1. Introduction

Three-dimensional shape analysis has been developing extremely rapidly over the course of the past decade, and much of the research in this field has been catalyzed by the many advances in 3D scanning and printing. Two of the central problems in shape analysis are the development of meaningful characterizations of shape similarities and efficient computational methods to compute point correspondences.

The mathematical framework of three-dimensional shape analysis is the following: the shapes in consideration are 2-manifolds embedded in $\mathbb{R}^3$ and the goal is to develop a meaningful metric on the space of Riemannian 2-manifolds that encodes a notion of shape similarity. Most importantly, we would like this metric to inform us about various point correspondences that can be used for various surface processing tasks.

In this paper, we will consider the Conformal Wasserstein Distance introduced by Lipman and Daubechies [LD11] and explore many of its interesting properties.

The structure of this paper is as follows: first we cover the necessary mathematical and algorithmic prerequisites. Next we summarize the conformal Wasserstein distance in both the continuous and discrete settings, discussing the optimal transportation framework, the main algorithm, and implementation details. Finally, we present illustrative examples that highlight what this distance is measures along with an analysis of various interesting properties that the conformal Wasserstein distance exhibits.

2. Necessary Prerequisites

2.1. Riemann Surfaces and Conformal Mappings. Recall that a Riemann surface is a 1-dimensional complex manifold. However, it is often easier to think of a Riemann surface as a real 2-manifold with a complex structure; as such, it is useful to understand under what conditions does a smooth real 2-manifold admit a complex structure. The following presentation not only gives insight into when real 2-manifolds admit a complex structure, but also give rise to simple and effective procedures for computing conformal parameterizations of surfaces. The easiest way to do this is by considering almost complex structures:

**Definition 2.1.** Let $M$ be a smooth real 2-manifold. An almost complex structure on $M$ is a linear map $\mathcal{J} : TM \to TM$ such that $\mathcal{J}(\mathcal{J}(X)) = -X$ for any tangent vector $X \in T_p M$ at any point $p \in M$. That is $\mathcal{J} \circ \mathcal{J} = - \mathrm{id}$ where $\mathrm{id} : TM \to TM$ is the identity mapping.

We can think about $\mathcal{J}$ as an analog of the complex number $i$ on the tangent planes of $M$ – that is $\mathcal{J}$ acts on tangent vectors by “rotating” them by $90^\circ$. It is known that every real orientable 2-manifold with a Riemannian metric admits an almost complex structure. Remarkably, it turns out that every almost complex 2-manifold admits a conformal structure (refer to Chapter 3.11 in [Jos13]! More importantly, given a 2-manifold with an almost complex structure there exists a unique complex structure such that the conformal maps characterized by the complex structure coincide with those characterized by the Cauchy-Riemann equation described below in Definition 2.2. Given that an almost-complex structure coincides with an actual complex structure, it is often easier to frame problems of conformal surface theory in the language of an almost complex structure, since it can be discussed in the invariant manifold language with ease.
Given this almost complex manifold framework we can begin to talk about what it means for a mapping to be conformal. A conformal mapping $f : C \to C$ is a (real) differentiable map that preserves angles, that is
\[ df(iX) = idf(X), \]
for all vectors $X$, equivalently $f$ is conformal if $df \neq 0$. Given this, we can define what it means to have a conformal mapping between (almost) complex surfaces:

**Definition 2.2.** Let $(M, J_M)$ and $(N, J_N)$ be two Riemann surfaces. We say that $f : M \to N$ is holomorphic or conformal if
\[ df(J_M X) = J_N df(X) \]
for every tangent vector $X$.

As a special case we have that $f : (M, J) \to C$ is a conformal mapping from the surface to the plane if
\[ df(JX) = idf(X) \]
for every tangent vector $X$.

Note that the requirement $df(JX) = idf(X)$ is really just the Cauchy-Riemann equation written in the language of manifolds with almost complex structures. Another important thing to note is that if we have a surface immersed in $\mathbb{R}^3$ then we have a canonical almost complex structure $J$ defined by $df(JX) = N \times df(X)$ where $N$ is the normal vector at $T_pM$ where $X$ is defined.

It turns out that conformal mappings are the right objects to look at when studying geometry processing tasks, since they preserve textures. However, it turns out that the global conformal mappings between surfaces are far too large of a class of deformations to differentiate isometric surfaces – examples of conformally equivalent surfaces that are not isometric are explained in the uniformization section. Nevertheless, as we will see throughout this paper, conformal mappings provide a powerful framework that can indeed be used for powerful discriminative shape analysis.

### 2.2. Uniformization Theory

Trying to compare any two surfaces with arbitrary topology becomes a very difficult and complicated task that often isn’t well defined; however, if we restrict ourselves to the smaller class simply-connected surfaces the uniformization theory of Riemann surfaces can simplify this problem dramatically. In a nutshell, the uniformization theory says that any simply connected compact (possibly with boundary) Riemann surface is conformally equivalent either the sphere or the unit disk.

Throughout this paper, let $D = \{ z \in C : |z| < 1 \}$ denote the unit disk. Lipman and Daubechies use the uniformization theory as the foundation for their conformal Wasserstein distance in the following way: Let $(M, g)$ be a simply-connected disk like Riemann surface, where $g$ is the Riemannian metric on $M$. Let $f : M \to D$ be the conformal mapping to the unit disk, which exists by the uniformization theory. Since $f$ is a conformal mapping we can write the metric on $D$ induced by $f$ in the following form:
\[ \hat{g} = f_* g = \mu(z) \delta_{ij} dx^i \otimes dx^j, \]
where $\mu(z) > 0$ is the conformal factor, or the scaling factor of $M$ at the point $z$. This conformal factor can be thought of as the density of the volume form on $M$. 

![Figure 1. An illustration of an almost complex manifold $(M, J)$, the action of $J$ on a tangent vector $X \in T_pM$, and a conformal parameterization $f : M \to C$ that satisfies the Cauchy-Riemann equation.](image)
Writing everything in this way we see that the uniformization theory simplifies shape analysis since if one wants to create a metric on the space of simply connected disk-like surfaces it suffices to create a metric on the space of the conformal factors, which can be normalized to become probability measures, on the disk.

**Remark 2.3.** Throughout the paper the authors use the hyperbolic density function:

\[ \mu^H(z) := (1 - |z|^2)\mu(z). \]

The reason for this is that the Möbius transformations form the group of *isometries* of the Poincaré disk. That is to say that the hyperbolic distance is invariant under Möbius transformations. This becomes important for ensuring that the optimal Möbius transformations coupling two neighborhoods together does not add any cost based on the specific mapping coupling them together. This also becomes a source of concern when extending the conformal Wasserstein distance (with neighborhoods) to the case of surfaces that are not topological disks.

2.3. **Möbius Transformations.** Another way in which conformal maps help simplify shape analysis is that all of the conformal mappings from the unit disk onto itself are given by Möbius transformations, a three parameter family \( M_D : \)

\[ M_D := \left\{ z \mapsto e^{i\theta} \frac{z - a}{1 - \overline{a}z} : a \in \mathbb{D} \text{ and } \theta \in S^1 \right\}. \]

These three degrees of freedom can loosely be thought of as shifts, rotations, and sphere inversions. One can intuitively think about Möbius transformations as mappings that take circles to circles or lines, and lines to lines or circles.

**Figure 2.** Rather than computing conformal mappings between surfaces to compare their conformal factors, we first apply the uniformization procedure to find a mapping from the surfaces onto the unit disk and then we can use Möbius transformations to study mappings or couplings between the surfaces.

3. **Definition of Conformal Wasserstein Distance**

3.1. **Comparison of metrics on the disc.** The Uniformization Theorem allows us to represent our surfaces \( \mathcal{M} \) and \( \mathcal{N} \) by absolutely continuous measures \( \mu \) and \( \nu \) on the disc \( \mathcal{D} \). At this point in deducing a theory of surface comparison, we are faced with two main obstacles. The first is technical: Since \( \mu \) and \( \nu \) were derived as conformal factors, there is a problem of uniqueness- If \( \mu \) is a conformal factor for \( \mathcal{M} \), so is \( m_\ast \mu \) for any Möbius transformation \( m \). Any method of comparing surfaces via conformal factors will have to address this- choosing the “best” representation of the conformal factor.

The second issue is more subtle: Given two absolutely continuous measures on the disc, there are many ways that we can compare them. One suggestion could be to compare their norm in some function space like \( L^p \), and in fact there are surface metrics which consider such a distance. However, Daubechies and Lipman elect to compare surfaces via a Kantorovich-style transportation energy. Their motivation appears to be found in ideas presented in [RTC00], in which the authors analyze it as a method of comparing visual data. They conclude that its allowance for partial matching and its ability to consider spatial-neighboring
relations between points make it a particularly attractive choice in quantifying when two images "look similar".

With the decision made to define our distance as a modified transportation distance, we need to determine an appropriate cost function which will behave nicely with the necessary Möbius transformation invariant structure.

3.2. Modified cost function. The standard Monge-Kantorovich framework for computing transportation distance is a functional of the form

\[ \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} C(z, w) d\pi(z, w) \]

For some appropriate cost function C. As we previously observed, however, our transportation should, in a sense, "quotient out" Möbius transformations, and look instead at the transportation distance between \( \nu \) and \( (m_{\mu, \nu})_*\mu \) for some optimally chosen Möbius transformation \( m_{\mu, \nu} \). In this case, we will have the distance between \( \mu \) and \( \nu \) written as

\[ \inf_{\pi \in \Pi((m_{\mu, \nu})_*\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} C(z, w) d\pi(z, w) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} C(m_{\mu, \nu}(z), w) d\pi(z, w) \]

The key jumping-off point is the observation that we can view this as a cost function which depends on the metrics being compared: Namely, given a simple cost function \( C \), we use some sort of modified cost function \( C_{\mu, \nu} \) inspired by the relation \( C_{\mu, \nu}(z, w) = C(m_{\mu, \nu}(z), w) \).

In fact the cost used by Daubechies and Lipman is written as follows. For a given scale \( R > 0 \),

\[ d_{\mu, \nu}^R(z_0, w_0) := \inf_{m \in M_D, m(z_0) = w_0} \int_{\Omega_{z_0, R}} |\mu(z) - (m^*(\nu(z)))| d\text{vol}_H(z) \]

where \( \Omega_{z_0, R} \) is a hyperbolic geodesic disc of radius \( R \), centered at \( z_0 \). It should be noted that, unlike the normal cost functions, \( d_{\mu, \nu}(z, w) \) is unrelated to the literal geodesic distance between \( z \) and \( w \) in \( \mathcal{D} \), rather it compares the metric on \( \mathcal{M} \) in a neighborhood of \( z \) to the metric on \( \mathcal{N} \) in a neighborhood of \( w \).

This new cost function \( d_{\mu, \nu}^R \) is well-behaved as a sort of metric. It is clearly symmetric, it is invariant under Möbius transformations, and it satisfies a more generalized form of the triangle inequality. Most importantly for the theory of Optimal Transport, we have the following result:

**Theorem 3.1.** If \( \mu \) and \( \nu \) are continuous from \( \mathcal{D} \) to \( \mathbb{R} \), then \( d_{\mu, \nu}^R(z, w) \) is a continuous function on \( \mathcal{D} \times \mathcal{D} \).

The proof of this relies heavily on the fact that we fully understand the class of Möbius transformations on the disc, and in fact the subclass of Möbius transformations which take \( z \) to \( w \) are a one parameter family (essentially unique up to rotations). To compare, imagine if the infimum in the definition of \( d_{\mu, \nu}^R \) was taken over all homeomorphisms of \( \mathcal{D} \) to itself which mapped \( z \) to \( w \). Here it is not clear if the infimum has a minimizer. Restricting ourselves to Möbius transformations allows us to easily pick a minimizing Möbius transformation depending continuously on \( \mu \) and \( \nu \).

With our cost function well established, we can move on to creating our actual distance on surfaces.

3.3. Optimal Transport. Armed with our modified cost function \( d_{\mu, \nu}^R \) we proceed as indicated above and define a distance on surfaces by

\[ d^R(\mathcal{M}, \mathcal{N}) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} d_{\mu, \nu}^R(z, w) d\pi(z, w). \]

The parameter \( R \) in the distance can be thought of as a feature-scale- the distance \( d^R \) will only be able to distinguish relations between features if \( R \) is greater than the average distance between features. If \( R \) is small, the distance will only be able to detect changes in the metric in a small neighborhood of specific points. Note that the continuity result on \( d^R \) allows us to apply our familiar tricks from Optimal Transport theory and guarantee the existence of a particular \( \pi^* \in \Pi(\mu, \nu) \) such that

\[ d^R(\mathcal{M}, \mathcal{N}) = \int_{\mathcal{D} \times \mathcal{D}} d_{\mu, \nu}^R(z, w) d\pi^*(z, w). \]
3.4. Properties of $d^R(\cdot, \cdot)$. Apart from some technical measure theoretic constructions and the triangle inequality for $d^R_{\mu, \nu}$, it is not terribly difficult to demonstrate that our surface distance $d^R$ satisfies the triangle inequality.

This categorizes it as a pseudo-metric- it is a trivial exercise to see that since the Uniformization Theorem only depends on the Riemannian metric on a surface $M$ that any isometric surface $M'$ would have the same conformal factor (up to a M"obius transformation) and therefore we would have $d^R(M, M') = 0$ for all $R$. The fact that this metric cannot see isometries is something which concerned us and is discussed later on in the experimental results.

However, the converse question is also of some interest: If $d^R(M, N) = 0$, does that mean that they are necessarily isometric? To this end, Daubechies and Lipman provide partial results. Essentially, if $M$ has too much symmetry, we will not be able to show that our optimal transport plan gives rise to an honest to goodness isometry even if $d^R(M, N) = 0$. However, if we ask that this does not occur, we actually can recover an isometry between $M$ and $N$, at least away from the boundary.

To make this more precise, we introduce the notion of $A-M$ fittability.

Definition 3.2. We say that a surface $M$ is $A-M$ fittable for $A > 0$ if there are no subdomains of $M$ of area larger than $A$ which are isometric.

This notion allows us to state the following theorem about $d^R$:

**Theorem 3.3.** Consider two surfaces $M$ and $N$ with conformal factors $\mu$ and $\nu$ such that $d^R(M, N) = 0$ for some $R > 0$. Then, for arbitrarily large $\rho > 0$, there is $m \in M_D$ and $A > 0$ such that if $M$ and $N$ are singly $A-M$ fittable then $\mu(m(z)) = \nu(z)$ for all $z \in \Omega_{0, \rho}$

Essentially that means that if our data sets are sufficiently asymmetrical in the sense of $A-M$ fittability, we can essentially think of $d^R$ as a metric on equivalence classes of isometric surfaces.

4. Algorithmic Implementation and Numerical Aspects

4.1. Discrete Surface Representation and Conformal Mappings. So far we have discussed Riemann surfaces and conformal mappings in the smooth setting, and significant care needs to be taken care of when we attempt to discretize these ideas. Discrete Differential Geometry (DDG) and the Discrete Exterior Calculus (DEC) provide a natural framework for dealing with discrete surfaces. We present a very brief overview of how we use to numerically implement the conformal Wasserstein distance.

We represent our discrete surfaces as a geometrically realized 2-simplicial complex embedded in $\mathbb{R}^3$.

Given the triangulation of a 2-manifold, it becomes very easy to discretize differential operators: given a smooth 0-form $\phi$, we discretize it by storing it’s values on each vertex that is $\hat{\phi} \in \mathbb{R}^{|V|}$. Similarly, given a smooth 1-form $\alpha$ we store

$$\hat{a}_e := \int_a$$
on each of the edges. Finally, we store integrated 2-forms on the faces. Since we store integrated values on the vertices, edges, and faces it becomes easy to discretize the exterior derivative since we have Stokes’ Theorem! The framework of discrete exterior calculus provides an elegant framework to discretize any differential operator. Similarly, there are other approximations of metric operators, such as the Hodge star or the Levi-Civita connection.

We now give a quick review of the Hodge star operator. The Hodge star provides a duality between $k$-forms and $(n-k)$-forms. The key insight comes from enumerating the number of $k$ forms. Note that $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, which tells us that the number of $k$-forms and the number of $(n-k)$ forms are the same, and therefore tells us that there should be some type of natural correspondence between $k$-forms and $(n-k)$-forms. An easy example of this is that a volume form on a manifold can either be expressed as a top-level $n$-form, or it can be expressed as a scalar density (a 0-form) at each point. Now we get to the precise definition: The Hodge star on $k$-forms is a linear isomorphism

$$\ast : \Omega^k(M) \rightarrow \Omega^{n-k}(M)$$

uniquely determined for a $k$-form $\beta$ as the $(n-k)$-form $\ast \beta$ that satisfies

$$\alpha \wedge \ast \beta = \langle \alpha, \beta \rangle \omega,$$
The Hodge duality of differential forms in \( \mathbb{R}^2 \). We have that
\[
\ast 1 = dx^1 \wedge dx^2, \quad \ast dx^1 = dx^2, \quad \ast (-dx^2) = dx^1, \quad \ast (dx^1 \wedge dx^2) = 1.
\]
On any flat space we have that
\[
\ast (dx^1 \wedge dx^2 \wedge \cdots \wedge dx^k) = dx^{k+1} \wedge dx^{k+2} \wedge \cdots \wedge dx^n,
\]
where \((i_1,i_2,\ldots,i_n)\) is any even permutation of \((1,2,\ldots,n)\). On manifolds the Hodge star operator acts in the same way, but it is weighted by the discrete metric (the edge lengths). There are many different discretizations of the Hodge star, and one of the most effective is methods uses finite element exterior calculus (FEEC) as seen in [APW06]. Given our definition of the Hodge star, it is not hard to see that \( \ast a(X) := a(JX) \) on a two manifold, where \( a \) is a 1-form and \( X \in T_pM \) for some \( p \in M \). This is the key connection between the almost complex structure and the Hodge star which makes computing conformal parameterizations of surfaces easy.

Given this representation of a discrete surface, one is now interested in finding conformal parameterizations of the surface over the unit disc. Here we give a very high-level overview of how one might compute the uniformization mapping of an arbitrary disk-like surface to the unit disk. For a map \( z : M \to \mathbb{C} \), on can define the following conformal energy:
\[
E_C(z) := \frac{1}{4} \| \ast dz - idz \|^2,
\]
where for complex 1-form \( a \) we define
\[
\langle a, a \rangle = \int_M \ast \overline{a} \wedge a \quad \text{and} \quad \| a \| = \sqrt{\langle a, a \rangle}.
\]
It is easy to show that \( \langle \cdot, \cdot \rangle \) defines an inner product on the space of complex 1-forms. Note that given the Hodge duality of differential forms in \( \mathbb{R}^2 \) we have that for any tangent vector \( X \in T_pM \) we have that \( \ast dz(X) = dz(JX) \). Note that \( E_C(z) \) measures how much the mapping \( z \) fails to be conformal. Now to find a conformal parameterization into the plane one simply needs to solve the following convex quadratic optimization problem:
\[
\begin{align*}
\text{minimize} & \quad E_C(z) \\
\text{subject to} & \quad \| z \| = 1, \\
& \quad \langle z, 1 \rangle = 0.
\end{align*}
\]
The first constraint ensures that \( z \) does not collapse all of \( M \) onto a single point and the second constraint ensure that the conformal parameterization is centered around the origin. This is really asking to find the best non-trivial almost conformal mapping. In terms of actually solving this convex optimization problem it becomes useful to write the conformal energy as
\[
E_C(z) = \frac{1}{2} \langle \Delta z, z \rangle - A(z),
\]
where \( A(z) \) is the signed area of the mapping \( z \), since the framework of DEC gives a very nice and fast discretization of the Laplace-Beltrami operator on discrete manifolds. Thinking of the conformal parameterization as a constrained minimization problem makes it very easy to impose the appropriate boundary conditions to ensure that the domain is mapped onto the unit disk. Note that there are many more interesting and complicated ways to find various approximations of discrete conformal mappings, but we instead just present the simplest overview of how one might be interested in computing these discrete conformal mappings.

Now that we’ve discussed the conformal Wasserstein distance with neighborhoods (cWn) in the smooth setting we now discuss the specific algorithm, sources of numerical error, and the issue of convergence.

4.2. Computing the conformal Wasserstein distance. There are five main steps to computing the conformal Wasserstein distance: (1) computing the uniformization of the surfaces \( M \) and \( N \) and building smooth interpolants of the conformal factors, (2) discretizing the resulting measures over a finite point sample, (3) approximating the measure-dependent transportation cost between points, and (4) solving the discrete
optimal transportation problem. In each step numerical error accumulates; we will discuss the main complications and sources of numerical error in each of these steps.

The full algorithm, with all it’s details, for computing the conformal Wasserstein distance is concisely presented:

**Algorithm 1** Discrete Approximation of Conformal Wasserstein Distance

1. **procedure** CONFORMALWnDISTANCE($\mathcal{M}, \mathcal{N}, R$)
2. Compute approximate uniformization of $\mathcal{M}$ and $\mathcal{N}$
3. $\mu \leftarrow$ smooth interpolation of conformal factors of $\mathcal{M}$
4. $\nu \leftarrow$ smooth interpolation of conformal factors of $\mathcal{N}$
5. Distribute $n$ points, $Z$, uniformly on $\mathcal{M}$ with respect to $\mu$
6. Distribute $n$ points, $W$, uniformly on $\mathcal{N}$ with respect to $\nu$
7. $\nu^{(Z)} \leftarrow$ Voronoi diagram of $\mathcal{D}$ with centers at $Z$
8. $\nu^{(W)}_M \leftarrow$ Voronoi diagram of $\mathcal{D}$ with centers at $W$
9. Normalize $\mu$ and $\nu$ to become probability measures
10. Let $\hat{\mu}$ be the discretized measure over $\nu^{(Z)}$
11. Let $\hat{\nu}$ be the discretized measure over $\nu^{(W)}$
12. **for** $(z, w) \in \nu^{(Z)}_M \times \nu^{(W)}_N$ **do**
13. Let $\Omega_{z,R}$ be the hyperbolic neighborhood around $z$ of radius $R$
14. Let $\{p_k\} \subseteq \Omega_{z,R}$ be the centers of a quadrature approximation of $\Omega_{z,R}$
15. Let $\{\alpha_k\}$ be the area corresponding to the cell $p_k$
16. $d_{\hat{\mu},\hat{\nu}}^{R}(z, w) \leftarrow \min_{\substack{m \in \mathcal{M} \forall z(w) \in \mathcal{N}}} \sum_{m(z)=w} \alpha_k |\mu(p_k) - \nu(m(p_k))|$
17. **end for**
18. Solve the discrete optimal transport problem between $\hat{\mu}$ and $\hat{\nu}$ described in Equation (*)
19. Set $\hat{d}^R(\mathcal{M}, \mathcal{N}) \leftarrow T_d(\hat{\mu}, \hat{\nu})$
20. **return** $\hat{d}^R(\mathcal{M}, \mathcal{N})$ and point correspondences between $\nu^{(Z)}_M$ and $\nu^{(W)}_N$
21. **end procedure**

The discrete optimal transportation problem we solve can be framed as the following linear program:

\[
\begin{align*}
& \text{minimize} & & \sum_{z,w} d_{\mu,\nu}^{R}(z, w) \pi_{z,w} \\
& \text{subject to} & & \sum_{z} \pi_{z,w} = \hat{\nu}(w), & \sum_{w} \pi_{z,w} = \hat{\mu}(z), \\
& & & \pi_{z,w} \geq 0 & \text{for all } (z,w) \in \nu^{(Z)}_M \times \nu^{(W)}_N.
\end{align*}
\]

Note that since we are solving a discrete optimal transportation problem between two point samples with the same number of points, mass will not be split. As such, the optimal transportation framework gives us point correspondences between the surfaces which can be used for various shape analysis tasks such as surface registration.

### 4.3. Effects of Numerical Error

We study the effects on the accuracy of the conformal parameterizations of the surfaces and the number of points used to discretize the measures on the approximation of the conformal Wasserstein distance. Throughout these tests we use two randomly selected meshes from the teeth database available at [http://www.wisdom.weizmann.ac.il/~ylipman/CPsurfcomp/](http://www.wisdom.weizmann.ac.il/~ylipman/CPsurfcomp/). The control which we compare various numerical approximations is computed using the minimum number of vertices of $\mathcal{M}$ and $\mathcal{N}$ (almost every vertex of the mesh) for the support of the discrete measures and best conformal parameterizations of the surfaces.

We begin by discussing the effects of the accuracy of the conformal parameterizations of the surfaces. We use the quasi-conformal error, described in [SSGH01], to determine how close the parameterization is to being conformal. Let $\eta : \mathcal{M} \to \mathcal{D}$ be any mapping; the quasi-conformal error $Q : \mathcal{M} \to \mathbb{R}$ is defined as the
ratio of the largest to smallest singular value of \(d\eta\), and we compute \(Q\) over each face in the discrete setting. An error of \(Q = 1\) is optimal, since it means that no shearing of the faces takes place. We implemented a total of 4 distinct uniformization procedures. The effects they have on the distance approximation are presented in Figure 3.

Next we compute the effects of increasing the sample size on the conformal Wasserstein distance error. It turns out that discrete measures approximate the continuous measures representing the conformal factors with a relatively small number of points. Again, the effects are summarized in Figure 3. It is important to note that the error induced from the conformal parameterization is much less significant when compared to the error induced by course discretizations of the underlying measure.

![Figure 3](image)

**Figure 3.** These figures show the general trend that the numerical error in the uniformization step has much less dramatic effect on the overall distance approximation when compared to the number of points used to discretize the optimal transportation cost.

### 4.4. Convergence

Lipman and Daubechies prove very general convergence results for the approximation scheme described above for the computation of the optimal transportation cost between arbitrary separable complete compact metric spaces. Furthermore, given the additional restriction that the cost function \((d_{\mu,\nu})\) is uniformly continuous they obtain convergence rates as a function of the modulus of continuity.

To recall, the approximation scheme the authors is based on Voronoi cells. Given two probability measures \(\mu,\nu\) on the unit disk, one begins by uniformly distributing two sets of \(n\) points, \(Z = \{z_1,\ldots,z_n\}\) and \(W = \{w_1,\ldots,w_n\}\), with respect to \(\mu\) and \(\nu\), respectively. One then builds up the corresponding Voronoi diagrams \(V_Z\) and \(V_W\), discretizes the measure on these diagrams to obtain two discrete measures \(\tilde{\mu}\) and \(\tilde{\nu}\). Finally, one approximates the continuous optimal transportation distance by computing \(T_{d_{\tilde{\mu},\tilde{\nu}}} (\tilde{\mu},\tilde{\nu})\).

Rather than going into all of the specifics of the proof, we provide the main result, briefly explain how this informs a better algorithm implementation, and briefly discuss the uniformization error in more depth than is presented in the paper.

The only notion necessary to understand the convergence results is the fill-distance. The fill-distance associated to a set \(Z \subseteq M\) and a metric tensor \(g\), is defined as

\[
\varphi_g(Z) := \sup \{ r > 0 : z \in M \text{ such that } B_g(z,r) \cap Z = \emptyset \}.
\]

This is just the radius of the largest geodesic ball that can be placed on \(M\) that doesn’t intersect \(Z\).

Let \(g\) and \(h\) be the metrics on the unit disk induced by the uniformizations of \(M\) and \(\mathcal{N}\), respectively. Let \(E\) denote the Euclidean metric. Let \(Z\) and \(W\) denote \(n\) uniformly distributed points on \(M\) from \(\mu\) and \(\mathcal{N}\) from \(\nu\), respectively. Let \(\mu_Z\) and \(\mu_W\) denote the Voronoi diagram based discretization of \(\mu\) and \(\nu\). Let \(\{p_k\}\) denote the quadrature samples used for discretizing the numerical integration, and finally let \(L\) denote the number of uniformly distributed points on \(S^1\) used to discretize the family of Möbius transformations that map points to points. It then follows that

\[
|d\hat{R}(M,N) - \hat{d}\hat{R}(M,N)| \leq C d_{\mu,\nu} (2 \max\{\varphi_g(Z),\varphi_h(W)\}) + C_1 \varphi_E (\{p_k\}) + \frac{C_2}{L} + \varepsilon_{\text{uniformization}},
\]
where $C_1, C_2$ are constants that depend only on $\mu, \nu, R$ and $\omega_{\mu, \nu}^R$ is the modulus of continuity of the neighborhood dissimilarity metric. This convergence follows from this bound since the metric $d^{R}_{\mu, \nu}$ is uniformly continuous.

Note that this result is very important since it tells us that it’s never a good idea to use random samples obtained uniformly from $\mu$ and $\nu$, but rather that we should just distribute $n$ points on $M$ and $N$ that minimizes the size of the largest Voronoi cell – that is choose the samples that minimize the fill distances above.

Note that the $\epsilon_{\text{uniformization}}$ term is an error that is only based on the differences between the true conformal factors $\mu$ and $\nu$ and the approximated conformal factors $\mu^{\text{approx}}$ and $\nu^{\text{approx}}$. Since the conformal Wasserstein distance $d^R$ is indeed a metric on the space of topological disks, we have that

$$
\epsilon_{\text{uniformization}} \leq T^R_d (\mu, \mu^{\text{approx}}) + T^R_d (\nu, \nu^{\text{approx}}).
$$

The authors state that this uniformization error is much smaller than the errors induced by the sampling and discretizations of the measures, but don’t go into any of the details. From the above numerical experiments, we see that the uniformization step does play a significant role in the error, especially if one is using a large number of sample points. It looks like this error scales linearly as a function of the quasi-conformal error. Furthermore, many uniformization procedures show that the quasi-conformal error scales linearly as a function of the number of points sampled from a smooth manifold – that is to say that as the mesh becomes more refined we are able to compute more accurate conformal parameterizations that better approximate the smooth case, and we have that the mean quasi-conformal error converges linearly as a function of $|V|$, the number of vertices (i.e. quasi-conformal error is $O(1/|V|) \approx O(\epsilon)$, where $\epsilon$ is the mean edge length). Although we haven’t done experiments to quantify exactly how the mesh quality effects $\epsilon_{\text{uniformization}}$ we can infer from the effects of the quasi-conformal error on the numerical error in Figure 3 and from the behavior of quasi-conformal error as a function of $|V|$ that as the mesh quality increases (becomes less coarse and the number of vertices on the mesh increases) that the $\epsilon_{\text{uniformization}}$ is on the order $O(1/|V|)$.

It is important to know that in practice it is often very difficult to artificially increase the mesh quality in a way that preserves the underlying geometry of the surface. Nevertheless, it is useful to understand the effects of all sources of error.

5. Examples and Discussion

Given our discussion of the conformal Wasserstein distance, we see that there are many intertwined computations that make it difficult to clearly understand what types of surfaces can this distance discriminate between, and what feature transformations have the most dramatic effect on the distances between surfaces. Lipman and Daubechies prove that $d^R(M, N) = 0$ if and only if $M$ and $N$ are isometric under additional restrictions on the space of surfaces – namely that there are no self-isometries on the scale of $R$ or larger. However, this assumption may not hold up in practice, and we show that this assumption is necessary for the conformal Wasserstein distance to be able to discriminate between non-isometric surfaces.

5.1. Continuity of the mappings. When studying the conformal Wasserstein distance, one might be concerned that the mappings between surfaces aren’t very smooth or continuous. One would hope that nearby points have locally similar conformal factors, which would implicitly some regularity on the point correspondences. However, given arbitrary set of meshes there is no reason to believe that the mappings should be smooth, a-priori; that is to say there is nothing which explicitly forces the minimizing optimal transportation plan $\pi$ not to jump around a lot. Here we present an example from the teeth dataset that show that, in general, the optimal transportation maps between surfaces need not be continuous.

5.2. Surfaces such that $d^R(M, N) \approx 0$. To try and understand the discriminative ability of the conformal Wasserstein distance (with neighborhoods) we constructed several artificial examples to elucidate some of the main features and drawbacks.

Throughout both of the papers [LD11, LPD13] most of the analysis is done under the additional assumption that the surfaces are not locally isometric. The simplest example of two surfaces that are not isometric but have almost zero distance are two surfaces with heights determined by two bimodal distributions with the centers of the peaks spread out: the point correspondences between these two surfaces is shown in Figure 5. This suggests that any local translations of surfaces cannot be differentiated by cWn. Formally,
we claim that the cWn is invariant under the following types of transformations: consider a geodesic ball centered at a local extrema of the Gaussian curvature that does not contain any other local extrema, denote it by $\Omega$, define a constant vector field $\vec{v} = c$ on $\Omega$ and $\vec{v} = 0$ on $\Omega^c$. Next we flow the mesh for some time against this vector field. An example of such a transformation is shown in Figure 5.

It is important to note that new features will develop if this transformation is done na"ively, so one can resample the mesh to remove the irregular features introduced by this transformation. This isn’t necessarily the best way to transform the meshes, but it worked reasonably well and I didn’t have much time to implement this. We then tested these types of transformations on the teeth dataset and observe that the conformal Wasserstein distance was unable to discern the difference between any of these two surfaces.

Next we looked at how sensitive this metric is to near-isometries. To test this we used 3 distinct hand models that have the fingers moved around and compared the top point correspondences and the distance between these surfaces. We find that the point correspondences between the surfaces are very meaningful and accurate, but the metric cannot discriminate between the surfaces. In some sense, this is a good thing since the surfaces are near isometric, but in another sense this is an undesirable feature for a distance claiming to be a metric on the space of surfaces.

5.3. Effects of Noise. A desirable feature that arises from the fact that the conformal Wasserstein distance considers neighborhood similarity is that the metric is able to match surfaces even in the presence of noise. Since the cost function is defined via integration over neighborhoods, we find that gaussian noise gets averaged out in some sense, and the point correspondences are still correct. However, since the distance
Figure 6. Top six point correspondences. The metric is unable to distinguish near isometries. The pairwise distances between these surfaces were all on the order of $10^{-4}$.

returned is based on the actual integrals over these neighborhoods, we find that the metric itself is not robust to noise. An illustrative example of this is shown in Figure 7.

Figure 7. Point correspondences induced by the conformal Wasserstein distance are robust to the effects of noise.

5.4. **Classification using the conformal Wasserstein distance.** Although this metric has several undesirable properties that arise when considering these artificial or constructed examples, Lipman and Daubechies were interested in using the metric to automatically and classify teeth from various species. Given the teeth dataset, we reproduced some of the results here and find that these results coincide with the results reported in [BLC+11]. We provide a ROC curve for the classification between two classes of teeth in Figure 8.
Figure 8. An ROC curve comparison three distinct differences. Here we see that there no real advantages to using the full blown conformal Wasserstein distance, rather than restricting to comparison neighborhoods.

References


