5$, since this parameter describes the distribution rather than the sample set. The number of variables (feature functions) in the corresponding non-linear optimization problem is then fixed for a given sample set, and is shown in the caption of figure 4.4.

As shown in Fig. 4.4, the SOT algorithm outperforms the discrete linear programming algorithm in every case (different sample sizes and different number of bins). From the computational perspective, the SOT algorithm uses no more than 30 variables in the optimization problem, while the discrete solver requires many more to yield solutions of similar quality.

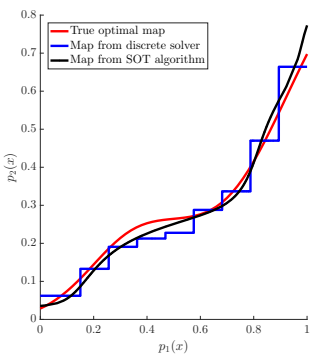
Notice that, with larger sample size, the value of $\mathbb{E}e(f)$ decreases, which is likely due to the component of $\mathbb{E}e(f)$ arising from the sample variance of the sample sets. Another observation is that the discrete solver requires larger bin numbers to achieve comparable $\mathbb{E}e(f)$ to the SOT solver. Hence the user needs to tune the number of bins in the solver to the different sample sizes. On the other hand, the SOT solver has as only parameter the bandwidth h, which gives uniformly good results for all sample sizes in our example.



(A) Starting distribution



(B) Destination distribution



(C) Optimal maps

FIGURE 4.3. One-dimensional synthetic example. Both sample sets have 100 sample points. The number of bins is 10 for both histograms. The SOT solver uses the mean shift methodology for feature selection with parameter $h = 0.1, 0.5$. (a) Gaussian mixture $\frac{1}{2}[N(\frac{2}{9}, \frac{2}{9}) + N(\frac{7}{9}, \frac{1}{9})]$. (b) Beta distribution Beta$(2,5)$. (c) True optimal map and solutions from the discrete solver and the SOT algorithm.

4.3 Optimal transport between general distributions

The prior sub-sections examined sample-based algorithms applied to problems with closed-form solutions. We now apply these algorithms to more challenging and practical problems to see the broad range of maps that the algorithm can generate.

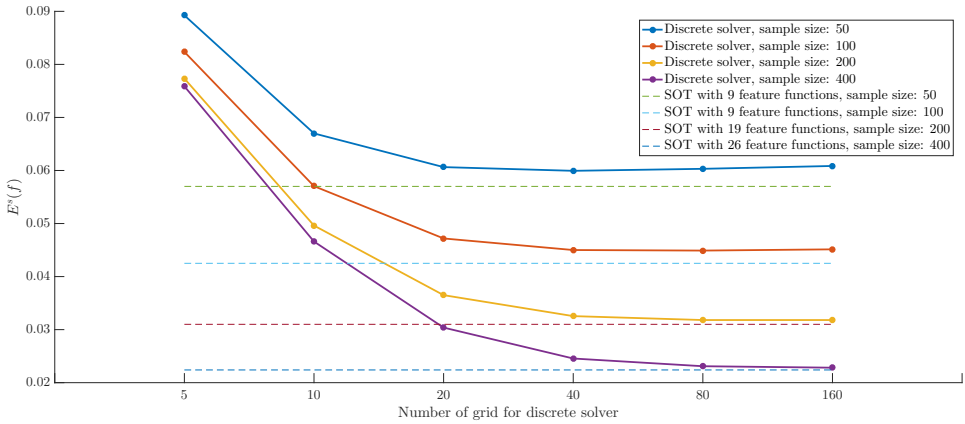


FIGURE 4.4. $E^s(f)$ Comparison between the discrete and the SOT solver. The $E^s(f)$ are averaged over 500 tests for each case. The sample sizes chosen for both sample sets are $50, 100, 200, 400$. In every case, the discrete solver is used with number of bins $5, 10, 20, 40, 80, 160$. The SOT solver is used with mean shift feature selection with band-width parameter $h = 0.1, 0.5$.

Optimal transport from normal to Gaussian mixture distributions

The first set of examples considered compute the two dimensional optimal transport from normal distributions to Gaussian mixtures.

One challenge that this scenario presents is that the optimal transport map needs to divide the single mode in the source distribution into two. Hence the solution maps a small high-density area to a large low-density one. This can potentially make the corresponding sample-based problem difficult, since with finite sample size, it is relatively hard to resolve low density areas well.

We present four examples with different parameters. In all four examples we set the starting distribution as a mean zero standard normal distribution and set the sample sizes of both sets to 200.

In the first example, the Gaussian mixture consists of two equal ratio normal distributions that are not faraway from each other (there is no clear separation in the given sample set). For the second Gaussian mixture, 75% percent of sample comes from the first normal distribution and 25% from the other, breaking the symmetry from the first example. In the third example, the Gaussian mixture is chosen such that the density in the area between the two modes is very low, so that the optimal transport has to optimally stretch the starting distribution into the destination. In the fourth example, the normal distributions have different shapes, which adds more structure to the resulting optimal transport. The resulting source and target distributions are shown in the first column of figure 4.5. The specific parameters are given in Table 4.2.

Example#	$\mathbf{m}_1$	$\mathbf{m}_2$	Σ_1	Σ_2	mixture ratio
1	$(-2,2)$	$(2,2)$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$50\% : 50\%$
2	$(-2,2)$	$(2,2)$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$75\% : 25\%$
3	$(-5,5)$	$(5,5)$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$50\% : 50\%$
4	$(-5,0)$	$(5,0)$	$\begin{pmatrix} 1 & -0.7 \\ -0.7 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix}$	$50\% : 50\%$

TABLE 4.2. Parameters of the Gaussian mixture models

For each of the examples, we first apply the preconditioning procedure [23] on both sample sets. We then use the SOT algorithm 7 with mean shift feature functions with $h = 0.25, 0.5, 1, 2$.

In figure 4.5, each row corresponds to the corresponding row in Table 4.2. The transported sample set $\tilde{\mu}_1$ is shown in the last column, together with the original target sample set. The resulting interpolant sample sets $\tilde{\mu}_{0.25}$,$\tilde{\mu}_{0.5}$ and $\tilde{\mu}_{0.75}$ are shown in the 2nd, 3rd and 4th columns respectively. We can see in all the examples the SOT algorithm's effectiveness in mapping normal to Gaussian mixture models.

Shape transforms

Optimal transport and barycenter problems have been used to solve shape transform problems [42]. The task is to find maps or barycenters for different shapes in two and three-dimensional spaces.

The application starts by viewing a shape as a uniform distribution supported on a given domain Ω. To transfer one shape to the other is equivalent to solving the optimal transport problem between the corresponding uniform distributions.

One of the advantages of the optimal transport methods is that they give one-to-one point maps between shapes and also intermediate shapes (interpolant and barycenter measures). This is particularly true in the sample-based setting, since the sample-based algorithm gives analytical map functions, which directly provide maps of arbitrary points without any further averaging scheme. Thus it is straightforward to derive intermediate shapes from the optimal solution.

An important issue to address is the representation of shapes in the sample-based setting. For source shapes, we generate sample sets through equal-distance grid points (e.g. Fig. 4.6). This simple process allows us to control the sample size by tuning the grid size. Choosing larger sample sizes allows us to represent more detailed structure in a shape. For instance, to represent the 'bird' shape in figure Fig. 4.7, we used 5000 sample points while for the ring (Fig. 4.6d) we used only 1000 sample points.

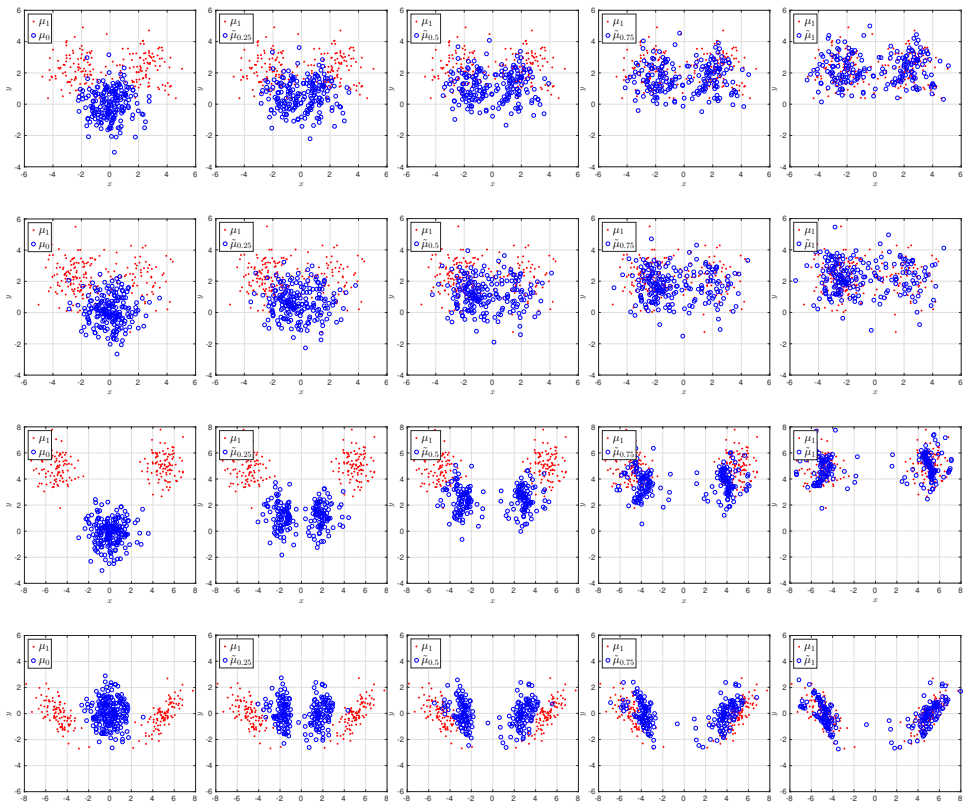


FIGURE 4.5. Sample-based optimal maps and interpolant maps at $t = 0.25, 0.5, 0.75$ from normal to 2-Gaussian mixture distributions by the SOT algorithm with mean shift feature functions. Source and mapped sample sets are displayed as blue circles, and target sample sets as red dots.

On the other hand, we need to develop a post-processing procedure to recover shapes from finite sample sets. We first generate a kernel density estimator $\tilde{p}(x)$ from the sample set. Then we construct the shape $\tilde{\Omega}$ by adopting as boundary a level set with relatively large $\tilde{p}(x)$.

In this article, we use Gaussian kernels for kernel density estimation. The bandwidth is chosen to be half of the *rules of thumb* bandwidth [21] to avoid blurring the shape's boundary. Then we define the following as the output shape:

$$\tilde{\Omega} = \{\mathbf{x} \,|\, \tilde{p}(\mathbf{x}) > \frac{1}{2} \max_{\mathbf{x}} \tilde{p}(\mathbf{x})\}. \tag{4.7}$$

In this section, all shape transforms are carried out using the preconditioning procedure in [23] and the SOT algorithm 7. The feature functions are chosen as Gaussian kernels with centers on a 7×7 grid with two bandwidths $h = 0.5, 1$.

In Fig. 4.6, we transfer an ellipse 4.6a to a ring 4.6d. Both sample sets consist of approximately 1000 grid points. We color the source sample sets with four different colors to show what each subset transforms into in the process. Interpolant sample sets are shown for $t = 1/3, 2/3$ (Figs. 4.6b and 4.6c).

To recover shapes from sample sets, we implement the post-processing step on all the sample sets. The corresponding shapes in Fig. 4.6 are recovered and shown in the first row of Fig. 4.7.

Three additional shape transforms are shown in Fig. 4.7. We can see that shapes with different orientations, different topologies and sharp corners can be effectively mapped into each other by the algorithm.

Color transfer

Next we apply sample-based algorithms to color transfer problems [38, 46, 13], which have as general objective to recolor a source image so that its colors resemble those of a target image. One can view the set of colors of an image as a distribution and find the optimal map between the source and the target, using the earth mover's distance (EMD) [39] as a quantification of the transfer required.

We follow the algorithm framework in [37], substituting the core optimal transport step by the SOT algorithm 7.

To represent an image, one assigns to each pixel a 5-dimensional vector $\mathbf{x} = (\mathbf{x}^s, \mathbf{x}^c)$, in which $\mathbf{x}^s = (x, y)$ represents the pixel location and $\mathbf{x}^c = (l, a, b)$ represents color in the CIELAB color space. Thus the sample set $\mathscr{S} = \{\mathbf{x}_i\}$ of all pixels is a full representation of the corresponding image.

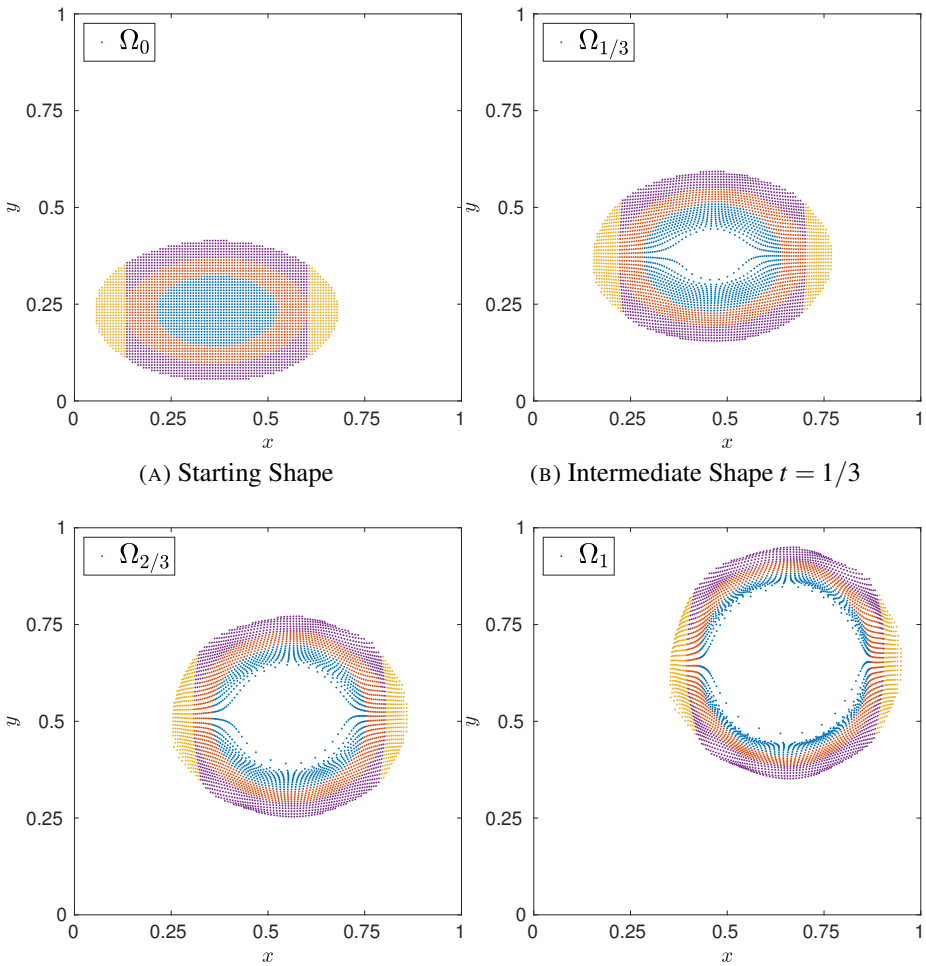
The first step of the color transfer meta-algorithm is Spatio-color Clustering. Using the super-pixel method [1], pixel set $\mathscr{S}$ is clustered into several subsets $\mathscr{S}_j$, where each $\mathscr{S}_j$ has mean vector $\bar{\mathbf{x}}_j = (\bar{x}_j, \bar{y}_j, \bar{l}_j, \bar{a}_j, \bar{b}_j)$ and empirical covariance matrix Σ_j. Then we define a new weighted sample set $\tilde{\mathscr{S}} = \{\bar{\mathbf{x}}_j\}$ with weights $|\mathscr{S}_j|/|\mathscr{S}|$ as a more compact approximation to an image (e.g. 4.8c).

There are two reasons for applying spatio-color clustering. First, since we only apply optimal transport on the color dimensions of sample sets, this step incorporates spatial information into the meta-algorithm. Second, compared to the original $\mathscr{S}$, $\tilde{\mathscr{S}}$ is a much smaller sample set. This makes the core optimal transport algorithm much faster since the computational cost scales with the sample size[2].

The second and core step is to apply an optimal transport algorithms to the weighted sample sets. Define $\tilde{\mathscr{S}}^c \in \mathbb{R}^3$ as the set of points in CIELAB color space projected from the elements in $\tilde{\mathscr{S}}$. To find an optimal map T between $\tilde{\mathscr{S}}^c_{\text{source}}$ and $\tilde{\mathscr{S}}^c_{\text{target}}$ of source and target images, we apply the preconditioning procedure along with the SOT algorithm.

The third step is image synthesis. As the optimal map f above only maps mean color vectors, one needs to construct a color map for all pixels in the original set $\mathscr{S}$ that incorporates geometrical information.

[2] This is true both for sample-based algorithms and for the regularized optimal transport in [37].



(A) Starting Shape (B) Intermediate Shape $t = 1/3$

(C) Intermediate Shape $t = 2/3$ (D) Target Shape

FIGURE 4.6. Optimal transport from an ellipse sample set (a) to a ring sample set (d). (b)(c) are intermediate sample sets with $t = 1/3, 2/3$. The sample set in (a) is colored to identify the trajectory of four subsets.

As in [44], we define a similarity metric $\omega_j(\mathbf{x})$ using Gaussian kernels:

$$\omega_j(\mathbf{x}) = \exp\left(-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}}_j)\tilde{\Sigma}_j^{-1}(\mathbf{x}-\bar{\mathbf{x}}_j)^T\right), \tag{4.8}$$

where $\tilde{\Sigma}_j$ is a weighted covariance matrix defined by

$$\tilde{\Sigma}_j = W\Sigma_j W, \quad W = \mathrm{diag}(\sigma^s, \sigma^s, \sigma^c, \sigma^c, \sigma^c). \tag{4.9}$$

Here σ^s, σ^c are parameters that control the strength of smoothing in the spatial and color spaces respectively.

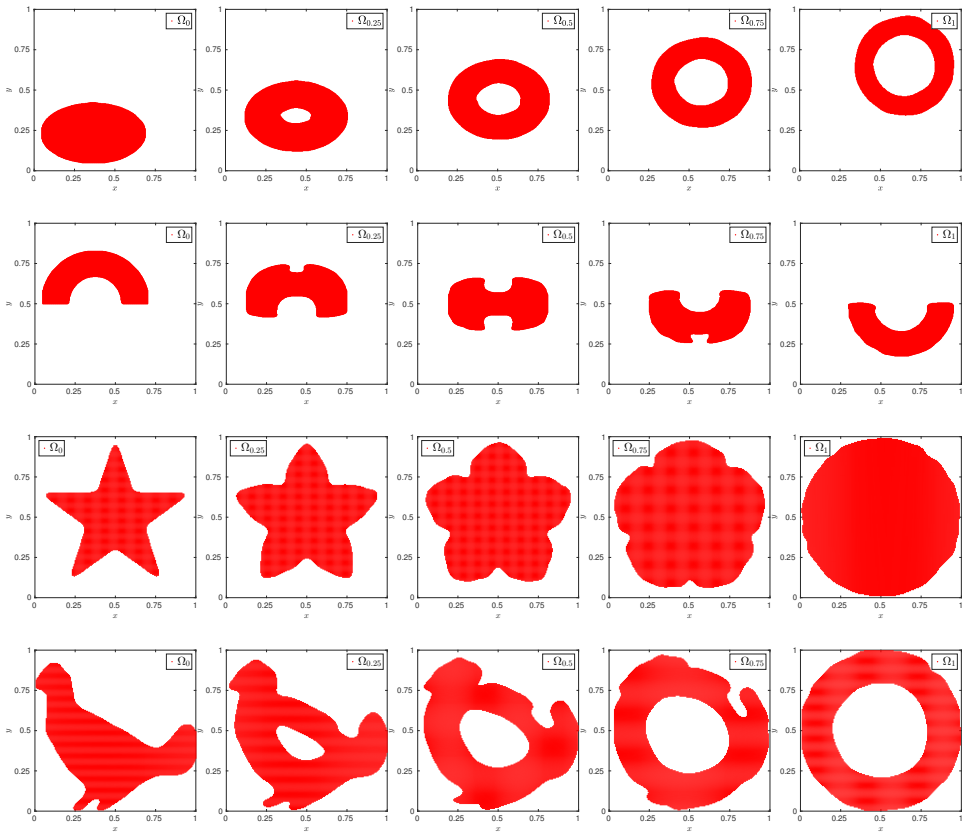


FIGURE 4.7. Each row shows a two-dimensional shape transform and intermediate shapes. The shapes are generated from the sample sets using the post-processing procedure. Intermediate shapes are shown for $t = 0.25, 0.5, 0.75$.

For each pixel $\mathbf{x}$, we define the map $\tilde{T}$:

$$\tilde{T}(\mathbf{x}^c) = \sum_j \omega_j(\mathbf{x}) T(\bar{\mathbf{x}}_j^c) \tag{4.10}$$

In the final step we apply iterative TMR filters [36] to the transferred image to restore sharp details of the original image.

While we used the framework introduced in [37], there are several differences in our implementation. First and foremost, we used the sample-based algorithm instead of the relaxed optimal transport. Although no relaxation is mentioned directly, adopting the sample-based setting itself can be considered as a relaxation, since in the sample-based setting we do not require source sample points to map to the exact locations of target sample points, and using a finite set of feature functions requires the distributions to match only in the desired subspaces.

The other difference is that we define optimal transport in CIELAB color space, because the Euclidean distance in CIELAB space better approximates the perceptual difference among colors [20].

For all the examples in this section, we use the super-pixel method with the number of super-pixels set to 2000 and the compactness parameter set to 2. For the SOT algorithm we use Gaussian feature functions on a $5 \times 5 \times 5$ grid with bandwidth $h = 1$. In the image synthesis step we set $\sigma^s = 10, \sigma^c = 1$.

In Fig. 4.8 we present step-by-step results of the full algorithm on the widely used parrot image, comparing it to existing methods. We can see that the final image Fig. 4.8e preserves sharp features in the source image and is free of the artificial defects in Figs. 4.8f and 4.8g. This is partially contributed by the post-processing step, although we can see that even without post-processing Fig. 4.8d is also very smooth in color. Comparing to Fig. 4.8h, the sample-based algorithm makes some different choices of color transfer. We can see that visually it is closer to the raw optimal transport in Fig. 4.8f, which suggests that the sample-based solution is closer to optimal in the optimal transport sense. In Figs. 4.8i to 4.8l, source, target and transferred sample sets are projected to *l*-*a* and *l*-*b* spaces. While both source and target sample sets have complex geometric features, the sample-based method is able to optimally transfer one to the other.

In Fig. 4.9 we present more examples of color transfer. The algorithm creates reasonable results in all three examples and the transferred sample set in *l*-*a* does agree with the target sample set. It is also worth noting that while the sample-based approach transfers the color distribution, it doesn't always create visually perfect results. For instance in the flower example the transferred figure has purple in the background due to the fact that the purple color has more weights in the target image than in the source image.

5 Conclusions and future work

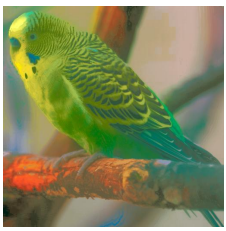
This article introduces a new family of optimal transport and barycenter algorithms for finite sample sets. Two features are unique to these algorithms. One is that they solve the optimal transport and barycenter problems iteratively, by approximating an adaptive number of interpolant measures using key properties 2.3 and 2.4. In each iteration, the local updates alternate with a global update, in which new sample sets are constructed as interpolants of the global map. The second feature is that the sample sets are compared with each other through a set of feature functions, whose gradients also define the function space to which the maps considered are constrained.

Through synthetic examples, we found that these algorithms converge to the optimal solution in just a handful of iterations. Also, the new algorithms give encouraging results in applications such as the shape transform and the color transfer problems.

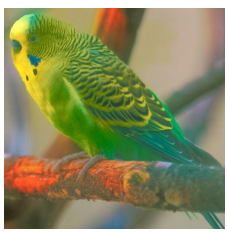
(A) Source Image

(B) Target Image

(C) Space-color Clusters



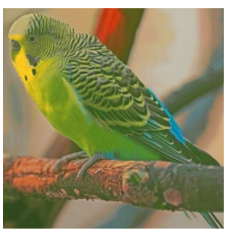
(D) Image Synthesis



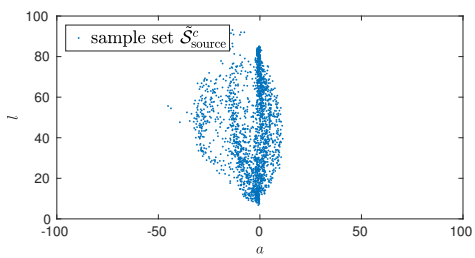
(E) Post-processing

(F) [34]

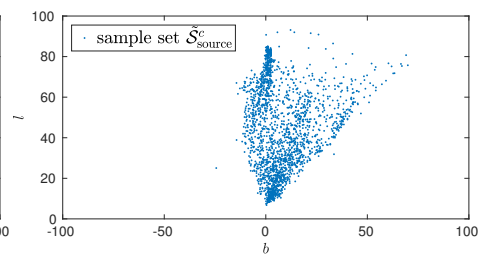
(G) [32]



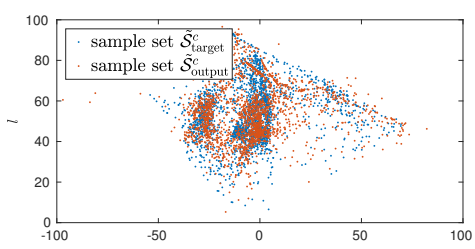
(H) [37]



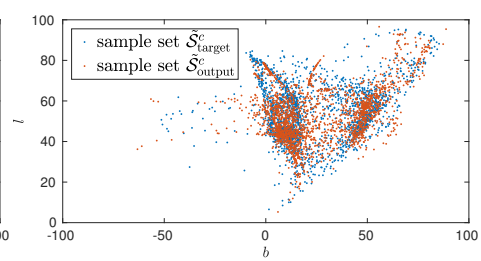
(I) Source Sample Set - *la* space



(J) Source Sample Set - *lb* space



(K) Target and Output Sample Sets - *la* space



(L) Target and Output Sample Sets - *lb* space

FIGURE 4.8. Color transfer of parrot images by the SOT algorithm.

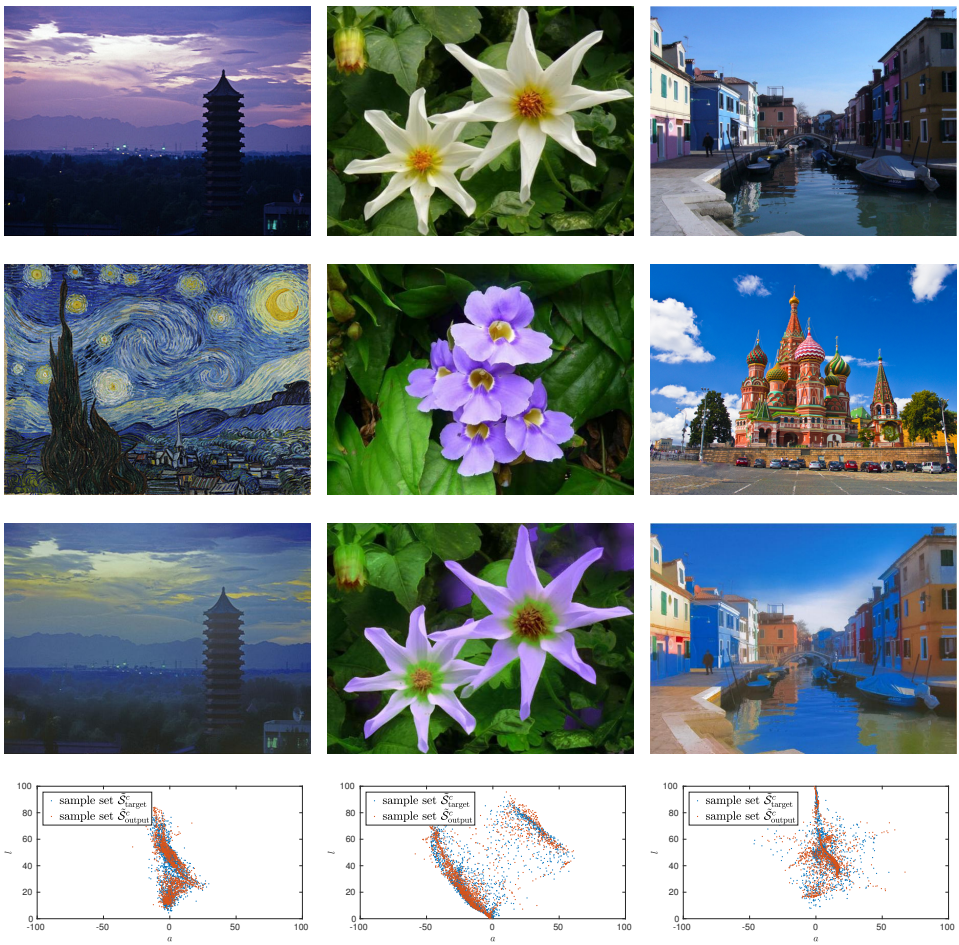


FIGURE 4.9. More Color Transfer Examples, with source images on the top row, targets on the second row and the results of transferring on the third. The bottom row displays the sample points of the target and those of the transformed source.

One possible extension is to develop other algorithms under the same theoretical framework: since we proved general convergence theorems for GDOT 2.7 and GDB 2.9, which are not limited to the proposed SOT (7) and CSB (8) algorithms, one can use these theorems to prove convergence for all those algorithms that meet the bare requirements of an L-descending map (2.5). For instance, a useful extension of the SOT algorithm consists of choosing adaptively the time variables t_i of the interpolant measures. Since the amount of improvement in the transportation cost of each interpolant segment can be different after applying a local update, it

is natural to put less effort into those segments of the interpolant that are closer to convergence. This improvement can potentially speed up the SOT algorithm.

Several improvements can be made in the sample-based formulation. An extension of the sample-based equivalence (3.1) consists of relaxing the equality of expected feature values, since exact equality is not a strict requirement in the presence of sample noise in the empirical expected values.

One important component of the algorithm is the selection of feature functions. While we looked into several possibilities, including moments and kernel functions, improvements could be made to better incorporate information from data, especially when the dimensionality of the sample space is high.

Other extensions under development include applications to constrained density estimation and the computation of the sample-based barycenter of probability measures conditioned on continuous variables.

Acknowledgment.

This work has been partially supported by grants from the division of mathematical sciences of the NSF and the ONR.

Bibliography

[1] Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence* **34** (2012), no. 11, 2274–2282.

[2] Agueh, M.; Carlier, G. Barycenters in the wasserstein space. *SIAM Journal on Mathematical Analysis* **43** (2011), no. 2, 904–924.

[3] Álvarez-Esteban, P. C.; del Barrio, E.; Cuesta-Albertos, J.; Matrán, C. A fixed-point approach to barycenters in wasserstein space. *Journal of Mathematical Analysis and Applications* **441** (2016), no. 2, 744–762.

[4] Ambrosio, L.; Gigli, N.; Savaré, G. *Gradient flows: in metric spaces and in the space of probability measures*, Springer Science & Business Media, 2008.

[5] Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein gan. *arXiv preprint arXiv:1701.07875* (2017).

[6] Benamou, J.-D.; Brenier, Y. A computational fluid mechanics solution to the monge-kantorovich mass transfer problem. *Numerische Mathematik* **84** (2000), no. 3, 375–393.

[7] Berger, A. L.; Pietra, V. J. D.; Pietra, S. A. D. A maximum entropy approach to natural language processing. *Computational linguistics* **22** (1996), no. 1, 39–71.

[8] Burkard, R. E.; Cela, E. *Linear assignment problems and extensions*, Springer, 1999.

[9] Chen, W.; Tabak, E. G. An adaptive linear programming methodology for data driven optimal transport .

[10] Cheng, Y. Mean shift, mode seeking, and clustering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **17** (1995), no. 8, 790–799.

[11] Conn, A. R.; Gould, N. I.; Toint, P. L. *Trust region methods*, vol. 1, Siam, 2000.

[12] Evans, L. C. Partial differential equations and monge-kantorovich mass transfer. *Current developments in mathematics* (1997), 65–126.

[13] Faridul, H. S.; Pouli, T.; Chamaret, C.; Stauder, J.; Reinhard, E.; Kuzovkin, D.; Trémeau, A.: Colour mapping: A review of recent methods, extensions and applications, in *Computer Graphics Forum*, Wiley Online Library, 2015 .

[14] Ferradans, S.; Papadakis, N.; Peyré, G.; Aujol, J.-F. Regularized discrete optimal transport. *SIAM Journal on Imaging Sciences* **7** (2014), no. 3, 1853–1882.

[15] Galichon, A. *Optimal Transport Methods in Economics*, Princeton University Press, 2016.

[16] Gangbo, W.; McCann, R. J. The geometry of optimal transportation. *Acta Mathematica* **177** (1996), no. 2, 113–161.

[17] Gangbo, W.; Swiech, A. Optimal maps for the multidimensional monge-kantorovich problem. *Communications on pure and applied mathematics* **51** (1998), no. 1, 23–45.

[18] Givens, C. R.; Shortt, R. M.; et al. A class of wasserstein metrics for probability distributions. *The Michigan Mathematical Journal* **31** (1984), no. 2, 231–240.

[19] Hubbard, J. H.; Hubbard, B. B. *Vector calculus, linear algebra, and differential forms: a unified approach*, 1999.

[20] Jain, A. K. *Fundamentals of digital image processing*, Prentice-Hall, Inc., 1989.

[21] Jones, M. C.; Marron, J. S.; Sheather, S. J. A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association* **91** (1996), no. 433, 401–407.

[22] Kantorovitch, L. V. On the translocation of masses. *Management Science* **5** (1958), no. 1, 1–4.

[23] Kuang, M.; Tabak, E. G. Preconditioning of optimal transport. *SIAM Journal of Scientific Computing* (2017).

[24] Laurence, P.; Pignol, R. J.; Tabak, E. G. Constrained density estimation. in *Quantitative Energy Finance*, pp. 259–284, Springer, 2014.

[25] McCann, R. J. A convexity principle for interacting gases. *Advances in mathematics* **128** (1997), no. 1, 153–179.

[26] Monge, G. *Mémoire sur la théorie des déblais et des remblais*, De l'Imprimerie Royale, 1781.

[27] Montavon, G.; Müller, K.-R.; Cuturi, M. Wasserstein training of boltzmann machines. *arXiv preprint arXiv:1507.01972* (2015).

[28] Moré, J. J. The levenberg-marquardt algorithm: implementation and theory. in *Numerical analysis*, pp. 105–116, Springer, 1978.

[29] Munkres, J. Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics* **5** (1957), no. 1, 32–38.

[30] Ni, K.; Bresson, X.; Chan, T.; Esedoglu, S. Local histogram based segmentation using the wasserstein distance. *International journal of computer vision* **84** (2009), no. 1, 97–111.

[31] Oberman, A. M.; Ruan, Y. An efficient linear programming method for optimal transportation. *arXiv preprint arXiv:1509.03668* (2015).

[32] Papadakis, N.; Provenzi, E.; Caselles, V. A variational model for histogram transfer of color images. *IEEE Transactions on Image Processing* **20** (2011), no. 6, 1682–1695.

[33] Peleg, S.; Werman, M.; Rom, H. A unified approach to the change of resolution: Space and gray-level. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **11** (1989), no. 7, 739–742.

[34] Pitié, F.; Kokaram, A. C.; Dahyot, R. Automated colour grading using colour distribution transfer. *Computer Vision and Image Understanding* **107** (2007), no. 1, 123–137.

[35] Prokhorov, Y. V. Convergence of random processes and limit theorems in probability theory. *Theory of Probability & Its Applications* **1** (1956), no. 2, 157–214.

[36] Rabin, J.; Delon, J.; Gousseau, Y. Removing artefacts from color and contrast modifications. *IEEE Transactions on Image Processing* **20** (2011), no. 11, 3073–3085.

[37] Rabin, J.; Ferradans, S.; Papadakis, N.: Adaptive color transfer with relaxed optimal transport, in *Image Processing (ICIP), 2014 IEEE International Conference on*, IEEE, 2014 pp. 4852–4856.

[38] Reinhard, E.; Adhikhmin, M.; Gooch, B.; Shirley, P. Color transfer between images. *IEEE Computer graphics and applications* **21** (2001), no. 5, 34–41.

[39] Rubner, Y.; Tomasi, C.; Guibas, L. J.: A metric for distributions with applications to image databases, in *Computer Vision, 1998. Sixth International Conference on*, IEEE, 1998 pp. 59–66.

[40] Sandhu, R.; Georgiou, T.; Reznik, E.; Zhu, L.; Kolesov, I.; Senbabaoglu, Y.; Tannenbaum, A. Graph curvature for differentiating cancer networks. *Scientific reports* **5** (2015), 12 323.

[41] Schölkopf, B.; Smola, A.; Müller, K.-R. Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation* **10** (1998), no. 5, 1299–1319.

[42] Solomon, J.; De Goes, F.; Peyré, G.; Cuturi, M.; Butscher, A.; Nguyen, A.; Du, T.; Guibas, L. Convolutional wasserstein distances: Efficient optimal transportation on geometric domains. *ACM Transactions on Graphics (TOG)* **34** (2015), no. 4, 66.

[43] Tabak, E. G.; Trigila, G. Data-driven optimal transport. *Commun. Pure. Appl. Math. doi* **10** (2014), 1002.

[44] Tai, Y.-W.; Jia, J.; Tang, C.-K.: Local color transfer via probabilistic segmentation by expectation-maximization, in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1, IEEE, 2005 pp. 747–754.

[45] Villani, C. *Optimal transport: old and new*, vol. 338, Springer Science & Business Media, 2008.

[46] Welsh, T.; Ashikhmin, M.; Mueller, K.: Transferring color to greyscale images, in *ACM Transactions on Graphics (TOG)*, vol. 21, ACM, 2002 pp. 277–280.

Received Month 200X.