Optimal Mass Transport and Its Applications

Xin Qi

Research Proficient Examination
Stony Brook University, Stony Brook, NY 11794, USA

Committee
Xianfeng David Gu, Jie Gao and Jing Chen

Abstract. Optimal Mass Transport Problem (OMT) was first introduced by Monge in 18th century [1], a problem concering about the optimal solution with minimal transportation cost to move a pile of objects from one place to another. The existence and uniqueness of the solution to relaxed OMT problems were proven by Kantorovich[2] based on linear programming methods, with a time complexity of $O(n^2)$. To increase the time efficiency, because of the amount of vertices in a high resolution 3D surface, our group introduced a practical optimal mass transport map based on Brenier's approach [3], to decrease the time complexity to $O(n)$.

Keywords: Optimal Mass Transport, Monge-Ampere Equation, Graphics Models, Parameterization

1 Introduction

1.1 Applications

2 Theoretical Background

In this section, we will review some basic mathematical theorems and results in comformal geometry and optimal mass transport.

2.1 Optimal Mass Transportation

Kantorovich and Monge Problems In the 18th cedntury, Monge first introduced a problem minimizing the inter-domain transportation cost while preserving measure quantities, which in modern language is shown as follows:

Problem 1 *Given two probability measures $\mu \in \mathbb{P}(X)$ and $\nu \in \mathbb{P}(Y)$, and a cost function $c : X \times Y \to [0, +\infty]$, solve*

$$\inf\{M(T) := \int c(x, T(x))d\mu(x) : T_{\#}\mu = \nu\}$$

where we recall that the measure denoted by $T_{\#}\mu$ is defined through $(T_{\#}\mu)(A) := \mu(T^{-1}(A))$ for every A, and is called image measure or fush-forward of μ through T.

However, because the constraint on T is not closed under weak convergence, this problem is difficult to solve and we will focus on the generalization proposed by Kantorovich[???]:

Problem 2 *Given $\mu \in \mathbb{P}(X), \nu \in \mathbb{P}(Y)$ and $c : X \times Y \to [0, +\infty)$ we consider the problem*

$$\inf\{K(\gamma) := \int_{X\times Y} c d\gamma : \gamma \in \Pi(\mu, \nu)\}$$

where $\Pi(\mu\nu)$ is the set of the so-called transport plans,

$$\Pi(\mu, \nu) = \{\gamma \in \mathbb{P}(X \times Y) : (\pi_x)_{\#}\gamma = \mu, (\pi_y)_{\#}\gamma = \nu\}$$

where π_x and π_y are the two projections of $X \times Y$ onto X and Y.

In our group, we have rewritten the problem in the following way: suppose X and Y are two metric spaces with probability measures μ and ν, and X and Y have equal total measures:

$$\int_X \mu = \int_Y \nu.$$

A map $T : X \to Y$ is *measure preserving* if for any measurable set $B \subset Y$, it satisfies

$$\mu(T^{(-1)}(B)) = \nu(B) \tag{1}$$

Denote transportation cost for sending $x \in X$ to $y \in Y$ by $c(x, y)$, then the total *transportation cost* is defined by

$$C(T) := \int_X c(x, T(x))d\mu(x). \tag{2}$$

Problem 3 *Given two metric spaces with probabilities measure $(X, \mu), (Y, \nu)$ with the transportation cost function $c : X \times Y \to \mathbb{R}$, the problem is to find the measure preserving map $T : X \to Y$ satisfying condition 1, which minimizes the transportation cost 2.*

The intrinsic connection between optimal mass transport map and convex geometry was discovered in late 1980's by Brenier[???] and we will follow this method to rewrite this problem as a convex geometry problem.

OMT and Convex Geometry If we take the measures in the section above as volumes as polyhedra(polyhedrons in 2D case), we will have a conclusion that optimal mass transportation map is a special area-preserving mapping, which is defined as follows in the language of differential geometry:

Definition 1 *(Area-preserving Mapping). Suppose $\phi : (S_1, \mathbf{g}_1) \to (S_2, \mathbf{g}_2)$ is a diffeomorphism, the pull back metric induced by ϕ on S_1 is $\phi^*\mathbf{g}_2$, if*

$$det(\mathbf{g}_1) = det(\phi^*\mathbf{g}_2)$$

then ϕ is an area-preserving mapping.

Minkowski Problem The classical Minkowski problem for convex polyhedron has influenced the development of convex geometry and differential geometry for decades. One version of the problem states as follows and shown in Figure 1:

Problem 4 ***(Minkowski problem for compact polytopes in*** $\mathbb{R}^n$.) *Suppose $n_1, n_2, ..., n_k$ are unit vectors which span $\mathbb{R}^n$ and $A_1, ..., A_k > 0$ so that $\sum_{i=1}^k A_i n_i = 0$. Find a compact convex polytope $P \subset \mathbb{R}^n$ with exactly k codimension-1 faces $F_1, ..., F_k$ so that n_i is the outward normal vector to F_i and the area of F_i is A_i.*

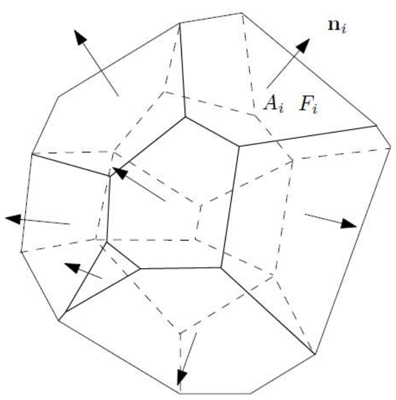

Fig. 1: Minkowski problem

Minkowski problem for unbounded convex polytopes was considered and solved by A.D.Alexandrov and A.Pogorelov. Infinite area unbounded faces are taken care of in the unbounded settings.

Theorem 1. *(Alexandrov) Suppose Ω is a compact convex polytope with non-empty interior in $\mathbb{R}^n$, $p_1, ..., p_k \subset \mathbb{R}^n$ are distinct k points and $A_1, ..., A_k > 0$ so that $\sum_{i=1}^k A_i = vol(\Omega)$. Then there exists a vector $h = (h_1, ..., h_k) \in \mathbb{R}^k$, unique up to adding the constant$(c, c, ..., c)$, sothat the piecewise linear convex function*

$$u(x) = \max_{1 \leq i \leq k} \{x \cdot p_i + h_i\}$$

satisfies $vol(\{x \in \Omega | \nabla u(x) = p_i\}) = A_i$

Alexandrov generallized Minkowski's result to non-compact convex polyhedra. As shown in Fig.2, given k planes $\pi_i :< x, p_i > +h_i$, one can construct a piecewise linear convex function

$$u(x) = \max_i \{< x, p_i > +h_i | i = 1, ..., k\}$$

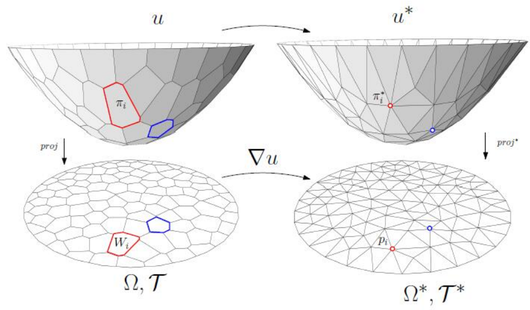

Fig. 2: A PL convex function induces a cell decomposition of Ω. Each cell is mapped to a point

whose graph is an infinite convex polyhedron. The PL convex function produces a convex cell decomposition $\{W_i\}$ of $\mathbb{R}^n$:

$$W_i = \{x| < x, p_i > +h_i \geq < x, p_j > +h_j, \forall j\} = \{x|\nabla u(x) = p_i\}$$

Alexandrov shows that the convex polyhedron is determined by the face normal, or equivalently the gradient $\{p_i\}$ and the projected area $\{A_i\}$.

Definition 2 *(Alexandrov map) The gradient map $\nabla u : x \to \nabla u(x)$ is the Alexandrov map.*

According to Monge-Brenier theory, the Alexandrov map is the unique Optimal Mass Transport map that minimizes the following mass transport energy

$$\int_{\Omega} ||x - f(x)||^2 dx$$

among all mass preserving maps $f : \Omega \to \{p_1, ..., p_k\}$, such that

$$Vol(f^{-1}(p_i)) = A_i.$$

The computation of the Alexandrov map is equivalent to computing the *power diagram* in computational geometry.

Power Diagram[???] The Power diagram is a generalization of Voronoi diagrams. For a finite set $M \subset E^d$, the *Voronoi diagram* of M associates each $p \in M$ with the convex region $R(p)$ of all points closer to p than to any other point in M: $R(p) = \{x \in E^d | d(x, p) < d(x, q), \forall q \in M - \{p\}\}$, where d denotes the Euclidean distance function. If we take the distance function(power function) to be $d(x, p)^2 - w(p)$ instead of the standard L_2 distance metric, then the Voronoi diagram will become a Power Voronoi diagram. In addition, we write the *power distance* from a point $x \in \mathbb{R}^2$ to p to be as follows:

$$POW(x, p_i) = \frac{1}{2}||x - p_i||^2 - \frac{1}{2}h_i$$

When h_i is positive, the intuitive meaning of the power distance is one half of the squred distance from x to the tangent point of x to the circle centered at p_i with radius $\sqrt{h_i}$. The Power Diagram is also a partition of the Euclidean plane into polygonal cells, although some sites may have empty power cells $\{W_i\}$, shown in Fig 3,

$$W_i = \{x | Pow(x, p_i) \leq Pow(x, p_j), \forall j\}$$

Comparing the W_is here and in the Alexandrov problem, it is obvious to see that computing a power diagram is equivalent to compute the Alexandrov map. The computation for discrete situation will be discussed in the next section.

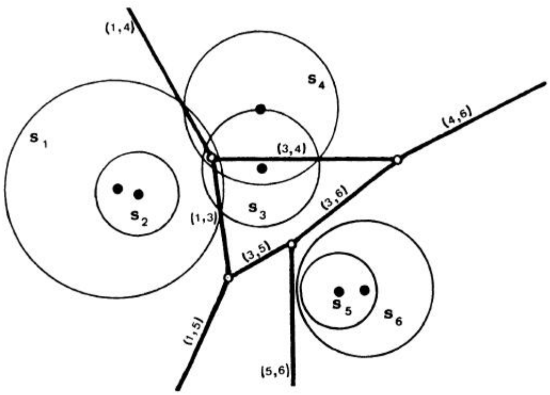

Fig. 3: Power diagram for 6 circles. A partition of the Euclidean plane.

Optimal Mass Transportation Map by Variational Principle The computation is based on the following *Generalized Alexandrov Theorem*:

Theorem 2. *Let Ω be a compact convex domain in $\mathbb{R}^n$ and $\{p_1, ..., p_k\}$ a set of distinct points in $\mathbb{R}^n$ and $\sigma : \Omega \to \mathbb{R}$ be a positive continuous function. Then for any $A_1, ..., A_k > 0$ with $\sum_{i=1}^k A_i = \int_\Omega \sigma(x)dx$, there exists $b = (b_1, ..., b_k) \in \mathbb{R}^k$, unique up to adding a constant $(c, ..., c)$, so that $\int_{W_i(b)\cap\Omega} \sigma(x)dx = A_i$ for all i. The vectors b are exactly minimum points of the convex function*

$$E(h) = \int_a^h \sum_{i=1}^k \int_{W_i(h)\cap\Omega} \sigma(x)dx dh_i - \sum_{i=1}^k h_i A_i$$

on the open convex set $H = \{h \in \mathbb{R}^k | vol(W_i(h) \cap \Omega) > 0, \ \forall\ i\}$. In fact, $E(h)$ restricted to $H_0 = H \cap \{h | \sum_{i=1}^k h_i = 0\}$ is strictly convex. Furthermore, ∇u_b minimizes the quadratic cost $\int_\Omega |x - T(x)|^2 \sigma dx$ among all transport maps $T : (\Omega, \sigma dx) \to (\mathbb{R}^n, \sum_{i=1}^k A_i \delta_{p_i})$

In 2D cases, we take σ to be the measure density $\rho : \Omega \to \mathbb{R}$, and A_i to be discrete measures $\mu = \{\mu_1, ..., \mu_k\}$. The computation on 2D is based on power diagram and power triangulation. Suppose, two voronoi cells $W_i(\mathbf{h}), W_j(\mathbf{h})$ are adjacent and share a common edge e_{ij}. The edge e_{ij} has a dual Delaunay edge $\bar{e}_{ij}$. The norm with respect to ρ is defined as

$$|e|_\rho = \int_3 \rho(x)dx$$

and $|e|$ is just the traditional Euclidean length. By direct computation, we have

$$\frac{\partial \omega_i}{\partial h_j} = \frac{\partial \omega_j}{\partial h_i} = \frac{|e_{ij}|_\rho}{|\bar{e}_{ij}|}$$

Therefore, the differential 1-form $\omega = \sum_{i=1}^k \omega_i dh_i$ is a closed 1-form, satisfying $d\omega = 0$. By Brunn-Minkowski inequality[???], the admissible space H is convex and $H \neq \phi$. In addition, $E(h)$ is convex in h and satisfies

$$\frac{\partial E(h)}{\partial h_i} = \int_{W_i(h)\cap\Omega} \sigma(x)dx.$$

The Hessian matrix of E is defined as follows:

$$\frac{\partial^2 E}{\partial h_i \partial h_j} = \frac{\partial \omega_i}{\partial h_j} = \frac{|e_{ij}|_\rho}{|\bar{e}_{ij}|}$$

and the diagonal elements are given by

$$\frac{\partial \omega_i}{\partial h_i} = -\sum_{i\neq j} \frac{\partial \omega_i}{\partial h_j}$$

the negative Hessian matrix is diagonal dominant, so E is concave on admissible space H. Therefore the gradient map ∇E is a diffeomorphism.

The information above gives a guideline of proof[???]. Some detailed info in how to obtain an Alexandrov map will be given in the following sections where computational algorithms are discussed.

Discrete Monge-Ampere equation (DMAE) Closely related to the optimal transport problem is the Monge-Ampere equation(MAE). Let Ω be a compact domain in $\mathbb{R}^n$, $g : \partial\Omega \to \mathbb{R}$ and $A : \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ be given. Then the Dirichlet problem for MAE is to find a function $w : \Omega \to \mathbb{R}$ so that

$$\begin{cases} det(Hess(w))(x) = A(x, w(x), \nabla w(x)) \\ w|_{\partial\Omega} = g \end{cases} \tag{3}$$

where $Hess(w)$ is the Hessian matrix of w. We are only interested in solving a discrete MAE in the simplest setting where $A(x, w(x), \nabla w(x)) = A(x) : \Omega \to \mathbb{R}$ so that $A(\Omega)$ is a finite set.

We can define the discrete Hessian determinant for piecewise linear function as following:

Definition 3 *Suppose*$(X,\mathscr{T})$ *is a domain in* $\mathbb{R}^n$ *with a convex cell decomposition* $\mathscr{T}$ *and* $w: X \to \mathbb{R}$ *is a convex function linear on each cell. Then the discrete Hessian determinant of* w *assigns each vertex* v *of* $\mathscr{T}$ *the volume of the convex hull of the gradients of the* w *at top-dimensional cells adjacent to* v.

Now, we will have the following Dirichlet problem:

Problem 5 ***(Dirichlet problem for discrete MAE)*** *Suppose* $\Omega = conv(v_1, ..., v_m)$ *is the convex hull of* $v_1, ..., v_m$ *in* $\mathbb{R}^n$. *Let* $p_1, ..., p_k$ *be in* $int(\Omega)$. *Given any* $g_1, ..., g_m \in \mathbb{R}$ *and* $A_1, ..., A_k > 0$, *find a convex subdivision* $\mathscr{T}$ *of* Ω *with vertices exactly* $\{v_1, ..., v_m, p_1, ..., p_k\}$ *and a PL convex function* $w : \Omega \to \mathbb{R}$ *linear on each cell of* $\mathscr{T}$ *so that*

(a) (Discrete Monge-Ampere Equation) the discrete Hessian determinant of w *at* p_i *is* A_i,
(b) (Dirichlet condition)$w(v_i) = g_i$.

Then, we have the following theorem for the solution to the equation above:

Theorem 3. *Suppose* $\Omega = conv(v_1, ..., v_m)$ *is an n-dimensional compact convex polytope in* $\mathbb{R}^n$ *so that* $v_i \notin conv(v_1, ..., v_{i-1}, v_{i+1}, ..., v_k)$ *for all* i *and* $p_1, ..., p_k$ *are in the interior of* Ω. *For any* $g_1, ..., g_k \in \mathbb{R}$ *and* $A_1, ..., A_k > 0$, *there exists a convex cell decomposition* $\mathscr{T}$ *having* v_i *and* p_j *as vertices and a piecewise linear convex function* $w : (\Omega, \mathscr{T}) \to \mathbb{R}$ *so that* $w(v_i) = g_i, i = 1, ..., m$ *and the discrete Hessian determinant of* w *at* p_j *is* A_j, $j = 1, ..., k$. *In fact, the solution* w *is the Legendre dual of* $\max\{x \cdot p_j + h_j, x \cdot v_i - g_i | j = 1, ..., k, i = 1, ..., m\}$ *and* h *is the unique minimal point of a strictly convex function.*

The following theorem presents the link of Optimal mass transport and the Monge-Ampere Equation[4]:

Theorem 4. *Let* μ *and* ν *be two compactly supported probability measures on* $\mathbb{R}^n$. *If* μ *is absolutely continuous with respect to the Lebesgue measure, then*

i. there exists a unique solution T *to the optimal transport problem with cost* $c(x, y) = |x - y|^2/2$;
ii. there exists a convex function $u : \mathbb{R}^n \to \mathbb{R}$ *such that the optimal map* T *is given by* $T(x) = \nabla u(x)$ *for* $\mu - a.e.\ x$

Furthermore, if $\mu(dx) = f(x)dx$ *and* $\nu(dy) = g(y)dy$, *then* T *is differentiable* $\mu - a.e.$ *and*

$$|det(\nabla T(X))| = \frac{f(x)}{g(T(x))} \quad for\ \mu - a.e.\ x \in \mathbb{R}^n.$$

In addition, the cost $|x - y|^2/2$ is equivalent to the cost $-x \cdot y$. For any transport map S, since we have the condition $S_{\#}\mu = \nu$, we have

$$\int_{\mathbb{R}^n} \frac{|S(x)|^2}{2} d\mu(x) = \int_{\mathbb{R}^n} \frac{|y|^2}{2} d\nu(y),$$

hence,

$$
\begin{aligned}
&\int_{\mathbb{R}^n} \frac{|x-S(x)|^2}{2} d\mu(x) \\
&= \int_{\mathbb{R}^n} \frac{|x|^2}{2} d\mu(x) + \int_{\mathbb{R}^n} \frac{|S(x)|^2}{2} d\mu(x) + \int_{\mathbb{R}^n} (-x \cdot S(x)) d\mu(x) \\
&= \int_{\mathbb{R}^n} \frac{|x|^2}{2} d\mu(x) + \int_{\mathbb{R}^n} \frac{|y|^2}{2} d\nu(y) + \int_{\mathbb{R}^n} (-x \cdot S(x)) d\mu(x)
\end{aligned}
$$

So, the two minimization problems

$$\min_{S_{\#}\mu=\nu} \int_{\mathbb{R}^n} \frac{|x-S(x)|^2}{2} d\mu(x)$$

and

$$\min_{S_{\#}\mu=\nu} \int_{\mathbb{R}^n} (-x \cdot S(x)) d\mu(x)$$

are equivalent.

Shape Distance Given a Riemannian surface $(S, \mathbf{g})$, we can compute a Riemann mapping $\phi : (S, \mathbf{g}) \to (\mathbb{D}, dzd\bar{z})$. Assume the conformal factor function is $\lambda : S \to \mathbb{R}$, such that

$$\mathbf{g} \circ \phi^{-1}(z) = e^{2\lambda(z)} dzd\bar{z}$$

Suppose the total area is π, such that

$$\int_{\mathbb{D}} e^{2\lambda(z)} dxdy = \pi,$$

we can find the unique optimal transportation map(Alexandrov map)

$$\tau : (\mathbb{D}, e^{2\lambda} dzd\bar{z}) \to (\mathbb{D}, d\omega d\bar{\omega}),$$

τ can be represented as a complex-valued function defined on the unit disk.

Definition 4 *(Shape Definition). Given two Riemannian surfaces, which are topological disks, $(S_1, \mathbf{g}_1)$ and $(S_2, \mathbf{g}_2)$, the Riemann mappings are ϕ_k, $k = 1, 2$ respectively. Let $\eta \in Mob(\mathbb{D})$ be a Mobius transformation, where $Mob(\mathbb{D})$ is the Mobius transformation group of the unit planar disk, then $\eta_k \circ \phi_k$ are still Riemann mappings. Each Riemann mapping $\eta_k \circ \phi_k$ determines a unique optimal transportation map $\tau_k(\phi_k, \eta_k)$. Then the distance between two surfaces is given by*

$$d(S_1, S_2) := \min_{\eta_1, \eta_2 \in Mob(\mathbb{D})} \int_{\mathbb{D}} |\tau_1(\phi_1, \eta_1) - \tau_2(\phi_2, \eta_2)|^2 dxdy$$

We will show that the Riemann mapping ϕ and the optimal transportation map η encodes all the Riemannian metric information of the original surface by the following lemma:

Lemma 1 *Suppose a Riemannian surface $(S, \mathbf{g})$ with a total area π, which is a topological disk, the Riemann mapping is $\phi : (S, \mathbf{g}) \to (\mathbb{D}, dzd\bar{z})$, the conformal factor induced by ϕ is $\lambda : S \to \mathbb{R}$, the optimal transportation map is $\eta : (\mathbb{D}, e^{2\lambda(z)dzd\bar{z}}) \to (\mathbb{D}, dzd\bar{z})$, then the Riemannian metric of the original surface is given by*

$$\mathbf{g} \circ \phi^{-1}(z) = det(J_\eta)dzd\bar{z}.$$

Proof. Because $\phi : (S, \mathbf{g}) \to (\mathbb{D}, dzd\bar{z})$ is conformal, according to the equation $\phi^*\mathbf{g}_2 = e^{2\lambda}\mathbf{g}_1$, we have

$$\mathbf{g} \circ \phi^{-1}(z) = e^{2\lambda(z)dzd\bar{z}}.$$

Then, since $\eta : (\mathbb{D}, e^{2\lambda(z)dzd\bar{z}}) \to (\mathbb{D}, dzd\bar{z})$ is an optimal transportation map, it is area-preserving. When we apply the definition of area-preserving mapping, $det(\mathbf{g}_1) = det(\phi^*\mathbf{g}_2)$, we have the following

$$e^{2\lambda} = det(J_\eta)$$

Therefore, we have proved the lemma 1.

Wasserstein Metric Space Again, we will suppose $(M, \mathbf{g})$ is a Riemannian manifold with a Riemannian metric $\mathbf{g}$. we will focus on the set:

$$P_p(M) := \{\mu \in P(M) : \int |x|^p d\mu < +\infty\}.$$

For $\mu, \nu \in P_p(M)$, we will define

$$W_p(\mu, \nu) := \inf_{T\#\mu=\nu} (\int_M d(x, T(x))^p d\mu(x))^{\frac{1}{p}}.$$

Then, The quantity W_p defined above is a distance over $P_p(M)$.[???]

Definition 5 *(Wasserstein Space). Let $P_p(M)$ denote the space of all probability measures μ on M with finite p^{th} moment, where $p \geq 1$. Suppose there exists some point $x_0 \in M$ that $\int_M d(x, x_0)^p d\mu(x) < +\infty$, where d is the geodesic distance induced by $\mathbf{g}$. We define the Wasserstein space of order p as the space $P_p(M)$, endowed with the distance W_p, written as $\mathbb{W}_p$.*

The following theorem is fundamental for current research.

Theorem 5. *The Wasserstein distance W_p is a Riemannian metric of the Wasserstein space $\mathbb{W}_p$*

Conformal Wasserstein Shape Space We now can construct a shape space framework based on optimal transportation and conformal mapping theories.

We consider all oriented metric surfaces $(M, \mathbf{g})$ with the disk topology, namely M is of genus 0 and with a single boundary ∂M. The two markers $\{p, q\} \subset M$, p is an interior point, q is a boundary point. We call $(M, \mathbf{g}, p, q)$ as a *marked metric surface.* All marked metric surfaces construct a set $\mathscr{M}$, $\mathscr{M}$:=$\{$*marked metric surfaces*$\}$.

We say two marked metric surfaces are equivalent, if there exists a *normalized isometric diffeomorphism* $\phi : (M_1, \mathbf{g}_1, p_1, q_1) \to (M_2, \mathbf{g}_2, p_2, q_2)$, such that ϕ preserves metrics $\phi^*\mathbf{g}_2 = \mathbf{g}_1$ and preserves markers $\phi(p_1) = p_2, \phi(q_1) = (1_2)$. The product of the normalized isometry diffeomorphism group and the scaling group is denoted as G, $G := \{\textit{normalized isometries}\} \oplus \{\textit{scaling}\}$.

Now, we can define the shape space to be a quotient space denoted as

$$S := \mathscr{M}/G.$$

In the following discussion, we always omit the markers $\{p, q\}$ in the normalized marked metric surface $(M, \mathbf{g}, p, q) \in S$ and assume the total area is π. According to *Riemann mapping theorem*, there is a unique conformal mapping $\phi : M \to \mathbb{D}$ with Euclidean metric $dx^2 + dy^2$, such that $\phi(p) = (0, 0)$ and $\phi(q) = (1, 0)$. Then $\mathbf{g} = e^{2\lambda(x,y)}(dx^2 + dy^2)$. ϕ push forward the area element on $(M, \mathbf{g})$ to the disk

$$\mu_{(M,g)} := e^{2\lambda(x,y)} dx \wedge dy.$$

Then we have an injective mapping $\Gamma : S \to \mathbb{W}_2(\mathbb{D})$, $\Gamma : (M, \mathbf{g}) \mapsto \mu_{(M,\mathbf{g})}$. The Wasserstein metric on the Wasserstein space $\mathbb{W}_2(\mathbb{D})$ is pulled back to S,

$$d_S((M_1, \mathbf{g}_1), (M_2, \mathbf{g}_2)) := W_2(\mu(M_1, \mathbf{g}_1), \mu(M_2, \mathbf{g}_2)).$$

We call the metric space (S, d_S) as the *conformal Wasserstein shape space.*

3 Computational Algorithms In Discrete Settings

In this section, we presents the algorithmic implementation details for conformal mapping and optimal mass transport map(OMT-Map) generation. Based on the OMT-Map algorithm, we introduce surface area-preserving parameterization algorithm on simply connected surfaces, and the computation for conformal Wasserstein distance between surfaces.

In discrete settings, surfaces are represented as discrete polyhedral patches. Suppose S is a topological surface, V is a set of points on S, (S, V) is called a *marked surface.* T is a triangulation of S, whose vertices are in V, then (S, T) is called a *triangular mesh.* In the following discussion, we use V, E and F to represent the set of *vertices, edges, and faces.* A piecewise linear Riemannian metric (PL metric) on (S, V) is a flat cone metric, whose cone points are in V, represented by edge lengths.

3.1 Conformal Mapping

Definition 6 *(Discrete Riemannian Metric). A discrete metric on a triangular mesh (S,T) is a function defined on the edges $d : E \to \mathbb{R}^+$, which satisfies the triangle inequality, on a face $[v_i, v_j, v_k]$,*

$$d_{ij} + d_{jk} > d_{ki}; d_{ki} + d_{ij} > d_{jk}; d_{ik} + d_{kj} > d_{ji}.$$

Definition 7 *(Discrete Gauss Curvature). The discrete Gauss curvature function on a mesh is defined on vertices, $K : V \to \mathbb{R}$, such that*

$$K(v) = \begin{cases} 2\pi - \sum_i \theta_i, & v \notin \partial S \\ \pi - \sum_i \theta_i, & v \in \partial S \end{cases}$$

where θ_i's are corner angles adjacent to the vertex v, and ∂S represents the boundary of the mesh.

Gauss-Bonnet theorem still holds on discrete surface, shown as follows:

$$\sum_i K(v_i) = 2\pi\chi(S)$$

where $\chi(S)$ is the Euler characteristic of S.

Usually, we choose the following triangulation to guarantee the quality of the discrete mesh.

Definition 8 *(Delaunay Triangulation). Aclosed discrete surface (S,T) with a discrete metric d, we say a triangulation T is Delaunay, if for any edge $[v_i, v_j]$ adjacent to two faces $[v_i, v_j, v_k]$ and $[v_j, v_i, v_l]$,*

$$\theta_k^{ij} + \theta_l^{ji} \leq \pi,$$

where θ_k^{ij} is the corner angle at v_k in $[v_i, v_j, v_k]$, and θ_l^{ji} is the angle at v_l in $[v_j, v_i, v_l]$.

Discrete Surface Yamabe flow The first algorithm presented here is the *discrete surface Yamabe flow* algorithm. We define the discrete conformal factor function as $u : V \to \mathbb{R}$, and conformal structure coefficient on edges $\eta : E \to \mathbb{R}^+$.

Definition 9 *(Discrete Yamabe Flow with Surgery). Given a surface (S,V) with a discrete metric d, given a target curvature function $\bar{K} : V \to \mathbb{R}$, $\bar{K}(v_i) \in (-\infty, 2\pi)$, and the total target curvature satisfies Gauss-Bonnet formula, the discrete Yamabe flow is defined as*

$$\frac{du(v_i)}{dt} = \bar{K}(v_i) - K(v_i),$$

under the constraint $\sum_{v_i \in V} u(v_i) = 0$. During the flow, the triangulation on (S,V) is updated to be Delaunay with respect to $d(t)$, for all time t.

The following theorem guarantees the existence of the solution to the Yamabe flow.

Theorem 6. *Suppose (S, V) is a closed connected surface and d is any discrete metric on (S, V). Then for any $\bar{K} : V \to (-\infty, 2\pi)$ satisfying Gauss-Bonnet formula, there exists a discrete metric $\bar{d}$, unique up to a scaling on (S, V), so that $\bar{d}$ is discrete conformal to d and the discrete curvature of $\bar{d}$ is $\bar{K}$. Furthermore, the $\bar{d}$ can be obtained by discrete Yamabe flow with surgery.*

It is known that that Yamabe flow is the negative gradient flow of the following Yamabe energy:

$$f(u_1, u_2, ..., u_n) = \int^{(u_1,u_2,...,u_n)} \sum_{v_i \in V} (\bar{K}(v_i) - K(v_i))du_i$$

The gradient of the energy is $\nabla f(u_1, u_2, ..., u_n) = (\bar{K}(1) - K(1), ..., \bar{K}(n) - K(n))^T$. The Hessian matrix is formulated as $H = (h_{ij})$, satisfying

$$h_{ij} = \begin{cases} -w_{ij} & v_i \sim v_j \; i \neq j \\ 0 & v_i \nsim v_j \; i \neq j \\ \sum_k w_{ik} & i = j \end{cases}$$

where w_{ij} is the cotangent edge weight defined as

$$w_{ij} := \begin{cases} cot\theta_k^{ij} + cot\theta_l^{ji} & [v_i, v_j] \notin \partial S \\ cot\theta_k^{ij} & [v_i, v_j] \in \partial S \end{cases}$$

Newton's method is used when optimizing the Yamabe energy to compute the conformal metric with prescribed curvature.

3.2 Discrete Optimal Mass Transport

Now, we will present the Discrete version of the Optimal Mass Transport theory.

Kantorovich's Approach. The main idea of Kantorovich's approach is to simplify the OMT problem to a linear programming problem with n^2 unknowns.

Suppose space X and Y are discretized to sample points, $X = \{x_1, x_2, ..., x_n\}$, $Y = \{y_1, y_2, ..., y_n\}$, the measures are dirac measures

$$\mu = \sum_{i=1}^{n} \mu_i \delta(x - x_i), \nu = \sum_{j=1}^{n} \nu_i \delta(y - y_i),$$

the transport plan f is represented as a matrix (f_{ij}), such that

$$\sum_{j=1}^{n} f_{ij} = 1, \sum_{i=1}^{n} f_{ij} = 1, f_{ij} \geq 0.$$

Algorithm 1 Discrete Surface Yamabe Flow

Require: The inputs include:

1. A triangular mesh Σ, embedded in $\mathbb{E}^3$
2. A target curvature $\bar{K}$, satisfying Gauss-Bonnet formula.

Ensure: : A discrete metric conformal to the original one, satisfying the target curvature $\bar{K}$.

1: Initialize the discrete conformal factor u as 0 and conformal structure coefficient η, such that $\eta(e)$ equals to the initial edge length of e.
2: **while** $\max_i |\bar{K}_i - K_i| > \epsilon$ **do**
3: compute the edge length from γ and η
4: Update the triangulation to be Delaunay using diagonal edge swap for each pair of adjacent faces
5: Compute the corner angle θ_i^{jk} from the edge length using cosine law
6: Compute the vertex curvature K
7: Compute the Hessian matrix H
8: Solve linear system $H\delta u = \bar{K} - K$
9: Update conformal factor $u \leftarrow u - \delta u$
10: **end while**
11: Output the result circle packing metric

All such matrices form a convex polytope, and the total transportation cost becomes a linear function

$$C(f) = \sum_{i,j} c(x_i, y_i) f_{ij}.$$

Obviously, this linear programming problem has n^2 unknowns f_{ij}.

Brenier's Approach Suppose μ has a compact support on X, define the domain

$$\Omega = supp\ \mu = x \in X|\mu(x) > 0,$$

assume Ω is a convex domain in X. The space Y is discretized to $Y = \{y_1, y_2, ..., y_k\}$ with Dirac measure $\nu = \sum_{j=1}^{k} \nu_j \delta(y - y_j)$.

We will define a height vector $\mathbf{h} = (h_1, h_2, ..., h_k)$ and for each $y_i \in Y$, we define a hyperplane defined on X,

$$\pi_i(\mathbf{h}) :< x, y_i > +h_i = 0.$$

Define a convex function as following:

$$u_{\mathbf{h}}(x) = \max_{i=1}^{k}\{< x, y_i > +h_i\}$$

We denote its graph by $G(\mathbf{h})$, which is an infinite convex polyhedron with supporting planes $\pi_i(\mathbf{h})$. The projection of $G(\mathbf{h})$ induces a polygonal partition of Ω,

$$\Omega = \cup_{i=1}^{k} W_i(\mathbf{h})$$

where each cell $W_i(\mathbf{h})$ is defined as the projection of a facet of the convex polyhedron $G(\mathbf{h})$ onto Ω, $W_i(\mathbf{h}) = \{x \in X | u_{\mathbf{h}}(x) =< x, y_i > +h_i\} \cap \Omega$. And the area of W_i is given by $w_i(\mathbf{h}) = \int_{W_i}(\mathbf{h})\mu(x)dx$. Since the convex function $u_{\mathbf{h}}$ on each cell $W_i(\mathbf{h})$ is a linear function $\pi_i(\mathbf{h})$, the gradient map

$$grad\ u_{\mathbf{h}} : W_i(\mathbf{h}) \to y_i,\ i = 1, 2, ..., k.$$

maps each $W_i(\mathbf{h})$ to a single point y_i.

Using the formular and definition above, we will present a theorem that plays a fundamental role for discrete optimal mass transport theory:

Theorem 7. *For any given measure ν, such that*

$$\sum_{j=1}^{n} \nu_j = \int_{\Omega} \mu, \nu_j > 0,$$

there must exist a height vector $\mathbf{h}$ unique up to adding a constant vector $(c, c, ..., c)$, the convex function $u_{\mathbf{h}}(x)$ induces the cell decomposition of $\Omega = \cup_{i=1}^{k} W_i(\mathbf{h})$, such that the folloing area-preserving constraints are satisfied for all cells,

$$\int_{W_i(\mathbf{h})} = \nu_i,\ i = 1, 2, ..., n.$$

Furthermore, the gradient map grad $u_{\mathbf{h}}$ optimizes the transportation cost

$$C(T) := \sum_{\Omega} |x - T(x)|^2 \mu(x)dx.$$

Recently, Gu [???] gives a novel proof for the existence and uniqueness based on the variational principle for the following discrete version of Brenier's approach.

Theorem 8. *Let Ω be a compact convex domain in $\mathbb{R}^n$, $\{p_1, ..., p_k\}$ be a set of distinct points in $\mathbb{R}^n$ and $\sigma : \Omega \to \mathbb{R}$ be a positive continuous function. Then for any $A_1, ..., A_k > 0$ with $\sum_{i=1}^{k} A_i = \int_{\Omega} \sigma(x)dx$, there exists $b = (b_1, ..., b_k) \in \mathbb{R}^k$, unique up to adding a constant $(c, c, ..., c)$, sothat $\int_{W_i(b)\cap\Omega} sigma(x)dx = A_i$ for all i. The vectors b are exactly minimum points of the convex function*

$$E(h) = \int_{a}^{h} \sum_{i=1}^{k} \int_{W_i(h)\cap\Omega} \sigma(x)dxdh_i - \sum_{i=1}^{k} h_i A_i$$

on the open convex set $H = \{h \in \mathbb{R}^k | vol(W_(h) \cap \Omega) > 0\ for\ all\ i\}$. Furthermore, ∇u_b minimizes the quadratic cost $\int_{\Omega} |x - T(x)|^2 \sigma(x)dx$ among all transport maps $T : (\Omega, \sigma dx) \to (\mathbb{R}^n, \sum_{i=1}^{k} A_i \delta_{p_i})$

Theorem 9. *(Discrete OMT) If Ω is convex, then the admissible space H_0 is convex, so is the energy $E(\mathbf{h})$, where*

$$E(\mathbf{h}) = \int_{\Omega} u_{\mathbf{h}}(x)\mu(x)dx - \sum_{i=1}^{k} \nu_i h_i.$$

In addition, the unique global minimum $\mathbf{h}_0$ *is an interior point of* H_0*. And the gradient map grad* $u_{\mathbf{h}}$ *induced by the minimum* $\mathbf{h}_0$ *is the unique optimal mass transport map, which minimizes the total transportation cost* $C(T)$*.*

Optimal Mass Transport Map Algorithm Assume Ω is a convex planar domain with measure density μ, $P = \{p_1, ..., p_k\}$ is a point set with measure $\nu = \{\nu_1, ..., \nu_k\}$, such that $\int_\Omega \mu(x)dx = \sum_{i=1}^k \nu_i$. In practice, the energy can be optimized using Newton's method, with the help of the computation of the energy gradient$\nabla E(\mathbf{h}) = (w_1(\mathbf{h}) - \nu_1), ..., w_k(\mathbf{h}) - \nu_k)^T$. The Hessian of $E(\mathbf{h})$ is given as following:

$$\frac{\partial^2 E(\mathbf{h})}{\partial h_i \partial h_j} = \begin{cases} \frac{\int_{e_{ij}} \mu(x)dx}{|y_j - y_i|} & W_i(\mathbf{h}) \cap W_j(\mathbf{h}) \cap \Omega \neq \emptyset \\ 0 & otherwise \end{cases}$$

Algorithm 2 Optimal Mass Transport Map

Require: The Input:

1. A convex planar domain with measure (Ω, μ);
2. A planar point set with measure $(P, \nu), \nu_i > 0, \int_\Omega u(x)dx = \sum_{i=1}^k \nu_i$

The Output: The unique discrete OMT-Map $f : (\Omega, \mu) \to (P, \nu)$

Scale and translate P, such that $P \subset \Omega$
$\mathbf{h} \leftarrow (0, 0, ..., 0)$
Compute the power diagram $D(\mathbf{h})$
Compute the dual power Delaunay triangulation $T(\mathbf{h})$
Compute the cell areas $\mathbf{w}(\mathbf{h}) = (w_1(\mathbf{h}), ..., w_k()\mathbf{h})$
while $||\nabla E|| < \epsilon$ **do**
  Compute ∇E
  Compute the Hessian matrix
  $\lambda \leftarrow 1$
  $\mathbf{h} \leftarrow \mathbf{h} - \lambda H^{-1} \nabla E(\mathbf{h})$
  Compute $D(\mathbf{h})$, $T(\mathbf{h})$, and $\mathbf{w}(\mathbf{h})$
  while $\exists w_i(\mathbf{h}) == 0$ **do**
    $\mathbf{h} \leftarrow \mathbf{h} + \lambda H^{-1} \nabla E(\mathbf{h})$
    $\lambda \leftarrow \frac{1}{2}\lambda$
    $\mathbf{h} \leftarrow \mathbf{h} - \lambda H^{-1} \nabla E(\mathbf{h})$
    Compute $D(\mathbf{h})$, $T(\mathbf{h})$, and $\mathbf{w}(\mathbf{h})$
  end while
end while
Output the result mapping $f : \Omega \to P, W_i(\mathbf{h}) \to p_i,\ i = 1, 2, ..., k.$

Area-preserving Parameterization for Topological Disks The area-preserving mapping is a generalization of the OMT mapping algorithm. Suppose S is a topological disk, with a Riemannian metric $\mathbf{g}$. By scaling, the total area of $(S, \mathbf{g})$ can be equal to π. According to Riemann Mapping theorem, there is a conformal mapping $\phi : (S, \mathbf{g}) \to (\mathbb{D}, dzd\bar{z})$, such that $\mathbf{g} = e^{2\lambda(z)}dzd\bar{z}$. Then we can find the OMT map $\tau : (\mathbb{D}, dzd\bar{z}) \to (\mathbb{D}, e^{2\lambda}dzd\bar{z})$, and the composition $\tau^{-1} \circ \phi : (S, \mathbf{g}) \to (\mathbb{D}, dzd\bar{z})$ gives the area-preserving mapping.

Algorithm 3 Topological Disk Area-preserving Parameterization

Require: The inputs: a triangular mesh M, which is a topological disk; three vertices $\{v_0, v_1, v_2\} \subset \partial M$
The output: The area-preserving parameterization $f : M \to \mathbb{D}$, which maps $\{v_0, v_1, v_2\}$ to $\{1, i, -1\}$ respectively.
1: Scale M such that the total area is π
2: Compute the conformal parameterization $\phi : M \to \mathbb{D}$, such that the images of $\{v_0, v_1, v_2\}$ are $\{1, i, -1\}$
3: For each vertex $v_i \in M$, define $p_i = \phi(v_i)$, ν_i to be $\frac{1}{3}$ of the total area of the faces adjacent to v_i. Set $P = \{p_i\}, \nu = \{\nu_i\}$
4: Compute the *Discrete Optimal Mass Transport Map*
5: Construct the mapping $\tau^{-1} \circ \phi : M \to \mathbb{D}$, which maps each vertex $v_i \in M$ to the centroid of $W_i(\mathbf{h}) \subset \mathbb{D}$

Conformal Wasserstein Distance The OMT-Map algorithm can also be generalized to compute the Wasserstein distance between surfaces. Given two topological disk surfaces $(M_1, g_1, p_1, q_1) \in S, (M_2, g_2, p_2, q_2) \in S$ with total area π, where S is the normalized marked metric space, p_1, p_2 are correspondent interior markers, and q_1, q_2 are markers on the boundary. We first compute the conformal maps $\phi_1 : M_1 \to \mathbb{D}_1$, and $\phi_2 : M_2 \to \mathbb{D}_2$, where $\mathbb{D}_1$ and $\mathbb{D}_2$ are unit planar disks with Euclidean metric $dx^2 + dy^2$, such that $\phi(p_1) = \phi(p_2) = (0, 0)$ and $\phi(q_1) = \phi(q_2) = (1, 0)$. Then we construct a convex planar domain (Ω, μ) from $\mathbb{D}_1$. Then we discretize $\mathbb{D}_2$ into a planar point set with measure (P, ν), where $\nu_i = \dfrac{1}{3} \sum_{[v_i, v_j, v_k] \in M} area([v_i, v_j, v_k])$. Using (Ω, μ) and (P, ν) as input for Algorithm 2, we can compute the OMT map $f : \Omega \to P, W_i(\mathbf{h}) \to p_i$. Therefore, the Wasserstein distance between M_1 and M_2 can be computed by

$$d_W(\mu, \nu) = \sum_{i=1}^{n} \int_{W_i} (x - p_i)^2 \mu(x) dx$$

Polar Factorization In Brenier's theory [3], one of the most important thing is about polar factorization.

Algorithm 4 Computing Wasserstein Distance for Two Surfaces

Require: The Inputs: Two topological disk surfaces:$(M_1, g_1, p_1, q_1), (M_2, g_2, p_2, q_2)$, parameters defined as above.

The Outputs: The Wasserstein distance between M_1 and M_2

1: Scale and normalize M_1 and M_2 such that the total area of each is π.
2: Compute the conformal maps $\phi_1 : M_1 \to \mathbb{D}_1$, and $\phi_2 : M_2 \to \mathbb{D}_2$ defined above.
3: Construct the convex planar domain (Ω, μ) from $\mathbb{D}_1$
4: Discretize $\mathbb{D}_2$ into a planar point set with measure (P, ν)
5: With (Ω, μ) and (P, ν) as inputs, compute the Optimal Mass Transport map f with Algorithm 2
6: Output the Wasserstein distance $d_W(\mu, \nu)$.

Theorem 10. *(Polar Factorization[3]). Let Ω_0 and Ω_1 be two convex subdomains of $\mathbb{R}^n$ with smooth boundaries, each with a positive density function μ_0, μ_1 respectively, and of the same total mass $\int_{\Omega_0} \mu_0 = \int_{\Omega_1} \mu_1$. Let $\phi : (\Omega_0, \mu_0) \to (\Omega_1, \mu_1)$ be an diffeomorphic mapping, then ϕ has a unique decomposition of the form*

$$\phi = (\nabla u) \circ s$$

where $u : \Omega_0 \to \mathbb{R}$ is a convex function, $s : (\Omega_0, \mu_0) \to (\Omega_0, \mu_0)$ is a measure-preserving mapping. This is called a polar factorization of ϕ with respect to μ_0.

In general, a diffeomorphism $\phi : (\Omega_0, \mu_0) \to (\Omega_1, \mu_1)$, where $\mu_1 = \phi_{\#}\mu_0$, can be decomposed to the composition of a measure preserving map $s : (\Omega_0, \mu_0) \to (\Omega_0, \mu_0)$ and a L^2 optimal mass transportation map $\nabla u : (\Omega_0, \mu_0) \to (\Omega_1, \mu_1)$, and the composition is unique.

Since $\nabla u : (\Omega_0, \mu_0) \to (\Omega_1, \mu_1)$ is the unique optimal mass transportation map [3], then according to *Polar Decomposition*, $\nabla u^* = (\nabla u)^{-1} : (\Omega_1, \mu_1) \to (\Omega_0, \mu_0)$ is also an optimal transportation map. So in the pipeline of our algorithm, we first compute the optimal mass transportation $(\nabla u)^{-1}$. The measure-preserving map s can be computed directly by $s = (\nabla u)^{-1} \circ \phi$.

Algorithm 5 Polar Factorization of Mapping

Require: Convex domains Ω_0 and Ω_1 in $\mathbb{R}^d$. A diffeomorphic mapping $\phi : (\Omega_0, \mu_0) \to (\Omega_1, \mu_1)$, satisfying $\mu_1 = \phi_{\#}\mu_0$.

Ensure: The polar factorization $\phi = \nabla u \circ s$, where s is measure-preserving and u is convex.

Compute the unique optimal mass transportation map $\nabla v : (\Omega_1, \mu_1) \to (\Omega_0, \mu_0)$ using Alg.2. The convex function u is the Legendre dual of v, $u = v^*$

Compute the composition $s = \nabla v \circ \phi$

4 Applications Based on Optimal Mass Transport

In this section, several applications based on or derived from the *Optimal Mass Transport* theory will be presented.

4.1 Area-Preservation Mapping using Optimal Mass Transport

As in paper [5],

5 Conclusions and Future Work

References

1. Monge, G.: Mémoire sur la théorie des déblais et des remblais. De l'Imprimerie Royale (1781)
2. Kantorovich, L.V.: On the transfer of masses. In: Dokl. Akad. Nauk. SSSR. Volume 37. (1942) 227–229
3. Brenier, Y.: Polar factorization and monotone rearrangement of vector-valued functions. Communications on pure and applied mathematics **44**(4) (1991) 375–417
4. De Philippis, G., Figalli, A.: The monge–ampère equation and its link to optimal transportation. Bulletin of the American Mathematical Society **51**(4) (2014) 527–580
5. Zhao, X., Su, Z., Gu, X.D., Kaufman, A., Sun, J., Gao, J., Luo, F.: Area-preservation mapping using optimal mass transport. IEEE transactions on visualization and computer graphics **19**(12) (2013) 2838–2847