J Sci Comput
DOI 10.1007/s10915-017-0600-y

Unbalanced and Partial L_1 Monge–Kantorovich Problem: A Scalable Parallel First-Order Method

Ernest K. Ryu[1] · **Wuchen Li**[1] · **Penghang Yin**[1] · **Stanley Osher**[1]

Received: 10 December 2016 / Revised: 26 September 2017 / Accepted: 6 November 2017

Abstract We propose a new algorithm to solve the unbalanced and partial L_1-Monge–Kantorovich problems. The proposed method is a first-order primal-dual method that is scalable and parallel. The method's iterations are conceptually simple, computationally cheap, and easy to parallelize. We provide several numerical examples solved on a CUDA GPU, which demonstrate the method's practical effectiveness.

Keywords Earth Mover's distance · Optimal transport · First-order methods · Primal-dual algorithm

1 Introduction

The Monge–Kantorovich problem [8], also named the Wasserstein metric or earth mover's distance, defines a metric between two densities on the probability set and is used in many applications including image processing, optical flow, computer vision, and statistics [3,13,17,23]. The original problem assumes that the total masses of the two given densities are equal, which often does not hold in practice. For instance, it is natural to compare two images of different intensities. Therefore, it is very useful to generalize the Wasserstein metric or the

This work is partially supported by ONR Grants N000141410683, N000141210838 and DOE Grant DE-SC00183838.

✉ Wuchen Li
wcli@math.ucla.edu

Ernest K. Ryu
eryu@math.ucla.edu

Penghang Yin
yph@math.ucla.edu

Stanley Osher
sjo@math.ucla.edu

[1] Department of Mathematics, University of California, Los Angeles, CA, USA

Published online: 15 November 2017

earth mover's distance to densities with unbalanced masses. There has been much work and interest in this direction [4,6,7,9,10,18,19], and we focus on the two such approaches, the unbalanced and partial L_1 Monge–Kantorovich problems, in this paper.

Following the idea in [14], we propose a scalable parallel method to solve the unbalanced and partial L_1 Monge–Kantorovich problems. Our algorithm uses a finite volume method to discretize the domain and then applies the Chambolle–Pock primal-dual method [5,21]. As a first-order method, our algorithm has the following advantages: (1) the "shrink" operator promotes sparsity; (2) each iteration is conceptually very simple and computationally very cheap (each iteration does not even solve a linear system); (3) it is easy to parallelize, and, in particular, can effectively utilize the computational power of CUDA GPUs; (4) the cost of each iteration scales well with the problem or discretization size.

A few algorithms have been proposed in this area. Ling et al. [15] and Rubner et al. [23] cast the discretized optimization problem into a linear program and solve it, in a similar setup. The disadvantage of this approach is, however, that the size of the linear program grows quadratically with the discretization size. Barrett and Prigozhin considers the same unbalanced and partial L_1 model. They approximate the L_1 norm with the L_r norm, solve the smooth approximation with ADMM, and let $r \downarrow 1$ [1]. We propose a different approach. Our approach handles the L_1 norm directly, which promotes sparsity [26], and it uses explicit updates, while ADMM requires computing the inverse of an elliptic operator every iteration.

This paper is organized as follows. In Sect. 2, we briefly review the unbalanced and partial L_1 Monge–Kantorovich problems. In Sect. 3, we propose a scalable parallel first-order method to solve these problems. In Sect. 4, we discuss the computational issues such as parallelization and parameter tuning of the proposed method. In Sect. 5, we show several numerical examples to demonstrate the method's effectiveness. In Sect. 6, we discuss existence and uniqueness of the unbalanced and partial L_1 Monge–Kantorovich problems, and propose how to regularize the problems to ensure the solution is unique.

2 L_1 Monge–Kantorovich Problem

In this section, we briefly review the balanced, unbalanced, and partial L_1 Monge–Kantorovich problem. The unbalanced and partial problems have been studied by [1,20]. Several related setups, which include both L_1, L_2 and L_2^2 cases, have been studied by [4,6,7,9,10,18,19].

Throughout this paper, assume $\Omega \subset \mathbb{R}^d$ is convex and compact. We write $\|\cdot\|$ for the standard Euclidean norm.

2.1 Balanced L_1 Monge–Kantorovich Problem

Let ρ^0 and ρ^1 be nonnegative densities supported on Ω with balanced mass, i.e.,

$$\int_\Omega \rho^0(\mathbf{x})\, d\mathbf{x} = \int_\Omega \rho^1(\mathbf{x})\, d\mathbf{x}.$$

The optimal transport map from ρ^0 to ρ^1 solves

$$\underset{T}{\text{minimize}} \quad \int_\Omega \|\mathbf{x} - T(\mathbf{x})\| \rho^0(\mathbf{x})\, d\mathbf{x}. \tag{1}$$

The optimization variable $T : \Omega \to \Omega$ is smooth, one-to-one, and transfers $\rho^0(\mathbf{x})$ to $\rho^1(\mathbf{x})$, i.e., T satisfies

$$\rho^0(\mathbf{x}) = \rho^1(T(\mathbf{x}))\det(\nabla T(\mathbf{x})).$$

The optimization problem (1) is nonlinear and nonconvex. We can relax (1) into a linear (convex) optimization problem:

$$W(\rho^0, \rho^1) = \begin{pmatrix} \underset{\mathbf{m}}{\text{minimize}} & \int_{\Omega\times\Omega} \|\mathbf{x}-\mathbf{y}\|\pi(\mathbf{x},\mathbf{y})\, d\mathbf{x}d\mathbf{y} \\ \text{subject to} & \pi(\mathbf{x},\mathbf{y}) \geq 0 \\ & \int_\Omega \pi(\mathbf{x},\mathbf{y})\, d\mathbf{y} = \rho^0(\mathbf{x}) \\ & \int_\Omega \pi(\mathbf{x},\mathbf{y})\, d\mathbf{x} = \rho^1(\mathbf{y}). \end{pmatrix} \tag{2}$$

The optimization variable π is a joint nonnegative measure on $\Omega \times \Omega$ having $\rho^0(\mathbf{x})$ and $\rho^1(\mathbf{y})$ as marginals. To clarify, $W(\rho^0, \rho^1)$ denotes the optimal value of (2).

The theory of optimal transport [8,25] remarkably points out that (2) is equivalent to the following flux minimization problem:

$$W(\rho^0, \rho^1) = \begin{pmatrix} \underset{\mathbf{m}}{\text{minimize}} & \int_\Omega \|\mathbf{m}(\mathbf{x})\|\, d\mathbf{x} \\ \text{subject to} & \nabla\cdot\mathbf{m}(\mathbf{x}) = \rho^0(\mathbf{x}) - \rho^1(\mathbf{x}) \\ & \mathbf{m}(\mathbf{x})\cdot\mathbf{n}(\mathbf{x}) = 0, \ \text{ for all } \begin{cases} \mathbf{x} \in \partial\Omega, \\ \mathbf{n}(\mathbf{x}) \text{ normal to } \partial\Omega. \end{cases} \end{pmatrix} \tag{3}$$

Although (3) and (2) are mathematically equivalent, (3) is much more computationally effective as its optimization variable $\mathbf{m}$ is much smaller when discretized. It is clear that (2) or (3) requires ρ^0 and ρ^1 have balanced mass; (2) by Fubini's theorem and (3) by the divergence theorem. Finally, $W(\rho^0, \rho^1)$ defines a metric on the set of probability measures and thus is called the 1-Wasserstein metric.

2.2 Unbalanced L_1 Monge–Kantorovich Problem

Let ρ^0 and ρ^1 be nonnegative densities supported on Ω with possibly unbalanced mass, i.e., we allow

$$\int_\Omega \rho^0(\mathbf{x})\, d\mathbf{x} \neq \int_\Omega \rho^1(\mathbf{x})\, d\mathbf{x}.$$

Without loss of generality, assume

$$\int_\Omega \rho^0(\mathbf{x})\, d\mathbf{x} \leq \int_\Omega \rho^1(\mathbf{x})\, d\mathbf{x}.$$

The *unbalanced* L_1 *Monge–Kantorovich problem* solves

$$U(\rho^0, \rho^1) = \begin{pmatrix} \text{minimize} & W(\rho^0, \tilde{\rho}^1) \\ \text{subject to} & 0 \leq \tilde{\rho}^1(\mathbf{x}) \leq \rho^1(\mathbf{x}) \\ & \int \rho^0(\mathbf{x})\, d\mathbf{x} = \int \tilde{\rho}^1(\mathbf{x})\, d\mathbf{x} \end{pmatrix}. \tag{4}$$

To idea is that we fully transport ρ^0, the smaller mass, to partially fill ρ^1, the larger mass, and $U(\rho^0, \rho^1)$ is the optimal (smallest) cost of doing so. If $\int_\Omega \rho^0(\mathbf{x})\, d\mathbf{x} > \int_\Omega \rho^1(\mathbf{x})\, d\mathbf{x}$, we simply flip the definition and interpretation. We can write (4) as a single equivalent optimization problem:

$$U(\rho^0,\rho^1)=\left(\begin{array}{ll}\underset{\mathbf{m},\tilde{\rho}^1}{\text{minimize}} & \int_\Omega \|\mathbf{m}(\mathbf{x})\|\,d\mathbf{x}\\ \text{subject to} & \nabla\cdot\mathbf{m}(\mathbf{x})=\rho^0(\mathbf{x})-\tilde{\rho}^1(\mathbf{x})\\ & \mathbf{m}(\mathbf{x})\cdot\mathbf{n}(\mathbf{x})=0,\ \text{for all}\ \begin{cases}\mathbf{x}\in\partial\Omega,\\ \mathbf{n}(\mathbf{x})\ \text{normal to}\ \partial\Omega\end{cases}\\ & 0\le\tilde{\rho}^1(\mathbf{x})\le\rho^1(\mathbf{x})\end{array}\right). \tag{5}$$

In (5), the constraint $\int \rho^0(\mathbf{x})\,d\mathbf{x}=\int\tilde{\rho}^1(\mathbf{x})\,d\mathbf{x}$ is enforced by $\nabla\cdot\mathbf{m}(\mathbf{x})=\rho^0(\mathbf{x})-\tilde{\rho}^1(\mathbf{x})$ and the zero-flux boundary condition on $\mathbf{m}$.

It is easy to see that when ρ^0 and ρ^1 have balanced mass the unbalanced case reduces to the balanced case, i.e., $U(\rho^0,\rho^1)=W(\rho^0,\rho^1)$ when $\int_\Omega\rho^0(\mathbf{x})\,d\mathbf{x}\neq\int_\Omega\rho^1(\mathbf{x})\,d\mathbf{x}$.

Note that U is not a metric as it satisfies neither the identity of indiscernibles nor the triangle inequality. However,

$$D(\rho^0,\rho^1)=U(\rho^0,\rho^1)+\lambda\int_\Omega|\rho^1(\mathbf{x})-\tilde{\rho}^{1,\star}(\mathbf{x})|\,d\mathbf{x},$$

where $\lambda\ge\max\{\|\mathbf{x}-\mathbf{y}\|\,|\,\mathbf{x},\mathbf{y}\in\Omega\}$ is a given constant and $\tilde{\rho}^{1,\star}$ is a minimizer of (4), defines a metric on nonnegative measures [1,10,18,19]. Of course, if ρ^0, ρ^1 have balanced mass, $D(\rho^0,\rho^1)$ reduces to the 1-Wasserstein metric $W(\rho^0,\rho^1)$.

2.3 Partial L_1 Monge–Kantorovich Problem

Let ρ^0 and ρ^1 be nonnegative densities supported on Ω with possibly unbalanced mass, and let

$$0<\gamma\le\min\left\{\int_\Omega\rho^0(\mathbf{x})\,d\mathbf{x},\int_\Omega\rho^1(\mathbf{x})\,d\mathbf{x}\right\}.$$

The *partial* L_1 *Monge–Kantorovich problem* solves

$$P_\gamma(\rho^0,\rho^1)=\left(\begin{array}{ll}\underset{\tilde{\rho}^0,\tilde{\rho}^1}{\text{minimize}} & W(\tilde{\rho}^0,\tilde{\rho}^1)\\ \text{subject to} & 0\le\tilde{\rho}^0(\mathbf{x})\le\rho^0(\mathbf{x})\\ & 0\le\tilde{\rho}^1(\mathbf{x})\le\rho^1(\mathbf{x})\\ & \gamma=\int_\Omega\tilde{\rho}^0(\mathbf{x})\,d\mathbf{x}=\int_\Omega\tilde{\rho}^1(\mathbf{x})\,d\mathbf{x}\end{array}\right). \tag{6}$$

The idea is that we partially transport a mass of γ from ρ^0 to partially fill a mass of γ of ρ^1, and $P_\gamma(\rho^0,\rho^1)$ is the optimal (smallest) cost of doing so. We can write (6) as a single equivalent optimization problem:

$$P_\gamma(\rho^0,\rho^1)=\left(\begin{array}{ll}\underset{\mathbf{m},\tilde{\rho}^0,\tilde{\rho}^1}{\text{minimize}} & \int_\Omega\|\mathbf{m}(\mathbf{x})\|\,d\mathbf{x}\\ \text{subject to} & \nabla\cdot\mathbf{m}(\mathbf{x})=\tilde{\rho}^0(\mathbf{x})-\tilde{\rho}^1(\mathbf{x})\\ & \mathbf{m}(\mathbf{x})\cdot\mathbf{n}(\mathbf{x})=0,\ \forall\text{for all}\ \begin{cases}\mathbf{x}\in\partial\Omega,\\ \mathbf{n}(\mathbf{x})\ \text{normal to}\ \partial\Omega\end{cases}\\ & 0\le\tilde{\rho}^0(\mathbf{x})\le\rho^0(\mathbf{x})\\ & 0\le\tilde{\rho}^1(\mathbf{x})\le\rho^1(\mathbf{x})\\ & \gamma=\int_\Omega\tilde{\rho}^0(\mathbf{x})\,d\mathbf{x}=\int_\Omega\tilde{\rho}^1(\mathbf{x})\,d\mathbf{x}\end{array}\right). \tag{7}$$

It is easy to see that when $\gamma=\int_\Omega\rho^0(\mathbf{x})\,d\mathbf{x}\le\int_\Omega\rho^1(\mathbf{x})\,d\mathbf{x}$, the partial case reduces to the unbalanced case, i.e., $P_\gamma(\rho^0,\rho^1)=U(\rho^0,\rho^1)$.

Robustness to noise is an interesting property of P_γ. Assume ρ^0 and ρ^1 are densities with (balanced) unit mass, and let δ^0 and δ^1 be their respective perturbations with small mass. Even though the perturbations have small mass, the difference between $W(\rho^0, \rho^1)$ and $W(\rho^0 + \delta^0, \rho^1 + \delta^1)$ can be large if the mass of δ^0 is far away from the mass of ρ^1 and δ^1. On the other hand, the difference between $P_\gamma(\rho^0, \rho^1)$ and $P_\gamma(\rho^0 + \delta^0, \rho^1 + \delta^1)$ is affected much less when $\gamma < 1$. This idea is illustrated in Figs. 1 and 2.

As before, U is not a metric. However, also as before,

$$D(\rho^0, \rho^1) = P_\gamma(\rho^0, \rho^1) + \lambda \int_\Omega |\rho^0(\mathbf{x}) - \tilde{\rho}^{0,\star}(\mathbf{x})| + |\rho^1(\mathbf{x}) - \tilde{\rho}^{1,\star}(\mathbf{x})| \, d\mathbf{x},$$

where $\lambda \geq \max\{\|\mathbf{x}-\mathbf{y}\| \mid \mathbf{x}, \mathbf{y} \in \Omega\}$ is a given constant and $\tilde{\rho}^{0,\star}$ and $\tilde{\rho}^{1,\star}$ are minimizers of (6), defines a metric on nonnegative measures [1]. Again, if $\gamma = \int_\Omega \rho^0(\mathbf{x}) \, dx = \int_\Omega \rho^1(\mathbf{x}) \, dx$, then $D(\rho^0, \rho^1)$ reduces to the 1-Wasserstein metric $W(\rho^0, \rho^1)$.

In what follows, we present a parallelizable first-order algorithm to solve $P_\gamma(\rho^0, \rho^1)$, the partial L_1 Monge–Kantorovich problem.

3 Algorithm

In this section, we derive and present the main algorithm for the partial L_1 Monge–Kantorovich problem, which can, of course, solve the unbalanced problem as a special case. We also provide a convergence proof.

3.1 Discretization

For notational simplicity, we will consider the case where $\Omega \subset \mathbb{R}^2$ and Ω is square. The following discussion does immediately generalize to higher dimensions and more complicated domains.

Also, we will use the same symbol to denote the discretizations and their continuous counterparts. Whether we are referring to the continuous variable or its discretization should be clear from the context.

Consider a $n \times n$ discretization of Ω with finite difference Δx in both x and y directions. Write the x and y coordinates of the points as $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$. So we are approximating the domain Ω with $\{x_1, \ldots, x_n\} \times \{y_1, \ldots, y_n\}$. Write $C(x, y)$ be the $\Delta x \times \Delta x$ cube centered at (x, y), i.e.,

$$C(x, y) = \{(x', y') \in \mathbb{R}^2 \mid |x' - x| \leq \Delta x/2 \,, \; |y' - y| \leq \Delta x/2\} \,.$$

We use a finite volume approximation for ρ^0, ρ^1, $\tilde{\rho}^0$, and $\tilde{\rho}^1$. Specifically, we write $\rho^0 \in \mathbb{R}^{n \times n}$ with

$$\rho^0_{ij} \approx \int_{C(x_i, y_j)} \rho^0(x, y) \, dxdy \,,$$

for $i, j = 1, \ldots, n$. The discretizations ρ^1, $\tilde{\rho}^0$, , $\tilde{\rho}^1 \in \mathbb{R}^{n \times n}$ are defined the same way.

Write $\mathbf{m} = (m_x, m_y)$ for both the continuous variable and its discretization. To be clear, the subscripts of m_x and m_y do not denote differentiation. We use the discretization $m_x \in \mathbb{R}^{(n-1) \times n}$ and $m_y \in \mathbb{R}^{n \times (n-1)}$. For $i = 1, \ldots, n-1$ and $j = 1, \ldots, n$

$$m_{x,ij} \approx \int_{C(x_i + \Delta x/2, y_j)} m_x(x, y) \, dxdy \,,$$

and for $i = 1, \ldots, n$ and $j = 1, \ldots, n-1$

$$m_{y,ij} \approx \int_{C(x_i, y_j + \Delta x/2)} m_y(x, y)\, dxdy\,.$$

In defining m_x and m_y, the center points are placed between the $n \times n$ grid points to make the finite difference operator symmetric.

Define the discrete divergence operator $\mathrm{div}(\mathbf{m}) \in \mathbb{R}^{n\times n}$ as

$$\mathrm{div}(\mathbf{m})_{ij} = \frac{1}{\Delta x}(m_{x,ij} - m_{x,(i-1)j} + m_{y,ij} - m_{y,i(j-1)})\,,$$

for $i, j = 1, \ldots, n$, where we mean $m_{x,0j} = m_{x,nj} = 0$ for $j = 1, \ldots, n$ and $m_{y,i0} = m_{y,in} = 0$ for $i = 1, \ldots, n$. This definition of $\mathrm{div}(\mathbf{m})$ makes the discrete approximation be consistent with the zero-flux boundary condition.

For $\Phi \in \mathbb{R}^{n\times n}$, define the discrete gradient operator $\nabla\Phi = ((\nabla\Phi)_x, (\nabla\Phi)_y)$ as

$$\begin{aligned}
(\nabla\Phi)_{x,ij} &= (1/\Delta x)\left(\Phi_{i+1,j} - \Phi_{i,j}\right) \quad \text{for } i = 1, \ldots, n-1,\ j = 1, \ldots, n \\
(\nabla\Phi)_{y,ij} &= (1/\Delta x)\left(\Phi_{i,j+1} - \Phi_{i,j}\right) \quad \text{for } i = 1, \ldots, n,\ j = 1, \ldots, n-1\,.
\end{aligned}$$

So $(\nabla\Phi)_x \in \mathbb{R}^{(n-1)\times n}$ and $(\nabla\Phi)_y \in \mathbb{R}^{n\times(n-1)}$, and the ∇ is the adjoint of $-\mathrm{div}$.

We will soon see that using ghost cells is convenient for both describing and implementing the method. So we define the variable $\tilde{\mathbf{m}} = (\tilde{m}_x, \tilde{m}_y) \in \mathbb{R}^{2\times n\times n}$ where

$$\begin{aligned}
\tilde{m}_{x,ij} &= \begin{cases} m_{x,ij} & \text{for } i < n \\ 0 & \text{for } i = n \end{cases} \\
\tilde{m}_{y,ij} &= \begin{cases} m_{y,ij} & \text{for } j < n \\ 0 & \text{for } j = n\,, \end{cases}
\end{aligned}$$

for $i, j = 1, \ldots, n$. We also define $\tilde{\nabla}\Phi = ((\tilde{\nabla}\Phi)_x, (\tilde{\nabla}\Phi)_y) \in \mathbb{R}^{2\times n\times n}$, where

$$\begin{aligned}
(\tilde{\nabla}\Phi)_{x,ij} &= \begin{cases} (\nabla\Phi)_{x,ij} & \text{for } i < n \\ 0 & \text{for } i = n \end{cases} \\
(\tilde{\nabla}\Phi)_{y,ij} &= \begin{cases} (\nabla\Phi)_{y,ij} & \text{for } j < n \\ 0 & \text{for } j = n\,, \end{cases}
\end{aligned}$$

for $i, j = 1, \ldots, n$. Finally, we write $\tilde{\mathbf{m}} = (\tilde{m}_x, \tilde{m}_y)$ and $\tilde{\mathbf{m}}_{ij} = (\tilde{m}_{x,ij}, \tilde{m}_{y,ij})$ and $(\tilde{\nabla}\Phi)_{ij} = ((\tilde{\nabla}\Phi)_{x,ij}, (\tilde{\nabla}\Phi)_{y,ij})$ for $i, j = 1, \ldots, n$

Using this notation, we write the discretization of (7) as

$$\begin{aligned}
\underset{\mathbf{m},\tilde{\rho}^0,\tilde{\rho}^1}{\text{minimize}} \quad & \|\mathbf{m}\|_{1,2} \\
\text{subject to} \quad & \mathrm{div}(\mathbf{m}) = \tilde{\rho}^0 - \tilde{\rho}^1 \\
& 0 \le \tilde{\rho}^0 \le \rho^0 \\
& 0 \le \tilde{\rho}^1 \le \rho^1 \\
& \gamma = \langle \mathbf{1}, \tilde{\rho}^0 \rangle = \langle \mathbf{1}, \tilde{\rho}^1 \rangle,
\end{aligned} \tag{8}$$

where $m_x \in \mathbb{R}^{(n-1)\times n}$, $m_y \in \mathbb{R}^{n\times(n-1)}$, $\tilde{\rho}^0 \in \mathbb{R}^{n\times n}$, and $\tilde{\rho}^1 \in \mathbb{R}^{n\times n}$ are the optimization variables. The inequalities are element-wise. Here, $\mathbf{1} \in \mathbb{R}^{n\times n}$ denotes the matrix filled with

1s and $\langle \cdot, \cdot \rangle$ denotes the inner product between $n \times n$ matrices treated as vectors. So

$$\langle \mathbf{1}, \tilde{\rho}^0 \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \tilde{\rho}_{ij}.$$

The boundary conditions are implicitly handled by the discretization. The objective is

$$\|\mathbf{m}\|_{1,2} = \sum_{i=1}^{n} \sum_{j=1}^{n} \|\mathbf{m}_{ij}\|_2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \sqrt{m_{x,ij}^2 + m_{y,ij}^2},$$

where we mean $m_{x,nj} = 0$ for $j = 1, \ldots, n$ and $m_{y,in} = 0$ for $i = 1, \ldots, n$.

3.2 Chambolle–Pock

Write

$$S(\rho, \gamma) = \{\tilde{\rho} \in \mathbb{R}^{n \times n} \mid 0 \leq \tilde{\rho} \leq \rho,\ \langle \mathbf{1}, \rho \rangle = \gamma\},$$

and we can write the constraints of (8) on $\tilde{\rho}^0$ and $\tilde{\rho}^1$ as

$$\tilde{\rho}^0 \in S(\rho^0, \gamma)$$
$$\tilde{\rho}^1 \in S(\rho^1, \gamma).$$

Define the Lagrangian

$$L(\mathbf{m}, \tilde{\rho}^0, \tilde{\rho}^1, \Phi) = \|\mathbf{m}\|_{1,2} + \langle \Phi, \operatorname{div}(\mathbf{m}) + \tilde{\rho}^1 - \tilde{\rho}^0 \rangle,$$

where $\Phi \in \mathbb{R}^{n \times n}$ is the Lagrange multiplier corresponding to the equality constraint $\operatorname{div}(\mathbf{m}) = \tilde{\rho}^0 - \tilde{\rho}^1$ of (8). Again, $\langle \cdot, \cdot \rangle$ denotes the inner product between $n \times n$ matrices treated as vectors. i.e.,

$$\langle A, B \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} B_{ij}.$$

Standard convex analysis states that $(\mathbf{m}^\star, \tilde{\rho}^{0\star}, \tilde{\rho}^{1\star})$ is a solution to (8) if and only if there is a $\Phi^\star$ such that $(\mathbf{m}^\star, \tilde{\rho}^{0\star}, \tilde{\rho}^{1\star}, \Phi^\star)$ is a saddle point of $L(\mathbf{m}, \Phi)$ on [22]. In other words, we can solve (8) by solving the minimax problem

$$\underset{\substack{m_x \in \mathbb{R}^{(n-1)\times n} \\ m_y \in \mathbb{R}^{n \times (n-1)} \\ \tilde{\rho}^0 \in S(\rho^0, \gamma) \\ \tilde{\rho}^1 \in S(\rho^1, \gamma)}}{\text{minimize}} \ \underset{\Phi \in \mathbb{R}^{n \times n}}{\text{maximize}} \ L(\mathbf{m}, \tilde{\rho}^0, \tilde{\rho}^1, \Phi). \tag{9}$$

Saddle point problems, such as (9), can be solved with the first-order primal-dual method of Chambolle and Pock [5,21]:

$$\mathbf{m}^{k+1} = \underset{\mathbf{m}}{\operatorname{argmin}} \left\{ \|\mathbf{m}\|_{1,2} + \langle \Phi^k, \operatorname{div}(\mathbf{m}) \rangle + \frac{1}{2\mu} \|\mathbf{m} - \mathbf{m}^k\|_2^2 \right\}$$

$$\tilde{\rho}^{0,k+1} = \underset{\tilde{\rho} \in S(\rho^0, \gamma)}{\operatorname{argmin}} \left\{ -\langle \Phi^k, \tilde{\rho} \rangle + \frac{1}{2\nu} \|\tilde{\rho} - \tilde{\rho}^{0,k}\|_2^2 \right\}$$

$$\tilde{\rho}^{1,k+1} = \underset{\tilde{\rho} \in S(\rho^1, \gamma)}{\operatorname{argmin}} \left\{ +\langle \Phi^k, \tilde{\rho} \rangle + \frac{1}{2\nu} \|\tilde{\rho} - \tilde{\rho}^{1,k}\|_2^2 \right\}$$

$$v^{k+1} = \operatorname{div}(2\mathbf{m}^{k+1} - \mathbf{m}^k) + 2\tilde{\rho}^{1,k+1} - \tilde{\rho}^{1,k} - 2\tilde{\rho}^{0,k+1} + \tilde{\rho}^{0,k}$$
$$\Phi^{k+1} = \operatorname*{argmax}_{\Phi} \left\{ \langle \Phi, v^{k+1} \rangle - \frac{1}{2\tau} \|\Phi - \Phi^k\|_2^2 \right\} \tag{10}$$

where $\mu, \nu, \tau > 0$ are step sizes. The meaning of $\|\cdot\|_2^2$ is standard; for $\mathbf{m}$ it is

$$\|\mathbf{m} - \mathbf{m}^k\|_2^2 = \sum_{i=1}^{n-1} \sum_{j=1}^{n} \left(m_{x,ij} - m_{x,ij}^k \right)^2 + \sum_{i=1}^{n} \sum_{j=1}^{n-1} \left(m_{y,ij} - m_{y,ij}^k \right)^2,$$

for Φ it is

$$\|\Phi - \Phi^k\|_2^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\Phi_{ij} - \Phi_{ij}^k \right)^2,$$

and for $\tilde{\rho}^0$ and $\tilde{\rho}^1$ it is the same as it is for Φ. These steps can be interpreted as a gradient descent in the primal variable $\mathbf{m}$ and a gradient ascent in the dual variable Φ.

3.3 Subproblems

The optimization problems that define (10) have closed-form or semi-closed-form solutions, and these efficient solutions to the subproblems allow algorithm (10) as a whole to be efficient.

We first simplify the $\mathbf{m}$ update. We have

$$\begin{aligned}
&\operatorname*{argmin}_{\mathbf{m}} \left\{ \|\mathbf{m}\|_{1,2} + \langle \Phi^k, \nabla \cdot \mathbf{m} \rangle + \frac{1}{2\mu} \|\mathbf{m} - \mathbf{m}^k\|_2^2 \right\} \\
&= \operatorname*{argmin}_{\mathbf{m}} \left\{ \sum_{ij} \left(\|\mathbf{m}_{ij}\|_{1,2} + \frac{1}{\Delta x} \Phi_{ij}^k (m_{x,ij} - m_{x,(i-1)j} + m_{y,ij} - m_{y,i(j-1)}) \right. \right. \\
&\quad \left. \left. + \frac{1}{2\mu} \|\mathbf{m}_{ij} - \mathbf{m}_{ij}^k\|_2^2 \right) \right\} \\
&= \operatorname*{argmin}_{\mathbf{m}} \left\{ \sum_{ij} \left(\|\mathbf{m}_{ij}\|_{1,2} - (\nabla \Phi^k)_{ij}^T \mathbf{m}_{ij} + \frac{1}{2\mu} \|\mathbf{m}_{ij} - \mathbf{m}_{ij}^k\|_2^2 \right) \right\},
\end{aligned}$$

where again, all out-of-bounds indicies are interpreted as zeros. This minimization has a closed form solution, which can be written concisely with $\tilde{\mathbf{m}}$ and $\tilde{\nabla}$:

$$\tilde{\mathbf{m}}_{ij}^{k+1} = \operatorname{shrink}_2(\tilde{\mathbf{m}}_{ij}^k + \mu(\tilde{\nabla} \Phi^k)_{ij}, \mu)$$

for $i, j = 1, \ldots, n$. The shrink operator shrink_2 is defined as

$$\operatorname{shrink}_2(v, \mu) = \begin{cases} (1 - \mu / \|v\|_2) v & \text{for } \|v\|_2 \geq \mu \\ 0 & \text{for } \|v\|_2 < \mu \,. \end{cases}$$

Note that shrink_2 maps from $\mathbb{R}^2$ to $\mathbb{R}^2$, given a fixed μ.

Next, we simplify the Φ update. We have

$$\operatorname*{argmax}_{\Phi} \left\{ \left\langle \Phi, v^{k+1} \right\rangle - \frac{1}{2\tau} \|\Phi - \Phi^k\|_2^2 \right\} = \operatorname*{argmax}_{\Phi} \left\{ \sum_{ij} \left(\Phi_{ij} v_{ij}^{k+1} - \frac{1}{2\tau} (\Phi_{ij} - \Phi_{ij}^k)^2 \right) \right\},$$

and last line of (10) simplifies to

$$\Phi_{ij}^{k+1} = \Phi_{ij}^{k} + v_{ij}^{k+1}$$

for $i, j = 1, \ldots, n$.

The updates for $\tilde{\rho}^0$ and $\tilde{\rho}^1$ have a semi-closed-form solution. After reorganization, the second and third lines of (10) becomes

$$\tilde{\rho}^{0,k+1} = P_{S(\rho^0,\gamma)}\left(\tilde{\rho}^{0,k} + \nu\Phi^k\right)$$
$$\tilde{\rho}^{1,k+1} = P_{S(\rho^1,\gamma)}\left(\tilde{\rho}^{1,k} - \nu\Phi^k\right),$$

where $P_{S(\rho^0,\gamma)}$ and $P_{S(\rho^1,\gamma)}$ are the projections onto $S(\rho^0, \gamma)$ and $S(\rho^1, \gamma)$, respectively. We can evaluate these projections via the following algorithm.

Projection algorithm

Input: $\sigma \in \mathbb{R}^{n\times n}$, $\rho \in \mathbb{R}_+^{n\times n}$, $\gamma \in (0, \langle \mathbf{1}, \rho\rangle]$, and $\varepsilon > 0$
Output: $P_{S(\rho,\gamma)}(\sigma)$

$\theta_{\min} = -\max_{i,j}\{\rho_{ij}\} + \min_{i,j}\{\sigma_{ij}\}$
$\theta_{\max} = \max_{i,j}\{\sigma_{ij}\}$
while $\theta_{\max} - \theta_{\min} \geq \varepsilon$
  $\theta_{\text{mid}} = (\theta_{\max} + \theta_{\min})/2$
  $\tilde{\rho}(\theta_{\text{mid}}) = \min\{\max\{\sigma - \theta_{\text{mid}}\mathbf{1}, 0\}, \rho\}$
  if $\gamma < \langle \mathbf{1}, \tilde{\rho}(\theta_{\text{mid}})\rangle$
    $\theta_{\min} = \theta_{\text{mid}}$
  else
    $\theta_{\max} = \theta_{\text{mid}}$
  end
end
$\tilde{\rho}(\theta_{\text{mid}}) = \min\{\max\{\sigma - \theta_{\text{mid}}\mathbf{1}, 0\}, \rho\}$

To clarify, the min and max are taken element-wise and $\rho \in \mathbb{R}_+^{n\times n}$ means ρ is positive element-wise.

Lemma 1 *The projection algorithm computes* $P_{S(\rho,\gamma)}(\sigma)$.

Proof Consider the constrained convex optimization problem defining $P_{S(\rho,\gamma)}$:

$$\begin{array}{ll} \underset{\tilde{\rho}}{\text{minimize}} & \frac{1}{2}\|\tilde{\rho} - \sigma\|_2^2 \\ \text{subject to} & 0 \leq \tilde{\rho} \\ & \tilde{\rho} \leq \rho \\ & \gamma = \langle \mathbf{1}, \tilde{\rho}\rangle, \end{array}$$

Then Lagrangian for this optimization problem is

$$\mathcal{L}(\tilde{\rho}; \lambda_1, \lambda_2, \theta) = \frac{1}{2}\|\tilde{\rho} - \sigma\|_2^2 - \langle \lambda_1, \tilde{\rho}\rangle + \langle \lambda_2, \tilde{\rho} - \rho\rangle - \theta(\gamma - \langle \mathbf{1}, \tilde{\rho}\rangle),$$

where $\lambda_1, \lambda_2 \in \mathbb{R}^{n\times n}$ and $\theta \in \mathbb{R}$ are Lagrange multipliers. The KKT conditions for the optimization problem are

$$0 \leq \tilde{\rho}, \quad \tilde{\rho} \leq \rho \tag{11}$$

$$\gamma = \langle \mathbf{1}, \tilde{\rho} \rangle \tag{12}$$
$$\lambda_1 \geq 0, \quad \lambda_2 \geq 0 \tag{13}$$
$$\lambda_1^T \tilde{\rho} = 0, \quad \lambda_2^T (\rho - \tilde{\rho}) = 0 \tag{14}$$
$$0 = \tilde{\rho} - \sigma - \lambda_1 + \lambda_2 + \theta \mathbf{1} \tag{15}$$

where all inequalities are element-wise. Write

$$\begin{aligned}
\tilde{\rho}(\theta) &= \min\{\max\{\sigma - \theta \mathbf{1}, 0\}, \rho\} \\
\lambda_1(\theta) &= -\min\{\sigma - \theta \mathbf{1}, 0\} \\
\lambda_2(\theta) &= \max\{\sigma - \theta \mathbf{1} - \rho, 0\},
\end{aligned}$$

where the max and min are taken element-wise. For any θ, we see that $\tilde{\rho}(\theta)$, $\lambda_1(\theta)$, and $\lambda_2(\theta)$ satisfy the KKT conditions (11), (13), (14), and (15). So if we find a $\theta^\star$ such that $\tilde{\rho}(\theta^\star)$ satisfies the final KKT condition (12), $\tilde{\rho}(\theta^\star)$ is a solution to the projection problem.

Define

$$\begin{aligned}
\theta_{\min} &= -\max_{i,j}\{\rho_{ij}\} + \min_{i,j}\{\sigma_{ij}\} \\
\theta_{\max} &= \max_{i,j}\{\sigma_{ij}\}.
\end{aligned}$$

Then

$$\begin{aligned}
\tilde{\rho}(\theta_{\min}) &= \rho \\
\tilde{\rho}(\theta_{\max}) &= 0
\end{aligned}$$

and we have

$$\begin{aligned}
\langle \mathbf{1}, \tilde{\rho}(\theta_{\min}) \rangle &= \langle \mathbf{1}, \rho \rangle \\
\langle \mathbf{1}, \tilde{\rho}(\theta_{\max}) \rangle &= 0.
\end{aligned}$$

Since $\langle \mathbf{1}, \tilde{\rho}(\theta) \rangle$ is a non-increasing function of θ, we can use bisection to find the $\theta^\star$ that satisfies $\langle \mathbf{1}, \tilde{\rho}(\theta) \rangle = \gamma$. □

3.4 Main Algorithm

We are now ready to state the main algorithm.

First-order Method for Partial L_1 Monge–Kantorovich Problem

Input: Discrete probabilities ρ^0, ρ^1, and γ
Initial guesses $\tilde{\rho}^{0,0}$, $\tilde{\rho}^{1,0}$, $\mathbf{m}^0$, Φ^0 and step size μ, ν, τ
Output: Optimal $\tilde{\rho}^{0,\star}$, $\tilde{\rho}^{1,\star}$, and $\mathbf{m}^\star$

for $k = 1, 2, \cdots$ (Iterate until convergence)

$$\tilde{\mathbf{m}}_{ij}^{k+1} = \text{shrink}_2(\tilde{\mathbf{m}}_{ij}^k + \mu(\tilde{\nabla}\Phi^k)_{ij}, \mu) \qquad \text{for } i, j = 1, \ldots, n$$
$$\tilde{\rho}^{0,k+1} = P_{S(\rho^0,\gamma)}(\tilde{\rho}^{0,k} + \nu\Phi^k)$$
$$\tilde{\rho}^{1,k+1} = P_{S(\rho^1,\gamma)}(\tilde{\rho}^{1,k} - \nu\Phi^k)$$
$$\Phi_{ij}^{k+1} = \Phi_{ij}^k + \tau(\text{div}(2\mathbf{m}^{k+1} - \mathbf{m}^k)_{ij} + 2\tilde{\rho}_{ij}^{1,k+1} - \tilde{\rho}_{ij}^{1,k} - 2\tilde{\rho}_{ij}^{0,k+1} + \tilde{\rho}_{ij}^{0,k}) \qquad \text{for } i, j = 1, \ldots, n$$

end

When the problem is unbalanced, then one of the projection steps for $\tilde{\rho}^0$ or $\tilde{\rho}^1$ can be eliminated. When the problem is balanced, both projection steps can be eliminated and the algorithm reduces to that of [14].

3.5 Convergence Analysis

We now show that the proposed primal-dual algorithm converges to the minimizer of (8).

Define the discrete Laplacian operator as $\nabla^2 = \text{div} \cdot \nabla$.

Theorem 2 *Assume* $\mu\tau/(1-2\nu\tau) < 1/\lambda_{\max}(-\nabla^2)$, *where* $\lambda_{\max}(-\nabla^2)$ *denotes the largest eigenvalue of the negative discrete Laplacian operator* $-\nabla^2$. *Then with iterations* (10)

$$\left(\mathbf{m}^k, \tilde{\rho}^{0,k}, \tilde{\rho}^{1,k}, \Phi^k\right) \to \left(\mathbf{m}^\star, \tilde{\rho}^{0,\star}, \tilde{\rho}^{1,\star}, \Phi^\star\right)$$

where $(\mathbf{m}^\star, \tilde{\rho}^{0,\star}, \tilde{\rho}^{1,\star}, \Phi^\star)$ *is a saddle point of* L *in* (9). *Define*

$$\begin{aligned} R^k = {} & (1/\mu)\|\mathbf{m}^{k+1} - \mathbf{m}^k\|_2^2 + (1/\nu)\|\tilde{\rho}^{0,k+1} - \tilde{\rho}^{0,k}\|_2^2 + (1/\nu)\|\tilde{\rho}^{1,k+1} - \tilde{\rho}^{1,k}\|_2^2 \\ & + (1/\tau)\|\Phi^{k+1} - \Phi^k\|_2^2 \\ & - 2\langle \Phi^{k+1} - \Phi^k, \text{div}(\mathbf{m}^{k+1} - \mathbf{m}^k) + \tilde{\rho}^{1,k+1} - \tilde{\rho}^{1,k} - \tilde{\rho}^{0,k+1} + \tilde{\rho}^{0,k}\rangle. \end{aligned} \tag{16}$$

Then $R^k \geq 0$ *and* $R^k = 0$ *if and only if* $(\mathbf{m}^k, \tilde{\rho}^{0,k}, \tilde{\rho}^{1,k}, \Phi^k)$ *is a saddle point of of* (9). *Finally,* R^k *monotonically converges to* 0.

Proof We check the conditions required in [5,21]. Let us rewrite L by

$$L(\mathbf{m}, \tilde{\rho}^0, \tilde{\rho}^1, \Phi) = G(\mathbf{m}, \tilde{\rho}^0, \tilde{\rho}^1) + \Phi^T K(\mathbf{m}, \tilde{\rho}^0, \tilde{\rho}^1) - F(\Phi)\,,$$

where $G(\mathbf{m}, \tilde{\rho}^0, \tilde{\rho}^1) = \|\mathbf{m}\|_{1,2}$,

$$K = \begin{bmatrix} \text{div} & -I & I \end{bmatrix},$$

and $F(\Phi) = 0$. Observe that G, F are convex functions and K is a linear operator. By Lemma 1 of [21] and an application of the Schur complement, the algorithm converges for $\mu\tau/(1-2\nu\tau) < 1/\lambda_{\max}(-\nabla^2)$.

The Chambolle–Pock methods can be interpreted as a proximal point method under a certain metric [11]. R^k is the fixed-point residual of the non-expansive mapping defined by the proximal point method and thus decreases monotonically to 0, cf., review paper [24]. □

4 Computational Considerations

The proposed method can be parallelized to run efficiently on GPUs. The $\mathbf{m}$ and Φ updates can be split over the indices (i, j) as follows:

```
(Main algorithm)
m_temp[i,j] = m[i,j]
m[i,j] = shrink(m[i,j]+mu/dx*(Phi[i+1,j]-Phi[i,j],Phi[i,j+1]
         -Phi[i,j]))
m_temp[i,j] = 2*m[i,j]-m_temp[i,j]
-------------------------------------------------------------
Synchronize over all i,j
-------------------------------------------------------------
```

```
Perform rho0 and rho1 update

-------------------------------------------------------------
divm[i,j] = m_temp_x[i,j]-m_temp_x[i-1,j]+m_temp_y[i,j]
            -m_temp_y[i,j-1]
Phi[i,j] = Phi[i,j] + tau*(divm[i,j]/dx+rho1_temp[i,j]
            -rho0_temp[i,j])
-------------------------------------------------------------
Synchronize over all i,j
-------------------------------------------------------------
```

This algorithmic structure can effectively utilize the parallel computing capabilities of GPUs (and even more so when with the use of ghost cells).

The $\tilde{\rho}^0$ and $\tilde{\rho}^1$ updates are the computational bottleneck of the proposed algorithm and are trickier to program. The hard work for these updates is finding the max, min, and sum of an array.

```
(rho0 update)
theta_min = min(t_rho0) + rho0_max
theta_max = max(t_rho0)
while (...)
  ...
  if ( sum(min(max(t_rho0-nu*Phi,0),rho0) > gamma)
    ...
  ...
end
...
```

These operations can be done via *parallel reduction*, which can effectively utilize the parallel computing capabilities of GPUs. Implementing an efficient parallel reduction for a CUDA GPU can be tricky, as it requires specific knowledge of the CUDA computing architecture. See [16] for a tutorial on this topic.

The bisection intervals for the $\tilde{\rho}^0$ and $\tilde{\rho}^1$ can be improved to reduce the number of bisection iterations. One approach we use is to remember an interval from the last iteration and check if it is still valid. If so, the past interval is used. Otherwise the method falls back to the valid but wider interval specified in the statement of the projection algorithm.

We perform the $\mathbf{m}$, $\tilde{\rho}^0$, and $\tilde{\rho}^1$ updates in parallel to achieve better concurrency.

We can use R^k, defined as (16), as a termination criterion. Computing R^k also can be done with parallel reduction.

In choosing the parameters μ, ν, and τ Theorem 2 provides an upper bound for $\mu\tau/(1 - 2\nu\tau)$. It does not, however, provide any guidance for choosing the individual values for μ, ν, and τ. As they represent the step sizes for variables of different scales, μ, ν, and τ should not be constrained to be equal. Indeed, we have empirically observed that the values of μ, ν, and τ must be different to get the best convergence rate and that a poor choice of μ, ν, and τ can slow down the rate of convergence significantly.

Theorem 2, however, does suggest suggest $\nu = 1/(4\tau)$ and $\mu = 1/(32(n-1)^2\tau)$ are reasonable values to use for $\Omega = [0, 1] \times [0, 1]$ and $\Delta x = 1/(n-1)$, since $\lambda_{\max}(-\nabla_{\mathbf{x}}) \leq 8(n-1)^2$. This is what we use in Sect. 5, and we report the specific τ we use for the experiment.

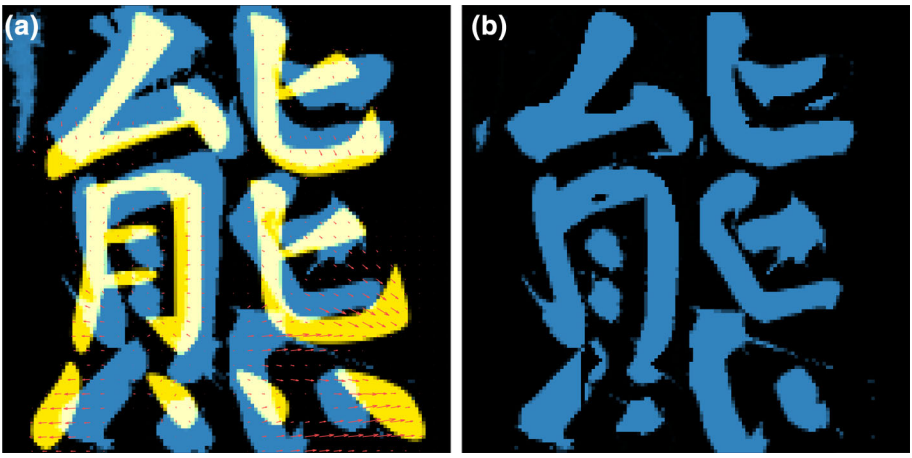

Fig. 1 Unbalanced problem: ρ^0 (blue) is a handwritten character with an ink spill on the top-left corner and has total mass 1; ρ^1 (yellow) is a computer font for the same character and has total mass 0.8. (The character means 'bear'.) We ran the method with $n = 128$, $\tau = 0.1$, and 2.5×10^5 iterations, which took 82.1 s time. $U(\rho^0, \rho^1) = 0.034$. **a** Transported mass. **b** The transported mass $\tilde{\rho}^0$. We can see that the ink spill on the top-left corner is ignored (Color figure online)

For the sake of scientific reproducibility, we release the code used for the experiments. For the convenience, we "mex"ed the CUDA code into a Matlab function.

5 Examples

In this section, we demonstrate several numerical results on $\Omega = [0, 1] \times [0, 1]$ with an $n \times n$ discretization. The initial values for $\mathbf{m}^0$ and Φ^0 are chosen as all zeros, and the initial values for $\tilde{\rho}^0$ and $\tilde{\rho}^1$ are chosen to be equal to ρ^0 and ρ^1. We implemented the method with CUDA C++ and ran it on a NVIDIA TITAN Xp graphics card. We describe the problem description and parameters in the figures' captions. For simplicity, we did not use the termination criterion R^k in these experiments; we simply ran the method up to a fixed iteration count. Rather, we demonstrate the convergence of R^k separately in Fig. 3 .

In Fig. 1, we consider the unbalanced L_1 Monge–Kantorovich problem between two Chinese characters. As a reference, the MNIST dataset uses 28×28 images of numbers to classify handwritten numerical digits [12]. However, Chinese characters are more complex, and 28×28 pixels are likely not enough to resolve the finer strokes and implement a character recognition algorithm. We use 128×128 pixels for Fig. 1.

In Fig. 1, more specifically, we transport a handwritten character with a ink spill to a computer font of the same character. The computer font is a reliable reference, while the handwritten character is noisy. So we let the reference character have mass 0.8, let the handwritten character have mass 1 and choose $\gamma = 0.8$. This way, the reference character must be fully filled, while 20% of the mass of the noisy handwritten character is ignored. In particular, the ink spill on the top left corner is ignored.

In Fig. 2, we solve a series of partial L_1 Monge–Kantorovich problems with different values of γ. Fig. 2a shows how $P_\gamma(\rho^0, \rho^1) = 0$ is possible when $\rho^0 \neq \rho^1$ if $\gamma < 1$. Fig. 2a–c entirely ignore the ink spill on the top-right corner, but Fig. 2d, which shows the

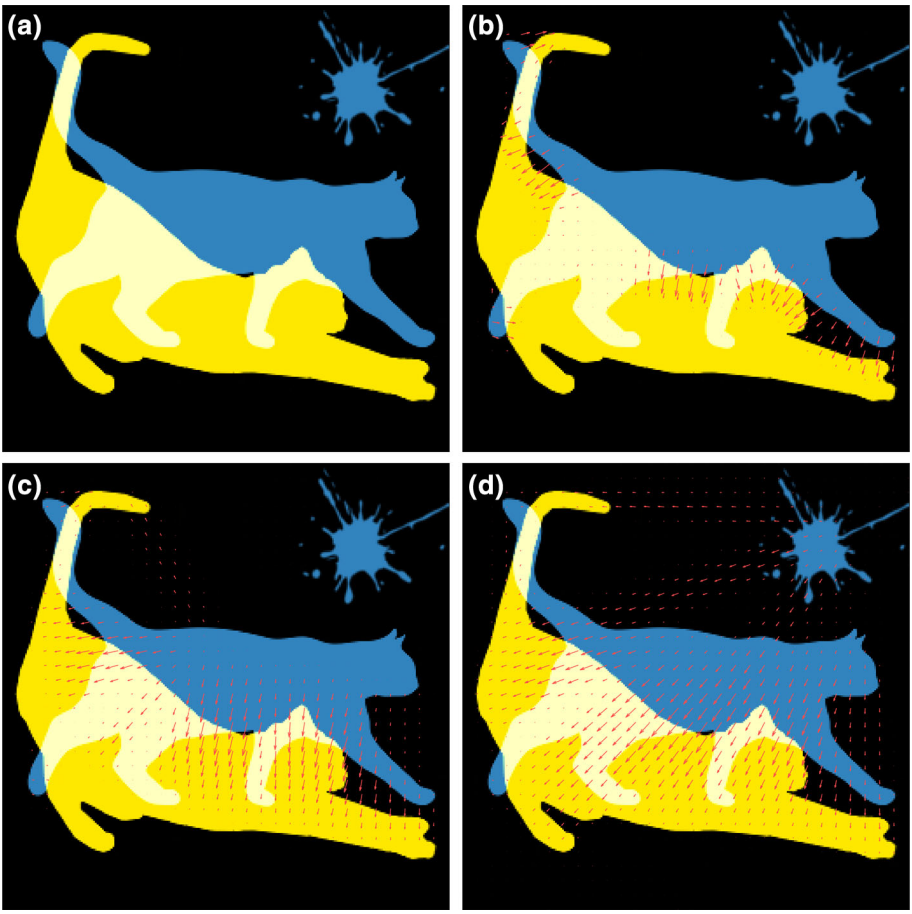

Fig. 2 Both ρ^0 (blue) and ρ^1 (yellow) are densities with mass 1. We ran the method with $n = 256$, $\tau = 0.01$, and 10^6 iterations, which took about 300 s for all 4 experiments. **a** $\gamma = 0.2$. No mass is transported as more than 0.2 mass overlaps. $P_\gamma(\rho^0, \rho^1) = 0$. **b** $\gamma = 0.5$. $P_\gamma(\rho^0, \rho^1) = 0.012$. **c** $\gamma = 0.8$. $P_\gamma(\rho^0, \rho^1) = 0.084$. **d** $\gamma = 1.0$. The entire mass is transported and the top-right ink spill is no longer ignored. $P_\gamma(\rho^0, \rho^1) = 0.197$ (Color figure online)

solution to the balanced, not partial, L_1 Monge–Kantorovich problem, cannot ignore the ink spill.

Figure 3 shows the convergence of the termination criterion R^k. The value of R^k decreases monotonically, as guaranteed by Theorem 2, until round-off errors become significant. We used single-precision floating-point numbers.

In Table 1, we show the rough number of iterations required for convergence. The setup is shown in Fig. 4. The circles of ρ^0 are centered at $(0, 0)$ and $(0, -1)$, and the circles of and ρ^1 are centered at $(-1, 1)$ and $(1, 1)$. So $P_\gamma(\rho^0, \rho^1)$ with $\gamma = 0.5$ should be roughly $1/\sqrt{2} \approx 0.71$. We roughly tuned the parameters μ, ν, and τ to get the best performance for each grid size. Finally, we ran the method until the computed $P_\gamma(\rho^0, \rho^1)$ was close enough to 0.71 and the flux looked good enough. The quantitative results are summarized in Table 1.

To the best of our knowledge, [1] is the only previous work that presents numerical methods for the partial L_1 Monge–Kantorovich problem. The performance of Table 1 is much faster than that of [1].

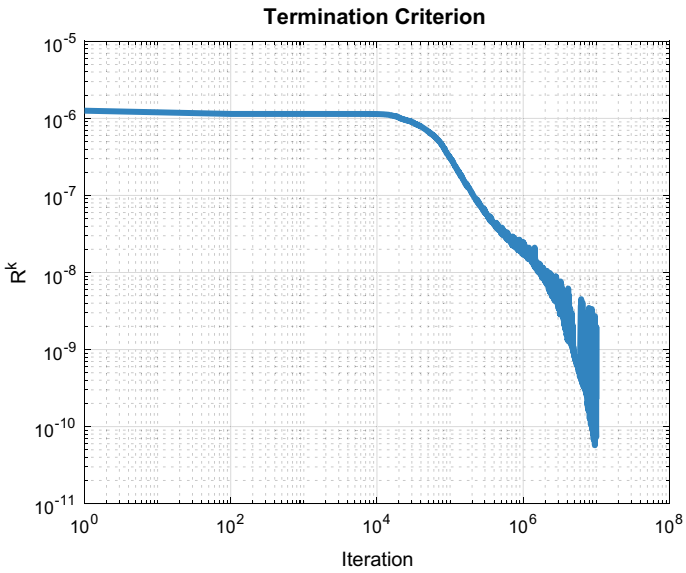

Fig. 3 Termination criterion R^k for the setup of Fig. 1 but with more iterations

Table 1 Run time as a function of grid size

Grids size	Run time	τ
32×32	21.8 s (0.5×10^5 iterations)	0.01
64×64	48.6 s (1×10^5 iterations)	0.01
128×128	148.5 s (2.5×10^5 iterations)	0.01
256×256	270.4 s (5×10^5 iterations)	0.01

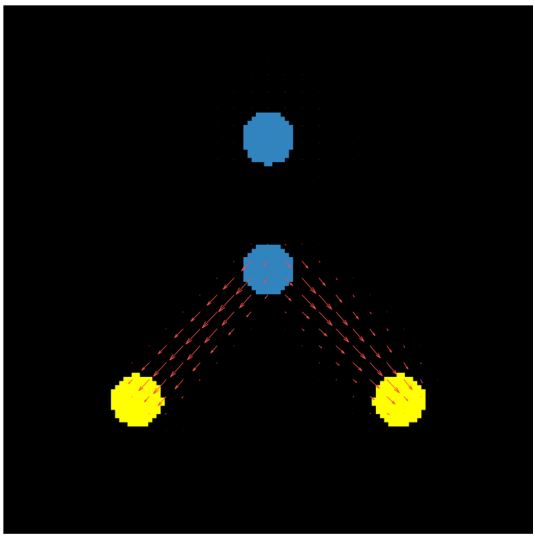

Fig. 4 A partial problem: ρ^0 (blue) and ρ^1 (yellow) of both have mass 1, and $\gamma = 0.5$. For Table 1, similar versions with $n = 32$, 64, 128, and 256 are used (Color figure online)

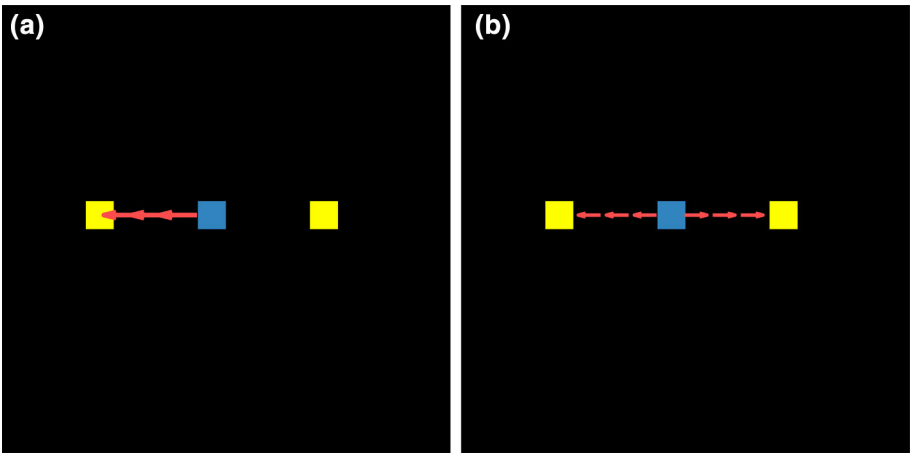

Fig. 5 An unbalanced 16×16 problem with ρ^0 (blue) of mass 0.5 and ρ^1 (yellow) of mass 1. Without regularization, the solution is not unique, and the left image shows one solution. **a** A solution to the unbalanced problem. **b** The unique solution to the regularized unbalanced problem (Color figure online)

6 Existence and Uniqueness

Solutions exist for both the continuous unbalanced and partial L_1 Monge–Kantorovich problems [1]. However, the solutions are not unique. To understand why, see the unbalanced problem of Fig. 5; the center delta mass of ρ^0 can be transported to either the left or right delta masses of ρ^1.

We can add regularization to remedy non-uniqueness. Instead of solving Problem (8), we can solve

$$\begin{aligned} \underset{\mathbf{m},\tilde{\rho}^0,\tilde{\rho}^1}{\text{minimize}} \quad & \|\mathbf{m}\|_{1,2} + (\varepsilon_1/2)\|\mathbf{m}\|_2^2 + (\varepsilon_2/2)\|\tilde{\rho}^2\|_2^2 + (\varepsilon_2/2)\|\tilde{\rho}^1\|_2^2 \\ \text{subject to} \quad & \operatorname{div}(\mathbf{m}) = \tilde{\rho}^0 - \tilde{\rho}^1 \\ & 0 \le \tilde{\rho}^0 \le \rho^0 \\ & 0 \le \tilde{\rho}^1 \le \rho^1 \\ & \gamma = \langle \mathbf{1}, \tilde{\rho}^0 \rangle = \langle \mathbf{1}, \tilde{\rho}^1 \rangle \end{aligned} \tag{17}$$

for some small $\varepsilon_1, \varepsilon_2 > 0$. The additional terms makes the objective strictly convex, and thereby make the solution unique.

We can solve the regularized problem (17) with only a slight modification to Algorithm (10). The $\mathbf{m}$, $\tilde{\rho}^0$, and $\tilde{\rho}^1$ updates change:

$$\begin{aligned} \mathbf{m}^{k+1} &= \underset{\mathbf{m}}{\operatorname{argmin}} \left\{ \|\mathbf{m}\|_{1,2} + (\varepsilon_1/2)\|\mathbf{m}\|_2^2 + \langle \Phi^k, \operatorname{div}(\mathbf{m}) \rangle + \frac{1}{2\mu}\|\mathbf{m} - \mathbf{m}^k\|_2^2 \right\} \\ \tilde{\rho}^{0,k+1} &= \underset{\tilde{\rho} \in S(\rho^0,\gamma)}{\operatorname{argmin}} \left\{ (\varepsilon_2/2)\|\tilde{\rho}\|_2^2 - \langle \Phi^k, \tilde{\rho} \rangle + \frac{1}{2\nu}\|\tilde{\rho} - \tilde{\rho}^{0,k}\|_2^2 \right\} \\ \tilde{\rho}^{1,k+1} &= \underset{\tilde{\rho} \in S(\rho^1,\gamma)}{\operatorname{argmin}} \left\{ (\varepsilon_2/2)\|\tilde{\rho}\|_2^2 + \langle \Phi^k, \tilde{\rho} \rangle + \frac{1}{2\nu}\|\tilde{\rho} - \tilde{\rho}^{1,k}\|_2^2 \right\} \end{aligned}$$

and

$$\tilde{\mathbf{m}}_{ij}^{k+1} = \frac{1}{1+\mu\varepsilon_1}\text{shrink}_2(\tilde{\mathbf{m}}_{ij}^{k} + \mu(\tilde{\nabla}\Phi^k)_{ij}, \mu)$$
$$\tilde{\rho}^{0,k+1} = P_{S(\rho^0,\gamma)}\left(1/(1+\varepsilon_2\nu)(\tilde{\rho}^{0,k} + \nu\Phi^k)\right)$$
$$\tilde{\rho}^{1,k+1} = P_{S(\rho^1,\gamma)}\left(1/(1+\varepsilon_2\nu)(\tilde{\rho}^{1,k} - \nu\Phi^k)\right)$$

The Φ update remains the same. See [14] for a discussion on a similar approach.

7 Conclusion

We proposed a scalable parallel primal-dual method algorithm to solve the unbalanced and partial L_1 Monge–Kantorovich problems. Our method leverages the structure of optimal transport, which converts the original problem into L_1-type minimization problem that is easier to discretize. We then apply the Chambolle–Pock primal-dual method [5,21] to obtain the main method. The subproblems of the proposed method have simple closed-form or semi-closed-form solutions, and we discuss their computational considerations, including how they can be parallelized to effectively utilize the computing power of a CUDA GPU. Finally, we provide numerical examples to demonstrate the method's effectiveness.

Possible interesting future directions include considering variations to the setup such as the L_2 Monge–Kantorovich problem [2] or the L_1 Monge–Kantorovich problem based a ground metric other than the Euclidean distance, such as the Manhattan distance [14,15]. Another interesting future direction is to explore the applications of unbalanced and partial L_1 Monge–Kantorovich problems in image processing.

Acknowledgements We would like to thank Professor Wilfrid Gangbo for many fruitful and inspirational discussions on the related topics. The Titan Xp used for this research was donated by the NVIDIA Corporation.

References

1. Barrett, J., Prigozhin, L.: Partial L^1 Monge–Kantorovich problem: variational formulation and numerical approximation. Interfaces Free Bound. **11**(2), 201–238 (2009)
2. Benamou, J.-D., Brenier, Y.: A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. Numer. Math. **84**(3), 375–393 (2000)
3. Benamou, J.-D., Carlier, G., Hatchi, R.: A numerical solution to Monge's problem with a Finsler distance as cost. M2AN (2016)
4. Caffarelli, L., McCann, R.: Free boundaries in optimal transport and Monge–Ampere obstacle problems. Ann. Math. **171**, 673–730 (2010)
5. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vis. **40**, 120–145 (2011)
6. Chizat, L., Peyre, G., Schmitzer, B., Vialard, F.-X.: Unbalanced optimal transport: geometry and Kantorovich formulation (2015). arXiv:1508.05216
7. Chizat, L., Schmitzer, B., Peyre, G., Vialard, F.-X.: An interpolating distance between optimal transport and Fischer-Rao. Found. Comput. Math. https://doi.org/10.1007/s10208-016-9331-y (2016)
8. Evans, L., Gangbo, W.: Differential equations methods for the Monge–Kantorovich mass transfer problem. Memoirs of AMS, no 653, vol. 137, (1999)
9. Figalli, A.: The optimal partial transport problem. Arch. Ration. Mech. Anal. **195**(2), 533560 (2010)
10. Hanin, L.G.: Kantorovich–Rubinstein norm and its application in the theory of Lipschitz spaces. Proc. Am. Math. Soc. **115**(2), 345–352 (1992)
11. He, B., Yuan, X.: Convergence analysis of primal-dual algorithms for a saddle-point problem: from contraction perspective. SIAM J. Imaging Sci. **5**(1), 119–149 (2012)

12. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)
13. Li, W.: A study of stochastic differential equations and Fokker–Planck equations with applications. Ph.D. thesis
14. Li, W., Ryu, E.K., Osher, S., Yin, W., Gangbo, W.: A parallel method for earth movers distance. J. Sci. Comput. https://doi.org/10.1007/s10915-017-0529-1 (2017)
15. Ling, H., Okada, K.: An efficient earth movers distance algorithm for robust histogram comparison. PAMI **29**, 840–853 (2007)
16. Luitjens, J.: Faster parallel reductions on Kepler. https://devblogs.nvidia.com/parallelforall/faster-parallel-reductions-kepler/. Accessed 15 July 2017
17. Métivier, L., Brossier, R., Mérigot, Q., Oudet, E., Virieux, J.: Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion. Geophys. J. Int. **205**(1), 345–377 (2016)
18. Piccoli, B., Rossi, F.: On properties of the generalized Wasserstein distance. Arch. Ration. Mech. Anal. **222**(3), 1339–1365 (2016)
19. Piccoli, B., Rossi, F.: Generalized Wasserstein distance and its application to transport equations with source. Arch. Ration. Mech. Anal. **211**(1), 335–358 (2014)
20. Pratelli, A.: Equivalence between some definitions for the optimal mass transport problem and for the transport density on manifolds. Ann. Mat. Pura Appl. **184**(2), 215–238 (2005)
21. Pock, T., Chambolle, A.: Diagonal preconditioning for first order primal-dual algorithms in convex optimization. In: International Conference on Computer Vision, IEEE, pp. 1762–1769 (2011)
22. Rockafellar, R.T.: Conjugate Duality and Optimization. Society for Industrial and Applied Mathematics, Philadelphia (1974)
23. Rubner, Y., Tomasi, C., Guibas, L.: The earth mover's distance as a metric for image retrieval. Int. J. Comput. Vis. **40**(2), 99–121 (2000)
24. Ryu, E.K., Boyd, S.: Primer on monotone operator methods. Appl. Comput. Math. **15**(1), 3–43 (2016)
25. Villani, C.: Topics in Optimal Transportation, vol. 58. American Mathematical Society, Providence (2003)
26. Yin, W., Osher, S., Goldfarb, D., Darbon, J.: Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing. SIAM J. Imaging Sci. **1**(1), 143–168 (2008)