

## SUPERVISED OPTIMAL TRANSPORT\*

ZIXUAN CANG<sup>†</sup>, QING NIE<sup>‡</sup>, AND YANXIANG ZHAO<sup>§</sup>

**Abstract.** Optimal transport, a theory for optimal allocation of resources, is widely used in various fields such as astrophysics, machine learning, and imaging science. However, many applications impose elementwise constraints on the transport plan which traditional optimal transport cannot enforce. Here we introduce supervised optimal transport (sOT), which formulates a constrained optimal transport problem where couplings between certain elements are prohibited according to specific applications. sOT is proved to be equivalent to an  $l^1$  penalized optimization problem, from which efficient algorithms are designed to solve its entropy regularized formulation. We demonstrate the capability of sOT by comparing it to other variants and extensions of traditional optimal transport in the color transfer problem. We also study the barycenter problem in sOT formulation, where we discover and prove a unique reverse and portion selection (control) mechanism. Supervised optimal transport is broadly applicable to applications in which a constrained transport plan is involved and the original unit should be preserved by avoiding normalization.

**Key words.** constrained transport plan, infinity cost matrix, unnormalized marginal distributions, entropic regularization, Dykstra algorithm

**MSC codes.** 49Q22, 90B06

**DOI.** 10.1137/22M1469171

**1. Introduction.** Optimal transport (OT) is a powerful tool for geometrically comparing and connecting measures. It seeks a globally optimal coupling between two probability distributions that minimizes the total coupling cost given a predefined finite cost [30, 24, 8, 40]. OT has been successfully applied in many fields recently, such as astrophysics [20], machine learning [15, 3, 14], and imaging science [4, 25]. The original OT is a linear programming problem which has a computational complexity of  $O(n^3)$  [40]. Recently, significant advancements in OT computation have been made which enable the application of OT to large scale practical problems, for example, the Sinkhorn algorithm [15], the Greenkhorn algorithm [2], and others [16, 22, 17, 29, 23].

However, there are limitations of OT that hinder its application to many problems, leading to several variants and extensions of OT. For example, unbalanced OT was introduced to couple nonprobability measures and reduce noise in a transport plan by replacing the original marginal constraints by soft divergence constraints [12, 11]. Partial OT generalizes OT to optimize the transport plan under the condition that a given fraction of mass is transported [6, 7, 10]. From the dynamics model

\*Received by the editors January 4, 2022; accepted for publication (in revised form) June 15, 2022; published electronically October 31, 2022.

<https://doi.org/10.1137/22M1469171>

**Funding:** The work of the first author was supported by a startup grant at North Carolina State University. The work of the second author was partially supported by NSF grant DMS1763272, a grant from the Simons Foundation (694598, QN), and NIH grant U01AR073159. The work of the third author was partially supported by a grant from the Simons Foundation through grant 357963 and by NSF grant DMS-2142500.

<sup>†</sup>Department of Mathematics, Center for Research in Scientific Computation, North Carolina State University, Raleigh, NC 27695 USA (zcang@ncsu.edu).

<sup>‡</sup>Corresponding author. Department of Mathematics, NSF-Simons Center for Multiscale Cell Fate Research, University of California, Irvine, Irvine, CA 92697 USA (qnie@uci.edu).

<sup>§</sup>Co-corresponding author. Department of Mathematics, George Washington University, Washington, DC 20052 USA (yxzhao@email.gwu.edu).

perspective, unnormalized OT was introduced to derive the transport dynamics between two marginals of different total mass with an external spatial-dependent or spatiotemporal-dependent mass source [21, 26]. In summary, these OT variants relax the marginal mass conservation constraint in the original OT to handle problems where the total masses of the two marginals do not match.

Another major limitation of OT is that there are natural constraints on the transport plans in many applications which cannot be handled by current OT methods. For instance, when ground transportation is blocked after a major natural disaster, many locations with a demand for resources cannot be safely reached by certain supply distribution locations. In the corresponding OT formulation, there should be constraints on the transport plan, causing some entries in the transport plan to be occluded as zero. This leads to a challenging OT problem since the total possible transported mass becomes an unknown due to the elementwise blockages in the transport plan.

Here, we introduce supervised optimal transport (sOT), which supervises the transport plan by enforcing a given elementwise constraint on the transport plan. sOT optimizes both the total transported mass and the transport plan simultaneously. Different from the OT problems with prescribed inequality constraints [6], the inequality constraints in sOT, arising due to the infinity entries in the cost matrix, are implicitly determined through the optimization of the transport plan. We show that sOT can be equivalently reformulated and linked to the unbalanced OT framework [11]. We further extend the standard OT barycenter problem into a sOT barycenter problem, in which an interesting and novel *reverse and portion selection mechanism* is discovered. We propose several new numerical methods for entropy regularized sOT based on Dykstra iteration.

We validate sOT and the proposed numerical algorithms in several numerical experiments. By applying it to an important problem in imaging science, the color transfer problem, we show the benefit and unique capability of sOT over other variants and extensions of traditional OT. More importantly, we prove the reverse and portion selection mechanism for the sOT barycenter problem, which is further validated in detail by numerical examples.

**2. Supervised optimal transport.** In this section, we define the sOT and derive an equivalent formulation upon which efficient algorithms are derived.

**2.1. Definition of sOT.** Let  $\mathbb{R}_+^n$  (and  $\mathbb{R}_{++}^n$ , respectively) denote the  $n$ -dimensional nonnegative (and positive, respectively) vector space. We define the probability simplex (and strictly positive probability simplex, respectively) as

$$(2.1) \quad \Sigma_+^n = \left\{ \mathbf{a} = (a_i)_i \in \mathbb{R}_+^n : \sum_i a_i = 1 \right\}, \quad \Sigma_{++}^n = \left\{ \mathbf{a} = (a_i)_i \in \mathbb{R}_{++}^n : \sum_i a_i = 1 \right\}.$$

The polytope of the couplings between  $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}_+^n \times \mathbb{R}_+^m$  is defined as

$$\mathbf{U}(\mathbf{a}, \mathbf{b}) = \{ \mathbf{P} \in \mathbb{R}_+^{n \times m} : \mathbf{P}\mathbf{1} = \mathbf{a}, \mathbf{P}^T\mathbf{1} = \mathbf{b} \},$$

where  $\mathbf{P}^T$  is the transpose of  $\mathbf{P}$  and  $\mathbf{1}$  is the all-ones matrix. The dimension of  $\mathbf{1}$  is determined by dimension consistency of matrix multiplication in the context. We further define the following polyhedra:

$$\mathbf{U}(\leq \mathbf{a}, \leq \mathbf{b}) = \{\mathbf{P} \in \mathbb{R}_+^{n \times m} : \mathbf{P}\mathbf{1} \leq \mathbf{a}, \mathbf{P}^T\mathbf{1} \leq \mathbf{b}\},$$

$$\mathbf{U}(= \mathbf{a}, \leq \mathbf{b}) = \{\mathbf{P} \in \mathbb{R}_+^{n \times m} : \mathbf{P}\mathbf{1} = \mathbf{a}, \mathbf{P}^T\mathbf{1} \leq \mathbf{b}\}.$$

We denote by  $\iota_{\mathcal{C}}$  the indicator of a set  $\mathcal{C}$ ,

$$\iota_{\mathcal{C}}(x) = \begin{cases} 0 & \text{if } x \in \mathcal{C}, \\ \infty & \text{otherwise.} \end{cases}$$

For  $\mathbf{P} = (P_{ij}) \in \mathbb{R}_+^{n \times m}$ , we define its entropy as

$$H(\mathbf{P}) = - \sum_{i,j} P_{ij} (\log P_{ij} - 1),$$

in which we use the convention  $0 \log 0 = 0$ . The Kullback–Leibler (KL) divergence between  $\mathbf{P} = (P_{ij}) \in \mathbb{R}_+^{n \times m}$  and  $\mathbf{Q} = (Q_{ij}) \in \mathbb{R}_{++}^{n \times m}$  is defined as

$$\text{KL}(\mathbf{P}|\mathbf{Q}) = \sum_{i,j} P_{ij} \log \left( \frac{P_{ij}}{Q_{ij}} \right) - P_{ij} + Q_{ij}.$$

For two vectors  $\mathbf{u} = (u_i), \mathbf{v} = (v_i)$  of the same dimension, we denote entrywise multiplication and division by

$$\mathbf{u} \odot \mathbf{v} = (u_i v_i)_i, \quad \mathbf{u} ./ \mathbf{v} = (u_i / v_i)_i.$$

For the standard Kantorovich's OT problem with discrete marginal measures  $\mathbf{a}, \mathbf{b} \in \Sigma_+^n$ , it reads

$$(2.2) \quad L_{\text{OT}}(\mathbf{a}, \mathbf{b}; \mathbf{C}) = \min_{\mathbf{P} \in \mathbf{U}(\mathbf{a}, \mathbf{b})} \langle \mathbf{P}, \mathbf{C} \rangle = \sum_{i,j} \mathbf{P}_{ij} \mathbf{C}_{ij}.$$

In the framework of sOT, the marginal measures  $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}_+^n \times \mathbb{R}_+^m$  do not necessarily have the same sum, and we are interested in the cost matrix  $\mathbf{C} = (C_{ij})$  that contains  $\infty$  entries. The  $\infty$ -pattern of  $\mathbf{C}$  is defined as the set [9]

$$(2.3) \quad \mathcal{P}_{\infty}(\mathbf{C}) = \{(i, j) : C_{ij} = \infty, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m\},$$

of positions of  $\mathbf{C}$  containing an infinity element. Similarly one can define the 0-pattern of a transport plan  $\mathbf{P}$ . By virtue of OT, the 0-pattern of the optimal plan  $\mathbf{P}^*$  must contain the  $\infty$ -pattern of  $\mathbf{C}$ , namely,  $\mathcal{P}_0(\mathbf{P}^*) \supseteq \mathcal{P}_{\infty}(\mathbf{C})$ . We further define a feasible set  $\mathcal{A}_{\mathbf{C}}$  for the *marginal blocked distribution*  $(\boldsymbol{\mu}, \boldsymbol{\nu})$  as follows:

$$(2.4) \quad \mathcal{A}_{\mathbf{C}} = \left\{ (\boldsymbol{\mu}, \boldsymbol{\nu}) \in [0, \mathbf{a}] \times [0, \mathbf{b}] \mid \exists \mathbf{P} \in \mathbf{U}(\mathbf{a} - \boldsymbol{\mu}, \mathbf{b} - \boldsymbol{\nu}) \text{ such that } \langle \mathbf{P}, \mathbf{C} \rangle < \infty \right\},$$

in which the inclusion  $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in [0, \mathbf{a}] \times [0, \mathbf{b}]$  is entrywise, namely,  $\mu_i \in [0, a_i], i = 1, 2, \dots, n$ , and  $\nu_j \in [0, b_j], j = 1, 2, \dots, m$ .

We define sOT as the following minimization problem:

$$(2.5) \quad L_{\text{sOT}}(\mathbf{a}, \mathbf{b}; \mathbf{C}) := \min_{(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathcal{B}} \min_{\mathbf{P} \in \mathbf{U}(\mathbf{a} - \boldsymbol{\mu}, \mathbf{b} - \boldsymbol{\nu})} \langle \mathbf{P}, \mathbf{C} \rangle,$$

where  $\mathcal{B} := \arg\min_{(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathcal{A}_{\mathbf{C}}} \|\boldsymbol{\mu}\|_1 + \|\boldsymbol{\nu}\|_1$ . In other words, we aim to find the OT plan  $\mathbf{P}$  which transports the most marginal density (blocks the least  $(\boldsymbol{\mu}, \boldsymbol{\nu})$ ) with minimal cost.

*Remark 2.1.* Note that when the cost matrix  $\mathbf{C}$  contains  $\infty$ ,  $(\boldsymbol{\mu}, \mathbf{v}) = (\mathbf{0}, \mathbf{0})$  may still be a feasible point in  $\mathcal{A}_{\mathbf{C}}$  (and therefore in  $\mathcal{B}$ ). For example, one can consider

$$\mathbf{C} = \begin{bmatrix} 1 & \infty & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 0.2 \\ 0.3 \\ 0.5 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0.4 \\ 0.3 \\ 0.3 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 0.1 & 0 & 0.1 \\ 0.1 & 0.1 & 0.1 \\ 0.2 & 0.2 & 0.1 \end{bmatrix},$$

then  $(\boldsymbol{\mu}, \mathbf{v}) = (\mathbf{0}, \mathbf{0}) \in \mathcal{A}_{\mathbf{C}}$  since  $\mathbf{P} \in \mathbf{U}(\mathbf{a}, \mathbf{b})$  and  $\langle \mathbf{P}, \mathbf{C} \rangle$  is finite. In this case, sOT reduces to the standard OT as  $\mathbf{U}(\mathbf{a}, \mathbf{b})$  is nonempty and the minimum is reached at some optimal  $\mathbf{P}^*$ .

In practical applications, the cost matrix  $\mathbf{C}$  could be sparse with respect to the  $\infty$  entries, which most likely leads to  $(\boldsymbol{\mu}, \mathbf{v}) = (\mathbf{0}, \mathbf{0}) \notin \mathcal{A}_{\mathbf{C}}$ . On the other hand,  $\mathcal{A}_{\mathbf{C}}$  is nonempty as  $(\boldsymbol{\mu}, \mathbf{v}) = (\mathbf{a}, \mathbf{b})$  is always an element in  $\mathcal{A}_{\mathbf{C}}$ . Therefore, we expect to find some  $(\boldsymbol{\mu}, \mathbf{v}) \in [0, \mathbf{a}] \times [0, \mathbf{b}]$  with smallest  $l^1$ -norm, and the associated OT plan  $\mathbf{P}^*$  over  $\mathbf{U}(\mathbf{a} - \boldsymbol{\mu}, \mathbf{b} - \mathbf{v})$ .

A problem related to sOT (2.5) is the partial transport problem [6], which finds the optimal plan to transport a given fraction of mass instead of the total amount of marginal mass. It is formulated as follows. Given marginal densities  $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}_{++}^n \times \mathbb{R}_{++}^m$ , not necessarily with the same total mass, the partial transport problem minimizes

$$(2.6) \quad \min_{\mathbf{P} \in \mathbf{U}(\leq \mathbf{a}, \leq \mathbf{b})} \{ \langle \mathbf{P}, \mathbf{C} \rangle : \langle \mathbf{P}, \mathbf{1} \rangle = \theta \},$$

in which  $\theta$  is the given fraction of mass. Different from partial OT, sOT does not require a given fraction of mass to be transported and instead optimizes the transported mass. Specifically, sOT can be rewritten as

$$(2.7) \quad \max_{\theta \in [0, \min(\mathbf{a}^T \mathbf{1}, \mathbf{b}^T \mathbf{1})]} \min_{\mathbf{P} \in \mathbf{U}(\leq \mathbf{a}, \leq \mathbf{b})} \{ \langle \mathbf{P}, \mathbf{C} \rangle : \langle \mathbf{P}, \mathbf{1} \rangle = \theta \},$$

where sOT performs an extra maximization over the transported mass  $\theta$ . Note that if the cost matrix  $\mathbf{C}$  does not contain  $\infty$  entries and  $\mathbf{a}^T \mathbf{1} = \mathbf{b}^T \mathbf{1}$ , the reformulated problem (2.7) degenerates to the standard OT problem (2.2) since  $\theta$  can be equal to the total mass, the largest possible value. On the other hand, if the cost matrix  $\mathbf{C}$  contains  $\infty$  entries, the partial transport problem (2.6) may not be well defined for a given fraction of mass  $\theta$  since the feasible set is potentially empty. In this case, sOT (2.7) remains wellposed as  $\theta = 0$  is always a feasible transported mass.

**2.2. Equivalent sOT formulations.** The sOT in the forms (2.5) or (2.7) is a double optimization problem which is computationally challenging. Here, we recast it to the form of a single optimization:

$$(2.8) \quad L_{\text{sOT}}(\mathbf{a}, \mathbf{b}; \mathbf{C}) = \min_{\substack{(\boldsymbol{\mu}, \mathbf{v}) \in \mathcal{A}_{\mathbf{C}} \\ \mathbf{P} \in \mathbf{U}(\mathbf{a} - \boldsymbol{\mu}, \mathbf{b} - \mathbf{v})}} \langle \mathbf{P}, \mathbf{C} \rangle + \gamma(\|\boldsymbol{\mu}\|_1 + \|\mathbf{v}\|_1)$$

for sufficiently large  $\gamma$  depending on  $\mathbf{C}$ . The equivalence between (2.5) and (2.8) is summarized in Lemma 2.3.

To begin the proof, we need a lemma characterizing the difference between  $\mathbf{P}_{\theta}$  and  $\mathbf{P}_{\bar{\theta}}$  for the maximal possible transported mass  $\bar{\theta}$  and a mass  $\theta$  smaller than but close to  $\bar{\theta}$ .



LEMMA 2.2. Let  $C = \{x : Ax \leq b\}$  be a bounded convex polyhedron where  $A = [a_1^T | \dots | a_r^T]^T \in \mathbb{R}^{r \times n}$  and  $b \in \mathbb{R}^r$ . Define

$$\bar{\theta} := \max_{\bar{x} \in C} \mathbf{1}^T \bar{x}, \quad C_{\bar{\theta}} = \operatorname{argmax}_{\bar{x} \in C} \mathbf{1}^T \bar{x},$$

and

$$C_{\theta} = \{x \in C : \mathbf{1}^T x = \theta\},$$

for any possible value of  $\theta$  that  $\mathbf{1}^T x$  can take. Then there exists a critical  $\theta_0 < \bar{\theta}$  such that for any fixed  $\theta \in (\theta_0, \bar{\theta}]$ , and any  $x \in C_{\theta}$ , there exists an  $\bar{x} \in C_{\bar{\theta}}$  such that  $\|x - \bar{x}\|_1 \leq \eta |\bar{\theta} - \theta|$ , where  $\eta$  is independent of  $\theta$ .

*Proof.* By definition,  $C_{\theta}$  is a bounded convex polyhedron, which can be reformulated as the convex hull of its extreme points:

$$C_{\theta} = \operatorname{conv} \left( x_{\theta}^{(1)}, x_{\theta}^{(2)}, \dots, x_{\theta}^{(s(\theta))} \right)$$

in which  $x_{\theta}^{(1)}, x_{\theta}^{(2)}, \dots, x_{\theta}^{(s(\theta))}$  are all extreme points of  $C_{\theta}$ . We denote by  $J_{\theta} := \{j : a_j^T x_{\theta}^{(k)} = b_j \text{ for some } 1 \leq k \leq s(\theta)\}$  the indices of constraints which are saturated in at least one extreme point of  $C_{\theta}$ . In other words,  $J_{\theta}$  are the indices of constraints which interact with hyperplane  $\mathbf{1}^T x = \theta$  on  $C$ . Letting  $E_C = \{x_1, \dots, x_m\}$  be the set of extreme points of  $C$  such that  $C = \operatorname{conv}(E_C)$ , we denote by  $E_j := \{x \in E_C : a_j^T x = b_j\}$  the extreme points of  $C$  saturating the  $j$ th constraint, and  $E_{J_{\theta}} = \cup_{j \in J_{\theta}} E_j$ .

Evaluating the linear function  $\mathbf{1}^T x$  over  $E_C$ , we know that the maximal value over  $\{\mathbf{1}^T x_1, \dots, \mathbf{1}^T x_m\}$  equals  $\bar{\theta}$ . Let the second maximal value over  $\{\mathbf{1}^T x_1, \dots, \mathbf{1}^T x_m\}$  be  $\theta_0$ . Note that in an extreme case where the second maximal value  $\theta_0$  cannot be attained, the maximal value  $\bar{\theta}$  is reached in the entire  $C$ , and the conclusion of the lemma is trivially held. So we only consider the nontrivial case where the second maximal value  $\theta_0$  is attained. It is evident that  $E_{J_{\theta}}$  is invariant for any  $\theta \in (\theta_0, \bar{\theta}]$ .

For any  $\theta \in (\theta_0, \bar{\theta})$ , we consider the extreme points  $\{x_{\theta}^{(j)}\}_{j=1}^{s(\theta)}$  of  $C_{\theta}$  in which  $s(\theta) \equiv s$ . As  $\theta \rightarrow \bar{\theta}$ , we have  $x_{\theta}^{(j)} \rightarrow \bar{x}_{\bar{\theta}}^{(j)}$  for  $j = 1, \dots, s$ . Here  $\{\bar{x}_{\bar{\theta}}^{(j)}\}_{j=1}^s$  are the extreme points of  $C_{\bar{\theta}}$ , some of which might be repeatedly counted. Now for any point  $x \in C_{\theta}$  such that

$$x = \lambda_1 x_{\theta}^{(1)} + \dots + \lambda_s x_{\theta}^{(s)}, \quad \{\lambda_j\}_{j=1}^s \in \Sigma_+^s,$$

we can take  $\bar{x}$  as

$$\bar{x} = \lambda_1 \bar{x}_{\bar{\theta}}^{(1)} + \dots + \lambda_s \bar{x}_{\bar{\theta}}^{(s)},$$

then

$$(2.9) \quad \|x - \bar{x}\|_1 \leq \max_{1 \leq j \leq s} \|x_{\theta}^{(j)} - \bar{x}_{\bar{\theta}}^{(j)}\|_1 \leq \|\mathbf{1}\|_2 \cdot \max_{1 \leq j \leq s} \|x_{\theta}^{(j)} - \bar{x}_{\bar{\theta}}^{(j)}\|_2.$$

Now we consider the angle  $\sigma_j$  made by  $\bar{x}_{\bar{\theta}}^{(j)} - x_{\theta}^{(j)}$  and  $\mathbf{1}$ ,

$$(2.10) \quad \sigma_j = \arccos \frac{\langle \bar{x}_{\bar{\theta}}^{(j)} - x_{\theta}^{(j)}, \mathbf{1} \rangle}{\|\bar{x}_{\bar{\theta}}^{(j)} - x_{\theta}^{(j)}\|_2 \|\mathbf{1}\|_2}.$$

It turns out that  $\sigma_j$  cannot be  $\pi/2$ . Otherwise,

$$0 = \langle \bar{x}_{\bar{\theta}}^{(j)} - x_{\theta}^{(j)}, \mathbf{1} \rangle = \bar{\theta} - \theta$$

implies that  $\theta = \bar{\theta}$ , which is the trivial case that  $\bar{\theta}$  is attained everywhere in  $C$ . Hence  $\{\cos \sigma_j\}_{j=1}^s$  must be bounded away from 0. Then we have

$$\cos \sigma_j \|\bar{x}_{\bar{\theta}}^{(j)} - x_{\theta}^{(j)}\|_2 \|\mathbf{1}\|_2 = \langle \bar{x}_{\bar{\theta}}^{(j)} - x_{\theta}^{(j)}, \mathbf{1} \rangle = \bar{\theta} - \theta,$$

which leads to

$$\|\bar{x}_{\bar{\theta}}^{(j)} - x_{\theta}^{(j)}\|_2 = \frac{1}{\cos \sigma_j \|\mathbf{1}\|_2} (\bar{\theta} - \theta).$$

Combining with (2.9), we have

$$\|x - \bar{x}\|_1 \leq \frac{1}{\min_{1 \leq j \leq s} \cos \sigma_j} |\bar{\theta} - \theta|.$$

The conclusion holds by taking  $\eta = (\min_{1 \leq j \leq s} \cos \sigma_j)^{-1}$ .

Last, we show that  $\cos \sigma_j$  is independent of  $\theta$ , or more specifically of  $x_{\theta}^{(j)}$ . To this end, we choose an arbitrary  $\tilde{\theta} \in (\theta, \bar{\theta})$ . Note that  $x_{\theta}^{(j)}$  and  $\bar{x}_{\bar{\theta}}^{(j)}$  saturate the same set of constraints since when  $\theta \rightarrow \bar{\theta}$ , the hyperplane  $\mathbf{1}^T x = \theta$  does not go through any extreme point of  $C$  ( $E_{J_{\theta}}$  is invariant for any  $\theta \in (\theta_0, \bar{\theta}]$ ); therefore the point

$$(2.11) \quad \tilde{x}_{\tilde{\theta}}^{(j)} := \frac{\tilde{\theta} - \theta}{\tilde{\theta} - \bar{\theta}} \bar{x}_{\bar{\theta}}^{(j)} + \frac{\bar{\theta} - \tilde{\theta}}{\bar{\theta} - \theta} x_{\theta}^{(j)},$$

which saturates the same set of constraints as  $x_{\theta}^{(j)}$  and  $\bar{x}_{\bar{\theta}}^{(j)}$ , is exactly the extreme point of  $C_{\tilde{\theta}}$  that falls on the line segment between  $x_{\theta}^{(j)}$  and  $\bar{x}_{\bar{\theta}}^{(j)}$ . Finally the linear relation (2.11) together with the definition of  $\sigma_j$  indicates that  $\cos \sigma_j$  is independent of  $\theta$ . The proof is completed. A schematic polyhedron is shown in Figure 1.  $\square$

**LEMMA 2.3.** *Given a cost matrix  $\mathbf{C}$  with  $\infty$ -pattern  $\mathcal{P}_{\infty}(\mathbf{C})$ , the two sOT formulations (2.5) and (2.8) are equivalent for sufficiently large  $\gamma$ .*

*Proof. Step I.* Let  $(\mu_{\text{opt}}, \nu_{\text{opt}})$  be an optimal pair of blocked measures for (2.5), namely, the system can at most transfer the amount of mass  $\bar{\theta} := \|\mathbf{a} - \mu_{\text{opt}}\|_1 =$

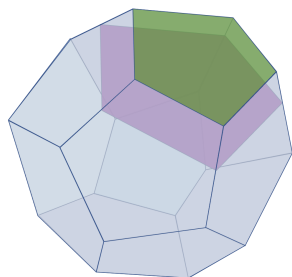


FIG. 1. A schematic polyhedron for Lemma 2.2. The green face is the hyperplane  $\mathbf{1}^T x = \bar{\theta}$ . The pink plane is the hyperplane  $\mathbf{1}^T x = \theta$ . When the pink plane is close enough to the green one, the set of extreme points  $E_{J_{\theta}}$  is invariant. See the proof of Lemma 2.2 for the definitions of the notation.

$\|\mathbf{b} - \mathbf{v}_{\text{opt}}\|_1$ . We denote by  $\mathbf{P}_{\bar{\theta}}^*$  a corresponding OT plan. Take any nonnegative and feasible  $\theta < \bar{\theta}$ ; if we can show that for a plan  $\mathbf{P}_{\theta}^*$  defined as

$$(2.12) \quad \mathbf{P}_{\theta}^* = \underset{\mathbf{P}}{\operatorname{argmin}} \{ \langle \mathbf{P}, \mathbf{C} \rangle : \mathbf{P} \in \mathbf{U}(\leq \mathbf{a}, \leq \mathbf{b}), \langle \mathbf{P}, \mathbf{1} \rangle = \theta \},$$

one has that

$$(2.13) \quad \langle \mathbf{P}_{\bar{\theta}}^* - \mathbf{P}_{\theta}^*, \mathbf{C} \rangle \leq 2\gamma \langle \mathbf{P}_{\bar{\theta}}^* - \mathbf{P}_{\theta}^*, \mathbf{1} \rangle,$$

for some  $\gamma > 0$ , then for any  $(\boldsymbol{\mu}, \mathbf{v})$  such that  $\theta = \|\mathbf{a} - \boldsymbol{\mu}\|_1 = \|\mathbf{b} - \mathbf{v}\|_1$ , it results in

$$\begin{aligned} & \|\boldsymbol{\mu}\|_1 + \|\mathbf{v}\|_1 - (\|\boldsymbol{\mu}_{\text{opt}}\|_1 + \|\mathbf{v}_{\text{opt}}\|_1) \\ &= \langle \boldsymbol{\mu} - \boldsymbol{\mu}_{\text{opt}}, \mathbf{1} \rangle + \langle \mathbf{v} - \mathbf{v}_{\text{opt}}, \mathbf{1} \rangle \\ &= 2\langle \mathbf{P}_{\bar{\theta}}^* - \mathbf{P}_{\theta}^*, \mathbf{1} \rangle \\ &\geq \gamma^{-1} \langle \mathbf{P}_{\bar{\theta}}^* - \mathbf{P}_{\theta}^*, \mathbf{C} \rangle, \end{aligned}$$

leading to

$$\langle \mathbf{P}_{\bar{\theta}}^*, \mathbf{C} \rangle + \gamma(\|\boldsymbol{\mu}_{\text{opt}}\|_1 + \|\mathbf{v}_{\text{opt}}\|_1) \leq \langle \mathbf{P}_{\theta}^*, \mathbf{C} \rangle + \gamma(\|\boldsymbol{\mu}\|_1 + \|\mathbf{v}\|_1),$$

which implies the optimality of  $(\boldsymbol{\mu}_{\text{opt}}, \mathbf{v}_{\text{opt}}, \mathbf{P}_{\bar{\theta}}^*)$  for (2.8), and consequently the equivalence holds.

*Step II.* We now prove that there exists a constant  $\gamma > 0$  such that (2.13) holds. Using Lemma 2.2, we know that there exists a critical  $\theta_0$  such that for any given  $\theta \in (\theta_0, \bar{\theta}]$  and  $\mathbf{P}_{\theta}^*$  defined in (2.12), we can find a feasible  $\mathbf{P}_{\bar{\theta}}$  satisfying  $\langle \mathbf{P}_{\bar{\theta}}, \mathbf{1} \rangle = \bar{\theta}$  such that

$$(2.14) \quad \|\mathbf{P}_{\bar{\theta}} - \mathbf{P}_{\theta}^*\|_1 \leq \eta|\bar{\theta} - \theta|.$$

Then

$$\begin{aligned} (2.15) \quad \langle \mathbf{P}_{\bar{\theta}}^* - \mathbf{P}_{\theta}^*, \mathbf{C} \rangle &\leq \langle \mathbf{P}_{\bar{\theta}} - \mathbf{P}_{\theta}^*, \mathbf{C} \rangle \\ &\leq \|\mathbf{P}_{\bar{\theta}} - \mathbf{P}_{\theta}^*\|_1 \cdot \|\mathbf{C}\|_{\infty} \\ &\leq \|\mathbf{C}\|_{\infty} \eta |\bar{\theta} - \theta| \\ &= \|\mathbf{C}\|_{\infty} \eta \langle \mathbf{P}_{\bar{\theta}}^* - \mathbf{P}_{\theta}^*, \mathbf{1} \rangle. \end{aligned}$$

On the other hand, for any feasible  $\theta \leq \theta_0$  and  $\mathbf{P}_{\theta}^*$  defined in (2.12), we simply have

$$\begin{aligned} (2.16) \quad \langle \mathbf{P}_{\bar{\theta}}^* - \mathbf{P}_{\theta}^*, \mathbf{C} \rangle &= \langle \mathbf{P}_{\bar{\theta}}^*, \mathbf{C} \rangle - \langle \mathbf{P}_{\theta}^*, \mathbf{C} \rangle \\ &\leq \|\mathbf{C}\|_{\infty} \cdot \bar{\theta} \\ &= \|\mathbf{C}\|_{\infty} \cdot \frac{\bar{\theta}}{\bar{\theta} - \theta} (\bar{\theta} - \theta) \\ &\leq \|\mathbf{C}\|_{\infty} \cdot \frac{\bar{\theta}}{\bar{\theta} - \theta_0} (\bar{\theta} - \theta) \\ &= \|\mathbf{C}\|_{\infty} \cdot \frac{\bar{\theta}}{\bar{\theta} - \theta_0} \langle \mathbf{P}_{\bar{\theta}}^* - \mathbf{P}_{\theta}^*, \mathbf{1} \rangle. \end{aligned}$$

Finally, combining the inequalities (2.15) and (2.16) and taking  $2\gamma = \max\{\eta, \bar{\theta}/(\bar{\theta} - \theta_0)\} \cdot \|\mathbf{C}\|_\infty$ , we prove the inequality (2.13), and therefore the optimality of  $(\boldsymbol{\mu}_{\text{opt}}, \boldsymbol{\nu}_{\text{opt}}, \mathbf{P}_\theta^*)$  for (2.8).  $\square$

Note that  $\mathcal{A}_{\mathbf{C}}$  is always nonempty; we can therefore rewrite the formulation (2.8) in a simpler form:

$$(2.17) \quad L_{\text{sOT}}(\mathbf{a}, \mathbf{b}; \mathbf{C}) = \min_{\mathbf{P} \in \mathbf{U}(\leq \mathbf{a}, \leq \mathbf{b})} \langle \mathbf{P}, \mathbf{C} \rangle + \gamma(\|\mathbf{a} - \mathbf{P}\mathbf{1}\|_1 + \|\mathbf{b} - \mathbf{P}^T\mathbf{1}\|_1).$$

**3. Entropic regularization of sOT.** The idea to regularize the standard OT problem by an entropic term can be traced back to the early work by Schrodinger [35]. This entropic regularization has been well motivated in economics for predicting flows of commodities or actors in a market, in which the smoothness of such flows can be guaranteed [1]. A recent work [15] provides a new motivation from the computational perspective that entropic regularization defines a strongly convex programming. Unlike the standard OT problem (2.2) which has multiple solutions, the entropic regularized OT problem has a unique solution, which corresponds to the optimizer of (2.2) with maximal entropy in the limit as the regularization parameter  $\epsilon$  vanishes. More importantly, the unique solution to the entropic regularized OT problem is simply a diagonal scaling of the matrix  $e^{-\mathbf{C}/\epsilon}$ . This diagonal scaling process can be efficiently implemented by the Sinkhorn algorithm [36, 38, 37], which has a linear rate of convergence [19].

We now consider the entropic regularization of sOT (2.8):

$$(3.1) \quad \min_{\substack{(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathcal{A}_{\mathbf{C}} \\ \mathbf{P} \in \mathbf{U}(\mathbf{a} - \boldsymbol{\mu}, \mathbf{b} - \boldsymbol{\nu})}} \langle \mathbf{P}, \mathbf{C} \rangle - \epsilon H(\mathbf{P}) + \gamma(\|\boldsymbol{\mu}\|_1 + \|\boldsymbol{\nu}\|_1)$$

or equivalently

$$(3.2) \quad \min_{\mathbf{P} \in \mathbf{U}(\leq \mathbf{a}, \leq \mathbf{b})} \langle \mathbf{P}, \mathbf{C} \rangle - \epsilon H(\mathbf{P}) + \gamma(\|\mathbf{a} - \mathbf{P}\mathbf{1}\|_1 + \|\mathbf{b} - \mathbf{P}^T\mathbf{1}\|_1).$$

It is well known that the unique solution  $\mathbf{P}_\epsilon^*$  of (3.2) converges to the optimal solution with maximal entropy within the set of all optimal solutions of the problem (2.17) [13].

Taking  $\mathbf{K} = \exp(-\mathbf{C}/\epsilon)$  as the Gibbs kernel, sOT problem (3.2) can be rewritten in terms of the KL divergence as

$$(3.3) \quad \min_{\mathbf{P} \in \mathbf{U}(\leq \mathbf{a}, \leq \mathbf{b})} \epsilon \text{KL}(\mathbf{P}|\mathbf{K}) + \gamma(\|\mathbf{a} - \mathbf{P}\mathbf{1}\|_1 + \|\mathbf{b} - \mathbf{P}^T\mathbf{1}\|_1),$$

or equivalently

$$(3.4) \quad \min_{\mathbf{P} \in \mathbb{R}_+^{n \times m}} \epsilon \text{KL}(\mathbf{P}|\mathbf{K}) + \gamma\|\mathbf{a} - \mathbf{P}\mathbf{1}\|_1 + \iota_{[0, \mathbf{a}]}(\mathbf{P}\mathbf{1}) + \gamma\|\mathbf{b} - \mathbf{P}^T\mathbf{1}\|_1 + \iota_{[0, \mathbf{b}]}(\mathbf{P}^T\mathbf{1}).$$

**3.1. Dykstra algorithm.** The entropic regularized sOT (3.4) fits into a more general form

$$(3.5) \quad \min_{\mathbf{P} \in \mathbb{R}_+^{n \times m}} B_g(\mathbf{P}|\mathbf{K}) + \hat{h}_1(\mathbf{P}) + \hat{h}_2(\mathbf{P}).$$

Here  $g$  is a given proper closed and strictly convex and differentiable function.  $B_g$  is the Bregman divergence (Bregman distance) defined as

$$(3.6) \quad B_g(\mathbf{P}|\mathbf{Q}) = g(\mathbf{P}) - g(\mathbf{Q}) - \langle \nabla g(\mathbf{Q}), \mathbf{P} - \mathbf{Q} \rangle.$$

In addition,  $\hat{h}_1$  and  $\hat{h}_2$  are two proper and lower semicontinuous convex functions.

Note that the Legendre transform of  $g$

$$g^*(y) = \max_x \langle x, y \rangle - g(x)$$

is also smooth and strictly convex. In particular one has that  $\nabla g$  and  $\nabla g^*$  are bijective functions such that  $\nabla g^* = (\nabla g)^{-1}$ .

Define the Bregman proximal operator of a convex function  $\phi$  as

$$(3.7) \quad \text{prox}_{\phi}^{B_g}(\mathbf{Q}) = \underset{\mathbf{P}}{\operatorname{argmin}} B_g(\mathbf{P}|\mathbf{Q}) + \phi(\mathbf{P}).$$

We assume that  $\phi$  is coercive so that  $\text{prox}_{\phi}^{B_g}(\mathbf{Q})$  is uniquely defined by strict convexity.

The Dykstra algorithm for problem (3.5) [31] reads as follows.

Dykstra algorithm for (3.5).

Input:  $\mathbf{P}^0 = \mathbf{K}$  and  $\lambda^{-1} = \lambda^0 = 0$ ;

General step: for any  $k = 0, 1, 2, \dots$ , execute the following steps:

$$(3.8) \quad \mathbf{P}^{2k+1} = \text{prox}_{\hat{h}_1}^{B_g} \left( \nabla g^* \left[ \nabla g(\mathbf{P}^{2k}) + \lambda^{2k-1} \right] \right);$$

$$(3.9) \quad \lambda^{2k+1} = \lambda^{2k-1} + \nabla g(\mathbf{P}^{2k}) - \nabla g(\mathbf{P}^{2k+1});$$

$$(3.10) \quad \mathbf{P}^{2k+2} = \text{prox}_{\hat{h}_2}^{B_g} \left( \nabla g^* \left[ \nabla g(\mathbf{P}^{2k+1}) + \lambda^{2k} \right] \right);$$

$$(3.11) \quad \lambda^{2k+2} = \lambda^{2k} + \nabla g(\mathbf{P}^{2k+1}) - \nabla g(\mathbf{P}^{2k+2}).$$

It is shown in [31] that the sequence  $\{\mathbf{P}^n\}_{n \geq 0}$  generated by the above Dykstra algorithm converges to the solution of the problem (3.5).

When taking  $g(\cdot)$  as the entropy function, the corresponding Bregman divergence  $B_g(\mathbf{P}|\mathbf{K}) = \text{KL}(\mathbf{P}|\mathbf{K})$  becomes the KL divergence. In this case, if  $\hat{h}_1$  and  $\hat{h}_2$  in (3.5) are of the special form as

$$\hat{h}_1(\mathbf{P}) = h_1(\mathbf{P}\mathbf{1}), \quad \hat{h}_2(\mathbf{P}) = h_2(\mathbf{P}^T\mathbf{1}),$$

then the problem (3.5) reduces to

$$(3.12) \quad \min_{\mathbf{P}} \text{KL}(\mathbf{P}|\mathbf{K}) + h_1(\mathbf{P}\mathbf{1}) + h_2(\mathbf{P}^T\mathbf{1}),$$

which is consistent with the sOT formulation (3.4), after dividing  $\epsilon$  over all terms of Eqn. (3.4).

In this case, the optimal solution  $\mathbf{P}$  has the decomposition

$$(3.13) \quad \mathbf{P} = \text{diag}(\mathbf{u})\mathbf{K}\text{diag}(\mathbf{v}),$$

which is a diagonal scaling of the initial Gibbs kernel  $\mathbf{K}$ , the same as the optimal solution for the regularized OT problem (which corresponds to  $h_1(x) = \iota_{\{x=\mathbf{a}\}}(x)$  and  $h_2(x) = \iota_{\{x=\mathbf{b}\}}(x)$  in (3.12)). Indeed, this decomposition (3.13) not only holds for

the optimal  $\mathbf{P}$ , but also holds for each iterate  $\mathbf{P}^n$  generated by Dykstra's algorithm for KL divergence. Therefore we assume that  $\mathbf{P}^n = \text{diag}(\mathbf{u}^n)\mathbf{K}\text{diag}(\mathbf{v}^n)$ . Then the Dykstra algorithm can be written in an implementable form given as follows [11].

Dykstra algorithm for KL divergence (implementable form).

Input:  $\mathbf{u}^0 = \mathbf{v}^0 = \mathbf{1}$ ;

General step: for any  $k = 0, 1, 2, \dots$ , execute the following steps:

$$(3.14) \quad \mathbf{u}^{2k+1} = \frac{\text{prox}_{h_1}^{\text{KL}}(\mathbf{K}\mathbf{v}^{2k})}{\mathbf{K}\mathbf{v}^{2k}}, \quad \mathbf{v}^{2k+1} = \mathbf{v}^{2k};$$

$$(3.15) \quad \mathbf{v}^{2k+2} = \frac{\text{prox}_{h_2}^{\text{KL}}(\mathbf{K}^T \mathbf{u}^{2k+1})}{\mathbf{K}^T \mathbf{u}^{2k+1}}, \quad \mathbf{u}^{2k+2} = \mathbf{u}^{2k+1}.$$

Note that in some of the literature, this implementable form of the Dykstra algorithm for KL divergence is called the *generalized Sinkhorn iteration* for problem (3.12).

**3.2. Dykstra algorithm for entropy regularized sOT problem.** The entropy regularized sOT problem (3.4) is a special case of the KL divergence problem (3.12) by taking

$$(3.16) \quad h_1(\mathbf{P}\mathbf{1}) = \frac{\gamma}{\epsilon} \|\mathbf{a} - \mathbf{P}\mathbf{1}\|_1 + \iota_{[0, \mathbf{a}]}(\mathbf{P}\mathbf{1}), \quad h_2(\mathbf{P}^T \mathbf{1}) = \frac{\gamma}{\epsilon} \|\mathbf{b} - \mathbf{P}^T \mathbf{1}\|_1 + \iota_{[0, \mathbf{b}]}(\mathbf{P}^T \mathbf{1}).$$

Indeed,  $h_1$  (and  $h_2$ , respectively) can be viewed as a regularization term to render  $\boldsymbol{\mu} = \mathbf{a} - \mathbf{P}\mathbf{1}$  ( $\boldsymbol{\nu} = \mathbf{b} - \mathbf{P}^T \mathbf{1}$ , respectively) as small as possible but within the range  $[0, \mathbf{a}]$  (and  $[0, \mathbf{b}]$ , respectively). In this case, the implementation of the Dykstra algorithm depends on the form of the proximal operator of  $h_i, i = 1, 2$  with respect to the KL divergence, which is given in the following lemma.

**LEMMA 3.1.** *Let  $h_i(\cdot) = \frac{\gamma}{\epsilon} \|\mathbf{a}_i - \cdot\|_1 + \iota_{[0, \mathbf{a}_i]}(\cdot), i = 1, 2$ , with  $\mathbf{a}_1 = \mathbf{a}, \mathbf{a}_2 = \mathbf{b}$ ; then the proximal operator of  $h_i$  with respect to the KL divergence is given as*

$$(3.17) \quad \text{prox}_{h_i}^{\text{KL}}(\mathbf{q}) = \min\{e^{\gamma/\epsilon} \mathbf{q}, \mathbf{a}_i\}, \quad i = 1, 2.$$

*Proof.* By definition of the proximal operator with respect to the KL divergence, we have

$$\text{prox}_{(\gamma/\epsilon)\|\mathbf{a}_i - \cdot\|_1 + \iota_{[0, \mathbf{a}_i]}(\cdot)}^{\text{KL}}(\mathbf{q}) = \underset{\mathbf{p} \in [0, \mathbf{a}_i]}{\text{argmin}} \text{KL}(\mathbf{p}|\mathbf{q}) + (\gamma/\epsilon)\|\mathbf{a}_i - \mathbf{p}\|.$$

If  $\mathbf{a}_i \leq \mathbf{q}$ , both  $\text{KL}(\cdot|\mathbf{q})$  and  $\|\mathbf{a}_i - \cdot\|$  decrease over domain  $[0, \mathbf{a}_i]$ , so the minimum is attained at  $\mathbf{p} = \mathbf{a}_i$ ; if  $\mathbf{a}_i \geq \mathbf{q}$ , taking the derivative of  $\text{KL}(\mathbf{p}|\mathbf{q}) + (\gamma/\epsilon)\|\mathbf{a}_i - \mathbf{p}\|$  with respect to  $\mathbf{p}$  and setting it to zero, we find  $\mathbf{p} = \min\{e^{\gamma/\epsilon} \mathbf{q}, \mathbf{a}_i\}$ . Combining the two cases yields the result.  $\square$

Inserting Lemma 3.1 into the Dykstra algorithm for KL divergence, we obtain the generalized Sinkhorn iteration for the entropy regularized sOT problem with the regularization terms in (3.16).

Generalized Sinkhorn algorithm for sOT

Input:  $\mathbf{u}^0 = \mathbf{v}^0 = \mathbf{1}$ ;

General step: for any  $k = 0, 1, 2, \dots$ , execute the following steps:

$$(3.18) \quad \mathbf{u}^{2k+1} = \frac{\min\{e^{\gamma/\epsilon} \mathbf{K} \mathbf{v}^{2k}, \mathbf{a}\}}{\mathbf{K} \mathbf{v}^{2k}} = \min\left\{e^{\frac{\gamma}{\epsilon}} \mathbf{1}, \frac{\mathbf{a}}{\mathbf{K} \mathbf{v}^{2k}}\right\}, \quad \mathbf{v}^{2k+1} = \mathbf{v}^{2k},$$

$$(3.19) \quad \mathbf{v}^{2k+2} = \frac{\min\{e^{\gamma/\epsilon} \mathbf{K}^T \mathbf{u}^{2k+1}, \mathbf{b}\}}{\mathbf{K}^T \mathbf{u}^{2k+1}} = \min\left\{e^{\frac{\gamma}{\epsilon}} \mathbf{1}, \frac{\mathbf{b}}{\mathbf{K}^T \mathbf{u}^{2k+1}}\right\}, \quad \mathbf{u}^{2k+2} = \mathbf{u}^{2k+1}.$$

**4. sOT barycenter.** Given a set  $\{\mathbf{b}_j\}_{j=1}^J$  of unbalanced marginal densities  $\mathbf{b}_j \in \mathbb{R}_+^m$  and a weight  $\lambda = (\lambda_1, \dots, \lambda_J) \in \text{int}(\Delta_J)$ , it is of practical interest to compute the weighted sOT barycenter of  $\{\mathbf{b}_j\}_{j=1}^J$ . This problem can be viewed as the generalization of the standard Wasserstein barycenter problem studied in.

We define the sOT barycenter problem in a similar manner as that for sOT. Let  $\mathbf{v} = (\mathbf{v}_j)_{j=1}^J \in (\mathbb{R}_+^m)^J$  denote the blocked marginal measure, and  $\|\mathbf{v}\|_1 = \sum_j \|\mathbf{v}_j\|_1$ . We define two sets, one of which is for the feasible blocked marginal density  $\mathbf{v}$ , and the other is for the feasible  $\mathbf{v}$  with minimal 1-norm:

$$\begin{aligned} \mathcal{F} &= \{\mathbf{v} : \|\mathbf{b}_1 - \mathbf{v}_1\|_1 = \dots = \|\mathbf{b}_J - \mathbf{v}_J\|_1, \text{ and } \exists \mathbf{P}_j \in \mathbf{U}(\mathbf{a}, \mathbf{b}_j - \mathbf{v}_j), \\ &\quad \text{such that } \langle \mathbf{P}_j, \mathbf{C} \rangle < \infty \text{ for some } \mathbf{a}\}, \\ \mathcal{G} &= \underset{\mathbf{v} \in \mathcal{F}}{\operatorname{argmin}} \|\mathbf{v}\|_1. \end{aligned}$$

Then the sOT barycenter problem is defined as

$$(4.1) \quad \min_{\mathbf{v} \in \mathcal{G}} \min_{\mathbf{P}_j \in \mathbf{U}(\mathbf{a}, \mathbf{b}_j - \mathbf{v}_j)} \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j, \mathbf{C} \rangle.$$

Similarly to the equivalence between various sOT formulations, we can show that (4.1) is equivalent to

$$(4.2) \quad \min_{\substack{\mathbf{v} \in \mathcal{F} \\ \mathbf{P}_j \in \mathbf{U}(\mathbf{a}, \mathbf{b}_j - \mathbf{v}_j)}} \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j, \mathbf{C} \rangle + \frac{\gamma}{J} \|\mathbf{v}\|_1$$

for sufficiently large  $\gamma$ . The equivalence between (4.1) and (4.2) is summarized in Theorem 4.1.

**THEOREM 4.1.** *Given a cost matrix  $\mathbf{C}$  with  $\infty$ -pattern  $\mathcal{P}_\infty(\mathbf{C})$ , the two formulations for the sOT barycenter problems (4.1) and (4.2) are equivalent for sufficiently large  $\gamma$ .*

*Proof.* The proof is similar to that of Lemma 2.3. Starting from an optimal  $\mathbf{v}^{\text{opt}} = (\mathbf{v}_j^{\text{opt}})_j$  for the formulation (4.1) and a corresponding optimal plan  $\mathbf{P}^{\bar{\theta},*} = (\mathbf{P}_j^{\bar{\theta},*})_j$  in which  $\bar{\theta} = \|\mathbf{b}_1 - \mathbf{v}_1^{\text{opt}}\|_1 = \dots = \|\mathbf{b}_J - \mathbf{v}_J^{\text{opt}}\|_1$ , and taking any nonnegative and feasible  $\theta < \bar{\theta}$ , if we can prove that for any plan  $\mathbf{P}^{\theta,*}$  defined as

$$(4.3) \quad \begin{aligned} \mathbf{P}^{\theta,*} &= \underset{\mathbf{P}}{\operatorname{argmin}} \left\{ \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j, \mathbf{C} \rangle : \mathbf{P}_j \in \mathbf{U}(=\mathbf{a}, \leq \mathbf{b}_j), \right. \\ &\quad \left. \mathbf{P}_1 \mathbf{1} = \dots = \mathbf{P}_J \mathbf{1} = \mathbf{a}, \langle \mathbf{P}_j, \mathbf{1} \rangle = \theta, \forall j \right\}, \end{aligned}$$

one has that

$$(4.4) \quad \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j^{\bar{\theta},*} - \mathbf{P}_j^{\theta,*}, \mathbf{C} \rangle \leq \gamma \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j^{\bar{\theta},*} - \mathbf{P}_j^{\theta,*}, \mathbf{1} \rangle \quad \text{for some } \gamma > 0,$$

then for any  $\mathbf{v} = (\mathbf{v}_j)_j$  such that  $\theta = \|\mathbf{b}_1 - \mathbf{v}_1\|_1 = \cdots = \|\mathbf{b}_J - \mathbf{v}_J\|_1$ , it implies that

$$\begin{aligned} \|\mathbf{v}\|_1 - \|\mathbf{v}^{\text{opt}}\|_1 &= \langle \mathbf{v} - \mathbf{v}^{\text{opt}}, \mathbf{1} \rangle \\ &= J \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j^{\bar{\theta},*} - \mathbf{P}_j^{\theta,*}, \mathbf{1} \rangle \\ &\geq J\gamma^{-1} \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j^{\bar{\theta},*} - \mathbf{P}_j^{\theta,*}, \mathbf{C} \rangle, \end{aligned}$$

leading to

$$J \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j^{\bar{\theta},*}, \mathbf{C} \rangle + \gamma \|\mathbf{v}^{\text{opt}}\|_1 \leq J \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j^{\theta,*}, \mathbf{C} \rangle + \gamma \|\mathbf{v}\|_1,$$

which implies the optimality of  $(\mathbf{v}^{\text{opt}}, \mathbf{P}^{\bar{\theta},*})$  for (4.2).

Now we prove that there exists a  $\gamma > 0$  such that (4.4) holds. Using Lemma 2.2 and taking the bounded convex polyhedron  $C$  to be the set of  $\mathbf{P} = (\mathbf{P}_j)_j$  defined by the constraints

$$\begin{cases} \mathbf{P}_j^T \mathbf{1} \leq \mathbf{b}_j, \mathbf{P}_j \geq 0, & j = 1 : J, \\ \mathbf{P}_j \mathbf{1} = \mathbf{P}_{j+1} \mathbf{1}, & j = 1 : J-1, \\ \mathbf{P}_j \geq 0, & j = 1 : J, \\ (\mathbf{P}_j)_{kl} = 0, & (k, l) \in \mathcal{P}_\infty(\mathbf{C}), \quad j = 1 : J, \end{cases}$$

and

$$C_\theta = \left\{ \mathbf{P} \in C : \langle \mathbf{P}, \mathbf{1} \rangle := \sum_{j=1}^J \langle \mathbf{P}_j, \mathbf{1} \rangle = J\theta \right\},$$

we know that there exists a critical  $\theta_0$  such that for any given  $\theta \in (\theta_0, \bar{\theta}]$  and  $\mathbf{P}^{\theta,*}$  defined in (4.3), we can find a feasible  $\mathbf{P}^{\bar{\theta}}$  satisfying  $\langle \mathbf{P}^{\bar{\theta}}, \mathbf{1} \rangle = J\bar{\theta}$  such that

$$(4.5) \quad \|\mathbf{P}^{\bar{\theta}} - \mathbf{P}^{\theta,*}\|_1 \leq J\eta|\bar{\theta} - \theta|.$$

Then

$$\begin{aligned} \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j^{\bar{\theta},*} - \mathbf{P}_j^{\theta,*}, \mathbf{C} \rangle &\leq \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j^{\bar{\theta}} - \mathbf{P}_j^{\theta,*}, \mathbf{C} \rangle \\ &\leq \frac{1}{J} \|\mathbf{P}^{\bar{\theta}} - \mathbf{P}^{\theta,*}\|_1 \cdot \|\mathbf{C}\|_\infty \\ &\leq \|\mathbf{C}\|_\infty \cdot \eta |\bar{\theta} - \theta| \\ &= \|\mathbf{C}\|_\infty \cdot \eta \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j^{\bar{\theta},*} - \mathbf{P}_j^{\theta,*}, \mathbf{1} \rangle. \end{aligned} \quad (4.6)$$



On the other hand, for any feasible  $\theta \leq \theta_0$  and  $\mathbf{P}^{\theta,*}$  defined in (4.3), we simply have

$$\begin{aligned}
 \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j^{\bar{\theta},*} - \mathbf{P}_j^{\theta,*}, \mathbf{C} \rangle &= \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j^{\bar{\theta},*}, \mathbf{C} \rangle - \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j^{\theta,*}, \mathbf{C} \rangle \\
 &\leq \|\mathbf{C}\|_{\infty} \cdot \bar{\theta} \\
 &= \|\mathbf{C}\|_{\infty} \cdot \frac{\bar{\theta}}{\bar{\theta} - \theta} (\bar{\theta} - \theta) \\
 &\leq \|\mathbf{C}\|_{\infty} \cdot \frac{\bar{\theta}}{\bar{\theta} - \theta_0} (\bar{\theta} - \theta) \\
 (4.7) \quad &= \|\mathbf{C}\|_{\infty} \cdot \frac{\bar{\theta}}{\bar{\theta} - \theta_0} \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j^{\bar{\theta},*} - \mathbf{P}_j^{\theta,*}, \mathbf{1} \rangle.
 \end{aligned}$$

Finally, combining the inequalities (4.6) and (4.7) and taking  $\gamma = \max\{\eta, \bar{\theta}/(\bar{\theta} - \theta_0)\} \cdot \|\mathbf{C}\|_{\infty}$ , we prove the inequality (4.4), and therefore the equivalence between two formulations.  $\square$

Noting that the feasible set  $\mathcal{F}$  is always nonempty, the formulation (4.2) can be recast into the form

$$(4.8) \quad \min_{\substack{(\mathbf{P}_j, \mathbf{a}) \\ \mathbf{P}_j \in \mathbf{U}(\mathbf{a}, \leq \mathbf{b}_j)}} \sum_{j=1}^J \lambda_j \langle \mathbf{P}_j, \mathbf{C} \rangle + \gamma \sum_{j=1}^J \|\mathbf{b}_j - \mathbf{P}_j^T \mathbf{1}\|_1.$$

Here we replace  $\frac{\gamma}{J}$  by  $\gamma$  in the equivalent formulation for the sake of simple notation.

**4.1. Entropic regularization of the sOT barycenter problem.** In this section, we consider the entropic regularization for the weighted sOT barycenter problem. To this end, we introduce the following notation:

$$\begin{aligned}
 (4.9) \quad \text{KL}_{\lambda}(\mathbf{P}|\mathbf{Q}) &:= \sum_{j=1}^J \lambda_j \text{KL}(\mathbf{P}_j|\mathbf{Q}_j), \\
 \text{where } \mathbf{P} = (\mathbf{P}_j)_j &\in (\mathbb{R}_+^{n \times m})^J, \mathbf{Q} = (\mathbf{Q}_j)_j \in (\mathbb{R}_{++}^{n \times m})^J,
 \end{aligned}$$

$$(4.10)$$

$$\hat{h}_1(\mathbf{P}) = \iota_{\mathcal{D}}(\mathbf{P}_1 \mathbf{1}, \dots, \mathbf{P}_J \mathbf{1}), \quad \mathcal{D} := \{(\mathbf{p}_1, \dots, \mathbf{p}_J) \in (\mathbb{R}_+^m)^J : \mathbf{p}_1 = \dots = \mathbf{p}_J\},$$

$$(4.11)$$

$$\hat{h}_2(\mathbf{Q}) = \frac{\gamma}{\epsilon} \|\mathbf{b} - \mathbf{Q}^T \mathbf{1}\|_1 + \iota_{[0, \mathbf{b}]}(\mathbf{Q}^T \mathbf{1}) := \sum_{j=1}^J \left( \frac{\gamma}{\epsilon} \|\mathbf{b}_j - \mathbf{Q}_j^T \mathbf{1}\|_1 + \iota_{[0, \mathbf{b}_j]}(\mathbf{Q}_j^T \mathbf{1}) \right),$$

$$(4.12)$$

$$h_1(\mathbf{p}) = \iota_{\mathcal{D}}(\mathbf{p}_1, \dots, \mathbf{p}_J) \quad \text{for } \mathbf{p} = (\mathbf{p}_j)_j,$$

$$(4.13)$$

$$h_2(\mathbf{q}) = \frac{\gamma}{\epsilon} \|\mathbf{b} - \mathbf{q}\|_1 + \iota_{[0, \mathbf{b}]}(\mathbf{q}) = \sum_{j=1}^J \left( \frac{\gamma}{\epsilon} \|\mathbf{b}_j - \mathbf{q}_j\|_1 + \iota_{[0, \mathbf{b}_j]}(\mathbf{q}_j) \right) \quad \text{for } \mathbf{q} = (\mathbf{q}_j)_{j=1}^J.$$

Note that  $\hat{h}_i(\mathbf{P}) = h_i(\mathbf{P} \mathbf{1})$ ,  $i = 1, 2$ . With the above notation, we can formulate the entropic regularized sOT barycenter problem as

$$(4.14) \quad \min_{\mathbf{P} \in (\mathbb{R}_+^{n \times m})^J} \sum_{j=1}^J \lambda_j \left( \langle \mathbf{P}_j, \mathbf{C} \rangle - \epsilon H(\mathbf{P}_j) \right) + \epsilon \hat{h}_1(\mathbf{P}) + \epsilon \hat{h}_2(\mathbf{P}),$$

or equivalently in terms of the KL divergence,

$$(4.15) \quad \min_{\mathbf{P} \in (\mathbb{R}_+^{n \times m})^J} \text{KL}_\lambda(\mathbf{P}|\mathbf{K}) + \hat{h}_1(\mathbf{P}) + \hat{h}_2(\mathbf{P}),$$

where  $\mathbf{K} = (\mathbf{K}_j)_{j=1}^J$  with  $\mathbf{K}_j = e^{-\mathbf{C}/\epsilon}$ ,  $j = 1, \dots, J$ .

To solve the entropic regularized sOT barycenter problem (4.15), we adopt the generic diagonal scaling algorithm introduced in [31] (also see [11]), in which each iterate  $\mathbf{P}$  has the diagonal scaling decomposition

$$(4.16) \quad \mathbf{P}^{(n)} = (\mathbf{P}_j^{(n)})_j = \left( \text{diag}(\mathbf{u}_j^{(n)}) \mathbf{K} \text{diag}(\mathbf{v}_j^{(n)}) \right)_j.$$

With a slight abuse of notation, we denote, consistent with  $\mathbf{P} = (\mathbf{P}_j)_j$ ,

$$\mathbf{u} = (\mathbf{u}_1, \dots, \mathbf{u}_J) \in (\mathbb{R}^n)^J, \quad \mathbf{v} = (\mathbf{v}_1, \dots, \mathbf{v}_J) \in (\mathbb{R}^m)^J.$$

Then the diagonal scaling algorithm reads as follows.

Input:  $\mathbf{u}^{(0)} = \mathbf{v}^{(0)} = \mathbf{1}$ ;

General step: for any  $n = 0, 1, 2, \dots$ , execute the following steps:

$$(4.17) \quad \mathbf{u}_j^{(2n+1)} = \frac{\left[ \text{prox}_{h_1}^{\text{KL}_\lambda}(\mathbf{K} \mathbf{v}_j^{(2n)}) \right]_j}{\mathbf{K} \mathbf{v}_j^{(2n)}}, \quad \mathbf{v}_j^{(2n+1)} = \mathbf{v}_j^{(2n)}, \quad \forall j;$$

$$(4.18) \quad \mathbf{v}_j^{(2n+2)} = \frac{\left[ \text{prox}_{h_2}^{\text{KL}_\lambda}(\mathbf{K}^T \mathbf{u}_j^{(2n)}) \right]_j}{\mathbf{K}^T \mathbf{u}_j^{(2n)}}, \quad \mathbf{u}_j^{(2n+2)} = \mathbf{u}_j^{(2n+1)}, \quad \forall j.$$

Note that one needs to compute the two proximal operators  $\text{prox}_{h_1}^{\text{KL}_\lambda}$  and  $\text{prox}_{h_2}^{\text{KL}_\lambda}$  for  $h_1$  and  $h_2$  defined in (4.12)–(4.13) to implement the diagonal scaling algorithm. The following lemma shows that the two proximal operators for the KL divergence can be computed in closed form. The derivation is similar to that of Proposition 5.1 in [31], so we omit the details.

LEMMA 4.2. For any  $\mathbf{p} = (\mathbf{p}_j)_j \in (\mathbb{R}^n)^J$ , and  $h_1$  and  $h_2$  defined in (4.12)–(4.13), one has

$$(4.19) \quad \left[ \text{prox}_{h_1}^{\text{KL}_\lambda}(\mathbf{p}) \right]_j = \mathbf{p}_1^{\lambda_1} \odot \dots \odot \mathbf{p}_J^{\lambda_J}, \quad \left[ \text{prox}_{h_2}^{\text{KL}_\lambda}(\mathbf{p}) \right]_j = \min \left\{ e^{\frac{\gamma}{\lambda_j \epsilon}} \mathbf{p}_j, \mathbf{b}_j \right\}.$$

Additionally, for any  $\mathbf{P} = (\mathbf{P}_j)_j \in (\mathbb{R}^{n \times m})^J$ , and  $\hat{h}_1$  and  $\hat{h}_2$  defined in (4.10)–(4.11), the two proximal operators  $\text{prox}_{h_1}^{\text{KL}_\lambda}$  and  $\text{prox}_{h_2}^{\text{KL}_\lambda}$  are related to (4.19) as

$$(4.20) \quad \begin{aligned} \left[ \text{prox}_{h_1}^{\text{KL}_\lambda}(\mathbf{P}) \right]_j &= \text{diag} \left( \frac{\left[ \text{prox}_{h_1}^{\text{KL}_\lambda}(\mathbf{P} \mathbf{1}) \right]_j}{\mathbf{P}_j \mathbf{1}} \right) \mathbf{P}_j, \\ \left[ \text{prox}_{h_2}^{\text{KL}_\lambda}(\mathbf{P}) \right]_j &= \mathbf{P}_j \text{diag} \left( \frac{\left[ \text{prox}_{h_2}^{\text{KL}_\lambda}(\mathbf{P}^T \mathbf{1}) \right]_j}{\mathbf{P}_j^T \mathbf{1}} \right), \end{aligned}$$

in which

$$\mathbf{P}\mathbf{1} := (\mathbf{P}_1\mathbf{1}, \dots, \mathbf{P}_J\mathbf{1}), \quad \mathbf{P}^T\mathbf{1} := (\mathbf{P}_1^T\mathbf{1}, \dots, \mathbf{P}_J^T\mathbf{1}).$$

With the proximal operators computed in (4.19), the diagonal scaling algorithm in a directly implementable form becomes as follows.

Input:  $\mathbf{u}^{(0)} = \mathbf{v}^{(0)} = \mathbf{1}$ ;

General step: for any  $n = 0, 1, 2, \dots$ , execute the following steps:

$$(4.21) \quad \mathbf{u}_j^{(n+1)} = \frac{\mathbf{a}^{(n)}}{\mathbf{K}\mathbf{v}_j^{(n)}}, \quad j = 1 : J, \quad \text{where } \mathbf{a}^{(n)} = \prod_j \left( \mathbf{K}\mathbf{v}_j^{(n)} \right)^{\lambda_j},$$

$$(4.22) \quad \mathbf{v}_j^{(n+1)} = \min \left\{ \frac{\mathbf{b}_j}{\mathbf{K}^T \mathbf{u}_j^{(n)}}, e^{\frac{\gamma}{\lambda_j \epsilon}} \right\}, \quad j = 1 : J.$$

**4.2. Log-domain implementation.** One drawback for the diagonal scaling algorithm (Sinkhorn algorithm) is that it suffers from numerical overflow when the regularization parameter  $\epsilon$  is too small compared to the entries of the cost matrix  $\mathbf{C}$ . This drawback is even more severe for the sOT problem as it will cause some entries of  $\mathbf{K} = e^{-\mathbf{C}/\epsilon}$  to be regarded as zero due to the numerical overflow, even though they should not. In other words, having more zero entries in  $\mathbf{K}$  because of the smallness of  $\epsilon$  will change the 0-pattern of  $\mathbf{K}$  and consequently the  $\infty$ -pattern of  $\mathbf{C}$ . Therefore it is necessary to implement the diagonal scaling algorithm for the sOT barycenter problem in the log-domain.

Using the log-sum-exp stabilization trick for the soft-minimization, and noting the primal-dual relation

$$(\mathbf{u}^{(n)}, \mathbf{v}^{(n)}) = \left( e^{\mathbf{f}^{(n)}/\epsilon}, e^{\mathbf{g}^{(n)}/\epsilon} \right),$$

the log-domain implementation for the diagonal scaling algorithm reads as follows.

Input:  $\mathbf{f}^{(0)} = \mathbf{g}^{(0)} = 0$ ;

General step: for any  $n = 0, 1, 2, \dots$  execute the following steps:

$$(4.23) \quad \begin{aligned} \mathbf{f}_j^{(n+1)} = & \sum_i \lambda_i \left[ \epsilon \log \left( e^{(\mathbf{f}_i^{(n)} \oplus \mathbf{g}_i^{(n)} - \mathbf{C})/\epsilon} \mathbf{1} \right) - \mathbf{f}_i^{(n)} \right] \\ & - \left[ \epsilon \log \left( e^{(\mathbf{f}_j^{(n)} \oplus \mathbf{g}_j^{(n)} - \mathbf{C})/\epsilon} \mathbf{1} \right) - \mathbf{f}_j^{(n)} \right], \end{aligned}$$

$$(4.24) \quad \mathbf{g}_j^{(n+1)} = \min \left\{ \epsilon \log(\mathbf{b}_j) - \epsilon \log \left( e^{(\mathbf{f}_j^{(n+1)} \oplus \mathbf{g}_j^{(n)} - \mathbf{C})^T/\epsilon} \mathbf{1} \right) + \mathbf{g}_j^{(n)}, \frac{\gamma}{\lambda_j} \right\}$$

for all  $j = 1 : J$ .

**4.3. Special case in which  $\lambda \in \partial(\Delta_J)$ .** In this subsection, we point out an important difference between the standard OT barycenter problem and the sOT one. For the sake of simplicity, we take  $J = 2$ .

Note that when  $\lambda = (0, 1)$ , the OT barycenter problem degenerates to the standard OT problem. More precisely,

$$\min_{\mathbf{a}} \lambda_1 L_{\text{OT}}(\mathbf{a}, \mathbf{b}_1) + \lambda_2 L_{\text{OT}}(\mathbf{a}, \mathbf{b}_2)$$

reduces to

$$\min_{\mathbf{a}} L_{\text{OT}}(\mathbf{a}, \mathbf{b}_2),$$

which leads to  $\mathbf{a} = \mathbf{b}_2$  and  $\mathbf{P}_2 = \text{diag}(\mathbf{b}_2)$ . Then  $\mathbf{P}_1$  is determined by the standard OT

$$L_{\text{OT}}(\mathbf{a}, \mathbf{b}_1) = \min_{\mathbf{P}_1 \in \mathbf{U}(\mathbf{a}, \mathbf{b}_1)} \langle \mathbf{P}_1, \mathbf{C} \rangle.$$

Additionally, the entropic OT barycenter problem reduces to the entropic OT problem. In other words,

$$\min_{\mathbf{a}} \lambda_1 L_{\text{OT}}^\epsilon(\mathbf{a}, \mathbf{b}_1) + \lambda_2 L_{\text{OT}}^\epsilon(\mathbf{a}, \mathbf{b}_2)$$

reduces to

$$\min_{\mathbf{a}} L_{\text{OT}}^\epsilon(\mathbf{a}, \mathbf{b}_2),$$

which leads to  $\mathbf{a} = \frac{\mathbf{K}\mathbf{b}_2}{\mathbf{K}^T \mathbf{1}}$  and  $\mathbf{P}_2 = \mathbf{K} \text{diag} \left( \frac{\mathbf{b}_2}{\mathbf{K}^T \mathbf{1}} \right)$ . Then  $\mathbf{P}_1$  is determined by the entropic OT

$$L_{\text{OT}}^\epsilon(\mathbf{a}, \mathbf{b}_1) = \min_{\mathbf{P}_1 \in \mathbf{U}(\mathbf{a}, \mathbf{b}_1)} \epsilon \text{KL}(\mathbf{P}_1 | \mathbf{K}).$$

However, such degeneration does not apply to the sOT barycenter problem,

$$\lim_{\lambda_1 \rightarrow 0} \left( \min_{\mathbf{a}} \lambda_1 L_{\text{sOT}}(\mathbf{a}, \mathbf{b}_1) + \lambda_2 L_{\text{sOT}}(\mathbf{a}, \mathbf{b}_2) \right) \neq \min_{\mathbf{a}} L_{\text{sOT}}(\mathbf{a}, \mathbf{b}_2).$$

To elucidate the idea, we take the cost matrix  $\mathbf{C}$  as in (5.2) with  $C_{\text{cut}} = 0.3$ , namely, any mass can only be transported within the distance no longer than  $C_{\text{cut}}$ . We take  $\mathbf{y} \in \mathbb{R}^{n+1}$  to be a uniform mesh over  $[0, 1]$  with  $h = 1/n$  being the mesh spacing and let  $\mathbf{b}_1 = \mathbf{1}$  (and  $\mathbf{b}_2 = \mathbf{1}$ , respectively) be uniform distribution on  $\mathbf{y}$  compactly supported over  $[0.1, 0.3]$  (and  $[0.7, 0.9]$ , respectively). For any  $\lambda = (\lambda_1, \lambda_2) \in \text{int}(\Delta_2)$ , the sOT barycenter  $\mathbf{a}$  must be the uniform distribution  $\mathbf{a} = \mathbf{1}$  on  $\mathbf{y}$  compactly supported over  $[0.4, 0.6]$ . The corresponding total cost is

$$\lambda_1 (0.3)^2 + \lambda_2 (0.3)^2 = (0.3)^2.$$

This is because any other possible transport plan will cause some mass, even only a bit, to be transported from either  $\mathbf{b}_1$  or  $\mathbf{b}_2$  to anywhere beyond  $[0.4, 0.6]$ , resulting in an infinite cost. On the other hand, when  $\lambda = (0, 1) \in \partial(\Delta_2)$ , the sOT barycenter is determined by

$$\min_{\mathbf{a}} L_{\text{sOT}}(\mathbf{a}, \mathbf{b}_2).$$

Since it is unrelated to  $L_{\text{sOT}}(\mathbf{a}, \mathbf{b}_1)$ , we can take  $\mathbf{a} = \mathbf{b}_2$  such that the total cost equals zero. Hence the degeneration leads to a discontinuity for the total cost. To

resolve this issue, we define the degenerate sOT problem for  $\tilde{\lambda} \in \partial(\Delta_J)$  as the limiting problem when  $\text{int}(\Delta_J) \ni \lambda \rightarrow \tilde{\lambda}$ ,

$$\min_{\mathbf{a}} \sum_j \tilde{\lambda}_j L_{\text{sOT}}(\mathbf{a}, \mathbf{b}_j) := \lim_{\text{int}(\Delta_J) \ni \lambda \rightarrow \tilde{\lambda}} \left( \min_{\mathbf{a}} \sum_j \lambda_j L_{\text{sOT}}(\mathbf{a}, \mathbf{b}_j) \right).$$

**4.4. Reverse and portion selection mechanism.** We can further perform theoretical analysis on the limiting behavior of the barycenter sOT problem. For the sake of simplicity, we still take  $J = 2$ . Let  $\mathbf{y} = (y_i)_i \in \mathbb{R}^n$  be the uniform mesh over  $[0, 1 - h]$  with  $h = 1/n$ . For the rest of this section, we take  $n = 1000$ .

To begin with, we define a cumulative sum inequality as follows. For any two vectors  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ , we say  $\mathbf{u}$  is cumulatively less than or equal to  $\mathbf{v}$  and denote by  $\mathbf{u} \leq_{\text{C}} \mathbf{v}$  if they satisfy

$$(4.25) \quad \sum_{j=1}^k u_j \leq \sum_{j=1}^k v_j, k = 1 : n-1; \quad \text{and} \quad \sum_{j=1}^n u_j = \sum_{j=1}^n v_j.$$

$\mathbf{u}$  is strictly cumulatively less than  $\mathbf{v}$  and we denote  $\mathbf{u} <_{\text{C}} \mathbf{v}$  if at least one inequality in (4.25) is strict for  $k = 1 : n-1$ . We denote  $\mathbf{u} =_{\text{C}} \mathbf{v}$  if all the inequalities in (4.25) are equality. It is clear that  $\mathbf{u} =_{\text{C}} \mathbf{v}$  if and only if  $\mathbf{u} = \mathbf{v}$ .

Let  $T^t$  be a periodic shift operator for any periodic function  $f(x)$  over  $[0, 1]$  such that  $(T^t f)(x) = f(x-t)$ . We take two nonnegative periodic functions  $b_1(x), b_2(x)$  over  $[0, 1]$  with compact support  $[0, 0.2]$  and let them satisfy the cumulative sum inequality

$$(4.26) \quad b_1|_{\mathbf{y} \cap [0, 0.2]} \leq_{\text{C}} b_2|_{\mathbf{y} \cap [0, 0.2]}.$$

We define two marginal distributions  $\mathbf{b}_1$  and  $\mathbf{b}_2$  as

$$(4.27) \quad \mathbf{b}_1 = (T^{0.1} b_1(x))|_{\mathbf{y}}, \quad \mathbf{b}_2 = (T^{0.7} b_2(x))|_{\mathbf{y}}.$$

The cost matrix  $\mathbf{C}$  is taken as

$$(4.28) \quad \mathbf{C}_{ij} = \begin{cases} |y_i - y_j|^2 & \text{if } |y_i - y_j| \leq 0.3, \\ \infty & \text{if } |y_i - y_j| > 0.3. \end{cases}$$

In the following lemma, we characterize the marginal distributions which are of the same amount of mass as  $\mathbf{b}_1$  and  $\mathbf{b}_2$  and can be completely transported to  $\mathbf{b}_1$  and  $\mathbf{b}_2$  given the cost matrix  $\mathbf{C}$  in (4.28).

**LEMMA 4.3.** *Given marginal distributions  $\mathbf{b}_1$  and  $\mathbf{b}_2$  as in (4.27), a marginal distribution  $\mathbf{a}$  can be completely transported to both  $\mathbf{b}_1$  and  $\mathbf{b}_2$  if and only if*

$$(4.29) \quad \mathbf{a} = (T^{0.4} a(x))|_{\mathbf{y}}$$

*in which  $a(x)$  is a nonnegative periodic function over  $[0, 1]$  with compact support  $[0, 0.2]$  and satisfies*

$$(4.30) \quad b_1(x)|_{\mathbf{y} \cap [0, 0.2]} \leq_{\text{C}} a(x)|_{\mathbf{y} \cap [0, 0.2]} \leq_{\text{C}} b_2(x)|_{\mathbf{y} \cap [0, 0.2]}.$$

*Proof.* We only consider the case for  $\mathbf{P}_2$ . That for  $\mathbf{P}_1$  is similar.

First, we show that if the cumulative sum inequality (4.30) holds, then there exists some plan  $\mathbf{P}_2$  to transport  $\mathbf{a}$  completely to  $\mathbf{b}_2$ . Here we use the northwest corner rule [32] to construct such a plan  $\mathbf{P}_2$ . If  $\mathbf{P}_2$  completely transports  $\mathbf{a}$  to  $\mathbf{b}_2$ ,

then all the entries not in the submatrix  $\mathbf{Q} = \mathbf{P}_2(401 : 600, 701 : 900) \in \mathbb{R}^{200 \times 200}$  must be 0 due to the compact support of  $\mathbf{a}$  and  $\mathbf{b}_2$ . Besides, since the entries of  $\mathbf{C}$  are  $\infty$  when  $|\mathbf{y}_i - \mathbf{y}_j| > 0.3$ , the strict upper triangular entries of  $\mathbf{Q}$  must also be 0. Now we apply the northwest corner rule to determine the lower triangular (including diagonal) entries of  $\mathbf{Q}$ . More precisely, the rule starts by giving the highest possible value to  $Q_{11} = (\mathbf{P}_2)_{401,701}$  by setting it to  $\min\{(\mathbf{a})_{401}, (\mathbf{b}_2)_{701}\}$ . At each step, the entry  $(\mathbf{P}_2)_{ij}$  is chosen to saturate either the  $i$ th row constraint or the  $j$ th column constraint, or both if possible. The indices  $i, j$  are then updated as follows:  $i$  is incremented in the first case,  $j$  is in the second, and both  $i$  and  $j$  are in the third case. The rule proceeds until  $(\mathbf{P}_2)_{600,900}$  receives a value. On the other hand, using the second half of the cumulative sum inequality (4.30), we have that each diagonal entry  $Q_{kk}$  of  $\mathbf{Q}$  must be chosen to saturate the corresponding row constraint or both the row and column constraints, but it cannot saturate only the corresponding column constraint. Since the total mass of  $\mathbf{a}$  is equal to that of  $\mathbf{b}_2$ , the possible excessive amount of mass due to the row saturation is eventually transported by the last row of  $\mathbf{Q}$ . Hence mass is completely transported.

Second, we show that if the cumulative sum inequality (4.30) is violated, then no plan  $\mathbf{P}_2$  can completely transport  $\mathbf{a}$  to  $\mathbf{b}_2$ . Let  $k$  be the smallest integer to break (4.30),

$$(4.31) \quad \sum_{j=1}^k a(y_j) > \sum_{j=1}^k b_2(y_j).$$

Assume there exists a plan  $\mathbf{P}_2$  which transports  $\mathbf{a}$  completely to  $\mathbf{b}_2$ . We still denote  $\mathbf{Q} = \mathbf{P}_2(401 : 600, 701 : 900)$ . Due to the compact support of  $\mathbf{a}, \mathbf{b}_2$  and the  $\infty$  entries in  $\mathbf{C}$ , the nonzero entries of  $\mathbf{P}_2$  can only lie in the lower triangular half (including the diagonal) of  $\mathbf{Q}$ . Consider the submatrix  $\mathbf{Q}_k = \mathbf{Q}(1 : k, 1 : k)$ . On one hand, since  $\mathbf{a}$  is completely transported, we have

$$\mathbf{Q}_k \mathbf{1} = [\mathbf{a}(y_1), \dots, \mathbf{a}(y_k)]^T, \quad \mathbf{Q}_k^T \mathbf{1} \leq [\mathbf{b}_2(y_1), \dots, \mathbf{b}_2(y_k)]^T.$$

However,  $\sum_i (\mathbf{Q}_k \mathbf{1})_i = \sum_{i,j} (\mathbf{Q}_k)_{ij} = \sum_j (\mathbf{Q}_k^T \mathbf{1})_j$  leads to  $\sum_{j=1}^k a(y_j) \leq \sum_{j=1}^k b_2(y_j)$ , causing a contradiction to inequality (4.31).  $\square$

Then we have the following result regarding the sOT barycenter of  $\mathbf{b}_1$  and  $\mathbf{b}_2$ .

**THEOREM 4.4.** *Let  $\mathbf{b}_1$  and  $\mathbf{b}_2$  be defined as in (4.27) with  $b_1$  and  $b_2$  satisfying (4.26). Given cost matrix  $\mathbf{C}$  defined in (4.28), we have*

$$(4.32) \quad \lim_{(\lambda_1, \lambda_2) \rightarrow (1,0)} \operatorname{argmin}_{\mathbf{a}} \left( \sum_{j=1}^2 \lambda_j L_{\text{sOT}}(\mathbf{a}, \mathbf{b}_j; \mathbf{C}) \right) = (T^{0.4} b_2)|_{\mathbf{y}},$$

$$(4.33) \quad \lim_{(\lambda_1, \lambda_2) \rightarrow (0,1)} \operatorname{argmin}_{\mathbf{a}} \left( \sum_{j=1}^2 \lambda_j L_{\text{sOT}}(\mathbf{a}, \mathbf{b}_j; \mathbf{C}) \right) = (T^{0.4} b_1)|_{\mathbf{y}}.$$

*Proof.* We only prove the case for  $(\lambda_1, \lambda_2) \rightarrow (0, 1)$ . The other case is similar. By virtue of Lemmas 4.1 and 4.3, all the candidates for the sOT barycenter of  $\mathbf{b}_1, \mathbf{b}_2$  are in the form of (4.29) with condition (4.30) satisfied. We denote by  $\mathcal{S}_{\mathbf{a}}$  the collection of all these candidates.

In the limit  $(\lambda_1, \lambda_2) \rightarrow (0, 1)$ ,  $\lambda_1 L_{\text{sOT}}(\mathbf{a}, \mathbf{b}_1)$  becomes zero. We only need to seek an optimal  $\mathbf{a}^*$  in  $\mathcal{S}_{\mathbf{a}}$  to minimize  $L_{\text{sOT}}(\mathbf{a}, \mathbf{b}_2)$ . Pick any  $\mathbf{a} \geq_{\mathbf{C}} \mathbf{b}_1$ , and let  $k$  be the smallest integer such that

$$(4.34) \quad \sum_{j=1}^k a(y_j) > \sum_{j=1}^k b_1(y_j).$$

Since  $\sum_{j=1}^n a(y_j) = \sum_{j=1}^n b_1(y_j)$ , there must be some integer  $l > k$  such that  $a(y_l) < b_1(y_l)$ . Let  $\mathbf{P}_{\mathbf{a}}^*$  be an OT plan for  $L_{\text{sOT}}(\mathbf{a}, \mathbf{b}_2)$ , and let  $(\mathbf{P}_{\mathbf{a}}^*)_{ki}$  be a nonzero entry on the  $k$ th row. Then for a sufficiently small  $\epsilon > 0$ , changing  $(\mathbf{P}_{\mathbf{a}}^*)_{ki} \rightarrow (\mathbf{P}_{\mathbf{a}}^*)_{ki} - \epsilon$  and  $(\mathbf{P}_{\mathbf{a}}^*)_{kl} \rightarrow (\mathbf{P}_{\mathbf{a}}^*)_{kl} + \epsilon$  causes a cost reduction. Therefore  $\mathbf{a}$  is not the optimal candidate in  $\mathcal{S}_{\mathbf{a}}$  unless  $\mathbf{a} =_{\mathbf{C}} \mathbf{b}_1$ , that is,  $\mathbf{a} = \mathbf{b}_1$ .  $\square$

*Remark 4.5.* In the standard OT barycenter problem, we have

$$(4.35) \quad \lim_{(\lambda_1, \lambda_2) \rightarrow (1, 0)} \operatorname{argmin}_{\mathbf{a}} \left( \sum_{j=1}^2 \lambda_j L_{\text{OT}}(\mathbf{a}, \mathbf{b}_j; \mathbf{C}) \right) = \mathbf{b}_1,$$

$$\lim_{(\lambda_1, \lambda_2) \rightarrow (0, 1)} \operatorname{argmin}_{\mathbf{a}} \left( \sum_{j=1}^2 \lambda_j L_{\text{OT}}(\mathbf{a}, \mathbf{b}_j; \mathbf{C}) \right) = \mathbf{b}_2.$$

However, the sOT barycenter problem, without considering the periodic translation of  $\mathbf{b}_1$  and  $\mathbf{b}_2$ , gives exactly opposite results as shown in Theorem 4.4. We call the limits (4.32) and (4.33) a reverse and portion selection mechanism.

**5. Numerical results.** In this section, we will present several numerical experiments to validate the proposed sOT problem. Note that in Lemmas 2.2 and 2.3, we only prove the existence of  $\lambda$  for the equivalence between the double minimization formulation (2.5) and the single minimization formulation (2.8), but there is no explicit evaluation of  $\gamma$ . For the numerical simulations, we will decide the value of  $\gamma$  by the following procedure: taking several values of  $\gamma = \gamma_1, \gamma_2, \dots$  in ascending order and running the Dykstra solver for each value  $\gamma_i$ , until the total transported mass  $\langle \mathbf{P}, \mathbf{1} \rangle$  becomes unchanged (within a certain accuracy) at some  $\gamma_j$ , then we choose  $\gamma_j$  to be the value of  $\gamma$  in the simulations.

Figure 2 shows an example for this procedure. In this test, we take  $\mathbf{C}$  as the truncated  $L^2$  distance defined in (5.2) with  $C_{\text{cut}} = 0.5$ . The two marginal densities  $\mathbf{a}$  and  $\mathbf{b}$  are given as follows:

$$(5.1) \quad \mathbf{a} = \frac{1}{D} \left( e^{-\frac{(x-0.2)^2}{0.1^2}} + 0.001 \right) \Big|_{h\mathbb{Z} \cap [0, 1]}, \quad \mathbf{b} = \frac{1}{D} \left( e^{-\frac{(x-0.8)^2}{0.1^2}} + 0.001 \right) \Big|_{h\mathbb{Z} \cap [0, 1]}$$

with  $h = \frac{1}{200}$ . We take  $\gamma = 0, 0.1, 0.2, 0.5, 100$ . For each value of  $\gamma$ , we calculate  $\langle \mathbf{P}, \mathbf{1} \rangle$ . We note that when  $\gamma = 0.5$  or larger,  $\langle \mathbf{P}, \mathbf{1} \rangle$  remains unchanged. Thereby we take  $\gamma = 0.5$ .

**5.1. Effect of  $\epsilon$  (regulator weight) for entropy regularized sOT.** In our first example, the two marginal densities are taken as one-dimensional (1D) discretized Gaussian distribution:

$$\mathbf{a} = \frac{1}{2D} \left( e^{-\frac{(x-0.2)^2}{0.1^2}} + 0.001 \right) \Big|_{h\mathbb{Z} \cap [0, 1]}, \quad \mathbf{b} = \frac{1}{D} \left( e^{-\frac{(x-0.8)^2}{0.1^2}} + 0.001 \right) \Big|_{h\mathbb{Z} \cap [0, 1]}$$

with  $h = \frac{1}{200}$ .  $D$  is a normalization constant such that  $\|\mathbf{a}\|_1 = \frac{1}{2}$  and  $\|\mathbf{b}\|_1 = 1$ . The cost matrix  $\mathbf{C} = (C_{ij}) \in \mathbb{R}^{200 \times 200}$  is taken as

$$(5.2) \quad C_{ij} = \begin{cases} |x_i - x_j|^2 & \text{if } |x_i - x_j| \leq C_{\text{cut}}, \\ \infty & \text{if } |x_i - x_j| > C_{\text{cut}} \end{cases}$$

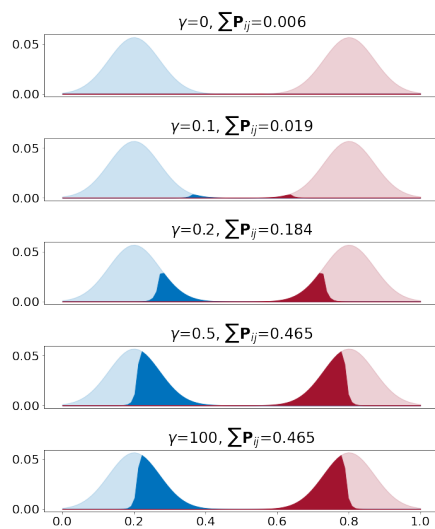


FIG. 2. The schematic for taking the value of  $\gamma$ . We take several values of  $\gamma$  in ascending order, run the Dykstra algorithm, and evaluate the total transported mass  $\langle \mathbf{P}, \mathbf{1} \rangle$  until it becomes unchanged within a certain accuracy. In this simulation, we take  $\epsilon = 0.01$ . Note that when  $\gamma = 0$ , there is still a tiny mass transported, which is due to the approximation of entry regularization.

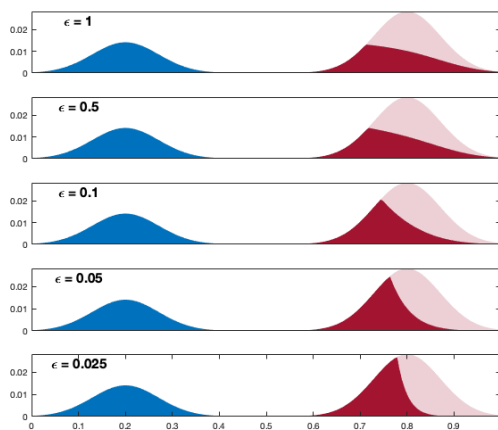


FIG. 3. The sOT solutions for various  $\epsilon$  as  $\epsilon \rightarrow 0$ . The blue region indicates the transported part of  $\mathbf{a}$ . The dark red region indicates the transported part of  $\mathbf{b}$ , and the light red region corresponds to the blocked part of  $\mathbf{b}$ .

with cutoff = 0.7. The value of  $\gamma$  is taken to be  $\gamma = 2$ . By taking various values of  $\epsilon = 1, 0.5, 0.1, 0.05, 0.025$  in Figure 3, we test the generalized Sinkhorn algorithm with  $h_1$  and  $h_2$  taken in the form of (3.16). For all values of  $\epsilon$ ,  $\mathbf{a}$  is completely transported. For the transported part of  $\mathbf{b}$ , namely, the row sum  $\mathbf{P}^T \mathbf{1}$  of the optimal plan  $\mathbf{P}$ , it spreads widely over the support of  $\mathbf{b}$ , with more mass in the region closer to the support of  $\mathbf{a}$ . As  $\epsilon$  becomes smaller, more mass is moved to the left half of the support of  $\mathbf{b}$ .



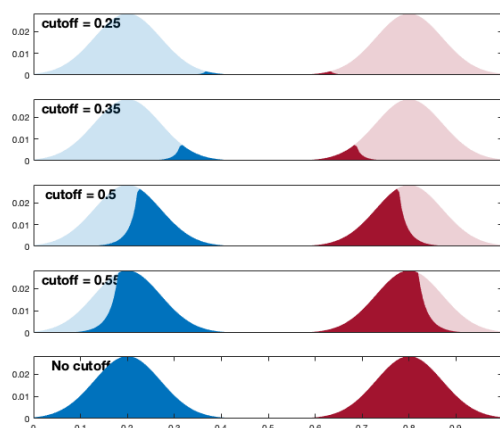


FIG. 4. The sOT solutions for various cutoff values. The dark blue region indicates the transported part of  $\mathbf{a}$ , and the light blue is for the blocked part of  $\mathbf{a}$ . The dark red region indicates the transported part of  $\mathbf{b}$ , and the light red region corresponds to the blocked part of  $\mathbf{b}$ .

**5.2. Effect of  $C_{\text{cut}}$  for entropy regularized sOT.** In the second example, we test the effect of  $\infty$  entries in the cost matrix  $\mathbf{C}$ . In particular, we construct  $\mathbf{C}$  as in (5.2) but take various cutoff values  $C_{\text{cut}} = 0.25, 0.35, 0.50, 0.55, 10$ . The marginal densities  $\mathbf{a}$  and  $\mathbf{b}$  are 1D Gaussian distributions in (5.1). Note that in this example, we take  $\|\mathbf{a}\|_1 = \|\mathbf{b}\|_1$  such that when there is no  $\infty$  entry in  $\mathbf{C}$ , the sOT problem reduces to the standard balanced OT problem. We fix  $\epsilon = 0.05$  and  $\gamma = 2$  in this example.

Figure 4 presents the solutions of the sOT problem with different values of cutoff parameter. When  $C_{\text{cut}} = 0.25$  is small, the majority of the mass is blocked; only a tiny amount on the supports of  $\mathbf{a}$  and  $\mathbf{b}$  within the separation of 0.25 is allowed to transport. When the  $C_{\text{cut}}$  becomes larger (fewer  $\infty$  entries in  $\mathbf{C}$ ), more mass is transported from the right corner of the support of  $\mathbf{a}$  to the left corner of the support of  $\mathbf{b}$ . When cutoff becomes large enough, for instance,  $C_{\text{cut}} = 10$ , all entries of  $\mathbf{C}$  become finite, and the sOT problem degenerates to the standard balanced OT problem, in which all mass in  $\mathbf{a}$  is completely transported to  $\mathbf{b}$ .

**5.3. The effect of  $C_{\text{cut}}$  on weighted sOT barycenter problem.** In this example, two unequal densities  $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{R}^{200}$  are taken as the discretization of

$$\mathbf{b}_1 = \frac{1}{D_1} (\chi_{[0.1, 0.3]} + 0.001) \Big|_{h\mathbb{Z} \cap [0, 1]}, \quad \mathbf{b}_2 = \frac{1}{D_2} \left( e^{-\frac{(x-0.8)^2}{0.1^2}} + 0.001 \right) \Big|_{h\mathbb{Z} \cap [0, 1]}$$

over a uniform mesh  $\{x_j = jh\}_{j=0}^{200}$  with  $h = \frac{1}{200}$ . The constants  $D_1, D_2$  are taken such that  $\|\mathbf{b}_1\|_1 = 1.0$  and  $\|\mathbf{b}_2\|_1 = 1.2$ . The cost matrix  $\mathbf{C} = (C_{ij}) \in \mathbb{R}^{200 \times 200}$  is taken as in (5.2) with various cutoff values. The value of  $\gamma$  is taken to be  $\gamma = 2$ . For each cutoff value, three pairs of weights  $(\lambda_1, \lambda_2) = (0.9, 0.1), (0.5, 0.5), (0.1, 0.9)$  are considered. Figure 5 depicts the numerical simulations for  $C_{\text{cut}} = 1.00, 0.50, 0.40, 0.30, 0.28, 0.26, 0.24, 0.22, 0.20$ , from top left, top middle, top right, to bottom right. In each plot, the opaque light blue (and red, respectively) represents the density  $\mathbf{b}_1$  (and  $\mathbf{b}_2$ , respectively), the transparent dark blue (and red, respectively) represents the transported mass of  $\mathbf{b}_1$  (and  $\mathbf{b}_2$ , respectively), and the transparent yellow is the weighted sOT barycenter.

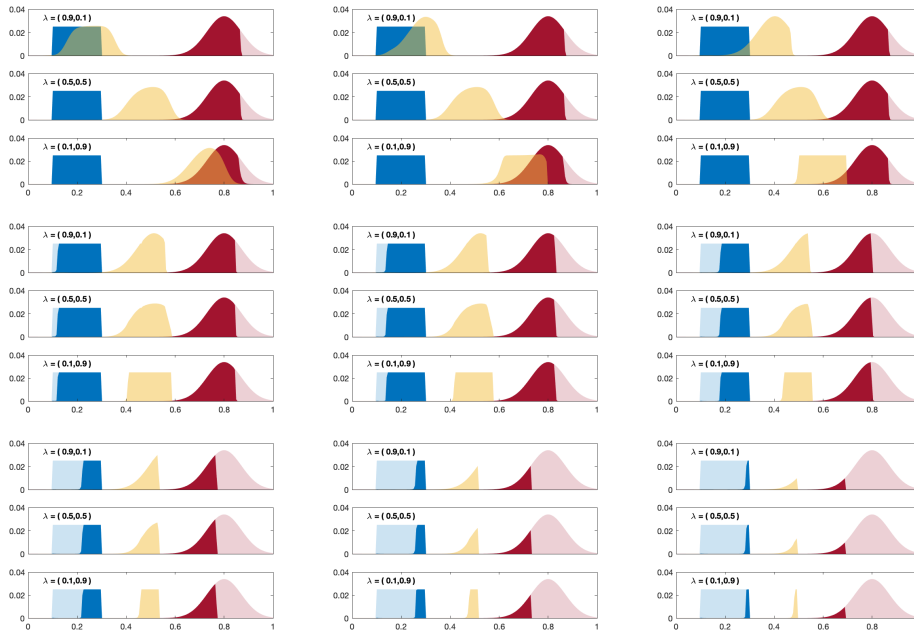


FIG. 5. The weighted sOT barycentric solutions for  $(\mathbf{b}_1, \mathbf{b}_2)$  for various values of  $C_{\text{cut}}$  and weights  $(\lambda_1, \lambda_2)$ . From top left, top middle, top right, to bottom right, with  $C_{\text{cut}} = 1.00, 0.50, 0.40, 0.30, 0.28, 0.26, 0.24, 0.20$ , respectively, each subfigure consists of three cases with  $(\lambda_1, \lambda_2) = (0.9, 0.1), (0.5, 0.5), (0.1, 0.9)$ .

For a large  $C_{\text{cut}} = 1.00$  in the top left subfigure, when  $\lambda_1$  is close to 1, the sOT barycenter  $\mathbf{a}$  is close to  $\mathbf{b}_1$ . However,  $\mathbf{b}_2$  is only partially transported as the mass of  $\mathbf{b}_2$  is more than that of  $\mathbf{b}_1$ . When  $\lambda_1$  approaches 0, the sOT barycenter becomes close to partially transported  $\mathbf{b}_2$ . This is similar to the standard barycentric problem, except that each density  $\mathbf{b}_j$  may be only partially transported.

For a small  $C_{\text{cut}} = 0.50$  in the top middle subfigure, when  $\lambda_1$  is close to 1, the sOT barycenter  $\mathbf{a}$ , compactly supported near the compact domain of  $\mathbf{b}_1$ , is surprisingly similar to  $\mathbf{b}_2$  (up to a translational shift). On the contrary, when  $\lambda_1$  is close to 0, the sOT barycenter  $\mathbf{a}$ , compactly supported near the compact domain of  $\mathbf{b}_2$ , resembles  $\mathbf{b}_1$ , up to a translational shift.

When taking further small value  $C_{\text{cut}} = 0.24$  as in the bottom left subfigure, each density  $\mathbf{b}_j$  is only allowed to transport within the distance  $= 0.24$ . Hence for any pair  $(\lambda_1, \lambda_2)$ , the sOT barycenter  $\mathbf{a}$  can only be compactly supported in between  $\mathbf{b}_1$  and  $\mathbf{b}_2$ . Besides, without considering the translational shift,  $\mathbf{a}$  is valued closely to the transported  $\mathbf{b}_2$  when  $\lambda_1 \rightarrow 1$ , while  $\mathbf{a}$  is valued closely to  $\mathbf{b}_1$  when  $\lambda_1 \rightarrow 0$ . The sOT barycenter solution in the bottom right subfigure for  $C_{\text{cut}} = 0.20$  is similar to that for  $C_{\text{cut}} = 0.24$ , except that more mass is blocked for  $\mathbf{b}_1$  and  $\mathbf{b}_2$ .

In Figure 6, we further present the resemblance phase diagram between the sOT barycenter  $\mathbf{a}$  and the marginal distribution  $\mathbf{b}_1$  on the  $\lambda_1$ - $C_{\text{cut}}$  plane (the left subfigure), and similarly that between  $\mathbf{a}$  and  $\mathbf{b}_2$  on the  $\lambda_2$ - $C_{\text{cut}}$  plane (the right subfigure). Here the resemblance between  $\mathbf{a}$  and  $\mathbf{b}_j$  is defined as

$$\text{Resem}(\mathbf{a}, \mathbf{b}_j) := \frac{\min_{0 \leq k \leq 200} \|\mathbf{b}_j - \text{circshift}(\mathbf{a}, k)\|_2}{\min_{0 \leq k \leq 200} \|\mathbf{b}_1 - \text{circshift}(\mathbf{b}_2, k)\|_2}, \quad j = 1, 2,$$

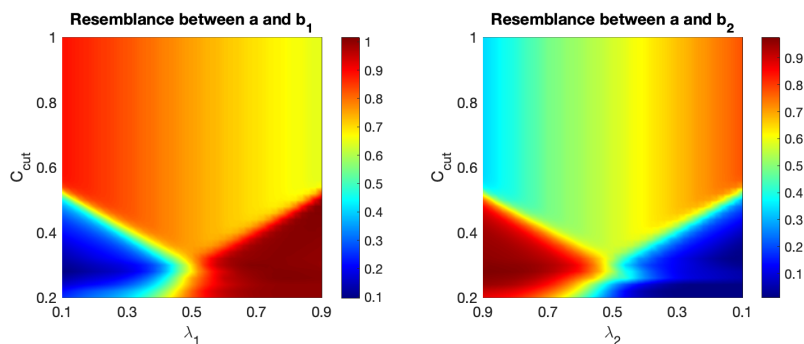


FIG. 6. The resemblance between the sOT barycenter  $\mathbf{a}$  and the marginal distributions  $\mathbf{b}_1, \mathbf{b}_2$ . When  $C_{\text{cut}}$  is small, it shows the reverse and portion selection mechanism as indicated by the limits in Theorem 4.4. When  $C_{\text{cut}}$  increases, the mechanism gradually changes back to normal as the standard OT barycenter problem.

in which `circshift` is a circular shift operator. The smaller the resemblance value, the greater the resemblance between the two densities. If  $\text{Resem}(\mathbf{x}, \mathbf{y}) = 0$ , then  $\mathbf{x} = \mathbf{y}$  up to a translational shift. It is evident that when  $C_{\text{cut}}$  is small, the resemblance of the sOT barycenter  $\mathbf{a}$  to  $\mathbf{b}_1$  (and  $\mathbf{b}_2$ , respectively) as  $(\lambda_1, \lambda_2) \rightarrow (1, 0)$  (and  $(\lambda_1, \lambda_2) \rightarrow (0, 1)$ , respectively) is reversed, compared to the standard OT barycenter problem in which  $\mathbf{a}$  resembles  $\mathbf{b}_1$  when  $\lambda_1 = 1$ , and  $\mathbf{a}$  resembles  $\mathbf{b}_2$  when  $\lambda_2 = 1$ . On the other hand, as the value of  $C_{\text{cut}}$  increases, the reverse effect is lessened. When  $C_{\text{cut}}$  becomes sufficiently large such that  $\mathbf{C}$  contains no  $\infty$  entries, the sOT barycenter problem degenerates to the standard OT barycenter problem, and the reverse mechanism turns back to normal.

**5.4. Color transfer.** Finally, we apply sOT to an important class of image processing problem, the color transfer problem. Specifically, color transfer imposes the color of a target image to an input image so that the output image has the same pattern and geometry as the input image but with the color palette from the target image. This can be viewed as transferring the histogram of pixels in the 3D color space of an image to another [33] which OT is powerful at. Direct application of conventional OT causes issues, and several OT based algorithms have been introduced to resolve these drawbacks. For example, adding regularization helps increase the robustness and eliminates outliers [18, 34]. Another issue is that transferring the entire color palette that is very different from the input image results in unrealistic looks. Fixing the amount of transferred mass a priori can help mitigate this issue but with the need of deciding a scale for each application case [7]. With sOT, we are able to directly control the similarity of transferred color by setting a distance threshold of transferred color in the color space. As a result, we control the color palette similarity of the output and the input image.

For an image, we represent the  $n$  pixels as a point cloud  $\mathbf{X} \in \mathbb{R}^{n \times 3}$  in the 3D color space (the RGB space). Given two images represented by  $\mathbf{X} \in \mathbb{R}^{n \times 3}$  (input image) and  $\mathbf{Y} \in \mathbb{R}^{m \times 3}$  (target image), the color transfer problem is formulated as coupling two uniform distributions  $\mathbf{a} \in \mathbb{R}_+^n$  and  $\mathbf{b} \in \mathbb{R}_+^m$  with the cost matrix  $\mathbf{C} \in \mathbb{R}_+^{n \times m}$  where  $C_{ij} = \|X_i - Y_j\|_2^2$ . In sOT, we use a modified cost matrix  $\tilde{\mathbf{C}}$  such that  $\tilde{C}_{ij} = C_{ij}$  for  $C_{ij} \leq C_{\text{cut}}$  and  $\tilde{C}_{ij} = \infty$  otherwise. When dealing with large images, a subsampling and upsampling is often implemented to improve efficiency [18]. We first obtain

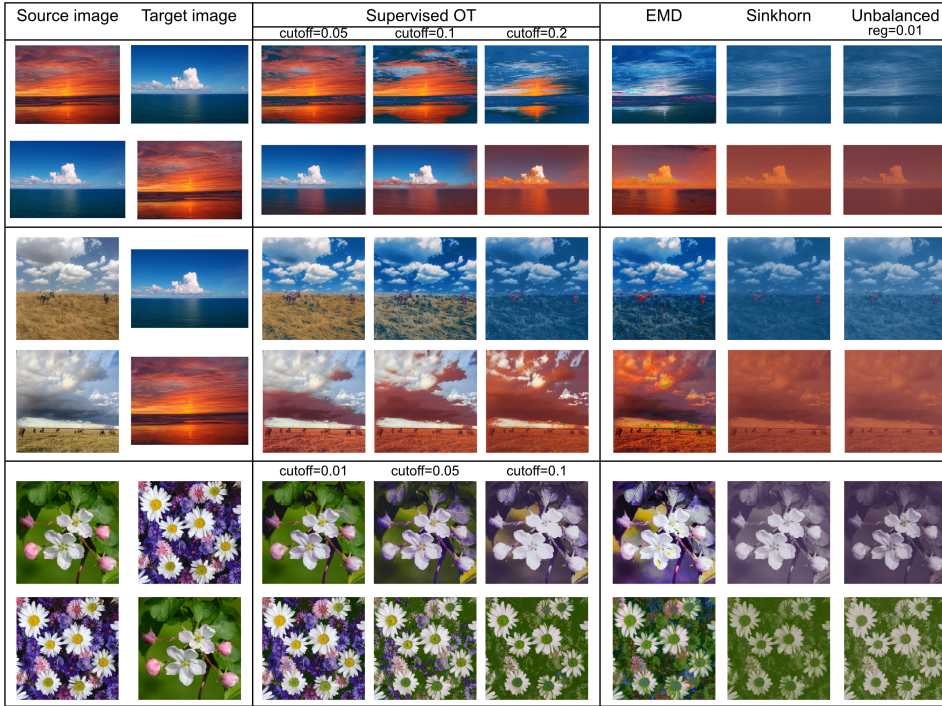


FIG. 7. The color transfer problem where the color palette of the target image is to be transferred to the source image and the output image keeps the geometry of the source image. The results of sOT with different cutoffs in the color space, earth mover's distance, entropy regularized OT (Sinkhorn), and unbalanced OT are shown.

subsampled images  $\mathbf{X}^s$  and  $\mathbf{Y}^s$  using the resize function from the PIL package [39] with the ANTIALIAS option. The OT map  $\mathbf{P}^s$  between the subsampled images is then determined by the sOT algorithm. The output image  $\mathbf{X}^{\text{out}} \in \mathbb{R}^{n \times 3}$  with the transferred color palette is constructed such that

$$(5.3) \quad \mathbf{X}_i^{\text{out}} = (\mathbf{P}^s \mathbf{Y}^s)_{N(i)} / a_{N(i)}^s + \mathbf{X}_{N(i)}^s (1 - \sum_j \mathbf{P}_{N(i),j}^s / a_{N(i)}^s) - \mathbf{X}_{N(i)}^s + \mathbf{X}_i,$$

where  $N(i)$  is the index of the pixel in the subsampled image  $\mathbf{X}^s$  that is the closest (in color space) to pixel  $i$  in the input image, and  $a_{N(i)}^s$  is the source distribution of the subsampled image. The color difference between the input and output images is determined by the color difference due to color transfer in the subsampled input and output images. We take  $\gamma = 2$  in this example.

The numerical results demonstrate that when the color palettes of the input and target images are considerably distant, applying color transfer produces unrealistic output images (Figure 7). In contrast, with the supervised cost matrix in sOT, the amount of transferred color can be controlled, producing more realistic output images (Figure 7). When the cutoff value in sOT equals infinity, the output image will converge to the result of regularized OT or earth mover's distance depending on whether entropy regularization was used.

**6. Conclusion and discussion.** In this work, we introduce the supervised optimal transport problem where the cost matrix  $\mathbf{C}$  can have  $\infty$  entries and the  $\infty$ -pattern of  $\mathbf{C}$  supervises and controls the transport plans. Also, the source and target distributions need not be normalized in sOT. These properties make sOT a method generally applicable to a large class of transportation problems where application-specific constraints are to be imposed on the transport plan and the original units are to be preserved for the original distributions. To apply sOT on large-scale real problems, we develop a fast numerical solver for sOT based on the Dykstra algorithm. We also extend the OT barycenter problem into a supervised one in the setup of sOT. By considering the sOT barycenter for two distributions  $(\mathbf{b}_1, \mathbf{b}_2)$ , a new reverse and portion selection mechanism is discovered, giving the barycenter  $\mathbf{a}$  opposite to that of the standard OT barycenter problem when the weight approaches the boundary of the unit simplex. The properties and effects of different parameter values of sOT are illustrated numerically with toy examples. We also demonstrate the reverse behavior of the sOT barycenter in an extensive numerical example. In an important problem in imaging science, we compare sOT to several other OT variants to demonstrate its unique utility of supervising the transport plan.

This work can be extended in several ways in the future. For example, a supervised Gromov–Wasserstein OT analogous to sOT can be developed and will enable the integration of multiple subsamples of the same system without known intersample correspondence.

In the current work, the formulations (2.8) and (2.17) are obtained from the discrete OT setting. Introducing the entropy regularization terms in (3.4), the optimal plan  $\mathbf{P}_\epsilon$  has a diagonal rescaling form, such that the Dykstra algorithm (Sinkhorn type) can be applied to improve the numerical efficiency significantly. On the other hand, it is also of practical interest to find an efficient solver for the sOT problem without entropy regularization. Inspired by [28], we may link (2.8) and (2.17) with the dynamical OT (Benamou–Breiner type [5]) formulation. Discrete sOT formulations (2.8) and (2.17) work for any “ground metric”  $\mathbf{C}$ . On the other hand, if the metric  $\mathbf{C}$  is homogeneous of degree one such as  $L^1$  metric  $\mathbf{C} = (\|x_i - y_j\|_1)_{ij}$ , the original OT problem can be reformulated as a minimal flux minimization problem [27]. Motivated by this, it is also interesting to consider, when taking some degree-one homogeneous metric  $\mathbf{C}$  in (2.17) with a certain  $\infty$ -pattern, whether there exists a minimal flux formulation for the sOT problem. Then we can apply well-established efficient solvers for  $L^1$  minimization problems to sOT provided that it can be reformulated as a minimal flux problem. Such a potential dynamical formulation for the sOT problem can be used to seek optimal transportation paths with constraints in applications such as continent movement in geology.

## REFERENCES

- [1] A. G. WILSON, *The use of entropy maximizing models, in the theory of trip distribution, mode split and route split*, J. Transport Econ. Pol., 3 (1969), pp. 108–126.
- [2] J. ALTSCHULER, J. WEED, AND P. RIGOLLET, *Near-linear time approximation algorithms for optimal transport via Sinkhorn iteration*, in Proceedings of the 31st International Conference on Neural Information Processing Systems, Curran Associates, Red Hook, NY, 2017, pp. 1961–1971.
- [3] M. ARJOVSKY, S. CHINTALA, AND L. BOTTOU, *Wasserstein generative adversarial networks*, in Proceedings of the International Conference on Machine Learning, PMLR, 2017, pp. 214–223.
- [4] M. BAUER, S. JOSHI, AND K. MODIN, *Diffeomorphic density matching by optimal information transport*, SIAM J. Imaging Sci., 8 (2015), pp. 1718–1751.

- [5] J. BENAMOU AND Y. BRENIER, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math., 84 (2000), pp. 375–393.
- [6] J. BENAMOU, G. CARLIER, M. CUTURI, L. NENNA, AND G. PEYRE, *Iterative Bregman projections for regularized transportation problems*, SIAM J. Sci. Comput., 37 (2015), pp. A1111–A1138.
- [7] N. BONNEEL AND D. COEURJOLLY, *SPOT: Sliced partial optimal transport*, ACM Trans. Graph., 38 (2019), pp. 1–13.
- [8] Y. BRENIER, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math., 44 (1991), pp. 375–417.
- [9] R. A. BRUALDI, *Combinatorial Matrix Classes*, Cambridge University Press, Cambridge, UK, 2006.
- [10] L. CHAPEL, M. Z. ALAYA, AND G. GASSO, *Partial optimal transport with applications on positive-unlabeled learning*, Adv. Neural Inf. Process Syst., 33 (2020), pp. 2903–2913.
- [11] L. CHIZAT, G. PEYRE, B. SCHMITZER, AND F.-X. VIALARD, *Scaling algorithms for unbalanced optimal transport problems*, Math. Comput., 26 (2013), pp. 2563–2609.
- [12] L. CHIZAT, G. PEYRÉ, B. SCHMITZER, AND F.-X. VIALARD, *Unbalanced optimal transport: Dynamic and Kantorovich formulations*, J. Funct. Anal., 274 (2018), pp. 3090–3123.
- [13] R. COMINETTI AND J. S. MARTIN, *Asymptotic analysis of the exponential penalty trajectory in linear programming*, Math. Program., 67 (1994), pp. 169–187.
- [14] N. COURTY, R. FLAMARY, D. TUIA, AND A. RAKOTOMAMONJY, *Optimal transport for domain adaptation*, IEEE Trans. Pattern Anal. Mach. Intell., 39 (2016), pp. 1853–1865.
- [15] M. CUTURI, *Sinkhorn distances: Lightspeed computation of optimal transport*, Adv. Neural Inf. Process Syst., 26 (2013), pp. 2292–2300.
- [16] M. CUTURI AND A. DOUCET, *Fast computation of Wasserstein barycenters*, in Proceedings of the International Conference on Machine Learning, PMLR, 2014, pp. 685–693.
- [17] P. DVURECHENSKY, A. GASNIKOV, AND A. KROSHNIN, *Computational optimal transport: Complexity by accelerated gradient descent is better than by Sinkhorn’s algorithm*, in Proceedings of the International Conference on Machine Learning, PMLR, 2018, pp. 1367–1376.
- [18] S. FERRADANS, N. PAPADAKIS, G. PEYRÉ, AND J.-F. AUJOL, *Regularized discrete optimal transport*, SIAM J. Imaging Sci., 7 (2014), pp. 1853–1882.
- [19] J. FRANKLIN, AND J. LORENTZ, *On the scaling of multidimensional matrices*, Linear Algebra Appl., 114 (1989), pp. 717–735.
- [20] U. FRISCH, S. MATARRESE, R. MOHAYAEI, AND A. SOBOLEVSKI, *A reconstruction of the initial conditions of the universe by optimal mass transportation*, Nature, 417 (2012), pp. 260–262.
- [21] W. GANBO, W. LI, S. OSHER, AND M. PUTHAWALA, *Unnormalized optimal transport*, J. Comput. Phys., 399 (2019), p. 108940.
- [22] W. GUO, N. HO, AND M. JORDAN, *Fast algorithms for computational optimal transport and Wasserstein barycenter*, in Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research 108, S. Chiappa and R. Calandra, eds., PMLR, 2020, pp. 2088–2097, <https://proceedings.mlr.press/v108/guo20a.html>.
- [23] W. GUO, N. HO, AND M. JORDAN, *Fast algorithms for computational optimal transport and Wasserstein barycenter*, in Proceedings of the International Conference on Artificial Intelligence and Statistics, PMLR, 2020, pp. 2088–2097.
- [24] L. KANTOROVITCH, *On the translocation of masses*, C. R. (Doklady) Acad. Sci. URSS (N.S.), 37 (1942), pp. 199–201.
- [25] J. KARLSSON AND A. RINGH, *Generalized Sinkhorn iterations for regularizing inverse problems using optimal mass transport*, SIAM J. Imaging Sci., 10 (2017), pp. 1935–1962.
- [26] W. LEE, R. LAI, W. LI, AND S. OSHER, *Generalized unnormalized optimal transport and its fast algorithms*, J. Comput. Phys., 436 (2021), p. 110041.
- [27] W. LI, E. RYU, S. OSHER, W. YIN, AND W. GANBO, *A parallel method for earth mover’s distance*, J. Sci. Comput., 75 (2018), pp. 182–197.
- [28] W. LI, P. YIN, AND S. OSHER, *Computations of optimal transport distance with fisher information regularization*, J. Sci. Comput., 75 (2018), pp. 1581–1595.
- [29] T. LIN, N. HO, AND M. JORDAN, *On efficient optimal transport: An analysis of greedy and accelerated mirror descent algorithms*, in Proceedings of the International Conference on Machine Learning, PMLR, 2019, pp. 3982–3991.
- [30] G. MONGE, *Memoire sur la theorie des deblais et des remblais*, Histoire de l’Academie Royale des Sciences, 1781.
- [31] G. PEYRE, *Entropic approximation of Wasserstein gradient flows*, SIAM J. Imaging Sci., 8 (2015), pp. 2323–2351.

- [32] G. PEYRÉ, AND M. CUTURI, *Computational optimal transport: With applications to data science*, Found. Trends Mach. Learn., 11 (2019), pp. 355–607, <https://doi.org/10.1561/22000000073>.
- [33] F. PITIE, A. C. KOKARAM, AND R. DAHYOT, *N-dimensional probability density function transfer and its application to color transfer*, in Proceedings of the 10th IEEE International Conference on Computer Vision, IEEE, 2005, pp. 1434–1439.
- [34] J. RABIN, S. FERRADANS, AND N. PAPADAKIS, *Adaptive color transfer with relaxed optimal transport*, in Proceedings of the IEEE International Conference on Image Processing, IEEE, 2014, pp. 4852–4856.
- [35] E. SCHRÖDINGER, *Über die umkehrung der naturgesetze*, Phys. Math., 144 (1931), pp. 144–153.
- [36] R. SINKHORN, *A relationship between arbitrary positive matrices and doubly stochastic matrices*, Ann. Math. Stat., 35 (1964), pp. 876–879.
- [37] R. SINKHORN, *Diagonal equivalence to matrices with prescribed row and column sums*, Amer. Math. Monthly, 74 (1967), pp. 402–405.
- [38] R. SINKHORN AND P. KNOPP, *Concerning nonnegative matrices and doubly stochastic matrices*, Pacific J. Math., 21 (1967), pp. 343–348.
- [39] P. UMESH, *Image processing in python*, CSI Commun., 23 (2012).
- [40] C. VILLANI, *Topics in Optimal Transportation*, AMS, Providence, RI, 2003.