

A Novel Sliced Fused Gromov-Wasserstein Distance

Moritz Piening*, Robert Beinert*

Institut für Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany
 piening@math.tu-berlin.de, beinert@math.tu-berlin.de

Abstract

The Gromov–Wasserstein (GW) distance and its fused extension (FGW) are powerful tools for comparing heterogeneous data. Their computation is, however, challenging since both distances are based on non-convex, quadratic optimal transport (OT) problems. Leveraging 1D OT, a sliced version of GW has been proposed to lower the computational burden. Unfortunately, this sliced version is restricted to Euclidean geometry and loses invariance to isometries, strongly limiting its application in practice. To overcome these issues, we propose a novel slicing technique for GW as well as for FGW that is based on an appropriate lower bound, hierarchical OT, and suitable quadrature rules for the underlying 1D OT problems. Our novel sliced FGW significantly reduces the numerical effort while remaining invariant to isometric transformations and allowing the comparison of arbitrary geometries. We show that our new distance actually defines a pseudometric for structured spaces that bounds FGW from below and study its interpolation properties between sliced Wasserstein and GW. Since we avoid the underlying quadratic program, our sliced distance is numerically more robust and reliable than the original GW and FGW distance; especially in the context of shape retrieval and graph isomorphism testing.

1 Introduction

The Gromov–Wasserstein (GW) distance (Mémoli 2011) and the Fused Gromov–Wasserstein (FGW) distance (Vayer et al. 2020a) extend the classical optimal transport (OT) framework (Villani 2003) to the comparison of heterogeneous data by modelling them as metric measure spaces (mm-spaces). While the resulting distances are powerful tools, their computation is costly and does not admit an exact solution. To accelerate the underlying non-convex, quadratic program, several algorithms have been proposed using regularization techniques (Peyré, Cuturi, and Solomon 2016; Peyré and Cuturi 2019), employing linearized distances (Beier, Beinert, and Steidl 2022; Nguyen and Tsuda 2023), quantizing the problem (Chowdhury, Miller, and Needham 2021) or using constrained optimization (Scetbon, Peyré, and Cuturi 2022; Scetbon et al. 2023). As a consequence, established solvers mainly compute upper bounds and numerical approximations of the actual GW and FGW distance.

*These authors contributed equally.

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

An alternative acceleration framework builds on 1D OT and random projections to define the so-called sliced GW distance (Vayer et al. 2019; Beinert, Heiss, and Steidl 2023). However, sliced GW is restricted to Euclidean geometry and lacks isometric invariance, limiting its practical use. As a remedy, we design a novel sliced GW distance fixing these issues and extend it to FGW. In particular, our new distance allows for an analytical computation outlined in Figure 1.

For our design, we turn towards the complementary literature on lower GW bounds (Mémoli 2011; Vayer et al. 2019, 2020b; Jin, Yu, and Zhang 2022). Since many of these rely on optimization heuristics, such as bi-convex relaxation (Vayer et al. 2020b) or sorting (Vayer et al. 2019; Beinert, Heiss, and Steidl 2023), and we are interested in exact computation, we turn our attention to the bounds in (Mémoli 2011). In particular, we are motivated by recent theoretical work proving their metricity on certain spaces (Mémoli and Needham 2022) and the practical effectiveness of such bounds in graph isomorphism testing (Weitkamp 2022; Weitkamp et al. 2024), where established GW solvers fail in practice. The aim behind graph isomorphism testing (Shervashidze et al. 2011; Grohe and Schweitzer 2020) is to verify or disprove whether two or more graphs are topologically equivalent or not.

In this paper, we focus on the strongest of Mémoli’s GW bounds—the third lower bound (TLB)—and start by extending TLB to the fused case. TLB itself is defined as a hierarchical OT problem, i.e., an OT problem whose costs are given by another OT problem. Indeed, formulations of such kinds recently have attracted attention from the OT community (Dukler et al. 2019; Alvarez-Melis and Fusi 2020). Despite interesting theoretical properties (Delon and Desolneux 2020; Bonet, Vauthier, and Korba 2025), these problems are usually computationally challenging. Therefore, recent work has focused on extending sliced Wasserstein (SW) techniques to the hierarchical case for acceleration (Piening and Beinert 2025b; Nguyen et al. 2025). Keeping this technique in mind, we tackle the computational burden of our new FGW lower bound by rewriting it as an Euclidean Wasserstein distance. To circumvent the limitation of slicing in high dimensions (Tanguy, Flamary, and Delon 2025), we propose to employ numerical quadrature schemes. All in all, we deduce an equivalent, but more efficient, lower bound. Apart from extending the literature on sliced hierarchical

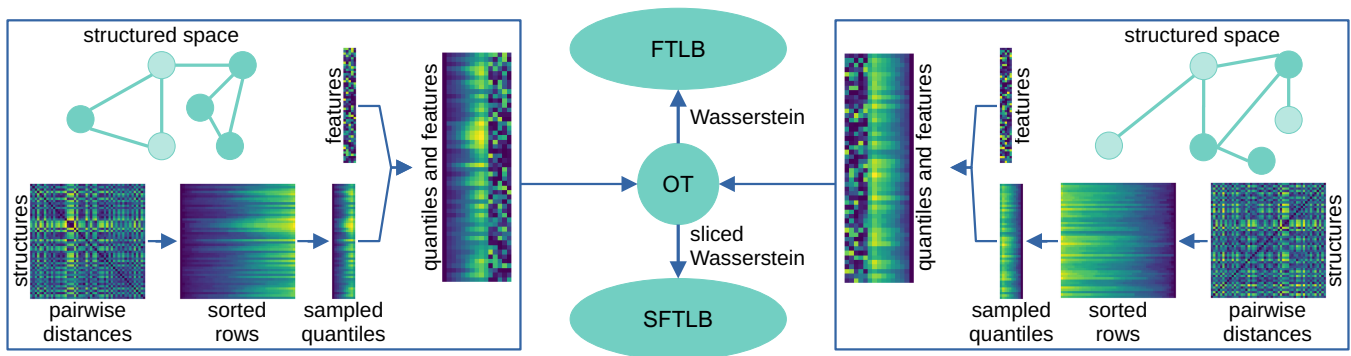


Figure 1: Visualization of our novel sliced fused Grovov–Wasserstein distance called SFTLB: Starting from two structured spaces like labeled graphs equipped with a graph distance, we sort the corresponding distance matrices row-wise. Subsequently, we determine samples of the quantile functions of the local distance distribution for each node. Afterward, we concatenate the sampled quantiles with the original features and interpret the outcome as an empirical Euclidean measure. Finally, we compute the Wasserstein or sliced Wasserstein distance between these measures.

OT, this also overcomes previous limitations of sliced GW, where sorting-based solvers might fail (Vayer et al. 2019; Beinert, Heiss, and Steidl 2023). We summarize our contributions as follows:

1. Towards the goal of an exact solver, we generalize Mémoli’s GW bound to FGW and recall its connection to a classical Wasserstein problem.
2. To overcome the limitations of sliced GW, we employ this connection and numerical quadrature schemes to define an effective novel sliced FGW distance between arbitrary structured spaces beyond Euclidean geometry.
3. We prove that our new distance is a pseudo-metric interpolating between TLB and SW. Additionally, our proof illustrates the dimensional dependence of SW.
4. Lastly, we provide numerical experiments for our sliced FGW, showcasing its efficiency as well as its applicability for shape comparison and graph isomorphism testing.

Proofs and additional remarks/experiments can be found in the extended version on arXiv (Piening and Beinert 2025a).

2 Optimal Transport-Based Distances

At the heart of OT is the calculation of optimal transfers from one measure to another. For this, let $\mathcal{P}(X)$ denote the space of Borel probability measures on a compact Polish space X . A measure $\xi \in \mathcal{P}(X)$ is translated to another compact Polish space Y via a mapping $T: X \rightarrow Y$ by the *push-forward* $T_{\#} \xi := \xi \circ T^{-1}$. Based on the canonical projections π_{\bullet} onto a component, e.g., $\pi_X: X \times Y \rightarrow X, (x, y) \mapsto x$, the set of *transport plans* between $\xi \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$ is defined as

$$\Gamma(\xi, \nu) := \{ \gamma \in \mathcal{P}(X \times Y) \mid \pi_{X, \#} \gamma = \xi, \pi_{Y, \#} \gamma = \nu \}.$$

In general, OT seeks to find an optimal transport plan minimizing a given loss function. For certain losses, this procedure yields a metric between probability measures.

Wasserstein Distance

For compact $Z \subset \mathbb{R}^d$ equipped with the Euclidean norm $\|\cdot\|$, $\zeta, \zeta' \in \mathcal{P}(Z)$, and $p \in [1, \infty)$, the *p-Wasserstein distance* is introduced as

$$W_p^p(\zeta, \zeta') := \inf_{\gamma \in \Gamma(\zeta, \zeta')} \int_{Z \times Z} \|z - z'\|^p d\gamma(z, z') \quad (1)$$

$\underbrace{\hspace{10em}}_{=:\mathcal{T}_p(\gamma)}$

and defines a metric¹ on $\mathcal{P}(Z)$. Its computation requires expensive optimization procedures since closed-form solutions are rarely available.

One exception is the special case $Z \subset \mathbb{R}$, where the optimal plan can be expressed using the *quantile function*:

$$q_{\zeta}: (0, 1) \rightarrow \mathbb{R}, s \mapsto \inf \{ z \in \mathbb{R} \mid \zeta((-\infty, z]) > s \}.$$

Then, the *p-Wasserstein distance* admits the closed-form

$$W_p^p(\zeta, \zeta') = \int_0^1 |q_{\zeta}(s) - q_{\zeta'}(s)|^p ds; \quad (2)$$

see (Villani 2003). For discrete measures with n and m support points, the integral in (2) reduces to a sum with $n + m - 1$ summands; see (Peyré and Cuturi 2019, Prop. 3.5) and (Villani 2003, Thm. 2.18).

Sliced Wasserstein Distance

To leverage the computational benefits of the 1D Wasserstein distance, we consider the *slicing operator*

$$\pi_{\theta}: \mathbb{R}^d \rightarrow \mathbb{R}, z \mapsto \theta \cdot z, \quad \theta \in \mathbb{S}^{d-1},$$

where $\mathbb{S}^{d-1} := \{ \theta \in \mathbb{R}^d \mid \|\theta\| = 1 \}$ denotes the sphere with surface area $A(\mathbb{S}^{d-1})$ and \cdot the Euclidean inner product. For compact $Z \subset \mathbb{R}^d$, $\zeta, \zeta' \in \mathcal{P}(Z)$, and $p \in [1, \infty)$, the *sliced p-Wasserstein (SW) distance* is introduced as

$$SW_p^p(\zeta, \zeta') := \frac{1}{A(\mathbb{S}^{d-1})} \int_{\mathbb{S}^{d-1}} W_p^p(\pi_{\theta, \#} \zeta, \pi_{\theta, \#} \zeta') d\theta$$

¹Symmetric, positive definite function with triangle inequality.

with respect to the surface measure on \mathbb{S}^{d-1} . The interest in this sliced distance is justified by the close relation to the classical Wasserstein distances. Namely, for compact $Z \subset \mathbb{R}^d$, both distances display a form of metric equivalence, i.e., there exists $C_Z > 0$ such that

$$SW_p^p(\zeta, \zeta') \leq W_p^p(\zeta, \zeta') \leq C_Z SW_p^{\frac{1}{d+1}}(\zeta, \zeta')$$

for any $\zeta, \zeta' \in \mathcal{P}(Z)$; see (Bonnotte 2013; Bonneel et al. 2015; Kolouri, Park, and Rohde 2016). Moreover, the spherical integral can be efficiently approximated via random projections using Monte Carlo or Quasi-Monte Carlo methods for low-dimensional data (Nguyen, Bariletto, and Ho 2024; Hertrich, Jahn, and Quellmalz 2025) as well as Gaussian approximations for high-dimensional data (Nadjahi, Vialard, and Peyré 2021). Note that there exist various extensions to non-Euclidean spaces (Kolouri et al. 2019; Quellmalz, Beinert, and Steidl 2023; Bonet, Drumetz, and Courty 2025), invariant data (Beckmann, Beinert, and Bresch 2025a,b), general probability divergences (Nadjahi et al. 2020; Hertrich et al. 2024) and kernels (Rux, Quellmalz, and Steidl 2025).

Gromov–Wasserstein Distance

To compare measures on distinct metric spaces, the GW distance extends classical OT by matching pairwise distances. For this, let $\mathbb{X} := (X, g, \xi)$ be a *metric measure (mm) space* consisting of a compact metric space (X, g) and $\xi \in \mathcal{P}(X)$. For mm-spaces $\mathbb{X} := (X, g, \xi)$ and $\mathbb{Y} = (Y, h, \nu)$, the *Gromov–Wasserstein (GW) distance* is then defined as

$$GW_p^p(\mathbb{X}, \mathbb{Y}) := \inf_{\gamma \in \Gamma(\xi, \nu)} \underbrace{\int \int_{(X \times Y)^2} |g(x, x') - h(y, y')|^p \times d\gamma(x', y') d\gamma(x, y)}_{=: \mathcal{D}_p(\gamma)}. \quad (3)$$

Here, each transport plan π induces an overall distortion \mathcal{D}_p between g and h . This expression defines a metric on the equivalence classes of isomorphic mm-spaces; see (Mémoli 2011). Since (3) is a non-convex, quadratic program, its minimization typically relies on entropic regularization and costly block-coordinate descent (Peyré and Cuturi 2019). As a consequence, a set of hierarchical lower bounds has been proposed in (Mémoli 2011; Mémoli and Needham 2022).

First Lower Bound Define the pointwise *p*-eccentricity $s_{\mathbb{X}, p} : X \rightarrow \mathbb{R}$ of an mm-space $\mathbb{X} := (X, g, \xi)$ as

$$s_{\mathbb{X}, p}^p(x) := \int_X g^p(x, x') d\xi(x').$$

Then, the first lower GW bound between \mathbb{X} and \mathbb{Y} is

$$FLB_p^p(\mathbb{X}, \mathbb{Y}) := W_p^p(s_{\mathbb{X}, p, \#} \xi, s_{\mathbb{Y}, p, \#} \nu),$$

which compares only the first-order moments of pairwise distances. This is an efficiently solvable 1D OT problem.

Second Lower Bound By matching the entire distribution of distances, the second lower GW bound reads as

$$SLB_p^p(\mathbb{X}, \mathbb{Y}) := W_p^p(g_{\#}(\xi \otimes \xi), h_{\#}(\nu \otimes \nu)),$$

where \otimes denotes the product measure. This lifts FLB by considering all pairwise distances at once. Again, characterization as 1D transport problem enables fast computation.

Third Lower Bound For each $(x, y) \in X \times Y$, the local distance distributions between \mathbb{X} and \mathbb{Y} is defined via

$$\begin{aligned} LD_p^p(x, y) &:= \inf_{\gamma \in \Gamma(\xi, \nu)} \int_{X \times Y} |g(x, x') - h(y, y')|^p d\gamma(x', y') \\ &= W_p^p(g(x, \cdot)_{\#} \xi, h(y, \cdot)_{\#} \nu). \end{aligned} \quad (4)$$

Decoupling the minimization of the inner and outer integral in (3), we arrive at the third lower GW bound:

$$TLB_p^p(\mathbb{X}, \mathbb{Y}) := \inf_{\gamma \in \Gamma(\xi, \nu)} \underbrace{\int_{X \times Y} LD_p^p(x, y) d\gamma(x, y)}_{=: \tilde{\mathcal{D}}_p(\gamma)}. \quad (5)$$

Note that the inner plans in (4) do not match the outer plan in (5). Similar to the other two cases, the estimation of LD benefits from accelerated 1D solutions. In contrast, TLB still requires a costly OT minimization.

Comparing all three lower bounds, we obtain the hierarchy (Mémoli 2011; Chowdhury and Mémoli 2019):

$$FLB_p(\mathbb{X}, \mathbb{Y}) \leq SLB_p(\mathbb{X}, \mathbb{Y}) \leq TLB_p(\mathbb{X}, \mathbb{Y}) \leq GW_p(\mathbb{X}, \mathbb{Y}).$$

Moreover, TLB is positive definite and thus defines a metric for certain classes of mm-spaces, whereas FLB and SLB fail to distinguish isomorphic mm-spaces (Mémoli and Needham 2022). To avoid the outer OT problem in (5), (Sato et al. 2020) rely on the maximum mean discrepancy (MMD) with respect to the 1D Wasserstein distance as kernel, which leads to an efficient computation scheme. In this paper, we follow a different approach by accelerating the computation of the outer OT problem using an efficient slicing, which actually leads to a metric equivalence with TLB.

Fused Gromov–Wasserstein Distance

The GW distance allows for a geometrically meaningful comparison between unlabeled, undirected graphs. However, due to the importance of node features in practical applications, practitioners often resort to a combination of Wasserstein and GW distances. For this, the mm-spaces are extended by a common compact feature space $Z \subset \mathbb{R}^d$. More precisely, a *structured space* $\mathbb{X}^S := (X \times Z, g, \xi)$ consists of a compact metric space (X, g) and $\xi \in \mathcal{P}(X \times Z)$. Structured spaces are also known as labelled spaces or structured objects (Vayer et al. 2020a). The space of all structured spaces is denoted as $\mathcal{S}(Z)$. Having the generalization of TLB and its slicing in mind, we study the following distance on $\mathcal{S}(Z)$, which is a specific variant of (Vayer et al. 2020a, Def. 8): for $\mathbb{X}^S := (X \times Z, g, \xi)$, and $\mathbb{Y}^S := (Y \times Z', h, \nu)$ with $Z = Z'$ as well as $\alpha \in [0, 1]$ and $p \in [1, \infty)$, the *fused GW (FGW) distance* reads as

$$FGW_{\alpha, p}^p(\mathbb{X}^S, \mathbb{Y}^S) := \inf_{\gamma \in \Gamma(\xi, \nu)} (1 - \alpha) \mathcal{D}_p(\pi_{X \times Y, \#} \gamma) + \alpha \mathcal{T}_p(\pi_{Z \times Z', \#} \gamma),$$

where \mathcal{D}_p denotes the distortion in (3), \mathcal{T}_p refers to the transport in (1), and α balances structure versus features. In the case $\alpha = 0$, we recover the original GW distance of the structure part; for $\alpha = 1$, we obtain the original Wasserstein distance of the label part. Notice that, for $\alpha \in (0, 1)$ and

$p \in [1, \infty)$, FGW only defines a semi-metric² on the space of isomorphic structure spaces, fulfilling the relaxed triangle inequality

$$\begin{aligned} & \text{FGW}_{\alpha,p}^p(\mathbb{Y}_1^S, \mathbb{Y}_2^S) \\ & \leq 2^{p-1}(\text{FGW}_{\alpha,p}^p(\mathbb{Y}_1^S, \mathbb{X}^S) + \text{FGW}_{\alpha,p}^p(\mathbb{X}^S, \mathbb{Y}_2^S)); \end{aligned}$$

see (Vayer et al. 2020a, Prop. 4).

3 Accelerating the Calculation of FGW

A Lower Bound Using TLB

In principle, any of the lower bounds for the GW distance may be directly extended to the FGW distance. Due to the metric properties of TLB and its tight affinity to the original GW distance, we solely focus on this bound. The basic idea is to replace the GW distortion in FGW by TLB. In this manner, for $\mathbb{X}^S := (X \times Z, g, \xi)$, $\mathbb{Y}^S := (Y \times Z', h, v)$ with compact $Z = Z' \subset \mathbb{R}^d$, $\alpha \in [0, 1]$, and $p \in [1, \infty)$, we introduce the *Fused Third Lower Bound* (FTLB) via

$$\text{FTLB}_{\alpha,p}^p(\mathbb{X}^S, \mathbb{Y}^S) := \inf_{\gamma \in \Gamma(\xi, v)} (1 - \alpha) \tilde{D}_p(\pi_{X \times Y, \#} \gamma) + \alpha \mathcal{T}_p(\pi_{Z \times Z', \#} \gamma), \quad (6)$$

where \tilde{D}_p relates to the mm-spaces $\mathbb{X} := (X, g, \pi_{X, \#} \xi)$ and $\mathbb{Y} := (Y, h, \pi_{Y, \#} v)$. Similar to GW and FGW, this includes TLB for $\alpha = 0$. Consequently, all subsequent results apply to the classical GW distance as well. Indeed, FTLB provides a lower bound since we minimize over multiple, uncoupled inner plans as opposed to finding a single coupled plan. Due to the similarity to FGW, many properties of the original distance carry over to FTLB.

Proposition 3.1. For $\mathbb{X}^S := (X \times Z, g, \xi)$, $\mathbb{Y}^S := (Y \times Z, h, v)$, $Z \subset \mathbb{R}^d$ compact, let $\xi^{\mathcal{F}} := \pi_{Z, \#} \xi$, $v^{\mathcal{F}} := \pi_{Z, \#} v$ and $\mathbb{X} := (X, g, \pi_{X, \#} \xi)$, $\mathbb{Y} := (Y, h, \pi_{Y, \#} v)$.

- (i) There exists $\gamma^* \in \Gamma(\xi, v)$ minimizing (6).
- (ii) $\text{FTLB}_{\alpha,p}^p(\mathbb{X}^S, \mathbb{Y}^S) \leq \text{FGW}_{\alpha,p}^p(\mathbb{X}^S, \mathbb{Y}^S)$.
- (iii) $\text{FTLB}_{\alpha,p}^p(\mathbb{X}^S, \mathbb{Y}^S) \rightarrow W_p(\xi^{\mathcal{F}}, v^{\mathcal{F}})$ as $\alpha \rightarrow 1$.
- (iv) $\text{FTLB}_{\alpha,p}^p(\mathbb{X}^S, \mathbb{Y}^S) \rightarrow \text{TLB}_p(\mathbb{X}, \mathbb{Y})$ as $\alpha \rightarrow 0$.
- (v) $\text{FTLB}_{\alpha,p}$ defines a pseudo-semi-metric³ on $\mathcal{S}(Z)$ with

$$\begin{aligned} & \text{FTLB}_{\alpha,p}^p(\mathbb{Y}_1^S, \mathbb{Y}_2^S) \\ & \leq 2^{p-1}(\text{FTLB}_{\alpha,p}^p(\mathbb{Y}_1^S, \mathbb{X}^S) + \text{FTLB}_{\alpha,p}^p(\mathbb{X}^S, \mathbb{Y}_2^S)) \end{aligned}$$

for all $\mathbb{X}^S, \mathbb{Y}_1^S, \mathbb{Y}_2^S \in \mathcal{S}(Z)$.

A Novel Lower Bound via Slicing

While lowering the computational burden, FTLB still requires solving an OT problem. For two discrete measures of cardinality n , this leads to a computational complexity of $\mathcal{O}(n^3)$ for an exact OT solution and $\mathcal{O}(n^2 \log n)$ for a regularized one (Peyré and Cuturi 2019). To achieve further acceleration, we aim to combine ideas from SW and numerical integration. For this, let $\mathbb{X}^S := (X \times Z, g, \xi)$ and $\mathbb{Y}^S := (Y \times Z', h, v)$ be equipped with discrete measures

$$\xi := \sum_{i=1}^n \xi_i \delta_{(x_i, z_i)} \quad \text{and} \quad v := \sum_{j=1}^m v_j \delta_{(y_j, z'_j)}. \quad (7)$$

²Metric without triangle inequality.

³Semi-metric d without definiteness ($d(x, x') = 0 \not\Rightarrow x = x'$).

For the special case $p = 2$, we start with considering LD_2 in (4), which is the key ingredient of the structure distortion \mathcal{D}_2 . The required 1D Wasserstein distance may here be analytically determined by exploiting (2), where the quantile functions can be explicitly computed using efficient sorting algorithms. To surmount the outer OT problem of FTLB, we instead propose to merely approximate (2) via some quadrature scheme, i.e.,

$$\text{LD}_2^2(x, y) \approx \sum_{k=1}^r w_k |q_{g(x, \cdot)} \# \xi(s_k) - q_{h(y, \cdot)} \# v(s_k)|^2$$

with weights $w_k > 0$, knots $s_k \in (0, 1)$, and small $r \in \mathbb{N}$. Note that we restrict ourselves to quadrature rules with positive weights. Using $\mathbf{q}_{\xi, x} := (\sqrt{w_k} q_{g(x, \cdot)} \# \xi(s_k))_{k=1}^r$ and $\mathbf{q}_{v, y} := (\sqrt{w_k} q_{h(y, \cdot)} \# v(s_k))_{k=1}^r$, we may also write

$$\text{LD}_2^2(x, y) \approx \|\mathbf{q}_{\xi, x} - \mathbf{q}_{v, y}\|^2.$$

In this manner, FTLB can be approximated by

$$\begin{aligned} \text{FTLB}_{\alpha,2}^2(\mathbb{X}^S, \mathbb{Y}^S) & \approx \inf_{\gamma \in \Gamma(\xi, v)} \sum_{i,j=1}^{n,m} \gamma_{ij} \left\| \frac{\sqrt{1-\alpha}(\mathbf{q}_{\xi, x_i} - \mathbf{q}_{v, y_j})}{\sqrt{\alpha}(z_i - z'_j)} \right\|^2 \\ & = W_2^2 \left(\underbrace{\sum_{i=1}^n \xi_i \delta_{\left(\frac{\sqrt{1-\alpha} \mathbf{q}_{\xi, x_i}}{\sqrt{\alpha} z_i}\right)}}_{=: \xi_{\alpha}^{\mathcal{Q}} \in \mathcal{P}(\mathbb{R}^{r+d})}, \underbrace{\sum_{j=1}^m v_j \delta_{\left(\frac{\sqrt{1-\alpha} \mathbf{q}_{v, y_j}}{\sqrt{\alpha} z'_j}\right)}}_{=: v_{\alpha}^{\mathcal{Q}} \in \mathcal{P}(\mathbb{R}^{r+d})} \right), \quad (8) \end{aligned}$$

where $\pi_{ij} := \pi((x_i, z_i), (y_j, z'_j))$. This approximation carries over to TLB. More precisely, for $\mathbb{X} := (X, g, \xi)$ and $\mathbb{Y} := (Y, h, v)$ with $\xi := \sum_{i=1}^n \xi_i \delta_{x_i}$ and $v := \sum_{j=1}^m v_j \delta_{y_j}$, we here obtain

$$\text{TLB}_2^2(\mathbb{X}, \mathbb{Y}) \approx W_2^2 \left(\underbrace{\sum_{i=1}^n \xi_i \delta_{\mathbf{q}_{\xi, x_i}}}_{=: \xi^{\mathcal{Q}} \in \mathcal{P}(\mathbb{R}^r)}, \underbrace{\sum_{j=1}^m v_j \delta_{\mathbf{q}_{v, y_j}}}_{=: v^{\mathcal{Q}} \in \mathcal{P}(\mathbb{R}^r)} \right). \quad (9)$$

Up to here, we have approximated TLB and FTLB using a quadrature rule for the inner 1D Wasserstein distances. To overcome the remaining outer OT minimization behind the Wasserstein distances in (8) and (9), we finally propose to replace these with the SW distance:

$$\begin{aligned} \text{SFTLB}_{\alpha,2}(\mathbb{X}^S, \mathbb{Y}^S) & := \text{SW}_2(\xi_{\alpha}^{\mathcal{Q}}, v_{\alpha}^{\mathcal{Q}}) \\ \text{STLB}_2(\mathbb{X}, \mathbb{Y}) & := \text{SW}_2(\xi^{\mathcal{Q}}, v^{\mathcal{Q}}). \quad (10) \end{aligned}$$

In general, the sliced lower bounds hold only approximately, i.e., $\text{SFTLB}_{\alpha,2}(\mathbb{X}^S, \mathbb{Y}^S) \lesssim \text{FGW}(\mathbb{X}^S, \mathbb{Y}^S)$, where the quality depends on the employed quadrature rule. Since $\xi_{\alpha}^{\mathcal{Q}}, v_{\alpha}^{\mathcal{Q}} \in \mathcal{P}(\mathbb{R}^{r+d})$ and $\xi^{\mathcal{Q}}, v^{\mathcal{Q}} \in \mathcal{P}(\mathbb{R}^r)$, the employed SW distances can be efficiently computed independently of the atom numbers n and m in (7). Moreover, many properties of the original FGW distance from Proposition 3.1 carry over to the sliced variant.

Proposition 3.2. For $\mathbb{X}^S := (X \times Z, g, \xi)$, $\mathbb{Y}^S := (Y \times Z, h, v)$, $Z \subset \mathbb{R}^d$ compact, let $\xi^{\mathcal{F}} := \pi_{Z, \#} \xi$, $v^{\mathcal{F}} := \pi_{Z, \#} v$ and $\mathbb{X} := (X, g, \pi_{X, \#} \xi)$, $\mathbb{Y} := (Y, h, \pi_{Y, \#} v)$. For $k, \ell \in \mathbb{N}$ with $\ell > 1$, let

$$c_{k,\ell} := \sqrt{\frac{A(\mathbb{S}^{k+\ell+1})}{A(\mathbb{S}^{\ell+1})} \frac{A(\mathbb{S}^{\ell-1})}{A(\mathbb{S}^{k+\ell-1})}}.$$

For an arbitrary quadrature rule, it holds:

- (i) SFTLB $_{\alpha,2}$ defines a pseudo-metric $\mathcal{S}(Z)$,
- (ii) SFTLB $_{\alpha,2}(\mathbb{X}^{\mathcal{S}}, \mathbb{Y}^{\mathcal{S}}) \rightarrow c_{r,d} \text{SW}_2(\xi^{\mathcal{F}}, v^{\mathcal{F}})$ as $\alpha \rightarrow 1$,
- (iii) SFTLB $_{\alpha,2}(\mathbb{X}^{\mathcal{S}}, \mathbb{Y}^{\mathcal{S}}) \rightarrow c_{d,r} \text{STLB}_2(\mathbb{X}, \mathbb{Y})$ as $\alpha \rightarrow 0$.

Given that $c_{k,\ell} \rightarrow 0$ for $k \rightarrow \infty$ and ℓ fixed (Gray 2004), the constants $c_{r,d}$ and $c_{d,r}$ illustrate the curse of dimensionality of the SW distance. In contrast to classical Wasserstein transport, the value of the SW distance for measures living on a subspace depends on the dimension of the underlying Euclidean space. Indeed, as we believe this to be of independent interest to the OT community, we state this dependence as a general result employed in the proof of Proposition 3.2.

Proposition 3.3. For $\zeta, \zeta' \in \mathcal{P}(A \times Z)$, $A \subset \mathbb{R}^r$ and $Z \subset \mathbb{R}^d$ compact, with $\pi_{A,\#} \zeta = \pi_{A,\#} \zeta' = \delta_0$, it holds

$$\text{SW}_2(\zeta, \zeta') = c_{r,d} \text{SW}_2(\pi_{Z,\#} \zeta, \pi_{Z,\#} \zeta').$$

Importantly, for empirical measures ξ and v with $n = m$ and $\xi_i = v_i = 1/n$ in (7), the rather simple, equispaced midpoint rule with $r = n$ becomes exact, i.e.,

$$\text{FTLB}_{\alpha,2}(\mathbb{X}^{\mathcal{S}}, \mathbb{Y}^{\mathcal{S}}) = \text{W}_2(\xi_{\alpha}^{\mathcal{Q}}, v_{\alpha}^{\mathcal{Q}})$$

instead of the approximation in (8). Consequently, SFTLB becomes an actual lower bound for the FGW distance. Moreover, SFTLB and FTLB are metrically equivalent.

Proposition 3.4. For $\mathbb{X}^{\mathcal{S}} := (X \times Z, g, \xi)$ and $\mathbb{Y}^{\mathcal{S}} := (Y \times Z, h, v)$ with bounded X and Y , i.e. $g(x, x') \leq D$ and $h(y, y') \leq D$ for fixed D , and compact $Z \subset \mathbb{R}^d$, let ξ and v in (7) be empirical measures with $n = m$ and $\xi_i = v_i = 1/n$. For the equispaced midpoint rule with $r = n$, it holds:

- (i) SFTLB $_{\alpha,2}(\mathbb{X}^{\mathcal{S}}, \mathbb{Y}^{\mathcal{S}}) \leq \text{FGW}_{\alpha,2}(\mathbb{X}^{\mathcal{S}}, \mathbb{Y}^{\mathcal{S}})$,
- (ii) there exists $C_{Z,D,r} > 0$ such that

$$\begin{aligned} \text{SFTLB}_{\alpha,2}(\mathbb{X}^{\mathcal{S}}, \mathbb{Y}^{\mathcal{S}}) &\leq \text{FTLB}_{\alpha,2}(\mathbb{X}^{\mathcal{S}}, \mathbb{Y}^{\mathcal{S}}) \\ &\leq C_{Z,D,r} \text{SFTLB}_{\alpha,2}^{\frac{1}{r+d+1}}(\mathbb{X}^{\mathcal{S}}, \mathbb{Y}^{\mathcal{S}}). \end{aligned} \quad (11)$$

Note that $C_{Z,D,r}$ can be chosen such that (11) holds true for all $\mathbb{X}^{\mathcal{S}}$ and $\mathbb{Y}^{\mathcal{S}}$ of the stated form. Furthermore, Proposition 3.4 can be directly generalized to empirical measures ξ and v with different atom numbers n and m by using a non-equispaced midpoint rule with $r = n + m - 1$, see the technical appendix on arXiv (Piening and Beinert 2025a).

4 Numerical Experiments

In the following, we study the computational runtime of our distances, the underlying transport plans, as well as their behaviour in applications like free-support barycenters, shape classification, and graph isomorphism testing. All experiments were conducted with 13th Gen Intel Core i5-13600K CPU and an NVIDIA GeForce RTX 3060 GPU (12 GB).

Computational and Runtime Analysis

For simplicity, we assume that the measures of the structured spaces $\mathbb{X}^{\mathcal{S}} := (X \times Z, g, \xi)$ and $\mathbb{Y}^{\mathcal{S}} := (Y \times Z, h, v)$ are empirical with the same number of atoms, i.e., $n = m$ and $\xi_i = v_j = 1/n$ in (7). The first step of FTLB and SFTLB is to sample the quantiles of the local distance distributions, i.e., to compute \mathbf{q}_{ξ, x_i} and \mathbf{q}_{v, y_j} in (9), which essentially requires a row-wise sorting of the distance matrices regarding g and h with total complexity $\mathcal{O}(n^2 \log n)$.

n	100	1000	10000
FGW	0.03 ± 0.02	0.45 ± 0.07	178.29 ± 5.22
FTLB	0.06 ± 0.03	0.10 ± 0.02	15.69 ± 0.43
SFTLB	0.04 ± 0.04	0.03 ± 0.01	1.18 ± 0.00

Table 1: Average computation time in seconds for 5 instances with graph sizes 100, 1000, and 10000. The remaining parameters have been chosen as $r = 10$ and $L = 50$.

FTLB In its second step, FTLB requires solving the outer OT problem in (9). Calculating the OT costs requires $\mathcal{O}((r+d)n^2)$ operations. The OT problem itself may be solved exactly in $\mathcal{O}(n^3)$ or approximately in $\mathcal{O}(n^2 \log n)$ using entropic regularization; see (Peyré and Cuturi 2019). This leads to an all-in-all complexity of $\mathcal{O}((r+d+n)n^2)$ (exact) or $\mathcal{O}((r+d+\log n)n^2)$ (approximated). In the experiments, we perform exact inner OT calculations with $r = n$.

SFTLB Here, we initially determine $\xi_{\alpha}^{\mathcal{Q}}$ and $v_{\alpha}^{\mathcal{Q}}$. Next, we employ a Monte Carlo scheme to calculate (10) by

$$\text{SFTLB}_{\alpha,2}^2(\mathbb{X}^{\mathcal{S}}, \mathbb{Y}^{\mathcal{S}}) \approx \frac{1}{L} \sum_{\ell=1}^L \text{W}_2^2(\pi_{\theta_{\ell},\#} \xi_{\alpha}^{\mathcal{Q}}, \pi_{\theta_{\ell},\#} v_{\alpha}^{\mathcal{Q}})$$

with uniformly distributed directions $\theta_{\ell} \in \mathbb{S}^{r+d-1}$. The 1D Wasserstein distance can be computed by sorting, yielding an all-in-all complexity of $\mathcal{O}(Ln(r+d) + (Ln+n^2) \log n)$, where we aim for a small r . Here, the Monte Carlo scheme converges in $\mathcal{O}(\sqrt{L})$ (Nadjahi et al. 2020).

The computational complexity of FTLB and SFTLB is dominated by the preliminary row-wise sorting and the employed OT solvers. To study the actual run-time speed-up, we generate 5 pairs of $n \times n$ distance matrices with $n \in \{100, 1000, 10000\}$ and 1D normal features. Using the POT toolbox (Flamary et al. 2021) for the OT computations, we then determine the corresponding FTLB, SFTLB, and FGW distance. The runtimes are recorded in Table 1, where the significant speed-up due to the slicing is clearly visible.

FTLB Transport Plans

Although the GW distance is well fitted to detect similarities between two single structures, it usually fails to detect these in a set of independent structures. The FGW distance can overcome this issue by assigning appropriate labels. This observation carries over to the outer transport plans of TLB and FTLB, which is illustrated in Figure 2, where the left two spirals are transported to the right two. Source and target are here modelled as point clouds in \mathbb{R}^2 equipped with the Euclidean distance and binary feature labels.

To compare our FTLB transport with FGW and its entropic version, we adapt an example from (Flamary et al. 2021); see Figure 3. Each image displays the source graph on the left and the target graph on the right. The transports are visualized using the color of the community assignment in the source. In this example, the FTLB plan captures the main structures similarly to the original FGW plan, indicating that our FTLB transport may be an efficiently-to-compute alternative in practice.

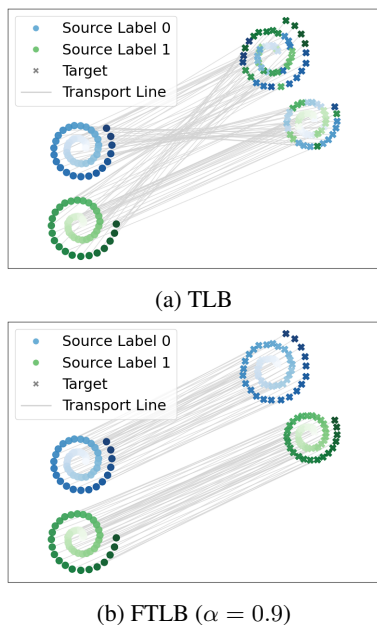


Figure 2: Comparison of TLB and FTLB transport plans. Crosses are colored based on the transported mass. FTLB achieves a more regular transport plan by integrating labels.

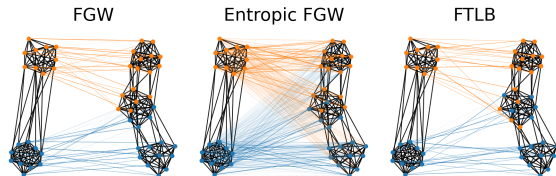


Figure 3: Comparison of FGW, entropic FGW ($\epsilon = 0.01$), FTLB transport plan for $\alpha = 0.5$. Target nodes are colored according to the transported mass. FGW and FTLB capture the same structural similarities.

Free-Support Euclidean Barycenters

GW barycenters are generalized Fréchet means that interpolate several mm-spaces \mathbb{Y}_k , $k = 1, \dots, K$ by solving

$$\arg \min \left\{ \sum_{k=1}^K \text{GW}_2^2(\mathbb{X}, \mathbb{Y}_k) \mid \mathbb{X} \text{ mm-space} \right\}. \quad (12)$$

The solutions of the GW barycenter problem can be completely characterized using multi-marginal OT (Beier, Beinert, and Steidl 2023; Beier and Beinert 2025); its numerical computation, however, remains challenging. Established free-support methods (Peyré and Cuturi 2019) minimize over the distance of \mathbb{X} producing non-Euclidean outputs; whereas fixed-support methods (Beier, Beinert, and Steidl 2023; Beier et al. 2025) limit the resolution of the barycenter. As an alternative, we replace GW in (12) by TLB, STLB ($r = 500, L = 500$) and the so-called Anchor Energy (AE) distance (Sato et al. 2020) for comparison. For each distance, we perform a gradient descent over the points $x_i \in \mathbb{R}^d$ in $\mathbb{X} = (\{x_i\}_{i=1}^n, d_E, \iota)$, where d_E is the Euclidean metric

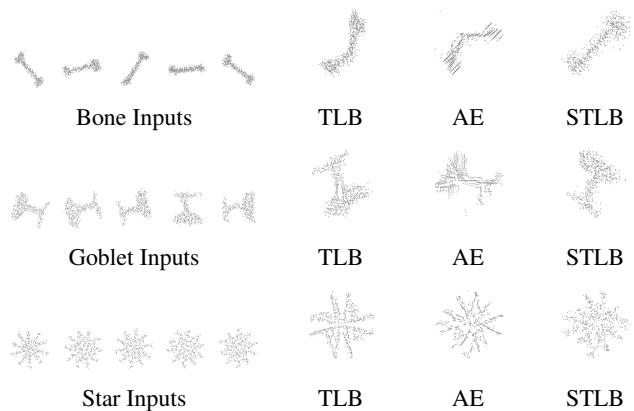


Figure 4: Euclidean 2D TLB, AE and STLB barycenters for five samples of different shape classes.

and ι the uniform measure. Figure 4 shows computed 2D barycenters for different upsampled shapes ($n = 500$) from (Beier, Beinert, and Steidl 2022), obtained via three identical random restart initializations, and 1000 steps (width: 0.1).

Classification of 2D and 3D Shapes

A key application of the GW distance is the comparison of shapes, which can be represented as Euclidean point clouds or as meshes equipped with geodesic distances. Building on experiments from (Mémoli 2011; Beier, Beinert, and Steidl 2022), we rely on the k-nearest neighbor (KNN) classification (Cover and Hart 1967). More precisely, we first precompute pairwise distance matrices with respect to SLB, TLB, STLB ($r = 10, L = 100$), AE (Sato et al. 2020), qGW (Chowdhury and Mémoli 2019), LGW (Beier, Beinert, and Steidl 2022) and GW. Second, we evaluate the classification accuracy by assigning each test point to the majority class among its three nearest neighbors. We report the mean accuracy over 1000 random 25%/75% training/test splits on the following datasets:

2D Shapes: 80 shapes from 4 classes (bone, goblet, star, horseshoe) represented as Euclidean point clouds, each consisting of 50 points (Beier, Beinert, and Steidl 2022);

Animals: 73 human/animal shapes from 7 classes (Sumner and Popović 2004) downsampled to 3D meshes with 50 vertices, equipped with the geodesic distance (Mémoli 2011; Beier, Beinert, and Steidl 2022);

FAUST-500 and FAUST-1000: 100 human meshes from 10 subjects (Bogo et al. 2014), downsampled and represented by 3D meshes with 500 and 1000 vertices equipped with the geodesic distance using PyVista (Sullivan and Kaszynski 2019) and NetworkX (Hagberg, Schult, and Swart 2008).

Table 2 summarizes classification accuracy, runtime, and nullspace types (Mémoli and Needham 2022): SLB vanishes for globally matching distance distributions (‘global’), TLB/STLB/AE for locally matching distributions (‘local’), and GW for isomorphic shapes (‘iso’). Lastly, LGW and

distance	type	2D shapes		animals		FAUST-500		FAUST-1000	
		acc. (%)	time (ms)	acc. (%)	time (ms)	acc. (%)	time (ms)	acc. (%)	time (ms)
SLB	global	97.9±2.2	0.6	100.0±0.1	0.7	31.2±5.4	48.3	35.2 ± 5.7	183.7
TLB	local	100.0±0.3	0.7	100.0±0.0	0.8	36.8±5.4	24.8	40.0 ± 6.0	88.3
STLB	local	99.3±1.2	1.3	99.2±1.1	1.5	35.8±5.5	11.6	40.1±5.7	23.1
AE	local	99.7±0.9	0.6	97.8±1.8	0.7	37.8±5.6	8.1	41.6±5.7	35.8
GW	iso	99.7±0.6	2.6	100.0±0.0	5.8	29.2±4.3	482.0	33.3±5.2	2398.3
LGW	approx	91.2±2.7	288.9	89.8±3.9	481.0	29.4±4.7	$7 \cdot 10^4$	30.2±5.1	10^5
qGW	approx	61.6±4.4	16.4	48.8±6.3	14.7	13.7±3.6	18.2	12.4±3.2	35.3

Table 2: KNN Shape classification: Mean accuracy (acc.) with standard deviation and mean runtime (time).

graph model	WS	BA	RR	WS	BA	RR
nodes	10	10	10	500	500	100
parameters	$k=4, p_E=0.1$	$m=5, p_B=0.5$	$R=3, p_B=0.5$	$k=250, p_E=0.1$	$m=250, p_B=0.1$	$R=3, p_B=0.5$
SLB	90.40 ± 2.62	61.50 ± 3.44	96.30 ± 1.62	50.00 ± 0.00	50.00 ± 0.00	100.00 ± 0.00
TLB	96.80 ± 2.99	99.50 ± 1.02	99.40 ± 1.80	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
STLB	99.00 ± 1.00	99.80 ± 0.40	99.80 ± 0.60	99.60 ± 0.80	99.40 ± 0.92	100.00 ± 0.00
GW	67.40 ± 3.83	73.30 ± 3.85	65.50 ± 4.03	65.00 ± 5.00	100.00 ± 0.00	100.00 ± 0.00
WL-D	100.00 ± 0.00	100.00 ± 0.00	50.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00	50.00 ± 0.00
FTLB	100.00 ± 0.00	96.00 ± 1.18	97.60 ± 1.50	100.00 ± 0.00	94.80 ± 1.60	98.60 ± 0.92
SFTLB	100.00 ± 0.00	98.20 ± 1.08	99.40 ± 0.92	100.00 ± 0.00	100.00 ± 0.00	99.80 ± 0.60
FGW	98.20 ± 1.66	90.30 ± 1.79	74.60 ± 2.69	100.00 ± 0.00	98.80 ± 1.33	69.80 ± 4.42
WL-F	100.00 ± 0.00	83.10 ± 3.21	81.30 ± 3.85	100.00 ± 0.00	99.20 ± 1.33	86.80 ± 3.92

Table 3: Mean accuracy (%) with standard deviation for graph isomorphism testing. The first five methods are label-agnostic, whereas the last four methods employ node features. Proposed methods are printed in bold.

qGW approximate GW (‘approx’)⁴. While most methods perform near-perfectly on ‘2D shapes’ and ‘animals’, accuracy drops on FAUST datasets. On these, ‘local’ methods (TLB, STLB, AE) consistently outperform other approaches. Although GW should yield perfect classifications in theory, the underlying non-convex quadratic minimization reduces the numerical accuracy. STLB shows a clear runtime advantage for large-scale datasets.

Detection of Graph Isomorphisms

An important theoretical property of the GW and FGW distance is that they vanish only for two isomorphic inputs (Mémoli 2011; Vayer et al. 2020a), which facilitates an application in graph isomorphism testing. In practice, the established GW and FGW solvers are however not accurate enough to give reliable results of this NP-hard problem (Grohe and Schweitzer 2020). Therefore, SLB has been proposed as a statistical test for graph isomorphisms (Weitkamp 2022; Weitkamp et al. 2024). However, SLB has less statistical power than TLB (Mémoli and Needham 2022); for this reason, we propose to use STLB and SFTLB instead. During the experiment, we rely on synthetic datasets of random graphs with fixed node numbers, where half of the graph pairs are isomorphic and half of the graph pairs are not. Here, we consider these graph generation models:

Watts–Strogatz (WS) Graph: The WS graph is governed by a neighborhood number k and an edge rewiring probability p_E (Watts and Strogatz 1998). We add node features according to a 1D standard normal distribution.

Barabasi–Albert (BA) Graph: The BA graph is governed by an edge generation number m (Barabási and Albert 1999). We assign node features in $\{0, 1\}$ based on a Bernoulli distribution with probability p_B .

Random Regular (RR) Graph: The RR graph is based on a regularity R (Bollobás 1998). We assign Bernoulli node features in $\{0, 1\}$ according to the probability p_B .

We compare the generated, unlabeled graph pairs using SLB, TLB, STLB, GW, and Weisfeiler-Lehman (WL) kernel test (Shervashidze et al. 2011). For the WL test, we use 5 iterations and set the labels to the node degree (WL-D). The labeled graph pairs are compared using $\alpha = 0.5$ for FTLB, SFTLB ($r = 5, L = 100$), and entropic FGW ($\varepsilon = 10^{-3}$), as well as with WL using 5 iterations and (binned) node features (WL-F). We classify two graphs as isomorphic if they are among the 50% most similar graph pairs in terms of the different distances or if the distance becomes exactly zero. In Table 3, we display the results of 10 experiment repetitions over 6 different random graph configurations.

5 Conclusion

We address the computational limits of GW and FGW by introducing new lower bounds. Extending TLB to the fused setting and linking it to a Wasserstein problem, we derive an efficient sliced FGW via numerical quadrature of 1D OTs. Our work bridges TLB and SW, clarifies SW’s dimensional behavior, and proves effective for shape and graph comparisons, thereby overcoming the limitations of sliced GW (Vayer et al. 2019). In future research, the calculation of local distance distribution required for AE, FTLB, and SFTLB may be further accelerated via sweep-line algorithms (Sato et al. 2020), which is not studied in this work.

⁴LGW/qGW depend heavily on hyperparameter choices balancing speed and performance, see our appendix for a discussion.

Acknowledgments

M.P. acknowledges funding from the German Research Foundation (DFG) within the project BIOQIC (GRK2260/289347353).

References

- Alvarez-Melis, D.; and Fusi, N. 2020. Geometric dataset distances via optimal transport. In *Advances in Neural Information Processing Systems*, volume 33, 21428–21439. Curran Associates.
- Barabási, A.-L.; and Albert, R. 1999. Emergence of scaling in random networks. *Science*, 286(5439): 509–512.
- Beckmann, M.; Beinert, R.; and Bresch, J. 2025a. Max-normalized Radon cumulative distribution transform for limited data classification. In *International Conference on Scale Space and Variational Methods in Computer Vision*, 241–254. Springer.
- Beckmann, M.; Beinert, R.; and Bresch, J. 2025b. Normalized Radon cumulative distribution transforms for invariance and robustness in optimal transport based image classification. arXiv:2506.08761.
- Beier, F.; and Beinert, R. 2025. Tangential fixpoint iterations for Gromov–Wasserstein barycenters. *SIAM Journal on Imaging Sciences*, 18(2): 1058–1100.
- Beier, F.; Beinert, R.; and Steidl, G. 2022. On a linear Gromov–Wasserstein distance. *IEEE Transactions on Image Processing*, 31: 7292–7305.
- Beier, F.; Beinert, R.; and Steidl, G. 2023. Multi-marginal Gromov–Wasserstein transport and barycentres. *Information and Inference: A Journal of the IMA*, 12(4): 2753–2781.
- Beier, F.; Piening, M.; Beinert, R.; and Steidl, G. 2025. Joint metric space embedding by unbalanced optimal transport with Gromov–Wasserstein marginal penalization. In *Proceedings of ICML’25*. OpenReview.net.
- Beinert, R.; Heiss, C.; and Steidl, G. 2023. On assignment problems related to Gromov–Wasserstein distances on the real line. *SIAM Journal on Imaging Sciences*, 16(2): 1028–1032.
- Bogo, F.; Romero, J.; Loper, M.; and Black, M. J. 2014. FAUST: Dataset and evaluation for 3D mesh registration. In *Proceedings of CVPR’14*. IEEE.
- Bollobás, B. 1998. *Random Graphs*. Cambridge: Cambridge University Press.
- Bonet, C.; Drumetz, L.; and Courty, N. 2025. Sliced-Wasserstein distances and flows on Cartan–Hadamard manifolds. *Journal of Machine Learning Research*, 26(32): 1–76.
- Bonet, C.; Vauthier, C.; and Korba, A. 2025. Flowing datasets with Wasserstein over Wasserstein gradient flows. In *Proceedings of ICML’25*. OpenReview.net.
- Bonneel, N.; Rabin, J.; Peyré, G.; and Pfister, H. 2015. Sliced and Radon Wasserstein barycenters of measures. *Journal of Mathematical Imaging and Vision*, 51: 22–45.
- Bonnotte, N. 2013. *Unidimensional and Evolution Methods for Optimal Transportation*. Ph.D. thesis, Université Paris Sud–Paris XI, Orsay, France. PhD Thesis.
- Chowdhury, S.; and Mémoli, F. 2019. The Gromov–Wasserstein distance between networks and stable network invariants. *Information and Inference: A Journal of the IMA*, 8(4): 757–787.
- Chowdhury, S.; Miller, D.; and Needham, T. 2021. Quantized Gromov–Wasserstein. In *Proceedings of ECML PKDD’21*, 811–827. Springer.
- Cover, T. M.; and Hart, P. E. 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1): 21–27.
- Delon, J.; and Desolneux, A. 2020. A Wasserstein-type distance in the space of Gaussian mixture models. *SIAM Journal on Imaging Sciences*, 13(2): 936–970.
- Dukler, Y.; Li, W.; Lin, A.; and Montúfar, G. 2019. Wasserstein of Wasserstein loss for learning generative models. In *Proceedings of ICML’19*, volume 97 of *Proceedings of Machine Learning Research*, 1716–1725. PMLR.
- Flamary, R.; Courty, N.; Gramfort, A.; Alaya, M. Z.; Boisbunon, A.; Chambon, S.; Chapel, L.; Corenflos, A.; Fatras, K.; Fournier, N.; Gautheron, L.; Gayraud, N. T. H.; Janati, H.; Rakotomamonjy, A.; Redko, I.; Rolet, A.; Schutz, A.; Seguy, V.; Sutherland, D. J.; Tavenard, R.; Tong, A.; and Vayer, T. 2021. POT: Python optimal transport. *Journal of Machine Learning Research*, 22(78): 1–8. Software available at <https://pythonot.github.io/>.
- Gray, A. 2004. *Tubes*. Basel: Birkhäuser, 2nd edition.
- Grohe, M.; and Schweitzer, P. 2020. The graph isomorphism problem. *Communications of the ACM*, 63(11): 128–134.
- Hagberg, A. A.; Schult, D. A.; and Swart, P. J. 2008. Exploring network structure, dynamics, and function using NetworkX. In Varoquaux, G.; Vaught, T.; and Millman, J., eds., *Proceedings of SciPy’08*, 11–15. Curvenote.
- Hertrich, J.; Jahn, T.; and Quellmalz, M. 2025. Fast summation of radial kernels via QMC slicing. In *Proceedings of the ICLR’25*. OpenReview.net.
- Hertrich, J.; Wald, C.; Altekruiger, F.; and Hagemann, P. 2024. Generative sliced MMD flows with Riesz kernels. In *Proceedings of the ICLR’24*. OpenReview.net.
- Jin, H.; Yu, Z.; and Zhang, X. 2022. Orthogonal Gromov–Wasserstein Discrepancy with Efficient Lower Bound. In *Proceedings of UAI’22*, volume 180 of *Proceedings of Machine Learning Research*, 917–927. PMLR.
- Kolouri, S.; Nadjahi, K.; Simsekli, U.; Badeau, R.; and Rohde, G. 2019. Generalized sliced Wasserstein distances. *Advances in Neural Information Processing Systems*, 32.
- Kolouri, S.; Park, S. R.; and Rohde, G. K. 2016. The Radon cumulative distribution transform and its application to image classification. *IEEE Transactions on Image Processing*, 25(2): 920–934.
- Mémoli, F. 2011. Gromov–Wasserstein distances and the metric approach to object matching. *Foundations of Computational Mathematics*, 11: 417–487.
- Mémoli, F.; and Needham, T. 2022. Distance distributions and inverse problems for metric measure spaces. *Studies in Applied Mathematics*, 149(4): 943–1001.

- Nadjahi, K.; Durmus, A.; Chizat, L.; Kolouri, S.; Shahrampour, S.; and Simsekli, U. 2020. Statistical and topological properties of sliced probability divergences. In *Advances in Neural Information Processing Systems*, volume 33, 15657–15669. Curran Associates.
- Nadjahi, K.; Vialard, F.-X.; and Peyré, G. 2021. Fast Approximation of the Sliced-Wasserstein distance Using concentration of random projections. In *International Conference on Machine Learning (ICML)*, volume 139 of *Proceedings of Machine Learning Research*, 7946–7956. PMLR.
- Nguyen, D. H.; and Tsuda, K. 2023. On a linear fused Gromov–Wasserstein distance for graph structured data. *Pattern Recognition*, 138: 109351.
- Nguyen, K.; Bariletto, N.; and Ho, N. 2024. Quasi-Monte Carlo for 3D Sliced Wasserstein. In *Proceedings of the ICLR'24*. OpenReview.net.
- Nguyen, K.; Nguyen, H.; Pham, T.; and Ho, N. 2025. Light-speed geometric dataset distance via sliced optimal transport. In *Proceedings of ICML'25*. OpenReview.net.
- Peyré, G.; and Cuturi, M. 2019. Computational optimal transport: with applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6): 355–607.
- Peyré, G.; Cuturi, M.; and Solomon, J. 2016. Gromov–Wasserstein averaging of kernel and distance matrices. In *Proceedings of ICML'16*, volume 48 of *Proceedings of Machine Learning Research*, 2664–2672. PMLR.
- Piening, M.; and Beinert, R. 2025a. A novel sliced fused Gromov–Wasserstein distance. arXiv:2508.02364.
- Piening, M.; and Beinert, R. 2025b. Slicing the Gaussian mixture Wasserstein distance. arXiv:2504.08544.
- Quellmalz, M.; Beinert, R.; and Steidl, G. 2023. Sliced optimal transport on the sphere. *Inverse Problems*, 39(4): 044003.
- Rux, N.; Quellmalz, M.; and Steidl, G. 2025. Slicing of radial functions: a dimension walk in the Fourier space. *Sampling Theory, Signal Processing, and Data Analysis*, 23(1): 1–40.
- Sato, R.; Cuturi, M.; Yamada, M.; and Kashima, H. 2020. Fast and robust comparison of probability measures in heterogeneous spaces. arXiv:2002.01615.
- Scetbon, M.; Klein, M.; Palla, G.; and Cuturi, M. 2023. Unbalanced low-rank optimal transport solvers. In *Advances in Neural Information Processing Systems*, volume 36, 52312–52325. Curran Associates.
- Scetbon, M.; Peyré, G.; and Cuturi, M. 2022. Linear-time Gromov–Wasserstein distances using low-rank couplings and costs. In *Proceedings of ICML'22*, volume 162 of *Proceedings of Machine Learning Research*, 19347–19365. PMLR.
- Shervashidze, N.; Schweitzer, P.; van Leeuwen, E. J.; Mehlhorn, K.; and Borgwardt, K. 2011. Weisfeiler–Lehman graph kernels. *Journal of Machine Learning Research*, 12: 2539–2561.
- Sullivan, B.; and Kaszynski, A. 2019. PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK). *Journal of Open Source Software*, 4(37): 1450.
- Sumner, R. W.; and Popović, J. 2004. Mesh data from deformation transfer for triangle meshes. Available online at <http://people.csail.mit.edu/sumner/research/deftransfer/data.html>. Accessed: 2025-06-03.
- Tanguy, E.; Flamary, R.; and Delon, J. 2025. Properties of discrete sliced Wasserstein losses. *Mathematics of Computation*, 94(353): 1411–1465.
- Vayer, T.; Chapel, L.; Flamary, R.; Tavenard, R.; and Courty, N. 2019. Sliced Gromov–Wasserstein. In *Advances in Neural Information Processing Systems*, volume 32, 10525–10535. Curran Associates. Correction: arxiv:1905.10124.
- Vayer, T.; Chapel, L.; Flamary, R.; Tavenard, R.; and Courty, N. 2020a. Fused Gromov–Wasserstein distance for structured objects. *Algorithms*, 13(9): 212.
- Vayer, T.; Redko, I.; Flamary, R.; and Courty, N. 2020b. CO-optimal transport. In *Advances in Neural Information Processing Systems*, volume 33, 17559–17570. Curran Associates.
- Villani, C. 2003. *Topics in Optimal Transportation*. Providence, RI: American Mathematical Society.
- Watts, D. J.; and Strogatz, S. H. 1998. Collective dynamics of 'small-world' networks. *Nature*, 393(6684): 440–442.
- Weitkamp, C. A. 2022. *Gromov–Wasserstein Distances and their Lower Bounds*. Doctoral thesis, Georg-August-Universität Göttingen.
- Weitkamp, C. A.; Proksch, K.; Taming, C.; and Munk, A. 2024. Distribution of distances based object matching: asymptotic inference. *Journal of the American Statistical Association*, 119(545): 538–551.